Inferential framework for nonstationary dynamics. I. Theory

Dmitri G. Luchinsky,[1,2] Vadim N. Smelyanskiy,[1] Andrea Duggento,[2] and Peter V. E. McClintock[2]
[1]*NASA Ames Research Center, Mail Stop 269-2, Moffett Field, California 94035, USA*
[2]*Department of Physics, Lancaster University, Lancaster LA1 4YB, United Kingdom*
(Received 30 January 2008; published 4 June 2008)

A general Bayesian framework is introduced for the inference of time-varying parameters in nonstationary, nonlinear, stochastic dynamical systems. Its convergence is discussed. The performance of the method is analyzed in the context of detecting signaling in a system of neurons modeled as FitzHugh-Nagumo (FHN) oscillators. It is assumed that only fast action potentials for each oscillator mixed by an unknown measurement matrix can be detected. It is shown that the proposed approach is able to reconstruct unmeasured (hidden) variables of the FHN oscillators, to determine the model parameters, to detect stepwise changes of control parameters for each oscillator, and to follow continuous evolution of the control parameters in the adiabatic limit.

DOI: 10.1103/PhysRevE.77.061105 PACS number(s): 02.50.Tt, 05.45.Tp, 05.10.Gg, 05.45.Xt

I. INTRODUCTION

Complex phenomena in nature and technology can often be modeled successfully by stochastic nonlinear dynamical systems, thereby facilitating the diagnosis of faults, the prognosis of future conditions, and control. Examples range from models of reactors [1] and helioseismology [2] to models used in physiology [3] and neuroscience [4]. The problem of inferring the parameters of such models from time-series data has therefore attracted much attention over the past decade [3,5–13]. In general, important control parameters of the systems in question vary in time so that the system dynamics is nonstationary. It is highly desirable, therefore, to extend the inferential framework to encompass almost-real-time tracking of time-varying parameters of nonstationary systems.

Most of the algorithms rely heavily on extensive numerical simulation [9,10,13], or require a large amount of data [5,7] (cf. econometric series analysis [14]), and cannot easily be adapted for parameter tracking in nonstationary stochastic nonlinear systems. More importantly, most earlier calculations of flows produce biased estimates because they lack a term related to the Jacobian of transformation from stochastic to deterministic variables. The term in question gives [15] a leading-order contribution to the inference results in the presence of strong dynamical noise.

We recently introduced an analytic solution of the dynamical inference problem [15,16] based on Bayesian statistics and a path-integral formulation of stochastic nonlinear dynamics. It allows for fast, unbiased estimation of the model parameters, provides optimal compensation for the dynamical noise, and paves the way to almost-real-time tracking of time-varying parameters. There are, however, two important features that have not hitherto been considered: measurement noise and nonstationarity of the dynamics. They are often important in practice.

In this paper, we demonstrate how the Bayesian framework can be extended to infer information encoded in time-varying control parameters of a nonlinear nonstationary system, almost in real time. In paper II [17], immediately following this paper, we consider an application of the scheme to a model of physiological signaling.

Such an inferential framework can have a wide range of interdisciplinary applications ranging from aerospace [18,19] to nanosensors. In particular, it can be especially advantageous in the analysis of signals from neuronal systems. Their dynamical details are known only approximately. Internal and measurement noises exert strong influences, and the time variation of the control parameters is directly related to information coding. We focus on physiological applications, therefore, and consider as an example the inference of time-varying control parameters from the measurements of the spiking dynamics of neural systems. The neural system is modeled by a set of FitzHugh-Nagumo (FHN) equations [20–23], a system that has been found useful in modeling nerve fibers [24] and certain muscle cells, e.g., in the heart tissue [25–27]. It has also been used intensively in studies of passive myelinated axons [28] and various forms of arrhythmia and cardiac activation evolution [29]. The highly nonlinear and nonstationary nature of the system dynamics makes it difficult to apply standard techniques for the reliable inference of control parameters.

We will show that our approach is able to decode the time evolution of the control parameters in a system of neurons modeled as FHN oscillators, including detection of their large stepwise changes for each oscillator and continuous variation in the adiabatic limit. Because the method is based on nonlinear dynamical inference, the parameter-tracking algorithm is effectively embedded into the learning inferential framework, enabling us to reconstruct both the unmeasured (hidden) variables of the FHN oscillators and the model parameters. To illustrate this point, we will reconstruct the system parameters assuming that the original parameters of the model are unknown, that only one coordinate of each oscillator is available for recording, and that these measurements are mixed by a measurement matrix.

The paper is organized as follows. Section II presents the theory of Bayesian dynamical inference in the presence of dynamical and measurement noises. An example of global optimization both in the parameter and path spaces is provided in Sec. III. In Sec. IV, the theory of Bayesian inference for a system of L FHN oscillators is presented, providing the basis for physiological applications. In Sec. V, the results are summarized and conclusions are drawn.

II. BAYESIAN INFERENTIAL FRAMEWORK FOR NONSTATIONARY DYNAMICS

A. A general approach

Let us consider the following problem. Given M-dimensional time-series data $\mathcal{Y}=\{\mathbf{y}_n \equiv \mathbf{y}(t_n)\}$ $(t_n=nh)$, how can one infer the time variation of the unknown model parameters and the unknown dynamical trajectory $\mathcal{M} =\{\mathbf{c}(t),\mathbf{b}(t),\hat{\mathbf{D}},\hat{\mathbf{M}},\{\mathbf{x}_n\}\}$? It is assumed that the underlying dynamics can be described by a set of L-dimensional ($L \geq M$) stochastic differential equations of the form

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}|\mathbf{c}) + \sqrt{\hat{\mathbf{D}}}\boldsymbol{\xi}(t), \tag{1}$$

and the observations $\mathcal{Y}$ are related to the actual unknown dynamical variables $\mathcal{X}=\{\mathbf{x}_n \equiv \mathbf{x}(t_n)\}$ via the following measurement equation:

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{x}|\mathbf{b}) + \sqrt{\hat{\mathbf{M}}}\boldsymbol{\eta}(t). \tag{2}$$

Here $\hat{X}$ is an $M\times L$ measurement matrix, $\boldsymbol{\xi}(t)$ and $\boldsymbol{\eta}(t)$ are L- and M-dimensional Gaussian white noises, and $\hat{\mathbf{D}}$ and $\hat{\mathbf{M}}$ are $(L\times L)$- and $(M\times M)$-dimensional dynamical and measurement diffusion matrices, respectively.

The problem is essentially stochastic and nonlinear and its solution is given by the so-called *posterior* density $\rho_{\text{post}}(\mathcal{M}|\mathcal{Y})$ of the unknown parameters $\mathcal{M}$ conditioned on observations. The latter can be calculated in general form via Bayes' theorem,

$$\rho_{\text{post}}(\mathcal{M}|\mathcal{Y}) = \frac{\ell(\mathcal{Y}|\mathcal{M})\rho_{\text{prior}}(\mathcal{M})}{\int \ell(\mathcal{Y}|\mathcal{M})\rho_{\text{prior}}(\mathcal{M})d\mathcal{M}}. \tag{3}$$

Here the *prior* density $\rho_{\text{prior}}(\mathcal{M})$ accumulates expert knowledge of the unknown parameters and the *likelihood* function $\ell(\mathcal{Y}|\mathcal{M})$ is the probability density to observe $\{\mathbf{y}_n(t)\}$ given choice $\mathcal{M}$ of the dynamical model. Thus within the Bayesian framework, the problem is reduced to the calculation of the likelihood function and optimization of the posterior distribution with respect to $\mathcal{M}$. If the sampling is dense enough, the problem can be conveniently solved using Euler midpoint discretization of Eqs. (1) and (2) in the form

$$\mathbf{x}_{n+1} = \mathbf{x}_n + h\mathbf{f}(\mathbf{x}_n^*|\mathbf{c}) + h\sqrt{\hat{\mathbf{D}}}\boldsymbol{\eta}_n,$$

$$\mathbf{y}_n = \mathbf{g}(\mathbf{x}_n|\mathbf{b}) + \sqrt{\hat{\mathbf{M}}}\boldsymbol{\eta}_n, \tag{4}$$

where $\mathbf{x}_n^*=(\mathbf{x}_{n+1}+\mathbf{x}_n)/2$. It was shown earlier (see, e.g., [15,30]) that for independent sources of white Gaussian noise in Eq. (4), the probability to observe $\mathbf{y}_{n+1}$ at each time step can be factorized and written in the form

$$\begin{aligned}\rho(\mathbf{y}_{n+1}|\mathbf{x}_n,\mathbf{c}) = \int & \frac{1}{\sqrt{(2\pi)^M|\hat{\mathbf{M}}|}}\exp\biggl(-\frac{1}{2}[\mathbf{y}_{n+1}-\mathbf{g}(\mathbf{x}_{n+1}|\mathbf{b})]^{\mathrm{T}} \\ & \times\hat{\mathbf{M}}^{-1}[\mathbf{y}_{n+1}-\mathbf{g}(\mathbf{x}_{n+1}|\mathbf{b})]\biggr) \\ & \times \frac{1}{\sqrt{(2\pi h)^L|\hat{\mathbf{D}}|}}\exp\biggl(-\frac{h}{2}[\dot{\mathbf{x}}_n-\mathbf{f}(\mathbf{x}_n^*|\mathbf{c})]^{\mathrm{T}} \\ & \times\hat{\mathbf{D}}^{-1}[\dot{\mathbf{x}}_n-\mathbf{f}(\mathbf{x}_n^*|\mathbf{c})]-\frac{h}{2}\nabla\cdot[\mathbf{f}(\mathbf{x}_n)|\mathbf{c}]\biggr)d\mathbf{x}_{n+1}.\end{aligned} \tag{5}$$

Summation over all the discretization points $n=0,\ldots,N-1$ yields the following result for the minus log-likelihood function $S=S_{\text{dyn}}+S_{\text{meas}}=-\ln\ell(\mathcal{Y}|\mathcal{M})$:

$$\begin{aligned}S = {} & \frac{N}{2}\ln|\hat{\mathbf{D}}| + \frac{h}{2}\sum_{n=0}^{N-1}\Bigl\{\nabla\cdot[\mathbf{f}(\mathbf{x}_n)|\mathbf{c}] \\ & + [\dot{\mathbf{x}}_n-\mathbf{f}(\mathbf{x}_n^*|\mathbf{c})]^{\mathrm{T}}\hat{\mathbf{D}}^{-1}[\dot{\mathbf{x}}_n-\mathbf{f}(\mathbf{x}_n^*|\mathbf{c})]\Bigr\} + \frac{N}{2}\ln|\hat{\mathbf{M}}| \\ & + \frac{1}{2}\sum_{n=1}^{N}[\mathbf{y}_n-\mathbf{g}(\mathbf{y}_n,\mathbf{x}_n|\mathbf{b})]^{\mathrm{T}}\hat{\mathbf{M}}^{-1}[\mathbf{y}_n-\mathbf{g}(\mathbf{y}_n,\mathbf{x}_n|\mathbf{b})] \\ & + (L+M)N\ln(2\pi h),\end{aligned} \tag{6}$$

where $\dot{\mathbf{x}}_n=\frac{\mathbf{x}_{n+1}-\mathbf{x}_n}{h}$. Here S_{dyn} and S_{meas} are the dynamical (first two terms) and measurement (next two terms) parts of the minus log-likelihood function. We note that S_{dyn} is the minus log-probability density in the space of dynamical paths and, in the limit of $N\to\infty$, $h\to 0$, $T=Nh=$ const, it coincides with the path-integral presentation obtained earlier in [31].

Equations (1)–(3) and (6) provide a general Bayesian framework for learning the state and the model of the system (1) as well as for learning the parameters of the measurement model (2) and for tracking the variation of the parameters of the system in time. It can readily be extended to encompass inertial measurement described by the following model:

$$\dot{\mathbf{y}} = \mathbf{g}(\mathbf{y},\mathbf{x}|\mathbf{b}) + \sqrt{\hat{\mathbf{M}}}\boldsymbol{\eta}(t).$$

In the latter case, S_{meas} has a form that is similar to S_{dyn}, as will be described in more detail elsewhere (see also [32]).

To find the general solution of the problem (1) and (2), one can iterate optimization of S in the space of dynamical paths $\{\mathbf{x}_n\}$ and in the space of parameters $\{\mathbf{c},\mathbf{b},\hat{\mathbf{D}},\hat{\mathbf{M}}\}$ (see [30]).

Let us assume that the optimal paths corresponding to the hidden dynamical variables $\{\mathbf{x}_n\}$ are found on the current step of iterations (see, for example, Sec. III). At the next step of iterations, the values of the model parameters $\{\mathbf{c},\mathbf{b},\hat{\mathbf{D}},\hat{\mathbf{M}}\}$ can be updated using the following equations (cf. with [15]):

$$\mathbf{f}(\mathbf{x}|\mathbf{c}) = \hat{\mathbf{F}}(\mathbf{x})\mathbf{c}, \quad \mathbf{g}(\mathbf{y},\mathbf{x}|\mathbf{b}) = \hat{\mathbf{G}}(\mathbf{y},\mathbf{x})\mathbf{b}, \tag{7}$$

where matrices $\hat{\mathbf{F}}$ and $\hat{\mathbf{G}}$ have the form

$$\hat{\mathbf{F}} = \left[\begin{pmatrix} \phi_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \phi_1 \end{pmatrix} \dots \begin{pmatrix} \phi_F & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \phi_F \end{pmatrix} \right], \tag{8}$$

$$\hat{\mathbf{G}} = \left[\begin{pmatrix} \psi_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \psi_1 \end{pmatrix} \dots \begin{pmatrix} \psi_G & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \psi_G \end{pmatrix} \right], \tag{9}$$

and $\{\phi_i\}$ and $\{\psi_i\}$ are the F- and G-dimensional sets of arbitrary base functions.

Choosing prior PDFs for $\mathbf{c}$ and $\mathbf{b}$ in the form of Gaussian distributions, and uniform prior PDFs for $\hat{\mathbf{D}}$ and $\hat{\mathbf{M}}$, the following equations can be obtained to update model parameters (cf. with [15]):

$$\langle \hat{\mathbf{D}} \rangle = \frac{h}{N} \sum_{n=0}^{N-1} (\dot{\mathbf{x}}_n - \hat{\mathbf{F}}_n \mathbf{c})(\dot{\mathbf{x}}_n - \hat{\mathbf{F}}_n \mathbf{c})^{\mathrm{T}}, \tag{10}$$

$$\langle \mathbf{c} \rangle = \hat{\Xi}_{\mathcal{X}}^{-1}(\hat{\mathbf{D}}) \mathbf{w}_{\mathcal{X}}(\hat{\mathbf{D}}), \tag{11}$$

$$\mathbf{w}_{\mathcal{X}}(\hat{\mathbf{D}}) = h \sum_{n=0}^{N-1} \left(\hat{\mathbf{F}}_n^{\mathrm{T}} \mathbf{D}^{-1} \dot{\mathbf{x}}_n - \frac{\mathbf{v}(\mathbf{x}_n)}{\mathbf{2}} \right), \tag{12}$$

$$\hat{\Xi}_{\mathcal{X}}(\hat{\mathbf{D}}) = h \sum_{n=0}^{N-1} \hat{\mathbf{F}}_n^{\mathrm{T}} \hat{\mathbf{D}}^{-1} \hat{\mathbf{F}}_n, \tag{13}$$

where $\hat{\mathbf{F}}_n \equiv \hat{\mathbf{F}}(\mathbf{x}_n)$, and the components of the vector $\mathbf{v}(\mathbf{x})$ are

$$v_m(\mathbf{x}) = \sum_{l=1}^{L} \frac{\partial F_{lm}(\mathbf{x})}{\partial x_l}, \quad m = 1, \dots, F. \tag{14}$$

The parameters of the measurement model can be estimated using the conditions $\frac{\partial S_{\text{meas}}}{\partial \mathbf{b}} = 0$ and $\frac{\partial S_{\text{meas}}}{\partial M_{nm}} = 0$, recovering the least-square results in the form

$$\langle \hat{\mathbf{M}} \rangle = \frac{1}{N} \sum_{n=1}^{N} [\mathbf{y}_n - \hat{\mathbf{G}}_n \mathbf{b}][\mathbf{y}_n - \hat{\mathbf{G}}_n \mathbf{b}]^{\mathrm{T}}, \tag{15}$$

$$\langle \mathbf{b} \rangle = \hat{\Theta}_{\mathcal{X},\mathcal{Y}}^{-1}(\hat{\mathbf{M}}) \mathbf{z}_{\mathcal{X},\mathcal{Y}}(\hat{\mathbf{M}}), \tag{16}$$

$$\mathbf{z}_{\mathcal{X},\mathcal{Y}}(\hat{\mathbf{M}}) = \sum_{n=1}^{N} [\hat{\mathbf{G}}_n^{\mathrm{T}} \hat{\mathbf{M}}^{-1} \mathbf{y}_n], \tag{17}$$

$$\hat{\Theta}_{\mathcal{X},\mathcal{Y}}^{-1}(\hat{\mathbf{M}}) = \sum_{n=0}^{N-1} \hat{\mathbf{G}}_n^{\mathrm{T}} \hat{\mathbf{M}}^{-1} \hat{\mathbf{G}}_n, \tag{18}$$

where $\hat{\mathbf{G}}_n \equiv \hat{\mathbf{G}}(\mathbf{y}_n, \mathbf{x}_n)$.

Equations (10)–(18), coupled with the optimization procedure in the paths' space, represent the general Bayesian framework for learning a nonlinear stochastic dynamical system from measurements that are corrupted by noise. Using this approach, we can develop a method of fast on-line tracking of the time-varying parameters of nonstationary systems, as described below.

B. The main idea of the inferential framework for nonstationary dynamics

The main idea of the method is to apply Eqs. (10)–(18) within a window moving along the time trace of the experimental data, resulting in time-resolved inference of the model parameters. The time resolution of the method is limited by the convergence of the model parameters, but can be improved substantially if we note that only a few parameters of the system have to be followed in time continuously, while the rest of the model parameters can be learned efficiently from a block of stationary data from the time series.

Indeed, in many practical applications, the majority of the system parameters remain constant and only a few control parameters vary in time. To achieve the best time resolution, we introduce a two-step procedure, in which the tracking of time-varying parameters is embedded into a Bayesian learning framework. As the first step, we select an interval of the experimental time traces during which the system dynamics is stationary and learn model parameters. In the second step, we assume that the majority of the parameters of the system remain constant and track only a few time-varying control parameters.

To clarify this idea, one has to take into account various characteristic time scales of the problem. The measured time series are characterized by the sampling time step h and the total measurement time $T_{\text{meas}} = nh$, where n is the number of measured points. The system dynamics has an intrinsic characteristic time scale τ_0 and characteristic time scales of slow τ_{slow} and fast τ_{fast} variation of the model parameters. For the FitzHugh-Nagumo model, τ_0 can be taken as equal to the width of the spike. The time resolution of the method is characterized by the convergence time of the inference $\tau_{\text{res}}(\mathbf{c})$. Note that $\tau_{\text{res}}(\mathbf{c})$ depends on the set of unknown model parameters.

For the method to be applicable, the characteristic time scale for slow variation of the model parameters has to be larger than measurement time, $\tau_{\text{slow}} > T_{\text{meas}}$. In this adiabatic approximation, slowly varying parameters can be assumed constant. In the first step, it is further assumed that there exists a time trace of length $T > \tau_{\text{res}}(\mathbf{c})$ where all the parameters of the system can be considered to remain constant. Equations (10)–(18) can then be used to learn the slowly varying parameters of the model together with parameters of the measurement model $\mathbf{b}$. These parameters, once inferred, are considered to be constant and known. In the second step, the set of model parameters is divided into known $\mathbf{c}^{(k)}$ and unknown $\mathbf{c}^{(u)}$ subsets. To infer fast-varying control parameters, one can use Eqs. (10)–(13) substituting $\dot{x}$ with $(\dot{x} - \hat{\mathbf{F}}_n \mathbf{c}^{(k)})$ and $\hat{\mathbf{F}}_n \mathbf{c}$ with $\hat{\mathbf{F}}_n \mathbf{c}^{(u)}$. The possibility of fast on-line tracking of the control parameters arises in this approach due to the fact that $\tau_{\text{res}}(\mathbf{c}^{(u)})$ can be made much shorter than $\tau_{\text{res}}(\mathbf{c})$.

In the context of the present research, we are interested mainly in the on-line tracking of the parameters of the physiological signals that can be modeled as a system of a set of FitzHugh-Nagumo oscillators mixed by a measurement matrix. In the following sections, we will introduce a specific example of a system that can be used to model physiological measurements. Our primary goal will be to establish whether the convergence of the model parameters is sufficiently fast to allow for the on-line tracking (decoding physiological information in real time) of the control parameters of the model. This will be demonstrated in paper II [17].

Next, we provide further arguments related to the convergence of the algorithm. We start by assuming that there is no measurement noise, so that we can avoid the need for optimization in the paths' space. We then provide a brief example of the optimization procedure in the paths' space.

C. Convergence of the model parameters

So let us neglect measurement noise, assuming that time traces of the state variables can be measured directly, that we have K blocks of data, and that we are interested only in the inference of the model parameters $\{c\}$. Even in this case, the convergence of the model parameters depends essentially on the specific system under consideration. However, a few general remarks may be helpful in clarifying the issues to be addressed. Note that each block of data can be measured independently and used at the step $k+1$ of inference ($k=1,\ldots,K$) provided that the results at previous steps are taken into account in the form of a prior distribution

$$p_k(\{c\}) \propto \exp\left[-\frac{1}{2}(\mathbf{c}-\mathbf{c}_k)^{\mathrm{T}}\hat{\boldsymbol{\Xi}}_k(\mathbf{c}-\mathbf{c}_k)\right]. \tag{19}$$

Equations (12) and (13) can be then written in the form (see [15])

$$\mathbf{w}_k = \hat{\boldsymbol{\Xi}}_{k-1}^{-1}\mathbf{c}_{k-1} + h\sum_{n\in N_k}\left(\hat{\mathbf{F}}_n^{\mathrm{T}}\hat{\mathbf{D}}^{-1}\dot{\mathbf{x}}_n - \frac{\mathbf{v}_n}{2}\right), \tag{20}$$

$$\hat{\boldsymbol{\Xi}}_k = \hat{\boldsymbol{\Xi}}_{k-1} + h\sum_{n\in N_k}\hat{\mathbf{F}}_n^{\mathrm{T}}\hat{\mathbf{D}}^{-1}\hat{\mathbf{F}}_n. \tag{21}$$

It is clear that the covariance matrix of the *posterior* distribution is a sum over all the blocks and has the structure of a Kronecker product,

$$\hat{\boldsymbol{\Xi}}_k = \hat{\boldsymbol{\Phi}} \otimes \hat{\mathbf{Q}}, \tag{22}$$

where

$$\hat{\boldsymbol{\Phi}} = \sum_{n\in N_1,\ldots,N_k}\begin{pmatrix}\psi_{1,n}\psi_{1,n} & \cdots & \psi_{1,n}\psi_{B,n}\\ \vdots & \ddots & \vdots\\ \psi_{B,n}\psi_{1,n} & \cdots & \psi_{B,n}\psi_{B,n}\end{pmatrix},$$

$\psi_{i,n}\equiv\psi_i(x_n)$, and $\hat{\mathbf{Q}}=\hat{\mathbf{D}}^{-1}$. Accordingly, all nonzero elements of this matrix grow in time as $T=hN$. On the other hand, the second term in Eq. (21) remains finite for a fixed number of points in one block N_k. Therefore, $\hat{\boldsymbol{\Xi}}_{k-1}$ approaches $\hat{\boldsymbol{\Xi}}_k$ for large T and c_k becomes

$$c_k \approx c_{k-1} + \hat{\mathbf{D}} \otimes \hat{\boldsymbol{\Phi}}_k^{-1}\sum_{n\in N_k}\left(\hat{\mathbf{F}}_n^{\mathrm{T}}\hat{\mathbf{D}}^{-1}\dot{\mathbf{x}}_n - \frac{\mathbf{v}_n}{2}\right). \tag{23}$$

The convergence of c_k is thus determined by the convergence of residuals in Eq. (23). Clearly, convergence of the residuals is proportional to the sum of eigenvalues of $\hat{\boldsymbol{\Phi}}_k^{-1}$: the presence of large eigenvalues slows down the convergence of all coefficients $\{c_k\}$.

At the same time, base functions related to the control parameters have a strong effect on the dynamics of the system and usually correspond to large eigenvalues of $\hat{\boldsymbol{\Phi}}$. Therefore, to achieve the best results in decoding nonstationary dynamics, one can use the general dynamical inference framework introduced above to learn most of the stationary model parameters in a preliminary analysis of the system. Next, by incorporating real-time inference into this inferential learning framework and excluding all but the most important nonstationary parameters from the tracking procedure, one can improve the time resolution of the method by orders of magnitude.

We note that to exclude the learned model parameters from further analysis, one has to separate the vector field into two parts $\mathbf{f}(\mathbf{x}|\mathbf{c})=\mathbf{f}'(\mathbf{x}|\mathbf{c}_{\text{known}})+\mathbf{f}''(\mathbf{x}|\mathbf{c}_{\text{unknown}})$ and to use $[\dot{\mathbf{x}}-\mathbf{f}'(\mathbf{x}|\mathbf{c}_{\text{known}})]$ and $\mathbf{f}''(\mathbf{x}|\mathbf{c}_{\text{unknown}})$ instead of $\dot{\mathbf{x}}$ and $\mathbf{f}(\mathbf{x}|\mathbf{c})$, respectively, in Eqs. (10) and (12). With this trivial modification, the method will allow for fast on-line tracking of the parameters of nonstationary nonlinear dynamical systems.

In the context of physiological applications, polynomial base functions and relatively small noise intensities are of special interest. It is clear that in this case the smallest elements of $\hat{\boldsymbol{\Phi}}$ correspond to the highest powers of polynomials. In particular, in the case of a symmetric one-dimensional (1D) potential, the contribution of the polynomials of order m to the diagonal terms of $\hat{\boldsymbol{\Phi}}$ will be proportional to $\langle x^m\rangle \propto D^m$. Accordingly, if the coefficients of the polynomials of power 3 can be learned before applying the on-line tracking procedure, the time resolution of the method can ideally be improved by three orders of magnitude for typical values of $D\approx 0.1$. This effect will be demonstrated in paper II [17] using as an example a system of mixed FHN oscillators. In the next section, we provide a brief example of the global optimization procedure, including optimization in the trajectory space, that can be used to learn model parameters in a general case before applying time-resolved methods.

III. GLOBAL NONLINEAR OPTIMIZATION IN THE SPACE OF DYNAMICAL TRAJECTORIES

A global optimization of the cost function in the space of the model trajectories of nonlinear stochastic dynamical systems is a long-standing problem in the theory of stochastic processes. A number of methods have been suggested earlier to solve this problem, including methods based on the Markov chain Monte Carlo (MCMC) [33], the extended Kalman filter [10], and the Langevin method of sampling the posterior [32]. In our earlier work, we generalized an extended Kalman filter approach [30] and the MCMC [34] to sample the paths of continuous nonlinear stochastic systems. Here we will describe briefly a method based on global nonlinear optimization utilizing an explicit analytic form of the gradient and the Hessian of the cost function and the conjugate-gradient method (see also [35]).

Indeed, let us consider the case in which $\mathbf{g}(\mathbf{y}_n,\mathbf{x}_n|\mathbf{b})=\hat{\boldsymbol{\Gamma}}\mathbf{x}_n$. In this case, the Hessian $\hat{\mathbf{H}}$ of S takes the form

$$\hat{\mathbf{H}} = \begin{pmatrix} \hat{\mathbf{B}}_{00} & \hat{\mathbf{B}}_{01} & \hat{\mathbf{0}} & \cdots & \hat{\mathbf{0}} & \hat{\mathbf{0}} \\ \hat{\mathbf{B}}_{10} & \hat{\mathbf{B}}_{11} & \hat{\mathbf{B}}_{12} & \cdots & \hat{\mathbf{0}} & \hat{\mathbf{0}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{\mathbf{0}} & \hat{\mathbf{0}} & \hat{\mathbf{0}} & \cdots & \hat{\mathbf{B}}_{n-1,n} & \hat{\mathbf{B}}_{nn} \end{pmatrix}, \tag{24}$$

where matrices $\hat{\mathbf{B}}_{nm}$ are given by the following expression:

$$\begin{aligned} \hat{\mathbf{B}}_{nm} = \frac{\partial S}{\partial \mathbf{x}_n \, \partial \mathbf{x}_m} &= \hat{\mathbf{\Gamma}}^{\mathrm{T}} \hat{\mathbf{M}}^{-1} \hat{\mathbf{\Gamma}} \delta_{nm} + \frac{h}{2} \frac{\partial}{\partial \mathbf{x}_n \, \partial \mathbf{x}_m} [\mathrm{tr} \hat{\mathbf{\Phi}}(\mathbf{x}_n|\mathbf{c})] \delta_{nm} \\ &+ \frac{1}{h} \hat{\mathbf{D}}^{-1} \delta_{nm} + h \left(\hat{\mathbf{I}}/h + \frac{\partial \mathbf{f}}{\partial \mathbf{x}_n} \right)^{\mathrm{T}} \hat{\mathbf{D}}^{-1} \left(\hat{\mathbf{I}}/h + \frac{\partial \mathbf{f}}{\partial \mathbf{x}_n} \right) \delta_{nm} \\ &- h [\dot{\mathbf{x}}_n - \mathbf{f}(\mathbf{x}_n^*|\mathbf{c})]^{\mathrm{T}} \hat{\mathbf{D}}^{-1} \left[\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}_n^2} \right] \delta_{nm} \\ &- h \left(\hat{\mathbf{I}}/h + \frac{\partial \mathbf{f}}{\partial \mathbf{x}_{n-1}} \right)^{\mathrm{T}} \hat{\mathbf{D}}^{-1} \delta_{m,n-1} \\ &- h \hat{\mathbf{D}}^{-1} \left(\hat{\mathbf{I}}/h + \frac{\partial \mathbf{f}}{\partial \mathbf{x}_n} \right) \delta_{m,n+1} \end{aligned} \tag{25}$$

and the gradient $\frac{\partial S}{\partial \mathbf{x}_n}$ of S has the form

$$\begin{aligned} \frac{\partial S}{\partial \mathbf{x}_n} &= -(\mathbf{y}_n - \hat{\mathbf{\Gamma}} \mathbf{x}_n)^{\mathrm{T}} \hat{\mathbf{M}}^{-1} \hat{\mathbf{\Gamma}} \, \partial \mathbf{x}_n \\ &\quad + \frac{h}{2} \frac{\partial}{\partial \mathbf{x}_n} \{ \nabla \cdot [\mathbf{f}(\mathbf{x}_n)|\mathbf{c}] \} + [\dot{\mathbf{x}}_{n-1} - \mathbf{f}(\mathbf{x}_{n-1}^*|\mathbf{c})]^{\mathrm{T}} \hat{\mathbf{D}}^{-1} \\ &\quad - [\dot{\mathbf{x}}_n - \mathbf{f}(\mathbf{x}_n^*|\mathbf{c})]^{\mathrm{T}} \hat{\mathbf{D}}^{-1} \left(\hat{\mathbf{I}} + h \frac{\partial \mathbf{f}(\mathbf{x}_n^*|\mathbf{c})}{\partial \mathbf{x}_n} \right). \end{aligned} \tag{26}$$

Given the form of the gradient $\frac{\partial S}{\partial \mathbf{x}_n}$ and of the Hessian $\hat{\mathbf{H}}$, global optimization in the paths' space can be performed especially efficiently.

Given a set of noisy observations $\mathcal{Y}$, we first minimize S with respect to $\mathcal{X}$ keeping the model parameters fixed. According to the standard conjugate gradient procedure [36], we do the following steps:

(i) Choose initial values for the state vector $\mathcal{X}_0$ and choose initial directions $\mathbf{d}_0 = -\nabla S(\mathcal{X}_0)$.

(ii) Update values of the coordinates using $\mathcal{X}_1 = \mathcal{X}_0 + \alpha \mathbf{d}_0$, where

$$\alpha = -[\mathbf{d}_0^{\mathrm{T}} \nabla S(\mathcal{X}_0)]/[\mathbf{d}_0^{\mathrm{T}} \hat{\mathbf{H}}(\mathcal{X}_0) \mathbf{d}_0].$$

(iii) Update direction, using

$$\mathbf{d}_1 = -\nabla S(\mathcal{X}_1) + \frac{|\nabla S(\mathcal{X}_1)|^2}{|\nabla S(\mathcal{X}_0)|^2} \mathbf{d}_0.$$

Once the conjugate gradient algorithm has converged to some global minimum $\mathcal{X}$ in the space of dynamical trajectories, we use this trajectory to infer model parameters by applying Eqs. (10)–(13).

Consider as an example a nonlinear system with a stable limit cycle in the form

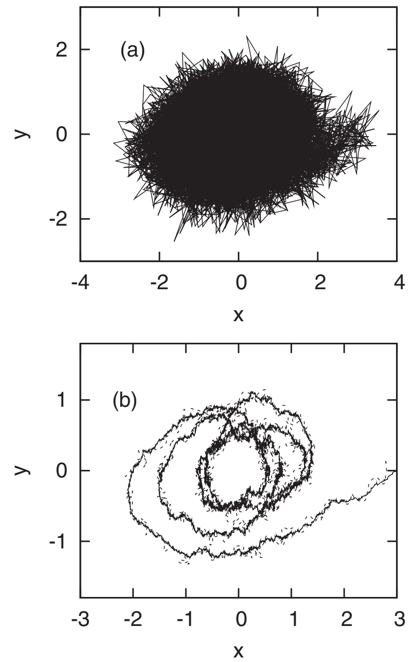

FIG. 1. (a) Example of corrupted-by-noise measurements (27) with intensity of dynamical noise $\langle \xi_x^2(t) \rangle = 0.1$ and $\langle \xi_y^2(t) \rangle = 0.2$ and the amplitude of measurement noise 0.4 in both coordinates. (b) Recovered stochastic dynamics of the system (27) (dotted line) is shown in comparison with the actual dynamical trajectory (solid line).

$$\begin{aligned} \dot{x}_1 &= x_2 - x_1^2 x_2 + \xi_1(t), \\ \dot{x}_2 &= -x_1 + 0.1(1 - x_1^2) x_2 + \xi_2(t). \end{aligned} \tag{27}$$

The state of the dynamical system is unknown. We assume for simplicity that both coordinates were measured with measurement noise of amplitude 0.4 in both coordinates, i.e.,

$$y_i(t) = x_i(t) + 0.4 \nu_i(t), \quad i = 1, 2.$$

Here the measurement matrix has the form $\hat{\mathbf{\Gamma}} = \hat{\mathbf{I}}$ and the measurement noise matrix has the form $\hat{\mathbf{M}} = \mathbf{0.16} \hat{\mathbf{I}}$.

We further assume that the vector field of Eq. (27) is unknown and model it using the following set of eight known base functions:

$$\Phi = \{1; x_1; x_2; x_1^2; x_2^2; x_1 x_2; x_1^3; x_1^2 x_2\}.$$

In explicit form, the model of the limit cycle system (27) is

$$\dot{x}_1 = \sum_{i=1}^{8} c_{2i-1} \phi_i + \xi_1(t), \quad \dot{x}_2 = \sum_{i=1}^{8} c_{2i} \phi_i + \xi_2(t). \tag{28}$$

We now apply the algorithm described in the previous section to infer both the unknown state and the vector field of this system. An example of noise-corrupted measurements of the system (27) is shown in Fig. 1(a). Our technique allows recovery of the stochastic dynamics of the system (27) as shown in Fig. 1(b) and to estimate model parameters. The results of the estimations are shown in Table I. It is evident that the optimization in the paths' space can be performed

TABLE I. Convergence of some coefficients of the system (28). We have used one block of data with 40 000 points.

Coefficients	True values	Inferred values	Updated values
c_3	1	1.46	1.08
c_4	0	−42.64	0.02
c_{10}	−1	−10.93	−1.07
c_{11}	0.1	−35.68	0.258
c_{15}	0	1.82	0.005
c_{16}	−0.1	−27.96	−0.17
D_{11}	0.04	325	0.045
D_{12}	0	6	0.006
D_{22}	0.04	318	0.03

efficiently in the presence of measurement noise. In practice, however, the conjugate gradient algorithm requires about 20 steps for convergence. The complexity of such an algorithm (calculated as a number of matrix operations) is much higher that the complexity of the calculations of the model parameters using Eqs. (10)–(13). Accordingly, for the fast on-line applications of the algorithm one should avoid global optimization in the space of dynamical trajectories by, e.g., suppressing measurement noise.

The main focus in the remaining part of the research will be the development of the fast on-line tracking method for the time-varying parameters. For the sake of simplicity of our further arguments, we therefore assume that measurement noise can be neglected. In the next section, we will introduce a specific example of a model that can be applied for the interpretation of physiological time-series data using our Bayesian inferential framework.

IV. SYSTEM OF FITZHUGH-NAGUMO OSCILLATORS

In the context of physiological applications, we consider the following dynamical model (see [17] for details of the numerical analysis of this model):

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}|\mathbf{c}) - \mathbf{q} + \sqrt{\hat{\mathbf{D}}}\boldsymbol{\xi}(t), \quad \mathbf{x} = (x_1, \ldots, x_L), \tag{29}$$

$$\dot{\mathbf{q}} = -\beta\mathbf{q} + \gamma\mathbf{x}, \tag{30}$$

representing a set of independent FitzHugh-Nagumo systems. The measurements are modeled by the following equation:

$$\mathbf{y} = \hat{\mathbf{X}}\mathbf{x}. \tag{31}$$

Note that the $\mathbf{q}$ coordinates are "hidden" or unobservable, while the $\mathbf{x}$ coordinates are accessible for measurement and are mixed by the measurement matrix $\hat{\mathbf{X}}$.

The main assumptions of the model (29)–(31) are that the measurement noise can be neglected together with the noise in Eq. (30). Under these assumptions, sampling in the paths' space can be avoided, thus paving the way to the fast on-line decoding of physiological information. Indeed, in this case Eq. (30) can be integrated,

$$\mathbf{q}(t) = \gamma\int_0^t d\tau e^{-\beta(t-\tau)}\mathbf{x}(\tau) + e^{-\beta t}\mathbf{q}(0). \tag{32}$$

Here $\mathbf{q}(0)$ is a set of initial coordinates that needs to be inferred along with the rest of the parameters. Therefore, the unobservable variables can be excluded from further consideration. According to the trapezoidal rule, the discrete version of Eq. (32) is

$$\mathbf{q}(t_k) = \gamma h\sum_{r=0}^{k} e^{-\beta(t_k-t_r)}\mathbf{x}(t_r) - \frac{h\gamma}{2}(\mathbf{x}(t_k) + e^{-\beta t_k}\mathbf{x}(t_0)) + e^{-\beta t_k}\mathbf{q}(0). \tag{33}$$

The resulting model and its discretization have the following form:

$$\mathbf{x}_{k+1} = h\mathbf{f}(\mathbf{x}_k^*|\mathbf{c}) - \gamma h^2\sum_{r=0}^{k} e^{-\beta(t_k-t_r)}\mathbf{x}(t_r) + \frac{h^2\gamma}{2}[\mathbf{x}(t_k) + e^{-\beta t_k}\mathbf{x}(t_0)] - he^{-\beta t_k}\mathbf{q}(0) + \boldsymbol{\Delta}_k + \mathcal{O}(h^2), \tag{34}$$

where $\boldsymbol{\Delta}_k = \int_{t_k}^{t_{k+1}} dt'\,\boldsymbol{\xi}(t')$ and $\mathbf{x}_k^* \equiv \frac{\mathbf{x}_{k+1}+\mathbf{x}_k}{2}$.

The FHN oscillator is a special case of the dynamics in Eqs. (29) and (30). It will be the subject of our numerical experiments,

$$\dot{v}_j = -v_j(v_j - \alpha_j)(v_j - 1) - q_j + \eta_j + d_j\xi_j,$$
$$\dot{q}_j = -\beta q_j + \gamma_j v_j,$$
$$\langle \xi_j(t)\xi_i(t')\rangle = \delta_{ij}d_i\delta(t-t'), \quad j = 1{:}L. \tag{35}$$

The system described in Eq. (35) represents the simplified dynamic of L noninteracting neurons [21], each of them labeled with j; v_j represents the membrane potential while q_j are slow recovery variables.

In practice, signals that are collected from biological systems are mixed with a measurement matrix; to tackle this problem, we assume that the measurement variable is y_i, which is a linear transformation of v_j,

$$y_i = X_{ij}v_j. \tag{36}$$

Here the mixing matrix X is an *unknown* quantity, therefore y_i contains all the accessible information. In Fig. 2, there is an example of y_i as in Eq. (36).

To write explicitly the system to be inferred, expressions of q_i from Eq. (33) and of v_i in Eq. (35) are plugged into Eq. (36). Within our inferential framework, this trajectory represents the output of the following model:

$$\dot{y}_i = A_{ij}y_j + B_{ijl}y_jy_l + C_{ijlm}y_jy_ly_m + \tilde{\eta}_i - h\sum_{r=0}^{k} e^{-\beta(t_k-t_r)}\Gamma_{il}y_l(t_r) - \sum_{r=0}^{k} e^{-\beta t_k}\tilde{z}_i + D_{ij}\xi_j + (t), \tag{37}$$

where $\tilde{z}_i$ are the components of the boundary condition $\mathbf{q}(t$

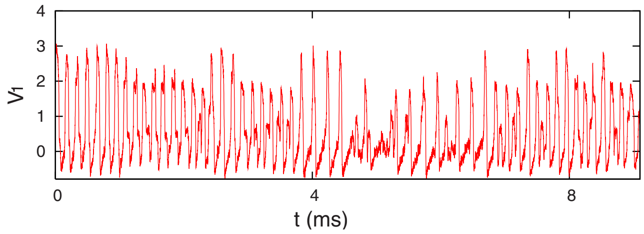

FIG. 2. (Color online) Example of a typical component $y(t)$ from Eq. (36) with mixing matrix $\binom{1\ 2}{2\ 1}$. Values of the parameters are $\alpha_1=\alpha_2=0.2$, $\eta_1=\eta_2=0.112$, $\beta=0.005\,105\,1$, $\gamma_1=\gamma_2=0.0051$, $d_1=0.001$, and $d_2=0.002$. Coefficients η_1 and η_2 change ranging from 0.05 to 0.25.

$=0$), and use of the following definitions was made:

$$A_{ij}=X_{im}\alpha_m(X^{-1})_{mj},$$

$$B_{ijl}=X_{im}(1+\alpha_m)(X^{-1})_{mj}(X^{-1})_{ml},$$

$$C_{jklm}=X_{ji}(X^{-1})_{ik}(X^{-1})_{il}(X^{-1})_{im},$$

$$\Gamma_{il}=X_{ij}\gamma_j(X^{-1})_{jl}. \tag{38}$$

In Eqs. (38), a sum over repeated indices is implied and all the indices range from 1 to L. Also, the diffusion matrix D_{ij} is expressed in terms of d_j as

$$D_{ji}=X_{ji}d_i. \tag{39}$$

Finally, Eqs. (38) contain the crucial model parameters $\tilde{\eta}_j$ that are the focus of our inference. They are related to the original model parameters η_j by

$$\tilde{\eta}_j=X_{ji}\eta_i. \tag{40}$$

We treat $y_j(t)$ in Eqs. (38) as measured variables. As a result of the inference procedure, we will recover the matrix elements of $A, B, C, \Gamma, D, \tilde{\eta}$.

The parameters of the modified and original dynamical models can be learned effectively using stationary blocks of the time-series data, as will be shown using numerical examples in paper II [17]. Once the constant parameters of the model have been learned, the algorithm will allow for very fast on-line tracking of the time-varying control parameters. Details of the convergence of the model parameters and of the time resolution of the parameter tracking will be provided in paper II, using as an example synthetic time-series data generated by the model (29)–(31).

V. CONCLUSION

Our Bayesian framework for the time-resolved inference of a nonstationary stochastic dynamical system allows for learning the parameters of the dynamical and measurement models from noise-corrupted time-series data with subsequent fast tracking of time-varying control parameters. Convergence of the method in the parameter space, and global optimization in the space of dynamical trajectories, are discussed. It is shown that to achieve the best time resolution, one has to embed the time tracking of nonstationary dynamics into an inferential learning framework that allows for preliminary inference of the model parameters in the stationary regime. Furthermore, one has to reduce the measurement noise to a low level to avoid global optimization in the trajectory space, which is necessarily time-consuming. In doing so, one can improve the time resolution of the method by several orders of magnitude. To apply this technique to the real time decoding of information from nonstationary physiological time-series data, we introduce a specific model of FHN oscillators mixed by an unknown measurement matrix. Next we show how this model can be reduced to allow for the fast on-line tracking of nonstationary parameters in a Bayesian inferential framework. A numerical investigation of this system is presented and discussed in paper II [17].

Note that for simplicity of the analysis, we have excluded dynamical noise from the equation for the recovery variable in Eq. (30) (cf., e.g., [23]). It is possible, however, to extend the proposed method to encompass the case of a stochastic linear differential equation for the hidden dynamical variable by adding a stochastic integral to the right-hand side of the reduced model (34). The corresponding extension of the method will be discussed in more detail elsewhere.

Finally, we emphasize the broad interdisciplinary applications of the method and we comment that it can readily be extended to take into account the effects of multiplicative and colored noise and of binary variables in the model.

ACKNOWLEDGMENTS

We are grateful to the Engineering and Physical Sciences Research Council (UK) and NASA for financial support, and to A. Stefanovska for valuable discussions.

[1] H. Konno and K. Hayashi, Ann. Nucl. Energy **23**, 35 (1996).
[2] J. Christensen-Dalsgaard, Rev. Mod. Phys. **74**, 1073 (2002).
[3] V. N. Smelyanskiy, D. G. Luchinsky, A. Stefanovska, and P. V. E. McClintock, Phys. Rev. Lett. **94**, 098101 (2005).
[4] E. Izhikevich, *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting* (MIT Press, Cambridge, MA, 2006).
[5] S. Siegert, R. Friedrich, and J. Peinke, Phys. Lett. A **243**, 275 (1998).
[6] P. E. McSharry and L. A. Smith, Phys. Rev. Lett. **83**, 4285 (1999).
[7] R. Friedrich *et al.*, Phys. Lett. A **271**, 217 (2000).
[8] J. P. M. Heald and J. Stark, Phys. Rev. Lett. **84**, 2366 (2000).
[9] R. Meyer and N. Christensen, Phys. Rev. E **62**, 3535 (2000).
[10] R. Meyer and N. Christensen, Phys. Rev. E **65**, 016206 (2001).
[11] J.-M. Fullana and M. Rossi, Phys. Rev. E **65**, 031107 (2002).
[12] M. Siefert, A. Kittel, R. Friedrich, and J. Peinke, Europhys.

Lett. **61**, 466 (2003).
[13] F. Watanabe, H. Konno, and S. Kanemoto, Ann. Nucl. Energy **31**, 375 (2004).
[14] W. Enders, *Applied Econometric Time Series*, 2nd ed. (Wiley, Hoboken, NJ, 2004).
[15] V. N. Smelyanskiy, D. G. Luchinsky, D. A. Timucin, and A. Bandrivskyy, Phys. Rev. E **72**, 026202 (2005).
[16] V. N. Smelyanskiy, D. G. Luchinsky, A. Stefanovska, and P. V. E. McClintock, Phys. Rev. Lett. **94**, 098101 (2005).
[17] A. Duggento *et al.*, following paper, Phys. Rev. E **77**, 061106 (2008).
[18] V. V. Osipov, D. G. Luchinsky, V. N. Smelyanskiy, and D. A. Timucin, *Proceedings of the AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, AIAA Conference Proceedings* (AIAA, Cincinnati, OH, 2007), p. 5823.
[19] D. G. Luchinsky *et al.*, *Proceedings of the AIAA Infotech@Aerospace 2007 Conference and Exhibit, AIAA Conference Proceedings* (AIAA, Robnert Park, CA, 2007), p. 2829.
[20] R. FitzHugh, Biophys. J. **1**, 445 (1961).
[21] J. Nagumo, S. Animoto, and S. Yoshizawa, Proc. IRE **50**, 2061 (1962).
[22] A. T. Winfree, *The Geometry of Biological Time* (Springer-Verlag, New York, 1980).
[23] E. V. Pankratova, A. V. Polovinkin, and B. Spagnolo, Phys. Lett. A **344**, 45 (2005).
[24] J. Keener and J. Sneyd, *Mathematical Physiology* (Springer-Verlag, New York, 1998).
[25] E. N. Best, Biophys. J. **27**, 87 (1979).
[26] J. Rogers and A. McCulloch, IEEE Trans. Biomed. Eng. **41**, 743 (1994).
[27] R. R. Aliev and A. V. Panfilov, J. Theor. Biol. **181**, 33 (1996).
[28] P. Chen, SIAM J. Math. Anal. **23**, 81 (1992).
[29] O. Berenfeld and S. Abboud, Med. Eng. Phys. **18**, 615 (1996).
[30] D. G. Luchinsky, V. N. Smelyanskiy, and J. Smith, *in Unsolved Problems of Noise and Fluctuations*, AIP Conf. Proc. No. 800, edited by L. Reggiani *et al.* (AIP, Melville, NY, 2005), pp. 539–545.
[31] R. Graham, Z. Phys. B **26**, 281 (1977).
[32] A. Apte, M. Hairer, A. M. Stuart, and J. Voss, Physica D **230**, 50 (2007).
[33] C. Calder, M. Lavine, P. Müller, and J. S. Clark, Ecology **84**, 1395 (2003).
[34] V. N. Smelyanskiy, D. G. Luchinsky, and M. Millonas, e-print arXiv:physics/0601004v1.
[35] D. G. Luchinsky, V. N. Smelyanskiy, M. Millonas, and P. V. E. McClintock, *in Noise in Complex Systems and Stochastic Dynamics III*, Vol. 5845 of *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE)*, edited by L. B. Kish, K. Lindenberg, and Z. Gingl (SPIE, Bellingham, WA, 2005), pp. 173–181.
[36] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, UK, 1992).