

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

© 2023

Chandradeep Chowdhury
ALL RIGHTS RESERVED

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.
Professor of Computer Science

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

COMMITTEE MEMBER: Theresa Anne Migler, Ph.D.
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 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.
- 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

	Page
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER	
1 Introduction	1
2 Background	2
3 Related work	3
4 Methods	4
5 Results	5
5.1 Theoretical results	5
5.1.1 Interference	5
5.1.2 Capacity	9
5.2 Empirical results	14
5.2.1 Fixed subset size	14
5.2.2 Bounded subset size	18
5.2.3 Application to Other Models	23
5.2.4 Capacity and JOIN	24
6 Conclusion	35
6.1 Future work	35
6.2 Closing thoughts	35
BIBLIOGRAPHY	36
APPENDICES	

LIST OF TABLES

Table	Page
-------	------

LIST OF FIGURES

Figure	Page
5.1 Capacity vs. Size of the set (n)	15
5.2 Capacity vs. Size of the subsets (r)	16
5.3 Capacity vs. Size of the subsets (r) for $n = 500$	17
5.4 Capacity vs. Size of the set (n)	19
5.5 Capacity vs. Size of the subsets (r)	20
5.6 Capacity vs. Size of the subsets (r)	21
5.7 Capacity vs. Size of the subsets (r)	22
5.8 Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1200 memories. Node value indicates number of interferences that occurred at each node	26
5.9 Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1400 memories. Node value indicates number of interferences that occurred at each node	27
5.10 Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1600 memories. Node value indicates number of interferences that occurred at each node	28
5.11 Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at capacity (1800 memories)	29
5.12 Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1200 memories .	30
5.13 Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1400 memories .	31
5.14 Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1600 memories .	32
5.15 Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at capacity (1800 memories)	33

Chapter 1

INTRODUCTION

Chapter 2

BACKGROUND

Chapter 3

RELATED WORK

Chapter 4

METHODS

Chapter 5

RESULTS

5.1 Theoretical results

5.1.1 Interference

We now formally define a 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

$$|U \cap W| \geq \frac{|W|}{k}. \quad (5.1)$$

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. 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 unintentional 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 [4]. This corresponds to the $k = 2$ case of our definition.

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 3. *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 . 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 $Y \sim \text{Hypergeometric}(n, r_u, r_w)$.

Proof. If $V = \{v_1, \dots, v_n\}$, 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 denotes the number of indices where both u, w have a 1. First note that

$$\mathbb{P}(Y = y) = \frac{\binom{r_u}{y} \binom{n-r_u}{r_w-y}}{\binom{n}{r_w}}. \quad (5.2)$$

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 $Y \sim \text{Hypergeometric}(n, r_u, r_w)$. Finally, to find the probability of U k -interfering with W we need to find $\mathbb{P}(Y \geq \lceil \frac{r_w}{k} \rceil)$ which is the sum of $\mathbb{P}(Y = y)$ from $y = \lceil \frac{r_w}{k} \rceil$ to $y = r_w$. \square

For brevity, we can reinterpret the above probability as the tail distribution function of Y at $\lceil \frac{r_w}{k} \rceil$,

$$\mathbb{P}\left(Y \geq \lceil \frac{r_w}{k} \rceil\right) = \mathbb{P}\left(Y > \left\lfloor \frac{r_w}{k} \right\rfloor\right) = \bar{F}_Y\left(\left\lfloor \frac{r_w}{k} \right\rfloor\right)$$

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 3 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 δ 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 [4]. 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 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 . If*

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

then

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

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. By 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 [4]. 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

$$\begin{aligned}
\bar{F}_Y \left(\left\lfloor \frac{r_w}{k} \right\rfloor \right) &= \sum_{y=\lceil \frac{r_w}{k} \rceil}^{r_w} \mathbb{P}(Y = y) \\
&= \sum_{y=\lceil \frac{r_w}{k} \rceil}^{r_w} \frac{\binom{r_u}{y} \binom{n-r_u}{r_w-y}}{\binom{n}{r_w}} \\
&\geq \sum_{y=\lceil \frac{r_w}{k} \rceil}^{r_w} \frac{\binom{r-\delta}{y} \binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}} \\
&\geq \sum_{y=\lceil \frac{r+\delta}{k} \rceil}^{\lfloor r-\delta \rfloor} \frac{\binom{r-\delta}{y} \binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}}
\end{aligned} \tag{5.3}$$

The first and second equalities follow from the definition of the tail distribution and lemma 3 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. We now formally define the capacity of a system of overlapping subsets with interference being the limiting factor.

Definition 5. *((r, T, k, δ)-Subset Capacity)* Given a set $V = \{v_1, \dots, v_n\}$, and parameters $r, T, k, \delta > 0$, the (r, T, k, δ) -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 interferences caused due to picking U .

We need the second restriction on the memories here since we want to apply lemma 4 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.

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, δ) -subset capacity of V is*

$$\left\lfloor \frac{T}{\bar{F}_Y(\lfloor \frac{r}{k} \rfloor)} + 1 \right\rfloor.$$

Remark. Since all subsets have fixed size r , note that the choice of δ is not relevant here.

Proof 1. Suppose we have already have $M - 1$ subsets in the universe. Pick a random subset U . From lemma 3, we know that the expected number of k -interferences of U with another arbitrary subset W from the universe is $\bar{F}_Y(\lfloor \frac{r}{k} \rfloor)$. Since there are $M - 1$ other subsets, the total expected number of k -interferences caused by picking U is $(M - 1)\bar{F}_Y(\lfloor \frac{r}{k} \rfloor)$.

From inequality 3 in the definition of capacity, we have

$$(M - 1)\bar{F}_Y\left(\left\lfloor \frac{r}{k} \right\rfloor\right) \leq T \implies M \leq \frac{T}{\bar{F}_Y\left(\left\lfloor \frac{r}{k} \right\rfloor\right)} + 1. \quad (5.4)$$

The (r, T, k, δ) -subset capacity of V then is the largest integer M that satisfies inequality 5.4. \square

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 3, we know that the expected number of k -interferences of U with W is $\bar{F}_Y(\lfloor \frac{r}{k} \rfloor)$. 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

$$2\bar{F}_Y\left(\left\lfloor \frac{r}{k} \right\rfloor\right).$$

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

$$2 \cdot \frac{M(M - 1)}{2} \bar{F}_Y \left(\left\lfloor \frac{r}{k} \right\rfloor \right) = M(M - 1) \bar{F}_Y \left(\left\lfloor \frac{r}{k} \right\rfloor \right).$$

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

$$\frac{M(M - 1)}{M} \bar{F}_Y \left(\left\lfloor \frac{r}{k} \right\rfloor \right) = (M - 1) \bar{F}_Y \left(\left\lfloor \frac{r}{k} \right\rfloor \right).$$

From inequality 3, we have

$$(M - 1) \bar{F}_Y \left(\left\lfloor \frac{r}{k} \right\rfloor \right) \leq T \implies M \leq \frac{T}{\bar{F}_Y \left(\left\lfloor \frac{r}{k} \right\rfloor \right)} + 1. \quad (5.5)$$

The (r, T, k, δ) -subset capacity of V is the largest integer M that satisfies inequality 5.5. \square

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, δ) -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 4. However as $\delta \rightarrow 0$, this expression converges to the expression in theorem 6.

Proof. Suppose we have M subsets U_1, \dots, U_M with sizes r_1, \dots, r_M . Pick two subsets U_i, U_j . From lemma 3, we know that the expected number of interferences caused by

this pair is

$$\bar{F}_Y\left(\left\lfloor \frac{r_j}{k} \right\rfloor\right) + \bar{F}_Y\left(\left\lfloor \frac{r_i}{k} \right\rfloor\right).$$

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

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

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

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

From inequality 3, we have

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

which implies

$$M \geq \frac{1}{T} \sum_{(i,j) \in \mathbb{Z} \times \mathbb{Z}, 1 \leq i, j \leq M, i \neq j} \left(\bar{F}_Y\left(\left\lfloor \frac{r_j}{k} \right\rfloor\right) + \bar{F}_Y\left(\left\lfloor \frac{r_i}{k} \right\rfloor\right) \right). \quad (5.6)$$

Using lemma 4 we get

$$\begin{aligned} M &\geq \frac{1}{T} \sum_{(i,j) \in \mathbb{Z} \times \mathbb{Z}, 1 \leq i, j \leq M, i \neq j} \left(2 \sum_{y=\lceil \frac{r+\delta}{k} \rceil}^{\lfloor r-\delta \rfloor} \frac{\binom{r-\delta}{y} \binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}} \right) \\ &= \frac{1}{T} \frac{M(M-1)}{2} \left(2 \sum_{y=\lceil \frac{r+\delta}{k} \rceil}^{\lfloor r-\delta \rfloor} \frac{\binom{r-\delta}{y} \binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}} \right), \end{aligned} \quad (5.7)$$

which implies

$$M \leq \frac{T}{\sum_{y=\lceil \frac{r+\delta}{k} \rceil}^{\lfloor r-\delta \rfloor} \frac{\binom{r-\delta}{y} \binom{n-r-\delta}{r-\delta-y}}{\binom{n}{r+\delta}}} + 1. \quad (5.8)$$

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

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 throughout our input range.

Then we 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 throughout 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 3. We believe

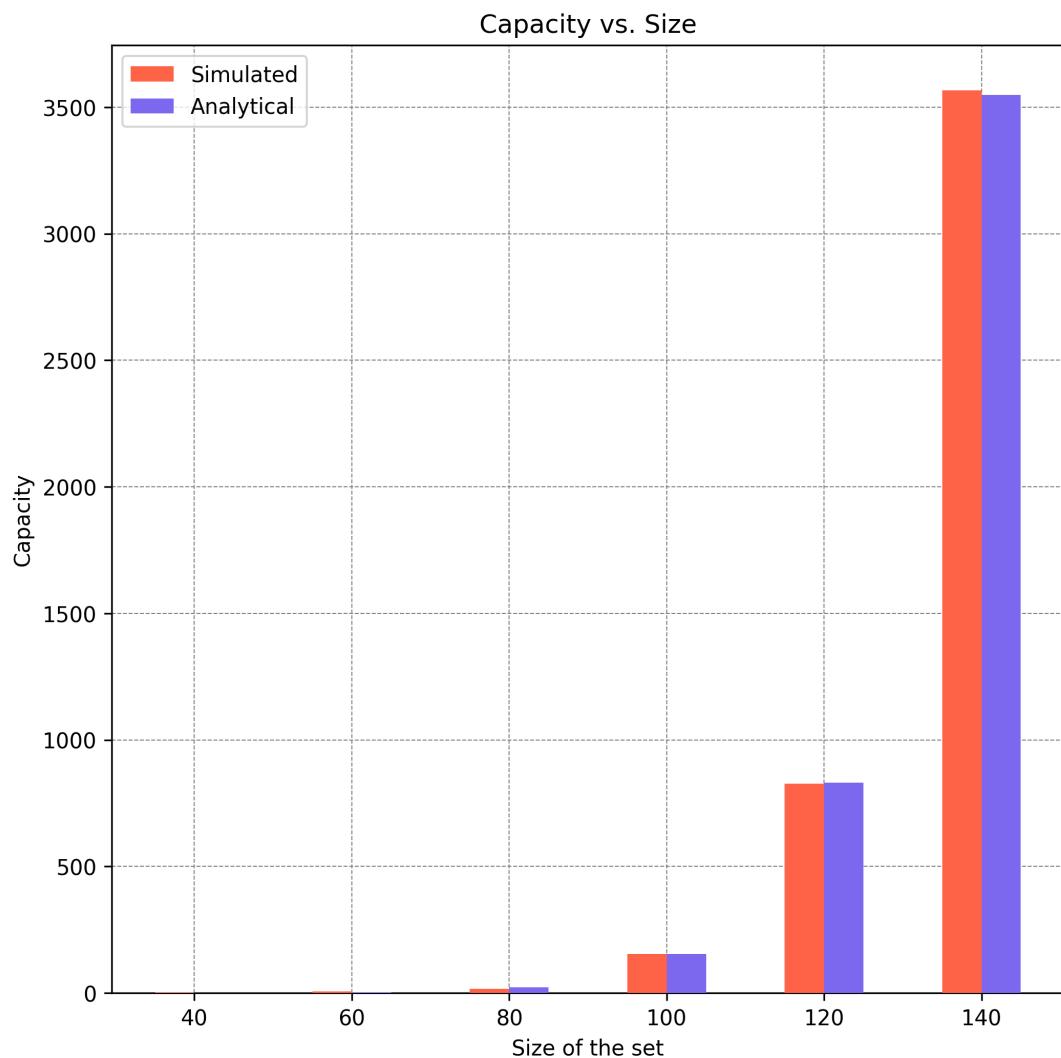


Figure 5.1: How capacity is affected by increasing the size of the set (n).

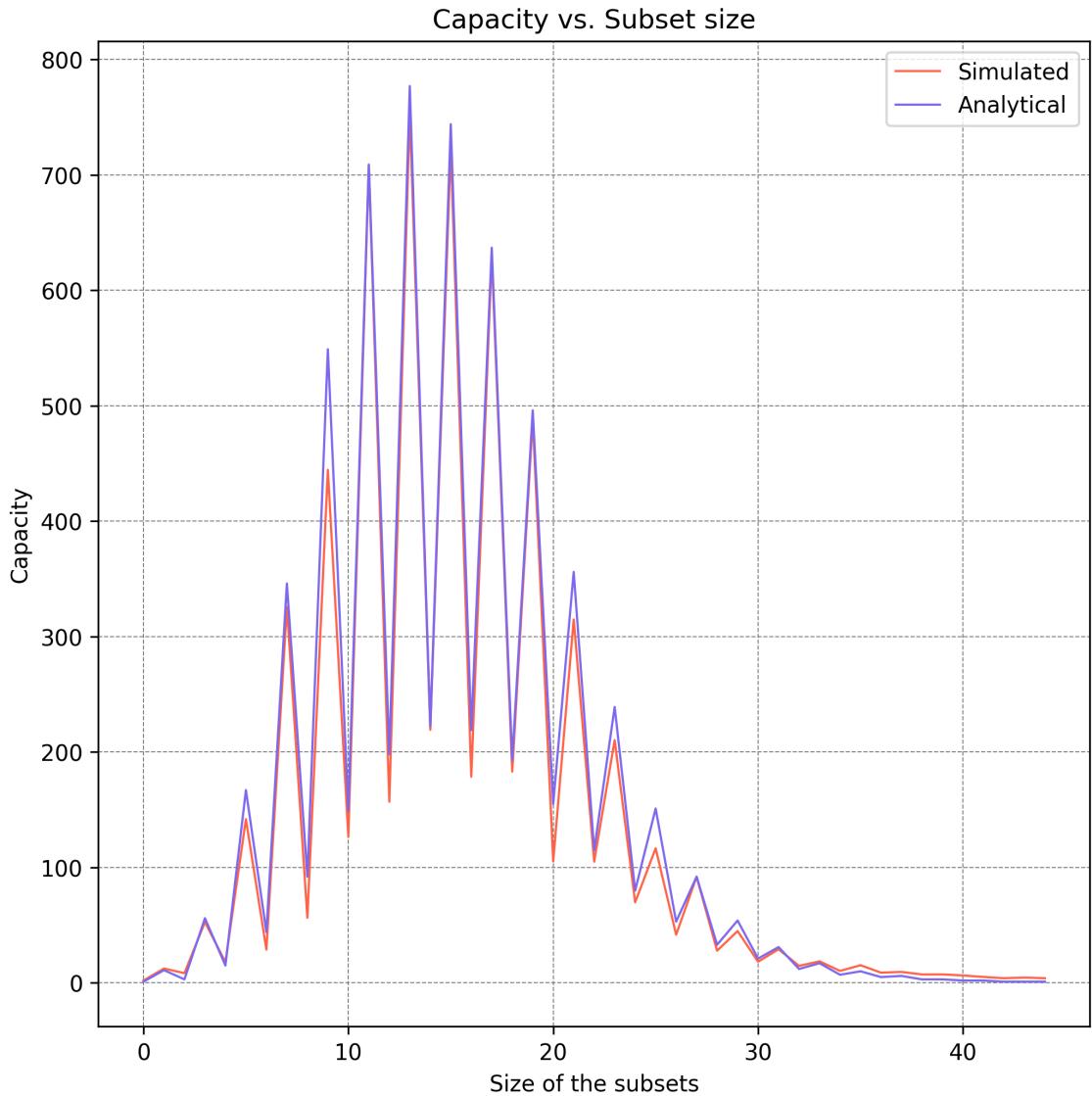


Figure 5.2: How capacity is affected by increasing the size of the subsets (r).

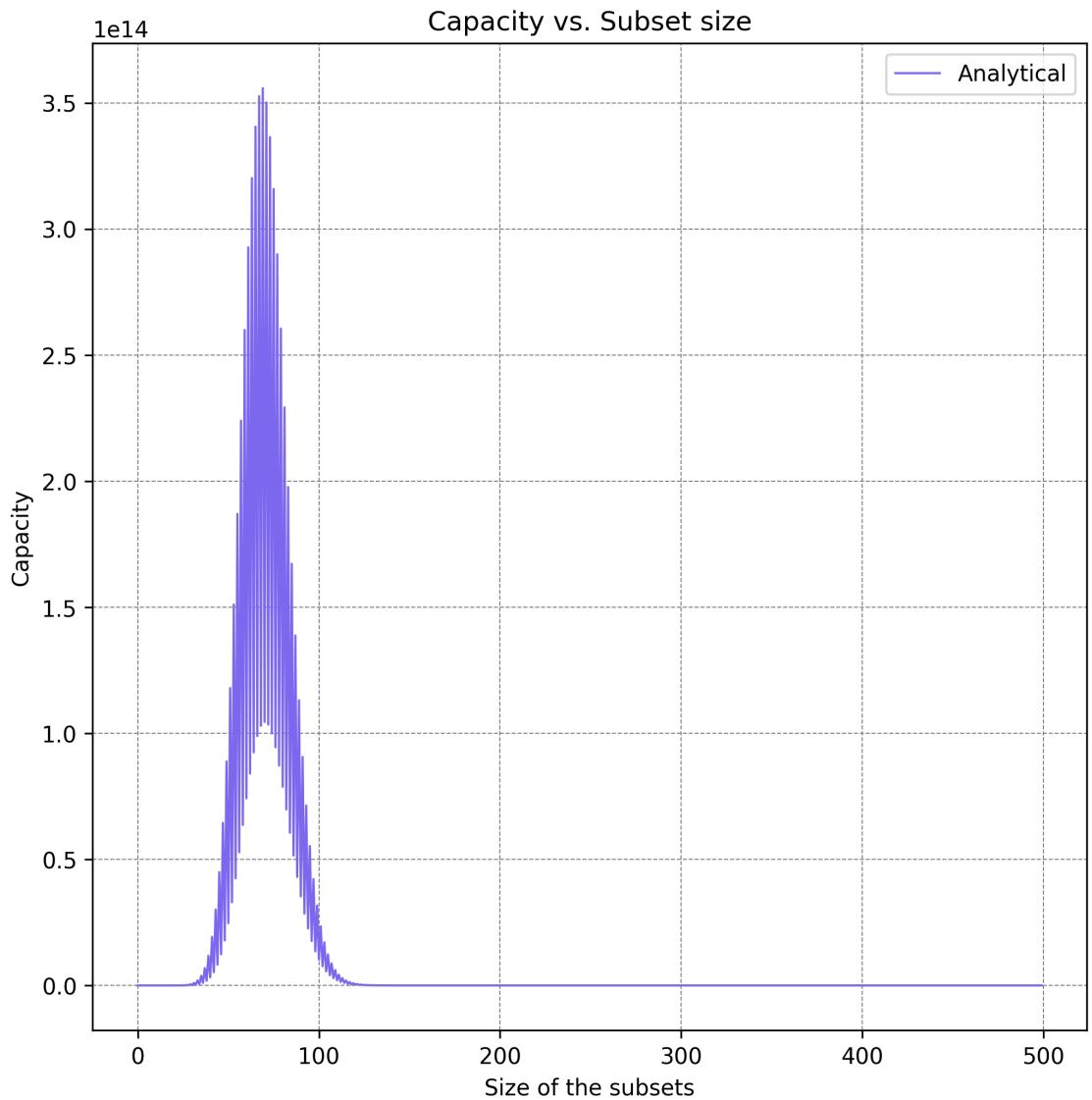


Figure 5.3: How capacity is affected by increasing the size of the subsets (r) when $n = 500$.

these peaks will reduce in intensity relative to the scale of the y axis as $n \rightarrow \infty$. 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.

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 comparison 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 expect 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 though the analytical bound from equation 3 bounds the simulated capacity, the bound is very loose and does not work as an approximation. We believe this is because of how the binomial coefficient varies with respect to its second parameter and that δ 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 5.4 and it was still a good approximation for the simulation with randomly drawn r 's.

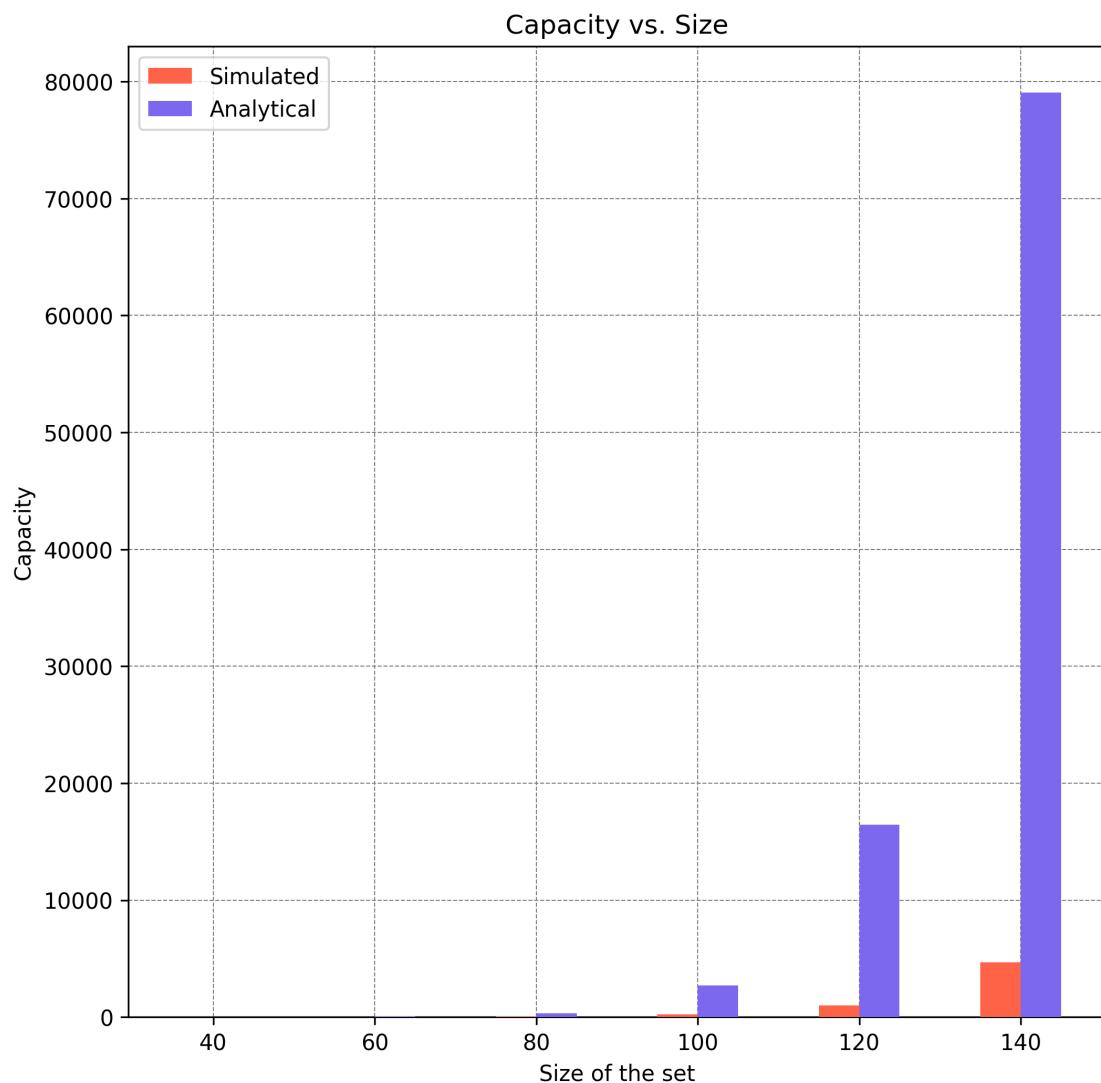


Figure 5.4: How capacity is affected by increasing the size of the subsets (r) when r is drawn from a normal distribution with mean r and standard deviation 1

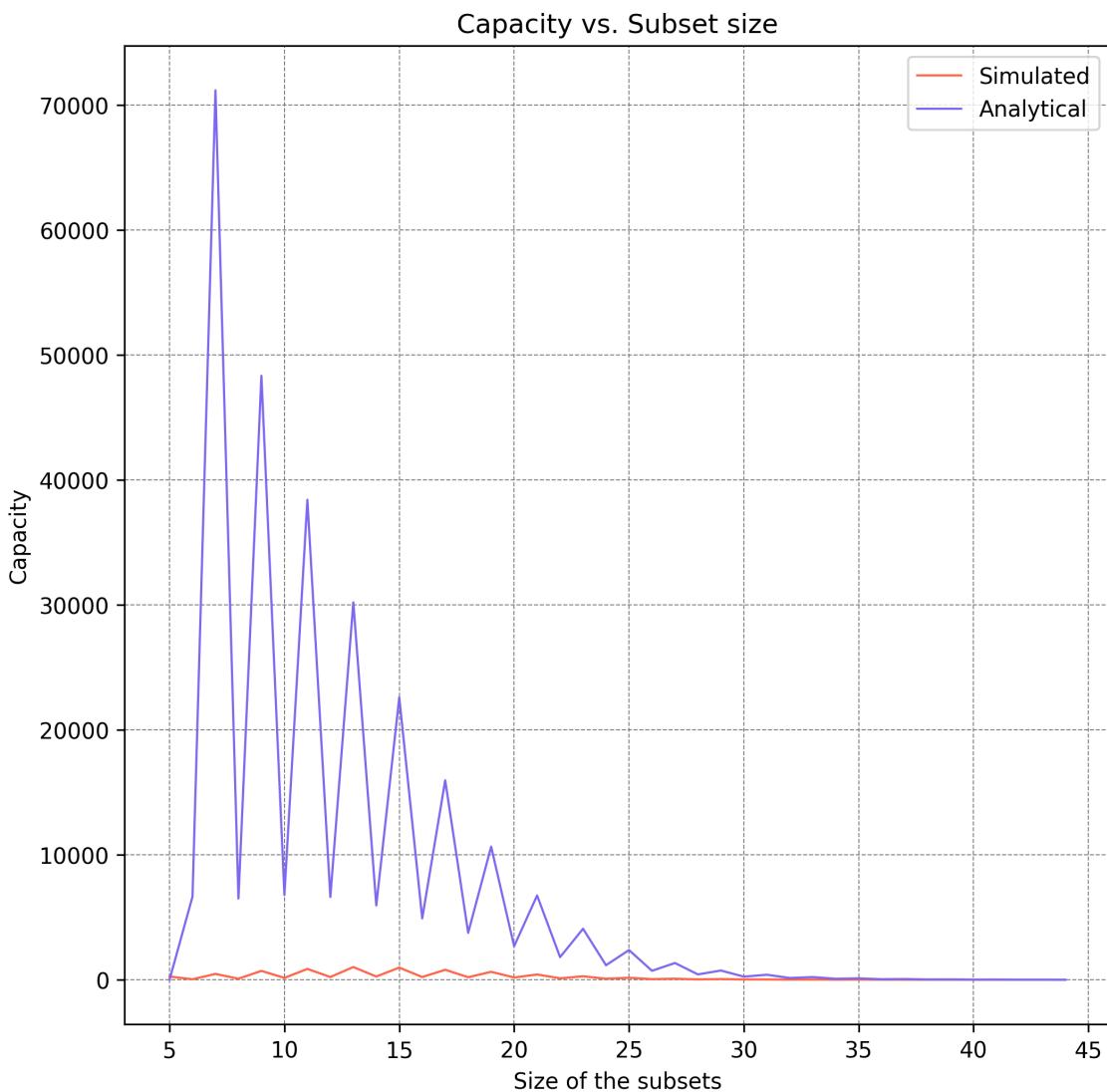


Figure 5.5: How capacity is affected by increasing the size of the subsets (r) when r is drawn from a normal distribution with mean r and standard deviation 1

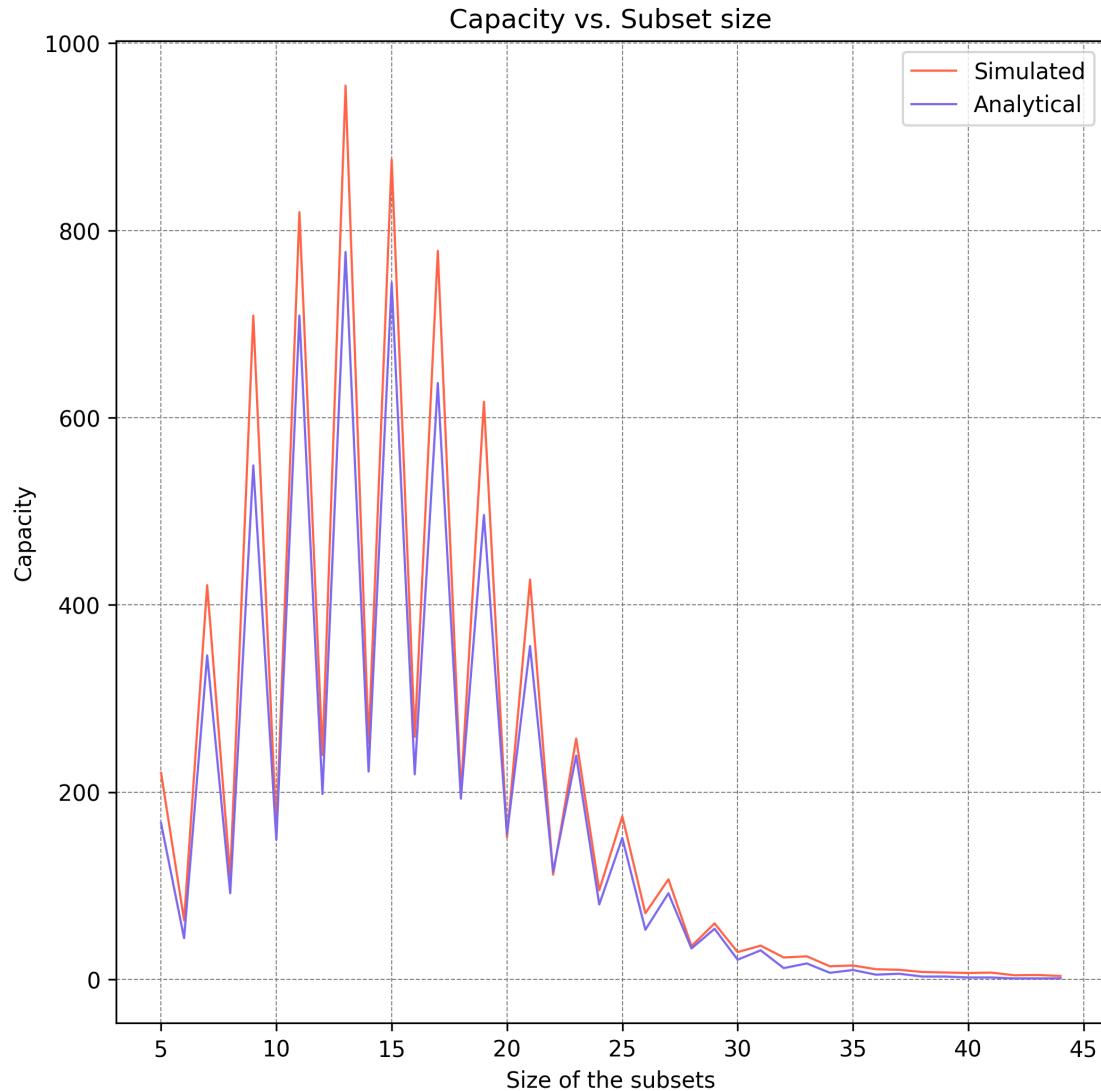


Figure 5.6: How capacity is affected by increasing the size of the subsets (r) comparing the results of 5.4 with the simulation where r is drawn from a normal distribution with mean r and standard deviation 1

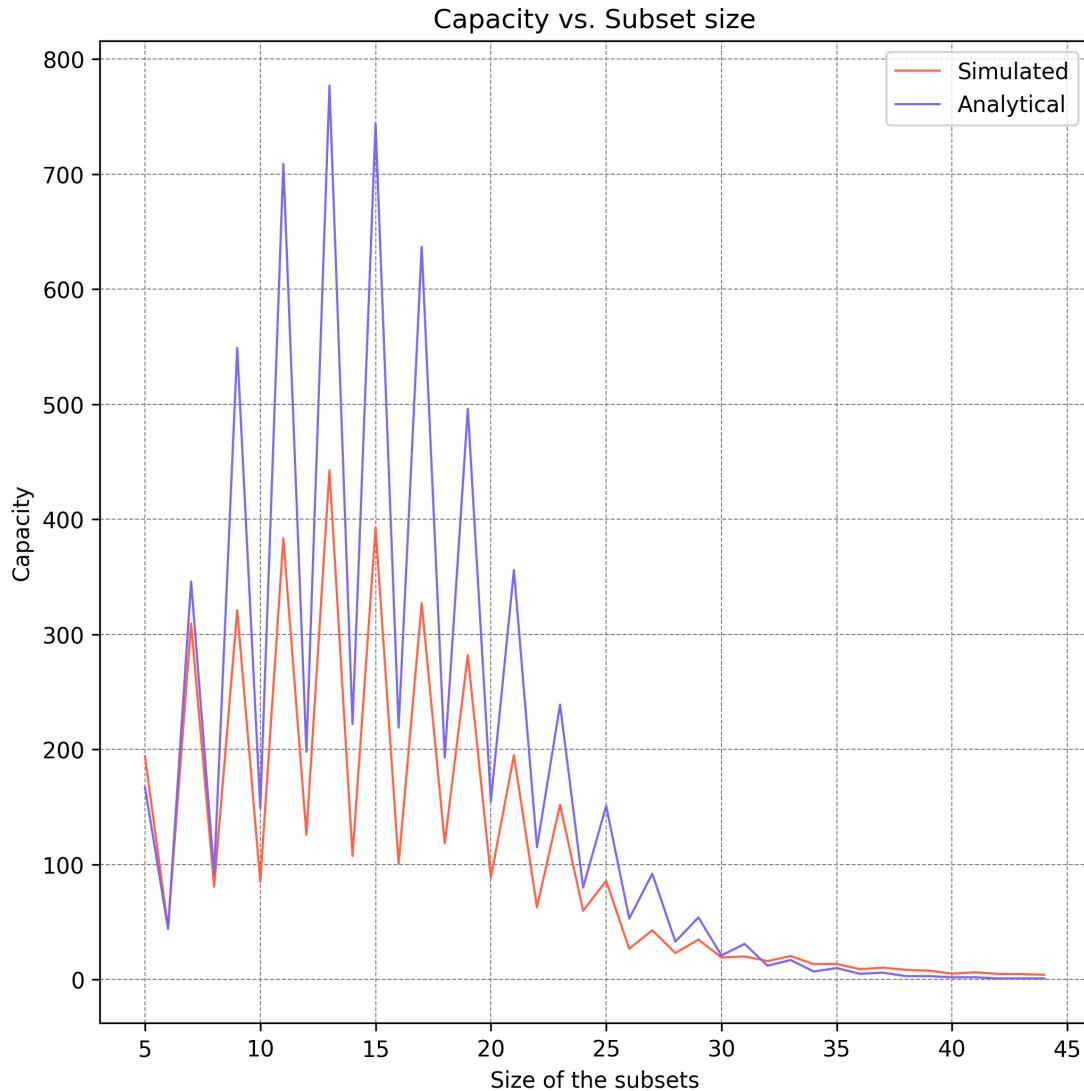


Figure 5.7: How capacity is affected by increasing the size of the subsets (r) comparing the results of 5.4 with the simulation where r is drawn from a normal distribution with mean r and standard deviation 2

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.

5.2.3 Application to Other Models

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 Calculus is with regards to the memory creation algorithm. We implicitly assume a random memory creation algorithm in lemma 3 while the Neuroidal model and Assembly Calculus use the JOIN and Project memory creation algorithms respectively [2] [4]. 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 3 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 and Project 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.

5.2.4 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 parameters.

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 approximation for the Neuroidal model’s capacity however it is effectively an 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 occurred 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 [3]. 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 5, which can be interpreted as total interference over 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

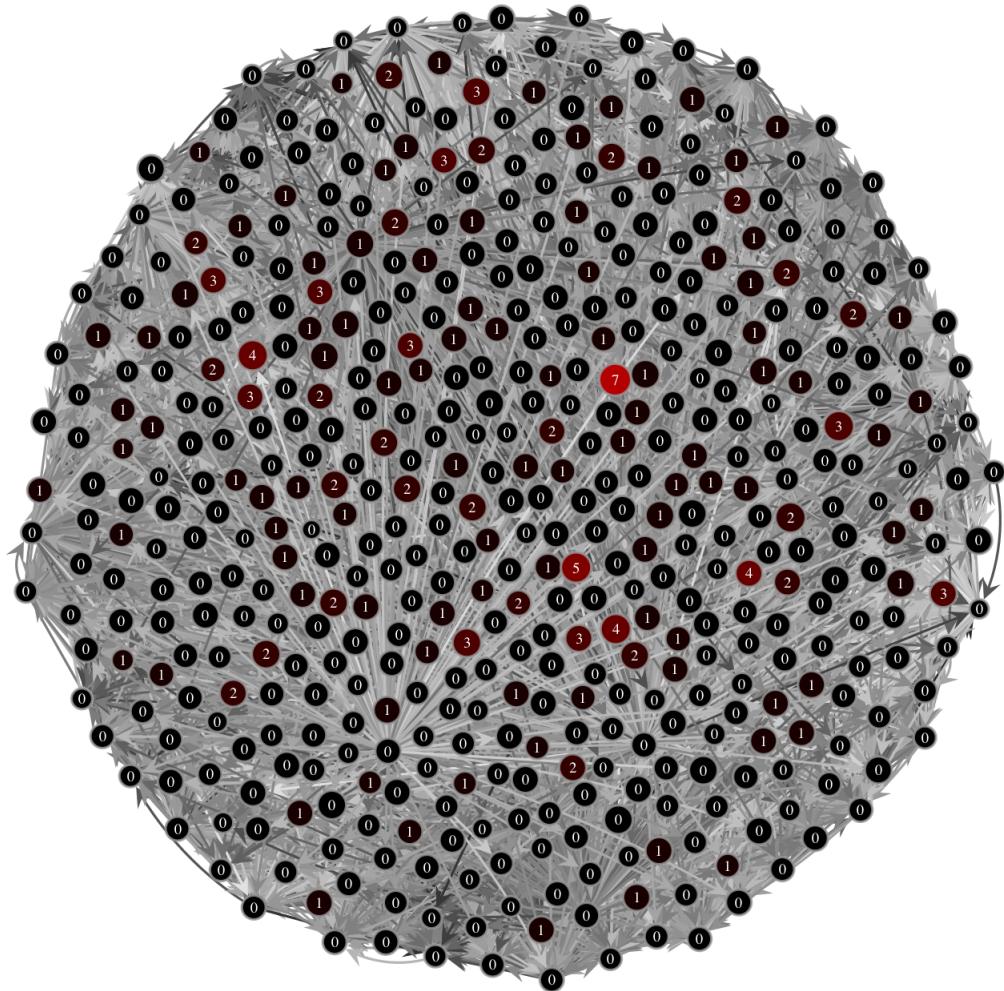


Figure 5.8: Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1200 memories. Node value indicates number of interferences that occurred at each node.

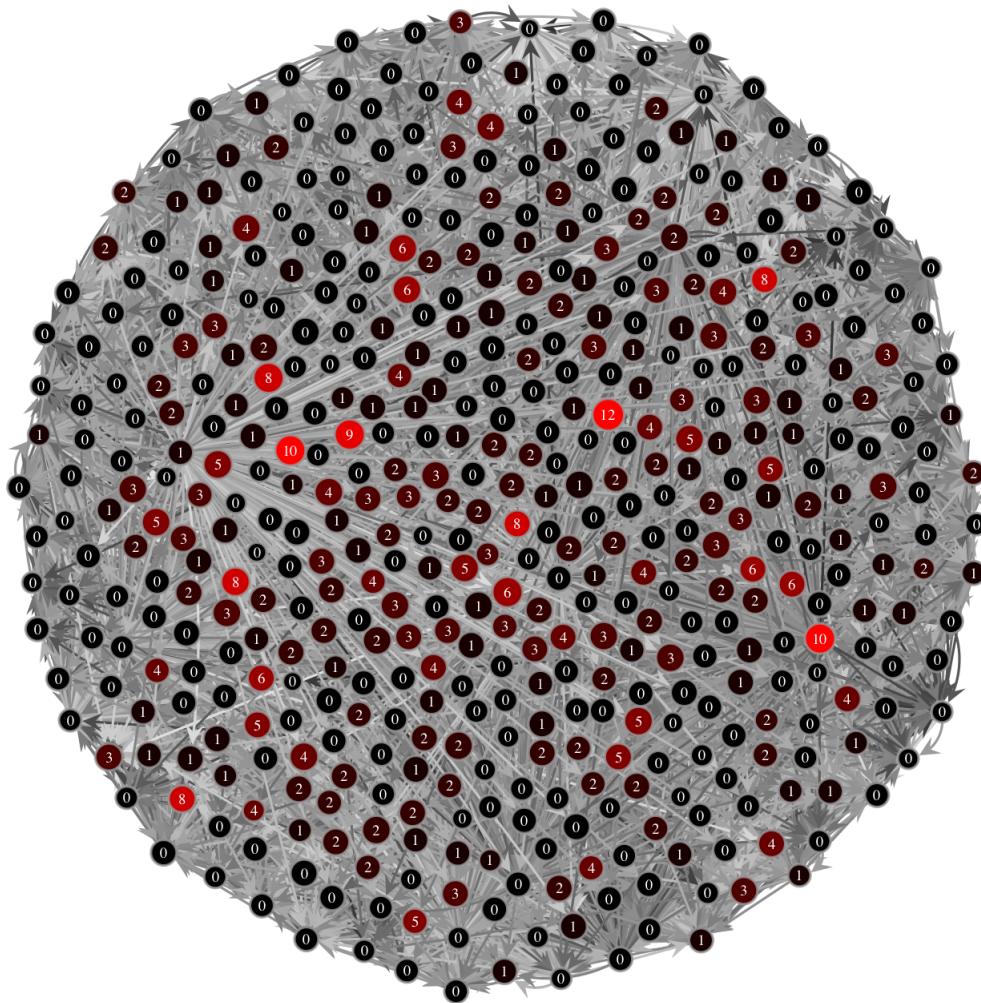


Figure 5.9: Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1400 memories. Node value indicates number of interferences that occurred at each node.

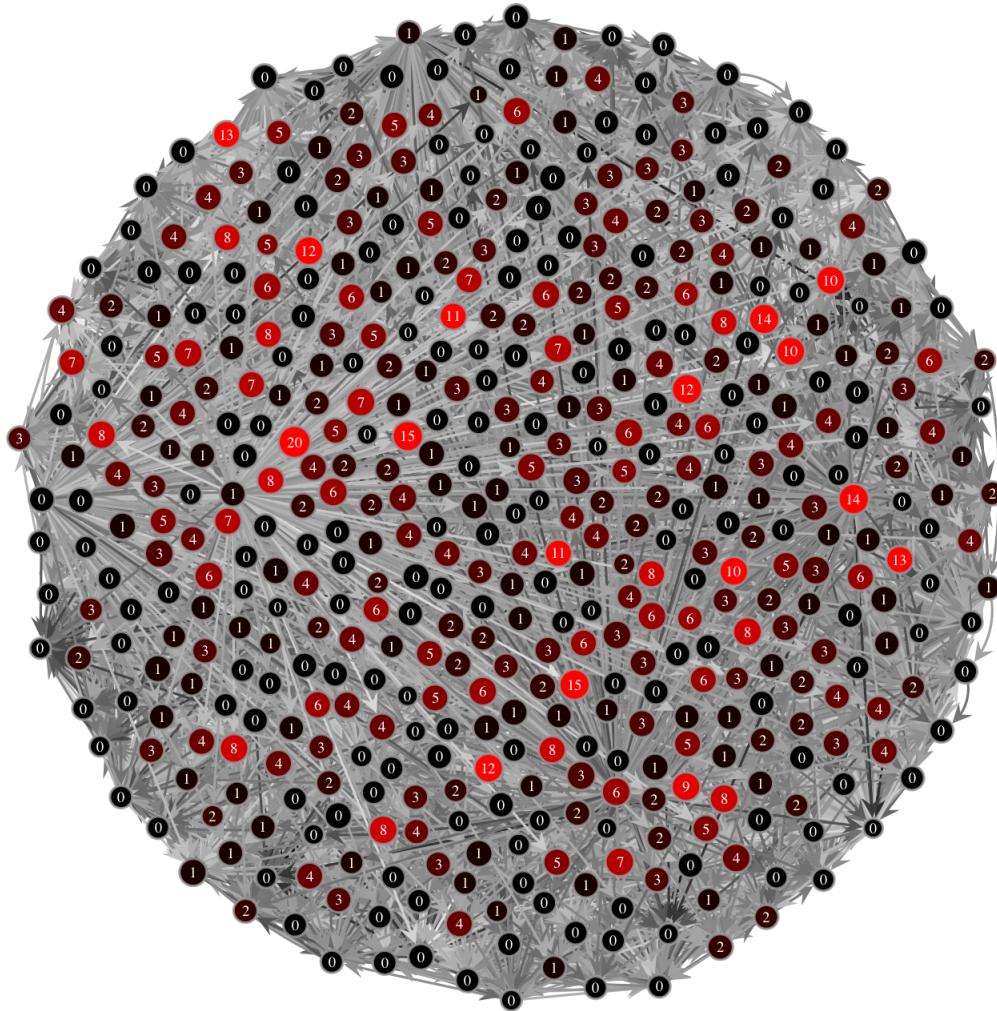


Figure 5.10: Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1600 memories. Node value indicates number of interferences that occurred at each node.

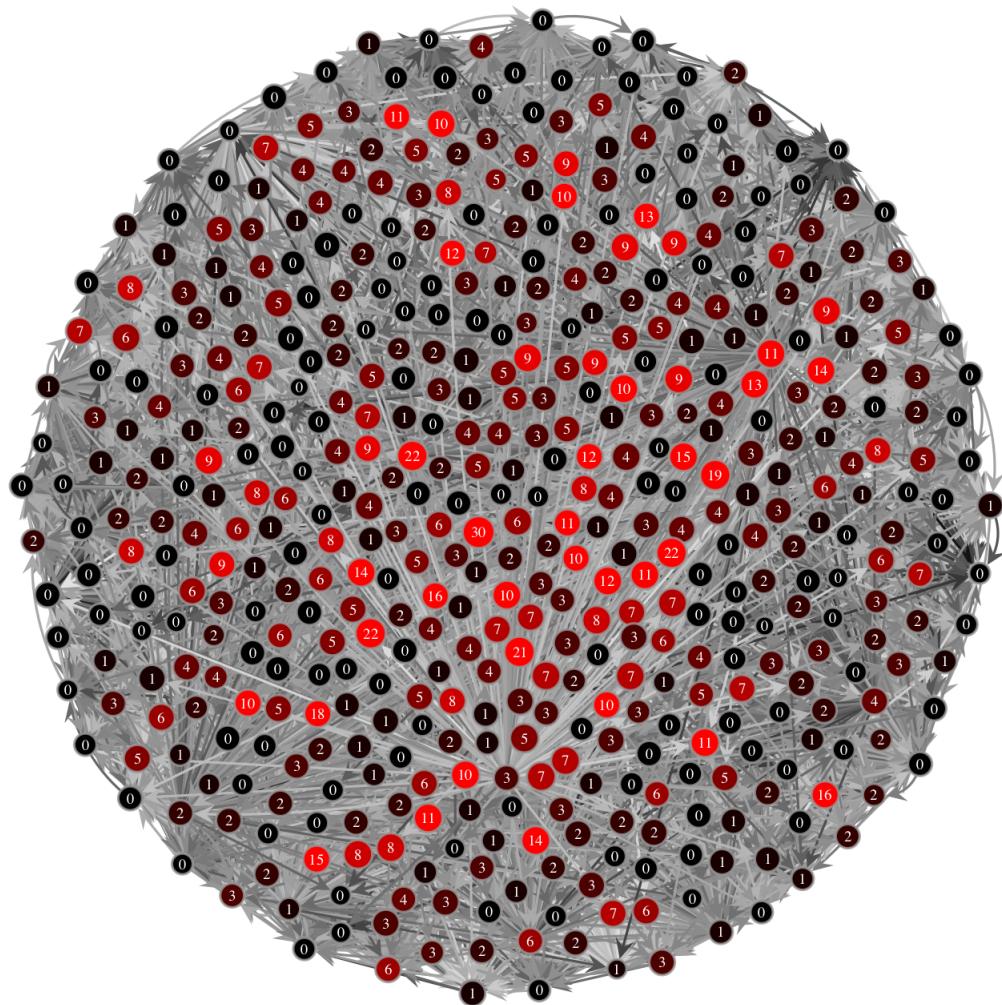


Figure 5.11: Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at capacity (1800 memories). Node value indicates number of interferences that occurred at each node.

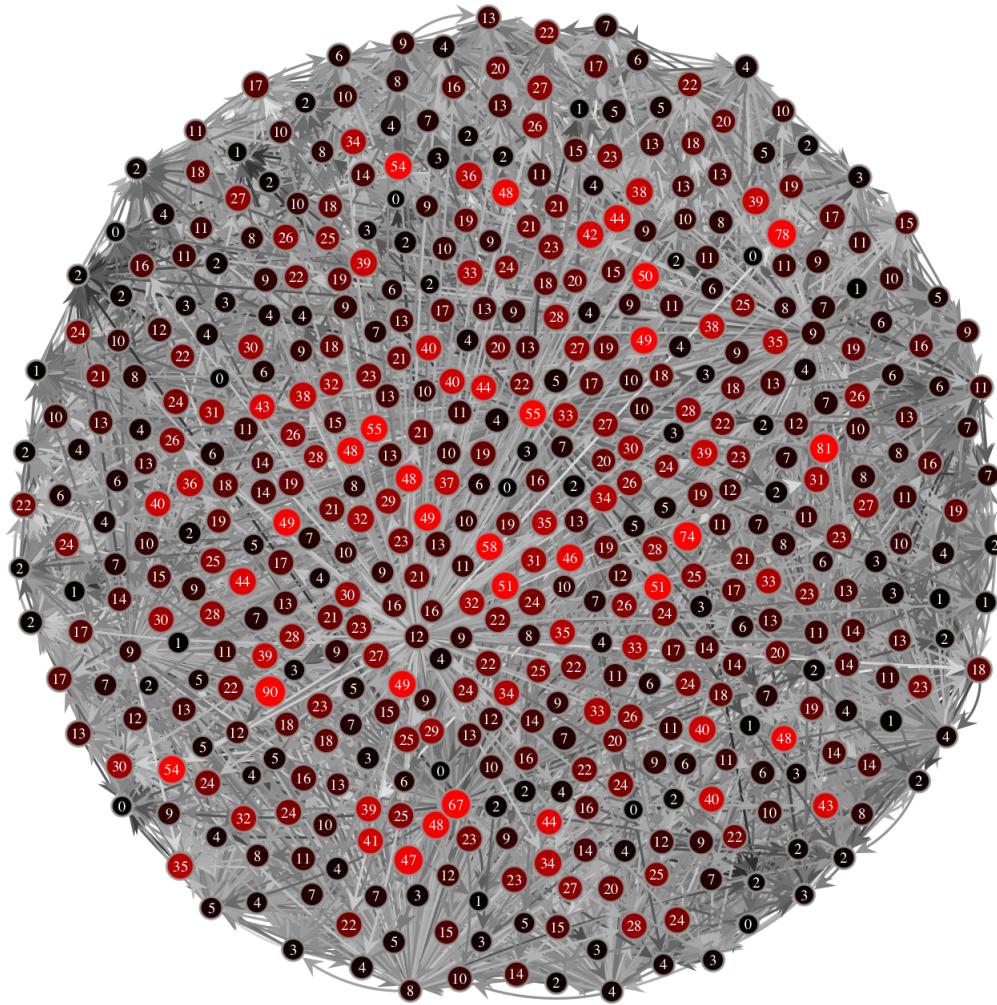


Figure 5.12: Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1200 memories. Node value indicates memory membership of the node.

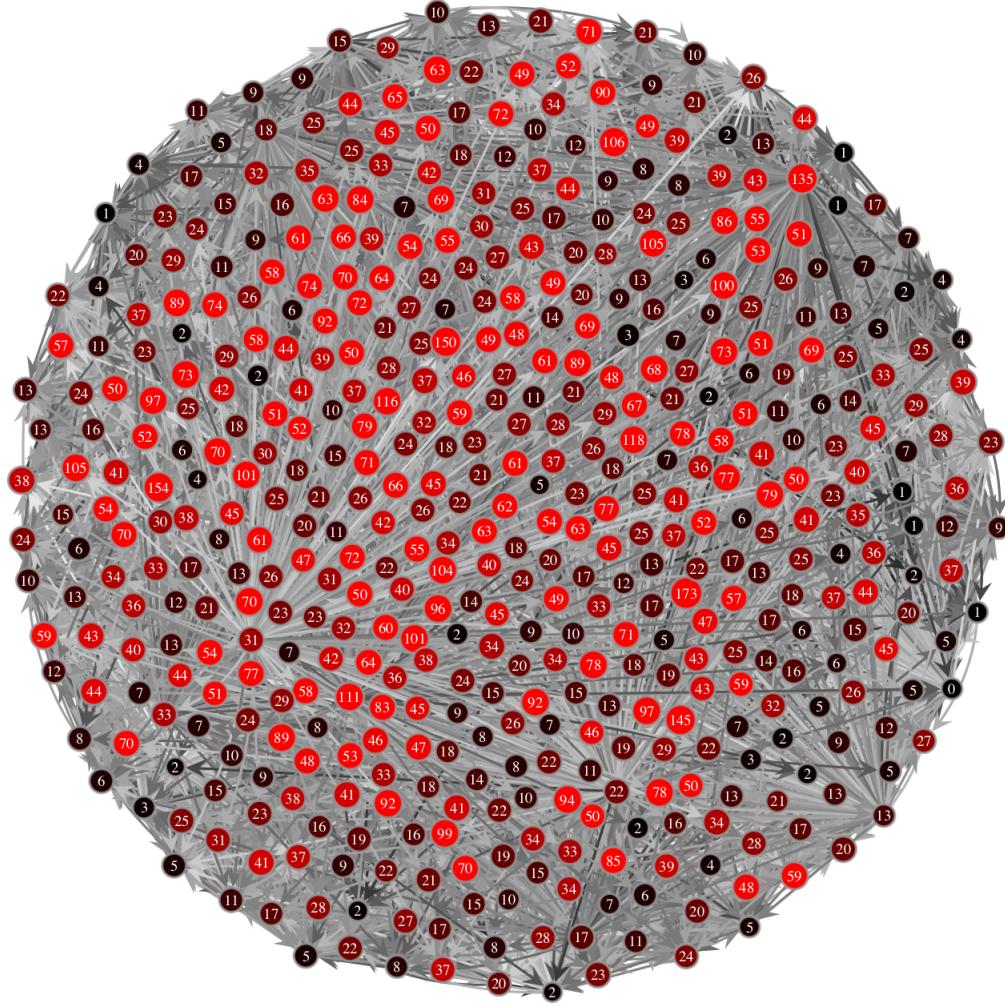


Figure 5.13: Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1400 memories. Node value indicates memory membership of the node.

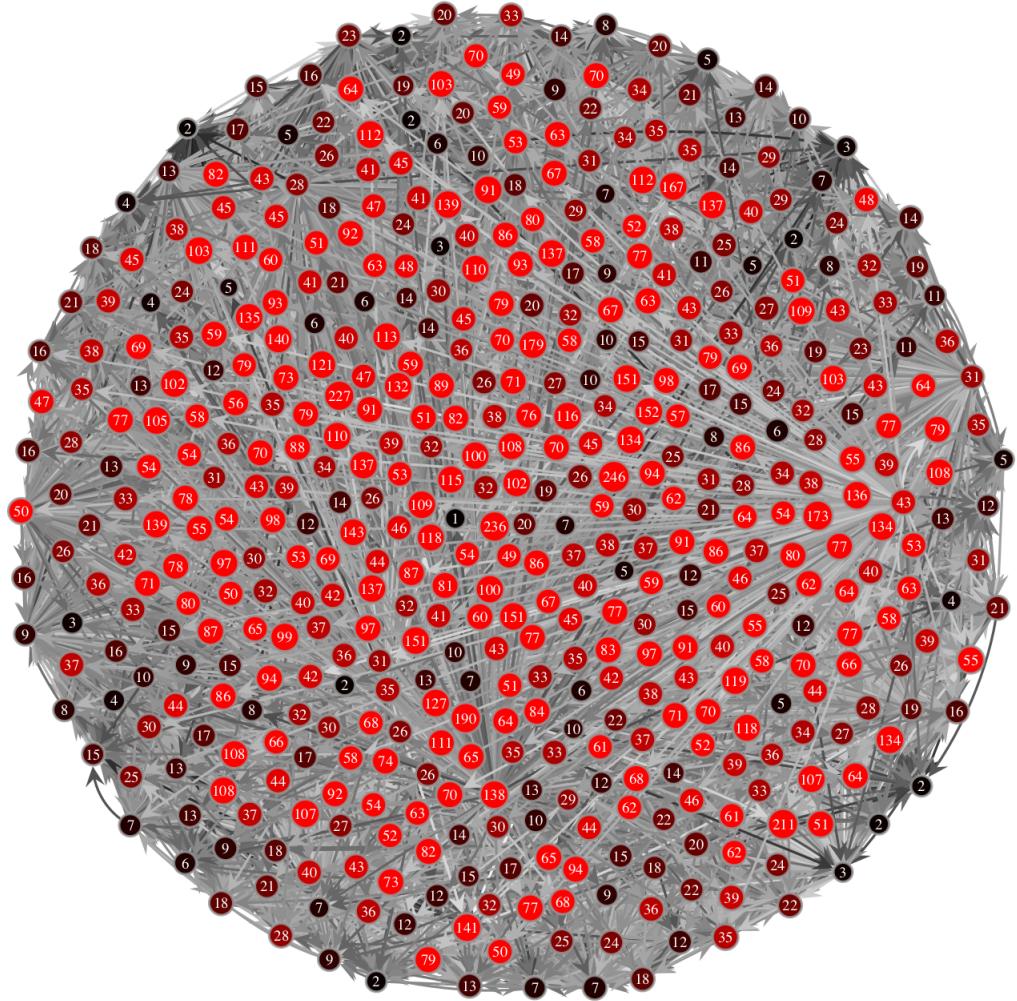


Figure 5.14: Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at 1600 memories. Node value indicates memory membership of the node.

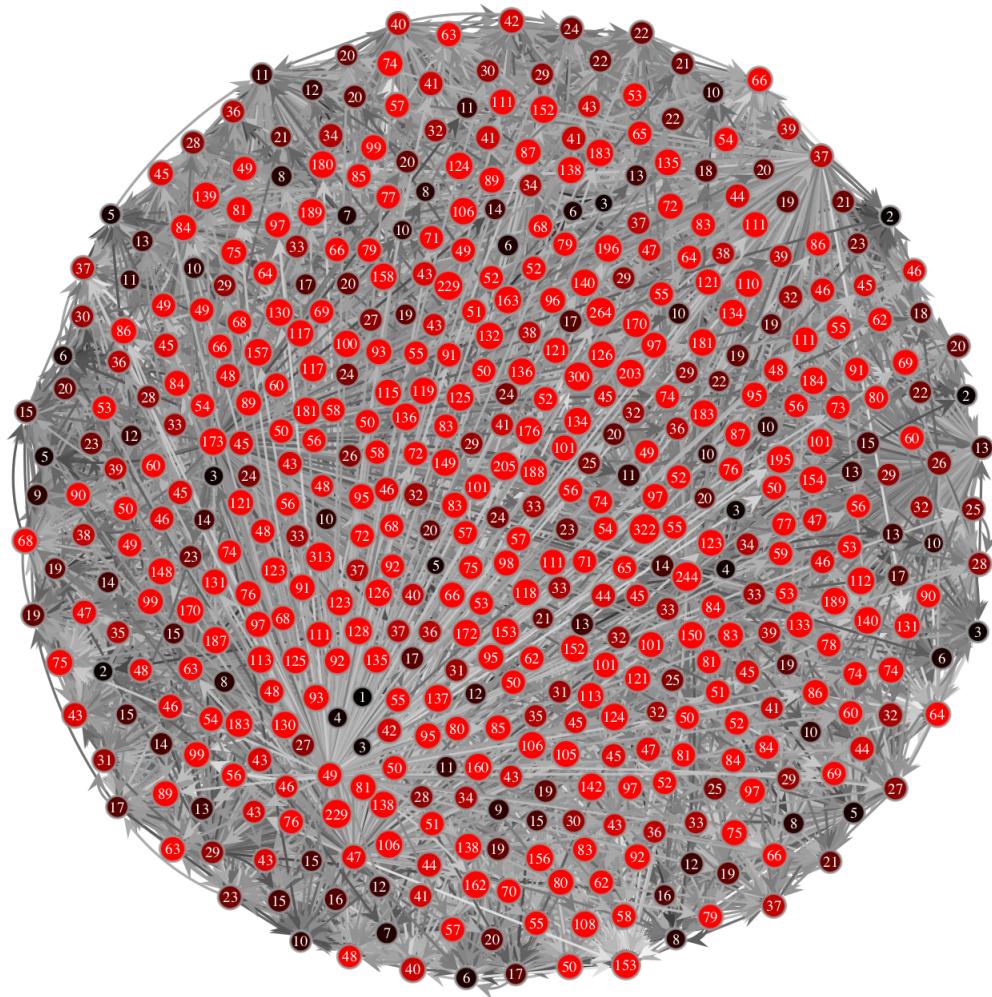


Figure 5.15: Neuroidal model ($n=500$, $d=128$, $w=16$, $r=40$) at capacity (1800 memories). Node value indicates memory membership of the node

indicating the number of interferences that occurred at that node and the number of memories the node belongs to respectively. The relative size of the node indicates the indegree of the node. We observe that the interference is not uniformly distributed across 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 contrast 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 3 to be compatible with JOIN. We believe this is really challenging and leave it as future work.

We also note that the number of memories a node is part of has no 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 is some correlation between the in-degree and the interference and could provide intuition regarding updating lemma 3.

Chapter 6

CONCLUSION

Inspired by advances in modern computational neuroscience, we rigorously defined and studied the notion of “capacity” and “interference” in a set both theoretically and empirically. 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 3 to find the expected interference in the case of other memory creation algorithms. The rest of the theorems will follow similarly to be able to find the capacity of the model.
- 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

We believe this study will inspire more computational neuroscientists to tackle the intriguing question of capacity as they develop models that slowly and steadily demistify the human brain.

BIBLIOGRAPHY

- [1] j4GGy (<https://math.stackexchange.com/users/186141/j4ggy>). Proof for the upper bound and lower bound for binomial coefficients. Mathematics Stack Exchange. URL:<https://math.stackexchange.com/q/1352350> (version: 2015-07-07).
- [2] 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.
- [3] P. R. Perrine. Neural Tabula Rasa: Foundations for Realistic Memories and Learning. Master’s thesis, California Polytechnic State University, San Luis Obispo, June 2023.
- [4] L. G. Valiant. Memorization and Association on a Realistic Neural Model. Neural Computation, 17(3):527–555, 2005.
- [5] 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.