Inference of stochastic nonlinear oscillators with applications to physiological problems

Vadim N. Smelyanskiy[a] and Dmitry G. Luchinsky[b]

[a]NASA Ames Research Center, Mail Stop 269-2, Moffett Field, CA 94035, USA
[b]Department of Physics, Lancaster University, Lancaster LA1 4YB, UK;

ABSTRACT

A new method of inferencing of coupled stochastic nonlinear oscillators is described. The technique does not require extensive global optimization, provides optimal compensation for noise-induced errors and is robust in a broad range of dynamical models. We illustrate the main ideas of the technique by inferencing a model of five globally and locally coupled noisy oscillators. Specific modifications of the technique for inferencing hidden degrees of freedom of coupled nonlinear oscillators is discussed in the context of physiological applications.

Keywords: Inference, time-series analysis, cardio-respiratory interaction, chaotic dynamics

1. INTRODUCTION

Coupled oscillators are ubiquitous in nature. They are used to describe observed phenomena intensively over the years in many areas of science and technology including e.g. physics,[1,2] chemistry[3] and biology.[4] In this approach a complex system is characterized by projecting it onto a specific dynamical model of coupled nonlinear oscillators. However, there are no general methods to infer parameters of stochastic nonlinear models from the measured time-series data. Furthermore, in a great number of important problems the model is not usually known exactly from "first principles" and one is faced with a rather broad range of possible parametric models to consider. In addition, the experimental data can be extremely skewed, whereby important "hidden" features of a model such as coupling coefficients between the oscillators can be very difficult to extract due to the intricate interplay between noise and nonlinearity.

As was pointed out by McSharry and co-authors,[5] deterministic inference techniques[6] consistently fail to yield accurate parameter estimates in the presence of noise. The problem becomes even more complicated when both measurement noise as well as intrinsic dynamical noise are present.[7] Various numerical schemes have been proposed recently to deal with different aspects of this inverse problem.[5,7–12] A standard approach to this problem is often based on optimization of a certain cost function (a *likelihood* function) at the values of the model parameters that best reconstruct the measurements. It can be further generalized using a Bayesian formulation of the problem.[7,9] Existing techniques usually employ numerical Monte Carlo techniques for complex optimization[11] or multidimensional integration[9] tasks. Inference results from noisy observations are shown to be very sensitive to the specific choice of the likelihood function.[5] Consequently, the *correct* choice of the *likelihood function* is one of the central questions in the inference of continuous-time noise-driven dynamical models considered here.

In this paper, we present an efficient technique of Bayesian inference of nonlinear noise-driven dynamical models from time-series data that avoids extensive numerical optimization. It also guarantees optimum compensation of noise-induced errors by invoking the likelihood function in the form of a path integral over the random trajectories of the dynamical system. The technique is verified on a broad range of dynamical models including system of five globally and locally coupled nonlinear oscillators.

A specific example of inferencing stochastic nonlinear model from skewed time-series data is considered in the context of physiological research. In particular, we refer to the situation when the variability of the cardiovascular signals is modelled in terms of coupled nonlinear oscillators.[13–16] At present there are no methods available

Further author information: (Send correspondence to V.N.S.)
V.N.S.: E-mail: Vadim.N.Smelyanskiy@nasa.gov, Telephone: 1 650 604 2261

to infer parameters of the nonlinear coupling between oscillators directly from experimental time series data. Furthermore, in many situations it is important to perform such inference using univariate time series. This rises another important issue in nonlinear time-series analysis related to the inference of hidden dynamical variables. If a technique of inferencing of coupling parameters from hidden dynamical variables could be found it could provide new effective tool for estimation of the state of autonomous nervous control[17] and risk stratification of cardiovascular diseases.[18] The corresponding problem of inference of the coupling parameters of two nonlinear oscillators perturbed by noise from univariate time-series data will be considered in this paper.

The paper is organized as follows. In the Sec. 2 the algorithm is introduced and its main features are compared with the results of earlier work. In the Sec. 3 the convergence of the algorithm is analyzed in the case of inference of coupled nonlinear stochastic oscillators. A modification of the algorithm that allows inference of hidden dynamical variables of two nonlinear coupled oscillators from univariate time-series data is considered in Sec. 4.

2. THEORY OF NONLINEAR INFERENCE OF NOISE-DRIVEN DYNAMICAL SYSTEMS

Consider N-dimensional dynamical system described by set of nonlinear Langevin equations

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}) + \varepsilon(t) = \mathbf{f}(\mathbf{x}) + \sigma\xi(t), \tag{1}$$

where $\varepsilon(t)$ is an additive stationary white, Gaussian vector noise process

$$\langle \xi(t) \rangle = 0, \quad \langle \xi(t)\, \xi^T(t') \rangle = \hat{\mathbf{D}}\, \delta(t - t'), \tag{2}$$

characterized by diffusion matrix $\hat{\mathbf{D}}$.

We assume that the trajectory $x(t)$ of this system is observed at sequential time instants $\{t_k; k = 0, 1, \ldots, K\}$ and a series $\mathcal{Y} = \{y_k \equiv y(t_k)\}$ thus obtained is related to the (unknown) "true" system states $\mathcal{X} = \{x_k \equiv x(t_k)\}$ through some conditional PDF $p_o(\mathcal{Y}|\mathcal{X})$.

The most general approach to dynamical model inference is based on Bayesian framework (cf.[7]). In the Bayesian model inference, two distinct PDFs are ascribed to the set of unknown model parameters: the *prior* $p_{\rm pr}(\mathcal{M})$ and the *posterior* $p_{\rm post}(\mathcal{M}|\mathcal{Y})$, respectively representing our state of knowledge about $\mathcal{M}$ before and after processing a block of data $\mathcal{Y}$. The prior acts as a *regularizer*, concentrating the parameter search to those regions of the model space favored by our expertise and any available auxiliary information. The two PDFs are related to each other via Bayes' theorem:

$$p_{\rm post}(\mathcal{M}|\mathcal{Y}) = \frac{\ell(\mathcal{Y}|\mathcal{M})\, p_{\rm pr}(\mathcal{M})}{\int \ell(\mathcal{Y}|\mathcal{M})\, p_{\rm pr}(\mathcal{M})\, {\rm d}\mathcal{M}}. \tag{3}$$

Here $\ell(\mathcal{Y}|\mathcal{M})$, usually termed the *likelihood*, is the conditional PDF of the measurements $\mathcal{Y}$ for a given choice $\mathcal{M}$ of the dynamical model. In practice, (3) can be applied iteratively using a sequence of data blocks $\mathcal{Y}, \mathcal{Y}'$, etc. The posterior computed from block $\mathcal{Y}$ serves as the prior for the next block $\mathcal{Y}'$, etc. For a sufficiently large number of observations, $p_{\rm post}(\mathcal{M}|\mathcal{Y}, \mathcal{Y}', \ldots)$ is sharply peaked at a certain most probable model $\mathcal{M} = \mathcal{M}^*$.

The main efforts in the research on stochastic nonlinear dynamical inference are focused on constructing the likelihood function that compensates noise induced errors and on introducing efficient optimization algorithms (cf.[5,7,9,11]).

No closed form expression for the likelihood function that provides optimal compensation of the noise-induced errors was introduced so far for continuous systems. The *ad hoc* likelihood function[5] and their generalization to the conditional PDF for stochastic trajectories in maps[7,9] do not compensate the error in continuous systems, since they are missing the main compensating term (see below). The problem of noise-induced errors in inference of continuous systems was considered in[11] and a general approach to constructing corresponding likelihood was outlined. However, the closed form expression for the likelihood that takes into account the leading compensating term was not found and instead an *ad hoc* expression for the likelihood function was used.

A common draw back of earlier research is the use of extensive numerical optimization. This problem will become increasingly important when complex systems with the large number (hundred or more) of model coefficients are investigated.

In the present paper we introduce a closed form of the likelihood function for continuous systems that provides optimal compensation for the noise-induced errors. We also suggest parametrization of the unknown vector field that reduces the problem of nonlinear dynamical inference to essentially linear one for a broad class of nonlinear dynamical systems. This allows us to write an efficient algorithm of Bayesian inference of nonlinear noise-driven dynamical models that avoids extensive numerical optimization and guarantees optimum compensation of noise-induced errors.

In what follows in this section we describe the likelihood function, the parametrization, and the corresponding algorithm.

2.1. The likelihood function

It was pointed out in[11] the probability density functional for the nonlinear dynamical stochastic systems in general is not known. Instead one can use the probability density functional for random trajectories in such systems. We note that the path-integral approach has also proved to be useful in nonlinear filtration of random signals (see e.g.[19])in the situations where standard spectral-correlation methods fail.

Therefore we write the expression for the likelihood in the form of a path integral over the random trajectories of the system:

$$\ell(\mathcal{Y}|\mathcal{M}) = \int_{\mathbf{x}(t_i)}^{\mathbf{x}(t_f)} p_o(\mathcal{Y}|\mathcal{X})\, \mathcal{F}_{\mathcal{M}}[\mathbf{x}(t)]\, \mathcal{D}\mathbf{x}(t), \tag{4}$$

which relates the dynamical variables $\mathbf{x}(t)$ of the system (1) to the observations $\mathbf{y}(t)$. Here we choose $t_i \ll t_0 < t_K \backslash t_f$ so that ℓ does not depend on the particular initial and final states $\mathbf{x}(t_i)$, $\mathbf{x}(t_f)$. The form of the probability functional $\mathcal{F}_{\mathcal{M}}$ over the system trajectory $\mathbf{x}(t)$ is determined by the properties of the dynamical noise $\xi(t)$.[20,21]

In the following we are focusing on the case of additive and stationary Gaussian white noise, as indicated in (1), (2). We consider a uniform sampling scheme $t_k = t_0 + hk$, $h \equiv (t_K - t_0)/K$ and assume that for each trajectory component $x_n(t)$ the measurement error ϵ is negligible compared with the fluctuations induced by the dynamical noise; that is, $\epsilon^2 \ll h(\hat{\mathbf{D}}^2)_{nn}$. Consequently, we use $p_o(\mathcal{Y}|\mathcal{X}) \to \prod_{k=0}^{K} \delta[\mathbf{y}_k - \mathbf{x}(t_k)]$ in (4). Using results from[20] for $\mathcal{F}_{\mathcal{M}}[\mathbf{x}(t)]$, the logarithm of the likelihood (4) takes the following form for sufficiently large K (small time step h):

$$-\frac{2}{K}\log \ell(\mathcal{Y}|\mathcal{M}) = \ln\det \hat{\mathbf{D}} + \frac{h}{K}\sum_{k=0}^{K-1}\Big[\, \mathrm{tr}\,\hat{\mathbf{\Phi}}(\mathbf{y}_k;\mathbf{c}) + (\dot{\mathbf{y}}_k - \mathbf{f}(\mathbf{y}_k;\mathbf{c}))^T\, \hat{\mathbf{D}}^{-1}\, \dot{\mathbf{y}}_k - \mathbf{f}(\mathbf{y}_k;\mathbf{c}))\Big] + N\ln(2\pi h), \tag{5}$$

here we introduce the "velocity" $\dot{\mathbf{y}}_k$ and matrix $\hat{\mathbf{\Phi}}(\mathbf{x})$

$$\dot{\mathbf{y}}_k \equiv h^{-1}(\mathbf{y}_{k+1} - \mathbf{y}_k), \quad (\hat{\mathbf{\Phi}}(\mathbf{x};\mathbf{c}))_{n\,n'} \equiv \partial f_n(\mathbf{x};\mathbf{c})/\partial x_{n'}.$$

It is the term $\mathrm{tr}\,\hat{\mathbf{\Phi}}(\mathbf{y}_k;\mathbf{c})$ that guarantees optimal compensation of the noise-induced errors in our technique and that distinguish our likelihood function from those introduced in earlier work. Formally this term appears in path integral as a Jacobian of transformation[22-24] from noise variables to dynamical ones. We emphasize, however, that this term is not a correction, but a leading term in inference as will be shown in the following sections.

Note, that the optimization of the log-likelihood function (5) is in general essentially nonlinear problem that requires extensive numerical optimization. Below we introduce parametrization that allows to avoid this problem for a broad class of nonlinear dynamical models. In particular, a vast majority of the model examples considered in the earlier work on the nonlinear dynamical inference can be solved using this technique. Moreover, a large number of important practical applications can be treated using the same approach.

2.2. Parametrization of the unknown vector field

We parameterize this system in the following way. The nonlinear vector field $\mathbf{f}(\mathbf{x})$ is written in the form

$$\mathbf{f}(\mathbf{x}) = \hat{\mathbf{U}}(\mathbf{x})\,\mathbf{c} \equiv \mathbf{f}(\mathbf{x};\mathbf{c}), \tag{6}$$

where $\hat{\mathbf{U}}(\mathbf{x})$ is an $N \times M$ matrix of suitably chosen basis functions $\{U_{nm}(\mathbf{x});\ n = 1:N,\ m = 1:M\}$, and $\mathbf{c}$ is an M-dimensional coefficient vector.

The choice of the base functions is not restricted to polynomials, $\phi_b(\mathbf{x})$ can be any suitable function. In general if we use B different base functions $\phi_b(\mathbf{x})$ to model the system (1) the matrix $\hat{\mathbf{U}}$ will have the following block structure

$$\hat{\mathbf{U}} = \left[\left[\begin{array}{cccc} \phi_1 & 0 & \dots & 0 \\ 0 & \phi_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi_1 \end{array}\right] \cdots \left[\begin{array}{cccc} \phi_2 & 0 & \dots & 0 \\ 0 & \phi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi_2 \end{array}\right] \cdots \left[\begin{array}{cccc} \phi_B & 0 & \dots & 0 \\ 0 & \phi_B & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi_B \end{array}\right]\right], \tag{7}$$

where we have B diagonal blocks of size $N \times N$ and $M = B \cdot N$.

An important feature of (6) for our subsequent development is that, while possibly highly nonlinear in $\mathbf{x}$, $\mathbf{f}(\mathbf{x};\mathbf{c})$ is strictly linear in $\mathbf{c}$.

Eqs. (5) and (6) are two main ingredients that allow to solve problem of nonlinear stochastic dynamical inference analytically as shown in the following section.

2.3. The algorithm

The vector elements $\{c_m\}$ and the matrix elements $\{D_{nn'}\}$ together constitute a set $\mathcal{M} = \{\mathbf{c}, \hat{\mathbf{D}}\}$ of unknown parameters to be inferred from the measurements $\mathcal{Y}$.

With the use of (6), substitution of the prior $p_{\rm pr}(\mathcal{M})$ and the likelihood $\ell(\mathcal{Y}|\mathcal{M})$ into (3) yields the posterior $p_{\rm post}(\mathcal{M}|\mathcal{Y}) = \text{const} \times \exp[-S(\mathcal{M}|\mathcal{Y})]$, where

$$S(\mathcal{M}|\mathcal{Y}) \equiv S_{\rm y}(\mathbf{c},\hat{\mathbf{D}}) = \frac{1}{2}\rho_{\rm y}(\hat{\mathbf{D}}) - \mathbf{c}^T\mathbf{w}_{\rm y}(\hat{\mathbf{D}}) + \frac{1}{2}\mathbf{c}^T\hat{\Xi}_{\rm y}(\hat{\mathbf{D}})\mathbf{c}. \tag{8}$$

Here, use was made of the definitions

$$\rho_{\rm y}(\hat{\mathbf{D}}) = h\sum_{k=0}^{K-1} \dot{\mathbf{y}}_k^T\,\hat{\mathbf{D}}^{-1}\,\dot{\mathbf{y}}_k + K\,\ln(\det\hat{\mathbf{D}}), \tag{9}$$

$$\mathbf{w}_{\rm y}(\hat{\mathbf{D}}) = \hat{\mathbf{\Sigma}}_{\rm pr}^{-1}\,\mathbf{c}_{\rm pr} + h\sum_{k=0}^{K-1}\left[\hat{\mathbf{U}}_k^T\,\hat{\mathbf{D}}^{-1}\,\dot{\mathbf{y}}_k - \frac{\mathbf{v}(\mathbf{y}_k)}{2}\right], \tag{10}$$

$$\hat{\Xi}_{\rm y}(\hat{\mathbf{D}}) = \hat{\mathbf{\Sigma}}_{\rm pr}^{-1} + h\sum_{k=0}^{K-1}\hat{\mathbf{U}}_k^T\,\hat{\mathbf{D}}^{-1}\,\hat{\mathbf{U}}_k, \tag{11}$$

where $\hat{\mathbf{U}}_k \equiv \hat{\mathbf{U}}(\mathbf{y}_k)$ and the components of vector $\mathbf{v}(\mathbf{x})$ are:

$$\mathrm{v}_m(\mathbf{x}) = \sum_{n=1}^{N}\frac{\partial U_{n\,m}(\mathbf{x})}{\partial x_n}, \quad m = 1:M. \tag{12}$$

The mean values of $\mathbf{c}$ and $\hat{\mathbf{D}}$ in the posterior distribution give the best estimates for the model parameters for a given block of data $\mathcal{Y}$ of length K and provide a global minimum to $S_{\rm y}(\mathbf{c},\hat{\mathbf{D}})$. We handle this optimization problem in the following way. Assume for the moment that $\mathbf{c}$ is known in (8). Then the posterior distribution

over $\hat{\mathbf{D}}$ has a mean $\hat{\mathbf{D}}'_{\rm post} = \hat{\mathbf{\Theta}}_{\rm y}(\mathbf{c})$ that provides a minimum to $S_{\rm y}(\mathbf{c}, \hat{\mathbf{D}})$ with respect to $\hat{\mathbf{D}} = \hat{\mathbf{D}}^T$. Its matrix elements are

$$\hat{\mathbf{\Theta}}_{\rm y}^{nn'}(\mathbf{c}) \equiv \frac{1}{K} \sum_{k=0}^{K-1} \left[\dot{\mathbf{y}}_k - \hat{\mathbf{U}}(\mathbf{y}_k)\,\mathbf{c}\right]_n \left[\dot{\mathbf{y}}_k - \hat{\mathbf{U}}(\mathbf{y}_k)\,\mathbf{c}\right]^T_{n'}. \tag{13}$$

Alternatively, assume next that $\hat{\mathbf{D}}$ is known, and note from (8) that in this case the posterior distribution over $\mathbf{c}$ is Gaussian. Its covariance is given by $\hat{\mathbf{\Xi}}_{\rm y}(\hat{\mathbf{D}})$ and the mean $\mathbf{c}'_{\rm post}$ minimizes $S_{\rm y}(\mathbf{c}, \hat{\mathbf{D}})$ with respect to $\mathbf{c}$

$$\mathbf{c}'_{\rm post} = \hat{\mathbf{\Xi}}_{\rm y}^{-1}(\hat{\mathbf{D}})\mathbf{w}_{\rm y}(\hat{\mathbf{D}}). \tag{14}$$

We repeat this two-step optimization procedure iteratively, starting from some prior values $\mathbf{c}_{\rm pr}$ and $\hat{\mathbf{\Sigma}}_{\rm pr}$.

It can be seen that the second term in the sum on the *rhs* of eq. (10) originating from $\mathrm{tr}\hat{\mathbf{\Phi}}(\mathbf{y}_k)$ does not vanish at the dynamical system attractors (1), unlike the term (5) $h\sum_{k=0}^{K-1} \hat{\mathbf{U}}_k^T\, \hat{\mathbf{D}}^{-1}\, \dot{\mathbf{y}}_k$ corresponding to the generalized least square optimization.[25] Therefore both types of terms are required to optimally balance the effect of noise effect in $\{\mathbf{y}_k\}$ (8) and provide the robust convergence. In the following section we analyze relative importance of both terms for the convergence of our algorithm.

3. NUMERICAL EXAMPLES

We verified the convergence and robustness of the algorithm on a broad range of dynamical systems. In this paper we will be specifically focused on the applications to the inference of coupled nonlinear oscillators.

3.1. Five coupled oscillators

Consider system of five locally and globally coupled van der Pol oscillators

$$\begin{aligned}
&\dot{x}_k = y_k, \\
&\dot{y}_k = \varepsilon_k(1 - x_k^2)y_k - \omega_k x_k + \sum_{\substack{j=1 \\ j\neq k}}^{5} \eta_{kj}x_j + \gamma_{k,k-1}x_{k-1}(t) + \gamma_{k,k+1}x_{k+1}(t) + \sum_{j=1}^{5} \sigma_{kj}\xi_j,
\end{aligned} \tag{15}$$

We assume for simplicity that there is no observational noise and that the observed signal is $\mathbf{y} = (y_1, y_2, y_3, y_4, y_5)$. We note that for the model of coupled oscillators (15) parameters of the equations $\dot{x}_k = y_k$ are known and do not have to be inferred. An example of a trajectory of (15) is shown in the figure 1(a) in projection on (x_1, x_2, x_3) subspace of the configuration space of this system. We chose the following base functions

$$\begin{aligned}
&\phi(1) = x_1; \quad \phi(2) = y_1; \quad \phi(3) = x_2; \quad \phi(4) = y_2; \quad \phi(5) = x_3; \quad \phi(6) = y_3; \\
&\phi(7) = x_4; \quad \phi(8) = y_4; \quad \phi(9) = x_5; \quad \phi(10) = y_5; \quad \phi(11) = x_1x_2; \quad \phi(12) = x_2x_3; \\
&\phi(13) = x_3x_4; \quad \phi(14) = x_4x_5; \quad \phi(15) = x_5x_1; \quad \phi(16) = x_1^2y_1; \quad \phi(17) = x_2^2y_2; \\
&\phi(18) = x_3^2y_3; \quad \phi(19) = x_4^2y_4; \quad \phi(20) = x_5^2y_5.
\end{aligned}$$

Together with the elements of the diffusion matrix we have to infer 115 model coefficients. Example of the convergence of the coefficients of the 5^{th} oscillator to their correct values is shown in the Fig. 1(b). Results of the corresponding convergence for the 4^{th} oscillator are summarized in the Table 1. It can be seen from the Table that accuracy of estimation of the model parameters is better then 1%. Of a special interest for us is the compensation of the noise-induced errors. In the figure Fig. 2 we compare results of inference of one of the coefficients of the system (1) ε_1 for two different diffusion matrices D and $2D$ where matrix D was chosen at random

$$\hat{\mathbf{D}} = \begin{bmatrix}
0.0621 & 1.9171 & 0.4307 & 0.0356 & 0.3113 \\
0.5773 & 1.3597 & 0.3648 & 1.7559 & 0.3259 \\
1.9421 & 0.1099 & 0.1535 & 0.7051 & 0.6268 \\
1.9010 & 1.1997 & 0.0148 & 1.4443 & 0.0588 \\
0.4561 & 0.7863 & 1.5776 & 1.9369 & 0.7153
\end{bmatrix}. \tag{16}$$

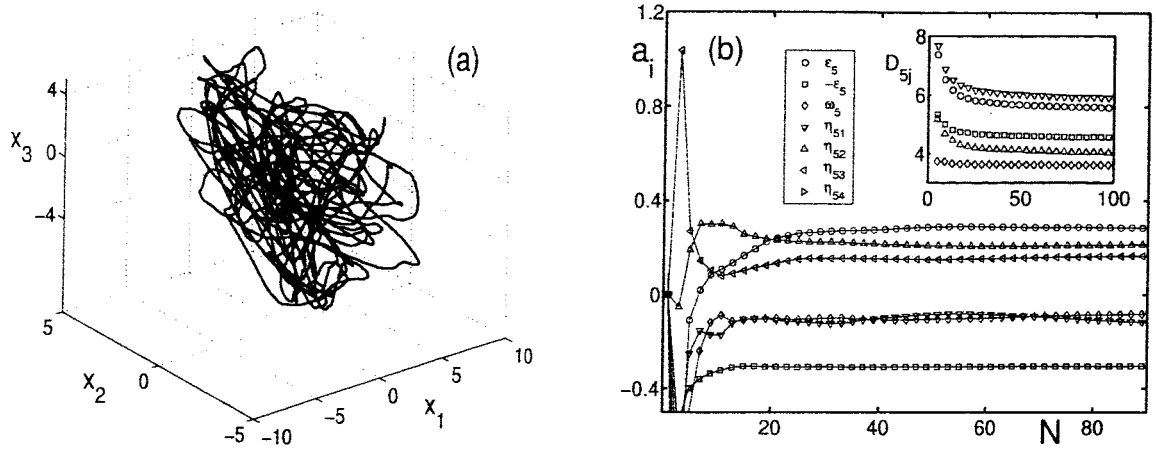

Figure 1. (a) A projection of a trajectory of system (15) on (x_1, x_2, x_3) subspace of its configuration space. (b) Convergence of the coefficients of the 5^{th} oscillator to the true values as a function of a number of blocks of data. We have 100 blocks of data with 800 points in each block and the sampling time $h = 0.02$. $a_1 = \epsilon_1$, $a_2 = -\epsilon_1$, $a_3 = -\omega_1$, $a_4 = \eta_{12}$, $a_5 = \eta_{13}$, $a_6 = \eta_{14}$, $a_7 = \eta_{15}$, $a_8 = \gamma_{15}$. The convergence of the five components of the diffusion matrix is shown in the insert.

Table 1. Convergence of the coefficients of the 4^{th} oscillator of system (15). We have used 200 blocks of data with 5000 points in each block. True values are shown in the first row, inferred values are shown in the second row, and corresponding standard deviations are shown in the last row. The accuracy of inference is within 1%.

coefficients	ε_4	ω_4	η_{41}	η_{42}	η_{43}	η_{44}	γ_{43}	γ_{45}	D_{41}	D_{42}	D_{43}
true value	0.2	-0.06	-0.075	0.24	-0.23	-0.2	0.064	0.095	0.575	1.032	0.833
inferred value	0.199	-0.062	-0.071	0.244	-0.228	-0.197	0.065	0.096	0.576	1.032	0.883
std	0.009	0.006	0.006	0.007	0.005	0.007	0.002	0.002	0.001	0.002	0.001

It can be shown that that without compensation term the estimator (14), (14) is reduced to the generalized least square (GLS) estimator. The Fig. 2 shows that the GLS estimator systematically overestimates the value of ε_1 and the larger is noise intensity the larger is the systematic error of the overestimation (see curves 1' and 2' for D and $2D$ correspondingly). By adding the term $\mathrm{tr}\,\hat{\mathbf{\Phi}}(\mathbf{y}_k;\mathbf{c})$ we obtain optimal compensation of the noise-induced errors as shown by the curves 1 and 2 obtained for the same noise intensities. To see the effect of the compensation analytically it is instructive to rewrite the sum in the eq. (10) in the integral form

$$\mathbf{w}_{\mathbf{y}}(\hat{\mathbf{D}}) = \hat{\mathbf{\Sigma}}_{\mathrm{pr}}^{-1}\,\mathbf{c}_{\mathrm{pr}} + \int_{x(t_0)}^{x(T)} \hat{\mathbf{U}}(\mathbf{y}(t))^T\,\hat{\mathbf{D}}^{-1}\,d\mathbf{y} - \frac{1}{2}\int_{t_0}^{T} \mathbf{v}(\mathbf{y}_k)dt, \tag{17}$$

It can be seen from eq. (17) that for the attractor localized in the phase space the first integral is finite, since initial and final points of integration belong to the attractor. The second integral is growing when the total time of inference is growing.

In particular, for a point attractor this integral is identically zero and the whole inference is due to the compensating term $\frac{1}{2}\int_{t_0}^{T}\mathbf{v}(\mathbf{y}_k)dt$. This result is intuitively clear, since for the point attractor in the absence of noise the system will stay forever in the same point and no inference can be done. It is only noise that forces the system to move about in the phase space and makes it possible to perform inference.

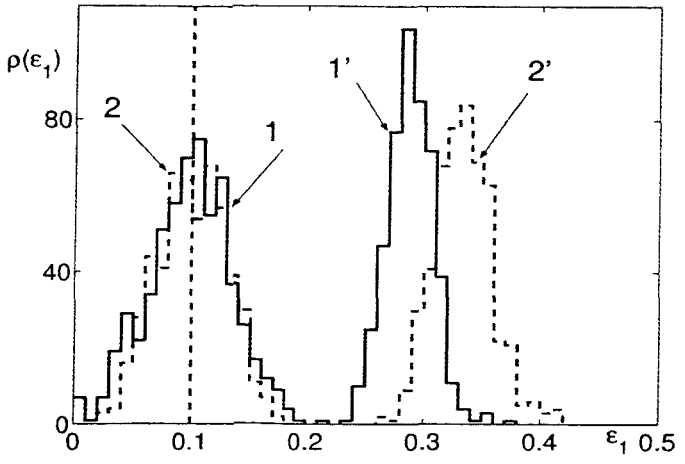

Figure 2. Results of inference of the ε_1 that were performed according to eqs. (10) – (14) (curves 1 and 2) are compared with the results of inference without compensating term tr $\hat{\Phi}(\mathbf{y}_k;\mathbf{c})$ (curves 1′ and 2′) for different noise matrices d: $d = D$ for (1) and (1'); $d = 2 * D$ for (2) and (2') where D is given in eq. (16).

4. INFERENCE OF TWO COUPLED OSCILLATORS FROM UNIVARIATE TIME-SERIES DATA

As we have mentioned in the introduction in many real experimental situations the model is not usually known exactly from "first principles" and in addition, the experimental data can be extremely skewed, whereby important "hidden" features of a model such as coupling coefficients can be very difficult to extract due to the intricate interplay between noise and nonlinearity.

A specific example of such experimental situation is inference of the strength, directionality and a degree of randomness of the cardiorespiratory interaction from the blood pressure signal. Such inference can provide valuable diagnostic information about the responses of the autonomous nervous system.[17,26] However, it is inherently difficult to dissociate a specific response from the rest of the cardiovascular interactions and the mechanical properties of the cardiovascular system in the intact organism.[27] Therefore a number of numerical techniques were introduced to address this problem using e.g. linear approximations,[28] or semi-quantitative estimations of either the strength of some of the nonlinear terms[29] or the directionality of coupling.[30,31] But the problem remains wide open because of the complexity and nonlinearity of the cardiovascular interactions.

It is important to notice that simultaneous measurements of the cardiovascular signals is performed in different parts of the system (see e.g.[15]). As a consequence the nonlinear characteristics of the oscillations are substantially modified in each signal and inference of nonlinear coupling parameters has to be performed preferably using univariate data e.g. blood pressure or blood flow signal only. The necessity to use univariate data in general poses serious limitations on the techniques of reconstruction and the problem become essentially nontrivial even in quasi-linear noise-free limit.[32–34]

In this section we investigate the possibility of extending our technique of reconstruction of coupled nonlinear stochastic oscillators to encompass the case of inference from the univariate time-series data in the context of physiological research. We note that this is a particular example of inference of hidden dynamical variables, which will be addressed elsewhere. An example of the actual signal of the central venous blood pressure (record 24 of the MGH/MF Waveform Database available at www.physionet.org). The main features of the blood signal data is the presence of the two oscillatory components at frequencies approximately $f_r = 0.2$ Hz and $f_c = 1.7$ Hz corresponding to the respiratory and cardiac oscillations. It can also be clearly seen from the spectra that the nonlinear terms including terms of nonlinear interaction (and cardiorespiratory interaction in particular) are very strong in this sample. We note that the relative intensity and position of the cardiac and respiratory components vary strongly from sample to sample with average frequency of the respiration being around 0.3 Hz and of the heart beat being around 1.1 Hz. To infer coupling parameters from the univariate blood pressure

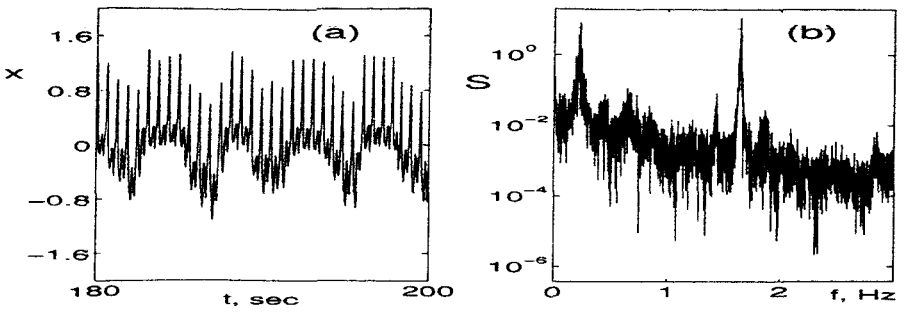

Figure 3. Example of the blood pressure signal (a) and of its spectrum (b) taken from the record 24 of the MGH/MF Waveform Database available at www.physionet.org.

signal an important simplifying assumption can be used. Namely it is assumed that the blood pressure signal can be represented as the sum of the oscillatory components with the main contributions coming from the oscillations of the respiration and heart.[15] Accordingly we chose our surrogate data as a sum of coordinates of two coupled van der Pol oscillators $s(t) = x_1(t) + x_2(t)$. It can be seen that the spectrum of $s(t)$ (Fig. 4 (c)) reproduces mentioned above main features of the real blood pressure signal.

$$\dot{x}_1 = y_1, \qquad \dot{y}_1 = \epsilon_1(1 - x_1^2)y_1 - \omega_1^2 x_1 + \alpha_1 x_2 + \sum_{i,j=1}^{2} \beta_{1,ij} x_i x_j + \sum_{\substack{i,j=1 \\ j \neq i}}^{2} \gamma_{1,ij} x_i y_j + \sum_{j=1}^{2} \sigma_{1j} \xi_j, \tag{18}$$

$$\dot{x}_2 = y_2, \qquad \dot{y}_2 = \epsilon_2(1 - x_2^2)y_2 - \omega_2^2 x_2 + \alpha_2 x_2 + \sum_{i,j=1}^{2} \beta_{2,ij} x_i x_j + \sum_{\substack{i,j=1 \\ j \neq i}}^{2} \gamma_{2,ij} x_i y_j + \sum_{j=1}^{2} \sigma_{2j} \xi_j, \tag{19}$$

$$\langle \xi_i(t) \rangle = 0, \qquad \langle \xi_i(t) \xi_j(t') \rangle = \delta_{ij} \delta(t - t').$$

Here noise matrix σ mixes zero-mean white Gaussian noises $\xi_j(t)$ and is related to the diffusion matrix $D = \sigma \cdot \sigma^T$.

To infer parameters of nonlinear coupling between cardiac and respiratory oscillations we decompose "measured" signal $s(t)$ on two oscillatory components using a combination of low- and high-pass Butterworth filters representing observations of mechanical cardiac and respiratory degrees of freedom on a discrete time grid with step $h = 0.02$ sec. Obtained in this way signals $s_0(t)$ and $s_1(t)$ are shown in the Fig. 4 (a) and (b) respectively.* To make this numerical experiment even more realistic the input signal $s(t)$ was filtered before decomposition (using high-pass Butterworth filter of the 2^{nd} order with cut-off frequency 0.0025 Hz) to model standard procedure of de-trending, which is used in time-series analysis of the cardiovascular signals to remove low frequency non-stationary trends. We now use standard embedding procedure to introduce an auxiliary two-dimensional dynamical system whose trajectory $\mathbf{z}(t) = (z_0(t), z_1(t))$ is related to the observations $\{\mathrm{s}(t_k)\}$ as follows

$$z_n(t_k) = \frac{s_n(t_k + h) - s_n(t_k)}{h},$$

where $n = 1, 2$. The corresponding simplified model of the nonlinear interaction between the cardiac and respiratory limit cycles has the form (cf. with [16])

$$\begin{aligned} \dot{z}_n = {} & b_{1,n} s_n + b_{2,n} s_{n-1} + b_{3,n} z_n + b_{4,n} z_{n-1} + b_{5,n} s_n^2 + b_{6,n} s_{n-1}^2 + b_{7,n} z_n^2 + b_{8,n} z_{n-1}^2 \\ & + b_{9,n} s_n s_{n-1} + b_{10,n} s_n z_n + b_{11,n} s_n z_{n-1} + b_{12,n} s_{n-1} z_n + b_{13,n} s_{n-1} z_{n-1} + b_{14,n} z_n z_{n-1} \\ & + b_{15,n} s_n^2 z_n^2 + b_{16,n} s_{n-1} z_n^2 + \xi_n(t), \quad n = 0, 1. \end{aligned} \tag{20}$$

where $\xi_n(t)$ is a Gaussian white noise with correlation matrix $Q_{n\,n'}$. We emphasize that a number of important parameters of the decomposition of the original signal (including the bandwidth, the order of the filters) have to

*Note the difference between actual oscillatory components of the original signal $s(t)$: $x_1(t)$ and $x_2(t)$ and the components obtained using spectral decomposition with Butterworth filters: $s_0(t)$ and $s_1(t)$.

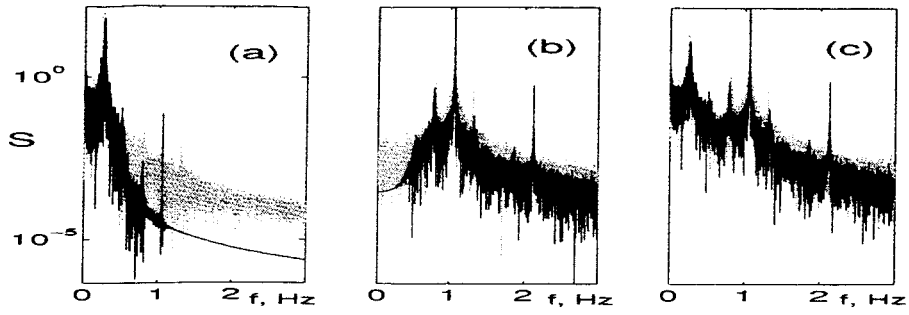

Figure 4. Comparison of the power spectra of the inferred $z(t)$ components of the signal (gray lines) with the original signal $s(t)$ (black lines): (a) a low-frequency component of the signal $s_0(t)$ obtained using low-pass Butterworth filter of the 5^{th} order with cut-off frequency 0.55 Hz (black line) is compared with the inferred signal $z_0(t)(grayline)$; (b) a high-frequency component of the signal $s_1(t)$ obtained high-pass Butterworth filter of the 4^{th} order with cut-off frequency 0.55 Hz (black line) is compared with the inferred signal $z_1(t)$ (gray line); (c) spectrum of the original signal $s(t) = x_1(t)+x_2(t)$ (black line) is compared with the spectrum of the inferred signal $z(t) = z_0(t) + z_1(t)$.

be selected to minimize the cost (8) and provide the best fit to the measured time series $\{s(t_k)\}$. The parameters of the model (20) can now be inferred directly from the univariate "measured" time series data $s(t)$. The comparison between the time series of the inferred and actual cardiac oscillations is shown in Fig. 4 and Fig. 5. The comparison of the inferred parameters with their actual values is summarized in the Tables 2. It can

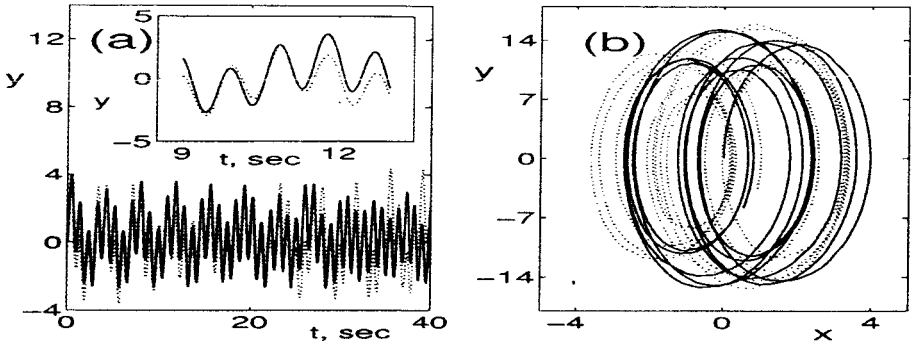

Figure 5. (a) Comparison of the inferred signal $x(t) = x_1(t)+x(2)$ (black solid line) with the original signal (black dotted line). In the insert fragments of both signals are compared with better resolution in time. (b) Comparison of the inferred phase space trajectory $(x(t), y(t))$ (black solid line) with the original one (black dotted line). To facilitate the comparison we have used the same initial conditions to generate phase space trajectory with exact parameters of the system and with inferred parameters of the system.

Table 2. Comparison of the non-zero parameters of the second equation of the model (20) inferred from the univariate time series data $s(t)$ with their actual values. We have used 1 block of data with 120000 points. True values are shown in the first row, inferred values are shown in the second row.

coefficients	$b_{2,1}$	$b_{2,2}$	$b_{2,3}$	$b_{2,4}$	$b_{2,5}$	$b_{2,6}$	$b_{2,9}$	$b_{2,14}$	$b_{2,16}$	D_1
true value	0.05	-45.0	-0.19	0.25	0	2.55	0.2	0.11	-0.25	0.2
inferred value	0.014	-44.73	-0.071	0.17	-0.081	1.25	0.415	0.14	-0.251	0.17

be seen from the Table that the inferred parameters give correct order of the magnitude for the actual values.

The inferred values can be further corrected taken into account attenuation of the filters at different frequencies. We emphasize, however, that the technique of spectral decomposition of the "measured" signal is in principal non-unique. Moreover, in the actual experimental situation the dynamics of the physiological oscillations is unknown and can be only very approximately modelled by the system of coupled oscillators. Furthermore, the only criterion for the goodness of the spectral decomposition is the coincidence of the original and inferred signal and spectrum. For these reasons the estimation of the model parameters with the accuracy better then the order of magnitude does not improve the quality of the inferred information as will be discussed in more details elsewhere.

In conclusion, we suggested new technique of inference of parameters nonlinear stochastic dynamical system. The technique does not require extensive global optimization, provides optimal compensation for noise-induced errors and is robust in a broad range of dynamical models. We illustrate the main ideas of the technique by inferencing 115 model coefficients of five globally and locally coupled noisy oscillators within accuracy 1%. It is demonstrated that our technique can be readily extended to solve selected problems of nonlinear stochastic inference of hidden dynamical variables in the context of the physiological modelling. We show in particular that the method allows one to estimate correct order of the magnitude of nonlinear coupling of two stochastic oscillators from univariate time series data. The framework of nonlinear Bayesian inference outlined in this paper can be further generalized to include errors of measurements and to solve problem of global inference hidden dynamical variables.

4.1. Acknowledgments

The work was supported by the NASA CICT/IS IDU Program (USA) Engineering and Physical Sciences Research Council (UK), the Russian Foundation for Fundamental Science, and INTAS.

REFERENCES

1. H. Haken, *Synergetics, An Introduction*, Springer, Berlin, 1983.
2. S. H. Strogatz, *Nonlinear Dynamics and Chaos*, Addison-Wesley, Reading, 1994.
3. Y. Kuramoto, *Chemical Oscillations, Waves, and Turbolence*, Springer-Verlag, Berlin, 1984.
4. A. T. Winfree, "The geometry of biological time," in *Springer-Verlag*, 1980, New York.
5. P. E. McSharry and L. A. Smith, "Better nonlinear models from noisy data: attractors with maximum likelihood," *Physical Review Letters* **83**, pp. 4285–4288, 1999.
6. H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, 1997.
7. R. Meyer and N. Christensen, "Fast bayesian reconstruction of chaotic dynamical systems via extended kalman filtering," *Phys. Rev. E* **65**, p. 016206, 2001.
8. J. P. M. Heald and J. Stark, "Estimation of noise levels for models of chaotic dynamical systems," *Phys. Rev. Lett.* **84**(11), pp. 2366–2369, 2000.
9. R. Meyer and N. Christensen, "Bayesian reconstruction of chaotic dynamical systems," *Phys. Rev. E* **62**, pp. 3535–3542, 2000.
10. J. Gradisek, S. Siegert, R. Friedrich, and I. Grabec, "Analysis of time series from stochastic processes," *Phys. Rev. E* **62**(3), pp. 3146–3155, 2000.
11. J.-M. Fullana and M. Rossi, "Identification methods for nonlinear stochastic systems," *Phys. Rev. E* **65**, p. 031107, 2002.
12. M. Siefert, A. Kittel, R. Friedrich, and J. Peinke, "On a quantitative method to analyze dynamical and measurement noise," *Europhys. Lett.* **61**(4), pp. 466–472, 2003.
13. P. J. Saul, D. T. Kaplan, and R. I. Kitney, "Nonlinear interactions between respiration and heart rate: Classical physiology or entrained nonlinear oscillators," in *Computers in Cardiology*, pp. 299–302, 1988 IEEE Comput. Soc. Press, (Washington), 1989.
14. I. Javorka, M. ans Zila, K. Javorka, and A. Calkovska, "Do the oscillations of cardiovascular parameters persist during voluntary apnea in humans?," *Physol. Res.* **51**, pp. 227–238, 2002.
15. A. Stefanovska and M. Bračič, "Physics of the human cardiovascular system," *Contemporary Physics* **40**, pp. 31–55, 1999.

16. A. Stefanovska, M. Bračič, S. Strle, and H. Haken, "The cardiovascular system as coupled oscillators?," *Physiol. Meas.* **22**, pp. 535–550, 2001.
17. S. C. Malpas, "Neural influences on cardiovascular variability: possibilities and pitfalls," *Am. J. Physiol.: Heart. Circ. Physiol.* **282**, pp. H6–H20, 2002.
18. P. van Leeuwen and H. Bettermann, "The status of nonlinear dynamics," *Herzschr Elektrophys* **11**, pp. 127–130, 2000.
19. A. K. Rosov, *Nonlinear Filtration of Signals*, Politechnika, St. Petersburg, 2002.
20. R. Graham, "Path integral formulation of general diffusion processes," *Z. Phys. B* **26**, pp. 281–290, 1977.
21. M. I. Dykman, "Large fluctuations and fluctuational transitions in systems driven by colored gaussian noise–a high frequency noise," *Phys. Rev. A* **42**, pp. 2020–2029, 1990.
22. R. Graham, *Tracts in Modern Physics*, vol. 66, ch. Quantum Statistics in Optics and Solid-State Physics. Springer-Verlag, New York, 1973.
23. G. E., "Functional-integral approach to parisi-wu stochastic quantization: Scalar theory," *Physical Review D* **28**, pp. 1922–1930, 1983.
24. A. J. McKane, "Noise-induced escape rate over a potential barrier: Results for a general noise," *Phys. Rev. A* **40**(7), pp. 4050–4053, 1989.
25. H. Theil, *Linear algebra and matrix methods in econometrix*, vol. I of *Hundhook of Econometrics*, ch. 1, pp. 5–65. North- Holland Publishing Company, 1983.
26. J. Hayano and F. Yasuma, "Hypothesis: respiratory sinus arrhythmia is an intrinsic resting function of cardiopulmonary system," *Cardiovascular Research* **58**, pp. 1–9, 2003.
27. D. Jordan, "Central nervous integration of cardiovascular regulation," in *Cardiovascular regulation*, D. Jordan and J. Marshall, eds., Portland Press, Cambridge, 1995.
28. J. A. Taylor, C. W. Myers, J. R. Halliwill, H. Seidel, and D. L. Eckberg, "Sympathetic restraint of respiratory sinus arrhythmia: implications for vagal-cardiac tone assessment in humans," *Am J Physiol Heart Circ Physiol* **280**(6), pp. H2804–2814, 2001.
29. J. Jamšek, I. A. Khovanov, D. G. Luchinsky, P. V. E. McClintock, and A. Stefanovska *Phys. Rev. E*, submitted , 2003.
30. M. G. Rosenblum, L. Cimponeriu, A. Bezerianos, A. Patzak, and R. Mrowka, "Identification of coupling direction: Application to cardiorespiratory interaction," *Phys. Rev. E.* **65**, p. 041909, 2002.
31. M. Paluš, V. Komárek, Z. Hrnčíř, and K. Štěbrová, "Synchronization as adjustment of information rates: Detection from bivariate time series," *Phys. Rev. E* **63**, p. 046211, 2001.
32. N. B. Janson, A. G. Balanov, V. S. Anishchenko, and P. V. E. McClintock, "Phase synchronization between several interacting processes from univariate data," *Phys. Rev. Lett.* **86**, pp. 1749–1752, 2001.
33. N. B. Janson, A. G. Balanov, V. S. Anishchenko, and P. V. E. McClintock, "Phase relationships between two or more interacting processes from one-dimensional time series: I. Basic theory," *Phys. Rev. E* **65**, pp. 036211/1–12, 2002.
34. N. B. Janson, A. G. Balanov, V. S. Anishchenko, and P. V. E. McClintock, "Phase relationships between two or more interacting processes from one-dimensional time series: II. Application to heart-rate-variability data," *Phys. Rev. E* **65**, pp. 036212/1–7, 2002.