From synaptic interactions to collective dynamics in random neuronal networks models: critical role of eigenvectors and transient behavior

E. Gudowska-Nowak[‡], M. A. Nowak[‡], D.R. Chialvo[†], J. K. Ochab[‡], and W. Tarnowski[‡]

[‡]*Marian Smoluchowski Institute of Physics and Mark Kac*

Complex Systems Research Center, Jagiellonian University,

S. Łojasiewicza 11, PL 30-348 Kraków, Poland. and

[†]*Center for Complex Systems & Brain Sciences (CEMSC³),*

Escuela de Ciencia y Tecnología, Universidad Nacional de San Martín,

& Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET),

25 de Mayo 1169, San Martín, (1650), Buenos Aires, Argentina.

(Dated: May 10, 2018)

1

Abstract

The study of neuronal interactions is currently at the center of several neuroscience big collaborative projects (including the Human Connectome, the Blue Brain, the Brainome, etc.) which attempt to obtain a detailed map of the entire brain matrix. Under certain constraints, mathematical theory can advance predictions of the expected neural dynamics based solely on the statistical properties of such synaptic interaction matrix. This work explores the application of free random variables (FRV) to the study of large synaptic interaction matrices. Besides recovering in a straightforward way known results on eigenspectra of neural networks, we extend them to heavy-tailed distributions of interactions. More importantly, we derive analytically the behavior of eigenvector overlaps, which determine stability of the spectra. We observe that upon imposing the neuronal excitation/inhibition balance, although the eigenvalues remain unchanged, their stability dramatically decreases due to strong non-orthogonality of associated eigenvectors. It leads us to the conclusion that the understanding of the temporal evolution of asymmetric neural networks requires considering the entangled dynamics of both eigenvectors and eigenvalues, which might bear consequences for learning and memory processes in these models. Considering the success of FRV analysis in a wide variety of branches disciplines, we hope that the results presented here foster additional application of these ideas in the area of brain sciences.

I. INTRODUCTION

Contemporary neuroscience is obsessed nowadays with detailed studies of the neuronal connections across the entire human brain. Large scale collaborative efforts [1] including the BRAIN Initiative in the USA, Brainome in China, and the BlueBrain in the European Union were launched with the objective of mapping at different resolution the connectivity of the entire brain. At a certain point theory will be desperately needed to analyze these very large maps, describing the adjacency matrix of the brain. The present work is an attempt to enter in this uncharted and challenging territory.

Under certain constraints, mathematical theory can advance predictions of the expected neural dynamics based solely on the statistical properties of their synaptic interaction matrix. In that sense randomly connected networks of neurons are one of the classical tools of theoretical neuroscience. Only recently it was observed that the fact that the matrix of the synaptic connections is not normal (i.e., the matrix does not commute with its transpose), has dramatic consequences for the temporal dynamics of stochastic equations, which can mimic the dynamics of the network [2–4]. In particular, the work of Marti et. al [5] shows that increasing the symmetry of the connectivity leads to a systematic slowing-down of the dynamics and *vice versa,* decreasing the symmetry of the matrix leads to the speeding of the dynamics. This non-normality of the matrix not only forces matrices to have complex spectrum (which challenges several traditional tools of random matrix theory), but more importantly, its study sheds new light on the role of Bell-Steinberger [6] matrix of overlaps between the left and right eigenvectors of the connectivity matrix.

Contemporarily, the pivotal role of overlaps is understood in the simplest case of the evolution of complex Ginibre matrix - either in Smoluchowski-Fokker-Planck formalism [7, 8] or in Langevin formalism [9], following the pioneering paper [10, 11]. The effects of the overlaps of the Ginibre matrix for the temporal autocorrelation function of randomly connected networks was addressed analytically in the latest paper [5], confirming the numerical simulations in the weakly coupled regime of synaptic models.

In this paper we study the non-normality aspects of the popular model with excitatory-inhibitory structure [12–14], proposed by Rajan and Abbott in [15]. An important ingredient of this model is the introduction of the balanced condition, which stabilize the wildly fluctu-ating spectra of the network. In a later paper [16], the numerical study of the full non-linear

dynamics in the Rajan-Abbott model has shown the emergence of a transition leading to synchronized (stationary or periodic) states. This phenomenon cannot be explained solely by the spectral features of the connectivity matrix, which motivates our study of missing non-spectral properties of non-normal networks, such as sensitivity to perturbations and transient dynamics induced by non-orthogonality of eigenvectors. Recently, it was hypothesized that the non-normality is universal in real complex networks [17].

Free random variables (hereafter FRV) theory is a relatively young mathematical theory, originating from the works of Voiculescu [18]. Partly due to the connection with large random matrices, it made in last decade a huge impact on physics [19], statistical inference [20], engineering of ICT technologies [21] and finances [22]. In brief, FRV can be viewed as a non-commutative probability theory for Big Data problems, where the information is hidden in statistical properties of eigenvalues and eigenvectors. As such, it is ideally suited for disentangling signals from noise in various kinds of complex systems. Another advantage comes form the fact, that at the operational level the formalism is simple and powerful, allowing very often to get results on the basis of "back-of-envelope" calculations.

From this perspective, it is rather bewildering, that FRV so far has not been broadly applied to the most challenging complex problem, i.e., the understanding of the brain. Thus, in this paper we consider FRV applications to understand the neuronal networks as represented by the synaptic strength matrix. Direct application of FRV not only allows us to recover in a straightforward way well known results in the literature [15], but also to address quantitatively such issues as stability of the network with respect to perturbation and extension the existing formalisms for the heavy-tailed distributions.

The paper is organized as follows. In Sec. II we discuss two important effects caused by the non-orthogonality of eigenvectors of non-normal matrices, namely high sensitivity of the spectra and the transient behavior of the linearized dynamics. We briefly describe free probability theory in Section III, showing how it allows one to calculate the spectral density and gives an access to the eigenvector non-orthogonality. In Section IV we reframe the model introduced by Rajan and Abbott in this language. Applying the theoretical toolbox explained in Appendices we recover and generalize their main results for the unbalanced network. In doing so, we uncover the analytic formulas for the one-point eigenvector correlation function for this model, crucial for the determining its stability. Since free random variables work also in the case of heavy, spectral tails [23], we also present results for the

spectra and eigenvectors of the Rajan-Abbott model adapted for the case of the Cauchy noise. We successfully confirm our analytic predictions with numerical simulations.

Further, in Section V, we show explicitly that the excitation/inhibition balance condition, not only tames the spectral outliers, but also exerts dramatic effects on non-orthogonality of eigenvectors, increasing the networks' eigenvalue condition number by several orders of magnitude. Section VI closes the paper with a summary of the main results and its implications. It also outlines main promising directions for further studies using the presented formalism.

II. NON-NORMALITY OF SYNAPTIC INTERACTIONS IN NEURAL NETWORKS

Adjacency matrices of directed networks and synaptic strength matrices are non-normal. This influences not only their spectra, as the eigenvalues can be complex, but also has a strong effect on the eigenvectors. A diagonalizable non-normal matrix possesses two eigenvectors: left and right for each eigenvalue. They satisfy eigenproblems

$$\langle L_i | \, X = \langle L_i | \, \lambda_i, \qquad X \, | R_i \rangle = \lambda_i \, | R_i \rangle \, . \tag{1}$$

We use here physicists' "bra-ket notation", where $| R_i \rangle$ is a column and $\langle L_i |$ is a row vector. The scalar product is denoted as $\langle L_i | R_j \rangle$ and we define conjugated left vector $| L_i \rangle = (\langle L_i |)^\dagger$.

Eigenvectors are normalized to $\langle L_i | R_j \rangle = \delta_{ij}$, but they are not orthogonal among themselves $\langle R_i | R_j \rangle \neq \delta_{ij} \neq \langle L_i | L_j \rangle$. Chalker and Mehlig introduced a matrix of scalar products of eigenvectors [10, 11]

$$O_{ij} = \langle L_i | L_j \rangle \, \langle R_j | R_i \rangle \, . \tag{2}$$

Below we describe two phenomena important in neural networks, in which the non-orthogonality of eigenvectors captured in the matrix of overlaps plays an essential role.

A. Synaptic plasticity seen as perturbations of a network

The synaptic strengths of real neuronal networks are not static [24]. Neural activity itself, in the course of time, allows neurons to form new connections, strengthening or weakening the existing synapses. This synaptic plasticity, on which biological learning is based, is not

captured in many models. Nonetheless, the change of the synaptic strengths in a short time interval can be treated as a small additive perturbation of the initial matrix. This results in reorganization of the spectrum on a complex plane.

Considering the perturbation of the matrix X by some ϵP, the change of the spectrum in the first order in ϵ reads

$$\delta\lambda_i = \epsilon \langle L_i| P |R_i\rangle \leq \epsilon\sqrt{\langle L_i|L_i\rangle \langle R_i|R_i\rangle}||P||_F. \tag{3}$$

The inequality follows from the Cauchy inequality and $||P||_F$ denotes the Frobenius norm $||P||_F^2 = \mathrm{Tr}PP^\dagger$. This inequality is saturated (equality holds) by the rank one Wilkinson matrix $P = |L_i\rangle \langle R_i|$. The inequality above shows that spectra of networks represented by non-normal matrices are more sensitive to changes in their connectivity. This enhanced sensitivity is driven by the non-orthogonality of eigenvectors. The quantity $\kappa(\lambda_i) = \sqrt{O_{ii}}$ is known in the numerical analysis community as the eigenvalue condition number [25, 26].

B. Eigenvector non-orthogonality in transient dynamics

Stability analysis and linear response of the dynamic systems with respect to external perturbations are among most popular methods of describing complex systems [36]. Let us consider dynamics obtained from linearization of the system in the vicinity of the fixed point

$$\frac{d}{dt} |\psi\rangle = (-\mu + X) |\psi\rangle + |\xi(t)\rangle. \tag{4}$$

Here ξ represents the external driving. Choosing it as a "spike" $|\xi(t)\rangle = \delta(t) |\psi(0)\rangle$, we formally solve the system for $t > 0$

$$|\psi(t)\rangle = \exp[(X - \mu)t] |\psi(0)\rangle. \tag{5}$$

The long-time dynamics is governed by the eigenvalue with the largest real part. However, if X is non-normal, this analysis is incomplete. The behavior of the linearized dynamics can be drastically different at its early stage. In particular, the system may initially move away from the fixed point. This sometimes invalidates the linear approximation and renders the fixed point unstable, even though the linearized dynamics predicts stability.

To describe this transient dynamics, we consider the squared Euclidean distance from the

fixed point, which is the squared norm of the solution (5)

$$D(t) = \langle \psi(t) | \psi(t) \rangle = e^{-2\mu t} \langle \psi(0) | e^{X^\dagger t} e^{Xt} | \psi(0) \rangle$$
$$= \sum_{i,j=1}^{N} \langle \psi(0) | L_i \rangle \langle R_i | R_j \rangle \langle L_j | \psi(0) \rangle \, e^{-2\mu t + t(\lambda_i + \lambda_j)}. \tag{6}$$

If we consider the "spike" $|\psi(0)\rangle$ as a particular versor on the N-hypersphere (real or complex), averaging over all directions uniformly distributed on the hypersphere leads to

$$\bar{D}(t) = e^{-2\mu t} \frac{1}{N} \mathrm{Tr} e^{X^\dagger t} e^{Xt} = e^{-2\mu t} \sum_{ij} e^{t(\lambda_i + \lambda_j)} O_{ij}. \tag{7}$$

We see that all elements of the overlaps of left and right eigenvectors drive the behavior of the squared distance. First, they enhance the contributions of the eigenmodes, which is responsible for amplification of the response to the external driving. Second, since the matrix is not diagonal, they couple different eigenmodes during the evolution. This results in an interference between eigenmodes, which is reflected as an oscillatory behavior of the squared norm of the solution (see also Fig. 7). Note that for normal matrices, such effects do not exist, since left and right vectors are orthogonal and the "coupling matrix" is an identity. Recently, the transient growth was proposed as an amplification mechanism of neural signals [2, 27, 28].

Usually the matrix X is modeled as random. We remark that the averaging over all initial conditions is equivalent to fixing an initial vector $|\psi(0)\rangle$ and averaging over the vectors $U |\psi(0)\rangle$, where U is uniformly distributed (according to the Haar measure) on the orthogonal (unitary) group. This implies that if a randomness in X is to be compatible with the average over initial conditions, the probability density function has to be invariant under unitary (orthogonal) transformations, $P(X) = P(UXU^\dagger)$.

III. THEORY OF FREE RANDOM VARIABLES

A. Spectral density and eigenvector correlations

Unitarily (and orthogonally) invariant random matrices in the large size limit are described by Free Probability. Its power relies on the easiness of obtaining analytical formulas, which are very good approximations even for a relatively small size of a matrix.

An important class of matrices, the so-called bi-unitarily invariant, which generalizes the Gaussian distribution (described in Sec. III B) is important in the models of neural networks. Despite the fact that they are genuinely non-Hermitian, due to enhanced symmetry the spectral problem is effectively one-dimensional because the spectrum is rotationally invariant on a complex plane. In this case, a powerful result holds in FRV, known as the Haagerup-Larsen theorem [29]. It states that the radial cumulative distribution function $F(r) = \int_0^r 2\pi\rho(r')r'dr'$, of the ensemble X can be inferred from the simple functional equation

$$S_{X^\dagger X}(F(r) - 1) = \frac{1}{r^2} \tag{8}$$

where $S_X(z)$ is so called S-transform for the ensemble X. In Appendix A we explain the probabilistic interpretation of S and we provide a simple example. Spectra of bi-unitarily invariant ensembles in large N limit are supported on either a disc or an annulus, a phenomenon dubbed "the single ring theorem" [30, 31]. The inner radius of the spectrum is deduced from the condition $F(r_{in}) = 0$, while the outer one is given by $F(r_{out}) = 1$.

The applicability of free probability to non-Hermitian matrices is not limited to spectra only. It gives also access to the local averages of the overlap matrix. The one-point function [32]

$$O(z) = \frac{1}{N^2} \left\langle \sum_{i=1}^N \delta^{(2)}(z - \lambda_i) \langle L_i|L_i \rangle \langle R_i|R_i \rangle \right\rangle, \tag{9}$$

associated with the diagonal elements of the overlap matrix can be calculated for any type of unitarily invariant probability [33]. For bi-unitarily invariant ensembles it takes remarkably simple form [34]

$$O(r) = \frac{1}{\pi r^2} F(r)(1 - F(r)). \tag{10}$$

The ratio of the one-point correlation function and the spectral density gives the conditional expectation of the squared eigenvalue condition number [34]

$$\mathbb{E}\left(\kappa^2(\lambda_i)|r = |\lambda_i|\right) = \frac{NO(r)}{\rho(r)}. \tag{11}$$

Recently, the two-point function associated with off-diagonal elements of the overlap matrix has become accessible within free probability [35].

B. Example: Ginibre-Girko ensemble

We conclude this section with an example of the above construction by considering the so called Ginibre-Girko matrix X, the entries of which are independently taken from the real/complex Gaussian distribution with zero mean and $1/N$ variance. According to (8), we need the S-transform for $X^\dagger X$. This matrix belongs to the Wishart ensemble [37]. Its S-transform reads $S_{X^\dagger X}(z) = \frac{1}{1+z}$ (see Appendix A). This completes the calculation, since now replacing $z \to F(r) - 1$ and using (8) we get

$$F(r) = r^2 \qquad (12)$$

The spectrum is therefore uniform, $\rho(r) = \frac{1}{2\pi r}\frac{dF(r)}{dr} = \frac{1}{\pi}$, on the unit disc ($F(r_{in}) = 0$, $F(r_{out}) = 1$), reproducing Ginibre-Girko result. The eigenvector correlator comes from (10), $O(r) = \frac{1}{\pi}(1 - r^2)$, in agreement with [10], where it was calculated using much more laborious techniques. In the next section, we show that the same computational simplicity is preserved when considering the ensembles taking into account physiological restrictions imposed on the neural networks models.

IV. REFRAMING RAJAN-ABBOTT MODEL

The strength of synapses between all pairs of N neurons in a network is represented by the weighted adjacency (synaptic) matrix. Contrary to the Ginibre matrices, the structure of its elements is more complicated. In the minimal model [15], there are two kinds of neurons with a fraction $f_E N$ representing excitatory (E), and $f_I N = (1 - f_E)N$ the remaining inhibitory (I) neurons. Their strengths are sampled from Gaussian ensembles, with means μ_i and variances σ_i^2/N, where $i = I, E$. The matricial representation of the synaptic strength matrix reads therefore $M + W$. Here the deterministic matrix M represents the average synaptic activity. In this model it is a rank one matrix with identical rows, each containing $f_E N$ consecutive means μ_E and followed by $f_I N$ consecutive means μ_I. The random part W models variability across the population. It is assumed to be of the form $W = X\Lambda$, where X is the Girko-Ginibre matrix and Λ is diagonal with its first $f_E N$ elements equal to σ_E and last $f_I N$ ones equal to σ_I.

Several studies [14, 38] show that the amount of excitation and inhibition of a neuron is the same (the so-called E/I balance) even on the scale of few milliseconds [12, 13]. To

incorporate this fact in the model, the balance condition is imposed on two levels. The global condition $f_E \mu_E + f_I \mu_I = 0$ means that neurons are balanced on average. This forces the last non-zero eigenvalue of M to vanish. Even in the case of a null spectrum of M, its non-normal character causes the eigenvalues of $M + X\Lambda$ differ much from that of $X\Lambda$. As a result a few eigenvalues lie far beyond the spectrum of $X\Lambda$ [15], see Fig. 1.

The local E/I balance is imposed on this model by demanding that the sum of strengths coupled independently to each neuron vanishes. Mathematically, the elements within each row sum to zero. This condition brings the outliers back to the disc of radius $R = \sqrt{f_I \sigma_I^2 + f_E \sigma_E^2}$ – now the spectra of W and $W + M$ are identical [15], see also Fig. 1 .

A. Rajan-Abbott results from FRV

Having known that the E/I balance causes the spectrum to be insensitive to the matrix of average strengths M, we consider a more general model of m types of neurons, each with multiplicity $f_k N$ and the synaptic strength variance σ_k^2 / N. The random part of the synaptic strength matrix can be written as $W = X\Lambda$, where X represents Ginibre-Girko ensemble and Λ is the diagonal matrix $\mathrm{diag}(\sigma_1 \mathbf{1}_{f_1 N}, ..., \sigma_m \mathbf{1}_{f_m N})$. The multiplicities are normalized as $\sum_{i=1}^{m} f_i = 1$. In Appendix A, using free probability, we obtain the algebraic equation for the radial cumulative distribution function $F(r)$

$$1 = \sum_{i=1}^{m} \frac{f_i \sigma_i^2}{r^2 - \sigma_i^2 (F(r) - 1)}. \tag{13}$$

Explicit solutions exist for $m = 2, 3, 4$ types of neurons, corresponding to the quadratic, cubic or quartic algebraic equation for $F(r)$, but other cases are easily tractable numerically. The case solved by Rajan and Abbott corresponds to the quadratic equation. Solution (13) is also equivalent to the diagrammatic construction of [39], but more explicit. The spectrum is always confined within the disc of radius $r_{out}^2 = \sum_{i=1}^{m} f_i \sigma_i^2$, as visible from the condition $F(r_{out}) = 1$.

We will argue in Sec. V that the presence of the deterministic matrix M and the balance condition exert a dramatic effect on the eigenvectors of the synaptic strength matrix. Knowing $F(r)$, free probability allows us to calculate via (10) also the eigenvector correlation function $O(r)$ for its random part W. In the case of the minimal model considered by
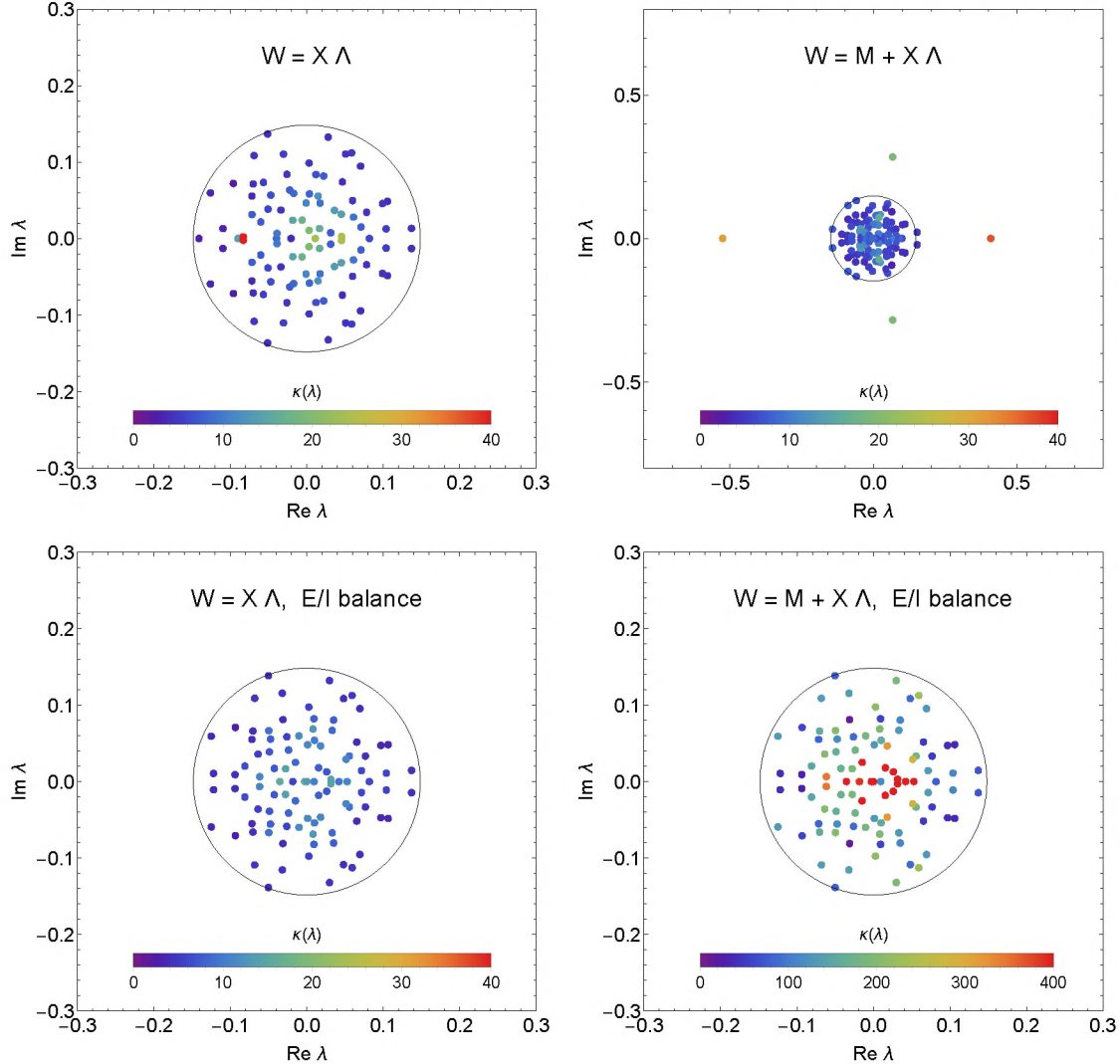
FIG. 1. Eigenvalues and their condition numbers of the matrix of variances $X\Lambda$ (top left), Rajan-Abbott model of neural network (top right), matrix of variances with E/I balance imposed (bottom left) and Rajan-Abbott model with E/I balance (bottom right). The same realization of the Gaussian matrix X was taken for all plots. We used parameters $\sigma_I = 0.3$, $\sigma_E = 0.1$, $f_I = 0.15$, $f_E = 0.85$, $\mu_I = 0.85$, $\mu_E = 0.15$. The matrix is $N = 100$ in size. We observed in many realizations of this scenario that the outliers of the unconstrained Rajan-Abbott model have higher condition number than the average of eigenvalues within the circle. We impose E/I balance by subtracting the $1/N$ of a sum of each row from any element from that row. The spectra in the panels on the left only slightly differ. The eigenvalues presented in bottom panels are exactly the same, but the presence of a highly non-normal matrix M causes that the eigenvalues on the bottom right are much worse conditioned. Note the tenfold ($\sqrt{N}$, as predicted by (20)) broader scale on the bottom right plot.

Rajan-Abbott, it reads explicitly

$$O_W(r) = \frac{1}{2\pi\sigma_E^4\sigma_I^4} \left((f_I - f_E)\sigma_I^2\sigma_E^2(\sigma_E^2\sigma_I^2) - r^2(\sigma_E^4 + \sigma_I^4) + (\sigma_E^2 + \sigma_I^2)\sqrt{K} \right), \qquad (14)$$

where

$$K = r^4(\sigma_E^2 - \sigma_I^2)^2 + \sigma_I^4\sigma_E^4 + 2r^2(f_E - f_I)\sigma_E^2\sigma_I^2(\sigma_E^2 - \sigma_I^2). \qquad (15)$$

This result is inaccessible within the framework of [39].

B.   Heavy-tailed noise

Cauchy noise, belonging to the regime of Lévy stable distributions, is used here as the simplest mechanism to mimic the non-Gaussianity of the realistic synaptic matrices. Since learning rules could change the initial random network structure into a small world network [40–42] by dynamic modification of synaptic weights, possibility of obtaining analytic benchmarks for heavy-tailed distributions is appealing. Spatial and temporal Lévy processes are omnipresent in biological time series, but the fact that they do not possess finite moments invalidates several standard tools of statistical analysis. In the case of matrices exhibiting heavy-tailed distributions of elements, the underlying mathematical structure is quite involved [43, 44]. Surprisingly, also in this case FRV offers a powerful shortcut.

Application of FRV techniques for the heavy-tailed noise with $\alpha = 1$ (spectral Cauchy distribution) leads (see Appendix C) to the simple result

$$\rho(r) = \frac{1}{2\pi r}\frac{dF(r)}{dr} = \frac{1}{\pi}\sum_{i=1}^{m}\frac{f_i}{(r^2 + \sigma_i^2)^2} \qquad (16)$$

$$O(r) = \frac{1}{\pi r^2}F(r)(1 - F(r)) = \frac{1}{\pi}\sum_{i=1}^{m}\frac{f_i}{r^2 + \sigma_i^2}\sum_{j=1}^{m}\frac{f_j\sigma_j^2}{r^2 + \sigma_j^2} \qquad (17)$$

In this case, the spectrum spreads over the whole complex plane, reflecting the large fluctuation of Lévy type noises. In the case of more realistic Lévy noises, one looses the simple analytic structure presented above, but the formalism stays – the resulting equations are usually of transcendental type, but can be easily solved numerically.

V.   NON-NORMALITY IN RAJAN-ABBOTT MODEL

Below we argue that imposing E/I balance not only confines the eigenvalues to a disc, but – more importantly – induces very strong non-orthogonality of eigenvectors. This in

12

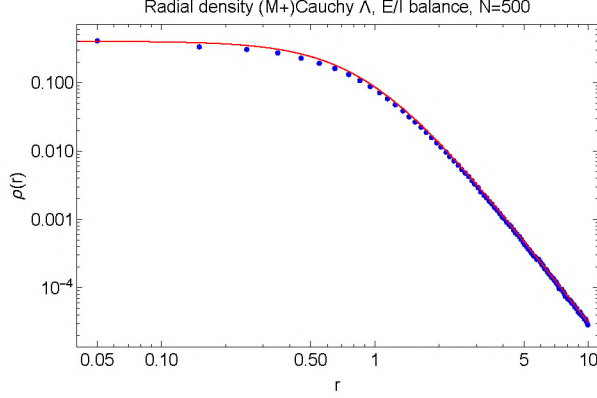Radial density (M+)Cauchy Λ, E/I balance, N=500

FIG. 2. Cross-check of the numerical results with the analytical prediction of the spectral density for the Cauchy synaptic matrix.

turn causes the spectra to be highly sensitive to perturbations and strengthen the transient effects.

Let us assume that the matrix W is diagonalizable. If we denote $|u\rangle = (1, 1, \ldots, 1)^T$, the E/I balance is equivalent to the fact that $|u\rangle$ is the right eigenvector of W to the eigenvalue $\lambda_1 = 0$. Let $\langle L_1|$ be the left eigenvector to this eigenvalue. For brevity we also denote $\langle m| = (\underbrace{\mu_1, \ldots, \mu_1}_{f_1 N \ times}, \ldots, \underbrace{\mu_m, \ldots, \mu_m}_{f_m N \ times})$, which allow us to write $M = |u\rangle \langle m|$. The spectral decomposition of W reads

$$W = 0 \cdot |u\rangle \langle L_1| + \sum_{j=2}^{N} |R_j\rangle \lambda_j \langle L_j| . \tag{18}$$

Since $\langle m|u\rangle = 0$, $\langle m|$ has a decomposition into the left eigenvectors of W, except for $\langle L_1|$, thus $\langle m| = \sum_{j=2}^{N} \langle L_j| \alpha_j$ with $\alpha_j = \langle m|R_j\rangle$. Hence, the total synaptic strength matrix is decomposed as

$$M + W = 0 \cdot |u\rangle \langle L_1| + \sum_{j=2}^{N} \left(|R_j\rangle + \frac{\alpha_j}{\lambda_j} |u\rangle \right) \lambda_j \langle L_j| . \tag{19}$$

We constructed explicitly the eigenvectors of the synaptic strength matrix. The left eigenvectors are not altered when M is taken into consideration due to the E/I balance. The bi-orthogonality condition $\langle L_i|R_j\rangle = \delta_{ij}$ leaves freedom of rescaling each pair of eigenvectors by a non-zero complex number $|R_j\rangle \to c_j |R_j\rangle$ and $\langle L_j| \to \langle L_j| c_j^{-1}$. These transformations allow us to set the length of left eigenvectors $\langle L_j|L_j\rangle = 1$. The diagonal elements of the new

overlap matrix now read

$$O'_{jj} = O_{jj} + 2\text{Re}\left(\frac{\langle m|R_j\rangle\langle R_j|u\rangle}{\lambda_j}\right) + N\frac{|\langle m|R_j\rangle|^2}{|\lambda_j|^2}, \tag{20}$$

where we have used $\langle u|u\rangle = N$. In the large N limit the last term will dominate. This shows that the condition numbers grow with the size of a matrix and the effect of the matrix of averages is stronger for eigenvalues close to the origin.

Analogous reasoning for the full overlap matrix leads to the conclusion that all its elements O_{ij} for $i,j \geq 2$ are affected by the E/I balance and the deterministic matrix. The dominant term in large N is given by

$$O'_{ij} - O_{ij} \sim N\langle L_i|L_j\rangle\frac{\langle m|R_i\rangle\langle R_j|m\rangle}{\lambda_i\bar{\lambda}_j}. \tag{21}$$
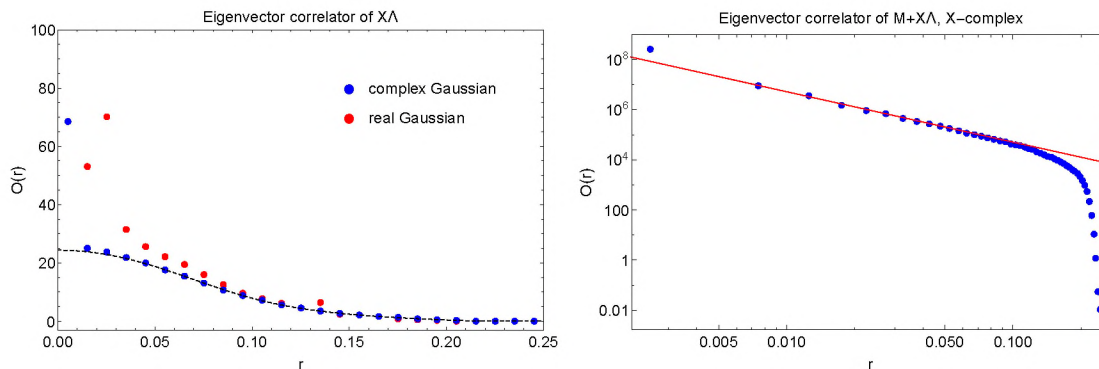

FIG. 3. (left) Eigenvector correlation function for the matrix of variances W with the E/I balance imposed. The random matrix was generated from complex and real Ginibre ensemble. The dashed line presents the analytical solutions from FRV. Numerical results (dots) were obtained by the diagonalization of 1500 matrices of size $N = 1000$. The discrepancies for real matrices come from the real eigenvalues. The fluctuations of the diagonal overlaps associated with them are so strong that the mean of their distribution does not exist [45]. (right) Eigenvector correlator of $M + X\Lambda$, where X is complex Ginibre. The solid line presents the power-law, $O(\lambda) \sim |\lambda|^{-2}$, predicted by (20). In both pictures we took parameters $\sigma_I = 0.4$, $\sigma_E = 0.1$, $f_I = 0.25$, $f_E = 0.75$. For the picture on the right we also set $\mu_E = 0.25$, $\mu_I = 0.75$.

To study the statistics of the eigenvalue condition numbers, we performed numerical simulations by diagonalizing matrices, the random part was generated from either real or complex Ginibre ensemble. The eigenvector correlation function is juxtaposed with (14) from
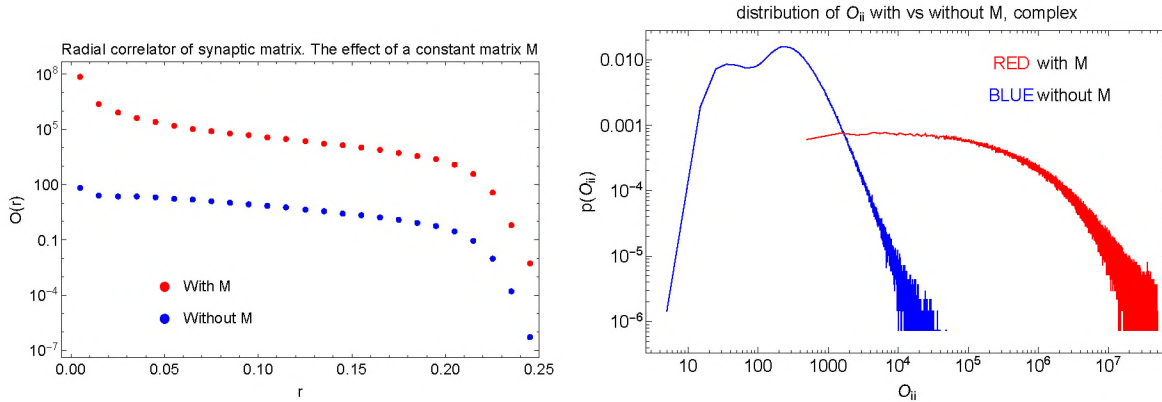
FIG. 4. (left): The eigenvector correlation function of the synaptic strength matrices with E/I balance condition. Despite the spectra are the same, the squared condition numbers differ much. (right) Distribution of squared eigenvalue condition numbers of the eigenvalues of the synaptic strength matrix with and without constant matrix M.

Free Probability, see Fig. 3. The presence of the matrix M and E/I balance is manifested in the scaling $O(r) \sim r^{-2}$ for small r, as observed in Fig. 3, in accordance with (20).

There is a visible mismatch between numerics for real matrices and the results from free probability, particularly evident for eigenvalues with small modulus. This fact is explained in the light of the recent result by Fyodorov [45], where he showed that the distribution of the overlap for Gaussian matrices is heavy-tailed. This distribution conditioned on real eigenvalues of the real Ginibre is so fat-tailed that even the mean does not exist, thus $O(z)$ can be considered only outside the real axis. Being aware of this fact, we have performed further simulations only for complex matrices which do not suffer from this problem.

We studied the effect of the deterministic matrix M by juxtaposing the eigenvector correlation function in Fig. 4 and noticed the significant increase in its magnitude. This enhancement of non-normality is visible not only on the level of mean value, but also on the full distribution of the overlap (see Fig. 4 (right)).

Above conclusions are strengthened by the similar study based on Cauchy synaptic matrices. Figure 2 shows perfect agreement of our predictions with the numerics. By construction of the Rajan-Abbott local condition, spectra are unchanged. This does not hold, however, for the squared eigenvalue condition numbers – they dramatically increase (several orders of magnitude, note the scales in Figures 5 and 6). Finally, the unperturbed eigenvector correlator approaches the predicted slope (compare the predicted slope 4 to the measured 3.84).
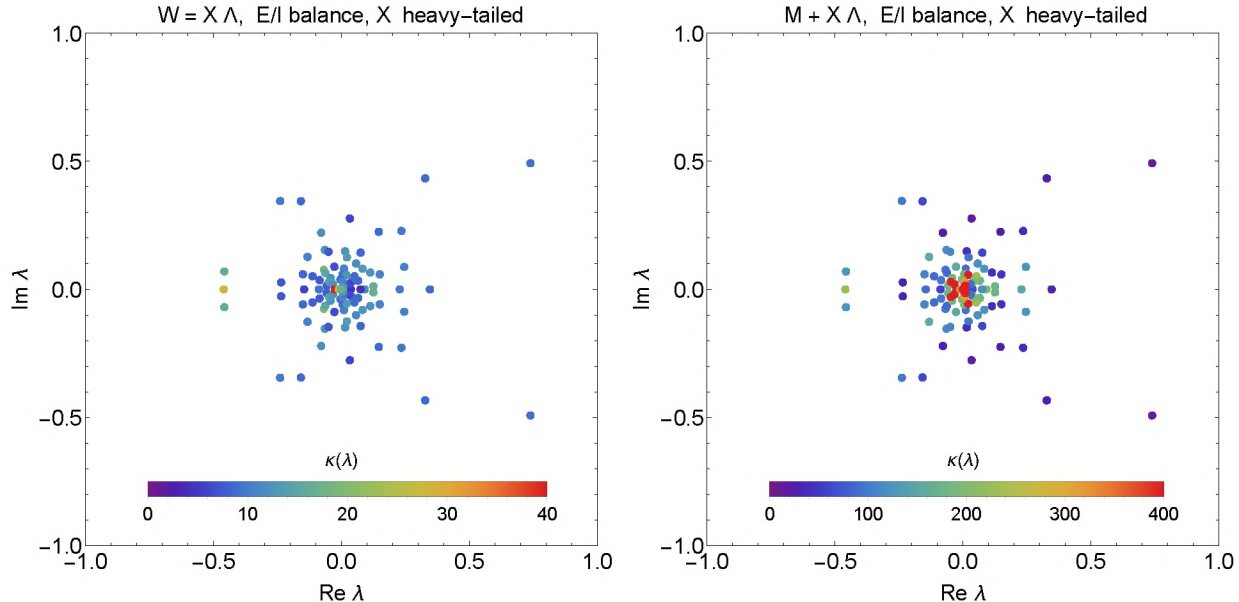
FIG. 5. (left) Eigenvalues and their condition numbers for the synaptic matrix, the random part of which is generated from the matrix Cauchy distribution without (left) and with(right) the deterministic connection, M, reflecting Dale's principle. Note the increase of condition numbers caused by an addition of M (the scale is resized by an order of magnitude). Matrices M and Λ are the same as in Fig. 1.

The perturbed correlator reproduces small r behavior (compare the predicted exponent 2 to the measured 2.03), whereas large r numerical simulations provide asymptotic slope 5.25, as compared to the predicted slope equal to 4.

The deterministic connections and the E/I balance causes an increase of all elements of the overlap matrix O_{ij}, as eq. (21) predicts. To elucidate importance of this fact, we studied the squared norm of the solution to the linearized dynamics (6) with $X = W$ and $X = M + W$. This dynamics is obtained by the linearization of the model considered in [16]. Results presented in Fig. 7 show that the deterministic connections in the network followed by the E/I balance significantly enhance the norm of the solution and the transient trajectories are present for almost all initial conditions. This is not the case if the connections were fully random. Moreover, the wild oscillations of the squared norm indicate the strong interference between the eigenmodes.

One expects these dramatic effects to be visible in the activity of individual neurons. We therefore studied the temporal dynamics of the components of the vector of neural activities
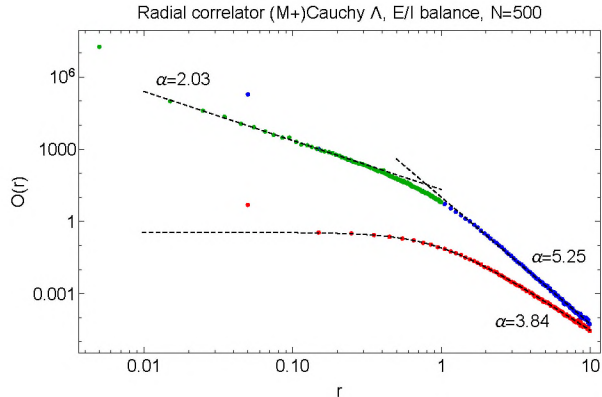
FIG. 6. Radial eigenvector correlator for Cauchy synaptic matrices. Green slope reflects the universal inverse squared behavior, for small r. Red slope approximates the analytic prediction ($\alpha = 4$) for unperturbed model. Blue slope shows the numerical fit to perturbed model.
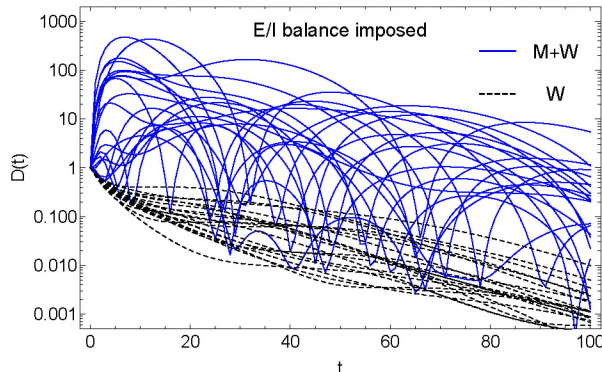

FIG. 7. Squared Euclidean distance from the fixed point in the linearized dynamics of [16]. The presence of M induces strong transient behavior and wild oscillations. These effects are caused by the strong non-normality. Numerical results were obtained for the minimal Rajan-Abbott model. The matrix is of size $N = 100$ with the same parameters as in Fig. 1. We chose $\mu = r_{out} + 0.02$ to ensure stability. Each curve corresponds to a single initial condition generated randomly from the set of vectors of unit norm.

(5) for randomly chosen initial conditions. The results, presented in Fig. 8, show that in the presence of M, the neuronal activity is not only transiently enhanced, but also more synchronized, as observed numerically in the full dynamics in [16]. This effect persistent in the non-linear model is observed as transient in the linearized dynamics.
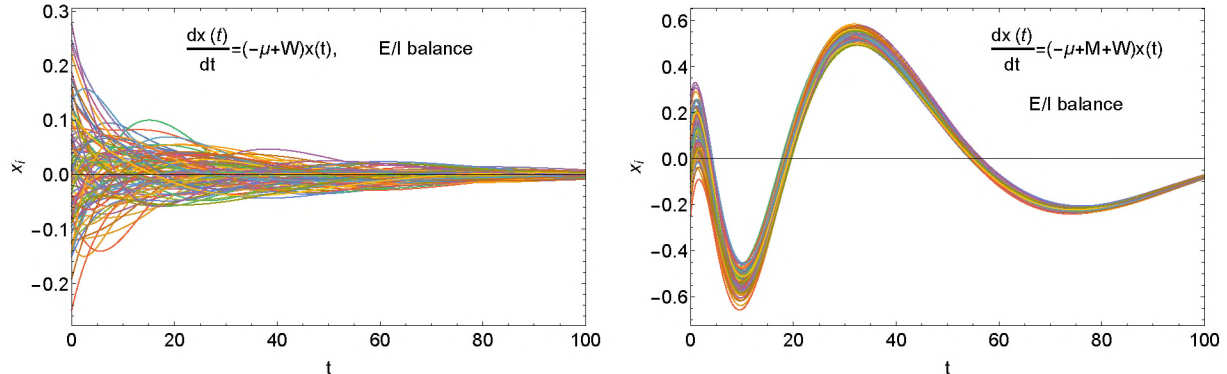
17

FIG. 8. Activity of each neuron in the linearized dynamics. In the right panel we can see the onset of collective dynamics driven by the matrix M and balance condition. Both simulations started from exactly the same initial condition randomly chosen from the N-dimensional hypersphere. Parameters are the same as in Fig. 7.

VI. DISCUSSION AND CONCLUSIONS

In this letter we explored the use of FRV in the study of large synaptic interaction matrices. Beside recovering in a straightforward way known results on the application of random matrices to neural networks, we have addressed the issue of large fluctuations, most probably very relevant to the dynamics of learning and memory in biological neural networks [46]. Using recent results on the properties of eigenvectors in non-normal matrices, we have quantitatively linked the strength of the fluctuation of the outliers to a certain eigenvector correlator. We presented our analysis for the simplest Gaussian case, nevertheless we also pointed the way how one can consider other distributions, e.g., heavy tails. The formalism stays the same, perhaps the only difference (modulo easy Cauchy case which we solved here analytically) is that in the case of more general pdf's one has to rely on numerical solutions.

From the present results it is clear that to understand the temporal evolution of non-normal matrix models requires considering the entangled dynamics of both eigenvectors and eigenvalues, contrary to the simple evolution of the spectra of normal matrices, for which the eigenvectors decouple in the presence of the spectral evolution [7, 8, 47]

Our results indicate, that for the balanced networks the sensitivity of eigenvalues to additive perturbation is dramatic and increases several orders of magnitude in the networks with heavy-tailed spectrum of adjacency matrices (small worlds). Since it is commonly accepted,

18

that spike-timing-dependent plasticity in small-world networks is a hypothetical learning mechanism (for a recent experimental study see [48]), one may worry, how synchronization of the network is possible at all. We emphasize here that the E/I balance is put into this model by hand. In the real brain the E/I balance is maintained on the scale of hundreds of milliseconds [14], and periods during which the balance is violated are not longer than few milliseconds [12, 13]. More complete models of neural networks must incorporate the E/I balance as a dynamical process.

Networks adapting to the changing external conditions may change their structure in a controlled way. The high sensitivity of eigenvalues to these changes in this case might be desired, because it can facilitate the adaptation. We hypothesize that such high sensitivity in the models with dynamical E/I balance can emerge through a process a kind of self-regulated criticality [49]. Although the specifics of such process is not certain yet, there is evidence both empirical[50–53] as well as theoretical [54–56] of its plausibility. In addition, the connection of E/I balance with criticality was already observed at the level of neuronal avalanches analysis in EEG or MEG data [57].

Since the balance condition leads to a dramatic increase of eigenvector overlaps – conditioning the spectra – which further take crucial part in driving mechanisms of temporal evolution of the networks, one needs a powerful, stabilizing mechanism preventing the transition to the chaotic behavior. Such chaotic behavior would imply sudden and drastic reorganization of the eigenvalues leading to unwanted dynamics of the neural network.

We envision one a priori mechanism, which can tame such a behavior – it is the transient behavior. This conclusion is consistent with the model of Molino et al. [16] for non-normal balanced networks, who have observed synchronization inexplicable by solely spectral properties of the networks. Transient behavior means that even stable trajectories may initially diverge before reaching the fixed point for long times. This implies that transient behavior is complementary to the stability analysis and may signal non-linear features already on the linear level [58]. Since analytic tools allowing the study of transient behavior for balanced networks are still missing, we have perform sample simulations, for Gaussian networks. Results are shown in Figures 7-8. These simulation confirm qualitatively the presence of transient behavior.

They raise however more quantitative questions: what are the statistical features of transient behavior in balanced neuronal networks? How the effects of transient behavior

scale with the size of the network? What are the time-scales in the transient behavior? How does the transient behavior depend on the type of an adjacency matrix? We hope to provide some analytic answers to these questions in the sequel to this work. Last, but not least, considering the success of FRV analysis in a variety of disciplines, we hope that the ideas presented in this paper may trigger more interdisciplinary interactions in the area of brain studies.

A   A guide through Free Random Variables

Free random variables can be viewed as a probability theory, where the basic random variable is represented by an infinite matrix. It is therefore most convenient to explain the cornerstones of free theory of probability using the concepts from classical theory of probability (CTP).

Let us consider the following problem. We have two random variables x_1 and x_2 drawn from independent probability distributions $p_1(x_1)$ and $p_2(x_2)$. The distribution of the random variable s being the sum of x_1 and x_2 reads therefore

$$p(s) = \int dx_1 dx_2 p_1(x_1) p_2(x_2) \delta(s - (x_1 + x_2)) = \int dx p_1(x) p_2(s - x) \tag{22}$$

One can easily unravel the convolution using the Fourier transform (characteristic function). Then $\hat{p}(k) \equiv \int p(s) e^{iks} ds = \hat{p}_1(k) \hat{p}_2(k)$, where $\hat{p}_i(k)$ are Fourier transforms corresponding to the original densities $p_i(x)$. Note that a characteristic function generates moments of the respective distribution. We can further simplify the problem if instead of characteristic functions we consider their natural logarithms $\phi(k) \equiv \ln \hat{p}(k)$. Then we get the addition

law, which linearizes the convolution

$$\phi_{1+2}(k) = \phi_1(k) + \phi_2(k) \tag{23}$$

Since ϕ is another generating function – this time for cumulants of the distribution – above relation means simple addition of the corresponding cumulants. The algorithm of convolution is therefore simple. First, knowing $p_i(x)$, we construct the $\phi_i(k)$. Then we perform the addition law (23). Finally, we reconstruct $p_{1+2}(s)$ from the $\phi_{1+2}(k)$, performing the first step in reversed order. A pedagogical and straightforward example is represented by the convolution of two independent Gaussian distributions $N_1(0, \sigma_1^2)$ and $N_2(0, \sigma_2^2)$. First step shows, that in both cases only one cumulant is non-vanishing, i.e., the second one, $\kappa_2^{(1)} = \sigma_1^2$ and $\kappa_2^{(2)} = \sigma_2^2$. The addition law and the last step of the logarithm immediately lead to the result, that the resulting distribution is also Gaussian, $N_{1+2}(0, \sigma^2 = \sigma_1^2 + \sigma_2^2)$.

In free probability, the notion of independence is replaced by the notion of freeness. Two large (infinite) matrices are mutually free if their eigenvectors are maximally decorrelated, e.g., matrices X and $UYU^\dagger$, where U is the Haar measure, are free.

The role of the characteristic function is played by the complex valued Green's function

$$G_X(z) = \int \frac{\rho_X(\lambda)}{z - \lambda} d\lambda \tag{24}$$

where $\rho_X(\lambda)$ is the average spectral density of the matrix X, playing here the role of probability density function in CTP. Indeed, expanding $G_X(z)$ around $z = \infty$ we get spectral moments $M_k^{(X)} = \int \lambda^k \rho_X(\lambda) d\lambda$. Note, that knowing $G_X(z)$ we can easily reconstruct $\rho_X(\lambda)$. Indeed,

$$-\frac{1}{\pi} \lim_{\epsilon \to 0} \Im G(z)|_{z=\lambda'+i\epsilon} = \lim_{\epsilon \to 0} \int \rho(\lambda) \frac{1}{\pi} \frac{\epsilon}{(\lambda' - \lambda)^2 + \epsilon^2} d\lambda = \int \rho(\lambda)\delta(\lambda - \lambda') d\lambda = \rho(\lambda') \tag{25}$$

The role of the generating function for free cumulants is played by the so-called R-transform, $R(z) = \sum_{k=1}^{\infty} \kappa_k z^{k-1}$. The crucial relation between $R(z)$ and $G(z)$ reads $R(G(z)) + \frac{1}{G(z)} = z$ or $G(R(z) + 1/z) = z$, i.e., $R(z)$ is – modulo the shift $1/z$ – the functional inverse of the Green's function. Let us come back to the problem of spectral addition. Imagine we have now the spectral measures $\rho_{X_i}(\lambda)$, corresponding to two matricial ensembles with the measures $P(X_i)dX_i$, where $i = 1, 2$. We are now asking, what is the spectral density of the ensemble $X_{1+2} = X_1 + X_2$. This is a highly non-trivial and non-linear problem, since X_1 and X_2 do not commute, but free calculus allows to solve this case in full analogy to CTP.

21

The algorithm is as follows: First, from ρ_i corresponding to X_i we construct matching $G_i(z)$ and $R_i(z)$. Then

$$R_{X_1+X_2} = R_1(z) + R_2(z) \tag{26}$$

which supersedes (23). Finally, we proceed in reverse order, reconstructing from $R_{X_1+X_2}(z)$ the Green's function $G_{X_1+X_2}(z)$, and finally the spectral density $\rho_{X_1+X_2}(\lambda)$. As an example, we consider the "Gaussian" in free theory, i.e., the spectral distribution the only non-vanishing cumulant of which is variance σ^2. Thus $R(z) = \sigma^2 z$. Reconstructing Green's function gives $\kappa_2 G + 1/G = z$, with obvious solution $G(z) = \frac{1}{2\sigma^2}(z - \sqrt{z^2 - 4\sigma^2})$. Taking imaginary part we reconstruct the celebrated Wigner semicircle $\rho(\lambda) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - \lambda^2}$. We see that the addition algorithm for two free Wigner semicircles mimics precisely the addition algorithm of two Gaussians.

Similarly to addition, one can consider multiplication laws for random variables $x_1 \cdot x_2$. In CTP, such problem is unravelled with the help of Mellin transform, see e.g., [59]. In free calculus, the role of the Mellin transform is played by S-transform, related to R-transform as $S_X(z)R_X(zS_X(z)) = 1$. The multiplication law reads

$$S_{X_1 X_2}(z) = S_{X_1}(z) \cdot S_{X_2}(z) \tag{27}$$

and the algorithm for multiplication follows the one for addition. However, one should be aware that the product of two symmetric (hermitian) matrices may be non-symmetric (non-hermitian). In such a case, the eigenvalues can appear on the whole complex plane, and the methods of $R(z)$ and $S(z)$ transforms, based on analyticity, require substantial modifications. Luckily, there exist one powerful case, governed by the Haagerup-Larsen theorem (known also as a "single ring" theorem), when analytic methods hold for complex spectra. If the complex matrix X can be decomposed as $X = PU$, where P is positive, U is Haar-measured and P and U are mutually free, the spectrum on the complex plane has a polar symmetry and the radial distribution can be easily read out from the singular values of X, i.e., the real eigenvalues of $X^\dagger X$. In mathematics, such ensembles are known as R-diagonal. To infer the information about the spectra and some correlations between left and right eigenvectors one needs only the explicit form of $S_{X^\dagger X}(z)$. In the case of the Ginibre ensemble X (i.e., where X_{ij} are drawn either from real or complex Gaussian distributions), this is particularly easy, since matrix $X^\dagger X$ is known as a Wishart ensemble.

To avoid obscure mathematics, let us recall that the Wishart ensemble is a free analogue of the Poisson distribution from classical probability [18]. This implies that all cumulants are the same, and if for convenience normalized to 1, its R transform is just, by definition, $R_{X^\dagger X}(z) = \sum_{i=1}^{\infty} z^{i-1} = \frac{1}{1-z}$. Using the above-mentioned functional relation between R and S transform we arrive at $S_{X^\dagger X} = \frac{1}{1+z}$. Similar techniques can be applied for generic randomness in Rajan-Abbott type models, as we show below.

B Rajan-Abbott model with Gaussian noise

We use the theorem from free probability, which states that the product of R-diagonal operator with any operator is R-diagonal [60], therefore W is subject to the Haagerup-Larsen theorem. Then $W^\dagger W = \Lambda X^\dagger X \Lambda \overset{tr}{=} X^\dagger X \Lambda^2$, where the last equation expresses the fact that the spectral properties are invariant under the cyclic permutations of matrices under the trace. The Green's function (resolvent) for Λ^2 reads therefore

$$G_{\Lambda^2}(z) = \sum_{i=1}^{m} \frac{f_i}{z - \sigma_i^2} \tag{28}$$

Substituting $z \to R_{\Lambda^2}(z) + \frac{1}{z}$ in (28) and using the fundamental FRV relation $G(R(z) + \frac{1}{z}) = z$ we arrive at

$$1 = \sum_{i=1}^{m} \frac{f_i}{z R_{\Lambda^2}(z) - z\sigma_i^2 + 1} \tag{29}$$

Now we replace in (29) $z \to t S_{\Lambda^2}(t)$ and using the relation between S and R transforms we arrive at

$$1 = \sum_{i=1}^{m} \frac{f_i}{1 + t - \sigma_i^2 t S_{\Lambda^2}(t)} \tag{30}$$

We note that $\frac{1}{1+t}$ is the S-transform for the Wishart ensemble (calculated above), and the multiplication law gives us the final S-transform for $W^\dagger W$, i.e., $\frac{1}{1+t} S_{\Lambda^2}(t) = S_{X^\dagger X}(t) S_{\Lambda^2}(t) = S_{W^\dagger W}(t)$, so we arrive at

$$1 + t = \sum_{i=1}^{m} \frac{f_i}{1 - \sigma_i^2 t S_{W^\dagger W}(t)}. \tag{31}$$

In the last step we substitute $t \to F(r) - 1$ and use the Haagerup-Larsen theorem, arriving at

$$F(r) = \sum_{i=1}^{m} \frac{f_i}{1 - \sigma_i^2 (F(r) - 1)/r^2}. \tag{32}$$

Subtracting $1 = \sum_i f_i$ from both sides finally gives the solution

$$1 = \sum_{i=1}^{m} \frac{f_i \sigma_i^2}{r^2 - \sigma_i^2(F(r) - 1)}. \tag{33}$$

C   Rajan-Abbott model with Cauchy noise

FRV calculus is a powerful technique and the range of its applications is not confined to the basin of attraction of the Gaussian type. In particular, for random matrices X belonging to the free Lévy class (spectrum behaving like $1/\lambda^{\alpha-1}$), the S-transform for the Wishart-Lévy matrix $X^\dagger X$ reads $S_{X^\dagger X}(t) = \frac{1}{t(1+t)} \left(\frac{t}{b}\right)^{t/\alpha}$, with $b = \exp[i\pi(\alpha/2 - 1)]$ [61]. Stability index $\alpha = 2$ reproduces the Gaussian case, but a simple form can be obtained also for the Cauchy disorder $\alpha = 1$. In this case $S_{X^\dagger X}(t) = -\frac{t}{1+t}$, and when applied to (30), yields

$$1 + t = \sum_{i=1}^{m} \frac{f_i}{1 + S_{W^\dagger W}(t)\sigma_i^2} \tag{34}$$

Final substitution $t \to F(r) - 1$ and the use of the Haagerup-Larsen theorem gives explicit, linear equation for arbitrary number of types of neurons

$$F(r) = \sum_{i=1}^{m} \frac{f_i}{1 + \sigma_i^2/r^2} \tag{35}$$

Contrary to the previous case, the spectrum is unbounded and stretches up to infinity. Explicitly, the spectral density and the eigenvector correlator read

$$\rho(r) = \frac{1}{2\pi r} \frac{dF(r)}{dr} = \frac{1}{\pi} \sum_{i=1}^{m} \frac{f_i}{(r^2 + \sigma_i^2)^2} \tag{36}$$

$$O(r) = \frac{1}{\pi r^2} F(r)(1 - F(r)) = \frac{1}{\pi} \sum_{i=1}^{m} \frac{f_i}{r^2 + \sigma_i^2} \sum_{j=1}^{m} \frac{f_j \sigma_j^2}{r^2 + \sigma_j^2} \tag{37}$$

In the case of arbitrary α, resulting transcendental equations can be easily solved numerically. Other type of randomness of the neural networks can also be modeled, e.g., by considering Student-Fisher spectral distributions.

---

[1] T.R. Insel, S.C. Landis, F.S. Collins, *Science* **340**, 687-688 (2013); *The Blue Brain Project* [online], http://bluebrainproject.epfl.ch (2005); D.C. Van Essen, S.M. Smith, D.M. Barch, T.E. Behrens, E. Yacoub, K. Ugurbil for the Wu-Minn HCP Consortium, *Neuroimage* **80**, 62-79 (2013).

[2] G. Hennequin, T. P. Vogels, W. Gerstner, *Physical Review E* **86**, 011909 (2012).

[3] S. Ganguli, D. Huh, H. Sompolinsky, *Proceedings of the National Academy of Sciences 105*, 18970-18975 (2008).

[4] Y. Ahmadian, F. Fumarola, K.D. Miller *Phys. Rev. E* **91**, 012820 (2015).

[5] D Martí, N Brunel & S Ostojic. Correlations between synapses in pairs of neurons slow down dynamics in randomly connected neural networks. arXiv:1707.08337 (2017).

[6] J. S. Bell & J. Steinberger, in: Oxford Int. Conf. on Elementary Particles 19/25 September 1965: Proceedings, Eds. T. R. Walsh, A. E. Taylor, R. G. Moorhouse and B. Southworth, (Rutheford High Energy Lab., Chilton, Didicot 1966), pp. 195–222.

[7] Z. Burda, J. Grela, M.A. Nowak, W. Tarnowski, P. Warchoł,*Phys. Rev. Lett.* **113**, 104102 (2014).

[8] E. Gudowska-Nowak, R.A. Janik, J. Jurkiewicz, M.A. Nowak, *Nuclear Physics B* **670**, 479-507 (2003).

[9] P. Bourgade, G. Dubach. *The distribution of overlaps between eigenvectors of Ginibre matrices*, arXiv:1801.01219v1 [math.PR].

[10] J. T. Chalker, B. Mehlig, *Phys. Rev. Lett.* **81**, 3367 (1998).

[11] B. Mehlig, J. T. Chalker, *J. Math. Phys* **41**, 3233 (2000).

[12] M. Wehr, A.M. Zador, *Nature* **426**, 442-446 (2003).

[13] M. J. Higley, D. Contreras, *Journal of Neuroscience* **26** (2) 448-457 (2006).

[14] B. Haider, A. Duque, A. R. Hasenstaub, D. A. McCormick, *Journal of Neuroscience* **26** (17) 4535-4545 (2006).

[15] K. Rajan, L. F. Abbott, *Phys. Rev. Lett.* **97** 188104 (2006).

[16] L.C.G. del Molino, K. Pakdaman, J. Touboul, G. Wainrib, *Physical Review E*, **88**, 042824 (2013).

[17] M. Asllani, T. Carletii, *Universality of non-normality in real complex networks*, erXiv:1803.11542 [nlin.AO].

[18] D. V. Voiculescu, A. Nica, K. Dykema, *Free Random Variables* CRM Monograph Series Providence RI (1992).

[19] R. Gopakumar, D.J. Gross, *Nuclear Physics B* **451**, 379-415 (1995).

[20] N. El Karoui, *The Annals of Statistics* **36**, 2757-2790 (2008); N.R. Rao, J.A. Mingo, R. Speicher, A. Edelman *The Annals of Statistics* **36**, 2850-2885 (2008).

[21] see, e.g., R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*, Cambridge Univ. Press (2011), and references therein.

[22] M. Potters, J.-P. Bouchaud, L. Laloux, *Acta Phys. Pol. B* **36**, 2767 (2005); Z. Burda, J. Jurkiewicz, M.A. Nowak, G. Papp, I. Zahed, *Physica A* **343**, 694 (2004); Z. Burda, A. Jarosz, M.A. Nowak, M. Snarska, *New Journal of Physics* **12**, 075036 (2010);Z. Burda, A. Jarosz, M.A. Nowak, J. Jurkiewicz, G. Papp, I. Zahed, *Quantitative Finance* **11**, 1103-1124 (2011).

[23] H. Bercovici and V. Pata, with an appendix by Ph. Biane, *Annals of Mathematics*, **149**, 1023 (1999).

[24] E.R. Kandel, J.J. Schwartz, T.M. Jessell, S.A. Siegelbaum, A.J. Hudspeth *Principles of Neural Science, 5th ed.* McGraw-Hill, New York (2012).

[25] J. H. Wilkinson, *Algebraic eigenvalue problem*, Oxford University Press (1965).

[26] L.N. Trefethen, M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press (2005).

[27] B. K. Murphy, K. D. Miller *Neuron*, 61(4):635 648, (2009).

[28] G. Hennequin,T.P. Vogels, W. Gerstner *Neuron*, 82(6), 1394-1406 (2014).

[29] U. Haagerup, F. Larsen, *Journ. Functional Analysis* **176** 331 (2000).

[30] J. Feinberg, A. Zee, *Nuclear Phys. B* **501**, 643 (1997).

[31] J. Feinberg, R. Scalettar, A. Zee *J. Math. Phys.* **42**, 5718 (2001).

[32] In the original formulation Chalker and Mehlig have chosen normalization $1/N$, and have shown that the correlator for the Ginibre-Girko ensemble grows like N. Our normalization $1/N^2$ ensures the finite limit for large N.

[33] R. A. Janik, W. Nörenberg, M. A. Nowak, G. Papp, I. Zahed *Phys. Rev. E* **60**, 2699 (1999).

[34] S. Belinschi, M.A. Nowak, R. Speicher, W. Tarnowski, *J. Phys. A: Math. Theor.* **50** 105204 (2017).

[35] M.A. Nowak, W. Tarnowski, *Probing non-orthogonality of eigenvectors in non-Hermitian matrix models: diagrammatic approach*, arXiv:1801.02526v2 [math-ph].

[36] J. Guckenheimer, Ph. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurctions of Vector Fields (Applied Mathematical Sciences)*, Springer, 2002.

[37] J. Wishart, *Biometrika* **A20**,32 (1928); T.W. Anderson, *Introduction to Multivariate Statistical Analy- sis*, Wiley, New York, (1958).

[38] M. N. Shadlen, W. T. Newsome, *Curr. Opin. Neurobiol.* **4** 569 (1994); T. W. Troyer, K.D. Miller, *Neural Comput.* **9** 971 (1997).

[39] Y. Wei, *Phys. Rev. E* **85** 066116 (2012).

[40] D. J. Watts, S. H. Strogatz, Collective dynamics of small-world networks, Nature, **393**, pp. 440 442, (1998).

[41] S. Yu, D. Huang, W. Singer, D. Nikolić, A Small World of Neuronal Synchrony, Cerebral Cortex **18**, pp. 28912901 (2008).

[42] A. J. Sadovsky, J. N. MacLean, "Scaling of Topologically Similar Functional Modules Defines Mouse Primary Auditory and Somatosensory Microcircuitry", *Journal of Neuroscience* **33** (35), 14048-14060 (2013).

[43] P. Cizeau, J. P. Bouchaud, Phys. Rev. E 50, 1810 (1994).

[44] Z. Burda, J. Jurkiewicz, M.A. Nowak, G. Papp, I. Zahed, Phys. Rev. E 75, 051126 (2007).

[45] Y.V. Fyodorov *On statistics of bi-orthogonal eigenvectors in real and complex Ginibre ensembles: combining partial Schur decomposition with supersymmetry* arXiv:1710.04699 [math-ph]

[46] J. Beggs, D. Plenz *Neuronal Avalanches in Neocortical Circuits.* Journal of Neuroscience **23** 11167M (2003).; M. Benayoun, J.D. Cowan, W. van Drongelen, E Wallace. *Avalanches in a Stochastic Model of Spiking Neurons* PLoS Comput Biol. **6**, e1000846. (2010).

[47] J. Grela, P. Warchoł *Full Dysonian dynamics of the complex Ginibre ensemble* arXiv:1804.09740 [math-ph]

[48] S.-Y. Kim, W. Lim, *Stochastic Spike Synchronization in A Small-World Neural Network with Spike-Timing-Dependent Plasticity*, arXiv:1704.03150 (2017).

[49] D. R. Chialvo, *Nature Physics* **6**, 744750 (2010)

[50] G. G. Turrigiano, S. B. Nelson *Nature reviews. Neuroscience* **5**, 97107 (2004).

[51] G. Liu *Nat. Neurosci.* **7**, 373379 (2004).

[52] J. Pu, H. Gong, X. Li, Q. Luo *Scientific Reports* **3**, 1081 (2013).

[53] W. L. Shew, H. Yang,S. Yu, R. Roy, D. Plenz *J. Neurosci* **31**, 5563 (2011).

[54] A. Levina, J. M. Herrmann, T. Geisel, *Nature Phys.* **3**, 857860 (2007).

[55] M. O. Magnasco, O. Piro O, G. A. Cecchi *Phys. Rev. Lett.* **102** 258102 (2009).

[56] E. Schneidman, M. J. Berry II, R. Segev, W. Bialek *Nature* **440**, 1007 (2006).

[57] S.-S. Poil, R. Hardstone, H.D. Mansvelder, K. Linkenkaer-Hansen, *The Journal of Neuroscience*, **32**, 9816 (2012).

[58] J. Grela, Phys. Rev. E **96**, 022316 (2017).

[59] B. Epstein *The Annals of Mathematical Statistics* **19**, 370 (1948).

[60] A. Nica, R. Speicher *Lectures on the Combinatorics of Free Probability* Cambridge University Press (2006).

[61] Z. Burda, J. Jurkiewicz, M. A. Nowak, G. Papp, I. Zahed, Physica A 299, 181 (2001).