A Dynamical Recurrent Neural Network Model for Visual Perception of Numerosity

A Thesis

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Bhavesh K Verma

IISER PUNE

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

December, 2022

Supervisor: Dr Rakesh Sengupta

Certificate

This is to certify that this dissertation entitled A Dynamical Recurrent Neural Network Model for Visual Perception of Numerositytowards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Bhavesh K Verma at SRU Warangal, under the supervision of Dr Rakesh Sengupta -Director, Center for Creative Cognition SRU, Warangal, during the academic year 2021-2022.

......

Dr Rakesh Sengupta

Bhavesh K Verma

Declaration

I hereby declare that the matter embodied in the report entitled A Dynamical Recurrent Neural Network Model for Visual Perception of Numerosity are the results of the work carried out by me at the Center for Creative Cognition, SRU Warangal under the supervision of Dr Rakesh Sengupta and the same has not been submitted elsewhere for any other degree.

......

Dr Rakesh Sengupta

Bhavesh K Verma

Abstract

Humans have an inherent ability to visually perceive small numerosities (the cardinality of a set). According to several studies, the LIP and other IPS regions of the brain include the neurological substrates for the processes related to numbers. Computational models play a crucial role in investigating the computations and dynamics behind the perception of numbers. Existing models of number perception simulate a few of the fundamental aspects of number perception, such as size and distance effects. However, most models usually work for a limited range of numbers and lack explanations for important behavioral characteristics like adaptation effects. In this study, we use a network of neurons with self-excitatory and mutual inhibitory properties to build a computational model of number perception. We assume that the network's mean activation at steady-state can encode numerosity when it increases monotonically with Setsize (the input to the network). By optimizing the total number of inhibition strengths required so that the combined monotonic regions cover the full stretch of numbers, we get three ranges of numbers (1:4, 5:17, and 21:50). This division of numbers into three parts closely matches the elbows in numerosity perception discovered in behavioral studies. Later in the study, we devised a method to decode the mean activation into a continuous scale of numbers ranging from 1 to 50. Furthermore, we suggest a mechanism for selecting inhibition strength based on current inputs, allowing the network to work for the entire range of numerosities. Our model provides novel perspectives on how our brain can generate various behavioral phenomena, such as the influences of continuous visual attributes and adaptation effects on perceived numerosity.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

Over the course of civilization, humans have developed many systems for symbolic representation of numbers, enabling us to easily communicate and store numerical information as well as perform sophisticated algebraic operations on them. Our aptness to work with numbers gives us an unparalleled ability to understand the world through mathematics and science. But how much do we know about the cognition of numbers? Number cognition research spans a wide range of disciplines, including neuroscience, cognitive linguistics, development psychology, cognitive psychology, and others. Every one of these disciplines has its own research interests regarding number cognition, focusing on different directions to approach the subject. One of the most fundamental ways to approach the subject is through the study of perception of numbers, which involves empirical methods to learn about how our brain perceives numbers and their neural underpinnings. Number perception is one of the most important cognitive tools available to humans. We perceive numbers through many of our senses, and from various kinds of inputs. Through our vision, we can perceive the numerosity of a set of objects presented in the visual field at once, or in a sequential manner, or by looking at man-made symbols of numbers like '5' and 'two'. Through our auditory sense, we can perceive the numerosity of a sequence of sounds or a sound assigned to a number. We have the ability to perceive numbers even through haptics. It's fascinating to see how often we compare and translate numbers from various modes of perception without paying much attention to them. We see a set of objects and can easily assign a word or symbol to their numerosity. Numbers in their symbolic form are such an integral part of our lives that it is hard to think about numbers independently from them. However, we have a biologically ancient ability known as ”number sense” that does not require formal training or an understanding of symbols (Burr and Ross, 2008). Humans have an innate ability to perceive the cardinality of sets; even non-verbal animals and children can perceive non-symbolic numbers(referred to as “numerosity”) when presented in the form of a small set of objects in the field of vision (Gelman and Gallistel, 1986; Xu and Spelke, 2000; Hauser et al., 2003).

Vision is considered the most dominant sense in humans, and it also appears to be the primary mode for perception of numerosity. Performance related to visual perception of numerosity has significant implications for mathematics-related performance (Piazza et al., 2010; Anobile et al., 2013) . Understanding about visual perception of numerosity can help us address many interesting problems from cognitive science. In the following section, we will explore various methods to investigate visual perception of numerosity and discuss key results from experimental studies, providing a context for our computational model.

1.1 Behavioral Studies

One of the oldest methods of studying visual number perception is through behavioral studies, specifically psychophysics, which is defined as ”the scientific study of the relationship between stimulus and sensation.” (Gescheider, 1997). Psychophysics studies involve experiment-based studies of animals and humans by presenting the subjects with some stimuli and recording their response. For example, a simple psychophysical experiment on humans involving enumeration of a random number of dots on a computer screen can help us understand how our accuracy and reaction time (the time duration between stimuli and response) vary with numerosity.

Many studies categorize numbers into different regimes based on the differences in behavioral observations such as enumeration accuracy, confidence, and response time. The range of numbers from 1 to 4 (sometimes 3 or 5) is called the “subitizing range”(Kaufman et al., 1949). Humans can rapidly enumerate a set of objects falling into this range with almost no error. Larger numbers (usually five onwards)are referred to as the estimation range. Large numbers can either be counted, which is a slow process with high accuracy, or estimated, which is a relatively quick process but comes with error (Whalen et al., 1999; Kaufman et al., 1949). The error increases proportionately to the number of items to be estimated, consistent with Weber's law. The Weber fraction, defined as the just noticeable difference or precision threshold divided by the mean, is almost constant(0.15) over an extensive range of base numerosities.

It is a matter of debate whether enumerating numbers in the subitizing range and estimation range invokes different processes. in the The tendency of subitizing to withstand disruption has led to the belief that subitizing is pre-attentive, or at the very least uses pre-attentive information. (Trick and Pylyshyn, 1994). However, several recent studies indicate that subitizing is indeed susceptible to attentional load.(Egeth et al., 2008; Burr et al., 2010; Xu and Liu, 2008). In support of distinct mechanisms for large and small numbers, some studies have suggested a link between subitizing and object individuation; a visuospatial mechanism that allows us to locate and track a limited set of objects (Piazza, 2011; Piazza et al., 2011).

Several behavioural studies have also called into question the existence of two distinct processes. accorind to Balakrishnan and Ashby (1991) there is no sharp discontinuity in reaction times between the estimation and subitizing ranges. Weber fractions for adult participants were reported by citeross2003visual to be approximately 0.25 over an extended range of values (8–60). The Weber Fraction of 0.25 means that the just noticeable difference for the numbers in the subitizing range is less than 1, which is the least count while talking about whole numbers. Hence, the errorless nature of subitizing may be merely a result of the combined effect of the resolution of estimation mechanisms and the descrete separation at low numbers. Dehaene and Changeux (1993) and Gallistel and Gelman (1992) have advanced similar ideas. On the other hand, a recent study Portley and Durgin (2019) presents the presence of a new elbow in the estimation range, further dividing large numbers into two ranges: below and above 20. In this thesis, we show that these behaviourally distinct ranges of numbers do not necessarily have distinct neural mechanisms, as the same neural network can work for both small and large numbers by adapting its internal parameters (here, inhibition strength).

1.2 Neuroscience and Brain Imaging Studies

As the field of neuroscience has advanced, it has found a crucial place in number perception research. Neuroscience studies focus on the neural basis of number perception. Experiments involving visual stimuli associated with numbers and measurement of neuronal response through electrodes are helpful in understanding the number-neurons(neurons tuned to particular numbers). Studying the correlation between damage or lesion of a particular brain region with behavioral observations helps us to understand the role of different brain areas in the context of number perception.

Nieder et al. (2002) have shown the existence of "number-neurons" in the lateral prefrontal cortex (IPFC) and intraparietal sulcus (IPS) by training monkeys to discriminate the numerosity of a sample stimulus using the Delayed Match-to-Sample(DMS) paradigm Nieder et al. (2006, 2002). The number-neurons produced a clear tuning function for numerosity by generating the highest firing rate for their preferred numerosity, which falls off with the numerical distance.

Roitman et al. (2007) examined the activity of individual LIP neurons in monkeys as they viewed arrays of dots on a computer screen and discovered that neurons responds proportionately to the number of elements in the display for an extended range of numerosities. The monotonic increase in neuronal activity with numerosity supports the presence of the integration stage, as suggested in some computational models. This result is consistent with our computational model, as instead of simulating the number-neurons tuned to individual numbers, we show that numerosity is encoded as the monotonic increase of the network activation.

Due to recent advancements in brain imaging techniques, it has also found its place in the study of number perception. Functional MRI (fMRI) measures changes in blood flow in the brain, allowing researchers to observe specific regions and structures of the brain that are active during a task. A simple brain imaging experiment can involve an MRI scan of subjects performing tasks related to numbers. By designing the experiment in such a way that interactions with other mental processes can be accounted for, these studies try to find the regions of the brain where neural activation correlates with number cognition.

Harvey et al. (2013) used high-field 7T fMRI to study the human parietal cortex neural populations tuned to numbers.The selectivity was obviously correlated with the number of items, not any of the other variables, according to this study, which systematically alter all other potential confounding factors, such as size, contrast, contour, and length. Similar to how V1 has a columnar organisation for tuning of attributes like orientation, the human intraparietal sulcus has a columnar organisation for number. They demonstrate that whereas the highest numerosity elicits the strongest responses from V1, number-selective neurons only fire when the preferred stimulus provided instead of following a proportional responce to the numerosity. This research demonstrates that the number selective response present in the parietal cortex is not driven by the early visual cortex (V1), proving that number is not derived from low-level visual features.

The studies involving Electroencephalography(EEG) are also very useful because of the temporal resolution they offer. EEG provides information about neural activities in the brain by measuring the potential difference across the electrodes attached to the scalp. Though EEG allows real-time recording of high-speed brain activities, it lacks precision for pinpointing which brain structures are active.

Fornaciai and Park (2017) captured EEG data as human volunteers looked at arrays with systematically varying non-numerical stimuli with either very few (1-4) or extremely many (100-400) dots. Regardless of the numerical range provided, a linear model that examined the effects of numerical and non-numerical cues on the visual-evoked potentials (VEPs) demonstrated considerable neuronal sensitivity to numerosity approximately 160–180 ms over right occipito-parietal locations. On the other hand, earlier neural responses (100 ms) displayed distinctly diverse patterns throughout the several tested number ranges. These findings suggest that variations in the early stages of visual analysis may be the cause of behavioural response patterns that differ in numerosity estimation across different numerical ranges. Similar to this, Fornaciai and Park (2018) showed that numerosity perception goes through at least two stages: the first stage involves the extraction of raw sensory information from a dot-array stimulus early (100 msec) in the visual stream (V2/V3), which is insufficient to serve as a basis for numerosity perception; and the second involves the segmentation of the stimulus into perceptual units, later (150 msec) in the visual stream(V3).These findings emphasize the active and constructive nature of perception, which involves the early visual cortex's multiple stages of processing numerical data before converting it into a format that is appropriate for numerosity perception. These findings strengthen the case for a brain system dedicated to numerosity processing, however they also suggest that Multiple early visual cortical mechanisms converge later in the visual stream to that numerosity processing stage.

1.3 Theories of Number Perception

The Theory of Number Sense states that humans and many other animals have a biologically inherent ability to perceive and manipulate non-symbolic numbers. It has been argued that the sense of numbers in humans consists of two distinct systems, namely the Approximate Number System(ANS) and the Parallel Individuation System. ANS refers to the cognitive system behind approximate estimation of large numerosities, without relying on symbols and language. **??**. The rapid cognition of the numerosity of objects and comparison between two sets of objects are facilitated by ANS. The ANS is usually represented as a number line where each number is placed as a gaussian distribution. The width of the distribution is seen as an indicator of ANS acuity- the degree of precision of the internal representation of numerosity, and the overlap between distributions signifies the confusion in the differentiation of numerosities. Piazza et al. (2004); Pica et al. (2004)

While ANS is associated with a noisy internal representation of numerosities greater than 5, the accurate and quick enumeration of smaller numbers is believed to be facilitated by the Parallel Individuation System, also called the Object Tracking System - a visuo-spatial mechanism that allows us to track and locate a limited set of objects Piazza et al. (2011); Piazza (2011); Cantlon et al. (2009); Feigenson et al. (2004). It has been suggested that the ANS and the parallel individuation system set the basis for a symbolic representation of numbers in the human brain. Experiments in children have shown that ANS acuity has an influence on mathematical abilities and can be a building block for numbers in general .Halberda et al. (2008); Butterworth (2005); Dehaene (2009) ,

While ANS is seen as working independently of sensory cues like size, density, and convex hull of the set of objects in the visual field, other opposing theories call for a process which extracts information of numerosity from these continuous attributes. The theory of the Sensory Integration System suggests that different sensory cues are used to generate an estimation of numerosity. Gebuis et al. (2016); Dakin et al. (2011); Morgan et al. (2014).

1.3.1 Evolutionary Perspective

The development of the ability to perceive numerosity could result from the adaptive advantages it provides to animals. The ability to perceive the number of predators or prey in a visual field can be vital for pre-estimating chances of success or failure, which can hugely impact survivability. Did our brain evolve to completely filter out the information of numerosity from the visual field so that we could have an accurate representation of the number of objects in a scene, or did it evolve to optimize for some other representations which were not just numerical but also involved an active contribution of other continous cues such as size and density. Consider a hypothetical situation where an animal has to choose between two sets of apples, the first with a smaller number of big apples and the second with a larger number of small apples. We can see that in such situations, both size and number can be essential variables for a good decision. Even if the animal has an inherent ability to detect numerosity independently of other factors, it can't make a good decision based on the information of numerosity alone. In this case, the brain would need to combine the knowledge of numerosity with size, which will require further processing before the final decision. What if the animal has a cognitive process which does not completely isolate the numerosity from the other useful variable and reaches the decision based on a quantity which is not an accurate representation of numerosity? Neuroimaging studies reveal that some brain activities reflect the association between numbers and space. For instance, regions of the parietal cortex show shared activation for both spatial and numerical processing.Dehaene (1992) We can ask which of the two mechanisms is more adaptive: the presence of an independent sense of number, or a mental representation of quantity which already takes account of other useful variables?

1.4 Computational Models

The neural mechanisms of visual perception are not yet well understood. The complexity of a single neuron is still unfolding, and we are far from a comprehensive understanding of how billions of neurons interact to create the complexities observed through experiments and subjective experiences. To better understand any complex system, it can be helpful to start with simpler models. Using computational models to understand visual perception allows us to work with a few variables and use mathematical tools to analyze the results. In this section, we'll talk about some of the pertinent network models for number perception, including those that use Hebbian learning, backpropagation-based supervised learning Verguts and Fias (2004), and unsupervised learning Stoianov and Zorzi (2012). Some network models do not involve learning as they generate interesting dynamics based on predefined network parameters.Sengupta et al. (2014) Inputs for these neural networks usually represent information from various stages of visual pathways. During network simulation, the activation of the neurons changes based on equations with biologically relevant parameters . Depending on the model type, the parameters mimic plasticity, learning rate, noise, decay constant, inhibition, or excitation strength between neurons. Activation patterns of some neurons are taken as the output. It is a common practice to compare the outputs from these networks to the patterns observed in Psychophysics and neuroscience experiments.

Not all computational models use network-based simulations to understand number perception; some use purely mathematical models. For example, the model of the mental number line by Dehaene et al. (1998) where numbers are represented as overlapping Gaussian distributions. In this model, a larger overlapping between the gaussian distributions signifies more confusion and less accuracy in number comparison tasks.

Dehaene and Changeux (1993) presented one of the earliest models of numerosity that makes use of a reinforcement-based supervised learning methodology. There are two main computational building blocks in this four-stage connectionist model: The normalization stage, which is the first block, ignores the items' irrelevant spatial characteristics; the classification stage, which is the second block, assigns a numerical label to the normalized activity. The normalisation stage is carried out by first buffering the visual input, then using a series of object detectors that are tuned for the size and position of the object. The second stage consists of nodes with reciprocal inhibition, which generates "winner take all" dynamics; when an input activates two detectors, they compete by inhibiting each other, and eventually only the closest match wins. Dehaene and Changeaux propose a possible solution to the problem of normalization across stimulus features. The model accounts for the behavioral findings from human experimental data, namely, size and distance effects, but has a limited range of enumeration of up to five items.

Verguts and Fias (2004) developed a numerosity detector model based on error backpropagation based learning. The three-layer neural network used in this model learns how to map input patterns to numerical categories. During the training phase, the excitation and inhibition strengths are adjusted, which makes the neurons in the hidden layer mimic 'integrating units'. The neurons at the output layer behave as number-selective neurons, which is compared to the activation patterns observed in biological number-neurons. Through this model, the authors hypothesize that the same network modules that support the nonsymbolic numerosities also aid in the construction of symbolic representations of numerosities during the developmental stage.

Stoianov and Zorzi (2012) trained a ‘deep’ networks to use pixel-by-pixel information from images through unsupervised learning. The network was only trained to efficiently code sensory data, but numerosity selectivity emerged as a statistical characteristic of the model’s final layer. Their model was able to satisfactorily explain adult human and monkey numerosity comparison task data in the higher number (estimation, not subitizing) range. The "number neurons" described by Nieder et al. (2002) and Roitman et al. (2007) are comparable to these numerosity detectors. Similar to humans, the simulations created using this model also followed Weber’s law with a weber fraction of 0.15. Both of the aforementioned models have intricate structures that enable the development of numerosity detectors.

Dakin et al. (2011) use spatial filtering to create summary statistics of the stimulus in an effort to simulate both numerosity and texture discrimination. It normalizes the amount of high spatial frequency content with the low frequency content in a stimulus, which are estimates obtained via filtering with a small and larger Laplacian of Gaussians, respectively. With a suitable choice of filter size, the high frequency response grows proportionally with set size(the number of elements). The low frequency estimator, on the other hand, exhibits hybrid behavior and grows only moderately with set size and area. It essentially implements a very fundamental principle that, with fixed element size, the number of items will correlate with the amount of borders: number judgments are made simply by looking at the high frequency content. They don’t address the sensitivity specifically, but since the model’s main noise limitation appears to be stimulus-based, one would anticipate square-root behavior rather than Weber-law behavior.

Many of the recently introduced models for number perception use backpropagation-based artificial neural networks, a network architecture gaining popularity in recent times due to its efficacy in various tasks like classification, decision-making, and learning (Nasr et al., 2019; Zanetti et al., 2019) These networks learn by modifying the weights between neurons to minimize the errors calculated using a cost function. Back propagation based artificial neural networks are usually criticized by the scientific community because these models frequently fall short of offering additional information about biological mechanisms. To acheive a better understanding of number perception, more computational models are needed that are biologically relevant and simulate the essential behavioural findings about visual number perception.

Chapter 2

Procedure

2.1 Network Architecture

Our model consists of an on-center off-surround recurrent network of 64 neurons, where every neuron has an excitatory connection with itself and inhibitory connections with all other neurons. The network is symmetric in nature and parameterized by three variables: self-excitation strength (α), mutual-inhibition strength (β), and a decay constant (λ). Network dynamics is governed by equation 1.

$$\frac{dx_i}{dt} = -\lambda x_i + \alpha F(x_i) - \beta \sum_{j=1, j\neq i}^{N} F(x_j) \ + I_i + \ Noise \tag{2.1}$$

$x_i(t)$ (written as x_i) stands for activation of the ith node at time t, $x_i(t)$ (in short x_i) is externally injected input current, which has value 1 when the ith node is being presented with an input for a short initial presentation time, and zero for the rest of the simulation. The activation function follows the formula:

$$F(x) = \begin{cases} 0 & for\ x \leqslant 0 \\ \frac{x}{1+x} & for\ x > 0 \end{cases} \tag{2.2}$$

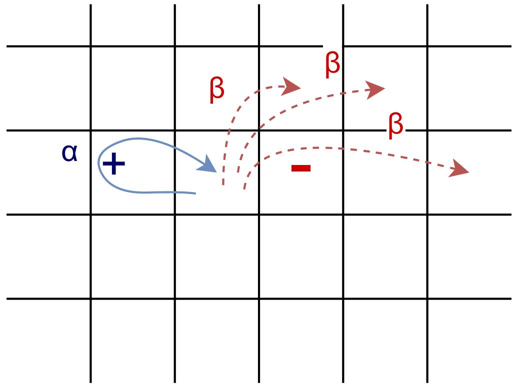

Figure 2.1: Network Diagram
Squares in the diagram represent neurons in the network. Our network consists of a single layer of 64 neurons. Each neuron has an excitatory connection (blue arrow) with itself and inhibitory connections (red arrows) with all other neurons in the network, where α denotes excitation strength and β denotes Inhibition strength.

TABLE 1 - Simulation Parameters	
Parameter	**value**
N (Number of neurons in the network)	64
α (strength of self excitation)	2.2
β (strength of mutual inhibition)	0.01-0.15
Duration of stimulus presentation (no. of time steps)	100
Total duration of simulation in time steps	5000

Table 2.1: Parameters for Network Dynamics Simulations

The noise is sampled from a normal distribution with a mean of 0 and a standard deviation of 0.03. Since changing the excitation parameter does not affect the overall nature of the data (Sengupta et al., 2014), self-excitation strength is fixed at the value of 2.2. We have kept an inventory of inhibition strengths to pick from, ranging from 0.01-0.15. The number of neurons receiving an input current during the initial presentation time is referred to as the "set size". The set size represents the numerosity of a stimulus. To simulate the neural dynamics, we have used Euler's method in Matlab software. During the presentation time, which is kept much smaller than the overall duration for the simulation (timesteps = 100 in comparison to 5000 for the full simulation), the nodes mapped to the setsize are injected with an input current (Ii) valued at 1.

When we simulate the network dynamics of the RNN network using equation(2.1), we obtain a pattern of activation across its 64 neurons as the network reaches a steady state. This steady state activation pattern of the network nodes depends both on the setsize (as an input corresponding to a numerosity) and the inhibition strength (as a network parameter). Using equation(2.3), we calculate the network's mean activation (MA) from the steady state activations and use it as the main output from of the RNN.

$$MA = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{2.3}$$

After taking the average of mean activations from 30 simulations for each combination of setsize and inhibition strength, we plot the mean activation against setsize for different inhibition strengths (figure 2.2).

2.2 Finding Minimal Number of Inhibition Strengths

When mean activation (output) varies monotonically with set size (input), it can be used as a basis for the internal representation of numerosity. We simulated ten inhibition strengths from 0.01 to 0.15, five of which are depicted in figure 2.2. out of ten inhibition strengths, We aimed to find the fewest number of inhibition strengths required such that most of the number between 1 to 50 belongs to at least one of the monotonic regions.first we include the inhibition strength of 0.01 in the list as it is the only candidate which covers numbers

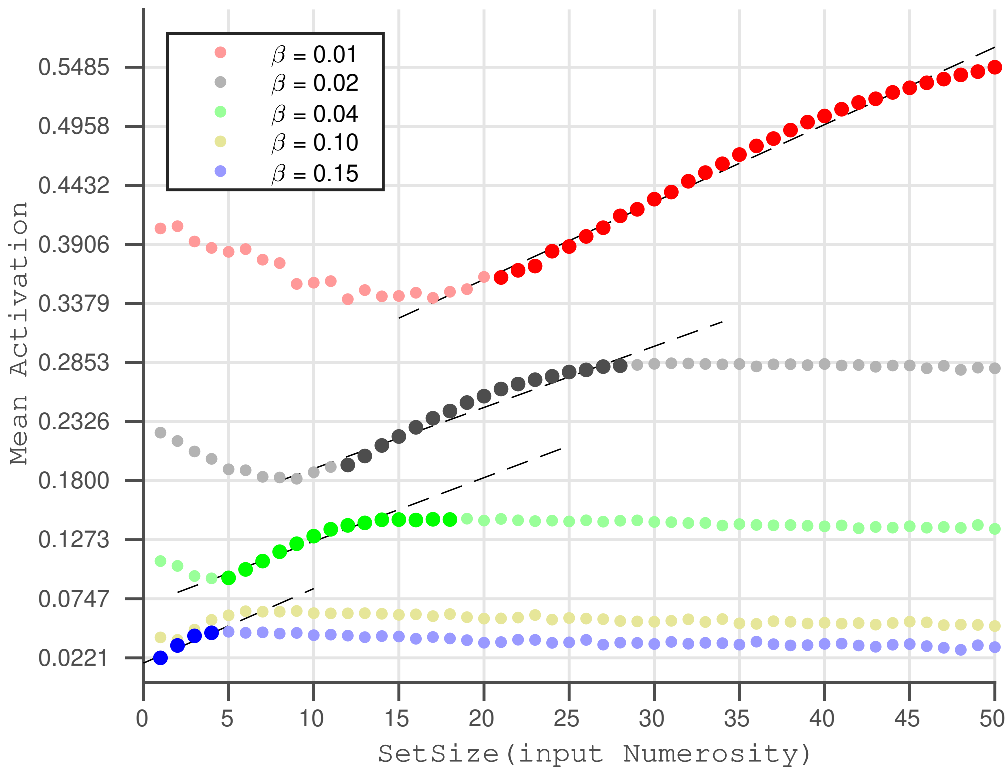

Figure 2.2: Input Numerosity Vs Mean Activation of Network
A plot between SetSize (Input Numerosity) and corresponding Mean Activation for five of the ten inhibition strengths. The darker spots denote the data points where Mean Activation increases monotonically with Setsize (input numerosity). The black dashed lines are the linear fit between the two variables, taking data only from the monotonic regions. We can observe that at strong inhibition ($\beta = 0.15$, green) mean activation increases monotonically with small numerosities while for weak inhibition ($\beta = 0.01$, red) mean activation increases monotonically with larger numerosities. There is very slight difference between curves for $\beta = 0.10$ and $\beta = 0.15$ and all other inhibition strengths between those value follow a similar pattern (not plotted here).

between 28 to 50. At extremely high inhibition ($\beta = 0.15$ *to* $\beta = 0.10$) there is little to no difference between mean activation patterns (green and yellow lines in figure 2.2), and they all have overlapping monotonic regions, starting at 1 and ending at numbers between 4 to 7. As there is no significant difference between these curve we keep all of them in hold for reconsideration. Now to cover the remaining middel regions from 5 to 20, we had three choices of inhibition strengths (0.02, 0.03 and 0.04), out of which $\beta = 0.04$ covers numbers from 5 to 18. as has the least overlap with the already selected regions and also have the least uncovered numbers (19 and 20) we chose it as the second selection for the list. Finally from the set of strong inhibitions, we chose the heighest value of $\beta = 0.15$ as it has least overlap with the regions from the already selected medium inhibition of $\beta = 0.04$.

2.3 Decoding the Mean Activation into Number Estimates

Sengupta et al. (2014) successfully demonstrate that the mean activation encodes information about the numerosity. By directly comparing the mean activation generated by various inputs, they simulate behaviorally significant outcomes like the size and distance effect. With this technique, two inputs can be compared without converting the corresponding mean activation into numerical estimates. In this section, we propose a method to decode mean activation to corresponding numerical estimates, which not only allows us to compare the magnitudes of two numerosities but also obtain number estimates for individual inputs. First, we run network dynamics simulations using all the combinations of set size (1-64) and inhibition strengths(0.01-0.15) 30 times and calculate an average of the resultant mean activation. (As for figure 2.2) For each inhibition strength, we take data points mapping setsize to mean activation solely from monotonic regions and fit a line between them. This gives us 15 functional relations ($LinearFunctions_{\beta}$) corresponding to each of the 15 inhibition strengths(β) used for simulations. By taking inverse of these linear equations we get linear quations ($InverseLinearFunctions_{\beta}$) mapping each mean activation to number estimates depending on the inhibition strength used for simulation. (Algorithm 1)

As we want an estimation of setsize from a given mean activation, we use the inverse of these linear functions (F_{β}^{-1}) to map mean activation to corresponding number estimation.

$$Number\ Estimate = InverseLinearFunctions_{\beta}(MA)$$
$$Estimation\ Error = Number\ Estimate - Setsize$$

For instance, if we acquire a mean activation of 0.53 by using a set size of 5 as input for a network with an inhibition strength of 0.1, To get the numerical estimate from the mean activation, we use the already obtained inverse function corresponding to inhibition strength (β) of 0.15.

$$4.82 = InverseLinearFunctions_{0.15}(0.53)$$
$$Estimation\ Error = 4.82 - 5 = -0.18$$

Similarly, we obtain ten different number estimates for the same input set size of 5 by employing ten different inhibition strengths. To demonstrate the results of this method, we plotted number estimations against setsize for three different inhibition strengths(0.01, 0.04, and 0.15). (figure 3.2)

To understand the general pattern of inhibition strength leading to the lowest amount of errors, we plotted inhibition strengths leading to three of the smallest estimation errors against input numerosity. (figure 3.1)

Algorithm 1 Decoding The Mean Activation For Numerosity Estimates

Input :

k (Setsize) : corresponding to numerosity from 1 to 50.

β **(Inhibition Strength) :** Set of 15 inhibition strengths from 0.01 to 0.15.

Output :

InverseLinearFunctions : Set of 15 linear functions corresponding to 15 inhibition strengths; mapping mean activation to numerosity.

1. $X_{k,\beta} \leftarrow SimulateNeuralDynamics_{\beta}(k)$ ▷ Simulates network dynamics based on equation(2.1) using Euler's method X : steady state activations of the 64 neurons

2.

$MA_{k,\beta} \leftarrow \frac{1}{N}\sum_{i=1}^{N} X_{k,\beta}(i)$ ▷ average activation of all neurons

3.

$[k_{\beta}^{min} : k_{\beta}^{max}] \leftarrow LargestMonotonicRegion(MA_{k,\beta})$ ▷ Function to find longest regions where MA increases monotonically with Setsize

4.

$[MA_{\beta}^{min} : MA_{\beta}^{max}] \leftarrow S([k_{\beta}^{min} : k_{\beta}^{max}])$ ▷ Calculate mean activations corresponding to monotonic regions

5.

$LinearFunctions_{\beta} \leftarrow LinearFit([k_{\beta}^{min} : k_{\beta}^{max}], [MA_{\beta}^{min} : MA_{\beta}^{max}])$ ▷ Fit lines to data from monotonic regions only. LinearFunctions map numerosity to mean activation depending on inhibition strength used for network dynamics

6.

$InverseLinearFunctions_{\beta} \leftarrow Inverse(LinearFunctions_{\beta})$ ▷ By finding inverse of the linear functions we get one equation for each of the inhibition strengths, these equations map mean activations to number estimates.

Algorithm 2 Finding The Most Suitable Inhibition Strength For A Given Input

Input :
k (setsize): input current to the network.

Output :
β^{best} : The most suitable inhibition strength for the given input k(Setsize).

1.
$MA_{\beta} \leftarrow S_{\beta}(k)$ ▷ Find a set of mean activation values using every inhibition strength(β)

2.
$MA^{-}_{\beta} \leftarrow MA_{\beta} - \Delta MA$
$MA^{+}_{\beta} \leftarrow MA_{\beta} + \Delta MA$ ▷ Add and substract small amounts to the mean activation

3.
$Estimate^{-}_{\beta} \leftarrow InverseLinearFunction_{\beta}(MA^{-}_{\beta})$
$Estimate^{+}_{\beta} \leftarrow InverseLinearFunction_{\beta}(MA^{+}_{\beta})$ ▷ Find number estimates from mean activation (see Algorithm(1))

4.
$\Delta Estimate_{\beta} \leftarrow Estimate^{+}_{\beta} - Estimate^{-}_{\beta}$ ▷ change in magnitude of estimates as a reult of small change in input

5.
$i \leftarrow IndexForMaximum(\Delta Estimate_{\beta})$ ▷ Find index of β which gives maximum $\Delta Estimate$

6.
$\beta^{best} \leftarrow \beta(i)$ ▷ Inhibition Strength which results in maximum change in number estimation with a variation in mean activation

Chapter 3

Result and Discussion

3.1 Emergence of Different Ranges of Numbers

By simulating the network dynamics using fifteen different levels of inhibition with all the inputs from Setsize 1 to 50, we get datapoints mapping Setsize to Mean Activation corresponding to each of the inhibition strengths. We took the average of 30 simulations for each datapoint and plotted the relationship between Setsize and the Mean Activation for five inhibition strengths (Figure 2.2). In the figure, we can see that the input numerosity and mean activation follow different relationships depending on the inhibition strength used for simulation. For a very strong inhibition ($\beta = 0.10$ to $\beta = 0.15$), the relationship is quite similar, where mean activation increases with Setsize until numbers 4 to 7, then slowly decreases with an almost flat curve for larger numerosities. Whereas for medium to low inhibition strengths ($\beta = 0.01, \beta = 0.02, \beta = 0.04$ in the plot), first mean activation decreases with an increase in Setsize then it increases monotonically for a range of numerosities (denoted by darker dots in the plot) followed by a flatter region for even larger numerosities. We observe that the monotonic regions cover larger numbers in cases of smaller inhibition strengths and vice versa. By minimizing the number of different inhibition strengths required so that their combined monotonic regions cover the full stretch of numerosity, we ended up with three inhibition strengths ($/beta = 0.01, \beta = 0.04, \beta = 0.15$) and three ranges of numbers (1:4, 5:17, and 21:50) corresponding to them. Behavioral studies demonstrate the presence of separate intervals of numbers based on the differences in their psychophysical properties (or numerical cognition effects), such as accuracy and reaction time Kaufman et al. (1949); Whalen et al. (1999). The differences between the subitizing range (from 1 to 4) and the estimation range (greater than 5) are among the most studied observations in number perception research. In a recent study, Portley and Durgin (2019) suggested the presence of another number-estimation elbow, further dividing the estimation range into two intervals, below and above number 20. Our model simulates the emergence of these number intervals with surprisingly similar boundaries as a result of the minimal coverage of the monotonic relation between network input (set size) and output (mean activation).

3.2 Estimation of Numerosity

When we examine the error patterns for all combinations of set size and inhibition strength (figure 3.2), we find that no inhibition strength produces a prediction that is accurate across the entire range of numerosity. Higher inhibition strengths make fewer errors in the case of small numbers, and lower inhibitions work best for larger numbers: For example, as seen in (figure 3.2), a high inhibition (0.15) network accurately estimates small numbers (1-5) and underestimates larger numbers (6-60). The low inhibition (0.04) network overestimates the small numbers (1-5) and gives accurate estimates for larger numbers (6-20). In figure 3.1, we observe that for larger numerosity, the inhibition strength leading to the least amount of error(big blue dots) is consistently equal to 0.01. While the inhibition strengths corresponding to the second and third lowest errors are progressively stronger ones (0.02 and 0.03). For the smaller numbers in the range of one to ten, the inhibition strengths responsible for the lowest errors are not very consistent. However, we see that for smaller numerosity, on average, a larger inhibition strength produces smaller errors.

In psychophysics studies, it is a common practice to check whether a stimulus is getting perceived more (Overestimation), or less (underestimation) than its real magnitude. In figure 3.2, we see that while using a high inhibition strength(0.15), our model underestimates larger numerosities(geater than 5). Whereas using a low inhibition strength(0.01), It overestimates smaller numbers(1-20). But when we look at the curve in the curve for (β=0.01) in a small numerosity range, we see an inverse relation between numerosity and estimation, i.e. with every increment in numerosity it produces smaller and smaller estimations. Such internal representations are not meaningful and can be easily discarded. When

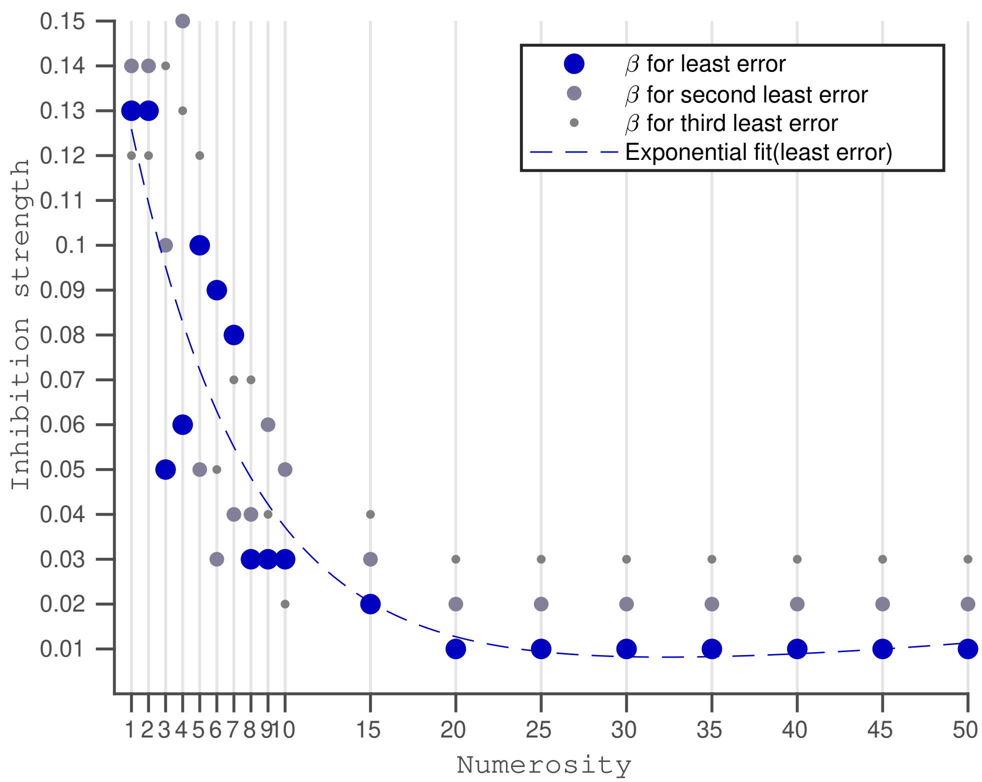

Figure 3.1: Numerosity Vs Inhibition Strengths Producing Lowest Error
Using fifteen network inhibition levels, we get fifteen different number estimates for an input. For each input numerosity (x-axis), we plot the three inhibition strengths that result in the three lowest errors (y-axis). We fitted an exponentially decreasing curve to show the pattern between input numerosity and the mean activations leading to the least amount of error. For smaller numbers, high inhibition strength results in a minor error in estimation, and vice versa. We can see that for input numerosity 1, an inhibition strength of 0.13 results in the lowest error, while for input numerosity 50, an inhibition strength of 0.01 results in the lowest error.

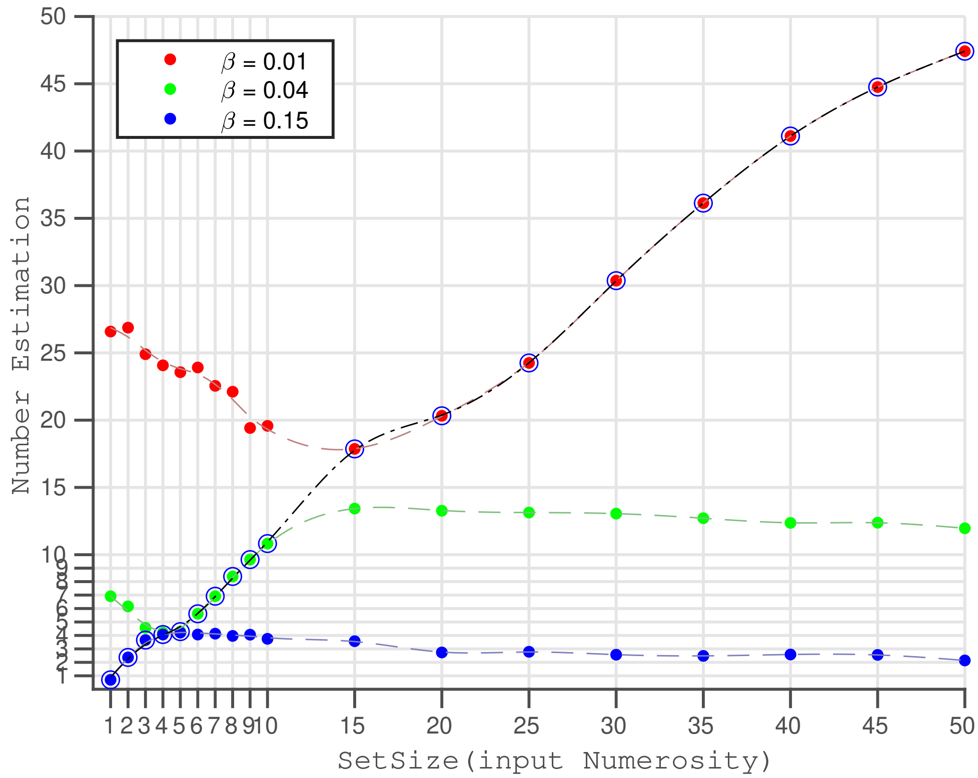

Figure 3.2: Input Numerosity Vs Number Estimations
For three inhibition strengths, SetSize(Input Numerosity) is plotted against the corresponding number estimates. Network with high inhibition(0.15) estimates small numbers (1-4) accurately but underestimates larger numbers. None of the inhibition strengths are effective across the entire range of numerosities. We get a curve (black) that can provide reasonable estimation for the entire range of numerosity by varying the inhibition strength depending on the input (Algorithm 2)

we ignore all the regions where input numerosity and number estimation have a negative correlation(negative slope), we see that for any inhibition strength, every numerosity either gets estimated correctly or underestimated. This result has some important implications for adaptation effects, analysed in depth in a subsequent section.

3.3 Regulation of Network Inhibition

In the previous section, we offered a technique for extracting numerosity information from the network's mean activation. We demonstrated that, depending on the input's numerosity, using some levels of network inhibition results in better number estimation than others. In this section, we will investigate an algorithm to control network inhibition, allowing the same network to generate the best possible estimates across the entire range of numbers.

It is unlikely to get a good number estimation if a combination of numerosity and inhibition strength does not produce a monotonic relationship between numerosity and mean activation. We got a set of three inhibition strengths by minimizing the number of different inhibition strengths required to ensure the monotonicity of mean activation for all possible inputs. When we categorize numbers based on which of the three inhibition strengths they match, we end up with three ranges of numbers.

Through (figure 2.2) and (figure 3.1), we can see the importance of monotonicity for estimation accuracy. In (figure 2.2), high inhibition strength (0.15) produces monotonicity for numerosities ranging from 1 to 4. Similarly, in (figure 3.1), we see that high inhibition strengths (0.15, 0.14, and 0.13) result in the lowest errors for those numerosities. When we assign the network inhibition strength based on the interval in which the input numerosity falls, we get a favourable estimate. For example, if we have the current input with numerosity of 2, we can use the inhibition strengths of 0.13 to 0.15, as with these inhibition strengths, the output of the network increases monotonically with input. Using these inhibition strengths gives us an estimation with the most minor error, as we see in figure 3.1; for numerosity 2, high inhibition strength produces the smallest errors. Although this illustration of selecting inhibition strength aids our understanding of the model, there is a fundamental reason why it cannot serve as a stand-alone model for network regulation. Here we have used a specific inhibition strength Based on the input's numerosity. However, because number estimation is the last step, the system does not know the current input's numerosity when choosing the inhibition level. We have implemented a sensitivity-based algorithm to choose a suitable inhibition strength rather than choosing the inhibition based on the input's numerosity.

To illustrate the concept, we use the set of three inhibition strengths obtained in the previous step, which guarantees that we have at least one inhibition strength where the condition of monotonicity is satisfied for every potential input to the network. Using the algorithm in appendix 2, we determine the sensitivity of the output (number estimate) to the change in input for a given input. We get three values for the sensitivities based on the three levels of inhibition in the network. The algorithm then selects the level of inhibition for which the sensitivity is the highest. For instance, for the input of setsize 2, the output from the network with an inhibition strength of 0.15 turns out to be the most sensitive compared to the other inhibition strengths. As a result, the algorithm uses an inhibition strength of 0.15 for this specific input. Finally, we obtain the estimate of 1.9 using the network inhibition strength of 0.15, which is more precise than the estimates we would have otherwise obtained using the other inhibition strengths.

We plot setsize against number estimations using the corresponding choices of inhibition strength for all possible inputs (Figure **??** The data points obtained using the selected inhibition strengths are shown in the figure as black circles. As anticipated, we discover that for the entire range of inputs, the estimations using the chosen inhibition strength produce the most precise number-estimations compared to other inhibition strengths.

To show that the mean activation of the network encodes the numerosity of the input, we have decoded the mean activation to a continuous scale of numbers. We used linear fitting to establish the connection between mean activation and numerosity. Using a curve fitting method to map mean activation to numerosity brings a limitation to our model. Since we only use data from monotonous regions, the quality of the fit varies with the length of the monotonous regions. As a result, the mapping of mean activation to a number estimate is less reliable because the subitizing range has a relatively shorter monotonous region. Due to this shortcoming, the final output does not reflect the observed difference between small and large numbers in terms of accuracy.

3.4 Role of Continuous Attributes

Number perception is also influenced by other visual attributes of the objects in a scene, such as density, size, and convex hull. Gebuis and Reynvoet (2012); **?** As discussed previously, The Sensory Integration Theory argues that the brain derives numerosity from these continuous cues (Gebuis et al., 2016; Dakin et al., 2011). Some studies have opposed the idea by arguing that sensory cues are insufficient in themselves to produce an accurate estimation of numerosity. Even if the sensory cues can't generate an accurate estimation of numerosity, it is possible to develop a preliminary idea about how small or big the numerosity could be.

We suggest a possibility that The information from the continuous visual attributes such as density and size can be used to extract an initial guess about the possible number range the input belongs to. This low-accuracy information about the magnitude of numerosity can be used for regulation of inhibition strength, thereby obtaining a more accurate estimation of the numerosity. For instance, while taking an input As our network model takes input from a filtered out information in visual stream, and does not care about topography, size and shape of individual spots, we need to add some further structure preceding the RNN network in the visual stream to computationally demonstrate this assumption.

3.5 Adaptation

Psychophysics experiments have shown that perception of numerosity are subject to adaptation (Fornaciai et al., 2016; Burr et al., 2011). when humans are exposed to multiple stimulus with certain numerosity ("adaptor") for an extended period, the apparent numerosity of a subsequent ("test") stimulus gets distorted. Adaptation is commonly thougth to be an hallmark of "primary" perceptual attributes of vision such as color, orientation, size, or density. Also like many of these primary perceptual attributes, perception of numerosity follows Weber's law, which led many researchers to consider it as a "primary visual feature"(Burr and Ross, 2008). In this study, we attempted to model high-level visual processing of numerosity that accepts input devoid of primary visual attributes. The variability of the network inhibition strength plays a vital role in our model because it allows for accurate internal representation of the whole range of numerosity. This section will discuss how the need for dynamic network parameters (here, inhibition strength) can relate to adaptation effects.

As discussed in previous sections, our model needs to employ different inhibition strengths for the network depending on the numerosity range. Our model is very simplistic in nature and is far from the complexity of biological networks with multiple degrees of freedom. We can assume that regulation of a particular or a combination of parameters across the network, such as inhibition strengths, can affect how accurately the numerosity gets internally represented in the brain. These processes might be reflected in behavioral observations because of the time and error associated with them. Assuming, with each instance of processing a certain range of numerosity, the network gets closer to the most appropriate inhibition strength; it would also increase the accuracy with each iteration. This way, we can draw a parallel between the adaptation of network parameters and the adaptation to numerosity.

It takes us relatively longer(increase in reaction time) to enumerate or compare a number when we switch from an adapted numerosity range. Additionally, the accuracy decreases immediately after the switch and then improves again with repeated exposure to the same range. The additional time required to select an appropriate network parameter might have a connection to the "switch cost" in the form of a delay in response. Using a network parameter suitable for the previously adapted numerosity might have a relationship with the observed drop in accuracy immediately after the switch. Exploring the possibility of such adaptation mechanisms in the higher levels of the number processing pathway can help us go beyond the conceptualization of numerosity as a primary visual feature.

In (figure 3.2) we see that the inhibition strength of 0.04 is suitable for numerosity between 6 and 20. Suppose the network is set to an inhibition of 0.04; let's call it as 'adapted' to numbers in the range of 6 to 17. When we suddenly provide an input corresponding to a larger numerosity (say 30), ideally, the network should shift to a lower inhibition strength (0.01) to give an accurate result. But if it uses the same inhibition strength of 0.04, it would underestimate the numerosity to be around 20 (on the green curve). Similarly, whenever we use an inhibition strength adapted to smaller numerosity on inputs with larger numerosity, we get an error in the form of underestimation. But when we use inhibition strength adapted to larger numerosity for smaller numerosity, we get an overestimation of numerosity. The plot fails to demonstrate the underestimation observed in behavioral experiments while switching from larger to smaller numerosities. But the regions in the plot which correspond to overestimation (numerosity from 1 to 10 on the red curve), the estimation of numerosity decreases with increase in input numerosity. As this does not follow the primary condition of monotonic increasing relationship we discard this region.

Castaldi et al. (2016) classified the brain activity recorded during numerosity perception before and after psychophysical adaptation to numerosity using multivariate pattern recognition. Using a support vector machine, they used BOLD responses from the IPS to classify numerosity successfully. Based on the observation that training the model with pre-adaptation responses did not classify numerosity while testing on post-adaptation data and vice-versa, they suggest that adaptation changes the neuronal representation of numerosity in the brain. Similarly, our model shows that neuronal representation of numerosity changes with inhibition strength. As shown in figure (), we get different curves mapping numerosities to mean activations for each inhibition strength. As we have suggested that adaptation is essentially achieved by changing network inhibition, our model also supports the claim by Castaldi et al. (2016) that adaptation changes the neural representation of numerosity. They showed that the degree of adaptation increases the discriminability of the IPS's BOLD response. In our model, the mean activation of the network is used as the basis for the magnitude of numerosity. When using the most appropriate inhibition strength, the mean network activation is most sensitive (has the highest slope) to the change in numerosity. We once again draw a comparison between adaptation and the control of network inhibition and by observing that the greater slope of the mean activation leads to greater discriminability.

Previously, we have discussed some of the possible ways to choose the inhibition of the network to suit individual inputs. We suggested that a method that can generate an approximate prediction of the range of numerosity can be used to choose inhibition strength for the network, which in turn gives a more accurate internal representation of numerosity. One of those ideas involved generating an approximation of numerosity by using continuous visual cues like size, density, and texture. In such a case, these continuous attributes indirectly control the internal representation of the numerosity. It is important to distinguish between this kind of influence and the suggestion made by the sensory integration theory, which holds that the knowledge of numerosity is entirely derived from the sensory cues. Here, only the level of network inhibition is preconditioned by the sensory cues so that it is prepared to work with a normalized and segmented input. Durgin (1995) claimed that adaptation to texture density affects the perception of numerosity. They have argued that our brain does not adapt to numerosity but to the texture density of a stimulus. Other studies have also shown that adaptation to other visual cues like size also affect numerosity perception (Zimmermann and Fink, 2016). But how does adaptation to density affect the perception of numerosity? Does it imply that our brain perceives numerosity using a sensory integration system? We argue that the influence on numerosity perception by adaptation to density need not imply the presence of a sensory integration system for number perception. If our assumption is true that sensory cues like texture and density are used to prepare the network to process numerosity, then it also explains why adaptation to texture influences numerosity related tasks.

3.6 Analog Output and Magnitude Perception

Computational models for number perception can roughly be divided into two groups based on how the numerosity is encoded. In the first group, the final layer of a network contains a set of neurons, where each neuron is assigned a specific number (Verguts and Fias, 2004)

(Stoianov and Zorzi, 2012). Selective activation of a neuron indicates the detection of a number assigned to it. These neurons are often compared with the number-neurons observed in the brain. Most of these models work for only a small range of numbers. The models in the other category often use a single output to simulate the encoding of numerosity (Dehaene and Changeux, 1993) (Sengupta et al., 2014). Here, an increase in the output's magnitude signifies an increase in the numerosity. These models are usually more flexible in terms of their working range with numbers. A common scale allows a simple explanation of numerosity comparison tasks, As the magnitude can be directly compared to choose a winner. Our model belongs to the second group, consisting an analog internal representation of numerosity. According to the A Theory of Magnitude (ATOM) model, space, time, and numbers all interact with one another (Walsh, 2003). The model proposes that a single analog magnitude system processes time, numbers, and space. Our model's ability to encode numerosity in an continuous quantity makes it useful for the creation of general models for magnitude perception.

.

Chapter 4

Conclusion

.

We have used a Recurrent Neural Network (RNN) with on-centre off-surround neural connections to computationally model various aspects of visual perception of numbers. The input for the RNN is inspired by the Normalized Object Location Map (OLM) (Dehaene and Changeux, 1993), and the mean of the steady-state activation of the network nodes is taken as the output from the network. The model successfully emulates some of the significant findings from the behavioral studies regarding Weber fraction, number comparison, and reaction time (Sengupta et al., 2014). In this study we have extended the model, first by introducing a method for decoding the mean activation into number estimate, then for regulating the network inhibition to achieve workability with an extended range of numerosity.

Behavioral studies have demonstrated that our capacity to perceive numerosity varies depending on which range of numerosity a stimulus belongs to. While we can quickly and accurately enumerate or discriminate numbers in the subitizing range (1 to 4), the numbers in the estimation range (greater than 5) require a longer time, and their discrimination follows Weber's law (Kaufman et al., 1949; Whalen et al., 1999). Recent findings suggest the presence of an additional elbow at number 20, giving us three ranges of numbers with a sudden change in number cognition Portley and Durgin (2019). We lack an understanding of the underlying mechanisms leading to these behavioral discontinuities. In our computational model, we use the mean steady-state activation of all the neurons in the network as the output from the network. We assume that, for the mean activation of the network to code for numerosity, it should follow a monotonic relationship with input numerosity. We observe that to keep the output from the network monotonous with the input numerosity, we need to employ multiple inhibition strengths (figure 2.2). By minimizing the total number of inhibition strengths required to maximally cover the entire range of numerosity, we arrive at three inhibition strengths ($\beta = 0.01, \beta = 0.04, \beta = 0.15$) and three corresponding ranges of numbers (1:4, 5:17, and 21:50), closely matching those observed in behavioral studies. The emergence of different ranges of numbers as a consequence of optimizing a network with limited capacity gives us a new perspective on possible mechanisms behind the elbows (or breaks) observed in numerosity perception. It shows that the discontinuity in our number cognition abilities does not necessitate the presence of dedicated neural networks for each range of numbers.

In the latter part of our study, we have suggested two potential methods to control the network inhibition so that the network can produce the best possible estimation for a broader range of inputs. In the first method, for a given input, we choose the level of inhibition for which the network output is most sensitive to a slight change in the input. We get reasonable number estimates for the entire range of numerosities by using the inhibition strengths chosen by this criteria, making our network one of the few computational models to work on an extended domain (1 to 50) of numerosities (figure 3.2). Numerous studies have demonstrated the possibility of deriving information on numerosity from continuous attributes in early visual pathways Gebuis et al. (2016); Dakin et al. (2011); Morgan et al. (2014). In the second method, we have argued that an inaccurate but faster estimation of numerosity from early visual pathways can help choose an inhibition strength appropriate to the current input. Further, this chosen inhibition strength can prepare the network to generate an accurate representation of numerosity from a normalized input: devoid of primary visual attributes. This way, we argue for a hybrid system that employs ideas from the Sensory Integration Theory (SIT) to get an initial guess of numerosity and Approximate Number System (ANS) theory to reach the final numerosity estimate using normalized information from deeper stages of the visual pathway.

Our model gives a computational account for some aspects of the adaptation effects observed in psychophysics studies. We suggest that adaptation to a specific range of numerosity can result from imperfect and delayed regulation of network inhibition. The extra time needed to find a suitable network parameter might be related to the switch-cost in the form of an increase in reaction time, and the drop in accuracy immediately after a switch might result from the use of a network parameter that is no longer suitable for the input. If our assumption is true that sensory cues like texture and density are used to prepare the network to reach a more accurate estimate, then it also explains why adaptation to texture influences numerosity-related tasks (Durgin, 1995).

Our model of number perception has a lot of room for growth; as the model represents numerosity in a continuous manner, further work is needed to explore compatible processes leading to the symbolic grounding of numbers. Based on the current findings, we have made some predictions about the roles of primary visual cues in this study. However, in order to computationally prove the predictions, a specific algorithm must be developed. We acknowledge that biological neural networks are not as simple as our model, and do not assert that the addressed behavioral observations are result of the dynamics of such simple networks. However, our model provides novel insights into the possible computations behind many fascinating aspects of the visual perception of numerosity, which can aid in developing a more biologically relevant model for visual number perception, and number cognition in general.

Bibliography

Anobile, G., Stievano, P. and Burr, D. C. (2013). Visual sustained attention and numerosity sensitivity correlate with math achievement in children. Journal of experimental child psychology *116*, 380–391.

Balakrishnan, J. and Ashby, F. G. (1991). Is subitizing a unique numerical ability? Perception & Psychophysics *50*, 555–564.

Burr, D., Anobile, G. and Turi, M. (2011). Adaptation affects both high and low (subitized) numbers under conditions of high attentional load. Seeing and Perceiving *24*, 141–150.

Burr, D. and Ross, J. (2008). A visual sense of number. Current biology *18*, 425–428.

Burr, D. C., Turi, M. and Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. Journal of Vision *10*, 20–20.

Butterworth, B. (2005). The development of arithmetical abilities. Journal of child psychology and psychiatry *46*, 3–18.

Cantlon, J. F., Platt, M. L. and Brannon, E. M. (2009). Beyond the number domain. Trends in cognitive sciences *13*, 83–91.

Castaldi, E., Aagten-Murphy, D., Tosetti, M., Burr, D. and Morrone, M. C. (2016). Effects of adaptation on numerosity decoding in the human brain. NeuroImage *143*, 364–377.

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. and Morgan, M. J. (2011). A common visual metric for approximate number and density. Proceedings of the National Academy of Sciences *108*, 19552–19557.

Dehaene, S. (1992). Varieties of numerical abilities. Cognition *44*, 1–42.

Dehaene, S. (2009). Origins of mathematical intuitions: The case of arithmetic. Annals of the New York Academy of Sciences *1156*, 232–259.

Dehaene, S. and Changeux, J.-P. (1993). Development of elementary numerical abilities: A neuronal model. Journal of cognitive neuroscience *5*, 390–407.

Dehaene, S., Dehaene-Lambertz, G. and Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. Trends in neurosciences *21*, 355–361.

Durgin, F. H. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. Journal of Experimental Psychology: Human perception and performance *21*, 149.

Egeth, H. E., Leonard, C. J. and Palomares, M. (2008). The role of attention in subitizing: Is the magical number 1? Visual Cognition *16*, 463–473.

Feigenson, L., Dehaene, S. and Spelke, E. (2004). Core systems of number. Trends in cognitive sciences *8*, 307–314.

Fornaciai, M., Cicchini, G. and Burr, D. (2016). Adaptation to number operates on perceived rather than physical numerosity. Cognition *151*, 63–67.

Fornaciai, M. and Park, J. (2017). Distinct neural signatures for very small and very large numerosities. Frontiers in Human Neuroscience *11*, 21.

Fornaciai, M. and Park, J. (2018). Early Numerosity Encoding in Visual Cortex Is Not Sufficient for the Representation of Numerical Magnitude. Journal of Cognitive Neuroscience *30*, 1788–1802.

Gallistel, C. R. and Gelman, R. (1992). Preverbal and verbal counting and computation. Cognition *44*, 43–74.

Gebuis, T., Kadosh, R. C. and Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. Acta psychologica *171*, 17–35.

Gebuis, T. and Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. Journal of Experimental Psychology: General *141*, 642.

Gelman, R. and Gallistel, C. R. (1986). The child's understanding of number, vol. 1,. Harvard University Press.

Gescheider, G. A. (1997). Psychophysical measurement of thresholds differential sensitivity. Psychophysics: the fundamentals *69*, 1–15.

Halberda, J., Mazzocco, M. M. and Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. Nature *455*, 665–668.

Harvey, B. M., Klein, B. P., Petridou, N. and Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. Science *341*, 1123–1126.

Hauser, M. D., Tsao, F., Garcia, P. and Spelke, E. S. (2003). Evolutionary foundations of number: spontaneous representation of numerical magnitudes by cotton–top tamarins. Proceedings of the Royal Society of London. Series B: Biological Sciences *270*, 1441–1446.

Kaufman, E. L., Lord, M. W., Reese, T. W. and Volkmann, J. (1949). The discrimination of visual number. The American journal of psychology *62*, 498–525.

Morgan, M. J., Raphael, S., Tibber, M. S. and Dakin, S. C. (2014). A texture-processing model of the 'visual sense of number'. Proceedings of the Royal Society B: Biological Sciences *281*, 20141137.

Nasr, K., Viswanathan, P. and Nieder, A. (2019). Number detectors spontaneously emerge in a deep neural network designed for visual object recognition. Science advances *5*, eaav7903.

Nieder, A., Diester, I. and Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. Science *313*, 1431–1435.

Nieder, A., Freedman, D. J. and Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. Science *297*, 1708–1711.

Piazza, M. (2011). Neurocognitive start-up tools for symbolic number representations. Space, time and number in the brain *55*, 267–285.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., Dehaene, S. and Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. Cognition *116*, 33–41.

Piazza, M., Fumarola, A., Chinello, A. and Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. Cognition *121*, 147–153.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D. and Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. Neuron *44*, 547–555.

Pica, P., Lemer, C., Izard, V. and Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. Science *306*, 499–503.

Portley, M. and Durgin, F. H. (2019). The second number-estimation elbow: Are visual numbers greater than 20 evaluated differently? Attention, Perception, & Psychophysics *81*, 1512–1521.

Roitman, J. D., Brannon, E. M. and Platt, M. L. (2007). Monotonic coding of numerosity in macaque lateral intraparietal area. PLoS biology *5*, e208.

Sengupta, R., Surampudi, B. R. and Melcher, D. (2014). A visual sense of number emerges from the dynamics of a recurrent on-center off-surround neural network. Brain research *1582*, 114–124.

Stoianov, I. and Zorzi, M. (2012). Emergence of a'visual number sense'in hierarchical generative models. Nature neuroscience *15*, 194–196.

Trick, L. M. and Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. Psychological review *101*, 80.

Verguts, T. and Fias, W. (2004). Representation of number in animals and humans: A neural model. Journal of cognitive neuroscience *16*, 1493–1504.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. Trends in cognitive sciences *7*, 483–488.

Whalen, J., Gallistel, C. R. and Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. Psychological science *10*, 130–137.

Xu, F. and Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. Cognition *74*, B1–B11.

Xu, X. and Liu, C. (2008). Can subitizing survive the attentional blink? An ERP study. Neuroscience Letters *440*, 140–144.

Zanetti, A., Testolin, A., Zorzi, M. and Wawrzynski, P. (2019). Numerosity Representation in InfoGAN: An Empirical Study. In International Work-Conference on Artificial Neural Networks pp. 49–60, Springer.

Zimmermann, E. and Fink, G. R. (2016). Numerosity perception after size adaptation. Scientific Reports *6*, 1–7.