# FOUNDATIONS OF MEMORY CAPACITY IN MODELS OF NEURAL COGNITION

A Thesis

presented to

the Faculty of California Polytechnic State University,

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Computer Science

by

Chandradeep Chowdhury

December 2023

COMMITTEE MEMBERSHIP

TITLE: Foundations of memory capacity in models of neural cognition

AUTHOR: Chandradeep Chowdhury

DATE SUBMITTED: December 2023

COMMITTEE CHAIR: Mugizi Robert Rwebangira, Ph.D.
Assistant Professor of Computer Science

COMMITTEE MEMBER: Rodrigo De Moura Canaan, Ph.D.
Assistant Professor of Computer Science

COMMITTEE MEMBER: Theresa Anne Migler, Ph.D.
Assistant Professor of Computer Science

ABSTRACT

Foundations of memory capacity in models of neural cognition

Chandradeep Chowdhury

A central problem in neuroscience is to understand how memories are formed as a result of the activities of neurons. Valiant's neuroidal model attempted to address this question by modeling the brain as a random graph and memories as subgraphs within that graph. However the question of memory capacity within that model has not been explored: how many memories can the brain hold? Valiant introduced the concept of interference between memories as the defining factor for capacity; excessive interference signals the model has reached capacity. Since then, exploration of capacity has been limited, but recent investigations have delved into the capacity of the Assembly Calculus, a derivative of Valiant's Neuroidal model. In this paper, we provide rigorous definitions for capacity and interference and present theoretical formulations for the memory capacity within a finite set, where subsets represent memories. We propose that these results can be adapted to suit both the Neuroidal model and Assembly calculus. Furthermore, we substantiate our claims by providing simulations that validate the theoretical findings. Our study aims to contribute essential insights into the understanding of memory capacity in complex cognitive models, offering potential ideas for applications and extensions to contemporary models of cognition.

# ACKNOWLEDGMENTS

Thanks to:

- My parents, and grandparents.
- My thesis committee members.
- My collaborators Patrick Perrine, and Shosei Anegawa.
- My roommate, Zachary Ward.
- Cal Poly Graduate Education Office, for supporting me with a tuition waiver for academic year 2022-2023
- Cal Poly Cares, for supporting me with a Grant.
- Cal Poly Housing Administration, for supporting me with emergency housing for two quarters.

TABLE OF CONTENTS

LIST OF FIGURES

# Chapter 1

# INTRODUCTION

Memories formed in the brain emerge from intricate patterns of neural activity involving specific subgroups of neurons. A pioneering model exploring this complex phenomenon is Valiant's Neuroidal model, introduced in his 2005 paper "Memorization and association in a realistic neural model", which represents memories as random subgraphs over a larger base graph modeling the connectivity of the brain [32]. Activating a certain proprotion of neurons in a memory causes the memory to fire, and effectively be retrieved from this network. Valiant provided two main algorithms - JOIN, for forming new memories, and LINK, for associating pre existing memories. One unique aspect of this model is that all memories, pre-existing and newly formed, must approximately have the same size to behave like 'equal citizens' in the system. To enforce this, Valiant introduced a set of six equations that the memory size must follow. For a given configuration of the system, there are unique integral solutions to this system of equations that Valiant referred to as the *replication factor*.

Valiant investigated two versions of the Neuroidal model - a disjoint version where memories do not intersect, and a more biologically plausible shared version where memories are allowed to intersect. The capacity of the disjoint version can be easily estimated - it is the number of neurons in the graph divided by the replication factor. To study the capacity of the shared memory model, Valiant introduced a notion of interference between memories. This refers to the unintended firing of one memory when another is activated, caused by overlapping subgroups of neurons. As interference accumulates from storing more memories, quality of retrieval degrades -

false firing escalates, signaling the network has hit its memory capacity. Valiant left quantitative characterization of this capacity for future work.

Since then, formal analysis of the neuroidal framework's storage capabilities remains limited. Valiant himself reinvestigated the capacity of the Neuroidal model with respect to LINK in 2017 [34]. However, he did not analyze the capacity with respect to JOIN, which we consider more interesting as it is primary memory creation tool in the model. Recently, Perrine empirically investigated the capacity of the Neuroidal model with respect to JOIN and provided some valuable insights into the problem [27]. There also have been recent empirical investigations into capacity in the context of the Assembly Calculus, a descendant of Valiant's model that uses Project and Merge, two more advanced memory formation operations inspired by LINK and JOIN respectively [35]. However, broader open questions persist regarding formulating general capacity theories spanning diverse random graph based models of cognition that capture essential interference phenomena governing information storage.

In service of this goal, we take foundational steps in this paper toward a rigorous capacity formulation for overlapping subset models in terms of expected interference between memories. We start by providing precise definitions for memory capacity and interference for a system of subsets over a finite base set. In contrast with the complex memory generation process used in Valiant's and other contemporary models, we initially consider random subset insertion to enable simpler closed-form solutions that can also be updated to account for the intricacies of different memory generation algorithms. Under simplifying assumptions, we derive expressions characterizing capacity, roughly defined as the maximum number of subsets that can be stored in the system before expected interference from adding more memories breaches intolerable thresholds. We also simulate the Neuroidal model inspired by Perrine's simulation in

his thesis and swap out JOIN with random subset insertion to empirically validate our results [27].

While mathematically convenient for an initial analysis, we understand random memory formation lacks biological plausibility. Therefore, we discuss strategies to adapt our interference calculations to represent specialized memory creation algorithms used in existing neural models, without compromising the generality of our overall capacity theory. As a case study, we analyze capacity in the Neuroidal model with respect to the JOIN operation. We simulate memory formation under JOIN, gaining preliminary insights into challenges to adapting our theory. Findings reveal uneven accumulation of interference on certain neurons, in contrast to the simplifying uniformity assumptions in our derivations.

Overall, this work initiates rigorous groundwork to elucidate the memory storage limitations of neural systems in light of interference. We substantiate our formulations with simulations that validate capacity findings under simplifying assumptions of random memory formation. The analytical capacity expressions and strategies proposed to handle complex memory creation algorithms offer potential springboards to tackle outstanding questions in exploring storage dynamics of contemporary cognitive models. They provide formal bases to investigate applications to long-standing frameworks like Valiant's neuroidal model and active areas like the Assembly Calculus.

Chapter 2

# BACKGROUND

In this chapter, we summarize the key characteristics of the Neuroidal model and Assembly Calculus relevant to our work. Both models are based on Erdős-Rényi $G(n, p)$ random graphs with nodes representing neurons and subgraphs representing memories.

## 2.1 Neuroidal model

In his 1994 paper, Valiant proposed the Neuroidal model to study memory formation and association in neural systems in a biologically plausible manner [31]. The Neuroidal model consists of a network of neuroids, which are simplified model neurons. Each neuroid is a threshold unit connected to other neuroids via directed, weighted synapses. Valiant constrained the neuroidal model using four key biological parameters observed in cortex: total neuron count, synapses per neuron, synaptic strength relative to threshold, and neuron switching times.

Later in 2005, Valiant developed algorithms for two basic cognitive functions, JOIN and LINK, within the neuroidal model while respecting realistic biological constraints [32]. JOIN implements memory formation - if representations for items A and B already exist in the network, JOIN modifies the network so a new representation C fires if and only if both A and B fire. LINK implements association - if A and B are already represented, LINK modifies the network so activation of A causes activation of B.

Now we summarize the key parameters that define the Neuroidal model:

- Number of neurons ($n$):
  - Specifies total neuron count in the network.
  - Values analyzed range from $10^5$ to $10^9$, consistent with biological observations
- Number of synapses per neuron ($d$):
  - Specifies expected number of synaptic connections from each neuron
  - Models network connectivity density
  - Values analyzed range from 16 to $10^6$
- Synaptic strength ($k$):
  - Relative to neuronal threshold required for neuron to fire
  - Specifies strength of each synapse as a fraction of threshold
  - Values analyzed range from 0.001 to 0.125 of the threshold
  - Consistent with typically observed weak synapse strengths
- Switching time:
  - The time it takes for a neuron to complete a cycle of causing an action potential in response to action potentials in its presynaptic neurons
  - Assumed to be 1 time unit for state changes, faster for threshold firings
- Threshold ($T$): Integer threshold value that summed synaptic inputs must exceed to cause neuron firing
- Neuron state ($s$): Each neuron has a state value indicating current mode (e.g. firing or non-firing)

- Synapse state ($q$):
  - State variable for each synapse
  - Used in algorithms to indicate intermediate synapse values
- Synapse weight ($w$):
  - Integer weight of each synapse, summed to determine neuron inputs
  - Initial weights set to $T/k$ based on synaptic strength parameter

We try our best to use the same notation in our results whenever possible however please note that there are some distinctions.

### 2.1.1 JOIN

We are primarily interested in JOIN as it the main memory formation tool of this model. Valiant discusses three variants of JOIN in his work - two-step disjoint JOIN, two-step shared JOIN, and one-step shared JOIN. The disjoint version is of little value to our analysis of capacity. Valiant did not make any claims of difference in biological plausibility between the one-step and two-step shared JOIN operations however we will focus on the one-step shared version as it easier to implement and computationally more efficient [32]. Further, it has already been succesfully simulated by Perrine in 2023 [27].

For the shared one-step case, Valiant suggests a slight change to the parameters - instead of a single synaptic strength parameter we now have two seperate parameters for memorization and association, $k_m$ and $k_a$ respectively that follow the relation $k_m = 2k_a$ [32]. The algorithm can be summarized as follows:

1. Fire memory A and B simultaneously

2. Propagate the synapse weights to all neuroids that have edges going from A or B to them

3. Collect the neuroids that have total number of edges greater than $k_m$

4. Group these neuroids together and call it memory C.

Note that since the synapse weight is $T/k_m$, by having greater than $k_m$ incoming edges from A or B, the total weight of the neuroid will be greater than $T$ causing it to fire.

This results in a set of approximately $r$ C neuroids that represent the conjunction of A and B - firing A or B will trigger at least half of these C neuroids to fire causing C to fire [32]. Subsequent JOIN operations treat C representations as first-class citizens like A and B.

Note that the original Neuroidal model is assumed to be bipartite, at least for the regions where the algorithms operate, an assumption we do not make in our analysis and instead work on a general random graph as we think that is more biologically plausible.

Valiant's original model assumes there are no edges between A and B since its a bipartite graph however we make no such assumption and it is possible that C could overlap with A or B. We think this still respects the intentions of JOIN as firing A or B will still fire C and it does not matter if it fires through edge weights or through overlap.

## 2.2 Assembly Calculus

Inspired by the Neuroidal model, Christos H. Papadimitriou, Santosh S. Vempala, Daniel Mitropolsky, and Wolfgang Maass proposed the Assembly Calculus, an updated model of neural computation, in their 2020 paper "Brain computation by assemblies of neurons" [25]. Assemblies were first introduced by Hebb way in his landmark 1949 paper "The first stage of perceptron: growth of the assembly" [11]. Assemblies are defined as large populations of neurons believed to imprint memories. The main difference between the memories in the Neuroidal model, later named arbsets by Valiant, and assemblies, are that arbsets are not better connected to each other than arbitrary sets unlike Hebbian assemblies [34].

The authors introduce a set of operations on assemblies. These operations correspond to properties of neuron assemblies observed in experiments, and can be shown, analytically and through simulations, to be realizable by generic, randomly connected populations of neurons with Hebbian plasticity and inhibition. Assemblies occupy a level of detail intermediate between the level of spiking neurons and synapses and that of the whole brain. They then argue that the resulting computational system can in principle be capable of carrying out arbitrary computations. The authors believe that something like it may underlie higher human cognitive functions like reasoning.

The main parameters of the model are:

- The number of neurons that are in a firing state in an area ($n$)
- The probability of recurrent (connections in a loop) and afferent (connections that carry sensory information towards the brain and spinal cord) synaptic connectivity ($p$)
- The maximum number of firing neurons in any area ($k$)

- Plasticity coefficient ($\beta$)

Typical values of these parameters used in the accompanying simulations with the 2020 paper are $n = 10^{6-7}, p = 10^{-3}, k = 10^{2-3}$, and $\beta = 0.1$ [25].

### 2.2.1 Project

Project is one of the two primary memory creation operations of the Assembly Calculus [25]. This operation creates a "copy" of an existing assembly $x$ in a new brain area $B$. It works by repeated firing of assembly $x$ while area $B$ is temporarily disinhibited. Through recurrent excitation and competition mediated by inhibition, a sequence of neuronal activity in area $B$ converges to form a stable new assembly $y$ that will fire whenever $x$ fires subsequently. This models experimentally observed phenomenon where assemblies activate associated assemblies in downstream areas. This process is described in great detail in the paper "Recurrent circuitry dynamically shapes the activation of piriform cortex" [8]. In the primary example used, the olfactory bulb is the upstream area and the piriform cortex is the downstream area. Projection underlies creation of new long-term memories.

### 2.2.2 Merge

Merge is the other main memory creation tool in the Assembly Calculus [25]. This operation combines two existing assemblies $x$ and $y$ into a single new assembly $z$ with strong reciprocal connections to $x$ and $y$. It is the most complex operation, requiring coordinated firing across 5 areas - those containing $x$, $y$, parents of $x$, parents of $y$, and the area A that will contain the new assembly $z$. Through repeated co-activation of $x$ and $y$ by their parents, recurrent excitation and inhibition between $A$ and the areas containing $x$ and $y$, the connectivities are adjusted and a new merged assembly $z$

forms in $A$. This operation is used to create merged neural representations supporting hierarchical combinations.

### 2.2.3 The NEMO model

Recently, Max Dabagia, Christos H Papadimitriou, and Santosh Vempala have introduced the NEMO model, an updated version of the Assembly calculus that claims to be even more biologically plausible by being able to create, process and manipulate temporal sequences of stimuli and memories [5]. The primary contribution of this model is the introduction of an updated version of Project called Sequence project, ab operation that can project a sequence of memories down.

## 2.3 Connection to this work

Despite the different nomenclature and goals of these models, they all share the same basic structure - a random graph representing the brain or brain area, subgraphs representing memories and nodes representing neurons. We believe that common definitions of interference and capacity and general theory will apply to all such models, although the exact formulation of expected interference between memories and capacity will vary between different memory creation algorithms.

# Chapter 3

# RELATED WORK

In this chapter, we go over works that have investigated the notion of capacity in random graph based models of cognition. To the best of our knowledge, there is a very limited amount of prior research carried out in this area.

One of the first attempts to estimate the capacity of the random graph based models of cognition was undertaken by Valiant himself [32]. Valiant concludes that it is complicated to derive the capacity of the model with shared memory representation, a sentiment we agree with. He accounts for interference by adding error rates for JOIN and LINK and the general noise rate $\sigma$, the total number of nodes active in the circuit at a given time, to his set of 6 equations governing JOIN and solves them assuming a reasonable bound on this value. Valiant also claimed that a single value for number of items that can be represented does not make sense for such a complicated model and it is more appropriate to simply bound the interference.

We assume a more optimistic stance regarding this and believe that it is possible to find an analytical formulation that will answer this question. We also think that Valiant agrees with us as he revisited this problem after a brief period of time, as discussed below, and made considerable progress in this area. While we are not able to solve the analytical capacity of the Neuroidal model with respect to shared JOIN in this paper, we believe we have laid significant groundwork for it and are optimistic it will be solved in the near future.

Valiant revisited the notion of capacity 4 years later, along with Vitaly Feldman, in their 2009 paper "Experience-Induced Neural Circuits That Achieve High Capacity"

[7]. He uses a relatively simple simulation of the Neuroidal model to analyze the capacity with a mix of tasks like memorization, association, inductive learning and hierarchical memory formation. He works with a rather loosely defined notion of interference in this paper that is however very similar to the idea introduced in 2005. The definition of capacity remains essentially the same as the point where there is too much interference or 'degradation'. The empirical results provided in this paper are quite interesting to us however the use of a bipartite Neouroidal model makes it slightly less biologically plausible. In our simulations we work with general sets and random graphs with no further assumption of structure.

Valiant again took on the challenge of capacity in his 2017 paper "Capacity of Neural Networks for Lifelong Learning of Composable Tasks", this time from a more theoretical standpoint [34]. In this paper, he analyzed the capacity of the Neuroidal model with respect to LINK. Valiant was able to successfuly derive theoretical estimates for the upper bound of the capacity. Valiant carries over the concept of interference from the 2005 paper and studies the evolution of the system until the associations created by LINK are no longer clearly defined and there is too much unintended excitation or firing of other memories, in other words, there is too much interference in the system.

This paper serves as the primary inspiration for our work and we follow very similar theoretical tools and style to arrive at our results. Of primary interest here, is the fact that Valiant analyzes capacity with regards to an operation that does not create new memories [34]. We assume that Valiant intended for the system to evolve by JOINing existing pairs of memories to form new memories as that is the primary memorization algorithm associated to the Neuroidal model. This raises the question of whether the capacity upper bound derived in this paper is actually realizable or will the system reach the point of excess interference before that. This is the core

reason for developing our theory with the memory formation algorithm as the central piece that determines the final formulation of interference and capacity.

In his 2023 paper, "Neural Tabula Rasa: Foundations for Realistic Memories and Learning", Perrine simulates the Neuroidal model and analyzes its capacity empirically with respect to the basic parameters of the model [27]. This paper inspired us to experimentally verify our theoretical results using an adapted version of the simulation code provided by Perrine. We also find the results in this paper intriguing, especially the behavior of the model where the capacity increases with a higher number of pre-existing starting memories in the model. This goes against our conventional wisdom and we try to explain it in section 5.3.1.

Also in 2023, Yi Xie, Yichen Li, and Akshay Rangamani explored the capacity of the Assembly Calculus with respect to the Project operation in their paper "Skip Connections Increase the Capacity of Associative Memories in Variable Binding Mechanisms" [35]. They work with a very similar interference driven definition of capacity without explicitly defining interference. The primary distinction with Neuroidal model based investigations into capacity is that they focus on interference between classes of memories rather than memories themselves. They define capacity as the number of classes when the within-class similarity is less than or equal to the between-class similarity. They were able to derive empirical results for this concept of capacity with respect to multiple basic parameters of the Assembly Calculus. Further, they also propose changes to some operations of the model to improve the capacity. We believe this is a very important work in this area and along with Perrine's paper signals the growing interest in the notion of capacity in models of cognition.

Chapter 4

# METHODS

In this chapter, we will discuss the methods we used to arrive at our results.

## 4.1 Theoretical methods

To derive the theoretical results, we used basic properties of real numbers, core principles from probability theory, and properties of common statistical distributions like normal, binomial and hypergeometric distributions. We start by focusing on a basic set with subsets being the main substructure of interest. As random graphs are a type of subsets, most of these general results will apply to the models discussed in Chapter 2 as long as the hypothesis is satisfied.

Before we can proceed further we need to formally define the notion of interference between subsets.

**Definition 1.** (***k*-Interference**) Given two sets $U, W$, and some number $k \in (0, |W|]$, we say $U$ *$k-$interferes* with $W$ if

$$|U \cap W| \geq \frac{|W|}{k}. \tag{4.1}$$

**Corollary 2.** *If $|U| = |W|$, then $U$ $k$-interferes with $W$ if and only if $W$ $k$-interferes with $U$.*

We restrict the upper range of $k$ to $|W|$ for convenience, as beyond that all values of $\frac{|W|}{k}$ will be less than 1.

This is a generalization of the notion of interference introduced by Valiant in 2005 [32]. Valiant defines a memory to be in a "firing" state if more than half the nodes in the memory are in a "firing" state. He then defines interference as the unintential firing of a memory $W$ when another memory $U$ is fired, which is possible if and only if more than half the nodes of $W$ are also present in $U$. This corresponds to the $k = 2$ case of our definition.

We now formally define the capacity of a system of overlapping subsets with interference being the limiting factor.

**Definition 3.** (**$(r, T, k, \delta)$-Subset Capacity**) Given a set $V = \{v_1, ..., v_n\}$, and parameters $r, T, k, \delta > 0$, the $(r, T, k, \delta)$-*subset capacity* of $V$ is the *maximum* number of subsets that can be picked subject to the conditions that for any randomly picked subset $U$,

1. $|U| \in [r - \delta, r + \delta]$,
2. $n >> 2(r + \delta)$,
3. $E[X_U] \leq T$ where $X_U$ is a random variable denoting the number of $k$-interferences caused due to picking $U$.

The first item simply accounts for the fact that the memories need not be exactly size $r$ however it should be bounded reasonably for us to make any claims regarding the capacity. We need the second restriction on the memories here since we want to apply lemma 5, that will be introduced later, to every pair. The third restriction here can be thought of as a stopping criteria as we stop picking the subsets once the expectation of interference reaches that threshold. In the context of models in computational neuroscience like the Neuroidal Model, this means that there will be

too much impact on the quality of memorization, that is too much noise and misfiring in the system if we add any further memories.

We believe these definitions of interference and capacity are rigorous and general enough to apply to a wide range of contemporary models of neural cognition.

### 4.2 Empirical methods

To validate our results empirically, we compare our theoretical results against the capacity of a simulated model. Perrine successfully implemented the Neuroidal model with shared one-step JOIN in 2023. The simulation uses the graph-tool python library which is a highly performant library written in C++ with wrappers for Python. Perrine also investigated the empirical capacity of the Neuroidal model under varying parameters. We build upon Perrine's codebase and make some key changes to achieve our goals.

First, we add the capability to the model to easily swap between memory creation algorithms. For our experiments, we use both the basic random memory creation and one-step shared JOIN algorithms. Future investigations into the capacity of random graph based models could be facilitated by adding support for more memory creation algorithms.

Second, we add the capability to the code to be able to generate visualizations of the model at regular intervals, this will allow us to visualize the model and develop deeper intuition into its behavior. This is especially critical for understanding JOIN and it's interference characteristics as these properties are not immediately obvious from the definition of the operation.

# Chapter 5

# RESULTS

In this chapter, we present the primary contributions of this work and provide discussion justifying their applicability to contemporary neuroscience.

## 5.1 Theoretical results

First we will go over the theoretical results of our analysis.

### 5.1.1 Interference

We are now interested in finding the probability of a randomly picked subset interfering with another randomly picked subset. We start with the case where they are randomly picked as we believe it is the simplest case. We will touch upon other possible cases in the Discussion section below when discussing models in Computational Neuroscience that use unique memory generation algorithms.

**Lemma 4.** *Given a set $V$ with $n$ items and two subsets $U, W$ of respective sizes $r_u, r_w$, denote the size of the intersection between them by the random variable $Y_{u,w}$. Then the probability of $U$ $k$-interfering with $W$ is*

$$\sum_{y=\lceil \frac{r_w}{k} \rceil}^{r_w} \frac{\binom{r_u}{y}\binom{n-r_u}{r_w-y}}{\binom{n}{r_w}}$$

*and $Y_{u,w} \sim Hypergeometric(n, r_u, r_w)$.*

*Proof.* If $V = \{v_1, ..., v_n\}$, we can represent the first randomly picked subset $U$ as a boolean vector $u$ of length $n$ defined by

$$u_i = \begin{cases} 1 & \text{if } v_i \in U \\ 0 & \text{if } v_i \notin U. \end{cases}$$

With this representation, $U$ will intersect another randomly picked subset $W$ at the indices where both boolean vectors $u, w$ have a 1. Then $Y_{u,w}$ is equivalent to the number of indices where both $u, w$ have a 1. First note that

$$\mathbb{P}(Y_{u,w} = y) = \frac{\binom{r_u}{y}\binom{n-r_u}{r_w-y}}{\binom{n}{r_w}}. \tag{5.1}$$

This follows from the fact that given the first vector $U$, we already know where the 1's are located. We can pick the $y$ intersecting 1's for the second vector in $\binom{r_u}{y}$ ways implicitly placing 0's in the remaining spots. We then fill the remaining $n - r_u$ indices corresponding to the 0's in the first vector with $r_w - y$ 1's in $\binom{n-r_u}{r_w-y}$ ways. Finally we divide by the total number of possible subsets $\binom{n}{r_w}$. Clearly, this is the probability mass function of the hypergeometric distribution with population size $n$, $r_u$ success states and $r_w$ draws. We conclude that $Y_{u,w} \sim Hypergeometric(n, r_u, r_w)$. Finally, to find the probability of $U$ $k$-interfering with $W$ we need to find $\mathbb{P}(Y_{u,w} \geq \lceil \frac{r_w}{k} \rceil)$ which is the sum of $\mathbb{P}(Y_{u,w} = y)$ from $y = \lceil \frac{r_w}{k} \rceil$ to $y = r_w$. □

For brevity, we can reinterpret the above probability as the tail distribution function of $Y_{u,w}$ at $\lfloor \frac{r_w}{k} \rfloor$,

$$\mathbb{P}\left(Y_{u,w} \geq \left\lceil \frac{r_w}{k} \right\rceil\right) = \mathbb{P}\left(Y_{u,w} > \left\lfloor \frac{r_w}{k} \right\rfloor\right) = \bar{F}_{Y_{u,w}}\left(\left\lfloor \frac{r_w}{k} \right\rfloor\right)$$

Recall from statistics that the expectation of a binary payoff, like intersection, that depends on a cutoff (in this case $\lceil \frac{r_w}{k} \rceil$) is equal to the probability of the variable being greater than or equal to the cutoff. Therefore the probability in lemma 4 is equal to the expected number of interferences of $U$ with $W$.

We then want to estimate the expected number of interferences when the sizes of the subsets are within a certain offset of $r$, say $\delta$ without being exactly equal to $r$. This approach will make our results more applicable to models like the Neuroidal Model that assume memory sizes follow some distribution [32]. The offset can be selected to best suit the distribution involved. For example if the sizes come from a discrete distribution like $\mathcal{B}(r/p, p)$, and if the variance $r(1-p)$ is more than 10, it makes sense to choose $\delta = 2\sqrt{r(1-p)}$ since roughly 95% of all values lie within $[r - 2\sigma, r + 2\sigma]$.

Generalizing this without any further assumptions is quite hard as the binomial coefficients do not vary nicely as a function of two variables over their domain. Instead we will make a reasonable assumption that will allow us to derive a reasonable lower bound for this expectation in terms of a general parameter instead of individual subset sizes.

**Lemma 5.** *Given a set $V$ with $n$ items and two subsets $U, W$ of respective sizes $r_u, r_w$, denote the size of the intersection between them by the random variable $Y_{u,w}$. If*

1. *$r_u, r_w \in [r - \delta, r + \delta]$ for some $r, \delta > 0$,*
2. *$n >> 2(r + \delta)$,*

*then*

$$\bar{F}_{Y_{u,w}}\left(\left\lfloor \frac{r_w}{k} \right\rfloor\right) \geq \sum_{y=\lceil \frac{r+\delta}{k} \rceil}^{\lfloor r-\delta \rfloor} \frac{\binom{r-\delta}{y}\binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}}$$

*Remark.* Before proceeding with the proof, we want to justify the second assumption made here. It is a known fact that bounding binomial coefficients above or below is hard due to the nature of how it varies with respect to the second argument. We know that $\binom{n}{k}$ reaches its maximum value at $\lceil \frac{n}{2} \rceil$ or $\lfloor \frac{n}{2} \rfloor$ and it is monotonically increasing at smaller values and decreasing at larger values. My making the assumption here we can ensure that our second argument is always a lot smaller than this maxima, and as such an increase in the second argument will only increase the value of the expression. This assumption is reasonable since models like the Neuroidal Model expect the memory sizes to be significantly smaller than the size of the model [32]. Also note that the binomial coefficient increases monotonically with respect to the first argument.

*Proof.* First note that $n > r_u, r_w$ and by extension $n > r$ since the size of a subset cannot exceed the size of the set. Then observe that

$$\begin{aligned}
\bar{F}_{Y_{u,w}}\left(\left\lfloor \frac{r_w}{k} \right\rfloor\right) &= \sum_{y=\lceil \frac{r_w}{k} \rceil}^{r_w} \mathbb{P}(Y_{u,w} = y) \\
&= \sum_{y=\lceil \frac{r_w}{k} \rceil}^{r_w} \frac{\binom{r_u}{y}\binom{n-r_u}{r_w-y}}{\binom{n}{r_w}} \\
&\geq \sum_{y=\lceil \frac{r_w}{k} \rceil}^{r_w} \frac{\binom{r-\delta}{y}\binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}} \\
&\geq \sum_{y=\lceil \frac{r+\delta}{k} \rceil}^{\lfloor r-\delta \rfloor} \frac{\binom{r-\delta}{y}\binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}}
\end{aligned} \tag{5.2}$$

The first and second equalities follow from the definition of the tail distribution and lemma 4 respectively. The third inequality follows from assumption 1. in the theorem

and the behavior of the binomial coefficient under varying arguments. The final inequality follows from the fact that since all terms in the sum are positive, reducing the number of terms will make the overall expression smaller.

□

### 5.1.2 Capacity

With the above lemmas in our arsenal we can now move on the main subject of this thesis.

Before deriving the capacity for the general case, let us consider the simpler case where all memories have the exact same size. This is valuable since it results in a much simpler expression and we can use this as an approximation for the more general case too. However note that we realize this scenario is not biologically plausible at all.

**Theorem 6.** *Given a set $V$ with $n$ items and the property that every picked subset will have size exactly $r$, the $(r, T, k, \delta)$-subset capacity of $V$ is*

$$\left\lfloor \frac{T}{\bar{F}_{Y_{u,w}}\left(\left\lfloor \frac{r}{k} \right\rfloor\right)} + 1 \right\rfloor.$$

*Remark.* Since all subsets have fixed size $r$, note that the choice of $\delta$ is not relevant here.

*Proof 1.* Suppose we have already have $M-1$ subsets in the universe. Pick a random subset $U$. From lemma 4, we know that the expected number of $k$-interferences of $U$ with another arbitrary subset $W$ from the universe is $\bar{F}_{Y_{u,w}}\left(\left\lfloor \frac{r}{k} \right\rfloor\right)$. Since there are

$M-1$ other subsets, the total expected number of $k$-interferences caused by picking $U$ is $(M-1)\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r}{k}\right\rfloor\right)$.

From inequality 3 in the definition of capacity, we have

$$(M-1)\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r}{k}\right\rfloor\right) \leq T \implies M \leq \frac{T}{\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r}{k}\right\rfloor\right)} + 1. \tag{5.3}$$

The $(r, T, k, \delta)$-subset capacity of $V$ then is the largest integer $M$ that satisfies inequality 5.3. □

We provide an alternate proof that, while less elegant, can be scaled to prove the general statement.

*Proof 2.* Suppose we have already have $M$ subsets in the universe. Pick two subsets $U, W$ without replacement. From lemma 4, we know that the expected number of $k$-interferences of $U$ with $W$ is $\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r}{k}\right\rfloor\right)$. Since we know all subsets have the same size, the expected number of $k$-interferences of $W$ with $U$ is the same. So the expected number of interferences caused by one pair is

$$2\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r}{k}\right\rfloor\right).$$

We know that there are $\binom{M}{2} = M(M-1)/2$ such pairs so the expected number of total interferences is

$$2 \cdot \frac{M(M-1)}{2}\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r}{k}\right\rfloor\right) = M(M-1)\bar{F}_{Y}\left(\left\lfloor\frac{r}{k}\right\rfloor\right).$$

Since there are $M$ subsets, the expected number of interferences by choosing picking one subset is

$$\frac{M(M-1)}{M}\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r}{k}\right\rfloor\right) = (M-1)\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r}{k}\right\rfloor\right).$$

From inequality 3, we have

$$(M-1)\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r}{k}\right\rfloor\right) \leq T \implies M \leq \frac{T}{\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r}{k}\right\rfloor\right)} + 1. \tag{5.4}$$

The $(r, T, k, \delta)$-subset capacity of $V$ is the largest integer $M$ that satisfies inequality 5.4. □

We will now tackle the general case using the same strategy as above.

**Theorem 7.** *Given a set $V$ with $n$ items, the $(r, T, k, \delta)$-subset capacity of $V$ is bounded above by*

$$\frac{T}{\sum_{y=\lceil\frac{r+\delta}{k}\rceil}^{\lfloor r-\delta\rfloor} \frac{\binom{r-\delta}{y}\binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}}} + 1$$

*Remark.* Note that we can only say it is bounded above and not the exact capacity as defined since we have to use lemma 5. However as $\delta \to 0$, this expression converges to the expression in theorem 6.

*Proof.* Suppose we have $M$ subsets $U_1, ..., U_M$ with sizes $r_1, ..., r_M$. Pick two subsets $U_i, U_j$. From lemma 4, we know that the expected number of interferences caused by this pair is

$$\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r_j}{k}\right\rfloor\right) + \bar{F}_{Y_{w,u}}\left(\left\lfloor\frac{r_i}{k}\right\rfloor\right).$$

We then sum over all possible pairings to get the expected number of total interferences:

$$\sum_{(i,j)\in\mathbb{Z}\times\mathbb{Z}, 1\leq i,j\leq M, i\neq j} \left(\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r_j}{k}\right\rfloor\right) + \bar{F}_{Y_{w,u}}\left(\left\lfloor\frac{r_i}{k}\right\rfloor\right)\right).$$

Since there are $M$ subsets, the expected number of interferences by picking one subset is

$$\frac{1}{M}\sum_{(i,j)\in\mathbb{Z}\times\mathbb{Z},1\leq i,j\leq M,i\neq j}\left(\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r_j}{k}\right\rfloor\right)+\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r_i}{k}\right\rfloor\right)\right).$$

From inequality 3, we have

$$\frac{1}{M}\sum_{(i,j)\in\mathbb{Z}\times\mathbb{Z},1\leq i,j\leq M,i\neq j}\left(\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r_j}{k}\right\rfloor\right)+\bar{F}_{Y_{w,u}}\left(\left\lfloor\frac{r_i}{k}\right\rfloor\right)\right)\leq T,$$

which implies

$$M\geq\frac{1}{T}\sum_{(i,j)\in\mathbb{Z}\times\mathbb{Z},1\leq i,j\leq M,i\neq j}\left(\bar{F}_{Y_{u,w}}\left(\left\lfloor\frac{r_j}{k}\right\rfloor\right)+\bar{F}_{Y_{w,u}}\left(\left\lfloor\frac{r_i}{k}\right\rfloor\right)\right). \tag{5.5}$$

Using lemma 5 we get

$$\begin{aligned}M&\geq\frac{1}{T}\sum_{(i,j)\in\mathbb{Z}\times\mathbb{Z},1\leq i,j\leq M,i\neq j}\left(2\sum_{y=\left\lceil\frac{r+\delta}{k}\right\rceil}^{\lfloor r-\delta\rfloor}\frac{\binom{r-\delta}{y}\binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}}\right)\\&=\frac{1}{T}\frac{M(M-1)}{2}\left(2\sum_{y=\left\lceil\frac{r+\delta}{k}\right\rceil}^{\lfloor r-\delta\rfloor}\frac{\binom{r-\delta}{y}\binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}}\right),\end{aligned} \tag{5.6}$$

which implies

$$M\leq\frac{T}{\sum_{y=\left\lceil\frac{r+\delta}{k}\right\rceil}^{\lfloor r-\delta\rfloor}\frac{\binom{r-\delta}{y}\binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}}}+1. \tag{5.7}$$

The expected $(r,T,k,\delta)$-subset capacity of $V$ should be bounded above by this expression and the tightness of the bound will depend on the parameter $\delta$. □

## 5.2 Empirical results

### 5.2.1 Fixed subset size

First we simulate the case for fixed subset size $r$.

We compare the average capacity of the simulation with the analytical result from Theorem 6 as a function of the size of the set. We fix $r = 20$, $k = 2$, $T = 0.1$. Figure 5.1 shows the results of this comparison. We see that the average simulated capacity is practically identical to the analytical capacity thoughout our input range.

Then we the compare the average capacity of the simulation with the analytical result from Theorem 6 as a function of the size of the subsets $r$. We fix $n = 100$, $k = 2$, $T = 0.1$. Figure 5.2 shows the results of this comparison. We see that the average simulated capacity is very close to the analytical capacity thoughout our input range and follows the general trend, even following the sharp decreases while going from odd numbers to even numbers. This is because the sets need intersections of size atleast $\lceil r/2 \rceil$ to interfere and as the size of the subset goes from an odd number to the next even number, this value remains the same while the size of the subsets increase leading to a higher probability of interference and lower capacity. One can also think of it as more terms being included in the sum in Lemma 4. We believe these peaks will reduce in intensity relative to the scale of the y axis as $n \to \infty$. Figure 5.3 shows the values of analytical capacity with same configuration as above but with n set to 500. We can already see that the graph has become a lot smoother. Unfortunately, it is impossible for us to simulate models of this size or bigger due to memory constraints.

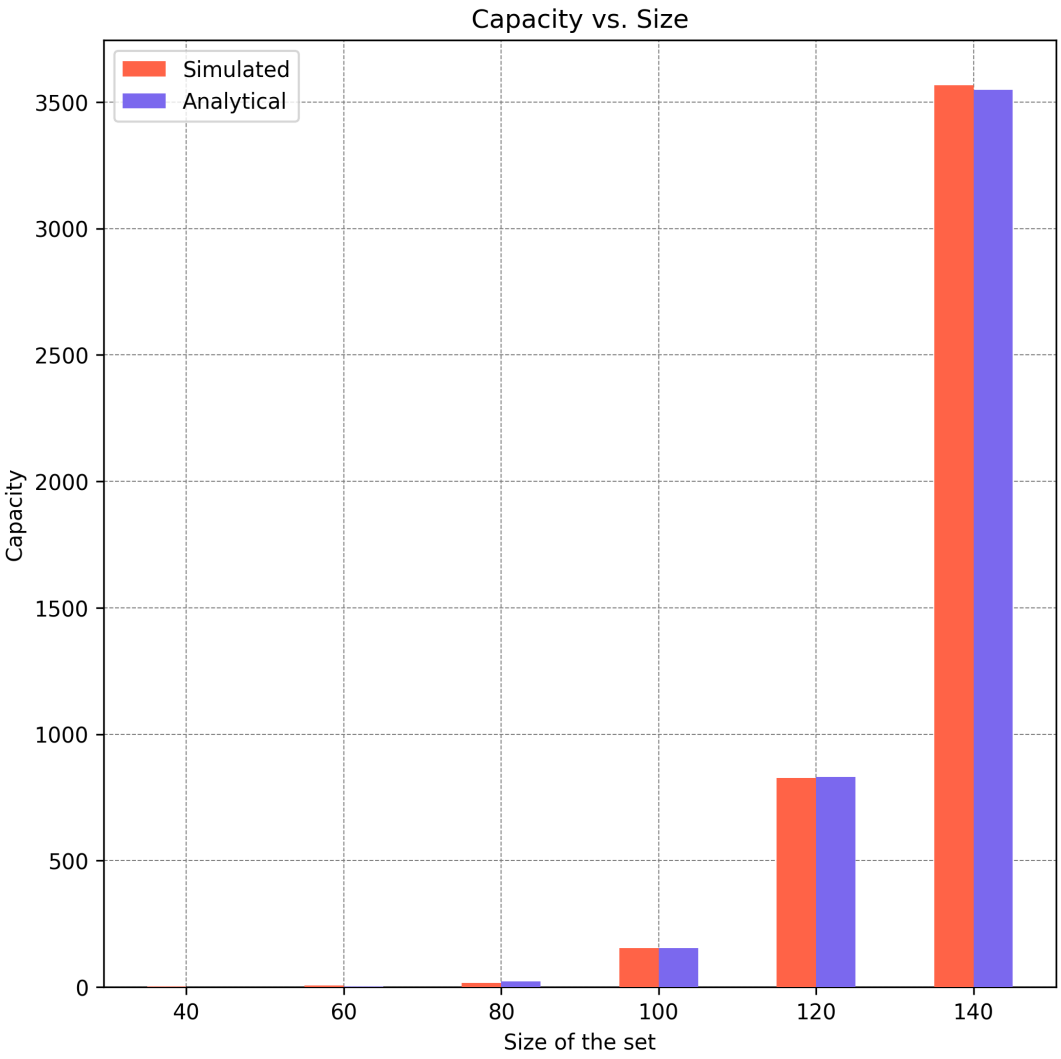


**Figure 5.1: Capacity vs. Size of the set ($n$).** How capacity is affected by increasing the size of the set ($n$). This figure compares the expression for fixed subset size against the simulation with fixed subset size.

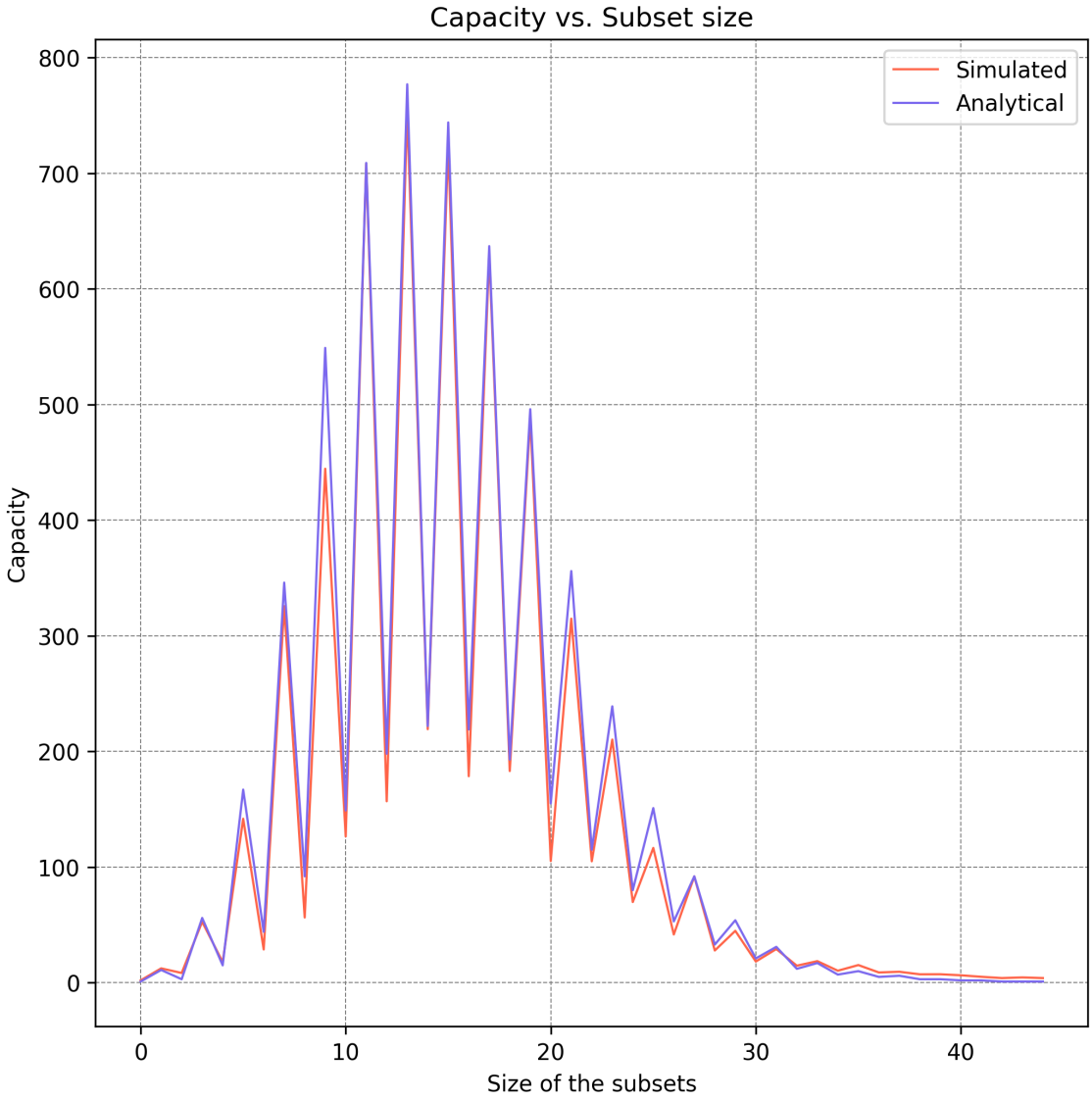


**Figure 5.2: Capacity vs. Size of the subsets ($r$).** How capacity is affected by increasing the size of the subsets ($r$). This figure compares the expression for fixed subset size against the simulation with fixed subset size.

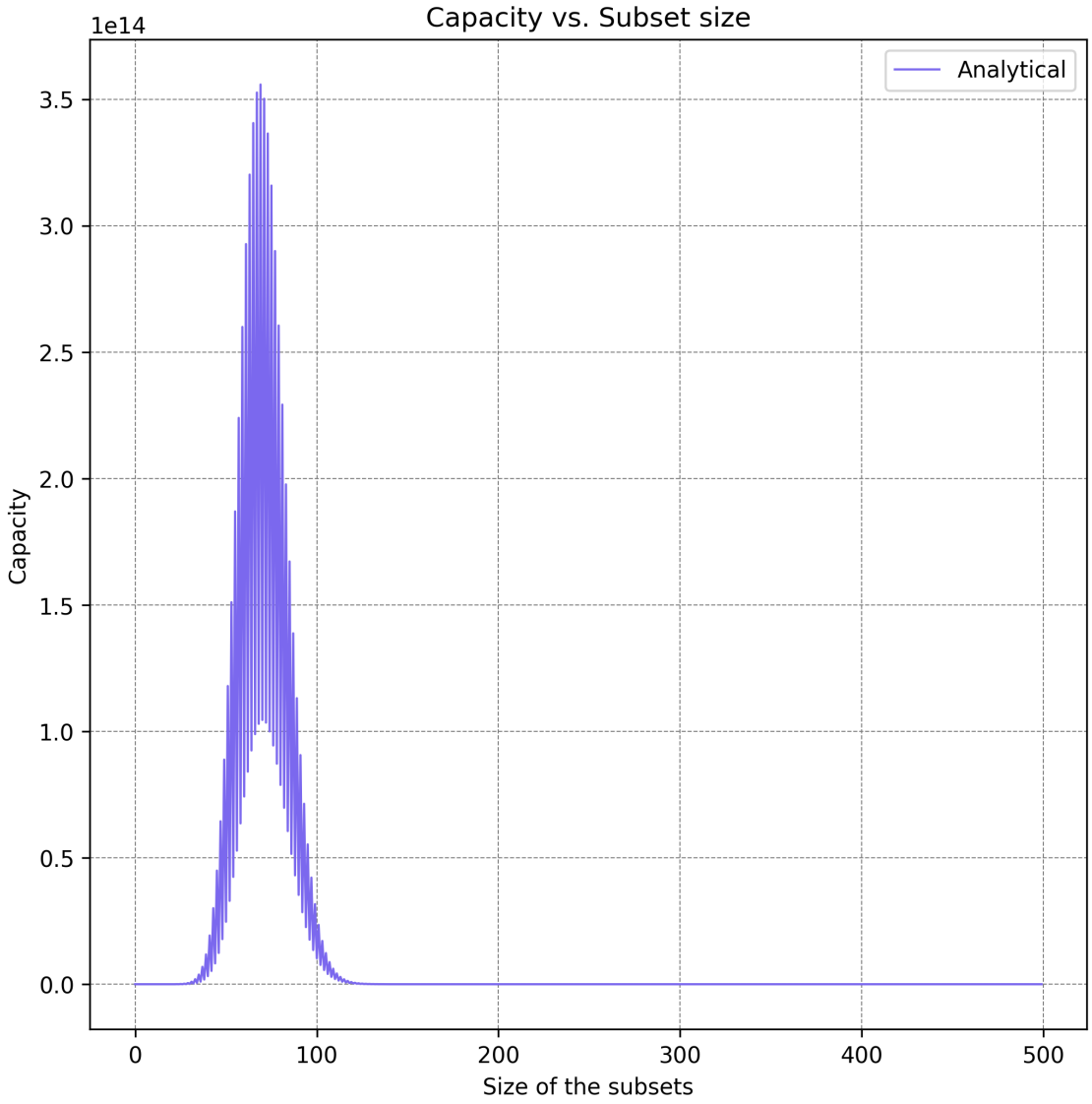


**Figure 5.3: Capacity vs. Size of the subsets ($r$) for $n = 500$.** How capacity is affected by increasing the size of the subsets ($r$) when $n = 500$.

### 5.2.2 Bounded subset size

Now we simulate the case where the subset sizes are not fixed but rather bounded above and below.

Like in the fixed case, we compare the simulated and analytical capacities against the size of the set and against the size of the subsets. For the comarison against size of the set, we set $r = 20$ and for the comparison against size of the subsets, we set $n = 100$. For both cases we fix $k = 2, T = 0.1$ and draw the $r$ values randomly from $\mathcal{N}(r, 1)$ followed by conversion to integer. Based on the Empirical Law, we except 95% of the values to lie within two standard deviations of $r$, so we choose $\delta = 2$. Figures 5.4 and 5.5 show the results of these experiments. We see that even thought the analytical bound from equation 3 bounds the simulated capacity, the bound is very loose and does cannot be used as an approximation. We believe this is because of how the binomial coefficient varies with respect to its second parameter and that $\delta$ here, even though only 2 is quite big with respect to $r$, and is reducing the second argument significantly in the term $\binom{n-r-\delta}{r-\delta-y}$. We believe that a larger $n$ and $r$ will make this bound tighter and applications like the Neuroidal model indeed use values of $n$ and $r$ that are orders of magnitudes larger. However, it is not feasible for us to simulate at such scale. We also compared the analytical results from equation 5.3 and it was still a good approximation for the simulation with randomly drawn r's. Figures 5.6 and 5.7 show the results of these simulations. As expected the simulated capacity is a lot closer to the analytical capacity when the standard deviation is low.

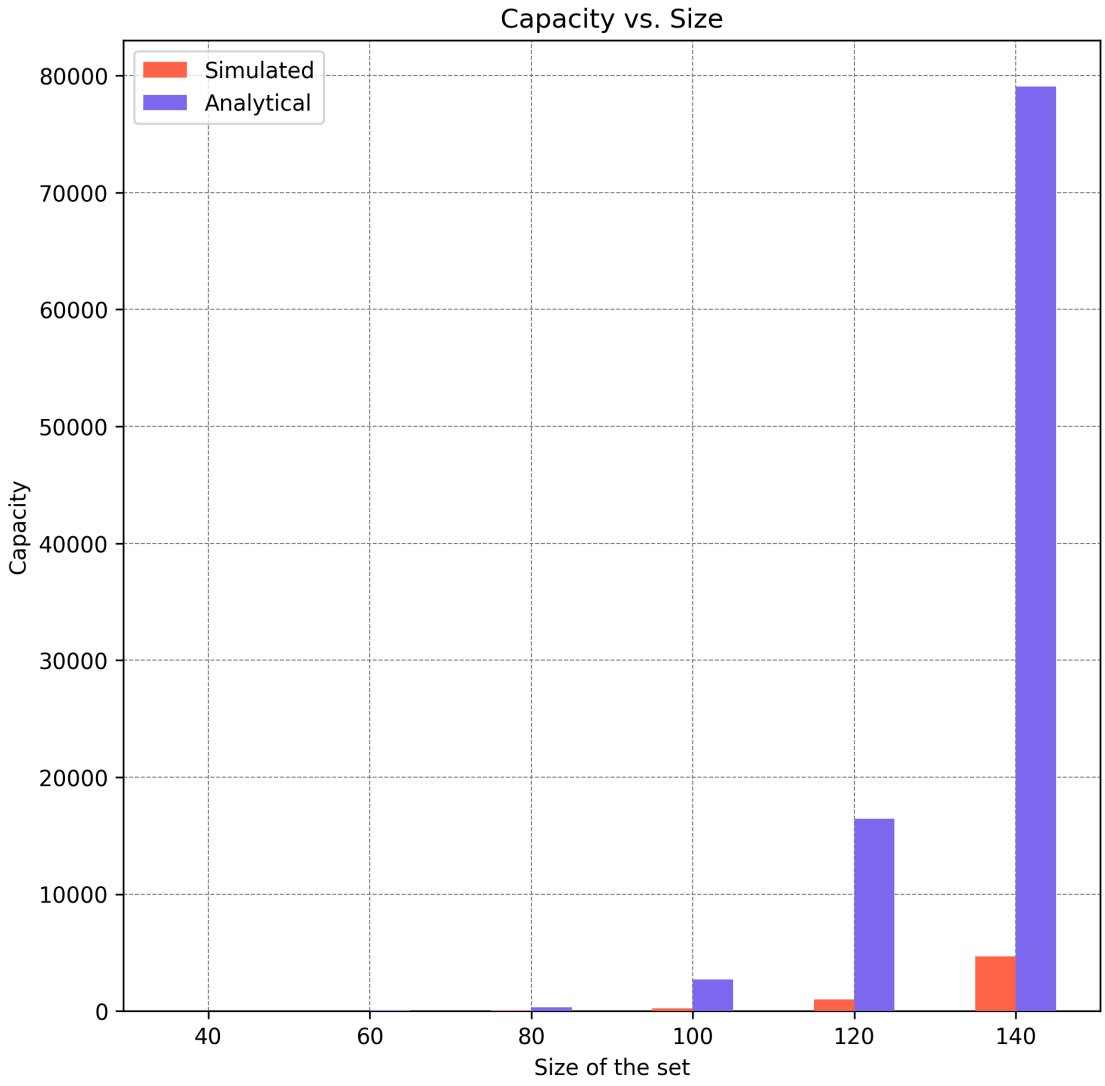


**Figure 5.4: Capacity vs. Size of the set ($n$) when $r \sim \mathcal{N}(r, 1)$.** How capacity is affected by increasing the size of the subsets ($r$) when $r$ is drawn from a normal distribution with mean $r$ and standard deviation 1. This figure compares the expression for bounded subset size against the simulation that draws memories from a distribution.

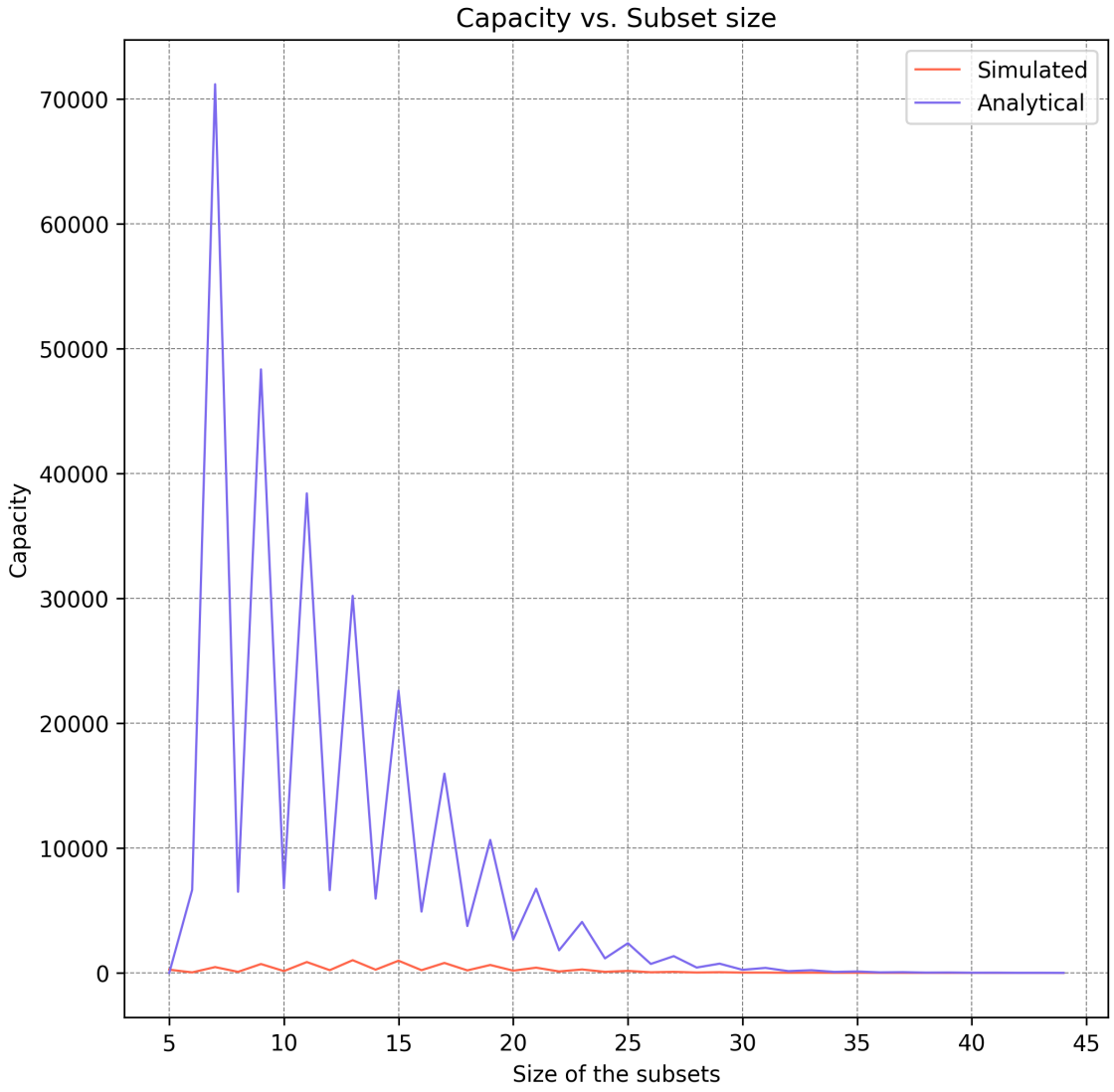


**Figure 5.5: Capacity vs. Size of the subsets ($r$) when $r \sim \mathcal{N}(r, 1)$.** How capacity is affected by increasing the size of the subsets ($r$) when $r$ is drawn from a normal distribution with mean $r$ and standard deviation 1. This figure compares the expression for bounded subset size against the simulation that draws memories from a distribution.

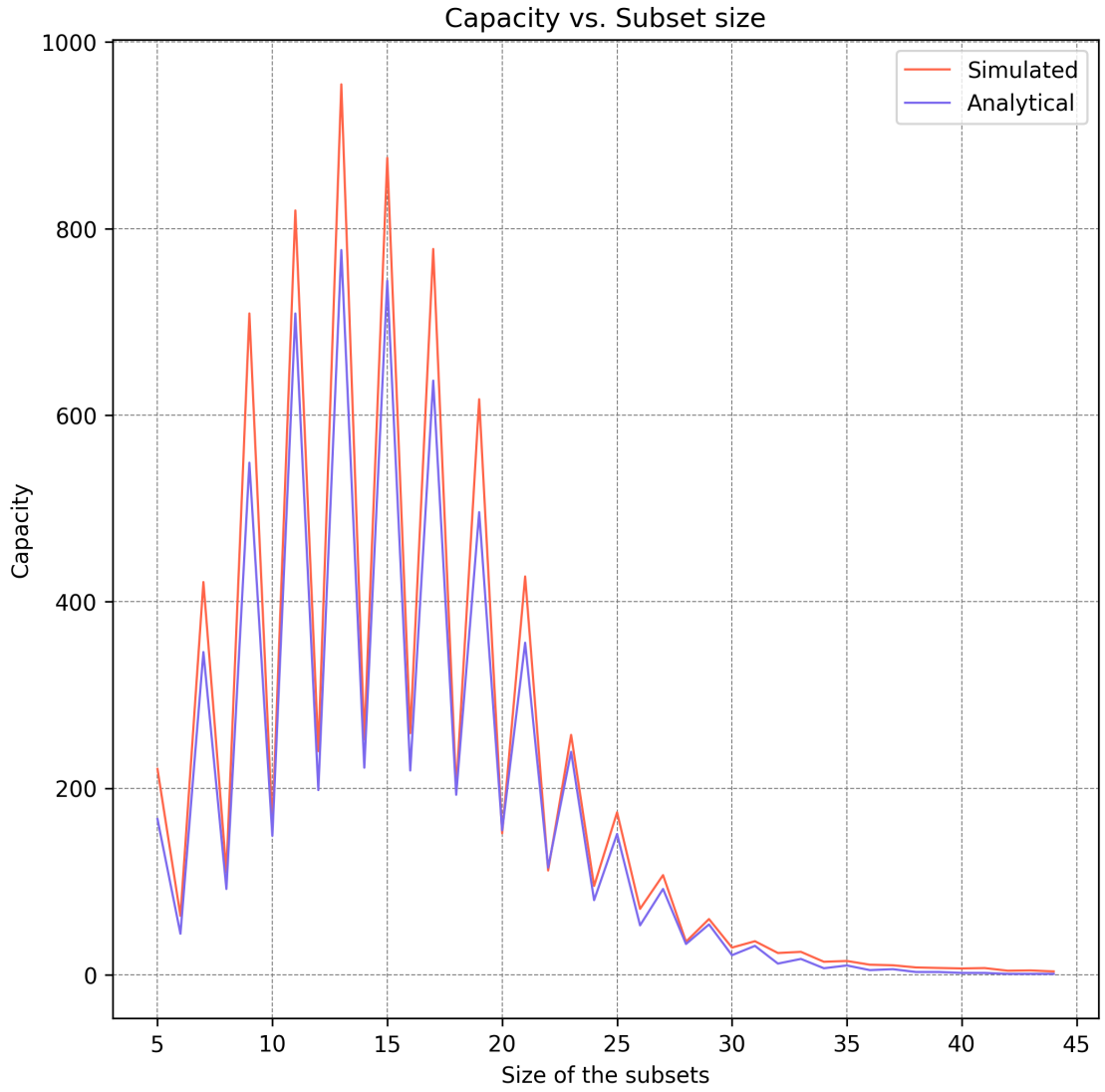


**Figure 5.6: Capacity vs. Size of the subsets ($r$) when $r \sim \mathcal{N}(r, 1)$ comparing exact formula vs simulation.** How capacity is affected by increasing the size of the subsets ($r$) comparing the results of 5.3 with the simulation where $r$ is drawn from a normal distribution with mean $r$ and standard deviation 1. This figure compares the expression for fixed subset size against the simulation that draws memories from a distribution.

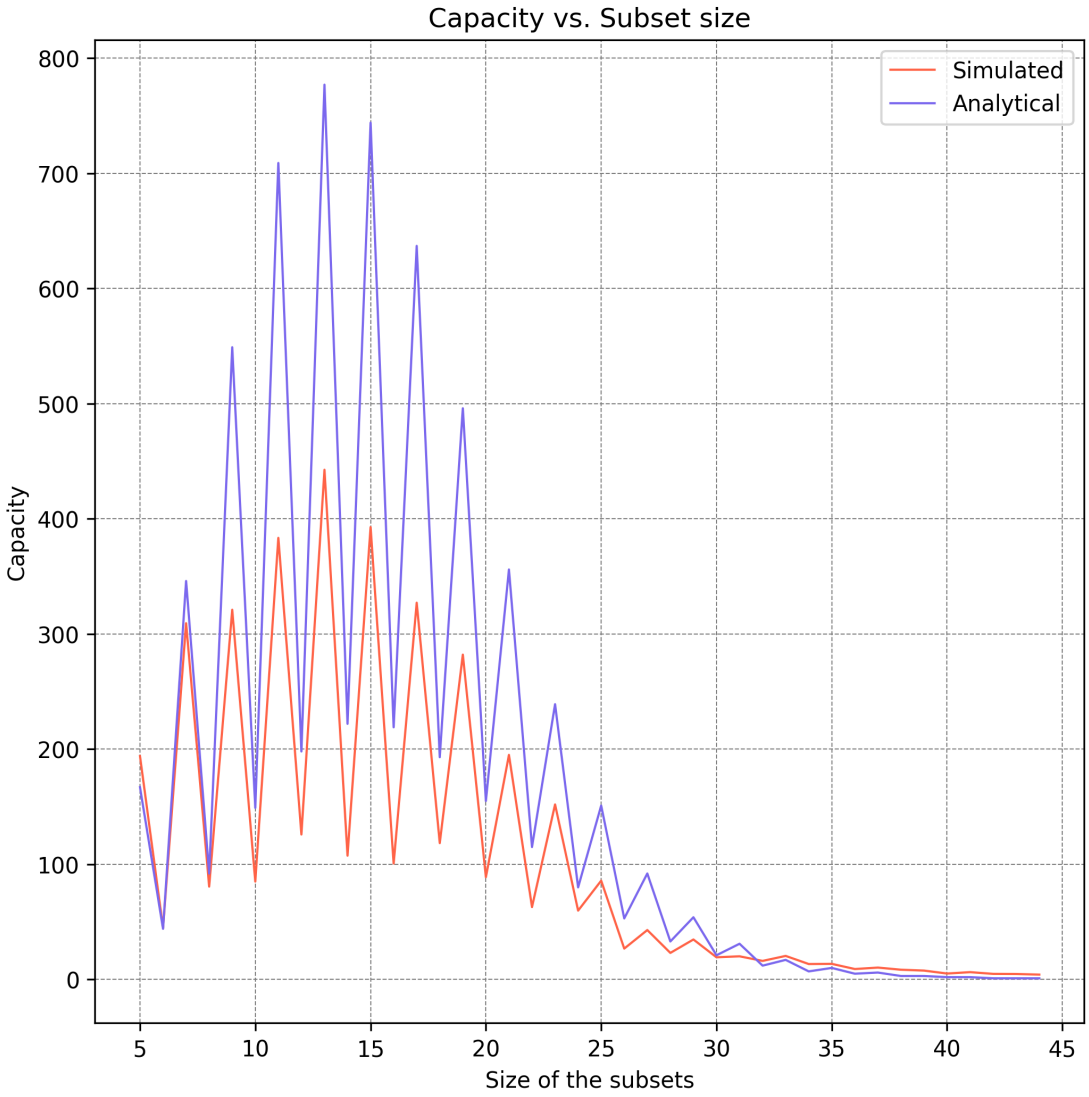


**Figure 5.7: Capacity vs. Size of the subsets ($r$) when $r \sim \mathcal{N}(r, 2)$.** How capacity is affected by increasing the size of the subsets ($r$) comparing the results of 5.3 with the simulation where $r$ is drawn from a normal distribution with mean $r$ and standard deviation 2. This figure compares the expression for fixed subset size against the simulation that draws memories from a distribution.

### 5.3 Discussion

We believe our theoretical framework for interference and capacity calculations can be extended to analyze and understand the behavior of graph-based models with more complex memory creation algorithms. The primary distinction between our simple model and more complicated models like the Neuroidal model and Assembly Caculus is with regards to the memory creation algorithm. We implicitly assume a random memory creation algorithm in lemma 4 while the Neuroidal model uses the JOIN operation and the Assembly Calculus uses the Project and Merge operations for memory creation [25, 32]. Since no other part of our theory makes any assumption about the process of memory formation, we believe that adjusting the calculation of expected interference between two memories in lemma 4 to incorporate the nuances of other memory creation algorithms and using it appropriately in Theorem 6 or 7 will give us an accurate representation of capacity in those models.

We maintain the generality of our interference calculation while allowing for variations in memory creation algorithms. Specifically, we can refine the lemma to account for the JOIN, Project and Merge operations, ensuring that our interference metric aligns with the unique characteristics of each model. This adaptability enables the application of our interference and capacity framework to a broader class of graph-based models in computational neuroscience.

The flexibility of our theoretical approach allows researchers to tailor interference calculations based on the specifics of memory creation algorithms in diverse neural network models. As a result, our capacity analysis can provide valuable insights into the limitations and efficiency of these models, enhancing our understanding of their memory storage capabilities.

### 5.3.1 Capacity and JOIN

In this section, we briefly discuss some insights we gained from our analysis of the JOIN operation with regards to capacity.

The JOIN algorithm is unique in the sense that it must follow a set of 6 basic constraints which ensure that the newly formed memories are roughly the same size as the memories that were JOINed and therefore the size of the memories ($r$) remains more or less fixed for a given graph size ($n$), expected degree ($d$) and edge weight ($w$). Note that edge degree and synaptic weight were not a factor in our theory as edges do not matter when memories are randomly inserted. However, we believe that an updated JOIN-compliant formula expected interference between two memories will involve these paramters.

First, we compare our results from equation 3 with results from our simulation of the Neuroidal Model. Valiant calculated the memory sizes for various parameter combinations, however the minimum graph size he considered was 100,000. Our simulation of the Neuroidal model cannot be scaled up that far so we use $n = 500, d = 128, w = 16$. We used patterns found in his results to estimate that $r = 40$ would be the appropriate value for this configuration. We start with 100 randomly generated memories which have negligible interference between them, essentially agreeing with our results from Theorem 7. We observed that newly formed memories had sizes around 40, validating that we are compliant with Valiant's definition of JOIN. Finally, we set the interference threshold $T = 0.1$ and interference parameter $k = 2$ as that is the only value of $k$ supported by the Neuroidal model.

Our equation produces a capacity of 3203070686178 while the Neuroidal model simulation only reaches around 250 memories over 20 runs before stopping due to excess interference. Therefore our equation in its current form cannot serve as a good approx-

imation for the Neuroidal model's capacity however it is effectively a very loose upper bound on most models like these as the memory formation is not capacity optimized which we think makes sense biologically. We believe random memory generation is quite good for capacity and memory formation algorithms that are even further optimized for capacity will be even more biologically implausible. As described in section 5.2.3, the theory will need to be updated to account for JOIN's behavior.

Finally, we made some attempts to gain some initial understanding of why memories formed by JOIN tend to interfere a lot more than randomly generated memories. We visualize the graphs and analyze the nodes with respect to two metrics: number of memories the nodes belong to and the number of interferences that have occured at those nodes. We choose the same parameters as before except we choose 1000 randomly inserted starting memories instead of 100. This will lead to overall higher capacity and allow us to analyze in more depth. Perrine has demonstrated that the capacity of the Neuroidal model scales up with the amount of starting memories [27]. We are able to verify this as our simulation ended up producing around 1800 memories, in particular, an additional 800 JOINed memories as compared to the 150 with 100 starting memories. A possible explanation for this intriguing behavior would be that since these starting memories are randomly generated and follow our results that predict significantly low interference among them, they are making it harder for the system to reach the threshold in constraint 3 in definition 3, since the simulation calculates the intereference rate using the number of total misfires at the state of the system over the number of total memories in the system.

Figures 5.8 through 5.11 and 5.12 through 5.15 show the progression of the system from 1200 memories till the system reaches capacity with the value inside the node indicating the number of interferences that occured at that node and the number of memories the node belongs to respectively. The relative size of the node indicates the

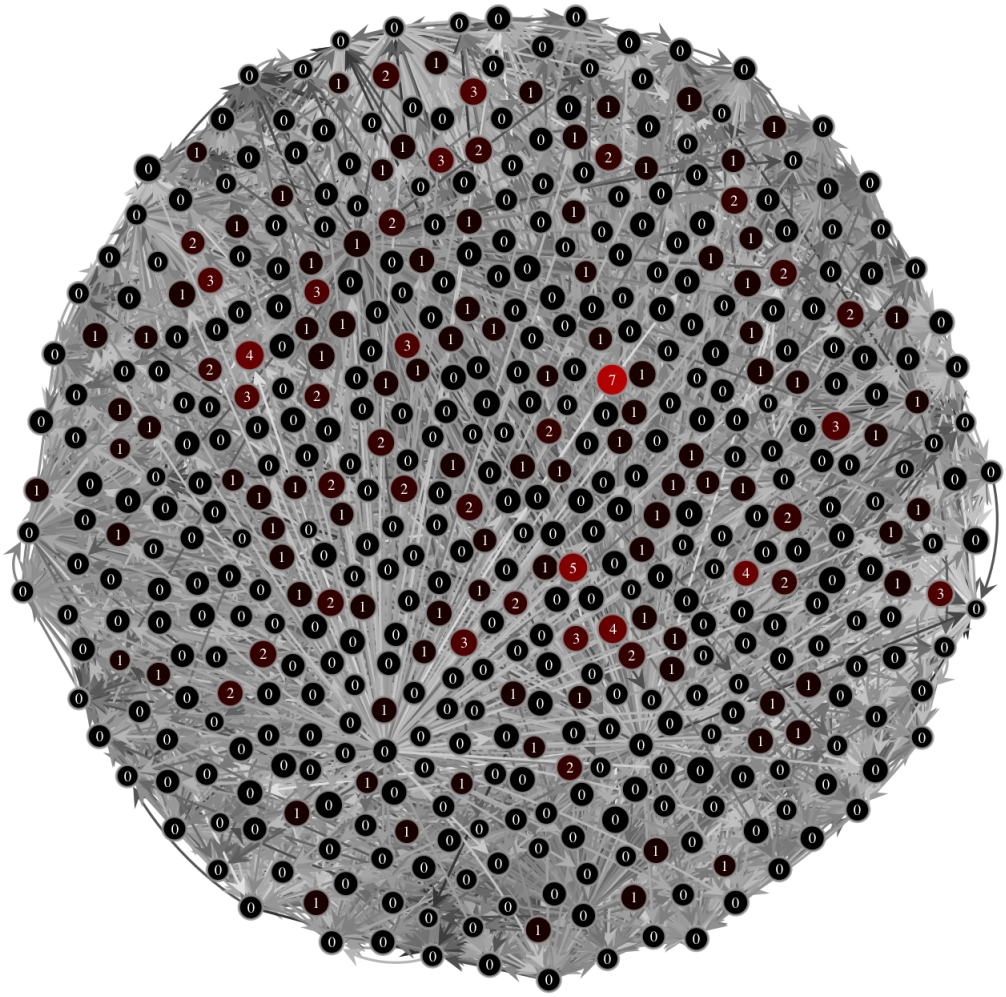

**Figure 5.8: Interference accumulation per node of Neuroidal model at 1200 memories.** Neuroidal model (n=500, d=128, w=16, r=40) at 1200 memories. Node value indicates number of interferences that occured at each node.

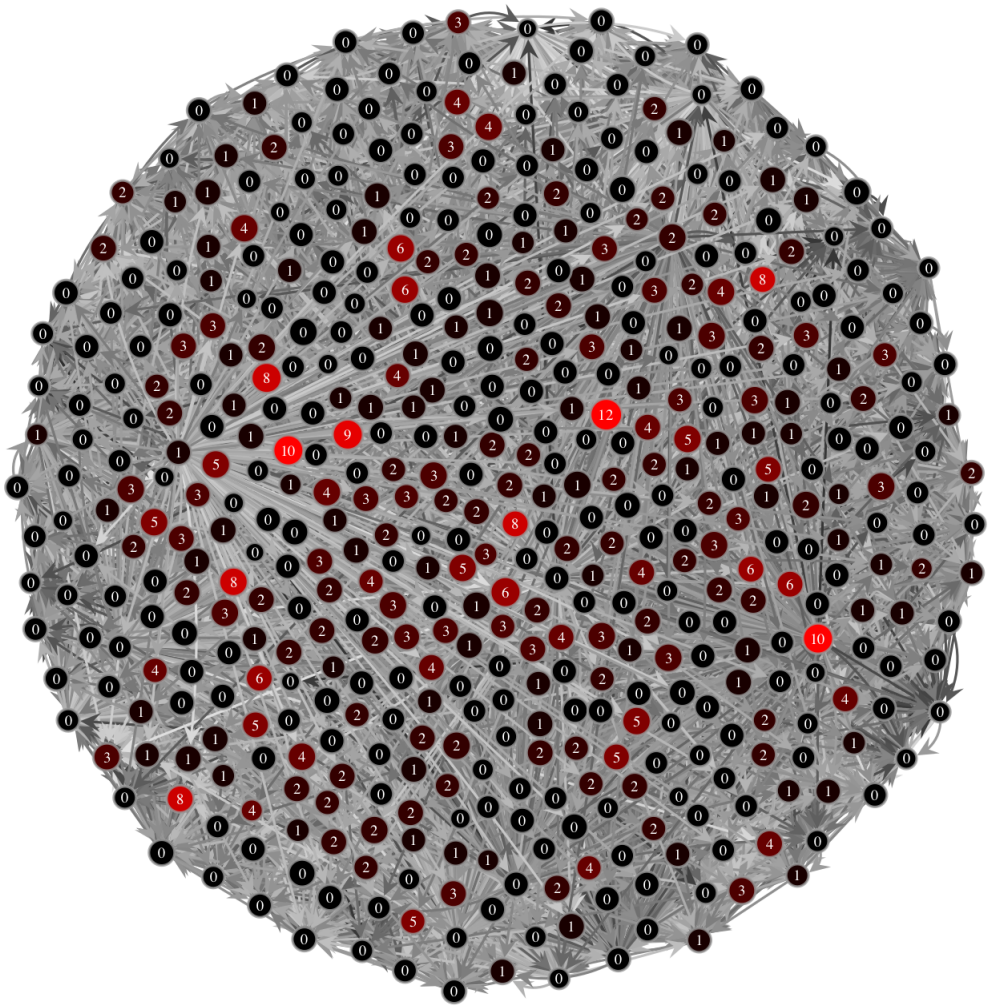

**Figure 5.9: Interference accumulation per node of Neuroidal model at 1400 memories.** Neuroidal model (n=500, d=128, w=16, r=40) at 1400 memories. Node value indicates number of interferences that occured at each node.

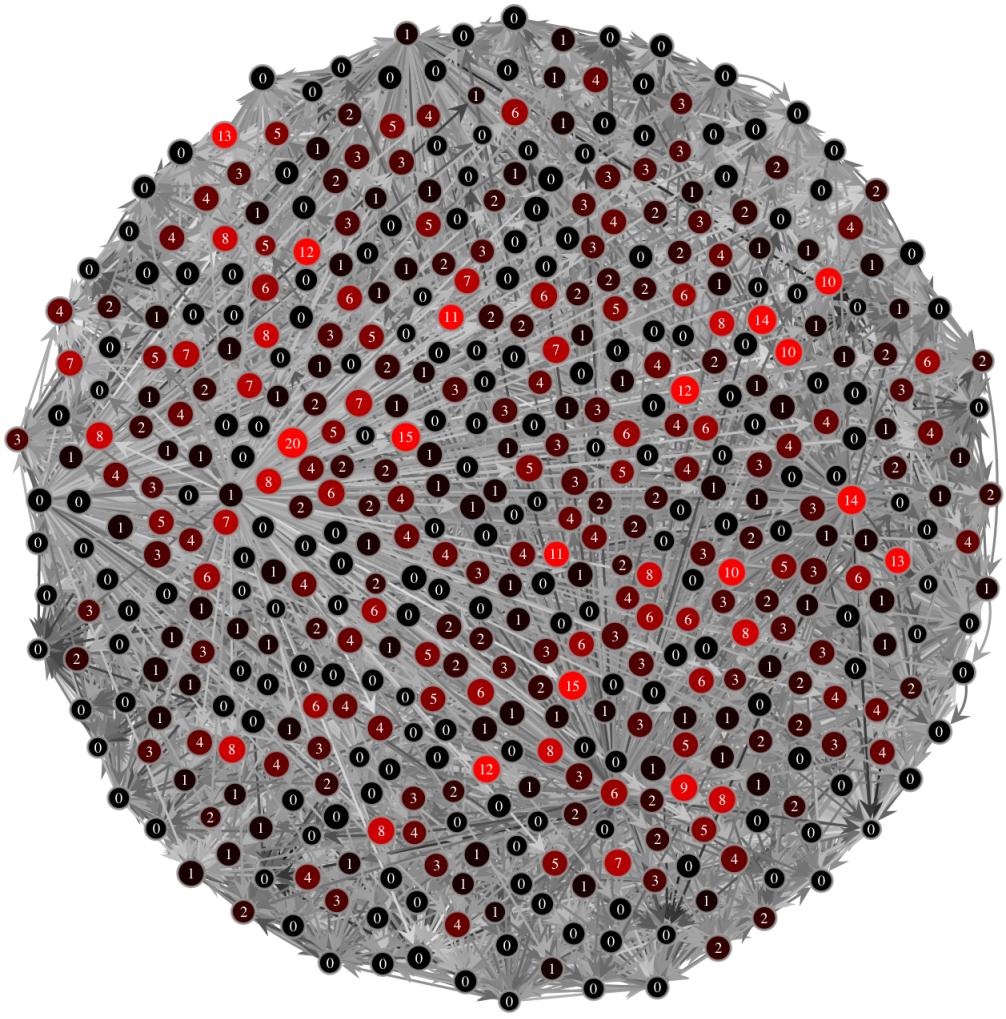

**Figure 5.10: Interference accumulation per node of Neuroidal model at 1600 memories.** Neuroidal model (n=500, d=128, w=16, r=40) at 1600 memories. Node value indicates number of interferences that occured at each node.

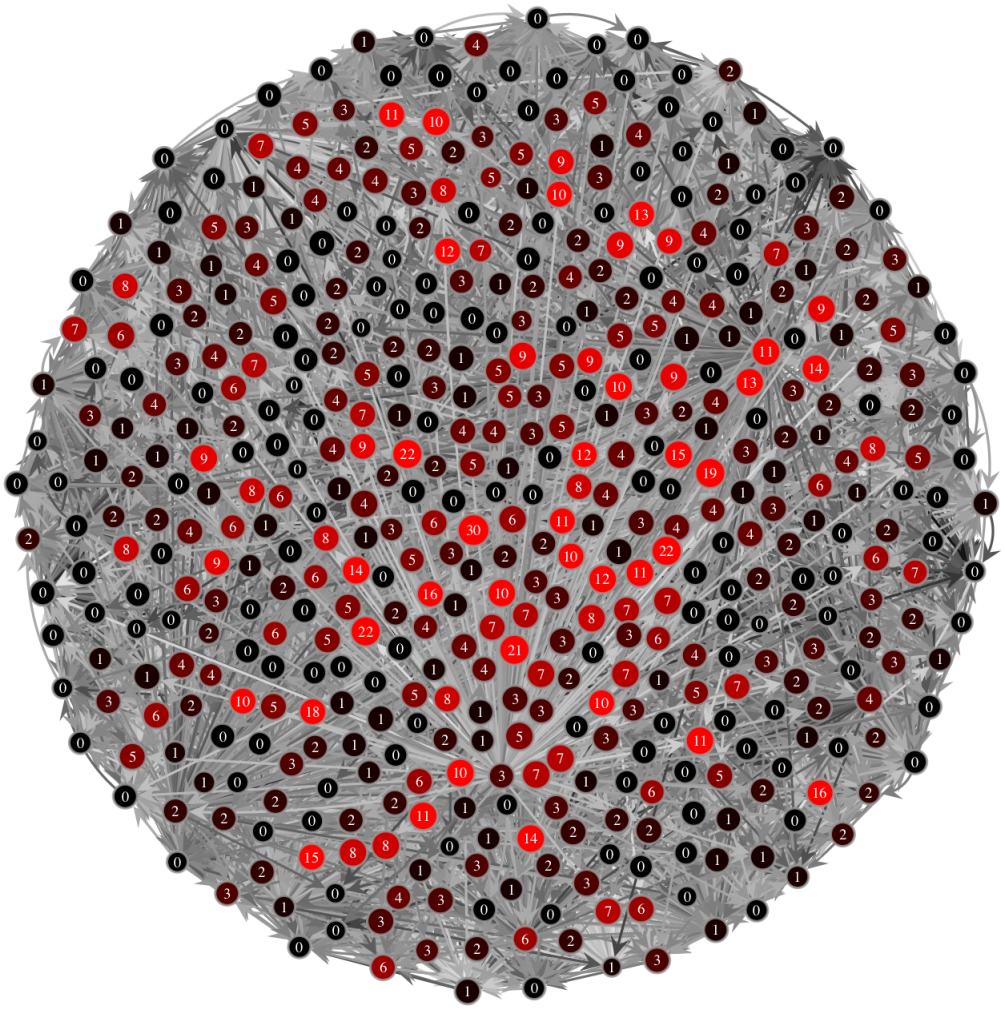

**Figure 5.11: Interference accumulation per node of Neuroidal model at capacity.** Neuroidal model (n=500, d=128, w=16, r=40) at capacity (1800 memories). Node value indicates number of interferences that occured at each node.

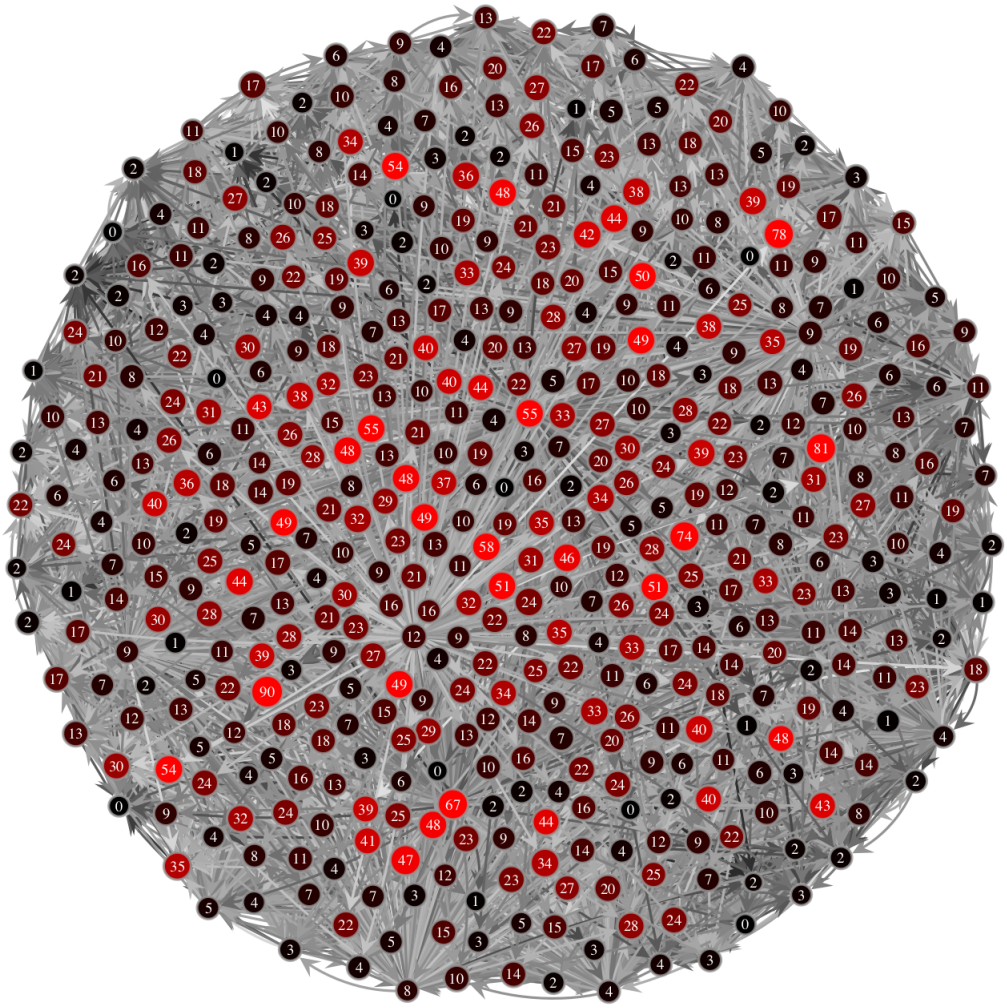

**Figure 5.12: Memory membership per node of Neuroidal model at 1200 memories.** Neuroidal model (n=500, d=128, w=16, r=40) at 1200 memories. Node value indicates memory membership of the node.

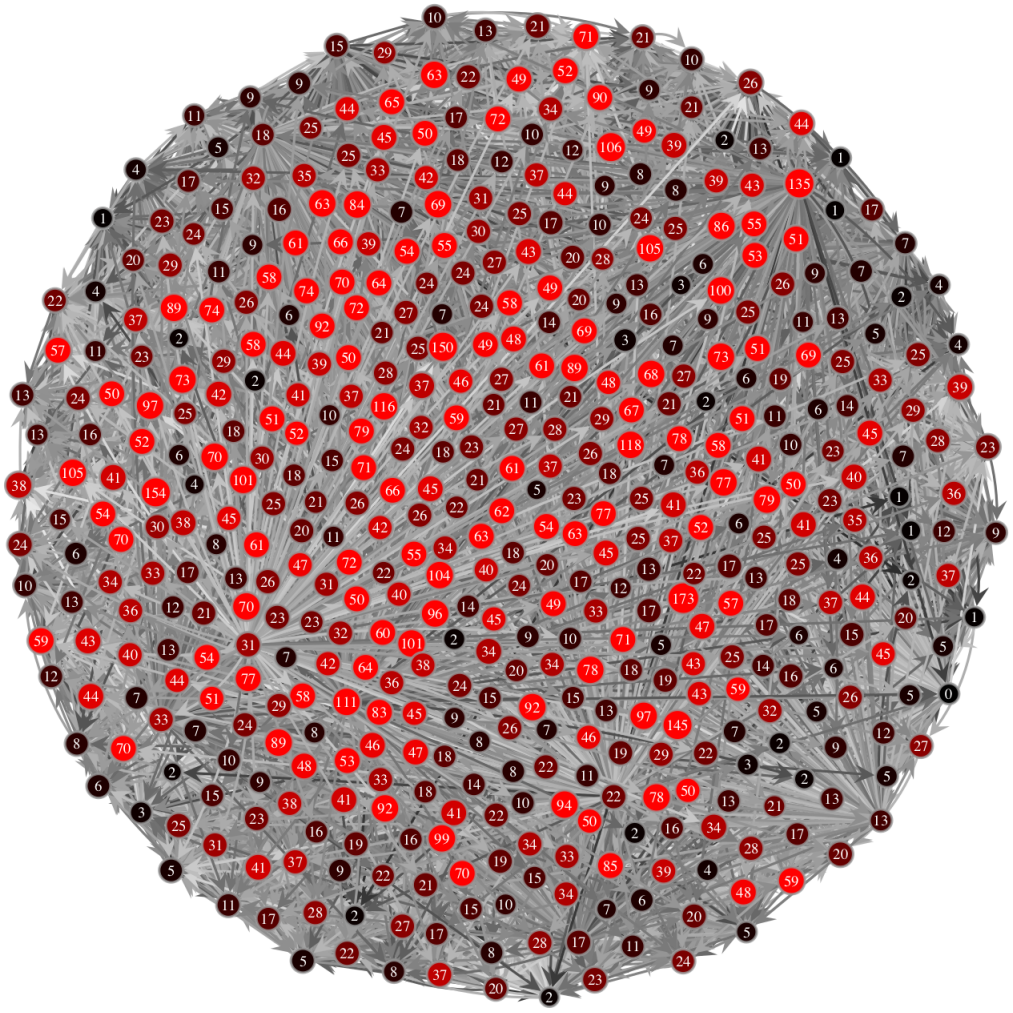

**Figure 5.13: Memory membership per node of Neuroidal model at 1400 memories.** Neuroidal model (n=500, d=128, w=16, r=40) at 1400 memories. Node value indicates memory membership of the node.

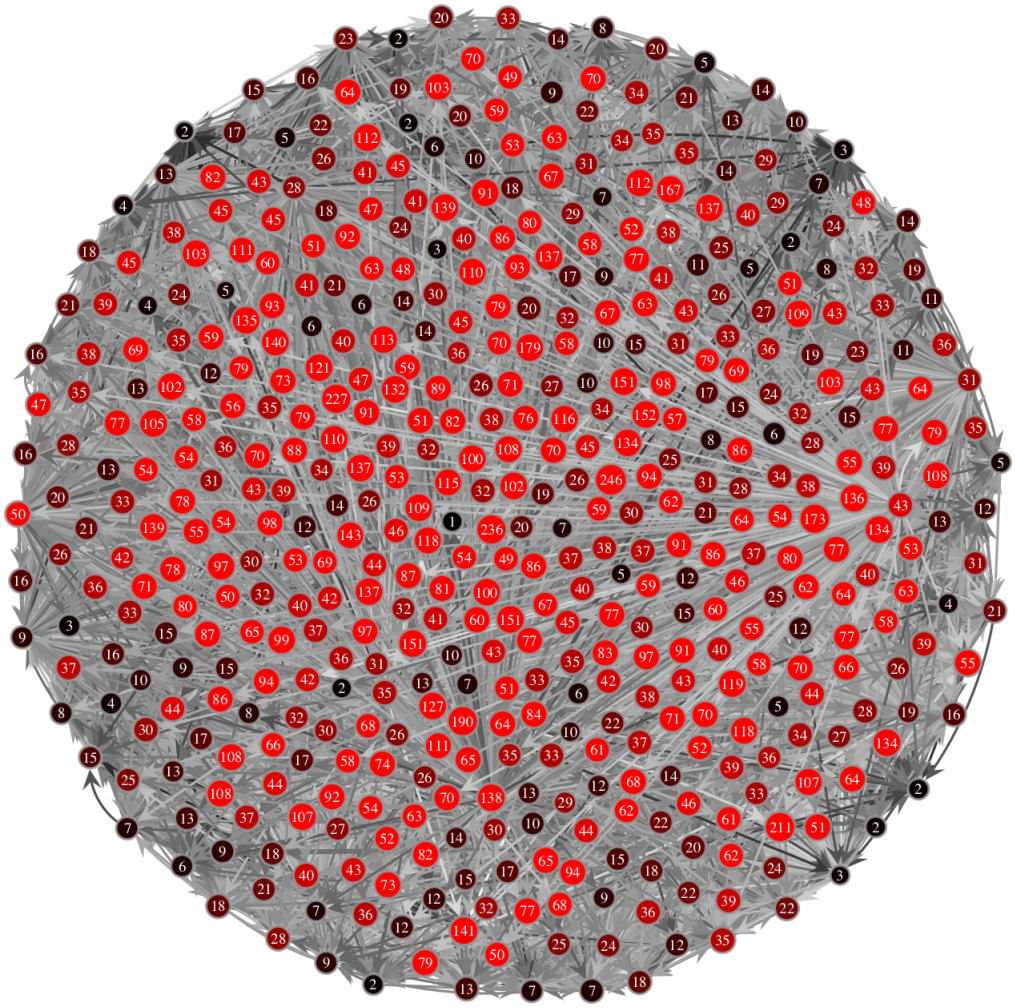

**Figure 5.14: Memory membership per node of Neuroidal model at 1600 memories.** Neuroidal model (n=500, d=128, w=16, r=40) at 1600 memories. Node value indicates memory membership of the node.

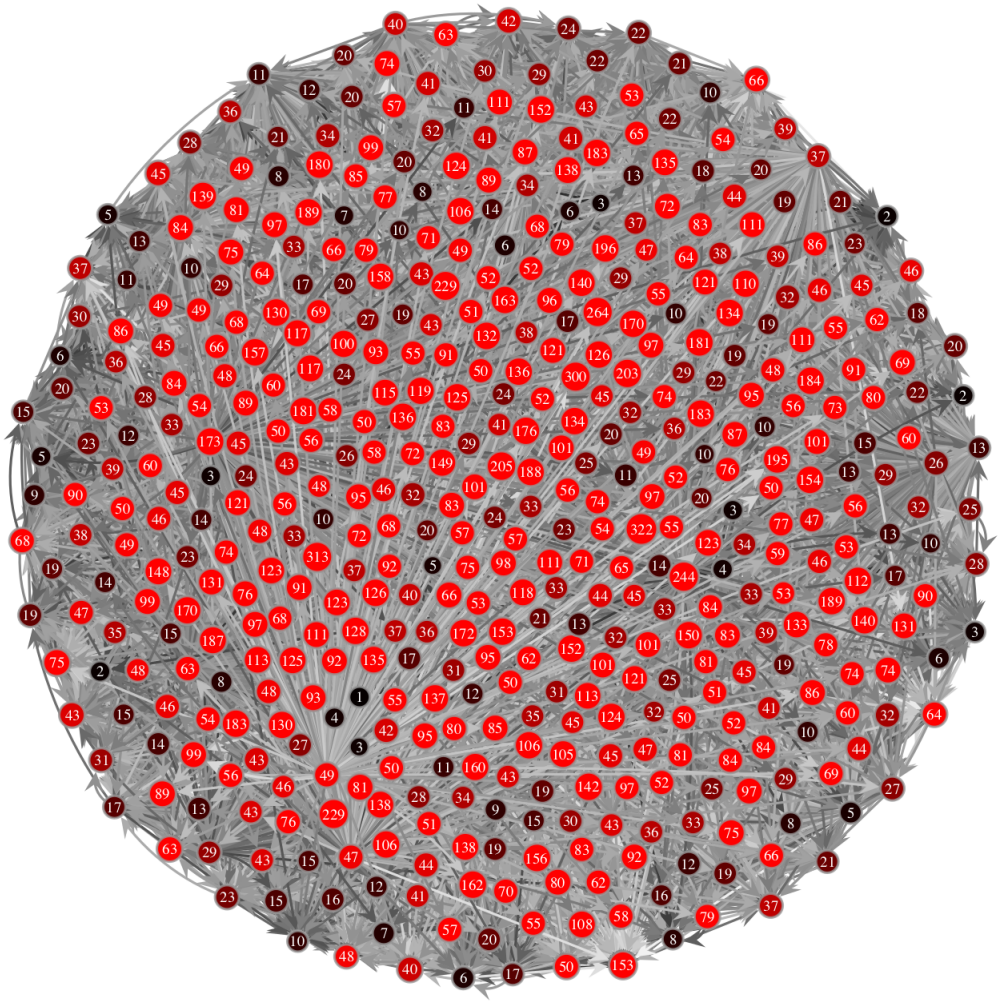

**Figure 5.15: Memory membership per node of Neuroidal model at capacity.** Neuroidal model (n=500, d=128, w=16, r=40) at capacity (1800 memories). Node value indicates memory membership of the node

indegree of the node. We observe that the interference is not uniformly distributed accross the nodes. And as we observe the pattern of evolution, we see that certain nodes end up with a lot of interference while many have negligible amount of interference. We also note that nodes that have already accumulated some interference have a higher chance of being a center of interference again. Practically, this indicates that some nodes are just innately more prone to interference than other nodes due to the initial edge layout. This is in constrast to our theory which assumes that the nodes are equal at the start and throughout the life of the system. We believe this is the primary factor that will need to be accounted for in attempts to update the theory and in particular, lemma 4 to be compatible with JOIN. We believe this is really challenging and leave it as future work.

We can also observe from the figures that the number of memories a node is part does not seem to have any correlation to the number of interferences at that node and therefore is not a good indicator of future interference at that node and cannot be used as a heuristic to update our estimate. There does seem to be some correlation between the in-degree and the interference and could provide intuition regarding updating lemma 4. As such, we feel analyzing the in-degree distribution will be key to finding the capacity for JOIN.

# Chapter 6

# CONCLUSION

Inspired by advances in modern computational neuroscience and growing interest in the capacity of the brain, we rigorously defined and studied the notion of "capacity" and "interference" in a set both theoretically and emprically. We also provided ideas to extend these results to more structured objects like graphs with more advanced algorithms for adding subobjects to the universe.

## 6.1 Future work

Here we discuss some potential future work building off this study:

- Adapt lemma 4 to find the expected interference in the case of other memory creation algorithms like
  - JOIN: This we believe will be the most challenging step as it involves deriving an estimate for the capacity based on the indegree distribution that is non-uniform throughout the graph and also over time.
  - Project: We believe that once we have an estimate for JOIN, it will be very easy to find an estimate for Project as these are very similar operations however with slightly different goals. The main thing to note here wil be that assemblies are more densely connected than arbsets and that will the key here.
  - Merge: We feel an estimate for merge would be very similar to those of JOIN and Project as it is essentially an amalgation of the two.

– Sequence Project: A analysis of the capacity of Sequence Project will be very interesting as it is the only algorithm we discussed that creates more than a single memory. We are excited to see how this will affect the analysis as well as final estimate.

We believe each one of these will provide considerable challenges due to their complex nature [5, 25, 32]. The rest of the theorems will follow similarly to be able to find the capacity of the model and the final expression should have the same general structure. We also think the estimates will follow similar trends against the number of neurons and number of memories.

- Instead of bounding the subset sizes, assume the subset sizes are drawn from a distribution with a given mean $r$ and find the expected subset capacity. This will involve finding the expectation of the hypergeometric PMF as a function of two random variables.

## 6.2 Closing thoughts

The study of capacity with regards to interference is really important as computational models of the brain need to keep the number of misfires low to be able to accumulate memories for a long period of time as well as maintain a high quality of retrieval. We believe this study will inspire more computational neuroscientists to tackle the intriguing question of capacity in contemporary models as well as upcoming models that will further demistify the human brain.

# BIBLIOGRAPHY


[1] C. M. Altimus, B. J. Marlin, N. E. Charalambakis, A. Colón-Rodríguez, E. J. Glover, P. Izbicki, A. Johnson, M. V. Lourenco, R. A. Makinson, J. McQuail, et al. The Next 50 Years of Neuroscience. Journal of Neuroscience, 40(1):101–106, 2020.

[2] M. Bear, B. Connors, and M. A. Paradiso. Neuroscience: Exploring The Brain. Jones & Bartlett Learning, 4th (enhanced) edition, 2020.

[3] G. Bi and M. Poo. Synaptic Modifications in Cultured Hippocampal Neurons: Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type. Journal of Neuroscience, 18(24):10464–10472, 1998.

[4] G. Buzsáki. Neural Syntax: Cell Assemblies, Synapsembles, and Readers. Neuron, 68(3):362–385, 2010.

[5] M. Dabagia, C. H. Papadimitriou, and S. S. Vempala. Computation with sequences in the brain. arXiv preprint arXiv:2306.03812, 2023.

[6] P. Erdős and A. Rényi. On the Evolution of Random Graphs. Matematikai Kutató Intézetének Közleményei: A Magyar Tudományos Akadémia, 5:17–61, 1960.

[7] V. Feldman and L. G. Valiant. Experience-Induced Neural Circuits that Achieve High Capacity. Neural Computation, 21(10):2715–2754, 2009.

[8] K. M. Franks, M. J. Russo, D. L. Sosulski, A. A. Mulligan, S. A. Siegelbaum, and R. Axel. Recurrent circuitry dynamically shapes the activation of piriform cortex. Neuron, 72(1):49–56, 2011.

[9] K. Fukushima. Neocognitron: A Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. Transactions of the Institute of Electronics, Information and Communication Engineers, 62(10):658–665, 1979.

[10] K. D. Harris, J. Csicsvari, H. Hirase, G. Dragoi, and G. Buzsáki. Organization of Cell Assemblies in the Hippocampus. Nature, 424(6948):552–556, 2003.

[11] D. O. Hebb. The first stage of perception: growth of the assembly. The Organization of Behavior, 4(60):78–60, 1949.

[12] D. O. Hebb. The Organization of Behavior. Wiley & Sons, New York, 1949.

[13] S. Herculano-Houzel. The Human Brain in Numbers: A Linearly Scaled-Up Primate Brain. Frontiers in Human Neuroscience, 3, 2009.

[14] J. J. Hopfield. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proceedings of the National Academy of Sciences, 79(8):2554–2558, 1982.

[15] D. H. Hubel and T. N. Wiesel. Receptive Fields of Single Neurones in The Cat's Striate Cortex. The Journal of Physiology, 148(3):574, 1959.

[16] D. Krotov and J. J. Hopfield. Large Associative Memory Problem in Neurobiology and Machine Learning. In 9th International Conference on Learning Representations, 2021.

[17] N. Lynch, C. Musco, and M. Parter. Computational Tradeoffs in Biological Neural Networks: Self-Stabilizing Winner-Take-All Networks. arXiv preprint arXiv:1610.02084, 2016. Unpublished Manuscript.

[18] W. Maass. Networks of Spiking Neurons: The Third Generation of Neural Network Models. Neural Networks, 10(9):1659–1671, 1997.

[19] W. Maass, C. H. Papadimitriou, S. Vempala, and R. Legenstein. Brain Computation: A Computer Science Perspective. Computing and Software Science: State of the Art and Perspectives, pages 184–199, 2019.

[20] G. Marcus. Deep Learning: A Critical Appraisal. arXiv preprint arXiv:1801.00631, 2018. Unpublished Manuscript.

[21] D. Marr. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. W. H. Freeman and Company, San Francisco, 1982.

[22] W. S. McCulloch and W. Pitts. A Logical Calculus of the Ideas Immanent in Nervous Activity. The Bulletin of Mathematical Biophysics, 5:115–133, 1943.

[23] C. H. Papadimitriou and A. D. Friederici. Bridging the Gap Between Neurons and Cognition Through Assemblies of Neurons. Neural Computation, 34(2):291–306, January 2022.

[24] C. H. Papadimitriou and S. S. Vempala. Unsupervised Learning through Prediction in a Model of Cortex. arXiv preprint arXiv:1412.7955, 2014. Unpublished Manuscript.

[25] C. H. Papadimitriou, S. S. Vempala, D. Mitropolsky, M. Collins, and W. Maass. Brain Computation by Assemblies of Neurons. Proceedings of the National Academy of Sciences, 117(25):14464–14472, 2020.

[26] T. P. Peixoto. The graph-tool Python library, Version 14, May 2017.

[27] P. R. Perrine. Neural Tabula Rasa: Foundations for Realistic Memories and Learning. Master's thesis, California Polytechnic State University, San Luis Obispo, June 2023.

[28] M. Reigl, U. Alon, and D. B. Chklovskii. Search for Computational Modules in the *C. elegans* Brain. BMC Biology, 2:1–12, 2004.

[29] E. L. Schwartz, editor. Computational Neuroscience. System Development Foundation Benchmark Series. MIT Press, Cambridge, August 1993.

[30] L. G. Valiant. A Theory of the Learnable. Communications of the ACM, 27(11):1134–1142, 1984.

[31] L. G. Valiant. Circuits of The Mind. Oxford University Press, Paperback edition, 1999.

[32] L. G. Valiant. Memorization and Association on a Realistic Neural Model. Neural Computation, 17(3):527–555, 2005.

[33] L. G. Valiant. A Quantitative Theory of Neural Computation. Biological Cybernetics, 95(3):205–211, 2006.

[34] L. G. Valiant. Capacity of Neural Networks for Lifelong Learning of Composable Tasks. In 58th Annual Symposium on Foundations of Computer Science, pages 367–378. IEEE, 2017.

[35] Y. Xie, Y. Li, and A. Rangamani. Skip connections increase the capacity of associative memories in variable binding mechanisms. Technical report, Center for Brains, Minds and Machines (CBMM), 2023.

[36] G. R. Yang and X.-J. Wang. Artificial Neural Networks for Neuroscientists: A Primer. Neuron, 107(6):1048–1070, 2020.