A Study of Combinatorial Games over Random Premises, a Model of Baggage Retrieval from Airports, and Union-Closed Families of Sets

विद्या वाचस्पति की
उपाधि की अपेक्षाओं की आंशिक पूर्ति में प्रस्तुत शोध प्रबंध

A thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy

द्वारा / By

ध्रुव भसीन / Dhruv Bhasin

पंजीकरण सं. / Registration No.: 20182008

शोध प्रबंध पर्यवेक्षक / Thesis Supervisor:
Prof. Moumanti Podder

IISER PUNE

भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान पुणे
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

2025

Certificate

This is to certify that this dissertation entitled A Study of Combinatorial Games over Random Premises, a Model of Baggage Retrieval from Airports, and Union-Closed Families of Sets towards the partial fulfilment of the degree of Doctor in Philosophy at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Dhruv Bhasin at Indian Institute of Science Education and Research under the supervision of Prof. Moumanti Podder, Assistant Professor at IISER Pune, during the academic years 2020-2025 .

Moumanti Podder

Prof. Moumanti Podder

Declaration by Student

Name of Student: Dhruv Bhasin

Reg. No.: 20182008
Thesis Supervisor(s): Prof. Moumanti Podder

Department: Mathematics
Date of joining program: 1 August, 2018

Date of Pre-Synopsis Seminar: 18 November, 2024

Title of Thesis: A Study of Combinatorial Games over Random Premises, a Model of Baggage Retrieval from Airports, and Union-Closed Families of Sets

I declare that this written submission represents my idea in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I declare that I have acknowledged collaborative work and discussions wherever such work has been included. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

The work reported in this thesis is the original work done by me under the guidance of Prof. Moumanti Podder.

Date: 24 April, 2025

DBhasin

Signature of the student

Acknowledgments

It is with immense gratitude that I begin by acknowledging the unwavering support and guidance of my Ph.D. supervisor, Prof. Moumanti Podder. Throughout this journey, her dedication to research has left a deep impression on me. I am continually inspired by her work ethic, clarity of thought, and passion for the subject, and I aspire to carry these values with me as I move forward in my career. This thesis simply would not have been possible without her encouragement and support, for which I am forever indebted.

I would like to sincerely thank Prof. Joel Spencer and Prof. Anup Biswas for their valuable insights, constructive feedback, and kind encouragement as members of my Research Advisory Committee. Their perspectives and expertise greatly contributed to shaping this work.

I am grateful to the National Board of Higher Mathematics (NBHM) for financial support during my PhD studies.

My heartfelt thanks go to my collaborators — Prof. Souvik Roy, Prof. Sayar Karmakar, Prof. Alessandro Neri, Prof. Somnath Basu, Prof. Siddhartha Lal, Dr. Siddhartha Patra, and Ms. Anurakti Gupta. Working alongside them has been an immensely enriching experience. Their intellectual generosity, patience, and collaborative spirit have taught me lessons that extend far beyond the confines of this thesis. I am equally grateful to Prof. Amit Hogadi for many stimulating and helpful discussions.

This journey would not have been as meaningful and exciting without the incredible friendships that sustained me through its many highs and lows. I am deeply thankful to Lovleen, Saurabh, Darshan, Priya, Sushant, Visakh, Namrata, Neeraj, Ramya, Ravi, Tumpa, Pranjal, Ronit, Hitendra, Shubham, Sujeet, Ajinkya, Tushar, and Bittu for their constant support, laughter, and companionship. Each one of them has made this long and often difficult journey not just bearable but truly memorable.

I am also grateful to the administrative staff at IISER, especially Suvarna Ma'am, Yogesh Sir, and Sayalee Ma'am, for their efficiency, kindness, and unwavering support. Their help ensured that I could focus on my research without being unduly burdened by administrative concerns.

Beyond academics, my experience at IISER was shaped by vibrant communities that brought immense joy and energy to daily life. I am grateful to all the members of the Zeher Sports Club, whose camaraderie and competitive spirit made life at IISER fun and rewarding. I also take pride in being a founding member of the Independent Students' Body, and I am thankful to my fellow founders — Aniket, Bagheera, Darshan, Pranjal, Priya, Raghav, and Suryo — for believing in the cause of student representation and I am thankful to all the students who participated in and supported this movement. I extend my gratitude to the then Registrar and the then Dean of Student Activities for being receptive to our initiatives and for helpful discussions.

I owe profound thanks to Bhumi for her constant encouragement, patience, and belief in me, even during the moments when my own confidence faltered. Her support has been a pillar of strength throughout these years.

I am deeply indebted to my sisters, Samridhi and Vridhi, and to my parents for their unconditional love, sacrifices, and unwavering belief in my potential. Their support has been the foundation upon which all my efforts are built.

Last, but not the least, I want to thank myself — for continuing to move forward despite every challenge life threw at me, for finding resilience in difficult times, and for believing that there is always light at the end of the tunnel.

Dhruv Bhasin

Abstract

This first five chapters of this thesis are concerned with the study of some *two-player combinatorial games* on random premises.

These chapter include

1. the *k-jump normal* and *k-jump misère* games on *rooted Galton-Watson trees*, in which the token is allowed to be moved from the vertex where it is currently located, to a descendant of this vertex that is at a distance at most k away from this vertex,

2. *percolation games* in their most general form, inspired by both *site percolation* and *bond percolation*, on infinite 2-dimensional lattice graphs.

In the former class of games, the probabilities of the various possible outcomes have been analysed and characterized, phase transition phenomena (pertaining to how the probability of draw changes from being 0 to being strictly positive as the underlying parameter(s) is / are allowed to vary) have been investigated, and sufficient conditions for the average duration of a game to be finite have been proposed. In the latter class of games, regimes (in terms of the values of the parameter(s) under consideration) have been found in which the probability of draw equals 0 – this, in turn, helps establish, with full mathematical rigour, the *ergodicity* of certain classes of *probabilistic cellular automaton* (PCA) that arise from the corresponding game rules.

The model studied in the sixth chapter involves passengers waiting around a conveyor belt, each having checked in precisely one bag, and it studies the interaction between these passengers and the bags appearing onto the belt according to a prespecified permutation. Passengers leave upon collecting their bags, while those still waiting move to positions vacated by co-passengers in front of them. We study the distribution of the time required for all passengers to retrieve their bags, as well as the distribution of the time required for the conveyor belt to become empty for the first time. We establish a fascinating connection between these quantities and the sequence of *telephone numbers* (where the n-th telephone number is the number of matchings on the complete labeled graph K_n).

In the final chapter, we study union-closed families of sets. Given a union-closed family $\mathscr{F}$ of subsets of the universe $[n]$, with $\mathscr{F}$ not equal to the power set of $[n]$, a new subset A can be added to it such that the resulting family remains union-closed. We construct a new family $\overline{\mathscr{F}}$ by adding to $\mathscr{F}$ all such A's, and call this the closure of $\mathscr{F}$. This work is dedicated to the study of various properties of such closures, including characterizing families whose closures equal the power set of $[n]$, providing a criterion for the existence of closure roots of such families etc.

Contents

Chapter 1

Two-player combinatorial games on random premises

1.1 Introduction

In this chapter, we give an overview of Chapters 2, 3, 4, 5. These chapters constitute a study of combinatorial games over random premises. These games are played between two players who alternate in moving a token on a randomly generated playing board which is usually infinite. This board is generated according to a predefined random process and the game is played on a realization of this process. Our game is perfect information in the sense that the players know the complete state of the game at every point.

The playing board is in the form of a graph whose vertices / edges are potentially labeled either 'trap' or 'target'. The game starts with a token at the 'origin' of the graph. The two players alternate turn moving the token to one of the 'neighbours' of the vertex where the token is currently located. Specific to each game, there are some predefined rules which decide when the game comes to an end. An example of such a rule is: moving the token to a vertex which is labeled a 'trap' results in a loss for the player who has made this move.

We now describe the random premises over which the games are played. These premises fall broadly into two categories:

1. Galton-Watson trees: These are random rooted trees defined by a given offspring distribution $\chi = (\chi_i)_{i=0}^{\infty}$, where χ_k for k represents the probability that a vertex produces k offspring. The process begins with a root vertex, say ϕ, which gives birth to k children with probability χ_k where $k \in \mathbb{N}_0$. This branching process is repeated independently at every subsequent vertex, resulting in a random tree. We study games played on an instance of this random tree structure. In the k-jump game (see Chapter 2), for any vertex v, we define $\mathrm{Out}_k(v)$ to be the set of descendants of v at distance of at most k from v. Initially, a token is kept at the root vertex. The players alternate in moving the token. A player is allowed to move the token from a vertex v to any vertex in $\mathrm{Out}_k(v)$. In the normal version of the game, the player who is unable to move the token loses, whereas, in the misère version of the game, the player who is unable to move the token wins. If the game continues indefinitely, it is declared a draw.

2. 2-dimensional infinite lattices: We begin with a lattice whose vertices are indexed by $\mathbb{Z}^2$. For each vertex $(x,y) \in \mathbb{Z}^2$, there is a set of vertices $\mathrm{Out}(x,y)$ such that there is an edge from (x,y) to (u,v) if and only if $(u,v) \in \mathrm{Out}(x,y)$. The random process is governed by four non-negative parameters p, q, r, and s satisfying $0 < p+q \leqslant 1$ and $0 < r+s \leqslant 1$. Each vertex, independently of the others, is labeled a trap with probability p, a target with probability q, and open with probability $(1-p-q)$. Similarly, each edge, independently of the others, is labeled a trap with probability r, a target with probability s, and open with probability

$(1-r-s)$. The game begins with a token at the origin. The players take turns moving the token from the current vertex (x,y) to a vertex in $\mathrm{Out}(x,y)$. If a player is forced to move the token to a vertex which is labeled a trap, she loses the game. Conversely, if a player moves the token to a vertex which is labeled a target, she wins the game. Similarly, if a player is forced to move the token along an edge that is a trap, she loses, while moving the token along a target edge results in a win. If the game continues indefinitely without a winner, it is declared a draw.

Given a realization of the random process at hand, we divide the vertices of this graph into three disjoint subsets denoted W, L and D (when we study the normal game, we denote these subsets to be NW, NL and ND and, similarly, when we study the misére game we denote these subsets to be MW, ML and MD.) We say that a vertex $v \in W$ if the game that starts from the vertex v then, it results in a win for the player who moves first. The sets L and D are defined analogously.

One of the important aspects of the analysis of these games we study is obtaining recurrence relations governed by the game rules. While we shall derive these relations formally for every game in the corresponding sections, we now give a brief idea of how these recurrence relations are obtained: suppose that we know which sets, among W, L and D, all the vertices in $\mathrm{Out}(v)$ belong to, for some vertex v. Suppose that in addition to this we know the label of each edge from v to u where $u \in \mathrm{Out}(v)$, i.e., whether it is trap, target or open and the label assigned to v itself (i.e., whether it is trap, target or open). This information along with games rules will tell us which of W, L and D the vertex v should belong to. As an example, suppose that all the vertices in $\mathrm{Out}(v)$ are labeled W, all the corresponding edges are open and that the vertex v is open. In this scenario, no matter which vertex in $\mathrm{Out}(v)$ the player who plays the first round moves the token to from v, the game will always result in a loss for her. Thus, $v \in L$.

Before delving into the study of specific games, we give a description of probabilistic cellular automata and their ergodicity, which share close relationships with the games we have studied on various infinite grids with vertex set $\mathbb{Z}^2$.

1.1.1 Probabilistic Cellular Automata

A PCA F defined on the lattice $\mathbb{Z}^d$ (and hence referred to as a d-dimensional PCA), for some $d \in \mathbb{N}$, comprises a finite set of states $\mathscr{A}$ that is called its *alphabet*, a finite set of indices $\mathscr{N} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m\} \subset \mathbb{Z}^d$ that is called its *neighborhood*, and a stochastic matrix $\varphi : \mathscr{A}^m \times \mathscr{A} \to [0,1]$ that is called its (random) *local update rule*. Given a *configuration* η in the *state space* $\Omega = \mathscr{A}^{\mathbb{Z}^d}$, we apply F to η, obtaining a (random) configuration $F\eta$, in which the state $F\eta(\mathbf{x})$ of the site

$\mathbf{x} \in \mathbb{Z}^d$ is a random variable whose probability distribution is given by

$$\mathbf{P}[F\eta(\mathbf{x}) = b \big| \eta(\mathbf{x}+\mathbf{y}_i) = a_i \text{ for all } 1 \leqslant i \leqslant m] = \varphi(a_1, a_2, \ldots, a_m; b) \text{ for all } b \in \mathscr{A}, \qquad (1.1.1)$$

for any $a_1, a_2, \ldots, a_m \in \mathscr{A}$. Here, by definition of stochastic matrices, for all $a_1, a_2, \ldots, a_m, b \in \mathscr{A}$, we have $\varphi(a_1, a_2, \ldots, a_m; b) \geqslant 0$ and $\sum_{b \in \mathscr{A}} \varphi(a_1, a_2, \ldots, a_m; b) = 1$. The updation from $\eta(\mathbf{x})$ to $F\eta(\mathbf{x})$ happens independently over all sites $\mathbf{x} \in \mathbb{Z}^d$. Since we consider discrete-time PCAs, it makes sense to indicate by η_t the configuration at time $t \in \mathbb{N}_0$, so that $\eta_{t+1} = F\eta_t$ for all $t \in \mathbb{N}_0$. We call a PCA *elementary* when it is defined on $\mathbb{Z}$ (i.e. $d = 1$) and $|\mathscr{N}| = |\mathscr{A}| = 2$. We refer the reader to [[74], §2] and [73] for excellent expositions on PCAs in general.

Next, we come to a brief discussion regarding the notion of ergodicity of a d-dimensional PCA. To begin with, we let $\mathscr{F}$ denote the σ-field that is generated by the cylinder sets of $\Omega = \mathscr{A}^{\mathbb{Z}^d}$, and we let $\mathbb{D}$ denote the set of all probability measures supported on Ω and defined with respect to $\mathscr{F}$. We emphasize here that every probability measure on Ω that is henceforth mentioned belongs to the set $\mathbb{D}$. We define $F^t\eta = F(F^{t-1}\eta)$ for $\eta \in \Omega$ and $t \in \mathbb{N}$ (in particular, $F^1\eta = F\eta$). In other words, $F^t\eta$ is the (random) configuration that is obtained by applying F sequentially t times to the initial configuration η. These definitions extend naturally to random η, and if η follows the probability distribution μ (that belongs to $\mathbb{D}$), we let $F^t\mu$ (simply written $F\mu$ when $t = 1$) denote the probability distribution of the (also random) configuration $F^t\eta$.

Definition 1.1.1. *We say that μ is a* stationary *or* invariant *measure for a PCA F if $F\mu = \mu$ (in other words, the pushforward measure induced by F is the same as the original measure). We call the PCA F* ergodic *if it has a* unique *stationary measure, say μ, which is* attractive*, i.e. for* every *probability measure ν on Ω, the sequence $F^t\nu$ converges weakly to μ as $t \to \infty$.*

A one-dimensional PCA is said to be *elementary* if the cardinality of its alphabet as well as its neighbourhood-marking set equals 2. In Chapter 7 of [37], two fundamental results pertaining to ergodicity of elementary PCAs have been proposed. We state these two results in terms of the notation used in this thesis. We let $\mathscr{A} = \{W, L\}$ denote the alphabet of *any* elementary PCA F, and we let

$$\theta_{i,j} = \mathbf{P}\left[F\eta(x) = L \big| \eta(x+y_1) = i, \eta(x+y_2) = j\right], \text{ for } i, j \in \mathscr{A},$$

where $\mathscr{N} = \{y_1, y_2\}$ forms the neighbourhood-marking set of F. Thus, the PCA F is completely specified by the parameters $\theta_{W,W}$, $\theta_{W,L}$, $\theta_{L,W}$ and $\theta_{L,L}$. We call F *symmetric* when $\theta_{W,L} = \theta_{L,W}$, in which case F is specified by only three parameters, namely, $\theta_{W,W}$, $\theta_{W,L}$ and $\theta_{L,L}$. The following are two well-known results when it comes to ergodicity properties of symmetric elementary PCAs, stated in Chapter 7 of [37]:

(a) the PCA F is ergodic when $\theta_{W,W}$, $\theta_{W,L}$ and $\theta_{L,L}$ satisfy the inequalities:

$$0 < \theta_{W,W}, \theta_{W,L}, \theta_{L,L} < 1, \tag{1.1.2}$$

$$\theta_{L,L} > \theta_{W,W} - 2\theta_{W,L}, \tag{1.1.3}$$

$$\theta_{L,L} > \theta_{W,W} - 2(1 - \theta_{W,L}); \tag{1.1.4}$$

(b) the PCA F is ergodic when $\theta_{W,W}$, $\theta_{W,L}$ and $\theta_{L,L}$ satisfy the inequality:

$$\max\{|\theta_{i,j} - \theta_{k,\ell}| : i, j, k, \ell \in \{W, L\}\} + 2\max\{|\theta_{L,L} - \theta_{W,L}|, |\theta_{W,W} - \theta_{W,L}|\} < 2. \tag{1.1.5}$$

It has been mentioned in Chapter 7 of [37] that these two regimes together cover more than 90% of the unit cube $[0,1]^3$ defined by the parameter-triple $(\theta_{W,W}, \theta_{W,L}, \theta_{L,L})$, and that the only region where no rigorous method for proving or disproving ergodicity for symmetric elementary PCAs, in general, is known is the union of the neighbourhoods of the points $(\theta_{W,W}, \theta_{W,L}, \theta_{L,L}) = (1,0,0)$ and $(\theta_{W,W}, \theta_{W,L}, \theta_{L,L}) = (1,1,0)$.

CAs are objects that are simple to define but have a fair amount of complexity in their properties. Computationally they are of great interest because of the fact that they can simulate any Turing machine [53]. It is worth mentioning that the celebrated Game of Life of John Conway can be seen as a CA [51].

As mentioned earlier, PCAs are CAs that have stochastic update rules. We refer the reader to [28, 70, 73, 74] which include a detailed survey of PCAs where they discuss different ways PCAs can arise in combinatorics, statistical physics, computational cell biology and theoretical computer science. Some of the primary motivations of studying PCAs are to investigate the fault tolerant capabilities of CAs ([77, 78]), classification of elementary CAs by using their robustness to errors as a discriminating criterion ([21, 69] and to investigate their connections with Gibbs potentials and Gibbs measures coming from statistical physics ([55, 68, 92]).

In [60], which is the primary motivation for our work, the idea of percolation games was introduced. It is also a two-player combinatorial game played on a realization of traps, targets and open sites on $\mathbb{Z}^2$. This realization is obtained in a similar manner as described in the previous section. The game studied in this paper is such that a player can move the token from (x,y) to one of $(x+1,y)$ or $(x,y+1)$ (i.e., $\text{Out}(x,y) = \{(x+1,y),(x,y+1)\}$). The authors build a connection between this game and a PCA $A_{p,q}$: the probability of draw in this game is zero if and only if $A_{p,q}$ is ergodic. A more direct connection is established between the game and the PCA $F_{p,q}$, which is envelope to $A_{p,q}$ (we refer the reader to [28] for an introduction to the notion of envelope PCAs), via the recurrence relations we have alluded to in the previous section. The main tool implemented for proving $\mu((x,y) \in D) = 0$, where μ is any stationary distribution for $F_{p,q}$, is the technique of weight function or potential function.

In [23], the connections made in [60] were used to come up with computer-generated weight functions for two different PCAs whose update rules can be described by $\eta_{t+1}(n) = \mathrm{BSC}_p(\mathrm{NAND}(\eta_t(n-1), \eta_t(n)))$ and $\eta_{t+1}(n) = \mathrm{NAND}(\mathrm{BSC}_p(\eta_t(n-1), \eta_t(n)))$. Here, BSC_p denotes a binary symmetric channel that takes a bit as input and flips it with probability p and leaves it unchanged with probability $1-p$. The weight functions obtained in this paper are for $p \in (0, \varepsilon)$ for some small $\varepsilon > 0$. The authors introduce the concept of polynomial linear programming and a suitable algorithm to obtain a suitable weight function for proving the ergodicity of these PCAs.

In Chapter 3, Chapter 4, and Chapter 5, we explore the method of weight functions to come up with ergodicity results for various PCAs in various regimes of parameter values. Alongside proving ergodicity of these PCAs, we are also able to establish that the probability draw is 0 in the corresponding games in the corresponding regimes of parameter values. In particular, our results in Chapter 4 are able to cover partially the regime that is not covered in Chapter 7 of [37]. Chapter 7 of [37] establishes, via more traditional means, the ergodicity of symmetric elementary PCAs (i.e. PCAs in which each of the alphabet and the neighbourhood-marking set is of cardinality 2, and which are parametrized by the triple $(\theta_{W,W}, \theta_{W,L}, \theta_{L,L})$ that belongs to the unit hypercube $[0,1]^3$), and nearly 90% of the parameter-space (which, as already mentioned, is $[0,1]^3$) is covered by these results. However, no general method has been found yet that is able to establish ergodicity when sufficiently small neighbourhoods of $(1,0,0)$ and $(1,1,0)$ are considered. Our results in Chapter 4 are able to cover quite a significant chunk of the neighbourhood around $(1,0,0)$.

Coming to Chapter 3, we see that one of the two PCAs considered has an alphabet of cardinality 2, a neighbourhood-marking set of cardinality 3, and is parametrized by the four-tuple $(\theta_{W,W,W}, \theta_{W,W,L}, \theta_{W,L,L}, \theta_{L,L,L})$ which belongs to the four-dimensional unit hypercube $[0,1]^4$. Our result establishes the ergodicity of this PCA over the entire parameter-space $[0,1]^4$ – a feat that is not, in general, possible to achieve using traditional means. In Chapter 5, yet another generalization is considered by allowing for heterogeneity of the neighbourhood-marking set depending on the parity of the cell or site concerned. In this work, as well, we are able to establish ergodicity over a significant portion of the parameter-space (which is, once again, the unit hypercube $[0,1]^3$) whereas general means for proving ergodicity rigorously are not always able to address certain neighbourhoods of this parameter-space.

Chapter 2

k-jump Galton-Watson Games

Preface

This chapter is based on the following paper:

- Bhasin D. and Podder M. "Combinatorial games on Galton-Watson trees involving several-generation-jump moves" – *Moscow Journal of Combinatorics and Number Theory*, Volume 13(1), Pages 1–58, 2024. DOI: https://doi.org/10.2140/cnt.2024.13.1.

2.1 Introduction

The simplest yet intriguing versions of the normal and misère games on rooted random trees were studied in [61]. Each game involves two players (henceforth addressed as P1 and P2) and a token, and requires visualizing a given rooted tree as a directed graph in which an edge between a parent vertex u and its child v is assumed to be directed from u to v. In each game, the players take turns to move the token along these directed edges. In a normal game, the first player to get stuck at a leaf vertex (i.e. unable to move the token any further) loses, whereas in a misère game, this same player wins. The two games share a fair amount of similarities in their analysis, but the objective of each player in a normal game is precisely the opposite of that in a misére game. While the authors of [61] provide an incredibly thorough analysis of these two games when played on rooted Galton-Watson (henceforth abbreviated as GW) trees, they also pose several open questions. Our work in this chapte delves deeper into the fascinating world of these two-player combinatorial games and asks: *what if we do not restrict the players to only single-generation moves?* In other words, instead of allowing each player, when it is her turn, to move the token from its current position u to a child of u, we now permit her to move the token from u to any descendant of u that is at a distance at most k away from u, where $k \geqslant 1$ is a pre-assigned positive integer. We call the corresponding versions of the normal and misére games the *k-jump normal* and *k-jump misère* games respectively.

2.1.1 Introduction to the games

We begin with a description of the rooted GW trees on which our games are played. A Galton-Watson branching process (henceforth denoted $\mathscr{T}_\chi$), introduced in [105] and independently studied in [17] as a model to investigate the extinction of ancestral family names, begins with a root ϕ giving birth to a random number X_0 of children where X_0 follows the *offspring distribution* χ (a probability distribution supported on the set $\mathbb{N}_0$ of non-negative integers). If $X_0 = 0$, we stop the process, whereas if $X_0 = m$ for some $m \in \mathbb{N}$, the children of ϕ are named $v_1, \ldots, v_m$, and v_i gives birth to X_i children with $X_1, \ldots, X_m$ i.i.d. χ. Thus the process continues, and the resulting tree is infinite with positive probability iff the expectation of χ exceeds 1. We refer the reader to [7], [5] and [6] for further reading on GW trees.

We now come to a formal description of the games studied in this chapter. Given any realization T of $\mathscr{T}_\chi$, any vertex u in T, and $i \in \mathbb{N}$, let $\Gamma_i(u)$ denote the set of all descendants v of u (excluding u itself) such that the distance between u and v is at most i. The vertex at which the token is placed at the beginning of a game is known as the *initial vertex*. The players P1 and P2 take turns to make *moves* (with P1 moving in the first round), where a move constitutes relocating the token from its current position, which is some vertex u in T, to a vertex $v \in \Gamma_k(u)$, where $k \in \mathbb{N}$ is fixed *a priori*. The player who is unable to make a move loses the k-jump normal game. Hence, in this game, each of P1 and P2 strives to relocate the token, obeying the rules of the game, from its current position to a leaf vertex of T, thereby making her opponent lose in the next round. On the other hand, the player who is unable to make a move wins the k-jump misère game. Therefore, in this game, each of P1 and P2 strives to force her opponent to relocate the token to a leaf vertex of T, thereby ensuring that she herself wins the game in the next round.

It is important to note here that a realization T of the random tree $\mathcal{T}_\chi$ is first generated and then revealed *in its entirety* to *both* P1 and P2, *before* the game begins. These games are thus *complete information* games. We also assume that P1 and P2 are both intelligent agents who play *optimally*, i.e. when a game is destined to end in a decision, the player who wins tries to end the game in as few rounds as possible, while the player who loses tries to prolong the game as much as possible.

2.1.2 Motivations for studying these games

The primary motivation for studying these games stems from our interest in examining how allowing each player more room to maneuver in each round ends up affecting the probability of each possible outcome (these outcomes being a win for P1, a loss for P1, and a draw for both players). It is also imperative that we view the k-jump versions of the games as broad generalizations of the versions studied in [61].

It turns out that it is rather hard to draw a direct, analytical comparison between the (1-jump) normal and misère games studied in [61] on one hand and the corresponding k-jump versions (for $k \geqslant 2$) on the other, even though our intuitions may suggest otherwise. Almost all such questions remain open and unexplored. For the commonly studied regime where the offspring distribution χ of the GW tree is Poisson(λ), we analytically compare the 1-jump normal game with the 2-jump normal game for sufficiently large values of λ in Theorem 2.1.9. Further comparisons can be drawn visually by plotting the curves corresponding to $\mathrm{n}\ell_1$, $\mathrm{n}\ell_2$ and $\mathrm{n}\ell_3$, the curves corresponding to nw_1, nw_2 and nw_3, and the curves corresponding to nd_1, nd_2 and nd_3, as functions of λ, when χ is Poisson(λ) (see Figures 2.1, 2.2 and 2.3). Here $\mathrm{n}\ell_k$, nw_k and nd_k respectively denote the probabilities of P1's loss, P1's win, and a draw in the k-jump normal game (see §2.1.3 for detailed definitions).

While direct comparisons seem difficult to deduce analytically for $k \geqslant 3$, it is illuminating to explore the many characteristics of the functions H_k (see Theorem 2.1.1) and J_k (see Theorem 2.1.2) whose minimum positive fixed points equal the probabilities $\mathrm{n}\ell_k$ and $\mathrm{m}\ell_k$ of P1 losing the k-jump normal game and the k-jump misère game respectively. It is instructive to examine how these func-

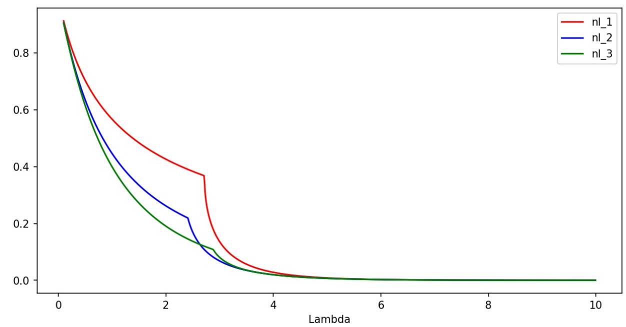


Figure 2.1: Comparing probabilities $n\ell_k$ of P1 losing as functions of λ, for $k = 1, 2, 3$

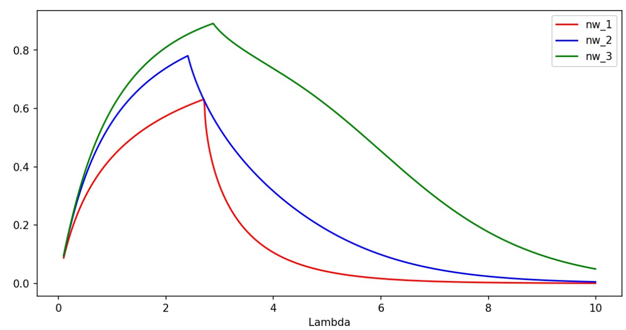


Figure 2.2: Comparing probabilities nw_k of P1 winning as functions of λ, for $k = 1, 2, 3$

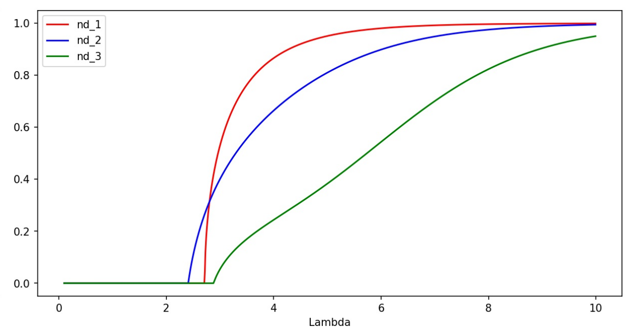


Figure 2.3: Comparing probabilities nd_k of draw as functions of λ, for $k = 1, 2, 3$

tions behave for various values of $k \in \mathbb{N}$. Analyzing these functions is the key to understanding the probabilities of the various possible outcomes of the k-jump games, and comparing and contrasting H_k (respectively J_k) for different values of k is instrumental in comparing and contrasting the games themselves.

For a quick appraisal of how H_k (respectively J_k) behaves as we vary k, and the pivotal roles they play in determining various characteristics of the probabilities $n\ell_k$, nw_k, nd_k (respectively $m\ell_k$, mw_k, md_k), we mention here some of our findings that have been described in detail later on. In §2.2.2, we show that H_k (and likewise, J_k, as mentioned in §2.3) is increasing on $[0, c_{k-1}] \subset [0, 1]$, where $\{[0, c_k]\}_k$ forms a sequence of steadily shrinking intervals (see Lemma 2.2.2). The proof is far more involved for higher values of k than when we consider $k = 1$ (in fact, for $k = 1$, we have H_1 and J_1 defined and increasing on the entire interval $[0, 1]$). In §2.6, it takes considerable work to show that when the offspring distribution χ is Poisson(λ) with $\lambda \geqslant 2$, the function H_2 is strictly convex on the interval $[0, c_2]$. Plotting H_2 for various values of λ (see, for example, Figure 2.4, where $\lambda = 5$ has been considered) seems to suggest that H_2 is, in fact, *not* convex on the interval $(c_2, c_1]$. When χ is Poisson(λ), the second assertion of Theorem 2.1.4 sheds light on the decay rate of $n\ell_k$ for *all* values of k as $\lambda \to \infty$, whereas Theorem 2.1.6 provides an even stronger result on the decay rate of $n\ell_k$ when $k = 2$. It is through a careful analysis of the derivative of H_k at the point c_k that we obtain Theorems 2.1.4 and 2.1.5, both of which shed light on the *phase transition* phenomenon pertaining to the draw probability nd_k when χ is Poisson(λ), i.e. how the value of nd_k evolves from 0 to strictly positive as λ is increased gradually. The magnitude of $H_k'(c_k)$ also plays a role, in Theorem 2.1.10, in determining if the expected duration of a k-jump normal game is finite. Finally, comparing H_2 with J_2 enables us to compare the probabilities of the various outcomes of the 2-jump normal game with those of the 2-jump misère game in Theorem 2.1.7, while comparing H_2 with H_1 allows a similar comparison between the 1-jump normal game and the 2-jump normal game in Theorem 2.1.9.

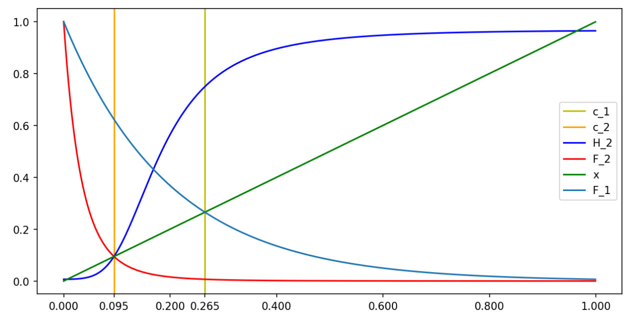


Figure 2.4: The function H_2 is *not* convex between $c_2 \approx 0.095$ and $c_1 \approx 0.265$, when χ is Poisson(5)

A second, and perhaps equally compelling, motivation arises from viewing these games as tools for understanding a generalized notion of *finite state tree automata* (henceforth abbreviated as FSTA). As such, the simplest FSTA is a state machine that comprises a *finite* set $\Sigma = \{1, 2, \ldots, r\}$ of *states* or *colours*, and a *rule* f_1, which is a function : $\mathbb{N}_0^r \to \Sigma$, such that if a vertex v in a rooted tree has n_i children that are in state i for each $i \in \Sigma$, then the state of v is given by $f_1(n_1, n_2, \ldots, n_r)$. Given a fixed $k \in \mathbb{N}$, a *generalized depth-k FSTA* (henceforth abbreviated as a k-GFSTA), with rule f_k, extends the above definition as follows: given a rooted tree T, a vertex v of T and an assignment $\sigma : \Gamma_k(v) \to \Sigma$ of states to the vertices of $\Gamma_k(v)$ (recall from above that $\Gamma_k(v)$ is the set of all descendants of v, other than v itself, that are at distance at most k away from v), the state of v, as dictated by this k-GFSTA, is given by $f_k\big((\sigma(w) : w \in \Gamma_k(v))\big)$. For instance, a rule $f_k : \mathbb{N}_0^r \to \Sigma$ may be considered such that if there are n_i vertices in $\Gamma_k(v)$ that are in state i for each $i \in \Sigma$, then the state of v is given by $f_k(n_1, n_2, \ldots, n_r)$. Denoting by NW_k the set of vertices v such that P1 wins the k-jump normal game that starts at v, by NL_k the set of vertices v such that P1 loses the k-jump normal game that starts at v, and by ND_k the set of vertices v such that the k-jump normal game that starts at v ends in a draw (see also §2.1.3 for these definitions), we obtain a k-GFSTA with the state space $\Sigma = \{\mathrm{NW}_k, \mathrm{NL}_k, \mathrm{ND}_k\}$ and the rule f_k defined by (2.2.1) and (2.2.2). Yet another k-GFSTA is obtained from the recurrence relations (2.3.1) and (2.3.2) arising from the k-jump misère game. It is evident that studying these games may pave the way for a deeper understanding of k-GFSTAs for large values of k, the associated *recursive distributional equations* and their *fixed points*.

A fixed point of an FSTA is a probability distribution ν on the state space Σ such that if the children of the root ϕ of a GW tree $\mathscr{T}_\chi$ are assigned i.i.d. states from Σ, each following the common law ν, then the *induced* (random) state (via the rule f_1) at ϕ also follows the law ν. Let $\mathscr{T}$ be the set of *all* possible rooted trees, and let, for any vertex v in a rooted tree T, $T(v)$ denote the subtree that comprises v and all its descendants. A map $\iota : \mathscr{T} \to \Sigma$ is called an *interpretation* of the FSTA if assigning the state or colour $\iota(T(v))$ to each vertex v of *any* arbitrary rooted tree T gives us a colouring of the *entire* tree T that is consistent with the rule f_1 of the FSTA. We call a fixed point ν of the FSTA *interpretable* if there exists an interpretation ι of the FSTA with $\iota(\mathscr{T}_\chi)$ following the law ν. Necessary and sufficient conditions for fixed points of a certain class of FSTAs to be interpretable were addressed in [62]. We mention here that the notion of interpretability ties in closely with the concept of *endogeny* (see, for instance, §2.4, and in particular, Definition 7, of [2], as well as [71], [72], [75], [89], [88]).

Let us understand how one may extend the above-mentioned notion of fixed points to the case of the 2-GFSTA obtained from the 2-jump normal game (with its rule f_2 given by (2.2.1) and (2.2.2) for $k = 2$). Let $\mathscr{C}_{0,1}$ denote the set of all vertices v such that v has at least one child whose state is NL_2, $\mathscr{C}_{0,2}$ the set of all vertices v such that v has no child in state NL_2 but at least one grandchild in state NL_2, and $\mathscr{C}_{1,2}$ the set of all vertices v such that every child of v is in state NW_2 and at least one grandchild of v is in state NL_2. It is immediate from these definitions that $\mathscr{C}_{1,2} \subset \mathscr{C}_{0,2}$. We mention here that these subsets are analogous to those defined in (2.2.6) (with the superscripts n, $n+1$ etc. removed).

We now describe the recurrence relations (these have been elaborated upon in §2.2 for the general case of *any* $k \in \mathbb{N}$) that tie the above-mentioned subsets together. A vertex v is in NW_2 if and only if it is either in $\mathscr{C}_{0,1}$ or $\mathscr{C}_{0,2}$, which is equivalent to saying that v either has a child that is in NL_2 or a child that is in $\mathscr{C}_{0,1}$. A vertex v is in NL_2 if and only if either v is childless or every child of v is in $\mathscr{C}_{1,2}$. In all other situations, v is in ND_2. We further note that v is in $\mathscr{C}_{0,1}$ if and only if at least one child of v belongs to NL_2. It is in $\mathscr{C}_{1,2}$ if and only if all its children are in NW_2 and at least one of its children is in $\mathscr{C}_{0,1}$, which is equivalent to saying that all children of v are in $\mathscr{C}_{0,1} \cup \mathscr{C}_{0,2}$ and at least one of them is in $\mathscr{C}_{0,1}$. Finally, v is in $\mathscr{C}_{0,2}$ if none of its children is in NL_2 but at least one of them is in $\mathscr{C}_{0,1}$.

We assign i.i.d. states from the state space $\Sigma = \{\mathscr{C}_{0,1}, \mathscr{C}_{1,2}, \mathscr{C}_{0,2} \setminus \mathscr{C}_{1,2}, \mathrm{NL}_2, \mathrm{ND}_2\}$ to the children of the root ϕ of the GW tree $\mathscr{T}_\chi$ according to the common law ν, where we set $p_{0,1} = \nu[\mathscr{C}_{0,1}]$, $p_{1,2} = \nu[\mathscr{C}_{1,2}]$, $p_{0,2} = \nu[\mathscr{C}_{0,2}]$ (so that $\nu[\mathscr{C}_{0,2} \setminus \mathscr{C}_{1,2}] = p_{0,2} - p_{1,2}$), $\mathrm{n}\ell_2 = \nu[\mathrm{NL}_2]$ and $\mathrm{nd}_2 = \nu[\mathrm{ND}_2]$. For ν to be a fixed point of this 2-GFSTA, the random state induced at ϕ must follow the law ν as well. This requires that the following equations, derived from the above-mentioned recurrence relations, hold:

$$\mathrm{n}\ell_2 = \sum_{m=0}^{\infty} p_{1,2}^m \chi(m) = G(p_{1,2}), \tag{2.1.1}$$

$$p_{0,1} = \sum_{m=1}^{\infty} [1 - (1 - \mathrm{n}\ell_2)^m] \chi(m) = 1 - G(1 - \mathrm{n}\ell_2), \tag{2.1.2}$$

$$p_{1,2} = \sum_{m=1}^{\infty} \left[(p_{0,1} + p_{0,2})^m - p_{0,2}^m\right] \chi(m) = G(p_{0,1} + p_{0,2}) - G(p_{0,2}), \tag{2.1.3}$$

$$p_{0,2} = \sum_{m=1}^{\infty} [(1 - \mathrm{n}\ell_2)^m - (1 - \mathrm{n}\ell_2 - p_{0,1})^m] \chi(m) = G(1 - \mathrm{n}\ell_2) - G(1 - \mathrm{n}\ell_2 - p_{0,1}), \tag{2.1.4}$$

where $\chi(m)$ denotes the probability of ϕ having m children, for $m \in \mathbb{N}_0$, and G indicates the probability generating function corresponding to χ. The above equations are used to solve for the fixed point ν. Note that we do not need a separate equation for nd_2 since $\mathrm{nd}_2 = 1 - \mathrm{n}\ell_2 - p_{0,1} - p_{0,2}$.

It is obvious that the sets $\mathscr{C}_{0,1}$, $\mathscr{C}_{1,2}$, $\mathscr{C}_{0,2} \setminus \mathscr{C}_{1,2}$, NL_2 and ND_2 that arise out of the 2-jump normal game itself provide an interpretation ι of the above 2-GFSTA with $\iota(\mathscr{T}_\chi)$ following the law ν. But there are questions pertaining to interpretability that remain open when the system of equations constituting (2.1.1), (2.1.2), (2.1.3) and (2.1.4) does not yield a unique solution for ν. For instance, Theorem 2.1.1 asserts that $\mathrm{n}\ell_2$ is the minimum positive fixed point of H_2, whereas Corollary 2.4.2 tells us that c_2 is also a fixed point of H_2. From Theorem 2.1.5, we see that when χ is Poisson(λ), we have $\mathrm{n}\ell_2 < c_2$ for all $\lambda > \lambda_c$ for $\lambda_c \approx 2.41$. So now we ask: is the probability distribution obtained by replacing $\mathrm{n}\ell_2$ by c_2 (and computing the corresponding values of $p_{0,1}$, $p_{0,2}$, $p_{1,2}$ and nd_2 from (2.1.1), (2.1.2), (2.1.3) and (2.1.4)) interpretable in the sense described earlier (see the definition for interpretability of fixed points of FSTAs in the previous page)? For higher values of k, H_k may have several fixed points other than these two, and understanding the interpretability of the corresponding probability distributions is also of interest to us.

A third motivation for investigating these games arises from our speculation that these games may serve as precursors to more complicated versions where, for example, each round involves choosing one of P1 and P2 uniformly randomly and then allowing her to make a move (where a move involves relocating the token from its current vertex to a child of that vertex), but ensuring that no player is chosen for more than k consecutive rounds, where k is a pre-fixed positive integer.

2.1.3 Notations and some definitions

Given a rooted tree T, we denote by $V(T)$ its vertex set. As previously mentioned, for $u \in V(T)$, we define $\Gamma_i(u)$, for $i \in \mathbb{N}$, as the set of descendants v of u (not including u itself) with $\rho(u,v) \leqslant i$, where ρ denotes the graph metric on T. As mentioned above in §2.1.2, we shall denote by G the probability generating function (pgf) of the offspring distribution χ of the GW tree $\mathscr{T}_\chi$, i.e. $G(x) = \sum_{i=0}^{\infty} x^i \chi(i)$ for any $x \in [0,1]$. All offspring distributions χ considered in this chapter satisfy $0 < \chi(0) < 1$.

Given $k \in \mathbb{N}$, we define NL_k (or simply NL when the value of k is clear from the context) to be the set of all vertices $v \in V(\mathscr{T}_\chi)$ such that if v is the initial vertex, then P1, who plays the first round, loses the k-jump normal game. Likewise, let NW_k (or simply NW) denote the set of all $v \in V(\mathscr{T}_\chi)$ such that if v is the initial vertex, then P1, playing the first round, wins the k-jump normal game. Let ND_k (or simply ND) denote the set of all $v \in V(\mathscr{T}_\chi)$ such that if v is the initial vertex, the k-jump normal game ends in a draw.

It is important to consider refinements of the above subsets of vertices in order to understand the probabilities of the game's outcomes better. For every $n \in \mathbb{N}$, we define $\mathrm{NL}_k^{(n)}$ (or $\mathrm{NL}^{(n)}$ when the value of k is unambiguous from the context) to be the subset of NL_k comprising vertices v such that if v is the initial vertex, the k-jump normal game lasts for less than n rounds. Likewise, $\mathrm{NW}_k^{(n)}$ (or simply $\mathrm{NW}^{(n)}$) is the subset of NW_k comprising vertices v such that the k-jump normal game starting at v ends in less than n rounds. We define $\mathrm{ND}^{(n)} = \mathrm{ND}_k^{(n)} = V(\mathscr{T}_\chi) \setminus \left[\mathrm{NL}_k^{(n)} \cup \mathrm{NW}_k^{(n)}\right]$, or in other words, $v \in \mathrm{ND}_k^{(n)}$ iff the k-jump normal game starting at v lasts for at least n rounds. We set $\mathrm{NL}_k^{(0)} = \mathrm{NW}_k^{(0)} = \emptyset$. By definition, we have $\mathrm{NL}_k^{(n)} \subset \mathrm{NL}_k^{(n+1)}$ and $\mathrm{NW}_k^{(n)} \subset \mathrm{NW}_k^{(n+1)}$ for all $n \in \mathbb{N}_0$.

We define $\mathrm{n}\ell_k$ to be the probability of the event that the root ϕ of $\mathscr{T}_\chi$ belongs to NL_k, whereas for each $n \in \mathbb{N}$, we define $\mathrm{n}\ell_k^{(n)}$ to be the probability of the event that ϕ belongs to $\mathrm{NL}_k^{(n)}$. Likewise, we define nw_k, $\mathrm{nw}_k^{(n)}$, nd_k and $\mathrm{nd}_k^{(n)}$. From above, we have $\mathrm{n}\ell_k^{(0)} = \mathrm{nw}_k^{(0)} = 0$. Once again, the subscript k is dropped whenever its value is clear from the context. From the last line of the previous paragraph, we get

$$\mathrm{n}\ell_k^{(n)} \leqslant \mathrm{n}\ell_k^{(n+1)} \text{ and } \mathrm{nw}_k^{(n)} \leqslant \mathrm{nw}_k^{(n+1)} \text{ for all } n \in \mathbb{N}_0. \tag{2.1.5}$$

The corresponding subsets for the k-jump misére games are denoted by ML_k, MW_k, MD_k, $\mathrm{ML}_k^{(n)}$, $\mathrm{MW}_k^{(n)}$ and $\mathrm{MD}_k^{(n)}$, and the corresponding probabilities by $\mathrm{m}\ell_k$, mw_k, md_k, $\mathrm{m}\ell_k^{(n)}$, $\mathrm{mw}_k^{(n)}$ and $\mathrm{md}_k^{(n)}$

(as above, the subscript is removed when the value of k is clear from the context).

2.1.4 Main results

We begin by introducing a couple of sequences of functions that are defined recursively. The first is $\{F_i\}_{i\in\mathbb{N}_0}$, where $F_0(x) = 1$ for $x \in [0,1]$, and $F_i : [0, c_{i-1}] \to [0,1]$, for all $i \in \mathbb{N}$, is defined as

$$F_i(x) = G(F_{i-1}(x) - x),\ x \in [0, c_{i-1}], \tag{2.1.6}$$

where recall that G is the pgf of χ, and c_i is the unique (as shown in Lemma 2.2.2) fixed point of F_i. The second sequence of functions $\{g_i\}_{i\in\mathbb{N}}$ is defined as follows. The function $g_1 : \mathbb{R}^2 \to \mathbb{R}$ is defined as $g_1(x,y) = x - y$, and having defined g_{i-1} for any $i \geqslant 2$, we define $g_i : \mathscr{D}_i \to \mathbb{R}$ as

$$g_i(x_0,x_1,\ldots,x_i) = G(g_{i-1}(x_0,x_2,\ldots,x_i)) - G(g_{i-1}(x_1,x_2,\ldots,x_i)),\ (x_0,x_1,\ldots,x_i) \in \mathscr{D}_i, \tag{2.1.7}$$

where the sets $\mathscr{D}_i$ are also recursively defined, as follows:

$$\mathscr{D}_i = \{(x_0,x_1,\ldots,x_i) : g_{i-1}(x_0,x_2,\ldots,x_i) \in [0,1], g_{i-1}(x_1,x_2,\ldots,x_i) \in [0,1]\}. \tag{2.1.8}$$

As an example, $\mathscr{D}_2 = \{(x_0,x_1,x_2) : x_2 \leqslant x_0 \leqslant x_2+1, x_2 \leqslant x_1 \leqslant x_2+1\}$. The motivation behind defining $\mathscr{D}_i$ this way is to simply ensure that the arguments $g_{i-1}(x_0,x_2,\ldots,x_i)$ and $g_{i-1}(x_1,x_2,\ldots,x_i)$ of the function G in (2.1.7) belong to the domain $[0,1]$ on which G is defined.

Theorem 2.1.1. *Consider the k-jump normal game for $k \in \mathbb{N}$. Define the function $H_k : [0, c_{k-1}] \to [0,1]$ as*

$$H_k(x) = G(g_k(F_0(x), F_1(x), \ldots, F_k(x))). \tag{2.1.9}$$

Then $\mathrm{n}\ell_k$ is the minimum positive fixed point of H_k. Moreover, $\mathrm{nw}_k = 1 - F_k(\mathrm{n}\ell_k)$.

Some discussions are in order regarding the functions defined above, as they form an integral part of the results in this chapter. First, we compare the findings of Theorem 2.1.9 with those of Theorem 1 (i) of [61]. Let us define the function $R(x) = 1 - G\big(1 - G(x)\big)$ for $x \in [0,1]$. According to the notation used in this chapter, Theorem 1 (i) of [61] states that $1 - \mathrm{n}\ell_1$ equals the maximum fixed point and nw_1 the minimum fixed point of $R(x)$ in $[0,1]$. From our definitions of g_1 and F_0, (2.1.6) and (2.1.9), we see that $H_1(x) = G\big(1 - G(1-x)\big)$. For any $y \in [0,1]$, we observe that $1 - y$

is a fixed point of $R(x)$ if and only if

$$1-y=1-G\big(1-G(1-y)\big)\iff y=G\big(1-G(1-y)\big)=H_1(y),$$

which is equivalent to y being a fixed point of $H_1(x)$. This observation immediately reveals that $\mathrm{n}\ell_1$ is the minimum fixed point of $H_1(x)$ in $[0,1]$ if and only if $1-\mathrm{n}\ell_1$ is the maximum fixed point of $R(x)$ in $[0,1]$, i.e. our conclusion about $\mathrm{n}\ell_1$ from Theorem 2.1.1 matches with what Theorem 1 (i) of [61] yields. Moreover, for every $y\in[0,1]$ that is a fixed point of $H_1(x)$, we observe that

$$1-F_1(y)=1-G(1-y)=1-G\big(1-H_1(y)\big)=R\big(1-G(1-y)\big)=R(1-F_1(y)),$$

showing that $1-F_1(y)$ is a fixed point of $R(x)$ in $[0,1]$. Conversely, under the assumption that G is a strictly increasing function on $[0,1]$ (which is true whenever $\chi(0)<1$), we observe that $1-F_1(y)$ is a fixed point of $R(x)$ in $[0,1]$, for $y\in[0,1]$, if and only if

$$\begin{aligned}&R\big(1-F_1(y)\big)=1-F_1(y)\iff G\big(1-G\big(1-F_1(y)\big)\big)=F_1(y)\\&\iff G\big(1-G\big(1-G(1-y)\big)\big)=G(1-y)\iff 1-G\big(1-G(1-y)\big)=1-y\iff H_1(y)=y,\end{aligned}$$

thus showing us that y is a fixed point of $H_1(x)$ in $[0,1]$. The last two observations tell us that y is a fixed point of $H_1(x)$ in $[0,1]$ if and only if $1-F_1(y)$ is a fixed point of $R(x)$ in $[0,1]$. Since F_1 is strictly decreasing on $[0,1]$ and $\mathrm{n}\ell_1$ has already been shown above to be the minimum fixed point of $H_1(x)$ in $[0,1]$, this establishes that $1-F_1(\mathrm{n}\ell_1)$ must be the minimum fixed point of $R(x)$ in $[0,1]$, thus showing that our conclusion about nw_1 from Theorem 2.1.1 matches with what Theorem 1 (i) of [61] yields.

We now try to give the reader an idea as to how the functions in (2.1.6), (2.1.7) and (2.1.9) come to be defined, without going into the actual technical details (which have been laid out fully in the proof of Theorem 2.1.1 in §2.2), and to this end, we focus on the case of $k=2$. This case has already been discussed, to some extent, in §2.1.2. In §2.1.2, along with NW_2, NL_2 and ND_2, we defined the subsets $\mathscr{C}_{0,1}$, $\mathscr{C}_{0,2}$ and $\mathscr{C}_{1,2}$, described the recurrence relations that tie these subsets to one another, and derived Equations (2.1.1), (2.1.2), (2.1.3) and (2.1.4) that relate the probabilities $\mathrm{n}\ell_2$, nd_2, $p_{0,1}$, $p_{0,2}$ and $p_{1,2}$ with each other (where $p_{i,j}$ is the probability that the root of $\mathscr{T}_\chi$ belongs to $\mathscr{C}_{i,j}$, for $0\leqslant i<j\leqslant 2$).

From the definition of F_0 and (2.1.6), we have $F_1(\mathrm{n}\ell_2)=G(1-\mathrm{n}\ell_2)$ and $F_2(\mathrm{n}\ell_2)=G\big(G(1-\mathrm{n}\ell_2)-\mathrm{n}\ell_2\big)$. Using the expression for F_1 and the definition of g_1, it becomes immediate from (2.1.2) that $p_{0,1}=g_1\big(F_0(\mathrm{n}\ell_2),F_1(\mathrm{n}\ell_2)\big)$. Using the expressions for F_1 and F_2 and Equations (2.1.2) and (2.1.4), we have

$$\begin{aligned} p_{0,2} &= G(1-\mathrm{n}\ell_2) - G(1-p_{0,1}-\mathrm{n}\ell_2) = F_1(\mathrm{n}\ell_2) - G(G(1-\mathrm{n}\ell_2)-\mathrm{n}\ell_2) \\ &= F_1(\mathrm{n}\ell_2) - F_2(\mathrm{n}\ell_2) = g_1\big(F_1(\mathrm{n}\ell_2), F_2(\mathrm{n}\ell_2)\big). \end{aligned}$$

From (2.1.2), (2.1.3), (2.1.4) and (2.1.7), we see that

$$\begin{aligned} p_{1,2} &= G\big(1-F_2(\mathrm{n}\ell_2)\big) - G\big(g_1\big(F_1(\mathrm{n}\ell_2), F_2(\mathrm{n}\ell_2)\big)\big) \\ &= G\big(g_1\big(F_0(\mathrm{n}\ell_2), F_2(\mathrm{n}\ell_2)\big)\big) - G\big(g_1\big(F_1(\mathrm{n}\ell_2), F_2(\mathrm{n}\ell_2)\big)\big) = g_2\big(F_0(\mathrm{n}\ell_2), F_1(\mathrm{n}\ell_2), F_2(\mathrm{n}\ell_2)\big). \end{aligned}$$

These ideas extend to the general case of arbitrary $k \in \mathbb{N}$, with the identity $p_{i,j,n} = g_{i+1}\big(F_{j-i-1}(\mathrm{n}\ell_k), F_{j-i}(\mathrm{n}\ell_k), F_{k-i+1}(\mathrm{n}\ell_k), \ldots, F_k(\mathrm{n}\ell_k)\big)$ being true for all $0 \leqslant i < j \leqslant k$ (see Lemma 2.2.3 and its proof for a better understanding of this fact). We now use the expression for $p_{1,2}$ derived above, (2.1.1) and (2.1.9) to conclude that $\mathrm{n}\ell_2$ is indeed a fixed point of $H_2(x)$ in $[0,1]$. Although the proof of Theorem 2.1.1 is rather involved when arbitrary values of k are considered, we hope that the above exposition, for $k = 2$, helps to shed some light on how our argument proceeds, how the recursive definitions of the functions in (2.1.6), (2.1.7) and (2.1.9) arise etc.

In order to state, for the case of k-jump misère games, the result that is analogous to Theorem 2.1.1, we introduce yet another sequence of functions $\{\gamma_i\}_{i\in\mathbb{N}}$ that bears significant resemblance to (2.1.7). Setting $\gamma_1 \equiv g_1$, for each $i \geqslant 2$ we define $\gamma_i : \mathscr{D}'_i \to \mathbb{R}$ recursively as

$$\gamma_i(x_0, x_1, \ldots, x_i) = G\left(\chi(0) + \gamma_{i-1}(x_0, x_2, \ldots, x_i)\right) - G\left(\chi(0) + \gamma_{i-1}(x_1, x_2, \ldots, x_i)\right), \tag{2.1.10}$$

where the sets $\mathscr{D}'_i$ are recursively defined as $\mathscr{D}'_i = \{(x_0, x_1, \ldots, x_i) : \chi(0) + \gamma_{i-1}(x_0, x_2, \ldots, x_i) \in [0,1], \chi(0) + \gamma_{i-1}(x_1, x_2, \ldots, x_i) \in [0,1]\}$.

Theorem 2.1.2. *Consider the k-jump misère game, $k \in \mathbb{N}$. Define the function $J_k : [0, c_{k-1}] \to [0,1]$ as*

$$J_k(x) = G\left(\chi(0) + \gamma_k(F_0(x), F_1(x), \ldots, F_k(x))\right) - \chi(0). \tag{2.1.11}$$

Then $\mathrm{m}\ell_k$ is the minimum positive fixed point of J_k. Moreover, $\mathrm{mw}_k = 1 - F_k(\mathrm{m}\ell_k) + \chi(0)$.

We mention here, for the convenience of the reader, that much of the proof of Theorem 2.1.2 unfolds the same way as that of Theorem 2.1.1, and the motivations behind the recursive defintions of the functions in (2.1.10) and (2.1.11) are very similar to those behind the recursive definitions of the functions in (2.1.7) and (2.1.9) respectively.

Theorem 2.1.3 provides bounds on $\mathrm{n}\ell_k$ and $\mathrm{m}\ell_k$, and necessary and sufficient conditions for the draw probabilities nd_k and md_k to be positive. Recall that c_k is the (unique, by Lemma 2.2.2) fixed point of F_k.

Theorem 2.1.3. *For every $k \in \mathbb{N}$, we have $\chi(0) < \mathrm{n}\ell_k \leqslant c_k$ and $\mathrm{m}\ell_k \leqslant \hat{c}_k$, where $\hat{c}_k$ is the unique point of intersection between $y = F_k(x)$ and $y = J_k(x) + \chi(0)$ in $(0, c_{k-1})$. Moreover, $\mathrm{nd}_k > 0$ if and only if $\mathrm{n}\ell_k < c_k$ and $\mathrm{md}_k > 0$ if and only if $\mathrm{m}\ell_k < \hat{c}_k$.*

It is worthwhile to note that when the offspring distribution χ of $\mathscr{T}_\chi$ has expectation bounded above by 1, $\mathscr{T}_\chi$ is finite almost surely, which, for any fixed k, forces the k-jump normal game starting at the root ϕ of $\mathscr{T}_\chi$ to end in a finite number of rounds almost surely. Consequently, the probability of draw in such a situation is 0. In particular, this tells us that when χ is Poisson(λ) for $\lambda \leqslant 1$, the probability that the k-jump normal game results in a draw is 0. Theorem 2.1.4 is of an asymptotic nature, asserting that when χ is Poisson(λ), the probability of the k-jump normal game ending in a draw eventually becomes strictly positive as we keep increasing λ. Evidently, this gives rise to a phase transition phenomenon in that, the probability $\mathrm{nd}_k = \mathrm{nd}_{k,\lambda}$ of the event that a k-jump normal game played on a rooted Galton-Watson tree with Poisson(λ) offspring results in a draw goes from being equal to 0 for $\lambda \leqslant 1$ to being strictly positive for all λ large enough.

Theorem 2.1.4. *Fix any $k \in \mathbb{N}$. When the offspring distribution χ is Poisson(λ), we have $\mathrm{nd}_k = \mathrm{nd}_{k,\lambda} > 0$ for all λ sufficiently large. We also have $\lambda^{k-1}\,\mathrm{n}\ell_k = \lambda^{k-1}\,\mathrm{n}\ell_{k,\lambda} \to 0$ as $\lambda \to \infty$.*

Theorem 2.1.5 provides a more nuanced insight into the phase transition phenomenon when $k = 2$:

Theorem 2.1.5. *For $\lambda \geqslant 2$, the function $H_2 = H_{2,\lambda}$ is strictly convex on the interval $[0, c_2] = [0, c_{2,\lambda}]$. The slope of $H_{2,\lambda}$ at $c_{2,\lambda}$ is strictly increasing as a function of λ, for all $\lambda \geqslant 1$. As a consequence of these two facts, there is a unique* critical $\lambda_c \approx 2.41$ *such that for all $2 \leqslant \lambda < \lambda_c$ we have $\mathrm{nd}_2 = \mathrm{nd}_{2,\lambda} = 0$, and for all $\lambda > \lambda_c$, we have $\mathrm{nd}_{2,\lambda} > 0$.*

The next result is an especially strong one as it sheds light on the rate of decay of $\mathrm{n}\ell_{k,\lambda}$, when $k = 2$ and the offspring distribution is Poisson(λ), as $\lambda \to \infty$.

Theorem 2.1.6. *We have $\lim_{\lambda \to \infty} \lambda^i\,\mathrm{n}\ell_{2,\lambda} = 0$ for* all $i \in \mathbb{N}$.

Theorem 2.1.7 compares the 2-jump normal game with the 2-jump misère game (see Proposition **??** for a more precise description of the values of λ for which the first of the three inequalities is shown to hold analytically), while Theorem 2.1.9 compares the 1-jump normal game with the 2-jump normal game, when all of these games are played on $\mathscr{T}_\chi$ with χ being Poisson(λ).

Theorem 2.1.7. *When $k = 2$ and χ is Poisson(λ), we have* $\mathrm{m}\ell_{2,\lambda} \leqslant \mathrm{n}\ell_{2,\lambda}$, $\mathrm{nd}_{2,\lambda} < \mathrm{md}_{2,\lambda}$ *and* $\mathrm{m}\ell_{2,\lambda} \leqslant \mathrm{n}\ell_{2,\lambda} < \mathrm{nw}_{2,\lambda}$ *for all λ sufficiently large.*

Remark 2.1.8. *From Theorems 2.1.6 and 2.1.7, we conclude that* $\lim_{\lambda\to\infty} \lambda^i \mathrm{m}\ell_{2,\lambda} = 0$ *for all* $i \in \mathbb{N}$.

Theorem 2.1.9. *When $k = 2$ and χ is Poisson(λ), we have* $\mathrm{n}\ell_{2,\lambda} \leqslant \mathrm{n}\ell_{1,\lambda}$, $\mathrm{nd}_{2,\lambda} < \mathrm{nd}_{1,\lambda}$ *and* $\mathrm{nw}_{1,\lambda} < \mathrm{nw}_{2,\lambda}$ *for all λ sufficiently large.*

Our final result goes back to general offspring distributions χ, and concerns itself with average durations of k-jump normal games. We conjecture, from the patterns noticed in its proof, that the second assertion of Theorem 2.1.10 can be extended to *any* $k \in \mathbb{N}$, though we cannot seem to provide an intuitive argument as to why such a relation should be true.

Theorem 2.1.10. *For any fixed k, when $\mathrm{n}\ell_k = c_k$ and $\max\left\{H_k'(c_k), \left|F_k'(c_k)\right|\right\} < 1$, the expected duration of the k-jump normal game is finite. Moreover, for $k = 2,3$, if $\mathrm{n}\ell_k = c_k$ and $\left|F_k'(c_k)\right| < 1$, then once again, the expected duration is finite.*

It is worthwhile to note here that while the condition $\mathrm{n}\ell_k = c_k$ alone does guarantee that $\mathrm{nd}_k = 0$ (by Theorem 2.1.3) and hence the k-jump normal game ends almost surely in a finite number of rounds, it does not automatically imply that the expected number of rounds is going to be finite as well.

2.1.5 A brief discussion of the literature on combinatorial games

Before we plunge into our exploration of the k-jump normal and k-jump misère games, the rich and variegated literature on combinatorial games that has developed over the past several decades deserves some delineation. This extremely broad class of games (see, for example, [47] and [48] for a general introduction) constitutes primarily two-player games with perfect information, no chance moves, and the possible outcomes being victory for one player (and loss for the other) and draw for both players. Aside from being utilized in studying mathematical problems that belong to complexity classes harder than NP, these games have intimate connections with disciplines such as mathematical logic, automata theory, complexity theory, graph and matroid theory, networks, error-correcting codes, online algorithms. Outside of mathematics, these games find applications in biology, psychology, economics, insurance, actuarial studies and political sciences.

[60] studies *percolation games* on oriented Euclidean lattices. Each site of $\mathbb{Z}^2$, independent of all other sites, is marked a "trap" or a "target" or "open" with probabilities p, q and $1-p-q$

respectively, and the two players take turns to move a token from its current position (x, y) to either $(x+1, y)$ or $(x, y+1)$. If a player moves to a target, she wins immediately, and if she moves to a trap, she loses immediately. The game's outcome can be interpreted in terms of the evolution of a one-dimensional discrete-time probablistic cellular automaton (PCA) – specifically, the game having no chance of ending in a draw is shown to be equivalent to the ergodicity of this PCA. [60] also establishes a connection between the *trapping game* (i.e. where $q = 0$) on directed graphs in higher dimensions and the hard-core model on related undirected graphs with reduced dimensions. [12] studies the trapping game on undirected graphs, where the players take turns to move the token from the vertex at which it is currently located to an adjacent vertex that has never been visited before, and the player unable to make a move loses (note the evident connection between this game and the normal game described above). The outcome of this game is shown to have close ties with maximum-cardinality matchings, and a draw in this game relates to the sensitivity of such matchings to boundary conditions. [104] studies a related, two-person zero-sum game called *exploration* on a *rooted distance model*, to analyze minimum-weight matchings in edge-weighted graphs. In a related game called *slither* ([4]), the players take turns to claim yet-unclaimed edges of a simple, undirected graph, such that the chosen edges, at all times, form a path, and whoever fails to move, loses. This too serves as a tool for understanding maximum matchings in graphs.

Bearing some resemblance to slither are the *maker-breaker positional games* ([58]), involving a set X, a collection $\mathscr{F}$ of subsets of X, and positive integers a and b. The players named *Maker* and *Breaker* take turns to claim yet-unclaimed elements of X, with Maker choosing a elements at a time and Breaker b elements at a time, until all elements of X are exhausted. Maker wins if she has claimed all elements of a subset in $\mathscr{F}$. When this game is played on a graph, the players take turns to claim yet-unclaimed edges, and Maker wins if the subgraph induced by her claimed edges satisfies a desired property (e.g. it is connected, or it forms a clique of a given size, a Hamiltonian cycle, a perfect matching or a spanning tree). The game is *unbiased* when $a = b$, and *biased* otherwise. This game has intimate connections with existential fragments of first order and monadic second order logic on graphs. [98] and [99] study the threshold probability $p = p_c$ beyond which Maker has a winning strategy when this game is played on Erdős-Rényi random graphs $G(n, p)$; [100] studies the game for Hamiltonian cycles on the complete graph K_n; [14] studies the game on random geometric graphs; [45] studies the *critical bias* b^* of the $(1 : b)$ biased game on $G(n, p(n))$ for $p(n) = \Theta(\ln n/n)$. In addition, [98] studies the game where Maker wins if she can claim a non-planar graph or a non-k-colourable graph. [32] indicates a deep connection between positional games on complete graphs and the corresponding properties being satisfied by a random graph, and this is consistent with *Erdős' probabilistic intuition*, which states that the course of a combinatorial game between two players playing optimally often resembles the evolution of a purely random process. Finally, the *Ehrenfeucht-Fraïssé games* comprise yet another extensive subclass of combinatorial games that play a pivotal role in our understanding of first and monadic second order logic on random rooted trees and random graphs (see, for example, [94, 95, 96, 65, 18, 81, 102, 80, 106, 67, 76, 107, 90, 83, 84, 59, 82]).

2.1.6 Organization of this chapter

Theorem 2.1.1 is proved in §2.2, with the two main parts addressed in §2.2.1 and §2.2.2 (Lemma 2.2.2, Lemma 2.2.3 and Equation (2.2.14) are proved in §2.10.1 of the Appendix). The (very similar) proof of Theorem 2.1.2 is briefly discussed in §2.3. Theorem 2.1.3 is proved in §2.4, Theorem 2.1.4 in §2.5 (Lemmas 2.5.1 and 2.5.2 proved in §2.10.2 of the Appendix), and Theorem 2.1.5 is proved in §2.6 (Lemmas 2.6.1 through 2.6.6 proved in §2.10.3 of the Appendix). Theorems 2.1.6 and 2.1.7 are proved in §2.7 (Lemma 2.7.1 proved in §2.10.4 of the Appendix), Theorem 2.1.9 is proved in §2.8, and the proof of Theorem 2.1.10 is covered in §2.9.

2.2 Proof of Theorem 2.1.1

Fix $k \in \mathbb{N}$ throughout §2.2, and hence the subscript k is dropped from notations used in this chapter (for instance, NW_k is replaced by NW, $\mathrm{NW}_k^{(n)}$ is replaced by $\mathrm{NW}^{(n)}$ etc.). We begin by deducing the two most fundamental recurrence relations pertaining to the k-jump normal game. For a vertex u to be in NW, P1 must be able to move the token, in the first round, to some descendant v of u in $\Gamma_k(u)$ such that, if we now consider the game that starts at v and P2 plays the first round, P2 loses. In other words, such a v must be in NL (note here that the symmetric roles of the two players is crucial). Thus

$$u \in \mathrm{NW} \Leftrightarrow \exists\, v \in \Gamma_k(u) \text{ such that } v \in \mathrm{NL} \Leftrightarrow \Gamma_k(u) \cap \mathrm{NL} \neq \emptyset. \tag{2.2.1}$$

For a vertex u to be in NL, either u is childless, in which case P1 is unable to make her very first move, or else *every* vertex v in $\Gamma_k(u)$ is such that, if P1 moves the token there, then the game that begins at v with P2 playing the first round is won by P2. In other words, every v in $\Gamma_k(u)$ must belong to NW. Thus we have

$$u \in \mathrm{NL} \Leftrightarrow \Gamma_1(u) = \emptyset \text{ or } v \in \mathrm{NW} \text{ for every } v \in \Gamma_k(u) \Leftrightarrow \Gamma_k(u) \subset \mathrm{NW}. \tag{2.2.2}$$

Next, we establish a compactness result which shows that, if a player is able to win the k-jump normal game on a rooted tree which is *locally finite* (i.e. every vertex of the tree has finite degree), then she can guarantee to do so within a finite number of rounds which can be specified in advance. Mathematically, this result can be stated as follows:

Lemma 2.2.1. *In any k-jump normal game, we have* $\mathrm{NW} = \bigcup_{n=1}^{\infty} \mathrm{NW}^{(n)}$ *and* $\mathrm{NL} = \bigcup_{n=1}^{\infty} \mathrm{NL}^{(n)}$.

This result is proven essentially the same way as Proposition 7 of [61], but we include a proof nonetheless (see §2.10.1 of the Appendix) for the sake of completeness of this work. We note here that the offspring distribution χ that we consider for $\mathscr{T}_\chi$ is supported on $\mathbb{N}_0$, hence the (random) number of children of any vertex of $\mathscr{T}_\chi$ is almost surely finite.

As a consequence of Lemma 2.2.1, we have $\mathrm{NL}^{(n)} \uparrow \mathrm{NL}$ and $\mathrm{NW}^{(n)} \uparrow \mathrm{NW}$ as $n \uparrow \infty$, which in turn yields

$$\lim_{n\to\infty} \mathrm{n}\ell^{(n)} = \mathrm{n}\ell \text{ and } \lim_{n\to\infty} \mathrm{nw}^{(n)} = \mathrm{nw}\,. \tag{2.2.3}$$

The two main parts of the proof of Theorem 2.1.1 are outlined as follows: in §2.2.1, we show that $\mathrm{n}\ell$ is a fixed point of the function H_k, and in §2.2.2, we prove that $\mathrm{n}\ell$ is, in fact, the minimum positive fixed point of H_k.

2.2.1 Showing that $\mathrm{n}\ell_k$ is a fixed point of H_k

For a vertex u to be in $\mathrm{NW}^{(n+1)}$ for any $n \in \mathbb{N}$, there must exist some $v \in \Gamma_k(u)$ such that, in the k-jump normal game that begins at v, the player who plays the first round loses in less than n rounds. In other words,

$$u \in \mathrm{NW}^{(n+1)} \Leftrightarrow \exists\, v \in \Gamma_k(u) \text{ with } v \in \mathrm{NL}^{(n)}\,. \tag{2.2.4}$$

For u to be in $\mathrm{NL}^{(n+1)}$, either u is childless, or every descendant $v \in \Gamma_k(u)$ must be such that the k-jump normal game that begins at v is won in less than n rounds by the player who plays the first round. Thus,

$$u \in \mathrm{NL}^{(n+1)} \Leftrightarrow \Gamma_1(u) = \emptyset \text{ or } v \in \mathrm{NW}^{(n)} \text{ for every } v \in \Gamma_k(u). \tag{2.2.5}$$

We emphasize here that for any $m, n \in \mathbb{N}$, the subsets $\mathrm{NW}^{(m)}$ and $\mathrm{NL}^{(n)}$ are mutually exclusive.

Figure 2.5 (for $k = 2$) is included here to help the reader visualize the more refined subsets or *classes* of vertices that we now introduce to carry out the full analysis. It may also be helpful for the reader to refer back to the discussion included right after the statement of Theorem 2.1.1 for the case of $k=2$, to keep in mind an outline of how we aim to proceed in the rest of §2.2.1.

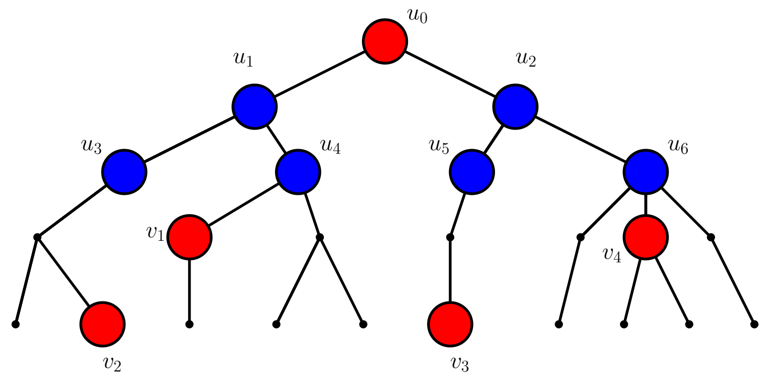


Figure 2.5: Here, $k=2$ and $u_0 \in \mathrm{NL}^{(n+2)}$ (indicated in red), so that $u_i \in \mathrm{NW}^{(n+1)}$ for all $1 \leqslant i \leqslant 6$ (indicated in blue). Since $u_1 \in \mathrm{NW}^{(n+1)}$ and its children u_3 and u_4 are in $\mathrm{NW}^{(n+1)}$, u_1 must have at least one grandchild, say v_1 (indicated in red) in $\mathrm{NL}^{(n)}$. Note that $v_1 \in \mathrm{NL}^{(n)}$ ensures $u_4 \in \mathrm{NW}^{(n+1)}$. Likewise, u_2 and its children u_5 and u_6 are in $\mathrm{NW}^{(n+1)}$, hence u_2 must have at least one grandchild, say v_4, in $\mathrm{NL}^{(n)}$. That $v_4 \in \mathrm{NL}^{(n)}$ ensures $u_6 \in \mathrm{NW}^{(n+1)}$. Let u_3 and u_5 have grandchildren v_2 and v_3 (respectively) in $\mathrm{NL}^{(n)}$, but no child in $\mathrm{NL}^{(n)}$. Then u_4 and u_6 are in $\mathscr{C}_{0,1,n}$, u_3 and u_5 are in $\mathscr{C}_{0,2,n}$, and u_1 and u_2 are in $\mathscr{C}_{1,2,n}$.

For $0 \leqslant i < j \leqslant k$, we define the subsets of vertices

$$\mathscr{C}_{i,j,n} = \{u : \Gamma_i(u) \subset \mathrm{NW}^{(n+1)}, \Gamma_{j-1}(u) \cap \mathrm{NL}^{(n)} = \emptyset, \Gamma_j(u) \cap \mathrm{NL}^{(n)} \neq \emptyset\}. \tag{2.2.6}$$

In other words, any vertex u in $\mathscr{C}_{i,j,n}$ satisfies the following conditions:

- *all* its descendants that are at distance at most i away from it are in $\mathrm{NW}^{(n+1)}$,
- *none* of its descendants at distance at most $j-1$ away from it is in $\mathrm{NL}^{(n)}$,
- and *at least one* of its descendants at distance precisely j away from it is in $\mathrm{NL}^{(n)}$.

Since $j \leqslant k$, the third condition above ensures, via (2.2.4), that $\mathscr{C}_{i,j,n} \subset \mathrm{NW}^{(n+1)}$, so that $\mathscr{C}_{i,j,n} \cap \mathrm{NL}^{(n)} = \emptyset$. The second condition above implies $\mathscr{C}_{i,j,n} \cap \mathscr{C}_{i',j',n} = \emptyset$ for all $0 \leqslant i < j \leqslant k$, $0 \leqslant i' < j' \leqslant k$ and $j \neq j'$. We let $p_{i,j,n}$ denote the probability of the event that the root ϕ of $\mathscr{T}_\chi$ belongs to $\mathscr{C}_{i,j,n}$.

From (2.2.4), we see that $u \in \mathrm{NW}^{(n+1)}$ iff $u \in \bigcup_{j=1}^{k} \mathscr{C}_{0,j,n}$, i.e. either u has a child $v \in \mathrm{NL}^{(n)}$, or u has a child v with at least one descendant w such that $\rho(v,w) \leqslant k-1$ and $w \in \mathrm{NL}^{(n)}$, which means that $v \in \bigcup_{j=1}^{k-1} \mathscr{C}_{0,j,n}$. Henceforth, given that the vertex u has m children, $m \in \mathbb{N}$, we name them $u_1, u_2, \ldots, u_m$. Thus

$$\begin{aligned}
\mathrm{nw}^{(n+1)} &= \sum_{m=1}^{\infty} \mathbf{P}\left[\text{at least one } u_t \in \bigcup_{j=1}^{k-1} \mathscr{C}_{0,j,n} \cup \mathrm{NL}^{(n)} \text{ for } 1 \leqslant t \leqslant m\right] \chi(m) \\
&= \sum_{m=1}^{\infty} \left[1 - \left(1 - \mathrm{n}\ell^{(n)} - \sum_{j=1}^{k-1} p_{0,j,n}\right)^m\right] \chi(m) = 1 - G\left(1 - \mathrm{n}\ell^{(n)} - \sum_{j=1}^{k-1} p_{0,j,n}\right). \quad (2.2.7)
\end{aligned}$$

From (2.2.5), we see that $u \in \mathrm{NL}^{(n+2)}$ if and only if either u is childless, or every child v of u as well as every vertex in $\Gamma_{k-1}(v)$ is in $\mathrm{NW}^{(n+1)}$. However, $v \in \mathrm{NW}^{(n+1)}$ iff some vertex in $\Gamma_k(v)$ is in $\mathrm{NL}^{(n)}$. Thus v must have a descendant w such that $\rho(v,w) = k$ and $w \in \mathrm{NL}^{(n)}$, i.e. $v \in \mathscr{C}_{k-1,k,n}$. Thus

$$\mathrm{n}\ell^{(n+2)} = \sum_{m=0}^{\infty} \mathbf{P}\left[u_t \in \mathscr{C}_{k-1,k,n} \text{ for all } 1 \leqslant t \leqslant m\right] \chi(m) = \sum_{m=0}^{\infty} p_{k-1,k,n}^m \chi(m) = G(p_{k-1,k,n}). \quad (2.2.8)$$

We now establish recurrence relations for the probabilities $p_{i,j,n}$. For a vertex u to be in $\mathscr{C}_{0,1,n}$, it must have at least one child in $\mathrm{NL}^{(n)}$, i.e.

$$p_{0,1,n} = \sum_{m=1}^{\infty} \left[1 - \left(1 - \mathrm{n}\ell^{(n)}\right)^m\right] \chi(m) = 1 - G\left(1 - \mathrm{n}\ell^{(n)}\right). \quad (2.2.9)$$

For $2 \leqslant j \leqslant k$, $u \in \mathscr{C}_{0,j,n}$ iff at least one child of u is in $\mathscr{C}_{0,j-1,n}$ and no child is in $\bigcup_{\ell=1}^{j-2} \mathscr{C}_{0,\ell,n} \cup \mathrm{NL}^{(n)}$:

$$\begin{aligned}
p_{0,j,n} &= \sum_{m=1}^{\infty} \mathbf{P}\left[u_t \notin \bigcup_{\ell=1}^{j-2} \mathscr{C}_{0,\ell,n} \cup \mathrm{NL}^{(n)}, 1 \leqslant t \leqslant m\right] \chi(m) - \sum_{m=1}^{\infty} \mathbf{P}\left[u_t \notin \bigcup_{\ell=1}^{j-1} \mathscr{C}_{0,\ell,n} \cup \mathrm{NL}^{(n)}, 1 \leqslant t \leqslant m\right] \chi(m) \\
&= \sum_{m=1}^{\infty} \left(1 - \mathrm{n}\ell^{(n)} - \sum_{\ell=1}^{j-2} p_{0,\ell,n}\right)^m \chi(m) - \sum_{m=1}^{\infty} \left(1 - \mathrm{n}\ell^{(n)} - \sum_{\ell=1}^{j-1} p_{0,\ell,n}\right)^m \chi(m) \\
&= G\left(1 - \mathrm{n}\ell^{(n)} - \sum_{\ell=1}^{j-2} p_{0,\ell,n}\right) - G\left(1 - \mathrm{n}\ell^{(n)} - \sum_{\ell=1}^{j-1} p_{0,\ell,n}\right). \quad (2.2.10)
\end{aligned}$$

Finally, for a vertex u to be in $\mathscr{C}_{i,j,n}$ for any $1 \leqslant i < j \leqslant k$, the following are necessary:

- *every* child v of u must be in $\mathrm{NW}^{(n+1)}$,
- *every* vertex in $\Gamma_{i-1}(v)$ must be in $\mathrm{NW}^{(n+1)}$, for *every* child v of u,
- *no* vertex in $\Gamma_{j-2}(v)$ is in $\mathrm{NL}^{(n)}$ for *any* child v of u,

- there exists *at least one* child v of u with *at least one* descendant w such that $\rho(v,w) = j-1$ and $w \in \mathrm{NL}^{(n)}$.

The first condition, along with (2.2.4), implies that every child v of u has a descendant in $\Gamma_k(v)$ that is in $\mathrm{NL}^{(n)}$. This, along with the second and third conditions, implies that each child of u must be in $\bigcup_{\ell=j-1}^{k} \mathscr{C}_{i-1,\ell,n}$. The fourth condition implies that at least one child of u must be in $\mathscr{C}_{i-1,j-1,n}$. Thus we have

$$p_{i,j,n} = \sum_{m=1}^{\infty} \mathbf{P}\left[u_t \in \bigcup_{\ell=j-1}^{k} \mathscr{C}_{i-1,\ell,n}, 1 \leqslant t \leqslant m\right] \chi(m) - \sum_{m=1}^{\infty} \mathbf{P}\left[u_t \in \bigcup_{\ell=j}^{k} \mathscr{C}_{i-1,\ell,n}, 1 \leqslant t \leqslant m\right] \chi(m)$$
$$= \sum_{m=1}^{\infty} \left(\sum_{\ell=j-1}^{k} p_{i-1,\ell,n}\right)^m \chi(m) - \sum_{m=1}^{\infty} \left(\sum_{\ell=j}^{k} p_{i-1,\ell,n}\right)^m \chi(m) = G\left(\sum_{\ell=j-1}^{k} p_{i-1,\ell,n}\right) - G\left(\sum_{\ell=j}^{k} p_{i-1,\ell,n}\right). \tag{2.2.11}$$

We now state a couple of lemmas, with their proofs deferred to §2.10.1 of the Appendix, the first of which is concerned with important properties of the function sequence $\{F_i\}_{i\in\mathbb{N}_0}$, whereas the second provides expressions for the probabilities $p_{i,j,n}$ using (2.2.9), (2.2.10) and (2.2.11) (once again, it helps if the reader recalls the discussion for $k = 2$ presented after the statement of Theorem 2.1.1). We then combine and consolidate the recurrence relations in (2.2.7) and (2.2.8) with the conclusion of Lemma 2.2.3 in order to obtain the final result.

Lemma 2.2.2. *Recall the functions F_i defined in (2.1.6). For each $i \in \mathbb{N}$, F_i is a strictly decreasing function on $[0, c_{i-1}]$, and consequently, c_i is uniquely defined. Moreover, $\chi(0) < c_i < c_{i-1}$ for each $i \in \mathbb{N}$.*

Lemma 2.2.3. *We have $p_{i,j,n} = g_{i+1}(F_{j-i-1}(\mathrm{n}\ell^{(n)}), F_{j-i}(\mathrm{n}\ell^{(n)}), F_{k-i+1}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)}))$ for $0 \leqslant i < j \leqslant k$, where F_is and g_is are as defined in (2.1.6) and (2.1.7) respectively.*

From (2.2.8), Lemma 2.2.3 and (2.1.9), we get

$$\mathrm{n}\ell^{(n+2)} = G(g_k(F_0(\mathrm{n}\ell^{(n)}), F_1(\mathrm{n}\ell^{(n)}), F_2(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)}))) = H_k(\mathrm{n}\ell^{(n)}). \tag{2.2.12}$$

Taking the limit as $n \to \infty$ and using (2.2.3), we conclude that $\mathrm{n}\ell$ is a fixed point of H_k (the continuity of the pgf G guarantees the continuity of H_k). By (2.2.7), Lemma 2.2.3 and (2.1.6), we have

$$\mathrm{nw}^{(n+1)} = 1 - G(1 - \mathrm{n}\ell^{(n)} - \sum_{j=1}^{k-1}\{F_{j-1}(\mathrm{n}\ell^{(n)}) - F_j(\mathrm{n}\ell^{(n)})\}) = 1 - F_k(\mathrm{n}\ell^{(n)}), \tag{2.2.13}$$

so that taking the limit as $n \to \infty$ and using (2.2.3), we have $\mathrm{nw} = 1 - F_k(\mathrm{n}\ell)$, as desired.

2.2.2 Showing that $\mathrm{n}\ell_k$ is the minimum positive fixed point of H_k

We begin with an outline for the contents of §2.2.2. The primary intention of §2.2.2 is to establish that $g_k(F_0(x), F_1(x), F_2(x), \dots, F_k(x))$, and consequently H_k, are both increasing on $[0, c_{k-1}]$, for every $k \in \mathbb{N}$. Once established, this fact aids, as follows, in concluding that $\mathrm{n}\ell = \mathrm{n}\ell_k$ is the smallest positive fixed point of H_k.

Let $0 < \eta \leqslant c_k$ be a fixed point of H_k. From (2.2.12), we have $\mathrm{n}\ell^{(2n)} = H_k^{(n)}(\mathrm{n}\ell^{(0)}) = H_k^{(n)}(0)$, where $H_k^{(n)}$ indicates the n-fold composition of H_k with itself. Note that,

1. since c_k is a fixed point of H_k by Corollary 2.4.2,
2. since the proof of Lemma 2.2.2 yields $0 < c_k < c_{k-1}$ and $H_k(0) = \chi(0)$ (which follows from the fact that $F_i(0) = 1$ for all i),
3. and since H_k is increasing on $[0, c_{k-1}]$,

we have $0 < \chi(0) \leqslant H_k^{(n)}(0) = \mathrm{n}\ell^{(2n)} \leqslant H_k^{(n)}(\eta) = \eta \leqslant H_k^{(n)}(c_k) = c_k < c_{k-1}$ for each $n \in \mathbb{N}$. Upon taking the limit as $n \to \infty$ and using (2.2.3), this yields $\mathrm{n}\ell = \lim_{n\to\infty} \mathrm{n}\ell^{(2n)} \leqslant \eta$, thus allowing us to conclude that $\mathrm{n}\ell$ is, indeed, the smallest positive fixed point of H_k.

We prove that $g_k(F_0(x), F_1(x), F_2(x), \dots, F_k(x))$ is increasing on $[0, c_{k-1}]$ by showing that the derivative

$$\frac{d}{dx} g_k\big(F_0(x), F_1(x), F_2(x), \dots, F_k(x)\big) = G'\big(g_{k-1}\big(F_0(x), F_2(x), F_3(x), \dots, F_k(x)\big)\big) \frac{d}{dx} g_{k-1}\big(F_0(x), F_2(x), F_3(x), \dots, F_k(x)\big) - G'\big(g_{k-1}\big(F_1(x), F_2(x), F_3(x), \dots, F_k(x)\big)\big) \frac{d}{dx} g_{k-1}\big(F_1(x), F_2(x), F_3(x), \dots, F_k(x)\big),$$

is non-negative for $x \in [0, c_{k-1}]$. We accomplish this by showing that, for each $x \in [0, c_{k-1}]$,

1. that $G'\big(g_{k-1}\big(F_0(x), F_2(x), F_3(x), \dots, F_k(x)\big)\big) \geqslant G'\big(g_{k-1}\big(F_1(x), F_2(x), F_3(x), \dots, F_k(x)\big)\big)$,
2. that $\frac{d}{dx} g_{k-1}\big(F_0(x), F_2(x), F_3(x), \dots, F_k(x)\big) \geqslant \frac{d}{dx} g_{k-1}\big(F_1(x), F_2(x), F_3(x), \dots, F_k(x)\big)$,
3. and the product $G'\big(g_{k-1}\big(F_0(x), F_2(x), F_3(x), \dots, F_k(x)\big)\big) \frac{d}{dx} g_{k-1}\big(F_0(x), F_2(x), F_3(x), \dots, F_k(x)\big)$ is non-negative.

We note that $G'(x) = \sum_{i=1}^{\infty} i\chi(i) x^{i-1} \geqslant 0$ for all $x \in [0,1]$, so that $G'\big(g_{k-1}\big(1, F_2(x), F_3(x), \dots, F_k(x)\big)\big) \geqslant 0$ for each $x \in [0, c_{k-1}]$. So, to show (3), it suffices to prove that $\frac{d}{dx} g_{k-1}\big(1, F_2(x), F_3(x), \ldots, F_k(x)\big)$ is non-negative. Furthermore, $G''(x) = \sum_{i=2}^{\infty} i(i-1)\chi(i)x^{i-2} \geqslant 0$ for all $x \in [0,1]$, showing that G' is increasing on $[0,1]$. Consequently, to show (1), it suffices to show that $g_{k-1}\big(1, F_2(x), F_3(x), \ldots, F_k(x)\big) \geqslant g_{k-1}\big(F_1(x), F_2(x), F_3(x), \ldots, F_k(x)\big)$.

Each of (1), (2) and (3) will be established if we show that for any $0 \leqslant i_1 < i_2 \leqslant k-j$ and all $x \in [0, c_{k-1}]$,

$$1 \geqslant g_j\big(F_{i_1}(x), F_{k-j+1}(x), F_{k-j+2}(x), \ldots, F_k(x)\big) \geqslant g_j\big(F_{i_2}(x), F_{k-j+1}(x), F_{k-j+2}(x), \ldots, F_k(x)\big) \geqslant 0, \tag{2.2.14}$$

and for $k \geqslant 2$, $1 \leqslant i \leqslant k-j$ and all $x \in [0, c_{k-1}]$, we have

$$\frac{d}{dx} g_j\big(F_0(x), F_{k-j+1}(x), F_{k-j+2}(x), \ldots, F_k(x)\big) \geqslant \max\Big\{0, \frac{d}{dx} g_j\big(F_i(x), F_{k-j+1}(x), F_{k-j+2}(x), \ldots, F_k(x)\big)\Big\}. \tag{2.2.15}$$

The proof of (2.2.14) is deferred to §2.10.1 of the Appendix, whereas (2.2.15) is proved below via induction on j. Setting $j = k-1$, $i_1 = 0$ and $i_2 = 1$ in (2.2.14), we obtain $g_{k-1}\big(F_0(x), F_2(x), F_3(x), \ldots, F_k(x)\big) \in [0,1]$ and $g_{k-1}\big(F_1(x), F_2(x), F_3(x), \ldots, F_k(x)\big) \in [0,1]$, so that $\big(F_0(x), F_1(x), F_2(x), \ldots, F_k(x)\big) \in \mathscr{D}_k$ due to (2.1.8). Consequently, by (2.1.9), we conclude that H_k is well-defined on $[0, c_{k-1}]$.

We now come to the inductive argument for proving (2.2.15). For $1 \leqslant i < k$, since $F_i'(x)$ and $F_k'(x)$ are both negative for $x \in [0, c_{k-1}]$ (due to Lemma 2.2.2), we have

$$\frac{d}{dx} g_1\big(F_0(x), F_k(x)\big) = -F_k'(x) \geqslant \max\big\{0, F_i'(x) - F_k'(x)\big\} = \max\Big\{0, \frac{d}{dx} g_1\big(F_i(x), F_k(x)\big)\Big\},$$

proving (2.2.15) for $j = 1$. Suppose (2.2.15) holds for some $j < k-1$. From (2.1.7), for $x \in [0, c_{k-1}]$, we have

$$\frac{d}{dx} g_{j+1}\big(F_0(x), F_{k-j}(x), F_{k-j+1}(x), \ldots, F_k(x)\big) = G'\big(g_j\big(F_0(x), F_{k-j+1}(x), \ldots, F_k(x)\big)\big) \frac{d}{dx} g_j\big(F_0(x), F_{k-j+1}(x), \\ \ldots, F_k(x)\big) - G'\big(g_j\big(F_{k-j}(x), F_{k-j+1}(x), \ldots, F_k(x)\big)\big) \frac{d}{dx} g_j\big(F_{k-j}(x), F_{k-j+1}(x), \ldots, F_k(x)\big) \geqslant 0, \tag{2.2.16}$$

1. since $G'\big(g_j\big(F_0(x), F_{k-j+1}(x), \ldots, F_k(x)\big)\big) \geqslant G'\big(g_j\big(F_{k-j}(x), F_{k-j+1}(x), \ldots, F_k(x)\big)\big) \geqslant 0$ due to (2.2.14) (setting $i_0 = 0$ and $i_2 = k-j$) and the fact (already shown above) that G' is increasing on $[0,1]$,

2. and since, by the induction hypothesis (2.2.15), we have $\frac{d}{dx}g_j\big(F_0(x),F_{k-j+1}(x),\ldots,F_k(x)\big) \geqslant \max\big\{0,\frac{d}{dx}g_j\big(F_{k-j}(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\big\}$.

Next, for $1 \leqslant i < k-j$, using (2.1.7) once again, we have

$$\begin{aligned}
&\frac{d}{dx}g_{j+1}\big(F_0(x),F_{k-j}(x),F_{k-j+1}(x),\ldots,F_k(x)\big) - \frac{d}{dx}g_{j+1}\big(F_i(x),F_{k-j}(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\\
&= \Big[G'\big(g_j\big(F_0(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\big)\frac{d}{dx}g_j\big(F_0(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\\
&- G'\big(g_j\big(F_{k-j}(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\big)\frac{d}{dx}g_j\big(F_{k-j}(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\Big]\\
&- \Big[G'\big(g_j\big(F_i(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\big)\frac{d}{dx}g_j\big(F_i(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\\
&- G'\big(g_j\big(F_{k-j}(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\big)\frac{d}{dx}g_j\big(F_{k-j}(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\Big]\\
&= G'\big(g_j\big(1,F_{k-j+1}(x),\ldots,F_k(x)\big)\big)\frac{d}{dx}g_j\big(1,F_{k-j+1}(x),\ldots,F_k(x)\big)\\
&- G'\big(g_j\big(F_i(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\big)\frac{d}{dx}g_j\big(F_i(x),F_{k-j+1}(x),\ldots,F_k(x)\big) \geqslant 0,
\end{aligned}$$

1. since $G'\big(g_j\big(1,F_{k-j+1}(x),\ldots,F_k(x)\big)\big) \geqslant G'\big(g_j\big(F_i(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\big) \geqslant 0$ due to (2.2.14) and the fact (already justified above) that G' is increasing on $[0,1]$,

2. and since, by the induction hypothesis (2.2.15), we have $\frac{d}{dx}g_j\big(1,F_{k-j+1}(x),\ldots,F_k(x)\big) \geqslant \max\big\{0,\frac{d}{dx}g_j\big(F_i(x),F_{k-j+1}(x),\ldots,F_k(x)\big)\big\}$.

This completes the proof of (2.2.15) by induction on j, as desired, and it brings us to the end of our proof that nℓ is the minimum positive fixed point of H_k, accomplishing the goal of §2.2.2.

2.3 Proof of Theorem 2.1.2

As the proof of Theorem 2.1.2 closely resembles that of Theorem 2.1.1, we only point out the major modifications. Instead of (2.2.1) and (2.2.2), we now have (omitting the subscript k, as k is fixed throughout §2.3):

$$u \in \mathrm{MW} \Leftrightarrow \Gamma_1(u) = \emptyset \text{ or } \exists\, v \in \Gamma_k(u) \text{ with } v \in \mathrm{ML}, \tag{2.3.1}$$

$$u \in \mathrm{ML} \Leftrightarrow \Gamma_1(u) \neq \emptyset \text{ and } v \in \mathrm{MW} \text{ for every } v \in \Gamma_k(u). \tag{2.3.2}$$

The recursions (2.2.4) and (2.2.5) are also accordingly replaced by

$$u \in \mathrm{MW}^{(n+1)} \Leftrightarrow \Gamma_1(u) = \emptyset \text{ or } \exists\, v \in \Gamma_k(u) \text{ with } v \in \mathrm{ML}^{(n)}, \tag{2.3.3}$$

$$u \in \mathrm{ML}^{(n+1)} \Leftrightarrow \Gamma_1(u) \neq \emptyset \text{ and } v \in \mathrm{MW}^{(n)} \text{ for every } v \in \Gamma_k(u). \tag{2.3.4}$$

Analogous to the classes of vertices $\mathscr{C}_{i,j,n}$ defined in (2.2.6), we now define, for $0 \leqslant i < j \leqslant k$,

$$\mathscr{D}_{i,j,n} = \{u : \Gamma_i(u) \subset \mathrm{MW}^{(n+1)}, \Gamma_{j-1}(u) \cap \mathrm{ML}^{(n)} = \emptyset, \Gamma_j(u) \cap \mathrm{ML}^{(n)} \neq \emptyset\}. \tag{2.3.5}$$

Then $\mathscr{D}_{i,j,n} \cap \mathscr{D}_{i',j',n} = \emptyset$ when $j \neq j'$, and (2.3.3) implies $\mathscr{D}_{i,j,n} \subset \mathrm{MW}^{(n+1)}$, so that $\mathscr{D}_{i,j,n} \cap \mathrm{ML}^{(n)} = \emptyset$. Let $q_{i,j,n}$ be the probability of the event that the root ϕ of $\mathscr{T}_\chi$ belongs to $\mathscr{D}_{i,j,n}$. From (2.3.3) and (2.3.5):

$$\begin{aligned}\mathrm{mw}^{(n+1)} &= \chi(0) + \sum_{m=1}^{\infty} \mathbf{P}[\text{at least one } u_t \in \bigcup_{j=1}^{k-1} \mathscr{D}_{0,j,n} \cup \mathrm{ML}^{(n)} \text{ for } 1 \leqslant t \leqslant m]\chi(m) \\ &= \chi(0) + \sum_{m=1}^{\infty} \{1 - (1 - \mathrm{m}\ell^{(n)} - \sum_{j=1}^{k-1} q_{0,j,n})^m\}\chi(m) = \chi(0) + 1 - G(1 - \mathrm{m}\ell^{(n)} - \sum_{j=1}^{k-1} q_{0,j,n}),\end{aligned} \tag{2.3.6}$$

where, as in §2.2, we denote the m children of ϕ by u_1, u_2, …, u_m. From (2.3.4) and (2.3.5), we have

$$\begin{aligned}\mathrm{m}\ell^{(n+2)} &= \sum_{m=1}^{\infty} \mathbf{P}[u_t \in \mathscr{D}_{k-1,k,n} \text{ or } \Gamma_1(u_t) = \emptyset, \text{ for each } 1 \leqslant t \leqslant m]\chi(m) \\ &= \sum_{m=1}^{\infty} (q_{k-1,k,n} + \chi(0))^m \chi(m) = G(q_{k-1,k,n} + \chi(0)) - \chi(0).\end{aligned} \tag{2.3.7}$$

The recurrence relations for $q_{i,j,n}$, $1 \leqslant i < j \leqslant k$, differs from (2.2.11) in that, for u to be in $\mathscr{D}_{i,j,n}$, each child v of u is either childless or in $\bigcup_{\ell=j-1}^{k} \mathscr{D}_{i-1,\ell,n}$, and at least one child of u must be in $\mathscr{D}_{i-1,j-1,n}$. Thus

$$\begin{aligned}q_{i,j,n} &= \sum_{m=1}^{\infty} \mathbf{P}[u_t \in \bigcup_{\ell=j-1}^{k} \mathscr{D}_{i-1,\ell,n} \text{ or } \Gamma_1(u_t) = \emptyset \text{ for each } 1 \leqslant t \leqslant m]\chi(m) - \sum_{m=1}^{\infty} \mathbf{P}[u_t \in \bigcup_{\ell=j}^{k} \mathscr{D}_{i-1,\ell,n} \text{ or } \\ &\Gamma_1(u_t) = \emptyset \text{ for each } 1 \leqslant t \leqslant m]\chi(m) = G(\chi(0) + \sum_{\ell=j-1}^{k} q_{i-1,\ell,n}) - G(\chi(0) + \sum_{\ell=j}^{k} q_{i-1,\ell,n}).\end{aligned} \tag{2.3.8}$$

Analogous to Lemma 2.2.3, we have the following relation: for every $0 \leqslant i < j \leqslant k$,

$$q_{i,j,n} = \gamma_{i+1}(F_{j-i-1}(\mathrm{m}\ell^{(n)}), F_{j-i}(\mathrm{m}\ell^{(n)}), F_{k-i+1}(\mathrm{m}\ell^{(n)}), \ldots, F_k(\mathrm{m}\ell^{(n)})), \tag{2.3.9}$$

where the functions γ_i are as defined in (2.1.10). From (2.3.7) and (2.3.9), we have $\mathrm{m}\ell^{(n+2)} = J_k\left(\mathrm{m}\ell^{(n)}\right)$, where J_k is as defined in (2.1.11). Taking the limit as $n \to \infty$, we see that $\mathrm{m}\ell$ is a fixed point of J_k. From (2.3.6) and (2.3.9), we have $\mathrm{mw}^{(n+1)} = \chi(0) + 1 - F_k\left(\mathrm{m}\ell^{(n)}\right)$, and taking the limit as $n \to \infty$, we get $\mathrm{mw} = \chi(0) + 1 - F_k(\mathrm{m}\ell)$. The approach to showing that $\mathrm{m}\ell$ is the minimum positive fixed point of J_k is nearly identical to that adopted in §2.2.2, and is therefore omitted.

2.4 Proof of Theorem 2.1.3

Fix any $k \in \mathbb{N}$. Recall that the objective of §2.4 is to show that $\chi(0) < \mathrm{n}\ell_k \leqslant c_k$, and that the probability of draw nd_k is strictly positive if and only if $\mathrm{n}\ell_k < c_k$, in case of k-jump normal games, where c_k is the unique fixed point of F_k in $(0, c_{k-1})$; on the other hand, $\mathrm{m}\ell_k \leqslant \hat{c}_k$, and the probability of draw md_k is strictly positive if and only if $\mathrm{m}\ell_k < \hat{c}_k$, in case of k-jump misère games, where $\hat{c}_k$ is the unique point of intersection between $y = F_k(x)$ and $y = J_k(x) + \chi(0)$ in $(0, c_{k-1})$ (recall the definitions of the functions $F_k(x)$ and $J_k(x)$ from (2.1.6) and (2.1.11) respectively).

Lemma 2.4.1. *For all $k \geqslant 2$, $j \leqslant k$ and $0 \leqslant i \leqslant k - j$, we have*

$$g_j\left(F_i(c_k), F_{k-j+1}(c_k), F_{k-j+2}(c_k), \ldots, F_k(c_k)\right) = F_{j+i-1}(c_k) - c_k. \tag{2.4.1}$$

Proof. When $j = 1$, the left side of (2.4.1) equals $g_1\left(F_i(c_k), F_k(c_k)\right) = F_i(c_k) - c_k$, since c_k is the fixed point of F_k. Suppose (2.4.1) holds for some $j < k$ and all $0 \leqslant i \leqslant k - j$. For $0 \leqslant i \leqslant k - j - 1$, we have

$g_{j+1}(F_i(c_k), F_{k-j}(c_k), F_{k-j+1}(c_k), \ldots, F_k(c_k)) = G(g_j(F_i(c_k), F_{k-j+1}(c_k), \ldots, F_k(c_k))) - G(g_j(F_{k-j}(c_k), F_{k-j+1}(c_k), \ldots, F_k(c_k))) = G\left(F_{j+i-1}(c_k) - c_k\right) - G\left(F_{k-1}(c_k) - c_k\right) = F_{j+i}(c_k) - F_k(c_k) = F_{j+i}(c_k) - c_k$. □

As an immediate consequence of Lemma 2.4.1, we get the following corollary:

Corollary 2.4.2. *For any $k \in \mathbb{N}$, c_k is a fixed point of H_k.*

Proof. We set $j = k$ and $i = 0$ in Lemma 2.4.1 to get $g_k(F_0(c_k), F_1(c_k), \ldots, F_k(c_k)) = F_{k-1}(c_k) - c_k$, which yields $H_k(c_k) = G(g_k(F_0(c_k), F_1(c_k), \ldots, F_k(c_k))) = G(F_{k-1}(c_k) - c_k) = F_k(c_k) = c_k$. □

By Theorem 2.1.1, $\mathrm{n}\ell_k$ is the minimum positive fixed point of H_k, and by Corollary 2.4.2 and Lemma 2.2.2, c_k is a positive fixed point of H_k. Hence $\mathrm{n}\ell_k \leqslant c_k$. The lower bound on $\mathrm{n}\ell_k$ in Theorem 2.1.3 follows simply from observing that if the root has no child, which happens with probability $\chi(0)$, then P1 loses the game starting at the root. From Theorem 2.1.1, we have $\mathrm{nd}_k = 1 - \mathrm{nw}_k - \mathrm{n}\ell_k = F_k(\mathrm{n}\ell_k) - \mathrm{n}\ell_k$. We already know from Lemma 2.2.2 that F_k is strictly decreasing on $[0, c_{k-1}]$ and $c_k \in (0, c_{k-1})$ is its unique fixed point, which is equivalent to saying that for $x \in [0, c_{k-1}]$, we have $F_k(x) - x$ strictly positive if and only if $x < c_k$. Therefore, nd_k is strictly positive if and only if $\mathrm{n}\ell_k < c_k$.

Note that $F_k(0) = 1$ (since the proof of Lemma 2.2.2 yields $F_i(0) = 1$ for all $i \in \mathbb{N}$) and $J_k(0) + \chi(0) = G(\chi(0)) < 1$ as $\chi(0) < 1$ (which also ensures that G is strictly increasing on $[0,1]$), implying that the curve $y = F_k(x)$ lies *above* the curve $y = J_k(x) + \chi(0)$ at $x = 0$. On the other hand, $F_k(c_{k-1}) = G(F_{k-1}(c_{k-1}) - c_{k-1}) = G(0)$ as c_{k-1} is the fixed point of F_{k-1}, whereas $J_k(c_{k-1}) + \chi(0) = G\big(\chi(0) + \gamma_k\big(1, F_1(c_{k-1}), \ldots, F_k(c_{k-1})\big)\big) > G(0)$ since $\chi(0) + \gamma_k\big(1, F_1(c_{k-1}), \ldots, F_k(c_{k-1})\big) > 0$, thus implying that the curve $y = F_k(x)$ lies *below* the curve $y = J_k(x) + \chi(0)$ at $x = c_{k-1}$. Lemma 2.2.2 shows that $F_{k-1}(x)$ is strictly decreasing on $[0, c_{k-1}]$, whereas an argument analogous to that used for showing that H_k is increasing on $[0, c_{k-1}]$ (as outlined in §2.2.2) can be employed to show that J_k, and hence $J_k + \chi(0)$, is increasing on $[0, c_{k-1}]$. Thus $y = F_k(x)$ and $y = J_k(x) + \chi(0)$ intersect at a *unique* point inside $(0, c_{k-1})$, which we call $\hat{c}_k$. From Theorem 2.1.2, we have $\mathrm{md}_k = 1 - \mathrm{mw}_k - \mathrm{m}\ell_k = F_k(\mathrm{m}\ell_k) - \{J_k(\mathrm{m}\ell_k) + \chi(0)\}$. Since $F_k(x)$ is strictly decreasing on $[0, c_{k-1}]$ and $J_k(x) + \chi(0)$ is increasing on $[0, c_{k-1}]$ and they intersect at $\hat{c}_k$, we must have $\mathrm{m}\ell_k \leqslant \hat{c}_k$ to ensure that $\mathrm{md}_k \geqslant 0$, and $\mathrm{md}_k > 0$ iff $\mathrm{m}\ell_k < \hat{c}_k$. This brings us to the end of the proof of Theorem 2.1.3.

2.5 Proof of Theorem 2.1.4

Throughout §2.5, we fix any $k \in \mathbb{N}$ and let the offspring distribution χ of the GW tree $\mathscr{T}_\chi$ be Poisson(λ). In order to emphasize the dependence of all functions and quantities involved on λ, we replace, from the third paragraph of §2.5 onward, all of G, F_i, H_k, g_i, c_i, $\mathrm{n}\ell_k$, nw_k and nd_k by G_λ, $F_{i,\lambda}$, $H_{k,\lambda}$, $g_{i,\lambda}$, $c_{i,\lambda}$, $\mathrm{n}\ell_{k,\lambda}$, $\mathrm{nw}_{k,\lambda}$ and $\mathrm{nd}_{k,\lambda}$ respectively (for all $1 \leqslant i \leqslant k$).

The proof of Lemma 2.2.2 shows that $F_i(0) = 1$ for all $i \in \mathbb{N}$, so that $H_k(0) = \chi(0) > 0$. The curve $y = H_k(x)$ thus lies *above* the curve $y = x$ at $x = 0$. By Theorem 2.1.1, we know that $x = \mathrm{n}\ell_k$ is the smallest positive value of x at which the curve $y = H_k(x)$ either touches or starts going *beneath* the curve $y = x$, so that the slope of $y = H_k(x)$ at $x = \mathrm{n}\ell_k$ has to be less than or equal to the slope of $y = x$. Therefore, we must have $H_k'(\mathrm{n}\ell_k) \leqslant 1$.

The goal of §2.5 is to establish that $H'_{k,\lambda}(c_{k,\lambda}) > 1$ for all λ sufficiently large. This ensures, via the conclusion drawn in the previous paragraph, that $\mathrm{n}\ell_{k,\lambda} \neq c_{k,\lambda}$. By Theorem 2.1.3, we conclude that $\mathrm{n}\ell_{k,\lambda} < c_{k,\lambda}$ and hence $\mathrm{nd}_{k,\lambda} > 0$ for all such values of λ. This would then conclude the proof of the first part of the statement of Theorem 2.1.4.

We outline here the salient steps of the argument employed to prove that $H'_{k,\lambda}(c_{k,\lambda}) > 1$ for all λ sufficiently large. In Lemma 2.5.1, we obtain an expression for the derivative of the function $g_{k,\lambda}\big(r_0(x), r_1(x), \ldots, r_k(x)\big)$ with respect to x, where $r_i(x)$, for $0 \leqslant i \leqslant k$, are differentiable, and $\big(r_0(x), r_1(x), \ldots, r_k(x)\big) \in \mathscr{D}_k$ (which is necessary because of how we define the function g_k in (2.1.7)). Letting $r_i(x)$ be the function $F_{i,\lambda}(x)$ for all $0 \leqslant i \leqslant k$, and using Lemma 2.5.2 that reveals a pattern in the derivatives of the functions $F_{i,\lambda}(x)$, we show (via (2.1.9)) that the leading term in the expansion of $H'_{k,\lambda}(c_{k,\lambda})$ is of the same order of magnitude as $\lambda^{2k} c_{k,\lambda}^2$, while the remaining terms are $O(\lambda^{2k-1} c_{k,\lambda}^2)$. Our final task is to show that

$$\lim_{\lambda \to \infty} \lambda^{k-1} c_{k,\lambda} = 0 \text{ and } \lim_{\lambda \to \infty} \lambda^{k} c_{k,\lambda} = \infty, \tag{2.5.1}$$

which allows us to conclude, in fact, that $\lim_{\lambda \to \infty} H'_{k,\lambda}(c_{k,\lambda}) = \infty$. Note that the second part of the statement of Theorem 2.1.4, asserting $\lambda^{k-1}\mathrm{n}\ell_{k,\lambda} \to 0$ as $\lambda \to \infty$, follows immediately from the first part of (2.5.1) and the fact that $\mathrm{n}\ell_{k,\lambda} \leqslant c_{k,\lambda}$ that we obtain from Theorem 2.1.3.

2.5.1 Understanding the behaviour of $c_{k,\lambda}$ as a function of λ

We begin the proof of (2.5.1) by attempting to understand how $c_{k,\lambda}$ behaves as a function of λ. The first task we accomplish in §2.5.1 is showing that $c_{k,\lambda}$ is, in fact, differentiable with respect to λ, for which we implement the well-known implicit function theorem.

To this end, we *redefine* the functions $F_{i,\lambda}(x)$ on the *extended* interval $[0,1]$ (instead of only on the sub-interval $[0, c_{i-1,\lambda}]$, as done in (2.1.6)) as follows:

$$F_{1,\lambda}(x) = e^{-\lambda x} \text{ and } F_{i+1,\lambda}(x) = \exp\{\lambda F_{i,\lambda}(x) - \lambda x - \lambda\} \text{ for all } x \in [0,1], \text{ for } i \in \mathbb{N}. \tag{2.5.2}$$

Note that these functions are well-defined. It is immediate that $0 < F_{1,\lambda}(x) < 1$ for all $x \in (0,1)$. We now show, via induction on i, that the inequalities $0 < F_{i,\lambda}(x) < 1$ hold for all $x \in (0,1)$, for each $i \in \mathbb{N}$. Suppose we have already shown that $0 < F_{i,\lambda}(x) < 1$ holds for every $x \in (0,1)$, for some $i \in \mathbb{N}$. This yields

$$\lambda F_{i,\lambda}(x) - \lambda x - \lambda < -\lambda x < 0 \implies 0 < F_{i+1,\lambda}(x) < 1 \text{ for all } x \in (0,1), \tag{2.5.3}$$

completing the inductive argument. Next, we note that $F'_{1,\lambda}(x) = -\lambda e^{-\lambda x} < 0$. We show, by induction on i, that $F'_{i,\lambda}(x) < 0$ for all $x \in (0,1)$, for each $i \in \mathbb{N}$. Suppose we have already shown that $F'_{i,\lambda}(x) < 0$ for all $x \in (0,1)$, for some $i \in \mathbb{N}$. We then have

$$F'_{i+1,\lambda}(x) = (\lambda F'_{i,\lambda}(x) - \lambda)F_{i+1,\lambda}(x) < 0 \text{ for all } x \in (0,1).$$

This completes the inductive argument. Let us define $f_i(\lambda,x) = F_{i,\lambda}(x) - x$ on $\Omega = (0,\infty)\times(0,1)$, so that $(\lambda, c_{i,\lambda})$ is a point on the curve $f_i(\lambda,x) = 0$, and $\frac{\partial}{\partial x} f_i(\lambda,x) = F'_{i,\lambda}(x) - 1 < 0$ for all $x \in (0,1)$. By the implicit function theorem, for every $\lambda > 0$, there exists an open $U \times V \subset \Omega$, containing $(\lambda, c_{i,\lambda})$, and a function $h : U \mapsto V$, differentiable on U, such that $c_{i,\lambda} = h(\lambda)$. This concludes our first task, i.e. showing that $c_{k,\lambda}$ is differentiable with respect to λ.

The second task we accomplish in §2.5.1 is showing that $c_{k,\lambda}$ is a strictly decreasing function of λ, by showing that $\frac{d}{d\lambda} c_{k,\lambda}$ is strictly negative for all $\lambda > 0$. For any function $f(x)$ that is defined and differentiable for all $x > 0$, and $0 < f(x) < 1$, we show, by induction on i, that

$$\frac{d}{d\lambda} F_{i,\lambda}(f(\lambda)) = A_i(f(\lambda)) + B_i(f(\lambda)) f'(\lambda) \text{ where } A_i(f(\lambda)) < 0 \text{ and } B_i(f(\lambda)) < 0, \tag{2.5.4}$$

for all $\lambda > 0$ and $i \in \mathbb{N}$. We note that $A_1(f(\lambda)) = -f(\lambda)e^{-\lambda f(\lambda)}$ and $B_1(f(\lambda)) = -\lambda e^{-\lambda f(\lambda)}$, so that the base case is verified. Assuming that (2.5.4) holds for some $i \in \mathbb{N}$, we have

$$\begin{aligned}
\frac{d}{d\lambda} F_{i+1,\lambda}(f(\lambda)) &= \left[F_{i,\lambda}(f(\lambda)) - f(\lambda) - 1 + \lambda \left\{ \frac{d}{d\lambda} F_{i,\lambda}(f(\lambda)) - f'(\lambda) \right\} \right] F_{i+1,\lambda}(f(\lambda)) \\
&= \left[F_{i,\lambda}(f(\lambda)) - f(\lambda) - 1 + \lambda \left\{ A_i(f(\lambda)) + B_i(f(\lambda)) f'(\lambda) - f'(\lambda) \right\} \right] F_{i+1,\lambda}(f(\lambda)) \\
&= \left[F_{i,\lambda}(f(\lambda)) - 1 - f(\lambda) + \lambda A_i(f(\lambda)) \right] F_{i+1,\lambda}(f(\lambda)) + \lambda \left[B_i(f(\lambda)) - 1 \right] F_{i+1,\lambda}(f(\lambda)) f'(\lambda),
\end{aligned}$$

so that

$$A_{i+1}(f(\lambda)) = \left[F_{i,\lambda}(f(\lambda)) - 1 - f(\lambda) + \lambda A_i(f(\lambda)) \right] F_{i+1,\lambda}(f(\lambda))$$

and

$$B_{i+1}(f(\lambda)) = \lambda \left[B_i(f(\lambda)) - 1 \right] F_{i+1,\lambda}(f(\lambda)).$$

These are both negative due to the induction hypothesis and because $0 < F_{i,\lambda}(f(\lambda)) < 1$ for all $\lambda > 0$ (due to (2.5.3)). This completes the proof by induction.

When $f(\lambda) = c_{k,\lambda}$, differentiating both sides of the identity $F_{k,\lambda}(c_{k,\lambda}) = c_{k,\lambda}$ (since $c_{k,\lambda}$, recall, is the unique fixed point of $F_{k,\lambda}$), we have

$$c'_{k,\lambda} = \frac{d}{d\lambda} c_{k,\lambda} = \frac{A_k(c_{k,\lambda})}{1 - B_k(c_{k,\lambda})}.$$

By (2.5.4), we see that the numerator is strictly negative whereas the denominator is strictly positive, thus ensuring that $c'_{k,\lambda} < 0$ for all $\lambda > 0$. This concludes our proof of the fact that $c_{k,\lambda}$ is strictly decreasing in λ for all $\lambda > 0$.

As an immediate consequence of this observation, we note that the limit $\lim_{\lambda \to \infty} c_{k,\lambda}$ exists as $c_{k,\lambda}$ is bounded below by 0 for all $\lambda > 0$. From (2.5.2) and the fact that $c_{k,\lambda}$ is the fixed point of $F_{k,\lambda}$, we obtain

$$F_{k,\lambda}(c_{k,\lambda}) = \exp\left\{\lambda F_{k-1,\lambda}(c_{k,\lambda}) - \lambda c_{k,\lambda} - \lambda\right\} = c_{k,\lambda} \iff F_{k-1,\lambda}(c_{k,\lambda}) - c_{k,\lambda} - 1 = \frac{\ln c_{k,\lambda}}{\lambda}. \tag{2.5.5}$$

If $\lim_{\lambda \to \infty} c_{k,\lambda} = c$ for some $c > 0$, then the right side will go to 0, while the left side remains bounded above by $-c$ since $F_{k-1,\lambda}(c_{k,\lambda}) < 1$ (due to (2.5.3)), yielding a contradiction. Therefore, we must have $\lim_{\lambda \to \infty} c_{k,\lambda} = 0$.

2.5.2 Understanding the behaviour of $F_{j,\lambda}(c_{k,\lambda})$ for all $1 \leqslant j \leqslant k-1$

Before we can establish the claims made in (2.5.1), we need to understand the behaviour of $F_{j,\lambda}(c_{k,\lambda})$ as a function of λ, as $\lambda \to \infty$, for each $1 \leqslant j \leqslant k-1$. To this end, note that, given *any* infinite sequence $\{\lambda_n\}_n$ of positive reals with $\lambda_n \to \infty$, since $0 < F_{k-1,\lambda_n}\left(c_{k,\lambda_n}\right) < 1$ for every $n \in \mathbb{N}$ due to (2.5.3), the Bolzano-Weierstrass Theorem guarantees the existence of an infinite subsequence $\{\lambda_{n_i}\}_i$ such that

$$\lim_{i \to \infty} F_{k-1,\lambda_{n_i}}(c_{k,\lambda_{n_i}}) \text{ exists and is in } [0,1] \implies \lim_{i \to \infty} \frac{\ln c_{k,\lambda_{n_i}}}{\lambda_{n_i}} \text{ exists and is in } [-1,0], \text{ due to (2.5.5).}$$

Suppose we assume that

$$\lim_{i \to \infty} \frac{\ln c_{k,\lambda_{n_i}}}{\lambda_{n_i}} = -c \text{ for some } 0 < c \leqslant 1. \tag{2.5.6}$$

In what follows, our aim is to show that (2.5.6) leads to a contradiction.

2.5.2.1 Proving that (2.5.6) leads to a contradiction

Given any $0 < \varepsilon < c$, (2.5.6) implies that there exists i_ε such that $c_{k,\lambda_{n_i}} < e^{(-c+\varepsilon)\lambda_{n_i}}$ for all $i \geqslant i_\varepsilon$. We then show, using an inductive argument with respect to the index j, that

$$F_{j,\lambda_{n_i}}\left(c_{k,\lambda_{n_i}}\right) > \exp\left\{-\sum_{t=1}^{j} \lambda_{n_i}^t e^{(-c+\varepsilon)\lambda_{n_i}}\right\} \text{ for all } i \geqslant i_\varepsilon, \text{ for } 1 \leqslant j \leqslant k. \tag{2.5.7}$$

For $i \geqslant i_\varepsilon$, we have

$$F_{1,\lambda_{n_i}}\left(c_{k,\lambda_{n_i}}\right) = e^{-\lambda_{n_i} c_{k,\lambda_{n_i}}} > \exp\left\{-\lambda_{n_i} e^{(-c+\varepsilon)\lambda_{n_i}}\right\},$$

so that (2.5.7) holds for $j = 1$, and the base case for the induction is thus verified. Assuming that (2.5.7) holds for some $j < k$ and using $e^{-x} - 1 > -x$ for all $x > 0$, we have, for all $i \geqslant i_\varepsilon$,

$$\begin{aligned}
F_{j+1,\lambda_{n_i}}\left(c_{k,\lambda_{n_i}}\right) &= \exp\left\{\lambda_{n_i} F_{j,\lambda_{n_i}}\left(c_{k,\lambda_{n_i}}\right) - \lambda_{n_i} c_{k,\lambda_{n_i}} - \lambda_{n_i}\right\} \\
&> \exp\left\{\lambda_{n_i} \exp\left\{-\sum_{t=1}^{j} \lambda_{n_i}^t e^{(-c+\varepsilon)\lambda_{n_i}}\right\} - \lambda_{n_i} e^{(-c+\varepsilon)\lambda_{n_i}} - \lambda_{n_i}\right\} \\
&\geqslant \exp\left\{-\lambda_{n_i} \sum_{t=1}^{j} \lambda_{n_i}^t e^{(-c+\varepsilon)\lambda_{n_i}} - \lambda_{n_i} e^{(-c+\varepsilon)\lambda_{n_i}}\right\} = \exp\left\{-\sum_{t=1}^{j+1} \lambda_{n_i}^t e^{(-c+\varepsilon)\lambda_{n_i}}\right\},
\end{aligned}$$

thus proving (2.5.7) by induction. Setting $j = k$, since $c_{k,\lambda_{n_i}}$ is the fixed point of $F_{k,\lambda_{n_i}}$, we have, for $i \geqslant i_\varepsilon$,

$$c_{k,\lambda_{n_i}} = F_{k,\lambda_{n_i}}\left(c_{k,\lambda_{n_i}}\right) > \exp\left\{-\sum_{t=1}^{k} \lambda_{n_i}^t e^{(-c+\varepsilon)\lambda_{n_i}}\right\}.$$

We know from §2.5.1 that the left side of the above inequality goes to 0 as $i \to \infty$, whereas the right side approaches 1, since $\lambda_{n_i}^t e^{(-c+\varepsilon)\lambda_{n_i}} \to 0$ as $\lambda_{n_i} \to \infty$ for every $1 \leqslant t \leqslant k$ (since $-c+\varepsilon < 0$). This yields the desired contradiction.

2.5.2.2 Concluding about the limit of $F_{j,\lambda}(c_{k,\lambda})$, for all $1 \leqslant j \leqslant k-1$

The contradiction obtained in §2.5.2.1 tells us that our assumption in (2.5.6) is wrong, which in turn implies that

$$\lim_{\lambda \to \infty} \frac{\ln c_{k,\lambda}}{\lambda} = 0. \tag{2.5.8}$$

By (2.5.5) and (2.5.8), we conclude that

$$\lim_{\lambda\to\infty} F_{k-1,\lambda}(c_{k,\lambda}) = 1. \tag{2.5.9}$$

Recall that in the proof of (2.2.14), we have shown that $F_{i,\lambda}(x) \geqslant F_{i+1,\lambda}(x)$ for all $0 \leqslant i \leqslant k-1$ and $x \in [0, c_{k-1,\lambda}]$. Using this fact and (2.5.9), we obtain

$$1 > F_{1,\lambda}(c_{k,\lambda}) \geqslant F_{2,\lambda}(c_{k,\lambda}) \geqslant \cdots \geqslant F_{k-1,\lambda}(c_{k,\lambda}) \text{ for } \lambda > 0 \implies \lim_{\lambda\to\infty} F_{i,\lambda}(c_{k,\lambda}) = 1 \text{ for } 1 \leqslant i \leqslant k-1. \tag{2.5.10}$$

2.5.3 The behaviour of $\lambda^i c_{k,\lambda}$, for $1 \leqslant i \leqslant k$, as functions of λ

We lay down the final steps of the proof of (2.5.1). Setting $i = 1$ in (2.5.10) and using (2.5.2), we have $\lim_{\lambda\to\infty} \lambda c_{k,\lambda} = 0$. We show, via induction on i, that

$$F_{i,\lambda}(c_{k,\lambda}) = \exp\{-\lambda^i c_{k,\lambda}(1 + O(\lambda^{-1}))\} \text{ for } 1 \leqslant i \leqslant k-1. \tag{2.5.11}$$

Assuming that (2.5.11) holds for some $i \leqslant k-2$, we have

$$\lim_{\lambda\to\infty} \lambda^i c_{k,\lambda} = 0 \tag{2.5.12}$$

due to (2.5.10), which further ensures that $c_{k,\lambda} = O(\lambda^{-i})$ for all λ large enough. Using this fact, (2.5.2) and a Taylor expansion, we have

$$\begin{aligned}
F_{i+1,\lambda}(c_{k,\lambda}) &= \exp\left\{\lambda\left[e^{-\lambda^i c_{k,\lambda}(1+O(\lambda^{-1}))} - c_{k,\lambda} - 1\right]\right\} \\
&= \exp\left\{\lambda\left[-\lambda^i c_{k,\lambda}(1+O(\lambda^{-1})) + O\left(\left\{\lambda^i c_{k,\lambda}(1+O(\lambda^{-1}))\right\}^2\right) - c_{k,\lambda}\right]\right\} \\
&= \exp\left\{-\lambda^{i+1} c_{k,\lambda}(1+O(\lambda^{-1}))\left[1 + O\left(\lambda^i c_{k,\lambda}(1+O(\lambda^{-1}))\right) + O(\lambda^{-i})\right]\right\} \\
&= \exp\left\{-\lambda^{i+1} c_{k,\lambda}\left[1 + O(\lambda^{-1}) + O(\lambda^i c_{k,\lambda}) + O(\lambda^{-i})\right]\right\}
\end{aligned} \tag{2.5.13}$$

for all λ sufficiently large. Note that each of the terms $O(\lambda^{-1})$, $O(\lambda^i c_{k,\lambda})$ (by (2.5.12)) and $O(\lambda^{-i})$ is $o(1)$ as $\lambda \to \infty$, so that the dominant term in the exponent of (2.5.13) is $-\lambda^{i+1} c_{k,\lambda}$. This fact, along with (2.5.10), yields $\lim_{\lambda\to\infty} \lambda^{i+1} c_{k,\lambda} = 0$. This in turn yields $\lambda^i c_{k,\lambda} = o(\lambda^{-1})$, so that the dominant term out of $O(\lambda^{-1})$, $O(\lambda^i c_{k,\lambda})$ and $O(\lambda^{-i})$ in (2.5.13) is $O(\lambda^{-1})$. This completes the proof of (2.5.11) by induction.

Combining (2.5.10) and (2.5.11) for $i = k-1$, we conclude that $\lim_{\lambda\to\infty} \lambda^{k-1} c_{k,\lambda} = 0$. Using this fact and setting $i = k-1$ in (2.5.13), we have $c_{k,\lambda} = F_{k,\lambda}(c_{k,\lambda}) = \exp\{-\lambda^k c_{k,\lambda}[1+O(\lambda^{-1})+O(\lambda^{k-1}c_{k,\lambda})]\}$, which leads to $\lambda^k c_{k,\lambda} \to \infty$ because of the final conclusion of §2.5.1. The two conclusions drawn in this paragraph bring us to the end of the proof of (2.5.1).

2.5.4 Stating the lemmas and connecting the dots

As promised in the outline of our argument chalked out right before (2.5.1), we now state two important lemmas (whose proofs are deferred to §2.10.2 of the Appendix).

Lemma 2.5.1. *Let $\{r_i\}_{0\leqslant i\leqslant k}$ be a sequence of functions defined and differentiable on an interval I, with $(r_i(x), r_{k-j+1}(x), r_{k-j+2}(x), \ldots, r_k(x)) \in \mathscr{D}_j$ (see (2.1.8)) for all $x \in I$ and all $0 \leqslant i < i+j \leqslant k$. Then*

$$\frac{d}{dx} g_{k,\lambda}(r_0(x), r_1(x), \ldots, r_k(x)) = \lambda^{k-1} \sum_{i=0}^{k-1} f_{k,i,\lambda}(r_0(x), r_1(x), \ldots, r_k(x))(r_i'(x) - r_k'(x)), \quad (2.5.14)$$

in which

$$f_{k,0,\lambda}(r_0(x), r_1(x), \ldots, r_k(x)) = \prod_{t=1}^{k-1} G_\lambda\left(g_{t,\lambda}\left(r_0(x), r_{k-t+1}(x), r_{k-t+2}(x), \ldots, r_k(x)\right)\right) \quad (2.5.15)$$

and

$$f_{k,i,\lambda}(r_0(x), r_1(x), \ldots, r_k(x)) = \prod_{t=1}^{k-i} G_\lambda\left(g_{t,\lambda}(r_i(x), r_{k-t+1}(x), r_{k-t+2}(x), \ldots, r_k(x))\right) \alpha_{k,i,\lambda}\left(r_0(x), r_1(x), \ldots, r_k(x)\right) \quad (2.5.16)$$

for $1 \leqslant i \leqslant k-1$, where $\alpha_{k,1,\lambda}(r_0(x), r_1(x), \ldots, r_k(x)) = -1$ and $\left|\alpha_{k,i,\lambda}(r_0(x), r_1(x), \ldots, r_k(x))\right|$ is bounded above by a constant $a_{k,i}$ that depends on k and i but not on λ nor on the functions $r_0, \ldots, r_k$, for $2 \leqslant i \leqslant k$.

Lemma 2.5.2. *For $i \geqslant 1$ and $x \in (0, c_{i-1})$, we have*

$$F_{i,\lambda}'(x) = -\lambda\left[\sum_{t=1}^{i-1} \lambda^{i-t} \prod_{j=t}^{i-1} F_{j,\lambda}(x) + 1\right] F_{i,\lambda}(x). \quad (2.5.17)$$

We set $r_i(x)$ to be the function $F_{i,\lambda}(x)$ for each i, and $x = c_{k,\lambda}$, in Lemma 2.5.1. Recall, from the outline chalked out right above §2.5.1, that we aim to show that the leading term in (2.5.14), in this case, will be of the same order of magnitude as $\lambda^{2k-1}c_{k,\lambda}$, so that the leading term of $H_{k,\lambda}'(c_{k,\lambda})$ is shown to be of the same order of magnitude as $\lambda^{2k}c_{k,\lambda}^2$. To this end, from (2.4.1), we have

$$G_\lambda(g_{k-i,\lambda}(F_{i,\lambda}(c_{k,\lambda}),F_{i+1,\lambda}(c_{k,\lambda}),\ldots,F_{k,\lambda}(c_{k,\lambda}))) = G_\lambda(F_{k-1,\lambda}(c_{k,\lambda})-c_{k,\lambda}) = F_{k,\lambda}(c_{k,\lambda}) = c_{k,\lambda},$$

so that from (2.5.16), using the fact that both $\left|\alpha_{k,i,\lambda}(r_0(x),r_1(x),\ldots,r_k(x))\right|$ and $G_\lambda(x)$ are $O(1)$, we obtain

$$f_{k,i,\lambda}(r_0(x),r_1(x),\ldots,r_k(x)) = O(c_{k,\lambda}) \text{ for } 1 \leqslant i \leqslant k-1. \tag{2.5.18}$$

Next, from Lemma 2.5.2 and (2.5.10), and using (2.5.18) in the second step, we deduce that

$$F'_{i,\lambda}(c_{k,\lambda}) = O\left(\lambda^i \prod_{j=1}^{i} F_{j,\lambda}(c_{k,\lambda})\right) = O(\lambda^i) \text{ as } \lambda \to \infty, \text{ for } 1 \leqslant i \leqslant k-1$$
$$\Longrightarrow \sum_{i=1}^{k-1} f_{k,i,\lambda}\left(F_{0,\lambda}(c_{k,\lambda}),F_{1,\lambda}(c_{k,\lambda}),\ldots,F_{k,\lambda}(c_{k,\lambda})\right) F'_{i,\lambda}(c_{k,\lambda}) = \sum_{i=1}^{k-1} O\left(\lambda^i c_{k,\lambda}\right). \tag{2.5.19}$$

Next, for any $0<\varepsilon<1$, Lemma 2.5.2, the fact that $c_{k,\lambda}$ is the fixed point of $F_{k,\lambda}$, and (2.5.10) together yield

$$O(\lambda^k c_{k,\lambda}) \geqslant -F'_{k,\lambda}(c_{k,\lambda}) = \lambda^k \prod_{j=1}^{k} F_{j,\lambda}(c_{k,\lambda}) + O\left(\sum_{t=2}^{k} \lambda^{k-t+1} \prod_{j=t}^{k} F_{j,\lambda}(c_{k,\lambda})\right) \geqslant \lambda^k c_{k,\lambda}(1-\varepsilon) + O(\lambda^{k-1}c_{k,\lambda}) \tag{2.5.20}$$

as $\lambda \to \infty$. From (2.5.18) and (2.5.20), we deduce that

$$\sum_{i=1}^{k-1} f_{k,i,\lambda}(F_{0,\lambda}(c_{k,\lambda}),F_{1,\lambda}(c_{k,\lambda}),\ldots,F_{k,\lambda}(c_{k,\lambda}))(-F'_{k,\lambda}(c_{k,\lambda})) = O(\lambda^k c_{k,\lambda}^2). \tag{2.5.21}$$

Finally, (2.5.15), (2.4.1) and (2.5.10) together yield, as $\lambda \to \infty$,

$$f_{k,0,\lambda}\left(F_{0,\lambda}(c_{k,\lambda}),F_{1,\lambda}(c_{k,\lambda}),\ldots,F_{k,\lambda}(c_{k,\lambda})\right) = \prod_{t=1}^{k-1} G_\lambda\left(F_{t-1,\lambda}(c_{k,\lambda})-c_{k,\lambda}\right) = \prod_{t=1}^{k-1} F_{t,\lambda}(c_{k,\lambda}) \to 1. \tag{2.5.22}$$

Hence, from (2.5.20) and (2.5.22), for any $0<\varepsilon<1$, as $\lambda \to \infty$, we have

$$f_{k,0,\lambda}(F_{0,\lambda}(c_{k,\lambda}),F_{1,\lambda}(c_{k,\lambda}),\ldots,F_{k,\lambda}(c_{k,\lambda}))(-F'_{k,\lambda}(c_{k,\lambda})) \geqslant (1-\varepsilon)\lambda^k c_{k,\lambda} + O(\lambda^{k-1}c_{k,\lambda}). \tag{2.5.23}$$

Substituting (2.5.19), (2.5.21) and (2.5.23) in (2.5.14), using the fact that $G_\lambda(x) = e^{\lambda(x-1)}$ so that

$G'_\lambda(x) = \lambda G_\lambda(x)$, and using Corollary 2.4.2, we have, as $\lambda \to \infty$,

$$
\begin{aligned}
H'_{k,\lambda}(c_{k,\lambda}) &= G'_\lambda(g_{k,\lambda}(F_{0,\lambda}(c_{k,\lambda}), F_{1,\lambda}(c_{k,\lambda}), \ldots, F_{k,\lambda}(c_{k,\lambda}))) \frac{d}{dx} g_{k,\lambda}(F_{0,\lambda}(x), F_{1,\lambda}(x), \ldots, F_{k,\lambda}(x))\big|_{x=c_{k,\lambda}} \\
&\geqslant \lambda H_{k,\lambda}(c_{k,\lambda}) \lambda^{k-1} \left[\sum_{i=1}^{k-1} O(\lambda^i c_{k,\lambda}) + O(\lambda^k c_{k,\lambda}^2) + (1-\varepsilon)\lambda^k c_{k,\lambda} + O(\lambda^{k-1} c_{k,\lambda}) \right] \\
&= \lambda^k c_{k,\lambda} \left[O(\lambda^{k-1} c_{k,\lambda}) + O(\lambda^k c_{k,\lambda}^2) + (1-\varepsilon)\lambda^k c_{k,\lambda} \right] = (1-\varepsilon)\lambda^{2k} c_{k,\lambda}^2 + O(\lambda^{2k-1} c_{k,\lambda}^2),
\end{aligned}
$$

so that the leading term of $H'_{k,\lambda}(c_{k,\lambda})$ is indeed of the same order of magnitude as $\lambda^{2k} c_{k,\lambda}^2$, and $H'_{k,\lambda}(c_{k,\lambda}) \to \infty$ due to the second assertion made in (2.5.1). This concludes the proof of Theorem 2.1.4.

2.6 Proof of Theorem 2.1.5

2.6.1 Showing strict convexity of $H_{2,\lambda}$ on $[0, c_{2,\lambda}]$ for $\lambda \geqslant 2$

We begin by stating the first of the three objectives we wish to achieve in §2.6. We fix $k = 2$, and we let the offspring distribution χ of $\mathscr{T}_\chi$ be Poisson(λ). We show that the curve $y = H_{2,\lambda}(x)$ is strictly convex for all $x \in [0, c_{2,\lambda}]$, whenever $\lambda \geqslant 2$ – we accomplish this by proving that the second derivative $H''_{2,\lambda}(x)$ is strictly positive for all $x \in [0, c_{2,\lambda}]$, for every $\lambda \geqslant 2$.

In order to keep the expression for $H''_{2,\lambda}(x)$ as uncluttered as possible, we set

$$\alpha(x) = G_\lambda(1 - F_{2,\lambda}(x)) \text{ and } \beta(x) = G_\lambda(F_{1,\lambda}(x) - F_{2,\lambda}(x)),$$

so that, from (2.1.9), we obtain

$$H_{2,\lambda}(x) = G_\lambda(\alpha(x) - \beta(x)).$$

Note, at the very outset, that since $F_{1,\lambda}(x) \leqslant 1$ (evident from (2.5.2)) and G_λ is increasing, hence $\alpha(x) \geqslant \beta(x)$ for $x \in [0, c_{1,\lambda}]$. From Lemma 2.5.2, which yields $F'_{1,\lambda}(x) = -\lambda F_{1,\lambda}(x)$ and $F'_{2,\lambda}(x) = -\lambda \left(\lambda F_{1,\lambda}(x) + 1\right) F_{2,\lambda}(x)$, we have

$$
\begin{aligned}
\alpha'(x) &= \lambda^2 \alpha(x) \left(\lambda F_{1,\lambda}(x) + 1\right) F_{2,\lambda}(x), \\
\beta'(x) &= \lambda^2 \beta(x) \left\{-F_{1,\lambda}(x) + \left(\lambda F_{1,\lambda}(x) + 1\right) F_{2,\lambda}(x)\right\}.
\end{aligned}
$$

Utilizing these expressions, we have

$$H'_{2,\lambda}(x) = \frac{d}{dx} G_\lambda(\alpha(x) - \beta(x)) = \lambda H_{2,\lambda}(x)(\alpha'(x) - \beta'(x)), \tag{2.6.1}$$

and substituting the expressions for $\alpha'(x)$ and $\beta'(x)$ in (2.6.1), then differentiating again,

$$\begin{aligned}
H''_{2,\lambda}(x) = \lambda^4 H_{2,\lambda}(x) \big[&\lambda^2 \{\alpha(x) - \beta(x)\}^2 \left(\lambda F_{1,\lambda}(x) + 1\right)^2 (F_{2,\lambda}(x))^2 + \lambda^2 (\beta(x))^2 (F_{1,\lambda}(x))^2 + \\
&2\lambda^2 \beta(x) \{\alpha(x) - \beta(x)\} \left(\lambda F_{1,\lambda}(x) + 1\right) F_{1,\lambda}(x) F_{2,\lambda}(x) + \lambda \{\alpha(x) - \beta(x)\} \left(\lambda F_{1,\lambda}(x) + 1\right)^2 (F_{2,\lambda}(x))^2 \\
&- \lambda \{\alpha(x) - \beta(x)\} F_{1,\lambda}(x) F_{2,\lambda}(x) - \{\alpha(x) - \beta(x)\} (\lambda F_{1,\lambda}(x) + 1)^2 F_{2,\lambda}(x) \\
&+ 2\lambda \beta(x) \left(\lambda F_{1,\lambda}(x) + 1\right) F_{1,\lambda}(x) F_{2,\lambda}(x) - \beta(x) F_{1,\lambda}(x) (\lambda F_{1,\lambda}(x) + 1) \big] \qquad (2.6.2) \\
= \lambda^4 H_{2,\lambda}(x) \big[&A_1 + A_2 + 2A_3 + A_4 - A_5 - A_6 + 2A_7 - A_8 \big], \qquad (2.6.3)
\end{aligned}$$

where

$$\begin{aligned}
A_1 &= \lambda^2 \{\alpha(x) - \beta(x)\}^2 \left(\lambda F_{1,\lambda}(x) + 1\right)^2 (F_{2,\lambda}(x))^2, \\
A_2 &= \lambda^2 (\beta(x))^2 (F_{1,\lambda}(x))^2, \\
A_3 &= \lambda^2 \beta(x) \{\alpha(x) - \beta(x)\} \left(\lambda F_{1,\lambda}(x) + 1\right) F_{1,\lambda}(x) F_{2,\lambda}(x), \\
A_4 &= \lambda \{\alpha(x) - \beta(x)\} \left(\lambda F_{1,\lambda}(x) + 1\right)^2 (F_{2,\lambda}(x))^2, \\
A_5 &= \lambda \{\alpha(x) - \beta(x)\} F_{1,\lambda}(x) F_{2,\lambda}(x), \\
A_6 &= \{\alpha(x) - \beta(x)\} (\lambda F_{1,\lambda}(x) + 1)^2 F_{2,\lambda}(x), \\
A_7 &= \lambda \beta(x) \left(\lambda F_{1,\lambda}(x) + 1\right) F_{1,\lambda}(x) F_{2,\lambda}(x), \\
A_8 &= \beta(x) F_{1,\lambda}(x) (\lambda F_{1,\lambda}(x) + 1),
\end{aligned}$$

and each A_i is non-negative (in fact, strictly positive except for A_1, A_3, A_4, A_5 and A_6 at $x = 0$). Thus, our aim now is to establish that the sum within the square brackets in (2.6.3) is strictly positive for each $x \in [0, c_{2,\lambda}]$, for every $\lambda \geqslant 2$. This is where Lemmas 2.6.2 through 2.6.6, all of whose proofs are deferred to §2.10.3 of the Appendix, come in. The objective each of them accomplishes is the collection of various terms from (2.6.3) and showing that their sums are strictly positive for $\lambda \geqslant 2$.

The primary idea we employ here is as follows: we split the interval $[0, c_{2,\lambda}]$ into three pairwise disjoint sub-intervals (in some cases, we may combine two consecutive sub-intervals), namely $[0, \delta_\lambda]$, $(\delta_\lambda, \gamma_\lambda]$ and $(\gamma_\lambda, c_{2,\lambda}]$, where we define γ_λ and δ_λ as follows:

$$F_{2,\lambda}(\gamma_\lambda) = \frac{1}{\lambda} \text{ and } F_{2,\lambda}(\delta_\lambda) = \frac{5}{4\lambda}. \tag{2.6.4}$$

Evidently, to be able to do the above, we require

$$0 < \delta_\lambda < \gamma_\lambda < c_{2,\lambda}. \tag{2.6.5}$$

Several aspects of the above paragraph need justification right away, before we can proceed. By Lemma 2.2.2, we know that $F_{2,\lambda}$ is strictly decreasing on $[0, c_{1,\lambda}]$, and $F_{2,\lambda}(0) = 1$. By definition of $c_{2,\lambda}$, we have $F_{2,\lambda}(c_{2,\lambda}) = c_{2,\lambda}$. Therefore, if we can show that

$$c_{2,\lambda} < \frac{1}{\lambda} < \frac{5}{4\lambda} < 1 \text{ for all } \lambda \geqslant 2, \tag{2.6.6}$$

of which the second and third inequalities are immediately seen to hold, we can conclude that γ_λ and δ_λ, defined via (2.6.4), exist and are unique, and that (2.6.5) holds as well. This is where Lemma 2.6.1 proves useful:

Lemma 2.6.1. *The function $\eta_\lambda = \lambda c_{2,\lambda}$ is strictly increasing for $\lambda \in (0, \lambda_0)$ and strictly decreasing for $\lambda \in (\lambda_0, \infty)$, where $\lambda_0 \approx 2.43634$. The maximum value of η_λ is ≈ 0.52839925.*

The proof of this lemma is deferred to §2.10.3. This lemma guarantees that for all $\lambda > 0$, we have $\eta_\lambda < 1 \implies c_{2,\lambda} < \frac{1}{\lambda}$, thus proving that (2.6.6) indeed holds.

Let us come back to chalking out an outline of our argument for showing that the expression in (2.6.3) is strictly positive for all $\lambda \geqslant 2$: for each $\lambda \geqslant 2$, on each of the sub-intervals $[0, \delta_\lambda]$, $(\delta_\lambda, \gamma_\lambda]$ and $(\gamma_\lambda, c_{2,\lambda}]$, we group the terms within the square brackets in the expression of (2.6.3) judiciously, so that the sum of the terms in each such group is strictly positive. We are now ready to state the lemmas that help accomplish this task.

Lemma 2.6.2. *For each $\lambda \geqslant 2.5$, for all $x \in (\gamma_\lambda, c_{2,\lambda}]$, we have $A_1 + A_4 - A_6 > 0$.*

Lemma 2.6.3. *For each $\lambda \geqslant 2$ and all $x \in (\gamma_\lambda, c_{2,\lambda}]$, we have $A_3 - A_5 > 0$.*

Lemma 2.6.4. *For $\lambda \geqslant 2$ and $x \in (\gamma_\lambda, c_{2,\lambda}]$, we have $A_3 + 2A_7 - A_8 > 0$.*

Before we state the last couple of lemmas, we note that if we add up the expressions to the left of the inequalities in Lemmas 2.6.2, 2.6.3 and 2.6.4, we can conclude that

$$A_1 + 2A_3 + A_4 - A_5 - A_6 + 2A_7 - A_8 > 0 \text{ for all } x \in (\gamma_\lambda, c_{2,\lambda}], \text{ for } \lambda \geqslant 2.5. \tag{2.6.7}$$

Lemma 2.6.5. *For $2 \leqslant \lambda < 2.5$ and $x \in (\gamma_\lambda, c_{2,\lambda}]$, we have $A_1 + A_4 - A_6 + 2A_3 - A_5 + 2A_7 - A_8 > 0$.*

Lemma 2.6.6. *For $\lambda \geqslant 2$ and $x \in [\delta_\lambda, \gamma_\lambda]$, we have $2A_3 - A_5 > 0$.*

Next, we note that as $F_{2,\lambda}$ is strictly decreasing (Lemma 2.2.2), we have $F_{2,\lambda}(x) \geqslant \frac{1}{\lambda}$ for $x \in [0, \gamma_\lambda]$, so that

$$A_4 - A_6 \geqslant 0 \text{ and } 2A_7 - A_8 > 0. \tag{2.6.8}$$

Likewise, we have $F_{2,\lambda}(x) \geqslant \frac{5}{4\lambda}$ for $x \in [0, \delta_\lambda]$, so that

$$A_4 - A_5 - A_6 \geqslant \frac{1}{4}\{\alpha(x) - \beta(x)\}\left(\lambda F_{1,\lambda}(x) + 1\right)^2 F_{2,\lambda}(x) - \lambda\{\alpha(x) - \beta(x)\}F_{1,\lambda}(x)F_{2,\lambda}(x) > 0 \tag{2.6.9}$$

by an application of the AM-GM inequality.

We are now ready to consolidate all of the findings above to achieve the desired conclusion. Combining (2.6.7) and Lemma 2.6.5, we obtain

$$A_1 + 2A_3 + A_4 - A_5 - A_6 + 2A_7 - A_8 > 0 \text{ for all } x \in (\gamma_\lambda, c_{2,\lambda}], \text{ for } \lambda \geqslant 2. \tag{2.6.10}$$

Adding the inequalities in Lemma 2.6.6 and (2.6.8) (for the shorter interval $(\delta_\lambda, \gamma_\lambda]$), we obtain

$$2A_3 - A_5 + A_4 - A_6 + 2A_7 - A_8 > 0 \text{ for all } x \in (\delta_\lambda, \gamma_\lambda], \text{ for } \lambda \geqslant 2. \tag{2.6.11}$$

Adding the inequality in (2.6.9) to the second inequality in (2.6.8) (again, for the shorter interval $[0, \delta_\lambda]$), we obtain

$$2A_7 - A_8 + A_4 - A_5 - A_6 > 0 \text{ for all } x \in [0, \delta_\lambda], \text{ for } \lambda \geqslant 2. \tag{2.6.12}$$

Combining the conclusions of (2.6.10), (2.6.11) and (2.6.12), we complete the proof of the desired claim that the expression in (2.6.3) is strictly positive for all $x \in [0, c_{2,\lambda}]$, for all $\lambda \geqslant 2$. This concludes the proof of the fact that $H_{2,\lambda}$ is strictly convex on $[0, c_{2,\lambda}]$, for all $\lambda \geqslant 2$.

2.6.2 Studying the behaviour of the slope of $H_{2,\lambda}$ at $c_{2,\lambda}$ as a function of λ

The second objective of §2.6 is to show that the slope of $H_{2,\lambda}$ at $x = c_{2,\lambda}$ is strictly increasing in λ for all $\lambda \geqslant 1$. We start by noting, since $c_{2,\lambda}$ is the fixed point of $F_{2,\lambda}$, that $\alpha(c_{2,\lambda}) = F_{1,\lambda}(c_{2,\lambda})$ whereas $\beta(c_{2,\lambda}) = c_{2,\lambda}$ (using (2.1.6)). Using these observations, the expression from (2.6.1), the conclusion of Corollary 2.4.2 and the notation $\eta_\lambda = \lambda c_{2,\lambda}$ introduced in Lemma 2.6.1, we obtain

$$H'_{2,\lambda}(c_{2,\lambda}) = \lambda H_{2,\lambda}(c_{2,\lambda})(\alpha'(c_{2,\lambda}) - \beta'(c_{2,\lambda})) = \lambda^2\eta_\lambda^2 e^{-2\eta_\lambda} + 2\lambda\eta_\lambda^2 e^{-\eta_\lambda} - \lambda\eta_\lambda^3 e^{-\eta_\lambda} - \eta_\lambda^3. \tag{2.6.13}$$

For ease of computation, we perform a term-by-term differentiation of the expression in (2.6.13) with respect to λ, and substitute from (2.10.6) the expression for η'_λ (the derivative of η_λ with respect to λ). This yields

$$\begin{aligned}
\frac{d}{d\lambda}[\lambda^2\eta_\lambda^2 e^{-2\eta_\lambda}] &= \frac{4\lambda\eta_\lambda^2 e^{-2\eta_\lambda} + 2\lambda^2\eta_\lambda^2 e^{-3\eta_\lambda} - 2\lambda^2\eta_\lambda^2 e^{-2\eta_\lambda} + 2\lambda^2\eta_\lambda^3 e^{-2\eta_\lambda}}{1+\lambda\eta_\lambda e^{-\eta_\lambda} + \eta_\lambda};\\
\frac{d}{d\lambda}[2\lambda\eta_\lambda^2 e^{-\eta_\lambda}] &= \frac{6\eta_\lambda^2 e^{-\eta_\lambda} + 4\lambda\eta_\lambda^2 e^{-2\eta_\lambda} - 4\lambda\eta_\lambda^2 e^{-\eta_\lambda} + 2\lambda\eta_\lambda^3 e^{-\eta_\lambda}}{1+\lambda\eta_\lambda e^{-\eta_\lambda} + \eta_\lambda};\\
\frac{d}{d\lambda}[-\lambda\eta_\lambda^3 e^{-\eta_\lambda}] &= \frac{-4\eta_\lambda^3 e^{-\eta_\lambda} - 3\lambda\eta_\lambda^3 e^{-2\eta_\lambda} + 3\lambda\eta_\lambda^3 e^{-\eta_\lambda} - \lambda\eta_\lambda^4 e^{-\eta_\lambda}}{1+\lambda\eta_\lambda e^{-\eta_\lambda} + \eta_\lambda};\\
\frac{d}{d\lambda}[-\eta_\lambda^3] &= \frac{-3\eta_\lambda^3 - 3\lambda\eta_\lambda^3 e^{-\eta_\lambda} + 3\lambda\eta_\lambda^3}{\lambda(1+\lambda\eta_\lambda e^{-\eta_\lambda} + \eta_\lambda)}.
\end{aligned}$$

The numerator of $\frac{d}{d\lambda}H'_{2,\lambda}(c_{2,\lambda})$ is then given by

$$\begin{aligned}
&\eta_\lambda^2[4\lambda^2 e^{-2\eta_\lambda} + 2\lambda^3 e^{-3\eta_\lambda} - 2\lambda^3 e^{-2\eta_\lambda} + 2\lambda^3\eta_\lambda e^{-2\eta_\lambda} + 6\lambda e^{-\eta_\lambda} + 4\lambda^2 e^{-2\eta_\lambda} - 4\lambda^2 e^{-\eta_\lambda} +\\
&2\lambda^2\eta_\lambda e^{-\eta_\lambda} - 4\lambda\eta_\lambda e^{-\eta_\lambda} - 3\lambda^2\eta_\lambda e^{-2\eta_\lambda} + 3\lambda^2\eta_\lambda e^{-\eta_\lambda} - \lambda^2\eta_\lambda^2 e^{-\eta_\lambda} - 3\eta_\lambda - 3\lambda\eta_\lambda e^{-\eta_\lambda} + 3\lambda\eta_\lambda]\\
&= \eta_\lambda^2\{2\lambda^3 e^{-2\eta_\lambda}[e^{-\eta_\lambda} - 1 + \eta_\lambda] + \lambda^2 e^{-\eta_\lambda}[8e^{-\eta_\lambda} - 3\eta_\lambda e^{-\eta_\lambda} - 4 + 5\eta_\lambda - \eta_\lambda^2]\\
&+ \lambda e^{-\eta_\lambda}[6 - 7\eta_\lambda] + 3\eta_\lambda(\lambda - 1)\}\\
&= \eta_\lambda^2\{2\lambda^3 e^{-2\eta_\lambda}[e^{-\eta_\lambda} - 1 + \eta_\lambda] + \lambda^2 e^{-\eta_\lambda}[8e^{-\eta_\lambda} - 4] + \lambda^2 e^{-\eta_\lambda}[5\eta_\lambda - 3\eta_\lambda e^{-\eta_\lambda} - \eta_\lambda^2]\\
&+ \lambda e^{-\eta_\lambda}[6 - 7\eta_\lambda] + 3\eta_\lambda(\lambda - 1)\},
\end{aligned}$$

and our aim now is to prove that this entire expression is strictly positive for $\lambda \geqslant 1$.

From Lemma 2.6.1, we have $8e^{-\eta_\lambda} - 4 \geqslant 8e^{-0.5284} - 4 \approx 0.7164 > 0$ and $6 - 7\eta_\lambda \geqslant 6 - 7 \cdot 0.5284 \approx 2.3012 > 0$. We also note that $5\eta_\lambda - 3\eta_\lambda e^{-\eta_\lambda} - \eta_\lambda^2 = 3\eta_\lambda(1 - e^{-\eta_\lambda}) + \eta_\lambda(1 - \eta_\lambda) + \eta_\lambda > 0$ and $e^{-\eta_\lambda} \geqslant 1 - \eta_\lambda$. Thus $\frac{d}{d\lambda}H'_{2,\lambda}(c_{2,\lambda}) > 0$ for $\lambda \geqslant 1$, hence proving the second part of Theorem 2.1.5.

2.6.3 Proving the third and final part of Theorem 2.1.5

The objective of §2.6.3 is to prove that there exists a critical value λ_c of λ such that $\mathrm{n}\ell_{2,\lambda} = c_{2,\lambda}$ (and hence, by Theorem 2.1.3, $\mathrm{nd}_{2,\lambda} = 0$) for all $2 \leqslant \lambda < \lambda_c$, and $\mathrm{n}\ell_{2,\lambda} < c_{2,\lambda}$ (and hence, $\mathrm{nd}_{2,\lambda} > 0$) for all $\lambda > \lambda_c$.

The argument can be outlined as follows. The convexity of $H_{2,\lambda}$ on $[0, c_{2,\lambda}]$ for $\lambda \geqslant 2$, as proved in §2.6.1, guarantees that $y = H_{2,\lambda}(x)$ intersects $y = x$ at most twice in $[0, c_{2,\lambda}]$. If the intersection happens twice, it must happen at the points $x = \mathrm{n}\ell_{2,\lambda}$ and $x = c_{2,\lambda}$, in which case $\mathrm{n}\ell_{2,\lambda} < c_{2,\lambda}$. If the intersection takes place only once, then this must be at the point $x = c_{2,\lambda}$, in which case $\mathrm{n}\ell_{2,\lambda} = c_{2,\lambda}$. Recall, from §2.5, the justification as to why $H'_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}) \leqslant 1$ for every $\lambda > 0$.

If we can show that $H'_{2,\lambda}(c_{2,\lambda})$ is strictly less than 1 at $\lambda = 2$ and it is strictly greater than 1 for some value of λ strictly exceeding 2, then using our conclusion from §2.6.2, we can deduce that there is precisely one $\lambda_c > 2$ such that $H'_{2,\lambda}(c_{2,\lambda}) < 1$ for $\lambda < \lambda_c$ and $H'_{2,\lambda}(c_{2,\lambda}) > 1$ for $\lambda > \lambda_c$. In the former case, $c_{2,\lambda}$ has to be the only point of intersection between $y = H_{2,\lambda}(x)$ and $y = x$ within the interval $[0, c_{2,\lambda}]$ (this follows by noticing that the curve $y = H_{2,\lambda}(x)$ travels from above $y = x$ to below $y = x$ at $x = \mathrm{n}\ell_{2,\lambda}$, and travels from below $y = x$ to above $y = x$ at $x = c_{2,\lambda}$ when $\mathrm{n}\ell_{2,\lambda} < c_{2,\lambda}$, so that in such a scenario, we would have $H'_{2,\lambda}(c_{2,\lambda}) > 1$). In the latter case, there have to be two points of intersection between $y = H_{2,\lambda}(x)$ and $y = x$ within the interval $[0, c_{2,\lambda}]$.

At $\lambda = 2$, we have $H'_{2,\lambda}(c_{2,\lambda}) \approx 0.721129 < 1$, and at $\lambda = 2.5$, we have $H'_{2,\lambda}(c_{2,\lambda}) \approx 1.06445 > 1$. This accomplishes the task we set for ourselves in the previous paragraph. Solving $H'_{2,\lambda_c}(c_{2,\lambda_c}) = 1$ numerically, we obtain $\lambda_c \approx 2.41$. This accomplishes what we set out to prove in §2.6.3, and concludes the proof of Theorem 2.1.5.

Remark 2.6.7. *Although our proof technique does not cover the range $1 < \lambda < 2$, plotting $H_{2,\lambda}$ on $[0, c_{2,\lambda}]$ for various values of λ in $(1,2)$ does seem to suggest that $H_{2,\lambda}$ is strictly convex on $[0, c_{2,\lambda}]$ for $\lambda \in (1,2)$ (see Figure 2.6). If this were to be proved analytically, it would give us a* unique *threshold for the phase transition phenomenon, showing that $\mathrm{nd}_{2,\lambda} = 0$ for all $0 < \lambda < \lambda_c$ and $\mathrm{nd}_{2,\lambda} > 0$ for all $\lambda > \lambda_c$.*

2.7 Proofs of Theorems 2.1.6 and 2.1.7

Recall that the two objectives we accomplish in §2.7 are 1. showing that $\lambda^i \mathrm{n}\ell_{2,\lambda} \to 0$ as $\lambda \to \infty$, for every $i \in \mathbb{N}$, which is a far stronger claim than the second part of Theorem 2.1.4 as far as $k = 2$ is concerned, 2. establishing the inequalities $\mathrm{m}\ell_{2,\lambda} \leqslant \mathrm{n}\ell_{2,\lambda}$, $\mathrm{nd}_{2,\lambda} < \mathrm{md}_{2,\lambda}$ and $\mathrm{m}\ell_{2,\lambda} \leqslant \mathrm{n}\ell_{2,\lambda} < \mathrm{nw}_{2,\lambda}$ for all λ sufficiently large, giving us a means to compare the probabilities of the various outcomes

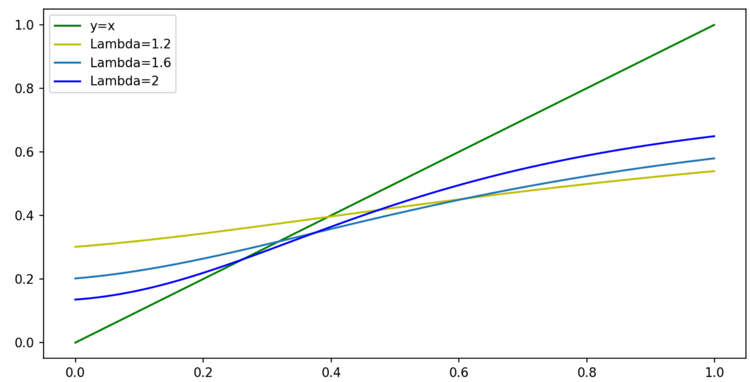


Figure 2.6: $H_{2,\lambda}(x)$ strictly convex for $x \in [0, c_{2,\lambda}]$, for $\lambda = 1.2$ and $\lambda = 1.6$

of the 2-jump normal game with those of the 2-jump misère game when both are played on $\mathscr{T}_\chi$ with χ being Poisson(λ).

2.7.1 Proof of Theorem 2.1.6

We outline the key idea for our argument. First and foremost, we note that the proof is already done for $i = 1$ due to Theorem 2.1.4. For any fixed $i \in \mathbb{N}$ with $i \geqslant 2$, to show that $\lambda^i \mathrm{n}\ell_{2,\lambda} \to 0$ as $\lambda \to \infty$, it suffices to show, for *any arbitrary* $c > 0$, that $\lambda^i \mathrm{n}\ell_{2,\lambda} < c \iff \mathrm{n}\ell_{2,\lambda} < c\lambda^{-i}$ for all λ sufficiently large.

For the same value of $i \geqslant 2$ and $c > 0$ considered above, by the second part of (2.5.1), we know that $\lambda^i c_{2,\lambda} > c \iff c_{2,\lambda} > c\lambda^{-i}$ for all λ large enough. Recall, from Theorem 2.1.5 and the discussion included towards the beginning of §2.5, that $\mathrm{n}\ell_{2,\lambda} < c_{2,\lambda}$ for all λ sufficiently large, and that the curve $y = H_{2,\lambda}(x)$ stays *above* the curve $y = x$ for all $x \in [0, \mathrm{n}\ell_{2,\lambda})$, then stays *beneath* it for all $x \in (\mathrm{n}\ell_{2,\lambda}, c_{2,\lambda})$, for all λ sufficiently large. Therefore, it suffices for us to show that the curve $y = H_{2,\lambda}(x)$ lies beneath the curve $y = x$ at $x = c\lambda^{-i}$, for all λ sufficiently large. In other words,

Lemma 2.7.1. *Fix $c > 0$ and $i \in \mathbb{N}$ with $i \geqslant 2$. Then $H_{2,\lambda}(c\lambda^{-i}) < c\lambda^{-i}$ for all λ sufficiently large.*

The proof of this lemma is deferred to §2.10.4 of the Appendix.

2.7.2 Proof of Theorem 2.1.7

We first show that $\mathrm{m}\ell_{2,\lambda} \leqslant \mathrm{n}\ell_{2,\lambda}$ for all λ sufficiently large. In fact, we prove that the inequality $\mathrm{m}\ell_{2,\lambda}^{(n)} < \mathrm{n}\ell_{2,\lambda}^{(n)}$ holds for all $n \in \mathbb{N}$, via induction on n (recall the definitions of these probabilities from §2.1.3). It is straightforward to see that $\mathrm{m}\ell_{2,\lambda}^{(1)} = 0$ and $\mathrm{n}\ell_{2,\lambda}^{(1)} = e^{-\lambda}$, thus verifying the base case for the inductive argument. Suppose $\mathrm{m}\ell_{2,\lambda}^{(n)} < \mathrm{n}\ell_{2,\lambda}^{(n)}$ for some $n \in \mathbb{N}$. From (2.2.12) and (2.1.9) (which is applied to $k = 2$ and Poisson(λ) offspring distribution), we have

$$\mathrm{n}\ell_{2,\lambda}^{(n+2)} = H_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)}) = G_\lambda\left(g_{2,\lambda}\left(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})\right)\right). \tag{2.7.1}$$

We also note (to be used later) that as $\mathrm{n}\ell_{2,\lambda}$ is a fixed point of $H_{2,\lambda}$ (by Theorem 2.1.1), we have

$$\mathrm{n}\ell_{2,\lambda} = G_\lambda\left(g_{2,\lambda}\left(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda})\right)\right) = \exp\left\{\lambda g_{2,\lambda}\left(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda})\right) - \lambda\right\}. \tag{2.7.2}$$

A relation analogous to (2.2.12) derived in §2.3, (2.1.11) (again, for $k = 2$ and Poisson(λ) offspring), the fact (mentioned in §2.3 and argued the same way as in §2.2.2) that $J_{2,\lambda}$ is increasing on $[0, c_{1,\lambda}]$, the fact (deduced from (2.1.7) and (2.1.10)) that $\gamma_2(1, F_{1,\lambda}(x), F_{2,\lambda}(x)) = e^{\lambda e^{-\lambda}} g_{2,\lambda}(1, F_{1,\lambda}(x), F_{2,\lambda}(x))$, and the induction hypothesis together yield

$$\mathrm{m}\ell_{2,\lambda}^{(n+2)} = J_{2,\lambda}(\mathrm{m}\ell_{2,\lambda}^{(n)}) \leqslant J_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)}) = G_\lambda\left(e^{\lambda e^{-\lambda}} g_{2,\lambda}\left(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})\right) + e^{-\lambda}\right) - e^{-\lambda}. \tag{2.7.3}$$

Therefore, to complete the inductive step of showing that $\mathrm{m}\ell_{2,\lambda}^{(n+2)} < \mathrm{n}\ell_{2,\lambda}^{(n+2)}$, it suffices to show that the expression in (2.7.3) is strictly less than that in (2.7.1) for all λ large enough.

Note that $G_\lambda'(x) = \lambda G_\lambda(x)$ for all $x \in [0,1]$ when we consider Poisson(λ) to be our offspring distribution. Since $G_\lambda''(x) = \sum_{i=2}^{\infty} i(i-1)\chi(i)x^i \geqslant 0$ for $x \in [0,1]$, hence G_λ is convex. Theorem 21.2 of [93] states that a continuously differentiable function f defined on an interval I of $\mathbb{R}$ is convex if and only if $f(y) - f(x) \geqslant f'(x)(y-x)$ for all $x, y \in I$. Applying this result with f replaced by G_λ, we obtain, from (2.7.1) and (2.7.3):

$$\begin{aligned}\mathrm{n}\ell_{2,\lambda}^{(n+2)} - \mathrm{m}\ell_{2,\lambda}^{(n+2)} \geqslant{}& -\lambda G_\lambda\left(e^{\lambda e^{-\lambda}} g_{2,\lambda}\left(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})\right) + e^{-\lambda}\right) \\ & \left[g_{2,\lambda}\left(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})\right)\left(e^{\lambda e^{-\lambda}} - 1\right) + e^{-\lambda}\right] + e^{-\lambda}.\end{aligned} \tag{2.7.4}$$

In what follows, we show that the expression on the right side of (2.7.4) is non-negative for all λ large enough.

2.7.2.1 Bound on the first factor of the first term of (2.7.4)

Since $J_{2,\lambda}$ is increasing on $[0, c_{1,\lambda}]$ and $\mathrm{n}\ell_{2,\lambda}^{(n)} \leqslant \mathrm{n}\ell_{2,\lambda}$ (evident from (2.1.5) and (2.2.3)), given any $\varepsilon > 0$, we have

$$\begin{aligned}
&G_\lambda\left(e^{\lambda e^{-\lambda}} g_{2,\lambda}\left(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})\right) + e^{-\lambda}\right) \leqslant G_\lambda\left(e^{\lambda e^{-\lambda}} g_{2,\lambda}\left(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda})\right) + e^{-\lambda}\right)\\
&= \exp\left\{e^{\lambda e^{-\lambda}}\lambda\left[g_{2,\lambda}\left(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda})\right) - 1\right] + \lambda e^{-\lambda} + \lambda e^{\lambda e^{-\lambda}} - \lambda\right\}\\
&= (\mathrm{n}\ell_{2,\lambda})^{e^{\lambda e^{-\lambda}}} \exp\{\lambda e^{\lambda e^{-\lambda}} + \lambda e^{-\lambda} - \lambda\}, \text{ using (2.7.2);}\\
&\leqslant \mathrm{n}\ell_{2,\lambda} \exp\left\{\lambda \sum_{i=1}^{\infty} \frac{(\lambda e^{-\lambda})^i}{i!} + \lambda e^{-\lambda}\right\}, \quad \text{since } e^{\lambda e^{-\lambda}} > 1 \text{ and } \mathrm{n}\ell_{2,\lambda} < 1 \text{ implies } (\mathrm{n}\ell_{2,\lambda})^{e^{\lambda e^{-\lambda}}} \leqslant \mathrm{n}\ell_{2,\lambda};\\
&= \mathrm{n}\ell_{2,\lambda} \exp\left\{\lambda^2 e^{-\lambda} \sum_{j=0}^{\infty} \frac{(\lambda e^{-\lambda})^j}{(j+1)!} + \lambda e^{-\lambda}\right\} \leqslant \mathrm{n}\ell_{2,\lambda} \exp\left\{\lambda^2 e^{-\lambda} \sum_{j=0}^{\infty} \frac{(\lambda e^{-\lambda})^j}{j!} + \lambda e^{-\lambda}\right\}\\
&\leqslant \mathrm{n}\ell_{2,\lambda} \exp\{\lambda^2 e^{-\lambda} e^{\lambda e^{-\lambda}} + \lambda e^{-\lambda}\} \leqslant \mathrm{n}\ell_{2,\lambda} \exp\{\lambda^2 e^{-\lambda} e^{e^{-1}} + \lambda e^{-\lambda}\} < (1+\varepsilon)\,\mathrm{n}\ell_{2,\lambda} \qquad (2.7.5)
\end{aligned}$$

for all λ sufficiently large, where, in the last line, we utilize the fact that the maximum value of $\lambda e^{-\lambda}$ is e^{-1}, and that both $\lambda^2 e^{-\lambda}$ and $\lambda e^{-\lambda}$ converge to 0 as $\lambda \to \infty$. It is crucial to note that how large λ needs to be for (2.7.5) to hold depends on ε alone, and *not* on n.

Note that, as a step in the derivation of (2.7.5), we obtain the bound $e^{\lambda e^{-\lambda}} - 1 \leqslant \lambda e^{-\lambda} e^{\lambda e^{-\lambda}} \leqslant \lambda e^{-\lambda} e^{e^{-1}}$. This will prove useful in obtaining a bound for the entire expression on the right side of (2.7.4).

2.7.2.2 Bound on the entire right side of (2.7.4)

Applying the bound from (2.7.5), the bound mentioned at the very end of §2.7.2.1, and the rather crude bound $g_{2,\lambda}(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})) \leqslant 1$ to (2.7.4), we obtain

$$\mathrm{n}\ell_{2,\lambda}^{(n+2)} - \mathrm{m}\ell_{2,\lambda}^{(n+2)} \geqslant -\lambda(1+\varepsilon)\,\mathrm{n}\ell_{2,\lambda}[\lambda e^{-\lambda} e^{e^{-1}} + e^{-\lambda}] + e^{-\lambda} = e^{-\lambda}[1 - (1+\varepsilon)\{e^{e^{-1}}\lambda^2\,\mathrm{n}\ell_{2,\lambda} + \lambda\,\mathrm{n}\ell_{2,\lambda}\}] \tag{2.7.6}$$

for all λ sufficiently large. Theorem 2.1.6 guarantees that both $\lambda^2\,\mathrm{n}\ell_{2,\lambda}$ and $\lambda\,\mathrm{n}\ell_{2,\lambda}$ converge to 0 as $\lambda \to \infty$, so that $(1+\varepsilon)\{e^{e^{-1}}\lambda^2\,\mathrm{n}\ell_{2,\lambda} + \lambda\,\mathrm{n}\ell_{2,\lambda}\}$ converges to 0 as well. Consequently, the final expression of (2.7.6) is strictly positive for all λ sufficiently large. Note, once again, that how large λ needs to be for this to happen depends on ε alone, and *not* on n. This completes the inductive argument, showing that for all λ sufficiently large, we have $\mathrm{m}\ell_{2,\lambda}^{(n)} < \mathrm{n}\ell_{2,\lambda}^{(n)}$ for every $n \in \mathbb{N}$. Taking the limit as $n \to \infty$, using (2.2.3) and an analogous relation for the misère games, we deduce that $\mathrm{m}\ell_{2,\lambda} \leqslant \mathrm{n}\ell_{2,\lambda}$, thus completing the proof of the first part of Theorem 2.1.7.

2.7.2.3 Proving that $\mathrm{nd}_{2,\lambda} < \mathrm{md}_{2,\lambda}$ for all λ large enough

Proving $\mathrm{nd}_{2,\lambda} < \mathrm{md}_{2,\lambda}$ for all λ sufficiently large establishes the second claim of Theorem 2.1.7. From §2.4, we know that $\mathrm{nd}_{2,\lambda} = F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}) - \mathrm{n}\ell_{2,\lambda}$, whereas $\mathrm{md}_{2,\lambda} = F_{2,\lambda}(\mathrm{m}\ell_{2,\lambda}) - \mathrm{m}\ell_{2,\lambda} - e^{-\lambda}$. The rest of §2.7.2.3 is dedicated to comparing these two expressions.

We set out to find a suitable upper bound on $\mathrm{nd}_{2,\lambda}$ for all λ sufficiently large. We start by obtaining suitable approximations to a couple of infinite series. From Theorem 2.1.6, we know that $\lambda\,\mathrm{n}\ell_{2,\lambda} \to 0$ as $\lambda \to \infty$, implying that for all λ sufficiently large,

$$\lambda\,\mathrm{n}\ell_{2,\lambda} < i+1 \iff \frac{(\lambda\,\mathrm{n}\ell_{2,\lambda})^{i+1}}{(i+1)!} < \frac{(\lambda\,\mathrm{n}\ell_{2,\lambda})^{i}}{i!} \text{ for all } i \in \mathbb{N}_0 \text{ and } \lim_{\to\infty} \frac{(\lambda\,\mathrm{n}\ell_{2,\lambda})^{i}}{i!} = 0,$$

so that by the well-known alternating series estimation theorem (see, for instance, §8.4 of [97]), we have

$$e^{-\lambda\,\mathrm{n}\ell_{2,\lambda}} - 1 \leqslant -\lambda\,\mathrm{n}\ell_{2,\lambda} + \frac{\lambda^2\,\mathrm{n}\ell_{2,\lambda}^2}{2}. \tag{2.7.7}$$

Applying Theorem 2.1.6 one more time, we know that each of $\lambda^2\,\mathrm{n}\ell_{2,\lambda}$, $\lambda^3\,\mathrm{n}\ell_{2,\lambda}$ and $\lambda\,\mathrm{n}\ell_{2,\lambda}$ approaches 0 in the limit as $\lambda \to \infty$. Consequently, $\left|-\lambda^2\,\mathrm{n}\ell_{2,\lambda} + \lambda^3\,\mathrm{n}\ell_{2,\lambda}^2/2 - \lambda\,\mathrm{n}\ell_{2,\lambda}\right| \to 0$ as $\lambda \to \infty$. Moreover,

$$\lim_{\lambda\to\infty} \frac{\lambda^3\,\mathrm{n}\ell_{2,\lambda}^2}{\lambda^2\,\mathrm{n}\ell_{2,\lambda}} = \lim_{\lambda\to\infty} \lambda\,\mathrm{n}\ell_{2,\lambda} = 0, \text{ again by Theorem 2.1.6}, \tag{2.7.8}$$

thus showing that $\lambda^3\,\mathrm{n}\ell_{2,\lambda}^2 = o(\lambda^2\,\mathrm{n}\ell_{2,\lambda})$ as $\lambda \to \infty$. Therefore, we have $-\lambda^2\,\mathrm{n}\ell_{2,\lambda} + \lambda^3\,\mathrm{n}\ell_{2,\lambda}^2/2 - \lambda\,\mathrm{n}\ell_{2,\lambda} < 0$ for all λ sufficiently large. Applying the alternating series estimation theorem and arguing the same way as above, we deduce that

$$\begin{aligned}\exp\left\{-\lambda^2\,\mathrm{n}\ell_{2,\lambda} + \frac{\lambda^3\,\mathrm{n}\ell_{2,\lambda}^2}{2} - \lambda\,\mathrm{n}\ell_{2,\lambda}\right\} &\leqslant 1 - \lambda^2\,\mathrm{n}\ell_{2,\lambda} + \frac{\lambda^3\,\mathrm{n}\ell_{2,\lambda}^2}{2} - \lambda\,\mathrm{n}\ell_{2,\lambda} \\ &\quad + \frac{1}{2}\left(-\lambda^2\,\mathrm{n}\ell_{2,\lambda} + \frac{\lambda^3\,\mathrm{n}\ell_{2,\lambda}^2}{2} - \lambda\,\mathrm{n}\ell_{2,\lambda}\right)^2. \end{aligned} \tag{2.7.9}$$

It is evident from (2.7.8) that the leading term of $-\lambda^2\,\mathrm{n}\ell_{2,\lambda} + \lambda^3\,\mathrm{n}\ell_{2,\lambda}^2/2 - \lambda\,\mathrm{n}\ell_{2,\lambda}$ is $-\lambda^2\,\mathrm{n}\ell_{2,\lambda}$, so that we can write $(-\lambda^2\,\mathrm{n}\ell_{2,\lambda} + \lambda^3\,\mathrm{n}\ell_{2,\lambda}^2/2 - \lambda\,\mathrm{n}\ell_{2,\lambda})^2 = O(\lambda^4\,\mathrm{n}\ell_{2,\lambda}^2)$. This observation, along with

Theorem 2.1.6, reveals that for any fixed $\varepsilon > 0$, for all λ large enough, we have

$$\lim_{\lambda\to\infty} \frac{\lambda^4 \mathrm{n}\ell_{2,\lambda}^2}{\mathrm{n}\ell_{2,\lambda}} = \lim_{\lambda\to\infty} \lambda^4 \mathrm{n}\ell_{2,\lambda} = 0 \implies \frac{\lambda^3 \mathrm{n}\ell_{2,\lambda}^2}{2} + \frac{1}{2}\left(-\lambda^2 \mathrm{n}\ell_{2,\lambda} + \frac{\lambda^3 \mathrm{n}\ell_{2,\lambda}^2}{2} - \lambda\, \mathrm{n}\ell_{2,\lambda}\right)^2 \leqslant \varepsilon\, \mathrm{n}\ell_{2,\lambda}\,. \tag{2.7.10}$$

We have now gathered all the bounds / approximations we need in order to proceed to find a suitable bound for $\mathrm{nd}_{2,\lambda}$: given any $\varepsilon > 0$, applying (2.7.7), followed by (2.7.9) and finally (2.7.10), we have

$$\begin{aligned}
\mathrm{nd}_{2,\lambda} &= F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}) - \mathrm{n}\ell_{2,\lambda} = \exp\{\lambda e^{-\lambda\, \mathrm{n}\ell_{2,\lambda}} - \lambda\, \mathrm{n}\ell_{2,\lambda} - \lambda\} - \mathrm{n}\ell_{2,\lambda} \\
&\leqslant \exp\left\{-\lambda^2 \mathrm{n}\ell_{2,\lambda} + \frac{\lambda^3 \mathrm{n}\ell_{2,\lambda}^2}{2} - \lambda\, \mathrm{n}\ell_{2,\lambda}\right\} - \mathrm{n}\ell_{2,\lambda} \\
&\leqslant 1 - \lambda^2 \mathrm{n}\ell_{2,\lambda} + \frac{\lambda^3 \mathrm{n}\ell_{2,\lambda}^2}{2} - \lambda\, \mathrm{n}\ell_{2,\lambda} + \frac{1}{2}\left(-\lambda^2 \mathrm{n}\ell_{2,\lambda} + \frac{\lambda^3 \mathrm{n}\ell_{2,\lambda}^2}{2} - \lambda\, \mathrm{n}\ell_{2,\lambda}\right)^2 - \mathrm{n}\ell_{2,\lambda} \\
&\leqslant 1 - \lambda^2 \mathrm{n}\ell_{2,\lambda} - \lambda\, \mathrm{n}\ell_{2,\lambda} - (1-\varepsilon)\, \mathrm{n}\ell_{2,\lambda}\,.
\end{aligned} \tag{2.7.11}$$

This completes our deduction of the desired upper bound on $\mathrm{nd}_{2,\lambda}$ for λ large enough.

We now come to the derivation of a suitable lower bound on $\mathrm{md}_{2,\lambda}$. We first obtain a rather simple bound, using the inequality $e^{-x} > 1 - x$ for all $x > 0$:

$$\begin{aligned}
&\mathrm{md}_{2,\lambda} = F_{2,\lambda}(\mathrm{m}\ell_{2,\lambda}) - \mathrm{m}\ell_{2,\lambda} - e^{-\lambda} = \exp\left\{\lambda e^{-\lambda\, \mathrm{m}\ell_{2,\lambda}} - \lambda\, \mathrm{m}\ell_{2,\lambda} - \lambda\right\} - \mathrm{m}\ell_{2,\lambda} - e^{-\lambda} \\
&\geqslant \exp\{-\lambda^2 \mathrm{m}\ell_{2,\lambda} - \lambda\, \mathrm{m}\ell_{2,\lambda}\} - \mathrm{m}\ell_{2,\lambda} - e^{-\lambda} \geqslant 1 - \lambda^2 \mathrm{m}\ell_{2,\lambda} - \lambda\, \mathrm{m}\ell_{2,\lambda} - \mathrm{m}\ell_{2,\lambda} - e^{-\lambda}.
\end{aligned} \tag{2.7.12}$$

Note that, since we have already proved while concluding §2.7.2.2 that $\mathrm{m}\ell_{2,\lambda} \leqslant \mathrm{n}\ell_{2,\lambda}$ for all λ large enough, we can further bound the expression on the right side of (2.7.12) to obtain $\mathrm{md}_{2,\lambda} \geqslant 1 - (\lambda^2 + \lambda + 1)\, \mathrm{n}\ell_{2,\lambda} - e^{-\lambda}$. However, this simple bound is not quite enough to compare with the bound obtained in (2.7.11) and arrive at our desired conclusion. This is the difficulty we overcome in what follows.

We know that $\mathrm{n}\ell_{2,\lambda}$ is a fixed point of $H_{2,\lambda}$ by Theorem 2.1.1, that $\mathrm{m}\ell_{2,\lambda}$ is a fixed point of $J_{2,\lambda}$ by Theorem 2.1.2, and that $J_{2,\lambda}$ is increasing on $[0, c_{1,\lambda}]$ (mentioned in §2.3 and argued the same way as in §2.2.2). Using the expressions for $J_{2,\lambda}$ and $H_{2,\lambda}$ that we obtain by considering $k = 2$ and Poisson(λ) offspring in (2.1.11) and (2.1.9), we have, given any $\varepsilon > 0$,

$$(\lambda^2 + \lambda + 1)\, \mathrm{m}\ell_{2,\lambda} + e^{-\lambda} = (\lambda^2 + \lambda + 1) J_{2,\lambda}(\mathrm{m}\ell_{2,\lambda}) + e^{-\lambda} \leqslant (\lambda^2 + \lambda + 1) J_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}) + e^{-\lambda}$$

$$
\begin{aligned}
&= (\lambda^2+\lambda+1)G_\lambda\left(e^{\lambda e^{-\lambda}} g_{2,\lambda}(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda})) + e^{-\lambda}\right) - (\lambda^2+\lambda)e^{-\lambda} \\
&= (\lambda^2+\lambda+1)\exp\left\{e^{\lambda e^{-\lambda}}\lambda\left[g_{2,\lambda}(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda})) - 1\right] + \lambda e^{\lambda e^{-\lambda}} + \lambda e^{-\lambda} - \lambda\right\} - (\lambda^2+\lambda)e^{-\lambda} \\
&< (\lambda^2+\lambda+1)\exp\left\{e^{\lambda e^{-\lambda}}\lambda\left[g_{2,\lambda}(1, F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}), F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda})) - 1\right] + \lambda e^{-\lambda}\right\} - (\lambda^2+\lambda)e^{-\lambda}\left(\text{as } e^{\lambda e^{-\lambda}} > 1\right); \\
&= (\lambda^2+\lambda+1)(\mathrm{n}\ell_{2,\lambda})^{e^{\lambda e^{-\lambda}}} e^{\lambda e^{-\lambda}} - (\lambda^2+\lambda)e^{-\lambda} < (\lambda^2+\lambda+1)(1+\varepsilon)\,\mathrm{n}\ell_{2,\lambda} - (\lambda^2+\lambda)e^{-\lambda}.
\end{aligned}
\tag{2.7.13}
$$

for all λ large enough. From (2.7.12) and (2.7.13), we obtain

$$\mathrm{md}_{2,\lambda} > 1 - (\lambda^2+\lambda+1)(1+\varepsilon)\,\mathrm{n}\ell_{2,\lambda} + (\lambda^2+\lambda)e^{-\lambda}, \tag{2.7.14}$$

so that from (2.7.11) and (2.7.14), we have

$$\mathrm{md}_{2,\lambda} - \mathrm{nd}_{2,\lambda} > -(\lambda^2+\lambda+2)\varepsilon\,\mathrm{n}\ell_{2,\lambda} + (\lambda^2+\lambda)e^{-\lambda}, \tag{2.7.15}$$

and our objective is to show that the expression on the right side of (2.7.15) is non-negative for all λ sufficiently large. Given any $\varepsilon > 0$, we have $\lambda^2+\lambda+2 < (1+\varepsilon)(\lambda^2+\lambda)$ for all λ large enough, and from Theorem 2.1.9 (which is yet to be proved), we have $\mathrm{n}\ell_{2,\lambda} \leqslant \mathrm{n}\ell_{1,\lambda}$ for all λ large enough. To establish (2.7.15), it thus suffices to prove that, given any $\varepsilon > 0$ such that $\varepsilon(1+\varepsilon) < 1$,

$$\frac{e^{-\lambda}}{\varepsilon(1+\varepsilon)} > \mathrm{n}\ell_{1,\lambda} \text{ for all } \lambda \text{ sufficiently large.} \tag{2.7.16}$$

To establish (2.7.16), we need to deduce certain properties of $\mathrm{n}\ell_{1,\lambda}$, which we do by examining the function $H_{1,\lambda}$ (evident from Theorem 2.1.1). Setting $k = 1$ and χ to be Poisson(λ) in (2.1.9), we obtain

$$H_{1,\lambda}(x) = e^{-\lambda e^{-\lambda x}} \implies \lambda^3(\lambda e^{-\lambda x} - 1)e^{-\lambda x - \lambda e^{-\lambda x}}. \tag{2.7.17}$$

Note that, using (2.5.2), for all $\lambda > e$, we have

$$F_{1,\lambda}\left(\frac{\ln\lambda}{\lambda}\right) = \exp\left\{-\lambda\cdot\frac{\ln\lambda}{\lambda}\right\} = \frac{1}{\lambda} < \frac{\ln\lambda}{\lambda},$$

and as Lemma 2.2.2 states that $F_{1,\lambda}$ is strictly decreasing on $[0,1]$ and its unique fixed point is $c_{1,\lambda}$, we conclude that

$$c_{1,\lambda} < \frac{\ln\lambda}{\lambda} \text{ for all } \lambda > e \implies \lambda e^{-\lambda x} > 1 \text{ for all } x \in [0, c_{1,\lambda}], \text{ for all } \lambda > e.$$

This finding, when applied to (2.7.17), reveals that $H_{1,\lambda}$ is strictly convex on $[0, c_{1,\lambda}]$ for all $\lambda > e$. Consequently, the curve $y = H_{1,\lambda}(x)$ intersects the line $y = x$ at at most two points inside the interval $[0, c_{1,\lambda}]$ – if there are two intersections, then these happen at $\mathrm{n}\ell_{1,\lambda}$ and $c_{1,\lambda}$, and if there is only one intersection, then this happens at $c_{1,\lambda}$ (in which case $\mathrm{n}\ell_{1,\lambda} = c_{1,\lambda}$). This conclusion, along with Theorem 2.1.4 whose proof guarantees that $\mathrm{n}\ell_{1,\lambda} < c_{1,\lambda}$ for all λ large enough, reveals that the curve $y = H_{1,\lambda}(x)$ stays *above* the line $y = x$ when $x \in [0, \mathrm{n}\ell_{1,\lambda})$, and *beneath* it when $x \in (\mathrm{n}\ell_{1,\lambda}, c_{1,\lambda})$, for λ sufficiently large.

We now note that $\lambda c_{1,\lambda} \to \infty$ as $\lambda \to \infty$, due to (2.5.1), so that $\frac{e^{-\lambda}}{\varepsilon(1+\varepsilon)} < c_{1,\lambda}$ for all λ sufficiently large. Therefore, to establish (2.7.16), it suffices to show that $y = H_{1,\lambda}(x)$ lies beneath $y = x$ at $x = \frac{e^{-\lambda}}{\varepsilon(1+\varepsilon)}$. Using $x > 1 - e^{-x}$ for all $x > 0$, we have

$$\begin{aligned} e^{\lambda} H_{1,\lambda}\left(\frac{e^{-\lambda}}{\varepsilon(1+\varepsilon)}\right) &= e^{\lambda} \exp\left\{-\lambda \exp\left\{-\lambda \cdot \frac{e^{-\lambda}}{\varepsilon(1+\varepsilon)}\right\}\right\} \\ &= \exp\left\{\lambda\left[1 - \exp\left\{-\frac{\lambda e^{-\lambda}}{\varepsilon(1+\varepsilon)}\right\}\right]\right\} < \exp\left\{\frac{\lambda^2 e^{-\lambda}}{\varepsilon(1+\varepsilon)}\right\}. \end{aligned}$$

Since the right side of the above inequality goes to 1 as $\lambda \to \infty$, and we chose ε above such that $\varepsilon(1+\varepsilon) < 1$, the right side of the above inequality is strictly less than $\frac{1}{\varepsilon(1+\varepsilon)}$ for all λ sufficiently large, thus proving that $H_{1,\lambda}\left(\frac{e^{-\lambda}}{\varepsilon(1+\varepsilon)}\right) < \frac{e^{-\lambda}}{\varepsilon(1+\varepsilon)}$. In other words, we have proved that $y = H_{1,\lambda}(x)$ lies beneath $y = x$ at $x = \frac{e^{-\lambda}}{\varepsilon(1+\varepsilon)}$, as desired. This concludes §2.7.2.3, and with it, the proof of the fact that $\mathrm{nd}_{2,\lambda} < \mathrm{md}_{2,\lambda}$ for all λ large enough.

2.7.2.4 Proving that $\mathrm{n}\ell_{2,\lambda} < \mathrm{nw}_{2,\lambda}$ for all λ large enough

Proving $\mathrm{n}\ell_{2,\lambda} < \mathrm{nw}_{2,\lambda}$ for all λ sufficiently large establishes the third and final claim made in the statement of Theorem 2.1.7. By Theorem 2.1.1, we have $\mathrm{nw}_{2,\lambda} = 1 - F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda})$, so that it suffices to show that $F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}) < 1 - \mathrm{n}\ell_{2,\lambda}$. From (2.5.2), we obtain

$$\begin{aligned} F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}) &= \exp\{\lambda e^{-\lambda\, \mathrm{n}\ell_{2,\lambda}} - \lambda\, \mathrm{n}\ell_{2,\lambda} - \lambda\} \\ &\leqslant \exp\left\{-\lambda^2 \mathrm{n}\ell_{2,\lambda} + \frac{\lambda^3 \mathrm{n}\ell_{2,\lambda}^2}{2} - \lambda\, \mathrm{n}\ell_{2,\lambda}\right\}, \text{ by (2.7.7);} \\ &\leqslant 1 - \lambda^2 \mathrm{n}\ell_{2,\lambda} + \frac{\lambda^3 \mathrm{n}\ell_{2,\lambda}^2}{2} - \lambda\, \mathrm{n}\ell_{2,\lambda} + \frac{1}{2}\left(-\lambda^2 \mathrm{n}\ell_{2,\lambda} + \frac{\lambda^3 \mathrm{n}\ell_{2,\lambda}^2}{2} - \lambda\, \mathrm{n}\ell_{2,\lambda}\right)^2, \text{ by (2.7.9);} \\ &\leqslant 1 - \lambda^2 \mathrm{n}\ell_{2,\lambda} - \lambda\, \mathrm{n}\ell_{2,\lambda} + \varepsilon\, \mathrm{n}\ell_{2,\lambda} \text{ by (2.7.10);} \end{aligned}$$

and this is less than $1-\mathrm{n}\ell_{2,\lambda}$ for all λ large enough since $\mathrm{n}\ell_{2,\lambda} = o\left(\lambda^2 \mathrm{n}\ell_{2,\lambda}\right)$. This concludes §2.7.2.4 and the proof of the entire Theorem 2.1.7.

2.8 Proof of Theorem 2.1.9

Recall that the objectives of §2.8 is to show that the inequalities $\mathrm{n}\ell_{2,\lambda} \leqslant \mathrm{n}\ell_{1,\lambda}$, $\mathrm{nd}_{2,\lambda} < \mathrm{nd}_{1,\lambda}$ and $\mathrm{nw}_{1,\lambda} < \mathrm{nw}_{2,\lambda}$ hold for all λ sufficiently large.

2.8.1 Showing that $\mathrm{n}\ell_{2,\lambda} \leqslant \mathrm{n}\ell_{1,\lambda}$ for all λ sufficiently large

We begin with the first inequality, and we establish this by first showing that $\mathrm{n}\ell_{2,\lambda}^{(n)} \leqslant \mathrm{n}\ell_{1,\lambda}^{(n)}$ for all $n \in \mathbb{N}$, which is proved by induction on n, then taking the limit as $n \to \infty$ and using (2.2.3).

Since $\mathrm{n}\ell_{1,\lambda}^{(1)} = \mathrm{n}\ell_{2,\lambda}^{(1)} = e^{-\lambda}$, the base case for the inductive argument is verified. Suppose $\mathrm{n}\ell_{2,\lambda}^{(n)} \leqslant \mathrm{n}\ell_{1,\lambda}^{(n)}$ for some $n \in \mathbb{N}$. From (2.2.12), we have $\mathrm{n}\ell_{1,\lambda}^{(n+2)} = H_{1,\lambda}(\mathrm{n}\ell_{1,\lambda}^{(n)})$ and $\mathrm{n}\ell_{2,\lambda}^{(n+2)} = H_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})$, so that our task now is to compare these two quantities. To this end, we observe, setting $k=2$ and considering Poisson(λ) offspring in (2.1.9),

$$\begin{aligned}
H_{2,\lambda}(x) &= G_\lambda\left(G_\lambda(1-F_{2,\lambda}(x)) - G_\lambda(F_{1,\lambda}(x)-F_{2,\lambda}(x))\right) \\
&= G_\lambda\left(e^{-\lambda F_{2,\lambda}(x)} - e^{\lambda F_{1,\lambda}(x)-\lambda F_{2,\lambda}(x)-\lambda}\right) = G_\lambda\left(e^{-\lambda F_{2,\lambda}(x)}\left[1-e^{\lambda F_{1,\lambda}(x)-\lambda}\right]\right).
\end{aligned}$$

Applying this, the induction hypothesis and the increasing nature of $H_{1,\lambda}$ on $[0,1]$ (as proved in §2.2.2), we obtain

$$\begin{aligned}
\mathrm{n}\ell_{1,\lambda}^{(n+2)} - \mathrm{n}\ell_{2,\lambda}^{(n+2)} &\geqslant H_{1,\lambda}\left(\mathrm{n}\ell_{2,\lambda}^{(n)}\right) - H_{2,\lambda}\left(\mathrm{n}\ell_{2,\lambda}^{(n)}\right) \\
&= G_\lambda\left(1-F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})\right) - G_\lambda\left(e^{-\lambda F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})}\left[1-e^{\lambda F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})-\lambda}\right]\right). \qquad (2.8.1)
\end{aligned}$$

We now construct a non-negative lower bound for the expression in (2.8.1).

Since $\mathrm{n}\ell_{2,\lambda}^{(n)} \leqslant \mathrm{n}\ell_{2,\lambda}$ (evident from (2.1.5) and (2.2.3)) and since $F_{2,\lambda}$ is strictly decreasing (by Lemma 2.2.2), given any $0 < \varepsilon < 1$, we have, using $e^{-x} - 1 > -x$ for all $x > 0$,

$$F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)}) \geqslant F_{2,\lambda}(\mathrm{n}\ell_{2,\lambda}) = \exp\{\lambda e^{-\lambda\,\mathrm{n}\ell_{2,\lambda}} - \lambda\,\mathrm{n}\ell_{2,\lambda} - \lambda\} \geqslant \exp\{-\lambda^2\,\mathrm{n}\ell_{2,\lambda} - \lambda\,\mathrm{n}\ell_{2,\lambda}\} > 1-\varepsilon \qquad (2.8.2)$$

for all λ sufficiently large, since both $\lambda^2\,\mathrm{n}\ell_{2,\lambda}$ and $\lambda\,\mathrm{n}\ell_{2,\lambda}$ converges to 0 as $\lambda \to \infty$ due to Theorem 2.1.6. Note, crucially, that how large λ needs to be for (2.8.2) to hold depends on ε alone, and not on n. Using (2.8.2) and $1-e^{-x}<x$ for all $x>0$, the expression on the right side of (2.8.1) is bounded below by

$$\begin{aligned}&G_\lambda\left(1-F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})\right)-G_\lambda\left(e^{-\lambda(1-\varepsilon)}\left[1-e^{\lambda F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})-\lambda}\right]\right)\\&\geqslant G_\lambda\left(1-F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})\right)-G_\lambda\left(e^{-\lambda(1-\varepsilon)}\lambda\{1-F_{1,\lambda}(\mathrm{n}\ell_{2,\lambda}^{(n)})\}\right),\end{aligned}$$

and this is strictly positive for all λ sufficiently large since $\lambda e^{-\lambda(1-\varepsilon)}\to 0$ as $\lambda\to\infty$ (and hence eventually becomes strictly less than 1) and G_λ is strictly increasing. This completes the inductive proof and brings us to the end of §2.8.1.

2.8.2 Showing that $\mathrm{nd}_{2,\lambda}<\mathrm{nd}_{1,\lambda}$ and $\mathrm{nw}_{1,\lambda}<\mathrm{nw}_{2,\lambda}$ for all λ sufficiently large

From Theorem 2.1.1, we have the following lower bound on $\mathrm{nd}_{1,\lambda}$:

$$\mathrm{nd}_{1,\lambda}=F_{1,\lambda}(\mathrm{n}\ell_{1,\lambda})-\mathrm{n}\ell_{1,\lambda}=e^{-\lambda\,\mathrm{n}\ell_{1,\lambda}}-\mathrm{n}\ell_{1,\lambda}\geqslant 1-\lambda\,\mathrm{n}\ell_{1,\lambda}-\mathrm{n}\ell_{1,\lambda},\tag{2.8.3}$$

and comparing this with the lower bound on $\mathrm{nd}_{2,\lambda}$ obtained in (2.7.11), we see that the objective of §2.8.2 will be accomplished if we can show that $\mathrm{n}\ell_{1,\lambda}<\lambda\,\mathrm{n}\ell_{2,\lambda}$ for all λ sufficiently large.

Recall that we estabilshed in §2.7.2.3 that $H_{1,\lambda}$ is strictly convex on $[0,c_{1,\lambda}]$ for $\lambda>e$, so that $y=H_{1,\lambda}(x)$ lies *above* $y=x$ on $[0,\mathrm{n}\ell_{1,\lambda})$, and *beneath* $y=x$ on $(\mathrm{n}\ell_{1,\lambda},c_{1,\lambda})$, for all λ sufficiently large. From (2.5.1), we have $\lambda c_{1,\lambda}\to\infty$, whereas Theorem 2.1.6 yields $\lambda^2\,\mathrm{n}\ell_{2,\lambda}\to 0$, so that $\lambda\,\mathrm{n}\ell_{2,\lambda}<c_{1,\lambda}$ for all λ sufficiently large. The goal set in the previous paragraph will, therefore, be accomplished, if we can show that $y=H_{1,\lambda}(x)$ lies beneath $y=x$ at $x=\lambda\,\mathrm{n}\ell_{2,\lambda}$.

Using $e^{-x}>1-x$ for $x>0$, we have

$$H_{1,\lambda}(\lambda\,\mathrm{n}\ell_{2,\lambda})=\exp\left\{-\lambda e^{-\lambda^2\mathrm{n}\ell_{2,\lambda}}\right\}<\exp\left\{-\lambda+\lambda^3\,\mathrm{n}\ell_{2,\lambda}\right\}<\lambda e^{-\lambda}=\lambda\,\mathrm{n}\ell_{2,\lambda}^{(1)}\leqslant\lambda\,\mathrm{n}\ell_{2,\lambda}$$

for all λ sufficiently large, since we have $\lambda^3\,\mathrm{n}\ell_{2,\lambda}\to 0$ by Theorem 2.1.6. This completes the proof of $\mathrm{nd}_{1,\lambda}>\mathrm{nd}_{2,\lambda}$ for all λ sufficiently large.

Since $\mathrm{nw}_{1,\lambda}=1-\mathrm{n}\ell_{1,\lambda}-\mathrm{nd}_{1,\lambda}$ and $\mathrm{nw}_{2,\lambda}=1-\mathrm{n}\ell_{2,\lambda}-\mathrm{nd}_{2,\lambda}$, it follows immediately from the previous conclusion and the conclusion drawn in §2.8.1 that $\mathrm{nw}_{1,\lambda}<\mathrm{nw}_{2,\lambda}$ for all λ sufficiently large.

2.9 Proof of Theorem 2.1.10

Recall that we intend to show that when $\mathrm{n}\ell_k = c_k$ and $\max\left\{H_k'(c_k), \left|F_k'(c_k)\right|\right\} < 1$, the expected duration of the k-jump normal game is finite. Here, unlike the previous few sections, we consider *any* offspring distribution χ that satisfies the restrictions discussed in §2.1.3.

Suppose $\mathrm{n}\ell_k = c_k$, $H_k'(c_k) \leqslant \gamma$ and $|F_k'(c_k)| \leqslant \gamma$ for some $\gamma < 1$. Given $0 < \varepsilon < 1-\gamma$, from (2.2.3) and the continuity of H_k' and F_k' (due to the continuity of G'), we know there exists $N \in \mathbb{N}$ such that $H_k'(x) \leqslant \gamma+\varepsilon$ and $|F_k'(x)| \leqslant \gamma+\varepsilon$ for all $\mathrm{n}\ell_k^{(N)} \leqslant x \leqslant \mathrm{n}\ell_k = c_k$. Using (2.2.12) and (2.1.5), for all $n \geqslant N$, we have

$$\mathrm{n}\ell_k - \mathrm{n}\ell_k^{(n+2)} = H_k(\mathrm{n}\ell_k) - H_k(\mathrm{n}\ell_k^{(n)}) = H_k'(\xi_n)(\mathrm{n}\ell_k - \mathrm{n}\ell_k^{(n)}) \leqslant (\varepsilon+\gamma)(\mathrm{n}\ell_k - \mathrm{n}\ell_k^{(n)}), \tag{2.9.1}$$

where $\mathrm{n}\ell_k^{(n)} < \xi_n < \mathrm{n}\ell_k$ (by the mean value theorem). Likewise, from (2.2.13) and (2.1.5), for all $n \geqslant N$, we have

$$\mathrm{nw}_k - \mathrm{nw}_k^{(n+2)} \leqslant \mathrm{nw}_k - \mathrm{nw}_k^{(n+1)} = F_k(\mathrm{n}\ell_k^{(n)}) - F_k(\mathrm{n}\ell_k) \leqslant (\varepsilon+\gamma)(\mathrm{n}\ell_k - \mathrm{n}\ell_k^{(n)}). \tag{2.9.2}$$

Denoting the (random) duration of the game (starting at the root ϕ of $\mathscr{T}_\chi$) by T, recalling the definition of $\mathrm{ND}_k^{(n)}$ from §2.1.3, and letting $C = \sum_{n=1}^{N+1}(1-\mathrm{nw}_k^{(n)} - \mathrm{n}\ell_k^{(n)})$, we have

$$\begin{aligned}
\mathbf{E}[T] &= \sum_{n=1}^{\infty} \mathbf{P}[T \geqslant n] = \sum_{n=1}^{\infty} \mathbf{P}[\phi \in \mathrm{ND}_k^{(n)}] = \sum_{n=1}^{\infty} \mathrm{nd}_k^{(n)} \\
&= C + \sum_{n=N+2}^{\infty} (1-\mathrm{nw}_k^{(n)} - \mathrm{n}\ell_k^{(n)}) = C + \sum_{n=N+2}^{\infty} \{(\mathrm{nw}_k - \mathrm{nw}_k^{(n)}) + (\mathrm{n}\ell_k - \mathrm{n}\ell_k^{(n)})\},
\end{aligned}$$

since $1-\mathrm{nw}_k - \mathrm{n}\ell_k = \mathrm{nd}_k = 0$ (by Theorem 2.1.3, as $\mathrm{n}\ell_k = c_k$). It converges as (2.9.1) and (2.9.2) guarantee exponential decay of the summands as $n \to \infty$.

To establish the second part of Theorem 2.1.10, we show that $|F_k'(c_k)| < 1 \implies H_k'(c_k) < 1$ when $k = 2,3$ (which then guarantees that the first part of Theorem 2.1.10 holds). Letting

$$a_0 = G'(1-c_2) = |F_1'(c_2)| \text{ and } a_1 = G'(F_1(c_2) - c_2),$$

we have (also making use of using Lemma 2.4.1)

$$F_2'(c_2) = -a_1(a_0+1) \text{ and } H_2'(c_2) = a_1^2(a_0^2 + 2a_0 - a_1 a_0 - a_1) < \{a_1(a_0+1)\}^2,$$

thus proving our claim for $k = 2$.

The proof is similar for $k=3$, albeit requiring more involved computations, as follows. Letting

$$b_0 = \left|F_1'(c_3)\right| = G'(1-c_3),\ b_1 = G'(F_1(c_3)-c_3) \text{ and } b_2 = G'(F_2(c_3)-c_3),$$

we have $F_3'(c_3) = -b_2(b_1b_0+b_1+1)$. This, along with Lemma 2.4.1, yields

$$\frac{d}{dx}G(g_2(1,F_2(x),F_3(x)))\big|_{x=c_3} = b_2[b_1^2b_0^2+2b_0b_1^2+b_0b_1+b_1^2-b_2b_1^2b_0-b_2b_1^2-b_2b_1],$$
$$\frac{d}{dx}G(g_2(F_1(x),F_2(x),F_3(x)))\big|_{x=c_3} = b_2[-b_1b_0+b_2b_1^2b_0+b_2b_1^2+2b_2b_1+b_2b_1b_0-b_2^2b_1b_0-b_2^2b_1-b_2^2].$$

These together give us

$$\begin{aligned}H_3'(c_3) = b_2^2[&b_1^2b_0^2+2b_0b_1^2+2b_0b_1+b_1^2+2b_1+1+\{-2b_2b_1^2b_0-2b_2b_1^2-3b_2b_1-b_2b_1b_0+b_2^2b_1b_0+b_2^2b_1\\ &+b_2^2-2b_1-1\}] < \{b_2[b_1b_0+b_1+1]\}^2\end{aligned}$$

The above inequality is obtained as follows. We use $F_2(c_3) < F_1(c_3)$ (since we have already shown that $F_i(x) \leqslant F_{i-1}(x)$ holds for $x \in [0, c_{i-1}]$, in the proof of Equation (2.2.14)) and the increasing nature of G' to deduce $b_2 < b_1$. This, in turn, yields $b_2^2b_1 < b_2b_1^2$, $b_2^2 < b_2b_1$ and $b_2^2b_1b_0 < b_2b_1^2b_0$.

Remark 2.9.1. *We have verified the inequality $H_k'(c_k) < |F_k'(c_k)|^2$ for a few higher values of k as well, and we conjecture that this is true for all $k \in \mathbb{N}$, but have been unable to discern a pattern in the expression for $H_k'(c_k)$ that is not too complicated to work with in order to prove this conjecture.*

2.10 Proofs of technical lemmas

2.10.1 Proofs of lemmas from §2.2

Proof of Lemma 2.2.1. As mentioned previously, the argument in this proof resembles that employed to prove Proposition 7 of [61]. We show that the sets $\widetilde{\mathrm{NL}} = \mathrm{NL} \setminus (\bigcup_{n=1}^{\infty} \mathrm{NL}^{(n)})$ and $\widetilde{\mathrm{NW}} = \mathrm{NW} \setminus (\bigcup_{n=1}^{\infty} \mathrm{NW}^{(n)})$ are empty. From (2.2.1), for u to be in $\widetilde{\mathrm{NW}}$, there exists some $v \in \Gamma_k(u) \cap \mathrm{NL}$. If v were in $\mathrm{NL}^{(n)}$ for some $n \in \mathbb{N}$, then the game starting at u is won by P1 (playing the first round) in less than $n+1$ rounds by moving the token from u to v in the first round. This implies $u \in \mathrm{NW}^{(n+1)}$, contradicting the assumption that $u \in \widetilde{\mathrm{NW}}$. Thus v must be in $\widetilde{\mathrm{NL}}$. Equivalently, we have $u \in \widetilde{\mathrm{NW}}$ iff $\Gamma_k(u) \cap \mathrm{NL} = \Gamma_k(u) \cap \widetilde{\mathrm{NL}} \neq \emptyset$. Moreover, since the players are assumed to play optimally, P1 moves the token from u to some $v \in \Gamma_k(u) \cap \widetilde{\mathrm{NL}}$ in the first round.

From (2.2.2), for v to be in $\widetilde{\mathrm{NL}}$, every vertex in $\Gamma_k(v)$ must be in NW. If for some $n \in \mathbb{N}$, $w \in \mathrm{NW}^{(n)}$ for every $w \in \Gamma_k(v)$, then no matter how P1 moves in the first round of the game starting at v, she loses in less than $n+1$ rounds, thus implying $v \in \mathrm{NL}^{(n+1)}$ and contradicting the assumption that $v \in \widetilde{\mathrm{NL}}$. Thus there exists some $w \in \Gamma_k(v)$ with $w \in \widetilde{\mathrm{NW}}$. Equivalently, $v \in \widetilde{\mathrm{NL}}$ iff $\Gamma_k(v) \subset \mathrm{NW}$ and $\Gamma_k(v) \cap \widetilde{\mathrm{NW}} \neq \emptyset$. Moreover, under optimal play, P1 moves the token from v to some $w \in \Gamma_k(v) \cap \widetilde{\mathrm{NW}}$ in the first round in her attempt to prolong the game as much as possible.

These observations reveal that if the game begins at $u_0 \in \widetilde{\mathrm{NL}}$, the token gets moved to some $u_1 \in \Gamma_k(u_0) \cap \widetilde{\mathrm{NW}}$ in the first round, to some $u_2 \in \Gamma_k(u_1) \cap \widetilde{\mathrm{NL}}$ in the second round, and so on. It continues till eternity, with each player *always* being able to make a move. The game thus ends in a draw, contradicting the definition of $\widetilde{\mathrm{NL}}$. Therefore, $\widetilde{\mathrm{NL}} = \emptyset$, and likewise, $\widetilde{\mathrm{NW}} = \emptyset$ as well. $\square$

Proof of Lemma 2.2.2. We show, for all $i \in \mathbb{N}$,

1. that F_i is a strictly decreasing function on $[0, c_{i-1}]$, where F_i is as defined in (2.1.6) and c_{i-1} is the unique fixed point of F_{i-1},

2. that $F_i(0) = 1$,

3. that c_i exists and is uniquely defined,

4. and that $\chi(0) < c_i < c_{i-1}$.

We prove these claims together, via induction on i. We note, at the very outset, that $G'(x) = \sum_{i=1}^{\infty} i\chi(i)x^i > 0$ for all $x > 0$, guaranteeing G is strictly increasing on $[0,1]$ – these facts are repeatedly utilized below.

Since $F_1'(x) = -G'(1-x) < 0$ for all $x \in [0,1)$ (as $\chi(0) < 1$), F_1 is strictly decreasing on $[0,1]$, proving (1) for $i = 1$. This also implies that $F_1(x) - x$ is strictly decreasing on $[0,1]$. Moreover, $F_1(0) = G(1) = 1 > 0$ (this verifies (2)), so that the curve $y = F_1(x) - x$ lies above the x-axis at $x = 0$, and $F_1(1) - 1 = \chi(0) - 1 < 0$, so that the curve $y = F_1(x) - x$ lies below the x-axis at $x = 1$. This fact, along with (1) for $i = 1$, guarantees that $F_1(x) - x$ has a unique root in $(0,1)$, which we denote by c_1. This proves (3) for $i = 1$. Finally, $F_1(\chi(0)) = G(1-\chi(0)) > G(0) = \chi(0)$ (as $\chi(0) < 1$). This, along with (1) and (3) for $i = 1$, allows us to conclude that $c_1 > \chi(0)$, thus proving (4) for $i = 1$. This completes verifying the base case for the induction.

Suppose we have shown that all of (1), (2), (3) and (4) hold for some $i \in \mathbb{N}$. From (2.1.6), we have $F_{i+1}'(x) = G'(F_i(x) - x)(F_i'(x) - 1)$. By the induction hypothesis, $F_i'(x) - 1 < F_i'(x) < 0$ for all $x \in [0, c_i] \subset [0, c_{i-1}]$, whereas $G'(F_i(x) - x) > 0$ for $x \in [0, c_i]$, thus proving that $F_{i+1}'(x) < 0$ and hence F_{i+1} is strictly decreasing on $[0, c_i]$. This proves (1).

By the induction hypothesis, $F_{i+1}(0) = G(F_i(0) - 0) = G(1) = 1$, proving (2). Note that this also implies that the curve $y = F_{i+1}(x) - x$ lies above the x-axis at $x = 0$, whereas $F_{i+1}(c_i) - c_i = G(F_i(c_i) - c_i) - c_i = G(0) - c_i = \chi(0) - c_i < 0$ (using the induction hypothesis pertaining to (4)), so that the curve $y = F_{i+1}(x) - x$ lies below the x-axis at $x = c_i$. This fact, along with (1) that we have already proved, guarantees that $F_{i+1}(x) - x$ has a unique root, c_{i+1}, in $(0, c_i)$ – this proves (3) and the second inequality of (4).

Due to (1) and as c_i is the fixed point of F_i in $[0, c_{i-1}]$, we know that the curve $y = F_i(x)$ lies strictly above the line $y = x$ for $x \in [0, c_i)$, and strictly below the line $y = x$ for $x \in (c_i, c_{i-1}]$. Since $\chi(0) \in [0, c_i)$ (because of the induction hypothesis pertaining to (4)), we conclude that $F_i(\chi(0)) > \chi(0)$. This, along with the strictly increasing nature of G on $[0,1]$, yields

$$F_{i+1}(\chi(0)) = G(F_i(\chi(0)) - \chi(0)) > G(0) = \chi(0). \tag{2.10.1}$$

Once again, by (1) and as c_{i+1} is the fixed point of F_{i+1} in $[0, c_i]$, we know that the curve $y = F_{i+1}(x)$ lies strictly above the line $y = x$ for $x \in [0, c_{i+1})$, and strictly below the line $y = x$ for $x \in (c_{i+1}, c_i]$. This, along with (2.10.1), implies that $\chi(0) \in [0, c_{i+1})$, thus completing the proof of (4). □

Proof of Lemma 2.2.3. Recall that we wish to prove $p_{i,j,n} = g_{i+1}(F_{j-i-1}(\mathrm{n}\ell^{(n)}), F_{j-i}(\mathrm{n}\ell^{(n)}), F_{k-i+1}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)}))$ for $0 \leqslant i < j \leqslant k$, where F_is and g_is are as defined in (2.1.6) and (2.1.7) respectively. We prove this via induction on i.

First, we note that for $i = 0$ and $j = 1$, the claim follows from (2.2.9). Suppose the claim holds for $p_{0,\ell,n}$ for $1 \leqslant \ell \leqslant j-1$, for some $2 \leqslant j \leqslant k$. From (2.2.10), the induction hypothesis and (2.1.6):

$$\begin{aligned}
p_{0,j,n} &= G\left(1 - \mathrm{n}\ell^{(n)} - \sum_{\ell=1}^{j-2} p_{0,\ell,n}\right) - G\left(1 - \mathrm{n}\ell^{(n)} - \sum_{\ell=1}^{j-1} p_{0,\ell,n}\right) \\
&= G\left(1 - \mathrm{n}\ell^{(n)} - \sum_{\ell=1}^{j-2} (F_{\ell-1}(\mathrm{n}\ell^{(n)}) - F_\ell(\mathrm{n}\ell^{(n)}))\right) - G\left(1 - \mathrm{n}\ell^{(n)} - \sum_{\ell=1}^{j-1} (F_{\ell-1}(\mathrm{n}\ell^{(n)}) - F_\ell(\mathrm{n}\ell^{(n)}))\right) \\
&= G\left(F_{j-2}(\mathrm{n}\ell^{(n)}) - \mathrm{n}\ell^{(n)}\right) - G\left(F_{j-1}(\mathrm{n}\ell^{(n)}) - \mathrm{n}\ell^{(n)}\right) \\
&= F_{j-1}(\mathrm{n}\ell^{(n)}) - F_j(\mathrm{n}\ell^{(n)}) = g_1\left(F_{j-1}(\mathrm{n}\ell^{(n)}) - F_j(\mathrm{n}\ell^{(n)})\right),
\end{aligned}$$

thus proving the claim for $i = 0$ and *all* $1 \leqslant j \leqslant k$.

Suppose the claim holds for all $p_{i-1,j,n}$ for all $i-1 < j \leqslant k$, for some $1 \leqslant i \leqslant k-1$. From

(2.2.11) and the induction hypothesis, for any j with $i < j \leqslant k$, we see that

$$\begin{aligned} p_{i,j,n} &= G\left(\sum_{\ell=j-1}^{k} p_{i-1,\ell,n}\right) - G\left(\sum_{\ell=j}^{k} p_{i-1,\ell,n}\right) \\ &= G\left(\sum_{\ell=j-1}^{k} g_i\left(F_{\ell-i}(\mathrm{n}\ell^{(n)}), F_{\ell-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right) \\ &\quad - G\left(\sum_{\ell=j}^{k} g_i\left(F_{\ell-i}(\mathrm{n}\ell^{(n)}), F_{\ell-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right). \end{aligned} \tag{2.10.2}$$

We analyze the two terms in (2.10.2) individually, to avoid cluttering. Using (2.1.7), we have

$$\begin{aligned} &G\left(\sum_{\ell=j-1}^{k} g_i\left(F_{\ell-i}(\mathrm{n}\ell^{(n)}), F_{\ell-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right) \\ &\quad = G\Bigg(\sum_{\ell=j-1}^{k} G\left(g_{i-1}\left(F_{\ell-i}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right) \\ &\qquad\quad - G\left(g_{i-1}\left(F_{\ell-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right)\Bigg) \\ &\quad = G\Bigg(G\left(g_{i-1}\left(F_{j-1-i}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right) \\ &\qquad\quad - G\left(g_{i-1}\left(F_{k-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right)\Bigg) \\ &\quad = G\left(g_i\left(F_{j-i-1}(\mathrm{n}\ell^{(n)}), F_{k-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right). \end{aligned} \tag{2.10.3}$$

The second term of (2.10.2) is analyzed as follows:

$$\begin{aligned} &G\left(\sum_{\ell=j}^{k} g_i\left(F_{\ell-i}(\mathrm{n}\ell^{(n)}), F_{\ell-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right) \\ &\quad = G\Bigg(\sum_{\ell=j}^{k} G\left(g_{i-1}\left(F_{\ell-i}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right) \\ &\qquad\quad - G\left(g_{i-1}\left(F_{\ell-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right)\Bigg) \end{aligned}$$

$$
\begin{aligned}
&= G\Bigg(G\left(g_{i-1}\left(F_{j-i}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right) \\
&\quad - G\left(g_{i-1}\left(F_{k-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right)\Bigg) \\
&= G\left(g_i\left(F_{j-i}(\mathrm{n}\ell^{(n)}), F_{k-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right). \qquad (2.10.4)
\end{aligned}
$$

Substituting from (2.10.3) and (2.10.4) in (2.10.2) and using (2.1.7), we obtain

$$
\begin{aligned}
p_{i,j,n} &= G\left(g_i\left(F_{j-i-1}(\mathrm{n}\ell^{(n)}), F_{k-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right) \\
&\quad - G\left(g_i\left(F_{j-i}(\mathrm{n}\ell^{(n)}), F_{k-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})\right)\right) \\
&= g_{i+1}(F_{j-i-1}(\mathrm{n}\ell^{(n)}), F_{j-i}(\mathrm{n}\ell^{(n)}), F_{k-i+1}(\mathrm{n}\ell^{(n)}), F_{k-i+2}(\mathrm{n}\ell^{(n)}), \ldots, F_k(\mathrm{n}\ell^{(n)})),
\end{aligned}
$$

thus completing the inductive step of the argument. This completes the proof of Lemma 2.2.3. $\square$

Proof of Equation (2.2.14). Recall that we intend to show that for any $0 \leqslant i_1 < i_2 \leqslant k-j$ and all $x \in [0, c_{k-1}]$,

$$
1 \geqslant g_j\big(F_{i_1}(x), F_{k-j+1}(x), F_{k-j+2}(x), \ldots, F_k(x)\big) \geqslant g_j\big(F_{i_2}(x), F_{k-j+1}(x), F_{k-j+2}(x), \ldots, F_k(x)\big) \geqslant 0.
$$

We prove this via induction on j. It is important to keep in mind that G is strictly increasing on $[0,1]$ and that $G(x) \in [0,1]$ for all $x \in [0,1]$, as these fact is used repeatedly in what follows.

First, we prove, via induction on i, that $F_i(x) \geqslant F_{i+1}(x)$ for all $x \in [0, c_i]$, for each $i \in \mathbb{N}_0$. the base case of $i = 0$ is immediate as $F_0(x) = 1$ and $F_1(x) = G(1-x) \in [0,1]$. Suppose $F_i(x) \geqslant F_{i+1}(x)$ for all $x \in [0, c_i]$, for some $i \in \mathbb{N}_0$. By (2.1.6) and the induction hypothesis, we have $F_{i+1}(x) = G(F_i(x) - x) \geqslant G(F_{i+1}(x) - x) = F_{i+2}(x)$ for all $x \in [0, c_{i+1}] \subset [0, c_i]$. When $0 \leqslant i_1 < i_2 < k$, we thus have

$$
g_1(F_{i_1}(x), F_k(x)) = F_{i_1}(x) - F_k(x) \geqslant F_{i_2}(x) - F_k(x) = g_1(F_{i_2}(x), F_k(x))
$$

for all $x \in [0, c_{k-1}]$, thus establishing (2.2.14) for $j = 1$.

We assume that (2.2.14) holds for some $j < k$, and now we prove it for $j+1$. For any $0 \leqslant i_1 < i_2 \leqslant k-(j+1)$ and $x \in [0, c_{k-1}]$, we obtain, using (2.1.7),

$$
g_{j+1}\big(F_{i_1}(x), F_{k-j}(x), F_{k-j+1}(x), \ldots, F_k(x)\big) - g_{j+1}\big(F_{i_2}(x), F_{k-j}(x), F_{k-j+1}(x), \ldots, F_k(x)\big)
$$

$$
\begin{aligned}
&= \left[G\left(g_j\left(F_{i_1}(x), F_{k-j+1}(x), \ldots, F_k(x)\right)\right) - G\left(g_j\left(F_{k-j}(x), F_{k-j+1}(x), \ldots, F_k(x)\right)\right)\right] \\
&\quad - \left[G\left(g_j\left(F_{i_2}(x), F_{k-j+1}(x), \ldots, F_k(x)\right)\right) - G\left(g_j\left(F_{k-j}(x), F_{k-j+1}(x), \ldots, F_k(x)\right)\right)\right] \\
&= G\left(g_j\left(F_{i_1}(x), F_{k-j+1}(x), \ldots, F_k(x)\right)\right) - G\left(g_j\left(F_{i_2}(x), F_{k-j+1}(x), \ldots, F_k(x)\right)\right)
\end{aligned}
$$

and this is non-negative by our induction hypothesis. For $0 \leqslant i \leqslant k-(j+1)$ and $x \in [0, c_{k-1}]$, using (2.1.7),

$$
\begin{aligned}
&g_{j+1}\left(F_i(x), F_{k-j}(x), F_{k-j+1}(x), \ldots, F_k(x)\right) \\
&\quad = G\left(g_j\left(F_i(x), F_{k-j+1}(x), \ldots, F_k(x)\right)\right) - G\left(g_j\left(F_{k-j}(x), F_{k-j+1}(x), \ldots, F_k(x)\right)\right)
\end{aligned}
$$

and this is non-negative because our induction hypothesis guarantees that $g_j\left(F_i(x), F_{k-j+1}(x), \ldots, F_k(x)\right) \geqslant g_j\left(F_{k-j}(x), F_{k-j+1}(x), \ldots, F_k(x)\right)$ (by setting $i_1 = i$ and $i_2 = k-j$), and G is increasing on $[0,1]$. It is also bounded above by 1 since $G\left(g_j\left(F_i(x), F_{k-j+1}(x), \ldots, F_k(x)\right)\right) \leqslant 1$, since $G(x) \leqslant 1$ for all $x \in [0,1]$. This completes the proof of Equation (2.2.14). □

2.10.2 Proofs of lemmas from §2.5

Proof of Lemma 2.5.1. Recall that, given a sequence of functions $\{r_i\}_{0 \leqslant i \leqslant k}$ defined and differentiable on an interval I of $\mathbb{R}$, with $(r_i(x), r_{k-j+1}(x), r_{k-j+2}(x), \ldots, r_k(x)) \in \mathscr{D}_j$ (defined as in (2.1.8)) for all $x \in I$ and all $0 \leqslant i < i+j \leqslant k$, Lemma 2.5.1 describes an expression for the derivative of $g_{k,\lambda}(r_0(x), r_1(x), \ldots, r_k(x))$ with respect to x, described via (2.5.14), (2.5.15) and (2.5.16), and this is what we establish now. Here, we focus on the offspring distribution Poisson(λ).

Recall that the pgf of Poisson(λ) is $G_\lambda(x) = e^{\lambda(x-1)}$, so that $G'_\lambda(x) = \lambda G_\lambda(x)$. We prove the lemma using induction on k. When $k=2$, we have

$$
\begin{aligned}
&\frac{d}{dx} g_{2,\lambda}(r_0(x), r_1(x), r_2(x)) = \frac{d}{dx} G_\lambda(r_0(x) - r_2(x)) - \frac{d}{dx} G_\lambda(r_1(x) - r_2(x)) \\
&= \lambda G_\lambda(r_0(x) - r_2(x))(r'_0(x) - r'_2(x)) - \lambda G_\lambda(r_1(x) - r_2(x))(r'_1(x) - r'_2(x)),
\end{aligned}
$$

which proves the base case for the induction.

Suppose (2.5.14), (2.5.15) and (2.5.16) hold for some $k \geqslant 2$. Using the induction hypothesis and (2.1.7), we have

$$
\frac{d}{dx} g_{k+1,\lambda}(r_0(x), r_1(x), \ldots, r_{k+1}(x))
$$

$$
\begin{aligned}
&= G'_\lambda\left(g_{k,\lambda}(r_0(x),r_2(x),\ldots,r_{k+1}(x))\right)\frac{d}{dx}g_{k,\lambda}\left(r_0(x),r_2(x),\ldots,r_{k+1}(x)\right)\\
&\quad - G'_\lambda\left(g_{k,\lambda}(r_1(x),r_2(x),\ldots,r_{k+1}(x))\right)\frac{d}{dx}g_{k,\lambda}\left(r_1(x),r_2(x),\ldots,r_{k+1}(x)\right)\\
&= \lambda^k G_\lambda\left(g_{k,\lambda}(r_0(x),r_2(x),\ldots,r_{k+1}(x))\right)f_{k,0,\lambda}\left(r_0(x),r_2(x),\ldots,r_{k+1}(x)\right)\left(r'_0(x)-r'_{k+1}(x)\right)\\
&\quad - \lambda^k G_\lambda\left(g_{k,\lambda}(r_1(x),r_2(x),\ldots,r_{k+1}(x))\right)f_{k,0,\lambda}\left(r_1(x),r_2(x),\ldots,r_{k+1}(x)\right)\left(r'_1(x)-r'_{k+1}(x)\right)\\
&\quad + \lambda^k\sum_{i=2}^{k}\Big[G_\lambda(g_{k,\lambda}(r_0(x),r_2(x),\ldots,r_{k+1}(x)))f_{k,i-1,\lambda}(r_0(x),r_2(x),\ldots,r_{k+1}(x))\\
&\quad - G_\lambda(g_{k,\lambda}(r_1(x),r_2(x),\ldots,r_{k+1}(x)))f_{k,i-1,\lambda}(r_1(x),r_2(x),\ldots,r_{k+1}(x))\Big]\left(r'_i(x)-r'_{k+1}(x)\right)\\
&= \lambda^k G_\lambda\left(g_{k,\lambda}\left(r_0(x),r_2(x),\ldots,r_{k+1}(x)\right)\right)\prod_{t=1}^{k-1}G_\lambda\left(g_{t,\lambda}\left(r_0(x),r_{k-t+2}(x),r_{k-t+3}(x),\ldots,r_{k+1}(x)\right)\right)\\
&\quad \left(r'_0(x)-r'_{k+1}(x)\right)\\
&\quad - \lambda^k G_\lambda\left(g_{k,\lambda}\left(r_1(x),r_2(x),\ldots,r_{k+1}(x)\right)\right)\prod_{t=1}^{k-1}G_\lambda\left(g_{t,\lambda}\left(r_1(x),r_{k-t+2}(x),r_{k-t+3}(x),\ldots,r_{k+1}(x)\right)\right)\\
&\quad \left(r'_1(x)-r'_{k+1}(x)\right)\\
&\quad + \lambda^k\sum_{i=2}^{k}\Bigg[G_\lambda\left(g_{k,\lambda}\left(r_0(x),r_2(x),\ldots,r_{k+1}(x)\right)\right)\prod_{t=1}^{k-(i-1)}G_\lambda\Big(g_{t,\lambda}\left(r_i(x),r_{k-t+2}(x),r_{k-t+3}(x),\right.\\
&\quad \left.\ldots,r_{k+1}(x)\right)\Big)\alpha_{k,i-1,\lambda}\left(r_0(x),r_2(x),\ldots,r_{k+1}(x)\right)\\
&\quad - G_\lambda\left(g_{k,\lambda}\left(r_1(x),r_2(x),\ldots,r_{k+1}(x)\right)\right)\prod_{t=1}^{k-(i-1)}G_\lambda\left(g_{t,\lambda}\left(r_i(x),r_{k-t+2}(x),r_{k-t+3}(x),\ldots,r_{k+1}(x)\right)\right)\\
&\quad \alpha_{k,i-1,\lambda}\left(r_1(x),r_2(x),\ldots,r_{k+1}(x)\right)\Bigg]\left(r'_i(x)-r'_{k+1}(x)\right)\\
&= \lambda^k\prod_{t=1}^{(k+1)-1}G_\lambda\left(g_{t,\lambda}\left(r_0(x),r_{(k+1)-t+1}(x),r_{(k+1)-t+2}(x),\ldots,r_{k+1}(x)\right)\right)\left(r'_0(x)-r'_{k+1}(x)\right)\\
&\quad - \lambda^k\prod_{t=1}^{(k+1)-1}G_\lambda\left(g_{t,\lambda}\left(r_1(x),r_{(k+1)-t+1}(x),r_{(k+1)-t+2}(x),\ldots,r_{k+1}(x)\right)\right)\left(r'_1(x)-r'_{k+1}(x)\right)\\
&\quad + \lambda^k\sum_{i=2}^{k}\prod_{t=1}^{(k+1)-i}G_\lambda\left(g_{t,\lambda}\left(r_i(x),r_{(k+1)-t+1}(x),r_{(k+1)-t+2}(x),\ldots,r_{k+1}(x)\right)\right)\\
&\quad \alpha_{k+1,i,\lambda}(r_0(x),\ldots,r_{k+1}(x))(r'_i(x)-r'_{k+1}(x)),
\end{aligned}
$$

where

$$\alpha_{k+1,i,\lambda}(r_0(x),\ldots,r_{k+1}(x)) = G_\lambda(g_{k,\lambda}(r_0(x),r_2(x),\ldots,r_{k+1}(x)))\alpha_{k,i-1,\lambda}(r_0(x),r_2(x),\ldots,r_{k+1}(x))$$

$$- G_\lambda(g_{k,\lambda}(r_1(x), r_2(x), \ldots, r_{k+1}(x)))\alpha_{k,i-1,\lambda}(r_1(x), r_2(x), \ldots, r_{k+1}(x))$$

is bounded by $a_{k+1,i} = 2a_{k,i-1}$ (using the induction hypothesis). $\square$

Proof of Lemma 2.5.2. Recall that our objective is to prove that $F'_{i,\lambda}(x) = -\lambda\left[\sum_{t=1}^{i-1}\lambda^{i-t}\prod_{j=t}^{i-1}F_{j,\lambda}(x)+1\right]F_{i,\lambda}(x)$ for all $x \in (0, c_{i-1})$ and $i \in \mathbb{N}$. We prove this by induction on i. When $i = 1$, we have $F'_{1,\lambda}(x) = -\lambda e^{-\lambda x} = -\lambda F_{1,\lambda}(x)$, proving the base case for the induction.

Suppose the lemma holds for some $i \geqslant 1$. Then, from (2.1.6) and the induction hypothesis, we have

$$\begin{aligned}
F'_{i+1,\lambda}(x) &= G'_\lambda(F_{i,\lambda}(x) - x)[F'_{i,\lambda}(x) - 1] \\
&= -\lambda G_\lambda(F_{i,\lambda}(x) - x)\left[\lambda\{\sum_{t=1}^{i-1}\lambda^{i-t}\prod_{j=t}^{i-1}F_{j,\lambda}(x)+1\}F_{i,\lambda}(x)+1\right] \\
&= -\lambda F_{i+1,\lambda}(x)\left[\sum_{t=1}^{i-1}\lambda^{i-t+1}F_{i,\lambda}(x)\prod_{j=t}^{i-1}F_{j,\lambda}(x)+\lambda F_{i,\lambda}(x)+1\right] \\
&= -\lambda\left[\sum_{t=1}^{i}\lambda^{(i+1)-t}\prod_{j=t}^{i}F_{j,\lambda}(x)+1\right]F_{i+1,\lambda}(x),
\end{aligned}$$

thus completing the inductive proof. $\square$

2.10.3 Proofs of lemmas from §2.6

Proof of Lemma 2.6.1. Recall that our objective here is to prove that $\eta_\lambda = \lambda c_{2,\lambda}$ is is strictly increasing for $\lambda \in (0, \lambda_0)$ and strictly decreasing for $\lambda \in (\lambda_0, \infty)$, where $\lambda_0 \approx 2.43634$. We accomplish this by showing that the derivative of η_λ with respect to λ is strictly positive, and we justify the differentiability of η_λ with respect to λ by recalling that $c_{2,\lambda}$ is differentiable with respect to λ, as proved in detail in §2.5.1.

To find an expression for the derivative of η_λ, we first find the derivative $c'_{2,\lambda}$ of $c_{2,\lambda}$ using the definition of $c_{2,\lambda}$ (i.e. $c_{2,\lambda} = F_{2,\lambda}(c_{2,\lambda})$). This yields

$$\begin{aligned}
c'_{2,\lambda} &= \frac{d}{d\lambda}F_{2,\lambda}(c_{2,\lambda}) = \frac{d}{d\lambda}\left[\exp\left\{\lambda e^{-\lambda c_{2,\lambda}} - \lambda c_{2,\lambda} - \lambda\right\}\right] \\
\implies c'_{2,\lambda} &= \left[e^{-\lambda c_{2,\lambda}} + \lambda\left\{-c_{2,\lambda} - \lambda c'_{2,\lambda}\right\}e^{-\lambda c_{2,\lambda}} - \left\{c_{2,\lambda} + \lambda c'_{2,\lambda}\right\} - 1\right]c_{2,\lambda}
\end{aligned}$$

$$\Longrightarrow c'_{2,\lambda} = \frac{\left[e^{-\lambda c_{2,\lambda}} - c_{2,\lambda} - 1 - \lambda c_{2,\lambda} e^{-\lambda c_{2,\lambda}}\right] c_{2,\lambda}}{1 + \lambda^2 c_{2,\lambda} e^{-\lambda c_{2,\lambda}} + \lambda c_{2,\lambda}}, \tag{2.10.5}$$

Next, we examine the behaviour of $c_{2,\lambda} e^{\lambda c_{2,\lambda}+1}$ as a function of λ. For all $\lambda > 0$, substituting from (2.10.5) and noting that $e^{-\lambda c_{2,\lambda}} < 1$, we have

$$\frac{d}{d\lambda}\left[c_{2,\lambda} e^{\lambda c_{2,\lambda}}\right] = \left[c'_{2,\lambda}\left(1 + \lambda c_{2,\lambda}\right) + c_{2,\lambda}^2\right] e^{\lambda c_{2,\lambda}} = \frac{(e^{-\lambda c_{2,\lambda}} - 1 - \lambda c_{2,\lambda}) c_{2,\lambda} e^{\lambda c_{2,\lambda}}}{1 + \lambda^2 c_{2,\lambda} e^{-\lambda c_{2,\lambda}} + \lambda c_{2,\lambda}} < 0.$$

Hence $c_{2,\lambda} e^{\lambda c_{2,\lambda}}$, and consequently, $c_{2,\lambda} e^{\lambda c_{2,\lambda}+1}$, is strictly decreasing in λ. Recall that (2.5.1) applied to $k = 2$ guarantees that $\lambda c_{2,\lambda} \to 0$ (and hence $c_{2,\lambda} \to 0$ as well) as $\lambda \to \infty$. Consequently,

$$\lim_{\lambda\to\infty} c_{2,\lambda} e^{\lambda c_{2,\lambda}+1} = 0.$$

On the other hand,

$$\lim_{\lambda\to 0} c_{2,\lambda} = \lim_{\lambda\to 0} F_{2,\lambda}(c_{2,\lambda}) = \lim_{\lambda\to 0} \exp\left\{\lambda\left(e^{-\lambda c_{2,\lambda}} - c_{2,\lambda} - 1\right)\right\} = 1 \Longrightarrow \lim_{\lambda\to 0} c_{2,\lambda} e^{\lambda c_{2,\lambda}+1} = e > 1.$$

From the three observations made above, we conclude that there exists a unique λ_0 with $c_{2,\lambda_0} e^{\lambda_0 c_{2,\lambda_0}+1} = 1$.

We now explore some useful characteristics of this λ_0. Using, again, $c_{2,\lambda} = F_{2,\lambda}(c_{2,\lambda})$, we have

$$\begin{aligned} c_{2,\lambda_0} &= \exp\left\{\lambda_0\left(e^{-\lambda_0 c_{2,\lambda_0}} - c_{2,\lambda_0} - 1\right)\right\} \\ \Longrightarrow c_{2,\lambda_0} e^{\lambda_0 c_{2,\lambda_0}+1} &= \exp\left\{\lambda_0 e^{-\lambda_0 c_{2,\lambda_0}} - \lambda_0 + 1\right\} = 1 \Longrightarrow c_{2,\lambda_0} = \frac{1}{\lambda_0}\ln\left(\frac{\lambda_0}{\lambda_0 - 1}\right) \end{aligned}$$

where the second step follows from how λ_0 has been defined above. Using the above derivation, we obtain

$$c_{2,\lambda_0} e^{\lambda_0 c_{2,\lambda_0}+1} = \frac{1}{\lambda_0}\ln\left(\frac{\lambda_0}{\lambda_0 - 1}\right) e \cdot \frac{\lambda_0}{\lambda_0 - 1} = 1 \Longrightarrow \lambda_0 \approx 2.43634.$$

We now come to the derivative of η_λ with respect to λ, and we make use of the observations made above:

$$\eta'_\lambda = c_{2,\lambda} + \lambda c'_{2,\lambda} = \frac{c_{2,\lambda}\left(1 + \lambda e^{-\lambda c_{2,\lambda}} - \lambda\right)}{1 + \lambda^2 c_{2,\lambda} e^{-\lambda c_{2,\lambda}} + \lambda c_{2,\lambda}} = \frac{c_{2,\lambda} \ln\left(c_{2,\lambda} e^{\lambda c_{2,\lambda}+1}\right)}{1 + \lambda^2 c_{2,\lambda} e^{-\lambda c_{2,\lambda}} + \lambda c_{2,\lambda}}. \tag{2.10.6}$$

Since $c_{2,\lambda}e^{\lambda c_{2,\lambda}+1}$ is strictly decreasing and takes the value 1 at $\lambda = \lambda_0$, we have $\ln\left(c_{2,\lambda}e^{\lambda c_{2,\lambda}+1}\right) > 0$ for $\lambda < \lambda_0$ and $\ln\left(c_{2,\lambda}e^{\lambda c_{2,\lambda}+1}\right) < 0$ for $\lambda > \lambda_0$. Hence $\eta'_\lambda > 0$ for $0 < \lambda < \lambda_0$ and $\eta'_\lambda < 0$ for $\lambda > \lambda_0$. Thus η_λ is strictly increasing for $\lambda \in (0, \lambda_0)$, strictly decreasing for $\lambda \in (\lambda_0, \infty)$, and its maximum value, attained at λ_0, is ≈ 0.52839925. □

Before we begin the proofs of Lemmas 2.6.2 through 2.6.6, we urge the reader to recall the definitions of $\alpha(x)$, $\beta(x)$, and A_i for $1 \leqslant i \leqslant 8$, from §2.6.

Proof of Lemma 2.6.2. Recall that our objective here is to prove, after taking out the common factor of $\{\alpha(x) - \beta(x)\}F_{2,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)^2$, that

$$\lambda^2 F_{2,\lambda}(x)\{\alpha(x) - \beta(x)\} + \lambda F_{2,\lambda}(x) > 1 \tag{2.10.7}$$

for all $x \in (\gamma_\lambda, c_{2,\lambda}]$ and all $\lambda \geqslant 2.5$. Differentiating the left side of this inequality with respect to x gives

$$\begin{aligned}
&\lambda^2 F'_{2,\lambda}(x)\{\alpha(x) - \beta(x)\} + \lambda^2 F_{2,\lambda}(x)\{\alpha'(x) - \beta'(x)\} + \lambda F'_{2,\lambda}(x) \\
&\quad = -\lambda^3\left(\lambda F_{1,\lambda}(x)+1\right)F_{2,\lambda}(x)\{\alpha(x) - \beta(x)\} + \lambda^2 F_{2,\lambda}(x)[\lambda^2\alpha(x)(\lambda F_{1,\lambda}(x)+1)F_{2,\lambda}(x) \\
&\qquad - \lambda^2\beta(x)\{-F_{1,\lambda}(x) + (\lambda F_{1,\lambda}(x)+1)F_{2,\lambda}(x)\}] - \lambda^2(\lambda F_{1,\lambda}(x)+1)F_{2,\lambda}(x) \\
&\quad = \lambda^3(\lambda F_{1,\lambda}(x)+1)F_{2,\lambda}(x)\{\alpha(x) - \beta(x)\}[\lambda F_{2,\lambda}(x) - 1] + \lambda^3 F_{1,\lambda}(x)F_{2,\lambda}(x)\{\lambda\beta(x) - 1\} - \lambda^2 F_{2,\lambda}(x).
\end{aligned} \tag{2.10.8}$$

In what follows, we show that this derivative is strictly negative for $x \in (\gamma_\lambda, c_{2,\lambda}]$, for $\lambda > 0$.

Since $F_{2,\lambda}$ is strictly decreasing on $[0, c_{1,\lambda}]$ by Lemma 2.2.2, and as $c_{2,\lambda}$ is the unique fixed point of $F_{2,\lambda}$, we have $F_{2,\lambda}(x) \geqslant x$ for all $0 \leqslant x \leqslant c_{2,\lambda}$. Thus

$$\beta(x) = G_\lambda(F_{1,\lambda}(x) - F_{2,\lambda}(x)) \leqslant G_\lambda(F_{1,\lambda}(x) - x) = F_{2,\lambda}(x). \tag{2.10.9}$$

Due to the strictly decreasing nature of $F_{2,\lambda}$, we also have

$$F_{2,\lambda}(x) < F_{2,\lambda}(\gamma_\lambda) = \frac{1}{\lambda} \text{ for } x > \gamma_\lambda, \tag{2.10.10}$$

so that from (2.10.9) and (2.10.10), we obtain

$$\lambda F_{2,\lambda}(x) - 1 < 0 \text{ and } \lambda\beta(x) - 1 < 0 \text{ for all } x \in (\gamma_\lambda, c_{2,\lambda}]. \tag{2.10.11}$$

Thus the derivative in (2.10.8) strictly negative for $x \in (\gamma_\lambda, c_{2,\lambda}]$, implying that the function on the left side of (2.10.7) is strictly decreasing for $x \in (\gamma_\lambda, c_{2,\lambda}]$.

Utilizing the above finding, for all $\lambda > 0$ and $\gamma_\lambda < x \leqslant c_{2,\lambda}$, we have

$$\begin{aligned}\lambda^2 F_{2,\lambda}(x)\{\alpha(x)-\beta(x)\}+\lambda F_{2,\lambda}(x) &\geqslant \lambda^2 F_{2,\lambda}(c_{2,\lambda})\{\alpha(c_{2,\lambda})-\beta(c_{2,\lambda})\}+\lambda F_{2,\lambda}(c_{2,\lambda})\\ &= \lambda^2 c_{2,\lambda}\{e^{-\lambda c_{2,\lambda}}-c_{2,\lambda}\}+\lambda c_{2,\lambda} = \lambda\eta_\lambda e^{-\eta_\lambda}-\eta_\lambda^2+\eta_\lambda.\end{aligned} \tag{2.10.12}$$

Differentiating the right side of (2.10.12) with respect to λ and substituting from (2.10.6), we have

$$\begin{aligned}&\eta_\lambda e^{-\eta_\lambda}+\lambda\eta_\lambda' e^{-\eta_\lambda}-\lambda\eta_\lambda\eta_\lambda' e^{-\eta_\lambda}-2\eta_\lambda\eta_\lambda'+\eta_\lambda'\\ &=\eta_\lambda e^{-\eta_\lambda}+\frac{\eta_\lambda\,(1+\lambda e^{-\eta_\lambda}-\lambda)}{\lambda\,(1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda)}\left[\lambda e^{-\eta_\lambda}-\lambda\eta_\lambda e^{-\eta_\lambda}-2\eta_\lambda+1\right]\\ &=\frac{\eta_\lambda[\lambda^2 e^{-\eta_\lambda}\{e^{-\eta_\lambda}-1+\eta_\lambda\}+2\lambda\eta_\lambda\{1-e^{-\eta_\lambda}\}+\{3\lambda e^{-\eta_\lambda}-2\eta_\lambda+1-\lambda\}]}{\lambda\,(1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda)}\\ &\geqslant\frac{\eta_\lambda\{3\lambda e^{-0.5284}-2\cdot 0.5284+1-\lambda\}}{\lambda\,(1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda)}=\frac{\eta_\lambda\{0.7686\lambda-0.0568\}}{\lambda\,(1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda)}>0 \text{ for all } \lambda\geqslant 2,\end{aligned}$$

where we use $e^{-x}-1 \geqslant -x$ for all $x \geqslant 0$, and that $\eta_\lambda < 0.5284$ for all $\lambda > 0$ from Lemma 2.6.1. Thus $\lambda\eta_\lambda e^{-\eta_\lambda}-\eta_\lambda^2+\eta_\lambda$ is strictly increasing for $\lambda \geqslant 2$. Since its value at $\lambda = 2.5$ is $\approx 1.0279 > 1$, hence we conclude that

$$\lambda\eta_\lambda e^{-\eta_\lambda}-\eta_\lambda^2+\eta_\lambda > 1 \text{ for all } \lambda \geqslant 2.5. \tag{2.10.13}$$

Combining (2.10.12) and (2.10.13), we conclude that (2.10.7) does hold for all $x \in (\gamma_\lambda, c_{2,\lambda}]$ and $\lambda \geqslant 2.5$, as desired. □

Proof of Lemma 2.6.3. Recall that our objective here is to show, after taking out the common factor of $\lambda\{\alpha(x)-\beta(x)\}F_{1,\lambda}(x)F_{2,\lambda}(x)$, that

$$\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1) > 1 \tag{2.10.14}$$

for all $x \in (\gamma_\lambda, c_{2,\lambda}]$ and $\lambda \geqslant 2$. Differentiating the left side of (2.10.14) with respect to x yields

$$\lambda\beta'(x)(\lambda F_{1,\lambda}(x)+1)+\lambda^2\beta(x)F_{1,\lambda}'(x) = \lambda^3\beta(x)[F_{2,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)^2-F_{1,\lambda}(x)(\lambda F_{1,\lambda}(x)+2)]. \tag{2.10.15}$$

For $x < c_{2,\lambda}$, we have

$$\frac{d}{dx}[F_{2,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)^2-F_{1,\lambda}(x)(\lambda F_{1,\lambda}(x)+2)]$$

$$
\begin{aligned}
&= \lambda(\lambda F_{1,\lambda}(x)+1)[F_{1,\lambda}(x)\{2-\lambda^2 F_{1,\lambda}(x)F_{2,\lambda}(x)-4\lambda F_{2,\lambda}(x)\}-F_{2,\lambda}(x)] \\
&\leqslant \lambda(\lambda F_{1,\lambda}(x)+1)[F_{1,\lambda}(x)\{2-\lambda^2 F_{1,\lambda}(c_{2,\lambda})F_{2,\lambda}(c_{2,\lambda})-4\lambda F_{2,\lambda}(c_{2,\lambda})\}-F_{2,\lambda}(x)] \text{ (by Lemma 2.2.2)} \\
&= \lambda(\lambda F_{1,\lambda}(x)+1)[F_{1,\lambda}(x)\{2-\lambda\eta_\lambda e^{-\eta_\lambda}-4\eta_\lambda\}-F_{2,\lambda}(x)]. \qquad (2.10.16)
\end{aligned}
$$

We need to examine the behaviour of $2-\lambda\eta_\lambda e^{-\eta_\lambda}-4\eta_\lambda$ as a function of λ. Substituting from (2.10.6),

$$
\frac{d}{d\lambda}[\lambda\eta_\lambda e^{-\eta_\lambda}] = \eta_\lambda e^{-\eta_\lambda}[1+(1-\eta_\lambda)\frac{1+\lambda e^{-\eta_\lambda}-\lambda}{1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda}] = \eta_\lambda e^{-\eta_\lambda}\cdot\frac{2+\lambda e^{-\eta_\lambda}-\lambda+\lambda\eta_\lambda}{1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda} > 0
$$

since $e^{-\eta_\lambda}-1+\eta_\lambda \geqslant 0$. Thus $\lambda\eta_\lambda e^{-\eta_\lambda}$ is strictly increasing in λ. Since its value is ≈ 2.25080 at $\lambda = 8.644$, we conclude that $\lambda\eta_\lambda e^{-\eta_\lambda} > 2$ for all $\lambda \geqslant 8.644$. This shows that $2-\lambda\eta_\lambda e^{-\eta_\lambda}-4\eta_\lambda < 0$ for all $\lambda \geqslant 8.644$.

On the other hand,

$$
\begin{aligned}
&\frac{d}{d\lambda}[\lambda\eta_\lambda e^{-\eta_\lambda}+4\eta_\lambda] = \eta_\lambda e^{-\eta_\lambda}\cdot\frac{2+\lambda e^{-\eta_\lambda}-\lambda+\lambda\eta_\lambda}{1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda}+\frac{4\eta_\lambda(1+\lambda e^{-\eta_\lambda}-\lambda)}{\lambda(1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda)} \\
&= \frac{\eta_\lambda}{\lambda(1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda)}[\{6\lambda e^{-\eta_\lambda}+4-4\lambda\}+\lambda^2 e^{-\eta_\lambda}\{e^{-\eta_\lambda}-1+\eta_\lambda\}] \\
&\geqslant \frac{\eta_\lambda[\{6\lambda e^{-0.5284}+4-4\lambda\}+\lambda^2 e^{-\eta_\lambda}\{e^{-\eta_\lambda}-1+\eta_\lambda\}]}{\lambda(1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda)} \text{ (by Lemma 2.6.1)} \\
&= \frac{\eta_\lambda[\{4-0.462715\lambda\}+\lambda^2 e^{-\eta_\lambda}\{e^{-\eta_\lambda}-1+\eta_\lambda\}]}{\lambda(1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda)}
\end{aligned}
$$

which is non-negative for all $\lambda \leqslant 4/0.462715 \approx 8.644$. This shows that $\lambda\eta_\lambda e^{-\eta_\lambda}+4\eta_\lambda$ is increasing for $\lambda \leqslant 8.644$, and at $\lambda = 2$ we have $4\eta_\lambda = 2.0957 > 2$. Therefore $\lambda\eta_\lambda e^{-\eta_\lambda}+4\eta_\lambda > 2$, and hence $2-\lambda\eta_\lambda e^{-\eta_\lambda}-4\eta_\lambda < 0$ for all $\lambda \leqslant 8.644$.

The above findings show that $2-\lambda\eta_\lambda e^{-\eta_\lambda}-4\eta_\lambda < 0$ for all $\lambda \geqslant 2$, so that the expression on the right side of (2.10.16) is strictly negative for all $\lambda \geqslant 2$. This implies that, via (2.10.16), that $F_{2,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)^2-F_{1,\lambda}(x)(\lambda F_{1,\lambda}(x)+2)$ is strictly decreasing for $x \leqslant c_{2,\lambda}$ and $\lambda \geqslant 2$. In the next paragraph, we focus on the value of this function at $c_{2,\lambda}$, as a function of λ.

When $x = c_{2,\lambda}$ and $\lambda \geqslant 2$, we have

$$
\begin{aligned}
&F_{2,\lambda}(c_{2,\lambda})(\lambda F_{1,\lambda}(c_{2,\lambda})+1)^2-F_{1,\lambda}(c_{2,\lambda})(\lambda F_{1,\lambda}(c_{2,\lambda})+2) \\
&< F_{2,\lambda}(c_{2,\lambda})(\lambda F_{1,\lambda}(c_{2,\lambda})+1)^2-F_{1,\lambda}(c_{2,\lambda})(\lambda F_{1,\lambda}(c_{2,\lambda})+1) \\
&= (\lambda F_{1,\lambda}(c_{2,\lambda})+1)\{c_{2,\lambda}(\lambda F_{1,\lambda}(c_{2,\lambda})+1)-F_{1,\lambda}(c_{2,\lambda})\}, \text{ since } F_{2,\lambda}(c_{2,\lambda}) = c_{2,\lambda}; \\
&\leqslant (\lambda F_{1,\lambda}(c_{2,\lambda})+1)\{0.5284F_{1,\lambda}(c_{2,\lambda})+c_{2,\lambda}-F_{1,\lambda}(c_{2,\lambda})\}, \text{ since } \eta_\lambda \leqslant 0.5284 \text{ by Lemma 2.6.1};
\end{aligned}
$$

$$= (\lambda F_{1,\lambda}(c_{2,\lambda})+1)\{-0.4716F_{1,\lambda}(c_{2,\lambda})+c_{2,\lambda}\}$$
$$= (\lambda F_{1,\lambda}(c_{2,\lambda})+1)\{-0.4716e^{-\lambda c_{2,\lambda}}+c_{2,\lambda}\}$$
$$\leqslant (\lambda F_{1,\lambda}(c_{2,\lambda})+1)\{-0.4716e^{-0.5284}+0.2619\} < 0, \tag{2.10.17}$$

where in the last step we use

1. the fact (proved in §2.5.1) that $c_{2,\lambda}$ is strictly decreasing in λ, and hence $c_{2,\lambda} \leqslant c_{2,2} \approx 0.2619$ for $\lambda \geqslant 2$,
2. and, once again, that $\eta_\lambda \leqslant 0.5284$, by Lemma 2.6.1.

Next, we focus on the behaviour of $F_{2,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)^2 - F_{1,\lambda}(x)(\lambda F_{1,\lambda}(x)+2)$ at $x = \gamma_\lambda$, as a function of λ. Since $F_{2,\lambda}(\gamma_\lambda) = \frac{1}{\lambda}$ (by definition of γ_λ in (2.6.4)), we have

$$-F_{1,\lambda}(\gamma_\lambda)(\lambda F_{1,\lambda}(\gamma_\lambda)+2) + F_{2,\lambda}(\gamma_\lambda)(\lambda F_{1,\lambda}(\gamma_\lambda)+1)^2 = \frac{1}{\lambda} > 0. \tag{2.10.18}$$

For each $\lambda \geqslant 2$, since we have proved above that $F_{2,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)^2 - F_{1,\lambda}(x)(\lambda F_{1,\lambda}(x)+2)$ is strictly decreasing when $x \in [0, c_{2,\lambda}]$, from (2.10.17) and (2.10.18) we conclude that $F_{2,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)^2 - F_{1,\lambda}(x)(\lambda F_{1,\lambda}(x)+2)$ is initially strictly positive, and then strictly negative, on the interval $(\gamma_\lambda, c_{2,\lambda}]$. This, along with (2.10.15), implies that $\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1)$ is initially strictly increasing, and then strictly decreasing, on $(\gamma_\lambda, c_{2,\lambda}]$. Therefore, for each $\lambda \geqslant 2$,

$$\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1) \geqslant \min\left\{\lambda\beta(c_{2,\lambda})(\lambda F_{1,\lambda}(c_{2,\lambda})+1), \lambda\beta(\gamma_\lambda)(\lambda F_{1,\lambda}(\gamma_\lambda)+1)\right\} \text{ for all } x \in (\gamma_\lambda, c_{2,\lambda}]. \tag{2.10.19}$$

The idea, now, is to consider the values of $\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1)$ at both $x = c_{2,\lambda}$ and $x = \gamma_\lambda$ to see which of them is the minimum, and then use this global minima of $\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1)$ on $(\gamma_\lambda, c_{2,\lambda}]$ to deduce that (2.10.14) holds.

At the very outset of §2.6.2, we have mentioned that $\beta(c_{2,\lambda}) = c_{2,\lambda}$. Substituting from (2.10.6), the derivative of the value of $\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1)$ at $x = c_{2,\lambda}$, with respect to λ, becomes

$$\begin{aligned}\frac{d}{d\lambda}[\lambda\beta(c_{2,\lambda})(\lambda F_{1,\lambda}(c_{2,\lambda})+1)] &= \frac{d}{d\lambda}\left[\lambda c_{2,\lambda}\left(\lambda e^{-\lambda c_{2,\lambda}}+1\right)\right] = \frac{d}{d\lambda}\left[\eta_\lambda\left(\lambda e^{-\eta_\lambda}+1\right)\right] \\ &= \eta_\lambda e^{-\eta_\lambda} + \frac{\eta_\lambda(1+\lambda e^{-\eta_\lambda}-\lambda)}{\lambda(1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda)}[\lambda e^{-\eta_\lambda} - \lambda\eta_\lambda e^{-\eta_\lambda}+1] \\ &= \frac{\eta_\lambda}{\lambda(1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda)}[\{3\lambda e^{-\eta_\lambda}+1-\lambda\} + \lambda^2 e^{-\eta_\lambda}\{e^{-\eta_\lambda}-1+\eta_\lambda\}],\end{aligned}$$

and this is strictly positive since $3e^{-\eta_\lambda} \geqslant 3e^{-0.5284} \approx 1.7686 > 1$ for all $\lambda > 0$, by Lemma 2.6.1. Thus $\lambda\beta(c_{2,\lambda})(\lambda F_{1,\lambda}(c_{2,\lambda})+1)$ is strictly increasing in λ, and its value at $\lambda = 2$ is $\approx 1.14446 > 1$. Thus

$$\lambda\beta(c_{2,\lambda})(\lambda F_{1,\lambda}(c_{2,\lambda})+1) \geqslant 1.14446 > 1 \text{ for all } \lambda \geqslant 2. \tag{2.10.20}$$

We now need to examine the behaviour of the value of $\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1)$ at $x = \gamma_\lambda$, as a function of λ. Via a similar application of the implicit function theorem as that used to justify the differentiability of $c_{k,\lambda}$ with respect to λ in §2.5.1, we conclude that γ_λ is differentiable with respect to λ as well. Writing $\theta_\lambda = \lambda\gamma_\lambda$, we then have

$$F_{2,\lambda}(\gamma_\lambda) = \exp\{\lambda e^{-\theta_\lambda} - \theta_\lambda - \lambda\} = \frac{1}{\lambda} \implies \theta'_\lambda = \frac{e^{-\theta_\lambda} - 1 + \frac{1}{\lambda}}{\lambda e^{-\theta_\lambda} + 1}.$$

Differentiating the value of $\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1)$ at $x = \gamma_\lambda$ with respect to λ and substituting from above, we have

$$\begin{aligned}
\frac{d}{d\lambda}[\lambda^2\beta(\gamma_\lambda)F_{1,\lambda}(\gamma_\lambda) + \lambda\beta(\gamma_\lambda)] &= \frac{d}{d\lambda}[\lambda^2 e^{\lambda e^{-\theta_\lambda}-1-\lambda-\theta_\lambda} + \lambda e^{\lambda e^{-\theta_\lambda}-1-\lambda}] \\
&= 2\lambda\beta(\gamma_\lambda)F_{1,\lambda}(\gamma_\lambda) + \lambda^2\{e^{-\theta_\lambda} - \lambda\theta'_\lambda e^{-\theta_\lambda} - 1 - \theta'_\lambda\}\beta(\gamma_\lambda)F_{1,\lambda}(\gamma_\lambda) \\
&\quad + \beta(\gamma_\lambda) + \lambda\{e^{-\theta_\lambda} - \lambda\theta'_\lambda e^{-\theta_\lambda} - 1\}\beta(\gamma_\lambda) \\
&= \lambda\beta(\gamma_\lambda)F_{1,\lambda}(\gamma_\lambda) + \beta(\gamma_\lambda) - \frac{\lambda\beta(\gamma_\lambda)}{\lambda e^{-\theta_\lambda}+1} \\
&= \lambda\beta(\gamma_\lambda)\cdot\frac{\lambda e^{-2\theta_\lambda} + e^{-\theta_\lambda} - 1}{\lambda e^{-\theta_\lambda}+1} + \beta(\gamma_\lambda) \\
&> \lambda\beta(\gamma_\lambda)\cdot\frac{\lambda e^{-2\eta_\lambda} + e^{-\eta_\lambda} - 1}{\lambda e^{-\theta_\lambda}+1} + \beta(\gamma_\lambda) \\
&\geqslant \lambda\beta(\gamma_\lambda)\cdot\frac{2e^{-2\cdot 0.5284} + e^{-0.5284} - 1}{\lambda e^{-\theta_\lambda}+1} + \beta(\gamma_\lambda) > 0,
\end{aligned}$$

for $\lambda \geqslant 2$, where we use $\gamma_\lambda < c_{2,\lambda} \implies \theta_\lambda < \eta_\lambda < 0.5284$ (from Lemma 2.6.1). This tells us that $\lambda\beta(\gamma_\lambda)(\lambda F_{1,\lambda}(\gamma_\lambda)+1)$ is strictly increasing for $\lambda \geqslant 2$, and its value at $\lambda = 2$ is $\approx 1.20824 > 1$. Thus

$$\lambda\beta(\gamma_\lambda)(\lambda F_{1,\lambda}(\gamma_\lambda)+1) \geqslant 1.20824 > 1 \text{ for all } \lambda \geqslant 2. \tag{2.10.21}$$

From (2.10.19), (2.10.20) and (2.10.21), we conclude that (2.10.14) holds for all $x \in (\gamma_\lambda, c_{2,\lambda}]$, for each $\lambda \geqslant 2$, completing our proof. □

Proof of Lemma 2.6.4. Recall that, after setting aside the common factor of

$\beta(x)F_{1,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)$, our objective here is to prove that

$$\lambda^2 F_{2,\lambda}(x)\{\alpha(x)-\beta(x)\}+2\lambda F_{2,\lambda}(x)>1 \tag{2.10.22}$$

for $x \in (\gamma_\lambda, c_{2,\lambda}]$ and $\lambda \geqslant 2$. Differentiating the left side with respect to x, we have

$$\lambda^3 F_{2,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)\{\alpha(x)-\beta(x)\}[\lambda F_{2,\lambda}(x)-1]+\lambda^2 F_{2,\lambda}(x)[\lambda F_{1,\lambda}(x)\{\lambda\beta(x)-2\}-2],$$

and using (2.10.11), we conclude that the above is strictly negative whenever $x > \gamma_\lambda$. Thus $\lambda^2 F_{2,\lambda}(x)\{\alpha(x)-\beta(x)\}+2\lambda F_{2,\lambda}(x)$ is strictly decreasing for $x \in [\gamma_\lambda, c_{2,\lambda})$, and its minimum is attained at $c_{2,\lambda}$. In other words,

$$\lambda^2 F_{2,\lambda}(x)\{\alpha(x)-\beta(x)\}+2\lambda F_{2,\lambda}(x) \geqslant \lambda^2 c_{2,\lambda}\{e^{-\lambda c_{2,\lambda}}-c_{2,\lambda}\}+2\lambda c_{2,\lambda} \text{ for all } x \in [\gamma_\lambda, c_{2,\lambda}). \tag{2.10.23}$$

Differentiating this minima with respect to λ and substituting from (2.10.6) yield

$$\begin{aligned}&\frac{d}{d\lambda}[\lambda^2 c_{2,\lambda}\{e^{-\lambda c_{2,\lambda}}-c_{2,\lambda}\}+2\lambda c_{2,\lambda}] = \frac{d}{d\lambda}[\lambda\eta_\lambda e^{-\eta_\lambda}-\eta_\lambda^2+2\eta_\lambda]\\ &= \frac{\eta_\lambda[2\lambda\{2e^{-\eta_\lambda}-1\}+2\{1-\eta_\lambda\}+2\lambda\eta_\lambda\{1-e^{-\eta_\lambda}\}+\lambda^2 e^{-\eta_\lambda}\{e^{-\eta_\lambda}-1+\eta_\lambda\}]}{\lambda(1+\lambda\eta_\lambda e^{-\eta_\lambda}+\eta_\lambda)},\end{aligned}$$

and to show that this is strictly positive for $\lambda \geqslant 2$, it suffices to show that the first two summands in the numerator are both strictly positive. This follows since $\eta_\lambda < 0.5284$, from Lemma 2.6.1. Thus $\lambda^2 c_{2,\lambda}\{e^{-\lambda c_{2,\lambda}}-c_{2,\lambda}\}+2\lambda c_{2,\lambda}$ is strictly increasing in λ. Its value at $\lambda = 2$ is $\approx 1.3939 > 1$, so that we conclude that

$$\lambda^2 c_{2,\lambda}\{e^{-\lambda c_{2,\lambda}}-c_{2,\lambda}\}+2\lambda c_{2,\lambda} > 1 \text{ for all } \lambda \geqslant 2. \tag{2.10.24}$$

From (2.10.23) and (2.10.24), we conclude that (2.10.22) holds, and the proof is complete. □

Proof of Lemma 2.6.5. It is important, in order to understand the inequalities we need to establish in this proof, to revisit the proofs of Lemmas 2.6.2, 2.6.3 and 2.6.4, for all $\lambda \geqslant 2$ and $x \in (\gamma_\lambda, c_{2,\lambda}]$. Note that the entire argument outlined in the proof of Lemma 2.6.2 extends to $\lambda \geqslant 2$, with the only difference being that we now consider the value of $\lambda\eta_\lambda e^{-\eta_\lambda}-\eta_\lambda^2+\eta_\lambda$ at $\lambda = 2$ instead of at $\lambda = 2.5$, and this value is ≈ 0.869957. Consequently, from (2.10.12) and the fact (shown in the proof of Lemma 2.6.2) that $\lambda\eta_\lambda e^{-\eta_\lambda}-\eta_\lambda^2+\eta_\lambda$ is strictly increasing for $\lambda \geqslant 2$, we conclude that

$$\lambda^2 F_{2,\lambda}(x)\{\alpha(x)-\beta(x)\}+\lambda F_{2,\lambda}(x) \geqslant 0.869957 \text{ for all } x \in (\gamma_\lambda, c_{2,\lambda}], \text{ for all } \lambda \geqslant 2. \tag{2.10.25}$$

This leads to

$$
\begin{aligned}
&\lambda^2\{\alpha(x)-\beta(x)\}^2(\lambda F_{1,\lambda}(x)+1)^2(F_{2,\lambda}(x))^2+\lambda\{\alpha(x)-\beta(x)\}(\lambda F_{1,\lambda}(x)+1)^2(F_{2,\lambda}(x))^2\\
&\quad-\{\alpha(x)-\beta(x)\}(\lambda F_{1,\lambda}(x)+1)^2F_{2,\lambda}(x)\\
&=\{\alpha(x)-\beta(x)\}^2(\lambda F_{1,\lambda}(x)+1)^2F_{2,\lambda}(x)[\lambda^2F_{2,\lambda}(x)\{\alpha(x)-\beta(x)\}+\lambda F_{2,\lambda}(x)-1]\\
&\geqslant -0.130043\{\alpha(x)-\beta(x)\}(\lambda F_{1,\lambda}(x)+1)^2F_{2,\lambda}(x) \text{ for all } x\in(\gamma_\lambda,c_{2,\lambda}], \text{ for } \lambda\geqslant 2. \quad (2.10.26)
\end{aligned}
$$

From (2.10.19), (2.10.20) and (2.10.21) in the proof of Lemma 2.6.3, we have $\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1)\geqslant 1.14446$ for all $\lambda\geqslant 2$, so that

$$
\begin{aligned}
&\lambda^2\beta(x)\{\alpha(x)-\beta(x)\}(\lambda F_{1,\lambda}(x)+1)F_{1,\lambda}(x)F_{2,\lambda}(x)-\lambda\{\alpha(x)-\beta(x)\}F_{1,\lambda}(x)F_{2,\lambda}(x)\\
&=\lambda\{\alpha(x)-\beta(x)\}F_{1,\lambda}(x)F_{2,\lambda}(x)[\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1)-1]\\
&\geqslant 0.14446\lambda\{\alpha(x)-\beta(x)\}F_{1,\lambda}(x)F_{2,\lambda}(x) \text{ for all } x\in(\gamma_\lambda,c_{2,\lambda}], \text{ for } \lambda\geqslant 2. \quad (2.10.27)
\end{aligned}
$$

Finally, from the proof of Lemma 2.6.4, we obtain $\lambda^2F_{2,\lambda}(x)\{\alpha(x)-\beta(x)\}+2\lambda F_{2,\lambda}(x)\geqslant 1.3939$, so that

$$
\begin{aligned}
&\lambda^2\beta(x)\{\alpha(x)-\beta(x)\}\left(\lambda F_{1,\lambda}(x)+1\right)F_{1,\lambda}(x)F_{2,\lambda}(x)+2\lambda\beta(x)F_{1,\lambda}(x)F_{2,\lambda}(x)\left(\lambda F_{1,\lambda}(x)+1\right)\\
&\quad-\beta(x)F_{1,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)\\
&=\beta(x)F_{1,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)[\lambda^2\{\alpha(x)-\beta(x)\}F_{2,\lambda}(x)+2\lambda F_{2,\lambda}(x)-1]\\
&\geqslant 0.3939\beta(x)F_{1,\lambda}(x)(\lambda F_{1,\lambda}(x)+1) \text{ for all } x\in(\gamma_\lambda,c_{2,\lambda}], \text{ for } \lambda\geqslant 2. \quad (2.10.28)
\end{aligned}
$$

To prove Lemma 2.6.5, it thus suffices to show that for $\lambda\geqslant 2$ and $x\in(\gamma_\lambda,c_{2,\lambda}]$, the expressions on the right sides of (2.10.26), (2.10.27) and (2.10.28) add up to a strictly positive quantity. In other words, we need to show, for all $x\in(\gamma_\lambda,c_{2,\lambda}]$ and each $\lambda\in[2,2.5)$, that

$$
\begin{aligned}
&0.14446\lambda\{\alpha(x)-\beta(x)\}F_{1,\lambda}(x)F_{2,\lambda}(x)+0.3939\beta(x)F_{1,\lambda}(x)(\lambda F_{1,\lambda}(x)+1)\\
&\qquad\qquad > 0.130043\{\alpha(x)-\beta(x)\}(\lambda F_{1,\lambda}(x)+1)^2F_{2,\lambda}(x).
\end{aligned}
$$

Separating the terms involving the factor $\alpha(x)$ from those that involve the factor $\beta(x)$, we find that this is equivalent to proving, for all $x\in(\gamma_\lambda,c_{2,\lambda}]$, for all $\lambda\in[2,2.5)$,

$$
\begin{aligned}
&\beta(x)[0.3939\lambda(F_{1,\lambda}(x))^2+0.3939F_{1,\lambda}(x)+0.130043\lambda^2(F_{1,\lambda}(x))^2F_{2,\lambda}(x)\\
&+0.115626\lambda F_{1,\lambda}(x)F_{2,\lambda}(x)+0.130043F_{2,\lambda}(x)]-\alpha(x)[0.130043\lambda^2(F_{1,\lambda}(x))^2F_{2,\lambda}(x)\\
&+0.115626\lambda F_{1,\lambda}(x)F_{2,\lambda}(x)+0.130043F_{2,\lambda}(x)]>0. \quad (2.10.29)
\end{aligned}
$$

The idea now is to demarcate the terms in (2.10.29) into a few different groups, and then show that the sum of the terms in each such group is strictly positive for $\lambda \geqslant 2$. In this endeavour, we make use of certain facts that we enumerate below:

1. Recall from Lemma 2.6.1 that η_λ increases for $\lambda \in [2, 2.43634]$ and decreases for $\lambda \in [2.43634, 2.5]$, and $\eta_2 \approx 0.523928 < 0.528322 \approx \eta_{2.5}$. Therefore, $\min\{\eta_\lambda : \lambda \in [2, 2.5)\} = \eta_2 \approx 0.523928$.

2. Since, substituting from (2.10.6), we have
$$\frac{d}{d\lambda}[\lambda e^{-\eta_\lambda} - \lambda] = \frac{e^{-\eta_\lambda} - 1 - \eta_\lambda}{1 + \lambda\eta_\lambda e^{-\eta_\lambda} + \eta_\lambda} < 0,$$
we conclude that $e^{\lambda e^{-\eta_\lambda} - \lambda}$ is strictly decreasing as a function of λ for all $\lambda > 0$. Its minimum value for $\lambda \in [2, 2.5]$, attained at $\lambda = 2.5$, is ≈ 0.358432.

3. Since, substituting from (2.10.6), we have
$$\frac{d}{d\lambda}[\lambda e^{-\eta_\lambda}] = \frac{e^{-\eta_\lambda}[1 + \lambda\eta_\lambda]}{1 + \lambda\eta_\lambda e^{-\eta_\lambda} + \eta_\lambda} > 0,$$
the function $\lambda e^{-\eta_\lambda}$ is strictly increasing in λ for all $\lambda > 0$, so that its minimum value for $\lambda \in [2, 2.5]$, attained at $\lambda = 2$, is ≈ 1.18438.

4. We make use of the definitions of $\alpha(x)$ and $\beta(x)$ from §2.6, namely that
$$\alpha(x) = e^{-\lambda F_{2,\lambda}(x)} \text{ and } \beta(x) = e^{\lambda F_{1,\lambda}(x) - \lambda F_{2,\lambda}(x) - \lambda}.$$

5. We also make use of the decreasing nature of $F_{2,\lambda}$ and $F_{1,\lambda}$ from Lemma 2.2.2, that $c_{2,\lambda}$ is the fixed point of $F_{2,\lambda}$, and the definition of γ_λ from (2.6.4). In particular, we make use of the inequalities
$$F_{2,\lambda}(\gamma_\lambda) > F_{2,\lambda}(x) \geqslant F_{2,\lambda}(c_{2,\lambda}) \text{ and } F_{1,\lambda}(x) = e^{-\lambda x} \geqslant e^{-\lambda c_{2,\lambda}} = e^{-\eta_\lambda}$$
for all $x \in (\gamma_\lambda, c_{2,\lambda}]$.

The first such group of terms from (2.10.29) that we consider is as follows: for $x \in (\gamma_\lambda, c_{2,\lambda}]$ and $\lambda \in [2, 2.5)$,

$$\beta(x)[0.3939\lambda(F_{1,\lambda}(x))^2 + 0.130043\lambda^2(F_{1,\lambda}(x))^2 F_{2,\lambda}(x)] - 0.130043\lambda^2\alpha(x)(F_{1,\lambda}(x))^2 F_{2,\lambda}(x)$$

$$\begin{aligned}
&\geqslant \lambda(F_{1,\lambda}(x))^2[0.3939\beta(x)+0.130043\lambda\beta(x)F_{2,\lambda}(c_{2,\lambda})-0.130043\lambda\alpha(x)F_{2,\lambda}(\gamma_\lambda)]\\
&=\lambda(F_{1,\lambda}(x))^2\left[0.3939\beta(x)+0.130043\beta(x)\lambda c_{2,\lambda}-0.130043\lambda\alpha(x)\cdot\frac{1}{\lambda}\right]\\
&=\lambda(F_{1,\lambda}(x))^2[0.3939\beta(x)+0.130043\beta(x)\eta_\lambda-0.130043\alpha(x)]\\
&\geqslant \lambda(F_{1,\lambda}(x))^2[0.3939\beta(x)+0.130043\beta(x)\cdot 0.523928-0.130043\alpha(x)]\\
&=\lambda(F_{1,\lambda}(x))^2 e^{-\lambda F_{2,\lambda}(x)}[0.462033e^{\lambda F_{1,\lambda}(x)-\lambda}-0.130043]\\
&\geqslant \lambda(F_{1,\lambda}(x))^2 e^{-\lambda F_{2,\lambda}(x)}[0.462033e^{\lambda e^{-\eta_\lambda}-\lambda}-0.130043]\\
&\geqslant \lambda(F_{1,\lambda}(x))^2 e^{-\lambda F_{2,\lambda}(x)}[0.462033\cdot 0.358432-0.130043]=0.035564\lambda(F_{1,\lambda}(x))^2\alpha(x)>0.
\end{aligned}\tag{2.10.30}$$

The second group of terms from (2.10.29) that we consider, for $x\in(\gamma_\lambda,c_{2,\lambda}]$ and $\lambda\in[2,2.5)$, is as follows:

$$\begin{aligned}
&\beta(x)[0.3939F_{1,\lambda}(x)+0.115626\lambda F_{1,\lambda}(x)F_{2,\lambda}(x)]-0.115626\lambda\alpha(x)F_{1,\lambda}(x)F_{2,\lambda}(x)\\
&\geqslant \beta(x)[0.3939F_{1,\lambda}(x)+0.115626\lambda F_{1,\lambda}(x)F_{2,\lambda}(c_{2,\lambda})]-0.115626\lambda\alpha(x)F_{1,\lambda}(x)F_{2,\lambda}(\gamma_\lambda)\\
&=\beta(x)[0.3939F_{1,\lambda}(x)+0.115626\lambda c_{2,\lambda}F_{1,\lambda}(x)]-0.115626\lambda\alpha(x)F_{1,\lambda}(x)\cdot\frac{1}{\lambda}\\
&=\beta(x)[0.3939F_{1,\lambda}(x)+0.115626\eta_\lambda F_{1,\lambda}(x)]-0.115626\alpha(x)F_{1,\lambda}(x)\\
&\geqslant \beta(x)[0.3939F_{1,\lambda}(x)+0.115626\cdot 0.523928F_{1,\lambda}(x)]-0.115626\alpha(x)F_{1,\lambda}(x)\\
&=F_{1,\lambda}(x)e^{-\lambda F_{2,\lambda}(x)}[0.454479e^{\lambda F_{1,\lambda}(x)-\lambda}-0.115626]\\
&\geqslant F_{1,\lambda}(x)e^{-\lambda F_{2,\lambda}(x)}[0.454479e^{\lambda e^{-\eta_\lambda}-\lambda}-0.115626]\\
&\geqslant F_{1,\lambda}(x)e^{-\lambda F_{2,\lambda}(x)}[0.454479\cdot 0.358432-0.115626]=0.047274F_{1,\lambda}(x)\alpha(x)>0.
\end{aligned}\tag{2.10.31}$$

We now add the remaining terms of (2.10.29) (i.e. the ones that have not been taken into account in the expressions on the left sides of (2.10.30) and (2.10.31)) with the expressions on the right sides of (2.10.30) and (2.10.31) to get

$$\begin{aligned}
&0.130043\beta(x)F_{2,\lambda}(x)-0.130043\alpha(x)F_{2,\lambda}(x)+0.035564\lambda(F_{1,\lambda}(x))^2\alpha(x)+0.047274F_{1,\lambda}(x)\alpha(x)\\
&=0.130043F_{2,\lambda}(x)\{\beta(x)-\alpha(x)\}+0.035564\lambda(F_{1,\lambda}(x))^2\alpha(x)+0.047274F_{1,\lambda}(x)\alpha(x)\\
&=0.130043F_{2,\lambda}(x)e^{-\lambda F_{2,\lambda}(x)}[e^{\lambda F_{1,\lambda}(x)-\lambda}-1]+0.035564\lambda(F_{1,\lambda}(x))^2\alpha(x)+0.047274F_{1,\lambda}(x)\alpha(x)\\
&\geqslant 0.130043F_{2,\lambda}(x)\alpha(x)[e^{\lambda e^{-\eta_\lambda}-\lambda}-1]+0.035564\lambda(F_{1,\lambda}(x))^2\alpha(x)+0.047274F_{1,\lambda}(x)\alpha(x)\\
&\geqslant 0.130043F_{2,\lambda}(x)\alpha(x)[0.358432-1]+0.035564\lambda(F_{1,\lambda}(x))^2\alpha(x)+0.047274F_{1,\lambda}(x)\alpha(x)\\
&=-0.083431F_{2,\lambda}(x)\alpha(x)+0.035564\lambda(F_{1,\lambda}(x))^2\alpha(x)+0.047274F_{1,\lambda}(x)\alpha(x)
\end{aligned}$$

$$\geqslant -0.083431F_{1,\lambda}(x)\alpha(x)+0.035564\lambda(F_{1,\lambda}(x))^2\alpha(x)+0.047274F_{1,\lambda}(x)\alpha(x)$$
$$\geqslant F_{1,\lambda}(x)\alpha(x)[0.035564\lambda F_{1,\lambda}(x)-0.036157]$$
$$\geqslant F_{1,\lambda}(x)\alpha(x)[0.035564\lambda e^{-\eta_\lambda}-0.036157]$$
$$\geqslant F_{1,\lambda}(x)\alpha(x)[0.035564\cdot 1.18438-0.036157]=0.005964F_{1,\lambda}(x)\alpha(x)>0, \tag{2.10.32}$$

where make use of the inequality $F_{2,\lambda}(x)\leqslant F_{1,\lambda}(x)$ for $x\in[0,c_{1,\lambda}]$ (as shown in the proof of (2.2.14)). From (2.10.30), (2.10.31) and (2.10.32), we conclude that the inequality in (2.10.29) does hold, and this brings us to the end of the proof. □

Proof of Lemma 2.6.6. Recall that, after taking out the common factor of $\lambda\{\alpha(x)-\beta(x)\}F_{1,\lambda}(x)F_{2,\lambda}(x)$, our objective here is to show that

$$2\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1)>1 \text{ for } x\in[\delta_\lambda,\gamma_\lambda], \text{ for all } \lambda\geqslant 2. \tag{2.10.33}$$

The idea now is to come up with suitable lower bounds on each of $\beta(x)$ and $\lambda F_{1,\lambda}(x)$, the first of which we accomplish by examining whether $\beta(x)$ exhibits monotonicity as a function of x. For the second, we recall from Lemma 2.2.2 that $F_{1,\lambda}$ is strictly decreasing on $[0,1]$, so that

$$F_{1,\lambda}(x)\geqslant F_{1,\lambda}(\gamma_\lambda) \text{ for } x\in[\delta_\lambda,\gamma_\lambda]. \tag{2.10.34}$$

Recalling from §2.6 the expression for the derivative of $\beta(x)$ with respect to x, we have, for $x\in[0,\gamma_\lambda]$:

$$\begin{aligned}\beta'(x)&=\lambda^2\beta(x)\left\{-F_{1,\lambda}(x)+\left(\lambda F_{1,\lambda}(x)+1\right)F_{2,\lambda}(x)\right\}\\&\geqslant\lambda^2\beta(x)[-F_{1,\lambda}(x)+(\lambda F_{1,\lambda}(x)+1)F_{2,\lambda}(\gamma_\lambda)]=\lambda\beta(x)>0,\end{aligned}$$

where we obtain the inequality since $F_{2,\lambda}$ is strictly decreasing on $[0,c_{1,\lambda}]$ (again from Lemma 2.2.2). This shows that $\beta(x)$ is a strictly increasing function of x for $x\in[0,\gamma_\lambda]$, so that

$$\beta(x)\geqslant\beta(\delta_\lambda) \text{ for } x\in[\delta_\lambda,\gamma_\lambda]. \tag{2.10.35}$$

Pur next aim is to obtain suitable expressions for $F_{1,\lambda}(\gamma_\lambda)$ and $\beta(\delta_\lambda)$. Recalling the definitions of $\alpha(x)$ and $\beta(x)$ from §2.6, and the definitions of δ_λ and γ_λ from (2.6.4), we have:

$$F_{2,\lambda}(\delta_\lambda)=\exp\{\lambda e^{-\lambda\delta_\lambda}-\lambda\delta_\lambda-\lambda\}=\frac{5}{4\lambda}$$

$$\Longrightarrow \beta(\delta_\lambda) = e^{\lambda F_{1,\lambda}(\delta_\lambda)-\lambda F_{2,\lambda}(\delta_\lambda)-\lambda} = \exp\{\lambda e^{-\lambda\delta_\lambda} - \frac{5}{4} - \lambda\} = \frac{5}{4\lambda}e^{\lambda\delta_\lambda - \frac{5}{4}}; \tag{2.10.36}$$

and

$$F_{2,\lambda}(\gamma_\lambda) = \exp\{\lambda F_{1,\lambda}(\gamma_\lambda) - \lambda\gamma_\lambda - \lambda\} = \frac{1}{\lambda} \Longrightarrow \lambda F_{1,\lambda}(\gamma_\lambda) = \lambda\gamma_\lambda + \lambda - \ln\lambda. \tag{2.10.37}$$

Combining our findings from (2.10.34), (2.10.35), (2.10.36) and (2.10.37), for $x \in (\delta_\lambda, \gamma_\lambda]$ and $\lambda \geqslant 2$, we have

$$\begin{aligned} 2\lambda\beta(x)(\lambda F_{1,\lambda}(x)+1) &\geqslant 2\lambda\beta(\delta_\lambda)(\lambda F_{1,\lambda}(\gamma_\lambda)+1) \\ &= \frac{5}{2}e^{\lambda\delta_\lambda-\frac{5}{4}}(\lambda\gamma_\lambda + \lambda - \ln\lambda + 1) > \frac{5}{2}e^{-\frac{5}{4}}(\lambda - \ln\lambda + 1). \end{aligned}$$

Since $\lambda - \ln\lambda$ is strictly increasing for $\lambda > 1$, we have $\frac{5}{2}e^{-\frac{5}{4}}(\lambda - \ln\lambda + 1) \geqslant \frac{5}{2}e^{-\frac{5}{4}}(2 - \ln 2 + 1) \approx 1.6523 > 1$ for $\lambda \geqslant 2$. This completes the proof of (2.10.33) and brings us to the end of the proof of Lemma 2.6.6. □

2.10.4 Proof of Lemma 2.7.1

Recall that our objective here is to prove, for given $c > 0$ and $i \in \mathbb{N}$ with $i \geqslant 2$, the inequality $H_{2,\lambda}(c\lambda^{-i}) < c\lambda^{-i}$ for all λ sufficiently large.

To begin with, we note that $c_{2,\lambda} > c\lambda^{-i}$ for all λ sufficiently large, since (2.5.1) tells us that $\lambda^2 c_{2,\lambda}$, and hence $\lambda^i c_{2,\lambda}$ as well, approaches ∞ as $\lambda \to \infty$.

From (2.5.2), and applying Taylor expansion, we have

$$F_{1,\lambda}(c\lambda^{-i}) = e^{-c\lambda^{-i+1}} \text{ and } F_{2,\lambda}(c\lambda^{-i}) = \exp\{\lambda e^{-c\lambda^{-i+1}} - c\lambda^{-i+1} - \lambda\} = \exp\{-c\lambda^{-i+2} + O(\lambda^{-i+1})\}.$$

Note that the second expression yields

$$\lim_{\lambda\to\infty} F_{2,\lambda}(c\lambda^{-2}) = e^{-c} \text{ and } \lim_{\lambda\to\infty} F_{2,\lambda}(c\lambda^{-i}) = 1 \text{ for each } i \geqslant 3.$$

Thus, given $0 < \varepsilon < e^{-c} < 1$, we have, for all λ sufficiently large,

$$F_{2,\lambda}(c\lambda^{-i}) > M > 0, \text{ with } M = e^{-c} - \varepsilon \text{ for } i = 2 \text{ and } M = 1 - \varepsilon \text{ for } i \geqslant 3. \tag{2.10.38}$$

Next, using the inequality $1-e^{-x} \leqslant x$ for $x \geqslant 0$, we have

$$1-e^{\lambda F_{1,\lambda}(c\lambda^{-i})-\lambda} \leqslant \lambda\left\{1-F_{1,\lambda}(c\lambda^{-i})\right\} = \lambda(1-e^{-c\lambda^{-i+1}}) \leqslant c\lambda^{-i+2}. \tag{2.10.39}$$

Thus, for all λ sufficiently large, using (2.10.38), (2.10.39) and the fact that G_λ is increasing, we have

$$\begin{aligned} H_{2,\lambda}\left(\frac{c}{\lambda^i}\right) &= G_\lambda\left[e^{-\lambda F_{2,\lambda}\left(\frac{c}{\lambda^i}\right)}\left\{1-e^{\lambda F_{1,\lambda}\left(\frac{c}{\lambda^i}\right)-\lambda}\right\}\right] \\ &< G_\lambda\left[e^{-\lambda M}\cdot\frac{c}{\lambda^{i-2}}\right] = \exp\left\{\frac{c}{\lambda^{i-3}}e^{-\lambda M}-\lambda\right\}. \end{aligned} \tag{2.10.40}$$

Since $i \geqslant 2$, we have $c\lambda^{-i+3}e^{-\lambda M} \leqslant c\lambda e^{-\lambda M} \to 0$, so that (2.10.40) is bounded above by $e^{-\lambda/2}$ for all λ sufficiently large, and this in turn is bounded above by $c\lambda^{-i}$ for all λ sufficiently large (since $\lambda^i e^{-\lambda/2} \to 0$ as $\lambda \to \infty$). This establishes the desired inequality $H_{2,\lambda}(c\lambda^{-i}) < c\lambda^{-i}$ for all λ sufficiently large.

Chapter 3

The three-neighbourhood game and its corresponding probabilistic cellular automata

Preface

This chapter is based on the following paper:

- Bhasin D., Karmakar S., Podder M. and Roy S. "On a class of probabilistic cellular automata with size-3 neighbourhood and their application in percolation games" – *Electronic Journal of Probability*, Volume 28, Pages 1–60, 2023. DOI: http://dx.doi.org/10.1214/23-EJP1046.

3.1 Introduction

In this chapter, we study the three-neighbourhood site percolation game and its corresponding probabilistic cellular automata. The most general setup of the *percolation games* we study was introduced in § 1.1. We now come to an overview of the percolation game we are concerned here, corresponding to the parameters p and q, with $(p,q) \in \mathscr{S}$, where

$$\mathscr{S} = \{(p', q') \in [0,1]^2 : 0 < p' + q' \leqslant 1\}, \tag{3.1.1}$$

are fixed *a priori*. We begin by assigning, independently to every vertex of $\mathbb{Z}^2$, a (random) label that reads *trap* with probability p, *target* with probability q, and *open* with the remaining probability $r = 1 - p - q$. A token is placed at a vertex of $\mathbb{Z}^2$, termed the *initial vertex*, at the start of the game, and two players take turns to make moves. A *move* involves relocating the token from its current position (x,y) (say) to any vertex in the set $\mathrm{Out}(x,y)$, where

$$\mathrm{Out}(x,y) = \{(x, y+1), (x+1, y+1), (x+2, y+1)\}. \tag{3.1.2}$$

A player loses if she is forced to relocate the token to a vertex labeled a trap, and wins if she is able to relocate the token to a vertex labeled a target. The game may also continue forever, with neither player being able to reach a target nor being able to force her opponent to fall into a trap, and in this case, we say that the game results in a draw.

This game shares close connections with a PCA and its envelope PCA, denoted $F_{p,q}$ and $\widehat{\mathrm{F}}_{p,q}$ respectively (see § 3.2.2). The primary goal of this chapter is to prove Theorem 3.1.1, stated below, the first assertion of which establishes ergodicity (see Definition 1.1.1) for the PCA $F_{p,q}$ (and thereby also for $\widehat{\mathrm{F}}_{p,q}$), while the second concerns itself with the outcome of draw in the percolation game described above.

Theorem 3.1.1. *For all $(p,q) \in \mathscr{S}$, the PCA $F_{p,q}$ is ergodic, and the probability of draw in the percolation game described above is* 0*.*

We give a brief outline of the essential constituents of the proof of this theorem. Recall the envelope PCA $\widehat{\mathrm{F}}_{p,q}$ corresponding to $F_{p,q}$, with alphabet $\{W, D, L\}$, alluded to above. For each $(p,q) \in \mathscr{S}$, it is shown, via connections explained in §3.2.3, that the probability of the event that the percolation game starting from the origin in $\mathbb{Z}^2$ culminates in a draw is equal to the probability that the symbol D occupies the origin in $\mathbb{Z}$ under a certain stationary distribution (see Definition 1.1.1) for $\widehat{\mathrm{F}}_{p,q}$. This stationary distribution for $\widehat{\mathrm{F}}_{p,q}$ is the same as the probability distribution of winning (for the first player), losing (for the first player) and draw positions (indicated, respectively, by the states W, L and D that are introduced at the beginning of §3.2.3) on *any* horizontal line $H_k = \{(x,k) : x \in \mathbb{Z}\}$ of $\mathbb{Z}^2$, for our percolation game. Next, Theorem 3.2.3 is used to establish that the probability of appearance of D at any given site in $\mathbb{Z}$ is 0 under every stationary distribution for $\widehat{\mathrm{F}}_{p,q}$, for each $(p,q) \in \mathscr{S}$, leading to the conclusion that the probability of draw in our percolation game is 0 for each $(p,q) \in \mathscr{S}$. Finally, Proposition 3.2.1 guarantees that the probability of draw in our percolation game is 0 if and only if the PCA $F_{p,q}$ is ergodic, yielding the conclusion that $F_{p,q}$ is ergodic for each $(p,q) \in \mathscr{S}$.

The key tool used in the proof of Theorem 3.2.3 is the technique of *weight functions*. In addition to proving the main result, Theorem 3.1.1, of this work, we include an extensive discussion on the class of percolation games to which this technique may be applicable, with possibly suitable modifications, to explore the 'regimes' (depending on the underlying parameters, such as (p,q) in our case), referred to as *critical regions*, in which such games have probability 0 of culminating in draws. We also provide insight into the class of *game theoretic* PCAs to which the technique of weight functions may be applicable in order to establish ergodicity properties, or to gain an understanding of the condition(s) to be imposed on the underlying parameter(s) of these PCAs that guarantee ergodicity. We also provide a formal game theoretic formulation of our problem, following which we establish connections between our work and the existing literature, speculate on the various directions in which future research can be carried out (for instance, by generalizing the set of actions, considering arbitrary mover-sequences, generalizing the event of draw itself, exploring possible monotonicity properties of the probabilities of draw, studying percolation games on higher dimensions, studying the values of such games etc.).

3.1.1 Organization of this chapter

The percolation game we investigate in this work is discussed and illustrated in §3.2.1, the PCA $F_{p,q}$ is formally introduced in §3.2.2, and its envelope $\widehat{\mathrm{F}}_{p,q}$, along with the deep ties it exhibits with the game via the game's recurrence relations, is elaborated on in §3.2.3. The results forming the crux of this chapter (apart from Theorem 3.1.1 that has already been stated) are stated in Proposition 3.2.1, Proposition 3.2.2 and Proposition 3.2.3, while the lemma required to prove Proposition 3.2.1 and Proposition 3.2.2 is discussed in §3.3. Mathematically, the most significant section of this chapter is §3.4, in which the method of weight functions is employed to prove Proposition 3.2.3, and the pivotal steps via which this is accomplished are outlined in §3.4.1 through §3.4.7. We go on to provide a formal, game theoretic formulation of our percolation game, along with a detailed description of relevant game theoretic terminology, in §3.5. In §3.6, we include discussions on several directions in which future research closely relating to this work can be carried out, such as examining the probabilities of draw in games where the set of actions is generalized, or where arbitrary mover-sequences are considered, or where the event of draw itself is generalized, percolation games on lattices of higher dimensions, possible monotonicity properties of the probabilities of draw, studying values of percolation games etc. Finally, in §3.7, we speculate on the various other settings (both in terms of percolation games and 'game theoretic' PCAs that we introduce in §3.7) in which we believe that the technique of weight functions can be applied to explore the probability of draw for the game under consideration / ergodicity of the PCA under consideration.

3.2 The principal objects studied in this chapter

3.2.1 Our percolation game

The permitted moves in our percolation game are illustrated via the directed (as indicated by the arrowheads), dashed, black lines in Figure3.1.

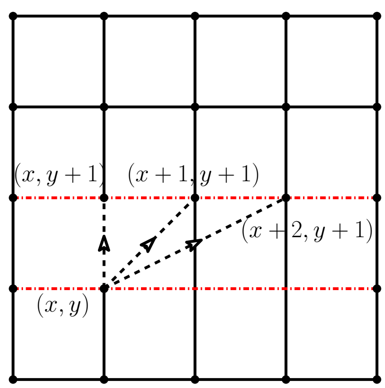


Figure 3.1: $\mathrm{Out}(x,y)$ in our percolation game

We recall that each vertex (or site) of $\mathbb{Z}^2$ is assigned, independent of all else, a label that is a trap with probability p, a target with probability q, and open with probability $r = 1 - p - q$, where $(p,q) \in \mathscr{S}$, with $\mathscr{S}$ as in (3.1.1). Starting from an initial vertex, the two players take turns to move the token from its current position, say (x,y), to any vertex in $\mathrm{Out}(x,y)$. A player wins if she is able to move the token to a target or if her opponent is forced to move it to a trap. We clarify here that, once $\mathbb{Z}^2$ has been endowed with a trap / target / open labeling, the assignment is revealed in its entirety to both the players *before* the game begins (hence, this is a *perfect information game*). Thereafter, to say that the corresponding game is won by a player is to assert that she has a strategy which, when employed, allows her to win no matter what strategy her opponent adopts. The game continues for as long as the token does not land on a site that is marked either a target or a trap, and this could happen indefinitely, leading to a draw. The primary question of interest to us is the same as that in [60], i.e. for what values of the parameters p and q the game exhibits a positive probability of draw. As previously mentioned, the proof pivots upon the connection this game has with the envelope PCA $\widehat{\mathrm{F}}_{p,q}$ (to be introduced formally in §3.2.3).

In this context, we mention [12] that studies the following two-player combinatorial game on any graph: starting from an initial vertex, the players take turns to move a token, where a move involves relocating the token from the vertex where it is currently situated to a neighbor of that vertex that has not yet been visited. A player who is unable to move loses the game. On $\mathbb{Z}^2$, in which odd and even sites, independently, are marked closed (i.e. forbidden from being visited by the token) with probabilities p and q respectively, it is shown that the game has probability 0 of ending in a draw provided closed sites of one parity are sufficiently rare compared to closed sites of the other parity. This question, however, remains open when the percolation parameters p and q are equal. Motivations for studying the games addressed in [12] include their deep connections with maximum-cardinality matchings in graphs, and in particular, the ways in which draws in these games relate to sensitivity of such matchings to boundary conditions.

3.2.2 Our PCA

We now come to a detailed description of the PCA $F_{p,q}$, with parameters $(p,q) \in \mathscr{S}$. This is a 1-dimensional PCA, with alphabet $\mathscr{A} = \{W, L\}$ and neighborhood $\mathscr{N} = \{0,1,2\}$, so that $F_{p,q}\eta(n)$ is a random variable whose probability distribution is a function of $\eta(n)$, $\eta(n+1)$ and $\eta(n+2)$ for all $n \in \mathbb{Z}$. More precisely, the stochastic matrix $\varphi_{p,q} : \mathscr{A}^3 \times \mathscr{A} \to [0,1]$ for this PCA is defined via the equations:

$$\varphi_{p,q}(W,W,W;b) = \begin{cases} p & \text{if } b = W, \\ 1-p & \text{if } b = L, \end{cases} \tag{3.2.1}$$

and

$$\varphi_{p,q}(a_0,a_1,a_2;b) = \begin{cases} 1-q & \text{if } b = W, \\ q & \text{if } b = L, \end{cases} \quad \text{for all } (a_0,a_1,a_2) \in \mathscr{A}^3 \setminus \{(W,W,W)\}. \tag{3.2.2}$$

The (stochastic) update rule for the automaton $F_{p,q}$ can be illustrated pictorially via Figure 3.2.

As defined in [74], given $\varepsilon > 0$, a PCA, say Φ, is said to be an *ε-perturbation* of a deterministic CA, say F, if Φ and F share the same alphabet, say S, and the same neighborhood, and the stochastic update rule φ of Φ satisfies $\varphi(a_1,a_2,\ldots,a_m; f(a_1,a_2,\ldots,a_m)) \geqslant 1-\varepsilon$ for all

$(a_1, a_2, \ldots, a_m) \in S^m$, where f is the local update rule for F. In other words, given any input $(a_1, a_2, \ldots, a_m) \in S^m$, the PCA Φ outputs the same symbol as F with probability at least $1-\varepsilon$, and alters it with the remaining probability. The PCA $F_{p,q}$ described above can be derived from a deterministic CA, F, and a stochastic noise, θ, via such a perturbation. The local update rule f of F is given by $f(a_0, a_1, a_2) = 1 - \max\{a_0, a_1, a_2\}$, for all $a_0, a_1, a_2 \in \mathscr{A}$, while the noise $\theta(a,b)$, for $a, b \in \{W, L\}^2$, transforms L into W with probability p (i.e. $\theta(L,W) = p$) and transforms W into L with probability q (i.e. $\theta(W,L) = q$). It is easily verified that $\varphi_{p,q}(a_0, a_1, a_2; b) = \theta(f(a_0, a_1, a_2), b)$ for all $a_0, a_1, a_2, b \in \mathscr{A}$. Moreover, we see that $\theta(a,b) = (1-\varepsilon)\delta_a(b) + \varepsilon g(b)$ for $a, b \in \mathscr{A}$, where $\varepsilon = p+q$, $\delta_a(b) = 1$ when $b = a$ and $\delta_a(b) = 0$ when $b \neq a$, and $g(0) = 1 - g(1) = \frac{p}{p+q}$. Thus, our PCA $F_{p,q}$ is obtained from the CA F via perturbation using the memoryless zero-range noise θ (we refer the reader to [74] for further reading on such noise-perturbed CAs).

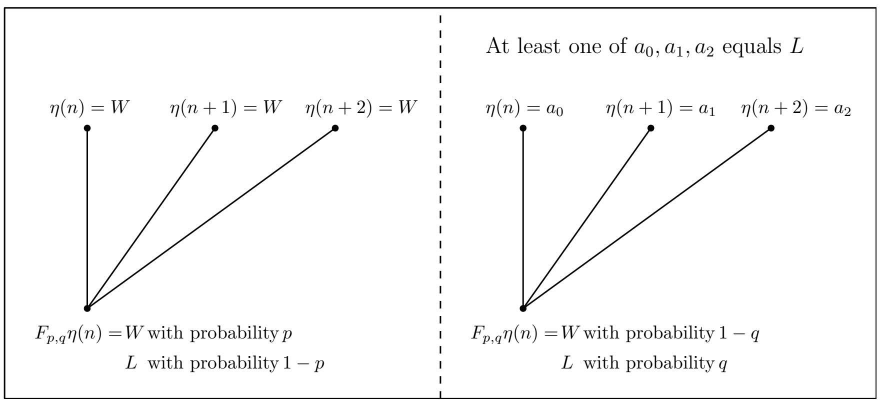


Figure 3.2: The stochastic rules that define our PCA $F_{p,q}$

The PCA $A_{p,q}$ that is studied in [60] is an elementary one that bears some resemblance with $F_{p,q}$ in that, following the notation in (1.1.1), its stochastic matrix $\varphi : \mathscr{A}^2 \times \mathscr{A}$ is defined by the equations:

$$\varphi(W,W;b) = \begin{cases} p & \text{if } b = W, \\ 1-p & \text{if } b = L, \end{cases} \quad \text{and} \quad \varphi(a_0, a_1; b) = \begin{cases} 1-q & \text{if } b = W, \\ q & \text{if } b = L, \end{cases} \quad \text{for all } (a_0, a_1) \in \mathscr{A}^2 \setminus \{(W,W)\}.$$

We draw the reader's attention to the primary contrast between $F_{p,q}$ and $A_{p,q}$: whereas in $A_{p,q}$, we draw upon information regarding the states of 2 consecutive sites, n and $n+1$, in order to decide the (random) updated state of the site n, in $F_{p,q}$, we draw upon information regarding the states of 3 consecutive sites, n, $n+1$ and $n+2$, to decide the (random) updated state of the site n. It is, therefore, expected that the latter will have a somewhat-more-involved underlying dependence structure, and establishing ergodicity results for $F_{p,q}$ indeed proves to be a considerably more challenging feat compared to that for $A_{p,q}$, at least as far as using the method of weight functions proposed and implemented in [60] is concerned.

Some discussion on how important the question of ergodicity (or the lack thereof) of PCAs is, and how difficult this question is to resolve under various circumstances, is now in order. It has been found to be notoriously difficult to construct a CA whose trajectories, starting from different initial conditions and evolving under repeated applications of the local update rule, remain distinguishable from each other if even the slightest positive noise is incorporated into the CA (for instance, see [74]). In other words, most CAs tend to forget their initial conditions under the influence of even the smallest amount of noise, and this is reflected in the ergodicity of the resulting PCA. The renowned *positive rates conjecture* states that *all* PCAs defined on $\mathbb{Z}$ and satisfying $\varphi(a_1, a_2, \ldots, a_{|\mathcal{N}|}; b) > 0$ for all $a_1, a_2, \ldots, a_{|\mathcal{N}|}, b \in \mathscr{A}$ (referred to as the *positive rates* condition) are ergodic. An extremely complicated example proposed by [50] refutes this conjecture, but the fascinating question still remains as to whether all sufficiently simple, naturally occurring 1-dimensional PCAs with positive rates are ergodic. There are, however, examples of d-dimensional PCAs for $d \geqslant 2$, such as Glauber dynamics for the Ising model at low temperatures, that are known to be non-ergodic. We note here that both $F_{p,q}$ and $A_{p,q}$ have positive rates as long as *both* p and q are strictly positive.

Multiple sources (see discussions in [29], [60] and [23]) reiterate that in general, even if the answer can be guessed from heuristics or simulations, *rigorously* proving whether a given PCA is ergodic or not is a notoriously difficult problem, and is shown to be algorithmically undecidable in [28] and [37]. Under the assumption of left-right symmetry (which guarantees $\varphi(L,W;W) = \varphi(W,L;W)$), an elementary PCA is determined by the parameters $\varphi(W,W;W)$, $\varphi(L,L;W)$ and $\varphi(L,W;W)$ (recall these notations from (1.1.1)). The many existing techniques that have been developed to study ergodicity can take care of ergodicity questions for such PCAs over more than 90% of the volume of the cube $[0,1]^3$ defined by these 3 parameters ([37]). However, when p and q are small, $A_{p,q}$ belongs to a domain of this cube where none of these techniques works, and this is where [60] comes in with their brilliant idea of employing weight functions. To the best of our knowledge, the problem of establishing ergodicity results for $F_{p,q}$, for *all* $(p,q) \in \mathscr{S}$, is an even more challenging one that has, so far, remained open, and we, in this chapter, utilize the method of weight functions to provide a concrete proof of the first assertion of Theorem 3.1.1.

Apart from being an interesting object to study in its own right, motivation to investigate $F_{p,q}$ can be found from the connections that such PCAs have with the problem of enumerating *directed animals*, as was first pointed out by [36], and further explored in [20]. The relation between PCAs and the problem of enumerating directed animals is elucidated upon in [73]. In [22], the generating functions for directed animals on square and triangular lattices, as well as on decorated square and triangular lattices, are obtained, further demonstrating the usefulness of these PCAs.

In [20], a directed graph G is considered, in which a vertex v is called a *successor* to a vertex u if there exists a directed edge *from u to v*. Given any subset S of vertices of G, a *directed animal A* with *source S* is a subset of vertices of G such that $S \subset A$, and every $v \in A \setminus S$ can be reached from S via a directed path all of whose vertices are in A. The number of vertices in A, denoted $|A|$, is referred to as the *area* of A. In order to count the number of possible directed animals on G with a given source S and a given area n, a suitable *generating function* $G_S^G(x) = x^{|S|} + \cdots$ is considered, where $\cdots$ represents a sum of monomials whose degrees are at least $|S|+1$.

Next, [20] defines a *particle system* or *gas occupation* on G as a map X from the set of vertices of G to $\{0,1\}$, with respect to which any vertex u is said to be *occupied* if $X_u = 1$. When X is random, this is referred to as a *random gas model*. A random colouring of the vertices of G with colours a and b is considered. Letting C_u denote the colour of the vertex u for each u in G, the random variables C_u are independent and identically distributed, $C_u = a$ with probability p and $C_u = b$ with probability $1-p$. The desired gas model is now defined with respect to this random colouring, as follows: for each vertex u of G,

$$X_u = \begin{cases} 0 & \text{if } C_u = a, \\ \prod_{v \text{ successor of } u} (1 - X_v) & \text{if } C_u = b. \end{cases} \tag{3.2.3}$$

One of the main results of [20] states that if R_u, for any vertex u of G, is the radius of convergence of the generating function $G_{\{u\}}^G(x)$, and $(1-p) \in [0, R_u)$, then $\mathbf{E}[X_u] = -G_{\{u\}}^G(-1+p)$. Therefore, knowledge about the random colouring scheme described above is valuable for the understanding of these generating functions.

It is fairly immediate that the above-mentioned random colouring scheme can be represented by a special case of our PCA $F_{p,q}$ when we consider $q = 0$, and in place of G, we consider either a 2-dimensional infinite directed lattice in which the successors of any site (x,y) are $(x,y+1)$, $(x+1,y+1)$ and $(x+2,y+1)$ (see Figure 3.3), or a 2-dimensional infinite directed lattice in which the successors of any site (x,y) are $(x+2,y)$, $(x+1,y+1)$ and $(x,y+2)$ (see Figure 3.4).

Referring to [23], we know that a *binary symmetric channel*, denoted BSC_δ where $0 < \delta < 1$ is a pre-specified parameter known as the *crossover probability*, takes as its input a bit and flips it with probability δ (and keeps it unchanged with probability $1-\delta$). For the case where $p = q = \delta$, we see that $F_{p,q}$ can be represented as

$$F_{\delta,\delta}\eta(n) = \mathrm{BSC}_\delta\left(\mathrm{NAND}\left(\eta(n), \eta(n+1), \eta(n+2)\right)\right), \tag{3.2.4}$$

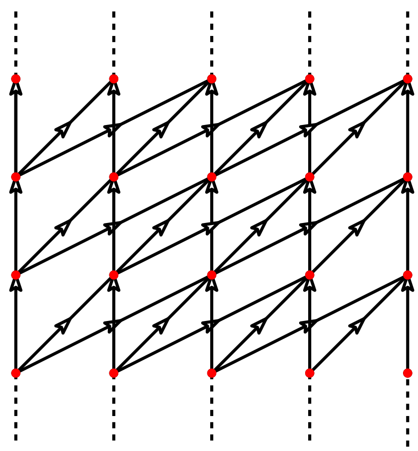


Figure 3.3: Lattice with (x,y) having successors $(x,y+1)$, $(x+1,y+1)$ and $(x+2,y+1)$

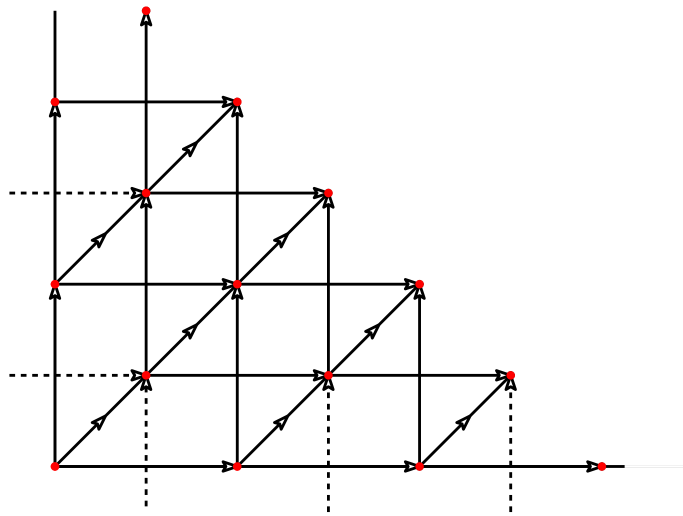


Figure 3.4: Lattice with (x,y) having successors $(x+2,y)$, $(x+1,y+1)$ and $(x,y+2)$

where the NAND function outputs the value 1 when all three inputs are 0, and outputs the value 0 when at least one of the three inputs equals 1. As outlined in § 3.2 of [23], the problem of *broadcasting of information on* 2*-dimensional grids* showcases another usefulness of such PCAs. Let us consider the problem of broadcasting information on the two 2-dimensional infinite directed grid shown in Figure 3.5. The origin O has one bit of information, and we are interested in broadcasting this bit to the entire infinite directed lattice. Let us associate with each vertex of the grid the index pair (t,i), with $t \geqslant 0$ and $0 \leqslant i \leqslant 2t$. At any time $t > 0$, each vertex with index (t,i), for $2 \leqslant i \leqslant 2t-2$, receives a bit from each of the vertices indexed $(t-1,i-2)$, $(t-1,i-1)$ and $(t-1,i)$, applies the NAND function to these bits, and then implements BSC_δ. The vertex indexed $(t,1)$ receives bits from the vertices $(t-1,0)$ and $(t-1,1)$, the vertex indexed $(t,2t-1)$ receives bits from the vertices $(t-1,2t-3)$ and $(t-1,2t-2)$, the vertex indexed $(t,0)$ receives a bit from the vertex $(t-1,0)$, and the vertex indexed $(t,2t)$ receives a bit from the vertex $(t-1,2t-2)$. From (3.2.4), it is evident that if we view the coordinate t as time, understanding this broadcasting problem is similar to studying the PCA $F_{\delta,\delta}$, except that broadcasting on such a grid involves bounded-length configurations and its behaviour is different at the boundary.

As a third motivation for studying the PCA $F_{p,q}$, we point out, as discussed right after (3.2.2), that $F_{p,q}$ can be obtained by perturbing a deterministic CA by a memoryless zero-range noise (once again, we refer the reader to [74]). The importance of noisy CAs, as a subclass of PCAs, lies in their use in studying the reliability of computations against the presence of noise. The study of low-noise PCAs presents us with the same kind of challenges that we face when studying low-

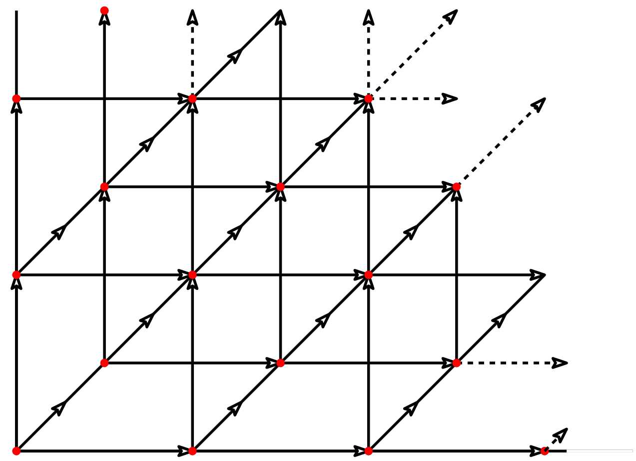


Figure 3.5: Broadcasting of information on this grid resembles the PCA $F_{p,q}$ with $p = q = \delta$

temperature models of statistical mechanics. In particular, studying the ergodicity, or lack thereof, of such PCAs, is intimately tied to the phenomenon of phase transition at low temperatures. This, along with the previous discussions, serve to motivate our investigation of the PCA $F_{p,q}$.

3.2.3 Our envelope PCA and its relation to our percolation game

We begin by deducing certain recurrence relations that arise naturally in the percolation game described in §3.2.1. Once $\mathbb{Z}^2$ has been endowed with the trap / target / open labeling, we define a site (x,y) to be in the class W if the game that begins with (x,y) as the initial vertex is won by the player who plays the first round. We define (x,y) to be in the class L if the game that begins with (x,y) as the initial vertex is lost by the player who plays the first round, and we define (x,y) to be in the class D if the game that begins with (x,y) as the initial vertex results in a draw. In particular, if (x,y) is a trap, then we place it in W, and if it is a target, we place it in L. The intuition behind these conventions is as follows: one may imagine an "unseen" round that takes place *before* the actual game begins, in which the player who is supposed to play the second round of the actual game moves the token from somewhere else to (x,y). Thus, if (x,y) is a trap (respectively a target), she loses (respectively wins) even before the game begins, implying that the player who plays the first round of the actual game wins (respectively loses).

For every $k \in \mathbb{Z}$, we denote by $H_k = \{(x,k) : x \in \mathbb{Z}\}$ the horizontal line containing all sites whose y-coordinate equals k. These lines have been illustrated in red in Figure 3.6. From the moves permitted in our game, it follows that for any $k \in \mathbb{Z}$, if all sites (x,y) that lie on the horizontal line H_{k+1} have already been categorized into the classes W, L and D, then this information, along with the pre-assigned labels of trap / target / open to the sites on H_k, is enough to determine the classes

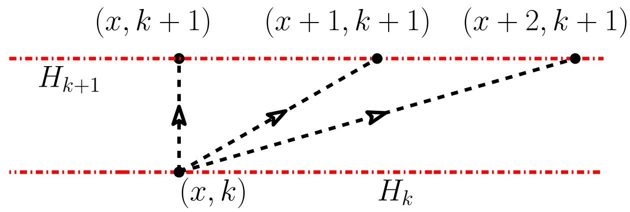


Figure 3.6: Illustrating deduction of the recurrence relations

to which the sites lying on H_k belong.

Using Figure 3.6 as a reference, we draw the following conclusions, assuming that (x, y) is our initial vertex:

1. If each vertex of $\text{Out}(x, y)$ belongs to W, then no matter which of these vertices the first player moves the token to from (x, y), the second player wins. We thus have the following two possibilities:

   (a) either the vertex (x, y) has been marked a trap and hence belongs to W, which happens with probability p,

   (b) or else the game that begins from (x, y) results in a loss for the first player, so that (x, y) is classified into L with the remaining probability $1 - p$.

2. If at least one of the vertices of $\text{Out}(x, y)$ belongs to L, the first player moves the token from (x, y) to *this* vertex, making the second player lose. We thus have the following two possibilities:

   (a) either (x, y) has been marked a target and hence belongs to L, which happens with probability q,

   (b) or else the game that begins from (x, y) results in a win for the first player, so that (x, y) is classified into W with the remaining probability $1 - q$.

3. The final scenario is where *none* of the vertices of $\text{Out}(x, y)$ belongs to L but at least one of them belongs to D. In this case, we have the following three possibilities:

   (a) either (x, y) has been marked a trap and hence belongs to W, which happens with probability p,

   (b) or (x, y) has been marked a target and hence belongs to L, which happens with probability q,

(c) or else the game that begins from (x,y) results in a draw (since the first player moves the token from (x,y) to the vertex in $\mathrm{Out}(x,y)$ which is in D), thus placing (x,y) in the class D with the remaining probability $r = 1-p-q$.

We further note that, conditioned on the classification of the vertices that lie on H_{k+1} into the classes W, L and D, the (random) class which a vertex lying on H_k gets sorted into via the above-mentioned rules is independent of all other vertices on H_k.

The above recurrence relations are the key to establishing a connection between our game and the envelope PCA $\widehat{\mathrm{F}}_{p,q}$ that we are now ready to define. Let us label each vertex with the set it belongs to. That is, we label a vertex W if it belongs to the set W and so on. For any $k \in \mathbb{Z}$, we identify H_k with $\mathbb{Z}$ by identifying (x,k) on H_k with x on $\mathbb{Z}$ for each $x \in \mathbb{Z}$. This allows us to represent the recurrence relations listed above via a PCA $\widehat{\mathrm{F}}_{p,q}$ that is endowed with the alphabet $\hat{\mathscr{A}} = \{W,L,D\}$, the neighborhood $\mathscr{N} = \{0,1,2\}$, and the stochastic matrix $\widehat{\varphi}_{p,q} : \hat{\mathscr{A}}^3 \times \hat{\mathscr{A}} \to [0,1]$ defined via the equations:

$$\widehat{\varphi}_{p,q}(W,W,W;b) = \begin{cases} p & \text{if } b = W, \\ 1-p & \text{if } b = L, \end{cases} \tag{3.2.5}$$

$$\widehat{\varphi}_{p,q}(a_0,a_1,a_2;b) = \begin{cases} 1-q & \text{if } b = W, \\ q & \text{if } b = L, \end{cases} \quad \text{for all } (a_0,a_1,a_2) \in \hat{\mathscr{A}}^3 \setminus \{W,D\}^3 \tag{3.2.6}$$

and

$$\widehat{\varphi}_{p,q}(a_0,a_1,a_2;b) = \begin{cases} p & \text{if } b = W, \\ q & \text{if } b = L, \\ r = 1-p-q & \text{if } b = D, \end{cases} \quad \text{for all } (a_0,a_1,a_2) \in \{W,D\}^3 \setminus (W,W,W). \tag{3.2.7}$$

We illustrate $\widehat{\varphi}_{p,q}$ in Figure 3.7. To clarify further, the classification of the vertices on H_{k+1} into W, L and D yields a configuration $\eta \in \hat{\mathscr{A}}^{\mathbb{Z}}$ via the identifications described above, and the classification of the vertices on H_k into W, L and D via the game's recurrence relations can then be represented by $\widehat{\mathrm{F}}_{p,q}\,\eta$.

If we endow $\hat{\mathscr{A}}$ with the total order $W \prec D \prec L$ and define $L - D = D$, then (similar to the way we interpret $F_{p,q}$ in §3.2.2) the envelope PCA $\widehat{\mathrm{F}}_{p,q}$ can be derived, via random perturbations, from a deterministic CA, $\widehat{\mathrm{F}}$, and a random noise $\widehat{\theta}$, as follows. The local update rule $\widehat{\mathrm{f}}$ of $\widehat{\mathrm{F}}$ is given by $\widehat{\mathrm{f}}(a_0,a_1,a_2) = 1 - \max\{a_0,a_1,a_2\}$, whereas the noise $\widehat{\theta}(a,b)$, for $a,b \in \hat{\mathscr{A}}$, transforms W to L with

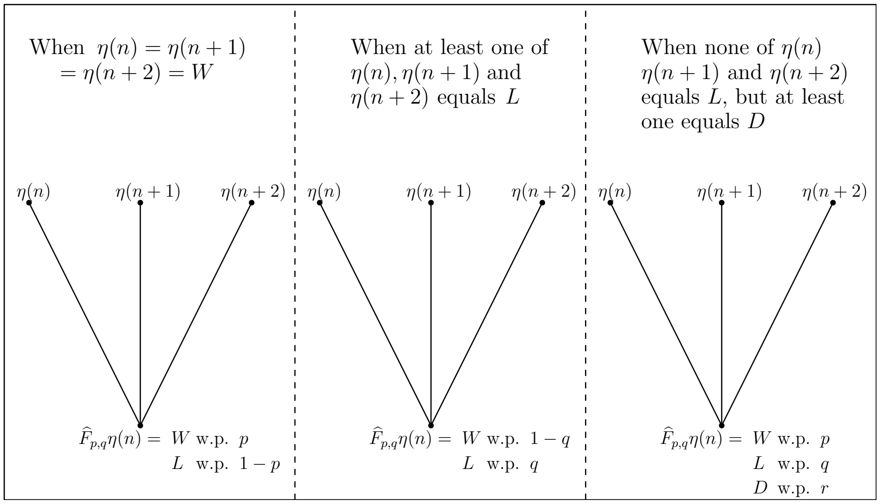


Figure 3.7: The stochastic rules that define our envelope PCA $\widehat{\mathrm{F}}_{p,q}$

probability q (i.e. $\widehat{\theta}(W,L) = q$) and keeps it unchanged with probability $1-q$, transforms L to W with probability p (i.e. $\widehat{\theta}(L,W) = p$) and keeps it unchanged with probability $1-p$, and transforms D to L with probability q (i.e. $\widehat{\theta}(D,L) = q$), D to W with probability p (i.e. $\widehat{\theta}(D,W) = p$), and keeps D unchanged with probability $r = 1-p-q$. It is now straightforward to check that $\widehat{\mathrm{F}}_{p,q}$ is obtained from $\widehat{\mathrm{F}}$ by injecting into it the memoryless zero-range noise $\widehat{\theta}$ (again, we refer the reader to [74] for a detailed reading on PCAs obtained via perturbations of deterministic CAs by random noises).

Although we have argued in §3.2.3 how the recurrence relations deduced from our percolation game give rise to $\widehat{\mathrm{F}}_{p,q}$, there is, in fact, another angle from which one may motivate the emergence of $\widehat{\mathrm{F}}_{p,q}$ (we mention here that the term "envelope" was introduced in [28]) – namely, the use of $\widehat{\mathrm{F}}_{p,q}$ in coupling two (random) configurations obtained via (possibly repeated) applications of the PCA $F_{p,q}$, starting from two different initial configurations. The symbol D is utilized to populate sites whose actual values may differ between these two coupled configurations.

We dwell here a little longer on the notion of envelope PCAs. In [28], a random algorithm known as a *perfect sampling procedure* is proposed which, when repeatedly applied, can estimate the (unique) stationary probability distribution corresponding to an ergodic PCA F with arbitrary precision. When the universe of F is finite (e.g. when $\mathbb{Z}/n\mathbb{Z}$ constitute the cells for F), a new PCA $\widehat{\mathrm{F}}$ on an extended alphabet, referred to as an *envelope PCA to F*, is introduced, which is then run on a single initial configuration to obtain a perfect sampling procedure for F. When the universe for F is infinite (e.g. when $\mathbb{Z}^d$ constitute the cells for F, making the universe uncountably infinite), [28] shows that an efficient perfect sampling procedure can be obtained when $\widehat{\mathrm{F}}$ is ergodic.

When the alphabet for F is $\mathscr{A} = \{W,L\}$, the alphabet for $\widehat{\mathrm{F}}$ is $\hat{\mathscr{A}} = \{W,L,D\}$. The universe E

and neighborhood $\mathscr{N}$ for $\widehat{\mathrm{F}}$ remain the same as those for F. Given a configuration $\eta = (\eta(x) : x \in E)$ in $\hat{\mathscr{A}}^E$, one thinks of η as a configuration in $\mathscr{A}^E$ as follows: if $x \in E$ is a cell such that it is not known which symbol from $\mathscr{A}$ occupies it, we imagine that it is occupied by the symbol D. It is to be noted that when $\eta \in \mathscr{A}^E$, i.e. η is a configuration which is devoid of the symbol D, the envelope PCA $\widehat{\mathrm{F}}$ acts on it in exactly the same way as F acts on it.

It is now time to connect the principal components of this work, namely the game, the PCA $F_{p,q}$, and its envelope $\widehat{\mathrm{F}}_{p,q}$, with one another. This connection is established via the following three results.

Proposition 3.2.1. *For every $(p,q) \in \mathscr{S}$, the percolation game with parameters p and q has probability 0 of culminating in a draw if and only if the PCA $F_{p,q}$ is ergodic.*

Proposition 3.2.2. *The PCA $F_{p,q}$ is ergodic if and only if the corresponding envelope PCA $\widehat{\mathrm{F}}_{p,q}$ is ergodic.*

Proposition 3.2.3. *For each $(p,q) \in \mathscr{S}$, the envelope PCA $\widehat{\mathrm{F}}_{p,q}$ admits no stationary distribution μ that assigns positive probability to the symbol D. To put it formally, the probability of the event $\eta(x) = D$, where η is a random configuration with law μ, is 0 for every $x \in \mathbb{Z}$.*

We implement the first assertion of Lemma 3.3.1 (that concerns itself with a couple of results regarding stochastic domination, as stated in §3.3 below) and the argument used in proving Proposition 2.1 of [60] to establish Proposition 3.2.2. Next, an argument identical to that used in proving Proposition 2.2 of [60] yields a proof of Proposition 3.2.1. We are now left with the task of proving Proposition 3.2.3.

3.3 An important lemma before we embark on a proof of Proposition 3.2.3

Recall from §3.2.3 that $\hat{\mathscr{A}} = \{W, D, L\}$, and borrowing from the definitions in §3.2.2, we let $\mathbb{D}$ denote the set of all probability measures on $\Omega = \hat{\mathscr{A}}^{\mathbb{Z}}$ that are defined with respect to the σ-field $\mathscr{F}$ generated by the cylinder sets of Ω.

Lemma 3.3.1. *Let μ and $\widetilde{\mu}$ be two probability distributions in $\mathbb{D}$. Let $\preceq$ denote the stochastic domination (on $\mathbb{D}$) with respect to the coordinate-wise total order induced by the ordering $W \prec D \prec L$, and let $\mu \preceq \widetilde{\mu}$. Then $\widehat{\mathrm{F}}_{p,q}\widetilde{\mu} \preceq \widehat{\mathrm{F}}_{p,q}\mu$.*

Let μ and $\widetilde{\mu}$ be two probability distributions in $\mathbb{D}$. Let $\trianglelefteq$ denote the stochastic domination (on $\mathbb{D}$) with respect to the coordinate-wise partial order induced by $W \triangleleft D \triangleright L$, and let $\mu \trianglelefteq \widetilde{\mu}$. Then $\widehat{\mathrm{F}}_{p,q}\,\mu \trianglelefteq \widehat{\mathrm{F}}_{p,q}\,\widetilde{\mu}$.

Proof. We begin with (deterministic) configurations η and $\widetilde{\eta}$ in $\hat{\mathscr{A}}^{\mathbb{Z}}$ such that $\eta(x) \preceq \widetilde{\eta}(x)$ for all $x \in \mathbb{Z}$. We consider two copies of $\mathbb{Z}^2$, in both of which we fix the *same* assignment, say ω, of trap / target / open labels to the vertices lying on H_k. In the first copy of $\mathbb{Z}^2$, we assign the state $\eta(x)$ to the site $(x,k+1)$ (that belongs to H_{k+1}), for each $x \in \mathbb{Z}$, whereas in the second copy of $\mathbb{Z}^2$, we assign the state $\widetilde{\eta}(x)$ to the site $(x,k+1)$, for each $x \in \mathbb{Z}$.

The state, say $\tau(x)$, assigned to the site (x,k) (that belongs to H_k), for each $x \in \mathbb{Z}$, in the first copy of $\mathbb{Z}^2$, is deduced using the states $\eta(x)$, $\eta(x+1)$ and $\eta(x+2)$, and the label $\omega(x,k)$, according to the following rules:

1. if $\omega(x,k)$ reads trap, then $\tau(x) = W$,

2. if $\omega(x,k)$ reads target, then $\tau(x) = L$,

3. if $\omega(x,k)$ reads open, and

   (a) $\eta(x) = \eta(x+1) = \eta(x+2) = W$, then $\tau(x) = L$,

   (b) at least one of $\eta(x)$, $\eta(x+1)$ and $\eta(x+2)$ equals L, then $\tau(x) = W$,

   (c) none of $\eta(x)$, $\eta(x+1)$ and $\eta(x+2)$ equals L, but at least one of them equals D, then $\tau(x) = D$.

The state, say $\widetilde{\tau}(x)$, assigned to (x,k), for each $x \in \mathbb{Z}$, in the second copy of $\mathbb{Z}^2$, is deduced using the states $\widetilde{\eta}(x)$, $\widetilde{\eta}(x+1)$ and $\widetilde{\eta}(x+2)$, and the label $\omega(x,k)$, following rules analogous to the above.

We now compare $\tau(x)$ and $\widetilde{\tau}(x)$ for each $x \in \mathbb{Z}$. Since the same assignment, ω, of trap / target / open labels to the vertices of H_k is used in each copy of $\mathbb{Z}^2$, we need only carry out this comparison when $\omega(x,k)$ is open. It suffices to consider $x = 0$.

1. Suppose $\widetilde{\eta}(0) = \widetilde{\eta}(1) = \widetilde{\eta}(2) = W$, which forces $\eta(0) = \eta(1) = \eta(2) = W$. In this case, $\tau(0) = \widetilde{\tau}(0) = L$.

2. Suppose $(\widetilde{\eta}(0), \widetilde{\eta}(1), \widetilde{\eta}(2)) \in \{W,D\}^3 \setminus \{(W,W,W)\}$. In this case, we either have $(\eta(0), \eta(1), \eta(2)) \in \{W,D\}^3 \setminus \{(W,W,W)\}$ or we have $(\eta(0), \eta(1), \eta(2)) = (W,W,W)$. In the former situation, $\tau(0) = \widetilde{\tau}(0) = D$, whereas in the latter situation, $\widetilde{\tau}(0) = D$ and $\tau(0) = L$.

3. Finally, suppose $(\widetilde{\eta}(0),\widetilde{\eta}(1),\widetilde{\eta}(2)) \in \{W,D,L\}^3 \setminus \{W,D\}^3$. The first possibility is that $(\eta(0),\eta(1),\eta(2)) \in \{W,D,L\}^3 \setminus \{W,D\}^3$, in which case $\tau(0) = \widetilde{\tau}(0) = W$. The second possibility is where $(\eta(0),\eta(1),\eta(2)) \in \{W,D\}^3 \setminus \{(W,W,W)\}$, in which case, $\widetilde{\tau}(0) = W$ whereas $\tau(0) = D$. The third and final possibility is where $(\eta(0),\eta(1),\eta(2)) = (W,W,W)$, in which case $\widetilde{\tau}(0) = W$ and $\tau(0) = L$.

These observations let us conclude that $\tau \succeq \widetilde{\tau}$. Given $\mu \preceq \widetilde{\mu}$, we let η and $\widetilde{\eta}$ denote coupled random configurations (defined on the same sample space) such that η follows μ, $\widetilde{\eta}$ follows $\widetilde{\mu}$ and $\eta \preceq \widetilde{\eta}$ almost surely. The above deduction applied to η and $\widetilde{\eta}$ completes the proof of the first assertion of Lemma 3.3.1.

To prove the second assertion of Lemma 3.3.1, we consider a very similar set-up as above, with the only distinction being that $\eta(x) \trianglelefteq \widetilde{\eta}(x)$ for all $x \in \mathbb{Z}$. Once again, we compare $\tau(x)$ and $\widetilde{\tau}(x)$ for each $x \in \mathbb{Z}$, and it suffices to consider $x = 0$ and $\omega(0,k)$ open:

1. Suppose $\eta(0) = \eta(1) = \eta(2) = W$ and $(\widetilde{\eta}(0),\widetilde{\eta}(1),\widetilde{\eta}(2)) \in \{W,D\}^3 \setminus \{(W,W,W)\}$. Then $\tau(0) = L$ whereas $\widetilde{\tau}(0) = D$.

2. When $(\eta(0),\eta(1),\eta(2)) \in \{W,D,L\}^3 \setminus \{W,D\}^3$, we either have $(\widetilde{\eta}(0),\widetilde{\eta}(1),\widetilde{\eta}(2)) \in \{W,D\}^3 \setminus \{(W,W,W)\}$ or $(\widetilde{\eta}(0),\widetilde{\eta}(1),\widetilde{\eta}(2)) \in \{W,D,L\}^3 \setminus \{W,D\}^3$. In the latter scenario, $\tau(0) = \widetilde{\tau}(0) = W$, while in the former, we have $\tau(0) = W$ and $\widetilde{\tau}(0) = D$.

These observations imply $\tau \trianglelefteq \widetilde{\tau}$, which, in turn, yields the conclusion in the second part of Lemma 3.3.1. □

3.4 The method of weight functions and the proof of Proposition 3.2.3

As discussed in §3.2.2, it is a non-trivial task to establish, rigorously, the ergodicity of PCAs in general. It has also been explicitly stated in [60] that coming up with a suitable weight function or potential function that serves our purpose of proving ergodicity is not an easy feat either.

Before we proceed further, we recall here the definitions of translation-invariant and reflection-invariant probability measures as discussed right after Definition 1.1.1. Given any 1-dimensional PCA with alphabet $\mathscr{A}$, a finite index set $S = \{y_1, y_2, \ldots, y_n\} \subset \mathbb{Z}$, and symbols a_1, a_2, $\ldots$, a_n that belong to $\mathscr{A}$, we call $(a_1 a_2 \ldots a_n)_S = \{\eta \in \mathscr{A}^{\mathbb{Z}} : \eta(y_i) = a_i \text{ for all } 1 \leqslant i \leqslant n\}$ a *cylinder set indexed by S*. When a probability measure μ is translation-invariant, and $y_i = k + i$ for all $1 \leqslant i \leqslant n$ and for some $k \in \mathbb{Z}$, we denote the measure $\mu((a_1 a_2 \ldots a_n)_S)$ of the cylinder set $(a_1 a_2 \ldots a_n)_S$ by simply

$\mu(a_1a_2\dots a_n)$, since the 'location' of S (i.e. the value of k) ceases to be relevant. For instance, if η is a random configuration following the law μ, then for *any* $x\in\mathbb{Z}$, we can let $\mu(D)$ indicate the probability of the event $\eta(x)=D$. Likewise, for *any* $x\in\mathbb{Z}$, we can let $\mu(WD)$ indicate the probability of the event that $\eta(x)=W$ and $\eta(x+1)=D$, and so on. When μ is both translation-invariant and reflection-invariant, $\mu(a_1a_2\dots a_n)=\mu(a_na_{n-1}\dots a_1)$ for all $a_1,\dots,a_n\in\mathscr{A}$. For instance, $\mu(LWD)=\mu(DWL)$, since the former is the probability of the event $\eta(-1)=L,\eta(0)=W,\eta(1)=D$ whereas the latter is the probability of the event $\eta(-1)=D,\eta(0)=W,\eta(1)=L$.

Lemma 3.4.1. *To prove Proposition 3.2.3, it suffices to show that under no translation-invariant and reflection-invariant stationary distribution for $\widehat{\mathrm{F}}_{p,q}$ can the symbol D appear anywhere with positive probability, for each $(p,q)\in\mathscr{S}$ (where $\mathscr{S}$ is as defined in (3.1.1)).*

Proof. Suppose there exists a stationary distribution μ for $\widehat{\mathrm{F}}_{p,q}$ with $\mu(D)>0$. To prove Lemma 3.4.1, it now suffices to show the existence of a translation-invariant and reflection-invariant stationary distribution for $\widehat{\mathrm{F}}_{p,q}$ under which the probability of occurrence of the symbol D is strictly positive.

Let δ_D indicate the configuration that assigns the state D to each $x\in\mathbb{Z}$. By the second assertion of Lemma 3.3.1, the sequence $\{\widehat{\mathrm{F}}^n_{p,q}\,\delta_D(D)\}_{n\in\mathbb{N}_0}$ is non-increasing, and $\widehat{\mathrm{F}}^n_{p,q}\,\delta_D(D)\geqslant\mu(D)>0$ for each $n\in\mathbb{N}_0$. Thus, any limit point of the Césaro sums of the sequence $\{\widehat{\mathrm{F}}^n_{p,q}\,\delta_D\}_{n\in\mathbb{N}_0}$ is a translation-invariant and reflection-invariant stationary distribution for $\widehat{\mathrm{F}}_{p,q}$ that assigns a strictly positive probability to the occurrence of D (because of our assumption that $\mu(D)>0$), yielding the desired conclusion. $\square$

In the rest of this chapter, we consider only μ that belongs to $\mathbb{D}$ (recall from §3.3) and that is both translation-invariant and reflection-invariant.

We now come to the actual construction of the weight function, which is accomplished via several steps. To begin with, for the sake of brevity, we let $\widehat{**}$ denote the set $\{W,D\}^2\setminus\{(W,W)\}$ and $\widehat{***}$ the set $\{W,D\}^3\setminus\{(W,W,W)\}$. We may think of $(\widehat{**})$ as representing the cylinder set in which $(\eta(W),\eta(L))\in\widehat{**}$, and $(\widehat{***})$ as representing the cylinder set in which $(\eta(0),\eta(1),\eta(2))\in\widehat{***}$. We write $(S_0S_1\dots S_k\widehat{***}S'_0S'_1\dots S'_{k'})$ (likewise, $(S_1S_2\dots S_k\widehat{**}S'_1S'_2\dots S'_{k'})$) to indicate the cylinder set in which $\eta(i)\in S_i$ for all $0\leqslant i\leqslant k$, $(\eta(k+1),\eta(k+2),\eta(k+3))\in\widehat{***}$, and $\eta(k+4+i)\in S'_i$ for all $0\leqslant i\leqslant k'$, for any subsets $S_0,\dots,S_k,S'_0,\dots,S'_{k'}$ of $\{W,D,L\}$. When $S_i=\{a_i\}$ is a singleton for some $0\leqslant i\leqslant k$, we replace S_i in the above notation by simply a_i (and likewise when S'_i is a singleton for some $0\leqslant i\leqslant k'$).

Keeping the reader's convenience in mind and before we plunge into the intricate technicalities of the weight function derivation, we briefly dwell here on how we plan to accomplish the task at hand, i.e. proving Proposition 3.2.3, using the method of weight functions. Following the notations introduced above, we call a cylinder set $(a_1, a_2, \dots, a_n)_S$ *D-inclusive* if $a_i = D$ for at least one $i \in \{1, \dots, n\}$. We envisage our weight function $w(\mu)$ to be a linear combination

$$w(\mu) = \sum_{i=1}^{s} c_i \mu(\mathscr{C}_i) \tag{3.4.1}$$

of cylinder sets $\mathscr{C}_1, \dots, \mathscr{C}_s$, each of which is D-inclusive, with $c_1, \dots, c_s$ being real constants (that are functions of the parameters p and q), and we want it to satisfy an inequality of the form

$$w(\widehat{\mathrm{F}}_{p,q}\mu) \leqslant w(\mu) - \sum_{i=1}^{s'} c'_i \mu(\mathscr{C}'_i), \tag{3.4.2}$$

where $c'_1, \dots, c'_{s'}$ are non-negative real constants (and are, once again, functions of p and q) and $\mathscr{C}'_1, \dots, \mathscr{C}'_{s'}$ are, once again, D-inclusive cylinder sets. When μ is stationary for $\widehat{\mathrm{F}}_{p,q}$, we have $w(\widehat{\mathrm{F}}_{p,q}\mu) = w(\mu)$, which then yields $\sum_{i=1}^{s'} c'_i \mu(\mathscr{C}'_i) = 0$. The constants $c'_1, \dots, c'_{s'}$ are such that, when $p + q > 0$, we can find a *non-empty* $P \subset \{1, \dots, s'\}$ such that $c'_i > 0$ whenever $i \in P$. This, in turn, yields $\mu(\mathscr{C}'_i) = 0$ for every $i \in P$ whenever μ is stationary and $p + q > 0$. From these, we infer, due to the good choice of the constants c_i made in the weight function in (3.4.1), that $\mu(D) = 0$.

We outline here a summary of how we proceed in the rest of §3.4. Keeping in mind the goal of showing $\mu(D) = 0$ for every probability distribution μ that is stationary for the envelope PCA $\widehat{\mathrm{F}}_{p,q}$, we attempt to

1. begin with an initial, reasonable guess for the weight function,

2. parallelly, come up with a suitable inequality (or equality) that is satisfied by this initial expression for the weight function,

3. take note of the terms on the right side of this inequality (or equality) that are non-negative and therefore need to be dealt with,

4. take note of the terms on the right side of this inequality (or equality) that are non-positive and have some power in negating the above-mentioned non-negative terms,

5. accordingly, introduce an adjustment into the initial weight function,

6. write down, possibly with *several* simplifications, how the above-mentioned adjustment affects the weight function inequality (or equality), and see if the updated weight function

inequality (or equality) has the desired form as stated in (3.4.2),

7. declare the current weight function as the final, desired weight function if the answer to the question in (3) is a yes,

8. otherwise, repeat the above procedure from (3) onward.

We conclude this introductory part of §3.4 by stating the final expression for the weight function that we deduce in the sequel:

$$\begin{aligned}w_3(\mu) = (1-p^2-pq-q)\mu(D)+2\mu(WD)-\mu(DWD)+2r(1-p^2)\mu(LWWD)-2pr\{\mu(LD)+\mu(LWD)\}\\ -2p^2r\{\mu(LDD)+\mu(LDWD)+\mu(LWDD)\}-4r\mu(LDWL)-2p^2r\{\mu(LDWW)+\mu(LWDW)\}.\end{aligned} \tag{3.4.3}$$

3.4.1 Computation of the probabilities of various cylinder sets under the pushforward measure induced by the action of $\widehat{\mathrm{F}}_{p,q}$

Recall from §3.2.2 the definition of the pushforward measure $\widehat{\mathrm{F}}_{p,q}\mu$. Ultimately, we would need to compute $\widehat{\mathrm{F}}_{p,q}\mu(\mathscr{C})$ for every cylinder set $\mathscr{C}$ that shows up in the expression for the final weight function in (3.4.3). However, in §3.4.1, we compute only the first few (i.e. the pushforward measures for (D), (WD), (DWD), $(LWWD)$) as these are the only cylinder sets that show up in the expression for our initial weight function (see (3.4.10)). The rest are computed as and when required (during the subsequent steps in which the weight function is appropriately tweaked to satisfy the desired criterion).

For *any* translation-invariant and reflection-invariant probability measure μ belonging to $\mathbb{D}$ (recall from §3.3), using (3.2.7), we have

$$\widehat{\mathrm{F}}_{p,q}\mu(D) = r\mu(\widehat{***}). \tag{3.4.4}$$

In the deduction of $\widehat{\mathrm{F}}_{p,q}\mu(\mathscr{C})$ where $\mathscr{C}$ is one of WD, DWD and $LWWD$, we make use (3.2.5), (3.2.6) and (3.2.7). Since these computations involve rather similar arguments, we explain in detail only two of them.

To compute $\widehat{\mathrm{F}}_{p,q}\mu(WD)$, we note that for $(\widehat{\mathrm{F}}_{p,q}\eta(W), \widehat{\mathrm{F}}_{p,q}\eta(L))$ to equal (WD) for some $\eta \in \hat{\mathscr{A}}^{\mathbb{Z}}$, we require $(\eta(1),\eta(2),\eta(3)) \in \widehat{***}$, and the event $\widehat{\mathrm{F}}_{p,q}\eta(1) = D$ occurs with probability r. If $\eta(0) = \eta(1) = \eta(2) = W$, forcing $\eta(3) = D$, the event $\widehat{\mathrm{F}}_{p,q}\eta(0) = W$ happens with probability p; if $\eta(0) = W$ and $(\eta(1),\eta(2),\eta(3)) \in \widehat{***} \setminus \{(WWD)\}$, then again the event $\widehat{\mathrm{F}}_{p,q}\eta(0) = W$ happens with probability p as $(\eta(0),\eta(1),\eta(2)) \in \widehat{***}$ in this case. If $\eta(0) = D$, the event $\widehat{\mathrm{F}}_{p,q}\eta(0) = W$ happens with probability p, and if $\eta(0) = L$, the event $\widehat{\mathrm{F}}_{p,q}\eta(0) = W$ happens with probability $(1-q)$. Combining all, we have

$$\widehat{\mathrm{F}}_{p,q}\mu(WD) = pr\mu(\{W,D\}\widehat{***}) + (1-q)r\mu(L\widehat{**}). \tag{3.4.5}$$

Arguing likewise, we obtain

$$\begin{aligned}\widehat{\mathrm{F}}_{p,q}\mu(LWWD) = (1-p)p^2r\mu(WWW\widehat{***}) + qp^2r[\mu(\widehat{***}\widehat{***}) + \mu(L\{W,D\}^2\widehat{***})] \\ + q(1-q)pr\mu(L\{W,D\}\widehat{***}) + q(1-q)^2r\mu(L\widehat{**}).\end{aligned} \tag{3.4.6}$$

We outline the argument to deduce $\widehat{\mathrm{F}}_{p,q}\mu(DWD)$. For $(\widehat{\mathrm{F}}_{p,q}\eta(0), \widehat{\mathrm{F}}_{p,q}\eta(1), \widehat{\mathrm{F}}_{p,q}\eta(2))$ to equal (DWD) for some $\eta \in \hat{\mathscr{A}}^{\mathbb{Z}}$, we require $(\eta(0), \eta(1), \eta(2)) \in \widehat{***}$ and $(\eta(2), \eta(3), \eta(4)) \in \widehat{***}$. Therefore, $(\eta(0), \eta(1)) \in \{W,D\}^2$, but we must avoid $\eta(0) = \eta(1) = \eta(2) = W$. Thus $(\eta(0), \eta(1), \eta(2), \eta(3), \eta(4))$ must belong to the set $\{W,D\}^2\widehat{***} \setminus WWW\widehat{**}$. Note that this forces $(\eta(1), \eta(2), \eta(3)) \in \{W,D\}^3$, so that $\widehat{\mathrm{F}}_{p,q}\eta(1) = W$ happens with probability p. Each of the events $\widehat{\mathrm{F}}_{p,q}\eta(0) = D$ and $\widehat{\mathrm{F}}_{p,q}\eta(2) = D$ occurs with probability r. Combining all, we have

$$\widehat{\mathrm{F}}_{p,q}\mu(DWD) = pr^2[\mu(\{W,D\}^2\widehat{***}) - \mu(WWW\widehat{**})]. \tag{3.4.7}$$

3.4.2 Important identities used in the derivation of the weight function

We use Tables 3.1, 3.2, 3.3 and 3.4 to illustrate the derivation of a few identities. In each of these tables, the first row indicates the indices of the coordinates in $\mathbb{Z}$, and the rows that follow represent events involving cylinder sets. To elucidate, in Table 3.1, the second and third rows represent respectively the events $(\eta(-1), \eta(0), \eta(1), \eta(2)) = (L,D,D,D)$ and $(\eta(-1), \eta(0), \eta(1), \eta(2)) = (L,D,D,W)$, and it is immediate that these two events are disjoint since they disagree on the symbol that occupies the coordinate 2. In fact, for any two distinct rows that are inside the same table, it can be seen that the corresponding events are mutually exclusive. To give the reader an understanding of how we make use of these tables, note that the union of the pairwise disjoint events listed in Table 3.1 forms a subset of the event $\eta(0) = D$. Thus, Table 3.1, along with the previously stated assumption that μ is translation-invariant and reflection-invariant (justified by Lemma 3.4.1), allows us to write

$$\begin{aligned}\mu(D) = \mu(LDDD) + \mu(LDDW) + \mu(LDWD) + \mu(LDWW) + \mu(DDD) + \mu(DDW) + \mu(WWD) \\ + \mu(DWD) + \mu(DDDL) + \mu(WDDL) + \mu(LWDD) + \mu(LWDW) + \mu(LDDL) + \mu(LDWL)\end{aligned}$$

-2	-1	0	1	2
	L	D	D	D
	L	D	D	W
	L	D	W	D
	L	D	W	W
	D	D	D	
	D	D	W	
W	W	D		
D	W	D		
D	D	D	L	
W	D	D	L	
L	W	D	D	
L	W	D	W	
	L	D	D	L
	L	D	W	L
	L	D	L	
L	D	D	L	
L	W	D	L	

Table 3.1: Decomposition of D to establish (3.4.8)

-1	0	1	2
D	D	D	D
D	D	D	W
D	D	W	D
D	D	W	W
W	D	D	D
W	D	D	W
W	D	W	D
W	D	W	W
L	D	D	D
L	D	D	W
L	D	D	L
L	D	W	D
L	D	W	W
L	D	W	L
	D	L	
D	D	D	L
W	D	D	L
D	D	W	L
W	D	W	L

Table 3.2: Decomposition of D to establish (3.4.9)

-1	0	1	2
D	W	D	D
D	W	D	W
W	W	D	D
W	W	D	W
L	W	D	D
L	W	D	W
D	W	D	L
W	W	D	L
L	W	D	L

Table 3.3: Decomposition of WD to establish (3.4.9)

-1	0	1	2
D	W	W	D
W	W	W	D
L	W	W	D

Table 3.4: Decomposition of WWD to establish (3.4.9)

$$
\begin{aligned}
&+\mu(LDL)+\mu(LDDL)+\mu(LWDL)\\
=\;&\underbrace{\mu(LDDD)+\mu(LDDW)+\mu(LDWD)+\mu(LDWW)+\mu(LWDD)+\mu(LWDW)}_{(1)}\\
&+\underbrace{\mu(DDDL)+\mu(WDDL)}_{(2)}+\underbrace{\mu(DDD)+\mu(DDW)+\mu(WWD)+\mu(DWD)}_{(3)}+2\mu(LDDL)\\
&+\underbrace{\mu(LDWL)+\mu(LWDL)}_{(4)}+\mu(LDL)\\
=\;&\underbrace{\mu(L\widehat{***})-\mu(LWWD)}_{\text{rewriting (1)}}\\
&+\underbrace{\mu(\widehat{***}L)-\mu(DWWL)-\mu(DDWL)-\mu(WDWL)-\mu(DWDL)-\mu(WWDL)}_{\text{rewriting (2)}}\\
&+\underbrace{\mu(\widehat{***})-\mu(WDD)-\mu(WDW)-\mu(DWW)}_{\text{rewriting (3)}}+2\mu(LDDL)+\underbrace{2\mu(LDWL)}_{\text{reflection-invariance on (4)}}+\mu(LDL)\\
=\;&\mu(\widehat{***})-\mu(WDD)-\mu(WDW)-\mu(DWW)+\underbrace{\mu(L\widehat{***})+\mu(\widehat{***}L)-\big[\mu(LWWD)+\mu(DWWL)\big]}_{\text{from the rewritten (1) and (2) in the previous step}}\\
&\underbrace{-\big[\mu(DDWL)+\mu(WDWL)\big]-\big[\mu(DWDL)+\mu(WWDL)\big]}_{\text{from the rewritten (2) in the previous step}}+2\mu(LDDL)+\mu(LDL)+2\mu(LDWL)\\
=\;&\mu(\widehat{***})-\mu(WDD)-\mu(WDW)-\mu(DWW)+\underbrace{2\mu(L\widehat{***})-2\mu(LWWD)}_{\text{reflection-invariance}}-\big[\mu(DDWL)+\mu(WDWL)\big]\\
&-\big[\mu(DWDL)+\mu(WWDL)\big]+2\mu(LDDL)+\mu(LDL)+2\mu(LDWL)\\
=\;&\mu(\widehat{***})-\mu(WDD)-\mu(WDW)-\mu(DWW)+2\mu(L\widehat{***})-2\mu(LWWD)-\big[\mu(DWL)-\mu(LDWL)\big]\\
&-\big[\mu(WDL)-\mu(LWDL)\big]+2\mu(LDDL)+\mu(LDL)+2\mu(LDWL)\\
=\;&\mu(\widehat{***})-\big[\mu(WDD)+\mu(WDW)+\mu(WDL)\big]-\big[\mu(DWW)+\mu(DWL)\big]+2\mu(L\widehat{***})-2\mu(LWWD)\\
&+2\mu(LDDL)+\mu(LDL)+4\mu(LDWL)\quad\text{(grouping the terms judiciously)}\\
=\;&\mu(\widehat{***})-\mu(WD)-\big[\mu(DW)-\mu(DWD)\big]+2\mu(L\widehat{***})-2\mu(LWWD)+2\mu(LDDL)+\mu(LDL)\\
&+4\mu(LDWL)\\
=\;&\mu(\widehat{***})-2\mu(WD)+\mu(DWD)-2\mu(LWWD)+2\mu(L\widehat{***})+2\mu(LDDL)+\mu(LDL)+4\mu(LDWL).
\end{aligned}
\tag{3.4.8}
$$

Likewise, Tables 3.2, 3.3 and 3.4 together yield

$$\mu(D)+\mu(WD)+\mu(WWD)$$

$$
\begin{aligned}
&= \mu(DDDD)+\mu(DDDW)+\mu(DDWD)+\mu(DDWW)+\mu(WDDD)+\mu(WDDW)+\mu(WDWD)\\
&\quad +\mu(WDWW)+\mu(LDDD)+\mu(LDDW)+\mu(LDDL)+\mu(LDWD)+\mu(LDWW)+\mu(LDWL)\\
&\quad +\mu(DL)+\mu(DDDL)+\mu(WDDL)+\mu(DDWL)+\mu(WDWL)+\mu(DWDD)+\mu(DWDW)\\
&\quad +\mu(WWDD)+\mu(WWDW)+\mu(LWDD)+\mu(LWDW)+\mu(DWDL)+\mu(WWDL)+\mu(LWDL)\\
&\quad +\mu(DWWD)+\mu(WWWD)+\mu(LWWD)\\
&= \big[\mu(DDDD)+\mu(DDDW)+\mu(DDWD)+\mu(DDWW)+\mu(DWDD)+\mu(DWDW)+\mu(DWWD)\big]\\
&\quad +\big[\mu(WDDD)+\mu(WDDW)+\mu(WDWD)+\mu(WDWW)+\mu(WWDD)+\mu(WWDW)+\mu(WWWD)\big]\\
&\quad +\big[\mu(LDDD)+\mu(LDDW)+\mu(LDWD)+\mu(LDWW)+\mu(LWDD)+\mu(LWDW)+\mu(LWWD)\big]\\
&\quad +\big[\mu(DDDL)+\mu(WDDL)+\mu(DDWL)+\mu(WDWL)+\mu(DWDL)+\mu(WWDL)\big]\\
&\quad +\mu(LDDL)+\mu(LDWL)+\mu(DL)+\mu(LWDL) \quad \text{(grouping the terms judiciously)}\\
&= \mu(D\widehat{***})+\mu(W\widehat{***})+\mu(L\widehat{***})+\mu(\widehat{***}L)-\mu(DWWL)+\mu(LDDL)+\mu(LD)+2\mu(LDWL)\\
&= \mu(\{W,D\}\widehat{***})+2\mu(L\widehat{***})-\mu(LWWD)+\mu(LD)+2\mu(LDWL)+\mu(LDDL). \qquad (3.4.9)
\end{aligned}
$$

3.4.3 The first step of composing the weight function

We start by defining

$$
w_0(\mu)=\mu(D)+2\mu(WD)-\mu(DWD)+2\mu(LWWD). \qquad (3.4.10)
$$

It is not straightforward to explain our intuition behind setting w_0 in (3.4.10) as our 'initial' choice of weight function (after which it gets tweaked and adjusted in several steps described in the sequel to yield the final weight function). Since we are ultimately interested in $\mu(D)$ when μ is stationary, and since each $\mathscr{C}_i$ in (3.4.1) is D-inclusive, it is not too far-fetched to entertain the possibility of starting with $\mathscr{C}_1=(D)_{\{0\}}$ (i.e. the cylinder set in which D occupies the origin). Thus, the right side of (3.4.2) contains $\mu(D)$, while the left contains $\widehat{\mathrm{F}}_{p,q}\mu(D)=r\mu(\widehat{***})$. When p and q are both small (intuitively, our task of showing $\mu(D)=0$ for μ stationary ought to become harder the smaller $p+q$ gets), $r\mu(\widehat{***})$ is nearly equal to $\mu(\widehat{***})$. The appearance of $2\mu(WD)$ in the right side of (3.4.2) implies, from (3.4.5), that $2(1-q)r\mu(L\widehat{***})$ appears in the left side of (3.4.2), and when p and q are both small, this is nearly the same as $2\mu(L\widehat{***})$. From (3.4.8), we see that $\mu(D)+2\mu(WD)-\mu(DWD)+2\mu(LWWD)$ serves as an upper bound for $\mu(\widehat{***})+2\mu(L\widehat{***})$. All these provide ample justification as to why we start with w_0 in (3.4.10) as our initial weight function.

Substituting the expressions from (3.4.4), (3.4.5), (3.4.7) and (3.4.6) into (3.4.10), we get

$$
\begin{aligned}
w_0(\widehat{\mathrm{F}}_{p,q}\mu) &= \widehat{\mathrm{F}}_{p,q}\mu(D) + 2\widehat{\mathrm{F}}_{p,q}\mu(WD) - \widehat{\mathrm{F}}_{p,q}\mu(DWD) + 2\widehat{\mathrm{F}}_{p,q}\mu(LWWD) \\
&= r\mu(\widehat{***}) + 2\{pr\mu(\{W,D\}\widehat{***}) + (1-q)r\mu(L\widehat{***})\} - pr^2[\mu(\{W,D\}^2\widehat{***}) \\
&\quad - \mu(WWW\widehat{**})] + 2\big\{(1-p)p^2r\mu(WWW\widehat{***}) + qp^2r[\mu(\widehat{***}\widehat{***}) + \mu(L\{W,D\}^2\widehat{***})] \\
&\quad + q(1-q)pr\mu(L\{W,D\}\widehat{***}) + q(1-q)^2r\mu(L\widehat{***})\big\} \\
&= r\mu(\widehat{***}) + 2\{pr\mu(\{W,D\}\widehat{***}) + (1-q)r\mu(L\widehat{***})\} - pr^2[\mu(\{W,D\}^2\widehat{***}) \\
&\quad - \mu(WWW\widehat{**})] + 2\big\{(1-p)p^2r\mu(WWW\widehat{***}) + qp^2r[\underbrace{\mu(\{W,D\}^3\widehat{***}) - \mu(WWW\widehat{***})}_{\text{splitting } \mu(\widehat{******})} \\
&\quad + \mu(L\{W,D\}^2\widehat{***})] + q(1-q)pr\mu(L\{W,D\}\widehat{***}) + q(1-q)^2r\mu(L\widehat{***})\big\} \\
&= r\mu(\widehat{***}) + 2\{pr\mu(\{W,D\}\widehat{***}) + (1-q)r\mu(L\widehat{***})\} - pr^2[\mu(\{W,D\}^2\widehat{***}) \\
&\quad - \mu(WWW\widehat{**})] + 2\big\{(1-p)p^2r\mu(WWW\widehat{***}) + qp^2r\big[\underbrace{\mu(\{W,D\}^2\widehat{***})}_{\text{adding } \mu(\{W,D\}^3\widehat{***}) \text{ and } \mu(L\{W,D\}^2\widehat{***})} \\
&\quad - \mu(WWW\widehat{***})\big] + q(1-q)pr\mu(L\{W,D\}\widehat{***}) + q(1-q)^2r\mu(L\widehat{***})\big\} \\
&= r\mu(\widehat{***}) + 2\{pr\mu(\{W,D\}\widehat{***}) + (1-q)r\mu(L\widehat{***})\} - pr^2[\mu(\{W,D\}^2\widehat{***}) \\
&\quad - \mu(WWW\widehat{**})] + 2\big\{(1-p)p^2r\mu(WWW\widehat{***}) + qp^2r\big[\underbrace{\mu(\{W,D\}\widehat{***}) - \mu(L\{W,D\}\widehat{***})}_{\text{splitting } \mu(\{W,D\}^2\widehat{***})} \\
&\quad - \mu(WWW\widehat{***})\big] + q(1-q)pr\mu(L\{W,D\}\widehat{***}) + q(1-q)^2r\mu(L\widehat{***})\big\} \\
&= r\mu(\widehat{***}) + 2\{pr\mu(\{W,D\}\widehat{***}) + (1-q)r\mu(L\widehat{***})\} - pr^2[\mu(\{W,D\}^2\widehat{***}) \\
&\quad - \mu(WWW\widehat{**})] + 2\big\{(1-p)p^2r\mu(WWW\widehat{***}) + qp^2r\big[\underbrace{\mu(\widehat{***}) - \mu(L\widehat{***})}_{\text{splitting } \mu(\{W,D\}\widehat{***})} \\
&\quad - \mu(L\{W,D\}\widehat{***}) - \mu(WWW\widehat{***})\big] + q(1-q)pr\mu(L\{W,D\}\widehat{***}) + q(1-q)^2r\mu(L\widehat{***})\big\} \\
&= \underbrace{(r+2qp^2r)\mu(\widehat{***})}_{\text{combining the terms involving } \mu(\widehat{***})} + pr\mu(\{W,D\}\widehat{***}) \\
&\quad + \left\{ pr\mu(\{W,D\}\widehat{***}) \underbrace{+2q(1-q)pr\mu(L\{W,D\}\widehat{***}) - 2qp^2r\mu(L\{W,D\}\widehat{***})}_{\text{combining the terms involving } \mu(L\{W,D\}\widehat{***})} \right\} \\
&\quad - pr^2[\mu(\{W,D\}^2\widehat{***}) - \mu(WWW\widehat{**})] + \underbrace{\big\{2(1-p)p^2r\mu(WWW\widehat{***}) - 2qp^2r\mu(WWW\widehat{***})\big\}}_{\text{combining terms involving } \mu(WWW\widehat{***})} \\
&\quad \underbrace{\big\{2q(1-q)^2r\mu(L\widehat{***}) - 2qp^2r\mu(L\widehat{***}) + 2(1-q)r\mu(L\widehat{***})\big\}}_{\text{combining the terms involving } \mu(L\widehat{***})}
\end{aligned}
$$

$$
\begin{aligned}
&= (r+2qp^2r)\mu(\widehat{***}) + pr\mu(\{W,D\}\widehat{***}) + \{\underbrace{pr\mu(\{W,D\}^2\widehat{***}) + pr\mu(L\{W,D\}\widehat{***})}_{\text{splitting } \mu(\{W,D\}\widehat{***})} \\
&\quad + 2qpr(1-q-p)\mu(L\{W,D\}\widehat{***})\} - pr^2\mu(\{W,D\}^2\widehat{***}) + pr^2\mu(WWW\widehat{**}) \\
&\quad + 2p^2r(1-p-q)\mu(WWW\widehat{***}) + 2r\{q(1-q)^2 - qp^2 + (1-q)\}\mu(L\widehat{**}*) \\
&= (r+2qp^2r)\mu(\widehat{***}) + pr\mu(\{W,D\}\widehat{***}) + (pr+2qpr^2)\mu(L\{W,D\}\widehat{***}) \\
&\quad + \underbrace{pr\mu(\{W,D\}^2\widehat{***}) - pr^2\mu(\{W,D\}^2\widehat{***})}_{\text{combining terms involving } \mu(\{W,D\}^2\widehat{***})} + pr^2\mu(WWW\widehat{**}) + 2p^2r^2\mu(WWW\widehat{***}) \\
&\quad + 2r\left\{q+q^3-2q^2-qp^2+1-q\right\}\mu(L\widehat{**}*) \\
&= (r+2qp^2r)\mu(\widehat{***}) + pr\mu(\{W,D\}\widehat{***}) + (pr+2qpr^2)\mu(L\{W,D\}\widehat{***}) \\
&\quad + pr(p+q)\mu(\{W,D\}^2\widehat{***}) + pr^2\mu(WWW\widehat{**}) + 2p^2r^2\mu(WWW\widehat{***}) \\
&\quad + 2r\left\{1+q^3-2q^2-qp^2\right\}\mu(L\widehat{**}*).
\end{aligned} \tag{3.4.11}
$$

Next, applying

1. the identity

$$
\begin{aligned}
\mu(\widehat{***}) + 2\mu(L\widehat{**}*) = \mu(D) + 2\mu(WD) - \mu(DWD) + 2\mu(LWWD) - 2\mu(LDDL) \\
- \mu(LDL) - 4\mu(LDWL),
\end{aligned} \tag{3.4.12}
$$

   obtained by rearranging the terms of (3.4.8),

2. and the identity

$$
\begin{aligned}
\mu(\{W,D\}\widehat{***}) = \mu(D) + \mu(WD) + \mu(WWD) - 2\mu(L\widehat{**}*) + \mu(LWWD) - \mu(LD) \\
- 2\mu(LDWL) - \mu(LDDL),
\end{aligned} \tag{3.4.13}
$$

   obtained by rearranging the terms of (3.4.9), we transform (3.4.11) as follows:

$$
\begin{aligned}
w_0(\widehat{\mathrm{F}}_{p,q}\mu) &= (r+2qp^2r)\mu(\widehat{***}) + 2r\{1+q^3-2q^2-qp^2\}\mu(L\widehat{**}*) + pr\mu(\{W,D\}\widehat{***}) \\
&\quad + (pr+2qpr^2)\mu(L\{W,D\}\widehat{***}) + pr(p+q)\mu(\{W,D\}^2\widehat{***}) + pr^2\mu(WWW\widehat{**}) + 2p^2r^2\mu(WWW\widehat{***}) \\
&= \underbrace{\{r\mu(\widehat{***}) + 2r\mu(L\widehat{**}*)\}}_{\text{to be bound using (3.4.12)}} + 2qp^2r\mu(\widehat{***}) + 2r\{q^3-2q^2-qp^2\}\mu(L\widehat{**}*) \underbrace{+pr\mu(\{W,D\}\widehat{***})}_{\text{to be bound using (3.4.13)}} \\
&\quad + (pr+2qpr^2)\mu(L\{W,D\}\widehat{***}) + pr(p+q)\mu(\{W,D\}^2\widehat{***}) + pr^2\mu(WWW\widehat{**}) + 2p^2r^2\mu(WWW\widehat{***}) \\
&= \underbrace{r\{\mu(D) + 2\mu(WD) - \mu(DWD) + 2\mu(LWWD) - 2\mu(LDDL) - \mu(LDL) - 4\mu(LDWL)\}}_{\text{(1) – bound obtained from (3.4.12)}} + 2qp^2r\mu(\widehat{***})
\end{aligned}
$$

$$+2r\{q^3-2q^2-qp^2)\}\mu(L\widehat{***})+\underbrace{pr\{\mu(D)+\mu(WD)+\mu(WWD)-2\mu(L\widehat{***})+\mu(LWWD)-\mu(LD)\}}_{\text{(2) – bound obtained from (3.4.13), continued next line within underbrace}}$$

$$\underbrace{-2\mu(LDWL)-\mu(LDDL)\}}_{\text{(2), continued from above}}+(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})+pr(p+q)\mu(\{W,D\}^2\widehat{***})+pr^2\mu(WWW\widehat{**})$$

$$+2p^2r^2\mu(WWW\widehat{***})$$

$$=\underbrace{(r+pr)\mu(D)+(2r+pr)\mu(WD)-r\mu(DWD)+pr\mu(WWD)+(2r+pr)\mu(LWWD)-(2r+pr)\mu(LDDL)}_{\text{combining terms from the bounds (1) and (2)}}$$

$$\underbrace{-r\mu(LDL)-(4r+2pr)\mu(LDWL)-pr\mu(LD)}_{\text{combining terms from the bounds (1) and (2)}}+2r\{q^3-2q^2-qp^2-p\}\mu(L\widehat{***})+2qp^2r\mu(\widehat{***})$$

$$+(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})+pr(p+q)\mu(\{W,D\}^2\widehat{***})+pr^2\mu(WWW\widehat{**})+2p^2r^2\mu(WWW\widehat{***})$$

$$=\big[\underbrace{\mu(D)}+(r+pr-1)\mu(D)\big]+\big[\underbrace{2\mu(WD)}+(2r+pr-2)\mu(WD)\big]+\big[\underbrace{-\mu(DWD)}-(r-1)\mu(DWD)\big]$$

$$+\big[\underbrace{2\mu(LWWD)}+(2r+pr-2)\mu(LWWD)\big]+pr\mu(WWD)-(2r+pr)\mu(LDDL)-r\mu(LDL)-$$

$$(4r+2pr)\mu(LDWL)-pr\mu(LD)+2r\{q^3-2q^2-qp^2-p\}\mu(L\widehat{***})+2qp^2r\mu(\widehat{***})+$$

$$(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})+pr(p+q)\mu(\{W,D\}^2\widehat{***})+pr^2\mu(WWW\widehat{**})+2p^2r^2\mu(WWW\widehat{***})$$

$$=\underbrace{w_0(\mu)}_{(3)}-[p(1-r)+q]\mu(D)-[p(1-r)+q]\mu(WD)+\underbrace{(r-1)\mu(WD)-(r-1)\mu(DWD)}_{(4)}+pr\mu(WWD)$$

$$+(pr-2p-2q)\mu(LWWD)-(2r+pr)\mu(LDDL)-r\mu(LDL)-(4r+2pr)\mu(LDWL)-pr\mu(LD)$$

$$+2r\{q^3-2q^2-qp^2-p\}\mu(L\widehat{***})+2qp^2r\mu(\widehat{***})+(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})$$

$$+pr(p+q)\mu(\{W,D\}^2\widehat{***})+pr^2\mu(WWW\widehat{**})+2p^2r^2\mu(WWW\widehat{***})$$

((3) obtained by combining underbraced terms from the previous step, & using (3.4.10))

$$=w_0(\mu)-[p(1-r)+q]\mu(D)-[p(1-r)+q]\mu(WD)+\underbrace{(r-1)\mu(LWD)+(r-1)\mu(WWD)}_{\text{rewriting (4)}}+pr\mu(WWD)$$

$$-[p(2-r)+2q]\mu(LWWD)-(2r+pr)\mu(LDDL)-r\mu(LDL)-(4r+2pr)\mu(LDWL)-pr\mu(LD)$$

$$+2r\{q^3-2q^2-qp^2-p\}\mu(L\widehat{***})+2qp^2r\mu(\widehat{***})+(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})$$

$$+pr(p+q)\mu(\{W,D\}^2\widehat{***})+pr^2\mu(WWW\widehat{**})+2p^2r^2\mu(WWW\widehat{***})$$

$$=w_0(\mu)\underbrace{-[p(1-r)+q]\mu(D)-[p(1-r)+q]\mu(WD)}_{(5)}-(p+q)\mu(LWD)\underbrace{-[p(1-r)+q]\mu(WWD)}_{(6)}$$

$$-[p(2-r)+2q]\mu(LWWD)-(2r+pr)\mu(LDDL)-r\mu(LDL)-(4r+2pr)\mu(LDWL)-pr\mu(LD)$$

$$+2r\{q^3-2q^2-qp^2-p\}\mu(L\widehat{***})+2qp^2r\mu(\widehat{***})+(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})$$

$$+pr(p+q)\mu(\{W,D\}^2\widehat{***})+pr^2\mu(WWW\widehat{**})+2p^2r^2\mu(WWW\widehat{***})$$

$$= w_0(\mu) \underbrace{-[p(1-r)+q][\mu(D)+\mu(WD)+\mu(WWD)]}_{\text{combining (5) and (6)}} -(p+q)\mu(LWD) - [p(2-r)+2q]\mu(LWWD)$$

$$- (2r+pr)\mu(LDDL) - r\mu(LDL) - (4r+2pr)\mu(LDWL) - pr\mu(LD) + 2qp^2r\mu(\widehat{***})$$
$$- 2r(p+2q^2+qp^2-q^3)\mu(L\widehat{***}) + (pr+2qpr^2)\mu(L\{W,D\}\widehat{***}) + pr(p+q)\mu(\{W,D\}^2\widehat{***})$$
$$+ pr^2\mu(WWW\widehat{**}) + 2p^2r^2\mu(WWW\widehat{***}). \tag{3.4.14}$$

One may note here how (3.4.14) lays down the first of the stepping stones which pave the way towards an inequality that resembles (3.4.2). We remark here that our ultimate goal would be to transform the equality in (3.4.14) into an inequality of the form (3.4.2). However, there are, at this point, several terms in the right side of (3.4.14) (such as $2qp^2r\mu(\widehat{***})$, $(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})$ etc.), other than $w_0(\mu)$, in which the coefficients are non-negative, and this needs to be remedied. This is what we accomplish, via several adjustments and suitable algebraic manipulations in-between, in §3.4.4, §3.4.5, §3.4.6 and §3.4.7.

In each of the steps involving the above-mentioned adjustments, the equality/ inequality (beginning from (3.4.14)) evolves. Suppose we have performed i adjustments so far, and the weight function currently under consideration is denoted by w_i. Let the inequality that we currently have (which could actually be an equality, in which case we refer to it as the current *weight function equality*) be given by

$$w_i(\widehat{\mathrm{F}}_{p,q}\mu) \leqslant w_i(\mu) - \sum_{j=1}^{s_i} \alpha_{i,j}\mu(\mathscr{E}_{i,j}), \tag{3.4.15}$$

where $\alpha_{i,1},\ldots,\alpha_{i,s_i}$ are real constants and $\mathscr{E}_{i,1},\ldots,\mathscr{E}_{i,s_i}$ are cylinder sets. Suppose the $(i+1)$-st adjustment is defined via the equation $w_{i+1}(\mu) = w_i(\mu) - \sum_{j=1}^{t_i}\beta_{i,j}\mu(\mathscr{G}_{i,j})$, in which $\beta_{i,1},\ldots,\beta_{i,t_i}$ are real constants and $\mathscr{G}_{i,1},\ldots,\mathscr{G}_{i,t_i}$ are cylinder sets. We can now rewrite (3.4.15) as follows:

$$w_{i+1}(\widehat{\mathrm{F}}_{p,q}\mu) + \sum_{j=1}^{t_i}\beta_{i,j}\widehat{\mathrm{F}}_{p,q}\mu(\mathscr{G}_{i,j}) \leqslant w_{i+1}(\mu) + \sum_{j=1}^{t_i}\beta_{i,j}\mu(\mathscr{G}_{i,j}) - \sum_{j=1}^{s_i}\alpha_{i,j}\mu(\mathscr{E}_{i,j})$$
$$\implies w_{i+1}(\widehat{\mathrm{F}}_{p,q}\mu) \leqslant w_{i+1}(\mu) + \sum_{j=1}^{t_i}\beta_{i,j}\mu(\mathscr{G}_{i,j}) - \sum_{j=1}^{t_i}\beta_{i,j}\widehat{\mathrm{F}}_{p,q}\mu(\mathscr{G}_{i,j}) - \sum_{j=1}^{s_i}\alpha_{i,j}\mu(\mathscr{E}_{i,j}). \tag{3.4.16}$$

This is the commonality shared by all the adjustments described in the sequel, and we refer back, several times, to (3.4.16) and use it to see how the inequality (or equality) evolves with each adjustment.

3.4.4 The second step of composing the weight function

As mentioned right after (3.4.14), a quick comparison of (3.4.14) with (3.4.2) reveals that the terms $2qp^2r\mu(\widehat{***})$, $(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})$, $pr(p+q)\mu(\{W,D\}^2\widehat{***})$, $pr^2\mu(WWW\widehat{**})$ and $2p^2r^2\mu(WWW\widehat{***})$ in the right side of (3.4.14) are all non-negative. To make sure that the final inequality is of the form given in (3.4.2), we have to use the existing non-positive terms in the right side of (3.4.14) to nullify these non-negative terms.

With the aim to tackle first the terms $(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})$ and $pr(p+q)\mu(\{W,D\}^2\widehat{***})$, and then the terms $pr^2\mu(WWW\widehat{**})$ and $2p^2r^2\mu(WWW\widehat{***})$, we consider the following first adjustment to the initial weight function in (3.4.10):

$$w_1(\mu) = w_0(\mu) - p(p+q)\mu(D) = w_0(\mu) - p(1-r)\mu(D). \tag{3.4.17}$$

Rewriting (3.4.14) as $w_0(\widehat{\mathrm{F}}_{p,q}\mu) = w_0(\mu) + A_0$, and using the idea presented in (3.4.16), we see that the adjustment of (3.4.17) would transform (3.4.14) into

$$\begin{aligned} w_1(\widehat{\mathrm{F}}_{p,q}\mu) &= w_1(\mu) + p(1-r)\mu(D) - p(1-r)\widehat{\mathrm{F}}_{p,q}\mu(D) + A_0 \\ &= w_1(\mu) + p(1-r)\mu(D) - p(p+q)\widehat{\mathrm{F}}_{p,q}\mu(D) + A_0. \end{aligned} \tag{3.4.18}$$

We include a brief discussion attempting to explain why we choose to subtract $p(1-r)\mu(D)$ in (3.4.17). At the very outset, we let the reader know that we do not claim here that this is the *only* way to begin the sequence of adjustments. Some other term out of the existing non-positive terms in the right side of (3.4.14) might very well have worked. However, a significant amount of work went into trying out adjustments using terms that seemed plausible enough to negate the above-mentioned non-negative terms, using the idea presented in (3.4.16), and the choice in (3.4.17) has the following advantages:

1. As seen from the much neater (3.4.18) and by recalling the coefficient of $\mu(D)$ in the right side of (3.4.14), the coefficient of $\mu(D)$ becomes $-q$ once (3.4.17) has been implemented, which is still non-positive and therefore creates no problems in the subsequent steps. This has been one of our main considerations in each adjustment step: that the adjustments are chosen such that they do not upset existing non-positive terms on the right side of the weight function inequality (or equality) and turn them non-negative.

2. We now come to the term $-p(p+q)\widehat{\mathrm{F}}_{p,q}\mu(D) = -p(p+q)r\mu(\widehat{***})$ in (3.4.18). The simple inequality $\mu(\{W,D\}^2\widehat{***}) \leqslant \mu(\widehat{***})$ tells us that $-p(p+q)r\mu(\widehat{***})$ would suffice to fully negate the non-negative term $pr(p+q)\mu(\{W,D\}^2\widehat{***})$.

3. Note that after accomplishing the task of negating $p(p+q)\mu(\{W,D\}^2\widehat{***})$, the term $-p(p+q)r\mu(\widehat{***})$ is also able to help in partially negating $(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})$ (we say "partially" because the order of magnitude of the coefficient $pr+2qpr^2$ is p when p and q are small, and this is larger than the order of magnitude of $p(p+q)r$), since $\mu(L\{W,D\}\widehat{***})+\mu(\{W,D\}^2\widehat{***})=\mu(\{W,D\}\widehat{***})\leqslant\mu(\widehat{***})$.

We hope that the above discussions provide some insight to the reader as to why we found $-p(1-r)\mu(D)$ a compelling choice in (3.4.17).

Presenting the reader with the final weight function, with all the adjustments combined together, may come across as abrupt and opaque. On the other hand, outlining the adjustments in the same chronological order in which we ourselves thought of them while developing the weight function implies that the power of each adjustment (in negating the existing non-negative terms in the right side of the weight function inequality / equality) is limited by our ability to carry out intricate computations and algebraic manipulations and to keep track of the non-negative terms that still remain. Therefore, it is sometimes the case that we carry out two different adjustments at two different stages of development of the weight function, but some of the terms used in both these adjustments are the same (up to the cylinder sets involved, but with different coefficients). For instance, in the final adjustment, i.e. (3.4.49), the term $-q\mu(D)$ appears, and we may very well have combined this with $-p(1-r)\mu(D)$ of (3.4.17), but in doing so, we would have been unable to point out effectively to the reader how the construction truly unfolds.

Applying the adjustment described in (3.4.17) to (3.4.14), as illustrated in (3.4.18), we obtain:

$$\begin{aligned}
w_1(\widehat{\mathrm{F}}_{p,q}\mu) &= w_1(\mu)+p(1-r)\mu(D)-p(1-r)\widehat{\mathrm{F}}_{p,q}\mu(D)-[p(1-r)+q][\mu(D)+\mu(WD)+\mu(WWD)]\\
&\quad -(p+q)\mu(LWD)-[p(2-r)+2q]\mu(LWWD)-(2r+pr)\mu(LDDL)-r\mu(LDL)\\
&\quad -(4r+2pr)\mu(LDWL)-pr\mu(LD)+2qp^2r\mu(\widehat{***})-2r(p+2q^2+qp^2-q^3)\mu(L\widehat{***})\\
&\quad +(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})+pr(p+q)\mu(\{W,D\}^2\widehat{***})+pr^2\mu(WWW\widehat{**})\\
&\quad +2p^2r^2\mu(WWW\widehat{***})\\
&= w_1(\mu)-q\mu(D)-[p(1-r)+q][\mu(WD)+\mu(WWD)]-(p+q)\mu(LWD)\\
&\quad -[p(2-r)+2q]\mu(LWWD)-(2r+pr)\mu(LDDL)-r\mu(LDL)-(4r+2pr)\mu(LDWL)\\
&\quad -pr\mu(LD)+2qp^2r\mu(\widehat{***})\underbrace{+pr^2\mu(WWW\widehat{**})+2p^2r^2\mu(WWW\widehat{***})}_{(1)}\\
&\quad \underbrace{-p(p+q)\widehat{\mathrm{F}}_{p,q}\mu(D)-2r(p+2q^2+qp^2-q^3)\mu(L\widehat{***})}_{(2)}
\end{aligned}$$

$$\underbrace{+(pr+2qpr^2)\mu(L\{W,D\}\widehat{***})+pr(p+q)\mu(\{W,D\}^2\widehat{***})}_{(3)}. \tag{3.4.19}$$

Combining the terms of (3.4.19) that have been highlighted by the underbraces numbered (2) and (3), using (3.4.4), and using the identities

1. $\mu(\widehat{***}) = \mu(L\widehat{***}) + \mu(L\{W,D\}\widehat{***}) + \mu(\{W,D\}^2\widehat{***})$,
2. $\mu(L\{W,D\}\widehat{***}) = \mu(L\widehat{***\{W,D\}}) + \mu(LWWWD) - \mu(LDWWW) = \mu(L\widehat{***}) - \mu(L\widehat{***}L) + \mu(LWWWD) - \mu(LDWWW)$,

we obtain

$$\begin{aligned}
&-p(p+q)\widehat{\mathrm{F}}_{p,q}\,\mu(D) - 2r(p+2q^2+qp^2-q^3)\mu(L\widehat{***}) + (pr+2qpr^2)\mu(L\{W,D\}\widehat{***})\\
&+pr(p+q)\mu(\{W,D\}^2\widehat{***})\\
=\;&\underbrace{-p(p+q)r\mu(\widehat{***})}_{\text{use (1)}}-2r(p+2q^2+qp^2-q^3)\mu(L\widehat{***}) + (pr+2qpr^2)\mu(L\{W,D\}\widehat{***})\\
&+pr(p+q)\mu(\{W,D\}^2\widehat{***})\\
=\;&\underbrace{-pr(p+q)\mu(L\widehat{***}) - pr(p+q)\mu(L\{W,D\}\widehat{***}) - pr(p+q)\mu(\{W,D\}^2\widehat{***})}_{\text{after applying (1)}}\\
&-2r(p+2q^2+qp^2-q^3)\mu(L\widehat{***}) + (pr+2qpr^2)\mu(L\{W,D\}\widehat{***}) + pr(p+q)\mu(\{W,D\}^2\widehat{***})\\
=\;&\underbrace{-pr(p+q)\mu(L\widehat{***}) - 2r(p+2q^2+qp^2-q^3)\mu(L\widehat{***})}_{\text{combining the terms involving } \mu(L\widehat{***})}\\
&\underbrace{-pr(p+q)\mu(L\{W,D\}\widehat{***}) + (pr+2qpr^2)\mu(L\{W,D\}\widehat{***})}_{\text{combining the terms involving } \mu(L\{W,D\}\widehat{***})}\\
&\text{(also, cancelling out the terms involving } \mu(\{W,D\}^2\widehat{***}) \text{ by each other)}\\
=\;&-r[p(p+q)+2(p+2q^2+qp^2-q^3)]\mu(L\widehat{***})\underbrace{+(pr^2+2qpr^2)\mu(L\{W,D\}\widehat{***})}_{\text{use (2)}}\\
=\;&-r[p(p+q)+2(p+2q^2+qp^2-q^3)]\mu(L\widehat{***})\\
&\underbrace{+(pr^2+2qpr^2)[\mu(L\widehat{***}) - \mu(L\widehat{***}L) + \mu(LWWWD) - \mu(LDWWW)]}_{\text{after applying (2)}}\\
=\;&-r[p(p+q)+2(p+2q^2+qp^2-q^3)]\mu(L\widehat{***}) + (pr^2+2qpr^2)[\mu(L\widehat{***}) - \mu(L\widehat{***}L)]\\
&+(pr^2+2qpr^2)\mu(LWWWD) - (pr^2+2qpr^2)\mu(LDWWW)\\
=\;&\underbrace{-r(p+4q^2+4qp^2-2q^3+2p^2+2q^2p)\mu(L\widehat{***}) - (pr^2+2qpr^2)\mu(L\widehat{***}L)}_{(4)\text{(we combined the terms involving } \mu(L\widehat{***}))}
\end{aligned}$$

$$\underbrace{+(pr^2+2qpr^2)\mu(LWWWD)-(pr^2+2qpr^2)\mu(LDWWW)}_{(5)}. \tag{3.4.20}$$

Next, we combine the terms of (3.4.20) grouped together by the underbrace numbered (5), with the terms of (3.4.19) highlighted by the underbrace numbered (1), and use the identities

1. $\mu(LWWWD) = \mu(WWWD) - \mu(DWWWD) - \mu(WWWWD)$,
2. $\mu(WWW\widehat{***}) = \mu(WWWD) + \mu(WWWWD) + \mu(WWWWWD) - \mu(WWW\widehat{**}L) - \mu(WWWDL)$,
3. $\mu(WWW\widehat{**}) = \mu(WWWDD) + \mu(WWWDW) + \mu(WWWWD)$,
4. $\mu(WWWDD) + \mu(WWWDW) = \mu(WWWD) - \mu(WWWDL)$,

to obtain:

$$
\begin{aligned}
&\underbrace{(pr^2+2qpr^2)\mu(LWWWD)}_{\text{use (1)}}-(pr^2+2qpr^2)\mu(LDWWW)+pr^2\mu(WWW\widehat{**})\underbrace{+2p^2r^2\mu(WWW\widehat{***})}_{\text{use (2)}}\\
=\ &\underbrace{(pr^2+2qpr^2)[\mu(WWWD)-\mu(DWWWD)-\mu(WWWWD)]}_{\text{after applying (1)}}\\
&-(pr^2+2qpr^2)\mu(LDWWW)+pr^2\mu(WWW\widehat{**})\\
&\underbrace{+2p^2r^2[\mu(WWWD)+\mu(WWWWD)+\mu(WWWWWD)-\mu(WWW\widehat{**}L)-\mu(WWWDL)]}_{\text{after applying (2)}}\\
=\ &\underbrace{(pr^2+2qpr^2)\mu(WWWD)}_{\text{(a)}}-(pr^2+2qpr^2)\mu(WWWWD)-(pr^2+2qpr^2)\mu(DWWWD)\\
&\underbrace{-(pr^2+2qpr^2)\mu(LDWWW)}_{\text{(c)}}+\underbrace{(pr^2+2qpr^2)\mu(WWW\widehat{**})-2qpr^2\mu(WWW\widehat{**})}_{\text{manipulating the coefficient of } \mu(WWW\widehat{**})}\underbrace{+2p^2r^2\mu(WWWD)}_{\text{(b)}}\\
&+2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)]-2p^2r^2\mu(WWW\widehat{**}L)\underbrace{-2p^2r^2\mu(WWWDL)}_{\text{(d)}}\\
=\ &\underbrace{(pr^2+2qpr^2+2p^2r^2)\mu(WWWD)}_{\text{combining (a) and (b)}}-(pr^2+2qpr^2)\mu(WWWWD)-\underbrace{(pr^2+2qpr^2+2p^2r^2)\mu(LDWWW)}_{\text{combining (c) and (d)}}\\
&+\underbrace{(pr^2+2qpr^2)\mu(WWW\widehat{**})}_{\text{use (3)}}-(pr^2+2qpr^2)\mu(DWWWD)+2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)]\\
&-2p^2r^2\mu(WWW\widehat{**}L)-2qpr^2\mu(WWW\widehat{**})\\
=\ &(pr^2+2qpr^2+2p^2r^2)\mu(WWWD)\underbrace{-(pr^2+2qpr^2)\mu(WWWWD)}_{\text{(e)}}-(pr^2+2qpr^2+2p^2r^2)\mu(LDWWW)
\end{aligned}
$$

$$+\underbrace{(pr^2+2qpr^2)\mu(WWWDD)+(pr^2+2qpr^2)\mu(WWWDW)+(pr^2+2qpr^2)\mu(WWWWD)}_{\text{(f): obtained by applying (3)}}$$

$$-(pr^2+2qpr^2)\mu(DWWWD)+2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)]-2p^2r^2\mu(WWW\widehat{**}L)$$

$$-2qpr^2\mu(WWW\widehat{**})$$

$$=(pr^2+2qpr^2+2p^2r^2)\mu(WWWD)\underbrace{-(pr^2+2qpr^2)\mu(WWWWD)+(pr^2+2qpr^2)\mu(WWWWD)}_{\text{combining (e) with the last term of (f)}}$$

$$-(pr^2+2qpr^2+2p^2r^2)\mu(LDWWW)+(pr^2+2qpr^2)\big[\underbrace{\mu(WWWDD)+\mu(WWWDW)}_{\text{use (4)}}\big]$$

$$-(pr^2+2qpr^2)\mu(DWWWD)+2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)]-2p^2r^2\mu(WWW\widehat{**}L)$$

$$-2qpr^2\mu(WWW\widehat{**})$$

$$=\underbrace{(pr^2+2qpr^2+2p^2r^2)\mu(WWWD)}_{\text{(g)}}-\underbrace{(pr^2+2qpr^2+2p^2r^2)\mu(LDWWW)}_{\text{(h)}}$$

$$+(pr^2+2qpr^2)[\underbrace{\mu(WWWD)-\mu(WWWDL)}_{\text{(i): obtained by applying (4)}}]-(pr^2+2qpr^2)\mu(DWWWD)$$

$$+2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)]-2p^2r^2\mu(WWW\widehat{**}L)-2qpr^2\mu(WWW\widehat{**})$$

$$=\underbrace{2pr^2(1+p+2q)\mu(WWWD)}_{\text{combining (g) with 1st term of (i)}}-\underbrace{2pr^2(1+2q+p)\mu(LDWWW)}_{\text{combining (h) with 2nd term of (i)}}+2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)]$$

$$-2p^2r^2\mu(WWW\widehat{**}L)-pr^2(1+2q)\mu(DWWWD)-2qpr^2\mu(WWW\widehat{**}). \tag{3.4.21}$$

Applying (3.4.20) and (3.4.21) to (3.4.19), we obtain:

$$w_1(\widehat{\mathrm{F}}_{p,q}\mu)=w_1(\mu)-q\mu(D)-[p(1-r)+q][\mu(WD)+\mu(WWD)]-(p+q)\mu(LWD)$$

$$-[p(2-r)+2q]\mu(LWWD)-(2r+pr)\mu(LDDL)-r\mu(LDL)-(4r+2pr)\mu(LDWL)-pr\mu(LD)$$

$$+2qp^2r\mu(\widehat{***})\underbrace{-r(p+4q^2+4qp^2-2q^3+2p^2+2q^2p)\mu(L\widehat{***})-pr^2(1+2q)\mu(L\widehat{***}L)}_{\text{terms grouped by underbrace numbered (4) of (3.4.20)}}$$

$$\underbrace{+2pr^2(1+p+2q)\mu(WWWD)-2pr^2(1+2q+p)\mu(LDWWW)+2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)]}_{\text{terms obtained from (3.4.21)}}$$

$$\underbrace{-2p^2r^2\mu(WWW\widehat{**}L)-pr^2(1+2q)\mu(DWWWD)-2qpr^2\mu(WWW\widehat{**})}_{\text{terms obtained from (3.4.21)}}$$

$$=w_1(\mu)-q\mu(D)-[p(1-r)+q][\mu(WD)+\mu(WWD)]-(p+q)\mu(LWD)-[p(2-r)+2q]\mu(LWWD)$$

$$-(2r+pr)\mu(LDDL)-r\mu(LDL)-(4r+2pr)\mu(LDWL)-pr\mu(LD)\underbrace{+2qp^2r\mu(\widehat{***})}$$

$$-r(p+4q^2+4qp^2-2q^3+2p^2+2q^2p)\mu(L\widehat{***})-pr^2(1+2q)\mu(L\widehat{***}L)$$

$$
\begin{aligned}
&-pr^2(1+2q)\mu(DWWWD)-2qpr^2\mu(WWW\widehat{**})\underbrace{+2pr^2(1+p+2q)\mu(WWWD)}\\
&-2pr^2(1+2q+p)\mu(LDWWW)\underbrace{+2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)]}\\
&-2p^2r^2\mu(WWW\widehat{**}L)\quad\text{(only rearrangements done in this step).}
\end{aligned}
\tag{3.4.22}
$$

Again, (3.4.22), much like (3.4.14), forms part of the foundation upon which the derivation of an inequality of the form (3.4.2) is built. Still, there are terms in the right side of (3.4.22), other than $w_1(\mu)$, indicated by underbraces in the final step of the derivation of (3.4.22), in which the coefficients are non-negative. Therefore, further adjustments are necessary.

3.4.5 The third step of composing the weight function

In order to make sure that we ultimately end up with an inequality of the form given in (3.4.2), we have to ensure that the non-negative terms highlighted by underbraces in the final expression for (3.4.22) are negated using some of the existing non-positive terms in the final expression for (3.4.22).

The second adjustment is carried out as follows:

$$
\begin{aligned}
w_2(\mu)=w_1(\mu)-\big[&2pr\{\mu(LD)+\mu(LWD)\}+2p^2r\{\mu(LDD)+\mu(LDWD)+\mu(LWDD)\}\\
&+4r\mu(LDWL)+2p\mu(LWWD)\big].
\end{aligned}
\tag{3.4.23}
$$

Again, as stated at the beginning of §3.4.4, a different choice of terms may very well have worked instead of those considered in (3.4.23). Writing (3.4.22) as $w_1(\widehat{\mathrm{F}}_{p,q}\mu)=w_1(\mu)+A_1$, we see, using the idea presented in (3.4.16), that applying (3.4.23) will transform (3.4.22) into

$$
\begin{aligned}
w_2(\widehat{\mathrm{F}}_{p,q}\mu)=w_2(\mu)+\{&2pr\{\mu(LD)+\mu(LWD)\}+2p^2r\{\mu(LDD)+\mu(LDWD)+\mu(LWDD)\}\\
&+4r\mu(LDWL)+2p\mu(LWWD)\}-\{2pr\{\widehat{\mathrm{F}}_{p,q}\mu(LD)+\widehat{\mathrm{F}}_{p,q}\mu(LWD)\}+2p^2r\{\widehat{\mathrm{F}}_{p,q}\mu(LDD)\\
&+\widehat{\mathrm{F}}_{p,q}\mu(LDWD)+\widehat{\mathrm{F}}_{p,q}\mu(LWDD)\}+4r\widehat{\mathrm{F}}_{p,q}\mu(LDWL)+2p\widehat{\mathrm{F}}_{p,q}\mu(LWWD)\}+A_1.
\end{aligned}
\tag{3.4.24}
$$

As in §3.4.4, we attempt to explain to the reader some of the motivations for choosing the above adjustment. Each of the terms used in (3.4.23) has its own significance, and these are, of course, detailed in the computations that follow for the reader to verify, so we focus on just the reason for our choice of the term $-2pr\mu(LD)$.

The terms in the right side of (3.4.22) that are non-negative are $2qp^2r\mu(\widehat{***})$, $2pr^2(1+p+$

$2q)\mu(WWWD)$ and $2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)]$. When p and q are small, the order of magnitude of the coefficient of the first of these terms is $2qp^2$, that of the second term is $2p$, and that of the third term is $2p^2$. We emphasize here, to the reader, that in our approach for constructing the weight function, we have tried to first annihilate the non-negative terms whose coefficients have higher orders of magnitude, and then addressed the non-negative terms whose coefficients have lower orders of magnitude. So, here, we first focus on $2pr^2(1+p+2q)\mu(WWWD)$, which can be thought of as being approximately equal to $2p\mu(WWWD)$ when p and q are small.

From (3.4.28), we know that one of the terms in the expansion of $\widehat{\mathrm{F}}_{p,q}\mu(LD)$ is $r^2\mu(WWWD)$. Therefore, $-2pr\widehat{\mathrm{F}}_{p,q}\mu(LD)$ in (3.4.24) will contribute $-2pr^3\mu(WWWD)$, which is approximately equal to $-2p\mu(WWWD)$ when p and q are small. This, then, aids in negating much of the non-negative term $2pr^2(1+p+2q)\mu(WWWD)$ – in fact, what is left of $2pr^2(1+p+2q)\mu(WWWD)$ after implementing this adjustment is a non-negative term whose coefficient is of a smaller order of magnitude.

At the same time, we note that on the right side of (3.4.24), the term $2pr\mu(LD)$ is being added, and we need to make sure that this does not upset any of the existing non-positive terms to the extent that they become non-negative. Note that there already exists a $-pr\mu(LD)$ on the right side of (3.4.22). Next, we note that

1. $-r(p+4q^2+4qp^2-2q^3+2p^2+2q^2p)\mu(L\widehat{***})$ on the right side of (3.4.22) supplies us with the terms $-pr\mu(LDDD)$, $-pr\mu(LDDW)$, $-pr\mu(LDWD)$ and $-pr\mu(LDWW)$,
2. the term $-(2r+pr)\mu(LDDL)$ provides us with $-pr\mu(LDDL)$,
3. the term $-(4r+2pr)\mu(LDWL)$ provides us with $-pr\mu(LDWL)$,
4. and the term $-r\mu(LDL)$ supplies us with $-pr\mu(LDL)$.

We see that $-pr\mu(LDDD)-pr\mu(LDDW)-pr\mu(LDWD)-pr\mu(LDWW)-pr\mu(LDDL)-pr\mu(LDWL)-pr\mu(LDL)=-pr\mu(LD)$. Thus, we have accounted for non-positive terms existing on the right side of (3.4.22) that are capable of negating, together, the $2pr\mu(LD)$ that gets introduced into the right side of (3.4.24) because of (3.4.23).

Once again, we hope that the above discussion is able to shed some light on why $-2pr\mu(LD)$ is a part of (3.4.23). Likewise, the presence of the other terms in (3.4.23) can be justified / motivated.

Applying the adjustment described in (3.4.23) to (3.4.22), we obtain (as shown in (3.4.24)):

$$w_2(\widehat{\mathrm{F}}_{p,q}\mu)=w_2(\mu)+[2pr\{\mu(LD)+\mu(LWD)\}+2p^2r\{\mu(LDD)+\mu(LDWD)+\mu(LWDD)\}$$
$$+4r\mu(LDWL)+2p\mu(LWWD)]-[2pr\{\widehat{\mathrm{F}}_{p,q}\mu(LD)+\widehat{\mathrm{F}}_{p,q}\mu(LWD)\}+2p^2r\{\widehat{\mathrm{F}}_{p,q}\mu(LDD)$$

$$
\begin{aligned}
&+\widehat{\mathrm{F}}_{p,q}\mu(LDWD)+\widehat{\mathrm{F}}_{p,q}\mu(LWDD)\}+4r\widehat{\mathrm{F}}_{p,q}\mu(LDWL)+2p\widehat{\mathrm{F}}_{p,q}\mu(LWWD)]-q\mu(D)\\
&-[p(1-r)+q][\mu(WD)+\mu(WWD)]-(p+q)\mu(LWD)-[p(2-r)+2q]\mu(LWWD)\\
&-(2r+pr)\mu(LDDL)-r\mu(LDL)-(4r+2pr)\mu(LDWL)-pr\mu(LD)+2qp^2r\mu(\widehat{***})\\
&-r(p+4q^2+4qp^2-2q^3+2p^2+2q^2p)\mu(L\widehat{***})-pr^2(1+2q)\mu(L\widehat{***}L)-pr^2(1+2q)\mu(DWWWD)\\
&-2qpr^2\mu(WWW\widehat{**})+2pr^2(1+p+2q)\mu(WWWD)-2pr^2(1+2q+p)\mu(LDWWW)\\
&+2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)]-2p^2r^2\mu(WWW\widehat{**}L)\\
={}&w_2(\mu)+\underbrace{2pr\{\mu(LD)+\mu(LWD)\}+2p^2r\{\mu(LDD)+\mu(LDWD)+\mu(LWDD)\}+4r\mu(LDWL)}\\
&\underbrace{+2p\mu(LWWD)-(p+q)\mu(LWD)-[p(2-r)+2q]\mu(LWWD)-(2r+pr)\mu(LDDL)-r\mu(LDL)}\\
&\underbrace{-(4r+2pr)\mu(LDWL)-pr\mu(LD)-r(p+4q^2+4qp^2-2q^3+2p^2+2q^2p)\mu(L\widehat{***})}-q\mu(D)\\
&-[p(1-r)+q][\mu(WD)+\mu(WWD)]+2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(L\widehat{***}L)\\
&-pr^2(1+2q)\mu(DWWWD)-2qpr^2\mu(WWW\widehat{**})+2pr^2(1+p+2q)\mu(WWWD)\\
&-2pr^2(1+2q+p)\mu(LDWWW)+2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)]-2p^2r^2\mu(WWW\widehat{**}L)\\
&-2pr\{\widehat{\mathrm{F}}_{p,q}\mu(LD)+\widehat{\mathrm{F}}_{p,q}\mu(LWD)\}-2p^2r\{\widehat{\mathrm{F}}_{p,q}\mu(LDD)+\widehat{\mathrm{F}}_{p,q}\mu(LDWD)+\widehat{\mathrm{F}}_{p,q}\mu(LWDD)\}\\
&-4r\widehat{\mathrm{F}}_{p,q}\mu(LDWL)-2p\widehat{\mathrm{F}}_{p,q}\mu(LWWD). \qquad (3.4.25)
\end{aligned}
$$

3.4.5.1 Step 1 of analyzing (3.4.25)

First, we combine the terms of (3.4.25) that have been highlighted using underbraces. To this end, we use the identities:

1. $\mu(L\widehat{***})=\mu(LD)+\mu(LWD)+\mu(LWWD)-\mu(LDL)-\mu(LDDL)-2\mu(LDWL)$,
2. $\mu(LDDD)+\mu(LDDW)=\mu(LDD)-\mu(LDDL)$,
3. $\mu(L\widehat{***})=\mu(LDDD)+\mu(LDDW)+\mu(LDWD)+\mu(LWDD)+\mu(LDWW)+\mu(LWDW)+\mu(LWWD)$,

and we obtain:

$$
\begin{aligned}
&\underbrace{-(p+q)\mu(LWD)}_{(1)}\underbrace{-\{p(2-r)+2q\}\mu(LWWD)}_{(3)}-(2r+pr)\mu(LDDL)-r\mu(LDL)\underbrace{-(4r+2pr)\mu(LDWL)}_{(5)}\\
&\underbrace{-pr\mu(LD)}_{(7)}-r(p+4q^2+4qp^2-2q^3+2p^2+2q^2p)\mu(L\widehat{***})\underbrace{+2pr\mu(LD)}_{(8)}\underbrace{+2pr\mu(LWD)}_{(2)}\\
&+2p^2r\{\mu(LDD)+\mu(LDWD)+\mu(LWDD)\}\underbrace{+4r\mu(LDWL)}_{(6)}\underbrace{+2p\mu(LWWD)}_{(4)}
\end{aligned}
$$

$$= \underbrace{-(p+q)\mu(LWD)+2pr\mu(LWD)}_{\text{combining (1) and (2)}}\underbrace{-\{p(2-r)+2q\}\mu(LWWD)+2p\mu(LWWD)}_{\text{combining (3) and (4)}}-(2r+pr)\mu(LDDL)$$

$$-r\mu(LDL)\underbrace{-(4r+2pr)\mu(LDWL)+4r\mu(LDWL)}_{\text{combining (5) and (6)}}\underbrace{-pr\mu(LD)+2pr\{\mu(LD)}_{\text{combining (7) and (8)}}\underbrace{-pr\mu(L\widehat{***})}_{\text{apply (1)}}$$

$$-2p^2r\underbrace{\mu(L\widehat{***})}_{\text{split using (3)}}-r(4q^2+4qp^2-2q^3+2q^2p)\mu(L\widehat{***})+2p^2r\{\mu(LDD)+\mu(LDWD)+\mu(LWDD)\}$$

$$= [-(p+q)+2pr]\mu(LWD)-(2q-pr)\mu(LWWD)-(2r+pr)\mu(LDDL)-r\mu(LDL)-2pr\mu(LDWL)$$

$$+pr\mu(LD)\underbrace{-pr[\mu(LD)+\mu(LWD)+\mu(LWWD)-\mu(LDL)-\mu(LDDL)-2\mu(LDWL)]}_{\text{after applying (1)}}$$

$$-2p^2r\big[\mu(LDDD)+\mu(LDDW)+\mu(LDWD)+\mu(LWDD)+\mu(LDWW)+\mu(LWDW)+\mu(LWWD)\big]$$

$$-r(4q^2+4qp^2-2q^3+2q^2p)\mu(L\widehat{***})+2p^2r\{\mu(LDD)+\mu(LDWD)+\mu(LWDD)\}$$

$$= \underbrace{[-(p+q)+2pr]\mu(LWD)}_{(9)}\underbrace{-(2q-pr)\mu(LWWD)}_{(11)}\underbrace{-(2r+pr)\mu(LDDL)}_{(13)}\underbrace{-r\mu(LDL)}_{(15)}\underbrace{-2pr\mu(LDWL)}_{(17)}$$

$$\underbrace{+pr\mu(LD)}_{(19)}\underbrace{-pr\mu(LD)}_{(20)}\underbrace{-pr\mu(LWD)}_{(10)}\underbrace{-pr\mu(LWWD)}_{(12)}\underbrace{+pr\mu(LDL)}_{(16)}\underbrace{+pr\mu(LDDL)}_{(14)}\underbrace{+2pr\mu(LDWL)}_{(18)}$$

$$-2p^2r\big[\underbrace{\mu(LDDD)+\mu(LDDW)}_{\text{apply (2)}}+\mu(LDWD)+\mu(LWDD)+\mu(LDWW)+\mu(LWDW)+\mu(LWWD)\big]$$

$$-r(4q^2+4qp^2-2q^3+2q^2p)\underbrace{\mu(L\widehat{***})}_{\text{split using (3)}}+2p^2r\{\mu(LDD)+\mu(LDWD)+\mu(LWDD)\}$$

$$= \underbrace{[-(p+q)+pr]\mu(LWD)}_{\text{adding (9) \& (10)}}\underbrace{-2q\mu(LWWD)}_{\text{adding (11) \& (12)}}\underbrace{-2r\mu(LDDL)}_{\text{adding (13) \& (14)}}\underbrace{-r(1-p)\mu(LDL)}_{\text{adding (15) \& (16)}}-2p^2r[\underbrace{\mu(LDD)-\mu(LDDL)}_{\text{after applying (2)}}$$

$$+\mu(LDWD)+\mu(LWDD)]-2p^2r[\mu(LDWW)+\mu(LWDW)+\mu(LWWD)]$$

$$-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDDW)+\mu(LDWD)+\mu(LWDD)+\mu(LDWW)$$

$$+\mu(LWDW)+\mu(LWWD)]+2p^2r\{\mu(LDD)+\mu(LDWD)+\mu(LWDD)\}$$

(also adding the terms indicated by underbraces numbered (17), (18), (19) & (20))

$$= -[p(1-r)+q]\mu(LWD)\underbrace{-2q\mu(LWWD)}_{(21)}\underbrace{-2r\mu(LDDL)}_{(24)}-r(1-p)\mu(LDL)\underbrace{+2p^2r\mu(LDDL)}_{(25)}$$

$$\underbrace{-2p^2r[\mu(LDD)+\mu(LDWD)+\mu(LWDD)]}_{(26)}\underbrace{-2p^2r[\mu(LDWW)+\mu(LWDW)]}_{(28)}\underbrace{-2p^2r\mu(LWWD)}_{(22)}$$

$$-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDDW)+\mu(LDWD)+\mu(LWDD)$$

$$\underbrace{+\mu(LDWW)+\mu(LWDW)]}_{(29)}\underbrace{-r(4q^2+4qp^2-2q^3+2q^2p)\mu(LWWD)}_{(23)}$$

$$\underbrace{+2p^2r\{\mu(LDD)+\mu(LDWD)+\mu(LWDD)\}}_{(27)}$$

$$= -[p(1-r)+q]\mu(LWD)\underbrace{-[2q+r(4q^2+4qp^2-2q^3+2p^2+2q^2p)]\mu(LWWD)}_{\text{adding (21), (22), (23)}}\underbrace{-2r(1-p^2)\mu(LDDL)}_{\text{adding (24) \& (25)}}$$
$$-r(1-p)\mu(LDL)-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDDW)+\mu(LDWD)+\mu(LWDD)]$$
$$\underbrace{-r(4q^2+4qp^2-2q^3+2p^2+2q^2p)[\mu(LDWW)+\mu(LWDW)]}_{\text{adding (28) \& (29)}} \quad \text{(also cancelling (26) by (27)).} \tag{3.4.26}$$

Incorporating the expression obtained in (3.4.26) (as indicated by the large square brackets in the expression below), we can now rewrite (3.4.25) as follows:

$$w_2(\widehat{\mathrm{F}}_{p,q}\mu) = w_2(\mu) + \Bigg[-[p(1-r)+q]\mu(LWD)-[2q+r(4q^2+4qp^2-2q^3+2p^2+2q^2p)]\mu(LWWD)$$
$$-2r(1-p^2)\mu(LDDL)-r(1-p)\mu(LDL)-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDDW)$$
$$+\mu(LDWD)+\mu(LWDD)]-r(4q^2+4qp^2-2q^3+2p^2+2q^2p)[\mu(LDWW)+\mu(LWDW)]\Bigg]-q\mu(D)$$
$$-[p(1-r)+q][\mu(WD)+\mu(WWD)]+2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(L\widehat{***}L)-pr^2(1+2q)\mu(DWWWD)$$
$$-2qpr^2\mu(WWW\widehat{**})+2pr^2(1+p+2q)\mu(WWWD)-2pr^2(1+2q+p)\mu(LDWWW)+2p^2r^2[\mu(WWWWD)$$
$$+\mu(WWWWWD)]-2p^2r^2\mu(WWW\widehat{**}L)\underbrace{-2pr\{\widehat{\mathrm{F}}_{p,q}\mu(LD)+\widehat{\mathrm{F}}_{p,q}\mu(LWD)\}-2p^2r\{\widehat{\mathrm{F}}_{p,q}\mu(LDD)}$$
$$\underbrace{+\widehat{\mathrm{F}}_{p,q}\mu(LDWD)+\widehat{\mathrm{F}}_{p,q}\mu(LWDD)\}-4r\widehat{\mathrm{F}}_{p,q}\mu(LDWL)-2p\widehat{\mathrm{F}}_{p,q}\mu(LWWD)}. \tag{3.4.27}$$

3.4.5.2 Step 2 of analyzing (3.4.25)

In §3.4.5.3, we deal with the terms of (3.4.27) that have been grouped using underbraces. To this end, we need to compute $\widehat{\mathrm{F}}_{p,q}\mu(LD)$, $\widehat{\mathrm{F}}_{p,q}\mu(LWD)$, $\widehat{\mathrm{F}}_{p,q}\mu(LDD)$, $\widehat{\mathrm{F}}_{p,q}\mu(LDWD)$, $\widehat{\mathrm{F}}_{p,q}\mu(LWDD)$ and $\widehat{\mathrm{F}}_{p,q}\mu(LDWL)$ (note that we have already computed $\widehat{\mathrm{F}}_{p,q}\mu(LWWD)$ in (3.4.6)). An argument similar to that adopted in (3.4.5) of §3.4.1 for computing $\widehat{\mathrm{F}}_{p,q}\mu(WD)$ can be used to deduce that

$$\widehat{\mathrm{F}}_{p,q}\mu(LD) = r^2\mu(WWWD)+qr\mu(\widehat{***}). \tag{3.4.28}$$

However, parts of the expression for each of the probabilities $\widehat{\mathrm{F}}_{p,q}\mu(LWD)$, $\widehat{\mathrm{F}}_{p,q}\mu(LDD)$, $\widehat{\mathrm{F}}_{p,q}\mu(LDWD)$, $\widehat{\mathrm{F}}_{p,q}\mu(LWDD)$ and $\widehat{\mathrm{F}}_{p,q}\mu(LDWL)$ will deliberately *not* be made explicit, in order to keep the subsequent mathematical expressions as concise as possible.

While computing $\widehat{\mathrm{F}}_{p,q}\mu(LWD)$, we first consider $\eta(0) = \eta(1) = \eta(2) = W$, so that $(\eta(3), \eta(4)) \in \widehat{**}$, the event $\widehat{\mathrm{F}}_{p,q}\eta(0) = L$ happens with probability $1-p$, and the event $\widehat{\mathrm{F}}_{p,q}\eta(1) = W$ happens with probability p. Next, we consider $\eta(0) = 1$, $\eta(1) \in \{W,D\}$ and $(\eta(2), \eta(3), \eta(4)) \in \widehat{***}$, so that $\widehat{\mathrm{F}}_{p,q}\eta(0) = L$ happens with probability q and $\widehat{\mathrm{F}}_{p,q}\eta(1) = W$ happens with probability p, and finally, we consider $\eta(1) = L$ and $(\eta(2), \eta(3), \eta(4)) \in \widehat{***}$, so that $\widehat{\mathrm{F}}_{p,q}\eta(0) = L$ happens with probability q and $\widehat{\mathrm{F}}_{p,q}\eta(1) = W$ happens with probability $1-q$. In each of these cases, $\widehat{\mathrm{F}}_{p,q}\eta(2) = D$ happens with probability r. Therefore

$$\widehat{\mathrm{F}}_{p,q}\mu(LWD) = (1-p)pr\mu(WWW\widehat{**}) + C_{LWD} + D_{LWD} \geqslant (1-p)pr\mu(WWW\widehat{**}) + C_{LWD}, \tag{3.4.29}$$

where

$$C_{LWD} = qpr\mu(L\{W,D\}\widehat{***}) + q(1-q)r\mu(L\widehat{***}), \tag{3.4.30}$$

and D_{LWD} is the component arising from the case where $(\eta(0), \eta(1), \eta(2)) \in \widehat{***}$. Similar arguments lead to

- $\widehat{\mathrm{F}}_{p,q}\mu(LDD) = (1-p)r^2\mu(WWWD\{W,D\}) + C_{LDD} \geqslant (1-p)r^2\mu(WWWD\{W,D\})$; (3.4.31)
- $\widehat{\mathrm{F}}_{p,q}\mu(LDWD) = (1-p)r^2p\mu(WWWD\{W,D\}^2) + C_{LDWD} \geqslant (1-p)r^2p\mu(WWWD\{W,D\}^2)$; (3.4.32)
- $\widehat{\mathrm{F}}_{p,q}\mu(LWDD) = (1-p)pr^2\mu(WWW\widehat{**}\{W,D\}) + C_{LWDD} \geqslant (1-p)pr^2\mu(WWW\widehat{**}\{W,D\})$; (3.4.33)

in which $C_{\mathscr{C}}$, where $\mathscr{C}$ is any of the cylinder sets (LDD), $(LDWD)$, $(LWDD)$, accounts for the contribution from the cases in which $(\eta(0), \eta(1), \eta(2)) \in \hat{\mathscr{A}}^3 \setminus \{(W,W,W)\}$.

While computing $\widehat{\mathrm{F}}_{p,q}\mu(LDWL)$, we first consider the case where $\eta(3) = \eta(4) = \eta(5) = W$, so that $(\eta(1), \eta(2)) \in \widehat{**}$, and the event $\widehat{\mathrm{F}}_{p,q}\eta(2) = W$ happens with probability p while the event $\widehat{\mathrm{F}}_{p,q}\eta(3) = L$ happens with probability $1-p$. Note that in this situation, no matter what the value of $\eta(0)$ is, the event $\widehat{\mathrm{F}}_{p,q}\eta(0) = L$ happens with probability q. The second possibility we take into account is where $\eta(0) = \eta(1) = \eta(2) = W$, which then forces $\eta(3) = D$, and $\widehat{\mathrm{F}}_{p,q}\eta(0) = L$ happens with probability $1-p$. If $\eta(4) \in \{W,D\}$, the events $\widehat{\mathrm{F}}_{p,q}\eta(2) = W$ and $\widehat{\mathrm{F}}_{p,q}\eta(3) = L$ happen with probabilities p and q respectively, whereas if $\eta(4) = L$, they happen with probabilities $1-q$ and q respectively. Combining all, we have

$$\begin{aligned}\widehat{\mathrm{F}}_{p,q}\mu(LDWL) = {} & qrp(1-p)\mu(\widehat{**}WWW) + (1-p)rpq\mu(WWWD\{W,D\}) \\ & + (1-p)r(1-q)q\mu(WWWDL) + C_{LDWL}\end{aligned}$$

$$
\begin{aligned}
&= qrp(1-p)\{\mu(DWWW)-\mu(LDWWW)+\mu(DWWWW)\}\\
&\quad +(1-p)rpq\{\mu(WWWD)-\mu(WWWDL)\}+(1-p)r(1-q)q\mu(WWWDL)+C_{LDWL}\\
&= \underbrace{qrp(1-p)\mu(DWWW)+(1-p)rpq\mu(WWWD)}+qrp(1-p)\mu(DWWWW)\\
&\quad \underbrace{-qrp(1-p)\mu(LDWWW)-(1-p)rpq\mu(WWWDL)+(1-p)r(1-q)q\mu(WWWDL)}\\
&\quad +C_{LDWL}\\
&\geqslant 2(1-p)rpq\mu(WWWD)+(1-p)prq\mu(WWWWD)+(1-p)rq(r-p)\mu(WWWDL),
\end{aligned}
\tag{3.4.34}
$$

where we make use of the reflection-invariance of μ, and C_{LDWL} takes into account the contributions from the situations not considered above.

Finally, instead of writing the entire expression of (3.4.6) in place of $\widehat{\mathrm{F}}_{p,q}\,\mu(LWWD)$, we write

$$
\widehat{\mathrm{F}}_{p,q}\,\mu(LWWD) = p^2r^2\mu(WWW\widehat{***})+C_{LWWD}, \tag{3.4.35}
$$

where

$$
C_{LWWD} = qp^2r\mu(\widehat{***})+qpr^2\mu(L\{W,D\}\widehat{***})+qr^2(1-q+p)\mu(L\widehat{***}). \tag{3.4.36}
$$

3.4.5.3 Step 3 of analyzing (3.4.25)

As mentioned at the start of §3.4.5.2, we deal with the terms grouped using underbraces in (3.4.27), using (3.4.28), (3.4.29), (3.4.31), (3.4.32), (3.4.33), (3.4.34), (3.4.35), and applying the identities

1. $\mu(WWW\widehat{***}) = \mu(WWWD)+\mu(WWWWD)+\mu(WWWWWD)-\mu(WWW\widehat{**}L)-\mu(WWWDL)$,
2. $\mu(WWW\widehat{**}) = \mu(WWWD)-\mu(WWWDL)+\mu(WWWWD)$,
3. $\mu(WWWD\{W,D\}) = \mu(WWWD)-\mu(WWWDL)$,
4. $\mu(WWWD\{W,D\}^2) = \mu(WWWD)-\mu(WWWDL)-\mu(WWWD\{W,D\}L)$,
5. $\mu(WWW\widehat{**}\{W,D\}) = \mu(WWWD)+\mu(WWWWD)-\mu(WWWDL)-\mu(WWWD\{W,D\}L)-\mu(WWWWDL)$,

Using (3.4.28), (3.4.29), (3.4.31), (3.4.32), (3.4.33), (3.4.34) and (3.4.35) in the first step, and subsequently applying the identities mentioned above, we find that the sum of the terms grouped using underbraces in (3.4.27) can be bounded above as follows:

$$
\begin{aligned}
& -2pr\{\widehat{\mathrm{F}}_{p,q}\,\mu(LD)+\widehat{\mathrm{F}}_{p,q}\,\mu(LWD)\}-2p^2r\{\widehat{\mathrm{F}}_{p,q}\,\mu(LDD)+\widehat{\mathrm{F}}_{p,q}\,\mu(LDWD)+\widehat{\mathrm{F}}_{p,q}\,\mu(LWDD)\}\\
& -4r\widehat{\mathrm{F}}_{p,q}\,\mu(LDWL)-2p\widehat{\mathrm{F}}_{p,q}\,\mu(LWWD)\\
\leqslant\; & -2pr^3\mu(WWWD)-2pqr^2\mu(\widehat{***})-2p^2r^2(1-p)\underbrace{\mu(WWW\widehat{**})}_{\text{use (2)}}-2prC_{LWD}\\
& -2p^2r^3(1-p)\underbrace{\mu(WWWD\{W,D\})}_{\text{use (3)}}-2p^3r^3(1-p)\underbrace{\mu(WWWD\{W,D\}^2)}_{\text{use (4)}}\\
& -2p^3r^3(1-p)\underbrace{\mu(WWW\widehat{**}\{W,D\})}_{\text{use (5)}}-8(1-p)r^2pq\mu(WWWD)-4(1-p)pr^2q\mu(WWWWD)\\
& -4(1-p)r^2q(r-p)\mu(WWWDL)-2p^3r^2\underbrace{\mu(WWW\widehat{***})}_{\text{use (1)}}-2pC_{LWWD}\\
=\; & -2pr^3\mu(WWWD)-2pqr^2\mu(\widehat{***})-2p^2r^2(1-p)\underbrace{[\mu(WWWD)-\mu(WWWDL)+\mu(WWWWD)]}_{\text{after applying (2)}}\\
& -2p^2r^3(1-p)\underbrace{[\mu(WWWD)-\mu(WWWDL)]}_{\text{after applying (3)}}\\
& -2p^3r^3(1-p)\underbrace{[\mu(WWWD)-\mu(WWWDL)-\mu(WWWD\{W,D\}L)]}_{\text{after applying (4)}}\\
& -2p^3r^3(1-p)\underbrace{[\mu(WWWD)-\mu(WWWDL)-\mu(WWWD\{W,D\}L)+\mu(WWWWD)-\mu(WWWWDL)]}_{\text{after applying (5)}}\\
& -8(1-p)r^2pq\mu(WWWD)-4(1-p)pr^2q\mu(WWWWD)-4(1-p)r^2q(r-p)\mu(WWWDL)\\
& -2p^3r^2\underbrace{[\mu(WWWD)+\mu(WWWWD)+\mu(WWWWWD)-\mu(WWWDL)}_{\text{after applying (1)}}\\
& \underbrace{-\mu(WWWD\{W,D\}L)-\mu(WWWWDL)]}_{\text{after applying (1)}}-2prC_{LWD}-2pC_{LWWD}\\
=\; & -[2pr^3+2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)+8(1-p)r^2pq+2p^3r^2]\mu(WWWD)\\
& -[2p^2r^2(1-p)+2p^3r^3(1-p)+4(1-p)pr^2q+2p^3r^2]\mu(WWWWD)-2p^3r^2\mu(WWWWWD)\\
& +[2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)-4(1-p)r^2q(r-p)+2p^3r^2]\mu(WWWDL)\\
& +[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L)+[2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)\\
& -2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD},
\end{aligned}
\tag{3.4.37}
$$

where, in the last step, we simply add all those terms that involve $\mu(\mathscr{C})$, for $\mathscr{C}$ being any of the cylinder sets $(WWWD)$, $(WWWWD)$, $(WWWWWD)$, $(WWWDL)$, $(WWWD\{W,D\}L)$ and

$(WWWWDL)$. Incorporating the inequality (3.4.37), we now see that the equality in (3.4.27) is transformed into an inequality as follows (it is worthwhile to note here that we encountered only weight function *equalities* in the various steps of adjustments carried out earlier, and this is the first time we encounter a weight function *inequality*):

$$
\begin{aligned}
&w_2(\widehat{\mathrm{F}}_{p,q}\mu) \leqslant w_2(\mu) - [p(1-r)+q]\mu(LWD) - [2q + r(4q^2+4qp^2-2q^3+2p^2+2q^2p)]\mu(LWWD)\\
&-2r(1-p^2)\mu(LDDL) - r(1-p)\mu(LDL) - r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDDW)\\
&+\mu(LDWD)+\mu(LWDD)] - r(4q^2+4qp^2-2q^3+2p^2+2q^2p)[\mu(LDWW)+\mu(LWDW)] - q\mu(D)\\
&-[p(1-r)+q][\mu(WD)+\mu(WWD)] - pr^2(1+2q)\mu(L\widehat{***}L)\\
&\underbrace{+2qp^2r\mu(\widehat{***}) - pr^2(1+2q)\mu(DWWWD) - 2qpr^2\mu(WWW\widehat{**}) + 2pr^2(1+p+2q)\mu(WWWD)}\\
&\underbrace{-2pr^2(1+2q+p)\mu(LDWWW) + 2p^2r^2[\mu(WWWWD)+\mu(WWWWWD)] - 2p^2r^2\mu(WWW\widehat{**}L)}\\
&\underbrace{-[2pr^3+2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)+8(1-p)r^2pq+2p^3r^2]\mu(WWWD)}\\
&\underbrace{-[2p^2r^2(1-p)+2p^3r^3(1-p)+4(1-p)pr^2q+2p^3r^2]\mu(WWWWD) - 2p^3r^2\mu(WWWWWD)}\\
&\underbrace{+[2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)-4(1-p)r^2q(r-p)+2p^3r^2]\mu(WWWDL)}\\
&\underbrace{+[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L) + [2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)}\\
&\underbrace{-2pqr^2\mu(\widehat{***}) - 2prC_{LWD} - 2pC_{LWWD}}, \qquad (3.4.38)
\end{aligned}
$$

where we let B_2 denote the sum of the terms of (3.4.38) that have been grouped using underbraces, and we let B_1 denote the remaining terms. In other words,

$$
\begin{aligned}
B_1 = w_2(\mu) &- [p(1-r)+q]\mu(LWD) - [2q + r(4q^2+4qp^2-2q^3+2p^2+2q^2p)]\mu(LWWD)\\
&-2r(1-p^2)\mu(LDDL) - r(1-p)\mu(LDL) - r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDDW)\\
&+\mu(LDWD)+\mu(LWDD)] - r(4q^2+4qp^2-2q^3+2p^2+2q^2p)[\mu(LDWW)+\mu(LWDW)] - q\mu(D)\\
&-[p(1-r)+q][\mu(WD)+\mu(WWD)] - pr^2(1+2q)\mu(L\widehat{***}L), \qquad (3.4.39)
\end{aligned}
$$

and

$$
\begin{aligned}
B_2 = {} & \underbrace{2qp^2r\mu(\widehat{***})} - pr^2(1+2q)\mu(DWWWD) - 2qpr^2\mu(WWW\widehat{**}) \underbrace{+2pr^2(1+p+2q)\mu(WWWD)}\\
&-2pr^2(1+2q+p)\mu(LDWWW) \underbrace{+2p^2r^2\mu(WWWWD)+2p^2r^2\mu(WWWWWD)} -2p^2r^2\mu(WWW\widehat{**}L)\\
&-[2pr^3+2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)+8(1-p)r^2pq+2p^3r^2]\mu(WWWD)
\end{aligned}
$$

$$
\begin{aligned}
&-[2p^2r^2(1-p)+2p^3r^3(1-p)+4(1-p)pr^2q+2p^3r^2]\mu(WWWWD)-2p^3r^2\mu(WWWWWD)\\
&\underbrace{+[2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)-4(1-p)r^2q(r-p)+2p^3r^2]\mu(WWWDL)}\\
&\underbrace{+[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L)+[2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)}\\
&-2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD}. \qquad (3.4.40)
\end{aligned}
$$

It is worthwhile to note here that all of the terms, other than $w_2(\mu)$, in B_1 are non-positive, whereas in B_2, the possibly non-negative (we write 'possibly' because the coefficient of $\mu(WWWDL)$ may or may not be non-negative, depending on the values of p and q) terms have been highlighted using underbraces in (3.4.40). We now have to make sure that these non-negative terms are negated using the existing non-positive terms on the right side of (3.4.40), and this is what we accomplish, to some extent, in §3.4.5.4.

3.4.5.4 Step 4 of analyzing (3.4.25)

We dedicate §3.4.5.4 to the analysis of B_2 in (3.4.40). Before we embark on this task, we perform a couple of rather intricate algebraic simplifications that are going to be of use while analysing B_2. The first of these is as follows, and this will be used in combining the coefficients of the various terms involving $\mu(WWWD)$ in the analysis of B_2:

$$
\begin{aligned}
&-2qpr^2+2pr^2(1+p+2q)-[2pr^3+2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)+8(1-p)r^2pq\\
&+2p^3r^2]\\
=&-2qpr^2+2pr^2+2p^2r^2+4pqr^2-2pr^3-2p^2r^2(1-p)-2p^2r^3(1-p)-4p^3r^3(1-p)\\
&-8(1-p)r^2pq-2p^3r^2\\
=&\,2pr^2-2pr^3+2p^2r^2-2p^2r^2(1-p)-2p^3r^2-2p^2r^3(1-p)-4p^3r^3(1-p)-8(1-p)r^2pq+2qpr^2\\
=&\,2pr^2(1-r)\underbrace{+2p^2r^2-2p^2r^2}_{\text{cancel each other}}\underbrace{+2p^3r^2-2p^3r^2}_{\text{cancel each other}}-2p^2r^3(1-p)-4p^3r^3(1-p)-8pqr^2+8p^2qr^2+2qpr^2\\
=&\,2p^2r^2+2pqr^2-2p^2r^3+2p^3r^3-4p^3r^3+4p^4r^3-6pqr^2+8p^2qr^2\\
=&\,2p^2r^2-2p^2r^3-2p^3r^3+4p^4r^3-4pqr^2+8p^2qr^2\\
=&\,2p^2r^2(p+q)-2p^3r^3+4p^4r^3-4pqr^2+8p^2qr^2\\
=&\,2p^3r^2+2p^2qr^2-2p^3r^3+4p^4r^3-4pqr^2+8p^2qr^2\\
=&\,2p^3r^2-2p^3r^3+4p^4r^3-4pqr^2+10p^2qr^2\\
=&\,2p^3r^2(p+q)+4p^4r^3-4pqr^2+10p^2qr^2
\end{aligned}
$$

$$
\begin{aligned}
&= \underbrace{2p^4r^2}_{(1)}\underbrace{+2p^3r^2q}_{(2)}+4p^4r^3\underbrace{-4pqr^2+10p^2qr^2}_{(3)}\\
&= \underbrace{2pqr^2[p^2-2+5p]+6p^4r^2-4p^4r^2}_{\text{combining (1), (2) and (3)}}+4p^4r^3\\
&= 2pqr^2[p^2-2+5p]+6p^4r^2-4p^4r^2(1-r)\\
&= 2pqr^2[p^2-2+5p]+6p^4r^2-4p^4r^2(p+q). \qquad (3.4.41)
\end{aligned}
$$

The second algebraic simplification we detail here is as follows, and this will be utilized in combining the coefficients of the various terms involving $\mu(WWWDL)$ in the analysis of B_2:

$$
\begin{aligned}
&2qpr^2-2pr^2(1+2q+p)+[2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)-4(1-p)r^2q(r-p)+2p^3r^2]\\
&= 2pr^2[q-1-2q-p+p(1-p)+pr(1-p)+2p^2r(1-p)+p^2]-4(1-p)r^2q(r-p)\\
&= -2pr^2[1+q+p-p+p^2-p(1-p-q)(1-p)-2p^2r(1-p)-p^2]-4(1-p)r^2q(r-p)\\
&= -2pr^2[1+q-p(1-p)^2+pq(1-p)-2p^2r(1-p)]-4(1-p)r^2q(r-p)\\
&= -2pr^2[1+q-p(1-2p+p^2)+pq(1-p)-2p^2r(1-p)]-4(1-p)r^2q(r-p)\\
&= -2pr^2[1+q-p+2p^2-p^3+pq(1-p)-2p^2r(1-p)]-4(1-p)r^2q(r-p)\\
&= -2pr^2[1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3]-4(1-p)r^2q(r-p). \qquad (3.4.42)
\end{aligned}
$$

In the analysis that follow, in many steps, we indicate using underbraces the terms that are to be combined in the next step, and we use the following identities:

1. $\mu(WWW\widehat{**})=\mu(WWWD)-\mu(WWWDL)+\mu(WWWWD)$,
2. $\mu(WWWWD)=\mu(WWWWWD)+\mu(DWWWWD)+\mu(LWWWWD)$,
3. $\mu(WWW\widehat{**}L)=\mu(WWWD\{W,D\}L)+\mu(WWWWDL)$.

The simplification of B_2 continues as follows:

$$
\begin{aligned}
B_2 = {}& 2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(DWWWD)\underbrace{-2qpr^2\mu(WWW\widehat{**})}_{\text{use (1)}}\underbrace{+2pr^2(1+p+2q)\mu(WWWD)}_{\text{term involving }\mu(WWWD)}\\
&-2pr^2(1+2q+p)\mu(LDWWW)+2p^2r^2\mu(WWWWD)+2p^2r^2\mu(WWWWWD)-2p^2r^2\mu(WWW\widehat{**}L)\\
&\underbrace{-[2pr^3+2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)+8(1-p)r^2pq+2p^3r^2]\mu(WWWD)}_{\text{term involving }\mu(WWWD)}\\
&-[2p^2r^2(1-p)+2p^3r^3(1-p)+4(1-p)pr^2q+2p^3r^2]\mu(WWWWD)-2p^3r^2\mu(WWWWWD)\\
&+[2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)-4(1-p)r^2q(r-p)+2p^3r^2]\mu(WWWDL)
\end{aligned}
$$

$$
\begin{aligned}
&+[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L)+[2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)\\
&-2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD}\\
={}&2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(DWWWD)\underbrace{-2qpr^2\mu(WWWWD)+2qpr^2\mu(WWWDL)}_{\text{obtained from (1)}}\\
&-2pr^2(1+2q+p)\mu(LDWWW)+2p^2r^2\mu(WWWWD)+2p^2r^2\mu(WWWWWD)\\
&-2p^2r^2\mu(WWW\widehat{**}L)\qquad\underbrace{+[2pqr^2\{5p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)}_{\text{summing terms involving }\mu(WWWD)\text{, including the one obtained from (1), and using (3.4.41)}}\\
&-[2p^2r^2(1-p)+2p^3r^3(1-p)+4(1-p)pr^2q+2p^3r^2]\mu(WWWWD)-2p^3r^2\mu(WWWWWD)\\
&+[2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)-4(1-p)r^2q(r-p)+2p^3r^2]\mu(WWWDL)\\
&+[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L)+[2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)\\
&-2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD}\\
={}&2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(DWWWD)\underbrace{-2qpr^2\mu(WWWWD)}_{\text{term involving }\mu(WWWWD)}+2qpr^2\mu(WWWDL)\\
&-2pr^2(1+2q+p)\mu(LDWWW)\underbrace{+2p^2r^2\mu(WWWWD)}_{\text{term involving }\mu(WWWWD)}+2p^2r^2\mu(WWWWWD)\\
&-2p^2r^2\mu(WWW\widehat{**}L)+[2pqr^2\{5p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)\\
&\underbrace{-[2p^2r^2(1-p)+2p^3r^3(1-p)+4(1-p)pr^2q+2p^3r^2]\mu(WWWWD)}_{\text{term involving }\mu(WWWWD)}-2p^3r^2\mu(WWWWWD)\\
&+[2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)-4(1-p)r^2q(r-p)+2p^3r^2]\mu(WWWDL)\\
&+[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L)+[2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)\\
&-2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD}\\
={}&2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(DWWWD)+2qpr^2\mu(WWWDL)-2pr^2(1+2q+p)\mu(LDWWW)\\
&+2p^2r^2\mu(WWWWWD)-2p^2r^2\mu(WWW\widehat{**}L)+[2pqr^2\{5p-2+p^2\}+6p^4r^2\\
&-4p^4r^2(p+q)]\mu(WWWD)\underbrace{-(6pqr^2-4p^2qr^2)\mu(WWWWD)-2p^3r^3(1-p)\mu(WWWWD)}_{\text{summing terms involving }\mu(WWWWD)}\\
&-2p^3r^2\mu(WWWWWD)+[2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)\\
&-4(1-p)r^2q(r-p)+2p^3r^2]\mu(WWWDL)+[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L)\\
&+[2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)-2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD}\\
={}&2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(DWWWD)+2qpr^2\mu(WWWDL)-2pr^2(1+2q+p)\mu(LDWWW)\\
&\underbrace{+2p^2r^2\mu(WWWWWD)}_{\text{term involving }\mu(WWWWWD)}-2p^2r^2\mu(WWW\widehat{**}L)+[2pqr^2\{5p-2+p^2\}
\end{aligned}
$$

$$
\begin{aligned}
&+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)-(6pqr^2-4p^2qr^2)\mu(WWWWD)\\
&\underbrace{-2p^3r^3(1-p)\mu(WWWWWD)}_{\text{use (2)}}\underbrace{-2p^3r^2\mu(WWWWWWD)}_{\text{term involving }\mu(WWWWWWD)}+[2p^2r^2(1-p)+2p^2r^3(1-p)\\
&+4p^3r^3(1-p)-4(1-p)r^2q(r-p)+2p^3r^2]\mu(WWWDL)\\
&+[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L)+[2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)\\
&-2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD}\\
=\;&2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(DWWWD)+2qpr^2\mu(WWWDL)-2pr^2(1+2q+p)\mu(LDWWW)\\
&-2p^2r^2\mu(WWW\widehat{**}L)+[2pqr^2\{5p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)\\
&-(6pqr^2-4p^2qr^2)\mu(WWWWD)\qquad\underbrace{+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)}_{\text{summing terms involving }\mu(WWWWWD)\text{, including that obtained from (2)}}\\
&\underbrace{-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}}_{\text{remaining terms obtained by applying (2)}}\\
&+[2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)-4(1-p)r^2q(r-p)+2p^3r^2]\mu(WWWDL)\\
&+[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L)+[2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)\\
&-2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD}\\
=\;&2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(DWWWD)\underbrace{+2qpr^2\mu(WWWDL)-2pr^2(1+2q+p)\mu(LDWWW)}_{\text{terms involving }\mu(WWWDL)}\\
&-2p^2r^2\mu(WWW\widehat{**}L)+[2pqr^2\{5p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)\\
&-(6pqr^2-4p^2qr^2)\mu(WWWWD)+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)\\
&-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&\underbrace{+[2p^2r^2(1-p)+2p^2r^3(1-p)+4p^3r^3(1-p)-4(1-p)r^2q(r-p)+2p^3r^2]\mu(WWWDL)}_{\text{term involving }\mu(WWWDL)}\\
&+[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L)+[2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)\\
&-2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD}\\
=\;&2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(DWWWD)-2p^2r^2\underbrace{\mu(WWW\widehat{**}L)}_{\text{split by (3)}}\\
&+[2pqr^2\{5p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)-(6pqr^2-4p^2qr^2)\mu(WWWWD)\\
&+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&\underbrace{-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL)-4(1-p)r^2q(r-p)\mu(WWWDL)}_{\text{summing terms involving }\mu(WWWDL)\text{ and using (3.4.42)}}\\
&+[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L)+[2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)
\end{aligned}
$$

$$
\begin{aligned}
&-2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD}\\
&=2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(DWWWD)\underbrace{-2p^2r^2\mu(WWWD\{W,D\}L)-2p^2r^2\mu(WWWWDL)}_{\text{to be added to underbraced terms below}}\\
&\quad+[2pqr^2\{5p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)\\
&\quad-(6pqr^2-4p^2qr^2)\mu(WWWWD)+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)\\
&\quad-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&\quad-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL)\\
&\quad-4(1-p)r^2q(r-p)\mu(WWWDL)\\
&\quad\underbrace{+[4p^3r^3(1-p)+2p^3r^2]\mu(WWWD\{W,D\}L)+[2p^3r^3(1-p)+2p^3r^2]\mu(WWWWDL)}_{\text{to be added to underbraced terms above}}\\
&\quad-2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD}\\
&=2qp^2r\mu(\widehat{***})-pr^2(1+2q)\mu(DWWWD)+[2pqr^2\{5p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)\\
&\quad-(6pqr^2-4p^2qr^2)\mu(WWWWD)+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)\\
&\quad-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&\quad-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL)\\
&\quad-4(1-p)r^2q(r-p)\mu(WWWDL)\\
&\quad\underbrace{-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)}_{\text{summing terms highlighted by underbraces in the previous step}}\\
&\quad-2pqr^2\mu(\widehat{***})-2prC_{LWD}-2pC_{LWWD}. \qquad (3.4.43)
\end{aligned}
$$

We now pause for a bit to write down explicitly $-2prC_{LWD}-2pC_{LWWD}$, where we borrow the mathematical expressions for C_{LWD} and C_{LWWD} from (3.4.30) and (3.4.36) derived in §3.4.5.2:

$$
\begin{aligned}
-2prC_{LWD}-2pC_{LWWD}=&-2qp^2r^2\mu(L\{W,D\}\widehat{***})-2pq(1-q)r^2\mu(L\widehat{***})-2qp^3r\mu(\widehat{***})\\
&-2qp^2r^2\mu(L\{W,D\}\widehat{***})-2pqr^2(1-q+p)\mu(L\widehat{***})\\
=&-4qp^2r^2\mu(L\{W,D\}\widehat{***})-2pqr^2(2-2q+p)\mu(L\widehat{***})-2qp^3r\mu(\widehat{***}).
\end{aligned}
\qquad (3.4.44)
$$

Substituting (3.4.44) in (3.4.43), we obtain:

$$
\begin{aligned}
B_2=&\underbrace{2qp^2r\mu(\widehat{***})}_{\text{term involving }\mu(\widehat{***})}-pr^2(1+2q)\mu(DWWWD)+[2pqr^2\{5p-2+p^2\}+6p^4r^2\\
&-4p^4r^2(p+q)]\mu(WWWD)-(6pqr^2-4p^2qr^2)\mu(WWWWD)
\end{aligned}
$$

$$
\begin{aligned}
&+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL)\\
&-4(1-p)r^2q(r-p)\mu(WWWDL)-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)\\
&-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)\underbrace{-2pqr^2\mu(\widehat{***})}_{\text{term involving }\mu(\widehat{***})}-4qp^2r^2\mu(L\{W,D\}\widehat{***})\\
&-2pqr^2(2-2q+p)\mu(L\widehat{***})\underbrace{-2qp^3r\mu(\widehat{***})}_{\text{term involving }\mu(\widehat{***})}\\
=\;&\underbrace{2pqr(p-r-p^2)\mu(\widehat{***})}_{\text{adding terms involving }\mu(\widehat{***})}-pr^2(1+2q)\mu(DWWWD)\\
&+[2pqr^2\{5p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)-(6pqr^2-4p^2qr^2)\mu(WWWWD)\\
&+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL)\\
&-4(1-p)r^2q(r-p)\mu(WWWDL)-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)\\
&-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)-4qp^2r^2\quad\underbrace{\mu(L\{W,D\}\widehat{***})}_{\text{split into }\mu(LWWWD)\text{ and the rest}}\\
&-2pqr^2(2-2q+p)\mu(L\widehat{***})\\
=\;&2pqr(p-r-p^2)\mu(\widehat{***})\underbrace{-pr^2(1+2q)\mu(DWWWD)}\\
&+[2pqr^2\{5p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)\underbrace{-(6pqr^2-4p^2qr^2)\mu(WWWWD)}\\
&+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL)\\
&-4(1-p)r^2q(r-p)\mu(WWWDL)-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)\\
&-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)\underbrace{-4qp^2r^2\mu(LWWWD)}-4qp^2r^2[\mu(L\{W,D\}\widehat{***})\\
&-\mu(LWWWD)]-2pqr^2(2-2q+p)\mu(L\widehat{***}).
\end{aligned}
\tag{3.4.45}
$$

We show here how we combine the terms highlighted by underbraces in the last step of (3.4.45):

$$
\begin{aligned}
&-pr^2(1+2q)\mu(DWWWD)-(6pqr^2-4p^2qr^2)\mu(WWWWD)-4qp^2r^2\mu(LWWWD)\\
=\;&\underbrace{-2p^2qr^2\mu(DWWWD)-pr^2\{1+2q(1-p)\}\mu(DWWWD)}_{\text{splitting the term involving }\mu(DWWWD)}\\
&\underbrace{-2p^2qr^2\mu(WWWWD)-6pqr^2(1-p)\mu(WWWWD)}_{\text{splitting the term involving }\mu(WWWWD)}-4p^2qr^2\mu(LWWWD)
\end{aligned}
$$

$$
\begin{aligned}
&= \underbrace{-2p^2qr^2\mu(DWWWD) - 2p^2qr^2\mu(WWWWD) - 2p^2qr^2\mu(LWWWD)}_{} \\
&\quad - pr^2\{1+2q(1-p)\}\mu(DWWWD) - 6pqr^2(1-p)\mu(WWWWD) - 2p^2qr^2\mu(LWWWD) \\
&= -2p^2qr^2\mu(WWWD) - pr^2\{1+2q(1-p)\}\mu(DWWWD) - 6pqr^2(1-p)\mu(WWWWD) \\
&\quad - 2p^2qr^2\mu(LWWWD).
\end{aligned}
\tag{3.4.46}
$$

Substituting (3.4.46) in (3.4.45) yields (we have highlighted the terms that have come from (3.4.46) using underbraces in the very first step of the computation below):

$$
\begin{aligned}
B_2 &= 2pqr(p-r-p^2)\mu(\widehat{\ast\ast\ast}) \underbrace{-2p^2qr^2\mu(WWWD) - pr^2\{1+2q(1-p)\}\mu(DWWWD)}_{} \\
&\quad \underbrace{-6pqr^2(1-p)\mu(WWWWD)}_{} + [2pqr^2\{5p-2+p^2\} + 6p^4r^2 - 4p^4r^2(p+q)]\mu(WWWD) \\
&\quad + 2p^2r^2(1-p)(1-pr)\mu(WWWWWD) - 2p^3r^3(1-p)\{\mu(LWWWWD) + \mu(DWWWWD)\} \\
&\quad - 2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL) \\
&\quad - 4(1-p)r^2q(r-p)\mu(WWWDL) - 2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L) \\
&\quad - 2p^2r^2(1-p)(1-pr)\mu(WWWWDL) \underbrace{-2p^2qr^2\mu(LWWWD)}_{} \\
&\quad - 4qp^2r^2[\mu(L\{W,D\}\widehat{\ast\ast\ast}) - \mu(LWWWD)] - 2pqr^2(2-2q+p)\mu(L\widehat{\ast\ast\ast}) \\
&= 2pqr(p-r-p^2)\mu(\widehat{\ast\ast\ast}) \underbrace{-2p^2qr^2\mu(WWWD)}_{} - pr^2\{1+2q(1-p)\}\mu(DWWWD) \\
&\quad - 6pqr^2(1-p)\mu(WWWWD) \underbrace{+[2pqr^2\{5p-2+p^2\} + 6p^4r^2 - 4p^4r^2(p+q)]\mu(WWWD)}_{} \\
&\quad + 2p^2r^2(1-p)(1-pr)\mu(WWWWWD) - 2p^3r^3(1-p)\{\mu(LWWWWD) + \mu(DWWWWD)\} \\
&\quad - 2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL) \\
&\quad - 4(1-p)r^2q(r-p)\mu(WWWDL) - 2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L) \\
&\quad - 2p^2r^2(1-p)(1-pr)\mu(WWWWDL) - 2p^2qr^2\mu(LWWWD) \\
&\quad - 4qp^2r^2[\mu(L\{W,D\}\widehat{\ast\ast\ast}) - \mu(LWWWD)] - 2pqr^2(2-2q+p)\mu(L\widehat{\ast\ast\ast}) \\
&= 2pqr(p-r-p^2)\mu(\widehat{\ast\ast\ast}) - pr^2\{1+2q(1-p)\}\mu(DWWWD) - 6pqr^2(1-p)\mu(WWWWD) \\
&\quad \underbrace{+[2pqr^2\{4p-2+p^2\} + 6p^4r^2 - 4p^4r^2(p+q)]\mu(WWWD)}_{\text{summing terms involving } \mu(WWWD)} \\
&\quad + 2p^2r^2(1-p)(1-pr)\mu(WWWWWD) - 2p^3r^3(1-p)\{\mu(LWWWWD) + \mu(DWWWWD)\} \\
&\quad - 2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL) \\
&\quad - 4(1-p)r^2q(r-p)\mu(WWWDL) - 2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L) \\
&\quad - 2p^2r^2(1-p)(1-pr)\mu(WWWWDL) - 2p^2qr^2\mu(LWWWD) - 4qp^2r^2[\mu(L\{W,D\}\widehat{\ast\ast\ast})
\end{aligned}
$$

$$-\mu(LWWWD)]-2pqr^2(2-2q+p)\mu(L\widehat{***}). \quad (3.4.47)$$

The above, i.e. the last step of (3.4.47), serves as our final, simplified expression for B_2. From (3.4.38), (3.4.39) and (3.4.47), we see that our weight function inequality has now transformed into:

$$\begin{aligned}
&w_2(\widehat{\mathrm{F}}_{p,q}\mu) \leqslant w_2(\mu)-[p(1-r)+q]\mu(LWD)-[2q+r(4q^2+4qp^2-2q^3+2p^2+2q^2p)]\mu(LWWD)\\
&-2r(1-p^2)\mu(LDDL)-r(1-p)\mu(LDL)-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDDW)\\
&+\mu(LDWD)+\mu(LWDD)]-r(4q^2+4qp^2-2q^3+2p^2+2q^2p)[\mu(LDWW)+\mu(LWDW)]-q\mu(D)\\
&-[p(1-r)+q][\mu(WD)+\mu(WWD)]-pr^2(1+2q)\mu(L\widehat{***}L)\\
&\underbrace{+2pqr(p-r-p^2)\mu(\widehat{***})}-pr^2\{1+2q(1-p)\}\mu(DWWWD)-6pqr^2(1-p)\mu(WWWWD)\\
&\underbrace{+[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)}\\
&\underbrace{+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)}-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL)-4(1-p)r^2q(r-p)\mu(WWWDL)\\
&-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)-2p^2qr^2\mu(LWWWD)\\
&-4qp^2r^2[\mu(L\{W,D\}\widehat{***})-\mu(LWWWD)]-2pqr^2(2-2q+p)\mu(L\widehat{***}). \quad (3.4.48)
\end{aligned}$$

The underbraces in the right side of (3.4.48) are intended to highlight the (possibly, depending on the values of p and q) non-negative terms that still remain. It is worthwhile to note, comparing (3.4.48) with (3.4.40), that we have already taken care of the (previously non-negative) terms involving $\mu(WWWWD)$, $\mu(WWWDL)$, $\mu(WWWD\{W,D\}L)$ and $\mu(WWWWDL)$, via the long and tedious algebraic manipulations that ultimately lead to (3.4.47).

3.4.6 The fourth step of composing the weight function

The third adjustment is carried out as follows:

$$w_3(\mu)=w_2(\mu)-2(q+p^2r)\mu(LWWD)-2p^2r\{\mu(LDWW)+\mu(LWDW)\}-q\mu(D). \quad (3.4.49)$$

Writing (3.4.48) as $w_2(\widehat{\mathrm{F}}_{p,q}\mu)\leqslant w_2(\mu)+A_2$, implementing the adjustment described in (3.4.49), and using the idea presented in (3.4.16), we obtain the weight function inequality

$$w_3(\widehat{\mathrm{F}}_{p,q}\mu)\leqslant w_3(\mu)+2(q+p^2r)\mu(LWWD)+2p^2r\{\mu(LDWW)+\mu(LWDW)\}+q\mu(D)$$

$$
\begin{aligned}
&-2(q+p^2r)\widehat{\mathrm{F}}_{p,q}\,\mu(LWWD)-2p^2r\{\widehat{\mathrm{F}}_{p,q}\,\mu(LDWW)+\widehat{\mathrm{F}}_{p,q}\,\mu(LWDW)\}\\
&-q\widehat{\mathrm{F}}_{p,q}\,\mu(D)+A_2.
\end{aligned}
\tag{3.4.50}
$$

As noted using underbraces in (3.4.48), we have to somehow negate the non-negative terms involving $\mu(WWWD)$, $\mu(WWWWWD)$ and $\mu(\widehat{***})$ using existing non-positive terms on the right side of (3.4.48).

1. As seen from (3.4.52), one of the terms in the expansion of $\widehat{\mathrm{F}}_{p,q}\,\mu(LDWW)$ involves $\mu(WWWD)$;

2. as seen from (3.4.53), one of the terms in the expansion of $\widehat{\mathrm{F}}_{p,q}\,\mu(LWDW)$ involves $\mu(WWW\widehat{**})$, which can be written as $\mu(WWWD)+\mu(WWWWD)-\mu(WWWDL)$;

3. and finally, one of the terms in the expansion of $\widehat{\mathrm{F}}_{p,q}\,\mu(LWWD)$, as seen in (3.4.6), involves $\mu(WWW\widehat{***})$, which can be written as $\mu(WWW\widehat{***}) = \mu(WWWD)+\mu(WWWWD)+\mu(WWWWWD)-\mu(WWWDL)-\mu(WWW\widehat{**}L)$.

Therefore, from (3.4.50), it can be hoped that $-2(q+p^2r)\widehat{\mathrm{F}}_{p,q}\,\mu(LWWD)-2p^2r\{\widehat{\mathrm{F}}_{p,q}\,\mu(LDWW)+\widehat{\mathrm{F}}_{p,q}\,\mu(LWDW)\}$ will aid in negating the non-negative terms involving $\mu(WWWD)$ and $\mu(WWWWD)$ highlighted in (3.4.48) using underbraces. From (3.4.4), we see that $\widehat{\mathrm{F}}_{p,q}\,\mu(D)=r\mu(\widehat{***})$, so that $-q\widehat{\mathrm{F}}_{p,q}\,\mu(D)$ may aid in negating the non-negative term involving $\mu(\widehat{***})$ highlighted in (3.4.48) using underbraces.

At the same time, we have to make sure that we can afford to introduce the non-negative terms $2(q+p^2r)\mu(LWWD)+2p^2r\{\mu(LDWW)+\mu(LWDW)\}+q\mu(D)$ in (3.4.50) without turning any of the existing non-positive terms in the right side of (3.4.48) non-negative:

1. the term $-[2q+r(4q^2+4qp^2-2q^3+2p^2+2q^2p)]\mu(LWWD)$ in the right side of (3.4.48) provides us with $-2(q+p^2r)\mu(LWWD)$,

2. whereas the term $-r(4q^2+4qp^2-2q^3+2p^2+2q^2p)[\mu(LDWW)+\mu(LWDW)]$ in the right side of (3.4.48) supplies us with $-2p^2r\{\mu(LDWW)+\mu(LWDW)\}$;

3. finally, the existing $-q\mu(D)$ term in the right side of (3.4.48) cancels out the $+q\mu(D)$ term.

We hope that the above two paragraphs serve to motivate the choice of our adjustment described in (3.4.49).

3.4.6.1 Step 1 of analyzing the effect of the adjustment in (3.4.49)

Incorporating (3.4.49) into (3.4.48), as illustrated in (3.4.50), we obtain (in any given step, we highlight, using underbraces, terms that are to be combined / manipulated algebraically to obtain the next step):

$$
\begin{aligned}
& w_3(\widehat{\mathrm{F}}_{p,q}\mu) \leqslant w_3(\mu) + 2(q+p^2r)\mu(LWWD) + 2p^2r\{\mu(LDWW)+\mu(LWDW)\} + q\mu(D) \\
& -[p(1-r)+q]\mu(LWD) - [2q+r(4q^2+4qp^2-2q^3+2p^2+2q^2p)]\mu(LWWD) - 2r(1-p^2)\mu(LDDL) \\
& -r(1-p)\mu(LDL) - r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDDW)+\mu(LDWD)+\mu(LWDD)] \\
& -r(4q^2+4qp^2-2q^3+2p^2+2q^2p)[\mu(LDWW)+\mu(LWDW)] - q\mu(D) \\
& -[p(1-r)+q][\mu(WD)+\mu(WWD)] - pr^2(1+2q)\mu(L\widehat{***}L) + 2pqr(p-r-p^2)\mu(\widehat{***}) \\
& -pr^2\{1+2q(1-p)\}\mu(DWWWD) - 6pqr^2(1-p)\mu(WWWWD) \\
& +[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD) \\
& +2p^2r^2(1-p)(1-pr)\mu(WWWWWD) - 2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\} \\
& -2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL) \\
& -4(1-p)r^2q(r-p)\mu(WWWDL) - 2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L) \\
& -2p^2r^2(1-p)(1-pr)\mu(WWWWDL) - 2p^2qr^2\mu(LWWWD) \\
& -4qp^2r^2[\mu(L\{W,D\}\widehat{***}) - \mu(LWWWD)] - 2pqr^2(2-2q+p)\mu(L\widehat{***}) \\
& -2(q+p^2r)\widehat{\mathrm{F}}_{p,q}\mu(LWWD) - 2p^2r\{\widehat{\mathrm{F}}_{p,q}\mu(LDWW)+\widehat{\mathrm{F}}_{p,q}\mu(LWDW)\} - q\widehat{\mathrm{F}}_{p,q}\mu(D) \\
= {} & w_3(\mu) - [p(1-r)+q]\mu(LWD) \underbrace{+2(q+p^2r)\mu(LWWD) + 2p^2r\{\mu(LDWW)+\mu(LWDW)\} + q\mu(D)} \\
& \underbrace{-[2q+r(4q^2+4qp^2-2q^3+2p^2+2q^2p)]\mu(LWWD)} -2r(1-p^2)\mu(LDDL) - r(1-p)\mu(LDL) \\
& \underbrace{-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDDW)+\mu(LDWD)+\mu(LWDD)]} \\
& \underbrace{-r(4q^2+4qp^2-2q^3+2p^2+2q^2p)[\mu(LDWW)+\mu(LWDW)] - q\mu(D)} \\
& -[p(1-r)+q][\mu(WD)+\mu(WWD)] - pr^2(1+2q)\mu(L\widehat{***}L) \\
& +2pqr(p-r-p^2)\mu(\widehat{***}) - pr^2\{1+2q(1-p)\}\mu(DWWWD) - 6pqr^2(1-p)\mu(WWWWD) \\
& +[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD) \\
& +2p^2r^2(1-p)(1-pr)\mu(WWWWWD) - 2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\} \\
& -2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL) \\
& -4(1-p)r^2q(r-p)\mu(WWWDL) - 2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L) \\
& -2p^2r^2(1-p)(1-pr)\mu(WWWWDL) - 2p^2qr^2\mu(LWWWD) - 4qp^2r^2[\mu(L\{W,D\}\widehat{***})
\end{aligned}
$$

$$
\begin{aligned}
&-\mu(LWWWD)]-2pqr^2(2-2q+p)\mu(L\widehat{***})-2(q+p^2r)\widehat{\mathrm{F}}_{p,q}\,\mu(LWWD)\\
&-2p^2r\{\widehat{\mathrm{F}}_{p,q}\,\mu(LDWW)+\widehat{\mathrm{F}}_{p,q}\,\mu(LWDW)\}-q\widehat{\mathrm{F}}_{p,q}\,\mu(D)\\
=\;& w_3(\mu)-[p(1-r)+q]\mu(LWD)\underbrace{-r(4q^2+4qp^2-2q^3+2q^2p)\mu(L\widehat{***})}_{\text{combining all terms indicated by underbraces above}}-2r(1-p^2)\mu(LDDL)\\
&-r(1-p)\mu(LDL)-[p(1-r)+q][\mu(WD)+\mu(WWD)]-pr^2(1+2q)\mu(L\widehat{***}L)\\
&+2pqr(p-r-p^2)\mu(\widehat{***})-pr^2\{1+2q(1-p)\}\mu(DWWWD)-6pqr^2(1-p)\mu(WWWWD)\\
&+[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)\\
&+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL)\\
&-4(1-p)r^2q(r-p)\mu(WWWDL)-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)\\
&-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)-2p^2qr^2\mu(LWWWD)\\
&-4qp^2r^2[\mu(L\{W,D\}\widehat{***})-\mu(LWWWD)]-2pqr^2(2-2q+p)\mu(L\widehat{***})\\
&\underbrace{-2(q+p^2r)\widehat{\mathrm{F}}_{p,q}\,\mu(LWWD)}\underbrace{-2p^2r\{\widehat{\mathrm{F}}_{p,q}\,\mu(LDWW)+\widehat{\mathrm{F}}_{p,q}\,\mu(LWDW)\}-q\widehat{\mathrm{F}}_{p,q}\,\mu(D)}.
\end{aligned}
\tag{3.4.51}
$$

3.4.6.2 Step 2 of analyzing the effect of the adjustment in (3.4.49)

Our next task is to compute the sum of the *last three terms*, highlighted by underbraces, in the last step of (3.4.51). To this end, we need to compute $\widehat{\mathrm{F}}_{p,q}\,\mu(LDWW)$ and $\widehat{\mathrm{F}}_{p,q}\,\mu(LWDW)$. As was the case for the computations carried out in §3.4.5.2, we are only concerned with parts of the expressions for these probabilities. While computing $\widehat{\mathrm{F}}_{p,q}\,\mu(LDWW)$, we consider only those cases in which $\eta(0)=\eta(1)=\eta(2)=W$ leads to the event that $(\widehat{\mathrm{F}}_{p,q}\,\eta(0),\widehat{\mathrm{F}}_{p,q}\,\eta(1),\widehat{\mathrm{F}}_{p,q}\,\eta(2),\widehat{\mathrm{F}}_{p,q}\,\eta(3))=(LDWW)$. Likewise, while computing $\widehat{\mathrm{F}}_{p,q}\,\mu(LWDW)$, we consider only those cases in which $\eta(0)=\eta(1)=\eta(2)=W$ leads to the event that $(\widehat{\mathrm{F}}_{p,q}\,\eta(0),\widehat{\mathrm{F}}_{p,q}\,\eta(1),\widehat{\mathrm{F}}_{p,q}\,\eta(2),\widehat{\mathrm{F}}_{p,q}\,\eta(3))=(LWDW)$. These considerations lead to

$$
\begin{aligned}
\bullet\,&\widehat{\mathrm{F}}_{p,q}\,\mu(LDWW)=(1-p)rp^2\mu(WWWD)+(1-p)r^2p\mu(WWWD\{W,D\}L)\\
&+(1-p)r^2(1+p-q)\mu(WWWDL)+C_{LDWW}\geqslant(1-p)rp^2\mu(WWWD)\\
&+(1-p)r^2p\mu(WWWD\{W,D\}L)+(1-p)r^2(1+p-q)\mu(WWWDL);
\end{aligned}
\tag{3.4.52}
$$

$$
\begin{aligned}
\bullet\,&\widehat{\mathrm{F}}_{p,q}\,\mu(LWDW)=(1-p)p^2r\mu(WWW\widehat{**})+(1-p)pr^2\mu(WWW\widehat{**}L)+C_{LWDW}\\
&\geqslant(1-p)p^2r\mu(WWW\widehat{**})+(1-p)pr^2\mu(WWW\widehat{**}L);
\end{aligned}
\tag{3.4.53}
$$

where C_{LDWW} and C_{LWDW} are the respective contributions from the cases where $(\eta(0),\eta(1),\eta(2)) \in \hat{\mathscr{A}}^3 \setminus \{(W,W,W)\}$. Finally, we rewrite (3.4.6) as

$$\widehat{\mathrm{F}}_{p,q}\,\mu(LWWD) = (1-p)p^2r\mu(WWW\widehat{***}) + D_{LWWD} \geqslant (1-p)p^2r\mu(WWW\widehat{***}), \quad (3.4.54)$$

where $D_{LWWD} = qp^2r\mu(\widehat{***}\widehat{***}) + qp^2r\mu(L\{W,D\}^2\widehat{***}) + q(1-q)pr\mu(L\{W,D\}\widehat{***}) + q(1-q)^2r\mu(L\widehat{***})$.

Using (3.4.4), (3.4.54), (3.4.52) and (3.4.53), and applying the identities

1. $\mu(WWW\widehat{***}) = \mu(WWWD) + \mu(WWWWD) + \mu(WWWWWD) - \mu(WWWDL) - \mu(WWWD\{W,D\}L) - \mu(WWWWDL)$,
2. $\mu(WWW\widehat{**}) = \mu(WWWD) - \mu(WWWDL) + \mu(WWWWD)$,
3. $\mu(WWW\widehat{**}L) = \mu(WWWD\{W,D\}L) + \mu(WWWWDL)$,

the sum of the last three terms, highlighted by underbraces, in (3.4.51) simplifies to

$$\begin{aligned}
&-2(q+p^2r)\widehat{\mathrm{F}}_{p,q}\,\mu(LWWD) - 2p^2r\{\widehat{\mathrm{F}}_{p,q}\,\mu(LDWW) + \widehat{\mathrm{F}}_{p,q}\,\mu(LWDW)\} - q\widehat{\mathrm{F}}_{p,q}\,\mu(D)\\
\leqslant\; &-2(q+p^2r)(1-p)p^2r\underbrace{\mu(WWW\widehat{***})}_{\text{use (1)}} - 2p^2r\{(1-p)rp^2\mu(WWWD) + (1-p)r^2p\mu(WWWD\{W,D\}L)\\
&+(1-p)r^2(1+p-q)\mu(WWWDL)\} - 2p^2r\{(1-p)p^2r\underbrace{\mu(WWW\widehat{**})}_{\text{use (2)}} + (1-p)pr^2\underbrace{\mu(WWW\widehat{**}L)}_{\text{use (3)}}\}\\
&-qr\mu(\widehat{***})\\
=\; &-2(q+p^2r)(1-p)p^2r[\underbrace{\mu(WWWD)+\mu(WWWWD)+\mu(WWWWWD)-\mu(WWWDL)}_{\text{after applying (1)}}\\
&\underbrace{-\mu(WWWD\{W,D\}L)-\mu(WWWWDL)}_{\text{after applying (1)}}] - 2p^4r^2(1-p)\mu(WWWD)\\
&-2p^3r^3(1-p)\mu(WWWD\{W,D\}L) - 2p^2r^3(1-p)(1+p-q)\mu(WWWDL)\\
&-2p^4r^2(1-p)[\underbrace{\mu(WWWD)-\mu(WWWDL)+\mu(WWWWD)}_{\text{after applying (2)}}]\\
&-2p^3r^3(1-p)[\underbrace{\mu(WWWD\{W,D\}L)+\mu(WWWWDL)}_{\text{after applying (3)}}]\\
&-qr\mu(\widehat{***})\\
=\; &-\big[2(q+p^2r)(1-p)p^2r + 2p^4r^2(1-p) + 2p^4r^2(1-p)\big]\mu(WWWD)\\
&-\big[2(q+p^2r)(1-p)p^2r + 2p^4r^2(1-p)\big]\mu(WWWWD) - 2(q+p^2r)(1-p)p^2r\mu(WWWWWD)
\end{aligned}$$

$$
\begin{aligned}
&+\big[2(q+p^2r)(1-p)p^2r-2p^2r^3(1-p)(1+p-q)+2p^4r^2(1-p)\big]\mu(WWWDL)\\
&+\big[2(q+p^2r)(1-p)p^2r-2p^3r^3(1-p)-2p^3r^3(1-p)\big]\mu(WWWD\{W,D\}L)\\
&+\big[2(q+p^2r)(1-p)p^2r-2p^3r^3(1-p)\big]\mu(WWWWDL)-qr\mu(\widehat{***})\\
&\text{(adding all terms involving } \mu(\mathscr{C}) \text{ for } \mathscr{C} \text{ any of } (WWWD),(WWWWD),(WWWWWD),\\
&(WWWDL),(WWWD\{W,D\}L),(WWWWDL))\\
=&-[2qp^2r(1-p-q+q)+6p^4r^2(1-p)]\mu(WWWD)-[2qp^2r(1-p)+4p^4r^2(1-p)]\mu(WWWWD)\\
&-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD)+[2qp^2r(1-p)+4p^4r^2(1-p)\\
&-2p^2r^3(1-p)(1+p-q)]\mu(WWWDL)+[2qp^2r(1-p)+2p^4r^2(1-p)\\
&-4p^3r^3(1-p)]\mu(WWWD\{W,D\}L)+[2qp^2r(1-p)+2p^4r^2(1-p)\\
&-2p^3r^3(1-p)]\mu(WWWWDL)-qr\mu(\widehat{***})\\
=&-[2qp^2r^2+2q^2p^2r+6p^4r^2(1-p)]\mu(WWWD)-[2qp^2r(1-p)+4p^4r^2(1-p)]\mu(WWWWD)\\
&-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD)+[2qp^2r(1-p)+4p^4r^2(1-p)\\
&-2p^2r^3(1-p)(1+p-q)]\mu(WWWDL)+[2qp^2r(1-p)+2p^4r^2(1-p)\\
&-4p^3r^3(1-p)]\mu(WWWD\{W,D\}L)+[2qp^2r(1-p)+2p^4r^2(1-p)\\
&-2p^3r^3(1-p)]\mu(WWWWDL)-qr\mu(\widehat{***}). \qquad (3.4.55)
\end{aligned}
$$

3.4.6.3 Step 3 of analyzing the effect of the adjustment in (3.4.49)

We begin with the following observation:

$$
\begin{aligned}
2p(p-r-p^2) &= 2p\{p-(1-p-q)-p^2\} = 2p\{2p-1+q-p^2\}\\
&= 2p\{q-(1-p)^2\} \leqslant 2p\{(1-p)-(1-p)^2\} \quad \text{(by (3.1.1))};\\
&\leqslant 2p^2(1-p) \leqslant 2p(1-p) \leqslant \frac{1}{2} < 1 \text{ for all } p \in [0,1]. \qquad (3.4.56)
\end{aligned}
$$

This leads to the conclusion that

$$
2pqr(p-r-p^2) \leqslant qr \text{ for all } (p,q) \in \mathscr{S}, \text{ with } \mathscr{S} \text{ as defined in (3.1.1).} \qquad (3.4.57)
$$

Incorporating (3.4.55) into (3.4.51), and writing $\mu(L\widehat{***})$ as the sum of its components, i.e. $\mu(LDDD), \ldots, \mu(LWWD)$, yields (once again, in each step of the computation below, we highlight using underbraces the terms that are to be manipulated to obtain the next step):

$$
w_3(\widehat{\mathrm{F}}_{p,q}\mu) \leqslant w_3(\mu)-[p(1-r)+q]\mu(LWD)-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDWD)
$$

$$
\begin{aligned}
&+\mu(LWDD)+\mu(LWWD)]-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDWW)+\mu(LWDW)+\mu(LDDW)]\\
&-2r(1-p^2)\mu(LDDL)-r(1-p)\mu(LDL)-[p(1-r)+q][\mu(WD)+\mu(WWD)]\\
&-pr^2(1+2q)\mu(L\widehat{***}L)\underbrace{+2pqr(p-r-p^2)\mu(\widehat{***})}-pr^2\{1+2q(1-p)\}\mu(DWWWD)\\
&-6pqr^2(1-p)\mu(WWWWD)+[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)\\
&+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL)\\
&-4(1-p)r^2q(r-p)\mu(WWWDL)-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)\\
&-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)-2p^2qr^2\mu(LWWWD)\\
&-4qp^2r^2[\mu(L\{W,D\}\widehat{***})-\mu(LWWWD)]-2pqr^2(2-2q+p)\mu(L\widehat{***})\\
&-[2qp^2r^2+2q^2p^2r+6p^4r^2(1-p)]\mu(WWWD)-[2qp^2r(1-p)+4p^4r^2(1-p)]\mu(WWWWD)\\
&-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD)+[2qp^2r(1-p)+4p^4r^2(1-p)\\
&-2p^2r^3(1-p)(1+p-q)]\mu(WWWDL)+[2qp^2r(1-p)+2p^4r^2(1-p)\\
&-4p^3r^3(1-p)]\mu(WWWD\{W,D\}L)+[2qp^2r(1-p)+2p^4r^2(1-p)-2p^3r^3(1-p)]\mu(WWWWDL)\\
&\underbrace{-qr\mu(\widehat{***})}\\
={}&w_3(\mu)-[p(1-r)+q]\mu(LWD)-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDWD)\\
&+\mu(LWDD)+\mu(LWWD)]-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDWW)+\mu(LWDW)+\mu(LDDW)]\\
&-2r(1-p^2)\mu(LDDL)-r(1-p)\mu(LDL)-[p(1-r)+q][\mu(WD)+\mu(WWD)]\\
&-pr^2(1+2q)\mu(L\widehat{***}L)\underbrace{-qr[1-2p(p-r-p^2)]\mu(\widehat{***})}_{\text{non-positive by (3.4.57)}}-pr^2\{1+2q(1-p)\}\mu(DWWWD)\\
&-6pqr^2(1-p)\mu(WWWWD)+[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)\\
&+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL)\\
&-4(1-p)r^2q(r-p)\mu(WWWDL)-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)\\
&-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)-2p^2qr^2\mu(LWWWD)-4qp^2r^2[\mu(L\{W,D\}\widehat{***})\\
&-\mu(LWWWD)]-2pqr^2(2-2q+p)\mu(L\widehat{***})-[2qp^2r^2+2q^2p^2r+6p^4r^2(1-p)]\mu(WWWD)\\
&-[2qp^2r(1-p)+4p^4r^2(1-p)]\mu(WWWWD)-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD)\\
&+[2qp^2r(1-p)+4p^4r^2(1-p)-2p^2r^3(1-p)(1+p-q)]\mu(WWWDL)\\
&+[2qp^2r(1-p)+2p^4r^2(1-p)-4p^3r^3(1-p)]\mu(WWWD\{W,D\}L)\\
&+[2qp^2r(1-p)+2p^4r^2(1-p)-2p^3r^3(1-p)]\mu(WWWWDL)\\
={}&w_3(\mu)-[p(1-r)+q]\mu(LWD)-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDWD)
\end{aligned}
$$

$$
\begin{aligned}
&+\mu(LWDD)+\mu(LWWD)]\underbrace{-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDWW)+\mu(LWDW)+\mu(LDDW)]}\\
&-2r(1-p^2)\mu(LDDL)-r(1-p)\mu(LDL)\underbrace{-[p(1-r)+q][\mu(WD)+\mu(WWD)]}\\
&-pr^2(1+2q)\mu(L\widehat{\ast\ast\ast}L)-qr[1-2p(p-r-p^2)]\mu(\widehat{\ast\ast\ast})-pr^2\{1+2q(1-p)\}\mu(DWWWD)\\
&-6pqr^2(1-p)\mu(WWWWD)\underbrace{+[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)}\\
&\underbrace{+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)}-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}\\
&\underbrace{-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)-p^3\}\mu(WWWDL)-4(1-p)r^2q(r-p)\mu(WWWDL)}\\
&\underbrace{-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)}\\
&-2p^2qr^2\mu(LWWWD)-4qp^2r^2[\mu(L\{W,D\}\widehat{\ast\ast\ast})-\mu(LWWWD)]-2pqr^2(2-2q+p)\mu(L\widehat{\ast\ast\ast})\\
&\underbrace{-[2qp^2r^2+2q^2p^2r+6p^4r^2(1-p)]\mu(WWWD)-[2qp^2r(1-p)+4p^4r^2(1-p)]\mu(WWWWD)}\\
&\underbrace{-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD)+[2qp^2r(1-p)+4p^4r^2(1-p)}\\
&\underbrace{-2p^2r^3(1-p)(1+p-q)]\mu(WWWDL)+[2qp^2r(1-p)+2p^4r^2(1-p)}\\
&\underbrace{-4p^3r^3(1-p)]\mu(WWWD\{W,D\}L)+[2qp^2r(1-p)+2p^4r^2(1-p)-2p^3r^3(1-p)]\mu(WWWWDL)}.
\end{aligned}
\tag{3.4.58}
$$

We collect all the terms that have been highlighted in the *last step* of (3.4.58) using underbraces, and denote their sum by C_2, whereas the sum of the remaning terms is denoted C_1. In other words,

$$
\begin{aligned}
C_1 = {} & w_3(\mu)-[p(1-r)+q]\mu(LWD)-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDWD)+\mu(LWDD)\\
&+\mu(LWWD)]-2r(1-p^2)\mu(LDDL)-r(1-p)\mu(LDL)-pr^2(1+2q)\mu(L\widehat{\ast\ast\ast}L)\\
&-qr[1-2p(p-r-p^2)]\mu(\widehat{\ast\ast\ast})-pr^2\{1+2q(1-p)\}\mu(DWWWD)-6pqr^2(1-p)\mu(WWWWD)\\
&-2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}-2p^2qr^2\mu(LWWWD)\\
&-4qp^2r^2[\mu(L\{W,D\}\widehat{\ast\ast\ast})-\mu(LWWWD)]-2pqr^2(2-2q+p)\mu(L\widehat{\ast\ast\ast}),
\end{aligned}
\tag{3.4.59}
$$

and

$$
\begin{aligned}
C_2 = {} & -r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDWW)+\mu(LWDW)+\mu(LDDW)]-[p(1-r)+q][\mu(WD)\\
&+\mu(WWD)]+[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)\\
&+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)\\
&-p^3\}\mu(WWWDL)-4(1-p)r^2q(r-p)\mu(WWWDL)-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)
\end{aligned}
$$

$$-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)-[2qp^2r^2+2q^2p^2r+6p^4r^2(1-p)]\mu(WWWD)$$
$$-[2qp^2r(1-p)+4p^4r^2(1-p)]\mu(WWWWD)-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD)$$
$$+[2qp^2r(1-p)+4p^4r^2(1-p)-2p^2r^3(1-p)(1+p-q)]\mu(WWWDL)+[2qp^2r(1-p)$$
$$+2p^4r^2(1-p)-4p^3r^3(1-p)]\mu(WWWD\{W,D\}L)+[2qp^2r(1-p)+2p^4r^2(1-p)$$
$$-2p^3r^3(1-p)]\mu(WWWWDL). \tag{3.4.60}$$

It is worthwhile to note that all terms in C_1, apart from $w_3(\mu)$, are non-positive.

Our task, now, is to see if we can turn each of the non-negative terms in C_2 non-positive by making use of the existing non-positive terms in (3.4.60). In the long computation that follows, the terms that are being dealt with in each step will be highlighted using underbraces. First, we note that, in

$$\begin{aligned}
C_2 = & -r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDWW)+\mu(LWDW)+\mu(LDDW)]-[p(1-r)+q][\mu(WD) \\
& +\mu(WWD)]+[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD) \\
& +2p^2r^2(1-p)(1-pr)\mu(WWWWWD)-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p) \\
& -p^3\}\mu(WWWDL)-4(1-p)r^2q(r-p)\mu(WWWDL)-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L) \\
& -2p^2r^2(1-p)(1-pr)\mu(WWWWDL)-[2qp^2r^2+2q^2p^2r+6p^4r^2(1-p)]\mu(WWWD) \\
& -[2qp^2r(1-p)+4p^4r^2(1-p)]\mu(WWWWD)-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD) \\
& +[2qp^2r(1-p)+4p^4r^2(1-p)-2p^2r^3(1-p)(1+p-q)]\mu(WWWDL)+[2qp^2r(1-p) \\
& +2p^4r^2(1-p)-4p^3r^3(1-p)]\mu(WWWD\{W,D\}L)+[2qp^2r(1-p)+2p^4r^2(1-p) \\
& -2p^3r^3(1-p)]\mu(WWWWDL) \\
= & \underbrace{-r(4q^2+4qp^2-2q^3+2q^2p)\mu(LDWW)-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)}_{A_1} \\
& \underbrace{-p^3\}\mu(WWWDL)-4(1-p)r^2q(r-p)\mu(WWWDL)+[2qp^2r(1-p)+4p^4r^2(1-p)}_{A_1} \\
& \underbrace{-2p^2r^3(1-p)(1+p-q)]\mu(WWWDL)}_{A_1}\underbrace{-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LWDW)+\mu(LDDW)]}_{A_2} \\
& \underbrace{-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)+[2qp^2r(1-p)+2p^4r^2(1-p)}_{A_2} \\
& \underbrace{-4p^3r^3(1-p)]\mu(WWWD\{W,D\}L)}_{A_2}\underbrace{-[2qp^2r(1-p)+4p^4r^2(1-p)]\mu(WWWWD)}_{A_3} \\
& \underbrace{+[2qp^2r(1-p)+2p^4r^2(1-p)-2p^3r^3(1-p)]\mu(WWWWDL)-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)}_{A_3}
\end{aligned}$$

$$
\begin{aligned}
&-[p(1-r)+q][\mu(WD)+\mu(WWD)]+[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)\\
&+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)-[2qp^2r^2+2q^2p^2r+6p^4r^2(1-p)]\mu(WWWD)\\
&-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD). \qquad (3.4.61)
\end{aligned}
$$

In what follows, we let A_i, for $i \in \{1,2,3\}$, denote the sum of the terms that have been highlighted in (3.4.61) using underbraces tagged A_i, and we simplify each A_i separately before incorporating these simplified expressions back into (3.4.61).

First, we simplify A_1, using the simple inequality $\mu(LDWW) \leqslant \mu(LDWWW)$, as follows:

$$
\begin{aligned}
A_1 &= \underbrace{-r(4q^2+4qp^2-2q^3+2q^2p)\mu(LDWW)}_{\text{split into two parts as shown below}}-2pr^2\{1+q-p+2p^2-2p^2r(1-p)+pq(1-p)\\
&\quad -p^3\}\mu(WWWDL)+\{2qp^2r(1-p)+4p^4r^2(1-p)-2p^2r^3(1-p)(1+p-q)\}\mu(WWWDL)\\
&\quad -4(1-p)r^2q(r-p)\mu(WWWDL)\\
&= \underbrace{-r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\mu(LDWW)-2qp^2r(1-p)\mu(LDWW)}_{\text{after splitting into two parts}}-2pr^2\{1+q-p+2p^2\\
&\quad -2p^2(1-p-q)(1-p)+pq(1-p)-p^3\}\mu(WWWDL)+\{2qp^2r(1-p)+4p^4r^2(1-p)\\
&\quad -2p^2r^3(1-p)(1+p-q)\}\mu(WWWDL)-4(1-p)r^2q(r-p)\mu(WWWDL)\\
&\leqslant -r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\mu(LDWW)\underbrace{-2qp^2r(1-p)\mu(LDWWW)}_{\text{cancel out}}-2pr^2\{1+q-p+2p^2\\
&\quad -2p^2(1-p)^2+2p^2q(1-p)+pq(1-p)-p^3\}\mu(WWWDL)\underbrace{+2qp^2r(1-p)\mu(WWWDL)}_{\text{cancel out}}\\
&\quad +\{4p^4r^2(1-p)-2p^2r^3(1-p)(1+p-q)\}\mu(WWWDL)-4(1-p)r^2q(r-p)\mu(WWWDL)\\
&= -r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\mu(LDWW)-2pr^2\{1+q-p+2p^2-2p^2+4p^3-2p^4\\
&\quad +2p^2q(1-p)+pq(1-p)-p^3\}\mu(WWWDL)+\{4p^4r^2(1-p)\\
&\quad -2p^2r^3(1-p)(1+p-q)\}\mu(WWWDL)-4(1-p)r^2q(r-p)\mu(WWWDL)\\
&= -r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\mu(LDWW)-2pr^2\{1+q-p+p^3+2p^2q(1-p)\\
&\quad +pq(1-p)\}\mu(WWWDL)-2pr^2(2p^3-2p^4)\mu(WWWDL)+\{4p^4r^2(1-p)\\
&\quad -2p^2r^3(1-p)(1+p-q)\}\mu(WWWDL)-4(1-p)r^2q(r-p)\mu(WWWDL)\\
&= -r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\mu(LDWW)-2pr^2\{1+q-p+p^3+2p^2q(1-p)\\
&\quad +pq(1-p)\}\mu(WWWDL)-4p^4r^2(1-p)\mu(WWWDL)+4p^4r^2(1-p)\mu(WWWDL)\\
&\quad -2p^2r^3(1-p)(1+p-q)\mu(WWWDL)-4(1-p)r^2q(r-p)\mu(WWWDL)\\
&= -r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\mu(LDWW)-2pr^2\{1+q-p+p^3+2p^2q(1-p)+pq(1-p)\}
\end{aligned}
$$

$$\mu(WWWDL) - 2p^2r^3(1-p)(1+p-q)\mu(WWWDL) - 4(1-p)r^2q(r-p)\mu(WWWDL). \tag{3.4.62}$$

Lemma 3.4.2. *The coefficient of $\mu(WWWDL)$ in (3.4.62) is bounded above by $-2p^2r^3(1-p)(1+p-q)$ for all values of $(p,q) \in \mathscr{S}$, where $\mathscr{S}$ is as defined in (3.1.1).*

Proof. As long as $r - p \geqslant 0$, i.e. $2p + q \leqslant 1$, the coefficient of $\mu(WWWDL)$ in (3.4.62) is the sum of three non-positive quantities, namely $-2pr^2\{1+q-p+p^3+2p^2q(1-p)+pq(1-p)\}$, $-2p^2r^3(1-p)(1+p-q)$ and $-4(1-p)r^2q(r-p)$, and hence it is obviously bounded above by the second of these non-positive quantities. When $r - p < 0$, the quantity $-4(1-p)r^2q(r-p)$ is actually non-negative, and we negate it by observing that

$$\begin{aligned}
&-2pr^2\{1+q-p+p^3+2p^2q(1-p)+pq(1-p)\}+4(1-p)r^2q(p-r)\\
&=-2pr^2\{1+q-p+p^3+2p^2q(1-p)+pq(1-p)\}+4(1-p)r^2qp-4(1-p)r^3q\\
&\leqslant -2pr^2\{1+q-p+p^3+2p^2q(1-p)+pq(1-p)\}+4(1-p)r^2qp\\
&=-2pr^2\{1+q-p+p^3+2p^2q(1-p)+pq(1-p)-2q(1-p)\}\\
&=-2pr^2\{1+q-p+p^3+2p^2q(1-p)+pq(1-p)-2q+2pq\}\\
&=-2pr^2\{r+p^3+2p^2q(1-p)+pq(1-p)+2pq\},
\end{aligned}$$

which is non-positive. This again leads to the conclusion that the combined coefficient of $\mu(WWWDL)$ in (3.4.62) is bounded above by $-2p^2r^3(1-p)(1+p-q)$. $\square$

Next, using the simple inequalities $\mu(WWWDDL) \leqslant \mu(LDDW)$ and $\mu(WWWDWL) \leqslant \mu(LWDW)$, we simplify A_2 (recall A_2 from (3.4.61)) as follows:

$$\begin{aligned}
A_2 &= -r(4q^2+4qp^2-2q^3+2q^2p)\{\mu(LDDW)+\mu(LWDW)\} - 2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)\\
&\quad + \{2qp^2r(1-p)+2p^4r^2(1-p)-4p^3r^3(1-p)\}\mu(WWWD\{W,D\}L)\\
&\leqslant \underbrace{-r\{4q^2+4qp^2-2q^3+2q^2p\}\{\mu(LDDW)+\mu(LWDW)\}}_{(1)}\underbrace{-2p^2r^2(1-p)(1-2pr)\mu(WWWD\{W,D\}L)}_{(2)}\\
&\quad \underbrace{+2p^4r^2(1-p)\mu(WWWD\{W,D\}L)-4p^3r^3(1-p)\mu(WWWD\{W,D\}L)}_{\text{combine with (2)}}\\
&\quad \underbrace{+2qp^2r(1-p)\{\mu(LDDW)+\mu(LWDW)\}}_{\text{combine with (1)}}\\
&= \underbrace{-r\{4q^2+4qp^2-2q^3+2q^2p-2qp^2(1-p)\}\{\mu(LDDW)+\mu(LWDW)\}}_{\text{(1) after combination}}
\end{aligned}$$

$$\underbrace{-2p^2r^2(1-p)[1-2pr-p^2+2pr]\mu(WWWD\{W,D\}L)}_{\text{(2) after combination}}$$

$$= -r\{4q^2+2qp^2(1+p)-2q^3+2q^2p\}\{\mu(LDDW)+\mu(LWDW)\}$$
$$-2p^2r^2(1-p)(1-p^2)\mu(WWWD\{W,D\}L). \tag{3.4.63}$$

Finally, using the simple inequality $\mu(WWWWDL) \leqslant \mu(WWWWD)$, we simplify A_3 (recall A_3 from (3.4.61)) as follows:

$$\begin{aligned}
A_3 = &-\{2qp^2r(1-p)+4p^4r^2(1-p)\}\mu(WWWWD)-2p^2r^2(1-p)(1-pr)\mu(WWWWDL)\\
&+\{2qp^2r(1-p)+2p^4r^2(1-p)-2p^3r^3(1-p)\}\mu(WWWWDL)\\
= &-2qp^2r(1-p)\mu(WWWWD)+2qp^2r(1-p)\mu(WWWWDL)-4p^4r^2(1-p)\mu(WWWWD)\\
&-[2p^2r^2(1-p)(1-pr)-2p^4r^2(1-p)+2p^3r^3(1-p)]\mu(WWWWDL)\\
\leqslant &-2p^2r^2(1-p)[1-pr-p^2+pr]\mu(WWWWDL)-4p^4r^2(1-p)\mu(WWWWD)\\
= &-2p^2r^2(1-p)(1-p^2)\mu(WWWWDL)-4p^4r^2(1-p)\mu(WWWWD).
\end{aligned} \tag{3.4.64}$$

Incorporating the expressions obtained from (3.4.62) (in conjunction with the bound obtained from Lemma 3.4.2), (3.4.63) and (3.4.64) into (3.4.61) yields (again, in each step, we highlight using underbraces the terms to be combined to obtain the next step):

$$\begin{aligned}
C_2 \leqslant\ &\underbrace{-r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\mu(LDWW)-2p^2r^3(1-p)(1+p-q)\mu(WWWDL)}_{\text{substituting from Lemma 3.4.2 and (3.4.62)}}\\
&\underbrace{-r(4q^2+2qp^2(1+p)-2q^3+2q^2p)[\mu(LWDW)+\mu(LDDW)]}_{\text{substituting from (3.4.63)}}\\
&\underbrace{-2p^2r^2(1-p)(1-p^2)\mu(WWWD\{W,D\}L)}_{\text{substituting from (3.4.63)}}\underbrace{-2p^2r^2(1-p)(1-p^2)\mu(WWWWDL)}_{\text{substituting from (3.4.64)}}\\
&\underbrace{-4p^4r^2(1-p)\mu(WWWWD)}_{\text{substituting from (3.4.64)}}-[p(1-r)+q][\mu(WD)+\mu(WWD)]\\
&+[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)\\
&-[2qp^2r^2+2q^2p^2r+6p^4r^2(1-p)]\mu(WWWD)-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD)\\
= &-r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\{\mu(LDWW)+\mu(LWDW)+\mu(LDDW)\}\\
&-[p(1-r)+q][\mu(WD)+\mu(WWD)]\underbrace{+[2pqr^2\{4p-2+p^2\}+6p^4r^2-4p^4r^2(p+q)]\mu(WWWD)}\\
&+2p^2r^2(1-p)(1-pr)\mu(WWWWWD)-2p^2r^3(1-p)(1+p-q)\mu(WWWDL)\\
&-2p^2r^2(1-p)(1-p^2)\mu(WWWD\{W,D\}L)-2p^2r^2(1-p)(1-p^2)\mu(WWWWDL)
\end{aligned}$$

$$
\begin{aligned}
&\underbrace{-[2qp^2r^2+2q^2p^2r+6p^4r^2(1-p)]\mu(WWWD)}-4p^4r^2(1-p)\mu(WWWWD)\\
&-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD)\\
=\ &-r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\{\mu(LDWW)+\mu(LWDW)+\mu(LDDW)\}\\
&\underbrace{-[p(1-r)+q][\mu(WD)+\mu(WWD)]}+[2pqr^2\{3p-2+p^2\}-2q^2p^2r\underbrace{+2p^5r^2}-4p^4qr^2]\mu(WWWD)\\
&\underbrace{+2p^2r^2(1-p)(1-pr)\mu(WWWWWWD)}-2p^2r^3(1-p)(1+p-q)\mu(WWWDL)\\
&-2p^2r^2(1-p)(1-p^2)\mu(WWWD\{W,D\}L)-2p^2r^2(1-p)(1-p^2)\mu(WWWWDL)\\
&\underbrace{-4p^4r^2(1-p)\mu(WWWWD)-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD)}. \qquad (3.4.65)
\end{aligned}
$$

We combine the terms highlighted by underbrace in the last step of (3.4.65) as follows:

$$
\begin{aligned}
&2p^2r^2(1-p)(1-pr)\mu(WWWWWWD)-[2qp^2r(1-p)+2p^4r^2(1-p)]\mu(WWWWWD)\\
&+2p^5r^2\mu(WWWD)-4p^4r^2(1-p)\mu(WWWWD)-[p(1-r)+q][\mu(WD)+\mu(WWD)]\\
=\ &2p^2r^2(1-p)(1-pr)\mu(WWWWWWD)-2p^4r^2(1-p)\mu(WWWWWD)-4p^4r^2(1-p)\mu(WWWWD)\\
&+2p^5r^2\mu(WWWD)-2qp^2r(1-p)\mu(WWWWWD)-[p(1-r)+q][\mu(WD)+\mu(WWD)]\\
&\text{(rearranging the terms of the previous expression)}\\
\leqslant\ &2p^2r^2(1-p)(1-pr)\mu(WWWWWWD)-2p^4r^2(1-p)\mu(WWWWWD)-4p^4r^2(1-p)\mu(WWWWWD)\\
&+2p^5r^2\mu(WWWD)-2qp^2r(1-p)\mu(WWWWWD)-[p^2+q(1+p)][\mu(WD)+\mu(WWD)]\\
&\text{(using the inequality } \mu(WWWWD)\geqslant\mu(WWWWWD))\\
=\ &2p^2r^2(1-p)(1-pr-3p^2)\mu(WWWWWWD)+2p^5r^2\mu(WWWD)-p^2\{\mu(WD)+\mu(WWD)\}\\
&-2qp^2r(1-p)\mu(WWWWWD)-q(1+p)\{\mu(WD)+\mu(WWD)\}. \qquad (3.4.66)
\end{aligned}
$$

Lemma 3.4.3. *The sum in* (3.4.66) *is non-positive for all* $(p,q)\in\mathscr{S}$, *where* $\mathscr{S}$ *is as defined in* (3.1.1). *Moreover, it is bounded above by* $-2qp^2r(1-p)\mu(WWWWWD)-q(1+p)\{\mu(WD)+\mu(WWD)\}$.

Proof. In this proof, we make use of the inequalities $\mu(WD)\geqslant\mu(WWD)\geqslant\mu(WWWD)\geqslant\mu(WWWWWD)$. If $1-pr-3p^2\leqslant 0$, each of the terms in (3.4.66), except $2p^5r^2\mu(WWWD)$, is non-positive, and we tackle this single non-negative term as follows:

$$
\begin{aligned}
2p^5r^2\mu(WWWD)-p^2\{\mu(WD)+\mu(WWD)\} &= 2p^5r^2\mu(WWWD)-2p^2\mu(WWWD)\\
&= -2p^2(1-p^3r^2)\mu(WWWD), \qquad (3.4.67)
\end{aligned}
$$

and every term in the above sum is non-positive. If $1-pr-3p^2>0$, both the first and the second terms of (3.4.66) are non-negative, while the rest are non-positive, and these non-negative terms are taken care of as follows:

$$
\begin{aligned}
&2p^2r^2(1-p)(1-pr-3p^2)\mu(WWWWWD)+2p^5r^2\mu(WWWD)-p^2\{\mu(WD)+\mu(WWD)\}\\
\leqslant\ &2p^2r^2(1-p)(1-pr-3p^2)\mu(WWWD)+2p^5r^2\mu(WWWD)-2p^2\mu(WWWD)\\
=\ &2p^2r^2[(1-p)(1-pr-3p^2)+p^3]\mu(WWWD)-2p^2\mu(WWWD)\\
=\ &2p^2r^2[1-p(1-p^2)-pr(1-p)-3p^2(1-p)]\mu(WWWD)-2p^2\mu(WWWD)\\
=\ &2p^2[r^2\{1-p(1-p^2)-pr(1-p)-3p^2(1-p)\}-1]\mu(WWWD), \qquad (3.4.68)
\end{aligned}
$$

and once again, we see that each term in the above expression is non-positive, since each of r^2 and $1-p(1-p^2)-pr(1-p)-3p^2(1-p)$ is bounded above by 1. From the final upper bounds in (3.4.67) and (3.4.68), we deduce that the expression in (3.4.66) is bounded above by $-2qp^2r(1-p)\mu(WWWWWD)-q(1+p)\{\mu(WD)+\mu(WWD)\}$. □

Incorporating (3.4.66) into (3.4.65) and using Lemma 3.4.3, we see that C_2 can be bounded above as follows:

$$
\begin{aligned}
C_2\leqslant\ &-r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\{\mu(LDWW)+\mu(LWDW)+\mu(LDDW)\}\\
&\underbrace{-q(1+p)[\mu(WD)+\mu(WWD)]}_{\text{split this into 2 parts, as below}}+[\underbrace{2pqr^2\{3p-2+p^2\}}_{\text{consider this separately}}-2q^2p^2r-4p^4qr^2]\mu(WWWD)\\
&-2qp^2r(1-p)\mu(WWWWWD)-2p^2r^3(1-p)(1+p-q)\mu(WWWDL)\\
&\underbrace{-2p^2r^2(1-p)(1-p^2)\mu(WWWD\{W,D\}L)-2p^2r^2(1-p)(1-p^2)\mu(WWWWDL)}_{\text{combine these two}}\\
=\ &-r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\{\mu(LDWW)+\mu(LWDW)+\mu(LDDW)\}\\
&-q(1+p-pr)[\mu(WD)+\mu(WWD)]\ \underbrace{-qpr[\mu(WD)+\mu(WWD)]}_{\text{bounded above by }-2pqr\mu(WWWD)}\ +2pqr^2\{3p-2+p^2\}\mu(WWWD)\\
&-[2q^2p^2r+4p^4qr^2]\mu(WWWD)-2qp^2r(1-p)\mu(WWWWWD)\\
&-2p^2r^3(1-p)(1+p-q)\mu(WWWDL)-2p^2r^2(1-p)(1-p^2)\mu(WWW\widehat{**}L)\\
\leqslant\ &-r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\{\mu(LDWW)+\mu(LWDW)+\mu(LDDW)\}\\
&-q(1+p-pr)[\mu(WD)+\mu(WWD)]\underbrace{-2qpr\mu(WWWD)+2pqr^2\{3p-2+p^2\}\mu(WWWD)}\\
&-[2q^2p^2r+4p^4qr^2]\mu(WWWD)-2qp^2r(1-p)\mu(WWWWWD)\\
&-2p^2r^3(1-p)(1+p-q)\mu(WWWDL)-2p^2r^2(1-p)(1-p^2)\mu(WWW\widehat{**}L)\\
=\ &-r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\{\mu(LDWW)+\mu(LWDW)+\mu(LDDW)\}
\end{aligned}
$$

$$
\begin{aligned}
& -q(1+p-pr)[\mu(WD)+\mu(WWD)]+\underbrace{[-2qpr+4p^2qr^2+2pqr^2\{p-2+p^2\}]\mu(WWWD)} \\
& -[2q^2p^2r+4p^4qr^2]\mu(WWWD)-2qp^2r(1-p)\mu(WWWWWD) \\
& -2p^2r^3(1-p)(1+p-q)\mu(WWWDL)-2p^2r^2(1-p)(1-p^2)\mu(WWW\widehat{**}L) \\
= & -r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\{\mu(LDWW)+\mu(LWDW)+\mu(LDDW)\} \\
& -q(1+p-pr)[\mu(WD)+\mu(WWD)]\underbrace{-2qpr[1-2pr]\mu(WWWD)-2pqr^2\{2-p-p^2\}\mu(WWWD)} \\
& -[2q^2p^2r+4p^4qr^2]\mu(WWWD)-2qp^2r(1-p)\mu(WWWWWD) \\
& -2p^2r^3(1-p)(1+p-q)\mu(WWWDL)-2p^2r^2(1-p)(1-p^2)\mu(WWW\widehat{**}L) \\
= & -r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\{\mu(LDWW)+\mu(LWDW)+\mu(LDDW)\} \\
& -q(1+p-pr)[\mu(WD)+\mu(WWD)]\underbrace{-2qpr[1-2pr]\mu(WWWD)}-2pqr^2(2+p)(1-p)\mu(WWWD) \\
& -[2q^2p^2r+4p^4qr^2]\mu(WWWD)-2qp^2r(1-p)\mu(WWWWWD) \\
& -2p^2r^3(1-p)(1+p-q)\mu(WWWDL)-2p^2r^2(1-p)(1-p^2)\mu(WWW\widehat{**}L). \qquad (3.4.69)
\end{aligned}
$$

We note that, by the AM-GM inequality, $2pr \leqslant p^2+r^2 \leqslant p+r \leqslant 1$, so that the underbraced term, and consequently, each term in the final expression of (3.4.69), is non-positive. Note that (3.4.69) also provides the final, simplified expression for C_2. From (3.4.59) and (3.4.69), we see that (3.4.58) now transforms into the weight function inequality

$$
\begin{aligned}
& w_3(\widehat{\mathrm{F}}_{p,q}\mu) \leqslant w_3(\mu)-[p(1-r)+q]\mu(LWD)-r(4q^2+4qp^2-2q^3+2q^2p)[\mu(LDDD)+\mu(LDWD) \\
& +\mu(LWDD)+\mu(LWWD)]-2r(1-p^2)\mu(LDDL)-r(1-p)\mu(LDL)-pr^2(1+2q)\mu(L\widehat{***}L) \\
& -qr[1-2p(p-r-p^2)]\mu(\widehat{***})-pr^2\{1+2q(1-p)\}\mu(DWWWD)-6pqr^2(1-p)\mu(WWWWD) \\
& -2p^3r^3(1-p)\{\mu(LWWWWD)+\mu(DWWWWD)\}-2p^2qr^2\mu(LWWWD)-4qp^2r^2[\mu(L\{W,D\}\widehat{***}) \\
& -\mu(LWWWD)]-2pqr^2(2-2q+p)\mu(L\widehat{**})-r(4q^2+2qp^2(1+p)-2q^3+2q^2p)\{\mu(LDWW) \\
& +\mu(LWDW)+\mu(LDDW)\}-q(1+p-pr)[\mu(WD)+\mu(WWD)]-2qpr[1-2pr]\mu(WWWD) \\
& -2pqr^2(2+p)(1-p)\mu(WWWD)-[2q^2p^2r+4p^4qr^2]\mu(WWWD)-2qp^2r(1-p)\mu(WWWWWD) \\
& -2p^2r^3(1-p)(1+p-q)\mu(WWWDL)-2p^2r^2(1-p)(1-p^2)\mu(WWW\widehat{**}L). \qquad (3.4.70)
\end{aligned}
$$

Finally, we have achieved a weight function inequality that is of the form presented in (3.4.2).

3.4.7 The desired conclusion

From (3.4.10), (3.4.17), (3.4.23) and (3.4.49), the final weight function turns out to be

$$\begin{aligned}
w_3(\mu) = \mu(D) + 2\mu(WD) - \mu(DWD) + 2\mu(LWWD) - p(p+q)\mu(D) - [2pr\{\mu(LD) + \mu(LWD)\} \\
+ 2p^2r\{\mu(LDD) + \mu(LDWD) + \mu(LWDD)\} + 4r\mu(LDWL) + 2p\mu(LWWD)] \\
- 2(q+p^2r)\mu(LWWD) - 2p^2r\{\mu(LDWW) + \mu(LWDW)\} - q\mu(D) \\
= (1-p^2-pq-q)\mu(D) + 2\mu(WD) - \mu(DWD) + [2-2p-2q-2p^2r]\mu(LWWD) \\
- 2pr\mu(LD) - 2pr\mu(LWD) - 2p^2r\{\mu(LDD) + \mu(LDWD) + \mu(LWDD)\} \\
- 4r\mu(LDWL) - 2p^2r\{\mu(LDWW) + \mu(LWDW)\},
\end{aligned}$$

which matches with the weight function stated in (3.4.3).

We now wish to draw the conclusion, using (3.4.70), that $\mu(D) = 0$ when μ is a translation-invariant and reflection-invariant stationary distribution for $\widehat{\mathrm{F}}_{p,q}$ (recall Lemma 3.4.1). Note that this conclusion, i.e. that $\mu(D) = 0$, is immediate when $r = 0$, i.e. $p+q = 1$, so that henceforth, we only consider $r > 0$, i.e. $p+q < 1$.

To begin with, we note that when μ is stationary for $\widehat{\mathrm{F}}_{p,q}$, we have $\widehat{\mathrm{F}}_{p,q}\mu = \mu$, so that $w_3(\widehat{\mathrm{F}}_{p,q}\mu) = w_3(\mu)$. When $q > 0$, we conclude that the coefficient $-qr[1-2p(p-r-p^2)]$ of $\mu(\widehat{***})$ in (3.4.70) is non-zero (see (3.4.56) that shows why $1-2p(p-r-p^2)$ is strictly positive). The last two sentences and (3.4.70) together imply that $\mu(\widehat{***}) = 0$ when μ is stationary for $\widehat{\mathrm{F}}_{p,q}$ and $q > 0$. Since $\mu(D) = \widehat{\mathrm{F}}_{p,q}\mu(D) = r\mu(\widehat{***})$ from (3.4.4), we conclude that $\mu(D) = 0$ when μ is stationary for $\widehat{\mathrm{F}}_{p,q}$ and $q > 0$.

We now come to the analysis when $q = 0$, which forces $p > 0$ from the definition of $\mathscr{S}$ in (3.1.1).

Lemma 3.4.4. *When $q = 0$, we obtain $\mu(D) = 0$ from (3.4.70) for any stationary distribution μ of $\widehat{\mathrm{F}}_{p,q}$.*

Proof. When $q = 0$, forcing $p > 0$, the coefficients $-[p(1-r)+q]$ of $\mu(LWD)$ and $-2p^2r^3(1-p)(1+p-q)$ of $\mu(WWWDL)$ in (3.4.70) are both strictly negative, which implies that $\mu(LWD) = \mu(WWWDL) = 0$. The former, via (3.4.29), yields $\mu(WWW\widehat{**}) = 0$, and this, together with the latter, yields

$$\mu(WWWD) = \mu(WWWDD) + \mu(WWWDW) + \mu(WWWDL) = 0.$$

We now focus on finding $\widehat{\mathrm{F}}_{p,q}\mu(WWWD)$.

In order for $(\widehat{\mathrm{F}}_{p,q}\eta(0),\widehat{\mathrm{F}}_{p,q}\eta(1),\widehat{\mathrm{F}}_{p,q}\eta(2),\widehat{\mathrm{F}}_{p,q}\eta(3))$ to equal $(WWWD)$, we must have $(\eta(3),\eta(4),\eta(5)) \in \widehat{***}$. If each of $\eta(0),\eta(1),\eta(2)$ belongs to $\{W,D\}$, then each of the events $\widehat{\mathrm{F}}_{p,q}\eta(i) = W$, for $i = 0,1,2$, happens with probability p. If $\eta(0) = L$ and $\eta(1),\eta(2) \in \{W,D\}$, then $\widehat{\mathrm{F}}_{p,q}\eta(0) = W$ happens with probability $1-q$ and each of $\widehat{\mathrm{F}}_{p,q}\eta(1) = W$ and $\widehat{\mathrm{F}}_{p,q}\eta(2) = W$ happens with probability p. If $\eta(1) = L$ and $\eta(2) \in \{W,D\}$, then each of $\widehat{\mathrm{F}}_{p,q}\eta(0) = W$ and $\widehat{\mathrm{F}}_{p,q}\eta(1) = W$ happens with probability $1-q$ and $\widehat{\mathrm{F}}_{p,q}\eta(2) = W$ happens with probability p. If $\eta(2) = L$, then each $\widehat{\mathrm{F}}_{p,q}\eta(i) = W$, for $i = 0,1,2$, happens with probability $1-q$. Combining everything, we have

$$\begin{aligned}\widehat{\mathrm{F}}_{p,q}\mu(WWWD) = p^3 r\mu(\{W,D\}^3\widehat{***}) + (1-q)p^2 r\mu(L\{W,D\}^2\widehat{***}) + (1-q)^2 pr\mu(L\{W,D\}\widehat{***})\\ + (1-q)^3 r\mu(L\widehat{***}).\end{aligned}$$

When $\mu(WWWD) = 0$, $q = 0$ (and hence $p > 0$), and μ is stationary for $\widehat{\mathrm{F}}_{p,q}$, we conclude from the previous paragraph that $\mu(\{W,D\}^3\widehat{***}) = \mu(L\{W,D\}^2\widehat{***}) = \mu(L\{W,D\}\widehat{***}) = \mu(L\widehat{***}) = 0$. Adding these, we get $\mu(\widehat{***}) = 0$, and via (3.4.4), this, once again, implies that $\mu(D) = 0$. $\square$

3.5 A formal game theoretic formulation of the problem

We provide a slightly more general game theoretic formulation of our games to facilitate the subsequent discussion on the connection of our games with various other games in the literature. Furthermore, we discuss some interesting open problems in this area. To this end, we shall consider $N = \{1,\ldots,n\}$ to be the (finite) set of players participating in the game, with each player $i \in N$ allowed to choose from a (finite) set of *actions* denoted by A_i. Let $T = \mathbb{N}$ be the set of *rounds* in the game, and let $T_0 = T \cup \{0\}$. A *mover-sequence* is an (infinite) sequence $\mu : T \to N$, such that $\mu(t)$, for $t \in T$, indicates the player who is supposed to make their move in round t of the game.

Given a mover-sequence μ, a *move-sequence* (also referred to as a *history*) is a (finite or infinite) sequence of actions $\underset{\sim}{a} = (a_1,a_2,\ldots)$ such that $a_k \in A_{\mu(k)}$ for each k. We denote by $\mathscr{A}_k$ the set of all move-sequences of length k, for each $k \in T_0$ (note that the move-sequence of length 0 is referred to as the *empty move-sequence* and denoted by $\underset{\sim}{a}^0$). We let $\mathscr{A}_\infty$ denote the set of all move-sequences of infinite length, and $\mathscr{A}$ denote the set of *all* move-sequences, both finite and infinite.

Each move-sequence $\underset{\sim}{a} \in \mathscr{A}$ is assigned a random variable $X_{\underset{\sim}{a}}$ taking values in $\mathbb{R}^n$, which denotes the (random) utilities of the players corresponding to the move-sequence $\underset{\sim}{a}$. The game starts once a realization $(x_{\underset{\sim}{a}})_{\underset{\sim}{a}\in\mathscr{A}}$ of values of the collection of random variables $(X_{\underset{\sim}{a}})_{\underset{\sim}{a}\in\mathscr{A}}$ has been fixed and revealed to all the players. [1] A *strategy* for player i is a collection of functions $\sigma_i := \{\sigma_i^k : k \in T \text{ such that } \mu(k) = i\}$, where $\sigma_i^k : \mathscr{A}_{k-1} \to A_i$. A collection of strategies $\sigma_N := (\sigma_1, \ldots, \sigma_n)$ is called a *strategy profile*.

Each strategy profile σ_N induces an infinite sequence of actions $\underset{\sim}{a}(\sigma_N) := (a_1, a_2, \ldots)$ in the following, natural way: $a_1 = \sigma^1_{\mu(1)}$, $a_2 = \sigma^2_{\mu(2)}(a_1))$, $a_3 = \sigma^3_{\mu(3)}(a_1, a_2)$, and so on. For $t \in T$, we denote by $\underset{\sim}{a}^t(\sigma_N)$ the (truncated) sequence $(a_1, \ldots, a_t)$ of $\underset{\sim}{a}(\sigma_N)$ that is the *game path* of length t induced by σ_N. The *stage utility* of player i at time t, $t \in T$, is given by $u_i^t(\sigma_N) := x_{\underset{\sim}{a}^t(\sigma_N)}(i)$. The *final utility* $u_i(\sigma_N)$ of player i corresponding to a strategy profile σ_N is a (real-valued) function of the stage utility sequence $(\underset{i}{\underset{\sim}{u}}^t(\sigma_N))_{t\in T}$.

We assume, in line with game theory, that each player $i \in N$ plays the game *optimally*. In other words, given a collection of strategies $\sigma_{N\setminus\{i\}}$ of the other players, i plays a *best response*: a strategy σ_i^* that maximizes her utility given $\sigma_{N\setminus\{i\}}$, i.e. $u_i(\sigma_i^*, \sigma_{N\setminus\{i\}}) \geqslant u_i(\sigma_i, \sigma_{N\setminus\{i\}})$ for all strategies σ_i of player i. A strategy σ_i for player i is *dominant* if it is a best response to *every* collection of strategies of the other players. In other words, no matter what strategies the other players adopt, playing σ_i is the best for player i, i.e. $u_i(\sigma_i, \sigma_{N\setminus\{i\}}) \geqslant u_i(\sigma_i', \sigma_{N\setminus\{i\}})$ for all strategies σ_i' of player i and all collections of strategies $\sigma_{N\setminus\{i\}}$ of the other players. [2] A strategy profile σ_N is said to be a *Nash equilibrium* if it is a best response to itself, i.e. for each $i \in N$, σ_i is a best response to $\sigma_{N\setminus\{i\}}$.

Several interesting games can be constructed by imposing restrictions and dependence structures on the random variables $(X_{\underset{\sim}{a}})_{\underset{\sim}{a}\in\mathscr{A}}$ (and accordingly, on the strategy-profiles). A class of such games are the ones where there are two players making alternative moves (that is, $\mu = (1,2,1,2,\ldots))$ and having the same set of actions (i.e. $A_1 = A_2$). For such games, there is an equivalence relation $\sim$ on $\mathscr{A}$ such that $X_{\underset{\sim}{a}^t} = X_{\underset{\sim}{b}^t}$ and $\sigma^t_{\mu(t)}(\underset{\sim}{a}^t) = \sigma^t_{\mu(t)}(\underset{\sim}{b}^t)$ whenever $\underset{\sim}{a}^t$ and $\underset{\sim}{b}^t$ are equivalent.

Combinatorial games played on directed graphs can be modeled via this approach. Let G be a directed graph such that each vertex of G has the same number of outgoing edges, and the outgoing edges from each vertex are identified with the elements of the set A of common actions (where $A = A_1 = A_2$).[3] Then, given an initial vertex, every strategy profile σ_N induces a (directed)

[1] It is worth emphasizing here that these aren't games with incomplete information as there is no randomness involved once the game begins.

[2] It is well-known that a dominant strategy for a player may not exist unless the game has some particular structure.

[3] One can model the games played on arbitrary directed graphs by (recursively) defining the set of actions A_i^t (for player i who is supposed to make their move in round t) as a function defined on the set of possible histories of length $t-1$. In fact, in addition to the utilities and the set of actions, which player is allowed to make a move in a given round

path in the graph G. We call two game paths $\underset{\sim}{a}^t$ and $\underset{\sim}{b}^t$ of length t equivalent if they lead to the same vertex in G. Suppose that the utilities are zero-sum, that is, the sum of the utilities of all the players corresponding to any realization of $X_{\underset{\sim}{a}^t}$ is zero. Let the final utility $u_i(\sigma_N)$ of any player $i \in N$ be the first non-zero value in the sequence of stage utilities $(u_i^t(\sigma_N))_{(t \in T_0)}$. In other words, the game ends whenever one of the players receive a non-zero stage utility and the utilities corresponding to that stage become the final utilities. We say that the game is *winning* for player i, for any $i \in N$, if player i has a dominant strategy which, when adopted, rewards her 1 as her final utility. The game ends in a draw if neither player has a winning dominant strategy.

Suppose, further, that $X_{\underset{\sim}{a}^t}(\mu(t))$ equals 1 with probability p, equals -1 with probability q, and equals 0 with probability $1-p-q$, for all $t \in T_0$ (here, we follow the convention that $\mu(0) = 2$). This leads to an (i.i.d.) percolation game on a k-regular tree when $|A| = k$ and each equivalence class is a singleton. [4] On the other hand, fixing an initial vertex arbitrarily (say, the origin), defining $A = \{(0,1),(1,1),(2,1)\}$ (respectively, $A = \{(0,1),(1,1)\}$), and defining two game paths $\underset{\sim}{a}^t = (a^1,\ldots,a^t)$ and $\underset{\sim}{b}^t = (b^1,\ldots,b^t)$ to be equivalent if $\sum_{s=1}^t a^s = \sum_{s=1}^t b^s$, one obtains the percolation game we have studied in this work (respectively, the percolation game considered in [60]).

Clearly, under such an optimality assumption, a player will play a dominant strategy whenever she has one. We explore what happens to the flow of the game if none of the players has a dominant strategy, limiting ourselves to the case where $|N| = 2$. Suppose player 1 plays a strategy σ_1. Since σ_1 is not dominant, there is a strategy σ_2 of player 2 which, if adopted by player 2, prevents player 1 from winning. Suppose player 2 wins corresponding to the strategy-profile (σ_1, σ_2). However, since σ_2 is not dominant, player 1 has a strategy, say σ_1', such that player 2 does not win when player 1 adopts σ_1'. Continuing in this manner, it follows that when both players play optimally (in other words, under any Nash equilibrium), neither player wins the game, allowing the game to continue forever with each player receiving zero utility at every stage of the game. Summarizing, under the assumption of optimal play, a draw implies continuation of the game for infinite time, which, in addition to requiring the existence of an infinite open path in the (directed) graph, requires that the path be *self-enforcing*, i.e. that neither of the players has an incentive to deviate from this path. The event of an infinite path in a random graph is a prime object of interest in percolation theory, and consequently, the event of a draw is an important object of study in percolation games.

can also be a function of the histories. In an even more general setup, all these three variables can be defined randomly corresponding to every history. See [46] (and §3.6.1 where we discuss) for a formal definition of such random games when the corresponding graph is a tree and the said variables are i.i.d over histories.

[4] See [52] for a formal definition of iid percolation games.

3.6 Relation with existing literature and open problems

Researchers mainly explore two types of problems in the context of combinatorial games played in random environments, the first of which is concerned with the probability of draw, while the second type focuses on the value of the game.

3.6.0.1 Generalizing the set of actions

For problems related to the probability of draw in percolation games on $\mathbb{Z}^2$ (or even on arbitrary $\mathbb{Z}^k$), the set of actions A can be generalized in three ways. The first of these is achieved by adding more vertices at the 'next' level, where the 'next' level refers to the value of the y-coordinate in the set A of actions being equal to 1 (for instance, $A = \{(0,1),\ldots,(k,1)\}$ for some $k \in \mathbb{N}$). The second possibility is the addition of vertices at levels that are 'higher' than the 'next' level, i.e. in which the value of the y-coordinate of at least one of the vertices in the set A of actions is strictly larger than 1 (for instance, $A = \{(0,1),(1,2),(2,1)\}$). The third option is to make A a (simple and structured) function $A(x,y)$ of $(x,y) \in \mathbb{Z}^2$. Such an instance has been dealt with in [15], and the percolation game considered therein has been shown to arise when edge-percolation is considered (i.e., instead of the vertices, it is the edges that are labeled, independently, as trap, target or open with probabilities p, q and $1-p-q$ respectively). The 'location dependent' action set $A(x,y)$ for the edge-percolation game studied in [15] depends (only) on the parity of the first coordinate of (x,y) (it is shown in [15] that $A(x,y) = \{(0,1),(1,1)\}$ when x is even, and $A(x,y) = \{(1,1),(1,2)\}$ when x is odd).

3.6.0.2 Arbitrary mover-sequence

Both [60] and this chapter consider an alternating mover-sequence (i.e., $\mu = (1,2,1,2\ldots)$). We would like to explore what happens (in particular, whether the probability of draw continues to be 0 or not) for other mover-sequences (that have been endowed with enough structure, such as when $\mu = (1,1,2,2,1,1,2,2,\ldots)$). Technically, this relates to a PCA whose update rule is dependent on time. To the best of our knowledge, ergodicity properties of such PCAs have not been explored in the literature, and we think it is an important problem as it generalizes the commonly studied notion of PCAs to a much broader perspective.

3.6.0.3 Generalizing the event of draw

Yet another class of open problems is concerned with generalizing the events of winning / losing / drawing. Let $u_N^t = (u_1^t,\ldots,u_n^t) \in \mathbb{R}^n$ denote a utility vector for all players in the stage game at time t (the *stage game* refers to the base game, so that the actual game is made up of repetitions of the stage game, and after each stage game is over, every player receives some utility). Given a set of sequences of utility vectors $W_i \subseteq \{(u_N^t)_{t \in T_0}\}$, let us call a game *winning* for player i if player i has a dominant strategy to ensure a utility sequence in W_i, and a game is said to *lead to a draw* if it is not winning for any of the players. A natural extension of the existing notion of winning (as defined in our work and in [60]) is obtained by defining W_i as the set of utility sequences for which i receives a cumulative utility of amount k, for some pre-fixed $k \in \mathbb{N}$, before any other player does. More formally, W_i contains those utility-vector sequences $(u_N^t)_{t \in T_0}$ for which there is $\hat{t} \in T_0$ such that $\sum_{t=0}^{\hat{t}} u_i^t = k$ and $\sum_{t=0}^{\bar{t}} u_j^t < k$ for all $j \in N \setminus \{i\}$ and all $\bar{t} < \hat{t}$. The existing notion of winning corresponds to the situation where $k = 1$. We think it is an interesting problem to explore what happens for higher values of k.

3.6.0.4 Percolation games on three (or higher) dimensions

The set-up for percolation games can be generalized to lattices in any dimension $k \in \mathbb{N}$. Each site $\mathbf{x} = (x_1, x_2, \ldots, x_k) \in \mathbb{Z}^k$ is assigned, independently, the label of trap with probability p, the label of target with probability q, and the label of open with probability $1 - p - q$, and A can be defined as a subset of $\mathbb{Z}^k$. An even more general set-up, on arbitrary, locally finite directed graphs is addressed in §1.2 of [60], but with $q = 0$ (i.e. there are no target-labeled vertices, and hence such a percolation game is referred to as a *trapping game*). The *dimension reduction method* is illustrated in [60], showing that if the directed graph under consideration satisfies certain conditions, then the existence of multiple Gibbs states for the hard-core model on a related graph would imply that the trapping game on the directed graph has a positive probability of culminating in a draw.

It is worthwhile to ponder if there are regime(s), as function(s) of the parameters p and q, in which percolation games played on lattices of higher dimensions have probability 0 of resulting in a draw, and whether, when A is suitably defined, the occurrence of draw in such a game can be connected with the ergodicity of a PCA, possibly deduced via the recurrence relations arising from the game, the same way as illustrated in this chapter. An immediate problem to consider can be where we set $A = \{(0,0,1), (1,0,1), (0,1,1)\}$ and the same equivalence relation as ours (that is, $(a_1, \ldots, a_k) \sim (b_1, \ldots, b_l)$ if $\sum_{j=1}^{k} a_j = \sum_{j=1}^{l} b_l$). This leads to a nice class of percolation games on $\mathbb{Z}^3$. Attempting to implement the technique of weight functions to find conditions on p and q that guarantee ergodicity of the PCA deduced from the recurrence relations of this game would be an interesting problem for the near future.

3.6.0.5 Monotonicity of draw probability

An important problem, in our opinion, of a slightly different flavour, is the analysis of monotonicity properties, if any, of the probabilities of draw in percolation games with respect to, in some sense, the degree of '*mixing*' of the actions / moves. Consider a class of percolation games on $\mathbb{Z}^2$, where each site (x,y) is assigned the label $X_{(x,y)}$, with $X_{(x,y)}$ i.i.d. over all $(x,y) \in \mathbb{Z}^2$ and $X_{(x,y)}$ equaling trap with probability p, target with probability q, and open with probability $1-p-q$. We characterize the games we are concerned with in this discussion by two parameters: k and ℓ, where $k \geqslant 2$ is the number of actions allowed for each player, and $1 \leqslant \ell \leqslant k$ is a mixing parameter. For every position of the game (x,y) in $\mathbb{Z}^2$ (i.e. (x,y) is where the token is currently located), the action a_i leads to the site $((x-1)\ell+i, y+1)$, for all $1 \leqslant i \leqslant k$. In other words, here, the set of outgoing edges $\mathrm{Out}(x,y)$ from (x,y) is $\{((x-1)\ell+i, y+1) : 1 \leqslant i \leqslant k\}$.

Note that, for any two '*neighboring*' game positions (x,y) and $(x+1,y)$, the cardinality of $\mathrm{Out}(x,y) \cap \mathrm{Out}(x+1,y)$ is precisely $k-\ell$. In particular, given k, when $\ell = 1$, we obtain the k-neighbor percolation game on the integer lattice $\mathbb{Z}^2$ (for instance, [60] considers $k=2$, and this chapter is concerned with $k=3$), and when $\ell = k$, we obtain the percolation game on a rooted k-regular tree. It has been shown in [63] that for arbitrary k, when $\ell = k$, there are positive values of p and / or q for which the probability of draw is strictly positive in the percolation game played on a rooted k-regular tree (more precisely, the probability of draw is strictly positive whenever $(1-p-q)(1-q)^{k-1} \leqslant (k+1)^{k-1}k^{-k}$). At the other extreme, for $k = 2,3$ and $\ell = 1$, it has been shown in [60] and this section that both p and q need to equal 0 for the probability of draw in the corresponding percolation game to be strictly positive. These findings raise the following important questions: does draw become less likely

1. as ℓ decreases (i.e. there is a greater degree of mixing) for a given k,

2. as k increases while ℓ remains fixed (for instance, when ℓ remains fixed at the value $\ell = 1$).

It is worth emphasizing that for a given, arbitrary configuration of trap / target / open labels (i.e. a given realization of the i.i.d. random variables $(X_{(x,y)})_{(x,y)\in\mathbb{Z}^2}$), no monotonicity can be found for the event of draw with respect to ℓ in the two situations mentioned above, implying that a coupling argument will not be sufficient to prove such results. We believe that a more involved combinatorial argument will be needed for addressing these questions, and currently, we are pursuing our research exploring these questions. Exploring such monotonicity questions for just the *trapping game* (where $q=0$) or just the *target game* (where $p=0$) is expected to be quite challenging.

3.6.1 Discussion on the value of a game

The *value* of a zero-sum game is a well-known quantity in game theory, but considering readers from various backgrounds, we provide its definition here. The value of a zero-sum game is said to exist when the utilities for player 1, obtained via the following two 'ways of thinking', turn out to be equal to each other, and the value is defined as the common utility (whenever it exists). In the first way of thinking, player 1 believes that player 2's only intention is to 'hurt' her, that is, whatever strategy σ_1 she plays, player 2 will play a strategy σ_2 that minimizes her utility given σ_1. Let $U^{\max\min}$ be the maximum utility player 1 can guarantee given this belief. Formally, $U^{\max\min} = \max_{\sigma_1 \in \Sigma_1} \min_{\sigma_2 \in \Sigma_2} u_1(\sigma_1, \sigma_2)$, where Σ_1 and Σ_2 are the sets of all possible strategies for players 1 and 2 respectively. In the second way of thinking, whatever strategy σ_2 is played by player 2, player 1 plans to play a strategy σ_1 that maximizes her utility given σ_2, and believes that player 2 will play the strategy that minimizes her utility when she plays according to this plan. Let $U^{\min\max} = \min_{\sigma_2 \in \Sigma_2} \max_{\sigma_1 \in \Sigma_1} u_1(\sigma_1, \sigma_2)$ be the utility player 1 receives according to this belief. The value of a game exists when $U^{\min\max} = U^{\max\min}$, and this common utility is called the value of the game.

In [52], a general notion of zero-sum percolation games on $\mathbb{Z}^d$ is defined, with two players making alternating moves. [5] In contrast to our notion of stage games where exactly one player makes a move in each round, [52] defines a stage game in which each round comprises a move by player 1 followed immediately by a move by player 2. Each of the players receives some utility at the end of each stage game. An n-stage game, for $n \in \mathbb{N}$, consists of n stage games played sequentially where the utility of each player is defined as the time-average (where time refers to the number of stages) of the total utility from all the stage games of that player. The utility function of an n-stage game can be defined using our framework by requiring that $u_i^t(\sigma_N) = 0$ for all $t \in 2\mathbb{N} - 1$ and all $t > n$, and defining $u_i(\sigma_N) = \frac{1}{n} \sum_{t \in T} u_i^t(\sigma_N)$.

For i.i.d. and oriented percolation games, [52] shows that as n tends to infinity, the sequence of (random) values of n-stage games converges almost surely to a constant, and the expected value converges at a rate of $O(\ln(n) n^{-1/2})$. They further show that the assumption of i.i.d. is necessary for this result. It is worth mentioning that the games we have considered in our work are i.i.d. and oriented. We refer the reader to [52] for a formal definition of i.i.d. and oriented percolation games.

In [46], a general class of two-player zero-sum games is considered where, in addition to the utilities (that are defined as the capacities corresponding to any history), in every round of the game, the mover and the set of available actions are both random and i.i.d. over the histories. The three

[5] Informally speaking, a percolation game requires, among other criteria, the (random) utility functions to be stationary and ergodic.

random variables, i.e. the chosen mover, the chosen set of available actions, and the utilities, need not be independent of each other corresponding to a given history. The games in [46] are played on (finitely branching) trees for each possible realization of the above-mentioned random variables. The utility for player 1 corresponding to a strategy profile is the infimum of the stage utilities along the game path induced by the strategy profile. In this setting, a natural measure space is defined over the games, and a (fixed point) characterization of the cumulative distribution function for the value of the game is provided (together with a number of corollaries and applications).

Although [3] does not deal with the value of a game nor its probability of draw, uses an approach involving ideas from percolation theory which makes it relevant for percolation games. In [3], random games are considered where the number of players is large, each player has two strategies, and the utilities are i.i.d. with possible ties (i.e. having an atomic distribution). In this setting, [3] explores the possibility of existence of a (pure) Nash equilibrium which, in contrast with mixed Nash equilibrium, may not always exist. In particular, asymptotic results regarding the (random) number of Nash equilibria and a central limit theorem for given bounds on the probability of ties are established. Furthermore, using tools from percolation theory, the geometry of the set of Nash equilibria is determined, and it is shown that the best response dynamics converges to a Nash equilibrium with high probability when the number of players is large and the probability of ties is positive but bounded above (such as in the case of potential games). [6]

3.7 The scope of the weight function technique

Let us define the critical region for a percolation game as the set of all values of $(p,q) \in \mathscr{S}$ (where $\mathscr{S}$ is as defined in (3.1.1)) for which the probability of draw is 0 (and hence, the corresponding PCA is ergodic). It follows from this chapter and [60] that the critical region for the percolation games considered in both these works is $\mathscr{S}$.

We begin with noting that the weight function technique provides only a sub-region, normally in terms of a lower bound, of the critical region of a percolation game. In case of our percolation games or that in [60], this sub-region turns out to be 'full', i.e. it covers all possible values of (p,q), and consequently, the weight function technique succeeds in providing a characterization of the *entire* critical region. In general, unless the critical region is full in the above-mentioned sense, the weight function technique itself will not be suffice to characterize the critical region, which, in our view, is the main weakness of this technique. Nevertheless, providing a sub-region of the

[6]More precisely, [3] establishes a connection between percolation and random oriented graphs through a coupling where the set of strategies that are accessible by the best response dynamics coincides with the connected component containing the origin for percolation on the hypercube.

critical region is, in general, an important contribution to the literature (in fact, it may well be the case that finding such a sub-region becomes the first breakthrough in answering questions about the critical regions in several open problems), and in view of that, we now proceed to describe some situations where one can apply this technique.

Recall that the weight function technique, when applicable, only shows that certain events have probability 0 under *any* stationary distribution corresponding to the PCA under consideration. One then needs to do further computations to deduce how this actually leads to the conclusion that the event of draw has probability 0 under *any* stationary distribution. The crucial step here is to link the ergodicity of the PCA under consideration with the event of draw, which may not be straightforward unless the PCA has some game theoretic structure. In what follows, we define a class of 'game theoretic' PCAs for which we believe that the weight function technique can be applied to establish ergodicity. For PCAs outside of this class, our guess is that the applicability of the weight function technique will depend on the particular structure of the PCA's stochastic update rule, but as of now, we are unable to specify what structural conditions on the update rule guarantee such applicability.

We call a d-dimensional deterministic CA F, with universe $\mathbb{Z}^d$ and a given neighborhood $\mathcal{N}$ that is a finite subset of $\mathbb{Z}^d$, *game-theoretic* if it satisfies the following criteria:

1. the alphabet $\widehat{\mathscr{A}}$ comprises the symbols W, D and L;
2. letting $\eta_t = \left(\eta_t(\mathbf{x}) : \mathbf{x} \in \mathbb{Z}^d\right)$ denote the configuration of states at time t, where $\eta_t(\mathbf{x})$ denotes the state of $\mathbf{x}$ at time t, the state $\eta_{t+1}(\mathbf{x})$ of *any* site $\mathbf{x} \in \mathbb{Z}^d$ at time $t+1$ is decided according to the following rules:
   (a) if $\eta_t(\mathbf{x}+\mathbf{y}) = W$ for *each* $\mathbf{y} \in \mathcal{N}$, then $\eta_{t+1}(\mathbf{x}) = L$,
   (b) if there exists at least one $\mathbf{y} \in \mathcal{N}$ for which $\eta_t(\mathbf{x}+\mathbf{y}) = L$, then $\eta_{t+1}(\mathbf{x}) = W$,
   (c) if there exists *no* $\mathbf{y} \in \mathcal{N}$ for which $\eta_t(\mathbf{x}+\mathbf{y}) = L$, but there exists at least one $\mathbf{z} \in \mathcal{N}$ such that $\eta_t(\mathbf{x}+\mathbf{z}) = D$, then $\eta_{t+1}(\mathbf{x}) = D$;
3. conditioned on η_t, the updates $\eta_{t+1}(\mathbf{x})$ happen independently over all $\mathbf{x} \in \mathbb{Z}^d$.

We obtain a *game-theoretic PCA* by perturbing the above CA in a manner similar to what we have considered in this work: given parameters p and q, each outcome (of the above-mentioned deterministic CA) that is not equal to W is flipped to W with probability p, and each outcome that is not equal to L is flipped to L with probability q. More formally, letting $\hat{\eta}_t = \left(\hat{\eta}_t(\mathbf{x}) : \mathbf{x} \in \mathbb{Z}^d\right)$ denote the (random) configuration of states at time t,

1. we let $\hat{\eta}_{t+1}(\mathbf{x})$ equal L with probability $1-p$ and W with probability p when $\eta_{t+1}(\mathbf{x}) = L$;

2. we let $\hat{\eta}_{t+1}(\mathbf{x})$ equal W with probability $1-q$ and L with probability q when $\eta_{t+1}(\mathbf{x}) = W$;

3. we let $\hat{\eta}_{t+1}(\mathbf{x})$ equal W with probability p, L with probability q and D with probability $1-p-q$ when $\eta_{t+1}(\mathbf{x}) = D$.

We observe, crucially, that for game-theoretic PCAs, both parts of Lemma 3.3.1 hold (and this can be proved in an identical manner as shown in this section). The first part of Lemma 3.3.1 leads to the conclusion, via a proof identical to that of Proposition 2.1 of [60], that F is ergodic if and only if the PCA A, obtained by restricting the alphabet $\hat{\mathscr{A}}$ to the sub-alphabet $\mathscr{A} = \{W, L\}$, is ergodic. Moreover, the unique stationary or limiting distribution μ of F, when F is ergodic, is the same as that of A, and therefore, μ must assign probability 0 to the symbol D, i.e. the probability, under μ, of the event that the symbol D occupies the site $\mathbf{x}$ is 0 for *every* $\mathbf{x} \in \mathbb{Z}^d$. Applying the technique of weight functions to the PCA F, provided the process of constructing the weight function is tractable either manually or with the aid of a computer, can reveal conditions on the parameter pair (p,q) under which *any* stationary distribution μ for F assigns probability 0 to the symbol D.

We believe that the method of weight function may even be of aid in case of percolation games (on lattices) whose recurrence relations do not necessarily yield a PCA, but something more general. For instance, suppose we consider the same random premise as in this work, i.e. each site of $\mathbb{Z}^2$ is assigned, independently, a label that reads trap with probability p, target with probability q, and open with probability $1-p-q$, but here, for each site $(x,y) \in \mathbb{Z}^2$, we set $\text{Out}(x,y) = \{(x+2,y),(x+2,y+2),(x,y+2)\}$. As before, two players take turns to make moves, where a move involves relocating the token from where it is currently located, say some site (x,y), to any of the sites in $\text{Out}(x,y)$. The outcome of this game is decided following the same rules as the percolation game that we study in this work This problem has also been mentioned in §3.6.0.1, when discussing the possibility of adding vertices at levels that are higher than the next level.

It is immediate that the recurrence relations arising from the above game cannot be represented via a PCA, but could be by a PCA of memory two as defined in [29]. A little pondering reveals that the only way to partition the lattice $\mathbb{Z}^2$ into "level sets" that are meaningful for our analysis of this game is to consider the diagonal lines $D_k = \{(x, k-x) : x \in \mathbb{Z}\}$, identify D_k with $\mathbb{Z}$ via the mapping $(x, k-x) \mapsto x$ for each $x \in \mathbb{Z}$, and consider the following discrete-time stochastic process $\{\eta_t\}_{t \in \mathbb{N}_0}$ where each η_t is a random configuration taking values in $\hat{\mathscr{A}}^{\mathbb{Z}} = \{W, D, L\}^{\mathbb{Z}}$. Conditioned on η_{t-1} and η_t, the random variables $\eta_{t+1}(x)$ are defined independently over all $x \in \mathbb{Z}$, with the distribution for $\eta_{t+1}(x)$ being a function of $\eta_t(x)$, $\eta_{t-1}(x+2)$ and $\eta_t(x+2)$, as follows:

1. if $\eta_t(x) = \eta_{t-1}(x+2) = \eta_t(x+2) = W$, then $\eta_{t+1}(x) = L$ with probability $1-p$ and $\eta_{t+1}(x) = W$ with probability p,

2. if at least one of $\eta_t(x)$, $\eta_{t-1}(x+2)$ and $\eta_t(x+2)$ equals L, we set $\eta_{t+1}(x) = W$ with probability $1-q$ and $\eta_{t+1}(x) = L$ with probability q,

3. if none of $\eta_t(x)$, $\eta_{t-1}(x+2)$ and $\eta_t(x+2)$ equals L but at least one of them equals D, we set $\eta_{t+1}(x) = W$ with probability p, $\eta_{t+1}(x) = L$ with probability q, and $\eta_{t+1}(x) = D$ with probability $1-p-q$.

It is evident that $\{\eta_t\}_{t\in\mathbb{N}_0}$ does not constitute a Markov chain. However, it is Markov of order 2 (i.e. retains as memory the configurations of the previous 2 time steps), and a probability distribution μ supported on $\hat{\mathscr{A}}^{\mathbb{Z}} \times \hat{\mathscr{A}}^{\mathbb{Z}}$ is said to be stationary or invariant for this stochastic process if the following is true: if the joint distribution of (η_{t-1}, η_t) is μ, then the joint distribution of (η_t, η_{t+1}) is μ as well. We then say that $\{\eta_t\}_{t\in\mathbb{N}_0}$ is ergodic if

1. it possesses a unique stationary distribution μ, and

2. if ν denotes *any* probability distribution supported on $\hat{\mathscr{A}}^{\mathbb{Z}} \times \hat{\mathscr{A}}^{\mathbb{Z}}$, and if the joint distribution of (η_0, η_1) is ν, then the joint distribution of (η_t, η_{t+1}) converges to μ as $t \to \infty$.

It should be possible to prove, along the same lines as Proposition 2.2 of [60] has been argued, that the percolation game described above has probability 0 of culminating in a draw if and only if the process $\{\eta_t\}_{t\in\mathbb{N}_0}$ is ergodic. Once again, it becomes important to investigate the critical region for this percolation game. We expect that the technique of weight functions is applicable in this set-up. Here, we must look for a suitable function w that is defined on the space of probability measures supported on $\hat{\mathscr{A}}^{\mathbb{Z}} \times \hat{\mathscr{A}}^{\mathbb{Z}}$, and we suspect that we have to consider measures of Carteisan products of cylinder sets (in other words, we have to consider events of the form $(\eta_{t-1}(x) = a_x \text{ for all } x \in A, \eta_t(y) = b_y \text{ for all } y \in B)$, where A and B are finite subsets of $\mathbb{Z}$, and $a_x, b_y \in \hat{\mathscr{A}}$ for all $x \in A$ and $y \in B$).

Chapter 4

Generalized percolation games on $\mathbb{Z}^2$ and corresponding probabilistic cellular automata

Preface

This chapter is based on the following paper:

- Bhasin D., Karmarkar S., Podder M., Roy S. "Generalized percolation games on the 2-dimensional square lattice, and ergodicity of associated probabilistic cellular automata" – submitted. Preprint available **here**.

4.1 Introduction

The study of *oriented Bernoulli bond percolation* on $\mathbb{Z}^d$ (i.e. the infinite graph whose vertex set is $\mathbb{Z}^d$ and in which two vertices $\mathbf{x} = (x_1, x_2, \ldots, x_d)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_d)$ are adjacent if and only if the Euclidean distance between them is 1) is now more than six decades old. For each $i \in \{1, 2, \ldots, d\}$, let $\mathbf{e}_i \in \mathbb{Z}^d$ denote the vertex in which the i-th coordinate equals 1, while all other coordinates equal 0. For each $\mathbf{x} \in \mathbb{Z}^d$, let the edge between $\mathbf{x}$ and $\mathbf{x} + \mathbf{e}_i$, denoted henceforth as $(\mathbf{x}, \mathbf{x} + \mathbf{e}_i)$, be directed *from* $\mathbf{x}$ *towards* $\mathbf{x} + \mathbf{e}_i$. Each such directed edge is now labeled *closed* with probability r, and *open* with probability $(1 - r)$, for some $r \in [0, 1]$. The question of interest is: for what values of r does there exist, with positive probability, an infinite path that begins from the origin $\mathbf{0}$ and consists only of open, directed edges? In other words, for which values of r does *percolation happen*? Introduced in [25], this problem has since been studied extensively in [54], [40], [39] etc. The existence of a *critical value* $r_c(d) \in (0, 1)$, such that percolation happens for $r < r_c(d)$ and percolation does *not* happen when $r > r_c(d)$, has been well-known. However, the exact value of $r_c(d)$ is seldom known. For instance, in [19], it has been shown, via rigorous methods, that $r_c(2) < 0.647$, while Monte Carlo simulations suggest that $r_c(2) \approx 0.6445$.

The work in this chapter falls into the setup of percolation games as studied described in §1.1 and began in an attempt to understand what happens when we incorporate an *adversarial element* into the set-up of oriented bond percolation mentioned above. In this chapter, we imagine a percolation game on $\mathbb{Z}^2$ with two players and a token that is placed at the origin, $(0, 0)$, at the beginning of the game. The two players take turns to make *moves*, and each such move constitutes a *round* of the game. Each player, when it is her turn to make a move, tries to relocate the token from where it is currently located, say $(x, y) \in \mathbb{Z}^2$, to one of $(x + 1, y)$ and $(x, y + 1)$ (that is, $\mathrm{Out}(x, y) = \{(x + 1, y), (x, y + 1)\}$), but relocation is only allowed along edges that are open. A player loses if she is unable to move, i.e. her opponent has succeeded in moving the token to some vertex (x, y) such that both the directed edges $\big((x, y), (x + 1, y)\big)$ and $\big((x, y), (x, y + 1)\big)$ are closed. The game may continue indefinitely (i.e. neither of the two players is able to clinch the victory within a finite number of rounds of the game), in which case we say that the game has resulted in a draw. Here, we ask the question: for what values of r does this game result in a draw with positive probability? Note how closely *this* question ties in with the question regarding the occurrence of percolation that we posed in the previous paragraph. The existence of an *infinite, open path* starting at the origin, i.e. an infinite path consisting of open, directed edges, is necessary for even the remotest possibility of a draw, but the existence of an *arbitrary* infinite, open path does not, necessarily, ensure the occurrence of a draw. In particular, for draw to happen, there must exist an infinite, open path that acts as an 'equilibrium path' – i.e. a path from which neither of the two players is willing to deviate under the common belief of rationality. We show, in this chapter, that for r as low as 0.157176, the probability of draw in the game described above is 0. A comparison of this result with the estimate of $r_c(2)$ mentioned in the previous paragraph reveals how fascinating the study of such games is, and how vastly the inclusion of an adversarial flavour impacts the original percolation process.

Even though the set-up described in the previous paragraph is what the work in this chapter commenced with, the contents of this chapter do not remain limited to this one-parameter percolation game. The first step towards generalizing this game can be brought about by classifying the edges of our graph into the following *three* categories, instead of two:

1. a player is *penalized* if she moves the token along a directed edge that has been labeled a *trap* (this is the equivalent of the label 'closed' in the previous set-up), which happens with probability r,

2. a player is *rewarded* if she moves the token along a directed edge that has been labeled a *target*, which happens with probability s,

3. and a player is neither rewarded nor punished if she moves the token along a directed edge that has been labeled *open*, which happens with the remaining probability $(1-r-s)$.

In other words, a player, when it is her turn to move, relocates the token, as before, from where it is currently located, say $(x,y) \in \mathbb{Z}^2$, to one of $(x+1,y)$ and $(x,y+1)$, but *this* game can come to an end in one of *two* different ways: either one of the players succeeds in moving the token along an edge that has been labeled a target, and thereby wins the game, or one of the players is forced to move the token along an edge that has been labeled a trap, and thereby loses the game. Note that the previously described one-parameter percolation game is a special case of this two-parameter percolation game, obtained by setting $s = 0$, and, as mentioned in the previous paragraph, we have shown in this chapter that the probability of draw in the one-parameter percolation game equals 0 whenever $r \geqslant 0.157176$. For the two-parameter set-up, we establish two different regimes of values of the parameter-pair (r,s) for which the probability of draw is guaranteed to be 0:

1. where $r = s > 0.10883$, and

2. where each of r and s is sufficiently small (how small they need to be comes out as a consequence of the various steps involved in the main, technical part of the proof, outlined in §4.9.2 – but it suffices for us to take each of r and s to be less than or equal to $1/50$), and $3(1-s)(2s-s^2)^2(1-6s+3s^2) \geqslant 4r$.

These results have been stated formally in Theorem 4.2.4 (the regime of values of (r,s) covered by (2) above is slightly smaller, i.e. the inequality is stronger, than what we actually deduce in (B1) of Theorem 4.2.4 – the stronger, and simpler, inequality has been stated here for the sake of simplicity and lucid reading).

Our endeavour to generalize the notion of percolation games does not end here – we go a step further and consider a *three-parameter* set-up which, in some sense, combines the notion of *site percolation* and *bond percolation*, as follows. Each vertex $(x,y) \in \mathbb{Z}^2$ is now assigned, independent of all else, a label that reads *trap* with probability p, *target* with probability q, and *open* with the remaining probability $(1-p-q)$. Additionally, each edge is assigned, independent of all else, a label that reads *trap* with probability r and *open* with the remaining probability $(1-r)$. The permitted moves by the players are the same as those described in the previous paragraph, but now, a player can clinch the victory in one of *three* possible ways:

1. either she succeeds in moving the token to a vertex labeled a target,

2. or she is able to force her opponent to move the token to a vertex labeled a trap,

3. or her opponent is compelled to move the token along an edge labeled a trap.

As before, the game continues for as long as one of the players does not emerge the victor, and the result is a draw if the game continues indefinitely. We establish a regime of values of (p,q,r), described in detail in Theorem 4.2.1, for which the probability of draw in this game is guaranteed to equal 0.

It is straightforward to see that draw is the only outcome possible when $p=q=r=0$, and it happens with probability 0 when $p+q=1$ or when $r=1$. As indicated in Chapter 7 of [37], an important and challenging question concerns itself with what happens to the probability of draw (and, as we shall introduce subsequently, to the ergodicity of a suitably defined probabilistic cellular automaton) for arbitrarily small values of the probabilities of trap and target (for both vertices and edges), i.e. when each of p, q and r is arbitrarily small. The crucial contribution of Theorem 4.2.1 (as also discussed in §4.3.3) lies in its coverage of a significant subset of values of (p,q,r) in a small neighbourhood of $(0,0,0)$, and pictorial illustrations of what this subset looks like have been provided in Figure 4.2 following the statement of Theorem 4.2.1. Theorem 4.2.1 provides four such regions, or *regimes*, each specified by some inequalities involving p,q, and r, such that the probability of draw equals 0 whenever (p,q,r) belongs to one of these regions. As these inequalities are somewhat involved, we provide here some simpler and stronger inequalities (in other words, some *sub-regimes* of the regimes covered by Theorem 4.2.1) to elucidate said contribution of the theorem to the reader. [1] A simpler (and stronger) version of the inequalities in (4.2.1) and (4.2.2) of Theorem 4.2.1 is given by

$$p \leqslant \frac{q}{2} \quad \text{and} \quad p + q \geqslant 2r, \tag{4.1.1}$$

and likewise, a simpler and stronger version of the inequalities in (4.2.1) and (4.2.5) of Theorem 4.2.1 is given by

$$q + r \leqslant p \quad \text{and} \quad 5q \geqslant 4r. \tag{4.1.2}$$

Note that (4.1.1) tells us that the probability of draw is 0 when $q = \varepsilon$, $p = \alpha\varepsilon$ and $r \leqslant (1+\alpha)\varepsilon/2$, where $\alpha \leqslant 1/2$ and ε is an arbitrarily small positive real. Similarly, (4.1.2) reveals that the probability of draw is 0 when $p = \varepsilon$, $q = \alpha\varepsilon$ and $r = \beta\varepsilon$ where $\alpha + \beta \leqslant 1$ and $5\alpha \geqslant 4\beta$, and ε is an arbitrarily small positive real. To demonstrate, pictorially, the 3-dimensional region that all (p,q,r) satisfying one of (4.1.1) and (4.1.2) cover, we refer the reader to Figure 4.1 (in each of (A) and (B) of Figure 4.1, we take each of p, q and r to be bounded above by $1/50 = 0.02$, although this could be optimized further, since the claims made in Theorem 4.2.1 assuredly hold when p, q and r are this small). Furthermore, the inequalities (4.2.1) and (4.2.3) of Theorem 4.2.1 together imply, as a special case, that draw happens with probability 0 when $p = q = r = \varepsilon$ for $\varepsilon > 0$ arbitrarily small. Yet another contribution of Theorem 4.2.1 stems from the fact that, setting $r = 0$, the inequalities in (4.2.1), (4.2.2) and (4.2.5) together imply that the probability of draw equals 0 whenever at least one of p and q is strictly positive – thereby allowing Theorem 1 of [60] to be derived as a special case of Theorem 4.2.1. These instances go to show how Theorem 4.2.1 establishes that the probability of draw is 0 for a fairly wide region of values of (p,q,r) arbitrarily close to $(0,0,0)$. Since the probability of draw is 1 at $(0,0,0)$, our result seems to indicate (and proves it under certain (relative) restrictions on p, q and r, such as when $p = q = r$) that a phase transition happens exactly at $p = q = r = 0$.

4.1.1 Why the three-parameter percolation game is an all-encompassing set-up

An obvious question to ask at this point is, why, in the set-up for the three-parameter percolation game, *target* is not included in the list of labels assigned to the edges of our graph. For any $(x,y) \in \mathbb{Z}^2$, if at least one of its outgoing edges, $\big((x,y),(x+1,y)\big)$ and $\big((x,y),(x,y+1)\big)$, has been labeled a target, a player knows that she will win the game the moment her opponent moves

[1] The actual conditions differ only with respect to second or higher order terms (i.e. terms of the form $p^i q^j r^k$ where $i,j,k \in \mathbb{N}_0$ and $i+j+k \geqslant 2$).

(a) 3D plot of (4.1.1)

(b) 3D plot of (4.1.2)

Figure 4.1: Regions covered by (4.1.1) and (4.1.2)

the token to (x,y), *unless* (x,y) has already been labeled a target. Consequently, seen from her opponent's perspective, except for the case where (x,y) has been labeled a target, it is as if (x,y) *acts* like a vertex that has been labeled a trap. This equivalence has been formalized in the following paragraph.

Consider the following policy of (random) assignment of labels to our graph: each vertex is labeled, independently, a trap with probability p', a target with probability q', and open with probability $(1-p'-q')$, and each edge is labeled, independently, a trap with probability r', a target with probability s', and open with probability $(1-r'-s')$. Let (σ,η) indicate a realized assignment of labels obtained by implementing this policy, with $\sigma = (\sigma((x,y)) : (x,y) \in \mathbb{Z}^2)$ denoting the tuple of vertex labels, and η denoting the tuple of edge labels, $\eta\big((x,y),(x+1,y)\big)$ and $\eta\big((x,y),(x,y+1)\big)$, for each $(x,y) \in \mathbb{Z}^2$. We now come up with a modified labeling, (σ',η'), as follows: if, for any $(x,y) \in \mathbb{Z}^2$, the label $\eta\big((x,y),(x+1,y)\big)$ equals a target, we set

$$\eta'\big((x,y),(x+1,y)\big) = \begin{cases} \text{trap with probability } r'(1-s')^{-1} \\ \text{open with probability } (1-r'-s')(1-s')^{-1} \end{cases}$$

$\big($an analogous re-labeling is applied if $\eta\big((x,y),(x,y+1)\big)$ is a target, instead$\big)$, and we set

$$\sigma'(x,y) = \begin{cases} \sigma(x,y) & \text{if } \sigma(x,y) \in \{\text{trap}, \text{target}\} \\ \text{trap} & \text{if } \sigma(x,y) = \text{open}. \end{cases}$$

As has been explained in §4.3.2, this modification reduces what was initially a four-parameter percolation game into a three-parameter one in which p indicates the probability of a vertex being labeled a trap, q indicates the probability of a vertex being labeled a target, and r indicates the probability of an edge being labeled a trap, with

$$p = p' + (1 - p' - q')\left(2s' - s'^2\right), \quad q = q' \quad \text{and} \quad r = \frac{r'}{1 - s'}.$$

This modification allows the reader to see why, when considering the most generalized framework for percolation games on the 2-dimensional infinite square lattice, it suffices to work with only *three* categories of labels instead of four (such as we do in this work, namely, 1. trap for vertices, 2. target for vertices, 3. and trap for edges).

4.1.2 Organization of the rest of the chapter

We describe here the organization of the rest of this chapter. In §4.2, we formally describe the two games, *generalized percolation games* and *bond percolation games*, that we study in this chapter, although the reader has already been accorded an informal introduction to these. The two main results of this chapter, pertaining to the probability of draw in each of these games, have been stated in Theorems 4.2.1 and 4.2.4. The probabilistic cellular automata (PCAs) arising as a result of the game rules have been formally described in §4.3, and the main result pertaining to their ergodicity properties has been stated in Theorem 4.3.1. We draw the reader's attention to §4.3.2, which shows how the bond percolation game (and correspondingly, the PCA, $\widehat{E}_{r',s'}$, arising out of its game rules) can be obtained as a special case of the generalized percolation game (and correspondingly, the PCA, $\widehat{G}_{p,q,r}$, arising out of its game rules). One of this work's most significant contributions (see also the less formal discussion preceding §4.1.1) has been outlined in detail in §4.3.3. We have dedicated §4.4 and §4.5 to discussions on the motivations propelling the study of these games and these PCAs, respectively. The game rules, and how they give rise to the PCAs, $\widehat{G}_{p,q,r}$ and $\widehat{E}_{r',s'}$, have been detailed in §4.6. The two most technical results of this work, namely, Theorems 4.7.1 and 4.7.2, have been stated in §4.7, along with the proofs of Theorems 4.2.1, 4.2.4 and 4.3.1 assuming the claims made in Theorems 4.7.1 and 4.7.2 to be true. The proof of Theorem 4.7.1 has been carried out in §4.8. We mention here that while the proof *specifically* of Theorem 4.7.1 has been outlined in detail, step by step, in §4.8.3 (the final weight function, the corresponding final weight function inequality, and the deduction of the claim made in Theorem 4.7.1, have been included in §4.8.2), a very general idea on how to construct a suitable weight function that serves our desired purpose has been outlined in §4.8.1. Likewise, the final weight functions, the corresponding weight function inequalities, and the deduction of the claims stated in Theorem 4.7.2, have been included in §4.9.1, while the details of the proofs have been covered in §4.9.2, §4.9.3

and §4.9.4.

4.2 Formal description of our games and the main results

We shall use $\mathbb{Z}^2$ to indicate both the infinite 2-dimensional square lattice graph and its set of vertices. Each edge of this graph is either of the form $\big((x,y),(x+1,y)\big)$ (directed from (x,y) towards $(x+1,y)$) or of the form $\big((x,y),(x,y+1)\big)$ (directed from (x,y) towards $(x,y+1)$), for $(x,y) \in \mathbb{Z}^2$. To each $(x,y) \in \mathbb{Z}^2$, we assign, independently, a label that reads *trap* with probability p, *target* with probability q, and *open* with probability $(1-p-q)$, while to each directed edge, we assign, independently, a label that reads *trap* with probability r and *open* with probability $(1-r)$. We assume that at least one of p, q and r is strictly positive, so that our parameter-space Θ is given by $\Theta = \{(p,q,r) \in [0,1]^3 : p+q+r > 0, p+q \leqslant 1\}$. The resulting graph, bearing (random) vertex-labels and (random) edge-labels, is referred to as a *random board*, on which our *generalized percolation games* are played. Any realization of vertex-labels and edge-labels, assigned as described above, is referred to as a *configuration* on $\mathbb{Z}^2$, and we *fix* a configuration *before* the game begins. A generalized percolation game is a *two-player combinatorial game*, which tells us that it is a game of *perfect information* – both players are fully aware of the environment (in this case, the environment is the configuration fixed before the game begins) on which the game is being played.

A token is placed at a vertex of $\mathbb{Z}^2$, referred to as the *initial vertex*, just before the game starts. Two players take turns to make moves, where a move involves relocating the token from where it is currently located, say $(x,y) \in \mathbb{Z}^2$, to one of $(x+1,y)$ and $(x,y+1)$. A player wins if she 1. either succeeds in moving the token to a vertex labeled a target, 2. or forces her opponent to move the token to a vertex labeled a trap, 3. or forces her opponent to move the token along an edge labeled a trap. The game continues for as long as the token stays *on* vertices labeled open and keeps being moved *along* edges labeled open, and this may happen indefinitely, leading to a draw. Of primary interest to us is the investigation of *regimes* of values of the parameter-triple (p,q,r) for which the probability of draw equals 0.

We assume that, if the game is destined to end in a finite number of rounds, then the player who is destined to win tries to achieve her victory as quickly as possible, while her opponent attempts to stall and prolong the game as much as possible. It is also crucial to note that the roles of the players are symmetric in this game – a fact that heavily impacts the simplicity of the recurrence relations that we shall study in §4.6.

Theorem 4.2.1. *There exist ε_1, ε_2 and ε_3, in $(0,1)$, such that whenever $(p,q,r) \in \Theta$ with $p \leqslant \varepsilon_1$, $q \leqslant \varepsilon_2$ and $r \leqslant \varepsilon_3$, the inequality in (4.2.1) holds, and (p,q,r) satisfies the constraints described in* *precisely one of* (4.2.2)*,* (4.2.3)*,* (4.2.4) *and* (4.2.5)*, the probability of the event that the generalized percolation game described above, with underlying parameters p, q and r, results in a draw, equals* 0*. Here,* (4.2.1)*,* (4.2.2)*,* (4.2.3)*,* (4.2.4) *and* (4.2.5) *are as follows:*

$$2(p+q)+6r^2+3r(p+2q) \geqslant 4r+(p+q)^2, \tag{4.2.1}$$

$$p(1-p) \leqslant q\{1-q-r(1-p-q)\}, \tag{4.2.2}$$

$$\left.\begin{aligned} &q\{1-q-r(1-p-q)\} < p(1-p) \leqslant \{q+r(1-p-q)\}\{1-q-r(1-p-q)\},\\ &p-2q+pq+qr-3pr-2p^2+3q^2 \leqslant 0, \end{aligned}\right\} \tag{4.2.3}$$

$$\left.\begin{aligned} &q\{1-q-r(1-p-q)\} < p(1-p) \leqslant \{q+r(1-p-q)\}\{1-q-r(1-p-q)\},\\ &p-2q+pq+qr-3pr-2p^2+3q^2 > 0,\\ &p+4q+9pr+5qr+6r^2 \geqslant 4r+p^2+4q^2+5pq, \end{aligned}\right\} \tag{4.2.4}$$

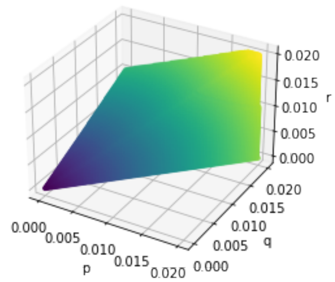
and

$$\left.\begin{aligned} &q+r(1-p-q) < p,\\ &6q+10pr+4qr+p^2+6r^2 \geqslant 4r+6pq+7q^2. \end{aligned}\right\} \tag{4.2.5}$$

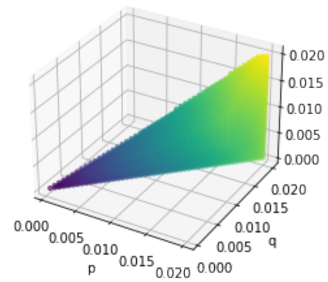
The values of ε_1, ε_2 and ε_3 (i.e. how small p, q and r need to be for Theorem 4.2.1 to hold) can be deduced from the various steps, outlined in §4.8.3, of the proof of Theorem 4.7.1 (which is instrumental in proving Theorem 4.2.1). Although better bounds on these are possible, we do not put in much effort to optimize such bounds, and we simply comment here that assuming each of ε_1, ε_2 and ε_3 to be equal to $1/50$ is enough.

In Figure 4.2, we plot the regions (of Θ) covered by the inequalities given by (4.2.1), (4.2.2), (4.2.3), (4.2.4) and (4.2.5). As the captions suggest, (A) of Figure 4.2 shows the coverage of (4.2.1) and (4.2.2), i.e. the set of all values of (p,q,r) that satisfy (4.2.1) and (4.2.2), (B) of Figure 4.2 shows the coverage of (4.2.1) and (4.2.3), and so on. The colours in each of these plots have been applied according to the distance (of the parameter-tuple under consideration) from the origin, with darker colours applied to those near the origin, and lighter colours applied to those farther away. Note that we consider only the segment from 0 to $1/50 = 0.02$ for each of the axes, since Theorem 4.2.1 assuredly holds when each of p, q and r lies in $[0,1/50]$ (along with satisfying (4.2.1) and precisely one of (4.2.2), (4.2.3), (4.2.4) and (4.2.5)). (E) of Figure 4.2 shows the union of all four regions demonstrated in (A), (B), (C) and (D) of Figure 4.2. In each of these five images, we see that the coloured region has a non-empty intersection with *every* neighbourhood around $(0,0,0)$. Since the 3D plots show only one 'surface' of the feasible regions, we demonstrate

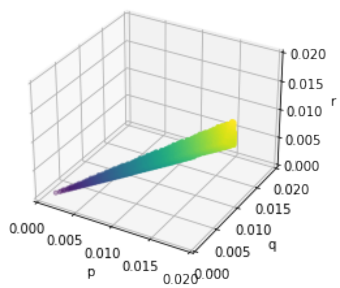
1. in (A) of Figure 4.3, the region covered by our result in Theorem 4.2.1 when the value of p



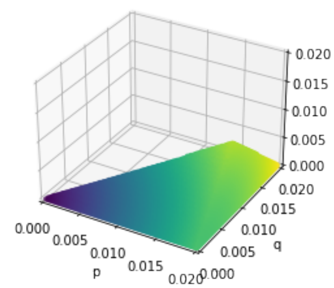
(a) 3D plot of (4.2.1) and (4.2.2)



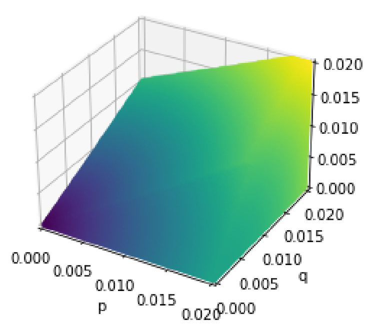
(b) 3D plot of (4.2.1) and (4.2.3)



(c) 3D plot of (4.2.1) and (4.2.4)

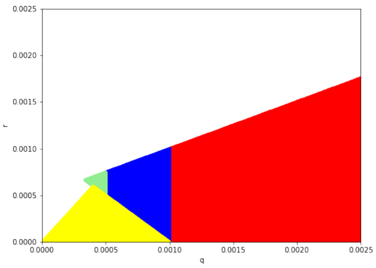


(d) 3D plot of (4.2.1) and (4.2.5)

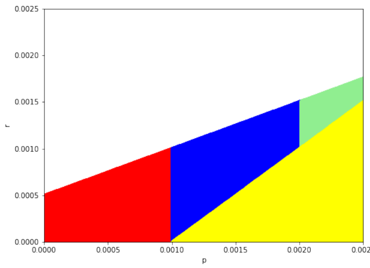


(e) 3D plot for the complete region covered by Theorem 4.2.1

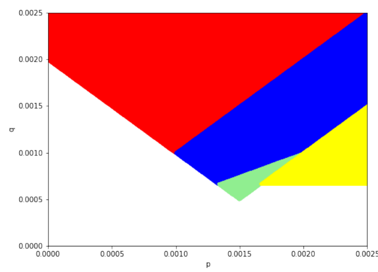
Figure 4.2: Regions covered by Theorem 4.2.1



(a) 2D plot for $p = 0.001$



(b) 2D plot for $q = 0.001$



(c) 2D plot for $r = 0.001$

Figure 4.3: 2D plots for regions covered by Theorem 4.2.1

is kept fixed at $p = 0.001$,

2. in (B) of Figure 4.3, the region covered by our result in Theorem 4.2.1 when the value of q is kept fixed at $q = 0.001$,

3. and finally, in (C) of Figure 4.3, the region covered by our result in Theorem 4.2.1 when the value of r is kept fixed at $r = 0.001$.

Note that in (A) of Figure 4.3, the four different colours used are meant to indicate the four distinct regions that are covered by Theorem 4.2.1, i.e. for $p = 0.001$, the region coloured red indicates the set of all values of (q,r) such that $(0.001,q,r)$ satisfies (4.2.1) and (4.2.2), the region coloured blue indicates the set of all values of (q,r) such that $(0.001,q,r)$ satisfies (4.2.1) and (4.2.3), the region coloured green indicates the set of all values of (q,r) such that $(0.001,q,r)$ satisfies (4.2.1) and (4.2.4), and the region coloured yellow indicates the set of all values of (q,r) such that $(0.001,q,r)$ satisfies (4.2.1) and (4.2.5). Similar interpretations apply to (B) and (C) of Figure 4.3. Notice that the four colour-coded regions in each of (A), (B) and (C) of Figure 4.3 are pairwise non-overlapping – this is as it should be, because no $(p,q,r) \in \Theta$ can satisfy all the constraints of more than one of (4.2.2), (4.2.3), (4.2.4) and (4.2.5).

Remark 4.2.2. *We draw the reader's attention to an important point here. The proof of Theorem 4.2.1 happens via Theorem 4.7.1, which, in turn, is proved via the technique of weight functions. The construction of a suitable weight function for proving Theorem 4.7.1, and thereby, Theorem 4.2.1, has been outlined in §4.8.3. As has been emphasized in §4.8.3, this weight function is not unique, and furthermore, small tweaks here and there in the long and involved process of constructing this weight function, step by step, could lead to very marginal improvements – we say "marginal" because such improvements are likely to happen via terms of the form $p^i q^j r^k$, for some $i, j, k \in \mathbb{N}_0$, such that $(i+j+k)$ equals 3 or more, and since we have already assumed that each of p, q and r is small, such terms make little difference to the regime obtained in Theorem 4.2.1.*

As explained in §4.1, the above-described three-parameter set-up is the most generalized of all the games studied in this work. Setting $r = 0$ reduces this set-up to that of the site percolation game studied in [60]. In fact, the main result (pertaining to the probability of draw in the site percolation game) of [60] follows as a corollary of Theorem 4.2.1:

Corollary 4.2.3. *Assuming that Theorem 4.2.1 holds, the generalized percolation game, with underlying parameters p, q and $r = 0$, results in a draw with probability 0 whenever $p+q > 0$.*

Proof. When $r = 0$, the first case to consider is where $p \leqslant q$, which, assuming both p and q to be bounded above by $1/2$, is equivalent to $p(1-p) \leqslant q(1-q)$ (since the function $f(x) = x(1-x)$ is strictly increasing for $x \in [0,1/2)$ and strictly decreasing for $x \in (1/2,1]$). Moreover, since $r = 0$, the inequality in (4.2.1) reduces to $2(p+q) \geqslant (p+q)^2$, which is indeed true since $p+q \leqslant 1$. Thus, in this case, $(p,q,0)$ satisfies the constraints of (4.2.1) and (4.2.2), leading to the conclusion, from Theorem 4.2.1, that the probability of draw in the corresponding generalized percolation game equals 0.

The second case to consider is where $p > q$. Note that, as above, (4.2.1) holds since $r = 0$. Furthermore, the second inequality of (4.2.5) reduces to

$$6q + p^2 \geqslant 6pq + 7q^2 \iff 6q(1-p-q) + (p-q)(p+q) \geqslant 0,$$

which is indeed true since $p+q \leqslant 1$ and $p > q$. Therefore, the constraints of (4.2.1) and (4.2.5) are satisfied by $(p,q,0)$ in this case, and by Theorem 4.2.1, we conclude that the probability of draw in the corresponding generalized percolation game equals 0.

This completes the proof of the claim that the main result of [60] follows as a special case of our Theorem 4.2.1. □

Next, we come to the following two-parameter set-up: each directed edge of $\mathbb{Z}^2$ is assigned, independent of all else, a label that reads *trap* with probability r', *target* with probability s', and *open* with the remaining probability $(1-r'-s')$, with (r',s') belonging to the parameter-space $\Theta' = \{(r',s') \in [0,1]^2 : 0 < r'+s' \leqslant 1\}$. We refer to the game played on this random board as the *bond percolation game*. The moves permitted in this game are the same as those described above for the generalized percolation game, and a player wins if she is 1. either able to move the token along an edge labeled a target, 2. or force her opponent to move the token along an edge labeled a trap. The game continues for as long as the token keeps getting moved along edges that have been labeled open, which may happen indefinitely, leading to a draw. Once again, our objective is to find regimes, in terms of values of the parameter-pair (r',s'), for which the probability of draw in this game equals 0, and in Theorem 4.2.4, we list three such regimes.

Theorem 4.2.4. *Let $(r',s') \in \Theta'$, where Θ' is as described above.*

(B1) There exist $\varepsilon_1, \varepsilon_2 \in (0,1)$ such that the probability of draw in the bond percolation game, with underlying parameters r' and s', equals 0 whenever $r' \leqslant \varepsilon_1$, $s' \leqslant \varepsilon_2$ and

$$\frac{3(2s'-s'^2)^2}{2} + \frac{8(2s'-s'^2)r'}{1-s'} + \frac{11r'^2}{2(1-s')^2} \geqslant \frac{2r'}{1-s'} + \frac{9(2s'-s'^2)^3}{2} + \frac{16(2s'-s'^2)^2 r'}{1-s'} + \frac{43(2s'-s'^2)r'^2}{2(1-s')^2} + \frac{5r'^3}{(1-s')^3}. \tag{4.2.6}$$

(B2) As long as $s'=0$ and $r' > 0.157175$, the probability of draw in the bond percolation game, with underlying parameters r' and s', equals 0.

(B3) As long as $r'=s' \geqslant 0.10883$, the probability of draw in the bond percolation game, with underlying parameters r' and s', equals 0.

We draw the reader's attention to the fact that, between them, the regimes described in (B1), (B2) and (B3) cover a significant subset of the parameter-space Θ'. In particular, (B1) covers a significant portion of Θ' that intersects the neighbourhood around $(r',s') = (0,0)$ – a region where older techniques in the literature fail to establish that the probability of draw equals 0. Furthermore, as is the case with the three-parameter generalized percolation game, (B1) seems to indicate (and proves it under the restriction that (r',s') satisfies the inequality given by (4.2.6)) that a phase transition happens exactly at $(0,0)$ (in the sense that, at $(r',s') = (0,0)$, the probability of draw equals 1, whereas for r', s' sufficiently small, at least one of them strictly positive and (r',s') satisfying (4.2.6)), the probability of draw equals 0).

4.3 Formal description of the probabilistic cellular automata we are concerned with

Recalling the definition of PCAs from §1.1.1, we present the specific PCAs we are concerned with in this work.

4.3.1 The *specific* PCAs we work with in this chapter

We begin by describing the PCA that represents the game rules arising from the generalized percolation games described in §4.2. This PCA, denoted $\widehat{G}_{p,q,r}$ (in order to emphasize its dependence on the parameter-triple (p,q,r)), is a 1-dimensional one (i.e. its universe is $\mathbb{Z}$), with

1. the alphabet $\hat{\mathscr{A}} = \{W, L, D\}$ (already, this choice of notations hints at the intimate connection between the game and this PCA: if $(x,y) \in \mathbb{Z}^2$ serves as the initial vertex for a generalized percolation game, we say that $(x,y) \in W$ if this game is won by the player who plays the first round, $(x,y) \in L$ if this game is lost by the player who plays the first round, and $(x,y) \in D$ if this game results in a draw),
2. the neighbourhood-marking set $\mathscr{N} = \{0,1\}$,
3. and stochastic update rules that can be described via the stochastic matrix $\widehat{\varphi}_{p,q,r} : \hat{\mathscr{A}}^2 \times \hat{\mathscr{A}} \to [0,1]$, defined as follows (see Figure 4.4 for a pictorial illustration):

$$\widehat{\varphi}_{p,q,r}(W,W,b) = \begin{cases} p & \text{if } b = W, \\ (1-p) & \text{if } b = L, \end{cases} \tag{4.3.1}$$

$$\widehat{\varphi}_{p,q,r}(a_0,a_1,b) = \begin{cases} p+(1-p-q)(1-r) & \text{if } b = W, \\ q+(1-p-q)r & \text{if } b = L, \end{cases} \quad \text{where } (a_0,a_1) \in \{(W,L),(L,W)\}, \tag{4.3.2}$$

$$\varphi_{p,q,r}(L,L,b) = \begin{cases} p+(1-p-q)(1-r^2) & \text{if } b = W, \\ q+(1-p-q)r^2 & \text{if } b = L, \end{cases} \tag{4.3.3}$$

$$\widehat{\varphi}_{p,q,r}(a_0,a_1,b) = \begin{cases} p & \text{if } b = W, \\ (1-p-q)(1-r) & \text{if } b = D, \\ (1-p-q)r+q & \text{if } b = L, \end{cases} \quad \text{where } (a_0,a_1) \in \{(W,D),(D,W)\}, \tag{4.3.4}$$

$$\widehat{\varphi}_{p,q,r}(a_0,a_1,b) = \begin{cases} p+(1-p-q)(1-r) & \text{if } b=W, \\ (1-p-q)r(1-r) & \text{if } b=D, \\ q+(1-p-q)r^2 & \text{if } b=L, \end{cases} \quad \text{where } (a_0,a_1) \in \{(L,D),(D,L)\}, \tag{4.3.5}$$

$$\widehat{\varphi}_{p,q,r}(D,D,b) = \begin{cases} p & \text{if } b=W, \\ (1-p-q)(1-r^2) & \text{if } b=D, \\ q+(1-p-q)r^2 & \text{if } b=L. \end{cases} \tag{4.3.6}$$

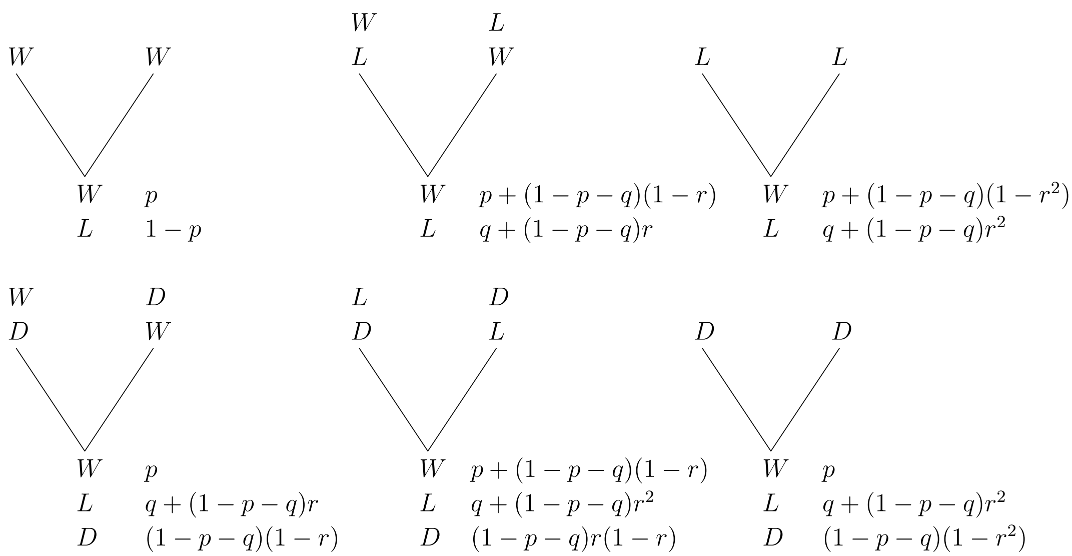


Figure 4.4: Illustrating the stochastic update rules for the PCA $\widehat{G}_{p,q,r}$

A detailed explanation as to why these stochastic update rules are equivalent to the game rules governing the generalized percolation game, has been provided in §4.6. When we restrict $\hat{\mathscr{A}}$ to the smaller alphabet, $\mathscr{A} = \{W,L\}$, we obtain yet another PCA, denoted $G_{p,q,r}$, with stochastic update rules described via the stochastic matrix $\varphi_{p,q,r} : \mathscr{A}^2 \times \mathscr{A} \to [0,1]$, where $\varphi_{p,q,r}(a_0,a_1,b) = \widehat{\varphi}_{p,q,r}(a_0,a_1,b)$ for all $a_0,a_1,b \in \mathscr{A}$ (in other words, $\varphi_{p,q,r}$ can be defined using (4.3.1), (4.3.2) and (4.3.3)). We refer to $\widehat{G}_{p,q,r}$ as the *envelope* to $G_{p,q,r}$ (see §4.7, where we include a discussion on the importance of invoking the notion of envelope PCAs).

We now describe the PCA, denoted $\widehat{E}_{r',s'}$, that represents the game rules arising from our bond percolation game, with the underlying parameters, r' and s', described in §4.2. Endowed with $\mathbb{Z}$ as its universe, $\widehat{E}_{r',s'}$ is, just like $\widehat{G}_{p,q,r}$, a 1-dimensional PCA with alphabet $\hat{\mathscr{A}}$ and neighbourhood-marking set $\mathscr{N}$, and its stochastic update rules are captured by the stochastic matrix $\widehat{\varphi}_{r',s'} : \hat{\mathscr{A}}^2 \times$

$\hat{\mathscr{A}} \to [0,1]$ defined as follows (see Figure 4.5 for a pictorial illustration):

$$\widehat{\varphi}_{r',s'}(W,W,b) = \begin{cases} 2s' - s'^2 & \text{if } b = W, \\ (1-s')^2 & \text{if } b = L, \end{cases} \tag{4.3.7}$$

$$\widehat{\varphi}_{r',s'}(a_0,a_1,b) = \begin{cases} 1 - r' + r's' & \text{if } b = W, \\ r' - r's' & \text{if } b = L, \end{cases} \quad \text{where } (a_0,a_1) \in \{(W,L),(L,W)\}, \tag{4.3.8}$$

$$\varphi_{r',s'}(L,L,b) = \begin{cases} 1 - r'^2 & \text{if } b = W, \\ r'^2 & \text{if } b = L, \end{cases} \tag{4.3.9}$$

$$\widehat{\varphi}_{r',s'}(a_0,a_1,b) = \begin{cases} (1-s')(1-r'-s') & \text{if } b = D, \\ r'(1-s') & \text{if } b = L, \\ 2s' - s'^2 & \text{if } b = W, \end{cases} \quad \text{where } (a_0,a_1) \in \{(W,D),(D,W)\}, \tag{4.3.10}$$

$$\widehat{\varphi}_{r',s'}(a_0,a_1,b) = \begin{cases} 1 - r' + r's' & \text{if } b = W, \\ r'(1-r'-s') & \text{if } b = D, \\ r'^2 & \text{if } b = L, \end{cases} \quad \text{where } (a_0,a_1) \in \{(L,D),(D,L)\}, \tag{4.3.11}$$

$$\widehat{\varphi}_{r',s'}(D,D,b) = \begin{cases} (1-r'-s')(1+r'-s') & \text{if } b = D, \\ r'^2 & \text{if } b = L, \\ 2s' - s'^2 & \text{if } b = W. \end{cases} \tag{4.3.12}$$

Once again, we refer the reader to §4.6 for a detailed explanation as to how the game rules associated with the bond percolation game give rise to the above-mentioned stochastic update rules. If we restrict $\hat{\mathscr{A}}$ to $\mathscr{A}$, we obtain the PCA $E_{r',s'}$ with the same neighbourhood-marking set $\mathscr{N}$, and with stochastic update rules captured by the stochastic matrix $\varphi_{r',s'} : \mathscr{A}^2 \times \mathscr{A} \to [0,1]$ such that $\varphi_{r',s'}(a_0,a_1,b) = \widehat{\varphi}_{r',s'}(a_0,a_1,b)$ for all $a_0,a_1,b \in \mathscr{A}$ (in other words, the equations defining $\varphi_{r',s'}$ are the same as (4.3.7), (4.3.8) and (4.3.9)). As is the case with $G_{p,q,r}$ and $\widehat{G}_{p,q,r}$, we refer to $\widehat{E}_{r',s'}$ as the envelope for the PCA $E_{r',s'}$.

We end §4.3 with the main results of this work pertaining to the PCAs described in §4.3.1:

Theorem 4.3.1. *The PCA $G_{p,q,r}$, as well as its envelope $\widehat{G}_{p,q,r}$, is ergodic whenever the underlying parameter-triple (p,q,r) belongs to the parameter-space Θ, where Θ is as defined in §4.2, and satisfies the constraints stated in Theorem 4.2.1 (in other words, when p, q and r are sufficiently*

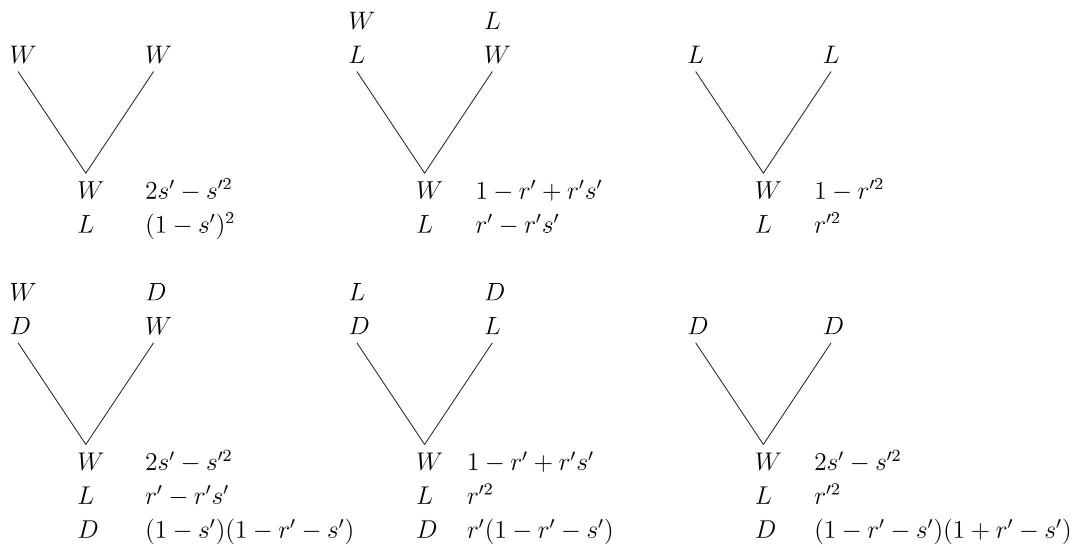


Figure 4.5: Illustrating the stochastic update rules for the PCA $\widehat{E}_{r',s'}$

small, the inequality (4.2.1) *holds, and* (p,q,r) *satisfies one of* (4.2.2), (4.2.3), (4.2.4) *and* (4.2.5)*).*

Likewise, the PCA $E_{r',s'}$*, as well as its envelope* $\widehat{E}_{r',s'}$*, is ergodic whenever the underlying parameter-pair* (r',s') *belongs to the parameter-space* Θ'*, where* Θ' *is as defined in §4.2, and* (r',s') *satisfies any one of the constraints stated in Theorem 4.2.4, i.e. either (B1) or (B2) or (B3).*

4.3.2 How the PCA $\widehat{E}_{r',s'}$ is obtained as a special case of the PCA $\widehat{G}_{p,q,r}$

Let us consider a bond percolation game with the underlying parameter-pair (r',s'), i.e. where r' indicates the probability of an edge being labeled a trap and s' indicates that of it being labeled a target. Letting η denote this random assignment of labels (i.e. η is the infinite tuple made up of $\eta\big((x,y),(x+1,y)\big)$ and $\eta\big((x,y),(x,y+1)\big)$ for all $(x,y) \in \mathbb{Z}^2$), we tweak η as follows:

1. if $\eta\big((x,y),(x+1,y)\big)$ (respectively, $\eta\big((x,y),(x,y+1)\big)$) equals a target, we re-label the edge $\big((x,y),(x+1,y)\big)$ (respectively, the edge $\big((x,y),(x,y+1)\big)$), independent of all else, to $\eta'\big((x,y),(x+1,y)\big)$ (respectively, to $\eta'\big((x,y),(x,y+1)\big)$), where

$$\eta'\big((x,y),(x+1,y)\big) = \begin{cases} \text{trap with probability } r'(1-s')^{-1} \\ \text{open with probability } (1-r'-s')(1-s')^{-1} \end{cases}$$

(the same distribution is adopted for $\eta'\big((x,y),(x,y+1)\big)$ if $\eta\big((x,y),(x,y+1)\big)$ equals a target), and we label the vertex (x,y) a trap, i.e. we set $\sigma'(x,y)$ to be equal to a trap;

2. if $\eta\big((x,y),(x+1,y)\big)$ (respectively, $\eta\big((x,y),(x,y+1)\big)$) is either open or a trap, we keep its label unchanged, i.e. we set $\eta'\big((x,y),(x+1,y)\big) = \eta\big((x,y),(x+1,y)\big)$ (respectively, $\eta'\big((x,y),(x,y+1)\big) = \eta\big((x,y),(x,y+1)\big)$);

3. for any $(x',y') \in \mathbb{Z}^2$ for which neither $\eta\big((x,y),(x+1,y)\big)$ nor $\eta\big((x,y),(x,y+1)\big)$ equals a target, we label (x,y) open, i.e. we set $\sigma'(x,y)$ to be open.

The modified labeling is now (σ',η'), where σ' is the tuple comprising $\sigma'(x,y)$ for all $(x,y) \in \mathbb{Z}^2$, and η' is the tuple comprising $\eta'\big((x,y),(x+1,y)\big)$ and $\eta'\big((x,y),(x,y+1)\big)$ for all $(x,y) \in \mathbb{Z}^2$. From the definition of (σ',η'), it is evident that the labels $\sigma'(x,y)$ are i.i.d. over all $(x,y) \in \mathbb{Z}^2$, with

$$\begin{aligned} p &= \mathbf{P}[\sigma'(x,y) = \text{trap}] \\ &= \mathbf{P}\left[\text{at least one of } \eta\big((x,y),(x+1,y)\big) \text{ and } \eta\big((x,y),(x,y+1)\big) \text{ equals a target}\right] \\ &= 1-(1-s')^2 = 2s' - s'^2, \end{aligned} \tag{4.3.13}$$

and likewise, the labels $\eta'(e)$ are i.i.d. over all directed edges e of our graph, with

$$\begin{aligned} r &= \mathbf{P}[\eta'(e) = \text{trap}] \\ &= \mathbf{P}[\eta(e) = \text{trap}] + \mathbf{P}[\eta'(e) = \text{trap} \,|\, \eta(e) = \text{target}]\, \mathbf{P}[\eta(e) = \text{target}] \\ &= r' + \frac{r'}{1-s'} \cdot s' = \frac{r'}{1-s'}. \end{aligned} \tag{4.3.14}$$

Moreover, it is easily verified that the tuples σ' and η' are independent of each other as well: for instance, we have

$$\begin{aligned} &\mathbf{P}\left[\sigma'(x,y) = \text{trap}, \eta'\big((x,y),(x+1,y)\big) = \text{trap}\right] \\ =\, &\mathbf{P}\left[\eta\big((x,y),(x+1,y)\big) = \text{target}, \eta'\big((x,y),(x+1,y)\big) = \text{trap}\right] \\ &+ \mathbf{P}\left[\eta\big((x,y),(x+1,y)\big) = \text{trap}, \eta\big((x,y),(x,y+1)\big) = \text{target}\right] \\ =\, &s' \cdot \frac{r'}{1-s'} + r' \cdot s' = \frac{r's'(2-s')}{1-s'} = \left(2s' - s'^2\right) \frac{r'}{1-s'} = pr, \end{aligned}$$

as desired. It is now straightforward to see that by choosing $q = 0$, p as in (4.3.13) and r as in (4.3.14), we can reduce (4.3.1) to (4.3.7), (4.3.2) to (4.3.8) and so on.

This transformation – from a generalized percolation game with the underlying parameter-triple $(p,0,r)$ to a bond percolation game with the underlying parameter-pair (r',s'), where p is as given by (4.3.13) and r is as given by (4.3.14) – as established above, lies at the heart of the proof that in the regime covered by (B1) of Theorem 4.2.4, the probability of draw equals 0.

4.3.3 A major contribution of this chapter in establishing ergodicity of elementary PCAs

Recall that a one-dimensional PCA is said to be *elementary* (see §1.1.1) if the cardinality of its alphabet as well as its neighbourhood-marking set equals 2. The definitions of $G_{p,q,r}$ and $E_{r',s'}$ in §4.3.1 immediately reveal that each of them is an elementary PCA. In Chapter 7 of [37], two fundamental results pertaining to ergodicity of elementary PCAs have been proposed. We first state these two results in terms of the notation used in this work. We let $\mathscr{A} = \{W, L\}$ (as has been defined in §4.3.1) denote the alphabet of *any* elementary PCA F, and we let

$$\theta_{i,j} = \mathbf{P}\left[F\eta(x) = L \middle| \eta(x+y_1) = i, \eta(x+y_2) = j\right], \text{ for } i, j \in \mathscr{A},$$

where $\{y_1, y_2\}$ forms the neighbourhood-marking set of F. Thus, the PCA F is completely specified by the parameters $\theta_{W,W}$, $\theta_{W,L}$, $\theta_{L,W}$ and $\theta_{L,L}$. We call F *symmetric* when $\theta_{W,L} = \theta_{L,W}$, in which case F is specified by only three parameters, namely, $\theta_{W,W}$, $\theta_{W,L}$ and $\theta_{L,L}$. The following are two well-known results when it comes to ergodicity properties of symmetric elementary PCAs, stated in Chapter 7 of [37]:

(a) the PCA F is ergodic when $\theta_{W,W}$, $\theta_{W,L}$ and $\theta_{L,L}$ satisfy the inequalities:

$$0 < \theta_{W,W}, \theta_{W,L}, \theta_{L,L} < 1, \tag{4.3.15}$$

$$\theta_{L,L} > \theta_{W,W} - 2\theta_{W,L}, \tag{4.3.16}$$

$$\theta_{L,L} > \theta_{W,W} - 2(1-\theta_{W,L}); \tag{4.3.17}$$

(b) the PCA F is ergodic when $\theta_{W,W}$, $\theta_{W,L}$ and $\theta_{L,L}$ satisfy the inequality:

$$\max\{|\theta_{i,j} - \theta_{k,\ell}| : i, j, k, \ell \in \{W, L\}\} + 2\max\{|\theta_{L,L} - \theta_{W,L}|, |\theta_{W,W} - \theta_{W,L}|\} < 2. \tag{4.3.18}$$

It has been mentioned in Chapter 7 of [37] that these two regimes together cover more than 90% of the unit cube $[0,1]^3$ defined by the parameter-triple $(\theta_{W,W}, \theta_{W,L}, \theta_{L,L})$, and that the only region where no method of proving or disproving ergodicity for symmetric elementary PCAs, in general, is known is the union of the neighbourhoods of the points $(\theta_{W,W}, \theta_{W,L}, \theta_{L,L}) = (1,0,0)$ and $(\theta_{W,W}, \theta_{W,L}, \theta_{L,L}) = (1,1,0)$.

We now come to the two symmetric elementary PCAs that are of interest to us in this section, namely $G_{p,q,r}$ and $E_{r',s'}$. We see, from (4.3.1), (4.3.2) and (4.3.3), that $\theta_{W,W} = (1-p)$, $\theta_{W,L} = q+(1-p-q)r$ and $\theta_{L,L} = q+(1-p-q)r^2$ for $G_{p,q,r}$, so that we have $(\theta_{W,W}, \theta_{W,L}, \theta_{L,L}) = (1,0,0)$

if and only if $p = q = r = 0$. Thus, the neighbourhood of $(1,0,0)$ is obtained when we consider as small a value of each of p, q and r as we desire. We then show, in Theorem 4.2.1, that if, in addition to $(\theta_{W,W}, \theta_{W,L}, \theta_{L,L})$ being in the neighbourhood of $(1,0,0)$ (equivalently, (p,q,r) being arbitrarily close to $(0,0,0)$), at least one of p, q and r is strictly positive and the constraints specified in Theorem 4.2.1 are satisfied, the PCA $G_{p,q,r}$ is ergodic – thereby rigorously proving the ergodicity of $G_{p,q,r}$ in a *considerable* chunk of the neighbourhood of $(1,0,0)$. Likewise, from (4.3.7), (4.3.8) and (4.3.9), we see that $\theta_{W,W} = (1-s')^2$, $\theta_{W,L} = r'(1-s')$ and $\theta_{L,L} = r'^2$ for the PCA $E_{r',s'}$, so that we have $(\theta_{W,W}, \theta_{W,L}, \theta_{L,L}) = (1,0,0)$ if and only if $r' = s' = 0$. Thus, the neighbourhood of $(1,0,0)$ is obtained when we consider as small a value of each of r' and s' as we desire. Part (B1) of Theorem 4.2.4 then shows that, in addition to $(\theta_{W,W}, \theta_{W,L}, \theta_{L,L})$ being in the neighbourhood of $(1,0,0)$ (equivalently, (r',s') being arbitrarily close to $(0,0)$), if at least one of r' and s' is strictly positive and (4.2.6) holds, the PCA $E_{r',s'}$ is ergodic – thereby rigorously proving the ergodicity of $E_{r',s'}$ in a considerable chunk of the neighbourhood of $(1,0,0)$.

In fact, even the regimes described by (B2) and (B3) of Theorem 4.2.4 go significantly beyond those that are obtained via ((a)) and ((b)) for the PCA $E_{r',s'}$. For instance, when $s' = 0$, the criterion in (4.3.15) fails to hold for $E_{r',s'}$ since $\theta_{W,W} = 1$, so that (B2) falls outside of the regime described in ((a)). On the other hand, $\max\{|\theta_{i,j} - \theta_{k,\ell}| : i,j,k,\ell \in \mathscr{A}\} = 1 - r'^2$ and $\max\{|\theta_{L,L} - \theta_{W,L}|, |\theta_{W,W} - \theta_{W,L}|\} = 1 - r'$ when $s' = 0$, so that for (4.3.18) to hold, we must have $1 - r'^2 + 2(1-r') = 3 - 2r' - r'^2 < 2 \iff 2r' + r'^2 > 1 \iff r' > 0.414214$. This tells us that the regime described in (B2), i.e. $s' = 0$ and $r' > 0.157175$, goes *far beyond* the regime obtained by implementing ((b)).

When $r' = s'$ (obviously, the common value for r' and s' in this case must be bounded above by 0.5, since $r' + s' \leqslant 1$), we have $\theta_{W,W} = (1-r')^2$, $\theta_{W,L} = r' - r'^2$ and $\theta_{L,L} = r'^2$. It is immediate that these values of $\theta_{W,W}$, $\theta_{L,W}$ and $\theta_{L,L}$ satisfy (4.3.15) when $r' \in (0,1)$. When examining if the inequality in (4.3.16) is satisfied or not, we observe that

$$\theta_{L,L} > \theta_{W,W} - 2\theta_{W,L} \iff r'^2 > (1-r')^2 - 2(r' - r'^2) \iff 2r'^2 - 4r' + 1 < 0 \iff r' > 1 - \frac{1}{\sqrt{2}} \approx 0.293.$$

When checking whether the inequality in (4.3.17) is satisfied or not, we obtain

$$\theta_{L,L} > \theta_{W,W} - 2(1 - \theta_{W,L}) \iff r'^2 > (1-r')^2 - 2(1 - r' + r'^2) \iff 2r'^2 + 1 > 0,$$

which is indeed true for all (r',s') belonging to the regime in (B3). We thus conclude that when $r' = s'$, we require $r' > 0.293$ for the inequalities in ((a)) to hold simultaneously, whereas our result in this work is able to establish ergodicity for $E_{r',s'}$ whenever $r' = s' > 0.10883$. We now focus on the regime covered by ((b)) when $r' = s'$. We observe that

$$\max\{|\theta_{W,W} - \theta_{W,L}|, |\theta_{W,W} - \theta_{L,L}, |\theta_{W,L} - \theta_{L,L}|\} = \max\{|r'(2r'-1)|, |2r'-1|, |(1-r')(2r'-1)|\} = 1-2r'$$

and

$$\max\{|\theta_{W,W} - \theta_{W,L}|, |\theta_{W,L} - \theta_{L,L}|\} = \max\{|r'(2r'-1)|, |(1-r')(2r'-1)|\} = (1-r')(1-2r'),$$

so that 4.3.18 reduces to

$$1-2r' + 2(1-r')(1-2r') < 2 \Longleftrightarrow 4r'^2 - 8r' + 1 < 0 \Longleftrightarrow r' > \frac{2-\sqrt{3}}{2} \approx 0.13397.$$

Once again, our result, in (B3), establishes ergodicity for $E_{r',s'}$ whenever $r' = s' > 0.10883$, a significant improvement on the lower bound that is provided by ((b)).

4.4 Motivations for studying generalized / bond percolation games, and a brief review of pertinent literature

4.4.1 In oligopolistic competitions

The bond percolation game studied in this work can be interpreted as a 'toy' set-up for modeling simplistic *oligopolistic competitions*. Imagine two companies, C_1 and C_2, locked in an oligopolistic competition: if a company sustains a short term capital loss that is beyond a certain value, it is forced to quit and leave the market, whereas if a company is able to make a capital profit beyond a certain margin, its opponent gives up and quits, leading to a *monopoly*. We consider a highly simplified model for such an oligopolistic market, in which we do not take into account possible collusion and cooperation between the two companies. This model comprises a directed graph in which each directed edge has been assigned a label, as follows: 1. either it is a *trap* edge, traveling along which is equivalent to sustaining a disastrous capital loss, 2. or it is a *target* edge, traversing which is equivalent to making a tremendous profit and dealing an irreparable blow to one's opponent, 3. or it is a *neutral* or *safe* edge, traversing which forces neither of the two companies to exit the market. This directed, labeled network is assumed to be visible, in its entirety, to both C_1 and C_2. The companies make their moves in alternate rounds, where a *move* involves traveling along a directed edge (u, v) (i.e. *from* the vertex u *to* the vertex v), where u is the vertex one's opponent reached at the end of the previous round. A company wins by either traversing a target edge, or by forcing its opponent to travel along a trap edge. A draw in this game would imply that neither company is able to perform decisively better than the other and eliminate the competition. An interesting future direction to explore, in the context of oligopolistic markets, would be where the directed, labeled network is *not* visible *a priori* to the two companies engaged in this competition, but instead, is generated *dynamically* (i.e. once a company has traveled along the directed edge (u,v), the label of the directed edge (v,w) is revealed for every vertex w *to* which an edge *from* v exists).

A similar applicability of our bond percolation game can be found in simplistically modeling TV debates in which two rival politicians, or the spokespersons of two rival political parties, have taken part. The lines of argument on a single topic of debate that can be pursued, are represented by the (directed) edges of a (possibly infinite) graph. Given vertices u, v and w such that there is a directed edge, (u,v), *from* u *to* v, and another, (v,w), *from* v *to* w, we assume that (v,w) represents a line of argument that counters the line of argument represented by (u,v). From prior social experiments and / or prior experience, both participants are already aware of

1. each line of argument, represented by a *target* edge, that is popular with the audience, so that traversing it leads to an increase in the *political capital* or *political goodwill*,

2. each line of argument, represented by a *trap* edge, that is *not* favoured by the audience (but a participant may be forced to pursue such a line of argument if their opponent manages to "corner" them suitably) and is likely to prove detrimental to the public image of the participant pursuing it, thereby depleting their political capital,

3. each line of argument, represented by a *neutral* or *safe* edge, that has little power in swaying the audience *either* in favour of *or* against the participant pursuing it.

If the anchor chooses to end the debate (and declare a winner) the moment one of the participants traverses an edge that is *not* safe, we get back the bond percolation game that we study in thiswork.

4.4.2 As an adversarial avatar of percolation

Another usefulness of our generalized / bond percolation game lies in incorporating an *adversarial* element into the notion of ordinary site and / or bond percolation on directed infinite graphs. Percolation, by itself, is a vast and ever-expanding area of research that spans a multitude of disciplines, including statistical physics, chemistry, computer science, biology and sociology (for instance, in the understanding of phenomena such as *polymeric gelation*, the spread of a forest fire or the propagation of an epidemic, the transportation of a fluid through a porous material etc.). Percolation games (more specifically, *site percolation games*) were first introduced in [60], where each *vertex* of the infinite 2-dimensional square lattice $\mathbb{Z}^2$ was assigned, independently, a label that read *trap* with probability p, *target* with probability q, and *open* with probability $(1-p-q)$. The two players were allowed the same moves as those permitted in our game in §4.2, and a player won if she could move the token to a vertex labeled a target or force her opponent to move the token to a vertex labeled a trap. It was shown that the probability of draw in this game is 0 whenever $(p+q)>0$. The set-up described for our generalized percolation game in §4.2 is a natural and important generalization of this model, as it additionally considers a random assignment of labels to the *edges* of $\mathbb{Z}^2$. The result presented in [60] was, in some sense, extended in [16], where the token was permitted to be moved from where it was currently located, say the vertex (x,y), to any one of $(x,y+1)$, $(x+1,y+1)$ and $(x+2,y+1)$. As in [60], it was shown in [16] that the probability of draw is 0 whenever at least one of p and q is strictly positive. A different direction of generalization of the above-mentioned site percolation game on $\mathbb{Z}^2$ was pursued in [15], by allowing the token to be moved from (x,y) to one of $(x,y+1)$ and $(x+1,y+1)$ if x is even, and from (x,y) to one of $(x+1,y+1)$ and $(x+2,y+1)$ if x is odd.

It is worthwhile to note here that when $s'=0$ and $r'>0$ in the set-up described for our bond percolation game in §4.2, the game is essentially a *normal* game that is being played on the infinite oriented 2-dimensional square lattice, $\mathbb{Z}^2$, once the process of bond percolation, with edge-deletion probability r', has been performed on it. This can be justified as follows: each edge of $\mathbb{Z}^2$ is, independently, retained with probability $(1-r')$ and deleted with probability r'. A move in the normal game played on this premise involves relocating the token from where it is currently located, say the site (x,y), to either $(x+1,y)$ (if the directed edge $\big((x,y),(x+1,y)\big)$ has not been deleted) or to $(x,y+1)$ (if the directed edge $\big((x,y),(x,y+1)\big)$ has not been deleted). A player loses if she is unable to make a move (i.e. there are no outgoing edges from the site where the vertex is currently located). The game continues indefinitely if neither player ever encounters a site from which both outgoing edges have been deleted, and this is possible only if there exists an infinite path starting from the initial vertex, i.e. if percolation happens. Thus, in such a scenario, the probability of draw is bounded above by the probability of occurrence of percolation, i.e. the probability of the existence of an infinite path that begins from the initial vertex. In this context, we refer the reader to [61], where the normal game, as well as the *misère* and the *escape* games have been studied on rooted Galton-Watson trees.

4.4.3 A brief discussion on Maker-Breaker percolation games

A different flavour of *two-player combinatorial games* pertaining to percolation, broadly referred to as the *Maker-Breaker percolation games*, was introduced in [34]. An $m \times n$ rectangular grid was considered, and two players, titled *Maker* and *Breaker*, took turns to make moves. When it was Maker's turn to move, she claimed m (as yet unclaimed) edges of the grid, and when it was Breaker's turn, she deleted b (as yet unclaimed) edges of the grid. Maker won if she managed to claim all edges constituting a *crossing path* joining the left boundary of the grid to its right boundary. Given $m, n \in \mathbb{N}$, [34] investigated the values of (m, b) for which this *two-player positional game*, titled the (m, b)*-crossing game*, was won by Maker. In [35], an infinite connected graph Λ was considered, and a vertex $v_0 \in \Lambda$ was specified. The Maker-Breaker percolation game was played on this premise (with the permitted moves being the same as those in [34]), and Breaker was said to win if at any point of time during the game, the connected component of Λ containing v_0 became finite. Two specific instances of Λ were considered: the 2-dimensional lattice $\mathbb{Z}^2$ and the d-regular tree (also known as the *Bethe lattice*, in which each vertex has degree d), and conditions involving m and b (and d, in the latter instance) were established under which Maker (respectively, Breaker) had a winning strategy. In [41], a critical threshold (lower bound) for b in terms of m was found such that if b were to exceed this value, Breaker would win the Maker-Breaker percolation game on $\mathbb{Z}^2$. Furthermore, [41] showed that when the game was played on $\mathbb{Z}^2$ after the usual bond percolation process with parameter p had been performed, and p was not too large compared to $\frac{1}{2}$, Breaker almost surely won the Maker-Breaker percolation game when $m = b = 1$ (referred to as the *unbiased* version). Continuing the work accomplished in [41], in [42], it is shown that when $p < 1$, Breaker almost surely wins the $(1, 1)$ Maker-Breaker percolation game on the same (random) premise as considered in [41], and it is further established that in the $(2, 1)$ Maker-Breaker percolation game on this premise, Maker almost surely wins whenever $p > 0.9402$, while Breaker almost surely wins whenever $p < 0.5278$. In [103], the (m, b)-crossing game was studied on a triangular grid that was m vertices across and n vertices high. In [42], the (m, b) Maker-Breaker percolation game, with $m = b = 1$ is played on the 2-dimensional square lattice $\mathbb{Z}^2$ after the usual bond percolation process with parameter p has been performed on it.

4.5 Motivation for studying the PCAs we consider in this chapter

PCAs find applications in various domains like fault-tolerant computation, classifying CAs based on robustness, connections to Gibbs potentials in statistical mechanics, and modeling complex systems in physics, chemistry, and biology. For a comprehensive survey on PCAs and their development, we refer the reader to [73] and [37]. The applications of PCAs are extensive, spanning probability, statistical mechanics, computer science, natural sciences, dynamical systems, and computational cell biology, as showcased in [70], [68], [44], [86], and many others.

But why are the *specific* PCAs we study, i.e. $G_{p,q,r}$, $\widehat{G}_{p,q,r}$, $E_{r',s'}$ and $\widehat{E}_{r',s'}$, of interest (independent of the significance they have in the context of our bond percolation games)? Focusing on $E_{r',s'}$, we see that it can be interpreted as a *learning model* for *social learning* in a system of *interacting particles* (these particles are often thought of as *players* or *agents*). The notion of social learning was introduced in [43], which studied how the speed of learning and market equilibrium were impacted by social networks and other institutions governing communication among market participants. In [8], a general framework in which agents, unaware of the payoffs from different actions, use their own past experience and the experience of their neighbours to guide their decision making. In [9], an approach to network formation is studied in which it is assumed that a link formed by one agent with another allows access, in part and in due course, to the benefits available to the latter agent via their own links. In [30], a model of social learning in a population of myopic, memoryless individuals is studied in which the agents are placed on $\mathbb{Z}$ and at each time-step, each agent performs an experiment using the technology they currently possess, then takes into account the outcome of their own experiment as well as the outcomes of the experiments performed by their neighbours.

Coming back to our PCA $E_{r',s'}$, we refer to Figure 4.6. Let each vertex or site on the integer line $\mathbb{Z}$ be inhabited by an agent, and at the beginning of each epoch (here, the epochs are indexed by the set $\mathbb{N}_0$ of non-negative integers), each agent can avail one of two technologies: W and L. During each epoch, each agent performs, using the technology it has chosen to adopt at the beginning of that epoch, an experiment that has two possible outcomes: success and failure. An agent using technology W has probability s' of achieving success, while an agent equipped with technology L has probability $(1-r')$ of achieving success. It is assumed that these outcomes occur independently for all agents on $\mathbb{Z}$, over all epochs in $\mathbb{N}_0$. At the beginning of each epoch, an agent looks at itself and its nearest neighbour to the right, and 1. updates its technology to L if and only if *both* of them suffered failures at the previous time-step, 2. or else, it updates its technology to

W. Figure 4.6 reveals that this is exactly how updates happen when we apply the stochastic update rules corresponding to $E_{r',s'}$ (recall these rules from §4.3.1). An understanding of the ergodicity, and subsequently, limiting distribution(s) of $E_{r',s'}$ will, therefore, reveal how the diffusion of the technologies, W and L, happens throughout this system of interacting agents as time approaches ∞.

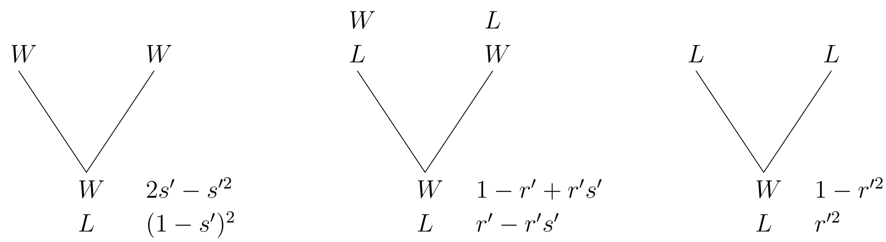


Figure 4.6: Illustrating the stochastic update rules of the PCA $E_{r',s'}$

Referring to Figure 1 of [73] and identifying W with the symbol 0 and L with the symbol 1, we see that $E_{r',s'}$ can be viewed as a generalization of the *noisy additive PCA*, since $E_{r',s'}$ allows us to introduce randomness into the update $\eta_{t+1}(x)$ that happens conditioned on $(\eta_t(x), \eta_t(x+1)) = (0,0)$ or on $(\eta_t(x), \eta_t(x+1)) = (1,1)$. We may even provide an interpretation to $E_{r',s'}$ that is similar to the interpretation of how we obtain noisy additive PCAs: given η_t, for each $x \in \mathbb{Z}$, we first set the value at x to be $\frac{1}{2}\eta_t(x) + \frac{1}{2}\eta_t(x+1)$, and then obtain $\eta_{t+1}(x)$ by

1. flipping a 0 to a 1 with probability $(1-s')^2$,
2. flipping a 1 to a 0 with probability $1 - r'^2$,
3. flipping a $\frac{1}{2}$ to a 1 with probability $r'(1-s')$ and to a 0 with probability $1 - r'(1-s')$.

Referring to Figure 7 of [73], we see that $E_{r',s'}$ can also be seen as a generalization of the *directed animals PCA*, allowing us to incorporate randomness into $\eta_{t+1}(x)$ when we condition on $(\eta_t(x), \eta_t(x+1))$ being an element from the set $\{(0,1),(1,0),(1,1)\}$. In [20], the *random gas model* on an *agreeable graph* has been described, in which a site (i.e. a vertex of the graph) cannot be occupied if any of its "children" or "neighbours" is occupied by a gas particle. Our PCA $E_{r',s'}$ admits a generalization of this model when we let W indicate the state of being unoccupied and L indicate the state of being occupied: assuming that both r' and s' are very small, and considering the underlying graph to be $\mathbb{Z}^2$ in which the "children" or "neighbours" of (x,y) are $(x+1,y)$ and $(x,y+1)$,

1. if both $(x+1,y)$ and $(x,y+1)$ are unoccupied, then (x,y) is occupied with probability $(1-s')^2$, and unoccupied otherwise,

2. if precisely one of $(x+1,y)$ and $(x,y+1)$ is occupied, then (x,y) is occupied with probability $r'(1-s')$ (which is much lower than the probability of occupation in the first case), and unoccupied otherwise,

3. if both of $(x+1,y)$ and $(x,y+1)$ are occupied, then (x,y) is occupied with probability r'^2 (which, in turn, is much lower than the probability of occupation in the second case), and unoccupied otherwise.

In some sense, we can see this as a less idealized, more practical version of the random gas model, in which gas particles do repel each other but not so strongly that the occupation of a "child" or "neighbour" to a given vertex completely prevents the occupation of that vertex (i.e. our PCA $E_{r',s'}$ allows for a random gas model *without hard constraints*), and where the higher the number of occupied "children" or "neighbours" of a vertex, the lower the chance of that vertex being occupied.

Very similar motivations justify the importance of studying the PCA $G_{p,q,r}$ as well.

We mention here that [23] shows the existence of a $\delta_0 > 0$ such that an elementary PCA (whose alphabet comprises the symbols W and L) with the NAND function and either a vertex-binary-symmetric-channel (BSC) noise with parameter δ or an edge-BSC noise with parameter δ is ergodic for all $\delta \in (0,\delta_0)$ (to elucidate, the BSC_δ operator applied to a symbol from $\{W,L\}$ keeps it intact with probability $(1-2\delta)$ and switches it to the other symbol with probability 2δ).

Finally, given that, of the two main objects we investigate in this section, one is a class of PCAs and the other a game that is closely tied to the process of percolation, it is only relevant that we draw attention of the reader to a few of the works in the literature dedicated to exploring the connection between PCAs and percolation. In [56], it is shown that *additive* PCAs can be interpreted as *oriented site percolation* models. In [57], a technique for proving the sharpness of phase transitions in percolation is implemented in the setting of *attractive* PCAs. Moreover, a correspondence between PCAs and *bootstrap percolation* is established in [57], thereby deducing exponential decay of the probability of remaining healthy above criticality for a class of bootstrap percolation models, with implications for related *kinetically constrained models*. It is further demonstrated that this correspondence provides an equivalence between the non-triviality of the phase transition in certain bootstrap percolation models and the stability, with respect to noise, of certain CAs.

4.6 Game rules: an analysis of each game via recurrence relations

Recall, from §4.1, our brief allusion to game rules as the bridge between our percolation games on one hand, and our PCA on the other. Deducing these game rules, or, in other words, recurrence relations that govern the games we study, is key to our analysis, and §4.6 is dedicated to this task.

We partition $\mathbb{Z}^2$ into the following three (random) subsets of vertices (the randomness arises from the random assignment of labels as described in §4.2):

1. we let W consist of all $(x,y) \in \mathbb{Z}^2$ such that the game with (x,y) as its initial vertex is won by the player who plays its first round;
2. we let L consist of all $(x,y) \in \mathbb{Z}^2$ such that the game with (x,y) as its initial vertex is lost by the player who plays its first round;
3. and we let D comprise all $(x,y) \in \mathbb{Z}^2$ such that the game with (x,y) as the initial vertex results in a draw.

In particular, in case of generalized percolation games, a vertex labeled a trap is placed in the set W, while a vertex labeled a target is placed in the set L. The intuition behind such a convention can be explained as follows. One may imagine an *unseen* round that takes place before the *actual* game begins, during which the player who is supposed to play the second round of the actual game moves the token to the initial vertex (x,y). If (x,y) is a trap, she loses immediately, allowing her opponent, i.e. the player who is supposed to play the first round of the actual game, to win even before the game begins, justifying the reason why we include each trap vertex in the set W. Likewise, the reason for including target vertices in the set L is justified.

For each $k \in \mathbb{Z}$, we set $S_k = \{(x,y) \in \mathbb{Z}^2 : x+y=k\}$ (i.e. S_k is the diagonal line, running from top-left to bottom-right, that constitutes all those vertices of $\mathbb{Z}^2$ whose coordinates sum to k). Right away, we state the following fact that becomes evident from the description of both of our games in §4.2:

1. once each vertex on S_{k+1} has been categorized into one of the subsets W, L, and D,
2. each edge between S_k and S_{k+1} has received its label (which is 'trap' or 'open' when we consider the generalized percolation game, and 'trap' or 'target' or 'open' when we consider the bond percolation game),

3. and each vertex on S_k has received its label (this is relevant only for the generalized percolation game, and the label is any one of 'trap', 'target' and 'open'),

we are in a position to uniquely determine which of W, L and D each vertex on S_k belongs to. How this determination happens is revealed via the recurrence relations, or game rules, described in §4.6.1 and §4.6.2.

4.6.1 Game rules for the generalized percolation game

Fix any $k \in \mathbb{Z}$, any $(x,y) \in S_k$, and set $\mathrm{Out}(x,y) = \{(x+1,y),(x,y+1)\}$. We assume that (x,y) serves as the initial vertex for a generalized percolation game, and we let P_1 denote the player who plays the first round, and P_2 the player who plays the second round of this game. Recalling the rules of the generalized percolation game described in §4.2, we are able to make the following observations:

1. When both vertices of $\mathrm{Out}(x,y)$ are in W, P_2 wins no matter how P_1 moves in the first round, unless (x,y) has been labeled a trap. Therefore, in this case, $(x,y) \in L$ with probability $(1-p)$ and $(x,y) \in W$ with probability p.

2. When $(x+1,y) \in L$ and $(x,y+1) \in W$ (an analogous situation arises when $(x+1,y) \in W$ and $(x,y+1) \in L$), P_1 wins (by moving the token from (x,y) to $(x+1,y)$ in the first round) if either (x,y) has been labeled a trap or (x,y) has been labeled open and the directed edge $\big((x,y),(x+1,y)\big)$ is also open, while P_2 is the winner otherwise. Therefore, in this case, $(x,y) \in W$ with probability $\{p+(1-p-q)(1-r)\}$, and $(x,y) \in L$ with probability $\{q+(1-p-q)r\}$.

3. When both vertices of $\mathrm{Out}(x,y)$ are in L, P_1 wins if (x,y) has been labeled a trap or (x,y) is open and at least one of the edges $\big((x,y),(x+1,y)\big)$ and $\big((x,y),(x,y+1)\big)$ is also open, while P_2 wins otherwise. Therefore, in this case, $(x,y) \in W$ with probability $\{p+(1-p-q)(1-r^2)\}$, and $(x,y) \in L$ with probability $\{q+(1-p-q)r^2\}$.

4. When $(x+1,y) \in W$ and $(x,y+1) \in D$ (an analogous situation arises when $(x+1,y) \in D$ and $(x,y+1) \in W$), P_1 wins if (x,y) has been labeled a trap, the game results in a draw if (x,y) is open and the edge $\big((x,y),(x,y+1)\big)$ has not been labeled a trap, and P_2 wins in all other cases. Therefore, we have $(x,y) \in W$ with probability p, $(x,y) \in D$ with probability $(1-p-q)(1-r)$, and $(x,y) \in L$ with probability $\{q+r(1-p-q)\}$.

5. When $(x+1,y) \in L$ and $(x,y+1) \in D$ (an analogous situation arises when $(x+1,y) \in D$ and $(x,y+1) \in L$), P_1 wins if (x,y) has been labeled a trap, or if both the vertex (x,y) and the edge $\big((x,y),(x+1,y)\big)$ are open. The game results in a draw if (x,y) is open, the edge $\big((x,y),(x+1,y)\big)$ has been labeled a trap and the edge $\big((x,y),(x,y+1)\big)$ is open, while in all other cases, P_2 wins. Therefore, we have $(x,y) \in W$ with probability $\{p+(1-p-q)(1-r)\}$, $(x,y) \in D$ with probability $(1-p-q)r(1-r)$, and $(x,y) \in L$ with probability $\{q+r^2(1-p-q)\}$.

6. Finally, when both vertices of $\mathrm{Out}(x,y)$ are in D, P_1 wins if (x,y) is a trap, the game results in a draw if (x,y) is open and at least one of the edges $\big((x,y),(x+1,y)\big)$ and $\big((x,y),(x,y+1)\big)$ is open, and in all other cases, P_2 wins. Therefore, $(x,y) \in W$ with probability p, $(x,y) \in D$ with probability $(1-p-q)(1-r^2)$ and $(x,y) \in L$ with probability $\{q+r^2(1-p-q)\}$.

For each $k \in \mathbb{Z}$, we identify S_k with a copy of the integer line $\mathbb{Z}$, by identifying the vertex $(x,k-x)$ on S_k with x on $\mathbb{Z}$. For an arbitrary but fixed $k \in \mathbb{Z}$, let $\omega(x) \in \hat{\mathscr{A}} = \{W,L,D\}$, referred to as the *state* of the vertex $(x,k+1-x)$, denote the subset (out of W, L and D) to which $(x,k+1-x)$ belongs, for each $x \in \mathbb{Z}$. Conditioned on the infinite tuple $\omega = (\omega(x) : x \in \mathbb{Z})$ that specifies the state of *each* vertex on S_{k+1}, the *random* state of the vertex $(x,k-x)$, lying on S_k, equals $\widehat{G}_{p,q,r}\omega(x)$ for each $x \in \mathbb{Z}$, as is evident from the recurrence relations described above, and from (4.3.1), (4.3.2), (4.3.3), (4.3.4), (4.3.5) and (4.3.6)). This is how the game rules governing the generalized percolation game give rise to the PCA $\widehat{G}_{p,q,r}$.

4.6.2 Game rules for the bond percolation game

The recurrence relations arising from the bond percolation game described in §4.2 can be deduced in much the same manner as those derived in §4.6.1 for the generalized percolation game. We let an arbitrarily chosen $(x,y) \in S_k$, for some $k \in \mathbb{Z}$, be the initial vertex, and assume, as before, that P_1 plays the first round of the game while P_2 plays the second. We then deduce the state of the vertex (x,y), i.e. which of the subsets W, L and D it belongs to, as follows:

1. When both vertices of $\mathrm{Out}(x,y)$ are in W, P_1 loses unless at least one of the edges $\big((x,y),(x+1,y)\big)$ and $\big((x,y),(x,y+1)\big)$ has been labeled a target, so that $(x,y) \in L$ with probability $(1-s')^2$ and $(x,y) \in W$ with probability $(2s'-s'^2)$.

2. When $(x+1,y) \in L$ and $(x,y+1) \in W$ (analogously, when $(x+1,y) \in W$ and $(x,y+1) \in L$), P_1 wins unless the edge $\big((x,y),(x+1,y)\big)$ has been labeled a trap and the edge $\big((x,y),(x,y+1)\big)$ is not a target. Hence, $(x,y) \in L$ with probability $r'(1-s')$, and $(x,y) \in W$ with probability $\{1-r'(1-s')\}$.

3. When both vertices of $\mathrm{Out}(x,y)$ are in L, P_1 wins unless both the edges $\big((x,y),(x+1,y)\big)$ and $\big((x,y),(x,y+1)\big)$ are traps. Therefore, $(x,y) \in L$ with probability r'^2 and $(x,y) \in W$ with probability $(1-r'^2)$.

4. When $(x+1,y) \in W$ and $(x,y+1) \in D$ (analogously, when $(x+1,y) \in D$ and $(x,y+1) \in W$), P_1 wins if at least one of the edges $\big((x,y),(x+1,y)\big)$ and $\big((x,y),(x,y+1)\big)$ has been labeled a target, the game results in a draw if $\big((x,y),(x+1,y)\big)$ is not a target and $\big((x,y),(x,y+1)\big)$ is open, and in every other case, P_2 wins. Thus, $(x,y) \in W$ with probability $\{1-(1-s')^2\}$, $(x,y) \in D$ with probability $(1-s')(1-r'-s')$, and $(x,y) \in L$ with probability $r'(1-s')$.

5. When $(x+1,y) \in L$ and $(x,y+1) \in D$ (analogously, when $(x+1,y) \in D$ and $(x,y+1) \in L$), P_1 wins as long as the edge $\big((x,y),(x+1,y)\big)$ is not a trap or at least one of the edges $\big((x,y),(x+1,y)\big)$ and $\big((x,y),(x,y+1)\big)$ is a target, the game results in a draw if $\big((x,y),(x+1,y)\big)$ is a trap and $\big((x,y),(x,y+1)\big)$ is open, and in all other cases, P_2 wins. Thus, $(x,y) \in W$ with probability $\{1-r'(1-s')\}$, $(x,y) \in D$ with probability $r'(1-r'-s')$, and $(x,y) \in L$ with probability r'^2.

6. When both vertices of $\mathrm{Out}(x,y)$ are in D, P_1 wins if at least one of $\big((x,y),(x+1,y)\big)$ and $\big((x,y),(x,y+1)\big)$ is a target, the game results in a draw if neither $\big((x,y),(x+1,y)\big)$ nor $\big((x,y),(x,y+1)\big)$ is a target but not both are traps, and P_2 wins otherwise. Thus, $(x,y) \in W$ with probability $\{1-(1-s')^2\}$, $(x,y) \in D$ with probability $(1-r'-s')(1+r'-s')$, and $(x,y) \in L$ with probability r'^2.

As argued in §4.6.1, the above-mentioned recurrence relations or game rules governing the bond percolation game are captured precisely by the stochastic update rules of the PCA $\widehat{E}_{r',s'}$ in (4.3.7), (4.3.8), (4.3.9), (4.3.10), (4.3.11) and (4.3.12).

4.7 Proofs of Theorems 4.2.1, 4.2.4 and 4.3.1, assuming Theorems 4.7.1 and 4.7.2 to be true

Let us consider a generalized percolation game on $\mathbb{Z}^2$. For each fixed realization of the random assignment of labels (as described in §4.2) to the vertices and edges of our graph $\mathbb{Z}^2$, each vertex $(x,y) \in \mathbb{Z}^2$ can be classified uniquely into one of the subsets W, D and L (dictated by the outcome of the game whose initial vertex is (x,y)), and we refer to this subset as the *state* of (x,y). Let $\boldsymbol{\eta}(x,y)$ denote the *random* state of (x,y) induced by the random assignment of labels to the vertices and edges of $\mathbb{Z}^2$. Since *all* vertices of $\mathbb{Z}^2$ bear labels that are i.i.d., and *all* edges of $\mathbb{Z}^2$ bear labels that are i.i.d. as well, the joint law μ_k of the random infinite tuple $(\boldsymbol{\eta}(x,y) : (x,y) \in S_k) = (\boldsymbol{\eta}(x,k-x) : x \in \mathbb{Z})$ is the same for *every* $k \in \mathbb{Z}$, so that we may denote this common law henceforth by μ (with no dependence on k). Referring to the game rules explained in §4.6.1 and how they can be represented by $\widehat{G}_{p,q,r}$ defined in §4.3, we can write $\widehat{G}_{p,q,r}\mu_{k+1} = \mu_k$, for each $k \in \mathbb{Z}$. Combining the two observations made above, we conclude that $\widehat{G}_{p,q,r}\mu = \mu$, thus proving that μ is a stationary distribution for $\widehat{G}_{p,q,r}$. Moreover, the i.i.d. nature of the assignment of labels to the vertices and edges of $\mathbb{Z}^2$ ensures that μ is both *translation-invariant* and *reflection-invariant*. By these two attributes, we mean the following:

1. Recall, from the last paragraph of §4.6.1, that we identify S_k with $\mathbb{Z}$ by mapping $(x,k-x)$ onto x for each $x \in \mathbb{Z}$. Consequently, μ can be thought of as a probability measure on $\Omega = \hat{\mathscr{A}}^{\mathbb{Z}} = \{W,L,D\}^{\mathbb{Z}}$. For any $y \in \mathbb{Z}$, we let $\mathfrak{T}^y : \Omega \to \Omega$ map any configuration $\eta = (\eta(x) : x \in \mathbb{Z})$ to the configuration $\mathfrak{T}^y\eta = (T^y\eta(x) : x \in \mathbb{Z})$ where $\mathfrak{T}^y\eta(x) = \eta(x+y)$ for each $x \in \mathbb{Z}$. A probability measure ν, defined with respect to the σ-field $\mathscr{F}$ generated by all the cylinder sets of Ω, is said to be *translation-invariant* if $\nu(B) = \nu(\mathfrak{T}^y B)$ for every $B \in \mathbb{D}$, where $\mathfrak{T}^y B = \{\mathfrak{T}^y\eta : \eta \in B\}$.

2. Let $\mathfrak{R} : \Omega \to \Omega$ map any configuration $\eta = (\eta(x) : x \in \mathbb{Z})$ to the configuration $\mathfrak{R}\eta = (\mathfrak{R}\eta(x) : x \in \mathbb{Z})$ where $\mathfrak{R}\eta(x) = \eta(-x)$ for each $x \in \mathbb{Z}$. A probability measure ν, defined with respect to $\mathscr{F}$, is said to be *reflection-invariant* if $\nu(B) = \nu(\mathfrak{R}B)$ for every $B \in \mathbb{D}$, where $\mathfrak{R}B = \{\mathfrak{R}\eta : \eta \in B\}$.

We now state what can be considered the most technically challenging result of ourwork (the measure μ in the statement of Theorem 4.7.1 need not be the same as the one alluded to in the previous paragraph):

Theorem 4.7.1. *Let $(p,q,r) \in \Theta$ satisfy the constraints stated in Theorem 4.2.1, i.e. each of p, q and r is sufficiently small, the inequality in (4.2.1) holds, and precisely one of (4.2.2), (4.2.3), (4.2.4) and (4.2.5) is true. If μ is any translation-invariant and reflection-invariant stationary distribution for $\widehat{G}_{p,q,r}$, then for any fixed but arbitrary $x \in \mathbb{Z}$, the probability, under μ, that x is assigned the state D, abbreviated as $\mu(D)$, equals* 0.

An analogous result is true in case of bond percolation games, and can be stated as follows:

Theorem 4.7.2. *Let $(r',s') \in \Theta'$ belong to any one of the three regimes described in Theorem 4.2.4, i.e. one of (B1), (B2) and (B3). If μ is any translation-invariant and reflection-invariant stationary distribution for $\widehat{E}_{r',s'}$, then for any fixed but arbitrary $x \in \mathbb{Z}$, the probability, under μ, that x is assigned the state D, abbreviated simply as $\mu(D)$, equals* 0.

The proofs of Theorems 4.7.1 and 4.7.2 are accomplished in §4.8 and §4.9 respectively, employing the technique of *weight functions* or *potential functions* that was first introduced in [60] and later explored in [23], [16] and [15].

We have yet to establish a connection between the ergodicity of either of the PCAs $G_{p,q,r}$ and $\widehat{G}_{p,q,r}$ (respectively, $E_{r',s'}$ and $\widehat{E}_{r',s'}$) and the event of draw in our generalized percolation game (respectively, bond percolation game). This is where $G_{p,q,r}$ (respectively, $E_{r',s'}$) plays an important role.

Theorem 4.7.3. *For each $(p,q,r) \in \Theta$, the generalized percolation game, with underlying parameters p, q and r, has probability 0 of resulting in a draw if and only if the PCA $G_{p,q,r}$ is ergodic. Likewise, for each $(r',s') \in \Theta'$, the bond percolation game, with underlying parameters r' and s', has probability 0 of resulting in a draw if and only if the PCA $E_{r',s'}$ is ergodic.*

Proof. The proof of Theorem 4.7.3 happens via the exact same argument as the proof of Proposition 2.2 of [60], and is therefore omitted from this work. □

Our next result is Lemma 4.7.4, and we note that it can be proved in exactly the same way as Proposition 2.1 of [60] (this lemma, in fact, gets used in proving one side of the two-way implication asserted upon in Theorem 4.7.3):

Lemma 4.7.4. *The PCA $\widehat{G}_{p,q,r}$ is ergodic if and only if the PCA $G_{p,q,r}$ is ergodic. Likewise, $\widehat{E}_{r',s'}$ is ergodic if and only if $E_{r',s'}$ is ergodic.*

Some discussion is now in order as to the philosophy behind invoking the notion of envelope PCAs. We let F denote a d-dimensional PCA endowed with the alphabet $\mathscr{A}_F = \{W, L\}$, some neighbourhood-marking set $\mathscr{N}_F$, and stochastic update rules given by some stochastic matrix φ_F. Introduced in [28], the envelope PCA $\widehat{\mathrm{F}}$ corresponding to F is a different PCA whose alphabet $\mathscr{A}_{\widehat{\mathrm{F}}} = \{W, L, D\}$ is obtained by appending $\mathscr{A}_F$ with the symbol D. The neighbourhood-marking set for $\widehat{\mathrm{F}}$ is, once again, $\mathscr{N}_F$, and the stochastic update rules for $\widehat{\mathrm{F}}$ are captured by the stochastic matrix $\varphi_{\widehat{\mathrm{F}}}$ such that, when we restrict $\mathscr{A}_{\widehat{\mathrm{F}}}$ to $\mathscr{A}_F$, the corresponding restriction of $\varphi_{\widehat{\mathrm{F}}}$ yields φ_F. The state spaces for F and $\widehat{\mathrm{F}}$ are respectively $\Omega_F = \mathscr{A}_F^{\mathbb{Z}^d}$ and $\Omega_{\widehat{\mathrm{F}}} = \mathscr{A}_{\widehat{\mathrm{F}}}^{\mathbb{Z}^d}$. A configuration $\eta = \left(\eta(\mathbf{x}) : \mathbf{x} \in \mathbb{Z}^d\right)$ in $\Omega_{\widehat{\mathrm{F}}}$ can be thought of as a configuration that actually belongs to Ω_F, but with an unknown symbol (which could equal either W or L) occupying each coordinate $\mathbf{x}$ of $\mathbb{Z}^d$ for which $\eta(\mathbf{x}) = D$. Thus, the symbol D acts as a placeholder whenever $\mathbf{x} \in \mathbb{Z}^d$ is occupied by a symbol from the original alphabet $\mathscr{A}_F$ that we are not aware of or are uncertain about. Given an ergodic PCA F with the unique stationary distribution μ, the motivation for introducing envelope PCAs lies in coming up with an efficient procedure, known as a *perfect sampling* procedure (see [87] and

[28]), for generating a configuration from μ. One such algorithm implements the *coupled from the past* (*CFTP*) method proposed in [87], but it is inefficient when the state space Ω_F is large (for instance, in this work, $\Omega_{G_{p,q,r}} = \Omega_{E_{r',s'}} = \{W,L\}^{\mathbb{Z}}$ is uncountably infinite). Even when the PCA F is ergodic, the corresponding envelope PCA $\widehat{\mathrm{F}}$ may or may not be ergodic, but if it is, then [28] shows that an efficient perfect sampling procedure can be achieved by running $\widehat{\mathrm{F}}$ on a single initial configuration. The envelope PCA $\widehat{\mathrm{F}}$ is used to couple two (random) configurations obtained via (possibly repeated) applications of F, and as mentioned above, the symbol D is used to populate the cells of $\mathbb{Z}^d$ at which these two coupled configurations may differ from one another.

We are now in a position to prove Theorem 4.2.1, assuming the conclusion of Theorem 4.7.1, as follows:

Proof of Theorem 4.2.1. Let $(p,q,r) \in \Theta$ satisfy the constraints stated in Theorem 4.2.1. Once Theorem 4.7.1 has been established, it allows us to conclude that $\mu[\{\boldsymbol{\eta}(x,k-x) = D\}] = 0$ for *any* $x,k \in \mathbb{Z}$, where μ is the joint law of $(\boldsymbol{\eta}(x,k-x) : x \in \mathbb{Z})$ (see the first paragraph of §4.7). Such a conclusion is possible because we showed, just before stating Theorem 4.7.1, that μ is a translation-invariant and reflection-invariant stationary distribution for $\widehat{G}_{p,q,r}$. In other words, for *any* vertex $(x,k-x)$ lying on *any* diagonal S_k, the probability that it belongs to D equals 0 under μ. Therefore, almost surely, there is *no* vertex in $\mathbb{Z}^2$ which, when serving as the initial vertex for the generalized percolation game, leads to a draw. This, along with Theorem 4.7.3 and Lemma 4.7.4, yields the conclusion of Theorem 4.2.1. □

The argument for proving Theorem 4.2.4, provided Theorem 4.7.2 holds, is identical to the above.

Proof of Theorem 4.3.1. As explained above, once Theorem 4.7.1 has been established, the conclusion of Theorem 4.2.1 follows. By Theorem 4.7.3, we then know that the PCA $G_{p,q,r}$ is ergodic whenever the underlying parameter-triple (p,q,r) belongs to Θ and satisfies the constraints stated in Theorem 4.2.1. By Lemma 4.7.4, we further conclude that the envelope PCA $\widehat{G}_{p,q,r}$ is ergodic for all such values of (p,q,r) as well. This concludes the proof of the first assertion stated in Theorem 4.3.1. The proof of the second assertion of Theorem 4.3.1 follows via an identical argument (in which we replace Theorem 4.2.1 by Theorem 4.2.4). □

All we have left to prove now are Theorems 4.7.1 and 4.7.2.

4.8 The proof of Theorem 4.7.1 by the technique of weight functions

It was stated in §4.7 that the crux of this work lies in the rather technical proof of Theorem 4.7.1 (and its counterpart in case of bond percolation games, Theorem 4.7.2). Before proceeding with the proof, we summarize here, for the reader's convenience, how the rest of §4.8 has been organized:

1. In §4.8.1, we outline the central ideas relevant to the step-by-step construction of our weight function – these ideas apply equally well to the construction of the weight functions that allow us to prove Theorem 4.7.2.

2. In §4.8.2, we state, without proof, the weight function we come up with for the generalized percolation game, the *weight function inequality* it satisfies, and finally, draw the desired conclusion from this inequality (thus proving Theorem 4.7.1) .

3. In §4.8.3, we show the details of each step leading to the construction of our weight function.

4.8.1 The principal ideas behind the construction of our weight functions

We present to the reader a very broad, but hopefully comprehensive, outline of how we aim to accomplish the proof of Theorem 4.7.1. We denote by $\mathscr{R}$ the subset of the parameter-space Θ (defined in §4.2) that consists of all those triples (p,q,r) that satisfy the constraints stated in Theorem 4.2.1 – in other words, p, q and r are sufficiently small, the inequality in (4.2.1) holds, and precisely one of (4.2.2), (4.2.3), (4.2.4) and (4.2.5) is true. Although, in the rest of §4.8.1, we focus on $(p,q,r) \in \mathscr{R}$ and work with the envelope PCA $\widehat{G}_{p,q,r}$ corresponding to generalized percolation games, this overview is equally pertinent to the construction of the weight functions corresponding to the three different regimes addressed in Theorem 4.7.2 for bond percolation games.

Recall, from §4.3.1, that the state space corresponding to our envelope PCA, $\widehat{G}_{p,q,r}$, is $\Omega = \hat{\mathscr{A}}^{\mathbb{Z}}$, where $\hat{\mathscr{A}} = \{W,L,D\}$ is its alphabet. Let $\mathscr{M}$ denote the space of all reflection-invariant and translation-invariant (recall these definitions from §4.7) probability measures on Ω, defined with respect to the σ-field $\mathscr{F}$ generated by all cylinder sets of Ω. The goal is to come up with a "suitable" real-valued function w, referred to as a *weight function* or *potential function*, defined on $\mathscr{M}$. As such, a weight function (that serves the purpose outlined below) need not be unique, but as far as this work is concerned, we construct w as a linear combination of the form

$$w(\mu) = \sum_{i=1}^{t} c_i \mu(\mathscr{C}_i), \tag{4.8.1}$$

where $t \in \mathbb{N}$, each of $\mathscr{C}_1, \ldots, \mathscr{C}_t$ is a cylinder set (thus belonging to $\mathscr{F}$), and each of $c_1, \ldots, c_t$ is a real-valued function (in fact, a polynomial) in the parameters p, q and r. We would like this function to satisfy an inequality, henceforth referred to as a *weight function inequality*, of the form

$$w\left(\widehat{G}_{p,q,r}\mu\right) \leqslant w(\mu) - \sum_{i=1}^{t'} c'_i \mu\left(\mathscr{C}'_i\right), \tag{4.8.2}$$

where $t' \in \mathbb{N}$, and $\mathscr{C}'_i$ is a cylinder set and c'_i is a real-valued function (once again, a polynomial) in p, q and r, for each $i \in \{1,2,\ldots,t'\}$. It must be ensured that

$$c'_i \geqslant 0 \text{ for each } i \in \{1,2,\ldots,t'\} \text{ for each } (p,q,r) \in \mathscr{R}. \tag{4.8.3}$$

Let $\mathscr{I}$ denote the set of all i, with $i \in \{1,2,\ldots,t'\}$, such that $c'_i > 0$ whenever $(p,q,r) \in \mathscr{R}$. If μ is stationary for $\widehat{G}_{p,q,r}$ (so that $\widehat{G}_{p,q,r}\mu = \mu$) and $\mathscr{I}$ is non-empty, by (4.8.3) we conclude that $\mu\left(\mathscr{C}'_i\right) = 0$ for each $i \in \mathscr{I}$ whenever $(p,q,r) \in \mathscr{R}$. If the cylinder sets $\mathscr{C}'_1, \ldots, \mathscr{C}'_{t'}$ have been chosen judiciously enough, suitable set-theoretic operations performed on $\mathscr{C}'_i$, for $i \in \mathscr{I}$, allow us to conclude that $\mu((D)_0) = 0$ whenever $(p,q,r) \in \mathscr{R}$ (here, $(D)_0$ indicates the cylinder set that consists of *all* configurations $\eta = (\eta(x) : x \in \mathbb{Z})$ with $\eta(0) = D$, i.e. in which the 0-th coordinate is occupied by the symbol D). This concludes the primary idea propelling the proof of Theorem 4.7.1.

From here onward, we make use of some simplified notations, as follows. In its most general form, a cylinder set is indicated by $(a_1,a_2,\ldots,a_\ell)_{x_1,x_2,\ldots,x_\ell}$, where $a_1,a_2,\ldots,a_\ell \in \hat{\mathscr{A}}$ and $x_1,x_2,\ldots,x_\ell \in \mathbb{Z}$, signifying that this cylinder set comprises all tuples $\eta = (\eta(x) : x \in \mathbb{Z})$ in which $\eta(x_i) = a_i$ for each $i \in \{1,2,\ldots,\ell\}$. When $x_1,x_2,\ldots,x_\ell$ are consecutive integers, i.e. $x_{i+1} = x_i + 1$ for each $i \in \{1,2,\ldots,\ell-1\}$, we have

$$\mu\left((a_1,a_2,\ldots,a_\ell)_{0,1,\ldots,\ell-1}\right) = \mu\left((a_1,a_2,\ldots,a_\ell)_{x_1,x_1+1,\ldots,x_1+\ell-1}\right)$$

as μ is translation-invariant. This allows us to abbreviate $\mu\left((a_1,a_2,\ldots,a_\ell)_{x_1,x_1+1,\ldots,x_1+\ell-1}\right)$ as simply $\mu((a_1,a_2,\ldots,a_\ell))$ for *any* $x_1 \in \mathbb{Z}$. We further remove the commas and the outer parentheses and shorten $\mu((a_1,a_2,\ldots,a_\ell))$ to $\mu(a_1a_2\ldots a_\ell)$. We call a cylinder set, of the form $(a_1,a_2,\ldots,a_\ell)_{x_1,x_2,\ldots,x_\ell}$, *D-inclusive* if $a_i = D$ for at least one $i \in \{1,2,\ldots,\ell\}$.

We emphasize here that even proceeding via the approach outlined above, one may be able to come up with *different* weight functions all of which serve the same purpose (i.e. satisfy an inequality of the form (4.8.2), along with the criterion stated in (4.8.3)). Since the ultimate goal is to draw a conclusion about $\mu(D)$, it makes sense 1. to consider each cylinder set $\mathscr{C}_i$ in (4.8.1) to be D-inclusive, and 2. to begin our step-by-step construction of the weight function by setting $\mathscr{C}_1 = (D)_0$ in (4.8.1). The subsequent choices for $\mathscr{C}_2$, $\mathscr{C}_3$ etc. will be motivated in the subsubsections under §4.8.3. As seen in §4.8.3, in every step, up to and including the penultimate step, of constructing our weight function, we obtain a weight function inequality that, although of the same form as (4.8.2), fails to satisfy (4.8.3). This is what leads us to carry out an "adjustment" of the weight function obtained until that step. In §4.8.1.1, we present to the reader the heuristics of how such an adjustment is carried out, and how it helps. However, the implementation of this idea becomes much clearer when the reader goes through the specific details laid out in the subsubsections of §4.8.3.

4.8.1.1 How each adjustment to the weight function is carried out

Let us denote by w_{i-1} the weight function we have constructed up until the start of the i-th adjustment (thus, the very first weight function we begin with is denoted by w_0, the weight function obtained after the first adjustment is denoted by w_1 and so on). The weight function inequality at the *beginning* of the i-th adjustment will then be of the form

$$w_{i-1}\left(\widehat{G}_{p,q,r}\mu\right) \leqslant w_{i-1}(\mu) - \sum_{t=1}^{k_{i-1}} \alpha_{i-1,t}\mu\left(\mathscr{C}_{i-1,t}\right), \tag{4.8.4}$$

for some $k_{i-1} \in \mathbb{N}$, coefficients $\alpha_{i-1,1}, \ldots, \alpha_{i-1,k_{i-1}}$ that are polynomials in p, q and r, and cylinder sets $\mathscr{C}_{i-1,1}, \ldots, \mathscr{C}_{i-1,k_{i-1}}$. However, if there exists at least one coefficient among $\alpha_{i-1,1}, \ldots, \alpha_{i-1,k_{i-1}}$ that is negative when $(p,q,r) \in \mathscr{R}$, the inequality (4.8.4) is unable to fulfill (4.8.3) when $(p,q,r) \in \mathscr{R}$, thereby necessitating an adjustment to the weight function w_{i-1} derived so far. We now define, for suitably chosen cylinder sets $\mathscr{D}_{i,1}$, ..., $\mathscr{D}_{i,\ell_i}$, and suitably chosen coefficients $\beta_{i,1}$, ..., β_{i,ℓ_i} that are polynomials in p, q and r, the updated / adjusted weight function

$$w_i(\mu) = w_{i-1}(\mu) - \sum_{t=1}^{\ell_i} \beta_{i,t}\mu\left(\mathscr{D}_{i,t}\right). \tag{4.8.5}$$

When incorporated into (4.8.4), w_i yields the updated weight function inequality (obtained at the *end* of the i-th adjustment):

$$\begin{aligned} & w_i\left(\widehat{G}_{p,q,r}\mu\right) + \sum_{t=1}^{\ell_i} \beta_{i,t}\widehat{G}_{p,q,r}\mu\left(\mathscr{D}_{i,t}\right) \leqslant w_i(\mu) + \sum_{t=1}^{\ell_i} \beta_{i,t}\mu\left(\mathscr{D}_{i,t}\right) - \sum_{t=1}^{k_{i-1}} \alpha_{i-1,t}\mu\left(\mathscr{C}_{i-1,t}\right) \\ \iff & w_i\left(\widehat{G}_{p,q,r}\mu\right) \leqslant w_i(\mu) + \sum_{t=1}^{\ell_i} \beta_{i,t}\mu\left(\mathscr{D}_{i,t}\right) - \sum_{t=1}^{k_{i-1}} \alpha_{i-1,t}\mu\left(\mathscr{C}_{i-1,t}\right) - \sum_{t=1}^{\ell_i} \beta_{i,t}\widehat{G}_{p,q,r}\mu\left(\mathscr{D}_{i,t}\right). \end{aligned} \tag{4.8.6}$$

Attempt is usually made to select $\mathscr{D}_{i,1}, \ldots, \mathscr{D}_{i,\ell_i}$ as subsets of elements from the set

$$\left\{\mathscr{C}_{i-1,t} : 1 \leqslant t \leqslant k_{i-1} \text{ and } \alpha_{i-1,t} > 0\right\}, \tag{4.8.7}$$

with the added restriction that if $\mathscr{D}_{i,t} \subseteq \mathscr{C}_{i-1,t'}$ for some $\mathscr{C}_{i-1,t'}$ belonging to the set in (4.8.7), then $\beta_{i,t} \leqslant \alpha_{i-1,t'}$, so that the coefficient of $\mu\left(\mathscr{C}_{i-1,t'}\right)$ in the right side of (4.8.6) remains non-positive. Attempt is also made to select $\mathscr{D}_{i,1}, \ldots, \mathscr{D}_{i,\ell_i}$ in such a manner that terms from the sum $-\sum_{t=1}^{\ell_i} \beta_{i,t} \widehat{G}_{p,q,r} \mu\left(\mathscr{D}_{i,t}\right)$ are able to aid in negating terms of the form $-\alpha_{i-1,t}\mu\left(\mathscr{C}_{i-1,t}\right)$ in which $\alpha_{i-1,t} < 0$ when $(p,q,r) \in \mathscr{R}$.

The above gives a broad overview of 1. why adjustments to the weight function, in several rounds, are required, and 2. how they are usually accomplished. The above idea is reiterated until the final weight function inequality is of the form given in (4.8.2) and satisfies the criterion in (4.8.3).

4.8.2 The final weight function obtained for the generalized percolation game when $(p,q,r) \in \mathscr{R}$

We state here, for the reader's convenience, the final weight functions obtained, and the corresponding weight function inequalities, of the form (4.8.2) (and obeying the criterion stated in (4.8.3)), that these weight functions satisfy when $(p,q,r) \in \mathscr{R}$. The final weight function that aids us in proving Theorem 4.7.1, and, in turn, Theorem 4.2.1, is

$$w(\mu) = \mu(D) + \mu(WD) + \mu(LWD) \tag{4.8.8}$$

when (4.2.2) is true, and

$$\begin{aligned} w(\mu) = {} & \mu(D) + \mu(WD) + \{4r - 2(p+q) + (p+q)^2 - 6r^2 - 3r(p+2q)\}\mu(LD) + (1-p-q)(1-r) \\ & (-2r + r^2 + 3pr + 3qr + r^3)\mu(LDW) + (1-p-q)(1-r)(-1+p+q+r^2+2qr)\mu(LDL) \\ & - [2 - (3 + p + q + r - 2q^2 - 2r^2 - 5qr + 2pq + pr + 3pqr + 4q^2r + 4qr^2 + 5pr^2 - 2p^2r \\ & - 5pqr^2 - 3p^2r^2 - 2q^2r^2)(1-p-q)(1-r)]\mu(LWD) \end{aligned} \tag{4.8.9}$$

when any one of (4.2.3), (4.2.4) and (4.2.5) is true. The weight function inequality that each of the two functions in (4.8.8) and (4.8.9) satisfies is given by

$$w(\widehat{G}_{p,q,r}\mu) \leqslant w(\mu) - [2 - (1-p-q)(1-r)\{2 + 2p + q + r - q^2 - r^2 - 3qr + pq + 2q^2r + 2qr^2 + pqr +$$

$$
\begin{aligned}
&2pr^2-p^2r-2pqr^2-p^2r^2-q^2r^2\}]\{\mu(WD)-\mu(WWWD)-\mu(DWDD)-\mu(LWD)\}\\
&-\{1-(1+p+q+pq+pr-q^2-qr-p^2r+q^2r)(1-p-q)(1-r)(1+r)\}\mu(DD)\\
&-r^2(1-r)^2(1-p-q)^2\mu(WLD)-r^2(1-r)^2(1-p-q)^2\mu(DLD). \qquad (4.8.10)
\end{aligned}
$$

Let us now deduce, from (4.8.10), our desired conclusion, i.e. that $\mu(D)=0$ whenever μ is a translation-invariant and reflection-invariant stationary distribution for $\widehat{G}_{p,q,r}$. To this end, we shall focus on the term involving $\mu(DD)$ in (4.8.10). Expanding its coefficient, we obtain:

$$
\begin{aligned}
&-1+(1+p+q+pq+pr-q^2-qr-p^2r+q^2r)(1-p-q)(1-r)(1+r)\\
=&-p^3r^3-p^2q\{1-r-r^2+r^3\}-p^2r\{2-r-2r^2-p\}-pq^2r(1-r^2)-pr^3-2q^2r^3-q^3r\{1+r-r^2\}\\
&-qr\{1-pr-2q-r^2-2qr\}-q^2(1-q)-p(p+q-r)-q^2-r^2. \qquad (4.8.11)
\end{aligned}
$$

Since (4.2.1) is assumed to hold no matter which of (4.2.2), (4.2.3), (4.2.4) and (4.2.5) (p,q,r) satisfies, we have

$$
2(p+q)-(p+q)^2+3r(p+2q)\geqslant 4r-6r^2 \iff p+q-r\geqslant r\left(1-3r-\frac{3p}{2}-3q\right)+\frac{1}{2}(p+q)^2,
$$

and this lower bound on the right side of the inequality above is strictly positive as long as p, q and r are sufficiently small but at least one of them is strictly positive. Note, furthermore, that if $p=0$, the term $-p(p+q-r)$ disappears from (4.8.11), but $-q^2-r^2$ remains, and this is strictly negative whenever at least one of q and r is non-zero. All these observations show that when p, q and r are sufficiently small and (4.2.1) holds, the expression in (4.8.11) is strictly negative as long as at least one of p, q and r is strictly positive (recall, from §4.2, that the condition $p+q+r>0$ is included in the definition of the parameter-space Θ). Consequently, when μ is stationary for $\widehat{G}_{p,q,r}$, we deduce, from (4.8.10) and the identity $\widehat{G}_{p,q,r}\mu=\mu$, that $\mu(DD)=0$.

Since, for any stationary μ, we have $\widehat{G}_{p,q,r}\mu=\mu$, hence $\mu(DD)=0 \implies \widehat{G}_{p,q,r}\mu(DD)=0$. Here, $\widehat{G}_{p,q,r}\mu(DD)$ is the probability of the event $\{\widehat{G}_{p,q,r}\eta(0)=\widehat{G}_{p,q,r}\eta(1)=D\}$, where η is a random configuration with law μ, and we consider all contributions to this event that arise from the event $\{\eta(1)=D\}$. Note that

$$
(1-p-q)^2(1-r)^2\{\mu(WDW)+2r\mu(WDL)+r^2\mu(LDL)\}\leqslant \widehat{G}_{p,q,r}\mu(DD),
$$

so that when $r>0$, we have

$$
\widehat{G}_{p,q,r}\mu(DD)=0 \implies \mu(WDW)=\mu(WDL)=\mu(LDL)=0.
$$

Moreover, we have $\mu(WDD) = \mu(LDD) = \mu(DDD) = 0$ since each of them is bounded above by $\mu(DD)$, and $\mu(DD) = 0$. Since $\mu(D) = \mu(WDW) + \mu(LDL) + \mu(DDD) + 2\mu(WDL) + 2\mu(WDD) + 2\mu(LDD)$ (using reflection invariance), we have $\mu(D) = 0$, as desired.

Note that, when $r = 0$, we are back to the set-up considered in [60], and it has already been established there that the associated PCA, $\widehat{G}_{p,q,0}$, is ergodic whenever $p + q > 0$.

4.8.3 Detailed construction of the weight function for generalized percolation games when $(p,q,r) \in \mathscr{R}$

Recall that $\mathscr{R} \subset \Theta$ comprises all those values of (p,q,r) that satisfy the constraints stated in Theorem 4.2.1, i.e. p, q and r are sufficiently small, the inequality in (4.2.1) holds, and precisely one of (4.2.2), (4.2.3), (4.2.4) and (4.2.5) is true. Throughout §4.8.3, it is assumed that the probability measure μ that we deal with is defined on the state space $\Omega = \hat{\mathscr{A}}^{\mathbb{Z}} = \{W, L, D\}^{\mathbb{Z}}$ of $\widehat{G}_{p,q,r}$, with respect to the σ-field $\mathscr{F}$ generated by all cylinder sets of Ω, and that μ is both reflection-invariant and translation-invariant. Recall, from the discussion in the paragraph preceding §4.8.1.1, that beginning the construction of our weight function with the cylinder set $\mathscr{C}_1 = (D)_0$ seems to be a sensible and natural choice. Since, at each step, we must obtain an inequality of the form given by (4.8.2), it is crucial that we write down the pushforward measure $\widehat{G}_{p,q,r}\mu(\mathscr{C}_i)$ for each cylinder set $\mathscr{C}_i$ that we incorporate into our gradually-unfolding weight function. To this end, from (4.3.4), (4.3.5) and (4.3.6), and using the notation introduced in §4.8.1, we deduce that the pushforward measure for our choice of $\mathscr{C}_1$ is given by:

$$\begin{aligned}\widehat{G}_{p,q,r}\mu(D) &= (1-p-q)(1-r)\{\mu(WD)+\mu(DW)\} + (1-p-q)r(1-r)\{\mu(LD)+\mu(DL)\} \\ &\quad + (1-p-q)(1-r^2)\mu(DD) \\ &= 2(1-p-q)(1-r)\mu(WD) + 2(1-p-q)r(1-r)\mu(LD) + (1-p-q)(1-r^2)\mu(DD),\end{aligned} \tag{4.8.12}$$

where, in the last step, we make use of the assumption that μ is reflection-invariant.

We now explain the intuition behind our choice of the second cylinder set, $\mathscr{C}_2$, as we attempt to shape up our weight function in the form given by (4.8.1). Note that the coefficient $2(1-p-q)r(1-r)$ of $\mu(LD)$ in (4.8.12) is always bounded above by 1 (in fact, by $1/2$), since $r(1-r) \leqslant 1/4$ and $(1-p-q) \leqslant 1$. The coefficient $(1-p-q)(1-r^2)$ of $\mu(DD)$ in (4.8.12) is obviously bounded above by 1. On the other hand, when p, q and r are all sufficiently small, the coefficient $2(1-p-q)(1-r)$ of $\mu(WD)$ in (4.8.12) is very close to 2. We can thus write, using (4.8.12) and the fact that $\mu(D) = \mu(WD) + \mu(LD) + \mu(DD)$, the following inequality: $\widehat{G}_{p,q,r}\mu(D) \leqslant \mu(D) + \mu(WD)$.

Comparing this with (4.8.2), it becomes evident that if $c_1 = 1$ and $\mathscr{C}_1 = (D)_0$ in (4.8.1), a sensible choice for our next cylinder set and the corresponding coefficient would be $\mathscr{C}_2 = (W,D)_{0,1}$ and $c_2 = 1$.

This now requires us to find the pushforward measure $\widehat{G}_{p,q,r}\mu(WD)$, which is the same as the probability of the event $\{\widehat{G}_{p,q,r}\eta(0) = W, \widehat{G}_{p,q,r}\eta(1) = D\}$, where $\eta = (\eta(x) : x \in \mathbb{Z})$ is a (random) configuration following the probability distribution μ. In order for the event $\{\widehat{G}_{p,q,r}\eta(1) = D\}$ to take place, we must have at least one of $\eta(1)$ and $\eta(2)$ equal to D. The probability of the event $\{\widehat{G}_{p,q,r}\eta(1) = D\}$ equals 1. $(1-p-q)(1-r)$ when conditioned on the event $\{\eta(1) = D, \eta(2) = W\}$, 2. $(1-p-q)r(1-r)$ when conditioned on the event $\{\eta(1) = D, \eta(2) = L\}$, 3. and $(1-p-q)(1-r^2)$ when conditioned on the event $\{\eta(1) = \eta(2) = D\}$. On the other hand, the probability of the event $\{\widehat{G}_{p,q,r}\eta(0) = W\}$ equals 1. p when conditioned on either of the events $\{\eta(0) = W, \eta(1) = D\}$ and $\{\eta(0) = \eta(1) = D\}$, 2. and $p + (1-p-q)(1-r)$ when conditioned on the event $\{\eta(0) = L, \eta(1) = D\}$. Conditioned on the event $\{\eta(1) = W, \eta(2) = D\}$, the probability of $\{\widehat{G}_{p,q,r}\eta(1) = D\}$ equals $(1-p-q)(1-r)$, while the probability of $\{\widehat{G}_{p,q,r}\eta(0) = W\}$ equals 1. p when conditioned on either of the events $\{\eta(0) = \eta(1) = W\}$ and $\{\eta(0) = D, \eta(1) = W\}$, 2. and $p + (1-p-q)(1-r)$ when conditioned on $\{\eta(0) = L, \eta(1) = W\}$. Finally, conditioned on the event $\{\eta(1) = L, \eta(2) = D\}$, the probability of $\{\widehat{G}_{p,q,r}\eta(1) = D\}$ equals $(1-p-q)r(1-r)$, while the probability of $\{\widehat{G}_{p,q,r}\eta(0) = W\}$ equals 1. $p + (1-p-q)(1-r)$ when conditioned on either of the events $\{\eta(0) = W, \eta(1) = L\}$ and $\{\eta(0) = D, \eta(1) = L\}$, 2. and $p + (1-p-q)(1-r^2)$ when conditioned on the event $\{\eta(0) = \eta(1) = L\}$. Combining all of the above, we obtain:

$$
\begin{aligned}
\widehat{G}_{p,q,r}\mu(WD) &= p(1-p-q)(1-r)\mu(WDW) + p(1-p-q)(1-r)\mu(DDW) + \{p + (1-p-q)(1-r)\} \\
&\quad (1-p-q)(1-r)\mu(LDW) + p(1-p-q)(1-r^2)\mu(WDD) + p(1-p-q)(1-r^2)\mu(DDD) \\
&\quad + \{p + (1-p-q)(1-r)\}(1-p-q)(1-r^2)\mu(LDD) + p(1-p-q)r(1-r)\mu(WDL) \\
&\quad + p(1-p-q)r(1-r)\mu(DDL) + \{p + (1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(LDL) \\
&\quad + p(1-p-q)(1-r)\mu(WWD) + p(1-p-q)(1-r)\mu(DWD) + \{p + (1-p-q)(1-r)\} \\
&\quad (1-p-q)(1-r)\mu(LWD) + \{p + (1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(WLD) + \\
&\quad \{p + (1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(DLD) + \{p + (1-p-q)(1-r^2)\}(1-p-q) \\
&\quad r(1-r)\mu(LLD) \\
&= p(1-p-q)(1-r)\mu(WDW) + p(1-p-q)(1-r)(2+r)\mu(DDW) + \{1-q-r+qr+2pr\} \\
&\quad (1-p-q)(1-r)\mu(LDW) + p(1-p-q)(1-r^2)\mu(DDD) + \{1-q-r^2+2pr+qr^2+pr^2\} \\
&\quad (1-p-q)(1-r)\mu(DDL) + (1-q-r+pr+qr)(1-p-q)r(1-r)\mu(LDL)
\end{aligned}
$$

$$
\begin{aligned}
&+p(1-p-q)(1-r)\mu(WWD)+p(1-p-q)(1-r)\mu(DWD)+(1-q-r+pr+qr)\\
&(1-p-q)(1-r)\mu(LWD)+(1-q-r+pr+qr)(1-p-q)r(1-r)\mu(WLD)+\\
&(1-q-r+pr+qr)(1-p-q)r(1-r)\mu(DLD)+(1-q-r^2+pr^2+qr^2)(1-p-q)\\
&r(1-r)\mu(LLD)\quad\text{(using reflection-invariance)}\\
=\ &\underbrace{p(1-p-q)(1-r)\mu(WDW)+p(1-p-q)(1-r)\mu(DDW)}_{(1)}\\
&\underbrace{+p(1-p-q)(1-r)(1+r)\mu(WDD)}_{(2)}\underbrace{+p(1-p-q)(1-r)\mu(LDW)}_{(1)}\\
&\underbrace{+(1-p-q-r+qr+2pr)(1-p-q)(1-r)\mu(LDW)}_{(3)}\underbrace{+p(1-p-q)(1-r^2)\mu(DDD)}_{(2)}\\
&\underbrace{+p(1-p-q)(1-r)(1+r)\mu(LDD)}_{(2)}\underbrace{+(1-p-q-r^2+pr+pr^2+qr^2)(1-p-q)}_{(3)}\\
&\underbrace{(1-r)\mu(LDD)+(1-q-r+pr+qr)(1-p-q)r(1-r)\mu(LDL)}_{(3)}\underbrace{+p(1-p-q)}_{(4)}\\
&\underbrace{(1-r)\mu(WWD)+p(1-p-q)(1-r)\mu(DWD)+p(1-p-q)(1-r)\mu(LWD)}_{(4)}\\
&+(1-p-q)^2(1-r)^2\mu(LWD)\underbrace{+(1-q-r^2+pr^2+qr^2)(1-p-q)r(1-r)\mu(WLD)}_{(5)}\\
&-r^2(1-r)^2(1-p-q)^2\mu(WLD)\underbrace{+(1-q-r^2+pr^2+qr^2)(1-p-q)r(1-r)\mu(DLD)}_{(5)}\\
&-r^2(1-r)^2(1-p-q)^2\mu(DLD)\underbrace{+(1-q-r^2+pr^2+qr^2)(1-p-q)r(1-r)\mu(LLD)}_{(5)}\\
=\ &\underbrace{p(1-p-q)(1-r)\mu(DW)}_{\text{combining terms underbraced by (1)}}\underbrace{+p(1-p-q)(1-r)(1+r)\mu(DD)}_{\text{combining terms underbraced by (2)}}\\
&\underbrace{+(1-p-q-r^2+pr+pr^2+qr^2)(1-p-q)(1-r)\mu(LD)}_{\text{combining terms underbraced by (3)}}-r(1-p-q)^2(1-r)^2\mu(LDW)\\
&-(1-r)^2(1-p-q)^2\mu(LDL)\underbrace{+p(1-p-q)(1-r)\mu(WD)}_{\text{combining terms underbraced by (4)}}+(1-p-q)^2(1-r)^2\mu(LWD)\\
&\underbrace{+(1-q-r^2+pr^2+qr^2)(1-p-q)r(1-r)\mu(LD)}_{\text{combining terms underbraced by (5)}}-r^2(1-r)^2(1-p-q)^2\mu(WLD)\\
&-r^2(1-r)^2(1-p-q)^2\mu(DLD)
\end{aligned}
$$

$$
\begin{aligned}
&= 2p(1-p-q)(1-r)\mu(WD) + p(1-p-q)(1-r)(1+r)\mu(DD) + (1-p-q+r+pr-qr \\
&\quad -r^2+pr^2+qr^2-r^3+pr^3+qr^3)(1-p-q)(1-r)\mu(LD) - r(1-p-q)^2(1-r)^2\mu(LDW) \\
&\quad -(1-r)^2(1-p-q)^2\mu(LDL) + (1-p-q)^2(1-r)^2\mu(LWD) - r^2(1-r)^2(1-p-q)^2 \\
&\quad \mu(WLD) - r^2(1-r)^2(1-p-q)^2\mu(DLD),
\end{aligned} \tag{4.8.13}
$$

where, in the last step, we combine terms using the reflection-invariance property of μ once more.

Recall that, so far, we have set $c_1 = c_2 = 1$, $\mathscr{C}_1 = (D)_0$ and $\mathscr{C}_2 = (W,D)_{0,1}$. The weight function, so far, equals $\mu(D)+\mu(WD)$, and the right side of the weight function inequality (of the form (4.8.2)) is given by

$$
\begin{aligned}
&\widehat{G}_{p,q,r}\mu(D) + \widehat{G}_{p,q,r}\mu(WD) \\
&= 2(1-p-q)(1-r)\mu(WD) + 2(1-p-q)r(1-r)\mu(LD) + (1-p-q)(1-r^2)\mu(DD) \\
&\quad + 2p(1-p-q)(1-r)\mu(WD) + p(1-p-q)(1-r)(1+r)\mu(DD) + (1-p-q+r+pr-qr \\
&\quad -r^2+pr^2+qr^2-r^3+pr^3+qr^3)(1-p-q)(1-r)\mu(LD) - r(1-p-q)^2(1-r)^2\mu(LDW) \\
&\quad -(1-r)^2(1-p-q)^2\mu(LDL) + (1-p-q)^2(1-r)^2\mu(LWD) - r^2(1-r)^2(1-p-q)^2 \\
&\quad \mu(WLD) - r^2(1-r)^2(1-p-q)^2\mu(DLD) \\
&= 2(1+p)(1-p-q)(1-r)\mu(WD) + (1+p)(1-p-q)(1-r)(1+r)\mu(DD) + (1-p-q+3r \\
&\quad +pr-qr-r^2+pr^2+qr^2-r^3+pr^3+qr^3)(1-p-q)(1-r)\mu(LD) - r(1-p-q)^2(1-r)^2\mu(LDW) \\
&\quad -(1-r)^2(1-p-q)^2\mu(LDL) + (1-p-q)^2(1-r)^2\mu(LWD) - r^2(1-r)^2(1-p-q)^2 \\
&\quad \mu(WLD) - r^2(1-r)^2(1-p-q)^2\mu(DLD).
\end{aligned} \tag{4.8.14}
$$

We now have to pause and focus on the coefficient of $\mu(LD)$ in (4.8.14) (we need not focus on the coefficients of $\mu(WD)$ and $\mu(DD)$ since we can readily see that the former is bounded above by 2 and the latter by 1). Writing

$$
pr-qr-r^2+pr^2+qr^2-r^3+pr^3+qr^3 = pr-qr-r^2(1-p-q)(1+r),
$$

we see that the coefficient of $\mu(LD)$ in (4.8.14) is given by

$$
\begin{aligned}
&1+2r-2(p+q)+(p+q)^2-3r^2-(p+q)r+3(p+q)r^2-(p+q)^2r \\
&\quad +pr(1-p-q)(1-r)-qr(1-p-q)(1-r)-r^2(1-p-q)^2(1-r)^2 \\
&= 1+2r-2(p+q)+(p+q)^2-3r^2-2qr+pr\{-1+(1-p-q)(1-r)\} \\
&\quad +r^2\{-(1-p-q)^2(1-r)^2+3p+4q-pq-q^2\}-p^2r-pqr
\end{aligned}
$$

$$\leqslant 1+2r-2(p+q)+(p+q)^2-3r^2-2qr \quad \text{when } p,q,r \text{ are sufficiently small,} \tag{4.8.15}$$

and when (4.2.1) holds, we have

$$2(p+q)-(p+q)^2-2r+3r^2+2qr \geqslant 4r-6r^2-3pr-6qr-2r+3r^2+2qr = r(2-3p-4q-3r) \geqslant 0$$

for all p and r sufficiently small, thus proving that the expression in (4.8.15), and consequently, the coefficient of $\mu(LD)$ in (4.8.14), is bounded above by 1.

We now continue with (4.8.14), and write (note that we make use of the decomposition $\mu(D) = \mu(WD)+\mu(LD)+\mu(DD)$ here):

$$\begin{aligned}
&\widehat{G}_{p,q,r}\mu(D)+\widehat{G}_{p,q,r}\mu(WD)\\
&= \mu(D)+\mu(WD)-2\{1-(1+p)(1-p-q)(1-r)\}\mu(WD)-\{1-(1+p)(1-p-q)(1-r)(1+r)\}\mu(DD)\\
&\quad -\{1-(1-p-q+3r+pr-qr-r^2+pr^2+qr^2-r^3+pr^3+qr^3)(1-p-q)(1-r)\}\mu(LD)\\
&\quad -r(1-p-q)^2(1-r)^2\mu(LDW)-(1-r)^2(1-p-q)^2\mu(LDL)+(1-p-q)^2(1-r)^2\mu(LWD)\\
&\quad -r^2(1-r)^2(1-p-q)^2\mu(WLD)-r^2(1-r)^2(1-p-q)^2\mu(DLD),
\end{aligned} \tag{4.8.16}$$

and given the observations made above regarding the coefficient of $\mu(LD)$ in (4.8.14), we conclude that, when p, q and r are all sufficiently small and (4.2.1) holds, the only term on the right side of (4.8.16), other than $\mu(D)$ and $\mu(WD)$, that has a non-negative coefficient is $(1-p-q)^2(1-r)^2\mu(LWD)$. Moreover, when p, q and r are all sufficiently small, we have $(1-p-q)^2(1-r)^2 = \Theta(1)$ (i.e. its order of magnitude is the same as the constant 1 when p, q and r approach 0). Hence, it seems sensible to set $\mathscr{C}_3 = (L,W,D)_{0,1,2}$ and $c_3 = 1$ as we continue with the construction of our weight function following the form given by (4.8.1).

The weight function we have obtained so far equals

$$w_0(\mu) = \mu(D)+\mu(WD)+\mu(LWD). \tag{4.8.17}$$

The corresponding weight function inequality, of the form given by (4.8.2), can be derived from (4.8.16) as follows:

$$\begin{aligned}
w_0(\widehat{G}_{p,q,r}\mu) &= \widehat{G}_{p,q,r}\mu(D)+\widehat{G}_{p,q,r}\mu(WD)+\widehat{G}_{p,q,r}\mu(LWD)\\
&= \mu(D)+\mu(WD)-2\{1-(1+p)(1-p-q)(1-r)\}\mu(WD)-\{1-(1+p)(1-p-q)(1-r)\\
&\quad (1+r)\}\mu(DD)-\{1-(1-p-q+3r+pr-qr-r^2+pr^2+qr^2-r^3+pr^3+qr^3)(1-p-q)\\
&\quad (1-r)\}\mu(LD)-r(1-p-q)^2(1-r)^2\mu(LDW)-(1-r)^2(1-p-q)^2\mu(LDL)+(1-p-q)^2
\end{aligned}$$

$$
\begin{aligned}
&(1-r)^2\mu(LWD) - r^2(1-r)^2(1-p-q)^2\mu(WLD) - r^2(1-r)^2(1-p-q)^2\mu(DLD)\\
&+\widehat{G}_{p,q,r}\mu(LWD)\\
={}&\mu(D)+\mu(WD)+\mu(LWD)-2\{1-(1+p)(1-p-q)(1-r)\}\mu(WD)\\
&-\{1-(1+p)(1-p-q)(1-r)(1+r)\}\mu(DD)-\{1-(1-p-q+3r+pr-qr-r^2+pr^2\\
&+qr^2-r^3+pr^3+qr^3)(1-p-q)(1-r)\}\mu(LD)-r(1-p-q)^2(1-r)^2\mu(LDW)\\
&-(1-r)^2(1-p-q)^2\mu(LDL)-\{1-(1-p-q)^2(1-r)^2\}\mu(LWD)-r^2(1-r)^2(1-p-q)^2\\
&\mu(WLD)-r^2(1-r)^2(1-p-q)^2\mu(DLD)+\widehat{G}_{p,q,r}\mu(LWD)\\
={}&w_0(\mu)-2\{1-(1+p)(1-p-q)(1-r)\}\mu(WD)-\{1-(1+p)(1-p-q)(1-r)(1+r)\}\\
&\mu(DD)-\{1-(1-p-q+3r+pr-qr-r^2+pr^2+qr^2-r^3+pr^3+qr^3)(1-p-q)(1-r)\}\\
&\mu(LD)-r(1-p-q)^2(1-r)^2\mu(LDW)-(1-r)^2(1-p-q)^2\mu(LDL)\\
&-\{1-(1-p-q)^2(1-r)^2\}\mu(LWD)-r^2(1-r)^2(1-p-q)^2\mu(WLD)\\
&-r^2(1-r)^2(1-p-q)^2\mu(DLD)+\widehat{G}_{p,q,r}\mu(LWD). \qquad (4.8.18)
\end{aligned}
$$

The expression in (4.8.18) necessitates the computation of $\widehat{G}_{p,q,r}\mu(LWD)$, which is the probability of the event $\{\widehat{G}_{p,q,r}\eta(0)=L, \widehat{G}_{p,q,r}\eta(1)=W, \widehat{G}_{p,q,r}\eta(2)=D\}$, where $\eta=(\eta(x):x\in\mathbb{Z})$ is a (random) configuration following the probability distribution μ. In order for the event $\{\widehat{G}_{p,q,r}\eta(2)=D\}$ to take place, at least one of the events $\{\eta(2)=D\}$ and $\{\eta(3)=D\}$ must occur. We shall reveal the terms in the expansion of $\widehat{G}_{p,q,r}\mu(LWD)$ step by step, and in each such step, we shall also see how to combine them with the existing terms (with non-positive coefficients) on the right side of (4.8.18).

1. The contribution to $\widehat{G}_{p,q,r}\mu(LWD)$ of the event $\{\eta(2)=L,\eta(3)=D\}$ is given by

$$
\begin{aligned}
&(1-p)\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(WWLD)+\{q+(1-p-q)r\}\{p+(1-p-q)\\
&(1-r^2)\}(1-p-q)r(1-r)\mu(WLLD)+\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}(1-p-q)\\
&r(1-r)\mu(LWLD)+\{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r^2)\}(1-p-q)r(1-r)\mu(LLLD)\\
&+\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(WDLD)+\{q+(1-p-q)r\}\\
&\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(DWLD)+\{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r)\}\\
&(1-p-q)r(1-r)\mu(LDLD)+\{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r^2)\}(1-p-q)r\\
&(1-r)\mu(DLLD)+\{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(DDLD). \qquad (4.8.19)
\end{aligned}
$$

A quick comparison of the coefficients in the various terms of (4.8.19) reveals that the coefficient of each of $\mu(LWLD)$, $\mu(LLLD)$, ..., $\mu(DDLD)$ is bounded above by the coefficient of $\mu(WLLD)$, and

$$
\begin{aligned}
&(1-p)\{p+(1-p-q)(1-r)\}-\{q+(1-p-q)r\}\{p+(1-p-q)(1-r^2)\}\\
&=(1-p)\{p+(1-p-q)(1-r)\}-\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)+r(1-p-q)(1-r)\}\\
&=(1-p-q)(1-r)\{p+(1-p-q)(1-r)\}-\{q+(1-p-q)r\}r(1-p-q)(1-r)\\
&=(1-p-q)(1-r)\{p+(1-p-q)(1-r-r^2)-qr\}
\end{aligned}
$$

is non-negative whenever p, q and r are all sufficiently small, so that the coefficient of $\mu(WWLD)$ exceeds that of $\mu(WLLD)$ in (4.8.19). For each cylinder set $\mathscr{C}$ that appears in the expression in (4.8.19), let us denote the corresponding coefficient by $\alpha_{\mathscr{C}}$ (for instance, the coefficient of $\mu(DLLD)$ is $\alpha_{DLLD}=\{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r^2)\}(1-p-q)r(1-r)$), and let us define

$$\alpha'_{\mathscr{C}}=\alpha_{\mathscr{C}}-\alpha_{WWLD}=\alpha_{\mathscr{C}}-(1-p)\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)$$

for each cylinder set $\mathscr{C}$, other than $(W,W,L,D)_{0,1,2,3}$, appearing in (4.8.19). From our discussion above, it is evident that when p, q and r are all small enough, each $\alpha'_{\mathscr{C}}\leqslant 0$ for each cylinder set $\mathscr{C}$, other than $(W,W,L,D)_{0,1,2,3}$, appearing in (4.8.19). We can now rewrite (4.8.19) as

$$
\begin{aligned}
&(1-p)\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(LD)+\alpha'_{WLLD}\mu(WLLD)\\
&+\alpha'_{LWLD}\mu(LWLD)+\alpha'_{LLLD}\mu(LLLD)+\alpha'_{WDLD}\mu(WDLD)+\alpha'_{DWLD}\mu(DWLD)\\
&+\alpha'_{LDLD}\mu(LDLD)+\alpha'_{DLLD}\mu(DLLD)+\alpha'_{DDLD}\mu(DDLD).
\end{aligned}
\tag{4.8.20}
$$

We combine the term involving $\mu(LD)$ in (4.8.20) with the term involving $\mu(LD)$ in (4.8.18), to obtain, when each of p, q and r is sufficiently small:

$$
\begin{aligned}
&-\{1-(1-p-q+3r+pr-qr-r^2+pr^2+qr^2-r^3+pr^3+qr^3)(1-p-q)(1-r)\}\mu(LD)\\
&+(1-p)\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(LD)\\
&=[-1+(1-p-q+4r-2qr-2r^2+pqr+3pr^2+2qr^2-r^3+pr^3+qr^3-p^2r^2-pqr^2)\\
&(1-p-q)(1-r)]\mu(LD)\\
&=[-1+\{1-p-q+4r-qr(2-p-2r)-r^2(2-3p)-r^3(1-p-q)-p^2r^2-pqr^2\}\\
&(1-p-q)(1-r)]\mu(LD)
\end{aligned}
\tag{4.8.21}
$$

$$\leqslant [-1+(1-p-q+4r)(1-p-q)(1-r)]\mu(LD) \quad \text{when } p,r \text{ are sufficiently small;}$$
$$= [-1+\{1+3r-4r^2-2(p+q)+(p+q)^2-(p+q)r-(p+q)r(1+p+q-4r)\}]\mu(LD)$$
$$\leqslant \{3r-4r^2-2(p+q)+(p+q)^2-(p+q)r\}]\mu(LD) \quad \text{when } r \text{ sufficiently small.} \tag{4.8.22}$$

When p, q and r are sufficiently small and the inequality in (4.2.1) holds, we have

$$\begin{aligned}
&2(p+q)-(p+q)^2-3r+4r^2+(p+q)r \\
\geqslant\, &4r-6r^2-3pr-6qr-3r+4r^2+(p+q)r = r(1-2p-5q-2r) \geqslant 0,
\end{aligned}$$

so that the coefficient of $\mu(LD)$ in (4.8.22) is non-positive, and consequently, the coefficient of $\mu(LD)$ in (4.8.21), is non-positive.

Incorporating (4.8.20) and (4.8.21) into (4.8.18), we obtain (the equality of (4.8.18) now becomes an inequality as we ignore all the terms with non-positive coefficients in (4.8.20)):

$$\begin{aligned}
w_0(\widehat{G}_{p,q,r}\mu) \leqslant w_0(\mu) &- 2\{1-(1+p)(1-p-q)(1-r)\}\mu(WD) - \{1-(1+p)(1-p-q) \\
&(1-r)(1+r)\}\mu(DD) - [1-\{1-p-q+4r-qr(2-p-2r)-r^2(2-3p) \\
&-r^3(1-p-q)-p^2r^2-pqr^2\}(1-p-q)(1-r)]\mu(LD) - r(1-p-q)^2(1-r)^2 \\
&\mu(LDW) - (1-r)^2(1-p-q)^2\mu(LDL) - \{1-(1-p-q)^2(1-r)^2\}\mu(LWD) \\
&-r^2(1-r)^2(1-p-q)^2\mu(WLD) - r^2(1-r)^2(1-p-q)^2\mu(DLD) + A_1,
\end{aligned} \tag{4.8.23}$$

where A_1 consists of all those terms in the expansion of $\widehat{G}_{p,q,r}\mu(LWD)$ that do not arise as a contribution from the event $\{\eta(2)=L, \eta(3)=D\}$.

2. The contribution to $\widehat{G}_{p,q,r}\mu(LWD)$ of the event $\{\eta(1)=\eta(2)=D\}$ (note that these terms are present in the quantity A_1 defined above) is given by:

$$\begin{aligned}
&\{q+(1-p-q)r\}p(1-p-q)(1-r)\mu(WDDW) + \{q+(1-p-q)r^2\}p(1-p-q) \\
&(1-r)\mu(DDDW) + \{q+(1-p-q)r^2\}p(1-p-q)(1-r)\mu(LDDW) + \{q+(1-p-q)r\} \\
&p(1-p-q)(1-r^2)\mu(WDDD) + \{q+(1-p-q)r^2\}p(1-p-q)(1-r^2)\mu(DDDD) \\
&+\{q+(1-p-q)r^2\}p(1-p-q)(1-r^2)\mu(LDDD) + \{q+(1-p-q)r\}p(1-p-q) \\
&r(1-r)\mu(WDDL) + \{q+(1-p-q)r^2\}p(1-p-q)r(1-r)\mu(DDDL) \\
&+\{q+(1-p-q)r^2\}p(1-p-q)r(1-r)\mu(LDDL).
\end{aligned} \tag{4.8.24}$$

A quick comparison of the coefficients of the various terms in (4.8.24) reveals that the coefficient of $\mu(WDDD)$, namely, $\{q+(1-p-q)r\}p(1-p-q)(1-r)(1+r)$, is the largest. Let $\alpha_{\mathscr{C}}$ denote the coefficient of $\mu(\mathscr{C})$ in (4.8.24), for every cylinder set $\mathscr{C}$ that appears in (4.8.24), and set

$$\alpha'_{\mathscr{C}} = \alpha_{\mathscr{C}} - \alpha_{WDDD} = \alpha_{\mathscr{C}} - \{q+(1-p-q)r\}p(1-p-q)(1-r)(1+r)$$

for each cylinder set, other than $(W,D,D,D)_{0,1,2,3}$, that appears in (4.8.24). By our observation above, we know that each $\alpha'_{\mathscr{C}} \leqslant 0$, and we can rewrite (4.8.24) as:

$$\begin{aligned}&\{q+(1-p-q)r\}p(1-p-q)(1-r)(1+r)\mu(DD) + \alpha'_{WDDW}\mu(WDDW) + \alpha'_{DDDW}\mu(DDDW)\\&+\alpha'_{LDDW}\mu(LDDW) + \alpha'_{DDDD}\mu(DDDD) + \alpha'_{LDDD}\mu(LDDD) + \alpha'_{WDDL}\mu(WDDL)\\&+\alpha'_{DDDL}\mu(DDDL) + \alpha'_{LDDL}\mu(LDDL). \qquad (4.8.25)\end{aligned}$$

Before proceeding with combining (4.8.25) with (4.8.23), we observe the following:

$$\begin{aligned}&\{(1-p-q)(1-r)(1+r)\}^{-1} = \{1-p-q-r^2+pr^2+qr^2\}^{-1} = \sum_{j=0}^{\infty}(p+q+r^2-pr^2-qr^2)^j\\&\geqslant 1+p+q+r^2-pr^2-qr^2+(p+q+r^2-pr^2-qr^2)^2\\&= 1+p+q+r^2+p^2+q^2+2pq+p^2r^4-2p^2r^2+2pqr^4-4pqr^2-2pr^4+pr^2+q^2r^4\\&\quad -2q^2r^2-2qr^4+qr^2+r^4 \qquad (4.8.26)\\&= (1+p+pq+pr-p^2r-pqr)+\{(p-r)^2+pr\}+p^2r(1-2r)+pr^2(1-2r^2)+pqr(1-4r)\\&\quad +q^2(1-2r^2)+qr^2(1-2r^2)+q+pq+p^2r^4+2pqr^4+q^2r^4+r^4\\&\geqslant 1+p+pq+pr-p^2r-pqr \text{ when } p,q,r \text{ are all sufficiently small.} \qquad (4.8.27)\end{aligned}$$

Combining the term involving $\mu(DD)$ in (4.8.25) with the term involving $\mu(DD)$ in (4.8.23), we obtain:

$$\begin{aligned}&-\{1-(1+p)(1-p-q)(1-r)(1+r)\}\mu(DD)+\{q+(1-p-q)r\}p(1-p-q)(1-r)(1+r)\mu(DD)\\&=[-1+(1+p+pq+pr-p^2r-pqr)(1-p-q)(1-r)(1+r)]\mu(DD), \qquad (4.8.28)\end{aligned}$$

and the coefficient of $\mu(DD)$ in (4.8.28) is non-positive by the inequality we have derived in (4.8.27). Incorporating (4.8.28) and (4.8.25) into (4.8.23) (and once again, as in the previous step, ignoring all terms of (4.8.25) that have non-positive coefficients), we obtain:

$$\begin{aligned} w_0(\widehat{G}_{p,q,r}\mu) \leqslant w_0(\mu) &- 2\{1-(1+p)(1-p-q)(1-r)\}\mu(WD) - \{1-(1+p+pq+pr-p^2r-pqr) \\ &(1-p-q)(1-r)(1+r)\}\mu(DD) - [1-\{1-p-q+4r-qr(2-p-2r)-r^2(2-3p) \\ &-r^3(1-p-q)-p^2r^2-pqr^2\}(1-p-q)(1-r)]\mu(LD) - r(1-p-q)^2(1-r)^2 \\ &\mu(LDW) - (1-r)^2(1-p-q)^2\mu(LDL) - \{1-(1-p-q)^2(1-r)^2\}\mu(LWD) \\ &- r^2(1-r)^2(1-p-q)^2\mu(WLD) - r^2(1-r)^2(1-p-q)^2\mu(DLD) + A_2, \end{aligned} \tag{4.8.29}$$

where A_2 consists of all those terms in the expansion of $\widehat{G}_{p,q,r}\mu(LWD)$ that do not arise as a contribution from either of the events $\{\eta(2)=L,\eta(3)=D\}$ and $\{\eta(1)=\eta(2)=D\}$.

3. Next, we find the contribution to $\widehat{G}_{p,q,r}\mu(LWD)$ of the events $\{\eta(1)=L,\eta(2)=W,\eta(3)=D\}$ and $\{\eta(0)=L,\eta(1)=W,\eta(2)=D\}$ (note that these contributions are a part of A_2 defined above):

$$\begin{aligned} &\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}(1-p-q)(1-r)\mu(WLWD) + \{q+(1-p-q)r^2\} \\ &\{p+(1-p-q)(1-r)\}(1-p-q)(1-r)\mu(DLWD) + \{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r)\} \\ &(1-p-q)(1-r)\mu(LLWD) + \{q+(1-p-q)r\}p(1-p-q)(1-r)\mu(LWDW) + \{q+(1-p-q)r\} \\ &p(1-p-q)(1-r^2)\mu(LWDD) + \{q+(1-p-q)r\}p(1-p-q)r(1-r)\mu(LWDL) \\ = &\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}(1-p-q)(1-r)\mu(LWD) + \alpha'_{DLWD}\mu(DLWD) + \alpha'_{LLWD} \\ &\mu(LLWD) + \{q+(1-p-q)r\}p(1-p-q)(1-r^2)\mu(LWD) + \alpha'_{LWDW}\mu(LWDW) + \alpha'_{LWDL}\mu(LWDL) \\ = &\{q+(1-p-q)r\}(1-p-q)(1-r)\{1+p-q-r+2pr+qr\}\mu(LWD) + \alpha'_{DLWD}\mu(DLWD) \\ &+\alpha'_{LLWD}\mu(LLWD) + \alpha'_{LWDW}\mu(LWDW) + \alpha'_{LWDL}\mu(LWDL), \end{aligned} \tag{4.8.30}$$

where

$$\begin{aligned} \alpha'_{DLWD} = \alpha'_{LLWD} &= -r(1-p-q)^2(1-r)^2\{p+(1-p-q)(1-r)\}, \\ \alpha'_{LWDW} &= -r\{q+(1-p-q)r\}p(1-p-q)(1-r), \\ \alpha'_{LWDL} &= -\{q+(1-p-q)r\}p(1-p-q)(1-r) \end{aligned}$$

are all non-positive coefficients. Before we proceed to incorporate (4.8.30) into (4.8.29), we observe, much like the derivation carried out for (4.8.27), that

$$\{(1-p-q)(1-r)\}^{-1} = (1-p-q-r+pr+qr)^{-1}$$

$$
\begin{aligned}
&\geqslant 1+(p+q+r-pr-qr)+(p+q+r-pr-qr)^2\\
&= 1+p+q+r+p^2+q^2+r^2+2pq+pr+qr-4pqr-2p^2r-2pr^2-2q^2r-2qr^2\\
&\quad +p^2r^2+q^2r^2+2pqr^2 \qquad (4.8.31)\\
&= (1-p-q^2-r^2-2qr+pq+pr+2pqr+3pr^2+2qr^2+2q^2r-p^2r-3pqr^2\\
&\quad -q^2r^2-2p^2r^2)+2p+q+r+2q^2+pq+qr(3-6p-4q-4r)+r^2(2-5p)\\
&\quad +p^2(1-r)+3p^2r^2+2q^2r^2+5pqr^2\\
&\geqslant (1-p-q^2-r^2-2qr+pq+pr+2pqr+3pr^2+2qr^2+2q^2r-p^2r-3pqr^2\\
&\quad -q^2r^2-2p^2r^2) \text{ for all } p,q,r \text{ sufficiently small.} \qquad (4.8.32)
\end{aligned}
$$

Combining the term involving $\mu(LWD)$ in (4.8.30) with the term involving $\mu(LWD)$ in (4.8.29), we obtain:

$$
\begin{aligned}
&\{q+(1-p-q)r\}(1-p-q)(1-r)\{1+p-q-r+2pr+qr\}\mu(LWD)\\
&\quad -\{1-(1-p-q)^2(1-r)^2\}\mu(LWD)\\
&= [-1+\{(q+r-pr-qr)(1+p-q-r+2pr+qr)+(1-p-q)(1-r)\}\\
&\quad (1-p-q)(1-r)]\mu(LWD)\\
&= [-1+(1-p-q^2-r^2-2qr+pq+pr+2pqr+3pr^2+2qr^2+2q^2r-p^2r-3pqr^2\\
&\quad -q^2r^2-2p^2r^2)(1-p-q)(1-r)]\mu(LWD), \qquad (4.8.33)
\end{aligned}
$$

and by (4.8.32), it is immediate that the coefficient of $\mu(LWD)$ in (4.8.33) is non-positive. Incorporating (4.8.33) and (4.8.30) into (4.8.29) (and, as before, ignoring all terms with non-positive coefficients in (4.8.30)), we get:

$$
\begin{aligned}
w_0(\widehat{G}_{p,q,r}\mu) \leqslant{} & w_0(\mu)-2\{1-(1+p)(1-p-q)(1-r)\}\mu(WD)-\{1-(1+p+pq+pr-p^2r-pqr)\\
& (1-p-q)(1-r)(1+r)\}\mu(DD)-[1-\{1-p-q+4r-qr(2-p-2r)-r^2(2-3p)\\
& -r^3(1-p-q)-p^2r^2-pqr^2\}(1-p-q)(1-r)]\mu(LD)-r(1-p-q)^2(1-r)^2\\
& \mu(LDW)-(1-r)^2(1-p-q)^2\mu(LDL)-[1-(1-p-q^2-r^2-2qr+pq+pr+2pqr\\
& +3pr^2+2qr^2+2q^2r-p^2r-3pqr^2-q^2r^2-2p^2r^2)(1-p-q)(1-r)]\mu(LWD)\\
& -r^2(1-r)^2(1-p-q)^2\mu(WLD)-r^2(1-r)^2(1-p-q)^2\mu(DLD)+A_3, \qquad (4.8.34)
\end{aligned}
$$

where A_3 is the sum of all those terms in the expansion of $\widehat{G}_{p,q,r}\mu(LWD)$ that do not arise as contributions from any of the events considered so far, i.e. $\{\eta(2)=L,\eta(3)=D\}$, $\{\eta(1)=$

$\eta(2)=D\}$, $\{\eta(0)=L, \eta(1)=W, \eta(2)=D\}$ and $\{\eta(1)=L, \eta(2)=W, \eta(3)=D\}$.

4. The contribution to $\widehat{G}_{p,q,r}\mu(LWD)$ of the event $\{\eta(1)=L, \eta(2)=D, \eta(3)=L\}$ equals

$$\begin{aligned}
&\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(WLDL)\\
&+\{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(DLDL)\\
&+\{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(LLDL)\\
=&\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(LDL)\\
&+\alpha'_{DLDL}\mu(DLDL)+\alpha'_{LLDL}\mu(LLDL),
\end{aligned}\tag{4.8.35}$$

where $\alpha'_{DLDL}=\alpha'_{LLDL}=-\{p+(1-p-q)(1-r)\}(1-p-q)^2r^2(1-r)^2$ are both non-positive coefficients. Combining the term involving $\mu(LDL)$ in (4.8.35) with the term involving $\mu(LDL)$ in (4.8.34), we obtain:

$$\begin{aligned}
&-(1-r)^2(1-p-q)^2\mu(LDL)+\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}(1-p-q)r(1-r)\mu(LDL)\\
&=-(1-p-q)(1-r)[(1-r)(1-p-q)-\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}r]\mu(LDL)\\
&=-(1-p-q)(1-r)[(1-r)(1-p-q)\{1-qr-r^2(1-p-q)\}-pqr-pr^2(1-p-q)]\mu(LDL),
\end{aligned}\tag{4.8.36}$$

which is clearly non-positive when p, q and r are sufficiently small. Incorporating (4.8.36) and (4.8.35) into (4.8.34) (and ignoring the terms with non-positive coefficients in (4.8.35)), we obtain:

$$\begin{aligned}
w_0(\widehat{G}_{p,q,r}\mu)\leqslant{}& w_0(\mu)-2\{1-(1+p)(1-p-q)(1-r)\}\mu(WD)-\{1-(1+p+pq+pr-p^2r-pqr)\\
&(1-p-q)(1-r)(1+r)\}\mu(DD)-[1-\{1-p-q+4r-qr(2-p-2r)-r^2(2-3p)\\
&-r^3(1-p-q)-p^2r^2-pqr^2\}(1-p-q)(1-r)]\mu(LD)-r(1-p-q)^2(1-r)^2\\
&\mu(LDW)-(1-p-q)(1-r)[(1-r)(1-p-q)\{1-qr-r^2(1-p-q)\}-pqr\\
&-pr^2(1-p-q)]\mu(LDL)-[1-(1-p-q^2-r^2-2qr+pq+pr+2pqr\\
&+3pr^2+2qr^2+2q^2r-p^2r-3pqr^2-q^2r^2-2p^2r^2)(1-p-q)(1-r)]\mu(LWD)\\
&-r^2(1-r)^2(1-p-q)^2\mu(WLD)-r^2(1-r)^2(1-p-q)^2\mu(DLD)+A_4,
\end{aligned}\tag{4.8.37}$$

where A_4 consists of all those terms in the expansion of $\widehat{G}_{p,q,r}\mu(LWD)$ that do not arise as contributions from any of the events considered so far, i.e. $\{\eta(2)=L, \eta(3)=D\}$, $\{\eta(1)=\eta(2)=D\}$, $\{\eta(0)=L, \eta(1)=W, \eta(2)=D\}$, $\{\eta(1)=L, \eta(2)=W, \eta(3)=D\}$ and $\{\eta(1)=$

$L, \eta(2) = D, \eta(3) = L\}$.

5. Recall that in the third step (i.e. two steps back), we have already taken into account the term involving $\mu(LWDL)$ appearing in the expansion of $\widehat{G}_{p,q,r}\mu(LWD)$. We now consider the remaining two terms that arise as contributions of the event $\{\eta(1) = W, \eta(2) = D, \eta(3) = L\}$ to $\widehat{G}_{p,q,r}\mu(LWD)$:

$$\begin{aligned} &(1-p)p(1-p-q)r(1-r)\mu(WWDL) + \{q + (1-p-q)r\}p(1-p-q)r(1-r)\mu(DWDL) \\ = &(1-p)p(1-p-q)r(1-r)\mu(WDL) - pr(1-p-q)^2(1-r)^2\mu(DWDL) \\ &-(1-p)p(1-p-q)r(1-r)\mu(LWDL), \end{aligned} \tag{4.8.38}$$

so that when we combine the term involving $\mu(WDL)$ in (4.8.38) with the term involving $\mu(LDW)$ in (4.8.37) (using reflection invariance), we obtain:

$$\begin{aligned} &-r(1-p-q)^2(1-r)^2\mu(LDW) + (1-p)p(1-p-q)r(1-r)\mu(WDL) \\ = &-r(1-p-q)(1-r)\{(1-p-q)(1-r) - p(1-p)\}\mu(WDL), \end{aligned} \tag{4.8.39}$$

and the coefficient of $\mu(WDL)$ in (4.8.39) is assuredly non-positive when p, q and r are sufficiently small. Incorporating (4.8.38) and (4.8.39) into (4.8.37) (and, as before, ignoring the terms with non-positive coefficients in (4.8.38)), we obtain:

$$\begin{aligned} w_0(\widehat{G}_{p,q,r}\mu) \leqslant\ & w_0(\mu) - 2\{1-(1+p)(1-p-q)(1-r)\}\mu(WD) - \{1-(1+p+pq+pr-p^2r-pqr) \\ &(1-p-q)(1-r)(1+r)\}\mu(DD) - [1-\{1-p-q+4r-qr(2-p-2r)-r^2(2-3p) \\ &-r^3(1-p-q)-p^2r^2-pqr^2\}(1-p-q)(1-r)]\mu(LD) - r(1-p-q)(1-r) \\ &\{(1-p-q)(1-r)-p(1-p)\}\mu(LDW) - (1-p-q)(1-r)[(1-r)(1-p-q) \\ &\{1-qr-r^2(1-p-q)\} - pqr - pr^2(1-p-q)]\mu(LDL) - [1-(1-p-q^2-r^2-2qr \\ &+pq+pr+2pqr+3pr^2+2qr^2+2q^2r-p^2r-3pqr^2-q^2r^2-2p^2r^2)(1-p-q)(1-r)] \\ &\mu(LWD) - r^2(1-r)^2(1-p-q)^2\mu(WLD) - r^2(1-r)^2(1-p-q)^2\mu(DLD) + A_5, \end{aligned} \tag{4.8.40}$$

where A_5 is the sum of all those terms in the expansion of $\widehat{G}_{p,q,r}\mu(LWD)$ that do not arise as contributions from any of the events considered so far, i.e. $\{\eta(2) = L, \eta(3) = D\}$, $\{\eta(1) = \eta(2) = D\}$, $\{\eta(0) = L, \eta(1) = W, \eta(2) = D\}$, $\{\eta(1) = L, \eta(2) = W, \eta(3) = D\}$, $\{\eta(1) = L, \eta(2) = D, \eta(3) = L\}$ and $\{\eta(1) = W, \eta(2) = D, \eta(3) = L\}$.

6. We now consider the contribution of the event $\{\eta(1)=L,\eta(2)=D,\eta(3)=W\}$ as well as that of the event $\{\eta(1)=W,\eta(2)=D,\eta(3)=W\}$ (note that we have already considered the contribution of $\{\eta(0)=L,\eta(1)=W,\eta(2)=D\}$, hence we must exclude the contribution arising from the intersection of these two events) to $\widehat{G}_{p,q,r}\mu(LWD)$:

$$
\begin{aligned}
&\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}(1-p-q)(1-r)\mu(WLDW)\\
&+\{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r)\}(1-p-q)(1-r)\mu(DLDW)\\
&+\{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r)\}(1-p-q)(1-r)\mu(LLDW)\\
&+(1-p)p(1-p-q)(1-r)\mu(WWDW)+\{q+(1-p-q)r\}p(1-p-q)(1-r)\mu(DWDW)\\
\leqslant\ &\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}(1-p-q)(1-r)\mu(DW)+\alpha'_{DLDW}\mu(DLDW)\\
&+\alpha'_{LLDW}\mu(LLDW)+\{p(1-p)-(q+r-pr-qr)(1-q-r+pr+qr)\}(1-p-q)(1-r)\\
&\mu(WWDW)+\alpha'_{DWDW}\mu(DWDW),
\end{aligned}
\tag{4.8.41}
$$

(here, we make use of the fact that the sum of $\mu(WLDW)$, $\mu(DLDW)$, $\mu(LLDW)$, $\mu(WWDW)$ and $\mu(DWDW)$ is bounded above by $\mu(DW)$) where

$$
\begin{aligned}
&\alpha'_{DLDW}=\alpha'_{LLDW}=-r\{p+(1-p-q)(1-r)\}(1-p-q)^2(1-r)^2,\\
&\alpha'_{DWDW}=-\{q+(1-p-q)r\}(1-p-q)^2(1-r)^2,
\end{aligned}
$$

each of which is a non-positive coefficient. It is worthwhile to note here the difference between (4.8.41) and similar equations previously derived, such as (4.8.20), (4.8.25) etc.: the coefficient of $\mu(WWDW)$ in (4.8.41) is not necessarily non-positive. It is, in fact, non-positive if and only if we have $(q+r-pr-qr)(1-q-r+pr+qr)\geqslant p(1-p)$, which, provided p, q and r are all sufficiently small (in particular, as long as each of p and $q+r(1-p-q)$ is bounded above by $1/2$), is true if and only if $p\leqslant q+r(1-p-q)$. The coefficient of $\mu(WWDW)$ in (4.8.41) is, thus, non-positive if and only if one of (4.2.2), (4.2.3) and (4.2.4) is true.

Next, we consider the contribution to $\widehat{G}_{p,q,r}\mu(LWD)$ of terms arising from the event $\{\eta(1)\neq L,\eta(2)=W,\eta(3)=D\}$:

$$
\begin{aligned}
&(1-p)p(1-p-q)(1-r)\mu(WWWD)+\{q+(1-p-q)r\}p(1-p-q)(1-r)\mu(DWWD)\\
&+\{q+(1-p-q)r\}p(1-p-q)(1-r)\mu(LWWD)+\{q+(1-p-q)r\}p(1-p-q)(1-r)\\
&\mu(WDWD)+\{q+(1-p-q)r^2\}p(1-p-q)(1-r)\mu(DDWD)+\{q+(1-p-q)r^2\}\\
&p(1-p-q)(1-r)\mu(LDWD)
\end{aligned}
$$

$$\leqslant p(1-p-q)^2(1-r)^2\mu(WWWD)+\{q+(1-p-q)r\}p(1-p-q)(1-r)\mu(WD)$$
$$+\alpha'_{DDWD}\mu(DDWD)+\alpha'_{LDWD}\mu(LDWD), \tag{4.8.42}$$

where $\alpha'_{DDWD}=\alpha'_{LDWD}=-pr(1-p-q)^2(1-r)^2$ (in order to obtain the last inequality, we make use of the fact that the sum of $\mu(WWWD)$, $\mu(DWWD)$, $\mu(LWWD)$, $\mu(WDWD)$, $\mu(DDWD)$ and $\mu(LDWD)$ is bounded above by $\mu(WD)$).

Before proceeding further, we observe, from (4.8.31), that

$$\begin{aligned}
&2\{(1-p-q)(1-r)\}^{-1}\geqslant 2+2p+2q+2r+2p^2+2q^2+2r^2+4pq+2pr+2qr-8pqr-4p^2r\\
&\quad -4pr^2-4q^2r-4qr^2+2p^2r^2+2q^2r^2+4pqr^2\\
&=\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2+pqr+2pr^2-p^2r-2pqr^2-p^2r^2-q^2r^2\}\\
&\quad +q+r+2p^2+3q^2+3r^2+3pq(1-3r)+pr(2-3p-6r)+qr(5-6q-6r)+3p^2r^2\\
&\quad +3q^2r^2+6pqr^2\\
&\geqslant\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2+pqr+2pr^2-p^2r-2pqr^2-p^2r^2-q^2r^2\}
\end{aligned} \tag{4.8.43}$$

whenever p, q and r are sufficiently small.

Combining the term involving $\mu(WD)$ in (4.8.42), the term involving $\mu(DW)$ in (4.8.41), and the term involving $\mu(WD)$ in (4.8.40), and making use of the reflection-invariance of μ, we obtain:

$$\begin{aligned}
&-2\{1-(1+p)(1-p-q)(1-r)\}\mu(WD)+\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}\\
&(1-p-q)(1-r)\mu(DW)+\{q+(1-p-q)r\}p(1-p-q)(1-r)\mu(WD)\\
&=[-2+(1-p-q)(1-r)\{2(1+p)+(q+r-pr-qr)(1-q-r+pr+qr)\\
&\quad +(q+r-pr-qr)p\}]\mu(WD)\\
&=[-2+(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2\\
&\quad +pqr+2pr^2-p^2r-2pqr^2-p^2r^2-q^2r^2\}]\mu(WD),
\end{aligned} \tag{4.8.44}$$

and by the inequality derived in (4.8.43), it is immediate that the coefficient of $\mu(WD)$ in (4.8.44) is non-positive. Incorporating (4.8.44), (4.8.41) and (4.8.42) into (4.8.40), and ignoring those terms from both (4.8.41) and (4.8.42) whose coefficients are *assuredly* non-positive, we obtain:

$$w_0(\widehat{G}_{p,q,r}\mu)\leqslant w_0(\mu)-[2-(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2$$

$$
\begin{aligned}
&+pqr+2pr^2-p^2r-2pqr^2-p^2r^2-q^2r^2\}]\mu(WD)-\{1-(1+p+pq+pr\\
&-p^2r-pqr)(1-p-q)(1-r)(1+r)\}\mu(DD)-[1-\{1-p-q+4r\\
&-qr(2-p-2r)-r^2(2-3p)-r^3(1-p-q)-p^2r^2-pqr^2\}(1-p-q)\\
&(1-r)]\mu(LD)-r(1-p-q)(1-r)\{(1-p-q)(1-r)-p(1-p)\}\mu(LDW)\\
&-(1-p-q)(1-r)[(1-r)(1-p-q)\{1-qr-r^2(1-p-q)\}-pqr\\
&-pr^2(1-p-q)]\mu(LDL)-[1-(1-p-q^2-r^2-2qr+pq+pr+2pqr+3pr^2\\
&+2qr^2+2q^2r-p^2r-3pqr^2-q^2r^2-2p^2r^2)(1-p-q)(1-r)]\mu(LWD)\\
&-r^2(1-r)^2(1-p-q)^2\mu(WLD)-r^2(1-r)^2(1-p-q)^2\mu(DLD)\\
&+\{p(1-p)-(q+r-pr-qr)(1-q-r+pr+qr)\}(1-p-q)(1-r)\mu(WWDW)\\
&+p(1-p-q)^2(1-r)^2\mu(WWWD)+A_6,
\end{aligned}
\tag{4.8.45}
$$

where A_6 is the sum of all those terms in the expansion of $\widehat{G}_{p,q,r}\mu(LWD)$ that do not arise as contributions from any of the events considered so far, i.e. $\{\eta(2)=L,\eta(3)=D\}$, $\{\eta(1)=\eta(2)=D\}$, $\{\eta(0)=L,\eta(1)=W,\eta(2)=D\}$, $\{\eta(1)=L,\eta(2)=D,\eta(3)=L\}$, $\{\eta(1)=W,\eta(2)=D,\eta(3)=L\}$, $\{\eta(1)=L,\eta(2)=D,\eta(3)=W\}$, $\{\eta(1)=W,\eta(2)=D,\eta(3)=W\}$ and $\{\eta(2)=W,\eta(3)=D\}$. It is helpful for the reader to note here that A_6 is made up of contributions to $\widehat{G}_{p,q,r}\mu(LWD)$ from the events $\{\eta(1)=L,\eta(2)=D,\eta(3)=D\}$ and $\{\eta(0)\neq L,\eta(1)=W,\eta(2)=\eta(3)=D\}$, where η follows the distribution μ.

7. The contribution to $\widehat{G}_{p,q,r}\mu(LWD)$ of the event $\{\eta(1)=L,\eta(2)=\eta(3)=D\}$ is given by:

$$
\begin{aligned}
&\{q+(1-p-q)r\}\{p+(1-p-q)(1-r)\}(1-p-q)(1-r)(1+r)\mu(WLDD)\\
&+\{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r)\}(1-p-q)(1-r)(1+r)\mu(DLDD)\\
&+\{q+(1-p-q)r^2\}\{p+(1-p-q)(1-r)\}(1-p-q)(1-r)(1+r)\mu(LLDD)\\
=&(q+r-pr-qr)(1-q-r+pr+qr)(1-p-q)(1-r)(1+r)\mu(LDD)\\
&-r\{p+(1-p-q)(1-r)\}(1-p-q)^2(1-r)^2(1+r)\{\mu(DLDD)+\mu(LLDD)\}.
\end{aligned}
\tag{4.8.46}
$$

The idea, now, is to split the term involving $\mu(LDD)$ in (4.8.46) into two suitable parts, one of which is to be negated by the existing term involving $\mu(DD)$ in (4.8.45), and the other to be negated by the existing term involving $\mu(LD)$ in (4.8.45). We decompose the term involving $\mu(LDD)$ in (4.8.46) into the following two parts:

$$(q+r-pr-qr)(1-q-r+pr+qr)(1-p-q)(1-r)(1+r)\mu(LDD)$$

$$= q(1-q-r+pr+qr)(1-p-q)(1-r)(1+r)\mu(LDD)$$
$$+ r(1-q-r+pr+qr)(1-p-q)^2(1-r)(1+r)\mu(LDD). \qquad (4.8.47)$$

Before proceeding further, we note that the only terms in the expansion of $\widehat{G}_{p,q,r}\mu(LWD)$ that we have not yet taken into account are those coming from the event $\{\eta(0) \in \{W,D\}, \eta(1) = W, \eta(2) = \eta(3) = D\}$, and the sum of these contributions equals

$$\begin{aligned}
&(1-p)p(1-p-q)(1-r)(1+r)\mu(WWDD) \\
&+ \{q+(1-p-q)r\}p(1-p-q)(1-r)(1+r)\mu(DWDD) \\
\leqslant\ & q(1-q-r+pr+qr)(1-p-q)(1-r)(1+r)\mu(WDD) \\
&+ \{p(1-p)-q(1-q-r+pr+qr)\}(1-p-q)(1-r)(1+r)\mu(WWDD) \\
&+ (pr+qr-q)(1-p-q)^2(1-r)(1+r)\mu(DWDD),
\end{aligned} \qquad (4.8.48)$$

where the inequality is obtained since we ignore the term involving $\mu(LWDD)$ (since its coefficient is non-positive). At this point, we observe, from (4.8.26), that:

$$\begin{aligned}
&\{(1-p-q)(1-r)(1+r)\}^{-1} \geqslant 1+p+q+r^2+p^2+q^2+2pq+p^2r^4-2p^2r^2+2pqr^4 \\
&-4pqr^2-2pr^4+pr^2+q^2r^4-2q^2r^2-2qr^4+qr^2+r^4 \\
=\ & (1+p+q+pq+pr-q^2-qr-p^2r+q^2r)+(p-r)^2+2q^2+pq(1-4r^2+2r^4) \\
&+qr(1-q+r-2qr-2r^3+qr^3)+pr(1+p+r-2pr-2r^3+pr^3)+r^4 \\
\geqslant\ & (1+p+q+pq+pr-q^2-qr-p^2r+q^2r) \text{ when } p,q,r \text{ are sufficiently small.}
\end{aligned} \qquad (4.8.49)$$

We now combine the first of the two terms in (4.8.47), the term involving $\mu(WDD)$ in (4.8.48), and the term involving $\mu(DD)$ in (4.8.45), using the inequality $\mu(WDD) + \mu(LDD) \leqslant \mu(DD)$:

$$\begin{aligned}
&-\{1-(1+p+pq+pr-p^2r-pqr)(1-p-q)(1-r)(1+r)\}\mu(DD) \\
&+q(1-q-r+pr+qr)(1-p-q)(1-r)(1+r)\mu(LDD) \\
&+q(1-q-r+pr+qr)(1-p-q)(1-r)(1+r)\mu(WDD) \\
&\leqslant [-1+(1+p+q+pq+pr-q^2-qr-p^2r+q^2r)(1-p-q)(1-r)(1+r)]\mu(DD),
\end{aligned} \qquad (4.8.50)$$

and by (4.8.49), we conclude that the coefficient of $\mu(DD)$ in (4.8.50) is non-positive.

We combine the second term appearing in (4.8.47) with the term involving $\mu(LD)$ in (4.8.45), and use the inequality $\mu(LDD) \leqslant \mu(LD)$, to obtain:

$$
\begin{aligned}
&-[1-\{1-p-q+4r-qr(2-p-2r)-r^2(2-3p)-r^3(1-p-q)-p^2r^2-pqr^2\}\\
&(1-p-q)(1-r)]\mu(LD)+r(1-q-r+pr+qr)(1-p-q)^2(1-r)(1+r)\mu(LDD)\\
=\;&[-1+\{1-p-q+4r-2qr+pqr+2qr^2-2r^2+3pr^2-r^3(1-p-q)-p^2r^2-pqr^2\}\\
&(1-p-q)(1-r)]\mu(LD)+r(1-q-r+pr+qr)(1-p-q)^2(1-r)(1+r)[\mu(LD)\\
&-\mu(LDL)-\mu(LDW)]\\
\leqslant\;&[-1+\{1-p-q+4r-2qr+pqr+2qr^2-2r^2+3pr^2+r(1-q-r+pr+qr)(1-p-q)(1+r)\}\\
&(1-p-q)(1-r)]\mu(LD)-r(1-q-r+pr+qr)(1-p-q)^2(1-r)(1+r)[\mu(LDL)+\mu(LDW)]\\
&\text{(when } p,q,r \text{ are sufficiently small)}\\
=\;&[-1+(1-p-q+5r-2r^2-pr-4qr-r^3+2qr^2+4pr^2+2pqr+q^2r-p^2r^3-p^2r^2-2pqr^3\\
&-pqr^2+2pr^3-q^2r^3+2qr^3)(1-p-q)(1-r)]\mu(LD)-r(1-q-r+pr+qr)(1-p-q)^2\\
&(1-r)(1+r)[\mu(LDL)+\mu(LDW)]\\
\leqslant\;&[-1+\{1-p-q+5r-2r^2-pr-4qr+2qr^2+4pr^2+2pqr+q^2r-r^3(1-2p-2q)\}(1-p-q)\\
&(1-r)]\mu(LD)-r(1-q-r+pr+qr)(1-p-q)^2(1-r)(1+r)[\mu(LDL)+\mu(LDW)]\\
\leqslant\;&[4r-2p-2q+p^2+2pq+q^2-7r^2-4pr-7qr+5pqr+12pr^2+4q^2r+13qr^2+2r^3-2p^2qr(1-r)\\
&-p^2r^2(5-4r)-3pq^2r(1-r)-pqr^2(13-6r)-6pr^3-q^3r(1-r)-q^2r^2(7-2r)-4qr^3]\mu(LD)\\
&-r(1-q-r+pr+qr)(1-p-q)^2(1-r)(1+r)[\mu(LDL)+\mu(LDW)]\\
&\text{(when } p,q \text{ are sufficiently small -- in particular, when } p+q\leqslant 1/2)\\
\leqslant\;&[4r-2(p+q)+(p+q)^2-6r^2-3pr-6qr-r^2(1-2r-6p-4q)-pr(1-5q-6r)-qr\\
&(1-4q-9r)]\mu(LD)-r(1-q-r+pr+qr)(1-p-q)^2(1-r)(1+r)[\mu(LDL)+\mu(LDW)]\\
&\text{(when } p,q,r \text{ are sufficiently small)}\\
\leqslant\;&[4r-2(p+q)+(p+q)^2-6r^2-3pr-6qr]\mu(LD)-r(1-q-r+pr+qr)(1-p-q)^2(1-r)(1+r)\\
&[\mu(LDL)+\mu(LDW)] \quad \text{(when } p,q,r \text{ are sufficiently small)}, \qquad (4.8.51)
\end{aligned}
$$

and the coefficient of $\mu(LD)$ in (4.8.51) is non-positive whenever (4.2.1) holds. Combining the terms involving $\mu(LDL)$ and $\mu(LDW)$ of (4.8.51) with those of (4.8.45), we obtain:

$$
\begin{aligned}
&-r(1-p-q)(1-r)\{(1-p-q)(1-r)-p(1-p)\}\mu(LDW)-(1-p-q)(1-r)[(1-r)(1-p-q)\\
&\{1-qr-r^2(1-p-q)\}-pqr-pr^2(1-p-q)]\mu(LDL)-r(1-q-r+pr+qr)(1-p-q)^2(1-r)\\
&(1+r)[\mu(LDL)+\mu(LDW)]\\
=\;&(1-p-q)(1-r)\{-2r+r^2+3pr+3qr+r^3-p^2r(1-r-r^2)-pqr(1-r-2r^2)-2pr^3-2pr^2
\end{aligned}
$$

$$
\begin{aligned}
&-q^2r(1-r^2)-2qr^3-qr^2\}\mu(LDW)+(1-p-q)(1-r)\{-1+p+q+r^2+2qr-2q^2r(1-r)\\
&-3qr^2(1-p)-pr^2(2-p)-pqr\}\mu(LDL)\\
\leqslant\ &(1-p-q)(1-r)(-2r+r^2+3pr+3qr+r^3)\mu(LDW)+(1-p-q)(1-r)(-1+p+q+r^2\\
&+2qr)\mu(LDL),
\end{aligned}
\tag{4.8.52}
$$

for all p, q and r sufficiently small. Moroever, the coefficient of each of $\mu(LDW)$ and $\mu(LDL)$ in (4.8.52) is non-positive when p, q and r are sufficiently small.

Incorporating (4.8.46), (4.8.48), (4.8.50), (4.8.51) and (4.8.52) into (4.8.45), we obtain:

$$
\begin{aligned}
w_0(\widehat{G}_{p,q,r}\mu) \leqslant\ & w_0(\mu)-[2-(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2\\
&+pqr+2pr^2-p^2r-2pqr^2-p^2r^2-q^2r^2\}]\mu(WD)-\{1-(1+p+q+pq+pr\\
&-q^2-qr-p^2r+q^2r)(1-p-q)(1-r)(1+r)\}\mu(DD)+\{4r-2(p+q)+(p+q)^2\\
&-6r^2-3r(p+2q)\}\mu(LD)+(1-p-q)(1-r)(-2r+r^2+3pr+3qr+r^3)\mu(LDW)\\
&+(1-p-q)(1-r)(-1+p+q+r^2+2qr)\mu(LDL)-[1-(1-p-q^2-r^2-2qr\\
&+pq+pr+2pqr+3pr^2+2qr^2+2q^2r-p^2r-3pqr^2-q^2r^2-2p^2r^2)(1-p-q)(1-r)]\\
&\mu(LWD)-r^2(1-r)^2(1-p-q)^2\mu(WLD)-r^2(1-r)^2(1-p-q)^2\mu(DLD)\\
&+\{p(1-p)-(q+r-pr-qr)(1-q-r+pr+qr)\}(1-p-q)(1-r)\mu(WWDW)\\
&+p(1-p-q)^2(1-r)^2\mu(WWWD)+\{p(1-p)-q(1-q-r+pr+qr)\}(1-p-q)\\
&(1-r)(1+r)\mu(WWDD)+(pr+qr-q)(1-p-q)^2(1-r)(1+r)\mu(DWDD).
\end{aligned}
\tag{4.8.53}
$$

We recall, for the reader's convenience, that the ultimate goal for us is to come up with a weight function inequality that satisfies (4.8.3), and although the weight function inequality in (4.8.53) is of the same form as (4.8.2), it does not satisfy (4.8.3) because of the last few terms. In particular, we note that:

1. the coefficient $\mu(WWWD)$ in (4.8.53) is non-negative;
2. the coefficient of $\mu(DWDD)$ in (4.8.53) may or may not be non-negative;
3. the coefficient of $\mu(WWDW)$ in (4.8.53) is positive whenever $p(1-p) > (q+r-pr-qr)(1-q-r+pr+qr)$, which is equivalent to $p > q+r(1-p-q)$ when p, q and r are sufficiently small (in particular, as long as each of p and $q+r(1-p-q)$ is bounded above by $1/2$, since the function $f(x) = x(1-x)$ is strictly increasing for $x \in [0,1/2)$ and strictly decreasing for $x \in (1/2,1]$), and this happens when (4.2.5) is true;

4. and the coefficient of $\mu(WWDD)$ in (4.8.53) is positive whenever $p(1-p) > q(1-q-r+pr+qr)$, which happens when one of (4.2.3), (4.2.4) and (4.2.5) is true.

At this point, our first attempt is to try to negate, to as large an extent as possible, the terms involving $\mu(WWWD)$ and $\mu(DWDD)$ in (4.8.53), using the term involving $\mu(WD)$ in (4.8.53). Here, we make use of the inequality (via reflection-invariance): $\mu(WWWD)+\mu(DWDD) = \mu(WWWD)+\mu(DDWD) \leqslant \mu(WD)$. First, we write $\mu(WD) = \mu(WWWD)+[\mu(WD)-\mu(WWWD)]$, and combine the term involving $\mu(WWWD)$ thus obtained from the term involving $\mu(WD)$ in (4.8.53), with the existing term involving $\mu(WWWD)$ in (4.8.53), to get:

$$
\begin{aligned}
&-[2-(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2+pqr+2pr^2-p^2r\\
&\quad -2pqr^2-p^2r^2-q^2r^2\}]\mu(WWWD)+p(1-p-q)^2(1-r)^2\mu(WWWD)\\
&=[-2(1-p-q)(1-r)\{(1-p-q)(1-r)\}^{-1}+(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr\\
&\quad +pq+2q^2r+2qr^2+pqr+2pr^2-p^2r-2pqr^2-p^2r^2-q^2r^2\}+p(1-p-q)^2(1-r)^2]\mu(WWWD)\\
&\leqslant (1-p-q)(1-r)[-2\{1+p+q+r+p^2+q^2+r^2+2pq+pr+qr-4pqr-2p^2r-2pr^2-2q^2r\\
&\quad -2qr^2+p^2r^2+q^2r^2+2pqr^2\}+\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2+pqr+2pr^2\\
&\quad -p^2r-2pqr^2-p^2r^2-q^2r^2\}+p(1-p-q)(1-r)]\mu(WWWD)\\
&\quad \text{(making use of the inequality in (4.8.31))}\\
&\leqslant (1-p-q)(1-r)[p-q-r-3p^2-3q^2-3r^2-4pq-3pr-5qr+4p^2r+6pr^2+6q^2r+6qr^2\\
&\quad +10pqr]\mu(WWWD). \qquad (4.8.54)
\end{aligned}
$$

The coefficient of $\mu(WWWD)$ in (4.8.54) may or may not be non-positive. Before proceeding further, we note, from (4.8.31), that

$$
\begin{aligned}
&2\{(1-p-q)(1-r)\}^{-1} \geqslant 2+2p+2q+2r+2p^2+2q^2+2r^2+4pq+2pr+2qr-8pqr-4p^2r-4pr^2\\
&\quad -4q^2r-4qr^2+2p^2r^2+2q^2r^2+4pqr^2\\
&=(2+2p+r-r^2-3qr+2pq+pr+3pr^2-2p^2r+2q^2r+3qr^2-2p^2r^2-2q^2r^2-4pqr^2)\\
&\quad +2q+r+2p^2+2q^2+3r^2+5qr+2pq+pr-7pr^2-2p^2r-6q^2r-7qr^2-8pqr+4p^2r^2+4q^2r^2\\
&\quad +8pqr^2\\
&=(2+2p+r-r^2-3qr+2pq+pr+3pr^2-2p^2r+2q^2r+3qr^2-2p^2r^2-2q^2r^2-4pqr^2)+\\
&\quad 2q+r+2p^2+2q^2+3r^2+qr(5-6q-7r+3qr)+2pq(1-4r+4r^2)+pr(1-7r-2p+4pr)\\
&\geqslant 2+2p+r-r^2-3qr+2pq+pr+3pr^2-2p^2r+2q^2r+3qr^2-2p^2r^2-2q^2r^2-4pqr^2\\
&\qquad (4.8.55)
\end{aligned}
$$

when p, q and r are sufficiently small. Next, we write $[\mu(WD)-\mu(WWWD)] = \mu(DWDD) + [\mu(WD)-\mu(WWWD)-\mu(DWDD)]$, and we combine the term involving $\mu(DWDD)$ thus obtained from the term involving $\mu(WD)$ in (4.8.53), with the existing term involving $\mu(DWDD)$ in (4.8.53), to get:

$$\begin{aligned}
& -[2-(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2+pqr+2pr^2-p^2r \\
& -2pqr^2-p^2r^2-q^2r^2\}]\mu(DWDD)+(pr+qr-q)(1-p-q)^2(1-r)(1+r)\mu(DWDD) \\
= {} & [-2+\{2+2p+r-r^2-3qr+2pq+pr+3pr^2-2p^2r+2q^2r+3qr^2-2p^2r^2-2q^2r^2-4pqr^2\} \\
& (1-p-q)(1-r)]\mu(DWDD), \qquad (4.8.56)
\end{aligned}$$

and the coefficient of $\mu(DWDD)$ in (4.8.56) is non-positive because of the inequality established in (4.8.55).

Incorporating (4.8.54) and (4.8.56) into (4.8.53), and eliminating the term involving $\mu(DWDD)$ since its coefficient in (4.8.56) has already been proven above to be non-positive, via the inequality in (4.8.55), when p, q and r are sufficiently small, we obtain:

$$\begin{aligned}
w_0(\widehat{G}_{p,q,r}\mu) \leqslant {} & w_0(\mu)-[2-(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2 \\
& +pqr+2pr^2-p^2r-2pqr^2-p^2r^2-q^2r^2\}]\{\mu(WD)-\mu(WWWD)-\mu(DWDD)\} \\
& -\{1-(1+p+q+pq+pr-q^2-qr-p^2r+q^2r)(1-p-q)(1-r)(1+r)\}\mu(DD) \\
& +\{4r-2(p+q)+(p+q)^2-6r^2-3r(p+2q)\}\mu(LD)+(1-p-q)(1-r)(-2r+r^2 \\
& +3pr+3qr+r^3)\mu(LDW)+(1-p-q)(1-r)(-1+p+q+r^2+2qr)\mu(LDL) \\
& -[1-(1-p-q^2-r^2-2qr+pq+pr+2pqr+3pr^2+2qr^2+2q^2r-p^2r-3pqr^2 \\
& -q^2r^2-2p^2r^2)(1-p-q)(1-r)]\mu(LWD)-r^2(1-r)^2(1-p-q)^2\mu(WLD) \\
& -r^2(1-r)^2(1-p-q)^2\mu(DLD)+\{p(1-p)-(q+r-pr-qr)(1-q-r+pr+qr)\} \\
& (1-p-q)(1-r)\mu(WWDW)+(1-p-q)(1-r)\{p-q-r-3p^2-3q^2-3r^2-4pq \\
& -3pr-5qr+4p^2r+6pr^2+6q^2r+6qr^2+10pqr\}\mu(WWWD)+\{p(1-p) \\
& -q(1-q-r+pr+qr)\}(1-p-q)(1-r)(1+r)\mu(WWDD) \\
= {} & w_0(\mu)-[2-(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2+pqr \\
& +2pr^2-p^2r-2pqr^2-p^2r^2-q^2r^2\}]\{\mu(WD)-\mu(WWWD)-\mu(DWDD)-\mu(LWD)\} \\
& -\{1-(1+p+q+pq+pr-q^2-qr-p^2r+q^2r)(1-p-q)(1-r)(1+r)\}\mu(DD) \\
& +\{4r-2(p+q)+(p+q)^2-6r^2-3r(p+2q)\}\mu(LD)+(1-p-q)(1-r)(-2r+r^2 \\
& +3pr+3qr+r^3)\mu(LDW)+(1-p-q)(1-r)(-1+p+q+r^2+2qr)\mu(LDL)-
\end{aligned}$$

$$[3-(3+p+q+r-2q^2-2r^2-5qr+2pq+pr+3pqr+4q^2r+4qr^2+5pr^2-2p^2r$$
$$-5pqr^2-3p^2r^2-2q^2r^2)(1-p-q)(1-r)]\mu(LWD)-r^2(1-r)^2(1-p-q)^2\mu(WLD)$$
$$-r^2(1-r)^2(1-p-q)^2\mu(DLD)+\{p(1-p)-(q+r-pr-qr)(1-q-r+pr+qr)\}$$
$$(1-p-q)(1-r)\mu(WWDW)+(1-p-q)(1-r)\{p-q-r-3p^2-3q^2-3r^2-4pq$$
$$-3pr-5qr+4p^2r+6pr^2+6q^2r+6qr^2+10pqr\}\mu(WWWD)+\{p(1-p)-q(1-q$$
$$-r+pr+qr)\}(1-p-q)(1-r)(1+r)\mu(WWDD). \tag{4.8.57}$$

The important change that takes place in the final expression of (4.8.57) is the combination of the term involving $\mu(LWD)$ that already exists in the previous step with the term involving $\mu(LWD)$ that comes out of the term involving $\{\mu(WD)-\mu(WWWD)-\mu(DWDD)\}$ (note, by reflection-invariance, that $\mu(WWWD)+\mu(DWDD)+\mu(LWD)=\mu(WWWD)+\mu(DDWD)+\mu(LWD)\leqslant \mu(WWD)+\mu(DWD)+\mu(LWD)=\mu(WD)$).

4.8.3.1 When (4.2.2) of Theorem 4.2.1 is true

Here, we assume that (p,q,r) satisfies (4.2.1) and (4.2.2). The latter allows us to deduce the following upper bound on the coefficient of $\mu(WWWD)$ in (4.8.54) (which is also the coefficient of $\mu(WWWD)$ in (4.8.57)):

$$\begin{aligned}
&(1-p-q)(1-r)\{p-q-r-3p^2-3q^2-3r^2-4pq-3pr-5qr+4p^2r+6pr^2+6q^2r+6qr^2+10pqr\}\\
=&(1-p-q)(1-r)\{p(1-p)-q-r-2p^2-3q^2-3r^2-4pq-3pr-5qr+4p^2r+6pr^2+6q^2r+6qr^2\\
&+10pqr\}\\
\leqslant&(1-p-q)(1-r)\{q-q^2-qr(1-p-q)-q-r-2p^2-3q^2-3r^2-4pq-3pr-5qr+4p^2r+6pr^2\\
&+6q^2r+6qr^2+10pqr\}\\
=&(1-p-q)(1-r)\{-r-2p^2-4q^2-3r^2-pq(4-11r)-pr(3-4p-6r)-qr(6-6r-7q)\}
\end{aligned}$$

which is non-positive for all p, q and r sufficiently small. Thus, the coefficient of $\mu(WWWD)$ in (4.8.54) is non-positive when p, q and r are sufficiently small and (p,q,r) belongs to the regime described in (4.2.2). Since (4.2.2) holds, the coefficient of $\mu(WWDD)$ in (4.8.57) is non-positive, and furthermore, we have $p(1-p)\leqslant q\{1-q-r(1-p-q)\}\leqslant\{q+r(1-p-q)\}\{1-q-r(1-p-q)\}$, proving that the coefficient of $\mu(WWDW)$ in (4.8.57) is non-positive as well. Consequently, (4.8.57) is of the form given by (4.8.2) and satisfies (4.8.3), thus fulfilling the objective we set out to achieve when (p,q,r) satisfies (4.2.1) and (4.2.2). Therefore, the construction of our desired weight function comes to an end at this step when (4.2.2) holds.

Note that, since the coefficient of each of $\mu(LD)$, $\mu(LDW)$, $\mu(LDL)$, $\mu(LWD)$, $\mu(WWDW)$, $\mu(WWDD)$ and $\mu(WWWD)$ in (4.8.57) is non-positive in this case, we can remove these terms from the right side of (4.8.57), and the resulting inequality would be exactly what appears in (4.8.10) (once we replace the notation w_0 by w).

4.8.3.2 When either (4.2.3) or (4.2.4) holds

The reason for considering these two possibilities together is that the way we proceed, to a large extent, is the same for both of them (in fact, much of this approach is also common to the scenario where (p,q,r) satisfies (4.2.5)).

The very first observation to make is that the coefficient of $\mu(WWDW)$ in (4.8.57) is non-positive when either of (4.2.3) and (4.2.4) is true. Next, recall, from the discussion in (3), that the inequality $p(1-p) \leqslant \{q+r(1-p-q)\}\{1-q-r(1-p-q)\}$, which is a part of both (4.2.3) and (4.2.4), becomes equivalent to $p \leqslant q+r(1-p-q)$ when $p \leqslant 1/2$ and $q+r(1-p-q) \leqslant 1/2$ (in other words, when p, q and r are sufficiently small). Applying the latter inequality, the coefficient of $\mu(WWWD)$ in (4.8.57) can be bounded above as follows:

$$\begin{aligned}
&(1-p-q)(1-r)\{p-q-r-3p^2-3q^2-3r^2-4pq-3pr-5qr+4p^2r+6pr^2+6q^2r+6qr^2+10pqr\}\\
&\leqslant (1-p-q)(1-r)\{q+r(1-p-q)-q-r-3p^2-3q^2-3r^2-4pq-3pr-5qr+4p^2r+6pr^2+6q^2r\\
&\quad +6qr^2+10pqr\}\\
&= (1-p-q)(1-r)\{-3p^2-3q^2-3r^2-4pq-4pr-6qr+4p^2r+6pr^2+6q^2r+6qr^2+10pqr\}\\
&= (1-p-q)(1-r)\{-3p^2-3q^2-3r^2-2pq(2-5r)-2pr(2-2p-3r)-6qr(1-q-r)\},
\end{aligned} \tag{4.8.58}$$

which is non-positive for all p, q and r sufficiently small. This shows us that the coefficient of $\mu(WWWD)$ in (4.8.57) is non-positive in this case. Only the coefficient of $\mu(WWDD)$ in (4.8.57) remains non-negative.

This is where the need for an adjustment to our weight function, along the lines of the idea outlined in §4.8.1.1, arises. There are quite a few terms with non-positive coefficients available to us on the right side of (4.8.57), but we should take into account the following aspects while deciding which of them to use for negating the existing (possibly) non-negative terms on the right side of (4.8.57):

1. which of the cylinder sets (namely, $(D,D)_{0,1}$, $(L,D)_{0,1}$, $(L,D,W)_{0,1,2}$ etc.) appearing on the right side of (4.8.57), with non-positive coefficients, can boast a significant contribution from $\mu(WWD)$, $\mu(WWWD)$, $\mu(WWDW)$ and $\mu(WWDD)$ when we consider their probabil-

ities under the pushforward measure (i.e. when we consider the expansion of $\widehat{G}_{p,q,r}\mu(DD)$, $\widehat{G}_{p,q,r}\mu(LD)$, $\widehat{G}_{p,q,r}\mu(LDW)$ etc.);

2. the order of magnitude of the coefficient of each of these terms assuming p, q and r to be sufficiently small (for instance, the coefficient of $\mu(LDW)$ in (4.8.57) is of the order of r, that of $\mu(LDL)$ is of the order of 1, and so on).

To this end, we note the following:

1. When we consider $\widehat{G}_{p,q,r}\mu(LD)$, which is the same as finding the probability of the event $\{\widehat{G}_{p,q,r}\eta(0) = L, \widehat{G}_{p,q,r}\eta(1) = D\}$ with η being a random configuration following the law μ, the contribution of the event $\{\eta(0) = \eta(1) = W, \eta(2) = D\}$ equals $(1-p)(1-p-q)(1-r)\mu(WWD)$, so that we can write

$$(1-p)(1-p-q)(1-r)\mu(WWD) \leqslant \widehat{G}_{p,q,r}\mu(LD). \tag{4.8.59}$$

2. When we consider $\widehat{G}_{p,q,r}\mu(LDW)$, which is the same as the probability of the event $\{\widehat{G}_{p,q,r}\eta(0) = L, \widehat{G}_{p,q,r}\eta(1) = D, \widehat{G}_{p,q,r}\eta(2) = W\}$ with η being a random configuration following the law μ, the contribution of the event $\{\eta(0) = \eta(1) = W, \eta(2) = D, \eta(3) \in \{W,D\}\}$ is given by $(1-p)(1-p-q)(1-r)p\{\mu(WWDW)+\mu(WWDD)\}$, so that we can write

$$(1-p)(1-p-q)(1-r)p\{\mu(WWDW)+\mu(WWDD)\} \leqslant \widehat{G}_{p,q,r}\mu(LDW). \tag{4.8.60}$$

3. When we consider $\widehat{G}_{p,q,r}\mu(LDL)$, which is the same as the probability of the event $\{\widehat{G}_{p,q,r}\eta(0) = L, \widehat{G}_{p,q,r}\eta(1) = D, \widehat{G}_{p,q,r}\eta(2) = L\}$ with η being a random configuration following the law μ, the contribution of the event $\{\eta(0) = \eta(1) = W, \eta(2) = D, \eta(3) \in \{W,D\}\}$ is given by $(1-p)(1-p-q)(1-r)\{q+(1-p-q)r\}\mu(WWDW) + (1-p)(1-p-q)(1-r)\{q+(1-p-q)r^2\}\mu(WWDD)$, so that we can write

$$\begin{aligned}&(1-p)(1-p-q)(1-r)\{q+(1-p-q)r\}\mu(WWDW) + (1-p)(1-p-q)(1-r)\\ &\{q+(1-p-q)r^2\}\mu(WWDD) \leqslant \widehat{G}_{p,q,r}\mu(LDL).\end{aligned} \tag{4.8.61}$$

4. Finally, when we consider $\widehat{G}_{p,q,r}\mu(LWD)$, which is the same as the probability of the event $\{\widehat{G}_{p,q,r}\eta(0) = L, \widehat{G}_{p,q,r}\eta(1) = W, \widehat{G}_{p,q,r}\eta(2) = D\}$ with η being a random configuration following the law μ, the contributions of the event $\{\eta(0) = \eta(1) = W, \eta(2) =$

$D, \eta(3) \in \{W, D\}\}$ and $\{\eta(0) = \eta(1) = \eta(2) = W, \eta(3) = D\}$ equal $(1-p)p(1-p-q)(1-r)\mu(WWDW) + (1-p)p(1-p-q)(1-r)(1+r)\mu(WWDD) + (1-p)p(1-p-q)(1-r)\mu(WWWD)$, so that we can write

$$\begin{aligned}&(1-p)p(1-p-q)(1-r)\mu(WWDW) + (1-p)p(1-p-q)(1-r)(1+r)\mu(WWDD)\\&+(1-p)p(1-p-q)(1-r)\mu(WWWD) \leqslant \widehat{G}_{p,q,r}\mu(LWD).\end{aligned} \tag{4.8.62}$$

We now define the adjusted / updated weight function, w_1, to be exactly as defined in (4.8.9). For the reader's convenience, we write here its rather long and involved expression, recalling w_0 from (4.8.17):

$$\begin{aligned}w_1(\mu) = w_0(\mu) &+ \{4r - 2(p+q) + (p+q)^2 - 6r^2 - 3r(p+2q)\}\mu(LD) + (1-p-q)(1-r)(-2r + r^2 + 3pr\\&+3qr + r^3)\mu(LDW) + (1-p-q)(1-r)(-1+p+q+r^2+2qr)\mu(LDL) - [3-(3+p+q+r\\&-2q^2 - 2r^2 - 5qr + 2pq + pr + 3pqr + 4q^2r + 4qr^2 + 5pr^2 - 2p^2r - 5pqr^2 - 3p^2r^2 - 2q^2r^2)\\&(1-p-q)(1-r)]\mu(LWD).\end{aligned} \tag{4.8.63}$$

Following the argument outlined in §4.8.1.1, the updated weight function in (4.8.63) transforms the current weight function inequality, in (4.8.57), as follows (the same way as we deduce (4.8.6)):

$$\begin{aligned}w_1(\widehat{G}_{p,q,r}\mu) \leqslant w_1(\mu) &- [2-(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2+pqr+\\&2pr^2 - p^2r - 2pqr^2 - p^2r^2 - q^2r^2\}]\{\mu(WD) - \mu(WWWD) - \mu(DWDD) - \mu(LWD)\}\\&-\{1-(1+p+q+pq+pr-q^2-qr-p^2r+q^2r)(1-p-q)(1-r)(1+r)\}\mu(DD)\\&-r^2(1-r)^2(1-p-q)^2\mu(WLD) - r^2(1-r)^2(1-p-q)^2\mu(DLD) + \{p(1-p)-\\&(q+r-pr-qr)(1-q-r+pr+qr)\}(1-p-q)(1-r)\mu(WWDW) + (1-p-q)(1-r)\\&\{p-q-r-3p^2-3q^2-3r^2-4pq-3pr-5qr+4p^2r+6pr^2+6q^2r+6qr^2+10pqr\}\\&\mu(WWWD) + \{p(1-p) - q(1-q-r+pr+qr)\}(1-p-q)(1-r)(1+r)\mu(WWDD)\\&+\{4r-2(p+q)+(p+q)^2-6r^2-3r(p+2q)\}\widehat{G}_{p,q,r}\mu(LD) + (1-p-q)(1-r)(-2r+\\&r^2+3pr+3qr+r^3)\widehat{G}_{p,q,r}\mu(LDW) + (1-p-q)(1-r)(-1+p+q+r^2+2qr)\\&\widehat{G}_{p,q,r}\mu(LDL) - [3-(3+p+q+r-2q^2-2r^2-5qr+2pq+pr+3pqr+4q^2r+4qr^2\\&+5pr^2-2p^2r-5pqr^2-3p^2r^2-2q^2r^2)(1-p-q)(1-r)]\widehat{G}_{p,q,r}\mu(LWD)\\\leqslant w_1(\mu) &- [2-(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2+pqr+\\&2pr^2 - p^2r - 2pqr^2 - p^2r^2 - q^2r^2\}]\{\mu(WD) - \mu(WWWD) - \mu(DWDD) - \mu(LWD)\}\\&-\{1-(1+p+q+pq+pr-q^2-qr-p^2r+q^2r)(1-p-q)(1-r)(1+r)\}\mu(DD)\end{aligned}$$

$$
\begin{aligned}
&-r^2(1-r)^2(1-p-q)^2\mu(WLD)-r^2(1-r)^2(1-p-q)^2\mu(DLD)+\{p(1-p)-\\
&(q+r-pr-qr)(1-q-r+pr+qr)\}(1-p-q)(1-r)\mu(WWDW)+(1-p-q)(1-r)\\
&\{p-q-r-3p^2-3q^2-3r^2-4pq-3pr-5qr+4p^2r+6pr^2+6q^2r+6qr^2+10pqr\}\\
&\mu(WWWD)+\{p(1-p)-q(1-q-r+pr+qr)\}(1-p-q)(1-r)(1+r)\mu(WWDD)\\
&+\{4r-2(p+q)+(p+q)^2-6r^2-3r(p+2q)\}(1-p)(1-p-q)(1-r)\mu(WWD)+\\
&(1-p-q)(1-r)(-2r+r^2+3pr+3qr+r^3)(1-p)(1-p-q)(1-r)p\{\mu(WWDW)\\
&+\mu(WWDD)\}+(1-p-q)(1-r)(-1+p+q+r^2+2qr)[(1-p)(1-p-q)(1-r)\\
&\{q+(1-p-q)r\}\mu(WWDW)+(1-p)(1-p-q)(1-r)\{q+(1-p-q)r^2\}\mu(WWDD)]\\
&-[3-(3+p+q+r-2q^2-2r^2-5qr+2pq+pr+3pqr+4q^2r+4qr^2+5pr^2-2p^2r\\
&-5pqr^2-3p^2r^2-2q^2r^2)(1-p-q)(1-r)]\{(1-p)p(1-p-q)(1-r)\mu(WWDW)+\\
&(1-p)p(1-p-q)(1-r)(1+r)\mu(WWDD)+(1-p)p(1-p-q)(1-r)\mu(WWWD)\}
\end{aligned}
\tag{4.8.64}
$$

where the inequality in the last step of the derivation of (4.8.64) is obtained by making use of (4.8.59), (4.8.60), (4.8.61) and (4.8.62).

Let us write down all those terms of (4.8.64) that we now need to combine suitably:

1. $\{p(1-p)-(q+r-pr-qr)(1-q-r+pr+qr)\}(1-p-q)(1-r)\mu(WWDW)$,

2. $(1-p-q)(1-r)\{p-q-r-3p^2-3q^2-3r^2-4pq-3pr-5qr+4p^2r+6pr^2+6q^2r+6qr^2+10pqr\}\mu(WWWD)$,

3. $\{p(1-p)-q(1-q-r+pr+qr)\}(1-p-q)(1-r)(1+r)\mu(WWDD)$,

4. $\{4r-2(p+q)+(p+q)^2-6r^2-3r(p+2q)\}(1-p)(1-p-q)(1-r)\mu(WWD)$,

5. $(1-p-q)(1-r)(-2r+r^2+3pr+3qr+r^3)(1-p)(1-p-q)(1-r)p\{\mu(WWDW)+\mu(WWDD)\}$,

6. $(1-p-q)(1-r)(-1+p+q+r^2+2qr)[(1-p)(1-p-q)(1-r)\{q+(1-p-q)r\}\mu(WWDW)+(1-p)(1-p-q)(1-r)\{q+(1-p-q)r^2\}\mu(WWDD)]$,

7. $-[3-(3+p+q+r-2q^2-2r^2-5qr+2pq+pr+3pqr+4q^2r+4qr^2+5pr^2-2p^2r-5pqr^2-3p^2r^2-2q^2r^2)(1-p-q)(1-r)]\{(1-p)p(1-p-q)(1-r)\mu(WWDW)+(1-p)p(1-p-q)(1-r)(1+r)\mu(WWDD)+(1-p)p(1-p-q)(1-r)\mu(WWWD)\}$.

Recall that when (p,q,r) satisfies the constraints of (4.2.3) or (4.2.4), it is only the coefficient of $\mu(WWDD)$ that is non-positive on the right side of (4.8.57) (other than $w_0(\mu)$ itself). Consequently, for this case, we need only focus on the terms involving $\mu(WWDD)$ in the above-mentioned list (appearing in (3), (5), (6) and (7)), and the term involving $\mu(WWD)$ (appearing in (4)). Combining the terms involving $\mu(WWDD)$ from (3), (5), (6) and (7), we get the following coefficient of $\mu(WWDD)$ (or rather, an upper bound on it), as long as p, q and r are sufficiently small:

$$\begin{aligned}
&(1-p-q)(1-r)[\{p(1-p)-q(1-q-r+pr+qr)\}(1+r)+(1-p-q)(1-r)(-2r+r^2+3pr+3qr\\
&\quad+r^3)(1-p)p+(-1+p+q+r^2+2qr)(1-p)(1-p-q)(1-r)\{q+(1-p-q)r^2\}-\{3-(3+p+q\\
&\quad+r-2q^2-2r^2-5qr+2pq+pr+3pqr+4q^2r+4qr^2+5pr^2-2p^2r-5pqr^2-3p^2r^2-2q^2r^2)\\
&\quad(1-p-q)(1-r)\}(1-p)p(1+r)]\\
&=[-p^5r(2+3r-2r^2-3r^3)-p^4qr^2(10+r-8r^2)-11p^4r^4-4p^4r^3-7p^3q^2r^2(1+r-r^2)-20p^3qr^4\\
&\quad-p^3q(1+3r-15r^2-9r^3-2p-pr)-p^3r(11+3r+r^2-13r^3-r^4-8p-6pr)-2p^2q^3r^3(2-r)\\
&\quad-2p^2q^3(1-2r)-p^2q^2r(13-7p-2r-15r^2+9r^3)-p^2qr^3(17-11r-2r^2)-p^2q(1-5q-5r)\\
&\quad-3p^2r^5-4p^2r^4-p^2r^2(1-7r)-p^2(1-8r-p-p^2)-3pq^3r(1+r-r^2)-pq^2r^4(3-r)-8pq^2r^3\\
&\quad-pq^2(7-6r-3q-11r^2)-4pqr^5-pqr^2(8-11r-5r^2)-3pqr-pr^3(5+2r-3r^2)-2q^3r^4\\
&\quad-q^3(1+r-3r^2-r^3)-q^2r^5-q^2r^2(7-5r^2)-2qr^4(2-r)-r^2(1-r-2p-5q-r^2+2qr+r^3)\\
&\quad+\{-2p^2+pq-3pr+p+3q^2+qr-2q\}](1-p-q)(1-r)\\
&\leqslant(p-2q+pq+qr-3pr-2p^2+3q^2)(1-p-q)(1-r).
\end{aligned}\tag{4.8.65}$$

The first scenario to consider is where the final expression in (4.8.65) is non-positive, i.e. the second inequality of (4.2.3) holds. Note that the current weight function inequality, obtained from (4.8.64) by updating the coefficient of $\mu(WWDD)$ to the expression in (4.8.65), and by combining the terms involving $\mu(WWDW)$ and those involving $\mu(WWWD)$ in the final step of (4.8.64) (note that this, in fact, *reduces* the already non-positive coefficients of $\mu(WWDW)$ and $\mu(WWWD)$ further), is given by:

$$\begin{aligned}
w_1(\widehat{G}_{p,q,r}\mu)\leqslant\ & w_1(\mu)-[2-(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2+pqr+\\
&2pr^2-p^2r-2pqr^2-p^2r^2-q^2r^2\}]\{\mu(WD)-\mu(WWWD)-\mu(DWDD)-\mu(LWD)\}\\
&-\{1-(1+p+q+pq+pr-q^2-qr-p^2r+q^2r)(1-p-q)(1-r)(1+r)\}\mu(DD)\\
&-r^2(1-r)^2(1-p-q)^2\mu(WLD)-r^2(1-r)^2(1-p-q)^2\mu(DLD)+(1-p-q)(1-r)\\
&[\{p(1-p)-(q+r-pr-qr)(1-q-r+pr+qr)\}+(1-p-q)(1-r)(-2r+r^2+3pr
\end{aligned}$$

$$
\begin{aligned}
&+3qr+r^3)(1-p)p+(-1+p+q+r^2+2qr)(1-p)(1-p-q)(1-r)\{q+(1-p-q)r\}\\
&-\{3-(3+p+q+r-2q^2-2r^2-5qr+2pq+pr+3pqr+4q^2r+4qr^2+5pr^2-2p^2r\\
&-5pqr^2-3p^2r^2-2q^2r^2)(1-p-q)(1-r)\}(1-p)p]\mu(WWDW)+(1-p-q)(1-r)\\
&[p-q-r-3p^2-3q^2-3r^2-4pq-3pr-5qr+4p^2r+6pr^2+6q^2r+6qr^2+10pqr-\\
&\{3-(3+p+q+r-2q^2-2r^2-5qr+2pq+pr+3pqr+4q^2r+4qr^2+5pr^2-2p^2r\\
&-5pqr^2-3p^2r^2-2q^2r^2)(1-p-q)(1-r)\}(1-p)p]\mu(WWWD)+(1-p-q)(1-r)\\
&(p-2q+pq+qr-3pr-2p^2+3q^2)\mu(WWDD)+\{4r-2(p+q)+(p+q)^2-6r^2\\
&-3r(p+2q)\}(1-p)(1-p-q)(1-r)\mu(WWD). \qquad (4.8.66)
\end{aligned}
$$

The inequality in (4.8.66) is evidently of the form given by (4.8.2), and satisfies (4.8.3). Therefore, our construction of the desired weight function comes to an end here in this case. The final weight function is as given by (4.8.63) (which is the same as (4.8.9)), and the final weight function inequality is (4.8.66). Since the coefficient of each of $\mu(WWDW)$, $\mu(WWDD)$, $\mu(WWWD)$ and $\mu(WWD)$ in (4.8.66) is non-positive when (4.2.1) and (4.2.3) hold, we may remove these terms from (4.8.66), which would yield the exact same weight function inequality as that given by (4.8.10) (once we replace the notation w_1 by simply w).

The second scenario to consider is where the final expression in (4.8.65) is positive, i.e. if the second inequality of (4.2.4) holds instead. In this case, we combine the coefficient of $\mu(WWD)$ in (4) with the expression in (4.8.65) (making use of the inequality $\mu(WWDD) \leqslant \mu(WWD)$ – we can use this since the coefficient of $\mu(WWD)$ in (4) is non-positive due to (4.2.1)), to get the following upper bound on the coefficient of $\mu(WWDD)$:

$$
\begin{aligned}
&(p-2q+pq+qr-3pr-2p^2+3q^2)(1-p-q)(1-r)+\{4r-2(p+q)+(p+q)^2-6r^2-3r(p+2q)\}\\
&(1-p)(1-p-q)(1-r)\\
&=(1-p-q)(1-r)\{-p^3-2p^2q+3p^2r+p^2-pq^2+6pqr+5pq+6pr^2-10pr-p+4q^2-5qr\\
&\quad-4q-6r^2+4r\}\\
&\leqslant(1-p-q)(1-r)\{-pr(1-3p-6q-6r)+p^2+5pq-9pr-p+4q^2-5qr-4q-6r^2+4r\}\\
&\leqslant(1-p-q)(1-r)(-p-4q+4r+p^2+4q^2+5pq-9pr-5qr-6r^2), \qquad (4.8.67)
\end{aligned}
$$

as long as p, q and r are sufficiently small, and this final expression is non-positive due to the third inequality of (4.2.4). Once incorporated into (4.8.66), the updated coefficient of $\mu(WWDD)$ from (4.8.67) transforms the weight function inequality into one that satisfies (4.8.3) (since now, each term on the right side of (4.8.66), apart from $w_0(\mu)$, has a non-positive coefficient). Since the coefficient of each of $\mu(WWDW)$, $\mu(WWDD)$ and $\mu(WWWD)$ in the updated weight function inequality is non-positive, we may remove all of these terms, leaving us with exactly the same weight function inequality as that given by (4.8.10) (once we replace the notation w_1 by simply w).

4.8.3.3 When (p,q,r) satisfies (4.2.5) of Theorem 4.2.1

The hardest case to consider is where (4.2.5) is true. Most of the approach adopted so far also applies to this case, i.e. the weight function is updated as in (4.8.63), the coefficient of $\mu(WWDD)$ is updated to the expression given by (4.8.65), and the weight function inequality is updated to (4.8.66). Note that in the regime given by (4.2.5), there is no guarantee that the coefficient of any of $\mu(WWDW)$, $\mu(WWDD)$ and $\mu(WWWD)$ in (4.8.66) is non-positive.

First, we simplify, by providing suitable upper bounds, the coefficients of $\mu(WWDW)$ and $\mu(WWWD)$ in (4.8.66), since otherwise, the expressions are too cluttered due of the presence of very small terms (third or higher order terms, i.e. terms of the form $p^i q^j r^k$ with $i+j+k \geqslant 3$, make little difference to us as p, q and r are assumed to be sufficiently small). Combining the terms involving $\mu(WWDW)$ from (1), (5), (6) and (7), we get the following upper bound on the coefficient of $\mu(WWDW)$:

$$\begin{aligned}
&(1-p-q)(1-r)[\{p(1-p)-(q+r-pr-qr)(1-q-r+pr+qr)\}+(1-p-q)(1-r)(-2r+r^2+3pr\\
&+3qr+r^3)(1-p)p+(-1+p+q+r^2+2qr)(1-p)(1-p-q)(1-r)\{q+(1-p-q)r\}-\{3-(3+p+\\
&q+r-2q^2-2r^2-5qr+2pq+pr+3pqr+4q^2r+4qr^2+5pr^2-2p^2r-5pqr^2-3p^2r^2-2q^2r^2)\\
&(1-p-q)(1-r)\}(1-p)p]\\
=\;&[-p^5r(2+r-3r^2)-p^4qr^2(9-8r)-p^4qr-11p^4r^3-7p^3q^2r^2(2-r)-20p^3qr^3-p^3qr(4-25r-7q)\\
&-p^3q(1-2p)-p^3r^2(7-14r-4p)-p^3r(8-6p)-2p^2q^3r^2(3-r)-2p^2q^3(1-3r)-9p^2q^2r^3\\
&-p^2q^2r(19-23r)-p^2qr^2(23-11r-r^2)-p^2q(1-12r)-p^2r^4-6p^2r^3-p^2(1-p-5r-p^2-8r^2)\\
&-3pq^3r(2-r)-pq^2r^3(4-r)-5pq^2r^2-pq^2(7-15r-5p-3q)-3pqr^4-pqr(13-9r-6r^2)\\
&-pr^2(6+r-2r^2)-2q^3r^3-q^3(1-3r^2)-q^2r^4-3q^2r^2(1-2r)-5q^2r-qr^3(5-2r)-2qr^2-r^4\\
&+\{-2p^2+pq+pr+p+3q^2+7qr-2q+r^3+2r^2-2r\}](1-p-q)(1-r)\\
\leqslant\;&(p-2q-2r+pq+pr+7qr-2p^2+3q^2+2r^2+r^3)(1-p-q)(1-r) \qquad (4.8.68)
\end{aligned}$$

provided p, q and r are small enough. Combining the terms involving $\mu(WWWD)$ from (2) and (7), we get the following upper bound on the coefficient of $\mu(WWWD)$:

$$(1-p-q)(1-r)\{p-q-r-3p^2-3q^2-3r^2-4pq-3pr-5qr+4p^2r+6pr^2+6q^2r+6qr^2+10pqr\}$$

$$
\begin{aligned}
&-\{3-(3+p+q+r-2q^2-2r^2-5qr+2pq+pr+3pqr+4q^2r+4qr^2+5pr^2-2p^2r-5pqr^2-3p^2r^2\\
&-2q^2r^2)(1-p-q)(1-r)\}(1-p)p(1-p-q)(1-r)\\
=\ &(1-p-q)(1-r)[-p^5r(2+r-3r^2)-p^4qr^2(9-8r)-p^4qr-11p^4r^3-7p^3q^2r^2(2-r)-22p^3qr^3\\
&-p^3qr(6-30r-7q)-2p^3q(1-p)-p^3r(4+12r-15r^2-6pr-4p)-2p^2q^3r^2(3-r)-2p^2q^3(1-3r)\\
&-13p^2q^2r^3-p^2q^2r(19-29r)-p^2qr^2(33-20r)-2pq^3r^3-6pq^3r(1-r)-pq^2r^2(15-6r)\\
&-pq^2(3-12r-2q)-p^2(1-4r-2q-p-p^2-3q^2-11qr-10r^2)-pr(1-3r-2r^2+9pr^2-6q-4p)\\
&-qr(1-6q-6r-12pr)+\{p-q-r-4p^2-3q^2-3r^2-6pq-4pr-4qr\}]\\
\leqslant\ &(1-p-q)(1-r)(p-q-r-4p^2-3q^2-3r^2-6pq-4pr-4qr) \qquad (4.8.69)
\end{aligned}
$$

when p, q and r are sufficiently small.

Incorporating (4.8.68) and (4.8.69) into (4.8.66), we obtain:

$$
\begin{aligned}
w_1(\widehat{G}_{p,q,r}\mu) \leqslant\ & w_1(\mu)-[2-(1-p-q)(1-r)\{2+2p+q+r-q^2-r^2-3qr+pq+2q^2r+2qr^2+pqr\\
&+2pr^2-p^2r-2pqr^2-p^2r^2-q^2r^2\}]\{\mu(WD)-\mu(WWWD)-\mu(DWDD)-\mu(LWD)\}\\
&-\{1-(1+p+q+pq+pr-q^2-qr-p^2r+q^2r)(1-p-q)(1-r)(1+r)\}\mu(DD)\\
&-r^2(1-r)^2(1-p-q)^2\mu(WLD)-r^2(1-r)^2(1-p-q)^2\mu(DLD)+(p-2q-2r+pq\\
&+pr+7qr-2p^2+3q^2+2r^2+r^3)(1-p-q)(1-r)\mu(WWDW)+(1-p-q)(1-r)\\
&(p-q-r-4p^2-3q^2-3r^2-6pq-4pr-4qr)\mu(WWWD)+(1-p-q)(1-r)(p-2q\\
&+pq+qr-2pr-2p^2+3q^2)\mu(WWDD)+\{4r-2(p+q)+(p+q)^2-6r^2-3r(p+2q)\}\\
&(1-p)(1-p-q)(1-r)\mu(WWD). \qquad (4.8.70)
\end{aligned}
$$

At this point, we could further subdivide our analysis into various cases depending on which of these three coefficients, i.e. the coefficients of $\mu(WWDW)$, $\mu(WWWD)$ and $\mu(WWDD)$, is / are non-negative, which would allow us coverage of a *marginally* better regime, while *significantly* lengthening and complicating the computations that follow. We, therefore, adopt the following idea irrespective of which among the coefficients of $\mu(WWDW)$, $\mu(WWDD)$ and $\mu(WWWD)$ in (4.8.70) is / are non-negative: this idea is to make use of the yet-untouched term involving $\mu(WWD)$, appearing in (4), in reducing further each of the coefficients of $\mu(WWDW)$, $\mu(WWDD)$ and $\mu(WWWD)$ in (4.8.70). Implementing the inequality $\mu(WWWD) \leqslant \mu(WWD)$, we can write, from (4.8.69) and keeping p, q and r sufficiently small wherever needed:

$$
\begin{aligned}
&(1-p-q)(1-r)(p-q-r-4p^2-3q^2-3r^2-6pq-4pr-4qr)\mu(WWWD)+\frac{1}{2}\{4r-2(p+q)+\\
&(p+q)^2-6r^2-3r(p+2q)\}(1-p)(1-p-q)(1-r)\mu(WWD)
\end{aligned}
$$

$$\leqslant (1-p-q)(1-r)(p-q-r-4p^2-3q^2-3r^2-6pq-4pr-4qr)\mu(WWWD)+\frac{1}{2}\{4r-2(p+q)+(p+q)^2-6r^2-3r(p+2q)\}(1-p)(1-p-q)(1-r)\mu(WWWD)$$

(the coefficient of $\mu(WWD)$ is non-positive due to (4.2.1))

$$= \Big\{-\frac{p^3}{2}-p^2q-\frac{3pr}{2}(1-p-2q-2r)-\frac{pq^2}{2}-6pr-\frac{5p^2}{2}-4pq-\frac{5q^2}{2}-7qr-2q-6r^2+r\Big\}(1-p-q)(1-r)\mu(WWWD)$$

$$\leqslant \Big(-6pr-\frac{5p^2}{2}-4pq-\frac{5q^2}{2}-7qr-2q-6r^2+r\Big)(1-p-q)(1-r)\mu(WWWD)$$

(when p,q,r are sufficiently small)

$$= \frac{1}{2}\Big\{\Big(-5pr-\frac{p^2}{2}+3pq+\frac{7q^2}{2}-2qr-3q-3r^2+2r\Big)-7pr-\frac{9p^2}{2}-11pq-\frac{17q^2}{2}-12qr-q-9r^2\Big\}(1-p-q)(1-r)\mu(WWWD)$$

$$\leqslant \frac{1}{2}\Big(-5pr-\frac{p^2}{2}+3pq+\frac{7q^2}{2}-2qr-3q-3r^2+2r\Big)(1-p-q)(1-r)\mu(WWWD). \qquad (4.8.71)$$

Likewise, implementing the inequality $\mu(WWDW)+\mu(WWDD)\leqslant\mu(WWD)$, using the expressions from from (4.8.68) and (4.8.65), and keeping p, q and r sufficiently small wherever needed, we can write:

$$(p-2q-2r+pq+pr+7qr-2p^2+3q^2+2r^2+r^3)(1-p-q)(1-r)\mu(WWDW)+(1-p-q)(1-r)(p-2q+pq+qr-3pr-2p^2+3q^2)\mu(WWDD)+\frac{1}{2}\{4r-2(p+q)+(p+q)^2-6r^2-3r(p+2q)\}(1-p)(1-p-q)(1-r)\mu(WWD)$$

$$\leqslant (p-2q-2r+pq+pr+7qr-2p^2+3q^2+2r^2+r^3)(1-p-q)(1-r)\mu(WWDW)+(1-p-q)(1-r)(p-2q+pq+qr-3pr-2p^2+3q^2)\mu(WWDD)+\frac{1}{2}\{4r-2(p+q)+(p+q)^2-6r^2-3r(p+2q)\}(1-p)(1-p-q)(1-r)\{\mu(WWDW)+\mu(WWDD)\}$$

$$= \Big\{-\frac{p^3}{2}-p^2q-\frac{pq^2}{2}-\frac{3pr}{2}(1-p-2q-2r)-5pr-\frac{p^2}{2}+3pq+\frac{7q^2}{2}-2qr-3q-3r^2+2r\Big\}(1-p-q)(1-r)\mu(WWDD)+\Big\{-\frac{p^3}{2}-p^2q-\frac{pq^2}{2}-\frac{3pr}{2}(1-p-2q-2r)-pr-\frac{p^2}{2}+\frac{7q^2}{2}+4qr-3q-r^2+3pq+r^3\Big\}(1-p-q)(1-r)\mu(WWDW)$$

$$\leqslant \Big(-5pr-\frac{p^2}{2}+3pq+\frac{7q^2}{2}-2qr-3q-3r^2+2r\Big)(1-p-q)(1-r)\mu(WWDD)+\Big(-pr-\frac{p^2}{2}+\frac{7q^2}{2}+4qr-3q-r^2+3pq+r^3\Big)(1-p-q)(1-r)\mu(WWDW)$$

$$
\begin{aligned}
&= \Big(-5pr - \frac{p^2}{2} + 3pq + \frac{7q^2}{2} - 2qr - 3q - 3r^2 + 2r\Big)(1-p-q)(1-r)\mu(WWDD) + (1-p-q)(1-r) \\
&\quad \Big\{\Big(-5pr - \frac{p^2}{2} + 3pq + \frac{7q^2}{2} - 2qr - 3q - 3r^2 + 2r\Big) - r(2-4p-6q-2r-r^2)\Big\}\mu(WWDW) \\
&\leqslant \Big(-5pr - \frac{p^2}{2} + 3pq + \frac{7q^2}{2} - 2qr - 3q - 3r^2 + 2r\Big)(1-p-q)(1-r)\{\mu(WWDD) + \mu(WWDW)\}.
\end{aligned} \tag{4.8.72}
$$

The final coefficient of each of $\mu(WWWD)$, $\mu(WWDD)$ and $\mu(WWDW)$ in (4.8.71) and (4.8.72) is non-positive due to the last inequality of (4.2.5).

Incorporating the updated coefficients of $\mu(WWDW)$, $\mu(WWDD)$ and $\mu(WWWD)$ from (4.8.71) and (4.8.72) into (4.8.66), we obtain:

$$
\begin{aligned}
w_1(\widehat{G}_{p,q,r}\mu) \leqslant\ & w_1(\mu) - [2 - (1-p-q)(1-r)\{2 + 2p + q + r - q^2 - r^2 - 3qr + pq + 2q^2r + 2qr^2 + pqr + \\
& 2pr^2 - p^2r - 2pqr^2 - p^2r^2 - q^2r^2\}]\{\mu(WD) - \mu(WWWD) - \mu(DWDD) - \mu(LWD)\} \\
& - \{1 - (1 + p + q + pq + pr - q^2 - qr - p^2r + q^2r)(1-p-q)(1-r)(1+r)\}\mu(DD) \\
& - r^2(1-r)^2(1-p-q)^2\mu(WLD) - r^2(1-r)^2(1-p-q)^2\mu(DLD) + \\
& \left(-5pr - \frac{p^2}{2} + 3pq + \frac{7q^2}{2} - 2qr - 3q - 3r^2 + 2r\right)(1-p-q)(1-r) \\
& \left\{\frac{1}{2}\mu(WWWD) + \mu(WWDD) + \mu(WWDW)\right\} \\
\leqslant\ & w_1(\mu) - [2 - (1-p-q)(1-r)\{2 + 2p + q + r - q^2 - r^2 - 3qr + pq + 2q^2r + 2qr^2 + pqr + \\
& 2pr^2 - p^2r - 2pqr^2 - p^2r^2 - q^2r^2\}]\{\mu(WD) - \mu(WWWD) - \mu(DWDD) - \mu(LWD)\} \\
& - \{1 - (1 + p + q + pq + pr - q^2 - qr - p^2r + q^2r)(1-p-q)(1-r)(1+r)\}\mu(DD) \\
& - r^2(1-r)^2(1-p-q)^2\mu(WLD) - r^2(1-r)^2(1-p-q)^2\mu(DLD),
\end{aligned}
$$

where the last step is obtained by eliminating the terms involving $\mu(WWDW)$, $\mu(WWDD)$ and $\mu(WWWD)$ from the previous step, which is allowed because the coefficient in each of these terms is non-positive when (4.2.5) is true. Note that the final inequality deduced above is the same the inequality stated in (4.8.10) (once we replace the notation w_1 by simply w), and (4.8.10) is of the same form as (4.8.2) and satisfies (4.8.3). This completes the construction of our weight function (the final form of which is as given by (4.8.9)), for the generalized percolation games, when the underlying parameter-triple (p,q,r) satisfies the constraints given by (4.2.5) of Theorem 4.2.1.

4.9 The proof of Theorem 4.7.2 by the technique of weight functions

The principal ideas involved in constructing our weight function for each of the three regimes stated in Theorem 4.2.4 are the same as those outlined in §4.8.1 (including, as outlined in §4.8.1.1, the ideas implemented for the iterative tweaking of the weight function until it satisfies an inequality of the form given by (4.8.2), along with the criterion stated in (4.8.3)). However, as becomes evident when one goes through the steps leading to the construction of the weight function for the proof of Theorem 4.7.1, outlined in §4.8.3, the details involved are rather *ad hoc* and *specific* to the regime of values of the parameter-pair (r', s') under consideration. Consequently, we dedicate §4.9.2 to the detailed construction of the weight function for the regime given by (B1), §4.9.3 to that for the regime given by (B2), and §4.9.4 to that for the regime given by (B3).

4.9.1 The final weight functions obtained for the bond percolation game when (r', s') belongs to one of the regimes given by (B1), (B2) and (B3)

As in §4.8.2, we state here, for the reader's convenience, the final weight function we have come up with, and the corresponding weight function inequality it satisfies, for (r', s') belonging to each of the regimes given by (B1), (B2) and (B3). When (r', s') satisfies the constraints of (B1), the weight function is given by

$$\begin{aligned} w(\mu) = {} & \mu(D) + \mu(WD) + \Big[1 - (1-r'-s')\Big\{2(2s'-s'^2)(1-s') + r' + (2s'-s'^2)^2(1-s') + \frac{2r'^2}{(1-s')} \\ & - (2s'-s'^2)^2 r' - \frac{5(2s'-s'^2)r'^2}{(1-s')}\Big\}\Big]\mu(LWD) - (1-r'-s')\Big\{-2(2s'-s'^2)(1-s') - 4r' \\ & + (2s'-s'^2)^2(1-s') + \frac{3r'^2}{(1-s')} + 2(2s'-s'^2)r' - 2(2s'-s'^2)^2 r' - \frac{6(2s'-s'^2)r'^2}{(1-s')}\Big\}\mu(LD) - \\ & (1-r'-s')\Big\{3r' + 2(2s'-s'^2)^2(1-s') + \frac{2r'^2}{(1-s')} - (2s'-s'^2)r' - 2(2s'-s'^2)^2 r' - \frac{4(2s'-s'^2)r'^2}{(1-s')} \\ & - \frac{r'^3}{(1-s')^2}\Big\}\mu(LDW) - (1-s')(1-r'-s')\Big\{(1-s')^2 - r'^2 + (2s'-s'^2)r'^2\Big\}\mu(LDL), \end{aligned} \tag{4.9.1}$$

and the corresponding weight function inequality is given by

$$w(\widehat{E}_{r',s'}\mu) \leqslant w(\mu) - \frac{(1-r'-s')\{(1-s')(2s'-s'^2)-r'\}^2}{1-s'}\mu(DD). \tag{4.9.2}$$

We now deduce, from (4.9.2), the conclusion we desire, i.e. $\mu(D) = 0$ for any translation-invariant and reflection-invariant probability distribution μ that is stationary for $\widehat{E}_{r',s'}$. As seen in §4.3.2, $\widehat{E}_{r',s'}$ is equivalent to $\widehat{G}_{p,q,r}$ if we set $q = 0$ and let p and r equal suitable functions of r' and s', as given by (4.3.13) and (4.3.14). From (4.3.14), we see that $r' = 0 \implies r = 0$, so that we are back to the set-up of [60], with $q = 0$ and $p = 2s' - s'^2$, when $r' = 0$. Our desired conclusion , therefore, follows from [60] when $r' = 0$. It thus suffices for us to consider $r' > 0$.

Since r' and s' are sufficiently small, we have $(1 - r' - s') > 0$, and because of (4.2.6), we have $(1-s')(2s'-s'^2)-r' > 0$ as well, so that the coefficient of $\mu(DD)$ in (4.9.2) is strictly positive. Consequently, when μ is stationary for $\widehat{E}_{r',s'}$, the parameters r' and s' are sufficiently small and they satisfy the inequality in (4.2.6), we have $\mu(DD) = 0$, which, in turn, implies that $\widehat{E}_{r',s'}\mu(DD) = 0$ as well. From (4.3.10), (4.3.11) and (4.3.12), we see that

$$(1-s')^2(1-r'-s')^2\mu(WDW)+2r'(1-s')(1-r'-s')^2\mu(WDL)+r'^2(1-r'-s')^2\mu(LDL) \leqslant \widehat{E}_{r',s'}\mu(DD),$$

so that, for r' and s' sufficiently small, $r' > 0$, and (r', s') satisfying (4.2.6), we have $\mu(WDW) = \mu(WDL) = \mu(LDW) = \mu(LDL) = 0$ when μ is any translation-invariant and reflection-invariant stationary distribution for $\widehat{E}_{r',s'}$. Moreover, we already know that $\mu(WDD) = \mu(DDW) = \mu(LDD) = \mu(DDL) = \mu(DDD) = 0$ as each of these is bounded above by $\mu(DD)$. Combining all of these, we obtain $\mu(D) = 0$, as desired.

When (r', s') belongs to the regime given by (B2), the weight function takes various forms depending on the subset of $(0.157175, 1]$ that the parameter r' belongs to. When $s' = 0$ and $r' \in [0.4564, 1]$, it suffices to consider the weight function

$$w(\mu) = \mu(D) + \mu(WD) + \mu(LWD) - 2r'\mu(WD), \tag{4.9.3}$$

and the corresponding weight function inequality is given by

$$\begin{aligned} w(\widehat{E}_{r',0}\mu) \leqslant{}& w(\mu) - r'^2\mu(DD) + (2r' - 6r'^2 + 4r'^3 - r'^4)\mu(LD) - (3r' - 3r'^2 + r'^3)\mu(LWD) \\ &- (1-r')^2(1-r'+2r'^2)\mu(LDL) - 2r'^2(1-r')^2\mu(LDD) - r'(1-r')^2(1+r')\mu(LDW) \\ &- r'(1-r')^4\mu(LLD) - r'(1-r')^3\mu(DLD). \end{aligned} \tag{4.9.4}$$

When $s' = 0$ and $r' \in (0.201383, 0.4564)$, the weight function that yields the desired conclusion is

$$w(\mu) = \mu(D) + \mu(WD) + \mu(LWD) - 2r'\mu(WD) - (3r' - 3r'^2 + r'^3)\mu(LWD), \tag{4.9.5}$$

and the corresponding weight function inequality is given by

$$\begin{aligned} w(\widehat{E}_{r',0}\mu) \leqslant\ & w(\mu) - r'^2\mu(DD) + (2r' - 14r'^2 + 23r'^3 - 14r'^4 + 3r'^6 - r'^7)\mu(LD) \\ & - (1-r')^2(1 - r' - 3r'^2 + 3r'^3 - r'^4)\mu(LDL) - r'(1-r')^2(1 - r' - 3r'^2 + 3r'^3 - r'^4)\mu(LDW) \\ & - r'(1-r')^2(r'^5 - 2r'^4 - r'^3 + 7r'^2 - 5r' + 1)\mu(LLD) - r'(1-r')^6\mu(DLD) \\ & - (3r' - 3r'^2 + r'^3)r'(1-r')^2\mu(LWD). \end{aligned} \tag{4.9.6}$$

For the final sub-regime, i.e. where $s' = 0$ and $r' \in (0.157175, 0.201383]$, the weight function is given by

$$\begin{aligned} w(\mu) = \ & \mu(D) + \mu(WD) + \mu(LWD) - 2r'\mu(WD) - (3r' - 3r'^2 + r'^3)\mu(LWD) - r'(1-r')^6\mu(LLWD) \\ & - r'(1-r')^6\mu(LLDW) - r'^2\mu(WDD) - (3r' - 3r'^2 + r'^3)r'(1-r')^2\mu(LWD) \\ & - r'(1-r')^4(1 - 3r' + r'^2 + r'^3 - 3r'^4 + 3r'^5 - r'^6)\mu(LWLD), \end{aligned} \tag{4.9.7}$$

and the corresponding weight function inequality is of the form

$$\begin{aligned} w(\widehat{E}_{r',0}\mu) \leqslant\ & w(\mu) + r'(2 - 16r' + 20r'^2 + 7r'^3 - 42r'^4 + 51r'^5 - 35r'^6 + 7r'^7 + 13r'^8 - 14r'^9 + 6r'^{10} - r'^{11}) \\ & \mu(LD) - r'^2\mu(DDD). \end{aligned} \tag{4.9.8}$$

When μ is a reflection-invariant and translation-invariant stationary measure for $\widehat{E}_{r',0}$, and (r', s') belongs to any one of the sub-regimes mentioned above, the relevant inequality out of (4.9.4), (4.9.6) and (4.9.8) yields $\mu(LD) = 0$. As μ is stationary for $\widehat{E}_{r',0}$, this implies $\widehat{E}_{r',0}\mu(LD) = 0$ as well. In order to deduce, from $\mu(LD) = \widehat{E}_{r',0}\mu(LD) = 0$, the conclusion that $\mu(D) = 0$ as well, we need to compute, partially, the pushforward probability $\widehat{E}_{r',0}\mu(LD)$, i.e. the probability of the event $\{\widehat{E}_{r',0}\eta(0) = L, \widehat{E}_{r',0}\eta(1) = D\}$ where η follows the law μ. It suffices for us to consider the contribution to $\widehat{E}_{r',0}\mu(LD)$ only from the event $\{\eta(1) = D\}$, which is given by

$$\begin{aligned} & r'(1-r')\mu(WDW) + r'(1-r')(1+r')\mu(WDD) + r'^2(1-r')\mu(WDL) + r'^2(1-r')\mu(DDW) \\ & + r'^2(1-r')(1+r')\mu(DDD) + r'^3(1-r')\mu(DDL) + r'^2(1-r')\mu(LDW) + r'^2(1-r')(1+r')\mu(LDD) \\ & + r'^3(1-p)\mu(LDL) \leqslant \widehat{E}_{r',0}\mu(LD). \end{aligned}$$

Since $\widehat{E}_{r',0}\mu(LD) = 0$, we conclude that each of $\mu(WDW)$, $\mu(WDD)$, $\mu(WDL)$, $\mu(DDW)$, $\mu(DDD)$, $\mu(DDL)$, $\mu(LDW)$, $\mu(LDD)$ and $\mu(LDL)$ equals 0. Adding these, we obtain $\mu(D) = 0$, as desired.

For the regime described in (B3), our weight function is given by

$$\begin{aligned} w(\mu) = \mu(D) + \mu(WD) &- (2r'^2 - 5r'^3 + 2r'^4)\mu(WDW) - (3r' - 10r'^2 + 10r'^3 - 4r'^4)\mu(WDL) \\ &- (7r'^2 - 7r'^3 + 2r'^4)\mu(DDW) - (2r' - 2r'^2 + 5r'^3 - 2r'^4)\mu(DDL) \\ &- (1 - 4r' + 6r'^2 - 5r'^3 + 2r'^4)\mu(LDL) - (4r'^2 - r'^3 + 2r'^4)\mu(WLD) \\ &- (3r'^2 + 3r'^3 - 2r'^4)\mu(LLD) - (2r' + 8r'^2 - 9r'^3 + 2r'^4)\mu(WWD), \end{aligned} \tag{4.9.9}$$

and the corresponding weight function inequality can be written as follows:

$$w(\widehat{E}_{r',r'}\mu) \leqslant w(\mu) + (1 - 10r' + 7r'^2 + 64r'^4 - 292r'^5 + 583r'^6 - 663r'^7 + 447r'^8 - 168r'^9 + 28r'^{10})\mu(LWD). \tag{4.9.10}$$

The coefficient of $\mu(LWD)$ in (4.9.10) is negative whenever $r' = s' > 0.10883$, so that, when μ is a translation-invariant and reflection-invariant stationary measure for $\widehat{E}_{r',r'}$, we conclude, from (4.9.10), that $\mu(LWD) = 0$. This, in turn, leads to $\widehat{E}_{r',r'}\mu(LWD) = 0$ since μ is stationary for $\widehat{E}_{r',r'}$. To obtain our desired conclusion (i.e. $\mu(D) = 0$), we now partially compute $\widehat{E}_{r',r'}\mu(LWD)$, which is the probability of the event $\{\widehat{E}_{r',r'}\eta(0) = L, \widehat{E}_{r',r'}\eta(1) = W, \widehat{E}_{r',r'}\eta(2) = D\}$, where η follows the distribution μ. We focus only on the contributions arising from the event $\{\eta(3) = D\}$. By (4.3.7) through (4.3.12), we obtain

$$\begin{aligned} \widehat{E}_{r',r'}\mu(LWD) \geqslant\ & (1-r')^2(2r'-r'^2)(1-r')(1-2r')\mu(WWWD) + r'(1-r')(2r'-r'^2)(1-r') \\ & (1-2r')\mu(LWWD) + r'(1-r')(2r'-r'^2)(1-r')(1-2r')\mu(DWWD) + r'(1-r') \\ & (1-r'+r'^2)(1-r')(1-2r')\mu(WLWD) + r'^2(1-r'+r'^2)(1-r')(1-2r')\mu(LLWD) \\ & + r'^2(1-r'+r'^2)(1-r')(1-2r')\mu(DLWD) + r'(1-r')(2r'-r'^2)(1-r')(1-2r') \\ & \mu(WDWD) + r'^2(2r'-r'^2)(1-r')(1-2r')\mu(LDWD) + r'^2(2r'-r'^2)(1-r')(1-2r') \\ & \mu(DDWD) + (1-r')^2(1-r'+r'^2)r'(1-2r')\mu(WWLD) + r'(1-r')(1-r'+r'^2)r' \\ & (1-2r')\mu(LWLD) + r'(1-r')(1-r'+r'^2)r'(1-2r')\mu(DWLD) + r'(1-r')(1-r'^2)r' \\ & (1-2r')\mu(WLLD) + r'^2(1-r'^2)r'(1-2r')\mu(LLLD) + r'^2(1-r'^2)r'(1-2r')\mu(DLLD) \\ & + r'(1-r')(1-r'+r'^2)r'(1-2r')\mu(WDLD) + r'^2(1-r'+r'^2)r'(1-2r')\mu(LDLD) \\ & + r'^2(1-r'+r'^2)r'(1-2r')\mu(DDLD) + (1-r')^2(2r'-r'^2)(1-2r')\mu(WWDD) \end{aligned}$$

$$
\begin{aligned}
&+r'(1-r')(2r'-r'^2)(1-2r')\mu(LWDD)+r'(1-r')(2r'-r'^2)(1-2r')\mu(DWDD)\\
&+r'(1-r')(1-r'+r'^2)(1-2r')\mu(WLDD)+r'^2(1-r'+r'^2)(1-2r')\mu(LLDD)\\
&+r'^2(1-r'+r'^2)(1-2r')\mu(DLDD)+r'(1-r')(2r'-r'^2)(1-2r')\mu(WDDD)\\
&+r'^2(2r'-r'^2)(1-2r')\mu(LDDD)+r'^2(2r'-r'^2)(1-2r')\mu(DDDD) \qquad (4.9.11)
\end{aligned}
$$

(the inequality arises from the fact that we did not take into account the contribution from tuples $(\eta(0),\eta(1),\eta(2),\eta(3))$ for which $\eta(3)\in\{W,L\}$). Since we consider $0<r'<0.5$ (when $r'=s'$, it makes sense to consider $r'\leqslant 0.5$ since $r'+s'\leqslant 1$, and when $r'+s'=1$, each edge of our graph would be labeled either a trap or a target, bringing the game to an end right after the first round), the coefficient of each term in the right side of (4.9.11) is strictly positive. Since $\widehat{E}_{r',r'}\mu(LWD)=0$ when μ is stationary for $\widehat{E}_{r',r'}$, we obtain, from (4.9.11), that $\mu(\mathscr{C})=0$ for each cylinder set comprising the expression in the right side of (4.9.11) (i.e. each of $\mu(WWWD)$, $\mu(LWWD)$, and so on, equals 0). Adding them all, we obtain $\mu(D)=0$, as desired.

4.9.2 Detailed construction of the weight function for the bond percolation game when (r',s') belongs to the regime given by (B1)

As in the case of §4.8.3, it is assumed, throughout §4.9.2, that μ is a translation-invariant and reflection-invariant measure on the state space $\Omega=\hat{\mathscr{A}}^{\mathbb{Z}}=\{W,L,D\}^{\mathbb{Z}}$ of $\widehat{E}_{r',s'}$, measurable with respect to the σ-field $\mathscr{F}$ generated by all cylinder sets of Ω. Recall, from §4.3.2, that the envelope PCA $\widehat{E}_{r',s'}$, corresponding to the bond percolation game with underlying parameter-pair (r',s'), identical to the envelope PCA $\widehat{G}_{p,q,r}$, corresponding to the generalized percolation game with underlying parameter-triple (p,q,r), provided we set (see (4.3.13) and (4.3.14))

$$
p=2s'-s'^2,\quad q=0\quad\text{and}\quad r=\frac{r'}{1-s'}. \qquad (4.9.12)
$$

With these values of p and r, the inequality in (4.2.6) boils down to

$$
\frac{3p^2}{2}+8pr+\frac{11r^2}{2}\geqslant 2r+\frac{9p^3}{2}+16p^2r+\frac{43pr^2}{2}+5r^3, \qquad (4.9.13)
$$

and proving the claim that $\nu(D)=0$ for any translation-invariant and reflection-invariant stationary distribution ν for the PCA $\widehat{E}_{r',s'}$, whenever (r',s') satisfies the constraints of (B1), is equivalent to showing, instead, that $\mu(D)=0$ for any translation-invariant and reflection-invariant stationary distribution μ for $\widehat{G}_{p,0,r}$, where p and r are as given by (4.9.12) and (p,r) satisfies the inequality in (4.9.13).

Recall that our computations in §4.8.3 were performed *without* making the assumption that $q = 0$, and hence we may proceed in the same manner as in §4.8.3, up to and including the derivation of (4.8.45). Setting $w_0(\mu) = \mu(D) + \mu(WD) + \mu(LWD)$, the weight function inequality obtained can be written as

$$\begin{aligned}
w_0(\widehat{G}_{p,0,r}\mu) \leqslant w_0(\mu) &- [2-(1-p)(1-r)\{2+2p+r-r^2+2pr^2-p^2r-p^2r^2\}]\mu(WD)\\
&-\{1-(1+p+pr-p^2r)(1-p)(1-r)(1+r)\}\mu(DD)-[1-\{1-p+4r\\
&-r^2(2-3p)-r^3(1-p)-p^2r^2\}(1-p)(1-r)]\mu(LD)-r(1-p)^2(1-r)\\
&(1-p-r)\mu(LDW)-(1-p)^2(1-r)[(1-r)\{1-r^2(1-p)\}-pr^2]\mu(LDL)\\
&-[1-(1-p-r^2+pr+3pr^2-p^2r-2p^2r^2)(1-p)(1-r)]\mu(LWD)\\
&-r^2(1-r)^2(1-p)^2\mu(WLD)-r^2(1-r)^2(1-p)^2\mu(DLD)+\{p-r+r^2(1-p)\}\\
&(1-p)^2(1-r)\mu(WWDW)+p(1-p)^2(1-r)^2\mu(WWWD)+A_6, \qquad (4.9.14)
\end{aligned}$$

where A_6 comprises the contribution to $\widehat{G}_{p,0,r}\mu(LWD)$ (which, recall, is the probability of the event $\{\widehat{G}_{p,0,r}\eta(0) = L, \widehat{G}_{p,0,r}\eta(1) = W, \widehat{G}_{p,0,r}\eta(2) = D\}$ when η follows the law μ) of the events $\{\eta(1) = L, \eta(2) = D, \eta(3) = D\}$ and $\{\eta(0) \neq L, \eta(1) = W, \eta(2) = \eta(3) = D\}$, where η follows the distribution μ. Next, substituting $q = 0$ in (4.8.46), the contribution to $\widehat{G}_{p,0,r}\mu(LWD)$ of the event $\{\eta(1) = L, \eta(2) = \eta(3) = D\}$ can be written as

$$\begin{aligned}
&r(1-r+pr)(1-p)^2(1-r)(1+r)\mu(LDD)-r\{p+(1-p)(1-r)\}(1-p)^2(1-r)^2(1+r)\\
&\{\mu(DLDD)+\mu(LLDD)\}\leqslant r(1-r+pr)(1-p)^2(1-r)(1+r)\mu(LDD). \qquad (4.9.15)
\end{aligned}$$

Combining (4.9.15) with the term involving $\mu(LD)$ in (4.9.14) yields

$$\begin{aligned}
&-[1-\{1-p+4r-r^2(2-3p)-r^3(1-p)-p^2r^2\}(1-p)(1-r)]\mu(LD)\\
&+r(1-r+pr)(1-p)^2(1-r)(1+r)\mu(LDD)\\
=\,&[-(1-p)(1-r)\{(1-p)(1-r)\}^{-1}+\{1-p+4r-r^2(2-3p)-r^3(1-p)-p^2r^2\}(1-p)(1-r)]\\
&\mu(LD)+r(1-r+pr)(1-p)^2(1-r)(1+r)\mu(LDD)\\
\leqslant\,&[-(1-p)(1-r)\{1+p+r+p^2+r^2+pr-2p^2r-2pr^2+p^2r^2\}+(1-p)(1-r)\{1-p+4r\\
&-r^2(2-3p)-r^3(1-p)-p^2r^2\}]\mu(LD)+r(1-r+pr)(1-p)^2(1-r)(1+r)\{\mu(LD)\\
&-\mu(LDW)-\mu(LDL)\}\\
=\,&(1-p)(1-r)\{-2p+4r-p^2-3r^2-2pr+2p^2r+6pr^2-r^3(2-3p)-3p^2r^2-p^2r^3\}\mu(LD)
\end{aligned}$$

$$
\begin{aligned}
&-r(1-r+pr)(1-p)^2(1-r)(1+r)\{\mu(LDW)+\mu(LDL)\}\\
&\leqslant (1-p)(1-r)\{-2p+4r-p^2-3r^2-2pr+2p^2r+6pr^2\}\mu(LD)\\
&\quad -r(1-r+pr)(1-p)^2(1-r)(1+r)\{\mu(LDW)+\mu(LDL)\} \qquad (4.9.16)
\end{aligned}
$$

when p is sufficiently small (it suffices for p to be bounded above by $2/3$). Note, here, that when (4.9.13) holds, we have

$$
\begin{aligned}
&-2p+4r-p^2-3r^2-2pr+2p^2r+6pr^2\\
\leqslant\ & -2p+3p^2+16pr+11r^2-9p^3-32p^2r-43pr^2-10r^3-p^2-3r^2-2pr+2p^2r+6pr^2\\
=\ & -2p+2p^2+14pr+8r^2-9p^3-30p^2r-37pr^2-10r^3\\
\leqslant\ & -2p+2p^2+14pr+\frac{1}{2}\left(3p^2+16pr+11r^2\right)^2-9p^3-30p^2r-37pr^2-10r^3\\
=\ & -2p+2p^2+14pr+9p^4+256p^2r^2+121r^4+96p^3r+66p^2r^2+352pr^3-9p^3-30p^2r-37pr^2-10r^3\\
=\ & -p(2-2p-14r+9p^2-9p^3)-p^2r(30-96p-161r)-pr^2(37-352r-161p)-r^3(10-121r),
\end{aligned}
$$

showing us that the coefficient of $\mu(LD)$ in (4.9.16) is non-positive when p and r are sufficiently small and (4.9.13) holds.

The contribution to $\widehat{G}_{p,0,r}\mu(LWD)$ arising from the event $\{\eta(0)\in\{W,D\},\eta(1)=W,\eta(2)=\eta(3)=D\}$ is given by

$$
p(1-p)^2(1-r)(1+r)\mu(WWDD)+pr(1-p)^2(1-r)(1+r)\mu(DWDD). \qquad (4.9.17)
$$

Since $\mu(WWDD)+\mu(DWDD)\leqslant\mu(WDD)\leqslant\mu(WD)$, we combine the term involving $\mu(WD)$ in (4.9.14) with the terms in (4.9.17) to write:

$$
\begin{aligned}
&-[2-(1-p)(1-r)\{2+2p+r-r^2+2pr^2-p^2r-p^2r^2\}]\mu(WD)+p(1-p)^2(1-r)(1+r)\mu(WWDD)\\
&\quad +pr(1-p)^2(1-r)(1+r)\mu(DWDD)\\
&=[-2(1-p)(1-r)\{(1-p)(1-r)\}^{-1}+(1-p)(1-r)\{2+2p+r-r^2+2pr^2-p^2r-p^2r^2\}]\mu(WD)\\
&\quad +p(1-p)^2(1-r)(1+r)\mu(WWDD)+pr(1-p)^2(1-r)(1+r)\mu(DWDD)\\
&\leqslant[-2(1-p)(1-r)\{1+p+r+p^2+r^2+pr-2p^2r-2pr^2+p^2r^2\}+(1-p)(1-r)\{2+2p+r-r^2\\
&\quad +2pr^2-p^2r-p^2r^2\}]\mu(WD)+p(1-p)^2(1-r)(1+r)\mu(WWDD)+pr(1-p)^2(1-r)(1+r)\\
&\quad \mu(DWDD) \quad \text{(substituting } q=0 \text{ in (4.8.31))}\\
&=(1-p)(1-r)\{-r-2p^2-3r^2-2pr+3p^2r+6pr^2-3p^2r^2\}\mu(WD)+p(1-p)^2(1-r)(1+r)\\
&\quad \mu(WWDD)+pr(1-p)^2(1-r)(1+r)\mu(DWDD)
\end{aligned}
$$

$$
\begin{aligned}
&= (1-p)(1-r)\{-r-2p^2-3r^2-2pr+3p^2r+6pr^2-3p^2r^2\}\{\mu(WD)-\mu(WWDD)-\mu(DWDD)\} \\
&\quad +(1-p)(1-r)(p-r-3p^2-3r^2-pr+2p^2r+6pr^2-3p^2r^2)\mu(WWDD) \\
&\quad +(1-p)(1-r)(-r-pr-2p^2-3r^2+7pr^2+2p^2r-4p^2r^2)\mu(DWDD) \\
&\leqslant (1-p)(1-r)\{-r-2p^2-3r^2-2pr+3p^2r+6pr^2-3p^2r^2\}\{\mu(WD)-\mu(WWDD)-\mu(DWDD)\} \\
&\quad +(1-p)(1-r)(p-r-3p^2-3r^2-pr+2p^2r+6pr^2)\mu(WWDD) \qquad (4.9.18)
\end{aligned}
$$

(we can remove the term involving $\mu(DWDD)$ in the last step since the coefficient of $\mu(DWDD)$ is evidently non-positive for p and r sufficiently small). Note that $\mu(WWDD)+\mu(DWDD)+\mu(WDL)\leqslant\mu(WD)$, so that we can combine the terms involving $\mu(LDW)$ and $\mu(LDL)$ in (4.9.14), the terms involving $\mu(LDW)$ and $\mu(LDL)$ in (4.9.16), and the term involving $\mu(WD)-\mu(WWDD)-\mu(DWDD)$ in (4.9.18) to get:

$$
\begin{aligned}
&-r(1-p)^2(1-r)(1-p-r)\mu(LDW)-(1-p)^2(1-r)[(1-r)\{1-r^2(1-p)\}-pr^2]\mu(LDL)-r(1-r \\
&\quad +pr)(1-p)^2(1-r)(1+r)\{\mu(LDW)+\mu(LDL)\}+(1-p)(1-r)\{-r-2p^2-3r^2-2pr+3p^2r+6pr^2 \\
&\quad -3p^2r^2\}\{\mu(WD)-\mu(WWDD)-\mu(DWDD)\} \\
&= (1-p)(1-r)\{-3r-2p^2-2r^2+pr+2p^2r+4pr^2+r^3-pr^3(2-p)-2p^2r^2\}\mu(LDW) \\
&\quad -(1-p)^2(1-r)(1-r^2+pr^2)\mu(LDL)+(1-p)(1-r)\{-r-2p^2-3r^2-2pr+3p^2r+6pr^2-3p^2r^2\} \\
&\quad \{\mu(WD)-\mu(WWDD)-\mu(DWDD)-\mu(WDL)\} \\
&\leqslant (1-p)(1-r)\{-3r-2p^2-2r^2+pr+2p^2r+4pr^2+r^3\}\mu(LDW)-(1-p)^2(1-r)(1-r^2+pr^2)\mu(LDL) \\
&\quad +(1-p)(1-r)\{-r-2p^2-3r^2-2pr+3p^2r+6pr^2-3p^2r^2\}\{\mu(WD)-\mu(WWDD)-\mu(DWDD) \\
&\quad -\mu(WDL)\}. \qquad (4.9.19)
\end{aligned}
$$

Finally, we bound above the coefficient of $\mu(LWD)$ in (4.9.14) as follows (once again, making use of (4.8.31), with $q=0$):

$$
\begin{aligned}
&-[1-(1-p-r^2+pr+3pr^2-p^2r-2p^2r^2)(1-p)(1-r)] \\
&= -(1-p)(1-r)\{(1-p)(1-r)\}^{-1}+(1-p)(1-r)(1-p-r^2+pr+3pr^2-p^2r-2p^2r^2) \\
&\leqslant (1-p)(1-r)\{1+p+r+p^2+r^2+pr-2p^2r-2pr^2+p^2r^2\} \\
&\quad +(1-p)(1-r)(1-p-r^2+pr+3pr^2-p^2r-2p^2r^2) \\
&= (1-p)(1-r)(-2p-r-p^2-2r^2+p^2r+5pr^2-3p^2r^2) \\
&\leqslant -(1-p)(1-r)\{2p+r+p^2+2r^2-p^2r-5pr^2\}. \qquad (4.9.20)
\end{aligned}
$$

Incorporating the term involving $\mu(LD)$ from (4.9.16), the term involving $\mu(WWDD)$ from

(4.9.18), the terms from (4.9.19), and the updated coefficient of $\mu(LWD)$ from (4.9.20) into (4.9.14), we obtain:

$$\begin{aligned} w_0(\widehat{G}_{p,0,r}\mu) \leqslant\ & w_0(\mu) + (1-p)(1-r)\{-r-2p^2-3r^2-2pr+3p^2r+6pr^2-3p^2r^2\}\{\mu(WD)-\mu(WWDD) \\ & -\mu(DWDD)-\mu(WDL)\} - \{1-(1+p+pr-p^2r)(1-p)(1-r)(1+r)\}\mu(DD) - (1-p) \\ & (1-r)\{2p-4r+p^2+3r^2+2pr-2p^2r-6pr^2\}\mu(LD) - (1-p)(1-r)\{3r+2p^2+2r^2-pr \\ & -2p^2r-4pr^2-r^3\}\mu(LDW) - (1-p)^2(1-r)(1-r^2+pr^2)\mu(LDL) - (1-p)(1-r) \\ & \{2p+r+p^2+2r^2-p^2r-5pr^2\}\mu(LWD) - r^2(1-r)^2(1-p)^2\mu(WLD) - r^2(1-r)^2(1-p)^2 \\ & \mu(DLD) + \{p-r+r^2(1-p)\}(1-p)^2(1-r)\mu(WWDW) + p(1-p)^2(1-r)^2\mu(WWWD) \\ & +(1-p)(1-r)(p-r-3p^2-3r^2-pr+2p^2r+6pr^2)\mu(WWDD). \qquad (4.9.21) \end{aligned}$$

The non-negative coefficients on the right side of (4.9.21), other than $w_0(\mu)$ itself, are those that correspond to $\mu(WWDW)$, $\mu(WWWD)$ and $\mu(WWDD)$.

We now update the weight function (the same way as we did in (4.8.63)) w_0 as follows:

$$\begin{aligned} w_1(\mu) = w_0(\mu) & - (1-p)(1-r)\{2p-4r+p^2+3r^2+2pr-2p^2r-6pr^2\}\mu(LD) - (1-p)(1-r) \\ & \{3r+2p^2+2r^2-pr-2p^2r-4pr^2-r^3\}\mu(LDW) - (1-p)^2(1-r)(1-r^2+pr^2)\mu(LDL) \\ & -(1-p)(1-r)\{2p+r+p^2+2r^2-p^2r-5pr^2\}\mu(LWD), \qquad (4.9.22) \end{aligned}$$

which is exactly the same as what we stated in (4.9.1) (once we replace w_1 by w and set p and r to be the functions of r' and s' as given by (4.9.12)). In order to understand how the update in (4.9.22) impacts the weight function inequality in (4.9.21), we need to compute $\widehat{G}_{p,0,r}\mu(LD)$, $\widehat{G}_{p,0,r}\mu(LDW)$, $\widehat{G}_{p,0,r}\mu(LDL)$ and $\widehat{G}_{p,0,r}\mu(LDW)$, but only partially, just as we did in (4.8.59), (4.8.60), (4.8.61) and (4.8.62):

$$\begin{aligned} & \bullet (1-p)^2(1-r)\mu(WWD) \leqslant \widehat{G}_{p,0,r}\mu(LD) \\ \iff & -(1-p)(1-r)\{2p-4r+p^2+3r^2+2pr-2p^2r-6pr^2\}\widehat{G}_{p,0,r}\mu(LD) \leqslant -(1-p)(1-r) \\ & \{2p-4r+p^2+3r^2+2pr-2p^2r-6pr^2\}(1-p)^2(1-r)\mu(WWD) \\ \iff & -(1-p)(1-r)\{2p-4r+p^2+3r^2+2pr-2p^2r-6pr^2\}\widehat{G}_{p,0,r}\mu(LD) \leqslant -(1-p)^3(1-r)^2 \\ & \{2p-4r+p^2+3r^2+2pr-2p^2r-6pr^2\}\mu(WWD) \qquad (4.9.23) \end{aligned}$$

(note that the second line in the derivation of (4.9.23) is true because the coefficient $-(1-p)(1-r)\{2p-4r+p^2+3r^2+2pr-2p^2r-6pr^2\}$ is non-positive when (4.9.13) holds and p and r are both sufficiently small);

$$
\begin{aligned}
&\bullet\, p(1-p)^2(1-r)\{\mu(WWDW)+\mu(WWDD)\} \leqslant \widehat{G}_{p,0,r}\mu(LDW)\\
\iff &-(1-p)(1-r)\{3r+2p^2+2r^2-pr-2p^2r-4pr^2-r^3\}\widehat{G}_{p,0,r}\mu(LDW) \leqslant -(1-p)(1-r)\\
&\{3r+2p^2+2r^2-pr-2p^2r-4pr^2-r^3\}p(1-p)^2(1-r)\{\mu(WWDW)+\mu(WWDD)\}\\
\iff &-(1-p)(1-r)\{3r+2p^2+2r^2-pr-2p^2r-4pr^2-r^3\}\widehat{G}_{p,0,r}\mu(LDW) \leqslant -p(1-p)^3(1-r)^2\\
&\{3r+2p^2+2r^2-pr-2p^2r-4pr^2-r^3\}\{\mu(WWDW)+\mu(WWDD)\}, \qquad (4.9.24)
\end{aligned}
$$

$$
\begin{aligned}
&\bullet\, r(1-p)^3(1-r)\mu(WWDW)+r^2(1-p)^3(1-r)\mu(WWDD) \leqslant \widehat{G}_{p,0,r}\mu(LDL)\\
\iff &-(1-p)^2(1-r)(1-r^2+pr^2)\widehat{G}_{p,0,r}\mu(LDL) \leqslant -(1-p)^2(1-r)(1-r^2+pr^2)\{r(1-p)^3(1-r)\\
&\mu(WWDW)+r^2(1-p)^3(1-r)\mu(WWDD)\}\\
\iff &-(1-p)^2(1-r)(1-r^2+pr^2)\widehat{G}_{p,0,r}\mu(LDL) \leqslant -r(1-p)^5(1-r)^2(1-r^2+pr^2)\mu(WWDW)\\
&-r^2(1-p)^5(1-r)^2(1-r^2+pr^2)\mu(WWDD); \qquad (4.9.25)
\end{aligned}
$$

and finally,

$$
\begin{aligned}
&\bullet\, p(1-p)^2(1-r)\{\mu(WWDW)+\mu(WWWD)\}+p(1-p)^2(1-r)(1+r)\mu(WWDD) \leqslant \widehat{G}_{p,0,r}\mu(LWD)\\
\iff &-(1-p)(1-r)\{2p+r+p^2+2r^2-p^2r-5pr^2\}\widehat{G}_{p,0,r}\mu(LWD) \leqslant -(1-p)(1-r)\{2p+r+p^2+2r^2\\
&-p^2r-5pr^2\}[p(1-p)^2(1-r)\{\mu(WWDW)+\mu(WWWD)\}+p(1-p)^2(1-r)(1+r)\mu(WWDD)]\\
\iff &-(1-p)(1-r)\{2p+r+p^2+2r^2-p^2r-5pr^2\}\widehat{G}_{p,0,r}\mu(LWD) \leqslant -p(1-p)^3(1-r)^2\{2p+r+p^2\\
&+2r^2-p^2r-5pr^2\}\{\mu(WWDW)+\mu(WWWD)\}-p(1-p)^3(1-r)^2(1+r)\{2p+r+p^2+2r^2\\
&-p^2r-5pr^2\}\mu(WWDD). \qquad (4.9.26)
\end{aligned}
$$

Implementing the same idea as that used to derive (4.8.6), we see that the update in (4.9.22) transforms the weight function inequality in (4.9.21) as follows, using the bounds obtained in (4.9.23), (4.9.24), (4.9.25) and (4.9.26):

$$
\begin{aligned}
w_1(\widehat{G}_{p,0,r}\mu) \leqslant\, & w_1(\mu)+(1-p)(1-r)\{-r-2p^2-3r^2-2pr+3p^2r+6pr^2-3p^2r^2\}\{\mu(WD)\\
&-\mu(WWDD)-\mu(DWDD)-\mu(WDL)\}-\{1-(1+p+pr-p^2r)(1-p)\\
&(1-r)(1+r)\}\mu(DD)-r^2(1-r)^2(1-p)^2\mu(WLD)-r^2(1-r)^2(1-p)^2\mu(DLD)\\
&+\{p-r+r^2(1-p)\}(1-p)^2(1-r)\mu(WWDW)+p(1-p)^2(1-r)^2\mu(WWWD)\\
&+(1-p)(1-r)(p-r-3p^2-3r^2-pr+2p^2r+6pr^2)\mu(WWDD)-(1-p)(1-r)
\end{aligned}
$$

$$
\begin{aligned}
&\{2p-4r+p^2+3r^2+2pr-2p^2r-6pr^2\}\widehat{G}_{p,0,r}\mu(LD)-(1-p)(1-r)\{3r+2p^2+2r^2\\
&-pr-2p^2r-4pr^2-r^3\}\widehat{G}_{p,0,r}\mu(LDW)-(1-p)^2(1-r)(1-r^2+pr^2)\widehat{G}_{p,0,r}\mu(LDL)\\
&-(1-p)(1-r)\{2p+r+p^2+2r^2-p^2r-5pr^2\}\widehat{G}_{p,0,r}\mu(LWD)\\
\leqslant\, & w_1(\mu)+(1-p)(1-r)\{-r-2p^2-3r^2-2pr+3p^2r+6pr^2-3p^2r^2\}\{\mu(WD)-\mu(WWDD)\\
&-\mu(DWDD)-\mu(WDL)\}-\{1-(1+p+pr-p^2r)(1-p)(1-r)(1+r)\}\mu(DD)-r^2(1-r)^2\\
&(1-p)^2\mu(WLD)-r^2(1-r)^2(1-p)^2\mu(DLD)+\{p-r+r^2(1-p)\}(1-p)^2(1-r)\mu(WWDW)\\
&+p(1-p)^2(1-r)^2\mu(WWWD)+(1-p)(1-r)(p-r-3p^2-3r^2-pr+2p^2r+6pr^2)\\
&\mu(WWDD)-(1-p)^3(1-r)^2\{2p-4r+p^2+3r^2+2pr-2p^2r-6pr^2\}\mu(WWD)-p(1-p)^3\\
&(1-r)^2\{3r+2p^2+2r^2-pr-2p^2r-4pr^2-r^3\}\{\mu(WWDW)+\mu(WWDD)\}-r(1-p)^5\\
&(1-r)^2(1-r^2+pr^2)\mu(WWDW)-r^2(1-p)^5(1-r)^2(1-r^2+pr^2)\mu(WWDD)-p(1-p)^3\\
&(1-r)^2\{2p+r+p^2+2r^2-p^2r-5pr^2\}\{\mu(WWDW)+\mu(WWWD)\}-p(1-p)^3(1-r)^2\\
&(1+r)\{2p+r+p^2+2r^2-p^2r-5pr^2\}\mu(WWDD). \qquad (4.9.27)
\end{aligned}
$$

We make use of the inequality $1/2\mu(WWWD)+1/2\mu(WWDD)+1/2\mu(WWDW)\leqslant\mu(WD)$, and we obtain the following:

1. the coefficient of $\mu(WWDD)$ is updated to

$$
\begin{aligned}
&(1-p)(1-r)(p-r-3p^2-3r^2-pr+2p^2r+6pr^2)-\frac{1}{2}(1-p)^3(1-r)^2\{2p-4r+p^2\\
&+3r^2+2pr-2p^2r-6pr^2\}-p(1-p)^3(1-r)^2\{3r+2p^2+2r^2-pr-2p^2r-4pr^2-r^3\}\\
&-r^2(1-p)^5(1-r)^2(1-r^2+pr^2)-p(1-p)^3(1-r)^2(1+r)\{2p+r+p^2+2r^2-p^2r-5pr^2\}\\
=\,&(1-p)(1-r)\Big[-p^5r^4(1-r)-p^5r^3-p^5r^2-p^5(3-5r)-p^4r^3-\frac{5p^4r}{2}(3-4r)-p^3r(4+12r\\
&-4r^2-r^3-10r^2)-10p^2r^5-\frac{p^2r^2}{2}(13-3r-6r^2)-pr^3(6+4r-5r^2)-r^5+\frac{7p^4}{2}+p^3\\
&+\frac{29p^2r}{2}-\frac{7p^2}{2}+20pr^2-9pr+r^4+\frac{5r^3}{2}-\frac{15r^2}{2}+r\Big]\\
\leqslant\,&\frac{1}{2}(1-p)(1-r)[2r-7p^2-15r^2-18pr+29p^2r+40pr^2+2p^3+5r^3+7p^4+2r^4]\\
=\,&\frac{1}{2}(1-p)(1-r)\Big[\Big\{\frac{9p^3}{2}+16p^2r-\frac{3p^2}{2}+\frac{43pr^2}{2}-8pr+5r^3-\frac{11r^2}{2}+2r\Big\}-\frac{11p^2}{2}-\frac{19r^2}{2}-10pr\\
&+13p^2r+\frac{37pr^2}{2}-\frac{5p^3}{2}+7p^4+2r^4\Big]\\
\leqslant\,&\frac{1}{2}(1-p)(1-r)\Big\{\frac{9p^3}{2}+16p^2r-\frac{3p^2}{2}+\frac{43pr^2}{2}-8pr+5r^3-\frac{11r^2}{2}+2r\Big\} \qquad (4.9.28)
\end{aligned}
$$

for p and r sufficiently small, and this final upper bound is non-positive because of (4.9.13);

2. the coefficient of $\mu(WWDW)$ is updated to

$$\begin{aligned}
&(1-p)^2(1-r)\{p-r+r^2(1-p)\}-\frac{1}{2}(1-p)^3(1-r)^2\{2p-4r+p^2+3r^2+2pr-2p^2r-6pr^2\}\\
&-p(1-p)^3(1-r)^2\{3r+2p^2+2r^2-pr-2p^2r-4pr^2-r^3\}-r(1-p)^5(1-r)^2(1-r^2+pr^2)\\
&-p(1-p)^3(1-r)^2\{2p+r+p^2+2r^2-p^2r-5pr^2\}\\
=&(1-p)^2(1-r)\Big[-p^4r^4-p^4r(6-3r-r^2)-p^3r^2(11-5r-4r^2)-\frac{p^3}{2}(1-5r-6p)-5p^2r^4-3p^2r^3\\
&-\frac{p^2}{2}(3-11r-8r^2)-\frac{pr^3}{2}(11-6r)-\frac{pr}{2}(6-11r)-\frac{r^2}{2}(3-5r+2r^2)\Big] \qquad (4.9.29)
\end{aligned}$$

which is evidently non-positive for p and r sufficiently small;

3. the coefficient of $\mu(WWWD)$ is updated to

$$\begin{aligned}
&p(1-p)^2(1-r)^2-\frac{1}{2}(1-p)^3(1-r)^2\{2p-4r+p^2+3r^2+2pr-2p^2r-6pr^2\}\\
&-p(1-p)^3(1-r)^2\{2p+r+p^2+2r^2-p^2r-5pr^2\}\\
=&(1-p)^2(1-r)^2\left\{2r-\frac{3p^2}{2}-\frac{3r^2}{2}-4pr+3p^2r+\frac{5pr^2}{2}+\frac{3p^3}{2}+4p^2r^2-5p^3r^2+p^4-p^4r\right\}\\
=&-p^6r^3-3p^6r(1-r)-\frac{p^5}{2}(1-2p-6r)-5p^5r^4-\frac{p^5r^2}{2}(19-24r)\\
&-p^4r^3(26-14r)-\frac{p^4}{2}(7-18r-13r^2)-\frac{21p^3r^4}{2}-p^3r(19-14r-11r^2)+\frac{9p^3}{2}-\frac{5p^2r^4}{2}\\
&-p^2r^2(30-18r)+16p^2r-\frac{3p^2}{2}-\frac{pr^3}{2}(38-11r)+\frac{43pr^2}{2}-8pr-\frac{3r^4}{2}+5r^3-\frac{11r^2}{2}+2r\\
\leqslant&\frac{9p^3}{2}+16p^2r-\frac{3p^2}{2}+\frac{43pr^2}{2}-8pr+5r^3-\frac{11r^2}{2}+2r \qquad (4.9.30)
\end{aligned}$$

which is non-positive precisely because of the inequality in (4.9.13).

A final step that we perform now is the deduction of a suitable upper bound for the coefficient of $\mu(DD)$ in (4.9.27) (using (4.8.31) with $q=0$):

$$\begin{aligned}
&-\{1-(1+p+pr-p^2r)(1-p)(1-r)(1+r)\}=-(1-p)(1-r)\{(1-p)(1-r)\}^{-1}\\
&+(1+p+pr-p^2r)(1-p)(1-r)(1+r)\\
\leqslant&-(1-p)(1-r)(1+p+r+p^2+r^2+pr-2p^2r-2pr^2+p^2r^2)\\
&+(1-p)(1-r)(1+p+r+2pr-p^2r+pr^2-p^2r^2)
\end{aligned}$$

$$= -(1-p)(1-r)(p^2+r^2-pr-p^2r-3pr^2+2p^2r^2)$$
$$= -(1-p)(1-r)\{p^2+r^2-2pr+pr(1-p-3r+2pr)\} \leqslant -(1-p)(1-r)(p-r)^2, \quad (4.9.31)$$

where the last inequality is true for p and r sufficiently small. Incorporating (4.9.28), (4.9.29), (4.9.30) and (4.9.31) into (4.9.27), we get the updated weight function inequality

$$\begin{aligned}
w_1(\widehat{G}_{p,0,r}\mu) \leqslant\ & w_1(\mu)+(1-p)(1-r)\{-r-2p^2-3r^2-2pr+3p^2r+6pr^2-3p^2r^2\}\{\mu(WD)\\
& -\mu(WWDD)-\mu(DWDD)-\mu(WDL)\}-(1-p)(1-r)(p-r)^2\mu(DD)\\
& -r^2(1-r)^2(1-p)^2\mu(WLD)-r^2(1-r)^2(1-p)^2\mu(DLD)\\
& +\left(\frac{9p^3}{2}+16p^2r-\frac{3p^2}{2}+\frac{43pr^2}{2}-8pr+5r^3-\frac{11r^2}{2}+2r\right)\mu(WWWD)\\
& +\frac{1}{2}(1-p)(1-r)\left(\frac{9p^3}{2}+16p^2r-\frac{3p^2}{2}+\frac{43pr^2}{2}-8pr+5r^3-\frac{11r^2}{2}+2r\right)\mu(WWDD)\\
& (1-p)^2(1-r)\Big[-p^4r^4-p^4r(6-3r-r^2)-p^3r^2(11-5r-4r^2)-\frac{p^3}{2}(1-5r-6p)\\
& -5p^2r^4-3p^2r^3-\frac{p^2}{2}(3-11r-8r^2)-\frac{pr^3}{2}(11-6r)-\frac{pr}{2}(6-11r)-\frac{r^2}{2}(3-5r+2r^2)\Big]\\
& \mu(WWDW).
\end{aligned}$$

Note that, at this point, we may remove all terms, except the one involving $\mu(DD)$, from the right side of the inequality above, since the coefficient of each of these terms is non-positive as long as p and r are sufficiently small and (4.9.13) holds. The final weight function inequality obtained is exactly what we have in (4.9.2) once we set p, q and r to be as in (4.9.12), and we replace w_1 by w. This concludes the construction of our weight function for the envelope PCA $\widehat{E}_{r',s'}$ (correspondingly, for $\widehat{G}_{p,q,r}$ with p, q and r are as given by (4.9.12)), when (r',s') belong to the regime in (B1).

4.9.3 Detailed construction of the weight function for the bond percolation game when (r',s') belongs to the regime given by (B2)

In the regime described by (B2), we set $s'=0$ and $r'>0$. We begin again, as in §4.8.3 and §4.9.2, with $c_1=1$ and $\mathscr{C}_1=(D)_0$, followed by $c_2=1$ and $\mathscr{C}_2=(W,D)_{0,1}$. We skip some of the details in the computation of the pushforward measures (i.e. $\widehat{E}_{r',0}\mu(D)$, $\widehat{E}_{r',0}\mu(WD)$ etc.) as these are special cases (obtained by setting $s'=0$) of what we have already computed in §4.9.2. In each computation that follows, we indicate by underbraces, wherever necessary, which terms are being combined together to proceed to the next step:

$$\widehat{E}_{r',0}\mu(D) = 2(1-r')\mu(WD) + (1+r')(1-r')\mu(DD) + 2r'(1-r')\mu(LD), \tag{4.9.32}$$

$$\begin{aligned}
\widehat{E}_{r',0}\mu(WD) = {} & \underbrace{(1-r')^2\mu(LDW) + r'(1-r')^2\mu(LDL) + (1-r')^2(1+r')\mu(LDD)}_{\text{combine using } \mu(LD)=\mu(LDW)+\mu(LDL)+\mu(LDD)} + (1-r')^2\mu(LWD) \\
& + \underbrace{r'(1-r')^2\mu(WLD) + r'(1-r')^2(1+r')\mu(LLD) + r'(1-r')^2\mu(DLD)}_{\text{combine using } \mu(LD)=\mu(WLD)+\mu(LLD)+\mu(DLD)} \qquad (4.9.33) \\
= {} & \underbrace{(1-r')^2(1+r')\mu(LD) - r'(1-r')^2\mu(LDW) - (1-r')^2\mu(LDL)}_{\text{obtained from the first combination}} + (1-r')^2\mu(LWD) \\
& + \underbrace{r'(1+r')(1-r')^2\mu(LD) - r'^2(1-r')^2\mu(WLD) - r'^2(1-r')^2\mu(DLD)}_{\text{obtained from the second combination}} \\
= {} & (1+r')^2(1-r')^2\mu(LD) + (1-r')^2\mu(LWD) - r'(1-r')^2\mu(LDW) - (1-r')^2\mu(LDL) \\
& - r'^2(1-r')^2\mu(WLD) - r'^2(1-r')^2\mu(DLD), \qquad (4.9.34)
\end{aligned}$$

and we can write $\widehat{E}_{r',0}\mu(LWD) = C_1 - C_2$, where

$$\begin{aligned}
C_1 = {} & r'(1-r')^2\mu(LDW) + r'(1-r')^2(1+r')\mu(LDD) + r'^2(1-r')^2\mu(LDL) + r'(1-r')^2\mu(LWD) \\
& + r'(1-r')^2\mu(WLD) + r'^2(1-r')^2(1+r')\mu(LLD) + r'^2(1-r')^2\mu(DLD) \qquad (4.9.35)
\end{aligned}$$

and

$$\begin{aligned}
C_2 = {} & r'(1-r')^3\mu(DLDW) + r'(1-r')^3\mu(LLDW) + r'(1-r')^3(1+r')\mu(DLDD) + r'(1-r')^3 \\
& (1+r')\mu(LLDD) + r'^2(1-r')^3\mu(DLDL) + r'^2(1-r')^3\mu(LLDL) + r'(1-r')^3\mu(LLWD) \\
& + r'(1-r')^3\mu(DLWD) + r'(1-r')^3\mu(LWLD) + r'(1-r')^3\mu(DWLD) + r'^2(1-r')^3 \\
& (1+r')\mu(LLLD) + r'^2(1-r')^3(1+r')\mu(DLLD) + r'^2(1-r')^3\mu(LDLD) + r'^2(1-r')^3\mu(DDLD).
\end{aligned} \tag{4.9.36}$$

It is worthwhile to note here that *each* term of *each* of C_1 and C_2 is non-negative. Another crucial observation is that, in $\widehat{E}_{r',0}\mu(LWD) = C_1 - C_2$, the coefficient corresponding to *each* term is of the order of r' (or smaller), so that when r' is small, each term ought to be negligible. We now set our initial choice of the weight function to be (just as in (4.8.17) of §4.8.3, as well as our initial choice in §4.9.2) $w_0(\mu) = \mu(D) + \mu(WD) + \mu(LWD)$. This yields the weight function inequality (which, in this case, is an identity):

$$w_0\left(\widehat{E}_{r',0}\mu\right) = \widehat{E}_{r',0}\mu(D) + \widehat{E}_{r',0}\mu(WD) + \widehat{E}_{r',0}\mu(LWD)$$

$$
\begin{aligned}
&= 2(1-r')\mu(WD)+(1+r')(1-r')\mu(DD)+2r'(1-r')\mu(LD)+(1+r')^2(1-r')^2\mu(LD)\\
&\quad +(1-r')^2\mu(LWD)-r'(1-r')^2\mu(LDW)-(1-r')^2\mu(LDL)-r'^2(1-r')^2\mu(WLD)\\
&\quad -r'^2(1-r')^2\mu(DLD)+r'(1-r')^2\mu(LDW)+r'(1-r')^2(1+r')\mu(LDD)+r'^2(1-r')^2\\
&\quad \mu(LDL)+r'(1-r')^2\mu(LWD)+r'(1-r')^2\mu(WLD)+r'^2(1-r')^2(1+r')\mu(LLD)\\
&\quad +r'^2(1-r')^2\mu(DLD)-C_2\\
&= \underbrace{2\mu(WD)}-2r'\mu(WD)+\underbrace{\mu(DD)}-r'^2\mu(DD)+\underbrace{\mu(LD)}+(2r'-4r'^2+r'^4)\mu(LD)\\
&\quad +\underbrace{\mu(LWD)}-(r'+r'^2-r'^3)\mu(LWD)-(1-r')^3(1+r')\mu(LDL)+r'(1-r')^3\mu(WLD)\\
&\quad +r'(1-r')^2(1+r')\mu(LDD)+r'^2(1-r')^2(1+r')\mu(LLD)-C_2\\
&= \underbrace{w_0(\mu)}_{\text{combining underbraced terms above}} -2r'\mu(WD)-r'^2\mu(DD)+\underbrace{(2r'-4r'^2+r'^4)\mu(LD)}\\
&\quad -(r'+r'^2-r'^3)\mu(LWD)-(1-r')^3(1+r')\mu(LDL)+\underbrace{r'(1-r')^3\mu(WLD)}\\
&\quad +\underbrace{r'(1-r')^2(1+r')\mu(LDD)}+\underbrace{r'^2(1-r')^2(1+r')\mu(LLD)}-C_2. \qquad (4.9.37)
\end{aligned}
$$

While of the same form as (4.8.2), the inequality in (4.9.37) does not satisfy (4.8.3): in the underbraced terms in the final expression of (4.9.37), the coefficients $r'(1-r')^3$, $r'(1-r')^2(1+r')$ and $r'^2(1-r')^2(1+r')$ are non-negative (in fact, strictly positive for $r' \in (0,1)$), and the coefficient $(2r'-4r'^2+r'^4)$ is also strictly positive for $r' \in (0, 0.53918)$. Therefore, further adjustments are necessary to our initial weight function.

Note, here, that $-2r'\mu(WD)$ is one of the non-positive terms in the right side of (4.9.37), and it is worthwhile to check whether $-2r'\widehat{E}_{r',0}\mu(WD)$ can be used to negate some of the positive terms (indicated by underbraces) in the final expression of (4.9.37). From (4.9.33), we see that $-2r'\widehat{E}_{r',0}\mu(WD)$ contributes the terms

$$
-2r'(1-r')^2\mu(LDW)-2r'^2(1-r')^2\mu(LDL)-2r'(1-r')^2(1+r')\mu(LDD),
$$

as well as the terms

$$
\begin{aligned}
&-2r'^2(1-r')^2\mu(WLD)-2r'^2(1-r')^2(1+r')\mu(LLD)-2r'^2(1-r')^2\mu(DLD)\\
=&-2r'^2(1-r')^2\mu(LD)-2r'^3(1-r')^2\mu(LLD), \qquad (4.9.38)
\end{aligned}
$$

and these can be used to negate, partially, the term $(2r'-4r'^2+r'^4)\mu(LD)$ on the right side of (4.9.37) when the coefficient $(2r'-4r'^2+r'^4)$ is positive. This serves as a good motivation for us to set the first adjustment to be $w_1(\mu) = w_0(\mu) - 2r'\mu(WD)$. This adjustment updates the weight function inequality of (4.9.37) to:

$$
\begin{aligned}
w_1\left(\widehat{E}_{r',0}\mu\right) &= w_1(\mu) + 2r'\mu(WD) - 2r'\mu(WD) - r'^2\mu(DD) + (2r' - 4r'^2 + r'^4)\mu(LD) - (r' + r'^2 \\
&\quad - r'^3)\mu(LWD) - (1-r')^3(1+r')\mu(LDL) + r'(1-r')^3\mu(WLD) + r'(1-r')^2(1+r')\mu(LDD) \\
&\quad + r'^2(1-r')^2(1+r')\mu(LLD) - C_2 - 2r(1-r')^2\mu(LDW) - 2r'^2(1-r')^2\mu(LDL) \\
&\quad - 2r'(1-r')^2(1+r')\mu(LDD) - 2r'(1-r')^2\mu(LWD) \underbrace{-2r'^2(1-r')^2\mu(WLD)}_{} \\
&\quad \underbrace{-2r'^2(1-r')^2(1+r')\mu(LLD) - 2r'^2(1-r')^2\mu(DLD)}_{} \\
&= w_1(\mu) - r'^2\mu(DD) \underbrace{+(2r' - 4r'^2 + r'^4)\mu(LD)}_{} \underbrace{-(r' + r'^2 - r'^3)\mu(LWD)}_{\text{(i)}} \\
&\quad \underbrace{-(1-r')^3(1+r')\mu(LDL)}_{\text{(ii)}} + r'(1-r')^3\mu(WLD) \underbrace{+r'(1-r')^2(1+r')\mu(LDD)}_{\text{(iii)}} \\
&\quad \underbrace{+r'^2(1-r')^2(1+r')\mu(LLD)}_{} - C_2 - 2r'(1-r')^2\mu(LDW) \underbrace{-2r'^2(1-r')^2\mu(LDL)}_{\text{(ii)}} \\
&\quad \underbrace{-2r'(1-r')^2(1+r')\mu(LDD)}_{\text{(iii)}} \underbrace{-2r'(1-r')^2\mu(LWD)}_{\text{(i)}} \\
&\quad \underbrace{-2r'^2(1-r')^2\mu(LD) - 2r'^3(1-r')^2\mu(LLD)}_{\text{combining the underbraced terms in the previous step using (4.9.38)}} \\
&= w_1(\mu) - r'^2\mu(DD) \quad \underbrace{+(2r' - 6r'^2 + 4r'^3 - r'^4)\mu(LD)}_{\text{combining underbraced terms involving } \mu(LD)} \quad \underbrace{-(3r' - 3r'^2 + r'^3)\mu(LWD)}_{\text{combining underbraced terms labeled (i)}} \\
&\quad \underbrace{-(1-r')^2(1+r'^2)\mu(LDL)}_{\text{combining underbraced terms labeled (ii)}} \quad + r'(1-r')^3\mu(WLD) \quad \underbrace{-r'(1-r')^2(1+r')\mu(LDD)}_{\text{combining underbraced terms labeled (iii)}} \\
&\quad \underbrace{+r'^2(1-r')^3\mu(LLD)}_{\text{combining underbraced terms involving } \mu(LLD)} \qquad -2r'(1-r')^2\mu(LDW) - C_2 \\
&= w_1(\mu) - r'^2\mu(DD) + (2r' - 6r'^2 + 4r'^3 - r'^4)\mu(LD) - (3r' - 3r'^2 + r'^3)\mu(LWD) \\
&\quad \underbrace{-(1-r')^2(1+r'^2)\mu(LDL) - r'(1-r')^2(1+r')\mu(LDD) - 2r'(1-r')^2\mu(LDW)}_{\text{obtain } -r'(1-r')^3\mu(LD) \text{ from these}} \\
&\quad \underbrace{+r'(1-r')^3\mu(WLD) + r'^2(1-r')^3\mu(LLD)}_{\text{obtain } r'(1-r')^3\mu(LD) \text{ from these}} - C_2 \\
&= w_1(\mu) - r'^2\mu(DD) + (2r' - 6r'^2 + 4r'^3 - r'^4)\mu(LD) - (3r' - 3r'^2 + r'^3)\mu(LWD) \\
&\quad \underbrace{-r'(1-r')^3\mu(LD)}_{} - (1-r')^2(1-r'+2r'^2)\mu(LDL) - 2r'^2(1-r')^2\mu(LDD) \\
&\quad - r'(1-r')^2(1+r')\mu(LDW) \underbrace{+r'(1-r')^3\mu(LD)}_{} - r'(1-r')^4\mu(LLD)
\end{aligned}
$$

$$
\begin{aligned}
&- r'(1-r')^3\mu(DLD) - C_2 \\
= {}& w_1(\mu) - r'^2\mu(DD) \underbrace{+(2r' - 6r'^2 + 4r'^3 - r'^4)\mu(LD)} - (3r' - 3r'^2 + r'^3)\mu(LWD) \\
&- (1-r')^2(1-r'+2r'^2)\mu(LDL) - 2r'^2(1-r')^2\mu(LDD) - r'(1-r')^2(1+r')\mu(LDW) \\
&- r'(1-r')^4\mu(LLD) - r'(1-r')^3\mu(DLD) - C_2.
\end{aligned} \tag{4.9.39}
$$

Note that (4.9.39) becomes the inequality in (4.9.4) if C_2, which consists of non-negative summands (as remarked right after (4.9.36)), is removed. The only problematic term on the right side of (4.9.39), which *may* prevent it from satisfying (4.8.3), has been indicated by an underbrace: the coefficient $(2r' - 6r'^2 + 4r'^3 - r'^4)$ of $\mu(LD)$ is strictly negative for $r' \in [0.4564, 1]$, so that if we were to consider the regime where $s' = 0$ and $r' \in [0.4564, 1]$, no further adjustments to our weight function would be required. Since, however, we *do* consider $r' < 0.4564$ in the regime described by (B2), we choose to make use of the term $-(3r' - 3r'^2 + r'^3)\mu(LWD)$, on the right side of (4.9.39), to negate, *partially*, the problematic, underbraced term $(2r' - 6r'^2 + 4r'^3 - r'^4)\mu(LD)$, following the idea outlined in §4.8.1.1.

We define our second adjustment to the weight function by setting $w_2(\mu) = w_1(\mu) - (3r' - 3r'^2 + r'^3)\mu(LWD)$, which is exactly the same as (4.9.5). Recall that $\widehat{E}_{r',0}\mu(LWD) = C_1 - C_2$, with C_1 and C_2 as defined in (4.9.35) and (4.9.36). As explained in (4.8.6), incorporating the new adjustment into (4.9.39) yields

$$
\begin{aligned}
w_2\left(\widehat{E}_{r',0}\mu\right) = {}& w_2(\mu) \underbrace{+(3r' - 3r'^2 + r'^3)\mu(LWD)} - r'^2\mu(DD) + (2r' - 6r'^2 + 4r'^3 - r'^4)\mu(LD) \\
&\underbrace{-(3r' - 3r'^2 + r'^3)\mu(LWD)} - (1-r')^2(1-r'+2r'^2)\mu(LDL) - 2r'^2(1-r')^2\mu(LDD) \\
&- r'(1-r')^2(1+r')\mu(LDW) - r'(1-r')^4\mu(LLD) - r'(1-r')^3\mu(DLD) \underbrace{-C_2}_{\text{(i)}} \\
&- (3r' - 3r'^2 + r'^3)C_1 \underbrace{+(3r' - 3r'^2 + r'^3)C_2}_{\text{(i)}} \\
= {}& w_2(\mu) - r'^2\mu(DD) + (2r' - 6r'^2 + 4r'^3 - r'^4)\mu(LD) \underbrace{-(1-r')^2(1-r'+2r'^2)\mu(LDL)}_{\text{(i)}} \\
&\underbrace{-2r'^2(1-r')^2\mu(LDD)}_{\text{(ii)}} \underbrace{-r'(1-r')^2(1+r')\mu(LDW)}_{\text{(iii)}} \underbrace{-r'(1-r')^4\mu(LLD)}_{\text{(iv)}} \underbrace{-r'(1-r')^3\mu(DLD)}_{\text{(v)}} \\
&\underbrace{-(3r' - 3r'^2 + r'^3)r'(1-r')^2\mu(LDW)}_{\text{(iii)}} \underbrace{-(3r' - 3r'^2 + r'^3)r'(1-r')^2(1+r')\mu(LDD)}_{\text{(ii)}} \\
&\underbrace{-(3r' - 3r'^2 + r'^3)r'^2(1-r')^2\mu(LDL)}_{\text{(i)}} - (3r' - 3r'^2 + r'^3)r'(1-r')^2\mu(LWD)
\end{aligned}
$$

$$
\begin{aligned}
&-(3r'-3r'^2+r'^3)r'(1-r')^2\mu(WLD)\underbrace{-(3r'-3r'^2+r'^3)r'^2(1-r')^2(1+r')\mu(LLD)}_{\text{(iv)}}\\
&\underbrace{-(3r'-3r'^2+r'^3)r'^2(1-r')^2\mu(DLD)}_{\text{(v)}}\qquad \underbrace{-(1-r')^3C_2}_{\text{combining underbraced terms labeled (i)}}\qquad \text{(using (4.9.35))}\\
=\,& w_2(\mu)-r'^2\mu(DD)+(2r'-6r'^2+4r'^3-r'^4)\mu(LD)\\
&\underbrace{-(1-r')^2(r'^5-3r'^4+3r'^3+2r'^2-r'+1)\mu(LDL)}_{\text{combining underbraced terms labeled (i)}}\underbrace{-r'^2(1-r')^2(r'^3-2r'^2+5)\mu(LDD)}_{\text{combining underbraced terms labeled (ii)}}\\
&\underbrace{-r'(1-r')^2(r'^3-3r'^2+4r'+1)\mu(LDW)}_{\text{combining underbraced terms labeled (iii)}}\underbrace{-r'(1-r')^2(r'^5-2r'^4+4r'^2-2r'+1)\mu(LLD)}_{\text{combining underbraced terms labeled (iv)}}\\
&\underbrace{-r'(1-r')^2(r'^4-3r'^3+3r'^2-r'+1)\mu(DLD)}_{\text{combining underbraced terms labeled (v)}}-(3r'-3r'^2+r'^3)r'(1-r')^2\mu(LWD)\\
&-(3r'-3r'^2+r'^3)r'(1-r')^2\mu(WLD)-(1-r')^3C_2\\
=\,& w_2(\mu)-r'^2\mu(DD)+(2r'-6r'^2+4r'^3-r'^4)\mu(LD)\\
&\underbrace{-(1-r')^2(r'^5-3r'^4+3r'^3+2r'^2-r'+1)\mu(LDL)-r'^2(1-r')^2(r'^3-2r'^2+5)\mu(LDD)}_{\text{(vi)}}\\
&\underbrace{-r'(1-r')^2(r'^3-3r'^2+4r'+1)\mu(LDW)}_{\text{(vi)}}\underbrace{-r'(1-r')^2(r'^5-2r'^4+4r'^2-2r'+1)\mu(LLD)}_{\text{(vii)}}\\
&\underbrace{-r'(1-r')^2(r'^4-3r'^3+3r'^2-r'+1)\mu(DLD)-(3r'-3r'^2+r'^3)r'(1-r')^2\mu(WLD)}_{\text{(vii)}}\\
&-(3r'-3r'^2+r'^3)r'(1-r')^2\mu(LWD)-(1-r')^3C_2. \qquad (4.9.40)
\end{aligned}
$$

Combining the terms that are grouped together using underbraces labeled (vi) in the final expression for (4.9.40), we obtain:

$$
\begin{aligned}
&-(1-r')^2(r'^5-3r'^4+3r'^3+2r'^2-r'+1)\mu(LDL)-r'^2(1-r')^2(r'^3-2r'^2+5)\mu(LDD)\\
&-r'(1-r')^2(r'^3-3r'^2+4r'+1)\mu(LDW)\\
=\,&-r'^2(1-r')^2(r'^3-2r'^2+5)\mu(LD)\\
&-\Big\{(1-r')^2(r'^5-3r'^4+3r'^3+2r'^2-r'+1)-r'^2(1-r')^2(r'^3-2r'^2+5)\Big\}\mu(LDL)\\
&-\left\{r'(1-r')^2(r'^3-3r'^2+4r'+1)-r'^2(1-r')^2(r'^3-2r'^2+5)\right\}\mu(LDW)\\
=\,&-r'^2(1-r')^2(r'^3-2r'^2+5)\mu(LD)-(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDL)\\
&-r'(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDW). \qquad (4.9.41)
\end{aligned}
$$

Combining the terms that are grouped together using underbraces labeled (vii) in the final expression for (4.9.40), we obtain:

$$
\begin{aligned}
& -r'(1-r')^2(r'^5-2r'^4+4r'^2-2r'+1)\mu(LLD)-r'(1-r')^2(r'^4-3r'^3+3r'^2-r'+1)\mu(DLD) \\
& -(3r'-3r'^2+r'^3)r'(1-r')^2\mu(WLD) \\
= & -(3r'-3r'^2+r'^3)r'(1-r')^2\mu(LD) \\
& -\left\{r'(1-r')^2(r'^5-2r'^4+4r'^2-2r'+1)-(3r'-3r'^2+r'^3)r'(1-r')^2\right\}\mu(LLD) \\
& -\left\{r'(1-r')^2(r'^4-3r'^3+3r'^2-r'+1)-(3r'-3r'^2+r'^3)r'(1-r')^2\right\}\mu(DLD) \\
= & -(3r'-3r'^2+r'^3)r'(1-r')^2\mu(LD)-r'(1-r')^2(r'^5-2r'^4-r'^3+7r'^2-5r'+1)\mu(LLD) \\
& -r'(1-r')^6\mu(DLD).
\end{aligned} \tag{4.9.42}
$$

Incorporating (4.9.41) and (4.9.42) into (4.9.40) yields:

$$
\begin{aligned}
w_2\left(\widehat{E}_{r',0}\mu\right) = {} & w_2(\mu)-r'^2\mu(DD)\underbrace{+(2r'-6r'^2+4r'^3-r'^4)\mu(LD)-r'^2(1-r')^2(r'^3-2r'^2+5)\mu(LD)} \\
& -(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDL)-r'(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDW) \\
& \underbrace{-(3r'-3r'^2+r'^3)r'(1-r')^2\mu(LD)}-r'(1-r')^2(r'^5-2r'^4-r'^3+7r'^2-5r'+1)\mu(LLD) \\
& -r'(1-r')^6\mu(DLD)-(3r'-3r'^2+r'^3)r'(1-r')^2\mu(LWD)-(1-r')^3C_2 \\
= {} & w_2(\mu)-r'^2\mu(DD)\underbrace{+(2r'-14r'^2+23r'^3-14r'^4+3r'^6-r'^7)\mu(LD)}_{\text{combining the underbraced terms above}} \\
& -(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDL)-r'(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDW) \\
& -r'(1-r')^2(r'^5-2r'^4-r'^3+7r'^2-5r'+1)\mu(LLD)-r'(1-r')^6\mu(DLD) \\
& -(3r'-3r'^2+r'^3)r'(1-r')^2\mu(LWD)-(1-r')^3C_2.
\end{aligned} \tag{4.9.43}
$$

Note that if we were to remove the term $-(1-r')^3C_2$, keeping in mind that C_2 consists of non-negative summands (as mentioned after (4.9.36)), then (4.9.43) yields the inequality in (4.9.5). Next, we note that

1. the polynomial $2r'-14r'^2+23r'^3-14r'^4+3r'^6-r'^7$ is strictly negative for $r' \in (0.201383, 1]$,
2. the polynomial $1-r'-3r'^2+3r'^3-r'^4$ is strictly positive for all $r' \in [0, 0.52779)$,
3. the polynomial $r'^5-2r'^4-r'^3+7r'^2-5r'+1$ is strictly positive for all $r' \geqslant 0$,

so that (4.9.43) satisfies the criterion in (4.8.3) as long as we have $r' \in (0.201383, 0.52779)$. However, we wish to go beyond this range of values of r', as indicated by (B2).

When $r' \leqslant 0.201382$, the only term we need be concerned about is the one indicated by an underbrace in the final expression of (4.9.43), i.e. $(2r' - 14r'^2 + 23r'^3 - 14r'^4 + 3r'^6 - r'^7)\mu(LD)$. We now motivate our approach for partially negating this term, though, as emphasized earlier to the reader, there is, truly, no *unique* way of proceeding. It is quite plausible that other, *similar* approaches yield somewhat better bounds, but intuitively, we do not see how our bounds can be *significantly* improved. It is the order of magnitude of the coefficient of $\mu(LD)$ in (4.9.43), i.e. $2r'$ (when r' is small), that prevents us from covering a regime larger than that in ((B2)), since the existing negative terms on the right side of (4.9.43) are collectively unable to negate this coefficient when r' is "too small". We explain this in the following paragraph.

On the right side of (4.9.43), we have, on one hand, negative terms involving $\mu(LDW)$ and $\mu(LDL)$, and on the other, negative terms involving $\mu(DLD)$ and $\mu(LLD)$. At a glance, what seem to be missing are negative terms involving $\mu(LDD)$ and $\mu(WLD)$ (had there been a couple of negative terms involving $\mu(LDD)$ and $\mu(WLD)$, with coefficients whose order of magnitude is r', for r' is small, on the right side of (4.9.43), these could have helped take care of $(2r' - 14r'^2 + 23r'^3 - 14r'^4 + 3r'^6 - r'^7)\mu(LD)$, since $\mu(LDW) + \mu(LDL) + \mu(LDD) = \mu(DLD) + \mu(LLD) + \mu(WLD) = \mu(LD)$). Next, among the terms of C_2 in (4.9.36), we have

1. the terms $2r'(1-r')^3\mu(DWLD)$ and $r'(1-r')^3\mu(LWLD)$, but no term involving $\mu(WWLD)$, which is why we cannot directly obtain from $-(1-r')^3C_2$ a negative term involving $\mu(WLD)$;

2. the terms $r'(1-r')^3(1+r')\mu(DLDD)$ and $r'(1-r')^3(1+r')\mu(LLDD)$, but no term involving $\mu(WLDD)$, which is why we cannot directly obtain from $-(1-r')^3C_2$ a negative term involving $\mu(LDD)$.

We shall, therefore, avail the aid of $-r'^2(3-3r'+r'^2)(1-r')^2\mu(LWD)$ and $-r'^2\mu(DD)$ that are present in the right side of (4.9.43), and the terms $-r'(1-r')^6\mu(LLWD)$ and $-r'(1-r')^6\mu(LLDW)$ from the expansion of $-(1-r')^3C_2$, to define the third adjustment to our weight function. It can be checked that the remaining terms of C_2, namely those involving $\mu(DLDW)$, $\mu(DLDL)$, $\mu(DDLD)$, $\mu(DDLD)$, $\mu(LLDL)$, $\mu(LDLD)$ and $\mu(LLLD)$, are either of no use in generating (via the probabilities of the corresponding cylinder sets induced by the pushforward measure $\widehat{E}_{r',0}\mu$) the terms that are missing above, or generate the missing terms with coefficients that are of an order of magnitude far too small to be of any significant use to us.

We set the new, adjusted weight function to be

$$w_3(\mu) = w_2(\mu) - r'(1-r')^6\{\mu(LLWD)+\mu(LLDW)\} - r'^2\mu(WDD) - (3r'-3r'^2+r'^3)r'(1-r')^2\mu(LWD). \tag{4.9.44}$$

To understand how (4.9.44) impacts the weight function inequality in (4.9.43), we need to compute, *partially*, the pushforward measures $\widehat{E}_{r',0}\mu(LLWD)$, $\widehat{E}_{r',0}\mu(LLDW)$ and $\widehat{E}_{r',0}\mu(WDD)$. Recall that $\widehat{E}_{r',0}\mu(LLWD)$ equals the probability of the event $\{\widehat{E}_{r',0}\eta(0) = \widehat{E}_{r',0}\eta(1) = L, \widehat{E}_{r',0}\eta(2) = W, \widehat{E}_{r',0}\eta(3) = D\}$ when η is a random configuration with law μ, and we consider the contributions to $\widehat{E}_{r',0}\mu(LLWD)$ arising from the following events *only*:

1. the contribution of the event $\{\eta(0) = \eta(1) = W, \eta(2) = L, \eta(3) = D\}$ to $\widehat{E}_{r',0}\mu(LLWD)$ is

$$\begin{aligned} &r(1-r')^2\mu(WWLDW) + r'(1-r')^2(1+r')\mu(WWLDD) + r'^2(1-r')^2\mu(WWLDL) \\ &= r'(1-r')^2\mu(WWLD) + r'^2(1-r')^2\mu(WWLDD) - r'(1-r')^3\mu(WWLDL); \end{aligned} \tag{4.9.45}$$

2. the contribution of the event $\{\eta(1) = \eta(2) = W, \eta(3) = L, \eta(4) = D\}$ to $\widehat{E}_{r',0}\mu(LLWD)$ is

$$\begin{aligned} &r'(1-r')^2\mu(WWWLD) + r'^2(1-r')^2\mu(LWWLD) + r'^2(1-r')^2\mu(DWWLD) \\ &= r'^2(1-r')^2\mu(WWLD) + r'(1-r')^3\mu(WWWLD); \end{aligned} \tag{4.9.46}$$

3. the contribution of the event $\{\eta(0) \in \{D, L\}, \eta(1) = W, \eta(2) = L, \eta(3) = \eta(4) = D\}$ to $\widehat{E}_{r',0}\mu(LLWD)$ is

$$r'^2(1-r')^2(1+r')\mu(DWLDD) + r'^2(1-r')^2(1+r')\mu(LWLDD). \tag{4.9.47}$$

Combining (4.9.45), (4.9.46) and (4.9.47), and letting C'_{LLWD} denote the contribution from all cases *not* considered above, we obtain:

$$\begin{aligned} \widehat{E}_{r',0}\mu(LLWD) = {} & \underbrace{r'(1-r')^2\mu(WWLD)}_{\text{(i)}} + \underbrace{r'^2(1-r')^2\mu(WWLDD)}_{\text{(ii)}} - r'(1-r')^3\mu(WWLDL) \\ & + \underbrace{r'^2(1-r')^2\mu(WWLD)}_{\text{(i)}} + r'(1-r')^3\mu(WWWLD) + \underbrace{r'^2(1-r')^2(1+r')\mu(DWLDD)}_{\text{(ii)}} \\ & + \underbrace{r'^2(1-r')^2(1+r')\mu(LWLDD)}_{\text{(ii)}} + C'_{LLWD} \end{aligned}$$

$$
\begin{aligned}
&= \underbrace{r'(1-r')^2(1+r')\mu(WWLD)}_{\text{summing underbraced terms labeled (i)}} - r'(1-r')^3\mu(WWLDL) \\
&\quad + \underbrace{r'^2(1-r')^2(1+r')\mu(WLDD) - r'^3(1-r')^2\mu(WWLDD)}_{\text{summing underbraced terms labeled (ii)}} + C_{LLWD}, \quad (4.9.48)
\end{aligned}
$$

where we set $C_{LLWD} = r'(1-r')^3\mu(WWWLD) + C'_{LLWD}$. Next, for computing $\widehat{E}_{r',0}\mu(LLDW)$, which is the probability of the event $\{\eta(0) = \eta(1) = L, \eta(2) = D, \eta(3) = W\}$ where η follows the distribution μ, we consider only the contribution arising from the event $\{\eta(0) = \eta(1) = W, \eta(2) = L, \eta(3) = D, \eta(4) = L\}$, allowing us to write

$$
\widehat{E}_{r',0}\mu(LLDW) = r'^2(1-r')^2\mu(WWLDL) + C_{LLDW}, \tag{4.9.49}
$$

where C_{LLDW} denotes the contribution from all the cases *not* considered above. For the computation of $\widehat{E}_{r',0}\mu(WDD)$, which is the probability of the event $\{\widehat{E}_{r',0}\eta(0) = W, \widehat{E}_{r',0}\eta(1) = \widehat{E}_{r',0}\eta(2) = D\}$ where η follows the distribution μ, we consider

1. the contribution from the event $\{\eta(0) = L, \eta(1) = \eta(2) = D\}$, given by

$$
\begin{aligned}
&(1-r')^3(1+r)\mu(LDDW) + (1-r')^3(1+r')^2\mu(LDDD) + r'(1-r')^3(1+r')\mu(LDDL) \\
&= r'(1-r')^3(1+r')\mu(LDD) + (1-r')^4(1+r')\mu(LDDW) + (1-r')^3(1+r')\mu(LDDD);
\end{aligned} \tag{4.9.50}
$$

2. and the contribution from the event $\{\eta(1) = L, \eta(2) = \eta(3) = D\}$, given by

$$
\begin{aligned}
&r'(1-r')^3(1+r')\mu(WLDD) + r'(1-r')^3(1+r')\mu(DLDD) + r'(1-r')^3(1+r')^2\mu(LLDD) \\
&= r'(1-r')^3(1+r')\mu(LDD) + r'^2(1-r')^3(1+r')\mu(LLDD).
\end{aligned} \tag{4.9.51}
$$

Combining (4.9.50) and (4.9.51), we obtain

$$
\begin{aligned}
\widehat{E}_{r',0}\mu(WDD) &= r'(1-r')^3(1+r')\mu(LDD) + (1-r')^4(1+r')\mu(LDDW) + (1-r')^3(1+r')\mu(LDDD) \\
&\quad + r'(1-r')^3(1+r')\mu(LDD) + r'^2(1-r')^3(1+r')\mu(LLDD) + C'_{WDD} \\
&= 2r'(1-r')^3(1+r')\mu(LDD) + C_{WDD},
\end{aligned} \tag{4.9.52}
$$

where C'_{WDD} is the contribution to $\widehat{E}_{r',0}\mu(WDD)$ of all the cases *not* considered above, and we set $C_{WDD} = C'_{WDD} + (1-r')^4(1+r')\mu(LDDW) + (1-r')^3(1+r')\mu(LDDD) + r'^2(1-r')^3(1+$

$r')\mu(LLDD)$. Finally, we write

$$\widehat{E}_{r',0}\mu(LWD) = r'(1-r')^2(1+r')\mu(WLDD) + r'(1-r')^2\mu(WWLD) + C_{LWD}, \tag{4.9.53}$$

where C_{LWD} captures the rest of the expression for $\widehat{E}_{r',0}\mu(LWD)$ that we have already explicitly computed earlier. It is important to keep in mind here that each term of C_{LWD} is non-negative.

From (4.9.43), (4.9.44), (4.9.48), (4.9.49), (4.9.52) and (4.9.53), setting $C_3 = (1-r')^3C_2 - r'(1-r')^6\mu(LLWD) - r'(1-r')^6\mu(LLDW) - 2r'(1-r')^6\mu(DWLD) - r'(1-r')^6\mu(LWLD) - r'(1-r')^6(1+r')\mu(DLDD) - r'(1-r')^6(1+r')\mu(LLDD)$, and implementing the same idea as that used to derive (4.8.6), we obtain:

$$\begin{aligned}
w_3\left(\widehat{E}_{r',0}\mu\right) = {} & w_3(\mu) \underbrace{+r'(1-r')^6\mu(LLWD) + r'(1-r')^6\mu(LLDW)}_{\text{(i)}} \underbrace{+r'^2\mu(WDD)}_{\text{(ii)}} \\
& \underbrace{+(3r'-3r'^2+r'^3)r'(1-r')^2\mu(LWD)}_{\text{(iii)}} \underbrace{-r'^2\mu(DD)}_{\text{(ii)}} + (2r' - 14r'^2 + 23r'^3 - 14r'^4 \\
& + 3r'^6 - r'^7)\mu(LD) - (1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDL) - r'(1-r')^2 \\
& (1-r'-3r'^2+3r'^3-r'^4)\mu(LDW) - r'(1-r')^2(r'^5-2r'^4-r'^3+7r'^2-5r'+1)\mu(LLD) \\
& - r'(1-r')^6\mu(DLD) \underbrace{-(3r'-3r'^2+r'^3)r'(1-r')^2\mu(LWD)}_{\text{(iii)}} \underbrace{-r'(1-r')^6\mu(LLWD)}_{\text{(i)}} \\
& \underbrace{-r'(1-r')^6\mu(LLDW)}_{\text{(i)}} - 2r'(1-r')^6\mu(DWLD) - r'(1-r')^6\mu(LWLD) - r'(1-r')^6 \\
& (1+r')\mu(DLDD) - r'(1-r')^6(1+r')\mu(LLDD) - C_3 - r'(1-r')^6\widehat{E}_{r',0}\mu(LLWD) \\
& - r'(1-r')^6\widehat{E}_{r',0}\mu(LLDW) - r'^2\widehat{E}_{r',0}\mu(WDD) - (3r'-3r'^2+r'^3)r'(1-r')^2\widehat{E}_{r',0}\mu(LWD) \\
= {} & w_3(\mu) + (2r' - 14r'^2 + 23r'^3 - 14r'^4 + 3r'^6 - r'^7)\mu(LD) \quad \underbrace{-r'^2\mu(LDD) - r'^2\mu(DDD)}_{\text{combining underbraced terms labeled (ii)}} \\
& \underbrace{-(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDL) - r'(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDW)}_{\text{(iv)}} \\
& \underbrace{-r'(1-r')^2(r'^5-2r'^4-r'^3+7r'^2-5r'+1)\mu(LLD) - r'(1-r')^6\mu(DLD)}_{\text{(v)}} \\
& \underbrace{-2r'(1-r')^6\mu(DWLD) - r'(1-r')^6\mu(LWLD)}_{\text{(v)}} \underbrace{-r'(1-r')^6(1+r')\mu(DLDD)}_{\text{(iv)}} \\
& \underbrace{-r'(1-r')^6(1+r')\mu(LLDD)}_{\text{(iv)}} \underbrace{-r'^2(1-r')^8(1+r')\mu(WWLD)}_{\text{(v)}} \underbrace{+r'^2(1-r')^9\mu(WWLDL)}_{\text{(iv)}}
\end{aligned}$$

$$\underbrace{-r'^3(1-r')^8(1+r')\mu(WLDD)}_{\text{(iv)}}+r'^4(1-r')^8\mu(WWLDD)-r'(1-r')^6C_{LLWD}$$
$$\underbrace{-r'^3(1-r')^8\mu(WWLDL)}_{\text{(iv)}}-r'(1-r')^6C_{LLDW}\underbrace{-2r'^3(1-r')^3(1+r')\mu(LDD)}_{\text{(iv)}}-r'^2C_{WDD}$$
$$\underbrace{-r'^3(3-3r'+r'^2)(1-r')^4(1+r')\mu(WLDD)}_{\text{(iv)}}\underbrace{-r'^3(3-3r'+r'^2)(1-r')^4\mu(WWLD)}_{\text{(v)}}$$
$$-r'^2(3-3r'+r'^2)(1-r')^2C_{LWD}-C_3, \tag{4.9.54}$$

where the last step is obtained by adding terms grouped by underbraces labeled (i) and (iii) in the previous step, and substituting from (4.9.48), (4.9.49), (4.9.52) and (4.9.53). Adding the terms indicated by underbraces labeled (iv) in the final expression of (4.9.54), we obtain

$$-(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDL)-r'(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDW)$$
$$\underbrace{-r'(1-r')^6(1+r')\mu(DLDD)-r'(1-r')^6(1+r')\mu(LLDD)-r'^3(1-r')^8(1+r')\mu(WLDD)}_{\text{(a)}}$$
$$-2r'^3(1-r')^3(1+r')\mu(LDD)\underbrace{-r'^3(3-3r'+r'^2)(1-r')^4(1+r')\mu(WLDD)}_{\text{(a)}}\underbrace{+r'^2(1-r')^9\mu(WWLDL)}_{\text{(b)}}$$
$$\underbrace{-r'^3(1-r')^8\mu(WWLDL)}_{\text{(b)}}$$
$$=-(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDL)-r'(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDW)$$
$$\underbrace{-r'^3(1-r')^4(1+r')(4-7r'+7r'^2-4r'^3+r'^4)\mu(LDD)}_{\text{combining underbraced terms labeled (a)}}$$
$$\underbrace{-r'(1-r')^4(1+r')(1-2r'-3r'^2+7r'^3-7r'^4+4r'^5-r'^6)\{\mu(DLDD)+\mu(LLDD)\}}_{\text{combining underbraced terms labeled (a)}}$$
$$-2r'^3(1-r')^3(1+r')\mu(LDD)\underbrace{+r'^2(1-r')^8(1-2r')\mu(WWLDL)}_{\text{combining underbraced terms labeled (b)}}$$
$$=-(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDL)-r'(1-r')^2(1-r'-3r'^2+3r'^3-r'^4)\mu(LDW)$$
$$\underbrace{-r'^3(1-r')^3(1+r')(2-r')(3-4r'+5r'^2-3r'^3+r'^4)\mu(LDD)}_{\text{summing the two terms involving } \mu(LDD) \text{ above}}$$
$$-r'(1-r')^4(1+r')(1-2r'-3r'^2+7r'^3-7r'^4+4r'^5-r'^6)\{\mu(DLDD)+\mu(LLDD)\}$$
$$+r'^2(1-r')^8(1-2r')\mu(WWLDL)$$
$$=\underbrace{-r'^3(1-r')^3(1+r')(2-r')(3-4r'+5r'^2-3r'^3+r'^4)\mu(LD)}$$

$$\underbrace{-(1-r')^2(1-r'-3r'^2-3r'^3+10r'^4-8r'^5+9r'^7-10r'^8+5r'^9-r'^{10})\mu(LDL)}$$
$$\underbrace{-r'(1-r')^2(1-r'-9r'^2+14r'^3-9r'^4+9r'^6-10r'^7+5r'^8-r'^9)\mu(LDW)}$$
$$-r'(1-r')^4(1+r')(1-2r'-3r'^2+7r'^3-7r'^4+4r'^5-r'^6)\{\mu(DLDD)+\mu(LLDD)\}$$
$$+r'^2(1-r')^8(1-2r')\mu(WWLDL)$$

(where the underbraced terms are obtained by combining the terms involving $\mu(LDL), \mu(LDW)$ and $\mu(LDD)$ from the previous step)

$$= -r'^3(1-r')^3(1+r')(2-r')(3-4r'+5r'^2-3r'^3+r'^4)\mu(LD)-\alpha_1\mu(LDL)-\alpha_2\mu(LDW)$$
$$-\alpha_3\{\mu(DLDD)+\mu(LLDD)\}+r'^2(1-r')^8(1-2r')\mu(WWLDL), \qquad (4.9.55)$$

where

1. the polynomial $\alpha_1=(1-r')^2(1-r'-3r'^2-3r'^3+10r'^4-8r'^5+9r'^7-10r'^8+5r'^9-r'^{10})$ is strictly positive for all $r'\in[0,0.435029)$,

2. the polynomial $\alpha_2=r'(1-r')^2(1-r'-9r'^2+14r'^3-9r'^4+9r'^6-10r'^7+5r'^8-r'^9)$ is strictly positive for all $r'\in[0,0.35678)$,

3. and the polynomial $\alpha_3=r'(1-r')^4(1+r')(1-2r'-3r'^2+7r'^3-7r'^4+4r'^5-r'^6)$ is strictly positive for all $r'\in[0,0.410819)$.

It is thus evident that, when we consider $0<r'\leqslant 0.201382$, each $\alpha_i>0$. Next, combining the terms indicated by underbraces labeled (v) in (4.9.54), we obtain:

$$-r'(1-r')^2(r'^5-2r'^4-r'^3+7r'^2-5r'+1)\mu(LLD)-r'(1-r')^6\mu(DLD)\underbrace{-2r'(1-r')^6\mu(DWLD)}_{\text{(c)}}$$
$$\underbrace{-r'(1-r')^6\mu(LWLD)-r'^2(1-r')^8(1+r')\mu(WWLD)-r'^3(3-3r'+r'^2)(1-r')^4\mu(WWLD)}_{\text{(c)}}$$
$$= -r'(1-r')^2(r'^5-2r'^4-r'^3+7r'^2-5r'+1)\mu(LLD)-r'(1-r')^6\mu(DLD)$$
$$\underbrace{-r'^2(1-r')^4(1-r'^2+3r'^3-3r'^4+r'^5)\mu(WLD)}_{\text{combining underbraced terms labeled (c)}}$$
$$\underbrace{-r'(1-r')^4(2-5r'+2r'^2+r'^3-3r'^4+3r'^5-r'^6)\mu(DWLD)}_{\text{combining underbraced terms labeled (c)}}$$
$$\underbrace{-r'(1-r')^4(1-3r'+r'^2+r'^3-3r'^4+3r'^5-r'^6)\mu(LWLD)}_{\text{combining underbraced terms labeled (c)}}$$

$$
\begin{aligned}
&= \underbrace{-r'(1-r')^2(r'^5-2r'^4-r'^3+7r'^2-5r'+1)\mu(LLD)-r'(1-r')^6\mu(DLD)}_{\text{(d)}} \\
&\quad \underbrace{-r'^2(1-r')^4(1-r'^2+3r'^3-3r'^4+r'^5)\mu(WLD)}_{\text{(d)}}-\beta_1\mu(DWLD)-\beta_2\mu(LWLD) \\
&= \underbrace{-r'^2(1-r')^4(1-r'^2+3r'^3-3r'^4+r'^5)\mu(LD)}_{\text{combining underbraced terms labeled (d)}} \\
&\quad \underbrace{-r'(1-r')^2(1-6r'+9r'^2-r'^3-7r'^4+11r'^5-10r'^6+5r'^7-r'^8)\mu(LLD)}_{\text{combining underbraced terms labeled (d)}} \\
&\quad \underbrace{-r'(1-r')^4(1-3r'+r'^2+r'^3-3r'^4+3r'^5-r'^6)\mu(DLD)}_{\text{combining underbraced terms labeled (d)}}-\beta_1\mu(DWLD)-\beta_2\mu(LWLD) \\
&= -r'^2(1-r')^4(1-r'^2+3r'^3-3r'^4+r'^5)\mu(LD)-\beta_3\mu(LLD)-\beta_2\{\mu(DLD)+\mu(LWLD)\} \\
&\quad -\beta_1\mu(DWLD),
\end{aligned}
\tag{4.9.56}
$$

in which

1. the polynomial $\beta_1 = r'(1-r')^4(2-5r'+2r'^2+r'^3-3r'^4+3r'^5-r'^6)$ is strictly positive for all $r' \in (0, 0.505225)$,
2. the polynomial $\beta_2 = r'(1-r')^4(1-3r'+r'^2+r'^3-3r'^4+3r'^5-r'^6)$ is strictly positive for all $r' \in (0, 0.387969)$,
3. the polynomial $\beta_3 = r'(1-r')^2(1-6r'+9r'^2-r'^3-7r'^4+11r'^5-10r'^6+5r'^7-r'^8)$ is strictly positive for all $r' \in (0, 0.265137)$.

Therefore, for $0 < r' \leqslant 0.201382$, each $\beta_i > 0$. Incorporating (4.9.55) and (4.9.56) into (4.9.54), we obtain

$$
\begin{aligned}
w_3(\widehat{E}_{r',0}\mu) &= w_3(\mu)+(2r'-14r'^2+23r'^3-14r'^4+3r'^6-r'^7)\mu(LD)\underbrace{-r'^2\mu(LDD)}-r'^2\mu(DDD) \\
&\quad -r'^3(1-r')^3(1+r')(2-r')(3-4r'+5r'^2-3r'^3+r'^4)\mu(LD)\underbrace{-\alpha_1\mu(LDL)} \\
&\quad \underbrace{-\alpha_2\mu(LDW)}-\alpha_3\{\mu(DLDD)+\mu(LLDD)\}+r'^2(1-r')^8(1-2r')\mu(WWLDL) \\
&\quad -r'^2(1-r')^4(1-r'^2+3r'^3-3r'^4+r'^5)\mu(LD)-\beta_3\mu(LLD)-\beta_2\{\mu(DLD)+\mu(LWLD)\} \\
&\quad -\beta_1\mu(DWLD)+r'^4(1-r')^8\mu(WWLDD)-r'(1-r')^6C_{LLWD}-r'(1-r')^6C_{LLDW} \\
&\quad -r'^2C_{WDD}-r'^2(3-3r'+r'^2)(1-r')^2C_{LWD}-C_3 \\
&= w_3(\mu)+(2r'-14r'^2+23r'^3-14r'^4+3r'^6-r'^7)\mu(LD)\underbrace{-r'^2\mu(LD)}-r'^2\mu(DDD)
\end{aligned}
$$

$$
\begin{aligned}
&-r'^3(1-r')^3(1+r')(2-r')(3-4r'+5r'^2-3r'^3+r'^4)\mu(LD)\underbrace{-(1-3r'-r'^2+2r'^3}\\
&\underbrace{+13r'^4-31r'^5+26r'^6+r'^7-28r'^8+34r'^9-21r'^{10}+7r'^{11}-r'^{12})\mu(LDL)}\\
&\underbrace{-r'(1-4r'-6r'^2+31r'^3-46r'^4+32r'^5-28r'^7+34r'^8-21r'^9+7r'^{10}-r'^{11})\mu(LDW)}\\
&-\alpha_3\{\mu(DLDD)+\mu(LLDD)\}+r'^2(1-r')^8(1-2r')\mu(WWLDL)\\
&-r'^2(1-r')^4(1-r'^2+3r'^3-3r'^4+r'^5)\mu(LD)-\beta_3\mu(LLD)-\beta_2\{\mu(DLD)+\mu(LWLD)\}\\
&-\beta_1\mu(DWLD)+r'^4(1-r')^8\mu(WWLDD)-r'(1-r')^6C_{LLWD}-r'(1-r')^6C_{LLDW}-r'^2C_{WDD}\\
&-r'^2(3-3r'+r'^2)(1-r')^2C_{LWD}-C_3\\
&\text{(the underbraced terms in the step above are obtained by combining the underbraced terms}\\
&\text{from the previous step)}\\
=\;& w_3(\mu)\underbrace{+(2r'-14r'^2+23r'^3-14r'^4+3r'^6-r'^7)\mu(LD)-r'^2\mu(LD)}-r'^2\mu(DDD)\\
&\underbrace{-r'^3(1-r')^3(1+r')(2-r')(3-4r'+5r'^2-3r'^3+r'^4)\mu(LD)}-\gamma_1\mu(LDL)-\gamma_2\mu(LDW)\\
&-\alpha_3\{\mu(DLDD)+\mu(LLDD)\}\underbrace{-r'^2(1-r')^4(1-r'^2+3r'^3-3r'^4+r'^5)\mu(LD)}\\
&-\beta_3\mu(LLD)-\beta_2\mu(DLD)-\beta_1\mu(DWLD)-\beta_2\mu(LWLD)+r'^4(1-r')^8\mu(WWLDD)\\
&+r'^2(1-r')^8(1-2r')\mu(WWLDL)-r'(1-r')^6C_{LLWD}-r'(1-r')^6C_{LLDW}-r'^2C_{WDD}\\
&-r'^2(3-3r'+r'^2)(1-r')^2C_{LWD}-C_3\\
=\;& w_3(\mu)\underbrace{+r'(2-16r'+21r'^2+4r'^3-39r'^4+50r'^5-35r'^6+7r'^7+13r'^8-14r'^9+6r'^{10}}\\
&\underbrace{-r'^{11})\mu(LD)}-r'^2\mu(DDD)-\gamma_1\mu(LDL)-\gamma_2\mu(LDW)-\alpha_3\{\mu(DLDD)+\mu(LLDD)\}\\
&-\beta_3\mu(LLD)-\beta_2\{\mu(DLD)+\mu(LWLD)\}-\beta_1\mu(DWLD)\underbrace{+r'^4(1-r')^8\mu(WWLDD)}\\
&\underbrace{+r'^2(1-r')^8(1-2r')\mu(WWLDL)}-r'(1-r')^6C_{LLWD}-r'(1-r')^6C_{LLDW}-r'^2C_{WDD}\\
&-r'^2(3-3r'+r'^2)(1-r')^2C_{LWD}-C_3\\
&\text{(once again, the underbraced terms in the step above are obtained by combining the}\\
&\text{underbraced terms from the previous step)} \qquad (4.9.57)
\end{aligned}
$$

where

1. the polynomial $\gamma_1=(1-3r'-r'^2+2r'^3+13r'^4-31r'^5+26r'^6+r'^7-28r'^8+34r'^9-21r'^{10}+7r'^{11}-r'^{12})$ is strictly positive for all $r'\in[0,0.345627)$,

2. the polynomial $\gamma_2=r'(1-4r'-6r'^2+31r'^3-46r'^4+32r'^5-28r'^7+34r'^8-21r'^9+7r'^{10}-$

r'^{11}) is strictly positive for all $r' \in (0, 0.238556)$,

3. and the polynomial $r'(2-16r'+21r'^2+4r'^3-39r'^4+50r'^5-35r'^6+7r'^7+13r'^8-14r'^9+6r'^{10}-r'^{11})$ is strictly negative for all $r' > 0.157175$.

Among the problematic terms, indicated by underbraces, on the right side of the final expression of (4.9.57), there is $r'^4(1-r')^8\mu(WWLDD)$, which we now take care of by introducing the adjustment $w_4(\mu) = w_3(\mu) - \beta_2\mu(LWLD)$. Note that $w_4(\mu)$, when expanded (by substituting the expression for $w_3(\mu)$), becomes exactly the weight function defined in (4.9.7).

We partially compute $\widehat{E}_{r',0}\mu(LWLD)$, which is the probability of the event $\{\widehat{E}_{r',0}\eta(0) = L, \widehat{E}_{r',0}\eta(1) = W, \widehat{E}_{r',0}\eta(2) = L, \widehat{E}_{r',0}\eta(3) = D\}$:

1. Suppose $\eta(0) = \eta(1) = W$, $\eta(2) = L$ and $\eta(3) = \eta(4) = D$. Then the event $\{\widehat{E}_{r',0}\eta(0) = L\}$ happens with probability 1, the event $\{\widehat{E}_{r',0}\eta(1) = W\}$ happens with probability $(1-r')$, the event $\{\widehat{E}_{r',0}\eta(2) = L\}$ happens with probability r'^2, and the event $\{\widehat{E}_{r',0}\eta(3) = D\}$ happens with probability $(1-r')(1+r')$. Thus, the contribution of this case to $\widehat{E}_{r',0}\mu(LWLD)$ is $r'^2(1-r')^2(1+r')\mu(WWLDD)$.

Therefore, letting C_{LWLD} denote the contribution from the cases not considered above, we obtain

$$\widehat{E}_{r',0}\mu(LWLD) = r'^2(1-r')^2(1+r')\mu(WWWLD) + C_{LWLD}. \tag{4.9.58}$$

Incorporating the new adjustment into (4.9.57), and emulating (4.8.6), we obtain:

$$\begin{aligned}
w_4\left(\widehat{E}_{r',0}\mu\right) &= w_4(\mu) + \beta_2\mu(LWLD) + r'(2-16r'+21r'^2+4r'^3-39r'^4+50r'^5-35r'^6+7r'^7+13r'^8 \\
&\quad -14r'^9+6r'^{10}-r'^{11})\mu(LD) - r'^2\mu(DDD) - \gamma_1\mu(LDL) - \gamma_2\mu(LDW) - \alpha_3\{\mu(DLDD) \\
&\quad +\mu(LLDD)\} - \beta_3\mu(LLD) - \beta_2\mu(DLD) - \beta_1\mu(DWLD) - \beta_2\mu(LWLD) \\
&\quad \underbrace{+r'^4(1-r')^8\mu(WWLDD)} + r'^2(1-r')^8(1-2r')\mu(WWLDL) - r'(1-r')^6C_{LLWD} \\
&\quad -r'(1-r')^6C_{LLDW} - r'^2C_{WDD} - r'^2(3-3r'+r'^2)(1-r')^2C_{LWD} - C_3 \\
&\quad \underbrace{-\beta_2r'^2(1-r')^2(1+r')\mu(WWLDD)} - \beta_2C_{LWLD} \\
&= w_4(\mu) + r'(2-16r'+21r'^2+4r'^3-39r'^4+50r'^5-35r'^6+7r'^7+13r'^8-14r'^9+6r'^{10}-r'^{11}) \\
&\quad \mu(LD) - r'^2\mu(DDD) - \gamma_1\mu(LDL) - \gamma_2\mu(LDW) - \alpha_3\{\mu(DLDD)+\mu(LLDD)\} - \beta_3\mu(LLD) \\
&\quad -\beta_2\mu(DLD) - \beta_1\mu(DWLD) \underbrace{-r'^3(1-r')^6(1-3r'+r'^3-2r'^4+2r'^6-r'^7)\mu(WWLDD)} \\
&\quad +r'^2(1-r')^8(1-2r')\mu(WWLDL) - r'(1-r')^6C_{LLWD} - r'(1-r')^6C_{LLDW} - r'^2C_{WDD} \\
&\quad -r'^2(3-3r'+r'^2)(1-r')^2C_{LWD} - C_3 - \beta_2C_{LWLD}
\end{aligned}$$

$$
\begin{aligned}
&= w_4(\mu) + r'(2 - 16r' + 21r'^2 + 4r'^3 - 39r'^4 + 50r'^5 - 35r'^6 + 7r'^7 + 13r'^8 - 14r'^9 + 6r'^{10} - r'^{11}) \\
&\quad \mu(LD) - r'^2\mu(DDD) \underbrace{-\gamma_1\mu(LDL)} - \gamma_2\mu(LDW) - \alpha_3\{\mu(DLDD) + \mu(LLDD)\} - \beta_3\mu(LLD) \\
&\quad - \beta_2\mu(DLD) - \beta_1\mu(DWLD) - \gamma_3\mu(WWLDD) \underbrace{+ r'^2(1-r')^8(1-2r')\mu(WWLDL)}_{\text{use } \mu(WWLDL) \leqslant \mu(LDL) \text{ in the next step}} \\
&\quad - r'(1-r')^6 C_{LLWD} - r'(1-r')^6 C_{LLDW} - r'^2 C_{WDD} - r'^2(3 - 3r' + r'^2)(1-r')^2 C_{LWD} - C_3 \\
&\quad - \beta_2 C_{LWLD} \\
&\leqslant w_4(\mu) + r'(2 - 16r' + 21r'^2 + 4r'^3 - 39r'^4 + 50r'^5 - 35r'^6 + 7r'^7 + 13r'^8 - 14r'^9 + 6r'^{10} \\
&\quad - r'^{11})\mu(LD) - r'^2\mu(DDD) \underbrace{-(1 - 3r' - 2r'^2 + 12r'^3 - 31r'^4 + 81r'^5 - 156r'^6 + 197r'^7} \\
&\quad \underbrace{-168r'^8 + 98r'^9 - 38r'^{10} + 9r'^{11} - r'^{12})\mu(LDL)} - \gamma_2\mu(LDW) - \alpha_3\{\mu(DLDD) \\
&\quad + \mu(LLDD)\} - \beta_3\mu(LLD) - \beta_2\mu(DLD) - \beta_1\mu(DWLD) - \gamma_3\mu(WWLDD) - r'(1-r')^6 C_{LLWD} \\
&\quad - r'(1-r')^6 C_{LLDW} - r'^2 C_{WDD} - r'^2(3 - 3r' + r'^2)(1-r')^2 C_{LWD} - C_3 - \beta_2 C_{LWLD} \\
&= w_4(\mu) + r'(2 - 16r' + 21r'^2 + 4r'^3 - 39r'^4 + 50r'^5 - 35r'^6 + 7r'^7 + 13r'^8 - 14r'^9 + 6r'^{10} - r'^{11}) \\
&\quad \mu(LD) - r'^2\mu(DDD) - \gamma_4\mu(LDL) - \gamma_2\mu(LDW) - \alpha_3\{\mu(DLDD) + \mu(LLDD)\} - \beta_3\mu(LLD) \\
&\quad - \beta_2\mu(DLD) - \beta_1\mu(DWLD) - \gamma_3\mu(WWLDD) - r'(1-r')^6 C_{LLWD} - r'(1-r')^6 C_{LLDW} \\
&\quad - r'^2 C_{WDD} - r'^2(3 - 3r' + r'^2)(1-r')^2 C_{LWD} - C_3 - \beta_2 C_{LWLD}, \qquad (4.9.59)
\end{aligned}
$$

where

1. the polynomial $\gamma_4 = 1 - 3r' - 2r'^2 + 12r'^3 - 31r'^4 + 81r'^5 - 156r'^6 + 197r'^7 - 168r'^8 + 98r'^9 - 38r'^{10} + 9r'^{11} - r'^{12}$ is strictly positive for all $r' \in [0, 0.345094)$,

2. and the polynomial $\gamma_3 = r'^3(1-r')^6(1 - 3r' + r'^3 - 2r'^4 + 2r'^6 - r'^7)$ is strictly positive for all $r' \in (0, 0.338338)$.

We now have our weight function inequality for $0.157175 < r' \leqslant 0.201382$, since the coefficient of each term in the right side of (4.9.59), apart from $w_4(\mu)$, is strictly negative for $r' \in (0.157175, 0.201382]$, thereby allowing (4.9.59) to satisfy (4.8.3). Note that (4.9.59) reduces to the inequality in (4.9.8) if we remove, apart from $w_4(\mu)$, $-r'^2\mu(DDD)$ and $r'(2 - 16r' + 20r'^2 + 7r'^3 - 42r'^4 + 51r'^5 - 35r'^6 + 7r'^7 + 13r'^8 - 14r'^9 + 6r'^{10} - r'^{11})\mu(LD)$, all other terms from the right side of (4.9.59). This brings us to the end of our construction of the weight functions for the regime described in (B2).

4.9.4 Detailed construction of the weight function for the bond percolation game when (r', s') belongs to the regime given by (B3)

We now demonstrate the step-by-step construction of our weight function for the regime given by (B3), i.e. for all (r', s') that satisfy the constraints $r' = s' > 0.10883$. The approach is the same the general idea outlined in §4.8.1, and the steps, while differing from those presented in §4.9.3 in their details, follow a similar line of reasoning. As in each of §4.9.2 and §4.9.3, we begin by setting $c_1 = 1$, $\mathscr{C}_1 = (D)_0$, $c_2 = 1$ and $\mathscr{C}_2 = (W, D)_{0,1}$ in (4.8.1). For $r' = s'$, we obtain, using reflection-invariance,

$$\begin{aligned}\widehat{E}_{r',r'}\mu(D) &= (1-r')(1-2r')\{\mu(WD)+\mu(DW)\} + (1-2r')\mu(DD) + r'(1-2r')\{\mu(LD)+\mu(DL)\}\\ &= 2(1-r')(1-2r')\mu(WD) + (1-2r')\mu(DD) + 2r'(1-2r')\mu(LD) \qquad (4.9.60)\end{aligned}$$

and

$$\begin{aligned}\widehat{E}_{r',r'}\mu(WD) = &(2r'-r'^2)(1-r')(1-2r')\mu(WWD) + (1-r'+r'^2)r'(1-2r')\mu(WLD)\\ &+ (1-r'+r'^2)(1-r')(1-2r')\mu(LWD) + (1-r'^2)r'(1-2r')\mu(LLD)\\ &+ (2r'-r'^2)(1-2r')\mu(WDD) + (2r'-r'^2)(1-r')(1-2r')\mu(DWD)\\ &+ (2r'-r'^2)(1-2r')\mu(DDD) + (1-r'+r'^2)(1-2r')\mu(LDD)\\ &+ (1-r'+r'^2)r'(1-2r')\mu(DLD) + (2r'-r'^2)(1-r')(1-2r')\mu(WDW)\\ &+ (2r'-r'^2)r'(1-2r')\mu(WDL) + (1-r'+r'^2)(1-r')(1-2r')\mu(LDW)\\ &+ (1-r'+r'^2)r'(1-2r')\mu(LDL) + (2r'-r'^2)(1-r')(1-2r')\mu(DDW)\\ &+ (2r'-r'^2)r'(1-2r')\mu(DDL). \qquad (4.9.61)\end{aligned}$$

Starting with our initial guess for the weight function, namely $w_0(\mu) = \mu(D) + \mu(WD)$, we obtain, using (4.9.60) and (4.9.61), the following weight function inequality:

$$\begin{aligned}w_0(\widehat{E}_{r',r'}\mu) &= w_0(\mu) \underbrace{-\mu(D) + \widehat{E}_{r',r'}\mu(D)}_{\text{use (4.9.60) and simplify}} - \mu(WD) + \widehat{E}_{r',r'}\mu(WD)\\ &= w_0(\mu) - (3r'-2r'^2)\mu(DW) - 2r'\mu(DD) - (1-r'+2r'^2)\mu(DL) + (1-r')(1-2r')\mu(WD)\\ &\quad + r'(1-2r')\mu(LD) - \mu(WD) \quad \underbrace{+\widehat{E}_{r',r'}\mu(WD)}_{\text{substitute using (4.9.61)}}\\ &= w_0(\mu) \underbrace{-(3r'-2r'^2)\mu(DW)}_{\text{substitute } \mu(DW)=\mu(WD)} - 2r'\mu(DD) \underbrace{-(1-r'+2r'^2)\mu(DL)}_{\text{(i)}} \underbrace{+(1-r')(1-2r')\mu(WD)}_{\text{(ii)}}\end{aligned}$$

$$\underbrace{+r'(1-2r')\mu(LD)}_{\text{(i)}}\underbrace{-\mu(WD)}_{\text{(ii)}}+(2r'-r'^2)(1-r')(1-2r')\mu(WWD)$$
$$+(1-r'+r'^2)r'(1-2r')\mu(WLD)+(1-r'+r'^2)(1-r')(1-2r')\mu(LWD)$$
$$+(1-r'^2)r'(1-2r')\mu(LLD)+(2r'-r'^2)(1-2r')\mu(WDD)+(2r'-r'^2)(1-r')$$
$$(1-2r')\mu(DWD)+(2r'-r'^2)(1-2r')\mu(DDD)+(1-r'+r'^2)(1-2r')\mu(LDD)$$
$$+(1-r'+r'^2)r'(1-2r')\mu(DLD)+(2r'-r'^2)(1-r')(1-2r')\mu(WDW)+(2r'-r'^2)$$
$$r'(1-2r')\mu(WDL)+(1-r'+r'^2)(1-r')(1-2r')\mu(LDW)+(1-r'+r'^2)r'(1-2r')\mu(LDL)$$
$$+(2r'-r'^2)(1-r')(1-2r')\mu(DDW)+(2r'-r'^2)r'(1-2r')\mu(DDL)$$
$$=w_0(\mu)\underbrace{-(3r'-2r'^2)\mu(WD)}-2r'\mu(DD)\underbrace{-(1-2r'+4r'^2)\mu(LD)}_{\text{combining terms labelled (i)}}\underbrace{-(3r'-2r'^2)\mu(WD)}_{\text{combining terms labelled (ii)}}$$
$$+(2r'-r'^2)(1-r')(1-2r')\mu(WWD)+(1-r'+r'^2)r'(1-2r')\mu(WLD)$$
$$+(1-r'+r'^2)(1-r')(1-2r')\mu(LWD)+(1-r'^2)r'(1-2r')\mu(LLD)$$
$$+(2r'-r'^2)(1-2r')\mu(WDD)+(2r'-r'^2)(1-r')(1-2r')\mu(DWD)$$
$$+(2r'-r'^2)(1-2r')\mu(DDD)+(1-r'+r'^2)(1-2r')\mu(LDD)+(1-r'+r'^2)$$
$$r'(1-2r')\mu(DLD)+(2r'-r'^2)(1-r')(1-2r')\mu(WDW)+(2r'-r'^2)r'(1-2r')\mu(WDL)$$
$$+(1-r'+r'^2)(1-r')(1-2r')\mu(LDW)+(1-r'+r'^2)r'(1-2r')\mu(LDL)$$
$$+(2r'-r'^2)(1-r')(1-2r')\mu(DDW)+(2r'-r'^2)r'(1-2r')\mu(DDL)$$
$$=w_0(\mu)\quad\underbrace{-(2r'-r'^2)(1-2r')\mu(WD)}_{\text{expand using }\mu(WD)=\mu(WDW)+\mu(WDL)+\mu(WDD)}\quad\underbrace{-2r'\mu(DD)}_{\text{expand using }\mu(DD)=\mu(DDW)+\mu(DDL)+\mu(DDD)}$$
$$\underbrace{-(1-r'+r'^2)(1-2r')\mu(LD)}_{\text{expand using }\mu(LD)=\mu(LDW)+\mu(LDL)+\mu(LDD)}\quad\underbrace{-((1-2r'+4r'^2)-(1-r'+r'^2)(1-2r'))\mu(LD)}_{\text{expand using }\mu(LD)=\mu(WLD)+\mu(LLD)+\mu(DLD)}$$
$$\underbrace{-(6r'-4r'^2-(2r'-r'^2)(1-2r'))\mu(WD)}_{\text{expand using }\mu(WD)=\mu(WWD)+\mu(LWD)+\mu(DWD)}+(2r'-r'^2)(1-r')(1-2r')\mu(WWD)$$
$$+(1-r'+r'^2)r'(1-2r')\mu(WLD)+(1-r'+r'^2)(1-r')(1-2r')\mu(LWD)$$
$$+(1-r'^2)r'(1-2r')\mu(LLD)+(2r'-r'^2)(1-2r')\mu(WDD)$$
$$+(2r'-r'^2)(1-r')(1-2r')\mu(DWD)+(2r'-r'^2)(1-2r')\mu(DDD)$$
$$+(1-r'+r'^2)(1-2r')\mu(LDD)+(1-r'+r'^2)r'(1-2r')\mu(DLD)$$
$$+(2r'-r'^2)(1-r')(1-2r')\mu(WDW)+(2r'-r'^2)r'(1-2r')\mu(WDL)$$
$$+(1-r'+r'^2)(1-r')(1-2r')\mu(LDW)+(1-r'+r'^2)r'(1-2r')\mu(LDL)$$
$$+(2r'-r'^2)(1-r')(1-2r')\mu(DDW)+(2r'-r'^2)r'(1-2r')\mu(DDL)$$
$$=w_0(\mu)\underbrace{-(2r'-r'^2)(1-2r')(\mu(WDW)+\mu(WDL)+\mu(WDD))}$$

$$
\begin{aligned}
&\underbrace{-2r'(\mu(DDW)+\mu(DDL)+\mu(DDD))}\underbrace{-(1-r'+r'^2)(1-2r')(\mu(LDW)+\mu(LDL)}\\
&\underbrace{+\mu(LDD))-((1-2r'+4r'^2)-(1-r'+r'^2)(1-2r'))(\mu(WLD)+\mu(LLD)+\mu(DLD))}\\
&\underbrace{-(6r'-4r'^2-(2r'-r'^2)(1-2r'))(\mu(WWD)+\mu(LWD)+\mu(DWD))}\\
&+(2r'-r'^2)(1-r')(1-2r')\mu(WWD)+(1-r'+r'^2)r'(1-2r')\mu(WLD)\\
&+(1-r'+r'^2)(1-r')(1-2r')\mu(LWD)+(1-r'^2)r'(1-2r')\mu(LLD)\\
&+(2r'-r'^2)(1-2r')\mu(WDD)+(2r'-r'^2)(1-r')(1-2r')\mu(DWD)+(2r'-r'^2)\\
&(1-2r')\mu(DDD)+(1-r'+r'^2)(1-2r')\mu(LDD)+(1-r'+r'^2)r'(1-2r')\mu(DLD)\\
&+(2r'-r'^2)(1-r')(1-2r')\mu(WDW)+(2r'-r'^2)r'(1-2r')\mu(WDL)\\
&+(1-r'+r'^2)(1-r')(1-2r')\mu(LDW)+(1-r'+r'^2)r'(1-2r')\mu(LDL)\\
&+(2r'-r'^2)(1-r')(1-2r')\mu(DDW)+(2r'-r'^2)r'(1-2r')\mu(DDL) \qquad (4.9.62)\\
={}& w_0(\mu)-(2r'^2-5r'^3+2r'^4)\mu(WDW)\underbrace{-(2r'-7r'^2+7r'^3-2r'^4)\mu(WDL)}\\
&-(7r'^2-7r'^3+2r'^4)\mu(DDW)-(2r'-2r'^2+5r'^3-2r'^4)\mu(DDL)-(5r'^2-2r'^3)\mu(DDD)\\
&\underbrace{-(r'-3r'^2+3r'^3-2r'^4)\mu(LDW)}_{\text{substituting } \mu(LDW)=\mu(WDL)}-(1-4r'+6r'^2-5r'^3+2r'^4)\mu(LDL)\\
&-(4r'^2-r'^3+2r'^4)\mu(WLD)-(3r'^2+3r'^3-2r'^4)\mu(LLD)-(4r'^2-r'^3+2r'^4)\mu(DLD)\\
&-(2r'+8r'^2-9r'^3+2r'^4)\mu(WWD)-(-1+8r'-5r'^2+3r'^3-2r'^4)\mu(LWD)\\
&-(2r'+8r'^2-9r'^3+2r'^4)\mu(DWD) \qquad (4.9.63)\\
={}& w_0(\mu)-(2r'^2-5r'^3+2r'^4)\mu(WDW)\underbrace{-(3r'-10r'^2+10r'^3-4r'^4)\mu(WDL)}_{\text{combining the highlighted terms}}\\
&-(7r'^2-7r'^3+2r'^4)\mu(DDW)-(2r'-2r'^2+5r'^3-2r'^4)\mu(DDL)-(5r'^2-2r'^3)\mu(DDD)\\
&-(1-4r'+6r'^2-5r'^3+2r'^4)\mu(LDL)-(4r'^2-r'^3+2r'^4)\mu(WLD)-(3r'^2+3r'^3\\
&-2r'^4)\mu(LLD)-(4r'^2-r'^3+2r'^4)\mu(DLD)-(2r'+8r'^2-9r'^3+2r'^4)\mu(WWD)\\
&-(-1+8r'-5r'^2+3r'^3-2r'^4)\mu(LWD)-(2r'+8r'^2-9r'^3+2r'^4)\mu(DWD).\\
&\qquad (4.9.64)
\end{aligned}
$$

The only problematic term in the above expression is $-(-1+8r'-5r'^2+3r'^3-2r'^4)\mu(LWD)$, since the coefficient of this term is strictly positive for sufficiently small values of r'. In order to take care of this term, we implement the idea outlined in §4.8.1.1 by updating our weight function as follows:

$$w_1(\mu)=w_0(\mu)-(2r'^2-5r'^3+2r'^4)\mu(WDW)-(3r'-10r'^2+10r'^3-4r'^4)\mu(WDL)$$

$$
\begin{aligned}
&-(7r'^2-7r'^3+2r'^4)\mu(DDW)-(2r'-2r'^2+5r'^3-2r'^4)\mu(DDL)\\
&-(1-4r'+6r'^2-5r'^3+2r'^4)\mu(LDL)-(4r'^2-r'^3+2r'^4)\mu(WLD)-(3r'^2+3r'^3-2r'^4)\mu(LLD)\\
&-(2r'+8r'^2-9r'^3+2r'^4)\mu(WWD),
\end{aligned}
\tag{4.9.65}
$$

which is exactly what we stated in (4.9.9).

In order to incorporate this adjustment into our current weight function inequality, following the idea used in the deduction of (4.8.6), we need to compute the pushforward probabilities $\widehat{E}_{r',r'}\mu(WDW)$, $\widehat{E}_{r',r'}\mu(WDL)$, $\widehat{E}_{r',r'}\mu(DDW)$, $\widehat{E}_{r',r'}\mu(DDL)$, $\widehat{E}_{r',r'}\mu(LDL)$, $\widehat{E}_{r',r'}\mu(WLD)$, $\widehat{E}_{r',r'}\mu(LLD)$ and $\widehat{E}_{r',r'}\mu(WWD)$. However, we need only *partially* compute these expressions, since we only care about the contributions arising from cylinder sets that are able to negate, to as large an extent as possible, the term $-(-1+8r'-5r'^2+3r'^3-2r'^4)\mu(LWD)$. For instance, $\widehat{E}_{r',r'}\mu(WDW)$ is the probability of the event $\{\widehat{E}_{r',r'}\eta(0)=W,\widehat{E}_{r',r'}\eta(1)=D,\widehat{E}_{r',r'}\eta(2)=W\}$ where η is a random configuration with law μ, and we need only consider the contributions to this pushforward probability that arise from the event $\{\eta(1)=D,\eta(2)=W,\eta(3)=L\}$. In fact, this is precisely the event whose contributions we consider to each of the pushforward probabilities mentioned above, and we thereby obtain the following lower bounds:

$$
\begin{aligned}
\widehat{E}_{r',r'}\mu(WDW) &\geqslant (1-r'+r'^2)(1-r')(1-2r')\{(2r'-r'^2)\mu(WDWL)+(1-r'+r'^2)\mu(LDWL)\\
&\qquad +(2r'-r'^2)\mu(DDWL)\}\\
&\geqslant (1-r'+r'^2)(1-r')(1-2r')(2r'-r'^2)\mu(DWL)\\
&= \underbrace{(1-r'+r'^2)(1-r')(1-2r')(2r'-r'^2)\mu(LWD)}_{\text{putting } \mu(DWL)=\mu(LWD)\text{, using reflection-invariance}};
\end{aligned}
\tag{4.9.66}
$$

$$
\begin{aligned}
\widehat{E}_{r',r'}\mu(WDL) \geqslant&(1-r')(1-2r')(r'-r'^2)\{(2r'-r'^2)\mu(WDWL)+(1-r'+r'^2)\mu(LDWL)\\
&+(2r'-r'^2)\mu(DDWL)\}\geqslant(1-r')(1-2r')(r'-r'^2)(2r'-r'^2)\mu(LWD);
\end{aligned}
\tag{4.9.67}
$$

$$
\begin{aligned}
\widehat{E}_{r',r'}\mu(DDW) \geqslant& (1-r'+r'^2)(1-r')(1-2r')\{(1-r')(1-2r')\mu(WDWL)+(r'-2r'^2)\mu(LDWL)\\
&+(1-2r')\mu(DDWL)\}\geqslant(1-r'+r'^2)(1-r')(1-2r')(r'-2r'^2)\mu(LWD);
\end{aligned}
\tag{4.9.68}
$$

$$
\widehat{E}_{r',r'}\mu(DDL) \geqslant (r'-r'^2)(1-r')(1-2r')\{(1-r')(1-2r')\mu(WDWL)+(r'-2r'^2)\mu(LDWL)
$$

$$+(1-2r')\mu(DDWL)\} \geqslant (r'-r'^2)(1-r')(1-2r')(r'-2r'^2)\mu(LWD); \tag{4.9.69}$$

$$\begin{aligned}\widehat{E}_{r',r'}\mu(LDL) &\geqslant (r'-r'^2)(1-r')(1-2r')\{(r'-r'^2)\mu(WDWL)+r'^2\mu(LDWL)+r'^2\mu(DDWL)\}\\ &\geqslant (r'-r'^2)(1-r')(1-2r')r'^2\mu(LWD);\end{aligned} \tag{4.9.70}$$

$$\begin{aligned}\widehat{E}_{r',r'}\mu(WLD) &\geqslant (1-r')(1-2r')(r'-r'^2)\{(1-r'+r'^2)\mu(WLWD)+(1-r'^2)\mu(LLWD)\\ &\quad+(1-r'+r'^2)\mu(DLWD)\} \geqslant (1-r')(1-2r')(r'-r'^2)(1-r'+r'^2)\mu(LWD);\end{aligned} \tag{4.9.71}$$

$$\begin{aligned}\widehat{E}_{r',r'}\mu(WWD) &\geqslant (1-r')(1-2r')(1-r'+r'^2)\{(1-r'+r'^2)\mu(WLWD)+(1-r'^2)\mu(LLWD)\\ &\quad+(1-r'+r'^2)\mu(DLWD)\} \geqslant (1-r')(1-2r')(1-r'+r'^2)^2\mu(LWD);\end{aligned} \tag{4.9.72}$$

and finally,

$$\begin{aligned}\widehat{E}_{r',r'}\mu(LLD) &\geqslant (1-r')(1-2r')(r'-r'^2)\{(r'-r'^2)\mu(WLWD)+r'^2\mu(LLWD)+r'^2\mu(DLWD)\}\\ &\geqslant (1-r')(1-2r')(r'-r'^2)r'^2\mu(LWD).\end{aligned} \tag{4.9.73}$$

Incorporating the adjustment introduced in (4.9.65) into the weight function inequality in (4.9.64) via the idea demonstrated in (4.8.6), and using the lower bounds obtained in (4.9.66), (4.9.67), (4.9.68), (4.9.69), (4.9.70), (4.9.71), (4.9.72), (4.9.73), we obtain

$$\begin{aligned}w_1(\widehat{E}_{r',r'}\mu) = & w_1(\mu)\underbrace{+(w_0(\widehat{E}_{r',r'}\mu)-w_0(\mu))}_{\text{substitute using (4.9.64)}}-(2r'^2-5r'^3+2r'^4)\widehat{E}_{r',r'}\mu(WDW)\\ &-(3r'-10r'^2+10r'^3-4r'^4)\widehat{E}_{r',r'}\mu(WDL)-(7r'^2-7r'^3+2r'^4)\widehat{E}_{r',r'}\mu(DDW)\\ &-(2r'-2r'^2+5r'^3-2r'^4)\widehat{E}_{r',r'}\mu(DDL)-(1-4r'+6r'^2-5r'^3+2r'^4)\widehat{E}_{r',r'}\mu(LDL)\\ &-(4r'^2-r'^3+2r'^4)\widehat{E}_{r',r'}\mu(WLD)-(3r'^2+3r'^3-2r'^4)\widehat{E}_{r',r'}\mu(LLD)\\ &-(2r'+8r'^2-9r'^3+2r'^4)\widehat{E}_{r',r'}\mu(WWD)+(2r'^2-5r'^3+2r'^4)\mu(WDW)\\ &+(3r'-10r'^2+10r'^3-4r'^4)\mu(WDL)+(7r'^2-7r'^3+2r'^4)\mu(DDW)\\ &+(2r'-2r'^2+5r'^3-2r'^4)\mu(DDL)+(1-4r'+6r'^2-5r'^3+2r'^4)\mu(LDL)\\ &+(4r'^2-r'^3+2r'^4)\mu(WLD)+(3r'^2+3r'^3-2r'^4)\mu(LLD)\end{aligned}$$

$$
\begin{aligned}
&+(2r'+8r'^2-9r'^3+2r'^4)\mu(WWD)\\
=&w_1(\mu)\underbrace{-(2r'^2-5r'^3+2r'^4)\mu(WDW)}_{\text{(i)}}\underbrace{-(3r'-10r'^2+10r'^3-4r'^4)\mu(WDL)}_{\text{(ii)}}\\
&\underbrace{-(7r'^2-7r'^3+2r'^4)\mu(DDW)}_{\text{(iii)}}\underbrace{-(2r'-2r'^2+5r'^3-2r'^4)\mu(DDL)}_{\text{(iv)}}\\
&-(5r'^2-2r'^3)\mu(DDD)\underbrace{-(1-4r'+6r'^2-5r'^3+2r'^4)\mu(LDL)}_{\text{(v)}}\\
&\underbrace{-(4r'^2-r'^3+2r'^4)\mu(WLD)}_{\text{(vi)}}\underbrace{-(3r'^2+3r'^3-2r'^4)\mu(LLD)}_{\text{(vii)}}\\
&-(4r'^2-r'^3+2r'^4)\mu(DLD)\underbrace{-(2r'+8r'^2-9r'^3+2r'^4)\mu(WWD)}_{\text{(viii)}}\\
&-(-1+8r'-5r'^2+3r'^3-2r'^4)\mu(LWD)-(2r'+8r'^2-9r'^3+2r'^4)\mu(DWD)\\
&-(2r'^2-5r'^3+2r'^4)\widehat{E}_{r',r'}\mu(WDW)-(3r'-10r'^2+10r'^3-4r'^4)\widehat{E}_{r',r'}\mu(WDL)\\
&-(7r'^2-7r'^3+2r'^4)\widehat{E}_{r',r'}\mu(DDW)-(2r'-2r'^2+5r'^3-2r'^4)\widehat{E}_{r',r'}\mu(DDL)\\
&-(1-4r'+6r'^2-5r'^3+2r'^4)\widehat{E}_{r',r'}\mu(LDL)-(4r'^2-r'^3+2r'^4)\widehat{E}_{r',r'}\mu(WLD)\\
&-(3r'^2+3r'^3-2r'^4)\widehat{E}_{r',r'}\mu(LLD)-(2r'+8r'^2-9r'^3+2r'^4)\widehat{E}_{r',r'}\mu(WWD)\\
&\underbrace{+(2r'^2-5r'^3+2r'^4)\mu(WDW)}_{\text{(i)}}\underbrace{+(3r'-10r'^2+10r'^3-4r'^4)\mu(WDL)}_{\text{(ii)}}\\
&\underbrace{+(7r'^2-7r'^3+2r'^4)\mu(DDW)}_{\text{(iii)}}\underbrace{+(2r'-2r'^2+5r'^3-2r'^4)\mu(DDL)}_{\text{(iv)}}\\
&\underbrace{+(1-4r'+6r'^2-5r'^3+2r'^4)\mu(LDL)}_{\text{(v)}}\underbrace{+(4r'^2-r'^3+2r'^4)\mu(WLD)}_{\text{(vi)}}\\
&\underbrace{+(3r'^2+3r'^3-2r'^4)\mu(LLD)}_{\text{(vii)}}\underbrace{+(2r'+8r'^2-9r'^3+2r'^4)\mu(WWD)}_{\text{(viii)}}\\
=&w_1(\mu)-(5r'^2-2r'^3)\mu(DDD)-(4r'^2-r'^3+2r'^4)\mu(DLD)\\
&-(-1+8r'-5r'^2+3r'^3-2r'^4)\mu(LWD)-(2r'+8r'^2-9r'^3+2r'^4)\mu(DWD)\\
&\underbrace{-(2r'^2-5r'^3+2r'^4)\widehat{E}_{r',r'}\mu(WDW)}_{\text{use (4.9.66)}}\underbrace{-(3r'-10r'^2+10r'^3-4r'^4)\widehat{E}_{r',r'}\mu(WDL)}_{\text{use (4.9.67)}}\\
&\underbrace{-(7r'^2-7r'^3+2r'^4)\widehat{E}_{r',r'}\mu(DDW)}_{\text{use (4.9.68)}}\underbrace{-(2r'-2r'^2+5r'^3-2r'^4)\widehat{E}_{r',r'}\mu(DDL)}_{\text{use (4.9.69)}}\\
&\underbrace{-(1-4r'+6r'^2-5r'^3+2r'^4)\widehat{E}_{r',r'}\mu(LDL)}_{\text{use (4.9.70)}}\underbrace{-(4r'^2-r'^3+2r'^4)\widehat{E}_{r',r'}\mu(WLD)}_{\text{use (4.9.71)}}
\end{aligned}
$$

$$\underbrace{-(3r'^2+3r'^3-2r'^4)\widehat{E}_{r',r'}\mu(LLD)}_{\text{use (4.9.73)}}\underbrace{-(2r'+8r'^2-9r'^3+2r'^4)\widehat{E}_{r',r'}\mu(WWD)}_{\text{use (4.9.72)}}$$

$$\begin{aligned}
\leqslant & w_1(\mu)-(5r'^2-2r'^3)\mu(DDD)-(4r'^2-r'^3+2r'^4)\mu(DLD) \\
& \underbrace{-(-1+8r'-5r'^2+3r'^3-2r'^4)\mu(LWD)}-(2r'+8r'^2-9r'^3+2r'^4)\mu(DWD) \\
& \underbrace{-(2r'^2-5r'^3+2r'^4)((1-r'+r'^2)(1-r')(1-2r')(2r'-r'^2)\mu(LWD))} \\
& \underbrace{-(3r'-10r'^2+10r'^3-4r'^4)((1-r')(1-2r')(r'-r'^2)(2r'-r'^2)\mu(LWD))} \\
& \underbrace{-(7r'^2-7r'^3+2r'^4)((1-r'+r'^2)(1-r')(1-2r')(r'-2r'^2)\mu(LWD))} \\
& \underbrace{-(2r'-2r'^2+5r'^3-2r'^4)((r'-r'^2)(1-r')(1-2r')(r'-2r'^2)\mu(LWD))} \\
& \underbrace{-(1-4r'+6r'^2-5r'^3+2r'^4)((r'-r'^2)(1-r')(1-2r')r'^2\mu(LWD))} \\
& \underbrace{-(4r'^2-r'^3+2r'^4)((1-r')(1-2r')(r'-r'^2)(1-r'+r'^2)\mu(LWD))} \\
& \underbrace{-(3r'^2+3r'^3-2r'^4)((1-r')(1-2r')(r'-r'^2)r'^2\mu(LWD))} \\
& \underbrace{-(2r'+8r'^2-9r'^3+2r'^4)((1-r')(1-2r')(1-r'+r'^2)^2\mu(LWD))} \\
= & w_1(\mu)-(5r'^2-2r'^3)\mu(DDD)-(4r'^2-r'^3+2r'^4)\mu(DLD)-(2r'+8r'^2-9r'^3+2r'^4)\mu(DWD) \\
& \underbrace{+(1-10r'+7r'^2+64r'^4-292r'^5+583r'^6-663r'^7+447r'^8-168r'^9+28r'^{10})\mu(LWD)}_{\text{obtained by combining the terms highlighted above}} \\
\leqslant & w_1(\mu)+(1-10r'+7r'^2+64r'^4-292r'^5+583r'^6-663r'^7+447r'^8-168r'^9+28r'^{10})\mu(LWD),
\end{aligned}$$

which is exactly what we stated in (4.9.10). This brings us to the conclusion of the construction of our weight function for (r',s') belonging to ((B3)),

Chapter 5

The parity-based site percolation game and its corresponding probabilistic cellular automata

Preface

This chapter is based on the following paper:

- Bhasin D., Karmarkar S., Podder M., Roy S. "Ergodicity of a generalized probabilistic cellular automaton with parity-based neighbourhoods" – submitted. Preprint available **here**.

5.1 Introduction

In the previous twochapters, we considered two special percolation games whose general model was given in §1.1. One of the common features in the games studied in §3 and §4 is that the set $\text{Out}(x,y)$ was constructed in a *homogeneous* manner. In this chapter, we consider the *parity-based site percolation* game, that is, a site percolation game for which the definition of the set $\text{Out}(x,y)$ depends on the parity of x. In particular, each vertex of $\mathbb{Z}^2$, independently of others, is labeled a trap with probability p, a target with probability q and open with probability $1-p-q$. In her turn, a player is allowed the token from the vertex (x,y) to any of the vertices in $\{(x,y+1),(x+1,y+1)\}$ if x is even and to any of the vertices in $\{(x+1,y+1),(x+2,y+2)\}$ if x is odd. While the method of analyzing this game is similar to that of the previous two chapters, the notion of parity brings an added complexity to weight function calculations. Additionally, the parity-dependent definitions of $\text{Out}(x,y)$ naturally motivate us to consider a broader version of PCAs, where the neighbourhood set of $\mathbf{x} \in \mathbb{Z}$ depend on the parity of $\mathbf{x}$. We call such PCAs, heterogeneous PCAs (hePCAs, in short).

This element of *inhomogeneity* (i.e. the scenario where the manner in which a vertex of $\mathbb{Z}^d$ updates its state differs from one vertex to another) is incorporated into a hePCA via the introduction of *multiple* neighbourhood sets. The most general way of accomplishing this is to define, for each $\mathbf{x} \in \mathbb{Z}^d$, the (finite) neighbourhood set $\mathscr{N}^{\mathbf{x}} = \{\mathbf{y}_1^{\mathbf{x}}, \mathbf{y}_2^{\mathbf{x}}, \dots, \mathbf{y}_m^{\mathbf{x}}\} \subset \mathbb{Z}^d$, allowing for the possibility that $\mathscr{N}^{\mathbf{x}}$ is different for different values of $\mathbf{x}$ (even the value of m, which is the cardinality of $\mathscr{N}^{\mathbf{x}}$, may depend on $\mathbf{x}$). The neighbours of the cell $\mathbf{x}$ are then the elements of the set $\mathbf{x} + \mathscr{N}^{\mathbf{x}}$, i.e. $\mathbf{x} + \mathbf{y}_i^{\mathbf{x}}$ for each $i \in \{1,2,\dots,m\}$. A hePCA updates a state $\eta \in \Omega = \mathscr{A}^{\mathbb{Z}^d}$ in an analogous way as PCA does (see §1.1.1), i.e., by using a stochastic matrix. In particular, attached to each tuple $(\eta(\mathbf{x}+\mathbf{y}_1^{\mathbf{x}}), \eta(\mathbf{x}+\mathbf{y}_2^{\mathbf{x}}), \dots, \eta(\mathbf{x}+\mathbf{y}_m^{\mathbf{x}}))$, is a probability distribution supported on the alphabet $\mathscr{A}$ which depicts the (random) state of the site $\mathbf{x}$ after the application of the hePCA to η.

The one-dimensional hePCA $E_{p,q}$ that we investigate in this chapter has the underlying alphabet set $=\{W,L\}$ and has two sets $\mathscr{N}_0 = \{0,1\}$ and $\mathscr{N}_1 = \{1,2\}$ associated to it. These sets are used to determine the collection $\mathscr{N}^x$ for each $x \in \mathbb{Z}$ in the following way: if x is even then $\mathscr{N}^x = \mathscr{N}_0$ and if x is odd then $\mathscr{N}^x = \mathscr{N}_1$, and using these sets, the neighbours of each cell are decided as described earlier. The parameters p and q are elements of the following set:

$$(p,q) \in \mathscr{S} = \{(p',q') \in [0,1]^2 : 0 < p'+q' \leqslant 1\},$$

and the update rule is as follows. If the neighbours x_1 and x_2 of a given cell x are both equal to 0 then x is updated to be 1 with probability $1-p$ and 0 with probability p and otherwise, x is updated to be 0 with probability $1-q$ and 1 with probability q. For a more formal definition of $E_{p,q}$, we refer the reader to 5.3.2.

In this chapter, we investigate the question of probability of draw of the parity-based game and that of ergodicity of our hePCA $E_{p,q}$.

Theorem 5.1.1. *The probability of draw in the parity-based games is* 0 *and the hePCA $E_{p,q}$ is ergodic for following ranges of p and q:*

1. $q = 0, p > p_0$, *where $p_0 \approx 0.215$ is the unique real root of* $p^5 - 5p^4 + 9p^3 - 8p^2 + 6p - 1$.
2. $q > \max\left\{\frac{-\sqrt{p^2+2}-p+2}{2}, 0\right\}, p \geq 0$.

5.2 Organization of this chapter

This chapter is organized as follows: in 5.3, we describe our model–percolation game we study is described in 5.3.1, our hePCA is described in 5.3.2, our *envelope* hePCA is introduced in 5.3.3 and finally we draw the connection between the game and hePCA in 5.3.4 where we also state our main result. We state and prove some preliminary lemmas in 5.4 and finally we prove our main theorem in 5.4.1.

5.3 Description of the model

Our model has two interrelated components, a percolation game and a generalized version of probabilistic cellular automaton. A percolation game is a combinatorial game played on a (possibly infinite) random graph and a probabilistic cellular automaton is a particular type of Markov process. We describe these components in what follows.

5.3.1 Description of our game

We consider a percolation game on $\mathbb{Z}^2$. To begin with, every vertex of $\mathbb{Z}^2$, independent of other vertices, is labelled a trap with probability p, a target with probability q and it is open with probability $1-p-q$. Here p and q are the parameters of the game and $(p,q) \in \mathscr{S}$.

The assignments of these labels are done to all the vertices simultaneously, and then the game is played between two players who take turns to move the token. Initially, a token is kept at the origin. As we have mentioned earlier, if the origin is labelled trap (or, target), we declare the player making the first move to be a winner (or, loser) of the game respectively. If the origin is labelled

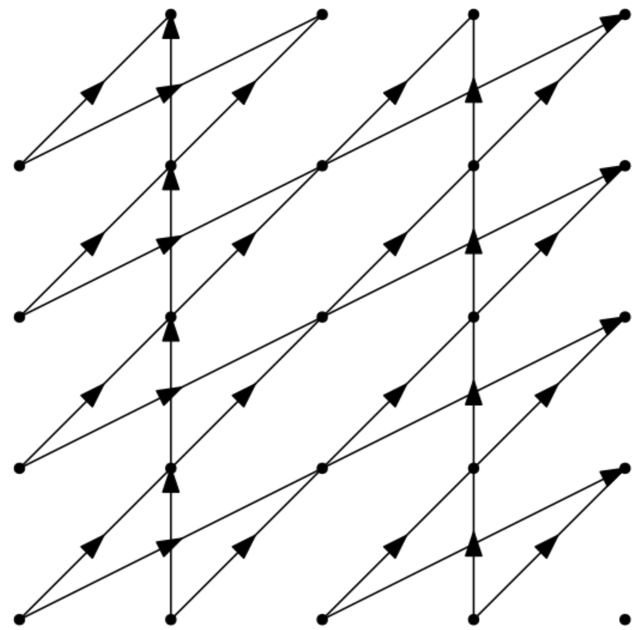


Figure 5.1: The graph on which the game is played

open then players alternate turns moving the token according to the following rule: if the token is at vertex (x,y) then the player, in turn, can move it to:

1. $(x,y+1)$ or $(x+1,y+1)$ if x is even.

2. $(x+1,y+1)$ or $(x+2,y+1)$ if x is odd.

We demonstrate the directed graph in which the game is being played in Figure 5.1.

A player wins the game if she is able to move her token to a target or if the other player is forced to move the token to a trap. The game continues as long as the token remains on an open node. If the game continues forever it is said to result in a draw. We note here that both the players decide their strategies after the graph has been assigned with the trap/target/open labelling and that they have access to the whole information of the graph and form their strategies accordingly.

5.3.2 Description of our hePCA

The definition of a heterogeneous PCA differs from the definition of a PCA in the notion of the neighbourhood. All other defining ingredients remain the same. In particular the principal defining components of a hePCA are: 1. a finite set $\mathscr{A}$ of symbols termed the *alphabet*, 2. a finite set of indices $\mathscr{N}^{\mathbf{x}} = \{\mathbf{y}_1^{\mathbf{x}}, \mathbf{y}_2^{\mathbf{x}}, \ldots, \mathbf{y}_m^{\mathbf{x}}\} \subset \mathbb{Z}^d$ (where d indicates the dimension of the CA or PCA under consideration), for every vertex $\mathbf{x}$, which is called the *neighbourhood* of the vertex $\mathbf{x}$ (of size

m), 3. and a local update rule $\varphi : \mathscr{A}^m \times \mathscr{A} \to [0,1]$. In addition to these essential ingredients the *state space* Ω equals $\mathscr{A}^{\mathbb{Z}^d}$, and the elements $\eta = (\eta(\mathbf{x}) \in \mathscr{A} : \mathbf{x} \in \mathbb{Z}^d)$ of Ω are termed *configurations*. While it is possible to define hePCAs in such a way that even the neighbourhood size, m, depends on the vertex $\mathbf{x}$, in this chapter, we restrict our attention to the case where the size of each neighbourhood remains constant. This enables us to use a common local update rule for all nodes. More formally, the update rules for a hePCA F are represented by a stochastic matrix $\varphi : \mathscr{A}^m \times \mathscr{A} \to [0,1]$ defined as follows:

$$\mathbf{P}[F\eta(\mathbf{x}) = b \big| \eta(\mathbf{x}+\mathbf{y}_i^{\mathbf{x}}) = a_i \text{ for all } 1 \leqslant i \leqslant m] = \varphi(a_1, a_2, \ldots, a_m, b) \text{ for all } b \in \mathscr{A}, \tag{5.3.1}$$

where, by definition of stochastic matrices, we have $\varphi(a_1, a_2, \ldots, a_m, b) \geqslant 0$ and $\sum_{b \in \mathscr{A}} \varphi(a_1, a_2, \ldots, a_m, b) = 1$, for all $a_1, a_2, \ldots, a_m, b \in \mathscr{A}$.

We now define the notion of ergodicity for a $d-$dimensional hePCA F. The definition is analogous to the usual definition for PCAs. Let $\mathscr{F}$ denote the $\sigma-$field that is generated by the cylinder sets of $\Omega = \mathscr{A}^{\mathbb{Z}^d}$, and let $\mathbb{D}$ denote the set of all probability measures over Ω and defined with respect to the sigma field $\mathscr{F}$. We put $F^1\eta = F\eta$ and for any natural number $t \geq 2$, we define $F^t\eta = F(F^{t-1}\eta)$ for $\eta \in \Omega$. This definition extend naturally to random η following the distribution $\mu \in \mathbb{D}$ and we let $F^t\mu$ (simply written $F\mu$ when $t = 1$ denote the probability distribution of the (also random) configuration $F^t\mu$.

We say that μ is a *stationary* measure for a hePCA F if $F\mu = \mu$. A hePCA F is said to be *ergodic* if it has unique stationary measure, such that for every probability measure ν on Ω, the sequence $F^t\nu$ converges weakly to μ as $t \to \infty$. That is, the hePCA forgets its initial state and converges to a unique measure.

We now introduce and explain the one-dimensional hePCA $E_{p,q}$ that we study in this chapter. The definition itself is the first among quite a few aspects in which the hePCA $E_{p,q}$ differs from more commonly studied PCAs in the literature. While $E_{p,q}$, as usual, is endowed with the alphabet $\mathscr{A} = \{0,1\}$ and stochastic local update rules captured by the stochastic matrix $\varphi_{p,q}$ defined below, we now need *two different* finite subsets of $\mathbb{Z}$, denoted $\mathscr{N}_1$ and $\mathscr{N}_0$, to serve as neighbourhoods in the definition of this hePCA. For $n \in \mathbb{Z}$, the neighbourhood we consider is $\mathscr{N}_1$ if n is odd, and $\mathscr{N}_0$ if n is even.

To elucidate, we set $\mathscr{N}_1 = \{1,2\}$ and $\mathscr{N}_0 = \{0,1\}$. Given a configuration $\eta = (\eta(n) : n \in \mathbb{Z}) \in \Omega$, where $\Omega = \mathscr{A}^{\mathbb{Z}}$ is the state space for the hePCA $E_{p,q}$, the (random) state $E_{p,q}\eta(n)$ of the site n under the application of $E_{p,q}$ is determined based (only) on the states $\eta(n+1)$ and $\eta(n+2)$ of sites $n+1$ and $n+2$, when n is odd, and on $\eta(n)$ and $\eta(n+1)$ of sites n and $n+1$, when n is even. It is worth noting that both $E_{p,q}\eta(2n-1)$ and $E_{p,q}\eta(2n)$ are functions of the states $\eta(2n)$

and $\eta(2n+1)$, and in particular, they are equidistributed.

For the hePCA that we are concerned with in this chapter, stochastic update rule $\varphi_{p,q}$ is defined as follows:

$$\varphi_{p,q}(W,W,b) = \begin{cases} p & \text{if } b = W, \\ 1-p & \text{if } b = L, \end{cases} \tag{5.3.2}$$

and

$$\varphi_{p,q}(a_0,a_1,b) = \begin{cases} 1-q & \text{if } b = W, \\ q & \text{if } b = L, \end{cases} \quad \text{for all } (a_0,a_1) \in \mathscr{A}^2 \setminus \{(W,W)\}. \tag{5.3.3}$$

Note that even though the definition of the stochastic rule is independent of the individual vertices, following 5.3.1, the way the individual vertices use this stochastic rule varies from vertex to vertex, and in our hePCA, $E_{p,q}$, the usage of this rule depends on the parity of the vertex.

5.3.3 Description of our envelope hePCA

The envelope $\widehat{E}_{p,q}$ to the hePCA $E_{p,q}$ is obtained by first extending the alphabet $\mathscr{A}$ to $\hat{\mathscr{A}} = \{W,D,L\}$, and then introducing the local update rules via the stochastic matrix $\hat{\varphi}_{p,q}$ defined as follows:

$$\widehat{\varphi}_{p,q}(W,W,b) = \begin{cases} p & \text{if } b = W, \\ 1-p & \text{if } b = L, \end{cases} \tag{5.3.4}$$

$$\widehat{\varphi}_{p,q}(a_0,a_1,b) = \begin{cases} 1-q & \text{if } b = W, \\ q & \text{if } b = L, \end{cases} \quad \text{for all } (a_0,a_1) \in \hat{\mathscr{A}}^2 \setminus \{W,D\}^2 \tag{5.3.5}$$

and

$$\widehat{\varphi}_{p,q}(a_0,a_1,b) = \begin{cases} p & \text{if } b = W, \\ q & \text{if } b = L, \\ r = 1-p-q & \text{if } b = D, \end{cases} \quad \text{for all } (a_0,a_1) \in \{W,D\}^2 \setminus (W,W). \tag{5.3.6}$$

The neighbourhoods $\mathscr{N}_1$ and $\mathscr{N}_0$ corresponding to odd and even vertices remain the same as in $E_{p,q}$. We reiterate here that this means that for an odd $n \in \mathbb{Z}$, $\widehat{E}_{p,q}\eta(n)$ depends on $\eta(n+1)$ and $\eta(n+2)$ whereas for even $n \in \mathbb{Z}$, $\widehat{E}_{p,q}\eta(n)$ depends on $\eta(n)$ and $\eta(n+1)$. In particular, $\widehat{E}_{p,q}\eta(2n-1), \widehat{E}_{p,q}\eta(2n)$ both depend on $\eta(2n)$ and $\eta(2n+1)$.

5.3.4 Connection between our game and the envelope hePCA

In this chapter, we delve into the underlying connection between the game we play and our hePCA. As stated in the previous section, recall that we play our game on $\mathbb{Z}^2$. Suppose every vertex of $\mathbb{Z}^2$ has been assigned to be a trap/target/open with corresponding probabilities. Throughout the section, we assume that both the players play optimally[1]. We define sets of vertices W, L, and D which we will need to establish this connection. We say that a vertex (x,y) is in W if the player who plays the first round wins the game when the token is placed at (x,y) at the beginning of the game. Similarly, we say that a vertex (x,y) is in L if the person who plays first, loses the game when the token is placed at (x,y) at the start of the game. Finally D is the set of vertices (x,y) such that when the token is initially kept at (x,y), the game is results in a draw. We follow the convention that if the vertex is a trap, then it is assigned a label W and if it is a target, it is assigned a label L.

For $k \in \mathbb{Z}$, let $H_k = \{(x,k) \in \mathbb{Z}^2 : x \in \mathbb{Z}\}$ be the horizontal line at height k. We claim that if, for any arbitrary realization of the graph into traps/targets/open sites, we know the labelling of vertices (into W, L and D) in the horizontal line H_{k+1}, then, using this information and the pre-assigned trap/target/open labelling, we can deduce the labelling of vertices in the horizontal line H_k. We note that since the neighbours of $(2x-1,y)$ and $(2x,y)$ are the same, it suffices to give arguments for $(2x,y)$:

1. Suppose both of the vertices $(2x,y+1)$ and $(2x+1,y+1)$ are labelled W, then no matter what move the first player makes, the second player is guaranteed to win the game. Hence, the vertex $(2x,y)$ should be labelled L if it is open. Otherwise, if it is a trap, it gets label W, which happens with probability p.

2. If at least one of the vertices $(2x,y+1)$ and $(2x+1,y+1)$ is labelled L then the first player can always move to this vertex in order to ensure victory in the game. This means that $(2x,y)$ should be labelled W if it is open. Otherwise, if it is a target, it gets label L, which happens with probability q.

3. If none of the vertices $(2x,y+1)$ and $(2x+1,y+1)$ is labelled L but at least one of them is labelled D, then the first player should move to this vertex to ensure a draw as the best case scenario. Hence, the vertex $(2x,y)$ should be labelled as D if it is open. Otherwise, it gets labelled W if it is a trap, which happens with probability p and it gets label L if it is a target which happens with probability q.

[1] For a formal definition of optimal play, we refer the reader to [16].

In what follows, we label the vertices with the set to which they belong. That is, if a vertex is in set W then, we label it with the symbol W. With this assignment of symbols in hand, it is clear that if $\eta \in \{W, L, D\}^{\mathbb{Z}}$ is a configuration of symbols on H_{k+1} then the random configuration on H_k is given by $\widehat{E}_{p,q}\eta$. Finally, the connection between the hePCA and the game we study is established via the following results.

Proposition 5.3.1. *The hePCA $E_{p,q}$ is ergodic if and only if the corresponding hePCA $\widehat{E}_{p,q}$ if ergodic.*

Proposition 5.3.2. *For every $(p,q) \in \mathscr{S}$, our game has probability 0 of ending in a draw if and only if the hePCA $E_{p,q}$ is ergodic.*

The proofs of Propositions 5.3.1 and 5.3.2 follow an analogous argument to Propositions 2.1 and 2.2, respectively, of [60], and Propositions 2.2 and 2.3, respectively, of [16].

Before stating the next theorem, we would like to delve on its importance and connections with the above stated propositions. We will show in the next section that for $\widehat{E}_{p,q}$ to be ergodic, it suffices to show that it has no stationary distribution which assigns positive probability to the occurrence of D. Once we have shown that the hePCA is ergodic, using the above propositions, it follows that our game has 0 probability of draw. From Proposition 5.3.1 and Proposition 5.3.2, it is easy to see that Theorem 5.1.1 and Theorem 5.3.3 are equivalent. In order to establish Theorem 5.1.1, we establish Theorem 5.3.3.

Theorem 5.3.3. *The envelope hePCA $\widehat{E}_{p,q}$ admits no stationary distribution μ that assigns positive probability to the occurrence of D, for following ranges of p and q:*

1. $q = 0, p > p_0$, *where* $p_0 \approx 0.215$ *is the unique real root of* $p^5 - 5p^4 + 9p^3 - 8p^2 + 6p - 1$.
2. $q > \max\left\{\frac{-\sqrt{p^2+2}-p+2}{2}, 0\right\}, p \geq 0.$

Remark 5.3.4. *We note that when $p = 0$, Theorem 5.3.3 tells us that whenever $q > \frac{2-\sqrt{2}}{2} \approx 0.2929$ the hePCA $\widehat{E}_{p,q}$ admits no stationary distribution which assigns a positive probability to the occurrence of D. The region covered by Theorem 5.3.3(ii) is shown in the figure* **??**.

Remark 5.3.5. *Note that the polynomial $p^5 - 5p^4 + 9p^3 - 8p^2 + 6p - 1$ assumes positive values for $p > p_0$. This polynomial, along with some other polynomials shows up in 5.4.30 where in order to obtain our result, we want all those polynomials to take positive values. It turns out that all the polynomials are positive in (p_0, p_1) where $p_1 \approx 0.555$ and is the second largest root of $p^3 - 2p^2 - p + 1$.*

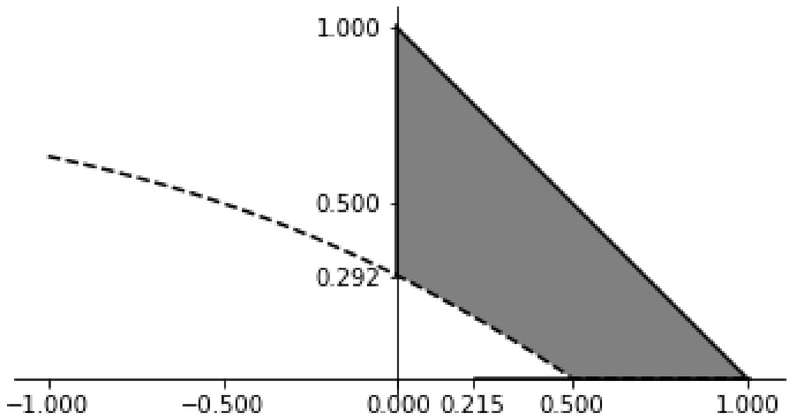


Figure 5.2: Region covered by Theorem 5.3.3 (dotted curve is not included in the region)

5.4 Proof of Theorem 5.3.3

For some $i \in \mathbb{Z}$ and $k \in \mathbb{N}$, let $S = \{i, \dots, i+k-1\} \subseteq \mathbb{Z}$ be a finite index set of size k and $a_1, a_2, \dots a_k \in \{W, L, D\}$ be some symbols. We define $(a_1, a_2, \dots, a_k)_i = \{\eta \in \{W, L, D\}^{\mathbb{Z}} : \eta(j) = a_{j-i+1} \ \forall j \in S\}$ as a cylinder set of length k indexed at i.

The lemma that we state now follows from the definition of the hePCA $\widehat{E}_{p,q}$ (in particular we use equation 5.3.1, the neighbourhoods of $\widehat{E}_{p,q}$ and the fact that the state of each site is updated independently under the action of $\widehat{E}_{p,q}$.

Lemma 5.4.1. *Let μ be a given distribution on $\{W, L, D\}^{\mathbb{Z}}$ and $\widehat{\varphi}_{p,q}$ be the local updation rule of the hePCA $E_{p,q}$ as describe in §5.3.3. Then the following holds:*

1. *If i is odd and $(b_1, \dots, b_{2k}) \in \{W, L, D\}^{2k}$, then*

$$\widehat{E}_{p,q}\mu(b_1, \dots, b_{2k})_i = \sum_{(a_1,\dots,a_{2k}) \in \{W,L,D\}^{2k}} f_1(a_1, \dots, a_{2k}, b_1, \dots, b_{2k}) \mu(a_1, \dots, a_{2k})_{i+1} \quad (5.4.1)$$

where

$$f_1(a_1, \dots, a_{2k}, b_1, \dots, b_{2k}) = \prod_{t=1}^{k} \widehat{\varphi}_{p,q}(a_{2t-1}, a_{2t}, b_{2t-1}) \widehat{\varphi}_{p,q}(a_{2t-1}, a_{2t}, b_{2t}).$$

2. *If i is odd and $(b_1, \dots, b_{2k-1}) \in \{W, L, D\}^{2k-1}$, then*

$$\widehat{E}_{p,q}\mu(b_1, \dots, b_{2k-1})_i = \sum_{(a_1 \dots a_{2k}) \in \{W,L,D\}^{2k}} f_2(a_1 \dots a_{2k}, b_1, \dots, b_{2k-1}) \mu(a_1 \dots a_{2k})_{i+1} \quad (5.4.2)$$

where

$$f_2(a_1 \ldots a_{2k}, b_1, \ldots, b_{2k-1}) = (\prod_{t=1}^{k-1} \widehat{\varphi}_{p,q}(a_{2t-1}, a_{2t}, b_{2t-1}) \widehat{\varphi}_{p,q}(a_{2t-1}, a_{2t}, b_{2t})) \widehat{\varphi}_{p,q}(a_{2k-1}, a_{2k}, b_{2k-1})$$

3. *If i is even and $(b_1, \ldots, b_{2k}) \in \{W, L, D\}^{2k}$, then*

$$\widehat{E}_{p,q}\mu(b_1, \ldots, b_{2k})_i = \sum_{(a_1,\ldots,a_{2k+2}) \in \{W,L,D\}^{2k+2}} f_3(a_1, \ldots, a_{2k+2}, b_1, \ldots, b_{2k}) \mu(a_1, \ldots, a_{2k+2})_i \tag{5.4.3}$$

where

$$f_3(a_1, \ldots, a_{2k+2}, b_1, \ldots, b_{2k}) =$$
$$\widehat{\varphi}_{p,q}(a_1, a_2, b_1)(\prod_{t=1}^{k} \widehat{\varphi}_{p,q}(a_{2t-1}, a_{2t}, b_{2t-1}) \widehat{\varphi}_{p,q}(a_{2t-1}, a_{2t}, b_{2t})) \widehat{\varphi}_{p,q}(a_{2k+1}, a_{2k}, b_{2k})$$

4. *If i is even and $(b_1, \ldots, b_{2k-1}) \in \{W, L, D\}^{2k-1}$, then*

$$\widehat{E}_{p,q}\mu(b_1, \ldots, b_{2k-1})_i = \sum_{(a_1,\ldots,a_{2k}) \in \{W,L,D\}^{2k}} f_4(a_1, \ldots, a_{2k}, b_1, \ldots, b_{2k-1}) \mu(a_1, \ldots, a_{2k})_i \tag{5.4.4}$$

where

$$f_4(a_1, \ldots, a_{2k}, b_1, \ldots, b_{2k-1}) = \widehat{\varphi}_{p,q}(a_1, a_2, b_1)(\prod_{t=1}^{k-1} \widehat{\varphi}_{p,q}(a_{2t-1}, a_{2t}, b_{2t-1}) \widehat{\varphi}_{p,q}(a_{2t-1}, a_{2t}, b_{2t}))$$

With this lemma in hand, we show that it is enough to prove Theorem 5.3.3 for distributions μ on $\{W, L, D\}^{\mathbb{Z}}$ satisfying:

1. $\mu(b_1 b_2 \ldots b_k)_i = \mu(b_1 b_2 \ldots b_k)_{i+2} \forall i \in \mathbb{Z}$.

2. $\mu(b_1 b_2 \ldots b_{2k})_i = \mu(b_{2k} b_{2k-1} \ldots b_1)_i \forall i \in \mathbb{Z}$.

3. $\mu(b_1 b_2 \ldots b_{2k+1})_i = \mu(b_{2k+1} b_{2k} \ldots b_1)_{i+1} \forall i \in \mathbb{Z}$.

4. If $i + j$ is even then, $\mu(b_1 b_2 \ldots b_j b_{j+1} \ldots b_k)_i = \mu(b_1 b_2 \ldots b_{j-1} b_{j+1} b_j b_{j+2} \ldots b_k)_i \forall i \in \mathbb{Z}$.

Lemma 5.4.2. *Let μ be a distribution on $\{W, L, D\}^{\mathbb{Z}}$ satisfying conditions (i)-(iv) above. Then $\widehat{E}_{p,q}\mu$ also satisfies the conditions (i)-(iv).*

Proof. We prove each of the properties separately in the following parts:

1. In this part, we show that $E_{p,q}\mu$ satisfies the first property. We begin with a cylinder set $(b_1 b_2 \dots b_{2k})_i$ indexed at an odd integer i. The first part of Lemma 5.4.1 gives us:

$$\widehat{E}_{p,q}\mu(b_1,\dots,b_{2k})_i = \sum_{(a_1,\dots,a_{2k})\in\{W,L,D\}^{2k}} f_1(a_1,\dots,a_{2k},b_1,\dots,b_{2k})\mu(a_1,\dots,a_{2k})_{i+1} \quad (5.4.5)$$

where

$$f_1(a_1,\dots,a_{2k},b_1,\dots,b_{2k}) = \prod_{t=1}^{k}\widehat{\varphi}_{p,q}(a_{2t-1},a_{2t},b_{2t-1})\widehat{\varphi}_{p,q}(a_{2t-1},a_{2t},b_{2t}).$$

Since μ satisfies the condition (i) above, we have that $\mu(a_1,\dots,a_{2k})_{i+1} = \mu(a_1,\dots,a_{2k})_{i+3}$. This allows us to rewrite equation 5.4.5 as

$$\begin{aligned}
\widehat{E}_{p,q}\mu(b_1,\dots,b_{2k})_i &= \sum_{(a_1,\dots,a_{2k})\in\{W,L,D\}^{2k}} f_1(a_1,\dots,a_{2k},b_1,\dots,b_{2k})\mu(a_1,\dots,a_{2k})_{i+1} \\
&= \sum_{(a_1,\dots,a_{2k})\in\{W,L,D\}^{2k}} f_1(a_1,\dots,a_{2k},b_1,\dots,b_{2k})\mu(a_1,\dots,a_{2k})_{i+3} \\
&= \widehat{E}_{p,q}\mu(b_1,\dots,b_{2k})_{i+2}
\end{aligned}$$

where in the last step, we use the first part of Lemma 5.4.1. Proofs of the other cases concerning the parities of i and k can be dealt with in a similar way.

2. In this part, we prove that $E_{p,q}\mu$ satisfies the second property. For this we begin with a cylinder set $(b_1,\dots,b_{2k})_i$ for an odd i and use Lemma 5.4.1 to conclude that

$$\widehat{E}_{p,q}\mu(b_1,\dots,b_{2k})_i = \sum_{(a_1,\dots,a_{2k})\in\{W,L,D\}^{2k}} f_1(a_1,\dots,a_{2k},b_1,\dots,b_{2k})\mu(a_1,\dots,a_{2k})_{i+1} \quad (5.4.6)$$

where

$$f_1(a_1,\dots,a_{2k},b_1,\dots,b_{2k}) = \prod_{t=1}^{k}\widehat{\varphi}_{p,q}(a_{2t-1},a_{2t},b_{2t-1})\widehat{\varphi}_{p,q}(a_{2t-1},a_{2t},b_{2t}).$$

From the fact that μ satisfies the property (ii) above, it follows that $\mu(a_1,\dots,a_{2k})_{i+1} =$

$\mu(a_{2k},\dots,a_1)_{i+1}$. We note that

$$\begin{aligned}
f_1(a_1,\dots,a_{2k},b_1,\dots,b_{2k}) &= \prod_{t=1}^{k}\widehat{\varphi}_{p,q}(a_{2t-1},a_{2t},b_{2t-1})\widehat{\varphi}_{p,q}(a_{2t-1},a_{2t},b_{2t})\\
&= \prod_{s=1}^{k}\widehat{\varphi}_{p,q}(a_{2k-2s+1},a_{2k-2s+2},b_{2k-2s+1})\widehat{\varphi}_{p,q}(a_{2k-2s+1},a_{2k-2s+2},b_{2k-2s+2})\\
&= f_1(a_{2k},\dots,a_1,b_{2k},\dots b_1)
\end{aligned}$$

This allows us to re-write equation 5.4.6 as follows:

$$\begin{aligned}
\widehat{E}_{p,q}\mu(b_1,\dots,b_{2k})_i &= \sum_{(a_1,\dots,a_{2k})\in\{W,L,D\}^{2k}} f_1(a_1,\dots,a_{2k},b_1,\dots,b_{2k})\mu(a_1,\dots,a_{2k})_{i+1}\\
&= \sum_{(a_{2k},\dots,a_1)\in\{W,L,D\}^{2k}} f_1(a_{2k},\dots,a_1,b_{2k},\dots,b_1)\mu(a_{2k},\dots,a_1)_{i+1}\\
&= \widehat{E}_{p,q}\mu(b_{2k},\dots,b_1)_i
\end{aligned}$$

as required. In the last step, we have used Lemma 5.4.1. The other case where i is even can be dealt analogously.

3. The proof that $\widehat{E}_{p,q}\mu$ satisfies the third property is analogous to proof of the previous part.

4. In this part, we prove that $\widehat{E}_{p,q}\mu$ satisfies last property. We begin with a cylinder set $(b_1b_2\dots b_{2k})_i$ with i being an odd integer. Then using Lemma 5.4.1, we see that

$$\widehat{E}_{p,q}\mu(b_1,\dots,b_{2k})_i = \sum_{(a_1,\dots,a_{2k})\in\{W,L,D\}^{2k}} f_1(a_1,\dots,a_{2k},b_1,\dots,b_{2k})\mu(a_1,\dots,a_{2k})_{i+1} \quad (5.4.7)$$

where

$$f_1(a_1,\dots,a_{2k},b_1,\dots,b_{2k}) = \prod_{t=1}^{k}\widehat{\varphi}_{p,q}(a_{2t-1},a_{2t},b_{2t-1})\widehat{\varphi}_{p,q}(a_{2t-1},a_{2t},b_{2t}).$$

Let $j\in\{1,\dots,2k\}$ be an odd integer. Using the fact that $i+j$ is an even number,it follows easily from the definition of f_1 that

$$f_1(a_1,\dots,a_{2k},b_1,\dots,b_{2k}) = f_1(a_1,\dots,a_{2k},b_1,\dots,b_{j-1},b_{j+1},b_j,b_{j+1},\dots,b_{2k})$$

This allows us to re-write equation 5.4.7 as

$$\begin{aligned}\widehat{E}_{p,q}\mu(b_1,\ldots,b_{2k})_i &= \sum_{(a_1,\ldots,a_{2k})\in\{W,L,D\}^{2k}} f_1(a_1,\ldots,a_{2k},b_1,\ldots,b_{2k})\mu(a_1,\ldots,a_{2k})_{i+1}\\ &= \sum_{(a_1,\ldots,a_{2k})\in\{W,L,D\}^{2k}} f_1(a_1,\ldots,a_{2k},b_1,\ldots,b_{j-1},b_{j+1},b_j,b_{j+1},\ldots,b_{2k})\mu(a_1,\ldots,a_{2k})_{i+1}\\ &= \widehat{E}_{p,q}\mu(b_1,\ldots,b_{j-1},b_{j+1},b_j,b_{j+1},\ldots,b_{2k})_i\end{aligned}$$

as required. We note here that the last step follows from Lemma 5.4.1. Other cases concerning the parities of i and k follow analogously.

This completes the proof of the lemma. $\square$

Remark 5.4.3. *Via an argument identical to that outlined at the very beginning of the proof of Proposition 2.3 of [60], and using Lemma 5.4.2, we conclude the following; to prove Theorem 5.3.3, it suffices to show that $\mu(D)_i = 0$ for each $i \in \{0,1\}$ for every stationary distribution μ of $\widehat{E}_{p,q}$ that satisfies the properties 1-4. Consequently, we confine ourselves to such distributions. In fact it is easy to see that such distributions μ satisfy $\mu(D)_0 = \mu(D)_1$.*

For the sake of brevity, we let $\widehat{**} = \{W,D\}^2 \setminus \{(W,W)\}$ and $\widehat{L*} = \{W,L,D\}^2 \setminus \{W,D\}^2$. Given $a_1,\ldots,a_k,b_1,\ldots,b_l \in \{W,L,D\}$ and $i \in \mathbb{N}$, we define $(a_1,\ldots,a_k,\widehat{**},b_1,\ldots,b_l)_i$ to be the cylinder set containing configurations η satisfying

1. $\eta(j) = a_{j-i+1}$ for $j \leq i+k-1$,
2. $(\eta(i+k),\eta(i+k+1)) \in \widehat{**}$,
3. $\eta(j) = a_{j-i+1}$ for $i+k+2 \leq j \leq i+l-1$.

In a similar fashion, we define $(a_1,\ldots,a_k,\widehat{L*},b_1,\ldots,b_l)_i$ to be the cylinder set containing configurations η satisfying

1. $\eta(j) = a_{j-i+1}$ for $j \leq i+k-1$,
2. $(\eta(i+k),\eta(i+k+1)) \in \widehat{L*}$,
3. $\eta(j) = a_{j-i+1}$ for $i+k+2 \leq j \leq i+l-1$.

5.4.1 Construction of our weight functions

The weight function method involves manipulations of linear combinations of measures of cylinder sets with coefficients polynomials in the parameters p and q. The goal is to come up with an inequality from where it is easy to infer the desired result, that is, $\mu(D)_0 = 0$. Formally speaking, given a distribution μ, a weight function is of the form $w(\mu) = \sum_{i=1}^{m} c_i(p,q)\mu(\mathscr{C}_i)_{a_i}$ where $(\mathscr{C}_i)_{a_i}$ is a cylinder set indexed at a_i (as defined at the beginning of §5.4), $a_i \in \{0,1\}$, c_i's are polynomials in p and q with real coefficients and m is the number of cylinder sets in the expression of our weight function. Our goal is to find a weight function $w(\mu) = \sum_{i=1}^{m} c_i(p,q)\mu(\mathscr{C}_i)_{a_i}$ such that the following inequality is satisfied: $w(\mu) - w(\widehat{E}_{p,q}\mu) \geq \sum_{i=1}^{k} d_i(p,q)\mu(\mathscr{C}_i')_{b_i}$ where d_i's are polynomials in p and q with real coefficients, $b_i \in \{0,1\}$ and $(\mathscr{C}_i')_{b_i}$ are cylinder sets indexed at b_i. We highlight here that the cylinder sets $(\mathscr{C}_i')_{b_i}$ can be different from the cylinder sets $(\mathscr{C}_i)_{a_i}$. Additionally, we want the d_i's to be such that whenever p and q are in our desired range, that is, the range in which we want to establish ergodicity, then $d_i(p,q) > 0$, for each i. When μ is a stationary measure, this allows to conclude that $\mu(\mathscr{C}_i')_{b_i} = 0$ for each i. We want the structure of the cylinder sets $(\mathscr{C}_i')_{b_i}$ to be such that from the information that all of them are of measure zero, we are able to infer that $\mu(D)_0 = \mu(D)_1 = 0$ which concludes the proof.

We construct this weight function in a step-by-step manner – beginning with an initial guess of the weight function and improving it incrementally. The idea is to start with an initial guess of the weight function, say w_0. We then analyze $w_0(\mu) - w_0(\widehat{E}_{p,q}\mu)$ and identify the 'troublesome' terms. Keeping these terms as the focus, we update our weight function so as to tackle these terms and 'improve' the weight function. We iterate the above procedure until the process of updation becomes computationally intractable. For a detailed explanation of the heuristics we follow while constructing this weight function, we refer the reader to 3.4. We would like to highlight here that there is no right choice of the initial weight function but rather it is obtained experimentally. The choices which we have made in this chapter are obtained by trial and error and are the best we could obtain keeping in mind the above stated process. In what follows, given a cylinder set $(a_1,\ldots,a_k)_i = \{\eta \in \{W,L,D\}^{\mathbb{Z}} : \eta(j) = a_{j-i+1} \ \forall j \in \{i,\ldots,i+k-1\}\}$, we will use Lemma 5.4.1 repeatedly to compute $\widehat{E}_{p,q}\mu(a_1,\ldots,a_k)_i$. As in [16] we will begin with a suitable initial weight function and then update it accordingly to obtain an inequality of the desired form as stated above.

Before delving into the construction of our weight function, we would like to highlight the additional difficulties encountered in constructing the weight function for the hePCA $\widehat{E}_{p,q}$ as compared to the PCAs in [60], [16]. Unlike the PCAs in [60], [16], the hePCA $\widehat{E}_{p,q}$ is neither reflection invariant nor translation invariant. The hePCA $\widehat{E}_{p,q}$, instead is 'two translation' invariant which is formally stated in property 1. Similarly, in place of reflection invariant property, $\widehat{E}_{p,q}$ satisfies the properties 2 and 3. We note here that $\widehat{E}_{p,q}$ satisfies property 4 which is not shared by the PCAs in [60], [16]. All of these differences are essentially coming from the parity dependent definitions of our hePCA. Because of these differences it is in fact difficult to construct the corresponding weight function covering a lot of 'range' in p and q.

We now begin by constructing the weight function for the first part of Theorem 5.3.3. For this case, since $q=0$, we denote $\widehat{E}_{p,0}$ by $\widehat{E}_p$. We begin with the initial weight function:

$$w_0(\mu) = 2\mu(DW)_1 + \mu(DD)_1 \tag{5.4.8}$$

Using Lemma 5.4.1,

$$\begin{aligned} w_0(\widehat{E}_p\mu) &= 2p(1-p)\mu(\widehat{**})_0 + (1-p)^2\mu(\widehat{**})_0 \\ &= (1-p^2)\mu(\widehat{**})_0 \end{aligned} \tag{5.4.9}$$

Re-writing $w_0(\mu)$ by re-arrangement of the terms, we obtain:

$$\begin{aligned} w_0(\mu) &= 2(\mu(WDW)_0 + \mu(DDW)_0 + \mu(LDW)_0) + \mu(WDD)_0 + \mu(DDD)_0 + \mu(LDD)_0 \\ &= 2(\mu(WDW)_0 + \mu(WDL)_0 + \mu(WDD)_0) + \mu(DDW)_0 + \mu(DDL)_0 + \mu(DDD)_0 + (\mu(LDW)_0 \\ &\quad + \mu(DDW)_0 + \mu(LDW)_0 + \mu(LDD)_0) - \mu(WDD)_0 \\ &= (2\mu(WD)_0 + \mu(DD)_0) + (\mu(LDW)_0 + \mu(DDW)_0 + \mu(LDW)_0 + \mu(LDD)_0) - \mu(WDD)_0 \\ &= \mu(\widehat{**})_0 + (2\mu(LDW)_0 + \mu(DDW)_0 + \mu(LDD)_0) - \mu(WDD)_0 \end{aligned} \tag{5.4.10}$$

Subtracting equation 5.4.9 from both sides:

$$\begin{aligned} w_0(\mu) - w_0(\widehat{E}_p\mu) &= \mu(\widehat{**})_0 + \mu(DDW)_0 + \mu(LDW)_0 + \mu(LDD)_0) - \mu(WDD)_0 - (1-p^2)\mu(\widehat{**}) \\ &= p^2\mu(\widehat{**})_0 + (2\mu(LDW)_0 + \mu(DDW)_0 + \mu(LDD)_0) - \mu(WDD)_0 \end{aligned}$$

In order to 'balance' the negative term: $-\mu(WDD)_0$, we update our weight function as follows:

$$w_1(\mu) = w_0(\mu) + \mu(WDD)_0 \tag{5.4.11}$$

Using Lemma 5.4.1, we obtain:

$$\begin{aligned} \widehat{E}_p\mu(WDD)_0 &= p(1-p)^2(\mu(WW\widehat{**})_0 + \mu(\widehat{****})_0) + (1-p)^2(\mu(\widehat{L\widehat{**}*})_0) \\ &= p(1-p)^2(\mu(WW\widehat{**})_0 + \mu(\widehat{****})_0) + (1-p)^2(\mu(LWWD)_0 + \mu(LWDW)_0 + \mu(LWDD)_0 \\ &\quad + \mu(LDWD)_0 + \mu(LDDW)_0 + \mu(LDDD)_0 + \mu(WLWD)_0 + \mu(WLDW)_0 + \mu(WLDD)_0 \end{aligned}$$

$$
\begin{aligned}
&+\mu(DLWD)_0+\mu(DLDW)_0+\mu(DLDD)_0+\mu(LLWD)_0+\mu(LLDW)_0+\mu(LLDD)_0)\\
&=p(1-p)^2(\mu(WW\widehat{**})_0+\mu(\widehat{****})_0)+(1-p)^2(\mu(LWWD)_0+\mu(LWDW)_0+\mu(LWDD)_0\\
&+\mu(LDWD)_0+\mu(LDDW)_0+\mu(LDDD)_0+\mu(WLWD)_0+\mu(DLWD)_0+\mu(LLWD)_0)
\end{aligned}
$$

Using equations 5.4.10 and 5.4.11, we obtain:

$$
\begin{aligned}
w_1(\mu)-w_1(\widehat{E}_p\mu) &= p^2\mu(\widehat{**})_0+(2\mu(LDW)_0+\mu(DDW)_0+\mu(LDD)_0)-\mu(WDD)_0+\mu(WDD_0)\\
&-p(1-p)^2(\mu(WW\widehat{**})_0+\mu(\widehat{****})_0)-(1-p)^2(\mu(LWWD)_0+\mu(LWDW)_0\\
&+\mu(LWDD)_0+\mu(LDWD)_0+\mu(LDDW)_0+\mu(LDDD)_0+\mu(WLWD)_0+\mu(DLWD)_0\\
&+\mu(LLWD)_0)\\
&= p^2\mu(\widehat{**})_0+(\mu(LDW)_0+\mu(DDW)_0)+(\mu(LDW)_0+\mu(LDD)_0)\\
&-p(1-p)^2(\mu(WW\widehat{**})_0+\mu(\widehat{****})_0)-(1-p)^2(\mu(LWWD)_0+\mu(LWDW)_0\\
&+\mu(LWDD)_0+\mu(LDWD)_0+\mu(LDDW)_0+\mu(LDDD)_0+\mu(WLWD)_0+\mu(DLWD)_0\\
&+\mu(LLWD)_0)\\
&= p^2\mu(\widehat{**})_0+(\mu(LDW)_0+\mu(DDW)_0)+(\mu(LDWW)_0+\mu(LDWL)_0+\mu(LDWD)_0\\
&+\mu(LDDW)_0+\mu(LDDL)_0+\mu(LDDD)_0)-p(1-p)^2(\mu(WW\widehat{**})_0+\mu(\widehat{****})_0)\\
&-(1-p)^2(\mu(LWWD)_0+\mu(LWDW)_0+\mu(LWDD)_0+\mu(LDWD)_0+\mu(LDDW)_0\\
&+\mu(LDDD)_0+\mu(WLWD)_0\\
&+\mu(DLWD)_0+\mu(LLWD)_0)\\
&= p^2\mu(\widehat{**})_0+(\mu(LDW)_0+\mu(DDW)_0)+(\mu(LDWL)_0)+\mu(LDDL)_0+(\mu(LWDW)_0\\
&+\mu(LDWD)_0+\mu(LDDW)_0+\mu(LDDD)_0-p(1-p)^2(\mu(WW\widehat{**})_0+\mu(\widehat{****})_0)\\
&-(1-p)^2(\mu(LWWD)_0+\mu(LWDW)_0+\mu(LWDD)_0+\mu(LDWD)_0+\mu(LDDW)_0\\
&+\mu(LDDD)_0+\mu(WLWD)_0+\mu(DLWD)_0+\mu(LLWD)_0)\\
&= p^2\mu(\widehat{**})_0+(\mu(LDW)_0+\mu(DDW)_0)+(\mu(LDWL)_0+\mu(LDDL)_0)\\
&+p(2-p)(\mu(LWDW)_0+\mu(LDWD)_0+\mu(LDDW)_0+\mu(LDDD)_0\\
&-p(1-p)^2(\mu(WW\widehat{**})_0+\mu(\widehat{****})_0)-(1-p)^2(\mu(LWWD)_0+\mu(LWDD)_0\\
&+\mu(WLWD)_0+\mu(DLWD)_0+\mu(LLWD)_0) \qquad (5.4.12)
\end{aligned}
$$

The terms $-p(1-p)^2(\mu(WW\widehat{**})_0+\mu(\widehat{****})_0)$ along with the terms $(\mu(LDW)_0+\mu(DDW)_0)$ suggest that we should update our weight function as follows:

$$
w_2(\mu)=w_1(\mu)-\mu(LDW)_0-\mu(DDW)_0 \qquad (5.4.13)
$$

This is because of the following equations obtained using Lemma 5.4.1:

$$\widehat{E}_p\mu(LDW)_0 = p(1-p)^2\mu(WW\widehat{**})_0 \tag{5.4.14}$$

$$\widehat{E}_p\mu(DDW)_0 = p(1-p)^2\mu(\widehat{****})_0 \tag{5.4.15}$$

Using equations 5.4.12, 5.4.13, 5.4.14 and 5.4.15, we obtain that

$$\begin{aligned} w_2(\mu) - w_2(\widehat{E}_p\mu) = {} & p^2\mu(\widehat{**})_0 + (\mu(LDWL)_0 + \mu(LDDL)_0) + p(2-p)(\mu(LWDW)_0 + \mu(LDWD)_0 \\ & + \mu(LDDW)_0 + \mu(LDDD)_0) - (1-p)^2(\mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1) \end{aligned} \tag{5.4.16}$$

Before updating our weight function further, we note that the equation 5.4.16 yields to us the weight function for the case $q = 0, p > 0.5$. In particular, we note that

$$\mu(WD)_0 \geq \mu(LWWD)_0 + \mu(LWD)_1 \tag{5.4.17}$$

and that

$$\mu(DD)_0 \geq \mu(LWDD)_0 \tag{5.4.18}$$

From equations 5.4.17 and 5.4.18 we obtain that

$$\mu(\widehat{**})_0 \geq \mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1 \tag{5.4.19}$$

Using this in equation5.4.16 we obtain that

$$\begin{aligned} w_2(\mu) - w_2(\widehat{E}_p\mu) = {} & p^2(\mu(\widehat{**})_0 - (\mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1) + (\mu(LDWL)_0 + \mu(LDDL)_0) \\ & + p(2-p)(\mu(LWDW)_0 + \mu(LDWD)_0 + \mu(LDDW)_0 + \mu(LDDD)_0) \\ & - (1-p)^2(\mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1) + p^2(\mu(LWWD)_0 + \mu(LWDD)_0 \\ & + \mu(LWD)_1) \\ = {} & p^2(\mu(\widehat{**})_0 - (\mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1) + (\mu(LDWL)_0 + \mu(LDDL)_0) \\ & + p(2-p)(\mu(LWDW)_0 + \mu(LDWD)_0 + \mu(LDDW)_0 + \mu(LDDD)_0) \\ & + (2p-1)(\mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1) \end{aligned} \tag{5.4.20}$$

All the terms of right hand side of of the above equation are positive when $p > 0.5$. Thus, when μ

is a stationary distribution, we obtain that

$$\mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1 = 0 \tag{5.4.21}$$

and that

$$\mu(\widehat{**})_0 - (\mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1) = 0 \tag{5.4.22}$$

From these two equations, it follows that $\mu(\widehat{**})_0 = 0$. Finally, from $\mu(D) = (1-p)\mu(\widehat{**})_0$, it follows that $\mu(D) = 0$, as required.

Continuing from where we left, equation 5.4.16 suggests us that we should update our weight function as follows:

$$w_3(\mu) = w_2(\mu) + \mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1 \tag{5.4.23}$$

Using Lemma 5.4.1, we obtain:

$$\widehat{E}_p\mu(LWWD)_0 = p^2(1-p)^2(\mu(WWWW\widehat{**})_0 + \mu(WW\widehat{****})_0) + (1-p)^2(\mu(WW\widehat{L}\widehat{***})_0)$$

$$\widehat{E}_p\mu(LWDD)_0 = p(1-p)^3(\mu(WW\widehat{****})_0)$$

$$\widehat{E}_p\mu(LWD)_1 = p(1-p)^2(\mu(WW\widehat{**})_0)$$

From equations 5.4.20, 5.4.23, 5.4.24, 5.4.24 and 5.4.24, it follows that

$$\begin{aligned}
w_3(\mu) - w_3(\widehat{E}_p\mu) = {} & p^2\mu(\widehat{**})_0 + (\mu(LDWL)_0 + \mu(LDDL)_0) + p(2-p)(\mu(LWDW)_0 + \mu(LDWD)_0 \\
& + \mu(LDDW)_0 + \mu(LDDD)_0) - (1-p)^2(\mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1) \\
& + (\mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1) - (p^2(1-p)^2(\mu(WWWW\widehat{**})_0 \\
& + \mu(WW\widehat{****})_0) + (1-p)^2(\mu(WW\widehat{L}\widehat{***})_0)) - p(1-p)^3(\mu(WW\widehat{****})_0) \\
& - p(1-p)^2(\mu(WW\widehat{**})_0) \\
= {} & p^2\mu(\widehat{**})_0 + (\mu(LDWL)_0 + \mu(LDDL)_0) + p(2-p)(\mu(LWDW)_0 + \mu(LDWD)_0 \\
& + \mu(LDDW)_0 + \mu(LDDD)_0 + \mu(LWWD)_0 + \mu(LWDD)_0 + \mu(LWD)_1) \\
& - (p^2(1-p)^2(\mu(WWWW\widehat{**})_0 + \mu(WW\widehat{****})_0) + (1-p)^2(\mu(WW\widehat{L}\widehat{***})_0)) \\
& - p(1-p)^3(\mu(WW\widehat{****})_0) - p(1-p)^2(\mu(WW\widehat{**}WW)_0 + \mu(WW\widehat{****})_0 \\
& + \mu(WW\widehat{**}\widehat{L*})_0)
\end{aligned}$$

$$
\begin{aligned}
&= p^2\mu(\widehat{**})_0 + (\mu(LDWL)_0 + \mu(LDDL)_0) + p(2-p)(\mu(LWDW)_0 + \mu(LDWD)_0 \\
&+ \mu(LDDW)_0 + \mu(LDDD)_0 + \mu(LWWD)_0 + \mu(LWDD)_0 + \mu(WLWD)_0 + \mu(LLWD)_0 \\
&+ \mu(DLWD)_0) - p^2(1-p)^2\mu(WWWW\widehat{**})_0 - 2p(1-p)^2\mu(WW\widehat{****})_0 \\
&- p(1-p)^2\mu(WW\widehat{**}WW)_0 - p(1-p)^2\mu(WW\widehat{**}\widehat{L*})_0 - (1-p)^2\mu(WW\widehat{L*}\widehat{**})_0 \\
&= p^2\mu(\widehat{**})_0 + (\mu(LDWL)_0 + \mu(LDDL)_0) + p(2-p)(\mu(\widehat{L*}\widehat{**})) \\
&\ - p^2(1-p)^2\mu(WWWW\widehat{**})_0 - 2p(1-p)^2\mu(WW\widehat{****})_0 - p(1-p)^2\mu(WW\widehat{**}WW)_0 \\
&\ - p(1-p)^2\mu(WW\widehat{**}\widehat{L*})_0 - (1-p)^2\mu(WW\widehat{L*}\widehat{**})_0 \\
&\geq p^2\mu(\widehat{**})_0 + (\mu(LDWL)_0 + \mu(LDDL)_0) + p(2-p)(\mu(WW\widehat{L*}\widehat{**})_0 + \mu(L\widehat{L*}\widehat{**})_1 \\
&+ \mu(D\widehat{L*}\widehat{**})_1) - p^2(1-p)^2\mu(WWWW\widehat{**})_0 - 2p(1-p)^2\mu(WW\widehat{****})_0 \\
&\ - p(1-p)^2\mu(WW\widehat{**}WW)_0 - p(1-p)^2\mu(WW\widehat{**}\widehat{L*})_0 - (1-p)^2\mu(WW\widehat{L*}\widehat{**})_0 \\
&= p^2\mu(\widehat{**})_0 + (\mu(LDWL)_0 + \mu(LDDL)_0) + p(2-p)(\mu(L\widehat{L*}\widehat{**})_1 + \mu(D\widehat{L*}\widehat{**})_1) \\
&- p^2(1-p)^2\mu(WWWW\widehat{**})_0 - 2p(1-p)^2\mu(WW\widehat{****})_0 - p(1-p)^2\mu(WW\widehat{**}WW)_0 \\
&- p(1-p)^2\mu(WW\widehat{**}\widehat{L*})_0 - (1+2p^2-4p)\mu(WW\widehat{L*}\widehat{**})_0 \qquad (5.4.24)
\end{aligned}
$$

We update the weight function as follows:

$$
w_4(\mu) = w_3(\mu) - p(2-p)\mu(L\widehat{L*}\widehat{**})_1 - p(2-p)\mu(D\widehat{L*}\widehat{**})_1 \qquad (5.4.25)
$$

Since $\widehat{E}_p\mu(LD)_1 = 0$, we can simplify equation 5.4.25 as follows:

$$
\begin{aligned}
\mu(L\widehat{L*}\widehat{**})_1 &= \mu(LLWWD)_1 + \mu(LLWDW)_1 + \mu(LLWDD)_1 + \mu(LLDWD)_1 + \mu(LLDDW)_1 + \mu(LLDDD)_1 \\
&+ \mu(LLLWD)_1 + \mu(LLLDW)_1 + \mu(LLLDD)_1 + \mu(LWLWD)_1 + \mu(LWLDW)_1 + \mu(LWLDD)_1 \\
&+ \mu(LDLWD)_1 + \mu(LDLDW)_1 + \mu(LDLDD)_1 \\
&= \mu(LLWWD)_1 + \mu(LLWDW)_1 + \mu(LLWDD)_1 + \mu(LLDWD)_1 + \mu(LLDDW)_1 + \mu(LLDDD)_1 \\
&+ \mu(LLLWD)_1 + \mu(LWLWD)_1
\end{aligned}
$$

and

$$
\begin{aligned}
\mu(D\widehat{L*}\widehat{**})_1 &= \mu(DLWWD)_1 + \mu(DLWDW)_1 + \mu(DLWDD)_1 + \mu(DLDWD)_1 + \mu(DLDDW)_1 + \mu(DLDDD)_1 \\
&+ \mu(DLLWD)_1 + \mu(DLLDW)_1 + \mu(DLLDD)_1 + \mu(DWLWD)_1 + \mu(DWLDW)_1 + \mu(DWLDD)_1 \\
&+ \mu(DDLWD)_1 + \mu(DDLDW)_1 + \mu(DDLDD)_1 \\
&= \mu(DWLWD)_1 + \mu(DDLWD)_1
\end{aligned}
$$

These two identities allow us to re-write equation 5.4.25 as follows:

$$w_4(\mu) = w_3(\mu) - p(2-p)(\mu(LLWWD)_1 + \mu(LLWDW)_1 + \mu(LLWDD)_1 + \mu(LLDWD)_1 + \mu(LLDDW)_1 + \mu(LLDDD)_1 + \mu(LLLWD)_1 + \mu(LWLWD)_1 + \mu(DWLWD)_1 + \mu(DDLWD)_1) \quad (5.4.26)$$

Using Lemma 5.4.1, we obtain the following:

$$\widehat{E}_p\mu(LLWWD)_1 = p^2(1-p)^3(\mu(WWWW\widehat{**})_0 + \mu(WW\widehat{****})_0) + (1-p)^3\mu(WW\widehat{L*}\widehat{**})_0$$

$$\widehat{E}_p\mu(LLWDW)_1 = p^2(1-p)^3(\mu(WW\widehat{**}WW)_0 + \mu(WW\widehat{****})_0) + p(1-p)^3\mu(WW\widehat{**}\widehat{L*})_0$$

$$\widehat{E}_p\mu(LLWDD)_1 = p(1-p)^4\mu(WW\widehat{****})_0$$

$$\widehat{E}_p\mu(LLDWD)_1 = p(1-p)^4\mu(WW\widehat{****})_0$$

$$\widehat{E}_p\mu(LLDDW)_1 = p(1-p)^4(\mu(WW\widehat{**}WW)_0 + \mu(WW\widehat{****})_0) + (1-p)^4\mu(WW\widehat{**}\widehat{L*})$$

$$\widehat{E}_p\mu(LLDDD)_1 = (1-p)^5\mu(WW\widehat{****})_0$$

$$\widehat{E}_p\mu(LLLWD)_1 = p(1-p)^4\mu(WWWW\widehat{**})_0$$

$$\widehat{E}_p\mu(LWLWD)_1 = p^2(1-p)^3\mu(WWWW\widehat{**})_0$$

$$\widehat{E}_p\mu(DWLWD)_1 = p^2(1-p)^3\mu(\widehat{**}WW\widehat{**})_0$$

$$\widehat{E}_p\mu(DDLWD)_1 = p(1-p)^4\mu(\widehat{**}WW\widehat{**})_0$$

Adding the previous ten equations and gathering coefficients of various cylinder sets, we obtain:

$$\begin{aligned} w_4(\widehat{E}_p\mu) \le & w_3(\widehat{E}_p\mu) - p(2-p)(p(1-p)^3(1+p)\mu(WWWW\widehat{**})_0 + (1-p)^3(1+p)\mu(WW\widehat{****})_0 \\ & + (1-p)^3\mu(WW\widehat{**}\widehat{L*})_0 + (1-p)^3\mu(WW\widehat{L*}\widehat{**})_0) \end{aligned} \tag{5.4.27}$$

Using 5.4.24, 5.4.26 and 5.4.27, we obtain the following:

$$\begin{aligned} w_4(\mu) - w_4(\widehat{E}_p\mu) \ge\ & p^2\mu(\widehat{**})_0 + (\mu(LDWL)_0 + \mu(LDDL)_0) + p(2-p)(\mu(L\widehat{L*}\widehat{**})_1 + \mu(D\widehat{L*}\widehat{**})_1) \\ & - p^2(1-p)^2\mu(WWWW\widehat{**})_0 - 2p(1-p)^2\mu(WW\widehat{****})_0 - p(1-p)^2\mu(WW\widehat{**}WW)_0 \\ & - p(1-p)^2\mu(WW\widehat{**}\widehat{L*})_0 - (1+2p^2-4p)\mu(WW\widehat{L*}\widehat{**})_0 - p(2-p)(\mu(L\widehat{L*}\widehat{**})_1 \\ & + \mu(D\widehat{L*}\widehat{**})_1) + p(2-p)(p(1-p)^3(1+p)\mu(WWWW\widehat{**})_0 \\ & + (1-p)^3(1+p)\mu(WW\widehat{****})_0 + (1-p)^3\mu(WW\widehat{**}\widehat{L*})_0 + (1-p)^3\mu(WW\widehat{L*}\widehat{**})_0) \\ =\ & p^2\mu(\widehat{**})_0 + (\mu(LDWL)_0 + \mu(LDDL)_0) - p^2(1-p)^2\mu(WWWW\widehat{**})_0 \\ & - 2p(1-p)^2\mu(WW\widehat{****})_0 - p(1-p)^2\mu(WW\widehat{**}WW)_0 - p(1-p)^2\mu(WW\widehat{**}\widehat{L*})_0 \\ & - (1+2p^2-4p)\mu(WW\widehat{L*}\widehat{**})_0 + p(2-p)(p(1-p)^3(1+p)\mu(WWWW\widehat{**})_0 \\ & + (1-p)^3(1+p)\mu(WW\widehat{****})_0 + (1-p)^3\mu(WW\widehat{**}\widehat{L*})_0 + (1-p)^3\mu(WW\widehat{L*}\widehat{**})_0) \end{aligned}$$

Finally, we update the weight function as follows:

$$w_5(\mu) = w_4(\mu) - \mu(LDWL)_0 - \mu(LDDL)_0 \tag{5.4.28}$$

Using Lemma 5.4.1, we obtain:

$$\widehat{E}_p\mu(LDWL)_0 = p(1-p)^3\mu(WW\widehat{**}WW)_0$$

$$\widehat{E}_p\mu(LDDL)_0 = (1-p)^4\mu(WW\widehat{**}WW)_0$$

From 5.4.28, 5.4.29 and 5.4.29, we see that:

$$\begin{aligned} w_5(\mu) - w_5(\widehat{E}_p\mu) \ge\ & p^2\mu(\widehat{**})_0 + (\mu(LDWL)_0 + \mu(LDDL)_0) - p^2(1-p)^2\mu(WWWW\widehat{**})_0 \\ & - 2p(1-p)^2\mu(WW\widehat{****})_0 - p(1-p)^2\mu(WW\widehat{**}WW)_0 - p(1-p)^2\mu(WW\widehat{**}\widehat{L*})_0 \\ & - (1+2p^2-4p)\mu(WW\widehat{L*}\widehat{**})_0 + p(2-p)(p(1-p)^3(1+p)\mu(WWWW\widehat{**})_0 \\ & + (1-p)^3(1+p)\mu(WW\widehat{****})_0 + (1-p)^3\mu(WW\widehat{**}\widehat{L*})_0 + (1-p)^3\mu(WW\widehat{L*}\widehat{**})_0) \\ & - (\mu(LDWL)_0 + \mu(LDDL)_0) + p(1-p)^3\mu(WW\widehat{**}WW)_0 + (1-p)^4\mu(WW\widehat{**}WW)_0 \end{aligned}$$

$$
\begin{aligned}
&= p^2\mu(\widehat{**})_0 - p^2(1-p)^2\mu(WWWW\widehat{**})_0 - 2p(1-p)^2\mu(WW\widehat{**}\widehat{**})_0 \\
&- p(1-p)^2\mu(WW\widehat{**}WW)_0 - p(1-p)^2\mu(WW\widehat{**}\widehat{L*})_0 - (1+2p^2-4p)\mu(WW\widehat{L*}\widehat{**})_0 \\
&+ p(2-p)(p(1-p)^3(1+p)\mu(WWWW\widehat{**})_0 + (1-p)^3(1+p)\mu(WW\widehat{**}\widehat{**})_0 \\
&+ (1-p)^3\mu(WW\widehat{**}\widehat{L*})_0 + (1-p)^3\mu(WW\widehat{L*}\widehat{**})_0) + p(1-p)^3\mu(WW\widehat{**}WW)_0 \\
&+ (1-p)^4\mu(WW\widehat{**}WW)_0
\end{aligned}
$$

We now note that:

$$
\begin{aligned}
\mu(\widehat{**})_0 &\geq \mu(WW\widehat{**})_0 + \mu(\widehat{L*}\widehat{**})_0 \\
&\geq \mu(WW\widehat{**}\widehat{**})_0 + \mu(WW\widehat{**}WW)_0 + \mu(WW\widehat{**}\widehat{L*})_0 + \mu(WW\widehat{L*}\widehat{**})_0 \qquad (5.4.29)
\end{aligned}
$$

This gives us

$$
\begin{aligned}
w_5(\mu) - w_5(\widehat{E}_p\mu) &\geq p^2\mu(\widehat{**})_0 - p^2(1-p)^2\mu(WWWW\widehat{**})_0 - 2p(1-p)^2\mu(WW\widehat{**}\widehat{**})_0 \\
&- p(1-p)^2\mu(WW\widehat{**}WW)_0 - p(1-p)^2\mu(WW\widehat{**}\widehat{L*})_0 - (1+2p^2-4p)\mu(WW\widehat{L*}\widehat{**})_0 \\
&+ p(2-p)(p(1-p)^3(1+p)\mu(WWWW\widehat{**})_0 + (1-p)^3(1+p)\mu(WW\widehat{**}\widehat{**})_0 \\
&+ (1-p)^3\mu(WW\widehat{**}\widehat{L*})_0 + (1-p)^3\mu(WW\widehat{L*}\widehat{**})_0) + p(1-p)^3\mu(WW\widehat{**}WW)_0 \\
&+ (1-p)^4\mu(WW\widehat{**}WW)_0 \\
&= p^2(\mu(\widehat{**}_0 - (\mu(WW\widehat{**}\widehat{**})_0 + \mu(WW\widehat{**}WW)_0 + \mu(WW\widehat{**}\widehat{L*})_0 + \mu(WW\widehat{L*}\widehat{**})_0)) \\
&+ p^2(\mu(WW\widehat{**}\widehat{**})_0 + \mu(WW\widehat{**}WW)_0 + \mu(WW\widehat{**}\widehat{L*})_0 + \mu(WW\widehat{L*}\widehat{**})_0) \\
&- p^2(1-p)^2\mu(WWWW\widehat{**})_0 - 2p(1-p)^2\mu(WW\widehat{**}\widehat{**})_0 - p(1-p)^2\mu(WW\widehat{**}WW)_0 \\
&- p(1-p)^2\mu(WW\widehat{**}\widehat{L*})_0 - (1+2p^2-4p)\mu(WW\widehat{L*}\widehat{**})_0 \\
&+ p(2-p)(p(1-p)^3(1+p)\mu(WWWW\widehat{**})_0 + (1-p)^3(1+p)\mu(WW\widehat{**}\widehat{**})_0 \\
&+ (1-p)^3\mu(WW\widehat{**}\widehat{L*})_0 + (1-p)^3\mu(WW\widehat{L*}\widehat{**})_0) + p(1-p)^3\mu(WW\widehat{**}WW)_0 \\
&+ (1-p)^4\mu(WW\widehat{**}WW)_0 \\
&= p^2(\mu(\widehat{**})_0 - (\mu(WW\widehat{**}\widehat{**})_0 + \mu(WW\widehat{**}WW)_0 + \mu(WW\widehat{**}\widehat{L*})_0 + \mu(WW\widehat{L*}\widehat{**})_0)) \\
&\qquad (5.4.30) \\
&+ p^2(1-p)^2(p^3-2p^2-p+1)\mu(WWWW\widehat{**}) + p^4(p-2)^2\mu(WW\widehat{**}\widehat{**})_0 \\
&+ (1-4p+6p^2-2p^3)\mu(WW\widehat{**}WW)_0 + p(p^4-5p^3+8p^2-4p+1)\mu(WW\widehat{**}\widehat{L*})_0 \\
&+ (p^5-5p^4+9p^3-8p^2+6p-1)\mu(WW\widehat{L*}\widehat{**})_0
\end{aligned}
$$

Using computer simulations we see that the coefficients of all the cylinder sets above are positive when $p \in (p_0, p_1)$ where $p_0 \approx 0.215$ is the unique real root of $p^5 - 5p^4 + 9p^3 - 8p^2 + 6p - 1$

and $p_1 \approx 0.555$ and is the second largest root of $p^3 - 2p^2 - p + 1$. For $p \in (p_0, p_1)$ and μ stationary, we obtain that

1. $\mu(WWWW\widehat{**})_0 = 0$.
2. $\mu(WW\widehat{\widehat{**}\widehat{**}})_) = 0$.
3. $\mu(WW\widehat{**}WW)_0 = 0$.
4. $\mu(WW\widehat{**}\widehat{L*})_0 = 0$.
5. $\mu(WW\widehat{L\widehat{**}*}) = 0$.
6. $\mu(\widehat{**})_0 = \mu(WW\widehat{\widehat{**}\widehat{**}})_0 + \mu(WW\widehat{**}WW)_0 + \mu(WW\widehat{**}\widehat{L*})_0 + \mu(WW\widehat{L\widehat{**}*})_0 = 0$

Finally, from $\mu(\widehat{**})_0 = 0$ and $E_p\mu(D) = (1-p)\mu(\widehat{**})_0$, it follows that $\mu(D) = 0$.

For the second part of Theorem 5.3.3, we begin with the initial weight function

$$\begin{aligned}
w_0(\mu) &= \mu(DD)_1 + \mu(LD)_1 + \mu(DL)_1 + \mu(WD)_1 + \mu(DW)_1 \\
&= \mu(DD)_1 + 2\mu(DL)_1 + 2\mu(DW)_1 \\
&= \mu(WDD)_0 + \mu(LDD)_0 + \mu(DDD)_0 + 2(\mu(WDL)_0 + \mu(LDL)_0 + \mu(DDL)_0) + 2(\mu(WDW)_0 \\
&+ \mu(LDW)_0 + \mu(DDW)_0)
\end{aligned} \tag{5.4.31}$$

Using Lemma 5.4.1, we obtain the following identities:

$$\widehat{E}_{p,q}\mu(DD)_1 = r^2\mu(\widehat{**})_0 \tag{5.4.32}$$

$$\widehat{E}_{p,q}\mu(DL)_1 = rq\mu(\widehat{**})_0 \tag{5.4.33}$$

$$\widehat{E}_{p,q}\mu(DW)_1 = rp\mu(\widehat{**})_0 \tag{5.4.34}$$

Adding equations 5.4.32, 5.4.33 and 5.4.34, we obtain that:

$$\begin{aligned}
w_0(\widehat{E}_{p,q}\mu) &= (r^2 + 2rp + 2rq)\mu(\widehat{**})_0 \\
&= (1-(p+q)^2)\mu(\widehat{**})_0 \\
&= (1-(p+q)^2)(\mu(WD)_0 + \mu(DW)_0 + \mu(DD)_0) \\
&= (1-(p+q)^2)(2\mu(WD)_0 + \mu(DD)_0) \\
&= (1-(p+q)^2)(2(\mu(WDW)_0 + \mu(LDW)_0 + \mu(DDW)_0) + \mu(DDW)_0 + \mu(DDL)_0 + \mu(DDD)_0
\end{aligned} \tag{5.4.35}$$

using equations 5.4.31 and 5.4.35, we obtain:

$$\begin{aligned}
w_0(\mu) - w_0(\widehat{E}_{p,q}\mu) &= (p+q)^2\mu(\widehat{**})_0 + \mu(LDD)_0 + \mu(DDL)_0 + 2\mu(LDL)_0 + 2\mu(LDW)_0 \\
&+ \mu(DDW)_0 - \mu(WDW)_0 \\
&= (p+q)^2\mu(\widehat{**})_0 + \mu(WLDD)_1 + \mu(LLDD)_1 + \mu(DLDD)_1 + \mu(DDL)_0 \\
&+ 2\mu(LDL)_0 + 2\mu(LDW)_0 + \mu(DDW)_0 - \mu(WWDD)_1 - \mu(LWDD)_1 - \mu(DWDD)_1 \\
&= (p+q)^2\mu(\widehat{**})_0 + \mu(LLDD)_1 + \mu(DLDD)_1 + \mu(DDL)_0 + 2\mu(LDL)_0 \\
&+ 2\mu(LDW)_0 + \mu(DDW)_0 - \mu(WWDD)_1 - \mu(DWDD)_1
\end{aligned}$$

We update our weight function as follows:

$$\begin{aligned}
w_1(\mu) = w_0(\mu) &- \mu(LLDD)_1 - \mu(DLDD)_1 - \mu(DDL)_0 - 2\mu(LDL)_0 \\
&- 2\mu(LDW)_0 - \mu(DDW)_0 + \mu(WWDD)_1 + \mu(DWDD)_1
\end{aligned}$$

Using Lemma 5.4.1 we have the following identities:

$$\widehat{E}_{p,q}\mu(LLDD)_1 = (1-p)^2r^2\mu(WW\widehat{**})_0 + q^2r^2\mu(\widehat{L\widehat{**}*})_0 + q^2r^2\mu(\widehat{****})_0 \tag{5.4.36}$$

$$\widehat{E}_{p,q}\mu(DLDD)_1 = qr^3\mu(\widehat{****})_0 \tag{5.4.37}$$

$$\widehat{E}_{p,q}\mu(DDL)_0 = qr^2\mu(\widehat{****})_0 \tag{5.4.38}$$

$$\widehat{E}_{p,q}\mu(LDL)_0 = (1-p)^2r\mu(WW\widehat{**})_0 + rq^2\mu(\widehat{L\widehat{**}*})_0 + rq^2\mu(\widehat{****})_0 \tag{5.4.39}$$

$$\widehat{E}_{p,q}\mu(LDW)_0 = (1-p)pr\mu(WW\widehat{**})_0 + pqr\mu(\widehat{L\widehat{**}*})_0 pqr\mu(\widehat{****})_0 \tag{5.4.40}$$

$$\widehat{E}_{p,q}\mu(DDW)_0 = pr^2\mu(\widehat{****})_0 \tag{5.4.41}$$

$$\widehat{E}_{p,q}\mu(WWDD)_1 = p^2r^2\mu(WW\widehat{**})_0 + (1-q)^2r^2\mu(\widehat{L***})_0 + p^2r^2\mu(\widehat{****})_0 \tag{5.4.42}$$

Using these equations and rewriting $\mu(\widehat{**})_0$ as $\mu(WW\widehat{**})_0 + \mu(\widehat{L***})_0 + \mu(\widehat{****})_0$, we obtain the following:

$$\begin{aligned}
w_1(\mu) - w_1(\widehat{E}_{p,q}\mu) &= ((p+q)^2 + (1-p)^2r^2 + (1-p)pr + (1-p)rq - p^2r^2)\mu(WW\widehat{**})_0 \\
&+ ((p+q)^2 + q^2r^2 + 2pqr + 2rq^2 - (1-q)^2r^2)\mu(\widehat{L***})_0 \\
&+ ((p+q)^2 + q^2r^2 + qr^3 + 2pqr + r^2p + 2rq^2 + r^2q - p^2r^2 - pr^3)\mu(\widehat{****})_0 \\
&= ((p+q)^2 + (1-p)^2r^2 + (1-p)pr + (1-p)rq - p^2r^2)\mu(WW\widehat{**})_0 \\
&+ (-2q^2 + 2q(2-p) + 2p - 1)\mu(\widehat{L***})_0 \\
&+ ((p+q)^2 + q^2r^2 + qr^3 + 2pqr + r^2p + 2rq^2 + r^2q - p^2r^2 - pr^3)\mu(\widehat{****})_0 \\
&\geq ((1-p)^2r^2 + (1-p)pr - p^2r^2)\mu(WW\widehat{**})_0 \\
&+ (-2q^2 + 2q(2-p) + 2p - 1)\mu(\widehat{L***})_0 \\
&+ (r^2p - p^2r^2 - pr^3)\mu(\widehat{****})_0 \\
&= (r^2(1-2p) + rp(1-p))\mu(WW\widehat{**})_0 + (-2q^2 + 2q(2-p) + 2p - 1)\mu(\widehat{L***})_0 \\
&+ r^2pq\mu(\widehat{****})_0 \\
&\geq (-r^2p + rp(1-p))\mu(WW\widehat{**})_0 + (-2q^2 + 2q(2-p) + 2p - 1)\mu(\widehat{L***})_0 \\
&+ r^2pq\mu(\widehat{****})_0 \\
&\geq rpq\mu(WW\widehat{**})_0 + (-2q^2 + 2q(2-p) + 2p - 1)\mu(\widehat{L***})_0 + r^2pq\mu(\widehat{****})_0
\end{aligned}$$

Using the quadratic formula, it is easy to see that the polynomial is positive in the following range:

1. $0 \leq p < \frac{1}{2}$, $q > \frac{-\sqrt{p^2+2}-p+2}{2}$.
2. $\frac{1}{2} \leq p < 1$, $0 < q < 1$.

When μ is a stationary distribution and p and q are in the above mentioned range, we can conclude that

1. $\mu(WW\widehat{**})_0 = 0$,
2. $\mu(\widehat{L***})_0 = 0$,

3. $\mu(\widehat{****})_0 = 0$.

This allows us to conclude that $\mu(\widehat{**})_0 = 0$. Finally, using $\widehat{E}_{p,q}\mu(D) = r\mu(\widehat{**})_0$, we see that $\widehat{E}_{p,q}\mu(D) = 0$. Since μ is stationary, we obtain the desired result of Theorem 5.3.3.

Chapter 6

Combinatorics of Conveyor Belts at Airports

Preface

This chapter is based on the following paper:

- Bhasin D., Gupta A., Podder M., Models for arrival of checked bags on conveyor belts at airports. Preprint in preparation.

6.1 Introduction

6.1.1 Overview and an informal description of our model

In this section, we study an interesting combinatorial model for the process of arrival and collection of checked bags from conveyor belts in airports. The moving components in our model are the passengers and the bags. We suppose that our n passengers are labeled 1 to n and are standing in a line adjacent to the belt. Each passenger has exactly one check-in bag. The bags come out in discrete time-steps according to a pre-determined fixed permutation. Whenever a passenger 'matches' with her bag, she collects her bag and leaves. Whenever a passenger still awaiting her bag sees a position 'open up' in 'front' of her because of a co-passenger leaving, she moves to that position. We shall introduce the details of this model formally in the §6.2.

Given a permutation of the checked bags, there is a natural statistic associated with it corresponding to our model – the time taken for everybody to receive their bags. In this section, we study this statistic and find tight lower and upper bounds for the same. In fact, this statistic has a trivial lower bound, namely n, because in order for each of the n passengers to get her bag, *all* the n bags must come out onto the belt. We enumerate all those permutations for which this lower bound is attained, and we show that the number of such permutations actually equals the well-studied *telephone numbers* (see, for example, [101] and [66]), which count the number of matchings in a complete labeled graph. We establish this relation by deriving an interesting bijection between the two. We show that the upper bound on the time required for all n passengers to collect their bags is $\frac{3n}{2}-1$ when n is even and $\frac{3n-1}{2}$ when n is odd.

Needless to say that this model, and the various ways it may be generalized, can be utilized by airline authorities in making the return of checked bags to their passengers a more efficient and smoother process. Even though we consider a fixed, specific order in which the passengers are assumed to stand next to the conveyor belt, our analysis and the results obtained in this section will prove useful even when a random permutation of the positions of the passengers around the belt is considered. One could then obtain the expected value of the (random) time required for all passengers to receive their bags, and attempt to minimize this value in order to provide the airlines' customers a quicker and more hassle-free experience. The simplifying assumptions imposed in our model could be gotten rid of one by one, in order to encompass more realistic scenarios, such as where each passenger is allowed to have a random number of bags checked in (with the support for the probability distribution of this random number being a finite set, such as $\{0,1,2\}$), where some passengers arrive *after* the checked bags have started appearing on the belt, where the inter-arrival times between the appearance of consecutive bags are continuous random variables (for instance, possessing an exponential distribution with a suitable rate parameter), and so on.

It is important to mention here that versions of this problem were first studied by Connor Albach, Jayadev Athreya, Andrew Lim, Qiaoxue Liu, Moumanti Podder and Qiubai Yu as part of the WXML (Washington Experimental Mathematics Lab) project at University of Washington in 2019. They had carried out numerical simulations computing the time taken for various permutations. In a version they studied, the passengers stood at fixed spots and were not allowed to move. In another version, whenever a position became vacant, all passengers moved one step forward simultaneously. Our model in this section differs in the way passengers are allowed to move, namely, a passenger moves only if the position in front of her is empty.

6.1.2 Literature review

Optimising time bounds on baggage handling at airports is a well studied problem and several models and techniques have been proposed and implemented to increase efficiency of checked baggage collection. In [49], the author describes the baggage handling into three types: baggage handling with short time from one flight to the other (transfer baggage), passengers to flight (outbound baggage) and flight to passengers (inbound baggage). The baggage handling process may be optimised at all stages, from checking in to outbound baggage handling. The work in [49] focused on the optimisation of the process with inflow of baggage from the aircraft to the conveyor belt. In [11], the authors focused on optimising the assignment of the bags from each arriving aircraft to an infeed area, which connects to baggage collection carousel. However, our work considers the mathematical models for collection of the baggage collection process and the analysis of time bounds there.

Given a graph $G = (V, E)$, we call a subset M of E a *matching* on G if $e \cap f = \emptyset$ for every $e, f \in M$, i.e. any two distinct edges in M are vertex-disjoint. The number of matchings on the labeled complete graph K_n, denoted $a(n)$, is referred to as the *telephone number* corresponding to K_n, for each $n \in \mathbb{N}$. The first few telephone numbers are $a(1) = 1, 2, 4, 10, 26, 76, 232, 764, 2620, 9496, \ldots$, etc.

It is well known that the sequence $\{a(n)\}_{n \in \mathbb{N}}$ of telephone numbers satisfies the following recurrence relation (see [31] and [66]):

$$a(n) = a(n-1) + (n-1)a(n-2),$$

and it is also known that $a(n)$ satisfies the following formula ([31], [66]):

$$a(n) = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{k} (2k-1)!! = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n!}{2^k (n-2k)! k!}.$$

Other than matchings on a complete labeled graph, the telephone numbers are also known to represent the number of involutions (see [31]) on n objects, and the number of standard Young tableaux (see [13]). In [79], a connection was established between telephone numbers and Hermite polynomials. Graphs are used to model molecules in chemical graph theory (see [101]) and the number of matchings of a given graph is said be its Hosoya index. Thus, if we consider the complete graph on n vertices, then its Hosoya index is equal to the n^{th} telephone number, $a(n)$.

6.2 The description of our model

Our model, which is inherently combinatorial in nature, is made up of two components – passengers and bags. The n passengers are labeled $1, 2, \ldots, n$. Henceforth, we denote the set $\{1, 2, \ldots, n\}$ by the notation $[n]$. Each passenger is assumed to have checked in precisely one bag, and the bag belonging to passenger i is also labeled i, for each $i \in [n]$. Our system comprises an array of order $2 \times n$, in which the first row is made up of the n passengers, and the second row is made up of their n checked bags (in some order). We refer to the first row as the *passenger-row* and the second row as the *bag-row*. These two rows interact with each other as follows. The bag-row is empty at time $t = 0$. We fix a permutation of bags $\sigma \in S_n$, where S_n indicates the set of all permutations of the elements of $[n]$. This permutation reveals the order in which the bags come out on to the conveyor belt. We assume that the bag labeled $\sigma(n)$ comes out first, followed by the bag labeled $\sigma(n-1)$, and so on. At time-step $t = 0$, in the passenger-row, the n passengers stand in an increasing order, represented by the map $\alpha_0^\sigma : [n] \to [n]$ with $\alpha_0^\sigma(i) = i$ (in other words, the i-th position is occupied by the passenger labeled i for each $i \in [n]$).

The configuration of passengers in the passenger-row of our array at time t is represented by the tuple $\alpha_t^\sigma = (\alpha_t^\sigma(i) : i \in [n])$, with $\alpha_t^\sigma(i) \in [n] \cup \{0\}$, whereas the configuration of bags in the bag-row at time t is represented by the tuple $\beta_t^\sigma = (\beta_t^\sigma(i) : i \in [n])$, with $\beta_t^\sigma(i) \in [n] \cup \{0\}$. Here, $\alpha_t^\sigma(i)$ either equals the label of the passenger occupying column i in the passenger-row of the array, or it equals 0 if the i-th column in the passenger-row is empty (and we similarly interpret the values of the coordinates of β_t^σ). Thus, our dynamical system is made up of $\{(\alpha_t^\sigma, \beta_t^\sigma)\}_{t=0}^{2n}$. Note that, since (as mentioned above) the bag-row is empty at time-step $t = 0$, hence $\beta_0^\sigma(i) = 0$ for each $i \in [n]$. The reason why we need only consider time-steps up to $t = 2n$ is because at time-step $2n$, the n^{th}

bag would either have already been picked up or it will be at the n^{th} passenger position which is the last position and hence it must get picked.

We now come to perhaps what can be considered the most important aspect of our model: a complete description of how the tuples α_t^σ and β_t^σ interact with each other, for each time-step t. Suppose we have all the information pertaining to our process up to and including time-step t, i.e. we know $\{(\alpha_s^\sigma, \beta_s^\sigma)\}_{s=0}^t$. Conditioned on this information, we define α_{t+1}^σ and β_{t+1}^σ in two steps, the first of which involves the definition of a couple of 'intermediate' functions, $\widetilde{\alpha}_{t+1}^\sigma$ and $\widetilde{\beta}_{t+1}^\sigma$, that are dependent, respectively, only on α_t^σ and β_t^σ:

$$\widetilde{\alpha}_{t+1}^\sigma(i) = \begin{cases} 0 & \text{if } \alpha_t^\sigma(i) \neq 0 \text{ and } \alpha_t^\sigma(i-1) = 0, \\ \alpha_t^\sigma(i) & \text{if } \alpha_t^\sigma(i) \neq 0 \text{ and } \alpha_t^\sigma(i-1) \neq 0, \\ \alpha_t^\sigma(i+1) & \text{if } \alpha_t^\sigma(i) = 0, \end{cases} \tag{6.2.1}$$

and

$$\widetilde{\beta}_{t+1}^\sigma(i) = \begin{cases} \beta_t^\sigma(i-1) & \text{if } i > 1, \\ \sigma(n-t+1) & \text{if } i = 1 \text{ and } n-t+1 \in [n]. \end{cases} \tag{6.2.2}$$

The significance of introducing the tuples $\widetilde{\alpha}_{t+1}^\sigma = (\widetilde{\alpha}_{t+1}^\sigma(i) : i \in [n])$ and $\widetilde{\beta}_{t+1}^\sigma = (\widetilde{\beta}_{t+1}^\sigma(i) : i \in [n])$ is that they represent how bags and passengers move in our model. They precisely capture the 'speed' with which these two entities move and when they are allowed to move.

Having defined $\widetilde{\alpha}_{t+1}^\sigma$ and $\widetilde{\beta}_{t+1}^\sigma$, we utilize the interactions between them in order to define α_{t+1}^σ and β_{t+1}^σ, as follows:

$$\alpha_{t+1}^\sigma(i) = \begin{cases} 0 & \text{if } \widetilde{\alpha}_{t+1}^\sigma(i) = \widetilde{\beta}_{t+1}^\sigma(i) \text{ or } \widetilde{\alpha}_{t+1}^\sigma(i) = \widetilde{\beta}_{t+1}^\sigma(i+1), \\ \widetilde{\alpha}_{t+1}^\sigma(i) & \text{otherwise}, \end{cases} \tag{6.2.3}$$

and

$$\beta_{t+1}^\sigma(i) = \begin{cases} 0 & \text{if } \widetilde{\alpha}_{t+1}^\sigma(i) = \widetilde{\beta}_{t+1}^\sigma(i) \text{ or } \widetilde{\alpha}_{t+1}^\sigma(i) = \widetilde{\beta}_{t+1}^\sigma(i+1), \\ \widetilde{\beta}_{t+1}^\sigma(i) & \text{otherwise}. \end{cases} \tag{6.2.4}$$

In words, our model can be described as follows. We assume that the conveyor belt is linear in shape, the chute through which the bags emerge is to the extreme left (see Figure 6.1). The passengers stand in a strictly increasing order next to the belt, with passenger 1 standing in the position $(1,1)$ (i.e. column 1 of the passenger-row) which is closest to the chute and passenger n standing in the position $(1,n)$ which is furthest from the chute. The conveyor belt is assumed to

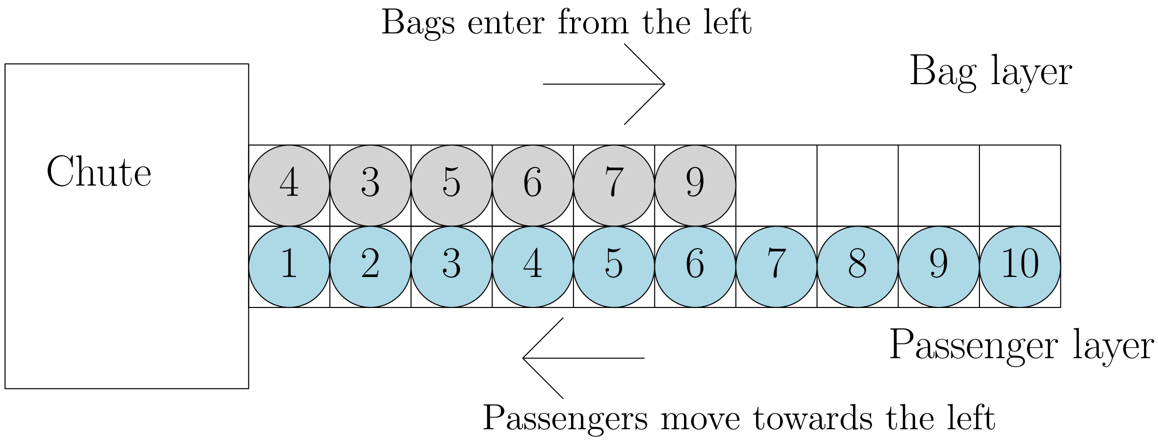


Figure 6.1: A visual representation of our model.

move in discrete time-steps (i.e. the bag labeled $\sigma(n)$ appears on the belt at time-step $t = 1$, the bag labeled $\sigma(n-1)$ appears at time-step $t = 2$ and so on). We say that a passenger 'matches' with her bag at some time-step t when one of the following scenarios arises at time-step $t-1$:

1. a passenger is in position $(1, i+1)$ (i.e. the $(i+1)$-st column of the passenger-row) while her bag is in position $(2, i)$ (i.e. the i-th column of the bag-row),

2. a passenger is in position $(1, i+1)$, the position $(1, i)$ (to her immediate left) is empty and her bag is in position $(1, i-1)$.

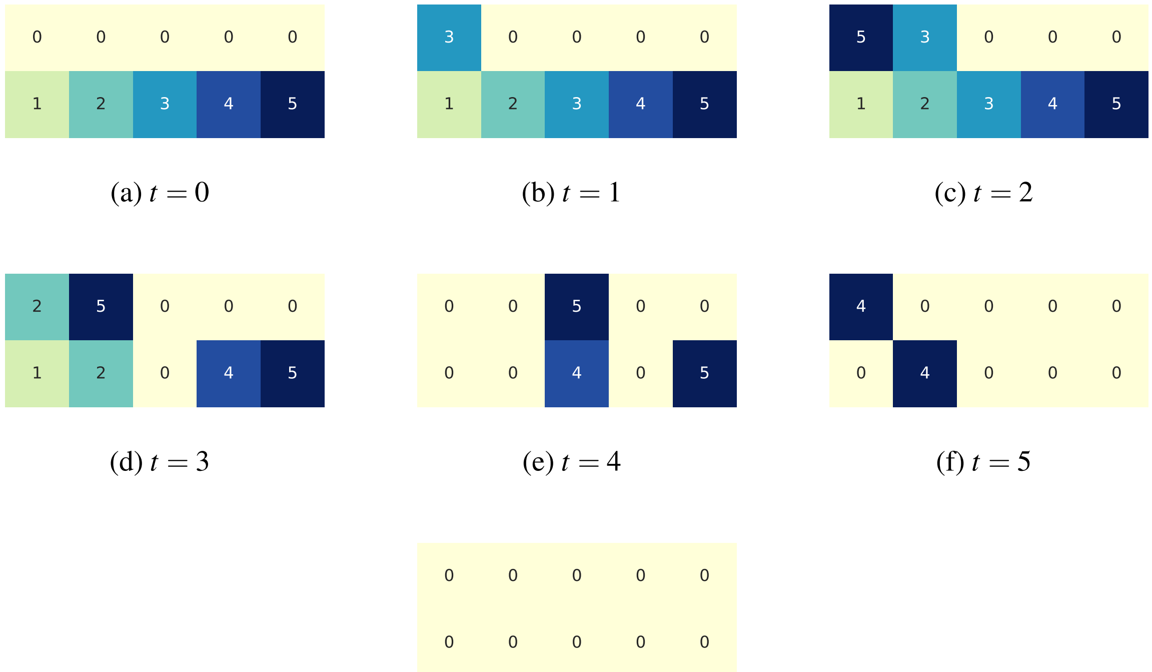
Once a passenger matches with her bag, she leaves. Otherwise, whenever there is an empty position to the left of a passenger, she moves to occupy this vacated position. We note that this intuition even works when the passenger is at the first position (closest to the chute), by imagining a bag to be at position '-1'.

We illustrate the dynamics of the model by an example in Figure 6.2, where we consider $n = 5$ and the permutation σ to be the tuple $(4, 1, 2, 5, 3)$.

6.2.1 Main results of this section

In this subsection, we state the main results of our work. We begin with some relevant definitions.

Definition 6.2.1. *Given a permutation of bags $\sigma \in S_n$, we denote by $T_{1,empty}(\sigma)$ the time taken for the belt to become empty for the first time after time-step $t = 1$, when the belt is fed the bags according to the permutation σ. We let $T(\sigma)$ be the time taken for every passenger to receive her bag, when the arrival of bags on the belt is initiated in accordance with σ. Let $T_{empty}(n) =$* $\max\{T(\sigma) : \sigma \in S_n\}$.



(a) $t = 0$ (b) $t = 1$ (c) $t = 2$

(d) $t = 3$ (e) $t = 4$ (f) $t = 5$

(g) $t = 6$

Figure 6.2: Running the process for $n = 5$ and the permutation of bags = $(4, 1, 2, 5, 3)$.

Definition 6.2.2. *Let $\mathscr{C}_{n,k}$ denote the set of all those permutations σ of the bags for which $T_{1,empty}(\sigma) = k$. Similarly, let $\mathscr{D}_{n,k}$ be the set of all those permutations σ of the bags that satisfy $T(\sigma) = k$.*

Equipped with these definitions, we now state our main results:

Theorem 6.2.3. *For all $n, k \in \mathbb{N}$, we have $|\mathscr{C}_{n+k,n}| = k!|\mathscr{C}_{n,n}|$.*

Theorem 6.2.3 gives us the value of $|\mathscr{C}_{n+k,n}|$ given the value of $|\mathscr{C}_{n,n}|$. In fact, as is suggested by the $k!$ factor in the right hand side of the identity presented in Theorem 6.2.3, we show that there is a bijection between $\mathscr{C}_{n+k,n}$ and $S_{[n+1,n+k]} \times \mathscr{C}_{n,n}$, where $S_{[n+1,n+k]}$ is the symmetric group of $[n+1, n+k] = \{n+1, n+2, \ldots, n+k\}$.

Theorem 6.2.4. *For $n \in \mathbb{N}$, the cardinality of $\mathscr{D}_{n,n}$ equals the n^{th} telephone number, i.e. $|\mathscr{D}_{n,n}| = a(n)$.*

We prove Theorem 6.2.4 by constructing an explicit bijection between the set of involutions of S_n and the set $\mathscr{D}_{n,n}$. We prove this bijection via an elaborate induction-based argument where we accomplish a step-by-step patching of information using Lemma 6.4.1 to obtain our desired result.

Theorem 6.2.5. *For all $n \geq 2, T_{empty}(n) = T_n$, where* $T_n = \begin{cases} \frac{3n}{2} - 1, & \text{if } n \text{ is even} \\ \frac{3n-1}{2}, & \text{if } n \text{ is odd} \end{cases}$.

Theorem 6.2.5 tells us that the upper bound on the time taken for any permutation of bags is T_n. We prove this theorem using an induction based proof as well.

6.3 Proof of Theorem 6.2.3

Let P_i be the i^{th} passenger and let B_j be the bag of the j^{th} passenger. We begin with the following lemma which tells that the j^{th} passenger can not 'act' before time j.

Lemma 6.3.1. *Given any permutation $\pi \in S_n$, let $t(\pi, j)$ denote the first time that passenger P_j moves or picks up her bag and leaves. Then $t(\pi, j) \geqslant j$ for each $j \in \{1, 2, \ldots, n\}$.*

Proof. It is straightforward to see this for $i = 1$, since either the first bag to come out is B_1, in which case P_1 leaves with her bag at time step 1, or else bag B_1 comes out at time step $t \geqslant 2$, and P_1 has to leave at t. Suppose we have proved this claim for $j = i$, for some $i \in [n]$. Let us consider P_{i+1}. There are two possible scenarios:

1. P_{i+1} remains at her initial position throughout and picks up her bag from there. This means that bag B_{i+1} has to traverse all the way up to the $(i+1)$-st position to be picked up, and this will take time at least $i+1$, given the speed of the conveyor belt. This yields $t(\pi, i+1) \geqslant i+1$ in this case.

2. P_{i+1} does *not* remain in her original / initial position when picking up her bag B_{i+1}. This means that she has moved at least one position to the left, which can only happen because either P_i herself has moved one position to the left, or has left after picking up her bag *before* B_{i+1} reaches P_{i+1}. However, P_{i+1} moves to the position to her left one time step *after* either P_i moves or picks up her bag and leaves. This means that $t(\pi, i+1) = t(\pi, i) + 1$. Consequently, implementing the induction hypothesis that states that $t(\pi, i) \geqslant i$, we conclude that $t(\pi, i+1) \geqslant i+1$. □

Using the above lemma, we obtain the following:

Lemma 6.3.2. *The i-th passenger P_i cannot pick up her bag B_i before time step i.*

Proof. This follows directly from Lemma 6.3.1, since for every permutation π of the bags, the time it takes for P_i to pick up B_i is bounded below by $t(\pi, i)$ by definition of $t(\pi, i)$. □

Theorem 6.2.3 is a direct consequence of the following:

Theorem 6.3.3. *The set $\mathscr{C}_{n+k,n}$ comprises permutations of the form $\pi = (i_1, i_2, \ldots, i_{n+k})$ where $(i_{k+1}, \ldots, i_{n+k})$ is a permutation of $[n]$ that belongs to $\mathscr{C}_{n,n}$, and $(i_1, \ldots, i_k)$ is a permutation of $n+1, \ldots, n+k$.*

Proof. We first establish the above claim for $k = 1$, for a better understanding of what is going on. Suppose $\pi = (i_1, i_2, \ldots, i_{n+1})$ is a permutation such that $i_1 \neq n+1$. Then we claim that π cannot belong to $\mathscr{C}_{n+1,n}$. To see this, we first note that since $i_1 \neq n+1$, hence $n+1 \in \{i_2, \ldots, i_{n+1}\}$, which means that by time step n, the bag B_{n+1} belonging to passenger P_{n+1} is out on the conveyor belt. Therefore, in order for $T_{1,\text{empty}}(\pi)$ to equal n, the bag B_{n+1} needs to be picked up by passenger P_{n+1} by time step n, and this cannot happen by Lemma 6.3.2. We thus conclude that $i_1 \neq n+1$ implies that such a permutation π cannot be in $\mathscr{C}_{n+1,n}$. On the other hand, when $i_1 = n+1$ and $(i_2, \ldots, i_n)$ is a permutation in $\mathscr{C}_{n,n}$, then by definition of $\mathscr{C}_{n,n}$, by time step n the conveyor belt becomes empty and bag $B_{i_1} = B_{n+1}$ is still not out on the belt. Hence such a permutation does belong to $\mathscr{C}_{n+1,n}$. This shows us that there is a bijection $\varphi : \mathscr{C}_{n,n} \to \mathscr{C}_{n+1,n}$ defined by

$$\varphi(\pi) = (n+1, i_1, i_2, \ldots, i_n) \text{ where } \pi = (i_1, i_2, \ldots, i_n) \in \mathscr{C}_{n,n},$$

thus yielding $\left|\mathscr{C}_{n+1,n}\right| = |\mathscr{C}_{n,n}|$. Next, we show that

$$\mathscr{C}_{n+k,n} = \{(i_1, i_2, \ldots, i_{n+k}) : (i_1, \ldots, i_k) \text{ a permutation of } n+1, \ldots, n+k \text{ and } (i_{k+1}, \ldots, i_{n+k}) \in \mathscr{C}_{n,n}\}.$$

If $\pi = (i_1, \ldots, i_{n+k})$ is a permutation of $1, \ldots, n+k$ that does not satisfy the above condition, then it must mean one of two things:

1. The first is that $(i_1, \ldots, i_k)$ is *not* a permutation of $n+1, \ldots, n+k$. This means that there exists some $1 \leqslant j \leqslant k$ such that $n+j \in \{i_{k+1}, \ldots, i_{n+k}\}$. This means that the bag B_{n+j} comes out onto the conveyor belt by time step n, and hence, for $(i_1, \ldots, i_{n+k})$ to be in $C_{n+k,n}$, it needs to be picked up by P_{n+j} within time step n. However, by Lemma 6.3.2, this is impossible, and hence, bag B_{n+j} stays on the belt until and including time step n, and thus the belt cannot become empty by time step n. This leads to a contradiction, which means that it *cannot* happen that $(i_1, \ldots, i_k)$ is *not* a permutation of $n+1, \ldots, n+k$.

2. The first does not happen, and the second is that $(i_{k+1}, \ldots, i_{n+k})$ is a permutation of $1, \ldots, n$ but is not in $\mathscr{C}_{n,n}$. In this case, even though only the bags $B_1, \ldots, B_n$ are the ones that come out onto the conveyor belt by time step n, the first time that the belt becomes empty is *not* equal to n, by definition in $\mathscr{C}_{n,n}$.

This tells us that $\mathscr{C}_{n+k,n} \subset \{(i_1,i_2,\ldots,i_{n+k}) : (i_1,\ldots,i_k)$ a permutation of $n+1,\ldots,n+k$, $(i_{k+1},\ldots,i_{n+k}) \in \mathscr{C}_{n,n}\}$. On the other hand, every permutation $(i_1,i_2,\ldots,i_{n+k})$ in which $(i_1,\ldots,i_k)$ is a permutation of $n+1,\ldots,n+k$ and $(i_{k+1},\ldots,i_{n+k}) \in \mathscr{C}_{n,n}$ is clearly in $\mathscr{C}_{n+k,n}$, since the only bags that come out onto the conveyor belt by time step n are $B_1,\ldots,B_n$, and the first time that the conveyor belt becomes empty is n because $(i_{k+1},\ldots,i_{n+k}) \in \mathscr{C}_{n,n}$. Thus we have the proof. □

Proof of Theorem 6.2.3. As a consequence of Theorem 6.3.3, we conclude that

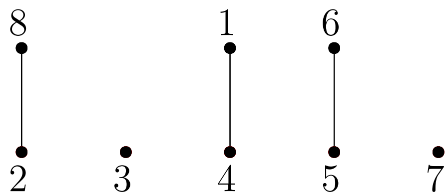
$$\begin{aligned}\left|\mathscr{C}_{n+k,n}\right| &= \text{no. of permutations } (i_1,\ldots,i_k) \text{ of } \{n+1,\ldots,n+k\} \\ &\quad \times \text{no. of permutations } (i_{k+1},\ldots,i_{n+k}) \text{ in } \mathscr{C}_{n,n} = k! \cdot |\mathscr{C}_{n,n}|,\end{aligned}$$

as claimed above. □

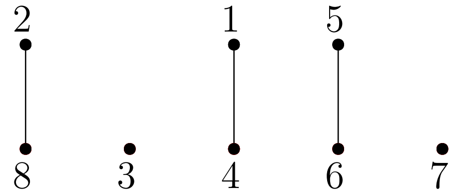
6.4 Proof of Theorem 6.2.4

In this section, we first explicitly describe the bijection used in proving Theorem 6.2.4 and then give the proof. We begin by reiterating here that it is easy to see that for n passengers, a lower bound for the time taken for every passenger to get her bag is n. In fact, the identity permutation attains this lower bound. It is then a natural question to count the number of permutations which attain this lower bound. Theorem 6.2.4 completely answers this question establishing a connection with telephone numbers (see Section 6.1.2). In particular, we prove that the number of permutations of bags which attain the lower bound of n is equal to the n^{th} telephone number, $a(n)$. We do this by constructing a bijection between the set of matchings on the complete labeled graph K_n, say $\mathscr{M}_n$, and the set of permutations of n bags which attain the lower bound, $\mathscr{D}_{n,n}$. Given such a matching, we construct a corresponding permutation of bags by following some steps. For example, when $n=8$ and the matching is $\{\{2,8\},\{3\},\{1,4\},\{5,6\},\{7\}\}$, we demonstrate the permutation which our bijection outputs, in Figure 6.3.

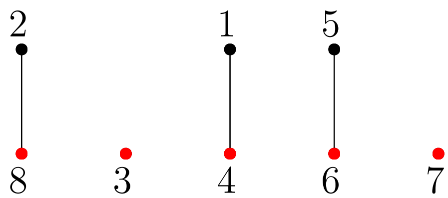
Formally, this bijection, $\phi : \mathscr{M}_n \to \mathscr{D}_{n,n}$ is defined as follows: We begin with a matching $M \in \mathscr{M}_n$ on labeled K_n. We first construct an *involution* $g_M : [n] \to [n]$ defined by $g_M(i) = j$ where j is the unique element of $[n]$ such that $\{i,j\} \in M$ (in what follows, we will identify a matching M with its corresponding involution g_M). We now recursively construct the permutation $\phi(M) = \sigma_M$. We start by putting $\sigma_M(n) = g_M(n)$ and assign remaining values to σ_M recursively as follows: suppose we have assigned the values of $\sigma_M(i+1),\sigma_M(i+2),\ldots,\sigma_M(n)$. Then, we let $A_i = [n] \setminus \{\sigma_M(i+1),\sigma_M(i+2),\ldots,\sigma_M(n)\}$ and put $a_i = \max A_i$. If $g_M(\sigma_M(i+1)) = a_i$, we define $\sigma_M(i) = a_i$, else, we define $\sigma_M(i) = g_M(a_i)$.



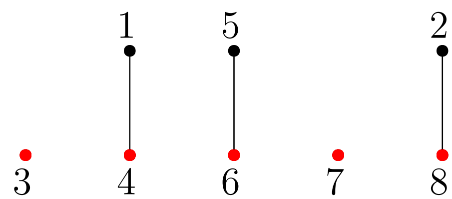
(a) We start with the matching and write it linearly like this.



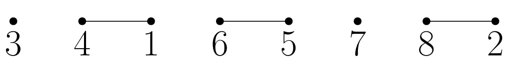
(b) All the edges are rearranged so that the bigger numbers are on the lower row.



(c) We now focus on the number on the lower row.



(d) The edges are now rearranged so that numbers in focus (of the lower row) are in increasing order.



(e) The edges are now rotated clockwise.

(f) We now forget the edges and obtain the output of our bijection.

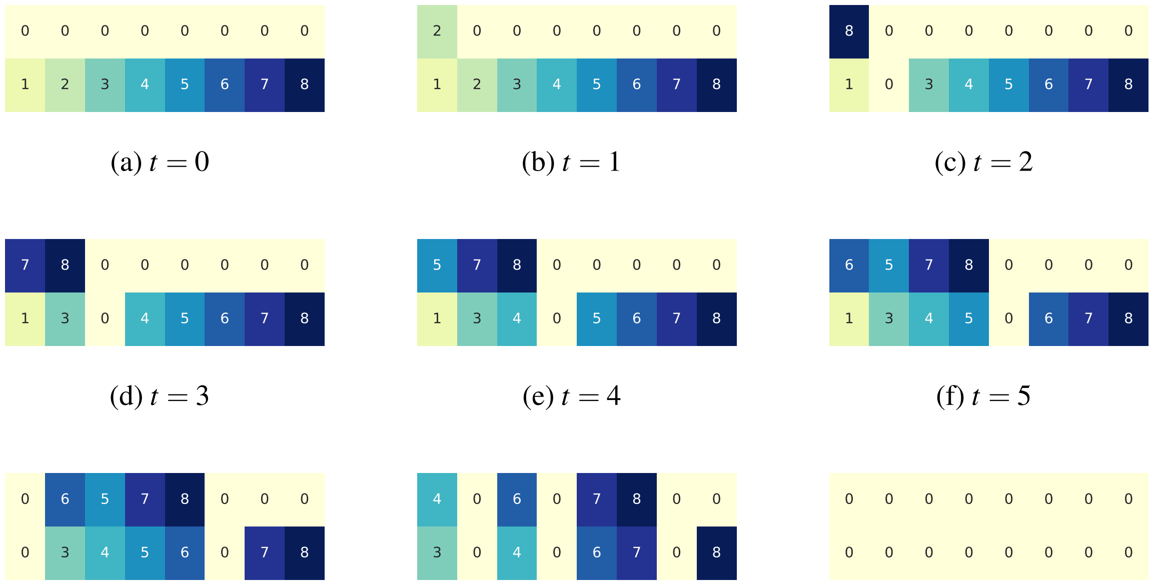
Figure 6.3: The permutation created by our bijection when the matching is as in Figure 6.3a

Given this map, the task at hand is twofold: (i) to show that this map is well-defined, that is, every permutation it outputs takes time n (ii) to establish that it is indeed a bijection, that is, to show that any permutation which takes time n, is uniquely obtained by a matching via this map. The crux of the idea for our proof is that we run the belt k many times where in the k^{th} run, we initialise the belt with the bags $\sigma_M(n-k+1),\ldots,\sigma_M(n)$ (in particular, after time $t=k$, no bag comes out) and we inductively gather information. Analogous to α_t^σ which was defined in Section 6.2, by $\alpha_{t,k}^\sigma(i)$ we mean the passenger occupying the i^{th} passenger position at time t, when the belt is initialised with the bags $\sigma_M(n-k+1),\sigma_M(n-k+2),\ldots,\sigma_M(n)$. We then use the following 'patching lemma' to step by step gather information and finally prove the required result.

Lemma 6.4.1. *For any $\sigma\in S_n, t\in[2n]$ and $k\in[n]$ satisfying $t\geq k$, we have, for any $i\in[t-k+1,n]$*

$$\alpha_t^\sigma(i)=\alpha_{t,k}^\sigma(i).$$

Lemma 6.4.1 is a straightforward consequence of the definition of our model but it comes in



(a) $t = 0$ (b) $t = 1$ (c) $t = 2$

(d) $t = 3$ (e) $t = 4$ (f) $t = 5$

(g) $t = 6$ (h) $t = 7$ (i) $t = 8$

Figure 6.4: Running the permutation $(3,4,1,6,5,7,8,2)$ obtained by the matching in Figure 6.3a

handy in proving our result by allowing induction to happen.

Definition 6.4.2. *Suppose $\sigma \in S_n$. Let $T(i,\sigma)$ be the time taken by the bag labeled $\sigma(i)$ to get picked when the belt is initialised with n bags according to σ. On the other hand, let $T_k(i,\sigma)$, such that $i \geq n-k+1$, be defined analogously when the belt gets initialised with bags $\sigma(n-k+1),\ldots,\sigma(n)$. We note that $T(\sigma) = \max\{T(i,\sigma) : i \in [n]\}$ and we define $T_k(\sigma) = \max\{T_k(i,\sigma) : i \in [n-k+1,n]\}$.*

Corollary 6.4.3. *For $i \in [n-k+1,n]$ and for $\sigma \in S_n$, we must have that $T(i,\sigma) = T_k(i,\sigma)$.*

Proof. Using Lemma 6.4.1, we see that $\alpha_t^\sigma(i) = \alpha_{t,k}^\sigma(i)$ for every $t \leq T_k(i,\sigma)$. This shows that $T(i,\sigma) = T_k(i,\sigma)$. □

Definition 6.4.4. *We define a function $f : [n] \times S_n \to \{0,1\}$ as follows:*

$$f(i,\sigma) = \begin{cases} 1 & \text{if } \sigma(i) \geq \sigma(j) \forall j \leq i \\ 0 & \text{otherwise} \end{cases} \tag{6.4.1}$$

We now state the following four statements which form the crux of our proof.

1. $P_1(k)$: Given $n \in \mathbb{N}, k \in [n]$ and $\sigma \in S_n$, we have $f(n-k+1,\sigma) = 1 \Rightarrow T(n-k+1,\sigma) = n$.

2. $P_2(k)$: Given $n \in \mathbb{N}, k \in [n]$ and $\sigma \in S_n$, we have $f(n-k+1.\sigma) = 0 \Rightarrow T(n-k+1,\sigma) < n$.

3. $P_3(k)$: Given $n \in \mathbb{N}, k \in [n]$ and $\sigma \in S_n$ if $f(n-k+1,\sigma) = 0$ then, for $k \le t \le n$, we have:

   (a) If $\sigma(n-k+1)$ has been picked up then, $\alpha_t^\sigma(t-k+1) = 0$.

   (b) If $\sigma(n-k+1)$ has not been picked up then, $\alpha_t^\sigma(t-k+1) \neq 0$.

4. $P_4(k)$: Given $n \in \mathbb{N}, k \in [n]$ and $\sigma \in S_n$, we have $f(n-k+2,\sigma) = 0 \Rightarrow f(n-k+1,\sigma) = 1$.

For technical reasons concerning P_4 which will be clear in a while, we extend the definition of f to $[n+1] \times S_n$ and put $f(n+1,\sigma) = 1$. We use the above statements to prove the following technical lemmas which are important for the proof of our theorem.

Lemma 6.4.5. *Let $n \in \mathbb{N}, k \in [n-1], \sigma \in S_n$. Suppose $\bigwedge_{j=1}^3 \bigwedge_{i=1}^k P_j(i)$ and $\bigwedge_{i=1}^{k+1} P_4(i)$ are true. Then, $\bigwedge_{j=1}^3 P_j(k+1)$ is also true.*

Proof. We first prove $P_1(k+1)$ and $P_2(k+1)$. We initialise the belt with k bags namely $\sigma(n-k+1),\ldots,\sigma(n)$. Using $\bigwedge_{i=1}^k P_3(i)$ and $\bigwedge_{i=1}^k P_1(i)$, we know that for $t \le n-1$, whenever a bag gets picked, the passenger position corresponding to the current position of the bag will be empty. Since, the number of bags that have been picked must be equal to the number of passengers that have left, we must have that all other positions are occupied by passengers at any given time. Recall that

$$A_k = [n] \setminus \{\sigma(n-k+1),\ldots,\sigma(n)\}.$$

The discussions in the previous paragraph yield , in particular, that at time $t \le n-1$, $\alpha_{t,k}^\sigma(i) = a_i$ for $i \le t-k$, where, recall that $a_i = \max A_i$. This equation can be seen by observing that the bag $\sigma(n-k+1)$ at time t will be at position $t-k+1$. We can show that a similar phenomenon happens for the case $t = n$. To see this, we study $\alpha_{n-1,k}$. Note that at $t = n-1$, we have shown that the bags $a_1,\ldots,a_{n-k}$ are at positions $1,\ldots,n-k$ respectively. Note also that for $i \le n-k-1$, $\alpha_{n,k}^\sigma(i) = \alpha_{n-1,k}^\sigma(i) = a_i$. We have the following two cases:

1. $\underline{f(n-k+1,\sigma) = 1}$. Using $P_1(k)$, we obtain that $\alpha_{n-1,k}^\sigma(n-k) \neq 0$. This clearly means that

$$\alpha_{n-1,k}^\sigma(n-k) = a_{n-k} = \alpha_{n,k}^\sigma(n-k).$$

2. $\underline{f(n-k+1,\sigma) = 0}$. By $P_4(k)$, this means that $f(n-k+2,\sigma) = 1$. Using $P_1(k)$, we obtain that $\alpha_{n-1,k}^\sigma(n-k+1) = a_{n-k}$ and using $P_2(k)$, we obtain that $\alpha_{n-1,k}^\sigma(n-k) = 0$. By the dynamics of our model, $\alpha_{n,k}^\sigma(n-k) = a_{n-k}$ as required.

This gives us that for $t \le n$ and $i \le t-k$,

$$\alpha^{\sigma}_{t,k}(i) = a_i. \tag{6.4.2}$$

Suppose $a_l = \sigma(n-k)$. Note that we have $\alpha^{\sigma}_{l+k,k}(l) = a_l$. This means that when the the belt is initiated with the bags $\sigma(n-k), \dots, \sigma(n)$, at time $t = l+k$, the passenger a_l will be at position l. But the bag $\sigma(n-k)$ will also be at the same position at the same time leading to $T(n-k,\sigma) = l+k \le n$. Also, $T(n-k,\sigma) = n \Leftrightarrow l = n-k \Leftrightarrow f(n-k,\sigma) = 1$ establishing $P_1(k+1)$ and $P_2(k+1)$. Note that from the previous line we can also conclude that $\alpha^{\sigma}_{l+k,k+1}(l) = 0$. We now prove $P_3(k+1)$. Thus, we proceed under the assumption that $f(n-k,\sigma) = 0$. Using $P_4(k+1)$, we obtain that $f(n-k+1,\sigma) = 1$. Note that we have already shown that $P_3(k+1)$ is true for $t \le T(n-k+1,\sigma)$. Suppose that $P_3(k+1)$ is true for some t with $T(n-k,\sigma) < t < n$. We know that $f(n-k+1,\sigma) = 1 \Rightarrow \alpha^{\sigma}_{t,k+1}(t-k+1) = \alpha^{\sigma}_{t,k}(t-k+1) \neq 0$ using $P_1(k)$ and Lemma 6.4.1. By the induction hypothesis, we know that $\alpha^{\sigma}_{t,k+1}(t-k) = 0$. By the rules of our model, this simply means that $\alpha^{\sigma}_{t+1,k+1}(t-k) = 0$ completing the proof. □

Lemma 6.4.6. *Let $n \in \mathbb{N}, k \in [2, n-1], \sigma \in S_n$. Suppose $\bigwedge_{i=1}^{k} P_j(i)$ for $j = 1,2,3,4$ is true. Then, $P_4(k+1)$ is also true.*

Proof. We are given that $f(n-k+1) = 0$ and we want to show that $f(n-k) = 1$. As in the proof of the previous lemma, we know that for $t \le n-1$, whenever a bag has been picked, the position corresponding to it is empty. This means that passengers are present at all other position. Also, note that $f(n-k+1,\sigma) = 0 \Rightarrow f(n-k+2,\sigma) = 1$ using $P_4(k)$. When the belt is initialised with k bags $\sigma(n-k+1), \dots, \sigma(n)$, this means that at time $t = n-1$, $\sigma(n-k+2)$ has not been picked and $\sigma(n-k+1)$ has been picked. Using this discussion, we obtain that

1. $\alpha^{\sigma}_{n-1,k}(i) = a_i$ for $i \le n-k-1$.
2. $\alpha^{\sigma}_{n-1,k}(n-k) = 0 = \alpha^{\sigma}_{n-1,k+1}(n-k) = \alpha^{\sigma}_{n-1}(n-k)$ using Lemma 6.4.1.
3. $\alpha^{\sigma}_{n-1,k}(n-k+1) = a_{n-k} = \alpha^{\sigma}_{n-1,k+1}(n-k+1) = \alpha_{n-1}(n-k+1,\sigma)$ using Lemma 6.4.1.

Note that 2 and 3 above mean that when the belt is initiated with all the bags, a_{n-k} will be at position $n-k$. Since $T(\sigma) = n$, a_{n-k} must be picked at time less than or equal to n. This means that $\sigma(n-k) = a_{n-k}$ (because the bag $\sigma(n-k)$ will be at position $n-k$ at time n). This proves that $f(n-k,\sigma) = n$.

□

Corollary 6.4.7. *Let $n \in \mathbb{N}, k \in [n-1], \sigma \in S_n$. Suppose that $\bigwedge_{i=1}^{k} P_j(i)$ for $j = 1,2,3,4$. Then, $\bigwedge_{j=1}^{4} P_j(k+1)$ is also true.*

Proof of Theorem 6.2.4. First of all,we show that ϕ is a well-defined map. That is, given an involution $g \in \mathscr{M}_n$, we need to show that $T(\phi(g)) = n$. We put $\phi(g) = \sigma$. Note that $f(-,\sigma)$ satisfies $\bigwedge_{i=1}^{n} P_4(i)$ by the definition of the map ϕ. Let us show that $P_1(1), P_2(1), P_3(1)$ are true for σ:

1. We are in the case where $\sigma(n) = n$. This clearly means that $T_1(n,\sigma) = n$.

2. This case is also straightforward.

3. Let $\sigma(n) = i \neq n$. Note that $T_1(n,\sigma) = i$. Note also that at time t, the bag will be or would have been (if it was not picked) at position t. The statement is clearly true for $t \leq i$. Suppose that it is true for some t such that $i < t < n$. We consider $t+1$. We are given that $\alpha^{\sigma}_{t,1}(t) = 0$ and $\alpha^{\sigma}_{t,1}(t+1) \neq 0$ (because only one bag has been picked). From this, $\alpha^{\sigma}_{t+1,1}(t+1) \neq 0$ clearly follows by the definition.

We can inductively establish $P_1(i), P_2(i), P_3(i)$ for all $i \neq n$ using Lemma 6.4.5 proving $T(\sigma) = n$. This proves that ϕ is well-defined. From the definition of ϕ it easily follows that it is injective. We now show that it is surjective. Suppose we have a $\sigma \in \mathscr{D}_{n,n}$, that is, $T(\sigma) = n$. $P_1(1), P_2(1), P_3(1)$ and $P_4(1)$ are clearly true for σ. We will now prove $P_4(2)$. Suppose that $f(n,\sigma) = 0$. We want to show that $f(n-1,\sigma) = 1$. This means that we are given that $\sigma(n) = i < n$ and we want to establish that $\sigma(n-1) = n$. However, note that $\sigma(n) = i$ gets picked up at time $i \Rightarrow$ the passenger labelled $i+1$ moves at time $i+1$ and inductively, we conclude that passenger n moves at time n and reaches $n-1$. But $\sigma(n-1)$ will be at position $n-1$ at time n. Since, $T(\sigma) = n$, this means that $\sigma(n-1) = n$ as required. Using Lemma 6.4.5, we have that $P_1(2), P_2(2), P_3(2)$ are true. Consequently, using Corollary 6.4.7 and induction, we obtain $\bigwedge_{j=1}^{4} \bigwedge_{i=1}^{n} P_j(i)$. Note that using $\bigwedge_{i=1}^{n} P_4(i)$, we construct the involution $g \in \mathscr{M}_n$ for which $\phi(g) = \sigma$ as follows: $g(\sigma(i+1)) = \sigma(i)$ and $g(\sigma(i)) = \sigma(i+1)$ if $f(i+1,\sigma) = 0$ else $g(\sigma(j)) = \sigma(j)$. This is clearly an involution for which $\phi(g) = \sigma$. $\square$

6.5 Proof of Theorem 6.2.5

One of the key lemmas in proving Theorem 6.2.5 is that

1. at time n, the passenger labeled n either leaves or moves for the first time,

2. she keeps on moving until she has picked her bag and left.

Formally, we present this lemma in the form of Lemma 6.5.5.

Using Lemma 6.5.5, it is easy that the time taken for the n^{th} passenger to get her bag is bounded above by T_n (see Theorem 6.2.5). This is because, since the n^{th} passenger starts moving at time n, and keeps on moving, at time T_n, the n^{th} passenger and the last bag will be at the same position (which is in the middle of the belt) and hence, the n^{th} passenger must leave.

The discussion in the preceding paragraph allows induction to come into the picture. We can now essentially 'remove' (this is made precise by the lemmas that follow) the n^{th} passenger and her bag. The only catch here is that, now, at the time-step when the bag of the n^{th} passenger was supposed to come out, no bag will come out. We analyze such 'lagged' permutations of bags as well and show that they also 'behave well' in Lemma 6.5.8. The following definition is needed for stating this lemma:

Definition 6.5.1. *Let $\sigma \in S_n$ be a permutation of bags. By inserting a lag at the i^{th} place, we mean the permutation of $[0,n]$, denoted $\widetilde{\sigma}_i$, defined by* $\widetilde{\sigma}_i(j) = \begin{cases} \sigma(j) & \text{if } j > i, \\ 0 & \text{if } j = i, \\ \sigma(j+1) & \text{if } 0 \leq j < i. \end{cases}$ *Moreover, we define $T(\widetilde{\sigma})$ to be the time taken for each passenger to get her bag when the belt is initiated according to $\widetilde{\sigma}$.*

With this lemmas in hand, we will be equipped to prove Theorem 6.2.5. We begin with the following definition and proofs of some lemmas:

Definition 6.5.2. *Let $\sigma \in S_n$ and a time step t and $p \in [n]$ be given. Suppose we initiate the belt with first k bags of σ. We say that a configuration of passengers is left aligning till position p at time t if at least one of the following holds:*

1. *For every $i \leq p$, we have $\alpha^{\sigma}_{t,k}(i) \neq 0$.*
2. *Let $m = \min\{i \leq p : \alpha^{\sigma}_{t,k}(i) = 0\}$. Then, $\forall i > m$ and $i \leq p-1, \alpha^{\sigma}_{t,k}(i) \neq 0 \Rightarrow \alpha^{\sigma}_{t,k}(i+1) = 0$.*

When p and t will be clear from the context, we will just say that the configuration of passengers is left-aligning.

The reason for the nomenclature as we will see is that left-aligning configurations makes sure that once the bags are past position p, then, in subsequent time steps, all the passengers who can move, *keep moving* until they can longer move in any future time step. Globally, in the end, this looks like as if all the passengers have been left-aligned

We now have:

Lemma 6.5.3. *Suppose we initiate the belt with first k bags of a permutation σ. Suppose at time t, the passengers till* $\min\{t-k,n\}$ *are left-aligning. Then, at time* $t+1$*, the passengers till position* $\min\{t-k+1,n\}$ *are left aligning.*

Proof. At time t, the bag $\sigma(n-k+1)$ is at position $t-k+1$. The passengers under consideration for us are $1,2,\ldots,t-k+1$. This means that only the passenger $t-k+2$ can potentially leave at time $t+1$. We now consider the following cases:

1. We first assume that for every $i \leq \min\{t-k,n\}$, we have $\alpha^{\sigma}_{t,k}(i) \neq 0$. Because of the discussion in the previous paragraph, we can conclude that for every $i \leq \min\{t-k,n\}$, we have $\alpha^{\sigma}_{t+1,k}(i) \neq 0$. This clearly means that at time step $t+1$, passengers till $\min\{t-k+1,n\}$ are left aligning.

2. In this case, we assume that there is an $i \leq \min\{t-k,n\}$ such that $\alpha^{\sigma}_{t,k}(i) = 0$. We let $m = \min\{i \leq \min\{t-k,n\} : \alpha^{\sigma}_{t,k}(i) = 0\}$. At time $t+1$, the minimum position which is empty will be m or $m+1$. This is because if $\alpha^{\sigma}_{t,k}(m+1) = 0$ then, clearly, $\alpha^{\sigma}_{t+1,k}(m) = 0$ and on the other hand, if $\alpha^{\sigma}_{t,k}(m+1) \neq 0$ then, $\alpha^{\sigma}_{t+1,k}(m) = \alpha^{\sigma}_{t,k}(m+1) \neq 0$ and $\alpha^{\sigma}_{t+1,k}(m+1) = 0$. We now observe that for $j > m$ and if we obtain such a $j-1 \geq m$ having $\alpha^{\sigma}_{t,k}(j-1) \neq 0$ and $\alpha^{\sigma}_{t,k}(j) \neq 0$ then it is a contradiction to our assumption. From this, we can conclude that at time step $t+1$, every passenger at position $j > m$ takes a step forward. That is , for every $j > m$ and $j \leq \min\{t-k,n\}$, we have $\alpha^{\sigma}_{t,k}(j) \neq 0 \Rightarrow \alpha^{\sigma}_{t+1,k}(j-1) = \alpha^{\sigma}_{t,k}(j)$. This clearly means that at time step $t+1$, the passengers till $\min\{t-k,n\}$ are left aligning. Note that if $t-k \geq n$ then there is nothing more to prove. Consequently, we now assume that $t-k < n$. Suppose that $\alpha^{\sigma}_{t,k}(t-k) \neq 0$. This means that $\alpha^{\sigma}_{t+1,k}(t-k) = 0$. From this, we can conclude that no matter what, $\alpha^{\sigma}_{t+1,k}(t-k+1)$ is, the passenger configuration till $t-k+1$ at time $t+1$ is left aligning. On the other hand, suppose that $\alpha^{\sigma}_{t,k}(t-k) = 0$. If $m = t-k$, there is nothing to prove. If $m < t-k$, then it is easy to see that $\alpha^{\sigma}_{t+1,k}(t-k-1) = 0$. This means that $(\alpha^{\sigma}_{t+1,k}(t-k), \alpha^{\sigma}_{t+1,k}(t-k+1)) \in \{(0,0),(0,1),(1,0)\}$. This finishes the proof.

□

Lemma 6.5.4. *Let $\sigma \in S_n$ and $k \in [2,n]$. Suppose that we initiate the belt with first k bags of σ. Then, at any time step $t \geq k-1$, the passenger configuration till $t-k+2$ is left aligning.*

Proof. We prove this lemma by induction on k. Let us first consider the base case, namely, $k=2$. We want to show that at every time step t, the passengers till t are left aligning. We divide the analysis into the following cases:

1. $\underline{\sigma(n) = \sigma(n+1)}$: In this case, at every time step t satisfying $1 \leq t < \sigma(n)$, none of the passengers have picked their bags. Consequently passengers till time step t are left aligning. At time step $t = \sigma(n)$, both the passengers $\sigma(n)$ and $\sigma(n-1)$ pick their bags and leave and hence, it is easy to see that passenger configuration till time step t is left aligning. By a repeated application of Lemma 6.5.3, we obtain the desired conclusion for this case.

2. $\underline{\sigma(n) > \sigma(n-1)+1}$: In this case, it is easy to see that the bag $\sigma(n-1)$ gets picked at time $\sigma(n)$. By our assumption in this case, $\sigma(n-1)$ gets picked first. By our assumption in this case, $\sigma(n-1)$ gets picked first. Consequently, for $t < \sigma(n-1)+1$, none of the passengers have picked their bags and hence, the configuration till t is left aligning. At time step $t = \sigma(n-1)+1$, we note that only the passenger $\sigma(n-1)$ has left and hence the passenger configuration till $\sigma(n-1)+1$ is left aligning. It is easy to see that for $\sigma(n-1)+1 < t < \sigma(n)$, the passenger position $t-1$ is empty and hence, passengers till position t are left aligning. At time step $t = \sigma(n)$, the passenger $\sigma(n)$ picks her bag and leaves. That is, the empty positions at time step $t = \sigma(n)$ are $t-1$ and t and hence, required passengers are left aligning. Using Lemma 6.5.3, we can conclude that this case is proved.

3. $\underline{\sigma(n) < \sigma(n-1)+1}$: We begin by noting that this case forces $\sigma(n) < \sigma(n-1)$ because $\sigma(n) = \sigma(n-1)$ is not possible. This means that the first bag to get picked is $\sigma(n)$. That is, for any $t < \sigma(n)$, no passenger has left and hence, passengers till t are left aligning. At time step $t = \sigma(n)$, the bag $\sigma(n)$ gets picked and hence, the passengers till position t are left aligning. For time steps t satisfying $\sigma(n) < t < \sigma(n-1)$ the only empty passenger position is t and passenger corresponding to the position of the bag $\sigma(n-1)$ is the passenger labelled t. Consequently, at time step t, the passenger $\sigma(n-1)$ picks her bag and leaves and hence, passengers till position t are left aligning. Using Lemma 6.5.3, this case is done.

We now proceed to the induction step. That is, we assume that for $t \geq k-1$, the passenger till $t-k+2$ are left aligning. We can conclude that for any time step t satisfying $k-1 \leq t < T(n-k,\sigma)$, the passenger configuration till $t-k+2$ is left aligning. This is because the bag $\sigma(n-k)$ plays no role in this range. Let us now consider $t = T(n-k,\sigma)$. We have to show that at this time step, the passenger configuration till $t-k+2$ is left aligning. Since $t = T(n-k,\sigma)$, we split our investigation into the following cases:

1. $\underline{\alpha^{\sigma}_{t-1,k+1}(t-k) = \sigma(n-k)}$: We consider the following sub-cases:

   (a) $\underline{\text{For every } i \leq t-k, \text{ we have } \alpha^{\sigma}_{t-1,k+1}(i) \neq 0}$: Since the bag $\sigma(n-k)$ is at position $t-k-1$ at time $t+1$, we can conclude that for every $i \leq t-k-1$, we have $\alpha^{\sigma}_{t,k+1}(i) \neq 0$.

This clearly means that at time step t, in this case, the passengers till $t-k+1$ are left aligning.

(b) There is an $i \leq t-k$ such that $\alpha^{\sigma}_{t-1,k+1}(i) = 0$: Since $\alpha^{\sigma}_{t-1,k+1}(t-k) = \sigma(n-k)$, in this case, by the induction hypothesis, we must have that

$$\alpha^{\sigma}_{t-1,k+1}(t-k-1) = 0 = \alpha^{\sigma}_{t-1,k+1}(t-k+1).$$

Since $t = T(n-k,\sigma)$, we must have that

$$\alpha^{\sigma}_{t,k+1}(t-k-1) = 0 = \alpha^{\sigma}_{t-1,k+1}(t-k).$$

Hence, the passengers till $t-k+1$ are left aligning.

2. $\alpha^{\sigma}_{t-1,k+1}(t-k+1) = \sigma(n-k)$ and $\alpha^{\sigma}_{t-1,k+1}(t-k) = 0$: From the assumption of this case, we can clearly conclude that

$$\alpha^{\sigma}_{t,k+1}(t-k) = 0 = \alpha^{\sigma}_{t,k+1}(t-k+1).$$

Clubbing this with the fact that passengers till $t-k-1$ are left aligning, we obtain that passengers till $t-k+1$ are left aligning.

We now have to consider $t > T(n-k,\sigma)$. We note that if $T(n-k+1,\sigma) \leq T(n-k,\sigma)$, by repeatedly using Lemma 6.5.3, we obtain the desired conclusion. So, we suppose that $T(n-k+1,\sigma) > T(n-k,\sigma)$. Repeatedly using Lemma 6.5.3, we obtain that for any t such that $T(n-k,\sigma) < t < T(n-k+1,\sigma)$, passengers till $t-k+1$ position are left aligning. This is because the bag $\sigma(n-k+1)$ plays no role in this time range. Now, at time step $t = T(n-k+1,\sigma)$, we note that if the bag $\sigma(n-k+1)$ was not there, then, by Lemma 6.5.3, the passenger position till $t-k+1$ would have been left aligning. Also, since $t = T(n-k+1,\sigma)$ and the bag $\sigma(n-k+1)$ would be at position $t-k+1$ at time t and the passenger $\sigma(n-k+1)$ would either be at $t-k$ with $t-k+1$ empty or at $t-k+1$. Clearly, removing this passenger, yields a left aligning configuration. By repeatedly using Lemma 6.5.3, we are done. $\square$

Lemma 6.5.5. *Let $\sigma \in S_n$, $k \in [n]$ and $i \in [k]$. Suppose that we initiate the belt with first k bags of σ. Then, at time $n+i-1$, the passenger n has either already left or is at position $n-i$. In other words, once passenger n starts moving (at time step n), she always keeps moving until no longer possible.*

Proof. We proceed by induction on k. We note that the base case easily follows from Lemma 6.3.1. We now assume that our claim is true for some k. Suppose that we initiate the belt with first $k+1$ bags of σ. We note that by Lemma 6.5.4, for $t \geq k$ and for any $i \geq t-k+1$, we must have that

$$\alpha^{\sigma}_{t,k}(i) = \alpha^{\sigma}_{t,k+1}(i) \tag{6.5.1}$$

We put $t^* = \max\{t : \exists i \geq t-k+1 \text{ such that } \alpha^{\sigma}_{t,k}(i) = n\}$. In words, t^* is the last time step for which n is to the right of the bag $\sigma(n-k+1)$. By the induction hypothesis and 6.5.1, we note that the passenger n satisfies our required condition for $t \leq t^*$. We have the following two cases:

1. $\underline{\alpha^{\sigma}_{t^*,k}(t^*-k+1) = n = \alpha^{\sigma}_{t^*,k+1}(t^*-k+1)}$: If $\sigma(n-k) = n$, our result is immediately achieved. On the other hand, we note that by Lemma 6.5.4, passengers till t^*-k+1 are left aligning and hence for $t > t^*$, we note that n keeps moving until possible, as required.

2. $\underline{\alpha^{\sigma}_{t^*,k}(t^*-k+2) = n}$: In this case, by the maximality of t^*, we must have that

$$\alpha^{\sigma}_{t^*,k}(t^*-k+1) = 0 = \alpha^{\sigma}_{t^*,k+1}(t^*-k+1).$$

   If $\sigma(n-k) = n$ then, our result is immediately achieved. Otherwise, $\alpha^{\sigma}_{t^*,k+1}(t^*-k+1) = 0$. Hence, using Lemma 6.5.4, we are done.

□

Definition 6.5.6. *Let $\sigma \in S_n$ be a permutation of bags. By inserting a lag at the i^{th} place, we mean the, obtained by inserting a 0 at the i^{th} position of σ.*

Lemma 6.5.7. *Let $\sigma \in S_n$, $j \in [n]$ and a time step t be given. Let $r(j,t)$ be the minimum number such that the passenger j has not picked up her bag by time t when the belt is initiated with $\tilde{\sigma}_{r(j,t)}$. Further suppose that, $\alpha^{\tilde{\sigma}_{r(j,t)}}_t(p(j,t)) = j$. Then at least one of the following holds:*

1. *For all $i \geq r(j,t)$, we have $\alpha^{\tilde{\sigma}_{r(j,t)}}_t(p(j,t)) = j$.*

2. *There is a $k(j,t) > r(j,t)$ such that*

   (a) *For every i satisfying $r(j,t) \leq i \leq k(j,t)$,*

   $$\alpha^{\tilde{\sigma}_{r(j,t)}}_t(p(j,t)) = j$$

   (b) *For every i satisfying $k(j,t) < i \leq n$,*

   $$\alpha^{\tilde{\sigma}_i}_t(p(j,t)) = 0 \text{ and } \alpha^{\tilde{\sigma}_i}_t(p(j,t)+1) = j$$

Also, if $r(j,t) > 0$ then, $T(\tilde{\sigma}_i^{-1}(j), \tilde{\sigma}_i) = t+1$ for every $i \geq r(j,t)$.

Proof. We will prove this lemma by strong induction on the label of the passengers. The result is clearly true for the passenger labelled 1. We assume that it is true for passengers having label $i \leq l$. We now consider the passenger labelled $l+1$. We want to show that for every time step t, the passenger $l+1$ satisfies the conditions laid out in the present lemma. By Lemma 6.3.1, the result is clearly true for some time step t. We want to prove that it is true for the time step $t+1$. Let us first focus on the passenger $l+1$ at time step t. Let $r(l+1,t)$ and $p(l+1,t)$ be the numbers obtained via the induction hypothesis. Note that if $r(l+1,t) > 0$ the, by the induction hypothesis, $T(l+1, \tilde{\sigma}_i) = t+1$ for every $i \geq r(j,t)$. Consequently, at time step 1, the passenger $l+1$ is not present with respect to any $\tilde{\sigma}$ and hence the conclusion of the lemma is trivially satisfied for this case. Therefore, we assume that $r(l+1,t) = 0$. We have the following cases:

1. $\underline{k(l+1,t) \text{ does not exist:}}$ This means that for every $i \geq 0$, we have $\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)) = l+1$. If $\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-1) = 0$ for every $i \geq 0$ then, it is easy to see that the required configuration is obtained. This is because of the observation that in this case $\alpha_{t+1}^{\tilde{\sigma}_i}(p(l+1,t)-1) = l+1$. We now put
$$i_* = \min\{i \geq 0 : \alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-1) \neq 0\}.$$
We note that if $i_* = 0$ then our required configuration is trivially obtained. We put $\alpha_t^{\tilde{\sigma}_{i_*}}(p(l+1,t)-1) = m$. Since $m < l+1$, we note, by the induction hypothesis, that for every $i \geq i_*$,
$$\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-1) = m$$
It follows from the definition of i_* that for every $i < i_*$, we have $\alpha_{t+1}^{\tilde{\sigma}_i}(p(l+1,t)-1) = l+1$. This is because in this range,
$$(\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-1), \alpha_t^{\tilde{\sigma}_i}(p(l+1,t))) = (0, l+1).$$
It now suffices to show that for $i \geq i_*$, we have
$$(\alpha_{t+1}^{\tilde{\sigma}_i}(p(l+1,t)-1), \alpha_{t+1}^{\tilde{\sigma}_i}(p(l+1,t)-1)) = (0, l+1).$$
For this purpose, we consider the number $r(m,t)$ provided to us by the induction hypothesis and investigate the following cases:

   (a) $\underline{r(m,t) = 0:}$ Since, we are under the assumption that $i_* > 0$, we get, by the induction hypothesis, that $p(m,t) = (p(l+1,t)-1)-1 = p(l+1,t)-2$ and that for every $i \geq i_*$, we have $\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-2)=0$. From this, our desired result follows because of the following:

$$(\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-2),\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-1),\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)))=(0,m,l+1)$$

(b) $\underline{r(m,t)>0}$: By the induction hypothesis, we note that for every $i\geq r(m,t)$, we have $T(\tilde{\sigma}_i^{-1}(m),\tilde{\sigma}_i)=t+1$ that is, our desired configuration follows.

2. $\underline{k(l,t)\text{ exists}}$: By the definition of $k(l,t)$, we have that for every $i>k(l+1,t)$,

$$(\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-1),\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)))=(0,l+1).$$

From this, we conclude that for $i>k(l+1,t)$, we have $\alpha_{t+1}^{\tilde{\sigma}_i}(p(l+1,t)-1)=l+1$. We define i_* as in the previous case. By analogous arguments as in the previous case, it suffices to consider the case $i_*>0$. We note that for $i\leq\min\{i_*-1,k(l+1,t)\}$, we have that

$$(\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-1),\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)))=(0,l+1).$$

From this, we conclude that $\alpha_{t+1}^{\tilde{\sigma}_i}(p(l+1,t)-1)=l+1$. It now suffices to show that for every $i>\min\{i_*-1,k(l+1,t)\}$. We show this by considering the following cases:

(a) $\underline{i\geq i_*}$: We further split this case into the following cases:

- $\underline{r(m,t)=0}$: In this case, we note that by the induction hypothesis, for every $i\geq i_*$, we have
$$(\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-2),\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-1))=(0,m)$$
From this, we obtain our required conclusion in this case.
- $\underline{r(m,t)>0}$: In this case, by the induction hypothesis, for every $i\geq r(m,t)$, we have $T(\tilde{\sigma}_i^{-1}(m),\tilde{\sigma}_i)=t+1$ that is, our desired conclusion follows.

(b) $\underline{i_*>i>k(l,t)}$: By the definitions of $k(l,t)$ and i_*, for i satisfying $i_*>i>k(l,t)$, we have that
$$(\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)-1),\alpha_t^{\tilde{\sigma}_i}(p(l+1,t)))=(0,0).$$
Consequently, $\alpha_{t+1}^{\tilde{\sigma}_i}(p(l+1,t)-1)=0$ as required.

We recall that we are in the case where $r(l+1,t)=0$. Let p be such that

$$\alpha_{t+1}^{\tilde{\sigma}_0}(p)=\alpha_{t+1}^{\tilde{\sigma}}(p)=l+1.$$

By the discussion so far, one of the following is true:

1. $\alpha_{t+1}^{\tilde{\sigma}_i}(p) = l+1$ for all $i \geq 0$.

2. There is a k such that for every $i \leq k$, we have $\alpha_{t+1}^{\tilde{\sigma}_i}(p) = l+1$ and for every $i > k$, we have $\alpha_{t+1}^{\tilde{\sigma}_i}(p) = 0$ and $\alpha_{t+1}^{\tilde{\sigma}_i}(p+1) = l+1$.

We put $\sigma(b) = l+1$. Suppose that we have updated the bags. If the bag labelled $l+1$ is at a position less than p then with respect to $\tilde{\sigma}_0$, then there is nothing to prove. Otherwise, we consider the following cases:

1. Suppose that with respect to $\tilde{\sigma}_0$, the bag labelled $l+1$ is at position p. This means that for every $i < b$, the passenger $l+1$ obtains her bag with respect to $\tilde{\sigma}_i$ and leaves. Note that for every $i \geq b$, the bags $l+1$ is at position $p-1$. This means that at time $t+2$, the bags labelled $l+1$ will be at position p with respect to $\tilde{\sigma}_i$ for every $i \geq b$. We also note that for every $i \geq b$, the passenger $l+1$ will be at either p or $p-1$. Consequently, at time $t+2$, the passenger $l+1$ will obtain her bag and leave, as required.

2. Suppose that with respect to $\tilde{\sigma}_0$, the bag labelled $l+1$ is at position $p+1$. This means that for every $i < b$, passenger $l+1$ obtains her bag with respect to $\tilde{\sigma}_i$ and leaves. Note that for every $i \geq b$, the bag $l+1$ is at position p. This means that for every $i \leq k$, passenger $l+1$ obtains her bag and leaves. Also, since the bag $l+1$ will be at position $p+1$ at time step $t+2$, we are done.

This completes the proof.

□

We obtain as a corollary of Lemma 6.5.7, the following:

Lemma 6.5.8. *Let $\sigma \in S_n$ be a permutation of bags. Then, for all $i \in [0,n]$, we have*

$$T(\sigma) = T(\tilde{\sigma}_0) \leq T(\tilde{\sigma}_i) \leq T(\tilde{\sigma}_n) = T(\sigma)+1.$$

Proof. Let us fix a passenger j, and an $i \in [0,n]$. Let us put $t = T(\tilde{\sigma}_i^{-1}(j), \tilde{\sigma}_j)$. Suppose that if possible $T(\tilde{\sigma}_0^{-1}, \tilde{\sigma}_0) > t$. This means that there is a p such that $\alpha_t^{\tilde{\sigma}_0}(p) = j$. By Lemma 6.5.7, we obtain that either $\alpha_t^{\tilde{\sigma}_i}(p) = j$ or that $\alpha_t^{\tilde{\sigma}_i}(p+1) = j$ which is not possible. Thus, we can conclude that

$$T(\tilde{\sigma}_0^{-1}(j), \tilde{\sigma}_j) \leq T(\tilde{\sigma}_i^{-1}(j), \tilde{\sigma}_j) \tag{6.5.2}$$

Now, suppose that $T(\tilde{\sigma}_n^{-1}(j), \tilde{\sigma}_j) < t$. By the definition of t, we have that there is a p such that $\alpha_{t-1}^{\tilde{\sigma}_i}(p) = j$. By Lemma 6.5.7, we obtain that either $\alpha_{t-1}^{\tilde{\sigma}_n}(p) = j$ or that $\alpha_{t-1}^{\tilde{\sigma}_n}(p+1) = j$ which is not possible. Thus, we can conclude that

$$T(\tilde{\sigma}_i^{-1}(j), \tilde{\sigma}_j) \leq T(\tilde{\sigma}_n^{-1}(j), \tilde{\sigma}_j) \tag{6.5.3}$$

Combing the above inequalities and taking the max over all passengers, we obtain that

$$T(\tilde{\sigma}_0) \leq T(\tilde{\sigma}_i) \leq T(\tilde{\sigma}_n).$$

Finally, our result follows from the observations that $T(\tilde{\sigma}_0) = T(\sigma)$ and that $T(\tilde{\sigma}_n) = T(\sigma) + 1$. □

Lemma 6.5.9. *Let $\sigma \in S_n$. Let σ' be defined as follows:*

$$\sigma'(i) = \begin{cases} \sigma(i) & \text{if } \sigma(i) \neq n \\ \\ 0 & \text{if } \sigma(i) = n \end{cases}$$

Then, $T(\sigma) = \max\{T(\sigma'), T(n, \sigma)\}$.

Proof. This a direct consequence of Lemma 6.5.5. □

Lemma 6.5.10. *Let $\sigma \in S_n$. Then, $T(n, \sigma) \leq T_n$. Moreover there exists at least one permutation for which $T(n, \sigma) = T_n$.*

Proof. This follows from Lemma 6.5.5. A boundary permutation is $\sigma : i \mapsto n - i + 1$. □

Proof of Theorem 6.2.5. We prove by induction on n. It is easy to check the result for the base cases base case of $n = 2$. We now assume that the result is true for some n. Let $\sigma \in S_{n+1}$. By Lemma 6.5.9, we obtain that $T(\sigma) = \max\{T(\sigma'), T(n+1, \sigma)\}$ (recall that σ' is obtained from σ by removing $(n+1)$ and leaving a gap there, see Lemma 6.5.9). It is easy to see that there is a $\gamma \in S_n$ and an $i \in [0, n]$ such that $\tilde{\gamma}_i = \sigma'$. Therefore, by Lemma 6.5.7, we obtain that

$$T(\sigma') = T(\tilde{\gamma}_i) \leq T(\tilde{\gamma}_0) + 1.$$

By the induction hypothesis, we obtain that

$$T(\tilde{\gamma}_0) + 1 \leq T_n + 1 \leq T_{n+1}.$$

Therefore, we have $T(\sigma') \leq T_{n+1}$. By Lemma 6.5.10, we have $T(n+1, \sigma) \leq T_{n+1} \Rightarrow T(n+1, \sigma) \leq T_{n+1}$. as required. Also, the permutation $\sigma : i \mapsto n-i+2$ satisfies that $T(\sigma) = T_{n+1}$. Hence proved. $\square$

Chapter 7

Union-Closed Families of Sets

Preface

This chapter is based on the following paper:

- Bhasin D., "Closures of union-closed families of sets" – submitted. Preprint available **here**.

7.1 Introduction

The union-closed sets conjecture is an easy-to-state, notoriously difficult problem that was first proposed by Péter Frankl(see [38]). Several research articles (see for e. g. [10, 26, 64, 85, 91]) and a Polymath Project (see [1]) have been devoted to the unraveling of this long-standing open problem.

Letting $[n]$ denote the finite set $\{1,2,\ldots,n\}$ which will serve as our universe and $2^{[n]}$ denote the power set of $[n]$, a family of subsets $\mathscr{F} \subseteq 2^{[n]}$ is said to be *union-closed over universe* $[n]$ if $[n] \in \mathscr{F}$, and for all $A, B \in \mathscr{F}$, we have $A \cup B \in \mathscr{F}$. We call each subset A of $[n]$ that is contained in the family $\mathscr{F}$ a *member-subset* of $\mathscr{F}$, and we let $|\mathscr{F}|$ denote the total number of member-subsets of $\mathscr{F}$. Then the union-closed sets conjecture can be stated as follows:

Conjecture 1. *Let $\mathscr{F}$ be a union-closed family of sets over the universe $[n]$. Then there exists some element $i \in [n]$ such that i belongs to at least half of the member-subsets of $\mathscr{F}$, i.e. $\sum_{A\in\mathscr{F}} \mathbf{1}_{i\in A} \geqslant \lfloor |\mathscr{F}|/2 \rfloor$.*

A detailed survey by Bruhn and Schaudt in [27] explains how the conjecture has travelled, both geographically and mathematically, as they put it. The problem has equivalent formulations in lattice theory (see [85]) and graph theory (see [26]), and the claim in the conjecture has been established for various lattice classes and graph classes. Some results have been obtained proving that when the size $|\mathscr{F}|$ of the family is sufficiently large (as a function of the size of the universe, n), then $\mathscr{F}$ satisfies the union-closed sets conjecture (for instance, see [64], where it has been established, using Boolean analysis, that there exists a constant c such that all union-closed families $\mathscr{F}$ with $|\mathscr{F}| \geq (\frac{1}{2} - c)2^{n-1}$ satisfy the conjecture).

In this section, we investigate several intriguing properties of union-closed families. In [24], the number of non-isomorphic union-closed families for $n = 7$ was computed, and an algorithm was devised which involved recursively adding a new set A to a union-closed family $\mathscr{F}$ such that the new family $\mathscr{F} \cup \{A\}$ remains union-closed. We collect all such sets A and construct a new family

$$\overline{\mathscr{F}} = \{A \in 2^{[n]} : \mathscr{F} \cup \{A\} \text{ is union-closed}\}. \tag{7.1.1}$$

We call this the *closure* of $\mathscr{F}$. The family $\overline{\mathscr{F}}$ is itself union-closed. We mention here (see also Lemma 7.2.2) that as long as a family $\mathscr{F}$ is a proper subset of $2^{[n]}$, there exists an $A \notin \mathscr{F}$ such that $\mathscr{F} \cup \{A\}$ remains union-closed. Consequently, for every union-closed $\mathscr{F}$ that is a proper subset of $2^{[n]}$, it is also a proper subset of its closure $\overline{\mathscr{F}}$ defined in (7.1.1).

Note, further, that since the universe $[n]$ itself is finite, one requires to repeat the operation of taking closures only a finite number of times before a given union-closed family $\mathscr{F}$ reaches the power set $2^{[n]}$. In other words, if we set $\overline{\mathscr{F}}^{(0)} = \mathscr{F}$ and $\overline{\mathscr{F}}^{(i)} = \overline{\overline{\mathscr{F}}^{(i-1)}}$ for each $i \in \mathbb{N}$, then there is a finite $k \in \mathbb{N}$ such that $\overline{\mathscr{F}}^{(k)} = 2^{[n]}$. This inspires us to introduce the following definition:

Definition 7.1.1. *A union-closed family $\mathscr{F}$ is said to be k-*dense*, for $k \in \mathbb{N}_0$, if k is the smallest non-negative integer such that $\overline{\mathscr{F}}^{(k)} = 2^{[n]}$. We call k the* density *of $\mathscr{F}$.*

Closures provide a natural parameter for a potential induction-based proof of Conjecture 1, namely the density of a family. The base case, $0-$dense families, is trivially true as $2^{[n]}$ is the only $0-$dense family. Thus, Conjecture 1 is equivalent to the following conjecture:

Conjecture 2. *Suppose $\mathscr{F}$ is a union-closed family such that $\overline{\mathscr{F}}$ satisfies Conjecture 1. Then $\mathscr{F}$ also satisfies Conjecture 1.*

The benefit of this formulation is that it gives us some extra information about the union-closed family for which we need to check the validity of Conjecture 1. If the structure of closures and its properties can be understood fairly well then it can hold a potential to give a better understanding of Conjecture 1.

We now describe the organization of this section. We mention at the very outset that throughout this work, it has been assumed that the empty set $\emptyset$ is not contained in *any* family $\mathscr{F}$ of subsets of $[n]$, including the cases where $\mathscr{F} = 2^{[n]}$. In §7.2, we state and prove several properties of closures and densities of union-closed families of $[n]$, many of which are further utilized in the proofs of results in §7.3, §7.4, §7.5. We conclude §7.2 by Theorem 7.2.8 which gives a criterion to obtain a lower bound on the density of a given union-closed family $\mathscr{F}$. We provide examples to show that for some families this bound can be achieved while for others it need not always be achieved.

In §7.3 we show that for each k there are at least $\binom{n}{k-1} f_{k-1}$ non-isomorphic union-closed families that are $k-$dense, where n is the size of the universe, $k \leqslant n-1$ and f_k is the number of labelled union-closed families over universe k. We achieve this by explicit construction. In §7.4, for each $n \geqslant 6$, we give example of an $(n-1)-$dense family $\mathscr{F}$ which is different from those discussed in §7.3. We give a complete description of its closures $\overline{\mathscr{F}}^{(k)}$ for $k \leq n-5$ and consequently show that it is $(n-1)-$dense. The method of proof used in §7.3 and §7.4 can potentially be used to find densities and description of closures of various union-closed families.

The goal of §7.5 is to give a criterion to check whether a given $1-$dense family $\mathscr{F}$ has a *closure root* or not, i.e, $\mathscr{H} \subseteq \mathscr{F}$ such that $\overline{\mathscr{H}} = \mathscr{F}$. We achieve this using the notion of *relative subsets* we introduce in §7.5. Intuitively, relative subsets can be thought of as subsets within a given family. Using this notion we also generalise Lemma 7.2.7 to Proposition 7.5.4. Having build up some basic results regarding relative subsets, in Theorem 7.5.13 , for a $1-$dense family $\mathscr{F}$, we construct a particular family such that $\mathscr{F}$ has a closure root $\Leftrightarrow$ that particular family is a closure root of $\mathscr{F}$. We conclude the section by Corollary 7.5.14 in which we give a non-trivial example of a $1-$dense family having no closure roots. The notion of relative subsets can serve as an effective tool in the study of closures of union-closed families.

7.2 Various properties of closures and densities of union-closed families

We begin by stating several properties of closures and densities of union-closed families, followed by a succinct characterization of 1-dense union-closed families using *up-sets*, and finally end with Theorem 7.2.8.

Lemma 7.2.1. *Let $\mathscr{F}$ be a union-closed family and $\overline{\mathscr{F}}$ be its closure, as defined in* (7.1.1)*. Then $\overline{\mathscr{F}}$ is union-closed.*

Proof. Fix any $A,B \in \overline{\mathscr{F}}$. We wish to show that $A \cup B \in \overline{\mathscr{F}}$. Fix any $C \in \mathscr{F}$. As $B \in \overline{\mathscr{F}}$, hence $B \cup C \in \mathscr{F} \cup \{B\}$, which implies that either $B \cup C = B$ or $B \cup C \in \mathscr{F}$. In the former case, $A \cup B \cup C = A \cup B \in \mathscr{F} \cup \{A \cup B\}$. In the latter case, $B \cup C \in \mathscr{F}$, which in turn implies, since $A \in \overline{\mathscr{F}}$, that $A \cup (B \cup C) \in \mathscr{F} \cup \{A\}$. If $A \cup B \cup C \in \mathscr{F}$ then we get our desired conclusion. Otherwise, we have $A \cup B \cup C = A$, which implies $B \subseteq A$ and hence $A \cup B = A$, which we already know is in $\overline{\mathscr{F}}$. $\square$

Lemma 7.2.2. *Let $\mathscr{F}$ be a union-closed family over a universe $[n]$. Suppose that $\mathscr{F} \neq 2^{[n]}$ then $\mathscr{F} \subsetneq \overline{\mathscr{F}}$.*

Proof. Choose A to be a *maximal* subset in the family $2^{[n]} \setminus \mathscr{F}$, i.e. if $B \in 2^{[n]} \setminus \mathscr{F}$ and $A \subseteq B$, then $A = B$. We claim that $\mathscr{F} \cup \{A\}$ is union-closed. To see this, choose any $C \in \mathscr{F}$. If $C \subset A$, then $C \cup A = A$, and if $C \not\subset A$, then A is a proper subset of $A \cup C$, which in turn, via the maximality of A, implies that $A \cup C \in \mathscr{F}$. Either way, we end up with $A \cup C \in \mathscr{F} \cup \{A\}$, as desired. $\square$

At this point, we define $\mathscr{A}_{k,n} = \{A \subseteq [n] : |A| = k\}$ to be the collection of all subsets of $[n]$ which have cardinality k.

Lemma 7.2.3. *Let $\mathscr{F}$ be a union-closed family over the universe $[n]$, with density k. Then $0 \leqslant k \leqslant n-1$.*

Proof. Recall the definition of $\overline{\mathscr{F}}^{(i)}$ for any $i \in \mathbb{N}_0$. From the definition of universe, we know that $\{[n]\} = \mathscr{A}_{n,n} \subseteq \overline{\mathscr{F}}^{(i)}$ for each i. Let t_i denote the smallest positive integer such that $\mathscr{A}_{t_i,n} \subseteq$

$\overline{\mathscr{F}}^{(i)}$. Consequently if a subset A belongs to $\overline{\mathscr{F}}^{(i)} \setminus \mathscr{A}_{t_i-1,n}$, then A has to be a maximal element of $2^{[n]} \setminus \overline{\mathscr{F}}^{(i)}$ and hence, as illustrated in the proof of Lemma 7.2.2, $A \in \overline{\mathscr{F}}^{(i+1)}$. Thus we conclude that $\mathscr{A}_{t_i-1,n} \subseteq \overline{\mathscr{F}}^{(i+1)}$. This in turn shows that $t_{i+1} \leqslant t_i - 1$ for each i with $t_i \geqslant 2$, and therefore, we must have $t_{n-1} = 1$. This implies that $\mathscr{A}_{t,n} \subset \overline{\mathscr{F}}^{(n-1)}$ for all $t \geqslant 1$, so that $\overline{\mathscr{F}}^{(n-1)} = 2^{[n]}$, thus concluding the proof. □

Therefore, given any family $\mathscr{F}$ over the universe $[n]$, we obtain $2^{[n]}$ in at most $n-1$ "steps" of closures. Now, we have:

Corollary 7.2.4. *Let $\mathscr{F}$ be a union-closed family over the universe $[n]$. Then $\mathscr{A}_{n-k} \subseteq \overline{\mathscr{F}}^{(k)}$.*

This follows from the inequality obtained in the proof of the previous lemma, namely $t_{i+1} \leqslant t_i - 1$ alongwith the observation that $t_1 \leqslant n-1$. We next show that you cannot do better than this $n-1$, i.e. there exists at least one family $\mathscr{F}$ with density precisely $n-1$.

Proposition 7.2.5. *Consider the family $\mathscr{F} = \{[1],[2],[3],\ldots,[n]\}$ over the universe $[n]$. Then $\mathscr{F}$ is $(n-1)$-dense.*

Proof. First of all, note that $\mathscr{F}$ is union-closed. Next, note that $\{1,2,3,\ldots,n-2,n\} \notin \mathscr{F}$ and $\{1,2,3,\ldots,n-2\} \in \mathscr{F}$. These together imply that

$$\{1,2,3,\ldots,n-3,n\} \notin \overline{\mathscr{F}}. \tag{7.2.1}$$

Next, we note that $\{1,2,3,\ldots,n-3\} \in \mathscr{F} \subset \overline{\mathscr{F}}$, and this, combined with (7.2.1), yields

$$\{1,2,3,\ldots,n-4,n\} \notin \overline{\mathscr{F}}^{(2)}. \tag{7.2.2}$$

Proceeding thus, by induction, we see that $\{n\} \notin \overline{\mathscr{F}}^{(n-2)}$, which implies that $\overline{\mathscr{F}}^{(n-2)} \neq 2^{[n]}$, and hence the density of $\mathscr{F}$ is at least $n-1$. This, along with Lemma 7.2.3, concludes the proof. □

Let $\mathscr{U}_n$ denote the set of all union-closed families over universe $[n]$. If we construct a graph G_n whose vertex set is $\mathscr{U}_n$, and draw a directed edge between two families $\mathscr{F}_1$ and $\mathscr{F}_2$, directed from $\mathscr{F}_1$ to $\mathscr{F}_2$, iff $\overline{\mathscr{F}_1} = \mathscr{F}_2$, then, by Lemma 7.2.1, every vertex has exactly one out-degree. By Lemma 7.2.2, there are no self-loops other than at the vertex $2^{[n]}$. By Lemma 7.2.3, the maximum length of a directed path in $\mathscr{G}_n$ is $n-1$. So, the set of all union-closed families over universe $[n]$ can be imagined as a tree with root $2^{[n]}$ and depth $n-1$.

Our next result focuses on a nice characterization of 1-dense families. To this end, we require the well known notion of up-sets(see [33]).

Definition 7.2.6. *A family $\mathscr{F}$ over a given finite or infinite universe is said to be an up-set if for every $A \in \mathscr{F}$ and $B \supseteq A$, we have $B \in \mathscr{F}$.*

Lemma 7.2.7. *Let $\mathscr{F} \neq 2^{[n]}$ be a union-closed family over the universe $[n]$. Then $\mathscr{F}$ is 1-dense if and only if $\mathscr{F}$ is an up-set.*

Proof. Suppose $\mathscr{F}$ is 1-dense. Let $A \in \mathscr{F}$ and let A be a proper subset of B. Since, $\mathscr{F}$ is 1-dense, we have $\overline{\mathscr{F}} = 2^{[n]}$, so that $B \setminus A \in \overline{\mathscr{F}}$. Therefore, $\mathscr{F} \cup \{B \setminus A\}$ is union-closed, and hence $B = (B \setminus A) \cup A \in \mathscr{F} \cup \{B \setminus A\}$. Since A is a non-empty, proper subset of B, hence $B \neq B \setminus A$, which yields $B \in \mathscr{F}$. Hence $\mathscr{F}$ is an up-set.

Conversely, suppose that $\mathscr{F}$ is an up-set. Let $B \in 2^{[n]} - \mathscr{F}$. Take any $A \in \mathscr{F}$. Note that $A \cup B \supseteq A$, so that $A \cup B \in \mathscr{F}$ as $\mathscr{F}$ is an up-set. Therefore, $B \in \overline{\mathscr{F}}$, thus proving that $\overline{\mathscr{F}} = 2^{[n]}$. □

It is well-known (see for e.g. Introduction of [10]) that up-sets satisfy Conjecture 1. By Lemma 7.2.7, we conclude that 1-dense families satisfy it as well. Using Lemma 7.2.7, we establish Theorem 7.2.8.

Theorem 7.2.8. *Let $\mathscr{F}$ be a k-dense union-closed family. If there exist subsets $A_1 \subsetneq A_2 \subsetneq \cdots \subsetneq A_r \subsetneq B_r$ of $[n]$ with $A_1, A_2, \ldots, A_r \in \mathscr{F}$ and $B_r \notin \mathscr{F}$, then $r < k$.*

Proof. We use induction on the density of $\mathscr{F}$. First, consider $\mathscr{F}$ that is 1-dense. By Lemma 7.2.7, it is an up-set. Hence $A \in \mathscr{F}$ and $B \notin \mathscr{F}$ imply that $A \nsubseteq B$. Therefore, $r = 0$. This concludes the proof of the claim for the base case.

Assume that the result holds for t-dense families for some $t \in \mathbb{N}$ with $t \leqslant n-2$. Let $\mathscr{F}$ be a $t+1$-dense family. Suppose there exist $A_1, \ldots, A_r$ and B_r as in the statement of Theorem 7.2.8. Now, $A_r \in \mathscr{F}$ and $A_{r-1} \cup (B_r \setminus A_r) \cup A_r = B_r \notin \mathscr{F}$ together imply that $A_{r-1} \cup (B_r \setminus A_r) \notin \overline{\mathscr{F}}$. Therefore we get $A_1 \subsetneq A_2 \subsetneq \cdots \subsetneq A_{r-1} \subsetneq A_{r-1} \cup (B_r \setminus A_r)$ with $A_1, A_2, \ldots, A_{r-1} \in \overline{\mathscr{F}}$ and $A_r \cup (B_r - A_r) \notin \overline{\mathscr{F}}$. Moreover, since $\mathscr{F}$ is $t+1$-dense, we know that $\overline{\mathscr{F}}$ is t-dense. By our induction hypothesis, we have $r-1 < t$ which yields $r < t+1$, as desired. □

Remark 7.2.9. *Letting $\mathscr{F} = \{[1],[2],[3],\ldots,[n]\}$, $A_i = \{[i]\}$ for $1 \leqslant i \leqslant n-2$ and $B_{n-2} = \{1,2,\ldots,n-2,n\}$, Theorem 7.2.8 yields the density of $\mathscr{F}$ to be at least $n-1$, which corroborates the conclusion of Proposition 7.2.5. This also leads to the following, far more general observation: if $\mathscr{F}$ is a union-closed family such that $[i] \in \mathscr{F}$ for all $i \in [n]$ and $\{1,2,\ldots,n-2,n\} \notin \mathscr{F}$, then $\mathscr{F}$ is $(n-1)$-dense.*

Theorem 7.2.8 shows that there cannot exist a "very long" chain of subsets of the type described in the statement of the theorem, especially when the density of $\mathscr{F}$ is "small" compared to n, the size of the universe. Let

$$s(\mathscr{F}) = \max\{r : \exists A_1 \subsetneq A_2 \subsetneq \cdots \subsetneq A_r \subsetneq B_r, A_i \in \mathscr{F}, B_r \notin \mathscr{F}\}.$$

For a $k-$*dense* family $\mathscr{F}$, Theorem 7.2.8 gives us the lower bound $s(\mathscr{F})+1 \leqslant k$. We note here that this bound is not always tight. As an instance, consider the union-closed family

$$\mathscr{F} = \{[n-2], \{1,2,4,5,6,\ldots,n\}, \{1,3,4,5,\ldots,n\}, \{2,3,\ldots,n\}, [n]\}.$$

It is straightforward to check that $2^{[n-2]} \subseteq \overline{\mathscr{F}}$ and $\{1,2,3,\ldots,n-3,n\} \notin \overline{\mathscr{F}}$. Setting $A_i = [i]$ for $1 \leqslant i \leqslant n-4$ and $B_{n-3} = \{1,2,3,\ldots,n-3,n\}$, we conclude, by Theorem 7.2.8, that $\overline{\mathscr{F}}$ is at least $(n-2)$-dense. This, along with Lemma 7.2.3, yields that $\mathscr{F}$ is $(n-1)$-dense. However, the only choice of subsets A and B such that $A \in \mathscr{F}$, $B \notin \mathscr{F}$ and $A \subsetneq B$ is $A = [n-2]$ and $B = [n-1]$, showing that $s(\mathscr{F}) = 1$, which is a lot smaller than $n-1$.

7.3 Many $k-$dense families

Let k be a fixed positive integer and $\mathscr{F}$ be a union -closed family over universe $[k]$. In this section, we will consider the family $\mathscr{H} = \mathscr{F} \cup \{[n]\}$ for $n \geqslant k+2$. Throughout this section, whenever $A \in 2^{[n]}$, by A_1, we will mean $A \cap [k]$ and by A_2, we will mean $A \cap ([n] \setminus [k])$. Our goal is to prove that $\mathscr{H}$ is $(k+1)-$dense. We begin with:

Lemma 7.3.1. *For any $A \in 2^{[n]}$ and for any $1 \leqslant t \leqslant k$, at least one of the following is true:*

(i) $A_2 = \emptyset$, $A_1 \in \overline{\mathscr{F}}^{(t)}$.
(ii) $A_2 = \emptyset$, $A_1 \notin \overline{\mathscr{F}}^{(t)}$.
(iii) $A_2 = [n] \setminus [k]$ *and* $\forall E \in \overline{\mathscr{F}}^{(t-1)}$ *with* $E \nsubseteq A$, $E \cup A \in \overline{\mathscr{H}}^{(t-1)}$.
(iv) $A_2 = [n] \setminus [k]$ *and* $\exists E \in \overline{\mathscr{F}}^{(t-1)}$ *with* $E \nsubseteq A$, $E \cup A \notin \overline{\mathscr{H}}^{(t-1)}$.
(v) $|A_1| \geqslant k-t+1$.
(vi) $A_2 \neq \emptyset, A_2 \neq [n] \setminus [k]$ *and* $|A_1| \leqslant k-t$.

where we put $\overline{\mathscr{F}}^{0} = \mathscr{F}$ *and* $\overline{\mathscr{H}}^{(0)} = \mathscr{H}$.

Proof. When $A_2 = \emptyset$ then A satisfies one of 1 or 2. When $A_2 = [n] \setminus [k]$ then A satisfies one of 3 or 4. When $A_2 \neq \emptyset, A_2 \neq [n] \setminus [k]$ then satisfies A satisfies 5 or 6. □

Using the previous Lemma, we prove the following:

Theorem 7.3.2. *Given $A \in 2^{[n]}$ and a positive integer $t \leqslant k$, if A satisfies any of the conditions stated in (i), (iii), (v) of Lemma 7.3.1 then $A \in \overline{\mathscr{H}}^{(t)}$. On the other hand,if A satisfies any of the conditions stated in (ii), (iv), (vi) of Lemma 7.3.1 then $A \notin \overline{\mathscr{H}}^{(t)}$.*

Proof. The proof is via induction on t. Let $\mathscr{C}_t = \{A \in \overline{\mathscr{H}}^{(t)} : A_2 = [n] \setminus [k]\}$. And put $M = \{\text{(i)}, \text{(iii)}, \text{(v)}\}$ and $N = \{\text{(ii)}, \text{(iv)}, \text{(vi)}\}$. For $q \in M$ let $P(l,q)$ be the statement that: if A satisfies q of Lemma 7.3.1 for $t = l$ then $A \in \overline{\mathscr{H}}^{(l)}$. And for $q \in N$ let $P(l,q)$ be the statement that: if A satisfies q of Lemma 7.3.1 for $t = l$ then $A \notin \overline{\mathscr{H}}^{(l)}$. Finally, let $P(l, \text{(vii)})$ be the statement that: $\mathscr{C}_l$ is an up-set. To prove this theorem, we will prove the following statement:

$$P(t) = \bigwedge_{q \in M \cup N \cup \{\text{(vii)}\}} P(t,q)$$

for $1 \leq t \leq k$. Let us first establish the base case, i.e., $P(1)$:

1. We need to show that $\mathscr{H} \cup \{A\}$ is union-closed. But this is true since $A \in \overline{\mathscr{F}}^{(1)}$ and $\mathscr{H} = \mathscr{F} \cup \{[n]\}$.

2. Since $A \notin \overline{\mathscr{F}}$, there is a $B \in \mathscr{F}$ such that $A \cup B \notin \mathscr{F} \cup \{A\}$. Since, $A \cup B \subseteq [k]$, we obtain that $A \cup B \notin \mathscr{H} \cup \{A\} \Rightarrow A \notin \overline{\mathscr{H}}$ as required.

3. We need to show that $\mathscr{H} \cup \{A\}$ is union-closed. Since $B \in \mathscr{H}$ means that either $B \in \mathscr{F}$ or $B = [n]$, we have that $A \cup B \in \mathscr{H} \cup \{A\}$ in both the cases.

4. When A satisfies (iv) of Lemma 7.3.1 then there is an $E \in \mathscr{F}$ with $E \nsubseteq A$ such that $E \cup A \notin \mathscr{H}$. But since $E \in \mathscr{F} \Rightarrow E \in \mathscr{H}$, we get that $A \notin \overline{\mathscr{H}}$.

5. Since $t = 1$, the only choice for A_1 is $[k]$. It is easy to see that $A \in \overline{\mathscr{H}}$.

6. Since $[k] \in \mathscr{H}$ and $A_2 \neq \emptyset$ and $A_2 \neq [n] \setminus [k]$, we have that $[k] \cup A \notin \mathscr{H}$ and hence, $A \notin \overline{\mathscr{H}}$.

7. Note that by $P(1, \text{(iii)})$ and $P(1, \text{(iv)})$, we have that when $A_2 = [n] \setminus [k]$ then $A \in \overline{\mathscr{H}} \Leftrightarrow \forall E \in \mathscr{F}$ with $E \nsubseteq A$, $E \cup A \in \mathscr{H}$. So suppose we have $A \in \mathscr{C}_1$ with $A \subseteq B$. Given $E \in \mathscr{F}$ with $E \nsubseteq B$, we also obtain that $E \nsubseteq A$. Therefore, $E \cup A \in \mathscr{H}$. Since, $A_2 = [n] \setminus [k]$, we have that $E \cup A = [n] \Rightarrow E \cup B = [n] \in \mathscr{H} \Rightarrow B \in \mathscr{C}_1$. And hence, $\mathscr{C}_1$ is an up-set.

Now, we suppose that $P(r)$ is true for some $1 \leqslant r < k$. Let us prove $P(r+1)$:

1. We are given that $A \in \overline{\mathscr{F}}^{(r+1)}$. We have to investigate $A \cup B$ for $B \in \overline{\mathscr{H}}^{(r)}$. By Lemma 7.3.1 and the induction hypothesis, the only cases are: when B satisfies (i), (iii), (v) of Lemma 7.3.1. We now consider all these cases:

(a) Suppose B satisfies (i) of Lemma 7.3.1. Thus, by the definition of closure, $A \cup B \in \overline{\mathscr{F}}^{(r)} \cup \{A\}$. If $A \cup B = A$ then there is nothing further to investigate. On the other hand, if $A \cup B \in \overline{\mathscr{F}}^{(r)}$ then by $P(r, \text{(i)})$, we have that $A \cup B \in \overline{\mathscr{H}}^{(r)}$, as required.

(b) Suppose B satisfies (iii) of Lemma 7.3.1. Thus by $P(r, \text{(iii)})$ we have that $B \in \mathscr{C}_r$. Therefore by $P(r, \text{(vii)})$, $\mathscr{C}_r$ is an up-set and hence $A \cup B \in \mathscr{C}_r \subseteq \overline{\mathscr{H}}^{(r)}$ as required.

(c) Suppose B satisfies (v) of Lemma 7.3.1. Therefore, $|B_1| \geqslant k - r + 1$. And hence, $|(A \cup B)_1| \geqslant k - r + 1 \Rightarrow A \cup B \in \overline{\mathscr{H}}^{(r)}$.

2. Since $A \notin \overline{\mathscr{F}}^{(r+1)}$, there is a $B \in \overline{\mathscr{F}}^{(r)}$ such that $A \cup B \notin \overline{\mathscr{F}}^{(r)} \cup \{A\}$. By $P(r, \text{(i)})$, we obtain that $B \in \overline{\mathscr{H}}^{(r)}$ and by $P(r, \text{(i)})$, we have $A \cup B \notin \overline{\mathscr{H}}^{(r)}$. Therefore, $A \notin \overline{\mathscr{H}}^{(r+1)}$ as required.

3. Suppose A satisfies (iii) of Lemma 7.3.1. We have to investigate $A \cup B$ for $B \in \overline{\mathscr{H}}^{(r)}$. By Lemma 7.3.1 and the induction hypothesis, the only cases are when B satisfies (i), (iii), (v) of Lemma 7.3.1. We now consider all these cases:

   (a) When B satisfies (i) of the Lemma 7.3.1 then $B \in \overline{\mathscr{F}}^{(r)}$ and thus, $B \cup A \in \overline{\mathscr{H}}^{(r)} \cup \{A\}$, as required.

   (b) Suppose B satisfies (iii) of the Lemma 7.3.1. Thus by $P(r, \text{(iii)})$ we have that $B \in \mathscr{C}_r$. Therefore by $P(r, \text{(vii)})$, $\mathscr{C}_r$ is an up-set and hence $A \cup B \in \mathscr{C}_r \subseteq \overline{\mathscr{H}}^{(r)}$ as required.

   (c) Suppose B satisfies (v) of Lemma 7.3.1. Therefore, $|B_1| \geqslant k - r + 1$. And hence, $|(A \cup B)_1| \geqslant k - r + 1 \Rightarrow A \cup B \in \overline{\mathscr{H}}^{(r)}$.

4. We are given that A satisfies (iv) of Lemma 7.3.1. Now, since $E \in \overline{\mathscr{F}}^{(r)}$, by $P(r, \text{(i)})$, we have that $E \in \overline{\mathscr{H}}^{(r)}$. Therefore, $A \notin \overline{\mathscr{H}}^{(r+1)}$ as required.

5. We are given that $|A_1| \geqslant k - (r+1) + 1 = k - r$. If $|A_1| \geqslant k - r + 1$ then by the induction hypothesis, $A \in \overline{\mathscr{H}}^{(r)}$ and hence there would be nothing further to investigate. Therefore, we only need to look at the case where $|A| = k - r$. If $B \in \overline{\mathscr{H}}^{(r)}$ is such that $B_1 \nsubseteq A_1$ then $|(A \cup B)_1| \geqslant k - r + 1$ and hence $A \cup B \in \overline{\mathscr{H}}^{(r)}$ by $P(r, \text{(v)})$. On the other hand, if $B_1 \subseteq A_1$ then by $P(r, \text{(vi)})$, either $B_2 = \emptyset$ or $B_2 = [n] \setminus [k]$. If $B_2 = \emptyset$ then $A \cup B = B$ and finally if $B_2 = [n] \setminus [k]$ then $B \in \mathscr{C}_r$ and since $\mathscr{C}_r$ is an up-set by $P(r, \text{(vii)})$, we get that $A \cup B \in \mathscr{C}_r \subseteq \overline{\mathscr{H}}^{(r)}$, as required.

6. Suppose A satisfies (vi) of Lemma 7.3.1. Using Corollary 7.2.4 we have, $\mathscr{A}_{k-r,k} \subseteq \overline{\mathscr{F}}^{(r)}$. Let us consider a set $B \in \mathscr{A}_{k-r,k}$ with $A_1 \subseteq B$. By $P(r, \text{(i)})$, this means that $B \in \overline{\mathscr{H}}^{(r)}$. And note that $|(A \cup B)_1| = k - r$ with $(A \cup B)_2 = A_2$. Thus, by $P(r, \text{(vi)})$, $A \cup B \notin \overline{\mathscr{H}}^{(r)}$ and clearly $A \cup B \neq A$ since $|A_1| \leqslant k - r - 1$. Therefore, $A \notin \overline{\mathscr{H}}^{(r+1)}$, as required.

7. Suppose $A \in \mathscr{C}_{r+1}$ and $A \subseteq B$. By $P(r+1,$(iii)$)$ and $P(r+1,$(iv)$)$, we have that when $A_2 = [n] \setminus [k]$ then

$$A \in \mathscr{C}_{r+1} \Leftrightarrow \forall E \in \overline{\mathscr{F}}^{(t)} \text{ with } E \nsubseteq A_1, E \cup A \in \overline{\mathscr{H}}^{(t)}.$$

We need to show that $B \in \mathscr{C}_{r+1}$. Suppose we have $E \in \overline{\mathscr{F}}^{(t)}$ with $E \nsubseteq B_1$. Since, $A_1 \subseteq B_1$, we also have that $E \nsubseteq A_1$. Therefore, $E \cup A \in \overline{\mathscr{H}}^{(t)}$. This means, by definition, that $E \cup A \in \mathscr{C}_r$. Using $P(r,(\text{vii}))$, this gives us that $E \cup B \in \mathscr{C}_r \subseteq \overline{\mathscr{H}}^{(r)}$, thus proving that $\mathscr{C}_{r+1}$ is an up-set.

□

Finally, let us show that:

Corollary 7.3.3. *$\mathscr{H}$ is $(k+1)-$dense*

Proof. Let us consider $2^{[n]} \setminus \overline{\mathscr{H}}^{(k)}$. Suppose $A \in 2^{[n]} \setminus \overline{\mathscr{H}}^{(k)}$. Then by $P(k,$(i)$)$ and $P(k,$(ii)$)$, $A \notin 2^{[k]}$. By $P(k,$(v)$)$, $A_1 = \emptyset$. By $P(k,$(vi)$)$, whenever $A_2 \neq \emptyset$ and $A_2 \neq [n] \setminus [k]$, we have $A \in 2^{[n]} \setminus \overline{\mathscr{H}}^{(k)}$. Thus, the only possibilities for $2^{[n]} \setminus \overline{\mathscr{H}}^{(k)}$ are either $2^{[n] \setminus [k]}$ or $2^{[n] \setminus [k]} \setminus \{[n] \setminus [k]\}$. In either case, $2^{[n]} \setminus \overline{\mathscr{H}}^{(k)}$ is a down set and hence $\overline{\mathscr{H}}^{(k)}$ is an up-set. And since $\overline{\mathscr{H}}^{(k)} \neq 2^{[n]}$, we have that by Lemma 7.2.7 $\overline{\mathscr{H}}^{(k)}$ is $1-$dense and hence $\mathscr{H}$ is $(k+1)-$dense. □

Remark 7.3.4. *If f_k is the number of labelled union-closed families with universe k the Corollary 7.3.3 tells us that are at least $\binom{n}{k-1} f_{k-1}$ labelled union-closed families that are $k-$dense(since there are $\binom{n}{k-1}$ possible choice for the universe of $\mathscr{F}$). In particular, the number of $(n-1)-$dense union-closed families is at least $\binom{n}{n-2} f_{n-2}$.*

Remark 7.3.5. *Suppose we are given integers c, k and n with $c < k \leqslant n+2$. Then consider the family $\mathscr{F} = \{[k-c], [k-c+1], \ldots, [k-1], [n]\}$. It is easy to see that $s(\mathscr{F}) = c$ and it follows from Corollary 7.3.3 that $\mathscr{F}$ is $k-$dense. Thus, for all possible feasible combinations of c, k and n, there exists a family $\mathscr{F}$ over universe $[n]$ having density k and $s(\mathscr{F}) = c$.*

7.4 Another example of an $(n-1)-$dense family

Let $\mathscr{F}$ be the union-closed family generated by the collection of subsets $\mathscr{B} = \{\{1,2\}, \{2,3\}, \cdots, \{n-1,n\}\}$, i.e.

$$\mathscr{F} = \left\{ \bigcup_{A \in \mathscr{C}} A : \mathscr{C} \subseteq \mathscr{B}, \mathscr{C} \text{ non-empty} \right\}.$$

Another way of describing $\mathscr{F}$ is as follows:

$$\mathscr{F} = \{A \subseteq [n] : A \neq \emptyset \text{ and } i \in A \text{ implies that either } i-1 \in A \text{ or } i+1 \in A\}.$$

Assume $n \geqslant 6$. This section is devoted to providing a description of the closures of $\mathscr{F}$ and consequently establishing that it is $(n-1)-$dense. We begin our discourse with the following lemma.

Lemma 7.4.1. *For any $A \notin \mathscr{F}$ and $|A| \leqslant n-k-1$ for some $1 \leqslant k \leqslant n-1$, at least one of the following is true:*

(i) $\{1\} \subseteq A \subseteq \{1,3,\cdots,n\}$;

(ii) $\{n\} \subseteq A \subseteq \{1,2,3,\cdots,n-2,n\}$;

(iii) $\{1,2\} \subseteq A \subseteq \{1,2,\cdots,n-1\}$ *and* $|A| = n-k-1$;

(iv) $\{n-1,n\} \subseteq A \subseteq \{2,3,\cdots,n\}$ *and* $|A| = n-k-1$;

(v) $A \subseteq \{2,3,\cdots,n-1\}$ *and* $|A| = n-k-1$;

(vi) $\{1,2,n-1,n\} \subseteq A$ *and* $|A| = n-k-1$;

(vii) $A \subseteq \{1,2,\ldots,n-2\}$, $A \cap \{1,2\} \neq \emptyset$, $|A| = n-k-2$ *and* $A \setminus \{\max A\} \notin \mathscr{F}$, *where* $\max A$ *indicates the largest element contained in A;*

(viii) $A \subseteq \{3,4,\ldots,n\}$, $A \cap \{n-1,n\} \neq \emptyset$, $|A| = n-k-2$ *and* $A \setminus \{\min A\} \notin \mathscr{F}$, *where* $\min A$ *indicates the smallest element contained in A;*

(ix) $|A| = n-k-2$, $A \cap \{1,2\} \neq \emptyset$ *and* $A \cap \{n-1,n\} \neq \emptyset$;

(x) $A \subseteq \{1,2,\ldots,n-2\}$, $|A| = n-k-2$ *and* $A \setminus \{\max A\} \in \mathscr{F}$;

(xi) $A \subseteq \{3,4,\ldots,n\}$, $|A| = n-k-2$ *and* $A \setminus \{\min A\} \in \mathscr{F}$;

(xii) $A \subseteq \{3,4,\ldots,n-2\}$, $|A| = n-k-2$ *and* $\exists j \in A$ *such that* $\min A < j < \max A$ *and* $j-1, j+1 \notin A$;

(xiii) $A \subseteq \{3,4,\ldots,n-2\}$, $|A| = n-k-2$ *and* $\forall j \in A$ *such that* $\min A < j < \max A$ *we have* $j-1 \in A$ *or* $j+1 \in A$;

(xiv) $|A| \leqslant n-k-3$

Proof. We split the proof into the following cases:

1. First, we consider $|A| = n-k-1$. If $1 \in A$ and $2 \notin A$, then (i) holds. If $1 \in A$, $2 \in A$ and $n-1 \notin A$, then (ii) holds. On the other hand, if $1,2,n-1 \in A$, then (iii) holds when $n \notin A$ and (iv) holds when $n \in A$. If $1,n \notin A$, then (v) holds. If $1 \notin A$ and $n \in A$, then (ii) holds when $n-1 \notin A$ and (iv) holds when $n-1 \in A$.

2. Next, we consider $|A| = n-k-2$. If $A \cap \{n-1,n\} = \emptyset$ and $A \cap \{1,2\} = \emptyset$ then A satisfies either (xii) or (xiii) . On the other hand, if $A \cap \{n-1,n\} = \emptyset$ and $A \cap \{1,2\} \neq \emptyset$ then (vii)

holds when $A \setminus \{\max A\} \notin \mathscr{F}$ and (x) holds when $A \setminus \{\max A\} \in \mathscr{F}$. Next, if $A \cap \{n-1,n\} \neq \emptyset$ and $A \cap \{1,2\} = \emptyset$ then (viii) holds when $A \setminus \{\min A\} \notin \mathscr{F}$ and (xi) holds when $A \setminus \{\min A\} \in \mathscr{F}$. And if $A \cap \{n-1,n\} \neq \emptyset$ and $A \cap \{1,2\} \neq \emptyset$, then (ix) holds.

3. Finally, when $|A| \leqslant n-k-3$, (xiv) holds. □

Equipped with Lemma 7.4.1, we now state Theorem 7.4.2 which provides a description of the closures of $\mathscr{F}$.

Theorem 7.4.2. *Given any positive integer $k \leqslant n-5$ and any subset A of $[n]$ with $A \notin \mathscr{F}$ and $|A| \leqslant n-k-1$, if A satisfies any of the conditions stated in (i), (ii), (vii), (viii), (ix), (xii), (xiv) of Lemma 7.4.1, then $A \notin \overline{\mathscr{F}}^{(k)}$. On the other hand, if A satisfies any of the conditions (iii), (iv), (v), (vi), (x), (xi), (xiii) then $A \in \overline{\mathscr{F}}^{(k)}$.*

Proof. Define $N = \{$(i), (ii), (vii), (viii), (ix), (xii), (xiv)$\}$ and $M = \{$(iii), (iv), (v), (vi), (x), (xi), (xiii)$\}$. For $q \in N$, define $P(l,q)$ be the statement that: if A satisfies statement q of Lemma 7.4.1 for $k = l$ then $A \notin \overline{\mathscr{F}}^{(l)}$. And for $q \in M$, define $P(l,q)$ be the statement that: if A satisfies statement q of Lemma 7.4.1 for $k = l$ then $A \in \overline{\mathscr{F}}^{(l)}$. Let

$$P(k) = \bigwedge_{q \in N \cup M} P(k,q).$$

To prove the theorem, we have to prove $P(k)$ for $1 \leq k \leq n-5$. The proof is via induction on k. Let us first establish the base case, i.e. $P(1)$.

1. When A satisfies (i) of Lemma 7.4.1, we have $A \cap \{3,4,\ldots,n\} = \{3,\ldots,n\} \setminus \{j\}$. If $j = 3$ then $A \cup \{3,4\} \notin \mathscr{F}$ and if $j > 3$ then $A \cup \{j-1,j\} \notin \mathscr{F}$, and either scenario leads to $A \notin \overline{\mathscr{F}}$.

2. Follows via an argument similar to the argument for $P(1,$(i)$)$, via symmetry.

3. When A satisfies (iii) of Lemma 7.4.1, there exists an i such that $A = [n-1] \setminus \{i\}$. If $i < n-1$ then the only supersets of A are $[n-1]$, $[n]$ and $A \cup \{n\}$, which are all in $\mathscr{F}$. If $i = n-1$ then $A = [n-2]$. Therefore, it suffices to consider $A \cup B$ for the member-subsets $B = \{n-2,n-1\}$ and $B = \{n-1,n\}$ of $\mathscr{F}$. In the former situation, we obtain $A \cup B = [n-1]$ and in the latter, $A \cup B = [n]$, thus proving the statement.

4. Follows via an argument similar to the argument for $P(1,$(iii)$)$, via symmetry.

5. When A satisfies (v) of Lemma 7.4.1, we have $A = \{2,3,\cdots,n-1\} \in \mathscr{F} \subseteq \overline{\mathscr{F}}$.

6. When A satisfies (vi) of Lemma 7.4.1, it is straightforward to check that all supersets of A are already in $\mathscr{F}$, thus yielding $A \in \mathscr{F}^{(1)}$.

7. When A satisfies (vii) of Lemma 7.4.1 then the only possibility for A is the set $A = \{1,3,\ldots,n-2\}$. And $A \cup \{n-1,n\} \notin \mathscr{F} \Rightarrow A \notin \overline{\mathscr{F}}$.

8. Follows via an argument similar to the argument for $P(1,$(vii)$)$, via symmetry.

9. Suppose A satisfies (ix) of Lemma 7.4.1. Since, $A \notin \mathscr{F}$, there is an $j \in A$ such that $j-1, j+1 \notin A$. If there is an $i \leqslant j-3$ such that $i \in A$, $i+1 \notin A$ then $A \cup \{i,i+1\} \notin \mathscr{F}$ therefore, $A \notin \overline{\mathscr{F}}^{(1)}$. Also, if $1 \notin A$ and $1 \leqslant j-3$, then the same argument works by considering $A \cup \{1,2\}$. Otherwise, suppose that for each $i \leqslant j-2$, $i \in A$. Since, $|A| = n-3$ and $A \cap \{n-1,n\} \neq \emptyset$, we must have $j \leqslant n-3$ and hence there is an $i \geqslant j+3$ such that $i \in A$ and ($i-1 \notin A$ or $i+1 \notin A$) which means that $A \cup \{i-1,i\} \notin \mathscr{F}$ or $A \cup \{i,i+1\} \notin \mathscr{F}$ respectively. In either case, $A \notin \overline{\mathscr{F}}$.

10. When A satisfies (x) of Lemma 7.4.1 then, the only such A is $A = \{1,2,\ldots n-4, n-2\}$. Note that $A \cup B$ for $B \in \mathscr{F}$ can either be $A, [n], [n-1], [n-2], \{1,2,\ldots n-4, n-2, n-1\}$ or $\{1,2,\ldots n-4, n-2, n-1, n\}$. All of these possible sets are in $\mathscr{F}$ and as a result, we have that $A \in \overline{\mathscr{F}}$ ($A \cup B$ cannot be $\{1,2,\ldots n-4, n-2, n\}$ since $n \in A \cup B$ means $n \in B \in \mathscr{F}$ which means that $n-1 \in B \subseteq A \cup B$).

11. Follows via an argument similar to the argument for $P(1,$(x)$)$, via symmetry.

12. For $k = 1$, there is no A satisfying (xii), hence there is nothing to check in this case.

13. For $k = 1$, there is no A satisfying (xiii), hence there is nothing to check in this case.

14. We have $A \notin \mathscr{F}$ and $|A| \leqslant n-4$. Suppose $|A| = 1$. If $A = \{i\}$ and $i \geq 4$ then consider $A \cup \{1,2\} \notin \mathscr{F}$ and hence $A \notin \overline{\mathscr{F}}$. Otherwise, if $i \leqslant 3$, then considering $A \cup \{5,6\}$ gives the required result, i.e. $A \notin \overline{\mathscr{F}}$. So, we can assume that $|A| \geq 2$. Now, since $A \notin \mathscr{F}$, we have a $j \in A$ such that $j-1, j+1 \notin A$. Since $|A| \geq 2$, we may assume, without loss of generality, that there is an $i \in A$ such that $i < j$. Now, if there is an $a \leqslant j-3$ such that $a \in A$ but $a+1 \notin A$ or $a-1 \notin A$, then we have $A \cup \{a,a+1\} \notin \mathscr{F}$ or $A \cup \{a-1,a\} \notin \mathscr{F}$ respectively, which gives $A \notin \overline{\mathscr{F}}$. So, we can assume that $a \leqslant j-2 \Rightarrow a \in A$. Now, if $j = \max A$ then because of the last assumption and the size of A, $j \leqslant n-3$. Therefore, $A \cup \{n-1,n\} \notin \mathscr{F}$. On the other hand if there is a $b > j$, then since $|A| = n-4$, there is a $c > j+2$ such that either $c-1 \notin A$ or $c+1 \notin A$, which in turn implies that either $A \cup \{c-1,c\}$ or $A \cup \{c,c+1\}$ is absent in $\mathscr{F}$, and hence $A \notin \overline{\mathscr{F}}$, as required.

Now, suppose that $P(t)$ is true for some $t \leqslant n-6$. Let us consider the $(t+1)$-st case. By Corollary 7.2.4 we have that $\mathscr{A}_{n-t,n} \subseteq \overline{\mathscr{F}}^{(t)}$.

1. Suppose A satisfies (i) of Lemma 7.4.1. First, consider $A = \{1\}$. We have $\{3,4\} \in \mathscr{F} \subseteq \overline{\mathscr{F}}^{(t)}$, thus yielding $A \cup \{3,4\} = \{1,3,4\}$. Moreover, we have $\{1,3,4\} \subseteq \{1,3,\cdots,n\}$ and $1 \in \{1,3,4\}$. Since $t \leqslant n-4$, we get $|\{1,3,4\}| = 3 \leqslant n-t-1$. Therefore, by $P(t,\text{(i)})$, we conclude that $A \cup \{3,4\} \notin \overline{\mathscr{F}}^{(t)}$, which gives $A = \{1\} \notin \overline{\mathscr{F}}^{(t+1)}$, as required. Now, let us consider the case where $A \neq \{1\}$. Then, there must exist an $i \in \{3,4,\cdots,n\}$ such that $i \in A$. Since $|A| \leqslant n-t-2 \leqslant n-3$, there must exist a $j \in \{3,4,\cdots,n\}$ such that $j \notin A$. Therefore, there exist consecutive $i,j \in \{3,4,\cdots,n\}$ such that $i \in A$ and $j \notin A$. Thus, by $P(t,\text{(i)})$, we have $\{i,j\} \cup A \notin \overline{\mathscr{F}}^{(t)}$ and hence $A \notin \overline{\mathscr{F}}^{(t+1)}$, as required.

2. Follows via an argument similar to the argument for $P(t+1,\text{(i)})$, via symmetry.

3. Suppose A satisfies (iii) of Lemma 7.4.1. We need to show that for each $B \in \overline{\mathscr{F}}^{(t)}$, we have $A \cup B \in \overline{\mathscr{F}}^{(t)} \cup \{A\}$. Since $|A| = n-t-2$ and $\mathscr{A}_{n-t,n} \subseteq \mathscr{A}$, we need only check the case where $|A \cup B| = n-t-1$. We must have $|A \cup B \setminus A| = |B \setminus A| = 1$. Let $B \setminus A = \{i\}$. If $i \neq n$ then by $P(t,\text{(iii)})$ of the induction hypothesis, we have $A \cup B \in \overline{\mathscr{F}}^{(t)}$. This yields $i = n$. Suppose $n-1 \notin A$. Since $B \setminus A = \{n\}$, this means that $n-1 \notin B$. Moreover, $|B| \leqslant |A \cup B| = n-t-1$. Therefore, by $P(t,\text{(ii)})$, we conclude that $B \notin \overline{\mathscr{F}}^{(t)}$, which leads to a contradiction since we started with $B \in \overline{\mathscr{F}}^{(t)}$. Therefore, we must have, $\{1,2,n-1,n\} \subset A \cup B$, and $|A \cup B| = n-t-1$, which, by $P(t,\text{(vi)})$, gives us $A \cup B \in \overline{\mathscr{F}}^{(t)}$, as required. Therefore, $A \in \overline{\mathscr{F}}^{(t+1)}$.

4. Follows via an argument similar to the argument for $P(t+1,\text{(iii)})$, via symmetry.

5. Suppose A satisfies (v) of Lemma 7.4.1. We need to check that for each $B \in \overline{\mathscr{F}}^{(t)}$, we have $A \cup B \in \overline{\mathscr{F}}^{(t)} \cup \{A\}$. As noted above in the argument for $P(t+1,\text{(iii)})$, it is enough to check this for the case when $|A \cup B| = n-t-1$. We have $B \setminus A = \{i\}$ for some i. Suppose $i \neq 1,n$. Then $A \cup B \subseteq \{2,3,\cdots,n-1\}$ and $|A \cup B| = n-t-1$ together yield, by $P(t,5)$, that $A \cup B \in \overline{\mathscr{F}}^{(t)}$. Without loss of generality, suppose $i = 1$. If $2 \notin A$, then since $B \setminus A = \{1\}$, we have $2 \notin B$. But, in that case, we have $1 \in B \subseteq \{1,3,4,\cdots,n\}$ and $|B| \leqslant n-t-1$, which together imply, by $P(t,\text{(i)})$, that $B \notin \overline{\mathscr{F}}^{(t+1)}$, which is a contradiction to our hypothesis. Therefore, we must have $2 \in A$. Therefore, $\{1,2\} \subset A \cup B \subseteq \{1,2,\cdots,n-1\}$ and $|A \cup B| = n-t-1$, so that by $P(t,\text{(iii)})$, we have $A \cup B \in \overline{\mathscr{F}}^{(t)}$, which in turn gives us $A \in \overline{\mathscr{F}}^{(t+1)}$.

6. Suppose A satisfies (vi) of Lemma 7.4.1. We need to check that for each $B \in \overline{\mathscr{F}}^{(t)}$, we have $A \cup B \in \overline{\mathscr{F}}^{(t)} \cup \{A\}$. As noted above in the argument for $P(t+1,\text{(iii)})$, it is enough to

check this for the case when $|A \cup B| = n-t-1$. In this case, $\{1,2,n-1,n\} \subseteq A \cup B$ and $|A \cup B| = n-t-1$ together imply, by $P(t,\text{(vi)})$, that $A \cup B \in \overline{\mathscr{F}}^{(t)}$. Therefore, $A \in \overline{\mathscr{F}}^{(t+1)}$, as required.

7. Suppose A satisfies (vii) of Lemma 7.4.1. Since $A - \{\max A\} \notin \mathscr{F}$, there exists a $j < \max A$ such that $j-1, j+1 \notin A$. Therefore, $A \cup \{\max A, \max A+1\} \notin \mathscr{F}$. Moreover, $A \neq A \cup \{\max A, \max A+1\}$. If $\max A = n-2$, then $A \cup \{\max A, \max A+1\} \notin \overline{\mathscr{F}}^{(t)}$, by $P(t,\text{(ix)})$. Otherwise, if $\max A < n-2$, then $A \cup \{\max A, \max A+1\} \notin \overline{\mathscr{F}}^{(t)}$, by $P(t,\text{(vii)})$. Either way, we have $A \notin \overline{\overline{\mathscr{F}}^{(t)}} = \overline{\mathscr{F}}^{(t+1)}$.

8. Follows via an argument similar to the argument for $P(t+1,\text{(vii)})$, via symmetry.

9. Suppose A satisfies (ix) of Lemma 7.4.1. Since $A \notin \mathscr{F}$, there exists a $j \in A$ such that $j-1, j+1 \notin A$. If there is an $i \leqslant j-3$ such that $i \in A$ and $i+1 \notin A$, then by $P(t,\text{(ix)})$, we have $A \cup \{i, i+1\} \notin \overline{\mathscr{F}}^{(t)} \Rightarrow A \notin \overline{\mathscr{F}}^{(t+1)}$. Moreover, if $1 \notin A$ and $1 \leqslant j-3$, then by considering $A \cup \{1,2\}$, the same argument works. Otherwise, suppose $i \in A$ for each $i \leqslant j-2$. Since $|A| = n-t-3 \leqslant n-4$ and $A \cap \{n, n+1\} \neq \emptyset$, we must have $j \leqslant n-3$ and hence there is an $i \geq j+3$ such that $i \in A$ and ($i-1 \notin A$ or $i+1 \notin A$) which means that $A \cup \{i-1, i\} \notin \mathscr{F}$ or $A \cup \{i, i+1\} \notin \mathscr{F}$ respectively. In either case, $A \notin \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(ix)})$.

10. Suppose A satisfies (x) of Lemma 7.4.1. We need to check that for each $B \in \overline{\mathscr{F}}^{(t)} \cup \{A\}$, we have $A \cup B \in \overline{\mathscr{F}}^{(t)}$. Since, $|A| = n-t-3$, we only need to check the cases $|A \cup B| = n-t-2$ and $|A \cup B| = n-t-1$. Since $B \in \overline{\mathscr{F}}^{(t)}$, we divide the proof into the following parts:

    (a) First, we consider the case where $B \in \mathscr{F}$. Note that if $\max B > \max A + 1$, then $\max B - 1 \notin A$ and hence we can assume that $|A \cup B| = n-t-1$, which in turn implies that $A \cup B = A \cup \{\max B - 1, \max B\}$. If $\max B \leqslant n-1$ and $1 \notin A$, then $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(v)})$, and if $\max B = n$ and $1 \notin A$, then $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(iv)})$. On the other hand, if $1 \in A$, then $2 \in A$ because $A \setminus \{\max A\} \in \mathscr{F}$. So, if $1 \in A$ and $\max B \leq n-1$, then $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(iii)})$, and if $1 \in A$ and $\max B = n$, then $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(vi)})$. If $\max B \in \{\max A - 1, \max A, \max A + 1\}$, then from $A \setminus \{\max A\} \in \mathscr{F}$, it follows that $A \cup B \in \mathscr{F} \subseteq \overline{\mathscr{F}}^{(t)}$. Finally, we consider the case where $\max B \leq \max A - 1$. We first assume that $|A \cup B| = n-t-1$. If $1 \notin A$ and $1 \notin B$, then $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(v)})$. Now, recall that if $1 \in A$, then $2 \in A$ and if $1 \in B$ then $2 \in B$. In both these cases, $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(iii)})$. On the other hand, if $|A \cup B| = n-t-2$, then it is straightforward to see that $A \cup B \setminus \{\max(A \cup B\} = A \cup B \setminus \{\max A\} \in \mathscr{F}$. Therefore, by $P(t,\text{(x)})$, we have $A \cup B \in \overline{\mathscr{F}}^{(t)}$.

(b) Next, we consider the case where B satisfies (x) of Lemma 7.4.1. Note that since $|B| = n-t-2$, we can assume that $|A \cup B| = n-t-1$. So, if $1 \notin A$ and $1 \notin B$, then $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(v)})$. And if $1 \in A$ or $1 \in B$, then $2 \in A \cup B$, and therefore, $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(iii)})$.

(c) We consider the case where B satisfies (xi) of Lemma 7.4.1. Since, $|B| = n-t-2$, we can assume that $|A \cup B| = n-t-1$. Suppose, if possible, $A \cup B \notin \overline{\mathscr{F}}^{(t)}$. By Lemma 7.4.1 and the induction hypothesis, $A \cup B$ either satisfies (i) or (ii). If it satisfies (i), then we must have $1 \in A \cup B$ and $2 \notin A \cup B$. But $1 \in A \cup B$ implies that $1 \in A$ (since $B \subseteq \{3,4,\ldots,n\}$), and $1 \in A$, in turn, implies that $2 \in A \subseteq A \cup B$. A similar argument works for the case where $A \cup B$ satisfies (ii).

(d) Finally, we consider the case where B satisfies (xiii) of Lemma 7.4.1. We may once again assume that $|A \cup B| = n-t-1$. Note that since $B \subseteq \{3,4,\ldots,n-2\}$, if $1 \notin A$ then $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(v)})$, and if $1 \in A$ then $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(iii)})$.

11. Follows via an argument similar to the argument for $P(t+1,\text{(x)})$, via symmetry.

12. If A satisfies (xii) of Lemma 7.4.1 and $\min A = 3$, then $\{\min A - 1, \min A\} \cup A \notin \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(vii)})$. On the other hand, if $\min A > 3$, then $\{\min A - 1, \min A\} \cup A \notin \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(xii)})$.

13. Suppose A satisfies (xiii) of Lemma 7.4.1. We need to check that for each $B \in \overline{\mathscr{F}}^{(t)}$, we have $A \cup B \in \overline{\mathscr{F}}^{(t)} \cup \{A\}$. Since, $|A| = n-t-3$, we only need to check the cases where $|A \cup B| = n-t-2$ and $|A \cup B| = n-t-1$. Since $B \in \overline{\mathscr{F}}^{(t)}$, we divide the proof into the following parts:

(a) First, we consider the case where $B \in \mathscr{F}$. Let us first suppose that $\min A - 1 < \min B < \max B < \max A$, which gives $\max A \cup B = \max A$ and $\min A \cup B = \min A$. Suppose, first that $|A \cup B| = n-t-2$. Consider $j \in A \cup B$ such that $\min A < j < \max A$. If $j \in A$ then $j-1 \in A \cup B$ or $j+1 \in A \cup B$ because of (xiii). On the other hand, if $j \in B$ then because $B \in \mathscr{F}$, we obtain that $j-1 \in B$ or $j+1 \in B$. Therefore, $A \cup B$ satisfies (xiii) and hence $A \cup \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(xiii)})$. Next, if $|A \cup B| = n-t-1$ then $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(v)})$. On the other hand, let us consider the case $\min A - 1 < \min B < \max B = \max A + 1$ and $A \cup B \notin \mathscr{F}$. If $|A \cup B| = n-t-2$ then, it is easy to see that $A \cup B \setminus \{\min(A \cup B)\} = A \cup B \setminus \{\min A\} \in \mathscr{F}$ and hence $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(xi)})$. And if $|A \cup B| = n-t-1$ then, $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t,\text{(v)})$. Analogous arguments go through for the case when $\min A - 1 = \min B < \max B < \max A + 1$. Now, if $\min A - 1 = \min B < \max B = \max A + 1$ then since A satisfies (xiii), it is easy to see that $A \cup B \in \overline{\mathscr{F}}^{(t)}$. Finally, suppose that

$\max B > \max A + 1$. In this case, $\max B - 1, \max B \notin A$ and since $B \in \mathscr{F}$, we do have that $\max B - 1 \in B$. Also, since we are assuming that $|A \cup B| \leqslant n - t - 1$, we have that $A \cup B = A \cup \{\max B - 1, \max B\}$. If $\max B \leqslant n - 1$ then, $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t, \text{(v)})$ and if $\max B = n$ then $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t, \text{(iv)})$. Analogous argument works for the case when $\min B < \min A - 1$.

(b) Next, we consider the case where B satisfies (x) of Lemma 7.4.1. It means that $|B| = n - t - 2$ therefore, we can assume that $|A \cup B| = n - t - 1$. If $\min B \geqslant 2$ then $A \cup B \in \overline{\mathscr{F}}^{(t)}$ by $P(t, \text{(v)})$ and on the other hand, if $\min B = 1$ then, since $B \setminus \{\max B\} \in \mathscr{F}$ and $t \leqslant n - 5 \Rightarrow |B| = n - t - 2 \geqslant 3$, we can conclude that $2 \in B$. But note that $1, 2 \notin A$ which give us that $|A \cup B| \geqslant n - k$.

(c) The case when B satisfies (xi) of Lemma 7.4.1 is analogous to the previous part, via symmetry.

(d) Finally, we consider the case when B satisfies (xiii) of Lemma 7.4.1. Since, in this case, we have that $|B| = n - t - 2$, we can again assume that $|A \cup B| = n - t - 1$ which, by $P(t, \text{(v)})$, gives us that $A \cup B \in \overline{\mathscr{F}}^{(t)}$.

14. Here, we have $A \notin \mathscr{F}$ and $|A| \leqslant n - t - 4$. Suppose $|A| = 1$. Therefore, if $A = \{i\}$ and $i \geq 4$, then consider $A \cup \{1, 2\} \notin \mathscr{F}$, which gives us $A \cup \{1, 2\} \notin \overline{\mathscr{F}}^{(t)}$ by $P(t, \text{(xiv)})$. Otherwise, if $i \leqslant 3$, then considering $A \cup \{5, 6\}$ gives us the required result, i.e. $A \notin \overline{\mathscr{F}}^{(t+1)}$. So now, we may assume that $|A| \geq 2$. Since $A \notin \mathscr{F}$, there exists a $j \in A$ such that $j - 1, j + 1 \notin A$. Since $|A| \geq 2$, we can assume, without loss of generality, that there is an $i \in A$ such that $i < j$. Now, if there is an $a \leqslant j - 3$ such that $a \in A$ but $a + 1 \notin A$ or $a - 1 \notin A$, then $A \cup \{a, a + 1\}$ or $A \cup \{a - 1, a\}$, respectively, satisfies (xiv). These are, therefore, not in $\overline{\mathscr{F}}^{(t)}$ by $P(t, \text{(xiv)})$, which gives $A \notin \overline{\mathscr{F}}^{(t+1)}$. So now, we may assume that $a \leqslant j - 2 \implies a \in A$. If $j = \max A$, then because of the last assumption and the size of A, we have $j \leqslant n - t - 3 \leqslant n - 4$. If $|A| = n - t - 4$, then $|A \cup \{n - 1, n\}| = n - t - 2$ and hence $A \cup \{n - 1, n\} \notin \overline{\mathscr{F}}^{(t)}$ by $P(t, \text{(ix)})$. If $|A| \leqslant n - t - 5$, then $|A \cup \{n - 1, n\}| \leqslant n - t - 3$ and hence $A \cup \{n - 1, n\} \notin \overline{\mathscr{F}}^{(t)}$ by $P(t, \text{(xiv)})$. On the other hand, if there is a $b > j$ with $b \in A$, then since $|A| = n - t - 4 \leqslant n - 5$, there is a $c > j + 2$ such that $c - 1 \notin A$ or $c + 1 \notin A$, which means that $A \cup \{c - 1, c\}$ or $A \cup \{c, c + 1\}$, respectively, satisfies (xiv), and hence $A \notin \overline{\mathscr{F}}^{(t+1)}$, as required.

This brings us to the end of the proof that $P(k)$ is true for all $k \leqslant n - 5$. $\square$

Let us note that Lemma 7.4.1, together with Theorem 7.4.2, describes all $\overline{\mathscr{F}}^{(k)}$'s for $k \leqslant n - 5$. This is because we know that $\mathscr{F} \subseteq \overline{\mathscr{F}}^{(k)}$ and $\mathscr{A}_{n-k} \subseteq \overline{\mathscr{F}}^{(k)}$. Therefore, if we have an $A \subseteq 2^{[n]}$ such that $A \notin \mathscr{F}$ and $|A| \leqslant n-k-1$, then by Lemma 7.4.1, it satisfies one of the statements (i), (ii), (iii), (iv), (v), (vi), (vii), (viii), (ix), (x), (xi), (xii), (xiii),(xiv), but then that statement together with the corresponding statement in Theorem 7.4.2 tells us whether $A \in \overline{\mathscr{F}}^{(k)}$ or $A \notin \overline{\mathscr{F}}^{(k)}$.

We now come to the proof of the fact that $\mathscr{F}$ is $(n-1)$-dense.

Corollary 7.4.3. *$\mathscr{F}$ is $(n-1)$-dense.*

Proof. First, we establish that $\{3\} \in \overline{\mathscr{F}}^{(n-4)}$. Since $\mathscr{A}_{5,n} \subseteq \overline{\mathscr{F}}^{(n-5)}$, we only need to investigate $\{3\} \cup B$ for all $B \in \overline{\mathscr{F}}^{(n-5)}$ such that $|\{3\} \cup B| \leqslant 4$. The cases where $|\{3\} \cup B| = 1,2$ are trivial. Let us consider the cases where $|\{3\} \cup B| = 3$. If $|B| = 3$, then $\{3\} \cup B = B \in \overline{\mathscr{F}}^{(n-5)}$. Therefore, let $|B| = 2$. By $P(n-5,$(xiv)$)$, this means that $B \in \mathscr{F}$. Therefore, let $B = \{i, i+1\}$. If $i \leqslant 4$, then $\{3\} \cup \{i, i+1\} \in \mathscr{F} \subseteq \overline{\mathscr{F}}^{(n-5)}$. If $i > 4$, then by $P(n-5,$(xi)$)$, we have $\{3\} \cup B \in \overline{\mathscr{F}}^{(n-5)}$. The only case that is left to consider is $|\{3\} \cup B| = 4$. Again, we may assume that $|B| = 3$ and $3 \notin B$. Therefore, B satisfies either (x) or (xi). If B satisfies (x), then $B \notin \mathscr{F}$ and $B \setminus \{\max B\} \in \mathscr{F}$. Since $|B| = 3$, this means that $B = \{i, i+1, j\}$ such that $i+1 < j \leqslant n-2$. If $i = 1$, then $\{3\} \cup B \in \overline{\mathscr{F}}^{(n-5)}$ by $P(n-5,$(iii)$)$; if $i \geq 4$ then $\{3\} \cup B \in \overline{\mathscr{F}}^{(n-5)}$ by $P(n-5,$(v)$)$; if $i = 2$ or $i = 3$, then $\{3\} \cup B = B$. On the other hand, if B satisfies (xi), then B is of the form $\{j, i-1, i\}$ where $3 \leqslant j < i-1$. If $i = n$, then $\{3\} \cup B \in \overline{\mathscr{F}}^{(k)}$ because of $P(n-5,$(iv)$)$, and if $i < n$ then $\{3\} \cup B \in F^{(k)}$ because of $P(n-5,$(v)$)$. Therefore, $\{3\} \in \overline{\mathscr{F}}^{(n-4)}$.
Since $\{1,3,4\} \cup \{4,5\} = \{1,3,4,5\} \notin \mathscr{F}^{(n-5)}$, by $P(n-5,$(i)$)$, we have $\{1,3,4\} \notin \overline{\mathscr{F}}^{(n-4)}$. Now, consider the subsets, $\{3\} \subseteq \{3,4\} \subseteq \{1,3,4\}$. Here, $\{3\}, \{3,4\} \in \overline{\mathscr{F}}^{(n-4)}$ and $\{1,3,4\} \notin \overline{\mathscr{F}}^{(n-4)}$. Therefore, by Theorem 7.2.8, $\overline{\mathscr{F}}^{(n-4)}$ is at least $3-$dense, which means that $\mathscr{F}$ is at least $(n-1)$-dense. This concludes the proof. $\square$

7.5 Relative subsets and closure roots

Definition 7.5.1. *Given a union-closed family $\mathscr{F}$ over the universe $[n]$, and $A, B \in \mathscr{F}$, we define A to be a* subset of B relative to $\mathscr{F}$*, and write $A \subseteq_{\mathscr{F}} B$, if at least one of the following happens:*

- *$A = B$.*
- *$B = [n]$*
- *$\exists C \in \mathscr{F}$ such that $C \neq B$ and $A \cup C = B$.*

We write $A \subsetneq_{\mathscr{F}} B$ if $A \subseteq_{\mathscr{F}} B$ and $A \neq B$, and $A \nsubseteq_{\mathscr{F}} B$ when A is not a subset of B relative to $\mathscr{F}$.

Note that $A \subseteq_{\mathscr{F}} B$ implies that $A \subseteq B$. It is straightforward to see that this definition coincides with the usual notion of subsets if $\mathscr{F} = 2^{[n]}$. To see that it actually differs from the usual notion of subsets when $\mathscr{F}$ is not $2^{[n]}$, it is enough to consider$\mathscr{F} = \{\{1\},\{1,2\},\{2,3\},\{1,2,3\}\}$ and observe that $\{1\} \not\subseteq_{\mathscr{F}} \{1,2\}$.

Lemma 7.5.2. *Let $\mathscr{F}$ be a 1-dense family. Consider any $A,B,C \in \mathscr{F}$. Then the following hold:*

1. *$A \subsetneq_{\mathscr{F}} B$ and $B \subseteq C$ together imply that $A \subsetneq_{\mathscr{F}} C$.*
2. *$A \subsetneq B$ and $B \subseteq C$ together imply that $B \subseteq_{\mathscr{F}} C$.*

Proof. We prove the two parts separately, as follows.

1. If $C = B$, the conclusion is immediate. Assume, therefore, that $B \subsetneq C$. By definition of the relation $A \subsetneq_{\mathscr{F}} B$, there exists $B_1 \in \mathscr{F}$ such that $A \cup B_1 = B$ and $B_1 \subsetneq B$. Note that on one hand, $(C \setminus B) \cup B_1 \subsetneq C$, and on the other, $B_1 \subseteq (C \setminus B) \cup B_1$. The latter implies that $(C \setminus B) \cup B_1 \in \mathscr{F}$ (since by Lemma 7.2.7, we know that $\mathscr{F}$ is an up-set.) Moreover, $A \cup (C \setminus B) \cup B_1 = (A \cup B_1) \cup (C \setminus B) = B \cup (C \setminus B) = C$. Together, these observations allow us to conclude that $A \subsetneq_{\mathscr{F}} C$.

2. If $B = C$, then there is nothing to prove. Otherwise, note that the set $A \in \mathscr{F}$ and from Lemma 7.2.7, we know that $\mathscr{F}$ is an up-set, and these two together imply that $A \cup (C \setminus B) \in \mathscr{F}$. Moreover, $A \neq B$ implies that $A \cup (C - B) \neq C$, and $C = B \cup (A \cup (C \setminus B))$. Combining all these, we conclude that $B \subseteq_{\mathscr{F}} C$. ☐

Corollary 7.5.3. *The relation $\subseteq_{\mathscr{F}}$ of relative subset is transitive when the family $\mathscr{F}$ is 1-dense.*

This follows from using the fact that $B \subseteq_{\mathscr{F}} C$ implies $B \subseteq C$, and using 1 of Lemma 7.5.2. However, this is not necessarily true when $\mathscr{F}$ is not 1-dense. As an instance, if we consider $\mathscr{F} = \{\{1\},\{2\},\{1,2\},\{1,3\},\{1,2,3\},\{1,2,3,4\}\}$, then it can be checked that $\{1\} \subseteq_{\mathscr{F}} \{1,2\} \subseteq_{\mathscr{F}} \{1,2,3\}$ but $\{1\} \not\subseteq_{\mathscr{F}} \{1,2,3\}$. It would be interesting to try to characterize union-closed families $\mathscr{F}$ such that $\subseteq_{\mathscr{F}}$ is transitive.

Proposition 7.5.4. *Let $\mathscr{F}$ and $\mathscr{H} \subseteq \mathscr{F}$ be two union-closed families over the universe $[n]$. Then $\overline{\mathscr{H}} \supseteq \mathscr{F}$ if and only if for all $A \in \mathscr{H}$ and $B \in \mathscr{F}$ with $A \subseteq_{\mathscr{F}} B$, we have $B \in \mathscr{H}$.*

Proof. Suppose $\overline{\mathscr{H}} \supseteq \mathscr{F}$. Consider any $A \in \mathscr{H}$ and $B \in \mathscr{F}$ with $A \subseteq_{\mathscr{F}} B$. If $A = B$ or $B = [n]$, the conclusion is immediate. Therefore, we consider the third possibility in Definition 7.5.1, i.e. there exists $C \in \mathscr{F}$ such that $A \cup C = B$ and $C \neq B$. Now, $C \in \mathscr{F} \subseteq \overline{\mathscr{H}}$, which implies that $\mathscr{H} \cup \{C\}$ is union-closed. Since $A \in \mathscr{H}$, we must have $B = A \cup C \in \mathscr{H} \cup \{C\}$, and $B \neq C$. Therefore, we must have $B \in \mathscr{H}$, as claimed.

We now prove the converse. Consider $C \in \mathscr{F}$. We want to show that $C \in \overline{\mathscr{H}}$, i.e. the family $\mathscr{H} \cup \{C\}$ is union-closed. Thus, for every $A \in \mathscr{H}$, we want to show that the subset $A \cup C$ either equals C or is in $\mathscr{F}$. If $A \cup C = A$ or $A \cup C = C$ then there's nothing further to check. Otherwise, $A \cup C = B$ for some B with $B \neq C$ and $C \in \mathscr{F}$, thus establishing that $A \subseteq_{\mathscr{F}} B$. Moreover, as $A \in \mathscr{H} \subseteq \mathscr{F}$ and $C \in \mathscr{F}$ and $\mathscr{F}$ is union-closed, hence $B \in \mathscr{F}$ as well. Therefore, the hypothesis of the proposition implies that $B \in \mathscr{H}$, hence completing the proof. $\square$

It is worth noting how Proposition 7.5.4 compares with Lemma 7.2.7. In Proposition 7.5.4, if we take $\mathscr{F} = 2^{[n]}$, then $\mathscr{H}$ has to be a 1-dense family and $\forall A \in \mathscr{H}$ and $B \in 2^{[n]}$ such that $A \subseteq B$, we have $B \in \mathscr{H}$, which implies that $\mathscr{H}$ is an up-set, thus corroborating the claim of Lemma 7.2.7. Thus Proposition 7.5.4 is a generalisation of Lemma 7.2.7 when we consider the notion of relative subsets. Let us recall that given a family $\mathscr{F}$, $A \in \mathscr{F}$ is an *inclusion-wise minimal* member subset if for all $B \in \mathscr{F}$ with $B \subseteq A$, we have $A = B$.

Definition 7.5.5. *Consider the family $\mathscr{G}$ of all inclusion-wise minimal member subsets of $\mathscr{F}$. When $\mathscr{F}$ is 1-dense, using Lemma 7.2.7 we see that*

$$\mathscr{F} = \{B \in 2^{[n]} : \exists A \in \mathscr{G} \text{ with } B \supseteq A\}.$$

We write $< \mathscr{G} >= \mathscr{F}$ and say that $\mathscr{G}$ is the generating set of $\mathscr{F}$.

Let us recall the notion of basis sets of a union-closed family(see [27]). A set $A \in \mathscr{F}$ is called a *basis set* if for all $X, Y \in \mathscr{F}$ with $A = X \cup Y$, we have either $X = A$ or $Y = A$. Let $\mathscr{B}$ be the set of all basis sets of $\mathscr{F}$. Then $\mathscr{B}$ is the minimal set with the property that the union-closed family generated by $\mathscr{B}$ is $\mathscr{F}$. We emphasize here that $\mathscr{G} \neq \mathscr{B}$ in general. For example, if the 1-dense family is $\{\{1,\},\{1,2\},\{1,3\},\{1,2,3\}\}$ then $\mathscr{G} = \{\{1\}\}$ but $\mathscr{B} = \{\{1\},\{1,2\},\{1,3\}\}$.

Definition 7.5.6. *Let $\mathscr{F}$ be a union-closed family over the universe $[n]$, and let $\mathscr{K} \subseteq \mathscr{F}$. Define the family*

$$< \mathscr{K} >_{\mathscr{F}} = \{B \in \mathscr{F} : \text{ there exists } A \in \mathscr{K} \text{ with } A \subseteq_{\mathscr{F}} B\}. \tag{7.5.1}$$

We say that $\mathscr{K}$ generates $< \mathscr{K} >_{\mathscr{F}}$ relative to $\mathscr{F}$.

We emphasise here that $\mathscr{K} \subseteq < K >_{\mathscr{F}}$. We shall eventually establish that when $\mathscr{G}$, as mentioned above, is the family of all inclusion-wise minimal member subsets of the 1−dense family $\mathscr{F}$, the family $< \mathscr{G} >_{\mathscr{F}}$ serves as a test candidate for determining whether there exists a union-closed family $\mathscr{H}$ such that $\overline{\mathscr{H}} = \mathscr{F}$.

Lemma 7.5.7. *Let $\mathscr{F}$ be a 1-dense union-closed family and $\mathscr{K} \subseteq \mathscr{F}$. Then, $< \mathscr{K} >_{\mathscr{F}}$ is union-closed.*

Proof. Fix any $A, B \in < \mathscr{K} >_{\mathscr{F}}$. Then there exist subsets A_o and B_o in $\mathscr{K}$ such that $A_o \subseteq_{\mathscr{F}} A$ and $B_o \subseteq_{\mathscr{F}} B$. If $A_o = A$ and $B_o = B$ and $A_o \cup B_o$ equals either A_o or B_o, then the conclusion is immediate since $A_o, B_o \in \mathscr{K} \subset < \mathscr{K} >_{\mathscr{F}}$. If $A_o = A$ and $B_o = B$ but $A_o \cup B_o$ equals neither A_o nor B_o, then since $B_o \in \mathscr{K} \subset \mathscr{F}$ and $B_o \neq A_o \cup B_o$, hence we conclude that $A_o \subset_{\mathscr{F}} A_o \cup B_o$. Finally, assume that either $A_o \subsetneq_{\mathscr{F}} A$ or $B_o \subsetneq_{\mathscr{F}} B$ or both. In this case, since $\mathscr{F}$ is a 1-dense union-closed family, A_o, A and $A \cup B$ all belong to $\mathscr{F}$, and $A_o \subsetneq_{\mathscr{F}} A \subset A \cup B$, by 1 of Lemma 7.5.2, we conclude that $A_o \subseteq_{\mathscr{F}} A \cup B \Rightarrow A \cup B \in < \mathscr{K} >_{\mathscr{F}}$. This concludes the proof. □

Given union-closed families $\mathscr{H} \subseteq \mathscr{F}$ and $A \in \mathscr{H}$, let us say that A is $\subseteq_{\mathscr{F}}$ *−wise minimal member-subset* of $\mathscr{H}$ if whenever $B \in \mathscr{H}$ and $B \subseteq_{\mathscr{F}} A$, we have that $B = A$.

Proposition 7.5.8. *Let $\mathscr{F}$ be a 1-dense family and $\mathscr{H} \subseteq \mathscr{F}$ be union-closed. Then, the following are equivalent:*

1. *$\overline{\mathscr{H}} \supseteq \mathscr{F}$.*

2. *For all $A \in \mathscr{H}$ and $B \in \mathscr{F}$ such that $A \subseteq_{\mathscr{F}} B$, we have $B \in \mathscr{H}$.*

3. *The set $\mathscr{K} \subseteq \mathscr{H}$ of all $\subseteq_{\mathscr{F}}$ −wise minimal member-subsets of $\mathscr{H}$ satisfies $< \mathscr{K} >_{\mathscr{F}} = \mathscr{H}$.*

Proof. That 1 and 2 are equivalent follows from Proposition 7.5.4.

We now show that 3 implies 2. Consider any $A \in < \mathscr{K} >_{\mathscr{F}} = \mathscr{H}$. Then there exists $A_o \in \mathscr{K}$ such that $A_o \subseteq_{\mathscr{F}} A$. Suppose there exists $B \in \mathscr{F}$ be such that $A \subseteq_{\mathscr{F}} B$. Then the relations $A_o \subseteq_{\mathscr{F}} A$ and $A \subseteq_{\mathscr{F}} B$ along with Corollary 7.5.3 imply that $A_o \subseteq_{\mathscr{F}} B$, therefore yielding $B \in < \mathscr{K} >_{\mathscr{F}}$.

Finally, we show that 2 implies 3. Consider any $B \in \mathscr{H}$. By minimality, there exists $A \in \mathscr{K}$ such that $A \subseteq_{\mathscr{F}} B$. This implies that $B \in < \mathscr{K} >_{\mathscr{F}}$. Conversely, consider any $B \in < \mathscr{K} >_{\mathscr{F}}$. Then there exists $A \in \mathscr{K}$ such that $A \subseteq_{\mathscr{F}} B$, which then implies that $B \in \mathscr{H}$, by the hypothesis in 2. □

7.5.1 Closure Roots

Definition 7.5.9. *For a union-closed family $\mathscr{F}$, we say that a union-closed family $\mathscr{H}$ is a* closure root *of $\mathscr{F}$ if $\overline{\mathscr{H}} = \mathscr{F}$.*

Closure roots need not exist nor do they need be unique: for example, $(n-1)$-dense families do not have closure roots, whereas every 1-dense family is a closure root of $2^{[n]}$. In the tree $\mathscr{G}_n$ described in §7.2, the families not having any closure root determine the leaves of the tree. In §7.5.1, we investigate the existence of closure roots of 1-dense families.

Remark 7.5.10. *Note that if $\mathscr{H}$ is a closure root of $\mathscr{F}$ then by Corollary 7.2.4, we have that $\mathscr{A}_{n-1} \subseteq \overline{\mathscr{H}} = \mathscr{F}$. Therefore, any family $\mathscr{F}$ with $\mathscr{A}_{n-1,n} \nsubseteq \mathscr{F}$ does not have a closure root. In particular, this shows that every family considered in §7.3 does not have a closure root thus giving at least $\binom{n}{k-1} f_{k-1}$ many leaves at level k of the tree $\mathscr{G}_n$. In Corollary 7.5.14, we give example of an $1-$dense family with $\mathscr{A}_{n-1} \subseteq \mathscr{F}$, which does not have a closure root.*

Lemma 7.5.11. *Let $\mathscr{F}$ be a 1-dense family, with $\mathscr{G}$ its generating set (as defined in Definition 7.5.5). Let $A, B \in <\mathscr{G}>_{\mathscr{F}}$ with $A \subseteq B$. Then $A \subseteq_{\mathscr{F}} B$.*

Proof. If $A = B$ or $B = [n]$, then the conclusion is immediate, therefore we assume that $A \subsetneq B \subsetneq [n]$. Note that as $B \in <\mathscr{G}>_{\mathscr{F}}$, there exists $C \in \mathscr{G}$ such that $C \subsetneq_{\mathscr{F}} B$, which in turn means that there exists $C_1 \in \mathscr{F}$ such that $C \cup C_1 = B$ and $C_1 \subsetneq B$. Now, $A \subsetneq B$ implies that $A \subsetneq C \cup C_1$, and $C \in \mathscr{G}$ implies that $A \nsubseteq C$, which in turn yields $\emptyset \neq A \setminus C \subsetneq C_1 \setminus C$. Let $a \in A \setminus C$. We then have

$$A \cup (C \cup (C_1 \setminus \{a\})) = B$$

where $C \subseteq C \cup (C_1 \setminus \{a\})$ and $C \in \mathscr{G}$ together imply that $C \cup (C_1 \setminus \{a\}) \in \mathscr{F}$, and since $a \notin C$, hence $C \cup (C_1 - \{a\}) \neq B$. These observations together yield $A \subseteq_{\mathscr{F}} B$, as desired. □

Theorem 7.5.12. *Let $\mathscr{F}$ be a 1-dense family with generating set $\mathscr{G}$. Let $\mathscr{H} \subseteq \mathscr{F}$ be a union-closed family such that $\overline{\mathscr{H}} \supseteq \mathscr{F}$. Then $\overline{\mathscr{H}} \supseteq \overline{<\mathscr{G}>}_{\mathscr{F}}$.*

Proof. Let $\mathscr{K}$ be the set of all $\subseteq_{\mathscr{F}}$-wise minimal elements of $\mathscr{H}$. Since $\overline{\mathscr{H}} \supseteq \mathscr{F}$, Proposition 7.5.8 yields $<\mathscr{K}>_{\mathscr{F}} = \mathscr{H}$. Let $A \in \overline{<\mathscr{G}>}_{\mathscr{F}}$ and $B \in <\mathscr{K}>_{\mathscr{F}}$. We wish to show that $A \cup B \in <\mathscr{K}>_{\mathscr{F}} \cup \{A\}$. If $A \cup B = B$ or $A \cup B = A$, then the conclusion is immediate. Henceforth, assume $A \nsubseteq B$ and $B \nsubseteq A$. Therefore, $B \subsetneq A \cup B$. There exists $C \in \mathscr{K}$ such that $C \subseteq_{\mathscr{F}} B$. If $C \neq B$, then $C \subsetneq_{\mathscr{F}} B \subseteq A \cup B$. Hence, by Lemma 7.5.2, we get $C \subsetneq_{\mathscr{F}} A \cup B$, implying that $A \cup B \in <\mathscr{K}>_{\mathscr{F}}$. Now, suppose $C = B$. There exists $A_o \in \mathscr{G}$ such that $A_o \subseteq B$. Suppose $A_o \neq B$. We get $A_o \subsetneq B \subseteq A \cup B$, implying that $B \subseteq_{\mathscr{F}} A \cup B$, again by Lemma 7.5.2. Therefore, we are left with $A_o = B$. Since $A \in \overline{<\mathscr{G}>}_{\mathscr{F}}$ and $A \cup B \neq A$, we get $A \cup B \in <\mathscr{G}>_{\mathscr{F}}$ and $B \subsetneq A \cup B$ with $B \in \mathscr{G}$. By Lemma 7.5.2, $B \subseteq_{\mathscr{F}} A \cup B$. Hence, $A \cup B \in <\mathscr{K}>_{\mathscr{F}} \cup \{A\}$. □

Equipped with Theorem 7.5.12, we characterise the 1-dense families that have closure roots, via Theorem 7.5.13.

Theorem 7.5.13. *Let $\mathscr{F}$ be a 1-dense family over universe $[n]$. Let $\mathscr{G}$ be the generating set of $\mathscr{F}$, as defined in Definition 7.5.5. Then $\mathscr{F}$ has a closure root if and only if $\overline{<\mathscr{G}>}_{\mathscr{F}} = \mathscr{F}$.*

Proof. If $\overline{<\mathscr{G}>_{\mathscr{F}}} = \mathscr{F}$ then the conclusion is immediate. Suppose there exists a closure root $\mathscr{H}$ of $\mathscr{F}$. By Theorem 7.5.12, we have $\overline{\mathscr{H}} \supseteq \overline{<\mathscr{G}>_{\mathscr{F}}} \supseteq \mathscr{F}$. But $\overline{\mathscr{H}} = \mathscr{F}$ by Definition 7.5.9, hence $\overline{<\mathscr{G}>_{\mathscr{F}}} = \mathscr{F}$. $\square$

We mention here that the proofs of Theorems 7.5.12 and 7.5.13 can not be generalised to general families because, as remarked earlier, transitivity of $\subseteq_{\mathscr{F}}$ does not necessarily hold when $\mathscr{F}$ is a k-dense family for some $k > 1$.

We now provide the example of a $1-$dense family $\mathscr{F}$ with $\mathscr{A}_{n-1,n} \subseteq \mathscr{F}$ which does *not* have a closure root.

Corollary 7.5.14. *Let $k \geq 2$ and $\mathscr{F}$ be the 1-dense family ,over universe $[2k+1]$, generated by $\mathscr{G}$ $= \{\{1,2\},\{2,3\},\ldots,\{2k,2k+1\}\}$. Then $\mathscr{F}$ does not have a closure root.*

Proof. Consider $A = \{1,3,5,\ldots,2k+1\}$. It is immediate that $A \notin \mathscr{F}$. Let $B \supsetneq A$ and $a \in B \setminus A$. Since $1 < a < 2k+1$, we have $\{a-1,a,a+1\} \subseteq B$. Since $\{a-1,a\} \subsetneq_{\mathscr{F}} \{a-1,a,a+1\} \subseteq B$ therefore, by Lemma 7.5.2, we have $\{a-1,a\} \subsetneq_{\mathscr{F}} B$, and as $\{a-1,a\} \in \mathscr{G}$, hence $B \in <\mathscr{G}>_{\mathscr{F}}$. Therefore, in particular $\forall C \in <\mathscr{G}>_{\mathscr{F}}$ with $A \cup C \supsetneq A$, we have $A \cup C \in <\mathscr{G}>_{\mathscr{F}}$, which implies that $A \in \overline{<\mathscr{G}>_{\mathscr{F}}}$. Therefore, by Theorem 7.5.13, $\mathscr{F}$ does not have a closure root.

$\square$

Remark 7.5.15. *Let $\mathscr{F}$ be the 1-dense family over universe $[2k]$ generated by $\mathscr{G}$ = $\{\{1,2\},\{2,3\},\ldots,\{2k-1,2k\}\}$ for $k \geq 3$. Then $\mathscr{F}$ has a closure root.*

To see this, let us first note that whenever $A \in <\mathscr{G}>_{\mathscr{F}}$ with $|A| \geqslant 3$ then the number of sets $\{i,i+1\}$ such that $\{i,i+1\} \subseteq A$ is at least 2. If $A = [2k]$ then the result is obvious, otherwise, by the definition of $<\mathscr{G}>_{\mathscr{F}}$, $A \in <\mathscr{G}>_{\mathscr{F}}$ means that there is a $\{i,i+1\} \in \mathscr{G}$ such that $\{i,i+1\} \subsetneq_{\mathscr{F}} A$. This means that there is a $B \in \mathscr{F}$ such that $\{i,i+1\} \cup B = A$ with $B \neq A$. This gives that $\{i,i+1\} \nsubseteq B$ and hence, since $B \in \mathscr{F}$ there is a $j \neq i$ such that $\{j,j+1\} \subseteq B \subseteq A$ hence proving our claim.

Now, let us consider $A \in \overline{<\mathscr{G}>_{\mathscr{F}}} \setminus \mathscr{F}$. Therefore, $\forall i \in A, i+1, i-1 \notin A$. If $A = \{i\}$ with $i \geqslant 4$ then $\{1,2\} \cup A \notin <\mathscr{G}>_{\mathscr{F}} \cup \{A\}$, otherwise, if $i \leqslant 3$ then $\{5,6\} \cup A \notin <\mathscr{G}>_{\mathscr{F}} \cup \{A\}$. So, we can have that $|A| \geqslant 2$. Suppose $1 \notin A$ and let $\min.A = t$. Then by the discussion above, $\{t-1,t\} \cup A \notin <\mathscr{G}>_{\mathscr{F}} \cup \{A\}$. Therefore, $1 \in A$. Now, suppose that for some $r \geqslant 2$, $\{1,3,\ldots,2r-1\} \subseteq A$ and $2r, 2r+1 \notin A$. But then, $\{2r-1,2r\} \cup A \notin <\mathscr{G}>_{\mathscr{F}} \cup \{A\}$(by the discussion in the previous paragraph). Therefore, by induction, $\{1,3,\ldots,2k-1\} \subseteq A$. But $\{2k-1,2k\} \cup A \notin <\mathscr{G}>_{\mathscr{F}}$ and hence, $A \notin \overline{<\mathscr{G}>_{\mathscr{F}}}$ which gives us that $\overline{<\mathscr{G}>_{\mathscr{F}}} = \mathscr{F}$. Thus, completing the argument.

We would like to conclude by noting that the family $<\mathscr{G}>_{\mathscr{F}}$ as described above need not always be $2-$dense. In fact, it can be of arbitrary density. As an example if we consider $\mathscr{F}$ to be the $1-$dense family over universe $[n]$ generated by $\mathscr{G}=\{[k]\}$ where $n\geqslant k+2$. Then it is easy to see that the family $<\mathscr{G}>_{\mathscr{F}}=\{[k],[n]\}$ and hence by Corollary 7.3.3, $<\mathscr{G}>_{\mathscr{F}}$ is $(k+1)-$dense. And finally if we let $\mathscr{G}=\mathscr{A}_{n-1,n}$ then $<\mathscr{G}>_{\mathscr{F}}=\mathscr{F}$ making $<\mathscr{G}>_{\mathscr{F}}$ to be $1-$dense.

7.6 Further Questions

Here, we state some further questions, which we believe are worth exploring:

1. Do 2-dense families satisfy Conjecture 1?
2. What is a characterization of 2-dense families?
3. Do $(n-1)$-dense families satisfy Conjecture 1?
4. What is a characterization of $(n-1)$-dense families?
5. Since it is known that union closed families $\mathscr{F}$ over universe $[n]$ satisfying $|\mathscr{F}|\geqslant 2^{n-1}$ satisfy Conjecture 1(see [64]), it can be an interesting problem to characterise families satisfying $|\overline{\mathscr{F}}|\geqslant 2^{n-1}$.
6. Suppose $\mathscr{F}$ satisfies Conjecture 1. Does $\overline{\mathscr{F}}$ also satisfy Conjecture 7.2.1? (An affirmative answer to this problem would mean that if $\mathscr{F}$ is an counterexample to Conjecture 1 then all its closure root are also counterexamples.)
7. As in [10], let $d_{\mathscr{F}}(x)=|\{A\in\mathscr{F}|x\in A\}$. Let $g(\mathscr{F})=\frac{1}{|\mathscr{F}|}\max\{d_{\mathscr{F}}(x)|x\in\mathscr{F}\}$. Let us put $a_{k,n}=\min\{g(\mathscr{F})|\mathscr{F}$ is k-dense and over universe $[n]\}$. It can be an interesting problem to study the numbers $a_{k,n}$. If it turns out that $a_{k-1,n}\leqslant a_{k,n}$ for all possible values of k and n then it would prove Conjecture 1 because $a_{0,n}\geqslant\frac{1}{2}$. We would like to note here that if Conjecture 1 is true then $a_{n-1,n}\to\frac{1}{2}$ as $n\to\infty$. This follows by considering the family $\mathscr{F}=2^{[n-2]}\cup\{[n]\}$ and noting that it is $(n-1)-$dense using Corollary 7.3.3. One natural way of trying to prove $a_{k-1,n}\leqslant a_{k,n}$ is to try and prove that for every family $\mathscr{F}$, $g(\mathscr{F})\geqslant g(\overline{\mathscr{F}})$. But this is not true in general as is seen by again considering the family $\mathscr{F}=2^{[n-2]}\cup\{[n]\}$ and applying Theorem 7.3.2.

Bibliography

[1] Polymath 11. Gowers's blog. https://gowers.wordpress.com/2016/01/21/. Accessed: 2025-04-21.

[2] David J Aldous and Antar Bandyopadhyay. A survey of max-type recursive distributional equations. *Annals of Applied Probability*, 15(2):1047–1110, 2005.

[3] Ben Amiet, Andrea Collevecchio, Marco Scarsini, and Ziwen Zhong. Pure nash equilibria and best-response dynamics in random games. *Mathematics of Operations Research*, 46(4):1552–1572, 2021.

[4] William N Anderson Jr. Maximum matching and the game of slither. *Journal of Combinatorial Theory, Series B*, 17(3):234–239, 1974.

[5] Krishna B Athreya and Peter Jagers. *Classical and modern branching processes*, volume 84. Springer Science & Business Media, 2012.

[6] Krishna B Athreya and Peter E Ney. *Branching processes*. Springer, Berlin, Heidelberg, 1972.

[7] Krishna B Athreya and AN Vidyashankar. Branching processes. *Stochastic processes: theory and methods*, 19:35–53, 2001.

[8] Venkatesh Bala and Sanjeev Goyal. Learning from neighbours. *The review of economic studies*, 65(3):595–621, 1998.

[9] Venkatesh Bala and Sanjeev Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229, 2000.

[10] Igor Balla, Béla Bollobás, and Tom Eccles. Union-closed families of sets. *Journal of Combinatorial Theory, Series A*, 120(3):531–544, 2013.

[11] Torben C Barth, Janus Timler Holm, Jakob Lindorff Larsen, and David Pisinger. Optimization of transfer baggage handling in a major transit airport. In *Operations Research Forum*, volume 2, pages 1–35. Springer, 2021.

[12] Riddhipratim Basu, Alexander E Holroyd, James B Martin, Johan Wästlund, et al. Trapping games on random boards. *The Annals of Applied Probability*, 26(6):3727–3753, 2016.

[13] Janet Simpson Beissinger. Similar constructions for young tableaux and involutions, and their application to shiftable tableaux. *Discrete mathematics*, 67(2):149–163, 1987.

[14] Andrew Beveridge, Andrzej Dudek, Alan Frieze, Tobias Müller, and Miloš Stojaković. Maker-breaker games on random geometric graphs. *Random structures & algorithms*, 45(4):553–607, 2014.

[15] Dhruv Bhasin, Sayar Karmakar, Moumanti Podder, and Souvik Roy. Ergodicity of a generalized probabilistic cellular automaton with parity-based neighbourhoods. *arXiv preprint arXiv:2212.01753*, 2022.

[16] Dhruv Bhasin, Sayar Karmakar, Moumanti Podder, and Souvik Roy. On a class of pca with size-3 neighborhood and their applications in percolation games. *Electronic Journal of Probability*, 28:1–60, 2023.

[17] Irénée-Jules Bienaymé. De la loi de multiplication et de la durée des familles. *Soc. Philomat. Paris Extraits, Sér*, 5(37-39):4, 1845.

[18] Tom Bohman, Alan Frieze, Tomasz Łuczak, Oleg Pikhurko, Clifford Smyth, Joel Spencer, and Oleg Verbitsky. First-order definability of trees and sparse random graphs. *Combinatorics, Probability and Computing*, 16(3):375–400, 2007.

[19] Béla Bollabás and Alan Stacey. Approximate upper bounds for the critical probability of oriented percolation in two dimensions based on rapidly mixing markov chains. *Journal of Applied Probability*, 34(4):859–867, 1997.

[20] Yvan Le Borgne and Jean-François Marckert. Directed animals and gas models revisited. *The Electronic Journal of Combinatorics*, 14(1):R71, 2007.

[21] Olivier Bouré, Nazim Fates, and Vincent Chevrier. Probing robustness of cellular automata through variations of asynchronous updating. *Natural Computing*, 11(4):553–564, 2012.

[22] Mireille Bousquet-Mélou. New enumerative results on two-dimensional directed animals. *Discrete Mathematics*, 180(1-3):73–106, 1998.

[23] Guy Bresler, Chenghao Guo, and Yury Polyanskiy. Linear programs with polynomial coefficients and applications to 1d cellular automata. *arXiv preprint arXiv:2204.06357*, 2022.

[24] Gunnar Brinkmann and Robin Deklerck. Generation of union closed sets and moore families. *arXiv preprint arXiv:1701.03751*, 2017.

[25] Simon R Broadbent and John M Hammersley. Percolation processes: I. crystals and mazes. In *Mathematical proceedings of the Cambridge philosophical society*, volume 53, pages 629–641. Cambridge University Press, 1957.

[26] Henning Bruhn, Pierre Charbit, Oliver Schaudt, and Jan Arne Telle. The graph formulation of the union-closed sets conjecture. *European Journal of Combinatorics*, 43:210–219, 2015.

[27] Henning Bruhn and Oliver Schaudt. The journey of the union-closed sets conjecture. *Graphs and Combinatorics*, 31:2043–2074, 2015.

[28] Ana Bušić, Jean Mairesse, and Irene Marcovici. Probabilistic cellular automata, invariant measures, and perfect sampling. *Advances in Applied Probability*, 45(4):960–980, 2013.

[29] Jérôme Casse and Irène Marcovici. Probabilistic cellular automata with memory two: invariant laws and multidirectional reversibility. *Annales Henri Lebesgue*, 3:501–559, 2020.

[30] Kalyan Chatterjee and Susan H Xu. Technology diffusion by learning from neighbours. *Advances in Applied Probability*, 36(2):355–376, 2004.

[31] S Chowla, IN Herstein, and WK Moore. On recursions connected with symmetric groups i. *Canadian Journal of Mathematics*, 3:328–334, 1951.

[32] Vašek Chvátal and Paul Erdös. Biased positional games. In *Annals of Discrete Mathematics*, volume 2, pages 221–229. Elsevier, 1978.

[33] Brian A Davey and Hilary A Priestley. *Introduction to lattices and order*. Cambridge university press, 2002.

[34] A Nicholas Day and Victor Falgas-Ravry. Maker–breaker percolation games i: crossing grids. *Combinatorics, Probability and Computing*, 30(2):200–227, 2021.

[35] A Nicholas Day and Victor Falgas-Ravry. Maker-breaker percolation games ii: Escaping to infinity. *Journal of Combinatorial Theory, Series B*, 151:482–508, 2021.

[36] Deepak Dhar. Exact solution of a directed-site animals-enumeration problem in three dimensions. *Physical review letters*, 51(10):853, 1983.

[37] RL Dobrušin, VI Kriukov, and AL Toom. *Stochastic cellular systems: ergodicity, memory, morphogenesis*. Manchester University Press, 1990.

[38] D. Duffus. In I. Rival, editor, *Graphs and Order*, page 525. Reidel, Dordrecht, Boston, 1985.

[39] Richard Durrett. Oriented percolation in two dimensions. *The Annals of Probability*, 12(4):999–1040, 1984.

[40] Richard Durrett and David Griffeath. Supercritical contact processes on z. *The Annals of Probability*, pages 1–15, 1983.

[41] Vojtěch Dvořák, Adva Mond, and Victor Souza. The maker-breaker percolation game on the square lattice. *arXiv preprint arXiv:2105.12864*, 2021.

[42] Vojtěch Dvořák, Adva Mond, and Victor Souza. The maker-breaker percolation game on a random board. *arXiv preprint arXiv:2402.17547*, 2024.

[43] Glenn Ellison and Drew Fudenberg. Rules of thumb for social learning. *Journal of political Economy*, 101(4):612–643, 1993.

[44] Wesam M Elsayed, Mohammed Elmogy, and B El-Desouky. Evolutionary behavior of dna sequences analysis using non-uniform probabilistic cellular automata model. *Cincia e Tcnica Vitivincola*, 32:137–148, 2017.

[45] Asaf Ferber, Roman Glebov, Michael Krivelevich, and Alon Naor. Biased games on random boards. *Random Structures & Algorithms*, 46(4):651–676, 2015.

[46] János Flesch, Arkadi Predtetchinski, and Ville Suomala. Random perfect information games. *Mathematics of Operations Research*, 48(2):708–727, 2023.

[47] Aviezri Fraenkel. Combinatorial games: selected bibliography with a succinct gourmet introduction. *The Electronic Journal of Combinatorics*, pages DS2–Aug, 2012.

[48] Aviezri S Fraenkel. Complexity, appeal and challenges of combinatorial games. *Theoretical Computer Science*, 313(3):393–415, 2004.

[49] Markus Frey. *Models and methods for optimizing baggage handling at airports*. PhD thesis, Technische Universität München, 2014.

[50] Peter Gács. Reliable cellular automata with self-organization. *Journal of Statistical Physics*, 103(1):45–267, 2001.

[51] Martin Gardner. The fantastic combinations of John Conway's new solitaire game life. *Sc. Am.*, 223:20–123, 1970.

[52] Guillaume Garnier and Bruno Ziliotto. Percolation games. *Mathematics of Operations Research*, 2022.

[53] Eric Goles and Martín Matamala. Uniform simulation of turing machines by cellular automata. In *Cellular Automata and Complex Systems*, pages 23–36. Springer, 1999.

[54] David Griffeath. The basic contact processes. *Stochastic Processes and their Applications*, 11(2):151–185, 1981.

[55] G Grinstein, C Jayaprakash, and Yu He. Statistical mechanics of probabilistic cellular automata. *Physical Review Letters*, 55(23):2527, 1985.

[56] I Hartarsky and R Szabó. Generalised oriented site percolation, probabilistic cellular automata and bootstrap percolation. *arXiv preprint arXiv*, 2021.

[57] Ivailo Hartarsky. Bootstrap percolation, probabilistic cellular automata and sharpness. *Journal of Statistical Physics*, 187(3):1–17, 2022.

[58] Dan Hefetz, Michael Krivelevich, Miloš Stojaković, and Tibor Szabó. *Positional games*. Springer, 2014.

[59] Alexander E Holroyd, Avi Levy, Moumanti Podder, and Joel Spencer. Second order logic on random rooted trees. *Discrete Mathematics*, 342(1):152–167, 2017.

[60] Alexander E Holroyd, Irène Marcovici, and James B Martin. Percolation games, probabilistic cellular automata, and the hard-core model. *Probability Theory and Related Fields*, 174(3):1187–1217, 2019.

[61] Alexander E Holroyd and James B Martin. Galton-watson games. *Random Structures & Algorithms*, 59(4):495–521, 2021.

[62] Tobias Johnson, Moumanti Podder, and Fiona Skerman. Random tree recursions: Which fixed points correspond to tangible sets of trees? *Random Structures & Algorithms*, 56(3):796–837, 2020.

[63] Sayar Karmakar, Moumanti Podder, Souvik Roy, and Soumyarup Sadhukhan. Phase transition in percolation games on rooted galton-watson trees. *arXiv preprint arXiv:2303.11402*, 2023.

[64] Ilan Karpas. Two results on union-closed families. *arXiv preprint arXiv:1708.01434*, 2017.

[65] Jeong Han Kim, Oleg Pikhurko, Joel H Spencer, and Oleg Verbitsky. How complex are random graphs in first order logic? *Random Structures & Algorithms*, 26(1-2):119–145, 2005.

[66] Donald E. Knuth. *The art of computer programming. Volume 3*. Addison-Wesley Series in Computer Science and Information Processing. 1973. Sorting and searching.

[67] Andrey Kupavskii and Maksim Zhukovskii. Short monadic second order sentences about sparse random graphs. *SIAM Journal on Discrete Mathematics*, 32(4):2916–2940, 2018.

[68] Joel L Lebowitz, Christian Maes, and Eugene R Speer. Statistical mechanics of probabilistic cellular automata. *Journal of statistical physics*, 59(1):117–170, 1990.

[69] Qin Lei, Jia Lee, Xin Huang, and Shuji Kawasaki. Entropy-based classification of elementary cellular automata under asynchronous updating: an experimental study. *Entropy*, 23(2):209, 2021.

[70] Pierre-Yves Louis and Francesca R Nardi. Probabilistic cellular automata. *Emergence, Complexity, Computation*, 27, 2018.

[71] Tibor Mach, Anja Sturm, and Jan M Swart. A new characterization of endogeny. *Mathematical Physics, Analysis and Geometry*, 21(4):30, 2018.

[72] Tibor Mach, Anja Sturm, and Jan M Swart. Recursive tree processes and the mean-field limit of stochastic flows. *Electronic Journal of Probability*, 25:1–63, 2020.

[73] Jean Mairesse and Irene Marcovici. Around probabilistic cellular automata. *Theoretical Computer Science*, 559:42–72, 2014.

[74] Irène Marcovici, Mathieu Sablik, and Siamak Taati. Ergodicity of some classes of cellular automata subject to noise. *Electronic Journal of Probability*, 24:1–44, 2019.

[75] James B Martin and Roman Stasiński. Minimax functions on galton–watson trees. *Combinatorics, Probability and Computing*, 29(3):455–484, 2020.

[76] AD Matushkin and ME Zhukovskii. First order sentences about random graphs: small number of alternations. *Discrete Applied Mathematics*, 236:329–346, 2018.

[77] Mark McCann and Nicholas Pippenger. Fault tolerance in cellular automata at high fault rates. *Journal of Computer and System Sciences*, 74(5):910–918, 2008.

[78] Mark McCann and Nicholas Pippenger. Fault tolerance in cellular automata at low fault rates. *Journal of Computer and System Sciences*, 79(7):1126–1143, 2013.

[79] Leo Moser. On solutions of xd= 1 in symmetric groups. *Canadian Journal of Mathematics*, 7:159–168, 1955.

[80] LB Ostrovsky and ME Zhukovskii. Monadic second-order properties of very sparse random graphs. *Annals of pure and applied logic*, 168(11):2087–2101, 2017.

[81] Oleg Pikhurko, Helmut Veith, and Oleg Verbitsky. The first order definability of graphs: upper bounds for quantifier depth. *Discrete applied mathematics*, 154(17):2511–2529, 2006.

[82] Moumanti Podder. The first order theory of $G(n, c/n)$. *European Journal of Combinatorics*, 78:214–235, 2019.

[83] Moumanti Podder and Joel Spencer. First order probabilities for galton–watson trees. In *A Journey Through Discrete Mathematics*, pages 711–734. Springer, 2017.

[84] Moumanti Podder and Joel Spencer. Galton-watson probability contraction. *Electronic Communications in Probability*, 22:Paper no. 20, 2017.

[85] Bjorn Poonen. Union-closed families. *Journal of Combinatorial Theory, Series A*, 59(2):253–268, 1992.

[86] E Popova, Y Staraselski, A Brahme, RK Mishra, and K Inal. Coupled crystal plasticity–probabilistic cellular automata approach to model dynamic recrystallization in magnesium alloys. *International Journal of Plasticity*, 66:85–102, 2015.

[87] James Gary Propp and David Bruce Wilson. Exact sampling with coupled markov chains and applications to statistical mechanics. *Random Structures & Algorithms*, 9(1-2):223–252, 1996.

[88] Balázs Ráth, Jan M Swart, and Márton Szőke. A phase transition between endogeny and nonendogeny. *Electronic Journal of Probability*, 27:1–43, 2022.

[89] Balázs Ráth, Jan M Swart, and Tamás Terpai. Frozen percolation on the binary tree is nonendogenous. *The Annals of Probability*, 49(5):2272–2316, 2021.

[90] AS Razafimahatratra and M Zhukovskii. Zero–one laws for k-variable first-order logic of sparse random graphs. *Discrete Applied Mathematics*, 276:121–128, 2020.

[91] Ian Roberts and Jamie Simpson. A note on the union-closed sets conjecture. *Australasian Journal of Combinatorics*, 47:265–267, 2010.

[92] Pal Rujan. Cellular automata and statistical mechanical models. *Journal of Statistical Physics*, 49(1):139–222, 1987.

[93] Carl P Simon, Lawrence Blume, et al. *Mathematics for economists*, volume 7. Norton New York, 1994.

[94] Joel Spencer. Threshold spectra via the ehrenfeucht game. *Discrete Applied Mathematics*, 30(2-3):235–252, 1991.

[95] Joel Spencer and Lubos Thoma. On the limit values of probabilities for the first order properties of graphs. *Contemporary trends in discrete mathematics*, 49:317–336, 1997.

[96] Joel H Spencer and Katherine St John. Random unary predicates: Almost sure theories and countable models. *Random Structures & Algorithms*, 13(3-4):229–248, 1998.

[97] James Stewart. *Essential calculus: Early transcendentals*. Cengage Learning, 2012.

[98] Miloš Stojaković. Games on graphs. In *International Conference on Conceptual Structures*, pages 31–36. Springer, 2014.

[99] Miloš Stojaković and Tibor Szabó. Positional games on random graphs. *Random Structures & Algorithms*, 26(1-2):204–223, 2005.

[100] Miloš Stojaković and Nikola Trkulja. Hamiltonian maker–breaker games on small graphs. *Experimental Mathematics*, 30(1):595–604, 2021.

[101] Robert F Tichy and Stephan Wagner. Extremal problems for topological indices in combinatorial chemistry. *Journal of computational biology*, 12(7):1004–1013, 2005.

[102] Oleg Verbitsky. The first order definability of graphs with separators via the ehrenfeucht game. *Theoretical computer science*, 343(1-2):158–176, 2005.

[103] Freddie Wallwork. Maker-breaker-crossing-game on the triangular grid-graph. *arXiv preprint arXiv:2201.01348*, 2022.

[104] Johan Wästlund. Replica symmetry of the minimum matching. *Annals of Mathematics*, pages 1061–1091, 2012.

[105] Henry William Watson and Francis Galton. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4:138–144, 1875.

[106] ME Zhukovskii. On infinite spectra of first order properties of random graphs. *Moscow Journal of Combinatorics and Number Theory*, 6(4):73–102, 2016.

[107] ME Zhukovskii. Logical laws for short existential monadic second-order sentences about graphs. *Journal of Mathematical Logic*, 20(02):2050007, 2020.