The Journal of Mathematical Neuroscience
a SpringerOpen Journal

RESEARCH   Open Access

Shifting Spike Times or Adding and Deleting Spikes—How Different Types of Noise Shape Signal Transmission in Neural Populations

Sergej O. Voronenko · Wilhelm Stannat · Benjamin Lindner

Abstract We study a population of spiking neurons which are subject to independent noise processes and a strong common time-dependent input. We show that the response of output spikes to independent noise shapes information transmission of such populations even when information transmission properties of single neurons are left unchanged. In particular, we consider two Poisson models in which independent noise either (i) adds and deletes spikes (AD model) or (ii) shifts spike times (STS model). We show that in both models *suprathreshold stochastic resonance* (SSR) can be observed, where the information transmitted by a neural population is increased with addition of independent noise. In the AD model, the presence of the SSR effect is robust and independent of the population size or the noise spectral statistics. In the STS model, the information transmission properties of the population are determined by the spectral statistics of the noise, leading to a strongly increased effect of SSR in some regimes, or an absence of SSR in others. Furthermore, we observe a high-pass filtering of information in the STS model that is absent in the AD model. We quantify information transmission by means of the lower bound on the mutual information rate and the spectral coherence function. To this end, we derive the signal–output cross-spectrum, the output power spectrum, and the cross-spectrum of two spike trains for both models analytically.

S.O. Voronenko (✉) · W. Stannat · B. Lindner
Bernstein Center for Computational Neuroscience, 10115 Berlin, Germany
e-mail: sergej@physik.hu-berlin.de

S.O. Voronenko · B. Lindner
Department of Physics, Humboldt University, 12489 Berlin, Germany

W. Stannat
Institut für Mathematik, TU Berlin, 10587 Berlin, Germany

1 Introduction

Neurons in the sensory periphery encode information about continuous time-dependent signals in sequences of action potentials. Hereby, upon repeated presentation of a stimulus, the response of the neuron is not perfectly reproducible but exhibits trial-to-trial variability. Processes, leading to such variability, are termed *noise* and can have various origins [1, 2]. How such noise processes affect the transmission of time-dependent signals in neurons can be studied in the framework of information theory [3, 4]. Within this framework, it has been shown, for instance, that the presence of noise can enhance the transmission of weak (subthreshold) signals in single neurons and neural models [5–7], an effect known as *stochastic resonance* and also observed outside biology [8, 9]. At the level of neural population coding, noise can also have a beneficial role for the transmission of strong (suprathreshold) signals [10, 11] by means of *suprathreshold stochastic resonance* (SSR), the mechanism of which is quite distinct from that of conventional stochastic resonance despite the similarity in their naming. Additionally, noise not only impacts the total transmitted information, but it also affects which frequencies of the sensory signal are preferably encoded by a neural system. The suppression of information about the input signal in certain frequency bands can be regarded as a form of information filtering [12–16]. Put differently, we may ask whether the neural system is preferentially encoding slow (low-frequency) components of a signal or fast (high-frequency) components of a signal, which can be quantified by the coherence function, as described below.

How noise affects information transmission in neural populations has been studied for a long time [11, 17, 18]. Of particular interest in the context of the information flow through a population are the correlations among neurons that have been observed in many experimental preparations, e.g. in the visual system [19–22], the somatosensory system [23], the olfactory system [24, 25], the barrel cortex of rats [26, 27], and in spinal motor neurons [28, and references therein]. Such correlations, either in membrane potential, in output spikes, or in spike counts of two cells, can be caused by a common input to both cells due to overlapping receptive fields. For instance, in the electrosensory system [29], the spontaneous activity of different neurons in the absence of the signal is uncorrelated and is driven by independent noise processes. In other systems, the output correlations are not caused by a stimulus. For example, in tangential neurons of the fly visual system, already the noise processes are correlated and lead, even in the absence of the sensory signal, to a spontaneous spiking activity that is correlated across different neurons [20] (for a detailed discussion of the noise sources see [30]). Other examples of neurons receiving common noise input are ganglion cells of the primate retina [21] or the projection neurons of the *Drosophila* olfactory system [25]. In the present study, we consider ensembles of neurons receiving highly correlated noise input as sketched in Fig. 1.

We consider two theoretical models of neural populations that exhibit strong spike train correlations among the neurons within the population, even in the absence of
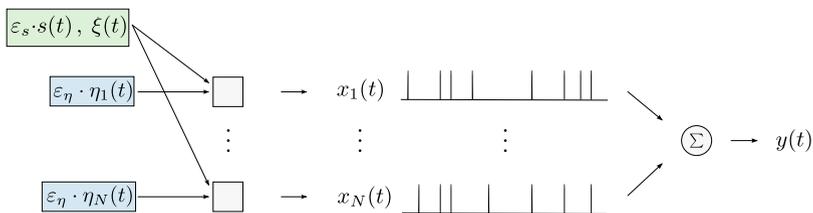
Fig. 1 A population of N uncoupled neurons is driven by common processes (*green box*) and independent noise processes (*blue boxes*). The common processes consist of a strong common noise $\xi(t)$ and a common sensory signal $s(t)$, which is multiplied with a small positive scaling parameter $\varepsilon_s \ll 1$. Independent noise processes $\eta_\mu(t)$ are multiplied with another small positive scaling parameter $\varepsilon_\eta \ll 1$. In the special case of $\varepsilon_s = 0$, $\varepsilon_\eta \neq 0$ the sensory signal is absent and the N neurons exhibit spontaneous activity. Due to the strong common noise ξ the spontaneous activity is highly correlated. In the special case of $\varepsilon_s \neq 0$, $\varepsilon_\eta = 0$ all neurons generate identical output spike trains encoding the sensory signal in the time-dependent firing rate. The output of the population is quantified by the sum $y(t)$ of the individual spike trains

a sensory signal. In this situation, we address the question of how the spike trains of different neurons may be decorrelated by independent noise processes and how this affects the transmission of a sensory signal. More specifically, we are interested in how independent noise influences the spikes of the output spike trains and study two extreme cases. In one case, we assume that independent noise adds and deletes spikes in the output spike trains (AD model) as illustrated in Fig. 2a. This is a likely effect of additional noise in an excitable neuron with low firing rate. In another case, we assume that independent noise shifts the spike times of the output spike trains (STS model) as illustrated in Fig. 2b. This scenario applies to neurons in a tonically firing regime, which generally do not fire with Poisson statistics. We construct the two models in such a way that they cannot be distinguished on a single neuron level. This allows us to ascribe any differences in the information transmission properties of the populations unambiguously to the different effects of the noise.

This work is organized as follows: First, we describe the methods by which we will study the effect of noise on signal transmission in a population of spiking neurons. Second, we introduce two models where independent noise either adds and deletes spikes, or shifts spike times in the output spike trains. In Sect. 4, we then derive the spectral statistics for the two models. These derivations can be skipped upon the first reading. In Sect. 5, we summarize the derived spectral statistics and proceed to study the effect of independent noise on information filtering and the total transmission of information in neural populations. We conclude with a summary and a discussion of our results in Sect. 6.

2 Methods

2.1 Spike Train Statistics & Ensemble Averages

In this paper, we study the transmission of a sensory time-dependent signal by a population of spiking neurons, which is illustrated in Fig. 1. We model the output spike trains of single neurons by stochastic point processes. The output of the μth

a

$x_1(t)$

$x_2(t)$

b

$x_1(t)$

$x_2(t)$
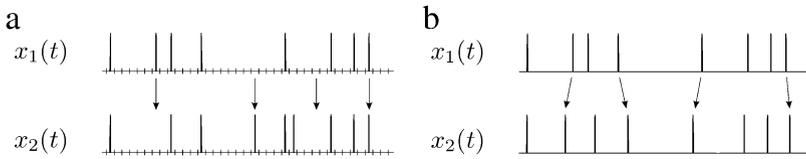
Fig. 2 Two neurons are driven by a strong common noise ξ, a weak common signal s, and independent noise processes η_μ. **a** Addition and deletion of spikes: Independent noise processes lead to addition and deletion of spikes by weakly modulating the threshold value $r_\mu(t)$ in Eq. (11) independently for both neurons. *The first arrow* indicates the deletion of a spike, *the second arrow* indicates the addition of a spike, and *arrows three and four* indicate time bins where there is no change in the spike trains. **b** Spike time shifting: Independent noise leads to shifting of spikes by weakly modulating the integrand $r_\mu(t)$ in Eq. (16) independently for both neurons, but no spikes are added or deleted. *The arrows* exemplify corresponding spikes in the two spike trains that have been shifted in time

stochastic point process can be described by the spike count $n_\mu(t)$. This function starts at 0 at $t = 0$ and is incremented by 1 at each spike time $t_{\mu,k}$, i.e. $n_\mu(t) = 0$ for $0 \le t \le t_{\mu,1}$, $n_\mu(t) = 1$ for $t_{\mu,1} \le t \le t_{\mu,2}$, and so forth. Equivalently, the output of a stochastic point process can be described by the derivative of $n_\mu(t)$. This derivative is called the spike train and is given by a sum of delta functions,

$$x_\mu(t) = \frac{d}{dt} n_\mu(t) = \sum_k \delta(t - t_{\mu,k}). \tag{1}$$

We study information transmission properties of the population by quantifying the amount of information about the input signal $s(t)$ encoded in the sum

$$y(t) = \sum_{\mu=1}^{N} x_\mu(t) \tag{2}$$

of the individual output spike trains.

We take into account different sources of variability: common noise $\xi(t)$, independent noise sources $\eta_\mu(t)$, and the stochastic signal $s(t)$ (cf. Fig. 1). Consequently, we can consider different ensemble averages, denoted by angular brackets $\langle \cdots \rangle$. Subscripts indicate over which processes we average and the absence of subscripts implies averaging over all involved processes. In mathematical terms this notation corresponds to the expectation with respect to the conditional distribution that is indicated by the subscripts, e.g. $\langle x_\mu(t) \rangle_\xi$ stands for the expectation of the process $x_\mu(t)$ with respect to the conditional distribution of ξ, conditioned on a realization of s and η, whereas $\langle x_\mu(t) \rangle$ stands for the total expectation. Note that $\langle x_\mu(t) \rangle_\xi$ is still a random process, unless a realization of s and η is fixed. Below, when analyzing correlation functions, e.g. Eq. (7), we will also consider averages over products of spike trains $\langle x_\mu(s) x_\nu(t) \rangle_\xi$, which in mathematical terms corresponds to

$$\frac{d}{ds} \frac{d}{dt} \langle n_\mu(\bar{s}) n_\nu(\bar{t}) \rangle_\xi \Big|_{\bar{s}=s, \bar{t}=t}.$$

This applies analogously to averages over the processes $s(t)$ and $y(t)$.

The instantaneous firing rate

$$\langle x_\mu(t) \rangle_\xi,$$

obtained by averaging the spike train *only* with respect to the common noise $\xi(t)$, will be an important quantity in our calculations. It still depends on the independent noise and the signal and is difficult to determine in experiments. More accessible is the average over all noise sources by repeated trials with a frozen stimulus and summation over all spike trains. In this way, we obtain (apart from a normalization factor $1/N$) the population rate

$$\rho(t) = \langle y(t) \rangle_{\xi,\eta}. \tag{3}$$

An example for a signal, spike trains, and the resulting population rate is shown in Fig. 3.

2.2 Information Transmission & Spectral Statistics

In the case of ergodic processes, the total amount of the information about a signal $s(t)$ transmitted by the output $y(t)$ can be quantified by the *mutual information rate* R [3], which is measured in bits per second. For Gaussian signals, a lower bound on the mutual information rate [4, 31, 32] is given by

$$R_{\mathrm{lb}} = -\int_0^\infty df \, \log_2\big(1 - C_{ys}(f)\big). \tag{4}$$

The coherence function $C_{ys}(f)$ between the input signal $s(t)$ and the output $y(t)$ is calculated from second-order spectral measures of input and output and is defined as

$$C_{ys}(f) = \frac{|S_{ys}(f)|^2}{S_{yy}(f)S_{ss}(f)}. \tag{5}$$

Here S_{yy} and S_{ss} are the summed-spike-train and signal power spectra, respectively, and S_{ys} is the signal–output cross-spectrum. The numerical estimation of spectra follows standard procedures [33]. In our analytical calculations we will use the Wiener–Khinchin theorem [34]

$$S_{z_1 z_2}(f) = \int_{-\infty}^\infty d\tau \, e^{i2\pi f\tau} K_{z_1 z_2}(\tau) \quad \text{with } z_{1,2} \in \{x_\mu, x_\nu, y, s\}, \tag{6}$$

that relates the spectra to the correlation functions in the time domain

$$K_{z_1 z_2}(\tau) = \lim_{t\to\infty}\big(\langle z_1(t+\tau)z_2(t)\rangle - \langle z_1(t+\tau)\rangle\langle z_2(t)\rangle\big). \tag{7}$$

The limit of large times in Eq. (7) ensures stationarity. For the summed spike train Eq. (2), the autocorrelation function with $z_1 = z_2 = y$ in Eq. (7), can be rewritten as

$$K_{yy}(\tau) = \lim_{t\to\infty}\sum_{k,l}\big(\langle x_k(t+\tau)x_l(t)\rangle - \langle x_k(t+\tau)\rangle\langle x_l(t)\rangle\big)$$

$$= NK_{xx}(\tau) + N(N-1)K_{x_\mu x_\nu}(\tau), \tag{8}$$

where $K_{xx}(\tau)$ is the spike train autocorrelation function and $K_{x_\mu x_\nu}(\tau)$ the cross-correlation function between two spike trains. Analogously, the signal–output cross-correlation function can be written as

$$K_{ys}(\tau) = \lim_{t\to\infty}\langle y(t+\tau)s(t)\rangle = \lim_{t\to\infty}\sum_k \langle x_k(t+\tau)s(t)\rangle$$

$$= N K_{xs}(\tau), \tag{9}$$

where $K_{xs}(\tau)$ is the cross-spectrum between the input signal and a single output spike train. Taking the Fourier transformation of Eqs. (8) and (9), using the Wiener–Khinchin theorem Eq. (6), and inserting the results into Eq. (5), yield the coherence function

$$C_{ys}(f) = \frac{N|S_{xs}(f)|^2}{(S_{xx}(f) + (N-1)S_{x_\mu x_\nu}(f))S_{ss}(f)}. \tag{10}$$

From Eq. (10) we see that for $N > 1$ the cross-spectrum of two spike trains, $S_{x_\mu x_\nu}$, appears in the denominator of the coherence function and gains significance as N becomes larger. Therefore, an essential theoretical problem is to calculate this cross-spectrum.

As outlined above, the coherence function allows one to estimate the total flow of information through the neural population. However, because $C_{ys}(f)$ enters in a monotonic fashion in Eq. (4), we can also regard the coherence as a frequency-resolved measure of information transfer. Reduction of the coherence in certain frequency bands can be regarded as a form of information filtering, which needs to be distinguished from power filtering. Hence, besides the lower bound R_{lb}, we will also inspect the frequency dependence of the coherence function.

3 Models

The models that we consider in this paper have the following assumptions in common:

(1) Poisson statistics of spontaneous activity;
(2) high correlations among neurons due to strong common noise input;
(3) encoding of a sensory signal in the time-dependent population rate.

For simplicity, we consider a linear encoding of a weak time-dependent signal. This will allow to use the lower bound on the mutual information rate as an approximation for the total transmitted information. Note that, although already a single Poisson process can show conventional stochastic resonance [35], with our linear encoding paradigm we exclude this possibility. In our models, the signal transmission in a single neuron is always degraded by noise.

In our theoretical model, we assume that all neurons fire to zeroth-order in complete synchrony and a weak noise input, which is independent for every neuron, leads to a decorrelation of the output spike trains. For simplicity, we assume that for each

neuron the independent noise process and the sensory signal are additive. Both, the sensory signal and the independent noise signals, are modeled by Gaussian processes with unit variance and zero mean.

The considered models can be regarded as *inhomogeneous Poisson processes* [36], which are rate-modulated by a common signal $\varepsilon_s s(t)$ and an independent noise $\varepsilon_\eta \eta(t)$. Such processes are examples of a *doubly stochastic process* [37] or a *Cox process* and are a special case of the *inhomogeneous Bernoulli process* [38]. The simplicity of the considered models will allow us to characterise the information transfer of weak time-dependent signals analytically. Note that the assumptions (1)–(3) made above describe, in good approximation, spiking in specific sensory systems, e.g. in tangential neurons of the fly visual system [20, 39, 40]. The additional modifications that make up the differences between our two models can be regarded as additional operations on the spike trains in the form of thinning (or the opposite of it) and the introduction of an operational time [37, 41].

Before we introduce in detail the two models sketched in Fig. 2, it is worth to note that, for weak stimuli and weak independent noise, these models possess the same signal–output cross-spectrum S_{xs}, the same power spectrum S_{xx}, and the same time-dependent output firing rate. Therefore, for $N = 1$ the coherence function and the information rate are identical for both models. The models are mainly distinguished by how independent noise affects the spikes of the output spike trains, which results in different cross-spectra $S_{x_\mu x_\nu}$ of two spike trains. This setup allows us to study how the response of spikes to noise affects information transmission in neural populations, while keeping all other potential influences on signal transmission unchanged.

### 3.1 Addition and Deletion Model (AD Model)

In the following, we introduce the model for a population of spiking neurons where independent noise adds or deletes spikes. First, we discretize the time axis into bins of width Δt. We generate a spike in the jth time bin in the μth spike train, whenever the following condition is fulfilled:

$$\xi(\Delta t \cdot j) < \Delta t \cdot r_\mu(\Delta t \cdot j) \quad \text{with} \tag{11}$$
$$r_\mu(t) = r_0 \cdot \big(1 + \varepsilon_s s(t) + \varepsilon_\eta \eta_\mu(t)\big) \quad \text{and} \quad r_0 = \text{const} > 0.$$

The common noise process ξ is uniformly distributed in $[0, 1]$ and uncorrelated in time. The spikes are assigned the height $1/\Delta t$ such that the discrete spike train reads

$$x_\mu(\Delta t \cdot j; \Delta t) = \sum_k \frac{1}{\Delta t} \Theta\big(\Delta t/2 - |\Delta t \cdot j - t_{\mu,k}|\big), \tag{12}$$

where $\Theta(z)$ is the Heaviside function (implementing the indicator function) and the second argument of x_μ indicates the time-discretized version of the spike train. Here $t_{\mu,k}$ is the midpoint of the time bin where the kth spike of the μth spike train was generated. In the limit $\Delta t \to 0$ the spike train $x_\mu(t; \Delta t)$ approximates the sum of δ-functions $x_\mu(t)$ given by Eq. (1).
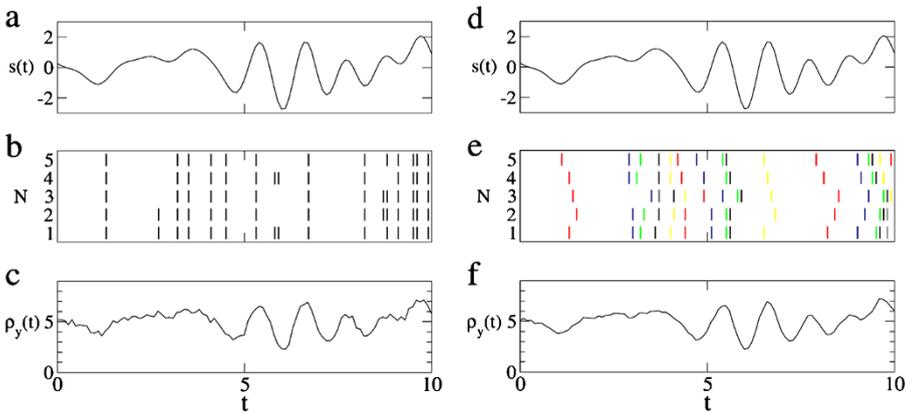
Fig. 3 Simulations of the population response for the AD model (**a–c**) and the STS model (**d–f**). Panels **a** and **d** show the same realization of an input signal. Panels **b** and **e** show 5 neurons of a population for the two models subject to the signal realization from panels **a** and **d** and the same noise realizations (not shown). In the AD model (panel **b**) independent noise leads to addition and deletion of spikes. In the STS model (panel **e**) independent noise leads to shifting of spikes, as indicated by *the coloring*. Panels **c** and **f** show the population response Eq. (3) to the signal shown in **a** and **d**, averaged over 5000 realizations of the noise processes. The model parameters were $r_0 = 1$ Hz, $f_\ell = 0.1$ Hz, $f_u = 1$ Hz, $\varepsilon_s = \varepsilon_\eta = 0.2$

We can compute the ensemble average of the spike train over the common noise ξ

$$\langle x_\mu(t; \Delta t)\rangle_\xi = \frac{1}{\Delta t} \int_0^1 d\xi\, \Theta\big(\Delta t \cdot r_\mu(t) - \xi\big)$$

$$= \begin{cases} 0 & \text{if } r_\mu(t) < 0, \\ r_\mu(t) & \text{if } 0 \le r_\mu(t) \le \frac{1}{\Delta t}, \\ \frac{1}{\Delta t} & \text{if } r_\mu(t) > \frac{1}{\Delta t}. \end{cases} \tag{13}$$

The average $\langle x_\mu(t)\rangle_\xi$ is conditioned on specific realizations of the processes s and η_μ. As we show explicitly in Appendix A in Eq. (57), averaging additionally over the independent noise and the signal, one finds in the limit of $\Delta t \to 0$

$$\langle x_\mu(t)\rangle = r_0 + \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big). \tag{14}$$

Throughout the paper, we will consider the limit $(\varepsilon_s^2 + \varepsilon_\eta^2) \to 0$, such that we can neglect correction terms like the one in the above equation.

In the left column of Fig. 3, we show how a sensory signal is encoded in the population firing rate $\rho_y(t)$ of a population of five AD neurons and how the output spike trains of the neurons are modulated by independent noise.

3.2 Spike-Time-Shifting Model (STS Model)

Next, we introduce the model for a population of spiking neurons where independent noise shifts the spike times of the output spike trains. To zeroth-order the N neurons

of the population generate identical spike trains

$$x_0(t) = \sum_k \delta(t - h_k), \tag{15}$$

which we model by a homogeneous Poisson process with mean firing rate r_0 and spike times h_k. For the μth neuron, the times h_k are transformed into new spike times $t_{k,m}$ via the transformation

$$r_0 \cdot h = \int_0^t d\tau r_\mu(\tau), \tag{16}$$

with $r_\mu(t)$ defined in Eq. (11). For a given spike time h_k, we integrate the right hand side of Eq. (16), until the integral attains the value $r_0 \cdot h_k$ [36]. The resulting integration boundary $t = t_{\mu,k}$ is then the kth spike time of the μth spike train $x_\mu(t)$. In general, due to the different independent noise processes η_μ, the output spike trains $x_\mu(t)$ will be different for each neuron. Hereby, each spike train is an inhomogeneous Poisson spike train with a time-dependent firing rate. The procedure described in this section is equivalent to the simulation of a perfect integrate-and-fire neuron with exponentially distributed thresholds [36]. The time t obtained after the transformation of the time axis h in Eq. (16) is also known as operational time [37, 41].

Although we do not model the underlying noise process explicitly, we think of the homogeneous spike trains in Eq. (15) as a result of a common noise process ξ, analogously to the AD model. By the average $\langle \cdots \rangle_\xi$, we will denote the average over different realizations of the homogeneous Poisson spike trains in Eq. (15).

For a homogeneous Poisson spike train that is transformed according to Eq. (16) with $r_\mu(t) > 0$, the average over the spike train for a fixed realization of the signal and the independent noise reads $\langle x_\mu(t) \rangle_\xi = r_\mu(t)$ [36]. For a process $r_\mu(t)$ that is not bound by zero this is not strictly fulfilled. Hence, ensemble averages over the spike train will contain correction terms that are proportional to the square root of the probability that $r_\mu(t)$ is smaller than zero, which we calculated in Appendix A in Eq. (52). Consequently, using Eq. (11), we obtain for the averaged spike train

$$\begin{aligned} \langle x_\mu(t) \rangle &= \langle r_\mu(t) \rangle + \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big) \\ &= r_0 + \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big), \end{aligned} \tag{17}$$

which in the limit $(\varepsilon_s^2 + \varepsilon_\eta^2) \to 0$ leads to the same mean firing rate as for the AD model Eq. (14) in the limit of $\Delta t \to 0$.

A simulation of five spike trains of the STS population, driven by a common noise process ξ, a common signal s, and independent noise processes η_μ, is shown Fig. 3e. Note that the modulation in Eq. (16) is very distinct from adding jitter to the single spike times, as is considered in [42–44], in that the modulation of the spike times presented here preserves the order of the spikes in each spike train. Other models that incorporate the deletion of spikes in a Poisson spike train [45] or a combination of deletion and shifting as in the thinning and shifting model [42, 44], differ from the models presented here in that the single spike trains of those models are homogeneous spike trains with constant rates. However, the models in the present paper

are designed such that the single spike trains have a prescribed time-dependent firing rate $\langle x(t) \rangle_\xi$, which still depends on the realization of the signal s and the individual noise η. The cross-correlations between spike trains are a consequence of the different implementations of the time-dependent firing rate and are not prescribed a priori as in [42, 44, 45]. Even if the deletion or shifting of spikes in the thinning and shifting model is performed on a rate-modulated mother process, the resulting process would not be equivalent to the AD model or STS model, in which the addition and deletion of spikes and the shifting of spike times are not independent of the signal realization. In particular, the thinning and shifting model of a population of daughter processes for which the stimulus is solely encoded in the firing rate of the mother process cannot exhibit suprathreshold stochastic resonance.

3.3 Modeling the Common Signal and the Independent Noise Processes

The sensory signal s and the independent noise sources η_μ are modeled by Gaussian stochastic processes with zero mean and unit variance. For simplicity, we choose for both, the signal and the independent noise, a flat power spectrum,

$$S_{ss}(f) = S_{\eta\eta}(f) = \begin{cases} \frac{1}{2(f_u - f_\ell)} & \text{if } |f| \in [f_\ell, f_u], \\ 0 & \text{else,} \end{cases} \tag{18}$$

where f_ℓ and f_u are lower and upper cutoff frequencies, respectively. Throughout the paper, we will consider a finite upper cutoff frequency and a non-vanishing lower cutoff frequency. As we will show in our analytical calculation below, the cross-spectrum for two spike trains of the STS model is finite only for $f_\ell > 0$. A realization of the common signal s is shown in Fig. 3a and 3d.

3.4 Simulations

In contrast to the AD model, the numerical measurement of the statistics of the STS model requires a careful choice of simulation parameters. Depending on the shape of the cross-spectrum between different spike trains for the STS model, one has to choose a large simulation time to ensure stationarity and a very small time discretization to be able to resolve correlations between spike trains on small time scales. Furthermore, the coherence function systematically depends on the number of realizations used for the numerical averaging of the spectral statistics. The values of the time discretization Δt, the total simulation time T, and the number of realizations N_{av} used for the numerical averaging of the spectral statistics are reported in Table 1.

4 Derivation of Spectral Measures

4.1 Input–Output Cross-spectrum

In this section, we calculate the spectral measures that are necessary to quantify information transmission properties of the populations. We start by considering the

Table 1 Parameters used in numerical simulations

Figure	Δt in seconds	T in seconds	N_{av}
Fig. 3	$1 \cdot 10^{-1}$	$1 \cdot 10^3$	$5 \cdot 10^3$
Fig. 4	$5 \cdot 10^{-6}$	$1 \cdot 10^2$	$5 \cdot 10^2$
Fig. 5	$1 \cdot 10^{-4}$	$1 \cdot 10^2$	$5 \cdot 10^4$
Fig. 6a	$5 \cdot 10^{-5}$	$3 \cdot 10^1$	$1 \cdot 10^5$
Fig. 6b	$3 \cdot 10^{-5}$	$2 \cdot 10^2$	$1 \cdot 10^4$
Fig. 6c	$1 \cdot 10^{-4}$	$3 \cdot 10^2$	$1 \cdot 10^4$
Fig. 6d	$3 \cdot 10^{-5}$	$1 \cdot 10^3$	$2 \cdot 10^3$
Fig. 7a STS	$5 \cdot 10^{-5}$	$2 \cdot 10^2$	$2 \cdot 10^4$
Fig. 7a AD	$5 \cdot 10^{-5}$	$3 \cdot 10^1$	$1 \cdot 10^5$
Fig. 7b STS	$5 \cdot 10^{-5}$	$1 \cdot 10^3$	$2 \cdot 10^3$
Fig. 7b AD	$1 \cdot 10^{-4}$	$3 \cdot 10^2$	$1 \cdot 10^4$
Fig. 8	$2 \cdot 10^{-4}$	$1 \cdot 10^3$	$2 \cdot 10^4$
Fig. 9	$2 \cdot 10^{-4}$	$1 \cdot 10^2$	–

input–output cross-correlation function

$$
\begin{aligned}
K_{xs}(\tau) &= \big\langle x_\mu(t+\tau)s(t) \big\rangle_{\xi,\eta_\mu,s} \\
&= \big\langle \big\langle x_\mu(t+\tau) \big\rangle_\xi s(t) \big\rangle_{\eta_\mu,s} \\
&= \big\langle r_\mu(t+\tau)s(t) \big\rangle_{\eta_\mu,s} + \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2+\varepsilon_\eta^2))} \big).
\end{aligned}
$$

The correction terms can be derived in complete analogy to the calculation in Appendix A. The first term in the above equation can be calculated using Eq. (11) and the fact that s and η_μ are Gaussian processes with unit variance and zero mean, which leads to

$$
K_{xs}(\tau) = r_0 \varepsilon_s K_{ss}(\tau) + \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2+\varepsilon_\eta^2))} \big).
$$

In the limit $(\varepsilon_s^2 + \varepsilon_\eta^2) \to 0$, keeping only the first-order term in ε_s, the correction term in the above equation can be neglected. Then, after a Fourier transformation, we find the input–output cross-spectrum

$$
S_{xs}(f) \approx r_0 \varepsilon_s S_{ss}(f), \tag{19}
$$

which is equal for both models.

4.2 Cross-spectrum for Two Spike Trains for the AD Model

The cross-correlation function between two spike trains is defined as

$$
K_{x_\mu x_\nu}(\tau) = \big\langle x_\mu(t+\tau)x_\nu(t) \big\rangle - \big\langle x_\mu(t+\tau) \big\rangle \big\langle x_\nu(t) \big\rangle, \tag{20}
$$

where $x_\mu(t)$ and $x_\nu(t)$ are different spike trains of a population with $\mu \neq \nu$. The ensemble averages in the above equation are taken over four stochastic processes: The common noise ξ, the common signal s, and the independent noise processes η_μ

and η_v. Employing Eq. (14), we can write the second term in Eq. (20) as

$$\langle x_\mu(t+\tau)\rangle\langle x_v(t)\rangle = r_0^2 + \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2+\varepsilon_\eta^2))}\big).$$

The first term in Eq. (20) can be interpreted as a probability density [4]. Choosing a discrete variant of the spike train $x(t; \Delta t)$ as introduced in Eq. (12), this leads to

$$\langle x_\mu(t+\tau; \Delta t) x_v(t; \Delta t)\rangle$$
$$= \frac{1}{(\Delta t)^2} \Pr\big(\text{spike in ST } \mu \text{ in } [t+\tau-\Delta t/2, t+\tau+\Delta t/2] \text{ and}$$
$$\text{spike in ST } v \text{ in } [t-\Delta t/2, t+\Delta t/2]\big)$$

(Pr stands for probability and ST stands for spike train). As we generate the spike trains in discrete time steps, we first consider the cross-correlation function between two spike trains with a finite time discretization $\tau = j \cdot \Delta t$ with $j = 0, \pm 1, \ldots$. Splitting the expression in the above equation into two parts, one for $\tau = 0$ and one for $|\tau| > \Delta t/2$, we obtain

$$\langle x_\mu(t+\tau; \Delta t) x_v(t; \Delta t)\rangle$$
$$= \frac{1}{(\Delta t)^2} \Theta\big(\Delta t/2 - |\tau|\big) \cdot P_{syn} + \frac{1}{(\Delta t)^2} \Theta\big(|\tau| - \Delta t/2\big) \cdot P_{asyn}(\tau), \quad (21)$$

with

$$P_{syn} = \Pr\big(\text{spike in ST } \mu \text{ in } [t-\Delta t/2, t+\Delta t/2] \text{ and}$$
$$\text{spike in ST } v \text{ in } [t-\Delta t/2, t+\Delta t/2]\big)$$

and

$$P_{asyn}(\tau) = \Pr\big(\text{spike in ST } \mu \text{ in } [t+\tau-\Delta t/2, t+\tau+\Delta t/2] \text{ and}$$
$$\text{spike in ST } v \text{ in } [t-\Delta t/2, t+\Delta t/2] | |\tau| > \Delta t/2\big).$$

Note that, due to stationarity of the stochastic signals and spike trains, the probabilities in Eq. (21) do not depend on t. As described in Sect. 3.1, the values of realizations of the process ξ at different times are independent of each other, which allows us to average both spike trains separately leading to

$$\frac{1}{(\Delta t)^2} P_{asyn}(\tau) = \big\langle \langle x_\mu(t+\tau; \Delta t)\rangle_\xi \langle x_v(t; \Delta t)\rangle_\xi \big\rangle_{s,\eta_\mu,\eta_v}$$
$$= \langle r_\mu(t+\tau) r_v(t)\rangle_{s,\eta_\mu,\eta_v} + \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2+\varepsilon_\eta^2))}\big).$$

Using Eq. (11) in the above equation and employing that $\eta_\mu(t)$ and $\eta_v(t)$ are independent Gaussian processes with zero mean we obtain in the limit $\Delta t \to 0$

$$\frac{1}{(\Delta t)^2} P_{asyn}(\tau) = r_0^2 \cdot \big(1 + \varepsilon_s^2 K_{ss}(\tau)\big) + \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2+\varepsilon_\eta^2))}\big). \quad (22)$$

From the definition of the AD model in Eq. (11), we can infer that the probability of observing a synchronous spike in two spike trains equals the probability that the

thresholds $\Delta t \cdot r_\mu(\Delta t \cdot j)$ and $\Delta t \cdot r_\nu(\Delta t \cdot j)$ are both higher than the realization of the common noise variable $\xi(\Delta t \cdot j)$. Then, dropping the time arguments, the probability of synchronous spiking can be expressed as an average over two theta functions

$$P_{syn} = \langle \Theta(\Delta t \cdot r_\mu - \xi)\Theta(\Delta t \cdot r_\nu - \xi)\rangle_{\xi, s, \eta_\mu, \eta_\nu}. \tag{23}$$

As is shown in Eq. (59) in Appendix B, we can write the above expression for weak sensory signals as

$$P_{syn} = \Delta t \cdot r_0 \cdot \left(1 - \frac{\varepsilon_\eta}{\sqrt{\pi}}\right) + \mathcal{O}\big(\Delta t(\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4}e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big)$$

$$+ \mathcal{O}\big(\sqrt{\Delta t}(\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4}e^{-(1 - r_0\Delta t)^2/(4r_0^2\Delta t^2(\varepsilon_s^2 + \varepsilon_\eta^2))}\big). \tag{24}$$

Inserting Eqs. (22) and (24) in Eq. (21), and taking the limit $\Delta t \to 0$, we obtain

$$\langle x_\mu(t + \tau)x_\nu(t)\rangle = r_0\left(1 - \frac{\varepsilon_\eta}{\sqrt{\pi}}\right)\delta(\tau) + r_0^2 + r_0^2\varepsilon_s^2 K_{ss}(\tau)$$

$$+ \mathcal{O}\big((1 + \delta(\tau))(\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4}e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big),$$

and for the cross-correlation function between two spike trains Eq. (20) we find

$$K_{x_\mu x_\nu}(\tau) = r_0\left(1 - \frac{\varepsilon_\eta}{\sqrt{\pi}}\right)\delta(\tau) + r_0^2\varepsilon_s^2 K_{ss}(\tau)$$

$$+ \mathcal{O}\big((1 + \delta(\tau))(\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4}e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big).$$

In the limit $(\varepsilon_s^2 + \varepsilon_\eta^2) \to 0$, keeping terms up to second-order in ε_s and ε_η, we can neglect the correction terms in the above equation and find the following cross-spectrum between two spike trains:

$$S_{x_\mu x_\nu}(f) \approx r_0 \cdot \left(1 - \frac{\varepsilon_\eta}{\sqrt{\pi}}\right) + r_0^2\varepsilon_s^2 S_{ss}(f). \tag{25}$$

The above equation shows that by adding and deleting spikes the weak independent noise sources lead to a decorrelation of the two spike trains with a uniform decrease of power at all frequencies proportional to ε_η. The analytical result for the cross-spectrum of two spike trains Eq. (25) for the AD model is compared with simulations in Fig. 4. Note that because the cross-correlation function between two spike trains is symmetric with respect to τ, the cross-spectrum is real-valued for all frequencies.

4.3 Cross-spectrum for Two Spike Trains for the STS Model

In this section, we calculate the cross-spectrum $S_{x_\mu x_\nu}$ between spike trains μ and ν for the STS model. We first consider the autocorrelation function

$$K_{xx}^0(t'' - t') = \lim_{t'', t' \to \infty}\big(\langle x_0(t'')x_0(t')\rangle_\xi - \langle x_0(t'')\rangle_\xi \langle x_0(t')\rangle_\xi\big)$$

of a homogeneous Poisson process with constant rate r_0, where we use a slightly different notation than in Eq. (7). The last term in the above equation equals r_0^2. The
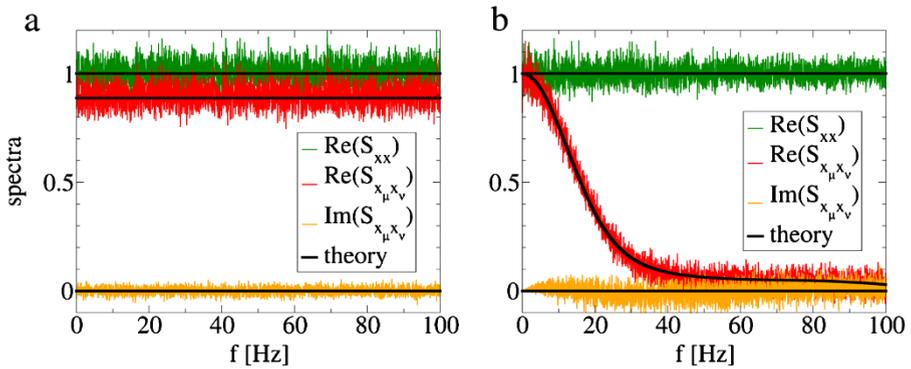
Fig. 4 **a** AD model: Simulation results of the spike train power spectrum (*green*) and the real and imaginary part of the cross-spectrum of two spike trains (*red* and *yellow*, respectively) are compared with the analytical results Eq. (39) and Eq. (25) (*black lines*). **b** STS model: Simulation results of the spike train power spectrum (*green*) and the real and imaginary parts of the cross-spectrum of two spike trains (*red* and *yellow*, respectively) are compared with the analytical results Eq. (39) and Eq. (38) (*black lines*). The model parameters were $r_0 = 1$ Hz, $f_\ell = 0.3$ Hz, $f_u = 100$ Hz, $\varepsilon_s = 0.5$, $\varepsilon_\eta = 0.2$

spike trains inside the first average can be expressed as derivatives of the spike count $n_0(t)$ as in Eq. (1), such that

$$K_{xx}^0\left(t'' - t'\right) = \lim_{t'',t'\to\infty} \frac{d}{d\bar{t}''}\frac{d}{d\bar{t}'}\left\langle n_0(\bar{t}'')n_0(\bar{t}')\right\rangle_\xi \bigg|_{\bar{t}''=t'',\bar{t}'=t'} - r_0^2. \qquad (26)$$

The power spectrum of a homogeneous Poisson process is constant $S_{xx}^0 = r_0$ and implies for the autocorrelation function of a homogeneous Poisson process

$$K_{xx}^0\left(t'' - t'\right) = r_0\delta\left(t'' - t'\right). \qquad (27)$$

Combining Eq. (27) with Eq. (26), we obtain

$$\lim_{t'',t'\to\infty} \frac{d}{d\bar{t}''}\frac{d}{d\bar{t}'}\left\langle n_0(\bar{t}'')n_0(\bar{t}')\right\rangle_\xi \bigg|_{\bar{t}''=t'',\bar{t}'=t'} = r_0\delta\left(t'' - t'\right) + r_0^2. \qquad (28)$$

Now we calculate the cross-correlation function between two spike trains

$$K_{x_\mu x_\nu}\left(t'' - t'\right) = \lim_{t'',t'\to\infty}\left(\left\langle x_\mu(t'')x_\nu(t')\right\rangle - \left\langle x_\mu(t'')\right\rangle\left\langle x_\nu(t')\right\rangle\right) \qquad (29)$$

of the full process, subject to an intrinsic noise ξ, independent noise processes η_μ and η_ν, and an input signal s. Employing Eq. (17), the last term in the above equation can be written as

$$\left\langle x_\mu(t'')\right\rangle\left\langle x_\nu(t')\right\rangle = r_0^2 + \mathcal{O}\left(\left(\varepsilon_s^2 + \varepsilon_\eta^2\right)^{1/4}e^{-1/(4(\varepsilon_s^2+\varepsilon_\eta^2))}\right). \qquad (30)$$

The first term of the cross-correlation function can be recast as before into

$$\lim_{t'',t'\to\infty}\left\langle x_\mu(t'')x_\nu(t')\right\rangle = \lim_{t'',t'\to\infty}\frac{d}{d\bar{t}''}\frac{d}{d\bar{t}'}\left\langle n_\mu(\bar{t}'')n_\nu(\bar{t}')\right\rangle_{\eta_\mu,\eta_\nu,s,\xi}\bigg|_{\bar{t}''=t'',\bar{t}'=t'}. \qquad (31)$$

The rate-modulated Poisson process generated by the STS model is related to a homogeneous Poisson process with constant rate by the time transformation Eq. (16). We use this relation to link the inhomogeneous to the homogeneous spike count via

$$n_\mu(t) = n_0\big(\gamma_\mu(t)\big) \quad \text{with} \quad \gamma_\mu(t) = \int_0^t dt' \big(1 + \varepsilon_s s(t') + \varepsilon_\eta \eta_\mu(t')\big). \tag{32}$$

Using the above relation and Eq. (28), we find

$$\begin{aligned}
&\frac{d}{d\bar{t}''} \frac{d}{d\bar{t}'} \big\langle n_\mu(\bar{t}'') n_\nu(\bar{t}') \big\rangle_\xi \Big|_{\bar{t}''=t'', \bar{t}'=t'} \\
&= \frac{d}{d\bar{t}''} \frac{d}{d\bar{t}'} \big\langle n_\mu(\bar{t}'') n_\nu(\bar{t}') \big\rangle_\xi \Big|_{\bar{t}''=\gamma_\mu(t''), \bar{t}'=\gamma_\nu(t')} \frac{d\gamma_\mu(t'')}{dt''} \frac{d\gamma_\nu(t')}{dt'} \\
&= \big(r_0 \delta\big(\gamma_\mu(t'') - \gamma_\nu(t')\big) + r_0^2\big) \frac{d\gamma_\mu(t'')}{dt''} \frac{d\gamma_\nu(t')}{dt'}.
\end{aligned}$$

Note that the above relation is valid only if r_μ is strictly larger than zero. Hence, we obtain for Eq. (31)

$$\begin{aligned}
&\lim_{t'',t'\to\infty} \big\langle x_\mu(t'') x_\nu(t') \big\rangle \\
&= \lim_{t'',t'\to\infty} \Big\langle \big(r_0 \delta\big(\gamma_\mu(t'') - \gamma_\nu(t')\big) + r_0^2\big) \\
&\quad \times \big(1 + \varepsilon_s s(t'') + \varepsilon_\eta \eta_\mu(t'')\big)\big(1 + \varepsilon_s s(t') + \varepsilon_\eta \eta_\nu(t')\big) \Big\rangle_{\eta_\mu, \eta_\nu, s} \\
&\quad + \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big),
\end{aligned} \tag{33}$$

where the correction term is proportional to the square root of the probability that $r_\mu(t) < 0$ computed in Eq. (52). Using $\langle s \rangle = \langle \eta_\mu \rangle = 0$, employing the relation Eq. (62) derived in Appendix C, and substituting the variables $t = t'$ and $t + \tau = t''$, we transform Eq. (33) into

$$\begin{aligned}
&\lim_{t'',t'\to\infty} \big\langle x_\mu(t'') x_\nu(t') \big\rangle \\
&= \lim_{t'',t'\to\infty} r_0 \big\langle \big\langle \delta\big(\gamma_\mu(t'') - \gamma_\nu(t')\big) \big\rangle_{\eta_\mu, \eta_\nu} \big(1 + \varepsilon_s s(t'')\big)\big(1 + \varepsilon_s s(t')\big) \big\rangle_s \\
&\quad + \varepsilon_s^2 r_0^2 K_{ss}(t'' - t') + r_0^2 + \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big).
\end{aligned} \tag{34}$$

Using the definition of $\gamma_\mu(t)$ Eq. (32), we can write the average over the delta function in Eq. (34) as

$$\big\langle \delta\big(\gamma_\mu(t+\tau) - \gamma_\nu(t)\big) \big\rangle_{\eta_\mu \eta_\nu} = \big\langle \delta\big(\tau + \varepsilon_s \bar{s}(t, \tau) + g(t, \tau)\big) \big\rangle_g \quad \text{with}$$

$$g(t, \tau) = \int_0^{t+\tau} dt_1 \varepsilon_\eta \eta_\mu(t_1) + \int_0^t dt_2 \varepsilon_\eta \eta_\nu(t_2) \quad \text{and}$$

$$\bar{s}(t, \tau) = \int_t^{t+\tau} dt_3 s(t_3).$$

The new stochastic variable g is a sum of two integrals over Gaussian variables and therefore also a Gaussian variable. The average of the delta function over realizations of g is then the probability that g attains the value $\tau + \varepsilon_s \bar{s}(t, \tau)$, and is given by

$$\left\langle \delta\left(\tau + \varepsilon_s \bar{s}(t, \tau) + g(t, \tau)\right)\right\rangle_g = \frac{1}{\sqrt{2\pi\sigma_g^2(t, \tau)}} \exp\left[-\frac{(\tau + \varepsilon_s \bar{s}(t, \tau))^2}{2\sigma_g^2(t, \tau)}\right], \quad (35)$$

where $\sigma_g^2(t, \tau)$ is the variance of $g(t, \tau)$. In Appendix D, Eq. (64), we show that

$$\hat{\sigma}_g^2 = \lim_{t \to \infty} \sigma_g^2(t, \tau) = \frac{\varepsilon_\eta^2}{\pi^2 f_u f_\ell}, \quad (36)$$

for our specific choice of a flat noise power spectrum, introduced in Sect. 3.3. Employing Eqs. (35), (34), and (30) in Eq. (29) and expanding up to second-order in ε_s, we obtain for the cross-correlation function for two spike trains

$$\begin{aligned}
K_{x_\mu x_\nu}(\tau) &= \frac{r_0}{\sqrt{2\pi\hat{\sigma}_g^2}} \exp\left[-\frac{\tau^2}{2\hat{\sigma}_g^2}\right] + \varepsilon_s^2 r_0^2 K_{ss}(\tau) \\
&\quad + \frac{r_0}{\sqrt{2\pi\hat{\sigma}_g^2}} \exp\left[-\frac{\tau^2}{2\hat{\sigma}_g^2}\right] \left(\frac{\varepsilon_s^2}{2\hat{\sigma}_g^2}\left(\frac{\tau^2}{\hat{\sigma}_g^2} - 1\right) \int_0^\tau dt' K_{ss}(t')(\tau - t') \right. \\
&\quad \left. + \varepsilon_s^2 K_{ss}(\tau) - \frac{2\varepsilon_s^2 \tau}{\hat{\sigma}_g^2} \int_0^\tau dt'' K_{ss}(t'') \right) \\
&\quad + \mathcal{O}(\varepsilon_s^4) + \mathcal{O}\left((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\right). \quad (37)
\end{aligned}$$

We note that the linear term in ε_s vanishes due to the zero mean of the Gaussian signal $\langle s \rangle = 0$. Equivalently, all higher-order odd terms in ε_s in Eq. (37) vanish due to the Gaussian nature of the signal (except for the correction term due to realizations of signal and individual noise that lead to $r_\mu(t) < 0$). From Eqs. (36) and (37) it can be seen that for a vanishing lower cutoff frequency of the independent noise spectrum ($f_\ell \to 0$), the variance $\hat{\sigma}_g^2$ diverges and as a consequence of this the cross-correlation between the two spike trains vanishes—only the part that is due to the signal (second term in Eq. (37)) still contributes.

After Fourier transforming Eq. (37) (neglecting the correction terms), we find the cross-spectrum for two spike trains in the STS population,

$$S_{x_\mu x_\nu}(f) \approx S_{x_\mu x_\nu}^0(f) + r_0^2 \varepsilon_s^2 S_{ss}(f) + \varepsilon_s^2 I(f) \quad \text{with} \quad (38)$$

$$S_{x_\mu x_\nu}^0(f) = r_0 \cdot \exp\left[-2f^2\pi^2\hat{\sigma}_g^2\right], \qquad \hat{\sigma}_g^2 = \frac{\varepsilon_\eta^2}{\pi^2 f_u f_\ell}, \quad \text{and}$$

$$I(f) = f^2 \int_{-\infty}^\infty df' \frac{S_{ss}(f')}{f'^2}\left(S_{x_\mu x_\nu}^0(f - f') - S_{x_\mu x_\nu}^0(f)\right).$$

In Fig. 4, the analytical result for the cross-spectrum for two spike trains of the STS model Eq. (38) is compared with simulations. As for the AD model the cross-

spectrum of two spike trains is real valued. In contrast to the AD model Eq. (25), the cross-spectrum of two spike trains for the STS model Eq. (38) exhibits a strong decrease at high frequencies, while it approaches the spike train power spectrum Eq. (39) at low frequencies. Note that, although we derived $S_{x_\mu x_\nu}$ only up to second-order in ε_s, the theory fits the simulation results very well even for $\varepsilon_s = 0.5$.

4.4 Single Spike Train Power Spectrum

In Appendix E in Eqs. (67) and (72), we derive the spike train power spectrum which in the limit of $\Delta t \to 0$ and $(\varepsilon_s^2 + \varepsilon_\eta^2) \to 0$ (keeping terms up to second-order in ε_s and ε_η) is equal for the AD and STS model

$$S_{xx}(f) \approx r_0 + r_0^2 \varepsilon_s^2 S_{ss}(f) + r_0^2 \varepsilon_\eta^2 S_{\eta\eta}(f). \tag{39}$$

For $\varepsilon_s = 0$ and $\varepsilon_\eta = 0$, the power spectrum is flat, as we would expect for homogeneous Poisson spike trains.

5 Information Transmission in Neural Populations

Here, we use the spectral measures derived in the previous section to study information transmission in two neural populations. The populations are constructed in such a way that they both encode the sensory signal in the time-dependent population firing rate, and both exhibit identical single-spike-train power spectra and identical signal–output cross-spectra. The main difference between the populations lies in the effect that independent noise has on the spikes of the output. In one population independent noise adds and deletes spikes (AD model), while in the other independent noise leads to spike-time-shifting (STS model). We quantify the total of the transmitted information about the sensory signal via the lower bound on the mutual information rate Eq. (4),

$$R_{\mathrm{lb}} = -\int_0^\infty df \, \log_2\big(1 - C_{ys}(f)\big),$$

and study information filtering by means of the coherence function Eq. (10),

$$C_{ys}(f) = \frac{N|S_{xs}(f)|^2}{(S_{xx}(f) + (N-1)S_{x_\mu x_\nu}(f))S_{ss}(f)}.$$

The input–output cross-spectrum Eq. (19) and the single spike train power spectrum Eq. (39) read

$$S_{xs}(f) \approx r_0 \varepsilon_s S_{ss}(f),$$
$$S_{xx}(f) \approx r_0 + r_0^2 \varepsilon_s^2 S_{ss}(f) + r_0^2 \varepsilon_\eta^2 S_{\eta\eta}(f),$$

while the different cross-spectra between two spike trains for the two different models are given by Eq. (25), and Eq. (38):

$$\text{AD model:} \quad S_{x_\mu x_\nu}(f) \approx r_0 \cdot \left(1 - \frac{\varepsilon_\eta}{\sqrt{\pi}}\right) + r_0^2 \varepsilon_s^2 S_{ss}(f),$$

STS model: $\quad S_{x_\mu x_\nu}(f) \approx S^0_{x_\mu x_\nu}(f) + r_0^2 \varepsilon_s^2 S_{ss}(f) + \varepsilon_s^2 I(f) \quad$ with

$$S^0_{x_\mu x_\nu}(f) = r_0 \cdot \exp\left[-2f^2\pi^2\hat{\sigma}_g^2\right], \qquad \hat{\sigma}_g^2 = \frac{\varepsilon_\eta^2}{\pi^2 f_u f_\ell}, \quad \text{and}$$

$$I(f) = f^2 \int_{-\infty}^{\infty} df' \frac{S_{ss}(f')}{f'^2} \left(S^0_{x_\mu x_\nu}(f - f') - S^0_{x_\mu x_\nu}(f) \right).$$

In all expression above, we considered the limits $\Delta t \to 0$ and $(\varepsilon_s^2 + \varepsilon_\eta^2) \ll 1$. If the sensory signal is weak compared to the noise processes driving the neurons, as is assumed throughout this paper, the coherence is much smaller than one. This allows us to employ an approximation for the lower bound on the mutual information rate,

$$R_{\mathrm{lb}} \approx \frac{1}{\ln(2)} \int_0^{\infty} df\, C_{ys}(f), \tag{40}$$

in the analytical calculations to obtain simpler expressions. In the subsequent sections, we will study information transmission in populations of AD neurons and STS neurons.

5.1 AD Population

Inserting the single spike train power spectrum Eq. (39), the input–output cross-spectrum Eq. (19), and the cross-spectrum for two spike trains Eq. (25) into Eq. (10), we find for the coherence function of the AD population

$$C_{ys}(f) \approx \frac{r_0 \varepsilon_s^2 S_{ss}(f)}{1 + r_0 \varepsilon_s^2 S_{ss}(f) + (r_0/N)\varepsilon_\eta^2 S_{ss}(f) - ((N-1)/(N\sqrt{\pi}))\varepsilon_\eta}. \tag{41}$$

Here, we used that signal and noise have equal power-spectra $S_{ss}(f) = S_{\eta\eta}(f)$, as described in Sect. 3.3. The coherence function for the AD model is plotted and compared with numerical simulations in Fig. 5. The only dependence of the coherence function Eq. (41) on frequency comes from the signal power spectrum $S_{ss}(f)$. Therefore, for a flat signal power spectrum the coherence function of the AD model is also flat for frequencies $f_\ell \leq |f| \leq f_u$. Consequently, a population of AD neurons can be referred to as a broadband filter of information, because the sum of the output spike trains contains equal amounts of information about different frequency bands of the signal.

Inserting the coherence Eq. (41) into Eq. (40) and employing Eq. (18), we obtain for the lower bound on the mutual information rate of the AD population

$$R_{\mathrm{lb}} \approx \frac{r_0 \varepsilon_s^2}{2\ln(2)} \cdot \left(1 + r_0 \frac{\varepsilon_s^2 + \varepsilon_\eta^2(1/N)}{2(f_u - f_\ell)} - \frac{(N-1)}{N\sqrt{\pi}}\varepsilon_\eta \right)^{-1}. \tag{42}$$

The approximate expression Eq. (42) is compared with simulations for two sets of parameters in Fig. 6.
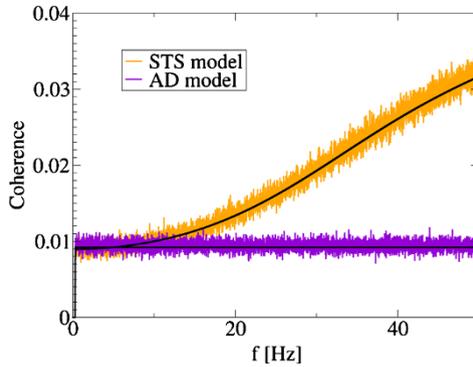
Fig. 5 Simulation results for the spectral coherence function for the AD model (*indigo*) and the STS model (*orange*) for $N = 5$ and a comparison with analytical results Eq. (41) and Eq. (46) (*black lines*). The AD population exhibits a flat coherence function and corresponds to a broadband filter of information. The coherence function of the STS population is monotonously increasing with the frequency and leads to a high-pass filter of information. The model parameters were $r_0 = 10$ Hz, $f_\ell = 0.3$ Hz, $f_u = 50$ Hz, $\varepsilon_s = 0.3$, $\varepsilon_\eta = 0.1$, and $N = 5$

For $N = 1$, the last term in the denominator of Eq. (42) vanishes and the lower bound of the mutual information rate can be simplified as

$$R_{\text{lb}}|_{N=1} \approx \frac{r_0 \varepsilon_s^2}{2\ln(2)} \cdot \left(1 + r_0 \frac{\varepsilon_s^2 + \varepsilon_\eta^2}{2(f_u - f_\ell)}\right)^{-1}. \tag{43}$$

From the above equation, it becomes evident that in a single neuron an increase of the independent noise level can only decrease the lower bound on the mutual information rate. For $N > 1$, additional independent noise ($\varepsilon_\eta > 0$) has a positive effect on information transmission and SSR is observed. The denominator of Eq. (42) is a quadratic function in ε_η and exhibits a minimum at a finite level of independent noise, resulting in a maximum of the lower bound on the mutual information rate. To study the behavior of R_{lb} for weak independent noise, we expand Eq. (42) with respect to ε_η and obtain

$$
\begin{aligned}
R_{\text{lb}} \approx R_{\text{lb}}^0 &+ \frac{R_{\text{lb}}^0}{1 + r_0 \varepsilon_s^2/(2(f_u - f_\ell))} \left(\frac{N-1}{\sqrt{\pi}N}\right) \varepsilon_\eta \\
&+ \frac{R_{\text{lb}}^0}{1 + r_0 \varepsilon_s^2/(2(f_u - f_\ell))} \left(\frac{((N-1)/(\sqrt{\pi}N))^2}{1 + r_0 \varepsilon_s^2/(2(f_u - f_\ell))} - \frac{r_0}{2N(f_u - f_\ell)}\right) \varepsilon_\eta^2
\end{aligned}
\tag{44}
$$

with

$$R_{\text{lb}}^0 = \frac{r_0 \varepsilon_s^2}{2\ln(2)} \left(1 + r_0 \frac{\varepsilon_s^2}{2(f_u - f_\ell)}\right)^{-1}. \tag{45}$$

The linear term in Eq. (44) is always positive. Hence, the population of AD neurons always profits from weak independent noise regardless of the specific choice of model parameters.
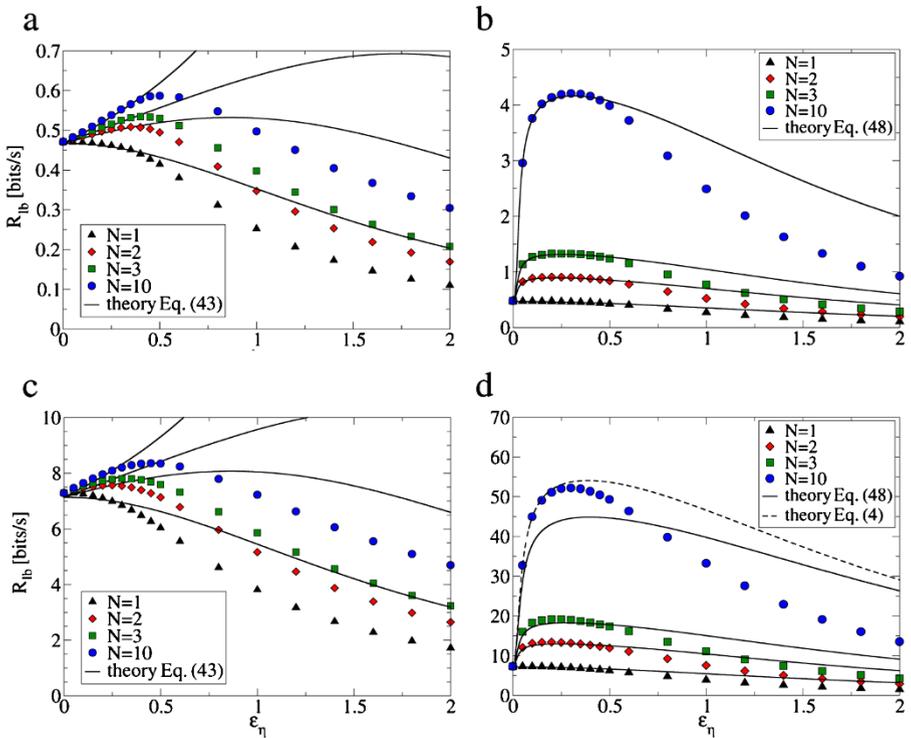
Fig. 6 The lower bound on the mutual information rate R_{lb} for the AD population (**a, c**) and the STS population (**b, d**) as a function of the independent noise level ε_η. The analytical results Eq. (42) and Eq. (47) (*black lines*) are compared with numerical simulations (*circles*) of the lower bound on the mutual information rate for various N for a weak sensory signal with $\varepsilon_s = 0.1$ (panels **a, b**) and a stronger sensory signal with $\varepsilon_s = 0.4$ (panels **c, d**). For $N = 1$ the lower bound on the mutual information rate is identical for both models and is degraded by independent noise. For $N > 1$ both models exhibit SSR, where a non-vanishing level of independent noise increases the lower bound on the mutual information rate. Hereby, the STS population profits significantly more from independent noise than the AD population. For $N = 10$ in panel **d** the theory Eq. (47) fails due to the linearization of the logarithm in Eq. (40). Here we plot Eq. (4), where we inserted the analytically calculated coherence function Eq. (46) and integrated numerically. We note that our analytical theory appropriately describes the increase of the mutual information for weak independent noise levels. The firing rate was $r_0 = 65$ Hz and the signal and noise cutoff frequencies were $f_\ell = 3 \cdot 10^{-2}$ Hz and $f_u = 100$ Hz

5.2 STS Population

Inserting the single spike train power spectrum Eq. (39), the input–output cross-spectrum Eq. (19), and the cross-spectrum for two spike trains Eq. (38) into Eq. (10), we find for the coherence function of the STS population

$$C_{ys}(f) \approx \frac{N r_0 \varepsilon_s^2 S_{ss}(f)}{1 + r_0 \varepsilon_\eta^2 S_{ss}(f) + N r_0 \varepsilon_s^2 S_{ss}(f) + ((N-1)/r_0)(S_{x_\mu x_\nu}^0(f) + \varepsilon_s^2 I(f))},$$

(46)

where $S^0_{x_\mu x_\nu}(f)$ and $I(f)$ are defined in Eq. (38). As for the AD model discussed above, we used that signal and noise have equal power-spectra $S_{ss}(f) = S_{\eta\eta}(f)$. Due to the frequency dependence of the cross-spectrum $S_{x_\mu x_\nu}$, the coherence function also depends strongly on the frequency and exhibits a monotone increase as shown in Fig. 5. Thus, the population of STS neurons can be regarded as a high-pass filter of information, similar to that observed for heterogeneous short-term plasticity [16] or coding by synchrony [13, 15].

In order to understand the high-pass filter effect in the coherence function as well as the stochastic resonance effect discussed below, we note that the cross-correlations between different spike trains contribute largely to the sum's output variability, in particular in the absence of intrinsic noise. This output variability is quantified by the output's power spectrum and appears in the denominator of the coherence function. With individual intrinsic noise, spike times of different neurons are slightly shifted, drastically reducing cross-correlations at high frequencies and thus the amount of the signal-unrelated variability in these frequency bands. Therefore, the coherence function increases with frequency.

Inserting Eq. (46) into Eq. (40) and inserting the noise and signal power spectrum Eq. (18), we find for the lower bound on the mutual information rate of the STS population

$$R_{\text{lb}} \approx \frac{N r_0 \varepsilon_s^2}{2 \ln(2)(f_u - f_\ell)}$$

$$\times \int_{f_\ell}^{f_u} df \left(1 + r_0 \frac{(\varepsilon_\eta^2 + N \varepsilon_s^2)}{2(f_u - f_\ell)} + \frac{N-1}{r_0} \left(S^0_{x_\mu x_\nu}(f) + \varepsilon_s^2 I(f) \right) \right)^{-1}. \quad (47)$$
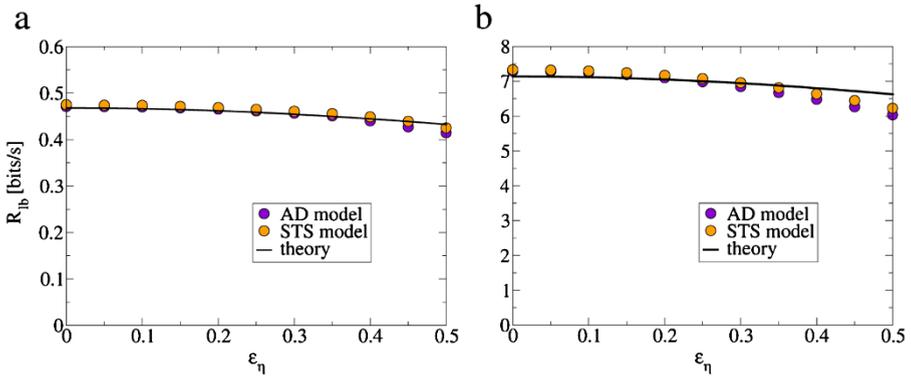
The lower bound on the mutual information rate for the STS population is compared with simulations for two sets of parameters in Fig. 6. We observe that for the given parameters the STS model shows a large SSR effect, while the AD model profits only weakly from additional noise.

For $N = 1$, the frequency dependent term in the integrand of Eq. (47) vanishes and the lower bound on the mutual information rate transforms into

$$R_{\text{lb}}|_{N=1} \approx \frac{r_0 \varepsilon_s^2}{2 \ln(2)} \cdot \left(1 + r_0 \frac{\varepsilon_s^2 + \varepsilon_\eta^2}{2(f_u - f_\ell)} \right)^{-1},$$

which is equal to Eq. (43) for the AD model. We compare the lower bound on the mutual information rate for $N = 1$ for the two models numerically in Fig. 7 for different signal strengths and independent noise levels. For sufficiently low levels of independent noise, there is no difference in the amount of transmitted information for the AD and the STS model on the level of a single neuron. By construction, from the observation of one single spike train it is impossible to distinguish between the two models.

For $N > 1$, additional noise can have a positive effect on information transmission, as illustrated in Fig. 6. Increasing ε_η leads to a decrease of $S^0_{x_\mu x_\nu}$ in the denominator of the integral Eq. (47), as already discussed in the beginning of this section in the context of the high-pass coherence function. However, an increase of ε_η also increases the second term in the denominator of Eq. (47), which is proportional to ε_η^2.

Fig. 7 Comparison of the lower bound on the mutual information rate between the AD model (*indigo*) and the STS model (*orange*) for $N = 1$ for a weak signal with $\varepsilon_s = 0.1$ (panel **a**) and a stronger signal with $\varepsilon_s = 0.4$ (panel **b**). The analytical prediction Eq. (43) (*black line*) is in good agreement with numerical simulations (*colored points*). As predicted by our theory, a single neuron transmits the same amount of information in both models. For strong independent noise and a strong sensory signal we observe a slight deviation of the theory from simulation results, as well as slight differences between the two models. These differences are due to nonlinear effects of the independent noise and signal on the time-dependent firing rate. Model and simulation parameters were as in Fig. 6.

Therefore, whether SSR is observed depends on the specific parameter values chosen. As for the AD model, we expand the lower bound on the mutual information rate Eq. (47) with respect to ε_η and obtain

$$R_{lb} \approx R_{lb}^0 + R_{lb}^0 \left(\frac{(4/3)((N-1)/N)((f_u^3 - f_\ell^3)/(f_u f_\ell))(1 + \varepsilon_s^2) - r_0/N}{2(f_u - f_\ell) + r_0 \varepsilon_s^2} \right) \varepsilon_\eta^2, \tag{48}$$

with R_{lb}^0 defined in Eq. (45). The above expansion illustrates that, when the independent noise vanishes, the lower bound on the mutual information rate is identical for the two models for arbitrary N. The second-order term in Eq. (48) can attain both negative and positive values depending on the choice of the model parameters. The condition that the second-order term becomes negative and that the lower bound on the mutual information rate at $\varepsilon_\eta = 0$ is a decreasing function of ε_η reads

$$\frac{r_0}{N-1} > \frac{4}{3} \frac{f_u^3 - f_\ell^3}{f_u f_\ell} (1 + \varepsilon_s^2).$$

If the above condition is fulfilled, the weak individual noise does not improve the information transmission of a sensory signal and no SSR is observed. Two examples are shown in Fig. 8. In contrast to the AD model, where SSR is always observed for $N > 1$, the occurrence of SSR in the STS model depends on the specific choice of the model parameters.

Using Eq. (44) and Eq. (48), we can find for $\varepsilon_\eta > 0$ and $N > 1$ a noise strength

$$\bar{\varepsilon}_\eta = \sqrt{\pi} \left(\frac{2\pi}{3 f_u f_\ell} \frac{(f_u^3 - f_\ell^3)}{(f_u - f_\ell)} - \frac{N-1}{N(1 + \varepsilon_s^2)} \right)^{-1},$$
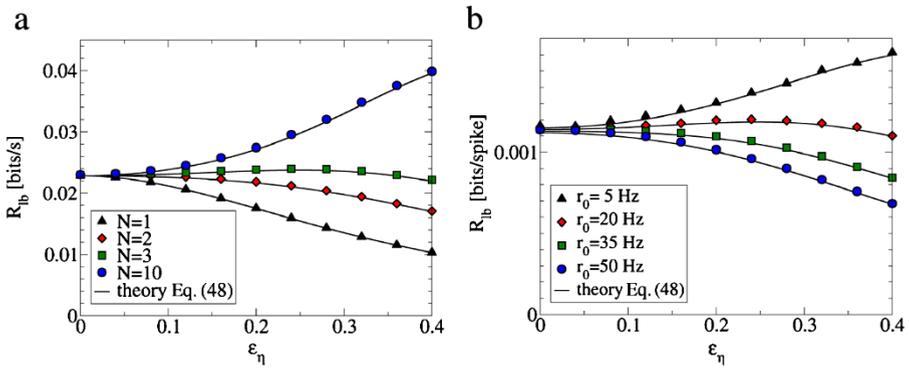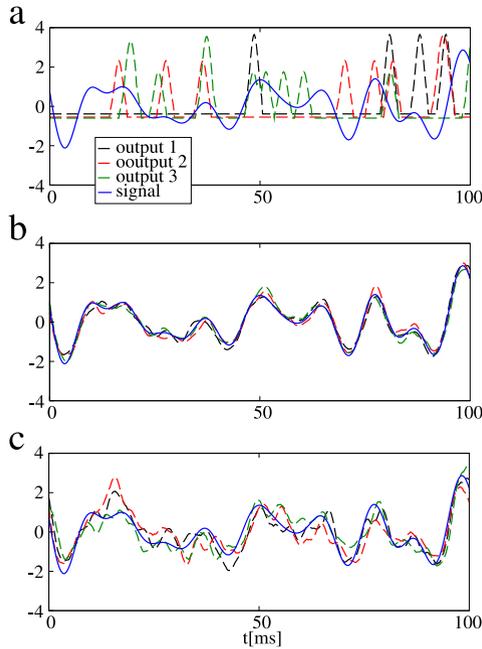
Fig. 8 The presence of SSR in the STS population is parameter dependent, in contrast to the AD population where SSR is always observed for $N > 1$. **a** The lower bound on the mutual information rate for the STS population is plotted for various population sizes as a function of the individual noise level ε_η. SSR is observed for $N \geq 3$ but not for $N = 2$. The model parameters were $r_0 = 20$ Hz, $f_\ell = 0.2$ Hz, $f_u = 1.5$ Hz, and $\varepsilon_s = 0.04$. **b** The lower bound on the mutual information rate for the STS population is plotted in units of bits per spike for various mean firing rates r_0. SSR is observed for $r_0 \leq 20$ Hz but not for $r_0 \geq 35$ Hz. The model parameters were $f_\ell = 0.2$ Hz, $f_u = 1.5$ Hz, $N = 3$, and $\varepsilon_s = 0.04$

for which the lower bound on the mutual information rate is equal for both models. From the above equation we can see that whether the STS population or the AD population transmits more information for a given value of independent noise is mainly determined by the noise and signal cutoff frequencies f_u and f_ℓ.

Finally, let us illustrate in Fig. 9 the stochastic resonance effect when it is most pronounced, namely, in the STS model for a large number of neurons ($N = 1000$) and a high cutoff frequency (except for N, all parameters as in Fig. 6d). In this situation, we consider the low-pass filtered summed output of the population for different levels of the intrinsic noise. Without intrinsic noise (Fig. 9a), the output, i.e. the sum of N perfectly synchronized spike trains, does not resemble the input signal very much. It is important to note that according to Eq. (19) and Eq. (9) an average over many such runs would yield a time series that tracks the input signal closely. However, single runs (red, black, green) in the absence of the intrinsic noise are strongly unreliable. The right amount of intrinsic noise (used in Fig. 9b) desynchronizes the N spike trains, reduces cross-correlations at high frequencies, and thus reduces output variability due to the common noise. Consequently, different realizations of the process for a frozen input signal look more similar and track the input signal reliably (cf. Fig. 9b). However, if we increase intrinsic noise to much higher levels, as in Fig. 9c, this noise itself starts to contribute significantly to the output variability and the reliability of signal transmission is diminished again.

6 Summary and Conclusions

In this paper, we investigated how the effect of noise on the output spikes influences information transmission properties of Poisson neurons. In particular, we considered two populations with strong common input, where in one case weak independent

Fig. 9 We illustrate the emergence of SSR in the STS model for the same parameters as in Fig. 6d. **a** A fixed realization of the signal (*blue*) and three realizations (different common noise realizations) of the output of the STS population for $N = 1000$ and $\varepsilon_\eta = 0$ (*black*, *green*, and *red*). For better visualization the output is convoluted with a Gaussian filter and all outputs and the signal are rescaled to unit variance and zero mean over the time window shown. For vanishing independent noise, the individual spike trains of the population are identical for a fixed realization of the signal and the common noise. In this case signal transmission is not improved by the large population size. **b** Same as in **a** but for $\varepsilon_\eta = 0.5$ (close to the point of stochastic resonance). The individual noise leads to shifting of spikes, such that the convoluted summed output is smoothed. Note that the three realizations of the output are all close to the input signal as well as to each other, indicating a reliable signal transmission. **c** Same as in **a** but for $\varepsilon_\eta = 2$ (far beyond the point of stochastic resonance). Note that our average over a comparatively short time window implies the suppression of long-term variability (corresponding to leaving out low-frequency components of the coherence function)

noise added and deleted spikes, while in the other it shifted spikes. In the limit of a weak sensory signal, we analytically derived the spectral statistics of both models and studied information filtering and the emergence of suprathreshold stochastic resonance (SSR). We showed that, even when single neurons of the AD model and STS model cannot be distinguished by their response statistics, the different effects of independent noise on spikes lead to qualitative and quantitative differences in information transmission on a population level.

In the AD model, the presence of the SSR effect is robust—whenever we consider a population with $N > 1$, a small amount of intrinsic noise has a beneficial effect on the signal transmission. In the STS model, the information transmission properties of the population are determined by the cutoff frequencies of the noise. Depending on the specific parameters, one finds a pronounced SSR in some regimes (exceeding the effect in the AD model by far) or no SSR effect in other regimes. Furthermore, we

observe a high-pass filtering of information in the STS model that is absent in the AD model.

There are a number of studies that explored theoretically the case of weakly correlated neurons and employed perturbation methods to relate output spike train correlations to input correlations [46–52]. In this paper, we have considered the opposite limit of strongly correlated spike trains that are only weakly *decorrelated* due to intrinsic noise sources. In this limit, we were not only able to derive comparatively simple expressions for the cross-correlation between two spike trains but were also able to explore analytically the consequences of these correlations for the transmission of time-dependent signals.

The question arises how the specific choice of the output, which is taken to be the sum of individual spike trains, affects the findings discussed above. The most general approach would be to study the multivariate mutual information between the input signal and the population of output spike trains. This quantity is hard to compute numerically and analytically, and its exact calculation is beyond the scope of this study. However, the mutual information between the input signal and the sum of outputs is a lower bound for the full multivariate mutual information, because the summation can only degrade the information content contained in the entire set of the output spike trains. Additionally, for vanishing individual noise, $\varepsilon_\eta = 0$, all output spike trains are identical and the information content of the population does not differ from the information content of the sum of identical spike trains. Therefore, if the mutual information between the input signal and the summed output increases with individual noise, i.e. exhibits suprathreshold stochastic resonance, the full multivariate mutual information increases as well.

The mutual information between the input signal and the summed output has been estimated here by its lower bound R_{lb}. In our setting with a weak signal that is encoded in the firing rate of the Poisson process, we expect that this bound is rather tight. In fact, for a single inhomogeneous Poisson process, the mutual information and its lower bound coincide in leading-order of the signal amplitude [53].

In this study, we inspected two simple and abstract models for the effect of a weak noise on neural spikes and its consequences on signal transmission by neural populations. We would like to emphasize that the pure limits of an AD model or an STS model approximate the behavior of biophysical neuron models. On one hand, it is plausible that in an excitable neuron model, in which the crossing of a threshold may be aided or prevented by a weak driving, addition and deletion of spikes as in our AD model can be observed. Stochastic oscillators, on the other hand, display a shifting of spike times due to a weak driving, as described by the phase response curve [54]. In between these limits, we expect a combination of both, addition and deletion as well as shifting of spikes. Indeed, such a combination has been observed experimentally [55]. Hence, a generalization of our framework to a Poisson process that includes both effects and allows one to tune gradually between the pure AD and STS models inspected in this paper would be certainly worth additional efforts in a future study.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

Appendix A: Mean Firing Rate of the AD and STS Model

For the AD model the average of the spike train over the intrinsic noise is given by Eq. (13) as

$$\langle x_\mu(t; \Delta t)\rangle_\xi = \frac{1}{\Delta t}\int_0^1 d\xi\, \Theta\big(\Delta t \cdot r_\mu(t) - \xi\big) = \begin{cases} 0 & \text{if } r_\mu(t) < 0, \\ r_\mu(t) & \text{if } 0 \le r_\mu(t) \le \frac{1}{\Delta t}, \\ \frac{1}{\Delta t} & \text{if } r_\mu(t) > \frac{1}{\Delta t}. \end{cases}$$

The average of the spike train over all stochastic processes can now be written as

$$\langle x_\mu(t; \Delta t)\rangle = \langle r_\mu(t)\rangle - \langle r_\mu(t)\Theta\big(-r_\mu(t)\big)\rangle$$
$$- \left\langle \left(r_\mu(t) - \frac{1}{\Delta t}\right)\Theta\left(r_\mu(t) - \frac{1}{\Delta t}\right)\right\rangle. \tag{49}$$

With $r_\mu(t) = r_0 \cdot (1 + \varepsilon_s s + \varepsilon_\eta \eta)$ from Eq. (11) the first term in the above equation gives

$$\langle r_\mu(t)\rangle = r_0 \tag{50}$$

as s and η are Gaussian processes with unit variance and zero mean. For the other terms in Eq. (49), we will show that they are of higher-order in Δt, ε_s, and ε_η. For the second term in Eq. (49), we can find an upper bound using the Cauchy–Schwarz inequality

$$\langle r_\mu(t)\Theta\big(-r_\mu(t)\big)\rangle^2 \le \langle r_\mu(t)^2\rangle\langle\Theta\big(-r_\mu(t)\big)^2\rangle$$
$$= r_0^2\big(1 + \varepsilon_s^2 + \varepsilon_\eta^2\big) \cdot \langle\Theta\big(-r_\mu(t)\big)\rangle. \tag{51}$$

The average in the last line of the above equation is the probability that $r_\mu(t)$ is smaller than zero and is given by

$$\langle\Theta\big(-r_\mu(t)\big)\rangle = \frac{1}{\sqrt{2\pi(\varepsilon_s^2 + \varepsilon_\eta^2)}}\int_1^\infty dz\, e^{-z^2/(2(\varepsilon_s^2 + \varepsilon_\eta^2))}$$
$$= \frac{1}{2}\,\mathrm{erfc}\left(\frac{1}{\sqrt{2(\varepsilon_s^2 + \varepsilon_\eta^2)}}\right)$$

$$\leq \frac{1}{\sqrt{2\pi(\varepsilon_s^2 + \varepsilon_\eta^2)}} \int_1^\infty dz z e^{-z^2/(2(\varepsilon_s^2 + \varepsilon_\eta^2))}$$

$$= \sqrt{\frac{\varepsilon_s^2 + \varepsilon_\eta^2}{2\pi}} e^{-1/(2(\varepsilon_s^2 + \varepsilon_\eta^2))}, \tag{52}$$

which gives for Eq. (51)

$$\langle r_\mu(t) \Theta(-r_\mu(t)) \rangle^2 \leq r_0^2 (1 + \varepsilon_s^2 + \varepsilon_\eta^2) \sqrt{\frac{\varepsilon_s^2 + \varepsilon_\eta^2}{2\pi}} e^{-1/(2(\varepsilon_s^2 + \varepsilon_\eta^2))}. \tag{53}$$

For the third term in Eq. (49) we can find an upper bound using again the Cauchy–Schwarz inequality

$$\left\langle \left(r_\mu(t) - \frac{1}{\Delta t} \right) \Theta \left(r_\mu(t) - \frac{1}{\Delta t} \right) \right\rangle^2$$

$$\leq \left\langle \left(r_\mu(t) - \frac{1}{\Delta t} \right)^2 \right\rangle \left\langle \Theta \left(r_\mu(t) - \frac{1}{\Delta t} \right)^2 \right\rangle$$

$$\leq \left(r_0^2(1 + \varepsilon_s^2 + \varepsilon_\eta^2) + \frac{1}{\Delta t^2} \right) \left\langle \Theta \left(r_\mu(t) - \frac{1}{\Delta t} \right) \right\rangle, \tag{54}$$

where in the last line we have dropped the mixed term that is always negative. Furthermore, note that the average in the last line of Eq. (54) is the probability that $r_\mu(t)$ is larger than $1/\Delta t$ and is given by

$$\left\langle \Theta \left(r_\mu(t) - \frac{1}{\Delta t} \right) \right\rangle = \frac{1}{\sqrt{2\pi(\varepsilon_s^2 + \varepsilon_\eta^2)}} \int_{(1/(r_0\Delta t) - 1)}^\infty dz e^{-z^2/(2(\varepsilon_s^2 + \varepsilon_\eta^2))}$$

$$\leq \frac{1}{\sqrt{2\pi(\varepsilon_s^2 + \varepsilon_\eta^2)}} \frac{r_0\Delta t}{1 - r_0\Delta t} \int_{(1/(r_0\Delta t) - 1)}^\infty dz z e^{-z^2/(2(\varepsilon_s^2 + \varepsilon_\eta^2))}$$

$$= \sqrt{\frac{(\varepsilon_s^2 + \varepsilon_\eta^2)}{2\pi}} \frac{r_0\Delta t}{1 - r_0\Delta t} e^{-(1 - r_0\Delta t)^2/(2r_0^2\Delta t^2(\varepsilon_s^2 + \varepsilon_\eta^2))}, \tag{55}$$

which gives for Eq. (54)

$$\left\langle \left(r_\mu(t) - \frac{1}{\Delta t} \right) \Theta \left(r_\mu(t) - \frac{1}{\Delta t} \right) \right\rangle^2$$

$$\leq \left(r_0^2(1 + \varepsilon_s^2 + \varepsilon_\eta^2) + \frac{1}{\Delta t^2} \right)$$

$$\times \sqrt{\frac{(\varepsilon_s^2 + \varepsilon_\eta^2)}{2\pi}} \frac{r_0\Delta t}{1 - r_0\Delta t} e^{-(1 - r_0\Delta t)^2/(2r_0^2\Delta t^2(\varepsilon_s^2 + \varepsilon_\eta^2))}. \tag{56}$$

Inserting Eqs. (50), (53), and (56) into Eq. (49), we obtain the mean firing rate for the AD model

$$\langle x_\mu(t; \Delta t)\rangle = r_0 + \mathcal{O}\left((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\right)$$

$$+ \mathcal{O}\left(\frac{(\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4}}{\sqrt{\Delta t}} e^{-(1-r_0\Delta t)^2/(4r_0^2\Delta t^2(\varepsilon_s^2 + \varepsilon_\eta^2))}\right).$$

In the limit of $\Delta t \to 0$, the last term in the above equation can be dropped and we obtain

$$\langle x_\mu(t)\rangle = r_0 + \mathcal{O}\left((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\right). \tag{57}$$

A similar estimation leads to the same formula for the STS model.

Appendix B: Probability for Synchronous Spikes in the AD Model

For the AD model, according to Eq. (23), the probability to observe two spikes in a time window Δt in two spike trains x_μ and x_v is given by

$$P_{\text{syn}} = \left\langle \Theta(\Delta t \cdot r_\mu - \xi)\Theta(\Delta t \cdot r_v - \xi)\right\rangle_{\xi, s, \eta_\mu, \eta_v}$$

$$= \int_0^1 d\xi \int_{-\infty}^{\infty} ds\, p(s) \int_{-\infty}^{\infty} d\eta_\mu\, p(\eta_\mu)$$

$$\times \int_{-\infty}^{\infty} d\eta_v\, p(\eta_v)\Theta(\Delta t \cdot r_\mu - \xi)\Theta(\Delta t \cdot r_v - \xi),$$

where $p(x)$ is a Gaussian distribution with unit variance and zero mean. Splitting the integration interval of the last integral in the above equation into two parts, such that for one interval $\eta_\mu \geq \eta_v$ and for the other $\eta_\mu < \eta_v$, we obtain

$$P_{\text{syn}} = \int_0^1 d\xi \int_{-\infty}^{\infty} ds\, p(s) \int_{-\infty}^{\infty} d\eta_\mu\, p(\eta_\mu)$$

$$\times \left(\int_{\eta_\mu}^{\infty} d\eta_v\, p(\eta_v)\Theta(\Delta t \cdot r_\mu - \xi) + \int_{-\infty}^{\eta_\mu} d\eta_v\, p(\eta_v)\Theta(\Delta t \cdot r_v - \xi)\right),$$

which after a change of the order of integration can be transformed into

$$P_{\text{syn}} = 2\int_0^1 d\xi \int_{-\infty}^{\infty} ds\, p(s) \int_{-\infty}^{\infty} d\eta_\mu\, p(\eta_\mu)\Theta(\Delta t \cdot r_\mu - \xi) \int_{\eta_\mu}^{\infty} d\eta_v\, p(\eta_v). \tag{58}$$

The average of the theta function over ξ reads

$$\int_0^1 d\xi\, \Theta\left(\Delta t \cdot r_\mu(t) - \xi\right) = \begin{cases} 0 & \text{if } r_\mu(t) < 0, \\ \Delta t \cdot r_\mu(t) & \text{if } 0 \leq r_\mu(t) \leq \frac{1}{\Delta t}, \\ 1 & \text{if } r_\mu(t) > \frac{1}{\Delta t}. \end{cases}$$

Using the above equation, Eq. (58) can now be written as

$$
P_{syn} = 2\Delta t \left\langle r_\mu \int_{\eta_\mu}^{\infty} d\eta_\nu \, p(\eta_\nu) \right\rangle - 2\Delta t \left\langle r_\mu \int_{\eta_\mu}^{\infty} d\eta_\nu \, p(\eta_\nu) \Theta(-r_\mu) \right\rangle
$$

$$
- 2\Delta t \left\langle \left(r_\mu - \frac{1}{\Delta t} \right) \int_{\eta_\mu}^{\infty} d\eta_\nu \, p(\eta_\nu) \Theta \left(r_\mu - \frac{1}{\Delta t} \right) \right\rangle.
$$

The order of the second and the third term in the above equation can be calculated analogously to the calculation in Appendix A which leads to

$$
P_{syn} = \Delta t \cdot r_0 \left(1 - \frac{\varepsilon_\eta}{\sqrt{\pi}} \right) + \mathcal{O}\left(\Delta t \cdot \left(\varepsilon_s^2 + \varepsilon_\eta^2 \right)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))} \right)
$$

$$
+ \mathcal{O}\left(\sqrt{\Delta t} \left(\varepsilon_s^2 + \varepsilon_\eta^2 \right)^{1/4} e^{-1/(4 r_0^2 \Delta t^2 (\varepsilon_s^2 + \varepsilon_\eta^2))} \right). \tag{59}
$$

Appendix C: Simplification of the Cross-correlation Function for the STS Model

In this section, we simplify Eq. (33) in Sect. 4.3. Therefore, we first consider

$$
\Gamma_1 = \lim_{t'' \to \infty} \left\langle \delta \left(\gamma_\mu(t'') - \gamma_\nu(t') \right) \left(1 + \varepsilon_s s(t'') + \varepsilon_\eta \eta_\mu(t'') \right) \right\rangle_{\eta_\mu}, \tag{60}
$$

where the average over η_μ is conditioned on fixed realizations of the signal $s(t)$ and the noise process $\eta_\mu(t)$. The variable $\gamma_\mu(t)$ is defined in Eq. (32). We first express the delta function in the above expression as a derivative of a Heaviside function, which leads to

$$
\Gamma_1 = \lim_{t'' \to \infty} \frac{d}{dt''} \left\langle \Theta \left(\gamma_\mu(t'') - \gamma_\nu(t') \right) \right\rangle_{\eta_\mu}
$$

$$
= \lim_{t'' \to \infty} \frac{d}{dt''} \frac{1}{\sqrt{2\pi \sigma(t'')^2}} \int_{-\infty}^{\alpha(t'',t')} dx \exp\left[-\frac{x^2}{2\sigma(t'')^2} \right] \quad \text{with} \tag{61}
$$

$$
\alpha(t'',t') = t'' - t' + \int_{t'}^{t''} dt_1 \varepsilon_s s(t_1) - \int_0^{t'} dt_2 \varepsilon_\eta \eta_\nu(t_2) \quad \text{and}
$$

$$
\sigma(t'')^2 = \varepsilon_\eta^2 \left\langle \int_0^{t''} dt_3 \int_0^{t''} dt_4 \eta_\mu(t_3) \eta_\mu(t_4) \right\rangle_{\eta_\mu}.
$$

As we will show in Eq. (63), the variance $\sigma(t'')^2$ becomes constant in the limit of large t''. Performing the derivative with respect to t'' in Eq. (61), we find

$$
\Gamma_1 = \lim_{t'' \to \infty} \frac{1}{\sqrt{2\pi \sigma^2}} \exp\left[-\frac{\alpha(t'',t')^2}{2\sigma^2} \right] \frac{d\alpha(t'',t')}{dt''}
$$

$$
= \lim_{t'' \to \infty} \left\langle \delta \left(\gamma_\mu(t'') - \gamma_\nu(t') \right) \right\rangle_{\eta_\mu} \left(1 + \varepsilon_s s(t'') \right).
$$

Combining the above equation with Eq. (60), we finally obtain the relation

$$\lim_{t''\to\infty}\langle\delta(\gamma_\mu(t'')-\gamma_\nu(t'))(1+\varepsilon_s s(t'')+\varepsilon_\eta\eta_\mu(t''))\rangle_{\eta_\mu}$$
$$=\lim_{t''\to\infty}\langle\delta(\gamma_\mu(t'')-\gamma_\nu(t'))\rangle_{\eta_\mu}(1+\varepsilon_s s(t'')).$$

Using the above result, we rewrite the average over the delta function in Sect. 4.3 in Eq. (33) as

$$\Gamma_2=\lim_{t'',t'\to\infty}\langle\delta(\gamma_\mu(t'')-\gamma_\nu(t'))$$
$$\times(1+\varepsilon_s s(t'')+\varepsilon_\eta\eta_\mu(t''))(1+\varepsilon_s s(t')+\varepsilon_\eta\eta_\nu(t'))\rangle_{\eta_\mu,\eta_\nu}$$
$$=\lim_{t'',t'\to\infty}\langle\langle\delta(\gamma_\mu(t'')-\gamma_\nu(t'))\rangle_{\eta_\mu}$$
$$\times(1+\varepsilon_s s(t''))(1+\varepsilon_s s(t')+\varepsilon_\eta\eta_\nu(t'))\rangle_{\eta_\nu}.$$

Expressing the delta function in the above equation as the derivative of a Heaviside function with respect to t' as in Eq. (61) and following the steps from the calculation of Γ_1, we find

$$\Gamma_2=\lim_{t'',t'\to\infty}\langle\delta(\gamma_\mu(t'')-\gamma_\nu(t'))\rangle_{\eta_\mu,\eta_\nu}(1+\varepsilon_s s(t''))(1+\varepsilon_s s(t')),$$

which leads to the result

$$\lim_{t'',t'\to\infty}\langle\delta(\gamma_\mu(t'')-\gamma_\nu(t'))$$
$$\times(1+\varepsilon_s s(t'')+\varepsilon_\eta\eta_\mu(t''))(1+\varepsilon_s s(t')+\varepsilon_\eta\eta_\nu(t'))\rangle_{\eta_\mu,\eta_\nu}$$
$$=\lim_{t'',t'\to\infty}\langle\delta(\gamma_\mu(t'')-\gamma_\nu(t'))\rangle_{\eta_\mu,\eta_\nu}(1+\varepsilon_s s(t''))(1+\varepsilon_s s(t')),\qquad(62)$$

used in the calculation of the cross-correlation function in Eq. (34).

Appendix D: Variance of the Integrated Independent Noise

To calculate the variance from Eq. (35), we first consider

$$\sigma^2(t)=\varepsilon_\eta^2\left\langle\int_0^t\int_0^t dt_1\,dt_2\eta(t_1)\eta(t_2)\right\rangle$$

for a Gaussian signal $\eta(t)$ with zero mean and unit variance. Performing the average in the above equation and performing a change of integration variables, we can write

$$\sigma^2(t)=2\varepsilon_\eta^2\int_0^t d\tau\,K_{\eta\eta}(\tau)(t-\tau).$$

Next, we express the autocorrelation function in the above equation by its Fourier transform according to Eq. (6) and find

$$\sigma^2(t) = 2\varepsilon_\eta^2 \int_0^t d\tau \int_{-\infty}^\infty df\, e^{-i2\pi f\tau} S_{\eta\eta}(f)(t-\tau)$$

$$= 2\varepsilon_\eta^2 \int_{-\infty}^\infty df\, S_{\eta\eta}(f)\left(\frac{t}{i2\pi} - \frac{e^{-i2\pi ft}}{4\pi^2 f^2} + \frac{1}{4\pi^2 f^2}\right),$$

where in the last line of the above equation we integrated over τ. Since the power spectrum $S_{\eta\eta}(f)$ is the Fourier transform of a real function, it is symmetric with respect to f, which leads to

$$\sigma^2(t) = \varepsilon_\eta^2 \int_0^\infty df\, S_{\eta\eta}(f) \frac{1-\cos(2\pi ft)}{2\pi^2 f^2}.$$

For a band-pass limited white noise with the power spectrum

$$S_{\eta\eta}(f) = \begin{cases} \frac{1}{2(f_u - f_\ell)} & \text{for } |f| \in [f_\ell, f_u], \\ 0 & \text{else,} \end{cases}$$

we obtain in the limit of large times

$$\lim_{t\to\infty} \sigma^2(t) = \frac{\varepsilon_\eta^2}{2\pi^2 f_u f_\ell}, \tag{63}$$

and for the variance in Eq. (35) for large t we find

$$\hat{\sigma}_g^2 = \lim_{t\to\infty}\left(\sigma^2(t+\tau) + \sigma^2(t)\right) = \frac{\varepsilon_\eta^2}{\pi^2 f_u f_\ell}. \tag{64}$$

Appendix E: Single Spike Train Power Spectrum

In this section, we calculate the single spike train power spectrum for the AD model. Therefore, we consider the autocorrelation function of the spike train

$$K_{xx}(\tau) = \langle x(t+\tau)x(t)\rangle - \langle x(t+\tau)\rangle\langle x(t)\rangle, \tag{65}$$

where we assume that the spike trains are stationary. Since a single spike train is considered and the average is taken over one independent noise process η, we will drop the subscript employed previously.

We first consider the AD model. As for the time-discrete cross-correlation function between two spike trains in Sect. 4.2, we can express Eq. (65) in terms of probability densities

$$K_{xx}(\tau) = \frac{1}{(\Delta t)^2}\Theta\left(\Delta t/2 - |\tau|\right)\cdot P_{\text{spike}}$$

$$+ \frac{1}{(\Delta t)^2}\Theta\left(|\tau| - \Delta t/2\right)\cdot P_{\text{two spikes}}(\tau) - r_0^2, \tag{66}$$

where P_{spike} is the probability to find a spike in a given time bin of width Δt and $P_{two\ spikes}(\tau)$ is the probability to find two spikes separated by the time interval τ with $|\tau| > \Delta t/2$. Analogously to Eq. (23) we find

$$P_{spike} = \langle \Theta(\Delta t \cdot r - \xi) \rangle$$
$$= \Delta t \cdot r_0 + \mathcal{O}\big(\Delta t \cdot (\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big)$$
$$+ \mathcal{O}\big(\sqrt{\Delta t}(\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4r_0^2 \Delta t^2 (\varepsilon_s^2 + \varepsilon_\eta^2))}\big),$$

where in the first line of the above equation we dropped the subscript and the time-argument for the parameter r, which is defined in Eq. (11). For the probability of asynchronous spikes with $|\tau| > \Delta t/2$ we find

$$\frac{1}{(\Delta t)^2} P_{two\ spikes}(\tau) = \langle x(t + \tau) x(t) \rangle$$
$$= \langle \langle x(t + \tau) \rangle_\xi \langle x(t) \rangle_\xi \rangle_{s,\eta}$$
$$= r_0^2 \big(1 + \varepsilon_s^2 K_{ss}(\tau) + \varepsilon_\eta^2 K_{\eta\eta}(\tau)\big)$$
$$+ \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big),$$

which in the limit of $\Delta t \to 0$ leads for the autocorrelation function Eq. (66) to

$$K_{xx}(\tau) = \delta(\tau) r_0 + r_0^2 \big(\varepsilon_s^2 K_{ss}(\tau) + \varepsilon_\eta^2 K_{\eta\eta}(\tau)\big)$$
$$+ \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big).$$

Taking the limit $(\varepsilon_s^2 + \varepsilon_\eta^2) \to 0$ in the above equation (keeping terms up to second-order in ε_s and ε_η), we find after a Fourier transformation the single spike train power spectrum for the AD model

$$S_{xx}(f) \approx r_0 + r_0^2 \big(\varepsilon_s^2 S_{ss}(f) + \varepsilon_\eta^2 S_{\eta\eta}(f)\big). \tag{67}$$

Next, we calculate the single spike train power spectrum for the STS model. Employing Eq. (17), the last term in Eq. (65) can be written as

$$\langle x(t + \tau) \rangle \langle x(t) \rangle = r_0^2 + \mathcal{O}\big((\varepsilon_s^2 + \varepsilon_\eta^2)^{1/4} e^{-1/(4(\varepsilon_s^2 + \varepsilon_\eta^2))}\big). \tag{68}$$

In analogy to Eq. (33), we can transform the first term of the autocorrelation function Eq. (65) into

$$\langle x(t + \tau) x(t) \rangle = r_0 \big\langle \big(\delta(\gamma(t + \tau) - \gamma(t)) + r_0^2\big)$$
$$\times \big(1 + \varepsilon_s s(t + \tau) + \varepsilon_\eta \eta(t + \tau)\big)\big(1 + \varepsilon_s s(t) + \varepsilon_\eta \eta(t)\big)\big\rangle_{\eta,s},$$

with the difference that we now consider a single spike train, and therefore only averages over one independent noise process η. The above equation can be rewritten

as

$$\langle x(t+\tau)x(t)\rangle = r_0\langle \delta\big(\gamma(t+\tau)-\gamma(t)\big)$$
$$\times \big(1+\varepsilon_s s(t+\tau)+\varepsilon_\eta \eta(t+\tau)\big)\big(1+\varepsilon_s s(t)+\varepsilon_\eta \eta(t)\big)\rangle_{\eta,s}$$
$$+ r_0^2 + r_0^2\varepsilon_s^2 K_{ss}(\tau) + r_0^2\varepsilon_\eta^2 K_{\eta\eta}(\tau). \tag{69}$$

After a Fourier transformation of the above equation, the autocorrelation functions of signal and noise are transformed into their respective power-spectra. Inserting the definition of $\gamma(t)$ Eq. (32), the first term of Eq. (69), which we denote by W, is Fourier transformed into

$$W = r_0\Bigg\langle \int_{-\infty}^{\infty} d\tau e^{i2\pi f\tau}\delta\big(G(\tau)\big)$$
$$\times \big(1+\varepsilon_s s(t+\tau)+\varepsilon_\eta \eta(t+\tau)\big)\big(1+\varepsilon_s s(t)+\varepsilon_\eta \eta(t)\big)\Bigg\rangle_{\eta,s} \quad \text{with} \quad (70)$$

$$G(\tau) = \int_t^{t+\tau} dt'\big(1+\varepsilon_s s(t')+\varepsilon_\eta \eta(t')\big).$$

Changing the integration variable in Eq. (70), we find

$$W = r_0\Bigg\langle \int_{-\infty}^{\infty} dG e^{i2\pi f\tau(G)}\delta(G)\big(1+\varepsilon_s s(t)+\varepsilon_\eta \eta(t)\big)\Bigg\rangle_{\eta,s}$$
$$= r_0\langle e^{i2\pi f\tau(0)}\big(1+\varepsilon_s s(t)+\varepsilon_\eta \eta(t)\big)\rangle_{\eta,s}.$$

For a strictly positive process $r(t)$ (see Eq. (11)), the only zero crossing of the integral $G(\tau)$ is at $\tau=0$. We can invert this relation to find $\tau(0)=0$, which leads to

$$W = r_0\langle\big(1+\varepsilon_s s(t)+\varepsilon_\eta \eta(t)\big)\rangle_{\eta,s} + \mathcal{O}\big((\varepsilon_s^2+\varepsilon_\eta^2)^{1/4}e^{-1/(4(\varepsilon_s^2+\varepsilon_\eta^2))}\big)$$
$$= r_0 + \mathcal{O}\big((\varepsilon_s^2+\varepsilon_\eta^2)^{1/4}e^{-1/(4(\varepsilon_s^2+\varepsilon_\eta^2))}\big). \tag{71}$$

The order of the correction term in the above equation is proportional to the square root of the probability that $r_\mu(t)<0$, which has been calculated in Appendix A Eq. (52).

Employing Eqs. (65), (68), (69), and (71) we find in the limit $(\varepsilon_s^2+\varepsilon_\eta^2)\to 0$ (keeping terms up to second-order in ε_s and ε_η) the single spike train power spectrum for the STS model

$$S_{xx}(f) \approx r_0 + r_0^2\varepsilon_s^2 S_{ss}(f) + r_0^2\varepsilon_\eta^2 S_{\eta\eta}(f). \tag{72}$$

References

1. Tuckwell HC: *Stochastic Processes in the Neuroscience*; 1989.
2. Faisal AA, Selen LPJ, Wolpert DM: **Noise in the nervous system**. *Nat Rev Neurosci* 2008, **9**:292.

3. Shannon R: **The mathematical theory of communication**. *Bell Syst Tech J* 1948, **27**:379.

4. Rieke F, Warland D, de Ruyter van Steveninck R, Bialek W: *Spikes: Exploring the Neural Code*. Cambridge: MIT Press; 1999.

5. Wiesenfeld K, Moss F: **Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs**. *Nature* 1995, **373**:33.

6. Hänggi P: **Stochastic resonance in biology**. *ChemPhysChem* 2002, **21**:285-290.

7. McDonnell MD, Ward LM: **The benefits of noise in neural systems: bridging theory and experiment**. *Nat Rev Neurosci* 2011, **12**:415.

8. Gammaitoni L, Hänggi P, Jung P, Marchesoni F: **Stochastic resonance**. *Rev Mod Phys* 1998, **70**:223.

9. Lindner B, Garcia-Ojalvo J, Neiman A, Schimansky-Geier L: **Effects of noise in excitable systems**. *Phys Rep* 2004, **392**:321.

10. Stocks NG: **Suprathreshold stochastic resonance in multilevel threshold systems**. *Phys Rev Lett* 2000, **84**:2310.

11. Stocks NG, Mannella R: **Generic noise-enhanced coding in neuronal arrays**. *Phys Rev E* 2001, **64**:030902.

12. Chacron MJ, Doiron B, Maler L, Longtin A, Bastian J: **Non-classical receptive field mediates switch in a sensory neuron's frequency tuning**. *Nature* 2003, **423**:77.

13. Middleton JW, Longtin A, Benda J, Maler L: **Postsynaptic receptive field size and spike threshold determine encoding of high-frequency information via sensitivity to synchronous presynaptic activity**. *J Neurophysiol* 2009, **101**:1160.

14. Lindner B, Gangloff D, Longtin A, Lewis JE: **Broadband coding with dynamic synapses**. *J Neurosci* 2009, **29**:2076.

15. Sharafi N, Benda J, Lindner B: **Information filtering by synchronous spikes in a neural population**. *J Comput Neurosci* 2013, **34**:285.

16. Droste F, Schwalger T, Lindner B: **Interplay of two signals in a neuron with heterogeneous synaptic short-term plasticity**. *Front Comput Neurosci* 2013, **7**:86

17. Knight BW: **Dynamics of encoding in a population of neurons**. *J Gen Physiol* 1972, **59**:734.

18. Gerstner W, Kistler WM: *Spiking Neuron Models*. Cambridge: Cambridge University Press; 2002.

19. Alonso JM, Usrey WM, Reid RC: **Precisely correlated firing in cells of the lateral geniculate nucleus**. *Nature* 1996, **383**:815.

20. Warzecha AK, Kretzberg J, Egelhaaf M: **Temporal precision of the encoding of motion information by visual interneurons**. *Curr Biol* 1998, **8**:359-368.

21. Trong PK, Rieke F: **Origin of correlated activity between parasol retinal ganglion cells**. *Nat Neurosci* 2008, **11**(11):1343-1351.

22. Churchland M, Byron M, Cunningham J, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM, Moore T, Chang SW, Snyder LH, Lisberger SG, Priebe NJ, Finn IM, Ferster D, Ryu SI, Santhanam G, Sahani M, Shenoy KV: **Stimulus onset quenches neural variability: a widespread cortical phenomenon**. *Nat Neurosci* 2010, **13**:369-378.

23. Steinmetz PN, Roy A, Fitzgerald PJ, Hsiao SS, Johnson KO, Niebur E: **Attention modulates synchronized neuronal firing in primate somatosensory cortex**. *Nature* 2000, **404**:187-190.

24. Stopfer M, Bhagavan S, Smith BH, Laurent G: **Impaired odour discrimination on desynchronization of odour-encoding neural assemblies**. *Nature* 1997, **390**:70-74.

25. Kazama H, Wilson RI: **Origins of correlated activity in an olfactory circuit**. *Nat Neurosci* 2009, **12**(9):1136-1144.

26. Poulet JFA, Petersen CCH: **Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice**. *Nature* 2008, **454**:881.

27. Gentet LJ, Avermann M, Matyas F, Staiger JF, Petersen CCH: **Membrane potential dynamics of gabaergic neurons in the barrel cortex of behaving mice**. *Neuron* 2010, **65**(3):422-435.

28. Binder MD, Powers RK: **Relationship between simulated common synaptic input and discharge synchrony in cat spinal motoneurons**. *J Neurophysiol* 2001, **86**:2266.

29. Neiman AB, Russell DF: **Two distinct types of noisy oscillators in electroreceptors of paddlefish**. *J Neurophysiol* 2004, **92**:492.

30. Warzecha AK, Rosner R, Grewe J: **Impact and sources of neuronal variability in the fly's motion vision pathway**. *J Physiol* 2013, **107**(1):26-40.

31. Bialek W, Deweese M, Rieke F, Warland D: **Bits and brains—information-flow in the nervous-system**. *Physica A* 1993, **200**:581.

32. Gabbiani F: **Coding of time-varying signals in spike trains of linear and half-wave rectifying neurons**. *Netw Comput Neural Syst* 1996, **7**:61.

33. Press WH, Teukolsky SA, Vetterling WT, Flannery BP: *Numerical Recipes: The Art of Scientific Computing*. 3rd edition. Cambridge: Cambridge University Press; 2007.
34. Gardiner CW: *Handbook of Stochastic Methods*; 1985.
35. Bezrukov SM, Vodyanoy I: **Stochastic resonance in non-dynamical systems without response thresholds**. *Nature* 1997, **385**:319-321.
36. Gabbiani F, Cox SJ: *Mathematics for Neuroscientists*. San Diego: Academic Press; 2010.
37. Cox DR, Isham V: *Point Processes*. London: Chapman & Hall; 1980.
38. Feller W: *An Introduction to Probability Theory and Its Applications*. *Volume 1*. 3rd edition. New York: Wiley; 1968.
39. Gestri G, Mastebroek HAK, Zaagman WH: **Stochastic constancy, variability and adaptation of spike generation—performance of a giant-neuron in the visual-system of the fly**. *Biol Cybern* 1980, **38**:31.
40. de Ruyter van Steveninck RR, Lewen GD, Strong SP, Koberle R, Bialek W: **Reproducibility and variability in neural spike trains**. *Science* 1997, **275**:1805.
41. Feller W: *An Introduction to Probability Theory and Its Applications*. *Volume 2*. 2nd edition. New York: Wiley; 1971.
42. Bauerle N, Grubel R: **Multivariate counting processes: copulas and beyond**. *ASTIN Bull* 2005, **35**(2):379.
43. Neiman AB, Russell DF, Rowe MH: **Identifying temporal codes in spontaneously active sensory neurons**. *PLoS ONE* 2011, **6**:e27380.
44. Trousdale J, Hu Y, Shea-Brown E, Josic K: **A generative spike train model with time-structured higher order correlations**. *Front Comput Neurosci* 2013, **7**:84.
45. Kuhn A, Aertsen A, Rotter S: **Higher-order statistics of input ensembles and the response of simple model neurons**. *Neural Comput* 2003, **15**(1):67-101.
46. Moreno-Bote R, Parga N: **Auto- and crosscorrelograms for the spike response of leaky integrate-and-fire neurons with slow synapses**. *Phys Rev Lett* 2006, **96**:028101.
47. de la Rocha J, Doiron B, Shea-Brown E, Josic K, Reyes A: **Correlation between neural spike trains increases with firing rate**. *Nature* 2007, **448**:802.
48. Shea-Brown E, Josić K, de la Rocha J, Doiron B: **Correlation and synchrony transfer in integrate-and-fire neurons: basic properties and consequences for coding**. *Phys Rev Lett* 2008, **100**:108102.
49. Ostojic S, Brunel N, Hakim V: **How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains**. *J Neurosci* 2009, **29**:10234.
50. Vilela RD, Lindner B: **A comparative study of three different integrate-and-fire neurons: spontaneous activity, dynamical response, and stimulus-induced correlation**. *Phys Rev E* 2009, **80**:031909.
51. Abouzeid A, Ermentrout B: **Correlation transfer in stochastically driven neural oscillators over long and short time scales**. *Phys Rev E* 2011, **84**(6):061914.
52. Schultze-Kraft M, Diesmann M, Grun S, Helias M: **Noise suppression and surplus synchrony by coincidence detection**. *PLoS Comput Biol* 2013, **9**(4):e1002904.
53. Bialek W, Zee A: **Coding and computation with neural spike trains**. *J Stat Phys* 1990, **59**:103.
54. Ermentrout GB, Terman DH: *Mathematical Foundations of Neuroscience*. New York: Springer; 2010.
55. Malyshev A, Tchumatchenko T, Volgushev S, Volgushev M: **Energy-efficient encoding by shifting spikes in neocortical neurons**. *Eur J Neurosci* 2013, **38**(8):3181-3188.