

RESEARCH

Open Access



# Data-driven inference for stationary jump-diffusion processes with application to membrane voltage fluctuations in pyramidal neurons

Alexandre Melanson<sup>1,2\*</sup>  and André Longtin<sup>1,3,4</sup>

\*Correspondence:

[melanson.alexandre@gmail.com](mailto:melanson.alexandre@gmail.com)

<sup>1</sup>Department of Physics, University of Ottawa, Ottawa, Canada

<sup>2</sup>Département de physique et d'astronomie, Université de Moncton, Moncton, Canada  
Full list of author information is available at the end of the article

## Abstract

The emergent activity of biological systems can often be represented as low-dimensional, Langevin-type stochastic differential equations. In certain systems, however, large and abrupt events occur and violate the assumptions of this approach. We address this situation here by providing a novel method that reconstructs a jump-diffusion stochastic process based solely on the statistics of the original data. Our method assumes that these data are stationary, that diffusive noise is additive, and that jumps are Poisson. We use threshold-crossing of the increments to detect jumps in the time series. This is followed by an iterative scheme that compensates for the presence of diffusive fluctuations that are falsely detected as jumps. Our approach is based on probabilistic calculations associated with these fluctuations and on the use of the Fokker–Planck and the differential Chapman–Kolmogorov equations. After some validation cases, we apply this method to recordings of membrane noise in pyramidal neurons of the electrosensory lateral line lobe of weakly electric fish. These recordings display large, jump-like depolarization events that occur at random times, the biophysics of which is unknown. We find that some pyramidal cells increase their jump rate and noise intensity as the membrane potential approaches spike threshold, while their drift function and jump amplitude distribution remain unchanged. As our method is fully data-driven, it provides a valuable means to further investigate the functional role of these jump-like events without relying on unconstrained biophysical models.

**Keywords:** Stochastic differential equations; Jump-diffusion processes; Membrane noise; Channel noise; Electric fish; Pyramidal neurons; Fokker–Planck equation; Chapman–Kolmogorov equation

## 1 Introduction

Complex systems are ubiquitous in many areas of science, including biology, neuroscience, climatology, engineering, as well as in finance and social sciences [1, 2]. The common feature uniting such vastly different systems is the nonlinear interaction between their numerous microscopic constituents. This collective activity leads to the emergence of macroscopic order that cannot be reduced to microscopic properties [3]. It is often these macroscopic variables that can be measured experimentally, allowing the emergent dynamics of

the system to be captured by low-dimensional data sets. The search for a macroscopic-level representation of the system thus relies on extracting dynamical information from observed time series [4].

For some systems, the high-dimensional, microscopic degrees of freedom can be well approximated by simple stochastic fluctuations. These fluctuations often participate as dynamical noise in the macroscopic evolution of the system. The low-dimensional representation of the system's dynamics can then be expressed as a stochastic dynamical system [5]. More specifically, the observed data are usually assumed to satisfy a Langevin-type stochastic differential equation (SDE). A common approach is to obtain the drift and diffusion functions of this equation by estimating the first and second Kramers–Moyal coefficient [6, 7]. The resulting model is completely data-driven and captures the core phenomenology of the original data without relying on knowledge or assumptions about the microscopic constituents of the observed system. This approach has been successfully applied in a variety of contexts, ranging from neuronal dynamics [8–12], heart rate variability [13, 14], turbulence [15, 16], calibration of optical tweezers [17], and others (see Ref. [5] for a review).

In all mentioned cases, however, the noise is assumed to be purely diffusive, i.e., random fluctuations with continuous sample paths. This description is incomplete if, in addition to diffusive fluctuations, large and abrupt events appear at random times throughout the time series. In this case, jump-diffusion stochastic processes provide a more appropriate framework to model these data. Jump-diffusion processes have been used in neuroscience as a model for the spatial [18–21] and temporal [22] organization of synaptic bombardment, in physics as a model for noise-driven transport in ratchet potentials [23–25], as well as in finance [26, 27] and soil moisture dynamics [28]. The Langevin approach is likely to fail if the observed system exhibits jump-diffusion characteristics, such as skewed distributions and sudden large jumps. In such cases, and especially when the microscopic dynamics are unknown, extracting a phenomenological model from the experimental data would provide a valuable tool to probe the dynamics of the observed system, its interaction with other systems, and the interplay between diffusive and jump noise sources in shaping the observed behavior.

In this paper, we present a novel, data-driven inference method that fits a jump-diffusion SDE to experimental time series. By detecting jumps through threshold-crossing and by calculating the contribution of diffusive fluctuations that are falsely detected as jumps, we iteratively estimate the drift function, noise intensity, jump rate, and jump amplitude distribution. The result of this semi-parametric method is a jump-diffusion SDE that successfully fits the original data. Our method is applicable in cases where these data are stationary, with additive diffusive noise and Poisson jumps. Note that other studies have attempted to infer jump-diffusion dynamics from data, but they rely on assuming a parametric form of the jump amplitude distribution [29–31], or consider only Lévy processes [32, 33].

We test our method with two validation cases, where realizations of known jump-diffusion processes with different characteristics are used as validation data. In both cases, by using only the simulated time series, we precisely recover the correct parameters and functions used in the original simulations. We also compare the autocorrelation functions (ACF) of the validation data and of fitted SDE. We then apply our method to recordings of intrinsic membrane voltage fluctuations in pyramidal neurons of electric fish. These

recordings contain sudden and unpredictable jump-like events that occur among more typical diffusive membrane noise. Although the exact biophysical origin of these fluctuations is unknown, we find that the recordings are well fitted by jump-diffusion processes. We evaluate goodness of fit quantitatively by comparing the observed and the estimated probability density functions (PDF) and power spectral densities (PSD). Interestingly, we find that some pyramidal cells increase their jump rate and noise intensity as the membrane potential approaches spike threshold, while their drift function and jump amplitude distribution remain unchanged.

In Sect. 2, we present the various steps involved in the inference procedure, including the detection scheme, the calculation relating to false positive, and the iterative approach. In Sect. 3, we validate this procedure against one pure diffusion and two jump-diffusion test cases, and then we apply it to neurophysiological recordings of membrane noise. This is followed by a discussion on the possible generalizations of our method and on future work with the experimental data (Sect. 4).

## 2 Methods

### 2.1 Definitions and overview

Let  $\{X(t)\}$  represent data in the form of a stationary time series. In what follows,  $\{X(t)\}$  is obtained either from experimental observations or from numerical simulation. In either case, the situation of interest here is when  $\{X(t)\}$  exhibits both diffusive fluctuations and abrupt events (henceforth called jumps). From  $\{X(t)\}$ , our goal is then to fit a jump-diffusion SDE of the form

$$dY(t) = F(Y(t)) dt + \sqrt{2D} dW(t) + dJ(t), \quad (1)$$

where  $F$  is the drift function,  $D$  is the noise intensity, and  $W(t)$  is a Wiener process (i.e., Brownian motion). Here  $J(t)$  is a compound Poisson process representing the jumps

$$J(t) = \sum_{i=0}^{N_{\lambda}(t)} B_i, \quad (2)$$

where  $N_{\lambda}(t)$  is a Poisson point process with rate  $\lambda$ , and the  $B_i$ 's are the independent and identically distributed jump amplitudes drawn from a distribution  $Q_B$ . For a small enough sampling interval, or time step  $\Delta t$ , jumps will occur with probability  $\Gamma_B \equiv \lambda \Delta t$ .

Although the method developed herein is applicable to a wide array of experimental data types, we do, however, impose certain conditions on the underlying dynamical process. Notably, we limit our analysis to systems where the dynamics are stationary (in the strict sense), the diffusion noise is additive, the jumps have positive amplitudes ( $B_i > 0$ ), and where the Poisson rate  $\lambda$  is constant in time and small enough so that  $\Gamma_B \ll 1$ . Furthermore, we assume here that  $F$  is continuous and that a single stable fixed point arises from the deterministic part of the dynamics, but our method could be generalized to multistable systems. Except for this restriction, we assume no particular shape for the drift function as long as it generates a stationary process. Finally, we assume that, on average, jump amplitudes are greater than the typical magnitude of diffusive increments:  $E[B_i] > (2D\Delta t)^{\frac{1}{2}}$  by one order of magnitude or more.

In terms of the time series itself, we assume that the data were sampled at a high-frequency, such that  $\Delta t$  can be assumed to be small with respect to the total duration

of the time series. Note that the value of  $\Delta t$  is set from the experiments that produced the data, and so is not a variable we can control. For experimental data, however, it is possible that the Markov property breaks down on the timescale of individual observations [5], but we assume that, in that case, the time series can be downsampled to a timescale where the Markov property then holds.

Furthermore, in the cases considered below, jumps appear unambiguously in the data, and so a jump-diffusion approach is warranted. In situations where this is not the case, the presence of jumps can be assessed from the fourth-order Kramers–Moyal coefficient, which is non-vanishing for processes with discontinuities [31, 34]. Lastly, note that we make no assumptions regarding the structure of the ACF of  $\{X(t)\}$ ; we only assume stationarity in these data. In fact, we compare the ACFs (and PSDs) of time series generated by the fitted SDE with those of the original data as a means of validating the inference procedure.

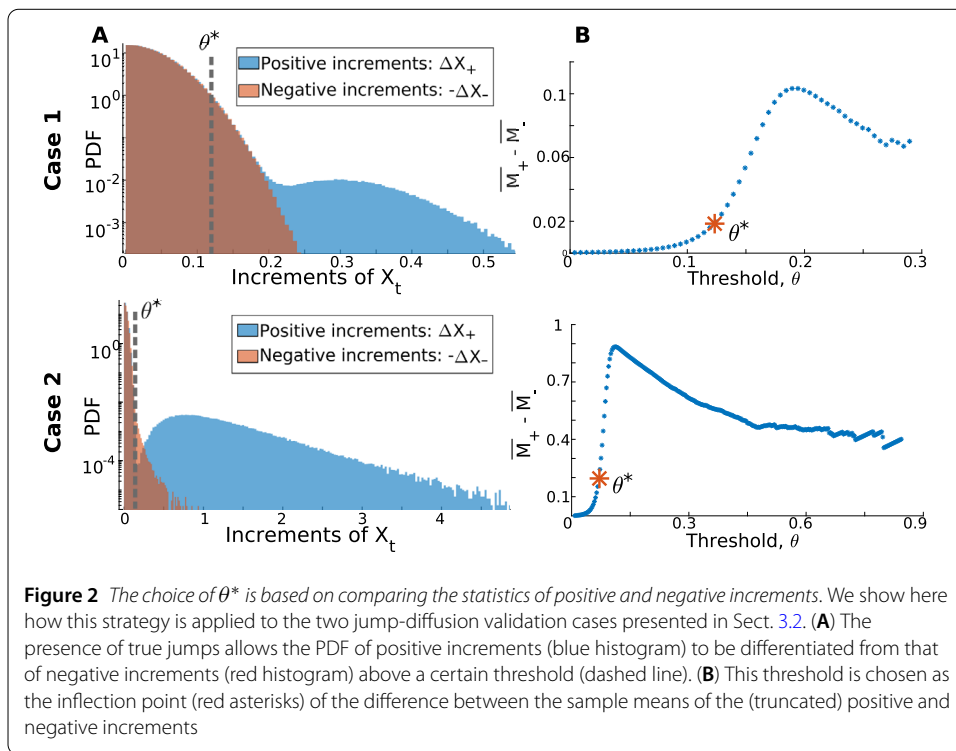
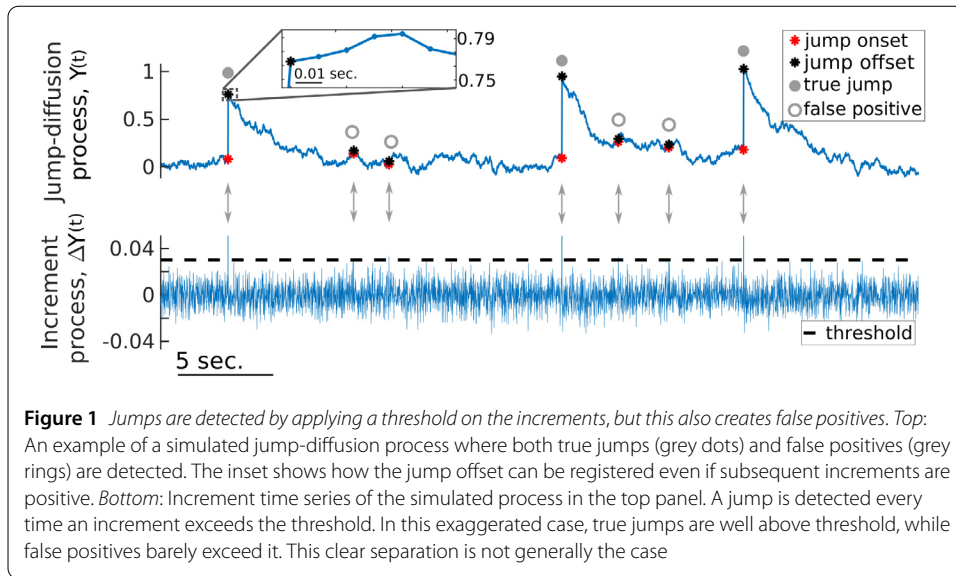
Below we develop a data-driven inference procedure that successfully generates the estimates  $\hat{D}$ ,  $\hat{\lambda}$ ,  $\hat{F}$ , and  $\hat{Q}_B$ , where the unknown functions  $F$  and  $Q_B$  are estimated non-parametrically. This inference procedure results in a fitted stochastic process  $Y(t)$  that is an adequate model of the original data. In this sense, we implicitly assume that the data  $\{X(t)\}$  are sampled from a realization of  $Y(t)$ . In our calculations, we thus associate the first-order equilibrium PDF of  $Y(t)$ ,  $P_Y$ , to the empirical PDF of  $\{X(t)\}$  (obtained by kernel density estimation)  $P_X$ .

Our approach is predicated on the detection of jumps in the data via the application of a threshold  $\theta^* > 0$  on the increments  $\Delta X(t + \Delta t) \equiv X(t + \Delta t) - X(t)$ . This procedure creates a pool of detected jumps with various amplitudes. Let  $Q_C$  be the empirical PDF estimated from these measured jump amplitudes. Also, let  $n$  be the total number of increments in the time series, and  $m$  be the number of increments whose value is greater than  $\theta^*$ . We define the (overall) jump detection probability as  $\Gamma_C \equiv \text{Prob}\{\text{detecting an increment larger than } \theta^* \text{ across an interval } \Delta t\}$ , which we estimate from the data as  $\Gamma_C = \frac{m}{n}$ .

An inherent challenge with this threshold-crossing approach is that, in addition to the true jumps generated by the compound Poisson process, we also unavoidably detect large diffusive fluctuations that are falsely identified as jumps, henceforth called false positives (FP; Fig. 1, top). A direct estimation of the true jump rate  $\lambda$  and of the true jump amplitude distribution  $Q_B$  is thus impossible because the detected jump pool consists of a mixture of true jumps and false positives. Our main contribution, and the central component of our inference procedure, is the calculation of FP-related statistics, namely the FP detection probability  $\Gamma_A \equiv \text{Prob}\{\text{detecting a diffusive increment larger than } \theta^* \text{ across a given time step } \Delta t\}$  and the distribution from which FP amplitudes are drawn  $Q_A$ . Once  $\Gamma_A$  and  $Q_A$  are calculated, we then extract  $\lambda$  (or, equivalently,  $\Gamma_B$ ) and  $Q_B$  from  $\Gamma_C$  and  $Q_C$ . Note that the subscripts “A”, “B”, and “C” will hereafter refer to FPs, true jumps, and both combined, respectively. More precisely, we measure quantities in “C” from the detected jump pool, we calculate FP statistics in “A”, and we seek the true jump statistics in “B”.

## 2.2 Choice of threshold

To ensure that a minimal number of true jumps are missed during the jump detection procedure, the threshold  $\theta^*$  should ideally be set as low as possible. If set too low, however, the number of FPs becomes so large that the statistics of true jumps, i.e.,  $\lambda$  and  $Q_B$ , cannot



be extracted from the statistical fluctuations of  $Q_C$  and  $\Gamma_C$ . We thus aim for an intermediate value of  $\theta^*$  that captures most true jumps while allowing a manageable number of FPs. This is done by relying on our assumption of positive jump amplitudes and, more precisely, by exploiting the asymmetry between positive and negative increment statistics. Note that the threshold depends implicitly on the time step of the original time series: a smaller  $\Delta t$  means that diffusive fluctuations are smaller, and thus a smaller value of  $\theta^*$  can be used.

Let  $\{\Delta X_+\}$  and  $\{\Delta X_-\}$  be the sets of positive and negative increments of  $\{X(t)\}$ , respectively (Fig. 2(A)). In addition, let  $M_+(\theta) = \{\Delta X_+ : \Delta X_+ > \theta\}$  and  $M_-(\theta) = \{-\Delta X_- : -\Delta X_- >$

$\theta$ ) be the reduced sets truncated by  $\theta$ , where  $\theta$  spans the common range of  $\{\Delta X_+\}$  and  $\{-\Delta X_-\}$ . We use the difference between the sample mean of these sets  $\overline{M}_+ - \overline{M}_-$  as a function of  $\theta$  to quantify the relative importance of true jumps with respect to diffusive fluctuations, over different increment sizes. The value of  $\theta$  for which  $\overline{M}_+ - \overline{M}_-$  is a maximum corresponds to the greatest separability between positive and negative increments. We find, however, that using the inflection point, i.e., where the second derivative becomes 0, located to the left of this maximum (Fig. 2(B), asterisks) is a better choice of threshold. This slightly lower value retains a greater range of true jumps, which is desirable, while avoiding the inclusion of an overwhelmingly large number of FPs. Choosing this inflection point rather than the maximum impacts primarily the estimation of  $\lambda$ , since it relies on the proper detection of true jumps in the time series. A slightly higher value of  $\theta^*$  will introduce a bias in  $\hat{\lambda}$ . For instance, in the first jump-diffusion validation case that we consider in Sect. 3.2, choosing the maximum as the threshold roughly yields a 1% increase in the error on  $\hat{\lambda}$  compared to when the inflection point is used. Our approach for setting  $\theta^*$  is thus motivated by the fact that it is advantageous to choose a value as small as possible for  $\theta^*$ , and the inflection point in the curve of  $\overline{M}_+ - \overline{M}_-$  provides a reliable way to achieve this.

### 2.3 Jump detection

Here we describe how the threshold is applied to the increments in order to generate the detected jump pool. We apply a detection scheme tailored specifically to handle two aspects that we observe in experimental data with jumps, and it is inspired by the method used in [35]. Firstly, if the data are resolved on a fine enough time scale, jumps may last longer than a single sampling interval. Secondly, jumps need not be followed immediately by negative increments. In data, and in some simulations as well, the diffusive increments following a jump can still be positive, but we seek a method for identifying when the jump actually ends (Fig. 1, inset). These two considerations shape the method used to calculate the FP amplitude distribution in Sect. 2.4.2.

When a given increment is larger than the threshold, a jump onset time  $T_{\text{on}}$  is registered, and if the next increment is below threshold, the associated offset time  $T_{\text{off}}$  is registered (even if this increment is positive). This defines a jump of duration  $T_{\text{off}} - T_{\text{on}} = \Delta t$ , where  $\Delta t$  is the sampling interval of the data. We henceforth refer to this type of jump as a singlet, as it spans the duration of a single time step. In contrast, if two or more successive increments are above threshold, then the jump is of duration  $2\Delta t$  or more, which we refer to as a doublet, triplet, and so forth. In other words, the jump offset time is only registered at the end of the sequence of above-threshold increments. With the onset and offset times identified, a jump amplitude is defined as the difference between  $X(T_{\text{off}})$  and  $X(T_{\text{on}})$ .

### 2.4 FP and true jump statistics

Here we present the calculations of the FP detection probability  $\Gamma_A$  and of the FP amplitude distribution  $Q_A$ . We then show how these quantities are used to extract the true jump rate  $\lambda$  and the true jump amplitude distribution  $Q_B$  from the detected jump probability  $\Gamma_C$  and the detected jump amplitude distribution  $Q_C$ . For the calculations in this subsection, we assume that the drift function  $F$  and noise intensity  $D$  are known. In the next subsection, we show how these calculations can be incorporated into an iterative scheme that allows the simultaneous estimation of all unknowns, including  $F$  and  $D$ . Furthermore, note

that the FP-related calculations involve only the diffusive part of Eq. (1) since, by definition, FPs occur during the purely diffusive segments between true jumps. The calculations in Sects. 2.4.1 and 2.4.2 thus pertain only to diffusive increments  $\Delta Y^{\text{diff}}(t)$ . Finally, note that we perform validation tests of the calculations of  $\Gamma_A$  and  $Q_A$  in Sect. 3.1.

#### 2.4.1 FP detection probability

As mentioned in Sect. 2.1, sampling a jump-diffusion process such as Eq. (1) at finite intervals and applying a threshold on the observed increments leads to the detection of FPs, i.e., diffusive (rather than true jump) increments larger than the threshold. Importantly, these FPs occur with a probability that depends on the value of the process at the start of the interval. Let this conditional detection probability be defined as

$$\begin{aligned}\alpha(y) &\equiv \text{Prob}\{\text{detecting an FP in the interval } [t, t + \Delta t], \text{ given that } Y(t) = y\} \\ &= \text{Prob}(\Delta Y^{\text{diff}}(t + \Delta t) > \theta^* | Y(t) = y).\end{aligned}\quad (3)$$

As  $\alpha$  does not depend explicitly on time, this definition relies on our assumption that  $Y(t)$  is stationary. The  $y$ -dependence arises from the drift function  $F$ . Indeed, if the drift function is positive (respectively, negative) at a given time, it biases diffusive fluctuations toward (away from) the threshold. This translates into an FP detection probability that assumes higher values when  $F(y) > 0$  than when  $F(y) < 0$ . We now turn to the explicit calculation of  $\alpha(y)$ .

Let  $\mathcal{E}_{\Delta Y|Y}(\xi | Y(t) = y)$  denote the PDF of  $\Delta Y^{\text{diff}}(t + \Delta t)$  conditioned on the value of the process at the start of the interval, and where  $\xi$  assumes the possible values of the increments. Note that, because the time step remains constant, it is always implied that the increments are defined across an interval  $\Delta t$ . Given that we assume  $\Delta t$  to be sufficiently small, we approximate  $\mathcal{E}_{\Delta Y|Y}$  as the short-time propagator of the Fokker–Planck equation [36, 37] associated with the diffusive part of jump-diffusion process (recall that what concerns us here are the purely diffusive segments between the true jumps of  $Y(t)$ ). We thus have  $\Delta Y^{\text{diff}}(t + \Delta t) \approx \mathcal{N}(F(Y(t))\Delta t, 2D\Delta t)$  and

$$\mathcal{E}_{\Delta Y|Y}(\xi | Y(t) = y) \approx \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left(-\frac{(\xi - F(y)\Delta t)^2}{4D\Delta t}\right), \quad (4)$$

that is, a Gaussian distribution with mean  $F(y)\Delta t$  and variance  $2D\Delta t$ .

For the test cases presented in Sect. 3 (with  $\Delta t = 10^{-4}$  s), we have validated this approximation by comparing it with numerical solutions of the associated Fokker–Planck equation solved at a finer temporal resolution ( $\Delta t/1000$ ) across the time step  $\Delta t$ . The numerical solutions were indeed well fitted with the approximation in Eq. (4) (not shown). Numerical integration was performed with a custom partial differential equation solver that implements a finite volume discretization along with the fully implicit Euler scheme. The advective term was treated with the upwind scheme, and a linear interpolation profile for the spatial derivative was applied to the diffusive term. The resulting algebraic equation was solved with the tridiagonal matrix algorithm [38].

Once the conditional PDF of the increments is evaluated with Eq. (4), we calculate the conditional FP detection probability, given that the process starts at  $y$ , as follows:

$$\alpha(y) = \int_{\theta^*}^{\infty} \mathcal{E}_{\Delta Y|Y}(\xi | Y(t) = y) d\xi, \quad (5)$$

that is, the probability of observing an increment larger than  $\theta^*$  starting at  $y$ . Finally, the unconditional FP detection probability is calculated based on the empirical PDF of  $\{X(t)\}$ ,  $P_X$ :

$$\Gamma_A = \int_{-\infty}^{\infty} \alpha(y) P_X(y) dy. \tag{6}$$

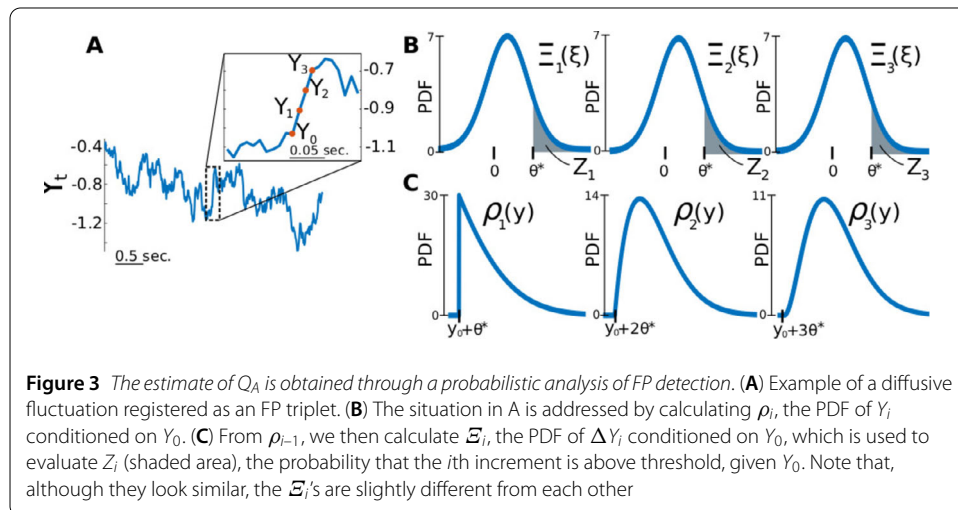
We validate these calculations in Sect. 3.1.

### 2.4.2 FP amplitude distribution

We now proceed with the calculation of  $Q_A$ , i.e., the distribution from which FP amplitudes are drawn. First, recall that our detection scheme allows for jumps of different durations (Sect. 2.3). As such, the detection of an FP implies either a succession of above-threshold increments (e.g., Fig. 3(A)) or, at least, a single above-threshold increment. Let  $T_{\text{on}}^{\text{FP}}$  denote the FP onset time, that is, the time at the start of the first above-threshold increment, and let  $Y_0 \equiv Y(T_{\text{on}}^{\text{FP}})$  be referred to as the starting value of the FP. Moreover, let  $\tau$  denote the FP duration, an integer multiple of  $\Delta t$ , such that  $\tau = \Delta t$  corresponds to an FP singlet,  $\tau = 2\Delta t$  to an FP doublet, and so forth.

In order to calculate  $Q_A$ , let us first identify the factors that influence FP amplitudes. Firstly, it must be noted that the FP amplitudes will exhibit a similar  $y$ -dependence as that discussed in the preceding section. Indeed, an increment starting at  $Y_0 \equiv Y(T_{\text{on}}^{\text{FP}}) = y$  will tend to be larger when  $F(y) > 0$  than when  $F(y) < 0$ . Secondly, the amplitude of an FP will also depend on its duration,  $\tau$ . For instance, the three increments of a triplet FP will summate and tend to have a larger amplitude than that of a singlet. The FP amplitudes  $A$  will thus covary with  $Y_0$  and with  $\tau$ , but note that  $\tau$  also depends on  $Y_0$ . Indeed, longer FPs will tend to occur where the drift function is more positive, and vice versa. To account for these dependencies, let us define the trivariate random variable  $\{A, \tau, Y_0\}$ , distributed according to its joint PDF  $P_{A,\tau,Y_0}(a, i\Delta t, y)$ , where we explicitly write  $\tau$  as an integer multiple of  $\Delta t$ . What we seek then is the marginal:

$$Q_A(a) = \int_{-\infty}^{\infty} \sum_{i=1}^{\infty} P_{A,\tau,Y_0}(a, i\Delta t, y) dy, \tag{7}$$





where the sum extends over all possible FP durations, and where  $a > 0$  represents all possible amplitudes. From the definition of conditional PDFs, we can expand the joint PDF as follows:

$$\begin{aligned}
 P_{A,\tau,Y_0}(a, i\Delta t, y) &= P_{A|\tau,Y_0}(a|i\Delta t, y)P_{\tau,Y_0}(i\Delta t, y) \\
 &= P_{A|\tau,Y_0}(a|i\Delta t, y)P_{\tau|Y_0}(i\Delta t|y)P_{Y_0}(y),
 \end{aligned}
 \tag{8}$$

where  $P_{\tau|Y_0}(i\Delta t|y) = \text{Prob}(\text{detecting an FP of duration } i\Delta t, \text{ given the starting value } y)$  is the conditional probability mass function of the FP duration  $\tau$ . We can thus write Eq. (7) as<sup>a</sup>

$$Q_A(a) = \int_{-\infty}^{\infty} \left( \sum_{i=1}^{\infty} P_{A|\tau,Y_0}(a|i\Delta t, y)P_{\tau|Y_0}(i\Delta t|y) \right) P_{Y_0}(y) dy.
 \tag{9}$$

The sum in the large parentheses is a function of  $y$ , and Eq. (9) is merely calculating its average with respect to the starting value  $Y_0$ . This sum can further be interpreted as a so-called mixture distribution: consider a collection of random variables, one of which is chosen according to a certain probability (its *mixture weight*) and is then realized according to its own PDF (its *mixture component*). The outcome of this experiment is itself a random variable whose PDF is called a *mixture distribution* and is expressed as a sum over the PDFs of the random variables in the collection, weighted by their respective probabilities. In our case, for a fixed value of  $Y_0$ , an FP duration is drawn according to a countable set of mixture weights  $P_{\tau|Y_0}$ , and an FP amplitude is then realized according to the associated mixture component  $P_{A|\tau,Y_0}$ . In practice, the sum will be truncated after the first few terms because the subsequent mixture weights become negligible.

From Eq. (9), we see that in order to arrive at the desired  $Q_A$ , the functions  $P_{A|\tau,Y_0}$ ,  $P_{\tau|Y_0}$ , and  $P_{Y_0}$  must first be calculated. Let us first consider the latter. Because  $Y_0$  represents, by definition, the value of  $Y(t)$  at the start of an above-threshold increment, we can express its PDF in terms of the joint PDF of  $\Delta Y^{\text{diff}}(t + \Delta t)$  and  $Y(t)$ :

$$\begin{aligned}
 P_{Y_0}(y) &= K \int_{\theta^*}^{\infty} P_{\Delta Y, Y}(\xi, y) d\xi \\
 &= K \int_{\theta^*}^{\infty} \Xi_{\Delta Y|Y}(\xi|Y(t) = y)P_Y(y) d\xi \\
 &= KP_Y(y) \int_{\theta^*}^{\infty} \Xi_{\Delta Y|Y}(\xi|Y(t) = y) d\xi \\
 &= KP_X(y)\alpha(y),
 \end{aligned}
 \tag{10}$$

where  $K$  is a normalization constant and where, in the last line, we have replaced  $P_Y$  by the empirical PDF of  $\{X(t)\}$ . Note that we integrate with  $\theta^*$  as a lower bound in order to enforce that  $Y_0$  is associated with the onset of an above-threshold increment. Let us now consider the calculation of  $P_{A|\tau,Y_0}$ .

In what follows, we simplify the notation by labeling time with the index  $i$ , such that  $i = 0$  represents the time  $T_{\text{on}}^{\text{FP}}$ ,  $i = 1$  the time  $T_{\text{on}}^{\text{FP}} + \Delta t$ ,  $i = 2$  the time  $T_{\text{on}}^{\text{FP}} + 2\Delta t$ , and so forth. With this notation,  $Y_i \equiv Y(T_{\text{on}}^{\text{FP}} + i\Delta t)$  denotes the  $i$ th point following the FP onset time, and  $\Delta Y_i^{\text{diff}} \equiv Y_i - Y_{(i-1)}$  the  $i$ th diffusive increment. Furthermore, our focus here is on

FPs of duration, say  $i\Delta t$ , which corresponds to a sequence of  $i$  successive above-threshold increments. As such, the forthcoming calculations involve PDFs that are implicitly conditioned on the event  $\{\Delta Y_n^{\text{diff}} > \theta^*, \forall n \leq i\}$ .

For an FP of duration  $\tau = i\Delta t$  starting at  $Y_0$ , we define its amplitude as  $A = Y_i - Y_0$ , and we seek the conditional PDF  $P_{A|\tau, Y_0}$ . For this purpose, let  $\rho_i(y) \equiv P_{Y_i|Y_0}(y|y_0)$  denote the PDF of  $Y_i, i > 1$ , conditioned on  $Y_0$ . Since  $A$  is expressed as the difference between  $Y_i$  and  $Y_0$ , we can directly write

$$P_{A|\tau, Y_0}(a|i\Delta t, y_0) = \rho_i(a + y_0). \tag{11}$$

The  $\rho_i$ 's, for  $i > 1$ , are evaluated sequentially based on the fact that  $Y_i = \Delta Y_i^{\text{diff}} + Y_{i-1}$ . The PDF of this sum, conditioned on  $Y_0$ , gives

$$\begin{aligned} \rho_i(y) &= \int_{-\infty}^{\infty} P_{\Delta Y_i, Y_{i-1}|Y_0}(\xi, y - \xi|y_0) d\xi \\ &= \int_{-\infty}^{\infty} P_{\Delta Y_i|Y_{i-1}, Y_0}(\xi|y - \xi, y_0) P_{Y_{i-1}|Y_0}(y - \xi|y_0) d\xi \\ &= \int_{-\infty}^{\infty} P_{\Delta Y_i|Y_{i-1}}(\xi|y - \xi) \rho_{i-1}(y - \xi) d\xi. \end{aligned} \tag{12}$$

To enforce the condition of above-threshold increments,  $\{\Delta Y_n^{\text{diff}} > \theta^*, \forall n \leq i\}$ , we evaluate  $P_{\Delta Y_i|Y_{i-1}}$  based on Eq. (4), but we truncate the distribution below  $\xi = \theta^*$ :

$$P_{\Delta Y_i|Y_{i-1}}(\xi|y) = K \begin{cases} \mathcal{E}_{\Delta Y_i|Y_{i-1}}(\xi|Y(t) = y) & \text{if } \xi > \theta^*, \\ 0 & \text{otherwise,} \end{cases} \tag{13}$$

where  $K$  is a normalization constant. With Eq. (13) and (12), we finalize the calculation of  $P_{A|\tau, Y_0}$  in Eq. (11). In Fig. 3(B), we see the representation of the  $\rho_i$  for an FP triplet. From  $\rho_i$ , we can also calculate the PDF of  $\Delta Y_i^{\text{diff}}$ , conditioned on  $Y_0$  (this will be useful in the calculation of  $P_{\tau|Y_0}$ ). Let this PDF be defined as  $\mathcal{E}_i(\xi) \equiv P_{\Delta Y_i|Y_0}(\xi|y_0)$ . We calculate it as a marginal over  $Y_{i-1}$ :

$$\begin{aligned} \mathcal{E}_i(\xi) &= \int_{-\infty}^{\infty} P_{\Delta Y_i, Y_{i-1}|Y_0}(\xi, y|y_0) dy \\ &= \int_{-\infty}^{\infty} \mathcal{E}_{\Delta Y_i|Y_{i-1}}(\xi|y) \rho_{i-1}(y) dy, \end{aligned} \tag{14}$$

where the dependence of  $\mathcal{E}_{\Delta Y_i|Y_{i-1}}$  on  $Y_0$  disappears because of the Markov property. In Fig. 3(C), we see the representation of the  $\mathcal{E}_i$ 's for an FP triplet.

We now turn to the calculation of  $P_{\tau|Y_0}$ , the probability of observing an FP of duration  $\tau$ , conditioned on the starting value  $Y_0$ . We are interested in the conditional probability of the event  $\{\tau = i\Delta t\}$ , where  $i$  is an integer. This event is equivalent to the intersection of the events  $E_1 \equiv \{\Delta Y_1^{\text{diff}} > \theta^*\}, E_2 \equiv \{\Delta Y_2^{\text{diff}} > \theta^*\}, \dots$  and  $E'_{i+1} \equiv \{\Delta Y_{i+1}^{\text{diff}} \leq \theta^*\}$ . In other words, we obtain an FP of duration  $i\Delta t$  when the first  $i$  increments are above threshold, but the  $(i + 1)$ th increment is below threshold.

By successively applying the definition of conditional probability, we can expand  $P_{\tau|Y_0}$  as follows:

$$\begin{aligned}
 P_{\tau|Y_0}(i\Delta t|y_0) &= \text{Prob}(\tau = i\Delta t|y_0) \\
 &= \text{Prob}\left[E'_{i+1} \cap \left(\bigcap_{n=1}^i E_n\right) \middle| y_0\right] \\
 &= \text{Prob}\left[E'_{i+1} \middle| \left(\bigcap_{n=1}^i E_n\right), y_0\right] \cdot \text{Prob}\left[E_i \middle| \left(\bigcap_{n=1}^{i-1} E_n\right), y_0\right] \\
 &\quad \cdot \text{Prob}\left[E_{i-1} \middle| \left(\bigcap_{n=1}^{i-2} E_n\right), y_0\right] \cdot \dots \cdot \text{Prob}(E_2|E_1, y_0) \cdot \text{Prob}(E_1|y_0).
 \end{aligned}$$

Let  $Z_i(y_0) \equiv \text{Prob}[E_i | (\bigcap_{n=1}^{i-1} E_n), y_0]$ ,  $i > 1$ , represent the probability that the  $i$ th increment is above threshold, given that the  $i - 1$  previous increments were also above threshold, and given the starting value  $y_0$ . These  $Z_i$ 's can be calculated from the  $\mathcal{E}_i$ 's of Eq. (14) as (Fig. 3(B), shaded area):

$$\begin{aligned}
 Z_i(y_0) &= \text{Prob}(\Delta Y_i^{\text{diff}} > \theta^* | \Delta Y_n^{\text{diff}} > \theta^*, \forall n < i; y_0) \\
 &= \int_{\theta^*}^{\infty} \mathcal{E}_i(\xi) d\xi.
 \end{aligned} \tag{15}$$

We now arrive at the desired probability mass function:

$$P_{\tau|Y_0}(i\Delta t|y_0) = \text{Prob}(\tau = i\Delta t|y_0) = (1 - Z_{i+1}(y_0)) \prod_{n=1}^i Z_n(y_0), \tag{16}$$

where we have used the fact that  $1 - Z_{i+1}$  is equal to the probability that the  $(i + 1)$ th increment is below threshold, and where we have defined  $Z_1(y_0) \equiv \text{Prob}(E_1|y_0) = \alpha(y_0)$ , i.e., the probability that the first increment after  $Y_0$  is above threshold. Once Eq. (10), (11), and (16) are evaluated, we apply Eq. (9) to obtain the desired  $Q_A$ . Using this approach, we obtain an excellent agreement between theory and simulations, as reported in Sect. 3.1.

### 2.4.3 True jump rate

Our estimate of the true jump rate  $\lambda$  relies on the knowledge of the overall jump detection probability  $\Gamma_C$  and on the FP detection probability  $\Gamma_A$  (both defined in Sect. 2.1). Recall that we calculate  $\Gamma_A$  from Eq. (6), while we estimate  $\Gamma_C$  directly from the data as  $m/n$ , where  $m$  is the number of time steps with  $\Delta X(t) > \theta^*$  (either from a true jump or an FP) and  $n$  is the total number of time steps in the data time series. On the other hand, from the definition of  $\Gamma_C$  we can write:

$$\begin{aligned}
 \Gamma_C &\equiv \text{Prob}\{\text{detecting an increment larger than } \theta^* \text{ across an interval } \Delta t\} \\
 &= \text{Prob}\{(\text{detecting an FP across } \Delta t) \cup (\text{detecting a true jump across } \Delta t)\} \\
 &= \Gamma_A + \Gamma_B - \Gamma_A \Gamma_B \\
 &= \Gamma_A + \lambda \Delta t - \Gamma_A \lambda \Delta t,
 \end{aligned} \tag{17}$$

where  $\Gamma_B = \lambda \Delta t$  is the probability of observing a true jump in an interval  $\Delta t$ . This is merely a statement of the addition law of probability, which would read: the probability of detecting a jump in an interval  $\Delta t$  is the sum of the probability of observing a true jump, plus that of observing an FP, minus the probability of observing both at the same time, where we use the fact that FPs and true jumps are independent events. For the test cases described in Sect. 3.2, isolating  $\lambda$  in Eq. (17) is accurate up to an error of 0.02%.

#### 2.4.4 True jump amplitude distribution

As in the previous subsection, we obtain an estimate for the true jump amplitude distribution  $Q_B$  based on the empirical PDF of jump amplitudes measured from the time series  $Q_C$  and on the calculated FP amplitude distribution  $Q_A$ . Because detected jumps are a mixture of true jumps and FPs, we can write, a priori,

$$Q_C = W_A Q_A + W_B Q_B, \tag{18}$$

where  $W_A = \frac{\Gamma_A}{\Gamma_C}$  is the probability that a detected jump is an FP, and  $W_B = \frac{\Gamma_B}{\Gamma_C}$  that it is a true jump. The subtlety here is that, contrary to FPs, true jumps are never detected on their own, as they always summate with a diffusive fluctuation. In other words, we never observe the  $B_i$ 's directly, but rather the  $B_i$ 's plus a diffusive increment. Over a short enough time step, diffusive increments are Gaussian variables and are approximately independent of each other. For the purpose of calculating  $Q_B$ , we will thus assume these increments are Gaussian with mean zero and variance  $2D\Delta t$ . Properly accounting for the  $y$ -dependence of the mean would be more precise, but would require  $Q_C$  to be broken down into a family of distributions parameterized by  $y$ , which would require a very large number of detected jumps in the data.

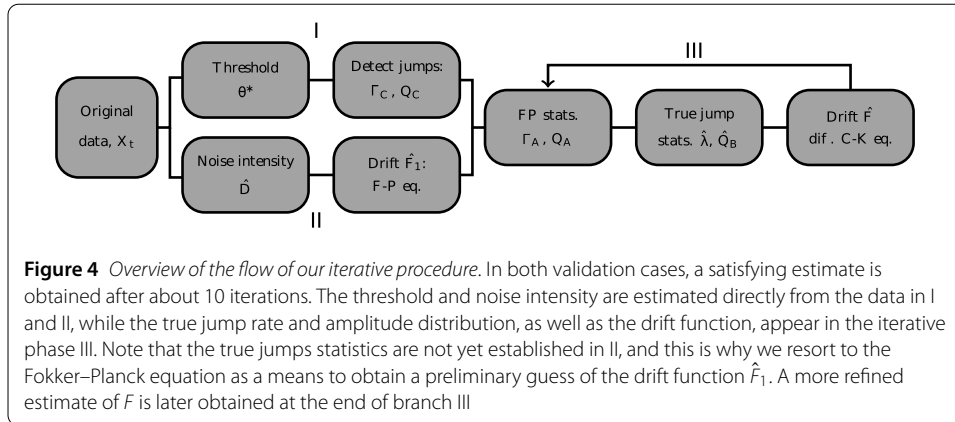
Let  $\tilde{\mathcal{E}}$  represent a Gaussian distribution with zero mean and variance  $2D\Delta t$ . The  $Q_B$  in Eq. (18) should thus be replaced by the convolution  $\tilde{\mathcal{E}} * Q_B$ . Furthermore,  $W_A$  must in fact be reduced by a factor  $(1 - \Gamma_B)$  to account for the probability that FPs can occur during the same interval as a true jump. This leads to

$$Q_C = \frac{\Gamma_A(1 - \Gamma_B)}{\Gamma_C} Q_A + \frac{\Gamma_B}{\Gamma_C} (\tilde{\mathcal{E}} * Q_B). \tag{19}$$

From this equation, we isolate the convolution term and apply the basic deconvolution algorithm [39] to extract  $Q_B$ . Let  $f$  be a measured, convolved signal, where the convolution kernel  $h$  is known. We seek the intact signal  $g$ , such that  $f = (h * g)$ . We compute an estimate of  $g$  at each iteration as  $g_{k+1} = g_k + [f - (h * g_k)]$ , with  $g_0 = f$ . The algorithm converges once the correct signal is reached, since the residual between  $f$  and  $(h * g_k)$  then becomes zero.

### 2.5 Iterative procedure, noise intensity, and drift function

We now turn to the problem of the simultaneous data-driven estimation of all the unknowns in Eq. (1). To this end, we incorporate the calculations of Sect. 2.4 in the iterative scheme depicted in Fig. 4, which consists of three main branches. The first two initial branches, I and II, are independent and are performed only once; this is followed by branch III where the iterations take place. In branch I, the threshold is set (Sect. 2.2) and then applied to the time series to yield the detected jump pool, from which  $\Gamma_C$  and  $Q_C$  are obtained (Sect. 2.3). The noise intensity is estimated in branch II, along with an initial



**Figure 4** Overview of the flow of our iterative procedure. In both validation cases, a satisfying estimate is obtained after about 10 iterations. The threshold and noise intensity are estimated directly from the data in I and II, while the true jump rate and amplitude distribution, as well as the drift function, appear in the iterative phase III. Note that the true jumps statistics are not yet established in II, and this is why we resort to the Fokker–Planck equation as a means to obtain a preliminary guess of the drift function  $\hat{F}_1$ . A more refined estimate of  $F$  is later obtained at the end of branch III

guess for the drift function  $\hat{F}_1$ , which are both used in the first iteration of branch III to calculate  $\hat{\Gamma}_B$  and  $\hat{\lambda}$  (Sect. 2.4). The last step in branch III uses  $\hat{D}$ ,  $\hat{\Gamma}_B$ , and  $\hat{\lambda}$  to estimate the drift function  $\hat{F}$ , which is fed back to the first step of branch III in order to iteratively refine the estimation procedure.

### 2.5.1 Noise intensity

As depicted in Fig. 4, the estimation of  $D$  does not rely on the value chosen for  $\theta^*$ . It does, however, still use the notion of applying a threshold on the increments. Indeed, calculating  $\hat{D}$  relies on partitioning  $\{X(t)\}$  into mostly jump-free segments, the length and number of which is sensitive to the value of the threshold used for detection: the lower the threshold, the more jumps are detected (some of which are FPs) and the shorter these partitions are, which, as explained below, can skew the estimate of  $D$ . A high threshold, on the other hand, leaves a significant number of true jumps in those segments. The goal here is thus to vary the threshold  $\theta$  in order to obtain the optimal estimate of  $D$ .

In the limit of an infinitesimally small sampling interval,  $\Delta t \rightarrow 0$ , the quadratic variation  $[Y^{\text{diff}}(t)]$  of a pure diffusion process converges to the so-called integrated variation, which, for additive and time-independent noise, gives [40, 41]

$$[Y^{\text{diff}}(t)] = \int_0^T 2D ds = 2DT, \tag{20}$$

where  $T = (n - 1)\Delta t$  is the total duration for the  $n$  samples of  $\{X(t)\}$ . We can, therefore, estimate  $D$  via the sample quadratic variation, also known as realized variance,  $RV(t)$  [40, 41]:

$$\hat{D} \approx \frac{1}{2T}RV(t) = \frac{1}{2T} \sum_{k=0}^{n-1} (X^{\text{diff}}(t_{k+1}) - X^{\text{diff}}(t_k))^2. \tag{21}$$

For instance, for a test diffusive process with  $F(y) = -0.2y$  and  $D = 0.15$  and sampled at  $\Delta t = 0.01$  s ( $N = 10^6$ ), Eq. (21) estimates  $D$  with an error of 0.002%. In contrast, applying Eq. (21) to a realization of a jump-diffusion process returns an overestimated  $\hat{D}$ , as expected due to large, non-diffusive, positive increments that populate  $\{X(t)\}$ . Also note that if the threshold is set low enough to detect the smallest true jumps, in general, it can also remove the largest diffusive increments, which are required to properly apply Eq. (21).

Simply removing the detected jumps from the sum in Eq. (21) would, therefore, yield an underestimated  $\hat{D}$ .

To circumvent this problem, we consider only the negative increments of  $\{X(t)\}$  in the calculation of  $\hat{D}$ , as they will remain essentially unaffected by the presence of positive jumps, with the exception of a short transient following the jump offset. Indeed, following each jump, we expect to see a brief period where the process is out of equilibrium. And since the jumps have positive amplitudes, negative increment statistics are biased toward negative values during this transient (e.g., Fig. 1). The calculation of  $\hat{D}$  is thus based on applying Eq. (21) to jump-free segments of  $\{X(t)\}$ , but only including negative increments, and neglecting the initial transient at the start of each segment (the duration of which is determined below). This is repeated for various values of  $\theta$ . The successful estimation of  $D$  based only on negative increments relies on our assumption that  $\Delta t$  is small, for in this case increments are approximately independent and distributed as  $\mathcal{N}(0, 2D\Delta t)$ . For larger  $\Delta t$ , the increment PDF can become asymmetric, meaning that the statistics of negative increments differ from those of positive ones, which would cause errors in our estimation of  $D$ .

Let  $T_{\text{off}}$  and  $T_{\text{on}}$  denote the jump offset and onset times, respectively. Note that the values of these times and the number of detected jumps all depend on the specific value of the threshold. Then the  $i$ th segment is defined by  $\{S(t)\}_i = \{X(t) : T_{\text{off}}(i) < t < T_{\text{on}}(i+1)\}$  and is of duration  $T_i = T_{\text{on}}(i+1) - T_{\text{off}}(i)$ , and let  $\{\Delta S(t)\}_i$  be its  $n_i$  increments. Out of these  $n_i$  increments, we keep only the  $n_i^-$  that are negative and that occur after the transient of approximate duration  $\Phi$ . We are thus left with the following subset of increments from each segment:

$$\{\Delta S(t)\}_i^- = \{\{\Delta S(t)\}_i : \{\Delta S(t)\}_i < 0, t > \Phi\}. \tag{22}$$

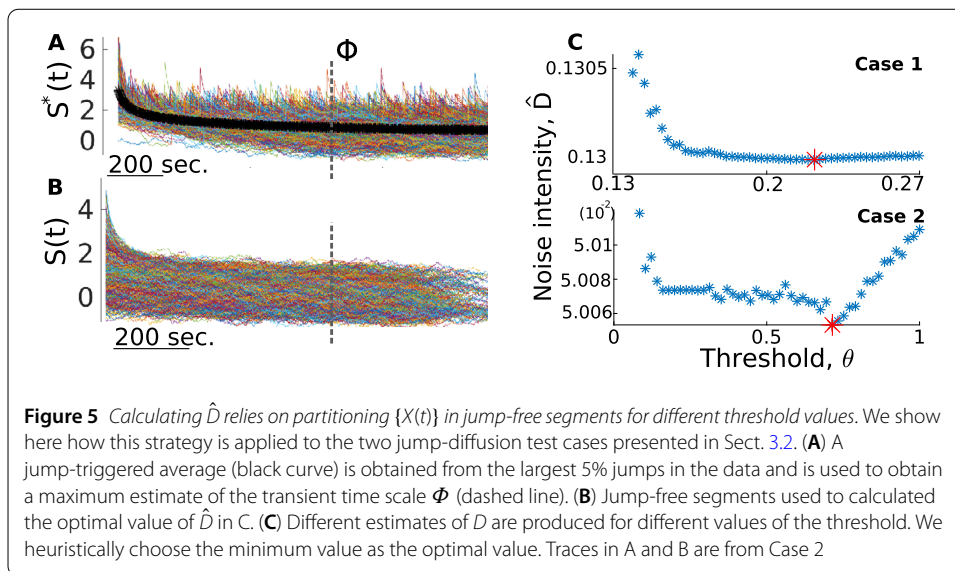
For each segment, we obtain an estimate  $\hat{D}_i$  as follows:

$$\hat{D}_i = \frac{1}{2T_i^-} \sum_{k=1}^{n_i^-} (\{\Delta S(t_k)\}_i^-)^2, \tag{23}$$

where  $T_i^- = n_i^- \Delta t$  is the effective duration of the combined  $n_i^-$  negative increments. We then calculate  $\hat{D}$  as an average of the  $\hat{D}_i$ 's, weighted by  $T_i/T$ .

More precisely, here are the steps taken in order to arrive at  $\hat{D}$ :

- Starting from the largest value of  $\{\Delta X(t)\}$ , lower the threshold until the largest 5% of jumps are detected, which are the ones with the most prominent transient.
- Let  $\{S^*(t)\}_i$  be the segments that follow these jumps (Fig. 5(A)). Average across them for each time step, creating a jump-triggered average trace (Fig. 5(A)), black line).
- Identify  $\Phi$  as the approximate moment when the jump-triggered average stabilizes, quantified here as when its derivative is less than 0.05 (results do not depend strongly on this particular value: changing  $\Phi$  by one order of magnitude on either side of the value used here yields estimates of  $D$  that differ by less than 0.2%). This gives us an estimate for the maximum time scale required for post-jump equilibrium.
- With  $\Phi$  determined, and for each value of the threshold, extract the  $\{S(t)\}_i$ 's (Fig. 5(B)) and apply Eq. (23) to their negative increments for which  $t > \Phi$  (neglecting segments for which  $T_i < \Phi$ ).



This scheme results in an estimate  $\hat{D}$  for each value of the threshold, and the lowest value is chosen as the optimal estimate (Fig. 5(C)). This is because this approach overestimates  $D$  on both ends of the range of threshold values, but for different reasons. For low  $\theta$ , many more jumps are detected, and this makes the  $\{S(t)\}_i$ 's shorter, which means that the associated  $\hat{D}$  is more prone to be biased by the residual of the transient. On the other hand, a very high threshold leaves significant jumps in the  $\{S(t)\}_i$ 's, which biases the statistics of negative increments. The optimal balance is reached somewhere in-between, where the segments are large enough so that the initial transient is negligible, and where the jumps that are inevitably left in the segments do not significantly alter the statistics of negative increments. The heuristic choice of an intermediate value, namely one that corresponds to the minimal  $\hat{D}$  estimate, gives excellent results in both validation cases (less than 0.1% error, see Table 2).

### 2.5.2 Drift function

Our estimation of the drift function  $F$  relies on the differential Chapman–Kolmogorov equation [42], which describes the evolution of the transition probability of a stochastic process where jumps occur alongside diffusive fluctuations. Let  $Y(t)$  be a jump-diffusion process with transition probability  $P_{Y|Y_0}$ . For the case of positive Poisson jumps and additive diffusive noise, the differential Chapman–Kolmogorov equation reduces to (see the Appendix)

$$\frac{\partial P_{Y|Y_0}(y, t|y_0, t_0)}{\partial t} = -\frac{\partial}{\partial y} [F(y)P_{Y|Y_0}(y, t|y_0, t_0)] + D \frac{\partial^2}{\partial y^2} P_{Y|Y_0}(y, t|y_0, t_0) - \lambda P_{Y|Y_0}(y, t|y_0, t_0) + \lambda \int_0^\infty Q_B(s)P_{Y|Y_0}(y-s, t|y_0, t_0) ds. \quad (24)$$

If  $Y(t)$  is assumed to have reached its equilibrium state, then the left-hand side vanishes, and in the right-hand side we can replace the transition probability with the first-order equilibrium PDF,  $P_Y$ ,<sup>b</sup> which we assume to be equal to the empirical PDF,  $P_X$ , of the measured time series  $\{X(t)\}$ . The drift function can then be evaluated from Eq. (24) if  $D$ ,  $\lambda$ , and  $Q_B$  are known (or estimated).

Note, however, that  $\hat{F}$  is required in the first step of branch III of the iterative procedure (Fig. 4), since the FP-related statistics,  $\Gamma_A$  and  $Q_A$ , are calculated based on the drift function. A preliminary estimate  $\hat{F}_1$  of the drift function is thus required. This particular estimate, which is needed only once throughout the inference procedure, is obtained by letting  $\lambda = 0$  in Eq. (24), such that it becomes the Fokker–Planck equation associated with the diffusive part of the stochastic process. The stationary solution of this Fokker–Planck equation can be used to establish a relation between the noise intensity, the drift function, and  $P_Y$  [37, 43]:

$$P_Y(y) = \frac{K}{\hat{D}} \exp\left(-\int \frac{\hat{F}_1(y)}{\hat{D}} dy\right), \quad (25)$$

where  $K$  is a normalization constant and where, again, we assume that  $P_Y = P_X$ . This first preliminary estimate is necessarily flatter than the true  $F$ , as the presence of jumps makes  $P_X$  wider than it would be if there were no jumps. Successive iterations gradually rectify this by incorporating estimates of  $\lambda$  and  $Q_B$  in Eq. (24).

### 3 Results

Here we present three applications of the method developed above. First, we validate the calculation of  $\Gamma_A$  and  $Q_A$  for the case of a purely diffusive process. Then we apply the full iterative scheme to two simulated jump-diffusion processes with different characteristics. Finally, we apply our inference method to electrophysiological recordings in pyramidal cells of electric fish.

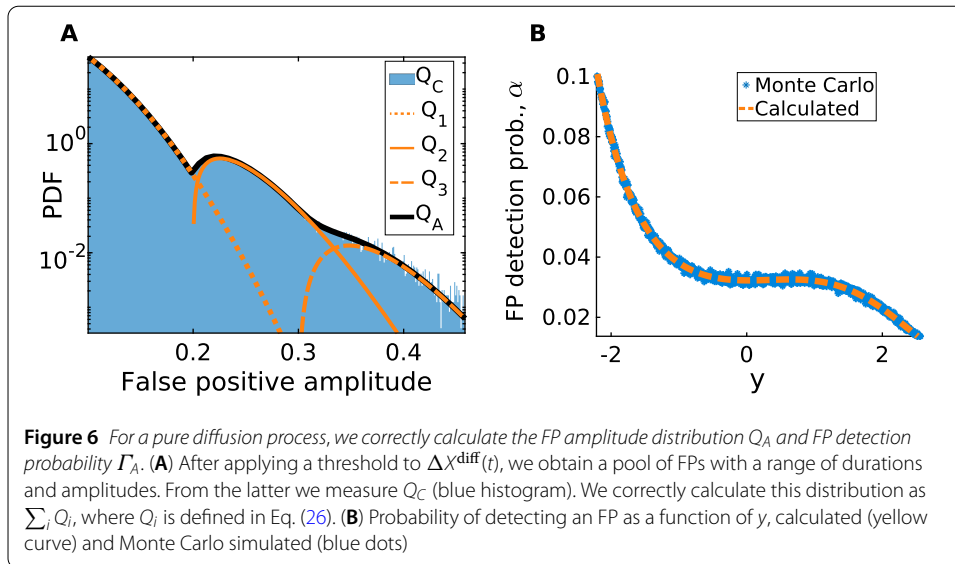
#### 3.1 Validation of the FP statistics calculations

To confirm that the calculations of  $Q_A$  and  $\Gamma_A$  are accurate, we start with a simple test case where we consider a time series  $\{X^{\text{diff}}(t)\}$  obtained from a simulated pure diffusion process. As there are no jumps here, the distribution of increments does not possess the necessary asymmetry to properly identify a threshold. For this test case only, we thus opt for a specific value,  $\theta^* = 0.1$ , that showcases the ability of our method to handle FPs of various durations. The results presented here, however, remain valid for a range of values of  $\theta^*$ . For the parameters used in this pure diffusion validation case (Table 1), this range extends from 0.025 up to 0.2. The upper limit is set by the fact that, beyond it, too few FPs are detected, which precludes any statistical calculations from being achieved. The lower limit, on the other hand, arises because too many FPs are detected, such that, for instance, they occur every other time step. In such a case,  $\Delta t$  is too large and the estimation of  $\lambda$  becomes imprecise due to the statistical fluctuations in the number of detected FPs.

Applying the threshold in this case leads to a detected jump pool comprised entirely of FPs. The goal now is to compare the measured  $Q_C$  and  $\Gamma_C$  with the calculated  $Q_A$  and  $\Gamma_A$ . If we obtain that  $\Gamma_C \approx \Gamma_A$  and that  $Q_C \approx Q_A$ , then we will effectively have shown that the true jump rate is zero,  $\lambda = 0$ , and that our method correctly calculates the FP amplitude distribution. In this pure diffusion test case, these calculations rely on the knowledge of the correct noise intensity  $D$  and the correct drift function  $F$ , but this will not be the case in subsequent sections.

With the particular values of  $D$  and  $F$  used here to simulate  $\{X^{\text{diff}}(t)\}$  (Table 1), we find that FPs are either singlets, doublets, or triplets, which contribute differently to the measured amplitude distribution  $Q_C$ . Indeed, the fact that longer FPs tend to have larger amplitudes and that FP durations are always multiples of  $\Delta t$  creates distinct modes in the





measured amplitude distribution (Fig. 6(A), blue histogram). By taking the sum out of the integral in Eq. (9), we can write  $Q_A(a) = \sum_{i=1}^{\infty} Q_i(a)$ , where

$$Q_i(a) = \int_{-\infty}^{\infty} P_{A|\tau, Y_0}(a|i\Delta t, y)P_{\tau|Y_0}(i\Delta t|y)P_{Y_0}(y) dy \tag{26}$$

are the individual distributions associated with FPs of duration  $i\Delta t$  (Fig. 6(A), yellow curves). These distributions are then summed to obtain  $Q_A$ , which is a precise match with  $Q_C$  for this purely diffusive example (Fig. 6(A), black curve).

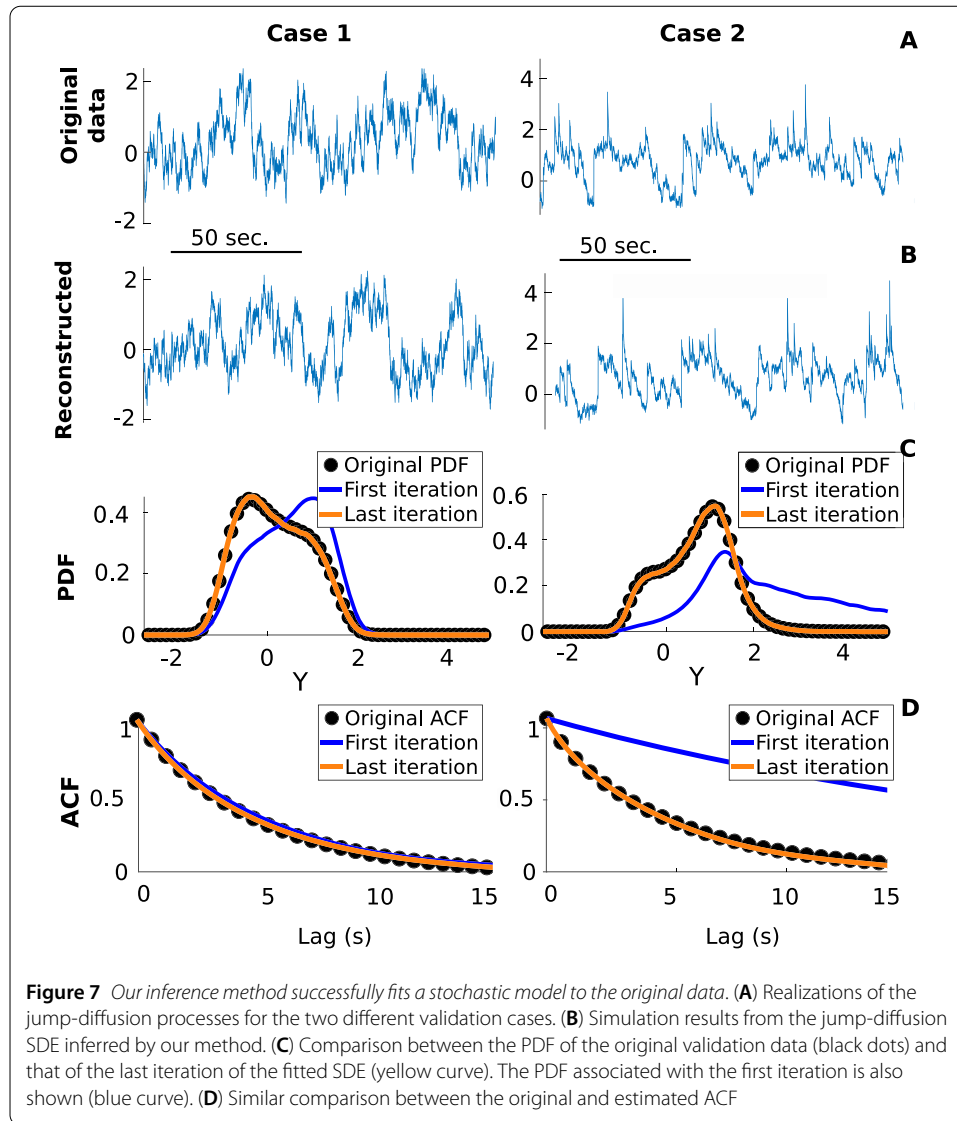
Furthermore, by applying Eq. (5) we calculate the  $y$ -dependent detection probability,  $\alpha(y)$  (Fig. 6(B), yellow curve). As expected, this function depends non-trivially on  $y$  and reflects the nonlinearity of the specific drift function used in this example. If multiplicative noise had been used, the noise intensity  $D(y)$  would also have influenced the shape of  $\alpha(y)$ . To validate this calculation of  $\alpha(y)$ , we run Monte Carlo simulations of the diffusion process

$$dY^{\text{diff}}(t) = F(Y^{\text{diff}}(t)) dt + \sqrt{2D}dW(t), \tag{27}$$

over a duration  $\Delta t$ , but with a time step of  $\Delta t/1000$  and with various initial conditions along the  $y$ -axis. For each initial condition, we evaluate the FP detection probability as the ratio between the number of Monte Carlo runs, where  $Y^{\text{diff}}(\Delta t) > Y^{\text{diff}}(0) + \theta^*$ , and the total number of Monte Carlo runs (Fig. 6(B), blue dots), the result of which precisely fits with the calculated  $\alpha(y)$ . Finally, we obtain the overall FP detection probability  $\Gamma_A$  from Eq. (6), which, in this pure diffusion test case, differs from  $\Gamma_C$  by only 0.06%.

### 3.2 Validation of the iterative scheme

Before applying the iterative procedure (Fig. 4) to real data, we first validate it against time series generated by numerically integrating Eq. (1) (using the Euler–Maruyama scheme). We consider two validation cases: in Case 1 the amplitude of the jumps is comparable to the diffusive fluctuations and jumps are sparser in time (i.e., occur at a lower rate) than in



Case 2, where jumps are much larger than the background noise and their rate is double that of Case 1 (Fig. 7(A)). The specific functions and parameters used to generate and analyze these validation data are shown in Table 1, which can be summarized as follows: low rate, low amplitude, high noise for Case 1, and high rate, high amplitude, low noise for Case 2. Preliminary tests with a linear drift function showed a successful fit between the fitted SDE and the numerical data. We now opt for a more general and arbitrary shape where the drift function is nonlinear and non-monotonic. The only restrictions are that it yields a single stable fixed point and that the resulting stochastic process is stationary. We thus restrict our analyses to drift functions that are mostly decreasing. Although the parameters and functions used for the simulations are known, they are not used in the inference procedure, only  $\{X(t)\}$  is. To assess the performance of the proposed method, we compare the estimated  $\hat{D}$ ,  $\hat{\lambda}$ ,  $\hat{F}$ , and  $\hat{Q}_B$  with their true values.

We choose these two specific cases because they challenge both the sensitivity and the robustness of our method. In Case 1, the statistics of  $\{X(t)\}$  are not too different from those of a pure diffusion process, which makes it easier to estimate  $F$ , but difficult to extract  $Q_B$

**Table 1** Parameters and functions for the validation cases

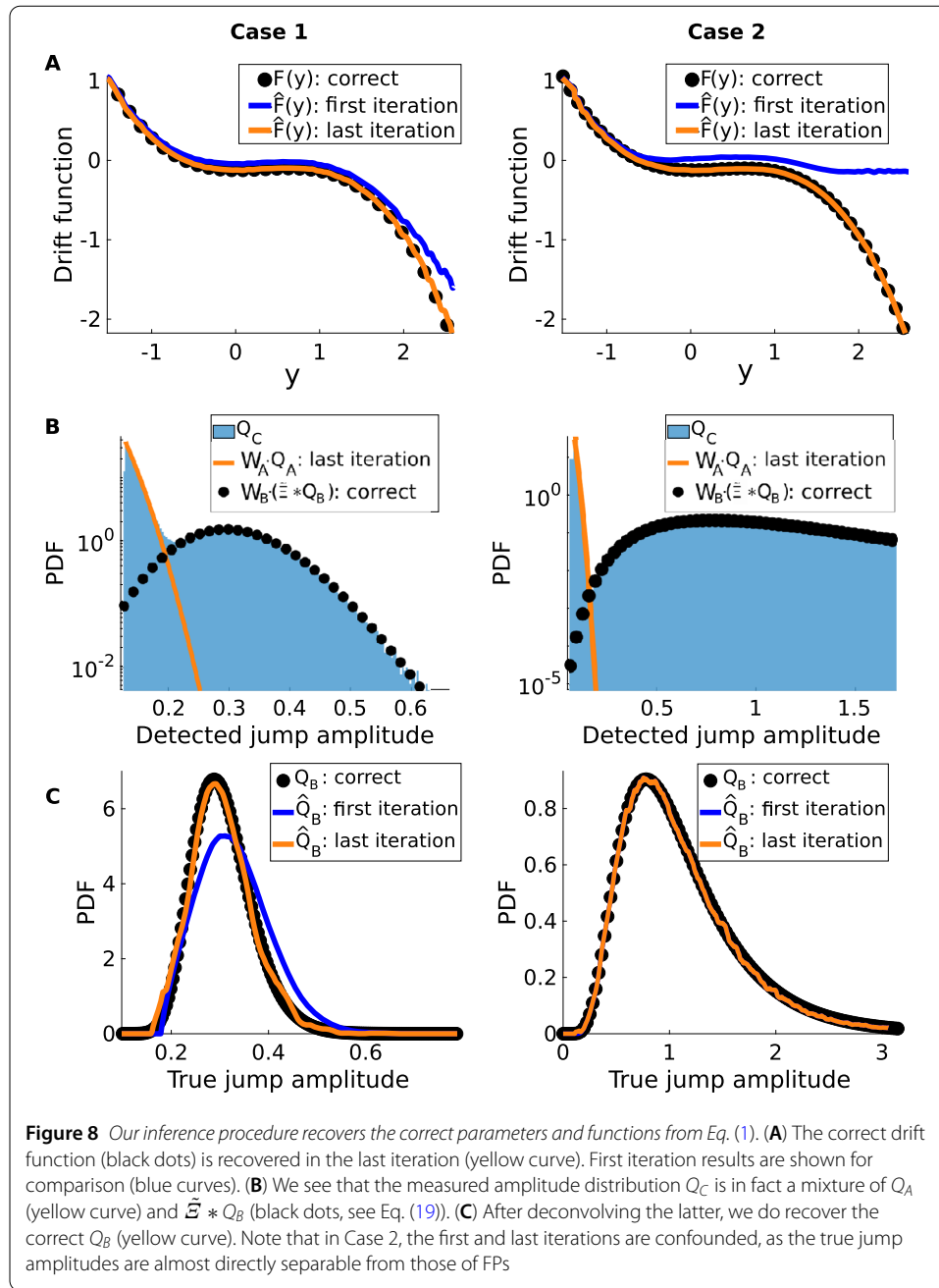
	Case 1	Case 2	Pure diffusion
$\lambda$	0.1	0.2	0
$D$	0.13	0.05	0.15
$\theta^*$	0.125	0.07	0.1
$\mu$	-1.2	1	N/A
$\sigma$	0.2	0.5	N/A
$Q_B$	Lognormal( $\mu, \sigma^2$ )		N/A
$a$	0.2	0.2	0.2
$F$	$-(a(y - 0.5)^3 + 0.5a(y - 0.7)^2 + 0.1)$		
$\Delta t$ (s)	$10^{-2}$	$10^{-2}$	$10^{-2}$

amidst the FPs. In contrast, the jumps in Case 2 are well separated from the diffusive fluctuations, which allows for a more direct estimation of  $Q_B$ . The presence of large jumps in this case, however, significantly alters  $\{X(t)\}$  and its PDF, making it harder to estimate  $F$ . In both cases, however, we find that the original SDEs can be precisely recovered by our method. For instance, simulating Eq. (1) with  $\hat{D}$ ,  $\hat{\lambda}$ ,  $\hat{F}$ , and  $\hat{Q}_B$  of the last iteration not only produces time series that resemble the originals (Figs. 7(A) and 7(B)), but also yields an excellent fit between the reconstructed and original PDFs (Fig. 7(C)), with an  $O(10^{-4})$  root-mean-square error (normalized by the range of  $\{X(t)\}$  and ACFs (Fig. 7(D)).

Inspecting the results from the last iteration, we indeed see that the correct drift function is recovered (Fig. 8(A)). This is done through the use of the so-called differential Chapman–Kolmogorov equation. This is then used to calculate the next  $Q_A$  and  $\Gamma_A$  (Fig. 8(B)), which allows the correct  $Q_B$  to be demixed from the measured  $Q_C$  (Fig. 8(C)). We also obtain low relative errors when comparing our estimates and the correct values of  $D$  and  $\lambda$  (Table 2). Note also that, although not shown here, we obtain the same fit quality between original and reconstructed when we apply our method to hybrid cases, where, for instance, the small jump amplitudes are paired with a higher rate instead of a lower one and vice-versa.

### 3.3 Application to experimental data

We now proceed with an application of our inference method to electrophysiological data published in Ref. [44]. These data consist of in vitro, intracellular recordings of membrane voltage fluctuations in pyramidal neurons of the weakly electric fish *Apteronotus leptorhynchus*. These fish are endowed with an active sensing mechanism whereby they generate a high frequency ( $\sim 700$  to 1000 Hz) oscillatory electric field around their body. This electric signal, along with any distortions caused by objects, preys, and conspecifics, is sensed by electroreceptors located on the fish's body. This information is then sent to the hindbrain, where it reaches the first stage of electrosensory processing, called the electrosensory lateral line lobe (ELL). The recordings in Ref. [44] were taken from neurons of the ELL, specifically in the centrolateral and centromedial segments (CLS and CMS, respectively). In order to isolate the impact of voltage-gated ion channels on membrane potential fluctuations, the ELL was treated with pharmacological agents (CNQX and APV) that block synaptic transmission onto the pyramidal cells. The resulting fluctuations are thus fully attributed to cell-intrinsic sources, which we refer to as membrane noise. The main source of this type of noise is often assumed to be the stochastic opening and closing of ion channels, i.e., channel noise [45]. In this case, we cannot rule out other potential contributions, as non-trivial soma-dendrite interactions have been observed in these cells

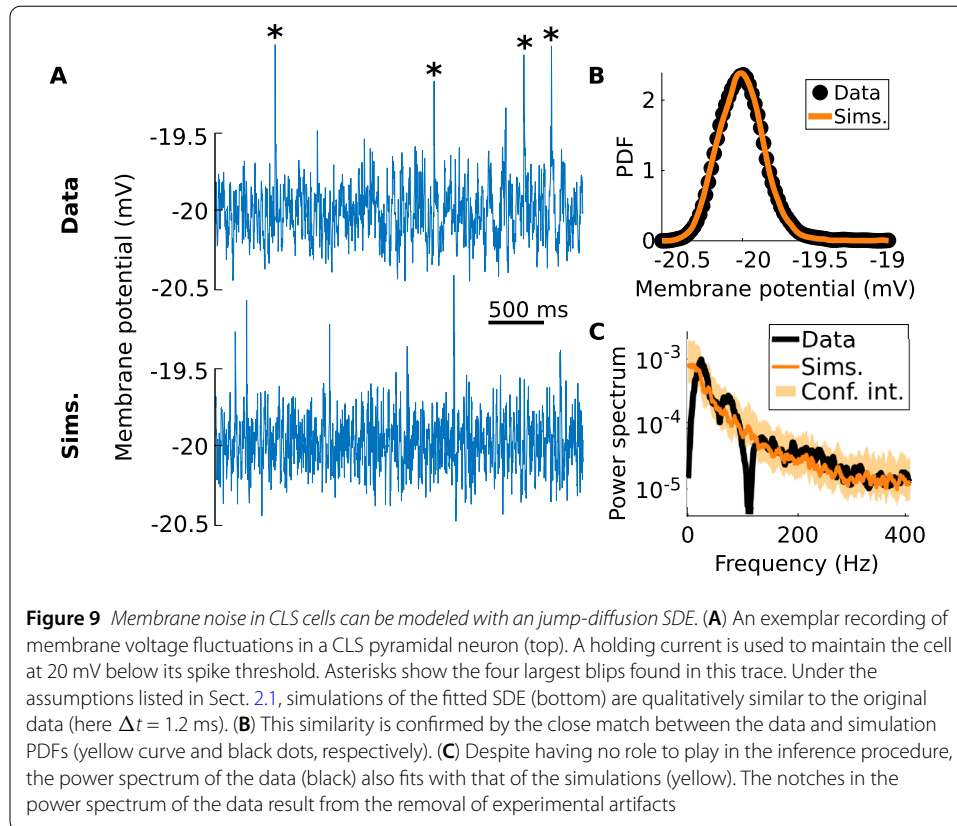


**Table 2** Comparison between estimated and correct parameters

	Case 1			Case 2			
	Correct	Estimated	Relative error (%)	Correct	Estimated	Relative error (%)	
D	0.13000	0.13003	0.023	D	0.05000	0.05005	0.1
$\lambda$	0.1000	0.0978	2.2	$\lambda$	0.2000	0.1960	2.13

[46]. Note also that, by imposing different holding currents on the cells, Ref. [44] recorded ongoing membrane noise at various levels of hyperpolarization relative to spike threshold.

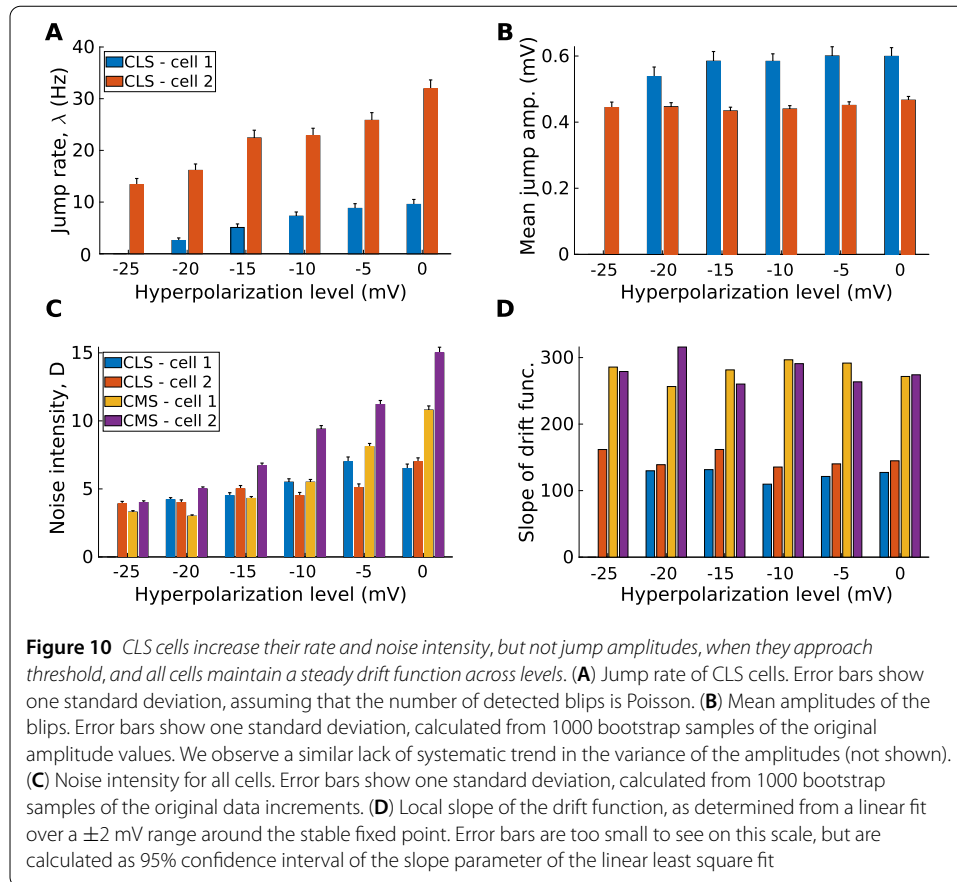
Of interest here is the presence of large, jump-like events, called blips, that abruptly depolarize the cells (Fig. 9(A), asterisks). Although Ref. [44] puts forth a hypothesis as to the



functional role for these blips, the mechanism underlying their occurrence is unknown. This, along with the limited amount of data, hinders the development of any meaningful mechanistic model of this phenomenon. The jump-diffusion inference approach developed here, however, is particularly well suited to circumvent this knowledge gap. Indeed, the resulting phenomenological models provide a useful tool for dynamically interpreting the available data without relying on poorly constrained biophysical mechanisms. For instance, we can address questions such as: Do certain parameters or functions of the model change as a function of the mean membrane potential?

Here we analyze recordings from two CMS and two CLS cells, each with five or six levels of hyperpolarization: from  $-25$  to  $0$  mV below threshold, with  $5$  mV steps between levels. Using these relative levels with respect to spike threshold is required to compare cells that might have different thresholds (e.g.,  $-67$  to  $-63$  mV for CMS cells [44]). After removing experimental artifacts (see Sect. 3.4), we obtain a total of 23 traces, each lasting approximately 10 s. Applying our inference method to these traces yields a good fit between the resulting simulations and the original data (Fig. 9(A)): the PDFs differ only by  $O(10^{-2})$  normalized root-mean-square errors, and the power spectra fall within 95% confidence intervals of each other (Figs. 9(B) and 9(C)).

Further insight can be gained by comparing the estimated SDE parameters and functions  $\hat{D}$ ,  $\hat{\lambda}$ ,  $\hat{Q}_B$ , and  $\hat{F}$  across all traces. We thus see, for instance, that CLS cells increase their jump rate (Fig. 10(A)), but not jump amplitudes (Fig. 10(B)), as they approach threshold. Note also that we could not measure any significant jump component for CMS cells. Instead, fluctuations in these cells are well described by pure diffusion. Furthermore, all cells show an increase in their diffusive noise intensity with depolarization, and this is



more prominent in CMS cells (Fig. 10(C)). Lastly, to compare the different drift functions with a scalar measure, we apply a linear fit to  $F$  (estimated as in the previous section) in the vicinity ( $\pm 0.2$  mV) of the stable fixed point. The slope parameter resulting from this fit can be interpreted as a measure of how wide or narrow the potential function is around the resting membrane voltage. Using this measure, we find no systematic intra-cell trend, but we do observe large differences between cell types: CLS cells have a wider potential function than CMS ones (Fig. 10(D)).

### 3.4 Data processing

For each cell, the raw data consist of a continuous, 60 to 70 second staircase-like trace, sampled at 20 kHz. Each step lasts  $\sim 10$  s and corresponds to a different holding current, which was applied such as to create 5 mV hyperpolarization from the previous level. In order to segment the recordings into different traces for each level, we first identify the transition times between different holding currents. This is done visually, as the transitions are unambiguous, and we omit  $\pm 0.5$  seconds around those times. At the  $-5$  and  $0$  mV level, a few spikes (1 to 4) occur in the recordings. They are manually removed from the traces along with the ensuing refractory period.

A conspicuous aspect of the resulting traces is the presence of slow, large amplitude (1 Hz,  $\sim 1$  mV) quasi-oscillations overlaid with the faster, stochastic fluctuations. The exact source of this slow component is unknown, but is possibly related to persistent sodium channels, which have been shown to populate the soma and proximal apical dendrites of

ELL cells [44] and to produce slow perithreshold oscillations in entorhinal stellate neurons [47]. In any case, these slow oscillations are outside the scope of the method presented here. A moving average filter (0.05 s window size) is thus applied to remove this low frequency content from the signal.

Line noise is removed at all 60 Hz multiples with a notch filter, but the data are also contaminated with artifacts in other frequency bands, potentially from interference with other sources. This is most prominent in the 900–3000 Hz, but also carries around lower frequencies, e.g., 100, 270, and 550 Hz. To account for this artifact, we opt for the combination of a low-pass filter with a 900 Hz cut-off, and 20 Hz wide band-stop filters centered on the other problematic frequencies. Electrophysiological recordings can be dominated by measurement noise at high frequencies [48]. In our case this is seen as a flattening out of the PSD above 1000 Hz, so the 900 Hz cut-off used here does not lead to the loss of important biological signals.

The end result of this processing chain are time series that exhibit fluctuations typical of jump-diffusion processes. We do observe, however, significant higher-order correlations on the smallest timescales ( $O(\Delta t)$ ,  $\Delta t = 50 \mu s$ ). To quantify these correlations, we use the notion of the Einstein–Markov timescale [5]. This is a measure of the timescale below which the Markov property no longer holds. Stochastic time series often show a departure from the Markov property on small timescales, possibly due to noise source correlation, the presence of an inertial component in the dynamics, or measurement noise [5]. Following [5] and [49], we estimate this Einstein–Markov timescale by finding the value of  $\tau$  that minimizes

$$\chi^2 = \iiint \frac{[P(x_1, x_2, x_3) - P(x_3|x_2)P(x_2, x_1)]^2}{\sigma^2} dx_1 dx_2 dx_3, \quad (28)$$

where  $x_1 = x(t)$ ,  $x_2 = x(t + \tau)$ ,  $x_3 = x(t + 2\tau)$ , and  $\sigma^2$  is the sum of the traces of the covariance matrices associated with the distributions in the numerator. For a proper Markov process,  $\chi^2 = 0$ ,  $\forall \tau$ . In this case, we find the minimum of  $\chi^2$  at 1.2 ms, indicating that the Einstein–Markov timescale of the data is over one order of magnitude larger than the sampling interval  $\Delta t$ . This means that, on the time scale of individual observations, the data evolve with a history dependence that is incompatible with a Markovian description. If, however, we look at the data on a coarser time scale, e.g., the Markov–Einstein time scale of 1.2 ms, then the Markov property is approximately satisfied. In that case only can we hope to use Eq. (1) as a valid model for these data. To account for this problem, we resample the data at a 1.2 ms interval ( $\sim 830$  Hz sampling rate) and obtain the final time series on which to apply our method (Fig. 9(A), top), with the time step equal to the Markov–Einstein time scale. Note that this situation is conceptually similar to how the Langevin model of diffusion (where the position of a particle is, by itself, not Markovian) reduces to the Einstein model (where the position is Markovian) only above a certain time scale [43].

#### 4 Discussion

In this study, we develop an iterative procedure that recovers the parameters and functions of a jump-diffusion SDE, based solely on a realization of the associated stochastic process. This approach is validated when the jumps are comparable in size to the diffusive fluctuations (Case 1), as well as when they are much larger than diffusive fluctuations (Case 2). We apply this method to membrane voltage fluctuations recorded in pyramidal

neurons of electric fish. Our analysis reveals that these data can indeed be represented as jump-diffusion processes. We find that pyramidal neurons increase their jump rate and noise intensity as they approach spike threshold, while their jump amplitudes and drift function remain unchanged.

Our approach relies on five main components: the use of the differential Chapman–Kolmogorov equation to estimate the drift function, the use of quadratic variation on jump-free segments to estimate the diffusive noise intensity, the detection of jumps via threshold-crossing of the increments, the modeling of detected jumps as a mixture of true jumps and FPs, and the calculation of FP statistics used to extract true jump statistics from the detected jump pool.

Although we estimate the drift function and the true jump amplitude distribution non-parametrically, we do limit our study to the case of additive diffusive noise, of constant jump rate, and of Poisson jumps. Relaxing the additive noise assumption would require an estimation scheme for the diffusion function  $D(y)$ . For purely diffusive processes, this function can be obtained directly through the estimation of the second Kramers–Moyal coefficient, which is defined in terms of the second conditional moment of the increments. Evaluating this moment simply requires the knowledge of the conditional PDF across time steps. For a Poisson jump-diffusion process, however, Ref. [31] has shown that the diffusion function can in fact be expressed in terms of the second conditional moment of the increments, the jump rate, and the second moment of the jump amplitudes. It should thus be possible to include the estimation of  $D(y)$  into the iterative portion of our method (Fig. 4). Indeed, estimates of the jump rate and of the amplitude distributions could be used at each pass to estimate the diffusion function. Furthermore, we have limited our analysis to noise intensities for which jump amplitudes are on average an order of magnitude or more larger than diffusive fluctuations. When diffusive fluctuations and jumps are of similar average magnitude, the number of detected FPs becomes too large and estimates of  $\lambda$  and  $Q_A$  become imprecise due to increased statistical fluctuations. A much finer temporal resolution would be necessary to address this particular case.

As for the assumption of constant jump rate, it should be possible to extract a rate function  $\lambda(y)$  as long as a  $y$ -dependent version of Eq. (17) can be written. This would require a long enough data time series such as to produce an estimate of  $\Gamma_C(y)$ . Relaxing the assumption of Poisson jumps, however, would be more difficult to do. The detection probability of true jumps,  $\Gamma_B = \lambda \Delta t$ , would obviously need to be modified with the appropriate expression. Moreover, the specific form of the differential Chapman–Kolmogorov used here, Eq. (24), relies on the assumption of Poisson jumps (see the Appendix) and would thus need to be extended in a manner that depends on the precise non-Poissonian nature of the jump process. More specifically, the last two terms in Eq. (24), which are originally defined based on the transition rates of the Poisson jump process, would now be derived from the modified  $\Gamma_B$ . Note that, for the special case of true jumps with zero-mean amplitudes, the drift function can be estimated directly from the first conditional moment of the increments, without relying on Eq. (24) [31].

#### 4.1 Membrane noise

The unusual characteristics of membrane noise observed in CLS neuron, initially reported in Ref. [44] and represented here as jump-diffusion SDEs, might be implicated in novel ways in electrosensory processing. The analysis we perform here is a first step toward investigating this possibility computationally.



The positive impact of noise on information processing in neural systems is widely recognized [50]. Although channel noise was initially thought to be too weak compared to synaptic noise to influence a neuron's output statistics [51, 52], it has since been shown to significantly impact neuronal reliability and action potential timing [47, 53–57]. Because it arises from the stochastic opening of voltage-gated ion channels, channel noise has been successfully modeled by populations of Markov chains with voltage-dependent transition rates [45]. In the quest for more computationally efficient models, however, various approximations have been used to model the collective behavior of these Markov chains as simple diffusion SDEs [57–60] (sometimes called the diffusion approximation for channel noise [61, 62], in reference to the diffusion approximation<sup>c</sup> for synaptic bombardment). One of these approaches, for instance, introduces a current noise term directly in the membrane equation [57, 63], effectively modeling the subthreshold voltage locally as an Ornstein–Uhlenbeck process. It is perhaps not surprising then that we obtain a good match between the observed CMS membrane noise (at various holding potential) and a pure diffusion SDE. In cases where membrane noise is more accurately described by multiplicative conductance noise [57], however, we might expect deviation from the simple SDE used here, similar as to how the diffusion approximation can misrepresent the subthreshold voltage distribution for certain types of synaptic drive [64].

We cannot completely exclude the possibility that small, hard to detect blips occur in CMS cells. Although our method proves capable of handling this type of situation (Sect. 3.2), the limited amount of available experimental data in this case precludes us from conclusively ruling out the existence of a jump component in models of CMS membrane noise. In addition, in both types of cells we find a positive correlation between the noise intensity and the holding potential (Fig. 10). This is consistent with how membrane potential variance has been observed to increase with depolarization in these same cells [44], as well as in rat neocortical pyramidal neurons [65]. Simple Markov models involving only  $\text{Na}^+$  and  $\text{K}^+$  channels are able to reproduce this correlation [66].

The presence of blips in CLS neurons suggests that ion channels co-activate to produce abrupt depolarizing currents. It was shown in [67] that sodium channels appear in clusters, or hot spots, in ELL pyramidal cells. This might allow local depolarization of the membrane to sufficiently couple channels within a cluster. Alternatively, perhaps channels are physically and functionally coupled through scaffolding protein complexes, as observed in [68]. Regardless of the exact coupling mechanism, the commonly used assumption of independence between channels [45, 57] is likely violated by these blips. The fact that we have successfully fitted a jump-diffusion SDE to CLS membrane noise suggests that a diffusion-like approximation could be applied in this case as well. Such a deductive approach would, however, require tentative descriptions of local channel coupling to be included in the kinetic schemes.

Given the unknown biophysical mechanism underlying the blips, our fitted jump-diffusion model is uniquely positioned to address questions related to their functional role. Future work will thus aim to incorporate our fitted jump-diffusion model as a membrane noise term in a more complete model of CLS cells [69]. By accounting for the synaptic input associated with electrosensory input, the resulting model could investigate the possibility that blips assist or influence spiking, perhaps through a stochastic resonance-like phenomenon. Stochastic resonance, and more generally stochastic facilitation, has been shown to be mediated by channel noise in models of auditory brain stem neurons

[51] and in modeled neuronal arrays [70], as well as to be mediated by synaptic noise in models of neocortical pyramidal neurons [71]. Since blips share a similar shape and amplitude as AMPA-driven excitatory post-synaptic potential, this begs the question of whether stochastic resonance is at play in the detection of weak electrosensory signals, such as small prey. This was indeed hypothesized, but not explicitly shown, in Ref. [44]: the voltage-dependence of membrane noise makes it impossible to vary the noise level, the rate of occurrence, or the amplitude of blips independently of the membrane potential. The value of this hypothesis, however, could be assessed by performing a numerical experiment with our fitted SDE representation of blip-laden membrane noise.

Lastly, although we apply our method here to recordings where synaptic input is completely blocked, it might be applicable to certain type of synaptic input patterns, such as correlated bombardment. We also hope to apply this method to fluctuations of the active-sensing rate of pulse-type electric fish, where jump-like events also occur [72].

### Appendix

The particular form of the differential Chapman–Kolmogorov equation that we use here, Eq. (24), relies on the assumption of Poisson jumps. For completeness, and to make explicit the dependence of this equation on Poisson statistics, we provide the following short derivation:

Let  $X(t)$  be a continuous-time Markov process, possibly with discontinuous sample paths, in one dimension and with transition probability  $p(x, t|x_0, t_0)$ . The Chapman–Kolmogorov equation then has the following differential form [42]:

$$\begin{aligned} \frac{\partial p(x, t|x_0, t_0)}{\partial t} = & -\frac{\partial}{\partial x} [F(x, t)p(x, t|x_0, t_0)] + \frac{\partial^2}{\partial x^2} [D(x, t)p(x, t|x_0, t_0)] \\ & + \int [W(x|\tilde{x}, t)p(\tilde{x}, t|x_0, t_0) - W(\tilde{x}|x, t)p(x, t|x_0, t_0)] d\tilde{x}, \end{aligned} \tag{29}$$

where

$$W(x|\tilde{x}, t) \equiv \lim_{\Delta t \rightarrow 0} \frac{p(x, t + \Delta t|\tilde{x}, t)}{\Delta t}. \tag{30}$$

The first and second term of Eq. (29) are associated with the continuous part of  $X(t)$  and are often referred to as the advective and diffusive terms in the parlance of partial differential equations [38]. The last term in Eq. (29), on the other hand, represents how probability mass is transferred through discontinuities in  $X(t)$ , as the function  $W$  is the instantaneous transition rate between different values of  $x$ . Indeed, sample path continuity is expressed as  $W(x|\tilde{x}, t) = 0$  [42], and in that case, Eq. (29) reduces to the Fokker–Planck equation. Note that, in the case where both the drift and diffusion functions vanish, Eq. (29) reduces to the so-called Master equation in statistical mechanics [36, 73].

In order to evaluate Eq. (30), we first introduce three mutually exclusive and collectively exhaustive events:  $E_0 = \{\text{no jumps in } [t, t + \Delta t]\}$ ,  $E_1 = \{\text{exactly one jump in } [t, t + \Delta t]\}$ , and  $E_2 = \{\text{two or more jumps in } [t, t + \Delta t]\}$ , such that  $E_i \cap E_j = \emptyset$ , for  $i \neq j$ , and  $\text{Prob}(\bigcup_i E_i) = 1$ . From the law of total probability, we get

$$p(x, t + \Delta t|\tilde{x}, t) = \sum_{i=1}^3 p(x, t + \Delta t|\tilde{x}, t; E_i) \cdot \text{Prob}(E_i). \tag{31}$$

Assuming that discontinuities occur through a Poisson process with parameter  $\lambda$ , we can use the well-known results [74, 75]:

$$\begin{aligned} \text{Prob}(E_0) &= 1 - \lambda \Delta t + o(\Delta t), \\ \text{Prob}(E_1) &= \lambda \Delta t + o(\Delta t), \\ \text{Prob}(E_2) &= o(\Delta t), \end{aligned} \tag{32}$$

where  $o(\Delta t)$  goes to zero faster than  $\Delta t$ :

$$\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0. \tag{33}$$

Inserting the sum in Eq. (31) into the limit of Eq. (30) results in the sum of three limits, associated with  $E_0$ ,  $E_1$ , and  $E_2$ :

$$W(x|\tilde{x}, t) \equiv \lim_{\Delta t \rightarrow 0} \sum_{i=0}^3 \frac{p(x, t + \Delta t|\tilde{x}, t; E_i) \cdot \text{Prob}(E_i)}{\Delta t}, \tag{34}$$

where the first term vanishes because of sample path continuity between jump events  $\lim_{\Delta t \rightarrow 0} p(x, t + \Delta t|\tilde{x}, t; E_0)/\Delta t = 0$ , while the third one vanishes as per Eq. (33)  $\lim_{\Delta t \rightarrow 0} \text{Prob}(E_2)/\Delta t = 0$ . If jump amplitudes are distributed according to  $Q_B$ , then this distribution can be defined as  $Q_B \equiv \lim_{\Delta t \rightarrow 0} p(x, t + \Delta t|\tilde{x}, t; E_1)$ , which results in

$$\begin{aligned} W(x|\tilde{x}, t) &= \lim_{\Delta t \rightarrow 0} \frac{p(x, t + \Delta t|\tilde{x}, t; E_1) \cdot \text{Prob}(E_1)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{p(x, t + \Delta t|\tilde{x}, t; E_1) \cdot (\lambda \Delta t + o(\Delta t))}{\Delta t} \\ &= \lambda Q_B(x - \tilde{x}). \end{aligned} \tag{35}$$

When inserted into Eq. (29), this yields

$$\begin{aligned} \frac{\partial p(x, t|x_0, t_0)}{\partial t} &= -\frac{\partial}{\partial x} [F(x, t)p(x, t|x_0, t_0)] + \frac{\partial^2}{\partial x^2} [D(x, t)p(x, t|x_0, t_0)] \\ &\quad - \lambda p(x, t|x_0, t_0) + \lambda \int Q_B(x - \tilde{x})p(\tilde{x}, t|x_0, t_0) d\tilde{x}. \end{aligned} \tag{36}$$

If we assume a time-homogeneous drift function, additive noise, and a positive support for  $Q_B$ , we recover Eq. (24).

**Acknowledgements**

The authors would like to thank Érik Harvey-Girard, Alexandre Payeur, Wilhelm Braun, and James Jun for helpful discussions, Curtis Marcoux and Anh-Tuan Trinh for providing the experimental data, and Leonard Maler for proofreading the manuscript.

**Funding**

Funding for this research was granted by the Natural Sciences and Engineering Research Council of Canada and the Ontario Graduate Scholarship program.

**Abbreviations**

SDE, stochastic differential equation; PDF, probability density function; ACF, autocorrelation function; PSD, power spectral density; FP, false positive; ELL, electrosensory lateral line lobe; CLS, centrolateral segment; CMS, centromedial segment.

**Availability of data and materials**

The dataset analysed during the current study are available from the corresponding author on request.

**Ethics approval and consent to participate**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Consent for publication**

Not applicable.

**Authors' contributions**

AM and AL conceptualized the initial direction of the project. AM derived the theory, performed the numerical simulations, analyzed the experimental data, and wrote the manuscript, under the supervision of AL. AM and AL proofread, edited, and approved the final version of the manuscript.

**Author details**

<sup>1</sup>Department of Physics, University of Ottawa, Ottawa, Canada. <sup>2</sup>Département de physique et d'astronomie, Université de Moncton, Moncton, Canada. <sup>3</sup>Centre for Neural Dynamics, University of Ottawa, Ottawa, Canada. <sup>4</sup>Brain and Mind Research Institute, University of Ottawa, Ottawa, Canada.

**Endnotes**

- <sup>a</sup> Alternatively, one could obtain Eq. (9) directly from Eq. (7) by invoking the law of total probability [74] for the set of mutually exclusive and exhaustive events  $\{\tau = i\Delta t, i \in \mathbb{N}^*\}$ .
- <sup>b</sup> This equilibrium PDF exists for stable processes and is defined by  $P_Y(y) = \lim_{t \rightarrow \infty} P_{Y|Y_0}(y, t|y_0, t_0)$  [43].
- <sup>c</sup> In this case we refer to how stochastic synaptic currents, modeled as shot noise and assumed to arrive at a high rate, can be approximated as a diffusion process.

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 7 September 2018 Accepted: 9 July 2019 Published online: 26 July 2019

**References**

1. Haken H. Information and self-organization: a macroscopic approach to complex systems. Berlin: Springer; 1988. OCLC: 649906343.
2. Holovatch Y, Kenna R, Thurner S. Complex systems: physics beyond physics. *Eur J Phys*. 2017;38(2):023002.
3. Nicolis G, Rouvas-Nicolis C. Complex systems. *Scholarpedia*. 2007;2(11):1473.
4. Clemson PT, Stefanovska A. Discerning non-autonomous dynamics. *Phys Rep*. 2014;542(4):297–368.
5. Friedrich R, Peinke J, Sahimi M, Reza Rahimi Tabar M. Approaching complexity by stochastic methods: from biological systems to turbulence. *Phys Rep*. 2011;506(5):87–162.
6. Siegert S, Friedrich R, Peinke J. Analysis of data sets of stochastic systems. *Phys Lett A*. 1998;243(5):275–80.
7. Lamouroux D, Lehnertz K. Kernel-based regression of drift and diffusion coefficients of stochastic processes. *Phys Lett A*. 2009;373(39):3507–12.
8. van Mourik AM, Daffertshofer A, Beek PJ. Estimating Kramers–Moyal coefficients in short and non-stationary data sets. *Phys Lett A*. 2006;351(1):13–7.
9. Prussek J, Stochastic LK. Qualifiers of epileptic brain dynamics. *Phys Rev Lett*. 2007;98(13):138103.
10. Deco G, Marti D, Ledberg A, Reig R, Sanchez Vives MV. Effective reduced diffusion-models: a data driven approach to the analysis of neuronal dynamics. *PLoS Comput Biol*. 2009;5(12):e1000587.
11. Hindriks R, Jansen R, Bijma F. Unbiased estimation of Langevin dynamics from time series with application to hippocampal field potentials in vitro. *Phys Rev E*. 2011;84(2):021133.
12. Hindriks R, Bijma F. Dynamics underlying spontaneous human alpha oscillations: a data-driven approach. *NeuroImage*. 2011;57(2):440–51.
13. Petelczyc M, Zebrowski JJ, Baranowski R. Kramers–Moyal coefficients in the analysis and modeling of heart rate variability. *Phys Rev E*. 2009;80:031127.
14. Petelczyc M, Zebrowski JJ, Orłowska-Baranowska E. A fixed mass method for the Kramers–Moyal expansion—application to time series with outliers. *Chaos*. 2015;25(3):033115.
15. Ragwitz M, Kantz H. Indispensable finite time corrections for Fokker–Planck equations from time series data. *Phys Rev Lett*. 2001;87(25):254501.
16. Tutkun M, Mydlarski L. Markovian properties of passive scalar increments in grid-generated turbulence. *New J Phys*. 2004;6(1):49.
17. Mousavi SM, Reihani SNS, Anvari G, Anvari M, Alinezhad HG, Tabar MRR. Stochastic analysis of time series for the spatial positions of particles trapped in optical tweezers. *Sci Rep*. 2017;7(1):4832.
18. Musila M, Generalized LP. Stein's model for anatomically complex neurons. *Biosystems*. 1991;25(3):179–91.
19. Giraudo MT, Sacerdote L. Jump-diffusion processes as models for neuronal activity. *Biosystems*. 1997;40(1):75–82.
20. Giraudo MT, Sacerdote L, Sirovich R. Effects of random jumps on a very simple neuronal diffusion model. *Biosystems*. 2002;67(1):75–83.
21. Sirovich R, Sacerdote L. Noise induced phenomena in jump diffusion models for single neuron spike activity. In: 2004 IEEE international joint conference on neural networks (IEEE cat. no. 04CH37541). vol. 4. 2004. p. 3025–8.

22. Sirovich R, Sacerdote L, Villa AEP. Cooperative behavior in a jump diffusion model for a simple network of spiking neurons. *Math Biosci Eng.* 2014;11(2):385–401.
23. Łuczka J, Bartussek R, Hänggi P. White-noise-induced transport in periodic structures. *Europhys Lett.* 1995;31(8):431.
24. Łuczka J, Czernik T, Hänggi P. Symmetric white noise can induce directed current in ratchets. *Phys Rev E.* 1997;56:3968–75.
25. Czernik T, Kula J, Łuczka J, Hänggi P. Thermal ratchets driven by Poissonian white shot noise. *Phys Rev E.* 1997;55:4057–66.
26. Kou SG. Jump-diffusion models for asset pricing in financial engineering. In: Birge JR, Linetsky V, editors. *Financial engineering. Handbooks in operations research and management science.* vol. 15. Amsterdam: Elsevier; 2007. p. 73–116.
27. Tankov P, Voltchkova E. *Jump-diffusion models: a practitioner's guide.* Banque et Marchés. 2009.
28. Daly E, Porporato A. Probabilistic dynamics of some jump-diffusion systems. *Phys Rev E.* 2006;73(2):026108.
29. Bandi FM, Nguyen TH. On the functional estimation of jump-diffusion models. *J Econom.* 2003;116(1):293–328.
30. Johannes M. The statistical and economic role of jumps in continuous-time interest rate models. *J Finance.* 2004;59(1):227–60.
31. Anvari M, Tabar MRR, Peinke J, Lehnertz K. Disentangling the stochastic behavior of complex time series. *Sci Rep.* 2016;6:35435.
32. Lee SS, Hannig J. Detecting jumps from Lévy jump diffusion processes. *J Financ Econ.* 2010;96(2):271–90.
33. Figueroa-López JE, Nisen J. Optimally thresholded realized power variations for Lévy jump diffusion models. *Stoch Process Appl.* 2013;123(7):2648–77.
34. Pawula RF. Approximation of the linear Boltzmann equation by the Fokker–Planck equation. *Phys Rev.* 1967;162:186–8.
35. Jun JJ, Longtin A, Maler L. Active sensing associated with spatial learning reveals memory-based attention in an electric fish. *J Neurophysiol.* 2016;115(5):2577–92.
36. Gillespie DT. *Markov processes: an introduction for physical scientists.* 1st ed. Boston: Academic Press; 1991.
37. Risken H, Frank T. *The Fokker–Planck equation: methods of solution and applications.* 2nd ed. Springer series in synergetics. Berlin: Springer; 1996.
38. Patankar SV. *Numerical heat transfer and fluid flow.* Series in computational methods in mechanics and thermal sciences. New York: Hemisphere; 1980. OCLC: 31743097.
39. Katsaggelos AK. Iterative image restoration algorithms. *Opt Eng.* 1989;28(7):287735.
40. Ait-Sahalia Y, Mancini L. Out of sample forecasts of quadratic variation. *J Econom.* 2008;147(1):17–33.
41. Hansen P, Lunde A. A realized variance for the whole day based on intermittent high-frequency data. *J Financ Econ.* 2005;3:525–54.
42. Gardiner CW. *Stochastic methods: a handbook for the natural and social sciences.* 4th ed. Springer series in synergetics. Berlin: Springer; 2009.
43. Gillespie DT, Simple SE. *Brownian diffusion: an introduction to the standard theoretical models.* 1st ed. Oxford: Oxford University Press; 2013.
44. Marcoux CM, Clarke SE, Nesse WH, Longtin A, Maler L. Balanced ionotropic receptor dynamics support signal estimation via voltage-dependent membrane noise. *J Neurophysiol.* 2016;115(1):530–45.
45. White JA, Rubinstein JT, Kay AR. Channel noise in neurons. *Trends Neurosci.* 2000;23(3):131–7.
46. Noonan L, Doiron B, Laing C, Longtin A, Turner RW. A dynamic dendritic refractory period regulates burst discharge in the electrosensory lobe of weakly electric fish. *J Neurosci.* 2003;23(4):1524–34.
47. Dorval AD, White JA. Channel noise is essential for perithreshold oscillations in entorhinal stellate neurons. *J Neurosci.* 2005;25(43):10025–8.
48. Diba K, Lester HA, Intrinsic KC. Noise in cultured hippocampal neurons: experiment and modeling. *J Neurosci.* 2004;24(43):9723–33.
49. Bahraminasab A, Ghasemi F, Stefanovska A, McClintock PVE, Friedrich R. Physics of brain dynamics: Fokker–Planck analysis reveals changes in EEG  $\delta$ – $\theta$  interactions in anaesthesia. *New J Phys.* 2009;11(10):103051.
50. McDonnell MD, Ward LM. The benefits of noise in neural systems: bridging theory and experiment. *Nat Rev Neurosci.* 2011;12(7):415–26.
51. Schmerl BA, McDonnell MD. Channel-noise-induced stochastic facilitation in an auditory brainstem neuron model. *Phys Rev E.* 2013;88(5):052722.
52. Burkitt AN. A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. *Biol Cybern.* 2006;95(1):1–19.
53. Faisal AA, Laughlin SB. Stochastic simulations on the reliability of action potential propagation in thin axons. *PLoS Comput Biol.* 2007;3(5):e79.
54. Schneidman E, Freedman B, Segev I. Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. *Neural Comput.* 1998;10(7):1679–703.
55. Faisal AA, Selen LPJ, Wolpert DM. Noise in the nervous system. *Nat Rev Neurosci.* 2008;9(4):292–303.
56. Chow C, Spontaneous WJA. Action potentials due to channel fluctuations. *Biophys J.* 1997;71:3013–21.
57. Goldwyn JH, Shea-Brown E. The what and where of adding channel noise to the Hodgkin–Huxley equations. *PLoS Comput Biol.* 2011;7(11):e1002247.
58. Fox RF, Lu Y. Emergent collective behavior in large numbers of globally coupled independently stochastic ion channels. *Phys Rev E.* 1994;49(4):3421–31.
59. Fox RF. Stochastic versions of the Hodgkin–Huxley equations. *Biophys J.* 1997;72(5):2068–74.
60. Goldwyn JH, Imenov NS, Famulare M, Shea-Brown E. Stochastic differential equation models for ion channel noise in Hodgkin–Huxley neurons. *Phys Rev E.* 2011;83(4 Pt 1):041908.
61. Linaro D, Storace M, Giugliano M. Accurate and fast simulation of channel noise in conductance-based model neurons by diffusion approximation. *PLoS Comput Biol.* 2011;7(3):e1001102.
62. Orio P, Soudry D. Simple fast and accurate implementation of the diffusion approximation algorithm for stochastic ion channels with multiple states. *PLoS ONE.* 2012;7(5):e36670.
63. Rowat P. Interspike interval statistics in the stochastic Hodgkin–Huxley model: coexistence of gamma frequency bursts and highly irregular firing. *Neural Comput.* 2007;19(5):1215–50.

64. Richardson MJE, Gerstner W. Statistics of subthreshold neuronal voltage fluctuations due to conductance-based synaptic shot noise. *Chaos*. 2006;16(2):026106.
65. Jacobson GA, Diba K, Yaron-Jakoubovitch A, Oz Y, Koch C, Segev I, et al. Subthreshold voltage noise of rat neocortical pyramidal neurones. *J Physiol*. 2005;564(Pt 1):145–60.
66. Steinmetz PN, Manwani A, Koch C, London M, Segev I. Subthreshold voltage noise due to channel fluctuations in active neuronal membranes. *J Comput Neurosci*. 2000;9(2):133–48.
67. Turner RW, Maler L, Deerinck T, Levinson SR, Ellisman MH. TTX-sensitive dendritic sodium channels underlie oscillatory discharge in a vertebrate sensory neuron. *J Neurosci*. 1994;14(11):6453–71.
68. Zhang J, Carver CM, Choveau FS, Shapiro MS. Clustering and functional coupling of diverse ion channels and signaling proteins revealed by super-resolution STORM microscopy in neurons. *Neuron*. 2016;92(2):461–78.
69. Bol K, Marsat G, Harvey-Girard E, Longtin A, Maler L. Frequency-tuned cerebellar channels and burst-induced LTD lead to the cancellation of redundant sensory inputs. *J Neurosci*. 2011;31(30):11028–38.
70. Ashida G, Kubo M. Suprathreshold stochastic resonance induced by ion channel fluctuation. *Physica D*. 2010;239(6):327–34.
71. Rudolph M, Destexhe A. Correlation detection and resonance in neural systems with distributed noise sources. *Phys Rev Lett*. 2001;86(16):3662–5.
72. Jun JJ, Longtin A, Maler L. Active sensing associated with spatial learning reveals memory-based attention in an electric fish. *J Neurophysiol*. 2016;115(5):2577–92.
73. Pathria RK, Beale PD. *Statistical mechanics*. 3rd ed. Amsterdam: Elsevier/Boston: Academic Press; 2011.
74. Papoulis A, Probability PSU. *Random variables and stochastic processes with errata sheet*. 4th ed. Boston: McGraw-Hill; 2002.
75. Hsu HP. *Schaum's outline of probability, random variables, and random processes*. 3rd ed. New York: McGraw-Hill; 2014.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---