

Chapter 9

On the Application of the SDLE to the Analysis of Complex Time Series

Jianbo Gao, Jing Hu, and Wen-wen Tung

Abstract Complex systems often generate highly nonstationary and multiscale signals because of nonlinear and stochastic interactions among their component systems and hierarchical regulations imposed by the operating environments. Rapid accumulation of such complex data in life sciences, systems biology, nano-sciences, information systems, and physical sciences has made it increasingly important to develop complexity measures that incorporate the concept of scale explicitly, so that different behaviors of signals on varying scales can be simultaneously characterized by the same scale-dependent measure. The scale-dependent Lyapunov exponent (SDLE) discussed here is such a measure and can be used as the basis for a unified theory of multiscale analysis of complex data. The SDLE can readily characterize deterministic low-dimensional chaos, noisy chaos, random $1/f^\alpha$ processes, random Levy processes, stochastic oscillations, and processes with multiple scaling behavior. It can also readily deal with many types of nonstationarity, detect intermittent chaos, and accurately detect epileptic seizures from EEG data and distinguish healthy subjects from patients with congestive heart failure from heart rate variability (HRV) data. More importantly, analyses of EEG and HRV data illustrate that commonly used complexity measures from information theory,

J. Gao (✉)

PMB Intelligence LLC, West Lafayette, IN 47996, USA

Mechanical and Materials Engineering, Wright State University, Dayton, Ohio 45435, USA

e-mail: jbgao.pmb@gmail.com

J. Hu

Affymetrix, Inc., 3380 Central Expressway, Santa Clara, CA 95051, USA

e-mail: jing.hu@gmail.com

W.-w. Tung

Department of Earth & Atmospheric Sciences, Purdue University,

West Lafayette, IN 47907, USA

e-mail: wwtung@purdue.edu

chaos theory, and random fractal theory can be related to the values of the SDLE at specific scales, and useful information on the structured components of the data is also embodied by the SDLE.

9.1 Introduction

Complex systems are usually comprised of multiple subsystems that exhibit both highly nonlinear deterministic, as well as, stochastic characteristics, and are regulated hierarchically. These systems generate signals that exhibit complex characteristics such as sensitive dependence on small disturbances, long memory, extreme variations, and nonstationarity [1]. Examples of such signals are abundant, including biological data such as heart rate variability (HRV) and EEG data [2], highly bursty traffic on the Internet [3–5], and highly varying stock prices and foreign exchange rates in financial markets [6, 7]. For illustration, in Fig. 9.1, we have shown an example of HRV data for a normal young subject [8]. Evidently, the signal is

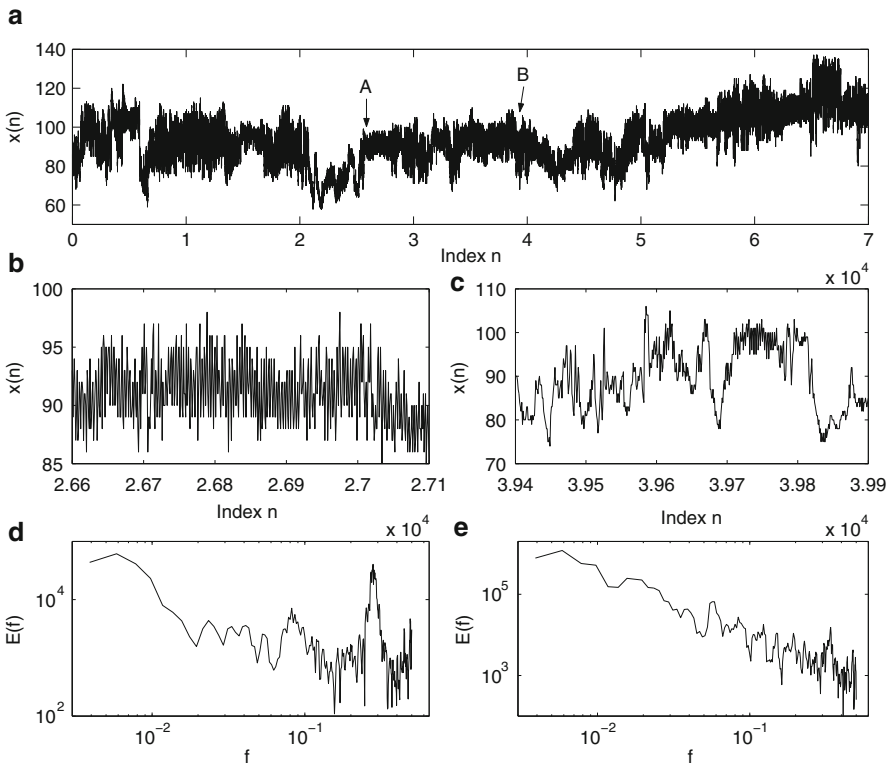


Fig. 9.1 (a) The HRV data for a normal subject; (b, c) the segments of signals indicated as A and B in (a); (d, e) power spectral density for the signals shown in (b,c)

highly nonstationary and multiscaled, appearing oscillatory for some period of time (Fig. 9.1b, d) and then varying as a $1/f$ process for another period of time (Fig. 9.1c, e).

While the multiscale nature of signals such as that shown in Fig. 9.1 cannot be fully characterized by existing methods, the nonstationarity of the data is even more troublesome, because it prevents direct application of spectral analysis or methods based on chaos theory and random fractal theory. For example, in order to reveal that the HRV data is of $1/f$ nature [9, 10] with anti-persistent long-range correlations [11, 12] and multifractality [13], time series such as shown in Fig. 9.1a have to be preprocessed to remove components (such as oscillatory ones) that do not conform to fractal scaling analysis. However, automated segmentation of complex biological signals to remove undesired components is a significant open problem, since it is closely related to the challenging task of accurately detecting transitions from normal to abnormal states in physiological data.

Rapid accumulation of complex data in life sciences, systems biology, nanosciences, information systems, and physical sciences has made it increasingly important to be able to analyze multiscale and nonstationary data. Since multiscale signals behave differently depending upon which scale the data are looked at, it is of fundamental importance to develop measures that explicitly incorporate the concept of scale so that different behaviors of the data on varying scales can be simultaneously characterized by the same scale-dependent measure. Here, we discuss such a measure, the scale-dependent Lyapunov exponent (SDLE), and develop a unified multiscale analysis theory of complex data.

This chapter is organized as follows. We first define the SDLE, then apply it to characterize low-dimensional chaos, noisy chaos, and random processes with power-law decaying power spectral density (so-called $1/f^\alpha$ processes). We then show how it can readily detect intermittent chaos, deal with nonstationarity, and apply it to characterize EEG and HRV data. Finally, we make a few concluding remarks, including a discussion of best practices for experimental data analysis.

9.2 SDLE: Definitions and Fundamental Properties

Chaos and random fractal theories have been used extensively in the analysis of complex data [2, 3, 6, 11–18]. Chaos theory shows that apparently irregular behaviors in a complex system may be generated by nonlinear deterministic interactions of only a few numbers of degrees of freedom while noise or intrinsic randomness does not play any role. Random fractal theory, on the other hand, assumes that the dynamics of the system are inherently random. One of the most important classes of random fractals is the set of $1/f^\alpha$ processes that display long-range correlations. Since the foundations of chaos theory and random fractal theory are entirely different, different conclusions may be drawn depending upon which theory is used to analyze the data. In fact, much of the research in the past has been devoted to determining whether a complex time series is generated

by a chaotic or a random system [19–30]. In this effort, $1/f^\alpha$ processes have distinguished themselves as the key counter examples invalidating commonly used tests for chaos [30–32]. Thus, successful classification of chaos and $1/f^\alpha$ processes based on scales may fundamentally change the practice of time series analysis—these theories will be used synergistically, instead of individually, to characterize the behaviors of signals on a wide range of scales.

SDLE is a generalization of two important concepts, the time-dependent exponent curves [24] and the finite size Lyapunov exponent (LE) [33]. It was first introduced by the authors in [34, 35], and has been further extended in [36, 37] and applied to study EEG [38], HRV [39, 40], and Earth’s geodynamo [41].

We assume that all that is known is a scalar time series $x(1), x(2), \dots, x(n)$. Regardless of whether the dynamics are chaotic or random, we use time-delay embedding [42–44] to form vectors of the form: $V_i = [x(i), x(i+L), \dots, x(i+(m-1)L)]$, where the embedding dimension m and the delay time L are chosen according to optimization criteria [24, 45]. When the time series is random, such a procedure transforms the self-affine stochastic process to a self-similar process in phase space. In this case, however, the specific values of m and L are not important, so long as $m > 1$.

After a proper phase space is reconstructed, we consider an ensemble of trajectories. We denote the initial separation between two nearby trajectories by ε_0 and their *average separation* at time t and $t + \Delta t$ by ε_t and $\varepsilon_{t+\Delta t}$, respectively. We then examine the relation between ε_t and $\varepsilon_{t+\Delta t}$, where Δt is small. When $\Delta t \rightarrow 0$, we have

$$\varepsilon_{t+\Delta t} = \varepsilon_t e^{\lambda(\varepsilon_t)\Delta t}, \quad (9.1)$$

where $\lambda(\varepsilon_t)$ is the SDLE. It is given by

$$\lambda(\varepsilon_t) = \frac{\ln \varepsilon_{t+\Delta t} - \ln \varepsilon_t}{\Delta t}. \quad (9.2)$$

Equivalently, we have a differential equation for ε_t ,

$$\frac{d\varepsilon_t}{dt} = \lambda(\varepsilon_t)\varepsilon_t. \quad (9.3)$$

Given a time series data, the smallest Δt possible is the sampling time τ .

To compute SDLE, we can start from an arbitrary number of shells,

$$\varepsilon_k \leq \|V_i - V_j\| \leq \varepsilon_k + \Delta\varepsilon_k, \quad k = 1, 2, 3, \dots, \quad (9.4)$$

where V_i, V_j are reconstructed vectors and ε_k (the radius of the shell) and $\Delta\varepsilon_k$ (the width of the shell) are arbitrarily chosen small distances ($\Delta\varepsilon_k$ is not necessarily a constant). We then monitor the evolution of all of the pairs of vectors (V_i, V_j) within a shell and take the average. As we will see shortly, as far as estimation of

the parameters corresponding to exponential or power-law divergence is concerned, taking logarithm and averaging can be exchanged; (9.2) can now be written as

$$\lambda(\varepsilon_t) = \frac{\langle \ln \|V_{i+t+\Delta t} - V_{j+t+\Delta t}\| - \ln \|V_{i+t} - V_{j+t}\| \rangle}{\Delta t} \tag{9.5}$$

where t and Δt are integers in units of the sampling time and the angle brackets denote the average within a shell.

To see why taking logarithm and averaging can be exchanged for the purpose of computing $\lambda(\varepsilon_t)$, let us consider a case involving $\varepsilon_1(t) = \varepsilon_1(0)e^{\lambda_1 t}$, $\varepsilon_2(t) = \varepsilon_2(0)e^{\lambda_2 t}$, where $\lambda_1 = \lambda_2$ is a positive constant. Let $\varepsilon(t)$ be the average of $\varepsilon_1(t)$ and $\varepsilon_2(t)$. Then it is clear that SDLE is simply λ whether one takes average first or takes logarithm first. In fact, when large t is concerned, if λ_1 is slightly larger than λ_2 , then taking logarithm first is beneficial, since, otherwise, the term $e^{\lambda_1 t}$ will dominate, and thus the presence of λ_2 will not be captured. Clearly, similar argument applies to the situation of power-law divergence.

Note that in the above computational procedure, the initial set of shells for computing SDLE serve as initial values of the scales; through evolution of the dynamics, they will automatically converge to the range of inherent scales—which are the scales that define (9.2) and (9.3). Also note that when analyzing chaotic time series, the condition

$$|j - i| \geq (m - 1)L \tag{9.6}$$

needs to be imposed when finding pairs of vectors within a shell, to eliminate the effects of tangential motions [24] and for an initial scale to converge to the inherent scales [35].

To better understand the notion of “inherent scales,” it is beneficial to discuss the notion “characteristic scale” (or “limiting scale”), ε_∞ , defined as the scale where SDLE is close to 0. If one starts from $\varepsilon_0 \ll \varepsilon_\infty$, then, regardless of whether the data is deterministically chaotic or simply random, ε_t will initially increase with time and gradually settle around ε_∞ . Consequentially, $\lambda(\varepsilon_t)$ will be positive before ε_t reaches ε_∞ . On the other hand, if one starts from $\varepsilon_0 \gg \varepsilon_\infty$, then ε_t will simply decrease, yielding negative $\lambda(\varepsilon_t)$, again regardless of whether the data are chaotic or random. When $\varepsilon_0 \sim \varepsilon_\infty$, then $\lambda(\varepsilon_t)$ will stay around 0—note, however, that ε_∞ may not be a single point but a function of time, such as a periodic function of time. These discussions make it clear that chaos can only be observed on scales much smaller than ε_∞ .

To better understand SDLE, we now point out a relation between SDLE and the largest positive LE λ_1 estimated for a true chaotic signal using, say, the Wolf et al.’s algorithm [21]. It is given by [35]

$$\lambda_1 = \int_0^{\varepsilon^*} \lambda(\varepsilon)p(\varepsilon)d\varepsilon, \tag{9.7}$$

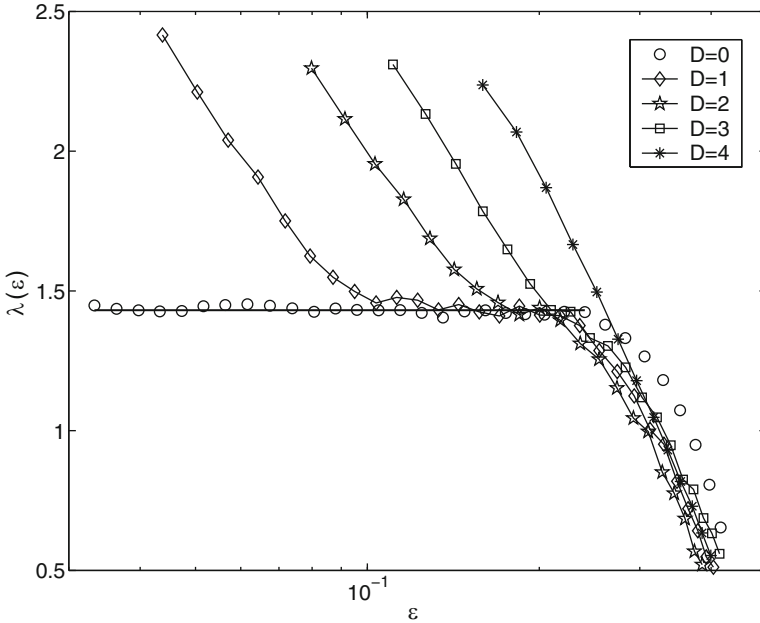


Fig. 9.2 $\lambda(\varepsilon)$ curves for clean and noisy Lorenz systems

where ε^* is a scale parameter (e.g., used for renormalization when using Wolf et al.’s algorithm [21]); $p(\varepsilon)$ is the probability density function for the scale ε given by

$$p(\varepsilon) = Z \frac{dC(\varepsilon)}{d\varepsilon}, \tag{9.8}$$

where Z is a normalization constant satisfying $\int_0^{\varepsilon^*} p(\varepsilon)d\varepsilon = 1$ and $C(\varepsilon)$ is the well-known Grassberger–Procaccia’s correlation integral [19]. Note that the lower bound for the integration is set to be zero here. In practice, on scales smaller than ε_{\min} , the probability $p(\varepsilon)$ will be zero. Therefore, one could replace the lower bound for the integration by ε_{\min} .

To understand the SDLE, it is instructive to apply it to characterize chaotic signals and $1/f^\alpha$ processes. First, we analyze the chaotic Lorenz system with stochastic forcing

$$\begin{aligned} dx/dt &= -16(x - y) + D\eta_1(t), \\ dy/dt &= -xz + 45.92x - y + D\eta_2(t), \\ dz/dt &= xy - 4z + D\eta_3(t). \end{aligned} \tag{9.9}$$

where $\eta_i(t)$, $i = 1, 2, 3$ are independent Gaussian noise forcing terms with zero mean and unit variance. When $D = 0$, the system is clean. Figure 9.2 shows five curves,

for the cases with $D = 0, 1, 2, 3, 4$. The computations are done with 10,000 points and $m = 4$, $L = 2$. We observe the following interesting features:

- For the clean chaotic signal, $\lambda(\varepsilon)$ fluctuates slightly around a constant. As is expected, this constant is the very largest positive LE, λ_1 ,

$$\lambda(\varepsilon) = \lambda_1. \quad (9.10)$$

The small fluctuation in $\lambda(\varepsilon)$ is due to the fact that the divergence rate on the Lorenz attractor varies from one region to another.

- When there is stochastic forcing, $\lambda(\varepsilon)$ is no longer a constant when ε is small but diverges to infinity as $\varepsilon \rightarrow 0$ according the following scaling law,

$$\lambda(\varepsilon) \sim -\gamma \ln \varepsilon, \quad (9.11)$$

where γ is a coefficient controlling the speed of loss of information. This feature suggests that entropy generation is infinite when the scale ε approaches zero.

- When the noise is increased, the part of the curve with $\lambda(\varepsilon) \sim -\gamma \ln \varepsilon$ shifts to the right. In fact, little chaotic signature can be identified when D is increased beyond 3.

Note that similar results to those shown in Fig. 9.2 have been observed in other model chaotic systems, including the Mackey–Glass delay differential equation with multiple positive Lyapunov exponents [46]. To simplify our discussion of HRV data analysis, we note that in order to resolve the behavior of $\lambda(\varepsilon)$ on ever smaller scales, longer and longer time series have to be used. More precisely, for a given dataset, if the smallest resolvable scale is ε_0 , in order to resolve a smaller scale ε_0/r , where $r > 1$, a larger dataset has to be used—the larger the dimension of the attractor, the longer the time series has to be.

Next we consider $1/f^\alpha$ processes. Such type of processes is ubiquitous in science and engineering (see [47] and references therein). Two important prototypical models for such processes are fractional Brownian motion (fBm) process [48] and ON/OFF intermittency with power-law distributed ON and OFF periods [47]. For convenience, we introduce the Hurst parameter $0 < H < 1$ through a simple equation

$$\alpha = 2H + 1. \quad (9.12)$$

Depending on whether H is smaller than, equal to, or larger than $1/2$, the process is said to have anti-persistent correlation, short-range correlation, and persistent long-range correlation [47]. Note that $D = 1/H$ is the fractal dimension of such processes and Kolmogorov's 5/3 law for the energy spectrum of fully developed turbulence [49] corresponds to $H = 1/3$.

It is well known that the variance of such stochastic processes increases with t as t^{2H} . Translating this into the average distance between nearby trajectories, we immediately have

$$\varepsilon_t = \varepsilon_0 t^H. \quad (9.13)$$

Using (9.2), we then have $\lambda(\varepsilon_t) \sim H/t$. Expressing t by ε_t , we obtain

$$\lambda(\varepsilon_t) \sim H\varepsilon_t^{-1/H}. \quad (9.14)$$

Equation(9.14) can be readily verified by calculating $\lambda(\varepsilon_t)$ from such processes.

SDLE also has distinct scaling laws for random Levy processes, stochastic oscillations, and complex motions with multiple scaling laws on different scale ranges. For the details, we refer to [34, 35].

Finally, we emphasize that $\lambda_1 > 0$ (say, computed by Wolf et al.'s algorithm [21]) is not a sufficient condition for chaos. This is evident from (9.7): any non-chaotic scalings of SDLE such as (9.11) and (9.14) will yield $\lambda_1 > 0$.

9.2.1 Detecting Intermittent Chaos by SDLE

Intermittent chaos is a type of complex motion where regular (i.e., periodic) and chaotic motions alternate. Note that the stretches of the time periods for regular motions could be considerably longer than those of chaotic motions. Exactly because of this, standard methods are unable to detect chaos in such motions. This, however, is a simple task for SDLE. To illustrate the idea, we examine the logistic map

$$x_{n+1} = ax_n(1 - x_n). \quad (9.15)$$

When $a = 3.8284$, we have intermittent chaos. An example of the time series is shown in Fig. 9.3a. We observe that time intervals exhibiting chaos are very short compared with those exhibiting periodic motions. Traditional methods for computing LE, being based on global average, are unable to quantify chaos in such an intermittent situation, since the laminar phase dominates, neither can FSLE, since it requires that divergence dominates most of the time. Interestingly, the SDLE curve shown in Fig. 9.3b clearly indicates existence of chaotic motions, since the plateau region extends almost one order of magnitude.

Why can SDLE even detect chaos in such a situation? The reason is that the oscillatory part of the data only affects the scale range where $\lambda(\varepsilon) \sim 0$. It cannot affect the positive portion of $\lambda(\varepsilon)$. This means SDLE has a nice scale separation property to automatically separate the regular from chaotic motions.

9.2.2 Dealing with Nonstationarity

To facilitate our discussion of HRV data below, we now consider complicated processes generated by the following two scenarios. One is to concatenate randomly $1/f^{2H+1}$ and oscillatory components. Another is to superimpose oscillatory components on $1/f^{2H+1}$ process at randomly chosen time intervals. Either scenario generates signals that appear quite similar to that shown in Fig. 9.1a. The $\lambda(\varepsilon)$ curves for such processes are shown in Fig. 9.4, for a wide range of the H parameter.

Fig. 9.3 (a) An intermittent time series generated by the logistic map with $a = 3.8284$ and $\sigma = 0$. (b) The SDLE curve for a time series of 10,000 points, with $m = 4, L = 1$, and a shell size of $(2^{-13.5}, 2^{-13})$. A plateau is clearly visible

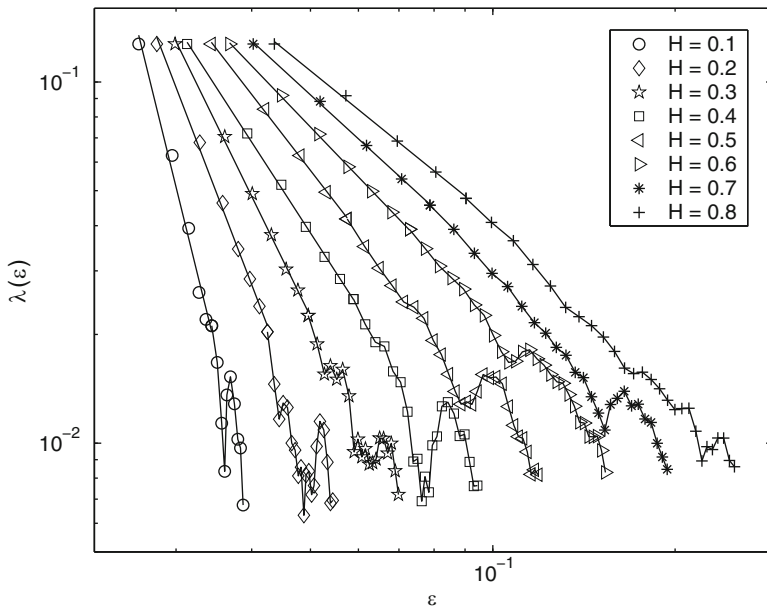
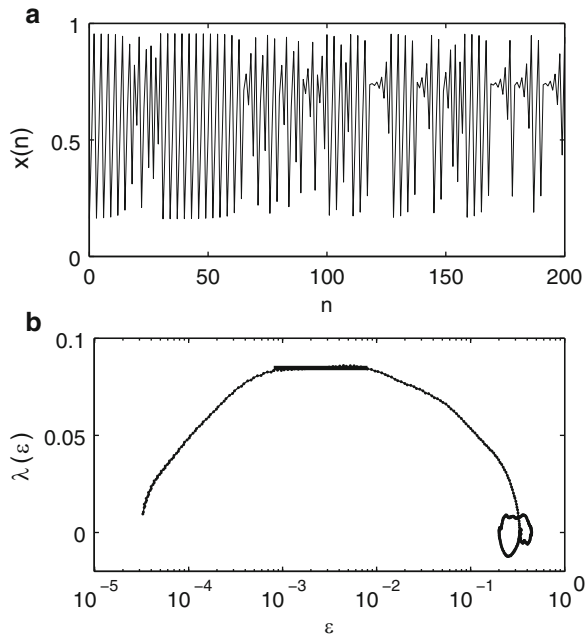


Fig. 9.4 $\lambda(\epsilon)$ vs. ϵ curves for the simulation data. Eight different H are considered. To put all the curves on one plot, the curves for different H (except the smallest one considered here) are arbitrarily shifted rightward

We observe well-defined power-law relations, consistent with (9.14), when $\lambda(\varepsilon) > 0.02$. This result clearly shows that oscillatory components in the signals can only affect the SDLE where $\lambda(\varepsilon)$ is close to 0. This is an illustration of the effect of scale isolation by the SDLE.

When we perturb chaotic data by similar procedures, will we still be able to detect chaos? The answer is yes. In fact, the intermittent chaos discussed above may be viewed as an example of such a procedure.

We are now ready to fully understand why the SDLE can deal with the types of nonstationarity discussed here. One type of nonstationarity causes shifts of the trajectory in phase space—the greater the nonstationarity, the larger the shifts. SDLE, however, cannot be affected much by shifts, especially large ones, since it is based on the coevolution of pairs of vectors within chosen small shells. The other type is related to oscillatory components. They only affect SDLE near where it is close to zero; therefore, they will not alter the distinct scaling for chaos and fractal processes.

9.3 Applications: Biological Data Analysis

To better understand the SDLE and appreciate its power, we now apply it to examine two types of physiological data, HRV and EEG.

9.3.1 EEG Analysis

EEG signals provide a wealth of information about brain dynamics, especially related to cognitive processes and pathologies of the brain such as epileptic seizures. To understand the nature of brain dynamics as well as to develop novel methods for the diagnosis of brain pathologies, a number of complexity measures have been used in the analysis of EEG data. These include the Lempel–Ziv (LZ) complexity [50], the permutation entropy [51], the LE [21], the Kolmogorov entropy [20], the correlation dimension D_2 [19, 52], and the Hurst parameter [53–55]. We compare the SDLE with these complexity measures or their close relatives.

The EEG signals analyzed here were measured intracranially by the Shands hospital at the University of Florida. Such EEG data are also called depth EEG and are considered cleaner and more free of artifacts than scalp (or surface) EEG. Altogether, we have analyzed seven patients' multiple-channel EEG data, each with a duration of a few hours, with a sampling frequency of 200 Hz. When analyzing EEG for epileptic seizure prediction/detection, it is customary to partition a long EEG signal into short windows of length W points and calculate the measure of interest for each window. The criterion for choosing W is such that the EEG signal in each window is fairly stationary, is long enough to reliably estimate the measure of interest, and is short enough to accurately resolve localized activities such as

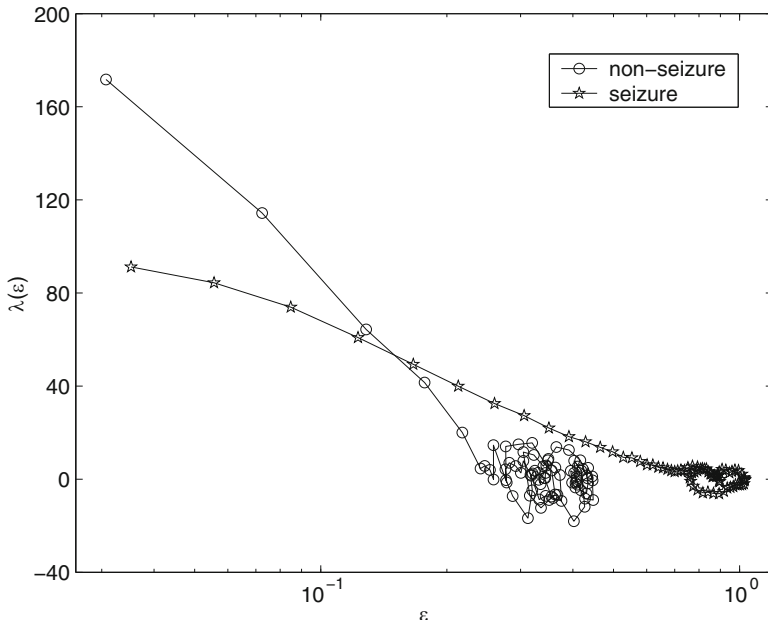


Fig. 9.5 Representative $\lambda(\epsilon)$ (per second) vs. ϵ for a seizure and non-seizure EEG segment

seizures. Since seizure activities usually last about 1–2 min, in practice, one often chooses W to be about 10 s. When applying methods from random fractal theory such as detrended fluctuation analysis (DFA) [53], it is most convenient when the length of a sequence is a power of 2. Therefore, we have chosen $W = 2,048$ when calculating various measures. We have found, however, that the variations of these measures with time are largely independent of the window size W . The relations among the measures studied here are the same for all the seven patients’ EEG data, so we illustrate the results based on only one patient’s EEG signals.

We have examined the variation of $\lambda(\epsilon)$ with ϵ for each segment of the EEG data. Two representative examples for seizure and non-seizure segments are shown in Fig. 9.5. We observe that on a specific scale ϵ^* , the two curves cross. Loosely, we may term any $\epsilon < \epsilon^*$ as small scale, while any $\epsilon > \epsilon^*$ as large scale. Therefore, on small scales, $\lambda(\epsilon)$ is smaller for seizure than for non-seizure EEG, while on large scales, the opposite is true. The variations of $\lambda_{\text{small}-\epsilon}$ and $\lambda_{\text{large}-\epsilon}$ with time for this patient’s data, where small $-\epsilon$ and large $-\epsilon$ stand for (more or less arbitrarily) chosen fixed small and large scales, are shown in Fig. 9.6a, b, respectively. We observe two interesting features: (1) the pattern of variation of $\lambda_{\text{small}-\epsilon}(t)$ is reciprocal of that of $\lambda_{\text{large}-\epsilon}(t)$. This result can be expected from Fig. 9.5. (2) The variations in $\lambda_{\text{small}-\epsilon}(t)$ and $\lambda_{\text{large}-\epsilon}(t)$ clearly indicate the two seizure events. Therefore, either $\lambda_{\text{small}-\epsilon}(t)$ or $\lambda_{\text{large}-\epsilon}(t)$ can be used to detect epileptic seizures accurately.

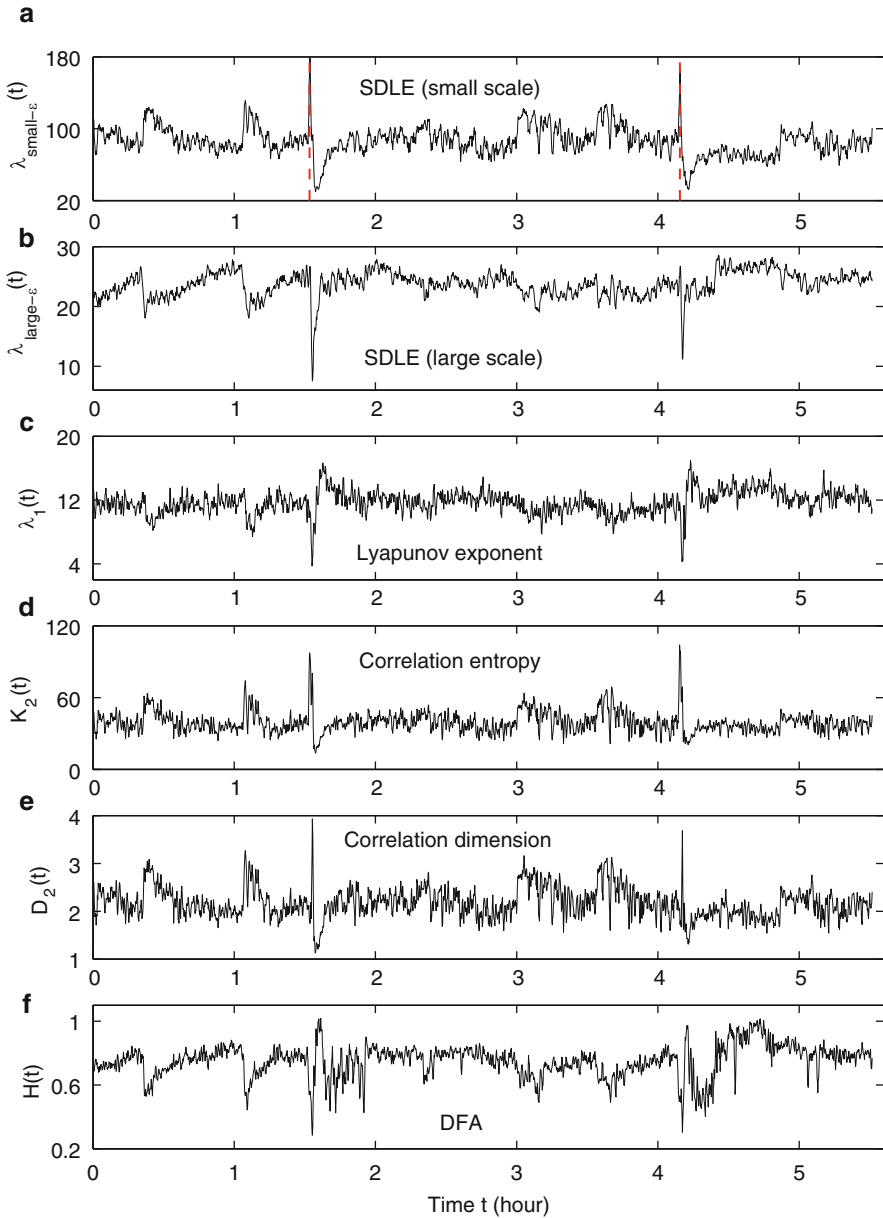


Fig. 9.6 The variation of (a) $\lambda_{\text{small-}\varepsilon}$, (b) $\lambda_{\text{large-}\varepsilon}$, (c) the LE, (d) the K_2 entropy, (e) the D_2 , and (f) the Hurst parameter with time for EEG signals of a patient. The vertical dashed lines in (a) indicate seizure occurrence times determined by medical experts

We now compare the SDLE with three commonly used measures from chaos theory, the largest positive LE [21], the correlation entropy [20], the correlation dimension [19], and one measure from random fractal theory, the Hurst parameter. We discuss the three measures from chaos theory first.

The LE is a dynamic quantity. It characterizes the exponential growth of an infinitesimal line segment, $\varepsilon_t \sim \varepsilon_0 e^{\lambda_1 t}$, $\varepsilon_0 \rightarrow 0$. It is often computed by the algorithm of Wolf et al. [21], by monitoring the exponential divergence between a reference and a perturbed trajectory. For truly chaotic signals, $1/\lambda_1$ gives the prediction time scale of the dynamics. Also, it is well known that the sum of all the positive Lyapunov exponents in a chaotic system equals the Kolmogorov–Sinai (KS) entropy. The KS entropy characterizes the rate of creation of information in a system. It is zero, positive, and infinite for regular, chaotic, and random motions, respectively. It is difficult to compute, however. Therefore, one usually computes the correlation entropy K_2 , which is a tight lower bound of the KS entropy. Similarly, the box-counting dimension, which is a geometrical quantity characterizing the minimal number of variables that are needed to fully describe the dynamics of a motion, is difficult to compute, and one often calculates the correlation dimension D_2 instead. Again, D_2 is a tight lower bound of the box-counting dimension. Both K_2 and D_2 can be readily computed from the correlation integral through the relation [19, 20].

$$C(m, \varepsilon) \sim \varepsilon^{D_2} e^{-mL\tau K_2} \quad (9.16)$$

where m and L are the embedding dimension and the delay time, τ is the sampling time, and $C(m, \varepsilon)$ is the correlation integral defined by

$$C(m, \varepsilon) = \frac{1}{N^2} \sum_{i,j=1}^N \theta(\varepsilon - \|V_i - V_j\|), \quad (9.17)$$

where $\theta(y)$ is the Heaviside step function taking values 1 or 0 depending on whether $y \geq 0$ or not, V_i and V_j are reconstructed vectors, N is the number of points in the time series, and ε is a prescribed small distance. Equation (9.16) means that in a plot of $\ln C(m, \varepsilon)$ vs. $\ln \varepsilon$ with m as a parameter, for truly low-dimensional chaos, one observes a series of parallel straight lines, with the slope being D_2 and the spacing between the lines estimating K_2 (where lines for larger m lie below those for smaller m). From these descriptions, one would expect that $\lambda_1(t)$ and $K_2(t)$ are similar, while $D_2(t)$ has little to do with either $\lambda_1(t)$ or $K_2(t)$. Surprisingly, from Fig. 9.6c–e, we observe that this is not the case: $\lambda_1(t)$ is similar to $D_2(t)$ but reciprocal of $K_2(t)$. In a moment, we shall explain how these puzzling relations may be understood based on $\lambda_{\text{small}-\varepsilon}(t)$ and $\lambda_{\text{large}-\varepsilon}(t)$.

Next we consider the calculation of the Hurst parameter H . As pointed out earlier, H characterizes the long-term correlations in a time series. There are many different ways to estimate H . We have chosen DFA [53], since it is more reliable [47] and has been used to study EEG [54, 55].

DFA works as follows: First, divide a given EEG data of length N , which is treated as a random walk process, into $\lfloor N/l \rfloor$ nonoverlapping segments (where $\lfloor x \rfloor$ denotes the largest integer that is not greater than x), each containing l points; then define the local trend in each segment to be the ordinate of a linear least-squares fit of the time series in that segment; finally compute the “detrended walk,” denoted by $x_l(n)$, as the difference between the original “walk” $x(n)$ and the local trend. One then examines the following scaling behavior:

$$F_d(l) = \left\langle \sum_{i=1}^l x_l(i)^2 \right\rangle \sim l^{2H} \quad (9.18)$$

where the angle brackets denote ensemble averages of all the segments. From our EEG data, we have found that the power-law fractal scaling breaks down around $l \approx 6$. This is caused by distinct time scales defined by the α rhythm [54] or the dendritic time constants [55]. Figure 9.6f shows $H(t)$ for our EEG data. We observe that the pattern of $H(t)$ is very similar to that of $\lambda_1(t)$ but reciprocal to $K_2(t)$ and $D_2(t)$. Such relations cannot be readily understood intuitively, since the foundations for chaos theory and random fractal theory are entirely different.

Let us now resolve all of the curious relations observed between $\lambda_1(t)$, $K_2(t)$, $D_2(t)$, and $H(t)$.

- Generally, entropy measures the randomness of a dataset. This pertains to small scale. Therefore, $K_2(t)$ should be similar to $\lambda_{\text{small}-\varepsilon}(t)$. This is indeed the case. We should point out that we have also calculated other entropy-related measures, such as the LZ complexity [50], which is closely related to the Shannon entropy and permutation entropy [51], and observed similar variations. Therefore, we can conclude that the variation of the entropy is represented by $\lambda_{\text{small}-\varepsilon}(t)$, regardless of how entropy is defined.
- To understand why $\lambda_1(t)$ calculated by the algorithm of Wolf et al. [21] corresponds to $\lambda_{\text{large}-\varepsilon}(t)$, we note that the algorithm of Wolf et al. [21] involves a scale parameter that whenever the divergence between a reference and a perturbed trajectory exceeds this chosen scale, a renormalization procedure is performed. When the algorithm of Wolf et al. [21] is applied to a time series with only a few thousand points, in order to obtain a well-defined LE, a fairly large-scale parameter has to be chosen. This is the reason that the LE and $\lambda_{\text{large}-\varepsilon}$ are similar. In fact, the scale we have chosen to calculate $\lambda_1(t)$ is even larger than that for calculating $\lambda_{\text{large}-\varepsilon}(t)$. This is the reason that the value of $\lambda_1(t)$ shown in Fig. 9.6c is smaller than that of $\lambda_{\text{large}-\varepsilon}(t)$ shown in Fig. 9.6b.
- It is easy to see that if one fits the $\lambda(\varepsilon)$ curves shown in Fig. 9.5 by a straight line, then the variation of the slope with time should be similar to $\lambda_{\text{small}-\varepsilon}(t)$ but reciprocal of $\lambda_{\text{large}-\varepsilon}(t)$. Such a pattern will be preserved even if one takes the logarithm of $\lambda(\varepsilon)$ first and then does the fitting. Such a discussion makes it clear that even if EEG is not ideally of the $1/f^{2H+1}$ type, qualitatively, the relation

$\lambda(\varepsilon) \sim \varepsilon^{-1/H}$ holds. This in turn implies $D_2 \sim 1/H$. With these arguments, it is clear that the seemingly puzzling relations among the measures considered here can be readily understood by the $\lambda(\varepsilon)$ curves. Most importantly, we have established that commonly used complexity measures can be related to the values of the SDLE at specific scales.

As we have pointed out, around the characteristic scale $\bar{\varepsilon}$, $\lambda(\varepsilon)$ is always close to 0. The pattern of $\lambda(\varepsilon)$ around $\bar{\varepsilon}$ is governed by the structured components in the data, such as the α , γ , β , and δ waves. From Fig. 9.5, we observe that the patterns for seizure and non-seizure EEG segments are very different. Such information is certainly helpful in predicting/detecting seizures. Since the numerous measures considered here are already very effective for this purpose, we will not pursue this issue further here. The issue becomes more important when distinguishing healthy subjects from patients with heart disease using the HRV data, as we will soon show.

9.3.2 HRV Analysis

HRV is an important dynamical variable of the cardiovascular function. Its most salient feature is the spontaneous fluctuation, even when the environmental parameters are maintained constant and no perturbing influences can be identified. Since the observation that HRV is related to various cardiovascular disorders [56], a number of methods have been proposed to analyze HRV data. They include methods based on simple statistics from time and frequency domain analyses (see [57] and references therein), as well as those derived from chaos theory and random fractal theory [10, 14, 15, 58–61]. We shall now show that the SDLE can readily characterize the hidden differences in the HRV under healthy and diseased conditions and shed much new light on the dynamics of the cardiovascular system.

We examine two types of HRV data, one for healthy subjects and another for subjects with the congestive heart failure (CHF), a life-threatening disease. The data were downloaded from the PhysioNet [8]. There are 18 healthy subjects and 15 subjects with CHF. Part of these datasets were analyzed by random fractal theory. In particular, 12 of the 15 CHF datasets were analyzed by wavelet-based multifractal analysis [13], for the purpose of distinguishing healthy subjects from CHF patients. For ease of comparison, we take the first 3×10^4 points of both groups of HRV data for analysis. In Fig. 9.7a, b, we have shown two typical $\lambda(\varepsilon)$ vs. ε curves, one for a healthy subject and another for a patient with CHF. We observe that for the healthy subject, $\lambda(\varepsilon)$ linearly decreases with $\ln \varepsilon$ before λ reaches around 0, or, before ε settles around the characteristic scale, $\bar{\varepsilon}$. Recall that this is a characteristic of noisy dynamics (Fig. 9.2). For the CHF case plotted in Fig. 9.7b, we observe that the $\lambda(\varepsilon)$ is oscillatory, with its value always close to 0, and hence, the only scale resolvable is around $\bar{\varepsilon}$. Since the length of the time series used in our analysis for the healthy and the CHF subjects is the same, the inability of resolving the $\lambda(\varepsilon)$ behavior on scales

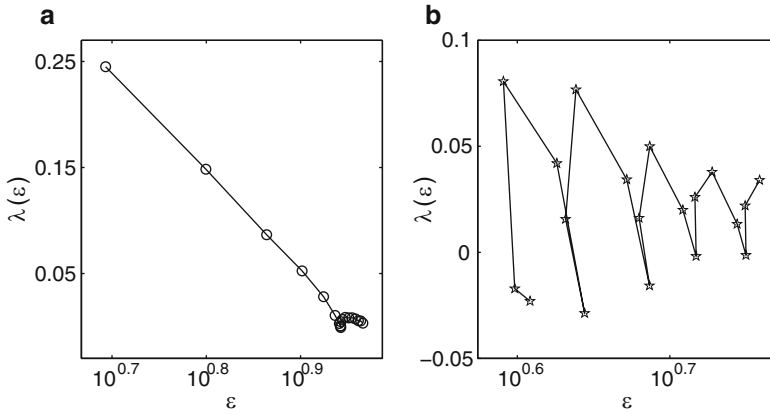


Fig. 9.7 $\lambda(\epsilon)$ (per beat) vs. ϵ (in semi-log scale) for HRV data of (a) a healthy subject and (b) a subject with CHF

much smaller than $\bar{\epsilon}$ for patients with CHF strongly suggests that the dimension of the dynamics of the cardiovascular system for CHF patients is considerably higher than that for healthy subjects.

We now discuss how to distinguish between healthy subjects and patients with CHF from HRV analysis. We have devised two simple measures, or features. One is to characterize how well the linear relation between $\lambda(\epsilon)$ and $\ln \epsilon$ can be defined. We have quantified this by calculating the error between a fitted straight line and the actual $\lambda(\epsilon)$ vs. $\ln \epsilon$ plots of Fig. 9.7a, b. The second feature is to characterize how well the characteristic scale $\bar{\epsilon}$ is defined. This is quantified by the ratio between two scale ranges, one is from the second to the sixth point of the $\lambda(\epsilon)$ curves and another is from the seventh to the 11th point of the $\lambda(\epsilon)$ curves. Now, each subject's data can be represented as a point in the feature plane, as shown in Fig. 9.8. We observe that for healthy subjects, feature 1 is generally very small but feature 2 is large, indicating that the dynamics of the cardiovascular system is like a nonlinear system with stochasticity, with resolvable small-scale behaviors and well-defined characteristic scale $\bar{\epsilon}$. The opposite is true for the patients with CHF: feature 1 is large but feature 2 is small, indicating that not only small-scale behaviors of the $\lambda(\epsilon)$ curves cannot be resolved, but also that the characteristic scale $\bar{\epsilon}$ is not well defined. Very interestingly, these two simple features separate completely the normal subjects from patients with CHF. In fact, each feature alone can almost perfectly separate the two groups of subjects studied here.

It is interesting to note that for the purpose of distinguishing normal HRV from CHF HRV, the features derived from SDLE are much more effective than other metrics including the Hurst parameter, the sample entropy, and multiscale entropy. For the details of the comparisons, we refer to [39].

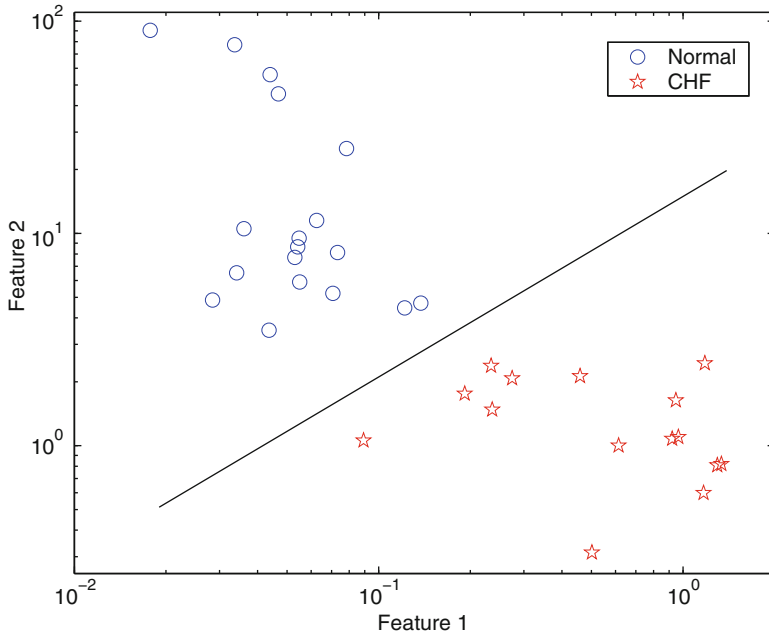


Fig. 9.8 Feature plane separating normal subjects from subjects with CHF

9.4 Concluding Remarks

In this chapter, we have discussed a multiscale complexity measure, the SDLE. We have shown that it can readily characterize low-dimensional chaos and random $1/f^\alpha$ processes, can readily detect intermittent chaos, can conveniently deal with nonstationarity, and can accurately detect epileptic seizures from EEG and distinguish healthy subjects from patients with CHF from HRV. More importantly, we have established that commonly used complexity measures can be related to the value of the SDLE at specific scales, and that the pattern of the SDLE around the characteristic scale $\bar{\epsilon}$ contains a lot of useful information on the structured components of the data that may greatly help detect significant patterns. Because of the ubiquity of chaos-like motions and $1/f^\alpha$ -type processes and the complexity of HRV and EEG data, our analyses strongly suggest that the SDLE is potentially important for clinical practice and provides a comprehensive characterization of complex data arising from a wide range of fields in science and engineering.

Our analyses have a number of important implications:

- To comprehensively characterize the complexity of complicated data such as HRV or EEG data, a wide range of scales has to be considered, since the complexity may be different on different scales. For this purpose, the entire $\lambda(\epsilon)$ curve, where ϵ is such that $\lambda(\epsilon)$ is positive, provides a good solution. This point

is particularly important when one wishes to compare the complexity between two signals—the complexity for one signal may be higher on some scales, but lower on other scales. The situation shown in Fig. 9.5 may be considered one of the simplest.

- For detecting important events such as epileptic seizures, $\lambda_{\text{small}-\varepsilon}$ and $\lambda_{\text{large}-\varepsilon}$ appear to provide better defined features than other commonly used complexity measures. This may be due to the fact that $\lambda_{\text{small}-\varepsilon}$ and $\lambda_{\text{large}-\varepsilon}$ are evaluated at fixed scales, while other measures are not. In other words, scale mixing may blur the features for events being detected, such as seizures.
- In recent years, there has been much effort in searching for cardiac chaos [14–18,58]. Due to the inability of unambiguously distinguishing deterministic chaos from noise by calculating the largest positive LE and the correlation dimension, it is still unclear whether the control mechanism of cardiovascular system is truly chaotic or not. Our analysis here highly suggests that if cardiac chaos does exist, it is more likely to be identified in healthy subjects than in pathological groups. This is because the dimension of the dynamics of the cardiovascular system appears to be lower for healthy than for pathological subjects. Intuitively, such an implication makes sense, because a healthy cardiovascular system is a tightly coupled system with coherent functions, while components in a malfunctioning cardiovascular system are somewhat loosely coupled and function incoherently.

As example applications, we have focused on the analyses of HRV and EEG data here. It is evident that SDLE will be useful for other kinds of complex data analyses, including financial time series and various kinds of physiological data. While much of the past as well as current research has been focused on determining whether some experimental data are chaotic or not, the scaling laws of SDLE suggest that it is often feasible to obtain the defining parameters of the data under study. That is, if the data is chaotic, then one should find out what kind of chaos it is; and if it is random, one can aim to find out what kind of random process that is, including its correlation structure. While in principle, SDLE is able to do so without pre-processing of the data under study, suitable detrending and denoising may help. A particularly simple and versatile procedure is the smooth adaptive filter developed by the authors, which has been successfully applied to recover chaos in heavy noise environment [62–64].

Acknowledgments This work is supported in part by NSF grants CMMI-1031958 and 0826119.

References

1. Fairley P (2004) The unruly power grid. *IEEE Spectrum* 41:22–27
2. Bassingthwaighte JB, Liebovitch LS, West BJ (1994) *Fractal physiology*. Oxford University Press, New York
3. Leland WE, Taquu MS, Willinger W, Wilson DV (1994) On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Trans Networking* 2:1–15

4. Gao JB, Rubin I (2001) Multiplicative multifractal modeling of long-range-dependent network traffic. *Int J Commun Syst* 14:783–801
5. Gao JB, Rubin I (2001) Multifractal modeling of counting processes of long-range-dependent network traffic. *Comput Commun* 24:1400–1410
6. Mandelbrot BB (1997) *Fractals and scaling in finance*. Springer, New York
7. Gao JB, Hu J, Tung WW, Zheng Y (2011) Multiscale analysis of economic time series by scale-dependent Lyapunov exponent. *Quantitative Finance* DOI:10.1080/14697688.2011.580774
8. Mit-bih normal sinus rhythm database & bidmc congestive heart failure database available at <http://www.physionet.org/physiobank/database/#ecg>.
9. Akselrod S, Gordon D, Ubel FA, Shannon DC, Barger MA, Cohen RJ (1981) Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science* 213:220–222
10. Kobayashi M, Musha T (1982) 1/f fluctuation of heart beat period. *IEEE Trans Biomed Eng* 29:456–457
11. Peng CK, Mietus J, Hausdorff JM, Havlin S, Stanley HE, Goldberger AL (1993) Long-range anticorrelations and non-gaussian behavior of the heartbeat. *Phys Rev Lett* 70:1343–1346
12. Ashkenazy Y, Ivanov PC, Havlin S, Peng CK, Goldberger AL, Stanley HE (2001) Magnitude and sign correlations in heartbeat fluctuations. *Phys Rev Lett* 86:1900–1903
13. Ivanov PC, Amaral L, Goldberger AL, Havlin S, Rosenblum MG, Struzik ZR, Stanley HE (1999) Multifractality in human heartbeat dynamics. *Nature* 399:461–465
14. Kaplan DT, Goldberger AL (1991) Chaos in cardiology. *J Cardiovasc Electrophysiol* 2:342–354
15. Goldberger AL, West BJ (1987) Applications of nonlinear dynamics to clinical cardiology. *Ann New York Acad Sci* 504:155–212
16. Garfinkel A, Spano ML, Ditto WL, Weiss JN (1992) Controlling cardiac chaos. *Science* 257:1230–1235
17. Fortrat JO, Yamamoto Y, Hughson RL (1997) Respiratory influences on non-linear dynamics of heart rate variability in humans. *Biol Cybernet* 77:1–10
18. Kaneko K, Tsuda I (2000) *Complex systems: chaos and beyond*. Springer, Berlin
19. Grassberger P, Procaccia I (1983) Characterization of strange attractors. *Phys Rev Lett* 50:346–349
20. Grassberger P, Procaccia I (1983) Estimation of the kolmogorov entropy from a chaotic signal. *Phys Rev A* 28:2591–2593
21. Wolf A, Swift JB, Swinney HL, Vastano JA (1985) Determining lyapunov exponents from a time series. *Physica D* 16:285–317
22. Sugihara G, May RM (1990) Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* 344:734–741
23. Kaplan DT, Glass L (1992) Direct test for determinism in a time-series. *Phys Rev Lett* 68:427–430
24. Gao JB, Zheng ZM (1994) Direct dynamical test for deterministic chaos and optimal embedding of a chaotic time series. *Phys Rev E* 49:3807–3814
25. Gao JB, Zheng ZM (1994) Direct dynamical test for deterministic chaos. *Europhys Lett* 25:485–490
26. Pei X, Moss F (1996) Characterization of low-dimensional dynamics in the crayfish caudal photoreceptor. *Nature* 379:618–621
27. Poon CS, Barahona M (2001) Titration of chaos with added noise. *Proc Natl Acad Sci USA* 98:7107–7112
28. Gaspard P, Briggs ME, Francis MK, Sengers JV, Gammons RW, Dorfman JR, Calabrese RV (1998) Experimental evidence for microscopic chaos. *Nature* 394:865–868
29. Dettmann CP, Cohen EGD (2000) Microscopic chaos and diffusion. *J Stat Phys* 101:775–817
30. Hu J, Tung WW, Gao JB, Cao YH (2005) Reliability of the 0-1 test for chaos. *Phys Rev E* 72:056207
31. Osborne AR, Provenzale A (1989) Finite correlation dimension for stochastic-systems with power-law spectra. *Physica D* 35:357–381

32. Provenzale A, Osborne AR (1991) Convergence of the k2 entropy for random noises with power law spectra. *Physica D* 47:361–372
33. Aurell E, Boffetta G, Crisanti A, Paladin G, Vulpiani A (1997) Predictability in the large: an extension of the concept of Lyapunov exponent. *Physica A* 30:1–26
34. Gao JB, Hu J, Tung WW, Cao YH (2006) Distinguishing chaos from noise by scale-dependent Lyapunov exponent. *Phys Rev E* 74:066204
35. Gao JB, Cao YH, Tung WW, Hu J (2007) Multiscale analysis of complex time series—integration of chaos and random fractal theory, and beyond. Wiley, New York
36. Gao JB, Hu J, Mao X, Tung WW (2012) Detecting low-dimensional chaos by the “noise titration” technique: possible problems and remedies. *Chaos, Solitons, and Fractals* 45:213–223
37. Gao JB, Tung WW, Hu J (2009) Quantifying dynamical predictability: the pseudo-ensemble approach (in honor of professor Andrew Majda’s 60th birthday). *Chinese Annals Math Ser B* 30:569–588
38. Gao JB, Hu J, Tung WW (2011) Complexity measures of brain wave dynamics. *Cogn Neurodyn* 5:171–182
39. Hu J, Gao JB, Tung WW, Cao YH (2010) Multiscale analysis of heart rate variability: a comparison of different complexity measures. *Annals Biomed Eng* 38:854–864
40. Hu J, Gao JB, Tung WW (2009) Characterizing heart rate variability by scale-dependent Lyapunov exponent. *Chaos (special issue on Controversial Topics in Nonlinear Science: Is the Normal Heart Rate Chaotic? it is one of the most downloaded papers in that issue)* 19:028506
41. Ryan DA, Sarson GR (2008) The geodynamo as a low-dimensional deterministic system at the edge of chaos. *Europhys Lett* 83:49001
42. Packard NH, Crutchfield JP, Farmer JD, Shaw RS (1980) Geometry from time-series. *Phys Rev Lett* 45:712–716
43. Takens F (1981) Detecting strange attractors in turbulence. In: Rand DA, Young LS (eds.) *Dynamical systems and turbulence, lecture notes in mathematics, vol 898*. Springer, Berlin, p 366
44. Sauer T, Yorke JA, Casdagli M (1991) Embedology. *J Stat Phys* 65:579–616
45. Gao JB, Zheng ZM (1993) Local exponential divergence plot and optimal embedding of a chaotic time series. *Phys Lett A* 181:153–158
46. Mackey MC, Glass L (1977) Oscillation and chaos in physiological control-systems. *Science* 197:287–288
47. Gao JB, Hu J, Tung WW, Cao YH, Sarshar N, Roychowdhury VP (2006) Assessment of long range correlation in time series: How to avoid pitfalls. *Phys Rev E* 73:016117
48. Mandelbrot BB (1982) *The fractal geometry of nature*. Freeman, San Francisco
49. Frisch U (1995) *Turbulence—the legacy of A.N. Kolmogorov*. Cambridge University Press
50. Lempel A, Ziv J (1976) On the complexity of finite sequences. *IEEE Trans Inform Theor* 22:75–81
51. Cao YH, Tung WW, Gao JB, Protopopescu VA, Hively LM (2004) Detecting dynamical changes in time series using the permutation entropy. *Phys Rev E* 70:046217
52. Martinerie J, Adam C, Quyen MLV, Baulac M, Clemenceau S, Renault B, Varela FJ (1998) Epileptic seizures can be anticipated by non-linear analysis. *Nat Med* 4:1173–1176
53. Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL (1994) On the mosaic organization of dna sequences. *Phys Rev E* 49:1685–1689
54. Hwa RC, Ferree TC (2002) Scaling properties of fluctuations in the human electroencephalogram. *Phys Rev E* 66:021901
55. Robinson PA (2003) Interpretation of scaling properties of electroencephalographic fluctuations via spectral analysis and underlying physiology. *Phys Rev E* 67:032902
56. Hon EH, Lee ST (1965) Electronic evaluations of the fetal heart rate patterns preceding fetal death: further observations. *Am J Obstet Gynecol* 87:814–826
57. Task Force of the European Society of Cardiology & the North American Society of Pacing & Electrophysiology (1996) Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Circulation* 93:1043–1065

58. Babyloyantz A, Destexhe A (1988) Is the normal heart a periodic oscillator? *Biol Cybern* 58:203–211
59. Bigger JT, Steinman RC, Rolnitzky LM, Fleiss JL, Albrecht P, Cohen RJ (1996) Power law behavior of rr-interval variability in healthy middle-aged persons, patients with recent acute myocardial infarction, and patients with heart transplants. *Circulation* 93:2142–2151
60. Ho KKL, Moody GB, Peng CK, Mietus JE, Larson MG, Levy D, Goldberger AL (1997) Predicting survival in heart failure cases and controls using fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics. *Circulation* 96:842–848
61. Pincus SM, Viscarello RR (1992) Approximate entropy: a regularity statistic for fetal heart rate analysis. *Obst Gynecol* 79:249–255
62. Hu J, Gao JB, Wang XS (2009) Multifractal analysis of sunspot time series: the effects of the 11-year cycle and fourier truncation. *J Stat Mech*. DOI: 10.1088/1742-5468/2009/02/P02066
63. Gao JB, Sultan H, Hu J, Tung WW (2010) Denoising nonlinear time series by adaptive filtering and wavelet shrinkage: a comparison. *IEEE Signal Process Lett* 17:237–240
64. Tung WW, Gao JB, Hu J, Yang L (2011) Recovering chaotic signals in heavy noise environments. *Phys Rev E* 83:046210