A study of Component-wise semi-Markov Process

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

Doctor of Philosophy

by

Yadav Ravishankar Kapildev

ID: 20173545

IISER PUNE

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

April 18, 2023

Certificate

Certified that the work incorporated in the thesis entitled "*A study of component-wise semi-Markov process*", submitted by *Yadav Ravishankar Kapildev* was carried out by the candidate, under my supervision in a collaboration with Dr. Subhamay Saha. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

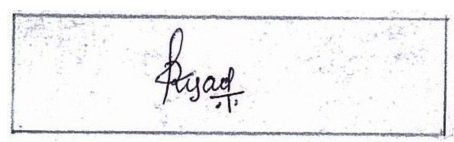
Anindya Goswami

Date: November 30, 2022 *Dr. Anindya Goswami*

Thesis Supervisor

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. 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.



Date: November 30, 2022 *Yadav Ravishankar Kapildev*

Roll Number: 20173545

This thesis is dedicated to Dr. Anindya Goswami, Dr. Shubhamay Shah, my parents as well as to all of my teachers.

Acknowledgements

I would like to express my deepest appreciation to my advisor, Dr. Anindya Goswami. He has been an incredible guide and source of inspiration. I feel privileged to be able to call myself one of his students. I appreciate the everyday advice he gave me, in addition to the professional guidance he provided. I am grateful for his encouraging words and the guidance I received during my unproductive times. Furthermore, I appreciate all of his encouragement and understanding throughout my PhD. Because he has always been there for me, even when things were tough. Furthermore, I would like to extend my sincere gratitude to Dr. Shubhamay Saha for the collaboration and all the fruitful discussions and suggestions. This endeavour would not have been possible without him.

I would also like to express my deepest gratitude to my RAC members, Prof. Gopal K. Basak and Dr. Anup Biswas, for their support, guidance, and fruitful discussions.

I am grateful to all the faculty members of the mathematics department at IISER Pune. I am thankful to Mrs Suvarna and Mr Yogesh, the administrative staff of the Mathematics Department, and the administrative staff of IISER Pune. I also appreciate the support: I received from the Department of Mathematics throughout my PhD years. I want to acknowledge CSIR and IISER Pune for their financial support.

In addition, I appreciate the encouragement, support, and guidance I received from Jancey Mathew and Santhosh B. throughout my schooling. I also appreciate the guidance and support from Mr. Vishwas Khare,Mr. Vishal Khare, Mrs. Varsha Rajput, Mrs. Vanita Prabhu, Mrs. Sherin Jose, Dr. Bhikha Lila Ghodadra, Dr. Sejal Shah, Dr. Rajendra G. Vyas, Dr. Bhadreshkumar Dave, and Dr. Ekta Shah during graduation and post-graduation.

These last five and a half years would only have been as educational and entertaining with an enlightening company of friends. I am indebted to my friends Tumpa Mahato, Yashwant Kumar, Rashmi Sharma, Prashant Yadav, Deepak Sharma, Ateek Shah, Umashankar

Rajput, Pranjal Vishwakarma, Mayuresh Konde, Prachi Joshi, Parents of Prachi Joshi and Mayuresh Konde, Visakh Narayanan, Namrata Aravind, Shubham namdeo, Rajeshwari, Dhruv Bhasin, Pavan Dighe, Sujeet Dhamore, Abhishek Anand, Keerti Tomar, Kundan Kumar, Girish Kulkarni, Rijubrata Kundu, Abdul Gaffar, Radhika Singh, Naveen Yadav, Avinash Shinde, Gaurishankar, Ajay Raj, Vijay Gupta, Karamveer, Viral, and all my seniors and folks of the IISER Pune.

As a last note, I'd want to thank my family and friends for the unwavering love, support, and encouragement they've given me throughout my life. In particular, my parents have provided the reliable backbone I've needed to overcome the many difficulties I've faced. I value your support and hope to earn your pride someday. Finally, I'd like to thank my brother Rohit, who held down the fort while I was away.

Contents

Abstract

Component-wise semi-Markov processes (CSM) constitute a larger class of pure jump processes which includes semi-Markov, and Markov pure jump processes. This thesis examines semi-Markov as well as CSM processes with dependent components. In order to better understand the interactions among components of CSM processes having bounded transition rates, we consider a family of stochastic flows using a system of SDEs driven by Poisson random measure (PRM), with an additional gaping parameter. More specifically, we have demonstrated that the proposed system of SDEs driven by a PRM does, in fact, has a unique solution. Then, we prove that a solution satisfies the desired law. Thus we establish a semimartingale representation of the homogeneous or nonhomogeneous semi-Markv process. Finally, we pick up an appropriate flow by fixing the gaping parameter. We derive expressions of the probabilities of meeting and merging of a pair of semi-Markov processes, solving the same equation but with different initial conditions. We also obtain a set of sufficient conditions for any two solutions merge eventually with probability one. The theoretical results are elaborated with the help of numerous numerical examples. An SMP with a specific law is what makes up each component of a vector-valued CSM. These parts might be governed by the same set of rules, or they might not. The current investigation of CSM focuses heavily on the junctures at which the constituent parts come together and form a whole. When the parts are unrelated to one another, questions about such occurrences can be answered right away. The questions become interesting, however, when the parts are driven by dependent or identical noises. We were able to derive the infinitesimal generator for CSMs with arbitrarily number of components driven by a single PRM. Additionally, we have defined correlated PRM and provided a semimartingale representation of a CSM driven by correlated PRM.

Notations

Let a and b be real numbers.

1. $a^+ = \max(0, a)$
2. $a \wedge b = \min(a, b)$
3. $a \vee b = \max(a, b)$

$\mathcal{B}(\mathbb{R}^d)$: The Borel sigma algebra on $\mathbb{R}^d$.
m_d : The Lebesgue measure on $\mathbb{R}^d$.
$\mathbb{N}_0$: The set of non-negative integers.
$\mathbb{R}_+$: The set of non-negative real numbers.
$\bar{\mathbb{R}}_+ := [0, \infty]$.
$\mathbb{D} := \{1, 2, \ldots, d\}$.
$\mathcal{P}(A)$: The power set of A.
$\mathcal{C}^k : \{f : \mathbb{R}^d \to \mathbb{R}^d \mid f \text{ is } k \text{ times differentiable}\}$.
$\mathcal{C}^\infty : \{f : \mathbb{R}^d \to \mathbb{R}^d \mid k \in \mathbb{N}_0 f \in \mathcal{C}^k\}$.

Preliminary

We recall a few standard definitions and results in this chapter along with some additional remarks. Readers may consult [8, 33, 36] for more details.

0.1 Measure theory concepts

Definition 0.1.1. *Let Ω be a nonempty set, and let $\mathcal{F}$ be a σ-algebra on Ω. Then the pair $(\Omega, \mathcal{F})$ is called a measurable space. If P is a measure on $(\Omega, \mathcal{F})$, then the triplet $(\Omega, \mathcal{F}, P)$ is called a **measure space**. If in addition, P is a probability measure i.e. $P(\Omega) = 1$, then $(\Omega, \mathbb{F}, \mathbb{P})$ is called a **probability space**. A measure space is **complete** if for any $A \in \mathcal{F}$ with $P(A) = 0 \implies \mathcal{P}(\mathcal{A}) \subset \mathcal{F}$.*

Definition 0.1.2. *a. Let $(\Omega_i, \mathcal{F}_i)$ be a measurable space for each $i = 1, 2$. Then a function $f : \Omega_1 \to \Omega_2$ is measurable, if*

$$f^{-1}(A) \in \mathcal{F}_1, \forall A \in \mathcal{F}_2.$$

*b. Let S be a polish space and $(\Omega, \mathcal{F}, P)$ be a probability space. A map $X : \Omega \to S$ is called a **random variable**, if,*

$$X^{-1}(A) \in \mathcal{F}, \forall A \in \mathcal{B}(S).$$

Theorem 0.1.3 (Monotone convergence Theorem)**.** *Let $(\Omega, \mathcal{F}, P)$ be a measure space and let $f_n : \Omega \to \mathbb{R}_+, n \geq 1$ be the sequence of non-negative $\mathcal{F}$-measurable functions such that $f_n \to f$ pointwise a.e. and $f_1 \leq f_2 \leq \ldots$, then*

$$\lim_{n\to\infty} \int f_n = \int f.$$

Theorem 0.1.4 (Dominated Convergence Theorem)**.** *Let $f_n : \Omega \to \mathbb{R}, n \geq 1$ is a sequence of measurable functions on a measure space $(\Omega, \mathcal{F}, P)$ such that $f_n \to f$ point wise a.e., and if there exists an integrable function g such that $|f_n(x)| \leq g(x)$ a.e.(P) for all n, then f is integrable and*

$$\lim_{n\to\infty} \int f_n = \int f.$$

0.2 Probability theory concepts

Definition 0.2.1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A family $\{\mathcal{F}_t\}_{0\leq t<\infty}$ sub-σ-algebras of $\mathcal{F}$ is called filtration if for $s < t, \mathcal{F}_s \subseteq \mathcal{F}_t$. For convenience, we will usually write $\mathbb{F}$ for the filtration $\{\mathcal{F}_t\}_{0\leq t<\infty}$. Also, the quadruplet $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is called a filtered probability space.*

Definition 0.2.2. *Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be the filtered probability space. A random variable $T \colon \Omega \to \bar{\mathbb{R}}_+$ is a **stopping time** if the event $\{T \leq t\} \in \mathcal{F}_t$, every $0 \leq t < \infty$.*

Definition 0.2.3 (Usual Hypothesis)**.** *A filtered complete probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is said to satisfy the **usual hypotheses** if*

(i) $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$;

(ii) $\mathcal{F}_t = \cap_{u>t}\mathcal{F}_u$, all t, $0 \leq t < \infty$; that is, the filtration $\mathbb{F}$ is right continuous.

Definition 0.2.4 (Stochastic Process)**.** *Let $\mathbb{T} \subset \mathbb{R}_+, S$ be a polish space . A **stochastic process** X on $(\Omega, \mathcal{F}, P)$ is a collection of S valued random variables $\{X_t\}_{t\in\mathbb{T}}$. The process X is said to be **adapted** to $\mathbb{F}$, a filtration of $\mathcal{F}$ if $X_t \in \mathcal{F}_t$ (that is, X is $\mathcal{F}_t$ measurable) for each $t \in \mathbb{T}$.*

Theorem 0.2.5. *Let S, T be stopping times. Then $S \wedge T = \min(S, T)$, $S \vee T = \max(S, T)$, $S + T$, αS, where $\alpha > 1$ are stopping times.*

Definition 0.2.6 (Expectation)**.** *Let X be a real valued random variable on a probability space $(\Omega, \mathcal{F}, P)$. Then the integral of X with respect to measure P is said to be **expectation** of X, denoted by EX,and defined by $EX := \int_\Omega X(\omega)P(d\omega)$.*

Definition 0.2.7. *Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. The **conditional expectation** of a non-negative random variable X with respect to $\mathcal{G}$ is a non-negative random variable, denoted by $E[X \mid \mathcal{G}]$ such that*

1. *$E[X \mid \mathcal{G}]$ is $\mathcal{G}$-measurable.*

2. *for every $A \in \mathcal{G}$,*

$$\int_A X dP = \int_A E[X \mid \mathcal{G}] dP \text{ a.s.}$$

The conditional expectation of any random variable X with respect to $\mathcal{G}$, if EX exists, is given by $E[X \mid \mathcal{G}] := E[X^+ \mid \mathcal{G}] - E[X^- \mid \mathcal{G}]$.

Properties of conditional expectation of real valued:

Proposition 0.2.8. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let Y be a measurable real valued random variable with $E|Y| < \infty$. Let $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$ be two sub σ-algebras contained in $\mathcal{F}$. Then*

1. *if X is $\mathcal{G}_1$-measurable then $E[X \mid \mathcal{G}_1] = X$.*
2. $E[Y \mid \mathcal{G}_1] = E[E[Y \mid \mathcal{G}_2] \mid \mathcal{G}_1]$.
3. $E[Y \mid \mathcal{G}_1] = E[E[Y \mid \mathcal{G}_1] \mid \mathcal{G}_2]$.
4. *For any bounded $\mathcal{G}_1$-measurable random variable U, $E[YU \mid \mathcal{G}_1] = UE[Y \mid \mathcal{G}_1]$.*

Definition 0.2.9. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\mathcal{G} \subset \mathcal{F}$ be a sub σ-algebra. For $B \in \mathcal{F}$, the* ***conditional probability*** *of B given $\mathcal{G}$, denoted by $P(B \mid \mathcal{G})$, is defined as*

$$P(B \mid \mathcal{G}) = E(\mathbb{1}_B \mid \mathcal{G}).$$

Thus $Z \equiv P(B \mid \mathcal{G})$ is a measurable function such that

$$P(A \cap B) = E(Z\mathbb{1}_A) \text{ for all } A \in \mathcal{G},$$

where $\mathbb{1}_A$ is the indicator function of the set A.

Note that, for a fixed $\mathcal{G}$, $B \mapsto P(B \mid \mathcal{G})(\omega)$ need not produce a probability measure on $(\Omega, \mathcal{F})$, P a.e. ω.

Definition 0.2.10. *Let X be a random variable on a probability space $(\Omega, \mathcal{F}, P)$ taking values in a complete, separable metric space $(S, \mathcal{B}(S))$. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. A* ***regular conditional distribution*** *of X given $\mathcal{G}$ is a function $Q : \Omega \times \mathcal{B}(S) \to [0, 1]$ such that*

1. *for each $\omega \in \Omega, Q(\omega, \cdot)$ is a probability measure on $(S, \mathcal{B}(S))$,*
2. *for each $B \in \mathcal{B}(S)$, the mapping $\omega \to Q(\omega, B)$ is $\mathcal{G}$-measurable, and*
3. *for each $B \in \mathcal{B}(S)$, $P[X \in B \mid \mathcal{G}](\omega) = Q(\omega, B)$, P a.e. ω.*

Definition 0.2.11. *Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be measurable spaces. Let Q be a mapping from $\Omega_1 \times \mathcal{F}_2$ into $\bar{\mathbb{R}}_+$. Then, Q is called a* ***transition kernel*** *from $(\Omega_1, \mathcal{F}_1)$ into $(\Omega_2, \mathcal{F}_2)$ if*

1. *the mapping $x \mapsto Q(x, B)$ is $\mathcal{F}_1$-measurable for every set B in $\mathcal{F}_2$, and*

2. *the mapping $B \mapsto Q(x, B)$ is measure on $(\Omega_2, \mathcal{F}_2)$ for every x in Ω_1.*

Remark 0.2.12. *A transition Kernel from $(\Omega, \mathcal{F})$ into $(\Omega, \mathcal{F})$ is called simply a transition kernel on $(\Omega, \mathcal{F})$. Such a kernel is called* ***Markov kernel*** *on $(\Omega, \mathcal{F})$ if $Q(x, \Omega) = 1$ for every x.*

Definition 0.2.13. *A general continuous-time* ***Markov process*** *is a process X on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and taking values in a polish space S, satisfying*

$$P[X_t \in A \mid \mathcal{F}_s] = P[X_t \in A \mid X_s] \tag{0.2.1}$$

for all $A \in \mathcal{B}(S)$ and for each $s < t$.

Definition 0.2.14. *Let X be a stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and taking values in a Polish space S. Also, f be a Borel measurable function on $(S, \mathcal{B}(S))$. Then X is said to be a* ***strong Markov process****, if for every stopping time τ which is adapted to $\mathcal{F}$, and $t > 0$ satisfies*

$$E[f(X_{\tau+t}) \mid \mathcal{F}_\tau] = E[f(X_{\tau+t}) \mid X_\tau].$$

Definition 0.2.15. *$X = \{X_t\}_{t \geq 0}$ defined on a complete probability space $(\Omega, \mathcal{F}, P)$ is an* ***semi-Markov process(SMP)*** *with state space $\mathcal{X} := \{1, 2, \dots\} \subset \mathbb{R}$ if*

1. *X is piece-wise constant r.c.l.l. process with discontinuities at $T_1 < T_2 < \cdots$, and*

2. *for each $n \geq 1, j \in \mathcal{X}$, and $y > 0$,*

$$\begin{aligned} &P[X_{T_{n+1}} = j, T_{n+1} - T_n \leq y \mid (X_0, T_0), (X_{T_k}, T_k) \ \forall 1 \leq k \leq n] \\ &= P[X_{T_{n+1}} = j, T_{n+1} - T_n \leq y \mid X_{T_n}] \end{aligned} \tag{0.2.2}$$

   where $T_0 \leq 0 < T_1$.

X is pure if $T_n \xrightarrow{a.s.} \infty$ as $n \to \infty$. If the right side of (0.2.2) is independent on n, then the SMP X is called time-homogeneous. Otherwise, X is called non-homogeneous SMP. Here, T_n is called n^{th} transition time of X.

Definition 0.2.16. *Let X be a random variable taking values in $\mathbb{N}_0$, it is understood that the relevant σ-algebra on $\mathbb{N}_0$ is the discrete σ-algebra of all subsets. Then X is said to have the* ***Poisson distribution*** *with mean c if*

$$P\{X = n\} = \frac{e^{-c} c^n}{n!}, n \in \mathbb{N}_0.$$

Definition 0.2.17. *A mapping $M : \Omega \times \mathcal{S} \to \mathbb{R}_+$ is called a random measure if $\omega \mapsto M(\omega)(A)$ is a random variable for each A in $\mathcal{S}$ and if $A \mapsto M(\omega)(A)$ is a measure on $(S, \mathcal{S})$ for each ω in Ω. We shall denote by $M(A)$ the former random variable: then , we may regard M as the collection of random variables $M(A), A \in \mathcal{S}$. We shall denote by $M(\omega)$ the latter measure $A \mapsto M(\omega)(A)$.*

Definition 0.2.18. *Let $(S, \mathcal{S})$ be the measurable space and let ν be a measure on it. A random measure $\wp$ on $(S, \mathcal{S})$ is said to **Poisson random measure(PRM)** with mean ν if it satisfies the following:*

1. *for every $A \in \mathcal{S}$, the random variable $\wp(A)$ has the Poisson distribution with mean $\nu(A)$, and*
2. *whenever $A_1, \ldots, A_n$ are in $\mathcal{S}$ and disjoint, the random variables $\wp(A_1), \ldots, \wp(A_n)$ are independent, this being true for every $n \geq 2$.*

Remark 0.2.19. *For each ω, the realization of the random measure is a well-defined deterministic measure. Let $\wp$ be a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d$, and it is a measure in $d+1$ dimensions. Since $\wp(\cdot)(A)$ is an integer-valued random variable for all $A \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^d)$, it is clear that $\wp(\omega)(\cdot)$ is a counting measure almost surely. Hence the $d+1$ dimensional measure $\wp(\omega)(dt, dv)$ is not absolutely continuous w.r.t. $m_{d+1}(dt, dv)$, the Lebesgue measure. Indeed $\wp(\omega)(dt, dv)$ is supported only on a countable set. So for any bounded set B, $\int_{[0,T]} \int_B g(t, v)\wp(\omega)(dt, dv)$, an integral that is meaningfull for each $\omega \in \Omega$, can also be written as $\sum_{[0,T]} \sum_B g(t, v)\wp(\omega)(\{t\}, \{v\})$, $\sum_{[0,T]} \int_B g(t, v)\wp(\omega)(\{t\}, dv)$, $\int_{[0,T]} \sum_B g(t, v)\wp(\omega)(dt, \{v\})$ or simply with slight abuse of notation $\int_{[0,T]} \int_B g(t, v)\wp(\omega)(\{t\}, dv)$ or $\int_{[0,T]} \int_B g(t, v)\wp(\omega)(dt, \{v\})$. We will follow this notation in the subsequent chapters.*

Theorem 0.2.20. *Let $f, g : \mathbb{R} \to \mathbb{R}$, δ_a a Dirac measure with atom at a, then*

$$f\left(\int g(x)\delta_a(dx)\right) = f \circ g(a) = \int f \circ g(x)\delta_a(dx),$$

where $f \circ g$ is composition of two function.

Theorem 0.2.21. *Let $\wp$ be a PRM on $(S, \mathcal{S})$. For a fixed $t > 0$ and $B \in \mathcal{S}$, $\wp(\omega)(\{t\} \times B) = 0$ P a.e. ω.*

We restate Theorem 3.4 (p-474) of [8] below.

Theorem 0.2.22. *Let E denote $(\mathbb{R}^d, \|\cdot\|)$ a d-dimensional Euclidean space and $(M(d, m), \|\cdot\|_M)$, the space of all $d \times m$ matrices with a norm $\|\cdot\|_M$. Consider a SDE of the form*

$$Z_t = Z_0 + \int_0^t a(Z_s)\ ds + \int_0^t b(Z_s)\ dW_s + \int_{[0,t]\times\mathbb{R}_+} J(Z_{s-}, v)\wp(ds, dv) \tag{0.2.3}$$

where W and $\wp$ are m-dimensional Wiener Process, and Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}_+$ having intensity m_2 respectively. The maps $a\colon E \to E$, $b\colon E \to M(d,m)$, $J\colon E \times \mathbb{R}_+ \to E$, are assumed to obey the following conditions.

1. *Lipschitz condition: There is a constant c in $\mathbb{R}_+$ such that $\|a(z) - a(z')\| \leq c\|z - z'\|$ and $\|b(z) - b(z')\|_M \leq c\|z - z'\|, \quad \forall z, z' \in E$.*

2. *There is a constant C in $\mathbb{R}_+$ such that $J(z,v) = 0$ for $v > C$ for all z in E.*

Then there exists almost surely unique solution to (0.2.3) that is piece-wise continuous, r.c.l.l. and locally bounded.

Theorem 0.2.23 (Theorem IX.3.8, pg 475, [8])**.** *Let X be a unique solution of (0.2.3), then for each time t, the process $\hat{X} = (X_{t+u})_{u\in\mathbb{R}_+}$ is conditionally independent of $\mathcal{F}_t$ given X_t; given that $X_t = y$, the conditional law of $\hat{X}$ is the same as the law of X under P^y. Here P^y is the conditional probability measure given $\{X_0 = y\}$.*

Theorem 0.2.24 (Theorem IX.3.9, pg 475, [8])**.** *Let X be a unique solution of the (0.2.3), then the process X is strong Markov: For every $\mathcal{F}$-stopping time T, the variable X_T is $\mathcal{F}_T$-measurable, and $\hat{X} = (X_{T+u})_{u\in\mathbb{R}_+}$ is conditionally independent of $\mathcal{F}_T$ given X_t; moreover, for y in E, on the event $\{X_T = y\}$, the conditional law of $\hat{X}$ given X_T is the same as the law of X under P^y.*

Definition 0.2.25. *Let $\mathbb{T} \subset \mathbb{R}_+$. A real-valued stochastic process $X = \{X_t\}_{t\in\mathbb{T}}$ on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is called an **$\mathcal{F}$-submartingale** if X is adapted to $\mathbb{F}$ each X_t is integrable, and $E(X_t - X_s \mid \mathcal{F}_s) \geq 0$ whenever $s < t$. It is called **$\mathcal{F}$-supermartingale** if $-X$ is an $\mathcal{F}$-submartingale, and **$\mathcal{F}$-martingale** if it is both $\mathcal{F}$-supermartingale and an $\mathcal{F}$-submartingale.*

Definition 0.2.26. *The process X is said to be the **Doobs property** for (S,T) provided S and T be stopping times with $S \leq T$, X_S and X_T be well defined and integrable, and $X_S = E[X_T \mid X_S]$. And the process X is said to be **Doobs martingale** on $[0,\eta]$ if η is a stopping time and X has the property for (S,T) for all stopping times S and T with $0 \leq S \leq T \leq \eta$.*

Definition 0.2.27 (Definition V.5.17, pg 219,[8])**.** *Let η be a stopping time. The process X is called a **local martingale on** $[0,\eta)$ if there exist an increasing sequence of stopping times T_n with limit η such that $(X_t - X_0)_{t\in\mathbb{R}_+}$ is a Doob martingale on $[0,T_n]$ for every n. If it is a local martingale on $\mathbb{R}_+$, then it is simply called a **local martingale**.*

Definition 0.2.28. *An adapted, r.c.l.l. process A is **finite variation process(FV)** if almost surely the paths of A are of finite variation on each compact interval of $[0,\infty)$.*

Definition 0.2.29 (Definition V.5.18,pg 220, [8]). *A stochastic process X on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is called a* ***semimartingale*** *if it can be decomposed as $X = L + V$, where L is a local martingale and V is locally of finite variation. And also L and V are adapted to the same filtration as of X.*

Lemma 0.2.30 (Itô formula, Theorem 2.32, pg 78, [33]). *Let X be a semimartingale and let f be a $\mathcal{C}^2$ real function. Then $f(X)$ is again a semimartingale, and the following formula holds*

$$f(X_t) - f(X_0) = \int_{0^+}^{t} f'(X_{s-})dX_s + \frac{1}{2}\int_{0^+}^{t} f''(X_{s-}d[X,X]^c_s \\ + \sum_{0<s<t} \{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\} \tag{0.2.4}$$

where $[X,Y]$ is the quadratic covariation of X and Y, $[X,X]^c_s$ is quadratic variation of path by path continuous part of $[X,X]$ and $\Delta X_t := X_t - X_{t-}$.

Definition 0.2.31. *[25, pg 79] Let C be the set of continuous maps from $\mathbb{R}^d$ into itself. Let $\{\phi_{s,t}; 0 \le s < t\}$ (or simply denoted by $\{\phi_{s,t}\}$ be a family of C-valued random variables. It is called a* ***stochastic flow of $\mathcal{C}^\infty$ maps*** *if it satisfies:*

1. *Maps $\phi_{s,t} : \mathbb{R}^d \to \mathbb{R}^d$ are $\mathcal{C}^\infty$ a.s. for any $s < t$.*
2. *$\phi_{t,u} \circ \phi_{s,t} = \phi_{s,u}$ a.s. for any $s < t < u$. Here $\phi \circ \psi$ is the composition of two maps ϕ, ψ of $\mathbb{R}^d$ into itself.*

Furthermore if it satisfies the following (1') and the above (2), then it is called a ***stochastic flow of diffeomorphisms****.*

1' Maps $\phi_{s,t} : \mathbb{R}^d \to \mathbb{R}^d$ are diffeomorphisms a.s. for any $s < t$.

A stochastic flow $\{\phi_{s,t}\}$ is called ***continuous*** *if $\phi_{s,t}(x), x \in \mathbb{R}^d$ and its derivatives $\partial^i \phi_{s,t}(x)$ are continuous with respect to $(x, s, t)(s < t)$ a.s. for any multi-index i.*

Chapter 1

Introduction

A semi-Markov process is a pure jump process such that the embedded chain is Markov, but the conditional distribution of the sojourn time need not be exponential. So a semi-Markov process does not enjoy the memoryless property of its Markov counterpart. In this thesis we consider a general class of semi-Markov processes on countable state-space, and having differentiable kernel such that the embedded Markov chain may or may not be homogeneous. The rate matrices may also have infinite trace. We only assume uniform bound on entries of the rate matrix.

The class of semi-Markov processes (SMP) subsumes the class of pure jump Markov processes. A component-wise semi-Markov (CSM) process, on the other hand, is a member of an even larger class of pure jump processes. Each component of a vector valued CSM is an SMP having a given law. These components may or may not be independent and the laws may or may not be identical.

It is known that for certain processes their law can be represented using the martingale formulation. Writing down a stochastic process on an Euclidean space as a solution of a stochastic differential equation (SDE) also gives another way of representing the law. Such SDEs may also be viewed as the semimartingale representation of the said process. Needless to mention, an SDE represents a stochastic flow too. We establish a family of semimartingale representations, involving Poisson random measure (PRM), of this class of semi-Markov processes. In particular, we have proved existence of almost sure unique solution to the proposed system of SDEs driven by a PRM. Then we show that a solution indeed possesses the desired semi-Markov law.

Solutions of an SDE with different initial conditions, together form a vector valued process where each component has identical law. Moreover, they being driven by the same noise, are not independent. This also gives rise to a CSM or system of semi-Markov processes, having dependent components, provided the SDE of an SMP has been considered in the first place. In the present study of CSM, our major focus is on the meeting and merging events of the components. Questions related to these events have immediate answers when the components are independent. However, when the components are driven by dependent or identical noises, the questions are worth pondering.

We suitably select a particular flow by fixing the gaping parameter, for which we study interactions of a pair of semi-Markov processes, solving the same equation but having two different initial conditions. We have obtained expressions of the meeting and merging probabilities in the next transition. A set of sufficient conditions are obtained under which any two solutions of the flow eventually meet or merge with probability one. Many numerical examples are considered for clarifying the intricacies and implications of the theoretical results. In one example the distribution of time of first meeting and merging are obtained and are compared with that of the holding times, for the purpose of illustration.

Apparently [26] and [37], which were presented at the International Congress of Mathematicians held at Amsterdam in 1954, are the first available literature that discuss the mathematical aspects of semi-Markov processes (SMP). In Lévy's work, [26], the definition of SMP was presented as a generalization of Markov chain. Around the same time, independent to Lévy's work, Smith [37] and then Tackas [38] have also introduced SMP. We provided a modern version of the definition, see Definition 0.2.15. We often call pure homogeneous SMP' as semi-Markov process or SMP only. SMP has been defined by many authors using the renewal processes unlike Definition 0.2.15. For example, in [34], Pyke has introduced the SMPs by specifying the conditional distribution of next state and holding time given the past states and past holding times. We notice that the σ-algebra generated by the transition times is identical to that generated by the initial time and past holding times. Thus the conditional distribution in [34] is identical to the conditional probability in Definition 0.2.15. In [34], a concept of regularity has been introduced for assuring that the chain is pure. Besides, classification of the states has also been studied there. On the other hand in [35], by considering the finite state SMPs, Pyke has derived expressions for the distribution functions of first passage times, as well as for the marginal distribution function. Furthermore, the limiting behavior of a Markov Renewal process has been discussed, and the stationary probabilities have been derived. Various aspects and approaches regarding limiting behavior has also been studied by Orey [31], around the same time. In [12], Bennet has studied some properties of sub-chain of SMP, which are obtained via regenerating points, if exists. SMP beyond the class of time homogeneity has been first studied in [18] and [28]. Various different aspects and generalisations of non-homogeneous SMP(NHSMP) has further been explored in [20, 22, 21, 23]. In these references several applications of NHSMP has also been emphasised.

It has been noted by several authors (see Nummelin [30], Athreya et al [2] and references therein) that an SMP can be augmented with the age process to obtain a jointly Markov process whose Feller property and infinitesimal generator can be derived (see [24, Chapter 2]). In a recent work [11] Elliott has presented a semimartingale representation of semi-Markov chain in contrast to the traditional description of a semi-Markov chain in terms of a renewal process. This presentation is different from that in [13] and [14] where a semimartingale representation appears using an integration with respect to a Poisson random measure(PRM). Such semimartingale representations are useful for studying several aspects including the stochastic flow of semi-Markov dynamics.

The non-homogeneous SMP, augmented with the age and transition count processes is represented as semi-martingales using a system of stochastic integral equations involving a Poisson random measure. The coefficients of the equation depend on a given transition rate function and an an additional gaping parameter. It is worth noting that neither the coefficients are compactly supported nor the intensity measure of the PRM is finite. Note that, compactly supported coefficient or finiteness of intensity measure are the standard assumptions under which an SDE involving PRM is studied commonly (see [8, 19]). So, we produce a self-contained proof of the existence and uniqueness of the solution to the SDE. This extends the results obtained for homogeneous semi-Markov process. Subsequently, we extend the results by showing that the state component of the solution is a pure non-homogeneous SMP with the given non-homogeneous transition rate function. We also derive the law of the bivariate process obtained from two solutions of the equation having two different initial conditions.

The CSM is a generalization of semi-Markov processes into a broader class of pure jump processes. The combination of state processes of more than one semi-Markov dynamics forms a semi-Markov system(SMS) or a component-wise semi-Markov (CSM) process having dependent or independent components. The SMSs [39, 41, 40], or CSMs [9, 10] with independent components have been introduced for modelling some random dynamics. However, a CSM with dependent components has not been studied in the literature yet. The CSM, studied in [10] possesses a well defined bounded transition rate function and hence that has been used to characterize the CSM. However, the definition of CSM does not imply existence of a rate function. In view of this it is important to find an alternative way of characterizing a general type of CSM. Needless to mention, the kernel characterization should be most suitable in this regard. We recall that the transition rate exists if and only if the kernel is almost everywhere differentiable. In that case, the rate can be expressed in terms of the kernel and vice versa. It is also easy to note that the knowledge of kernels of all individual components of CSM is sufficient to characterize the CSM, provided the components are independent to each other. However, in this thesis we consider an extension of CSM, appearing in [10], by dropping the independence condition. We further allow the state-space of each component be non-identical and at most countable. So, we propose characterization of CSM using a novel notion of kernel. The way we define the kernel, is broad enough to include both the dependent and independent component cases and both the homogeneous and non-homogeneous cases. As per our knowledge, this is the first effort in the literature to characterize a general CSM using a kernel based approach. Then we derived the infinitesimal generator formula for CSM with arbitrary components driven by one PRM.

The study of meeting and coalescence of stochastic processes is an active branch of probability theory. Some of the earliest instances of such study dates back to Arratia [1], and Harris [17] where they have considered merging of one dimensional Brownian flow. On the other hand mixing for a class of non Markov flows have been investigated by Melbourne and Terhesiu [29]. However, to the best of our knowledge, questions regarding meeting and merging have not been addressed in the literature for stochastic flow of SMPs.

In [4] for the stability analysis of Markov modulated diffusions, the merging of Markov chains has been crucially used. In view of this, we believe that the study of meeting and merging of multiple semi-Markov particles are relevant for investigating stability properties of a diffusion that is modulated by semi-Markov processes.

We give an outline of the remaining chapter here. In Chapter 2, we look at a broad category of semimartingale representations of SMPs with a fixed instantaneous transition rate. An expression for the conditional probability of meeting and merging in the next transition, merging at a meeting time, is derived. We also construct a set of sufficient conditions under which a pair of SMPs will eventually meet and merge with probability one. The work presented in this chapter 3 results from a collaboration with Dr. Subhamay Saha. We have given a semimartingale representation of a class of semi-Markov processes; this representation is more general than the class introduced in Chapter 2, as it includes non-homogeneous semi-Markov processes. Again we have considered a particular pair of solutions of SDE (3.2.6)-(3.2.8) with two distinct initial conditions and investigate the various event of the meeting and merging. We derived an expression of the conditional probability of meeting in the next transition, coherent meeting and merging in the next transition and eventual meeting and merging. Also, the number of transitions required to encounter a meeting is shown to have all moments finite. In Chapter 4, the theoretical results are elaborated with the help of several numerical examples. We have considered numerical examples of homogeneous SMP and non-homogeneous Markov processes, where we calculated the probability of meeting in the next transition and the expected first meeting time for homogeneous SMP and the probability of coherent meeting in the next transition. We have provided an algorithm for simulating homogeneous SMP and non-homogeneous Markov processes. In Chapter 5, we introduced the definition of general CSM on countable state space; also its associated kernels and showed that it satisfies the transition kernel definition. We also defined the marginal of l^{th} component. Further, We have introduced the correlated PRM; with the help of this, we have given a semi-Martingale representation of a general CSM and computed the infinitesimal generator of CSM with d components driven by a single PRM.

Chapter 2

Homogeneous semi-Markov processes

2.1 Introduction

Apparently [26] and [37], which were presented at the International Congress of Mathematicians held at Amsterdam in 1954, are the first available literature that discuss the mathematical aspects of semi-Markov processes (SMP). In Lévy's work, [26], the definition of SMP was presented as a generalization of Markov chain. Around the same time, independent to Lévy's work, Smith [37] and then Tackas [38] have also introduced SMP. In Chapter 1 Definition 0.2.15, we provided a modern version of the definition.

As the study of non-homogeneous or impure SMP is excluded from this chapter, from now we will call 'pure homogeneous SMP' as semi-Markov process or SMP only. We also recall that in Chapter 1 of [16], a book by Boris Harmalov, a stepped SMP is introduced and in subsequent chapters further generalizations to continuous state space appears. We confine ourselves to the study of SMPs on a countable state space.

SMP has been defined by many authors using the renewal processes unlike Definition 0.2.15. For example, in [34], Pyke has introduced the SMPs by specifying the conditional distribution of next state and holding time given the past states and past holding times. We notice that the σ-algebra generated by the transition times is identical to that generated by the initial time and past holding times. Thus the conditional distribution in [34] is identical to the conditional probability in Definition 0.2.15. In [34], a concept of regularity has been introduced for assuring that the chain is pure. Besides, classification of the states has also been studied there. On the other hand in [35], by considering the finite state SMPs, Pyke has derived expressions for the distribution functions of first passage times, as well as for the marginal distribution function. Furthermore, the limiting behavior of a Markov Renewal process has been discussed, and the stationary probabilities have been derived. Various aspects and approaches regarding limiting behavior has also been studied by Orey [31], around the same time. In [12], Bennet has studied some properties of sub-chain of SMP, which are obtained via regenerating points, if exists.

It has been noted by several authors (see Nummelin [30], Athreya et al [2] and references therein) that an SMP can be augmented with the age process to obtain a jointly Markov process whose Feller property and infinitesimal generator can be derived (see [24, Chapter 2]). In a recent work [11] Elliott has presented a semimartingale dynamics of semi-Markov chain in contrast to the traditional description of a semi-Markov chain in terms of a renewal process. This presentation is different from that in [13] and [14] where a semimartingale representation appears using an integration with respect to a Poisson random measure(PRM). Such semimartingale representations are useful for studying several aspects including the stochastic flow of semi-Markov dynamics.

A study of merging for a couple of renewal processes on the the same probability space appears in [27]. In this work, Lindvall studied coupling events for monotonic hazard rate case only. In this connection a previous work by Brown [6] on comparisons of such renewal processes is worth mentioning. More recently, for studying dietary contamination dynamics, Bouguet [5] applied the notion of merging of renewal processes. However, as per our knowledge, questions regarding meeting and merging have not been addressed in the literature for SMPs with the general bounded measurable transition rates. We address this question in this Chapter using the SDE representation [15] of SMPs. However, as per our knowledge, questions regarding meeting and merging have not been addressed in the literature for stochastic flow of SMPs.

Here, we consider a wide class of semimartingale representations of an SMP with a given instantaneous transition rate. With the help of an additional gaping parameter, given a semi-Markov dynamics, we could consider a family of stochastic flows. The law of a single solution of course do not depend on the gaping parameter. However, we show, the joint distribution of a couple of solutions with different initial conditions do depend on the additional gaping parameter. Then we suitably select a particular flow, for which the investigation of meeting and merging of two solutions of the same SDE, starting with two different states, becomes convenient. Although the study of meeting and merging event of a finite-state continuous-time irreducible Markov chain is straightforward, that is not the case for semi-Markov counterpart. We show with an example, that the meeting time need not be a merging time for a pair of SMPs. We derive an expression of the conditional probability of merging at a meeting time. A set of sufficient conditions are also obtained under which a pair of SMPs eventually merge with probability one.

2.2 Semimartingle Representation for Homogeneous Semi-Markov Processes

As the study of non-homogeneous or impure SMP is excluded from this chapter, from now we will call 'pure homogeneous SMP' as semi-Markov process or SMP only. Let $(\Omega, \mathcal{F}, P)$ be the underlying probability space and $\mathcal{X}$ the state space, a countable subset of $\mathbb{R}$. Endow the set $\mathcal{X}_2 := \{(i,j) \in \mathcal{X}^2 \;\mid\; i \neq j\}$ with a total order $\prec$. Let $\mathcal{B}(\mathbb{R}^d)$ denote the Borel σ-algebra on $\mathbb{R}^d$ and m_d denote the Lebesgue measure on $\mathbb{R}^d$. Let $\lambda := (\lambda_{ij})$ denote a matrix in which the i^{th} diagonal element is $\lambda_{ii}(y) := -\sum_{j\in\mathcal{X}\setminus\{i\}} \lambda_{ij}(y)$ and for each $(i,j) \in \mathcal{X}_2$, $\lambda_{ij}\colon [0,\infty) \to (0,\infty)$ is a bounded measurable function such that

(A1) $C := \sum_{i\in\mathcal{X}} \sum_{j\in\mathcal{X}\setminus\{i\}} \|\lambda_{ij}\|_\infty < \infty$, and

(A2) $\lim_{y\to\infty} \gamma_i(y) = \infty$, where $\gamma_i(y) := \int_0^y \lambda_i(y')dy'$, where $\lambda_i(y) := |\lambda_{ii}(y)|$.

For each $(i,j) \in \mathcal{X}_2$, we consider another measurable function $\tilde{\lambda}_{ij}\colon [0,\infty) \to (0,\infty)$ and a collection of generic intervals such that $\tilde{\lambda}_{ij}(y) \le \|\lambda_{ij}\|_\infty$ for almost every $y \ge 0$ and

$$\lambda_{ij}(y) \le \tilde{\lambda}_{ij}(y), \text{ and } \Lambda_{ij}(y) = \left(\sum_{(i',j')\prec(i,j)} \tilde{\lambda}_{i'j'}(y)\right) + \Big[0, \lambda_{ij}(y)\Big) \tag{2.2.1}$$

for each $y \ge 0$, where $a + B = \{a + b \mid b \in B\}$ for $a \in \mathbb{R}, B \subset \mathbb{R}$. From (2.2.1), it is clear that for every $y \ge 0$, $\{\Lambda_{ij}(y)\colon (i,j) \in \mathcal{X}_2\}$ is a collection of disjoint intervals which is denoted by Λ. We define h_Λ and g_Λ on $\mathcal{X} \times \mathbb{R}_+ \times \mathbb{R}$ as

$$h_\Lambda(i,y,v) := \sum_{j\in\mathcal{X}\setminus\{i\}} (j-i)\mathbb{1}_{\Lambda_{ij}(y)}(v) \tag{2.2.2}$$

$$g_\Lambda(i,y,v) := y \sum_{j\in\mathcal{X}\setminus\{i\}} \mathbb{1}_{\Lambda_{ij}(y)}(v) \tag{2.2.3}$$

where $\mathbb{R}_+$ denotes the set of non-negative real numbers. We consider the following system of stochastic differential equations in X and Y

$$X_t = X_0 + \int_{0^+}^t \int_{\mathbb{R}} h_\Lambda(X_{u-}, Y_{u-}, v)\wp(du, dv) \tag{2.2.4}$$

$$Y_t = Y_0 + t - \int_{0^+}^t \int_{\mathbb{R}} g_\Lambda(X_{u-}, Y_{u-}, v)\wp(du, dv) \tag{2.2.5}$$

for $t > 0$, where the domain of integration $\int_{0+}^t$ is $(0,t]$, and the PRM $\wp(du,\ dv)$ is on $\mathbb{R}_+ \times \mathbb{R}$ with intensity $m_2(du, dv)$, and defined on the probability space $(\Omega, \mathcal{F}, P)$. We also assume that $\{\wp((0,t]\times dv)\}_{t\geq 0}$ is adapted to $\{\mathcal{F}_t\}_{t\geq 0}$, a filtration of $\mathcal{F}$ satisfying the usual hypothesis. Evidently, $\tilde{\lambda}$ controls the left end points of the intervals in Λ and so can be utilized to regulate relation between solutions to (2.2.4)-(2.2.5) with different initial conditions. Indeed a specific choice namely, $\tilde{\lambda}_{ij} = \|\lambda_{ij}\|_\infty$ a.e. simplifies the relation between the intervals $\Lambda_{ij}(y)$ with different values of i, j and y. Thanks to the independence of $\wp$ measure of disjoint sets, the system (2.2.4)-(2.2.5) defines a non-Brownian stochastic flow with independent increments. Some special cases of such a system have been considered by many authors following [13]. However, as per our knowledge, system (2.2.4)-(2.2.5) of this generality has not been studied before. Although some special cases give rise to solutions with law identical to that of (2.2.4)-(2.2.5), but they correspond to a particular flow. For example, in [14] and [32] $\tilde{\lambda}$ has been taken identical to λ. Here in (2.2.1) we consider a larger family of flows by introducing the additional parameter $\tilde{\lambda}$. Evidently, $\tilde{\lambda}$ controls the left end points of the intervals in Λ and so can be utilized to regulate relation between solutions to (2.2.4)-(2.2.5) with different initial conditions. Indeed a specific choice of $\tilde{\lambda}$, namely, $\tilde{\lambda}_{ij} = \|\lambda_{ij}\|_\infty$ simplifies the relation between the intervals $\Lambda_{ij}(y)$ with different values of i, j and y. We need the following lemma for proving the subsequent theorem that asserts the existence and uniqueness of solution to this general system.

Lemma 2.2.1. *For each fixed $\omega \in \Omega$, consider the set $\mathcal{D} := \{s' \in (0, \infty) \mid \wp(\omega)(\{s'\}\times E) > 0\}$, where $E \in \mathcal{B}(\mathbb{R})$, and $\wp$ is a Poisson random measure with intensity m_2. If $m_1(E) < \infty$, then set $\mathcal{D}$ has no limit point in $\mathbb{R}$ almost surely.*

Proof. Evidently, from Remark 0.2.19, $\wp(\omega)(\mathcal{D} \times E) = \wp(\omega)((0,\infty) \times E) = \infty$ if $m_1(E) > 0$. Hence, $\mathcal{D}$ is non-empty iff $m_1(E) > 0$. If $m_1(E) < \infty$, for any natural number n, $\wp(\omega)([0,n] \times E)$ is a Poisson random variable with mean $n \times m_1(E)$. Hence $P(\wp([0,n] \times E) < \infty) = 1$. Thus $\mathcal{D} \cap [0,n]$ is finite with probability 1 for each $n \geq 1$. Hence $P(\bigcap_{n>1}\{\omega \mid \mathcal{D} \cap [0,n] \text{ is finite}\}) = 1$. Therefore, $\mathcal{D}$ has no limit point in $\mathbb{R}$ w.p. 1.

Theorem 2.2.2. *There exists a unique strong solution $(X,Y) = \{(X_t,Y_t)\}_{t\geq 0}$ to the coupled system of stochastic integral equations (2.2.4)-(2.2.5). Furthermore, almost surely X and Y have r.c.l.l. piece-wise constant and piece-wise linear paths respectively.*

Remark 2.2.3. *We recall Theorem 0.2.22. Thus for proving the first part of Theorem 2.2.2 it is enough to rewrite Equations (2.2.4)-(2.2.5) in the form of (0.2.3). To this end we embed $\mathcal{X}$ in $\mathbb{R}$ by identifying that with the set of natural numbers and take $d = 2$, $Z_t = (X_t, Y_t)$. For each $i \in \mathcal{X}$, $y \geq 0$, we also set $a((i,y)) = (0,1)$, $b((i,y)) = O_{2\times 2}$, the null matrix of order 2, and $J((i,y),v) = (h_\Lambda, -g_\Lambda)(i,y,v)$. Clearly Condition (1) is valid as a and b are constant functions in this case. For verifying (2), we note that (2.2.2) and (2.2.3) imply that for each i and almost every y, h_Λ and g_Λ are sums of functions which are non-zero only on the intervals $\Lambda_{ij}(y)$ for $j \in \mathcal{X}\setminus\{i\}$. Furthermore, $\Lambda_{ij}(y)$ is contained in $\left[0, \sum_{\mathcal{X}_2} \|\lambda_{ij}\|_\infty\right]$ for each i,j and almost every y. Hence the support of $J(i,y,\cdot)$ is contained in $\left[0, \sum_{\mathcal{X}_2} \|\lambda_{ij}\|_\infty\right]$ which is a finite interval by (A1). Thus Condition (2) is also true. Hence the first part of*

Theorem 2.2.2 follows from Theorem 3.4 (p-474) of [8]. However, the second part, which asserts a specific property of the solution, should be justified separately. It turns out that for this cause, it is essential to first spell out (2.2.4)-(2.2.5)*, in the similar line of what appears in [8]. In view of this, we include below a self contained proof of the first part before attempting to prove the second part.*

Proof. (Theorem 2.2.2) By Assumption (A1), $C = \sum_{\mathcal{X}_2} \|\lambda_{ij}\|_\infty$ is finite. Hence for almost every $y \geq 0$, the total absolute length of the generic intervals $\sum_{(i,j)\in\mathcal{X}_2} |\Lambda_{ij}(y)| \leq C$. For each $\omega \in \Omega$, we define the set $\mathcal{D} := \{s \in \ (0,\ \infty) | \wp(\omega)(\{s\} \times [0,\ C]) > 0\}$ be the time coordinate of the point masses of a realisation of the PRM $\wp(\omega)$. Interval $[0, C]$ has finite Lebesgue measure which is the vertical section of the intensity measure of $\wp(\omega)$. By Lemma 2.2.1, $\mathcal{D}$ has no limit points in $\mathbb{R}$ with probability 1. Thus we can enumerate $\mathcal{D}$, say $\mathcal{D} = \{\sigma_n\}_{n=1}^{\infty}$, where $\sigma_1 < \cdots < \sigma_n < \sigma_{n+1} < \cdots$ for each ω. For each $n \in \mathbb{N}$, $\sigma_n \colon \Omega \to (0, \infty]$ and $\{\sigma_n \leq t\} = \{\omega \mid \wp((0,t] \times [0,C]) \geq n\} \in \mathcal{F}_t$, as $\wp((0,t] \times [0,C])$ is $\mathcal{F}_t$ measurable. Hence σ_n is a stopping time for each $n \geq 1$.

For a fixed ω, we plan to construct a solution to equations (2.2.4)-(2.2.5) on the time interval $[0,\ \sigma_1]$. Then we extend this solution to the time interval $(\sigma_1,\ \sigma_2]$, and so on. Since

$$\wp(\omega)([0,\ \sigma_1) \times\ [0,\ C]) = 0,$$

for $t \in [0,\ \sigma_1)$

$$X_t(\omega) = X_0 + \int_{(0,t]} \int_{[0,C]} h_\Lambda(X_{u-},\ Y_{u-},\ v)\wp(\omega)\ (du,\ dv) = X_0$$

and

$$Y_t(\omega) = Y_0 + t - \int_{(0,t]} \int_{[0,C]} g_\Lambda(X_{u-},\ Y_{u-},\ v)\wp(\omega)\ (du,\ dv) = Y_0 + t.$$

This gives unique solution on $[0, \sigma_1)$. Moreover by using above, at $t = \sigma_1$,

$$\begin{aligned} X_{\sigma_1}(\omega) =& X_0 + \int_{[0,C]} h_\Lambda(X_0,\ Y_{\sigma_1 -},\ v)\wp(\omega)(\{\sigma_1\} \times dv), \\ Y_{\sigma_1}(\omega) =& Y_0 + \sigma_1 - \int_{[0,C]} g_\Lambda(X_0,\ Y_{\sigma_1 -},\ v)\wp(\omega)(\{\sigma_1\} \times dv). \end{aligned}$$

Hence this is the unique solution in the time interval $[0,\ \sigma_1]$. Continuing in the similar way we can construct solution for each consecutive interval $(\sigma_n,\ \sigma_{n+1}]$, where $n \geq 1$. Now we recall that σ_n is increasing and diverges to infinity with probability 1, due to Lemma 2.2.1. Therefore, these intervals cover the entire positive real time-axis. Hence, the solution is globally determined with probability 1.

Furthermore, for a fixed ω, $X_t(\omega) = X_{\sigma_n(\omega)}$ for all $t \in [\sigma_n(\omega), \sigma_{n+1}(\omega))$. Hence X is an r.c.l.l. and piece-wise constant process almost surely. Next we show the piece-wise linear feature of Y. First we note that $\int_{[0,C]} g_\Lambda(X_{t-},\ Y_{t-},\ v)\wp(\omega)(\{t\} \times dv)$ is zero for all $t \in (\sigma_n, \sigma_{n+1})$ for every $n \geq 1$. Let $n_1 := \min\{l \geq 1\colon \int_{[0,C]} g_\Lambda(X_{\sigma_l-},\ Y_{\sigma_l-},\ v)\wp(\omega)(\{\sigma_l\} \times dv) \neq 0\}$. Then $t = \sigma_{n_1}$ is the first time when the integral $\int_{[0,C]} g_{\tilde{\lambda}}(X_{t-},\ Y_{t-},\ v)\wp(\omega)(\{t\} \times dv)$ is non-zero. Consequently, $Y_t = Y_0 + t$ for all $t \in [0, \sigma_{n_1})$ and hence $Y_{\sigma_{n_1}-} = Y_0 + \sigma_{n_1}$ and

$$\begin{aligned} 0 \neq& \int_{(0,\sigma_{n_1}]} \int_{[0,C]} g_\Lambda(X_{t-},\ Y_{t-},\ v)\wp(\omega)(\{t\} \times dv) \\ =& \int_{[0,C]} g_\Lambda(X_{\sigma_{n_1}-},\ Y_{\sigma_{n_1}-},\ v)\wp(\omega)(\{\sigma_{n_1}\} \times dv) = Y_{\sigma_{n_1}-} \end{aligned}$$

using the fact that $\wp(\omega)(\{\sigma_{n_1}\} \times [0, C]) = 1$ and $g_\Lambda(i, y, v) \in \{0, y\}$. Thus from (2.2.5) and above expressions

$$\begin{aligned} Y_{\sigma_{n_1}} =& Y_0 + \sigma_{n_1} - \int_{(0,\sigma_{n_1}]} \int_{[0,C]} g_\Lambda(X_{t-},\ Y_{t-},\ v)\wp(\omega)(\{t\} \times dv) \\ =& Y_{\sigma_{n_1}-} - Y_{\sigma_{n_1}-} = 0. \end{aligned}$$

Thus $Y_{\sigma_{n_1}} = 0$. In general, for every $m \geq 1$, we set

$$n_{m+1} := \min\left\{ l > n_m\colon \int_{[0,C]} g_\Lambda(X_{\sigma_l-},\ Y_{\sigma_l-},\ v)\wp(\omega)(\{\sigma_l\} \times dv) \neq 0 \right\}. \tag{2.2.6}$$

In other words, for every $t \geq 0$,

$$\int_{\mathbb{R}} g_\Lambda(X_{t-},\ Y_{t-},\ v)\wp(\{t\} \times dv) = \begin{cases} Y_{\sigma_{n_m}-} & \text{, if } t = \sigma_{n_m} \text{ for some } m \geq 1 \\ 0 & \text{, otherwise.} \end{cases} \tag{2.2.7}$$

Then by summarising above observations, one gets from (2.2.5) that for every $t \in [0, \infty)$

$$\begin{aligned} Y_t =& Y_0 + t - \sum_{\{r \geq 1 | \sigma_{n_r} \leq t\}} Y_{\sigma_{n_r}-} \\ =& Y_0 + t - \sum_{r=1}^{\infty} (Y_{\sigma_{n_r}-}) 1_{[\sigma_{n_r}, \infty)}(t) \end{aligned} \tag{2.2.8}$$

holds with probability 1. Hence Y is r.c.l.l, and piece-wise linear. □

2.3 Semi-Markov Law of the Solution

Definition 2.3.1. *The sequence of transition times $\{T_n\}_{n \geq 1}$ is given by $T_n := \inf\{t > T_{n-1} \vee 0 : X_t \neq X_{t-}\}$ where $T_0 := -Y_0$. We define the holding time $\tau_n := T_n - T_{n-1}$ for all $n \geq 1$. The number of transitions until time t is denoted by N_t which is given by $\max\{m \geq 0 \mid T_m \leq t\}$.*

From the above definition

$$X_u - X_{u-} = \int_{\mathbb{R}} h_\Lambda\, (X_{u-}, Y_{u-}, v)\wp(\{u\} \times dv) \tag{2.3.1}$$

is non-zero if and only if $u = T_n$ for some positive integer n.

Remark 2.3.2. *From (2.2.2) and (2.2.3) it is evident that for each $i \in \mathcal{X}$, $y > 0$, the maps $h_\Lambda(i, y, \cdot)$ and $g_\Lambda(i, y, \cdot)$ have identical supports. Hence the integral $\int_{[0,C]} g_\Lambda(X_{t-}, Y_{t-}, v)\wp(\{t\}\times dv)$ is nonzero if and only if $\int_{[0,C]} h_\Lambda(X_{t-}, Y_{t-}, v)\wp(\{t\} \times dv)$ is nonzero since $Y_{t-} > 0$ for all $t \geq 0$. We also recall that it is shown in the above proof that Y jumps only at time t when $\int_{[0,C]} g_\Lambda(X_{t-},\ Y_{t-},\ v)\wp(\{t\} \times dv)$ is nonzero. Therefore, X and Y jump simultaneously as X jumps only at time t when $\int_{[0,C]} h_\Lambda(X_{t-},\ Y_{t-},\ v)\wp(\{t\} \times dv)$ is nonzero. In other words, the sequence $\{\sigma_{n_m}\}_{m\geq 1}$, where n_m is as in (2.2.6), gives the times of consecutive jumps of X. Again, under (A1), due to Lemma 2.2.1, $\sigma_{n_m} \to \infty$ almost surely. Thus all the jump times of X are included in $\{\sigma_{n_m}\}_{m\geq 1}$ which is a sub-sequence in $\mathcal{D}$. Hence, $\sigma_{n_m} = T_m$ for all $m \geq 1$.*

Lemma 2.3.3. *For each $n \in \mathbb{N}$, (i) $Y_{T_n} = 0$, and (ii) $Y_{T_n-} = T_n - T_{n-1}$. Also (iii) $Y_t = t - T_{N_t}$, where N_t is as in Definition 2.3.1.*

Proof. (i) By (2.2.8) and Remark 2.3.2 we obtain, for all $n \in \mathbb{N}$,

$$\begin{aligned} Y_{T_n} =& Y_0 + T_n - \sum_{\{r\geq 1 | T_r \leq T_n\}} Y_{T_r-} \\ =& \left(Y_0 + T_n - \sum_{\{r\geq 1 | T_r < T_n\}} Y_{T_r-} \right) - Y_{T_n-} \\ =& Y_{T_n-} - Y_{T_n-} = 0 \end{aligned}$$

as (2.2.8) implies (by replacing $\leq$ by $<$) $Y_{t-} = Y_0 + t - \sum_{\{r\geq 1 | T_r < t\}} Y_{T_r-}$. Alternatively, by taking the left limit $t \uparrow T_n$ in (2.2.8),

$$Y_{T_n-} = Y_0 + T_n - \sum_{\{r\geq 1 | T_r < T_n\}} Y_{T_r-}. \tag{2.3.2}$$

Hence we have shown above that

$$Y_t = 0 \text{ iff } t = T_n \text{ for some } n \in \mathbb{N}. \tag{2.3.3}$$

(ii) We obtain for $n \geq 2$, by adding and subtracting T_{n-1} on the right side of (2.3.2),

$$Y_{T_n-} = T_n - T_{n-1} + \left(Y_0 + T_{n-1} - \sum_{r=1}^{n-1} Y_{T_r-} \right).$$

Again using (2.2.8) with $t = T_{n-1}$, the above is equal to $T_n - T_{n-1} + Y_{T_{n-1}} = T_n - T_{n-1}$, since $Y_{T_{n-1}} = 0$ from (2.3.3). To complete the proof, we should show $Y_{T_1-} = T_1 - T_0$. This is true as

$$\begin{aligned} Y_{T_1-} =& Y_0 + T_1 - \sum_{\{r\geq 1 | T_r < T_1\}} Y_{T_r-} \\ =& - T_0 + T_1 \end{aligned}$$

using $T_0 = -Y_0$ from Definition 2.3.1.

(iii) From (2.2.8), and Definition 2.3.1, and part (ii) of the lemma, we have

$$\begin{aligned} Y_t =& Y_0 + t - \sum_{\{r\geq 1 | T_r \leq t\}} (T_r - T_{r-1}) \\ =& - T_0 + t - \sum_{r=1}^{N_t} (T_r - T_{r-1}) = t - T_{N_t}. \end{aligned}$$

This completes the proof. □

Theorem 2.3.4. *Let $Z = (X, Y) = \{(X_t,\ Y_t)\}_{t\geq 0}$ be the unique strong solution to (2.2.4)-(2.2.5). Then the following hold.*

i. The process Z is a strong Markov process.

ii. The embedded chain for the pure jump process X is Markov.

Proof. We have already seen in Remark 2.2.3 that the finite support condition on the integrands in (2.2.4)-(2.2.5) hold true. Indeed all conditions mentioned in Remark 2.2.3 are true. So by Theorem IX.3.8, and IX.3.9 of [8] (p-475), the process Z is strong Markov, i.e., $\hat{Z} = (Z_{T+u})_{u\in\mathbb{R}_+}$ is conditionally independent of $\mathcal{F}_T$ given Z_T, where T is any $\{\mathcal{F}_t\}_t$ stopping time. We also recall that for each $n(\geq 0) \in \mathbb{Z}$, T_n is a stopping time. Hence, due to the strong Markov property, $Z_{T_{n+1}}$ is conditionally independent to the σ-algebra $\mathcal{F}_{T_n}$ given Z_{T_n}. That is, $\{(X_{T_n}, Y_{T_n})\}_n$ is Markov. Finally, due to (2.3.3), the embedded chain, $\{X_{T_n}\}_n$ is Markov. □

Theorem 2.3.5. *Let $(X, Y) = \{(X_t,\ Y_t)\}_{t\geq 0}$ be the solution to (2.2.4)-(2.2.5), then $\{X_t\}_{t\geq 0}$ is an SMP.*

Proof. We have already seen in the proof of Theorem 2.2.2, that X is an r.c.l.l process. Next, we need to show (0.2.2), i.e., for each $n \geq 0$,

$$P[X_{T_{n+1}} = j, T_{n+1} - T_n \leq y \mid X_{T_0}, T_0, X_{T_1}, T_1, \ldots, X_{T_n}, T_n] = P[X_{T_{n+1}} = j, T_{n+1} - T_n \leq y \mid X_{T_n}].$$

We note that

$$\begin{aligned}
&P(X_{T_{n+1}} = j, T_{n+1} - T_n \leq y \mid (X_{T_0}, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n)) \\
=&P(T_{n+1} - T_n \leq y \mid (X_{T_0}, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n)) \\
&\times P(X_{T_{n+1}} = j \mid (X_{T_0}, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n), \{T_{n+1} - T_n \leq y\}).
\end{aligned} \tag{2.3.4}$$

Each of the two probabilities on the right side is further simplified below. For almost every $\omega \in \Omega$, equation (2.2.7), Remark 2.3.2 and Lemma 2.3.3 imply that for any $n \geq 0$

$$\int_{(T_n, T_n+t]} \int_{\mathbb{R}} g_\Lambda(X_{T_n}, u - T_n, v) \wp(du, dv) = \begin{cases} 0, & \text{for } t < T_{n+1} - T_n \\ T_{n+1} - T_n, & \text{for } t = T_{n+1} - T_n. \end{cases}$$

Hence, by a suitable change of variable, almost surely $T_{n+1} - T_n$ is the first occurrence of a non-zero value of the following map

$$t \mapsto \int_{(0,t]} \int_{\mathbb{R}} g_\Lambda(X_{T_n}, u, v) \wp(T_n + du, dv)$$

and that occurs at $t = T_{n+1} - T_n$. Again, since $\wp(T_n + du, dv)$ is independent to $\mathcal{F}_{T_n}$ we obtain, $T_{n+1} - T_n$ is conditionally independent to $\mathcal{F}_{T_n}$ given X_{T_n}. Thus

$$\begin{aligned}
&P\left(T_{n+1} - T_n \leq y \mid (X_{T_0}, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n)\right) \\
&= P(T_{n+1} - T_n \leq y \mid X_{T_n}).
\end{aligned} \tag{2.3.5}$$

By substituting $u = T_{n+1}$ in Equation (2.3.1), and using Lemma 2.3.3 we get

$$X_{T_{n+1}} = X_{T_n} + \int_{\mathbb{R}} h_\Lambda(X_{T_n}, T_{n+1} - T_n, v) \wp(\{T_{n+1}\} \times dv), \tag{2.3.6}$$

as $X_{T_{n+1}-} = X_{T_n}$ and $Y_{T_{n+1}-} = T_{n+1} - T_n$. Thus using (2.3.6)

$$\begin{aligned}
&P\left(X_{T_{n+1}} = j \mid (X_{T_0}, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n), \{T_{n+1} - T_n \leq y\}\right) \\
&\quad = P\Bigg(\int_{\mathbb{R}} h_\Lambda(X_{T_n}, T_{n+1} - T_n, v) \wp(\{T_n + (T_{n+1} - T_n)\} \times dv) = j - X_{T_n} \Bigg| \\
&\qquad\qquad (X_{T_0}, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n), \{T_{n+1} - T_n \leq y\} \Bigg).
\end{aligned}$$

Again, using the independence of $\wp(T_n + du, dv)$ to $\mathcal{F}_{T_n}$ and conditional independence of $T_{n+1} - T_n$ to $\mathcal{F}_{T_n}$ given X_{T_n} we conclude, the above expression is equal to

$$\begin{aligned}
&P\left(\int_{\mathbb{R}} h_\Lambda(X_{T_n}, T_{n+1} - T_n, v) \wp(\{T_n + (T_{n+1} - T_n)\} \times dv) = j - X_{T_n} \,\middle|\, X_{T_n}, \{T_{n+1} - T_n \leq y\} \right) \\
&= P\left(X_{T_{n+1}} = j \mid X_{T_n}, \{T_{n+1} - T_n \leq y\}\right)
\end{aligned} \tag{2.3.7}$$

using (2.3.6). Thus, using the simplifications (2.3.7) and (2.3.5) in (2.3.4), we obtain

$$\begin{aligned}
&P(X_{T_{n+1}} = j, T_{n+1} - T_n \leq y \mid (X_{T_0}, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n)) \\
&= P\left(T_{n+1} - T_n \leq y \mid X_{T_n}\right) P\left(X_{T_{n+1}} = j \mid X_{T_n}, \{T_{n+1} - T_n \leq y\}\right) \\
&= P\left(X_{T_{n+1}} = j, T_{n+1} - T_n \leq y \mid X_{T_n}\right).
\end{aligned}$$

Hence, X is an SMP. □

2.4 Expression of Transition Kernel

In this section we derive an expression of the transition kernel. For each $i \in \mathcal{X}$, we define a function $F(\cdot \mid i)\colon [0,\infty) \to [0,1]$ as

$$F(y \mid i) := 1 - e^{-\gamma_i(y)} \tag{2.4.1}$$

where $\gamma_i(y)$ is as in (A2). Since, $\gamma_i(y)$, being an integral of a bounded Lebesgue measurable function, is absolutely continuous in y, and hence differentiable almost everywhere. Let $f(y \mid i)$ be the almost everywhere derivative of $F(y \mid i)$. We also define a matrix $p := (p_{ij}(y))_{\mathcal{X}\times\mathcal{X}}$, such that

$$p_{ij}(y) := \begin{cases} \dfrac{\lambda_{ij}(y)}{-\lambda_{ii}(y)}, & \text{if } j \neq i \\ 0, & \text{if } j = i. \end{cases} \tag{2.4.2}$$

This ensures that p is a transition probability matrix for each $y \geq 0$. The following proposition asserts that p gives the conditional probability of selecting a state at the time of transition given the holding time and location of the previous state. Furthermore, the map $F(y \mid i)$ as in (2.4.1) is also asserted as the conditional cumulative distribution function of the holding time given the state is i.

Proposition 2.4.1. *Let $Z = (X, Y)$ be the solution to* (2.2.4)-(2.2.5)*, then the following hold.*

i. *$F(\cdot \mid i)$ is the conditional cumulative distribution function of the holding time of the process X.*

ii. *For $i \neq j$, $p_{ij}(y) = P[X_{T_{n+1}} = j \mid X_{T_n} = i, Y_{T_{n+1}-} = y]$.*

Proof. Using (2.2.3) and (2.2.7), the conditional probability of no transition in the next y

unit time, given that the state at the n^{th} transition is i, is given by,

$$\begin{aligned}
&P[X_t = X_{t-}, \forall t \in (T_n, T_n + y] \mid X_{T_n} = i] \\
&= P\left[\int_{\mathbb{R}} g_\Lambda(X_{u-},\ u - T_n,\ v)\wp(\{u\} \times dv) = 0, \forall u \in (T_n, T_n + y] \mid X_{T_n} = i\right] \\
&= P\left[\wp\left\{(u, v) \in (T_n, T_n + y] \times \mathbb{R}_+ \mid v \in \bigcup_{j \neq X_{T_n}} \Lambda_{X_{T_n}, j}(u - T_n)\right\} = 0 \mid X_{T_n} = i\right] \\
&= e^{-\gamma_i(y)} \qquad (2.4.3)
\end{aligned}$$

since, the intensity of $\wp$ is Lebesgue measure, and the Lebesgue measure of $\{(u, v) \in (T_n, T_n + y] \times \mathbb{R}_+ \mid v \in \cup_{j \neq i}\Lambda_{i,j}(u - T_n)\}$ is $\int_{T_n}^{T_n+y} \sum_{j \neq i} \lambda_{ij}(u - T_n)du$ which is equal to $\gamma_i(y)$ (see (A2)). Using (2.4.3), the conditional cumulative distribution function at y of τ_{n+1}, the holding time after the n^{th} transition, given the n^{th} state, is

$$\begin{aligned}
P[\tau_{n+1} \leq y \mid X_{T_n} = i] &= 1 - P[X_t = X_{t-}, \forall t \in (T_n, T_n + y] \mid X_{T_n} = i] \\
&= 1 - e^{-\gamma_i(y)}
\end{aligned}$$

for all $y \geq 0$ and $i \in \mathcal{X}$. Thus (i) follows from (2.4.1).

We note that, for $j \neq i$, $P[X_{T_{n+1}} = j \mid X_{T_n} = i, Y_{T_{n+1}-} = y]$ is the conditional probability of the event that the $(n+1)$th state is j, given that $T_{n+1} = T_n + y$ and the nth state is i. Using (2.3.6), the above is the conditional probability that a Poisson point mass appears in $\{T_n + y\} \times \Lambda_{ij}(y)$ given that the point mass lies somewhere in $\{T_n + y\} \times \cup_{j \neq i}\Lambda_{ij}(y)$ and no transition of X occurs during $(T_n, T_n + y)$. If these three events are denoted by A, B, and C respectively, then the conditional probability $P(A \mid B \cap C)$ can be simplified as $P(A \mid B)$ because C is independent to both A and B. Thus using the Lebesgue intensity of $\wp$,

$$\begin{aligned}
&P\left[X_{T_{n+1}} = j \mid X_{T_n} = i, Y_{T_{n+1}-} = y\right] \\
&= P\left[\wp(\{T_n + y\} \times \Lambda_{ij}(y)) = 1 \mid \wp(\{T_n + y\} \times \bigcup_{j \neq i} \Lambda_{ij}(y)) = 1\right] \\
&= \frac{|\Lambda_{ij}(y)|}{|\bigcup_{j \neq i} \Lambda_{ij}(y)|} \\
&= \frac{\lambda_{ij}(y)}{\lambda_i(y)}
\end{aligned}$$

for every $y \geq 0, j \neq i$. Thus (ii) follows from (2.4.2). □

Remark 2.4.2. *We note that under Assumptions (A1) and (A2),* $F(y \mid i) < 1$ *for all* $y \geq 0$ *and* $\lim_{y \to \infty} F(y \mid i) = 1$. *Thus, the holding times are unbounded but finite almost surely. By dropping (A1), one may include a class of SMPs having bounded holding times. However,* *we exclude that class from our discussion. It is also important to note that the SMPs having discontinuous cdf of holding time are also not considered in the present setting. Nevertheless, the present study subsumes countable-state continuous time Markov chains and the processes having age dependent transitions as appears in [14]. Moreover, in Theorem 2.4.4, we obtain the transition kernel that is homogeneous in time. In other words, we have excluded the time-inhomogeneous SMP from the present study too.*

Proposition 2.4.3. *We have, for almost every $y \geq 0$,*

$$p_{ij}(y)\frac{f(y \mid i)}{1 - F(y \mid i)} = \begin{cases} \lambda_{ij}(y), & \text{for } i \neq j, \\ 0, & \text{for } i = j. \end{cases}$$

Proof. By differentiating both sides of (2.4.1), we obtain $f(y \mid i) = \lambda_i(y)e^{-\gamma_i(y)}$ for a.e. $y \geq 0$. This is equal to $\lambda_i(y)(1 - F(y \mid i))$ using (2.4.1). Hence, for a.e. $y \geq 0$, and $i \in \mathcal{X}$

$$\frac{f(y \mid i)}{1 - F(y \mid i)} = \lambda_i(y). \tag{2.4.4}$$

If $i \neq j$, for a.e. $y \geq 0$, using (2.4.2)

$$p_{ij}(y)\frac{f(y \mid i)}{1 - F(y \mid i)} = -\lambda_{ii}(y) \times \frac{\lambda_{ij}(y)}{-\lambda_{ii}(y)} = \lambda_{ij}(y).$$

The case for $i = j$ follows from (2.4.2) directly. □

Theorem 2.4.4. *Let X be an SMP as in Theorem 2.3.5. Then, the associated kernel is given by*

$$P[X_{T_{n+1}} = j, T_{n+1} - T_n \leq y \mid X_{T_n} = i] = \int_0^y e^{-\gamma_i(s)}\lambda_{ij}(s)\, ds,$$

which is denoted by $Q_{ij}(y)$ for every $y > 0$, and $i \neq j$.

Proof. Using Proposition 2.4.1(i) and (ii) and Lemma 2.3.3

$$\begin{aligned} P\left[X_{T_{n+1}} = j, Y_{T_{n+1}-} \leq y \mid X_{T_n} = i\right] &= \mathbb{E}\left[P\left(X_{T_{n+1}} = j, Y_{T_{n+1}-} \leq y \mid X_{T_n} = i, Y_{T_{n+1}-}\right) \mid X_{T_n} = i\right] \\ &= \int_0^\infty \mathbb{1}_{[0,y]}(s) P\left[X_{T_{n+1}} = j \mid X_{T_n} = i, Y_{T_{n+1}-} = s\right] f(s \mid i)\, ds \\ &= \int_0^y p_{ij}(s) f(s \mid i)\, ds. \end{aligned}$$

For each $i \neq j$, using Proposition 2.4.3 and (2.4.1), the right side of above can be rewritten as

$$\int_0^y (1 - F(s \mid i))\lambda_{ij}(s)\, ds = \int_0^y e^{-\gamma_i(s)}\lambda_{ij}(s)\, ds = Q_{ij}(y).$$

□

Proposition 2.4.5. *Let X be an SMP as in Theorem 2.3.5. Then, λ is the instantaneous transition rate matrix.*

Proof. The rate of transition from state i to j at age y is given by

$$\begin{aligned}
&\lim_{h\to 0}\frac{1}{h}\left[P(X_{T_{n+1}}=j, T_{n+1}-T_n\in(y,y+h]\mid X_{T_n}=i,\{T_{n+1}-T_n>y\})\right]\\
&=\lim_{h\to 0}\frac{1}{h}\frac{P(X_{T_{n+1}}=j, T_{n+1}-T_n\in(y,y+h]\mid X_{T_n}=i)}{P(T_{n+1}-T_n>y\mid X_{T_n}=i)}\\
&=\lim_{h\to 0}\frac{1}{h}\frac{P(X_{T_{n+1}}=j, T_{n+1}-T_n\le y+h\mid X_{T_n}=i)-P(X_{T_{n+1}}=j, T_{n+1}-T_n\le y\mid X_{T_n}=i)}{1-P(T_{n+1}-T_n\le y\mid X_{T_n}=i)}.
\end{aligned}$$

Using Theorem 2.4.4, the above limit is equal to $\frac{\frac{d}{dy}Q_{ij}(y)}{1-F(y|i)}$ which can further be simplified as $\lambda_{ij}(y)$. □

Remark 2.4.6. *We have obtained $Q_{ij}(y)=\int_0^y p_{ij}(s)f(s\mid i)\,ds$ in the proof of Theorem 2.4.4, which expresses $Q_{ij}(\cdot)$ in terms of the $p_{ij}(\cdot)$, and $f(\cdot\mid i)$. These parameters give the age dependent transition probabilities and the conditional holding time densities. In an alternative conditioning, the kernel can also be expressed as*

$$Q_{ij}(y)=P\left[X_{T_{n+1}}=j\mid X_{T_n}=i\right]P\left[T_{n+1}-T_n\le y\mid X_{T_n}=i, X_{T_{n+1}}=j\right],$$

which is the product of transition probabilities of embedded chain and the conditional cdf of holding time given the current and the next state. This representation of the kernel is more general as that does not require absolute continuity of holding time CDF. For this reason, this factorisation of kernel is more popular in the literature. We also recall that the transition kernel Q characterizes an SMP. We have seen in the above two resuls that if Q is differentiable, the instantaneous transition rate matrix λ exists, and each of Q and λ can be expressed in terms of the other.

2.5 Homogeneous Component-wise Semi-Markov Process

Notation 2.5.1. *Fix $i,j\in\mathcal{X}$ and $y_1,y_2\ge 0$. Let $Z^1=(X^1,Y^1)$ and $Z^2=(X^2,Y^2)$ be the strong solutions of* (2.2.4)-(2.2.5) *with initial conditions*

$$X_0^1=i, Y_0^1=y_1, \text{ and } X_0^2=j, Y_0^2=y_2$$

respectively. The jump times of $Z:=(Z^1,Z^2)$ is denoted by $\{T_n\}_{n\in\mathbb{N}_0}$ and given by $T_0:=0$ and $T_n:=\inf\{t>T_{n-1}\colon t\in T^1\cup T^2\}$ for all $n\ge 1$ where T^l denotes the collection of transition times of X^l for each $l=1,2$.

The above notation is adopted henceforth. We impose the following restriction on $\tilde{\lambda}$.

(A3) For all $(i,j) \in \mathcal{X}_2$, and for almost every $y \geq 0$, set $\tilde{\lambda}_{ij}(y) = \|\lambda_{ij}\|_\infty$.

Remark 2.5.2. *In the preceding subsection we have seen that the law of the solution to* (2.2.4)-(2.2.5) *depends only on the λ matrix and the initial position and does not depend on the choice of $\tilde{\lambda}$. Hence, (A3) does not impose any condition on the law of Z^1 or Z^2 separately. However, the law of Z depends on the choice of $\tilde{\lambda}$. Therefore, (A3) selects a specific flow from the family specified in* (2.2.1). *We select that, as the absence of (A3) significantly complicates the relations between the intervals $\Lambda_{ij}(y)$ with different values of i, j and y and thus ramifies the relation between Z_1 and Z_2. On the other hand (A3) implies a very simple relation, namely $\cup_{y\geq 0}\Lambda_{ij}(y)$ are disjoint for different values of i and j. This helps us to compute expressions of various parameters related to the law of $Z = (Z^1, Z^2)$. This assumption is central for our study of mixing and merging times of Z.*

All the subsequent results, which assume (A3), do hold under a relaxed condition that for an $\alpha \geq 1$, $\tilde{\lambda}_{i'j'}(y) = \alpha\|\lambda_{i'j'}\|_\infty$ for all $y \geq 0$ and $(i',j') \in \mathcal{X}_2$. Even the proof of Theorem 2.2.2 also works after replacing C by αC. Nevertheless, this relaxation is artificial as that does not enlarge the scope of stochastic flow under consideration. So, for avoiding cumbersome notations arising due to an artificial relaxation, we follow (A3) only.

Since, Z^1 and Z^2 as in Notation 2.5.1 are Markov, $Z = (Z^1, Z^2)$ is also Markov. It has state and age components $X = (X^1, X^2)$ and $Y = (Y^1, Y^2)$ respectively. While each of X^1 and X^2 is semi-Markov, the pure jump process X is not. Rather, X is a component-wise semi-Markov process (CSM) and the Markov process Z is called the augmented process of CSM X. A CSM with independent components has been introduced for modelling financial assets in [10]. However, a CSM with dependent components has not been studied in the literature yet. Since, for our case, the components of the CSM X are driven by a single Poisson random measure, they are not independent. In view of this, it is interesting to derive the law of X by finding the generator of Z. To this end, we recall Itô's lemma for r.c.l.l. semimartingales. Let $\varphi\colon (\mathcal{X} \times \mathbb{R}_+)^2 \to \mathbb{R}$ be bounded and continuously differentiable in its continuous variables. Using the expression of J in Remark 2.2.3, we write

$$
\begin{aligned}
& d\varphi(Z_t^1, Z_t^2) - \left(\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2}\right)\varphi(Z_t^1, Z_t^2)dt \\
&= \varphi(Z_t^1, Z_t^2) - \varphi(Z_{t-}^1, Z_{t-}^2) \\
&= \varphi\left(Z_{t-}^1 + \int_{\mathbb{R}_+} J(Z_{t-}^1, v)\wp(dt,dv), Z_{t-}^2 + \int_{\mathbb{R}_+} J(Z_{t-}^2, v)\wp(dt,dv)\right) - \varphi(Z_{t-}^1, Z_{t-}^2) \\
&= \int_{\mathbb{R}_+} \left[\varphi(Z_{t-}^1 + J(Z_{t-}^1, v), Z_{t-}^2 + J(Z_{t-}^2, v)) - \varphi(Z_{t-}^1, Z_{t-}^2)\right] \wp(dt,dv) \\
&= \left(\int_{\mathbb{R}_+} \left[\varphi(Z_{t-}^1 + J(Z_{t-}^1, v), Z_{t-}^2 + J(Z_{t-}^2, v)) - \varphi(Z_{t-}^1, Z_{t-}^2)\right] dv\right) dt + dM_t
\end{aligned}
$$

where M is the martingale obtained by integration wrt the compensated PRM $\wp(dt, dv) - dtdv$. We get the third equality by using Therorem 0.2.20. For simplifying the above integral term, we impose (A3) and divide the derivation in two complementary cases.

Case1: Assume $X^1_{t-} \neq X^2_{t-}$. Now under (A3), the intervals $\Lambda_{X^1_{t-}j_1}(Y^1_{t-})$ and $\Lambda_{X^2_{t-}j_2}(Y^2_{t-})$ are disjoint for every j_1 and j_2. Thus by considering these intervals where the integrand is non-zero constants, we get

$$\begin{aligned}
&\int_{\mathbb{R}_+} \left[\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v)) - \varphi(Z^1_{t-}, Z^2_{t-})\right] dv \\
&= \int_{\overset{2}{\underset{k=1}{\cup}}\left(\underset{j \neq X^k_{t-}}{\cup} \Lambda_{X^k_{t-}j}(Y^k_{t-})\right)} \left[\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v)) - \varphi(Z^1_{t-}, Z^2_{t-})\right] dv \\
&= \sum_{j \neq X^1_{t-}} \left[\varphi(j, 0, Z^2_{t-}) - \varphi(Z^1_{t-}, Z^2_{t-})\right] |\Lambda_{X^1_{t-}j}(Y^1_{t-})| \\
&\quad + \sum_{j \neq X^2_{t-}} \left[\varphi(Z^1_{t-}, j, 0) - \varphi(Z^1_{t-}, Z^2_{t-})\right] |\Lambda_{X^2_{t-}j}(Y^2_{t-})|
\end{aligned}$$

where $|I|$ is the length of the interval I.
Case2: Assume that $X^1_{t-} = X^2_{t-} = i$ say. Also recall that under (A3), the intervals $\Lambda_{ij}(y_1)$ and $\Lambda_{ij}(y_2)$ are having identical left end points. So, $\Lambda_{X^1_{t-}j_1}(Y^1_{t-})$ and $\Lambda_{X^2_{t-}j_2}(Y^2_{t-})$ are not disjoint when $j_1 = j_2$. Thus

$$\begin{aligned}
&\int_{\mathbb{R}_+} \left[\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v)) - \varphi(Z^1_{t-}, Z^2_{t-})\right] dv \\
&= \int_{\underset{j \neq i}{\cup}\left(\Lambda_{ij}(Y^1_{t-}) \cup \Lambda_{ij}(Y^2_{t-})\right)} [\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v)) - \varphi(Z^1_{t-}, Z^2_{t-})] dv \\
&= \sum_{j \neq i} [\varphi(j, 0, Z^2_{t-}) - \varphi(Z^1_{t-}, Z^2_{t-})] |\Lambda_{ij}(Y^1_{t-}) \setminus \Lambda_{ij}(Y^2_{t-})| \\
&\quad + \sum_{j \neq i} [\varphi(Z^1_{t-}, j, 0) - \varphi(Z^1_{t-}, Z^2_{t-})] |\Lambda_{ij}(Y^2_{t-}) \setminus \Lambda_{ij}(Y^1_{t-})| \\
&\quad + \sum_{j \neq i} [\varphi(j, 0, j, 0) - \varphi(Z^1_{t-}, Z^2_{t-})] |\Lambda_{ij}(Y^1_{t-}) \cap \Lambda_{ij}(Y^2_{t-})|.
\end{aligned}$$

Thus by combining the expressions under both the cases, the generator $\mathcal{A}$ of (Z^1, Z^2) is given by

$$\begin{aligned}\mathcal{A}\varphi(z_1,z_2) = &\left(\frac{\partial}{\partial y_1}+\frac{\partial}{\partial y_2}\right)\varphi(z_1,z_2) + \sum_{j\notin\{i_1\}}(\lambda_{i_1j}(y_1)-\delta_{i_1i_2}\lambda_{i_2j}(y_2))^+[\varphi(j,0,z_2)-\varphi(z_1,z_2)]\\ &+\sum_{j\notin\{i_2\}}(\lambda_{i_2j}(y_2)-\delta_{i_1i_2}\lambda_{i_1j}(y_1))^+[\varphi(z_1,j,0)-\varphi(z_1,z_2)]\\ &+\delta_{i_1i_2}\sum_{j\notin\{i_1,i_2\}}(\lambda_{i_1j}(y_1)\wedge\lambda_{i_2j}(y_2))[\varphi(j,0,j,0)-\varphi(z_1,z_2)]. \end{aligned} \tag{2.5.1}$$

where $z_1=(i_1,y_1)$, $z_2=(i_2,y_2)$, δ_{ij} is Kronecker delta, $a^+=\max(0,a)$ and $a\wedge b=\min(a,b)$. This leads to the following theorem.

Theorem 2.5.3. *Under (A3), the infinitesimal generator $\mathcal{A}$ of the augmented process $Z=(Z^1,Z^2)$ is given by* (2.5.1) *where $Z^1=(X^1,Y^1)$ and $Z^2=(X^2,Y^2)$ are as in Notation 2.5.1.*

From the above derivation of generator of Z, it is not difficult to guess its expression when Z_1 and Z_2 are driven by two different independent Poisson random measures. Indeed, the intervals $\Lambda_{X^1_{t-}j_1}(Y^1_{t-})$ and $\Lambda_{X^2_{t-}j_2}(Y^2_{t-})$ can be treated as if they are disjoint for any value of X^1_{t-} and X^2_{t-}. Thus an expression like (2.5.1) can be obtained where $\delta_{i_1i_2}$ should be replaced by zero irrespective of i_1,i_2. This is agreeing with (2.10) of [10]. In that case $\mathcal{A}\varphi(z_1,z_2)=\mathcal{A}_1\varphi(\cdot,z_2)(z_1)+\mathcal{A}_2\varphi(z_1,\cdot)(z_2)$, where $\mathcal{A}_i$ denotes the infinitesimal generator of Z^i.

Remark 2.5.4. *Although given a $\omega\in\Omega$ for two different families $\tilde\lambda^1$ and $\tilde\lambda^2$ one obtains two different solutions paths for SDE* (2.2.4)-(2.2.5)*, the law does not differ. Indeed it is evident from SDE* (2.2.4)-(2.2.5) *that the law of (X,Y) does not depend on the choice of $\tilde\lambda$ and depends only on the λ matrix and the initial position. Hence, (A3) imposes no condition on the laws of Z^1 and Z^2 separately. However, the law of Z depends on the choice of $\tilde\lambda$. Therefore, (A3) selects a specific flow from the family specified in* (2.2.1)*. We select that, as the absence of (A3) significantly complicates the relations between the intervals $\Lambda_{ij}(y)$ with different values of i, j and y and thus ramifies the relation between Z_1 and Z_2. On the other hand (A3) implies a very simple relation, namely $\cup_{y\ge0}\Lambda_{ij}(y)$ are disjoint for different values of i and j. This helps us to compute expressions of various probabilities related to meeting and merging times of X^1 and X^2. This assumption is central for our study.*

2.6 Meeting and Merging at the Next Transition

Definitions 2.6.1. *Given an $\{\mathcal{F}_t\}_{t\ge0}$ stopping time T, the time τ of subsequent meeting by the processes X^1 and X^2 is given by $\tau:=\inf\{t>T: X^1_t=X^2_t, \min(Y^1_t,Y^2_t)=0\}$. If*

$\{\tau < \infty\}$, *then* X^1 *and* X^2 *are said to* ***meet eventually***. *If at a meeting time* τ, *their transition counts [see Definition 2.3.1]* N^1_τ *and* N^2_τ *are identical, then the said meeting is called* ***coherent***. *The merging time of* X^1 *and* X^2 *is given by* $\tau' := \inf\{t' \geq 0 \mid X^1_t = X^2_t, \forall t \geq t'\}$ *and if* $\{\tau' < \infty\}$, *then they are said to* ***merge***.

The nature of meeting and merging for a semi-Markov family is more involved than those for the Markovian special case. We clarify this in the next section.

Markov pure jump process, although a special case of (2.2.4)-(2.2.5), deserves a separate mention due to its importance. Hence we first consider a special case where λ is independent of the age variable y and satisfies (A1). Evidently, (A2) holds too. Furthermore, assume that $\tilde{\lambda}_{ij}(y) = \lambda_{ij}$, a constant function for each $(i,j) \in \mathcal{X}_2$. Hence (2.2.4) reduces to

$$X_t = X_0 + \int_{0^+}^t \tilde{h}(X_{s-}, v)\wp(ds, dv) \tag{2.6.1}$$

where $\tilde{h}(i,v) := h_\Lambda(i,y,v) = \sum_{j \in \mathcal{X}\setminus\{i\}} (j-i)\mathbb{1}_{\Lambda_{ij}(y)}(v)$ is constant in y, as the intervals $\Lambda_{ij}(y)$, do not vary with y variable. Uniqueness result of (2.6.1) implies the following.

Theorem 2.6.2. *Let* X^1 *and* X^2 *be strong solutions of SDE* (2.6.1) *with initial states* $X^1_0 = i$ *and* $X^2_0 = j$ *respectively. Then, if* X^1 *and* X^2 *meet, they merge at the first meeting.*

Proof. For a $\omega \in \Omega$, if there exists a $t' > 0$ such that $X^1_{t'} = X^2_{t'} = k$ for some $k \in \mathcal{X}^1$, then using (2.6.1) for $t > t'$, both X^1 and X^2 solve

$$X_t = X_{t'} + \int_{t'^+}^t \tilde{h}(X_{s-}, v)\wp(ds, dv) = k + \int_{t'^+}^t \tilde{h}(X_{s-}, v)\wp(ds, dv).$$

Now using almost sure uniqueness of the strong solution of the above SDE, X^1 and X^2 would be identical from time t' onward. Thus X^1 and X^2 merge at their first meeting time.

It is interesting to note that, if λ is constant, the merging time of X^1 and X^2, as given in Theorem 2.6.2, is a stopping time. This is because, merging and meeting times coincide, and the latter is a stopping time. This consequence is not valid for a general semi-Markov family. Indeed, if X^1 and X^2 are as in Notation 2.5.1, at the meeting time they may have unequal ages and those age variables appear in the SDE (2.2.4)-(2.2.5). So, the mere uniqueness of the SDE does not imply merging at the first meeting time. We produce below an example of a meeting event which is not the merging of a semi-Markov family.

Example 2.6.3. *Let* $\mathcal{X} = \{1, 2\}$, *with* $(1,2) \prec (2,1)$; *also* $\lambda_{12}(y) = \lambda_{21}(y) = \frac{y}{1+y}$, *and* $\tilde{\lambda}_{12}(y) = \tilde{\lambda}_{21}(y) = \sup_{(0,\infty)} |\frac{y}{1+y}| = 1$ *for all* $y \geq 0$. *Thus for every* $(i,j) \in \mathcal{X}_2$, $\Lambda_{ij}(y) = [i-1, i-1+\frac{y}{1+y})$. *We further assume that* $Z^l = (X^l, Y^l)$ *is the strong solution of* (2.2.4)-(2.2.5) *with above parameters and initial conditions* $(X^l_0, Y^l_0) = (l, \mathbb{1}_{\{2\}}(l))$ *for* $l = 1, 2$ *respectively. Now fix a sample* $\omega \in \Omega$ *such that* $\wp(\omega)|_{[0,3/2]\times[0,2]} = \delta_{(1,3/2)} + \delta_{(3/2,1/2)}$, *the addition of two Dirac measures at* $(1, 3/2)$ *and* $(3/2, 1/2)$ *respectively. Then none of the processes has*

[1]Note that, it is not necessary that X^1 and X^2 transit to state k at the same time.

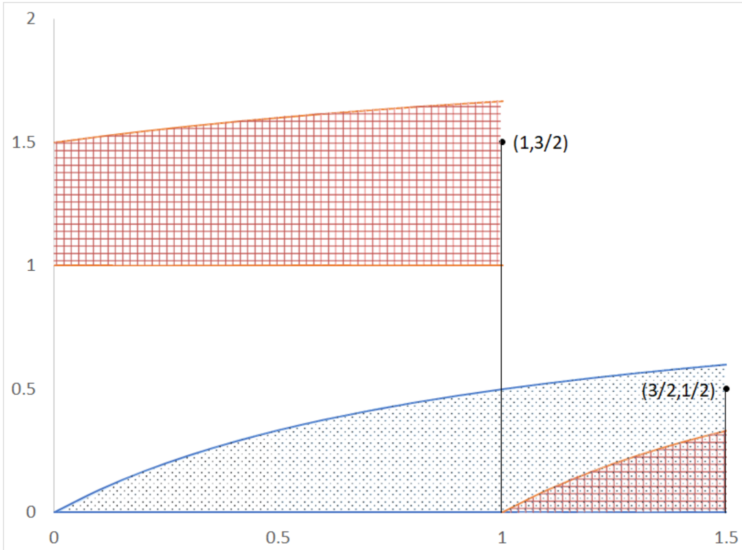


Figure 2.1: The t and v variables are plotted along horizontal and vertical axes. The point masses are shown by black dots. The intervals relevant for transitions of the first and second processes are plotted vertically and shown in blue and red respectively.

transition until time $t = 1$. Hence, for both $l = 1, 2$,

$$X^l_{1-} = l, \quad and, \quad Y^l_{1-} = Y^l_0 + 1 - \int_{(0,1)} \int_{\mathbb{R}} g_\Lambda(X^l_{u-}, Y^l_{u-}, v)\wp(du, dv)(\omega) = \mathbb{1}_{\{2\}}(l) + 1 = l.$$

Then from (2.2.4)

$$X^l_1 = X^l_{1-} + \int_{\mathbb{R}} h_\Lambda(X^l_{1-}, Y^l_{1-}, v)\wp(\{1\}, dv)(\omega) = l + h_\Lambda(l, l, 3/2).$$

Therefore, using (2.2.2) *and the intervals* $\Lambda_{12}(1), \Lambda_{21}(2)$, *we get* $X^1_1 = 1 + (2-1)\mathbb{1}_{[0,1/2)}(3/2) = 1$ *and* $X^2_1 = 2 + (1-2)\mathbb{1}_{[1,1+2/3)}(3/2) = 1$. *Thus,* $t = 1$ *is a meeting time. However, this is not a merging time, because at* $t = 3/2$, X^1 *and* X^2 *separate, which is shown below. We note that until* $t = 3/2$, X^1 *and* X^2 *are at state 1 since* $t = 0$, *and* $t = 1$ *respectively. So, while the pre-transition state* $X^l_{3/2-}$ *is* 1 *for each* $l = 1, 2$, *the pre-transition ages* $Y^1_{3/2-}$, *and* $Y^2_{3/2-}$ *are* $3/2$ *and* $1/2$ *respectively. Consequently,*

$$X^l_{3/2} = 1 + \int_{\mathbb{R}} h_\Lambda(1, Y^l_{3/2-}, v)\wp(\{3/2\}, dv)(\omega) = 1 + \mathbb{1}_{\Lambda_{12}(Y^l_{3/2-})}(1/2) = \begin{cases} 2, & for\ l = 1 \\ 1, & for\ l = 2 \end{cases}$$

since, $1/2 \in \Lambda_{12}(3/2) = [0, \frac{3/2}{1+3/2}) = [0, 3/5)$ *and* $1/2 \notin \Lambda_{12}(1/2) = [0, \frac{1/2}{1+1/2}) = [0, 1/3)$.

Theorem 2.6.4. *Assume (A3). Let* $Z^1 = (X^1, Y^1)$ *and* $Z^2 = (X^2, Y^2)$ *be as in Notation 2.5.1 where* $i \neq j$. *The probability of* X^1 *and* X^2 *meeting in the next transition is*

$$\int_0^\infty e^{-\int_0^y (\lambda_i(y_1+t)+\lambda_j(y_2+t))dt}(\lambda_{ij}(y_1 + y) + \lambda_{ji}(y_2 + y))dy.$$

Proof. In this proof we will utilise that for every $y' \geq 0$ and $y'' \geq 0$, $\cup_{k\neq i}\Lambda_{ik}(y')$ is disjoint to $\cup_{k\neq j}\Lambda_{jk}(y'')$ when $i \neq j$. This is consequence of definitions of the intervals in (2.2.1), and (A3). Non-meeting event in the next transition of X^1 and X^2, happens in two ways.
Case 1: X^1 has the first transition to a state which is different from X^2_0 before X^2 transits for the first time. This event can be written as $\mathcal{E} := \{X^1_{T_1-} \neq X^1_{T_1}, X^1_{T_1} \neq X^2_{T_1}\}$. We will make use of $P(\mathcal{E} \mid \mathcal{F}_0) = E[P(\mathcal{E} \mid T_1) \mid \mathcal{F}_0]$ and, the expression of conditional density η_{T_1} of T_1 given $\{X^1_{T_0} = i, X^2_{T_0} = j, Y^1_{T_0} = y_1, Y^2_{T_0} = y_2\}$. Clearly,

$$P(\mathcal{E}|T_1 = y) = \frac{m_1(\underset{k\notin\{i,j\}}{\cup}\Lambda_{ik}(y_1+y))}{m_1(\underset{k\notin\{i\}}{\cup}\Lambda_{ik}(y_1+y) \cup \underset{k\notin\{j\}}{\cup}\Lambda_{jk}(y_2+y))} = \frac{\lambda_i(y_1+y) - \lambda_{ij}(y_1+y)}{\lambda_i(y_1+y) + \lambda_j(y_2+y)}.$$

Moreover,

$$\eta_{T_1}(y) = e^{-m_2(B)}(\lambda_i(y_1+y) + \lambda_j(y_2+y)),$$

where $B := \underset{t\in[0,y)}{\cup}\left(\{t\} \times \left((\underset{k\notin\{i\}}{\cup}\Lambda_{ik}(y_1+t)) \cup (\underset{k\notin\{j\}}{\cup}\Lambda_{jk}(y_2+t))\right)\right)$. Indeed, the event of no transition of X^1 and X^2 until first y unit of time, is equivalent to $\{\wp(B) = 0\}$, the non-occurrence of Poisson point mass in B. Clearly, $P(\{\wp(B) = 0\} \mid X^1_{T_0} = i, X^2_{T_0} = j, Y^1_{T_0} = y_1, Y^2_{T_0} = y_2)$ is equal to $e^{-m_2(B)}$, and $m_2(B) = \int_0^y(\lambda_i(y_1+t) + \lambda_j(y_2+t))dt$. Hence

$$\begin{aligned} P(\mathcal{E} \mid \mathcal{F}_0) &= \int_0^\infty P(\mathcal{E} \mid T_1 = y)\eta_{T_1}(y)dy \\ &= \int_0^\infty e^{-\int_0^y(\lambda_i(y_1+t)+\lambda_j(y_2+t))dt}[\lambda_i(y_1+y) - \lambda_{ij}(y_1+y)]dy. \end{aligned} \tag{2.6.2}$$

Similarly for Case 2, i.e., X^2 has the first transition to a state, different from X^1_0, before X^1 transits for the first time is given by,

$$P(X^2_{T_1-} \neq X^2_{T_1}, X^2_{T_1} \neq X^1_{T_1} \mid \mathcal{F}_0) = \int_0^\infty e^{-\int_0^y(\lambda_i(y_1+t)+\lambda_j(y_2+t))dt}[\lambda_j(y_2+y) - \lambda_{ji}(y_2+y)]dy. \tag{2.6.3}$$

Hence the total probability (denoted by $a'_{(i,j,y_1,y_2)}$) of not meeting in the next transition is sum of the probabilities appearing in (2.6.2), and (2.6.3).
Using $\phi_1(y) := e^{-\int_0^y(\lambda_i(y_1+t)+\lambda_j(y_2+t))dt}\,(\lambda_i(y_1+y) + \lambda_j(y_2+y))$,

$$\begin{aligned} a'_{(i,j,y_1,y_2)} &= \int_0^\infty \left(\phi_1(y) - e^{-\int_0^y(\lambda_i(y_1+t)+\lambda_j(y_2+t))dt}\,[(\lambda_{ij}(y_1+y) + \lambda_{ji}(y_2+y))]\right)dy \qquad (2.6.4) \\ &= 1 - \int_0^\infty e^{-\int_0^y(\lambda_i(y_1+t)+\lambda_j(y_2+t))dt}(\lambda_{ij}(y_1+y) + \lambda_{ji}(y_2+y))dy \end{aligned}$$

as $\int_0^\infty \phi_1(y)dy = 1$. Hence $1 - a'_{(i,j,y_1,y_2)}$, the probability of meeting of X^1 and X^2 in the next transition has the desired expression. □

Definition 2.6.5. *Let $Z^1 = (X^1, Y^1)$ and $Z^2 = (X^2, Y^2)$ be the strong solutions of* (2.2.4)-(2.2.5) *with two different sets of initial conditions. Let $\mathcal{P}(k, y)$ denote the regular conditional probability of merging of X^1 and X^2 at a meeting time given meeting occurred in finite time, k is the meeting state, and y is the age of the chain which arrives at k prior to the meeting time.*

Theorem 2.6.6. *Assume (A3). Let $Z^1 = (X^1, Y^1)$ and $Z^2 = (X^2, Y^2)$ be the strong solutions of* (2.2.4)-(2.2.5) *with two different sets of initial conditions. If at a time instant t', $X^1_{t'} = X^2_{t'} = k \in \mathcal{X}$ with $Y^1_{t'} \wedge Y^2_{t'} = 0$ and $Y^1_{t'} \vee Y^2_{t'} = y > 0$, then the probability $\mathcal{P}(k, y)$ that X^1, and X^2 are merged at t' is given by,*

$$\mathcal{P}(k, y) = \int_0^\infty e^{-\int_0^{y'} \sum\limits_{k' \neq k} (\lambda_{kk'}(t) \vee \lambda_{kk'}(y+t))dt} \left[\sum_{k' \neq k} \lambda_{kk'}(y') \wedge \lambda_{kk'}(y + y') \right] dy' \tag{2.6.5}$$

where $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$.

Proof. Let t' denote a meeting time of X^1 and X^2. It is given that t' is finite, with $X^1_{t'} = k = X^2_{t'}$, $Y^1_{t'} \wedge Y^2_{t'} = 0$ and $Y^1_{t'} \vee Y^2_{t'} = y > 0$. Let ϑ denote the duration both the processes stay at k before either of them transit to some other state. Clearly, the event of no transition of X^1 and X^2 for next y' unit of time after t', is equivalent to the event where no Poisson point mass belongs to the set $B := \bigcup\limits_{t \in [0, y')} \left(\{t' + t\} \times \bigcup\limits_{k' \notin \{k\}} (\Lambda_{kk'}(t) \cup \Lambda_{kk'}(y + t)) \right)$. Evidently, this event occurs with probability $e^{-m_2(B)}$, as m_2 is the intensity of the Poisson random measure. Since, simultaneous occurrence of this event and the event of a Poisson point mass lying on the line segment $\{t' + y'\} \times \bigcup\limits_{k' \notin \{k\}} (\Lambda_{kk'}(y') \cup \Lambda_{kk'}(y + y'))$ is equivalent to the occurrence of $\{\vartheta = y'\}$, the expression of conditional density η_ϑ of ϑ is given by $\eta_\vartheta(y') = e^{-m_2(B)} m_1(\bigcup\limits_{k' \notin \{k\}} A_{k'}(y'))$, where $A_{k'}(y') := \Lambda_{kk'}(y') \cup \Lambda_{kk'}(y + y')$ for every $k' \neq k$. As $\tilde{\lambda}_{i'j'}(y)$ is set as constant $\|\lambda_{i'j'}\|_\infty$ for almost every y (Assumption (A3)), due to the definitions of the intervals in (2.2.1), for almost every $y \geq 0$ and $y' \geq 0$ the collection $\{A_{k'}(y')\}_{k' \in \mathcal{X} \setminus \{k\}}$ is disjoint. Moreover, due to (A3) the left end points of the intervals $\Lambda_{kk'}(y')$ and $\Lambda_{kk'}(y+y')$ are common (see (2.2.1)). Thus the Lebesgue measures of $\Lambda_{kk'}(y') \cup \Lambda_{kk'}(y+y')$ and $\Lambda_{kk'}(y') \cap \Lambda_{kk'}(y + y')$ are $\lambda_{kk'}(y') \vee \lambda_{kk'}(y + y')$ and $\lambda_{kk'}(y') \wedge \lambda_{kk'}(y + y')$ respectively. Thus $\eta_\vartheta(y') = \exp(-\int_0^{y'} \sum\limits_{k' \neq k} (\lambda_{kk'}(t) \vee \lambda_{kk'}(y + t))dt) \sum_{k' \neq k} \lambda_{kk'}(y') \vee \lambda_{kk'}(y + y')$.

We consider two cases regarding the transition of X^1 and X^2, at $t' + \vartheta$ which are (i) simultaneous, and (ii) non-simultaneous. Case (ii) implies that X^1 and X^2 will depart in the next transition. So, under case (ii), t' is not a merging time. Consequently, case (i) is necessary for t' to be the merging time. We show that case (i) is a sufficient condition too. We recall that at $t' + \vartheta$ the Poisson point mass (which is responsible for the transition) lies in only one of the members of the disjoint family $\{A_{k'}(\vartheta)\}_{k'}$ with probability one. Therefore, under case (i), at $t' + \vartheta$, X^1 and X^2 enter into an identical state and the ages Y^1 and Y^2

become zero and therefore, the uniqueness of the SDE (2.2.4)-(2.2.5) implies merging at time t'.

For case (i) to occur, the point mass must lie in $\{t'+\vartheta\} \times \cup_{k'\notin\{k\}} \Lambda_{kk'}(\vartheta) \cap \Lambda_{kk'}(y+\vartheta)$. On the other hand if the point mass lies in $\{t'+\vartheta\} \times \cup_{k'\notin\{k\}} (A_{k'}(\vartheta) \setminus \Lambda_{kk'}(\vartheta) \cap \Lambda_{kk'}(y+\vartheta))$, then the transition at $t'+\vartheta$ is of case (ii). Hence, for almost every y and y', the conditional probability of merging given $\vartheta = y'$ is equal to $\frac{\sum_{k'\neq k} \lambda_{kk'}(y') \wedge \lambda_{kk'}(y+y')}{\sum\limits_{k'\neq k} (\lambda_{kk'}(y') \vee \lambda_{kk'}(y+y'))}$. Thus

$$P\left(X_t^1 = X_t^2, \forall t \geq t' \mid \{t' < \infty\}, X_{t'}^1 = X_{t'}^2 = k, Y_{t'}^1 \wedge Y_{t'}^2 = 0, Y_{t'}^1 \vee Y_{t'}^2 = y\right)$$

is equal to

$$\int_0^\infty P\left(X_t^1 = X_t^2, \forall t \geq t' \mid \{t' < \infty\}, X_{t'}^1 = X_{t'}^2 = k, Y_{t'}^1 \wedge Y_{t'}^2 = 0, Y_{t'}^1 \vee Y_{t'}^2 = y, \vartheta = y'\right) \eta_\vartheta(y') dy'$$
$$= \int_0^\infty e^{-\int_0^{y'} \sum\limits_{k'\neq k} (\lambda_{kk'}(t) \vee \lambda_{kk'}(y+t)) dt} \left(\sum_{k'\neq k} \lambda_{kk'}(y') \wedge \lambda_{kk'}(y+y')\right) dy'.$$

This completes the proof. □

Remark 2.6.7. *It is interesting to note that for Markov special case, where the transition rate matrix λ is independent of the age variable y, a direct calculation gives that $\mathcal{P}(k,y) = 1$. This makes Theorem 2.6.2, a corollary of the above theorem. On the other hand by considering the two-state semi-Markov chain given in Example 2.6.3, one can obtain for each $k = 1,2$, $\lim_{y\to\infty} \mathcal{P}(k,y) = \int_0^\infty e^{-y'} \frac{y'}{1+y'} dy' < \frac{1}{2}\int_0^1 e^{-y'} dy' + \int_1^\infty e^{-y'} dy' = \frac{1+e^{-1}}{2} < 1$. This further clarifies that a meeting time for the flow in Example 2.6.3 need not be a merging time. Below we show that the chance of merging for a general semi-Markov chain increases to 1 as y decreases to zero, provided that the transition rate is continuous at zero.*

Proposition 2.6.8. *Assume (A3) and that $y \mapsto \lambda(y)$ is continuous at zero. As y tends to zero, $\mathcal{P}(k,y)$ converges to 1.*

Proof. From Theorem 2.6.6

$$\lim_{y\to 0} \mathcal{P}(k,y) = \lim_{y\to 0} \int_0^\infty e^{-\int_0^{y'} \sum\limits_{k'\neq k} (\lambda_{kk'}(t) \vee \lambda_{kk'}(y+t)) dt} \left(\sum_{k'\neq k} \lambda_{kk'}(y') \wedge \lambda_{kk'}(y+y')\right) dy'.$$

Due to the continuity of λ in y, the integrand converges pointwise to

$$\psi(y') := \sum_{k'\neq k} \lambda_{kk'}(y') \exp\left(-\int_0^{y'} \sum_{k'\neq k} \lambda_{kk'}(t) dt\right).$$

The integrand is also uniformly dominated by ψ, which is integrable on $[0,\infty)$. Indeed $\int_{[0,\infty)} \psi(y') dy' = 1$. Thus the result follows using dominated convergence theorem. □

2.7 Eventual Meeting, Merging, and Time

It is important to note that the strict positivity of entries of the rate matrix, as assumed in this chapter, implies irreducibility of the process. It is also known that mere irreducibility of a Markov chain does not ensure the convergence. However, the meeting event of two chains may take place even if the chains do not converge. The discrete time Markov chain on two states having zero probability of transition to the same state constitutes an example where chains with different initial states never meet due to its periodicity. Nevertheless, the same phenomena is untrue for its continuous time version. Indeed, if two such chains (Markov/semi-Markov), having bounded transition rate and driven by the same noise (the Poisson random measure) start from two different states, they meet surely at the next transition. In this chapter, due to the consideration of processes having bounded transition rate, the discrete time scenario is excluded. Thus an ergodicity assumption is not needed for assuring eventual meeting. The next theorem establishes eventual meeting of Markov special case under finiteness assumption of the state space.

Theorem 2.7.1. *Let X^1 and X^2 be as in Theorem 2.6.2.*

1. *The conditional probability of meeting in the next transition given $\mathcal{F}_0$ is $\frac{(\lambda_{ij}+\lambda_{ji})}{\lambda_i+\lambda_j}$.*
2. *If $\mathcal{X}$ is finite, X^1 and X^2 eventually meet with probability 1.*

Proof. Recall $X^1_0 = i$, $X^2_0 = j$, and the sequence $\{T_n\}$ from Notation 2.5.1. By applying Theorem 2.6.4 for the Markov special case, we can write the conditional probability of meeting in the next transition of X^1 and X^2, given the initial conditions as

$$\int_0^\infty e^{-\int_0^y(\lambda_i+\lambda_j)dt}(\lambda_{ij}+\lambda_{ji})dy = \frac{(\lambda_{ij}+\lambda_{ji})}{(\lambda_i+\lambda_j)}\int_0^\infty e^{-(\lambda_i+\lambda_j)y}(\lambda_i+\lambda_j)dy = \frac{(\lambda_{ij}+\lambda_{ji})}{(\lambda_i+\lambda_j)}.$$

Hence the part (1) is proved. Since, $\lambda_{ij} > 0$ for all $i \neq j$ and $\mathcal{X}$ is finite, $\min_{i,j}\frac{(\lambda_{ij}+\lambda_{ji})}{\lambda_i+\lambda_j} > 0$. Thus $\max_{i,j} a_{(i,j)} < 1$ where $a_{(i,j)}$ denotes the probability of not meeting in the next transition. Now since $\{T_n\}$ is a sequence of stopping times, using Theorem 2.3.4(1), we get

$$E\left[\mathbb{1}_{\{X^1_{T_n}\neq X^2_{T_n}\}} \mid \mathcal{F}_{T_{n-1}}\right] = a_{(X^1_{T_{n-1}},X^2_{T_{n-1}})} \le \max_{i,j} a_{(i,j)} < 1. \tag{2.7.1}$$

The event of never meeting of processes X^1 and X^2 is identical to the repeated occurrence of $\{X^1_{T_n} \neq X^2_{T_n}\}$ for all $n \ge 1$. Hence, using the fact (thanks to (A2)) that the chains experience infinitely many transitions with probability 1, the probability of never meeting, $P\left(X^1_t \neq X^2_t, \forall t \ge 0 \mid \mathcal{F}_0\right)$ matches with $\lim_{N\to\infty} P\left(\cap_{n=1}^N \{X^1_{T_n} \neq X^2_{T_n}\} \mid \mathcal{F}_0\right)$. Next if

$$E\left[\prod_{n=1}^N \mathbb{1}_{\{X^1_{T_n}\neq X^2_{T_n}\}} \mid \mathcal{F}_0\right] \le \max_{i,j} a_{(i,j)} E\left[\prod_{n=1}^{N-1} \mathbb{1}_{\{X^1_{T_n}\neq X^2_{T_n}\}} \mid \mathcal{F}_0\right], \tag{2.7.2}$$

holds for all $N \geq 1$, using that repeatedly, we get

$$P\left(\cap_{n=1}^{N}\{X_{T_n}^1 \neq X_{T_n}^2\} \mid \mathcal{F}_0\right) = E\left[\prod_{n=1}^{N} \mathbb{1}_{\{X_{T_n}^1 \neq X_{T_n}^2\}} \mid \mathcal{F}_0\right] \leq \left(\max_{i,j} a_{(i,j)}\right)^N$$

for all $N \geq 1$. The right side clearly vanishes as $N \to \infty$, and thus $P\left(X_t^1 \neq X_t^2, \forall t \geq 0 \mid \mathcal{F}_0\right)$ is zero as desired, provided (2.7.2) holds. Finally (2.7.2) is shown using (2.7.1) below

$$\begin{aligned} E\left[\prod_{n=1}^{N} \mathbb{1}_{\{X_{T_n}^1 \neq X_{T_n}^2\}} \mid \mathcal{F}_0\right] &= E\left[E\left[\prod_{n=1}^{N} \mathbb{1}_{\{X_{T_n}^1 \neq X_{T_n}^2\}} \mid \mathcal{F}_{T_{N-1}}\right] \mid \mathcal{F}_0\right] \\ &= E\left[\left(\prod_{n=1}^{N-1} \mathbb{1}_{\{X_{T_n}^1 \neq X_{T_n}^2\}}\right) E\left[\mathbb{1}_{\{X_{T_N}^1 \neq X_{T_N}^2\}} \mid \mathcal{F}_{T_{N-1}}\right] \mid \mathcal{F}_0\right] \\ &\leq \max_{i,j} a_{(i,j)} E\left[\prod_{n=1}^{N-1} \mathbb{1}_{\{X_{T_n}^1 \neq X_{T_n}^2\}} \mid \mathcal{F}_0\right] \end{aligned}$$

for all $N \geq 1$. Hence the proof of part(2) is complete. □

In the above proof, the second part of the theorem has been proved using the first part. However, the former has been proved in Lemma 3.5 of [4] without utilizing part 1, under identical assumption in a different approach.

It is important to note that for ensuring almost sure eventual meeting, we have assumed finiteness of $\mathcal{X}$ in the above theorem, whereas in the proof we have used $\min_{i,j} \frac{(\lambda_{ij}+\lambda_{ji})}{\lambda_i+\lambda_j} > 0$ only. In the following lemma we show that under (A1), these conditions are equivalent.

Lemma 2.7.2. *Let λ be a transition rate matrix of a Markov chain obeying (A1). If $\mathcal{X}$ is infinite, $\inf_{i,j} \frac{(\lambda_{ij}+\lambda_{ji})}{\lambda_i+\lambda_j}$ is zero.*

Proof. Fix a $j \in \mathcal{X}$. Since, due to Assumption (A1), $\sum_{i=1}^{\infty} \lambda_i < \infty$, given $\epsilon > 0$ there exists an $i_{\epsilon,j}$ such that $\lambda_i < \epsilon\lambda_j \; \forall i > i_{\epsilon,j}$. So we get an inequality $\lambda_{ij} \leq \lambda_i < \epsilon\lambda_j$ for all $i > i_{\epsilon,j}$. Using this inequality we have the following relation,

$$\frac{\lambda_{ij}+\lambda_{ji}}{\lambda_i+\lambda_j} < \frac{\epsilon\lambda_j+\lambda_{ji}}{\lambda_i+\lambda_j} < \frac{\epsilon\lambda_j+\lambda_{ji}}{\lambda_j} = \epsilon + \frac{\lambda_{ji}}{\lambda_j}, \tag{2.7.3}$$

for all $i > i_{\epsilon,j}$. For each j we also have $\lambda_j = \sum_{i \in \mathcal{X}\setminus\{j\}} \lambda_{ji} < \infty$. Hence, there exists a $i_{j,\epsilon}^*$ such that for all $i > i_{j,\epsilon}^*$ we have $\lambda_{ji} < \epsilon\lambda_j$. Now, using (2.7.3), we get for each $i > \max(i_{\epsilon,j}, i_{j,\epsilon}^*)$,

$$\frac{\lambda_{ij}+\lambda_{ji}}{\lambda_i+\lambda_j} < \epsilon + \frac{\epsilon\lambda_j}{\lambda_j} = 2\epsilon.$$

Since ϵ is arbitrary, the above implies that for each $j \in \mathcal{X}$,

$$\lim_{i\to\infty} \frac{\lambda_{ij} + \lambda_{ji}}{\lambda_i + \lambda_j} = 0. \tag{2.7.4}$$

Similarly by interchanging the roles of i and j in the above argument, one obtains

$$\lim_{j\to\infty} \frac{\lambda_{ij} + \lambda_{ji}}{\lambda_i + \lambda_j} = 0 \tag{2.7.5}$$

for each $i \in \mathcal{X}$. Hence from (2.7.4), and (2.7.5), we conclude $\inf\limits_{i,j}(\frac{\lambda_{ij}+\lambda_{ji}}{\lambda_i+\lambda_j}) = 0$. □

Next we wish to investigate the eventual meeting event for semi-Markov family. Clearly, in view of Theorem 2.7.1(2), a condition like $\inf_{(i,j)\in\mathcal{X}_2,y_1,y_2,y} \frac{(\lambda_{ij}(y_1+y)+\lambda_{ji}(y_2+y))}{\lambda_i(y_1+y)+\lambda_j(y_2+y)} > 0$ is needed for this purpose. However, finiteness of $\mathcal{X}$ is not enough to ensure that. We consider the following assumption.

(A4) $\mathcal{X}$ is finite and $\sup_{(i,j)\in\mathcal{X}_2,y_1\geq 0,y_2\geq 0} \left\| 1 - \frac{(\lambda_{ij}(y_1+\cdot)+\lambda_{ji}(y_2+\cdot))}{\lambda_i(y_1+\cdot)+\lambda_j(y_2+\cdot)} \right\|_{L^\infty} < 1.$

Theorem 2.7.3. *Assume (A1)-(A4) and that X^1 and X^2 are as in Notation 2.5.1. Then X^1 and X^2 eventually meet with probability 1.*

Proof. Using $\phi_2(y) := 1 - \frac{(\lambda_{ij}(y_1+y)+\lambda_{ji}(y_2+y))}{\lambda_i(y_1+y)+\lambda_j(y_2+y)}$, we rewrite (2.6.4) as

$$a'_{(i,j,y_1,y_2)} = \int_0^\infty \phi_1(y)\phi_2(y)dy \leq \|\phi_1\|_{L^1}\|\phi_2\|_{L^\infty} = \|\phi_2\|_{L^\infty}.$$

Now by a direct application of (A4), we get that supremum of $\|\phi_2\|_{L^\infty}$ over all $(i,j) \in \mathcal{X}_2, y_1 \geq 0, y_2 \geq 0$ is less than 1, which implies that

$$\sup_{(i,j)\in\mathcal{X}_2,y_1\geq 0,y_2\geq 0} a'_{(i,j,y_1,y_2)} < 1. \tag{2.7.6}$$

Again as in the proof of Theorem 2.7.1, the total probability of never meeting is the probability of intersection of occurrence of not meeting in next transition for every transition, and (A2) ensures almost sure infinite transitions. Moreover, since (Z^1, Z^2) is strong Markov (Theorem 2.3.4(1)) and $\{T_n\}_{n\geq 1}$ are stopping times $P\left(\{X^1_{T_n} \neq X^2_{T_n}\} \mid \mathcal{F}_{T_{n-1}}\right) = a'_{(X^1_{T_{n-1}},X^2_{T_{n-1}},Y^1_{T_{n-1}},Y^2_{T_{n-1}})}$ which is not more than the left side of (2.7.6). Therefore, in the similar line of the proof of Theorem 2.7.1, we get

$$E\left[\prod_{n=1}^N \mathbb{1}_{\{X^1_{T_n}\neq X^2_{T_n}\}} \mid \mathcal{F}_0\right] \leq \left(\sup_{(i,j)\in\mathcal{X}_2,y_1\geq 0,y_2\geq 0} a'_{(i,j,y_1,y_2)}\right)^N \tag{2.7.7}$$

and $P\left(X^1_t \neq X^2_t, \forall t \geq 0 \mid \mathcal{F}_0\right) = \lim_{N\to\infty} E\left[\prod_{n=1}^N \mathbb{1}_{\{X^1_{T_n}\neq X^2_{T_n}\}} \mid \mathcal{F}_0\right]$. This limit is zero from (2.7.6) and (2.7.7). Thus the probability of never meeting is zero.

Under (A1)-(A4), the pair (X^1, X^2) not only surely meet, the expected number of transitions needed for meeting is also finite. A rather stronger result is shown below.

Theorem 2.7.4. *Assume (A1)-(A4) and that X^1 and X^2 are as in Notation 2.5.1. If N denotes the number of collective transitions until the first meeting time of X^1 and X^2, then $E[N^r] < \infty$ for any $r \geq 1$.*

Proof. For the sake of brevity, we write $a'_{(Z^1_{\mathcal{T}_n}, Z^2_{\mathcal{T}_n})}$ for $a'_{(X^1_{\mathcal{T}_n}, X^2_{\mathcal{T}_n}, Y^1_{\mathcal{T}_n}, Y^2_{\mathcal{T}_n})}$, a notation that appears in the proof of Theorem 2.6.4. Since N denotes the number of collective transitions until the first meeting time, using the above notation and (2.7.7), we get for all $n \geq 0$

$$\begin{aligned} P(N = n+1) =& E\left[\left(\prod_{r=1}^{n} \mathbb{1}_{\{X^1_{\mathcal{T}_r} \neq X^2_{\mathcal{T}_r}\}}\right) E\left(\mathbb{1}_{\{X^1_{\mathcal{T}_{n+1}} = X^2_{\mathcal{T}_{n+1}}\}} \mid \mathcal{F}_{\mathcal{T}_n}\right)\right] \\ \leq& \left(\sup_{z^1,z^2} a'_{(z^1,z^2)}\right)^n \left(1 - \inf_{z^1,z^2} a'_{(z^1,z^2)}\right), \end{aligned}$$

by following the convention that product and intersection of an empty family are 1 and empty set respectively. Thus the r^{th} raw moment, $E[N^r]$ is

$$\sum_{n=1}^{\infty} n^r P(N = n) \leq \left(1 - \inf_{z^1,z^2} a'_{(z^1,z^2)}\right) \left(\sup_{z^1,z^2} a'_{(z^1,z^2)}\right)^{-1} \sum_{n=1}^{\infty} n^r \left(\sup_{z^1,z^2} a'_{(z^1,z^2)}\right)^n .$$

The infinite series on the right converges provided $\sup_{z^1,z^2} a'_{(z^1,z^2)} < 1$ which is ensured in (2.7.6) due to the assumption (A4). To be more precise, that series sum is expressed as $Li_{-r}(\sup_{z^1,z^2} a'_{(z^1,z^2)})$ where $Li_r(z)$ is polylogarithm function of order r and with argument z. Thus we conclude that N has finite moments. $\square$

We end this section with the final result below. That requires essential infimum of at least one entry of each row of λ to be nonzero.

Theorem 2.7.5. *Assume (A1)-(A4) and that X^1 and X^2 are as in Notation 2.5.1. Further assume that for each $k \in \mathcal{X}$ there is at least one $k' \in \mathcal{X} \setminus \{k\}$ such that $\|\lambda_{kk'}{}^{-1}\|_{L^\infty} < \infty$. Then X^1 and X^2 eventually merge with probability 1.*

Proof. Since (A1)-(A4) hold, Theorem 2.7.3 ensures eventual meeting with probability 1. Hence $\mathcal{T}_1 < \infty$ with probability 1, where $\mathcal{T}_1$ denotes the first meeting time (see Definition 2.6.1). If $\mathcal{T}_1$ is NMT (not a merging time), X^1 and X^2 separate at the next transition and again due to Theorem 2.7.3, they meet at $\mathcal{T}_2$, say, which is again finite almost surely. By repeating this argument, if X^1 and X^2 never merge, we obtain an infinite sequence $\{\mathcal{T}_n\}_n$ of meeting times where each of them are finite almost surely. Using $k_n := X^1_{\mathcal{T}_n} = X^2_{\mathcal{T}_n}$ and $y_n :=$

$\max(Y^1_{\mathcal{T}_n}, Y^2_{\mathcal{T}_n})$, we get $P\left(\{\mathcal{T}_n \text{ is NMT}\} \mid \mathcal{F}_{\mathcal{T}_n}\right) = P\left(\{\mathcal{T}_n \text{ is NMT}\} \mid k_n, y_n\right) = 1 - \mathcal{P}(k_n, y_n)$, since $\{\mathcal{T}_n\}_{n\geq 1}$ is a sequence of stopping times and (Z^1, Z^2) is strong Markov. Therefore,

$$\begin{aligned} E\left[\prod_{n=1}^{N} \mathbb{1}_{\{\mathcal{T}_n \text{ is NMT}\}}\right] =& E\left[\prod_{n=1}^{N-1} \mathbb{1}_{\{\mathcal{T}_n \text{ is NMT}\}} E\left[\mathbb{1}_{\{\mathcal{T}_N \text{ is NMT}\}} \mid \mathcal{F}_{\mathcal{T}_N}\right]\right] \\ \leq& \left(1 - \inf_{k\in\mathcal{X}, y\geq 0} \mathcal{P}(k,y)\right) E\left[\prod_{n=1}^{N-1} \mathbb{1}_{\{\mathcal{T}_n \text{ is NMT}\}}\right]. \end{aligned}$$

Since the event of never merging can be expressed as $\cap_{n\geq 1}\{\mathcal{T}_n \text{ is NMT}\}$, an upper bounded of its probability can be obtained by using the above inequality repeatedly, i.e.,

$$P\left(\mathcal{T}_n \text{ is NMT}, \forall n \geq 1\right) = \lim_{N\to\infty} E\left[\prod_{n=1}^{N} \mathbb{1}_{\{\mathcal{T}_n \text{ is NMT}\}}\right] \leq \lim_{N\to\infty} \left(1 - \inf_{k\in\mathcal{X}, y\geq 0} \mathcal{P}(k,y)\right)^N. \tag{2.7.8}$$

This confirms that the probability of never merging is zero, provided $\inf_{k\in\mathcal{X}, y\geq 0} \mathcal{P}(k,y) > 0$. Since, (A3) holds, from Theorem 2.6.6,

$$\begin{aligned} \mathcal{P}(k,y) \geq& \int_0^\infty e^{-\sum_{k'\neq k} \|\lambda_{kk'}\|_\infty y'} \sum_{k'\neq k} \|{\lambda_{kk'}}^{-1}\|_\infty^{-1} dy' \\ \text{or,} \inf_{k\in\mathcal{X}, y\geq 0} \mathcal{P}(k,y) \geq& \left(\int_0^\infty e^{-Cy'} dy'\right) \min_{k\in\mathcal{X}} \sum_{k'\neq k} \|{\lambda_{kk'}}^{-1}\|_\infty^{-1}. \end{aligned}$$

Since for each $k \in \mathcal{X}$, there is a $k' \in \mathcal{X} \setminus \{k\}$ such that $\|{\lambda_{kk'}}^{-1}\|_{L^\infty} < \infty$, and $\mathcal{X}$ is finite, the right side of above inequality is positive. Thus $\inf_{k\in\mathcal{X}, y\geq 0} \mathcal{P}(k,y) > 0$ as desired. □

2.8 Conclusion

In this chapter we make use of a particular type of semimartingale representation of a class of semi-Markov processes. We have then studied various aspects of a pair of solutions having two different initial conditions. Several questions regarding the meeting and merging of stochastic flow of SMP have been answered by considering a solution pair. We have obtained explicit expressions of probabilities of many relevant events in terms of the transition rate matrix.

The study of eventual meeting and merging in Section 7 is carried out for finite state-space case. These results could be examined for certain infinite state cases, like birth-death processes, or more generally, where all entries of λ, except k nearest neighbours of diagonal are zero. The present study which has been carried out for the time-homogeneous case, is worth investigating for the time non-homogeneous case. We study that in the next chapter.

Chapter 3

Non-Homogeneous semi-Markov processes

3.1 Introduction

In this chapter also we confine ourselves to the study of SMPs on a countable state space. The continuous-time discrete-state non-homogeneous Markov chains are included in this class. The class we consider in this chapter is such that the embedded discrete time Markov chain is allowed to be non-homogeneous. However, a more general non-homogeneity that appears in a continuous time Markov process having time dependent transition rate is not included. SMP beyond the class of time homogeneity has been first studied in [18] and [28]. Various different aspects and generalisations of non-homogeneous SMP(NHSMP) has further been explored in [20, 22, 21, 23]. In these references several applications of NHSMP has also been emphasised.

The non-homogeneous SMP, augmented with the age and transition count processes is represented as semi-martingales using a system of stochastic integral equations involving a Poisson random measure. The coefficients of the equation depend on a given transition rate function and an additional gaping parameter. It is worth noting that neither the coefficients are compactly supported nor the intensity measure of the PRM is finite. Note that, compactly supported coefficient or finiteness of intensity measure are the standard assumptions under which an SDE involving PRM is studied commonly. So, we produce a self-contained proof of the existence and uniqueness of the solution to the SDE in this chapter. This extends the results obtained in Section 2.2. Subsequently, we extend the results presented in Sections 2.3-2.5 by showing that the state component of the solution is a pure non-homogeneous SMP with the given non-homogeneous transition rate function. We also derive the law of the bivariate process obtained from two solutions of the equation having two different initial conditions.

In the study of stochastic flow on a discrete state space, meeting and merging events are of significant importance. Investigation of meeting and merging of homogeneous semi-Markov processes (SMP) appears in the preceding chapter, which is not an off-the-shelf problem. In contrast to the traditional description of a semi-Markov chain in terms of a renewal process, a semimartingale representation (see [11, 13, 14] for more details) has been crucially used in that chapter. In the present chapter, we extend many results of sixth and seventh sections of earlier chapter for a larger class of SMPs where the embedded Markov chain may not be homogeneous. Following Chapter 2, with the help of an additional gaping parameter, we have obtained explicit formulae for probabilities of various meeting and merging related events of a generalized semi-Markov stochastic flow.

The study of meeting and coalescence of stochastic processes is an active branch of probability theory. Some of the earliest instances of such study dates back to Arratia [1], and Harris [17] where they have considered merging of one dimensional Brownian flow. On the other hand mixing for a class of non Markov flows have been investigated by Melbourne and Terhesiu [29]. However, to the best of our knowledge, questions regarding meeting and merging have not been addressed in the literature for stochastic flow of non-homogeneous SMPs. In [4] for the stability analysis of Markov modulated diffusions, the merging of Markov chains has been crucially used. In view of this, we believe that the study of meeting and merging of multiple semi-Markov particles are relevant for investigating stability properties of a diffusion that is modulated by semi-Markov processes.

It is shown in Chapter 2 that for a homogeneous semi-Markov flow, unlike the Markov counterpart, a pair of solutions may not merge when they meet. Moreover, in general the merging time is not a stopping time, although the meeting time is. The study becomes even more involved in the absence of time homogeneity. In this chapter we explain this with an example of a non-homogeneous Markov flow on binary state-space, where a pair of chains meet almost surely but never merge. The absence of homogeneity for the semi-Markov flow ramifies the analysis further.

We derive an expression of the conditional probability of meeting in the next transition. Using the notion of coherent meeting, where both of the chains have identical transition count at the meeting time, we obtain a lower bound of probability of merging in the next transition. We further provide with a sufficient condition under which should a meeting is non coherent, is not a merging. Moreover, we show that under such circumstances, the lower bounds are the actual merging probabilities. A set of sufficient conditions are also obtained under which a pair of SMPs eventually meet or merge with probability one. The number of transitions required to encounter a meeting is shown to have all moments finite. Many numerical examples are considered which highlight the intricacies and implications of the theoretical results. In one of the numerical examples the distribution of time of first meeting and merging are obtained and are compared with that of the holding times, for the purpose of illustration.

3.2 Semimartingle Representation for Non-homogeneous Semi-Markov Processes

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be the underlying filtered probability space satisfying the usual hypothesis and $\mathcal{X} = \{1, 2, \ldots\}$ the state space. We wish to construct a semi-Markov process on this state space with a given transition rate function. To this end we first embed $\mathcal{X}$ in $\mathbb{R}$ and endow with a total order $\prec_1$, which in turn induces a total order $\prec_2$ on $\mathcal{X}_2 := \{(i,j) \in \mathcal{X}^2 \mid i \neq j\}$ by following lexicographic order. For each $y \in \mathbb{R}_+$, $n \in \mathbb{N}_0$, $\lambda(y,n) := (\lambda_{ij}(y,n))$ denote a matrix in which the i^{th} diagonal element is negative of $\lambda_i(y,n) := \sum_{j\in\mathcal{X}\setminus\{i\}} \lambda_{ij}(y,n)$ and for each $(i,j) \in \mathcal{X}_2$, $\lambda_{ij} \colon \mathbb{R}_+ \times \mathbb{N}_0 \to \mathbb{R}_+$ is a bounded measurable function satisfying the following two assumptions.

(B1). If $c_i := \sum\limits_{j\in\mathcal{X}\setminus\{i\}} \|\lambda_{ij}(\cdot,\cdot)\|_{L^\infty_{(\mathbb{R}_+\times\mathbb{N}_0)}}$, $c := \sup_i c_i < \infty$.

(B2). For each n in $\mathbb{N}_0$, and i in $\mathcal{X}$, $\lim_{y\to\infty} \gamma_i(y,n) = \infty$, where $\gamma_i(y,n) := \int_0^y \lambda_i(y',n)dy'$.

For each $(i,j) \in \mathcal{X}_2$, we consider another measurable function $\tilde{\lambda}_{ij} \colon \mathbb{R}_+ \times \mathbb{N}_0 \to \mathbb{R}_+$ such that for each $y \in \mathbb{R}_+$, and $n \in \mathbb{N}_0$

$$\lambda_{ij}(y,n) \leq \tilde{\lambda}_{ij}(y,n), \tag{3.2.1}$$

and also for almost every $y \in \mathbb{R}_+$ and $n \in \mathbb{N}_0$

$$\tilde{\lambda}_{ij}(y,n) \leq \|\lambda_{ij}(\cdot,\cdot)\|_{L^\infty_{(\mathbb{R}_+\times\mathbb{N}_0)}}. \tag{3.2.2}$$

Now for each $y \in \mathbb{R}_+$, and $n \in \mathbb{N}_0$, with the help of $\lambda(y,n)$ and $\tilde{\lambda}(y,n) := (\tilde{\lambda}_{ij}(y,n))$, we introduce a disjoint collection of intervals $\Lambda := \{\Lambda_{ij}(y,n) \colon (i,j) \in \mathcal{X}_2\}$, by

$$\Lambda_{ij}(y,n) = \left(\sum_{(i',j')\prec_2(i,j)} \tilde{\lambda}_{i'j'}(y,n) \right) + \Big[0, \lambda_{ij}(y,n)\Big) \tag{3.2.3}$$

where $a + B = \{a + b \mid b \in B\}$ for $a \in \mathbb{R}, B \subset \mathbb{R}$. Clearly, for each y and n, the interval $\Lambda_{ij}(y,n)$ is of length $\lambda_{ij}(y,n)$. If $C_i := c_i + \sum_{k\prec_1 i} c_k$, according to (B1) and (3.2.2), the set $\Lambda_i(y,n) := \cup_{j\in\mathcal{X}\setminus\{i\}} \Lambda_{ij}(y,n)$ is contained in the finite interval $[0, C_i]$, for almost every $y \in \mathbb{R}_+$ and each n. Using the above intervals we define $h_\Lambda \colon \mathcal{X} \times \mathbb{R}_+ \times \mathbb{N}_0 \times \mathbb{R} \to \mathbb{R}$ as

$$h_\Lambda(i,y,n,v) := \sum_{j\in\mathcal{X}\setminus\{i\}} (j-i)\mathbb{1}_{\Lambda_{ij}(y,n)}(v) \tag{3.2.4}$$

and $g_\Lambda \colon \mathcal{X} \times \mathbb{R}_+ \times \mathbb{N}_0 \times \mathbb{R} \to \mathbb{R}$ as

$$g_\Lambda(i,y,n,v) := \sum_{j\in\mathcal{X}\setminus\{i\}} \mathbb{1}_{\Lambda_{ij}(y,n)}(v). \tag{3.2.5}$$

These functions are piece-wise constant in v variable. Using these, we consider the following system of coupled stochastic integral equations in X, Y and N

$$X_t = X_0 + \int_{0+}^{t} \int_{\mathbb{R}} h_\Lambda(X_{u-}, Y_{u-}, N_{u-}, v) \wp(du, dv) \tag{3.2.6}$$

$$Y_t = Y_0 + t - \int_{0+}^{t} (Y_{u-}) \int_{\mathbb{R}} g_\Lambda(X_{u-}, Y_{u-}, N_{u-}, v) \wp(du, dv) \tag{3.2.7}$$

$$N_t = \int_{0+}^{t} \int_{\mathbb{R}} g_\Lambda(X_{u-}, Y_{u-}, N_{u-}, v) \wp(du, dv) \tag{3.2.8}$$

for $t \geq 0$, where $\int_{0+}^{t}$ is integration over the interval $(0, t]$, $\wp(du,\ dv)$ is a PRM on $\mathbb{R}_+ \times \mathbb{R}$ with intensity $m_2(du, dv)$, and defined on the probability space $(\Omega, \mathcal{F}, P)$. We also assume that $\{\wp((0, t] \times dv)\}_{t \geq 0}$ is adapted to $\{\mathcal{F}_t\}_{t \geq 0}$. For $t \geq 0$, we rewrite the above coupled integral equation (3.2.6)-(3.2.8) as a vector form given by,

$$Z_t = Z_0 + \int_0^t a(Z_{u-}) du + \int_{0+}^{t} \int_{\mathbb{R}} J(Z_{u-}, v) \wp(du, dv), \tag{3.2.9}$$

where for each $t \geq 0$, $Z_t = (X_t, Y_t, N_t)$, $N_0 = 0$, $z = (i, y, n)$, $a(z) = (0, 1, 0)$, and $J(z, v)$ is the vector $(h_\Lambda(z, v), -y g_\Lambda(z, v), g_\Lambda(z, v))$. Note that the assumption (B1) is weaker than (A1).

Remark 3.2.1. *The existing results on the general theory of SDE are not directly applicable for assuring existence and uniqueness of solution to the system of integral equations* (3.2.6)-(3.2.8)*. We recall that Theorem 3.4 (p-474) of [8] assumes compact support of the integrands whereas Theorem IV.9.1 (p-231) of [19] assumes finiteness of the intensity measure on the complement of a neighbourhood of origin. For the system* (3.2.6)-(3.2.8) *neither the integrands are compactly supported nor the intensity of Poisson random measure is finite in the complement of any bounded set. For this reason we produce an original proof of existence and uniqueness of* (3.2.6)-(3.2.8)*.*

Theorem 3.2.2. *Under assumption (B1), there exists an increasing sequence of stopping times $\{T_n\}_{n=1}^{\infty}$ such that the following hold.*

1. *The coupled system of stochastic integral equations* (3.2.6)-(3.2.8) *has a unique strong solution $(X, Y, N) = (X_t, Y_t, N_t)_{t \in [0, \tau)}$ where $\tau = \lim_{n \to \infty} T_n$.*

2. *Almost surely X, Y, and N have r.c.l.l. paths respectively and they jump only at T_n, for each n. While X and N are piece-wise constants, Y is piece-wise linear.*

3. *We set $T_0 := -Y_0$. For each $n \in \mathbb{N}$, (i) $Y_{T_n} = 0$, (ii) $Y_{T_n-} = T_n - T_{n-1}$, and (iii) $N_{T_n} = n$. Also (iv) $Y_t = t - T_{N_t}$ for all $t \in \mathbb{R}_+$.*

Proof. We recall that due to Assumption (B1) and (3.2.2), for almost every $y \in \mathbb{R}_+$, and each i, n the union of intervals $\{\Lambda_{ij}(y,n) \mid j \in \mathcal{X} \setminus \{i\}\}$ is contained in $[0, C_i]$. For each $\omega \in \Omega$ and $i \in \mathcal{X}$ we define the set $\mathcal{D}_i := \{s \in (0,\infty) \mid \wp(\omega)(\{s\} \times [0, C_i]) > 0\}$, the collection of time coordinates of the point masses of a realisation of the PRM $\wp(\omega)$ in which the second component is not more than C_i. Since, the interval $[0, C_i]$ has finite Lebesgue measure, by Lemma 2.2.1, $\mathcal{D}_i$ has no limit points in $\mathbb{R}$ with probability 1. Thus for each $i \in \mathcal{X}$ we can enumerate $\mathcal{D}_i$, in increasing order

$$\mathcal{D}_i = \{\sigma_l^i\}_{l=1}^{\infty}, \text{ where } \sigma_1^i < \cdots < \sigma_l^i < \sigma_{l+1}^i < \cdots \text{ for each } \omega. \tag{3.2.10}$$

For each $l \in \mathbb{N}, \sigma_l^i \colon \Omega \to (0,\infty]$ and $\{\sigma_l^i \leq t\} = \{\omega \mid \wp((0,t] \times [0, C_i]) \geq l\} \in \mathcal{F}_t$, as $\wp((0,t] \times [0, C_i])$ is $\mathcal{F}_t$ measurable. Hence σ_l^i is a stopping time for each $i \in \mathcal{X}$ and $l \geq 1$. As $\mathcal{D}_i$ has no limit points, $\sigma_l^i \uparrow \infty$ as $l \to \infty$ almost surely.

Here we plan to construct an increasing sequence of stopping times $\{\bar{\sigma}_m\}_{m=1}^{\infty}$ so that we can first define a solution to equations (3.2.6)-(3.2.8) on the time interval $[0,\ \bar{\sigma}_1]$ and then on the next time interval $(\bar{\sigma}_1,\ \bar{\sigma}_2]$, and so on. For a fixed $\omega \in \Omega$, X, Y at time $t = 0$ are $X_0 \in \mathcal{X}$, and $Y_0 \in \mathbb{R}_+$ respectively. Using (3.2.10), we consider the increasing sequence $\mathcal{D}_{X_0}$ and call the first element of this sequence as $\bar{\sigma}_1$. Clearly $\bar{\sigma}_1$ is a stopping time. Moreover,

$$\wp(\omega)\left([0,\ \bar{\sigma}_1) \times\ [0,\ C_{X_0}]\right) = 0,$$

and thus for $t \in [0,\ \bar{\sigma}_1)$

$$\begin{aligned}
X_t(\omega) &= X_0 + \int_{0+}^{t} \int_{0}^{C_{X_0}} h_\Lambda(X_{u-},\ Y_{u-}, N_{u-},\ v)\wp(\omega)(du,\ dv) = X_0,\\
Y_t(\omega) &= Y_0 + t - \int_{0+}^{t} \int_{0}^{C_{X_0}} (Y_{u-}) g_\Lambda(X_{u-},\ Y_{u-}, N_{u-},\ v)\wp(\omega)(du,\ dv) = Y_0 + t,\\
N_t(\omega) &= \int_{0+}^{t} \int_{0}^{C_{X_0}} g_\Lambda(X_{u-},\ Y_{u-}, N_{u-},\ v)\wp(\omega)(du,\ dv) = 0.
\end{aligned}$$

Note that in the above integrations the domain $\mathbb{R}$ has been replaced by the compact set $[0, C_{X_0}]$, as the integrands vanish outside this interval. Hence at $t = \bar{\sigma}_1$,

$$\begin{aligned}
X_{\bar{\sigma}_1}(\omega) &= X_0 + \int_{0}^{C_{X_0}} h_\Lambda(X_0,\ Y_{\bar{\sigma}_1 -}, N_0,\ v)\wp(\omega)(\{\bar{\sigma}_1\} \times dv),\\
Y_{\bar{\sigma}_1}(\omega) &= Y_0 + \bar{\sigma}_1 - \int_{0}^{C_{X_0}} (Y_{\bar{\sigma}_1 -}) g_\Lambda(X_0,\ Y_{\bar{\sigma}_1 -}, N_0,\ v)\wp(\omega)(\{\bar{\sigma}_1\} \times dv),\\
N_{\bar{\sigma}_1}(\omega) &= \int_{0}^{C_{X_0}} g_\Lambda(X_0,\ Y_{\bar{\sigma}_1 -}, N_0,\ v)\wp(\omega)(\{\bar{\sigma}_1\} \times\ dv).
\end{aligned}$$

Hence the solution is unique in the time interval $[0,\ \bar{\sigma}_1]$. Again by using (3.2.10), we have an increasing sequence $\mathcal{D}_{X_{\bar{\sigma}_1}}$. As $\mathcal{D}_{X_{\bar{\sigma}_1}} \cap (\bar{\sigma}_1, \infty)$ is nonempty almost surely, there is the first element of the set, denoted by $\bar{\sigma}_2$.

We would next argue that $\bar\sigma_2$ is a $\{\mathcal{F}_t\}$-stopping time. First we observe $\{\bar\sigma_2 \le t\} = (\{\bar\sigma_1 \ge t\} \cap \{\bar\sigma_2 \le t\}) \cup (\{\bar\sigma_1 < t\} \cap \{\bar\sigma_2 \le t\})$. The first event $\{\bar\sigma_1 \ge t\} \cap \{\bar\sigma_2 \le t\}$ is empty set and the second event $\{\bar\sigma_1 < t\} \cap \{\bar\sigma_2 \le t\}$ can be written as $\bigcup_{s\in\mathbb{Q}, s<t} (\{\wp(\omega)([0,s]\times[0,C_{X_0}]) = 1\} \cap \{\wp(\omega)((s,t]\times[0,C_{X_s}]) \neq 0\}) \in \mathcal{F}_t$. Thus we have,

$$\wp(\omega)\left((\bar\sigma_1,\ \bar\sigma_2)\times\ [0,\ C_{X_{\bar\sigma_1}}]\right) = 0$$

and for $t \in (\bar\sigma_1,\ \bar\sigma_2)$, as before

$$X_t(\omega) = X_{\bar\sigma_1} + \int_{\bar\sigma_1+}^{t}\int_0^{C_{X_{\bar\sigma_1}}} h_\Lambda(X_{u-},\ Y_{u-}, N_{u-},\ v)\wp(\omega)\ (du,\ dv) = X_{\bar\sigma_1},$$

$$Y_t(\omega) = Y_{\bar\sigma_1} + (t-\bar\sigma_1) - \int_{\bar\sigma_1+}^{t}\int_0^{C_{X_{\bar\sigma_1}}} (Y_{u-})g_\Lambda(X_{u-},\ Y_{u-}, N_{u-},\ v)\wp(\omega)\ (du,\ dv) = Y_{\bar\sigma_1} + (t-\bar\sigma_1),$$

$$N_t(\omega) = N_{\bar\sigma_1} + \int_{\bar\sigma_1+}^{t}\int_0^{C_{X_{\bar\sigma_1}}} g_\Lambda(X_{u-},\ Y_{u-}, N_{u-},\ v)\wp(\omega)(du,\ dv) = N_{\bar\sigma_1}.$$

Hence by using above equalities, at $t = \bar\sigma_2$,

$$X_{\bar\sigma_2}(\omega) = X_{\bar\sigma_1} + \int_0^{C_{X_{\bar\sigma_1}}} h_\Lambda(X_{\bar\sigma_1},\ Y_{\bar\sigma_2-}, N_{\bar\sigma_1},\ v)\wp(\omega)(\{\bar\sigma_2\}\times dv),$$

$$Y_{\bar\sigma_2}(\omega) = Y_{\bar\sigma_1} + (\bar\sigma_2-\bar\sigma_1) - \int_0^{C_{X_{\bar\sigma_1}}} (Y_{\bar\sigma_1-})g_\Lambda(X_{\bar\sigma_1},\ Y_{\bar\sigma_2-}, N_{\bar\sigma_1},\ v)\wp(\omega)(\{\bar\sigma_2\}\times dv),$$

$$N_{\bar\sigma_2}(\omega) = N_{\bar\sigma_1} + \int_0^{C_{X_{\bar\sigma_1}}} g_\Lambda(X_{\bar\sigma_1},\ Y_{\bar\sigma_2-}, N_{\bar\sigma_1},\ v)\wp(\omega)(\{\bar\sigma_2\}\times\ dv).$$

Continuing in the similar way we can construct a solution in a unique manner for each consecutive interval $(\bar\sigma_m,\ \bar\sigma_{m+1}]$, where $m \ge 2$. Again, with the similar argument $\{\bar\sigma_m\}_{m\ge0}$ are $\mathbb{F}_t$-stopping times. Moreover, for a fixed ω, $X_t(\omega) = X_{\bar\sigma_m(\omega)}$ for all $t \in [\bar\sigma_m(\omega), \bar\sigma_{m+1}(\omega))$. Hence X is an r.c.l.l. and piece-wise constant process almost surely on $[0, \lim_{m\to\infty}\bar\sigma_m]$. By this, Part (1) and (2) would have followed if $\{\bar\sigma_m\}_{m\ge1}$ were the jump times, which may not be true. For this reason, now we select an appropriate sub-sequence of $\bar\sigma_m$ which are the jump times. To this end we first note that $\int_0^{C_{X_{t-}}} g_\Lambda(X_{t-},\ Y_{t-}, N_{t-}, v)\wp(\omega)(\{t\}\times dv)$ is zero for all $t \in (\bar\sigma_m, \bar\sigma_{m+1})$ for every $m \ge 1$. Using this and $X_{\bar\sigma_m-} = X_{\bar\sigma_{m-1}}$, $N_{\bar\sigma_m-} = N_{\bar\sigma_{m-1}}$ we introduce

$$l_1 := \min\{m \ge 1\colon \int_0^{C_{X_{\bar\sigma_{m-1}}}} g_\Lambda(X_{\bar\sigma_{m-1}},\ Y_{\bar\sigma_m-}, N_{\bar\sigma_{m-1}},\ v)\wp(\omega)(\{\bar\sigma_m\}\times dv) \neq 0\}. \tag{3.2.11}$$

From (3.2.4) and (3.2.5) it is evident that h_Λ and g_Λ have identical supports in v variable. Hence the integral
$\int_I h_\Lambda(X_{t-}, Y_{t-}, N_{t-}v)\wp(\{t\}\times dv)$ is non-zero if and only if $\int_I g_\Lambda(X_{t-}, Y_{t-}, N_{t-}v)\wp(\{t\}\times dv)$ is non-zero, where I is any interval. Then $t = \bar\sigma_{l_1}$ is the first time when both the integrals

$$\int_{[0,C_{X_{t-}}]} g_\Lambda(X_{t-},\ Y_{t-}, N_{t-},\ v)\wp(\omega)(\{t\}\times dv) \quad \text{and} \quad \int_{[0,C_{X_{t-}}]} h_\Lambda(X_{t-},\ Y_{t-}, N_{t-},\ v)\wp(\omega)(\{t\}\times dv)$$

are non-zero. Consequently, $Y_t = Y_0 + t$, $X_t = X_0$, and $N_t = 0$, for all $t \in [0, \bar{\sigma}_{l_1})$. Hence $X_{\bar{\sigma}_{l_1}-} = X_0$, $Y_{\bar{\sigma}_{l_1}-} = Y_0 + \bar{\sigma}_{l_1}$ and $N_{\bar{\sigma}_{l_1}-} = 0$. Furthermore, at $t = \bar{\sigma}_{l_1}$,

$$\begin{aligned}
0 \neq & \int_{0+}^{\bar{\sigma}_{l_1}} \int_0^{C_{X_{\bar{\sigma}_{l_1-1}}}} g_\Lambda(X_{t-},\ Y_{t-}, N_{t-},\ v) \wp(\omega)(dt, dv) \\
= & \int_0^{C_{X_{\bar{\sigma}_{l_1-1}}}} g_\Lambda(X_{\bar{\sigma}_{l_1-1}},\ Y_{\bar{\sigma}_{l_1}-}, N_{\bar{\sigma}_{l_1-1}},\ v) \wp(\omega)(\{\bar{\sigma}_{l_1}\} \times dv) = 1,
\end{aligned}$$

using (3.2.11) and the facts that $\wp(\omega)(\{\bar{\sigma}_{l_1}\} \times [0, C_{X_{\bar{\sigma}_{l_1-1}}}]) = 1$ and $g_\Lambda(i, y, n, v) \in \{0, 1\}$ for every i, y, n, and v. Thus from (3.2.8) and above expressions

$$N_{\bar{\sigma}_{l_1}} = \int_{0+}^{\bar{\sigma}_{l_1}} \int_0^{C_{X_{\bar{\sigma}_{l_1-1}}}} g_\Lambda(X_{t-},\ Y_{t-}, N_{t-},\ v) \wp(\omega)(dt, dv) = 1.$$

Similarly, using (3.2.7) we have $Y_{\bar{\sigma}_{l_1}} = 0$. To see this, we write $Y_{\bar{\sigma}_{l_1}} = Y_{\bar{\sigma}_{l_1-}} + (Y_{\bar{\sigma}_{l_1}} - Y_{\bar{\sigma}_{l_1}-})$, i.e.,

$$\begin{aligned}
Y_{\bar{\sigma}_{l_1}} = & Y_{\bar{\sigma}_{l_1-}} - \int_0^{C_{X_{\bar{\sigma}_{l_1-1}}}} (Y_{\bar{\sigma}_{l_1}-}) g_\Lambda(X_{\bar{\sigma}_{l_1}-}, Y_{\bar{\sigma}_{l_1}-}, N_{\bar{\sigma}_{l_1}-},\ v) \wp(\omega)(\{\bar{\sigma}_{l_1}\} \times dv) \\
= & Y_{\bar{\sigma}_{l_1-}} - Y_{\bar{\sigma}_{l_1-}} = 0.
\end{aligned}$$

In general, for every $n \geq 1$, we set

$$l_{n+1} := \min \left\{ m > l_n \colon \int_0^{C_{X_{\bar{\sigma}_{m-1}}}} g_\Lambda(X_{\bar{\sigma}_{m-1}},\ Y_{\bar{\sigma}_m -}, N_{\bar{\sigma}_{m-1}},\ v) \wp(\omega)(\{\bar{\sigma}_m\} \times dv) \neq 0 \right\}. \tag{3.2.12}$$

In other words, for every $t \geq 0$,

$$\int_{\mathbb{R}} g_\Lambda(X_{t-},\ Y_{t-}, N_{t-}\ v) \wp(\{t\} \times dv) = \begin{cases} 1 & , \text{ if } t = \bar{\sigma}_{l_n} \text{ for some } n \geq 1 \\ 0 & , \text{ otherwise.} \end{cases} \tag{3.2.13}$$

From (3.2.13) and (3.2.8) we get

$$N_t = \sum_{\{r \geq 1 | \bar{\sigma}_{l_r} \leq t\}} 1 = \sum_{r=1}^{\infty} \mathbb{1}_{[\bar{\sigma}_{l_r}, \infty)}(t). \tag{3.2.14}$$

Moreover, from (3.2.13) and (3.2.7) we get

$$Y_t = Y_0 + t - \sum_{\{r \geq 1 | \bar{\sigma}_{l_r} \leq t\}} Y_{\bar{\sigma}_{l_r}-} = Y_0 + t - \sum_{r=1}^{\infty} (Y_{\bar{\sigma}_{l_r}-}) \mathbb{1}_{[\bar{\sigma}_{l_r}, \infty)}(t). \tag{3.2.15}$$

Furthermore, as the support of g_Λ and h_Λ are identical, from (3.2.12) we have that the integral

$$\int_{\mathbb{R}} h_\Lambda(X_{t-},\ Y_{t-}, N_{t-}\ v) \wp(\{t\} \times dv)$$

is nonzero if and only if $t = \bar{\sigma}_{l_n}$ for some $n \geq 1$. Therefore, if $n \geq 1$ is such that $\bar{\sigma}_{l_n} \leq t < \bar{\sigma}_{l_{n+1}}$, then $X_t = X_{\bar{\sigma}_{l_n}}$. Thus we can write

$$X_t = X_0 + \sum_{\{r\geq 1 | \bar{\sigma}_{l_r} \leq t\}} (X_{\bar{\sigma}_{l_r}} - X_{\bar{\sigma}_{l_{r-1}}}) = X_0 + \sum_{r=1}^{\infty} (X_{\bar{\sigma}_{l_r}} - X_{\bar{\sigma}_{l_{r-1}}}) \mathbb{1}_{[\bar{\sigma}_{l_r},\infty)}(t).$$

Hence X and N are r.c.l.l., and piece-wise constant and Y is r.c.l.l., and piece-wise linear. We denote $T_n := \bar{\sigma}_{l_n}$ for each $n \geq 1$. From above, it is evident that $\{T_n\}_{n\geq 1}$ is the desired sequence of stopping times at which the processes X, Y, and N jump and the properties in parts (1) and (2) hold.

Next for proving part (3), we first rewrite (3.2.15)

$$Y_t = Y_0 + t - \sum_{\{r\geq 1 | T_r \leq t\}} Y_{T_r -}$$

and thus for all $n \in \mathbb{N}$,

$$\begin{aligned} Y_{T_n} &= Y_0 + T_n - \sum_{\{r\geq 1 | T_r \leq T_n\}} Y_{T_r -} \\ &= \left(Y_0 + T_n - \sum_{\{r\geq 1 | T_r < T_n\}} Y_{T_r -} \right) - Y_{T_n -} \\ &= Y_{T_n -} - Y_{T_n -} = 0. \end{aligned}$$

Thus (i) holds. For showing (ii) we first recall that $Y_{T_1 -} = Y_0 + T_1$, which is same as $T_1 - T_0$. Now for $n > 1$ using $Y_{T_{n-1}} = 0$ we get

$$\begin{aligned} Y_{T_n -} &= \left(Y_0 + T_n - \sum_{\{r\geq 1 | T_r < T_n\}} Y_{T_r -} \right) \\ &= (T_n - T_{n-1}) + \left(Y_0 + T_{n-1} - \sum_{\{r\geq 1 | T_r \leq T_{n-1}\}} Y_{T_r -} \right) \\ &= (T_n - T_{n-1}) + Y_{T_{n-1}} = (T_n - T_{n-1}). \end{aligned}$$

Hence (ii) is shown. Again, (3.2.14) implies that $N_t = \max\{r : T_r \leq t\}$. That is $N_{T_n} = \max\{r : T_r \leq T_n\} = n$, which proves (iii). Finally, (iv) follows using (3.2.15) and (ii) as below

$$Y_t = Y_0 + t - \sum_{\{1\leq r \leq N_t\}} (T_r - T_{r-1}) = Y_0 + t - (T_{N_t} - T_0) = t - T_{N_t}. \quad \square$$

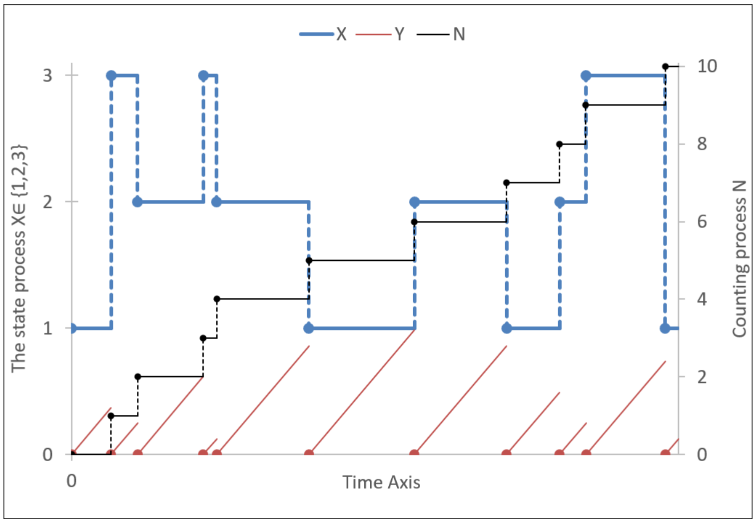


Figure 3.1: Illustration of a sample path of the joint process (X, Y, N) where the state space is $\mathcal{X} = \{1, 2, 3\}$. The state, age, and counting processes are plotted in blue, red, and black against the horizontal time axis.

Remark 3.2.3. *The jump times $\{T_n\}_{n\geq 1}$ of the process X are called the transition times. We need to prove that this sequence diverges to infinity almost surely for establishing that the solution is globally determined with probability 1. This is accomplished in the following Theorem.*

Theorem 3.2.4. *There exists a unique strong solution $Z = (X_t, Y_t, N_t)_{t\geq 0}$ of the coupled SDE* (3.2.6)-(3.2.8)

Proof. Let $\{T_n\}_{n\geq 1}$ be as in Theorem 3.2.2. Now we will show that T_n diverges. Let $\tau := \lim_{n\to\infty} T_n$. Clearly, for a fixed $\epsilon > 0$, the event $\{\tau < \infty\}$ is a subset of $\cup_{n_0\geq 1}\{T_n - T_{n-1} < \epsilon,\ \forall n \geq n_0\}$. For the sake of brevity, we denote $\{T_n - T_{n-1} < \epsilon\}$ as A_n for each n. Therefore, we will show $P(\cap_{n\geq n_0} A_n) = 0$ for each n_0 which is enough for proving $P(\tau < \infty) = 0$. To this end we first compute the conditional probability of the event A_n given the observations till T_{n-1}, using the properties of Poisson random measure. Indeed, using part (3) of Theorem 3.2.2, occurrence of A_n is same as having Poisson random measure of $\cup_{0<y<\epsilon}\left(\{T_{n-1} + y\} \times \Lambda_{X_{T_{n-1}}}(y, n-1)\right)$ positive. Thus using the Lebesgue intensity of $\wp$, we have

$$
\begin{aligned}
&P\left(A_n \mid T_{n-1}, X_{T_{n-1}}, \ldots, T_0, X_0\right) \\
&= P\left(\wp\left(\bigcup_{0<y<\epsilon}\left(\{T_{n-1} + y\} \times \Lambda_{X_{T_{n-1}}}(y, n-1)\right)\right) \neq 0 \mid T_{n-1}, X_{T_{n-1}}, \ldots, T_0, X_0\right) \\
&= \left(1 - e^{-\int_0^\epsilon \lambda_{X_{T_{n-1}}}(y,n-1)dy}\right) \\
&\leq \left(1 - e^{-\epsilon c}\right) \qquad (3.2.16)
\end{aligned}
$$

which is a deterministic constant. We recall that the last inequality is due to Assumption (B1). Next we note that if

$$E\left[\prod_{n_0\le n\le l} \mathbb{1}_{A_n}\right] \le \left(1-e^{-\epsilon c}\right) E\left[\prod_{n_0\le n\le l-1} \mathbb{1}_{A_n}\right] \tag{3.2.17}$$

holds for all $l \ge n_0$, using that repeatedly, we get

$$P\left(\cap_{n_0\le n} A_n\right) \le P\left(\cap_{n_0\le n\le l} A_n\right) = E\left[\prod_{n_0\le n\le l} \mathbb{1}_{A_n}\right] \le \left(1-e^{-\epsilon c}\right)^{l-n_0+1}$$

for all $l \ge n_0$. The right side clearly vanishes as $l \to \infty$, and thus $P(\cap_{n_0\le n} A_n) = 0$, as desired, provided (3.2.17) holds. Finally we show (3.2.17) below using the property of conditional expectation, and inequality (3.2.16)

$$\begin{aligned}
E\left[\prod_{n_0\le n\le l} \mathbb{1}_{A_n}\right] &= E\left[E\left[\prod_{n_0\le n\le l} \mathbb{1}_{A_n}\middle| T_{l-1}, X_{T_{l-1}}, \dots, T_0, X_0\right]\right] \\
&= E\left[\left(\prod_{n_0\le n\le l-1} \mathbb{1}_{A_n}\right) E\left[\mathbb{1}_{A_l}\middle| T_{l-1}, X_{T_{l-1}}, \dots, T_0, X_0\right]\right] \\
&\le \left(1-e^{-\epsilon c}\right) E\left[\prod_{n_0\le n\le l-1} \mathbb{1}_{A_n}\right]
\end{aligned}$$

for all $l \ge n_0$. Hence the proof is complete. □

The above theorem essentially asserts that the jump process X is pure. In the next section we show that X is a semi-Markov process.
The results in earlier chapter can be easily extended for NHSMP. However, for the sake of completeness we give full details of this extension in the following Theorems 3.3.1, 3.4.1, 3.4.4 and Proposition 3.4.3, 3.4.5.

3.3 Semi-Markov Law of the Solution

Theorem 3.3.1. *Let* $Z = (X, Y, N) = \{(X_t,\ Y_t, N_t)\}_{t\ge0}$ *be a unique strong solution to* (3.2.6)-(3.2.8)*. Then the following hold.*

i. The process $\{X_t\}_{t\ge0}$ *is a pure SMP.*

ii. The embedded chain $\{X_{T_n}\}_{n\ge1}$ *is a discrete time Markov chain.*

Proof. First we will prove that X is SMP. We have already seen in the proof of Theorem 3.2.2, that X is an r.c.l.l. process. Next, we need to show (0.2.2), i.e., for each $n \in \mathbb{N}$, $y \in \mathbb{R}_0$, $j \in \mathcal{X}$

$$P[X_{T_{n+1}} = j, T_{n+1} - T_n \leq y \mid X_0, T_0, X_{T_1}, T_1, \ldots, X_{T_n}, T_n] \\ = P[X_{T_{n+1}} = j, T_{n+1} - T_n \leq y \mid X_{T_n}]. \tag{3.3.1}$$

We note that the left side is equal to

$$P(T_{n+1} - T_n \leq y \mid (X_0, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n)) \\ \times P(X_{T_{n+1}} = j \mid (X_0, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n), \{T_{n+1} - T_n \leq y\}). \tag{3.3.2}$$

Each of the two conditional probabilities is further simplified below. For almost every $\omega \in \Omega$, Equation (3.2.13) and Theorem 3.2.2 part (3)(iii)-(iv) imply that for any $n \geq 1$

$$\int_{T_n+}^{T_n+t} (u - T_n) \int_{\mathbb{R}} g_\Lambda(X_{T_n}, u - T_n, n, v)\, \wp(du, dv) = \begin{cases} 0, & \text{for } t < T_{n+1} - T_n \\ T_{n+1} - T_n, & \text{for } t = T_{n+1} - T_n. \end{cases}$$

Hence, by a suitable change of variable, almost surely $T_{n+1} - T_n$ is the first occurrence of a non-zero value of the following map

$$t \mapsto \int_{0+}^{t} u \int_{\mathbb{R}} g_\Lambda(X_{T_n}, u, n, v)\, \wp(T_n + du, dv)$$

and that occurs at $t = T_{n+1} - T_n$. Again, since $\wp(T_n + du, dv)$ is independent to $\mathcal{F}_{T_n}$ we obtain, $T_{n+1} - T_n$ is conditionally independent to $\mathcal{F}_{T_n}$ given X_{T_n}. Thus for all $n \geq 1$

$$P\left(T_{n+1} - T_n \leq y \mid (X_0, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n)\right) \\ = P(T_{n+1} - T_n \leq y \mid X_{T_n}). \tag{3.3.3}$$

By substituting t equal to T_n and T_{n+1} in Equation (3.2.6), and using Theorem 3.2.2 part (3) (ii)-(iii) we get

$$X_{T_{n+1}} = X_{T_n} + \int_{\mathbb{R}} h_\Lambda(X_{T_n}, T_{n+1} - T_n, n, v)\, \wp(\{T_{n+1}\} \times dv), \tag{3.3.4}$$

as $X_{T_{n+1}-} = X_{T_n}$, $Y_{T_{n+1}-} = T_{n+1} - T_n$ and $N_{T_{n+1}-} = n$. Thus using (3.3.4)

$$P\left(X_{T_{n+1}} = j \mid (X_0, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n), \{T_{n+1} - T_n \leq y\}\right) \\ = P\Bigg(\int_{\mathbb{R}} h_\Lambda(X_{T_n}, T_{n+1} - T_n, n, v)\, \wp(\{T_n + (T_{n+1} - T_n)\} \times dv) = j - X_{T_n} \Bigg| \\ (X_0, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n), \{T_{n+1} - T_n \leq y\} \Bigg).$$

Again, using the independence of $\wp(T_n + du, dv)$ to $\mathcal{F}_{T_n}$ and conditional independence of $T_{n+1} - T_n$ to $\mathcal{F}_{T_n}$ given X_{T_n} we conclude, the right side expression is equal to

$$P\bigg(\int_{\mathbb{R}} h_\Lambda(X_{T_n}, T_{n+1} - T_n, n, v)\, \wp(\{T_n + (T_{n+1} - T_n)\} \times dv) = j - X_{T_n} \;\bigg|\; X_{T_n}, \{T_{n+1} - T_n \le y\}\bigg)$$

which is again equal to $P\left(X_{T_{n+1}} = j \mid X_{T_n}, \{T_{n+1} - T_n \le y\}\right)$ using (3.3.4). Thus, using this simplification and (3.3.3) in (3.3.2), we obtain for all $n \ge 1$

$$\begin{aligned} &P(X_{T_{n+1}} = j, T_{n+1} - T_n \le y \mid (X_0, T_0), (X_{T_1}, T_1), \ldots, (X_{T_n}, T_n)) \\ &= P\left(T_{n+1} - T_n \le y \mid X_{T_n}\right) P\left(X_{T_{n+1}} = j \mid X_{T_n}, \{T_{n+1} - T_n \le y\}\right) \\ &= P\left(X_{T_{n+1}} = j, T_{n+1} - T_n \le y \mid X_{T_n}\right). \end{aligned}$$

Hence, part (i) is proved.
For every $n \ge 1$, taking $y \to \infty$ on both sides in Equation (3.3.1), we get

$$P[X_{T_{n+1}} = j \mid X_{T_0}, T_0, X_{T_1}, T_1, \ldots, X_{T_n}, T_n] = P[X_{T_{n+1}} = j \mid X_{T_n}].$$

Hence the part (ii). □

Theorem 3.3.2. *Let $Z = (X, Y, N) = \{(X_t, Y_t, N_t)\}_{t \ge 0}$ be a unique strong solution to (3.2.6)-(3.2.8) with an additional condition that $\mathcal{X}$ is finite, i.e., $\mathcal{X} = \{1, 2, \ldots, k\}$ for some natural number k. Then the process Z is a strong Markov process.*

Proof. We note that for each $z := (i, y, n) \in \mathcal{X} \times \mathbb{R}_+ \times \mathbb{N}_0$, the support of the functions $h_\Lambda(z, \cdot)$, and $g_\lambda(z, \cdot)$ are contained in $\Lambda_i(y, n)$. On the other hand, for almost every $y > 0$, $\Lambda_i(y, n)$ is contained in the interval $[0, \sum_{i \in \mathcal{X}} \|\lambda_i\|_{L^\infty_{(\mathbb{R}_+ \times \mathbb{N}_0)}})$, which is a finite interval due to (B1) provided $\mathcal{X}$ is finite. Hence by following the line of argument of Theorem 2.3.5 of Chapter 2, which uses [8, Theorem IX.3.9], we obtain the result. □

3.4 Expression of Transition Kernel

For each $i \in \mathcal{X}, n \in \mathbb{N}_0$, we define a function $F(\cdot \mid i, n)\colon [0, \infty) \to [0, 1]$ as

$$F(y \mid i, n) := 1 - e^{-\gamma_i(y,n)} \tag{3.4.1}$$

where $\gamma_i(y, n)$ is as in (B2). Clearly, $F(\cdot \mid i, n)$ is differentiable almost everywhere. To see this, we note that $\gamma_i(y, n)$ is an integral of a bounded Lebesgue measurable function, and thus is absolutely continuous in y. Let $f(y \mid i, n)$ be the almost everywhere derivative of

$F(y \mid i,n)$. We also define for each $y \in \mathbb{R}_+$, $n \in \mathbb{N}_0$, a matrix $p(y,n) := (p_{ij}(y,n))_{\mathcal{X}\times\mathcal{X}}$, such that

$$p_{ij}(y,n) := \begin{cases} \dfrac{\lambda_{ij}(y,n)}{\lambda_i(y,n)} \mathbb{1}_{(0,\infty)}(\lambda_i(y,n)), & \text{if } j \neq i \\ \mathbb{1}_{\{0\}}(\lambda_i(y,n)), & \text{if } j = i. \end{cases} \tag{3.4.2}$$

Since, $\lambda_{ij}(y,n) \geq 0$ and $\sum_{j\in\mathcal{X}\setminus\{i\}} \lambda_{ij}(y,n) = \lambda_i(y,n)$, for each $n \in \mathbb{N}_0$ and $y \in \mathbb{R}_+$, $p(y,n)$ is a transition probability matrix. The following theorem asserts that $p(y,n)$ gives the conditional probability of selecting a state at the time of $n+1^{\text{st}}$ transition given the age y and location i of the previous state. Furthermore, the map $F(\cdot \mid i,n)$ as in (3.4.1) is also asserted as the conditional cumulative distribution function of the holding time at the nth state given that is i.

Theorem 3.4.1. *Let $Z = (X,Y,N)$ be a solution to* (3.2.6)-(3.2.8)*, $i \in \mathcal{X}$, $y \in \mathbb{R}_+$ and $n \geq 1$ then the following hold.*

i. $F(\cdot \mid i,n)$, as in (3.4.1)*, is the conditional cumulative distribution function of the holding time of the process X.*

ii. For all $j \neq X_{T_n}$, $p_{X_{T_n}j}(T_{n+1}-T_n,n) = P[X_{T_{n+1}} = j \mid X_{T_n}, T_{n+1}-T_n]$ almost surely, where $p_{ij}(\cdot,\cdot)$ is as in (3.4.2).

Proof. We recall that, the intensity of $\wp$ is Lebesgue measure, and the Lebesgue measure of $\{(u,v) \in (T_n, T_n+y) \times \mathbb{R}_+ \mid v \in \Lambda_i(u-T_n,n)\}$ is $\int_{T_n}^{T_n+y} \lambda_i(u-T_n,n)du$ which is equal to $\gamma_i(y,n)$ (see (B2)). Using (3.2.5) and (3.2.13), the conditional probability of no transition in the next y unit time, given that the n^{th} transition happens now to state i, is given by $e^{-\gamma_i(y,n)}$. Using the above fact, the conditional cumulative distribution function at y of the holding time after the n^{th} transition, given the state i, is

$$\begin{aligned} P[T_{n+1}-T_n \leq y \mid X_{T_n} = i] &= 1 - P[X_t = X_{t-}, \forall t \in (T_n, T_n+y] \mid X_{T_n} = i] \\ &= 1 - e^{-\gamma_i(y,n)} \end{aligned}$$

for all $y \in \mathbb{R}_+$ and $i \in \mathcal{X}$. Thus (i) follows from (3.4.1).

We note that, for $j \neq i$, $P[X_{T_{n+1}} = j \mid X_{T_n} = i, T_{n+1}-T_n = y]$ is the conditional probability of the event that the $n+1^{\text{st}}$ state is j, given that $T_{n+1} = T_n + y$ and the nth state is i. Using (3.3.4), the above is the conditional probability that a Poisson point mass appears in $\{T_n+y\} \times \Lambda_{ij}(y,n)$ given that the point mass lies somewhere in $\{T_n+y\} \times \Lambda_i(y,n)$ and no transition of X occurs during (T_n, T_n+y). If these three events are denoted by A, B, and C respectively, then the conditional probability $P(A \mid B \cap C)$ can be simplified as $P(A \mid B)$

because C is independent to both A and B. Thus using the Lebesgue intensity of $\wp$,

$$\begin{aligned}
&P\left[X_{T_{n+1}} = j \mid X_{T_n} = i, T_{n+1} - T_n = y\right] \\
&= P\left[\wp(\{T_n + y\} \times \Lambda_{ij}(y,n)) = 1 \mid \wp(\{T_n + y\} \times \Lambda_i(y,n)) = 1\right] \\
&= \frac{|\Lambda_{ij}(y,n)|}{|\Lambda_i(y,n)|} \\
&= \frac{\lambda_{ij}(y,n)}{\lambda_i(y,n)}
\end{aligned}$$

for every $y \in \mathbb{R}_+$, and $j \neq i$, provided $\lambda_i(y,n) \neq 0$. Thus from (3.4.2), we get $p_{ij}(y,n)$ is equal to
$P\left[X_{T_{n+1}} = j \mid X_{T_n} = i, T_{n+1} - T_n = y\right]$ when $\lambda_i(y,n) \neq 0$. Next we note that $\lambda_i(y,n) = 0$ if and only if
$\frac{d}{dy}F(y \mid i,n) = 0$, i.e., the density of $T_{n+1} - T_n$ is zero at y. Hence (ii) holds. □

Remark 3.4.2. *Using* (3.4.1)*, we note that under Assumptions (B1) and (B2),* $F(y \mid i,n) < 1$ *for all* $y \in \mathbb{R}_+$ *and* $\lim_{y\to\infty} F(y \mid i,n) = 1$ *using* (3.4.1)*. Thus, the holding times are unbounded but finite almost surely. By dropping (B1), one may include a class of SMPs having bounded holding times. However, we exclude that class from our discussion. It is also important to note that the SMPs having discontinuous cdf of holding times are also not considered in the present setting. Nevertheless, the present study subsumes countable-state continuous time Markov chains and the processes having age dependent transitions as appear in [14, 32].*

Proposition 3.4.3. *We have, for almost every* $y \geq 0$ *and* $n \geq 1$*,*

$$p_{ij}(y,n)\frac{f(y \mid i,n)}{1 - F(y \mid i,n)} = \begin{cases} \lambda_{ij}(y,n), & \text{for } i \neq j, \\ 0, & \text{for } i = j. \end{cases}$$

Proof. By differentiating both sides of (3.4.1), we obtain $f(y \mid i,n) = \lambda_i(y,n)e^{-\gamma_i(y,n)}$ for every $y \in \mathbb{R}_+$. This is equal to $\lambda_i(y,n)(1 - F(y \mid i,n))$ using (3.4.1). Hence, for every $y \in \mathbb{R}_+$, $n \geq 1$ and $i \in \mathcal{X}$

$$\frac{f(y \mid i,n)}{1 - F(y \mid i,n)} = \lambda_i(y,n). \tag{3.4.3}$$

If $i \neq j$, for every $y \in \mathbb{R}_+$, using (3.4.2)

$$p_{ij}(y,n)\frac{f(y \mid i,n)}{1 - F(y \mid i,n)} = \frac{\lambda_{ij}(y,n)}{\lambda_i(y,n)} \times \lambda_i(y,n)\mathbb{1}_{(0,\infty)}(\lambda_i(y,n)) = \lambda_{ij}(y,n)$$

as $0 \leq \lambda_{ij}(y,n) \leq \lambda_i(y,n)$. The case for $i = j$ follows from (3.4.2) and (2.4.4) directly as $p_{ii}(y,n)\dfrac{f(y \mid i,n)}{1 - F(y \mid i,n)}$ is equal to $\lambda_i(y,n)\mathbb{1}_{\{0\}}(\lambda_i(y,n))$ which is zero. □

Theorem 3.4.4. *Let X be a SMP as in Theorem 3.3.1. Then, the associated kernel is given by*

$$P[X_{T_{n+1}} = j, T_{n+1} - T_n \le y \mid X_{T_n} = i] = \int_0^y e^{-\gamma_i(s,n)} \lambda_{ij}(s,n)\ ds,$$

which is denoted by $Q_{ij}(y,n)$ for every $y > 0$, $n \ge 1$ and $i \ne j$.

Proof. Using Theorem 3.4.1 (i) and (ii) and Theorem 3.2.2 part (3) (ii)

$$\begin{aligned}
&P\left[X_{T_{n+1}} = j, T_{n+1} - T_n \le y \mid X_{T_n} = i\right] \\
&= \mathbb{E}\left[P\left(X_{T_{n+1}} = j, T_{n+1} - T_n \le y \mid X_{T_n} = i, T_{n+1} - T_n\right) \mid X_{T_n} = i\right] \\
&= \int_0^\infty \mathbb{1}_{[0,y]}(s) P\left[X_{T_{n+1}} = j \mid X_{T_n} = i, T_{n+1} - T_n = s\right] f(s \mid i,n)\, ds \\
&= \int_0^y p_{ij}(s,n) f(s \mid i,n)\, ds.
\end{aligned}$$

For each $i \ne j$, using Proposition 3.4.3 and (3.4.1), the right side of above can be rewritten as

$$\int_0^y (1 - F(s \mid i,n))\lambda_{ij}(s,n)\, ds \;=\; \int_0^y e^{-\gamma_i(s,n)} \lambda_{ij}(s,n)\, ds = Q_{ij}(y,n).$$

□

Proposition 3.4.5. *Let X be a SMP as in Theorem 3.3.1. Then, $\lambda(n,y)$ is the instantaneous transition rate matrix.*

Proof. The rate of transition from state i to j at age y is given by

$$\begin{aligned}
&\lim_{h\to 0} \frac{1}{h}\left[P(X_{T_{n+1}} = j, T_{n+1} - T_n \in (y, y+h] \mid X_{T_n} = i, \{T_{n+1} - T_n > y\})\right] \\
&= \lim_{h\to 0} \frac{1}{h} \frac{P(X_{T_{n+1}} = j, T_{n+1} - T_n \in (y, y+h] \mid X_{T_n} = i)}{P(T_{n+1} - T_n > y \mid X_{T_n} = i)} \\
&= \lim_{h\to 0} \frac{1}{h} \frac{P(X_{T_{n+1}} = j, T_{n+1} - T_n \le y + h \mid X_{T_n} = i) - P(X_{T_{n+1}} = j, T_{n+1} - T_n \le y \mid X_{T_n} = i)}{1 - P(T_{n+1} - T_n \le y \mid X_{T_n} = i)}.
\end{aligned}$$

Using Theorem 3.4.4, the above limit is equal to $\frac{\frac{d}{dy}Q_{ij}(y,n)}{1-F(y|i,n)}$ which can further be simplified as $\lambda_{ij}(y,n)$. □

Example 3.4.6. *Let $\mathcal{X} = \{1,2,3\}$ and k_1, k_2, k_3 be some real constants and for all i, j in $\mathcal{X}$, $n \ge 0$ and $y \ge 0$,*

$$\lambda_{ij}(y,n) = \begin{cases} \frac{y \sin^2(k_i y)}{1+y} & \text{if } j - i \text{ modulo } 3 = 1 \\ \frac{y \cos^2(k_i y)}{1+y} & \text{if } j - i \text{ modulo } 3 = 2 \\ 0 & \text{else.} \end{cases}$$

Hence, from (3.4.1) $\lambda_i(y,n) = \frac{y}{1+y}$ *and* $F(y \mid i,n) = 1 - \exp\left(-\int_0^y \frac{y'}{1+y'}dy'\right) = 1 - e^{-y+\ln(1+y)} = 1-(1+y)e^{-y}$. *Consequently, for each* $y \geq 0$, $i \in \mathcal{X}$, $n \in \mathbb{N}_0$, $f(y \mid i,n) = ye^{-y}$, *the density of Gamma distribution with shape parameter 2 and rate parameter 1 and using* (3.4.2) *for* $j \neq i$

$$p_{ij}(y,n) = \begin{pmatrix} 0 & \sin^2(k_1 y) & \cos^2(k_1 y) \\ \cos^2(k_2 y) & 0 & \sin^2(k_2 y) \\ \sin^2(k_3 y) & \cos^2(k_3 y) & 0 \end{pmatrix}.$$

Remark 3.4.7. *We have obtained* $Q_{ij}(y,n) = \int_0^y p_{ij}(s,n) f(s \mid i,n)\, ds$ *in the proof of Theorem 3.4.4, which expresses* $Q_{ij}(\cdot,n)$ *in terms of the* $p_{ij}(\cdot,n)$, *and* $f(\cdot \mid i,n)$. *These parameters give the age dependent transition probabilities and the conditional holding time densities. In an alternative conditioning, the kernel can also be expressed as* $Q_{ij}(y,n) = P\left[X_{T_{n+1}} = j \mid X_{T_n} = i\right] P\left[T_{n+1} - T_n \leq y \mid X_{T_n} = i, X_{T_{n+1}} = j\right]$, *which is the product of transition probabilities of embedded chain and the conditional cdf of holding time given the current and the next states. Generally an SMP is characterized using the transition kernel* Q. *Although, instantaneous transition rate matrix* λ *also characterizes an SMP,* Q *is considered more fundamental as* λ *exists only if* Q *is differentiable. In that case, each of* Q *and* λ *can be expressed in terms of another, which is evident from the above two results.*

3.5 Non-homogeneous Component-wise Semi-Markov Process

It is interesting to note that although a pair of Markov processes form a Markov process again, a pair of SMPs do not form a SMP. Hence, the process whose components are SMPs needs a separate study. On the other hand, such processes arise naturally by solving (3.2.6)-(3.2.8) with two different initial conditions. While the law of SMP can be identified by its transition kernel, the same is not obvious for a pair of correlated SMPs. In view of this, we identify the law of a pair of SMPs by deriving the infinitesimal generator of its augmented process.

Definition 3.5.1. *A pure jump process* X *on a countable state space* $\mathcal{X}$ *is called a component-wise Semi-Markov Process (CSM) if there is a bijection* $\Gamma : \mathcal{X} \to \prod_{i=1}^d \mathcal{X}_i$, *such that each component of* $\Gamma(X)$ *is a semi-Markov process, where* d *is a positive integer and for each* $i \leq d$, $\mathcal{X}_i$ *is an at-most countable non-empty set.*

Without loss of generality, we assume that $\mathcal{X} = \prod_{i=1}^d \mathcal{X}_i$ and X^i is a semi-Markov process on $\mathcal{X}_i$ for each $i \leq d$. Here Γ is the identity map. Next we consider a specific CSM process of dimension 2.

Notation 3.5.2. *Fix* $i,j \in \mathcal{X}$ *and* $y_1, y_2 \geq 0$. *Let* $Z^1 = (X^1, Y^1, N^1)$ *and* $Z^2 = (X^2, Y^2, N^2)$ *be the strong solutions of* (3.2.6)-(3.2.8) *with two different initial conditions. At a fixed time*

$s(>0)$ we denote

$$i = X^1_s, y_1 = Y^1_s, n_1 = N^1_s \tag{3.5.1}$$

and

$$j = X^2_s, y_2 = Y^2_s, n_2 = N^2_s \tag{3.5.2}$$

respectively. We also denote (Z^1, Z^2) as Z. The successive transition times of X^1 and X^2 are denoted by $\{T^1_n\}$ and $\{T^2_n\}$ respectively. Let $\tau(t)$ denote the time of next transition after a given time t of either of the chains, thus $\tau(t) := T^1_{N^1_t+1} \wedge T^2_{N^2_t+1}$. If $\tau^{(n)}$ denotes composition of n number of map τ, then $\{\tau^{(n)}(s)\}_n$ gives a sequence of stopping times, representing the successive transitions of the combined process (X^1, X^2) after time s.

For the sake of computing some specific parameters connected to the law of Z, we limit ourselves to the following choice of $\tilde{\lambda}$.

(B3) For almost every $y \geq 0$, $n \in \mathbb{N}_0$ and $(i,j) \in \mathcal{X}_2$, $\tilde{\lambda}_{ij}(y,n) = \|\lambda_{ij}(\cdot,\cdot)\|_{L^\infty_{(\mathbb{R}_+\times\mathbb{N}_0)}}$.

We next extend Theorem 2.5.3 in the present settings below. As before $Z = (Z^1, Z^2)$ is Markov. It has state, component-wise age and component-wise number of transitions $X = (X^1, X^2)$, $Y = (Y^1, Y^2)$ and $N = (N^1, N^2)$ respectively. The Markov process Z is called the augmented process of CSM X. Under (B3), the infinitesimal generator $\mathcal{A}$ of the augmented process $Z = (Z^1, Z^2)$ is obtained below using Itô's lemma for r.c.l.l. semimartingales. The arguments are analogous to those in the proof of Theorem 2.5.3. The details are produced below for the sake of completeness. Let $\varphi\colon (\mathcal{X} \times \mathbb{R}_+ \times \mathbb{N}_0)^2 \to \mathbb{R}$ be bounded and continuously differentiable in its continuous variables. Then using (3.2.9)

$$\begin{aligned}
&d\varphi(Z^1_t, Z^2_t) - \left(\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2}\right)\varphi(Z^1_t, Z^2_t)dt \\
&= \varphi(Z^1_t, Z^2_t) - \varphi(Z^1_{t-}, Z^2_{t-}) \\
&= \varphi\left(Z^1_{t-} + \int_{\mathbb{R}_+} J(Z^1_{t-}, v)\wp(dt,dv), Z^2_{t-} + \int_{\mathbb{R}_+} J(Z^2_{t-}, v)\wp(dt,dv)\right) - \varphi(Z^1_{t-}, Z^2_{t-}) \\
&= \int_{\mathbb{R}_+} \left[\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v)) - \varphi(Z^1_{t-}, Z^2_{t-})\right]\wp(dt,dv) \\
&= \left(\int_{\mathbb{R}_+} \left[\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v)) - \varphi(Z^1_{t-}, Z^2_{t-})\right] dv\right) dt + dM_t
\end{aligned}$$

where M is the martingale obtained by integration wrt the compensated Poisson random measure $\wp(dt,dv) - dtdv$. We get the third equality by using the Theorem 0.2.20. For simplifying the above integral term, we impose (B3) and divide the derivation in two complementary cases.

Case 1: Assume $X^1_{t-} \neq X^2_{t-}$. Now under (B3), the intervals $\Lambda_{X^1_{t-}j_1}(Y^1_{t-}, N^1_{t-})$ and $\Lambda_{X^2_{t-}j_2}(Y^2_{t-}, N^2_{t-})$ are disjoint for every $j_1 \in \mathcal{X} \setminus \{X^1_{t-}\}, j_2 \in \mathcal{X} \setminus \{X^2_{t-}\}$ and N^1_{t-}, N^2_{t-}. Thus by considering these intervals where the integrand is non-zero constants, we get

$$
\begin{aligned}
&\int_{\mathbb{R}_+} \left[\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v)) - \varphi(Z^1_{t-}, Z^2_{t-})\right] dv \\
&= \int_{\bigcup_{k=1}^{2}\left(\bigcup_{j\in\mathcal{X}\setminus\{X^k_{t-}\}} \Lambda_{X^k_{t-}j}(Y^k_{t-}, N^k_{t-})\right)} \left[\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v)) - \varphi(Z^1_{t-}, Z^2_{t-})\right] dv \\
&= \sum_{j\in\mathcal{X}\setminus\{X^1_{t-}\}} \left[\varphi(j, 0, N^1_{t-} + 1, Z^2_{t-}) - \varphi(Z^1_{t-}, Z^2_{t-})\right] |\Lambda_{X^1_{t-}j}(Y^1_{t-}, N^1_{t-})| \\
&\quad + \sum_{j\in\mathcal{X}\setminus\{X^2_{t-}\}} \left[\varphi(Z^1_{t-}, j, 0, N^2_{t-} + 1) - \varphi(Z^1_{t-}, Z^2_{t-})\right] |\Lambda_{X^2_{t-}j}(Y^2_{t-}, N^2_{t-})|
\end{aligned}
$$

Case 2: Assume that $X^1_{t-} = X^2_{t-} = i$ say. Also recall that under (B3), the intervals $\Lambda_{ij}(y_1, n_1)$ and $\Lambda_{ij}(y_2, n_2)$ are having identical left end points for almost every $y_1, y_2 \geq 0$ and $n_1, n_2 \in \mathbb{N}_0$. So, $\Lambda_{X^1_{t-}j_1}(Y^1_{t-}, N^1_{t-})$ and $\Lambda_{X^2_{t-}j_2}(Y^2_{t-}, N^2_{t-})$ are not disjoint when $j_1 = j_2$. Therefore,

$$
\begin{aligned}
&\int_{\mathbb{R}_+} \left[\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v)) - \varphi(Z^1_{t-}, Z^2_{t-})\right] dv \\
&= \int_{\bigcup_{j\in\mathcal{X}\setminus\{i\}}\left(\Lambda_{ij}(Y^1_{t-}, N^1_{t-}) \cup \Lambda_{ij}(Y^2_{t-}, N^2_{t-})\right)} [\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v)) - \varphi(Z^1_{t-}, Z^2_{t-})] dv \\
&= \sum_{j\in\mathcal{X}\setminus\{i\}} [\varphi(j, 0, N^1_{t-} + 1, Z^2_{t-}) - \varphi(Z^1_{t-}, Z^2_{t-})] |\Lambda_{ij}(Y^1_{t-}, N^1_{t-}) \setminus \Lambda_{ij}(Y^2_{t-}, N^2_{t-})| \\
&\quad + \sum_{j\in\mathcal{X}\setminus\{i\}} [\varphi(Z^1_{t-}, j, 0, N^2_{t-} + 1) - \varphi(Z^1_{t-}, Z^2_{t-})] |\Lambda_{ij}(Y^2_{t-}, N^2_{t-}) \setminus \Lambda_{ij}(Y^1_{t-}, N^1_{t-})| \\
&\quad + \sum_{j\in\mathcal{X}\setminus\{i\}} [\varphi(j, 0, N^1_{t-} + 1, j, 0, N^2_{t-} + 1,) - \varphi(Z^1_{t-}, Z^2_{t-})] |\Lambda_{ij}(Y^1_{t-}, N^1_{t-}) \cap \Lambda_{ij}(Y^2_{t-}, N^2_{t-})|.
\end{aligned}
$$

Hence by combining the expressions under both the cases,

$$\begin{aligned}\mathcal{A}\varphi(z_1,z_2) &:= \left(\frac{\partial}{\partial y_1}+\frac{\partial}{\partial y_2}\right)\varphi(z_1,z_2)+\int_{\mathbb{R}_+}[\varphi(z_1+J(z_1,v),z_2+J(z_2,v))-\varphi(z_1,z_2)]\,dv \\ &= \left(\frac{\partial}{\partial y_1}+\frac{\partial}{\partial y_2}\right)\varphi(z_1,z_2) \\ &\quad+\sum_{j\in\mathcal{X}\setminus\{i_1\}}(\lambda_{i_1j}(y_1,n_1)-\delta_{i_1i_2}\lambda_{i_2j}(y_2,n_2))^+[\varphi(j,0,n_1+1,z_2)-\varphi(z_1,z_2)] \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.5.3) \\ &\quad+\sum_{j\in\mathcal{X}\setminus\{i_2\}}(\lambda_{i_2j}(y_2,n_2)-\delta_{i_1,i_2}\lambda_{i_1j}(y_1,n_1))^+[\varphi(z_1,j,0,n_2+1)-\varphi(z_1,z_2)] \\ &\quad+\delta_{i_1,i_2}\sum_{j\in\mathcal{X}\setminus\{i_1,i_2\}}(\lambda_{i_1j}(y_1,n_1)\wedge\lambda_{i_2j}(y_2,n_2))[\varphi(j,0,n_1+1,j,0,n_2+1)-\varphi(z_1,z_2)] \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.5.4)\end{aligned}$$

where $z_1=(i_1,y_1,n_1)$, $z_2=(i_2,y_2,n_2)$. This leads to the following Theorem.

Theorem 3.5.3. *Under (B3), the infinitesimal generator $\mathcal{A}$ of the augmented process $Z=(Z^1,Z^2)$ is given by* (3.5.3) *where $Z^1=(X^1,Y^1,N^1)$ and $Z^2=(X^2,Y^2,N^2)$ are as in Notation 3.5.2.*

The following notation is used in computing the meeting and merging probabilities throughout the chapter.

3.6 Meeting and Merging at the Next Transition

We follow Notation 3.5.2 throughout this section. By closely following Chapter 2, we define the meeting and merging related time instances and events below. We also recall the Definition 2.6.1 for the meeting, coherent meeting and merging related events and times.

Consider a specific case of the coupled integral equation (3.2.6)-(3.2.8) where the transition rate matrix λ is independent of the age variable y and satisfies (B1). Evidently, (B2) holds too. Furthermore, assume that $\tilde{\lambda}_{i'j'}(y,n)=\tilde{\lambda}_{i'j'}(n)$, a constant function for each $(i',j')\in\mathcal{X}_2$ and $n\geq 1$. Hence (3.2.6) and (3.2.8) reduces to

$$X_t = X_0+\int_{0^+}^t \tilde{h}(X_{u-},N_{u-},v)\wp(du,dv), \qquad (3.6.1)$$

$$N_t = \int_{0^+}^t \tilde{g}(X_{u-},N_{u-},v)\wp(du,dv), \qquad (3.6.2)$$

where $\tilde{h}(i,n,v) := h_\Lambda(i,y,n,v) = \sum_{j\in\mathcal{X}\setminus\{i\}}(j-i)\mathbb{1}_{\Lambda_{ij}(y,n)}(v)$, $\tilde{g}(i,n,v) := g_\Lambda(i,y,n,v) = \sum_{j\in\mathcal{X}\setminus\{i\}}\mathbb{1}_{\Lambda_{ij}(y,n)}(v)$ are constant in y, as the intervals $\Lambda_{ij}(y,n)$, defined using $\tilde{\lambda}_{ij}(y,n)$ do not vary with y variable but depends on n. It is evident that the strong solution of (3.6.1)-(3.6.2) gives a continuous time non-homogeneous Markov chain X on $\mathcal{X}$. In view of Theorem 3.2.2, the age of the Markov process at time t is given by $Y_t := t - T_{N_t}$, where $T_0 = 0$ and $\{T_n\}_{n\geq 1}$ denotes the consecutive transition times of X.

Theorem 3.6.1. *Let (X^1, N^1) and (X^2, N^2) be strong solutions of SDE (3.6.1)-(3.6.2) with initial conditions $X^1_0 = i$ and $X^2_0 = j$ respectively. A coherent meeting of X^1 and X^2, is a merging event.*

Proof. We will make use of the comments in Remark 0.2.19 regarding integration wrt PRM for each sample point $\omega \in \Omega$. For a $\omega \in \Omega$, if there exists a $t' > 0$ such that $X^1_{t'} = X^2_{t'} = k$ and also if $N^1_{t'} = N^2_{t'} = n$, for some $k \in \mathcal{X}$, $n \in \mathbb{N}_0$ then using (3.6.1)-(3.6.2), both X^1 and X^2 solve

$$X_t = X_{t'} + \int_{t'+}^{t} \tilde{h}(X_{u-}, N_{u-}, v)\wp(du, dv) = k + \int_{t'+}^{t} \tilde{h}(X_{u-}, N_{u-}, v)\wp(du, dv)$$
$$N_t = N_{t'} + \int_{t'+}^{t} \tilde{g}(X_{u-}, N_{u-}, v)\wp(du, dv) = n + \int_{t'+}^{t} \tilde{g}(X_{u-}, N_{u-}, v)\wp(du, dv)$$

for all $t > t'$. Now using almost sure uniqueness of the strong solution of the above SDE, (X^1, N^1) and (X^2, N^2) would be identical from time t' onward. Thus (X^1, N^1) and (X^2, N^2) merge at time t'. □

It is interesting to note that if they meet at t' and $N^1_{t'} \neq N^2_{t'}$, then t' cannot be assured as the time of merging for a given arbitrary rate parameter. Indeed for certain choice of parameters, the probability of t' being merging time is zero. We produce an example below.

Example 3.6.2. *Let $\mathcal{X} = \{1, 2\}$, with $(1,2) \prec_2 (2,1)$; also $\lambda_{12}(y,n) = 1 + r(n)$ and $\lambda_{21}(y,n) = 1$, where $r(n)$ is 0 or 1 if n is even or odd respectively. Thus $\tilde{\lambda}_{12}(y,n) = 2$ and $\tilde{\lambda}_{21}(y,n) = 1$ for all $y \geq 0$. Hence, for every $n \in \mathbb{N}_0$, $\Lambda_{12}(y,n) = [0, 1 + r(n))$ and $\Lambda_{21}(y,n) = [2,3)$. We further assume that $Z^l = (X^l, N^l)$ is the strong solution of* (3.6.1)-(3.6.2) *with above parameters and initial conditions $X^l_0 = l$ for $l = 1,2$ respectively. Now for the purpose of illustration, fix a sample $\omega \in \Omega$ such that $\wp(\omega)|_{[0,3]\times[0,3]} = \delta_{(1,5/2)} + \delta_{(3/2,3/2)}$, the addition of two Dirac measures at $(1, 5/2)$ and $(3/2, 3/2)$ respectively. Then none of the processes has transition before time $t = 1$. Hence, for both $l = 1, 2$, $X^l_{1-} = l$, and $N^l_{1-} = 0$. Then from* (3.6.1)

$$X^l_1 = X^l_{1-} + \int_{\mathbb{R}} \tilde{h}(X^l_{1-}, N^l_{1-}, v)\wp(\{1\}, dv)(\omega) = l + \tilde{h}(l, 0, 5/2).$$

Therefore, using (3.2.4) *and the intervals $\Lambda_{12}(Y^1_{1-}, 0), \Lambda_{21}(Y^2_{1-}, 0)$, we get $X^1_1 = 1 + (2-1)\mathbb{1}_{[0,1+r(0))}(5/2) = 1$ and $X^2_1 = 2 + (1-2)\mathbb{1}_{[2,3)}(5/2) = 1$. Hence, $t = 1$ is a meeting time.*

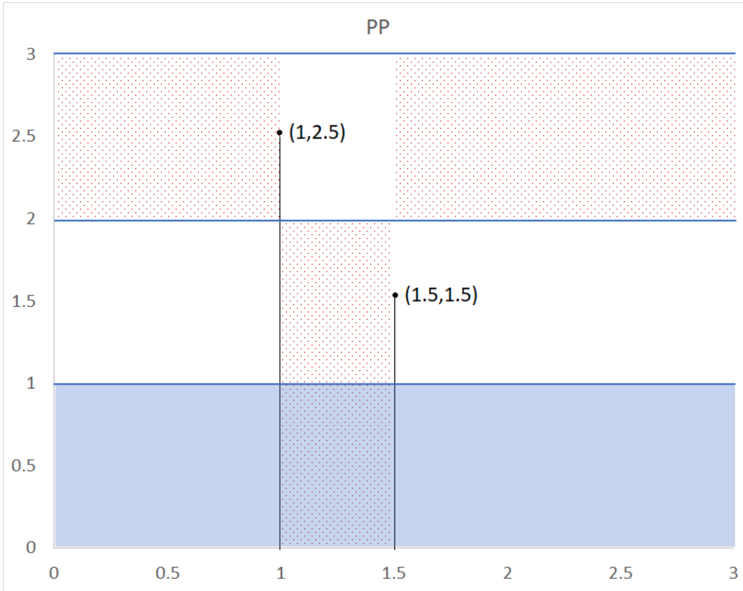


Figure 3.2: The t and v variables are plotted along horizontal and vertical axes. The point masses are shown by black dots. The intervals relevant for transitions of the first and second processes are plotted vertically and shown in blue and red respectively.

However, using (3.2.5)*, we get* $N_1^1 = 0 \neq 1 = N_1^2$ *and so merging at* $t = 1$ *is not guaranteed (Theorem 3.6.1). Indeed, at* $t = 3/2$*,* X^1 *and* X^2 *separate, which is shown below. We note that for* $t = 3/2$*, the pre-transition state* $X^l_{3/2-}$ *is* 1 *for each* $l = 1, 2$*, on the other hand* $N^1_{3/2-}$*, and* $N^2_{3/2-}$ *are* 0 *and* 1 *respectively. Consequently,*

$$X^l_{3/2} = 1 + \int_{\mathbb{R}} \tilde{h}(1, N^l_{3/2-}, v)\wp(\{3/2\}, dv)(\omega) = 1 + \mathbb{1}_{\Lambda_{12}(Y^l_{3/2-}, N^l_{3/2-})}(3/2) = \begin{cases} 1, & \text{for } l = 1 \\ 2, & \text{for } l = 2 \end{cases}$$

since, $3/2 \notin \Lambda_{12}(Y^1_{3/2-}, 0) = [0, 1 + r(0)) = [0, 1)$ *and* $3/2 \in \Lambda_{12}(Y^2_{3/2-}, 1) = [0, 1 + r(1)) = [0, 2)$*. Therefore, the meeting time* $t = 1$ *is not a merging time for the above mentioned realisation. Indeed, half is the conditional probability of separation in the next transition after* $t = 1$ *given* $\wp(\omega)|_{[0,1]\times[0,3]} = \delta_{(1,5/2)}$*.*

Interestingly, due to the binary nature of the state space and the initial conditions, no meetings are coherent. In particular, if at any time t *both the chains are at state 1, then* N_t^1 *and* N_t^2 *are even and odd respectively. So, the probability of getting separated in the next transition is half. Hence, using the recurrence of state 1, the probability of merging, given the initial conditions, is zero.*

The above example clearly indicates that for the binary state space, conditions $X_0^1 \neq X_0^2$ and $N_0^1 = N_0^2 = 0$ imply that at any time t, N_t^1 and N_t^2 are either both odds or evens, provided $X_t^1 \neq X_t^2$. In other words, $X_t^1 \neq X_t^2$ implies $|N_t^1 - N_t^2|$ is even. That need not be the case when the state space is not binary. In particular, for a non-binary state space, an event $\{X_t^1 \neq X_t^2\} \cap \{|N_t^1 - N_t^2| = 1\}$ may occur for some $t > 0$ with a positive probability. Given this event the conditional probability of coherent meeting at the next transition may be positive. Finally a coherent meeting may become a merging.

We provide with a lower bound of this probability in the following Theorem.

Theorem 3.6.3. *Let* (X^1, N^1) *and* (X^2, N^2) *be strong solutions of SDE* (3.6.1)-(3.6.2) *with initial conditions as in* (3.5.1)-(3.5.2) *and* $i \neq j$.

1. *The conditional probability of meeting in the next transition given* $\mathcal{F}_s$ *is* $\frac{(\lambda_{ij}(n_1)+\lambda_{ji}(n_2))}{\lambda_i(n_1)+\lambda_j(n_2)}$.
2. *If* $|n_2 - n_1| = 1$, *then the conditional probability of merging in next transition given* $\mathcal{F}_s$ *is at least*

$$\alpha(i, n_1, j, n_2) = \begin{cases} \frac{\lambda_{ij}(n_1)}{\lambda_i(n_1)+\lambda_j(n_2)} & \text{if } n_1 + 1 = n_2, \\ \frac{\lambda_{ji}(n_2)}{\lambda_i(n_1)+\lambda_j(n_2)} & \text{if } n_1 = n_2 + 1. \end{cases} \tag{3.6.3}$$

The first part of Theorem 3.6.3 can be viewed as a generalization of Theorem 2.7.1(1) of Chapter 2, where two solutions of homogeneous Markov case have been studied. On the other hand this part is a special case when we would derive an expression of the meeting probability in the next transition for the non-homogeneous semi-Markov case in the following theorem. Following theorem not only extends the first part of Theorem 3.6.3 but also Theorem 2.6.4 of Chapter 2. We present a proof of both parts of Theorem 3.6.3 after the proof of Theorem 3.6.4. The proof of Theorem 3.6.4 is very close to that of Theorem 2.6.4 of Chapter 2 in spirit. However, since the model assumptions are very different, we produce the proof with full details.

Theorem 3.6.4. *Assume (B1)-(B3). Let* $Z^1 = (X^1, Y^1, N^1)$ *and* $Z^2 = (X^2, Y^2, N^2)$ *be as in* (3.5.1) *and* (3.5.2) *respectively where* $i \neq j$. *The conditional probability that* X^1 *and* X^2 *meet in the next transition given* $\mathcal{F}_s$ *is*

$$\int_0^\infty e^{-\int_0^y (\lambda_i(y_1+t,n_1)+\lambda_j(y_2+t,n_2))dt} (\lambda_{ij}(y_1 + y, n_1) + \lambda_{ji}(y_2 + y, n_2))dy.$$

Proof. We recall Notation 3.5.2 regarding the initial conditions and the collective transition time sequence of X^1 and X^2. That is the processes X^1, X^2 at time s are at states i, j with ages y_1, y_2 and having transited n_1, n_2 times respectively. The event of no meeting in the next transition of X^1 and X^2, has two following sub-cases.
Case 1: After time s, but before X^2 transits, X^1 transits to a state different from X_s^2. This event can be written as $\mathcal{E} = \{X^1_{T^1_{n_1+1}} \neq X^2_{T^2_{n_2}}, T^1_{n_1+1} < T^2_{n_2+1}\}$. We will make use of

$$P(\mathcal{E} \mid \mathcal{F}_s) = E[P(\mathcal{E} \mid \mathcal{F}_s, T^1_{n_1+1} \wedge T^2_{n_2+1}) \mid \mathcal{F}_s]$$

and, the expression of conditional density $\eta_{T^1_{n_1+1}\wedge T^2_{n_2+1}}$ of $T^1_{n_1+1}\wedge T^2_{n_2+1}$ given $\mathcal{F}_s$. Clearly, $P(\mathcal{E}\mid T^1_{n_1+1}\wedge T^2_{n_2+1}=s+y,\mathcal{F}_s)$ is

$$\frac{m_1(\underset{k\in\mathcal{X}\setminus\{i,j\}}{\cup}\Lambda_{ik}(y_1+y,n_1))}{m_1(\underset{k\in\mathcal{X}\setminus\{i\}}{\cup}\Lambda_{ik}(y_1+y,n_1)\cup\underset{k\in\mathcal{X}\setminus\{j\}}{\cup}\Lambda_{jk}(y_2+y,n_2))}=\frac{\lambda_i(y_1+y,n_1)-\lambda_{ij}(y_1+y,n_1)}{\lambda_i(y_1+y,n_1)+\lambda_j(y_2+y,n_2)}.$$

Moreover $\eta_{T^1_{n_1+1}\wedge T^2_{n_2+1}}(s+y)=e^{-m_2(B)}(\lambda_i(y_1+y,n_1)+\lambda_j(y_2+y,n_2))$, where

$$B:=\underset{t\in[0,y)}{\cup}\left(\{s+t\}\times\left((\underset{k\in\mathcal{X}\setminus\{i\}}{\cup}\Lambda_{ik}(y_1+t,n_1))\cup(\underset{k\in\mathcal{X}\setminus\{j\}}{\cup}\Lambda_{jk}(y_2+t,n_2))\right)\right).$$

Indeed, given $\mathcal{F}_s$, the event of no transition of X^1 and X^2 from s to $s+y$ unit of time, is equivalent to $\{\wp(B)=0\}$, the non-occurence of Poission point mass in B. Clearly, $P(\wp(B)=0)$ is equal to $e^{-m_2(B)}$ and $m_2(B)=\int_0^y(\lambda_i(y_1+t,n_1)+\lambda_j(y_2+t,n_2))dt$. Hence the probability of the event in Case 1 is

$$\begin{aligned}P(\mathcal{E}\mid\mathcal{F}_s)&=\int_0^\infty P(\mathcal{E}\mid\mathcal{F}_s,T^1_{n_1+1}\wedge T^2_{n_2+1}=s+y)\eta_{T^1_{n_1+1}\wedge T^2_{n_2+1}}(s+y)dy\\&=\int_0^\infty e^{-\int_0^y(\lambda_i(y_1+t,n_1)+\lambda_j(y_2+t,n_2))dt}[\lambda_i(y_1+y,n_1)-\lambda_{ij}(y_1+y,n_2)]dy.\end{aligned}\tag{3.6.4}$$

Case 2: Similar to Case-1, the concerned event is the transition of X^2 before X^1 to a state, different from X^1_s. The conditional probability of this event given $\mathcal{F}_s$ is written below

$$\begin{aligned}&P(X^2_{T^2_{n_2+1}}\neq X^1_{T^1_{n_1}},T^2_{n_2+1}<T^1_{n_1+1}\mid\mathcal{F}_s)\\&=\int_0^\infty e^{-\int_0^y(\lambda_i(y_1+t,n_1)+\lambda_j(y_2+t,n_2))dt}[\lambda_j(y_2+y,n_2)-\lambda_{ji}(y_2+y,n_2)]dy.\end{aligned}\tag{3.6.5}$$

Hence the total probability (denoted by $a'_{(z_1,z_2)}$) of not meeting in the next transition is sum of the probabilities appearing in (3.6.4), and (3.6.5), where $z_1=(i,y_1,n_1),z_2=(j,y_2,n_2)$. Using

$$\begin{aligned}\phi_1(y)&:=e^{-\int_0^y(\lambda_i(y_1+t,n_1)+\lambda_j(y_2+t,n_2))dt}\left(\lambda_i(y_1+y,n_1)+\lambda_j(y_2+y,n_2)\right),\\a'_{(z_1,z_2)}&=\int_0^\infty\left(\phi_1(y)-e^{-\int_0^y(\lambda_i(y_1+t,n_1)+\lambda_j(y_2+t,n_2))dt}\left[\lambda_{ij}(y_1+y,n_1)+\lambda_{ji}(y_2+y,n_2)\right]\right)dy\qquad(3.6.6)\\&=1-\int_0^\infty e^{-\int_0^y(\lambda_i(y_1+t,n_1)+\lambda_j(y_2+t,n_2))dt}[\lambda_{ij}(y_1+y,n_1)+\lambda_{ji}(y_2+y,n_2)]dy\end{aligned}$$

as $\int_0^\infty\phi_1(y)dy=1$. Since the probability of meeting of X^1 and X^2 in the next transition is $1-a'_{(z_1,z_2)}$, the result follows from the above expression of a'. □

Proof. [of Theorem 3.6.3] We recall the initial condition at time s from Notation 3.5.2. As a direct application of Theorem 3.6.4 the conditional probability of meeting in the next transition of X^1 and X^2, given $\mathcal{F}_s$ can be written as

$$\begin{aligned}&\int_0^\infty e^{-\int_0^y(\lambda_i(n_1)+\lambda_j(n_2))dt}(\lambda_{ij}(n_1)+\lambda_{ji}(n_2))dy\\&=\frac{(\lambda_{ij}(n_1)+\lambda_{ji}(n_2))}{(\lambda_i(n_1)+\lambda_j(n_2))}\int_0^\infty e^{-(\lambda_i(n_1)+\lambda_j(n_2))y}(\lambda_i(n_1)+\lambda_j(n_2))dy\\&=\frac{(\lambda_{ij}(n_1)+\lambda_{ji}(n_2))}{(\lambda_i(n_1)+\lambda_j(n_2))}\end{aligned}$$

as the integral on the right side is 1. Therefore, the part (1) is true.

Next we wish to compute the probability of an event where merging occurs in the next transition given that $X_s^1=i, N_s^1=n_1, X_s^2=j, N_s^2=n_2$. We are also provided with $i\neq j$ and $|n_2-n_1|=1$. Merging in the next transition is guaranteed if that becomes a coherent meeting (see Theorem 3.6.1). We first consider the case $n_2=n_1+1$. Given this, for a coherent meeting, X^1 must transit before X^2 does after the time s, and the transition must be to the state j. Indeed, then at time $t=T^1_{n_1+1}$, X^1 and X^2 meet with $N_t^1=n_1+1=n_2=N_t^2$, and they stay merged by Theorem 3.6.1. Following a very similar argument as in Theorem 3.6.4, the conditional probability of the above mentioned coherent meeting is computed as follows,

$$\begin{aligned}&\int_0^\infty P(\{X^1_{T^1_{n_1+1}}=X^2_{T^2_{n_2}}, T^1_{n_1+1}<T^2_{n_2+1}\}\mid \mathcal{F}_s, T^1_{n_1+1}\wedge T^2_{n_2+1}=s+y)\eta_{T^1_{n_1+1}\wedge T^2_{n_2+1}}(s+y)dy\\&=\int_0^\infty \frac{\lambda_{ij}(n_1)}{\lambda_i(n_1)+\lambda_j(n_2)}e^{-(\lambda_i(n_1)+\lambda_j(n_2))y}(\lambda_i(n_1)+\lambda_j(n_2))dy\\&=\frac{\lambda_{ij}(n_1)}{\lambda_i(n_1)+\lambda_j(n_2)}.\end{aligned}$$

Similarly, if $n_1=n_2+1$, the event $\{X^2_{T^2_{n_2+1}}=X^1_{T^1_{n_1}}, T^2_{n_2+1}<T^1_{n_1+1}\}$ having probability $\frac{\lambda_{ji}(n_2)}{\lambda_i(n_1)+\lambda_j(n_2)}$, implies merging in next transition due to Theorem 3.6.1. Hence the proof is complete for both the cases. □

Next we wish to compute the conditional probability of occurrence of merging in the next transition after time s, given (3.5.1)-(3.5.2), provided $i\neq j$ and $|n_2-n_1|=1$. We first consider the case $n_2=n_1+1$. In the proof of Theorem 3.6.4, we have seen that meeting at the next transition occurs in two ways. One of them is coherent where X^1 transits to j before X^2 leaves j after the time s. Indeed then at time $t=T^1_{n_1+1}$, $X_t^1=X_t^2$, $N_t^1=n_1+1=n_2=N_t^2$, $Y_t^1=0$, and $Y_t^2=y_2+(t-s)$. The conditional probability of this coherent meeting given $\mathcal{F}_s$ is

$$\int_0^\infty e^{-\int_0^y(\lambda_i(y_1+u,n_1)+\lambda_j(y_2+u,n_2))du}\lambda_{ij}(y_1+y,n_1)dy, \tag{3.6.7}$$

whose derivation is analogous with that in Case 1 of Theorem 3.6.4. This meeting time becomes merging time only if subsequently, both the chains transit simultaneously to an identical state. To see this, we first note that non-simultaneous transition separates. On the other hand, in addition to identical N values, at the time of simultaneous transition, state and age variables of both the chains also become identical. Consequently, they merge due to the uniqueness result of (3.2.6)-(3.2.8) (see Theorem 3.2.2(1)). An expression (see Theorem 2.6.6, Chapter 2) of the conditional probability of such simultaneous transition given the fact that X^1 has already transited to j at $s+y$ before X^2 could leave j after s, is

$$\int_0^\infty e^{-\int_0^{y'} \sum_{j'\neq j} (\lambda_{jj'}(u,n_2)\vee\lambda_{jj'}(y_2+y+u,n_2))du} \left[\sum_{j'\neq j} \lambda_{jj'}(y',n_2)\wedge\lambda_{jj'}(y_2+y+y',n_2)\right] dy' \quad (3.6.8)$$

and denoted by $\mathcal{P}(j, y_2+y, n_2)$. Thus the probability of the combined event given $\mathcal{F}_s$ is

$$\int_0^\infty \mathcal{P}(j, y_2+y, n_2) e^{-\int_0^y (\lambda_i(y_1+u,n_1)+\lambda_j(y_2+u,n_2))du} \lambda_{ij}(y_1+y, n_1) dy. \quad (3.6.9)$$

This gives a lower bound of the merging probability given $\mathcal{F}_s$, should $i \neq j$ and $n_2 = n_1+1$. Analogously, if $n_1 = n_2+1$, an expression of a lower bound of merging probability similar to (3.6.9) can be obtained with two modifications. First of all, in view of the requirement of second chain's transition to the state of the first, the last multiplicative term in the integrand of (3.6.9) should be replaced by $\lambda_{ji}(y_2+y, n_2)$. Secondly, the variables j, y_2 and n_2 should be replaced by i, y_1 and n_1 in (3.6.8) to get $\mathcal{P}(i, y_1+y, n_1)$. This gives the conditional probability of simultaneous transition given the fact that X^2 has already transited to i at $s+y$ before X^1 could leave i after s. Thus we have proved the following result.

Theorem 3.6.5. *Assume (B1)-(B3) and that X^1 and X^2 are as in Notation 3.5.2. If $|n_2 - n_1| = 1$ and $i \neq j$, then the conditional probability that X^1 and X^2 merge in the next transition given $\mathcal{F}_s$ is at least*

$$\alpha(z_1,z_2) = \begin{cases} \int_0^\infty \mathcal{P}(j, y_2+y, n_2) e^{-\int_0^y (\lambda_i(y_1+u,n_1)+\lambda_j(y_2+u,n_2))du} \lambda_{ij}(y_1+y, n_1) dy & \text{if } n_1+1 = n_2, \\ \int_0^\infty \mathcal{P}(i, y_1+y, n_1) e^{-\int_0^y (\lambda_i(y_1+u,n_1)+\lambda_j(y_2+u,n_2))du} \lambda_{ji}(y_2+y, n_2) dy & \text{if } n_1 = n_2+1. \end{cases}$$

The above theorem should be viewed as an extension of part (2) of Theorem 3.6.3 to the case of non-homogeneous semi-Markov processes. To see this, we first note that for Markov special case $\mathcal{P}$ term is 1, as λ does not vary with the age variable. Finally, for the same reason the integration in the above expression of α can easily be calculated and matched with the expression in Theorem 3.6.3. A more precise assertion under an additional assumption is stated below.

Theorem 3.6.6. *Assume (B1)-(B3) and that X^1 and X^2 are as in Notation 3.5.2. Also assume that the infimum $\inf\{\|\lambda_{i'}(\cdot,n_1') - \lambda_{i'}(\cdot,n_2')\|_{L^\infty(\mathbb{R}_+)} \mid n_1'-n_2' \neq 0, i' \in \mathcal{X}\}$ is positive. If $i \neq j$, then the conditional probability that X^1 and X^2 merge in the next transition given $\mathcal{F}_s$ is $\alpha(z_1,z_2)$, provided $|n_2-n_1| = 1$, and zero otherwise.*

Proof. From the proof of Theorem 3.6.5, we know that $\alpha(z_1, z_2)$ gives the conditional probability of occurrence of a coherent meeting which is also merging given $\mathcal{F}_s$ with $|n_2 - n_1| = 1$ and $i \neq j$. Thus it is enough to prove that a meeting is not merging if it is not coherent, provided infimum of $\|\lambda_{i'}(\cdot, n_1') - \lambda_{i'}(\cdot, n_2')\|_{L^\infty(\mathbb{R}_+)}$ over all $i' \in \mathcal{X}$, and $n_1' \neq n_2'$ is positive.

Similar to (3.6.7), the conditional probability of non-coherent meeting in the next transition can be expressed. That meeting becomes merging if the pair never separate. We also recall that (B2) implies that the chains must transit infinitely many times almost surely. Therefore, if the pair never separate, they must transit simultaneously to the identical states infinitely many often and the difference of N values stay unchanged in successive transitions. At the time of non-coherent meeting, $N^1 - N^2 (\neq 0)$ takes value, due to the initial conditions, $n_2 - n_1 - 1$ (or $n_2 - n_1 + 1$) if the first (or the second) chain transits to another's state before the other one leaves.

Assume that the next transition is non-coherent meeting, led by transition of the second chain before the first's departure. An expression similar to (3.6.8) of the conditional probability of having $m+1$th transition, after s, non-separated given the fact that the previous transition has been simultaneous and the common state has been i', is

$$\int_0^\infty e^{-\int_0^{y'} \sum_{j' \neq i'} (\lambda_{i'j'}(u, n_2+m) \vee \lambda_{i'j'}(u, n_1+m-1))du} \left[\sum_{j' \neq i'} \lambda_{i'j'}(y', n_2+m) \wedge \lambda_{i'j'}(y', n_1+m-1) \right] dy' \tag{3.6.10}$$

for every $m \geq 2$. As $(n_2+m) - (n_1+m-1) = n_2 - n_1 + 1 \neq 0$, $\|\lambda_{i'}(\cdot, n_1+m-1) - \lambda_{i'}(\cdot, n_2+m)\|_{L^\infty(\mathbb{R}_+)}$ is nonzero. Consequently, (3.6.10) is strictly smaller than 1, since

$$\sum_{j' \neq i'} (\lambda_{i'j'}(\cdot, n_2+m) \vee \lambda_{i'j'}(\cdot, n_1+m-1)) - \sum_{j' \neq i'} \lambda_{i'j'}(\cdot, n_2+m) \wedge \lambda_{i'j'}(\cdot, n_1+m-1)$$

is positive and its $L^\infty(\mathbb{R}_+)$ norm is bounded away from zero. This implies that the conditional probability of repeated occurrence of non-separation, given the transition following s is a non-coherent meeting, is zero. Thus we have proved that the meeting is not a merging with probability 1 if it is non-coherent. □

3.7 Eventual Meeting, Merging, and Time

Having obtained in Theorem 3.6.4, an expression of conditional probability of meeting in the next transition for a general semi-Markov case, we next find a sufficient condition for sure occurrence of eventual meeting. Before addressing this for semi-Markov case, we first investigate the Markov special case in the next Theorem. The proof of this result is in the similar line of proof of Theorem 2.6.4(2) of Chapter 2, where the result is stated for homogeneous Markov case.

Theorem 3.7.1. *Assume (B1)-(B3), Let* (X^1, N^1) *and* (X^2, N^2) *be strong solutions of SDE* (3.6.1)-(3.6.2) *with two different initial conditions. If* $\mathcal{X} = \{1, 2, \ldots, k\}$ *is finite and* $\inf_{i,j,n_1,n_2} \frac{(\lambda_{ij}(n_1)+\lambda_{ji}(n_2))}{\lambda_i(n_1)+\lambda_j(n_2)} > 0$*, then* X^1 *and* X^2 *eventually meet with probability 1.*

Proof. Let $s > 0$ be the initial time. Since $\inf_{i,j,n_1,n_2} \frac{(\lambda_{ij}(n_1)+\lambda_{ji}(n_2))}{\lambda_i(n_1)+\lambda_j(n_2)} > 0$, the supremum of $a_{(i,j,n_1,n_2)} := 1 - \frac{(\lambda_{ij}(n_1)+\lambda_{ji}(n_2))}{\lambda_i(n_1)+\lambda_j(n_2)}$ is less than 1. From Theorem 3.6.3(1) we know that $a(i, j, n_1, n_2)$ denotes the conditional probability of not meeting in the next transition. We also recall that $\{\tau^{(n)}(s)\}$ gives a sequence of stopping times, representing the transitions of combined process (X^1, X^2). Now using Theorem 3.3.2, we get

$$E\left[\mathbb{1}_{\{X^1_{\tau^{n+1}(s)} \neq X^2_{\tau^{(n+1)}(s)}\}} \mid \mathcal{F}_{\tau^{(n)}(s)}\right] = a_{(X^1_{\tau^{(n)}(s)}, X^2_{\tau^{(n)}(s)}, N^1_{\tau^{(n)}(s)}, N^2_{\tau^{(n)}(s)})} \leq \sup_{i,j,n_1,n_2} a_{(i,j,n_1,n_2)} < 1. \tag{3.7.1}$$

Therefore, using (3.7.1), we get

$$\begin{aligned}
&E\left[\prod_{n=1}^{m} \mathbb{1}_{\{X^1_{\tau^{(n)}(s)} \neq X^2_{\tau^{(n)}(s)}\}} \mid \mathcal{F}_s\right] \\
&= E\left[E\left[\prod_{n=1}^{m} \mathbb{1}_{\{X^1_{\tau^{(n)}(s)} \neq X^2_{\tau^{(n)}(s)}\}} \mid \mathcal{F}_{\tau^{(m-1)}(s)}\right] \mid \mathcal{F}_s\right] \\
&= E\left[\left(\prod_{n=1}^{m-1} \mathbb{1}_{\{X^1_{\tau^{(n)}(s)} \neq X^2_{\tau^{(n)}(s)}\}}\right) E\left[\mathbb{1}_{\{X^1_{\tau^{(m)}(s)} \neq X^2_{\tau^{(m)}(s)}\}} \mid \mathcal{F}_{\tau^{(m-1)}(s)}\right] \mid \mathcal{F}_s\right] \\
&\leq \sup_{i,j,n_1,n_2} a_{(i,j,n_1,n_2)} E\left[\prod_{n=1}^{m-1} \mathbb{1}_{\{X^1_{\tau^{(n)}(s)} \neq X^2_{\tau^{(n)}(s)}\}} \mid \mathcal{F}_s\right]
\end{aligned}$$

for all $m \geq 1$. Using this repeatedly, we get for all $m \geq 1$

$$P\left(\cap_{n=1}^{m} \{X^1_{\tau^{(n)}(s)} \neq X^2_{\tau^{(n)}(s)}\} \mid \mathcal{F}_s\right) = E\left[\prod_{n=1}^{m} \mathbb{1}_{\{X^1_{\tau^{(n)}(s)} \neq X^2_{\tau^{(n)}(s)}\}} \mid \mathcal{F}_s\right] \leq \left(\sup_{i,j,n_1,n_2} a_{(i,j,n_1,n_2)}\right)^m. \tag{3.7.2}$$

The left side value is the probability of the event of not meeting of processes X^1 and X^2 till time $\tau^{(m)}(s)$. Hence, using the fact (thanks to (B2)) that the chains experience infinitely many transitions with probability 1, the probability of never meeting, $P(X^1_t \neq X^2_t, \forall t \geq s \mid \mathcal{F}_s)$ is the limit of left side of (3.7.2) as $m \to \infty$. Since, right side of (3.7.2) vanishes as $m \to \infty$, $P(X^1_t \neq X^2_t, \forall t \geq s \mid \mathcal{F}_s)$ is zero as desired. □

Remark 3.7.2. *In Theorem 3.7.1, we assume* $\mathcal{X}$ *is finite and* $\inf_{i,j,n_1,n_2} \frac{(\lambda_{ij}(n_1)+\lambda_{ji}(n_2))}{\lambda_i(n_1)+\lambda_j(n_2)} > 0$ *in addition to (B1)-(B3) for the SDE* (3.6.1) - (3.6.2)*. Below, in Lemma 3.7.3, by considering* *a homogeneous Markov special case, we show that under (B1) the positivity of the infimum indeed implies finiteness of $\mathcal{X}$. Hence, a separate mention of finiteness of $\mathcal{X}$ in Theorem 3.7.1 is redundant.*

Lemma 3.7.3. *Let λ be a transition rate matrix obeying (B1) and being constant in n and y variables. If $\mathcal{X}$ is infinite, $\inf_{i,j} \frac{(\lambda_{ij}+\lambda_{ji})}{\lambda_i+\lambda_j}$ is zero.*

Proof. Given $\epsilon > 0$, for each $i \in \mathcal{X}$, there is a $j^*_{i,\epsilon} \in \mathcal{X}$ such that $\frac{\lambda_{ij}}{\lambda_i} < \epsilon$ for all $j \geq j^*_{i,\epsilon}$. This is due to the fact that $\sum_{j \in \mathcal{X}\setminus\{i\}} \lambda_{ij} < \infty$. Hence using $\lambda_i > 0, \lambda_j > 0$,

$$\frac{\lambda_{ij}+\lambda_{ji}}{\lambda_i+\lambda_j} < \frac{\lambda_{ij}}{\lambda_i} + \frac{\lambda_{ji}}{\lambda_j} < \epsilon + \frac{\lambda_{ji}}{\lambda_j}$$

for all $j \geq j^*_{i,\epsilon}$ and for all $i \in \mathcal{X}$. Therefore,

$$\inf_{i,j} \frac{\lambda_{ij}+\lambda_{ji}}{\lambda_i+\lambda_j} < \epsilon + \inf_i \inf_{j \geq j^*_{i,\epsilon}} p_{ji} \tag{3.7.3}$$

where $p_{ji} := \frac{\lambda_{ji}}{\lambda_j}$ gives the probability of transition from j to i. Let if possible there is a $n \in \mathbb{N}$ such that $\inf_i \inf_{j \geq j^*_{i,\epsilon}} p_{ji} > \frac{1}{n}$. Then we choose $i_1, i_2, \ldots, i_{n+1}$ distinctly from $\mathcal{X}$ and set

$$j^* = \max_{1 \leq k \leq n+1} j^*_{i_k,\epsilon}.$$

Then $p_{j^* i_k} > \frac{1}{n}$ $\forall k \leq n+1$. Hence contradiction. Thus $\inf_i \inf_{j \geq j^*_{i,\epsilon}} p_{ij} = 0$. By applying this to (3.7.3) and by noting that ϵ is an arbitrary positive number, we conclude that $\inf_{i,j} \frac{\lambda_{ij}+\lambda_{ji}}{\lambda_i+\lambda_j} = 0$. □

From Theorem 3.7.1 (2), one can infer that a suitably extended but a similar condition on the rate matrix may guarantee an eventual meeting event for the semi-Markov family. One such condition is presented below in (B4).

(B4) $\mathcal{X} = \{1, 2, \ldots, k\}$ is a finite state space, and

$$\sup_{(i,j) \in \mathcal{X}_2, y_1, y_2, n_1, n_2 \geq 0} \left\| 1 - \frac{(\lambda_{ij}(y_1 + \cdot, \cdot) + \lambda_{ji}(y_2 + \cdot, \cdot))}{\lambda_i(y_1 + \cdot, \cdot) + \lambda_j(y_2 + \cdot, \cdot)} \right\|_{L^\infty} < 1.$$

Theorem 3.7.4. *Assume (B1)-(B4) and that X^1 and X^2 are as in Notation 3.5.2. Then X^1 and X^2 eventually meet with probability 1.*

Proof. Using $\phi_2(y) := 1 - \frac{(\lambda_{ij}(y_1+y,n_1)+\lambda_{ji}(y_2+y,n_2))}{\lambda_i(y_1+y,n_1)+\lambda_j(y_2+y,n_2)}$, and Hölder inequality, from (3.6.6) we get

$$a'_{(z_1,z_2)} = \int_0^\infty \phi_1(y)\phi_2(y)dy \leq \|\phi_1\|_{L^1}\|\phi_2\|_{L^\infty} = \|\phi_2\|_{L^\infty}.$$

Now by a direct application of (B4), we get that supremum of $\|\phi_2\|_{L^\infty}$ over all $(i,j) \in \mathcal{X}_2, y_1 \geq 0, y_2 \geq 0$ is less than 1, which implies that

$$\sup_{(i,j)\in\mathcal{X}_2,y_1,y_2,n_1,n_2\geq 0} a'_{(z_1,z_2)} < 1. \tag{3.7.4}$$

Again, the total probability of never meeting is the probability of intersection of non-occurrence of meeting in all transitions. On the other hand, (B2) ensures almost sure infinite transitions. Moreover, since (Z^1, Z^2) is strong Markov (Theorem 3.3.2) and $\{T_n := \tau^{(n)}(s)\}_{n\geq 1}$ are stopping times $P\left(\{X^1_{T_n} \neq X^2_{T_n}\} \mid \mathcal{F}_{T_{n-1}}\right) = a'_{(Z^1_{T_{n-1}},Z^2_{T_{n-1}})}$ which is upper bounded by a value less than 1 (see (3.7.4)). Therefore, in the similar line of the proof of Theorem 3.7.1, we get

$$E\left[\prod_{n=1}^m \mathbb{1}_{\{X^1_{T_n}\neq X^2_{T_n}\}} \mid \mathcal{F}_s\right] \leq \left(\sup_{(i,j)\in\mathcal{X}_2,y_1,y_2,n_1,n_2\geq 0} a'_{(z_1,z_2)}\right)^m. \tag{3.7.5}$$

Thus from (3.7.4) and (3.7.5), $P\left(X^1_t \neq X^2_t, \forall t \geq s \mid \mathcal{F}_s\right) = \lim_{N\to\infty} E\left[\prod_{n=1}^N \mathbb{1}_{\{X^1_{T_n}\neq X^2_{T_n}\}} \mid \mathcal{F}_s\right] = 0$. In other words, the probability of never meeting is zero. □

The above theorem asserts that the waiting time for the first meeting is a finite stopping time almost surely, provided (B1)-(B4) hold. A comprehensive study of its distribution is rather involved, even in the homogeneous Markov special case. However, the tail property of number of required transitions for meeting is considerably straight forward. We produce a result below which is an immediate extension of Theorem 2.7.4 of Chapter 2.

Theorem 3.7.5. *Assume (B1)-(B4) and that X^1 and X^2 are as in Notation 3.5.2. If $\bar{N}$ denotes the number of collective transitions, until X^1 and X^2 meet after time s, then $E[\bar{N}^r|\mathcal{F}_s]$ is finite for any $r \geq 1$.*

Proof. Since $\bar{N}$ denotes the number of collective transitions until the meeting after s, using (3.7.5), we get for all $n \geq 0$

$$\begin{aligned} P(\bar{N} = n+1|\mathcal{F}_s) =& E\left[\left(\prod_{r=1}^n \mathbb{1}_{\{X^1_{T_r}\neq X^2_{T_r}\}}\right) E\left(\mathbb{1}_{\{X^1_{T_{n+1}}=X^2_{T_{n+1}}\}} \mid \mathcal{F}_{T_n}\right) |\mathcal{F}_s\right] \\ \leq& \left(\sup_{z^1,z^2} a'_{(z^1,z^2)}\right)^n \left(1 - \inf_{z^1,z^2} a'_{(z^1,z^2)}\right), \end{aligned}$$

where $T_n := \tau^{(n)}(s)$, by following the convention that product and intersection of an empty family are 1 and empty set respectively. Thus the r^{th} raw moment, $E[\bar{N}^r|\mathcal{F}_s]$ is

$$\sum_{n=1}^{\infty} n^r P(\bar{N} = n|\mathcal{F}_s) \leq \left(1 - \inf_{z^1,z^2} a'_{(z^1,z^2)}\right)\left(\sup_{z^1,z^2} a'_{(z^1,z^2)}\right)^{-1} \sum_{n=1}^{\infty} n^r \left(\sup_{z^1,z^2} a'_{(z^1,z^2)}\right)^n.$$

The infinite series on the right converges provided $\sup_{z^1,z^2} a'_{(z^1,z^2)} < 1$ which is ensured in (3.7.4) due to the assumption (B4). Thus we conclude that $\bar{N}$ has finite moments. □

We note that if the pair of SMPs do not separate in the subsequent transition after meeting, they merge, provided either the meeting being coherent or the rate function $\lambda(y, n)$ is independent of n. This is because the state and age become identical in the subsequent simultaneous transition after meeting and hence merging is assured from the uniqueness of the driving SDE. In fact this phenomena has been used to prove Theorem 3.6.5 for coherent case and also the almost sure eventual merging in Theorem 2.7.5 of Chapter 2 for homogeneous case. However, for the non-homogeneous case, the event of subsequent few number of non-separations after a non-coherent meeting, does not imply a merging. Therefore, the argument of Theorem 2.7.5 of Chapter 2 cannot be extended directly. On the other hand calculation or estimation of the probability of eventual coherent meeting in any reasonable generality is hard. In this connection we recall that probability of coherent meeting is zero for the binary state-space case, (see Example 3.6.2). Despite these intricacies, it is not hard to propose some sufficient conditions under which the eventual merging is guaranteed. We explain this in the following remark.

Remark 3.7.6. *Imagine that the transition law λ is such that the conditional probability of an eventual coherent meeting given $\mathcal{F}_s$, is one. Further assume that the conditional probability of subsequent separation, given a coherent meeting has occurred, is upper bounded by a number less than 1. Then as the coherent meeting will occur repeatedly with probability one, the argument of Theorem 2.7.5 of Chapter 2 can be mimicked. Indeed the probability of repeated occurrence of separation following all consecutive coherent meeting becomes zero. In other words, the pair eventually merge almost surely.*

We have not expressed or estimated the conditional probability of coherent meeting in terms of λ. So, a sufficient condition, expressed algebraically, on λ is missing from the above comment. Coherent meeting in the next transition is possible only if the difference of the counts of transitions of both the chains is 1 prior to the meeting. This condition appears rather restrictive. However, if the sequence of functions $\{\lambda(\cdot, n)\}_n$ is periodic in n, there are other meetings which play the same role as the coherent meetings do in the above analysis. This allows one to extend the scope of all the results related to merging for the special case of periodic rates. Moreover, some assumptions on λ can also be simplified. The following remark explains these.

Remark 3.7.7. *Given a transition rate function λ, define $n^* \in \mathbb{N}_0$ as the smallest values such that $\{\lambda(\cdot, n)\}_n$ is periodic on $\{n \in \mathbb{N}_0 \mid n \geq n^*\}$. If $n^* < \infty$, let $\nu \in \mathbb{N}$ denote the* *periodicity. Note that if* $\lambda(\cdot, n)$ *is constant in* n*, then* $n^* = 0$*, and* $\nu = 1$*. On the other extreme, when* $\{\lambda(\cdot, n)\}_n$ *possesses no periodicity,* $n^* = +\infty$*, due to the convention that minimum of an empty set is* $+\infty$*. Next we list some observations below. Owing to the directness of their justification, we omit the proofs.*

1. *The assertions in the part (2) of Theorem 3.6.3 and in Theorem 3.6.5 still hold if the conditions* $n_1 + 1 = n_2$ *and* $n_2 + 1 = n_1$ *are extended as* $n_2 - n_1 = 1(\text{mod } \nu)$ *and* $n_1 - n_2 = 1(\text{mod } \nu)$ *respectively, provided* $n^* \le n_1, n_2 < \infty$*.*

2. *If* $n^* \le n_1, n_2 < \infty$*, the assertion of Theorem 3.6.6 holds true for a relaxed constraint on infimum, namely* $\inf\{\|\lambda_{i'}(\cdot, n_1') - \lambda_{i'}(\cdot, n_2')\|_{L^\infty(\mathbb{R}_+)} \mid n_1' - n_2' \neq 0(\text{mod } \nu), i' \in \mathcal{X}\}$ *is positive. In the statement the condition* $|n_2 - n_1| = 1$ *may also be replaced by* $|n_2 - n_1| = 1(\text{mod } \nu)$*.*

3. *The positivity condition on the infimum in Theorem 2.7.1, holds true if* $\mathcal{X}$ *is finite, the constant function* $\lambda(\cdot, n)$ *is positive for each* n*, and* n^* *is finite.*

4. *The supremum in (A4) is strictly less than 1 if* $\lambda(\cdot, n)$ *is bounded away from zero for each* n*, and* n^* *is finite.*

Remark 3.7.8. *Consider a pair of solutions* (X^1, N^1) *and* (X^2, N^2) *of SDE* (3.6.1)-(3.6.2) *with initial conditions* $X_0^1 = i$ *and* $X_0^2 = j$ *respectively. Clearly the discrete time process* $W := (W^1, W^2)$ *is a Markov chain, where* $W_n^l := (X_{T_n}^l, N_{T_n}^l)$*,* $T_n = \tau^n(0)$ *for each* $l = 1, 2$ *and* $n \ge 0$*. Now if the given* λ *is such that* n^* *and* ν *are finite, and* $n\%\nu$ *denotes* n *mod* ν*, using the map* $(i, n) \mapsto \pi(i, n) := (i, n \wedge n^* + \max(0, n - n^*)\%\nu) \in \mathcal{X} \times \{0, 1, \ldots, n^* + \nu - 1\}$*, we obtain a new bi-variate chain* $\bar{W} := (\bar{W}^1, \bar{W}^2)$ *where* $\bar{W}^l := \{\pi(W_n^l)\}_{n \in \mathbb{N}_0}$*. Under finiteness assumption on* $\mathcal{X}$*,* $\bar{W}$ *is a finite state homogeneous Markov chain. Furthermore, if eventual merging is an almost sure event, the states of* $\bar{W}$ *corresponding to the non-meeting instances of* (X^1, X^2) *are transient. If in addition,* $\inf\{|\lambda_{i'}(n_1') - \lambda_{i'}(n_2')| : \text{either } n_1' - n_2' \neq 0(\text{mod } \nu) \text{ or } n_1, n_2 \le n^*, \text{ and } i' \in \mathcal{X}\}$ *is positive, every state of* $\bar{W}$*, where* $\bar{W}^1$ *mismatch with* $\bar{W}^2$ *is transient, in view of Theorem 3.6.6. Thus for every initial condition, the set of transient states of* $\bar{W}$ *includes* $(\mathcal{X} \times \{0, 1, \ldots, n^* + \nu - 1\})^2 \setminus \{(i, n', i, n') \mid i \in \mathcal{X}, n' \ge n^*\}$*.*

3.8 Conclusion

In this chapter, we have explored a stochastic differential equation (SDE) representation of a broad category of semi-Markov processes on a countable state-space, where the kernel is differentiable and the underlying Markov chain may or may not be homogeneous. The system of SDEs is driven by a Poisson random measure with Lebesgue intensity. The coefficients are chosen depending on the given transition rate function and an additional gaping parameter. Since, the coefficients are not compactly supported and also the intensity measure is not finite, the existence result is not straightforward. We have first proved the local existence and then established the global existence of a unique strong solution. We then show that the state component of the solution is a pure semi-Markov process with the given transition rate function and the other two components are the age and transition count processes. Although the law of a single solution does not depend on the additional gaping parameter, the joint distribution of a couple of solutions with different initial conditions does depend on that. Under a simplified assumption on the gaping parameter, we derive the law of the bivariate process by calculating the infinitesimal generator of the augmented process.

To the best of our knowledge, the SDE under consideration or its any generalizations have not been studied in the literature. The approach of the proof of existence is significantly original. The SDE also gives a semi-Martingale representation of a semi-Markov process which need not be time-homogeneous. This representation is also not present in the literature. The detailed proof of the fact that the solution gives a semi-Markov process with desired transition kernel is valuable for all future study of this representation. This is a vital contribution of this chapter. Finally, the semimartingale representation has been used to generate a correlated semi-Markov system with multiple members.

In veiw of the immense applicability of the correlated semi-Markov system, the formulation and the results presented in this chapter are important. It is evident that SDE (3.2.6)-(3.2.8) generates a semi-Markov flow. We have studied the associated flow by investigating meeting and merging related events for solutions starting with two different initial conditions. The pair of the state components, thus obtained, form a component-wise semi-Markov process having dependent components. Some sufficient conditions for almost sure meeting and merging of the components have also been obtained. In next chapter we have illustrated and complemented some of the theoretical findings of this and earlier chapter by providing with several numerical examples.

Chapter 4

Numerical Examples

4.1 Numerical Results

4.1.1 Example of a Homogeneous SMP

Let $\mathcal{X} = \{1, 2\}$, with $(1, 2) \prec_2 (2, 1)$. Furthermore, assume for each $y \geq 0$ and $n \in \mathbb{N}_0$, $\lambda_{12}(y, n) = \frac{y}{1+y} = \lambda_{21}(y, n)$, and hence (A3) implies $\tilde{\lambda}_{12}(y, n) = 1 = \tilde{\lambda}_{21}(y, n)$, $\Lambda_{12}(y, n) = [0, \frac{y}{1+y})$, and $\Lambda_{21}(y, n) = [1, 1 + \frac{y}{1+y})$. Since, these intervals do not depend on n, (2.2.4)-(2.2.5) have a unique strong solution for a given initial condition and the X component constitutes a homogeneous SMP. Let (X^l, Y^l) denote a solution with $X^l_0 = l, Y^l_0 = 0$ for each $l = 1, 2$ respectively. If τ denotes the time of first meeting of X^1 and X^2, $P(\tau > t \mid \mathcal{F}_0) = P(\{T^1_1 > t\} \cap \{T^2_1 > t\} \mid \mathcal{F}_0)$. Since $\mathcal{F}_0 = \sigma\{X^1_0, X^2_0, Y^1_0, Y^2_0\}$, and $X^1_0 \neq X^2_0$, the events $\{T^1_1 > t\}$ and $\{T^2_1 > t\}$ are independent. Moreover, using the formula ye^{-y} of holding time density function at each state (see Chapter 2, Theorem 2.4.1(i) and $Y^1_0 = Y^2_0 = 0$, we get

$$P(\tau > t \mid \mathcal{F}_0) = \left(\int_t^\infty ye^{-y}dy\right)^2 = e^{-2t}(1+t)^2.$$

Thus the expected first meeting time $E(\tau \mid \mathcal{F}_0)$ is

$$\int_0^\infty P(\tau > t \mid \mathcal{F}_0)dt = \int_0^\infty e^{-2t}(1+t)^2 dt = \frac{-1}{4}e^{-2t}(2t^2 + 6t + 5)\Big|_0^\infty = 5/4. \tag{4.1.1}$$

Similarly the expected holding time at each state can be calculated as $\int_0^\infty y^2 e^{-y}dy = \Gamma(3) = 2$. It is important to note that this example does not satisfy the sufficient condition, as stated in Chapter 2, Theorem 2.7.5, for eventual merging of homogeneous semi-Markov flow. Indeed $\|\lambda_{12}^{-1}\|_{L^\infty} = \infty = \|\lambda_{21}^{-1}\|_{L^\infty}$. Nevertheless, the finiteness of norm of the reciprocal has been used in the proof of Chapter 2, Theorem 2.7.5 only for showing that $\mathcal{P}(k, y)$ is bounded away from zero. For this particular example, an exact expression of $\mathcal{P}(k, y, n)$ can be obtained and shown to be bounded away from zero. Indeed, from Chapter 2 Theorem 2.6.6, for all $y \geq 0$

$$\mathcal{P}(k, y) = \int_0^\infty e^{-\int_0^{y'} \frac{y+t}{1+y+t} dt} \frac{y'}{1+y'} dy' = 1 + \frac{y}{1+y} e \int_{-\infty}^{-1} \frac{e^t}{t} dt \geq 1 + e \int_{-\infty}^{-1} \frac{e^t}{t} dt \approx 0.403653.$$

Thus eventual merging happens almost surely. The combined process $((X^1, Y^1), (X^2, Y^2))$, being sampled 9×10^5 times using Monte Carlo simulation, gives the sample mean of merging time (For algorithm see chapter 4 section 2.1). That turns out to be 2.00, correct up to 2 decimal places. Using these samples the plots of the empirical probability density functions for meeting and merging times are also obtained and compared with the plot of the holding time density function in Figure 4.1.

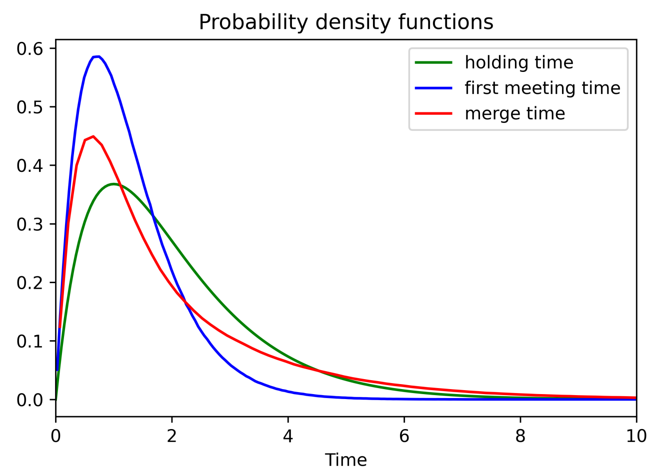


Figure 4.1: For the numerical example in subsection 4.1.1, true probability density function of holding time, and estimated probability density functions of first meeting time and merging time are plotted in green, blue and red respectively.

4.1.2 Example of a Non-Homogeneous Markov process

Assume $\mathcal{X} = \{1, 2, 3\}$, and the transition rate function for each $y \geq 0$ is given by

$$\lambda(y, n) = \begin{cases} 2A & \text{if } n \text{ is odd} \\ A & \text{if } n \text{ is even} \end{cases}; \text{ where, } A = \begin{pmatrix} -1 & \frac{2}{3} & \frac{1}{3} \\ 1 & -2 & 1 \\ \frac{1}{3} & \frac{2}{3} & -1 \end{pmatrix}. \tag{4.1.2}$$

n	$\Lambda_{12}(\cdot,n)$	$\Lambda_{13}(\cdot,n)$	$\Lambda_{21}(\cdot,n)$	$\Lambda_{23}(\cdot,n)$	$\Lambda_{31}(\cdot,n)$	$\Lambda_{32}(\cdot,n)$
Even	$\left[0,\frac{2}{3}\right)$	$\left[\frac{4}{3},\frac{5}{3}\right)$	$[2,3)$	$[4,5)$	$\left[6,\frac{19}{3}\right)$	$\left[\frac{20}{3},\frac{22}{3}\right)$
Odd	$\left[0,\frac{4}{3}\right)$	$\left[\frac{4}{3},2\right)$	$[2,4)$	$[4,6)$	$\left[6,\frac{20}{3}\right)$	$\left[\frac{20}{3},8\right)$

Table 4.1: For the numerical example in Subsection 4.1.2, the intervals $\Lambda_{ij}(\cdot,\cdot)$ are presented for both even and odd values of n and each $(i,j)\in\mathcal{X}_2$.

Then using (B3) and following the lexicographic ordering, we get the intervals, listed in Table 4.1. As the parameters do not depend on y variable, the intervals also do not. For this example, the SDE (3.6.1)-(3.6.2) has a unique solution for a given initial condition. The X component of the solution gives a non-homogeneous Markov process on $\{1,2,3\}$ with rate matrix λ as define in (4.1.2). In the fourth column of Table 4.2, we list the values of

Sr.No.	(X_0^1,X_0^2)	(N_0^1,N_0^2)	$\alpha(X_0^1,N_0^1,X_0^2,N_0^2)$	estimated value
1	(1,2)	(even, odd)	$2/15$	0.13
2	(1,2)	(odd, even)	$1/4$	0.25
3	(1,3)	(even, odd)	$1/9$	0.11
4	(1,3)	(odd, even)	$1/9$	0.11
5	(2,3)	(even, odd)	$1/4$	0.25
6	(2,3)	(odd, even)	$2/15$	0.13

Table 4.2: Theoretical and estimated probabilities of Merging in the next transition for all initial conditions

α as in (3.6.3) for a pair of solutions with all possible initial conditions. These values are the probabilities of merging in the next transition, thanks to Theorem 3.6.6 and Remark 3.7.7(2). These values are also contrasted with the estimated merging probabilities listed on the 5th column. This relative frequency based estimators' values have been rounded to two decimal places and computed using 9×10^5 number of samples generated by Monte Carlo simulation of pair of solutions corresponding to each initial condition(For algorithm see chapter 4 section 2.2).

4.2 Algorithms for the Simulation Studies

We have simulated the system of stochastic differential equations for semi-Markov processes.

4.2.1 Algorithm for dynamics of homogeneous SMP

```
from scipy.stats import poisson
import random as rd
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import scipy as sp
import math
from statistics import mean
simu=900000 ## Total number of Simulations
meet=[];merg=[]; merge_time=[];first_meet=[]; // Global Variables
for k in range(9000000):
  ## Generation of Poisson Points
    T=30;grid=1;A=[];N_part=0
    N=poisson.rvs(mu=2*T,size=grid)
  ## Total number of Poisson point Masses(PPM)
    N_total=sum(N); u=[];z=[];
    for j in range(grid):
        for i in range(N[j]):
        ## this is coordinate in time direction
          u.append(T*rd.random()+j*T);
        ## this is coordinate in space direction
          z.append(2*rd.random());
    v=np.array(u);   ## converted the list into array
    k=np.argsort(v);
    u=v[k];  ## sorted in ascending order
    B=np.column_stack((u,z)); ## Locus of PPM.
 ## Simulation of First SMP
    X_1=[[0,1,0]]; ## initial condition of first SMP.
 ## First transition of SMP due to first PPM
    X_1.append([u[0],X_1[0][1],u[0]]);
    ## First guess of (time, state, age) due to first PPM
    c=X_1[1][2]; ##storing pre transition age value in variable c.
    if (X_1[0][1] == 1 and (c/(1+c)) > B[0][1]):
    ## comparing PPM with first interval
        X_1[1][1]=2;X_1[1][2]=0; ## correcting the first guess depending
          on transition
    elif (X_1[0][1]==2 and 1<=B[0][1] and B[0][1] < 1+(c/(1+c))): ##
        comparing PPM with second interval
        X_1[1][1]=1;X_1[1][2]=0;  ## correcting the first guess depending
          on transition

    for i in range(1,N_total):
```

```
        X_1.append([u[i],X_1[i][1],X_1[i][2]+u[i]-u[i-1]]);## First guess
            of (time, state, age) due ith PPM
        c=X_1[i+1][2]; ##storing pre transition age value in variable c.
        if (X_1[i][1] == 1 and (c/(1+c)) > B[i][1]):## comparing PPM with
            first interval
               X_1[i+1][1]=2;X_1[i+1][2]=0;## correcting the ith guess
                   depending on transition
        elif (X_1[i][1]==2 and 1<=B[i][1] and B[i][1] < 1+(c/(1+c))): ##
            comparing PPM with second interval
               X_1[i+1][1]=1;X_1[i+1][2]=0; ## correcting the ith guess
                   depending on transition
    # print(len(X_1))
 ## Simulation of Second SMP
    X_2=[[0,2,0]]; ## initial condition of second SMP.
 ## First transition of SMP due to first PPM
    X_2.append([u[0],X_2[0][1],u[0]]); ## First guess of (time, state, age
        ) due to first PPM
    c=X_2[1][2]; ##storing pre-transition age value in variable c.
    if (X_2[0][1] == 1 and (c/(1+c)) > B[0][1]):
        X_2[1][1]=2;X_2[1][2]=0; ## correcting the first guess depending
            on transition
    elif (X_2[0][1]==2 and 1<=B[0][1] and B[0][1] < 1+(c/(1+c))):
        X_2[1][1]=1;X_2[1][2]=0; ## correcting the first guess depending
            on transition
    for i in range(1,N_total):
        X_2.append([u[i],X_2[i][1],X_2[i][2]+u[i]-u[i-1]]); c=X_2[i+1][2];
        if (X_2[i][1] == 1 and (c/(1+c)) > B[i][1]): ## comparing PPM with
             first interval
               X_2[i+1][1]=2;X_2[i+1][2]=0; ## correcting the ith guess
                   depending on transition
        elif (X_2[i][1]==2 and 1<=B[i][1] and B[i][1] < 1+(c/(1+c))): ##
            comparing PPM with second interval
               X_2[i+1][1]=1;X_2[i+1][2]=0; ## correcting the ith guess
                   depending on transition
    # print(len(X_2))
## Storing the time of meeting of two Simulations
    tmeet=[]; ## Reinitialise after each simulation
    for i in range(1,N_total):
        if (X_1[i-1][1]!=X_2[i-1][1] and X_1[i][1]==X_2[i][1]):
          tmeet.append([X_1[i],X_2[i]]);
          meet.append([X_1[i],X_2[i]]);
    first_meet.append(tmeet[0][0][0]);
## Storing the merging time of two Simulations
    for i in range(N_total):
          ## Checking merging condition
        if (X_1[i][2]== 0 and 0 == X_2[i][2] and X_1[i][1]==X_2[i][1]):
          merg.append([X_1[i],X_2[i]]); ## Storing whole merging event
          merge_time.append(tmeet[(len(tmeet)-1)][0][0]); ## Storing the
              merging time.
          break;  ##Since after merging meeting is prohibited, if merged
              the last meeting time is merging time.
```

```
print(len(merg)); ## Number of times merged out off 9 lakhs simulation.
print(max(merge_time))
print(max(first_meet))
## Graph plot
columns=200; ## Bins in histogram
x = np.arange(0, max(max(merge_time),max(first_meet)), 0.05)
y_1=x*np.exp(-x)
pf = np.array(first_meet)
y_2, binedges=np.histogram(pf,bins=columns)
bincenters = (0.5 * (binedges[1:] + binedges[:-1]))
bin_width=(binedges[1:]- binedges[:-1])
y_2_=y_2/(len(first_meet)*bin_width)
y_2=y_2_.astype(np.float)
pg = np.array(merge_time)
y_3, binedges_1=np.histogram(pg,bins=columns);
bincenters_1 = (0.5 * (binedges_1[1:] + binedges_1[:-1]))
bin_width_1=(binedges_1[1:]- binedges_1[:-1])
y_3_=y_3/(len(merge_time)*bin_width_1);
y_3=y_3_.astype(np.float)
plt.plot(x,y_1,c='green',label='holding time')
plt.plot(bincenters, y_2_, '-', c='blue',label='first meeting time')
plt.plot(bincenters_1, y_3_, '-', c='red',label='merge time')
plt.title('Probability density functions')
plt.xlabel('Time')
plt.ylabel('')
plt.legend()
plt.xlim([0, 10])
plt.show
plt.savefig('2S_HSMP.png',format="png", dpi = 600) ## Saving the image 2
    H_HSMP with 600 dpi in png format
print("Average first Meet time"mean(first_meet))
print("Average of Merge time"mean(merge_time))
```

4.2.2 Algorithm for Non-Homogeneous Markov Process

```
## program for simulating multiple merging of two non-homogeneous Markov
    Process
from scipy.stats import poisson
import random as rd
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import math
simu=900000 ## Total number of Simulations
meet=[];merg=[]; merge_time=[];first_meet=[];nco_meet=[];nco_time=[];
for k in range(simu):
  ## Generation of Poisson Points
    T=30;grid=1;A=[];
    N=poisson.rvs(mu=8*T,size=grid)
```

```
    N_total=sum(N); u=[];z=[]; ## Total number of Poisson point Masses(PPM
        )
    for j in range(grid):
        for i in range(N[j]):
          u.append(T*rd.random()+j*T);  ## this is coordinate in time
              direction
          z.append(8*rd.random());    ## this is coordinate in space
              direction
    v=np.array(u);   ## converted the list variable into array
    k=np.argsort(v);
    u=v[k];  ## sorted in ascending order
    B=np.column_stack((u,z)); ## Locus of PPM.
  ## Simulation of First NHMP
    X_1=[[0,1,0]]; ## initial condition of first NHMP.
    for i in range(N_total):
      X_1.append([u[i],X_1[i][1],X_1[i][2]]);## guess of (time, state, age
          ) due to ith PPM
      if ((X_1[i][2])%2==1): ## when pre-transition count is odd

        if (X_1[i][1] == 1 and (4/3 > B[i][1])):## comparing PPM with
            first interval
              X_1[i+1][1]=2; X_1[i+1][2]+=1;## correcting the first guess
                  depending on transition
        elif (X_1[i][1]==1 and 4/3<=B[i][1] and B[i][1] < 2): ## comparing
             PPM with second interval
              X_1[i+1][1]=3;X_1[i+1][2]+=1; ## correcting the first guess
                  depending on transition
        elif (X_1[i][1]==2 and 2<=B[i][1] and B[i][1] < 4): ## comparing
            PPM with third interval
              X_1[i+1][1]=1;X_1[i+1][2]+=1; ## correcting the first guess
                  depending on transition
        elif (X_1[i][1]==2 and 4<=B[i][1] and B[i][1] < 6): ## comparing
            PPM with fourth interval
              X_1[i+1][1]=3;X_1[i+1][2]+=1; ## correcting the first guess
                  depending on transition
        elif (X_1[i][1]==3 and 6<=B[i][1] and B[i][1] < 20/3): ##
            comparing PPM with fifth interval
              X_1[i+1][1]=1;X_1[i+1][2]+=1; ## correcting the first guess
                  depending on transition
        elif (X_1[i][1]==3 and 20/3<=B[i][1] and B[i][1] < 8): ##
            comparing PPM with sixth interval
              X_1[i+1][1]=2;X_1[i+1][2]+=1; ## correcting the first guess
                  depending on transition

      elif ((X_1[i][2])%2==0): ## when pre-transition count is even

        if (X_1[i][1] == 1 and (2/3 > B[i][1])):   ## comparing PPM with
            first interval
              X_1[i+1][1]=2;X_1[i+1][2]+=1;## correcting the first guess
                  depending on transition
        elif (X_1[i][1]==1 and 4/3<=B[i][1] and B[i][1] < 5/3): ##
```

```
            comparing PPM with second interval
               X_1[i+1][1]=3;X_1[i+1][2]+=1; ## correcting the first guess
                   depending on transition
        elif (X_1[i][1]==2 and 2<=B[i][1] and B[i][1] < 3): ## comparing
            PPM with third interval
               X_1[i+1][1]=1;X_1[i+1][2]+=1; ## correcting the first guess
                   depending on transition
        elif (X_1[i][1]==2 and 4<=B[i][1] and B[i][1] < 5): ## comparing
            PPM with fourth interval
               X_1[i+1][1]=3; X_1[i+1][2]+=1;## correcting the first guess
                   depending on transition
        elif (X_1[i][1]==3 and 6<=B[i][1] and B[i][1] < 19/3): ##
            comparing PPM with fifth interval
               X_1[i+1][1]=1;X_1[i+1][2]+=1; ## correcting the first guess
                   depending on transition
        elif (X_1[i][1]==3 and 20/3<=B[i][1] and B[i][1] < 22/3): ##
            comparing PPM with sixth interval
               X_1[i+1][1]=2; X_1[i+1][2]+=1;## correcting the first guess
                   depending on transition

## Simulation of Second NHMP
 X_2=[[0,2,0]]; ## initial condition of first NHMP.
 for i in range(N_total):
   X_2.append([u[i],X_2[i][1],X_2[i][2]]);## Guess of (time,state,age)
       due to ith PPM
   if ((X_2[i][2])%2==1): ## When pre-transition count is odd
      if (X_2[i][1] == 1 and (4/3 > B[i][1])):## comparing PPM with
            first interval
               X_2[i+1][1]=2; X_2[i+1][2]+=1;## correcting the ith guess
                   depending on transition
        elif (X_2[i][1]==1 and 4/3<=B[i][1] and B[i][1] < 2): ## comparing
             PPM with second interval
               X_2[i+1][1]=3;X_2[i+1][2]+=1; ## correcting the ith guess
                   depending on transition
        elif (X_2[i][1]==2 and 2<=B[i][1] and B[i][1] < 4): ## comparing
            PPM with third interval
               X_2[i+1][1]=1;X_2[i+1][2]+=1; ## correcting the ith guess
                   depending on transition
        elif (X_2[i][1]==2 and 4<=B[i][1] and B[i][1] < 6): ## comparing
            PPM with fourth interval
               X_2[i+1][1]=3;X_2[i+1][2]+=1; ## correcting the ith guess
                   depending on transition
        elif (X_2[i][1]==3 and 6<=B[i][1] and B[i][1] < 20/3): ##
            comparing PPM with fifth interval
               X_2[i+1][1]=1;X_2[i+1][2]+=1; ## correcting the ith guess
                   depending on transition
        elif (X_2[i][1]==3 and 20/3<=B[i][1] and B[i][1] < 8): ##
            comparing PPM with sixth interval
               X_2[i+1][1]=2;X_2[i+1][2]+=1; ## correcting the ith guess
                   depending on transition
```

```
      elif ((X_2[i][2])%2==0): ## when pre-transition count is even

        if (X_2[i][1] == 1 and (2/3 > B[i][1])):   ## comparing PPM with
            first interval
              X_2[i+1][1]=2;X_2[i+1][2]+=1;## correcting the ith guess
                  depending on transition
        elif (X_2[i][1]==1 and 4/3<=B[i][1] and B[i][1] < 5/3): ##
            comparing PPM with second interval
              X_2[i+1][1]=3;X_2[i+1][2]+=1; ## correcting the ith guess
                  depending on transition
        elif (X_2[i][1]==2 and 2<=B[i][1] and B[i][1] < 3): ## comparing
            PPM with third interval
              X_2[i+1][1]=1;X_2[i+1][2]+=1; ## correcting the ith guess
                  depending on transition
        elif (X_2[i][1]==2 and 4<=B[i][1] and B[i][1] < 5): ## comparing
            PPM with fourth interval
              X_2[i+1][1]=3; X_2[i+1][2]+=1;## correcting the ith guess
                  depending on transition
        elif (X_2[i][1]==3 and 6<=B[i][1] and B[i][1] < 19/3): ##
            comparing PPM with fifth interval
              X_2[i+1][1]=1;X_2[i+1][2]+=1; ## correcting the ith guess
                  depending on transition
        elif (X_2[i][1]==3 and 20/3<=B[i][1] and B[i][1] < 22/3): ##
            comparing PPM with sixth interval
              X_2[i+1][1]=2; X_2[i+1][2]+=1;## correcting the ith guess
                  depending on transition
    ## Storing the time of meeting of two Simulations
    tmeet=[]; ## Reinitialise after each simulation
    for i in range(1,N_total):
        if (X_1[i-1][1]!=X_2[i-1][1] and X_1[i][1]==X_2[i][1]): ## pre-
            transition states are different and meeting happens
           tmeet.append([X_1[i],X_2[i]]);
           meet.append([X_1[i],X_2[i]]);
    first_meet.append(tmeet[0][0][0]);
    ## Merging in the next transition
    if (tmeet!=[]):
      first_meet.append(tmeet[0][0][0]);
    if (X_1[0][2]==0 and X_2[0][2]==1):
      for i in range(1,N_total):
        if(X_1[i][1]==X_2[0][1] and X_1[i][2]==1 and X_2[i][2]==1):
         merge_in_next+=1;
         break;
    elif(X_1[0][2]==1 and X_2[0][2]==0):
      for i in range(1,N_total):
       if(X_1[0][1]==X_2[i][1] and X_1[i][2]==1 and X_2[i][2]==1):
         merge_in_next+=1;
         break;
print("Total␣Simulation",simu)
print("Number␣of␣Times␣Merge␣in␣Next␣Transition",merge_in_next)
print("Probability␣of␣Merging␣in␣next␣transition",merge_in_next/simu)
columns=100; ## Bins in the histogram
```

```
if (first_meet != []):
  pf = np.array(first_meet)
  y_2, binedges=np.histogram(pf,bins=2*columns)
  bincenters = (0.5 * (binedges[1:] + binedges[:-1]))
  bin_width=(binedges[1:]- binedges[:-1])
  y_2_=y_2/(len(first_meet)*bin_width)
  y_2=y_2_.astype(np.float)
  plt.plot(bincenters, y_2_, '-', c='blue',label='first meet time')
plt.title('probability density function')
plt.xlabel('Time')
plt.ylabel('')
plt.legend()
plt.xlim([-0.1,7])
plt.savefig('1_2_3S_NHMP.png',format="png", dpi=600)
plt.show()
print("Average of first Meet ",sum(first_meet)/len(first_meet))
```

Chapter 5

Component-wise Semi-Markov Process

5.1 General CSM

Definition 5.1.1. *A pure jump process X on a countable state space $\mathcal{X}$ is called a Component-wise Semi-Markov Process (CSM) if there is a bijection $\Gamma : \mathcal{X} \to \prod_{l=1}^{d} \mathcal{X}^l$, such that each component of $\Gamma(X)$ is a semi-Markov process, where d is a positive integer and for each $l \in \mathbb{D} := \{1, 2, \ldots, d\}$, $\mathcal{X}^l$ is an at-most countable non-empty set.*

Component-wise semi-Markov process (CSM) on finite state-space with independent components has first been introduced in [10]. The CSM is a generalization of semi-Markov processes into a broader class of pure jump processes. The CSM, studied in [10] possesses a well defined bounded transition rate function and hence that has been used to characterize the CSM. However, the definition of CSM does not imply existence of a rate function. In view of this it is important to find an alternative way of characterizing a general type of CSM. Needless to mention, the kernel characterization should be most suitable in this regard. We recall that the transition rate exists if and only if the kernel is almost everywhere differentiable. In that case, the rate can be expressed in terms of the kernel and vice versa. It is also easy to note that the knowledge of kernels of all individual components of CSM is sufficient to characterize the CSM, provided the components are independent to each other. However, in this chapter we consider an extension of CSM, appearing in [10], by dropping the independence condition. We further allow the state-space of each component be non-identical and at most countable. So, we propose characterization of CSM using a novel notion of kernel. The way we define the kernel, is broad enough to include both the dependent and independent component cases and both the homogeneous and non-homogeneous cases. As per our knowledge this is the first effort in the literature to characterize a general CSM using a kernel-based approach.

5.1.1 Kernel associated with general CSM

Without loss of generality, we assume that $\mathcal{X} = \prod_{l=1}^{d} \mathcal{X}^l$ and X^l is a semi-Markov process on $\mathcal{X}^l$ for each $l \in \mathbb{D}$. Here Γ is the identity map. The process $X = (X^1, \ldots, X^d)$ is the CSM. Let $x = (x^1, x^2, \ldots, x^d)$ and $y = (y^1, y^2, \ldots, y^d)$ be the state and age of CSM, where $x^l \in \mathcal{X}^l$ and $y^l \geq 0$. Assume that $\{T_n\}_{n\in\mathbb{N}_0}$ is an increasing sequence of positive numbers such that $T_0 := 0$ and $T_n := \inf\{t > T_{n-1} : t \in T^1 \cup \cdots \cup T^d\}$, where T^l denotes the collection of transition times of X^l for each $l \in \mathbb{D}$. We further define the component-wise age process $Y = (Y^1, \ldots, Y^d)$ of the CSM recursively. For each $l \in \mathbb{D}$, $Y^l_{T_0} = 0$ and

$$Y^l_{T_m} = \begin{cases} Y^l_{T_{m-1}} + (T_m - T_{m-1}), & \text{if } X^l_{T_m} = X^l_{T_{m-1}} \\ 0, & \text{otherwise} \end{cases}$$

for all $m \geq 1$. It is evident from above that for each $m \geq 1$, Y_{T_m} is measurable wrt the σ- algebra generated by $\{X_{T_m}, X_{T_{m-1}}, \cdots, X_{T_0}, T_m, \cdots, T_0\}$. Furthermore, for a CSM the following holds, for each $y_2 \in \mathbb{R}^d_+$, $x_2 \in \prod_{i=1}^{d} \mathcal{X}_i$

$$\begin{aligned} &P\left[\bigcap_{l\in\mathbb{D}}\left(\{X^l_{T^l_{n^l_m+1}} = x^l_2\} \cap \{T^l_{n^l_m+1} - T_m \leq y^l_2\}\right) \mid X_{T_m}, Y_{T_m}\right] \\ &= P\left[\bigcap_{l\in\mathbb{D}}\left(\{X^l_{T^l_{n^l_m+1}} = x^l_2\} \cap \{T^l_{n^l_m+1} - T_m \leq y^l_2\}\right) \mid X_{T_m}, X_{T_{m-1}}, \cdots, X_{T_0}, T_m, \cdots, T_0\right] \end{aligned}$$

where $n^l_m := N^l_{T_m}$, N^l_t is the number of transitions in l component till time t. For each $m \geq 1$, $y_1, y_2 \in \mathbb{R}^d_+$, $x_1, x_2 \in \prod_{i=1}^{d} \mathcal{X}_i$ the transition kernel of CSM is given by,

$$Q^m_{x_1,y_1}(x_2, y_2) := P\left[\bigcap_{l\in\mathbb{D}}\left(\{X^l_{T^l_{n^l_m+1}} = x^l_2\} \cap \{T^l_{n^l_m+1} - T_m \leq y^l_2\}\right) \mid X_{T_m} = x_1, Y_{T_m} = y_1\right]. \tag{5.1.1}$$

Definition 5.1.2. *If the transition kernel Q^m for a general CSM is constant in m then we call the CSM homogeneous, otherwise non-homogeneous.*

There are two natural questions. (i) Does $Q^m_{x_1,y_1}(x_2, y_2)$ coincide with the kernel function of semi-Markov process if $d = 1$? (ii) Does Q^m result in a unique kernel if $d > 1$? The answer to the first is affirmative and is evident. Indeed if $d = 1$, there is only one component. Thus $T^1 = T$ and $T^1_{N^1_{T_m}+1} = T_{m+1}$. Hence, for $x_1, x_2 \in \mathcal{X}$ and $y_1, y_2 \geq 0, m \in \mathbb{N}_0$ from (5.1.1) we get

$$Q^m_{x_1,y_1}(x_2, y_2) = P\left[\left(\{X_{T_{m+1}} = x_2\} \cap \{T_{m+1} - T_m \leq y_2\}\right) \mid X_{T_m} = x_1, Y_{T_m} = y_1\right].$$

The right side is indeed the well known kernel function for a semi-Markov process. The second question is also answered affirmatively in the following theorem.

Theorem 5.1.3. *Let $X = (X^1, X^2, \dots, X^d)$ be a CSM on $\mathcal{X}$, and for each $m \geq 1$, $x_1, x_2 \in \mathcal{X}$, $y_1, y_2 \in {\mathbb{R}_+}^d$, $Q^m_{x_1,y_1}(x_2, y_2)$ be defined by (5.1.1). Also denote $E := \mathcal{X} \times {\mathbb{R}_+}^d$, and $\mathcal{E} := \mathcal{B}(\mathcal{X} \times {\mathbb{R}_+}^d)$, the Borel σ algebra on E. Then the map $(x_1, y_1, x_2, y_2) \mapsto Q^m_{x_1,y_1}(x_2, y_2)$ determines a unique regular conditional distribution $\mathcal{Q}^m$ satisfying*

1. *for each (x_1, y_1), $B \mapsto \mathcal{Q}^m(x_1, y_1, B)$ is a probability measure on $(E, \mathcal{E})$ such that $\mathcal{Q}^m(x_1, y_1, \{x_2\} \times \prod_{l=1}^d (-\infty, y_2^l]) = Q^m_{x_1,y_1}(x_2, y_2)$;*
2. *for each $B \in \mathcal{E}$, $(x_1, y_1) \mapsto \mathcal{Q}^m(x_1, y_1, B)$ is measurable wrt $(E, \mathcal{E})$.*

For a fixed $m \in \mathbb{N}$, the measurability of Q^m w.r.t. y_1 is due to the definition 5.1.1. Therefore, the above theorem can be proved in the similar line of the proof of [7, Theorem 7.2.2, page 225]. We omit the details.

Several different transition related probabilities can be obtained from the kernel. For the purpose of illustration we present expressions of only couple of them here.

Definition 5.1.4 (Marginal kernel of l^{th} component)**.** *Given X is a CSM, then the Marginal kernel of l^{th} component is denoted by $\bar{Q}^{m,l}_{x_1,y_1}(x', r)$ and is defined as below*

$$\bar{Q}^{m,l}_{x_1,y_1}(x', r) := \sum_{l' \notin \{l\}} \lim_{y^{l'} \to \infty} \sum_{x_2^{l'} \in \mathcal{X}^{l'}} Q^m_{x_1,y_1}(x_2, y_2)\Big|_{y_2^l = r, x_2^l = x'}.$$

Remark 5.1.5. *Using the definition 5.1.4 of marginal kernel of l^{th} component, $\bar{Q}^{m,l}_{x_1,y_1}(x', r)$ is equal to $P(\{X^l_{T^l_{n^l_m+1}} = x'\} \cap \{T^l_{n^l_m+1} - T_m \leq r\} \mid X_{T_m} = x_1, Y_{T_m} = y_1)$, the conditional probability of transition of lth component to x' within next r unit of time given the CSM is at x_1 with the component-wise ages being y_1 at the mth transition time.*

Definition 5.1.6. *Let $\tau^l(t)$ denote the remaining duration after time t before the l^{th} component of X has the next transition. In fact $\tau^l(t) = T^l_{N^l_t+1} - t$.*

Let $F_{\tau^l(T_m)}(\cdot \mid x_1, y_1)$ denote the conditional c.d.f. of $\tau^l(T_m)$ given $X_{T_m} = x_1$ and $Y_{T_m} = y_1$. For $r \geq 0$,

$$\begin{aligned}
&F_{\tau^l(T_m)}(r \mid x_1, y_1) \\
&= P(\tau^l(T_m) \leq r \mid X_{T_m} = x_1, Y_{T_m} = y_1) \\
&= P(T^l_{n^l_{m+1}} - T_m \leq r \mid X_{T_m} = x_1, Y_{T_m} = y_1) \\
&= \sum_{x' \in \mathcal{X}^l} P(\{X^l_{T^l_{n^l_m+1}} = x'\} \cap \{T^l_{n^l_m+1} - T_m \leq r\} \mid X_{T_m} = x_1, Y_{T_m} = y_1) \\
&= \sum_{x' \in \mathcal{X}^l} \bar{Q}^{m,l}_{x_1,y_1}(x', r). \qquad (5.1.2)
\end{aligned}$$

Note that the way we define T_m, that gives the time of mth transition of the CSM process. More precisely, the transition instance of CSM where multiple components transit together is also recorded as a single transition of the CSM. Therefore, $m = \sum_{l\in\mathbb{D}} n_m^l$ holds only if no two states transit together at or before time T_m. Let $l(t)$ denote the set of components of X, where the subsequent transition of CSM happens after time t. In fact if all components of a CSM process are independent, $l(t)$ is singleton with probability one. However, the dependent components may transit together in finite time with positive probability. At those occasions $l(t)$ fail to remain almost sure singleton.

5.2 Semimartingale Representation of CSM Process

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be the underlying filtered probability space satisfying the usual hypothesis. Given a right stochastic matrix $W = (w_{ll'})_{1\leq l,l'\leq d}$ and d number of independent Poisson random measures (PRM) $\wp_1, \ldots, \wp_d$ on $\mathbb{R}^2 \times [0,1]$ with intensity m_3, define $\bar{\wp}_l$ as

$$\bar{\wp}_l(A) = \sum_{l'=1}^{d} \wp_{l'}(A \times [0, w_{ll'}]), \quad l \in \mathbb{D} \tag{5.2.1}$$

where A is any measurable subset of $\mathbb{R}^2$. It is easy to see that $\{\bar{\wp}_l\}_{l\in\mathbb{D}}$ are PRM on $\mathbb{R}^2$. Notice that $\bar{\wp}_1, \ldots, \bar{\wp}_d$ are not independent to each other. We also assume that $\wp_1, \ldots, \wp_d$ are such that $\{\bar{\wp}_l((0,t] \times dv)\}_{t\geq 0}$ is adapted to $\{\mathcal{F}_t\}_{t\geq 0}$. We further assume that and $\mathcal{X}^l = \{1, 2, \ldots\}$ the state space. We wish to construct a semi-Markov process on this state pace with a given transition rate function. The transition rate function, under consideration is allowed to yield a non-homogeneous embedded Markov chain. To this end we first embed this set in $\mathbb{N}$ and endow with its usual total order $\prec_1$, which in turn induces a total order $\prec_2$ on $\mathcal{X}_2^l := \{(i,j) \in \mathcal{X}^l \times \mathcal{X}^l \mid i \neq j\}$ by following lexicographic order. Let $\lambda^l(y,n) := (\lambda_{ij}^l(y,n))$ denote a matrix in which the i^{th} diagonal element is $\lambda_{ii}^l(y,n) := -\sum_{j\in\mathcal{X}^l\setminus\{i\}} \lambda_{ij}^l(y,n)$ and for each $(i,j) \in \mathcal{X}_2^l$, $\lambda_{ij}^l \colon \mathbb{R}_+ \times \mathbb{N}_0 \to \mathbb{R}_+$ is a bounded measurable function satisfying the following two assumptions.

D1. $c^l := \sup_i c_i^l$, where $c_i^l := \|\lambda_i^l(\cdot,\cdot)\|_{L^\infty_{(\mathbb{R}_+\times\mathbb{N}_0)}}$, and $\lambda_i^l(y,n) := |\lambda_{ii}^l(y,n)|$.

D2. For each $n \geq 0$ and $l \in \mathbb{D}$, $\lim_{y\to\infty} \gamma_i^l(y,n) = \infty$, where $\gamma_i^l(y,n) := \int_0^y \lambda_i^l(y',n)dy'$.

Let $C_i^l := c_i^l + \sum_{k\prec_1 i} c_k^l$. For each $(i,j) \in \mathcal{X}_2^l$ and $l \in \mathbb{D}$ we consider another measurable function $\tilde{\lambda}_{ij}^l \colon \mathbb{R}_+ \times \mathbb{N}_0 \to \mathbb{R}_+$ such that for each $y \geq 0$, $l \in \mathbb{D}$ and $n \in \mathbb{N}_0$

$$\lambda_{ij}^l(y,n) \leq \tilde{\lambda}_{ij}^l(y,n), \tag{5.2.2}$$

and also for almost every $y \geq 0$ and $n \in \mathbb{N}_0$

$$\tilde{\lambda}^l_{ij}(y,n) \leq \|\lambda^l_i(\cdot,\cdot)\|_{L^\infty_{(\mathbb{R}_+\times\mathbb{N}_0)}}. \tag{5.2.3}$$

Now for each $y \geq 0$, and $n \in \mathbb{N}_0$, with the help of $\lambda^l(y,n)$ and $\tilde{\lambda}^l(y,n) := (\tilde{\lambda}^l_{ij}(y,n))$, we introduce a collection of disjoint intervals $\Lambda^l := \{\Lambda^l_{ij}(y,n) \colon (i,j) \in \mathcal{X}^l_2\}$, by

$$\Lambda^l_{ij}(y,n) = \left(\sum_{(i',j')\prec_2 (i,j)} \tilde{\lambda}^l_{i'j'}(y,n) \right) + \Big[0, \lambda^l_{ij}(y,n)\Big) \tag{5.2.4}$$

where $a + B = \{a + b \mid b \in B\}$ for $a \in \mathbb{R}, B \subset \mathbb{R}$. Clearly, for each i,j,n and l, the interval $\Lambda^l_{ij}(y,n)$ is of length $\lambda^l_{ij}(y,n)$ for each $y \geq 0$, and for almost every $y \geq 0$ the union $\Lambda^l_i(y,n) := \cup_{j\in\mathcal{X}\setminus\{i\}}\Lambda^l_{ij}(y,n)$ is contained in a finite interval $\big[0, C^l_i\big]$, according to (**D1**). Using the above intervals we define $h^l_\Lambda \colon \mathcal{X} \times \mathbb{R}_+ \times \mathbb{N}_0 \times \mathbb{R} \to \mathbb{R}$ as

$$h^l_\Lambda(i,y,n,v) := \sum_{j\in\mathcal{X}^l\setminus\{i\}} (j-i)\mathbb{1}_{\Lambda^l_{ij}(y,n)}(v), \tag{5.2.5}$$

and $g^l_\Lambda \colon \mathcal{X} \times \mathbb{R}_+ \times \mathbb{N}_0 \times \mathbb{R} \to \mathbb{R}$ as

$$g^l_\Lambda(i,y,n,v) := \sum_{j\in\mathcal{X}^l\setminus\{i\}} \mathbb{1}_{\Lambda^l_{ij}(y,n)}(v). \tag{5.2.6}$$

These functions are piece-wise constant in v variable. Using these, we consider the following system of coupled stochastic integral equations in X^l, Y^l, and N^l for $t > 0$, where $\bar{\wp}_l$ is defined before.

$$X^l_t = X^l_0 + \int_{0+}^t \int_{\mathbb{R}} h^l_\Lambda(X^l_{u-}, Y^l_{u-}, N^l_{u-}, v)\bar{\wp}_l(du,dv) \tag{5.2.7}$$

$$Y^l_t = Y^l_0 + t - \int_{0+}^t (Y^l_{u-}) \int_{\mathbb{R}} g^l_\Lambda(X^l_{u-}, Y^l_{u-}, N^l_{u-}, v)\bar{\wp}_l(du,dv) \tag{5.2.8}$$

$$N^l_t = \int_{0+}^t \int_{\mathbb{R}} g^l_\Lambda(X^l_{u-}, Y^l_{u-}, N^l_{u-}, v)\bar{\wp}_l(du,dv). \tag{5.2.9}$$

We assume the vector notation for $t > 0$,

$$Z^l_t = Z^l_0 + \int_0^t A(Z^l_{u-})du + \int_{[0,t]\times\mathbb{R}^+} J^l(Z^l_{u-}, v)\bar{\wp}_l(du,dv), \quad l \in \mathbb{D} \tag{5.2.10}$$

where $Z^l_t = \big(X^l_t, Y^l_t, N^l_t\big)'$, $Z^l_0 = \big(X^l_0, Y^l_0, N^l_0\big)'$, $A = \big(0,1,0\big)'$, $J^l = \big(h^l_\Lambda, -yg^l_\Lambda, g^l_\Lambda\big)'$ and $\bar{\wp}_l$ are PRM as in (5.2.1).

Remark 5.2.1. *For each $l \in \mathbb{D}$, $Z^l = (X^l, Y^l, N^l)$ exists uniquely from Theorem 3.2.2 and Theorem 3.2.4.*

Remark 5.2.2. *If $W = I_{d\times d}$, then from* (5.2.1) $\bar{\wp}_l(A) = \wp_l(A \times [0,1])$. *As* $\{\wp_l \mid l \in \mathbb{D}\}$ *are independent PRMs on* $\mathbb{R}^2 \times [0,1]$, *the collection* $\{\bar{\wp}_l \mid l \in \mathbb{D}\}$ *becomes independent too. Thus* (5.2.10), *produces a CSM process with independent components. Such CSM have been used for modelling financial asset price dynamics in [9] and [10].*

Remark 5.2.3. *For a fixed j_1 we set $W = (w_{ij})$ where* $w_{ij} = \begin{cases} 1, & \text{if } j = j_1 \\ 0, & \text{otherwise} \end{cases}$.
Hence, from (5.2.1) $\bar{\wp}_l(A) = \wp_{j_1}(A \times [0,1])$ *for each* $l \in \mathbb{D}$. *Thus in* (5.2.10) *a single PRM drives each component of* $Z = (Z^1, \ldots, Z^d)$. *In addition to this, if* $J^1, J^2, \ldots, J^d$ *are all identical, Z can be viewed as the vector of solutions of the same equation with possibly different initial conditions. A CSM* $X = (X^1, \ldots, X^d)$ *obtained from such solutions have been closely studied in preceding chapters.*

5.3 Infinitesimal Generator of CSM

Notation 5.3.1. *Fix $i_l \in \mathcal{X}^l$ and $y_l \geq 0$, $l \in \mathbb{D}$. Let $Z^l = (X^l, Y^l, N^l), l \in \mathbb{D}$ be the strong solutions of* (5.2.7)-(5.2.9) *with initial conditions for fixed time* $s(> 0)$

$$i_l = X^l_s, y_l = Y^l_s, n_l = N^l_s, l \in \mathbb{D}. \tag{5.3.1}$$

We also denote $(Z^1, Z^2, \cdots, Z^d)$ *as* Z, $(X^1, \ldots, X^d)$ *as* X, $(Y^1, \ldots, Y^d)$ *as* Y *and* $(N^1, \ldots, N^d)$ *as* N, *where Z is the augmented process, X is CSM and Y, N are its age and transition counts.*

Next we restrict ourselves to the following choice of $\tilde{\lambda}$ for the ease of deriving the law of Z.

D3. For almost every $l \in \mathbb{D}$ and $y \geq 0$ and $(i', j') \in \mathcal{X}^l_2$, $\tilde{\lambda}_{i'j'}(y) = \|\lambda^l_{i'j'}(\cdot, \cdot)\|_{L^\infty_{\mathbb{R}_+ \times \mathbb{N}_0}}$.

Infinitesimal Generator of CSM with two component driven by one PRM is presented in Chapter 2 (2.5.1)(both components are homogeneous SMP's) and Chapter 3(3.5.3)(both components are non-homogeneous SMP's), in the next subsection we calculate the infinitesimal generator of CSM with three component where each component is non-homogeneous SMP driven by one PRM.

5.3.1 CSM with 3 components and single PRM

Let us consider the case for $d = 3$ and assume for each $l = 1, 2, 3$ PRMs $\bar{\wp}_1 = \bar{\wp}_2 = \bar{\wp}_3 = \wp$, $\mathcal{X}^1 = \mathcal{X}^2 = \mathcal{X}^3$ and $J^1 = J^2 = J^3 = J$. Also Since, Z^1, Z^2 and Z^3 as in Notation

5.3.1 are Markov, $Z = (Z^1, Z^2, Z^3)$ is also Markov. It has state, age components and transition counts $X = (X^1, X^2, X^3)$, $Y = (Y^1, Y^2, Y^3)$ and $N = (N^1, N^2, N^3)$ respectively. While each of X^1,X^2 and X^3 is semi-Markov, the pure jump process X is not. Rather, X is a component-wise semi-Markov process (CSM) and the Markov process Z is called the augmented process of CSM X. A CSM with independent components has been introduced for modelling financial assets in [10]. However, a CSM with dependent components has not been studied in the literature yet. Since, for our case, the components of the CSM X are driven by a single Poisson random measure, they are not independent. In view of this, it is interesting to derive the law of X by finding the generator of Z. To this end, we recall Itô's lemma for r.c.l.l. semi-martingales. Let $\varphi\colon (\mathcal{X} \times \mathbb{R}_+ \times \mathbb{N}_0)^3 \to \mathbb{R}$ be bounded and continuously differentiable in its continuous variables, then

$$
\begin{aligned}
&d\varphi(Z_t^1, Z_t^2, Z^3) - \left(\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2} + \frac{\partial}{\partial y_3}\right)\varphi(Z_t^1, Z_t^2, Z^3)dt \\
&= \varphi(Z_t^1, Z_t^2, Z_t^3) - \varphi(Z_{t-}^1, Z_{t-}^2, Z_{t-}^3) \\
&= \varphi\left(Z_{t-}^1 + \int_{\mathbb{R}_+} J(Z_{t-}^1, v)\wp(dt, dv), Z_{t-}^2 + \int_{\mathbb{R}_+} J(Z_{t-}^2, v)\wp(dt, dv), Z_{t-}^3 + \int_{\mathbb{R}_+} J(Z_{t-}^3, v)\wp(dt, dv)\right) \\
&- \varphi(Z_{t-}^1, Z_{t-}^2, Z_{t-}^3) \\
&= \int_{\mathbb{R}_+} \left[\varphi(Z_{t-}^1 + J(Z_{t-}^1, v), Z_{t-}^2 + J(Z_{t-}^2, v), Z_{t-}^3 + J(Z_{t-}^3, v)) - \varphi(Z_{t-}^1, Z_{t-}^2, Z_{t-}^3)\right] \wp(dt, dv) \\
&= \left(\int_{\mathbb{R}_+} \left[\varphi(Z_{t-}^1 + J(Z_{t-}^1, v), Z_{t-}^2 + J(Z_{t-}^2, v), Z_{t-}^3 + J(Z_{t-}^3, v))) - \varphi(Z_{t-}^1, Z_{t-}^2, Z_{t-}^3)\right] dv\right) dt + dM_t
\end{aligned}
$$

where M is the martingale obtained by integration w.r.t the compensated Poisson random measure $\wp(dt, dv) - dtdv$. And the third equality holds true due to Theorem 0.2.20. To simplify the above integral term, we impose (D3) and divide the derivation into five disjoint and exhaustive cases. Those are based on equality of the states of the components before transition. To be more precise, the cases where all three components are different, any two are identical and different from the only remaining one, and where all three are at identical state are considered.

Case 1: Assume X_{t-}^1, X_{t-}^2 and X_{t-}^3 are all distinct. Then under (D3), $\Lambda_{X_{t-}^1 j_1}(Y_{t-}^1, N_{t-}^1)$, $\Lambda_{X_{t-}^2 j_2}(Y_{t-}^2, N_{t-}^2)$ and $\Lambda_{X_{t-}^3 j_2}(Y_{t-}^3, N_{t-}^3)$ are disjoint for every $j_1 \neq X_{t-}^1$, $j_2 \neq X_{t-}^2$ and $j_3 \neq X_{t-}^3$.

Thus, by considering these intervals where the integrand is non-zero constants, we get

$$\begin{aligned}
&\int_{\mathbb{R}_+}[\varphi(Z^1_{t-}+J(Z^1_{t-},v),Z^2_{t-}+J(Z^2_{t-},v),Z^3_{t-}+J(Z^3_{t-},v))-\varphi(Z^1_{t-},Z^2_{t-},Z^3_{t-})]dv\\
&=\int_{\bigcup_{k=1}^{3}\left(\underset{j\in\mathcal{X}^l\setminus\{X^k_{t-}\}}{\cup}\Lambda_{X^k_{t-}j}(Y^k_{t-},N^k_{t-})\right)}[\varphi(Z^1_{t-}+J(Z^1_{t-},v),Z^2_{t-}+J(Z^2_{t-},v),Z^3_{t-}+J(Z^3_{t-},v))\\
&\qquad-\varphi(Z^1_{t-},Z^2_{t-},Z^3_{t-})]dv\\
&=\sum_{j\neq X^1_{t-}}[\varphi((j,0,N^1_{t-}+1),Z^2_{t-},Z^3_{t-})-\varphi(Z^1_{t-},Z^2_{t-},Z^3_{t-})]|\Lambda_{X^1_{t-}j}(Y^1_{t-},N^1_{t-})|\\
&\qquad+\sum_{j\neq X^2_{t-}}[\varphi(Z^1_{t-},(j,0,N^2_{t-}+1),Z^3_{t-})-\varphi(Z^1_{t-},Z^2_{t-},Z^3_{t-})]|\Lambda_{X^2_{t-}j}(Y^2_{t-},N^2_{t-})|\\
&\qquad+\sum_{j\neq X^3_{t-}}[\varphi(Z^1_{t-},Z^2_{t-},(j,0,N^3_{t-}+1))-\varphi(Z^1_{t-},Z^2_{t-},Z^3_{t-})]|\Lambda_{X^3_{t-}j}(Y^3_{t-},N^3_{t-})|.
\end{aligned}$$

Case 2: Assume that $X^1_{t-}=X^2_{t-}=i\neq X^3_{t-}$. Now under (D3), the intervals $\Lambda_{X^1_{t-}j_1}(Y^1_{t-},N^1_{t-})$, $\Lambda_{X^2_{t-}j_2}(Y^2_{t-},N^2_{t-})$ have identical left end points for every $j_1,j_2\neq i$ and $\Lambda_{X^3_{t-}j_3}(Y^3_{t-},N^3_{t-})$ is disjoint with $\Lambda_{X^1_{t-}j_1}(Y^1_{t-},N^1_{t-})$ and $\Lambda_{X^2_{t-}j_2}(Y^2_{t-},N^2_{t-})$ for every $j_1,j_2\neq i$. Thus, by considering the intervals where the integrand is non-zero constants, we get

$$\begin{aligned}
&\int_{\mathbb{R}_+}[\varphi(Z^1_{t-}+J(Z^1_{t-},v),Z^2_{t-}+J(Z^2_{t-},v),Z^3_{t-}+J(Z^3_{t-},v))-\varphi(Z^1_{t-},Z^2_{t-},Z^3_{t-})]dv\\
&=\int_{\bigcup_{k=1}^{3}\left(\underset{j\in\mathcal{X}^l\setminus\{X^k_{t-}\}}{\cup}\Lambda_{X^k_{t-}j}(Y^k_{t-},N^k_{t-})\right)}[\varphi(Z^1_{t-}+J(Z^1_{t-},v),Z^2_{t-}+J(Z^2_{t-},v),Z^3_{t-}+J(Z^3_{t-},v))\\
&\quad-\varphi(Z^1_{t-},Z^2_{t-},Z^3_{t-})]dv\\
&=\sum_{j\neq i}[\varphi((j,0,N^1_{t-}+1),Z^2_{t-},Z^3_{t-})-\varphi(Z^1_{t-},Z^2_{t-},Z^3_{t-})]|\Lambda_{ij}(Y^1_{t-},N^1_{t-})\setminus\Lambda_{ij}(Y^2_{t-},N^2_{t-})|\\
&+\sum_{j\neq i}[\varphi(Z^1_{t-},(j,0,N^2_{t-}+1),Z^3_{t-})-\varphi(Z^1_{t-},Z^2_{t-},Z^3_{t-})]|\Lambda_{ij}(Y^2_{t-},N^2_{t-})\setminus\Lambda_{ij}(Y^1_{t-},N^1_{t-})|\\
&+\sum_{j\neq i}[\varphi((j,0,N^1_{t-}+1),(j,0,N^2_{t-}+1),Z^3_{t-})-\varphi(Z^1_{t-},Z^2_{t-},Z^3_{t-})]|\Lambda_{ij}(Y^1_{t-},N^1_{t-})\cap\Lambda_{ij}(Y^2_{t-},N^2_{t-})|\\
&+\sum_{j\neq X^3_{t-}}[\varphi(Z^1_{t-},Z^2_{t-},(j,0,N^3_{t-}+1))-\varphi(Z^1_{t-},Z^2_{t-},Z^3_{t-})]|\Lambda_{X^3_{t-}j}(Y^3_{t-},N^3_{t-})|.
\end{aligned}$$

Case 3: Assume that $X^1_{t-}=X^3_{t-}=i\neq X^2_{t-}$. Again under (D3), the intervals $\Lambda_{X^1_{t-}j_1}(Y^1_{t-},N^1_{t-})$, $\Lambda_{X^3_{t-}j_3}(Y^3_{t-},N^3_{t-})$ have identical left end points for every $j_1,j_3\neq i$ and $\Lambda_{X^2_{t-}j_2}(Y^2_{t-},N^2_{t-})$ is disjoint with intervals $\Lambda_{X^1_{t-}j_1}(Y^1_{t-},N^1_{t-})$ and $\Lambda_{X^3_{t-}j_3}(Y^3_{t-},N^3_{t-})$ for every $j_1,j_3\neq i$, and $j_2\neq X^2_{t-}$.

Thus, by considering these intervals where the integrand is non-zero constants, we get

$$\int_{\mathbb{R}_+} \left[\varphi(Z^1_{t-} + J(Z^1_{t-},v), Z^2_{t-} + J(Z^2_{t-},v), Z^3_{t-} + J(Z^3_{t-},v)) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})\right] dv$$
$$= \int_{\bigcup\limits_{k=1}^{3}\left(\underset{j\in\mathcal{X}^l\setminus\{X^k_{t-}\}}{\cup}\Lambda_{X^k_{t-}j}(Y^k_{t-},N^k_{t-})\right)} [\varphi(Z^1_{t-} + J(Z^1_{t-},v), Z^2_{t-} + J(Z^2_{t-},v), Z^3_{t-} + J(Z^3_{t-},v))$$
$$- \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})]dv$$
$$= \sum_{j\neq i}[\varphi((j,0,N^1_{t-}+1), Z^2_{t-}, Z^3_{t-}) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})]|\Lambda_{ij}(Y^1_{t-},N^1_{t-}) \setminus \Lambda_{ij}(Y^3_{t-},N^3_{t-})|$$
$$+ \sum_{j\neq i}[\varphi(Z^1_{t-}, Z^2_{t-}, (j,0,N^3_{t-}+1)) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})]|\Lambda_{ij}(Y^3_{t-},N^3_{t-}) \setminus \Lambda_{ij}(Y^1_{t-},N^1_{t-})|$$
$$+ \sum_{j\neq i}[\varphi((j,0,N^1_{t-}+1), Z^2_{t-}, (j,0,N^3_{t-}+1)) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})]|\Lambda_{ij}(Y^1_{t-},N^1_{t-}) \cap \Lambda_{ij}(Y^3_{t-},N^3_{t-})|$$
$$+ \sum_{j\neq X^2_{t-}} \left[\varphi(Z^1_{t-}, (j,0,N^2_{t-}+1), Z^3_{t-}) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})\right] |\Lambda_{X^2_{t-}j}(Y^2_{t-},N^2_{t-})|.$$

Case 4: Assume that $X^1_{t-} \neq X^2_{t-} = X^3_{t-} = i$. Now under (D3), the intervals $\Lambda_{X^2_{t-}j_2}(Y^2_{t-},N^2_{t-})$, $\Lambda_{X^3_{t-}j_3}(Y^3_{t-},N^3_{t-})$ have identical left end points for every $j_2, j_3 \neq i$ and $\Lambda_{X^1_{t-}j_1}(Y^1_{t-},N^1_{t-})$ is disjoint with the intervals $\Lambda_{X^2_{t-}j_2}(Y^2_{t-},N^2_{t-})$ and $\Lambda_{X^3_{t-}j_3}(Y^3_{t-},N^3_{t-})$ for every $j_1 \neq X^1_{t-}$, $j_2, j_3 \neq i$. Thus, by considering these intervals where the integrand is non-zero constants, we get

$$\int_{\mathbb{R}_+} \left[\varphi(Z^1_{t-} + J(Z^1_{t-},v), Z^2_{t-} + J(Z^2_{t-},v), Z^3_{t-} + J(Z^3_{t-},v)) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})\right] dv$$
$$= \int_{\bigcup\limits_{k=1}^{3}\left(\underset{j\in\mathcal{X}^l\setminus\{X^k_{t-}\}}{\cup}\Lambda_{X^k_{t-}j}(Y^k_{t-},N^k_{t-})\right)} [\varphi(Z^1_{t-} + J(Z^1_{t-},v), Z^2_{t-} + J(Z^2_{t-},v), Z^3_{t-} + J(Z^3_{t-},v))$$
$$- \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})]dv$$
$$= \sum_{j\neq i}[\varphi((Z^1_{t-}, (j,0,N^2_{t-}+1), Z^3_{t-}) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})]|\Lambda_{ij}(Y^2_{t-},N^2_{t-}) \setminus \Lambda_{ij}(Y^3_{t-},N^3_{t-})|$$
$$+ \sum_{j\neq i}[\varphi(Z^1_{t-}, Z^3_{t-}, (j,0,N^3_{t-}+1)) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})]|\Lambda_{ij}(Y^3_{t-},N^3_{t-}) \setminus \Lambda_{ij}(Y^2_{t-},N^2_{t-})|$$
$$+ \sum_{j\neq i}[\varphi(Z^1_{t-}, (j,0,N^2_{t-}), (j,0,N^3_{t-}+1)) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})]|\Lambda_{ij}(Y^2_{t-},N^2_{t-}) \cap \Lambda_{ij}(Y^3_{t-},N^3_{t-})|$$
$$+ \sum_{j\neq X^1_{t-}} \left[\varphi((j,0,N^1_{t-}+1), Z^2_{t-}, Z^3_{t-}) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})\right] |\Lambda_{X^1_{t-}j}(Y^1_{t-},N^1_{t-})|.$$

Case 5: Assume that $X^1_{t-} = X^2_{t-} = X^3_{t-} = i$. Then under (D3), the intervals $\Lambda_{ij}(Y^1_{t-},N^1_{t-})$, $\Lambda_{ij}(Y^2_{t-},N^2_{t-})$ and $\Lambda_{ij}(Y^3_{t-},N^3_{t-})$ are having identical left end points. So, $\Lambda_{X^1_{t-}j_1}(Y^1_{t-},N^1_{t-})$,

$\Lambda_{X^2_{t-}j_2}(Y^2_{t-}, N^2_{t-})$ and $\Lambda_{X^3_{t-}j}(Y^3_{t-}, N^3_{t-})$ are not disjoint when $j_1 = j_2 = j_3$. Thus

$$\begin{aligned}
&\int_{\mathbb{R}_+} [\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v), Z^3_{t-} + J(Z^3_{t-}, v)) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})]dv \\
&= \int_{(\bigcup_{k=1}^{3}(\bigcup_{j\neq i}\Lambda_{ij}(Y^k_{t-}, N^k_{t-})))} [\varphi(Z^1_{t-} + J(Z^1_{t-}, v), Z^2_{t-} + J(Z^2_{t-}, v), Z^3_{t-} + J(Z^3_{t-}, v)) \\
&\qquad\qquad - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})]dv
\end{aligned}$$

$$\begin{aligned}
&= \sum_{j\neq i} [\varphi((j, 0, N^1_{t-} + 1), Z^2_{t-}, Z^3_{t-}) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})] \times \\
&\qquad |\Lambda_{ij}(Y^1_{t-}, N^1_{t-}) \setminus (\Lambda_{ij}(Y^2_{t-}, N^2_{t-}) \cup \Lambda_{ij}(Y^3_{t-}, N^3_{t-}))| \\
&+ \sum_{j\neq i} [\varphi(Z^1_{t-}, (j, 0, N^2_{t-} + 1), Z^3_{t-}) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})] \times \\
&\qquad |\Lambda_{ij}(Y^2_{t-}, N^2_{t-}) \setminus (\Lambda_{ij}(Y^1_{t-}, N^1_{t-}) \cup \Lambda_{ij}(Y^3_{t-}, N^3_{t-}))| \\
&+ \sum_{j\neq i} [\varphi(Z^1_{t-}, Z^2_{t-}, (j, 0, N^3_{t-} + 1)) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})] \times \\
&\qquad |\Lambda_{ij}(Y^3_{t-}, N^3_{t-}) \setminus (\Lambda_{ij}(Y^1_{t-}, N^1_{t-}) \cup \Lambda_{ij}(Y^2_{t-}, N^2_{t-}))| \\
&+ \sum_{j\neq i} [\varphi((j, 0, N^1_{t-} + 1), (j, 0, N^2_{t-} + 1), Z^3_{t-}) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})] \times \\
&\qquad |\Lambda_{ij}(Y^1_{t-}, N^1_{t-}) \cap \Lambda_{ij}(Y^2_{t-}, N^2_{t-}) \setminus \Lambda_{ij}(Y^3_{t-}, N^3_{t-})| \\
&+ \sum_{j\neq i} [\varphi((j, 0, N^1_{t-} + 1), Z^2_{t-}, (j, 0, N^3_{t-} + 1)) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})] \times \\
&\qquad |\Lambda_{ij}(Y^1_{t-}, N^1_{t-}) \cap \Lambda_{ij}(Y^3_{t-}, N^3_{t-}) \setminus \Lambda_{ij}(Y^2_{t-}, N^2_{t-})| \\
&+ \sum_{j\neq i} [\varphi(Z^1_{t-}, (j, 0, N^2_{t-} + 1), (j, 0, N^3_{t-} + 1)) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})] \times \\
&\qquad |\Lambda_{ij}(Y^2_{t-}, N^2_{t-}) \cap \Lambda_{ij}(Y^3_{t-}, N^3_{t-}) \setminus \Lambda_{ij}(Y^1_{t-}, N^1_{t-})| \\
&+ \sum_{j\neq i} [\varphi((j, 0, N^1_{t-} + 1), (j, 0, N^2_{t-} + 1), (j, 0, N^3_{t-} + 1)) - \varphi(Z^1_{t-}, Z^2_{t-}, Z^3_{t-})] \times \\
&\qquad |\Lambda_{ij}(Y^1_{t-}, N^1_{t-}) \cap \Lambda_{ij}(Y^2_{t-}, N^2_{t-}) \cap \Lambda_{ij}(Y^3_{t-}, N^3_{t-})|.
\end{aligned}$$

Thus, by combining the expressions under all the cases, the generator $\mathcal{A}$ of (Z^1, Z^2, Z^3) is given by

$$
\begin{aligned}
&\mathcal{A}\varphi(z_1, z_2, z_3) \\
&= \left(\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2} + \frac{\partial}{\partial y_3}\right)\varphi(z_1, z_2, z_3) \\
&\quad + \sum_{j\notin\{i_1\}} [\varphi((j, 0, n_1 + 1), z_2, z_3) - \varphi(z_1, z_2, z_3)]\,[\lambda_{i_1 j}(y_1, n_1) - \delta_{i_1,i_2}(1 - \delta_{i_1,i_3})\lambda_{i_2 j}(y_2, n_2) \\
&\qquad\qquad - \delta_{i_1,i_3}(1 - \delta_{i_1,i_2})\lambda_{i_3 j}(y_3, n_3) - \delta_{i_1,i_2}\delta_{i_1,i_3}(\lambda_{i_2 j}(y_2, n_2) \vee \lambda_{i_3 j}(y_3, n_3))]^+ \\
&\quad + \sum_{j\notin\{i_2\}} [\varphi(z_1, (j, 0, n_2 + 1), z_3) - \varphi(z_1, z_2, z_3)]\,[\lambda_{i_2 j}(y_2, n_2) - \delta_{i_1,i_2}(1 - \delta_{i_2,i_3})\lambda_{i_1 j}(y_1, n_1) \\
&\qquad\qquad - \delta_{i_2,i_3}(1 - \delta_{i_1,i_2})\lambda_{i_3 j}(y_3, n_3) - \delta_{i_1,i_2}\delta_{i_2,i_3}(\lambda_{i_1 j}(y_1, n_1) \vee \lambda_{i_3 j}(y_3, n_3))]^+ \\
&\quad + \sum_{j\notin\{i_3\}} [\varphi(z_1, z_2, (j, 0, n_3 + 1)) - \varphi(z_1, z_2, z_3)]\,[\lambda_{i_3 j}(y_3, n_3) - \delta_{i_1,i_3}(1 - \delta_{i_2,i_3})\lambda_{i_1 j}(y_1, n_1) \\
&\qquad\qquad - \delta_{i_2,i_3}(1 - \delta_{i_1,i_3})\lambda_{i_2 j}(y_2, n_2) - \delta_{i_1,i_3}\delta_{i_2,i_3}(\lambda_{i_1 j}(y_1, n_1) \vee \lambda_{i_2 j}(y_2, n_2))]^+ \\
&\quad + \delta_{i_1,i_2} \sum_{j\notin\{i_1,i_2\}} [\varphi((j, 0, n_1 + 1), (j, 0, n_2 + 1), z_3) - \varphi(z_1, z_2, z_3)] \times \\
&\qquad\qquad [(\lambda_{i_1 j}(y_1, n_1) \wedge \lambda_{i_2 j}(y_2, n_2)) - \delta_{i_1,i_3}\lambda_{i_3 j}(y_3, n_3)]^+ \\
&\quad + \delta_{i_1,i_3} \sum_{j\notin\{i_1,i_3\}} [\varphi((j, 0, n_1 + 1), z_2, (j, 0, n_3 + 1)) - \varphi(z_1, z_2, z_3)] \times \\
&\qquad\qquad [(\lambda_{i_1 j}(y_1, n_1) \wedge \lambda_{i_3 j}(y_3, n_3)) - \delta_{i_1,i_2}\lambda_{i_2 j}(y_2, n_2)]^+ \\
&\quad + \delta_{i_2,i_3} \sum_{j\notin\{i_2,i_3\}} [\varphi(z_1, (j, 0, n_2 + 1), (j, 0, n_3 + 1)) - \varphi(z_1, z_2, z_3)] \times \\
&\qquad\qquad [(\lambda_{i_2 j}(y_2, n_2) \wedge \lambda_{i_3 j}(y_3, n_3)) - \delta_{i_1,i_3}\lambda_{i_1 j}(y_1, n_1)]^+ \\
&\quad + \delta_{i_1,i_2}\delta_{i_1,i_3} \sum_{j\notin\{i_1,i_2,i_3\}} [\varphi((j, 0, n_1 + 1), (j, 0, n_2 + 1), (j, 0, n_3 + 1)) - \varphi(z_1, z_2, z_3)] \times \\
&\qquad\qquad [\lambda_{i_1 j}(y_1, n_1) \wedge \lambda_{i_2 j}(y_2, n_2) \wedge \lambda_{i_3 j}(y_3, n_3)] \qquad (5.3.2)
\end{aligned}
$$

where $z_1 = (i_1, y_1, n_1)$, $z_2 = (i_2, y_2, n_2)$, $z_3 = (i_3, y_3, n_3)$, and $\delta_{i,j} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise.} \end{cases}$

5.3.2 CSM with arbitrary components and single PRM

Let $z = (z_1, z_2, \dots, z_d)$ and $z' = (z'_1, z'_2, \dots, z'_d)$, where $z_l = (x_l, y_l, n_l), z'_l = (x'_l, y'_l, n'_l)$ are in $\mathcal{X}^l \times \mathbb{R}_+ \times \mathbb{N}_0$. For any given z and z' we define

$$\begin{aligned} i(z,z') :=& \{l \in \mathbb{D} \mid x'_l \neq x_l, y'_l = 0, n'_l = n_l + 1\},\\ x(z,z') :=& \{x_l \mid l \in i(z,z')\},\\ \text{and } x'(z,z') :=& \{x'_l \mid l \in i(z,z')\}. \end{aligned}$$

Evidently, $x(z,z') = x'(z,z') = \emptyset$, the empty set if $i(z,z')$ is empty. Set

$$\delta_A := \begin{cases} 1 & \text{if } A \text{ is singleton} \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore assume that $\max_\emptyset v = 0$ for any $v : \mathbb{D} \to \mathbb{R}_+$. If $\mathcal{X}^1 = \cdots = \mathcal{X}^d = \mathcal{X}$ and $\lambda^1 = \lambda^2 = \cdots = \lambda^d = \lambda$, using the above notations we define for any $\phi : \prod_{l=1}^d (\mathcal{X} \times \mathbb{R}_+ \times \mathbb{N}_0) \to \mathbb{R}$

$$\begin{aligned} \mathcal{A}\phi(z) = \sum_{l=1}^d \frac{\partial}{\partial y_l}\phi(z) + \sum_{z'} \delta_{x(z,z')}\delta_{x'(z,z')} \left(\prod_{l \in \mathbb{D} \setminus i(z,z')} \delta_{\{z_l, z'_l\}} \right) \times \\ \left(\min_{l \in i(z,z')} \lambda_{x_l, x'_l}(y_l, n_l) - \max_{\substack{\{l \in \mathbb{D} \setminus i(z,z') \\ \mid x_l \in x(z,z')\}}} \lambda_{x_l, x'_l}(y_l, n_l) \right)^+ (\phi(z') - \phi(z)) \end{aligned} \tag{5.3.3}$$

provided ϕ has first order partial derivatives w.r.t y_l for each $l \in \mathbb{D}$.

Theorem 5.3.2. *Assume (D1)-(D3). Further assume that $\bar{\wp}_1 = \cdots = \bar{\wp}_d = \wp$, $\mathcal{X}^1 = \cdots = \mathcal{X}^d$ and $\lambda^1 = \cdots = \lambda^d = \lambda$. Then the infinitesimal generator $\mathcal{A}$ of the solution to* (5.2.7)-(5.2.9) *is given by* (5.3.3).

Proof. The partial differential operators appearing in the first additive term is due to the time like continuous growth of age variable during non-occurrence of transitions. We would justify the remaining terms below.

It is evident from the derivation of (5.3.2) that in the expression of the generator the increment factor $\phi(z') - \phi(z)$ should be multiplied by the corresponding rate. The rate of the increment of course depends on z and z', the states before and after the transition respectively. As, due to a transition the age and the transition count becomes zero and increases by one respectively, the set $i(z,z')$ denotes the set of components which observe transitions. In addition to that, the components of z and z' other than those in $i(z,z')$ remain identical. So the rate possesses a factor $\prod_{l \in \mathbb{D} \setminus i(z,z')} \delta_{\{z_l, z'_l\}}$, which is one if and only if $z_l = z'_l$ for all $l \in \mathbb{D} \setminus i(z,z')$ and zero otherwise. We also observe that due to (D3) for almost every y and every n, the collection $\{\Lambda_{i,j}(y,n) \mid i \neq j\}$ contains disjoint intervals. So, for every transition, being caused due to a single point mass, the prior states of the transiting components are all identical. Also for the same reason the future states of all transiting components are identical. Therefore, the factor $\delta_{x(z,z')}\delta_{x'(z,z')}$ appears in the rate.

It is important to note that under the condition of $\delta_{x(z,z')} = 1$, the members of $i(z,z')$ are not the only components for which the x components of z coincide to those for $i(z,z')$. That is $i(z,z')$ could be a proper subset of $\{l \in \mathbb{D} \mid x_l \in x(z,z')\}$. In other words, components in $i(z,z')$ undergo transition and those in $\{l \in \mathbb{D} \mid x_l \in x(z,z')\} \setminus i(z,z')$ do not undergo transition. Given $\delta_{x(z,z')} = 1$, this can happen only if the Possion point mass appears in $\cap_{l\in i(z,z')}\Lambda_{x_l,x'_l}(y_l,n_l)$ and do not appear in $\cup(\Lambda_{x_l,x'_l}(y_l,n_l) \mid x_l \in x(z,z'), l \notin i(z,z'))$. The second set is empty if the union is over an empty family. As the left end points of all the intervals in the expression of the intersection and union are coincident (thanks to (D3)), the Lebesgue measure of these two sets are $\min\limits_{l\in i(z,z')} \lambda_{x_l,x'_l}(y_l,n_l)$ and $\max_{\{l\in\mathbb{D}\setminus i(z,z') \mid x_l\in x(z,z')\}} \lambda_{x_l,x'_l}(y_l,n_l)$ respectively. Note that this is valid even if the second set is empty due to our convention. Therefore, as the left end points are identical, the length of the set after subtracting the latter from the former is $\left(\min\limits_{l\in i(z,z')} \lambda_{x_l,x'_l}(y_l,n_l) - \max_{\substack{\{l\in\mathbb{D}\setminus i(z,z')\\ \mid x_l\in x(z,z')\}}} \lambda_{x_l,x'_l}(y_l,n_l)\right)^+$. This completes the proof.

Remark 5.3.3. *Note that the expression of the infinitesimal generator involves a summation over $z' \in (\mathcal{X} \times \mathbb{R}_+ \times \mathbb{N}_0)^d$. Although this appears as an uncountable sum, only at most countable number of terms may survive. Indeed, due to the definition of δ_A and $i(z,z')$, the factor $\prod_{l\in\mathbb{D}\setminus i(z,z')} \delta_{\{z_l,z'_l\}}$ is nonzero only if y'_l is either zero or the same as y_l for each $l \in \mathbb{D}$. Now, $\mathbb{D}$ being finite and $\mathcal{X}$ being at most countable, the summation has only a countable many nonzero terms. In fact this has at most $(2^d - 1) \times card(\mathcal{X})$ many nonzero terms.*

Chapter 6

Conclusions

In this thesis, we have considered semimartingale representations of certain classes of homogeneous and non-homogeneous semi-Markov processes. These representations are via stochastic differential equations involving Poisson random measures. The combination of state processes of two solutions (having two different initial conditions) of this equation forms a semi-Markov system(SMS) or a component-wise semi-Markov (CSM) process having dependent components. The SMSs [39, 41, 40], or CSMs [9, 10] with independent components have been introduced by several authors for modelling some random dynamics. However, a CSM with dependent components has not been studied in the literature yet. The law of the CSM has been calculated in terms of the infinitesimal generator of the augmented process. As per our knowledge, the present thesis is the first work on the correlated semi-Markov system. The immense applicability of semi-Markov processes and semi-Markov systems is well known. So, the formulation and the results related to the correlated semi-Markov system, presented in this thesis have significantly expanded the scope of further theoretical and applied studies. In particular this has opened up the possibility of studying a system generated by a semi-Markov flow. The questions related to the meeting and merging events of multiple particles of a semi-Markov flow have far reaching implications. In view of these, the present thesis appears as a stepping stone for several future studies.

There are many promising future research directions. We list some of those below.

- We have shown that the probability of eventual merging is one under some conditions on the model parameters. We have also produced an example of parameter values which does not fulfil the condition but a pair of solutions are shown to merge with probability one. In view of this, a further study of eventual merging under more relaxed condition for homogeneous or non-homogeneous semi-Markov flow appears interesting. For the non-homogeneous case transition rate matrix $\lambda(y, n)$ may be considered periodic in n.
- We recall that for studying stability of Markov Modulated diffusion, in [3, Lemma 3.3], the meeting time of a pair of Markov chains has been crucially utilized for constructing a merged pair. For the homogeneous Markov case, the meeting of a pair of solution

of the same SDE is identical to merging, which is not true for the semi-Markov case. Therefore, the argument presented in [3] is not applicable for studying the stability of semi-Markov Modulated diffusion. On the other hand the study of merging of semi-Markov solutions, presented in this thesis has enhanced our understanding of the related dynamics. We believe, the results and the line of arguments appearing in this thesis will be useful in the study of stability of semi-Markov Modulated diffusion.

- In this thesis we have seen that a pair of solutions of the SDE for semi-Markov process constitute a CSM. In principle, any number of such solutions together form a CSM process with dependent components. Nevertheless, these processes have differentiable kernels, which is a restriction on the general class of CSM processes. In this thesis we have proposed definition of a much wider class of homogeneous and non-homogeneous CSM. We have described the law using a kernel function. However, the expressions of various important transitions parameters in terms of the kernel is still in the dark. For example, further investigation is needed to express the conditional probability of one step transition in terms of the kernel function. An expression of conditional distribution of the holding time in terms of the kernel function is also important for many relevant studies.

- The law of a pair of solutions, that constitutes a CSM, has been presented in this thesis in terms of the infinitesimal generator. Moreover, a compact expression of the same when $d(\gg 2)$ number of solutions are considered has also been obtained in the fifth Chapter. However, the questions related to the meeting and merging of $d(> 2)$ number of solutions is unaddressed in this thesis. Further research is needed for answering these questions.

- Instead of considering the flow of a semi-Markov dynamics, one may consider the flow of a CSM with correlated components (correlated PRMs). Asymptotic analysis of such dynamics could be of great interest among researchers. This thesis just introduces such dynamics and shows the existence but does not study the asymptotic properties.

Other than the above mentioned theoretical studies, many applied branches may also be benefited from the investigations presented in this thesis. However, this thesis contains no study of applications. The potential application could be in studying and controlling queueing network or modelling investment strategy in quantitative finance, to name a few. Briefly we describe a queueing design that can be modelled and studied using the semi-Markov flow, introduced in this thesis.

Imagine there are finite or countably many service stations in a system and the total number of servers in the system is finite and fixed. Customers arrive at each station with a fixed arrival rate. A customer when arrives at a station is either served immediately by all, some or none of the servers in the station and then the customer leaves. The servers, which have served on an arrival of a customer, are transferred to another randomly chosen station together. In a station which of the servers serve, are decided based on a predetermined preference order depending on the duration (or experience) of the servers in that station.

Such queueing and service design can be modeled by the stochastic flow of semi-Markov process. In particular, each station is a state in the statespace, each server is a solution of the SDE, each arrival of customer corresponds to the appearance of Poisson point mass and each service is a transition. In view of this, the merging event is same as assemblage of all servers together in a single station and become indistinguishable.

Bibliography

[1] R. A. Arratia. "Coalescing Brownian motions on the line". PhD thesis. Univ. Wisconsin, Madison, 1979.

[2] K. B. Athreya, D. McDonald, and P. Ney. "Limit Theorems for Semi-Markov Processes and Renewal Theory for Markov Chains". In: *The Annals of Probability* 6.5 (1978), pp. 788–797. DOI: 10.1214/aop/1176995429. URL: https://doi.org/10.1214/aop/1176995429.

[3] Gopal K. Basak, Arnab Bisi, and Mrinal K. Ghosh. "Stability of a Random Diffusion with Linear Drift". In: *Journal of Mathematical Analysis and Applications* 202 (1996), pp. 604–622.

[4] Gopal K. Basak, Arnab Bisi, and Mrinal K. Ghosh. "Stability and Functional Limit Theorems for Random Degenerate Diffusions". In: *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)* 61.1 (1999), pp. 12–35. ISSN: 0581572X. URL: http://www.jstor.org/stable/25051226.

[5] Bouguet, Florian. "Quantitative speeds of convergence for exposure to food contaminants". In: *ESAIM: PS* 19 (2015), pp. 482–501. DOI: 10.1051/ps/2015002. URL: https://doi.org/10.1051/ps/2015002.

[6] Mark Brown. "Bounds, Inequalities, and Monotonicity Properties for Some Specialized Renewal Processes". In: *The Annals of Probability* 8.2 (1980), pp. 227–240. DOI: 10.1214/aop/1176994773. URL: https://doi.org/10.1214/aop/1176994773.

[7] Y.S. Chow et al. *Probability Theory: Independence, Interchangeability, Martingales*. Springer Texts in Statistics. Springer New York, 1997. ISBN: 9780387982281. URL: https://books.google.co.in/books?id=MwDvAAAAMAAJ.

[8] E. Çınlar. *Probability and Stochastics*. Graduate Texts in Mathematics. Springer New York, 2011. ISBN: 9780387878591. URL: https://link.springer.com/book/10.1007/978-0-387-87859-1.

[9] Milan Kumar Das, Anindya Goswami, and Tanmay S. Patankar. "Pricing derivatives in a regime switching market with time inhomogenous volatility". In: *Stochastic Analysis and Applications* 36.4 (2018), pp. 700–725. DOI: 10.1080/07362994.2018.1448996. eprint: https://doi.org/10.1080/07362994.2018.1448996. URL: https://doi.org/10.1080/07362994.2018.1448996.

[10] Milan Kumar Das, Anindya Goswami, and Nimit Rana. "Risk sensitive portfolio optimization in a jump diffusion model with regimes". In: *SIAM Journal on Control and Optimization* 56.2 (2018), pp. 1550–1576.

[11] Robert J Elliott. "The semi-martingale dynamics and generator of a continuous time semi-Markov chain". In: *Journal of Stochastic Analysis* 1.1 (2020), p. 1.

[12] Bennett L. Fox. *Semi-Markov Processes: A Primer*. Santa Monica, CA: RAND Corporation, 1968.

[13] Mrinal K. Ghosh and Anindya Goswami. "Risk minimizing option pricing in a semi-Markov modulated market". In: *SIAM J. Control. Optim.* 48 (2009), pp. 1519–1541.

[14] Mrinal K. Ghosh and Subhamay Saha. "Stochastic Processes with Age-Dependent Transition Rates". In: *Stochastic Analysis and Applications* 29.3 (2011), pp. 511–522.

[15] Anindya Goswami, Subhamay Saha, and Ravishankar Kapildev Yadav. "Semimartingle Representation of a class of Semi-Markov Dynamics". In: *to appear in Journal of Theoretical Probability* (2023). DOI: 10.48550/ARXIV.2207.06132. URL: https://arxiv.org/abs/2207.06132.

[16] Boris Harlmov. *Continuous semi-Markov processes*. doi: https://doi.org/10.1002/9780470610923. John Wiley & Sons, Ltd, 2008. ISBN: 9780470610923. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470610923.

[17] Theodore E. Harris. "Coalescing and noncoalescing stochastic flows in R_1". In: *Stochastic Processes and their Applications* 17.2 (1984), pp. 187–210. ISSN: 0304-4149. DOI: https://doi.org/10.1016/0304-4149(84)90001-2. URL: https://www.sciencedirect.com/science/article/pii/0304414984900012.

[18] Jan M. Hoem. "Inhomogeneous Semi-Markov Processes, Select Actuarial Tables, and Duration-Dependence in Demography". In: 1972.

[19] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*. North-Holland, Amsterdam, 1981, p. 419. ISBN: 978-0444861726.

[20] J. Janssen and R. de Dominicis. "Finite non-homogeneous semi-Markov processes: Theoretical and computational aspects". In: *Insurance: Mathematics and Economics* 3.3 (1984), pp. 157–165. ISSN: 0167-6687. DOI: https://doi.org/10.1016/0167-6687(84)90057-X. URL: https://www.sciencedirect.com/science/article/pii/016766878490057X.

[21] Jacques Janssen and Raimondo Manca. *Applied semi-Markov processes*. Springer Science & Business Media, 2006.

[22] Jacques Janssen and Raimondo Manca. "Numerical Solution of non-Homogeneous Semi-Markov Processes in Transient Case*". In: *Methodology And Computing In Applied Probability* 3.3 (Sept. 1, 2001), pp. 271–293. DOI: 10.1023/A:1013719007075. URL: https://doi.org/10.1023/A:1013719007075.

[23] Jacques Janssen and Raimondo Manca. *Semi-Markov risk models for finance, insurance and reliability*. Springer Science & Business Media, 2007.

[24] S. Kotz, I.I. Gikhman, and A.V. Skorokhod. *The Theory of Stochastic Processes II*. Classics in Mathematics. Springer Berlin Heidelberg, 1977. ISBN: 9783540202851. URL: https://link.springer.com/book/9783540202851.

[25] Hiroshi Kunita. *Stochastic flows and jump-diffusions*. Springer, 2019.

[26] Paul Levy. "Systemes Semi-Markoviens 'aau plus une infinite denombrable d'etats possibles". In: *Proc. Int. Congr.Math., Amsterdam* 2.1188 (1954), p. 294.

[27] Torgny Lindvall. "On coupling of renewal processes with use of failure rates". In: *Stochastic Processes and their Applications* 22.1 (1986), pp. 1–15. ISSN: 0304-4149. DOI: https://doi.org/10.1016/0304-4149(86)90109-2. URL: https://www.sciencedirect.com/science/article/pii/0304414986901092.

[28] A. Iosifescu Manu. "Non homogeneous semi-Markov processes". In: *Stud. Lere. Mat* (1972), pp. 529–533.

[29] Ian Melbourne and Dalia Terhesiu. "Renewal theorems and mixing for non Markov flows with infinite measure". In: *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* 56.1 (2020), pp. 449–476. DOI: 10.1214/19-AIHP968. URL: https://doi.org/10.1214/19-AIHP968.

[30] Esa Nummelin. "Limit theorems for $\alpha-$recurrent semi-Markov processes". In: *Advances in Applied Probability* 8.3 (1976), pp. 531–547. DOI: 10.2307/1426143.

[31] Steven Orey. "Change of Time Scale For Markov Processes". In: *Transactions of the American Mathematical Society* 99.3 (1961), pp. 384–397. ISSN: 00029947. URL: http://www.jstor.org/stable/1993552.

[32] Tanmay Patankar. "Asset Pricing in a Semi-Markov Modulated Market with Time-dependent Volatility". MA thesis. IISER Pune, 2016.

[33] PE Plotter. *Stochastic integration and differential equation*. Vol. 21. 2005.

[34] Ronald Pyke. "Markov renewal processes: definitions and preliminary properties". In: *The Annals of Mathematical Statistics* (1961), pp. 1231–1242.

[35] Ronald Pyke. "Markov renewal processes with finitely many states". In: *The Annals of Mathematical Statistics* (1961), pp. 1243–1259.

[36] H.L. Royden and P. Fitzpatrick. *Real Analysis*. Prentice Hall, 2010. ISBN: 9780131437470.

[37] Walter L Smith. "Regenerative stochastic processes". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 232.1188 (1955), pp. 6–31.

[38] Lajos Takács. "Some investigations concerning recurrent stochastic processes of a certain type". In: *Maygyar Tud. Akad. Mad. Kutató int. Kozl* 3 (1954), pp. 115–128.

[39] Aglaia Vasileiou and P.-C. G. Vassiliou. "An inhomogeneous semi-Markov model for the term structure of credit risk spreads". In: *Advances in Applied Probability* 38.1 (2006), pp. 171–198. DOI: 10.1239/aap/1143936146.

[40] P.-C. G. Vassiliou and A. A. Papadopoulou. "Non-homogeneous semi-Markov systems and maintainability of the state sizes". In: *Journal of Applied Probability* 29.3 (1992), pp. 519–534. DOI: 10.2307/3214890.

[41] P.-C.G. Vassiliou. "Exotic Properties of Non Homogeneous Markov and Semi-Markov Systems". In: *Communications in Statistics - Theory and Methods* 42.16 (2013), pp. 2971–2990. DOI: 10.1080/03610926.2012.698782. eprint: https://doi.org/10.1080/03610926.2012.698782. URL: https://doi.org/10.1080/03610926.2012.698782.