

DARPA Y1 (Q1-Q4) Task 2 Supplement report

Graph/Network Structures and Processes on These Structures

1. SPIKING PROCESSES ON A NEURAL NETWORK WITH EXCITATORY AND INHIBITORY NODES

1.1. Introductory Remarks. Previously we introduced a neural network $N(E, I)$. The excitatory layer of the network E has nodes and connections that are defined by the vertices and edges of the graph $G_{\mathbb{Z}_N^2, p_d}$ respectively. The inhibitory layer I has $\frac{N^2}{4}$ nodes that are connected in an all to all fashion and can be thought of as a fully connected graph $K_{\frac{N^2}{4}}$ whose vertices are the nodes. Moreover, each inhibitory node is connected to four excitatory nodes at random in such a way that no two inhibitory nodes share any excitatory neighbors. Due to the convenient parallel with graph theory we occasionally borrow notation, and so by $V(E)$, $V(I)$ we mean the set of excitatory and inhibitory nodes, etc.

Each node in both E and I can take on one of two states: active or inactive. Let $\chi_v(t)$ define the potential function for node v in either layer at time t such that $\chi_v(t) = 1$ if v is active and $\chi_v(t) = 0$ if v is inactive at time t . The state of a node is completely determined at every time step by the state of its neighbors. To define this more formally, let $A^E(t)$ denote the set of active vertices in E at time t and similarly $A^I(t)$ denote the set of active inhibitors at time t . Furthermore, define $A^E(0)$ as a random subset of excitatory nodes that became active with probability p independently of all others and $A^I(0) = \emptyset$. Then for a vertex in E we say its state at time $t + 1$ is

$$\chi_v(t + 1) = \mathbb{1} \left(\sum_{u \in N(v) \cap V(E)} \chi_u(t) \geq k \right)$$

Similarly, for a vertex in I we have

$$\chi_v(t + 1) = \mathbb{1} \left(\sum_{u \in N(v) \cap V(E)} \chi_u(t) \geq \ell \right)$$

In both cases $\mathbb{1}$ is the indicator function and $N(v)$ denotes the subset of nodes in the closed neighborhood of v (i.e the node v and its neighbors). Both k and ℓ are nonnegative integers that specify the number of active neighbors any given vertex needs to become active on the next time step in E and I respectively.

1.2. New basic results on criticality over random graphs and their mean field models. Our new results (obtained after the March report) on limiting critical percolation values are indicated by bold entries in the following table. The entries of this Table are defined similarly to the ones of Table 1 in our previous report. The rather technical proofs to be published in a mathematical journal are omitted.

Table II
Critical Probabilty in Random Graphs

Neighborhood size (k)	RM	RM	MF	MF
	$\lambda = 0$	$\lambda > 0$	$\lambda = 0$	$\lambda > 0$
0	0	0	0	0
1	0	0	0	0
2	0	0	~ 0.131	$x_2(\lambda)$
3	1	1	0.5	$x_3(\lambda)$
4	1	1	$\sim 1 - 0.131$	$x_4(\lambda)$
5	1	1	1	$\begin{cases} \mathbf{1} & \text{for } \lambda \leq \ln(5) \\ \mathbf{x}_5(\lambda) & \text{for } \lambda > \ln(5) \end{cases}$

1.3. A fire together process. We conclude our review of $N(E, I)$ by presenting the inhibitory firing function we had introduced that causes all of the inhibitors to fire together once $m \in [0, \frac{N^2}{4}]$ inhibitory vertices are active during a time step. In other words an inhibitory node $v \in V(I)$ fires at time $t + 1$ if

$$F_v(t + 1) = \mathbb{1} \left(\sum_{u \in N(v); u, v \in V(I)} \chi_u(t) \geq m \right)$$

but v did not fire at time t . Notice that active inhibitory nodes fire simultaneously since they are in all to all connection with each other. At the time of firing, the inhibitory node sets the activity of all excitatory nodes connected to it and itself to 0. That is to say in a firing step the following nodes get inactive: (i) all inhibitory nodes, and (ii) those excitatory nodes which were connected to an active inhibitory node that was firing at that step. After the inhibitory firing occurs both layers carry on by propagating activity (or the lack thereof) with whichever excitatory nodes were left in tact.

Before this network architecture, we had also shown that *assuming there is no inhibition*, for any $\lambda \geq 0$ with the $k = 2$ activation rule, E has an asymptotic critical probability $p_c = 0$. Therefore there always exists a system large enough such that for any choice of $p_{in} > 0$, regardless of how arbitrarily small, with high probability (whp) all of the vertices in E will eventually become active when the inhibitors are not allowed to fire. When we introduce the inhibitory firing rule, with a proper choice of parameters, periodic behaviour became a possibility. By choosing parameters in such a way that the firing process extinguishes vertices while maintaining $\frac{|A^E|}{N^2} > p_c$ in the specified network, the activation in both layers likely begins growing once more until the next firing. For the sake of specificity

Definition 1. Let Δt be a lapse in time such that there exists finite t where $|A^I(t)| = 0$, $|A^I(t + \Delta t)| = 0$, and $\nexists \Delta t' \in [0, \Delta t)$ such that $|A^I(t + \Delta t')| = 0$. We will call the sequence $\{A^E(t), A^E(t + 1), \dots, A^E(t + \Delta t)\}$ a **spike** or a **cycle** of length Δt .

As we will highlight this indicates oscillatory behaviour in this dynamical system for a certain, wide range of parameters.

2. DISCUSSION OF THE SPIKING PROCESS OVER $N(E, I)$

To precisely uncover where this oscillatory regime lives in the parameter space, we begin with an analysis of what it means for the system to fire. This reveals strict bounds between the excitatory activation density and the system's potential to fire.

Proposition 1. *Given a network $N(E, I)$, $\ell = 1, \dots, 4$, and $m \in [0, \frac{N^2}{4}]$.*

- *If $|A^E(t)| > N^2 - (5 - \ell)(\frac{N^2}{4} - m + 1)$ then the system will necessarily fire*
- *If $\ell m \leq |A^E(t)| \leq N^2 - (5 - \ell)(\frac{N^2}{4} - m + 1)$ then the system has the potential to fire*
- *and if $|A^E(t)| < \ell m$ then the system will not fire.*

Proof. The lower bound, ℓm , follows by definition. To see the upper bound we first note that the system will fire regardless of whether inhibitors are active from exactly ℓ neighbors or more. It follows that if every active inhibitor had all four excitatory neighbors active and every inactive inhibitor had $(\ell - 1)$ active excitatory neighbors (i.e the maximum number of active excitatory neighbors possible while the inhibitor remains inactive) then this is the maximal $|A^E(t)|$ possible for each value of $|A^I(t)|$. We know the system will fire when $|A^I(t)| \geq m$. As such $|A^I(t)| = m - 1$ is the maximum number of active inhibitors allowable without the system firing. By the above, if $|A^I(t)| = m - 1$ then $|A^E(t)| \leq N^2 - (4 - x) \left(\frac{N^2}{4} - (m - 1) \right)$ where $4 - x = \ell - 1$. A little algebra reveals to us that in such a case $|A^E(t)| \leq N^2 - (5 - \ell) \left(\frac{N^2}{4} - m + 1 \right)$. Clearly if $|A^I(t)| < m - 1$ then $|A^E(t)|$ can never achieve the upper bound we just found and the inequality becomes strict. Combining these observations and taking their contrapositive unravels the fact that if $|A^E(t)| > N^2 - (5 - \ell) \left(\frac{N^2}{4} - m + 1 \right)$ then $|A^I(t)| > m - 1$ and so the system necessarily fires.

Thus we have seen that by definition if $|A^E(t)| < \ell m$ it can never fire and if $|A^E(t)| > N^2 - (5 - \ell) \left(\frac{N^2}{4} - m + 1 \right)$ then the system will surely fire. \square

We should note that the above proposition gives us a sufficient condition for firing regardless of λ , but as is made clear in Figures 1 and 2 this barely begins to address the phenomena that arise from this spiking process. In the discussion that follows one should assume the case when $k = 2$ unless explicitly stated otherwise.

All of the data we've produced in the spirit of what formed Figures 1 and 2 display variations of the same interesting behaviour: sharp horizontal lines that split cyclic and acyclic behaviour, a sharp vertical line that corresponds to m where cycles cannot form, seeming independence of cycle length from p_{in} , and fractal like acyclic slivers.

We refer to the horizontal line under which no cycles remained in Figure 1 as $p_c^{low}(\lambda, \ell, N)$ and the horizontal line over which no cycles remained as $p_c^{high}(\lambda, \ell, N)$.

Proposition 2. *For every $\lambda \geq 0$, $\ell = 1, \dots, 4$ and positive integer N , there exist lower and upper critical probabilities, $p_c^{low}(\lambda, \ell, N)$ and $p_c^{high}(\lambda, \ell, N)$ respectively, such that for any initial activation $p_{in} < p_c^{low}$ or $p_{in} > p_c^{high}$ the system cannot exhibit oscillatory dynamics. Furthermore, for any initial activation $p_c^{low} \leq p_{in} \leq p_c^{high}$ the system may have oscillatory dynamics.*

Proof. To show the existence of $p_c^{low}(\lambda, \ell, N)$ we begin by noting that between inhibitory firings our network $N(E, I)$ reduces to an instance of the non-monotonic percolation model we have studied in the past since, when considered on its own, E is precisely that. As we have previously shown, when $k = 2$ in such models there exists a critical probability p_c where the system percolates whp when it is initialized above p_c and when it is initialized below p_c whp either gets frozen at low activation

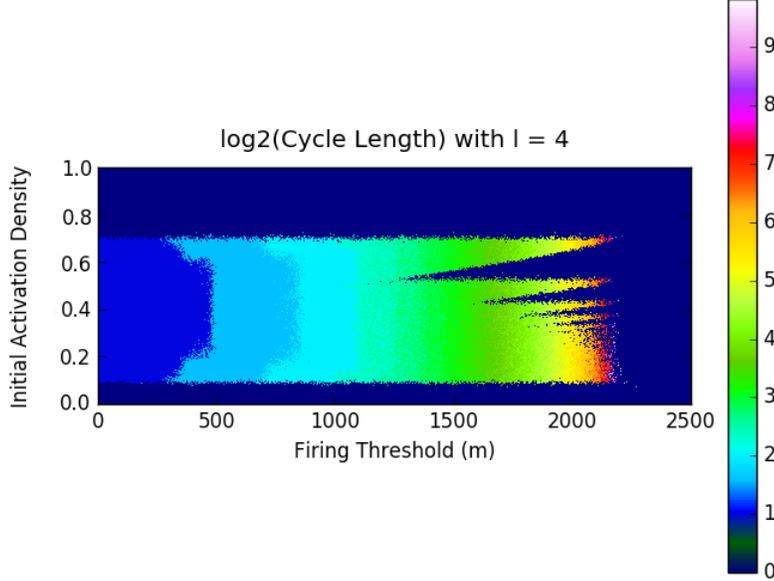


FIGURE 1. A fine grained heat map of $\log_2(\text{cycle length})$ after 1000 time steps when $k = 2$, $\ell = 4$, $\lambda = 0$, and $|V(E)| = 10,000$. We ran computer simulations for each value of $m \in [0, 2500]$ and every p_{in} from 0 to 1 in increments of 0.001.

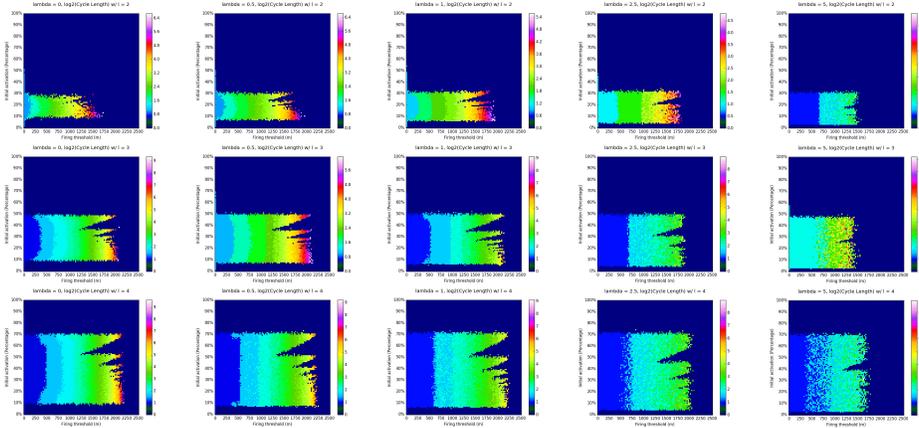


FIGURE 2. Heat maps much like 1, but $\lambda = 0, 0.5, 1, 2.5, 5$ from left to right and $\ell = 2, 3, 4$ from top to bottom.

density with small active clusters or dies out completely. It follows that when $\frac{\ell m}{N^2}$ in $N(E, I)$ is greater than p_c in an equally sized percolation model, by Proposition

1 the inhibitors cannot fire and so the existence of p_c in non-monotonic percolation is exactly equivalent to $p_c^{low}(\lambda, \ell, N)$. In the case when $\frac{\ell m}{N^2} \leq p_c$ the system either immediately fires, our network needs to build back up and behaves like percolation, but clearly is below p_c and whp won't grow or the system does not immediately fire and is equivalent to percolation, but is below p_c and so it doesn't percolate whp. In all cases we have seen then that when $p_{in} < p_c$, whp, the system will fire at most once and thus $p_c^{low}(\lambda, \ell, N)$ exists.

With this in mind we consider $p_c^{high}(\lambda, \ell, N)$. Arguing by contradiction, suppose that there doesn't exist $p_c^{high}(\lambda, \ell, N)$. This implies that regardless of what value of p_{in} we begin the system with it can whp begin building its activation density back up after an inhibitory firing and exhibit these oscillatory dynamics. Clearly we have a contradiction since if we choose p_{in} such that the remaining activation density following the first inhibitory firing is below $p_c^{low}(\lambda, \ell, N)$ the system will not be able to create another cycle whp and so $p_c^{high}(\lambda, \ell, N)$ exists. \square

Two things are now worth noting, first the claim in the above proposition that for any initial activation $p_c^{low} \leq p_{in} \leq p_c^{high}$ the system may have oscillatory dynamics follows from the negation of what we have just proven. Secondly, Proposition 2 needed to account for asymptotic network sizes where p_c^{low} approaches 0 and so of course p_c^{high} in turn must approach 1. In practice with finite network sizes p_c^{high} is sizably less than 1 as is evidenced by Figures 1 and 2 where $N = 100$.

Remark 1. *The following are some properties of $p_c^{low}(\lambda, \ell, N)$:*

- For fixed λ and N , $p_c^{low}(\lambda, \ell, N)$ does not depend on ℓ
- For fixed N and ℓ , $p_c^{low}(\lambda, \ell, N)$ is decreasing in λ

Proof. Consider a case where $p_{in} < p_c^{low}(\lambda, \ell, N)$ and suppose we have p_c of an equivalent percolation model of the same size. When $\ell m > p_c$ then by Proposition 1, since E behaves like percolation when there isn't inhibitory firing, $p_c^{low}(\lambda, \ell, N)$ is not dependant of ℓ . When $\ell m \leq p_c$ then we might still experience an inhibitory firing. If the system doesn't fire, then p_c is still equivalent to $p_c^{low}(\lambda, \ell, N)$. If the system does fire, then the activation density has only decreased and E still will likely not be able to increase its activation density because that density is below p_c . Therefore regardless of ℓ , $p_c^{low}(\lambda, \ell, N)$ will be the same given λ and N .

The second statement follows from the observation that as λ increases, the expected degree of a vertex does so too. Therefore the average number of active vertices adjacent to any given vertex is increasing. It follows that any vertex, and as a result the whole system as well, may get active with higher probability. Thus as λ increases, $p_c^{low}(\lambda, \ell, N)$ decreases since E is able to percolate more easily. \square

We can see these properties in Figure 2, where if we focus on a single column (i.e fixed λ and N) we can see that p_c^{low} remains fixed and, if we focus on a single row (i.e fixed ℓ and N), p_c^{low} decreases from left to right as λ increases. For an even clearer view of the decrease in p_c^{low} with λ , consider Figure 3 where we have put side by side a heatmap generated from $\lambda = 0$ and $\lambda = 10$ both with $N = 100$ and $\ell = 4$.

Remark 2. *The following are some properties of $p_c^{high}(\lambda, \ell, N)$:*

- For fixed λ and N , $p_c^{high}(\lambda, \ell, N)$ is increasing in ℓ
- For fixed N and ℓ , $p_c^{high}(\lambda, \ell, N)$ never increases in λ

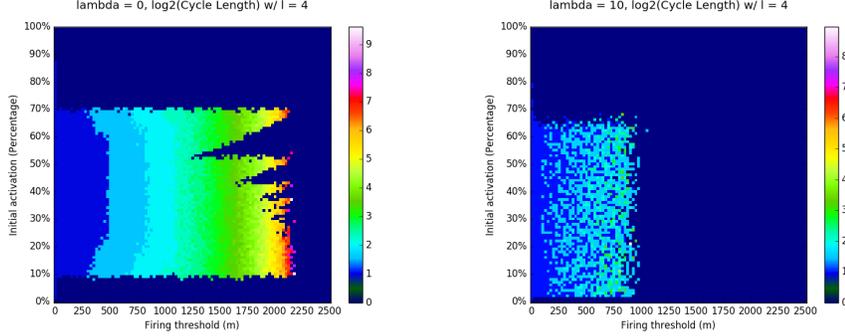


FIGURE 3. Heatmaps as we have seen above. We plot $\log_2(\text{cycle length})$ from data generated using $\lambda = 0$ on the left and $\lambda = 10$ on the right.

Proof. We saw above that when $p_{in} > p_c^{high}$ and the inhibitors fire, the excitatory activation density plummets below p_c^{low} and is unlikely to be able to pick itself back up. When we increase ℓ the affect that this has on the system is to make it harder for inhibitors to activate by definition. It follows that, when all else is the same, at a majority of the possible activation configurations of nodes in E fewer inhibitors will be active at higher values of ℓ relative to low values of ℓ and so $p_c^{high}(\lambda, \ell, N)$ increases as ℓ does.

For us to see the impossibility of increasing in p_c^{high} with increase in λ , given N and ℓ , consider networks $N(E, I)$ and $N(\bar{E}, \bar{I})$ with $\lambda = 0$ and $\lambda > 0$ respectively and all else the same. As we increase λ more excitatory vertices, and as a result more inhibitory vertices, are able to become active during a time step. It follows that at a time t if $A^{\bar{E}}(t) \equiv A^E(t)$ and $A^{\bar{I}}(t) \equiv A^I(t)$ we will necessarily see $|A^{\bar{E}}(t+1)| \geq |A^E(t+1)|$ and $|A^{\bar{I}}(t+1)| \geq |A^I(t+1)|$ before we account for any effects brought on by an inhibitory firing. As λ grows however, so do the number of configurations of active excitatory vertices that result in $|A^{\bar{E}}(t+1)| > |A^E(t+1)|$ and $|A^{\bar{I}}(t+1)| > |A^I(t+1)|$. We have seen that when $p_{in} > p_c^{high}$ the excitatory layer whp will experience a short lived growth¹ before being over extinguished from an inhibitory firing and falling below p_c^{low} . It follows that, for $\lambda > 0$, $p_{in} > p_c^{high}(0, \ell, N) \geq p_c^{high}(\lambda, \ell, N)$ and that for sufficiently large λ we will likely have $p_c^{high}(0, \ell, N) > p_c^{high}(\lambda, \ell, N)$. \square

Similarly to the previous remark, Figure 2 demonstrates both properties of p_c^{high} . Note the large jumps in p_c^{high} between values of ℓ (i.e rows). The second property we mentioned about p_c^{high} , the effect that λ has given ℓ and N , are there but much harder too see without having big jumps in λ side by side; we see precisely this sharp contrast in Figure 3.

We now turn our attention to what we previously referred to as the seeming independence of cycle length from p_{in} . Intuitively speaking this phenomenon arises because, regardless of how densely the system is initialized, if we have chosen parameters that allow for oscillatory dynamics then once the inhibitors fire and

¹This may be exactly the first time step after initialization or a tight rope walk of excitatory activation for a few time steps where the system is at the cusp of a debilitating firing.

activity needs to build again the network “forgets” the state it was in prior to firing and behaves similarly to a system initialized at its new $A^E(t)$. It is for this reason that we call this phenomenon “memorylessness”. Let $p^{inhib} = \frac{4m}{N^2}$ represent the density of active inhibitory vertices in I needed for the system to system to fire.

Proposition 3. (*Memoryless behaviour.*) *For every fixed $\lambda \geq 0$, $\ell = 1, \dots, 4$ and positive integer N , the cycle length $c(\lambda, N, \ell, p^{inhib})$ in the oscillatory region of parameter space depends on p^{inhib} but does not depend on initialization probability p_{in} .*

Proof. The value of p^{inhib} controls the reset; we see this very clearly in all of the heatmaps above. Notice that as we increase its value the cycles in the oscillatory region get longer. This follows from the fact that the higher p^{inhib} is, the higher the number of excitatory vertices that get deactivated are. Thus, the overall density decreases more sharply with inhibitory firing as we increase p^{inhib} . Clearly the lower density of active excitatory vertices that the process continues from, the more activation steps are needed to reach p^{inhib} again. In other words, as p^{inhib} increases, so too does the length of our cycles. It follows that we achieve maximum cycle length when p^{inhib} is chosen in such a way that the density of active vertices the process continues with after the reset is just slightly above $\sim p_c^{low}(\lambda, \ell, N)$. If we increase p^{inhib} too much then the density of active vertices the process continues with after inhibitory firing will eventually fall below $p_c^{low}(\lambda, \ell, N)$ and the system dies out whp.

When we choose values of p_{in} such that $p_c^{low} \leq p_{in} \leq p_c^{high}$, we only effect the length of the first cycle. After the first inhibitory firing occurs the system will continue the activation process with a density of active excitatory vertices that is determined by the reset parameter p^{inhib} , regardless of p_{in} . \square

This behaviour can be clearly seen in Figures 1 and 2. The phenomenon of cycle length $c(\lambda, N, \ell, p^{inhib})$ depending on p^{inhib} but not p_{in} , for fixed λ and N , manifests itself here as the vertical stripes in both figures. Also notice that the “very long” cycle regime that lies on the right border of the oscillatory region undergoes a sharp transition to acyclic behaviour.

At this point we should note the important (but subtle) distinction about p_{in} having no sway on the cycle length $c(\lambda, N, \ell, p^{inhib})$ inside of the oscillatory region of parameter space. It is clear that, for sufficiently high values of p^{inhib} due to the fractal like “bays” that cut into regions of parameter space that allow for oscillatory dynamics, p_{in} plays a vital role as to whether we are in the oscillatory region or not. By considering the bounds derived in Proposition 1, we can begin to draw the line that divides values of p^{inhib} for which we have long orbits and none at all.

Remark 3. *For every fixed λ , ℓ , and N the network $N(E, I)$ cannot display oscillatory dynamics for m such that $\ell m > p_c^{high}(\lambda, \ell, N)$.*

Proof. Suppose that $N(E, I)$ is as we have stated. By Proposition 1 we know that the system will never fire for $\frac{|A^E(t)|}{N^2} < \frac{\ell m}{N^2}$. Therefore the system necessarily must have activation density above $p_c^{high}(\lambda, \ell, N)$ before it ever fires. As we have seen, its activity will fall below $p_c^{low}(\lambda, \ell, N)$ when it eventually fires and not be able to pick itself back up. \square

A quick look at Figures 1, 2, and 3 suggests that once we have determined a maximal m that can produce oscillations in a given network the value $\frac{\ell m}{N^2}$ can be a rather good heuristic value for p_c^{high} and vice versa.

3. COMPUTER SIMULATIONS OF THE SPIKING PROCESS

3.1. Introductory Remarks. The current implementation of generalized bootstrap percolation is done in python and was designed to offer as much flexibility as possible so as to allow for unexpected research directions that may arise. As of now the model allows us to specify the number of excitatory vertices in the lattice N^2 , the number of excitatory vertices that should be mapped to any inhibitory vertex², the density of excitatory vertices we would like active in the initial time step p_{in} , the parameter λ that influences the probability of adding a random edge between excitatory vertices, the constant in the activation rule for excitatory vertices k , the constant in the activation rule for inhibitory vertices l , and the constant in the firing rule for the inhibitory vertices m . It additionally offers the option of excluding some of these parameters by letting us specify whether we want to add random edges between excitatory vertices, whether we want inhibitory vertices and, if so, whether to map inhibitory vertices to the excitatory vertices at random or in some predefined way.

In the discussion that follows, the reader should assume $N = 100$, $k = 2$, and $\ell = 4$ unless it is explicitly stated otherwise.

3.2. Findings from Computer Simulations on Variations of λ . Data gathered from simulations as described have matched all of the propositions highlighted in the mathematical analysis above. However, multiple observations remain open problems and a rigorous understanding eludes us. This is in part due to the complicated problem of describing the evolution of $A(t)$ for any value of λ . In what follows we describe some of the facets of the spiking process that remain to be rigorously understood and provide our findings and intuitions as to why they arise.

In Proposition 3 it was shown that the length of the cycles formed in $N(E, I)$ is independent of p_{in} and in Remark 4 we give a rough bound on m where cycles can form. Looking at any of the heatmaps above, this leaves unaddressed the obvious acyclic fractal like bays that jut into the oscillatory region of our parameter space and seem to become larger and more pronounced as we increase m and p_{in} .

The presence of these bays are intuitively consolable once we understand the general behaviour of these cellular systems. First we note that it makes sense that for small t we can treat a system with activation density $\frac{|A^E(t)|}{N^2}$ similarly, but not exactly, like a system that is initialized with $p_{in} = \frac{|A^E(t)|}{N^2}$ and is otherwise identical. It follows from this that a system initialized inside of the biggest, top-most bay in Figure 1 may have enough active excitatory vertices to immediately jump up above a density equal to p_c^{high} which as we have seen will likely result in a firing of too many inhibitory nodes and the system's density plummets below p_c^{low} . By inductively applying this intuition, it seems plausible that more bays appear at smaller initialization densities where the system jumps into a bigger bay.

Our simulations of systems initialized inside of different bays supports this intuition because in every instance of the experiment we have noticed the activation

²Note that this also dictates that we have N^2 divided said number of inhibitory vertices in total.

jumps from its initial bay, to a larger bay, until it eventually jumps above p_c^{high} , and causes a massive inhibitory firing. It is important to note that we do not necessarily observe a neat skipping from bay to bay in perfect increasing order, we often observe jumps from a small bay over multiple bays and into a much larger one.

As we can observe in Figure 2, the number of bays and how pronounced they are changes with λ . Exploring data that resulted in Figure 2, but for many more values of λ , provides intriguing insights into not only these effects on the behaviour and presence of bays, but also the effects λ has on multiple facets of the system such as the lengths of possible cycles and the presence of cycles at higher values of m .

Figure 4 in conjunction with 2 makes clear the heavy influence λ has on the bays. We notice that as we increase $\lambda \in [0, 1]$ the bays seem to become gradually more frequent, as we increase $\lambda \in [2.5, \infty)$ the bays become less frequent until they disappear outright. By the same intuition used to explain the presence of bays, this behaviour is expected. For small values of λ the system is able to jump into the larger bay from more initializations in the parameter space. For large values of λ each active excitatory vertex has a much larger influence over activation in the next time step, this forces the system to take massive leaps in activation density and so initializations immediately jump above p_c^{high} . As we see in Figure 4, by $\lambda = 10$ there isn't a single bay left. Furthermore, as we will see in Figure 6, the bays at different values of λ play functionally the same role and, when increasing λ is thought of as a chronological process, these bays are all in fact a gradual transformation of one another.

A great benefit of looking at these heat maps is that they make obvious when cycles of different lengths move to new locations in the parameter space. As we steadily increase λ it seems that progressively fewer initialization parameters result in extremely large cycles, but that the length of the largest observed cycle remains relatively unperturbed. This can be explained by the same forces that affect the bays as we vary λ ; in general it takes fewer steps for the activation density to climb back up to the inhibitory firing threshold. Perhaps more perplexing is the observation that as λ increases it seems that the emergence of small cycles of length ≈ 4 appear in progressively more sets of initialization parameters. We can see this rather clearly in Figure 4.

Changes in λ also have an effect on the presence of cycles at high values of m . This effect is clear when we look at large lapses in the value of λ , but are also present in subtle changes in λ . We have observed this for $\ell = 2, \dots, 4$, but demonstrate it with $\ell = 4$ in Figure 5 because it is easiest to see.

In Figure 5 we plotted side by side the highest value of m that display cycles, at different values of p_{in} , at multiple values of λ . We easily see that from $\lambda = 0$ to $\lambda = 1$ the ability to form cycles extended towards higher values of m . Then, from $\lambda = 1$ to $\lambda = 2.5$ the ability to form cycles recedes behind even $\lambda = 0$ and that trend continues. Plots like this one but with values of λ^3 that lay between any two outlines shown in Figure 5 further support this, but are cluttered and harder to read. Though, still difficult to see without being able to interact with the plot, in Figure 6 we plot on a 3D surface the outline from data on every value of λ that we have. This demonstrates the gradual receding in cycle formation with respect to m quite nicely.

³We have made these plots for $\lambda = 0.01, 0.025, \dots, 0.1, \dots$, etc.

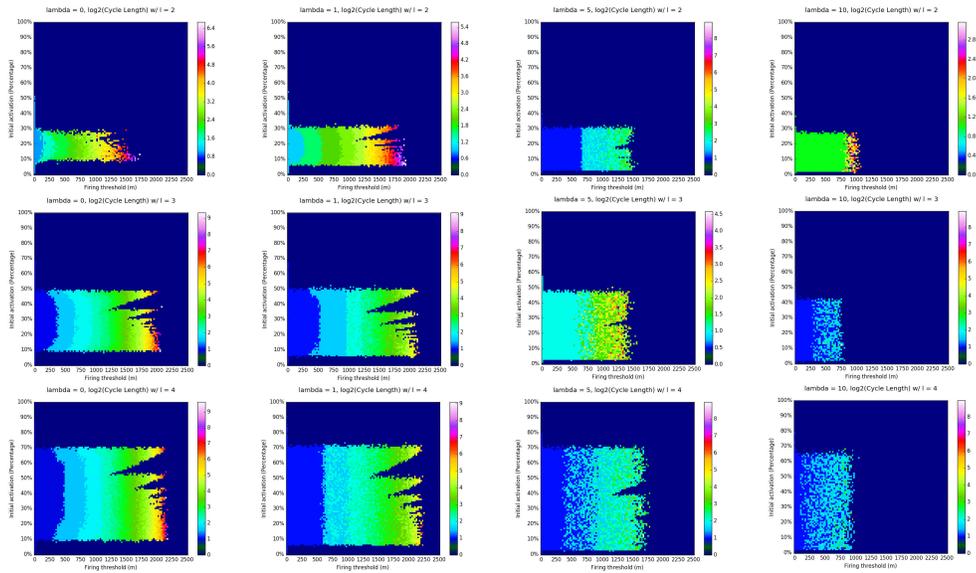


FIGURE 4. Heatmaps arranged like in Figure 2, but with $\lambda = 0, 1, 5, 10$ from left to right.

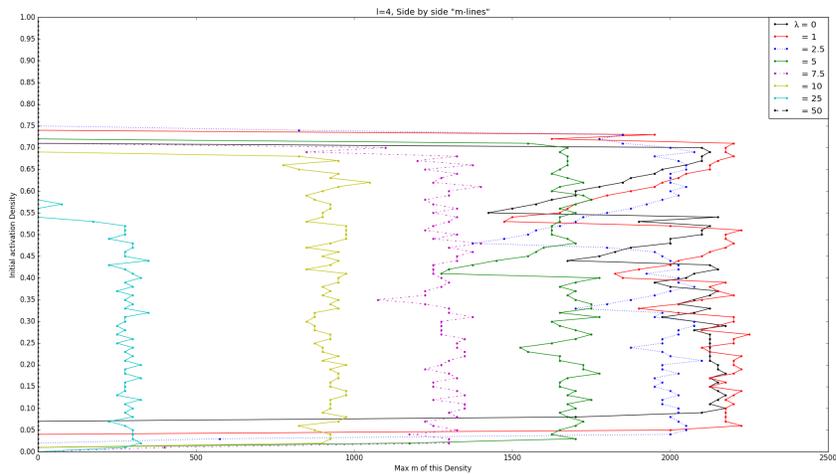


FIGURE 5. An “outline” of the values of m that can produce cycles. These are made by plotting the highest recorded m containing a spike at each p_{in} .

4. OSCILLATORY REGIMES AND CHAOS IN $N(E, I)$

We begin this discussion with a disclaimer that strictly speaking, chaotic regimes are impossible in any finite sized network. This however doesn't have any negative

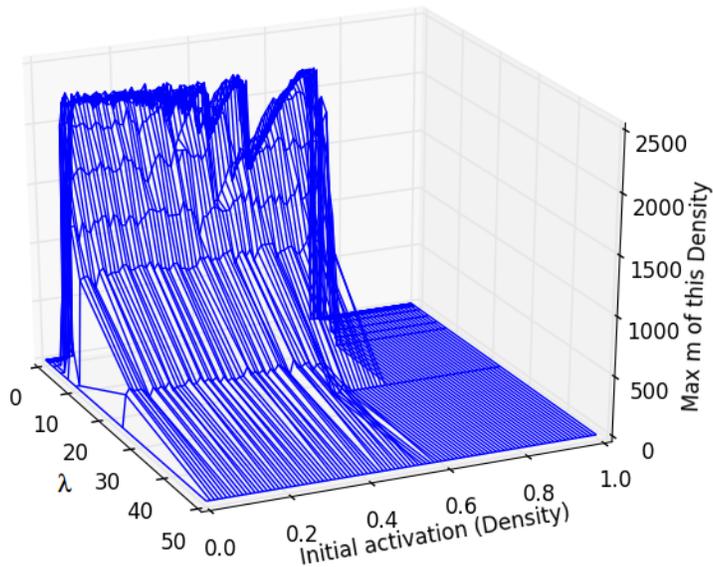


FIGURE 6. The outline of cycle forming regions on a 3D surface. Contains the outlines of numerous lambda between 0 and 50.

connotations for our use of these networks because, as will be discussed below, the regimes we call “chaotic” would take so much time to repeat themselves that, by all practical standards they are chaotic and behave chaotically.

This calls for a more precise definition of what we are referring to when we describe a regime as chaotic. To begin, we note that our investigations into the presence and nature of oscillations in our system has revealed that when the neural network $N(E, I)$ is left to oscillate unperturbed it settles into perfectly periodic oscillations—after hundreds of thousands of times steps in certain cases—for a majority of the parameter space, but that in select regions exact patterns didn’t form after even 10,000,000 time steps. More carefully stated.

Definition 2. *Suppose there exists mutually unique spikes $\{S_1, S_2, \dots, S_n\}$ of length $\{\Delta t_1, \Delta t_2, \dots, \Delta t_n\}$ such that for some $t \in [0, \infty)$, $S_k = \{A^E(t + \sum_{j=1}^{k-1} \Delta t_j + y(\sum_{j=1}^n \Delta t_j)), \dots, A^E(t + \sum_{j=1}^k \Delta t_j + y(\sum_{j=1}^n \Delta t_j))\}$ for all integer $y \geq 0$ and $k \in [1, n]$. We say our system is **n-periodic** if there doesn’t exist a sequence of less than n spikes that satisfy this property.*

We found that for a majority of parameters n is finite and relatively small. Extending this further

Definition 3. *We say our system is **chaotic** if for some very large T , there does not exist a time $t \leq T$ for which our system is n -periodic.*

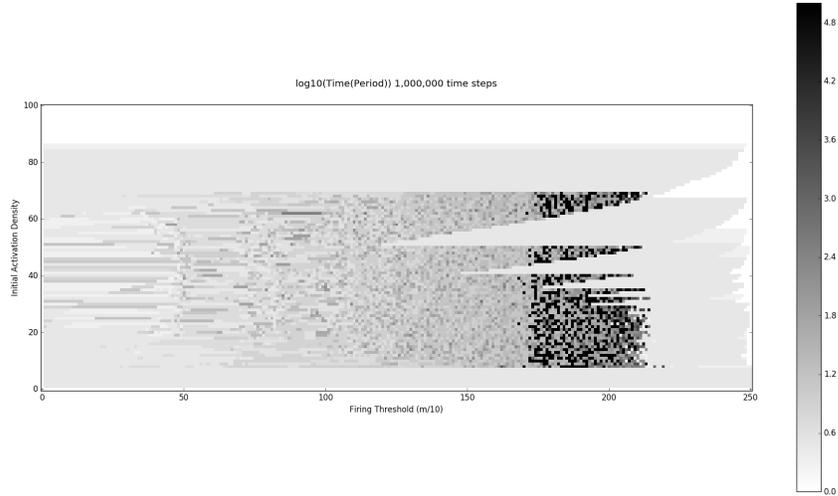


FIGURE 7. Heatmap measuring \log_{10} of the actual amount of time steps in an n -period after one million time steps when initialized at that initial density and m . The largest (darkest) values shown represent parameters where periods didn't form in 1,000,000 time steps, but we cut the values down appropriately (while remaining visibly different) so as to not make the information from actual periods difficult to see.

Figures 7 and 8 contains two heatmaps similar to what we have seen, but instead of helping us compare the raw length of spikes that form like above, in Figure 7 we are showing in each cell the \log_{10} of the actual time span an n -period takes and in Figure 8 we have in each cell the actual value of n itself. The former acts to “update” the heatmaps we have been looking at and gives us a gauge of just how long periods are with those parameters as opposed to just how long individual cycles are. The latter begins to give us a glimpse as to how spikes may bifurcate in the parameter space.

In both of these heatmaps, every experiment is generated using the same random seed. We checked for periods by comparing the excitatory layer in the last 10,000 time steps for exactly matching configurations of that layer. Since the system is deterministic, a matching configuration necessarily means a period. The plot in Figure 7, along with multiple other heuristics we've used, serves to support our choice in only checking the last 10,000 time steps for periods on the excitatory layer since other than “infinite” period length, the next largest period found was $\approx 5,000$ time steps long.

As we would expect from what we have seen above, in Figure 7 as m increases so do the time steps in each period in general. What is particularly striking is that there seems to be almost a sharp threshold at $m \approx 1700$ at which our system can produce chaotic oscillations. This is also confirmed in Figure 8 which additionally

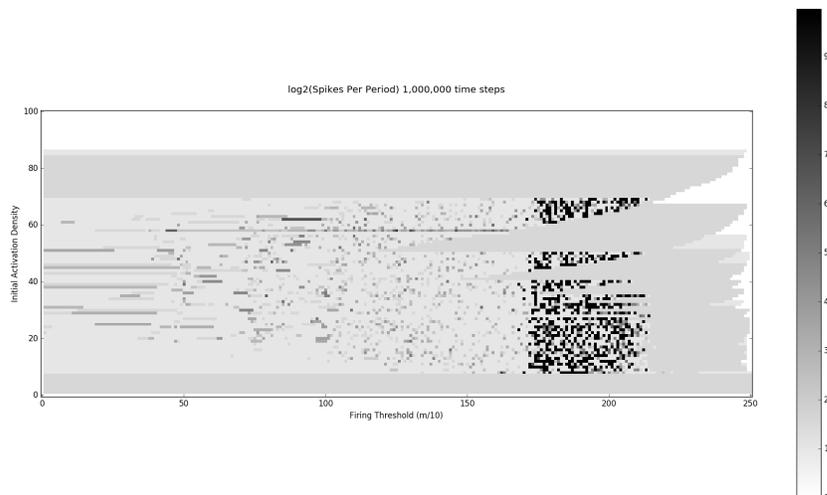


FIGURE 8. Heatmap measuring \log_2 of n itself under the same conditions. The largest (darkest) values shown represent parameters where periods didn't form in 1,000,000 time steps, but we cut the values down appropriately (while remaining visibly different) so as to not make the information from actual periods difficult to see.

informs us that in general, long periods do not mean there were a huge number of spikes inside of them. We are currently analyzing the raw numbers plotted in this to try and understand if there is some structure nested within.

Looking at these values more closely, we took our random seed and recorded time series for the tail end of 10,000,000 time steps of the density of active inhibitors at each time step when $p_0 = 0.5$ and for every value of $m \in [0, 2500]$. By plotting the last 1000 activation densities in this time series against each m we begin to etch out a bifurcation diagram for $N(E, I)$; we see this in Figure 9. This is to say, in Figure 9 that approximately when $m < 500$ the network is stuck at a single density, when m is between about 500 and 800 it goes from one density to another and then fires, etc.

To interpret Figure 9 it is worth noting that plotted against each m is every excitatory density visited by the network in the last 1000 time steps. Therefore, if there are multiple spikes in a given period, the configuration visited immediately following each individual spike will appear next to each other, but likely not in the exact same place. Due to the way the excitatory layer develops, this means that every step following this until the next spike will also be at roughly a similar density. This is why we see slight "fuzz" at m as low as 500. This is also why we see those thick/fuzzy bifurcation points that appear from m greater than roughly 1000. A better way of saying this is that this bifurcation diagram shows us, in some sense, the "spike doubling" rather than the period doubling and that when

we experience a bifurcation in the actual period (i.e the actual period doubles and not just the number of configurations visited by individual spikes) then the diagram here will become fuzzy since at each spike in a period the figure will plot gradually more points near one another. Summarizing this, we see in Figure 9 a gauge of how chaotic our network is by how unable we are to discern the steps taken in the excitatory layer.

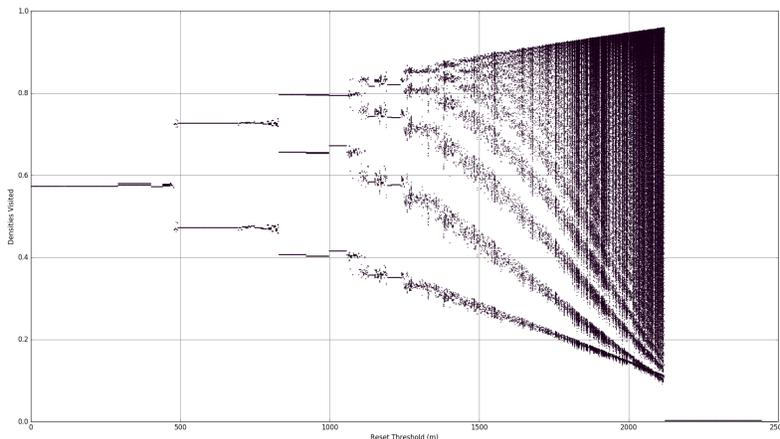


FIGURE 9. A kind of the bifurcation diagram for $N(E, I)$.

To more accurately see how chaotic a specific trajectory is, we embed these time series into two dimensions using a delay-time embedding. As we can see in Figure 10, periodic regimes have distinct closed trajectories and chaotic trajectories are such a tangled mess that we cannot discern one trajectory from the other.

By using metrics such as the delay embedding, or exhaustively checking the actual excitatory layer⁴, we are able to get an accurate count of exactly how many values of m produce chaotic oscillations when we run the system at different time scales and collect our data accordingly. In Figure 11 we plot this value for when we run the system for 50000, 100000, 1000000, and 10,000,000 time steps. Due to the large nature of the numbers involved we do this in a \log_{10} vs \log_{10} scale. Notice that as we run the system for longer periods of time, fewer values of m are chaotic. To understand the rate at which periodic regimes appear we fit a line to these data points and see that, if this trend continues, we will have chaotic values of m until we run the system for over a hundred trillion time steps. We conclude from this that it is safe to assume that our chaotic regimes are more than just a transient.

As was mentioned above, deriving precise probability distributions for these systems is not feasible, but we can reinforce these empirical findings about the oscillatory regimes in $N(E, I)$ by first deriving the expected spiking behavior using some simplifying mean field assumptions. This type of analysis is not suited for making claims about phenomena that arise from the topology of the network as we would

⁴We have done both and they yield the same results.

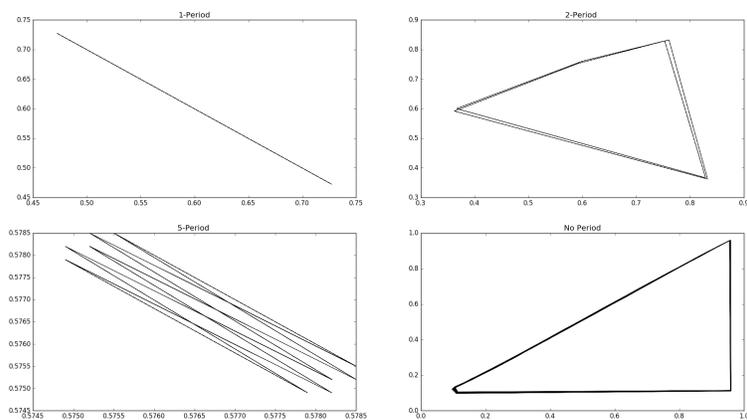


FIGURE 10. Comparison of the last 5000 time steps of a *1-periodic* regime, a *2-periodic* regime, a *5-periodic* regime, and a chaotic regime. Closed loops represent the trajectory of a spike. In the chaotic example there are so many in close proximity that it takes on the appearance of a single, thick, line.

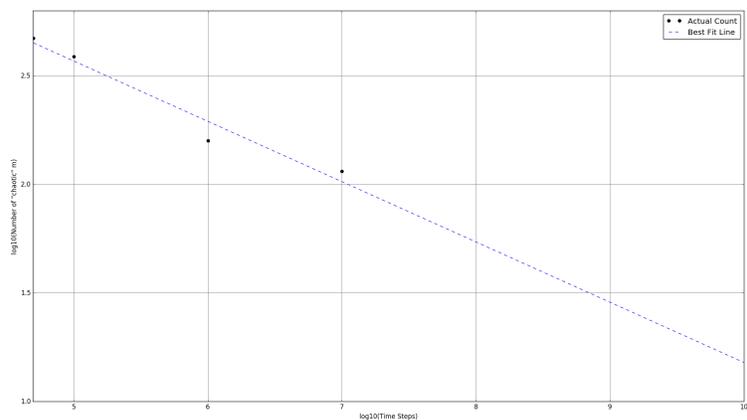


FIGURE 11. Number of chaotic parameters relative to amount of time the system is allowed to run. We plot the best fit line to predict how many parameters will be chaotic after a given time scale if the trend continues.

have wished above, but they are great when we are making qualitative claims about the networks dynamic behaviour.

We noted strict bounds in Proposition 1 for when the network is able to fire, but these bounds left a wide range of excitatory densities that have the potential

to fire. To shed some light on the probability distribution of firing when excitatory activation density is between the extremes highlighted above we humor the naive assumption that the probability of some excitatory vertex being active can be described by a Bernoulli trial with parameter p_E and that activation of vertices are mutually independent. Allow this we quickly find that

$$Pr\left(|A^E(t)| = x \mid N^2, p_E\right) = \binom{N^2}{x} p_E^x (1 - p_E)^{N^2 - x}$$

for $x \in \{0, \dots, N^2\}$. Furthermore the probability that an inhibitory vertex is active becomes $p_I = \sum_{j=\ell}^4 \binom{4}{j} p_E^j (1 - p_E)^{4-j}$ and so

$$Pr\left(|A^I(t)| \geq m \mid N^2, p_E\right) = \sum_{q=m}^{\frac{N^2}{4}} \binom{\frac{N^2}{4}}{q} p_I^q (1 - p_I)^{\frac{N^2}{4} - q}$$

By taking the derivative of this probability mass function and finding the value of p_E where the probability of firing is maximized we are able to find a curve that describes the excitatory densities which are most likely to fire at each m . We then combine this with our knowledge that our system necessarily loses at least $\frac{\ell * m}{N^2}$ excitatory density upon firing to find a curve that describes the maximal density the network will fall to after firing from the density derived from the mean field assumptions. We have just predicted a pre-firing curve and a post-firing curve.

Extending this analysis further, by our mean field assumptions we are able to derive a distribution of the excitatory density on time step $t+1$ knowing the current density at t which becomes a particularly potent tool using our derived post-firing densities. Let IA be the set of vertices that were inactive at time t but active at time $t+1$. Similarly, let AI be the set of vertices that were active at time t but inactive at time $t+1$. Then clearly $|A^E(t+1)| = |A^E(t)| + |IA| - |AI|$.

Notice that for some excitatory vertex to be in IA it needed to be inactive and also have at least k active neighbors in time t . Let p_{IA} be the probability that an active vertex is in IA . Since we know $|A^E(t)|$, then $p_{IA} = \sum_{j=k}^4 \frac{\binom{|A^E(t)|}{j} \binom{N^2 - |A^E(t)|}{4-j}}{\binom{N^2}{4}}$.

It follows then that

$$Pr\left(|IA| = x \mid N^2, |A^E(t)|\right) = \binom{N^2 - |A^E(t)|}{x} p_{IA}^x (1 - p_{IA})^{N^2 - |A^E(t)| - x}$$

We know every variable needed to compute p_{IA} and this is a known distribution so we easily find that

$$\mathbf{E}(|IA|) = (N^2 - |A^E(t)|) p_{IA}$$

In the same vein note that in order for some excitatory vertex to be in AI it needed to be active and also have at least $4 - (k - 2)$ inactive neighbors in time t . Note that for $k < 2$ this shows us that $|AI|$ is always 0. Let p_{AI} be the probability that an inactive vertex is in AI . Then since we know $|A^E(t)|$ we find that, for

$k \geq 2$, $p_{AI} = \sum_{j=6-k}^4 \frac{\binom{N^2 - |A^E(t)|}{j} \binom{|A^E(t)|}{4-j}}{\binom{N^2}{4}}$. Therefore

$$Pr\left(|AI| = x \mid N^2, |A^E(t)|\right) = \binom{|A^E(t)|}{x} p_{AI}^x (1 - p_{AI})^{|A^E(t)| - x}$$

once again using our knowledge of p_{AI} and the above distribution we can find that

$$\mathbf{E}(|AI|) = |A^E(t)| p_{AI}$$

Putting what we have found together then, since we know $|A^E(t)|$,

$$\begin{aligned} \mathbf{E}(|A^E(t+1)|) &= |A^E(t)| + \mathbf{E}(|IA|) - \mathbf{E}(|AI|) \\ &= |A^E(t)| + (N^2 - |A^E(t)|)p_{IA} - |A^E(t)|p_{AI} \end{aligned}$$

We then compute these values and overlay the results onto the actual bifurcation diagram to produce Figure 12.

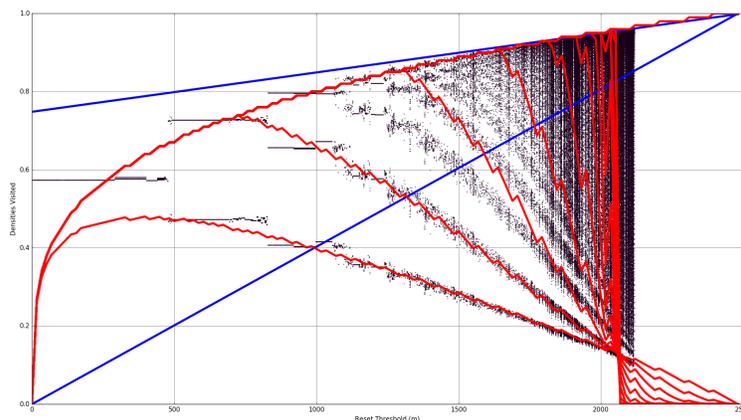


FIGURE 12. Results of the above mean field analysis in red. Strict upper and lower bound on activation density able to fire in blue. We put this over the actual bifurcation diagram to see how well it predicts the behavior.

The curves displayed in Figure 12 were computed using $N = 24$ instead of the actual size of our excitatory layer due to the machine approximation error that results from computing these distributions for large N . Regardless, these approximations fit extremely well and only deviate from our collected data on two accounts. First, and perhaps the biggest discrepancy in this approximation, is the fact that it predicts that the system will be unable to produce oscillations earlier than our empirical data suggests, but this is not surprising due to the fact that by ignoring topology, as we do in the mean field assumptions, it is in some sense “easier” to fire. If we suppose that this range of m not accounted for in this simplified analysis continued in a similar vein, then these results are incredibly telling. Second our actual data suggests that a single spike will contain more steps than our simplified model suggests. Again, knowing that we ignored the topology helps reconcile this, but when we also consider Figures 13 and 14 we see that as m increases so does the deviation in the individual spikes and the average length.

Couple these observations with the the variance in each time step, the fact that the post-firing curve is the maximal such curve, and also the fact that pre-firing curve is found by ignoring the topology of the network, and we can begin to imagine what this may look like if we managed to plot multiple spikes and we also begin to see why the chaotic regions of the bifurcation diagram arise where they are.

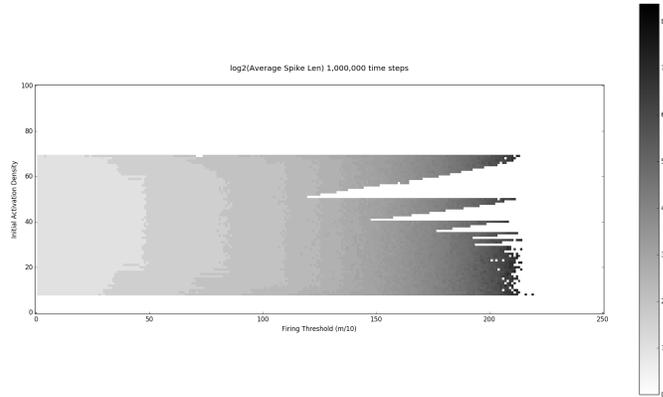


FIGURE 13. Average Spike Length of spikes that occurred in the final 50,000 time steps of a 1,000,000 time step run.

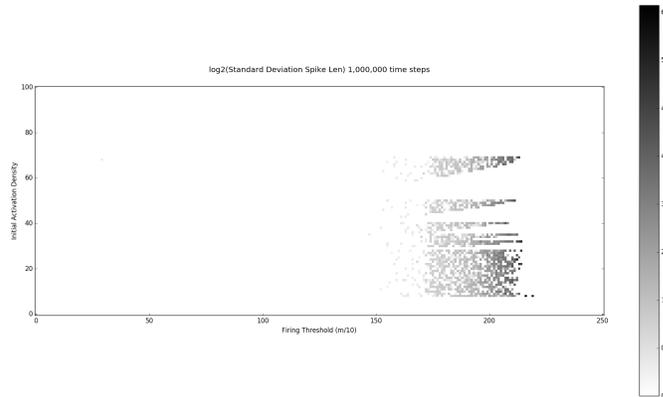


FIGURE 14. Standard Deviation of Spike Length of spikes that occurred in the final 50,000 time steps of a 1,000,000 time step run.

4.1. Initial Results on External Stimuli. A major motivation in learning to navigate our parameter space as accurately as possible is to use these networks for computational memories via chaos control. By putting together the chaotic behaviour of the system at certain values of m with the observation that a fixed value of m can undergo significant changes in regime with small changes in λ makes for a promising application in robust encoding. En route to this goal we have begun experimenting with the introduction of noise to targeted regions of the excitatory lattice E .

In Figure 15 we see side by side the delay-embeddings of the system when $p_0 = 0.5$ and $m = 1760$. The embeddings are from just before, immediately following, 10,000 time steps following, and 50,000 time steps following the injection of stimuli. We injected additive stimuli to the “top-right” corner of E for 100 time steps after the system had been given 40,000 time steps to out grow its first transient and allowed it to run for a grand total of 100,000 time steps and settle back down after the injection. In this example the stimuli took the form of activating an excitatory vertex that would normally be inactive based on a Bernoulli trial with 0.3 chance of activating.

Notice in Figure 15 that when we inject the noise, the system is “bumped” out of chaos and into a more tame period, then quickly returns to its chaotic state. This system living on the verge of criticality is precisely the behaviour we want. It implies that we can let it idle in a chaotic state and spur specific oscillations in response to stimuli. It is important we note that the 0.3 parameter on the Bernoulli trial nor the region that we injected noise into were chosen for any deep reasons and that this would still hold for other configurations. As affirmation of this we include Figure 16 which is produced from exactly the same experiment other than the fact that we applied stimuli to the “bottom left” corner of E ; notice that the behaviour is virtually the same, but that the system is “bumped” into a different trajectory than when we inject the same amount of noise into the opposite corner of E .

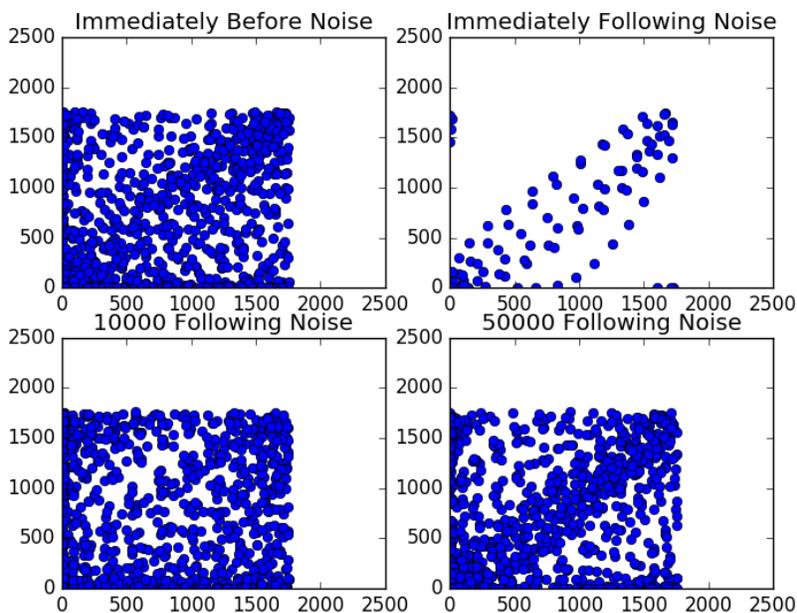


FIGURE 15. Delay-embedding at four time frames relative to the noise. We injected noise in the “top right” corner of E with $m = 1760$.

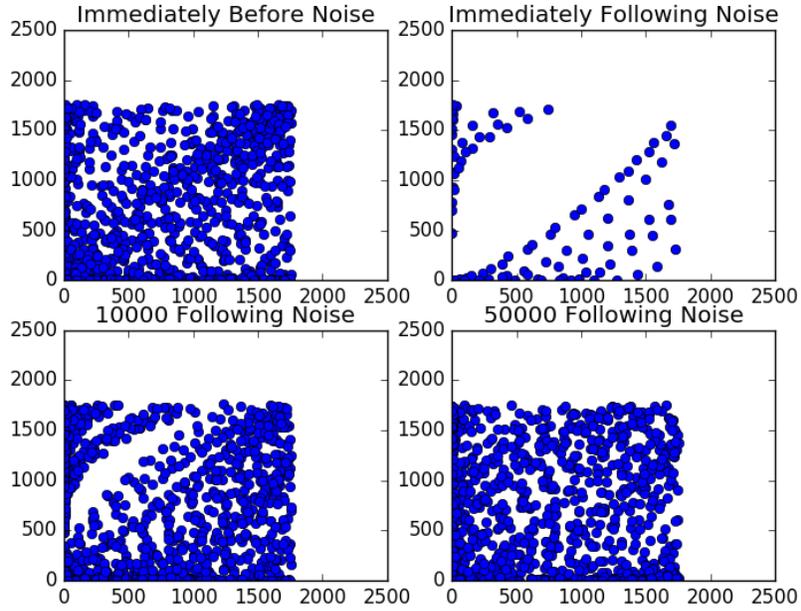


FIGURE 16. Delay-embedding at four time frames relative to the noise. We injected the same amount of noise as in Figure 15 but in the “bottom left” corner of E for $m = 1760$.

Explorations into the injection of noise is still in its infantile stages, but we are optimistic that this will make for even more robust, interesting, and controlled oscillations throughout the parameter space.