

# Statistical structure of neural spiking under non-Poissonian or other non-white stimulation

Tilo Schwalger<sup>1,2</sup> · Felix Droste<sup>2,3</sup> · Benjamin Lindner<sup>2,3</sup>

Received: 19 November 2014 / Revised: 24 February 2015 / Accepted: 27 March 2015 / Published online: 5 May 2015  
© Springer Science+Business Media New York 2015

**Abstract** Nerve cells in the brain generate sequences of action potentials with a complex statistics. Theoretical attempts to understand this statistics were largely limited to the case of a temporally uncorrelated input (Poissonian shot noise) from the neurons in the surrounding network. However, the stimulation from thousands of other neurons has various sorts of temporal structure. Firstly, input spike trains are temporally correlated because their firing rates can carry complex signals and because of cell-intrinsic properties like neural refractoriness, bursting, or adaptation. Secondly, at the connections between neurons, the synapses, usage-dependent changes in the synaptic weight (short-term plasticity) further shape the correlation structure of the effective input to the cell. From the theoretical side, it is poorly understood how these correlated stimuli, so-called colored noise, affect the spike train statistics. In particular, no standard method exists to solve the associated first-passage-time problem for the interspike-interval statistics with an arbitrarily colored noise. Assuming that input fluctuations are weaker than the mean neuronal drive, we derive simple

formulas for the essential interspike-interval statistics for a canonical model of a tonically firing neuron subjected to arbitrarily correlated input from the network. We verify our theory by numerical simulations for three paradigmatic situations that lead to input correlations: (i) rate-coded naturalistic stimuli in presynaptic spike trains; (ii) presynaptic refractoriness or bursting; (iii) synaptic short-term plasticity. In all cases, we find severe effects on interval statistics. Our results provide a framework for the interpretation of firing statistics measured *in vivo* in the brain.

**Keywords** Interspike-interval statistics · Stochastic integrate-and-fire neuron · Non-renewal process · Temporal correlations · Spontaneous activity

## 1 Introduction

The biophysics of action potential generation in a single nerve cell is well understood in the framework of the celebrated equivalent circuit model by Hodgkin and Huxley (Koch 1999). However, there is a huge difference between the properties of neurons in isolation as described in the Hodgkin-Huxley model and *in vivo*, i.e. embedded in a network of other neurons. In particular, spontaneous neural activity is largely caused by the massive quasi-random input from surrounding cells (Destexhe et al. 2003), which also strongly affects a cell's computational properties (Brunel et al. 2001; London et al. 2010). Hence, when modeling single neuron activity, it is vital to properly account for the characteristics of the spike trains that constitute this input.

Theoretical studies often assume that input spike trains have no temporal structure but are completely random in time with a spiking statistics given by a Poisson process. This assumption has been successfully used to explain and

---

Action Editor: Brent Doiron

✉ Tilo Schwalger  
tilo.schwalger@epfl.ch

<sup>1</sup> Brain Mind Institute,  
École Polytechnique Fédérale de Lausanne (EPFL) Station 15,  
CH-1015, Lausanne, Switzerland

<sup>2</sup> Bernstein Center for Computational Neuroscience,  
Haus 2, Philippstr 13, 10115 Berlin, Germany

<sup>3</sup> Department of Physics, Humboldt Universität zu Berlin,  
Newtonstr 15, 12489, Berlin, Germany

self-consistently determine the irregular, Poisson-like firing patterns of cortical neurons and the emergence of network oscillations (see, e.g. Brunel (2000)). However, in biological neural networks the Poisson assumption is rarely strictly fulfilled (Baddeley et al. 1997). Sources that induce temporal correlations are signal-related processes (Baddeley et al. 1997), neuronal refractoriness (Câteau and Reyes 2006) and bursting (Bair et al. 1994), adaptation (Wang 1998) and network-generated oscillations (Buzsáki and Draguhn 2004). Also the synapse with its conductance dynamics (Koch 1999) as well as short-term synaptic plasticity (Fortune and Rose 2001) contribute to correlate the effective input seen by the postsynaptic cell. As a consequence of both the nonlinear dynamics and the temporally structured driving, the interspike-interval (ISI) statistics of the driven cell is complex (Bair et al. 1994; Baddeley et al. 1997; Compte et al. 2003; Nawrot et al. 2007).

Mathematically, it is convenient to neglect any correlations in the input because it implies that the postsynaptic ISI can be modeled as the first-passage time of a Markov process. For the important class of integrate-and-fire models with white noise (Burkitt 2006) efficient numerical schemes (Richardson 2008) and in simple cases even explicit analytical expressions exist (Gerstein and Mandelbrot 1964) that fit the ISI histograms of some neurons surprisingly well (Gerstein and Mandelbrot 1964; Fisch et al. 2012). The challenge of analyzing neural activity driven by temporally correlated fluctuations, i.e. by a colored noise, is that we need to consider a non-Markovian process, for which no standard techniques exist to compute the first-passage-time (i.e. ISI) statistics. For the special case of a colored noise that can be represented by one or a few stochastic differential equations, researchers have used a Markovian embedding (Hänggi and Jung 1995), i.e. an extension of the phase space by the degrees of the noise dynamics ending with a first-passage-time problem in a higher-dimensional space. Analytical approaches using this trick have been largely limited to the case of exponentially correlated noise (for an exception, see Bauermeister et al. (2013)), that can be mimicked by an Ornstein-Uhlenbeck process (one additional degree of freedom). Using perturbation techniques (small or large correlation time or small noise intensity) approximations have been worked out for the firing rate (Brunel and Sergi 1998; Moreno-Bote and Parga 2004; Alijani and Richardson 2011), the spike train's auto-correlation (Brenner et al. 2002; Moreno-Bote and Parga 2006), the ISI density (Lindner 2004; Schwalger and Schimansky-Geier 2008) and ISI correlations (Lindner 2004), and the dynamical response (Brunel et al. 2001; Alijani and Richardson 2011) and applied to experimental data (Fisch et al. 2012). However, the diverse temporal correlation patterns of in-vivo input mentioned above cannot be captured by an exponentially correlated noise.

To account for the variety of correlation structures, a theory is needed that characterizes the firing statistics for an *arbitrary* input correlation function. In this paper, we put forward a set of explicit formulas that relate the correlation structure of the input noise to the ISI statistics of a perfect integrate-and-fire (PIF) neuron, which is a canonical model for the mean-driven firing regime (Section 2.1). The analysis is based on the assumption that the correlated noise input can be represented by a multi-dimensional Ornstein-Uhlenbeck process (Markovian embedding) and that the noise is weak compared to the mean driving current. This allows us to formulate the first-passage-time problem in terms of a multi-dimensional Fokker-Planck equation, which is solved for arbitrary dimensions of the Ornstein-Uhlenbeck process using weak noise approximation techniques. The details of the rather lengthy calculations are presented in the Appendix A. In Section 2.2, we apply our general results to generic situations of correlated input and compare the ISI statistics to the commonly studied case of uncorrelated Poisson inputs. In Section 2.3, we investigate whether our findings for the PIF model also hold for more realistic nonlinear integrate-and-fire models. For the suprathreshold firing regime and weak noise, we use the phase response curve to derive expressions that generalize some of the results for the PIF model. For the exponential integrate-and-fire model, we use these analytical expressions as well as simulation results for the subthreshold regime to show that similar effects of non-Poissonian inputs also occur under more general conditions. We finally discuss further applications of the theory (Section 3).

## 2 Results

### 2.1 Theoretical framework and analytical results

We consider a perfect integrate-and-fire (PIF) neuron, which represents a reasonable approximation of neural spike generation in the mean-driven firing regime. The membrane potential  $V(t)$  obeys

$$C_m \dot{V} = I_0 + I_{\text{syn}}(t), \quad \text{if } V = V_T : \quad V \rightarrow 0 \text{ mV}. \quad (1)$$

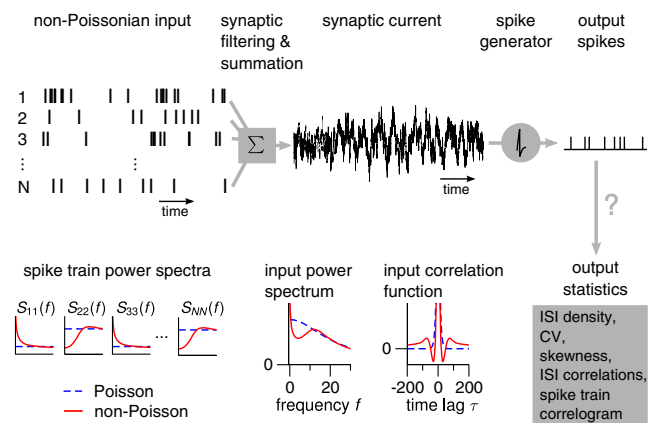
where  $C_m$  is the membrane capacity and  $I_0$  is a constant membrane current. Instances at which  $V(t)$  reaches the threshold  $V_T = 15$  mV and is reset to  $V = 0$  mV define the spike times  $\{t_i\}$  and thus a sequence of ISIs  $T_i = t_i - t_{i-1}$ ,  $i = 1, 2, \dots$ , as well as the output spike train  $x_{\text{out}}(t) = \sum_i \delta(t - t_i)$ . The synaptic input current is modeled as  $I_{\text{syn}}(t) = \sum_{\ell=1}^N (j_\ell * x_{\text{in},\ell})(t)$ , where  $x_{\text{in},\ell}(t) = \sum_j \delta(t - t_j^{(\ell)})$ ,  $\ell = 1, \dots, N$ , is the input spike

train of the  $\ell^{th}$  presynaptic neuron with spike times  $t_j^{(\ell)}$ . The convolution  $(j_\ell * x_{in,\ell})(t) = \int dt' j_\ell(t')x_{in,\ell}(t - t')$  of the spike train with the post-synaptic current  $j_\ell(t)$  approximates the synaptic filtering by the conductance dynamics of synapse  $\ell$ . The statistics of stationary input spike trains can be characterized by the input firing rates  $\nu_\ell = \langle x_{in,\ell}(t) \rangle$  and the input covariance matrix  $C_{k\ell}(\tau) = \langle x_{in,k}(t)x_{in,\ell}(t + \tau) \rangle - \nu_k\nu_\ell$ , or equivalently, by the spectral statistics  $S_{k\ell}(f) = \int d\tau e^{2\pi if\tau} C_{k\ell}(\tau)$ . For the case of the commonly used Poisson statistics, the spike train power spectrum is simply a constant equal to  $S_{\ell\ell}(f) = \nu_\ell$ . In general, however, a spike train has a frequency-dependent power spectrum reflecting the presence of temporal correlations (Fig. 1, left).

If the number of inputs is large,  $I_{syn}(t)$  is approximately Gaussian with an auto-correlation function that is shaped by both the spatio-temporal correlations  $C_{k\ell}(\tau)$  of presynaptic spike trains (Câteau and Reyes 2006; Lindner 2006) and the synaptic filter (Fig. 1, middle). In this Gaussian approximation, the dynamics Eq. (1) can be cast into the form  $\dot{V} = \mu + \sigma\eta(t)$ , where  $\mu$  subsumes the constant parts of the input,  $\sigma^2$  is the variance of the synaptic drive and  $\eta(t)$  is a zero-mean, unit-variance Gaussian process with correlation function  $C_{in}(\tau)$  and power spectrum  $S_{in}(f) = \int dt e^{2\pi ift} C_{in}(t)$ , cf. Appendix A.1. In the case of a flat power spectrum (white noise), the model generates a renewal spike train with an inverse Gaussian ISI statistics. In contrast, we are concerned here with the case of a correlated (colored) noise  $\eta(t)$  with an arbitrary correlation function and a non-flat power spectrum leading to a nonrenewal spike train, i.e. the ISI sequence displays serial correlations. One can show that regardless of the input's correlation structure, the output firing rate of our simple model is always  $r = \mu/V_T$  (Bauermeister et al. 2013). However, the ISI probability density and its higher moments as well as the serial correlations among ISIs are strongly shaped by input correlations.

### 2.1.1 Markovian embedding

The voltage dynamics  $V(t)$  is non-Markovian because  $\eta(t)$  is correlated in time. Apart from a few special cases (e.g. Salinas and Sejnowski (2002), Lindner (2004), and Droste and Lindner (2014)), exact solutions of the associated FPT problem are not known for non-Markovian processes. To obtain approximations of the FPT statistics, we assume here that the noise is weak as expressed by the small parameter  $\epsilon = \sigma/\mu \ll 1$ . In particular, this excludes the singular case of purely white noise, for which the variance diverges. Furthermore, we assume that there exists a (possibly high-dimensional) Markovian embedding (Hänggi and Jung 1995) such that  $\eta(t)$  can be represented as a projection of a multivariate Ornstein-Uhlenbeck process



**Fig. 1** Illustration of the colored-noise problem for neurons receiving massive spike train input through  $N$  synapses. Presynaptic spike trains have temporal structure, e.g. due to refractoriness, bursting, or rate modulations, which is expressed by non-flat power spectra  $S_{kk}(f)$ ,  $k = 1, \dots, N$ , for each single spike train (red lines). This is contrasted to a Poisson statistics corresponding to flat (constant) Poisson spectra (blue dashed line). Each spike train is filtered by the synaptic conductance dynamics and subsequently summed to produce a total input current that drives the neuron. The input current is approximately Gaussian with a power spectrum (or auto-correlation function) formed by single spike train statistics, cross-correlations and synaptic filter. The goal of this study is to link this input statistics to the spiking statistics of the postsynaptic neuron (output statistics)

(OUP):

$$\eta(t) = \mathbf{b}^T \mathbf{Y}(t), \quad \dot{\mathbf{Y}} = \mathbf{A} \mathbf{Y} + \mathbf{B} \boldsymbol{\xi}(t). \tag{2}$$

Here,  $\mathbf{b}$  is a constant projection vector,  $\mathbf{Y}$  is a  $d$ -dimensional OUP,  $\boldsymbol{\xi}(t)$  is a vector of Gaussian white noise processes obeying  $\langle \xi_i(t)\xi_j(t') \rangle = \delta_{i,j}\delta(t - t')$ , and  $\mathbf{A}$  and  $\mathbf{B}$  are constant matrices determining the drift and diffusion strength of the OUP, respectively. To ensure a finite variance of the OUP, we require that the eigenvalues of  $\mathbf{A}$  have negative real parts.

The Markovian embedding is possible if the correlation function of the noise can be approximated by a sum of exponential functions and damped harmonic oscillations with decay rates and oscillation frequencies given by the (complex) eigenvalues of  $\mathbf{A}$ . In the frequency domain, this amounts to approximating the power spectrum of the noise by a sum of Lorentzian functions. Thus, by allowing arbitrarily large dimensions of the OUP, the class of attainable correlation functions or power spectra is large. In particular, many biologically relevant correlation functions may be approximated by a sufficiently large number of exponentials, possibly with complex-valued rates. In Section 2.2, we show that even power spectra with an algebraic decay proportional to  $1/f^\gamma$  are captured by our approach because such noise arises from infinitely many exponentially correlated processes with a continuum of correlation times (see e.g. Sobie et al. (2011)).

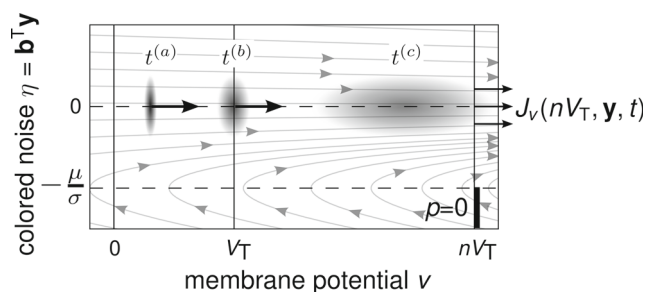
Explicitly, the correlation function and the power spectrum corresponding to Eq. (2) are given by  $C_{in}(\tau) = \mathbf{b}^T \Sigma e^{|\tau|A^T} \mathbf{b}$  and  $S_{in}(f) = -2\mathbf{b}^T [A^2 + (2\pi f)^2 \mathbb{I}_d]^{-1} A \Sigma \mathbf{b}$ , respectively (Risken 1984). Here,  $e^{\tau A}$  is the matrix exponential function,  $\mathbb{I}_d$  is the identity matrix and  $\Sigma$  is the covariance matrix of  $\mathbf{Y}$ . Its elements  $\Sigma_{ij} = \langle Y_i Y_j \rangle$  can be computed from Eq. (34) given in the Appendix A.2.

Fortunately, the inverse problem, i.e. finding matrices A and B that yield a desired correlation function  $C_{in}(\tau)$ , does not need to be solved. Indeed, the Markovian representation (2) is only an intermediate step that allows us to formulate the evolution of the system by the multi-variable Fokker-Planck equation (FPE) (Risken 1984)

$$\frac{\partial p}{\partial t} = -(\mu + \sigma \mathbf{b}^T \mathbf{y}) \frac{\partial p}{\partial v} - \sum_{i,j=1}^d \left[ A_{ij} \frac{\partial}{\partial y_i} (y_j p) - D_{ij} \frac{\partial^2 p}{\partial y_i \partial y_j} \right]. \tag{3}$$

Here,  $p(v, \mathbf{y}, t)$  is the probability density for the joint process  $[V(t), \mathbf{Y}(t)]$  (Fig. 2) and  $D = \frac{1}{2} \mathbf{B}^T \mathbf{B}$  is the diffusion matrix. Remarkably, we will see that the final expressions for the ISI statistics can be re-expressed solely in terms of the correlation function  $C_{in}(\tau)$ . Thus, an explicit construction of the OUP, i.e. of the Markovian embedding Eq. (2), is not necessary for the application of our theory.

In order to calculate the statistics of the stationary ISI sequence  $\{T_i\}$ , we shift the time origin to the zeroth spike,  $t_0 = 0$ . The statistics of  $t_n = \sum_{i=1}^n T_i$  is given by the  $n^{th}$ -order interval density  $P_n(t)$ . Because the right hand side of Eq. (1) does not depend on  $V$ , the time of the  $n$ -th spike is equivalent to the first-passage time from  $V = 0$  to the  $n$ -fold threshold  $V = nV_T$  (Middleton et al. 2003; Lindner 2004). For an ensemble of trajectories with  $V(0) = 0$ , the first-passage-time density is equal to the total probability



**Fig. 2** Schematic of the probability flow in the space  $(v, \mathbf{y})$ . Initially, the probability density  $p(v, \mathbf{y}, 0)$  is concentrated at  $v = 0$  and proportional to the distribution upon firing  $p_{\text{spike}}(\mathbf{y})$ . Snapshots of  $p(v, \mathbf{y}, t)$  are schematically depicted at three different times  $(0 < t^{(a)} < t^{(b)} < t^{(c)})$ . The density moves towards the threshold for  $t > 0$  and is not reset upon crossing the hyperplane  $v = V_T$ . The p.d.f. of the  $n$ th spike time,  $P_n(t)$ , is equal to the total flux  $\int d^d y J_v(nV_T, \mathbf{y}, t)$  through the hyperplane  $v = nV_T$ . The boundary condition that prevents recrossings with negative velocity  $\mu + \sigma \mathbf{b}^T \mathbf{y} < 0$  (region below dashed line) can be neglected for  $\sigma/\mu \ll 1$

flux at time  $t$  across the threshold (Fig. 2) if trajectories that have reached the threshold are removed from the ensemble. Thus,  $P_n(t) = \int d^d y J_v(nV_T, \mathbf{y}, t)$ , where  $J_v(v, \mathbf{y}, t) = (\mu + \sigma \mathbf{b} \cdot \mathbf{y}) p(v, \mathbf{y}, t)$  is the  $v$ -component of the probability flux. Besides the natural boundary conditions  $[p(v, \mathbf{y}, t) = 0 \text{ for } v \rightarrow -\infty \text{ and } y_i \rightarrow \pm\infty]$ , the removal of trajectories can be realized by requiring that  $p(nV_T, \mathbf{y}, t) = 0$  for all  $\mathbf{y}$  for which  $\mu + \sigma \mathbf{b} \cdot \mathbf{y} < 0$ . The latter ensures that there is no re-flux of probability into the domain  $v < nV_T$  from above threshold (cf. Brunel and Sergi (1998)). However, for weak noise,  $\sigma/\mu \ll 1$ , negative velocities  $\dot{V} = \mu + \sigma \mathbf{b} \cdot \mathbf{Y}$  are highly unlikely, so that this boundary condition can be safely ignored. The initial condition corresponds to the state of the ensemble conditioned on a spike at  $t = 0$ . Hence,  $p(v, \mathbf{y}, 0) = \delta(v) P_{\text{spike}}(\mathbf{y})$ , where  $P_{\text{spike}}(\mathbf{y}) = (1 + \epsilon \mathbf{b} \cdot \mathbf{y}) p_s(\mathbf{y})$  is the stationary density of the OUP upon firing and  $p_s(\mathbf{y}) = \exp\left(-\frac{1}{2} \mathbf{y}^T \Sigma^{-1} \mathbf{y}\right) / \sqrt{(2\pi)^d |\Sigma|}$  is the stationary density of the OUP at an arbitrary time.

### 2.1.2 ISI density and spike auto-correlations

As shown in Appendix A.2, the  $n^{th}$ -order interval density  $P_n(t)$  can be re-expressed in terms of the characteristic function  $q(\ell, \mathbf{k}, t) = \int d\mathbf{v} e^{i\ell v} \int d^d y e^{i\mathbf{k} \cdot \mathbf{y}} p(v, \mathbf{y}, t)$  instead of  $p(v, \mathbf{y}, t)$ . This function obeys a tractable first-order differential equation (Fourier-transformed FPE) (see Appendix A.2). In the solution, the correlation function  $C_{in}(\tau) = \mathbf{b}^T \Sigma e^{|\tau|A^T} \mathbf{b}$  can be identified and re-substituted. The result is

$$P_n(t) = \frac{r}{2\sqrt{4\pi\epsilon^2 h^3(t)}} \exp\left[-\frac{(rt-n)^2}{4\epsilon^2 h(t)}\right] \times \left\{ \frac{[(n-rt)g(t) + 2h(t)]^2}{2h(t)} - \epsilon^2 [g^2(t) - 2h(t)C_{in}(t)] \right\},$$

$$g(t) = r \int_0^t dt' C_{in}(t'), \quad h(t) = r \int_0^t dt' g(t'). \tag{4}$$

For  $n = 1$  this formula yields the ISI density  $P_1(t)$ . Our result includes the known special cases for exponentially correlated noise (Lindner 2004) and narrow-band noise (Bauermeister et al. 2013). It also allows to relate the output correlation function  $C_{out}(\tau) = \langle S(t)S(t+\tau) \rangle - r^2$  to the input correlation function  $C_{in}(\tau)$  by the formula (Cox and Lewis 1966a)

$$C_{out}(\tau) = r \left( \delta(\tau) + \sum_{n=1}^{\infty} P_n(|\tau|) - r \right). \tag{5}$$

In Figs. 4D and 5C, below, we truncated the sum at  $n = 4$ .



### 2.1.3 Cumulants and correlations of ISIs

Using the moment-generating function  $\bar{P}_n(s) = \int_0^\infty dt e^{-st} P_n(t)$ , the  $k$ -th cumulant of  $t_n$  is given by  $\kappa_{k,n} = (-1)^k d^k \ln \bar{P}_n/ds^k|_{s=0}$  (Risken 1984). The moment generating function can be rewritten as  $\bar{P}_n(s) = \mu (1 - i\mathbf{e}^T \nabla_{\mathbf{k}}) \varphi(nV_T, \mathbf{k}, s)|_{\mathbf{k}=0}$ , where  $\varphi(v, \mathbf{k}, s) = \int_0^\infty dt e^{-st} \int d^d y e^{i\mathbf{k}\cdot\mathbf{y}} p(v, \mathbf{y}, t)$  obeys a transformed FPE, which still contains mixed derivatives. To solve this equation for weak noise, we use the perturbation expansion  $\varphi = \sum_{k=0}^\infty \varphi^{(k)} \epsilon^k$ . This yields a hierarchy of first-order equations for the coefficients  $\varphi^{(k)}$ , which can be solved iteratively (see Appendix A.2). Up to a correction of order  $\epsilon^6$ , the second and third cumulant of the  $n$ th-order interval are obtained as

$$\kappa_{2,n} \approx \hat{\kappa}_{2,n} = 2r^{-2}\epsilon^2 \left\{ h_n + \epsilon^2 \left[ g_n^2 + C_{in}(n/r)h_n \right] \right\} \quad (6)$$

$$\kappa_{3,n} \approx 12r^{-3}\epsilon^4 g_n h_n, \quad (7)$$

using the shorthand  $g_n = g(n/r)$  and  $h_n = h(n/r)$ .

To quantify the variability of ISIs, we consider the (squared) coefficient of variation

$$C_V^2 = r^2 \hat{\kappa}_{2,1} + \mathcal{O}(\epsilon^6) \quad (8)$$

$$\approx 2\epsilon^2 h_1 = \frac{\sigma^2}{r^2 V_T^2} \int df S_{in}(f) \text{sinc}^2(f/r) \quad (9)$$

where  $\text{sinc}(x) = \sin(\pi x)/(\pi x)$ . This shows that the variance of ISIs is determined by the total power of the sinc-filtered input noise. Hence the color of the noise (i.e. the form of  $S_{in}(f)$ ) and the output firing rate  $r = 1/\langle T \rangle$  (affecting the width of the filter but also the prefactor) directly influence the CV.

Particular effects of input correlations become apparent in higher-order cumulants of the ISI density. Convenient is the comparison to the cumulants of the inverse Gaussian density, which is the exact solution for pure uncorrelated (white noise) input (Gerstein and Mandelbrot 1964). Here we consider a rescaled version of the skewness involving the third cumulant:  $\alpha_s = \langle T \rangle \kappa_{3,1} / (3\kappa_{2,1}^2)$ . This quantity is equal to unity for any inverse Gaussian ISI density and is smaller or larger unity for a less or more skewed distribution, respectively (Schwalger et al. 2010). For weak noise, we find from Eqs. (6) and (7)

$$\alpha_s \approx \frac{g(\langle T \rangle)}{h(\langle T \rangle)} = \frac{2 \int df S_{in}(f) \text{sinc}(2f/r)}{\int df S_{in}(f) \text{sinc}^2(f/r)}. \quad (10)$$

That is, the skewness is the ratio of the areas under the functions  $2S_{in}(f)\text{sinc}(2f/r)$  and  $S_{in}(f)\text{sinc}^2(f/r)$ . For a constant power spectrum, these areas are equal yielding  $\alpha_s = 1$  as expected for a white-noise-driven neuron. A less or more skewed ISI density can be obtained by weighting the minima or maxima of  $\text{sinc}(2f/r)$  by the input spectrum, respectively.

Correlations between ISIs can be quantified by the serial correlation coefficient (SCC)  $\rho_n = (\langle T_i T_{i+n} \rangle - \langle T_i \rangle^2) / ((\langle T_i^2 \rangle - \langle T_i \rangle^2))$ , where  $n$  is the lag between ISIs. The SCC is given by the second-order cumulants (Lindner 2004), yielding

$$\rho_n = \frac{\kappa_{2,n+1} - 2\kappa_{2,n} + \kappa_{2,n-1}}{2\kappa_{2,1}} \quad (11)$$

$$= \frac{\hat{\kappa}_{2,n+1} - 2\hat{\kappa}_{2,n} + \hat{\kappa}_{2,n-1}}{2\hat{\kappa}_{2,1}} + \mathcal{O}(\epsilon^4) \quad (12)$$

$$= \frac{\int df S_{in}(f) \text{sinc}^2(f\langle T \rangle) e^{-2\pi i f \langle T \rangle n}}{\int df S_{in}(f) \text{sinc}^2(f\langle T \rangle)} + \mathcal{O}(\epsilon^2). \quad (13)$$

The leading-order approximation (third line) shows that the SCC is proportional to the inverse Fourier transform of the filtered spectrum  $S_{in}(f)\text{sinc}^2(f\langle T \rangle)$  evaluated at discrete time lags  $t = n\langle T \rangle$ . This is equivalent to the correlation function of the input noise smoothed by a moving average of window size  $\langle T \rangle$ .

As shown in the Appendix A.2, the leading-order approximations for the CV and the SCC, Eqs. (9) and (13), are consistent with exact analytical results for the variability of the spike count on large time scales, which have been obtained previously (Middleton et al. 2003; Sobie et al. 2011; Moreno-Bote et al. 2014). Indeed, the Fano factor  $F(t)$ , defined as the variance to mean ratio of the spike count in a time window of length  $t$ , can be approximated for general  $t > 0$  as

$$F(t) = [\{t\}(1 - \{t\}/\langle T \rangle) + 2\epsilon^2 \langle T \rangle h(t)]/t. \quad (14)$$

Here,  $\{t\} = t - \langle T \rangle \lfloor t/\langle T \rangle \rfloor$  denotes the difference between  $t$  and the largest multiple of  $\langle T \rangle$  equal or less than  $t$ . The long-time limit of Eq. (14) is exact and yields  $\lim_{t \rightarrow \infty} F(t) =$

$C_V^2 \left( 1 + 2 \sum_{n=1}^\infty \rho_n \right)$ , which recovers a well-known identity for spike trains (Cox and Lewis 1966b).

## 2.2 Applications

By means of our theory, we now discuss the spiking statistics for three important mechanisms of temporal input correlations. We consider an excitatory and an inhibitory population of  $N_E = 800$  and  $N_I = 200$  presynaptic neurons, respectively. Each input spike generates an exponential postsynaptic current of the form  $j_k(t) = J_k e^{-t/\tau_k} \theta(t)$ ,  $k = E, I$  with synaptic time constants  $\tau_E = 4$  ms and  $\tau_I = 8$  ms ( $\theta(t)$  denotes the Heaviside step function). In the case of dynamic synapses, the  $\ell$ th synapse has its own time-dependent coefficient  $J_{k,\ell}(t) = \bar{J}_k F_\ell(t) D_\ell(t)$  with a constant prefactor  $\bar{J}_k$  and facilitation and depression variables (see Appendix A.6); for static synapses  $J_{k,\ell}(t) = \bar{J}_k$ . If not stated otherwise, we use the base current  $I_0 = 0.02$  (here and in the following, we state all current amplitudes

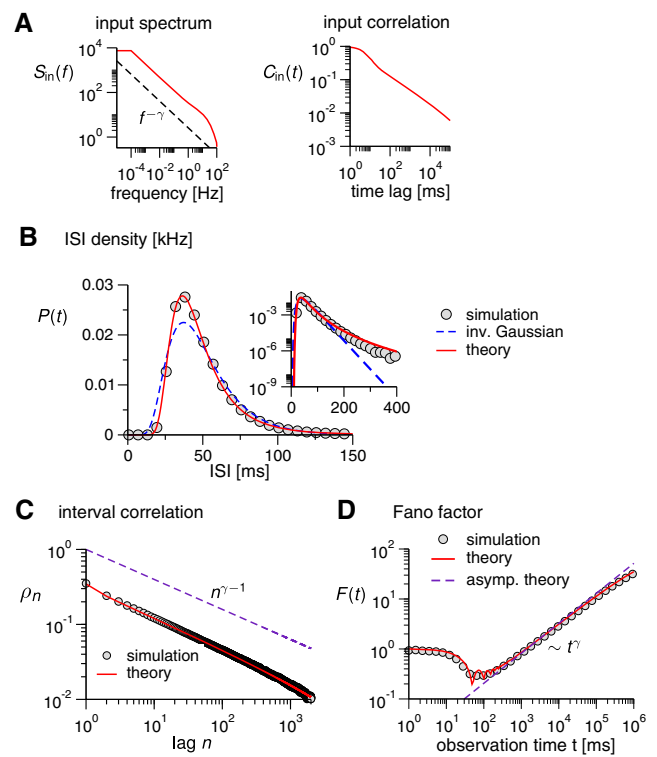
and synaptic weights  $J_k$  in units of  $C_m V_T/\text{ms}$ ). Note that the synaptic filter alone leads to a colored noise input even if input spike trains are uncorrelated (Poissonian) and short-term plasticity is absent. Most previous studies of colored noise in neurons (Brunel and Sergi 1998; Fourcaud and Brunel 2002; Moreno-Bote and Parga 2006; Lindner 2004; Schwalger and Schimansky-Geier 2008) have investigated exactly this case of Poissonian input spike trains filtered by the conductance dynamics of static synapses. In contrast, we focus in the following on additional sources of input correlations that have a more drastic effect on postsynaptic ISI statistics.

### 2.2.1 Gaussian stimuli with complex temporal statistics

A number of studies have reported membrane potential fluctuations that indicate synaptic input with a power law spectrum (Pozzorini et al. 2013; Destexhe et al. 2003). Direct injection of power-law currents have also been used to mimic the statistics of natural stimuli (Wang et al. 2003). How does such a powerlaw statistics shape the spiking statistics of a neuron? To answer this question, we consider Poissonian input with a common rate modulation (signal). Specifically, the firing rate of the excitatory input streams is  $r_{\mathcal{E}}(t) = \nu_{\mathcal{E}}(1 + \varepsilon s(t))$ , where  $s(t)$  is a unit-variance Gaussian process with power spectrum

$$S_s(f) = \mathcal{N} \begin{cases} f_1^{-\gamma}, & |f| \leq f_1 \\ f^{-\gamma}, & f_1 < |f| < f_2 \\ 0, & |f| \geq f_2 \end{cases}, \quad \gamma \in (0, 1). \quad (15)$$

Here,  $\mathcal{N}$  is a normalization constant (see Appendix A.4). The power spectrum exhibits a power-law decay with exponent  $\gamma$  in the range between  $f_1$  and  $f_2$ , which implies long-range correlations that decay like  $t^{\gamma-1}$ . A standard calculation yields the total input spectrum  $S_{\text{in}}(f)$  (see Appendix A.4), which is illustrated in Fig. 3A. Although, strictly speaking, such a scale-free spectrum cannot be achieved by a finite-dimensional OUP, the input spectrum can nevertheless be used in our final formulas, provided the variance is sufficiently small. Figures 3B–D show good agreement between theory (red lines) and stochastic simulations (circles). The ISI density is significantly more skewed than an inverse Gaussian density with the same mean and variance (corresponding to white noise input), as observed previously for positive input correlations (Schwalger et al. 2010). The signature of the powerlaw correlations in the input is apparent in the serial correlations of ISIs that decay like  $n^{\gamma-1}$ . Likewise, the Fano factor increases algebraically like  $t^{\gamma}$ , resembling “fractal” behavior found in auditory neurons (Lowen and Teich 1992; Peterson et al. 2014). This



**Fig. 3** Firing statistics in response to a complex signal modelled as a common rate modulation of excitatory presynaptic neurons. **(A)** The modulation has a power-law spectrum that scales like  $1/f^\gamma$ . **(B)** The ISI density is more peaked (simulation – circles; theory – red solid line) than an inverse Gaussian density with the same mean and variance (dashed blue line). **(C)** The serial correlation coefficients (red line – Eq. (11)) reveal long-range correlations, which decay slowly like  $n^{\gamma-1}$  (dashed blue line). **(D)** The Fano factor (red line – Eq. (14)) increases like  $t^\gamma$  for an intermediate range of observation times  $t$ , during which spikes are counted (dashed line – Eq. (139)). Parameters:  $\gamma = 0.6$ ,  $J_{\mathcal{E}} = 0.0015$ ,  $J_{\mathcal{I}} = -0.003$ ,  $\bar{\nu}_{\mathcal{E}} = \bar{\nu}_{\mathcal{I}} = 5$  Hz,  $\varepsilon = 0.5$ ,  $f_1 = 10^{-4}$  Hz,  $f_2 = 100$  Hz (corresponding to  $\varepsilon = \sigma/\mu = 0.62$ )

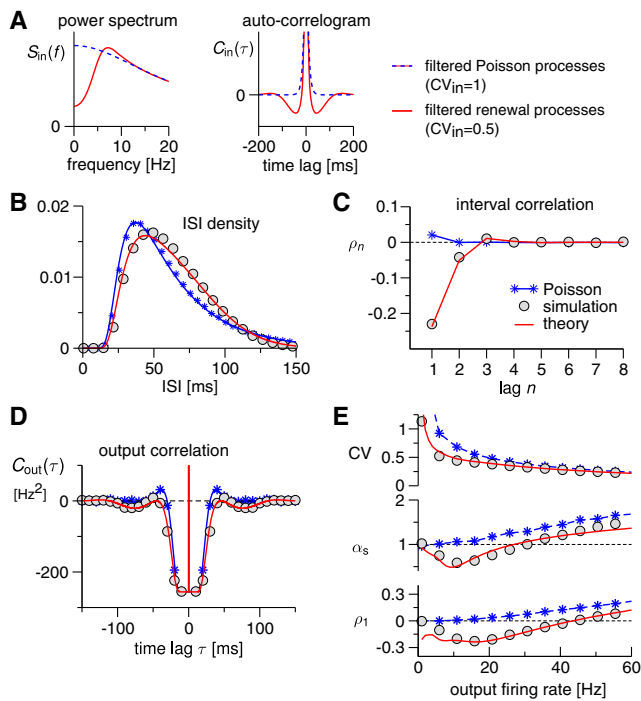
scaling behavior can be understood by an asymptotic analysis of Eqs. (13) and (14) (see Appendix A.4).

### 2.2.2 Presynaptic refractoriness

Even in the absence of correlated stimuli, neuronal spiking is correlated in time. In contrast to Poisson statistics, neurons cannot exhibit arbitrarily small ISIs due to a refractory period after each spike. This causes correlations between subsequent spikes. A simple description of refractoriness is provided by a renewal point process with reduced probability for short ISIs (Gerstner and Kistler 2002). Such processes display often a reduced variability ( $C_{V,\text{in}} < 1$ ), indicating a spike train that is more regular than a Poisson process ( $C_{V,\text{in}} = 1$ ). In the following, we model presynaptic input spike trains by renewal processes with an inverse Gaussian ISI density (Gerstein and Mandelbrot 1964; Cox and Miller 1965), for which the exact form of the

input power spectrum can be inferred (see Appendix A.5). This process can be parametrized by input firing rate and input CV allowing us to continuously vary the intensity and regularity of the input.

For regular presynaptic spike trains ( $C_{V, in} < 1$ ), the spectrum of the total synaptic input current exhibits a dip at low frequencies and a maximum near the firing rate of presynaptic neurons (Fig. 4A) – effects of the refractory period that are seen in a large fraction of cortical cells (Bair et al. 1994; Franklin and Bair 1995; Compte et al. 2003). In the corresponding input correlation function the refractory period becomes manifest by a negative trough (Fig. 4A). Such input, as predicted by our theory, yields a less skewed ISI density than a Poisson-driven neuron (Fig. 4B). Although the spike correlations in the output does not differ drastically from the case of Poisson driving (Fig. 4D), correlations among intervals display pronounced short-term negative correlations (Fig. 4C) – a feature commonly observed for neurons with adaptation currents driven by fast (uncorrelated) noise (Chacron et al. 2000; Liu and Wang 2001;



**Fig. 4** Spiking statistics in response to presynaptic input that is more regular than Poisson ( $C_{V, in} = 0.5$ ), modeling input refractoriness. (A) Neuron driven by 800 excitatory and 200 inhibitory neurons (inverse Gaussian renewal processes with rate  $\nu_E = \nu_I = 5\text{ Hz}$  and  $C_{V, in} = 0.5$ ). Summed input has a correlation function and power spectrum as indicated by solid red lines. For comparison: synaptically filtered Poisson input (blue line and stars). (B) ISI density for  $C_{V, in} = 0.5$  (red) is less skewed compared to Poisson input (blue). (C) SCCs are markedly different (theory, solid lines; simulation, symbols). Parameters:  $I_0 = 0.016$ ,  $J_E = 0.004$ ,  $J_I = -0.008$  (corresponding to  $\epsilon = 1.18$ ) (E) With increasing output firing rate (controlled by  $I_0$ ), the CV decreases, the SCC  $\rho_1$  changes sign and the ISI density changes from less ( $\alpha_s < 1$ ) to more skewed ( $\alpha_s > 1$ )

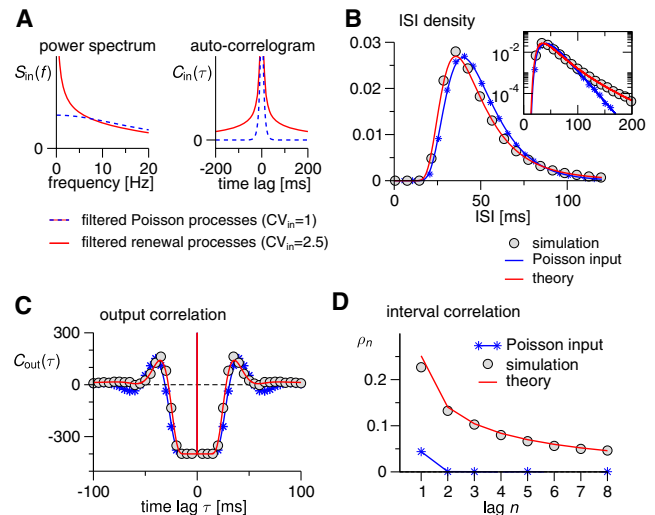
Schwalger et al. 2010; Schwalger and Lindner 2013). However, here it is entirely due to an ubiquitous property of presynaptic spike trains, namely, a refractory period.

Effects of refractoriness on the input side are restricted to small to moderate postsynaptic firing rates. At high output rates the synaptic low-pass filter overshadows refractory effects and induces positive correlations over multiple ISIs. Accordingly, our theory predicts a transition from negative to positive serial correlations and from less to more skewed ISI distributions (Fig. 4E).

2.2.3 Presynaptic burstiness

Another important source of temporal correlations is bursting activity, which is a widespread phenomenon in the nervous system. Bursting leads to spike patterns that are more irregular than a Poisson process because of the presence of short intraburst intervals and long interburst intervals. We mimic burstiness by a superposition of renewal processes with  $C_{V, in} > 1$ .

For irregular spike input (here  $C_{V, in} = 2.5$ ), the input spectrum has increased low-frequency power as compared to the Poisson case (Fig. 5A, red vs. blue line). Such power spectra are observed in cortical cells (Bair et al. 1994; Baddeley et al. 1997; Compte et al. 2003). The burst-like input leads to extended positive input correlations, which, similar to the case of positive power-law correlations, result in a more skewed ISI distribution and pronounced



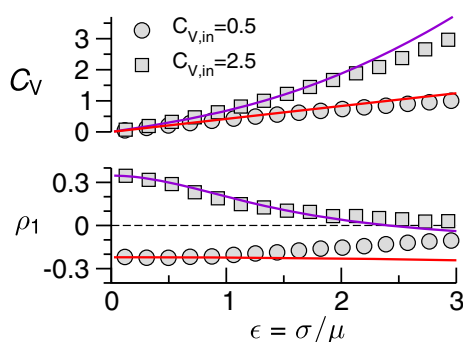
**Fig. 5** Spiking statistics in response to presynaptic input that is more irregular than Poisson ( $C_{V, in} = 2.5$ ), modeling burstiness. (A) Input exhibits increased power at low frequencies and positive correlations (red lines). While the ISI density (B) and the spike train auto-correlogram  $C_{out}(\tau)$  (C) show only slight differences compared to a Poisson-driven neuron, the ISI sequence shows pronounced positive correlations that are largely absent in Poisson-driven case (D). Parameters:  $J_E = 0.003$ ,  $J_I = -0.006$  (corresponding to  $\epsilon = 0.75$ )

positive serial correlations (Fig. 5B,D); effects of this are not pronounced at the level of the spike train output correlation function (Fig. 5C). In contrast, ISI correlations in the presence of Poisson input are very small — they are still present due to synaptic filtering.

Our approximations for the moments of the ISI statistics are based on a Gaussian approximation of the input shot-noise and on a perturbation expansion for weak noise ( $\epsilon \ll 1$ ). To test the range of validity of the perturbation theory, we compared the theory for CV and ISI correlations with simulations for different values of the perturbation parameter  $\epsilon = \sigma/\mu$  (Fig. 6). Remarkably, we find excellent agreement even beyond  $\epsilon = 1$ . As expected, the perturbation theory becomes worse for even larger input variability ( $\epsilon \gtrsim 2$ ).

### 2.2.4 Neurons with synaptic short-term plasticity

It is known that input spectra can be changed by synaptic short-term dynamics (Lindner et al. 2009; Merkel and Lindner 2010; Rosenbaum et al. 2012; Droste et al. 2013). In particular, facilitating (depressing) synapses lead to a low-pass (high-pass) filtering of the input power. Even if we assume completely uncorrelated presynaptic spike trains, we thus expect synaptic dynamics to have similar effects as the presynaptic refractoriness or bursting discussed above. As an example for the difference in postsynaptic spiking statistics, we inspect the interval correlations that are induced by dynamic synapses. To this end, we use in our standard setup synapses endowed with a facilitation (depression) dynamics, using an established model of short-term plasticity (Dittman et al. 2000; Merkel and Lindner 2010). Doing so, we find that for Poissonian spike trains impinging on facilitating synapses, the postsynaptic cell shows marked



**Fig. 6** Range of validity of the weak-noise perturbation theory. Coefficient of variation and ISI correlation coefficient  $\rho_1$  for  $N_E = 800$  excitatory and  $N_I = 200$  inhibitory input spike trains are shown as a function of the weak-noise parameter  $\epsilon = \sigma/\mu$ . Red (violet) lines and circles (squares) correspond to theory and simulations with regular,  $C_{V,\text{in}} = 0.5$ , (irregular,  $C_{V,\text{in}} = 2.5$ ) presynaptic input, respectively. Further parameters are the same as in Fig. 5 except for the synaptic weights  $J_E = -J_I/2$ , which have been varied in the range  $J_E \in [0.0005, 0.0125]$  for  $C_{V,\text{in}} = 0.5$  and  $J_E \in [0.0005, 0.0117]$  for  $C_{V,\text{in}} = 2.5$  to obtain different noise strength  $\epsilon$

positive ISI correlations (cf. Fig. 7B, red line and circles) due to the low-pass filtering of the input. Conversely, depressing synapses cause pronounced negative correlations (Fig. 7B, green line and squares) because depression reduces input power at low frequencies. Our theory (lines in Fig. 7B) is in excellent agreement with the simulations (symbols).

### 2.3 Beyond the perfect integrate-and-fire model

A crucial assumption that permitted a systematic perturbation solution was the voltage-independence of the membrane current  $I_0$ . This is justified if the mean drive is so strong that it dominates any variations of the membrane current due to voltage changes. Here, we briefly discuss the significance of our results if this assumption is relaxed; for a similar analysis for white-noise-driven IF neurons and neurons endowed with an adaptation current, see Schwalger et al. (2013). We extend the subthreshold dynamics by a voltage-dependent term:

$$\dot{V} = \gamma_L \Psi(V) + \mu + \sigma \eta(t). \quad (16)$$

Spikes are registered when  $V(t)$  reaches the value  $V_{\text{spike}} \geq V_T$ , whereupon  $V$  is reset to zero. The nonlinear IF model Eq. (16) is a more general description of neural dynamics that can accurately model passive membrane properties (leak) and spike initiation dynamics. In particular, the exponential integrate-and-fire (EIF) model (Fourcaud-Trocmé et al. 2003) given by the choice  $\Psi(V) = -V + \Delta_T \exp\left(\frac{V-V_T}{\Delta_T}\right)$  fits neural data of many neurons remarkably well (Badel et al. 2008).

The model Eq. (16) can operate in two different dynamical regimes: Firstly, if  $\gamma_L \Psi(V) + \mu > 0$  for all  $V < V_{\text{spike}}$ , the deterministic dynamics ( $\sigma = 0$ ) generates periodic spiking (supra-threshold or mean-driven regime). For the EIF model, this condition reads  $\gamma_L < \mu/(V_T - \Delta_T)$ . Secondly, for  $\gamma_L > \mu/(V_T - \Delta_T)$ , the noiseless dynamics reaches a steady state below threshold, i.e. the neuron is quiescent unless it is excited by the noise (sub-threshold or fluctuation-driven regime). We will separately consider these two cases.

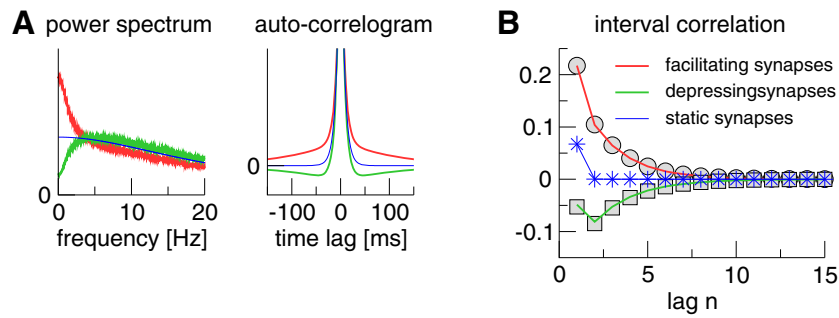
#### 2.3.1 Mean-driven regime

If  $\gamma_L < \mu/(V_T - \Delta_T)$ , the mean ISI is approximately given by the deterministic ISI

$$\langle T \rangle = \int_0^{V_{\text{spike}}} \frac{dv}{\gamma_L \Psi(v) + \mu}. \quad (17)$$

Furthermore, the dynamics is associated with a phase-response curve (PRC)  $Z(t')$  that describes the change of the ISI  $\delta T = -\alpha Z(t')$  caused by a weak current pulse  $\alpha \delta(t - t')$  at a given time  $t'$  (“phase”) in the interval  $(0, \langle T \rangle)$





**Fig. 7** Interval correlations caused by short-term synaptic plasticity. (A) Facilitating (depressing) synapses lead to a low-pass (high-pass) filtering of the input power, shown by red (green) line, when compared to a setup with static synapses (blue line). (B) For facilitation (depression), ISIs exhibit positive (negative) correlations. Parameters:

$\tau_F = 100\text{ms}$ ,  $\tau_D = 200\text{ms}$ ,  $v_E = v_I = 10\text{Hz}$ ,  $\bar{J}_E = 0.005$ ,  $\bar{J}_I = -0.010$ .  $F_0 = 0.05$ ,  $\Delta = 0.175$ ,  $D_\ell(t) \equiv 1$  (facilitation,  $\epsilon = 0.37$ );  $F_0 = 0.4$ ,  $F_\ell(t) \equiv F_0$  (depression,  $\epsilon = 0.39$ ). Amplitudes of static synapses are tuned to the average amplitudes of the depressing synapses ( $J_E = 0.00062$ ,  $J_I = -0.00125$ )

(Ermentrout and Terman 2010). For the nonlinear IF model the PRC reads

$$Z(t) = 1 / [\gamma_L \Psi(V_0(t)) + \mu] \tag{18}$$

(see, e.g. (Schwalger and Lindner 2013)). Here,  $V_0(t)$  is the deterministic solution of Eq. (16) with  $V_0(0) = 0$ . As shown in Appendix A.7, the CV and the SCC can be related to the power spectrum  $S_{in}(f)$  of a weak colored noise perturbation  $\sigma \eta(t)$  by

$$C_V^2 \approx \sigma^2 \int df S_{in}(f) |\tilde{Z}(f)|^2 \tag{19}$$

$$\rho_n \approx \frac{\sigma^2}{C_V^2} \int df S_{in}(f) |\tilde{Z}(f)|^2 e^{-2\pi i f (T)n}, \quad n \geq 1. \tag{20}$$

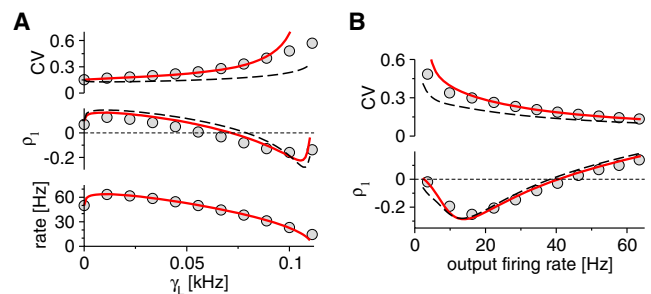
Here,  $\tilde{Z}(f) = \frac{1}{\langle T \rangle} \int_0^{\langle T \rangle} dt Z(t) e^{2\pi i f t}$ . Note that for the PIF model (i.e.  $\gamma_L = 0$ ), we obtain  $|\tilde{Z}_{PIF}(f)|^2 = \text{sinc}^2(f/r)/\mu^2$ . This exactly recovers the leading-order terms of CV and SCC, Eqs. (9) and (13).

Figure 8 shows the estimation of the CV and  $\rho_1$  using the PRC  $Z(t)$ , which agrees well with simulation data. We have also tried an even simpler approximation in this case by using the formulas for an effective PIF model with the same mean ISI given by Eq. (17) (see also Schwalger et al. (2013)). This approximation (dashed lines in Fig.7) already yields a good agreement. Most importantly, the effect of a presynaptic refractory period on postsynaptic variability and ISI correlations remains qualitatively the same as for the PIF neuron (cf. Fig. 4E and Fig. 8B): with increasing output firing rate, ISI correlations turn from negative to positive and output variability decreases.

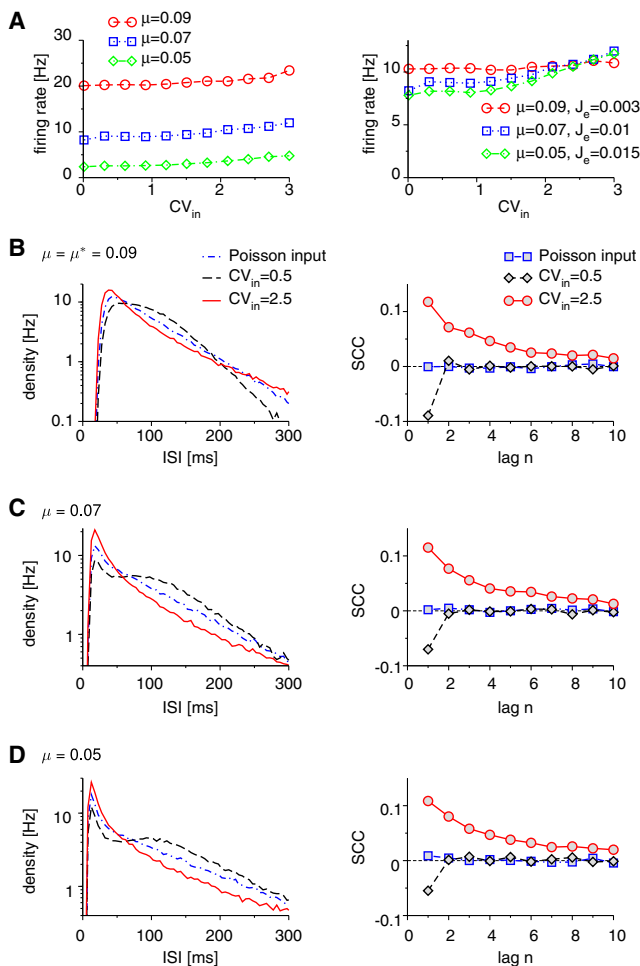
### 2.3.2 Fluctuation-driven regime

To see whether similar effects of non-Poissonian stimuli also hold in the fluctuation-driven (excitable) regime, we performed simulations of the EIF model with subthreshold

currents  $\mu \leq \mu^* \equiv \gamma_L (V_T - \Delta_T)$  (Fig. 9). In contrast to the mean-driven regime, where the firing rate is mainly determined by the deterministic ISI, Eq. (17), the firing rate in the fluctuation-driven regime may also depend on the properties of the noise. To study the sensitivity of the spiking statistics to regularity of the input spike trains, we again model input spike trains by inverse Gaussian renewal processes with varying  $C_{V,in}$  and fixed firing rates. Figure 9A (left panel) shows that the firing rate in this setup depends only weakly on the regularity of the input, whereas the base current  $\mu$  has a much stronger effect. The weak increase of the firing rate with  $C_{V,in}$  is caused by an increasing low frequency power due to higher input irregularity (cf. Fig. 4A and Fig. 5A). According to our main focus, we further investigated effects on higher-order ISI statistics at different values of  $\mu \leq \mu^*$ . To avoid trivial effects of the firing rate, we chose for each



**Fig. 8** ISI statistics for the exponential integrate-and-fire (EIF) model, Eq. (16), in the mean-driven regime. The neuron is driven by inverse Gaussian renewal processes each with  $r = 5\text{Hz}$  and  $C_V = 0.5$  (cf. Fig. 4). Solid red lines depict the theory, Eqs. (19, 20). The black dashed line shows the PIF-theory using an effective base current  $\mu_{\text{eff}} = V_T/\langle T \rangle$ , where  $\langle T \rangle$  is given by Eq. (17). **A** The leak  $\gamma_L$  is varied from  $\gamma_L = 0$  (limit of the PIF model) to the maximal value, at which the dynamics becomes subthreshold. ( $J_E = 0.005$ ,  $J_I = -0.01$ ,  $\mu = 0.1 V_T/\text{ms}$ ; corresponding to  $\epsilon = 0.24$ ) **B** The output firing rate is varied by the parameter  $\mu$  ( $J_E = 0.004$ ,  $J_I = -0.008$ ,  $\gamma_L = 0.01 \text{ms}^{-1}$ ). Other parameters:  $\Delta_T = 0.1 V_T$ ,  $V_{\text{spike}} = 2 V_T$ ,  $\tau_E = 4\text{ms}$ ,  $\tau_I = 8\text{ms}$



**Fig. 9** Firing rate, ISI density and serial correlation coefficient of the EIF model, Eq. (16), for the fluctuation-driven regime,  $\mu \leq \mu^* \equiv \gamma_L(V_T - \Delta_T)$ . The neuron receives balanced excitatory-inhibitory inputs modeled as inverse Gaussian renewal processes. **A** The firing rate depends only weakly on the regularity  $CV_{in}$  of the input spike trains. *Left*: Firing rate for different  $\mu$  and constant synaptic weights  $J_E = 0.01$ . *Right*: To obtain comparable firing rates for different  $\mu$  in panel **B–D**,  $J_E$  was chosen such that the firing rate was about 10 Hz at  $CV_{in} = 2.5$ . **B–D** ISI density (*left*) and SCC (*right*) for Poisson input (blue lines and symbols,  $CV_{in} = 1$ ) or inverse Gaussian input with  $CV_{in} = 0.5$  (dashed black lines) or  $CV_{in} = 2.5$  (red, solid lines). Note the logarithmic axis scaling in the left panels. **A** ISI statistics at the bifurcation point  $\mu = \mu^* = 0.09V_T/ms$  ( $J_E = 0.003$  corresponding to  $\epsilon(CV_{in} = 0.5) = 0.16$ ,  $\epsilon(CV_{in} = 2.5) = 0.17$ ). Below the bifurcation point: **B**  $\mu = 0.07V_T/ms$ ,  $J_E = 0.01$  ( $\epsilon(CV_{in} = 0.5) = 0.68$ ,  $\epsilon(CV_{in} = 2.5) = 0.72$ ); **C**  $\mu = 0.05V_T/ms$ ,  $J_E = 0.015$  ( $\epsilon(CV_{in} = 0.5) = 1.42$ ,  $\epsilon(CV_{in} = 2.5) = 1.50$ ). Other parameters:  $\gamma_L = 0.1ms^{-1}$ ,  $\Delta_T = 0.1V_T$ ,  $V_{spike} = 2V_T$ ,  $\tau_E = 4ms$ ,  $\tau_I = 8ms$ ,  $\nu_E = \nu_I = 5Hz$ ,  $N_E = 800$ ,  $N_I = 200$ ,  $J_I = -2J_E$

value of  $\mu$  synaptic weights that yielded comparable firing rates (Fig. 9A (right panel)). Directly at the bifurcation point  $\mu = \mu^*$  from sub- to suprathreshold firing, the ISI densities were unimodal and more (less) skewed than the case of Poisson input for irregular (regular) input (Fig. 9B) – similar to the findings for the mean-driven firing regime. Furthermore,

**Table 1** Values of the modified skewness  $\alpha_s$ , corresponding to the ISI densities shown in Fig. 9B–D

base current	$CV_{in} = 0.5$	$CV_{in} = 1$ (Poisson)	$CV_{in} = 2.5$
$\mu = \mu^* = 0.09$	0.84	0.98	1.13
$\mu = 0.07$	0.68	0.73	0.85
$\mu = 0.05$	0.65	0.67	0.79

Note that  $\alpha_s$  increases with input CV. Decreasing  $\mu$  leads to a transition from an inverse Gaussian distribution ( $\alpha_s = 1$ ) to a gamma distribution ( $\alpha_s = 2/3$ )

the ISI sequence showed significant correlations that were positive over several lags for irregular input ( $CV_{in} = 2.5$ ) and was negative at lag one for regular input ( $CV_{in} = 0.5$ ). Poisson input did not result in any significant serial correlations. Qualitatively, these results persist in the subthreshold case,  $\mu < \mu^*$ . In addition, we also observed that the ISI density can become bimodal for regular input (Fig. 9C, D).

The ISI densities become more skewed with increasing input irregularity, which is confirmed by the corresponding values of  $\alpha_s$  (Table 1). Furthermore, the skewness of the ISI density under Poisson stimulation approaches the value  $\alpha_s = 2/3$  for decreasing  $\mu$ . This corresponds to a gamma distribution, which is expected far below the bifurcation point,  $\mu \ll \mu^*$ , where the ISI statistics becomes Poissonian.

### 3 Discussion

In this paper, we put forward a theory for the spiking statistics of a mean-driven neuron stimulated by an arbitrary colored Gaussian noise. Within the developed framework, all statistics (ISI density, auto-correlogram, CV, skewness and SCC) can be expressed as nonlinear functions of two simple integrals of the input correlation function or, equivalently, of the input power spectrum. The leading order of the theory permits to understand the interaction between input time scales (e.g. the correlation time of the input stimulus) and the postsynaptic firing rate. We confirmed the theory for different scenarios and found excellent quantitative agreement as long as output CVs were below 0.6. Our theory unites and simplifies previous results for special cases of input statistics (Lindner 2004; Schwalger et al. 2010; Bauermeister et al. 2013). For the SCC and CV, we can extend our theory to non-perfect IF models in the mean-driven regime, replacing the sinc function in our formulas by a filter function that is determined by the phase-response curve of the respective IF model. Furthermore, simulations of the fluctuation-driven (excitable) regime indicate pronounced effects of non-Poissonian stimuli. How to extend the theory to the fluctuation-driven regime thus constitutes an important open problem.

Applying our theory, we found that distinct sources of input correlations can lead to similar ISI statistics, all showing pronounced differences to the case of a purely Poissonian drive. For instance, both presynaptic refractoriness and synaptic short-term depression lead to negative interval correlations and less skewed ISI distributions than with Poisson input. Conversely, rate-coded stimuli, presynaptic bursting, and synaptic facilitation cause positive ISI correlations and more skewed ISI densities. Our theory thus shows clearly the limitations for inferring neural dynamics from spike statistics only. Negative correlations, for instance, have often been found for neurons driven by fast fluctuations and adaptation currents, but, as we have shown, can be caused already by a much more basic feature, namely, presynaptic refractoriness. Despite this ambiguity, our formulas may still be applicable to draw conclusions about stimuli and neural parameters (see Bauermeister et al. (2013) and Fisch et al. (2012) for examples from the sensory periphery).

In a recent study (Peterson et al. 2014), negative ISI correlations observed in auditory nerve fibres have been attributed to synaptic depression of the ribbon synapse that generates the input to a nerve fibre. In that case, the spike train was modeled as a thinned-out version of stochastic transmitter release events of the synapse. The transmitter release events already carried negative inter-event interval correlations, which were inherited by the thinned-out spike train. Our theory for negative correlations due to synaptic depression provides an alternative explanation based on the temporal structure of synaptic input and its effect on a dynamic spike generator.

Relations between input and output spike statistics as developed in this paper may be also instrumental for novel approaches to recurrent neural networks. In the simple situation of a completely homogeneous and sparse network, the statistics of the postsynaptic spike train (forming the input to other cells) has to coincide with the statistics of presynaptic spike trains (Lerchner et al. 2006; Dummer et al. 2014). Specifically, in our theory one should be able to find a solution for which  $C_{in}(t) = C_{out}(t)$ . For heterogeneous and non-sparse networks, more self-consistency conditions should be met, not only for the auto- but also for the cross-correlation functions among neurons. In this sense, our paper constitutes only a first step in a theory of neural networks that incorporates the correlation statistics more accurately than the classical approach that uses the diffusion approximation.

**Acknowledgments** Research was supported by the European Research Council (Grant Agreement no. 268689, MultiRules), the BMBF (FKZ: 01GQ1001A), and the research training group GRK1589/1.

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

## Appendix A

### A.1 Gaussian approximation of the synaptic input current

We would like to approximate the synaptic input current

$$I_{syn}(t) = \sum_{k=1}^N \sum_{t_i^{(k)} < t} j_k(t - t_i^{(k)}) \tag{21}$$

by a Gaussian process with the same mean and correlation function as  $I_{syn}(t)$ . To this end, let us rewrite Eq. (21) in terms of the delta-spike trains  $x_k(t) = \sum_{t_i^{(k)} < t} \delta(t - t_i^{(k)})$ ,  $k = 1, \dots, N$  (where  $t_i^{(k)}$  denotes the  $i$ -th spike time at the  $k$ -th synapse):

$$I_{syn}(t) = \sum_{k=1}^N \int j_k(t - t') x_k(t') dt' \tag{22}$$

In the case of stationary input, mean and correlation function are given by

$$\langle I_{syn}(t) \rangle = \sum_{k=1}^N v_k \int j_k(t) dt \tag{23}$$

and

$$C_I(\tau) = \sum_{k=1}^N \sum_{l=1}^N \int dt_1 C_{kl}(t_1) \int dt_2 j_l(t_2 + \tau) j_k(t_1 + t_2) \tag{24}$$

Here,  $v_k = \langle x_k(t) \rangle$  is the stationary firing rate of the  $k$ -th input spike train and  $C_{kl}(\tau)$  is the correlation function between spike trains defined by

$$C_{kl}(\tau) = \langle x_k(t) x_l(t + \tau) \rangle - v_k v_l \tag{25}$$

It is convenient to formulate Eq. (24) in the Fourier domain. Defining the power spectrum as the Fourier transform of the correlation function

$$S_I(f) = \int d\tau e^{2\pi i f \tau} C_I(\tau) \tag{26}$$

we find

$$S_I(f) = \sum_{k=1}^N \sum_{l=1}^N \tilde{j}_k^*(f) \tilde{j}_l(f) S_{kl}(f) \tag{27}$$

Here and in the following,  $\tilde{\phantom{x}}$  and  $^*$  denote Fourier transform and complex conjugation, respectively. Using these approximations, we can write the current balance Eq. (1) as

$$\dot{V} = \mu + \zeta(t), \tag{28}$$

where

$$\mu = \frac{1}{C_m} \left( I_0 + \sum_{k=1}^N v_k \tilde{j}_k(0) \right) \tag{29}$$

and  $\zeta(t)$  being a zero mean Gaussian process with power spectrum

$$S_\zeta(f) = \frac{1}{C_m^2} S_I(f). \quad (30)$$

If  $S_\zeta(f)$  is not constant,  $\zeta(t)$  will be referred to as colored noise (in contrast to white noise, which has a flat power spectrum).

## A.2 Derivation of the interspike interval statistics

Our approach assumes that the colored noise can be represented as  $\zeta(t) = \mathbf{b} \cdot \mathbf{Y}(t)$ , where  $\mathbf{b}$  is a constant  $d$ -dimensional vector and  $\mathbf{Y}(t)$  is a  $d$ -dimensional Ornstein-Uhlenbeck. Specifically, the neuron model can be written as

$$\dot{V} = \mu + \mathbf{b} \cdot \mathbf{Y} \quad (31a)$$

$$\dot{\mathbf{Y}} = \mathbf{A}\mathbf{Y} + \mathbf{B}\boldsymbol{\xi}(t) \quad (31b)$$

with the additional fire-and-reset rule that  $V$  is reset to  $V = 0$  whenever  $V$  reaches the threshold  $V_T$ .  $\mathbf{A}$  and  $\mathbf{B}$  are drift and diffusion matrices, respectively. The eigenvalues of  $\mathbf{A}$  must have negative real part and  $\boldsymbol{\xi}(t)$  is a vector of Gaussian white noise processes that obey  $\langle \xi_i(t)\xi_j(t') \rangle = \delta_{i,j}\delta(t-t')$ . Furthermore, we assume that the elements of  $\mathbf{b}$  have the same physical dimensions as  $\mu$  and that  $\mathbf{Y}$  has non-dimensional units.

The variance of  $\zeta(t)$  is given by  $\sigma^2 = \mathbf{b}^T \boldsymbol{\Sigma} \mathbf{b}$ , where  $\boldsymbol{\Sigma}$  is the covariance matrix of the OUP with elements

$$\Sigma_{ij} = \langle Y_i Y_j \rangle. \quad (32)$$

The covariance matrix is related to the drift matrix  $\mathbf{A}$  and the diffusion matrix

$$\mathbf{D} = \frac{1}{2} \mathbf{B}^T \mathbf{B} \quad (33)$$

by the Lyapunov equation (see e.g. Risken 1984; van Kampen 1992)

$$\mathbf{A}\boldsymbol{\Sigma} + \boldsymbol{\Sigma}\mathbf{A}^T = -2\mathbf{D}. \quad (34)$$

This system of linear equations can be used to solve for the elements of the covariance matrix  $\Sigma_{ij}$ .

### Non-dimensional model

It is useful to non-dimensionalize the model Eq. (31) by measuring time in units of the mean ISI  $V_T/\mu$  and voltage in units  $V_T$ . This corresponds to the introduction of dimensionless variables

$$\hat{V} = \frac{V}{V_T}, \quad \hat{t} = \frac{\mu}{V_T} t \quad (35)$$

and dimensionless parameters

$$\hat{\mathbf{b}} = \frac{\mathbf{b}}{\sigma}, \quad \hat{\mathbf{A}} = \frac{V_T}{\mu} \mathbf{A}, \quad \hat{\mathbf{B}} = \sqrt{\frac{V_T}{\mu}} \mathbf{B}. \quad (36)$$

This yields the non-dimensionalized Langevin equation

$$\dot{\hat{V}} = 1 + \epsilon \hat{\mathbf{b}} \cdot \mathbf{Y}, \quad (37a)$$

$$\dot{\mathbf{Y}} = \hat{\mathbf{A}}\mathbf{Y} + \hat{\mathbf{B}}\boldsymbol{\xi}(t). \quad (37b)$$

The non-dimensional firing threshold is  $\hat{V}_T = 1$ . To make the small parameter in our perturbation calculation explicit, we have introduced the dimensionless parameter  $\epsilon = \sigma/\mu \ll 1$ . In the following, the non-dimensional model Eq. (37) will be considered. For notational simplicity, we will omit in the following the hats on top of the non-dimensional quantities.

The stationary auto-correlation function of the noise  $\eta(t) = \mathbf{b} \cdot \mathbf{Y}(t)$  is for  $\tau \geq 0$  given by (see e.g. (Riskin 1984))

$$C_{\text{in}}(t) \equiv \langle \eta(t)\eta(t+\tau) \rangle = \mathbf{b}^T e^{\tau\mathbf{A}} \boldsymbol{\Sigma} \mathbf{b}, \quad (38)$$

where  $e^{\tau\mathbf{A}}$  is the matrix exponential defined by its series representation  $e^{\tau\mathbf{A}} = \sum_{k=0}^{\infty} \frac{1}{k!} (\tau\mathbf{A})^k$ . Note that by our rescaling of  $\mathbf{b}$ , Eq. (36), the variance of  $\eta(t)$  is  $C_{\text{in}}(0) = \langle \eta^2 \rangle = 1$ . Furthermore, taking the Fourier transform of the correlation function yields the input power spectrum

$$S_{\text{in}}(f) = -2\mathbf{b}^T \left[ \mathbf{A}^2 + (2\pi f)^2 \mathbb{I}_d \right]^{-1} \mathbf{A} \boldsymbol{\Sigma} \mathbf{b}, \quad (39)$$

where  $\mathbb{I}_d$  is the  $d \times d$  identity matrix.

### Fokker-Planck equation

The stochastic dynamics (37) with fire-and-reset rule can be reformulated in terms of probability currents and densities. The probability current that describes the drift of  $V(t)$  is given by

$$J_v(v, \mathbf{y}, t) = (1 + \epsilon \mathbf{b} \cdot \mathbf{y}) p(v, \mathbf{y}, t), \quad (40)$$

where  $p(v, \mathbf{y}, t) = \langle \delta(v - V(t)) \delta(\mathbf{y} - \mathbf{Y}(t)) \rangle$  is the probability density of  $V(t)$  and  $\mathbf{Y}(t)$  at time  $t$ . The  $d$ -dimensional OUP corresponds to  $d$  probability currents

$$J_i(v, \mathbf{y}, t) = \sum_{j=1}^d (A_{ij} y_j - D_{ij} \frac{\partial}{\partial y_j}) p(v, \mathbf{y}, t) \quad (41)$$

Here,  $i = 1, \dots, d$  and the non-dimensionalized diffusion matrix is defined by  $\mathbf{D} = \frac{1}{2} \hat{\mathbf{B}}^T \hat{\mathbf{B}}$ . Because trajectories cannot leave the system, the probability is conserved, which can be expressed by the continuity equation

$$\frac{\partial p}{\partial t} + \frac{\partial J_v}{\partial v} + \sum_{i=1}^d \frac{\partial J_i}{\partial y_i} = J_v(1, \mathbf{y}, t) [\delta(v) - \delta(v-1)]. \quad (42)$$



The right hand side represents the reinjection and outflux of probability at the reset  $v = 0$  (source) and the threshold  $v = 1$  (sink), respectively, corresponding to the reset of trajectories. With the definition of the currents, Eqs. (40) and (41), the last equation can be written as a partial differential equation for  $p(v, \mathbf{y}, t)$ . Because the source and sink terms are sharply localized, we can omit these terms in the regions  $v < 0$ ,  $0 < v < 1$  and  $v > 0$  and we arrive at the Fokker-Planck equation (FPE), Eq. (3) with non-dimensionalized parameters  $\mu = 1$  and  $\sigma = \epsilon$ . The source and sink terms can be realized by imposing appropriate boundary conditions at  $v = 0$  and  $v = 1$ . The sink term at  $v = 1$  ensures that no probability can be found beyond the threshold and thus there is no backflux of probability from above the threshold, i.e.  $J_v(1, \mathbf{y}, t) = 0$  for  $1 + \epsilon \mathbf{b} \cdot \mathbf{y} < 0$ . With Eq. (40) this implies the condition (cf. (Brunel and Sergi 1998))

$$p(1, \mathbf{y}, t) = 0 \quad \forall \mathbf{y} : 1 + \epsilon \mathbf{b} \cdot \mathbf{y} < 0. \tag{43}$$

Furthermore, integration of Eq. (42) from  $v = -\epsilon$  to  $v = \epsilon$  (and letting  $\epsilon \rightarrow 0$ ), reveals that  $J_v(v, \mathbf{y}, t)$  suffers a jump at  $v = 0$  by an amount  $J_v(1, \mathbf{y}, t)$ . With Eq. (40) this implies the condition

$$p(0+, \mathbf{y}, t) - p(0-, \mathbf{y}, t) = p(1, \mathbf{y}, t), \tag{44}$$

where  $v = 0-$  and  $v = 0+$  denote the left- and right-sided limits. Finally, probability should not leak across infinite boundaries and the probability density must be normalized:

$$\lim_{v \rightarrow -\infty} p(v, \mathbf{y}, t) = 0, \quad \lim_{y_i \rightarrow \pm\infty} p(v, \mathbf{y}, t) = 0, \tag{45}$$

$$\int dv \int d^d y p(v, \mathbf{y}, t) = 1. \tag{46}$$

In addition to the boundary conditions, the FPE must be supplied with an initial condition. Because we aim at the statistics of the interspike intervals, it is advantageous to consider an ensemble of neurons that have just fired a spike at time  $t = 0$ . This corresponds to preparing the system in a state, in which the membrane potential has been reset to zero, i.e.  $V(0) = 0$ , and the OUP is distributed according to the distribution of  $\mathbf{Y}$  sampled at the moments of firing. If we denote this distribution by  $p_{\text{spike}}(\mathbf{y})$ , the initial condition for the FPE must be of the form  $p(v, \mathbf{y}, 0) = \delta(v) p_{\text{spike}}(\mathbf{y})$ .

### Distribution upon firing

In order to determine  $p_{\text{spike}}(\mathbf{y})$ , we first seek the stationary solution  $p_s(v, \mathbf{y})$  of the FPE. Setting  $\partial p / \partial t = 0$ , we obtain the stationary FPE

$$0 = (1 + \epsilon \mathbf{b}^T \mathbf{y}) \frac{\partial p_s}{\partial v} + \sum_{i,j=1}^d \left( A_{ij} \frac{\partial}{\partial y_i} (y_j p_s) - D_{ij} \frac{\partial^2 p_s}{\partial y_i \partial y_j} \right), \tag{47}$$

which is still difficult to solve analytically due to the discontinuous boundary condition (43) at  $v = 1$ . However, for weak noise,  $\epsilon \ll 1$ , the velocity  $1 + \epsilon \mathbf{b} \cdot \mathbf{y}$ , is practically always positive, so that the boundary condition (43) can be safely ignored. A positive velocity also precludes trajectories from entering the region  $v < 0$ . As a consequence, the system is restricted to the region  $0 \leq v \leq 1$  and has the periodic boundary condition

$$p_s(0, \mathbf{y}) = p_s(1, \mathbf{y}), \quad \mathbf{y} \in \mathbb{R}^d \tag{48a}$$

$$\lim_{y_i \rightarrow \pm\infty} p_s(v, \mathbf{y}) = 0, \quad v \in [0, 1] \tag{48b}$$

$$\int d^d y \int_0^1 dv p_s(v, \mathbf{y}) = 1. \tag{48c}$$

A solution of Eq. (47) is given by the stationary distribution of the OUP, i.e.

$$p_s(v, \mathbf{y}) = p_s(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left(-\frac{1}{2} \mathbf{y}^T \Sigma^{-1} \mathbf{y}\right), \tag{49}$$

which clearly fulfills the boundary conditions (48).

The distribution upon firing  $p_{\text{spike}}(\mathbf{y})$  is proportional to the stationary probability current  $J_v^{(s)}(1, \mathbf{y}) = (1 + \epsilon \mathbf{b} \cdot \mathbf{y}) p_s(1, \mathbf{y})$  across the threshold (Lindner 2004). In fact, the probability per unit time that a trajectory crosses the threshold through the surface element  $d\mathbf{S}(\mathbf{y})$  at the point  $\mathbf{y}$  on the hyperplane defined by  $v = 1$  is given by  $J_v^{(s)}(1, \mathbf{y}) dy_1 \cdots dy_d$ . Thus, the probability density upon firing must be proportional to  $J_v^{(s)}(1, \mathbf{y})$ . The factor of proportionality can be found from the normalization condition, which yields  $p_{\text{spike}}(\mathbf{y}) = (1 + \epsilon \mathbf{b} \cdot \mathbf{y}) p_s(1, \mathbf{y})$  and hence the initial distribution

$$p(v, \mathbf{y}, 0) = \delta(v) (1 + \epsilon \mathbf{b} \cdot \mathbf{y}) p_s(\mathbf{y}). \tag{50}$$

### Interspike interval density

An interspike interval is equivalent to the first-passage-time (FPT) with respect to the threshold  $V_T = 1$ , i.e. the time a system that has been prepared according to a spike at  $t = 0$  needs to reach the threshold for the first time. A standard approach to obtain the probability density of the FPT (i.e. the ISI density) is to consider an ensemble in which trajectories are *not* reset upon threshold crossing but are taken out of the system. As a consequence, the probability to find a trajectory with  $v < 1$ ,

$$G(t) = \int d^d y \int_{-\infty}^1 dv p(v, \mathbf{y}, t), \tag{51}$$

is the probability that a trajectory survives until time  $t$ . Here,  $p(v, \mathbf{y}, t)$  is the time-dependent solution of the FPE, for which the reset condition (44) has been dropped now. The

FPT density is given by  $P(t) = -dG/dt$ , or

$$P(t) = - \int d^d y \int_{-\infty}^1 dv \partial_t p(v, \mathbf{y}, t). \quad (52)$$

Using the continuity Eq. (42) and the natural boundary conditions (45) yields

$$P(t) = \int d^d y J_v(1, \mathbf{y}, t) = \mathcal{P}(t, 1), \quad (53)$$

where

$$\mathcal{P}(t, v) = \int d^d y (1 + \epsilon \mathbf{b}^T \mathbf{y}) p(v, \mathbf{y}, t) \quad (54)$$

is the FPT density with respect to a general threshold  $V_T = v$ . Eq. (53) simply expresses that the FPT is equal to the total probability current through the threshold at time  $t$ . The ISI density can be rewritten as

$$P(t) = \frac{1}{2\pi} \int d\ell e^{-i\ell v} \tilde{P}(\ell, t) \Big|_{v=1}, \quad (55)$$

where  $\tilde{P}(t, \ell) = \int dv e^{i\ell v} \mathcal{P}(t, v)$  is the Fourier transform with respect to  $v$ . Using Eq. (54), we thus have

$$\tilde{P}(t, \ell) = \int dv e^{i\ell v} \int d^d y (1 + \epsilon \mathbf{b}^T \mathbf{y}) p(v, \mathbf{y}, t). \quad (56)$$

This function can be further related to the characteristic function

$$q(\ell, \mathbf{k}, t) = \int dv e^{i\ell v} \int d^d y e^{i\mathbf{k} \cdot \mathbf{y}} p(v, \mathbf{y}, t), \quad (57)$$

where  $\mathbf{k}$  is a  $d$ -dimensional vector with elements  $k_i$ ,  $i = 1, \dots, d$ . Indeed, comparing (56) with (57), we find the relation

$$\tilde{P}(t, \ell) = q(\ell, \mathbf{k}, t) - i\epsilon [\nabla_{\mathbf{k}} q(\ell, \mathbf{k}, t)]^T \mathbf{b} \Big|_{\mathbf{k}=\mathbf{0}}, \quad (58)$$

where  $\nabla_{\mathbf{k}} q$  denotes the gradient of  $q$  with respect to  $\mathbf{k}$ , i.e.  $\nabla_{\mathbf{k}} q$  is a vector with elements  $\partial_i q \equiv \partial q / \partial k_i$ . Thus, if the characteristic function  $q(\ell, \mathbf{k}, t)$  is known, one can derive the ISI density via (55) and (58). In order to compute  $q(\ell, \mathbf{k}, t)$ , it is not necessary to calculate the probability density  $p(v, \mathbf{y}, t)$  explicitly from the FPE, which does not permit a closed form solution for the initial conditions (50). However, applying the Fourier transform to the FPE directly yields a first-order differential equation for  $q(\ell, \mathbf{k}, t)$ :

$$\partial_t q - (\mathbf{k}^T \mathbf{A} + \epsilon \ell \mathbf{b}^T) \nabla_{\mathbf{k}} q = (-\mathbf{k}^T \mathbf{D} \mathbf{k} + i\ell) q \quad (59)$$

with initial condition

$$q(\ell, \mathbf{k}, 0) = (1 + i\epsilon \mathbf{b}^T \Sigma \mathbf{k}) \exp(-\mathbf{k}^T \mathbf{D} \mathbf{k}). \quad (60)$$

This equation can be solved by the method of characteristics (see e.g. (Kamke 1965)). The characteristic equations are

$$\dot{\mathbf{k}} = -\mathbf{A}^T \mathbf{k} - \epsilon \ell \mathbf{b} \quad (61)$$

$$\dot{q} = (-\mathbf{k}^T \mathbf{D} \mathbf{k} + i\ell) q. \quad (62)$$

The solution of the first equation is

$$\mathbf{k}(t, \mathbf{c}) = e^{-t\mathbf{A}^T} \mathbf{c} - \epsilon \ell \int_0^t dt' e^{-(t-t')\mathbf{A}^T} \mathbf{b}, \quad (63)$$

which can be solved for the initial vector  $\mathbf{c}$ :

$$\mathbf{c}(t, \mathbf{k}) = e^{t\mathbf{A}^T} \mathbf{k} + \epsilon \ell \int_0^t dt' e^{t'\mathbf{A}^T} \mathbf{b}. \quad (64)$$

The solution of Eq. (62) is given by

$$q(t, \mathbf{c}) = Q_0(\mathbf{c}) \times \exp \left( -\frac{1}{2} \mathbf{k}^T \Sigma \mathbf{k} + i\ell t - \epsilon \ell \mathbf{b}^T \Sigma \int_0^t dt' \mathbf{k}(t', \mathbf{c}) \right), \quad (65)$$

where the Lyapunov equation (34) has been used. The prefactor  $Q_0(\mathbf{c})$  can be determined from the initial condition (60) as  $Q_0 = 1 + i\epsilon \mathbf{b}^T \Sigma \mathbf{c}$ . The integration in Eq. (65) has to be performed using the expression (63), followed by a re-substitution of the vector  $\mathbf{c} = \mathbf{c}(t, \mathbf{k})$  according to Eq. (64). As a result we obtain the characteristic function

$$q(\ell, \mathbf{k}, t) = (1 + i\epsilon \mathbf{b}^T \Sigma e^{t\mathbf{A}^T} \mathbf{k} + i\epsilon^2 \ell g(t)) \times \exp \left( -\frac{1}{2} \mathbf{k}^T \Sigma \mathbf{k} + i\ell t - \epsilon \ell \mathbf{b}^T \Sigma \int_0^t dt' e^{t'\mathbf{A}^T} \mathbf{k} - \epsilon^2 \ell^2 h(t) \right),$$

where

$$g(t) = \int_0^t dt' C_{\text{in}}(t') \quad h(t) = \int_0^t dt'' \int_0^{t''} dt' C_{\text{in}}(t'), \quad (66)$$

Using Eq. (58) leads to

$$\tilde{P}(t, \ell) = [1 + \epsilon^2 C_{\text{in}}(t) + 2i\epsilon^2 \ell g(t) - \epsilon^4 \ell^2 g^2(t)] e^{i\ell t - \epsilon^2 \ell^2 h(t)},$$

and

$$P(t) = \frac{1}{2\sqrt{4\pi\epsilon^2 h^3(t)}} \exp \left[ -\frac{(t-1)^2}{4\epsilon^2 h(t)} \right] \times \left\{ \frac{[(1-t)g(t) + 2h(t)]^2}{2h(t)} - \epsilon^2 [g^2(t) - 2h(t)C_{\text{in}}(t)] \right\}, \quad (67)$$

which constitutes the final result for the ISI density. Let us provide two alternative expressions for  $h(t)$  and  $g(t)$ . First, if  $\mathbf{A}$  is invertible we can write these functions in terms of the drift matrix:

$$h(t) = \mathbf{b}^T \Sigma (\mathbf{A}^T)^{-2} \left[ e^{t\mathbf{A}^T} - t\mathbf{A}^T - \mathbb{I}_d \right] \mathbf{b} \quad (68)$$

$$g(t) = \mathbf{b}^T \Sigma (\mathbf{A}^T)^{-1} \left[ e^{t\mathbf{A}^T} - \mathbb{I}_d \right] \mathbf{b}. \quad (69)$$

Second, using the power spectrum of the noise

$$S_{in}(f) = \int_{-\infty}^{\infty} d\tau C_{in}(\tau)e^{2\pi if\tau}, \tag{70}$$

we find

$$h(t) = \frac{t^2}{2} \int_{-\infty}^{\infty} df S_{in}(f)\text{sinc}^2(ft),$$

$$g(t) = t \int_{-\infty}^{\infty} df S_{in}(f)\text{sinc}(2ft),$$

where  $\text{sinc}(x) = \sin(\pi x)/(\pi x)$ .

*Cumulants of the n-th order intervals*

Apart from the ISI density the ISI statistics can be characterized by its cumulants (semi-invariants). We will calculate these cumulants also for the  $n$ th-order intervals because they will be used to find an expression for the SCC and the calculation does not pose any extra difficulties. In fact, as argued in the main text, the time  $t_n$  of the  $n$ -th threshold crossing, i.e. the  $n$ th-order interval, is equal to the first-passage time of the free process with respect to the  $n$ th-fold threshold  $V_T = n$  (given the same noise realization). The probability density of the  $n$ th-order interval is thus given by  $\mathcal{P}(t, n)$  (cf. Eq. (54)). The  $k$ -th cumulants of the  $n$ th-order intervals is given by (van Kampen 1992)

$$\kappa_k(n) = (-1)^k \left. \frac{d^k \ln \bar{P}_n}{ds^k} \right|_{s=0}, \tag{71}$$

where

$$\bar{P}_n(s) = \int_0^\infty dt e^{-st} \mathcal{P}(t, n). \tag{72}$$

is the Laplace transform of the  $n$ -th-order interval density, which plays the role of a moment generating function. Introducing the Laplace-Fourier transform of the probability density

$$\varphi(v, \mathbf{k}, s) = \int_0^\infty dt e^{-st} \int d^d y e^{i\mathbf{k}\cdot\mathbf{y}} p(v, \mathbf{y}, t), \tag{73}$$

the moment generating function can be rewritten as

$$\bar{P}_n(s) = [\varphi - i\epsilon \mathbf{b} \cdot \nabla_{\mathbf{k}} \varphi]_{\mathbf{k}=0, v=n}. \tag{74}$$

The function  $\varphi$  satisfies the transformed FPE

$$\begin{aligned} \partial_v \varphi - \mathbf{k}^T \mathbf{A} \nabla_{\mathbf{k}} \varphi &= -(s + \mathbf{k}^T \mathbf{D} \mathbf{k}) \varphi + \delta(v) \exp\left(-\frac{\mathbf{k}^T \Sigma \mathbf{k}}{2}\right) \\ &+ i\epsilon \mathbf{b}^T \left( \partial_v \nabla_{\mathbf{k}} \varphi + \delta(v) \Sigma \mathbf{k} \exp\left(-\frac{\mathbf{k}^T \Sigma \mathbf{k}}{2}\right) \right). \end{aligned} \tag{75}$$

The mixed derivative  $\partial_v \nabla_{\mathbf{k}} \varphi$  makes an exact solution infeasible. However, using the perturbation expansion

$$\varphi = \varphi^{(0)} + \epsilon \varphi^{(1)} + \epsilon^2 \varphi^{(2)} + \dots, \tag{76}$$

yields a hierarchy of first-order partial differential equations for the coefficients  $\varphi^{(k)}(v, \mathbf{k}, s)$ . Specifically, we obtain

$$\partial_v \varphi^{(m)} - \mathbf{k}^T \mathbf{A} \nabla_{\mathbf{k}} \varphi^{(m)} = -(s + \mathbf{k}^T \mathbf{D} \mathbf{k}) \varphi^{(m)} + I_m, \tag{77}$$

where the inhomogeneities  $I_m$  are given by

$$I_0 = \delta(v) \exp\left(-\frac{\mathbf{k}^T \Sigma \mathbf{k}}{2}\right) \tag{78}$$

$$I_1 = i\mathbf{b}^T \partial_v \nabla_{\mathbf{k}} \varphi^{(0)} + i\delta(v) \mathbf{b}^T \Sigma \mathbf{k} \exp\left(-\frac{\mathbf{k}^T \Sigma \mathbf{k}}{2}\right) \tag{79}$$

$$I_m = i\mathbf{b}^T \partial_v \nabla_{\mathbf{k}} \varphi^{(m-1)}, \quad m \geq 2. \tag{80}$$

If we insert the ansatz (76) into the formula (74), we obtain a perturbation series for the moment generating function:

$$\bar{P}_n(s) = \bar{P}_n^{(0)}(s) + \epsilon \bar{P}_n^{(1)}(s) + \epsilon^2 \bar{P}_n^{(2)}(s) + \dots, \tag{81}$$

where  $\bar{P}_n^{(0)}(s) = \varphi^{(0)}(n, 0, s)$  and

$$\bar{P}_n^{(m)}(s) = \left[ \varphi^{(m)} - i\mathbf{b}^T \nabla_{\mathbf{k}} \varphi^{(m-1)} \right]_{\mathbf{k}=0, v=n} \tag{82}$$

for  $m \geq 1$ . Equation (77) can be again solved by the method of characteristics. To this end, the solution surface  $(v, \mathbf{k}, \varphi^{(m)}(v, \mathbf{k}, s))$  is parametrized by a parameter  $\tau$  and  $d$  initial constants  $c_i, i = 1, \dots, d$ . The total derivative of  $\varphi^{(m)}$  with respect to  $\tau$  is then  $d\varphi^{(m)}/d\tau = \frac{dv}{d\tau} \partial_v \varphi^{(m)} + \frac{d\mathbf{k}^T}{d\tau} \nabla_{\mathbf{k}} \varphi^{(m)}$ . Comparison with Eq. (77) yields the characteristic equations

$$\frac{dv}{d\tau} = 1, \tag{83}$$

$$\frac{d\mathbf{k}}{d\tau} = -\mathbf{A}^T \mathbf{k}, \tag{84}$$

$$\frac{d\varphi^{(m)}}{d\tau} = -(s + \mathbf{k}^T \mathbf{D} \mathbf{k}) \varphi^{(m)} + I_m. \tag{85}$$

The integration of the first two equations results in

$$v = \tau, \quad \mathbf{k} = e^{-\tau \mathbf{A}^T} \mathbf{c}. \tag{86}$$

To solve the third equation, we make the ansatz

$$\varphi^{(m)}(\tau, \mathbf{c}) = \exp\left(-s\tau - \frac{1}{2} \mathbf{k}^T \Sigma \mathbf{k}\right) \psi^{(m)}(\tau, \mathbf{c}). \tag{87}$$

Using this ansatz in Eq. (85) yields

$$\frac{d\psi^{(m)}}{d\tau} = -\left( \mathbf{k}^T \mathbf{D} \mathbf{k} - \frac{1}{2} \frac{d}{d\tau} (\mathbf{k}^T \Sigma \mathbf{k}) \right) \psi^{(m)} + \hat{I}_m(\tau, \mathbf{c}) \tag{88}$$

where  $\hat{I}_n(\tau, \mathbf{c}) = I_n \exp\left(s\tau + \frac{1}{2}\mathbf{k}^T \Sigma \mathbf{k}\right)$ . In terms of the lower-order functions  $\psi^{n-1}$ , the inhomogeneous parts can be written as

$$\hat{I}_0 = \delta(\tau), \tag{89}$$

$$\hat{I}_1 = i\mathbf{b}^T \left[ \partial_v \nabla_{\mathbf{k}} \psi^{(0)} - \Sigma \mathbf{k} \partial_v \psi^{(0)} - s \left( \nabla_{\mathbf{k}} \psi^{(0)} - \Sigma \mathbf{k} \psi^{(0)} \right) + \delta(\tau) \Sigma \mathbf{k} \right] \tag{90}$$

$$\hat{I}_m = i\mathbf{b}^T \left[ \partial_v \nabla_{\mathbf{k}} \psi^{(m-1)} - \Sigma \mathbf{k} \partial_v \psi^{(m-1)} - s \left( \nabla_{\mathbf{k}} \psi^{(m-1)} - \Sigma \mathbf{k} \psi^{(m-1)} \right) \right], \quad m \geq 2. \tag{91}$$

In order to regard these expressions as functions of  $\tau$  and  $\mathbf{c}$ , the vector  $\mathbf{k}$  has to be understood as  $\mathbf{k}(\tau, \mathbf{c}) = e^{-\tau A^T} \mathbf{c}$ . It turns out that in Eq. (88), the prefactor in front of  $\psi^{(m)}$  vanishes because in light of Eqs. (34) and (84)

$$\begin{aligned} \frac{d}{d\tau} (\mathbf{k}^T \Sigma \mathbf{k}) &= \dot{\mathbf{k}}^T \Sigma \mathbf{k} + \mathbf{k}^T \Sigma \dot{\mathbf{k}} = -(\mathbf{k}^T A \Sigma \mathbf{k} + \mathbf{k}^T \Sigma A^T \mathbf{k}) \\ &= -\mathbf{k}^T (A \Sigma + \Sigma A^T) \mathbf{k} = 2\mathbf{k}^T D \mathbf{k}. \end{aligned} \tag{92}$$

Thus,  $\psi^{(m)}$  can be easily integrated:

$$\psi^{(m)}(\tau, \mathbf{c}) = \int_{-\infty}^{\tau} d\tau' \hat{I}_m(\tau', \mathbf{c}). \tag{93}$$

For the evaluation of this formula, the derivatives of the lower-order functions  $\psi^{(m-1)}(\tau, \mathbf{c})$  with respect to  $v$  and  $\mathbf{k}$  are needed. To this end, it is useful to apply the chain rule: for a given scalar function  $f(\tau, \mathbf{c})$  we can write

$$(\nabla_{\mathbf{k}} f)^T = [\partial_{\tau} f \nabla_{\mathbf{k}} \tau + (\nabla_{\mathbf{k}} \mathbf{c}) \nabla_{\mathbf{c}} f]^T = (\nabla_{\mathbf{c}} f)^T e^{\tau A^T} \tag{94}$$

and

$$\partial_v f = \partial_{\tau} f + (\nabla_{\mathbf{c}} f)^T A^T \mathbf{c}. \tag{95}$$

In Eq. (94), we made use of  $\nabla_{\mathbf{k}} \tau = 0$  and  $\nabla_{\mathbf{k}} \mathbf{c} = \left(e^{\tau A^T}\right)^T = e^{\tau A}$ , where  $\nabla_{\mathbf{k}} \mathbf{c}$  is a matrix with elements  $(\nabla_{\mathbf{k}} \mathbf{c})_{ij} = \partial c_j / \partial k_i$ . In Eq. (95), the relation  $\partial_v \mathbf{c} = \partial_v e^{v A^T} \mathbf{k} = A^T \mathbf{c}$  was employed.

For  $m \geq 1$ , the perturbation coefficient  $\bar{P}_n^{(m)}(s)$  can be expressed in terms of  $\psi^{(m-1)}(\tau, \mathbf{c})$  by using Eqs. (82), (93) and (95):

$$\begin{aligned} \bar{P}_n^{(m)}(s) &= e^{-sn} \left[ \psi^{(m)} - i\mathbf{b}^T \nabla_{\mathbf{k}} \psi^{(m-1)} \right]_{\mathbf{c}=0, \tau=n}, \\ &= e^{-sn} \left\{ \int_{-\infty}^{\tau} d\tau' i\mathbf{b}^T \left[ \partial_v \nabla_{\mathbf{k}} \psi^{(m-1)} - s \nabla_{\mathbf{k}} \psi^{(m-1)} - i\mathbf{b}^T \nabla_{\mathbf{k}} \psi^{(m-1)} \right] \right\}_{\mathbf{c}=0, \tau=n}, \\ &= -ise^{-sn} \left[ \int_0^n d\tau' \mathbf{b}^T \nabla_{\mathbf{k}} \psi^{(m-1)}(\tau', \mathbf{c}) \right]_{\mathbf{c}=0}. \end{aligned} \tag{96}$$

**Zerth-order** For  $m = 0$ , we obtain from Eq. (93)

$$\psi^{(0)}(\tau, \mathbf{c}) = \theta(\tau). \tag{97}$$

Equation (87) then yields the leading-order of the moment-generating function

$$\bar{P}_n^{(0)}(s) = \varphi^{(0)}(n, 0, s) = e^{-sn}. \tag{98}$$

The derivatives of  $\psi^{(0)}$  simply read

$$\partial_v \psi^{(0)} = \delta(\tau), \quad \nabla_{\mathbf{k}} \psi^{(0)} = 0. \tag{99}$$

**First-order** Because the gradient in Eq. (96) equals zero for  $m = 1$  (Eq. (99)), the first-order contribution to  $\bar{P}_n(s)$  vanishes:

$$\bar{P}_n^{(1)}(s) = 0. \tag{100}$$

From Eqs. (90) and (93) we obtain

$$\psi^{(1)}(\tau, \mathbf{c}) = i\theta(\tau)s \int_0^{\tau} d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c}. \tag{101}$$

Using Eq. (94), the gradient of  $\psi^{(1)}$  reads

$$\left(\nabla_{\mathbf{k}} \psi^{(1)}\right)^T = i\theta(t)s \int_0^{\tau} d\tau_1 \mathbf{b}^T \Sigma e^{\tau_1 A^T}. \tag{102}$$

Furthermore, the derivative with respect to  $v$  is

$$\begin{aligned} \partial_v \psi^{(1)} &= i\theta(t)s \left( \mathbf{b}^T \Sigma e^{-\tau A^T} \mathbf{c} + \int_0^{\tau} d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} A^T \mathbf{c} \right) \\ &= i\theta(t)s \mathbf{b}^T \Sigma \mathbf{c}. \end{aligned}$$

and the mixed derivative takes the form

$$\left(\nabla_{\mathbf{k}} \partial_v \psi^{(1)}\right)^T = i\theta(t)s \mathbf{b}^T \Sigma e^{\tau A^T}. \tag{103}$$

**Second-order** For  $m = 2$ , we find the first non-vanishing correction of the moment-generating function due to noise:

$$\bar{P}_n^{(2)}(s) = s^2 e^{-sn} \int_0^n d\tau_2 \int_0^{\tau_2} d\tau_1 C_{in}(\tau_1). \tag{104}$$

Here,  $C_{in}(t)$  is the noise correlation function as defined in Eq. (38). Furthermore, we obtain

$$\begin{aligned} \psi^{(2)} &= -\theta(\tau)s \int_0^{\tau} d\tau_2 \left\{ \mathbf{b}^T \Sigma e^{\tau_2 A^T} \mathbf{b} - \left( \mathbf{b}^T \Sigma e^{-\tau_2 A^T} \mathbf{c} \right) \left( \mathbf{b}^T \Sigma \mathbf{c} \right) - s \left[ \int_0^{\tau_2} d\tau_1 \mathbf{b}^T \Sigma e^{\tau_1 A^T} \mathbf{b} - \left( \mathbf{b}^T \Sigma e^{-\tau_2 A^T} \mathbf{c} \right) \int_0^{\tau_2} d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c} \right] \right\} \end{aligned} \tag{105}$$



and the gradient

$$\begin{aligned}
 (\nabla_{\mathbf{k}}\psi^{(2)})^T &= -\theta(\tau)s \int_0^\tau d\tau_2 \left\{ -(\mathbf{b}^T \Sigma \mathbf{c}) \mathbf{b}^T \Sigma e^{(\tau-\tau_2)A^T} \right. \\
 &\quad - (\mathbf{b}^T \Sigma e^{-\tau_2 A^T} \mathbf{c}) \mathbf{b}^T \Sigma e^{\tau A^T} \\
 &\quad + s (\mathbf{b}^T \Sigma e^{-\tau_2 A^T} \mathbf{c}) \int_0^{\tau_2} d\tau_1 \mathbf{b}^T \Sigma e^{(\tau-\tau_1)A^T} \\
 &\quad \left. + s \left( \int_0^{\tau_2} d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c} \right) (\mathbf{b}^T \Sigma e^{(\tau-\tau_2)A^T}) \right\}. \tag{106}
 \end{aligned}$$

The derivative with respect to  $v$  is more involved: using Eq. (95) first leads to

$$\begin{aligned}
 \partial_v \psi^{(2)} &= -\theta(\tau)s \left\{ C_{in}(\tau) - (\mathbf{b}^T \Sigma e^{-\tau A^T} \mathbf{c}) (\mathbf{b}^T \Sigma \mathbf{c}) \right. \\
 &\quad - s \left[ \int_0^\tau d\tau_1 C_{in}(\tau - \tau_1) - (\mathbf{b}^T \Sigma e^{-\tau A^T} \mathbf{c}) \int_0^\tau d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c} \right] \\
 &\quad + \int_0^\tau d\tau_2 \left[ -(\mathbf{b}^T \Sigma \mathbf{c}) \mathbf{b}^T \Sigma e^{-\tau_2 A^T} A^T \mathbf{c} \right. \\
 &\quad \left. - (\mathbf{b}^T \Sigma e^{-\tau_2 A^T} \mathbf{c}) \mathbf{b}^T \Sigma A^T \mathbf{c} \right. \\
 &\quad \left. + s (\mathbf{b}^T \Sigma e^{-\tau_2 A^T} \mathbf{c}) \int_0^{\tau_2} d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} A^T \mathbf{c} \right. \\
 &\quad \left. + s \left( \int_0^{\tau_2} d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c} \right) (\mathbf{b}^T \Sigma e^{-\tau_2 A^T} A^T \mathbf{c}) \right\}. \tag{107}
 \end{aligned}$$

Using the identity  $e^{-\tau A^T} A^T = -\frac{d}{d\tau} e^{\tau A^T}$  and integrating by parts simplifies the expression to

$$\begin{aligned}
 \partial_v \psi^{(2)} &= -\theta(\tau)s \left\{ \mathbf{b}^T \Sigma e^{\tau A^T} \mathbf{b} - (\mathbf{b}^T \Sigma \mathbf{c})^2 \right. \\
 &\quad - (\mathbf{b}^T \Sigma A^T \mathbf{c}) \int_0^\tau d\tau_1 (\mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c}) \\
 &\quad - s \left[ \int_0^\tau d\tau_1 \mathbf{b}^T \Sigma e^{\tau_1 A^T} \mathbf{b} \right. \\
 &\quad \left. - (\mathbf{b}^T \Sigma \mathbf{c}) \int_0^\tau d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c} \right] \right\}. \tag{108}
 \end{aligned}$$

Taking the derivative of the last equation with respect to  $\mathbf{k}$  results in

$$\begin{aligned}
 (\nabla_{\mathbf{k}} \partial_v \psi^{(2)})^T &= -\theta(\tau)s \left\{ -2(\mathbf{b}^T \Sigma \mathbf{c}) \mathbf{b}^T \Sigma e^{\tau A^T} \right. \\
 &\quad + \left[ \int_0^\tau d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c} \right] \mathbf{b}^T \Sigma (s \mathbb{I}_d - A^T) e^{\tau A^T} \\
 &\quad \left. + (\mathbf{b}^T \Sigma (s \mathbb{I}_d - A^T) \mathbf{c}) \int_0^\tau d\tau_1 \mathbf{b}^T \Sigma e^{\tau_1 A^T} \right\}, \tag{109}
 \end{aligned}$$

where  $\mathbb{I}_d$  denotes the  $d \times d$  identity matrix.

**Third-order** Because the gradient  $\nabla_{\mathbf{k}} \psi^{(2)}$  vanishes for  $\mathbf{c} = 0$ , we find

$$\bar{P}_n^{(3)}(s) = 0. \tag{110}$$

Using the derivatives in Eqs. (106), (108) and (109) we obtain the third-order perturbation solution

$$\begin{aligned}
 \psi^{(3)} &= -i\theta(\tau) \frac{s}{\mu^4} \int_0^\tau d\tau_3 \left\{ -2(\mathbf{b}^T \Sigma \mathbf{c}) C_{in}(\tau_3) \right. \\
 &\quad + \left( \int_0^{\tau_3} d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c} \right) (\mathbf{b}^T \Sigma (2s \mathbb{I}_d - A^T) e^{\tau_3 A^T} \mathbf{b}) \\
 &\quad + (\mathbf{b}^T \Sigma (2s \mathbb{I}_d - A^T) \mathbf{c}) \left( \int_0^{\tau_3} d\tau_1 C_{in}(\tau_1) \right) \\
 &\quad - (\mathbf{b}^T \Sigma e^{-\tau_3 A^T} \mathbf{c}) [C_{in}(\tau_3) - (\mathbf{b}^T \Sigma \mathbf{c})^2 \\
 &\quad - (\mathbf{b}^T \Sigma A^T \mathbf{c}) \int_0^{\tau_3} d\tau_1 (\mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c}) \\
 &\quad - s \left( \int_0^{\tau_3} d\tau_1 C_{in}(\tau_1) - (\mathbf{b}^T \Sigma \mathbf{c}) \int_0^{\tau_3} d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c} \right)] \\
 &\quad - s^2 \int_0^{\tau_3} d\tau_2 \left[ (\mathbf{b}^T \Sigma e^{-\tau_2 A^T} \mathbf{c}) \int_0^{\tau_2} d\tau_1 C_{in}(\tau_3 - \tau_1) \right. \\
 &\quad \left. + \left( \int_0^{\tau_2} d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c} \right) C_{in}(\tau_3 - \tau_2) \right] \\
 &\quad + s (\mathbf{b}^T \Sigma e^{-\tau_3 A^T} \mathbf{c}) \int_0^{\tau_3} d\tau_2 [C_{in}(\tau_2) - (\mathbf{b}^T \Sigma e^{-\tau_2 A^T} \mathbf{c}) (\mathbf{b}^T \Sigma \mathbf{c}) \\
 &\quad - s \left( \int_0^{\tau_2} d\tau_1 C_{in}(\tau_1) - (\mathbf{b}^T \Sigma e^{-\tau_2 A^T} \mathbf{c}) \int_0^{\tau_2} d\tau_1 \mathbf{b}^T \Sigma e^{-\tau_1 A^T} \mathbf{c} \right)] \left. \right\}. \tag{111}
 \end{aligned}$$

Differentiating this expression with respect to  $\mathbf{k}$ , setting  $\mathbf{c} = 0$ , and taking the scalar product with the vector  $\mathbf{b}$  leads to

$$\begin{aligned}
 (\nabla_{\mathbf{k}} \psi^{(3)})^T \mathbf{b} \Big|_{\mathbf{c}=0} &= i\theta(\tau)s \int_0^\tau d\tau_3 \left\{ 2C_{in}(\tau_3) C_{in}(\tau) \right. \\
 &\quad + C_{in}(\tau - \tau_3) C_{in}(\tau_3) + \int_0^{\tau_3} d\tau_1 [C_{in}(\tau - \tau_1) C'_{in}(\tau_3) + C_{in}(\tau_1) C'_{in}(\tau)] \\
 &\quad - 2s \int_0^{\tau_3} d\tau_1 [C_{in}(\tau - \tau_1) C_{in}(\tau_3) + C_{in}(\tau_1) C_{in}(\tau) + C_{in}(\tau_1) C_{in}(\tau - \tau_3)] \\
 &\quad + s^2 \int_0^{\tau_3} d\tau_2 \int_0^{\tau_2} d\tau_1 [C_{in}(\tau - \tau_2) C_{in}(\tau_3 - \tau_1) \\
 &\quad + C_{in}(\tau - \tau_1) C_{in}(\tau_3 - \tau_2) + C_{in}(\tau - \tau_3) C_{in}(\tau_2 - \tau_1)] \left. \right\} \tag{112}
 \end{aligned}$$

Here,  $C'_{in}(\tau)$  denotes the derivative of the correlation function with respect to  $\tau$ .

**Fourth-order** Finally, using Eqs. (96) and (112) yields the second non-vanishing correction:

$$\begin{aligned}
 \bar{P}_n^{(4)}(s) &= e^{-sn} s^2 \int_0^n d\tau_4 \int_0^{\tau_4} d\tau_3 \left\{ 2C_{in}(\tau_3) C_{in}(\tau_4) \right. \\
 &\quad + C_{in}(\tau_4 - \tau_3) C_{in}(\tau_3) + \int_0^{\tau_3} d\tau_1 [C_{in}(\tau_4 - \tau_1) C'_{in}(\tau_3) + C_{in}(\tau_1) C'_{in}(\tau_4)] \\
 &\quad - 2s \int_0^{\tau_3} d\tau_1 [C_{in}(\tau_4 - \tau_1) C_{in}(\tau_3) + C_{in}(\tau_1) C_{in}(\tau_4) + C_{in}(\tau_1) C_{in}(\tau_4 - \tau_3)] \\
 &\quad + s^2 \int_0^{\tau_3} d\tau_2 \int_0^{\tau_2} d\tau_1 [C_{in}(\tau_4 - \tau_2) C_{in}(\tau_3 - \tau_1) \\
 &\quad + C_{in}(\tau_4 - \tau_1) C_{in}(\tau_3 - \tau_2) + C_{in}(\tau_4 - \tau_3) C_{in}(\tau_2 - \tau_1)] \left. \right\}. \tag{113}
 \end{aligned}$$

To summarize, the moment-generating function reads up to fourth order in  $\epsilon$ :

$$\bar{P}_n(s) = e^{-sn} \left( 1 + \epsilon^2 s^2 \int_0^n d\tau_2 \int_0^{\tau_2} d\tau_1 C_{in}(\tau_1) \right) + \epsilon^4 \bar{P}_n^{(4)}(s). \tag{114}$$

This formula satisfies the general property  $\bar{P}_n(0) = 1$  of a moment-generation function, which expresses the normalization  $\int_0^\infty dt P_n(t) = 1$  of the  $n$ th-order interval density.

**Cumulants.** Knowing the moment-generating function, Eq. (114), we can compute the cumulants of the  $n$ th-order intervals by means of Eq. (71) up to fourth-order in  $\epsilon$ . As expected, the first cumulant, or mean  $n$ th-order interval, is given by

$$\kappa_{1,n} \equiv \langle t_n \rangle = n, \tag{115}$$

which is consistent with the fact that  $\langle t_n \rangle = \langle T_1 \rangle + \dots + \langle T_n \rangle = n \langle T_i \rangle = n$  (recall that  $\langle T_i \rangle = 1$  in the non-dimensionalized system).

The second cumulant is equal to the variance  $\langle \Delta t_n^2 \rangle$  of the  $n$ th-order intervals, where  $\Delta t_n = t_n - \langle T_i \rangle$ . From the formula for the cumulants, Eq. (71), it can be seen that only the coefficients of the quadratic terms  $s^2$  in Eqs. (113) and (114) enter the expression for the second cumulant ( $\bar{P}_n(s)$  has no linear terms). Thus, we find

$$\begin{aligned} \kappa_{2,n} \equiv \langle \Delta t_n^2 \rangle &= 2\epsilon^2 \int_0^n d\tau_2 \int_0^{\tau_2} d\tau_1 C_{in}(\tau_1) \\ &+ 2\epsilon^4 \int_0^n d\tau_4 \int_0^{\tau_4} d\tau_3 \left\{ 2C_{in}(\tau_3)C_{in}(\tau_4) \right. \\ &+ C_{in}(\tau_4 - \tau_3)C_{in}(\tau_3) \\ &\left. + \int_0^{\tau_3} d\tau_1 [C_{in}(\tau_4 - \tau_1)C'_{in}(\tau_3) + C_{in}(\tau_1)C'_{in}(\tau_4)] \right\} \\ &= 2\epsilon^2 h(n) + 2\epsilon^4 [g^2(n) + C_{in}(n)h(n)]. \end{aligned} \tag{116}$$

The third cumulant only involves the cubic terms in Eq. (113), hence

$$\begin{aligned} \kappa_{3,n} \equiv \langle \Delta t_n^3 \rangle &= 12\epsilon^4 \int_0^n d\tau_3 \int_0^{\tau_3} d\tau_2 \int_0^{\tau_2} d\tau_1 [C_{in}(\tau_3 - \tau_1)C_{in}(\tau_2) \\ &+ C_{in}(\tau_1)C_{in}(\tau_3) + C_{in}(\tau_1)C_{in}(\tau_3 - \tau_2)] \\ &= 12\epsilon^4 g(n)h(n). \end{aligned} \tag{117}$$

*Coefficient of variation and serial correlation coefficient*

**Formulas for a given noise correlation function** The coefficient of variation is given by the ratio of standard deviation and mean of the 1st-order interval (i.e. ISI). If we take

into account that we measure time in units of the mean ISI, we find

$$Cv = \frac{\sqrt{\langle \Delta t_1^2 \rangle}}{\langle t_1 \rangle} = \sqrt{\kappa_{2,1}}, \tag{118}$$

where  $\kappa_{2,1}$  is given by Eq. (116) with  $n = 1$ . Furthermore, we can compute the SCC by using formula (116) for arbitrary  $n$ :

$$\rho_n = \frac{\langle \Delta t_{n+1}^2 \rangle - 2\langle \Delta t_n^2 \rangle + \langle \Delta t_{n-1}^2 \rangle}{2\langle \Delta t_1^2 \rangle}. \tag{119}$$

If we further restrict ourselves to the leading order, we can derive a much simplified expression for the SCC in terms of the noise correlation function. In fact, writing  $\langle \Delta t_n^2 \rangle = 2\epsilon^2 \int_0^n d\tau_2 \int_0^{\tau_2} d\tau_1 C_{in}(\tau_1) + \mathcal{O}(\epsilon^4)$  and accounting for the symmetry of  $C_{in}(\tau)$ , we obtain from Eq. (119)

$$\rho_n = \frac{\int_0^1 d\tau \int_0^1 d\tau' C_{in}(n + \tau - \tau')}{\int_0^1 d\tau \int_0^1 d\tau' C_{in}(\tau - \tau')} + \mathcal{O}(\epsilon^2). \tag{120}$$

In terms of the noise power spectrum, Eq. (70), this expression can be rewritten as

$$\rho_n = \frac{\int df S_{in}(f) \text{sinc}^2(f) e^{-2\pi i f n}}{\int df S_{in}(f) \text{sinc}^2(f)} + \mathcal{O}(\epsilon^2). \tag{121}$$

**Formulas for a given matrix representation of the Ornstein-Uhlenbeck process** To elucidate the general structure of serial correlations, it is instructive to express the SCC in terms of the drift matrix  $A$  and covariance matrix  $\Sigma$  of the OUP. In leading order of  $\epsilon$ , the variance of the  $n$ th-order intervals reads

$$\langle \Delta t_n^2 \rangle = 2\epsilon^2 \left( \mathbf{b}^T \Sigma (A^T)^{-2} \left[ e^{nA^T} - nA^T - \mathbb{I}_d \right] \mathbf{b} \right) + \mathcal{O}(\epsilon^4),$$

where we assume that the drift matrix  $A$  is invertible. Setting  $n = 1$ , we obtain the squared coefficient of variation:

$$Cv^2 = 2\epsilon^2 \mathbf{b}^T \Sigma (A^T)^{-2} \left[ e^{A^T} - A^T - \mathbb{I}_d \right] \mathbf{b} + \mathcal{O}(\epsilon^4). \tag{122}$$

Using Eq. (119), the leading order of the SCC can now be written as

$$\rho_n = \frac{\mathbf{b}^T \Sigma (A^T)^{-2} \left( \mathbb{I}_d - e^{A^T} \right)^2 e^{(n-1)A^T} \mathbf{b}}{2\mathbf{b}^T \Sigma (A^T)^{-2} \left[ e^{A^T} - A^T - \mathbb{I}_d \right] \mathbf{b}} + \mathcal{O}(\epsilon^2). \tag{123}$$

Apparently, the SCC has the simple exponential structure  $\rho_n = \mathbf{c}_1^T e^{(n-1)A^T} \mathbf{c}_2$  with constant vectors  $\mathbf{c}_1$  and  $\mathbf{c}_2$ . That means that the SCC is a linear combination of the entries in the matrix exponential function  $e^{(n-1)A^T}$ . In fact, let  $Q$  be the constant matrix made up by the (generalized) eigenvectors of  $A^T$ . Then,  $A^T$  can be diagonalized (or, more generally, transformed into its Jordan normal form) through the representation  $A^T = QJQ^{-1}$ . Thus, the SCC can be rewritten as

$$\rho_n = \mathbf{c}_1^T Q e^{(n-1)J} Q^{-1} \mathbf{c}_2 = \mathbf{d}_1^T e^{(n-1)J} \mathbf{d}_2, \tag{124}$$

where the vectors  $\mathbf{d}_1 = \mathbf{Q}\mathbf{c}_1$  and  $\mathbf{d}_2 = \mathbf{Q}^{-1}\mathbf{c}_2$  are constant. In the special case, where  $\mathbf{A}^T$  is diagonalizable (i.e.  $\mathbf{A}^T$  possesses  $d$  linearly-independent eigenvectors corresponding to the eigenvalues  $\lambda_i, i = 1, \dots, d$ ), the matrix exponential  $e^{(n-1)\mathbf{J}}$  is diagonal with the diagonal elements  $e^{\lambda_i(n-1)}$ . Thus, the SCC is simply a superposition of the exponentials  $e^{\lambda_i(n-1)}$ . Real eigenvalues contribute to exponential decays with “rate”  $1/\lambda_i$ . On the other hand, complex eigenvalues lead to damped oscillations, which, however, might appear rather irregular as  $e^{\lambda_i(n-1)}$  is sampled at discrete integer values (Bauermeister et al. 2013). In any case, the SCC decays to zero for  $n \rightarrow \infty$  because, by assumption, the real parts of the eigenvalues of the drift matrix are negative.

Cumulative correlations and Fano factor

The exponential structure of the SCC, Eq. (123), also allows us to calculate the cumulative effect of serial correlation as expressed by the sum of all SCCs. Applying the formula for the geometric series,  $\sum_{n=1}^{\infty} e^{(n-1)\mathbf{A}^T} = (\mathbb{I}_d - e^{\mathbf{A}^T})^{-1}$ , we find

$$\sum_{n=1}^{\infty} \rho_n = \frac{\mathbf{b}^T \Sigma (\mathbf{A}^T)^{-2} (\mathbb{I}_d - e^{\mathbf{A}^T}) \mathbf{b}}{2\mathbf{b}^T \Sigma (\mathbf{A}^T)^{-2} [e^{\mathbf{A}^T} - \mathbf{A}^T - \mathbb{I}_d] \mathbf{b}} + \mathcal{O}(\epsilon^2). \tag{125}$$

This sum also determines the long-term variability of the spike train. Indeed, the fundamental relation (Cox and Lewis 1966a)

$$\lim_{t \rightarrow \infty} F(t) = \langle T \rangle \lim_{f \rightarrow 0} S_{\text{out}}(f) = C_V^2 \left( 1 + 2 \sum_{n=1}^{\infty} \rho_n \right) \tag{126}$$

relates the Fano factor  $F(t)$  of the spike count or the power spectrum  $S_{\text{out}}(f)$  of the output spike train to the second order ISI statistics as expressed by the CV and the sum of SCCs. The Fano factor is defined as the variance to mean ratio of the spike count  $N(t)$  in the time window  $(0, t)$ , i.e.  $F(t) = \langle \Delta N(t)^2 \rangle / \langle N(t) \rangle$  (note that, in this definition, the origin of the time window is arbitrary; it does not have to coincide with a spike time). The power spectrum of the output spike train  $x_{\text{out}}(t) = \sum_i \delta(t - t_i)$  is defined as the Fourier transform of the spike train auto-correlation function, i.e.  $S_{\text{out}}(f) = \int d\tau e^{2\pi i f \tau} \langle x_{\text{out}}(t) x_{\text{out}}(t + \tau) \rangle$ . Using the expression for the squared CV, Eq. (122), and the summed SCC, Eq. (125), the long-time limit of the Fano factor is

$$\lim_{t \rightarrow \infty} F(t) = -2\epsilon^2 \mathbf{b}^T \Sigma (\mathbf{A}^T)^{-1} \mathbf{b} + \mathcal{O}(\epsilon^4). \tag{127}$$

In order to obtain an approximation for the Fano factor for arbitrary length of the time window  $t$ , we follow (Middleton et al. 2003): If  $t$  is large and  $\epsilon \ll 1$ , the spike count  $N(t)$  can be approximated by the free membrane potential  $V(t)$  with  $V(0) \in [0, 1)$ . Indeed, for weak

noise we can assume  $\dot{V}(t) > 0$ , so that the  $n$ -th spike time  $t_n$  is equivalent to the crossing of the  $n$ -fold threshold, i.e.  $V(t_n) = n$ . This implies that the deviation  $\delta V(t) = V(t) - N(t)$  is bounded (in fact,  $0 \leq \delta V(t) < 1$ ) and hence,  $N(t) = V(t) + \mathcal{O}(1)$ . This relation still holds if we assume the specific initial condition  $V(0) = 0$ . The trajectory of  $V(t)$  can then be obtained by integrating Eq. (37a):

$$V(t) = t + \epsilon \int_0^t dt' \eta(t'). \tag{128}$$

Thus, the mean and variance are

$$\begin{aligned} \langle V(t) \rangle &= t, \\ \langle \Delta V(t)^2 \rangle &= \epsilon^2 \int_0^t dt' \int_0^t dt'' C_{\text{in}}(t' - t'') = 2\epsilon^2 h(t), \end{aligned}$$

where  $h(t)$  has been defined in Eq. (66). For large times, the Fano factor is hence given by

$$\begin{aligned} F_{\text{large}}(t) &= \frac{\langle \Delta V(t)^2 \rangle}{\langle V(t) \rangle} + \mathcal{O}(t^{-1}) \approx 2\epsilon^2 \frac{h(t)}{t} \\ &= 2\epsilon^2 \mathbf{b}^T \Sigma \left( \frac{(\mathbf{A}^T)^{-2} [e^{t\mathbf{A}^T} - t\mathbf{A}^T - \mathbb{I}_d]}{t} \right) \mathbf{b}. \end{aligned} \tag{129}$$

In the exact limit  $t \rightarrow \infty$  we re-obtain the leading-order term of Eq. (127), which shows the consistency of the two approaches.

For small time windows, the spike train can be regarded as periodic given the weakness of the noise. This argument holds irrespective of the specific correlation function of the noise, which allows us to apply the same approximation as in (Middleton et al. 2003):

$$F_{\text{small}}(t) = \frac{\{t\}(1 - \{t\})}{t}, \tag{130}$$

where  $\{t\} = t - \lfloor t \rfloor$  is the difference between  $t$  and the largest integer  $\lfloor t \rfloor$  not exceeding  $t$  (sawtooth function). Following (Middleton et al. 2003), a uniform approximation on the whole range of  $t$  can be obtained by adding the small and large time approximations of the Fano factor:

$$F(t) \approx F_{\text{small}}(t) + F_{\text{large}}(t). \tag{131}$$

A.3 Model of the synaptic input current

To test our theory (Figs. 3–9), we assume that the pool of presynaptic neurons can be split up into  $N_{\mathcal{E}}$  excitatory and  $N_{\mathcal{I}} = N - N_{\mathcal{E}}$  inhibitory neurons. Furthermore, the postsynaptic synaptic currents induced by each spike are modeled by the exponential impulse response functions  $j_{\mathcal{E}}(t) = J_{\mathcal{E}} e^{-t/\tau_{\mathcal{E}}} \theta(t)$  and  $j_{\mathcal{I}}(t) = J_{\mathcal{I}} e^{-t/\tau_{\mathcal{I}}} \theta(t)$  with synaptic time

constants  $\tau_{\mathcal{E}}$  and  $\tau_{\mathcal{I}}$  (synaptic filtering). This corresponds to the simple first order kinetics

$$\tau_i \dot{I}_i = -I_i + J_i \tau_i \sum_{k=1}^{N_i} x_{i,k}(t), \quad i = \mathcal{E}, \mathcal{I} \tag{132}$$

where  $x_{\mathcal{E}/\mathcal{I},k}(t) = \sum_j \delta(t - t_j^{(\mathcal{E}/\mathcal{I},k)})$  is the spike train arriving at the  $k$ -th excitatory/inhibitory synapse.  $J_{\mathcal{E}}$  and  $J_{\mathcal{I}}$  denote the synaptic efficacies. The mean synaptic current is thus  $\langle I_{\text{syn}}(t) \rangle = N_{\mathcal{E}} J_{\mathcal{E}} \tau_{\mathcal{E}} \nu_{\mathcal{E}} + N_{\mathcal{I}} J_{\mathcal{I}} \tau_{\mathcal{I}} \nu_{\mathcal{I}}$ , where we assumed that excitatory and inhibitory neurons fire with rate  $\nu_{\mathcal{E}}$  and  $\nu_{\mathcal{I}}$ , respectively.

In frequency domain, the synaptic filter becomes

$$|\tilde{j}_i(f)|^2 = \frac{J_i^2 \tau_i^2}{1 + (2\pi \tau_i f)^2} \tag{133}$$

for  $i = \mathcal{E}, \mathcal{I}$ . In the special case, where the single presynaptic spike trains are independent and have power spectra  $S_{\mathcal{E}}(f)$  and  $S_{\mathcal{I}}(f)$ , respectively, we find the power spectrum of the synaptic current from Eq. (27):

$$S_I(f) = N_{\mathcal{E}} |\tilde{j}_{\mathcal{E}}(f)|^2 S_{\mathcal{E}}(f) + N_{\mathcal{I}} |\tilde{j}_{\mathcal{I}}(f)|^2 S_{\mathcal{I}}(f). \tag{134}$$

#### A.4 Signal encoded in a common rate modulation of Poissonian inputs

To model the encoding of a continuous signal  $s(t)$ , we assume that  $s(t)$  represents the rate modulation of an inhomogeneous Poisson process. Specifically, we assume that the  $k$ -th presynaptic spike train is a Poisson process with rate  $r_k(t) \equiv \langle x_k(t) \rangle = \nu_k(1 + \varepsilon_k s(t))$ .

Here,  $s(t)$  is common to all inputs and is modeled as a mean-zero Gaussian process with power spectrum  $S_s(f)$ . For the correlation matrix, Eq. (25), we find

$$S_{kl}(f) = \nu_k \delta_{kl} + \varepsilon_k \varepsilon_l \nu_k \nu_l S_s(f).$$

The power spectrum  $S_I(f)$  of the total input current is then given by Eq. (27), and the normalized spectrum is  $S_{\text{in}}(f) = S_I(f) / \int df S_I(f)$ . Specifically, we find for homogeneous excitatory and inhibitory neurons, where only the excitatory population carries a signal ( $\varepsilon_{\mathcal{E}} = \varepsilon > 0, \varepsilon_{\mathcal{I}} = 0$ ):

$$\begin{aligned} \langle I_{\text{syn}} \rangle &= N_{\mathcal{E}} J_{\mathcal{E}} \tau_{\mathcal{E}} \nu_{\mathcal{E}} + N_{\mathcal{I}} J_{\mathcal{I}} \tau_{\mathcal{I}} \nu_{\mathcal{I}} \\ S_I(f) &= N_{\mathcal{E}} \nu_{\mathcal{E}} |\tilde{j}_{\mathcal{E}}(f)|^2 + N_{\mathcal{I}} \nu_{\mathcal{I}} |\tilde{j}_{\mathcal{I}}(f)|^2 \\ &\quad + \varepsilon^2 \nu_{\mathcal{E}}^2 N_{\mathcal{E}}^2 |\tilde{j}_{\mathcal{E}}(f)|^2 S_s(f). \end{aligned} \tag{135}$$

**Powerlaw signal** A signal spectrum with a powerlaw decay can be modeled by Eq. (15) (cf. (Sobie et al. 2011)), where  $\mathcal{N} = \frac{1}{2}(1 - \gamma) / (f_2^{1-\gamma} - \gamma f_1^{1-\gamma})$  is a normalization constant ensuring unit variance, i.e.  $\int df S_s(f) = 1$ . The steepness of the decay is quantified by the exponent  $\gamma$ .

For  $\gamma = 0$ , we recover a flat spectrum, i.e. white noise with a cut-off frequency  $f_2$ . For  $\gamma = 1$ , we obtain  $1/f$ -noise (flicker noise), which possesses strongly increased power at low frequencies. For the sake of the following calculations, we restrict  $\gamma$  to the range  $[0, 1)$ . Furthermore, we assume the lower cut-off to be small,  $f_1 \ll 1$ . Note that in the limit  $f_1 \rightarrow 0$ , we have to require  $\gamma < 1$  in order to have a signal with a finite-variance. Taking this limit allows us to analytically derive the statistical properties of the spike train on long times  $t$  in the range  $1 \ll t \ll f_1^{-1}$ . Nevertheless, in simulations we keep a small lower cut-off  $f_1 > 0$  in order to resolve the low-frequency components.

We are interested in the asymptotic behaviors

$$F(t) = \epsilon^2 t \int df S_{\text{in}}(f) \text{sinc}^2(ft) \tag{136}$$

$$\rho_n \propto \int df S_{\text{in}}(f) \text{sinc}^2(f) e^{-2\pi i f n} \tag{137}$$

for  $1 \ll t \ll f_1^{-1}$ , and  $1 \ll n \ll f_1^{-1}$ , respectively. We will exploit the property that  $\text{sinc}^2(x)$  is negligibly small for  $x > 1$ . Thus, in Eq. (136) only frequencies  $f < 1$  contribute to the integral. Similarly, we can exploit the fact that for large  $n$ , the term  $e^{-2\pi i f n}$  oscillates rapidly for  $f > 1$ . Thus, only small frequencies  $f \ll 1$  contribute to the integral in Eq. (137). This allows us to make the following approximations: First, if  $f_2 > 1$ , we can replace the upper cut-off frequency by infinity, i.e.  $f_2 \rightarrow \infty$ , without changing the integrals in Eqs. (136) and (137). Secondly, if we assume a time scale separation between spiking and fast synaptic filtering dynamics,  $\tau_{\mathcal{E}}, \tau_{\mathcal{I}} \ll 1$ , we can replace the synaptic filter by

$$|\tilde{j}_k(f)|^2 \equiv \frac{J_k^2 \tau_k^2}{1 + (2\pi \tau_k f)^2} \approx J_k^2 \tau_k^2, \quad k = \mathcal{E}, \mathcal{I}, \quad f < 1$$

because only frequencies  $f < 1$  matter in the above integrals. From Eq. (135) then follows that the normalized spectrum has the form

$$S_{\text{in}}(f) = \frac{\sigma_s^2 S_s(f) + S_0}{\sigma_I^2} \tag{138}$$

with  $\sigma_s = \varepsilon_{\mathcal{E}} J_{\mathcal{E}} \tau_{\mathcal{E}} N_{\mathcal{E}} \nu_{\mathcal{E}}$  and  $S_0 = \sum_{k=\mathcal{E}, \mathcal{I}} J_k^2 \tau_k^2 N_k \nu_k$ . Using Eqs. (136) and (138), we find for the Fano factor

$$\begin{aligned} F(t) &= \frac{2\epsilon^2 \sigma_s^2 t}{\sigma_I^2} \int_0^\infty S_s(f) \text{sinc}^2(ft) df + \epsilon^2 \frac{S_0}{\sigma_I^2} \\ &\approx \frac{2\epsilon^2 \sigma_s^2 t \mathcal{N}}{\sigma_I^2} \left( \int_{f_1}^\infty \frac{\text{sinc}^2(ft)}{f^\gamma} df + \frac{\int_0^{f_1} \text{sinc}^2(ft) df}{f_1^\gamma} \right) + \epsilon^2 \frac{S_0}{\sigma_I^2} \\ &\simeq \frac{2\epsilon^2 \sigma_s^2 \mathcal{N}}{\sigma_I^2} t^\gamma \int_{f_1 t}^\infty x^{-\gamma} \text{sinc}^2(x) dx + \mathcal{O}(1), \quad t \gg 1. \end{aligned}$$



If the lower cut-off frequency  $f_1$  is still small enough such that  $f_1 t \ll 1$ , we can set the lower integration boundary to zero. This yields

$$F(t) \simeq \frac{2\epsilon^2 \sigma_s^2 \mathcal{N} \mathcal{C}_1}{\sigma_I^2} t^\gamma + \mathcal{O}(1), \quad 1 \ll t \ll 1/f_1, \quad (139)$$

where  $\mathcal{C}_1 = 2(2\pi)^\gamma \Gamma(-1 - \gamma) \sin(\frac{\pi\gamma}{2})$ .

To get the behavior of the SCC for large lags  $n$ , we can again set  $f_1$  to zero. Ignoring prefactors we find

$$\begin{aligned} \rho_n &\propto \int_0^\infty df \left[ \sigma_s^2 S_s(f) + S_0 \right] \text{sinc}^2(f) \cos(2\pi f n) \\ &\propto \int_0^\infty df f^{-\gamma} \text{sinc}^2(f) \cos(2\pi f n) \\ &\propto (n+1)^{1+\gamma} - 2n^{1+\gamma} + (n-1)^{1+\gamma} \end{aligned}$$

For large  $n$ , we can approximate  $(n \pm 1)^{1+\gamma} \approx n^{1+\gamma} \pm (1+\gamma)n^\gamma + \frac{\gamma(1+\gamma)}{2}n^{\gamma-1}$ , hence

$$\rho_n \propto n^{\gamma-1}, \quad 1 \ll n \ll f_1^{-1}.$$

### A.5 Inverse Gaussian input

To model non-Poissonian input spike trains with a CV smaller or larger than unity, we assume presynaptic spike trains to be renewal processes with an inverse Gaussian ISI density:

$$P(T; \nu, C_V) = \frac{1}{\sqrt{2\pi C_V^2 \nu T^3}} \exp \left[ -\frac{(\nu T - 1)^2}{2C_V^2 \nu T} \right]. \quad (140)$$

This distribution is parametrized by the rate  $\nu$  and the coefficient of variation  $C_V$ . The spike train power spectrum is given by (Stratonovich 1967)

$$S_{IG}(f; \nu, C_V) = \frac{1 - |\tilde{P}(f; \nu, C_V)|^2}{|1 - \tilde{P}(f; \nu, C_V)|^2},$$

where  $\tilde{P}(f; \nu, C_V)$  is the Fourier transform of the inverse Gaussian, Eq. (140), given by

$$\tilde{P}(f; \nu, C_V) = \exp \left[ \frac{1 - \sqrt{1 - 4\pi i C_V^2 f / \nu}}{C_V^2} \right]$$

(Holden 1976; Bulsara et al. 1994; Chacron et al. 2005). The power spectrum of the synaptic current is then, according to Eq. (27):

$$\begin{aligned} S_I(f) &= N_E |\tilde{j}_E(f)|^2 S_{IG}(f; \nu_E, C_V) \\ &\quad + N_I |\tilde{j}_I(f)|^2 S_{IG}(f; \nu_I, C_V). \end{aligned}$$

### A.6 Synapses with short-term plasticity

Following (Dittman et al. 2000; Merkel and Lindner 2010), we use a deterministic model for the synaptic dynamics. A presynaptic spike at the  $\ell$ th synapse causes a postsynaptic response with the history-dependent weight  $A_\ell(t) = F_\ell(t)D_\ell(t)$ , where facilitation is governed by

$$\begin{aligned} F_\ell(t) &= F_0 + ((1 - F_0)^{-1} + F_{C,\ell}^{-1}(t))^{-1}, \\ \dot{F}_{C,\ell}(t) &= -F_{C,\ell}(t)/\tau_F + \Delta \cdot x_\ell(t), \end{aligned}$$

and where depression dynamics obey

$$\dot{D}_\ell(t) = (1 - D_\ell(t))/\tau_D - F_\ell(t^-)D_\ell(t^-) \cdot x_\ell(t).$$

Here,  $x_\ell(t)$  is the presynaptic spike train,  $F_0$  is the intrinsic release probability,  $\Delta$  quantifies the strength of facilitation and  $\tau_F$  ( $\tau_D$ ) is the timescale of facilitation (depression). Above, we have compared the limit cases of pure facilitation and pure depression, which corresponds to setting  $D_\ell(t) \equiv 1$  (pure facilitation) or  $F_\ell(t) \equiv F_0$  (pure depression).

### A.7 Nonlinear integrate-and-fire models

Using the PRC, Eq. (18), the deviation of the  $i$ th ISI from its mean is given by

$$\delta T_i \equiv T_i - \langle T \rangle = -\sigma \int_0^{(T)} dt Z(t) \eta(t_{i-1} + t).$$

This allows us to compute the serial correlations:

$$\begin{aligned} \langle \delta T_i \delta T_{i+k} \rangle &= \sigma^2 \int_0^{(T)} dt \int_0^{(T)} dt' Z(t) Z(t') \langle \eta(t_{i-1} + t) \eta(t_{i+k-1} + t') \rangle \\ &= \sigma^2 \int_0^{(T)} dt \int_0^{(T)} dt' Z(t) Z(t') C_{in}(t_{i+k-1} - t_{i-1} + t' - t) \\ &= \sigma^2 \int_0^{(T)} dt \int_0^{(T)} dt' Z(t) Z(t') C_{in}(t_k + t' - t). \end{aligned}$$

If  $C_{in}(t)$  does not change rapidly, we can approximate  $t_k \approx k\langle T \rangle$  in the last equation. Hence

$$\begin{aligned} \langle \delta T_i \delta T_{i+k} \rangle &\approx \sigma^2 \int_0^{(T)} dt \int_0^{(T)} dt' Z(t) Z(t') C_{in}(k\langle T \rangle + t' - t) \\ &= \sigma^2 \langle T \rangle^2 \int df S_{in}(f) |\tilde{Z}(f)|^2 e^{-2\pi i f \langle T \rangle k}, \end{aligned}$$

where  $\tilde{Z}(f) = \frac{1}{\langle T \rangle} \int_0^{(T)} dt Z(t) e^{2\pi i f t}$ . From this follow Eqs. (19) and (20).

The sum of all SCCs evaluates to

$$\sum_{n=1}^\infty \rho_n = \frac{1}{2} \left( \frac{\sigma^2 \tilde{Z}(0)^2}{C_V^2 \langle T \rangle} S(0) - 1 \right),$$

where  $\tilde{Z}(0)$  is equal to the average PRC obtained by averaging over the interval  $[0, \langle T \rangle]$ . Furthermore, using Eq. (126)

we obtain the long-time limit of the Fano factor or, equivalently, the spike train power spectrum in the zero-frequency limit as follows

$$\lim_{t \rightarrow \infty} F(t) = \langle T \rangle \lim_{f \rightarrow 0} S_{\text{out}}(f) = \frac{\sigma^2 \tilde{Z}(0)^2}{\langle T \rangle} S(0).$$

This equation implies that for any neural oscillator driven by weak colored noise, the spike count diffusion coefficient  $D_{\text{eff}} = S_{\text{out}}(0)/2$  is proportional to the intensity  $D_{\eta} = S(0)/2$  of the input noise, irrespective of the nonlinearity.

## References

- Alijani, A., & Richardson, M. (2011). Rate response of neurons subject to fast or frozen noise: from stochastic and homogeneous to deterministic and heterogeneous populations. *Physical Review E*, *84*, 011,919–1.
- Baddeley, R., Abbott, L.F., Booth, M.C.A., Sengpiel, F., Freeman, T., Wakeman, E.A., & Rolls, E.T. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proceedings of the Royal Society of London B*, *264*, 1775.
- Badel, L., Lefort, S., Brette, R., Petersen, C.C., Gerstner, W., & Richardson, M.J. (2008). Dynamic I-V curves are reliable predictors of naturalistic pyramidal-neuron voltage traces. *Journal of Neurophysiology*, *99*(2), 656.
- Bair, W., Koch, C., Newsome, W., & Britten, K. (1994). Power spectrum analysis of bursting cells in area MT in the behaving monkey. *Journal of Neurophysiology*, *14*, 2870.
- Bauermeister, C., Schwalger, T., Russell, D., Neiman, A., & Lindner, B. (2013). Characteristic effects of stochastic oscillatory forcing on neural firing statistics: theory and application to paddlefish electroreceptor afferents. *PLoS Computational Biology*, *9*(8), e1003,170.
- Brenner, N., Agam, O., Bialek, W., & de Ruyter van Steveninck R. (2002). Statistical properties of spike trains: universal and stimulus-dependent aspects. *Physical Review E*, *66*, 031,907.
- Brunel, N. (2000). Sparsely connected networks of spiking neurons. *Journal of Computational Neuroscience*, *8*, 183.
- Brunel, N., & Sergi, S. (1998). Firing frequency of leaky integrate-and-fire neurons with synaptic currents dynamics. *Journal of Theoretical Biology*, *195*, 87.
- Brunel, N., Chance, F.S., Fourcaud, N., & Abbott, L.F. (2001). Effects of synaptic noise and filtering on the frequency response of spiking neurons. *Physical Review Letters*, *86*, 2186.
- Bulsara, A., Lowen, S.B., & Rees, C.D. (1994). Cooperative behavior in the periodically modulated Wiener process: noise-induced complexity in a model neuron. *Physical Review E*, *49*, 4989.
- Burkitt, A.N. (2006). A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biological Cybernetics*, *95*, 1–19.
- Buzsáki, G., & Draguhn, A. (2004). Neural oscillations in cortical networks. *Science*, *304*, 1926.
- Câteau, H., & Reyes, A.D. (2006). Relation between single neuron and population spiking statistics and effects on network activity. *Physical Review Letters*, *96*, 058,101.
- Chacron, M.J., Longtin, A., St-Hilaire, M., & Maler, L. (2000). Suprathreshold stochastic firing dynamics with memory in P-type electroreceptors. *Physical Review Letters*, *85*, 1576.
- Chacron, M.J., Longtin, A., & Maler, L. (2005). Delayed excitatory and inhibitory feedback shape neural information transmission. *Physical Review E*, *72*, 051,917.
- Compte, A., Constantinidis, C., Tegner, J., Raghavachari, S., Chafee, M.V., Goldman-Rakic, P.S., & Wang, X.J. (2003). Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task. *Journal of Neurophysiology*, *90*(5), 3441.
- Cox, D.R., & Lewis, P.A.W. (1966a). *The statistical analysis of series of events*. London: Chapman and Hall.
- Cox, D.R., & Lewis, P.A.W. (1966b). *The statistical analysis of series of events*. London: Chapman and Hall. chap 4.6.
- Cox, D.R., & Miller, H.D. (1965). *The theory of stochastic processes*: Chapman and Hall.
- Destexhe, A., Rudolph, M., & Paré, D. (2003). The high-conductance state of neocortical neurons *in vivo*. *Nature Reviews Neuroscience*, *4*(9), 739–751.
- Dittman, J.S., Kreitzer, A.C., & Regehr, W.G. (2000). Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. *Journal of Neuroscience*, *20*, 1374.
- Droste, F., & Lindner, B. (2014). Integrate-and-fire neurons driven by asymmetric dichotomous noise. *Biological Cybernetics*, *108*(6), 825–843.
- Droste, F., Schwalger, T., & Lindner, B. (2013). Interplay of two signals in a neuron with heterogeneous synaptic short-term plasticity. *Frontiers in Computational Neuroscience*, *7*, 86.
- Dummer, B., Wieland, S., & Lindner, B. (2014). Self-consistent determination of the spike-train power spectrum in a neural network with sparse connectivity. *Front Comp Neurosci*, *8*, 104.
- Ermentrout, G.B., & Terman, D.H. (2010). *Mathematical foundations of neuroscience*: Springer.
- Fisch, K., Schwalger, T., Lindner, B., Herz, A., & Benda, J. (2012). Channel noise from both slow adaptation currents and fast currents is required to explain spike-response variability in a sensory neuron. *Journal of Neuroscience*, *34*(48), 17,332–17.
- Fortune, E.S., & Rose, G.J. (2001). Short-term synaptic plasticity as a temporal filter. *Trends in Neurosciences*, *24*, 381.
- Fourcaud, N., & Brunel, N. (2002). Dynamics of the firing probability of noisy integrate-and-fire neurons. *Neural Computation*, *14*, 2057.
- Fourcaud-Trocme, N., Hansel, D., van Vreeswijk, C., & Brunel, N. (2003). How spike generation mechanisms determine the neuronal response to fluctuating inputs. *Journal of Neuroscience*, *640*(37), 11,628–11.
- Franklin, J., & Bair, W. (1995). The effect of a refractory period on the power spectrum of neuronal discharge. *SIAM Journal on Applied Mathematics*, *55*, 1074.
- Gerstein, G.L., & Mandelbrot, B. (1964). Random walk models for the spike activity of a single neuron. *Biophysical Journal*, *4*, 41.
- Gerstner, W., & Kistler, W.M. (2002). *Spiking neuron models: single neurons, populations, plasticity*. Cambridge: Cambridge University Press.
- Hänggi, P., & Jung, P. (1995). Colored noise in dynamical systems. *Advances in Chemical Physics*, *89*, 239.
- Holden, A.V. (1976). *Models of the stochastic activity of neurones*. Berlin: Springer.
- Kamke, E. (1965). *Differentialgleichungen, Lösungsmethoden und Lösungen II: partielle Differentialgleichungen erster Ordnung für eine gesuchte Funktion*. Leipzig: Geest & Portig.
- van Kampen, N.G. (1992). *Stochastic processes in physics and chemistry*. North-Holland, Amsterdam.
- Koch, C. (1999). *Biophysics of computation: information processing in single neurons*: Oxford University Press.
- Lerchner, A., Ursta, C., Hertz, J., Ahmadi, M., Ruffiot, P., & Enemark, S. (2006). Response variability in balanced cortical networks. *Neural Computation*, *18*(3), 634.

- Lindner, B. (2004). Interspike interval statistics of neurons driven by colored noise. *Physical Review E*, *69*, 022,901.
- Lindner, B. (2006). Superposition of many independent spike trains is generally not a Poisson process. *Physical Review E*, *73*, 022,901.
- Lindner, B., Gangloff, D., Longtin, A., & Lewis, J.E. (2009). Broad-band coding with dynamic synapses. *Journal of Neuroscience*, *29*(7), 2076–2088.
- Liu, Y.H., & Wang, X.J. (2001). Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron. *Journal of Computational Neuroscience*, *10*, 25.
- London, M., Roth, A., Beeren, L., Häusser, M., & Latham, P.E. (2010). Sensitivity to perturbations *in vivo* implies high noise and suggests rate coding in cortex. *Nature*, *466*(7302), 123.
- Lowen, S.B., & Teich, M.C. (1992). Auditory-nerve action potentials form a nonrenewal point process over short as well as long time scales. *Journal of the Acoustical Society of America*, *92*, 803.
- Merkel, M., & Lindner, B. (2010). Synaptic filtering of rate-coded information. *Physical Review E*, *92*(4 Pt 1), 041,921–041.
- Middleton, J.W., Chacron, M.J., Lindner, B., & Longtin, A. (2003). Firing statistics of a neuron model driven by long-range correlated noise. *Physical Review E*, *68*, 021,920.
- Moreno-Bote, R., & Parga, N. (2004). Role of synaptic filtering on the firing response of simple model neurons. *Physical Review Letters*, *92*(2), 028102.
- Moreno-Bote, R., & Parga, N. (2006). Auto- and crosscorrelations for the spike response of leaky integrate-and-fire neurons with slow synapses. *Physical Review Letters*, *96*(2), 028,101. doi:10.1103/PhysRevLett.96.028101.
- Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., & Pouget, A. (2014). Information-limiting correlations. *Nature Neuroscience*, *17*(10), 1410.
- Nawrot, M.P., Boucsein, C., Rodriguez-Molina, V., Aertsen, A., Grun, S., & Rotter, S. (2007). Serial interval statistics of spontaneous activity in cortical neurons *in vivo* and *in vitro*. *Neurocomp*, *70*, 1717.
- Peterson, A.J., Irvine, D.R.F., & Heil, P. (2014). A model of synaptic vesicle-pool depletion and replenishment can account for the interspike interval distributions and nonrenewal properties of spontaneous spike trains of auditory-nerve fibers. *Journal of Neuroscience*, *34*(45), 15,097.
- Pozzorini, C., Naud, R., Mensi, S., & Gerstner, W. (2013). Temporal whitening by power-law adaptation in neocortical neurons. *Nature Neuroscience*, *16*(7), 942–948.
- Richardson, M.J.E. (2008). Spike-train spectra and network response functions for non-linear integrate-and-fire neurons. *Biological Cybernetics*, *99*(4–5), 381.
- Risken, H. (1984). *The Fokker-Planck equation*. Berlin: Springer.
- Rosenbaum, R., Rubin, J., & Doiron, B. (2012). Short term synaptic depression imposes a frequency dependent filter on synaptic information transfer. *PLoS Computational Biology*, *8*(6).
- Salinas, E., & Sejnowski, T.J. (2002). Integrate-and-fire neurons driven by correlated stochastic input. *Neural Computation*, *14*, 2111.
- Schwalger, T., & Lindner, B. (2013). Patterns of interval correlations in neural oscillators with adaptation. *Frontiers in Computational Neuroscience*, *7* (164).
- Schwalger, T., & Schimansky-Geier, L. (2008). Interspike interval statistics of a leaky integrate-and-fire neuron driven by Gaussian noise with large correlation times. *Physical Review E*, *77*, 031,914–9.
- Schwalger, T., Fisch, K., Benda, J., & Lindner, B. (2010). How noisy adaptation of neurons shapes interspike interval histograms and correlations. *PLoS Computational Biology*, *6*(12), e1001,026. doi:10.1371/journal.pcbi.1001026.
- Schwalger, T., Miklody, D., & Lindner, B. (2013). When the leak is weak – how the first-passage statistics of a biased random walk can approximate the ISI statistics of an adapting neuron. *European Physical Journal Spec Topics*, *222*(10), 2655.
- Sobie, C., Babul, A., & de Sousa R. (2011). Neuron dynamics in the presence of 1/f noise. *Physical Review E*, *83*(5), 051,912.
- Stratonovich, R.L. (1967). *Topics in the theory of random noise, vol 1*. New York: Gordon and Breach.
- Wang, X.J. (1998). Calcium coding and adaptive temporal computation in cortical pyramidal neurons. *Journal of Neurophysiology*, *79*(3), 1549–1566.
- Wang, X.J., Liu, Y., Sanchez-Vives, M.V., & McCormick, D.A. (2003). Adaptation and temporal decorrelation by single neurons in the primary visual cortex. *Journal of Neurophysiology*, *89*(6), 3279–3293.