Chaos in random recurrent neural networks

Tuan Pham

1. Overall description

The goal of my project is to quantitatively describe the dependence of chaotic/regular regimes on parameters of recurrent random neural networks with random connectivity, partially via replicating the main results of stability in recurrent neural networksin [1, 2, 3, 4]. Particularly, the main parameter is the gain of the nonlinear activation function of artificial neural units. Lastly, I used the analytical results from [3] to examine to dependence of order/chaos of the system on the existence of bias terms.

Disclaimer: I was struggling to understand the mean-field theory derivation from [1] so I combined some additional sources [2, 3, 4] to produce analytical results in discrete-time system instead, then ran simulations in continuous time to compare the chaotic/regular dependence on the parameters.

2. Concept introduction

2.1. Single unit construction and activation

A biological neuron could be modelled by many different ways, from simple binary-switching units to detailed compartmental conductance-based models. For the sake of simplicity, a neuron input-output response could be modelled with a hyperbolic tangent (Equation 1) to scale the input from $(-\infty, \infty)$ monotonically to a bounded range between $[-1, 1]$.

$$\phi(x) = \tanh(gx + b) \tag{1}$$

The rationale is that small inputs would result in inactive state (nearer to -1) whereas larger inputs would lead to more active state (nearer to 1). The continuous range, when scaled from $[-1, 1]$ to $[0, 1]$, could be interpreted in many ways, for example the probability of spiking in a given time step, or a normalized firing rate. Additionally, the continuous range and smoothness are easier to work with than applying conditions for spiking like binary-switching models.

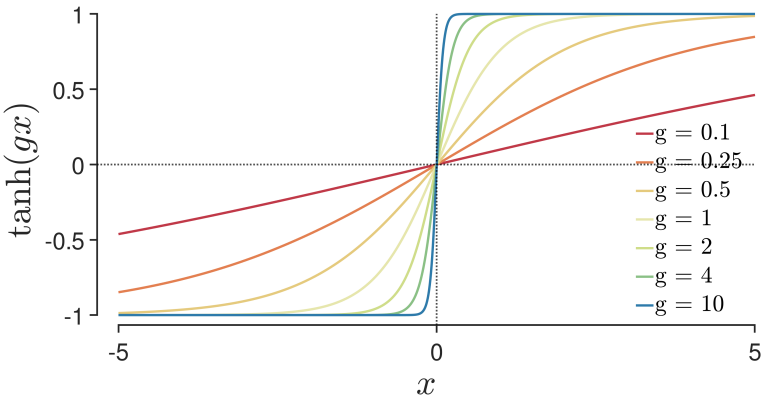

Figure 1: *Example of* $\tanh(gx)$ *for different gains (colors) but similar bias* $b = 0$

As for the parameters of the activation function, the gain g could be interpreted as the neuron's input resistance, whereas the bias b could be analogous to the opposite of a threshold for spiking or activated. Higher gain and larger bias would be appropriate to mimic the activation of a highly excitable neuron. For the sake of simplicity, I will ignore the bias terms for now ($b = 0$) and will integrate in the analysis later on with some modifications. Examples of the activation functions for different gain parameters are shown in Figure 1. It will be shown later on that the activation gain parameter

could be combined with a synaptic variance parameter to reduce the parameter dimension for analysis.

2.2. Recurrent neural network

Assume we have N neurons, each with its continuous local field input h_i and their activated state $S_i = \phi(h_i)$, with $i = \overline{1, N}$. A recurrent network connects the presynaptic neuron j (source) with postsynaptic neuron i (target) with a weight (synaptic efficay) of $\mathbf{J}_{ij}$, so that the input from $j \to i$ is $\mathbf{J}_{ij}S_j$. For simplicity, we will only have 2 constraints on the weight matrix:

- no self-connectivity, meaning $\mathbf{J}_{ii} = 0 \ \forall i$;
- $\mathbf{J}_{ij} \sim \mathcal{N}\left(0, J^2/N\right)$, independently from each other.

It should be noted that this is overly simplified - as in biological neural networks, there are other important constraints on excitatory and inhibitory connectivity to be the least, and proportions of excitatory and inhibtory neurons, as well as synaptic dynamics.

The governing equations for the local field variables are the system of the following Equation 2 (illustrated in Figure 2), without considering external outputs. The first term is analogous to a resting "leaky" term, whereas the latter corresponds to the weighted-sum of the inputs from presynaptic neurons.

$$\dot{h_i} = -h_i(t) + \sum_{j=1}^{N} \mathbf{J}_{ij} S_j = -h_i(t) + \sum_{j=1}^{N} \mathbf{J}_{ij} \phi(h_j) \tag{2}$$

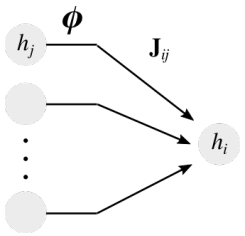

Figure 2: *Inputs to each unit of a recurrent network*

3. Properties

3.1. Reduction of gain parameters $gJ \leftarrow g$

Although in [1], the parameters of investigation include both the nonlinear gain g and the variability of the synaptic weights J, more specifically the product gJ. However, we could easily reduce such product to just one variable g and set $J = 1$. More specifically, divide Equation 2 by J, we'd have:

$$\frac{d(h_i/J)}{dt} = \frac{h_i}{J} + \sum_{j=1}^{N} \frac{\mathbf{J}_{ij}}{J} \tanh\left[(gJ) \times \frac{h_j}{J}\right] \tag{3}$$

Because $\mathbf{J}_{ij} \sim \mathcal{N}\left(0, J^2/N\right) \implies \mathbf{J}_{ij}/J \sim \mathcal{N}\left(0, 1/N\right)$. Now, we can make the following substitutions to Equation 3 to have Equation 2 with just one parameter g:

$$\begin{cases} gJ & \leftarrow g \\ h_i/J & \leftarrow h_i \\ \mathbf{J}_{ij}/J & \leftarrow \mathbf{J}_{ij} \sim \mathcal{N}\left(0, 1/N\right) \end{cases}$$

3.2. Jacobian

The derivative of the activation function with respect to the local field variable is

$$\phi'(x) = \frac{\partial \phi}{\partial x} = g\left(1 - \tanh^2(x)\right) = g\left(1 - \phi^2(x)\right)$$

Hence, the Jacobian of system could be expressed as Equation 4, with $\overrightarrow{h}$ as a column vector of the local field variables, and the element-wise multiplication operator $\circ$ defined as followed for a matrix $\mathbf{J}$ and a row vector $\overrightarrow{h}^T$: $\left[\mathbf{J} \circ \overrightarrow{h}^T\right]_{ij} = \mathbf{J}_{ij} \cdot h_j$.

$$Df = g\mathbf{J} \circ \left(1 - \phi^2\left[\overrightarrow{h}^T\right]\right) - \mathbf{I} \tag{4}$$

The explicit expression is useful for numerical calculation of the maximal Lyapunov exponent, as in [5], and implimented in MATLAB on [6]. Due to time constraints, I am not covering the summary of the methods outlined in [5].

4. Chaos analysis

4.1. Chaos analysis in discrete time

Although the problem is set in continuous time as originally in [1], it seems that it is easier to analyze in discrete time (like in [2]). Hence I will be analyzing the stability of the system in discrete time ($dt = 1$), then show simulation results of the continuous version (small dt). With $dt = 1$, $\dot{h} = h(t+1) - h(t)$, then Equation 2 becomes:

$$h_i(t+1) = \sum_{j=1}^{N} \mathbf{J}_{ij}\phi(h_j) \tag{5}$$

Local chaos hypothesis: To summarize from [3, 4], as $N \to \infty$, without the constraints of symmetric coupling (as in $\mathbf{J}_{ij} \neq \mathbf{J}_{ji}$), the local field variables $h_i(t)$ can be treated as independent from each other, and from the random weights in $\mathbf{J}$. In this case, they would be independent Gaussian processes. And we would just need to focus on one variable in the thermodynamic limit.

Because there are no bias terms and $\langle \mathbf{J} \rangle = 0$, we would expect $\langle h(t) \rangle = 0$, as in zero-mean. So now we'd focus on the second moment, which is also the variance (because mean is 0), of the local field variable: $\nu(t) = \langle h^2(t) \rangle$.

$$h_i \sim \mathcal{N}(0, \nu(t)) \tag{6}$$

Additionally, we also define $q(t) = \langle \phi^2(h(t)) \rangle = \text{var}(\phi(h(t))$ (because mean is also 0). Let's call x as a random variable with normally distributed with unit variance zero mean, $x \sim \mathcal{N}(0, 1)$ and $p_{\mathcal{N}}(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$. Because of Equation 6 and $x\sqrt{\nu(t)} \sim \mathcal{N}(0, \nu(t))$, we can rewrite $q(t)$ as following:

$$q(t) = \int_{-\infty}^{\infty} \phi^2\left(x\sqrt{\nu(t)}\right) p_{\mathcal{N}}(x) dx \tag{7}$$

Because h_i and $\mathbf{J}_{ij}$ are independent and that their means are 0, we'd have $\text{var}(\mathbf{J}_{ij}\phi(h_j)) = \text{var}(\mathbf{J}_{ij})q(t) = q(t)/N$. Combined with Equation 5 and Equation 7:

$$\nu(t+1) = \langle h^2(t+1) \rangle = N \times \frac{1}{N} q(t) = q(t) = \int_{-\infty}^{\infty} \phi^2 \left(x\sqrt{\nu(t)} \right) p_{\mathcal{N}}(x) dx \tag{8}$$

Let's call $\mathcal{M}(\nu(t))$ as the left-hand side of Equation 8. Then we have a 1-D discrete mapping $\nu(t+1) = \mathcal{M}(\nu(t))$ and we could attempt to investigate the existence of a fixed point ν^* as $t \to \infty$. The fixed point would satisfy $\nu^* = \mathcal{M}(\nu^*)$. We could make assumption about the fixed point and do a Taylor expansion to reach an analytic approximation. However, the actual values of ν^* are less important, so we could evaluate it graphically for ease (Figure 3, left).

Additionally, according to [2], the maximum Lyapunov exponent could be expressed as function of the fixed point, as in Equation 9. The partial proof of this is shown below.

$$\lambda = \frac{1}{2} \log \int_{-\infty}^{\infty} \left[\phi' \left(\sqrt{\nu^*} x \right) \right]^2 p_{\mathcal{N}}(x) dx \tag{9}$$

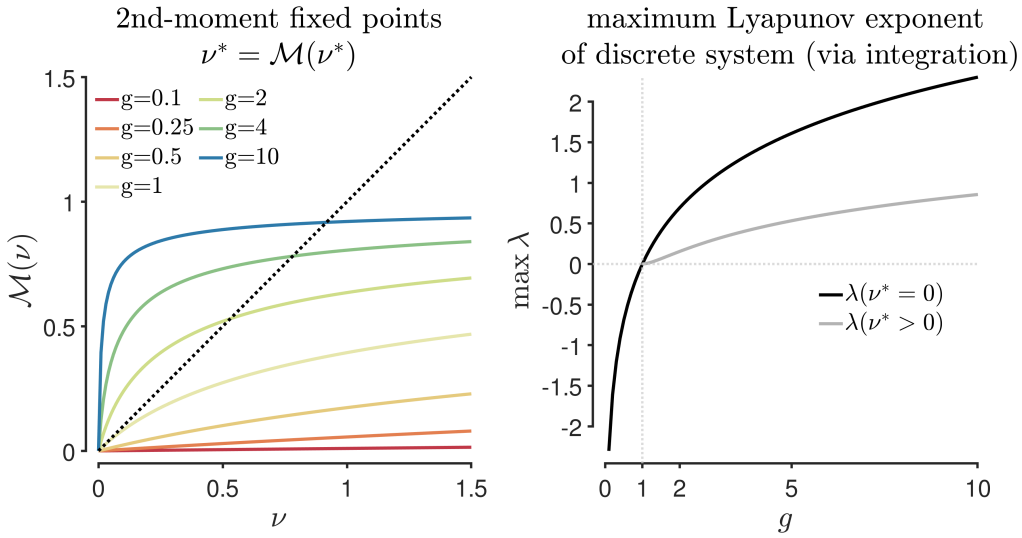

Figure 3: *Analysis of second moment fixed points and the corresponding maximal Lyapunov exponent of the discrete version.* (**Left**) Second moment fixed points following Equation 8 for different gains g (colors). The fixed points are at the intersection of the identity line and the colored curves. (**Right**) The maximum Lyapunov exponent as obtained via integration in Equation 9, as a function the nonlinear gain g. The black line corresponds to $\nu^* = 0$, whereas the gray line corresponds to $\nu^* > 0$ (only for $g > 1$).

As illustrated in Figure 3 (left), it can be seen that for $g \leq 1$, there is only one fixed point, which is $\nu^* = 0$, which is actually stable as the maximal Lyapunov exponent calculated via Equation 9 are negative (Figure 3, right). However, for $g > 1$, aside from $\nu^* = 0$ as a fixed point, there is another nonzero fixed point within the unit square. Both of the maximal Lyapunov exponents corresponding to these 2 fixed points for each g are positive, indicating chaotic regime.

NOTE: I am uncertain whether the actual magnitudes of Lyapunov exponents between discrete time and continuous time system agree

Partial Proof of Equation 9

To calculate the Lyapunov exponent, we need to define two initial conditions $h^{(1)}(0)$ and $h^{(2)}(0)$, and observe the difference between the corresponding trajectories. We also define the expected squared difference between them $d_{12}^2(t) = \left\langle \left(h^{(1)}(t) - h^{(2)}(t) \right)^2 \right\rangle$.

The maximum Lyapunov exponent λ could then be estimated by:

$$\lambda = \lim_{\substack{t \to \infty \\ d_{12}(0) \to 0}} \frac{1}{t} \log \frac{|d_{12}(t)|}{|d_{12}(0)|} = \lim_{\substack{t \to \infty \\ d_{12}(0) \to 0}} \frac{1}{2t} \log \frac{d_{12}^2(t)}{d_{12}^2(0)} \tag{10}$$

Additionally, we also define the correlation between the 2 trajectories as $\rho_t = \rho(t)$ (Equation 11). With $\nu = \nu^*$ (lose the superscript for short) is the fixed point of $\nu(t) = \left\langle h^2(t) \right\rangle$ (Equation 6 and Equation 8), we can express $d_{12}^2(t)$ as Equation 12. And we could rewrite Equation 10 as Equation 13:

$$\rho_t = \rho(t) = \left\langle h^{(1)}(t) h^{(2)}(t) \right\rangle \tag{11}$$

$$d_{12}^2(t) = \left\langle \left[h^{(1)}(t) \right]^2 \right\rangle + \left\langle \left[h^{(2)}(t) \right]^2 \right\rangle - 2 \left\langle h^{(1)}(t) h^{(2)}(t) \right\rangle = 2\nu - 2\rho_t \tag{12}$$

$$\implies \lambda = \lim_{\substack{t \to \infty \\ \rho_0 \to \nu}} \frac{1}{2t} \log \frac{\nu - \rho_t}{\nu - \rho_0} \tag{13}$$

With the correlation definition in Equation 11 and mean-field theory results in Equation 6, we have:

$$\begin{aligned} \rho_t = \left\langle h^{(1)}(t) h^{(2)}(t) \right\rangle &= \left\langle \phi\left(h^{(1)}(t-1) \right) \phi\left(h^{(2)}(t-1) \right) \right\rangle \\ &= \int \mathrm{D}x \int \mathrm{D}y \; \phi(\sqrt{\nu} x) \; \phi\left(Y(x,y) \right) \\ &\qquad \text{where } \mathrm{D}x = p_{\mathcal{N}}(x)dx \text{ and } x, y \sim \mathcal{N}(0,1), x \perp\!\!\!\perp y \end{aligned}$$

We need the random variable defined as $Y(x,y)$ to have $\langle Y \rangle = 0, \mathrm{var}(Y) = \nu$. More importantly, the correlation between x and Y needs to satisfy $\langle XY \rangle = \left\langle (x\sqrt{\nu}) Y \right\rangle = \rho_{t-1}$. A way to achieve that is:

$$Y(x,y) = \frac{\rho_{t-1}}{\sqrt{\nu}} x + \sqrt{\frac{\nu^2 - \rho_{t-1}^2}{\nu}} y$$

Indeed, as $x, y \sim \mathcal{N}(0,1)$ and $x \perp\!\!\!\perp y$, those conditionsn are met:

1. $\langle Y \rangle = c_1 \langle x \rangle + c_2 \langle y \rangle = 0$, since $\langle x \rangle = \langle y \rangle = 0$
2. $\mathrm{var}(Y) = \frac{\rho_{t-1}^2}{\nu} \mathrm{var}(x) + \frac{\nu^2 - \rho_{t-1}^2}{\nu} \mathrm{var}(y) = \nu$, since $\mathrm{var}(x) = \mathrm{var}(y) = 1$
3. $\langle XY \rangle = \sqrt{\nu} \langle xY \rangle = \rho_{t-1} \left\langle x^2 \right\rangle + \sqrt{\nu^2 - \rho_{t-1}^2} \langle xy \rangle = \rho_{t-1}$, since $\left\langle x^2 \right\rangle = \mathrm{var}(x) = 1$ and $x \perp\!\!\!\perp y \to \langle xy \rangle = 0$

So we can define $\rho_t = R(\rho_{t-1})$ and $R^n \equiv R \circ R \circ \cdots \circ R$ (chained n times).

$$\rho_t = R(\rho_{t-1}) = \int \mathrm{D}x \int \mathrm{D}y \; \phi(\sqrt{\nu} x) \; \phi\left(\frac{\rho_{t-1}}{\sqrt{\nu}} x + \sqrt{\frac{\nu^2 - \rho_{t-1}^2}{\nu}} y \right) \tag{14}$$

$$= R \circ R(\rho_{t-2}) = \cdots = R^t(\rho_0) \tag{15}$$

One thing to notice is $\nu = R(\nu)$, because $R(\nu) = \int \mathrm{D}x\ \phi^2(x\sqrt{\nu}) \int \mathrm{D}y = \int \mathrm{D}x\ \phi^2(x\sqrt{\nu})$ (because $\int \mathrm{D}y = \int p_{\mathcal{N}}(y)dy = 1$). Additionally, because of Equation 7 and Equation 8, we would have $q(t) = \nu(t+1) \to \nu^* \equiv \nu$. Hence $\nu = q = \int \mathrm{D}x\ \phi^2(x\sqrt{\nu}) = R(\nu)$. This means $\nu = R^t(\nu)\ \forall t > 0$. Combine this with Equation 15, we can rewrite Equation 13 as:

$$\lambda = \lim_{\substack{t \to \infty \\ \rho_0 \to \nu}} \frac{1}{2t} \log \frac{R^t(\nu) - R^t(\rho_0)}{\nu - \rho_0} = \lim_{t \to \infty} \frac{1}{2t} \log \left. \frac{\partial R^t(\rho)}{\partial \rho} \right|_{\rho=\nu} = \frac{1}{2} \log \left. \frac{\partial R(\rho)}{\partial \rho} \right|_{\rho=\nu} \tag{16}$$

The last equality in Equation 16 is because $\left. \frac{\partial R^t(\rho)}{\partial \rho} \right|_{\rho=\nu} = \left(\left. \frac{\partial R(\rho)}{\partial \rho} \right|_{\rho=\nu} \right)^t$, as can be obtained via chain rule, and the fact that $\nu = R^t(\nu)\ \forall t > 0$ (see above).

However, I was **not** able to prove Equation 17, as claimed by both [2] and [3]. This is why this is just a partial proof.

$$\left. \frac{\partial R(\rho)}{\partial \rho} \right|_{\rho=\nu} = \int \mathrm{D}x\ \left[\phi'(x\sqrt{\nu})\right]^2 \tag{17}$$

Accepting Equation 17 and combining Equation 16, one could easily infer Equation 9.

4.2. Simulations of continuous system

4.2.1. Coarse scan of g

I ran continuous time (as in $dt < 1$) simulation for $N = 200$ until $t = 100$. For each value of g, I had 10 realizations of the network ($\mathbf{J} \sim \mathcal{N}(0, 1)$) and initial conditions h_0. I used the implemented package from [6] to calculate the Lyapunov exponent. The results are shown in Figure 4.

First of all, as expected, for $g \leq 1$, the solutions seem to decay to $h_i = 0$. This agrees with $\nu^* = 0$, as discussed above. The sign of the maximal Lyapunov exponents also seem to agree generally with the discrete case (Figure 3, right). This supports that this fixed point is stable.

Secondly, for many of $g > 1$, the activity of the network tends towards chaotic in the simulations, as also supported by the positivity of the maximal Lyapunov exponents (Figure 4, right). This, in the gross part, agrees with the positivity observed in the discrete case (Figure 3).

Disregarding the mismatch between actual magnitudes between the discrete and continuous versions, the signs, which in turn indicate the regularity (order) of the system, do seem to agree, with the exception of g near 1, particularly the range of $1 < g < 2$. There might be more points in this range that could correspond to negative λ. Additionally, this could be due to the possibility that N is not large enough.

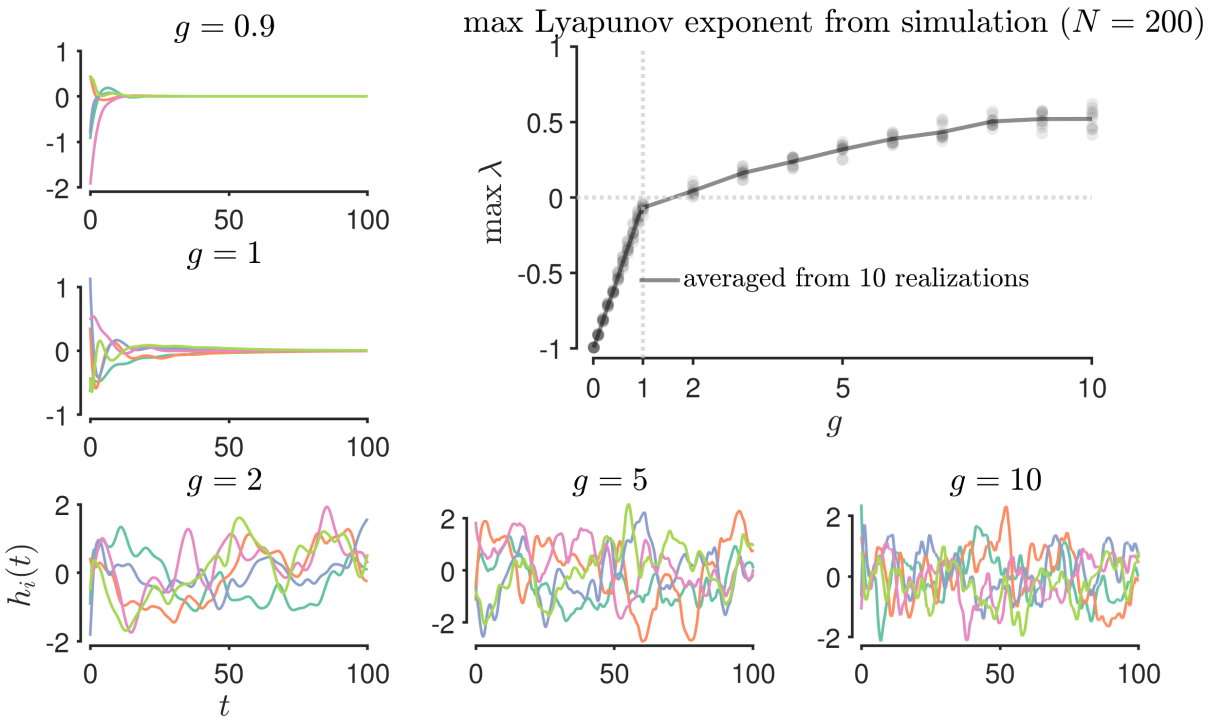

Figure 4: *Lyapunov exponents of the continuous timed simulations.* The small panels on the left and bottom show examples of the simulations for different values of g. Plotted are 5 randomly chosen local field variables $h_i(t)$ in 200 neurons. The top right panel shows the maximal Lyapunov exponents at the end of the simulation ($t = 100$) estimated via [6], averaged across 10 different realizations of the weight matrix and initial conditions.

4.2.2. Finer scan of $g \approx 1$

Next I inspected whether λ could have some dependence on the network sizes N with the gain parameters near unity ($g \approx 1$). Because the implementation of [6] took a long time to run for large N (the problem scales $O(N^2)$ for the implementation), and of course for longer T, I only ran 5 realizations for each (N, g), and with $N \in \{100, 200, 300\}$ and 7 values of $g \in [0.95, 2]$. Running with larger N was not practical for my laptop unfortunately.

The results are shown in Figure 5. Similar to the case of $N = 200$ in Figure 4, $g = 1$ is not exactly the critical point - the transition between negative to positive Lyapunov exponents seems to be some value around $g \approx 1.5$. Running a simple t-test with zero-mean λ as the null hypothesis (Figure 5 bottom), because the sample size is too small ($n = 5$), I also plotted the significance level $\alpha = 0.1$ for observation. For $N = 300$, it seems that the null hypothesis could be rejected, except for $g = 1.5$. And for all N's investigated here, the null hypothesis could not be rejected $g = 1.5$. This seems to supported that the critical value of g is different from the theoretical estimate ($g = 1$) for the investigated N's.

Again, it is possible that N needs to be larger (≈ 1000) and the simulation needs to run longer T (maybe ≈ 1000 for confident estimation). Other Lyapunov exponent estimation methods or certain modifications of the method [6] might be able to achieve better estimation.

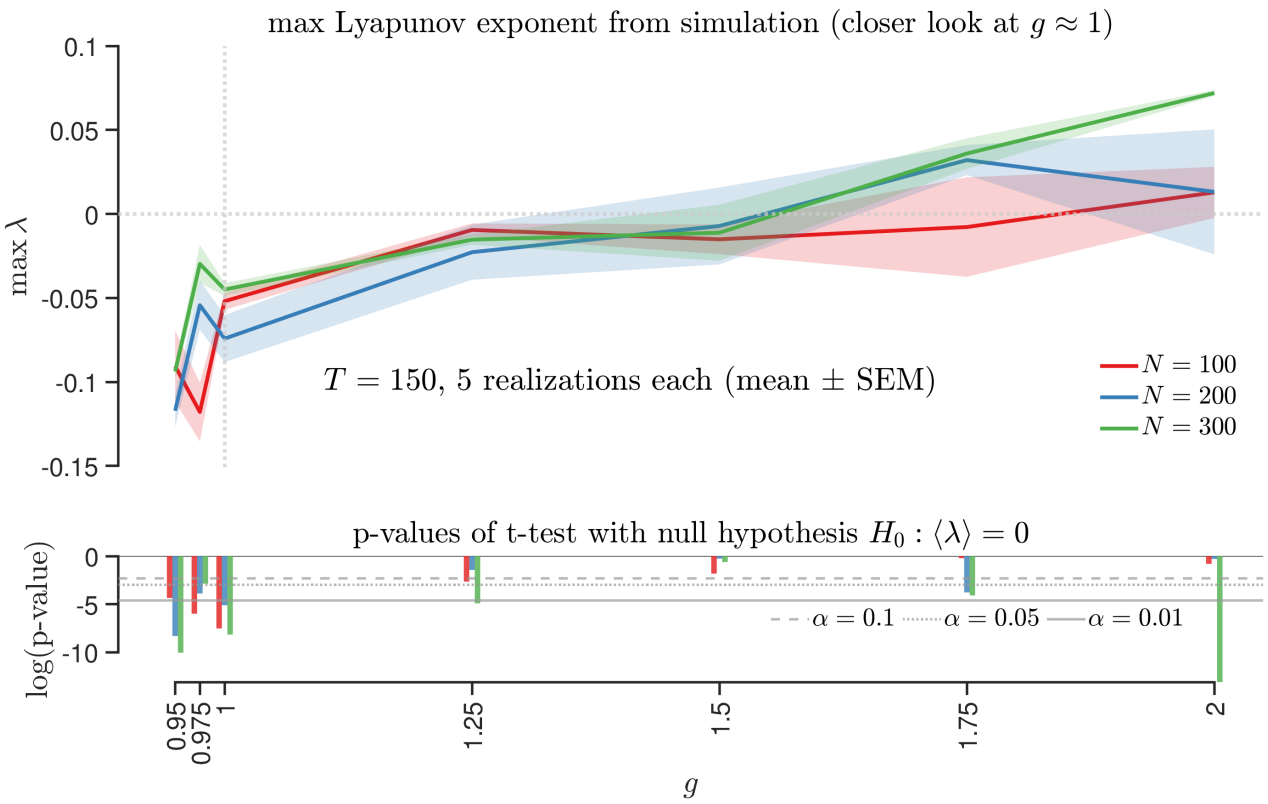

Figure 5: *Lyapunov exponents for $g \approx 1$.* (**Top**) Mean λ (line) with SEM (shading) ($n = 5$) as a function of the gain parameter g (like Figure 4, but with a closer look around $g \approx 1$), with different network sizes $N = 100, 200, 300$ (colors). (**Bottom**) log of p-values of t-test for λ with null hypothesis as zero mean λ, along with horizontal lines of different significance levels α.

NOTE: I was not able to investigate period doubling using a simple method as tracking $d(t) = \left|\vec{h}(t) - \vec{h}(T)\right|$, a bit similar to a previous homework.

5. Inclusion of bias terms b

For simplicity, we still assume $J = 1$ and discrete time system but consider each neuron has its own bias b_i independent from each other but identically distributed as $b_i \sim \mathcal{N}\left(\beta, \sigma_\beta^2\right)$. However, instead of defining the bias inside the activation function like in Equation 1, we will instead redefine $h_i(t+1)$ to include the bias terms as followed:

$$\begin{cases} h_i(t+1) & = \sum_{j=1}^{N} \mathbf{J}_{ij} S_j(t) + b_i \\ S_i(t) & = \phi(h_i(t)) = \tanh(g \cdot h_i(t)) \end{cases} \tag{18}$$

I won't be deriving the following but will summarize from [3] below. The derivation is more generalized with the inclusion of information about the bias distribution, but is quite similar to the derivations from Equation 5 to Equation 8 using the local chaos hypothesis in the thermodynamic limit. I will focus in this section on the results of the regimes for chaos/order

when the biases are considered, stemming from such analytical derivations instead.

Define the following moments:

$$\begin{cases} \mu(t) & = \langle h(t) \rangle \\ \nu(t) & = \mathrm{var}(h(t)) = \langle h^2(t) \rangle - \mu^2(t) \\ m(t) & = \langle S(t) \rangle \\ q(t) & = \langle S^2(t) \rangle \end{cases}$$

With the local hypothesis, the evolutions of $\mu(t)$ and $\nu(t)$ in discrete time are as followed:

$$\begin{cases} \mu(t+1) & = \beta \\ \nu(t+1) & = q(t) + \sigma_\beta^2 \end{cases}$$

In the thermodynamic limit, the moments of $S(t)$ could be given as followed:

$$\begin{cases} m(t) & = \int \mathrm{D}x \ \phi(\sqrt{\nu(t)}x + \mu(t)) \\ q(t) & = \int \mathrm{D}x \ \phi^2(\sqrt{\nu(t)}x + \mu(t)) \end{cases} \qquad \text{where } x \sim \mathcal{N}(0,1) \text{ and } \mathrm{D}x = p_{\mathcal{N}}(x)dx$$

Since we simplified $\mathbf{J}_{ij} \sim \mathcal{N}(0,1)$, $\mu(t) = \mu = \beta$. The fixed point for $\nu(t) \to \nu$, with:

$$\begin{cases} \nu = q + \sigma_\beta^2 \\ q = \int \mathrm{D}x \ \phi^2\left(x\sqrt{\nu} + \beta\right) \end{cases} \implies \nu = \sigma_\beta^2 + \int \phi^2\left(x\sqrt{\nu} + \beta\right) \mathrm{D}x \tag{19}$$

The expected maximal Lyapunov exponent is quite similar to Equation 9, as it can be expressed as Equation 20.

$$\lambda = \frac{1}{2} \log \int_{-\infty}^{\infty} \left[\phi'\left(\sqrt{\nu}x + \beta\right)\right]^2 \mathrm{D}x \tag{20}$$

Additionally, because the derivative of $\phi(x)$ w.r.t x is an even function and λ includes the integral from both $x \to \infty$ and $x \to -\infty$, it can easily be seen that $\lambda(\beta)$ is an even function w.r.t. β: $\lambda(\beta) = \lambda(-\beta)$. So for simplicity, we will only focus on $\beta \geq 0$.

Using Equation 19, we could easily find the fixed points depending the parameter set (g, β, σ_β) like Figure 4 (left). And similarly, we could use Equation 20 to find out the corresponding λ, similarly to Figure 4 (right). Instead of looking at the actual values of λ, we just need to focus on the sign of λ. Particularly, $\lambda < 0$ would indicate order/regularity of the corresponding fixed point (light blue in small panels of Figure 6); while $\lambda > 0$ would indicate chaos of the corresponding fixed point (light red in small panels of Figure 6).

The dependencies of the system chaos/order boundaries on the gain and bias parameters are shown in Figure 6. First off, as expected, larger g leads to chaos more readily. Secondly, larger $|\beta|$ and/or σ_β increases the area of the "order" regime. In other words, larger gains tend the system towards chaos, whereas the existence and variability of the bias terms tend the system towards order.

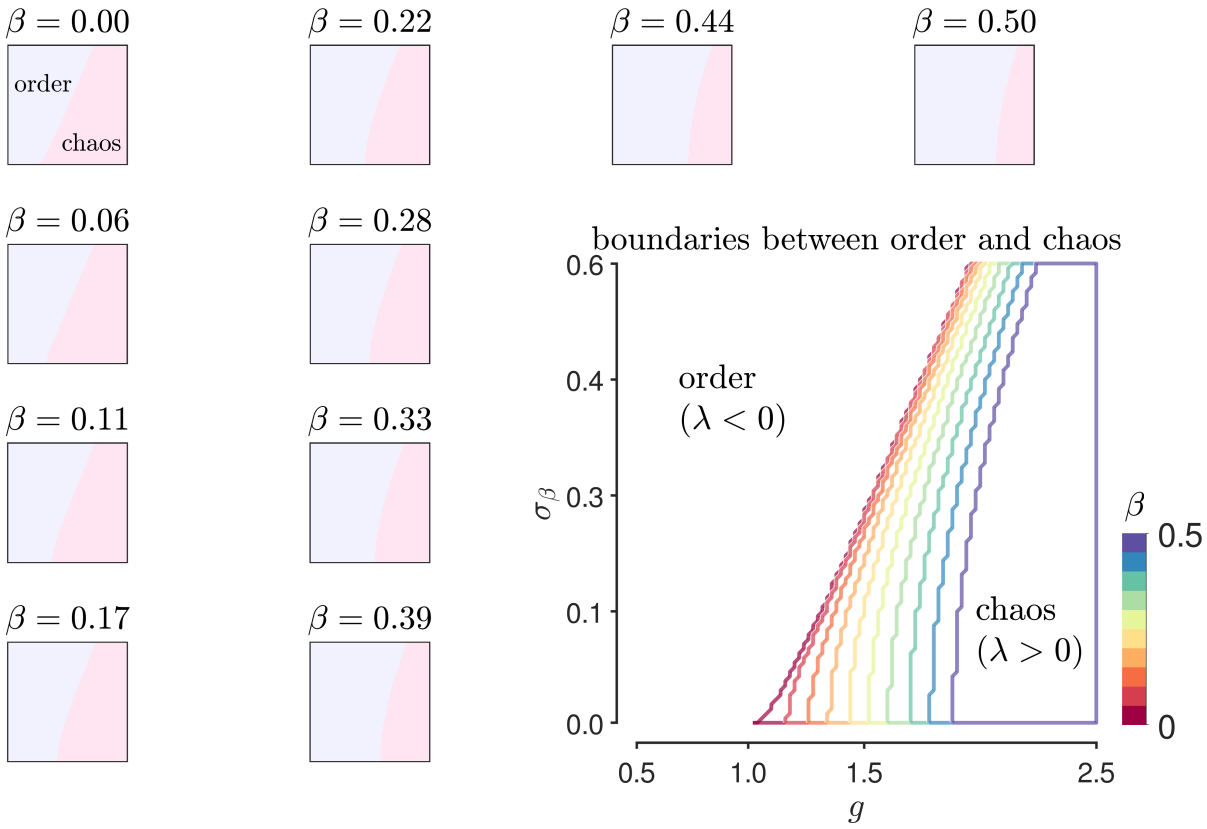

Figure 6: *Order and chaos depending the gain of the activation function and properties of the bias distribution.* Each small panel shows the signs of λ, corresponding to a given β, x-axis is g and y-axis is σ_β. Light blue: order, $\lambda < 0$. Light red: chaos, $\lambda > 0$. The bottom large panel shows the boundaries between order and chaos. The colors represent different β values.

References

[1] H. Sompolinsky, A. Crisanti, and H. J. Sommers. Chaos in random neural networks. *Phys. Rev. Lett.*, 61:259–262, Jul 1988.

[2] L. Molgedey, J. Schuchhardt, and H. G. Schuster. Suppressing chaos in neural networks by noise. *Physical Review Letters*, 69(26):3717–3719, December 1992.

[3] B. Cessac. Increase in Complexity in Random Neural Networks. *Journal de Physique I*, 5(3):409–432, March 1995.

[4] B. Doyon, B. Cessac, M. Quoy, and M. Samuelides. Mean-field equations, bifurcation map and chaos in discrete time, continuous state, random neural networks. *Acta Biotheoretica*, 43(1-2):169–175, June 1995.

[5] Alan Wolf, Jack B. Swift, Harry L. Swinney, and John A. Vastano. Determining lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, 16(3):285 – 317, 1985.

[6] Vasiliy Govorukhin. *Calculation Lyapunov Exponents for ODE*, Retrieved May 19, 2020. MATLAB Central File Exchange.