Nonlinear noise reduction using reference data

Karsten Sternickel,[1] Arndt Effern,[1,2] Klaus Lehnertz,[2] Thomas Schreiber,[3] and Peter David[1]
[1]*Institut für Strahlen- und Kernphysik, University of Bonn, Nussallee 14-16, 53115 Bonn, Germany*
[2]*Department of Epileptology, Medical Center, University of Bonn, Sigmund-Freud Strasse 25, 53105 Bonn, Germany*
[3]*Max-Planck-Institute for Physics of Complex Systems, Nöthnitzer Strassee 38, 01187 Dresden, Germany*
(Received 12 April 2000; published 23 February 2001)

We introduce a method to clean uncorrelated deterministic and stochastic noise components from time series. It combines recently developed techniques for nonlinear projection with properties of the wavelet transform to extract noise in state space. The method requires that time series are generated by a dynamical system which is at least approximately deterministic and that they are recorded together with a reference signal. Its efficiency was tested on both simulated signals and measured magnetic fields of the heart. Convincing results are obtained even at low signal-to-noise ratios.

DOI: 10.1103/PhysRevE.63.036209 PACS number(s): 05.45.Tp, 05.40.Ca, 87.19.Hh, 87.50.Mn

Time series representing the sampling values of an observed dynamical process allow one to analyze the underlying dynamics. Generally, time series are contaminated with noise so that a suitable filtering is required prior to further data processing [1,2]. In many physical experiments the simultaneous recording of reference time series is possible. The reference is provided by an additional sensor **R** that records mainly the noise whereas the main sensor **S** records both the signal and the noise. In general, **R** and **S** will record different projections of the noise due to their spatial separation. This setup is used in various fields, e.g., in optical [3], acoustical [4], or geophysical [5] applications. It is particularly helpful in the case of high-amplitude deterministic disturbances with discrete frequency peaks in the power spectrum; see, e.g., [6,7]. In typical field experiments these deterministic disturbances are mainly caused by the 50 or 60 Hz interference and its harmonics. If the noise is uncorrelated with the signal, we can write

$$\mathbf{S}=(s_1,\ldots,s_N) \quad \text{with} \quad s_n=x_n+\epsilon_n^s, \tag{1}$$

$$\mathbf{R}=(r_1,\ldots,r_N) \quad \text{with} \quad r_n=\epsilon_n^r,$$

where x_n denotes the pure signal and ϵ_n^s and ϵ_n^r the superimposed noise at time n, respectively. Both ϵ_n^s and ϵ_n^r can represent deterministic ($^{\mathrm{det}}\epsilon_n^s$, $^{\mathrm{det}}\epsilon_n^r$) as well as stochastic ($^{\mathrm{stoch}}\epsilon_n^s$, $^{\mathrm{stoch}}\epsilon_n^r$) disturbances. Usually, the two sensors have different sensitivities and ϵ_n^s and ϵ_n^r are not identical. In addition, their spatial distance makes a matching of ϵ_n^s and ϵ_n^r very difficult. Thus, a simple subtraction may lead to artifacts in signal reconstruction. A suitable strategy is to determine coupling factors in the time or frequency domain. A common method to determine these factors is the application of Wiener filters that utilize signal averages in the frequency domain [1]. This, however, requires initial knowledge about the signal that is often not available at high noise levels.

Considering these problems, an approach independent of the different sensitivities of the sensors appears to be advantageous. In Ref. [8] a geometric approach for noise reduction in deterministic chaotic data was introduced. It is based on a state space reconstruction and carries out local projections onto a noise-free subspace. These projections are usually performed by initially fixing the dimension of the subspace and by estimating the preferred direction of trajectories using a singular value decomposition [9]. However, data limitations can make it difficult to identify the principal directions for small neighborhoods in relatively high-dimensional spaces. With a suitable set of given basis functions, only the coefficients need to be determined, which is much less demanding on the data. The discrete wavelet transform [10] provides a good choice, suitable for smooth functions as well as for singularities. Moreover, hard or soft thresholding of wavelet coefficients is well suited for signal recovery [11] even in state space [12].

In this article we show that local projections in state space using wavelets represents an improved method for nonlinear noise reduction. This holds true even for high noise amplitudes if the experimental setup permits the recording of a reference signal. To demonstrate the efficiency of this method we have used it on simulated and measured cardiographic data.

Let $\vec{s}_n=(s_n,\ldots,s_{n-(m-1)\tau})\in\mathcal{S}$ and $\vec{r}_n=(r_n,\ldots,r_{n-(m-1)\tau})\in\mathcal{R}$ be vectors in the reconstructed state space of **S** and **R** where τ is a time delay and m an embedding dimension. Appropriate choices of m and τ enable unfolding of the trajectories in $\mathcal{S}$ and $\mathcal{R}$.

For the first step of our denoising procedure we assume x_n and ϵ_n^s are statistically independent, while ϵ_n^s and ϵ_n^r are strongly correlated. This correlation need not be linear, but we assume that there is a strict functional relationship. Furthermore, we assume x_n and $^{\mathrm{det}}\epsilon_n^s$ to be confined to different subspaces $\mathcal{S}^{\mathrm{x}}$ and $\mathcal{S}^{\mathrm{det}}$ in $\mathcal{S}$ with dimensions $d_{S^{\mathrm{x}}}\ll m$ and $d_{S^{\mathrm{det}}}\ll m$. In order to identify $\mathcal{S}^{\mathrm{x}}$ and $\mathcal{S}^{\mathrm{det}}$ we determine k nearest neighbors $(\vec{r}_{n_1},\ldots,\vec{r}_{n_k})$ of $\vec{r}_n$ in $\mathcal{R}$. Every $\vec{r}_n$ has a temporally corresponding $\vec{s}_n$ in $\mathcal{S}$ and so do its neighbors. If we now treat x_n as a distortion of the deterministic noise trajectories in $\mathcal{S}$ temporally corresponding neighbors of $\vec{r}_n$ are mainly displaced in $\mathcal{S}^{\mathrm{x}}$. We take this as a criterion to identify $\mathcal{S}^{\mathrm{x}}$ and $\mathcal{S}^{\mathrm{det}}$. Note that different properties of the

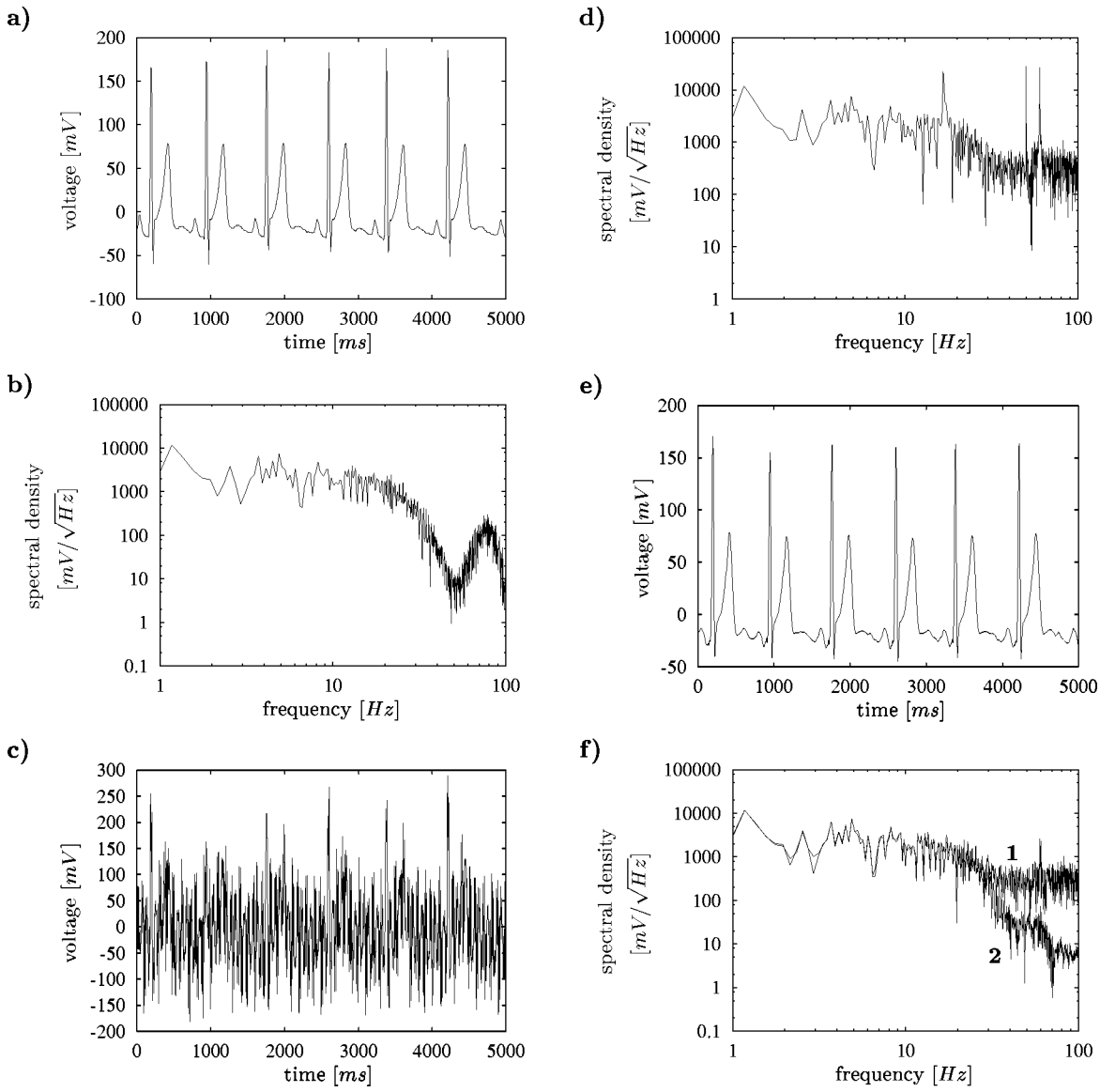

FIG. 1. (a) shows the original ECG and (b) its corresponding Fourier spectrum. The effect of the 50 Hz notch filter is obvious. (c) depicts the noise-contaminated signal. The superimposed noise causes a general increase of the Fourier spectrum as well as sharp peaks due to deterministic noise components (d). The result of noise reduction is shown in (e). The cleaned time series is almost free of artifacts. In (f) line 1 illustrates the effect of the first noise reduction step [the sharp noise peaks disappeared, cf. (d)] whereas line 2 shows the complete correction.

applied measuring systems keep the relationship of neighbors of $\vec{r}_n$ in $\mathcal{S}$ unchanged: different amplitude offsets in **S** and **R** lead to tangential shifts of all trajectories, phase shifts between **S** and **R** lead to spatial rotation, and different amplitudes in the two time series lead to spatial expansion.

To separate $\mathcal{S}^{\mathrm{x}}$ and $\mathcal{S}^{\mathrm{det}}$ we assume that both are spanned by different translations and dilatations of a mother wavelet. Let $\vec{\psi}_n=(\psi_{n,1},\psi_{n,2},\ldots,\psi_{n,m})$ denote the wavelet transform of $\vec{s}_n$ and $\vec{\psi}_{n_1},\ldots,\vec{\psi}_{n_k}$ the wavelet transforms of its neighbors. To identify $\mathcal{S}^{\mathrm{det}}$, the center of mass of each $\vec{\psi}_n$ and its neighbors is calculated. Here $C_i(n)=\langle\vec{\psi}_n,\vec{\psi}_{n_1},\ldots,\vec{\psi}_{n_k}\rangle_i$ defines the ith center of mass component and $\sigma_i^2(n)$ the corresponding variance. We expect that the ratio $C_i(n)/\sigma_i^2(n)$ is smaller for $\mathcal{S}^{\mathrm{x}}$ than for $\mathcal{S}^{\mathrm{det}}$. Thus, the shrinking condition

$$\tilde{\psi}_{n,i}=\begin{cases}\psi_{n,i}, & |C_i(n)|\geqslant 2\lambda_R\dfrac{\sigma_i(n)}{\sqrt{k+1}}\\ C_i(n), & \text{else}\end{cases}\qquad(2)$$

defines a projection of $\mathcal{S}$ onto $\mathcal{S}^{\mathrm{det}}$. Here λ_R denotes a cutoff value that depends on specific properties of the measured data (1). The inverse wavelet transform of $\tilde{\psi}_n=(\tilde{\psi}_{n,1},\tilde{\psi}_{n,2},\ldots,\tilde{\psi}_{n,m})$ allows reconstruction of $\epsilon_n^{\mathrm{det}}$ so that a simple subtraction from **S** results in a time series $\hat{\mathbf{S}}$. Its state

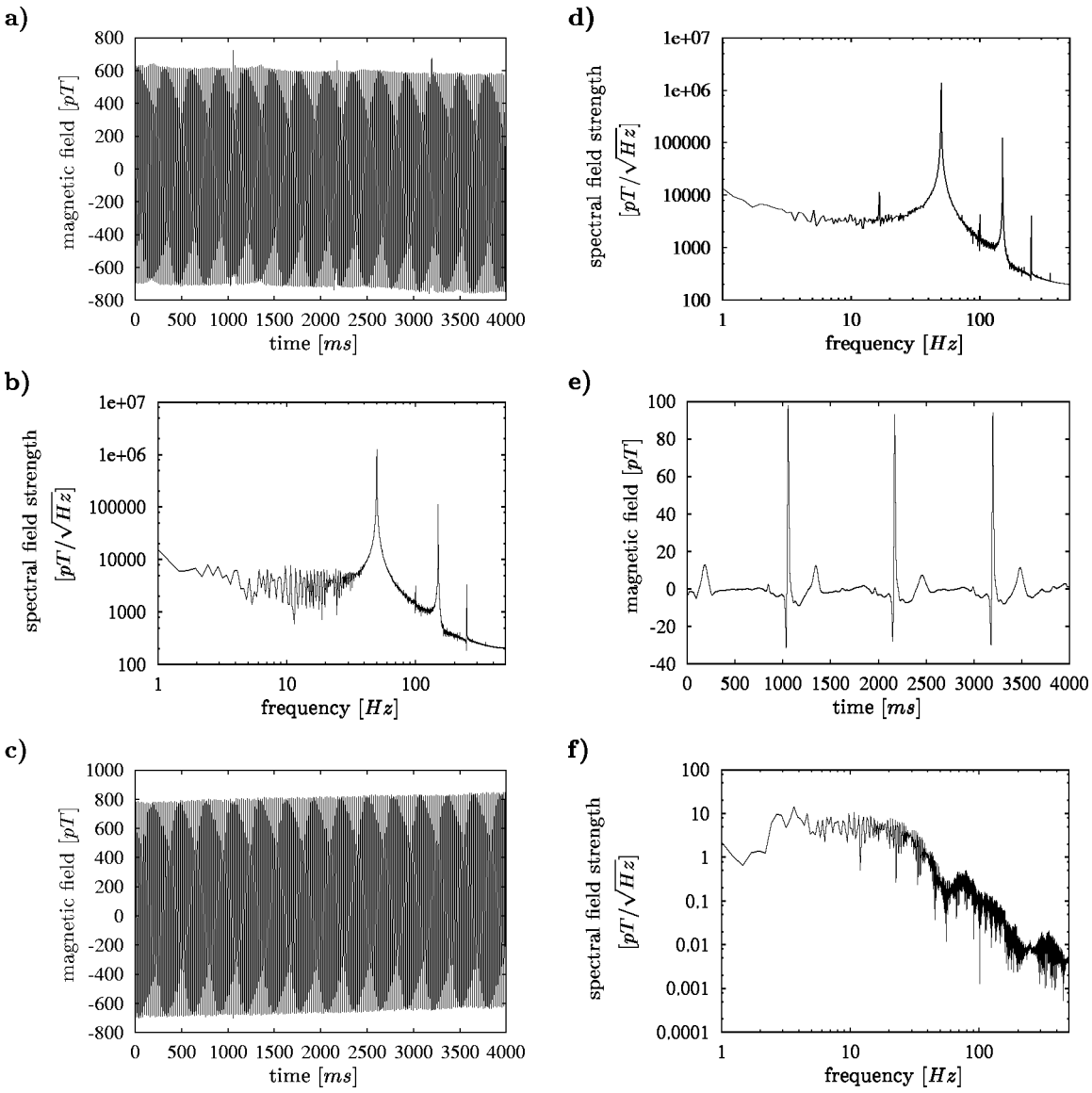

FIG. 2. (a) shows the MCG sequence recorded outside a shielding room. Only the main component of the heart signal (R wave) is visible. The width of the 50 Hz peak in the corresponding Fourier spectrum (b) actually prohibits the application of a notch filter. (c) exhibits the reference signal and (d) its Fourier spectrum. The noise reduced MCG sequence (e) even shows details of the heartbeat and an almost noise-free baseline. The corresponding Fourier spectrum is given in (f).

space $\hat{\mathcal{S}}$ is now dominated by the determinism of the signal x_n, and the remaining disturbances are caused by stochastic noise components $\epsilon_n^{\text{stoch}}$ only.

In the second step, $\epsilon_n^{\text{stoch}}$ is eliminated using our procedure once again, but this time we determine the nearest neighbors $(\vec{s}_{n_1}, \ldots, \vec{s}_{n_k})$ in $\hat{\mathcal{S}}$ and calculate their wavelet transforms. For the shrinking condition (2) we now use λ_S instead of λ_R to define a projection of $\hat{\mathcal{S}}$ onto a noise reduced subspace $\tilde{\mathcal{S}}$. Finally, the inverse wavelet transform recovers the delay vectors from which the cleaned time series $\tilde{\mathbf{S}}$ can be reconstructed.

To demonstrate the efficiency of this noise reduction scheme we first apply it to simulated signals. We use 5 s electrocardiographic (ECG) data of a healthy person recorded at 200 Hz as the pure signal recorded by the main sensor. The ECG was prefiltered by a 50 Hz notch filter and a second-order low-pass filter at 100 Hz [Figs. 1(a) and 1(b)]. We add white noise with an amplitude variance of 30% referred to the ECGs variance. The deterministic noise had frequency peaks at $16\frac{2}{3}$ Hz, 50 Hz (power supply and subharmonics), and 60 Hz (signal analysis systems) with an amplitude variance of 100% [see Figs. 1(c) and 1(d)]. Generally, these frequencies are the strongest disturbances to overlay physiological measurements. We generate the reference time series by creating noise using the same parameters as mentioned above, but additionally with variations in amplitude and a constant phase shift for the deterministic noise components. We define a noise reduction coefficient by

$$\rho = \sqrt{\frac{\langle (s_i - x_i)^2 \rangle}{\langle (\tilde{s}_i - x_i)^2 \rangle}}, \tag{3}$$

where $x_i, s_i \in \mathbf{S}$, and $\tilde{s}_i \in \tilde{\mathbf{S}}$ denote the ith components of the pure signal, the contaminated time series, and the noise reduced signal, respectively. The applied filter settings consist of an embedding dimension of $m=128$, a time delay of $\tau=1$ sample point, the number of nearest neighbors of $k=20$, and the coiflet 4 mother wavelet [13]. Empirically, we determine thresholding parameters $\lambda_R=0.75$ and $\lambda_S=1.0$ as most appropriate, resulting in $\rho=14.56$. Figures 1(e) and 1(f) show that the baseline between the heartbeats (as a good indicator for the noise reduction quality) is almost noise-free. Therefore, our method performs both signal preservation and considerable noise reduction.

As an example of measured data, we analyze 5 s of magnetocardiographic (MCG) data [cf. Figs. 2(a) and 2(b)] recorded with a high-T_c superconducting quantum interference device (SQUID) outside a shielding room [14]. Two axial gradiometers of first order with 7 cm baseline were mounted at a distance of 7 cm one above the other. In this setup the top gradiometer records the reference signal [cf. Figs. 2(c) and 2(d)]. To increase the signal-to-noise ratio SNR this measurement generally is performed in a shielding room although these rooms are expensive and immobile. In an unshielded environment it is difficult to detect any signals originating from the observed object.

Figures 2(e) and 2(f) show the time series along with its corresponding Fourier spectrum resulting from the presented noise reduction procedure using the following parameters: a sampling rate of 1000 Hz, an embedding dimension of $m=512$, a time delay of $\tau=1$ sample point, the number of nearest neighbors of $k=20$, the coiflet 2 mother wavelet, and thresholding parameters $\lambda_R=0.5$ and $\lambda_S=0.5$. In the reconstructed MCG even small details of the heartbeat are revealed. Again the baseline between the heartbeats is almost noise free.

In summary, the technique described substantially eliminates deterministic and stochastic noise from measured time series. This holds true even if the noise amplitude is of the same order of magnitude as (or higher than) the signal itself. Even for physiological data recorded in an unshielded environment its success could be demonstrated. Hence, the proposed noise reduction method represents an excellent alternative to similar techniques already in use [15]. Further improvements are currently being developed to accelerate our method for real-time application in order to control experimental setups.

We are grateful to Marcel Bick who kindly lent us the MCG data recorded at the Forschungszentrum Jülich, and to R. Andrzejak, P. Grassberger, F. Mormann, and C. Rieke for helpful comments on earlier versions of this manuscript. A.E. and K.L. acknowledge support from the Deutsche Forschungsgemeinschaft.

[1] S.D. Stearns and D.R. Hush, *Digital Signal Analysis* (Prentice Hall, Englewood Cliffs, NJ, 1994).
[2] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, 1997).
[3] H. Adams, D. Reinert, P. Kalkert, and W. Urban, Appl. Phys. B: Photophys. Laser Chem. **34**, 179 (1984).
[4] S.J. Kim and K.Y. Song, AIAA J. **37**, 1180 (1999).
[5] T.D. Gamble, W.M. Goubau, and J. Clarke, Geophysics **44**, 53 (1979).
[6] G.C. Lauchle, J.R. MacGillivray, and D.C. Swanson. J. Acoust. Soc. Am. **101**, 341 (1997).
[7] K.-H. Weck and F. Weinel, ETEP, **8**, 299 (1998).
[8] P. Grassberger, R. Hegger, H. Kantz, C. Schaffrath, and T. Schreiber, Chaos **41**, 127 (1993).
[9] T. Schreiber and D.T. Kaplan, Chaos **6**, 87 (1995).
[10] S.G. Mallat, IEEE Trans. Pattern Anal. Mach. Intell. **11**, 674 (1989).
[11] D.L. Donoho, I.M. Johnstone, and B.W. Silverman, IEEE Trans. Inf. Theory **41**, 613 (1995).
[12] A. Effern, K. Lehnertz, T. Grunwald, T. Schreiber, P. David, and C.E. Elger, Physica D (to be published).
[13] C.S. Burrus, R.A. Copinath, and H. Guo, *Wavelets and Wavelet Transforms* (Prentice Hall, Englewood Cliffs, NJ, 1998).
[14] H. Weinstock, *SQUID Sensors: Fundamentals, Fabrication and Applications* (Kluwer, Dordrecht, 1996).
[15] M. Roy, V.R. Kumar, B.D. Kulkarni, J. Sanderson, M. Rhodes, and M. vander Stappen, AIChE J. **45**, 2461 (1999).