Correlations between synapses in pairs of neurons slow down dynamics in randomly connected neural networks

Daniel Martí, 1 Nicolas Brunel, 2, 3 and Srdjan Ostojic 1

1 Laboratoire de Neurosciences Cognitives, INSERM U960, Ecole Normale Supérieure - PSL Research University, Paris, France
2 Departments of Statistics and Neurobiology, University of Chicago, Chicago, USA
3 Departments of Neurobiology and Physics, Duke University, Durham, USA

(Dated: July 27, 2017)

Networks of randomly connected neurons are among the most popular models in theoretical neuroscience. The connectivity between neurons in the cortex is however not fully random, the simplest and most prominent deviation from randomness found in experimental data being the overrepresentation of bidirectional connections among pyramidal cells. Using numerical and analytical methods, we investigated the effects of partially symmetric connectivity on dynamics in networks of rate units. We considered the two dynamical regimes exhibited by random neural networks: the weak-coupling regime, where the firing activity decays to a single fixed point unless the network is stimulated, and the strong-coupling or chaotic regime, characterized by internally generated fluctuating firing rates. In the weak-coupling regime, we computed analytically for an arbitrary degree of symmetry the auto-correlation of network activity in presence of external noise. In the chaotic regime, we performed simulations to determine the timescale of the intrinsic fluctuations. In both cases, symmetry increases the characteristic asymptotic decay time of the autocorrelation function and therefore slows down the dynamics in the network.

INTRODUCTION

The dynamics and function of a network of neurons is to a large extent determined by its pattern of synaptic connections. In the mammalian brain, cortical networks exhibit a complex connectivity that to a first approximation can be regarded as random. This connectivity structure has motivated the study of networks of neurons connected through a random synaptic weight matrix with independent, identically distributed (i.i.d.) entries, which have become a central paradigm in theoretical neuroscience [1–3]. Randomly connected networks of ‘firing rate’ units exhibit a chaotic phase [1], which can be exploited as a substrate for complex computations [4–6]. Networks of randomly connected spiking neurons also exhibit rich dynamics that can account for the highly irregular spontaneous activity observed in the cortex in vivo [2, 3, 7, 8]. Importantly, these models are to a large extent amenable to a mathematical analysis, which allows for a thorough understanding of the mechanisms underlying their dynamics.

Detailed analyses of experimental data on cortical connectivity have however identified patterns of connectivity that strongly deviate from the i.i.d. assumption [9–13]. The most prominent of such deviations is the overrepresentation of reciprocal connections [9, 10, 14], and the fact that synapses of bidirectionally connected pairs of neurons are in average stronger that synapses of unidirectionally connected pairs. These observations are consistent with a partially symmetric connectivity structure, intermediate between full symmetry and full asymmetry. How partial symmetry in the connectivity impacts network dynamics is not yet understood, in part because such partial symmetry renders the mathematical analyses more challenging [15]. Here, we study the impact of partial symmetry in the connectivity structure on dynamics of a simple network model consisting of interacting rate units. Depending on the overall strength of coupling, such a network can display either a stable or a chaotic regime of activity, as in the random asymmetric case [1].

We examined how the degree of symmetry in the network influences the temporal dynamics in both regimes. For the stable regime, we exploited recent results from random matrix theory [16, 17] to derive analytical expressions for the autocorrelation functions. These expressions demonstrate that increasing the symmetry in the network leads to a slowing down of the dynamics. Numerical simulations in the chaotic regime show a similar effect, with time scales increasing far more substantially with symmetry than in the fixed point regime.

I. DESCRIPTION OF THE MODEL

We consider a network of $N$ fully connected neurons, each described by an activation variable (synaptic current) $x_i, i = 1, \ldots, N$, obeying

$$\frac{dx_i}{dt} = -x_i + g \sum_{j=1}^{N} J_{ij} \phi(x_j),$$

where $g$ is a gain parameter that modulates the strength of recurrent connections, and where $\phi(\cdot)$ is the input-output transfer function that transforms activations $x_i$ into firing rates. This transformation is non-linear and we model it as $\phi(x) = \tanh(x)$ for mathematical convenience (see [18–21] for studies of network models with different choices of $\phi$). The elements $J_{ij}$ of the connectivity matrix are drawn from a Gaussian distribution with zero mean,
variance $1/N$, and correlation

$$[J_{ij}J_{ji}]_J = \eta/N,$$

with the square brackets $[.]_J$ denoting average over realizations of the random connections. The parameter $\eta$ is the correlation coefficient between the two weights connecting pairs of neurons, and quantifies the degree of symmetry of the connections. For $\eta = 0$ the the elements $J_{ij}$ and $J_{ji}$ are independent and the connectivity matrix is fully asymmetric, whereas for $\eta = 1$ the connectivity matrix is symmetric. In the following we will consider the full range in between, $\eta \in [0, 1]$.

II. DYNAMICAL REGIMES OF THE NETWORK

For fully asymmetric matrices, previous work has shown that the network activity described by (1) undergoes a phase transition at $g = 1$ in the limit of large $N$ [1]. For $g < 1$ the activity for all units decays to 0, which is the unique stable fixed point of the dynamics, while for $g > 1$ the activity is chaotic. Such a transition can be partially understood by assessing the stability of the fixed point at $x_i = 0$ for $i = 1, \ldots, N$. If we linearize Eq. (1) around this fixed point we obtain the stability matrix, with components

$$M_{ij} = -\delta_{ij} + gJ_{ij}.$$ (2)

The eigenvalues of $M_{ij}$ are therefore those of the matrix $J_{ij}$, scaled by the gain $g$ and shifted along the real axis by $-1$. In the limit $N \to \infty$, for a connectivity matrix $J_{ij}$ whose entries are i.i.d. Gaussian random variables of zero mean and variance $1/N$, eigenvalues are uniformly distributed in the unit disk of the complex plane [22–24]. This implies that the eigenvalues of the stability matrix have negative real part as long as $g < 1$, and therefore that the fixed point at 0 is stable in that range.

An analogous transition occurs when connections are partially symmetric. The presence of correlations among weights deforms the spectrum of eigenvalues into an ellipse, elongating its major radius by a factor of $1 + \eta$ and shortening the minor radius by a factor $1 - \eta$ [25–28] (Figure 1a). This property is usually referred to as the elliptic law. For the network described by (1) such a deformation causes the fixed point at $x_i = 0$ for $i = 1, \ldots, N$ to lose its stability at $g = 1/(1 + \eta)$ (Figure 1b,c). In other words, symmetry lowers the critical coupling.

Our goal is to characterize how the degree of symmetry in the connections affects the network activity on each side of the instability: the relaxation response of the network at low gains and the chaotic self-generated activity observed at strong gains. Our description of the network activity will be based on the average autocorrelation function,

$$C(\tau) = \frac{1}{N} \sum_{i=1}^{N} \langle [x_i(t)x_i(t+\tau)] \rangle_J,$$ (3)

where the average is over both the population and the realizations of the connectivity matrix $[1]$, and where we are assuming for now that the system is stationary.

III. DYNAMICS IN THE FIXED-POINT REGIME

A. Derivation of the autocorrelation function

In the fixed-point regime, the activity decays to zero unless the network is stimulated by external inputs. To characterize the dynamics of the network in this regime, we induce network activity by feeding each neuron with independent Gaussian white noise. The amplitude of this noise is assumed to be small enough so that the synaptic activation of all neurons lies within the linear range of their input-to-rate transfer function. Under these conditions, $\phi(x)$ can be approximated by its first order Taylor expansion $\phi(0) + \phi'(x)|_{x=0} x = x$, and the dynamical equations become

$$\frac{dx(t)}{dt} = (-1 + gJ)x(t) + \sigma \xi(t),$$ (4)

where $x(t) = (x_1(t), \ldots, x_N(t))^T$, $1$ is the identity matrix, $J$ is the connectivity matrix, and $\xi(t) = (\xi_1(t), \ldots, \xi_N(t))^T$ is a vector of independent white noise sources of zero mean and unit variance: $\langle \xi_i(t) \rangle = 0$, $\langle \xi_i(t)\xi_j(t') \rangle = \delta_{ij}\delta(t-t')$, with angle brackets representing averages over noise realizations. The parameter $\sigma$ is the standard deviation of the white noise injected to neurons.

The time scales displayed by a linear system like (4) are strongly affected by the real part of the eigenvalues of the system’s stability matrix and, in particular, get longer as eigenvalues get closer to the imaginary axis. To disentangle this type of slowing down from the effects due to symmetry alone, we vary the parameter $\eta$ while keeping the spectral gap fixed. By spectral gap we mean the distance between the spectrum of eigenvalues of the stability matrix $M_{ij}$, Eq. (2), and the imaginary axis (see Fig. 2a). From the elliptic law, the eigenvalue of the stability matrix with largest real part is $z = -1 + g(1 + \eta)$, and we can keep the spectral gap at $\delta$ by setting the gain to $g = (1 - \delta)/(1 + \eta)$.

The system described by (4) is linear and can be solved by diagonalizing the connectivity matrix. The matrix $J$ admits a set of right eigenvectors $\{R_1, \ldots, R_N\}$ that obey $JR_i = \lambda_iR_i$ for $i = 1, \ldots, N$. These eigenvectors are in general complex-valued and, except for the symmetric case $\eta = 1$, not orthogonal to one another, which implies that $J$ cannot be diagonalized through a unitary transformation. Matrices of this kind are called non-normal and do not commute with their transpose conjugate: $JJ^T \neq J^TJ$ [29]. Even if non-normal matrices cannot be diagonalized by an orthogonal set of eigenvectors, it is always possible to form a biorthogonal basis by extending the set of right eigenvectors with the set of
Figure 1.  a Spectrum of eigenvalues of a Gaussian random matrix with zero mean, variance $1/N$, and correlation coefficient $\eta$ between weights connecting neuronal pairs.  b Time evolution of the firing rate of five arbitrary neurons, for the fixed-point regime (top) and the chaotic regime (bottom), and for two different values of $\eta$ (left and right panels). Next to each trajectory is the eigenspectrum of the corresponding linearized system along with the real and imaginary axes, which we include for reference. The initial firing rate was the same for all panels. All connectivity matrices were derived from a single realization of a Gaussian random matrix. To generate connectivity matrices with different $\eta$, we used the upper right and the lower left triangular portions of the Gaussian random matrix to create a symmetric and antisymmetric matrix, which we then combined to yield a $J$ with the desired symmetry parameter $\eta$ (for details, see text surrounding Eq.(D4), in Appendix D). c Activity regimes as a function of the gain and the degree of symmetry in the connections. The green squares indicate the parameter values used in b.

Figure 2. a The distance of the eigenspectrum to the imaginary axis, or spectral gap, is kept fixed at a value $\delta$ independently of the symmetry parameter $\eta$. The color curves represent the right halves of the boundaries of the spectrum of eigenvalues for $\eta = 0, 0.4, 0.8$ (values indicated next to each curve). b,c Analytical prediction and numerical estimate of the average autocorrelation, for different values of the symmetry parameter $\eta$, indicated in the key. Each subplot corresponds to a particular spectral gap $\delta$. Insets show the same curves in a log-log scale. The numerical estimate of $C(\tau)$ was derived from Eq. (3), using simulated neuronal activity generated by Eq. (4), with $N = 10000$ units, and averaging over time, units, and 200 different realizations of the connectivity matrix. d Comparison of our analytical prediction with three alternative semi-numerical predictions (see text for details).

left eigenvectors, which obey $L_i^\dagger J = \lambda_i L_i^\dagger$. This extended basis is biorthogonal in the sense that $L_i^\dagger R_j = \delta_{ij}$. We can summarize all these properties in a compact way by defining the square matrices $R$ and $L$ that result from adjoining in columns the set of, respectively, right and left eigenvectors, and by introducing the diagonal matrix $\Lambda$ that contains the eigenvalues $\lambda_i$ of $J$ in its diagonal entries. In this notation the biorthogonality condition is $L_i^\dagger R_j = \delta_{ij}$ and the eigenvalue equations for the right and left eigenvectors read $J R = R \Lambda$, $L^\dagger J = \Lambda L^\dagger$. 
We can now write the formal solution of (4):

\[ x(t) = \sigma \int_{-\infty}^{t} e^{-(1+gJ)(t-s)} \xi(s) \, ds \]

\[ = \sigma R \int_{-\infty}^{t} e^{-(1+gA)(t-s)} R^{-1} \xi(s) \, ds , \]

where in the last equality we used the basis of right eigenvectors to write \( J = \mathbf{R} \mathbf{A} \mathbf{R}^{-1} \) and we implicitly expanded the exponential in its power series to obtain the final result. From this expression we can derive the population-average autocorrelation for a particular realization of the connectivity:

\[ C_J(\tau) = \frac{1}{N} \langle x(t)x(t+\tau) \rangle = \frac{1}{N} \text{Tr}(x(t+\tau)x(t)) \]

\[ = \frac{\sigma^2}{N} \int_{0}^{\infty} e^{-2u - \tau} \text{Tr}\{ \mathbf{R}' \mathbf{R} e^{\sigma \Lambda(u+\tau)} \mathbf{L} \mathbf{L}^{-1} e^{\sigma \Lambda^\dagger u} \} \, du . \]

(5)

In the second equality we used the cyclicity of the trace, and in the last line we changed the integration variable to \( u = t - s \) and we used the biorthogonality condition to write \( R^{-1} = \mathbf{L}' \mathbf{L} \). The average over noise amounts to applying the identity \( \langle \xi(t)\xi(t') \rangle = \sigma^2 \delta(t - t') \). Note that the \( \sigma^2 \) appears as an overall factor, so we can set \( \sigma = 1 \) without loss of generality.

We can simplify (5) introducing the so-called overlap matrix, with components

\[ O_{ij} = (\mathbf{L} \mathbf{L})_{ij} (\mathbf{R}' \mathbf{R})_{ji} , \]

and which characterizes the correlations between left and right eigenvectors [16]. Equation (5) then becomes

\[ C_J(\tau) = \frac{1}{N} \int_{0}^{\infty} e^{-2u - \tau} \sum_{j=1}^{N} \sum_{i=1}^{N} e^{\sigma \lambda_i (u+\tau)} O_{ij} e^{\sigma \lambda_j u} \, du . \]

(7)

If the connectivity matrix were normal, the overlap would be the identity matrix and the autocorrelation (7) would just be a sum of independent contributions—one per eigenvalue. These contributions are coupled for non-normal matrices.

We can make further analytical progress by studying the autocorrelation (7) in the limit \( N \rightarrow \infty \), in which the differences of \( C_J(\tau) \) across realizations of the connectivity matrix disappear. In that limit sums over indices are replaced with integrals over eigenvalues, while the overlap matrix is replaced with the local average of the overlap, defined as

\[ D(z_1, z_2) = \lim_{N \rightarrow \infty} \frac{4}{N} \left[ \sum_{l=1}^{N} \sum_{j=1}^{N} O_{ij} \delta^2(z_1 - \lambda_i) \delta^2(z_2 - \lambda_j) \right] . \]

(8)

Here \( z = x + iy \) are complex numbers and we defined the complex Dirac delta as \( \delta^2(z) \equiv (1/2) \delta(x) \delta(y) \) so that it satisfies the normalization condition \( \int \delta^2(z) \, d^2z = 1 \). After taking the limit \( N \rightarrow \infty \), equation (7) becomes

\[ C(\tau) = \int_{0}^{\infty} e^{-2u - \tau} A(u, \tau) \, du . \]

(9)

where we defined

\[ A(u, \tau) = \frac{1}{4} \int \int e^{g(z_1 + z_2) u + g z_1 \tau} D(z_1, z_2) \, dz_1 dz_2 . \]

(10)

Each of the integrals in Eq. (10) is over complex values, and involves the expression of \( D(z_1, z_2) \) for the ensemble of Gaussian random matrices with partial symmetry, which was derived using diagrammatic techniques by [16] and whose functional form can be found in Appendix A. We used the result of [16] to evaluate the double complex integral \( A(u, \tau) \) in Eq. (10). The details of the evaluation are given in Appendix A, and the result is

\[ A(u, \tau) = A_1(u, \tau) + A_2(u, \tau) , \]

(11)

with

\[ A_1(u, \tau) = (1 + \eta^2) I_0 \left( g \psi(u, \tau; \eta) \right) \]

\[ - 2 \eta \left( 1 + \frac{2(1 - \eta)^2}{\psi(u, \tau; \eta)^2} \right) I_2 \left( g \psi(u, \tau; \eta) \right) , \]

(12)

\[ A_2(u, \tau) = \frac{1}{g^2 u (u + \tau)} \sum_{k=1}^{\infty} \eta^k k^2 I_k \left( 2 g \sqrt{\eta u} \right) \]

\[ \times I_k \left( 2 g \sqrt{\eta (u + \tau)} \right) , \]

(13)

where \( I_k(\cdot) \) is the modified Bessel function of order \( k \), and where in Eq. (12) we defined

\[ \psi(u, \tau; \eta) = 2 \sqrt{(1 + \eta^2) u (u + \tau) + \eta^2} . \]

The autocorrelation is finally computed from Eq. (9), integrating numerically over \( u \).

The analytical prediction given by Eqs. (9) and Eqs. (11)–(13) matches with the autocorrelation estimated from numerical simulations (Figure 2b), although for long time lags the numerical estimate becomes noisy due to finite-size effects. To check the validity of our prediction also at long time lags, we compared our analytical prediction with three alternative derivations (Figure 2c). One of such derivations consists of computing the autocorrelation for large but finite \( N \), by computing numerically the eigenvalues and eigenvectors of randomly generated matrices, evaluating the time integral of Eq. (7), which gives

\[ C_J(\tau) = - \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} O_{ij} e^{-(1-\sigma \lambda_i) \tau} + g (\lambda_i + \lambda_j) , \]

(14)

and then by averaging \( C_J(\tau) \) over multiple realizations of the connectivity matrix. Another derivation is based on dynamical mean-field theory [15, 30], which gives rise to a set of integro-differential equations involving...
$C(\tau)$ and which can be solved numerically (Appendix D).
And finally, we numerically computed the inverse Fourier
transform of the power spectrum derived by [31] for this
same system. [31] used a perturbative method to de-
rive the system of integro-differential equations (D25)–
(D26), which they solved for the correlation and reponse
functions by using a Laplace transform. All derivations
yield the same result, except for the deviations we ob-
serve when applying Eq. (14) at long $\tau$ and which are
caused by finite-size effects.

Our results show that an increase in symmetry tends to
spread autocorrelations toward longer time lags, and that
this effect gets larger the closer the system gets to the
onset of chaos. An intuitive explanation for this slowing
down is that the deformation of the eigenspectrum caused
by symmetry increases the density of eigenvalues with
small imaginary parts, thereby enlarging the contribution
of low-frequency modes.

B. Behavior at long time lags

While equations (9)–(13) are exact, they provide little
analytical insight into how the autocorrelation depends
on parameters. A more explicit dependence can be ob-
ained by evaluating $C(\tau)$ in the limit of long $\tau$. We
relegate the details of the calculation to Appendix B and
summarize the main results here. The analysis shows
that, in the fixed point regime, there exist two subregimes
of activity that differ in how the asymptotic decay rate of
the autocorrelation depends on the symmetry parameter
$\eta$ and the spectral gap $\delta$. For small values of $\eta$ and $\delta$,
the autocorrelation decays as a pure exponential at long
$\tau$ (regime I):

$$C(\tau) = F_1(\eta, \delta) e^{-\tau G_1(\eta, \delta)}, \quad (15)$$

with

$$F_1(\eta, \delta) = \frac{\delta^{-1/2}(1 - \eta)^2}{2\sqrt{2}(1 - \delta)}, \quad (16)$$

$$G_1(\eta, \delta) = \frac{1 - \eta}{1 + \eta} \sqrt{2\delta - \delta^2}. \quad (17)$$

Conversely, for sufficiently large values of $\eta$ and $\delta$
autocorrelation for long $\tau$ can be approximated by a power
multiplied by an exponential decay (regime II):

$$C(\tau) = \tau^{-3/2} F_{II}(\eta, \delta) e^{-\tau G_{II}(\eta, \delta)}, \quad (18)$$

with

$$F_{II}(\eta, \delta) = \frac{1}{4\sqrt{\pi}} \left( \frac{1 + \eta}{1 - \delta} \right)^{3/2} \left[ \frac{2\eta^{-1/4}(1 + \eta^2)}{\delta(1 - \eta) - [1 - \sqrt{\eta}]^2} \right. \left. \frac{\eta^{5/6}(1 + \eta)}{(1 - \sqrt{\eta})^2 + 2\sqrt{\eta} \delta} \right], \quad (19)$$

$$G_{II}(\eta, \delta) = \frac{(1 - \sqrt{\eta})^2 + 2\sqrt{\eta}}{1 + \eta}. \quad (20)$$

A comparison between the asymptotic expression in
Eq. (18) and the full expression for the autocorrelation
function reveals however that the power law is not ob-
erved in practice because the range below the cutoff falls
below the values of $\tau$ where the asymptotic approxima-
tion starts matching the exact expression.

Figure 3a shows the exact parameter region of each
asymptotic regime, after transforming the spectral gaps
into gains. In both regimes the autocorrelation’s asymp-
totic decay rate matches the exact result for time lags
longer than a few time units (see Figure 3a, lateral pan-
els). It seems therefore reasonable to associate the time
scale of the autocorrelation with the inverse of $G_{II}(\eta, \delta)$
(see Eqs. (15) and (18)), where the subindex I, II is chosen
according to the subregime found at the parameter values
$(\eta, \delta)$. The asymptotic time scale of the autocorrelation
increases monotonically with symmetry regardless of the
subregime the system enters in (Figure 3b), although
this dependence is convex in the exponential subregime
and concave in the power-law-with-cutoff regime (in Fig-
ure 3b see the curves split by the red dots, which mark
the boundary between subregimes). Note also that as
the spectral gap $\delta$ shrinks to 0 the system enters into the
exponential regime and timescales diverge as $\delta^{-1/2}$,
according to Eq. (17).

C. Effect of overlaps

In this section we characterize the effects that the non-
normality of the connectivity matrix $J$ has on the au-
tocorrelation. A simple way to illustrate these effects is
by comparing the autocorrelation we derived in Sec-
tion IIIA with the autocorrelation we would obtain if we
assumed that the eigenvectors of $J$ were orthogonal. If
that were the case, the autocorrelation could be com-
puted as a sum of decoupled contributions associated
with the different eigenvalues and, in particular, the dis-
tribution of eigenvectors would play no role in the result
(see Appendix III C for details). Figure 4a shows the
predicted autocorrelations, both including and excluding
the contribution from the overlap, Eq. (6). As expected,
both predictions coincide for $\eta = 1$ and they increasingly
depart from each other for decreasing values of $\eta$. To
better characterize this difference, we show the variance,
$C(0)$, as a function of the symmetry parameter for sev-
eral values of the spectral gap (Fig. 4b). Remarkably,
the variance decreases with symmetry, but the opposite
occurs when we remove the contribution from eigenvector
overlaps (see ‘with’ and ‘without’ overlap curves in Fig.
4b). In both cases the variance increases as spectral
gaps get smaller, which is consistent with the fact that
the restoring drive towards the fixed point gets weaker
as the spectral gap gets smaller, but this effect is much
weaker when overlaps are not taken into account [32].

The overlap also contributes to the overall time scale
of the autocorrelation, which we define by the quan-
tity $\hat{\tau} = \int_0^{\infty} t C(t) \, dt / \int_0^{\infty} C(t) \, dt$. This definition guar-
antees that for an exponential autocorrelation $C(\tau) \propto \exp(-|\tau|/\tau_0)$ the overall time scale is exactly $\tau_0$, and provides a rough estimate of a natural time scale for autocorrelations with more complex dependences. The numerical evaluation of $T$ shows that the overall time scale is systematically smaller if the contribution of the eigenvalues is removed (Fig. 4c). Unsurprisingly, either with or without the overlap contribution the timescale gets longer as the spectral gap gets smaller. Note also that the overall time scale $T$ varies non-monotonically with the symmetry parameter (Fig. 4c), unlike the asymptotic dependence shown in Fig. 3b.

IV. DYNAMICS IN THE CHAOTIC REGIME

In the chaotic regime, the network generates its own fluctuating activity without the need for external noise. Recall that chaotic activity emerges as soon as the largest of the real parts of the eigenspectrum, given by $-\delta$ and usually called spectral abscissa, becomes positive. We follow the strategy of the previous section and we keep the spectral abscissa fixed while we vary the symmetry parameter $\eta$.

The evolution of firing activities shown in Figure 5a suggests that in the chaotic regime the self-generated fluctuations get slower as $\eta$ increases. This slowing is accompanied by an increasing tendency of firing rates to linger around the extreme values of their dynamical range, as reflected by an increasingly bimodal distribution of currents $x$ and rates $\phi(x)$ when $\eta$ increases (Figure 5b). We quantified the slowing down of the fluctuations with the population-average autocorrelation. For $\eta = 0$ the autocorrelation can be derived self-consistently in the limit of infinitely large networks, using the dynamical mean-field approach (Sompolinsky et al. 1, Rajan et al. 18, see also Appendix D for a general derivation). Unfortunately, this method does not lead to a closed-form solution for the autocorrelation as soon as $\eta > 0$ (see Appendix D) and we have to resort to numerical estimates, summarized in Figure 5c for several values of $\eta$. For completeness we also include the autocorrelation functions for fixed gain, rather than fixed spectral abscissa (Fig. 5d).

The numerical estimates show that the time scale associated with the autocorrelation increases strongly as a function of $\eta$ and is considerably longer than in the fixed-point regime (Fig. 5e). Such a slowing is rather insensitive to whether we fix the spectral abscissa or the gain, despite the fact that the variance $C(0)$ varies far more strongly when gain is fixed (Fig. 5d).

Quite strikingly, for $\eta = 1$ fluctuations become slower as time goes by, and our initial assumption that the activity is stationary does not hold. The population-averaged autocorrelation $C(\tau, t+\tau) = [x(t)x(t+\tau)]_{J,N}$ at different points in time shows that the characteristic timescale of the autocorrelation grows with $t$ (Fig. 5d), a signature of aging dynamics [33]. For lower values of $\eta$, the dependence on the auto-correlation on the two timescales is less clear. Due to strong finite-size effects, it is difficult to determine from simulations alone whether aging appears also when the connectivity is not fully symmetric.

V. DISCUSSION

In this work we examined the effect of partially symmetric connectivity on the dynamics of randomly connected networks composed of rate units. We have derived an analytical expression for the autocorrelation function in the regime of linear fluctuations around the fixed point, and shown that increasing the symmetry of the connectivity leads to a systematic slowing-down of the dynamics. Numerical simulations confirm that a similar phenomenon takes place in the chaotic regime of the net-
The impact of the degree of symmetry of the connectivity matrix on the dynamics of neural networks has been a long-standing question in theoretical neuroscience. Theorists initially focused on fully symmetric networks of binary spin-like neurons [34] for which tools from equilibrium statistical mechanics could be readily applied [35]. After these initial studies, the realization that brain networks are not symmetric led physicists to investigate the dynamics of networks whose connectivity matrix has a random antisymmetric component. It was found that departures from full symmetry destroy spin-glass states, while retrieval states in associative memory models were found to be robust to the presence of weak asymmetry [15, 36, 37].

Theorists also studied fully asymmetric networks, using rate models [1], networks of binary neurons [38] and networks of spiking neurons [3]. In all these models, chaotic states were shown to be present for sufficiently strong coupling. In networks of spiking neurons, chaotic states are characterized by strongly irregular activity of the constituent neurons, with self-generated fluctuations that evolve on fast time scales. Motivated by experimental findings, recent studies have considered synaptic connectivity matrices where bidirectionally connected pairs are overrepresented with respect to a random network. In contrast with our model, in which no structure exists beyond the level of pairs of neurons, these studies have considered structured connectivity matrices in which partial symmetry is a consequence of a larger-scale structure. [39] considered a connectivity clustered into groups of highly connected neurons and demonstrated that clustered connectivity could lead to slow firing-rate dynamics generated by successive transitions between up and down states within individual clusters. An overrepresentation bidirectional connections can also arise in networks with broad in- and out-degree distributions, which affect the dynamics and the stability of asynchronous states in such networks [40]. Other works have considered connectivities with non-trivial second-order connectivity statistics, and studied the resulting network dynamics. [41] analyzed how the presence of connectivity patterns involving two connections (not only bidirectionally connected pairs) affected the tendency for a neuronal network to synchronize, while [42] focused on the oscillatory activity generated by partially antisymmetric, delayed interactions. Taking a completely different approach, [43] showed that maximizing the number of patterns stored in a network entails an overrepresentation of bidirectionally connected pairs of neurons, which suggests that partially symmetric connectivity may be a signature of optimal information storage.

An important ingredient in our analysis is the fact that partially symmetric interaction matrices are non-normal, i.e., they are not diagonalizable by a set of mutually orthogonal eigenvectors. The influence of non-normal connectivity on network dynamics has recently received a considerable attention in the neuroscience community [32, 44–48]. Particularly relevant to our study is the work by [32], who quantified the effects of non-normality on the amplitude of the autocorrelation function in random networks. Here we extend their results by studying the full temporal shape of the autocorrelation function and by characterizing how this shape is affected by the partial symmetry of connections. Note that our analytical results are also valid for the case of anti-symmetric connectivity \((-1 \leq \eta < 0\), where the eigenspectrum gets elongated along the imaginary axis [42, 47].

The present work is also related to models of disordered systems and spin glasses [49]. Most studies in that field were inspired by physical phenomena and considered fully symmetric interaction matrices. In that context, a major result has been the discovery of aging, the phenomenon by which dynamics become slower the longer the system evolves [33, 50, 51]. This phenomenon has been observed in a broad class of complex systems characterized by configuration spaces with extremely rugged energy landscapes, composed of many local minima surrounded by high barriers. In these systems a random initial condition is very likely to set the system far from a stationary state and initiate a very slow relaxation to-
Figure 5. Effects of symmetry in the chaotic regime. a) Firing rates of five arbitrary units in a network of size $N = 10^4$, for three different values of $\eta$ (value indicated next to each inset). b) Histograms of the currents $x$ and firing activities $\phi(x)$, for different values of $\eta$. Samples were taken every 40 time units from the simulated activity of a network of $10^4$ units for 10 different realizations of the connectivity matrix. The spectral gap is $\delta = -0.4$ (i.e., the spectral abscissa is 0.4). c) Population-average autocorrelation, for different values of the symmetry parameter $\eta$ and $\delta = -0.4$. The autocorrelation was estimated from the simulated activity of a network of $10^4$ units, using 200 different draws of the connectivity matrix. The estimated standard error of the mean is shown in semitransparent shade. d) As in c, but keeping the gain fixed at $g = 1.5$ instead of fixing $\delta$. For this panel we estimated the autocorrelation using 10 independent realizations of the connectivity matrix. Notice the different axis ranges with respect to c. e) Time scale of the network fluctuations, estimated from the width of the autocorrelation function at half of its maximum value, using the simulation results from c and d. The error bars indicate the standard error of the mean. f) Non-stationarity of the network activity for $\eta = 1$. The autocorrelation function depends on two time scales: the time lag $\tau$ and the time $t$ since the simulation started with an arbitrary initial condition. The autocorrelation is estimated from simulated neuronal activity of size $N = 10^4$ for $T = 3030$ time units, using 5 different trials with different realizations of the connectivity matrix.

Towards a fixed point. The relaxation takes infinitely long for $N \to \infty$ because, loosely speaking, the longer the system evolves, the deeper it wanders in the valleys of the energy landscape, and the harder it becomes for it to find configurations of lower energy [33].

Whether fully symmetric interactions are necessary to observe aging does not seem to be entirely understood, as to the best of our knowledge only a few works seem to have considered partially symmetric coupling [15, 52, 53]. Fully asymmetric networks have received more attention, but they do not exhibit any aging phenomena. Here we interpolated between fully asymmetric and fully symmetric networks, and have been able to obtain mathematical results only in linear networks, in the non-chaotic regime. Interestingly, we found that the partially symmetric case is mathematically more complex than the symmetric or asymmetric limits. This can be seen in the form of autocorrelation function (11), which simplifies considerably when $\eta = 0$ or $\eta = 1$, but also in the Dynamical Mean Field Theory (Appendix D), where a coupling between the auto-correlation and the response function appears for $\eta > 0$. This additional complexity results from the fact that the influence of a single neuron’s activity on all the other neurons is fed back through couplings that are correlated with the neuron’s activity, due to the partial symmetry of the connections. More specifically, the inputs received by neuron $i$ are given by terms $\sum_j J_{ij} \phi(x_j)$, which are themselves influenced by the activity of neuron $i$. As a result, neuron $i$ influences its own activity by an amount proportional to the sum $\sum_j J_{ij} J_{ji} \phi(x_i)$, a random number of mean $\eta \phi(x_i)$. The effect of this feedback loop is that the individual input terms exhibit correlated fluctuations. When $\eta = 0$, the inputs received by neurons are uncorrelated and their sum can be approximated by a Gaussian random variable whose mean and variance can be determined self-consistently [1, 18]. At the other extreme, when $\eta = 1$, the inputs received by neurons are correlated, but the dynamics of the network can be described as a relaxation of an energy function and the standard machinery of statistical mechanics can be used. For other values of $\eta$, none of these analytical strategies can be applied and the analysis becomes more complex. Demonstrating analytically whether aging dynamics are present in partially symmetric, non-linear networks seems an outstanding open problem.

Our results on the autocorrelation function in the linear network are closely related to recent results published by [31], who used a different set of methods to compute the power spectrum of the network activity, i.e., the Fourier transform of the autocorrelation function of the same model we investigated. Unlike [31], we obtained the autocorrelation directly in real time, although our re-
results are fully consistent with theirs in that we obtain the same two regimes with the same asymptotic timescales, depending on the symmetry and the gain (or leak, in their case. cfr. Fig. 1 of [31] with Fig. 3a).

Our work provides a potential bridge between two seemingly unrelated observations in neuroscience. The first is the observation of strong correlations between the synaptic strengths in pairs of cortical pyramidal cells, the main excitatory neuronal type in cerebral cortex, by multiple groups using in vitro electrophysiological recordings [9–12]. These correlations are a consequence of two features of the connectivity: First, there exists an overrepresentation of bidirectionally connected pairs, compared to an Erdős-Rényi network with the same connection probability. For instance, Song et al (2005) found a connection probability of $c = 0.116$ in pairs of neurons whose somas are less than 100 $\mu$m apart, while the probability that a pair of such neurons are connected bidirectionally is approximately $4c^2$. This degree of overrepresentation has been found in multiple cortical areas, except in barrel cortex where no such overrepresentation exists [54]. Second, synaptic connections in bidirectionally connected pairs are on average stronger than those in unidirectionally connected pairs, and are significantly correlated [10]. These observations lead to estimates of $\eta \sim 5$, a value which our model predicts would lead to a significant increase in autocorrelation time scales compared to a random asymmetric connectivity.

The second is the observation of long time scales in the autocorrelations of neuronal activity from in vivo electrophysiological recordings (see e.g. [55]). Interestingly, the time scales of these autocorrelations increase from sensory to higher level areas such as the prefrontal cortex. Several mechanisms have been proposed to account for this phenomenon: differences in the level of expression of slow NMDA receptors [56], or increase in the strength of recurrent connectivity [57]. Our results suggest that this increase in time scale could also be due to an increase in the degree of symmetry of cortical connectivity. This would be consistent with the study of [14], who showed that the overrepresentation of bidirectionally connected pairs of neurons is significantly stronger in prefrontal cortex than in visual cortex.

From a neuroscience point of view, the model considered here is an extremely simplified model of cortical networks because it lacks the fundamental constraint that neurons are either excitatory or inhibitory, and because it does not constrain firing rates to be positive. These simplifications were made for the sake of mathematical tractability. A few recent studies have investigated how these two constraints influence the dynamics of such networks [19–21, 58–60]. Extending those works to connectivity with segregated excitation and inhibition and partial symmetry is an important direction for future work that might be facilitated by recent developments in random matrix theory [61, 62].

Acknowledgments

We thank Johnatan Aljadeff for his comments on a previous version of the manuscript. The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme FP7/2007–2013/ under REA grant agreement 301671. This has also been funded by the Programme Emergences of City of Paris, and the program “Investissements d’Avenir” launched by the French Government and implemented by the ANR, with the references ANR-10-LABX-0087 IEC and ANR-11-IDEX-0001-02 PSL* Research University.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A: Derivation of the double complex integral

We summarize here the derivation of the double complex integral of Eq. (9). Before doing so, we sketch the derivation of the local density of the overlap done by [16], as this will let us introduce some notation and pave the way for our calculation.

For a complex variable $z = x + iy$, with $x$ and $y$ real and with conjugate $\bar{z} = x - iy$, we define the Wirtinger derivatives $\partial/\partial z = (\partial/\partial x - i\partial/\partial y)/2$, $\partial/\partial \bar{z} = (\partial/\partial x + i\partial/\partial y)/2$, which obey $\partial z/\partial z = \partial \bar{z}/\partial \bar{z} = 1$ as well as $\partial z/\partial \bar{z} = \partial \bar{z}/\partial z = 0$. The complex differential is defined to be $d^2z \equiv dz d\bar{z} = 2 dz dy$, where the factor 2 comes from the Jacobian. We also define the complex Dirac delta so that it obeys the relation $\int d^2z \delta^2(z) = 1$, which implies $\delta^2(z) = (1/2)\delta(x)\delta(y)$ given our convention for the complex differential. Two useful identities for the delta function in the complex plane are

$$\delta^2(z) = \frac{1}{2\pi} \frac{\partial}{\partial \bar{z}} \bar{z} = \frac{1}{2\pi} \frac{\partial}{\partial z} \bar{z}, \quad (A1)$$

which can be checked by integrating over $d^2z$ and applying the complex version of Green’s theorem:

$$\int \left( \frac{\partial v}{\partial z} + \frac{\partial \bar{v}}{\partial \bar{z}} \right) d^2z = i \int (v d\bar{z} - \bar{v} dz), \quad (A2)$$

where $v$ and $\bar{v}$ are to be considered independent functions.

The resolvent, defined for any matrix $J$ as $(zI - J)^{-1}$, is a key quantity in the analysis of random matrices because it can be ensemble-averaged using standard methods and can be related to quantities of interest. So, for example, the empirical density of eigenvalues of a given $J$,

$$\rho_J(z) = \frac{1}{N} \sum_{i=1}^{N} \delta(x - Re \lambda_i) \delta(y - Im \lambda_i) = \frac{2}{N} \sum_{i=1}^{N} \delta^2(z - \lambda_i),$$

can be expressed thanks to the identities (A1) as

$$\rho_J(z) = \frac{1}{\pi} \frac{\partial}{\partial z} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{z - \lambda_i} = \frac{1}{\pi} \frac{\partial}{\partial z} \frac{1}{N} \text{Tr}(zI - J)^{-1}.$$
large \( N \), where all empirical densities converge to the average density

\[
\rho(z) = [\rho_J(z)]_j = \frac{1}{\pi} \frac{\partial}{\partial z} \left[ \frac{1}{N} \text{Tr}(z \mathbf{1} - \mathbf{J})^{-1} \right]_j.
\]

Deriving the average density, therefore, amounts to computing the function \( G(z) = \left[ \text{Tr}(z \mathbf{1} - \mathbf{J})^{-1} / N \right]_j \) in the large \( N \) limit.

The local density of the overlap can be derived in a similar manner using the spectral decomposition \((z \mathbf{1} - \mathbf{J})^{-1} = \sum_{i=1}^{N} \mathbf{R}_i (z - \lambda_i)^{-1} \mathbf{L}_i^\dagger \) where \( \mathbf{R}_i \) and \( \mathbf{L}_i \) are the right and left eigenvectors of \( \mathbf{J} \). If we substitute the definition of the overlap matrix into Eq. (8) and we use the identities (A1) we obtain

\[
D(z_1, z_2) = \frac{4}{N} \left[ \sum_{i,j=1}^{N} (\mathbf{L}_i^\dagger \mathbf{L}_j)(\mathbf{R}_i^\dagger \mathbf{R}_j) \delta^2(z_1 - \lambda_i) \delta^2(z_2 - \lambda_j) \right]_j
\]

\[
= \frac{1}{N\pi^2} \frac{\partial}{\partial z_1} \frac{\partial}{\partial z_2} \left[ \sum_{i,j=1}^{N} \text{Tr} \mathbf{R}_i \mathbf{L}_j^\dagger \mathbf{L}_j - \frac{1}{z_1 - \lambda_i} \frac{1}{z_2 - \lambda_j} \right]_j,
\]

and the problem reduces to computing the quantity

\[
G(z_1, z_2) = \left[ \frac{1}{N} \text{Tr} \frac{1}{z_1 \mathbf{1} - \mathbf{J}} \frac{1}{z_2 \mathbf{1} - \mathbf{J}} \right]_j.
\]

The expression of \( G(z_1, z_2) \) for the ensemble of Gaussian random matrices with partial symmetry was derived by [17]. The basic idea behind their calculation is to expand resolvents in power series, average over the disorder term by term, and organize the sums so that a recursive relation can be established and ultimately solved (for a thorough description of the method, see also [48]). The result is a complex function that takes the value

\[
G(z_1, z_2) = \frac{1}{\pi} \left[ \left( 1 - \eta^2 \right)^2 + \eta (z_1^2 + z_2^2) - (1 + \eta^2) z_1 z_2 \right] - 1
\]

when both \( z_1 \) and \( z_2 \) lie inside the ellipse centered at the origin and which has major and minor radii \( 1 + \eta \) and \( 1 - \eta \), respectively. We will call this ellipse \( E_\eta \) for later convenience. When \( z_2 \) lies outside \( E_\eta \) we have instead

\[
G(z_1, z_2) = \frac{h_1 h_2}{1 - h_1 h_2},
\]

where

\[
h_1 = \frac{z_1 - \sqrt{z_1^2 - 4\eta}}{2\eta}, \quad h_2 = \frac{z_2 - \sqrt{z_2^2 - 4\eta}}{2\eta}.
\]

Right on the ellipse \( E_\eta \), \( |h_1| = 1 \). When both \( z_1 \) and \( z_2 \) lie outside the ellipse, the function \( G(z_1, z_2) \) is analytic on \( z_1 \) and \( z_2 \). This analyticity implies, from (A3), that the local density of the overlap vanishes outside the ellipse.

We now proceed to compute \( A(u, \tau) \), Eq. (10). Inserting the identity (A3) into (10) leads to

\[
A(u, \tau) = \frac{1}{4\pi^2} \int \int e^{\theta(z_1 + z_2)u + g z_1 \tau} \frac{\partial}{\partial z_1} \frac{\partial}{\partial z_2} G(z_1, z_2) d^2z_1 d^2z_2.
\]

Because the exponential prefactor is analytic in \( z_1 \) and \( z_2 \), it commutes with the two partial derivatives. We can therefore apply Green’s theorem twice to obtain

\[
A(u, \tau) = \frac{1}{4\pi^2} \int_{E_\eta} \int e^{\theta(z_1 + z_2)u + g z_1 \tau} G(z_1, z_2) d^2z_1 d^2z_2,
\]

where both contour integrals are around the ellipse \( E_\eta \), at whose boundary \( G(z_1, z_2) \) stops being analytic. To compute \( A(u, \tau) \) we follow the approach of [17] and use the linear transformation \( w = (z - \eta \bar{z})/(1 - \eta^2) \) (or, equivalently, \( z = w + \eta \bar{w} \)) to reshape the contour of integration from the ellipse \( E_\eta \) into the unit circle. Applying this transformation to both \( z_1 \) and \( z_2 \), the surface integrals in (A5) become contour integrals on the unit circle \( |w|^2 = w \bar{w} = 1 \). On this contour we can replace every \( \bar{w} \) in the integrand by \( w^{-1} \) and we can use the standard tools of complex analysis to carry out the integrals.

We describe in more detail our derivation in the following.

We start performing the integral over \( z_2 \), expressing Eq. (A4) in terms of \( w_1, \bar{w}_1 \), and \( w_2, \bar{w}_2 \), replacing all \( \bar{w}_2 \) by \( w_2^{-1} \). After some simplifications, we obtain

\[
G(z_1(w_1), z_2(w_2)) = -1 + \frac{1}{\eta(\alpha_+ - \alpha_-)} \times \left( \frac{1 - \eta \alpha_+}{w_2 - \alpha_+} - \frac{1 - \eta \alpha_+}{w_2 - \alpha_-} \right),
\]

where we defined the poles

\[
\alpha_{\pm}(w_1) = \frac{w_1 + \eta \alpha_1 \pm \sqrt{(w_1 + \eta \alpha_1)^2 - 4\eta}}{2\eta}.
\]

These poles depend on \( w_1 \) and can be shown to map the unit disk onto an annulus of inner radius 1 and outer radius \( 1/\eta \) (see Fig. 6). This information will be relevant when we use residue calculus.

The double integral (A5) has to be regularized because the integrand diverges at \( z_1 = z_2 \). Our regularization consists of first integrating \( w_2 \) on the unit circle while constraining \( w_1 \) to be on a concentric circle of smaller radius \( |w_1| = 1 - \epsilon \), with \( \epsilon > 0 \) small. Once the integral over \( w_2 \) is done, we take the limit \( \epsilon \to 0 \) and perform the second integral over \( w_1 \).

Under this regularization, we decompose the double integral (A5) as

\[
A(u, \tau) = \lim_{\epsilon \to 0} \frac{1}{4\pi^2} \int_{|w_1| = 1 - \epsilon} e^{\theta(w_1 + \eta \bar{w}_1)(u + \tau)} A(w_1, u) \times (dw_1 + \eta d\bar{w}_1).
\]
where we used $z_1 = w_1 + \eta \bar{w}_1$ and we defined

$$A(w_1, u) = \oint_{|w_2|=1} e^{(\xi^2 + \eta w_2)u} G(z_1(w_1), z_2(w_2)) \times \left( -\frac{1}{w_2^2} + \eta \right) \, dw_2. \quad (A9)$$

Note that here we used $\bar{w}_2 = w_2^{-1}$ to express the integrand and the differential $d\bar{z}_2 = dw_2 + \eta d\bar{w}_2$ in terms of $w_2$ only. The integrand of Eq. (A9) contains one singularity inside the contour of integration, at $w_2 = 0$. This singularity is associated with the essential singularity from the exponential, $e^{1/w_2}$, as well as with the pole of second order $1/w_2^2$. Because we are assuming that $|w_1| < 1$, the poles of $G(z_1(w_1), z_2(w_2))$ at $w_2 = \pm \alpha$ lie outside the contour and therefore do not contribute to the integral. We do thus left with the task of computing the residue at the origin. We do that by expanding the integrand in Laurent series around $w_2 = 0$, using the relations

$$e^{(\xi^2 + \eta w_2)u} = \sum_{k=-\infty}^{\infty} \left( z/\sqrt{\eta} \right)^k I_k(2\sqrt{\eta}u), \quad \text{for } |z| \neq 0,$$

$$\left( z - z_0 \right)^{-1} = -\frac{1}{z_0} \sum_{k=0}^{\infty} \left( z/z_0 \right)^k, \quad \text{for } |z| < |z_0|.$$ We use the last power series to expand the terms $(w_2 - \alpha \pm)^{-1}$ in $G(z_1(w_1), z_2(w_2))$, Eq. (6). This power series converges because $|w_2| < |\alpha \pm|$ when $|w_1| < 1$, as we assuming in our regularization scheme. After expanding, applying Cauchy’s residue theorem, and taking the limit $\epsilon \to 0$, we obtain:

$$A(w_1, u) = \frac{2\pi i}{\sqrt{\eta}} \sum_{k=0}^{\infty} w_1^{-k} \eta^{-k/2} k I_k(2\sqrt{\eta}u),$$

with $I_k(z)$ being the modified Bessel function of order $k$. The final step is to compute the integral in (A8) with the same strategy we used for $A(w_1, u, \tau)$. In this case we express the integrand in terms of $w_1$ only, and we expand the exponential factor in (A8) with the identity

$$e^{(\xi^2 + \eta z - 1)u} = \sum_{k=-\infty}^{\infty} \left( z/\sqrt{\eta} \right)^k I_k(2\sqrt{\eta}u).$$

We then pick the residue from the expansion and apply Cauchy’s theorem. The result is

$$A(u, \tau) = A_1(u, \tau) + A_2(u, \tau),$$

with

$$A_1(u, \tau) = \sum_{k=-\infty}^{\infty} \eta^{k/2} I_k(2\sqrt{\eta}u) \left[ \left( \eta^2 \right) I_k(2\eta(1 + \eta)u) - \eta \left( I_{k-2}(2\eta(1 + \eta)u) + I_{k+2}(2\eta(1 + \eta)u) \right) \right] \quad (A10)$$

and $A_2(u, \tau)$ given by Eq. (13).

The expression (A10) for $A_1(u, \tau)$ can be further simplified with the identity [63]

$$\sum_{k=-\infty}^{\infty} e^{ik\alpha} J_k(w) J_{k+\nu}(z) = \int_{\tau} \left( \frac{z - we^{i\alpha}}{z - we^{-i\alpha}} \right)^{\nu/2} \times J_{\nu} \left( \sqrt{w^2 + z^2 - 2wz \cos \alpha} \right), \quad (A11)$$

which we can transform into a more convenient expression for our problem, using $J_{\nu}(iz) = i^\nu J_{-\nu}(z)$ and taking $\alpha = \pi + (1/2) \ln \eta$ so that $e^{i\alpha} = -\sqrt{\eta}$. The identity (A11) then becomes

$$\sum_{k=\infty}^{\infty} \eta^{k/2} I_k(w) I_{k+\nu}(z) = \left( \frac{z + w\eta^{-1/2}}{z + w\eta^{1/2}} \right)^{\nu/2} \times I_{\nu} \left( \sqrt{w^2 + z^2 + wz(\eta^{1/2} + \eta^{-1/2})} \right), \quad (A12)$$

which allows to rewrite Eq. (A10) as the final expression (12).

The series $A_2(u, \tau)$ does not seem to have a closed expression for general $\eta$. For $\eta = 1$, however, we can exploit the identity

$$\sum_{k=-\infty}^{\infty} I_k(w) I_{k-\nu}(z) = I_\nu(w+z)$$

to conclude that

$$A_2(u, \tau) = -I_0(2g(2u + \tau)) + I_2(2g(2u + \tau)) \quad \text{for } \eta = 1.$$  

Appendix B: Evaluation of the time integral for long $\tau$

The exact average autocorrelation is given in (9) as a time integral that can be decomposed as

$$C(\tau) = \int_0^{\infty} e^{-\tau} \left[ A_1(u, \tau) + A_2(u, \tau) \right] \, du,$$

where $A_1(u, \tau)$ and $A_2(u, \tau)$ are defined in Eqs. (12)–(13). To gain more analytical insight we will evaluate $C(\tau)$ when $\tau$ is sufficiently large, a limit that allows us to invoke Laplace’s method and approximate the integral with a closed-form expression [64]. Before applying the limit it is convenient to express this integral in terms of a new variable $\xi \equiv u/\tau$, which is well defined for $\tau > 0$. With this definition

$$C(\tau) = \int_0^{\infty} e^{-\tau(2\xi+1)} \left[ A_1(u, \xi) + A_2(u, \xi) \right] \, \tau d\xi,$$

$$\equiv C_1(\tau) + C_2(\tau). \quad (B1)$$

which we split in the two terms composing the integrand. We start with the asymptotic dependence of $C_1(\tau)$, ignoring for the moment $C_2(\tau)$. The integrand of $C_1(\tau)$ contains $I_{0}(\cdot)$ and $I_{2}(\cdot)$, whose argument is large in the long-$\tau$ limit. We can therefore use the asymptotic expansion of the modified Bessel functions of order $\nu$, $I_{\nu}(x) = \frac{e^{x}}{\sqrt{2\pi x}} \left[ 1 - \frac{2x}{8x} + O(x^{-2}) \right]$ for $x \gg 1$. (B2)

At this order, and defining

$$\psi(\xi; \eta) \equiv \psi(u(\xi), \tau; \eta) / \tau = 2\sqrt{(1 + \eta^2)\xi} \left( \xi + 1 \right) + \eta,$$
Figure 7. Sign of $\xi^*$, at which the exponent in Eq. (B3) is largest, as a function of the spectral gap and the degree of symmetry. This diagram is equivalent to that shown in Figure 3a, after transforming $\delta$ into its associated gain $g = (1 - \delta)/(1 + \eta)$.

we obtain

$$C_1(\tau) = \sqrt{2\pi} \int_0^\infty \frac{\exp\left\{ -\tau \left[ 2(\eta + 1) - g\psi(\xi; \eta) \right] \right\}}{\sqrt{g\psi(\xi; \eta)}} d\xi,$$

$$C_1(\tau) = \left[ (1 + \eta^2) \left( 1 + \frac{1}{8\tau g\psi(\xi; \eta)} \right) - 2\eta \left( 1 + \frac{2(1 - \eta)^2}{\psi(\xi; \eta)^2} \right) \right] d\xi.$$  \hspace{1cm} (B3)

with $g = (1 - \delta)/(1 + \eta)$. This integral is of the form

$$A(\tau) = \int_0^\infty f(\xi) e^{\tau b(\xi)} d\xi.$$  \hspace{1cm} (B4)

In the limit of large $\tau$, only the infinitesimal interval around the maximum of $b(\xi)$ contributes to the integral because the contribution of the remaining intervals is exponentially suppressed. In our particular case the maximum of $b(\xi)$ is at

$$\xi^* = \frac{1}{2} \left( -1 + \frac{1 - \eta}{(1 + \eta)\sqrt{2\delta - \delta^2}} \right).$$  \hspace{1cm} (B5)

We distinguish two cases. For large enough values of $\eta$ and $\delta$, $\xi^*$ is negative (Figure 7), which means that in the integration range $[0, \infty)$ of Eq. (B3) the maximum value of $b(\xi)$ is at $\xi = 0$. Conversely, for low values of $\eta$ and $\delta$, the maximum of $b(\xi)$ occurs within $(0, \infty)$. These two cases will lead to different time dependences and will be studied separately in the following.

For values of $\eta$ and $\delta$ such that $\xi^* < 0$, the limit $\tau \to \infty$ of Eq. (B4) can be approximated by [64, pp. 266-268]

$$A(\tau) \sim \lim_{\epsilon \to 0} \int_0^\epsilon \frac{f(0) e^{\tau b(0) + b'(0) s}}{\tau b'(0)} ds,$$

$$\sim \int_0^\infty f(0) e^{\tau b(0) + b'(0) s} ds \sim - \frac{f(0) e^{\tau b(0)}}{\tau b'(0)},$$  \hspace{1cm} (B6)

where $b'(0)$ denotes the derivative of $b(\xi)$ evaluated at $\xi^* = 0$. Applying this approximation to (B3), we obtain

$$C_1(\tau) = \frac{\tau^{-3/2} \eta^{-1/4}}{2\sqrt{\pi}} \frac{(1 + \eta^2)}{\delta(1 + \eta) - 1 - \sqrt{\eta^2}(1 - \frac{1}{\delta})^{3/2}} \times \exp \left\{ -\tau \left( 1 - \sqrt{\eta^2} + 2\delta \sqrt{\eta} \right) \right\}.$$  \hspace{1cm} (B7)

Conversely, if $\eta$ and $\delta$ are such that $\xi^* > 0$, the maximum $\xi^*$ falls within the integration region and we can approximate the large-$\tau$ limit of the integral (B4) by [see 64, p. 267]

$$A(\tau) \sim \lim_{\epsilon \to 0} \int_{\xi^* - \epsilon}^{\xi^* + \epsilon} f(\xi^*) e^{\tau (b(\xi^*) + (s - \xi^*)^2 b''(\xi^*)/2)} ds,$$

$$\sim \int_{-\infty}^{\infty} f(\xi^*) e^{\tau (b(\xi^*) + (s - \xi^*)^2 b''(\xi^*)/2)} ds,$$

$$\sim \sqrt{2\pi} \frac{f(\xi^*) e^{\tau b(\xi^*)}}{\tau b''(\xi^*)}.$$  \hspace{1cm} (B8)

with $b''(\xi^*)$ denoting the second derivative of $b(\xi)$ evaluated at $\xi^*$. In this case Eq. (B3) is approximately given by

$$C_1(\tau) = \frac{\delta^{-1/2}(1 - \eta)^2}{2\sqrt{2}(1 - \delta)} \exp \left\{ -\tau \frac{1 - \eta}{1 + \eta} \sqrt{2\delta - \delta^2} \right\}.$$  \hspace{1cm} (B9)

We now turn to the asymptotic dependence of $C_2(\tau)$. The original form for $C_2(\tau)$ is

$$C_2(\tau) = \int_0^\infty \frac{e^{-(2\xi + 1) \tau}}{\sqrt{\xi}} \sum_{k=0}^\infty \eta^{k+2} (k + 1)^2 \frac{I_k + 1}{\eta^2 \xi^2} d\xi.$$  \hspace{1cm} (B8)

The integrand contains a product of Bessel functions that grows exponentially with $\tau$ (see Eq. (B2)), but this growth is kept in check by the exponential prefactor. To see this, we introduce the scaled modified Bessel function,

$$I'_s(x) \equiv e^{-x^2} I_s(x),$$

in terms of which (B8) becomes

$$C_2(\tau) = \int_0^\infty \frac{e^{-(2\xi + 1) \tau}}{\xi} \sum_{k=0}^\infty \eta^{k+1} (k + 1)^2 \frac{I_k + 1}{\eta^2 \xi^2} d\xi.$$  \hspace{1cm} (B9)

The integral converges because the exponential decays to zero for large $\tau$ and trumps the power-law decay of $I_k(\xi)$. Equation (B9) also has the same form as (B4) and contains exactly the same exponent as in Eq. (B7); because the exponential attains its maximum at $\xi = 0$, we can use the approximation (B6). To evaluate at $\xi = 0$ we use the power expansion $I'_k(2\sqrt{\xi}) \equiv (g \tau \sqrt{\eta})^k e^{(k+1)}$, and identify $f(\xi)$ in (B6) with

$$f(\xi) \equiv \sum_{k=0}^\infty \eta^{(k+1)/2} (k + 1)^2 \frac{e^{-3/2 \eta^{1/4} \xi^2 \tau} I'_{k+1}(2\sqrt{\eta}(\xi + 1))}{\xi + 1}.$$  \hspace{1cm} (B10)

All the terms of $f(\xi)$ with $k > 0$ vanish at $\xi = 0$, and equation (B6) reads in this case

$$C_2(\tau) = \frac{\tau^{-3/2} \eta^{5/6}}{(1 - \delta)^{3/2}} \frac{1}{1 - \sqrt{\eta^2} + 2\delta \sqrt{\eta}} \frac{1}{1 + \eta} \exp \left\{ -\tau \left( 1 - \sqrt{\eta^2} + 2\delta \sqrt{\eta} \right) \right\}.$$  \hspace{1cm} (B11)

Summing this contribution to that in Eq. (B7) leads to the result reported in Eqs. (18)–(20).
Appendix C: Autocorrelation without overlaps

Here we compute the autocorrelation ignoring the effect of the overlaps between eigenvectors. In this case, we need to compute the individual contribution of a single eigenvalue to the autocorrelation, and then sum over the contributions of all eigenvalues. We start with the one-dimensional version of Eq. (4)

$$\frac{dx}{dt} = \alpha x + \sigma \xi(t),$$  \hspace{1cm} (C1)

where the parameter $\alpha$ would be the (single) eigenvalue of the system, assumed to have negative real part to prevent $x(t)$ to blow, and where $\xi(t)$ is a source of standard Gaussian white noise. The solution of (C1) is

$$x(t) = \sigma \int_{-\infty}^{t} e^{\alpha(t-s)} \xi(s) \, ds,$$

from which we can derive the autocorrelation:

$$\langle x(t)x(t+\tau) \rangle = \sigma^2 \int_{-\infty}^{t} \int_{-\infty}^{t+\tau} e^{\alpha(2t+\tau-s-u)} \langle \xi(s)\xi(u) \rangle \, du \, ds$$

$$= -\sigma^2 e^{\alpha \tau} \equiv \sigma^2 C_a(\tau).$$  \hspace{1cm} (C2)

The eigenvalue $\alpha$ determines both the time scale and the amplitude of the autocorrelation. We set the overall factor $\sigma^2$ to 1, without without loss of generality.

The average autocorrelation for the high-dimensional system in the absence of overlaps is the sum of (C2) over all the eigenvalues. In the large $N$ limit we would have

$$C(\tau) = \frac{1}{N} \sum_{i=1}^{N} C_a(\tau) \frac{N \to \infty}{N} \int C_a(\tau) \rho(\alpha) \, d\alpha,$$  \hspace{1cm} (C3)

where $\rho(\alpha)$ is the probability density of eigenvalues and where the integral is on the complex plane. For the system (4) and for the connectivity matrices we consider, the density of the eigenvalues $\alpha$ is uniform and has support on an ellipse centered at $z = -1$ with major radius $g(1+\eta)$ and minor radius $g(1-\eta)$. The integral (C3) can be computed in that case and reads

$$C(\tau) = \frac{-1}{\pi g^2 (1-\eta^2)} \int_0^{\pi} \frac{e^{\alpha \tau}}{2\alpha} \, d\alpha,$$  \hspace{1cm} (C4)

where we used Eq. (C2) and where the prefactor is the constant value that $\rho(\alpha)$ takes on the elliptic support $E$. To evaluate the integral we use the parametrization

$$\alpha = -1 + r(1+\eta) \cos \theta + ir(1-\eta) \sin \theta$$

and integrate over $r \in [0, g]$ and $\theta \in [0, 2\pi]$. Noting that

$$\int_E d\alpha = (1-\eta^2) \int_0^{2\pi} \int_0^g r \, dr \, d\theta,$$

Eq. (C4) becomes

$$C(\tau) = \frac{1}{\pi g^2} \int_0^{2\pi} \int_0^g r \exp \left\{-\frac{\tau}{1-r \psi(\theta)} \right\} \, dr \, d\theta,$$  \hspace{1cm} (C5)

where for convenience we defined

$$\psi(\theta) \equiv (1+\eta) \cos \theta + i(1-\eta) \sin \theta.$$

The integral (C5) is hard to compute, but we can make progress by taking the derivative of $C(\tau)$ with respect to $\tau$

$$C'(\tau) = -\frac{e^{-\tau}}{2\pi g^2} \int_0^{2\pi} \int_0^g e^{\tau \psi(\theta)} r \, dr \, d\theta,$$  \hspace{1cm} (C6)

which is easier to evaluate. Eq. (C6) can be integrated over $r$ by parts, yielding an integral over $\theta$ only that, excluding prefactors, reads

$$B(g, \tau) \equiv \int_0^{2\pi} \left\{ \frac{e^{\tau \psi(\theta)}}{\tau \psi(\theta)} - \frac{e^{\tau \psi(\theta)}}{\tau^2 \psi^2(\theta)} - 1 \right\} \, d\theta.$$

Again, this integral is hard to compute but we can use the same trick we used before, noting that the partial derivative of $A(g, \tau)$ with respect to $g$ simplifies considerably:

$$\frac{\partial B(g, \tau)}{\partial g} = g \int_0^{2\pi} e^{\tau \psi(\theta)} \, d\theta = 2\pi g I_0(2g\tau \sqrt{\eta}),$$

where in the last equation we used [65, 3.937.2, p. 496]. We recover the expression for $A(g, \tau)$ by integrating along $g$, with initial condition $A(0, \tau) = 0$:

$$B(g, \tau) = 2\pi \int_0^g x I_0(2x\tau \sqrt{\eta}) \, dx = \frac{2\pi}{4\tau^2 \eta} \int_0^{2\tau \sqrt{\eta}} y I_0(y) \, dy.$$

The last integral can be computed with the help of the recurrence relation $z I_0(z) = z I_1(z) + I_0(z)$. An integration by parts of term $z I_1(z)$ leads to the final identity $\int x I_0(x) \, dx = x I_1(x)$ and therefore to

$$B(g, \tau) = \frac{\pi g}{\tau \sqrt{\eta}} I_1(2\tau g \sqrt{\eta}).$$

Equation (C6) then reads

$$C'(\tau) = -\frac{e^{-\tau}}{2\tau g \sqrt{\eta}} I_1(2\tau g \sqrt{\eta}),$$  \hspace{1cm} (C7)

which we have to integrate to recover $C(\tau)$. Such an integration is subject to the initial condition $C(0)$:

$$C(0) = \frac{1}{2\pi g^2} \int_0^{2\pi} \int_0^g r \, dr \, d\theta$$

$$= -\frac{1}{2\pi g^2} \int_0^{2\pi} \left\{ \frac{g}{\psi(\theta)} + \frac{1}{\psi^2(\theta)} \ln \left[ 1 - g \psi(\theta) \right] \right\} \, d\theta,$$

which can be evaluated numerically.

Appendix D: Summary of the dynamic mean field derivation

The starting point of the calculation is the moment generating functional for the state variables $x_i(t)$ obeying Eq. (1). We consider the more general case where the activation variable is driven by recurrent inputs as well as independent external white noise:

$$\dot{x}_i(t) = -x_i(t) + g \sum_{j=1}^{N} J_{ij} r_j(t) + \sigma \xi_i(t), \hspace{1cm} i = 1, \ldots, N$$  \hspace{1cm} (D1)
where we defined \( r_j(t) \equiv \phi(x_j(t)) \) to lighten the notation.

The white noise sources \( \xi(t) \) have zero mean and unit variance. The moment generating functional for such a system can be shown to be \([66, 67]. See 68, 69 for a more pedagogical description.\]

\[
Z[l, \tilde{I}; J] = \int \mathcal{D}x(t) \mathcal{D}\tilde{x}(t) \exp \left\{ -S[x, \tilde{x}; J] \right\} \\
+ \sum_{i=1}^{N} \int \tilde{l}_i(t) x_i(t) \, dt + \sum_{i=1}^{N} \int \tilde{l}_i(t) \tilde{x}_i(t) \, dt, \\
\]

where \( \mathcal{D}x(t) \mathcal{D}\tilde{x}(t) = \prod_{i=1}^{N} \mathcal{D}x_i(t) \mathcal{D}\tilde{x}_i(t) \) is the functional measure for all possible paths for all variables, and where we introduced the action

\[
S[x, \tilde{x}; J] = \sum_{i=1}^{N} \int \tilde{x}_i(t) \\
\times \left\{ \tilde{x}_i(t) + x_i(t) - g \sum_{j=1}^{N} J_{ij} r_j(t) - \frac{\sigma^2}{2} \tilde{x}_i(t) \right\} \, dt \\
\tag{D2}
\]

In this definition we assume that the auxiliary fields \( \tilde{x}(t) \) are purely imaginary, so that we don’t have to write explicit imaginary units all along. By construction the generating functional satisfies the normalization condition \( Z[0, 0; J] = 1 \). The fact that \( Z[0, 0; J] \) does not depend on \( J \) allows us to compute the quenched average directly on \( Z \)[70],

\[
Z[l, \tilde{I}] = \int \frac{Z[l, \tilde{I}; J]}{Z[0, 0; J]} \, dP(J) = \int Z[l, \tilde{I}; J] \, dP(J), \\
\tag{D3}
\]

which simplifies considerably the average, now reduced to computing \( \exp(-S[x, \tilde{x}, J]) \). To do so, we use the decomposition of partially symmetric connectivity matrices

\[
J_{ij} = J_{ij}^0 + k J_{ij}^s, \\
\tag{D4}
\]

where \( J_{ij}^0 = J_{ji}^0 \), \( J_{ij}^s = -J_{ji}^s \), and where both \( J_{ij}^0 \) and \( J_{ij}^s \) are Gaussian random variables with zero mean and variance

\[
\left[ (J_{ij}^0)^2 \right]_J = \left[ (J_{ij}^s)^2 \right]_J = \frac{1}{N} \frac{1 - k^2}{1 + k^2},
\]

so that \( \left[ J_{ij}^2 \right]_J = J^2 / N \). With these matrix decompositions, the correlation between bidirectional weights is [15]

\[
\left[ J_{ij} J_{ji} \right]_J = \frac{1}{N} \frac{1 - k^2}{1 + k^2},
\]

which must equal \( \eta/N \) by our definition of \( \eta \). This leads to the relation \( k^2 = (1 - \eta)/(1 + \eta) \). To integrate over the disorder we use the Gaussian measures:

\[
dP(J^s) = \prod_{i<j} dP(J_{ij}^s) \propto \exp \left\{ -\frac{N}{1 + \eta} \sum_{i<j} (J_{ij}^s)^2 \right\} \, dJ^s, \\
dP(J^0) = \prod_{i<j} dP(J_{ij}^0) \propto \exp \left\{ -\frac{N}{1 + \eta} \sum_{i<j} (J_{ij}^0)^2 \right\} \, dJ^0,
\]

with \( dJ^s = \prod_{i<j} dJ_{ij}^s \) and \( dJ^0 = \prod_{i<j} dJ_{ij}^0 \). We will ignore the contribution of diagonal elements of the synaptic matrix because it is negligible in the limit of large \( N \). We can now integrate out the terms linear in \( J_{ij} \) that appear in Eq. (D2), by separating symmetric and antisymmetric components. Excluding prefactors and time integrals, these terms are of the form

\[
L(J, t) \equiv \sum_{i,j} \tilde{x}_i(t) J_{ij} r_j(t) = \sum_{i,j} \tilde{x}_i(t) [J_{ij}^0 + k J_{ij}^s] r_j(t) = \sum_{i,j} \left\{ J_{ij}^0 \left[ \tilde{x}_i(t) r_j(t) + \tilde{x}_j(t) r_i(t) \right] + k J_{ij}^s \left[ \tilde{x}_i(t) r_j(t) - \tilde{x}_j(t) r_i(t) \right] \right\},
\]

so that

\[
\int \exp \left\{ g \int L(J, t) \, dt \right\} \, dP(J^s) dP(J^0) = \exp \left\{ \frac{g^2}{2N} \sum_{i,j} \int \left\{ \tilde{x}_i(t) r_j(t) \tilde{x}_i(t') r_j(t') \right\} \, dt \right\},
\]

which we used the property that, for a Gaussian variable \( z \) of zero mean and variance \( \sigma^2 \), the expected value of \( \exp(\lambda z) \) is \( \exp(\lambda^2 \sigma^2/2) \), which can be checked by completing the square in the exponential.

Putting back all the pieces together, the average generating functional, Eq. (D3), is therefore

\[
Z[l, \tilde{I}] = \int \mathcal{D}x(t) \mathcal{D}\tilde{x}(t) \exp \left\{ -S_0[x(t), \tilde{x}(t)] \right\} \\
+ \frac{\sigma^2}{2} \tilde{x} \cdot \tilde{x} + \tilde{l} \cdot \tilde{x} + l \cdot \tilde{x} + \frac{g^2}{2N} \sum_{i,j} \int \left\{ \tilde{x}_i(t) r_j(t) \tilde{x}_i(t') r_j(t') \right\} \, dt \right\},
\tag{D5}
\]

where we defined the free action

\[
S_0[x, \tilde{x}] \equiv \sum_{i=1}^{N} \int \tilde{x}_i(t) \left[ \tilde{x}_i(t) + x_i(t) \right] \, dt. \\
\tag{D6}
\]

and we introduced the notation

\[
f \cdot g \equiv \sum_{i=1}^{N} \int f_i(t) g_i(t) \, dt.
\]

As a result of averaging out the disorder, we obtained a coupling involving four fields with different indices and at different times. To proceed it is convenient to introduce auxiliary fields that involve terms local in space (i.e., with the same index) but not in time:

\[
q_1(t, t') = \frac{g^2}{N} \sum_{i=1}^{N} \tilde{x}_i(t) \tilde{x}_i(t'), \quad q_2(t, t') = \frac{g^2}{N} \sum_{i=1}^{N} r_i(t) r_i(t'), \quad q_3(t, t') = \frac{g^2}{N} \sum_{i=1}^{N} \tilde{x}_i(t) r_i(t'), \quad q_4(t, t') = \frac{g^2}{N} \sum_{i=1}^{N} r_i(t) \tilde{x}_i(t'),
\]

with \( x_i(t) = \tilde{x}_i(t) \) and \( r_i(t) = r_i(t) \).
so that Equation (D5) now reads

$$Z[l, \tilde{l}] = \int Dx(t) D\tilde{x}(t) \left( \prod_{\alpha=1}^{4} \int \frac{N}{g^2} Dq_{\alpha} \right)$$

$$\times \exp \left\{ \frac{1}{2\pi} \int D\tilde{q}_1(t, t') \right. - \sum_{i=1}^{N} \tilde{x}_i(t) \tilde{x}_i(t') \left. \right\} dt dt'$$

$$\times \exp \left\{ \frac{1}{2\pi} \int D\tilde{q}_2(t, t') \right. - \sum_{i=1}^{N} \tilde{x}_i(t) \tilde{x}_i(t') \left. \right\} dt dt'$$

$$\times \exp \left\{ \frac{1}{2\pi} \int D\tilde{q}_3(t, t') \right. - \sum_{i=1}^{N} \tilde{x}_i(t) \tilde{x}_i(t') \left. \right\} dt dt'$$

$$\times \exp \left\{ \frac{1}{2\pi} \int D\tilde{q}_4(t, t') \right. - \sum_{i=1}^{N} \tilde{x}_i(t) \tilde{x}_i(t') \left. \right\} dt dt'$$

$$+ \frac{N}{g^2} \int \left\{ \frac{4}{g^2} \sum_{\alpha=1}^{N} \tilde{q}_\alpha(t, t') q_{\alpha}(t, t') \right. - \frac{1}{2} \left[ q_1(t, t') q_1(t, t') + q_2(t, t') q_2(t, t') + q_3(t, t') q_3(t, t') + q_4(t, t') q_4(t, t') \right]$$

$$+ \frac{g^2}{2} \sum_{i=1}^{N} \left[ q_1(t, t') \tilde{x}_i(t) \tilde{x}_i(t') \right. + q_2(t, t') r_i(t) r_i(t')$$

$$\left. + q_3(t, t') \tilde{x}_i(t) r_i(t') + q_4(t, t') r_i(t) \tilde{x}_i(t') \right] \right\} dt dt'$$

Equation (D8) can now be expressed as [5, 30]

$$Z[l, \tilde{l}] = \int DQ e^{Nf(q, \hat{q}, \bar{x})},$$

where

$$f(q, \hat{q}, \bar{x}) \equiv G(q, \hat{q}) + \frac{1}{N} \log \int DX \exp \left[ \mathcal{L}(q, \hat{q}, \bar{x}) \right],$$

$$G(q, \hat{q}) \equiv \frac{1}{g^2} \int \left\{ \sum_{\alpha=1}^{N} q_{\alpha} \hat{q}_{\alpha} + \frac{1}{2} \left[ q_1 q_2 + \eta q_3 q_4 \right] \right\} dt dt',$n

$$\mathcal{L}(q, \hat{q}, \bar{x}) \equiv -S_0[q, \hat{q}] + \frac{g^2}{2} \bar{x} \cdot \bar{x} + \bar{l} \cdot \bar{x}$$

$$- \sum_{i=1}^{N} \int \left[ \hat{q}_1(t, t') \tilde{x}_i(t) \tilde{x}_i(t') + \hat{q}_2(t, t') r_i(t) r_i(t')$$

$$+ \hat{q}_3(t, t') \tilde{x}_i(t) r_i(t') + \hat{q}_4(t, t') r_i(t) \tilde{x}_i(t') \right] dt dt'.$$
The auxiliary fields defined in Eqs. (D11)–(D14) are related to physically observable quantities. First, $q_i^2(t, t')$ is related to the population-averaged autocorrelation function

$$C(t, t') \equiv \frac{1}{N} \sum_{i=1}^{N} \langle r_i(t) r_i(t') \rangle,$$

by $q_i^2(t, t') = g^2 C(t, t')$.

Second, the auxiliary fields $q_i^0(t, t')$ and $q_i^1(t, t')$ are related to the so-called response function, which characterizes the response of the system when it is perturbed by a weak field. More specifically, in our context the response function at site $t$ would be

$$G(t, t') \equiv \frac{\delta \langle r_i(t) \rangle}{\delta h_i(t')} \bigg|_{h_i = 0}, \quad (D17)$$

where $h_i(t')$ is a time-dependent external field, and angular brackets denote the average over the effective action $S[x, \tilde{x}] = S_0[x, \tilde{x}] + S_{\text{int}}[x, \tilde{x}]$ that appears in Eq. (D15). Note that from the definition of response function $G(t, t')$ has to be 0 whenever $t < t'$, due to causality. To see the link between $G(t, t')$ and $q_i^0(t, t')$ and $q_i^1(t, t')$, we add an external field $h_i(t)$ for each neuron in Eq. (D1), and evaluate (D17). With the new field the action becomes $S_h[x, \tilde{x}] = S_0[x, \tilde{x}] + S_{\text{int}}[x, \tilde{x}]$ that appears in Eq. (D15). From the original stochastic system (D1) and its associated Martin-Siggia-Rose-Janssen-deDominicis (MSRJD) action (D2), we infer that the equation of motion associated with the action (D19) is

$$\dot{x}(t) = -x(t) + \eta g^2 \int_{-\infty}^{t} G(t, s) r(s) \, ds + \varphi(t), \quad (D20)$$

where $\varphi(t)$ is a source of noise with autocorrelation

$$\langle \varphi(t) \varphi(t') \rangle = \Gamma(t, t') = g^2 C(t, t') + \sigma^2 \delta(t - t').$$

This relation has to be consistent with the dynamics generated by Eq. (D20), that is, the noise $\varphi(t)$ has to be such that the firing activity $r(t)$ has autocorrelation $C(t, t')$.

We can go further and write a self-consistent relation involving the two-point functions $C(t, t')$ and $G(t, t')$. A starting point to derive them are the identities

$$\frac{\delta \langle x(t) \rangle}{\delta x(t')} = 0, \quad \frac{\delta \langle \dot{x}(t) \rangle}{\delta x(t')} = 0, \quad \frac{\delta \langle x(t) \rangle}{\delta \dot{x}(t')} = 0, \quad \frac{\delta \langle \dot{x}(t) \rangle}{\delta \dot{x}(t')} = 0.$$
and we define the autocorrelation and response function of the activation field \( x(t) \)

\[
\Delta(t, t') \equiv \langle x(t)x(t') \rangle, \quad R(t, t') \equiv \langle x(t)\tilde{x}(t') \rangle.
\]

The last equation in (D21) and the first equation in (D22) then become, respectively,

\[
\frac{\partial}{\partial t} \Delta(t, t') = -\Delta(t, t') + \sigma^2 R(t', t) \eta \phi^2 \int_0^{t'} G(t,s) \phi(s) x(t') \, ds \\
+ g^2 \int_0^{t'} R(t', s) C(t,s) \, ds,
\]

\[
\frac{\partial}{\partial t} R(t, t') = -R(t, t') + \delta(t-t') \\
+ \eta \phi^2 \int_0^{t'} G(t,s) G(s,t') \, ds,
\]

where in (D24) we have used \( \langle \tilde{x}(t) \tilde{x}(t') \rangle = 0 \). It can be shown that the remaining identities in Eqs. (D21)–(D22), which involve \( \delta S/\delta x \), do not provide additional information [69]. Note that \( \Delta(t, t') \) has a cusp at \( t = t' \) due to the term \( \sigma^2 R(t', t) \), which from (D22) we know must be of the form \( R(t, t') \propto \Theta(t-t') \), with \( \Theta(t) \) being the step function. More specifically,

\[
\left[ \frac{\partial}{\partial t} \Delta(t, t') \right]_{t'=t} = \sigma^2 [R(t', t)]_{t'=t} - \sigma^2.
\]

Moreover, the symmetry of \( \Delta(t, t') \) around \( t = t' \) implies \( \lim_{t' \to t^-} \partial_t \Delta(t, t') = -\lim_{t' \to t^+} \partial_t \Delta(t, t') \), which leads to the relation \( \lim_{t' \to t^-} \partial_t \Delta(t, t') = -\sigma^2/2 \). The amplitude of external noise thus determines the slope of the autocorrelation of \( x(t) \) at zero time lag. This is the only dependence on \( \sigma^2 \) of the solutions of (D23)–(D24).

Equations (D23)–(D24) cannot be solved in a closed-form except for \( \eta = 0 \) [1], but perturbative solutions can be found by expanding the nonlinearity \( r(t) = \phi(x) \) in power series of \( x(t) \) and then solving the resulting hierarchy of equations, which involve correlations and response functions of increasingly larger order. The problem becomes unwieldy except for the linear case where \( r(t) = x(t) \). In that case, \( C(t, t') = \Delta(t, t') \), \( G(t, t') = R(t, t') \), and Eqs. (D23)–(D24) form a closed system of integro-differential equations:

\[
\frac{\partial}{\partial t} \Delta(t, t') = -\Delta(t, t') + \sigma^2 R(t', t) + \eta \phi^2 \int_0^{t'} R(t', s) \Delta(s, t) \, ds \\
+ g^2 \int_0^{t'} R(t', s) G(t,s) \, ds,
\]

\[
\frac{\partial}{\partial t} R(t, t') = -R(t, t') + \delta(t-t') \\
+ \eta \phi^2 \int_0^{t'} R(t', s) R(s,t') \, ds.
\]
[31] B. Bravi, P. Sollich, and M. Opper, J. Phys. A-Math. 49, 194003 (2016).
[32] G. Hennequin, T. P. Vogels, and W. Gerstner, Phys. Rev. E 86, 011909 (2012).
[33] J.-P. Bouchaud, L. F. Cugliandolo, J. Kurchan, and M. Mézard, in Spin glasses and random fields, edited by A. P. Young (World Scientific, Singapore, 1998) Chap. 6, pp. 161–223.
[34] J. J. Hopfield, P. Natl. Acad. Sci. USA 79, 2554 (1982).
[35] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A 32, 1007 (1985).
[36] J. A. Hertz, P. Natl. Acad. Sci. USA 79, 2554 (1982).
[37] B. Derrida, E. Gardner, and A. Zippelius, Europhys. Lett. 4, 167 (1987).
[38] C. van Vreeswijk and H. Sompolinsky, Science 274, 1724 (1996).
[39] A. Litwin-Kumar and B. Doiron, Nat. Neurosci. 15, 1498 (2012).
[40] A. Roxin, Front. Comput. Neurosci. 5, 8 (2011).
[41] L. Zhao, B. Beverlin, T. Netoff, and D. Nykamp, Front. Comput. Neurosci. 5, 28 (2011).
[42] C. Bimbard, E. Ledoux, and S. Ostojic, Phys. Rev. E 94, 062207 (2016).
[43] N. Brunel, Nat. Neurosci. 19, 749 (2016).
[44] S. Ganguli, D. Huh, and H. Sompolinsky, P. Natl. Acad. Sci. USA 105, 18970 (2008).
[45] B. K. Murphy and K. D. Miller, Neuron 61, 635 (2009).
[46] M. S. Goldman, Neuron 61, 621 (2008).
[47] G. Hennequin, T. P. Vogels, and W. Gerstner, Neuron 82, 1394 (2014).
[48] Y. Ahmadian, F. Fumarola, and K. D. Miller, Phys. Rev. E 91, 012820 (2015).
[49] M. Mézard, G. Parisi, and M. Virasoro, Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications, Vol. 9 (World Scientific, 1987).
[50] L. F. Cugliandolo and J. Kurchan, Phys. Rev. Lett. 71, 173 (1993).
[51] L. F. Cugliandolo and D. S. Dean, J. Phys. A-Math. Gen. 28, 4213 (1995).
[52] G. Iori and E. Marinari, J. Phys. A-Math. Gen. 30, 4489 (1997).
[53] E. Marinari and D. A. Stariolo, J. Phys. A-Math. Gen. 31, 5021 (1998).
[54] S. Lefort, C. Tomm, J. C. F. Sarria, and C. C. Petersen, Neuron 61, 301 (2009).
[55] J. D. Murray, A. Bernacchia, D. J. Freedman, R. Romo, J. D. Wallis, X. Cai, C. Padou-Schioppa, T. Pasternak, H. Seo, D. Lee, and X. J. Wang, Nat. Neurosci. 17, 1661 (2014).
[56] H. Wang, G. G. Stradtman, X. J. Wang, and W. J. Gao, P. Natl. Acad. Sci. USA 105, 16791 (2008).
[57] G. N. Elston, Cereb. Cortex 13, 1124 (2003).
[58] K. Rajan and L. F. Abbott, Phys. Rev. Lett. 97, 188104 (2006).
[59] S. Ostojic, Nat. Neurosci. 17, 594 (2014).
[60] J. Aljadeff, M. Stern, and T. Sharpee, Phys. Rev. Lett. 114, 088101 (2015).
[61] A. Kuczala and T. O. Sharpee, Phys. Rev. E 94, 050101 (2016).
[62] J. Aljadeff, D. Renfrew, and M. Stern, J. Math Phys. 56, 103502 (2015).
[63] A. P. Prudnikov, I. Brychkov, and O. I. Marichev, Integrals and Series / Vol 2 - Special Functions/ (Gordon and Breach Science Publishers, 1992).
[64] C. M. Bender and S. A. Orszag, Advanced Mathematical Methods for Scientists and Engineers (Springer Verlag, 1999).
[65] I. S. Gradshteyn and I. M. Ryzhik, Table of integrals, series, and products, seventh ed. (Elsevier/Academic Press, Amsterdam, 2007) pp. xlviii+1171, translated from the Russian, Translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger, With one CD-ROM (Windows, Macintosh and UNIX).
[66] P. C. Martin, E. Siggia, and H. Rose, Phys. Rev. A 8, 423 (1973).
[67] H.-K. Janssen, Zeitschrift für Physik B Condensed Matter 23, 377 (1976).
[68] C. Chow and M. Buice, J. Math. Neurosci. 5, 8 (2015).
[69] L. F. Cugliandolo, “Out of equilibrium dynamics of complex systems,” (2013), lecture Notes.
[70] C. De Dominicis, Phys. Rev. B 18, 4913 (1978).