

Synaptic filtering of rate-coded information

Matthias Merkel and Benjamin Lindner

Max-Planck-Institut für Physik komplexer Systeme, Nöthnitzer Str. 38, 01187 Dresden, Germany

(Received 3 December 2009; revised manuscript received 26 February 2010; published 29 April 2010)

In this paper, we analytically examine the influence of synaptic short-term plasticity (STP) on the transfer of rate-coded information through synapses. STP endows each presynaptic input spike with an amplitude that depends on previous input spikes. We develop a method to calculate the spectral statistics of this amplitude modulated spike train (postsynaptic input) for the case of an inhomogeneous Poisson process. We derive in particular analytical approximations for cross-spectra, power spectra, and for the coherence function between the postsynaptic input and the time-dependent rate modulation for a specific model. We give simple expressions for the coherence in the limiting cases of pure facilitation and pure depression. Using our analytical results and extensive numerical simulations, we study the spectral coherence function for postsynaptic input resulting from a single synapse or from a group of synapses. For a single synapse, we find that the synaptic coherence function is largely independent of frequency indicating broadband information transmission. This effect is even more pronounced for a large number of synapses. However, additional noise gives rise to frequency-dependent information filtering.

DOI: [10.1103/PhysRevE.81.041921](https://doi.org/10.1103/PhysRevE.81.041921)

PACS number(s): 87.19.fo, 87.19.lc, 87.18.Sn

I. INTRODUCTION

Synapses constitute the connections between nerve cells (neurons) in the brain. However, they are not just bare conveyers of information from one neuron to another, but play a far more active role [1]. First of all, slow changes in their efficacy (i.e., in their postsynaptic amplitude), called long-term plasticity, are believed to be the physiological basis for learning and memory in the brain. Furthermore, synapses also display changes in their efficacy on shorter time scales (~ 100 ms to 1 s). The latter effect is referred to as short-term plasticity (STP) and is the subject of our work. Depending on the most recent spiking history, synapses modify the amplitudes of the action potentials they transmit. Repeated presynaptic spiking may increase the postsynaptic amplitude (facilitation) or decrease it (depression). Both kinds of STP can coexist in a single synapse because they rely on a variety of different physiological mechanisms [2,3].

A number of phenomenological models describing STP have been established in the past [4–10]. In this paper, we focus on facilitation-depression (FD) models as proposed and studied in [8,11–13].

The fact that synapses are active elements in the neural dynamics suggests, that they may also play an important role in information processing in the brain. STP may be important for synaptic gain control, for the detection of transients and of bursts, and may be linked to working memory [6,9,14–16]. In addition, synaptic filtering properties have been examined in terms of the average postsynaptic amplitude depending on the presynaptic firing rate [6,8,14,17]. The amount information about the presynaptic spiking history carried by the synaptic strength has been studied in [18].

Recently, it was shown that information transmission about a rate-coded time-dependent stimulus across an ensemble of dynamic synapses is broadband, i.e., independent of frequency [13]. More precisely, it was shown by numerical simulations that neither facilitation nor depression introduces a frequency-dependent filtering of the spectral input-

output coherence. In this work, we study this problem analytically. We consider an ensemble of independent dynamic synapses driven by rate-modulated Poisson processes and derive expressions for the power spectra, cross-spectra, and coherence functions. We investigate these statistics in detail for a single FD synapse and discuss their basic features for the physiologically accessible part of the parameter space. For an ensemble of many synapses, we furthermore inspect the effect of an additional synaptic noise on information transfer. Our results contribute to a deeper understanding of the conditions for broadband coding by dynamic synapses.

Our study can also be regarded as a contribution to the theory of point processes. The main problem in this paper is the calculation of the spectral statistics of a spike train with time-dependent amplitudes which are functionals of the underlying point process itself. We would like to point out that our results may also be of interest for other areas where pulse trains with variable amplitudes are encountered, as for instance, in the time series of earthquakes or in drop-outs of lasers.

II. MODEL AND MEASURES

We consider a population of facilitating and depressing synapses, by which a neuron (the “target neuron”) receives independent Poissonian spike train input from other neurons (the “input neurons”); see Fig. 1. For a part of this input ensemble (the “signal neurons”) a time-dependent stimulus is encoded in the instantaneous firing rate of the input spike trains. The transfer of this stimulus is affected by the nature (facilitating or depressing) of the “signal synapses” connecting the signal neurons to the target neuron. The remaining “noise synapses” receive Poissonian input spike trains with constant rates. They may have a different STP character and they will also contribute to the total postsynaptic input to the target neuron.

We would like to stress in this context, that we use the term “postsynaptic *input*” for the amplitude-modulated input

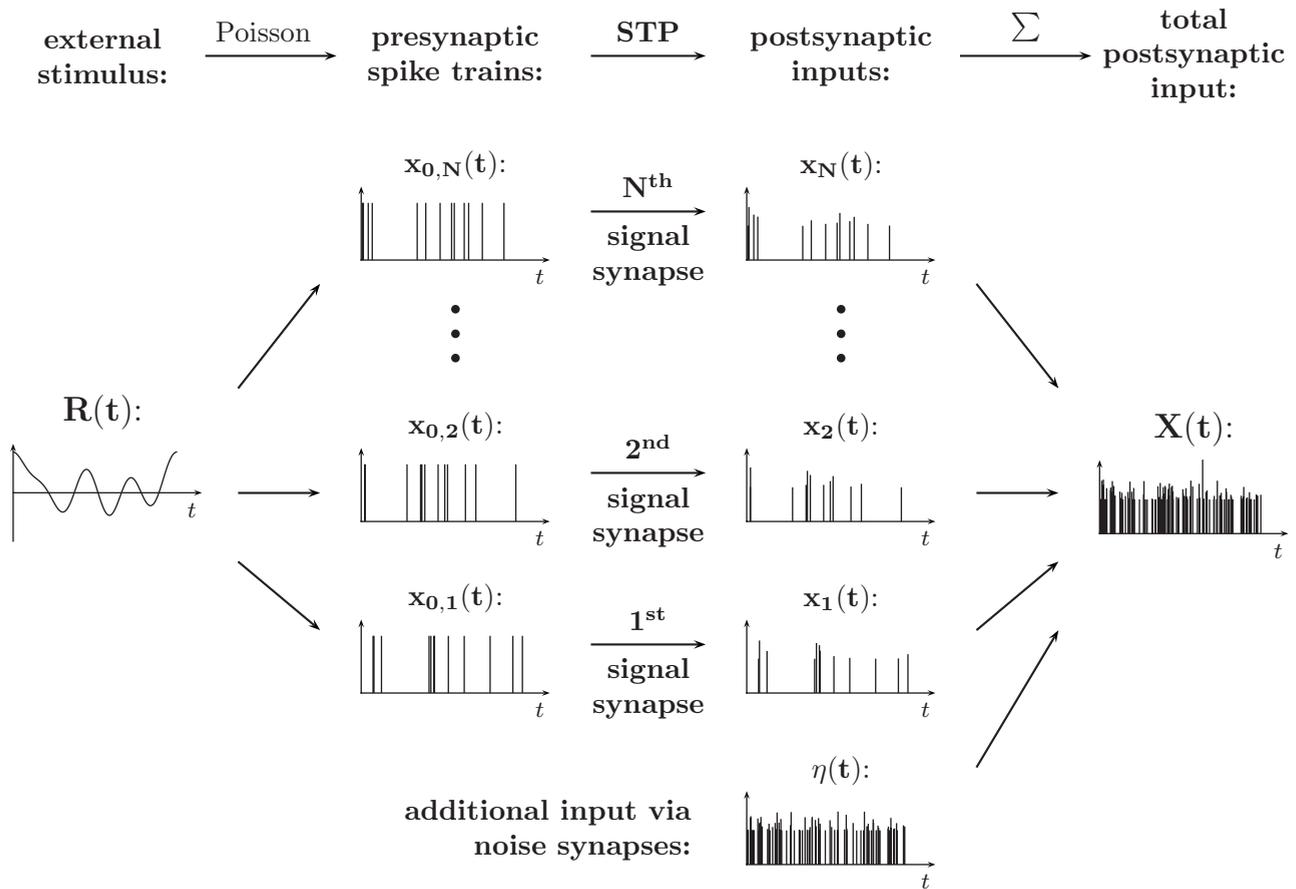


FIG. 1. Scheme illustrating Eqs. (1)–(3) and (9): The external stimulus $R(t)$ modulates the rate of Poissonian spike trains $x_{0,i}(t)$. Each of these spike trains is modulated by a dynamic synapse according to the FD model turning it into $x_i(t)$. Finally, the total postsynaptic input $X(t)$ to the target neuron is given by the sum of these modulated spike trains with an additional noise $\eta(t)$. The additional noise accounts for input from synapses that are not involved in the transmission of the stimulus signal R .

that a synapse provides to the target neuron. Similarly, the “total postsynaptic input” is the sum over all the synaptic inputs that the target neuron receives. In this paper, we do not consider the output spike train of the target neuron, which is often termed “postsynaptic spike train.” However, in Sec. II E, we briefly show how to extend our results about the total postsynaptic input to the membrane conductance and the subthreshold membrane voltage of the target neuron.

Note that in the following, we strictly distinguish between the notions of rate and frequency. *Rate* refers to a neuronal firing rate, whereas *frequency* is solely used for the argument of the Fourier transform.

A. Time-dependent signals as rate modulation

We write the presynaptic spike train associated to the i th signal neuron as

$$x_{0,i}(t) = \sum_j \delta(t - t_{i,j}). \quad (1)$$

We assume that the Poissonian spike trains from different signal neurons are statistically independent from each other except for a common rate modulation—the external stimulus $R(t)$. This is clearly a simplification, since it is known, that

correlations between cortical spike trains are common [19]. The time-dependent firing rate $\nu(t)$ is

$$\nu(t) = \begin{cases} r[1 + \varepsilon R(t)] & \text{for } 1 + \varepsilon R(t) \geq 0 \text{ and} \\ 0 & \text{else.} \end{cases} \quad (2)$$

Here, r is the baseline rate, which is weakly modulated by the external stimulus $R(t)$ with small amplitude ε . We assume that $R(t)$ is a band-limited Gaussian signal with $\langle R \rangle = 0$, $\langle R^2 \rangle = 1$, and a limiting frequency of 50 Hz. Generally, we neglect the possibility of negative rates, i.e., $1 + \varepsilon R(t) < 0$. Note that for the amplitudes used in this paper ($\varepsilon \leq 0.2$), the probability of $1 + \varepsilon R(t) < 0$ at an arbitrary instance t is less than 3×10^{-7} .

B. Short-term plasticity

For each of the considered synapses in the population, we use a deterministic FD model in order to describe short-term plasticity. The corresponding dynamical system is similar to those used in [8,11–13]. Generally, facilitation and depression are modeled via different variables F and D , respectively. The product of these variables is the synaptic amplitude A .

The postsynaptic input from the synapse corresponding to the i th input neuron is given by (see also Fig. 1)

$$x_i(t) = A_i(t^-)x_{0,i}(t). \quad (3)$$

Here and in the following, $f(t^\pm)$ refers to $\lim_{\eta \rightarrow 0} f(t \pm \eta)$.

The synaptic amplitude $A_i(t)$ is the product of a facilitation variable $F_i(t)$ and a depression variable $D_i(t)$,

$$A_i(t) = F_i(t)D_i(t). \quad (4)$$

The dynamics of the facilitation and depression variables, $F_i(t)$ and $D_i(t)$, depend on the respective presynaptic spiking history $x_{0,i}(t)$ of each synapse. For a *static synapse* we would have

$$A_i(t) = A_0 = \text{const}, \quad (5)$$

i.e., the postsynaptic input would, apart from a factor, coincide with the presynaptic spike train.

1. Facilitation

The facilitation dynamics models the increase of the release probability of synaptic vesicles due to an increase of residual calcium in the presynaptic terminal [8,11]. $F_C(t)$ corresponds to the concentration of a calcium bound molecule within the presynaptic terminal. Upon spiking, this concentration is increased by the constant Δ . However, this concentration decays with a rate $1/\tau_F$. These features are expressed by the following differential equation:

$$\frac{dF_C(t)}{dt} = -\frac{F_C(t)}{\tau_F} + \Delta \times x_0(t). \quad (6)$$

The calcium bound molecule may bind to the release site altering the release probability of neurotransmitter into the synaptic cleft. This release probability is our facilitation variable $F(t)$, given by

$$F(t) = F_0 + \frac{1}{(1 - F_0)^{-1} + F_C^{-1}(t)}. \quad (7)$$

The constant F_0 is the baseline release probability, i.e., the release probability in the limit of low presynaptic spiking rates. If $F_C \ll 1 - F_0$, we obtain $F \approx F_0 + F_C$. If, on the other hand, F_C is large, F approaches 1.

The facilitation dynamics is illustrated in Fig. 2. As can be seen in the amplitude dynamics (lower panel), a rapid succession of presynaptic spikes evokes an increase of the synaptic amplitude, whereas during large pauses of the presynaptic spike train, the synaptic amplitude drops again.

2. Depression

As in [11–13], we simplify the approach chosen in [8]. We assume that depression is due to a refractoriness of synaptic release sites. Release sites enter a refractory state upon neurotransmitter release. Our depression variable $D(t)$ is the fraction of sites that are ready for immediate release. Sites recover from the refractory state into the release-ready state with the rate $1/\tau_D$ (in [8], τ_D is defined differently). This is expressed by

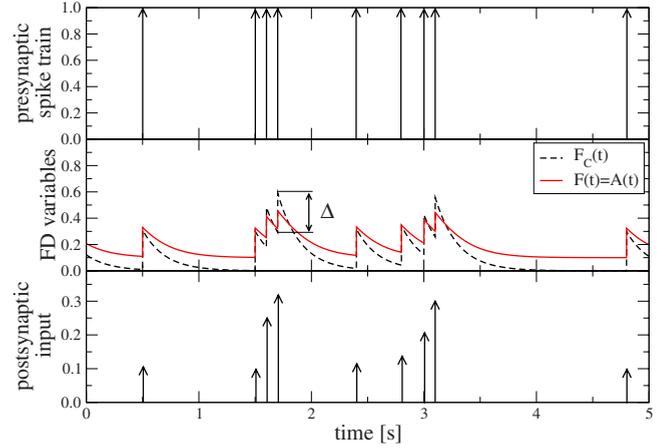


FIG. 2. (Color online) Illustration of facilitation-only dynamics as described by Eq. (3) with $A(t)=F(t)$, and Eqs. (6) and (7). Upper panel: presynaptic spike train $x_0(t)$. Middle panel: dynamics of $F_C(t)$ and facilitation dynamics $F(t)$. Lower panel: postsynaptic input $x(t)$. For illustration purposes, the presynaptic spike times were chosen manually. Parameters: $F_0=0.1$, $\Delta=0.3$, and $\tau_F=200$ ms.

$$\frac{dD(t)}{dt} = \frac{1 - D(t)}{\tau_D} - F(t^-)D(t^-)x_0(t). \quad (8)$$

The initial value for D is chosen such that D varies between 0 and 1. The depression dynamics is illustrated in Fig. 3. In contrast to the facilitation dynamics, a rapid succession of presynaptic spikes leads to a decrease in the synaptic amplitude, whereas for long pauses in the presynaptic input, the synaptic amplitude recovers. The combined effect of both, facilitation and depression, is illustrated in Fig. 4.

C. Total postsynaptic input

The total postsynaptic input $X(t)$ to the target neuron includes the postsynaptic input $x_i(t)$ from the N signal synapses

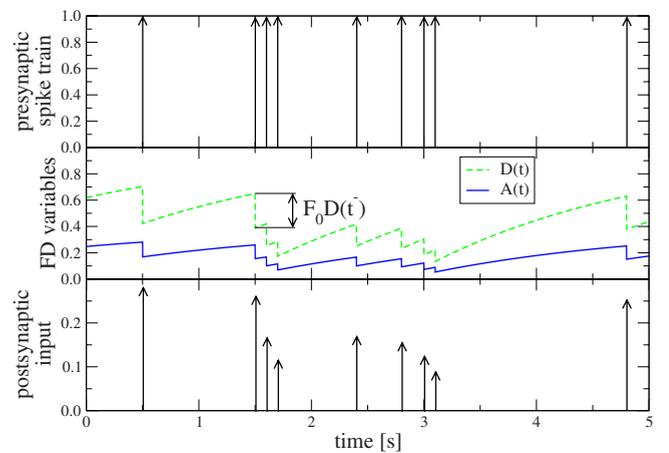


FIG. 3. (Color online) Illustration of depression-only dynamics as described by Eq. (3) with $A(t)=F_0 \cdot D(t)$ and Eq. (8) with $F(t)=F_0$. Upper panel: presynaptic spike train $x_0(t)$. Middle panel: depression dynamics $D(t)$ and synaptic amplitude $A(t)$. Lower panel: postsynaptic input $x(t)$. Parameters: $F_0=0.4$ and $\tau_D=2000$ ms; spike times as in Fig. 2.

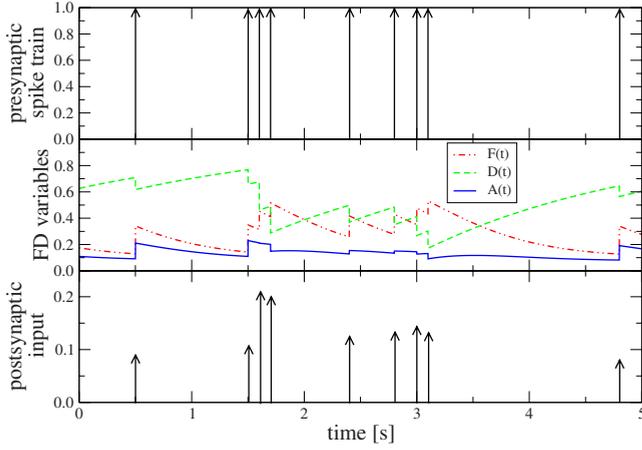


FIG. 4. (Color online) Illustration of the full FD dynamics as described by Eq. (3) with $A(t)=F(t) \cdot D(t)$ and Eqs. (6)–(8). Upper panel: presynaptic spike train $x_0(t)$. Middle panel: facilitation dynamics $F(t)$, depression dynamics $D(t)$, and synaptic amplitude $A(t)$. Lower panel: postsynaptic input $x(t)$. Parameters: $F_0=0.1$, $\Delta=0.3$, $\tau_F=500$ ms, and $\tau_D=2000$ ms; spike times as in Fig. 2.

and the total synaptic input $\eta(t)$ by the noise synapses (see also Fig. 1):

$$X(t) = \eta(t) + \sum_{i=1}^N x_i(t). \quad (9)$$

The noise $\eta(t)$ could likewise include intrinsic fluctuations of the target neuron, for instance, channel noise of the postsynaptic membrane. Below, we will however mainly consider the case, in which the noise statistics is shaped by either facilitation or depression, i.e., the case where the noise arises from presynaptic input.

D. Spectral measures

We define the correlation function $K_{ab}(\tau)$ between two stochastic processes $a(t)$ and $b(t)$ by

$$K_{ab}(\tau) = \langle a(t)b(t+\tau) \rangle - \langle a(t) \rangle \langle b(t) \rangle. \quad (10)$$

K_{ab} does not depend on the absolute time t , because we assume stationarity (i.e., we ignore any kind of transient).

The spectrum $S_{ab}(f)$ is the Fourier transformation of the correlation function $K_{ab}(\tau)$,

$$S_{ab}(f) = \int_{-\infty}^{\infty} e^{2\pi i f \tau} K_{ab}(\tau) d\tau. \quad (11)$$

If $a=b$, then $K_{ab}(\tau)$ is the autocorrelation and $S_{ab}(f)$ is the power spectrum of $a(t)$. If $a \neq b$, then $K_{ab}(\tau)$ is the cross-correlation and $S_{ab}(f)$ is the cross-spectrum of $a(t)$ and $b(t)$.

Another way to calculate the spectra is given by [20]

$$\delta(f-f') S_{ab}(f) = \langle \tilde{a}(f) \tilde{b}^*(f') \rangle, \quad (12)$$

where \tilde{b}^* denote the complex conjugate of \tilde{b} , and \tilde{a} denote the Fourier transform of a ,

$$\tilde{a}(f) = \int_{-\infty}^{\infty} e^{-2\pi i f t} a(t) dt. \quad (13)$$

Correlations between two signals a and b can be quantified by the coherence function given by

$$C_{ab}(f) = \frac{|S_{ab}(f)|^2}{S_{aa}(f)S_{bb}(f)}. \quad (14)$$

This function attains values between zero and one and measures the normalized linear correlation in the frequency domain.

E. Relation to the conductance and membrane potential

In this paper, we are mainly interested in the coherence $C_{RX}(f)$ between rate modulation $R(t)$ and the total postsynaptic input $X(t)$ and in the cross- and power spectra $S_{RX}(f)$ and $S_{XX}(f)$. However, we briefly want to point out relations to statistics which is more easily accessible in experiments than the postsynaptic input X , as for example the total membrane conductance G of the target neuron or its membrane voltage V . In voltage-clamp experiments, $G(t)$ is proportional to the excitatory postsynaptic current (EPSC).

According to a common model for conductance-based synapses [21], the conductance G is the low-pass filtered postsynaptic input X . Hence, the cross- and power spectra $S_{RG}(f)$ and $S_{GG}(f)$ are low-pass filtered versions of $S_{RX}(f)$ and $S_{XX}(f)$. We conclude that $C_{RG}(f)$ equals $C_{RX}(f)$.

One may use a conductance-based leaky integrator model in order to characterize the *subthreshold* membrane voltage V . For a low variance of the conductance G , which is a realistic assumption for many neuronal systems, one can approximate the conductance-based leaky integrator by a current-based neuron model via the effective time constant approximation [22,23] (for critical evaluations of this approximation and for improved approximation schemes, see [24,25]). With these assumptions, the cross- and power spectra $S_{RV}(f)$ and $S_{VV}(f)$ are twice low-pass filtered versions of $S_{RX}(f)$ and $S_{XX}(f)$; and furthermore, the coherence $C_{RV}(f)$ is well approximated by $C_{RX}(f)$. However, as soon as a spiking mechanism for the target neuron is taken into account, non-trivial consequences arise [13].

III. THEORY

The main interest of this paper is the study of effects of STP on the information transmission across the synapse. To this end, we aim at the calculation of the coherence function between the external stimulus R and the total postsynaptic input X to the target neuron.

In Appendix A, we find

$$\begin{aligned} \frac{1}{C_{RX}(f)} &= \frac{1}{N} \times \frac{1}{C_{R_x}(f)} + \frac{N-1}{N} \times \frac{1}{C_{R\langle x \rangle}(f)} \\ &+ \frac{1}{N} \times \frac{1}{C_{R_x}(f)} \times \frac{S_{\eta\eta}(f)}{NS_{xx}(f)}. \end{aligned} \quad (15)$$

This is one of the central formulas used in this work, because it illustrates the influence of the number of synapses N , the single synapse coherence $C_{R_x}(f)$, and noise $S_{\eta\eta}(f)$ on the coherence $C_{RX}(f)$. The coherence $C_{R\langle x \rangle}(f)$ between R and $\langle x \rangle_{x_0|R}$ turns out to be approximately one (see Appendix A 2).

Note that additional noise enters Eq. (15) only via its power spectrum $S_{\eta\eta}$.

In the following sections, we first present our results for a single synapse (see Sec. IV) and then discuss the general case including multiple synapses and additional noise (see Sec. V). In this section, we start with some general considerations about $C_{R_x}(f)$ and $C_{R_X}(f)$.

A. Information rate and coherence function

In this section, we treat the single-synapse situation (i.e., $N=1$) from a more formal point of view. We consider the mean information rate $\mathcal{R}_{\text{info}}$ (i.e., the mutual information that is transmitted per unit of time) over a single synapse in the absence of external noise η . First of all, we ask for $\mathcal{R}_{\text{info}}(R, x_0)$ between $R(t)$ and the presynaptic spike train $x_0(t)$. For a rate-modulated Poisson process, this information rate reads to the lowest significant order in ε [using Eq. (18) from [26]; compare also [27]]

$$\mathcal{R}_{\text{info}}(R, x_0) \approx \frac{1}{2} \int_{-\infty}^{\infty} \log_2[1 + \varepsilon^2 r S_{RR}(f)] df \approx \frac{\varepsilon^2 r}{2 \ln 2}. \quad (16)$$

Note that without STP, the postsynaptic response x equals the presynaptic spike train x_0 . Hence, $\mathcal{R}_{\text{info}}(R, x_0)$ is also the mean information rate between $R(t)$ and the postsynaptic response in a situation *without* STP.

In this paper, we are interested in the effect of STP on the information transmission of a synapse. In a strict sense, STP as modeled in Eqs. (3)–(8) and in the absence of external noise does not change the mean information rate:

$$\mathcal{R}_{\text{info}}(R, x) = \mathcal{R}_{\text{info}}(R, x_0) \quad \text{for } \eta \equiv 0. \quad (17)$$

This can be seen as follows. On the one hand, the information carried by the postsynaptic input $x(t)$ cannot be more than that carried by the presynaptic spike train $x_0(t)$, because $x(t)$ depends exclusively on x_0 [according to Eqs. (3)–(8)]. Put differently, x cannot contain additional information on the rate modulation $R(t)$ which is not yet present in $x_0(t)$. On the other hand, x contains not less information than x_0 , because one can reconstruct x_0 from x by simply extracting only the spike times from x and ignoring the amplitudes (this could be realized, for instance, by a spike detector with arbitrarily small threshold). Therefore, whatever information is carried by x_0 , it must also be carried by x . Hence, in the absence of additional noise and for the deterministic STP model treated here, synaptic plasticity does not change the basic information content of a single spike train, although the amplitudes of the spikes are modulated.

However, the dynamics of the neuron that is driven by the postsynaptic input x may profit or suffer from this amplitude modulation. One might also regard a dynamic synapse as preprocessing instance for the neuron in the sense that it prepares the spike train for neuronal processing. In this work, we want to consider a case, where the neuron's dynamics depend mainly linearly on the synaptic input it receives. Therefore, we are interested in the linear information transmission properties of a dynamic synapse. A quantity that is

well suited to describe this linear information transmission is the *coherence function* $C(f)$ as defined in Eq. (14). It quantifies, how well a transmission can be expressed by a linear filter—and in addition, for Gaussian signals R , it gives a lower bound on the mean information rate $\mathcal{R}_{\text{info}}(R, x)$ [28–31],

$$\mathcal{R}_{\text{info}}(R, x) \geq \mathcal{R}_{\text{info, LB}}(R, x) = - \int_0^{\infty} \log_2[1 - C_{R_x}(f)] df. \quad (18)$$

Note that x does not need to be Gaussian for the lower bound condition in Eq. (18) to hold; according to [30], it suffices that the stimulus R is Gaussian—as we assume in this work.

First, we want to consider the coherence function $C_{R_{x_0}}(f)$ between R and the presynaptic spike train x_0 . Following [30,32], we can write to the lowest significant order in ε :

$$C_{R_{x_0}}(f) \approx \varepsilon^2 r S_{RR}(f). \quad (19)$$

Putting this into Eq. (18) and comparing to Eq. (16), one sees

$$\mathcal{R}_{\text{info, LB}}(R, x_0) \approx \frac{1}{2} \int_{-\infty}^{\infty} \log_2[1 + \varepsilon^2 r S_{RR}(f)] df \approx \mathcal{R}_{\text{info}}(R, x_0). \quad (20)$$

Hence, for a Poissonian spike train with constant amplitude, the lower bound coincides with the information rate itself. Using Eq. (18), we can conclude that—to the lowest significant order in ε —the integral of the single-synapse coherence $C_{R_x}(f)$ cannot become greater than the integral of the coherence of the presynaptic spike train $C_{R_{x_0}}(f)$:

$$\int_0^{\infty} C_{R_x}(f) df \leq \int_0^{\infty} C_{R_{x_0}}(f) df. \quad (21)$$

Indeed, we observe that even $C_{R_x}(f) \leq C_{R_{x_0}}(f)$ for all $f > 0$ (see below).

In this work, we want to examine the coherence function $C_{R_x}(f)$ between $R(t)$ and the postsynaptic input $x(t)$ in dependence on frequency. We use $C_{R_x}(f)$ in order to judge, whether dynamic synapses can be considered as preprocessing units for those neurons which evaluate mainly the *linear* information content of their postsynaptic input. If, for example, at a certain frequency f_1 , $C_{R_x}(f_1)$ is considerably less than $C_{R_{x_0}}(f_1)$ while at another frequency f_2 , $C_{R_x}(f_2)$ is close to $C_{R_{x_0}}(f_2)$, frequency-dependent information filtering would come into play.

We emphasize, that Eq. (17) holds true for a single synapse neglecting stochasticity of transmitter release. For the total synaptic input, the coherence and the mutual information is expected to be reduced, for instance, by an external noise $\eta(t)$. The extent of this reduction is addressed in Sec. V B.

B. Main idea for the calculation of C_{R_x}

We present a way to calculate the coherence function $C_{R_x}(f)$ according to Eq. (14). In order to do this, we show how to derive the power spectrum $S_{xx}(f)$ as well as the cross-spectrum $S_{R_x}(f)$.

Using Eqs. (10) and (11) and stationarity, we can calculate $S_{xx}(f)$ for $f \neq 0$ by the Fourier transformation of

$$\langle x(t)x(t+\tau) \rangle = \langle \langle x(t)x(t+\tau) \rangle_{x_0} \rangle_R \quad (22)$$

and $S_{Rx}(f)$ for $f \neq 0$ by the Fourier transformation of

$$\langle R(t)x(t+\tau) \rangle = \langle R(t) \langle x(t+\tau) \rangle_{x_0} \rangle_R, \quad (23)$$

where $\langle \cdot \rangle_{x_0}$ indicates averaging over the x_0 ensemble for frozen R and $\langle \cdot \rangle_R$ indicates averaging over the R ensemble.

The main difficulty of the derivation is the calculation of the averages

$$\langle x(t)x(t+\tau) \rangle_{x_0} \quad \text{and} \quad \langle x(t+\tau) \rangle_{x_0} \quad (24)$$

to the lowest significant order in ε . Once expressions for these averages have been found, the subsequent averaging over the R ensemble and the Fourier transformation are elementary.

For the calculation of the averages in Eq. (24), we adopt the following strategy: First, we attempt to express x explicitly in terms of the unmodulated spike train x_0 . Second, we calculate these averages, now only involving x_0 . In the following, we describe both steps in more detail.

Concerning the first step, we can write the averages in Eq. (24) in terms of F , D , and x_0 , because, $x(t) = F(t^-)D(t^-)x_0(t)$ [cf. Eqs. (3) and (4)]. What remains is the derivation of explicit formulas expressing F and D in terms of $x_0(t)$. Unfortunately, because of the static nonlinearity in $F(F_C)$ in the facilitation dynamics and because of the term $F \times D$ in the depression dynamics, this appears to be disproportionately complicated. For this reason, we linearize the F and D dynamics as outlined below.

For the facilitation dynamics, we linearize the dependency $F(F_C)$ as discussed in Appendix C. In our linear approximation, F is a constant plus a convolution of x_0 :

$$F(t) \approx F_{\text{lin}}(t) = F_{0,\text{lin}} + \Delta_{\text{lin}} \int_{-\infty}^t e^{-(t-t')/\tau_F} x_0(t') dt', \quad (25)$$

where $F_{0,\text{lin}}$ and Δ_{lin} are constants that depend on the model parameters and r . For linear facilitation models as used for example in [13], $F_{0,\text{lin}}$ and Δ_{lin} correspond to F_0 and Δ , respectively.

For the depression dynamics, we also use a linearized equation in which, however, the input enters in a multiplicative fashion (cf. Appendix D). The solution is for this reason more complicated and reads

$$D(t) \approx D_{\text{lin}}(t) = \int_{-\infty}^t e^{-(t-t')/\tau_D} e^{t'+t} \left(\frac{1}{\tau_D} - x_0(t') \right) \times \left[\langle D_{\text{lin}} \rangle F_{\text{lin}}(t'^-) + \langle F_{\text{lin}} D_{\text{lin}} \rangle - 2 \langle F_{\text{lin}} \rangle \langle D_{\text{lin}} \rangle \right] dt', \quad (26)$$

where we define the term $e(t_a, t_b)$ as follows:

$$e(t_a, t_b) = \exp \left[\ln(1 - F_1) \int_{t_a}^{t_b} x_0(t') dt' \right], \quad t_a \leq t_b. \quad (27)$$

The constants $F_1 = \langle F_{\text{lin}} \rangle$, $\langle D_{\text{lin}} \rangle$, and $\langle F_{\text{lin}} D_{\text{lin}} \rangle$ depend merely on model parameters and r (cf. Appendix D).

Regarding the explicit dependencies of F and D on x_0 , we find, that the averages in Eq. (24) take the form of *sums of multiple convolution integrals* over averages of the kind

$$\langle x_0(t_1)x_0(t_2) \cdots e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \cdots \rangle_{x_0}, \quad (28)$$

where the time instants $t_1, t_2, \dots, t_j, t_{j+1}, \dots$ may in general be completely arbitrary (including, that any two of these time variables may be identical), except for:

$$t_j \leq t_{j+1}, \quad t_{j+2} \leq t_{j+3}, \quad \dots$$

For the sake of illustration of how the averages Eq. (28) appear in the calculation of spectral measures, let us consider the simple case of pure depression ($F(t) \equiv F_0$). For the cross-spectrum $S_{Rx}(f)$, we obtain

$$S_{Rx}(f) = F_0 \int_{-\infty}^{\infty} e^{2\pi i f \tau} \langle R(t) \langle D_{\text{lin}}(t+\tau^-) x_0(t+\tau) \rangle_{x_0} \rangle_R d\tau. \quad (29)$$

Inserting the explicit solution for $D_{\text{lin}}(t)$, we find for the average $\langle D_{\text{lin}}(t^-) x_0(t) \rangle_{x_0}$ (using that for pure depression $F_{\text{lin}} = F_0$)

$$\begin{aligned} \langle D_{\text{lin}}(t^-) x_0(t) \rangle_{x_0} &= \frac{1}{\tau_D} \int_{-\infty}^{t^-} e^{-(t-t')/\tau_D} \langle x_0(t) e(t'^+, t^-) \rangle_{x_0} dt' \\ &\quad - (\langle D_{\text{lin}} \rangle F_0 + \langle F_{\text{lin}} D_{\text{lin}} \rangle - 2 \langle F_{\text{lin}} \rangle \langle D_{\text{lin}} \rangle) \\ &\quad \times \int_{-\infty}^{t^-} e^{-(t-t')/\tau_D} \langle x_0(t') x_0(t) e(t'^+, t^-) \rangle_{x_0} dt' \\ &= \frac{1}{\tau_D} \int_{-\infty}^{t^-} e^{-(t-t')/\tau_D} \langle x_0(t) e(t'^+, t^-) \rangle_{x_0} dt'. \end{aligned} \quad (30)$$

This way, we express the average $\langle D_{\text{lin}}(t^-) x_0(t) \rangle_{x_0}$ in terms of $\langle x_0(t) e(t'^+, t^-) \rangle_{x_0}$, which is of the kind described by Eq. (28).

For the more involved case of full FD dynamics, the derivation of the power spectrum requires the calculation of averages as in Eq. (28) with up to eight x_0 factors.

Given, that we can reduce the spectral measures to averages like in Eq. (28), we now have to clarify how to actually calculate these averages. Our method of solution for this problem (outlined in Appendix B) is to express averages like in Eq. (28) in terms of averages of products of another inhomogeneous Poissonian δ spike train \hat{x}_0

$$\langle \hat{x}_0(t_1) \hat{x}_0(t_2) \cdots \rangle_{\hat{x}_0}, \quad (31)$$

where the firing rate of \hat{x}_0 is

$$\hat{\nu}(t) = \nu(t)(1 - F_1)^k.$$

These terms can be calculated following Stratonovich [20]. For the special cases of either pure facilitation or pure de-

pression, the required averages can be calculated by hand. In the general case of both facilitation and depression present, we used a Python script to compute the formulas for the averages.

Note that the average of *any* product of F , D , and x_0 can be derived in that way—for instance, the average synaptic amplitude

$$\langle A \rangle \approx \langle \langle F_{\text{lin}}(t) D_{\text{lin}}(t) \rangle_{x_0} \rangle_R.$$

Remarkably, the resulting expression for $\langle A \rangle$ turns out to be the same as that derived in [13], although different approximations have been made.

C. Contributions to C_{RX}

It is convenient to introduce the following dimensionless single synapse spectra (approximating the spectra to the lowest significant order in ε)

$$\hat{s}_{xx}(f) \approx \frac{S_{xx}(f)}{r}, \quad \hat{s}_{Rx}(f) \approx \frac{S_{Rx}(f)}{\varepsilon r S_{RR}(f)}. \quad (32)$$

Without short-term plasticity, they become exactly one for all frequencies: $\hat{s}_{x_0x_0}(f) = \hat{s}_{Rx_0}(f) = 1$ (this can be derived as explained in Sec. III B or as previously described in [30,32]). Short-term plasticity turns them into frequency-dependent functions.

We express Eq. (15) by means of these dimensionless spectra. Using $C_{R\langle x \rangle}(f) \approx 1$ (see Appendix A 2), we obtain to the most significant order in ε

$$\begin{aligned} \frac{1}{C_{RX}(f)} &\approx \frac{1}{N} \times \frac{1}{\varepsilon^2 r S_{RR}(f)} \times \frac{\hat{s}_{xx}(f)}{|\hat{s}_{Rx}(f)|^2} + \frac{N-1}{N} \\ &+ \frac{1}{N} \times \frac{1}{\varepsilon^2 r S_{RR}(f)} \times \frac{S_{\eta\eta}(f)}{Nr |\hat{s}_{Rx}(f)|^2}. \end{aligned} \quad (33)$$

A number of interesting results are apparent in this equation. First, a small ε leads to a large first and third term in Eq. (33). Hence, a weak influence of the external stimulus on the rate modulation diminishes linear information transmission. Similarly, a low firing rate r combined with a low power spectrum of the external stimulus S_{RR} decreases the linear information transmission. Finally, a large number of input neurons N improves linear information transmission, in particular when $N\varepsilon^2 r S_{RR} \geq 1$. Moreover, one can always bring the coherence arbitrarily close to 1 by increasing N sufficiently. In the following two sections, we investigate the effect of each of the three terms in Eqs. (15) and (33).

IV. RESULTS FOR A SINGLE SYNAPSE

Without noise [$\eta(t) \equiv 0$] and for $N=1$, only the first term in Eq. (15) remains. In this section, we present our results for $C_{Rx}(f)$ and spectra $S_{xx}(f)$ and $S_{Rx}(f)$. We calculate everything to the lowest significant order in ε (results that have been approximated in such a way are marked by “ \approx ”). Simulations show that for $\varepsilon \leq 0.2$, this is sufficient to characterize the spectra (see Figs. 5, 7–10, and 13–16). This can be justified by the fact that all terms in S_{xx} and in $|S_{Rx}|^2$ vanish that

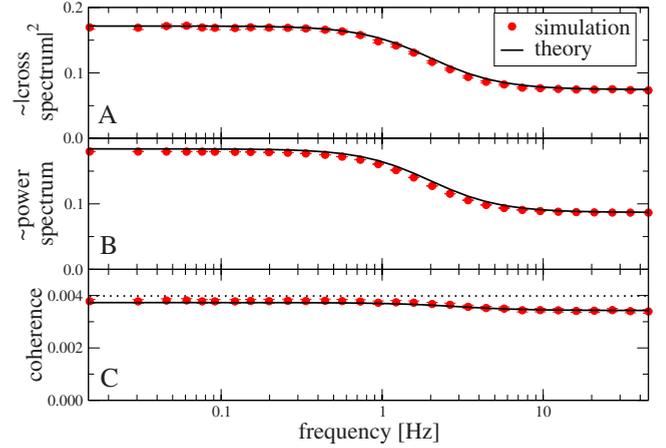


FIG. 5. (Color online) $|\hat{s}_{Rx}(f)|^2$, $\hat{s}_{xx}(f)$, and $C_{Rx}(f)$ for the case of facilitation only. The parameter values are $F_0=0.1$, $\Delta=0.3$, $r=10$ Hz, $\tau_F=80$ ms, and $\varepsilon=0.2$. Circles show the results of a simulation without any approximations made (error bars within symbol size); while the solid lines illustrate the theoretical predictions expressed in Eqs. (38), (41), and (34), respectively. The spectra show pronounced low-pass behavior while the coherence function is rather flat. The dotted coherence curve shows the coherence for the case without any synaptic plasticity [i.e., $F(t)=D(t)=1$].

are odd with respect to ε . This is a direct consequence of the two facts that on the one hand, ε accompanies R as factor; and on the other hand, the distribution of R is symmetric under sign reversal. In addition, we always assume $f \neq 0$ (i.e., we will ignore the DC components of the spectra).

For convenience, we will use the dimensionless versions $\hat{s}_{xx}(f)$ and $\hat{s}_{Rx}(f)$ of the spectra as defined in Eq. (32). The coherence without noise is then given by Eq. (33) with $N=1$ and $S_{\eta\eta}=0$,

$$C_{Rx}(f) \approx \varepsilon^2 r S_{RR}(f) \frac{|\hat{s}_{Rx}(f)|^2}{\hat{s}_{xx}(f)}. \quad (34)$$

Note that for a static synapse,

$$C_{Rx}(f) = C_{Rx_0}(f) \approx \varepsilon^2 r S_{RR}(f), \quad (35)$$

because in that case, $\hat{s}_{x_0x_0}(f) = \hat{s}_{Rx_0}(f) = 1$ (see above). We discuss the special cases of pure facilitation and pure depression, before we make some remarks about the general case.

A. Pure facilitation

For the pure facilitation case, we use

$$A_i(t) = F_{\text{lin},i}(t) \quad (36)$$

instead of Eq. (4). In other words, we ignore the depression variable ($D(t) \equiv 1$) and use the linearized F -dynamics from Appendix C. The facilitation-only dynamics is exemplified in Fig. 2. The results are discussed in the following and illustrated in Figs. 5–7.

For the cross-spectrum and its absolute square, we obtain

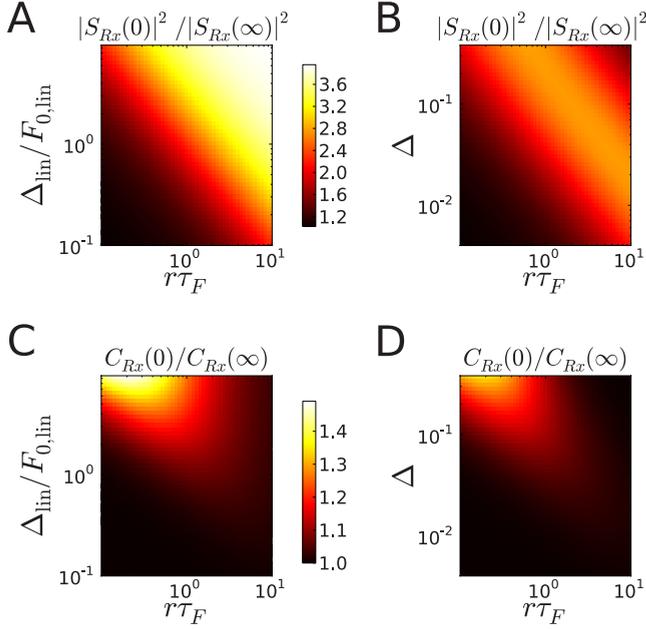


FIG. 6. (Color online) Theoretical results for $|S_{Rx}(0)|^2/|S_{Rx}(\infty)|^2$ and for $C_{Rx}(0)/C_{Rx}(\infty)$. (A) $|S_{Rx}(0)|^2/|S_{Rx}(\infty)|^2$ depending on $r\tau_F$ and the facilitation intensity expressed by the parameters for linear facilitation dynamics $\Delta_{lin}/F_{0,lin}$. (B) $|S_{Rx}(0)|^2/|S_{Rx}(\infty)|^2$ for $F_0=0.04$, depending on $r\tau_F$ and Δ . (C) $C_{Rx}(0)/C_{Rx}(\infty)$ depending on $r\tau_F$ and $\Delta_{lin}/F_{0,lin}$. (D) $C_{Rx}(0)/C_{Rx}(\infty)$ for $F_0=0.04$, depending on $r\tau_F$ and Δ . The Δ axes in (B) and (D) are scaled such that Δ/F_0 ranges from 0.1 to 10 like $\Delta_{lin}/F_{0,lin}$ in (A) and (C).

$$\hat{s}_{Rx}(f) \approx F_1 + \frac{\Delta_{lin} r \tau_F}{1 - 2\pi i f \tau_F}, \quad (37)$$

and

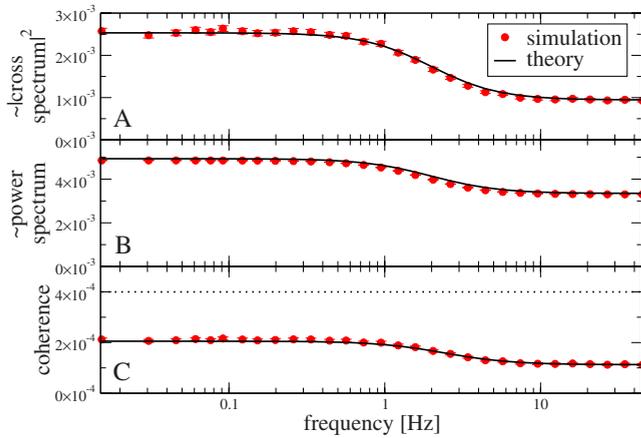


FIG. 7. (Color online) $|\hat{s}_{Rx}(f)|^2$, $\hat{s}_{xx}(f)$, and $C_{Rx}(f)$ for the case of facilitation only. The parameter values are $F_0=0.01$, $\Delta=0.3$, $r=1$ Hz, $\tau_F=80$ ms, and $\varepsilon=0.2$. Circles show the results of a simulation without any approximations made (error bars within symbol size); while the solid lines illustrate the theoretical predictions expressed in Eqs. (38), (41), and (34), respectively. The spectra and the coherence function show pronounced low-pass behavior. The dotted coherence curve shows the coherence for the case without any synaptic plasticity [i.e., $F(t)=D(t)=1$].

$$|\hat{s}_{Rx}(f)|^2 \approx \frac{1}{1 + (2\pi f \tau_F)^2} \left[(F_1 + \Delta_{lin} r \tau_F)^2 + (2\pi f \tau_F)^2 F_1^2 \right], \quad (38)$$

where

$$F_1 = F_{0,lin} + \Delta_{lin} r \tau_F. \quad (39)$$

The absolute square of the cross-spectrum $|\hat{s}_{Rx}(f)|^2$ can be understood as the sum of a constant plus a Lorentzian low-pass spectrum, i.e., pure facilitation has a low-pass effect. The low- and high-frequency limits of the squared cross-spectrum are given by

$$|\hat{s}_{Rx}(f)|^2 \approx \begin{cases} (F_{0,lin} + 2\Delta_{lin} r \tau_F)^2 & \text{for } f \rightarrow 0 \text{ and} \\ (F_{0,lin} + \Delta_{lin} r \tau_F)^2 & \text{for } f \rightarrow \infty. \end{cases} \quad (40)$$

Hence, the low-pass behavior is most pronounced for a high ratio $\Delta_{lin} r \tau_F / F_{0,lin}$ (i.e., for high rates and for high Δ_{lin} compared to $F_{0,lin}$). The ratio between the two limits is shown in Fig. 6(A) as a function of $r\tau_F$ and $\Delta_{lin}/F_{0,lin}$ and illustrates the parameter region of low-pass behavior of the cross spectrum. In terms of the original parameters Δ and $r\tau_F$ [compare Fig. 6(B)], we see a similar effect: For high input rate and a large increment Δ , the low-pass behavior is most pronounced. An exception to this is observed in the limit of very large Δ and $r\tau_F$, which is due to the saturation effect in the nonlinearity $F(F_C)$ in Eq. (7): for a very high-input rate and high Δ , the effective increment Δ_{lin} decreases with growing Δ and, thus, leads to a weaker facilitation effect and a less pronounced low-pass behavior of the cross spectrum.

Remarkably, the expression for the power spectrum is similar to that of the absolute square of the cross-spectrum. To the lowest significant order in ε , $\hat{s}_{xx}(f)$ equals $|\hat{s}_{Rx}(f)|^2$ up to a shift of $\frac{1}{2}\Delta_{lin}^2 r \tau_F$:

$$\hat{s}_{xx}(f) \approx |\hat{s}_{Rx}(f)|^2 + \frac{1}{2}\Delta_{lin}^2 r \tau_F. \quad (41)$$

Both cross- and power spectra predicted by these formulas show good agreement with the results of numerical simulations (see, for example, Figs. 5 and 7, (A) and (B), respectively). There is a slight deviation in the power spectrum, which is overestimated by Eq. (41). We have verified that this deviation is due to the linearization approximation for the F dynamics outlined in Appendix C.

The coherence $C_{Rx}(f)$ as calculated from cross- and power spectra reads

$$\begin{aligned} C_{Rx}(f) &\approx \varepsilon^2 r S_{RR}(f) \left[1 + \frac{[1 + (2\pi f \tau_F)^2] \Delta_{lin}^2 r \tau_F / 2}{(F_1 + \Delta_{lin} r \tau_F)^2 + (2\pi f \tau_F)^2 F_1^2} \right]^{-1} \\ &= C_{Rx_0}(f) \left[1 + \frac{[1 + (2\pi f \tau_F)^2] \Delta_{lin}^2 r \tau_F / 2}{(F_1 + \Delta_{lin} r \tau_F)^2 + (2\pi f \tau_F)^2 F_1^2} \right]^{-1}. \end{aligned} \quad (42)$$

The latter relation reflects the fact that the coherence of dynamic synapse $C_{Rx}(f)$ is always smaller than that of the static synapse $C_{Rx_0}(f)$, because the inverted square bracket is always smaller than one.

In order to find conditions for broadband coding, we inspect the ratio between low and high-frequency limits of the coherence. This ratio is given by

$$\frac{C_{Rx}(0)}{C_{Rx}(\infty)} = \frac{2 + r\tau_F/(F_{0,\text{lin}}/\Delta_{\text{lin}} + r\tau_F)^2}{2 + r\tau_F/(F_{0,\text{lin}}/\Delta_{\text{lin}} + 2r\tau_F)^2}. \quad (43)$$

Note that this ratio only depends on two dimensionless parameters: $\Delta_{\text{lin}}/F_{0,\text{lin}}$ and $r\tau_F$. The first characterizes the strength of facilitation approaching zero in the absence of facilitation (i.e., for $\Delta_{\text{lin}}=0$). The second parameter, $r\tau_F$, is the ratio of two time scales: the decay time constant of facilitation and the mean interspike interval of the input spike train. The coherence is broadband if one or both of the following conditions are met:

$$\frac{1}{2}r\tau_F \ll (F_{0,\text{lin}}/\Delta_{\text{lin}} + r\tau_F)^2 \quad \text{and} \quad r\tau_F \ll F_{0,\text{lin}}/\Delta_{\text{lin}}. \quad (44)$$

Closer inspection of these inequalities reveals that indeed for most parameter values, the synapse performs broadband coding. The ratio in Eq. (43) shown in Fig. 6(C) as a function of $\Delta_{\text{lin}}/F_{0,\text{lin}}$ and $r\tau_F$ is close to one unless for strong facilitation and for intermediate values of the time scale ratio $r\tau_F$, where the coherence differs by a moderate factor of 1.5 between the low and high-frequency limits. The effect is less pronounced in terms of the original parameters Δ and $r\tau_F$ as depicted in Fig. 6(D) for $F_0=0.04$. An example for a low-pass coherence is shown in Fig. 7, where the coherence suffers a considerable overall reduction compared to the case of a static synapse (dotted line). Note that in this case, the ratio between Δ and F_0 is extremely large—typical values are rather $\Delta/F_0 \leq 5$ [2,8,12].

B. Pure depression

For the pure depression case, we use

$$A_i(t) = F_0 D_{\text{lin},i}(t), \quad (45)$$

instead of Eq. (4) [the facilitation variable is set to $F(t)=F_0$]. Note that the parameter F_0 not only scales the amplitude in Eq. (45), but also the multiplicative decrease in the D variable [compare Fig. 3 and Eq. (8)]. The results calculated to the lowest significant order in ε , are illustrated in Fig. 8 and discussed in the following.

We obtain for the cross-spectrum and its absolute square:

$$\hat{s}_{Rx}(f) \approx \frac{F_0}{\beta} \left(1 - \frac{F_0 r \tau_D / \beta}{1 - 2\pi i f \tau_D / \beta} \right) \quad (46)$$

and

$$|\hat{s}_{Rx}(f)|^2 \approx \frac{F_0^2}{\beta^4} \times \frac{1 + (2\pi f \tau_D / \beta)^2 \beta^2}{1 + (2\pi f \tau_D / \beta)^2}, \quad (47)$$

with

$$\beta = 1 + F_0 r \tau_D. \quad (48)$$

The absolute square of the cross-spectrum $|\hat{s}_{Rx}(f)|^2$ can be regarded as a constant minus a Lorentzian spectrum leading

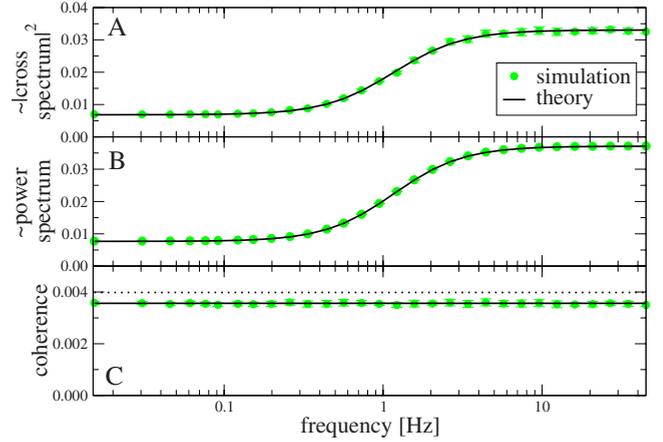


FIG. 8. (Color online) $|\hat{s}_{Rx}(f)|^2$, $\hat{s}_{xx}(f)$, and $C_{Rx}(f)$ for the case of depression only. The parameter values are $F_0=0.4$, $r=10$ Hz, $\tau_D=300$ ms, and $\varepsilon=0.2$. Circles show the results of a simulation without any approximations made (error bars within symbol size); while the solid lines illustrate the theoretical predictions expressed in Eqs. (47), (50), and (34). The spectra show pronounced high-pass behavior while the coherence function is perfectly flat. The dotted coherence curve shows the coherence for the case without any synaptic plasticity [i.e., $F(t)=D(t)=1$].

to a high-pass effect of the transfer function for pure depression. The low- and high-frequency limits are given by

$$|\hat{s}_{Rx}(f)|^2 \approx \begin{cases} F_0^2/\beta^4 & \text{for } f \rightarrow 0 \text{ and} \\ F_0^2/\beta^2 & \text{for } f \rightarrow \infty. \end{cases} \quad (49)$$

Hence, the high-pass behavior is most pronounced for large $F_0 r \tau_D$ (i.e., for strong depression and high rates).

The power spectrum $\hat{s}_{xx}(f)$ shows (up to a constant factor) exactly the same behavior,

$$\hat{s}_{xx}(f) \approx |\hat{s}_{Rx}(f)|^2 \left[1 - \frac{F_0^2 r \tau_D}{2\beta} \right]^{-1}. \quad (50)$$

This results in a coherence function, that is does not depend on frequency at all

$$C_{Rx}(f) \approx \varepsilon^2 r S_{RR}(f) \left[1 - \frac{F_0^2 r \tau_D}{2\beta} \right] = C_{Rx_0}(f) \left[1 - \frac{F_0^2 r \tau_D}{2\beta} \right]. \quad (51)$$

The broadband coherence for pure depression is a nontrivial result. As in the case of pure facilitation, one can also verify for the case of pure depression that the coherence $C_{Rx}(f)$ is lower than that of a static synapse $C_{Rx_0}(f)$. The reduction compared to the static case can be neglected

$$C_{Rx}(f) \approx C_{Rx_0}(f), \quad \frac{F_0^2 r \tau_D}{2\beta} \ll 1. \quad (52)$$

Hence, a substantial reduction of the coherence is only expected for high F_0 and moderate to large $r\tau_D$.

C. General case

For the general case, we use the full coupled FD dynamics according to Eq. (4), which is illustrated in Fig. 4. For

both—facilitation and depression—present, our results for power and cross-spectra become very involved.

To the lowest significant order in ε , we obtain the following general structure for the absolute square of the cross-spectrum,

$$|\hat{s}_{Rx}(f)|^2 \approx \hat{s}_{c,\infty} + \sum_{k=1}^4 \frac{\hat{s}_{c,k}}{1 + (2\pi f \tau_k)^2}. \quad (53)$$

Here, $\hat{s}_{c,\infty}$ is a constant indicating the high-frequency limit of $|\hat{s}_{Rx}(f)|^2$. The constants $\hat{s}_{c,k}$ (with $k=1, \dots, 4$) are factors for Lorentzian terms while the τ_k are the corresponding time constants. The expressions for these four time constants are simple

$$\begin{aligned} \tau_1 &= \tau_F, \\ \tau_2 &= \frac{\tau_D}{1 + F_1 r \tau_D}, \\ \tau_3 &= \tau_F/2, \\ \tau_4 &= \frac{\tilde{\tau}}{1 + F_1 r \tilde{\tau}}, \end{aligned} \quad (54)$$

where $F_1 = F_{0,\text{lin}} + \Delta_{\text{lin}} r \tau_F$ and $\tilde{\tau} = [\tau_F^{-1} + \tau_D^{-1}]^{-1}$. The expressions for the constants $\hat{s}_{c,\infty}$ and $\{\hat{s}_{c,k}\}$ are too cumbersome to be presented here.

The general expression of the power spectrum is of the same kind as $|\hat{s}_{Rx}(f)|^2$,

$$\hat{s}_{xx}(f) \approx \hat{s}_{p,\infty} + \sum_{k=1}^4 \frac{\hat{s}_{p,k}}{1 + (2\pi f \tau_k)^2}. \quad (55)$$

The time constants τ_k of the Lorentzians are the same as for $|\hat{s}_{Rx}(f)|^2$. However, the coefficients in Eq. (55) differ from those in Eq. (53), are rather lengthy (the complete formula amounts to ~ 100 kB of Python code), and will not be stated here explicitly.

As discussed previously [13], one can find parameter ranges where facilitation or depression will dominate and where the spectral statistics are similar to the cases of pure facilitation and depression discussed above. However, we can also find parameter sets, where new spectral features are observed, in particular if the time scales of facilitation and depression differ significantly. Examples are displayed in Figs. 9 and 10, they show a minimum or a maximum in the power and cross-spectra, which are predicted by theory and confirmed by simulations. Substantial deviations are seen in Fig. 10(A). Closer inspection reveals that these deviations are mainly due to the D_{lin} approximation (cf. discussion in Appendix D).

In order to characterize the changes induced by STP over a large part of the parameter space, we consider again the ratio between low and high frequency limits of power and cross-spectra (see Fig. 11). We find that the low-pass property of both, $|\hat{s}_{Rx}(f)|^2$ and $\hat{s}_{xx}(f)$, behave similar over the whole domain shown. For moderate $r\tau_F$ and low $r\tau_D$, we find a strong low-pass behavior. In contrast to that, for high $r\tau_D$, we find a strong high-pass behavior. The results of the simu-

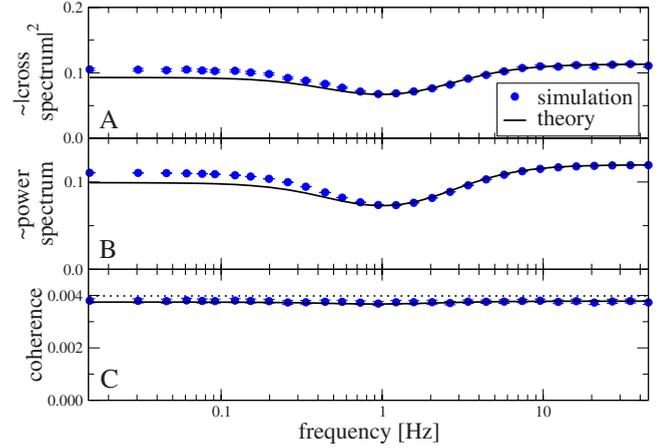


FIG. 9. (Color online) $|\hat{s}_{Rx}(f)|^2$, $\hat{s}_{xx}(f)$, and $C_{Rx}(f)$ for the general case (facilitation and depression). The parameter values are $F_0 = 0.1$, $\Delta = 0.3$, $r = 10$ Hz, $\tau_F = 300$ ms, $\tau_D = 100$ ms, and $\varepsilon = 0.2$. Circles show the results of a simulation without any approximations made (error bars within symbol size); while the solid lines illustrate the theoretical predictions. The spectra show a pronounced minimum while the coherence function is rather flat. There is only a slight deviation of the spectra between theory and simulations for low frequencies that cancel out for the coherence. The dotted coherence curve shows the coherence for the case without any synaptic plasticity [i.e., $F(t) = D(t) = 1$].

lations are in general in a good accordance with the theory—except for the aforementioned deviation, due to the D_{lin} approximation, which comes into play for high $r\tau_D$.

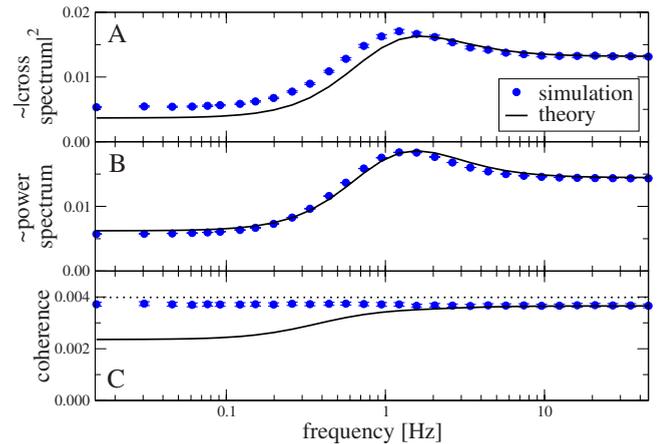


FIG. 10. (Color online) $|\hat{s}_{Rx}(f)|^2$, $\hat{s}_{xx}(f)$, and $C_{Rx}(f)$ for the general case (facilitation and depression). The parameter values are $F_0 = 0.1$, $\Delta = 0.3$, $r = 10$ Hz, $\tau_F = 100$ ms, $\tau_D = 500$ ms, and $\varepsilon = 0.2$. The circles show the results of a simulation without any approximations made (error bars within symbol size); while the solid lines illustrate the theoretical predictions. The spectra show a pronounced maximum, while the simulated coherence is flat again. For the power spectrum, theory and simulation are in good accordance. However, the cross-spectrum shows a deviations for low frequencies, which leads to deviations in the coherence function. However, the coherence function as determined from simulations is rather flat. The dotted coherence curve shows the coherence for the case without any synaptic plasticity [i.e., $F(t) = D(t) = 1$].

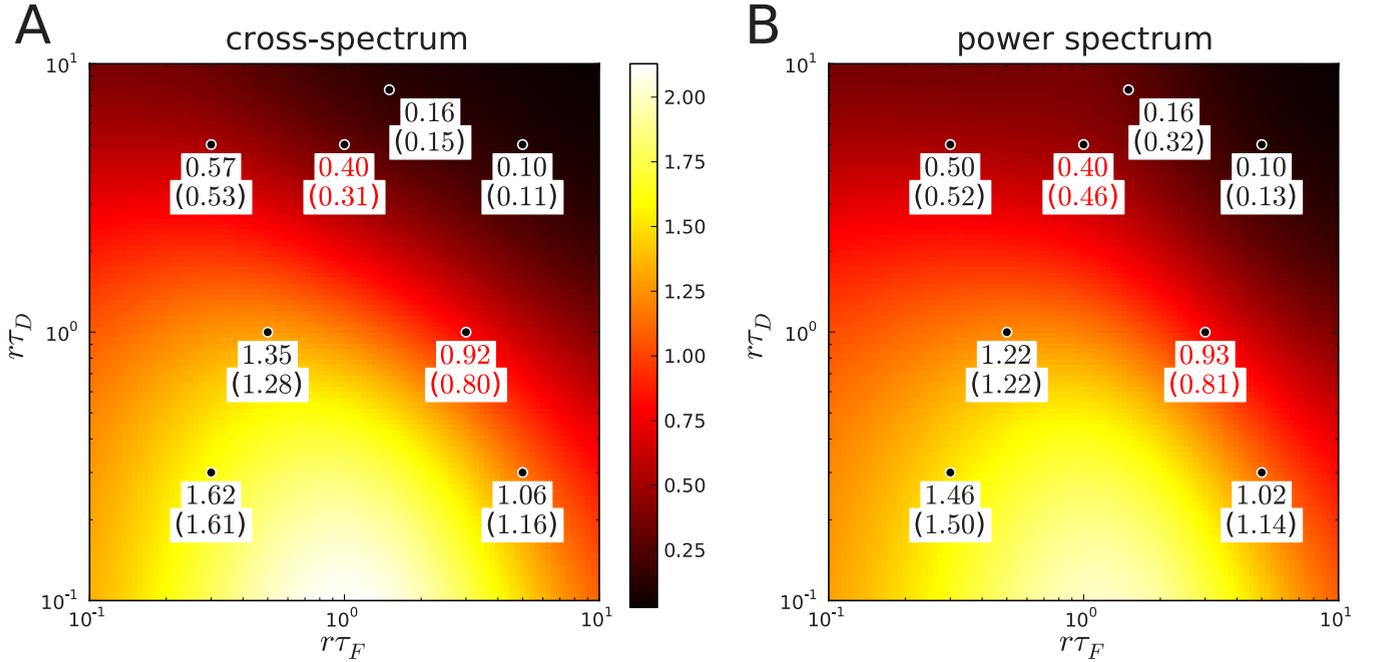


FIG. 11. (Color online) Theoretical results for (A) $|S_{Rx}(0)|^2/|S_{Rx}(\infty)|^2$ and (B) $S_{xx}(0)/S_{xx}(\infty)$. The low-pass property is shown in dependence of $r\tau_F$ and $r\tau_D$ for a large part of the physiological parameter space. The dots with the annotations show simulation results for the ratios between the values at two different frequencies (simulation uncertainties on the left side: less than 3%; on the right side: less than 0.1%), while the values in parentheses are the corresponding theoretical results. The values of the spectra have been calculated at the frequencies 1.5 mHz and 25 Hz. They have been chosen so that according to Eqs. (53)–(55), the spectra do not vary much outside of this interval. The dots with the red (gray) annotations correspond to the simulations shown in Figs. 9 and 10. Here, the frequency values 15 mHz and 27 Hz have been used. We simulated at constant $r=1$ Hz varying τ_F and τ_D . For theory and simulations, we used the parameter values $F_0=0.1$ and $\Delta=0.3$.

Turning to the coherence function, the theory predicts for a large part of the physiological parameter space broadband coding, i.e., the coherence between low and high frequencies deviate not more than 10% (see Fig. 12). For low $r\tau_D$, facilitation seems to dominate and consequently, we find a mild low-pass behavior which is also confirmed by simulations (compare the numbers in Fig. 12). At large $r\tau_D$ and moderate $r\tau_F$, the theory predicts a strong high-pass behavior of the coherence, which is not confirmed by simulations. An example for such a parameter set is shown in Fig. 10(C), which illustrates how the deviations in cross- and power spectra between theory and simulations lead to a strong error in their ratio, i.e., in the coherence function. We find from simulations, that in this region of parameter space, the coherence is rather flat. The above simulations and analytical calculations were carried out for fixed $F_0=0.1$ and $\Delta=0.3$. We have observed, that for smaller values of Δ , differences between low- and high-frequency values are even smaller as well as the general discrepancies between theory and simulations. Put differently, the parameter set considered in Fig. 10 represents an extreme example.

In conclusion, while for parameter regimes of dominating facilitation or depression, we find low-pass and high-pass behavior, respectively, for the cross-spectrum, the coherence function for a single synapse is largely independent of frequency.

V. RESULTS FOR MANY SYNAPSES AND ADDITIONAL NOISE

Let us state again Eq. (15),

$$\frac{1}{C_{RX}(f)} = \frac{1}{N} \times \frac{1}{C_{Rx}(f)} + \frac{N-1}{N} \times \frac{1}{C_{R(x)}(f)} + \frac{1}{N} \times \frac{1}{C_{Rx}(f)} \times \frac{S_{\eta\eta}(f)}{NS_{xx}(f)}.$$

So far we have seen that the first term, which is essentially determined by the coherence for a single synapse, is largely independent of frequency. In this section, we want to examine the influence of the second and the third term in the above equation, reflecting the effects of a multitude of synapses and of an additional noise. In particular, we want to know whether the broadband property of $C_{RX}(f)$ is maintained or even enhanced for many synapses and/or in the presence of noise.

A. Many synapses without additional noise

In general, averaging over many independent spike trains with common rate modulation improves the coherence with this rate modulation and thus, the coherence increases with N . Here we ask how the coherence increases with N and how its frequency dependence is affected by N . To this end, we

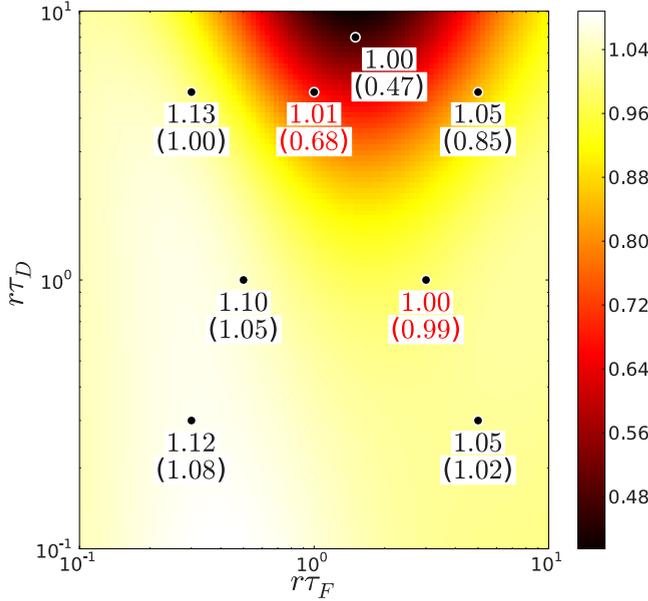


FIG. 12. (Color online) Theoretical results for the low-pass property of the coherence function $C_{RX}(f)$ expressed by the ratio of the low- and the high-frequency limit: $C_{RX}(0)/C_{RX}(\infty)$. The low-pass property is shown in dependence of $r\tau_F$ and $r\tau_D$ for a large part of the physiological parameter space. The dots with the annotations show simulation results for the ratios between the values at two different frequencies (simulation uncertainties: less than 3%), while the values in parentheses are the corresponding theoretical results. The dots with the red (gray) annotations correspond to the simulations shown in Figs. 9 and 10. Theory and simulation parameters are the same as in Fig. 11.

consider Eq. (15) without additional noise [$S_{\eta\eta}(f)=0$],

$$\frac{1}{C_{RX}(f)} = \frac{1}{N} \times \frac{1}{C_{Rx}(f)} + \frac{N-1}{N} \times \frac{1}{C_{R(x)}(f)}. \quad (56)$$

We would like to recall that $C_{R(x)}(f) \approx 1$ [see Eq. (A7) in Sec. A 2].

It is instructive to consider the two limiting cases of small and large N . If N is small, we find that we can neglect the second term in Eq. (56),

$$C_{RX}(f) \approx NC_{Rx}(f), \quad NC_{Rx} \ll 1. \quad (57)$$

Hence, for small N , the multiple synapse coherence grows proportionally with the number of synapses. Therefore, any low-pass or high-pass behavior in $C_{RX}(f)$ with respect to frequency (i.e., deviations from broadband coding) is preserved in this case.

In the opposite limit of large N , we obtain,

$$C_{RX}(f) \approx 1 - \frac{1}{NC_{Rx}(f)}, \quad NC_{Rx} \gg 1. \quad (58)$$

Thus, the coherence approaches 1 by a N^{-1} correction term. This describes the saturation in the limit of an infinite number of synapses in which a perfect noiseless signal transfer is achieved. The very same saturation effect is responsible for a suppression of a possible frequency dependence. In other words, the relative difference between low- and high-

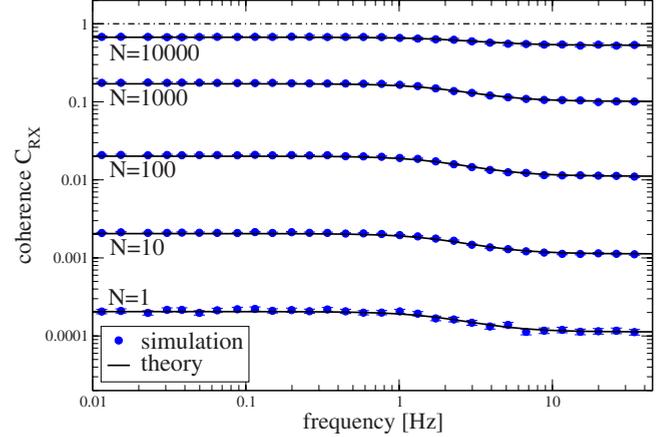


FIG. 13. (Color online) Influence of the number N of synapses on the coherence function $C_{RX}(f)$. The synapses are purely facilitating with parameters as in Fig. 7. We assume zero noise $\eta=0$. Simulations (circles) agree well with the theoretical curves (black solid lines) corresponding to Eq. (56). For $N \leq 10^3$, the coherence function is proportional to N preserving the low-pass property [compare Eq. (57)]. However, for $N=10^4$, we see that $C_{RX}(f)$ approaches 1 and starts to flatten out.

frequency limits of the coherence decreases for growing N . This indicates an enhancement of the broadband property for many synapses. In particular, the absolute and relative differences $\delta C = |C(f_1) - C(f_2)|$ and $\delta C / C(f_1)$ decrease with N (for arbitrary f_1 and f_2):

$$\delta C_{RX} \approx \frac{\delta C_{Rx}}{NC_{Rx}^2(f_1)}, \quad NC_{Rx} \gg 1 \quad (59)$$

and

$$\frac{\delta C_{RX}}{C_{RX}(f_1)} \approx \frac{1}{NC_{Rx}(f_1) - 1} \times \frac{\delta C_{Rx}}{C_{Rx}(f_1)} \ll \frac{\delta C_{Rx}}{C_{Rx}(f_1)}, \quad NC_{Rx} \gg 1. \quad (60)$$

In order to further illustrate the N dependence of $C_{RX}(f)$, we consider a situation of N facilitating input synapses. The model parameters are chosen as in Fig. 7, so that the single synapse coherence $C_{Rx}(f)$ exhibits a mild low-pass behavior. Simulations and theory for different N are shown in Fig. 13. With the parameters given, we have $C_{Rx} \sim 10^{-4}$. For $N \ll 1/C_{Rx} \sim 10^4$, the coherence function $C_{RX}(f)$ is proportional to N . In a logarithmic plot, this manifests itself in the fact that the curves for $N=1$ to 1000 are shifted versions of each other. However, if $N \gtrsim 1/C_{Rx} \sim 10^4$, the coherence function $C_{RX}(f)$ approaches one and becomes more flat.

B. Many synapses and additional noise

Additional noise may arise from spike train input (not modulated by the signal) stemming from other neurons. These inputs will enter via synapses that we refer to as *noise synapses*; equivalently, the rate-modulated spike trains enter via the *signal synapses*. Noise synapses and signal synapses may share the same STP character, i.e., both may be facilitation- or depression-dominated. They may however

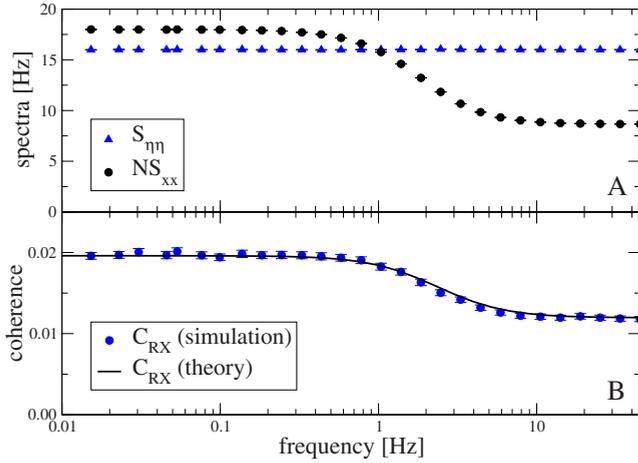


FIG. 14. (Color online) Influence of noise η on the coherence function $C_{RX}(f)$. We simulate $N=10$ facilitating input synapses with $F_0=0.1$, $\Delta=0.3$, $r=10$ Hz, $\tau_f=80$ ms, and $\varepsilon=0.2$. The noise η originates in 10 static synapses with amplitude $A_{\text{Noise}}=0.4$ each of which is driven by a Poissonian process with constant rate $r_{\text{Noise}}=10$ Hz. The resulting coherence is well predicted by Eq. (15) (black line). In this situation, a low-pass power spectrum of the input synapses together with a flat power spectrum of the additional noise leads to a low-pass coherence function.

also differ, e.g., signal synapses may be depressing while noise synapses may be facilitating.

For the coherence, we can again distinguish between the limiting cases of small N and large N . For small N , we obtain,

$$C_{RX}(f) \approx NC_{Rx}(f) \left[1 + \frac{S_{\eta\eta}(f)}{NS_{xx}(f)} \right]^{-1} \quad \text{for } NC_{Rx} \left[1 + \frac{S_{\eta\eta}}{NS_{xx}} \right]^{-1} \ll 1, \quad (61)$$

while for large N , the coherence reads,

$$C_{RX}(f) \approx 1 - \frac{1}{NC_{Rx}(f)} \left[1 + \frac{S_{\eta\eta}(f)}{NS_{xx}(f)} \right] \quad \text{for } NC_{Rx} \left[1 + \frac{S_{\eta\eta}}{NS_{xx}} \right]^{-1} \gg 1. \quad (62)$$

In both expressions, it becomes obvious that additional noise always reduces the coherence between the total postsynaptic input and the rate modulation. Furthermore, let us assume that $S_{\eta\eta}$ results from Poissonian spike train input without rate modulation, but with similar baseline rate r and via similar FD dynamics for the involved synapses. Then, the additional noise terms in Eqs. (61) and (62) do not introduce any additional frequency dependence of the coherence. In these cases, the broadband coherence of synaptic filtering is maintained.

We may, however, also consider, that noise arrives via static synapses or via dynamic synapses which are different from the synapses that transmit the signal $R(t)$. In either case, the ratio $S_{\eta\eta}(f)/[NS_{xx}(f)]$, if equal or larger than one, introduces a frequency dependence into the coherence function. This is illustrated for static noise synapses in Fig. 14, which

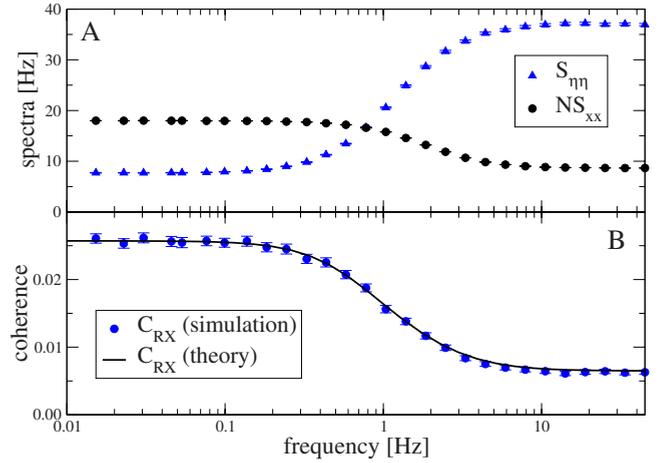


FIG. 15. (Color online) Influence of noise η on the coherence function $C_{RX}(f)$. We simulate $N=10$ facilitating input synapses with parameters as in Fig. 14. The noise η originates in 100 depressing synapses with $F_{0,\text{Noise}}=0.4$ and $\tau_{D,\text{Noise}}=300$ ms each of which is driven by a Poissonian process with constant rate $r_{\text{Noise}}=10$ Hz. The resulting coherence is well predicted by Eq. (15). Comparing to Fig. 14, a high-pass noise power spectrum leads to an even stronger low-pass behavior of the coherence function.

lead in conjunction with facilitating signal synapses to a low-pass coherence function. The reduction of coherence at high frequencies results essentially from the frequency-independent noise. Similarly, for static noise synapses and depressing signal synapses, one observes a high-pass coherence function (not shown). The above filtering effects are even more enhanced, if noise and signal synapses have opposite STP character. For instance, facilitating signal synapses and depressing noise synapses lead to a strong low-pass filtering of information as quantified by the coherence function (see Fig. 15). The opposite case is illustrated in Fig. 16: depressing signal synapses and facilitating noise synapses will lead to a pronounced high-pass filtering of the coherence.

We note that this filtering effect is largest for a single signal synapse. As discussed in the previous section, also in the presence of additional noise, the coherence increases but also flattens out (with respect to frequency) by increasing the number of signal synapses. To see a significant filter effect on the coherence without reducing its magnitude too much, the number of signal synapses should be intermediate, i.e., it should be roughly about $N \sim [1 + S_{\eta\eta}/(NS_{xx})]/C_{Rx}$.

VI. SUMMARY AND DISCUSSION

In this paper, we developed a method to calculate spectral measures of spike trains transmitted by dynamic synapses showing facilitation and depression. Mathematically, dynamic synapses assign a history-dependent amplitude to each spike. Besides this amplitude modulation, we also considered a weak modulation of the firing rate by a time-dependent signal. We described, how to compute time-dependent averages of these postsynaptic responses. These measures can then be used in order to obtain an expression for the coher-

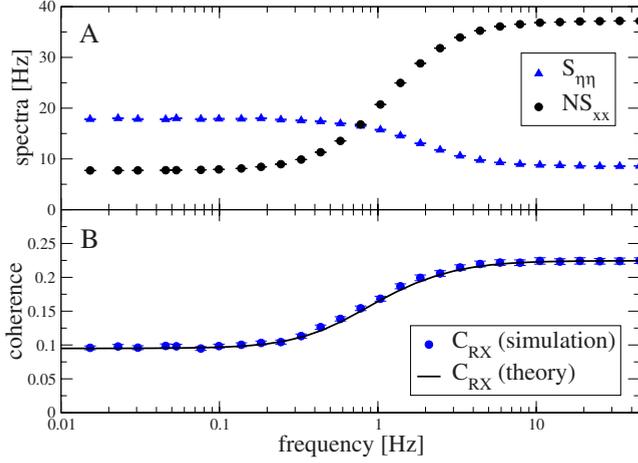


FIG. 16. (Color online) Influence of noise η on the coherence function $C_{RX}(f)$. We simulate $N=100$ depressing input synapses with $r=10$ Hz, $F_0=0.4$, and $\tau_D=300$ ms. The noise η originates in 10 facilitating synapses with parameters as in Fig. 14; each of which is driven by a Poissonian process with constant rate $r_{Noise}=10$ Hz.

ence function between the amplitude-modulated spike train and the rate modulation. By means of the coherence function, we estimated the effect of STP on information transfer through the synapse. Although we treated the rate modulation in linear response theory, assuming a small relative amplitude $\varepsilon \ll 1$, the obtained expressions were also shown to be valid for moderate amplitudes ($\varepsilon=0.2$).

We systematically examined spectral information measures and found that if the presynaptic population is modulated with a time-dependent signal, STP does not filter the information content about such a rate modulation (broadband coding). We thus confirmed the numerical results from [13].

In examining this effect, we observed that for a single synapse, the coherence function is flat for various parameter sets chosen from the physiological parameter space. Furthermore, analytical arguments revealed that a growing number of synapses does not only increase the information transmission, but also decreases any frequency-dependent information filtering effect.

However, for an extended scheme with additional noise (i.e., background synaptic activity), we showed that substantial information filtering becomes possible. This kind of filtering is particularly pronounced if the short-term plasticity character of signal and noise synapses differ (e.g., one is facilitating and the other one is depressing or vice versa). The synaptic background input can switch the information transfer from broadband coding to frequency-dependent filtering.

Our results make predictions about spectral measures that are accessible in experiments. For example, according to Sec. II E, the coherence C_{RV} of the subthreshold membrane potential in the absence of postsynaptic spiking is to a good approximation given by the coherence of the total synaptic input C_{RX} , which we studied in this paper. Broadband coding for the subthreshold membrane voltage under Poissonian rate modulation can be verified or falsified experimentally. Furthermore, according to Sec. II E, power and cross-spectra

would be simple low-pass-filtered versions of the synaptic input spectra which we discussed here. Fitting, for instance, power spectra to experimental data obtained for Poissonian stimulation may allow for an estimate of the parameters characterizing STP. Note, that experiments *in vitro* are closer to the single-synapse scenario, which we focused on in Sec. IV. This is so, because upon stimulation, a number of presynaptic input fibers receive one and the same spike train [11,12].

In our study, we made a number of simplifying assumptions. First of all, we considered Poisson statistics for the presynaptic spike trains. If the firing statistics is not exactly but close to Poissonian statistics, we do not expect any drastic changes of the main conclusions of the present study. Our results do not apply, however, to pacemakerlike or bursting presynaptic inputs (for a study on the latter in the context of STP, see [33]).

Second, we assumed statistical independence of the presynaptic spike trains (apart from the common rate modulation). Weak correlations between input spike trains (as observed *in vitro* [34]) will presumably diminish the beneficial averaging across synapses. In this case, the effective number N_{eff} of statistically independent synapses is smaller than their actual number. For a study on the interplay between input cross-correlations and STP, see [35].

A third limitation of our results is due to neglecting the stochasticity of transmitter release. In a more detailed model, the synaptic amplitude A would be drawn from a probability distribution. We expect that the difference between stochastic synaptic response and the deterministic response can qualitatively be treated as noise. Therefore, the coherence C_{RX} will be reduced when considering stochastic synapses. However, for many release sites with similar parameters contributing to transmission, this additional transmission noise is weak.

The analytical methods developed in this paper can in principle be used to treat a broad range of models, for instance the model by Markram and Tsodyks [7], models using multiple facilitation [9] or multiple depression components [5]. These may be interesting subjects of future investigations.

APPENDIX A: EXPRESSIONS FOR SPECIFIC COHERENCE FUNCTIONS

In this section, we derive Eq. (15) for C_{RX} and show that $C_{R(x)}$ equals one to leading order in ε .

1. The coherence C_{RX}

Using Eq. (12), we write for the cross spectrum $S_{R(x)}(f)$ between R and $\langle x(t) \rangle_{x_0}$

$$\begin{aligned} \delta(f-f')S_{R(x)}(f) &= \left\langle \tilde{R}(f) \widetilde{\langle x \rangle_{x_0}}^*(f') \right\rangle_R = \left\langle \tilde{R}(f) \tilde{x}^*(f') \right\rangle_{x_0,R} \\ &= \delta(f-f')S_{RX}(f). \end{aligned} \quad (\text{A1})$$

Hence,

$$S_{R\langle x \rangle}(f) = S_{R_x}(f), \quad (\text{A2})$$

i.e., the cross spectrum between stimulus R and response equals the cross spectrum between stimulus R and average response $\langle x \rangle_{x_0}$.

Similarly, we put the definition of X [cf. Eq. (9)] into Eq. (12) with $a=b=X$ using statistical independence among the $x_{0,i}$ and statistical independence between $x_{0,i}$ and η ,

$$S_{XX}(f) = N(N-1)S_{\langle x \rangle \langle x \rangle}(f) + NS_{xx}(f) + S_{\eta\eta}(f). \quad (\text{A3})$$

Analogously, we find

$$S_{RX}(f) = NS_{R_x}(f). \quad (\text{A4})$$

The formulas (A3) and (A4) relate the many-synapses spectra to the single-synapse versions.

We combine Eqs. (A2)–(A4) with the definition of the coherence [cf. Eq. (14)] and obtain

$$\frac{1}{C_{RX}(f)} = \frac{N-1}{N} \frac{1}{C_{R\langle x \rangle}(f)} + \frac{1}{N} \frac{1}{C_{R_x}(f)} + \frac{1}{N} \frac{1}{C_{R_x}(f)} \frac{S_{\eta\eta}(f)}{NS_{xx}(f)}. \quad (\text{A5})$$

2. The coherence $C_{R\langle x \rangle}$

We want to evaluate the coherence $C_{R\langle x \rangle}(f)$. To this end, we consider the dependence of the average modulated spike train $\langle x(t) \rangle_{x_0}$ on the external stimulus R . As R enters this dependency only via the rate modulation, it is always accompanied by the factor ε [cf. Eq. (2)]. To the lowest significant order in ε , we obtain a linear relationship in Fourier space:

$$\widetilde{\langle x \rangle}_{x_0}(f) \approx \varepsilon \widetilde{R}(f) \chi(f), \quad (\text{A6})$$

where $\chi(f)$ is the transfer function characterizing this linear transformation. There is no absolute term, because for $\varepsilon=0$, stationarity has to be fulfilled.

Hence, to the lowest order in ε , we obtain a *deterministic linear relation between $R(t)$ and $\langle x(t) \rangle_{x_0}$* , for which the coherence is one

$$C_{R\langle x \rangle}(f) \approx 1, \quad (\text{A7})$$

for all frequencies f .

APPENDIX B: CALCULATION OF POISSONIAN AVERAGES

We show how to calculate averages as in Eq. (28),

$$\langle x_0(t_1)x_0(t_2) \cdots e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \cdots \rangle, \quad (\text{B1})$$

where the times $t_1, t_2, \dots, t_j, t_{j+1}, \dots$ are completely arbitrary (including, that any two of these time variables may be identical), except for [cf. Eq. (27)]

$$t_j \leq t_{j+1}, \quad t_{j+2} \leq t_{j+3}, \quad \dots \quad (\text{B2})$$

We define the term $e(t_a, t_b)$ as

$$e(t_a, t_b) = \exp \left[\ln(1 - F_1) \int_{t_a}^{t_b} x_0(t') dt' \right]. \quad (\text{B3})$$

We recall that the time-dependent rate of x_0 is

$$\nu(t) \approx r[1 + \varepsilon R(t)] \quad (\text{B4})$$

In the simple cases where Eq. (B1) does not contain any $e(\cdot, \cdot)$ -term, we use Eqs. (6.28) and (6.32) from [20] in order to obtain the following scheme,

$$\langle x_0(t_1) \rangle_{x_0} = \nu(t_1),$$

$$\langle x_0(t_1)x_0(t_2) \rangle_{x_0} = \nu(t_1)\nu(t_2) + \delta(t_1 - t_2)\nu(t_1),$$

$$\cdots = \cdots, \quad (\text{B5})$$

which directly yields an expression for Eq. (B1) in terms of r , ε , and R .

For the general case, where $e(\cdot, \cdot)$ terms may be present, there are two principal steps required. First, we express Eq. (B1) in the following subsection by a product of averages of the kind

$$\langle x_0(t_1)x_0(t_2) \cdots e^k(t_s, t_e) \rangle, \quad (\text{B6})$$

where the instants t_1, t_2 are of arbitrary order, but all of them are between t_s and t_e

$$t_s < t_1, t_2, \dots < t_e \quad (\text{B7})$$

and the exponent $k \leq 0$.

Second, in the subsequent subsection, we show how to calculate the average in Eq. (B6).

1. Reduction of Eq. (B1) to terms as in Eq. (B6)

For a case distinction between different orderings of the times appearing in Eq. (B1), we use the Heaviside-function. For instance, for two different times, we can use the identity

$$1 = \Theta(t_1 - t_2) + \Theta(t_2 - t_1) \quad \text{for } t_1 \neq t_2, \quad (\text{B8})$$

which, as a prefactor of the average Eq. (B1), formalizes the distinction between the cases $t_1 > t_2$ and $t_2 > t_1$ corresponding to the two different terms on the right-hand side (rhs) in Eq. (B8). (The case $t_1 = t_2$ may demand additional considerations [36], not presented here.) In general, the products of Heaviside functions cycle through all possibilities of ordering the times. Formally, this can be written as follows:

$$1 = \sum_{\sigma \in P_n} \prod_{i=1}^{n-1} \Theta(t_{\sigma(i)} - t_{\sigma(i+1)}), \quad (\text{B9})$$

where P_n is the set of all permutations of the sequence of integers from 1 to n . As an example, let us consider the computation of $\langle e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \rangle$,

$$\begin{aligned} & \langle e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \rangle \\ &= \Theta(t_{j+2} - t_{j+1}) \langle e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \rangle \\ & \quad + \Theta(t_j - t_{j+3}) \langle e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \rangle \\ & \quad + \Theta(t_j - t_{j+2}) \Theta(t_{j+3} - t_{j+1}) \langle e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \rangle \\ & \quad + \Theta(t_{j+2} - t_j) \Theta(t_{j+1} - t_{j+3}) \langle e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \rangle \\ & \quad + \Theta(t_j - t_{j+2}) \Theta(t_{j+3} - t_j) \Theta(t_{j+1} - t_{j+3}) \\ & \quad \times \langle e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \rangle \end{aligned}$$

$$\begin{aligned}
 & + \Theta(t_{j+2} - t_j)\Theta(t_{j+1} - t_{j+2})\Theta(t_{j+3} - t_{j+1}) \\
 & \times \langle e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \rangle, \quad (\text{B10})
 \end{aligned}$$

where we have already used Eq. (B2) reducing the number of terms on the rhs from 24 to six. For each term, the respective prefactor determines uniquely the time ordering, which is essential to calculate the average properly. For instance, for the term on the first line, we can assume the order $t_j \leq t_{j+1} \leq t_{j+2} \leq t_{j+3}$, and obtain for the average (omitting the Heaviside function)

$$\langle e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \rangle = \langle e(t_j, t_{j+1}) \rangle \times \langle e(t_{j+2}, t_{j+3}) \rangle, \quad (\text{B11})$$

because the averages are taken in nonoverlapping intervals and thus, the two $e(\cdot, \cdot)$ terms are statistically independent. The last term in Eq. (B10) is only nonvanishing for $t_j \leq t_{j+2} \leq t_{j+1} \leq t_{j+3}$, which yields for this specific case

$$\begin{aligned}
 & \langle e(t_j, t_{j+1})e(t_{j+2}, t_{j+3}) \rangle \\
 & = \langle e(t_j, t_{j+2})e^2(t_{j+2}, t_{j+1})e(t_{j+1}, t_{j+3}) \rangle \\
 & = \langle e(t_j, t_{j+2}) \rangle \times \langle e^2(t_{j+2}, t_{j+1}) \rangle \times \langle e(t_{j+1}, t_{j+3}) \rangle. \quad (\text{B12})
 \end{aligned}$$

This example illustrates, that powers of the $e(\cdot, \cdot)$ term arise due to overlapping intervals. If there are additional $x_0(t_i)$ prefactors in the average in Eq. (B1), then we have to do more case distinctions for the respective time instances, which can be carried out by means of Heaviside functions. For example

$$\begin{aligned}
 \langle x_0(t_1)e(t_2, t_3) \rangle & = \Theta(t_2 - t_1) \times \langle x_0(t_1) \rangle \langle e(t_2, t_3) \rangle \\
 & + \Theta(t_3 - t_1)\Theta(t_1 - t_2) \times \langle x_0(t_1)e(t_2, t_3) \rangle \\
 & + \Theta(t_1 - t_3) \times \langle x_0(t_1) \rangle \langle e(t_2, t_3) \rangle, \quad (\text{B13})
 \end{aligned}$$

where we recall, that $t_2 \leq t_3$ by virtue of Eq. (B3).

The techniques presented here can be generalized in order to split averages of the type of Eq. (B1) into products of the type of Eq. (B6). For the necessary case distinctions, one can also use computer algebra software.

2. Calculation of terms as in Eq. (B6)

Here, we outline the proof for

$$\langle x_0(t_1)x_0(t_2) \cdots e^k(t_s, t_e) \rangle = C_k \langle \hat{x}_0(t_1)\hat{x}_0(t_2) \cdots \rangle, \quad (\text{B14})$$

where \hat{x}_0 is a Poissonian δ -spike train with the rate $\hat{\nu}(t)$ defined by

$$\hat{\nu}(t) = \nu(t)(1 - F_1)^k \approx r[1 + \varepsilon R(t)](1 - F_1)^k \quad (\text{B15})$$

and

$$C_k = \exp\left(-r[1 - (1 - F_1)^k] \times \left[(t_e - t_s) + \varepsilon \int_{t_s}^{t_e} R(t') dt'\right]\right). \quad (\text{B16})$$

The terms $\langle \hat{x}_0(t_1)\hat{x}_0(t_2) \cdots \rangle$ are *moment functions* of the Poisson process $\hat{x}_0(t)$. These can be calculated according to the scheme [cf. (6.28) and (6.32) in [20]]

$$\begin{aligned}
 \langle \hat{x}_0(t_1) \rangle_{\hat{x}_0} & = \hat{\nu}(t_1), \\
 \langle \hat{x}_0(t_1)\hat{x}_0(t_2) \rangle_{\hat{x}_0} & = \hat{\nu}(t_1)\hat{\nu}(t_2) + \delta(t_1 - t_2)\hat{\nu}(t_1), \\
 & \cdots = \cdots, \quad (\text{B17})
 \end{aligned}$$

using the time-dependent rate $\hat{\nu}$ given by Eq. (B15). Equation (B14) together with the scheme in Eq. (B17) allows for the analytical calculation of the averages in Eq. (B6).

In the following, we prove Eq. (B14). The spike train within the interval $[t_s, t_e]$ can be regarded as the limit case of a temporally discretized function. More precisely, in terms of the equally-sized intervals I_j of width $\Delta T = (t_e - t_s)/N_d$, the spike counts $n_j = \int_{I_j} x_0(t') dt'$, and the indicator functions χ_j for the intervals I_j , we write,

$$x_0(t) = \lim_{N_d \rightarrow \infty} \frac{1}{\Delta T} \sum_{j=0}^{N_d-1} \chi_j(t) n_j. \quad (\text{B18})$$

For the averages of the spike count in the interval I_j , we write $\lambda_j = \langle n_j \rangle$. Using this, we transform the average Eq. (B6) as follows:

$$\begin{aligned}
 \langle x_0(t_1)x_0(t_2) \cdots e^k(t_s, t_e) \rangle_{x_0} & = \lim_{N_d \rightarrow \infty} \left\langle \left(\frac{1}{\Delta T} \sum_{i_1=0}^{N_d-1} \chi_{i_1}(t_1) n_{i_1} \right) \left(\frac{1}{\Delta T} \sum_{i_2=0}^{N_d-1} \chi_{i_2}(t_2) n_{i_2} \right) \cdots \prod_{j=0}^{N_d-1} (1 - F_1)^{n_j k} \right\rangle_{x_0} \\
 & = \lim_{N_d \rightarrow \infty} \left[\frac{1}{\Delta T} \sum_{i_1=0}^{N_d-1} \chi_{i_1}(t_1) \frac{1}{\Delta T} \sum_{i_2=0}^{N_d-1} \chi_{i_2}(t_2) \cdots \left\langle n_{i_1} n_{i_2} \cdots \prod_{j=0}^{N_d-1} (1 - F_1)^{n_j k} \right\rangle_{x_0} \right] \\
 & = \lim_{N_d \rightarrow \infty} \left[\frac{1}{\Delta T} \sum_{i_1=0}^{N_d-1} \chi_{i_1}(t_1) \frac{1}{\Delta T} \sum_{i_2=0}^{N_d-1} \chi_{i_2}(t_2) \cdots \prod_{j=0}^{N_d-1} \langle n_{i_1}^{\delta_{i_1 j}} n_{i_2}^{\delta_{i_2 j}} \cdots (1 - F_1)^{n_j k} \rangle_{x_0} \right]. \quad (\text{B19})
 \end{aligned}$$

For the last step, we used, that the n_j are statistically independent for different intervals I_j . The average appearing on the last line can be rewritten as

$$\begin{aligned}
 \langle n_j^m (1 - F_1)^{n_j k} \rangle_{x_0} &= \sum_{s=0}^{\infty} P(n_j = s) s^m (1 - F_1)^{s k} = \sum_{s=0}^{\infty} \frac{e^{-\lambda_j} \lambda_j^s}{s!} s^m (1 - F_1)^{s k} = \sum_{s=0}^{\infty} \frac{e^{-\lambda_j} [\lambda_j (1 - F_1)^k]^s}{s!} s^m \\
 &= e^{-\lambda_j [1 - (1 - F_1)^k]} \sum_{s=0}^{\infty} \frac{e^{-\lambda_j (1 - F_1)^k} [\lambda_j (1 - F_1)^k]^s}{s!} s^m = e^{-\lambda_j [1 - (1 - F_1)^k]} \sum_{s=0}^{\infty} \frac{e^{-\hat{\lambda}_j} \hat{\lambda}_j^s}{s!} s^m = e^{-\lambda_j [1 - (1 - F_1)^k]} \sum_{s=0}^{\infty} P(\hat{n}_j = s) s^m \\
 &= e^{-\lambda_j [1 - (1 - F_1)^k]} \langle \hat{n}_j^m \rangle_{\hat{x}_0},
 \end{aligned} \tag{B20}$$

where \hat{x}_0 is the aforementioned Poissonian δ -spike train with the rate as defined in Eq. (B15) and \hat{n}_j and $\hat{\lambda}_j = \langle \hat{n}_j \rangle$ are the corresponding spike count and its average, respectively. Inserting Eq. (B20) into Eq. (B19) finally yields Eq. (B14).

APPENDIX C: LINEARIZATION OF F DYNAMICS

We want to approximate the function $F(F_C)$ given in Eq. (7) by a linear relationship,

$$F_{\text{lin}}(F_C) = F_{0,\text{lin}} + m F_C, \tag{C1}$$

which obeys the linear facilitation dynamics,

$$\frac{dF_{\text{lin}}(t)}{dt} = \frac{F_{0,\text{lin}} - F_{\text{lin}}(t)}{\tau_F} + \Delta_{\text{lin}} x_0(t), \tag{C2}$$

where we used Eq. (6) and $\Delta_{\text{lin}} = m\Delta$. The integration of this linear equation yields Eq. (25). If one demands, that the mean square difference between the true and the approximated dynamics becomes minimal, i.e.,

$$f(m, F_{0,\text{lin}}) = \langle [F(F_C) - F_{\text{lin}}(F_C)]^2 \rangle \rightarrow \min, \tag{C3}$$

one arrives (considering, for simplicity, a constant rate) at the following expressions for the effective parameters of the linear dynamics:

$$\Delta_{\text{lin}} = \frac{\Delta(1 - F_0)^2}{\gamma^2} \left[1 - \frac{2\Delta}{3\gamma} + \frac{\Delta^2(1 + 3r\tau_F)}{2\gamma^2} \right], \tag{C4}$$

$$\begin{aligned}
 F_{0,\text{lin}} &= F_0 + \frac{(\Delta r \tau_F)^2 (1 - F_0)}{\gamma^2} + \frac{\Delta^2 r \tau_F (1 - F_0)^2}{6\gamma^3} \\
 &\quad \times \left[1 - \frac{\Delta(1 + 9r\tau_F)}{\gamma} \right],
 \end{aligned} \tag{C5}$$

and

$$\gamma = 1 - F_0 + \Delta r \tau_F. \tag{C6}$$

We sketch briefly how to derive these expressions. First, we need the first four central moments of F_C defined by

$$Q_k = \langle [F_C - \langle F_C \rangle]^k \rangle, \quad \langle F_C \rangle = \Delta r \tau_F. \tag{C7}$$

These can be obtained using the explicit solution of the linear F_C dynamics [cf. Eq. (6)]:

$$F_C(t) = \Delta \int_{-\infty}^t e^{-(t-t')/\tau_F} x_0(t') dt', \tag{C8}$$

and the Stratonovich scheme Eq. (B5) with constant rate $\nu(t) = r$. The central moments read

$$Q_1 = 0, \quad Q_2 = \frac{1}{2} \Delta^2 r \tau_F,$$

$$Q_3 = \frac{1}{3} \Delta^3 r \tau_F, \quad Q_4 = \frac{1}{4} \Delta^4 r \tau_F (3r\tau_F + 1).$$

The exact minimization of Eq. (C3) is not feasible. Instead, we use a Taylor expansion of $F(F_C)$ around the mean value of F_C ,

$$F(F_C) = \sum_{k=0}^{\infty} \frac{\partial^k F}{k!} f_C^k, \tag{C9}$$

where we define

$$f_C = F_C - \langle F_C \rangle \quad \text{and} \quad \partial^k F = \left. \frac{d^k F}{dF_C^k} \right|_{\langle F_C \rangle}.$$

Insertion of Eq. (C9) into Eq. (C3) yields

$$\begin{aligned}
 f(m, F_{0,\text{lin}}) &= \left\langle \left[\partial^0 F - F_{0,\text{lin}} - m \langle F_C \rangle + (\partial^1 F - m) f_C \right. \right. \\
 &\quad \left. \left. + \sum_{k=2}^{\infty} \frac{\partial^k F}{k!} f_C^k \right]^2 \right\rangle.
 \end{aligned} \tag{C10}$$

Standard calculus results in the following expressions for the minimum:

$$m_{\text{min}} = \partial^1 F + \frac{1}{Q_2} \sum_{k=2}^{\infty} \frac{\partial^k F}{k!} Q_{k+1}, \tag{C11}$$

and

$$F_{0,\text{lin},\text{min}} = \partial^0 F - \partial^1 F \langle F_C \rangle + \sum_{k=2}^{\infty} \frac{\partial^k F}{k!} \left(Q_k - \langle F_C \rangle \frac{Q_{k+1}}{Q_2} \right). \tag{C12}$$

Using terms up to $k=3$ results in the expressions in Eqs. (C4) and (C5).

APPENDIX D: APPROXIMATION OF D DYNAMICS

Assuming linear facilitation dynamics, we approximate the depression dynamics given by Eq. (8) linearizing the product $F_{\text{lin}}(t)D(t)$. With the deviations $\delta F = F_{\text{lin}} - \langle F_{\text{lin}} \rangle$ and $\delta D = D - \langle D \rangle$, we write

$$\begin{aligned} F_{\text{lin}}(t)D(t) &= \langle F_{\text{lin}} \rangle \langle D \rangle + \langle F_{\text{lin}} \rangle \delta D + \langle D \rangle \delta F + \delta F \delta D \\ &\approx \langle F_{\text{lin}} D \rangle + \langle F_{\text{lin}} \rangle \delta D + \langle D \rangle \delta F \end{aligned} \quad (\text{D1})$$

where in the last line, we approximated $\delta F \delta D \approx \langle \delta F \delta D \rangle$. This leads to the following approximate D_{lin} dynamics with $D_{\text{lin}} \approx D$:

$$\begin{aligned} \frac{dD_{\text{lin}}(t)}{dt} &= \frac{1 - D_{\text{lin}}(t)}{\tau_D} - [\langle D_{\text{lin}} \rangle F_{\text{lin}}(t^-) + \langle F_{\text{lin}} \rangle D_{\text{lin}}(t^-) \\ &\quad + \langle F_{\text{lin}} D_{\text{lin}} \rangle - 2\langle F_{\text{lin}} \rangle \langle D_{\text{lin}} \rangle] x_0(t). \end{aligned} \quad (\text{D2})$$

The explicit solution of this differential equation is Eq. (26).

We derive the averages $\langle F_{\text{lin}} \rangle$, $\langle D_{\text{lin}} \rangle$, and $\langle F_{\text{lin}} D_{\text{lin}} \rangle$ for a constant spiking rate $\nu(t) = r$ as explained in Sec. III B:

$$\langle F_{\text{lin}} \rangle = F_{0,\text{lin}} + \Delta_{\text{lin}} r \tau_F, \quad (\text{D3})$$

$$\langle D_{\text{lin}} \rangle = \frac{1 - r \tau_D [\langle F_{\text{lin}} D_{\text{lin}} \rangle - \langle F_{\text{lin}} \rangle \langle D_{\text{lin}} \rangle]}{1 + \langle F_{\text{lin}} \rangle r \tau_D}, \quad (\text{D4})$$

and

$$\begin{aligned} \langle F_{\text{lin}} D_{\text{lin}} \rangle &= \langle F_{\text{lin}} \rangle \langle D_{\text{lin}} \rangle - \frac{\Delta_{\text{lin}} r \tilde{\tau}}{1 + \langle F_{\text{lin}} \rangle r \tilde{\tau}} \left[\langle F_{\text{lin}} D_{\text{lin}} \rangle \right. \\ &\quad \left. + \frac{1}{2} \langle D_{\text{lin}} \rangle \Delta_{\text{lin}} r \tau_F \right], \end{aligned} \quad (\text{D5})$$

where

$$\tilde{\tau} = [\tau_F^{-1} + \tau_D^{-1}]^{-1}. \quad (\text{D6})$$

For $\langle D_{\text{lin}} \rangle$ and $\langle F_{\text{lin}} D_{\text{lin}} \rangle$, these are self-consistent equations. The solutions are

$$\langle D_{\text{lin}} \rangle = \frac{1 + r \tilde{\tau} (\langle F_{\text{lin}} \rangle + \Delta_{\text{lin}})}{\text{denom.}}, \quad (\text{D7})$$

and

$$\langle F_{\text{lin}} D_{\text{lin}} \rangle = \frac{\langle F_{\text{lin}} \rangle (1 + \langle F_{\text{lin}} \rangle r \tilde{\tau}) - \frac{1}{2} \Delta_{\text{lin}}^2 r \tilde{\tau} \tau_F}{\text{denom.}}, \quad (\text{D8})$$

where the denominator of both expressions is given by

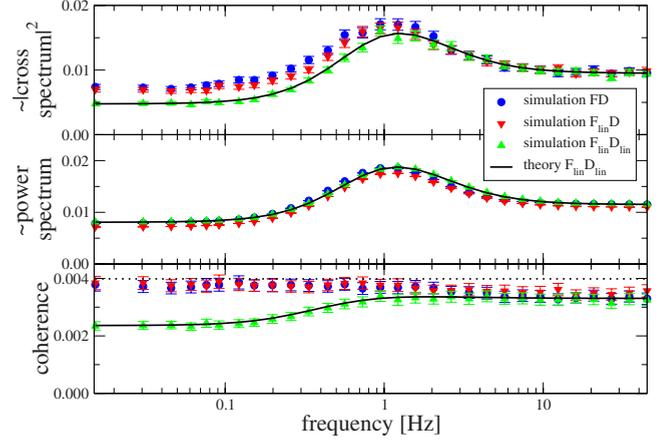


FIG. 17. (Color online) $|\hat{s}_{Rx}(f)|^2$, $\hat{s}_{xx}(f)$, and $C_{Rx}(f)$. Comparison between the results of approximated and original dynamics. “simulation FD”: simulation without any approximations made; “simulation $F_{\text{lin}}D$ ”: simulation, where the approximated facilitation dynamics has been used; “simulation $F_{\text{lin}}D_{\text{lin}}$ ”: simulation, where approximated facilitation and depression dynamics has been used; “theory $F_{\text{lin}}D_{\text{lin}}$ ”: theoretical results for the approximated dynamics. The dotted coherence curve shows the coherence for the case without any synaptic plasticity [i.e., $F(t) = D(t) = 1$]. The parameters are the same as in Fig. 10.

$$\begin{aligned} \text{denom.} &= (1 + \langle F_{\text{lin}} \rangle r \tilde{\tau})(1 + \langle F_{\text{lin}} \rangle r \tau_D) + \Delta_{\text{lin}} r \tilde{\tau} \\ &\quad \times \left(1 - \frac{1}{2} \Delta_{\text{lin}} r \tau_F r \tau_D \right). \end{aligned} \quad (\text{D9})$$

In the main text, we use the abbreviation $F_1 = \langle F_{\text{lin}} \rangle$.

Using the approximation described in this appendix, we derived the single synapse spectra and the coherence function. In most situations, this leads to good agreements with simulations of the original dynamics (cf. Figs. 9, 11, and 12), however, for large $r\tau_D$ and moderate $r\tau_F$, we find substantial deviations. For the example shown in Fig. 10, these deviations are clearly due to the D_{lin} approximation as can be seen in Fig. 17: simulations where only the facilitation dynamics is approximated coincide fairly well with the original dynamics. However, as soon as the depression dynamics is approximated, substantial deviations are observed, in particular in the coherence function for low frequencies.

- [1] L. F. Abbott and W. G. Regehr, *Nature (London)* **431**, 796 (2004).
 [2] S. A. Fisher, T. M. Fischer, and T. J. Carew, *Trends Neurosci.* **20**, 170 (1997).
 [3] R. S. Zucker and W. G. Regehr, *Annu. Rev. Physiol.* **64**, 355 (2002).
 [4] A. Mallart and A. R. Martin, *J. Physiol. (London)* **193**, 679 (1967).

- [5] J. A. Varela, K. Sen, J. Gibson, J. Fost, L. F. Abbott, and S. B. Nelson, *J. Neurosci.* **17**, 7926 (1997).
 [6] L. F. Abbott, J. A. Varela, K. Sen, and S. B. Nelson, *Science* **275**, 221 (1997).
 [7] H. Markram, Y. Wang, and M. Tsodyks, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5323 (1998).
 [8] J. S. Dittman, A. C. Kreitzer, and W. G. Regehr, *J. Neurosci.* **20**, 1374 (2000).

- [9] K. Magleby, *Synaptic Function* (Wiley, New York, 1987), pp. 21–56.
- [10] A. W. Liley and K. A. K. North, *J. Neurophysiol.* **16**, 509 (1953).
- [11] J. E. Lewis and L. Maler, *J. Neurophysiol.* **88**, 1695 (2002).
- [12] J. E. Lewis and L. Maler, *J. Neurophysiol.* **91**, 1064 (2004).
- [13] B. Lindner, D. Gangloff, A. Longtin, and J. E. Lewis, *J. Neurosci.* **29**, 2076 (2009).
- [14] M. Tsodyks and H. Markram, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 719 (1997).
- [15] J. E. Lisman, *Trends Neurosci.* **20**, 38 (1997).
- [16] G. Mongillo, O. Barak, and M. Tsodyks, *Science* **319**, 1543 (2008).
- [17] E. S. Fortune and G. J. Rose, *Trends Neurosci.* **24**, 381 (2001).
- [18] G. Fuhrmann, I. Segev, H. Markram, and M. Tsodyks, *J. Neurophysiol.* **87**, 140 (2002).
- [19] E. Salinas and T. J. Sejnowski, *Nat. Rev. Neurosci.* **2**, 539 (2001).
- [20] R. L. Stratonovich, *Topics in the Theory of Random Noise* (Gordon and Breach, New York, 1967), Vol. I.
- [21] W. Gerstner and W. M. Kistler, *Spiking Neuron Models* (Cambridge University Press, Cambridge, England, 2002).
- [22] A. N. Burkitt, *Biol. Cybern.* **85**, 247 (2001).
- [23] A. N. Burkitt, *Biol. Cybern.* **95**, 1 (2006).
- [24] M. J. Richardson and W. Gerstner, *Neural Comput.* **17**, 923 (2005).
- [25] L. Wolff and B. Lindner, *Phys. Rev. E* **77**, 041913 (2008); *Neural Comput.* **22**, 94 (2010).
- [26] W. Bialek and A. Zee, *J. Stat. Phys.* **59**, 103 (1990).
- [27] I. Goychuk, *Phys. Rev. E* **64**, 021909 (2001).
- [28] R. B. Stein, A. S. French, and A. V. Holden, *Biophys. J.* **12**, 295 (1972).
- [29] A. Borst and F. Theunissen, *Nat. Neurosci.* **2**, 947 (1999).
- [30] F. Gabbiani, *Network Comput. Neural Syst.* **7**, 61 (1996).
- [31] F. Rieke, R. de Ruyter Van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, MA, 1997).
- [32] F. Gabbiani and C. Koch, *Methods in Neuronal Modeling: From Synapses to Networks* (MIT Press, Cambridge, MA, 1998), pp. 360–313.
- [33] M. J. E. Richardson, O. Melamed, G. Silberberg, W. Gerstner, and H. Markram, *J. Comput. Neurosci.* **18**, 323 (2005).
- [34] E. Zohary, M. N. Shadlen, and W. T. Newsome, *Nature (London)* **370**, 140 (1994).
- [35] J. de la Rocha, R. Moreno, and N. Parga, *Neurocomputing* **58-60**, 313 (2004).
- [36] M. Merkel, Diploma thesis, Technical University Dresden, 2009.