

# Equations from Harris et al. (2023)

## I. Introduction

This section introduces the study of spatiotemporal dynamics in large populations of neurons. It highlights the difficulty of understanding the relationship between network connectivity and dynamics. The authors propose using **Random Matrix Theory (RMT)** to study the stability of the network Jacobian's eigenspectrum. The key innovation is examining the combined effects of **sparsity** (probability  $\alpha$  of connection) and **Dale's Law** (separate excitatory and inhibitory populations) on network stability, generalizing previous results for fully connected or single-population sparse networks.

## II. Network Model and Analysis

The paper describes neural network dynamics using a rate model. The network state is governed by:

$$\dot{x}_i(t) = -\frac{x_i(t)}{\tau} + \sum_{j=1}^N w_{ij}\phi(x_j(t)) \quad (1)$$

where  $x_i(t)$  is the current,  $\tau$  is the time constant,  $W$  is the connectivity matrix, and  $\phi$  is the firing rate function. The equilibria  $\mathbf{x}^*$  are solutions to:

$$\mathbf{x}^* = \tau W \phi(\mathbf{x}^*) \quad (2)$$

For networks obeying Dale's Law, a homogeneous equilibrium requires equal row sums:

$$\sum_{j=1}^N w_{ij} = N\mu_r \quad (3)$$

The homogeneous equilibrium  $x_0^*$  satisfies:

$$x_0^* = \tau N\mu_r \phi(x_0^*) \quad (4)$$

Stability is analyzed via the Jacobian of the linearized system:

$$J(\mathbf{x}^*) = \left[ -\frac{1}{\tau} \mathbb{I}_N + W \Phi'(\mathbf{x}^*) \right] \quad (5)$$

### III. Eigenvalue Spectral Properties

This section investigates how Dale’s law, E-I imbalance, and sparsity affect the Jacobian’s eigenspectrum.

#### A. Eigenvalues of a random matrix

Refers to Girko’s circular law for i.i.d. matrices (radius  $\mathcal{R} = \sigma\sqrt{N}$ ) and the existence of an outlier  $\lambda_O = \mu N$  if the mean is nonzero.

#### B. Sparse random matrices

A general class of sparse random matrices is defined, incorporating a deterministic component  $M$  (low-rank structure) and a random component  $AD$ , masked by sparsity matrix  $S$ :

$$W = S \circ (AD + M) \quad (6)$$

Here,  $\circ$  denotes the Hadamard product. The mean connectivity weights relate to the normalized mean  $\tilde{\mu} = \mu/\sqrt{N}$ , normalized standard deviation  $\tilde{\sigma} = \sigma/\sqrt{N}$ , and sparsity  $\alpha$ :

$$\mathbb{E}(w_{ij}) = \alpha\tilde{\mu} \quad (7)$$

This predicts the location of the eigenvalue outlier:

$$\lambda_O = \alpha\tilde{\mu}N \quad (8)$$

The variance of the weights is derived as:

$$\text{Var}(w_{ij}) = \alpha(1 - \alpha)\tilde{\mu}^2 + \alpha\tilde{\sigma}^2 \quad (9)$$

Consequently, the radius of the eigenspectral disc for sparse networks scales non-linearly with sparsity:

$$\mathcal{R} = \sqrt{N[\alpha(1 - \alpha)\tilde{\mu}^2 + \alpha\tilde{\sigma}^2]} \quad (10)$$

Notably, unlike fully connected networks, the radius depends on population **means** ( $\tilde{\mu}$ ) in addition to variances.

### C. Eigenvalues of sparse random matrices obeying Dale's law

The model is extended to distinct excitatory ( $f$ ) and inhibitory ( $1-f$ ) populations. The standard deviation matrix  $D$  becomes:

$$D = \text{diag}(\underbrace{\tilde{\sigma}_e, \dots, \tilde{\sigma}_e}_{Nf \text{ times}}, \underbrace{\tilde{\sigma}_i, \dots, \tilde{\sigma}_i}_{N(1-f) \text{ times}}) \quad (11)$$

The deterministic low-rank perturbation vectors  $\mathbf{u}$  and  $\mathbf{v}$  are defined as:

$$\mathbf{u} = (1, \dots, 1)^\top, \quad \mathbf{v} = (\underbrace{\tilde{\mu}_e, \dots, \tilde{\mu}_e}_{Nf \text{ times}}, \underbrace{\tilde{\mu}_i, \dots, \tilde{\mu}_i}_{N(1-f) \text{ times}})^\top \quad (12)$$

### D. Implications on network dynamics

#### D. Eigenspectral properties of sparse random matrices that obey Dale's law

The mean and variance of the weights in the matrix  $W$  are derived by combining the sparse component ( $S$ ) with the population specific statistics:

$$\mathbb{E}(W) = f\mu_{se} + (1-f)\mu_{si} \quad (13)$$

$$\mathbb{V}ar(W) = f\sigma_{se}^2 + (1-f)\sigma_{si}^2 \quad (14)$$

where the statistics for the excitatory ( $k = e$ ) and inhibitory ( $k = i$ ) populations are defined in terms of the normalized means  $\tilde{\mu}_k$ , normalized variances  $\tilde{\sigma}_k$ , and sparsity  $\alpha$ :

$$\mu_{sk} = \alpha\tilde{\mu}_k \quad (15)$$

$$\sigma_{sk}^2 = \alpha(1-\alpha)\tilde{\mu}_k^2 + \alpha\tilde{\sigma}_k^2 \quad (16)$$

Using these, the location of the eigenvalue outlier  $\lambda_O$  and the radius of the eigenspectral disc  $\mathcal{R}$  are:

$$\lambda_O = N[f\mu_{se} + (1-f)\mu_{si}] \quad (17)$$

$$\mathcal{R} = \sqrt{N[f\sigma_{se}^2 + (1-f)\sigma_{si}^2]} \quad (18)$$

**1. Nonuniform spectral density of eigenvalue distribution** The spectral density of sparse, balanced (or unbalanced) matrices is nonuniform and depends on the difference in variances between the populations. It is defined as:

$$\rho_{RA}(z) = \begin{cases} \frac{1}{\pi N \sigma_{si}^2} \left[ 1 - \frac{g}{2} \mathcal{H}_f \left( g \frac{|z|^2}{N \sigma_{si}^2} \right) \right] & |z| \leq \mathcal{R} \\ 0 & |z| > \mathcal{R} \end{cases} \quad (19)$$

where the asymmetry parameter  $g$  is:

$$g = 1 - \sigma_{si}^2 / \sigma_{se}^2 = 1 - \frac{(1 - \alpha) \tilde{\mu}_i^2 + \tilde{\sigma}_i^2}{(1 - \alpha) \tilde{\mu}_e^2 + \tilde{\sigma}_e^2} \quad (20)$$

and the helper function  $\mathcal{H}_f(x)$  is:

$$\mathcal{H}_f(x) = \frac{2f - 1 + x}{\sqrt{1 + x(4f - 2 + x)}} + 1 \quad (21)$$

Alternatively, the density can be expressed in terms of precisions  $P_{sk} = 1/\sigma_{sk}^2$ . Let  $\Delta f = 2f - 1$ :

$$\rho(z) = \begin{cases} \frac{1}{2\pi N} [\Sigma P - \Delta P \mathcal{H}_{\Delta f}(\Delta P |z|^2)] & |z| \leq \mathcal{R} \\ 0 & |z| > \mathcal{R} \end{cases} \quad (22)$$

where:

$$\mathcal{H}_{\Delta f}(x) = \frac{x - \Delta f N}{\sqrt{(x - \Delta f N)^2 + N^2(1 - \Delta f^2)}} \quad (23)$$

**2. Local eigenvalue outliers: A zero row-sum condition** We observe that a small number of “local” eigenvalues can escape the spectral disc. To control these, previous work on fully connected matrices utilized a projection operator.

The matrix  $W$  is conceptually decomposed into a random component and a deterministic low-rank structure  $M = \mathbf{u}\mathbf{u}^\top$  which encodes the excitatory/inhibitory population means. For fully connected balanced matrices, the outlier  $\lambda_O = 0$ , and the stability is determined by the random radius. Outliers are controlled by a projection operator  $P$ :

$$P = \mathbb{I}_N - \frac{\mathbf{u}\mathbf{u}^\top}{N} \quad (24)$$

such that the connectivity matrix is:

$$W = ADP + \mathbf{u}\mathbf{v}^\top \quad (25)$$

This ensures the random component has a zero row sum:

$$ADPu = \mathbf{0} \quad (26)$$

And the eigenvector property is preserved for the full matrix:

$$(ADP + \mathbf{u}\mathbf{v}^\top)\mathbf{u} = \mathbf{0} + \mathbf{u}\mathbf{v}^\top\mathbf{u} = (\mathbf{v} \cdot \mathbf{u})\mathbf{u} \quad (27)$$

The left eigenvectors  $l_k$  of  $ADP$  satisfy:

$$l_k\mathbf{u}\mathbf{v}^\top = \mathbf{0} \quad (28)$$

Thus sharing eigenvalues with the full system:

$$l_k(ADP + \mathbf{u}\mathbf{v}^\top) = \lambda_k l_k \quad (29)$$

**Balance vs. Imbalance:** In the **balanced** case,  $\mathbf{v} \cdot \mathbf{u} = 0$ , so the projection naturally kills the low-rank term. However, in the **unbalanced** case ( $\mathbf{v} \cdot \mathbf{u} \neq 0$ ), applying  $P$  to the entire matrix  $W$  would enforce a global zero row sum, inadvertently removing the structural imbalance (the  $\lambda_O$  outlier) that defines the network's state. To retain the imbalance while removing local random outliers, the projection  $P$  must be applied **only** to the random component  $AD$ , leaving the mean structure  $\mathbf{u}\mathbf{v}^\top$  intact.

**3. Local eigenvalue-outliers: A sparse zero row-sum condition** For sparse matrices, the algebraic projection operator  $P$  defined above is problematic: it calculates updates based on all columns, which “fills in” the zero-entries of  $W$ , destroying the sparsity pattern  $S$ . To ensure the sparsity pattern is preserved, we instead enforce a ZRS numerically by subtracting the row-average from *only* the nonzero elements.

This is succinct defined as:

$$W = S \circ (AD + \mathbf{u}\mathbf{v}^\top) - B \quad (30)$$

where  $B_{ij} = S_{ij}\bar{W}_i$ , and  $\bar{W}_i$  is the average of the non-zero elements in row  $i$ :

$$\bar{W}_i = \sum_j W_{ij} / \sum_j S_{ij} \quad (31)$$

For sparse unbalanced matrices, a **partial SZRS** can be applied to just the random component  $J = AD$ :

$$\bar{J}_i = \sum_j J_{ij} / \sum_j S_{ij} \quad (32)$$

**Partial SZRS:** This distinction is critical for **unbalanced, sparse** networks. A full SZRS (Eq 30) would force the row sums to zero, effectively balancing the network and removing the large outlier  $\lambda_O$  caused by the E-I imbalance. By applying the correction only to the interaction of the sparse mask and the random component ( $J = AD$ ), the **Partial SZRS** (Eq 32) controls the radius of the bulk spectrum (removing local outliers) while preserving the global mean structure ( $S \circ \mathbf{uv}^\top$ ) responsible for the primary outlier  $\lambda_O$ .

## IV. Discussion and Conclusions

The paper synthesizes sparsity, Dale’s Law, and structural balance into a unified framework for analyzing neural network stability.

### Key Factors Influencing Stability

1. **Non-linear Radius Scaling:** Unlike fully connected networks where radius refers only to variances, in sparse networks, the radius  $\mathcal{R}$  depends non-linearly on the interactions between sparsity  $\alpha$  and the population means  $\mu_k$ .
2. **Non-uniform Density:** A uniform circular law only holds if the sparse variances satisfy a specific symmetry condition ( $\sigma_{se}^2 = \sigma_{si}^2$ ). In most biological cases (different E/I stats), the density is non-uniform, with eigenvalues concentrating differently depending on the “precisions” of the E and I distributions.
3. **Local Outliers:** “Rogue” eigenvalues can escape the bulk even in balanced networks. The proposed **SZRS** (for balanced) and **Partial SZRS** (for unbalanced) conditions theoretically and numerically eliminate these outliers without disrupting the gross network structure (sparsity) or the intended E/I imbalance.

### Implications for Dynamics

- **Excitatory Dominated:** System diverges; not physiologically relevant for spontaneous activity.
- **Inhibitory Dominated / Balanced:** The transition to chaos is driven by the bulk spectrum crossing the stability line. The non-uniform density implies that fewer eigenvalues lie at the “edge” of the disc compared to the center, leading to less complex dynamics near the transition point than previously thought.

### **Limitations**

The analysis assumes homogeneous equilibria ( $\mathbf{x}^* = 0$ ). Heterogeneous fixed points introduce correlations between the Jacobian and the connectivity matrix that require dynamical mean-field theory (DMFT) to analyze. Additionally, the model assumes instantaneous synaptic transmission (rate model) rather than conductance-based or spiking dynamics.