P H Y S I C A L R E V I E W L E T T E R S

Inference of a Nonlinear Stochastic Model of the Cardiorespiratory Interaction

V. N. Smelyanskiy,[1] D. G. Luchinsky,[2,3] A. Stefanovska,[2,4] and P. V. E. McClintock[2]

[1]*NASA Ames Research Center, MS 269-2, Moffett Field, CA 94035, USA*
[2]*Department of Physics, Lancaster University, Lancaster LA1 4YB, United Kingdom*
[3]*Mission Critical Technologies, Inc., NASA Ames Research Center, MS 269-3, Moffett Field, CA 94035, USA*
[4]*Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, 1000 Ljubljana, Slovenia*

We reconstruct a nonlinear stochastic model of the cardiorespiratory interaction in terms of a set of polynomial basis functions representing the nonlinear force governing system oscillations. The strength and direction of coupling and noise intensity are simultaneously inferred from a univariate blood pressure signal. Our new inference technique does not require extensive global optimization, and it is applicable to a wide range of complex dynamical systems subject to noise.

Heart rate variability (HRV) is an important dynamical phenomenon in physiology. Altered HRV is associated with cardiovascular diseases and increased mortality [1], and its parameters are starting to be used for diagnostic tests. However, signals from the human cardiovascular system (CVS), being derived from a living organism, arise through the interaction of many dynamical degrees of freedom and processes with different time scales [2]. Thus HRV is attributable to the mutual interaction of a large number of oscillatory processes. Among them, the effect of respiration on heart rate has been the most intensively studied and a number of physiological mechanisms have been proposed [3], including, e.g., modulation of the cardiac filling pressure during respiratory movements [4] and baroreceptor feedback control [5].

An important feature of these processes is that they are nonlinear, time-varying, and subject to fluctuations [6–8]. For such systems, deterministic techniques fail to yield accurate parameter estimates [9]. Additionally, models of the cardiovascular interactions are not usually known exactly from first principles and one is faced with a rather broad range of possible parametric models to consider [5,10]. Inverse approaches, in which dynamical properties are analyzed from measured data, have recently been considered. A variety of numerical techniques have been introduced to analyze cardiorespiratory interactions using, e.g., linear approximations [11], estimations of either the strength of some of the nonlinear terms [12], the occurrence of cardiorespiratory synchronization [13], or the directionality of coupling [14]. Hitherto, modeling approaches have not been used interactively in conjunction with time-series analysis methods. Rather, the latter have each focused on a particular dynamical property, e.g., synchronization, or nonlinearities, or directionality.

In this Letter we introduce an approach to the problem that combines mathematical modeling of system dynamics and extraction of model parameters directly from measured time series. In this way we estimate simultaneously the strength, directionality of coupling, and noise intensity in the cardiorespiratory interaction. The technique reconstructs the nonlinear system dynamics in the presence of fluctuations. In addition, the method provides optimal compensation of dynamical noise-induced errors for continuous systems while avoiding extensive numerical optimization. We demonstrate the approach by using a univariate blood pressure (BP) signal for reconstruction of a nonlinear stochastic model of the cardiorespiratory interaction. The results are verified by analysis of data synthesized from the inferred model.

The problems faced in the analysis of CVS variability are common, not only to all living systems, but also to all complex systems subject to fluctuations, e.g., molecular motors [15] or coupled matter-radiation systems in astrophysics [16]. Yet there are no general methods for the dynamical inference of stochastic nonlinear systems. Thus the technique introduced in this Letter will be of wide applicability.

We analyze central venous blood pressure data, record 24 of the MGH/MF Waveform Database available at www.physionet.org. Its spectrum, shown in Fig. 1(a), exhibits two basic frequencies corresponding to respiratory, $f_r \approx 0.2$ Hz, and cardiac, $f_c \approx 1.7$ Hz, oscillations; the higher frequency peaks are the second, third, and fourth
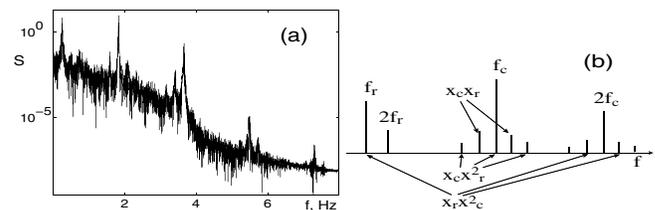

FIG. 1.   (a) Power spectrum of the BP data after Butterworth filtration: low-pass of the fourth order, with a cutoff frequency of 3 Hz, and high-pass of the second order with cutoff frequency of 0.03 Hz. (b) Summary of the main combinatorial frequencies of the cardiac and respiratory components. Correspondences between the nonlinear interaction terms of the model (1) and the observed frequencies are shown by arrows.

098101-1

harmonics of the cardiac oscillation. We note that the relative intensity and position of these peaks vary from subject to subject, with the average frequencies for healthy subjects at rest being around 0.2 and 1.1 Hz for respiration and heart rate, respectively.

We must bear in mind that CVS power spectra also contain lower frequency components [1,17]. In practice, parametric modeling is usually restricted to a specific part of the power spectrum. Because our interest here centers on the cardiorespiratory interaction, we select for study the frequency range that includes the main harmonics of cardiac and respiratory oscillations f_c and f_r and their combinational frequencies as shown in Fig. 1(b). In addition, we assume that the two higher basic frequency components observed in all CVS signals [8,18] can be separated. Hence the blood pressure signal can be approximated as a sum of the cardiac and respiratory oscillatory components $s(t) = s_c(t) + s_r(t)$. Accordingly, we use a combination of zero-phase forward and reverse digital filtering based on

Butterworth filters to decompose [19] it into two-dimensional time series $\{\mathbf{s}(t_k) = [s_c(t_k), s_r(t_k)], t_k = kh, k = 0:K\}$ representing the cardiac and respiratory oscillatory contributions, respectively. A 18 000-point window out of the original signal, sampled at 360 Hz, was resampled at 90 Hz yielding a signal for inference of length 500 s, with a step size of $h = 1/90$ s.

Following the suggestion of coupled oscillators [8,18], we now choose the simplest model that can reproduce this type of oscillation: two nonlinearly coupled systems with limit cycles on a plane

$$\dot{x}_r = a_1 x_r + y_r, \dot{y}_r = \alpha_i \phi_i(\mathbf{x}, \mathbf{y}) + \sqrt{D_{1j}}\xi_j,$$
$$\dot{x}_c = a_2 x_c + y_c, \dot{y}_c = \beta_i \phi_i(\mathbf{x}, \mathbf{y}) + \sqrt{D_{2j}}\xi_j. \quad (1)$$

Here $\xi_j(t)$ are zero-mean white Gaussian noises, the summation is taken over repeated indexes $i = 1, \ldots, 22$ and $j = r, c$, and the base functions are

$$\phi = \{1, x_r, x_c, y_r, y_c, x_r^2, x_c^2, y_r^2, y_c^2, x_r y_r, x_c y_c, x_r^3, x_c^3, x_r^2 y_r, x_c^2 y_c, x_r y_r^2, x_c y_c^2, y_r^3, y_c^3, x_r x_c, x_r^2 x_c, x_r x_c^2\}, \quad (2)$$

including nonlinear coupling terms up to third order. We assume that measurement noise can be neglected. The two dynamical variables of the model (1), $x_r(t)$ and $x_c(t)$ correspond to $\mathbf{s}(t) = \{s_r(t), s_c(t)\}$, introduced above. Using (1) the remaining two dynamical variables $\mathbf{y}(t) = \{y_r(t), y_c(t)\}$ can be related to observations $\{\mathbf{s}(t_k)\}$ as

$$b_n y_n(t_k) = \frac{s_n(t_k + h) - s_n(t_k - h)}{2h} + a_n s_n(t_k), \quad (3)$$

where $n = r, c$. Parametric representation (1) with a special form of embedding (3) allows one to infer a wide class of dynamical models including, e.g., the van der Pol and FitzHugh-Nagumo models. Furthermore, it allows physiological interpretation of model parameters.

Using (3) we can reduce the original problem of characterizing the cardiorespiratory interaction to that of inferring the set of unknown parameters $\mathcal{M} = \{\mathbf{c}, \hat{\mathbf{D}}\}$ of the coupled stochastic nonlinear differential equations

$$\dot{\mathbf{y}} = \hat{\mathbf{U}}(\mathbf{s}, \mathbf{y})\mathbf{c} + \sqrt{\hat{\mathbf{D}}}\xi(t). \quad (4)$$

Here $\xi(t)$ is a two-dimensional Gaussian white noise with independent components mixed with unknown correlation matrix $\hat{\mathbf{D}}$. The matrix $\hat{\mathbf{U}}$ will have the following block structure

$$\hat{\mathbf{U}} = \left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} x_r & 0 \\ 0 & x_r \end{bmatrix} \cdots \begin{bmatrix} x_r x_c^2 & 0 \\ 0 & x_r x_c^2 \end{bmatrix}\right]. \quad (5)$$

The vector of unknown coefficients $\mathbf{c} = \{\alpha_1, \beta_1, \ldots, \alpha_{22}, \beta_{22}\}$ with length $M = 2B$, has $B = 22$ diagonal blocks of size 2×2 formed by the basis functions (2).

The model parameters can be obtained by use of our novel stochastic dynamical inference method, based on the

Bayesian technique. Details, and a comparison with the results of earlier research, are given elsewhere [20]. Here we describe briefly the main steps in applying the method to infer cardiorespiratory interactions. First, one defines the so-called *likelihood* function $\ell(\mathbf{y}|\mathcal{M})$: the probability density to observe the dynamical variables $\mathbf{y}(t)$ under the condition that the underlying dynamical model (4) has a given set of parameters $\mathcal{M}$. We suggest that, for a uniform sampling scheme and a sufficiently small time step h, one can use results from [21] to write the logarithm of the likelihood function as

$$-\frac{2}{K}\log\ell(\mathbf{y}|\mathcal{M}) = \ln\det\hat{\mathbf{D}} + \frac{h}{K}\sum_{k=0}^{K-1}[\mathbf{v}(\mathbf{y}_k)\mathbf{c}$$
$$+ (\dot{\mathbf{y}}_k - \hat{\mathbf{U}}_k\mathbf{c})^T \hat{\mathbf{D}}^{-1}(\dot{\mathbf{y}}_k - \hat{\mathbf{U}}_k\mathbf{c})]$$
$$+ N\ln(2\pi h). \quad (6)$$

Here $\hat{\mathbf{U}}_k \equiv \hat{\mathbf{U}}(\mathbf{y}_k)$, $\dot{\mathbf{y}}_k \equiv h^{-1}(\mathbf{y}_{k+1} - \mathbf{y}_k)$, and the vector $\mathbf{v}(\mathbf{x})$ has components

$$v_m(\mathbf{x}) = \sum_{n=1}^{N} \frac{\partial U_{nm}(\mathbf{x})}{\partial x_n}, \quad m = 1:M.$$

Note that the form of (6) differs from the cost function in the method of least-squares: the term involving $\mathbf{v}$ provides optimal compensation of noise-induced errors [20]. In the next step, one has to summarize *a priori* expert knowledge about the model parameters in the so-called *prior* probability density function (PDF), $p_{\mathrm{pr}}(\mathcal{M})$. We assume $p_{\mathrm{pr}}(\mathcal{M})$ to be Gaussian with respect to the elements of $\mathbf{c}$ and uniform with respect to the elements of $\hat{\mathbf{D}}$.

Finally, one can use the measured time series $\mathbf{y}$ to improve the *a priori* estimation of the model parameters.

The improved knowledge is summarized in the *posterior* conditional PDF $p_{post}(\mathcal{M}|\mathbf{y})$, which is related to the *prior* PDF via Bayes' theorem

$$p_{post}(\mathcal{M}|\mathbf{y}) = \frac{\ell(\mathbf{y}|\mathcal{M})p_{pr}(\mathcal{M})}{\int \ell(\mathbf{y}|\mathcal{M})p_{pr}(\mathcal{M})d\mathcal{M}}. \qquad (7)$$

For a sufficiently large number of observations, p_{post} is sharply peaked at a certain most probable model $\mathcal{M} = \mathcal{M}^*$, providing a solution to the inference problem. To find it, we substitute the prior $p_{pr}(\mathcal{M})$ and the likelihood $\ell(\mathbf{y}|\mathcal{M})$ into (7) and perform the optimization by differentiation of the resulting expression with respect to $\hat{\mathbf{D}}_{\mathbf{y}}^{nn'}$ and c_m, yielding

$$\hat{\mathbf{D}}_{post}^{nn'}(\mathbf{c}) \equiv \frac{1}{K} \sum_{k=0}^{K-1} [\dot{\mathbf{y}}_k - \hat{\mathbf{U}}_k \mathbf{c}]_n [\dot{\mathbf{y}}_k - \hat{\mathbf{U}}_k \mathbf{c}]_{n'}^T, \qquad (8)$$

$$\mathbf{c}_{post}'(\hat{\mathbf{D}}) = \hat{\Xi}_{\mathbf{y}}^{-1}(\hat{\mathbf{D}})\mathbf{w}_{\mathbf{y}}(\hat{\mathbf{D}}), \qquad \hat{\mathbf{U}}_k \equiv \hat{\mathbf{U}}(\mathbf{y}_k). \qquad (9)$$

Here, use was made of the definitions

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

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

We repeat this two-step optimization procedure iteratively, starting from arbitrary prior values $\mathbf{c}_{pr}$ and $\hat{\Sigma}_{pr}$. We emphasize that a number of important parameters of the decomposition of the original signal (e.g., the bandwidth, order of the filters, and scaling parameters a_{ki}) have to be selected to provide the best fit to the measured time series $\{\mathbf{s}(t_k)\}$. The parameters of the model (4) can now be inferred directly from the measured time series of blood pressure, yielding the values shown in the first row of Table I. The spectra of the inferred, $x_r(t)$, and the measured, $s_r(t)$, cardiac oscillations are compared in Fig. 2. Similar results are obtained for the respiratory oscillations. In particular, the parameters of the nonlinear coupling and of the noise

intensity of the cardiac oscillations are $\beta_{20} = 2.2$, $\beta_{21} = 0.27$, $\beta_{22} = -8.67$, and $D_{22} = 8.13$; here we use a double-indexing scheme for the coefficients of the linear expansion (2), the scheme being evident from the caption in Table I. There is clearly a close resemblance between spectral peaks at the basic and combinational frequencies, $nf_c + mf_r$. A similarly close resemblance is found for respiratory oscillations, $s_r(t)$ and $x_r(t)$, respectively.

The frequency content can be reproduced from a univariate signal $s(t)$ because for $f_r \ll f_c$ it can be written in the form: $s(t) \approx s_r(t) + A_c(t)\cos[f_c t + \theta_c(t)] + \ldots$, where $A_c(t)$, $\theta_c(t)$ are slow amplitude and phase and the omitted terms oscillate at multiples of f_c. Fast-oscillating terms in this expansion correspond to a cardiac signal $s_c(t)$ and this ensures the validity of the signal decomposition $s(t) = s_r(t) + s_c(t)$, with components corresponding to weakly coupled nonlinear oscillators.

To validate these results we consider a synthesized signal $x(t) = x_r(t) + x_c(t)$ where $x_r(t)$, $x_c(t)$ are obtained using numerical simulations of the model (1) with parameters taken from the inference. Repetition of the full inference procedure, using the synthesized univariate signal $x(t)$ as a time-series data input $s(t)$, gives us the cardiac oscillator parameters $\beta_{20} = 6.32$, $\beta_{21} = 0.49$, $\beta_{22} = 6.03$, and $D_{22} = 3.44$. These differ from the values in the first row of Table I, but provide a correct estimation of the order of magnitude of the absolute values of the measured parameters. The main source of error here is the fact that we have to reconstruct the state of multidimensional system using the univariate signal.

If the state of the system was known, the accuracy of inference could be arbitrarily high [20]. We illustrate this point by using the synthesized time series $\{x_r(t), x_c(t), y_r(t), y_c(t)\}$ as bivariate data for two coupled oscillators to infer parameters of the model (1). The results, summarized in the second row of Table I, show that parameter

TABLE I. Coefficients corresponding to the last three base functions in (2), $\{x_r x_c, x_r^2 x_c, x_r x_c^2\}$, with $\{\alpha_i\}$ corresponding to the respiration coupling to cardiac rhythm and $\{\beta_i\}$ to the cardiac oscillation coupling to respiration. The top row gives coefficients inferred from measured data. The middle row represents coefficients inferred from synthesized data, obtained as an average of 100 nonoverlapped 1600 s blocks. Each block includes 160 000 points with a sampling time 0.01 sec. The estimation error is shown in the bottom line.

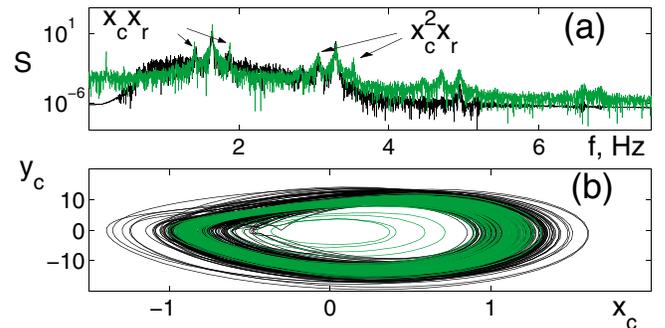| α_{20}	β_{20}	α_{21}	β_{21}	α_{22}	β_{22}	D_{11}	D_{22}
0.12	2.20	0.048	0.27	-0.066	-8.67	0.18	8.13
0.12	2.41	0.048	0.28	-0.070	-8.61	0.18	8.14
2.9%	9.3%	1.8%	5.6%	5.2%	0.7%	0.2%	0.2%

FIG. 2 (color). (a) Power spectra of cardiac oscillations obtained from measured data (black line) and from the synthesized model signal (green line). Arrows summarize combinational frequencies recovered in our analysis, corresponding to the nonlinear cardiorespiratory interaction. (b) Limit cycles of the cardiac oscillations $[x_c(n), y_c(n)]$ obtained from measured data (black line) and the synthesized signal (green line).

values can be estimated with relative error $<10\%$. Estimation of the noise intensity now has a relative error $<4\%$. Accuracy of the estimation can be further improved by increasing the total observation time. The decomposition problem could, of course, be eliminated by using bivariate cardiovascular data, which are now commonly available.

The relative magnitudes of the parameters obtained, $|\beta_i| > |\alpha_i|$, indicate that respiration influences cardiac activity more strongly than vice versa, consistent with the results of methods specifically developed for detecting the coupling directionality of interacting oscillators [14], and with direct physiological observations. Furthermore, the presence of nonzero quadratic terms is consistent with recent results obtained by time-phase bispectral analysis [12]. The frequency and amplitude variability of the main oscillatory components [8] is implicitly captured within the coupling terms and noise. We find that the present model class is able to reproduce, not only the coupling directionality, but also to a large extent the 1:7 and 1:8 cardiorespiratory synchronization properties of the measured data, as will be discussed in detail elsewhere.

In summary, we have solved a long-standing problem in physiology: inference of a nonlinear model of cardiorespiratory interactions in the presence of fluctuations. Our technique estimates simultaneously the strength and directionality of coupling, and the noise intensity in the cardiorespiratory interaction, directly from measured time series. It has also demonstrated stable and reliable inference of a broad class of models with high accuracy [20] and can, in principle, be applied to any physiological signal. Although our technique is only a first step in the development of path-integral-based dynamical inference of stochastic nonlinear models, it can readily be extended to encompass frequencies below that of respiration, nervous feedback control mechanisms, time-delay functions, nonpolynomial basis functions, parametric, and nonwhite noise. Our solution is facilitated by an analytic derivation of the likelihood function that optimally compensates noise-induced errors in continuous dynamical systems. It has enabled us to effect the first application of nonlinear stochastic inference to identify a dynamical model from real data.

[1] A. J. Camm *et al.*, Circulation **93**, 1043 (1996).
[2] A. T. Winfree, *The Geometry of Biological Time* (Springer-Verlag, Berlin, 1980); L. Glass and M. C. Mackey, *From Clocks to Chaos: The Rhythms of Life* (Princeton University Press, Princeton, NJ, 1988).
[3] D. L. Eckberg, J. Physiol.–London **548**, 339 (2003).
[4] M. B. Visscher, A. Rupp, and F. H. Scott, Am. J. Physiol. **70**, 586 (1924).
[5] R. W. deBoer, J. M. Karemaker, and J. Strackee, Am. J. Physiol. **253**, H680 (1987).
[6] J. P. Saul, D. T. Kaplan, and R. I. Kitney, *Computers in Cardiology* (IEEE Comput. Soc. Press, Washington, 1988), pp. 299–302.
[7] K. H. Chon, T. J. Mullen, and R. J. Cohen, IEEE Trans. Biomed. Eng. **43**, 530 (1996); K. Suder, F. R. Drepper, M. Schiek, and H. H. Abel, Am. J. Physiol.: Heart. Circ. Physiol. **275**, H1092 (1998).
[8] A. Stefanovska and M. Bračič, Contemp. Phys. **40**, 31 (1999).
[9] E. J. Kostelich, Physica D (Amsterdam) **58**, 138 (1992); P. E. McSharry and L. A. Smith, Phys. Rev. Lett. **83**, 4285 (1999).
[10] M. Clynes, J. Appl. Physiol. **15**, 863 (1960); G. Baselli, S. Cerutti, A. Malliani, and M. Pagani, IEEE Trans. Biomed. Eng. **35**, 1033 (1988); B. J. TenVoorde *et al.*, in *Computer Analysis of Cardiovascular Signals*, edited by M. Di Renzo *et al.* (IOS Press, Amsterdam, 1995); H. Seidel and H. Herzel, in *Modeling the Dynamics of Biological Systems*, edited by E. Mosekilde and O. G. Mouritsen (Springer, New York, 1996), pp. 205–229; S. Cavalcanti and E. Belardinelli, IEEE Trans. Biomed. Eng. **43**, 982 (1996); K. Kotani *et al.*, Phys. Rev. E **65**, 051923 (2002).
[11] R. D. Berger, J. P. Saul, and R. J. Cohen, Am. J. Physiol.: Heart. Circ. Physiol. **256**, H142 (1989); J. A. Taylor *et al.*, Am. J. Physiol.: Heart. Circ. Physiol. **280**, H2804 (2001); R. Mukkamala and R. J. Cohen, Am. J. Physiol.: Heart Circ. Physiol. **281**, H2714 (2001); S. Lu, K. H. Ju, and K. H. Chon, IEEE Trans. Biomed. Eng. **48**, 1116 (2001).
[12] J. Jamšek, A. Stefanovska, P. V. E. McClintock, and I. A. Khovanov, Phys. Rev. E **68**, 016201 (2003); J. Jamšek, A. Stefanovska, and P. V. E. McClintock, Phys. Med. Biol. **49**, 4407 (2004).
[13] C. Schäfer, M. G. Rosenblum, J. Kurths, and H. H. Abel, Nature (London) **392**, 239 (1998); N. B. Janson, A. G. Balanov, V. S. Anishchenko, and P. V. E. McClintock, Phys. Rev. Lett. **86**, 1749 (2001).
[14] M. G. Rosenblum *et al.*, Phys. Rev. E **65**, 041909 (2002); M. Paluš and A. Stefanovska, Phys. Rev. E **67**, 055201(R) (2003).
[15] K. Visscher, M. J. Schnitzer, and S. M. Block, Nature (London) **400**, 184 (1999).
[16] J. Christensen-Dalsgaard, Rev. Mod. Phys. **74**, 1073 (2002).
[17] A. Stefanovska and P. Krošelj, Open Syst. Inf. Dyn. **4**, 457 (1997); J. A. Taylor, D. L. Carr, C. W. Myers, and D. L. Eckberg, Circulation **98**, 547 (1998); A. Stefanovska, D. G. Luchinsky, and P. V. E. McClintock, Physiol. Meas. **22**, 551 (2001).
[18] A. Stefanovska, M. Bračič Lotrič, S. Strle, and H. Haken, Physiol. Meas. **22**, 535 (2001).
[19] Alternative approaches could include, e.g., empirical mode decomposition, Karhunen-Levé decomposition, or independent component analysis.
[20] V. N. Smelyanskiy *et al.*, cond-mat/0409282.
[21] R. Graham, Z. Phys. B **26**, 281 (1977).