**Eigenvectors and eigenvalues in biology:** Principal Component Analysis (PCA)

mRNA expression levels of XPB1 & GATA3; M=105 breast tumor cells

- estrogen receptor -
- estrogen receptor +


- Principal Component Analysis determines the directions of maximum variation in a data cloud, from the eigenvectors of a symmetric correlation matrix. 
  \[ C_{\alpha\beta} = \frac{1}{M} \sum_{j=1}^{M} r_j^\alpha r_j^\beta \equiv \langle r_j^\alpha r_j^\beta \rangle \]

- Eigenvalues of the correlation matrix in the remaining directions determine the residual. 
  \[ |\vec{\lambda}|^2 = \lambda_1 + \lambda_2, \quad R = \lambda_2 \]

- It can be easier to determine epistatic interactions between genes by looking along the principal components
Eigenvalues and Eigenvectors in Neural Networks

- Non-Hermitian, i.e., asymmetric matrices, with complex eigenvalues, arise naturally in simple models of networks containing both excitatory and inhibitory neurons (or species!)

- Localized eigenvectors dominate the eigenvalue spectra for sparse random neural networks.

- An intricate eigenvalue spectrum controls the spontaneous activity and induced response. Directed rings of neurons lead to a hole centered on the origin in the density of states in the complex plane.

- All states are extended on the rim of this hole, while the states outside the hole are localized.

Physical Review E93, 042310 (2016)
**Eigenvectors and eigenvalues in biology: rabbits vs. sheep**

\[ \frac{dx}{dt} = 3x(1 - x / 3) \]
\[ \frac{dy}{dt} = 2y(1 - y / 2) \]

**decoupled model:**

**two logistic equations**

linearize about the fixed point at (3,2)

\[ x'(t) = x(t) - 3, \quad y'(t) = y(t) - 2 \]

\[ \begin{pmatrix} dx'(t) / dt \\ dy'(t) / dt \end{pmatrix} \approx \begin{pmatrix} -3 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} \]

\[ x'(t) = x'(0)e^{-\lambda_1 t}, \quad y'(t) = y'(0)e^{-\lambda_2 t} \]

two real eigenvalues:

\[ \lambda_1 = -3, \quad \lambda_2 = -2, \] stable fixed point

x(t) = number of rabbits
y(t) = number of sheep

**mutualism**

\[ x(t) = \text{number of rabbits} \]
\[ y(t) = \text{number of sheep} \]
**Eigenvectors and eigenvalues in biology: rabbits vs. sheep**

\[
\frac{dx}{dt} = 3x(1 - x/3 - 2y/3)
\]

\[
\frac{dy}{dt} = 2y(1 - y/2 - x/2)
\]

**or…. two coupled inhibitory neurons**

\[
\nu_i = \text{firing rate deviation from the background rate of the } i^{th} \text{ neuron}
\]

\[
\tau \frac{dv_1}{dt} = -v_1 + \tanh \left[ M_{11}v_1 + M_{12}v_2 \right]
\]

\[
\tau \frac{dv_2}{dt} = -v_2 + \tanh \left[ M_{21}v_1 + M_{22}v_2 \right]
\]

\[
M = \begin{pmatrix}
0 & -s \\
-s & 0
\end{pmatrix}
\]

**Eigenvectors and eigenvalues in biology: Rabbits vs. Sheep**

\[
\frac{dx}{dt} = 3x(1 - x/3 - 2y/3) \\
\frac{dy}{dt} = 2y(1 - y/2 - x/2)
\]

Four fixed points are obtained: (0,0), (0,2), (3,0), and (1,1).

linearize about the fixed point at (1,1)

\[x'(t) = x(t) - 1, \quad y'(t) = y(t) - 1\]

\[
\begin{pmatrix}
\frac{dx'(t)}{dt} \\
\frac{dy'(t)}{dt}
\end{pmatrix} \approx
\begin{pmatrix}
-1 & -2 \\
-1 & -1
\end{pmatrix}
\begin{pmatrix}
x'(t) \\
y'(t)
\end{pmatrix}
\]

two real eigenvalues control dynamics:

\[\lambda_1 = -1 + \sqrt{2}, \quad \lambda_2 = -1 - \sqrt{2}\]

due to interactions, there is now one

stable and one unstable eigendirection

---

Rabbits vs. Foxes: complex eigenvalues lead to oscillations...

II. LOTKA-VOLterra equation

\[ \frac{dY_1}{dt} = c_1XY_1 - c_2Y_1Y_2 \]
\[ \frac{dY_2}{dt} = c_2Y_1Y_2 - c_3Y_2 \]

2 fixed point: \( (0, 0), \left( \frac{c_3}{c_2}, \frac{c_1X}{c_2} \right) \)

Stability matrix: \( M(Y_1, Y_2) = \begin{pmatrix} c_1X - c_2Y_2 & -c_2Y_1 \\ c_2Y_2 & c_2Y_1 - c_3 \end{pmatrix} \)

1\textsuperscript{st} fixed point: \( M(0, 0) = \begin{pmatrix} c_1X & 0 \\ 0 & -c_3 \end{pmatrix} \)

\( \text{eigenvalues are} \quad c_1X, \quad -c_3 \quad \rightarrow \text{Saddle Point} \)

2\textsuperscript{nd} fixed point: \( M\left( \frac{c_3}{c_2}, \frac{c_1X}{c_2} \right) = \begin{pmatrix} 0 & -c_3 \\ c_1X & 0 \end{pmatrix} \)

\( \text{eigenvalues are} \quad \pm i\sqrt{c_3c_1X} \quad \rightarrow \text{Center} \)

or…. coupled excitatory & inhibitory neurons
II. LOTKA-VOLterra EQUATION

\[
\begin{align*}
\frac{dY_1}{dt} &= c_1 XY_1 - c_2 Y_1 Y_2 \\
\frac{dY_2}{dt} &= c_2 Y_1 Y_2 - c_3 Y_2
\end{align*}
\]

Neutrally stable limit cycles

\[
\lambda \pm = \pm i \sqrt{c_3 c_1 X}
\]
Random matrix theory applied to N-species ecology models (\(N >> 1\))

1. Assume each species in isolation would obey a stable logistic equation with stable eigenvalue -1 then switch on random interactions of either sign

\[
\frac{dx_i}{dt} = x_i(1-x_i) - \sum_{j=1}^{N} B_{ij} x_i x_j; \quad \text{Let } x_i'(t) = x_i(t) - x^* = x - 1
\]

2. \[
\frac{dx_i'}{dt} \approx \sum_{j=1}^{N} A_{ij} x_j' \quad \tilde{x}'(t) \text{ is an N-component vector of species deviations from the logistic fixed point } (x_1^*, x_2^*, \ldots, x_N^*) = (1, 1, \ldots, 1)
\]

3. \(\tilde{A} \approx -\tilde{I} - \tilde{C}\), where \(\tilde{C}\) is an N-component interaction matrix with zero mean for each element and each with standard deviation \(\sigma\)

The spectrum of \(\tilde{C}\) is a uniform distribution of complex eigenvalues in unit circle in the complex plane of radius \(\sigma\sqrt{N}\).

Universal density of states for large \(N\)!

"Girko's Law"

Any ecological system becomes unstable for sufficiently large \(N\)!
Visual stimulus $s(t)$ transferred from retinal neurons $\rightarrow$ LGN $\rightarrow$ V1 region of the visual cortex

Spike rate $r(t)$ depends on orientation of bar moving across the visual field

spike rate tuning curve $r(t)$

signal $s(t)$: orientation in degrees

Pathway from the retina through the lateral geniculate nucleus (LGN) to the primary visual cortex

*Dayan and Abbott: Theoretical Neuroscience*
Random matrix models of the brain (H. Sompolinsky, L. Abbott et alia)

- Random neural connections can be formed during development, with many stochastic attachments of axons and dendrites to other neurons.

- Over time, pruning and strengthening/weakening of connections allow neural circuits to "learn" various functions.

- The spectra and eigenfunctions of completely random neural networks with a mixture of inhibitory and excitatory connections, can describe neural activity during the early stages of development.

Girko’s Law

The simplest models of neural networks assume long range connectivity between individual neurons in the brain, leading to synaptic matrices $M(i,j)$ with statistical properties independent of the separation $r_{ij} = |r_i - r_j|$ in three dimensions.

The eigenvalue spectrum of $M(i,j)$ controls the spontaneous activity and induced response of the network, and much is known when its elements are chosen from simple random matrix ensembles. Vary asymmetry, $M(i,j) \neq M(j,i)$.

What happens when $M(i,j) \sim \exp[-|r_i - r_j| / \xi]$, where $\xi$ is a length spanning spanning as many as 50 neurons? On scales larger than $\xi$, the relevant random matrices are banded about the diagonal.

Will localized eigenfunctions dominate the dynamics?

“all states are localized in 1d disordered systems”

---

Sensory inputs, possibly after a processing step, are sent via feed forward couplings into a circular ring of $N$ neurons. Note that $M(1,2)$ and $M(2,1)$ can not only be unequal, but also of opposite sign, if one direction is excitatory and the other inhibitory.

$\nu_i =$ firing rate deviation from background of the $i^{th}$ neuron in recurrent network

$u_j =$ input firing rate of the $j^{th}$ neuron in the input (feed forward) network

$$\tau \frac{d\nu_i}{dt} = -\nu_i + \tanh \left[ \sum_{j=1}^{N} M_{ij} \nu_j + h_i \right], \quad h_i = \sum_{j=1}^{N} W_{ij} u_j$$

$$\tau \frac{d\nu_i}{dt} \approx -\nu_i + \sum_{j=1}^{N} M_{ij} \nu_j + h_i \quad \text{(linear approximation)}$$
Non-Hermitian neural networks with random excitatory \((M(i,j) > 0)\) and inhibitory \((M(i,j) < 0)\) connections

\[
M = -\sum_{j=1}^{N} \left[ s_j^+ e^g | j \rangle \langle j+1 | + s_j^- e^{-g} | j+1 \rangle \langle j | \right]
\]

\(g\) provides a systematic clockwise \((g > 0)\) or counterclockwise \((g < 0)\) directional bias

Study eigenvalues and eigenvectors of directed, banded non-Hermitian random matrices

\[
M = \begin{pmatrix}
0 & s_1^+ e^g & 0 & \cdots & s_N^- e^{-g} \\
s_1^- e^{-g} & 0 & s_2^+ e^g & 0 \\
0 & s_2^- e^{-g} & \ddots & \ddots & 0 \\
\vdots & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
s_N^+ e^g & \cdots & 0 & s_{N-1}^- e^{-g} & 0
\end{pmatrix}
\]

\(s_j^+, s_j^- = \pm 1, \) indep. random variables;

**Set** \(g = 0\) **for now** \(\rightarrow\) random sign model of J. Feinberg and A. Zee, PRE 59 6433 (1999)
Eigenvalue distribution in the complex plane $\lambda = \lambda_1 + i\lambda_2$

Result of exact diagonalization of $10,000 \times N$ matrices with $N = 5000$ and $g = 0$

How localized are the eigenfunctions??
What does “localization” mean?

Eigenfunctions within circle on right side are highly localized w/real eigenvalues.

Eigenfunctions in an annulus closer to the origin are more extended.

Localization length diverges near the origin:

\[ \lambda = \lambda_1 + i\lambda_2, \quad \xi(\lambda_1, \lambda_2) \sim \frac{1}{(|\lambda_1| + |\lambda_2|)\sqrt{\lambda_1^2 + \lambda_2^2}} \]

Graphs showing distribution of eigenvalues and localization patterns.
What is the effect of the bias parameter $g$?

$$M = \begin{pmatrix} 0 & s_1^+ e^g & 0 & \ldots & s_N^+ e^g \\ s_1^- e^{-g} & 0 & s_2^+ e^g & 0 \\ 0 & s_2^- e^{-g} & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & s_{N-1}^+ e^g \\ s_N^+ e^g & \ldots & 0 & s_{N-1}^- e^{-g} & 0 \end{pmatrix}$$

$s_j^+ = \pm 1, \quad s_j^- = \pm 1$ with equal probability

$0 \leq g < \infty \quad$ (no Dale's law for now)

As $g$ increases from 0, it tunes down the amount of feedback in a “feed clockwise” recurrent network...

Similar layered neural nets used for image & sound classification, etc. in machine learning algorithms.

Many layers $\rightarrow$ “deep learning”
Effect of a directional bias around the chain (g > 0)

\[ N = 5000, \quad g = 0.0 \]
Effect of a directional bias around the chain ($g > 0$)

$N = 5000$, $g = 0.002$
Effect of a directional bias around the chain ($g > 0$)

$N = 5000, \ g = 0.01$
Effect of a directional bias around the chain ($g > 0$)

$N = 5000, \ g = 0.05$

As the network becomes increasing feed-clockwise, a gap or hole appears in the eigenvalue spectrum in the complex plane...
Effect of a directional bias around the chain ($g > 0$)

$N = 5000, \ g = 0.1$
Effect of a directional bias around the chain ($g > 0$)

$N = 5000, \ g = 0.2$
Effect of a directional bias around the chain ($g > 0$)

$N = 5000, \quad g = 0.5$

These states move around the ring – they delocalize…
Localization lengths and effect of boundary conditions

Define inverse participation ratio

\[ IPR \equiv \sum_{j} |\phi_j|^4 / \sum_{j} |\phi_j|^2 \]

\[ IPR \sim \text{inverse localization length} \]

extended state: \( \phi_j \sim 1 / \sqrt{N}, \forall j \)

\[ IPR \equiv \sum_{j} |\phi_j|^4 / \sum_{j} |\phi_j|^2 \sim 1 / N \ll 1 \]

localized state, \( \phi_j \sim \exp[-|x_j - x_0|/\xi_{loc}] \)

\[ IPR \equiv \sum_{j} |\phi_j|^4 / \sum_{j} |\phi_j|^2 = O(1) \]

Eigenvalue spectrum for \( g = 0 \)
(or, for any \( g \) with open boundary conditions!)

\[ \text{Localization length diverges on the rim of the hole when } g > 0 \rightarrow \text{extended states} \]

Eigenvalue spectrum for \( g = 0.1 \)
with periodic boundary conditions
What about Dale’s law? All neurons must be purely excitatory or inhibitory....

\[ M = -\sum_{j=1}^{N} \left[ s_j^+ e^g |j+1\rangle \langle j| + s_j^- e^{-g} |j\rangle \langle j+1| \right] \]

Replace 2N random variables with only N of them...

\[ G = -\sum_{k=1}^{N} \sigma_k \left[ e^g |k+1\rangle \langle k| + e^{-g} |k-1\rangle \langle k| \right] \]

The spectra and eigenfunctionsof M and G are essentially identical! The spectral properties are determined in both cases by above/below diagonal products such as 

\[ M(j, j+1) \cdot M(j+1, j) = s_j^+ s_j^- \] and \[ G(j, j+1) \cdot G(j+1, j) = \sigma_j \sigma_{j+1} \]

which have identical statistics!!
Large g spectra: perturbation theory about a “delay ring”

\[ M = \begin{pmatrix}
0 & s_1^+ e^g & 0 & \ldots & s_N^+ e^{-g} \\
0 & s_1^- e^{-g} & 0 & \ldots & 0 \\
0 & 0 & s_2^+ e^{-g} & \ddots & \ddots \\
\vdots & \vdots & \ddots & \ddots & 0 \\
s_N^- e^g & \ldots & 0 & s_{N-1}^+ e^{-g} & 0
\end{pmatrix} \]

After a similarity transformation, 
\( M \rightarrow M' = A + B \)

\[ N = 4 \]
\[ A = e^g \begin{pmatrix}
0 & 0 & 0 & b_4 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0
\end{pmatrix} \cdot \begin{pmatrix}
0 & 0 & 0 & b_4 \\
0 & 0 & 0 & 0 \\
0 & b_2 & 0 & 0 \\
0 & 0 & b_3 & 0
\end{pmatrix} \]

Eigenvectors of A are plane waves
\[ \nu_j^R = e^{ik_j}, \quad \nu_j^L = e^{-ik_j} \]
\[ k = 2\pi n/N, \quad n = 0, 1 \ldots (N - 1) \]

\[ \lambda_k = e^{g+ik} + e^{-g}[e^{-ik}(b_1 + b_2 + \ldots + b_{n-1} + b_N)/N] = \text{random walk} \]
Large $g$ limit: Plane wave states, all eigenfunctions delocalized

- Trajectories of eigenvalues for $N=100$ and values of $g$ decreasing from 1 down to zero.
- Eigenvalues "flow" in the complex plane.
- Motion stops once eigenvalues localize.
The gap rimmed by extended states is robust…

\[ s^+_j = \pm 1, \quad s^-_j = \pm 1, \quad N=1000, \]
\[ g = 0.1, \text{ but with diagonal randomness} \]
Energy gap and rings of extended states also appear for coupled neural clusters.

1000 triangular neural clusters, obeying Dale’s law, and coupled together to form a ring.

“Band theory” for neural networks?

Layered neural network with tunable back propagation.
Sparse non-Hermitian random matrices and a model of short term memory

Recent proposals have suggested that recurrent networks could store information in their transient dynamics, even if the network does not have information-bearing attractor states. Downstream readout networks can then be trained to extract relevant functions of the past input stream to guide future actions.

A useful analogy is the surface of a liquid. Even though this surface has no attractors, (save the trivial one in which it is flat), transient ripples on the surface can nevertheless encode information about when and where past objects that were thrown in.

Memory traces in dynamical systems
S. Ganguli, D. Huh & H. Sompolinsky
PNAS, 105, 18970 (2008)
Fisher memory curves for neural nets

\[ x_i(n) = \sum_{j=1}^{N} M_{ij} x_j(n-1) + v_i s(n) + \xi_i(n); \quad \langle \xi_i(n-k) \xi_j(n-k') \rangle = \varepsilon \delta_{ij} \delta_{kk}, \]

where \( x_i(n) \) is the firing rate of a neuron placed at the \( i \)th site at discrete time \( n \), \( v_i \) is the feedforward connection between a signal input \( s(n) \) and the \( i \)th neuron, and \( \xi_i(n) \) represents Gaussian white noise.

- Ability to recover the signal \( s(n) \) downstream depends on the properties of the matrix \( M_{ij} \)
- What happens when the excitatory and inhibitory connections in the recurrent network of neurons in, say, the higher auditory cortex are chosen at random? Can we improve the efficiency by adjusting the \( \{v_j\} \)?
- Do localized eigenvalues and eigenfunctions of \( M_{ij} \) play a role?
Conversion of temporal to spatial information.

Three scalar signals: a base signal, $s(k)$, and two more signals obtained by perturbing $s$ by the addition of an identical pulse centered at time $n-10$ and $n-20$. At time $n$, the temporal structure of each signal is encoded in the spatial distribution of the network state $x(n)$.

As both perturbations recede into the past, both perturbed memory traces decay, and the three distributions become identical.

Given a probability distribution of firing rates $\tilde{x}(n)$ at time $n$ in a recurrent network $P(\tilde{x}(n) \mid \tilde{s})$ conditioned on a signal history $\tilde{s} = \{s(n-k) \mid k \geq 0\}$, the Fisher memory curve $J(k)$ can be extracted from the Kullback-Leibler divergence between $P(\tilde{x}(n) \mid \tilde{s})$ and $P(\tilde{x}(n) \mid \tilde{s} + \delta\tilde{s})$ as

$$J(k) = \left\langle \frac{-\partial^2}{\partial s_k \partial s_k} \ln P(\tilde{x}(n) \mid \tilde{s}) \right\rangle_{P(\tilde{x}(n)\mid \tilde{s})}$$

$J(k)$, the Fisher memory curve, gives the information retained about a pulse entering the network $k$ steps in the past.

$$J_{tot} = \sum_{k=1}^{\infty} J(k)$$
Kulhbeck-Leibler Information

\[ D_{KL}(p(x|s)||p(x|s + \delta s)) = - \int_{-\infty}^{\infty} dx \ p(x|s) \log \frac{p(x|s + \delta s)}{p(x|s)} \]

Expand for small $\delta s$...

\[ D_{KL}(p(x|s)||p(x|s + \delta s)) = - \int_{-\infty}^{\infty} dx \ p(x|s) \log \frac{p(x|s + \delta s)}{p(x|s)} \]
\[ = - \int_{-\infty}^{\infty} dx \ p(x|s) \log \left[ 1 + \frac{1}{p(x|s)} \frac{\partial p(x|s)}{\partial s_i} \delta s_i + \frac{1}{2 p(x|s)} \frac{\partial^2 p(x|s)}{\partial s_i \partial s_j} \delta s_i \delta s_j + O(\delta s^3) \right] \]

\[ J_{ij} = \left\langle - \frac{\partial^2}{\partial s_i \partial s_j} \log p(x(n)|s) \right\rangle_{p(x(n)|s)} . \]
Fisher memory curves \& 1d sparse random matrices

\[ J(k) \equiv J_{kk} = \nu^T M^{kT} C_n^{-1} M^k \nu \]

\[ C_n = \varepsilon \sum_{k=0}^{\infty} \left[ M^k M^{kT} \right] \]

*For normal networks, (i.e., \([W,W^T] = 0\)),

\[ J(k) = \sum_{j=1}^{N} \nu_j^2 |\lambda_j|^{2k} \left(1 - |\lambda_j|^2 \right) \]

& dependence on eigenvectors of \(W\) drops out;

also, \(J_{tot} = \sum_{k=1}^{\infty} J(k) = 1\)

*However for non-normal networks, we must optimize \(J_{tot} = \nu^T J^s \nu\), where the spatial Fisher memory matrix is \(J_{ij}^s = \sum_{k=0}^{\infty} \left[ M^{kT} C_n^{-1} M^k \right]_{ij}; J_{tot} \leq N\).

Maximize \(J_{tot}\) with the principal eigenvector of \(J_{ij}^s\).

\(J_{ij}^s\) is symmetric with *real* eigenvalues.

What is the role of localization?
Example: delay ring vs. delay line

If the connectivity matrix $W$ is non-normal, then $J_{\text{tot}}$ depends on the feedforward connection $v$. Following Ganguli et al, we compare $J_{\text{tot}}$ for

1. Without optimization, where each component of $v$ is drawn from a uniform distribution $[0, 1]$ and then normalized and

2. With optimization, where we use the principal eigenvector of $J_{\text{ij}}$.

$$M = \begin{pmatrix}
0 & 0.99 & 0 & \ldots & 0 \\
0 & 0 & 0.99 & 0 \\
0 & 0 & \ddots & \ddots & 0 \\
\vdots & 0 & \ddots & 0.99 \\
0.99 & \ldots & 0 & 0 & 0
\end{pmatrix}$$

delay ring is a normal matrix

delay ring is non-normal
Fisher memory curves and localization in one-dimension non-Hermitian random matrices

How does localization manifest itself?

The difference between $J_{tot}$ with optimal and non-optimal feedforward connections, consistent with non-normality of the recurrent network when $f = 0.5$ and $a = 0.0$
Non-Hermitian Localization in Neural Networks

- Non-Hermitian matrices, with complex eigenvalue spectra, arise naturally in simple models of complex ecosystems, and neural networks.

- Striking departures from the conventional wisdom about localization arise in the one-dimensional non-Hermitian random matrices that describe sparse neural and ecological networks.

- An intricate eigenvalue spectrum controls the spontaneous activity and induced response. Directed rings of neurons lead to a hole centered on in the density of states in the complex plane.

- All states are extended on the rim of this hole, while the states outside the hole are localized.

Thank you!