


ARIANE ERNST<sup>1</sup>, CHRISTOF SCHÜTTE, STEPHAN SIGRIST,  
STEFANIE WINKELMANN<sup>2</sup>

## **Variance of filtered signals: characterization for linear reaction networks and application to neurotransmission dynamics**

---

<sup>1</sup>  0000-0002-7347-6670

<sup>2</sup>  0000-0002-0114-7819

Zuse Institute Berlin  
Takustr. 7  
14195 Berlin  
Germany

Telephone: +49 30 84185-0  
Telefax: +49 30 84185-125

E-mail: [bibliothek@zib.de](mailto:bibliothek@zib.de)  
URL: <http://www.zib.de>

ZIB-Report (Print) ISSN 1438-0064  
ZIB-Report (Internet) ISSN 2192-7782

# Variance of filtered signals: characterization for linear reaction networks and application to neurotransmission dynamics

Ariane Ernst<sup>\*a</sup>, Christof Schütte<sup>a,b</sup>, Stephan Sigrist<sup>c,d</sup>, and Stefanie Winkelmann<sup>a</sup>

<sup>a</sup>Zuse Institute Berlin, Berlin, Germany

<sup>b</sup>Freie Universität Berlin, Faculty of Mathematics and Computer Science, Berlin, Germany

<sup>c</sup>Freie Universität Berlin, Faculty of Biology, Chemistry, Pharmacy, Berlin, Germany

<sup>d</sup>NeuroCure Cluster of Excellence, Charité Universitätsmedizin Berlin, Berlin, Germany

June 28, 2021

## Abstract

Neurotransmission at chemical synapses relies on the calcium-induced fusion of synaptic vesicles with the presynaptic membrane. The distance to the calcium channels determines the release probability and thereby the postsynaptic signal. Suitable models of the process need to capture both the mean and the variance observed in electrophysiological measurements of the postsynaptic current. In this work, we propose a method to directly compute the exact first- and second-order moments for signals generated by a linear reaction network under convolution with an impulse response function, rendering computationally expensive numerical simulations of the underlying stochastic counting process obsolete. We show that the autocorrelation of the process is central for the calculation of the filtered signal's second-order moments, and derive a system of PDEs for the cross-correlation functions (including the autocorrelations) of linear reaction networks with time-dependent rates. Finally, we employ our method to efficiently compare different spatial coarse graining approaches for a specific model of synaptic vesicle fusion. Beyond the application to neurotransmission processes, the developed theory can be applied to any linear reaction system that produces a filtered stochastic signal.

**Key words:** linear reaction networks, cross-correlation, neurotransmission models, LTI filter, chemical master equation, vesicle fusion dynamics

---

<sup>\*</sup>ariane.ernst@zib.de

# 1 Introduction

Neurotransmission is the process by which a chemical or electrical signal is passed from a neuron to a target cell. The junction points at which it occurs are called *synapses*. In the presynaptic terminal of chemical synapses, the incoming action potential is translated into the release of neurotransmitters that rapidly diffuse over the synaptic cleft and elicit a postsynaptic response by binding to receptors in the target cell membrane [1]. Before being released into the cleft, the molecules are contained in small vesicles that can localize to the presynaptic cell membrane in the close perimeter of voltage-gated  $\text{Ca}^{2+}$ -channels (the *active zone*) and are functionally prepared – we say *primed* – for release [2, 3]. The arriving signal prompts the opening of the channels, leading to an influx of calcium-ions that bind to receptors on the vesicle surface. This triggers the fusion of vesicles with the presynaptic membrane, thereby releasing neurotransmitters into the synaptic cleft [4, 5].

Dynamical imaging of the neurotransmission process is not trivial due to the diffraction limit. Analyses of neuronal signaling therefore often combine microscopy techniques with electrophysiological measurements, which use electrodes in order to capture the postsynaptic electric signal [6–9]. Various approaches have been developed in order to describe the dynamic behaviour and classify synapses according to characteristic parameters [10–14]. Stochastic models play an especially important role, as the vesicle release process has been observed to be ‘reliably unreliable’: in most synapses, arrivals of action potentials provoke fusion in only 20% of cases and spontaneous release may occur in the absence of a stimulus [15–17]. The release probability of a single vesicle grows nonlinearly with cooperative binding of the  $\text{Ca}^{2+}$ -ions, where cooperativity values (i.e. binding steps) have been shown to range from 0 to 5 [18, 19].

Kobbersmed et al. [20] were first to introduce a stochastic vesicle fusion model which takes these aspects as well as active zone geometry into account. Based on a realistic spatial vesicle distribution, they optimized different models to electrophysiological data of repeated stimulation of the *drosophila* neuromuscular junction. In these models, the vesicle dynamics are formulated as a Markov jump process involving first-order reactions for the docking of vesicles to the membrane (priming), the binding of calcium ions to the vesicle’s surface and finally the fusion of the vesicle with the membrane. They showed that generally, accurate replication of short-term plasticity and signal variances could only be achieved in models where not only the release probability, but also the number of releasable vesicles at a given moment was  $\text{Ca}^{2+}$ -dependent. The best results were obtained with the so-called *unpriming model*, which additionally includes a reaction for the undocking of the vesicle from the membrane happening at a rate that depends on the calcium concentration. In order to reproduce the postsynaptic current measured in electrophysiological experiments, Kobbersmed et al. calculated the convolution of the process’ components with the response current of a single vesicle. The optimization of model parameters was done by solving deterministic reaction rate equations, and signal variances were determined afterwards based on Monte Carlo simulations of the underlying stochastic process.

In this paper, we take up the unpriming model of [20] and show that expensive and imprecise stochastic simulations may be circumvented due to the linearity of the reaction system, even if the output – namely the postsynaptic current resulting from the convolution – is given by a non-linear response function. We consider this postsynaptic current as the output signal of a linear time-invariant (LTI) filter and find that the second-order moment of such filtered signals can be expressed by means of the autocorrelation function of the underlying stochastic process. Motivated by this observation, we derive a characteristic partial differential equation (PDE) for the cross-correlation function (including the

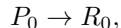
autocorrelation) of the process described by a general biochemical reaction network with time-dependent linear propensity functions. The time evolution of such a reaction system can be characterized by the chemical master equation, which has been solved analytically for certain subsets of linear reaction systems [21]. To our knowledge, the autocorrelation functions have so far only been researched for systems with time-independent linear rates (e.g. in [22] via means of noise power spectra analysis<sup>1</sup>). However, in this work we are concerned specifically with non-equilibrium reaction systems. We show that, in case of a linear reaction network, the resulting PDEs – together with the well-known ordinary differential equations (ODEs) for the first- and second-order moments of the process – form a closed system of equations that can be solved directly, avoiding stochastic simulations of the process. The solution then determines the first- and second-order moments of the output signal of interest. We apply this analytical approach to neurotransmission dynamics and examine a spatial extension of the vesicle fusion model, comparing expectations and variances of the total junction current depending on different levels of spatial coarse-graining. Furthermore, we demonstrate consistency of our results with experimental data.

Herein we are concerned with the *forward problem*: compute the total current based a  $\text{Ca}^{2+}$ -dependent vesicle fusion model and its spatial extension. The consequent next step is the associated *inverse problem*: Identify the key parameters of the fusion model and its spatial components based on experimental data on the total current. Finding a robust solution to such inverse problems is generally very difficult if the forward problem involves complex stochastic dynamics and a large number of involved parameters [24]. The inherent ill-posedness of the inverse problem requires advanced regularization in combination with highly precise computations of expectation values of the stochastic dynamics. The results to be presented herein provide a decisive advantage in this regard: we deliver a deterministic method for computing precise expectation values of the forward problem, thus allowing to avoid stochastic simulation and their intrinsic uncertainty and computational effort.

We start in Sec. 2 by introducing the model for stochastic vesicle fusion dynamics including the filtered output current as the quantity of central interest. This serves as a motivation for the investigations of Sec. 3, where we analyse filtered signals for more general chemical reaction systems and derive the characteristic PDEs for the cross-correlation function (including the autocorrelation) of the corresponding reaction jump process. The application to the spatially extended fusion model follows in Section 4.

## 2 Motivation: Neurotransmission dynamics

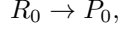
In [20], Kobbersmed et al. describe the neurotransmission process by means of the reaction model depicted in Figure 1, where vesicle priming, fusion and the binding of  $\text{Ca}^{2+}$ -ions are given by Markov jump processes. Each process describes the dynamics of a set of release sites located at a specific distance to the calcium channel. A release site can either be empty (state  $P_0$ ), or there is a vesicle attached to it, which itself can be in different states depending on the number of calcium-ions bound to the vesicle’s fusion sensor (states  $R_m$  with  $m = 0, \dots, 5$  referring to the number of bound calcium-ions). The process by which a vesicle attaches to an empty release site is called *priming* and it is described by the reaction




---

<sup>1</sup>The autocorrelation is related to the power spectral density via the Fourier transform [23].

which occurs at rate  $k_{\text{rep}} > 0$  for each release site and turns the empty release site into a vesicle-docked release site of status  $R_0$ .<sup>2</sup> The model also includes the backward reaction,

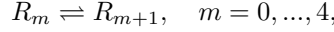


called *unpriming*, which detaches the vesicle from the release site again. This happens, for each release site, at rate  $u \cdot r$  for a constant  $u \geq 0$  and a rate  $r$  that depends on the calcium concentration  $[\text{Ca}^{2+}]$  according to

$$r = 1 - \frac{[\text{Ca}^{2+}]^n}{[\text{Ca}^{2+}]^n + (k_M)^n},$$

where  $k_M > 0$  is a Michaelis-Menten constant and  $n \in \mathbb{N}$  is the cooperativity exponent.

The  $\text{Ca}^{2+}$ -binding state of a vesicle changes according to the forward- and backward-reactions



with the corresponding calcium-dependent forward rates

$$\gamma_{\text{on}}^{(m)} = (5 - m) [\text{Ca}^{2+}] k_{\text{on}}$$

and the backward rates given by

$$\gamma_{\text{off}}^{(m)} = (m + 1) \beta^m k_{\text{off}}$$

for constants  $k_{\text{on}}, k_{\text{off}}, \beta \geq 0$  that refer to  $\text{Ca}^{2+}$ -binding/unbinding rate and cooperativity factor, respectively.

The central mechanism is the fusion of a vesicle with the membrane, which can happen from each status  $R_m$  and is modeled by the reactions



where  $F$  refers to the cumulative number of fused vesicles. It resets the release site to the empty state  $P_0$  and occurs with rate

$$\gamma_{\text{f}}^{(m)} = L^+ \alpha^m.$$

Here,  $L^+ > 0$  is the basal fusion rate constant and  $\alpha = \left(\frac{k_{\text{f}}}{L^+}\right)^{\frac{1}{5}}$  with  $k_{\text{f}}$  denoting the fusion rate of a fully activated vesicle.

The calcium-concentration  $[\text{Ca}^{2+}]$ , which determines some of the reaction rates above, depends on the distance  $d \geq 0$  of the considered release sites to the calcium channel and changes with time  $t \geq 0$ :

$$[\text{Ca}^{2+}] = [\text{Ca}^{2+}](t, d). \tag{2}$$

The spreading of  $[\text{Ca}^{2+}]$  within the active zone is given by a PDE<sup>3</sup>, taking into account the external calcium concentration (determining the inflow of calcium into the cell) and the times a stimulus is applied (determining when the calcium channel opens) [25].

<sup>2</sup>That is, the total propensity for such a reaction to occur is given by  $k_{\text{rep}} > 0$  times the number of empty release sites.

<sup>3</sup>The model can therefore be regarded as a hybrid approach between a jump process and a diffusion PDE [24]. Note that the approximation of the calcium ion population as a continuous concentration is essential for the linearity of the reaction propensities, which will become of great importance in the following sections.

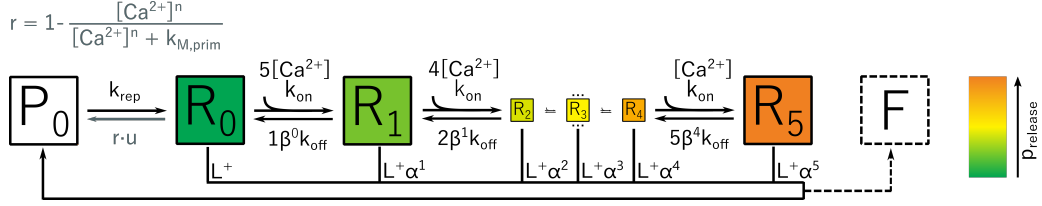


Figure 1: **Model of synaptic vesicle fusion.** Each release site can either be empty (state  $P_0$ ), or there is a vesicle docked to it which itself switches between calcium-binding states  $R_0, \dots, R_5$ . From each of these states, fusion can occur, leading back to an empty release site and an increase in fused vesicles  $F$ . A docked vesicle can detach again, thereby resetting the release site to the empty state (unpriming reaction, see gray arrow) [20].

### The stochastic process and output current

For each location on the membrane with fixed distance  $d$  to the calcium channel, the resulting reaction dynamics are modeled by a Markov jump process of the form

$$\mathbf{X}(t) = (P_0(t), R_0(t), \dots, R_5(t))$$

with  $P_0(t)$  denoting the number of empty release sites and  $R_m(t)$  giving the number of docked vesicles with binding state  $m$  at time  $t$ . In addition to the reaction jump process  $(\mathbf{X}(t))_{t \geq 0}$  describing the status of the release sites and docked vesicles, we introduce the stochastic process  $(F(t))_{t \geq 0}$  which counts the number of fusion events given by the reactions (1). That is,  $F(t)$  is a monotonically increasing process with initial value  $F(0) = 0$ , which rises by one each time that a fusion event occurs, see Eq. (11) in Sec. 3.2 for a more precise definition. Alternatively, the fused vesicle  $F$  may itself be treated as a chemical species, leading directly to an extended reaction jump process of the form

$$\hat{\mathbf{X}}(t) = (P_0(t), R_0(t), \dots, R_5(t), F(t)) \quad (3)$$

as considered in [20]. However, for our general investigations of Section 3, the special role of the component  $F(t)$  as a counting process plays an important role.

The waiting times between jumps in the system's state are exponentially distributed according to the rates on the arrows depicted in Figure 1. Several of these rates are implicitly time-dependent by their connection to the calcium-concentration (2). Note that the sum over the number of empty release sites  $P_0$  and vesicles in states  $R_0, \dots, R_5$  is constant over time.

The postsynaptic response current resulting from the dynamics of the process  $\mathbf{X}(t)$  can be calculated by convolving the functional derivative  $f(t)$  of the counting process  $F(t)$  with the *impulse response function*  $g : \mathbb{R}_+ \rightarrow \mathbb{R}$  (given by the mini current evoked by a single vesicle, see Figure 2), leading to the stochastic *output current*  $\mathcal{C}(t)$  given by

$$\mathcal{C}(t) := (f * g)(t) = \int_0^\infty f(s)g(t-s)ds, \quad (4)$$

see [26]. It is this output current  $\mathcal{C}(t)$  which is measured in electrophysiological experiments and for which first- and second-order moments are analyzed. In terms of signal processing theory, the output current  $\mathcal{C}(t)$  corresponds to the output signal of a linear time-invariant (LTI) filter with input signal  $f$  and impulse response  $g$  [27]. This invites to investigate such filtered signals for more general reaction networks, which will be done in the following section.

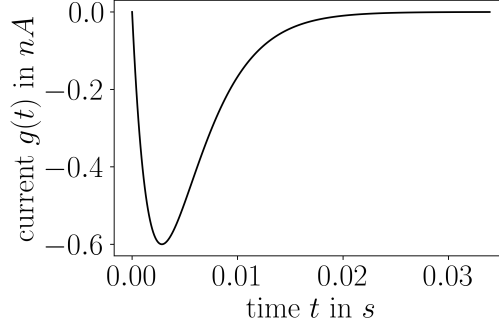


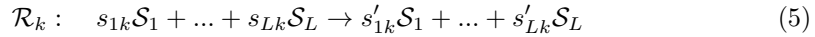
Figure 2: **Mini current.** Miniature excitatory postsynaptic current (mEPSC)  $g(t)$  evoked by a single vesicle.

### 3 Filtered signals for linear reaction systems

As output signals from LTI filters are of major interest when analyzing neurotransmission dynamics, we study such signals in the following, with the aim to derive characteristic equations for their first- and second-order moments. We start in Section 3.1 by introducing the general reaction network under consideration. For the case of linear propensity functions, we summarize the well-known moment equations of such networks, which will be employed in the subsequent analysis. In Section 3.2, we derive the moment equations for the filtered signal and show that the second-order moments are determined by the autocorrelation function of the underlying stochastic process. A characteristic system of PDEs for the cross-correlation function (including the autocorrelation) will be derived in Section 3.3, revealing that in case of a linear reaction network, this system is closed and can be used to directly determine the signal's moments, thereby avoiding expensive stochastic simulations of the underlying reaction process.

#### 3.1 The reaction network

The standard reaction network under consideration consists of  $L \in \mathbb{N}$  chemical species  $\mathcal{S}_1, \dots, \mathcal{S}_L$  which randomly interact through  $K \in \mathbb{N}$  reaction channels  $\mathcal{R}_1, \dots, \mathcal{R}_K$ . Each reaction is given by an equation of the form



with stoichiometric coefficients  $s_{lk}, s'_{lk} \in \mathbb{N}_0$ . The state of the system is defined by a vector  $\mathbf{x} = (x_1, \dots, x_L) \in \mathbb{N}_0^L$  with  $x_l$  denoting the number of particles of species  $\mathcal{S}_l$ . Each time that reaction  $\mathcal{R}_k$  takes place, this leads to a jump in the system state of the form  $\mathbf{x} \mapsto \mathbf{x} + \boldsymbol{\nu}^{(k)}$ , where  $\boldsymbol{\nu}^{(k)} = (\nu_1^{(k)}, \dots, \nu_L^{(k)}) \in \mathbb{Z}^L$  is the state-change vector of reaction  $\mathcal{R}_k$  with entries  $\nu_l^{(k)} := s'_{lk} - s_{lk}$ . The rate at which a jump induced by reaction  $\mathcal{R}_k$  occurs depends on the system's state  $\mathbf{x}$  and is given by the propensity function  $\lambda_k : \mathbb{N}_0^L \rightarrow \mathbb{R}_+$ . That is,  $\lambda_k(\mathbf{x})$  is the probability per unit of time for reaction  $\mathcal{R}_k$  to take place under the condition that the system is in state  $\mathbf{x}$ .

Given an initial state  $\mathbf{X}(0) = \mathbf{x}_0$ , the components of the reaction network define a Markov jump process  $(\mathbf{X}(t))_{t \geq 0}$ , with  $\mathbf{X}(t) \in \mathbb{N}_0^L$  being the system state at time  $t \geq 0$ . Letting  $P(\mathbf{x}, t)$  denote the probability to find this process  $\mathbf{X}(t)$  at time  $t$  in state  $\mathbf{x}$ , the



dynamics are fully characterized by the chemical master equation

$$\frac{d}{dt}P(\mathbf{x}, t) = \sum_{k=1}^K [\lambda_k(\mathbf{x} - \boldsymbol{\nu}_k)P(\mathbf{x} - \boldsymbol{\nu}_k) - \lambda_k(\mathbf{x})P(\mathbf{x})]. \quad (6)$$

### Moment equations for linear reaction networks

Let  $\mathcal{K}_0$  denote the index set of zero-order reactions with propensities of the form

$$\lambda_k(\mathbf{x}, t) = \gamma_k(t) \quad (7)$$

for a function  $\gamma : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , while  $\mathcal{K}_1$  is the index set of first-order reactions with propensity functions of the form

$$\lambda_k(\mathbf{x}, t) = \gamma_k(t)x_{l_k} \quad (8)$$

for some index  $l_k \in \{1, \dots, L\}$ . The system is called *linear reaction network* if there are only zero- and first-order reactions, while reactions of higher order are excluded, i.e., it holds that  $\mathcal{K}_0 \cup \mathcal{K}_1 = \{1, \dots, K\}$ .

For a linear reaction network with propensity functions given by (7) and (8), respectively, the first-order moments  $\mu_l(t) := \mathbb{E}[X_l(t)]$  evolve according to

$$\frac{d}{dt}\mu_l(t) = \sum_{k \in \mathcal{K}_0} \gamma_k(t)\nu_l^{(k)} + \sum_{k \in \mathcal{K}_1} \gamma_k(t)\mu_{l_k}(t)\nu_l^{(k)} \quad (9)$$

which results from the master equation (6) [28, 29]. For the (mixed) second-order moments  $c_{l,l'}(t) := \mathbb{E}[X_l(t)X_{l'}(t)]$  of a linear reaction network we have

$$\begin{aligned} \frac{d}{dt}c_{l,l'}(t) = & \sum_{k \in \mathcal{K}_0} \gamma_k(t) \left( \nu_l^{(k)}\mu_{l'}(t) + \nu_{l'}^{(k)}\mu_l(t) + \nu_l^{(k)}\nu_{l'}^{(k)} \right) \\ & + \sum_{k \in \mathcal{K}_1} \gamma_k(t) \left( \nu_l^{(k)}c_{l',l_k}(t) + \nu_{l'}^{(k)}c_{l,l_k}(t) + \nu_l^{(k)}\nu_{l'}^{(k)}\mu_{l_k}(t) \right). \end{aligned} \quad (10)$$

Together with (9) this forms a closed system of equations for the first- and second-order moments of a linear reaction system. Equations (9) and (10) will be needed in Section 3.3 in order to determine the autocorrelation function of a linear reaction network.

### 3.2 Moments of filtered signals

Given a stochastic reaction system with linear reaction propensities, we are interested in the signal produced by a set of *key reactions*  $\mathcal{R}_k$ ,  $k \in \mathcal{K}^* \subset \{1, \dots, K\}$ , for a given index subset  $\mathcal{K}^*$ . In the context of the vesicle fusion model from Sec. 2, these key reactions are given by the fusion reactions (1), while all intermediate steps such as docking and undocking of calcium ions are not counted as key reactions. Let  $0 \leq T_1 \leq T_2 \leq \dots$  denote the random time points where any of these key reactions  $\mathcal{R}_k$ ,  $k \in \mathcal{K}^*$ , takes place. Note that the random numbers  $T_j$  are *not* independent of each other, and they satisfy  $T_j < T_{j+1}$  almost surely for all  $j$ . Given these jump times, we define the counting process  $(F(t))_{t \geq 0}$  by

$$F(t) := \sum_{j=1}^{\infty} \mathbf{1}_+(t - T_j), \quad \mathbf{1}_+(t) := \begin{cases} 1, & t \geq 0, \\ 0, & t < 0, \end{cases} \quad (11)$$

which means that  $F(t)$  describes the cumulative number of occurrences of key reactions  $\mathcal{R}_k$ ,  $k \in \mathcal{K}^*$ , by time  $t$ . Per definition, the process  $F$  is piecewise-constant and monotonically increasing in time by jumps of size 1, and it satisfies  $F(0) = 0$  almost surely.

**Remark 1.** In accordance with our notes in Section 2, we can introduce an additional species  $\mathcal{S}_{L+1} = F$  into the reaction network defined in Section 3.1 and replace the stoichiometric equation (5) of each key reaction  $\mathcal{R}_k$ ,  $k \in \mathcal{K}^*$ , by

$$\mathcal{R}_k : s_{1k}\mathcal{S}_1 + \dots + s_{Lk}\mathcal{S}_L \rightarrow s'_{1k}\mathcal{S}_1 + \dots + s'_{Lk}\mathcal{S}_L + F, \quad (12)$$

while keeping the remaining reactions unaltered. This means that the system state  $\mathbf{x} = (x_1, \dots, x_L)$  is extended by a component  $x_{L+1}$  referring to the number of particles of this additional species  $F$  of cumulative key reaction events. We obtain an extended Markov jump process of the form  $\hat{\mathbf{X}}(t) = (X_1(t), \dots, X_L(t), F(t))$ , in analogy to the one given in (3). Setting  $F(0) = 0$ , we get the counting process defined in (11). In the context of the vesicle fusion model of Sec. 2,  $F$  refers to the species of fused vesicles, and Eq. (12) is a generalization of the fusion reaction (1).

The functional derivative  $f$  of the counting process  $F$  is given by a sum of dirac delta functions shifted by the time points  $T_j$ , i.e., we have  $F(t) = \int_{-\infty}^t f(s)ds$  with

$$f(t) := \sum_{j=1}^{\infty} \delta(t - T_j),$$

where  $\delta$  is the standard Dirac delta. We consider the process  $f$  as the *input signal* of an LTI filter, which determines the *output signal*  $\mathcal{C}(t) = (f * g)(t)$  via convolution with a given impulse response function  $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ , as defined in (4). In the following, we investigate the first- and second-order moments  $\mathbb{E}[\mathcal{C}(t)]$  and  $\mathbb{E}[\mathcal{C}^2(t)]$  of this central quantity  $\mathcal{C}(t)$ .

### First- and second-order moments of the output signal $\mathcal{C}(t)$

Let  $\mu_F(t) := \mathbb{E}[F(t)]$  denote the expectation of the counting process  $F(t)$ . For the first-order moment  $\mu_{\mathcal{C}}(t) := \mathbb{E}[\mathcal{C}(t)]$  of the output signal  $\mathcal{C}(t)$  we have

$$\begin{aligned} \mu_{\mathcal{C}}(t) &= \mathbb{E} \left[ \int f(t-s)g(s)ds \right] \\ &= \mathbb{E} \left[ \int \frac{d}{dt} F(t-s)g(s)ds \right] \\ &\stackrel{(*)}{=} \int \mathbb{E} \left[ \frac{d}{dt} F(t-s) \right] g(s)ds \\ &\stackrel{(**)}{=} \int \frac{d}{dt} \mathbb{E}[F(t-s)] g(s)ds \\ &= \int \mu'_F(t-s)g(s)ds \end{aligned}$$

for  $\mu'_F := \frac{d}{dt}\mu_F$ , where  $(*)$  follows from Fubini's theorem and  $(**)$  from the Leibniz rule. This means that we can express the first-order moment  $\mu_{\mathcal{C}}(t)$  of the output signal  $\mathcal{C}(t)$  by means of  $\mu_F(t)$  according to

$$\mu_{\mathcal{C}} = \mu'_F * g. \quad (13)$$

As we can understand  $F(t)$  as the component of an extended reaction jump process with additional species  $F$  (see Remark 1), its expectation  $\mu_F$  is determined by the standard moment-ODE (9).

In order to determine also the second-order moments  $\mathbb{E}[\mathcal{C}^2(t)]$  of the output signal, we define, for any function  $a : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ , the twofold convolution

$$(a * g * g)(t) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(s, s') g(t-s) g(t-s') ds ds'.$$

With this definition, we can write

$$\begin{aligned} \mathbb{E}[\mathcal{C}^2(t)] &= \mathbb{E} \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s) f(s') g(t-s) g(t-s') ds ds' \right] \\ &= (a_f * g * g)(t) \end{aligned}$$

for  $a_f(s, s') := \mathbb{E}[f(s)f(s')]$ .<sup>4</sup> Noting that  $a_f(s, s') = \frac{\partial}{\partial s} \frac{\partial}{\partial s'} a_F(s, s')$  for  $a_F(s, s') := \mathbb{E}[F(s)F(s')]$ , we finally get

$$\mathbb{E}[\mathcal{C}^2(t)] = \left( \left( \frac{\partial}{\partial s} \frac{\partial}{\partial s'} a_F \right) * g * g \right) (t). \quad (14)$$

That is, the second-order moments of the output signal  $\mathcal{C}(t)$  can be expressed by means of the autocorrelation function  $a_F$  of  $F(t)$ . In the following subsection, we will derive a characteristic equation for the autocorrelation function of a general linear reaction network.

### 3.3 Characteristic equation for cross-correlation functions

While the evolution equations for the first- and second-order moments of the reaction jump process  $\mathbf{X}(t)$  are well-known, see Equations (9) and (10), the corresponding cross-correlations

$$a_{l,l'}(t, s) := \mathbb{E}[X_l(t)X_{l'}(s)], \quad l, l' \in \{1, \dots, L\}, \quad t \neq s,$$

have not explicitly been investigated in the literature. However, as we have seen before, these cross-correlation functions are essential to analytically determine the second-order moments of the filtered signal  $\mathcal{C}(t)$ . The following theorem delivers a characteristic equation for the cross-correlation functions of a linear reaction network.

**Theorem 1.** *For a linear reaction network the cross-correlation functions satisfy*

$$\frac{\partial}{\partial t} a_{l,l'}(t, s) = \sum_{k \in \mathcal{K}_0} \gamma_k(t) \nu_l^{(k)} \mu_{l'}(s) + \sum_{k \in \mathcal{K}_1} \gamma_k(t) \nu_l^{(k)} a_{l_k, l'}(t, s) \quad (15)$$

for  $t \geq s \geq 0$ , with initial value  $a_{l,l'}(s, s) = c_{l,l'}(s) = \mathbb{E}[X_l(s)X_{l'}(s)]$ .

The proof can be found in the Appendix, page 23.

This means that, together with (9) and (10), we have a closed system of equations for the first-order moments  $\mu_l(t)$ , the covariances  $\text{cov}(X_l(t), X_{l'}(t)) = c_{l,l'}(t) - \mu_l(t)\mu_{l'}(t)$  and the cross-correlations  $a_{l,l'}(t, s)$  of the reaction jump process  $\mathbf{X}(t)$  (or, likewise, for the extended process  $\tilde{\mathbf{X}}(t)$ ) for a linear reaction network.

<sup>4</sup>A remark on the first- and second-order moments of the input signal  $f$  is given in the Appendix 6.2.

**Remark 2** (Non-linear reaction systems). *For second-order reactions with propensities of the form  $\lambda_k(\mathbf{x}, t) = \gamma(t)x_l x_{l'}$  the term on the right-hand side of (15) includes third-order moments of the form  $\mathbb{E}[X_l(t)X_{l'}(t)X_{l''}(s)]$ , so the overall system is not closed anymore.*

In the following, we will illustrate our result by an example in which we will also consider the normalized autocorrelation coefficient given by

$$\kappa_{l,l}(t, s) := \frac{a_{l,l}(t, s) - \mu_l(t)\mu_l(s)}{\sigma_l(t)\sigma_l(s)} \quad (16)$$

for  $\sigma_l(t) := \sqrt{\mathbb{V}[X_l(t)]} = \sqrt{\mathbb{E}[X_l^2(t)] - \mu_l^2(t)} = \sqrt{c_{l,l}(t) - \mu_l^2(t)}$ .

**Example 1** (Reduced linear system). *Consider the simple linear reaction network with two species  $\mathcal{S}_1, \mathcal{S}_2$  and two reactions*



The key reaction we are interested in is given by  $\mathcal{R}_2$ , i.e., we have  $\mathcal{K}^* = \{2\}$ , and  $F(t)$  is counting the occurrences of this second reaction. In order to apply our results, we use the model extension introduced in Remark 1, meaning that we replace the second reaction by  $\mathcal{R}_2 : \mathcal{S}_2 \rightarrow \mathcal{S}_1 + F$  and consider the extended process  $\hat{\mathbf{X}}(t) = (X_1(t), X_2(t), F(t))$  with the additional species  $\mathcal{S}_3 = F$ . Let  $\gamma_1(t)$  and  $\gamma_2(t)$  denote the reaction rates for  $\mathcal{R}_1$  and  $\mathcal{R}_2$ . We compare two scenarios in order to demonstrate the behaviour of the system for constant as well as for time-dependent rates. In both cases, we set the initial state of the extended system to  $\hat{\mathbf{X}}(0) = (10, 0, 0)$  and the rate of the second reaction  $\mathcal{R}_2$  to  $\gamma_2(t) = 5$  for all  $t$ . The scenarios differ in how we set the rate for the first reaction  $\mathcal{R}_1$ :

$$\begin{aligned} \gamma_1(t) &= 2 \quad (\text{Scenario 1}), \\ \gamma_1(t) &= 0.1 + 2 \cdot \frac{1}{\sqrt{2\pi \cdot 0.05^2}} e^{-\frac{(t-0.5)^2}{2 \cdot 0.05^2}} \quad (\text{Scenario 2}). \end{aligned}$$

Samples of the stochastic process  $\hat{\mathbf{X}}(t)$  are created via a modified version of Gillespie's direct method with a maximal timestep in order to update the time-dependent rates [30]. We display the average system behaviour and standard deviation for the two scenarios, calculated from the moment-ODEs (9) and (10) and via Monte-Carlo simulations of the process, in Figure 3a and 3b. As one would expect, in the case of constant rates (Scenario 1), the means of species  $\mathcal{S}_1$  and  $\mathcal{S}_2$  quickly equilibrate, leading to an approximately linear increase in  $\mu_F(t) = \mathbb{E}[F(t)]$  after a short time delay. While the standard deviation of both  $X_1(t)$  and  $X_2(t)$  converges to a constant value with equilibration, it increases over time for  $F(t)$ . In the second scenario, similar behavior arises far from the peak of  $\gamma_1$  at time  $t = 0.5$ , but the temporary increase in  $\gamma_1(t)$  during the peak leads to nonlinear changes in the average system state as well as the standard deviations.

Using Theorem 1, we compute the autocorrelation function  $a_{l,l}(s, t) = \mathbb{E}[X_l(t)X_l(s)]$  for each species of the extended system between all time points  $s$  and  $t$  by solving the PDE (15). The autocorrelation function  $a_{F,F}$  is displayed in Figure 4 for both scenarios. In Scenario 1 (Fig. 4a), the linear increase of  $F$  after the equilibration period causes  $a_{F,F}(s, t)$  to be dependent on the product  $s \cdot t$  of the time points. Denoting the slope of the mean  $\mu_F(t)$  in Figure 3a by  $m$ , one can approximate the autocorrelation as  $a_{F,F}(s, t) \approx m^2 \cdot s \cdot t$ . In the second scenario, due to the fact that the counting process  $F(t)$  has low values for all time points before the passage of the peak in  $\gamma_1(t)$ , the autocorrelation is small whenever  $t$  or  $s$  is in that domain. The temporary increase in  $\gamma_1(t)$  causes an

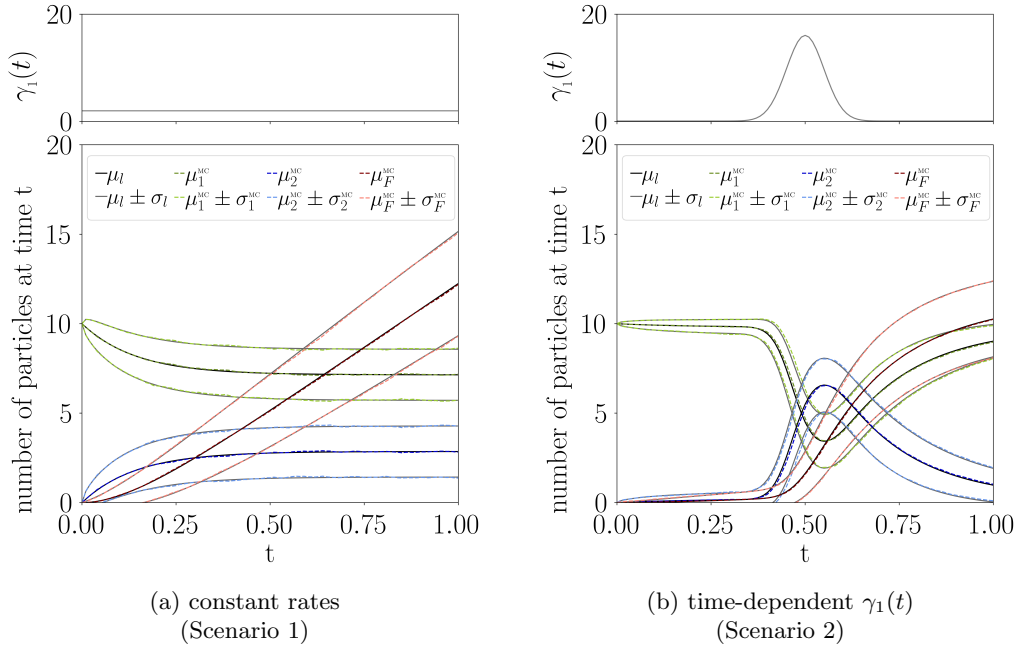


Figure 3: **Average system behavior.** Mean  $\mu_l(t) = \mathbb{E}[X_l(t)]$  and standard deviation  $\sigma_l(t) = \sqrt{\mathbb{V}[X_l(t)]}$  of the number of particles of species  $\mathcal{S}_l$  from moment-ODEs (solid) and Monte-Carlo simulation ( $\mu_l^{\text{MC}}(t), \sigma_l^{\text{MC}}(t)$ , dashed) for the reduced system of Example 1 and (a) constant reaction rates (Scenario 1) and (b) time-dependent  $\gamma_1(t)$  (Scenario 2).  $10^3$  MC-steps for each scenario.

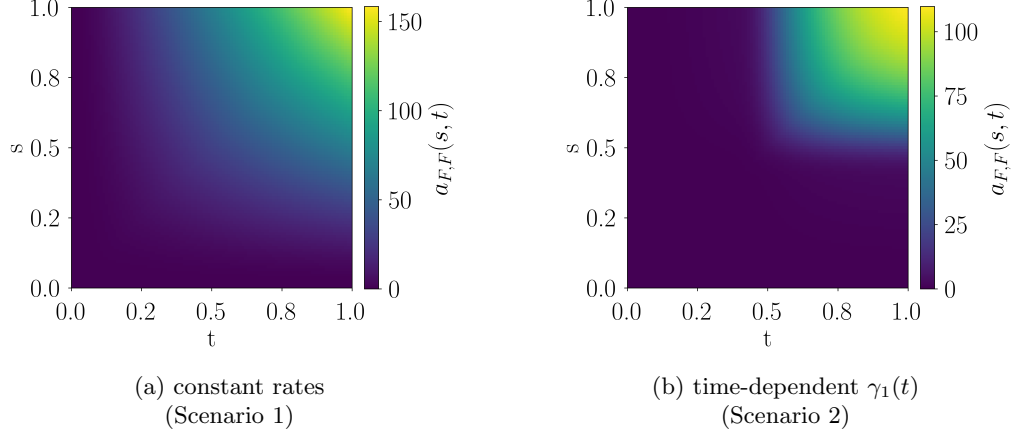


Figure 4: **Autocorrelation for species  $F$ .** Autocorrelation function  $a_{F,F}(t, s)$  as solution of the PDE (15) for (a) constant reaction rates (Scenario 1) and (b) time-dependent  $\gamma_1(t)$  (Scenario 2) of Example 1.

asymptotic, approximately logistic growth in the number of key reaction events  $F$  (Figure 3b), which is reflected in similar autocorrelation values when both  $s$  and  $t$  are close to 1.

Since  $a_{i,i}(s, t)$  is in absolute units, its scale varies according to the respective number of particles. In order to show the stochastic dependency independent of the system scale and to make a direct comparison between different scenarios feasible, we also compute the autocorrelation coefficient  $\kappa_{i,i}(s, t)$  from Eq. (16), shown for species  $F$  in Figure 5. Note especially the low autocorrelation coefficient between timepoints before and after the peak for the second scenario in Figure 5b.

Circling back to the initial motivation of our calculations, we now compute the first- and second-order moment of the output signal generated by the second reaction,

$$\mathcal{C}(t) = (f * g)(t), \quad (17)$$

where  $f(t)$  is the functional derivative of the stochastic process  $F(t)$ , and the impulse response function  $g$  is defined as the simple step function

$$g(t) := \begin{cases} -1, & \text{if } t \in [0, 0.2) \\ 0, & \text{otherwise.} \end{cases}$$

For both scenarios we compare the deterministic solution that results from solving Eq. (13) and Eq. (14) and using Thm. (1) to Monte-Carlo estimations, observing that the results agree very well in both cases, see Figure 6. In order to keep the relative error in the single-digit range, more than  $10^4$  MC-steps were required for each scenario. Solving the deterministic equations instead reduced the computational runtime by a factor of 200. This example thus illustrates that, in case of a linear reaction network, numerically expensive MC-simulations for estimating first- and second-order moments of stochastic filtered signals can be avoided by using our result from Thm. (1).

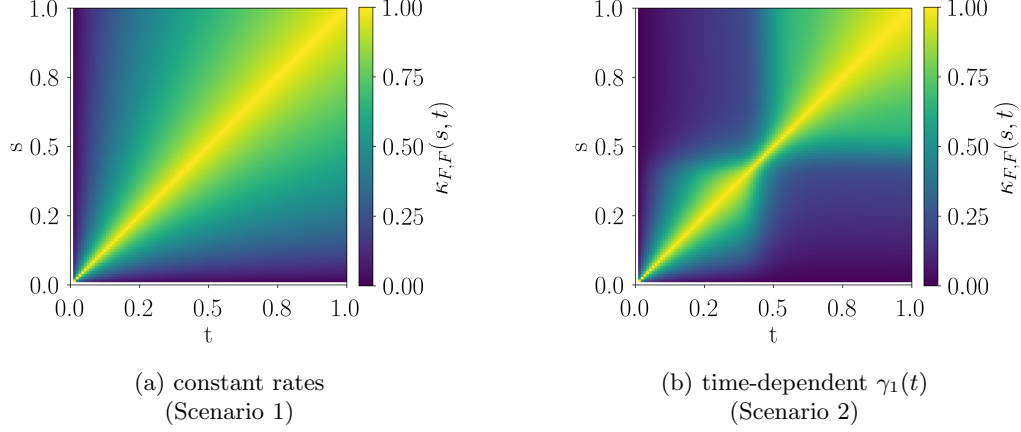


Figure 5: **Autocorrelation coefficient for species  $F$ .** Normalized autocorrelation coefficient  $\kappa_{F,F}(t, s)$  given in Eq. (16) for (a) constant reaction rates (Scenario 1) and (b) time-dependent  $\gamma_1(t)$  (Scenario 2) of Example 1.

## 4 Application to vesicle fusion model

In the neurotransmission process as modeled by Kobbersmed et al., the measured post-synaptic current is actually the sum of the currents resulting from multiple release sites placed along the active zone. In the following, we therefore model the overall neurotransmission process by combining the dynamics at different spatial locations. In Section 4.1 we introduce the spatially extended vesicle fusion model. Subsequently, we analyze the dynamics for a spatially fully resolved and a spatially coarse-grained setting in Sections 4.2 and 4.3, respectively, thereby using the results from Sec. 3 in order to determine the moments of the total junction current. We close our investigations by an analysis of the current's peak size in Section 4.4.

### 4.1 Spatial extension of the vesicle fusion model

The calcium concentration a release site will be exposed to during an action potential varies depending on the distance to the calcium channel. This determines the number of primed vesicles and the overall vesicle fusion dynamics at the different release sites. We define a set of distances  $d_i$ ,  $i \in \{1, \dots, N_d\}$ , at which the release sites can be located. (Note that there can be several release sites located at the same distance.) For each distance  $d_i$  we consider a process

$$\hat{\mathbf{X}}^{(i)}(t) = \left( P_0^{(i)}(t), R_0^{(i)}(t), \dots, R_5^{(i)}(t), F^{(i)}(t) \right) \quad (18)$$

as the one defined in (3), with  $F^{(i)}(t)$  referring to the number of fused vesicles (or, equivalently, the cumulative number of fusion events) at this location at time  $t$ . The calcium concentration  $[\text{Ca}^{2+}](t, d_i)$ , which determines the jump rates for the process  $\hat{\mathbf{X}}^{(i)}$  according to the rules of the vesicle fusion model described in Sec. 2 is a function of time  $t$  and distance  $d_i$  to the calcium channel, see Eq. (2).

We denote by  $\mathcal{C}_i(t)$  the output current at location with distance  $d_i$ , i.e.,

$$\mathcal{C}_i(t) = (f^{(i)} * g)(t),$$

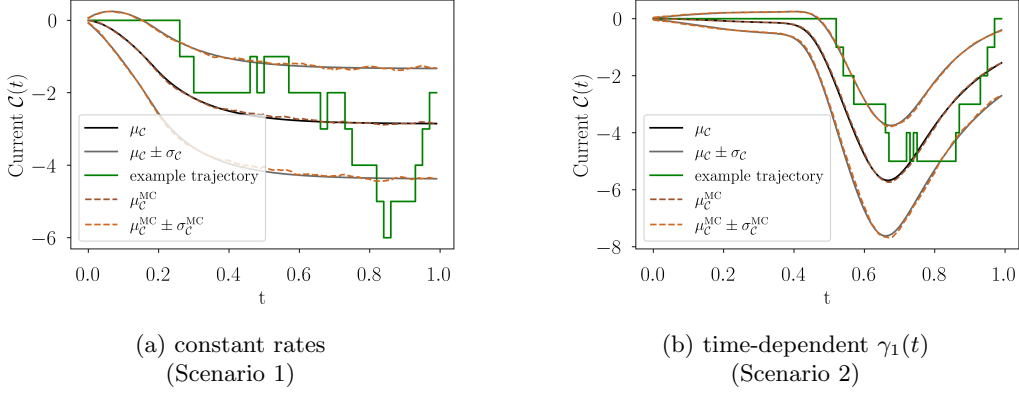


Figure 6: **Output current for reduced linear system.** Mean  $\mu_C(t) = \mathbb{E}[C(t)]$  and standard deviation  $\sigma_C(t) = \sqrt{\mathbb{V}[C(t)]}$  of the current  $C(t)$  given in (17) computed by solving the PDE (15) and using Eq. (14) (solid black) in comparison to Monte-Carlo estimations (dashed brown) for (a) constant rates (Scenario 1) and (b) time-dependent  $\gamma_1(t)$  (Scenario 2) given in Example 1. For both scenarios, an exemplary trajectory (green) illustrates the stochastic temporal evolution of the current  $C(t)$ . The MC-estimations result from  $10^3$  MC-steps for each scenario.

see (4), with  $f^{(i)}(t)$  denoting the input signal at distance  $d_i$ , i.e.,  $f^{(i)}$  is given by the functional derivative of the counting process  $F^{(i)}$ . The impulse response  $g(t)$  is the miniature excitatory postsynaptic current depicted in Figure 2. The *total junction current* evoked by all release sites together is then given by adding up all distance-specific output currents [26]:

$$C_{total}(t) := \sum_{i=1}^{N_d} C_i(t) = \int_0^t \sum_i f^{(i)}(s) g(t-s) ds.$$

Since the  $C_i(t)$  are stochastically independent of each other, we have

$$\mu_{C_{total}}(t) = \mathbb{E}[C_{total}(t)] = \sum_{i=1}^{N_d} \mathbb{E}[C_i(t)] \quad (19)$$

$$\sigma_{C_{total}}^2(t) = \mathbb{V}[C_{total}(t)] = \sum_{i=1}^{N_d} \mathbb{V}[C_i(t)] = \sum_{i=1}^{N_d} (\mathbb{E}[C_i^2(t)] - \mathbb{E}[C_i(t)]^2) \quad (20)$$

for the first and second cumulant of  $C_{total}(t)$ .

We consider the case of two presynaptic stimuli in the form of two spikes in the calcium concentration over a certain time interval  $[0, T]$  for a suitable  $T > 0$ . We want to analyse the behavior of a system of 180 release sites for different external calcium concentrations  $[\text{Ca}^{2+}]_{\text{ext}}$ , which regulate the amount of calcium inflow while the channel is open and thereby determine the function  $[\text{Ca}^{2+}](t, d_i)$ . For the analysis in the following sections, this calcium flow will be calculated in advance by numerically solving the corresponding PDE [25]. The stochastic processes (18) will be simulated using the temporal Gillespie algorithm [30].



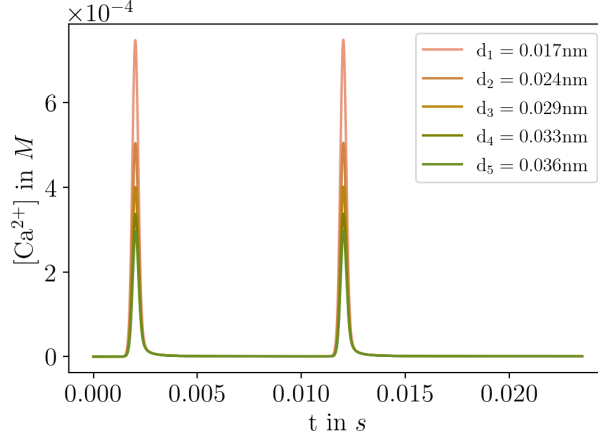


Figure 7: **Calcium concentration at example locations.** Calcium concentration  $[Ca^{2+}](t, d_i)$  over time  $t$  at five different distances  $d_i$  from the calcium channel for two stimuli at  $t_1 = 0.002s$  and  $t_2 = 0.012s$  with an external calcium concentration  $[Ca^{2+}]_{ext} = 0.75mM$ .

## 4.2 One release site at each location

Let us first consider the case discussed in [20], where each of the 180 release sites is placed at a different distance  $d_i$  from the calcium channel, so that the number of locations is  $N_d = 180$ , and a two-pulse stimulus is applied to the system.

### Distribution of distances

According to [20], the probability density function (pdf) of the distribution of vesicles (at the *drosophila* neuromuscular junction) is the integrated Rayleigh distribution

$$\rho(d) = \frac{\sqrt{2}}{\sqrt{\pi} \cdot \varsigma^3} \cdot d^2 \cdot e^{-\frac{d^2}{2\varsigma^2}} \quad (21)$$

where  $d$  is the distance from the calcium channel and  $\varsigma = 76.51nm$ . For a depiction of  $\rho(d)$  see the red curve in Figure 10 further below. Using inverse transform sampling, the release site distances can then be generated from the cumulative distribution function of  $\rho(d)$ .

The application of two consecutive stimuli at  $t_1 = 0.002s$  and  $t_2 = 0.012s$  then results in unique calcium concentration evolutions for each location, which is shown in Figure 7 for the first five  $d_i$ . Since many of the rates in the model are calcium-dependent (see Figure 1), this means that will also vary according to the release site's location.

### Autocorrelation at different locations

Solving the PDE (15) for all locations and using (16), we can now compute the autocorrelation functions  $a_F^{(i)}(t, s) := a_{F^{(i)}, F^{(i)}}(t, s)$  and the normalized autocorrelation coefficients  $\kappa_F^{(i)}(t, s) := \kappa_{F^{(i)}, F^{(i)}}(t, s)$  for each  $d_i$ . Figure 8 displays them for an example location  $d_1 = 0.118nm$ . The sharpness of the calcium peaks causes the areas with similar values of  $a_F^{(1)}(t, s)$  or  $\kappa_F^{(1)}(t, s)$  to have very sharp edges compared to the case with a broader peak

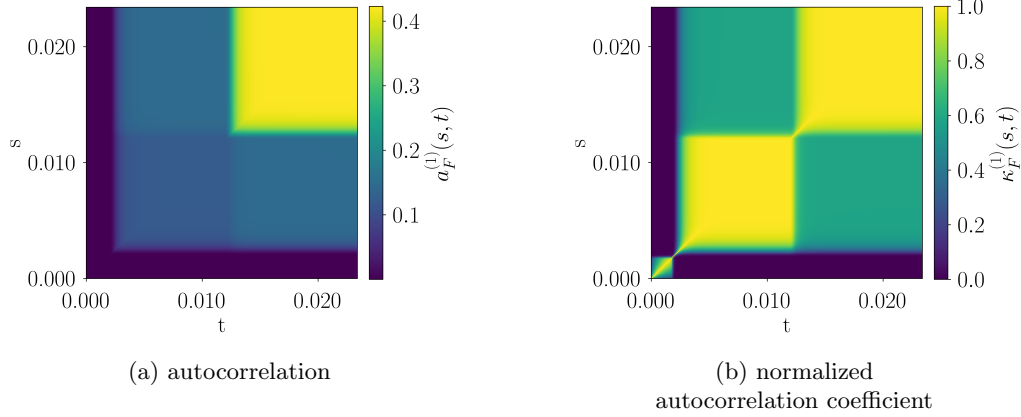


Figure 8: **Autocorrelation and normalized autocorrelation coefficient for an example location.** (a) Autocorrelation  $a_F^{(1)}(t, s)$  and (b) normalized autocorrelation coefficient  $\kappa_F^{(1)}(t, s)$  of species  $F^{(1)}$  (= number of fused vesicles at distance  $d_1$ ) for two stimuli as shown in Figure 7 from solution of the PDE (15). The distance for this location was  $d_1 = 0.118nm$ .

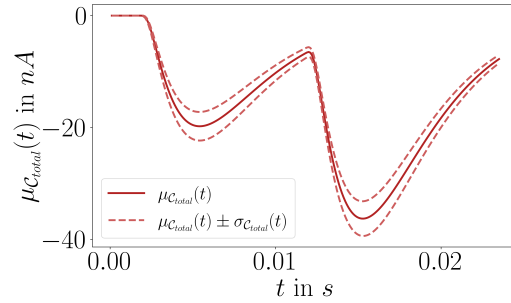


Figure 9: **Total junction current for two stimuli.** Mean  $\mu_{C_{total}}(t) = \mathbb{E}[C_{total}(t)]$  and standard deviation  $\sigma_{C_{total}}(t) = \sqrt{\mathbb{V}[C_{total}(t)]}$  of the total junction current for 180 locations ( $N_d = 180$ ) as in [20]. Stimuli were as shown in Figure 7.

in the reduced system (Scenario 2 of Example 1, Figures 4b and 5b). However, interestingly, the overall behaviour of both  $a_F^{(1)}(t, s)$  and  $\kappa_F^{(1)}(t, s)$  is consistent with the reduced linear system of Example 1, despite the larger number of intermediate states (additional species) and more complex time dependence in the vesicle fusion model, which suggests to reduce the vesicle fusion model skipping intermediate reaction steps.

### Total current

We can now compute first- and second-order moments  $\mathbb{E}[C_i(t)]$  and  $\mathbb{E}[C_i^2(t)]$  for the individual distances and, by Eqs. (19) and (20), the first two cumulants for the total junction current,  $\mu_{C_{total}} = \mathbb{E}[C_{total}(t)]$  and  $\sigma_{C_{total}}^2 = \mathbb{V}[C_{total}(t)]$ . The total junction current and standard deviation resulting from 180 vesicles that are distributed according to Eq. (21) is shown in Figure 9.

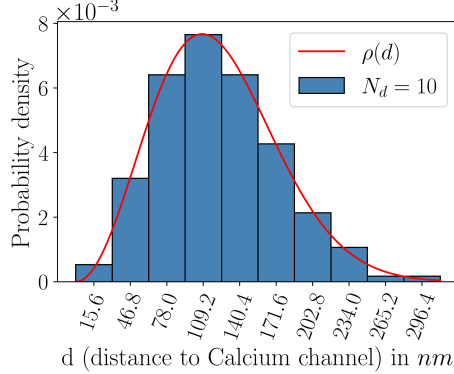


Figure 10: **Coarse grained probability density and integrated Rayleigh distribution**  $\rho(d)$ . Probability densities for the subintervals were computed after grouping Rayleigh-distributed 180 release sites into  $N_d = 10$  evenly spaced bins. The integrated Rayleigh distribution  $\rho(d)$  (red) is given in Equation (21).

### 4.3 Spatial coarse graining

The approach described in Section 4.2 requires the solution of 180 systems of differential equations, each consisting of 7 equations for the first-order moments (one for each species),  $7 \times 7$  equations for the second-order moments and  $n_{\Delta t}$  equations for the autocorrelation, where  $n_{\Delta t}$  is the number of time steps for the chosen time discretization  $\Delta t$  (so  $n_{\Delta t} \cdot \Delta t = t_{max}$ ). For example, for  $n_{\Delta t} = 100$  this equates to  $(7 + 7 \times 7 + 100) = 156$  equations for each distance and approximately  $180 \times 156 \approx 2.8 \cdot 10^3$  equations in total. Thus, it is favorable to choose the time discretization  $\Delta t$  as large as possible without significant loss of accuracy to keep the number of equations per system low. In order to further reduce the resulting computational cost, we introduce a spatial coarse graining (CG) approach for the locations: the maximum distance to the calcium channel is divided into  $N_d$  equally spaced subintervals (bins) with midpoints  $d_i$ . The number of release sites within each bin is determined from the integrated Rayleigh distribution  $\rho$  (Eq. (21)). An example plot of the probability densities for  $N_d = 10$  together with the integrated Rayleigh distribution is depicted in Figure 10.

The computation is then very similar to that in Section 4.2, but with multiple release sites at each of the new equidistant locations  $d_i$ . This is easily implemented by scaling the initial conditions at each  $d_i$  according to the number of release sites that were grouped into interval  $i$ . We compare the total current and its variance for different values of  $N_d$  with the results without CG in Figure 11. The first-order moment  $\mu_{C_{total}}(t)$  can be approximated sufficiently well even with only  $N_d = 3$ , while the variance computation requires at least  $N_d = 10$  subintervals to achieve a good agreement. Thus, the number of systems of differential equations can be brought down to 10 instead of 180. This also means that the system can be modelled accurately with a CG approach in both the stochastic as well as the deterministic case.

With 1000 MC-steps for the system with  $N_d = 180$ , we still observe significant statistical uncertainty in the variance. Thus, when interested in high accuracy (e.g., when solving stochastic inverse problems) our proposed deterministic technique is advantageous regarding accuracy and numerical effort. Our method is especially preferable for

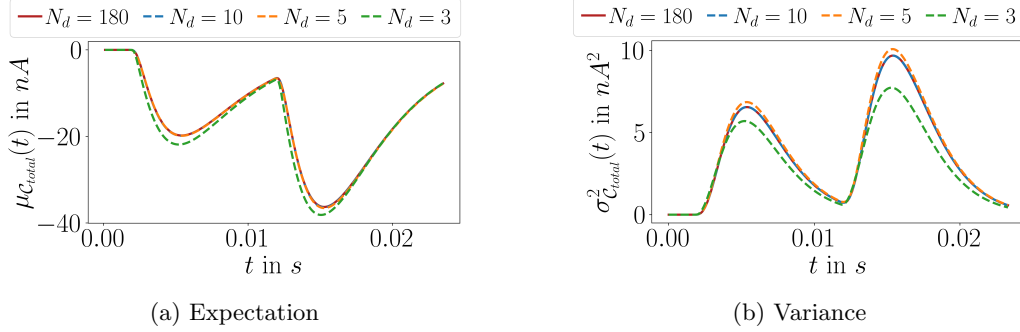


Figure 11: **Comparison of different degrees of spatial coarseness.** (a) Expectation  $\mu_{C_{total}}(t) = \mathbb{E}[C_{total}(t)]$  of total junction current and (b) its variance  $\sigma_{C_{total}}^2(t) = \mathbb{V}[C_{total}(t)]$  for different numbers of coarse graining bins  $N_d$  (dashed) compared to the original system with no coarse graining (solid red,  $N_d = 180$ ). The system consisted of 180 release sites grouped in  $N_d$  locations with 2 applied stimuli as shown in Figure 7 and time discretization  $\Delta_t = 1e-4$ .

determining a favorable CG approach since the analytically exact results allow for direct comparison between different levels of coarseness.

#### 4.4 Peak analysis: Local minima of the mean current

Of special interest are the peaks (i.e., the local minima) of the resulting total current. While in [20], the mean and variances of the random local minima of  $C_{total}$  are analyzed (see Appendix 6.3), we will here consider the local minima of the mean current  $\mu_{C_{total}}$ . To that end, we recursively define the time points  $T^{(n)} \geq 0$  where the gradient of  $\mu_{C_{total}}$  is zero by setting  $T^{(0)} := 0$  and

$$T^{(n)} := \min \left\{ t > T^{(n-1)} : \frac{d}{dt} \mu_{C_{total}}(t) = 0 \right\}$$

for  $n \geq 1$ . Then,

$$\mathcal{M}_n := \mu_{C_{total}}(T^{(2n-1)}) \quad (22)$$

is the  $n$ th local minimum of  $\mu_{C_{total}}$  for  $n \geq 1$ , and

$$\mathcal{V}_n := \sigma_{C_{total}}^2(T^{(2n-1)}) \quad (23)$$

is the variance at the corresponding time point. Figure 12 shows the variance  $\mathcal{V}_1$  versus the first local minimum  $\mathcal{M}_1$  at time  $T^{(1)}$  for different external calcium concentrations  $[Ca^{2+}]_{ext}$  in comparison to the corresponding values calculated from experimental data by [20]. The experimental data points are averaged over 6 different animals, with 9 repetitions for each animal. Due to the low number of samples in the experiment and the considerable variation between the different animals (indicated by the large standard deviations in Figure 12), it is very difficult to perform a sound analysis of the agreement between our results and the observed data. Nevertheless, we find that all our calculated values lie within the confidence area given by the statistical mean plus/minus the standard deviation. Most importantly, one can observe an approximately parabolic pattern in the relation between mean and variance, just as described in [20].

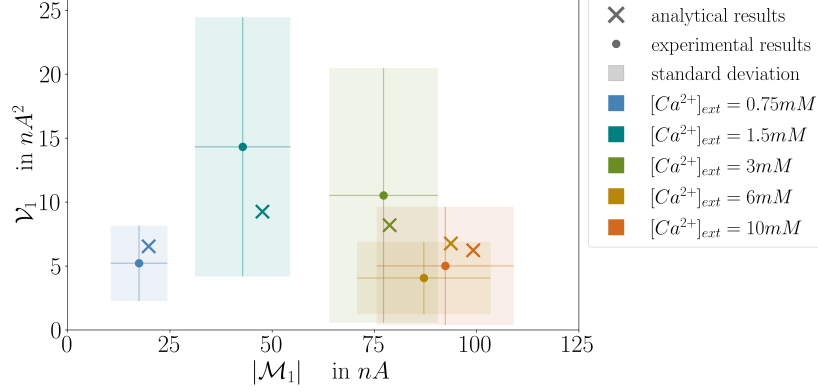


Figure 12: **Peak analysis.** Variance  $\mathcal{V}_1 = \sigma_{\mathcal{C}_{total}}^2(T^{(1)})$  versus first peak  $\mathcal{M}_1 = \mu_{\mathcal{C}_{total}}(T^{(1)})$  of the mean total current. The crosses refer to the values calculated analytically using (22) and (23). The points were calculated from experimental data and are the averages of 6 animals, with 9 repetitions per animal.

## 5 Conclusion

In this work, the process of neurotransmission served as a motivation to analyse moments of filtered signals for stochastic linear reaction networks with time-dependent reaction rates. We examined the total postsynaptic current, which is the quantity measured and explored in electrophysiological experiments, from the mathematical perspective, considering it as the output signal from a linear time-invariant filter. We showed that the second-order moments of such output signals may be quantified by the autocorrelation function of the underlying stochastic process. Although moment equations for stochastic reaction dynamics given by chemical master equations are well-studied in the literature, analogue equations for the corresponding cross-correlations outside of equilibrium have not been investigated so far. As one main result of the present work, we derived a characteristic PDE for the cross-correlation functions of linear reaction networks, which completes the system of moment equations and allows to analytically quantify the variance of the output signals of interest. By solving these equations, the filtered signal's variance can be calculated without any stochastic simulations of the underlying reaction process, thereby reducing numerical effort.

We demonstrated this by means of an exemplary reduced reaction system and via application to neurotransmission dynamics. For the latter, a spatial extension of the model for vesicle-fusion dynamics in the active zone was investigated. Our findings show the possibility of a drastic spatial coarse-graining: The large amount of individual release sites located along the presynaptic membrane may be grouped according to their distance to the calcium channel into a small number of equally spaced bins. This reduced system delivers a very good approximation for the original solution while significantly reducing the number of equations to be solved. Finally, we showed that the relation between mean and variance of the total postsynaptic current's peaks is in accordance with the results from experimental studies.

Our general result for the cross-correlation (Theorem 1) can be applied to any linear reaction network with time-dependent rates. This is of special significance for solving inverse problems such as parameter estimation for stochastic reaction systems: As demonstrated in the example of the vesicle fusion model, these systems can be very complex

and involve a great amount of different parameters that produce a random outcome, which makes finding a robust solution to the inverse problem extremely difficult. Here, we propose an accurate and fast method to solve the corresponding forward problem by finding the first- and second-order moments deterministically. This renders imprecise and computationally expensive stochastic simulations redundant and is therefore of great advantage in parameter estimation and model optimization. In the future, our approach may find application not only in the analysis of other time-dependent biochemical reaction networks, but also in any other area where stochasticity and impulse responses are relevant, e.g. particle detection or macroeconomic modeling. As we demonstrated for vesicle fusion models, our approach can be especially useful in the efficient and accurate exploration and comparison of coarse-graining approaches.

In the context of neurotransmission dynamics our results may be used to efficiently investigate the effect of varying vesicle distributions at synapses of different species or to optimize other neurotransmission models via spatial coarse graining or model reduction. Further research into the mathematical part of this work could include the investigation of the effect of other (e.g. time-variant) filters. Furthermore, it may be interesting to research the possibility of extending our approach to approximate the cross-correlation of nonlinear reaction networks, e.g. by a stepwise linearization of the reaction rates.

## Acknowledgements

We would like to thank Alexander Walter for fruitful discussions. This research has been partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through grant CRC 1114/2 and under Germany’s Excellence Strategy – The Berlin Mathematics Research Center MATH+ (EXC-2046/1 project ID: 390685689).

## References

- [1] Thomas C. Südhof. The synaptic vesicle cycle. *Annual Review of Neuroscience*, 27:509–547, July 2004.
- [2] Matthijs Verhage and Jakob B. Sørensen. Vesicle docking in regulated exocytosis. *Traffic*, 9:1414–1424, 2008.
- [3] Pascal S. Kaeser and Wade G. Regehr. The readily releasable pool of synaptic vesicles. *Current Opinion in Neurobiology*, 43:63–70, 2017.
- [4] Rafael Fernández-Chacón, Andreas Königstorfer, Stefan H. Gerber, Jesús García, Maria F. Matos, Charles F. Stevens, Nils Brose, Josep Rizo, Christian Rosenmund, and Thomas C. Südhof. Synaptotagmin I functions as a calcium regulator of release probability. *Nature*, 410:41–49, March 2001.
- [5] J. Troy Littleton and Hugo J. Bellen. Synaptotagmin controls and modulates synaptic-vesicle fusion in a  $Ca^{2+}$ -dependent manner. *Trends in Neurosciences*, 18:177–183, 1995.
- [6] Massimo Scanziani and Michael Häusser. Electrophysiology in the age of light. *Nature*, 461:930–939, 2009.
- [7] Tina Ghelani and Stephan J. Sigrist. Coupling the structural and functional assembly of synaptic release sites. *Frontiers in Neuroanatomy*, 12, 2018.

- [8] Hong Zhan, Joseph Bruckner, Ziheng Zhang, and Kate O'Connor-Giles. Three-dimensional imaging of *Drosophila* motor synapses reveals ultrastructural organizational patterns. *Journal of Neurogenetics*, 30:237–246, October 2016.
- [9] Nadine Ehmann, David Oswald, and Robert J. Kittel. *Drosophila* active zones: From molecules to behaviour. *Neuroscience Research*, 127:14–24, February 2018.
- [10] Chuang-Chung J. Lee, Mihai Anton, Chi-Sang Poon, and Gregory J. McRae. A kinetic model unifying presynaptic short-term facilitation and depression. *Journal of Computational Neuroscience*, 26:459, December 2008.
- [11] Elham Bayat Mokhtari, J. Josh Lawrence, and Emily F. Stone. Data driven models of short-term synaptic plasticity. *Frontiers in Computational Neuroscience*, 12, 2018.
- [12] Calvin Zhang and Charles S. Peskin. Analysis, simulation, and optimization of stochastic vesicle dynamics in synaptic transmission. *Communications on Pure and Applied Mathematics*, 0, October 2019.
- [13] Bin Pan and Robert S. Zucker. A general model of synaptic transmission and short-term plasticity. *Neuron*, 62:539–554, May 2009.
- [14] Marie Levakova, Massimiliano Tamborrino, Susanne Ditlevsen, and Petr Lansky. A review of the methods for neuronal response latency estimation. *Biosystems*, 136:23–34, 2015.
- [15] Yukiko Goda and Thomas C Südhof. Calcium regulation of neurotransmitter release: reliably unreliable? *Current opinion in cell biology*, 9:513–518, 1997.
- [16] Ege T. Kavalali. Neuronal  $\text{Ca}^{2+}$  signalling at rest and during spontaneous neurotransmission. *The Journal of Physiology*, 598:1649–1654, 2020.
- [17] Bernard Katz. The release of neural transmitter substances. *Liverpool University Press*, pages 5–39, 1969.
- [18] Victor Matveev, Richard Bertram, and Arthur Sherman. Calcium cooperativity of exocytosis as a measure of  $\text{Ca}^{2+}$  channel domain overlap. *Brain Research*, 1398:126–138, June 2011.
- [19] FA Dodge Jr. and Rami Rahamimoff. Co-operative action of calcium ions in transmitter release at the neuromuscular junction. *The Journal of physiology*, 193:419–432, 1967.
- [20] Janus RL Kobbersmed, Andreas T. Grasskamp, Meida Jusyte, Mathias A. Böhme, Susanne Ditlevsen, Jakob Balslev Sørensen, and Alexander M. Walter. Rapid regulation of vesicle priming explains synaptic facilitation despite heterogeneous vesicle:  $\text{Ca}^{2+}$  channel distances. *Elife*, 9:e51032, 2020.
- [21] Tobias Jahnke and Wilhelm Huisinga. Solving the chemical master equation for monomolecular reaction systems analytically. *Journal of Mathematical Biology*, 54(1):1–26, 2007.
- [22] Patrick B. Warren, Sorin Tănase-Nicola, and Pieter Rein ten Wolde. Exact results for noise power spectra in linear biochemical reaction networks. *The Journal of Chemical Physics*, 125:144904, 2006.

- [23] Petrus MT Broersen. *Automatic autocorrelation and spectral analysis*. Springer Science & Business Media, 2006.
- [24] David Schnoerr, Guido Sanguinetti, and Ramon Grima. Approximation and inference methods for stochastic biochemical kinetics—a tutorial review. *Journal of Physics A: Mathematical and Theoretical*, 50(9):093001, 2017.
- [25] Victor Matveev, Arthur Sherman, and Robert S. Zucker. New and corrected simulations of synaptic facilitation. *Biophysical Journal*, 83:1368–1373, 2002.
- [26] John M. Bekkers. Convolution of mini distributions for fitting evoked synaptic amplitude histograms. *Journal of Neuroscience Methods*, 130:105–114, 2003.
- [27] Robert M. Gray and Lee D. Davisson. *An introduction to statistical signal processing*. Cambridge University Press, 2004.
- [28] Chetan Gadgil, Chang Hyeong Lee, and Hans G. Othmer. A stochastic analysis of first-order reaction networks. *Bulletin of Mathematical Biology*, 67:901–946, 2005.
- [29] Stefanie Winkelmann and Christof Schütte. *Stochastic Dynamics in Computational Biology*. Springer, 2020.
- [30] Nataša Djurdjevac Conrad, Luzie Helfmann, Johannes Zonker, Stefanie Winkelmann, and Christof Schütte. Human mobility and innovation spreading in ancient times: a stochastic agent-based simulation approach. *EPJ Data Science*, 7:1–22, 2018.



## 6 Appendix

### 6.1 Proof of Theorem 1

For fixed  $l, l' \in \{1, \dots, L\}$  and fixed  $s \geq 0$  we have, for  $t \geq s$ ,

$$\begin{aligned}
& \frac{d}{dt} a_{l,l'}(t, s) \\
&= \frac{d}{dt} \mathbb{E}[X_l(t) X_{l'}(s)] \\
&= \frac{d}{dt} \sum_{x, y \in \mathbb{N}_0} xy \mathbb{P}[X_l(t) = x, X_{l'}(s) = y] \\
&= \frac{d}{dt} \sum_{y \in \mathbb{N}_0} \sum_{\mathbf{x} \in \mathbb{N}_0^L} x_l y \mathbb{P}[\mathbf{x}(t) = \mathbf{x}, X_{l'}(s) = y] \\
&= \sum_{k=1}^K \sum_{y \in \mathbb{N}_0} \sum_{\mathbf{x} \in \mathbb{N}_0^L} x_l y \left( \lambda_k(\mathbf{x} - \boldsymbol{\nu}^{(k)}) \mathbb{P}[\mathbf{x}(t) = \mathbf{x} - \boldsymbol{\nu}^{(k)}, X_{l'}(s) = y] - \lambda_k(\mathbf{x}) \mathbb{P}[\mathbf{x}(t) = \mathbf{x}, X_{l'}(s) = y] \right) \\
&= \sum_{k \in \mathcal{K}_0} \sum_{y \in \mathbb{N}_0} \sum_{\mathbf{x} \in \mathbb{N}_0^L} x_l y \left( \gamma_k(t) \mathbb{P}[\mathbf{x}(t) = \mathbf{x} - \boldsymbol{\nu}^{(k)}, X_{l'}(s) = y] - \gamma_k(t) \mathbb{P}[\mathbf{x}(t) = \mathbf{x}, X_{l'}(s) = y] \right) \\
&\quad + \sum_{k \in \mathcal{K}_1} \sum_{y \in \mathbb{N}_0} \sum_{\mathbf{x} \in \mathbb{N}_0^L} x_l y \left( \gamma_k(t) (x_{l_k} - \nu_{l_k}^{(k)}) \mathbb{P}[\mathbf{x}(t) = \mathbf{x} - \boldsymbol{\nu}^{(k)}, X_{l'}(s) = y] - \gamma_k(t) x_{l_k} \mathbb{P}[\mathbf{x}(t) = \mathbf{x}, X_{l'}(s) = y] \right) \\
&= \sum_{k \in \mathcal{K}_0} \sum_{y \in \mathbb{N}_0} \sum_{\mathbf{x} \in \mathbb{N}_0^L} \gamma_k(t) y \left( (x_l + \nu_l^{(k)}) \mathbb{P}[\mathbf{x}(t) = \mathbf{x}, X_{l'}(s) = y] - x_l \mathbb{P}[\mathbf{x}(t) = \mathbf{x}, X_{l'}(s) = y] \right) \\
&\quad + \sum_{k \in \mathcal{K}_1} \sum_{y \in \mathbb{N}_0} \sum_{\mathbf{x} \in \mathbb{N}_0^L} \gamma_k(t) y \left( (x_l + \nu_l^{(k)}) x_{l_k} \mathbb{P}[\mathbf{x}(t) = \mathbf{x}, X_{l'}(s) = y] - x_l x_{l_k} \mathbb{P}[\mathbf{x}(t) = \mathbf{x}, X_{l'}(s) = y] \right) \\
&= \sum_{k \in \mathcal{K}_0} \sum_{y \in \mathbb{N}_0} \sum_{\mathbf{x} \in \mathbb{N}_0^L} \gamma_k(t) y \nu_l^{(k)} \mathbb{P}[\mathbf{x}(t) = \mathbf{x}, X_{l'}(s) = y] \\
&\quad + \sum_{k \in \mathcal{K}_1} \sum_{y \in \mathbb{N}_0} \sum_{\mathbf{x} \in \mathbb{N}_0^L} \gamma_k(t) y \nu_l^{(k)} x_{l_k} \mathbb{P}[\mathbf{x}(t) = \mathbf{x}, X_{l'}(s) = y] \\
&= \sum_{k \in \mathcal{K}_0} \gamma_k(t) \nu_l^{(k)} \sum_{y \in \mathbb{N}_0} y \mathbb{P}[X_{l'}(s) = y] \\
&\quad + \sum_{k \in \mathcal{K}_1} \gamma_k(t) \nu_l^{(k)} \sum_{y \in \mathbb{N}_0} \sum_{x_{l_k} \in \mathbb{N}_0} y x_{l_k} \mathbb{P}[X_{l_k}(t) = x_{l_k}, X_{l'}(s) = y] \\
&= \sum_{k \in \mathcal{K}_0} \gamma_k(t) \nu_l^{(k)} \mathbb{E}[X_{l'}(s)] + \sum_{k \in \mathcal{K}_1} \gamma_k(t) \nu_l^{(k)} \mathbb{E}[X_{l_k}(t) X_{l'}(s)].
\end{aligned}$$

Using the definitions of  $\mu_l(t)$  and  $a_{l,l'}(t, s)$  we obtain (15).

### 6.2 First- and second-order moments of the input signal

As in Sec. 3.2, let  $0 \leq T_1 \leq T_2 \leq \dots$  denote the random time points where any of the key reactions  $\mathcal{R}_k$ ,  $k \in \mathcal{K}^*$ , happens. For the first-order moment of the input signal  $f$  we have

$$\mathbb{E}[f(t)] = \sum_{j=1}^{\infty} \mathbb{E}[\delta(t - T_j)] = \sum_{j=1}^{\infty} \int_0^{\infty} \delta(t - s) \rho_j(s) ds = \sum_{j=1}^{\infty} \rho_j(t)$$

where  $\rho_j$  is the probability density (exponential distribution, continuous function) of the jump time  $T_j$ . For the second-order moments we get

$$\begin{aligned}
& \mathbb{E}[f(t)f(t')] \\
&= \mathbb{E} \left[ \left( \sum_{i=1}^{\infty} \delta(t - T_i) \right) \left( \sum_{j=1}^{\infty} \delta(t' - T_j) \right) \right] \\
&= \sum_{i=1}^{\infty} \mathbb{E} [\delta(t - T_i) \delta(t' - T_i)] + \sum_{\substack{i,j=1 \\ j \neq i}}^{\infty} \mathbb{E} [\delta(t - T_i) \delta(t' - T_j)] \\
&= \sum_{i=1}^{\infty} \int_0^{\infty} \delta(t - s) \delta(t' - s) \rho_i(s) ds + \sum_{\substack{i,j=1 \\ j \neq i}}^{\infty} \int_0^{\infty} \int_0^{\infty} \delta(t - s) \delta(t' - s') \rho_{ij}(s, s') ds ds' \\
&= \sum_{i=1}^{\infty} \delta(t - t') \rho_i(t) + \sum_{\substack{i,j=1 \\ j \neq i}}^{\infty} \int_0^{\infty} \delta(t' - s') \rho_{ij}(t, s') ds' \\
&= \sum_{i=1}^{\infty} \delta(t - t') \rho_i(t) + \sum_{\substack{i,j=1 \\ j \neq i}}^{\infty} \rho_{ij}(t, t')
\end{aligned}$$

where  $\rho_{ij}$  is the joint density function of two jump times  $T_i, T_j$ ,  $i \neq j$ . (Especially, we have  $\rho_{ij}(t, t') = 0$  for  $t' < t$  and  $i < j$ .)

### 6.3 Peak analysis: Comparison

In [20], an analysis of the peaks and peak variances resembling the one in 4.4 is performed. However, there is an important difference between our method and theirs in the order of taking the expectation and the local minima: The authors consider the first minimum of each of the *individual realisations*  $\mathcal{C}_{total}$  and calculate expectation and variances afterwards.<sup>5</sup> Formally, they split the time interval into two intervals  $I_1 := [0, T']$  and  $I_2 := [T', T]$  for a certain time point  $T' \in [0, T]$ . One can then define the mean

$$\tilde{\mathcal{M}}_n := \mathbb{E} \left[ \min_{t \in I_n} \mathcal{C}_{total}(t) \right] \quad (24)$$

and the variance

$$\tilde{\mathcal{V}}_n := \mathbb{V} \left[ \min_{t \in I_n} \mathcal{C}_{total}(t) \right] \quad (25)$$

of the temporal minimum of the output current  $\mathcal{C}_{total}$  over the time interval  $I_n$  for  $n = 1, 2$ . In [20], they estimate  $\tilde{\mathcal{M}}_1$  and  $\tilde{\mathcal{V}}_1$  for the case of two stimuli by means of MC-simulations. These are different from  $\mathcal{M}_1$  and  $\mathcal{V}_1$  defined in (22) and (23), respectively, where

$$\mathcal{M}_1 = \mu_{\mathcal{C}_{total}}(T^{(1)}) = \mathbb{E}[\mathcal{C}_{total}(T^{(1)})] = \min_{t \in I_1} \mathbb{E}[\mathcal{C}_{total}(t)]$$

for  $I_1 := [0, T^{(2)}]$ , and

$$\mathcal{V}_1 = \sigma_{\mathcal{C}_{total}}^2(T^{(1)}) = \mathbb{V}[\mathcal{C}_{total}(T^{(1)})].$$

This difference holds true for all other  $n$ , as well.

<sup>5</sup>This is standard practice in biology in order to allow for better comparison between different animals.

**Remark 3** (Alternative definition by means of gradient). *In analogy to the definitions in Sec. 4.4, we can recursively define the (random) time points  $\tilde{T}^{(n)} \geq 0$  where the gradient of the current is zero by setting  $\tilde{T}^{(0)} := 0$  and*

$$\tilde{T}^{(n+1)} := \min \left\{ t > \tilde{T}^{(n)} : \frac{d}{dt} \mathcal{C}_{total}(t) = 0 \right\}.$$

*By this definition, we can set*

$$\tilde{\mathcal{M}}_n := \mathbb{E} \left[ \mathcal{C}_{total}(\tilde{T}^{(2n-1)}) \right]$$

$$\tilde{\mathcal{V}}_n := \mathbb{V} \left[ \mathcal{C}_{total}(\tilde{T}^{(2n-1)}) \right].$$

*However, this all relies on the assumption that there are no saddle points. Moreover, it can happen that  $\frac{d}{dt} \mathcal{C}_{total}(t) = 0$  holds for other time points which are local minima/maxima or saddle points but not the 'peaks' we want to identify by (24). On the other hand, this approach can be advantageous since one does not need to assign the number of peaks in advance and no prior assumptions on their time points are made.*