From phase space to frequency domain: A time-frequency analysis for chaotic time series

Junfeng Sun,* Yi Zhao, Tomomichi Nakamura, and Michael Small
Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China
(Received 23 August 2006; revised manuscript received 20 November 2006; published 31 July 2007)

Time-frequency analysis is performed for chaotic flow with a power spectrum estimator based on the phase-space neighborhood. The relation between the reference phase point and its nearest neighbors is demonstrated. The nearest neighbors, representing the state recurrences in the phase space reconstructed by time delay embedding, actually cover data segments with similar wave forms and thus possess redundant information, but recur with no obvious temporal regularity. To utilize this redundant recurrence information, a neighborhood-based spectrum estimator is devised. Then time-frequency analysis with this estimator is performed for the Lorenz time series, the Rössler time series, experimental laser data, and colored noise. Features revealed by the spectrogram can be used to distinguish noisy chaotic flow from colored noise.

DOI: 10.1103/PhysRevE.76.016220 PACS number(s): 05.45.−a, 05.10.−a, 02.70.Hm

I. INTRODUCTION

In order to obtain the inherent properties of a system from the observed time series, a variety of methods have been proposed and are widely applied, such as surrogate tests [1,2], wavelets [3], Fourier transforms [4,5], and approaches based on time delay embedding [6]. Among these methods, approaches based on time delay embedding may be the most popular framework for analyzing chaotic time series. Based on Taken's embedding theory [7,8], some measures such as Lyapunov exponents [9] and correlation dimensions [10] have been proposed to characterize the global features of dynamical systems. However, few studies of the local time pattern of chaotic time series have been reported; nevertheless, this is important for some purposes, such as to reveal the degree of chaoticity of a sequence.

Spectra analysis provides an alternative framework for chaotic time series analysis [4,5,11–13]. With methods based on Fourier transformation, the relation between the spectra and the topology as the corresponding dynamical system bifurcates to chaos has been studied. Spectrum bands of some period-doubling bifurcation sequences (e.g., the Rössler time series) merge as the dynamic system bifurcates to chaos [4,13]. Another typical spectrum of chaotic data (e.g., the Lorenz time series) is broadband and falls off via an exponential law. This spectrum falloff pattern has been utilized to distinguish chaotic sequences from colored noise with power-law spectra [11]. However, other researchers have argued that a chaotic sequence cannot be well distinguished from either colored noise [14] or quasiperiodic motion (with singular power spectra) by its finite-time power spectra [5]. This is especially true when the chaotic data are contaminated by observational noise.

For a chaotic signal with complicated evolution (e.g., for the Lorenz time series, the time interval between peaks and the amplitude of the sequence both vary with no obvious regularity), the simple frequency domain representation may obscure information related to timing. Spectrum analysis usually only adopts the spectral amplitude, while neglecting the phase information. Consequently, confusion will occur between any two signals with the same spectral amplitudes. A time-frequency joint analysis is therefore desirable to better unveil these features [15]. However, few studies of the time-frequency analysis for chaotic sequence have been reported. Chandre *et al.* [3] performed a time-frequency analysis of Hamiltonian systems based on a ridge (instantaneous frequency curve) extraction from a wavelet decomposition of a single-trajectory coordinate. The ridge pattern can reveal the phase-space structures (resonance transitions, trappings, etc.) and give a characterization of weak or strong chaos.

State recurrence is one important feature of chaotic systems. In the phase space reconstructed by time delay embedding, the state recurrences of a reference phase point turn out to be its nearest neighbors, which can provide redundant information but recur with no temporal regularity as we will demonstrate later. Wavelet analysis, as with other conventional time-frequency analysis methods (e.g., periodogram [16]), utilizes only one segment of consecutive data and neglects temporally isolated state recurrences beside this data segment. So a time-frequency analysis which can utilizes all state recurrences is desirable. The present paper focuses on (i) demonstrating that nearest neighbors can provide redundant information for chaotic signal analysis and processing, (ii) proposing a spectrum estimator which can utilize all the neighbors, and (iii) performing a time-frequency analysis to (noisy) chaotic flow with the proposed spectrum estimator and extracting some features that can be used to distinguish the (noisy) chaotic data from colored noise.

We will demonstrate that the nearest neighbors cover segments of data with similar wave forms to that of the corresponding reference phase point. However, the recurrence time of each neighbor appears to be irregular; nonetheless, the mean recurrence time obeys a scaling law [17]. In order to utilize these state recurrences, the nearest neighbors must be grouped according to their Euclidean distance to the reference point in phase space. Other techniques, such as local projection noise reduction [18,19] and nonlinear prediction [20] that utilize nearest neighbors, frequently demonstrate positive results. Analogously, aiming to use the state recurrence of a chaotic system, we propose a neighborhood-based spectrum estimator (NSE) to estimate the corresponding power spectra of the reference phase point. The NSE first performs eigenvalue decomposition to the covariance matrix of the neighbors and then estimates the power spectra of the

*sun.junfeng@polyu.edu.hk

reference point by applying the Blackman-Tukey (BT) estimator [21]. Thus, the NSE utilizes the long-term state recurrence of chaotic systems and relates time delay embedding information (deterministic dynamics) to the frequency domain.

The organization of this paper is as follows. In Sec. II, the relationship between the reference point and its nearest neighbors is demonstrated, and the principle of neighborhood-based spectrum estimation (NSE) is presented. In Sec. III, time-frequency analysis with NSE is performed for the Lorenz time series, the Rössler time series, and experimental laser data. The spectrogram, the ridge of main frequencies, and "hidden" frequency are used to characterize the features of these time series. In Sec. IV, NSE is applied to colored noise. It is shown that colored noise can be distinguished from (noisy) chaotic flow based on their respective main ridge patterns. Finally, a conclusion and discussion are given in Sec. V.

II. PRINCIPLE OF THE METHOD

A. Nearest neighbors

Let $\{z_n\}_{n=1}^{L}$ denote a chaotic time series with L samples. The phase points can be reconstructed by time delay embedding—i.e., $\{\mathbf{z}_n\}_{n=1}^{L-(d-1)\tau}$:

$$\mathbf{z}_n = [z_n, z_{n+\tau}, z_{n+2\tau}, \ldots, z_{n+(d-1)\tau}]^{\mathrm{T}},$$

where d is embedding dimension, τ is time delay, and $(\cdot)^{\mathrm{T}}$ denotes the transpose of a real matrix. The near neighborhood of the reference point $\mathbf{z}_n$ is defined as

$$\mathbf{N}_n \triangleq \{\mathbf{z}_k : \|\mathbf{z}_k - \mathbf{z}_n\| < \varepsilon, 1 \leqslant k \leqslant L-(d-1)\tau\}$$

and arranged as $\mathbf{N}_n = \{\mathbf{z}_{k_1}, \mathbf{z}_{k_2}, \ldots, \mathbf{z}_{k_N}\}$, $k_1 < k_2 < \cdots < k_N$, where $N = |\mathbf{N}_n|$ is the number of neighbors and ε is the neighborhood radius (note that $\mathbf{z}_n \in \mathbf{N}_n$). Furthermore, the recurrence time of $\mathbf{z}_n$ can be simply defined as $T_n(i) = k_{i+1} - k_i, i = 1, \ldots, N-1$ [17].

Considering a chaotic time series generated from the Lorenz system [22],

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= (r - z)x - y, \\ \dot{z} &= xy - bz, \end{aligned} \tag{1}$$

where $(\sigma, r, b) = (10, 28, 8/3)$. Note that all the Lorenz time series used in this paper are 10 000 points sampled from the x component with time interval 0.04 and these fixed parameters, unless stated otherwise.

Figure 1 demonstrates the relationship between the reference point $\mathbf{z}_{2963}$ (randomly selected) and its first ten nearest neighbors with subscript $k = 192, 2659, 3485, 4387, 4388, 5376, 5415, 6763, 6764, 7235$. The reference point $\mathbf{z}_n$ covers a segment of time series $[z_n, z_{n+1}, z_{n+2}, \ldots, z_{n+(d-1)\tau}]^{\mathrm{T}}$ with the length of embedding window $L_w = (d-1)\tau + 1$. For clarity, let $\mathbf{s}_n$ denote this associated segment of the time series. If $\tau = 1$, $\mathbf{s}_n$ is the same as $\mathbf{z}_n$. It can be observed that the corresponding wave forms of the neighbors are similar to each other, but the recurrence time seems irregular. From the viewpoint of signal processing, these similar wave-form segments contain much redundant information relative to the reference one. There are some neighbors that are adjacent in time—for example, $k = 4387$ and 4388. The adjacent neighbors that lie on the same recurrence trajectory provide only one new sample; primarily they serve to increase the weight of the corresponding state recurrence within the neighborhood.

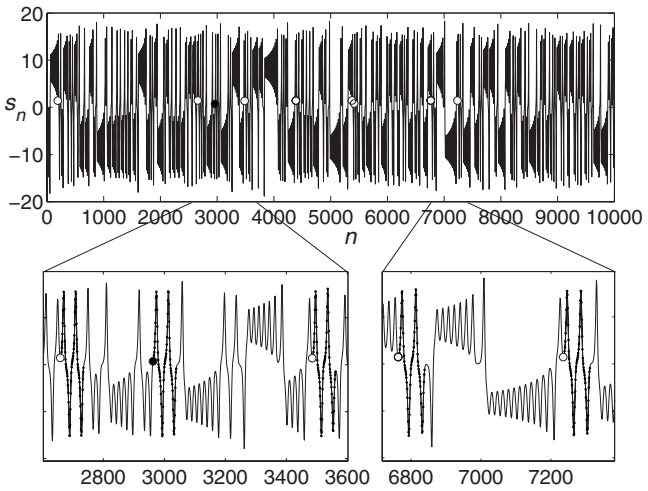

FIG. 1. The Lorenz time series and the first ten nearest neighbors of reference point $\mathbf{z}_{2963}$. ●, the sample z_{2963}; ○, the samples z_k corresponding to the first ten nearest neighbors. The bottom panels are enlargements of short segments. Each segment marked with small dots corresponds to one neighbor in phase space.

As aforementioned, several recurrence points, which lie on one same recurrence trajectory, may be included in the near neighborhood. To investigate the recurrence time of state recurrence trajectory, we define a second type of neighborhood by selecting only one point from each recurrence trajectory, and the selected one is the nearest one to the reference point among the neighbors on that recurrence trajectory. A histogram of the recurrence time of the second type of neighbors in all neighborhoods is shown in Fig. 2. It indicates that (i) the recurrence time varies over a large range

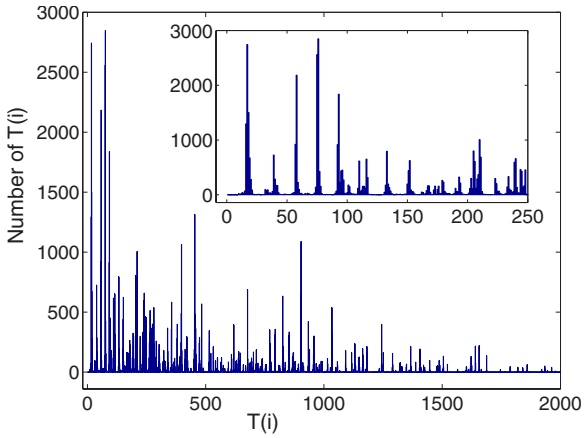

FIG. 2. (Color online) Histogram of the recurrence time of the first ten nearest neighbors of the second type. The small panel is a local enlargement of the main figure.

and (ii) the state recurrence seems to appear more frequently after some periodic constant time that corresponds to the peaks of the histogram.

Some applications of recurrence time have been reported in recent years. A similar second type of neighborhood was defined by excluding so-called "sojourn points," and the recurrence time statistics of chaotic time series was revealed to obey a scaling law [17].[1] Further, the recurrence time statistics has been applied to detect nonstationarity and state transitions [23]. On the other hand, conventional linear techniques (e.g., classical Fourier transform) neglect some scattered state recurrences and just utilizing one segment of consecutive data. Consequently, these techniques usually obtain poor results in analyzing chaotic data, while some methods (e.g., local projection noise reduction), specifically designed for chaotic data, utilize the neighbors and thus achieve better results.

B. Neighborhood-based spectrum estimation

For the neighborhood $\mathbf{N}_n$, we define an $L_w \times N$ neighborhood matrix as $\mathbf{D}_n=[\mathbf{x}_{k_1},\mathbf{x}_{k_2},\ldots,\mathbf{x}_{k_N}]$, with notation $\mathbf{x}_{k_i}=\mathbf{s}_{k_i}-\bar{\mathbf{s}}_n$, where $\bar{\mathbf{s}}_n=\langle \mathbf{s}_{k_i}\rangle$ is the center. First, an eigenvalue decomposition to the covariance matrix—i.e., $\mathbf{C}_n=\frac{1}{N}\mathbf{D}_n\mathbf{D}_n^{\mathrm{T}}$—of the neighborhood $\mathbf{N}_n$ is performed,

$$\mathbf{C}_n\mathbf{u}_i-\lambda_i\mathbf{u}_i=0, \tag{2}$$

where λ_i is the ith eigenvalue, and $\mathbf{u}_i=[u_i(1),u_i(2),\ldots,u_i(L_w)]^{\mathrm{T}}$ is its associated eigenvector. Then with the discrete-time Fourier transform of eigenvector $\mathbf{u}_i$,

$$V_i(\omega)=\sum_{p=1}^{L_w} u_i(p)e^{-jwp}, \tag{3}$$

the NSE can be expressed as

$$P_{NSE}(\omega)=\frac{1}{L_w}\sum_{i=1}^{L_w}\lambda_i|V_i(\omega)|^2. \tag{4}$$

The NSE is derived from the BT spectrum estimator [21]. The difference is that the BT estimator utilizes the covariance matrix generated from only one segment of consecutive data, while NSE uses the covariance matrix estimated from the data segments covered by the temporally scattered nearest neighbors. Thus, NSE can capture the long-time state recurrence of chaotic data. If the neighborhood contains only the reference point, the NSE reduces to the BT estimator.

The eigenvalues are arranged in descending order—that is, $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_{L_w}$. If the chaotic time series is contaminated by observational noise, the local phase space can be divided into two orthogonal subspaces: (i) the noise subspace spanned by $[\mathbf{u}_{M+1},\ldots,\mathbf{u}_{L_w}]$, assumed to contain components from the noise process only, and (ii) the signal subspace

[1]The recurrence time statistics in [17] is related to the radius of neighborhood, while the histogram in Fig. 2 is related to the first ten nearest neighbors of the second type. Due to difference in statistics, Fig. 2 does not show the scaling law revealed in [17].

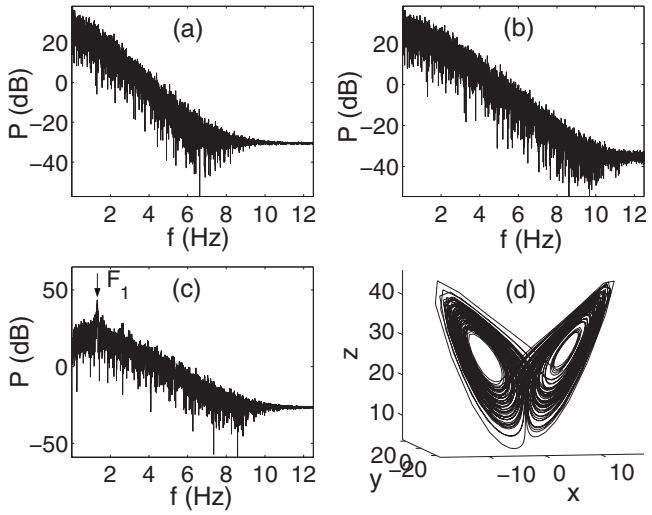

FIG. 3. Power spectra of the Lorenz system estimated by a periodogram. The y-axis label P denotes the power spectra of each sequence. (a), (b), and (c) are the power spectra of the time series measured from x, y, and z components, respectively; (d) the strange attractor of the Lorenz system.

constructed with $[\mathbf{u}_1,\ldots,\mathbf{u}_M]$ containing the "clean signal" and a certain, small, amount of noise [19], where M is the minimum embedding dimension of the dynamic system [24]. Just adopting those components in the signal subspace, a principal component version of the NSE is

$$P_{PC\text{-}NSE}(\omega)=\frac{1}{L_w}\sum_{i=1}^{M}\lambda_i|V_i(\omega)|^2. \tag{5}$$

Obviously, the principal component version of the NSE has the ability to suppress the observational noise for spectrum estimation of contaminated chaotic data.

Furthermore, for each reference point, we define the *main frequency* ω_m as

$$P(\omega_m)=\max P(\omega_l),\quad \omega_m\in\{\omega_l\}, \tag{6}$$

where ω_l is the frequencies with local maximum power amplitude—i.e.,

$$\left.\frac{dP(\omega)}{d\omega}\right|_{\omega=\omega_l}=0,\qquad \left.\frac{d^2P(\omega)}{d\omega^2}\right|_{\omega=\omega_l}<0. \tag{7}$$

Then the *main frequency* will form a *main ridge* as the reference point moves along the phase trajectory. We observe in the following sections that this main ridge shows different characteristic patterns for different types of data.

III. TIME-FREQUENCY ANALYSIS OF CHAOTIC TIME SERIES

In this section, time-frequency analysis with NSE is presented for the Lorenz time series, the Rössler time series, and experimental laser data.

The Lorenz system is a typical chaotic system with two scrolls. Figure 3 shows the power specta of the Lorenz system estimated by a periodogram [16], a classical method

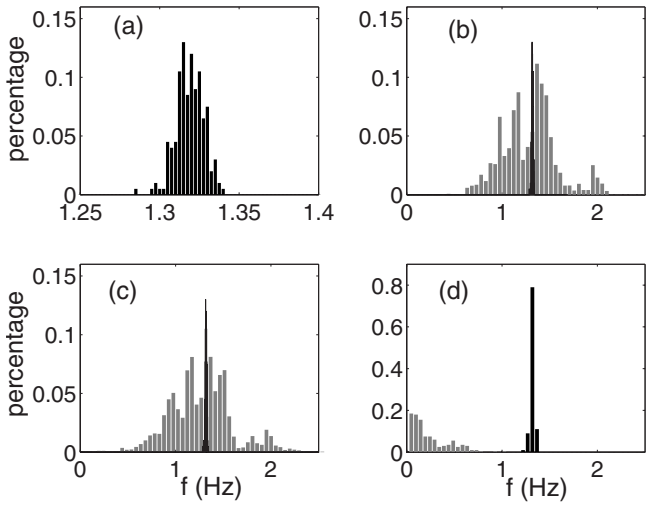

FIG. 4. (a) The histogram of the spectra peaks [as that in Fig. 3(c)] of 200 Lorenz sequences (10 000 points each sequence) measured from the z component with different initial conditions; (b) the gray bars denote the histogram of the main frequencies in Fig. 5(a), and the black bars indicate the histogram in (a); (c) the gray bars denote the histogram of the main frequencies in Fig. 5(b), and the black bars indicate the histogram in (a); (d) the histogram of the spectra peaks of 200 Lorenz sequences (500 points each sequence) estimated by a periodogram; the gray and the black bars, respectively, denote the histograms of the spectra peaks of sequences measured from x and z components simultaneously.

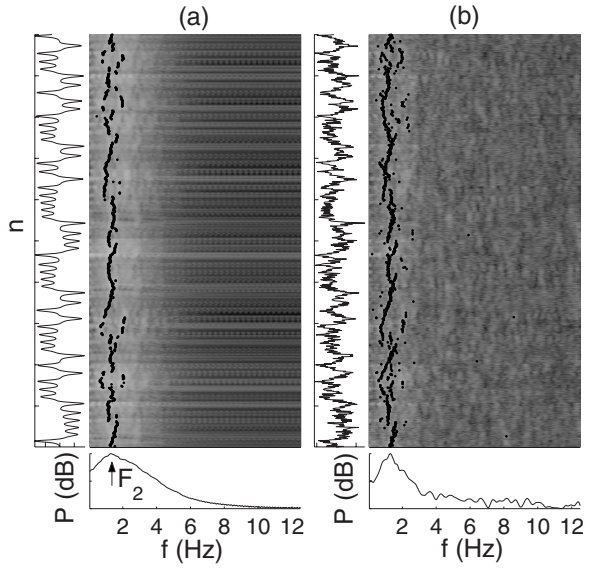

FIG. 5. Spectrogram of the Lorenz time series, $r=28$. (a) The clean Lorenz time series and (b) the noisy Lorenz time series with 5 dB additive white noise. For each subfigure, the left panel is the time series, the top-right panel is the corresponding spectrogram estimated by NSE, and the bottom panel is the average of the spectrogram over time. The black points are the main frequencies. This pattern is followed in all following spectrogram figures, unless otherwise stated.

based on Fourier transformation. We can observe that the power spectra of the x and y components are broadband and similar to each other, while the power spectra of the z component have a peak. This spectra peak, which is indicated by $\downarrow F_1$ in Fig. 3(c) and named the *hidden frequency* in Refs. [25,26], can reveal the frequency related to the principal oscillation of the Lorenz system. We note that this frequency is not a particular case of this sequence. The spectrum peak universally exists with small deviation (1.305–1.330 Hz with 95% confidence), as Fig. 4(a) indicates. Though this oscillation exists in the x and y components simultaneously as the dynamics evolves, the periodogram spectra of x and y fail to reveal it. The time interval between peaks and the amplitude of the x and y sequences both vary with the phase state switching between the two scrolls with no obvious regularity and thus can be considered as frequency modulation and amplitude modulation, respectively. The periodogram spectra cannot capture this complicated modulation and therefore fail to reveal the principal oscillation.

In contrast, time-frequency analysis of the same Lorenz time series used in Fig. 3(a) with NSE can reveal the principal oscillation. We overembed the time series with time delay $\tau=4$ (determined by the first minimum of mutual information [27]) and embedding dimension $d=20$, and use the first 20 nearest neighbors in NSE analysis. A 1000-sample segment of the spectrogram for each case is illustrated in Fig. 5. We can observe that (i) the spectra are broadband and the energy is primarily distributed in the low-frequency region and (ii) the main ridge is formed by many short disjointed curves (even for the Lorenz time series contaminated with 5 dB white noise), which vary slowly around a frequency related to the principal oscillation. The bottom panel is the average of the spectrogram over time, which can be considered as the energy distribution versus frequency. The frequency corresponding to the maximum peak of this curve, indicated by $\uparrow F_2$, is approximately equal to the hidden frequency indicated by $\downarrow F_1$ in Fig. 3(c). Histograms of the main frequencies in Fig. 5 are shown in Figs. 4(b) and 4(c), respectively. The hidden frequency is located at the center of the main frequencies. This implies that the main frequencies contain the information of the principal oscillation of chaotic system. Similar results can also be obtained with data measured from the y component of the Lorenz system.

Similar wave forms covered by the neighbors can enhance their common structure—i.e., the principal oscillation—while they may simultaneously "average" out the substructures and noise. Thus, even for the noisy Lorenz time series with 5 dB white noise, the principal oscillation can be extracted. In time-frequency analysis with NSE, we will mainly focus on the pattern of the main ridge. Time-frequency analysis with a short-time periodogram [15] has also been performed of sequences simultaneously measured from the x and z components. As Fig. 4(d) indicates, the spectrum peaks of sequences measured from the z component are located at the center region of the histogram in Figs. 4(b) and 4(c), while the spectrum peaks of sequences measured from the x component tend to be close to zero frequency. Note that a different window length has been adopted for the short-time periodogram, but the pattern of the main frequencies is similar to that shown in Fig. 4(d).

The Lorenz system with different values of the parameter r has been widely studied [28]. We further apply NSE to (noisy) Lorenz time series generated with different values of

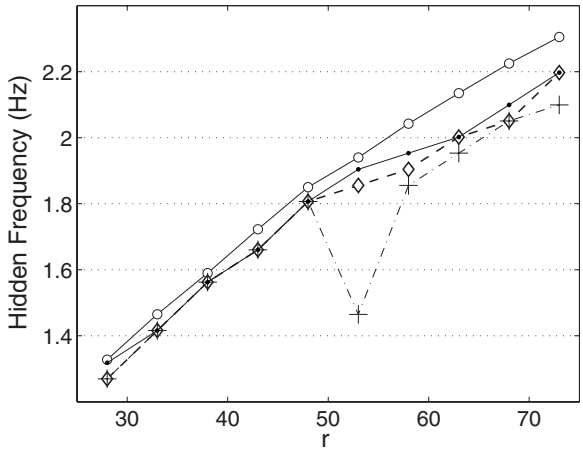

FIG. 6. Hidden frequency of the Lorenz time series. ○, by a periodogram with a clean Lorenz time series of the z component; ●, by NSE with a clean series of the x component; ◇, by NSE with a noisy series of the x component (10 dB additive white noise); +, by NSE with a noisy series of the x component (5 dB additive white noise).

the parameter r while fixing the parameters $(\sigma,b)=(10,8/3)$. When the Lorenz system is chaotic (e.g., $r=33$ and 38), the main ridge is formed with unconnected short curves, which is similar to the case of $r=28$. While when the Lorenz system is nonchaotic (e.g., $r=18$ and 148), the system is almost periodic and the main frequencies are approximately constant. Figure 6 shows that the hidden frequencies can be detected by NSE even when the Lorenz series is contaminated by 5 dB additive white noise (only one mismatch for the case of noisy data with 5 dB noise). We set $d=20$ for the cases of $r=28,33,38,43,48,53$ and $d=10$ for the cases of $r=58,63,68,73$. This is because as the hidden frequency increases with r, the period of oscillation becomes small and thus the reference phase point will cover more cycles, which leads to fewer well-matched neighbors. If the number of appropriate neighbors is too small, the reliability of NSE will be reduced.

The length of the embedding window, $L_w=(d-1)\tau+1$, is a trade-off between the reliability and the frequency resolution of the estimated spectra. To get a better reliability, L_w should be set relatively shorter to ensure more appropriate neighbors. On the other hand, to obtain a spectrum with higher resolution, L_w should be set relatively longer. With sampling time interval $\Delta t=0.04$ and $L_w=77$ ($d=20$, $\tau=4$), the physical resolution of the spectra estimated by Fourier transformation is $\frac{1}{\Delta t L_w}\approx 0.32$ Hz. This resolution cannot provide an accurate detection of the hidden frequency (about 1.32 Hz when $r=28$). Padding zeros at the end of the data is a common strategy adopted in the implementation of discrete Fourier transformation, so as to obtain a higher computational resolution—i.e., smaller frequency interval between the bins calculated by fast Fourier transformation (FFT). In this paper, we pad $(512-L_w)$ zeros to the end of $\mathbf{u}_i$ and Eq. (3) is implemented by 512-point FFT. With this strategy, the computational resolution is $\frac{1}{0.04\times 512}\approx 0.05$ Hz, which can be considered as an appropriate result of interpolation. But padding zeros to individual realizations does not increase the physical resolution of the estimated spectra, and the location of the spectra peak may depart from the hiddern frequency. Further, we refer the reader to Refs. [21,29] for more discussions of frequency resolution.

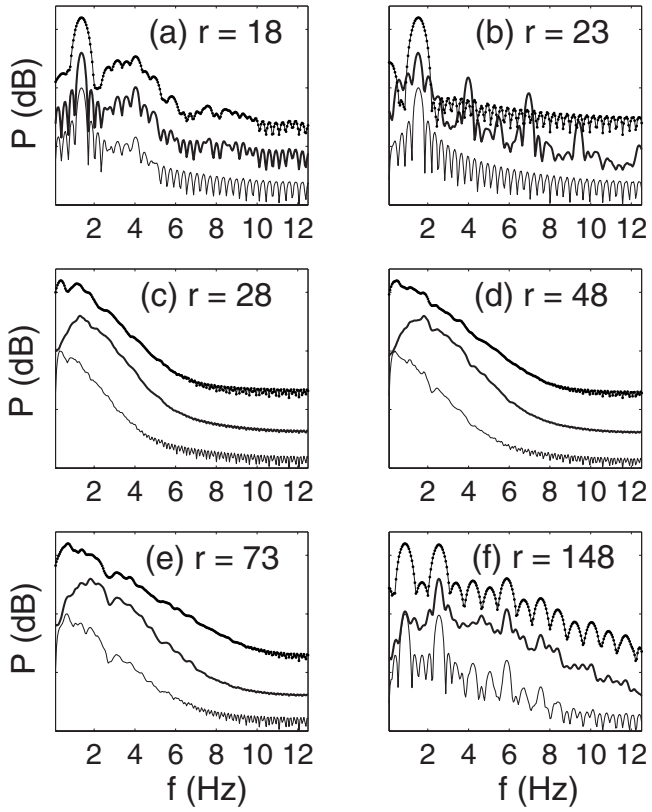

FIG. 7. Energy distribution versus frequency of the Lorenz time series measured from the x component. The maximum Lyapunov exponents corresponding to (a)–(f) are −0.23, −0.05, 0.90, 1.24, 1.53, and 0.00, respectively. In each panel, the three curves correspond to the normalized average of the spectrogram by the short-time periodogram, NSE, and BT estimator from top to bottom. Each curve is offset vertically for clarity. The scale in the vertical axis is therefore arbitrary.

Figure 7 shows the energy distribution versus frequency for the Lorenz series. When the system has nonpositive maximum Lyapunov exponent [Figs. 7(a), 7(b), and 7(f)], the time series seems pseudoperiodic. NSE can detect a fundamental frequency as well as the short-time periodogram and BT estimator, though some harmonics are different, while for the Lorenz system with positive maximum Lyapunov exponent [Figs. 7(c)–7(e)], only NSE can reveal the frequency (i.e., the hidden frequency) related to the principal oscillation, which has been verified in Fig. 6.

The Rössler system is another typical chaotic system with periodic motion superimposed on chaotic behavior [30],

$$\dot{x}=-(y+z),$$
$$\dot{y}=x+0.2y,$$
$$\dot{z}=0.2+xz-Cz. \tag{8}$$

When $C=4.6$, the Rössler system is chaotic [4]. Its power spectra contain almost periodic δ-function peaks and broad

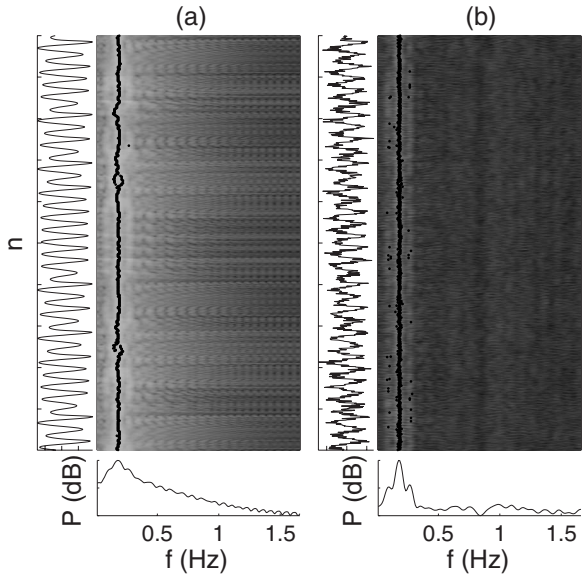

FIG. 8. Spectrogram of the Rössler time series by NSE. (a) The clean time series measured from the x component of the Rössler system and (b) the noisy Rössler time series with 5 dB additive white noise. The sample time interval is 0.2, $\tau=7$, and $d=12$.

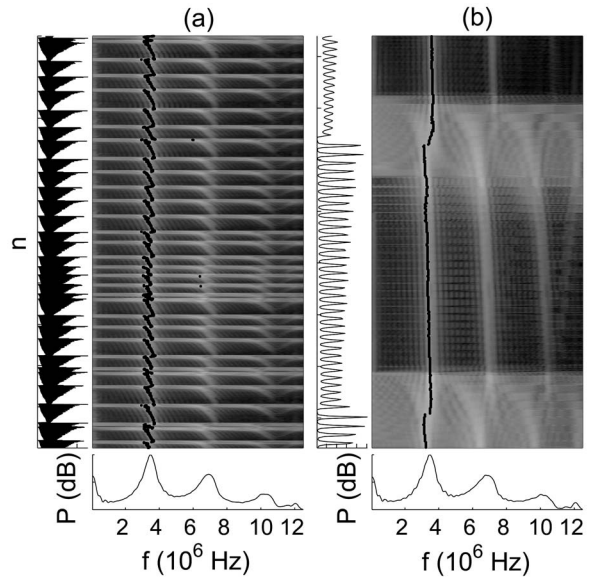

FIG. 9. Spectrogram of laser time series by NSE (10 000 samples). (a) 10 000—point data and their spectrogram and (b) an enlargement of (a) from point 2921 to point 3320.

background components. The sharp peaks are due to periodic motion, and the broad components are the result of amplitude modulation.

Figure 8(a) shows that the Rössler time series measured form the x component is something like a pseudoperiodic one with amplitude fluctuation and the main frequency varies very little with time. The peak of the curve of the energy distribution corresponds to the fundamental frequency of the Rössler system. For the noisy Rössler time series [Fig. 8(b)], the main ridge is obvious and "stationary" with only a few scattered points.

Further we apply NSE to an experimental laser time series [31], which was obtained from the Santa Fe time series competition data [32]. To better match the neighbors, we form the neighborhood with $\tau=1$, and $d=80$. From Fig. 9, we can see that the main frequency decreases little during the pulse boosting, due to amplitude modulation, and then jumps back as the oscillation collapses.

For comparison, time-frequency analyses with short-time periodogram and BT estimator are performed of the (noisy) Rössler time series, and laser data, similar main ridge patterns to that generated by the NSE, are obtained, respectively.

IV. DISTINGUISHING CHAOTIC TIME SERIES FROM COLORED NOISE

It is difficult to distinguish (noisy) chaotic data from colored noise by their spectra falloff patterns [14]. Chaotic flow has scattered state recurrences, while colored noise does not possess this deterministic feature. Here, we will demonstrate that time-frequency analysis with NSE can reveal this difference and thus can be an alternative method to distinguish them.

We take a pink noise and a surrogate sequence as examples. The pink noise (10 000 points) is generated by a special case of a first-order autoregressive process [AR(1)] [33], $X_{n+1}=\beta X_n+(1-\beta)\epsilon_n$, where $\beta=0.69$ and $\epsilon_n \sim N(0,1)$ is a Gaussian process [more discussion about AR(1) can be found in Ref. [34]]. As Fig. 10 indicates, the spectra of the clean Lorenz data have a long exponential-law scaling region (marked by A). As the Lorenz data are contaminated by observational noise, the exponential-law region (as that indicated by B and C for the cases with 10 dB and 5 dB white noise, respectively) becomes less obvious and difficult to be distinguished from that of pink noise (marked by E), while the time-frequency analysis with NSE is sensitive to this difference. As Fig. 11(a) indicates, the main frequency of pink noise varies along time with no regularity, while the main ridge pattern of the noisy chaotic data with 5 dB white noise [Fig. 5(b)] exhibits more long-term temporal structure.

Surrogate tests provide a powerful technique to detect determinism in time series [1]. Here, the surrogate data are generated by shuffling the phase of the original noisy Lorenz data [the one used in Fig. 5(b)] [2]. The power spectra of the surrogate data (marked by D in Fig. 10) are similar to those of the original data (marked by C). However, due to the phase shuffling, the surrogate data do not possess the deterministic features of the original noisy Lorenz data, and thus their main ridge patterns are clearly distinct [Fig. 11(b) vs Fig. 5(b)]. As we discussed in Sec. I, time-frequency joint analysis can reveal some information that is obscured by just a single finite-time frequency representation.

The histograms of the main frequencies related to Fig. 11 are illustrated in Fig. 12. Comparing with Fig. 4, we can observed that the main frequencies of both the pink noise and the surrogate data are mainly distributed in a region near zero frequency, while the main frequencies of (noisy) chaotic data are located in a region relatively far from zero frequency.

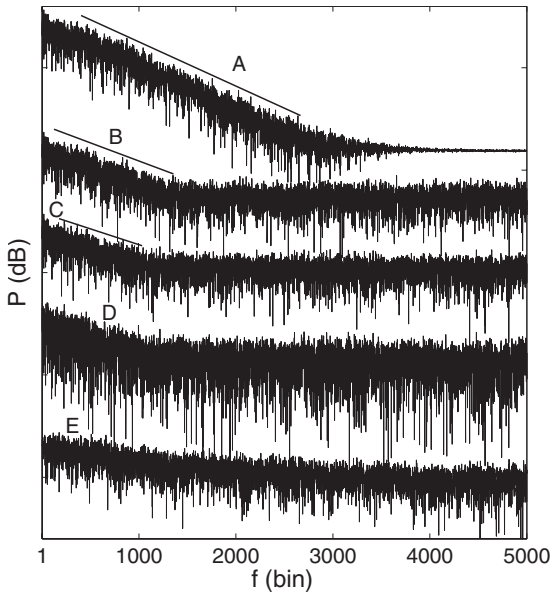

FIG. 10. Power spectra estimated by periodogram. From top down, the spectra indicated by *A*, *B*, *C*, *D*, and *E* correspond to the clean Lorenz time series used in Fig. 5(a), the noisy Lorenz time series with 10 dB white noise, the noisy Lorenz time series (with 5 dB white noise) used in Fig. 5(b), a phase shuffled surrogate data of the noisy Lorenz time series used in Fig. 5(b), and the pink noise generated by AR(1). Each time series has 10 000 points. The three lines indicate the exponential-law scaling regions of the corresponding spectra. Here, the frequency bins calculated by discrete Fourier transformation are not scaled to the real frequency with units of Hz.

In summary, for the chaotic Lorenz time series, the main ridge has many short unconnected curves, which vary around the hidden frequency. We believe that this main ridge pattern is a characteristic of chaotic time series. For the chaotic Rössler time series, the wave form is pseudoperiodic, yielding a main ridge that varies smoothly and slowly. In the time domain, the wave-form variation of the Lorenz time series (switching between two scrolls) seems more "complex" than that of the Rössler time series (evolving around one focus). Therefore, the main ridge of the Rössler time series is more "regular." For noisy chaotic flow, the principal oscillation can be extracted with the nearest neighbors, and thus the main ridge reserves some characteristics of the corresponding clean data, while for the pink noise and surrogate data, there is no deterministic feature, and thus the main ridge is irregular, which is distinct from that of (noisy) chaotic flow. The difference in main ridge pattern can be used to distinguish them. NSE is designed to investigate (noisy) chaotic flow in the viewpoint of time-frequency analysis. Various methods from other viewpoints, such as the 0-1 test [35,36] and method based on scale-dependent Lyapunov exponent [37], have been developed to investigate whether a nonlinear time series is deterministically chaotic or stochastic. We refer readers interested in this topic to Refs. [35–37] for more discussions.

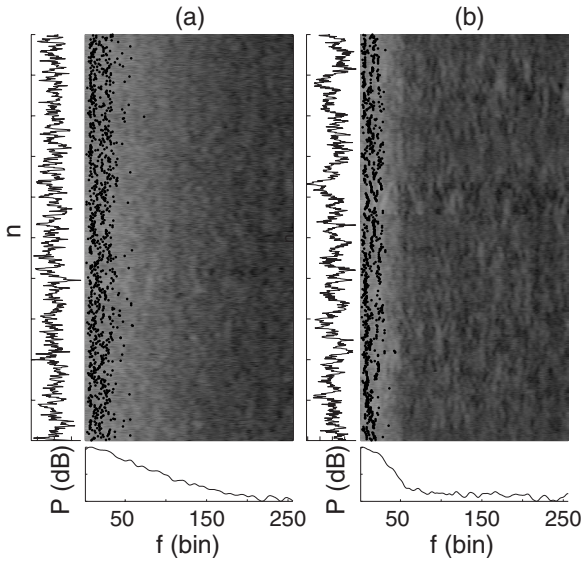

FIG. 11. (a) Spectrogram of pink noise by NSE and (b) spectrogram of the surrogate data used in Fig. 10. The frequency bins calculated by FFT are not scaled to the real frequency with units of Hz.

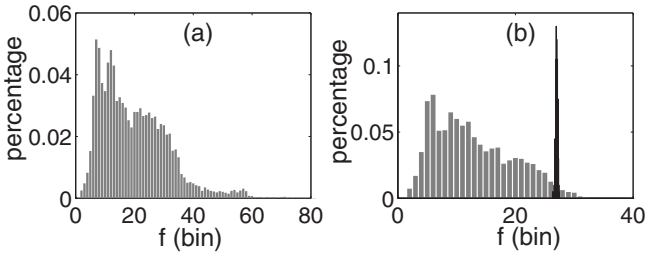

FIG. 12. (a) Histogram of the main frequencies in Fig. 11(a). (b) The gray bars denote the histogram of the main frequencies in Fig. 11(b), and for comparison, the histogram in Fig. 4(a) is plotted by black bars.

V. CONCLUSION AND DISCUSSION

We performed a time-frequency analysis for chaotic time series. First, chaotic data were overembedded, and the relation between the reference phase point and its nearest neighbors was demonstrated. Neighbors represent the state recurrences of the reference point and cover data segments with similar wave forms, but recur with no obvious temporal regularity. To apply these state recurrences, a neighborhood-based power spectrum estimator was devised for chaotic flow, bridging time delay embedding and the frequency domain. Then time-frequency analysis with NSE was performed for (noisy) Lorenz time series. We found that NSE can reveal the frequency related to the principal oscillation of the dynamical system, which is hidden in the spectra estimated by the periodogram method. Furthermore, NSE was applied to the Rössler time series and experimental laser data. We observed that the pattern of main frequencies has similar characteristics: they vary slowly around a frequency related to the principal oscillation of the system. We further applied NSE to pink noise and phase shuffled surrogate data. The results show that their main ridge patterns are distinct from that of (noisy) chaotic flow, thus providing an alternative method to distinguish colored noise from (noisy) chaotic

flow, though for some real or more chaotic systems, a distinction may not be that easy.

On the one hand, NSE can reveal some meaningful features that classical methods fail to uncover; on the other hand, NSE also may "average" out some substructures. So NSE can be adopted together with other methods to make a comprehensive understanding of the dynamical system. Further study of time-frequency analysis with NSE is needed: to verify the ability of distinguishing chaotic data from colored noise systematically and to extract some measures from the main ridge patterns. Application of NSE for noise reduction of contaminated chaotic data also appears to be promising.

ACKNOWLEDGMENT

This research was funded by Hong Kong University Grants Council, Grant Competitive Earmarked Research Grant (CERG) No. PolyU 5269/06E.

[1] M. Small and C. K. Tse, IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **50**, 63 (2003).
[2] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, Physica D **58**, 77 (1992).
[3] C. Chandre, S. Wiggins, and T. Uzer, Physica D **181**, 171 (2003).
[4] J. Crutchfield, D. Farmer, N. Packard, R. Shaw, G. Jones, and R. Donnelly, Phys. Lett. **76A**, 1 (1980).
[5] R. S. Dumont and P. Brumer, J. Chem. Phys. **88**, 1481 (1988).
[6] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. (Cambridge University Press, Cambridge, England, 2003).
[7] F. Takens, Lect. Notes Math. **898** (1981).
[8] T. Sauer, J. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).
[9] A. Wolf, J. B. Swift, H. L. Swinney, and A. Vastano, Physica D **16**, 285 (1985).
[10] J. D. Farmer, E. Ott, and J. A. Yorke, Physica D **7**, 152 (1983).
[11] D. Sigeti and W. Horsthemke, Phys. Rev. A **35**, 2276 (1987).
[12] V. Brunsden and P. Holmes, Phys. Rev. Lett. **58**, 1699 (1987).
[13] J. D. Farmer, Phys. Rev. Lett. **47**, 179 (1981).
[14] A. A. Tsonis and J. B. Blsner, Nature (London) **358**, 217 (1992).
[15] *Time Frequency Signal Analysis and Processing: A Comprehensive Reference*, edited by B. Boashash (Elsevier, Amsterdam, 2003).
[16] S. M. Kay, *Modenern Spectral Estimation* (Prentice-Hall, Englewood Cliffs, NJ, 1988).
[17] J. B. Gao, Phys. Rev. Lett. **83**, 3178 (1999).
[18] R. Cawley and G. H. Hsu, Phys. Rev. A **46**, 3057 (1992).
[19] R. Hegger, H. Kantz, and L. Matassini, IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **48**, 1454 (2001).
[20] H. D. I. Abarbanel, T. A. Carroll, L. M. Pecora, J. J. Sidorowich, and L. S. Tsimring, Phys. Rev. E **49**, 1840 (1994).
[21] M. H. Hayes, *Statistical Digital Signal Processing and Modeling* (Wiley, New York, 1996).
[22] E. Lorenz, J. Atmos. Sci. **20**, 130 (1963).
[23] J. B. Gao, Phys. Rev. E **63**, 066202 (2001).
[24] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, Phys. Rev. A **45**, 3403 (1992).
[25] G. J. Ortega, Phys. Rev. Lett. **77**, 259 (1996).
[26] J. Chen, J. Ko, J. Lih, H. Su, and R. Hsu, Phys. Lett. A **238**, 134 (1998).
[27] A. M. Fraser and H. L. Swinney, Phys. Rev. A **33**, 1134 (1986).
[28] C. Sparrow, IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **30**, 533 (1983).
[29] S. J. Orfanidis, *Introduction to Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1996).
[30] O. E. Rossler, Phys. Lett. **57A**, 397 (1976).
[31] U. Hubner, N. B. Abraham, and C. O. Weiss, Phys. Rev. A **40**, 6354 (1989).
[32] http://www-psych.stanford.edu/~andreas/Time-Series/SantaFe.html
[33] O. Miramontes and P. Rohani, Physica D **166**, 147 (2002).
[34] J. Gao, J. Hu, W.-W. Tung, Y. Cao, N. Sarshar, and V. P. Roychowdhury, Phys. Rev. E **73**, 016117 (2006).
[35] G. A. Gottwald and I. Melbourne, Proc. R. Soc. London, Ser. A **460**, 603 (2004).
[36] J. Hu, W.-W. Tung, J. Gao, and Y. Cao, Phys. Rev. E **72**, 056207 (2005).
[37] J. B. Gao, J. Hu, W.-W. Tung, and Y. H. Cao, Phys. Rev. E **74**, 066204 (2006).