Single neuron properties shape chaotic dynamics in random neural networks

Samuel P. Muscinelli and Wulfram Gerstner

School of Computer and Communication Sciences and School of Life Sciences
École polytechnique fédérale de Lausanne (EPFL)
Station 15, CH-1015 Lausanne, Switzerland

Tilo Schwalger
Bernstein Center for Computational Neuroscience, 10115 Berlin, Germany and
Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany

Abstract

Cortical neurons exhibit rich dynamics, shaped by the presence of history-dependent biophysical mechanisms such as adaptation over different time scales. On the other hand, cortical circuits are highly recurrent which can result in chaotic network dynamics even if network nodes have simple dynamics. We analyze how the dynamics of nodes (such as single neurons) and recurrent connections interact to shape the effective dynamics in large randomly connected networks. A novel dynamical mean-field theory for networks of multi-dimensional rate units shows that, despite the nonlinear interactions between units, the recurrent network dynamics inherit, in the chaotic phase, the qualitative features of the linear response properties of the single units. For the case of two-dimensional rate units with strong adaptation, we find that the network exhibits a state of "resonant chaos", whose resonance frequency can be predicted from the properties of isolated units. The correlation time of the chaotic network scales with the adaptation timescale of a single unit. Our theory for network analysis can be applied to numerous other mechanisms at the level of single nodes, such as synaptic filtering, refractoriness or spike synchronization. These results advance our understanding of the interaction between single-neuron mechanisms and recurrent connectivity, a fundamental step toward the description of biologically realistic network models in the brain, or, more generally, networks of other physical or man-made complex dynamical units.
I. INTRODUCTION

The rich dynamics of large random networks of neuron-like elements can be exploited to learn complex tasks involving generation of temporal patterns [1–7]. In these approaches, the complex behavior of the network mainly arises from the random interactions, whereas the dynamics of single elements are typically given by simple first-order differential equations. The simplicity of single elements allows to quantitatively determine the chaotic phase of the coupled elements using dynamical mean-field theory (DMFT) [8]. A fascinating question is whether a similar chaotic phase also exists in brain-like networks, that are subject to additional biological constraints. To address this question, previous work has focused on incorporating more realistic connectivity structure into the DMFT framework [7, 9, 10]. Individual cortical neurons, however, exhibit rich multi-dimensional internal dynamics [11–14] that are inconsistent with first-order equations. Despite the biological relevance, a theoretical understanding of the emergent activity patterns in networks of more realistic multi-dimensional dynamical elements, is largely lacking. In particular, beyond a few simple cases [8, 15–17], a self-consistent mean-field theory for fluctuations such as the auto-correlation function or the power spectrum of the network activity is still an unsolved theoretical problem. Here, we develop a theoretical framework that extends DMFT to multi-dimensional rate neurons and allows to predict the emergent network dynamics based on single-neuron response properties.

Groups of neurons are often described using firing rate models. While these models discard information on the exact spike-timing of single neurons, they have the advantage of being accessible to an analytical characterization of their dynamics. However, commonly-used one-dimensional rate models cannot fully capture the dynamics of the mean activity of a population of spiking neurons. For example, classic rate models fail to account for rapid synchronization of neurons in response to a stimulus [18–20], an effect that is readily observed even in simple spiking neuron models, such as the leaky integrate-and-fire model [21–24]. To capture rapid synchronization after stimulus onset, in rate models it is necessary to consider at least two equations per rate neuron [25–27]. If we also want to account for additional biophysical mechanisms present at the single neuron level such as refractoriness [28], spike-frequency adaptation (SFA) [29–31], synaptic filtering [32, 33], subthreshold resonance [34] or for the effect of dendritic compartments [35, 36], multiple auxiliary variables and equations
must be added.

As an important application of our theoretical framework, we consider SFA, the decrease of a neuron’s firing rate in response to a sustained stimulus. SFA is present in neurons at all stages of sensory processing, and is believed to play a crucial role for efficient coding of external stimuli [11]. Moreover, SFA over multiple timescales represents an efficient solution for information transmission of sensory signals whose statistics change dynamically [12, 14, 37]. The interest for adaptation is motivated not only by biological realism, but also by recurrent neural network models that exploit adaptation to solve particular tasks. For example, adaptation has been proposed to play a role in sequential memory retrieval [38], slow activity propagation [39], perceptual bistability [40] and decision making [41]. Moreover, SFA has beneficial consequences both for reservoir computing approaches [5] and for spiking neuron-based machine learning architectures [42]. Here, we show that introducing adaptation in a rate model, a random network of such rate neurons can enter a “resonant” chaotic state qualitatively different from the chaotic behavior of the network without adaptation.

This paper is organized as follows: In section II we present the microscopic network model, for which in section III we discuss the fixed point stability. In section IV we present the mean-field theory description, that allows us to study the chaotic dynamics (section V). In each of these sections, we first present the general formalism and then apply the result to the case of adaptation. Finally, in sections V B and V C we analyze the consequences of the temporal correlation structure of the chaotic phase for the network with adaptation, focusing on the correlation time and on the response to external input, respectively.

II. MICROSCOPIC MODEL

We are interested in studying the dynamics of a randomly connected recurrent network of multi-dimensional firing-rate units where each unit is described by a set of $D$ variables $x^1_i, \ldots, x^D_i$. We assume that the first variable $x^1_i$ is an activation variable that defines the output rate $y$ via a nonlinear function $\phi$, i.e. $y_i(t) = \phi(x^1_i(t))$. More precisely, $\phi(x^1_i(t))$ should be interpreted as the deviation from some reference rate corresponding to the long-term average firing rate. Because in our model the average of $x^1_i$ is zero, we can thus assume $\phi(0) = 0$. Note that $\phi(x^1_i(t))$ can take positive and negative values.

The remaining $D - 1$ variables are auxiliary variables. In isolation, each unit obeys a
system of $D$ first-order linear differential equations

$$\dot{x}_i^\alpha(t) = \sum_{\beta=1}^{D} A^{\alpha\beta} x_i^\beta(t) ,$$  \hspace{1cm} (1)$$

where the dot denotes the temporal derivative. In what follows, subscripts (in Latin letters) indicate the index of the unit in the network and run from 1 to $N$, while superscripts (in Greek letters) indicate the index of the variable in the rate model and run from 1 to $D$. The matrix $A$ is assumed to be non-singular, have real entries and to have eigenvalues with negative real parts. We assume that the rate $\phi(x_j^1(t))$ is the only signal that unit $j$ uses to communicate with other units. Conversely, the signals coming from other units only influence the variable $x_i^\alpha$, i.e. the rate of unit $j$ is directly coupled only to one variable of unit $i$. Without loss of generality, we will assume $\alpha = 1$. Unit $i$ receives input from all the other units, via a set of random connections $J_{ij}$, sampled i.i.d. from a Gaussian distribution with mean zero and variance $g^2/N$. When incorporating these assumptions, the network equations read

$$\dot{x}_i(t) = \sum_{\beta=1}^{D} A^{\alpha\beta} x_i^\beta(t) + \delta^{\alpha1} \left( \sum_{j=1}^{N} J_{ij} \phi(x_j^1(t)) + I_i(t) \right) J_{ij} \sim \mathcal{N}(0, g^2/N)$$  \hspace{1cm} (2)$$

where $\delta^{\alpha\beta}$ is the Kronecker delta symbol. The external input $I_i(t)$ is assumed to have stationary statistics and zero mean.

**Network with adaptation.** As a biologically relevant two-dimensional example, we consider rate units that undergo firing rate adaptation. We have only one auxiliary variable that mediates adaptation and that we will indicate as $a_i(t) := x_i^2(t)$. To ease our notation we drop the superscript 1 in the first variable and write $x_i$ instead of $x_i^1$. The adaptation variable $a_i(t)$ of neuron $i$ is driven by the neuron activation $x_i(t)$ and provides negative feedback onto $x_i$. The network equations (Eq. (2)) for the adaptation case become

$$\dot{x}_i(t) = -x_i(t) + \sum_{j=1}^{N} J_{ij} \phi(x_j(t)) - a_i(t) + I_i(t) \hspace{1cm} (3)$$

$$\dot{a}_i(t) = -\gamma a_i(t) + \gamma \beta x_i(t) , \hspace{1cm} (4)$$

where the parameter $\gamma > 0$ can be interpreted as the ratio of the timescales of $x$ and $a$ ($\gamma = \tau_x/\tau_a$), while $\beta > 0$ is a parameter that controls the strength of adaptation. In
simulations, we choose $\phi(x)$ as a piecewise-linear function given by

$$
\phi_{PL}(x) = \begin{cases} 
-1 & \text{for } x < -1 \\
x & \text{for } -1 < x < 1 \\
1 & \text{for } x > 1 
\end{cases}
$$

(5)

unless stated otherwise.

III. STABILITY OF THE FIXED POINT

General theory. Equation (2) is a system of $N \cdot D$ coupled nonlinear differential equations that become intractable for large $N$. However, because $\phi(0) = 0$, the system has a fixed point at the origin $\{x^0_i = 0\}_{i=1}^{D}$, the stability of which can be studied owing to the clustered structure of the system. The Jacobian at the fixed point is given by

$$
B = \begin{pmatrix} 
A^{11}I_N + \phi'(0)J & A^{12}I_N & \cdots & A^{1D}I_N \\
A^{21}I_N & A^{22}I_N & \cdots & A^{2D}I_N \\
& \cdots & \cdots & \cdots \\
A^{D1}I_N & A^{D2}I_N & \cdots & A^{DD}I_N 
\end{pmatrix},
$$

(6)

where $J$ is the random connectivity matrix and $I_N$ is the $N$-dimensional identity matrix. The matrix $B$ is of size $ND \times ND$ and it therefore admits $ND$ eigenvalues. Since all the blocks of $B$ commute with each other, we can apply the result of [43] to find a relation between the eigenvalues of $J$, $A$ and $B$

$$
\lambda_J = \frac{\prod_{i=1}^{D} (\lambda_B - \lambda_A^i)}{\phi'(0) \prod_{j=1}^{B-1} (\lambda_B - \lambda_A^{-j})},
$$

(7)

where $A^-$ is the matrix obtained by removing the first column and the first row from the matrix $A$. This expression is valid for all the eigenvalues of $B$ that are not coincident with those of $A^-$. Equation (7) can be transformed into a degree-$D$ polynomial equation in $\lambda_B$, so that for every value of $\lambda_J$ we obtain $D$ eigenvalues of $B$, as expected. From now on we will assume that, without loss of generality, $\phi'(0) = 1$.

In the $N \to \infty$ limit, the eigenvalues $\lambda_J$ are known to be uniformly distributed on a disk in the complex plane, centered at zero and of radius $g$ [44]. If one can invert Eq. (7), it becomes computationally fast to compute the eigenvalues of the Jacobian in the $N \to \infty$
limit without finite-size effects. Whether one can obtain an explicit inverse formula depends on the dimensionality and on the entries of the matrix A.

**Network with adaptation.** For the two-dimensional model defined by Eqs. (3, 4), we can invert Eq. (7), and obtain an expression for the eigenvalues of the Jacobian

\[ \lambda_B(\lambda_J) = \frac{1}{2} \left(-1 - \gamma + \lambda_J \pm \sqrt{(\lambda_J - 1 + \gamma)^2 - 4\gamma\beta}\right). \] (8)

Using the mapping in Eq. (8), we find the critical value of \( g \) for which the stability of the fixed point is lost

\[ g_c(\gamma, \beta) = \begin{cases} \sqrt{1 - \gamma(\gamma + 2\beta) + 2\sqrt{\gamma^2(2\gamma + 2\beta + 2)}}, & \beta > \beta_H(\gamma) \\ 1 + \beta, & \beta \leq \beta_H(\gamma) \end{cases} \] (9)

where \( \beta_H(\gamma) = -1 - \gamma + \sqrt{2\gamma^2 + 2\gamma + 1} \). We will see below that the critical value of \( g \) can also be calculated from dynamical mean-field theory (see appendix C). Examples of the eigenvalue spectrum of \( B \) are shown in the insets of Fig. 1.

If \( g < g_c(\gamma, \beta) \), the network exhibits transient dynamics before it settles at the zero fixed point (Fig. 1a,b). From Eq. 9 we notice that \( g_c(\gamma, \beta) \geq 1 \), since both \( \gamma \) and \( \beta \) are positive. In the limit \( \gamma \to 0 \) or \( \beta \to 0 \) we retrieve the same dynamical regime as for the network without adaptation, for which \( g_c = 1 \) [8]. If adaptation increases due to an increase in either \( \gamma \) or \( \beta \), the value of \( g_c \) also increases (see Fig. 2a). Hence, adaptation increases the range of stability.

The bifurcation that characterizes the loss of stability depends on two parameters, viz. the ratio of timescales \( \gamma \) and the strength of the adaptation \( \beta \). To further characterize the bifurcation at \( g = g_c(\gamma, \beta) \), we can study the imaginary part of the critical eigenvalue, i.e. the one with real part equal to zero at \( g = g_c(\gamma, \beta) \). If the adaptation strength \( \beta \) has a value \( \beta \leq \beta_H(\gamma) \), then the imaginary part of the critical eigenvalue is equal to zero corresponding to a saddle-node bifurcation at \( g = g_c(\gamma, \beta) \). On the other hand, if \( \beta > \beta_H(\gamma) \), then the critical eigenvalue is a pair of complex-conjugate, purely imaginary eigenvalues, a signature of a Hopf bifurcation. Therefore, we introduce the curve \( \beta = \beta_H(\gamma) \), which separates the positive quadrant of the \( \gamma - \beta \) plane in two regions: one in which the systems becomes unstable at the critical value \( g_c(\gamma, \beta) \) via a saddle-node bifurcation, and another one in which the instability occurs via a Hopf bifurcation (Fig 2a). In the Hopf-bifurcation region,
the imaginary part of the critical eigenvalues can be computed analytically:

$$\text{Im}(\lambda_c^B) = \sqrt{-\gamma^2 + \sqrt{\beta^2(\beta + 2\gamma + 2)}} =: 2\pi f_m. \quad (10)$$

The parameter $f_m$ is the frequency of low-amplitude oscillations close to the bifurcation, if $N < \infty$. In the finite-$N$ case, we find numerically that these low-amplitude oscillations are stable. When $N \to \infty$, however, we find that chaotic dynamics onset right above the bifurcation (see section V A). The frequency $f_m$ is monotonic in $\beta$ but non-monotonic in $\gamma$ (Fig. 2b), indicating that a slower adaptation variable (smaller $\gamma$) does not necessarily correspond to slower oscillations. Finally, if $g > g_c(\gamma, \beta)$, the network exhibit self-sustained, irregular fluctuations (Fig. 1c,d) that will be characterized in the next sections.

IV. MEAN-FIELD THEORY

*General theory.* The dynamics of the $ND$-dimensional dynamical system in Eq. (2) for large $N$ is too high-dimensional to be studied at the microscopic level. In contrast, using dynamical mean-field theory [8], we can find properties of the network dynamics that are independent of the specific connectivity realization. Following [8], we approximate the network input to a representative unit $i$ with a Gaussian process $\eta$ and substitute the average over time, initial conditions and network realizations with the average over realizations of $\eta$. This approximation is valid in the large-$N$ limit, in which neurons become independent [45, 46]. Notice that we are interested in the regime in which the number of auxiliary variables per unit $D$ remains finite. In the mean-field description, the activity of each individual unit in the network follows a realization of the following system of $D$ stochastic differential equations, to which we refer to as mean-field equations (see appendix B for more details)

$$\dot{x}^\alpha(t) = \sum_{\beta=1}^{D} A^{\alpha\beta} x^\beta(t) + \delta^{\alpha1} (\eta(t) + I(t)) , \quad (11)$$

where $\eta(t)$ is a Gaussian process with mean zero. Its auto-correlation needs to be determined self-consistently by imposing

$$\langle \eta(t)\eta(s) \rangle = g^2 \langle \phi(x^1(t))\phi(x^1(s)) \rangle . \quad (12)$$

Thanks to the mean-field approximations, we reduced a $ND$-dimensional, deterministic, nonlinear system to a D-dimensional, stochastic, linear system. The effect of the nonlinear-
FIG. 1. Microscopic network dynamics with single-scale adaptation. a: Evolution of the rate $\phi(x_i)$ in time. Fifteen randomly chosen units are shown, out of $N = 1000$ units. Network parameters are $\gamma = 0.2$, $\beta = 0.5$ and $g = 0.96g_c(\gamma, \beta)$. For these parameters the network is in the Hopf-bifurcation regime, below the bifurcation, and hence it relaxes to the stable fixed point at zero. Inset: Eigenvalue spectrum of the Jacobian at the fixed point, in the complex plane. The dashed line indicates $\text{Re}(\lambda) = 0$. b: Initial transient to the stable fixed point, as in a, but in the saddle-node bifurcation regime, with $\gamma = 1$, $\beta = 0.1$ and $g = 0.96g_c(\gamma, \beta)$. c: Same as a, but above the Hopf bifurcation, with $\gamma = 0.2$, $\beta = 0.5$ and $g = 1.3g_c(\gamma, \beta)$. The fixed point is unstable and the network exhibits irregular, self-sustained fluctuations (resonant chaotic state). d: Same as b, but above the saddle-node bifurcation, with $\gamma = 1$, $\beta = 0.1$ and $g = 1.3g_c(\gamma, \beta)$. The system exhibits irregular activity similar to the chaotic system described in [8].
FIG. 2. **Stability of the fixed point and local properties.** a: Critical value of the coupling \( g_c \) (color code, right) for different adaptation parameters \( \gamma \) (horizontal axis) and \( \beta \) (vertical axis). The curve \( \beta_H(\gamma) \) (solid black line) separates the regions of the \( \gamma - \beta \) plane in which for increasing \( g \) we encounter a Hopf bifurcation (above \( \beta_H(\gamma) \)) or a saddle-node bifurcation (below \( \beta_H(\gamma) \)). Cross and filled circle: parameters used in Fig. 1. Left inset: dependence of \( g_c \) on \( \beta \) for fixed \( \gamma = 0.9 \). Top inset: dependence of \( g_c \) on \( \gamma \) for fixed \( \beta = \beta_H(\gamma = 0.9) \). Blue line: Hopf bifurcation; red line: saddle-node bifurcation. b: Resonance frequency \( f_0 \) for different adaptation parameters \( \gamma, \beta \). Notice that in the non-resonant region the resonance frequency is not defined. Left inset: square-root increase of \( f_m \) as a function of \( \beta \) for fixed \( \gamma = 0.9 \). Top inset: non-monotonic behavior of \( f_m \) as a function of \( \gamma \), for fixed \( \beta = \beta_H(\gamma = 0.9) \).

The mathematical structure allows us to write, in the frequency domain

\[
\tilde{x}^1(f) = \tilde{\chi}_0(f) \left( \tilde{\eta}(f) + \tilde{I}(f) \right) ,
\]

where \( \tilde{\chi}_0(f) \) is the linear response function (susceptibility) of the mean-field system (Eq. (11)), which is equal to the linear response function of an uncoupled single neuron in the microscopic description (Eq. (11)). For the linear dynamics given by Eq. (11), the linear response function \( \tilde{\chi}_0(f) \) is given by

\[
\tilde{\chi}_0(f) = \left( (2\pi if - A)^{-1} \right)^{11} ,
\]

where the upper indices 11 indicates the first element of the first row of the matrix inside the square brackets.
In what follows, we assume that the network is in the stationary regime. Since also the input is assumed to be stationary and zero-mean, the mean of all variables is equal to zero. The second-order statistics must be determined self-consistently. In the frequency domain, this requires a self-consistent determination of the power spectral density (or simply power spectrum) $S_x(f)$ of the activation variable $x^1$, defined by $\delta(f - f')S_x(f) = \langle (\tilde{x}^1)^*(f)\tilde{x}^1(f') \rangle$, where the average is over the Gaussian process $\eta$. Notice that $S_x(f)$ is the Fourier transform of the autocorrelation function of $x^1(t)$. Using the squared modulus of the linear response function $\tilde{\chi}_0(f) = |\tilde{\chi}_0(f)|^2$, the power spectrum can be expressed as

$$S_x(f) = \tilde{G}(f) (S_\eta(f) + S_I(f)),$$

where $S_\eta(f)$ and $S_I(f)$ denote the power spectral densities of $\eta(t)$ and $I(t)$, respectively. Importantly, $S_\eta(f)$ depends implicitly on $S_x(f)$ through the self-consistency condition

$$S_\eta(f) = g^2 S_{\phi(x^1)}(f).$$

(16)

The factor $\tilde{G}(f)$ can be expressed as a function of the entries of the matrix $A$ as

$$\tilde{G}(f) = \left[ (A^2 + (2\pi f)^2 I_D)^{-1} \right]^{11} = \frac{[\text{adj}(2\pi i f I_D - A)]^{11}}{\prod_{i=1}^{D} |2\pi i f - \lambda_A^i|^2},$$

(17)

where adj $(2\pi i f I_D - A)$ is the adjoint matrix of $(2\pi i f I_D - A)$. In appendix C, we show that knowing $\tilde{G}(f)$ is sufficient to compute the critical value of $g$. We find that $g_c$ is determined by the simple relation

$$g_c^2 \max_f \tilde{G}(f) = 1.$$  

(18)

Network with adaptation. For the network with adaptation defined by Eqs. (3, 4), the mean-field equations read

$$\dot{x}(t) = -x(t) - a(t) + \eta(t) + I(t)$$

(19)

$$\dot{a}(t) = -\gamma a(t) + \gamma \beta x(t),$$

(20)

with $\langle \eta(t) \rangle = 0$ and $\langle \eta(t + \tau)\eta(t) \rangle = g^2 \langle \phi(x(t + \tau))\phi(x(t)) \rangle$. The self-consistent equations (Eqs. (15, 16)), reduce to

$$S_x(f) = \tilde{G}(f) (g^2 S_{\phi(x)}(f) + S_I(f)).$$

(21)

The factor $\tilde{G}(f)$ can be calculated explicitly, yielding

$$\tilde{G}(f) = \frac{\gamma^2 + \omega^2}{\omega^4 + (1 + \gamma^2 - 2\beta \gamma)\omega^2 + \gamma^2(1 + \beta)^2},$$

(22)
with $\omega = 2\pi f$. In order to solve Eq. (21), we need to compute $S_{\phi(x)}(f)$ as a functional of $S_x(f)$, which is known to be a hard problem and not possible in general. However, as discussed in appendix D, the effect of the nonlinearity $\phi$ can be evaluated numerically. Since the transformation can be computed analytically in an integral form (appendix D), the computation of $S_{\phi(x)}(f)$ is fast.

In the next section we show how the qualitative features of the dynamics of the network in the fluctuating regime can be predicted by the properties of the single unit linear response function, as summarized in the factor $\tilde{G}(f)$

V. CHAOTIC REGIME

The traditional approach in the DMFT literature is to consider the time-domain version of Eq. (15). The time-domain mean-field system can be obtained by applying an inverse Fourier transform, which leads to a differential equation of order $2D$. Unfortunately, for $D > 1$ the resulting dynamics is no longer conservative, which precludes the determination of the initial conditions. We propose an alternative approach to find a self-consistent solution to Eq. (15) in the Fourier domain. This approach is based on an iterative map, the fixed point of which is the self-consistent solution. Similar iterative methods have been proposed previously both in the context of spiking [15, 47] and rate-based networks [48]. Here, we that such a method allows to robustly solve the mean-field theory and to qualitatively understand several features of the network dynamics.

In the frequency domain, the linear transform associated with $\chi_0(f)$ is simple, whereas the nonlinearity $\phi(x)$ is difficult to handle. Concretely, we need to express $S_{\phi(x)}$ as a functional of $S_x(f)$. While this calculation cannot be carried out analytically for a general nonlinearity, it is possible to compute it numerically or semi-analytically. The idea of our iterative method is to start with an arbitrary initial power spectral density $S_{\phi(x)}^{(0)}(f)$, which we choose to be constant (white noise). We then apply multiple iterations each consisting of a linear step followed by a nonlinear one. At each iteration, the linear step is simply a multiplication by $g^2|\tilde{\chi}^{11}_1(f)|^2$ and it allows us to compute $(S_x)^{(n+1)}(f)$. The nonlinear step afterwards transforms $(S_x)^{(n+1)}(f)$ into $S_{\phi(x)}^{(n+1)}(f)$. Depending on the nonlinearity $\phi$ that we consider, this step can introduce numerical errors of variable magnitude. Due to the importance of understanding the effect of the nonlinearity on the second order statistics,
the discussion on how to numerically or semi-analytically implement the nonlinear step is treated separately in appendix D.

Here we focus on the main results for the network with adaptation.

A. Resonant chaos in networks with adaptation

We apply the iterative method described above to solve the mean-field equations for the network with adaptation (Eq. (21)). We find, in the absence of external input \( I = 0 \), that if \( g < g_c(\gamma, \beta) \), the power spectrum converges to zero, \( S_x(f) \to 0 \), at all frequencies, hence the mean-field variable \( x \) is constantly equal to zero. This is consistent with the presence of a stable fixed point at zero and it indicates that, in the thermodynamic limit, this fixed point solution is the only possible one.

On the other hand, if \( g > g_c(\gamma, \beta) \), the mean-field network is characterized by a nonzero, continuous power spectral density. This is an indication that, at the microscopic level, the network is in a chaotic state. However, we stress that a more rigorous proof of chaos would require the computation of the maximum Lyapunov exponent of the network. In contrast to a network without adaptation \([8]\), we find that in the presence of adaptation the network can be in two qualitatively different chaotic regimes. For very weak and/or fast adaptation, the chaotic fluctuations are qualitatively the same as for the network without adaptation, i.e. the power spectrum is broad-band with maximum at \( f = 0 \) (Fig. 3a). We refer to this regime as to the non-resonant regime. On the other hand, for strong and/or slow adaptation, the mean-field network settles in a new regime, that we refer to as resonant regime, characterized by an autocorrelation that decays to zero via damped oscillations and, equivalently, by a power spectrum that exhibits a resonance band around a nonzero resonance frequency \( f_p \) (Fig. 3b). The decaying autocorrelation function and the continuous power spectral density are an indication that also this regime corresponds to microscopic chaos. This new dynamical state, that we refer to as resonant chaos, is qualitatively different from the one of the non-resonant regime and from the one of the non-adaptive network.

Strikingly, whether the network settles in the resonant or in the non-resonant regime can be predicted purely based on the single-unit adaptation properties. More precisely, if \( \beta < \beta_H(\gamma) \), the function \( \tilde{G}(f) \) is monotonically decreasing with the frequency \( f \), i.e. it exhibits a low-pass characteristic (Fig. 3a). This low-pass behavior of the single neuron is
FIG. 3. **Dynamical regimes in the mean-field description**

**a:** Resonant chaos. Power spectral density of the mean-field network (solid line) compared with microscopic simulations (light blue, dashed) for $\gamma = 0.25$, $\beta = 1$ and $g = 2g_c(\gamma, \beta)$. The dashed, dark blue line indicates the power spectral density of a network of independent neurons, driven by white noise with $\langle \eta(t)\eta(t') \rangle = g^2 \delta(t-t')$ and with the same adaptation parameters. Inset: Normalized mean-field autocorrelation $C_x(\tau)$ for the same parameters, plotted against the time lag in units of $\tau_x$.

**b:** Known chaotic regime. With $\gamma = 1$, $\beta = 0.1$ and $g = 2g_c(\gamma, \beta)$.

**c:** Maximum-power frequency $f_p$ of the recurrent network plotted against $\gamma$, for different $\beta$. Crosses depict results obtained from microscopic simulations, circles show the semi-analytical prediction based on the iterative method and dashed lines shows the theory based on the single neuron response function. For $\gamma = 0$ all curves start at $f_p = 0$.

**d:** Distributions $P(x)$ of the activation $x$ from microscopic simulation ($N = 2000$, solid lines) and theoretical prediction (dashed lines). The adaptation parameter were $\gamma = 0.25$ and $\beta = 1$.

reflected by a power spectrum of the network that is also dominated by low frequencies, albeit less broad. The network power spectrum corresponds exactly to the non-resonant regime discussed above.

In contrast, if $\beta > \beta_H(\gamma)$, the single neuron response amplitude $\tilde{G}(f)$ exhibits a maximum
at a nonzero frequency $f_0 = \frac{1}{2\pi} \sqrt{-\gamma^2 + \sqrt{\beta\gamma^2(\beta + 2\gamma + 2)}}$. Such a resonance peak is typical of a band-pass filter (Fig. 3b). The frequency $f_0$ is identical to $f_m$, which is derived from the imaginary part of the critical eigenvalue at the Hopf bifurcation (Eq. (10)). The single-neuron linear response characteristics are qualitatively preserved in the fluctuating activity of the recurrent network, which also exhibit a power spectral density dominated by a nonzero frequency $f_p$. Interestingly, we find that $f_p = f_0$, i.e. the resonance frequency is not affected by the introduction of recurrent connections (Fig. 3c). We notice that the non-resonant and resonant regimes are consistent with the fixed point stability analysis of the network in the microscopic description. Indeed, the resonant and non-resonant regimes match the regions in which we observe Hopf or saddle-node bifurcations, respectively.

Using simulations of the full microscopic network, we verify that the mean-field description is a good approximation of the system for large but finite $N$. In Fig. 3d we show that the probability density of the activation variable $x$ measured from the microscopic simulations matches the Gaussian distribution predicted by the mean-field theory, with relatively small finite-size effects close to the criticality. Moreover, the mean-field solution provides a good description of the system for a wide range of adaptation parameters $\gamma, \beta$ (Fig. 3c).

**FIG. 4. Effect of iterations on the power spectral density.** a: Evolution of the power spectral density $S_x$ over iterations, for the network with adaptation. The initial power spectral density is a constant. b: Evolution of the variance of $x$ and of the height of the peak of $S_x$, over iterations. The adaptation parameters for both panels are $\gamma = 0.25$, $\beta = 1$, $g = 2g_c(\gamma, \beta)$.
Iterative dynamics. We numerically study the evolution of the power spectrum during iteration towards the mean-field solution (Fig. 4). In the resonant regime, the iterative map leads first to a sharpening of the resonance band, and in the later iterations to the formation of harmonics (Fig. 4a). Notice that since we initialize the power spectrum \( S_{\phi(x)} \) to a constant, we have \( S_{x}^{(1)} \propto \tilde{G}(f) \), and thereafter the resonance frequency never changes. We verified that the iterative method is robust to different initializations by varying the peak position and amplitude of the initial power spectrum (Fig. 4b). To visualize the evolution of the power spectrum over iterations, we consider the variance of \( x \) (total area under \( S_{x}(f) \)) and the maximum amplitude of the spectrum itself. For all tested parameters, the iterative dynamics converge to the same solution, but the speed of convergence depends on the initialization (Fig. 4b).

B. Oscillation coherence and correlation time

FIG. 5. Correlation time and effect of recurrent connections. a: Correlation time (blue solid line) and Q-factor (dashed line) as a function of the connectivity strength. The weakest connectivity level plotted is \( g = 1.1g_{c}(\gamma, \beta) \). Adaptation parameters: \( \gamma = 0.1 \) and \( \beta = 1 \). The dash-dotted horizontal line indicates the Q-factor of a single unit with the same adaptation parameters, driven by with noise. b: Correlation time (blue) and Q-factor (black, dash-dotted line) as a function of the adaptation timescale \( \tau_{a} = \gamma^{-1}\tau_{x}[\tau_{x}] \). Both the recurrent network (solid line) and the single unit driven by white noise (dashed line) scale with \( \tau_{a}. \beta = 1 \) and \( g = 1.5g_{c}(\gamma, \beta) \).
While the resonance frequency in the resonant regime seems to depend solely on the single-neuron properties, the introduction of recurrent connections increases the coherence of the stochastic oscillations, i.e. to decrease the width of the resonance band. The narrower the resonance band, the more coherent the oscillatory behavior will be. To quantify the increase of the oscillation coherence, we measure the quality factor ($Q$-factor) of the stochastic oscillations, defined as

$$Q = \frac{f_p}{\Delta f},$$

(23)

where $\Delta f$ is the frequency width of the power spectrum $S_x(f)$ at the half-maximum. Intuitively, for a narrow-band oscillation, the quality factor quantifies the number of oscillation cycles during the characteristic decay time of the autocorrelation function. For a single neuron driven by white noise ($\langle \eta(t)\eta(t') \rangle = \delta(t-t')$), the single-neuron power spectrum of $x$ is proportional to $\tilde{G}(f)$. Compared to this reference shape, we find that the $Q$-factor increases in the recurrent network (Fig. 5a), corresponding to a sharper resonance peak in the power spectrum. When approaching the criticality from the chaotic phase, $g \to g_c(\gamma, \beta)^+$, the quality factor diverges (Fig. 5a) i.e. the dynamics approach regular oscillations.

While the $Q$-factor measures the decay time constant of the autocorrelation function relative to the mean oscillation period, it is also interesting to consider the absolute correlation time of the activity. A common measure of correlation time of a stochastic process is given by the normalized first moment (center of mass) of the absolute value of the autocorrelation function (e.g. [49]),

$$t_c = \frac{\int_{0}^{\infty} \tau |C_x(\tau)| d\tau}{\int_{0}^{\infty} |C_x(\tau)| d\tau},$$

(24)

Due to the increase of the $Q$-factor, the correlation time also diverges when $g \to g_c(\gamma, \beta)$ (Fig. 5a).

In the regime of slow adaptation, a single unit driven by white noise can have a larger correlation time than a recurrent network (Fig. 5b). This is due to the fact that in this regime the correlation time of the single unit driven by white noise is dominated by the long tail of the autocorrelation. The introduction of recurrent connections increases the oscillatory component, giving a larger “weight” to the short time lags and therefore decreasing $t_c$. Nevertheless, the correlation time scales with the timescale of adaptation $\tau_a$ for both the single unit driven by white noise and the recurrent network (Fig. 5b). Note that the $Q$-factor goes to zero for very large adaptation timescale ($\gamma \to 0$), so that the dominant contribution
to the correlation time in this regime is the non-oscillatory one.

C. Response to an external stimulus

FIG. 6. Response of the mean-field network to an oscillatory input. a: Effect of the external input on the power spectral density $S_x(f)$. In the example, $\gamma = 0.25$, $\beta = 1$, $g = 2g_c(\gamma, \beta)$, $f_I = 0.12$, while $A_I = 5 \cdot 10^{-1}$ (blue) and $A_I = 0$ (gray). Simulation (solid blue) and theory (dashed blue) are superimposed. b: Amplitude of the oscillatory ($A_{osc}$, blue solid line) and chaotic ($A_{chaos}$, black dashed line) components of the power spectral density, as a function of the ratio between input frequency $f_I$ and the resonance frequency predicted by the single unit properties $f_0$. The gray dash-dotted vertical line indicates $f_I = f_0$. Network parameters: $\gamma = 0.25$, $\beta = 1$, $g = 2g_c(\gamma, \beta)$ and $A_I = 2$.

In order to go beyond the study of the spontaneous activity of the network, we consider its response to an external drive. An interesting class of external drives are oscillatory signals, since we can study the locking of the network to the external signal. Similarly to previous approaches [50], we provide oscillatory input to each unit in the microscopic network, randomizing the phase

$$I_i(t) = A_I \cos(2\pi f_I t + \theta_i),$$

where $\theta_i \sim U(0, 2\pi)$. The corresponding power spectral density of the input is given by

$$S_I(f) = (A_I^2/4) \cdot (\delta(f - f_I) + \delta(f + f_I)).$$

Thanks to the phase randomization, the network
still reaches a stationary state and the mean $\langle x(t) \rangle$ remains at zero. The presence of the input affects the dynamics of the mean-field network, quantified by the power spectral density (Fig. 6a). A sharp peak at the driving frequency $f_I$ and multiples thereof are elicited by the external input, while the nearby frequencies are suppressed (notice the log-scale in Fig 6a). For $f_I > f_p$, as in the example, the relative peaks of the spectrum are slightly shifted toward larger values. The opposite happens if $f_I < f_p$. Notice that both this shift and the suppression of spontaneous activity are nonlinear effects due the recurrent dynamics. As an additional nonlinear effect, the network activity also exhibits harmonics of the external input.

The formation of a sharp peak together with the suppression of other modes are indications that at the microscopic level the network is driven towards a limit cycle while chaotic activity is suppressed. To quantify this effect we assume, similarly to [50], that the power spectral density $S_x(f)$ in the presence of the input can be separated into a chaotic component and an oscillatory component

$$S_x(f) = S_{\text{chaos}}(f) + \sum_{k=1}^{\infty} b_k \left( \delta(f - kf_I) + \delta(f + kf_I) \right),$$

(26)

where we included the multiples of the driving frequency in order to account for the harmonics. The total variance of $x$ can then be split in two contributions

$$\text{Var}(x) = \int_{-\infty}^{\infty} S_{\text{chaos}}(f) df + 2 \sum_{k=1}^{\infty} b_k =: A_{\text{chaos}} + A_{\text{osc}}.$$  

(27)

Both $A_{\text{chaos}}$ and $A_{\text{osc}}$ exhibit a non-monotonic dependence on $f_I$ (Fig. 6b). $A_{\text{osc}}$ depends smoothly on $f_I$, reaching its largest value around $f_0$. On the other hand, $A_{\text{chaos}}$ is close to zero for input frequencies that are close to $f_0$. While also the network without adaptation exhibits such a non-monotonic dependence [50], in our case this effect is more pronounced due to the non-monotonic power spectral density induced by adaptation.

VI. DISCUSSION

We studied how the dynamics of a random network of rate neurons is shaped by the properties of single neurons, in particular by the presence of history-dependent mechanisms such as adaptation. To this end, we generalized DMFT, a well-established theoretical tool [8], to the case of multi-dimensional rate units. This allowed us to reduce the high-dimensional,
deterministic network model to a low dimensional system of stochastic differential equations. Standard approaches to solve the mean-field equations were not fruitful in the multi-dimensional setting. However, the mean-field solution could be found efficiently in a semi-analytical way using an iterative approach. This approach has the advantage of highlighting how recurrent connections sharpen the response function of single neurons, i.e. how bands of preferred frequencies become narrower (see also appendix A).

We studied in detail the important case of neuronal adaptation, using a two-dimensional rate model. We showed that adaptation extends the stability region of a recurrent network of rate units because the transition from a stable fixed point to a fluctuating regime happens at $g = g_c \geq 1$, i.e. for higher coupling strength than for the network without adaptation. Crucially, above the criticality and for slow adaptation, the dynamics settles in a state of “resonant chaos” that, unlike the chaotic activity of networks of rate units without adaptation, is dominated by a nonzero resonance frequency. We observed that the resonance frequency can be computed from the single unit properties and it is therefore independent of the connectivity strength $g$. Consistent with this result, the eigenvalue spectrum of the Jacobian at the fixed point corresponds to a Hopf bifurcation in the microscopic system. On the other hand, the presence of recurrent connections increases the coherence of the oscillations and therefore influence the correlation time. The increase is maximal near the criticality and decreases with $g$, for $g > g_c(\gamma, \beta)$. Indeed, as it is typical of critical behavior, the correlation time in the chaotic phase diverges when approaching the criticality. In the presence of adaptation, this happens because the oscillations get more coherent and the system approaches a limit cycle. Furthermore, we showed that adaptation enhances the range of chaos suppression [50, 51] by periodic perturbations.

It is interesting to observe that for slow adaptation there seems to be two separate contributions to the correlation time of the network activity: an oscillatory component, related to the resonance frequency, and a long tail that scales with the adaptation timescale. For finite $\tau_a$, the correlation time diverges when $g \to g_c$ due to the oscillatory component. For $\tau_a \to \infty$, the correlation time also diverges, but this is due to the long tail, since the both the resonance frequency and the Q-factor go to zero for large $\tau_a$, yielding a finite and therefore sub-dominant contribution to the correlation time. Such multi-scale structure of the autocorrelation could be advantageous for network computations that require expressive dynamics over multiple timescales, as it is often the case in motor control. The long correlation
time could also provide an explanation for the evidences of a beneficial role of adaptation in tasks requiring memory over long time lags ([5, 42]). The relation between correlation time and computational properties, however, is more complex and requires further investigation.

Current mean-field theories for spiking neural networks [52] are self-consistent only with respect to mean activities (firing rates), whereas second-order statistics such as auto-correlation function (ACF) or power spectral density (PSD) of inputs and outputs are inconsistent [15]. A self-consistent calculation of the ACF or PSD via DMFT for networks of spiking neurons is known to be a hard theoretical problem, which remains unsolved. However, the DMFT for spiking networks has been solved numerically using an iterative Monte-Carlo procedure [15, 47, 53]. An analytical solution of the self-consistency problem would require an analytical expression for the map from the input ACF to the output ACF. However, even for simple integrate-and-fire models it is unknown how to compute the second-order statistics for colored noise inputs analytically (for an explicit series expression in the case of a perfect integrate-and-fire model, see [54]). In this paper, the rate-based modeling framework allowed us to put forward explicit expressions for the map of ACFs (or PSDs). For a general nonlinearity $\phi(x)$, this map takes the form of an infinite series (according to Eq. (A1) in appendix A). However, for polynomial nonlinearities the series simplifies to a finite sum, e.g. Eq. (D1) in appendix D which permits a closed-form analytical expression for the iterative map. Therefore, our study offers a unique method for the calculation of the ACF in biologically constrained random neural networks, and thus represents a promising step towards a self-consistent mean-field theory beyond first-order rate models [8, 9].

Our framework can readily be used to study the network dynamics in the presence of multiple adaptation variables. Indeed, spike-frequency adaptation (SFA) is known to have multiple timescales that are power-law distributed [12, 14]. In appendix A, we provide two additional examples of multi-dimensional rate models that are instantiations of multi-scale adaptation (Fig. 7). Interestingly, if no coupling is present between different adaptation variables, the linear response function exhibits only one band of preferred frequencies. On the other hand, in the presence of coupling, the linear response function can be multimodal and this can be reflected also in the recurrent dynamics (as shown in Fig. 7). Since our approach requires only the knowledge of the factor $\tilde{G}(f)$, it can be applied even to truly power-law units (i.e. infinite-dimensional), that cannot be described using a finite set of
differential equations. In this paper we considered the case in which the network input affects only one of the internal variables of the rate model. The theoretical framework can be generalized to the case in which each variable receives input from the network with an independent set of random connections, leading to a matrix equation, analogous to Eq. \cite{15}, for the matrix of power spectral densities.

We carried out our analysis for the case of balanced network, i.e. in which the mean of the input is zero. There are no conceptual obstacle in extending the analysis to include more recent developments. For example, the case of non-balanced input was considered in \cite{9, 10}, in which it was shown that a transition to chaos can be observed in many different network architectures and in some cases also for unbounded nonlinearities. We expect this observation to hold when considering multi-dimensional rate models. The introduction of white noise, e.g. to mimic spiking noise, is also an interesting extension that has been studied in the absence of adaptation \cite{9, 55}. It is interesting to notice that the introduction synaptic delay can result a bifurcation structure similar to the one induced by adaptation \cite{56}. As another example, \cite{7} extended the formalism to include the effect of low-rank perturbation of the random connectivity, showing that the resulting dynamics becomes effectively low dimensional and that it allows the network to perform complex tasks. It is an open question how the properties of the multi-dimensional rate model would shape this effective low-dimensional dynamics.

Appendix A: Qualitative study of the iterative map

For a qualitative understanding of the effect of the iterations on the power spectral density, we exploit the fact that \( x^1 \) is a Gaussian process, for which the following formula holds \cite{57}

\[
C_{\phi(x^1)}(\tau) = \sum_{n=0}^{\infty} \frac{1}{n!} \left( \langle \frac{d^n \phi}{d(x^1)^n} \rangle \right)^2 C_{x^1}^n(\tau),
\]

where the angular brackets indicate the mean over the statistics of \( x^1 \). Eq. \eqref{A1} gives the effect of a nonlinearity \( \phi \) on a the autocorrelation of a Gaussian process \( x^1 \). By truncating the series after the first term, we get

\[
C_{\phi(x^1)}(\tau) \simeq \left( \langle \phi'(x^1) \rangle \right)^2 C_{x^1}(\tau).
\]

21
Fourier transforming this equation we get an approximation of the power spectral density of $\phi(x^1)$

$$S_{\phi(x^1)}(f) \simeq s_1 \left( \int_{-\infty}^{\infty} S_x(f') df' \right) S_x(f), \quad (A3)$$

where we rewrote $((\phi'(x^1)))^2 =: s_1 \left( \int_{-\infty}^{\infty} S_x(f') df' \right)$ to highlight the fact that the coefficient that multiplies $S_x(f)$ depends on the area under the power spectral density, i.e. on the variance of $x^1$. We stress that retaining only the first term in Eq. (A1) is different than considering a linear approximation of $\phi$, since the dependence of the coefficient on the variance would not appear in that case.

Using this approximation, we can express the power spectral density at the $n^{th}$ iteration of the iterative method, as a function of the initial power spectral density $S_{\phi(x^1)}^{(0)}(f)$ from which we started to iterate. We obtain

$$(S_x)^{(n)}(f) = \left( \prod_{k=1}^{n-1} s_1^{(k)} \right) \left( g^2 \tilde{G}(f) \right)^n S_{\phi(x^1)}^{(0)}(f), \quad (A4)$$

where $s_1^{(k)} := s_1 \left( \int_{-\infty}^{\infty} (S_x)^{(n)}(f') df' \right)$. If we take $S_{\phi(x^1)}^{(0)}(f)$ to be constant and we define $a_n = \left( \prod_{k=1}^{n-1} s_1^{(k)} \right)$, we can rewrite the above expression as

$$(S_x)^{(n)}(f) = a_n \left( g^2 \tilde{G}(f) \right)^n. \quad (A5)$$

If $g > g_c$, there will be a range of frequencies for which $g^2 \tilde{G}(f) > 1$, which implies that its $n^{th}$ power diverges when $n$ grows. In a purely linear network, this phenomenon would lead to a blow-up of the power spectral density, in agreement with the fact that activity in a linear network is unbounded for $g > g_c$. If $\phi$ is a compressive nonlinearity however, the coefficient $a_n$ will tend to zero for growing $n$, counterbalancing the unbounded growth of $\left( g^2 \tilde{G}(f) \right)^n$. Notice that this constraint on the coefficient $a_n$ is necessary independently of the truncation of the series in Eq. (A1), since all the neglected terms are positive and would not provide a different mechanism for contrasting the growth of the first term. Based on Eq. (A5), we would predict that all the modes for which $\tilde{G}(f) > 1/g^2$ will get amplified over multiple iterations, while all the other modes will get suppressed. While this is a highly simplified description, the suppression and the amplification of modes is clearly visible when observing the evolution of the power spectrum over iterations (Fig. 4a) and when comparing the dynamics of the self-consistent solution (Fig. 7c,f) to the corresponding linear response function (Fig. 7b,e). When truncating the series after the first order however, the mean-field
network does not admit a self-consistent solution, for which we need to retain also higher order terms. The presence of those terms will be reflected, among others, in the interference among amplified modes.

As an example of higher-order term, consider the next term in the series in Eq. (A1), given by

$$\frac{1}{2} \left( \langle \phi''(x^1) \rangle \right)^2 (C_{x1}(\tau))^2 \xrightarrow{FT} \xrightarrow{FT} \frac{s_2}{2} \left( \int_{-\infty}^{\infty} S_x(f') df' \right) (S_x * S_x)(f)$$

where $s_2$ is defined analogously to $s_1$. In general, higher-order terms will contain convolutions of the power spectral density with itself, which are responsible for the creation of higher harmonics. To qualitatively understand this effect, consider the case in which $S_x(f)$ is a Dirac $\delta$-function with support in $f_0$. In this case, the two-fold convolution of $S_x(f)$ with itself is again equal to a Dirac $\delta$-function, but centered in $2f_0$. A similar argument can be given for resonant power spectral densities, which implies that a self-consistent solution should exhibit harmonics of the fundamental resonance frequency. Note that in this paper we considered odd functions, for which

**Appendix B: Mean-field theory derivation**

In this section, we extend the derivation of the dynamical mean-field theory (DMFT) to the case of the network of multi-dimensional rate units. Since there are no additional
FIG. 7. **Two examples of multi-dimensional rate models a-b-c:** Analysis of a three-dimensional rate model. Eigenvalue spectra (a) corresponding to the coupling values $g_1 = 1.28$, $g_2 = 1.4$ and $g_3 = 2$. The dashed line indicates the imaginary axis. In b we plot the linear response function of the single unit $\tilde{G}(f)$ (solid line), and the instability threshold corresponding to the three coupling values $g_1$, $g_2$ and $g_3$ (dashed lines). In c we plot the solution of the mean field theory obtained with the iterative method for the three values of $g$, $g_1 = 1.5$, $g_2 = 2$ and $g_3 = 3$. **d-e-f:** Same as a,b,c, but for a four-dimensional rate model.

complication with respect to the standard case, we report here only the main steps. For a
review of the path-integral approach to DMFT, see e.g. ([45][46]). The moment-generating functional corresponding to our differential equations is

$$Z[j, \tilde{j}](J) = \int D\mathbf{x} D\tilde{\mathbf{x}} \exp \left[ S_0[\mathbf{x}, \tilde{\mathbf{x}}] - (\tilde{x}^1)^T J \phi(x^1(t)) + j^T \mathbf{x} + \tilde{j}^T \tilde{\mathbf{x}} \right] ,$$  \hspace{1cm} (B1)

where

$$S_0[\mathbf{x}, \tilde{\mathbf{x}}] := \tilde{x}^T (I_D \partial_t - A) \mathbf{x}$$  \hspace{1cm} (B2)

and we introduced the notation $\tilde{x}^T \mathbf{x} = \sum_\alpha \sum_i \int \tilde{x}^\alpha_i(t)x^\alpha_i(t)dt$. The integral is over paths and bold symbols indicate vectors, over both the network space and the rate model space, so that $D\mathbf{x} := \prod_\alpha \prod_i Dx^\alpha_i$.

We are interested in properties that are independent of the particular realization of the coupling matrix $J$. In order to extract those properties, we average over the quenched disorder by defining the averaged generating function

$$\bar{Z}[j, \tilde{j}] := \int \prod_{ij} dJ_{ij} N \left( 0, \frac{g^2}{N}, J_{ij} \right) Z[j^x, \tilde{j}^x](J) .$$  \hspace{1cm} (B3)

The average over each $J_{ij}$ can be computed by noticing that the terms corresponding to different $J_{ij}$ factorize and the integral can be solved using the square-completion method. Since the details of this calculation are analogous to the one-dimensional case, we directly report the result

$$\bar{Z}[j^x, \tilde{j}^x] = \int D\mathbf{x} D\tilde{\mathbf{x}} \exp \left[ S_0[\mathbf{x}, \tilde{\mathbf{x}}] + j^T \mathbf{x} + \tilde{j}^T \tilde{\mathbf{x}} \right] \times$$

$$\times \exp \left[ \frac{1}{2} \int_{-\infty}^{\infty} \left( \sum_i \tilde{x}^1_i(t)\tilde{x}^1_i(t') \right) \left( \frac{g^2}{N} \sum_j \phi(x^1_j(t))\phi(x^1_j(t')) \right) \right] .$$  \hspace{1cm} (B4)

We now aim to decouple the interaction term in the last line by introducing the auxiliary field

$$Q_1(t, s) := \frac{g^2}{N} \sum_j \phi(x^1_j(t))\phi(x^1_j(s)) .$$  \hspace{1cm} (B5)

We rewrite the averaged generating functional as a field theory for two auxiliary fields $Q_1, Q_2$. The result is, following the same steps for the one-dimensional case,

$$\bar{Z}[j, \tilde{j}] = \int DQ_1 DQ_2 \exp \left( -\frac{N}{g^2} Q^T_1 Q_2 + N \ln Z[Q_1, Q_2] + j^T Q_1 + \tilde{j}^T Q_2 \right)$$

$$Z[Q_1, Q_2] := \int D\mathbf{x} D\tilde{\mathbf{x}} \exp \left( S_0[\mathbf{x}, \tilde{\mathbf{x}}] + \frac{1}{2}(\tilde{x}^1)^T Q_1 \tilde{x}^1 + \phi(x^1)^T Q_2 \phi(x^1) \right) .$$  \hspace{1cm} (B6)
where we extended our notation to \( Q_1^T Q_2 := \int \int Q_1(s, t)Q_2(s, t) ds dt \). The crucial observation to make is that essentially all factors associated to different units factorized yielding the factor \( N \). For this reason, the integration is now only over all rate model indices but over only one unit index. The remainder is the problem of one unit, characterized by \( D \) variables, interacting with two external fields \( Q_1, Q_2 \).

The final step is to perform a saddle-point approximation, i.e. replace \( Q_1, Q_2 \) by their values that make the action stationary. After this step, the averaged generating functional reduces to

\[
\bar{Z}^* \propto \int \mathcal{D}x \mathcal{D}\tilde{x} \exp \left( S_0[x, \tilde{x}] + \frac{g^2}{2} (\tilde{x}^1)^T C_{\phi(x^1)} \tilde{x}^1 \right),
\]

(B7)

This is the statistical field theory corresponding to \( D \) linearly interacting variables, with \( x^1 \) that receives a Gaussian noise whose autocorrelation is given by \( C_{\phi(x^1)} \). Writing the corresponding differential equations results in our mean-field description (Eq. (11)).

**Appendix C: Fixed point stability in the mean-field network**

Here we consider the full matrix of linear response functions (see below), to conclude that the only quantity that matters for the stability at the fixed point is \( \tilde{G}(f) \).

Starting from the microscopic network equations (Eq. (2)), we derive a set of differential equations, that we write in matrix form

\[
(I_D \partial_\tau - A) \chi_{ik}(\tau) = \sum_{j=1}^{N} J_{ij} \Delta_1 \chi_{jk}(\tau) + \delta_{ik} I_D \delta(\tau),
\]

(C1)

where \( \Delta_1 = \delta^{i_1} \delta^{j_1} \) is a matrix whose only nonzero element is \([\Delta_1]^{11} = 1\). \( \chi_{ik}(\tau) \) is a \( D \) by \( D \) matrix, whose component are defined as \( \chi_{ik}^{\alpha\beta}(\tau) = \frac{\delta \chi_{ik}(\tau)}{\delta h^\beta_k(0)} \), where \( h^\beta_k \) is a small perturbation given to the variable \( x^\beta_k \) at time \( \tau = 0 \). Notice that in deriving Eq. (C1), we have assumed stationarity and that \( \phi'(0) = 1 \). We now Fourier transform Eq. (C1) and get

\[
(2\pi i f I_D - A) \tilde{\chi}_{ik}(f) = \sum_{j=1}^{N} J_{ij} \Delta_1 \tilde{\chi}_{jk}(f) + \delta_{ik} I_D
\]

(C2)

Inverting the matrix \((2\pi i f I_D - A)\) and recognizing the linear response function of the single unit \( \tilde{\chi}_0(f) \), we obtain

\[
\tilde{\chi}_{ik}(f) = \sum_{j=1}^{N} J_{ij} \tilde{\chi}_0(f) \Delta_1 \tilde{\chi}_{jk}(f) + \delta_{ik} \tilde{\chi}_0(f),
\]

(C3)
where \( \tilde{\chi}_0(f) \) is a \( D \times D \) matrix whose elements are \( \tilde{\chi}_0^{\alpha\beta}(f) \), defined in section [V].

Since in the mean-field approximation the mean of the linear response function is zero, we look for the second moments ([9]). We multiply every element of the matrix equation (Eq. (C3)) by its complex conjugate and average over the quenched disorder. We obtain

\[
|\tilde{\chi}(f)|^2 = g^2 |\tilde{\chi}_0(f)\Delta_1 \tilde{\chi}(f)|^2 + \tilde{G}(f),
\]

where the absolute value is intended element-wise. Due to the structure of the matrix \( \Delta_1 \), we have that

\[
|\tilde{\chi}_0(f)\Delta_1 \tilde{\chi}(f)|^2 = \tilde{G}(f) \Delta_1 |\tilde{\chi}(f)|^2,
\]

as it can be verified simply by using the definition of \( \Delta_1 \). Finally, we can solve for \( |\tilde{\chi}(f)|^2 \)

\[
|\tilde{\chi}(f)|^2 = \left( I_D - g^2 \tilde{G}(f) \Delta_1 \right)^{-1} \tilde{G}(f).
\]

Since the only nonzero eigenvalue of the matrix \( \tilde{G}(f) \Delta_1 \) is \( |\tilde{\chi}^{11}_0(f)|^2 \), the stability condition for the fixed point is given by

\[
g^2 \max_f \tilde{G}(f) < 1.
\]

**Appendix D: Effect of nonlinearities on second-order statistics**

In this section, we provide some additional details on how to compute the effect of nonlinearities on the second order statistics (autocorrelation or power spectral density) of a Gaussian process. We consider three cases of interest: polynomials, piecewise linear functions and arbitrary nonlinear functions. To simplify our notation, we drop the superscript of \( \phi \) and consider a generic Gaussian process \( x \).

The effect of polynomial nonlinearities can be expressed in closed form in time domain. This can be seen by considering again the infinite series expression (Eq. [A1]), valid for stationary Gaussian processes \( x \)

\[
C_{\phi(x)}(\tau) = \sum_{n=0}^{\infty} \left( \left\langle \frac{d^n \phi}{dx^n} \right\rangle \right)^2 C_x(\tau),
\]

where the angular brackets indicate the average over the statistics of \( x \). In the case in which \( \phi \) is a polynomial of degree \( p \), only the terms in the sum up to \( p \) are nonzero. As an example, we can compute the effect of a cubic approximation of the hyperbolic tangent, i.e. \( \phi(x) \simeq \phi_3(x) := x - \frac{x^3}{3} \)

\[
C_{\phi_3(x)}(\tau) = (1 + C_x^2(0) - 2C_x(0)) C_x(\tau) + \frac{2}{3} C_x^3(\tau).
\]
As expected, the effect of the nonlinearity depends on $C_x(0)$ i.e. on the variance of $x$ itself. Notice that the coefficient of the first term is compressive (i.e. smaller than one) only if $C_x(0)$ is smaller than one itself. This type of behavior is expected since $\phi_3$ is unbounded.

Another interesting case are piecewise linear nonlinearities. In this case, we use Price’s theorem twice to get

$$\frac{\partial^2 C_{\phi(x)}(t)}{\partial (C_x(t))^2} = C_{\phi''(x)}(t).$$

(D3)

For a piecewise linear $\phi$, the second derivative $\phi''$ is a sum of Dirac’s delta functions with variable coefficients. More precisely, we consider

$$\phi_{PL}(x) = \Theta(x_1 - x)c_0x + \sum_{p=1}^{P-1} \Theta(x - x_p)\Theta(x_{p+1} - x)c_px_p + \Theta(x - x_P)c Px,$$

(D4)

where $x_p$ are the points in which the first derivative is discontinuous, $c_p$ are some arbitrary coefficients and $\Theta(\cdot)$ is the Heaviside function. The second derivative of $\phi_{PL}$ is given by

$$\phi''_{PL}(x) = \sum_{p=1}^{P}(c_p - c_{p-1})\delta(x - x_p).$$

(D5)

The delta functions allow us to compute the correlation function $C_{\phi''_{PL}}(t)$ explicitly

$$C_{\phi''_{PL}}(t) = \sum_{p,p'=1}^{P} \frac{(c_p - c_{p-1})(c_{p'} - c_{p'-1})}{2\pi C_x(0)\sqrt{1 - \rho^2(t)}} \times$$

$$\times \exp \left( -\frac{x_p^2 + x_{p'}^2 - 2\rho(t)x_px_{p'}}{2C_x(0)(1 - \rho^2(t))} \right),$$

(D6)

where we defined $\rho(t) := \frac{C_x(t)}{C_x(0)}$. Inserting Eq. (D6) in Eq. (D3) and integrating twice with respect to $C_x(t)$ we get

$$C_{\phi_{PL}(x)}(t) = f_\phi(0; C_x(0)) + f_{\phi'}(0; C_x(0))C_x(t)$$

$$+ \sum_{p,p'=1}^{P} \int_0^{C_x(t)} \int_0^{\sigma'} \frac{(c_p - c_{p-1})(c_{p'} - c_{p'-1})}{2\pi C_x(0)\sqrt{1 - \frac{\sigma^2}{C_x(0)}}} \times$$

$$\times \exp \left( -\frac{x_p^2 + x_{p'}^2 - 2\frac{\sigma}{C_x(0)}x_px_{p'}}{2C_x(0)(1 - \frac{\sigma^2}{C_x(0)})} \right) d\sigma d\sigma'.$$

(D7)

In the case in which $\phi$ is an odd function, the term $f_\phi(0; C_x(0))$ is equal to zero. For the specific case of the piecewise linear approximation of the hyperbolic tangent considered in
this paper, i.e.

\[
\phi_{PL}(x) = \begin{cases} 
-1 & \text{for } x < -1 \\
 x & \text{for } -1 < x < 1 \\
 1 & \text{for } x > 1 
\end{cases},
\]  

(D8)

the expression in Eq. (D7) reduces to

\[
C_{\phi_{PL}(x)}(t) = \text{Erf}^2 \left( \frac{1}{\sqrt{2C_x(0)}} \right) C_x(t) + \frac{2}{\pi C_x(0)} \int_0^{C_x(t)} \int_0^{\sigma'} \frac{1}{\sqrt{1 - \frac{\sigma'^2}{C_x^2(0)}}} \times \\
\times \exp \left( -\frac{1}{C_x(0) (1 - \frac{\sigma^2}{C_x^2(0)})} \right) \sinh \left( \frac{\sigma}{C_x^2(0) (1 - \frac{\sigma^2}{C_x^2(0)})} \right) d\sigma d\sigma' . \]  

(D9)

For the piecewise linear function, an alternative approach based on the infinite series in Eq. (A1) and on Hermite polynomials was proposed by [58].

For an arbitrary nonlinear function, we can use two methods. The first method is a semi-analytical approach that relies on the integral form of the autocorrelation of the rate \( C_{\phi(x)}(\tau) \) as a functional of the autocorrelation \( C_x(\tau) \) of \( x \) (15)

\[
C_{\phi(x)}(\tau) = \int \int \phi \left( \sqrt{C_x(0) - \frac{C_x^2(\tau)}{C_x(0)}} x + \frac{C_x(\tau)}{\sqrt{C_x(0)}} z \right) \phi \left( \sqrt{C_x(0)} z \right) DxDz , \]  

(D10)

where \( Dx = e^{-x^2/2} dx \). Notice that a slightly different version of this formula was already proposed in [8]. Therefore, to obtain the effect of \( \phi \) on the power spectral density, one should 1) inverse Fourier transform \( S_x(f) \) to get \( C_x(\tau) \) 2) apply Eq. (D10), by computing the two integrals numerically 3) Fourier transform \( C_{\phi(x)}(\tau) \) to get \( S_{\phi(x)}(f) \). Practically, this procedure requires the application of the fast Fourier transform algorithm and the numerical evaluation of two integrals.

The second method is fully numerical and it can be useful in cases in which the integrals in the first method are expensive to evaluate numerically. This method consists in approximating the power spectral density \( S_{\phi(x)} \) via Monte Carlo sampling. More precisely, we sample multiple realizations in frequency domain of the Gaussian process with zero mean and power spectral density \( S_x(f) \). We then transform each sample to time domain and apply the nonlinearity \( \phi(x) \) to each sample \( x(t) \) individually. Finally, we transform back to Fourier
domain and get $S_{\phi(x)}$ by averaging. Despite being computationally more expensive than the closed form expressions, this sampling method provides a solution of the mean-field theory for an arbitrary nonlinearity and it is computationally much cheaper than running the full microscopic simulation.

[1] W. Maass, T. Natschläger, and H. Markram, Neural Computation, 2531 (2002).
[2] H. Jaeger and H. Haas, Science 304, 78 (2004).
[3] D. Sussillo and L. F. Abbott, Neuron 63, 544 (2009).
[4] R. Laje and D. Buonomano, Nat. Neurosci. 16, 925 (2013).
[5] W. Nicola and C. Clopath, Nat. Commun. 8, 2208 (2017).
[6] B. DePasquale, C. J. Cueva, K. Rajan, G. S. Escola, and L. F. Abbott, PLOS ONE 13, 1 (2018).
[7] F. Mastrogiuseppe and S. Ostojic, Neuron 99, 609 (2018).
[8] H. Sompolinsky, A. Crisanti, and H. J. Sommers, Physical Review Letters 61, 259 (1988).
[9] J. Kadmon and H. Sompolinsky, Phys. Rev. X 5, 041030 (2015).
[10] F. Mastrogiuseppe and S. Ostojic, PLoS Comput. Biol. 13, 1 (2017).
[11] J. Benda and A. V. M. Herz, Neural Computation 15, 2523 (2003).
[12] B. Lundstrom, M. Higgs, W. Spain, and A. Fairhall, Nature Neuroscience 11, 1335 (2008).
[13] G. La Camera, A. Rauch, D. Thurbon, H.-R. Luscher, W. Senn, and S. Fusi, J. Neurophysiology 96, 3448 (2006).
[14] C. Pozzorini, R. Naud, S. Mensi, and W. Gerstner, Nat. Neurosci. 16, 942 (2013).
[15] S. Wieland, D. Bernardi, T. Schwalger, and B. Lindner, Phys. Rev. E 92, 040901 (2015).
[16] A. van Meegen and B. Lindner, Phys. Rev. Lett. (2018).
[17] T. Schwalger, M. Deger, and W. Gerstner, PLoS Comput. Biol. 13, e1005507 (2017).
[18] Z. F. Mainen and T. J. Sejnowski, Science 268, 1503 (1995).
[19] W. Bair and C. Koch, Neural Computation 8, 1185 (1996).
[20] F. Devalle, A. Roxin, and E. Montbrió, PLOS Comput. Biol. 13, 1 (2017).
[21] B. W. Knight, J. Gen. Physiology 59, 734 (1972).
[22] P. Konig, A. K. Engel, and W. Singer, Trends Neurosci 19, 130 (1996).
[23] W. Gerstner, Neural Computation 12, 43 (2000).
[24] R. Brette and E. Guignon, Neural Computation 12, 279 (2003).
[25] M. Mattia and P. Del Giudice, Phys. Rev. E xx, xx (2002).
[26] E. S. Schaffer, S. Ostojic, and L. F. Abbott, PLoS Comput. Biol. 9, e1003301 (2013).
[27] E. Montbrió, D. Pazó, and A. Roxin, Phys. Rev. X 5, 021028 (2015).
[28] M. J. Berry and M. Meister, J. Neurosci. 18, 2200 (1998).
[29] R. Naud and W. Gerstner, PLOS Comput. Biol. 8, e1002711 (2012).
[30] M. Deger, T. Schwalger, R. Naud, and W. Gerstner, Phys. Rev. E 90, 062704 (2014).
[31] T. Schwalger, M. Deger, and W. Gerstner, PLoS Comput. Biol. 13, e1005507 (2017).
[32] N. Fourcaud and N. Brunel, Neural Computation 14, 2057 (2002).
[33] T. Schwalger and L. Schimansky-Geier, Phys. Rev. E 77, 031914 (2008).
[34] M. Richardson, N. Brunel, and V. Hakim, J. Neurophysiology 89, 2538 (2003).
[35] S. Ostojic, G. Szapiro, E. Schwartz, B. Barbour, N. Brunel, and V. Hakim, J. Neurosci. 35, 7056 (2015).
[36] J. Doose, G. Doron, M. Brecht, and B. Lindner, Journal of Neuroscience 36, 11120 (2016).
[37] A. L. Fairhall, G. Lewen, W. Bialek, and R. van Steveninck, Nature 412, 787 (2001).
[38] G. Deco and E. T. Rolls, J. Neurophysiol. 94, 295 (2005).
[39] H. Setareh, M. Deger, and W. Gerstner, PLoS Comput. Biol. 14, 1 (2018).
[40] A. Shpiro, R. Moreno-Bote, N. Rubin, and J. Rinzel, J. Comput. Neurosci. 27, 37 (2009).
[41] P. Theodoni, G. Kovács, M. W. Greenlee, and G. Deco, J. Neurosci. 31, 234 (2011).
[42] G. Bellec, D. Salaj, A. Subramoney, R. Legenstein, and W. Maass, arXiv (2018).
[43] J. R. Silvester, The Mathematical Gazette 84, 460467 (2000).
[44] V. Girko, Theory of Probability & Its Applications 29, 694 (1985).
[45] J. Schücker, S. Goedeke, D. Dahmen, and M. Helias, arXiv (2016).
[46] A. Crisanti and H. Sompolinsky, arXiv e-prints (2018).
[47] B. Dummer, S. Wieland, and B. Lindner, Front. Comp. Neurosci. 8, 104 (2014).
[48] M. Stern, H. Sompolinsky, and L. F. Abbott, Phys. Rev. E 90, 062710 (2014).
[49] K. Lindenber, B. J. West, and J. Masoliver, “First passage time problems for non-markovian processes,” in Noise in Nonlinear Dynamical Systems, Vol. 1, edited by F. Moss and P. V. E. McClintock (Cambridge University Press, 1989) pp. 110–160.
[50] K. Rajan, L. F. Abbott, and H. Sompolinsky, Phys. Rev. E 82, 011903 (2010).
[51] T. Schwalger, A. Dzhanoev, and A. Loskutov, Chaos 16, 023109 (2006).
[52] N. Brunel, J. Comput. Neurosci. 8, 183 (2000).
[53] A. Lerchner, C. Ursta, J. Hertz, M. Ahmadi, P. Ruffiot, and S. Enemark, Neural Comput. 18, 634 (2006).
[54] T. Schwalger, F. Droste, and B. Lindner, J. Comput. Neurosci. 39, 29 (2015).
[55] J. Schuecker, S. Goedeke, and M. Helias, Phys. Rev. X 8, 041029 (2018).
[56] C. Bimbard, E. Ledoux, and S. Ostojic, Phys. Rev. E 94, 062207 (2016).
[57] R. L. Stratonovich, Topics in the Theory of Random Noise, Vol. 1 (Gordon and Breach, 1967).
[58] A. Kruscha and B. Lindner, Phys. Rev. E 94, 022422 (2016).