POLITECNICO DI TORINO
Repository ISTITUZIONALE

Estimating input parameters from intracellular recordings in the Feller neuronal model

Original
Estimating input parameters from intracellular recordings in the Feller neuronal model / Bibbona, E., Lansky, P., Sirovich, R.. - In: PHYSICAL REVIEW E, STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS. - ISSN 1539-3755. - 81:(2010), pp. 031916-1-031916-13. [10.1103/PhysRevE.81.031916]

Availability:
This version is available at: 11583/2646684 since: 2016-11-18T15:33:50Z

Publisher:
American Physical Society

Published
DOI:10.1103/PhysRevE.81.031916

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

26 November 2026

Estimating input parameters from intracellular recordings in the Feller neuronal model

Enrico Bibbona*
Department of Mathematics G. Peano, University of Torino, Via Carlo Alberto 10, 10123 Torino, Italy

Petr Lansky†
Institute of Physiology, Academy of Sciences of the Czech Republic, Videnska 1083, 142 20 Prague 4, Czech Republic

Roberta Sirovich‡
Department of Mathematics G. Peano, University of Torino, Via Carlo Alberto 10, 10123 Torino, Italy

(Received 10 November 2009; published 25 March 2010)

We study the estimation of the input parameters in a Feller neuronal model from a trajectory of the membrane potential sampled at discrete times. These input parameters are identified with the drift and the infinitesimal variance of the underlying stochastic diffusion process with multiplicative noise. The state space of the process is restricted from below by an inaccessible boundary. Further, the model is characterized by the presence of an absorbing threshold, the first hitting of which determines the length of each trajectory and which constrains the state space from above. We compare, both in the presence and in the absence of the absorbing threshold, the efficiency of different known estimators. In addition, we propose an estimator for the drift term, which is proved to be more efficient than the others, at least in the explored range of the parameters. The presence of the threshold makes the estimates of the drift term biased, and two methods to correct it are proposed.

DOI: 10.1103/PhysRevE.81.031916 PACS number(s): 87.10.−e, 87.19.L−, 02.50.−r, 05.40.−a

I. INTRODUCTION

Stochastic leaky integrate-and-fire (LIF) models are among the most popular mathematical descriptors for the neuronal activity [1–4]. These are simple models that reproduce with a reasonable degree of approximation the response of a neuron or complex neuronal models [5]. They appear in many variants; in all of them the membrane potential is described as a stochastic process, whereas the spike generation is due to the crossing of a threshold level by the process. The most common stochastic LIF model is based on the Ornstein-Uhlenbeck process [6–9]. In that model the synaptic transmission is state independent and the membrane potential is not limited from below (nor from above, apart that if the threshold is reached a spike is generated and the process is reset). As a consequence it may happen that the membrane potential reaches unrealistic low values. In order to prevent such events, modifications that better characterize the process of synaptic transmission by inclusion of reversal potentials were proposed and explored [10,11]. Here we consider one of the variants of the LIF model with inhibitory reversal potential in which the membrane potential is described by the so-called Feller process. This model with multiplicative noise includes state-dependent inputs and ensures that the membrane potential fluctuations are limited by an inaccessible lower boundary as the effect of the inhibitory postsynaptic potential decreases if the membrane potential gets closer to it.

The quality of a model, however, should be judged on the basis of a quantitative comparison with real data. Starting from [12–15], for many years only few papers were dedicated to this problem [16–20], but recently attention to this topic began to grow. Attempts to establish a benchmark test that permits a comparison between the predictive capability of different single neuron models over a broad range of firing rates and firing patterns were suggested in [21] and two competitions, the neural prediction challenge [22] and the quantitative single neuron model competition [23], took place. This effort aims at prediction of the output from the knowledge of the input. From a statistical point of view, the first step in the direction of comparing output data with output produced by the models is an efficient estimation of the parameters. The input signal is the most relevant parameter of the model [24,25]. Experimental data are essentially of two kinds—extracellular and intracellular recordings. In the former case just the sequence of firing times is registered, and parameter (signal) estimation for data of that kind was studied, for example, in [26–29] and a review can be found in [30]. In the latter case the trajectory of the membrane potential is sampled at discrete times. Statistical methods for the estimation of the parameters from discretized trajectories are available for a large class of processes [31,32]. The application to the stochastic neuronal models is not obvious due to the presence of the absorbing threshold which has been shown to bring a bias in the maximum likelihood estimator of the drift parameter [33].Only few references [34–36] provide statistical methods dedicated to parameter estimation for randomly stopped processes, but they are restricted to very special cases.

In the present paper we focus on the Feller model and estimation of its input parameters from a sampled trajectory. Some studies were already performed outside the neuronal context but only in the absence of the threshold [37–39]. Thus, we have several different estimators whose performances should be compared in a parameter range expected

*enrico.bibbona@unito.it
†lansky@biomed.cas.cz
‡roberta.sirovich@unito.it

 ©2010 The American Physical Society

for neuronal data and in the presence of the threshold.

The Feller process was first introduced in [40] as an example of a singular diffusion process and it has many applications apart from neuronal modeling. Well known is a model of short term interest rates as proposed in [41] from where the mathematical finance community adapted its name to Cox Ingersoll Ross process. More recently it has been used to model nitrous oxide emission from soil [42] or, in the framework of survival analysis theory, to model the individual hazard rate [43].

In Sec. II we review the relevant properties of the Feller process. Some methods for the estimation of the input parameters are reviewed and a different one is introduced in Sec. III. In Sec. IV a complete comparison of the different estimators both in the absence and in the presence of the absorbing threshold is performed. Finally methods to improve the estimates in the presence of the threshold are proposed.

II. MODEL

One of the most common models for the neuronal depolarization is the Ornstein-Uhlenbeck diffusion process U_t which can be given in the form of the following stochastic differential equation:

$$dU_t = \left(-\frac{U_t - u_0}{\tau_U} + \mu_U \right) dt + \sigma_U dW_t, \quad U_0 = u_0, \tag{1}$$

where u_0 is the resetting and resting potential, $\tau_U > 0$ is the membrane time constant, μ_U is the drift parameter, $\sigma_U > 0$ is the diffusion coefficient, and W_t is a standard Brownian motion (Wiener process).

The physiology of the cell does not allow the membrane potential to reach too low values and a better model should include a lower bound of the depolarization. In [11] the following model was proposed. The evolution of the membrane potential is described by the process Y_t in the form of the following stochastic differential equation:

$$dY_t = \left(-\frac{Y_t - y_0}{\tau} + \mu_Y \right) dt + \sigma \sqrt{Y_t - V_I^Y} dW_t, \quad Y_0 = y_0, \tag{2}$$

where $y_0 > V_I^Y$ is the resetting and resting potential, $\tau > 0$ is the membrane time constant, μ_Y is the drift parameter, and $\sigma > 0$ is the diffusion coefficient. Comparing Eqs. (1) and (2) we can see their identical behavior with vanishing noise. Under the presence of the noise its amplitude in Eq. (2) changes with the actual value of the depolarization, whereas in Eq. (1) it remains constant. The main difference between the models is caused by the new parameter V_I^Y which is the minimum value allowed for the membrane potential and it is called inhibitory reversal potential. If the parameters satisfy condition $2\mu_Y \tau + 2V_I^Y - 2y_0 \geq \sigma^2 \tau$ the process starting above V_I^Y never reaches that value [40]. In [44] the qualitative behavior of the Ornstein-Uhlenbeck and Feller models was compared showing that if parameters are suitably chosen, the two models provide rather similar spike trains (i.e., the sequence of firing times). In [19] it was shown that as long as intracellular recordings are available, true data are in some cases much better fitted by model (2) with respect to a simple Ornstein-Uhlenbeck.

A translated form $X_t = Y_t - V_I^Y$ of the process allows us to get rid of one of the parameters y_0 and V_I^Y. If parameters are mapped as follows,

$$\mu_Y = \mu - \frac{x_0}{\tau},$$

$$V_I^Y = -x_0 + y_0,$$

the Feller process X_t is the solution of the stochastic differential equation,

$$dX_t = \left(-\frac{X_t}{\tau} + \mu \right) dt + \sigma \sqrt{X_t} dW_t, \quad X_0 = x_0, \tag{3}$$

where $x_0 > 0$ is the resetting point and the inhibitory reversal potential is shifted to the zero level. For $2\mu \geq \sigma^2$ the zero value is an inaccessible barrier (an entrance boundary according to Feller's classification [40]); this is a reasonable assumption in the context of neuronal modeling [27] and it holds true in all parametric range explored in this paper.

The conditional expectation of process (3) is

$$\mathbb{E}(X_t | X_0, \mu) = X_0 e^{-t/\tau} + \mu\tau(1 - e^{-t/\tau}). \tag{4}$$

Conditional variance and covariance are

$$\begin{aligned} \mathrm{Var}(X_t | X_0, \mu, \sigma^2) &= \sigma^2 v(X_t | X_0, \mu) \\ &= \sigma^2 \frac{\tau}{2} (1 - e^{-t/\tau}) [\mu\tau(1 - e^{-t/\tau}) + 2X_0 e^{-t/\tau}] \end{aligned} \tag{5}$$

and

$$\mathrm{Cov}(X_t, X_s | X_0, \mu, \sigma^2) = e^{-(t-s)/\tau} \mathrm{Var}(X_s | X_0, \mu, \sigma^2) \tag{6}$$

for $s < t$. The process is time homogeneous and the transition density function $f(x_t, t | x_0)$ is [40,41]

$$f(x, t | x_0) = c e^{-r-s} \left(\frac{s}{r} \right)^{q/2} I_q(2\sqrt{rs}), \tag{7}$$

where

$$c = \frac{2}{\tau\sigma^2(1 - e^{-t/\tau})}, \quad q = \frac{2\mu}{\sigma^2} - 1,$$

$$r = c x_0 e^{-t/\tau}, \quad s = cx,$$

and $I_q(\cdot)$ is the modified Bessel function [45].

Generation of the action potentials is not a part of membrane potential model (3). To make the cell fire, a firing threshold is imposed at level $S > x_0$. The first time the process reaches the boundary level, an action potential is elicited and the membrane potential is instantaneously reset to x_0. Then the evolution restarts anew according to the same law. The times between each consecutive couple of spikes (called interspike intervals) follow the distribution of the random variable $T = \inf\{t \geq 0 | X_t \geq S\}$ that is often referred to as the *first-passage time* across the barrier. The most important

characteristic of the random variable T is its mean [44]

$$\mathbb{E}(T) = \frac{S - X_0}{\mu} + \sum_{j=1}^{\infty} \frac{\tau(S^{j+1} - X_0^{j+1})}{(j+1)\prod_{k=0}^{j} (\mu\tau + k\tau\sigma^2/2)} \tag{8}$$

as it is inversely proportional to the neuronal firing rate. We will use it to check the precision of our numerical experiments. The properties of the model firing are often related to the position of asymptotic depolarization (3) with respect to S and are called subthreshold or suprathreshold regimens. Let us remark that the presence of the threshold is a constraint on the process and its characteristics, such as the accessibility of the states, the transition density, and the moments, are altered.

Following the construction of the model as diffusion approximation of a model with discontinuous trajectories, the parameters could be divided into two classes: parameters characterizing the input and intrinsic parameters characterizing the neuron irrespectively of the incoming signal. However, for model (3), such a classification of the parameters is not so straightforward as for the Ornstein-Uhlenbeck model. Indeed, S and x_0 remain independent on the input, but the membrane time constant is input dependent (for details see [11]). Nevertheless, we assume that these three parameters are known and we focus on the estimation of the remaining two parameters, μ and σ, which characterize the input.

III. ESTIMATORS FOR THE INPUT PARAMETERS

All the estimation methods presented here are designed to work on a sample made of a single trajectory $\{X_i\}_{i=1}^n$ of process (3) recorded at discrete times $t_i = ih$ with constant sampling interval h. Some of them lead to explicit estimators and others to quasilikelihood functions to be minimized or maximized numerically. The presence of the threshold is not accounted for in the estimation methods, however in numerical experiments we thoroughly investigate its effect. All the estimation methods and their main properties are summarized in Table I.

A. Least-squares method

The least-squares (LS) method consists of minimizing the squared deviation between the observed data and their unconditional mean and variance given by Eqs. (4) and (5) (cf. Appendix A 1 for further details). The estimators we get are the following:

$$\hat{\mu}_{\mathrm{LS}} = \frac{\sum_{i=1}^{n} [(X_i - X_0 e^{-ih/\tau})(1 - e^{-ih/\tau})]}{\tau \sum_{i=1}^{n} (1 - e^{-ih/\tau})^2}, \tag{9}$$

TABLE I. Summary of estimation methods and properties in the absence of the threshold. An asterisk means that the estimator would be unbiased if the true value of μ was known. C means consistency, AN means asymptotic normality, O is optimality in the sense of [48], and AE denotes asymptotic efficiency.

Method	Estimator	Properties
Least squares	$\hat{\mu}_{\mathrm{LS}}$	Unbiased
	$\hat{\sigma}^2_{\mathrm{LS}}$	*
Conditional least squares	$\hat{\mu}_{\mathrm{CLS}}$	Unbiased
	$\hat{\sigma}^2_{\mathrm{CLS}}$	*
Gauss-Markov	$\hat{\mu}_{\mathrm{GM}}$	Unknown
Bibby-Sørensen	$\hat{\mu}_{\mathrm{BS}}$	C, AN
	$\hat{\sigma}^2_{\mathrm{BS}}$	C, AN
Optimal estimating function	$\hat{\mu}_{\mathrm{OEF}}$	C, AN, O
Maximum likelihood	$\hat{\mu}_{\mathrm{ML}}$	C, AN, AE
	$\hat{\sigma}^2_{\mathrm{ML}}$	C, AN, AE

$$\hat{\sigma}^2_{\mathrm{LS}} = \frac{\sum_{i=1}^{n} [X_i - \mathbb{E}(X_i|X_0, \hat{\mu}_{\mathrm{LS}})]^2 v(X_i|X_0, \hat{\mu}_{\mathrm{LS}})}{\sum_{i=1}^{n} v^2(X_i|X_0, \hat{\mu}_{\mathrm{LS}})}, \tag{10}$$

where the functions $\mathbb{E}(X_i|X_0, \hat{\mu}_{\mathrm{LS}})$ and $v(X_i|X_0, \hat{\mu}_{\mathrm{LS}})$ are given by Eqs. (4) and (5). For the estimation of μ, the residuals $e_i = X_i - \mathbb{E}(X_i|X_0, \mu)$ are neither independent nor identically distributed. Thus, the asymptotic normality of the estimates is not obvious. However, as one can directly verify using Eq. (4), the estimator $\hat{\mu}_{\mathrm{LS}}$ is unbiased. Its variance is given by formula (A1). Estimator (10) would be unbiased if the true value of the parameter μ was used instead of $\hat{\mu}_{\mathrm{LS}}$.

B. Conditional least-squares method

In order to avoid correlation between the residuals, the LS method may be improved by replacing unconditional mean and variance with conditional ones (cf. Appendix A 2 for further details). We call this method conditional least squares (CLS). The resulting estimators are

$$\hat{\mu}_{\mathrm{CLS}} = \frac{\sum_{i=1}^{n} (X_i - X_{i-1} e^{-h/\tau})}{n\tau(1 - e^{-h/\tau})}, \tag{11}$$

$$\hat{\sigma}^2_{\mathrm{CLS}} = \frac{\sum_{i=1}^{n} [X_i - \mathbb{E}(X_i|X_{i-1}, \hat{\mu}_{\mathrm{CLS}})]^2 v(X_i|X_{i-1}, \hat{\mu}_{\mathrm{CLS}})}{\sum_{i=1}^{n} v^2(X_i|X_{i-1}, \hat{\mu}_{\mathrm{CLS}})}, \tag{12}$$

where the functions $\mathbb{E}(X_i|X_{i-1}, \hat{\mu}_{\mathrm{CLS}})$ and $v(X_i|X_{i-1}, \hat{\mu}_{\mathrm{CLS}})$ are given by Eqs. (4) and (5). Estimator (11) is unbiased as is easily proved by direct evaluation of its moments using Eqs.

(4) and (5) and its variance is given by Eq. (A3). Estimator (12) would be unbiased if the true value of the parameter μ was used instead of $\hat{\mu}_{\mathrm{CLS}}$. When the estimated value is used, a correcting factor $n/(n-1)$ improves the estimates (as we shall see in Sec. IV; cf. also Appendix A 2).

As mentioned, the advantage of the CLS method for the estimation of μ with respect to the LS one is that the residuals $\eta_i = X_i - \mathbb{E}(X_i|X_{i-1},\mu)$ are uncorrelated [it can be proved by applying formula (6)]. However, they still have different variances [cf. formula (A5)].

C. Estimators based on martingale estimating functions

One of the most powerful methods to build estimators for discretely observed diffusion processes consists of constructing some martingale estimating functions [31,37,38]. The estimators are shown to be consistent and asymptotically normal. Applying this method to Feller model (3) the following estimators are obtained:

$$\hat{\mu}_{\mathrm{BS}} = \frac{\sum_{i=1}^{n} [(X_i - e^{-h/\tau} X_{i-1})/X_{i-1}]}{\tau(1-e^{-h/\tau}) \sum_{i=1}^{n} X_{i-1}^{-1}}, \tag{13}$$

$$\hat{\sigma}^2_{\mathrm{BS}} = \frac{2}{\tau(1-e^{-h/\tau})} \times \frac{\sum_{i=1}^{n} \{[X_i - e^{-h/\tau} X_{i-1} - \hat{\mu}_{\mathrm{BS}} \tau(1-e^{-h/\tau})]^2 / X_{i-1}\}}{\sum_{i=1}^{n} \{[\hat{\mu}_{\mathrm{BS}} \tau(1-e^{-h/\tau}) + 2X_{i-1}e^{-h/\tau}]/X_{i-1}\}}. \tag{14}$$

The subscript BS stands for the authors in [37,38]. Analogously to the CLS case, we consider a correcting factor $n/(n-1)$ for estimator (14) [cf. also Eq. (A6)].

D. Gauss-Markov method

Let us introduce an estimator for the parameter μ aiming to improve CLS method. We call this method Gauss-Markov (GM). The estimator $\hat{\mu}_{\mathrm{CLS}}$ given in Eq. (11) is the arithmetic mean of the n quantities

$$\hat{\mu}_i = \frac{X_i - X_{i-1}e^{-h/\tau}}{\tau(1-e^{-h/\tau})},$$

each of them with the same expectation μ but different variances given by [cf. formula (A5)]

$$\mathrm{Var}(\hat{\mu}_i) = \frac{\mathrm{Var}(\eta_i)}{[\tau(1-e^{-h/\tau})]^2},$$

where $\eta_i = X_i - \mathbb{E}(X_i|X_{i-1},\mu)$. If these variances were known, the estimator for μ with the smallest variance would be the weighted mean,

$$\hat{\mu}_{\mathrm{GM}} = \frac{\sum_i \hat{\mu}_i p_i}{\sum_i p_i}, \tag{15}$$

with weights $p_i = \sigma^2/\mathrm{Var}(\eta_i)$ (cf. [46,47]; Gauss-Markov theorem). Here the variances depend on the parameter we are estimating, thus a direct application of the Gauss-Markov theorem is not possible. However we can build an estimator in two stages: we first estimate μ using $\hat{\mu}_{\mathrm{CLS}}$ [cf. formula (11)] and then we approximate the weights using the estimated value of μ in Eq. (A5). The following weights are obtained:

$$p_i = \left[(\tau x_0 - \hat{\mu}_{\mathrm{CLS}}\tau^2)(1-e^{-h/\tau})e^{-ih/\tau} + \frac{\hat{\mu}_{\mathrm{CLS}}\tau^2}{2}(1-e^{-2h/\tau}) \right]^{-1}. \tag{16}$$

Finally, inserting Eq. (16) into Eq. (15), the estimator takes form

$$\hat{\mu}_{\mathrm{GM}} = \frac{\sum_i (X_i - X_{i-1}e^{-h/\tau}) p_i}{\tau(1-e^{-h/\tau}) \sum_i p_i}. \tag{17}$$

An analogous estimator for σ^2 is not available.

E. Optimal estimating function method

In [48], an optimality criterion based on the minimality of the asymptotic variance was established for a class of estimating functions. In [37,39] it was shown that in the case of estimation of μ for Feller process such optimality is achieved by finding the zeros of the following optimal estimating function (OEF):

$$F(\mu) = \sum_{i=1}^{n} \frac{\tau(1-e^{-h/\tau})}{\mathrm{Var}(X_i|X_{i-1},\mu,\sigma^2)} [X_i - \mathbb{E}(X_i|X_{i-1},\mu)], \tag{18}$$

where $\mathbb{E}(X_i|X_{i-1},\mu)$ and $\mathrm{Var}(X_i|X_{i-1},\mu,\sigma^2)$ are given by Eqs. (4) and (5). Function (18) is a weighted sum of the residuals η_i where the weights are proportional to the inverses of the variances of X_i conditioned by X_{i-1}. The estimate of μ is obtained numerically and it is denoted by $\hat{\mu}_{\mathrm{OEF}}$. Due to the linearity of the conditional variance in σ^2 to find zeros of Eq. (18) no knowledge of σ^2 is needed. Consistency and asymptotic normality are proved in [37]. Let us note that for small sampling intervals ($h \to 0$) the conditional variance $\mathrm{Var}(X_i|X_{i-1},\mu,\sigma^2) \approx \sigma^2 X_{i-1} h$ and that if we apply such approximation to function (18) we can found that the zeros are given by $\hat{\mu}_{\mathrm{BS}}$. For small sampling step the optimal estimating function (OEF) method gives results very close to the BS one.

F. Maximum likelihood method (ML)

The transition density is known in analytic form (7) for Feller model (3) and thus maximum likelihood method is feasible. The log-likelihood function is the following:

TABLE II. Values of the parameters used in the numerical procedures.

	Name	Symbol	Value	Units
Known constants	Reversal potential		0	mV
	Resetting potential	x_0	10	mV
	Threshold	S	20	mV
	Time constant	τ	22–90	ms
	Discretization step	h	0.01	ms
Parameters to be estimated	Drift	μ	0.4–1.4	mV ms^{-1}
	Infinitesimal variance	σ^2	0.0081–0.0992	mV ms^{-1}

$$\mathrm{L}(x_i, \ldots, x_n|\mu,\sigma) = \log\left(\prod_{i=1}^{n} f(x_i,h|x_{i-1})\right), \tag{19}$$

where $f(x_i,h|x_{i-1})$ is given by Eq. (7). Estimates are obtained by numerical maximization of the log-likelihood and they are denoted by $\hat{\mu}_{\mathrm{ML}}$ and $\hat{\sigma}^2_{\mathrm{ML}}$. Consistency, asymptotic normality and efficiency of ML estimators is proved in [31,49].

IV. RESULTS

To compare the above summarized procedures with the GM method introduced in Sec. III D and the effect of the absorbing threshold we perform a set of numerical experiments. For that purpose not 1 but 10 000 trajectories have been simulated. The estimation procedures introduced in Sec. III are then applied to each simulated trajectory. The sample averages of the estimates [that we denote by avg$(\hat{\mu}_*)$ and avg$(\hat{\sigma}^2_*)$, where $*$ stands for one of the subscripts LS, CLS, BS, GM, ML, or OEF] are computed together with the sample variances [denoted by lowercase var$(\hat{\mu}_*)$] used to calculate the confidence intervals under the assumption of normality of the estimators (which is empirically confirmed for all the explored ranges of the parameters). Parameters are chosen in a biologically compatible ranges in correspondence with the experimental values obtained in [17,18] for Ornstein-Uhlenbeck model. For example, σ is chosen in such a way that the two models have the same variance at the resting level [44]. In the numerical experiments we explore different parameter settings that cover both the so-called subthreshold $(S>\mu\tau)$ and suprathreshold $(S<\mu\tau)$ regimens. The values of the parameters we considered are summarized in Table II. To check the reliability of the simulation procedure we compare $\mathbb{E}(T)$ given by Eq. (8) with its estimates obtained from the simulations (see Fig. 1).

A. Comparison of estimators

1. Estimation of μ

The results on the estimators of μ in the absence and in the presence of the threshold are illustrated in Figs. 2 and 3, respectively. To show the properties of the estimators we plot confidence intervals for the bias, avg$(\hat{\mu}_*)-\mu$. Let us remark that both $\hat{\mu}_{\mathrm{ML}}$ and $\hat{\mu}_{\mathrm{OEF}}$ give estimates which are indistinguishable from those given by $\hat{\mu}_{\mathrm{BS}}$. This first result is not surprising as both BS and OEF estimators converge for small sampling interval to the one coming from the continuous likelihood function as it was proved in [37] and also the discrete likelihood for $h\to 0$ converges to the continuous one (cf. [49]). We have $h=0.01$ ms and the resemblance between $\hat{\mu}_{\mathrm{BS}}$, $\hat{\mu}_{\mathrm{OEF}}$, and $\hat{\mu}_{\mathrm{ML}}$ is almost perfect. Thus for the detailed comparison only four methods (LS, CLS, BS, and GM) remain and in the rest of the paper whenever we present results on the BS estimators they equally apply to OEF and ML.

As illustrated in Fig. 2, all the estimators except for $\hat{\mu}_{\mathrm{BS}}$ are unbiased. This result confirms theoretical conclusions about the estimators $\hat{\mu}_{\mathrm{LS}}$, $\hat{\mu}_{\mathrm{CLS}}$, and $\hat{\mu}_{\mathrm{GM}}$. As $\hat{\mu}_{\mathrm{BS}}$ is con-

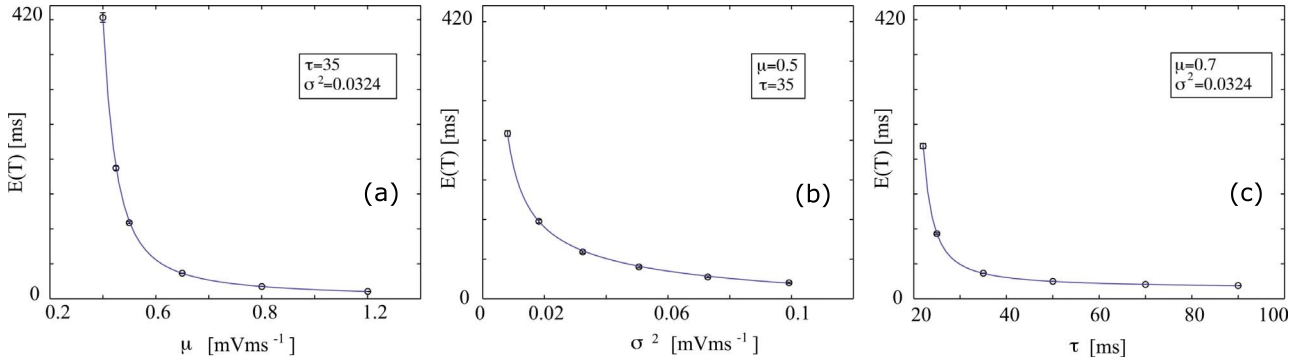

FIG. 1. (Color online) Expectation of the first-passage time of the process X in dependency on (a) μ for fixed $\tau=35$ ms and $\sigma^2=0.0324$ mV ms^{-1}, (b) σ^2 for fixed $\tau=35$ ms and $\mu=0.5$ mV ms^{-1}, and (c) τ for fixed $\mu=0.7$ mV ms^{-1} and $\sigma^2=0.0324$ mV ms^{-1}. The continuous line is calculated using Eq. (8); the confidence intervals (not visible) are obtained by simulations.

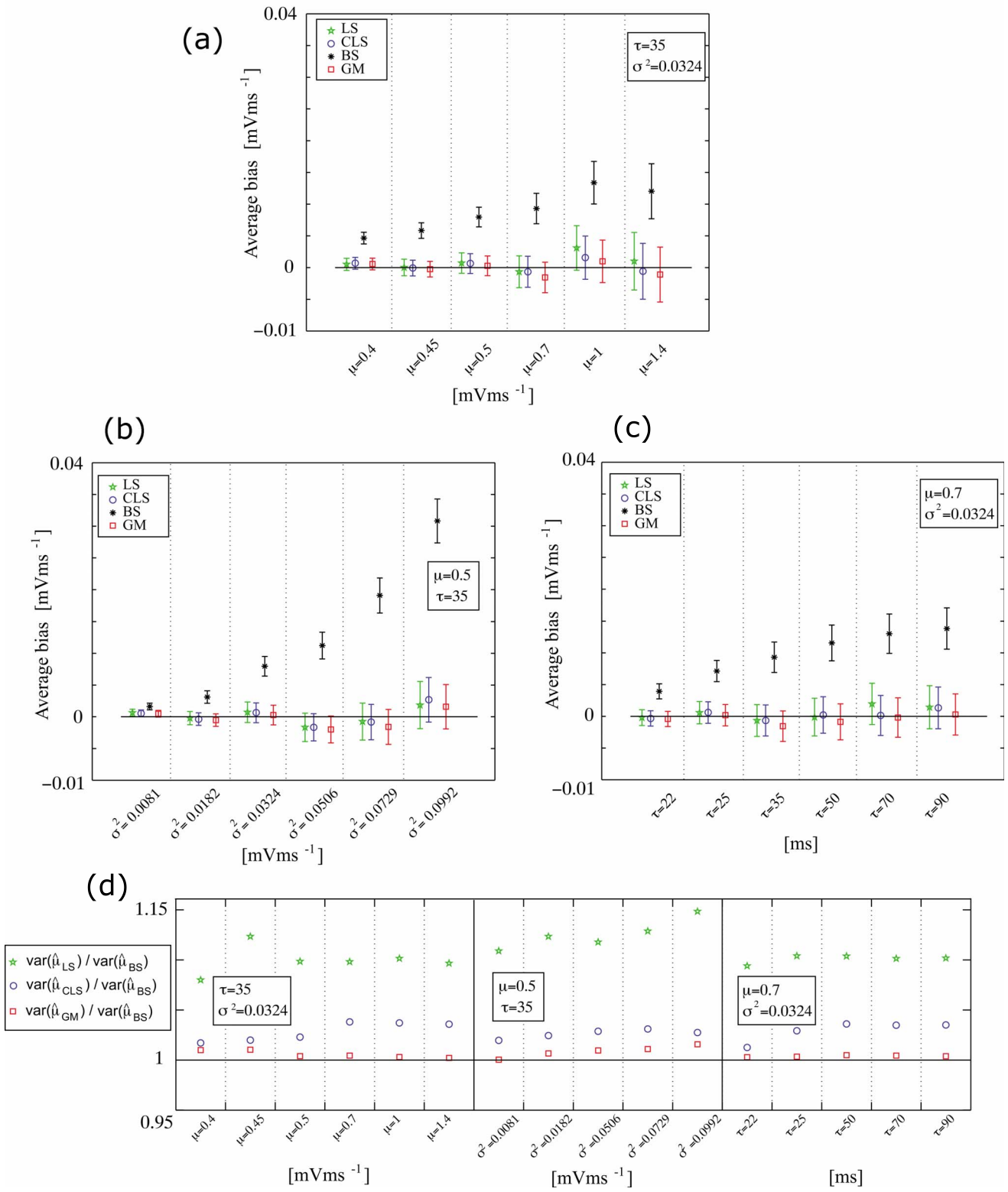

FIG. 2. (Color online) Estimation of μ in the absence of the threshold. Average bias and its confidence interval in dependency on (a) μ for fixed τ=35 ms and σ^2=0.0324 mV ms^{-1}, (b) σ^2 for fixed τ=35 ms and μ=0.5 mV ms^{-1}, and (c) τ for fixed μ=0.7 mV ms^{-1} and σ^2=0.0324 mV ms^{-1}. (d) Sample variance ratios $\mathrm{var}(\hat{\mu}_*)/\mathrm{var}(\hat{\mu}_{\mathrm{BS}})$.

cerned, it was proved that it is asymptotically unbiased, but the sample sizes used here (the mean lengths of the trajectories go from a minimum of 1000 points to a maximum of 42 000 points) are apparently not sufficient to reach the asymptotic properties. In both cases the bias seems to increase with μ, σ^2, and τ. This is related to the length of the trajectories because, due to the simulation procedure described in Appendix B, even in the absence of the threshold

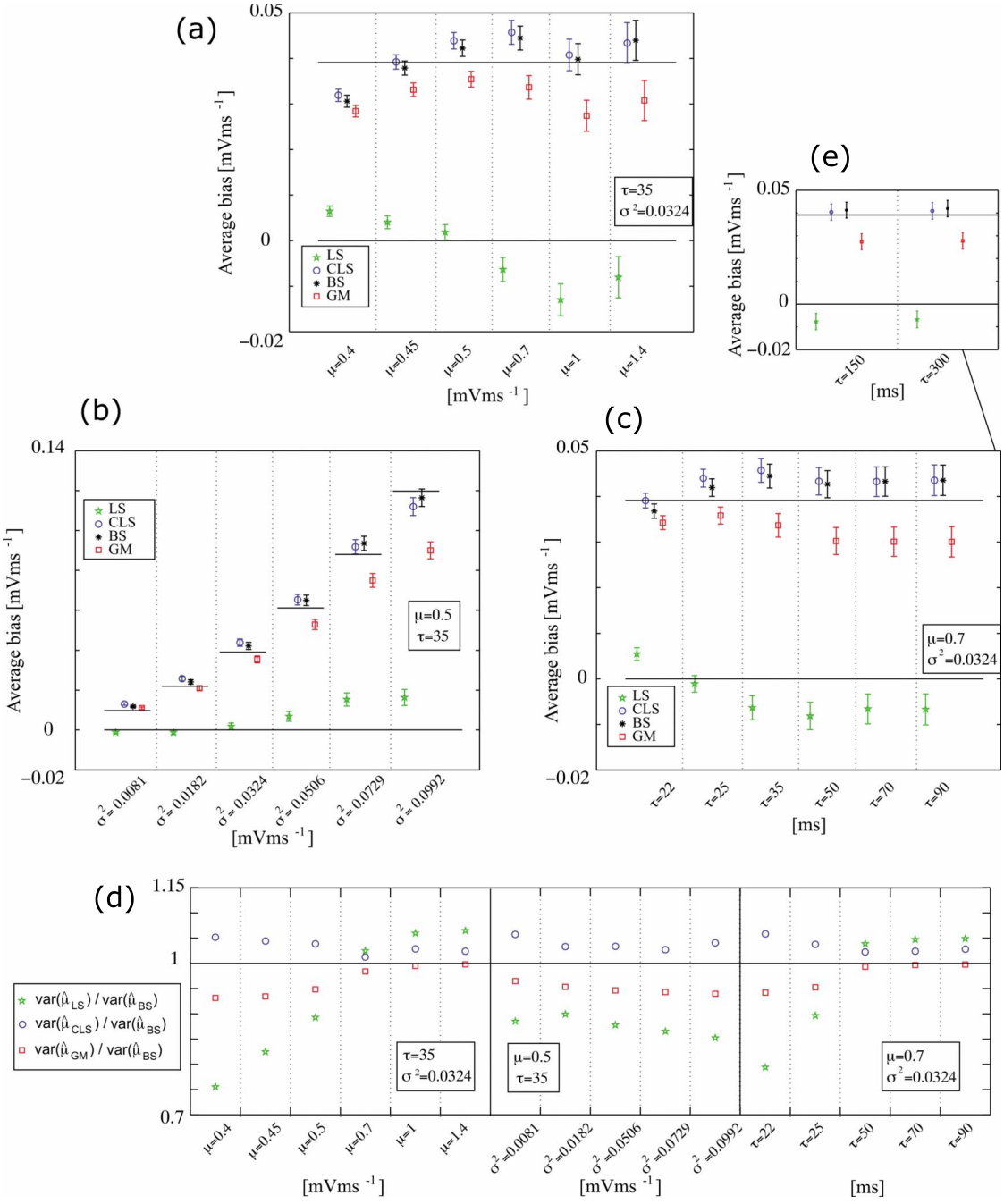

FIG. 3. (Color online) Estimation of μ in the presence of the threshold. Average bias and its confidence interval in dependency on (a) μ for fixed $\tau=35$ ms and $\sigma^2=0.0324$ mV ms^{-1}, (b) σ^2 for fixed $\tau=35$ ms and $\mu=0.5$ mV ms^{-1}, and (c)–(e) τ for fixed $\mu=0.7$ mV ms^{-1} and $\sigma^2=0.0324$ mV ms^{-1} [the values in (e) are out of the neuronal range]. The horizontal lines in panels (a)–(c) and (e) are the asymptotic approximation for the bias given in Eq. (20). (d) Sample variance ratios $\mathrm{var}(\hat{\mu}_*)/\mathrm{var}(\hat{\mu}_{\mathrm{BS}})$.

the trajectories have the same length as in its presence and they get shorter when increasing the parameters. Hence in Figs. 2(a) and 2(c), the process goes from a subthreshold regimen ($\mu\tau<20$), where trajectories are longer, to a suprathreshold regimen ($\mu\tau>20$), where trajectories are shorter and strongly driven by the deterministic component. A similar remark holds true for Fig. 2(b): here $\mu\tau=S$ and the length of the trajectories is shorter for larger values of the variability coefficient σ^2. Let us stress that this result for $\hat{\mu}_{\mathrm{BS}}$ is not in contradiction with its asymptotic consistency proved in [37]. Obviously, all the plotted confidence intervals increase their size when the estimates are calculated on shorter trajectories, suggesting that the variance of all the estimators increases.

In Fig. 2(d) the variances of the estimators are compared as they reflect the precision of the estimating procedures. Note that even if variances (A1) and (A3) are available analytically, they are not used in the figure because they just hold when the length n of the trajectories is fixed. For each parameter setting the estimators show sample variances

ordered as follows: $\mathrm{var}(\hat{\mu}_{\mathrm{BS}}) < \mathrm{var}(\hat{\mu}_{\mathrm{GM}}) < \mathrm{var}(\hat{\mu}_{\mathrm{CLS}}) < \mathrm{var}(\hat{\mu}_{\mathrm{LS}})$. Thus BS estimator has been considered as a common reference since it has the smallest sample variance and the ratios $\mathrm{var}(\hat{\mu}_*)/\mathrm{var}(\hat{\mu}_{\mathrm{BS}})$ are plotted. We can see that $\hat{\mu}_{\mathrm{GM}}$ shows sample variances very near to that of $\hat{\mu}_{\mathrm{BS}}$. On the other hand $\hat{\mu}_{\mathrm{LS}}$ has a variance approximately 10–15 % larger than these two. In conclusion, $\hat{\mu}_{\mathrm{GM}}$ is both unbiased and with small variability in all the explored range of the parameters.

In the presence of the threshold, all the estimators for μ are biased (cf. Fig. 3), and the bias is larger than those that $\hat{\mu}_{\mathrm{BS}}$ shows in the absence of the threshold. The estimators $\hat{\mu}_{\mathrm{BS}}$, $\hat{\mu}_{\mathrm{CLS}}$, and $\hat{\mu}_{\mathrm{GM}}$ show a similar behavior: the bias is always positive, weakly dependent on μ [cf. Fig. 3(a)], and settles to a constant value for larger values of τ [cf. Fig. 3(c)]. Moreover it increases linearly with σ^2 [cf. Fig. 3(b)], with a smaller slope for $\hat{\mu}_{\mathrm{GM}}$. A qualitatively different bias is shown by $\hat{\mu}_{\mathrm{LS}}$: for small values of σ^2 the bias is negative and changes to positive as σ^2 increases [cf. Fig. 3(b)], and it is always smaller than the bias of the other estimators. We can see that imposing the absorbing threshold on the process has a substantial effect on the quality of estimation of μ. Therefore whenever the data are collected under this condition, the estimate has to be corrected or at least taken with care.

In Fig. 3(d) the variances of the estimators are compared using the same procedure as in Fig. 2(d). The presence of the threshold changes the ranking of the variances. Indeed in the subthreshold regimen $\hat{\mu}_{\mathrm{LS}}$ gives the estimates with the smallest variability (up to a 30% gain) while in the suprathreshold range of the parameters it is the worst one (but just 5% worse, then the best two, $\hat{\mu}_{\mathrm{BS}}$ and $\hat{\mu}_{\mathrm{GM}}$). Concluding, $\hat{\mu}_{\mathrm{LS}}$ has the smaller bias and in the subthreshold regimen it also has the smallest variance.

2. Estimation of σ^2

The results about the estimators of σ^2 in the absence of the threshold are illustrated in Fig. 4. The estimator $\hat{\sigma}^2_{\mathrm{LS}}$ is omitted from the figure since it is strongly biased and the comparison with $\hat{\sigma}^2_{\mathrm{BS}}$ and $\hat{\sigma}^2_{\mathrm{CLS}}$ is trivial. The error in the estimation of σ^2 by LS method can be predicted by closed form calculations (see Appendix C for the details). The remaining two estimators [we used them with the correcting factors presented in formula (A4) and (A6)] are unbiased in all the explored range of the parameters. Their sample variances increase on short trajectories for larger values of μ, σ^2, and τ. As shown in Fig. 4(d), $\hat{\sigma}^2_{\mathrm{CLS}}$ has larger variance with respect to $\hat{\sigma}^2_{\mathrm{BS}}$ in all the explored range of the parameters. Hence $\hat{\sigma}^2_{\mathrm{BS}}$ is the best choice: it is both unbiased and with smallest variability.

The presence of the threshold does not influence the quality of the estimation of σ^2. This holds despite the fact that the estimates are based on biased estimates of μ probably due to the small effect of this bias on the result.

B. Corrections for bias

As shown in Fig. 3, the presence of the threshold induces biases on $\hat{\mu}_{\mathrm{CLS}}$, $\hat{\mu}_{\mathrm{BS}}$, and $\hat{\mu}_{\mathrm{GM}}$ which are larger than the bias shown by $\hat{\mu}_{\mathrm{LS}}$. However those always positive biases show regular trends with the parameters: weakly dependent on μ, linearly increasing with σ^2, and for larger τ they settle to constant values [cf. Figs. 3(c) and 3(e) for larger values of τ] out of the neuronal range. We propose here two methods to find a correction of the bias. The first one, based on analytical results for the limiting case $\tau \to \infty$, is derived only for $\hat{\mu}_{\mathrm{CLS}}$. The second one, based on simulations, can be applied to all the estimators.

Let us consider CLS estimator (11) in the limit for $\tau \to \infty$. The estimator tends to $(X_n - x_0)/(nh)$, where X_n is the last measured value of X_t and h is the sampling interval. In the presence of the threshold and with small h, the last point X_n is very close to the threshold and nh approximates the first passage time T. The CLS estimator converges then to $(S - x_0)/T$. If moreover $\mu = \sigma^2/4$, the Feller process X_t can be transformed into a Wiener process $W_t = 2\sqrt{X_t}$ and in that case we have a closed form expression for the expectation of $1/T$. Transforming back to the Feller process we get

$$\mathbb{E}(\hat{\mu}_{\mathrm{CLS}}) \sim \mathbb{E}\left(\frac{S - x_0}{T}\right) = \frac{\sigma^2 (S - x_0)}{4(\sqrt{S} - \sqrt{x_0})^2}$$

and the bias is

$$\mathbb{E}(\hat{\mu}_{\mathrm{CLS}}) - \mu \sim \frac{\sigma^2}{4}\left[\frac{S - x_0}{(\sqrt{S} - \sqrt{x_0})^2} - 1\right]. \tag{20}$$

Despite that in the considered parameter range τ is not large enough to allow the identification $\hat{\mu}_{\mathrm{CLS}} \sim (S - x_0)/T$ and μ is much larger than $\sigma^2/4$, approximation (20) still holds in many of the illustrated cases. In Figs. 3(a), 3(c), and 3(e) a horizontal line is plotted in correspondence of bias (20). For larger values of μ and τ the formula fits the estimates not only for $\hat{\mu}_{\mathrm{CLS}}$ but also for $\hat{\mu}_{\mathrm{BS}}$. Also $\hat{\mu}_{\mathrm{GM}}$ seems to settle to a constant value but smaller than Eq. (20). Applying Eq. (20) as a correction, we always get improved estimates.

Alternatively, let us introduce the following method which again uses the knowledge of σ^2. Replacing σ^2 by its estimate creates no problem as it can be estimated with no bias even in the presence of the threshold. Let us denote by μ_1 the biased estimate obtained by one of the GM, CLS, or BS methods. Then one can simulate a sample of trajectories in the presence of the threshold with $\mu = \mu_1$. From such simulations one estimates again μ getting the value μ_2. Since the bias of the estimator depends weakly on the true value of the parameter μ, the quantity $b = \mu_2 - \mu_1$ gives the bias (or at least an approximation) for both the simulated trajectories and the original sample of real data. Hence the corrected estimate for μ is simply $\mu_1 - b$. Despite it appears complicated, this method provides unbiased estimates. To get the best possible estimate for parameter μ in the presence of a threshold we suggest the use of GM estimator corrected as explained above. It has the smallest variance and it is unbiased in the absence of the threshold.

V. CONCLUSIONS

Real data coming from intracellular recordings of the membrane potential of some neurons were analyzed using different statistical techniques and in order to answer different questions in papers [17–20,50]. These are attempts to use

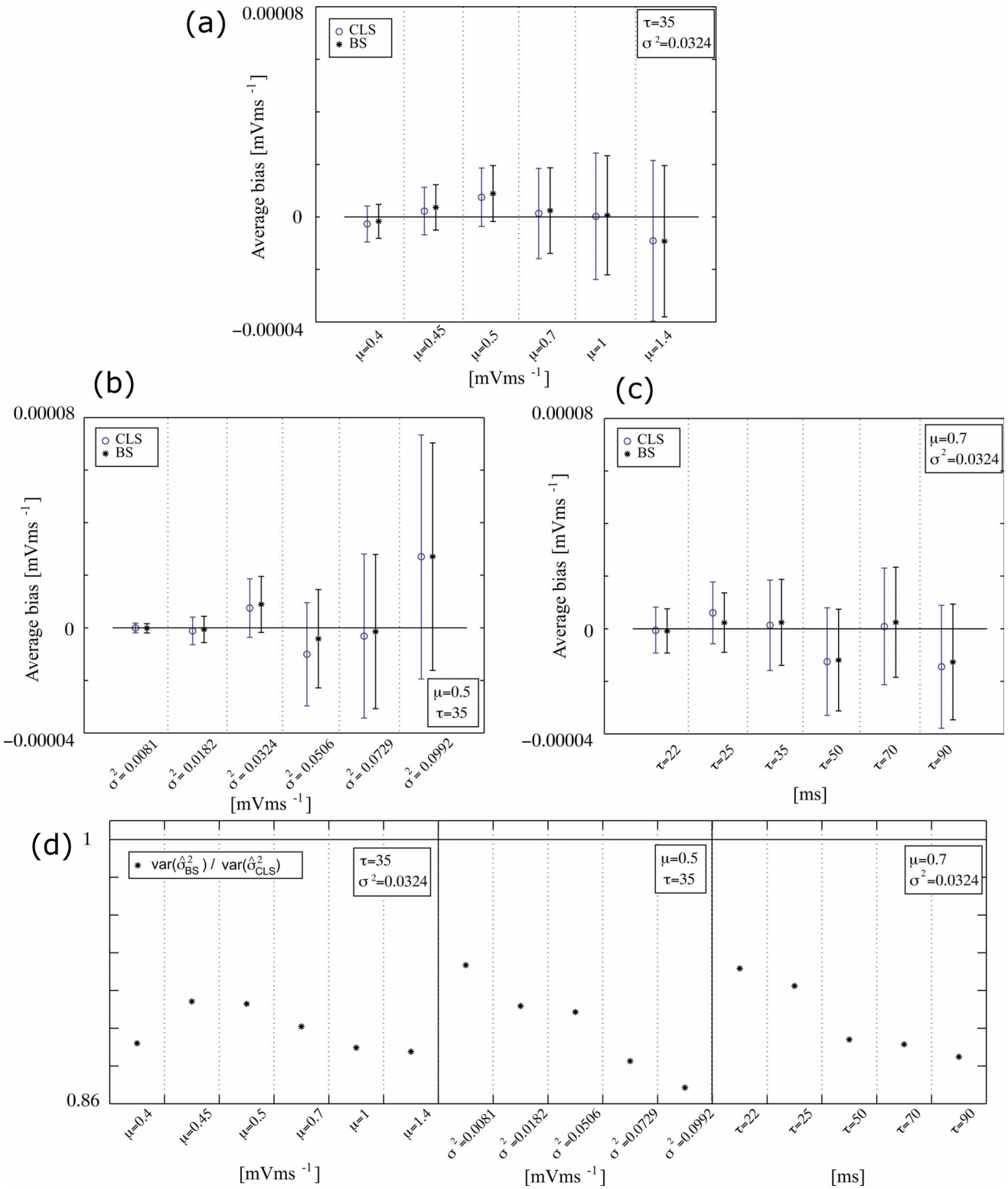

FIG. 4. (Color online) Estimation of σ^2 in the absence of the threshold. Average bias and its confidence interval in dependency on (a) μ for fixed τ=35 ms and σ^2=0.0324 mV ms^{-1}, (b) σ^2 for fixed τ=35 ms and μ=0.5 mV ms^{-1}, and (c) τ for fixed μ=0.7 mV ms^{-1} and σ^2=0.0324 mV ms^{-1}. (d) Sample variance ratio var($\hat{\sigma}^2_{\mathrm{BS}}$)=var($\hat{\sigma}^2_{\mathrm{CLS}}$).

experimental data in validation of mathematical models instead of investigating their qualitative behavior only. For the data recorded *in vivo* the only direct way to deduce the incoming signal is via the estimates of the input parameters. It was confirmed in [18] that these parameters change if the neuron is exposed to some type of stimulation. Simultaneously it was shown in [19] that there are conditions under which the Feller model fits the data better than the Ornstein-

Uhlenbeck. Therefore, the efficient estimation of parameters for the Feller model is worth studying.

In the present paper we reviewed LS, CLS, BS, ML, and OEF estimation methods for the input parameters μ and σ^2 of the Feller neuronal model and we introduced GM method for μ. We compared their performances on simulated samples both in the absence and in the presence of the threshold. In the absence of the threshold the GM estimator for μ is clearly the best one, while for σ^2 BS estimator is the one with the smallest variance. Whatever the method, the presence of the threshold brings a bias into the estimate of μ while it leaves the estimate of σ^2 unaffected. LS estimator for μ has the smallest bias and in the subthreshold regimen it also has the smallest variance. However GM and BS estimators still have a small variance (the smallest one in the suprathreshold regimen) and their biases may be strongly reduced by means of the two procedures that we have proposed.

The LIF neuronal models assume that the spiking mechanism is due to crossing of the firing threshold. Then, the phenomenon has to be also taken into account in the estimation methods. The present paper contributes to this aim at least showing that if the threshold is neglected a systematic error in the estimation procedure for μ is introduced. We quantify the bias and provide two methods for reducing it.

ACKNOWLEDGMENTS

This study was supported by the Center for Neurosciences LC554, by Grant No. AV0Z50110509, and Grant Agency of the Academy of Sciences of the Czech Republic (Project No. IAA101120604). Istituto Nazionale di Alta Matematica F. Severi is acknowledged for the financial support for the stay of P.L. at Torino University.

APPENDIX A: DETAILS ON ESTIMATION METHODS

1. Least-squares method

In order to derive LS estimators (9) and (10) the procedure is the following. First minimize the function

$$F_1(\mu) = \sum_{i=1}^{n} [X_i - \mathbb{E}(X_i|X_0,\mu)]^2$$

with respect to μ and get $\hat{\mu}_{\mathrm{LS}}$ as in Eq. (9); then using Eq. (9) instead of μ in Eq. (4) minimize

$$F_2(\sigma^2) = \sum_{i=1}^{n} ([X_i - \mathbb{E}(X_i|X_0,\hat{\mu}_{\mathrm{LS}})]^2 - \mathbb{E}\{[X_i - \mathbb{E}(X_i|X_0,\hat{\mu}_{\mathrm{LS}})]^2\})^2$$

with respect to σ^2 and get $\hat{\sigma}^2_{\mathrm{LS}}$ as in Eq. (10).

The variance of estimator (9) is given by

$$\mathrm{Var}(\hat{\mu}_{\mathrm{LS}}|X_0,\mu,\sigma^2) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(1-e^{-jh/\tau})(1-e^{-ih/\tau})\mathrm{Cov}(X_j,X_i|X_0,\mu,\sigma^2)}{\tau^2\left[\sum_{k=1}^{n}(1-e^{-kh/\tau})^2\right]^2}. \tag{A1}$$

Substituting Eq. (6) and summing, an explicit expression can be written but it is omitted here for the sake of simplicity.

2. Conditional least-squares method

In order to avoid correlation between the residuals, the LS method can be improved by replacing unconditional mean and variance with conditional ones. The procedures are then to minimize the following quantity

$$F_3(\mu) = \sum_{i=1}^{n} [X_i - \mathbb{E}(X_i|X_{i-1},\mu)]^2 \tag{A2}$$

with respect to μ getting estimator $\hat{\mu}_{\mathrm{CLS}}$ of Eq. (11), then minimize

$$F_4(\sigma^2) = \sum_{i=1}^{n} ([(X_i - \mathbb{E}(X_i|X_{i-1},\hat{\mu}_{\mathrm{CLS}})]^2 - \mathbb{E}\{[(X_i - \mathbb{E}(X_i|X_{i-1},\hat{\mu}_{\mathrm{CLS}})]^2|X_{i-1}\})^2$$

with respect to σ^2, and get $\hat{\sigma}^2_{\mathrm{CLS}}$ of Eq. (12). The variance of estimator (11) is given by

$$\mathrm{Var}(\hat{\mu}_{\mathrm{CLS}}|X_0,\mu,\sigma^2) = \frac{\sigma^2\left[e^{-h/\tau}(1-e^{-nh/\tau})(X_0-\mu\tau)+\frac{n\mu\tau}{2}(1-e^{-2h/\tau})\right]}{n^2\tau(1-e^{h/\tau})^2}. \tag{A3}$$

From the results of our simulations it is apparent that estimator (12) is biased on shorter trajectories when the estimated value of μ is used in its expression. Introducing the correcting factor $n/(n-1)$ we get the estimator

$$\hat{\sigma}^2_{\mathrm{CLSc}} = \frac{n}{n-1}\hat{\sigma}^2_{\mathrm{CLS}}, \tag{A4}$$

which proves to be unbiased [the values of corrected estimator (A4) are those plotted in Fig. 4].

The variances of the residuals η_i in the CLS method are

$$\begin{aligned}\mathrm{Var}(\eta_i) &= \mathrm{Var}[X_i - e^{-h/\tau}X_{i-1} - \mu\tau(1-e^{-h/\tau})|X_0,\mu] \\ &= \mathrm{Var}(X_i|X_0,\mu) - e^{-2h/\tau}\,\mathrm{Var}(X_{i-1}|X_0,\mu) \\ &= \sigma^2\left[(\tau x_0 - \mu\tau^2)(1-e^{-h/\tau})e^{-ih/\tau} + \frac{\mu\tau^2}{2}(1-e^{-2h/\tau})\right]\end{aligned} \tag{A5}$$

as can be seen by using Eqs. (4)–(6).

3. Estimators based on martingale estimating functions

From the results of our simulations it is apparent that estimator (14) is biased on shorter trajectories when the estimated value of μ is used in its expression. Introducing the correcting factor $n/(n-1)$ we get the estimator

$$\hat{\sigma}^2_{\mathrm{BSc}} = \frac{n}{n-1}\hat{\sigma}^2_{\mathrm{BS}}, \tag{A6}$$

which proves to be unbiased [the values of corrected estimator (A6) are those plotted in Fig. 4].

APPENDIX B: NUMERICAL METHODS

The trajectories for the numerical experiment were simulated according to the following procedure. The Wagner-Platen scheme [51] is implemented to build the time discrete trajectory $X_0, X_1, \ldots, X_n$ with constant time increment h. A trajectory in the presence of the threshold is generated at first. It means that we simulate the process until it hits the barrier for the first time. To detect the first hitting time we proceed as follows: starting from $X_i = x_i$, below S, we generate the point X_{i+1}, then we check if X_{i+1} lies below S. If not, then the time $(i+1)h$ is returned as the first passage time. If X_{i+1} is below S, due to the continuity of the process it is still possible that a hitting of the threshold occurred between the times ih and $(i+1)h$. To account for this possibility, we use the algorithm proposed in [52]. Finally, if $X_{i+1} < S$ and no hidden passage occurs in between, we proceed simulating the next point of the trajectory. As a consequence of the presence of the threshold, the lengths of the trajectories are different. Since the properties of the estimators depend on the length of the sampled path, we proceed as follows: for each trajectory simulated in the presence of the threshold, we simulate a trajectory in the absence of the threshold with the same length. In this way the effect of the different lengths on the estimates is the same in both cases and we test the effect of the presence threshold only. To check the reliability of the simulation procedure we compare $\mathbb{E}(T)$ given by Eq. (8) with its estimates obtained from the simulations (see Fig. 1). The estimation methods introduced in Secs. III E and III F are performed by means of the standard Nelder-Mead simplex minimization algorithm, with starting point $\hat{\mu}_{\mathrm{GM}}$.

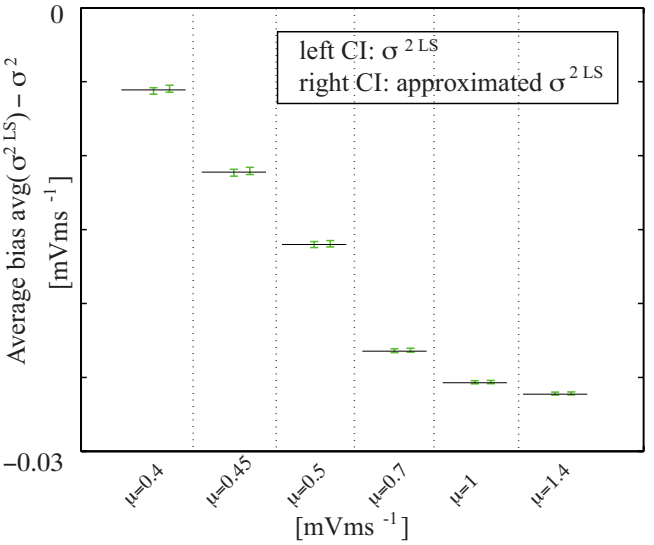

FIG. 5. (Color online) Average bias of $\hat{\sigma}^2_{\mathrm{LS}}$ in the absence of the threshold. The left confidence intervals are obtained from $\hat{\sigma}^2_{\mathrm{LS}}$ [Eq. (10)] and the right confidence intervals are obtained from its approximation (C1). The horizontal continuous lines are given by the approximated expectation of $\hat{\sigma}^2_{\mathrm{LS}}$ in Eq. (C2).

APPENDIX C: ON THE BIAS OF $\hat{\sigma}^2_{\mathrm{LS}}$

As shown in Fig. 5, estimator (10) can be approximated substituting the true value of the parameter μ in the term derived from the variance as follows:

$$\hat{\sigma}^2_{\mathrm{LS}} \sim \frac{\sum_{i=1}^{n} [X_i - \mathbb{E}(X_i|X_0, \hat{\mu}_{\mathrm{LS}})]^2 v(X_i|X_0, \mu)}{\sum_{i=1}^{n} v^2(X_i|X_0, \mu)}. \tag{C1}$$

Let us rewrite the term $[X_i - \mathbb{E}(X_i|X_0, \hat{\mu}_{\mathrm{LS}})]^2$ as

$$\begin{aligned}[X_i - \mathbb{E}(X_i|X_0, \hat{\mu}_{\mathrm{LS}})]^2 &= [X_i - X_0 e^{-ih/\tau} - \mu\tau(1 - e^{-ih/\tau}) - (\hat{\mu}_{\mathrm{LS}} - \mu)\tau(1 - e^{-ih/\tau})]^2 \\ &= [X_i - \mathbb{E}(X_i|X_0, \mu)]^2 - 2(\hat{\mu}_{\mathrm{LS}} - \mu)\tau(1 - e^{-ih/\tau})[X_i - \mathbb{E}(X_i|X_0, \mu)] + (\hat{\mu}_{\mathrm{LS}} - \mu)^2\tau^2(1 - e^{-ih/\tau})^2.\end{aligned}$$

Hence we obtain

$$\hat{\sigma}^2_{\mathrm{LS}} \sim \frac{\sum_{i=1}^{n} [X_i - \mathbb{E}(X_i|x_0,\mu)]^2 v(X_i|X_0,\mu)}{\sum_{i=1}^{n} v^2(X_i|X_0,\mu)} - 2(\hat{\mu}_{\mathrm{LS}} - \mu)$$
$$\times \frac{\sum_{i=1}^{n} \{\tau(1-e^{-ih/\tau})[X_i - \mathbb{E}(X_i|X_0,\mu)]\} v(X_i|X_0,\mu)}{\sum_{i=1}^{n} v^2(X_i|X_0,\mu)} + (\hat{\mu}_{\mathrm{LS}} - \mu)^2 \frac{\sum_{i=1}^{n} [\tau^2(1-e^{-ih/\tau})^2] v(X_i|X_0,\mu)}{\sum_{i=1}^{n} v^2(X_i|X_0,\mu)}.$$

Computing the expectation we get

$$\mathbb{E}(\hat{\sigma}^2_{\mathrm{LS}}) \sim \sigma^2 - 2\frac{\sum_{i=1}^{n} [\tau(1-e^{-ih/\tau})\mathrm{Cov}(\hat{\mu}_{\mathrm{LS}}, e_i|X_0,\mu,\sigma^2)] v(X_i|X_0,\mu)}{\sum_{i=1}^{n} v^2(X_i|X_0,\mu)} + \mathrm{Var}(\hat{\mu}_{\mathrm{LS}}|X_0,\mu,\sigma^2)\frac{\sum_{i=1}^{n} [\tau^2(1-e^{-ih/\tau})^2] v(X_i|X_0,\mu)}{\sum_{i=1}^{n} v^2(X_i|X_0,\mu)},$$

where $e_i = X_i - \mathbb{E}(X_i|X_0,\mu)$ are the residuals. Being the variance of $\hat{\mu}_{\mathrm{LS}}$ given by Eq. (A1),

$$\mathrm{Cov}(\hat{\mu}_{\mathrm{LS}}, e_i|X_0,\mu,\sigma^2) = \frac{\mathrm{Cov}\left\{\left[\sum_{j=1}^{n} e_j(1-e^{-jh/\tau})\right], e_i|X_0,\mu,\sigma^2\right\}}{\tau\sum_{k=1}^{n}(1-e^{-kh/\tau})^2} = \frac{\sum_{j=1}^{n}(1-e^{-jh/\tau})\mathrm{Cov}(X_j,X_i|X_0,\mu,\sigma^2)}{\tau\sum_{k=1}^{n}(1-e^{-kh/\tau})^2},$$

and using Eqs. (5) and (6), we get

$$\mathbb{E}(\hat{\sigma}^2_{\mathrm{LS}}) \sim \sigma^2 - 2\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}[(1-e^{-ih/\tau})(1-e^{-jh/\tau})\mathrm{Cov}(X_j,X_i|X_0,\mu,\sigma^2)]v_i}{\sum_{k=1}^{n}(1-e^{-kh/\tau})^2\sum_{i=1}^{n} v_i^2}$$
$$+ \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}[(1-e^{-jh/\tau})(1-e^{-ih/\tau})\mathrm{Cov}(X_j,X_i|X_0,\mu,\sigma^2)]\sum_{k=1}^{n}[(1-e^{-kh/\tau})^2]v_k}{\left[\sum_{k=1}^{n}(1-e^{-kh/\tau})^2\right]^2\sum_{i=1}^{n} v_i^2}. \quad \text{(C2)}$$

This result completely fits the simulated data plotted in Fig. 5. A perfect fit is also shown for varying τ and σ. Let us remark that this result holds for fixed lengths n of the trajectories. Hence, for each considered set of parameters, the trajectories used to derive the confidence intervals in Fig. 5 have not been simulated according to the procedure described in Appendix B but with length fixed to the expected first passage time.

[1] H. C. Tuckwell, *Introduction to Theoretical Neurobiology* (Cambridge University Press, Cambridge, 1988), Vols. 1–2.
[2] W. Gerstner and W. M. Kistler, *Spiking Neuron Models* (Cambridge University Press, Cambridge, 2002).
[3] A. N. Burkitt, Biol. Cybern. **95**, 1 (2006).
[4] M. J. E. Richardson, Phys. Rev. E **76**, 021919 (2007).
[5] W. M. Kistler, W. Gerstner, and J. L. vanHemmen, Neural Comput. **9**, 1015 (1997).
[6] L. M. Ricciardi and L. Sacerdote, Biol. Cybern. **35**, 1 (1979).
[7] A. R. Bulsara, T. C. Elston, C. R. Doering, S. B. Lowen, and K. Lindenberg, Phys. Rev. E **53**, 3958 (1996).
[8] B. Lindner, M. J. Chacron, and A. Longtin, Phys. Rev. E **72**, 021911 (2005).
[9] B. Lindner, L. Schimansky-Geier, and A. Longtin, Phys. Rev. E **66**, 031916 (2002).
[10] F. B. Hanson and H. C. Tuckwell, J. Theor. Neurobiol. **2**, 127

(1983).
[11] P. Lansky and V. Lanska, Biol. Cybern. **56**, 19 (1987).
[12] H. C. Tuckwell and W. Richter, J. Theor. Biol. **71**, 167 (1978).
[13] P. Lansky, Math. Biosci. **67**, 247 (1983).
[14] J. Inoue, S. Sato, and L. M. Ricciardi, Biol. Cybern. **73**, 209 (1995).
[15] S. Shinomoto, Y. Sakai, and S. Funahashi, Neural Comput. **11**, 935 (1999).
[16] L. Paninski, J. Pillow, and E. Simoncelli, Neurocomputing **65-66**, 379 (2005).
[17] P. Lansky, P. Sanda, and J. He, J. Comput. Neurosci. **21**, 211 (2006).
[18] P. Lansky, P. Sanda, and J. He, J. Physiol. (Paris) (to be published).
[19] R. Höpfner, Math. Biosci. **207**, 275 (2007).
[20] P. Jahn, Ph.D. thesis, Johannes Gutenberg University of Mainz, http://ubm.opus.hbz-nrw.de/volltexte/2009/1939/
[21] R. Jolivet, A. Kobayashi, A. Rauch, R. Naud, S. Shinomoto, and W. Gerstner, J. Neurosci. Methods **169**, 417 (2008).
[22] http://neuralprediction.berkeley.edu/
[23] http://lcn.epfl.ch/QuantNeuronMod/
[24] J. Inoue and S. Doi, BioSystems **87**, 49 (2007).
[25] R. D. Vilela and B. Lindner, J. Theor. Biol. **257**, 90 (2009).
[26] L. Paninski, J. Pillow, and E. Simoncelli, Neural Comput. **16**, 2533 (2004).
[27] S. Ditlevsen and P. Lansky, Phys. Rev. E **73**, 061910 (2006).
[28] S. Ditlevsen and P. Lansky, Phys. Rev. E **76**, 041906 (2007).
[29] P. Mullowney and S. Iyengar, J. Comput. Neurosci. **24**, 179 (2008).
[30] P. Lansky and S. Ditlevsen, Biol. Cybern. **99**, 253 (2008).
[31] B. L. S. Prakasa Rao, *Statistical Inference for Diffusion Type Processes*, Kendall's Library of Statistics, Vol. 8 (Edward Arnold, London, 1999).
[32] Y. A. Kutoyants, *Statistical Inference for Ergodic Diffusion Processes*, Springer Series in Statistics (Springer-Verlag, London, 2004).
[33] E. Bibbona, P. Lansky, L. Sacerdote, and R. Sirovich, Phys. Rev. E **78**, 011918 (2008).
[34] M. A. Girshick, F. Mosteller, and L. J. Savage, Ann. Math. Stat. **17**, 13 (1946).
[35] B. Ferebee, J. Appl. Probab. **20**, 94 (1983).
[36] M. Sørensen, Int. Statist. Rev. **51**, 93 (1983).
[37] B. M. Bibby and M. Sørensen, Bernoulli **1**, 17 (1995).
[38] B. M. Bibby and M. Sørensen, Theory of Stochastic Processes **2**, 49 (1996).
[39] L. Overbeck and T. Rydén, J. Econ. Theory **13**, 430 (1997).
[40] W. Feller, Ann. Math. **54**, 173 (1951).
[41] J. C. Cox, J. E. Ingersoll, and S. A. Ross, Econometrica **53**, 385 (1985).
[42] A. R. Pedersen, Scand. J. Stat. **27**, 385 (2000).
[43] O. Aalen and H. Gjessing, Lifetime Data Anal **10**, 407 (2004).
[44] P. Lansky, L. Sacerdote, and F. Tomassetti, Biol. Cybern. **73**, 457 (1995).
[45] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Applied Mathematics Series (U.S. Government, Washington, D.C., 1964), Vol. 55.
[46] D. G. Luenberger, *Optimization by Vector Space Methods* (Wiley, New York, 1969).
[47] T. Kariya and H. Kurata, *Generalized Least Squares*, Wiley Series in Probability and Statistics (Wiley, Chichester, 2004).
[48] V. P. Godambe and C. C. Heyde, Int. Statist. Rev. **55**, 231 (1987).
[49] D. Dacunha-Castelle and D. Florens-Zmirou, Stochastics **19**, 263 (1986).
[50] R. W. Berg, S. Ditlevsen, and J. Hounsgaard, PLoS ONE **3**, e3218 (2008).
[51] P. E. Kloeden and E. Platen, *Numerical Solution Of Stochastic Differential Equations* (Springer-Verlag, Berlin, 1992), Vol. 23.
[52] M. T. Giraudo and L. Sacerdote, Commun. Stat.-Simul. Comput. **28**, 1135 (1999).