

Transitions between asynchronous and synchronous states: a theory of correlations in small neural circuits

Diego Fasoli^{1,2}  · Anna Cattani^{1,3} · Stefano Panzeri¹

Received: 11 April 2017 / Revised: 4 September 2017 / Accepted: 10 October 2017 / Published online: 10 November 2017
© The Author(s) 2017. This article is an open access publication

Abstract The study of correlations in neural circuits of different size, from the small size of cortical microcolumns to the large-scale organization of distributed networks studied with functional imaging, is a topic of central importance to systems neuroscience. However, a theory that explains how the parameters of mesoscopic networks composed of a few tens of neurons affect the underlying correlation structure is still missing. Here we consider a theory that can be applied to networks of arbitrary size with multiple populations of homogeneous fully-connected neurons, and we

focus its analysis to a case of two populations of small size. We combine the analysis of local bifurcations of the dynamics of these networks with the analytical calculation of their cross-correlations. We study the correlation structure in different regimes, showing that a variation of the external stimuli causes the network to switch from asynchronous states, characterized by weak correlation and low variability, to synchronous states characterized by strong correlations and wide temporal fluctuations. We show that asynchronous states are generated by strong stimuli, while synchronous states occur through critical slowing down when the stimulus moves the network close to a local bifurcation. In particular, strongly positive correlations occur at the saddle-node and Andronov-Hopf bifurcations of the network, while strongly negative correlations occur when the network undergoes a spontaneous symmetry-breaking at the branching-point bifurcations. These results show how the correlation structure of firing-rate network models is strongly modulated by the external stimuli, even keeping the anatomical connections fixed. These results also suggest an effective mechanism through which biological networks may dynamically modulate the encoding and integration of sensory information.

Action Editor: Bard Ermentrout

This research was supported by the Autonomous Province of Trento, Call “Grandi Progetti 2012”, project “Characterizing and improving brain mechanisms of attention—ATTEND”, and by the Future and Emerging Technologies (FET) program within the Seventh Framework Program for Research of the European Commission, under FET-Open grant FP7-600954 (VISUALISE). The funders had no role in study design, data collection and analysis, decision to publish, interpretation of results, or preparation of the manuscript.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10827-017-0667-3>) contains supplementary material, which is available to authorized users.

✉ Diego Fasoli
diego.fasoli@iit.it

¹ Laboratory of Neural Computation, Center for Neuroscience and Cognitive Systems @UniTn, Istituto Italiano di Tecnologia, 38068 Rovereto, Italy

² Center for Brain and Cognition, Computational Neuroscience Group, Universitat Pompeu Fabra, 08002 Barcelona, Spain

³ Department of Biomedical and Clinical Sciences “L. Sacco”, University of Milan, Milan, Italy

Keywords Stochastic neural networks · Graded firing-rate model · Finite-size effects · Bifurcation analysis · Synchronous and asynchronous states · Propagation of chaos · Critical slowing down

1 Introduction

The study of correlations (or in general of statistical dependencies) among neurons is a topic of central importance to systems neuroscience. Statistical dependencies, which

are also termed functional connectivity (Friston 2011), have been studied extensively at multiple scales, from interactions among neurons within a network, to interactions among macroscopic circuits (Sporns 2006; Bressloff 2009; Renart et al. 2010; Pernice et al. 2011; Trousdale et al. 2012; Buice and Chow 2013). There are several reasons why studying statistical interactions among neurons is important. Firstly, the structure of both second-order and higher-order statistical dependencies among neurons, and how these dependencies are modulated by either external stimuli or internal factors such as neuromodulation or attention, is key to understanding the information encoding capabilities of neural populations (Abbott and Dayan 1999; Pola et al. 2003; Pillow et al. 2008; Cohen and Maunsell 2009; Moreno-Bote et al. 2014). Secondly, measuring and understanding statistical dependencies is crucial to making inferences about how different neurons or areas exchange and integrate information (Singer 1993; Tononi et al. 1994; David et al. 2004; Rogers et al. 2007; Friston 2011). Thirdly, statistical dependencies among the activities of different neurons are useful to infer the underlying network structure (Friston et al. 2013; Gilson et al. 2016).

From the theoretical point of view, it is of particular interest to understand whether and how these statistical interactions can be modulated dynamically by changes in parameters, such as the strength of the external input to the network or by other network characteristics. This knowledge can help to better interpret patterns of correlations observed experimentally, and to understand how the brain can implement a dynamic qualitative change in information processing or transmission despite the relatively slow time scales of changes of anatomical connectivity (Womelsdorf et al. 2007; Akam and Kullmann 2010; Battaglia et al. 2012; Besserve et al. 2015). In their pioneering work (Ginzburg and Sompolinsky 1994), Ginzburg and Sompolinsky developed a first theory of correlations among neurons in a neural network model with binary firing rates. They proved that, in the limit of large network size, the population averaged activities can switch from asynchronous states (a regime in which correlations among neurons become weak and vanish as $1/N$, where N is the size of the network) to synchronous states characterized by strong correlations.

The emergence of asynchronous states of uncorrelated neurons has been reported also by many other studies of large neural systems, and is key to the formulation of their mean-field approximation (Samuelides and Cessac 2007; Touboul et al. 2012; Baladron et al. 2012; Baladron Pezoa et al. 2012). On the other hand, synchronous regimes have been reported in large networks typically when the network undergoes critical slowing down, a phenomenon that happens when a system becomes increasingly sensitive to external perturbations (Kéfi et al. 2013). In this situation, the state variables undergo large and asymmetric

fluctuations, with a strong increase in the cross- and auto-correlation functions (Scheffer et al. 2009; Kuehn 2013). Critical slowing down occurs at some (but not all) of the bifurcation points of the network's dynamics, where small parameter variations cause profound qualitative changes in its dynamics. For example, in Ginzburg and Sompolinsky (1994) the authors showed the formation of critical slowing down in large networks of binary neurons when they approached a saddle-node or an Andronov-Hopf bifurcation, which corresponded to catastrophic transitions and the emergence of oscillatory activity, respectively.

The most established theories of correlation were developed in the large-network size limit (Ginzburg and Sompolinsky 1994; Bressloff 2009; Renart et al. 2010; Buice and Chow 2013) and can be in practice applied to macroscopic networks composed of at least few thousands of neurons. However, and somewhat counter-intuitively, the cross-correlation structure of small neural networks containing only a few tens of neurons can be much more difficult to study mathematically than that of large networks. This is mainly due to the impossibility to apply the powerful methods of statistical analysis, such as the law of large numbers and the central limit theorem, to small neural circuits. This fact prevents these theories to be able to describe networks of neurons at the mesoscopic and microscopic circuit level encompassing, for example, networks of few tens of cells in invertebrates (Williams and Herrup 1988). This small network level has been investigated in recent years both theoretically (Ingber 1992; Freeman 2000a, b; Wright et al. 2003; Bohland et al. 2009) and experimentally (Buzsáki et al. 2012; Einevoll et al. 2013), as it is a useful scale to link neural activity to brain function (Buzsáki and Draguhn 2004).

In this article we fill this gap by developing a theory of correlations in small neural circuits composed of homogeneous populations of fully-connected neurons. Unlike the analytical results introduced in Ginzburg and Sompolinsky (1994) and Renart et al. (2010), which were based on neurons with binary firing rates, we will consider the more complex and more biologically realistic case of graded firing rates. While Refs. (Ginzburg and Sompolinsky 1994; Renart et al. 2010) considered both fully-connected circuits with homogeneous synaptic weights and systems with random connectivity, for simplicity in this work we focus only on the case of networks with fully-connected topology (a possible extension to the case of random networks is discussed in Section 4.5). Moreover, unlike previous work, we will explicitly compute how cross-correlations relate to the whole bifurcation structure of the network. To do so, we study the dynamics of neural circuits of arbitrary size taking advantage of a mathematical approach that does not rely on statistical averaging (Fasoli et al. 2016). These networks of arbitrary size, when studied in the deterministic case, revealed a surprisingly rich set of local bifurcations

of dynamics that could be studied analytically (Fasoli et al. 2016). By extending this previous work to include stochastic perturbations to the dynamics, we computed analytically the correlations among all neurons in the network, we studied their dependence on the network parameters and related them to the bifurcations of the dynamics.

We found that such finite-size networks displayed both asynchronous and synchronous regimes, with important qualitative and quantitative differences with respect to the large network size limit of Ref. (Ginzburg and Sompolinsky 1994). We proved that asynchronous states may occur also in small networks for strongly depolarizing or strongly hyperpolarizing stimuli, extending and generalizing previous analytical calculations (Fasoli et al. 2015) valid only when the network has a regular topology. We then proved the emergence in finite-size networks of critical slowing down at the saddle-node and Andronov-Hopf bifurcations, that was previously observed in the large-size limit of binary networks (Ginzburg and Sompolinsky 1994). Moreover, we proved that at the branching points or pitchfork bifurcations, which are characterized by a spontaneous symmetry-breaking of neural activity with spontaneous formation of heterogeneous activity within homogeneous populations, the inhibitory neurons undergo critical slowing down characterized by strong anti-correlation. This phenomenon was not found in large-size networks (Ginzburg and Sompolinsky 1994).

Interestingly, our formalism predicts the formation of synchronous and asynchronous states in networks composed of an arbitrary number of neural populations without calculating explicitly their cross-correlation structure (see the Online Resource 1 and Online Resource 2). However, our approach also allows explicit calculations in networks composed of a few neural populations. For exemplary purposes, in the main text we focus on the case of two neural populations, and we extensively validate the closed-form expression of the cross-correlations through numerical simulations.

2 Materials and methods

2.1 The stochastic firing-rate network model

Here we describe the stochastic firing-rate finite-size neural network model that we use in this article. This model is based on a number of assumptions and simplifications (described below) that represent the best compromise we could find between biological plausibility and mathematical tractability.

A cortical column can be thought of as a network of neural masses distributed vertically across layers, and therefore it is composed of several populations of excitatory and inhibitory neurons (see for example (Binzegger et al.

2004)). Our theory can be used to study such cortical architectures, but the complexity of the resulting formulas increases considerably with the number of populations. Thus for exemplary purposes, in the main text we focus on the case of a network made of two fully homogeneous neural populations, one excitatory (E) and one inhibitory (I), which is commonly considered a good approximation of a single neural mass (Grimbert 2008). The structure of this network is schematized in the left-hand panel of Fig. 1. In the Online Resource 1 we generalize these results to an arbitrary number \mathfrak{P} of populations (an example of network structure for $\mathfrak{P} = 8$ is shown in Fig. 1, right panel).

The populations contain an arbitrary finite number of neurons which are interconnected through synaptic connections with arbitrarily strong weights. In order to make the network analytically tractable, we assume (as it is often made when considering local cortical circuits Grimbert 2008; Deco et al. 2008) that the axonal delays are negligible. Moreover, whenever a neural population α projects synapses to a population β (with $\alpha, \beta = E, I$), we assume that each neuron in population α sends connections to each neuron in population β (avoiding self-connections in the case $\alpha = \beta$). Furthermore, as often done in theoretical neuronal network studies (Brunel and Hakim 1999; Steyn-Ross et al. 2004; Touboul et al. 2012), we describe random fluctuations in the network by means of a white noise component in the external input to the network.

In more detail, we consider a graded rate model to describe the dynamics of single neurons by means of the following system of stochastic differential equations:

$$\frac{dV_i(t)}{dt} = -\frac{1}{\tau_i} V_i(t) + \frac{1}{M_i} \sum_{j=0}^{N-1} J_{ij} \mathcal{A}_j(V_j(t)) + I_i(t) + \sigma_i^{\mathcal{B}} \frac{d\mathcal{B}_i(t)}{dt}, \quad i = 0, \dots, N - 1, \tag{1}$$

where N is the number of neurons in the network, $V_i(t)$ is the membrane potential of the i th neuron at the time instant t , and τ_i is its membrane time constant. The normalization factor M_i represents the number of incoming connections to the i th neuron, while J_{ij} is the weight of the synaptic connection from the j th (presynaptic) neuron to the i th (postsynaptic) neuron. $\mathcal{A}_j(\cdot)$ is an algebraic activation function which converts the membrane potential V into the corresponding firing rate $v = \mathcal{A}(V)$, according to the formula:

$$\mathcal{A}_j(V) = \frac{v_j^{\max}}{2} \left[1 + \frac{\frac{\Lambda_j}{2} (V - V_j^T)}{\sqrt{1 + \frac{\Lambda_j^2}{4} (V - V_j^T)^2}} \right]. \tag{2}$$

Here v_j^{\max} is the maximum firing rate of the neuron, V_j^T is the threshold of the activation function, and Λ_j is its slope parameter. The latter represents the “speed” with

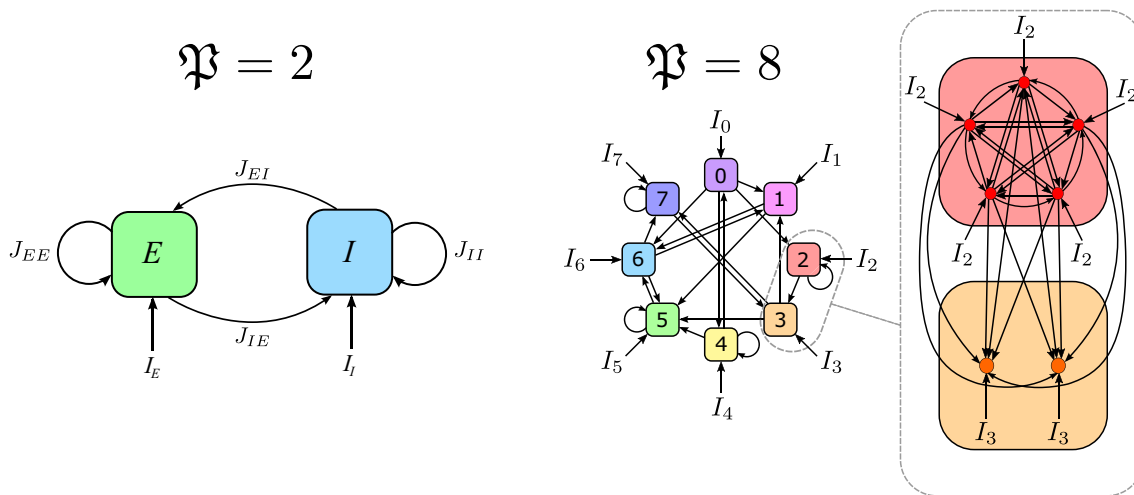


Fig. 1 Examples of the neural network architectures considered here. The left panel shows the case of $\mathfrak{P} = 2$ populations that we study in the main text. For simplicity we identify the two populations by the letters E, I , in order to distinguish between populations of excitatory and inhibitory neurons. In the case $\mathfrak{P} = 2$ we consider non-zero synaptic weights, therefore the network has a fully-connected architecture

which the neuron switches between low rates ($v_j \approx 0$) and high rates ($v_j \approx v_j^{\max}$). Moreover, in Eq. (1) $I_i(t)$ is a deterministic external input (i.e. the stimulus) to the i th neuron, while $\sigma_i^{\mathcal{B}} \frac{d\mathcal{B}_i(t)}{dt}$ is a white noise input with normal distribution and standard deviation $\sigma_i^{\mathcal{B}}$. In order to apply linear response theory, we will assume that $\sigma^{\mathcal{B}}$ is small enough to neglect second-order corrections to the perturbative expansion of the cross-correlations. This assumption will hold if the standard deviation of the fluctuations in the membrane potential is smaller than the minimum radius of curvature of the activation function of Eq. (2). The latter is a function of the parameters of the activation function: $r(\mu) = \left| \frac{(1 + \mathcal{A}'(\mu)^2)^{3/2}}{\mathcal{A}''(\mu)} \right|$, where \mathcal{A}' and \mathcal{A}'' are the first and second-order derivatives of \mathcal{A} with respect to V , while μ is the stationary membrane potential in absence of noise. The functions $\mathcal{B}_i(t)$ are arbitrarily correlated Brownian motions, which represent the source of stochasticity of the model. The model defined in Eq. (1) can be considered a stochastic perturbation to the firing-rate finite-size network model that we previously analyzed in Fasoli et al. (2016).

We define N_E (respectively N_I) to be the size of the excitatory (respectively inhibitory) population, with $N_E + N_I = N$, and we rearrange the neurons so that the structural connectivity of the network can be written as follows:

$$J = \begin{bmatrix} \mathfrak{J}_{EE} & \mathfrak{J}_{EI} \\ \mathfrak{J}_{IE} & \mathfrak{J}_{II} \end{bmatrix}, \quad \mathfrak{J}_{\alpha\beta} = \begin{cases} J_{\alpha\alpha} (\mathbb{1}_{N_\alpha} - \text{Id}_{N_\alpha}), & \text{for } \alpha = \beta \\ J_{\alpha\beta} \mathbb{1}_{N_\alpha, N_\beta}, & \text{for } \alpha \neq \beta, \end{cases} \quad (3)$$

($M_E = M_I = N - 1$). The right panel shows an example of a network composed of $\mathfrak{P} = 8$ neural populations. In particular, we highlighted the structure of the synaptic connections within and between two neural populations. The extension of the theory to the case of an arbitrary number of neural populations has been developed in the [Online Resource 1](#)

for $\alpha, \beta = E, I$. The real numbers $J_{\alpha\beta}$ are free parameters that describe the strength of the synaptic connections from the population β to the population α . We have $J_{EE}, J_{IE} \geq 0$ and $J_{EI}, J_{II} \leq 0$. Moreover, $\mathbb{1}_{N_\alpha, N_\beta}$ is the $N_\alpha \times N_\beta$ all-ones matrix (here we use the simplified notation $\mathbb{1}_{N_\alpha} \stackrel{\text{def}}{=} \mathbb{1}_{N_\alpha, N_\alpha}$), while Id_{N_α} is the $N_\alpha \times N_\alpha$ identity matrix. From our rearrangement of the neurons, we also obtain that the external input currents are organized into two vectors $\mathbf{I}_{E,I}$ such that:

$$\mathbf{I}_\alpha(t) = I_\alpha(t) \mathbf{1}_{N_\alpha},$$

where $\mathbf{1}_{N_\alpha} \stackrel{\text{def}}{=} \mathbb{1}_{N_\alpha, 1}$ is the $N_\alpha \times 1$ all-ones vector. The same subdivision between populations is performed for the parameters $M, \tau, v^{\max}, \Lambda, V^T$.

We also assume that the covariance structure of the white noise $\sigma_i^{\mathcal{B}} \frac{d\mathcal{B}_i(t)}{dt}$ is given by the matrix:

$$\Sigma^{\mathcal{B}} = \begin{bmatrix} \Sigma_{EE}^{\mathcal{B}} & \Sigma_{EI}^{\mathcal{B}} \\ \Sigma_{IE}^{\mathcal{B}} & \Sigma_{II}^{\mathcal{B}} \end{bmatrix}, \quad \Sigma_{\alpha\beta}^{\mathcal{B}} = \begin{cases} (\sigma_\alpha^{\mathcal{B}})^2 [\text{Id}_{N_\alpha} + C_{\alpha\alpha}^{\mathcal{B}} (\mathbb{1}_{N_\alpha} - \text{Id}_{N_\alpha})], & \text{for } \alpha = \beta \\ \sigma_\alpha^{\mathcal{B}} \sigma_\beta^{\mathcal{B}} C_{\alpha\beta}^{\mathcal{B}} \mathbb{1}_{N_\alpha, N_\beta}, & \text{for } \alpha \neq \beta, \end{cases} \quad (4)$$

where $C_{\alpha\beta}^{\mathcal{B}}$ are arbitrary parameters that quantify the cross-correlation between the white noise sources. We observe that $C_{EI}^{\mathcal{B}} = C_{IE}^{\mathcal{B}}$ since $\Sigma^{\mathcal{B}}$ must be symmetric in order to be a true covariance matrix, and that $\Sigma^{\mathcal{B}}$ determines the covariance structure of the white noise since $\text{Cov} \left(\sigma_i^{\mathcal{B}} \frac{d\mathcal{B}_i(t)}{dt}, \sigma_j^{\mathcal{B}} \frac{d\mathcal{B}_j(s)}{ds} \right) = [\Sigma^{\mathcal{B}}]_{ij} \delta(t - s)$. Equation (4) represents the most general covariance matrix of the noise under our assumption of fully homogeneous

neural populations. Since noise correlations can be interpreted as the amount of shared inputs between nearby neurons (Renart et al. 2010), which is not uniform in the brain, we perform our study for all the possible values of noise correlation, ranging from zero (independent stimuli) to one (identical stimuli to all the neurons).

Since we study the case of two neural populations, we can take advantage of the detailed bifurcation analysis performed in Fasoli et al. (2016) (see also Fig. 2), which we

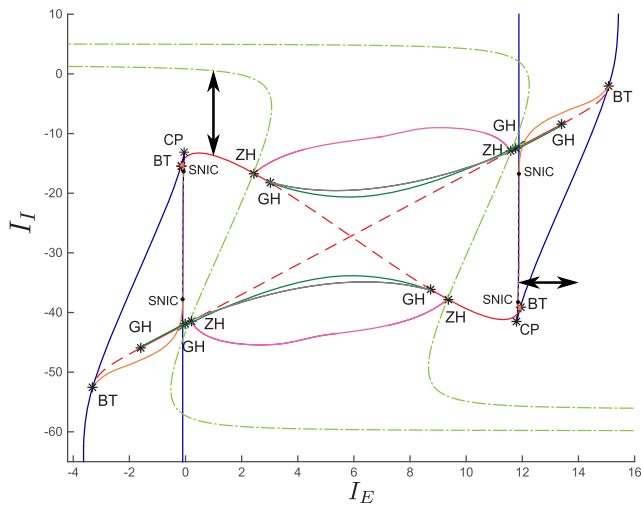


Fig. 2 Codimension two bifurcation diagram in the $I_E - I_I$ plane for the two-population network. This diagram was obtained in Fasoli et al. (2016) for the values of the parameters reported in Table 1. The blue curves represent the saddle-node bifurcations (LP for short in Figs. 3, 5) on the primary branch of stationary solutions of Eq. (1), with cusp bifurcations (CP). The red curves correspond to the Andronov-Hopf bifurcations (H for short in Figs. 4, 5) on the primary branch, which in turn are divided into supercritical (plain) and subcritical (dashed) portions. The supercritical/subcritical portions are bounded by a generalized Hopf bifurcation (GH), and Bogdanov-Takens bifurcations (BT). The latter are the contact points among saddle-node, Andronov-Hopf and homoclinic bifurcation curves on the primary branch (hyperbolic-saddle/saddle-node homoclinic bifurcations are represented by plain/dashed orange curves). Saddle-node on invariant circle bifurcations (SNIC) correspond to the contact points between the saddle-node and the homoclinic curves. GH generates limit point of cycles curves, represented by dark green lines, that collapse into the homoclinic curves. The gray lines represent the torus bifurcations, while the light green dot-dashed curves correspond to the branching-point bifurcations (BP for short in Figs. 3, 4 and 5). The purple curves represent the Andronov-Hopf bifurcations that originate from the secondary branches, which meet the branching-point curves and the other Andronov-Hopf curves at the zero-Hopf bifurcations (ZH). The double-headed black arrows represent the ranges in which we varied the stimuli $I_{E,I}$ in order to study the behavior of the cross-correlation. In more detail, on the horizontal arrow the network switches from an asynchronous state to critical slowing down near a saddle-node bifurcation (see also Fig. 3). Moreover, on the vertical arrow the network switches from positively correlated activity at the Andronov-Hopf bifurcation curve, to anti-correlated activity in the inhibitory population at the branching-point curve (see also Fig. 4). Adapted from Fasoli et al. (2016) with permission of the authors

will use to determine where the cross-correlation undergoes the most interesting variations.

2.2 Cross-correlation

To quantify statistical dependencies among neurons, we use Pearson cross-correlation, the simplest and most used measure of functional connectivity (David et al. 2004):

$$\text{Corr}(V_i(t), V_j(t)) \stackrel{\text{def}}{=} \frac{\text{Cov}(V_i(t), V_j(t))}{\sqrt{\text{Var}(V_i(t)) \text{Var}(V_j(t))}}, \quad (5)$$

where $\text{Var}(V_i(t)) = \text{Cov}(V_i(t), V_i(t))$. By applying linear response theory around a stable fixed point of the dynamics, we obtain the following analytical expression of the covariance structure of the rate model (1) at the first perturbative order in $\sigma_i^{\mathcal{B}}$:

$$\text{Cov}(V_i(t), V_j(t)) = \sum_{k,l=0}^{N-1} [\Sigma^{\mathcal{B}}]_{kl} \int_0^t \Phi_{ik}(t-s) \Phi_{jl}(t-s) ds. \quad (6)$$

In Eq. (6), the matrix $\Sigma^{\mathcal{B}}$ is given by Eq. (4), $\Phi(t) = e^{\mathcal{J}t}$ is the fundamental matrix of the system at time t , and \mathcal{J} is its Jacobian matrix (which depends on J). Note that this linearized equation is equivalent to that derived in Riskin and Frank (1996) (see Eq. (3.45) therein) for a set of linearly coupled Ornstein-Uhlenbeck processes.

When applied to our connectivity matrix (see Eq. (3)), Eqs. (5) and (6) provide a very cumbersome expression of the cross-correlation. Thus, for simplicity, in this article we consider only the infinite-time limit $t \rightarrow +\infty$ (i.e. the stationary state of the probability distribution of the membrane potentials), even if correlations may be calculated at any finite t , if desired.

It is also important to note that under our weak-noise assumption, as proven in Fasoli et al. (2015), the correlations between membrane potentials are equivalent to the correlations between the firing rates of the neurons (which are the quantities normally measured and computed by neuroscientists), as summarized by the following equation:

$$\text{Corr}(v_i(t), v_j(t)) \approx \text{Corr}(V_i(t), V_j(t)). \quad (7)$$

Thus in our paper we will quantify and plot without loss of generality correlations between membrane potentials, but this measure will reflect also the correlations between the firing rates.

2.3 Numerical simulations

To validate our analytical approach, we compared it with numerical evaluations of the correlations in the same

network, and we expressed all the results in simulation units (see Section 3). The numerical results were obtained by integrating the neural equations (1) with the Euler-Maruyama scheme, using the parameters reported in Table 1. We used an integration time step of $\Delta t = 10^{-3}$ (in simulation units), and the equations were integrated with a Monte Carlo method over 5,000 repetitions of the network dynamics in the temporal interval $t = [0, 30]$. Cross-correlations reach a stationary solution after a time period of the order of $\frac{1}{\min_{i=0, \dots, N-1} |\Re(\lambda_i)|}$, where $\Re(\lambda_i)$ is the real part of the i th eigenvalues of \mathcal{J} . Whenever the network is close to a local bifurcation, the real part of (at least) one of its eigenvalues goes to zero, therefore the duration of the transient regime is determined by that eigenvalue. For this reason, the temporal evolution of the cross-correlations slows down. We assumed that at $t = 30$ the transient regime of the correlation has already passed (so that the correlation has already converged to its equilibrium solution), an assumption confirmed a posteriori by the good agreement between the analytical and numerical results.

To conclude, we observe that the numerical integration schemes generally display a loss of stability at the bifurcation points of the network. For this reason, we stabilized the Euler-Maruyama scheme by choosing a small noise amplitude, $\sigma_E^{\mathcal{B}} = \sigma_I^{\mathcal{B}} = 10^{-4}$.

3 Results

In this section we explicitly calculate the cross-correlation structure of the firing-rate network model introduced in Section 2.1, and we study how it depends on the strength

Table 1 Values of the parameters of the two-population network

Population Sizes and Memb. Time Consts.	Synaptic Weights
$N_E = 8$	$J_{EE} = 10$
$N_I = 2$	$J_{EI} = -70$
$\tau_E = \tau_I = 1$	$J_{IE} = 70$
	$J_{II} = -34$
Activation Functions	Brownian Motions
$v_E^{\max} = v_I^{\max} = 1$	$\sigma_E^{\mathcal{B}} = \sigma_I^{\mathcal{B}} = 10^{-4}$
$\Lambda_E = \Lambda_I = 2$	$C_{EE}^{\mathcal{B}} = C_{II}^{\mathcal{B}} = C_{EI}^{\mathcal{B}} = 0$
$V_E^T = V_I^T = 2$	

We used these parameters to generate all the figures in the article, apart from Figs. 5 and 6 where the cross-correlation was evaluated for different values of $C_{\alpha\beta}^{\mathcal{B}}$. The ratio $N_E/N_I = 4$ matches the proportion between excitatory and inhibitory neurons in real cortical circuits (see Markram et al. 2004). Our theory can be applied to networks of arbitrary size, but in the article we analyze only the case of small networks ($N = 10$ in this example), see text

of the external stimulus. For simplicity, throughout the main text we consider only the case of two neural populations (for a Python implementation, see the Online Resource 3), but we report the theory for an arbitrary number of populations in the Online Resource 1 and Online Resource 2.

The usefulness and novelty of our approach consists in a convenient decomposition of the spectrum of the Jacobian matrix into “intra-population” eigenvalues, which are generated by the fully-connected architecture within each neural population, and “inter-population” eigenvalues, which depend on the synaptic connections among the populations. According to Eq. (6), the cross-correlations depend on the fundamental matrix of the network, $\Phi(t) = e^{\mathcal{J}t}$. We define $\mu_{E,I}$ to be the stationary membrane potentials in the two populations in absence of noise, namely the zeroth-order approximation to the mean membrane potentials of the stochastic network in the stationary regime. In the Online Resource 1 (see Eq. (S28)) we calculated $\Phi(t)$ in terms of the intra-population eigenvalues $\lambda_{E,I}$ and the inter-population eigenvalues $\lambda_{0,1}^{\mathcal{R}}$ of the Jacobian matrix \mathcal{J} , evaluated at the stationary solutions:

$$\begin{aligned} \lambda_E &= -\left[\frac{1}{\tau_E} + \frac{J_{EE}}{M_E} \mathcal{A}'_E(\mu_E) \right], \\ \lambda_I &= -\left[\frac{1}{\tau_I} + \frac{J_{II}}{M_I} \mathcal{A}'_I(\mu_I) \right], \\ \lambda_{0,1}^{\mathcal{R}} &= \frac{\mathcal{Y} + \mathcal{Z} \pm \sqrt{(\mathcal{Y} - \mathcal{Z})^2 + 4\mathcal{X}}}{2}, \end{aligned} \tag{8}$$

where:

$$\begin{aligned} \mathcal{X} &= \frac{N_E N_I}{M_E M_I} J_{EI} J_{IE} \mathcal{A}'_E(\mu_E) \mathcal{A}'_I(\mu_I), \\ \mathcal{Y} &= -\frac{1}{\tau_E} + \frac{N_E - 1}{M_E} J_{EE} \mathcal{A}'_E(\mu_E), \\ \mathcal{Z} &= -\frac{1}{\tau_I} + \frac{N_I - 1}{M_I} J_{II} \mathcal{A}'_I(\mu_I), \end{aligned} \tag{9}$$

and in terms of the functions:

$$\begin{aligned} K_x &\stackrel{\text{def}}{=} \frac{M_E \left(\lambda_x^{\mathcal{R}} + \frac{1}{\tau_E} \right) - (N_E - 1) J_{EE} \mathcal{A}'_E(\mu_E)}{N_I J_{EI} \mathcal{A}'_I(\mu_I)} \\ &= \frac{N_E J_{IE} \mathcal{A}'_E(\mu_E)}{M_I \left(\lambda_x^{\mathcal{R}} + \frac{1}{\tau_I} \right) - (N_I - 1) J_{II} \mathcal{A}'_I(\mu_I)}, \quad x=0, 1. \end{aligned} \tag{10}$$

If the parameters of the network are such that $\lambda_{E,I}$ and $\lambda_{0,1}^{\mathcal{R}}$ have negative real part, by applying Eqs. (6) and (S28) we obtain the following expression of the covariance matrix of the membrane potentials:

$$\Sigma^V \stackrel{\text{def}}{=} \lim_{t \rightarrow +\infty} [\text{Cov}(V_i(t), V_j(t))]_{v_i, j} = \begin{bmatrix} \Sigma_{EE}^V & \Sigma_{EI}^V \\ [\Sigma_{EI}^V]^T & \Sigma_{II}^V \end{bmatrix}, \tag{11}$$

where T denotes the transpose of a matrix, while the blocks $\Sigma_{\alpha\beta}^V$ are given by the following formulas:

$$\begin{aligned} \Sigma_{\alpha\alpha}^V &= (\sigma_\alpha^V)^2 \text{Id}_{N_\alpha} + \varsigma_{\alpha\alpha}^V (\mathbb{I}_{N_\alpha} - \text{Id}_{N_\alpha}), \\ \Sigma_{EI}^V &= \varsigma_{EI}^V \mathbb{I}_{N_E, N_I}, \\ (\sigma_E^V)^2 &= (\sigma_E^{\mathcal{B}})^2 \left[\gamma_{EE}^{EE} \Delta_E - \left(1 - \frac{1}{N_E}\right) \Theta_E \right] \\ &\quad + (\sigma_I^{\mathcal{B}})^2 \gamma_{EE}^{II} \Delta_I + 2\sigma_E^{\mathcal{B}} \sigma_I^{\mathcal{B}} \gamma_{EE}^{EI} C_{EI}^{\mathcal{B}}, \\ (\sigma_I^V)^2 &= (\sigma_I^{\mathcal{B}})^2 \left[\gamma_{II}^{II} \Delta_I - \left(1 - \frac{1}{N_I}\right) \Theta_I \right] \\ &\quad + (\sigma_E^{\mathcal{B}})^2 \gamma_{II}^{EE} \Delta_E + 2\sigma_E^{\mathcal{B}} \sigma_I^{\mathcal{B}} \gamma_{II}^{EI} C_{EI}^{\mathcal{B}}, \\ \varsigma_{EE}^V &= (\sigma_E^{\mathcal{B}})^2 \left(\gamma_{EE}^{EE} \Delta_E + \frac{\Theta_E}{N_E} \right) \\ &\quad + (\sigma_I^{\mathcal{B}})^2 \gamma_{EE}^{II} \Delta_I + 2\sigma_E^{\mathcal{B}} \sigma_I^{\mathcal{B}} \gamma_{EE}^{EI} C_{EI}^{\mathcal{B}}, \\ \varsigma_{II}^V &= (\sigma_I^{\mathcal{B}})^2 \left(\gamma_{II}^{II} \Delta_I + \frac{\Theta_I}{N_I} \right) \\ &\quad + (\sigma_E^{\mathcal{B}})^2 \gamma_{II}^{EE} \Delta_E + 2\sigma_E^{\mathcal{B}} \sigma_I^{\mathcal{B}} \gamma_{II}^{EI} C_{EI}^{\mathcal{B}}, \\ \varsigma_{EI}^V &= (\sigma_E^{\mathcal{B}})^2 \gamma_{EI}^{EE} \Delta_E + (\sigma_I^{\mathcal{B}})^2 \gamma_{EI}^{II} \Delta_I \\ &\quad + \sigma_E^{\mathcal{B}} \sigma_I^{\mathcal{B}} \gamma_{EI}^{EI} C_{EI}^{\mathcal{B}}, \\ \Delta_\alpha &= \frac{1}{N_\alpha} + C_{\alpha\alpha}^{\mathcal{B}} \left(1 - \frac{1}{N_\alpha}\right), \\ \Theta_\alpha &= \frac{1}{2\lambda_\alpha} \left(1 - C_{\alpha\alpha}^{\mathcal{B}}\right). \end{aligned} \tag{12}$$

The functions γ are defined as below:

$$\begin{aligned} \gamma_{EE}^{EE} &= \frac{1}{(K_1 - K_0)^2} \left[\frac{2K_0K_1}{\lambda_0^{\mathcal{R}} + \lambda_1^{\mathcal{R}}} - \frac{1}{2} \left(\frac{K_0^2}{\lambda_1^{\mathcal{R}}} + \frac{K_1^2}{\lambda_0^{\mathcal{R}}} \right) \right], \\ \gamma_{EE}^{II} &= \frac{1}{(K_1 - K_0)^2} \left[\frac{2}{\lambda_0^{\mathcal{R}} + \lambda_1^{\mathcal{R}}} - \frac{1}{2} \left(\frac{1}{\lambda_0^{\mathcal{R}}} + \frac{1}{\lambda_1^{\mathcal{R}}} \right) \right], \\ \gamma_{II}^{EE} &= \frac{K_0^2 K_1^2}{(K_1 - K_0)^2} \left[\frac{2}{\lambda_0^{\mathcal{R}} + \lambda_1^{\mathcal{R}}} - \frac{1}{2} \left(\frac{1}{\lambda_0^{\mathcal{R}}} + \frac{1}{\lambda_1^{\mathcal{R}}} \right) \right], \\ \gamma_{II}^{II} &= \frac{1}{(K_1 - K_0)^2} \left[\frac{2K_0K_1}{\lambda_0^{\mathcal{R}} + \lambda_1^{\mathcal{R}}} - \frac{1}{2} \left(\frac{K_0^2}{\lambda_0^{\mathcal{R}}} + \frac{K_1^2}{\lambda_1^{\mathcal{R}}} \right) \right], \\ \gamma_{EE}^{EI} &= \frac{1}{(K_1 - K_0)^2} \left[\frac{1}{2} \left(\frac{K_0}{\lambda_1^{\mathcal{R}}} + \frac{K_1}{\lambda_0^{\mathcal{R}}} \right) - \frac{K_0 + K_1}{\lambda_0^{\mathcal{R}} + \lambda_1^{\mathcal{R}}} \right], \end{aligned}$$

$$\begin{aligned} \gamma_{II}^{EI} &= \frac{K_0K_1}{(K_1 - K_0)^2} \left[\frac{1}{2} \left(\frac{K_0}{\lambda_0^{\mathcal{R}}} + \frac{K_1}{\lambda_1^{\mathcal{R}}} \right) - \frac{K_0 + K_1}{\lambda_0^{\mathcal{R}} + \lambda_1^{\mathcal{R}}} \right], \\ \gamma_{EI}^{EE} &= \frac{K_0K_1}{(K_1 - K_0)^2} \left[\frac{K_0 + K_1}{\lambda_0^{\mathcal{R}} + \lambda_1^{\mathcal{R}}} - \frac{1}{2} \left(\frac{K_0}{\lambda_1^{\mathcal{R}}} + \frac{K_1}{\lambda_0^{\mathcal{R}}} \right) \right], \\ \gamma_{EI}^{II} &= \frac{1}{(K_1 - K_0)^2} \left[\frac{K_0 + K_1}{\lambda_0^{\mathcal{R}} + \lambda_1^{\mathcal{R}}} - \frac{1}{2} \left(\frac{K_0}{\lambda_0^{\mathcal{R}}} + \frac{K_1}{\lambda_1^{\mathcal{R}}} \right) \right], \\ \gamma_{EI}^{EI} &= \frac{1}{(K_1 - K_0)^2} \left[K_0K_1 \left(\frac{1}{\lambda_1^{\mathcal{R}}} + \frac{1}{\lambda_0^{\mathcal{R}}} - \frac{2}{\lambda_0^{\mathcal{R}} + \lambda_1^{\mathcal{R}}} \right) \right. \\ &\quad \left. - \frac{K_0^2 + K_1^2}{\lambda_0^{\mathcal{R}} + \lambda_1^{\mathcal{R}}} \right]. \end{aligned} \tag{13}$$

In Eq. (12), σ_α^V is the standard deviation of the membrane potentials in the population α , while $\varsigma_{\alpha\alpha}^V$ is the covariance between any pair of potentials in the same population α , and ς_{EI}^V is the covariance between any pair of potentials in two different populations.

From the above equations, the correlation matrix is obtained simply by normalizing the covariance matrix (11) as follows (see Eq. (5)). In particular, we call $C_{\alpha\beta}^V$:

$$C_{\alpha\beta}^V \stackrel{\text{def}}{=} \frac{\varsigma_{\alpha\beta}^V}{\sigma_\alpha^V \sigma_\beta^V}$$

the entries of the correlation matrix that represent the correlation between any pair of membrane potentials in populations α, β .

In the following sections we will study how the above obtained analytical expression of the cross-correlation varies when changing the network parameters. This will be used to show that, depending on the values of the parameters, the network can switch from asynchronous to synchronous states. These regimes have radically contrasting properties, which will be discussed in detail in Sections 3.1 and 3.2.

3.1 Asynchronous states

We will first examine the existence of network states characterized by small cross-neuron correlations. Networks are said to be in an *asynchronous regime* when they show uncorrelated activity (Ecker et al. 2010; Renart et al. 2010; Tetzlaff et al. 2012; Grytskyy et al. 2013). In mathematics, a regime characterized by statistically independent (though interacting) units is called *local chaos* (Boltzmann 1872; Samuelides and Cessac 2007; Touboul et al. 2012; Baladron et al. 2012; Fasoli et al. 2015). Note that however in our model the asynchronous state and local chaos are equivalent at the first perturbative order, since neurons are jointly normally distributed under our weak-noise assumption.

The characterizing feature of an asynchronous state is the weak cross-correlations between the membrane potentials. Moreover, in our model asynchrony generally occurs with small fluctuations of the potentials, as we prove below. The most known and straightforward way to generate an asynchronous state with negligible correlations across neurons is using a network of large or infinite size (Ginzburg and Sompolinsky 1994; Samuelides and Cessac 2007; Touboul et al. 2012; Baladron et al. 2012; Baladron Pezoa et al. 2012). Indeed, if $C_{EE}^{\mathcal{B}} = C_{II}^{\mathcal{B}} = C_{EI}^{\mathcal{B}} = 0$, from Eqs. (12) and (13) it follows that $\varsigma_{\alpha\beta}^V \rightarrow 0$ and $(\sigma_{\alpha}^V)^2 \rightarrow \left(-\frac{1}{2\lambda_{\alpha}}\right) \left(\sigma_{\alpha}^{\mathcal{B}}\right)^2 \approx \frac{1}{2\tau_{\alpha}} \left(\sigma_{\alpha}^{\mathcal{B}}\right)^2$ (for $\alpha, \beta = E, I$) in the thermodynamic limit $N_{E,I} \rightarrow \infty$. In other words, in infinite-size networks with independent Brownian motions, the membrane potentials are independent too, leading to local chaos with small fluctuations (indeed, under our weak-noise assumption, their standard deviation $\sigma_{\alpha}^V \approx \frac{\sigma_{\alpha}^{\mathcal{B}}}{\sqrt{2\tau_{\alpha}}}$ is small). Local chaos is usually invoked to justify the mean-field description of large neural networks and is compatible with recent findings in visual cortex (Ecker et al. 2010; Renart et al. 2010; Tetzlaff et al. 2012).

Interestingly, also finite-size networks can however experience decorrelated activity. In Fasoli et al. (2015) the authors showed that, for any N , weak correlations occur for strongly depolarizing or strongly hyperpolarizing external inputs, if the Brownian motions are independent. This phenomenon can be proven to occur in networks with any topology from general considerations about its Jacobian matrix (see the Online Resource 1). However, this approach is only qualitative, therefore it does not provide any explicit formula of the cross-correlation. In Fasoli et al. (2015), explicit expressions of the correlation structure were obtained for networks with regular topology, through the analytical calculation of the eigenquantities of the synaptic connectivity matrix. In this work we extended the analytical calculations to multi-population networks (an explicit example is shown in Eqs. (11), (12), (13) for the two-population case), whose topology is generally irregular since the corresponding graph has non-uniform synaptic weights.

By taking advantage of the analytical expressions of the cross-correlation, we observe that the formation of asynchronous states can be proven for the two-population case as a consequence of $\gamma_{\alpha\alpha}^{\alpha\alpha} \rightarrow -\frac{1}{2\lambda_{\alpha}}$, $\gamma_{\beta\beta}^{\alpha\alpha} \rightarrow 0$ (with $\alpha \neq \beta$) and $\gamma_{EI}^{\alpha\alpha} \rightarrow 0$ for $|I_{E,I}| \rightarrow \infty$, which in turn is due to the saturation of the activation function $\mathcal{A}(V)$. For the same reason, the standard deviations $\sigma_{E,I}^V$ of the membrane potentials in the two populations decrease with the input. Interestingly, the input-driven reduction of both the correlation and the variance of the neural responses is supported by experimental evidence (Tan et al. 2014; Ponce-Alvarez et al. 2015). Figure 3 shows an example of formation of an

asynchronous regime, which is obtained for the values of the parameters in Table 1 and for strong stimuli ($I_E > 13$, $I_I = -35$).

The theory developed in this article can also be applied when the noise sources in Eq. (1) are correlated. However, given arbitrary (i.e. non-specifically tuned) parameters, the firing-rate network model of Eq. (1) generally does not undergo the formation of asynchronous states when the Brownian motions are not independent. In particular, since $\gamma_{\alpha\alpha}^{EI} \rightarrow 0$ and $\gamma_{EI}^{EI} \rightarrow \frac{1}{2\sqrt{\lambda_E\lambda_I}}$ for $|I_{E,I}| \rightarrow \infty$, we get $(\sigma_{\alpha}^V)^2 \rightarrow \frac{1}{2\tau_{\alpha}} \left(\sigma_{\alpha}^{\mathcal{B}}\right)^2$ and $C_{\alpha\beta}^V \rightarrow C_{\alpha\beta}^{\mathcal{B}}$ (see also Figs. 5 and 6, where we plot the cross-correlations for different values of $C^{\mathcal{B}}$). In other words, in the considered small-size firing-rate model, for strong stimuli the correlation between the membrane potentials converges to that between the Brownian motions, and again the fluctuations of the membrane potentials have small standard deviation. Note, however, that other network models (such as the networks of non-leaky integrate-and-fire neurons considered in Moreno-Bote 2014 and Moreno-Bote et al. 2014) can have very weak correlations even for correlated inputs if the connectivity matrix is specifically tuned for that.

3.2 Synchronous states

We now examine the existence of network states characterized by large correlations among neurons, called *synchronous states* (Ginzburg and Sompolinsky 1994; De La Rocha et al. 2007; Harris and Thiele 2011).

Large correlations can be generated in two different ways. The most straightforward is by increasing the correlation between the Brownian motions, while a more subtle one takes advantage of the phenomenon called *critical slowing down* (Scheffer et al. 2009; Kéfi et al. 2013). Contrary to asynchronous states, the most important features of critical slowing down are large temporal fluctuations of the membrane potentials and strong cross-correlations (even though the Brownian motions are independent). Critical slowing down typically occurs at the bifurcation points of the system. Near local bifurcations, the real part of one of the eigenvalues tends to zero, therefore the system becomes increasingly slow in recovering from small perturbations. As a consequence, the system has a longer memory for perturbations, and its dynamics is characterized by larger stochastic fluctuations and stronger correlations. In Fasoli et al. (2016) we performed a detailed bifurcation analysis in the two-population case and for the values of the parameters in Table 1, obtaining the entangled set of local and global bifurcations shown in Fig. 2. Local bifurcations occur when a parameter variation causes the stability of an equilibrium point to change, therefore they are studied through the eigenvalues of the Jacobian matrix. Local

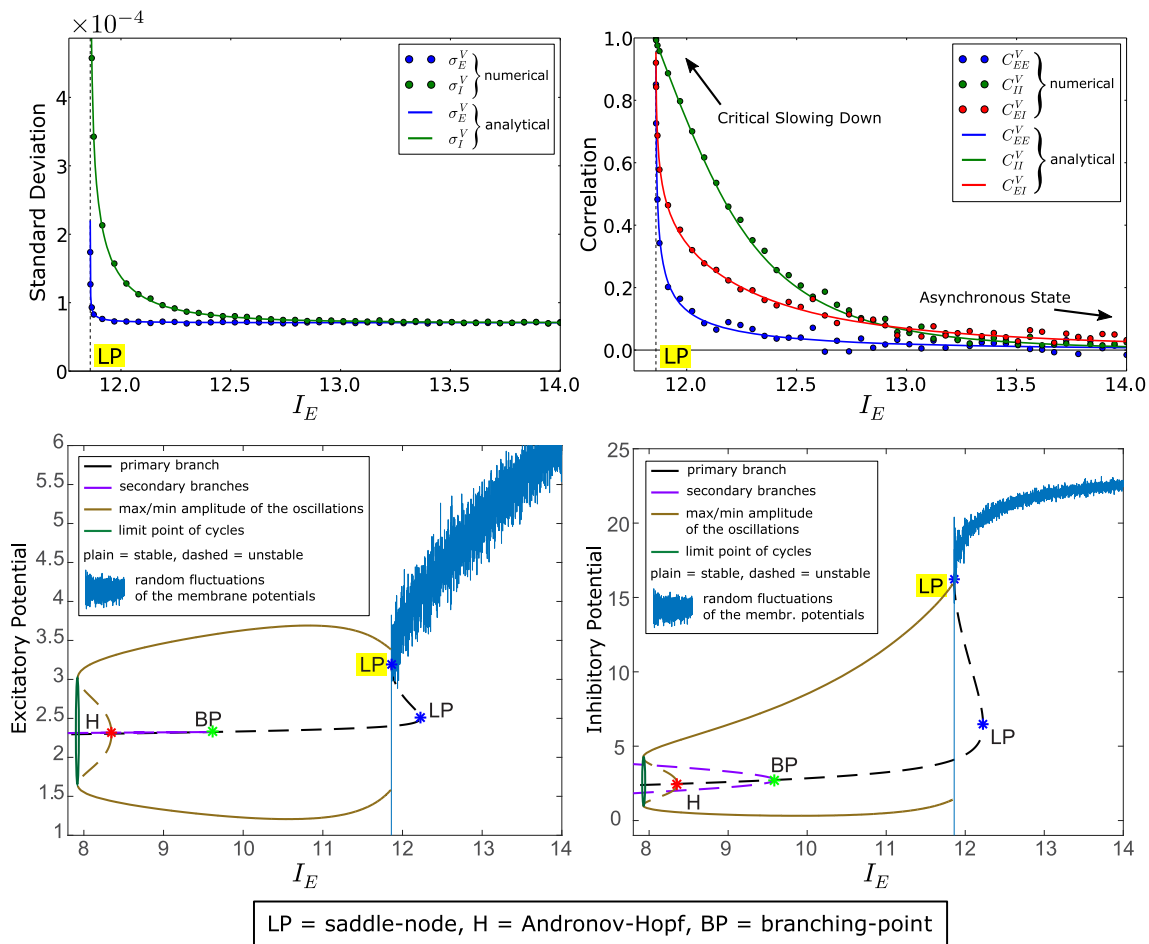


Fig. 3 Transition between asynchronous and synchronous states near a saddle-node bifurcation. The top panels show a good agreement between the numerical approximations of the standard deviation and correlation (left and right panel respectively), and the corresponding analytical formulas (see Eqs. (12) and (13)). The numerical approximations have been obtained through the methods described in Section 2.3. For large inputs ($I_E > 13$, see also Fig. 2) we observe the formation of an asynchronous state, which is characterized by weak correlation and low variability. On the other hand, near a saddle-node bifurcation ($I_E \approx 11.86$, see the highlighted LP), we observe the formation of a synchronous state characterized by strong correlations and

wide temporal fluctuations (critical slowing down). The bottom panels show numerical simulations of the fluctuations of the membrane potentials in the excitatory and inhibitory population (left and right panel respectively), calculated at $t = 30$ for different values of I_E and superposed to the codimension one bifurcation diagram of the network. The fluctuations are displayed at $3,000\times$ actual size in the excitatory and inhibitory population, in order to make them visible on the bifurcation diagrams. The reader may verify the agreement between the standard deviations (top-left panel) and the envelope of the fluctuations of the membrane potentials

bifurcations can be of codimension one or two, depending on the number of parameters (e.g. $I_{E,I}$) that must be varied for the bifurcation to occur. As shown in Fig. 2, the local bifurcations of codimension one the network undergoes are saddle-node, Andronov-Hopf and branching-point bifurcations, while those of codimension two are cusp, Bogdanov-Takens, generalized Hopf and zero-Hopf bifurcations. As discussed in Fasoli et al. (2016), for $N_I > 2$ other kinds of local bifurcations of codimension two may occur due to spontaneous-symmetry breaking (e.g. the double-Hopf bifurcation). Nevertheless, for simplicity, in the main text we restrict our discussion to the case $N_I = 2$.

Similarly to Kuehn (2013), we study the behavior of the correlation only at the local bifurcations of the network, and in particular we consider only those of codimension one. These bifurcations are studied in the following paragraphs for the case of two neural populations, and in Section (S6) of the Online Resource 1 for the case of an arbitrary number of populations. Our theory can also be used to study the behavior of the correlation near local bifurcations of codimension two, but due to the high variety of the bifurcations the system exhibits, a complete study is beyond the purpose of this article. We note that there are also global bifurcations (for example the homoclinic, limit point of cycles, and

torus¹ curves that are global bifurcations of codimension one, and saddle-node on invariant circle points that represent the only global bifurcations of codimension two). However, to our knowledge no analytical method is known for studying the global bifurcations of Eq. (1). We therefore restrict our analysis to local bifurcations only.

Saddle-node bifurcations (catastrophic transitions)

Saddle-node bifurcations occur in many dynamical systems. They represent tipping points at which tiny perturbations can cause an abrupt and discontinuous change of the equilibrium point of the system. In particular, in neuroscience some authors proposed that the whole-cortex activity may undergo saddle-node bifurcations during anesthetic administration at the edge between conscious and unconscious states (Steyn-Ross et al. 2004).

In Fasoli et al. (2016) we proved that, in the two-population case, the network undergoes a saddle-node bifurcation whenever one of the eigenvalues $\lambda_{0,1}^{\mathcal{R}}$ in Eq. (8) tends to zero. The saddle-node bifurcations are described by the blue curves in Fig. 2. In Fasoli et al. (2016) we also proved that a necessary condition for the formation of these bifurcations is:

$$\frac{N_E - 1}{N - 1} J_{EE} \frac{v_E^{\max} \Lambda_E}{4} \tau_E > 1, \tag{14}$$

or in other words sufficiently strong self-excitatory weights are required. From Eq. (13) we observe that for $\lambda_0^{\mathcal{R}} \rightarrow 0^-$ or $\lambda_1^{\mathcal{R}} \rightarrow 0^-$ the functions Υ diverge, therefore the terms proportional to $\frac{1}{2\lambda_{E,I}}$ in Eq. (12) become negligible. This implies $\zeta_{\alpha\alpha}^V \sim (\sigma_{\alpha}^V)^2 \rightarrow \infty$ and $\zeta_{EI}^V \sim \sigma_E^V \sigma_I^V$, therefore $C_{\alpha\beta}^V \sim 1$ between every population. Thus, when the network is close to a saddle-node bifurcation, we observe the emergence of critical slowing down. Moreover, we obtain a simple relation between the variances of the two neural populations, namely $\sigma_I^V \sim K \sigma_E^V$, where $K \stackrel{\text{def}}{=} \lim_{\lambda_0^{\mathcal{R}} \rightarrow 0^-} K_0 =$

$$\lim_{\lambda_1^{\mathcal{R}} \rightarrow 0^-} K_1 = \frac{N_E J_{IE} \mathcal{A}'_E(\mu_E)}{\frac{M_I}{\tau_I} - (N_I - 1) J_{II} \mathcal{A}'_I(\mu_I)}.$$

The reader can also verify that $\zeta_{EI}^V > 0$ for $\lambda_{\alpha}^{\mathcal{R}} \rightarrow 0$ as a consequence of $K > 0$, which in turn is due to $J_{IE} > 0$ and $J_{II} < 0$. An example of critical slowing down obtained for $I_E \approx 11.86$, $I_I = -35$ and $C_{EE}^{\mathcal{B}} = C_{II}^{\mathcal{B}} = C_{EI}^{\mathcal{B}} = 0$ is reported in Fig. 3. We observe that this phenomenon occurs even if there is no correlation between the Brownian motions (i.e.

¹More precisely, the torus bifurcation is a local bifurcation of the Poincaré map of a limit cycle of the network (Kuznetsov 1998). For this reason the torus bifurcation corresponds to a change of stability of the fixed points of the Poincaré map, and not to a change of stability of the equilibrium points of Eq. (1). Therefore, the torus bifurcation cannot be studied through the eigenvalues of the Jacobian matrix of the network.

$C_{\alpha\beta}^{\mathcal{B}} = 0$), therefore it is entirely a consequence of the neural interactions mediated by the synaptic connections.

Andronov-Hopf bifurcations (oscillations)

Andronov-Hopf bifurcations correspond to the emergence of neural oscillations, which are a phenomenon often seen in cortical activity and which is thought to play a key role in many cognitive processes (Ward 2003). In the two-population case, the network undergoes an Andronov-Hopf bifurcation whenever $\lambda_{0,1}^{\mathcal{R}}$ in Eq. (8) are complex-conjugate purely imaginary. The Andronov-Hopf bifurcations are described by the red curves in Fig. 2. In Fasoli et al. (2016) we proved that a necessary condition for the formation of these bifurcations is:

$$\frac{v_E^{\max} \Lambda_E}{4\mathfrak{J}} > 1 \quad \text{and} \quad b^2 - 4ac > 0 \tag{15}$$

where:

$$\begin{aligned} \mathfrak{J} &= \frac{-b - \sqrt{b^2 - 4ac}}{2a}, \\ a &= \left(\frac{N_E - 1}{N - 1} J_{EE} \right)^2 - \frac{N_E N_I (N_E - 1)}{(N - 1)^2 (N_I - 1)} \frac{J_{EE} J_{EI} J_{IE}}{J_{II}}, \\ b &= -\frac{2}{\tau_E} \frac{N_E - 1}{N - 1} J_{EE} + \frac{N_E N_I}{(N - 1)(N_I - 1)} \frac{J_{EI} J_{IE}}{J_{II}} \left(\frac{1}{\tau_E} + \frac{1}{\tau_I} \right), \\ c &= \frac{1}{\tau_E^2}. \end{aligned}$$

The mechanism of formation of critical slowing down at the Andronov-Hopf bifurcations is similar to that described in the previous paragraph for the saddle-node-bifurcations, therefore we discuss it only briefly. Whenever the network approaches an Andronov-Hopf bifurcation, we get $\lambda_0^{\mathcal{R}} + \lambda_1^{\mathcal{R}} \rightarrow 0^-$, which causes the terms Υ to diverge (see Eq. (13)). For this reason the variance of the membrane potentials diverges as well, while the cross-correlation tends to one, similarly to the case of the saddle-node bifurcations. This proves that the network undergoes critical slowing down also at the Andronov-Hopf bifurcations. The main difference with the case of the saddle-node-bifurcations is represented by the inter-population correlation C_{EI}^V , which does not tend to one near the Andronov-Hopf bifurcations (this phenomenon is discussed in more detail in Section (S6.2.2) of the Online Resource 1). An example obtained for $I_E = 1$, $I_I \approx -13.67$ and $C_{EE}^{\mathcal{B}} = C_{II}^{\mathcal{B}} = C_{EI}^{\mathcal{B}} = 0$ is shown in Fig. 4.

Branching-point bifurcations (spontaneous symmetry-breaking)

In the deterministic model (i.e. for $\sigma_{\alpha}^{\mathcal{B}} = 0$), because of the homogeneity assumption of Section 2.1, neurons within each population are expected to have identical dynamics. This means that, in the absence of noise, the

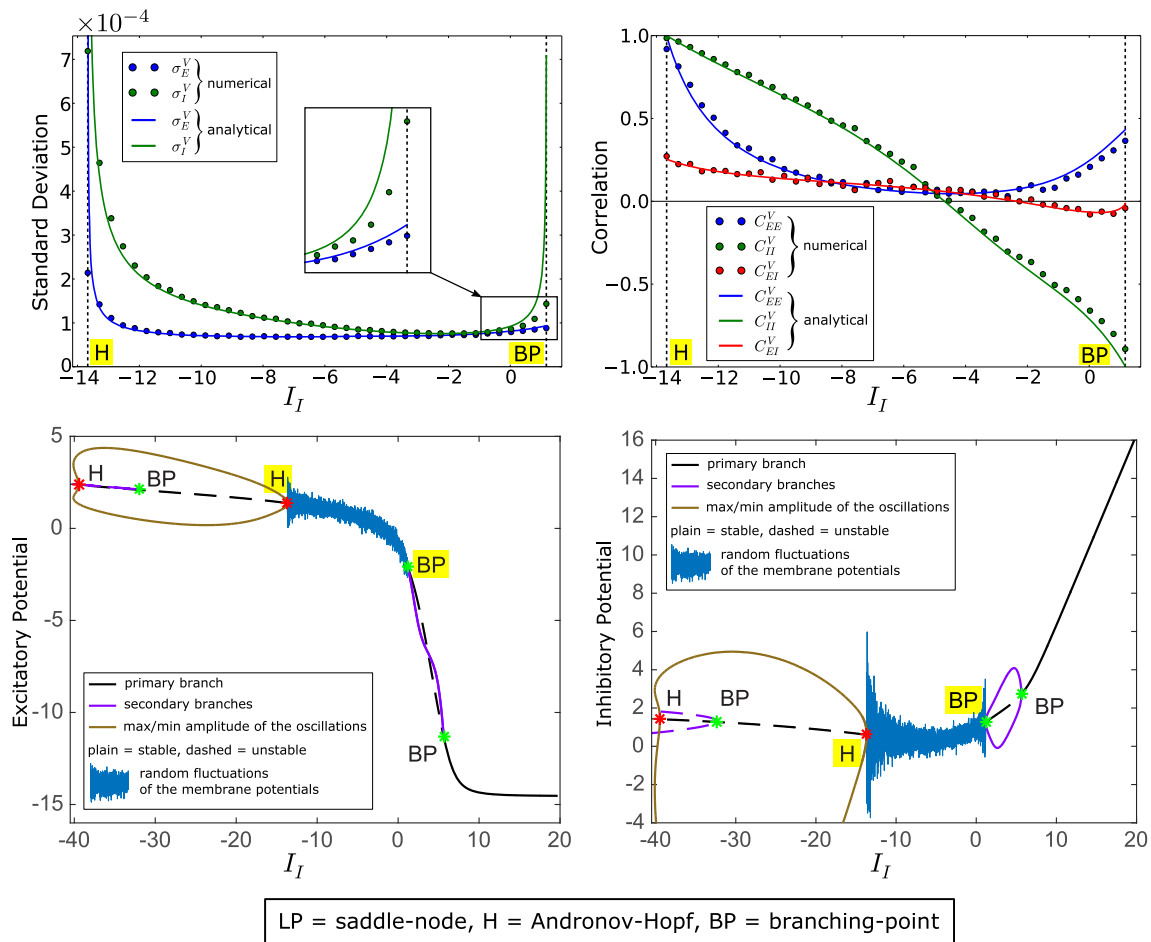


Fig. 4 Fluctuations and cross-correlations of the membrane potentials between Andronov-Hopf and branching-point bifurcations. The simulations are similar to those of Fig. 3), but now we set $I_E = 1$ and we vary the input to the inhibitory population (see Fig. 2), obtaining a transition between an Andronov-Hopf bifurcation ($I_I \approx -13.67$, see the highlighted H) and a branching-point bifurcation ($I_I \approx 1.165$, highlighted BP). We obtain a good agreement between numerical and analytical correlations for any current I_I in the range, while the standard deviations display a good agreement only when I_I is sufficiently

far from the bifurcation points. At the Andronov-Hopf and branching-point bifurcations the standard deviations predicted by the analytical formulas are larger than those obtained numerically. This suggests that generally second-order corrections to Eqs. (12) and (13) play a stronger role when the network undergoes these local bifurcations. Nevertheless, the first-order approximation describes qualitatively the increase of the standard deviation that characterizes critical slowing down

network dynamics is invariant under transformations in the group $S_{N_0} \times \dots \times S_{N_{\mathbb{P}-1}}$, where S_{N_α} is the permutation group on N_α items (also known as *symmetric group*). When we include noise ($\sigma_\alpha^{\mathcal{B}} > 0$), we introduce a small *explicit symmetry-breaking* into Eq. (1). However, the behavior of a nearly symmetric dynamical system is more similar to that of an idealized symmetric system than that of a completely asymmetric one (Stewart et al. 2003). Therefore, if the degree of explicit heterogeneity introduced by the noise is not too strong, it is legitimate to study Eq. (1) as a perturbation of the corresponding deterministic system. However, symmetry-breaking may occur also in the deterministic model. Indeed, at the branching-point bifurcations

we observe the formation of a *spontaneous symmetry-breaking* (Fasoli et al. 2016) because some of the neurons within a given inhibitory population become dynamically distinct from the others. In other words, we observe the formation of an heterogeneous inhibitory population, even if the neural equations (1) for $\sigma_\alpha^{\mathcal{B}} = 0$ do not contain any term that breaks explicitly the symmetry. Interestingly, this phenomenon is a consequence of the finite size of the network, therefore it does not occur in the thermodynamic limit (Fasoli et al. 2016).

In Fasoli et al. (2016) we also proved that, in the two-population case, branching-point bifurcations occur whenever $\lambda_I = 0$ (see the light green dot-dashed curves in

Fig. 2) and that a necessary condition for their formation is:

$$\frac{\tau_I |J_{II}| v_I^{\max} A_I}{4(N-1)} > 1. \quad (16)$$

This means that sufficiently strong self-inhibitory weights are required for the bifurcations to occur. According to Eq. (12), for $C_{II}^{\mathcal{B}} < 1$ and $\lambda_I \rightarrow 0^-$ only the variance of the inhibitory neurons diverges. As a consequence, in the case $C_{II}^{\mathcal{B}} < 1$ we get $(\sigma_I^V)^2 \sim -(\sigma_I^{\mathcal{B}})^2 \left(1 - \frac{1}{N_I}\right) \Theta_I$ and $\zeta_{II}^V \sim (\sigma_I^{\mathcal{B}})^2 \frac{\Theta_I}{N_I}$, from which we conclude that $C_{II}^V \sim \frac{1}{1-N_I}$. According to Fasoli et al. (2015), this is the lower bound of the correlation between fully-connected neurons in a homogeneous population with size N_I . Since $\frac{1}{1-N_I} < 0$ for $N_I \geq 2$, at the branching-point bifurcations the inhibitory neurons are maximally anti-correlated (in particular, the correlation tends to -1 only for $N_I = 2$). From these results we conclude that, contrary to the saddle-node and Andronov-Hopf bifurcations, at the branching points critical slowing down occurs only in the inhibitory population. This is confirmed by Fig. 4, which shows an example obtained for $I_E = 1$, $I_I \approx 1.165$ and $C_{EE}^{\mathcal{B}} = C_{II}^{\mathcal{B}} = C_{EI}^{\mathcal{B}} = 0$. Intuitively, the membrane potentials become anti-correlated because the inhibitory neurons follow different branches of stationary solutions beyond the branching-point (see the codimension one bifurcation diagram in the bottom-right panel of Fig. 4). Therefore while the potential of one neuron increases due to noise fluctuations, the potential of the other neuron decreases and viceversa, resulting in a negative correlation.

On the other hand, for $C_{II}^{\mathcal{B}} = 1$ and $\lambda_I \rightarrow 0^-$, from Eq. (12) we get:

$$\begin{aligned} (\sigma_I^V)^2 &= \zeta_{II}^V = (\sigma_E^{\mathcal{B}})^2 \gamma_{II}^{EE} \Delta_E + (\sigma_I^{\mathcal{B}})^2 \gamma_{II}^{II} \\ &\quad + 2\sigma_E^{\mathcal{B}} \sigma_I^{\mathcal{B}} \gamma_{II}^{EI} C_{EI}^{\mathcal{B}}, \end{aligned}$$

therefore now $(\sigma_I^V)^2$ does not diverge anymore and $C_{II}^V = 1$ (see also Figs. 5 and 6). To conclude, for $C_{II}^{\mathcal{B}} = 1$ and $\lambda_I = 0$, Eq. (12) gives an indeterminate form $\frac{0}{0}$ for the variance $(\sigma_I^V)^2$ and the covariance ζ_{II}^V , which is represented by the empty circles in the middle panels of Fig. 5 and in the bottom-central panel of Fig. 6. This result can be intuitively interpreted as the consequence of the competition between the positive correlation introduced by the Brownian motions and the anti-correlation generated by the branching point.

4 Discussion

We developed a theory of correlations in a multi-population graded-firing-rate network model of arbitrary size. This theory, taking advantage of mathematical methods (Fasoli et al. 2015; Fasoli et al. 2016) not based on statistical averages, allows a rigorous analytical understanding of how correlations depend on the parameters and structure of networks of arbitrary size. These networks include small-size circuits of a few tens of neurons, such as for example neural circuits in some invertebrates (Williams and Herrup 1988), and are not restricted to large systems as in previous work (Ginzburg and Sompolinsky 1994; Bressloff 2009; Renart et al. 2010; Buice and Chow 2013). Importantly, our formalism allowed us to investigate the interplay between network dynamics and network statistics, by combining the non-linear analysis of the bifurcations of the network with a linear analysis of the covariance matrix.

Through our formalism, explicit calculations of the cross-correlations are possible only for networks composed of a few neural populations. In particular, for exemplary purposes, in the main text we focused on the case of a two-population network, composed of excitatory and inhibitory neurons, and we extensively validated the closed-form expression of its correlation structure through numerical simulations. However, it is important to observe that the more abstract formalism introduced in the [Online Resource 1](#) does not rely on the explicit calculation of the cross-correlations, and therefore predicts the formation of synchronous and asynchronous states in networks composed of an arbitrary number of neural populations.

In the following we discuss the advances of our results with respect to previous work, and the implications of our work to better understand neural network dynamics.

4.1 Progress with respect to previous modeling work

A first advance with respect to some previous theories of correlations based on rate models made of binary neurons (Ginzburg and Sompolinsky 1994; Renart et al. 2010), was that we were able to introduce a biologically more realistic network composed of graded neurons with continuous firing rate, without losing the possibility to derive analytical expressions for the network correlations. However, a more important advance with respect to these previous works was that we could consider correlations among populations of arbitrary size rather than of large size. In the context of neuroscience, this is important because some networks (e.g. the nervous system of some invertebrates such as rotifers and nematodes) are composed only of a few tens of cells (Williams and Herrup 1988), and because small networks have, as shown in this work, different relationships between

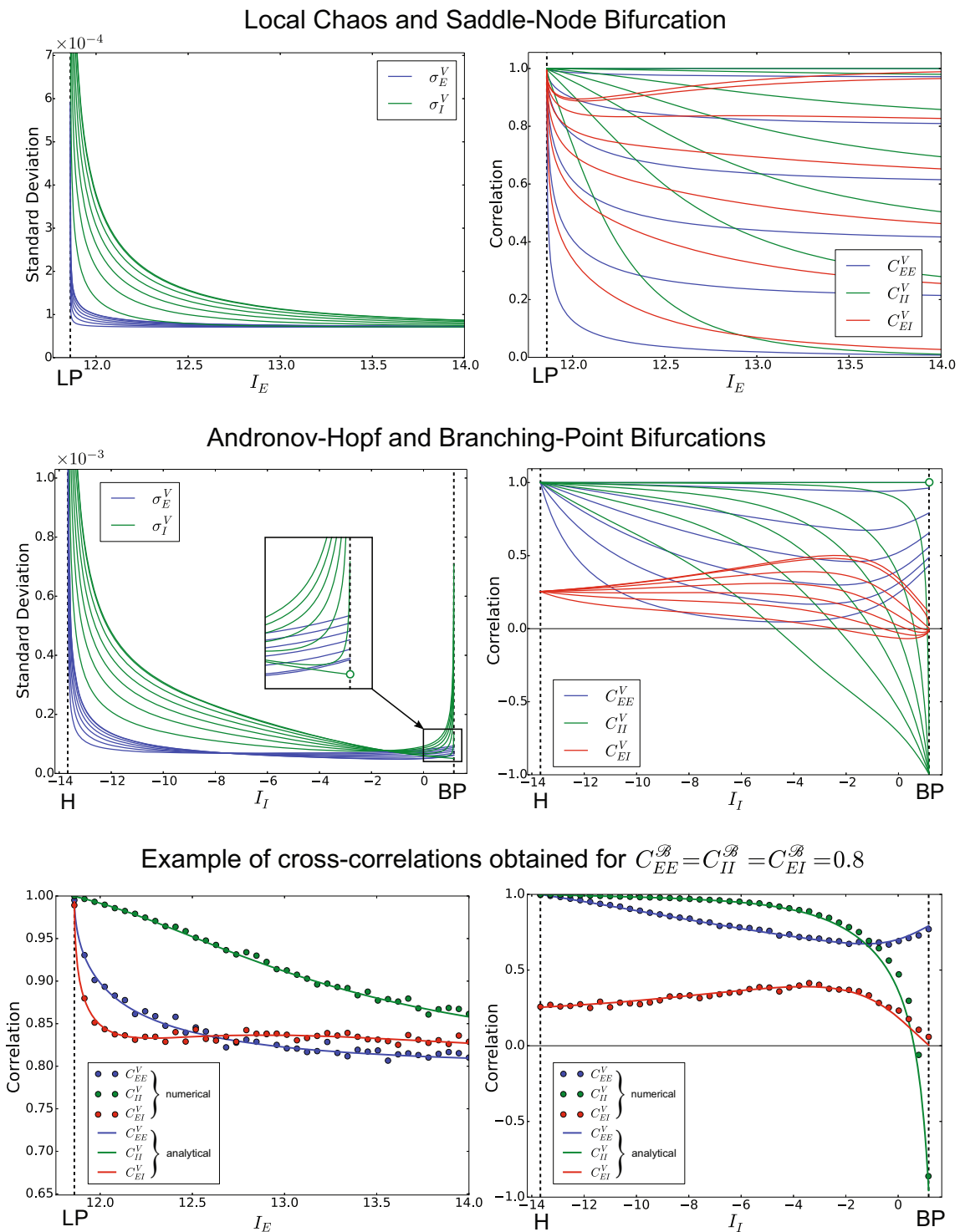
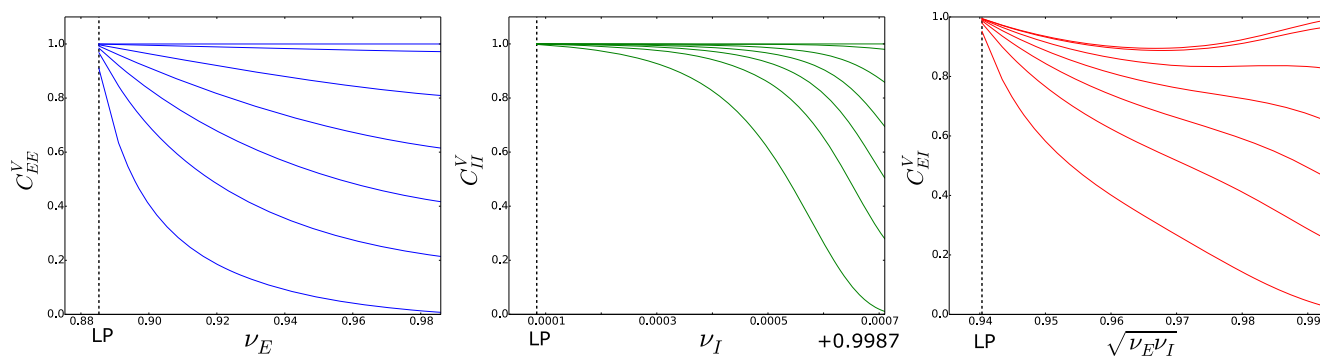


Fig. 5 Fluctuations and cross-correlations of the membrane potentials as a function of the input and of the noise correlation. The top panels show the standard deviation (left) and the cross-correlation (right) of the membrane potentials when the network is close to a saddle-node bifurcation (similarly to Fig. 3), for different values of the noise correlation. The curves have been obtained from Eqs. (12) and (13) for $C_{EE}^{\mathcal{B}} = C_{II}^{\mathcal{B}} = C_{EI}^{\mathcal{B}} = 0, 0.2, 0.4, 0.6, 0.8, 0.97, 1$. The panels show that the noise correlation increases both the standard deviation

and the cross-correlation, for $\sigma_{E,I}^{\mathcal{B}}$ fixed (see Table 1). The middle panels show similar results for the neural states between Andronov-Hopf and branching-point bifurcations (compare with Fig. 4). The only difference is observed close to the branching-point bifurcation, where $\sigma_{E,I}^{\mathcal{B}}$ decrease with the noise correlation. The bottom panels show the comparison between the analytical and numerical cross-correlations in the case $C_{EE}^{\mathcal{B}} = C_{II}^{\mathcal{B}} = C_{EI}^{\mathcal{B}} = 0.8$

Local Chaos and Saddle-Node Bifurcation



Andronov-Hopf and Branching-Point Bifurcations

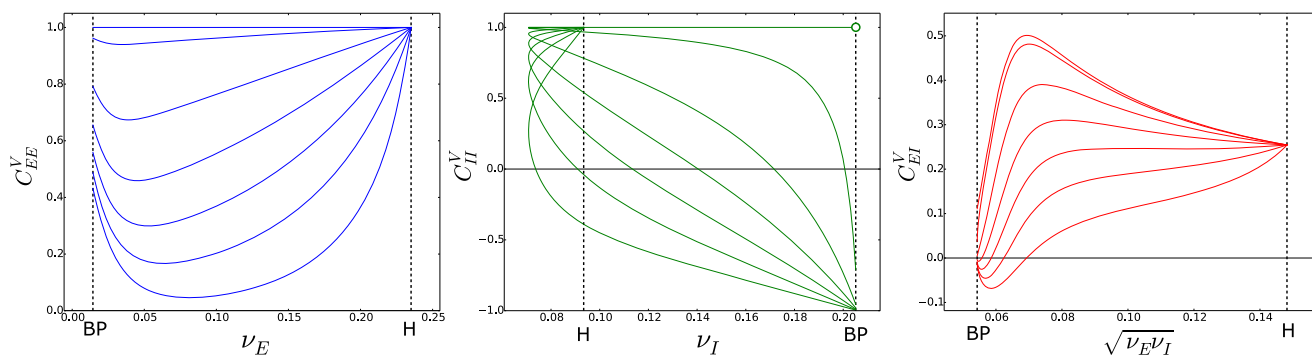


Fig. 6 Cross-correlations of the membrane potentials as a function of the firing rate and of the noise correlation. The figure shows the cross-correlation of the membrane potentials as a function of the geometric mean of the firing rates $\sqrt{\nu_\alpha \nu_\beta} = \sqrt{\mathcal{A}_\alpha(\mu_\alpha) \mathcal{A}_\beta(\mu_\beta)}$

for $\alpha, \beta = E, I$. Similarly to Fig. 5, the curves have been obtained from Eqs. (12) and (13) for $C_{EE}^{\mathcal{B}} = C_{II}^{\mathcal{B}} = C_{EI}^{\mathcal{B}} = 0, 0.2, 0.4, 0.6, 0.8, 0.97, 1$. The panels show a non-monotonic dependence of the correlation on the firing rates

dynamics and correlations than the large scale ones. In particular, the main difference between small and large networks is the formation of spontaneous-symmetry breaking in the inhibitory populations. This phenomenon, which we discuss in the next subsections, increases considerably the complexity of the dynamics in small networks, and determines relationships between dynamics and correlations that are not predicted by large-network studies.

In the recent work (Fasoli et al. 2015), the authors were able to begin investigating some properties of correlations in small networks. However, in this earlier work, they could only investigate neural circuits with regular topology, while here we extended their theory to multi-population networks with irregular topology. Importantly, and unlike in Fasoli et al. (2015), here we studied the relation between the correlation structure of the network and the bifurcation points of the neural dynamics. Specifically, we studied the behavior of the correlation in terms of the stimuli I_α , and this analysis revealed the ability of the network to switch dynamically from asynchronous regimes, characterized by weak correlation and low variability, to

synchronous regimes, characterized by strong correlations and wide temporal fluctuations of the state variables.

Note that the mathematical bases for systematically studying critical slowing down at bifurcations up to codimension two were laid out in general terms in Kuehn (2013). Here, we show that such deep mathematical concepts can be realized and found in neural networks.

4.2 New insights into the relation between network dynamics and correlation structure

These mathematical advances with respect to previous work allowed us to reach a set of novel insights into the relation between the correlation structure of the network and the bifurcation points of the neural dynamics. In particular, we found how transitions between synchronous and asynchronous states relate to bifurcations.

The asynchronous regime can be observed in large networks driven by independent sources of noise (Samuelides and Cessac 2007; Touboul et al. 2012; Baladron et al. 2012; Baladron Pezoa et al. 2012). Here, however, we proved

that in the small-size firing-rate network model considered here asynchronous states can be generated dynamically by strong stimuli. In the firing-rate model we study, this phenomenon occurs for arbitrary (i.e. not specifically tuned) network parameters, provided the input is strong enough and the Brownian motions are independent. This is a consequence of the saturation of the activation function which avoids blowing-up solutions. However, asynchronous states can occur through this mechanism also in networks with non-saturating functions for specific values of the network parameters. Interestingly, the decrease of both the variance and the cross-correlation of the neural responses with the input is supported by experimental evidence (Tan et al. 2014; Ponce-Alvarez et al. 2015). For arbitrary (i.e. non-specifically tuned) parameters, the small-size firing-rate network model of Eq. (1) in general does not undergo the formation of asynchronous states when the Brownian motions are correlated. However, it is important to note that, unlike our case of small-size circuits, in large neural networks several mechanisms have been shown to be able to decorrelate network activity even in the presence of correlated inputs (see e.g. van Vreeswijk and Sompolinsky 1998; Renart et al. 2010; Tetzlaff et al. 2012; Hennequin et al. 2016).

Importantly, we found that the synchronous regime occurs near the bifurcation points of the network, which in this model can be analytically determined (Fasoli et al. 2016). In particular, in the present article we considered the local bifurcations of codimension one, namely the saddle-node, Andronov-Hopf and branching-point bifurcations. Contrary to the strongly positive correlations that occur at the saddle-node and Andronov-Hopf bifurcations, at the branching points we have observed the emergence of strong *anti-correlations* between inhibitory neurons.

The emergence of strong correlations at any of the local bifurcations of the network is a finite-size effect, and does not require correlated sources of noise. Indeed, for a network with independent Brownian motions, in Fasoli et al. (2015) the authors proved that the neurons are strongly synchronized at a time instant t_N that depends on the size of the network. Strong correlations are very unlikely to occur in large networks after short time intervals, since $t_N \rightarrow \infty$ in the limit $N \rightarrow \infty$. However, exceptions may arise in sparsely-connected networks (see Section 4.5), or if the Brownian motions are correlated.

It is also important to observe that, in the case of networks made of two populations, Eqs. (11), (12), (13) represent a mathematical description of a multidimensional continuum of states, ranging from asynchronous to synchronous states, and corresponding to varying levels of spontaneous fluctuations and cross-correlation in neural population activity. The existence of this continuum was first proposed in Harris and Thiele (2011), where the authors

observed that the multidimensional nature of the continuum emerges from analyzing the structure of the fluctuations and of the cross-correlation under several different behavioral and experimental conditions. These conditions can be effectively modeled for example by varying the strength of the synaptic weights (see Section 4.4), or by varying the strength of the input parameters $I_{E,I}$ in order to describe the presence or absence of external stimulation.

4.3 Spontaneous symmetry-breaking as the origin of anti-correlations

We proved that at the branching-point bifurcations the inhibitory neurons become strongly anti-correlated as a consequence of spontaneous symmetry-breaking. More generally, other kinds of spontaneous symmetry-breaking can occur in the network, depending on its symmetries. For example, in the case of two identical inhibitory populations, two different symmetries may be broken: the symmetry between neurons in a given population, and that between the two populations. In the latter case, the two populations would behave differently from each other, while keeping their corresponding neurons homogeneous. This phenomenon is also characterized by strongly positive intra-population correlations and strongly negative inter-population correlations (result not shown), reinforcing the idea of a general relationship between spontaneous symmetry-breaking and anti-correlations. In Fasoli et al. (2016) we described possible extensions of our formalism to spatially extended networks with more complex symmetries, therefore spontaneous symmetry-breaking is likely to affect also the cross-correlation structure of large-scale neural models.

Negative correlations have been observed in resting-state fMRI experiments, for example during cognitive tasks performed by human subjects (Fox et al. 2005), and also in the frontolimbic circuit of awake rats (Liang et al. 2012), but their origin and functional role are still poorly understood. Our findings suggest branching-point bifurcations and spontaneous symmetry-breaking as a potential neurobiological basis of this phenomenon.

4.4 How pharmacological manipulations may affect the correlation structure of the network

One potential application of our formalism is to model the effect on neural correlations of pharmacological application of drugs that act as agonist or antagonists of the major neurotransmitters (Curtis et al. 1971; Krogsgaard-Larsen et al. 1980; Corda et al. 1992; Chen et al. 1992; Cunningham and Jones 2000; Garcia et al. 2010). The effect of these drugs can be effectively modeled by varying the synaptic strengths in the model; in other words their

effect on network dynamics and correlations can be studied through a bifurcation analysis in terms of the parameters $J_{\alpha\beta}$. Our results suggest that pharmacological manipulations that alter synaptic weights of small networks will change their dynamics profoundly. For example whenever for a set of synaptic weights the network does not satisfy the conditions (14), (15), (16), the corresponding bifurcations become forbidden for any pair of stimuli (I_E, I_I). It is therefore natural to speculate that our formalism may add to theoretical models of the effect of drugs on neural dynamics (Foster et al. 2008).

4.5 Future directions

We studied the cross-correlation structure of multi-population networks near local bifurcations of codimension one. Furthermore, our theory can be easily extended to the analysis of local bifurcations of larger codimension.

Another possible extension of our theory is the study of correlations in sparse networks. In Fasoli et al. (2015) the authors showed that, when the number of connections per neuron does not diverge for $N \rightarrow \infty$, asynchronous states in general do not occur in the thermodynamic limit for weak stimuli (compare with van Vreeswijk and Sompolinsky (1998), where the authors considered the case of sparse networks with infinite connections per neuron). Therefore, sufficiently sparse networks cannot rely on their size for generating asynchronous states, but uncorrelated activity can still be generated through strong stimuli or with special combinations of the network's parameters. Moreover, in Fasoli et al. (2016) we showed that in sparse networks the branching-point bifurcations are more likely to occur, resulting in a considerable increase of the complexity of the bifurcation diagrams.

It is also possible to study correlations in small neural circuits with random synaptic weights, extending the results obtained in Ginzburg and Sompolinsky (1994) and Renart et al. (2010) for large random networks of binary neurons. Their correlation structure can be calculated straightforwardly from the fundamental matrix Φ , by applying the formalism developed in Fasoli et al. (2015) (see Eqs. (4.3) and (4.6) therein). The bifurcation structure of networks with heterogeneous weights has been studied only in the limit of large systems (Hermann and Touboul 2012), and is still unexplored in the case of small neural circuits.

The effect of temporally correlated afferent currents on neural activity has been studied extensively in integrate-and-fire neural network models, see e.g. Brunel and Sergi (1998), Sakai et al. (1999), Moreno et al. (2002), and Renart et al. (2003). We observe that, by following the techniques described in Hennequin et al. (2016), also our finite-size firing-rate model can be easily extended to

include temporally correlated noise sources. Due to the non-linear interplay between spatial and temporal correlations, we expect to observe strong quantitative deviations from Eqs. (12) and (13), especially at the bifurcation points of the network. However, in our model these deviations do not affect qualitatively the properties of critical slowing down. In other words, in our model the explosion of the variance of the stochastic fluctuations, and the formation of arbitrarily strong cross- and auto-correlations, are expected to occur regardless of the temporal correlations of the noise sources. Moreover, by extending our perturbative approach to the second order, we would be able to quantify rigorously how temporal correlations affect the skewness of the stochastic fluctuations of the membrane potentials, which is expected to be strongly non-zero at the bifurcation points of the network (Scheffer et al. 2009).

The dependence of network correlations on the neuron's firing rates has been investigated extensively in recent years (De La Rocha et al. 2007; Ecker et al. 2014; Goris et al. 2014). In particular, for pairs of unconnected cortical neurons receiving correlated inputs *in vitro*, and for model integrate-and-fire neurons, in De La Rocha et al. (2007) the authors reported that correlations increase with the geometric mean of the firing rates. However, in our model we generally observed a non-monotonic dependence of the correlation on the firing rates, that depended on which dynamical state the network was in. In particular, we found non-monotonicity when the stimulus made the network switch between Andronov-Hopf and branching-point bifurcations, or when, given a strong correlation between the Brownian motions, the network was close to a saddle-node bifurcation (see the bottom panels and the top-right panel of Fig. 6, respectively). This result shows that the relation between firing rates and correlations are expected to be more complicated in neuronal networks than in pairs of non-interacting neurons. A consequence of the possibly non-monotonic and dynamical state-dependent relationship between rates and correlations is that the mean firing rates and correlations can act as separate information channels for the encoding of the strength of the external stimuli. A rigorous analysis of the encoding capability of our model may be performed by calculating analytically the Cramér-Rao bound (Abbott and Dayan 1999). In particular, it would be interesting to evaluate how the mean firing rates and the correlation structure of the network contribute to the analytical expression of the Cramér-Rao bound, and to determine the differences in the encoding capability of the network between synchronous and asynchronous regimes. In future work, this approach will allow us to determine systematically whether correlations are detrimental or helpful for the encoding of sensory information in firing-rate network models.

Online resource 1

Theory of correlations for multi-population networks In this supplemental text we extend our analysis of correlations to networks composed of an arbitrary number of neural populations. In particular, we prove that asynchronous states are elicited by strong stimuli, while synchronous states occur at the local bifurcations of the network, regardless of the number of neural populations. Moreover, we introduce a mathematical formalism for calculating the fundamental matrix Φ of the network, from which the correlation structure can be derived from Eq. (6). For the sake of clarity, we also implemented this formalism with Python in the [Online Resource 2](#).

Online resource 2

Python script 1 In this script we implemented the formalism, described in the [Online Resource 1](#), for the calculation of the fundamental matrix Φ of multi-population neural networks.

Online resource 3

Python script 2 This script performs a comparison between the numerical simulations (described in Section 2.3) and the analytical formulas of the variance and cross-correlation (Eqs. (11)–(13)), in the specific case of the two-population network considered in the main text.

Acknowledgements We thank the members of the IIT's Neural Computation laboratory for useful feedback.

Author Contributions DF and SP designed the study, with the contribution of AC. DF performed the study through analytical calculations and numerical simulations. AC performed the numerical study of the bifurcations. DF, AC and SP wrote the paper.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abbott, L.F., & Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural Computation*, 11(1), 91–101.
- Akam, T., & Kullmann, D.M. (2010). Oscillations and filtering networks support flexible routing of information. *Neuron*, 67(2), 308–320.
- Baladron, J., Fiasoli, D., Faueras, O., Touboul, J. (2012). Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons. *The Journal of Mathematical Neuroscience*, 2(1), 10.
- Baladron Pezoa, J., Fiasoli, D., Faueras, O. (2012). Three applications of GPU computing in neuroscience. *Computing in Science & Engineering*, 14(3), 40–47.
- Battaglia, D., Witt, A., Wolf, F., Geisel, T. (2012). Dynamic effective connectivity of inter-areal brain circuits. *PLoS Computational Biology*, 8(3), e1002438.
- Besserve, M., Lowe, S.C., Logothetis, N.K., Schölkopf, B., Panzeri, S. (2015). Shifts of gamma phase across primary visual cortical sites reflect dynamic stimulus-modulated information transfer. *PLoS Biology*, 13(9), e1002257.
- Binzegger, T., Douglas, R.J., Martin, K.A. (2004). A quantitative map of the circuit of cat primary visual cortex. *The Journal of Neuroscience*, 24(39), 8441–8453.
- Bohland, J.W., Wu, C., Barbas, H., Bokil, H., Bota, M., Breiter, H.C., Cline, H.T., Doyle, J.C., Freed, P.J., Greenspan, R.J., Haber, S.N., Hawrylycz, M., Herrera, D.G., Hilgetag, C.C., Huang, Z.J., Jones, A., Jones, E.G., Karten, H.J., Kleinfeld, D., Köster, R., Lester, H.A., Lin, J.M., Mensh, B.D., Mikula, S., Panksepp, J., Price, J.L., Safdieh, J., Saper, C.B., Schiff, N.D., Schmahmann, J.D., Stillman, B.W., Svoboda, K., Swanson, L.W., Toga, A.W., Van Essen, D.C., Watson, J.D., Mitra, P.P. (2009). A proposal for a coordinated effort for the determination of brainwide neuroanatomical connectivity in model organisms at a mesoscopic scale. *PLoS Computational Biology*, 5, e1000334.
- Boltzmann, L. (1872). Weitere studien über das wärmegleichgewicht unter gasmolekülen [Further studies on thermal equilibrium between gas molecules]. *Wien Ber.*, 66, 275–370.
- Bressloff, P.C. (2009). Stochastic neural field theory and the system-size expansion. *SIAM Journal on Applied Mathematics*, 70(5), 1488–1521.
- Brunel, N., & Hakim, V. (1999). Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Computation*, 11(7), 1621–1671.
- Brunel, N., & Sergi, S. (1998). Firing frequency of leaky integrate-and-fire neurons with synaptic current dynamics. *Journal of Theoretical Biology*, 195(1), 87–95.
- Buice, M.A., & Chow, C.C. (2013). Dynamic finite size effects in spiking neural networks. *PLoS Computational Biology*, 9(1), e1002872.
- Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, 304, 1926–1929.
- Buzsáki, G., Anastassiou, C.A., Koch, C. (2012). The origin of extracellular fields and currents – EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, 13, 407–420.
- Chen, H.S., Pellegrini, J.W., Aggarwal, S.K., Lei, S.Z., Warach, S., Jensen, F.E., Lipton, S.A. (1992). Open-channel block of n-methyl-D-aspartate (NMDA) responses by memantine: therapeutic advantage against NMDA receptor-mediated neurotoxicity. *The Journal of Neuroscience*, 12(11), 4427–4436.
- Cohen, M.R., & Maunsell, J.H. (2009). Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience*, 12(12), 1594–1600.
- Corda, M.G., Orlandi, M., Lecca, D., Giorgi, O. (1992). Decrease in GABAergic function induced by pentylentetrazol kindling in rats: antagonism by MK-801. *The Journal of Pharmacology and Experimental Therapeutics*, 262(2), 792–800.
- Cunningham, M.O., & Jones, R.S.G. (2000). The anticonvulsant, lamotrigine decreases spontaneous glutamate release but increases

- spontaneous GABA release in the rat entorhinal cortex *in vitro*. *Neuropharmacology*, 39(11), 2139–2146.
- Curtis, D.R., Duggan, A.W., Felix, D., Johnston, G.A.R., McLennan, H. (1971). Antagonism between bicuculline and GABA in the cat brain. *Brain Research*, 33(1), 57–73.
- David, O., Cosmelli, D., Friston, K.J. (2004). Evaluation of different measures of functional connectivity using a neural mass model. *NeuroImage*, 21(2), 659–673.
- De La Rocha, J., Doiron, B., Shea-Brown, E., Josić, K., Reyes, A. (2007). Correlation between neural spike trains increases with firing rate. *Nature*, 448(7155), 802–806.
- Deco, G., Jirsa, V.K., Robinson, P.A., Breakspear, M., Friston, K.J. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Computational Biology*, 4(8), e1000092.
- Ecker, A.S., Berens, P., Keliris, G.A., Bethge, M., Logothetis, N.K., Tolias, A.S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science*, 327(5965), 584–587.
- Ecker, A.S., Berens, P., Cotton, R.J., Subramaniyan, M., Denfield, G.H., Cadwell, C.R., Smirnakis, S.M., Bethge, M., Tolias, A.S. (2014). State dependence of noise correlations in macaque primary visual cortex. *Neuron*, 82(1), 235–248.
- Einevoll, G.T., Kayser, C., Logothetis, N.K., Panzeri, S. (2013). Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nature Reviews Neuroscience*, 14, 770–785.
- Fasoli, D., Faugeras, O., Panzeri, S. (2015). A formalism for evaluating analytically the cross-correlation structure of a firing-rate network model. *The Journal of Mathematical Neuroscience*, 5, 6.
- Fasoli, D., Cattani, A., Panzeri, S. (2016). The complexity of dynamics in small neural circuits. *PLoS Computational Biology*, 12(8), e1004992.
- Foster, B.L., Bojak, I., Liley, D.T.J. (2008). Population based models of cortical drug response: insights from anaesthesia. *Cognitive Neurodynamics*, 2(4), 283–296.
- Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27), 9673–9678.
- Freeman, W.J. (2000a). Mesoscopic neurodynamics: from neuron to brain. *The Journal of Physiology*, 94, 303–322.
- Freeman, W.J. (2000b). *Neurodynamics: an exploration in mesoscopic brain dynamics*. London: Springer.
- Friston, K.J. (2011). Functional and effective connectivity: a review. *Brain Connectivity*, 1(1), 13–36.
- Friston, K., Moran, R., Seth, A.K. (2013). Analysing connectivity with Granger causality and dynamic causal modelling. *Current Opinion in Neurobiology*, 23(2), 172–178.
- Garcia, P.S., Kolesky, S.E., Jenkins, A. (2010). General anesthetic actions on GABA(A) receptors. *Current Neuropharmacology*, 8(1), 2–9.
- Gilson, M., Moreno-Bote, R., Ponce-Alvarez, A., Ritter, P., Deco, G. (2016). Estimation of directed effective connectivity from fMRI functional connectivity hints at asymmetries of cortical connectome. *PLoS Computational Biology*, 12(3), e1004762.
- Ginzburg, I., & Sompolinsky, H. (1994). Theory of correlations in stochastic neural networks. *Physical Review E*, 50, 3171–3191.
- Goris, R.L., Movshon, J.A., Simoncelli, E.P. (2014). Partitioning neuronal variability. *Nature Neuroscience*, 17(6), 858–865.
- Grimbert, F. (2008). Mesoscopic models of cortical structures. PhD thesis, University of Nice - Sophia Antipolis.
- Grytskyy, D., Tetzlaff, T., Diesmann, M., Helias, M. (2013). A unified view on weakly correlated recurrent networks. *Frontiers in Computational Neuroscience*, 7, 131.
- Harris, K.D., & Thiele, A. (2011). Cortical state and attention. *Nature Reviews Neuroscience*, 12(9), 509–523.
- Hennequin, G., Ahmadian, Y., Rubin, D.B., Lengyel, M., Miller, K.D. (2016). Stabilized supralinear network dynamics account for stimulus-induced changes of noise variability in the cortex. bioRxiv preprint.
- Hermann, G., & Touboul, J. (2012). Heterogeneous connections induce oscillations in large-scale networks. *Physical Review Letters*, 109(1), 018702.
- Ingber, L. (1992). Generic mesoscopic neural networks based on statistical mechanics of neocortical interactions. *Physical Review A*, 45(4), R2183–R2186.
- Kéfi, S., Dakos, V., Scheffer, M., Van Nes, E.H., Rietkerk, M. (2013). Early warning signals also precede non-catastrophic transitions. *Oikos*, 122, 641–648.
- Krogsgaard-Larsen, P., Honoré, T., Hansen, J.J., Curtis, D.R., Lodge, D. (1980). New class of glutamate agonist structurally related to ibotenic acid. *Nature*, 284, 64–66.
- Kuehn, C. (2013). A mathematical framework for critical transitions: normal forms, variance and applications. *Journal of Nonlinear Science*, 23(3), 457–510.
- Kuznetsov, Y.A. (1998). *Elements of applied bifurcation theory* Vol. 112. New York: Springer.
- Liang, Z., King, J., Zhang, N. (2012). Anticorrelated resting-state functional connectivity in awake rat brain. *NeuroImage*, 59(2), 1190–1199.
- Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nature Reviews Neuroscience*, 5(10), 793–807.
- Moreno, R., de la Rocha, J., Renart, A., Parga, N. (2002). Response of spiking neurons to correlated inputs. *Physical Review Letters*, 89, 288101.
- Moreno-Bote, R. (2014). Poisson-like spiking in circuits with probabilistic synapses. *PLoS Computational Biology*, 10(7), e1003522.
- Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., Pouget, A. (2014). Information-limiting correlations. *Nature Neuroscience*, 17, 1410–1417.
- Pernice, V., Staude, B., Cardanobile, S., Rotter, S. (2011). How structure determines correlations in neuronal networks. *PLoS Computational Biology*, 7(5), e1002059.
- Pillow, J.W., Shlens, J., Paninski, L., Sher, A., Litke, A.M., Chichilnisky, E.J., Simoncelli, E.P. (2008). Spatio-temporal correlations and visual signaling in a complete neuronal population. *Nature*, 454(7206), 995–999.
- Pola, G., Thiele, A., Hoffmann, K.P., Panzeri, S. (2003). An exact method to quantify the information transmitted by different mechanisms of correlational coding. *Network: Computation in Neural Systems*, 14(1), 35–60.
- Ponce-Alvarez, A., He, B.J., Hagmann, P., Deco, G. (2015). Task-driven activity reduces the cortical activity space of the brain experiment and whole-brain modeling. *PLoS Computational Biology*, 11(8), e1004445.
- Renart, A., Brunel, N., Wang, X.-J. (2003). *Mean-field theory of recurrent cortical networks: from irregularly spiking neurons to working memory*. Boca Raton: CRC Press.
- Renart, A., De La Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., Harris, K.D. (2010). The asynchronous state in cortical circuits. *Science*, 327(5965), 587–590.
- Risken, H., & Frank, T. (1996). *The Fokker-Planck equation: methods of solution and applications*. Berlin: Springer.
- Rogers, B.P., Morgan, V.L., Newton, A.T., Gore, J.C. (2007). Assessing functional connectivity in the human brain by fMRI. *Magnetic Resonance Imaging*, 25(10), 1347–1357.
- Sakai, Y., Funahashi, S., Shinomoto, S. (1999). Temporally correlated inputs to leaky integrate-and-fire models can reproduce spiking statistics of cortical neurons. *Neural Networks*, 12(7), 1181–1190.

- Samuelides, M., & Cessac, B. (2007). Random recurrent neural networks dynamics. *The European Physical Journal Special Topics*, 142(1), 89–122.
- Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V., Held, H., Van Nes, E.H., Rietkerk, M., Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, 461(7260), 53–59.
- Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annual Review of Physiology*, 55(1), 349–374.
- Sporns, O. (2006). Small-world connectivity, motif composition, and complexity of fractal neuronal connections. *Bio Systems*, 85(1), 55–64.
- Stewart, I., Elmhirst, T., Cohen, J. (2003). *Bifurcation, symmetry and patterns, chapter symmetry-breaking as an origin of species*, (pp. 3–54). Basel: Birkhäuser.
- Steyn-Ross, M.L., Steyn-Ross, D.A., Sleight, J.W. (2004). Modelling general anaesthesia as a first-order phase transition in the cortex. *Progress in Biophysics and Molecular Biology*, 85(2-3), 369–385.
- Tan, A.Y.Y., Chen, Y., Scholl, B., Seidemann, E., Priebe, N.J. (2014). Sensory stimulation shifts visual cortex from synchronous to asynchronous states. *Nature*, 509(7499), 226–229.
- Tetzlaff, T., Helias, M., Einevoll, G.T., Diesmann, M. (2012). Decorrelation of neural-network activity by inhibitory feedback. *PLoS Computational Biology*, 8(8), e1002596.
- Tononi, G., Sporns, O., Edelman, G.M. (1994). A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences of the United States of America*, 91(11), 5033–5037.
- Touboul, J., Hermann, G., Faugeras, O. (2012). Noise-induced behaviors in neural mean field dynamics. *SIAM Journal on Applied Dynamical Systems*, 11(1), 49–81.
- Trousdale, J., Hu, Y., Shea-Brown, E., Josić, K. (2012). Impact of network structure and cellular response on spike time correlations. *PLoS Computational Biology*, 8(3), e1002408.
- van Vreeswijk, C., & Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Computation*, 10(6), 1321–1371.
- Ward, L.M. (2003). Synchronous neural oscillations and cognitive processes. *Trends in Cognitive Sciences*, 7, 553–559.
- Williams, R.W., & Herrup, K. (1988). The control of neuron number. *Annual Review of Neuroscience*, 11, 423–453.
- Womelsdorf, T., Schoffelen, J.-M., Oostenveld, R., Singer, W., Desimone, R., Engel, A.K., Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. *Science*, 316, 1609–1612.
- Wright, J.J., Rennie, C.J., Lees, G.J., Robinson, P.A., Bourke, P.D., Chapman, C.L., Gordon, E., Rowe, D.L. (2003). Simulated electrocortical activity at microscopic, mesoscopic, and global scales. *Neuropsychopharmacology*, 28(Suppl. 1), S80–S93.