

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Nonlinear Analysis of Surface EMG Signals

Min Lei and Guang Meng

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/49986>

1. Introduction

The aim of this chapter is to answer the essence of SEMG and to explore the potential use of nonlinear analysis as a tool in the clinical and biomechanical applications. The technical tools include nonlinear time series test, time series analysis based on chaos theory, multifractal analysis.

In Section 2, we discuss the two methods of nonlinear time series test: surrogate data test method and Volterra-Wiener-Korenberg (VWK) model test method. Theoretically, the two methods can detect the nonlinearity of the data indirectly. The surrogate data method is used to analyze the SEMG. The result shows that the SEMG has deterministic nonlinear components. Meanwhile, we introduce the VWK model test method and compare it with the surrogate data method. The nonlinearity of SEMG during muscle fatigue can be detected by the VWK.

In Section 3, we describe the time series analysis based on chaos theory. The chaos definition and chaotic characteristics are discussed. The embedding theory of the attractor reconstruction is investigated for the dynamical system. From the view of the fractal structure of the chaotic attractor, the correlation dimension is used to test the chaotic characteristics of the SEMG during arm movements. The Largest Lyapunov exponent is also studied. Then, we investigate the influence of measure noise, internal noise and sampling interval on the principal components of chaotic time series. The symplectic principal component analysis is given. We illustrate the feasibility of this method and give the embedding dimension of the action surface EMG signal.

In Section 4, the self-affine fractal definition and nature are described. The power spectrum and frequency relationship is used to calculate the self-affine fractal dimension of the time series, such as SEMG. Then, the multifractal dimension is given for the SEMG.

The conclusion and future research are shown in Section 5. Here, it is necessary to note that this chapter is actually the result of many years work. The new methods presented here

build on a broad and strong foundation of nonlinear time series analysis and chaotic dynamical theory.

2. Detecting nonlinearity of the surface EMG signals

In many areas of science and engineering, it is a critical issue to determine whether an observed time series is purely stochastic, or deterministic nonlinear, even chaotic. One may know about the intrinsic properties of the observed phenomenon by distinguishing between nonlinear deterministic dynamics and noisy dynamics from a time series. In this section, we review and discuss the surrogate data test method[1] and Volterra-Wiener-Korenberg (VWK) model test method[2] for identifying the nonlinearity of a time series. These methods have been successfully used to detect and characterize nonlinear dynamics from recordings in biology and medicine[2-5].

Surrogate analysis is currently an important empirical technique of testing nonlinearity for a time series. The aim is to test whether the dynamics are consistent with linearly filtered noise or a nonlinear dynamical system[1, 6]. The basic idea of the surrogate data method is to first specify some kind of linear stochastic process as a null hypothesis that mimics “linear properties” of the original data. According to the null hypothesis, surrogate data sets are generated. Then, a discriminating statistic is calculated for the original and for each of the surrogate data sets. If the statistic of the original data is significantly different from those of surrogate data sets, the null hypothesis can be rejected within a certain confidence level. It shows that the original data is from a nonlinear dynamical system. The method is demonstrated for numerical time series generated by known chaotic systems and applied to the analysis of SEMG.

VWK test method is a kind of nonlinear detection of time series based on linear and nonlinear Volterra-Wiener-Korenberg model [2, 5]. That is, it first produces the linear and nonlinear predicted data from the original time series and then compares their information criterions to detect the nonlinearity of the raw data. VWK test technique is capable of robust and highly sensitive statistical detection of deterministic dynamics, including chaotic dynamics, in experimental time series. This method is superior to other techniques when applied to short time series, either continuous or discrete, even when heavily contaminated with noise or in the presence of strong periodicity. Later, an extension of the Volterra algorithm (called the numerical titration algorithm) was given to detect and quantify chaos in noisy time series[7]. Here, the surrogate data method and VWK test approach are used to analyze the nonlinearity of surface EMG signals.

2.1. Surrogate data test method

Surrogate data method includes two parts: a null hypothesis and a test statistic. The null hypothesis is a specific process which may or may not adequately explain an origin of the data. The test statistic provides a quantitative description to demonstrate the data sources.

2.1.1. Null hypotheses and algorithms[1]

The null hypotheses usually specify some certain properties of the original data that reflect some structure characteristics of the dynamical system, such as mean and variance, and possibly also the Fourier power spectrum. Different null hypotheses describe different specific dynamical systems. In terms of the corresponding null hypothesis, the surrogate data can be generated so as to test the corresponding specific dynamical system class.

Null hypothesis 1 The observed data is produced by independent and identically distributed (IID) random variables.

For this hypothesis, the corresponding surrogate data can be generated by shuffling the time-order of the original time series so that it has the same mean, variance and amplitude distribution as the original data. But any temporal correlations of the original data are destroyed in the surrogate data. Schienkman and LeBaron[8] applied this hypothesis to analyze stock market returns. Breeden and Packard also used this hypothesis to demonstrate that a time series of quasar data which were sampled nonuniformly in time has some dynamics structure[9].

The algorithm of the null hypothesis is that one first create gaussian random numbers from 1 to N , where N is the length of the original data x . Then, the original data x is permuted by the random numbers to generate the surrogate data.

Null hypothesis 2 The observed data is produced by the Ornstein-Uhlenbeck process.

The surrogate data generated by the Ornstein-Uhlenbeck process is a sequence that has the simplest time correlation. The Ornstein-Uhlenbeck process can be given as follows.

$$x_t = a_0 + a_1 x_{t-1} + \sigma e_t \quad (1)$$

where e_t is a Gaussian random with zero mean and unit variance. The coefficients a_0 , a_1 , and σ work together to determine the mean, variance, and autocorrelation time of the time series x_t . In this case, its autocorrelation function is exponential form. Let $\lambda = -\log a_1$, $\langle \cdot \rangle$ denotes an average over time t . That is,

$$A(\tau) \equiv \frac{\langle x_t \cdot x_{t-\tau} \rangle - \langle x_t \rangle^2}{\langle x_t^2 \rangle - \langle x_t \rangle^2} = e^{-\lambda|\tau|} \quad (2)$$

In order to generate the surrogate data that is consistent with this hypothesis, the algorithm is that one first calculates the mean μ , variance γ and autocorrelation $A(1)$ (in Eq. (2)) from the original data x . Then, the coefficients in Eq. (1) can be estimated: $a_1 = A(1)$, $a_0 = \mu(1 - a_1)$, and $\sigma^2 = \gamma(1 - a_1^2)$. The Gaussian e_t can be generated by a pseudorandom number generator. Finally, the surrogate data can be produced by iterating Eq. (1).

Null hypothesis 3 The observed data is produced by the linear autocorrelated gaussian process with the mean and variance of the original time series.

The hypothesis has been usually used to test whether the original time series contains nonlinear components. It can be described by using a linear autoregressive (AR) model.

$$x_t = a_0 + \sum_{k=1}^q a_k x_{t-k} + \sigma e_t \quad (3)$$

There are the two algorithms to produce the surrogate data in accord with this hypothesis. One algorithm is to directly use Eq.3. That is, the coefficients are firstly identified by using the original data. Then, the surrogate data is generated by repeatedly iterating Eq.3. However, the performance of this algorithm is very unstable. If the values of the coefficients are mis-estimated slightly, this algorithm may lead to the iterative results which easily diverge to infinity. The alternative algorithm is that a surrogate data is generated by randomizing the phases of a Fourier transform. According to the Weiner-Khintchine theorem, the two algorithms are equivalent in essence[1, 10]. The surrogate data has the same Fourier spectrum as the original data. Meanwhile, the algorithm based on the Fourier transform is stabler in the numerical calculation than the first algorithm. The following is the steps of this algorithm.

Let an observed data as $x(n)$. The Fourier transform of $x(n)$ is computed as follows[3]:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-2\pi i n k / N} \quad (4)$$

The Fourier transform has a complex amplitude at each frequency. One can randomize the phases of the Fourier transform by multiplying $e^{i\varphi}$,

$$X'(k) = X(k) \cdot e^{i\varphi} \quad (5)$$

where φ is independently chosen for each frequency from the $[0, 2\pi]$. In order to the inverse Fourier transform to be real (no imaginary components), the phases must satisfy the antisymmetric condition $\varphi(k) = -\varphi(N-k)$. Meanwhile, $\varphi(0) = 0$, $\varphi(N/2) = 0$ (when N is even), so that

$$X'(k) = \overline{X'(N-k)} \quad (6)$$

This point can be easily proved[11].

Proof:

According to the nature of DFT of a real time series $x(n)$, if $x(n) \in R$, then

$$\phi(k) = -\phi(-k) \quad (7)$$

where $\phi(k)$ is the phase angle of $X(k)$. $k = 0, 1, \dots, N-1$. N is the period of Fourier Transform. Then, for $k = 0, k = N/2$ (N is even), there are

$$\phi(0) = -\phi(0), \quad (8)$$

$$\phi(N/2) = -\phi(-N/2) = -\phi(N/2) \quad (9)$$

In order to ensure that the inverse Fourier transform results are real values, there must be

$$\phi(0) = 0, \quad \phi(N/2) = 0.$$

In practical, if the data length N is odd, $\phi(f_1)=0$, $\phi(f_i)=-\phi(f_k)$, $i=2\sim(N+1)/2$, $k=N\sim(N+1)/2+1$; If N is even, $\phi(f_1)=0$, $\phi(f_{N/2+1})=0$, $\phi(f_i)=-\phi(f_k)$, $i=2\sim N/2$, $k=N\sim N/2+2$. Thus, the surrogate data $x'(n)$ given by the inverse Fourier transform is a sequence of real numbers.

$$x'(n) = \frac{1}{N} \sum_{k=0}^{N-1} X'(k) \cdot e^{2\pi i n k / N} \quad (10)$$

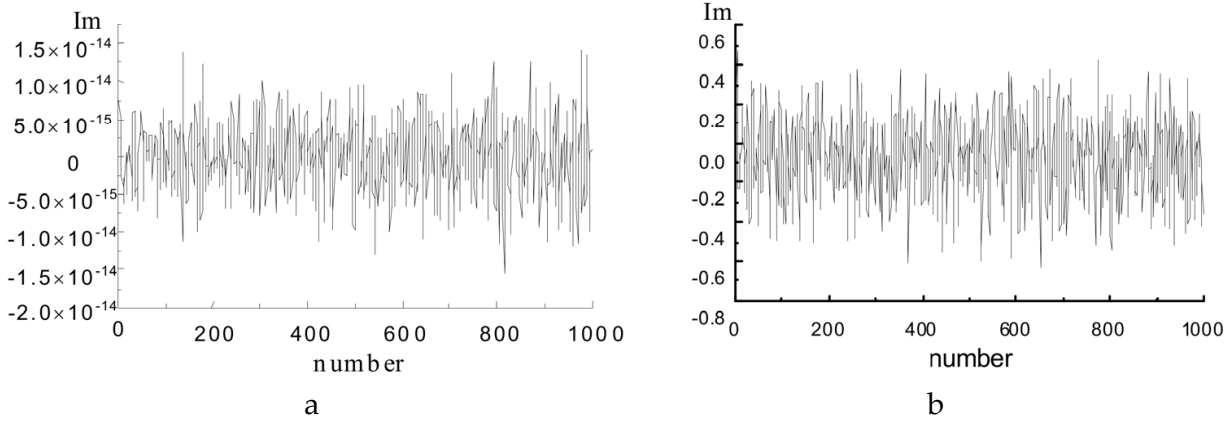


Figure 1. The imaginary components of surrogate data by our (a) and previous (b) FT algorithm

Thus, there are no imaginary components (see Fig. (1a)). The values of the imaginary parts are very little (magnitude 10^{-14}) so that they can be regarded as computing precision errors. The surrogate data has the same Fourier transform spectrum as the original data by using this algorithm. The reproduced “pure” frequencies are very well. Fig.(1b) shows that the previous FT algorithm[1] cannot make the imaginary components of Fourier inverse transform to be zero. So, if one only uses the real part of Fourier inverse transform as surrogate data and omit its imaginary components, the obtained surrogates would have the two limitations[1, 12].

Null hypothesis 4 The observed data is produced by the static nonlinear transform of linear gaussian process.

The static nonlinear transform is that the observation or measure function is nonlinear. The static means that the measure data x_t only depends on the state y_t of the dynamic process at the time t , not on derivatives or values in the past. Let h be a measure function, then

$$x_t = h(y_t) \quad (11)$$

The generated surrogate data not only contain the linear correlated characteristic, but also can reflect the static, monotonic nonlinearity of the original data. Strictly speaking, time series in this class are nonlinear. But this nonlinearity is not from the dynamics. This hypothesis can be used to indicate whether the nonlinearity is from the dynamical system or the amplitude distribution (i.e. the measure process).

For generating surrogate data corresponding to this null hypothesis, an algorithm is described. The aim is to shuffle the time-order of the data x_t and to preserve the linear correlations of the underlying time series $y_t = h^{-1}(x_t)$. The first step is to make a Gaussian time series y_t , where each element is generated independently from a Gaussian pseudorandom number generator. Next, we rescale y_t in accordance with the time-order of the original data x_t . The re-ordered y_t has a time series which “follows” the static, monotonic nonlinearity of the original data. Then, the data y'_t is created by using the above FT algorithm to deal with the re-ordered y_t . Finally, the raw data x_t is rescaled in terms of the time-order of the data y'_t to generate the surrogate data x'_t . The “underlying” time series (y_t and y'_t) are Gaussian and have the same Fourier power spectrum. The produced x'_t matches the amplitude distribution of the raw data x_t .

2.1.2. Test statistics[6]

The test statistic is a value which estimates some certain aspects of the time series. To compare the raw data to its surrogate data sets, a suitable test statistic must be selected. A useful statistic should be pivotal and independent of the way that surrogate data sets are generated. In other words, for every data set z and every realization z_i of any $F_i \in F_\phi$ their test statistics should be different, i.e.

$$T(z) \neq T(z_i) \quad (12)$$

where F_ϕ represents the null hypothesis process. Meanwhile, the distribution of T under the null hypothesis does not depend on μ or σ . Here we give two discriminating statistics as follows:

$$T = \overline{(x - \bar{x})^4} / \overline{(x - \bar{x})^2}^2 \quad (13)$$

$$T = \frac{1}{n-1} \sum_{i=1}^{n-1} |(x(i) - \bar{x}) \cdot (x(i+1) - \bar{x})| / \sqrt{\overline{(x - \bar{x})^2}} \quad (14)$$

where “ $\bar{}$ ” denotes the average of the data. The mean μ and variance σ have no effect on the T value in Eq. (13). Therefore, some linear structure characteristics can be determined except for the mean and variance. The T value in Eq.(14) can judge if the surrogate data are consist with the raw data in the view of the correlation with the mean and variance. The T value in Eq. (15) is a simple skewed difference statistic that is both rapidly computable and often quite powerful.

$$T = \left\langle (x_{t+m} - x_t)^3 \right\rangle / \left\langle (x_{t+m} - x_t)^2 \right\rangle \quad (15)$$

where $\langle \cdot \rangle$ is mean operator, m is time delay. This statistic T provides a more significant rejection of the hypothesis of the static nonlinear filter of an underlying linear process. Informally, this statistic indicates the asymmetry between rise and fall times in the time series.

2.1.3. Performance of surrogate data method based on our FT algorithm[3, 13, 14]

The surrogate data method is suitable to detect the nonlinearity of a short, noisy time series. Here, a Gaussian data and a Logistic chaotic time series are used to study the performance of surrogate data method. For a two-sided test, the probability of rejecting the null hypothesis is given by the confidence level p , the surrogate data sets B must be at least as large as $B_{\min} = 2/(1 - p) - 1$. For 95% confidence level, there should be 39 sets of surrogate data.

A Gaussian data x is a random time series with zero mean and unit variance produced by the pseudorandom generator. The data length is 1000 points. According to the null hypothesis 3, 39 sets of surrogate data are generated by using our above FT algorithm. The T value is calculated by Eq. (13) and Eq. (14), respectively. In the Figure (2a and b), there are no statistical discrepancy between the test statistic T of the raw data x and those of surrogate data.

The statistic T values of the raw data are on the range of the empirical distribution of T given by the surrogate data. The results show that the generated surrogate data has the same Fourier transform spectrum as the raw data besides the same mean and variance as the raw data because the T value in Eq. (14) is a measure of the time irreversibility of the data. The null hypothesis 3 is accepted at the confidence level 95%. The raw data is consistent with the stochastic process of the null hypothesis 3. The surrogate data produced by the above FT algorithm is equivalent to the raw data. The generated surrogate data reflects the null hypothesis 3.

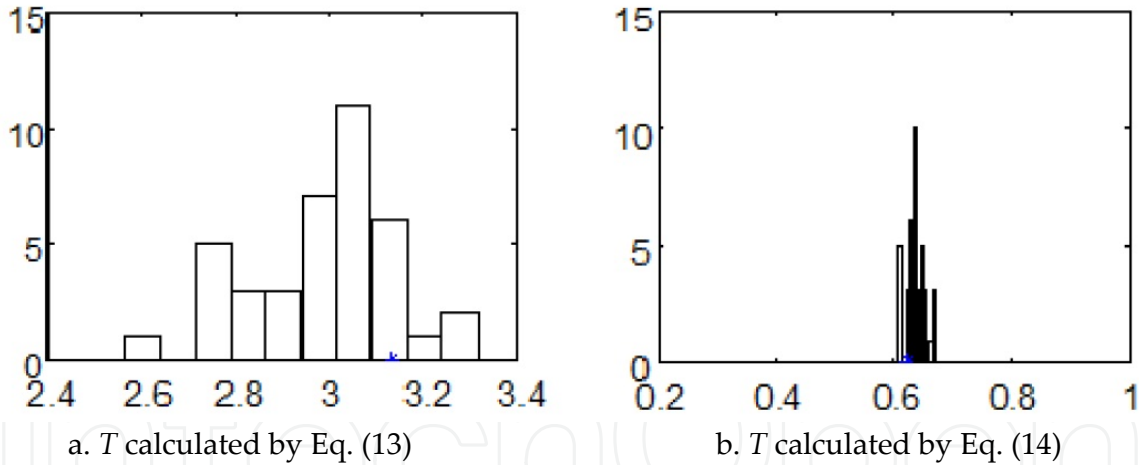


Figure 2. The histogram is T distribution of surrogate data given by FT algorithm, * is T value of the original data, where abscissa is T , ordinate is the number of surrogate data sets

To further test that the surrogate data based on the above FT algorithm can be used to detect nonlinearity of a time series, we apply the Logistic chaotic system to produce a chaotic time series as follows.

$$x_{t+1} = \alpha x_t (1 - x_t) + e \quad (16)$$

where $\alpha = 3.9$, $x_0 \in [0, 1]$, e is white noise with mean 0, variance 0.001². If e takes part in the evolution process of the above equation, it is called as interior noise (or dynamic noise); otherwise it is called as measure noise.

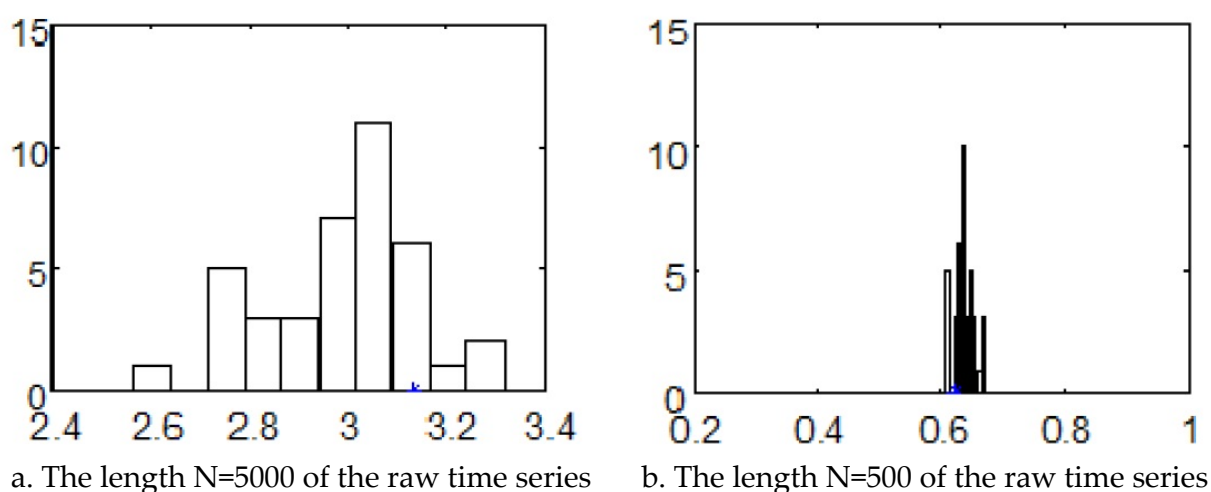


Figure 3. T calculated by Eq. (13), histogram is T distribution of surrogate data, * is T value of the original data, where abscissa is T , ordinate is the number of surrogate data sets

For the case without noise $e = 0$, we use Eq. (16) to compute the two Logistic chaotic time series with the length of 5000 points and 500 points. 39 surrogate data generated by the above FT algorithm contain the linear properties of the original data in terms of the null hypothesis 3. In Fig.3, we can see the obvious difference between the original data and its surrogate data, regardless of the length of 5000 points or 500 points. The null hypothesis can be rejected in 95% confidence level. The original data has nonlinear components. The results show that the data length has little effect on the surrogate data method based on the above FT algorithm. In Figure 4, we study the nonlinear test of Logistic chaotic time series with measure noise and interior noise, respectively. The data length is 1000 points. According to the null hypothesis, 39 sets of surrogate data are generated. The statistic T for the original data is significantly different from the values obtained for the surrogate data sets. The null hypothesis 3 can also be rejected in 95% significance. The nonlinearity of the original data can be detected. To sum up, the length and noise has no impact on the surrogate data method based on our FT algorithm.

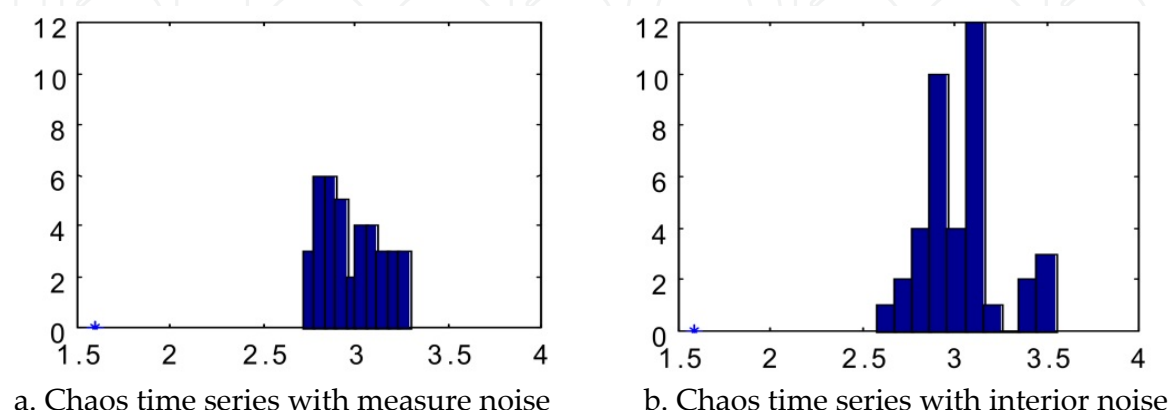


Figure 4. T calculated by Eq.(13), histogram is T distribution of surrogate data, * is T value of the original data, where abscissa is T , ordinate is the number of surrogate data sets

2.1.4. Nonlinear test of surface EMG signal based on surrogate data[3, 13, 14]

The nature of SEMG plays an important role in neuromuscular disorder diagnosis, muscle fatigue monitoring, prosthesis control, etc. Here the analyzed data are collected from physiological instruments. Humid surface electrode and AD12-16LG collecting card of physiology signal are used in the whole experiment that was done at Hua Shan Hospital in Shanghai. The data are sampled at 1kHz for the action surface EMG (ASEMG) [3] and the fatigue surface EMG (FSEMG) when one hand carries a 1kg heavy thing [15] (see Fig. 5). The length of data for ASEMG is 1000 points during the beginning of action because this time span contains the information of the forearm movement. In the case of carrying a 1kg heavy thing, the length of FSEMG data is also 1000 points when the arm has been fatigue.

For these surface EMG signals, 39 surrogate data are produced by the null hypothesis 2. The surrogate data analysis is given for the action surface EMG signal and the fatigue surface EMG signal, respectively (see Fig. 6). The results show that for action surface EMG signal and fatigue surface EMG signal, their T values are obviously different from those of surrogate data in terms of Eq.13. The null hypothesis 2 can be rejected in 95% degree of confidence. The action surface EMG signal and fatigue surface EMG signal is not stochastic signal produced by a linear stochastic process, but contains nonlinear components. However, this result could not ensure that this nonlinearity must be from the dynamic system.

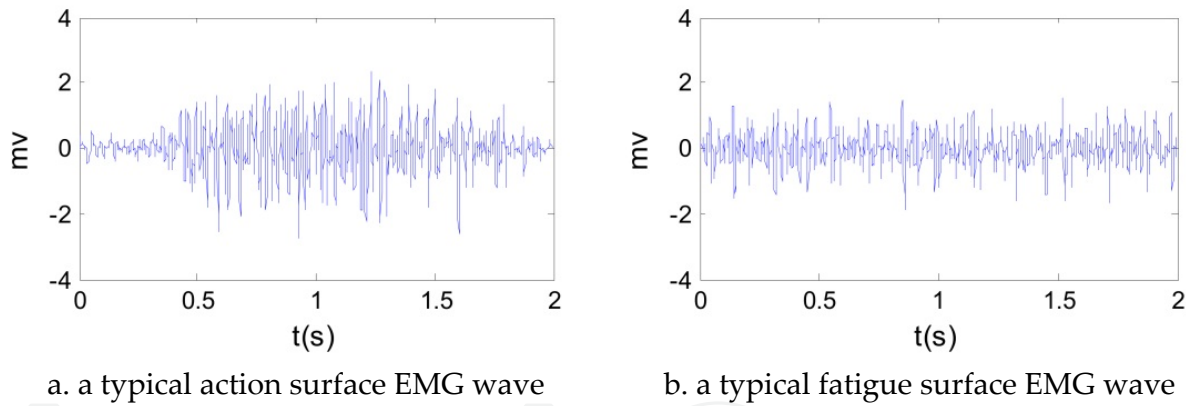


Figure 5. The surface EMG signals

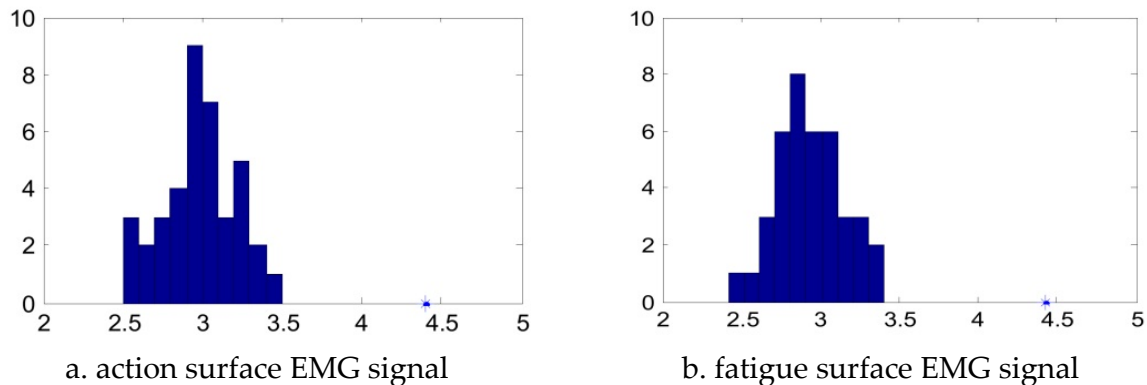


Figure 6. Surrogate data analysis of surface EMG signal. * is T_{orig} , histogram is T_{surr} distribution of surrogate data

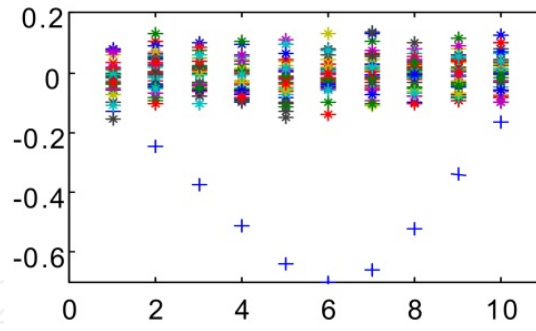


Figure 7. Surrogate data test of surface EMG signal during movement, where surrogate data sets are 39 sets; * is T value of surrogate data by the null hypothesis 4, + is T value of EMG signal, where T is calculated by Eq. 15

In order to test that the nonlinear components are intrinsic deterministic, we further assume that ASEMG is stochastic signal consistent with the null hypothesis 4. Fig.7 gives the T values of ASEMG and surrogate data calculated by Eq. 15. This statistic indicates the asymmetry between rise and fall times in the time series. From this figure, we can see that there is the difference between data and surrogates, and the null hypothesis 4 is rejected in 95% credibility. This result shows that the nonlinearity of ASEMG is intrinsic and deterministic.

2.2. Volterra-Wiener-Korenberg test method

2.2.1. Volterra-Wiener-Korenberg test model[2]

For a dynamic system, an observed time series $\{y_n\}_{n=1}^N$ can be treated as a closed loop Volterra series by utilizing the feedback of y_n . Suppose the time series is univariate. A discrete Volterra-Wiener-Korenberg series can be calculated as follows:

$$\begin{aligned} y_n^{cala} &= a_0 + a_1 y_{n-1} + a_2 y_{n-2} + \cdots + a_k y_{n-k} + a_{k+1} y_{n-1}^2 + a_{k+2} y_{n-1} y_{n-2} + \cdots + a_{M-1} y_{n-k}^d \\ &= \sum_{m=0}^M a_m z_m(n) \end{aligned} \quad (17)$$

where the memory k and combination degree d correspond to the embedding dimension and the degree of nonlinearity of the model, respectively. The coefficients a_m are recursively estimated through a Gram-Schmidt procedure from linear and nonlinear autocorrelations of the data itself with a total dimension $M=(k+d)!/(d!k!)$.

There is the following information criterion in accordance with the parsimony principle:

$$C(r) = \log \varepsilon(r) + r/N \quad (18)$$

$$\varepsilon(k, d)^2 \equiv \frac{\sum_{n=1}^N (y_n^{calc}(k, d) - y_n)^2}{\sum_{n=1}^N (y_n - \bar{y})^2} \quad (19)$$

where $r \in [1, M]$ is the number of polynomial terms of the truncated Volterra expansions from the given pair $\{k, d\}$, $\varepsilon(k, d)^2$ is a normalized variance of the error residuals, $\bar{y} = \frac{1}{N} \sum_{n=1}^N y_n$. For $d=1$, VWK model is linear, whereas the model is nonlinear for $d>1$.

For each data series, there is the following numerical procedure to search for the optimal pair $\{k_{\text{opt}}, d_{\text{opt}}\}$:

1. when $d=1$, search for k_{opt} which minimizes $C(r)$.
2. with $k=k_{\text{opt}}$, increasing $d>1$, search for d_{opt} which minimizes $C(r)$.
3. calculate $C^{\text{lin}}(r)$ with $d=1$ and $k=M-1$, and $C^{\text{nl}}(r)$ with $d=d_{\text{opt}}$ and $k=k_{\text{opt}}$.
4. Compare $C^{\text{lin}}(r)$ and $C^{\text{nl}}(r)$, if $C^{\text{nl}}(r)$ is obviously smaller than $C^{\text{lin}}(r)$, then the original system dynamics is nonlinear, the obtained time series is nonlinear, even chaos; otherwise, the original system dynamics is linear, the raw data is linear.

Note that when k_{opt} is rather large, M is quite large, too, then the computational time will rapidly go up. In this case, k and d should be adjusted synchronously to search for k_{opt} and d_{opt} so as to make $C^{\text{nl}}(r) < C^{\text{lin}}(r)$. Furthermore, one can obtain the corresponding linear and nonlinear models for surrogate data generated by the FT algorithm according to the null hypothesis 3 so that $C_{\text{orig}}^{\text{lin}}(r), C_{\text{surr}}^{\text{lin}}(r), C_{\text{surr}}^{\text{nl}}(r) > C_{\text{orig}}^{\text{nl}}(r)$ in the statistical sense.

2.2.2. Analysis of SEMG based on VWK method[15, 16]

Here, the VWK method is used to deal with the surface EMG signals in Fig.5. For the action surface EMG signal, $C^{\text{lin}}(r)$ is almost equal to $C^{\text{nl}}(r)$, i.e. $C^{\text{lin}}(r) \approx C^{\text{nl}}(r)$, this technique can hardly detect its nonlinearity (see Fig.8a). For the fatigue surface EMG signal, $C^{\text{nl}}(r)$ is distinctly smaller than $C^{\text{lin}}(r)$ so that its nonlinear component can be detected (see Fig.8b). So VWK technique can effectively detect the nonlinear dynamic speciality of fatigue surface EMG signal but fails to test the nonlinearity of the action surface EMG signal. In other words, VWK technique can not be used directly to deal with the action surface EMG signal.

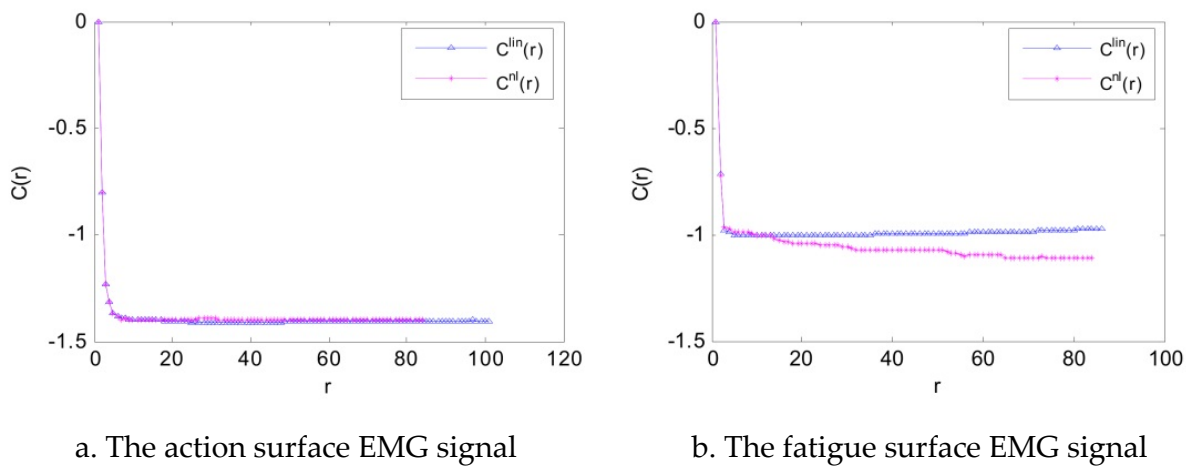


Figure 8. VWK test analysis of surface EMG signal

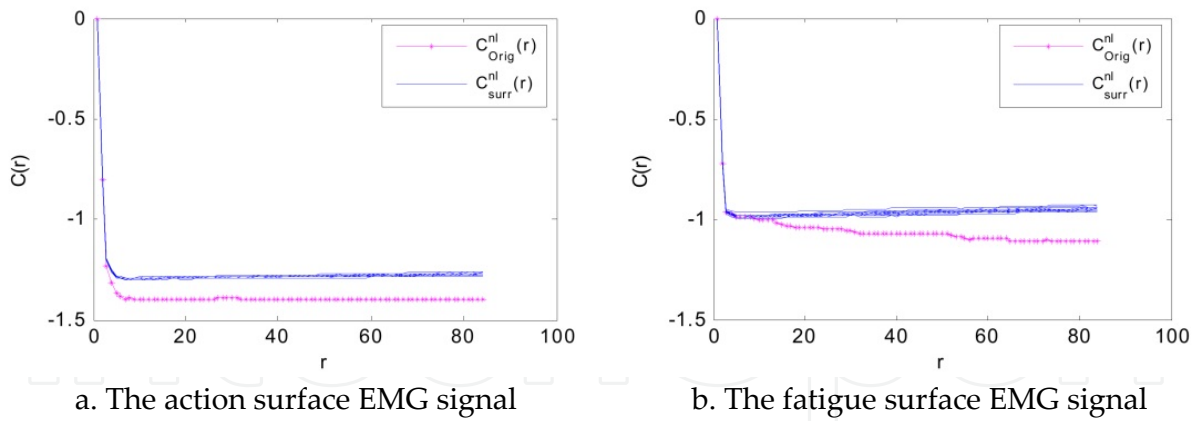


Figure 9. VWK combined with surrogate analysis of surface EMG signal. solid is $C_{surr}^{nl}(r)$, * is $C_{orig}^{nl}(r)$

2.2.3. Analysis of SEMG based on VWK method with surrogate data[16]

In order to detect the nonlinearity of the action surface EMG signal, 39 FT-based surrogate data are used according to the null hypothesis 3. The generated surrogate data contain the linear properties of the raw data. Figure 9 is the analysis of surface EMG signal based on VWK with surrogate data. We can see that no matter whether it is the action or fatigue EMG signal, $C_{orig}^{nl}(r)$ is always smaller than $C_{surr}^{nl}(r)$. The null hypothesis 3 can be rejected in 95% significance. The results illuminate that the action and fatigue surface EMG signals contain nonlinear dynamic properties.

3. Analysis of the surface EMG signals based on chaos theory

The discovery of chaotic phenomena is the third major breakthrough in the 20th century physics scientific community following the creation of relativity and quantum mechanics. It organically combines the two major theoretical systems of determinism and probabilism that have long been debated to create a scientific model of a new paradigm, so that people can use some simple rules to explain seemingly stochastic information in the past[17-19]. The practical significance that finds chaotic phenomena is to recognize that a deterministic nonlinear system can have inherent uncertainty. Perhaps a system has only a few degrees of freedom, but it can produce complex, similar to the random output signal. In the past, one could only denote a random-looking data as a random process from the view of the traditional time series analysis. The statistical methods or random time series models were used to analyze the data. Since the chaotic phenomenon was discovered by Lorenz[17], people have begun to reunderstand and restudy these random-looking signals so as to reveal the inherent deterministic mechanisms of these signals. That is, it is to explore that the systems which generate these signals may contain essentially deterministic characteristics. Chaos phenomenon breaks the path that the regularity is found in a lot of completely different systems. This will lead to a revolution in the field of influence of various disciplines. It is chaos to lead people to explore the complexity in nature.

At present, the idea of Chaos has been introduced into the analysis of time series to create the field of chaotic time series analysis. Since the inception of chaotic time series analysis, it has quickly been penetrated into other disciplines and engineering fields. Thus it becomes the most active branch of the modern nonlinear dynamics. This section describes the chaos definition and the phase space reconstruction of chaotic time series, discusses some parameters that are used to analysis chaotic time series, such as the correlation dimension and Lyapunov exponent, study the principal component analysis methods based on SVD, and propose the symplectic principal component method based on symplectic geometry. Then we use these methods to investigate the surface EMG signals.

3.1. Chaos and its definition

Chaos is “order in disorder”. The order means its deterministic nature. The disorder means that the final results can be unpredictable for a long time. As a scientific concept, chaos generally denotes that the long-term dynamical behavior of a deterministic nonlinear system manifests as a random-looking behavior. Mathematically speaking, “chaos” has not been a unified strict definition. For the definition of chaos, there are at least nine different definitions, where the three definitions given by Li-Yorke, Devaney, Marotto are more commonly used. Here describes the definition of chaos by Li-Yorke[18].

Li-Yorke Theorem: Let $f(x)$ as a continuous self-map in $[a, b]$. If $f(x)$ has a periodic point with period 3, then for any positive integer $n=1, 2, 3, \dots$, there is a periodic point with period n .

This is the famous period 3 theorem. It becomes a milestone in the development history of chaos theory and promotes the creation and development of chaos theory. From this theorem, the first formal mathematical definition of the chaos is given.

Chaos definition: Let $f(x)$ as a continuous self-map in closed interval I , i.e.

$$f: I \rightarrow I \subset \mathbb{R}^m, \quad y = f(x) \quad (20)$$

where $x, y \in I$. If $m=1$, f is one-dimensional mapping. If $m \neq 1$, f is multi-dimensional mapping. Denote the n times iteration of f as $f^n(x)$. If Eq.20 satisfies the following conditions, then it has chaotic motion:

1. The period of periodic point of f has no upper bound.
2. There is an uncountable set $S \subset I$, which satisfies the following conditions:

$$\limsup_{n \rightarrow \infty} |f^n(x) - f^n(y)| > 0, \quad \forall x, y \in S, \quad x \neq y \quad (21)$$

$$\liminf_{n \rightarrow \infty} |f^n(x) - f^n(y)| = 0, \quad \forall x, y \in S \quad (22)$$

$$\limsup_{n \rightarrow \infty} |f^n(x) - f^n(p)| > 0, \quad \forall x \in S, \quad \forall p \in P(f) \quad (23)$$

where $P(f) \triangleq \{x \mid x \text{ is a periodic point of } f\}$.

This definition explains “existence” of chaos in mathematics. According to the above theorem and the definition, the description of chaotic motion is different from the general periodic and quasi-periodic motion. Its motion is not a single periodic orbit but an envelope for a bunch of tracks, where the infinite number of countable stable periodic orbits and uncountable stable aperiodic orbits are embedded densely. Meanwhile, there is at least one unstable aperiodic orbit. Overall, the chaos not only contains some inherent regularity, but also shows that the system has ergodicity. That is, the system has a long-term unpredictability. In other words, the long-term behavior of the system can not be predicted if the system displays the so-called “sensitive dependence on initial conditions”. The meaning of this definition is that the aperiodicity of chaotic system is exhibited accurately. For a dynamical system, the observable behaviour was called stochastic in the past. In fact, it can be random-looking, i.e. “stochastic behaviour occurring in a deterministic system”. Therefore, it is challenging to quantitatively describe the nature of chaotic dynamics and distinguish between the so-called random and chaotic motions from a time series, especially from an experimental time series. At present, chaotic time series analysis methods have been widely attention in fields of mathematics, physics, biology, biomedicine, robotics, geology, engineering, economics, finance, and so on.

3.2. Phase space reconstruction theory

3.2.1. Phase space reconstruction

Phase space reconstruction is generally the first step of chaotic time series analysis from a time series data. The dynamic characteristic of the system can be explored through phase space reconstruction of the original time series so that the mechanism of the original system can be revealed from the original time series[20]. It has been proved by the so-called Takens’ embedding theorem[21]. According to the theorem, the reconstructed phase space can maintain the invariance of geometry for the original dynamical system[22], such as the characteristic value of the fixed point, the fractal dimension of the attractor, the Lyapunov exponent in the phase space orbit, and so on.

Definition 1: Let (N, ρ) and (N_1, ρ_1) as two metric spaces. If the mapping $\varphi: N \rightarrow N_1$ satisfies ① φ is a surjection; ② $\rho(x, y) = \rho_1(\varphi x, \varphi y)$, then (N, ρ) and (N_1, ρ_1) are called as the isometric isomorphism.

Definition 2: If (N_1, ρ_1) is isometric isomorphism with a subspace (N_0, ρ_2) of another metric space (N_2, ρ_2) , then (N_1, ρ_1) can be embedded into (N_2, ρ_2) .

Theorem 1: Let M be a compact manifold of dimension m . For pairs (φ, x) , $\varphi: M \rightarrow M$ a smooth diffeomorphism and $x: M \rightarrow R$ a smooth function, it is a generic property that the map $\phi_{(\varphi, x)}: M \rightarrow R^{2m+1}$ is an embedding, where $\phi_{(\varphi, x)}(y) = \{x(y), x(\varphi(y)), \dots, x(\varphi^{2m}(y))\}$.

In terms of the above definitions and theorem, $\phi_{(\varphi, x)}(y) = \{x(y), x(\varphi(y)), \dots, x(\varphi^{2m}(y))\}$ is a subspace of R^{2m+1} . $\phi_{(\varphi, x)}$ and the subspace are isometric isomorphic. Then the manifold M of dimension m can be embedded into R^{2m+1} . In other words, $\phi_{(\varphi, x)}$ is an embedding of

$M \rightarrow R^{2m+1}$. For a practical time series $\{x_t\}$, the state of the original system is equivalent to the m -dimensional manifold M . In fact, $\{x_t\}$ is a signal observed in the m -dimensional manifold M . If let $\varphi: y_t \rightarrow y_{t-\tau_d}$, φ is a smooth diffeomorphism. y_t denotes the state of the system M at time t . τ_d is the delay time. The signal observed in M at time t consist of $\{x_t, x_{t-\tau_d}, \dots, x_{t-2m\tau_d}\}$, where $x_t = x(y_t)$, $x_{t-\tau} = x(y_{t-\tau_d}) = x(\varphi(y))$, \dots , $x_{t-2m\tau_d} = x(\varphi^{2m}(y))$. $\phi_{(\varphi, x)} = (x_t, x_{t-\tau_d}, \dots, x_{t-2m\tau_d})$ is an embedding of $M \rightarrow R^{2m+1}$. The manifold M is diffeomorphic with $\{x_t, x_{t-\tau_d}, \dots, x_{t-2m\tau_d}\}$. If the embedding dimension is greater than 2 times the dimension m of the attractor of the original system, the phase space with the base of the practical signal delay time coordinates is equivalent to the state space of the original system. That is, Takens' embedding theorem states that if the time series is indeed composed of scalar measurements of the state from a dynamical system, then under certain genericity assumptions, a one-to-one image of the original set $\{x\}$ is given by the time-delay embedding, provided d is large enough. At present, the delay coordinate method has been widely used to give the phase space reconstruction from the original signal.

For a time series $x(t)$ observed by the measure function h , i.e.

$$x(t) = h(Y) \quad (24)$$

the vector \overline{X}_t can be constructed as follows,

$$\overline{X}_t = (x(t), x(t - \tau_d), x(t - 2\tau_d), \dots, x(t - (d-1)\tau_d))^T \quad (25)$$

where τ_d is an integer multiple of the sampling interval τ , called as the lag time or delay time. d is an embedding dimension, $d \geq 2m+1$. m is an attractor dimension of the original system.

3.2.2. Problems in phase space reconstruction

Takens' embedding theorem offers in the absence of noise, the possibility of reconstructing n -dimensional dynamics from one-dimensional infinite data of one observable-measurable system. This means that in the case of any delay time, a time series can always be embedded into the state space of the system, and when the embedding dimension is sufficiently large, reconstructed space and embedded space is almost one-to-one correspondence. Therefore, one can reconstruct a phase space from an experimental time series so as to estimate dynamical invariants of the time series, such as dimensions, Lyapunov exponents, entropies[21, 23, 24] and so on. However, the embedding theorem does not directly answer how to choose embedding dimension d and delay time t . In practical application, the experimental data is always limited and noisy so that the estimation of the above parameters presents some difficulties[25, 26]. Accuracy of the phase space reconstruction is critically important to the estimation of invariant measures characterizing system behavior. The choice of delay time τ_d and embedding dimension d always has a great impact on the phase space reconstruction.

Some researchers have studied the choice of delay time τ_d [26-30]. If the delay time τ_d is too small, the reconstructed attractor will be crowded around the main diagonal, which is called as redundancy. If τ_d is too large, the dynamic shape of the attractor will be broken, which is called as irrelevance, and the phase space reconstruction is no longer representative of the true dynamics in the real system[28]. In normal circumstances, in order to make the elements of \bar{Y}_t independent, τ_d is the same for all the embedding dimension d [27-29]. The autocorrelation function method and mutual information technique[30] have been most commonly used to give the delay time τ_d although the issue of the delay time choice has still been completely resolved.

For the embedding dimension d , there are three methods that are usually used to choose the appropriate embedding dimension, including the correlation dimension, singular value decomposition(SVD), the false neighbors[21, 31, 32]. The correlation dimension method is to estimate appropriate dimension d in terms of the correlation theorem[8, 21, 33]. By increasing the embedding dimension, one notes an appropriate dimension d when the value of the correlation dimension stops changing. Broomhead and King[31] used the singular value decomposition (SVD) technique to determine an appropriate embedding dimension d directly from the raw time series. The false neighbor method is based on the fact that choosing a too low embedding dimension results in points, which are far apart in the original phase space, being moved closer together in the reconstruction space[32]. Besides, there are also some other methods and modified extensions developed based on the above methods. However, there are still problems on how to determine the appropriate embedding dimension from a scalar time series[34-38].

3.3. Correlation dimension

If a system is chaotic, the strange attractor in a region of the phase space constitutes an infinite hierarchy of self-similar structure, i.e. a fractal structure. One can use quantitative measures to define the fractal nature. The correlation dimension is a useful measurement. Grassberger and Proccacia give a kind of computation method, called GP algorithm[33, 39, 40].

3.3.1. GP algorithm of correlation dimension

Let X_1, X_2, \dots, X_n be a point of the attractor in phase space. $C_l(X_j)$ is denoted as a hypersurface sphere with the radius l at the reference point X_j . $\mu[C_l(X_i)]$ is the probability that $X_i (i=1, \dots, n)$ falls into $C_l(X_j)$, as follows.

$$\mu[C_l(X_j)] = \frac{1}{n} \sum_{i=1}^n \theta(l - \|X_i - X_j\|) \quad (26)$$

where $\|\bullet\|$ is Euclidean norm. $\theta(r)$ is Heaviside function whose value is 1 if $r \geq 0$, otherwise, zero.

Then a correlation integral function is defined as

$$C(l) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \theta(l - \|X_i - X_j\|) \quad (27)$$

with $l \rightarrow 0$, there are a scaling relation $C(l) \propto l^{D_2}$ between the correlation integral $C(l)$ and l . A correlation dimension D_2 is defined.

$$D_2 = \lim_{l \rightarrow 0} \frac{\ln C(l)}{\ln(l)} \quad (28)$$

In practical computation, D_2 is the slope of $\log C$ vs $\log l$ curve over a selected straight line range.

3.3.2. Correlation dimension theorem[41]

Theorem 2: Let a map $G: R^n \rightarrow R^n$. A is an attractor of map G that has only a finite number of periodic points of period P . Under a natural probability measure μ , the correlation dimension of A is $D_2(\mu)$. For a measure function h of A , $h: R^n \rightarrow R$, define a delay coordinate map $F_h: R^n \rightarrow R^d$ as

$$F_h(X) = [h(X), h(G^{-1}(X)), \dots, h(G^{-(d-1)}(X))] \quad (29)$$

where $d \geq P$. $F_h(\mu)$ is a natural probability measure in R^d of $F_h(A)$. If $d \geq D_2(\mu)$, then for almost every h , $D_2(F_h(\mu)) = D_2(\mu)$.

The theorem says that with the embedding dimension increasing, the slope of corresponding correlation integral curve will converge to the correlation dimension D_2 of the original system attractor. Therefore, the optimal embedding dimension can be estimated by using the correlation dimension D_2 . That is, if the embedding dimension $d \geq D_2$, the slope of the correlation integral curve is equal to the correlation dimension. This also indicates that the dimension estimation actually does not have to meet the requirements of the embedding theorem on the embedding dimension $d \geq 2D_2 + 1$. When the embedding dimension $d \geq D_2$, the reconstructed attractor can contain the fractal structural feature of the original system attractor to reflect the chaotic characteristics of the original system. Correlation dimension has been widely used in the analysis of chaotic time series.

3.3.3. Chaotic test based on correlation dimension

Chaos has a fractal structure so that the corresponding correlation dimension D_2 is a fractional value. The estimation of correlation dimension D_2 from a time series can be used to determine whether the time series is chaotic. If D_2 is fractional, the original time series can have chaotic features, otherwise, it cannot be chaotic. According to the correlation dimension theorem, when the embedding dimension d of the reconstructed phase space is increased to a certain value, the correlation dimension D_2 will be saturated. Then, the

optimal embedding dimension d will be given from a time series. The corresponding correlation dimension D_2 is called as the correlation dimension of this time series.

Lorenz chaotic time series is given by the state variable x of Lorenz system as follows.

$$\begin{cases} \frac{dx}{dt} = \sigma(y - x) \\ \frac{dy}{dt} = rx - y - xz \\ \frac{dz}{dt} = -bz + xy \end{cases} \quad (30)$$

where $\sigma=10$, $b=8/3$, $r=28$, initial conditions: $x(0)=5$, $y(0)=5$, $z(0)=15$. The sampling interval $\tau=0.1$. The sampling points $N=1000$. For delay time $\tau_d=\tau$, the corresponding correlation dimension values are given in Table 1 when the embedding dimension d is increased from 2 to 12. From this table, we can see that the correlation dimension of the time series is about 2.07. The result shows that the reconstructed attractor has a fractal structure to reflect the chaotic feature of the system. The time series can reconstruct the state space of the original system when the embedding dimension $d=6$.

Logistic chaotic time series $\{x_n\}_{n=1}^N$ is given by Logistic system in Eq.16 with $\alpha=3.9$ and $e=0$. The length N is 1000 points. Here, $\tau_d=1$ (i.e. discrete time series interval). With increasing the embedding dimension d , the corresponding correlation dimension is 0.97 (see Table 2). The optimal embedding dimension is 2.

For finite sampling number (e.g. $N=1000$), the reconstructed attractor will be broken when the embedding dimension d is increased continuously to a higher value. The estimation of correlation dimension will fail during computation. Therefore, embedding dimension d should not be unlimitedly increased.

d	2	3	4	5	6	7	8	9	10	11	12
D_2	1.8009	1.9284	1.9718	2.0389	2.0737	2.0966	2.086	2.0788	2.0705	2.0760	2.0753

Table 1. The analysis of correlation dimensions of Lorenz chaos time series

d	1	2	3	4	5	6	7	8	9	10	11
D_2	0.9598	0.9718	0.9689	0.9713	0.9718	0.9591	0.9774	0.9621	0.9827	0.9826	0.9839

Table 2. The analysis of correlation dimensions of Logistic chaos time series

3.3.4. The analysis of surface EMG signal based on correlation dimension

From the above analysis, we can see that the surface EMG signal has deterministic nonlinear component. Here, the correlation dimension is further used to study whether its nonlinear component are chaotic. Figure 10a shows a raw data for forearm pronation. Figure 10b gives the correlation integral curve of the data under the embedding dimension from 2 to 12. In

the reconstructed phase space, the delay time τ_d is chosen as the sampling interval. With the increase of embedding dimension, the straight line segments of the computed correlation integral curves will tend to be parallel and keep unchange in the range.

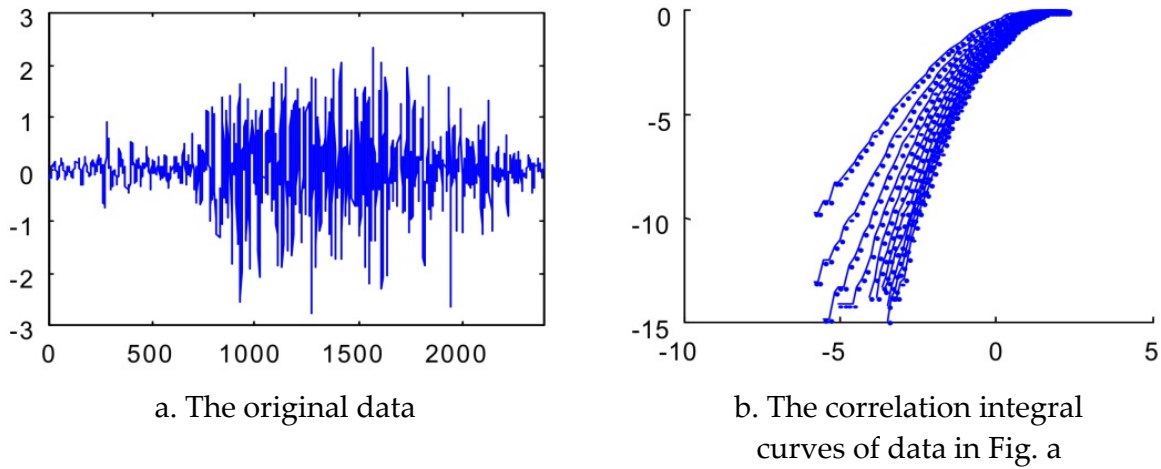


Figure 10. The correlation dimension analysis of surface EMG signal

The corresponding slope value is the correlation dimension of the surface EMG signal, about 3.8050 ± 0.0560 (see Table 3). The result indicates that the surface EMG signal during movement has fractal feature to reveal the implied chaotic motion behavior. Table 3 shows that the reconstructed attractor contains the nature of the raw system when the embedding dimension is over 6.

d	2	3	4	5	6	7	8	9	10	11	12
D_2	1.9315	2.7037	3.2808	3.4837	3.8047	3.7597	3.8511	3.8864	3.8129	3.7160	3.8039

Table 3. The analysis of correlation dimension of surface EMG signal during movement

3.3.5. Study of surrogate data test method based on correlation dimension

The correlation dimension is a quantitative index that describes the fractal structure of chaotic attractor. It measures the freedom degree and complexity of the system. For the raw data and all data of $F \in F_\phi$, in the case of the same embedding dimension, the corresponding test results will be obviously significant[42]. Here, the correlation dimension is used as a test statistic to analyze the surface EMG signal. According to the null hypothesis 3, 39 sets of surrogate data are produced in confidence level 95%. In order to quickly obtain the correlation dimension in a variety of circumstances, the linear parts of all correlation integral curves are taken as the same. Though these values are not accurate, this test algorithm is great effective to test the chaotic fractal feature of the experimental data.

When the sampling interval $\tau=1$, Figure 11a and b show the results of the surrogate data test analysis for Lorenz chaotic time series based on correlation dimension. There are significant differences between the original data and its surrogate data in $m = 5$ (see Fig. 11a). This result explains that the null hypothesis can be rejected in confidence level 95%. The

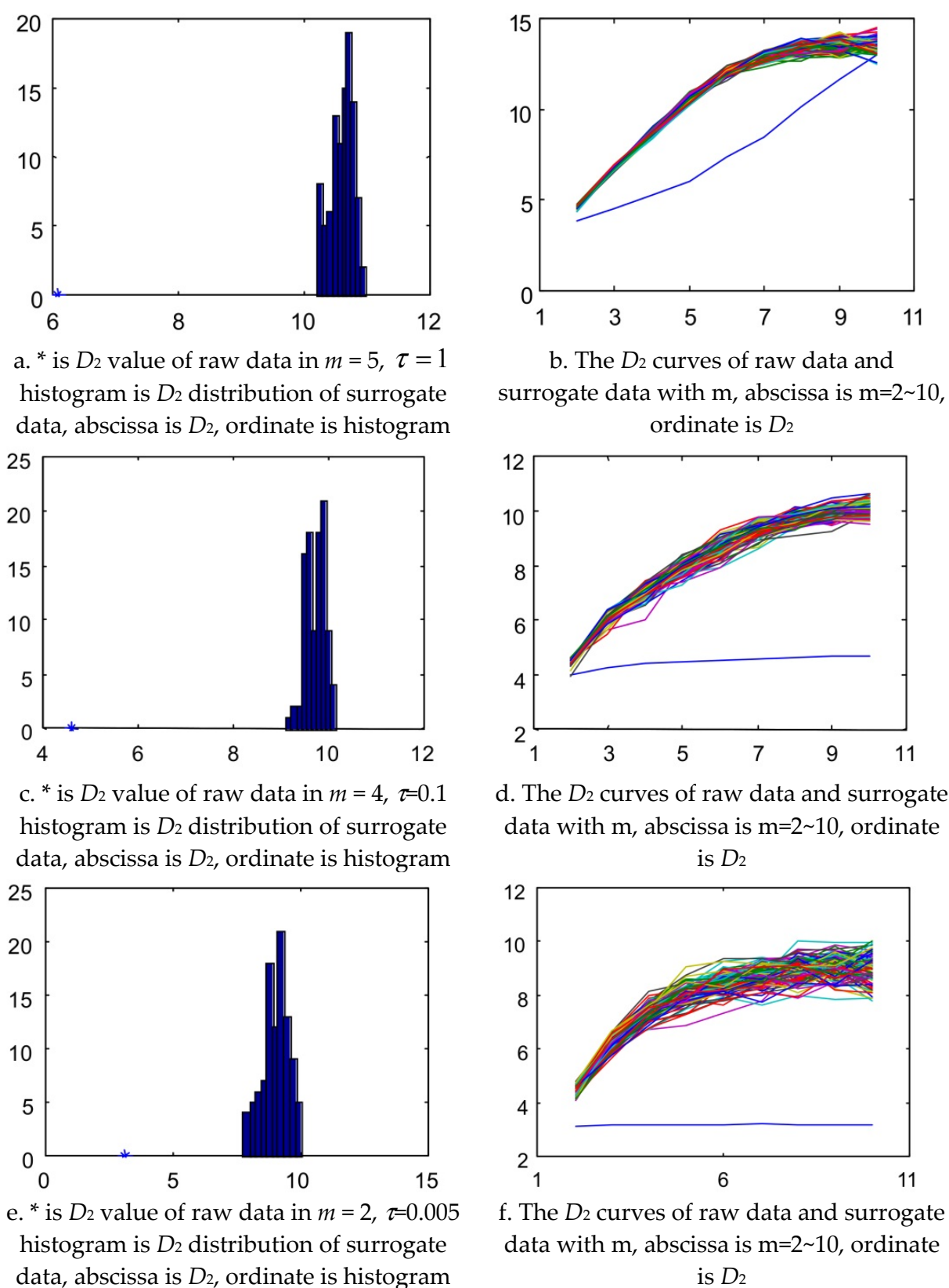


Figure 11. The surrogate data test analysis based on correlation dimension for Lorenz chaos time series by sampling intervals

differences disappear between the raw data and its surrogate data in $m = 10$ (see Fig. 11b). This illustrates that the reconstructed attractor appears broken. The reconstructed phase space is similar to that of the surrogate data with linear stochastic noise characteristics.

Figure 11c and d show the results with $\tau=0.1$. The differences between the raw data and its surrogate data can be seen in Fig. 11c ($m = 4$). When $m>2$, these differences become larger as m increases. Figure 11e and f show the results with $\tau=0.005$. Figure 11e shows the surrogate data test histogram in $m = 2$. The correlation dimension curves of the raw data and its surrogate data are given in $m = 2 \sim 10$ (see Fig. 11f). Even in the case of oversampling, the correlation dimension as test statistic can also make the surrogate data method very effective.

3.3.6. The surface EMG signal analysis based on the surrogate data and correlation dimension

Figure 12 shows the surrogate data analysis for the surface EMG signal in Fig. 12a based on correlation dimension. When $m = 6$, the correlation dimension value of the raw data is different from those of its surrogate data generated by the null hypothesis 3. The correlation dimension curves of the raw data and its surrogate data are given when $m = 2 \sim 8$ in Fig. 12b. We can see the differences between the original data and its surrogate data. The null hypothesis 3 can be rejected in confidence level 95%. The result indicates that the surface EMG signal has deterministic nonlinear components, even chaotic.

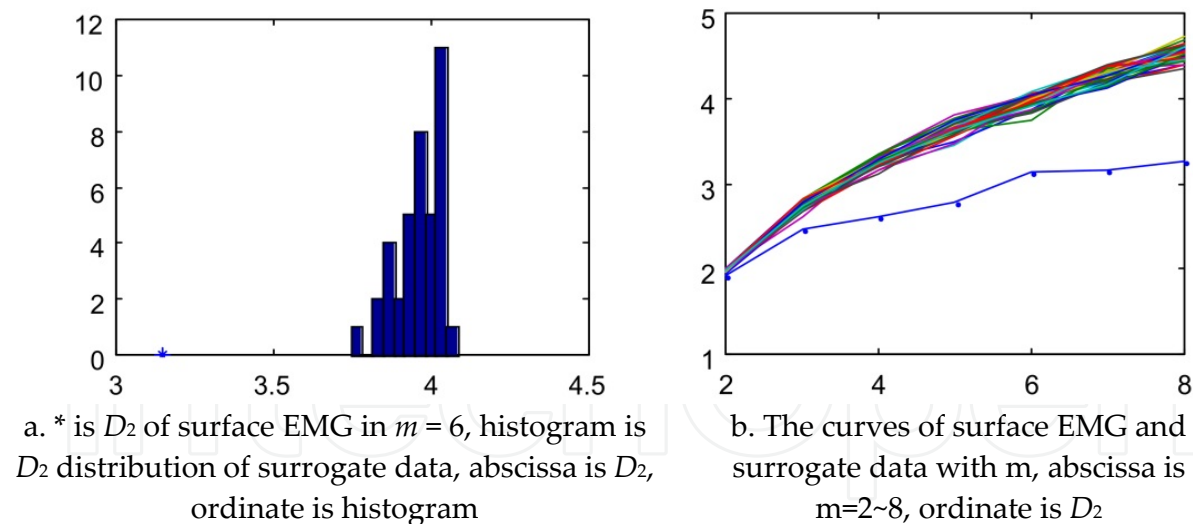


Figure 12. The surrogate data test analysis based on correlation dimension for surface EMG signal.

3.4. Largest Lyapunov exponent

The Lyapunov exponent method is to directly identify whether a system is chaotic. If the system is chaotic, the Lyapunov exponent is positive. Otherwise, the Lyapunov exponent is negative. For this, the Lyapunov exponent can be used to test the chaotic feature of a signal under study. The first algorithms developed computed the whole Lyapunov spectrum by

Wolf et al. [43] and Sano et al. [44]. Meanwhile, the largest Lyapunov exponent is sufficient for assessing the presence of chaos. At present, there are many algorithms to estimate the largest Lyapunov exponent from a time series, such as an algorithm given by Rosenstein et al.[45]. This algorithm is aimed specifically at estimating the largest Lyapunov exponent from short data.

3.4.1. Algorithm of largest Lyapunov exponent estimation[45]

For a short time series, Rosenstein et al. present a robust estimation algorithm of the largest Lyapunov exponent. First, the attractor is reconstructed, refer to Eq. 25. Next, the algorithm locates the closest neighbor of each point X_i on the trajectory, with respect to the Euclidian distance. Then, one defines the distance between two neighboring points at instant $n=0$ by:

$$d_i(0) = \min_{X_j} \|X_j - X_i\| \quad (31)$$

where $\|\bullet\|$ is the Euclidian norm. Here, the temporal separation of the nearest neighbors should be greater than the mean period of the time series.

$$|i - j| > \text{mean period} \quad (32)$$

According to time, the average distance between two neighboring vectors can be simply

$$d_i(n) = \|X_{j+n} - X_{i+n}\| \quad (33)$$

Assume that the system is controlled by the largest Lyapunov exponent only. Then, the distance between two neighbor points obey the following relationship:

$$d(t) = Ce^{\lambda t} \quad (34)$$

For $t = n\Delta t$, there is:

$$d_i(n) \approx C_i e^{\lambda n \Delta t} \quad (35)$$

$$\lambda = \frac{1}{n \cdot \Delta t} \ln \frac{d_i(n)}{C_i} \quad (36)$$

$$n \cdot \lambda = \frac{1}{\Delta t} (\ln d_i(n) - \ln C_i) \quad (37)$$

$$\begin{aligned} n \langle \lambda \rangle &= \frac{1}{\Delta t} \langle \ln d_i(n) \rangle - \frac{1}{\Delta t} \langle \ln C_i \rangle \\ &= \frac{1}{\Delta t} \langle \ln d_i(n) \rangle - b \end{aligned} \quad (38)$$

where $b = \frac{1}{\Delta t} \langle \ln C_i \rangle$.

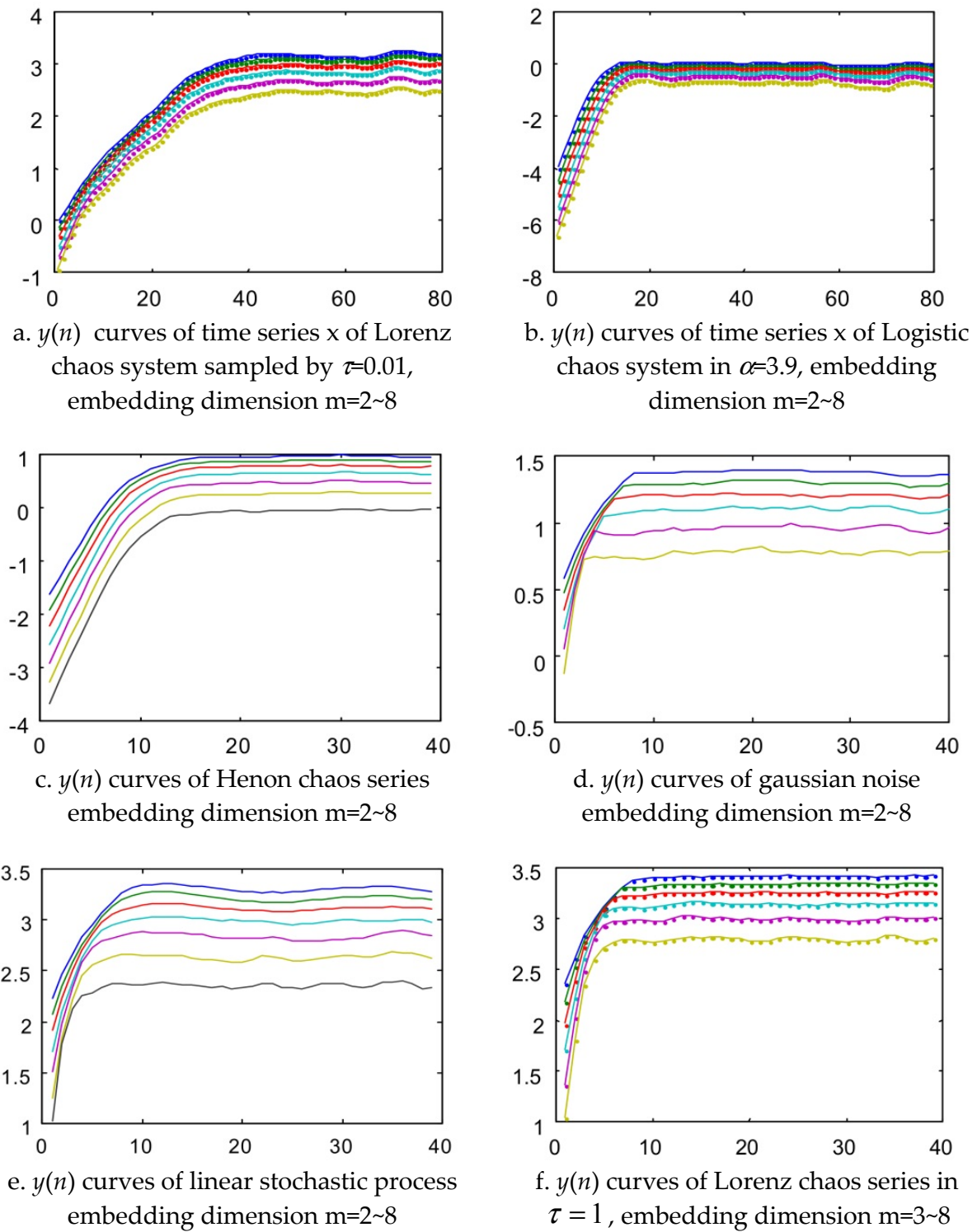


Figure 13. $y(n)$ curves of signals, abscissa is n , ordinate is $y(n)$

Then, the Lyapunov exponent can be given by using a least-squares fit to the “average” line:

$$y(n) = \frac{1}{\Delta t} \langle \ln d_i(n) \rangle \quad (39)$$

where $y(n) = n\langle\lambda\rangle + b$. The largest Lyapunov exponent λ is the slope of $y(n)$ in the above equation.

This method is deduced directly from the largest Lyapunov exponent definition. The accurate evaluation of λ depends on the full use of the data. In practice, the curve $y(n)$ will tend to saturation. The largest Lyapunov exponent λ is given by computing the slope of the linear part in the curve $y(n)$.

3.4.2. Chaos test based on largest Lyapunov exponent

In general, if the signal is chaotic, the slope of the curve $y(n)$ will be independent of the embedding dimension. Otherwise, if the signal is not chaotic, the slope of the curve $y(n)$ will depend on the embedding dimension. When the embedding dimension m is chosen from 2 to 8, the Lyapunov exponent of the curve $y(n)$ of the signal is shown in Figure 13. For a chaotic signal, a good illustration is given (see Figure 13a, b and c). The $y(n)$ curves are different from those of a non-chaotic signal (compare with Figure 13d and e). However, even for chaotic signals, the $y(n)$ curves are not always parallel. For example, in the case of undersampling ($\tau = 1$), the $y(n)$ curves of Lorenz chaotic time series are similar to those of linear stochastic process (compare with Figure 13e and f). In the literature[23], the $y(n)$ curves of the Ikeda chaotic time series are also not parallel.

3.4.3. The analysis of surface EMG signal based on the largest Lyapunov exponent

Figure 14 gives the curves of Lyapunov exponent $y(n)$ for the surface EMG signal. The $y(n)$ curves are not very parallel for the surface EMG signal. It is difficult to distinguish the curves of $y(n)$ for the surface EMG signal from those of Figure 13d, f and g. The surface EMG signal can not be determined as chaotic, or as stochastic. But it can be a high-dimensional system.

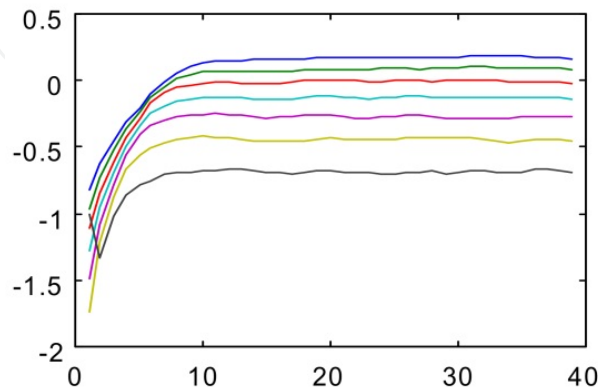


Figure 14. $y(n)$ curves of surface EMG signal, embedding dimension $m=2\sim 8$, abscissa is n , ordinate is $y(n)$

3.5. Principal component analysis

Broomhead and King[31] proposed the idea of singular system analysis that determines an appropriate embedding dimension d directly from the raw time series. It provides its convenience for the further analysis of the given system. Numerical experience, however, led several authors to express some doubts about reliability of singular system analysis in the attractor reconstruction[46-48]. Palus and Dvorak[37] explain why singular-value decomposition(SVD), the heart of the singular system analysis and by nature a linear method, may become misleading technique when it is used in nonlinear dynamics studies that reconstruction parameters are time-delay, embedding dimension (or embedding windows). For this, we propose a novel nonlinear analysis method based symplectic geometry, called symplectic principal component analysis(SPCA)[49].

3.5.1. Principle and algorithm of principal component analysis

Let a time series x_1, x_2, \dots, x_n be the measured signal by sampling interval t_s , n is the number of samples. According to Takens' embedding theorem, a trajectory matrix X can be given by time delay coordinates method, refer to Eq. 25($\tau=1$):

$$X = \begin{bmatrix} X_1^T \\ X_2^T \\ \vdots \\ X_m^T \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_d \\ x_2 & x_3 & \cdots & x_{d+1} \\ \vdots & \vdots & \cdots & \vdots \\ x_m & x_{m+1} & \cdots & x_n \end{bmatrix} \quad (40)$$

where d is embedding dimension. $m=n-d+1$ is the number of points in d -dimension reconstruction attractor, $X_i^T, i=1, \dots, m$, denotes a point in the attractor. For the matrix X , there are a $m \times d$ orthogonal matrix V and a $d \times d$ orthogonal matrix. The matrix X can be decomposed as follows.

$$X = VSU^T \quad (41)$$

where S is $d \times d$ diagonal matrix, whose elements are defined

$$S_{ij} = \delta_{ij} \sqrt{\sigma_i m}, \quad i, j = 1, 2, \dots, d \quad (42)$$

Since the matrix V is orthogonal, then

$$(V^T \cdot V)_{ij} = \delta_{ij} \quad (43)$$

Meanwhile, for the matrix U , there are

$$(U^T \cdot U)_{ij} = (U \cdot U^T)_{ij} = \delta_{ij} \quad (44)$$

In order to facilitate the calculation, Broomhead et al. applies the covariance matrix C of the matrix X to replace the matrix X . The details are as follows:

$$C = \frac{1}{m} \cdot X^T \cdot X \quad (45)$$

$$C_{ij} = \frac{1}{m} \sum_{k=1}^m x_{i+k-1} x_{j+k-1}, \quad i, j = 1, 2, \dots, d \quad (46)$$

Its values reflect the degree of correlation between the time delay coordinate variable i and j .

$$C = \frac{1}{m} \cdot X^T \cdot X = \frac{1}{m} \cdot U \cdot S \cdot S \cdot U^T = \frac{1}{m} \cdot U \cdot S^2 \cdot U^T \quad (47)$$

$$U^T \cdot C \cdot U = \frac{1}{m} \cdot (XU)^T \cdot XU = \frac{1}{m} \cdot S^2 \quad (48)$$

Let $Y=XU$, then:

$$U^T \cdot C \cdot U = \frac{1}{m} \cdot Y^T \cdot Y = \frac{1}{m} \cdot S^2 \quad (49)$$

where $U^T C U$ is the covariance matrix of the matrix Y . Its elements are zero, except that the diagonal elements are equal to σ . This means that the variables i and j of the matrix Y are independent. The coordinate system is orthogonal, which is constituted from the variables of the matrix Y after the above transformed. The σ is called the principal component or singular value in accordance with the order of the largest to the smallest. The orthogonal vector U_i corresponding to the principal component σ is called the principal axis. The principal component describes the distribution of the signal energy. That is, the value of the principal component reflects the projection of the signal energy in the corresponding principal axis. In the different principal axes, a distribution value is given as $\sigma_i / \sum_{i=1}^d \sigma_i$, where $\sum_{i=1}^d \sigma_i$ is the total energy of the signal. If $\sigma_{i+1} \approx \dots \approx \sigma_d$, the distribution values are called a noise floor. The distribution can be used to estimate the dimension of the dynamical system that generates a time series or to filter out the noise. Let \bar{U}_i be the principal axes corresponding to the principal components over the noise floor. Zero vector describes the principal axes corresponding to the noise floor. Thus, for $\sigma > \text{noise floor}$, a new coordinate transform matrix is made up of:

$$\bar{U} = [\bar{U}_1, \bar{U}_2, \dots, \bar{U}_i, 0, \dots, 0] \quad (50)$$

In order to filter out noise, the trajectory matrix X is first projected into the coordinate system U .

$$Y = XU \quad (51)$$

The variables in the matrix Y are independent. Then, the original coordinate system is updated by using the matrix Y :

$$\bar{X} = Y\bar{U}^T = (XU) \cdot \bar{U}^T \quad (52)$$

That is, \bar{X} is a new time delay coordinate system.

3.5.2. Influence of noise on the principal component spectrum of chaotic time series

Figure 15a shows the principal component spectrum of Logistic attractor from a Logistic chaotic time series without noise. The principal component spectrum has not a significant noise level. When the interior noise is Gaussian noise with zero mean and 0.001^2 variance, the principal component spectrum is given for Logistic attractor. Figure 15c and d give the principal component spectrum of Logistic attractor with the measurement noise $\sigma^2=0.001^2$ and $\sigma^2=0.8^2$, respectively. It can be seen that the Logistic attractor with the internal noise has the same principal component spectrum as the attractor without noise. The curves of the principal component spectrum are also slanting. The total energy is significantly distributed into each principal axes. The principal components are declining with the index i so that there is no noise floor. It is difficult to choose an appropriate embedding dimension d . For the larger measurement noise, the corresponding principal component spectrum of Logistic attractor slant into a floor area with increasing the embedding dimension. In the floor area, the principal components keep unchanged and do not decline with the index i , called noise floor. Broomhead and King[31] have suggested that this noise floor can be used to determine the embedding dimension and filter out noise from the data. The signal energy will be focused on the truncated principal components and the corresponding principal axes

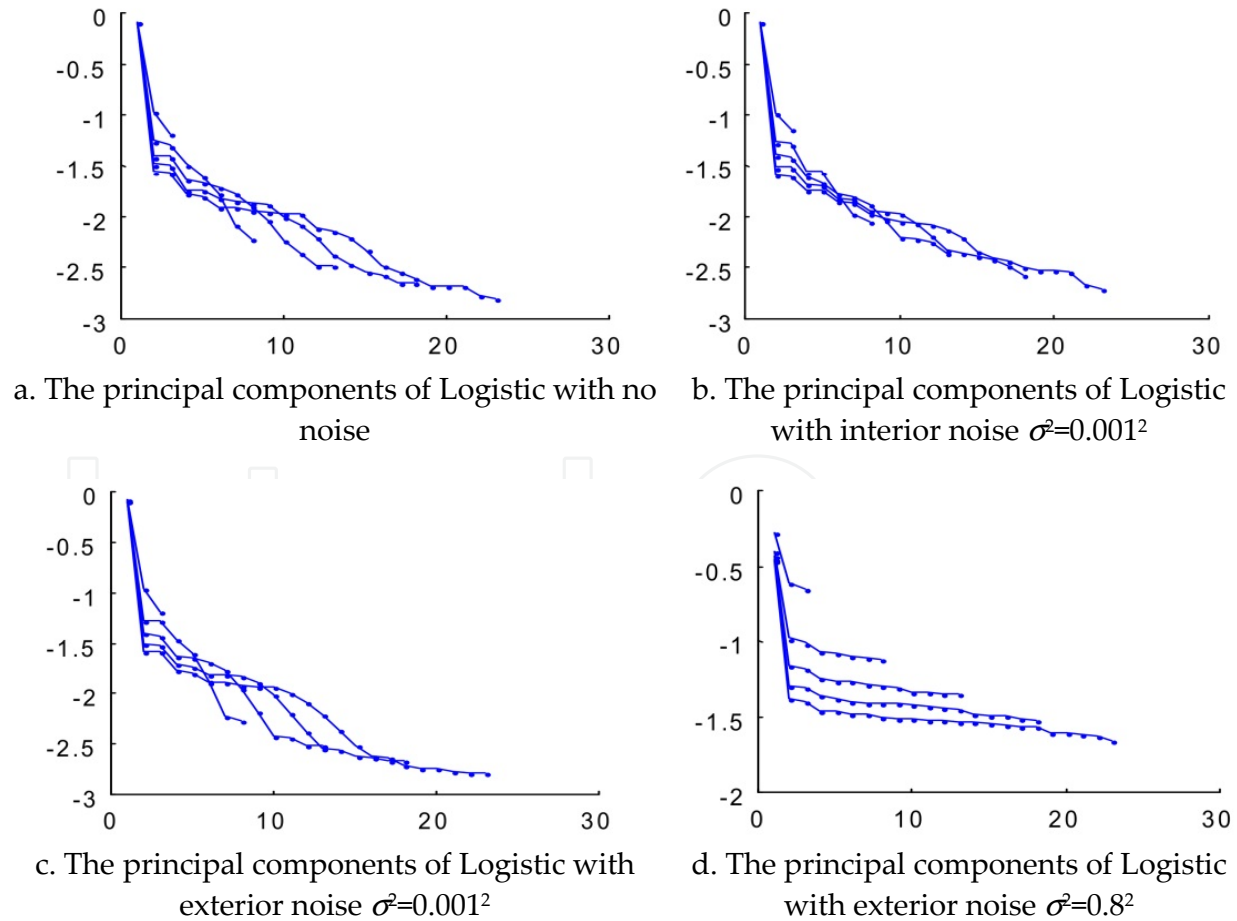


Figure 15. The principal component analysis of Logistic chaos series with different noises based on SVD, $d=3 : 2 : 23$, abscissa is d , ordinate is $\log(\sigma_i / \text{tr}(\sigma_i))$

when the principal components above noise floor are only held. The number of the principal components above the noise floor is the optimal embedding dimension.

Besides, the new coordinate system corresponding to the principal axes can eliminate the noise floor to reduce the noise from the data. However, the truncated position of the principal components depends on the signal-noise-ratio, especially for the measurement noise. The principal components of the chaotic time series based on SVD spectrum more easily subject to the measurement noise so that the embedding dimension estimation is directly affected. For the smaller noise, there is the more number of principal components above the noise floor. For the larger noise, the number of the corresponding principal components will be reduced. Here, the above calculation accuracy is $2.2204e-016$, which does not consider the numerical calculation error.

3.5.3. Influence of sampling interval on the principal component spectrum of chaotic time series[49]

The Lorenz chaotic system is considered to give the state variable x in order to study the influence of sampling interval on the principal component spectrum. The principal component spectrum slant and have no floor for the chaotic time series x with $\tau=0.005$ (see Fig. 16a). When $\tau=0.1$ (see Fig. 16b), the principal component spectrum are basically similar

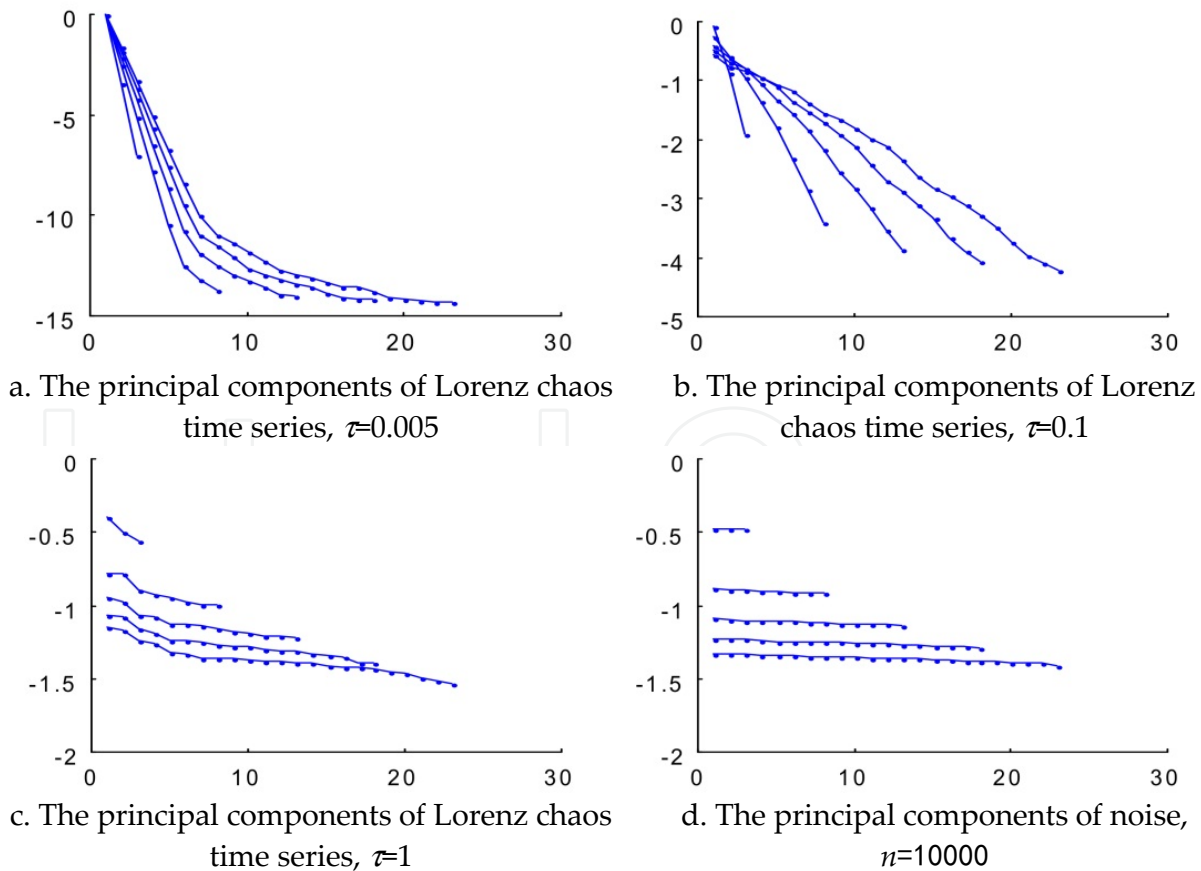


Figure 16. The principal component analysis of gaussian noise and Lorenz chaos time series by different sampling intervals based on SVD, $d=3 : 2 : 23$, abscissa is d , ordinate is $\log(\sigma_i / \text{tr}(\sigma_i))$

to those in the Figure 16a. When $\tau=1$, each line is separated from each other and tends to horizontal line in the case of different embedding dimensions(see Fig. 16c). It shows that the distribution of the total energy has little difference in each principal axis, like the Gaussian noise (see Fig. 16d). For the Gaussian noise, its principal component spectrum curves are horizontal lines, where $N=10000$. It shows that every principal component is equal to each other. The energy distributes into every principal axis averagely. Therefore, it can be seen that sampling interval affects the determination of embedding dimension. When the sampling interval is not undersampling, the determination of embedding dimension depends the amount of signal-noise-ratio. In the case of undersampling, the chaotic time series is similar to noise so that the embedding dimension seems to be estimated as 1.

3.6. Symplectic principal component analysis

The symplectic geometry is a kind of phase space geometry. Its nature is nonlinear. It can describe the system structure, especially nonlinear structure, very well. It has been used to study various nonlinear dynamical systems[50-52] since Feng Kang[53] has proposed a symplectic algorithm for solving symplectic differential. However, from the view of data analysis, few literatures have employed symplectic geometry theory to explore the dynamics of the system. Our previous works have proposed the estimation of the embedding dimension based on symplectic geometry from a time series[49, 54-56]. Subsequently, Niu et al. have used our method to evaluate sprinter's surface EMG signals[57]. Xie et al[58] have proposed a kind of symplectic geometry spectra based on our work. Subsequently, we show that SPCA can well represent chaotic time series and reduce noise in chaotic data[59, 60].

In SPCA, a fundamental step is to build the multidimensional structure (attractor) in symplectic geometry space. Here, in terms of Taken's embedding theorem, we first construct an attractor in phase space, i.e. the trajectory matrix X from a time series. That is, for a measured data (the observable of the system under study) x_1, x_2, \dots, x_n recorded with sampling interval t_s , the corresponding d -dimension reconstruction attractor, $X_{m \times d}$ can be given (refer to Eq.40). Then we describe the symplectic principal component analysis (SPCA) based on symplectic geometry theory and give its corresponding algorithm.

3.7. Symplectic principal component method

SPCA is a kind of PCA approaches based on symplectic geometry. Its idea is to map the investigated complex system in symplectic space and elucidate the dominant features underlying the measured data. The first few larger components capture the main relationship between the variables in symplectic space. The remaining components are composed of the less important components or noise in the measured data. In symplectic space, the used geometry is called symplectic geometry. Different from Euclid geometry, symplectic geometry is the even dimensional geometry with a special symplectic structure. It is dependent on a bilinear antisymmetric nonsingular cross product—symplectic cross product:

$$[x, y] = \langle x, Jy \rangle \quad (53)$$

$$\text{where, } J = J_{2n} = \begin{bmatrix} 0 & +I_n \\ -I_n & 0 \end{bmatrix} \quad (54)$$

When $n = 1$, $x = [x_1, x_2]$, $y = [y_1, y_2]$,

$$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (55)$$

$$\begin{aligned} [x, y] &= [x_1 \ x_2] J \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \end{aligned} \quad (56)$$

The measurement of symplectic space is area scale. In symplectic space, the length of arbitrary vectors always equals zero and without signification, and there is the concept of orthogonal cross-course. In symplectic geometry, the symplectic transform is the nonlinear transform in essence, which is also called canonical transform, since it has measure preserving characteristics and can keep the natural properties of the original data unchanged. It is fit for nonlinear dynamics systems.

The symplectic principal components are given by symplectic similar transform. It is similar to SVD-based PCA. The corresponding eigenvalues can be obtained by symplectic QR method. Here, we first construct the autocorrelation matrix $A_{d \times d}$ of the trajectory matrix $X_{m \times d}$. Then the matrix A can be transformed as a Hamilton matrix M in symplectic space.

Definition 1 Let S is a matrix, if $JSJ^{-1} = S^{-*}$, then S is a symplectic matrix.

Definition 2 Let H is a matrix, if $JHJ^{-1} = -H^*$, then H is a Hamilton matrix.

Theorem 1 Any $d \times d$ matrix can be made into a Hamilton matrix. Let a matrix as A , so $\begin{pmatrix} A & 0 \\ 0 & -A^T \end{pmatrix}$ is a Hamilton matrix. (Proof refers to appendix A)

Theorem 2 Hamilton matrix M keeps unchanged at symplectic similar transform. (Proof refers to appendix A)

Theorem 3 Let $M \in C^{2d \times 2d}$ as Hamilton matrix, so e^M is symplectic matrix.

Theorem 4 Let $S \in C^{2d \times 2d}$ as symplectic matrix, there is $S = QR$, where Q is symplectic unitary matrix, R is upper triangle matrix.

Theorem 5 The product of symplectic matrixes is also a symplectic matrix. (Proof refers to appendix A)

Theorem 6 Suppose Household matrix H is :

$$H = H(k, \omega) = \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix} \quad (57)$$

where $P = I_n - \frac{2\varpi\varpi^*}{\varpi^*\varpi}$, $\varpi = (0, \dots, 0; \omega_k, \dots, \omega_d)^T \neq 0$

so, H is symplectic unitary matrix. ϖ^* is ϖ conjugate transposition. (Proof refers to appendix A)

For Hamilton matrix M , its eigenvalues can be given by symplectic similar transform and the primary $2d$ dimension space can be transformed into d dimension space to resolve^[17-19], as follows:

i. Let $N = M^2$

$$M^2 = \begin{bmatrix} A^T & G \\ F & -A \end{bmatrix}^2 \quad (58)$$

ii. Construct a symplectic matrix Q ,

$$Q^T N Q = \begin{bmatrix} B & R \\ 0 & B^T \end{bmatrix} \quad (59)$$

where B is up Hessenberg matrix ($b_{ij}=0, i>j+1$). The matrix Q may be a symplectic Household matrix H . If the matrix M is a real symmetry matrix, M can be considered as N . Then one can get an upper Hessenberg matrix (referred to equ. 13), namely,

$$\begin{aligned} H M H' &= \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -A' \end{pmatrix} \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix}' \\ &= \begin{pmatrix} P A P' & 0 \\ 0 & -P A' P' \end{pmatrix} \\ &= \begin{pmatrix} B & 0 \\ 0 & -B' \end{pmatrix} \end{aligned} \quad (60)$$

where H is the symplectic Householder matrix.

iii. Calculate eigenvalues $\lambda(B) = \{\mu_1, \mu_2, \dots, \mu_d\}$ by using symplectic QR decomposition method; if M is a real symmetry matrix, the eigenvalues of A is equal to those of B :

$$\mu = \lambda(B) = \lambda(A) \quad (61)$$

$$\lambda(A) = \mathcal{K}(X) \quad (62)$$

iv. These eigenvalues $\mu = \{\mu_1, \mu_2, \dots, \mu_d\}$ are sorted by descending order, that is

$$\mu_1 > \mu_2 > \dots > \mu_k >> \mu_{k+1} \geq \dots \geq \mu_d \quad (63)$$

Thus the calculation of $2d$ dimension space is transformed into that of that of d dimension space. The μ is the symplectic principal component spectrums of A with relevant symplectic

orthonormal bases. In the so-called noise floor, values of μ_i , $i = k+1, \dots, d$, reflect the noise level in the data[49, 55]. The corresponding matrix Q denotes symplectic eigenvectors of A .

3.7.1. Proposed algorithm of symplectic principal component method

For a measured data x_1, x_2, \dots, x_n , our proposed algorithm consists of the following steps:

1. Reconstruct the attractor $X_{m \times d}$ from the measured time series, where d is the embedding dimension of the matrix X , and $m = n-d+1$.
2. Remove the mean values X_{mean} of each row of the matrix X .
3. Build the real $d \times d$ symmetry matrix A , that is,

$$A = (X - X_{mean})'(X - X_{mean}) \quad (64)$$

Here, d should be larger than the dimension of the system in terms of Taken's embedding theorem.

4. Calculate the symplectic principal components of the matrix A by QR decomposition, and choose the Householder matrix H instead of the transform matrix Q . It is easy to prove that H is a symplectic unitary matrix (Proof refers to appendix A) and H can be constructed from real matrix (refer to appendix B).
5. Construct the corresponding principal eigenvalue matrix W according to the number k of the chosen symplectic principal components of the matrix A , where $W \subseteq Q$. That is, when $k=d$, $W=Q$, otherwise $W \subset Q$. In use, k can be chosen according to Eq.63.
6. Get the transformed coefficients $S = \{S_1, S_2, \dots, S_m\}$, where

$$S_i = W' X_i, \quad i = 1, \dots, m \quad (65)$$

7. Reestimate the X_s from S ,

$$X_{si} = W S_i \quad (66)$$

Then the reestimation data $x_{s1}, x_{s2}, \dots, x_{sm}$ can be given.

8. For the noisy time series, the first estimation of data is usually not good. Here, one can go back to the step (6) and let $X_i = X_s$ in Eq.(65) to do step (6) and (7) again. Generally, the second estimated data will be better than the first estimated data.

Besides, it is necessary to note that for the clean time series, the step (8) is unnecessary to handle.

3.7.2. Performance evaluation

SPCA, like PCA, can not only represent the original data by capturing the relationship between the variables, but also reduce the contribution of errors in the original data. Here, the performance analysis of SPCA is studied from the two views, i.e. representation of chaotic signals and noise reduction in chaotic signals.

Representation of chaotic signals

We first show that for the clean chaotic time series, SPCA can perfectly reconstruct the original data in a high-dimensional space. We first embed the original time series to a phase space. Considering the dimension of the Lorenz system(see Eq. 30) is 3, d of the matrix A is chosen as 8 in our SPCA analysis. To quantify the difference between the original data and the SPCA-filtered data, we employ the root-mean-square error (RMSE) as a measure:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [x(i) - \hat{x}(i)]^2} \quad (67)$$

where $x(i)$ and $\hat{x}(i)$ are the original data and estimated data, respectively.

When $k = d$, the RMSE values are lower than 10^{-14} (see Figure 17). In Figure 17, the original data are generated by Eq. 30. The estimated data is obtained by SPCA with $k=d$. The results show that the SPCA method is better than the PCA. Since the real systems are usually unknown, it is necessary to study the effect of sampling time, data length, and noise to the SPCA approach. From the Figure 17 and 18, we can see that the sampling time and data length have less effect on SPCA method in the case of free-noise.

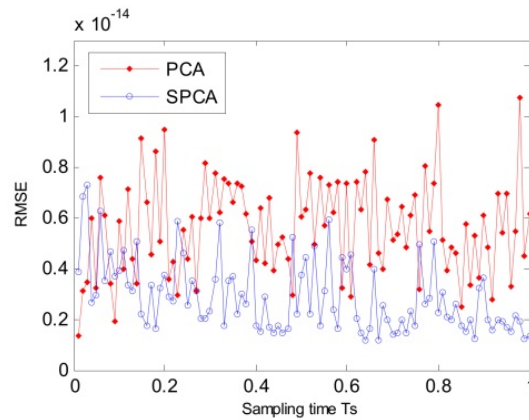


Figure 17. (Color online) RMSE vs. Sampling time curves for the SPCA and PCA.

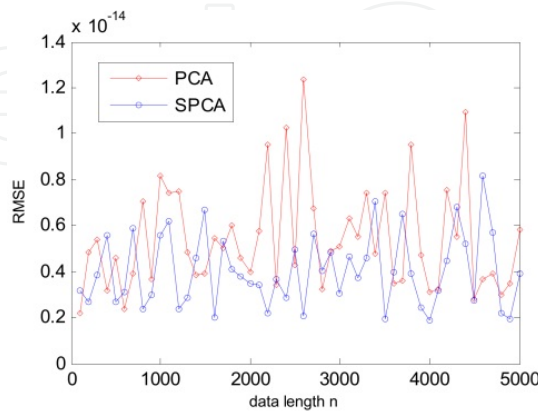


Figure 18. (Color online) RMSE vs. data length curves for the SPCA and PCA.

For analyzing noisy data, we use the percentage of principal components (PCs) to study the occupancy rate of each PC in order to reduce noise. The percentage of PCs is defined by

$$P_i = \frac{\mu_i}{\sum_{i=1}^d \mu_i} \times 100\% \quad (68)$$

where d is the embedding dimension, μ_i is the i -th principal component value. From the Figure 19, we find that the first largest symplectic principal component (SPC) of the SPCA is a little larger than that of the PCA. It is almost possessed of all the proportion of the symplectic principal components. This shows that it is feasible for the SPCA to study the principal component analysis of time series.

Next, we study the reduced space spanned by a few largest symplectic principal components (SPCs) to estimate the chaotic Lorenz time series (see Fig. 20). In Figure 20, the data x is given with a sampling time of 0.01 from chaotic Lorenz system. The estimated data is calculated by the first three largest SPCs. The average error and standard deviation between the original data and the estimated data is $-6.55e-16$ and $1.03e-2$, respectively. The

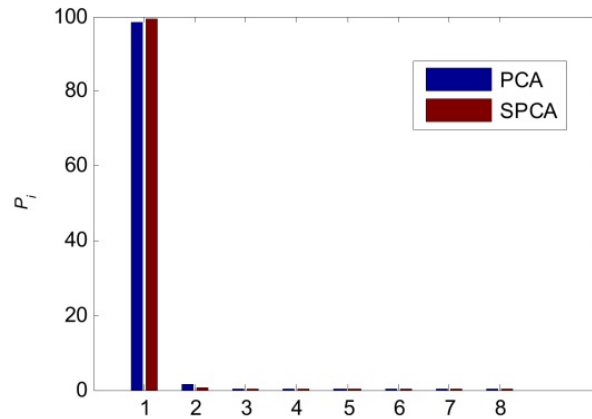


Figure 19. (Color online) The percentage of principal components for the SPCA and PCA.

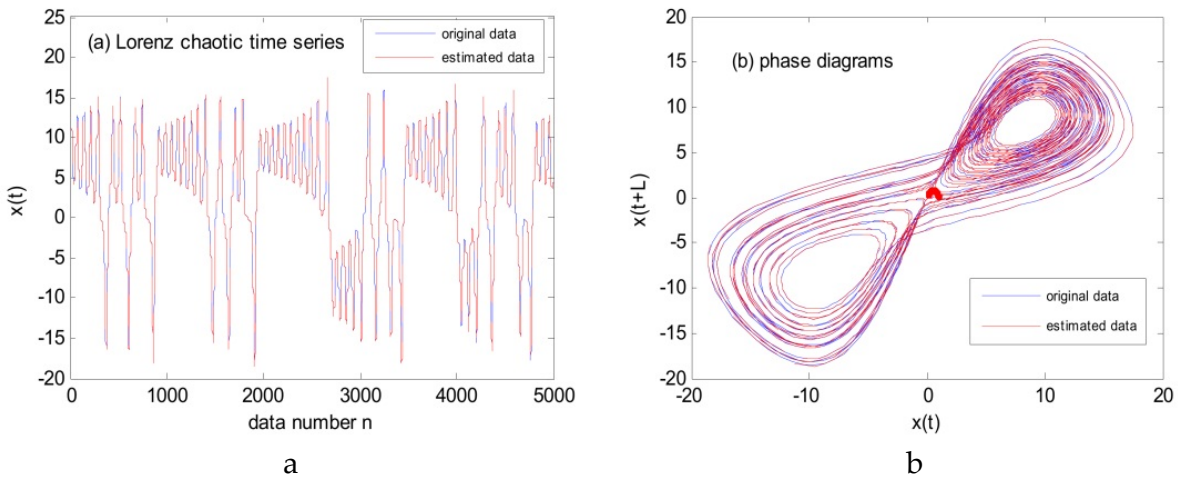


Figure 20. (Colour online) Chaotic signal reconstructed by the proposed SPCA algorithm with $k=3$, where (a) the time series of the original Lorenz data x without noise and the estimated data; (b) phase diagrams with $L=11$ for the original Lorenz data x without noise and the estimated data. The sampling time $t_s = 0.01$.

estimated data is very close to the original data not only in time domain (see Figure 20a) but also in phase space (see Figure 20b). We further explore the effect of sampling time in different number of PCs. When the PCs number $k=1$ and $k=7$, respectively, the SPCA and PCA give the change of RMSE values with the sampling time in Figure 21. We can see that the RMSE values of the SPCA are smaller than those of the PCA. The sampling time has less impact on the SPCA than the PCA. In the case of $k=7$, the data length has also less effect on the SPCA than the PCA (see Fig. 22).

Comparing with PCA, the results of SPCA are better in the above Figures. We can see that the SPCA method keep the essential dynamical character of the primary time series generated by chaotic continuous systems. These indicate that the SPCA can reflect intrinsic nonlinear characteristics of the original time series. Moreover, the SPCA can elucidate the dominant features underlying the observed data. This will help to retrieve dominant patterns from the noisy data. For this, we study the feasibility of the proposed algorithm to reduce noise by using the noisy chaotic Lorenz data.

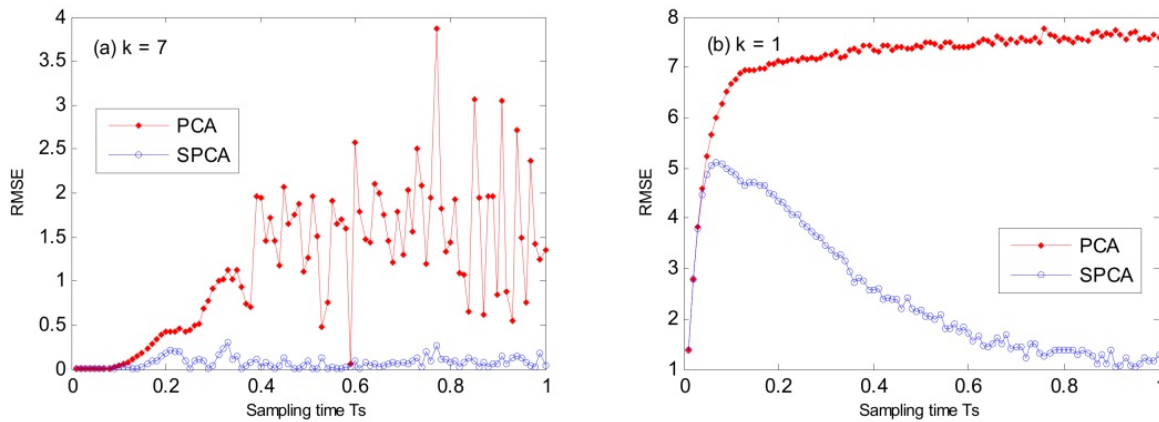


Figure 21. The RMSE values vs. the sampling time for the SPCA and PCA, where (a) the PCs number $k=7$; (b) $k=1$.

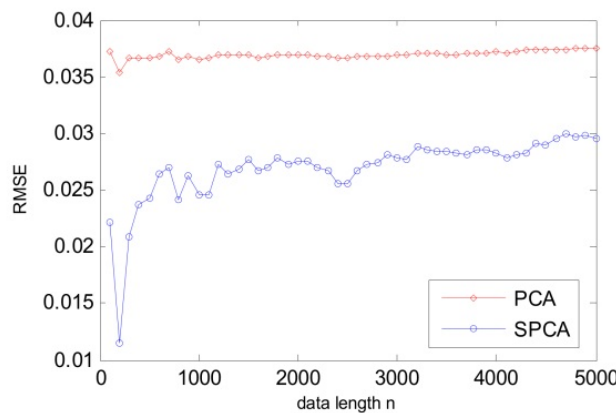


Figure 22. The RMSE vs. the data length for the SPCA and PCA, where $k=7$. The sampling time is 0.1.

Noise reduction in chaotic signals

For the noisy Lorenz data x , the phase diagrams of the noisy and clean data are given in Figure 23a and 23b. The clean data is the chaotic Lorenz data x with noise-free (see Eq. 30).

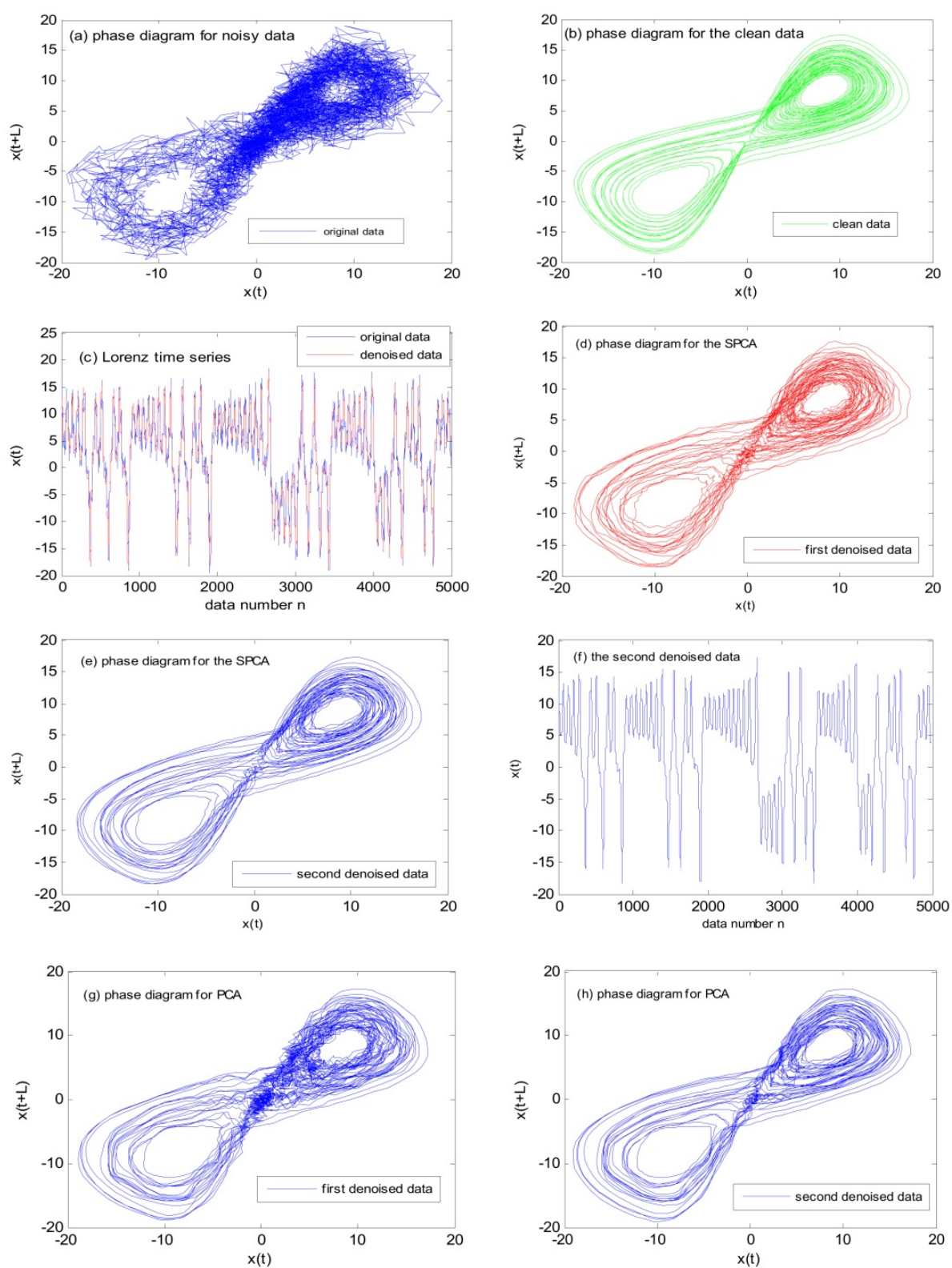


Figure 23. The noise reduction analysis of the proposed SPCA algorithm and PCA for the noisy Lorenz time series, where $L=11$.

The noisy data is the chaotic Lorenz data x with Gaussian white noise of zero mean and one variance (see Eq. 30). The sampling time is 0.01. The time delay L is 11 in Figure 23. It is obvious that noise is very strong. The first denoised data is obtained in terms of the proposed SPCA algorithm (see Figure 23c- f). Here, we first build an attractor X with the embedding dimension of 8. Then the transform matrix W is constructed when $k=1$. The first denoised data is generated by Eq.(65) and (66). In Figure 23c, the first denoised data is compared with the noisy Lorenz data x from the view of time field. Figure 23d shows the corresponding phase diagram of the first denoised data. Compared with Fig. 23a, the first denoised data can basically give the structure of the original system. In order to obtain better results, this denoised data is reduced noise again by the step (8). We can see that after the second noise reduction, the results are greatly improved in Fig. 23e and 23f, respectively. The curves of the second denoised data are better than those of the first denoised data whether in time domain or in phase space by contrast with Fig. 23c and 23d. Figure 23g shows that the PCA technique gives the first denoised result. We refer to our algorithm to deal with the first denoised data again by the PCA (see Figure 23h). Some of noise has been further reduced but the curve of PCA is not better than that of SPCA in Figure 23e. The reason is that the PCA is a linear method indeed. When nonlinear structures have to be considered, it can be misleading, especially in the case of a large sampling time (see Figure 24). The used program code of the PCA comes from the TISEAN tools (<http://www.mpi-pks-dresden.mpg.de/~tisean>).

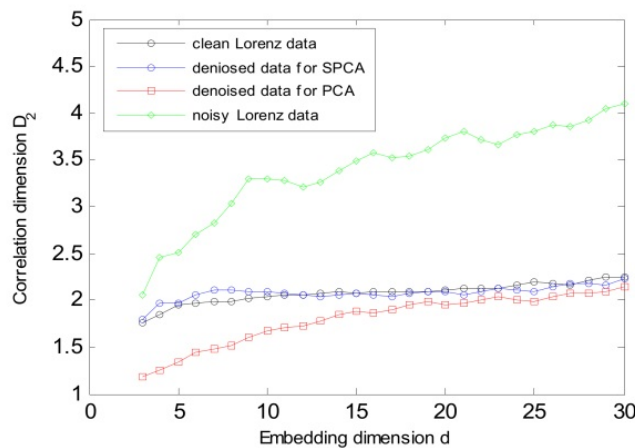


Figure 24. (Color online) D_2 vs. embedding dimension d

Figure 24 shows the variation of correlation dimension D_2 with embedding dimension d in the sampling time of 0.1 for the clean, noisy, and denoised Lorenz data. We can observe that for the clean and SPCA denoised data, the trend of the curves tends to smooth in the vicinity of 2. For the noisy data, the trend of the curve is constantly increasing and has no platform. For the PCA denoised data, the trend of the curve is also increasing and trends to a platform with 2. However, this platform is smaller than that of SPCA. It is less effective than the SPCA algorithm. This indicates that it is difficult for the PCA to describe the nonlinear structure of a system, because the correlation dimension D_2 manifests nonlinear properties of chaotic systems. Here, the correlation dimension D_2 is estimated by the Grassberger-Procaccia's algorithm[33, 40].

3.7.3. Estimation of embedding dimension based on symplectic geometry

In terms of Eq. 63, the values of μ_i , $i=k+1, \dots, d$, are far smaller than μ_k . These values form a noise floor. Therefore, the embedding dimension of the reconstruction system can be determined by the noise floor. Here, the noise and nonlinear time series are used to investigate the feasibility of the embedding dimension estimation based on symplectic geometry.

For noise (which is generally regarded as Gauss white noise with mean value 0 and variance 1 in practical systems), symplectic geometry spectrums of this noise give the even distribution of its total energy (see Fig. 25a). From this figure, we can see that the symplectic geometry spectrums of noise can reflect the characteristic of noise very well when $N=1000$. This shows SG method can reflect noise level in the condition of short data length. For the time series of state variable x in Logistic chaos system without noise interference, the symplectic geometry spectrums (see Fig. 25b) are slant in the beginning then turn into plane area with the increase of index i . In other words, the distribution of total energy on the different axes is obviously different and with increasing the embedding dimension, the slants of symplectic geometry spectrums transit into noise floor. So one can determine embedding dimension from the number of symplectic geometry spectrums over noise floor, in which its determining criterion is similar to that in [37]. From Fig. 25b, the embedding dimension of Logistic chaotic time series can be estimated at 4 because the symplectic geometry spectrums begin to turn into noise floor at index 5. In a similar way, for Lorenz chaos time series without noise, when sampling interval $\tau=0.005$, the embedding dimension can be estimated at 6 (see Fig. 25c).

Comparison of the results of our method (see Fig. 25b and 25c) and the results of SVD method (see Fig. 25d and 25e) shows that in SG method, the position of the noise floor is determined by the intrinsic dynamical structure of the nonlinear dynamic system rather than the numerical accuracy of the input data and the computation precision, but in SVD method, the noise floor was determined reversely[8, 37, 61].

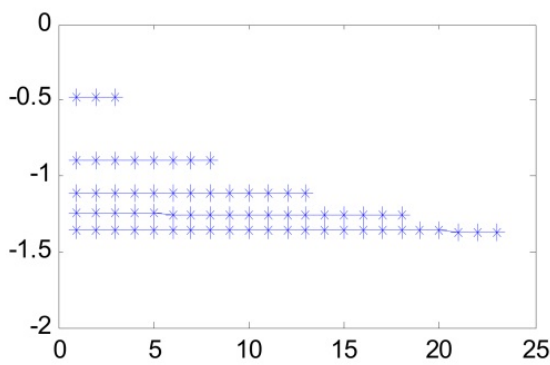
In a word, the numerical experiments discuss that for the nonlinear dynamic systems, SG method can give the appropriate embedding dimension from their time series but SVD method cannot. So SG method is fit to deal with nonlinear systems.

3.7.4. Robustness of the embedding dimension estimation based on symplectic geometry

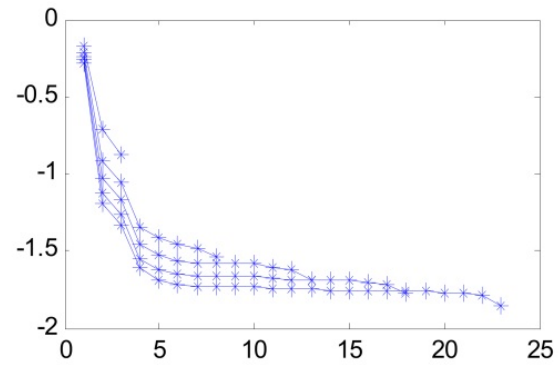
It is well known that the recent methods about embedding dimension are almost more or less subjective, or are affected by changes of the data length, noise, time lag, or sampling time, etc. Here, it is necessary that the robustness of the SG method is studied.

The effect of data length

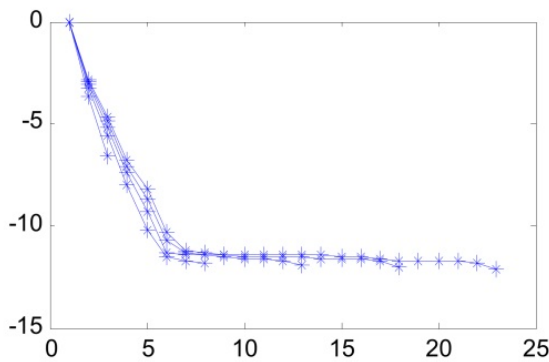
In order to avoid the effect of the characteristics of the nonlinear system, this paper only considers and uses the noise to analyze the effect of data length. For Gauss white noise with mean value 0 and variance 1, when $N=1000$, the SG method can give better results than the SVD method (see Fig. 26a) because the total energy is distributed equably (see Fig. 25a).



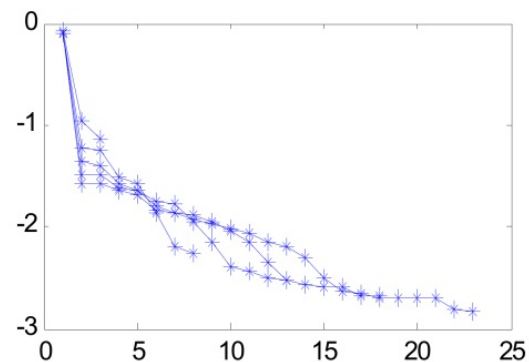
a. The symplectic geometry spectrums of Gauss white noise with mean value 0 and variance 1



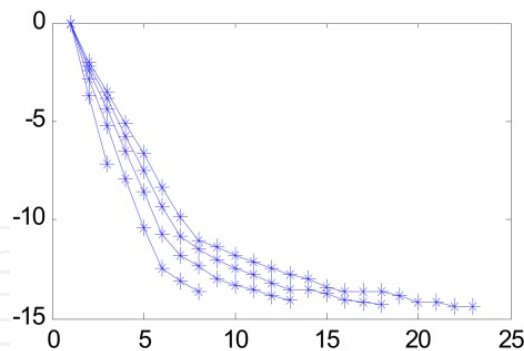
b. The symplectic geometry spectrums of Logistic chaotic series with no noise



c. The symplectic geometry spectrums of Lorenz chaotic series with no noise, $\tau=0.005$



d. The SVD principal components of Logistic chaotic series with no noise



e. The SVD principal components of Lorenz chaotic series with no noise, $\tau=0.005$

Figure 25. The study of embedding dimension based on symplectic geometry algorithm, $N=1000$, $d=3, 8, 13, 18, 23$, abscissa is d , ordinate is $\log(\mu_i / \text{tr}(\mu_i))$

And yet when N is rather large, e.g. $N=10000$, the SVD method can just have the similar results (see Fig. 26b) with Fig. 25a. These show that the SG method is more robust to changes of the data length than the SVD method. Then the SG method is fitter to the analysis of short time series.

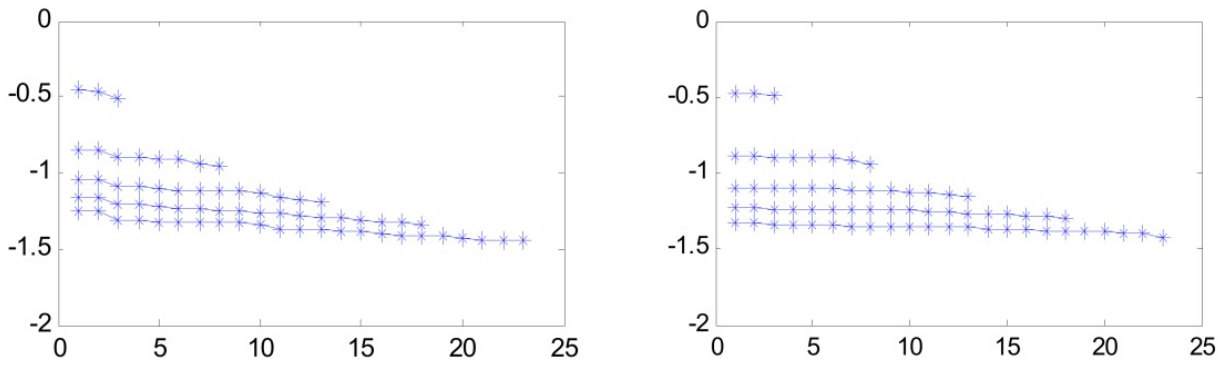
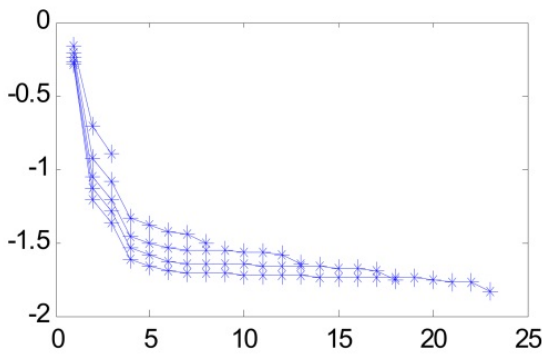
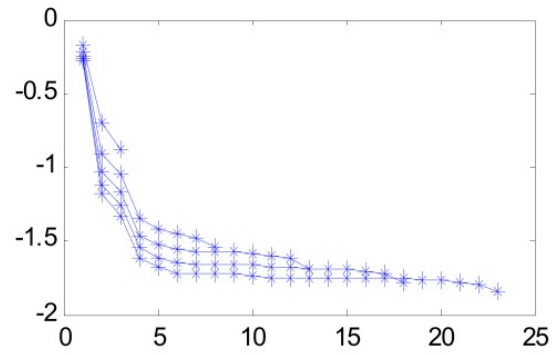


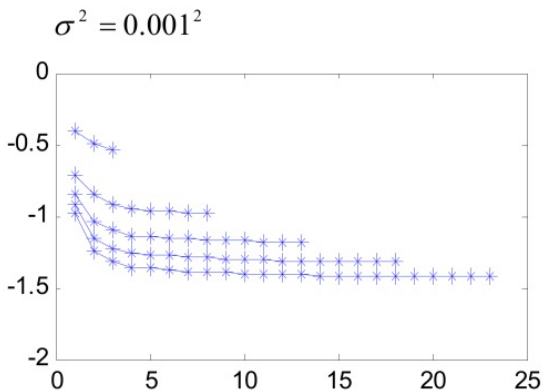
Figure 26. The analysis of SVD principal components of noise with different data length, $d=3, 8, 13, 18, 23$, abscissa is d , ordinate is $\log(\mu_i / \text{tr}(\mu_i))$



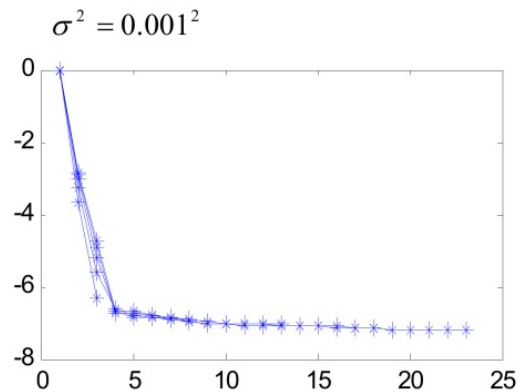
a. The symplectic geometry spectra of Logistic chaos series with interior noise, $\sigma^2 = 0.001^2$



b. The symplectic geometry spectra of Logistic chaos series with exterior noise, $\sigma^2 = 0.001^2$



c. The symplectic geometry spectra of Logistic chaos series with exterior noise, $\sigma^2 = 0.8^2$



d. The symplectic geometry spectra of Lorenz chaos series with exterior noise, $\sigma^2 = 0.01^2$

Figure 27. The study of symplectic geometry spectrum analysis in different noises, $N=1000$, $d=3, 8, 13, 18, 23$, abscissa is d , ordinate is $\log(\mu_i / \text{tr}(\mu_i))$

The effect of noise

At present, there are many estimators of appropriate embedding dimension, but it has gradually been realized that such estimators are useful only for low-dimensional noise-free

systems; such systems, however, seem hardly to occur in the real life. Therefore, this paper studies the robustness of the SG method under noise. For the signal obtained from the real system, it is always contaminated by noise (inner noise or/and outer noise). Although contaminated by inner or/and outer noise, the embedding dimension of Logistic system can always be noted at 4 by using the SG method because the noise floor begins at the embedding dimension 5 (see Fig. 27a and 27b). These show either inner noise or outer noise has little impact on the symplectic geometry spectrums. On the further increase of noise, the position of noise floor is obviously raised from the Figure 27c, but the appropriate embedding dimension 2 can still be obtained. In the similar way, for Lorenz chaos time series without and with noise, when sampling interval $\tau=0.005$, the embedding dimension is 6 without noise and 3 with noise, respectively (see Fig. 25c and Fig. 27d). These results show that the SG method is useful for Lorenz system with noise, too. Meanwhile, we find that the SG method can obtain the results similar to nonlinear high singular spectrum algorithm[62]. Thus, it further shows that the SG method can reflect intrinsic nonlinear characteristics of the raw data.

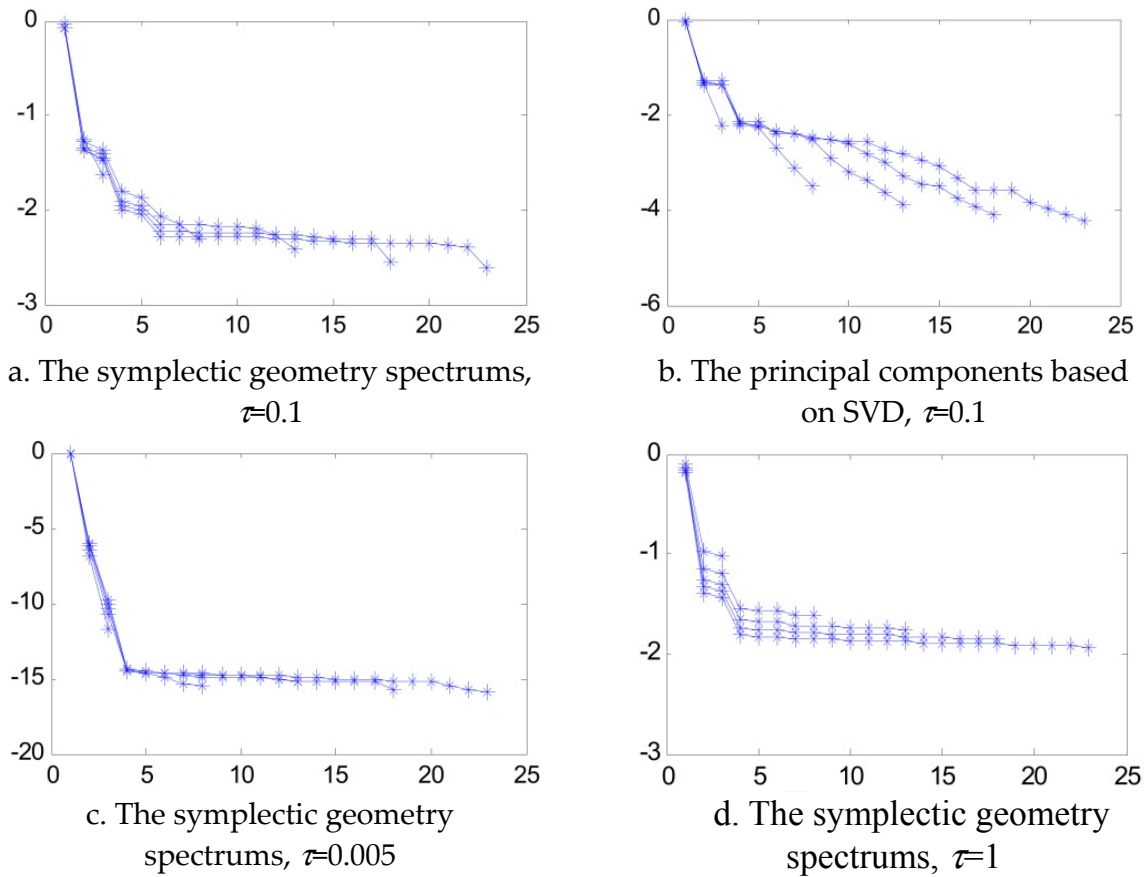


Figure 28. The symplectic geometry spectrum analysis of Lorenz chaos series by different sampling intervals, $N=1000$, $d=3, 8, 13, 18, 23$, abscissa is d , ordinate is $\log(\mu_i / \text{tr}(\mu_i))$

The effect of sampling interval

For the changes of the sampling interval from $\tau=0.005$ to $\tau=0.1$, this paper finds that the embedding dimension can be estimated at 6 from the corresponding symplectic geometry

spectrums of Lorenz chaos time series (see Fig. 25c and Fig. 28a), although the position of noise floor is constantly driven up. However, in the same condition, SVD method cannot give the appropriate embedding dimension (see Fig. 25e and 28b), the results of which are similar to the results of the literature[61]. Besides, no matter the sampling interval is over sampling or under sampling, SG method can always give the appropriate embedding dimension d of Lorenz chaos time series (see Fig. 28c and 28d) because the correlation dimension m of Lorenz system is 2.07, in general, if $d > m$, d is viable.

3.7.5. Analysis of the surface EMG signal based on symplectic geometry

For the action surface EMG signal (ASEMG) collected from a normal person, SVD method cannot give its appropriate embedding dimension (see Fig. 29a). The method based on correlation theory can do it but costs much time for computation. Here, SG method can fast obtain its embedding dimension. Figure 29b is the symplectic geometry spectrums of action surface EMG signal. The embedding dimension can be chosen as 6, which is the same as that of correlation dimension analysis[3]. This further shows that the SG method has stronger practicability for the small sets of experiment data.

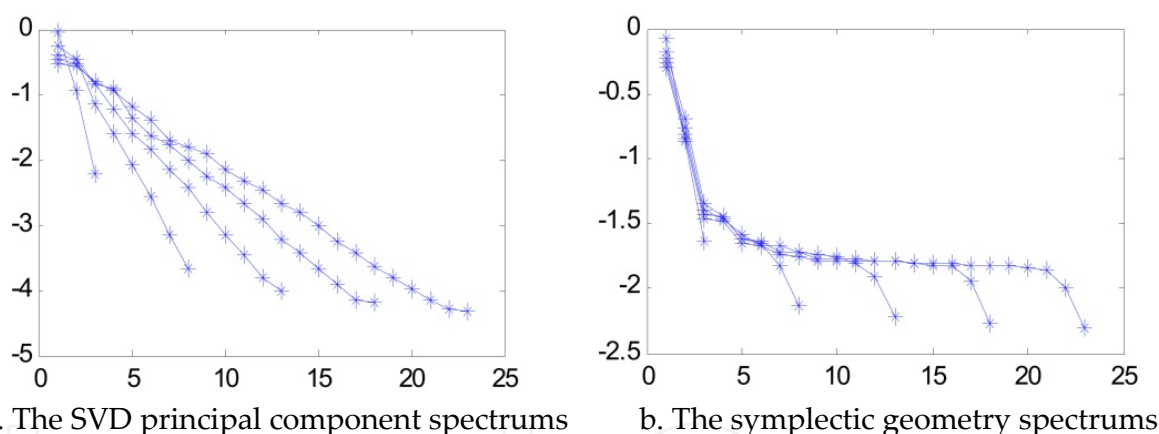


Figure 29. The analysis of action surface EMG signal, $d=3, 8, 13, 18, 23$, abscissa is d , ordinate is $\log(\mu_i / \text{tr}(\mu_i))$

4. Study of nonlinear dynamical systems based on multifractal theory

Fractal is a kind of geometry structures that have similarity in structure, form or function between the local and the whole. In nature, almost every object is very complex and performs a self-organization phenomenon that is a spatiotemporal structure or state phenomenon by forming spontaneously. From the view of geometry structure, this object has its own self-similarity properties in many parts, called a multifractal system. This structure can often be characterized by a set of coefficients, such as multifractal dimension, wavelet multifractal energy dimension. The multifractal theory reflects the complexity and richness of the nature in essence.

4.1. Self-affine fractal[63-65]

Definition 1 Let the mapping $S: R^n \rightarrow R^n$. S is defined by

$$S(x) = T(x) + b \quad (69)$$

where T is a linear transformation on R^n . b is a vector in R^n . Thus, S is a combination of a translation, rotation, dilation and, perhaps, a reflection, called an affine mapping. Unlike similarities, affine mappings contract with differing ratios in different directions.

Theorem 1 Consider the iterated function system given by the contractions $\{S_1, \dots, S_m\}$ on $D \subset R^n$, so that

$$|S_i(x) - S_i(y)| \leq C_i |x - y|, \quad \forall x, y \in D \quad (70)$$

with $\exists C_i < 1$ for each i . Then there is a unique attractor F , i.e. a non-empty compact set such that

$$F = \bigcup_{i=1}^m S_i(F) \quad (71)$$

Moreover, if we define a transformation S on the class ϕ of non-empty compact sets by

$$S(E) = \bigcup_{i=1}^m S_i(E) \quad (72)$$

For $E \in \phi$, and write S^k for the k th iterate of S (so $S^0(E) = E$ and $S^k(E) = S(S^{k-1}(E))$ for $k \geq 1$), then

$$F = \bigcap_{k=1}^{\infty} S^k(E) \quad (73)$$

for every set $E \in \phi$ such that $S_i(E) \subset E$ for all i .

If an IFS consists of affine contractions $\{S_1, \dots, S_m\}$ on R^n , the attractor F guaranteed by Theorem 1 is termed a self-affine set.

Since self-affine time series have a power-law dependence of the power-spectral density function on frequency, self-affine time series exhibit long-range persistence. For a practical data, one can use the relationship of power spectrum and frequency to determine if the data has the self-affine fractal characteristic.

4.2. Spectrum analysis[65, 66]

Let a time series be $x(t)$, $t \in [0, T]$. Its spectrum is given by

$$X(f, T) = \int_0^T x(t) e^{2\pi i f t} dt \quad (74)$$

where f is frequency. The power spectrum of f is defined by

$$S(f) = \frac{1}{T} |X(f, T)|^2 \quad (75)$$

If the power spectrum obeys a power law

$$S(f) = K \cdot f^{-\beta} \quad (76)$$

for large f , the time series $x(t)$ has the self-affine fractal characteristic. The self-affine fractal dimension D is given

$$D = (5 - \beta) / 2 \quad (77)$$

$S(f)$ is plotted as a function of f with log-log scaling. β is the negative of the slope of the best-fit straight line in the range of large f . Note that the value of β is a measure of the strength of persistence in a time series. $\beta > 1$ reflects strong persistence and nonstationary. $1 > \beta > 0$ describes weak persistence and stationary. $\beta = 0$ shows uncorrelated stationary. $\beta < 0$ indicates antipersistence and stationary. In all cases, however, a self-affine time series with a non-zero β has long-range (as well as short-range) persistence and anti-persistence. For small β , the correlations with large lag are small but are non-zero. This can be contrasted with time series that are not self-affine; these may have only short-range persistence (either strong or weak).

Although the self-affine mapping are varied in a continuous way, the dimension of the self-affine set need not change continuously. Unfortunately, the self-affine fractal situation is much more complicated. It is quite difficult to obtain a general formula for the dimension of self-affine sets. It is not enough that only one fractal dimension is used to describe the self-affine fractal time series. The multifractal dimensions have been proposed to describe this kind of the time series[67-72].

4.3. Multifractal dimension

For a measured time series of a multifractal system, its trajectory in phase space is often attracted to a bounded fractal object called strange attractor for which a whole set of dimension D_q has been introduced which generalize the concept of the Hausdorff dimension. Let X_1, \dots, X_n be a point of the attractor in the phase space. The probability that the trajectory point is found within a ball of radius l around one of the inhomogeneously distributed points of the trajectory is denoted by

$$\mu(C_i) = \frac{1}{n} \sum_{i=1}^n \theta(l - \|X_i - X_j\|) \quad (78)$$

where $\theta(X)$ is the Heaviside step function. If $X \geq 0$, $\theta(X) = 1$; otherwise, $\theta(X) = 0$.

The q -order correlation integral is defined by

$$C_q(l) = \left(\frac{1}{n} \sum_{j=1}^n (\mu(C_i))^q \right)^{1/(q-1)} \quad (79)$$

The multifractal dimension D_q can be computed by the following equation:

$$D_q = \lim_{l \rightarrow 0} \frac{\ln(C_q(l))}{\ln(l)} = \begin{cases} \frac{1}{q-1} \lim_{l \rightarrow 0} \frac{\ln \sum_{i=1}^{N(l)} [\mu(C_i)]^q}{\ln(l)} & q \neq 1 \\ \lim_{l \rightarrow 0} \frac{\sum_{i=1}^{N(l)} \mu(C_i) \ln \mu(C_i)}{\ln(l)} & q = 1 \end{cases} \quad (80)$$

The above D_q is the multifractal dimension method based on Grassberger and Procaccia. The generalized correlation integral $C^q(l)$ which can be obtained from an experimental time series yields in a plot $\ln C^q(l)$ vs $\ln l$ straight lines with slopes D_q . For $q=0$, the D_0 is called the topological dimension, fractal dimension or capacity dimension. For $q=1$, the D_1 is called the information dimension. For $q=2$, the D_2 is called correlation dimension. The function D_q is monotonically decreasing with q and gives information about the inhomogeneity of the attractor. For simple fractals, called monofractals, such as a homogeneous attractor, the multifractal dimension D_q is constant. In the general case of multifractal objects, the values of D_q monotonically decrease as q increase[67]. The shape of the D_q can be considered a criterion confirming that the object is a nonuniform fractal. Furthermore, it can be determined if the object is a nonlinear, complex structure by using the multifractal dimension of the signal.

4.4. Analysis of surface EMG signal based on multifractal dimension

The surface EMG signal is a complicated physiological signal. Its distribution is clearly uneven (see Figure 30). When the surface EMG signal is studied by using the fractal method, one should first determine if the surface EMG signal is fractal. Then, its corresponding fractal dimension D can be estimated by Eq.(77) under a certain resolution. Figure 30 shows the self-affine fractal analysis of the surface EMG signals from Channel 1 during finger flexion, finger tension, forearm pronation and forearm supination (the results of Channel 2 are similar to those of Channel 1). It can be seen that the surface EMG signals have self-affine fractal characteristics. The results explain the physiological mechanism of the surface EMG signals.

In view of self-affine fractal characteristics, only one single fractal dimension is not easy to characterize the dynamics of surface EMG signals for different actions (see Table 4). There is little difference for the self-affine fractal dimensions of the four actions, where each type of action signals was chosen 100 sets of the data. The data length is 1000 points. In other words, it is difficult to identify the surface EMG signals of the different actions by using a single fractal dimension. The multifractal dimension values should be used to describe the action surface EMG signals during the arm movements.

	Finger flexion	Finger tension	Forearm pronation	Forearm supination
Channel 1	-0.2402±0.0725	-0.2571±0.0947	-0.0280±0.3250	0.0692±0.1418
Channel 2	0.0738±0.5734	-0.3199±0.2842	-0.0901±0.2591	-0.2343±0.2134

Table 4. The self-affine D of surface EMG signals during movements

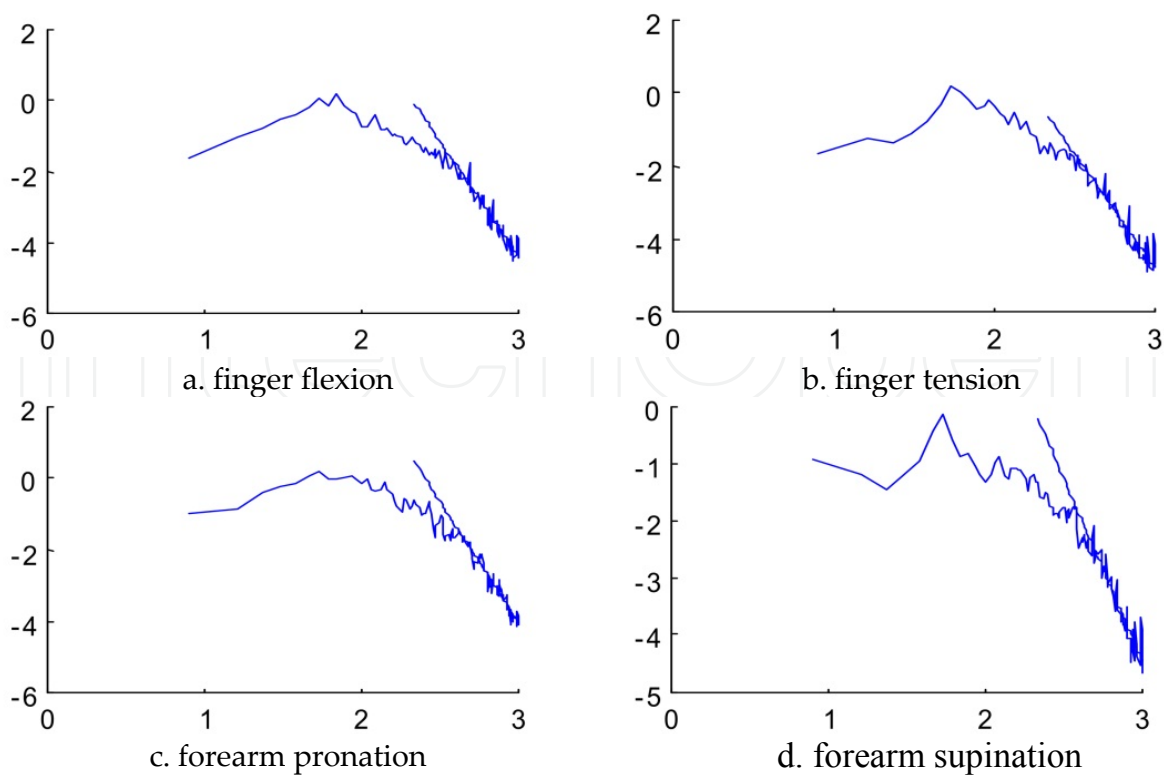
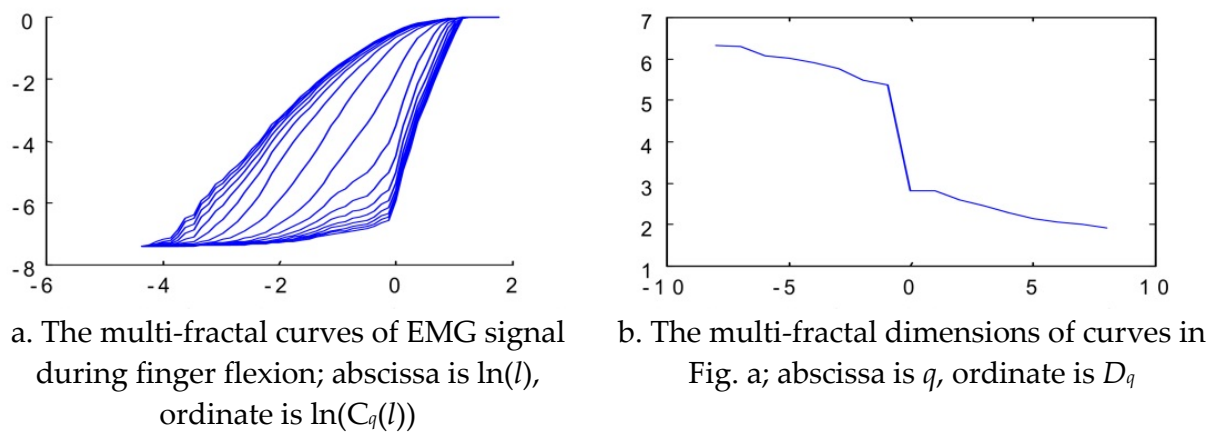
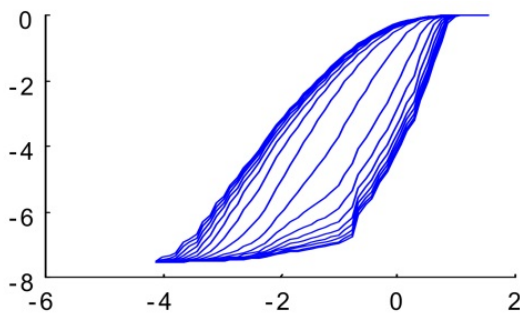


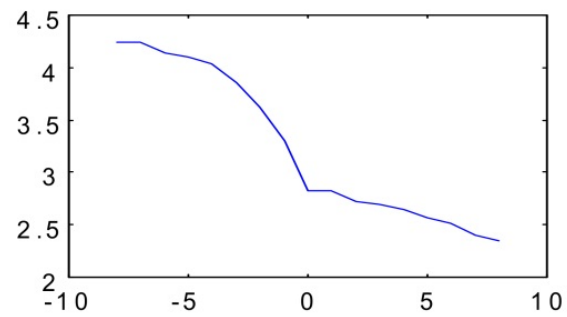
Figure 30. The analysis study of self-affine of surface EMG signal: Curve is power spectrum of surface EMG signal, line is straight line fit related part of curve; Abscissa is $\lg(f)$, ordinate is $\lg(\text{Psd})$

Here, we use the above multifractal dimension theory to analyze the action surface EMG signals. For the surface EMG signals of the four actions in channel 1, the multifractal analysis results are shown in Figure 31. The results of channel 2 are omitted since they are similar to those of channel 1. In the figure, the D_q - q curves are calculated under $q = 8, 7, 6, 5, 4, 3, 2, 1, 0, -1, -2, -3, -4, -5, -6, -7, -8$. It can be seen that the D_q - q curves have a certain range. The results indicate that the surface EMG signals are non-uniform fractal structure signals. These are consistent with the results of the above self-affine fractal analysis. The parameter values with q can be used to classify the data. In theory, it will be more reasonable that multifractal dimensions are used to describe the surface EMG signals. However, the actual calculation process of the multifractal dimensions is very time-consuming. For the surface EMG signals, it is extremely difficult to meet the requirements of real-time classification.

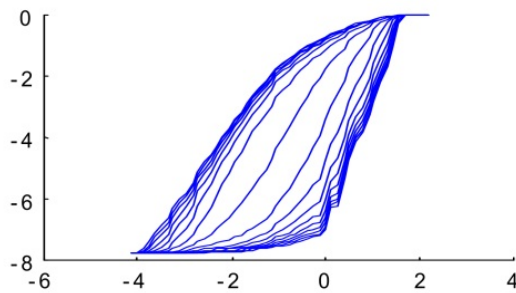




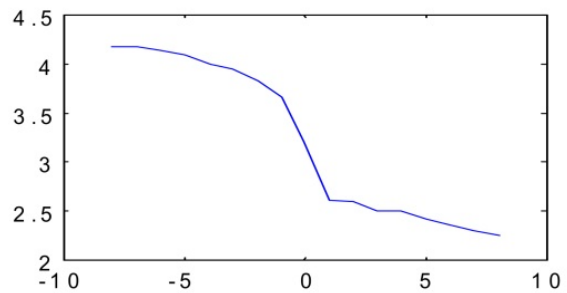
c. The multi-fractal curves of EMG signal during finger tension; abscissa is $\ln(l)$, ordinate is $\ln(C_q(l))$



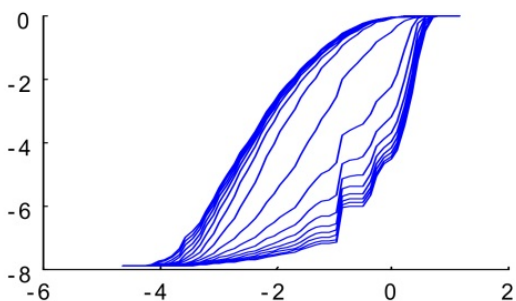
d. The multi-fractal dimensions of curves in Fig. c; abscissa is q , ordinate is D_q



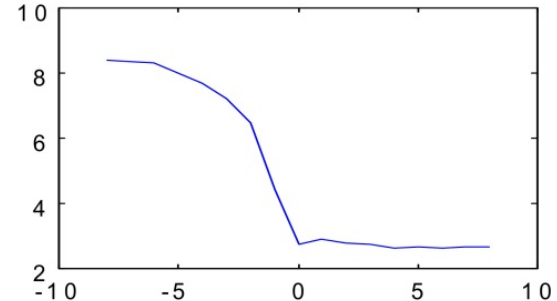
e. The multi-fractal curves of EMG signal during forearm pronation; abscissa is $\ln(l)$, ordinate is $\ln(C_q(l))$



f. The multi-fractal dimensions of curves in Fig. e; abscissa is q , ordinate is D_q



g. The multi-fractal curves of EMG signal during forearm pronation; abscissa is $\ln(l)$, ordinate is $\ln(C_q(l))$



h. The multi-fractal dimensions of curves in Fig. g; abscissa is q , ordinate is D_q

Figure 31. The multi-fractal analysis of surface EMG signals during movements

5. Conclusion and future research

In order to investigate whether the essence of the surface EMG signal is stochastic or deterministic nonlinear (even chaotic), some emerging nonlinear time series analysis approaches are discussed in this chapter. These techniques are based on detecting and describing deterministic structure in the signal, such as surrogate data method, VWK model method, chaotic analysis method, symplectic geometry method, fractal analysis method, and so on.

The surrogate data method and VWK model method are used to detect the surface EMG signal for arm movement and muscle fatigue. The results show that the surface EMG signal has deterministic nonlinear components. Moreover, our algorithm of surrogate data based on the null hypothesis 3 is proved that can completely satisfy the requirement of the null hypothesis 3. The VWK method with surrogate data can illuminate that not only the action but also fatigue surface EMG signals contain nonlinear dynamic properties.

Chaotic analysis techniques are reviewed and applied to investigate the surface EMG signals. The results show that the surface EMG signals have high-dimension chaotic dynamics by using correlation dimension and largest Lyapunov exponent techniques. For the estimation of embedding dimension, symplectic principal component analysis method is introduced and discussed. In comparison with correlation dimension algorithm and SVD analysis, symplectic geometry analysis is both very simple and reliable. The results show that symplectic geometry method is useful for determining of the system attractor from the experiment data.

The fractal theory is applied to study the fractal feature of the action surface EMG signal collected from forearm of normal person. The results show that the action surface EMG signal possesses the self-affine fractal characteristic. So, it is difficult to describe the surface EMG signals by using a single fractal dimension. The multifractal dimensions are used to analyze the action surface EMG signals during the arm movements. The results indicate that the surface EMG signals are non-uniform fractal structure signals. The multifractal dimension values can be used to identify the surface EMG signals for different movements.

For the nonlinear characteristics of the surface EMG signals, chaos and fractal theories will play the leading role in the nonlinear study of the surface EMG signals. The related methods need to be further researched and developed although these techniques have been applied to analyze the surface EMG signals. It provides a new way for the study of the quantitative analysis of physiology and pathology, sports medicine, clinical medical diagnostics and bionics of robot limb motion.

Appendix A

Theorem 1 proof:

$$\begin{aligned}
 & \because J \begin{pmatrix} A & 0 \\ 0 & -A^* \end{pmatrix} J^{-1} \\
 &= \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -A^* \end{pmatrix} \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}^{-1} \\
 &= \begin{pmatrix} -A^* & 0 \\ 0 & A \end{pmatrix} \\
 &= - \begin{pmatrix} A & 0 \\ 0 & -A^* \end{pmatrix}^*
 \end{aligned} \tag{81}$$

\therefore According Definition 2, $\begin{pmatrix} A & 0 \\ 0 & -A^* \end{pmatrix}$ is a Hamilton matrix. \square

Theorem 2 proof:

Let S as a symplectic transform matrix, M as a Hamilton matrix. Then S^{-1} is also symplectic matrix. According Definition 1 and 2, there is

$$\begin{aligned} & J(SMS^{-1})J^{-1} \\ &= JSJ^{-1}JMJ^{-1}JS^{-1}J^{-1} \\ &= S^{-*}(-M^*)S^* \\ &= -(SMS^{-1})^* \end{aligned} \quad (82)$$

$\therefore SMS^{-1}$ is also a Hamilton matrix.

$$\therefore SMS^{-1} \sim M \quad (83)$$

So Hamilton matrix M keeps unchanged at symplectic similar transform. \square

Theorem 5 proof:

Let S_1, S_2, \dots, S_n as symplectic matrix, respectively. According Definition 1, there are

$$\begin{aligned} JS_1J^{-1} &= S_1^{-*} \\ JS_2J^{-1} &= S_2^{-*} \\ &\dots \\ JS_nJ^{-1} &= S_n^{-*} \end{aligned} \quad (84)$$

$$\begin{aligned} & J(S_1S_2 \dots S_n)J^{-1} \\ &= JS_1J^{-1}JS_2J^{-1}J \dots J^{-1}JS_nJ^{-1} \\ &= S_1^{-*}S_2^{-*} \dots S_n^{-*} \\ &= (S_1S_2 \dots S_n)^{-*} \end{aligned} \quad (85)$$

So the product of symplectic matrixes is also a symplectic matrix.

Theorem 6 proof:

In order to prove that H matrix is symplectic matrix, we only need to prove $H^*JH = J$.

$$\begin{aligned} H^*JH &= \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix}^* J \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix} \\ &= \begin{pmatrix} 0 & P^*P \\ -P^*P & 0 \end{pmatrix} \end{aligned} \quad (86)$$

$$\because P = I_n - \frac{2\varpi\varpi^*}{\varpi^*\varpi}$$

$$\therefore P^* = P$$

$$P^*P = P^2$$

$$\begin{aligned} &= \left(I_n - \frac{2\varpi\varpi^*}{\varpi^*\varpi} \right) \left(I_n - \frac{2\varpi\varpi^*}{\varpi^*\varpi} \right) \\ &= I_n - \frac{4\varpi\varpi^*}{\varpi^*\varpi} + \frac{4\varpi(\varpi^*\varpi)\varpi^*}{(\varpi^*\varpi)(\varpi^*\varpi)} \\ &= I_n \end{aligned} \quad (87)$$

where $\varpi = (0, \dots, 0; \omega_k, \dots, \omega_n)^T \neq 0$.

Plugging Eq.(87) into Eq.(86), we have:

$$H^*JH = J \quad (88)$$

$\therefore H$ is symplectic matrix.

$$\begin{aligned} H^*H &= \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix}^* \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix} \\ &= \begin{pmatrix} P^*P & 0 \\ 0 & P^*P \end{pmatrix} \\ &= I_{2n} \end{aligned} \quad (89)$$

$\therefore H$ is also unitary matrix.

$\therefore H$ is symplectic unitary matrix. \square

Appendix B

Theorem 7 suppose x and y are two unequal n dimension vectors, and $\|x\|_2 = \|y\|_2$, so there is elementary reflective array $H = 1 - 2\varpi\varpi^T$, which make $Hx = y$, where $\varpi = \frac{x-y}{\|x-y\|_2}$.

It can be easily deduced from theorem 5, for non zero n dimension vector $x = (x_1, x_2, \dots, x_n)^T$, notes $\alpha = \|x\|_2$, there is

$$Hx = \alpha e_1 \quad (90)$$

$$H = 1 - 2\varpi\varpi^T \quad (91)$$

$$e_1 = (1, 0, \dots, 0)^T \quad (92)$$

$$\varpi = \frac{1}{\rho}(x - \alpha e_1) \quad (93)$$

$$\rho = \|x - \alpha e_1\|_2 \quad (94)$$

Then $\|\omega\|_2 = 1$, and H is elementary reflective array.

It's easy to testify, elementary reflective array H is symmetry matrix ($H^T = H$), orthogonal matrix ($H^T H = I$) and involution matrix ($H^2 = I$).

For real symmetrical matrix A , Householder matrix H can be constructed as follows[73].
Notes A :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & A_{12}^{(1)} \\ \alpha_{21}^{(1)} & A_{22}^{(1)} \end{pmatrix} \quad (95)$$

First, suppose $\alpha_{21}^{(1)} \neq 0$, otherwise this column will be skipped and the next column will be considered until the i^{th} column of $\alpha_{2i}^{(1)} \neq 0$. Set first column vector of A :

$$S^{(1)} = (a_{11}^{(1)}, a_{21}^{(1)}, \dots, a_{n1}^{(1)})^T = (a_{11}, a_{21}, \dots, a_{n1}) \quad (96)$$

select elementary reflective array $H^{(1)}$:

$$H^{(1)} = I - 2\varpi^{(1)}(\varpi^{(1)})^T \quad (97)$$

where

$$\begin{cases} \alpha_1 = \|S^{(1)}\|_2 \\ E^{(1)} = (1, 0, \dots, 0)^T & n \times 1 \\ \rho_1 = \|S^{(1)} - \alpha_1 E^{(1)}\| \\ \varpi^{(1)} = \frac{1}{\rho_1} (S^{(1)} - \alpha_1 E^{(1)}) \end{cases} \quad (98)$$

so, after $H^{(1)}$ transform, A is changed to a matrix with the first column is all zero except the first element is α_1 , namely:

$$H^{(1)} A = \begin{pmatrix} \sigma_1 & a_{12}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix} = A^{(2)} \quad (99)$$

Second, the same method is adopted to the second column vector of $A^{(2)}$, let

$$S^{(2)} = (0, a_{22}^{(2)}, \dots, a_{n2}^{(2)})^T \quad (100)$$

construct $H^{(2)}$ matrix :

$$H^{(2)} = I - 2\omega^{(2)}(\omega^{(2)})^T \quad (101)$$

where,

$$\begin{cases} \alpha_2 = \|S^{(2)}\|_2 \\ E^{(2)} = (0, 1, 0, \dots, 0)^T \\ \rho_2 = \|S^{(2)} - \alpha_2 E^{(2)}\| \\ \varpi^{(2)} = \frac{1}{\rho_2} (S^{(2)} - \alpha_2 E^{(2)}) \end{cases} \quad n \times 1 \quad (102)$$

Using $H^{(2)}$, the second column of $A^{(2)}$ can be changed to all zero vector except the first and second elements, namely:

$$H^{(2)} A^{(2)} = A^{(3)} \quad (103)$$

Householder matrix H can be obtained by repeating above mentioned method until $A^{(n)}$ becomes an upper triangle matrix:

$$H = H^{(n)} H^{(n-1)} \dots H^{(1)} \quad (104)$$

Author details

Min Lei and Guang Meng

Institute of Vibration, Shock and Noise, State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai, P R China

Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 10872125 and No.69675002), Science Fund for Creative Research Groups of the National Natural Science Foundation of China(no. 50821003), State Key Lab of Mechanical System and Vibration , Project supported by the Research Fund of State Key Lab of MSV, China (Grant no. MSV-MS-2010-08) , Key Laboratory of Hand Reconstruction, Ministry of Health, Shanghai, People's Republic of China, Shanghai Key Laboratory of Peripheral Nerve and Microsurgery, Shanghai, People's Republic of China, Science and Technology Commission of Shanghai Municipality (no.06ZR14042). We also thank Chinese Academy of Engineering GU Yudong and Professor Zhang Kaili very much for providing related data and valuable discussions.

6. References

- [1] J Theiler, S Eubank, A Longtin, et al., Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, 1992. 58: 77-94.
- [2] M Barahona and C-S Poon, Detection of nonlinear dynamics in short, noisy time series. *Nature*, 1996. 381: 215-217.
- [3] M Lei, Z Z Wang, and Z J Feng, Detecting nonlinearity of action surface EMG signal. *Phys. Lett. A*, 2001. 290(5-6): 297-303.

- [4] M Palus and D Hoyer, Detecting nonlinearity and phase synchronization with surrogate data. *IEEE Engineering in Medical and Biology*, 1998. 17(6): 40-45.
- [5] C-S Poon and K M Christopher, Decrease of cardiac chaos in congestive heart failure. *Nature*, 1997. 389: 492-495.
- [6] J Theiler and D Prichard, Constrained-realization Monte-Carlo method for hypothesis testing. *Physica D*, 1996. 94: 221-235.
- [7] C S Poon and M Barahona, Titration of chaos with added noise. *Proceedings of The National Academy of Sciences of The United States of America*, 2001. 98(13): 7107-7112.
- [8] J A Scheinkman and B LeBaron., Nonlinear dynamics and stock returns. *Journal of Business*, 1989. 62(3): 311-337.
- [9] J L Breeden and N H Packard, Nonlinear analysis of data sampled nonuniformly in time. *Physica D*, 1992. 58(1-4): 273-283.
- [10] A R Osborne and A Provenzale, Finite correlation dimension for stochastic systems with power-law spectra. *Physica D*, 1989. 35(3): 357-381.
- [11] M Lei, *Nonlinear Time Series Analysis and Its Applied Study in Surface EMG Signal*. 2000, Shanghai Jiao Tong University (in chinese).
- [12] T Schreiber and A Schmitz, Improved surrogate data for nonlinearity tests. *Phys. Rev. Lett.*, 1996. 77(4): 635-638.
- [13] M Lei and Z Z Wang, Application of surrogate data method in electromyograph analysis. *Journal of Shanghai Jiao Tong University*, 2000. 34(11): 1594-1597 (in Chinese).
- [14] M Lei and Z Z Wang, Study of the surrogate data method for nonlinearity of time series. *Journal of Electronics and Information Technology*, 2001. 23(3): 248-254 (in Chinese).
- [15] M Lei, Z J Feng, and Z Z Wang, The Volterra-Wiener-Korenberg model nonlinear test method. *Journal of Shanghai Jiao Tong University*, 2003. 37(2): 269-272 (in Chinese).
- [16] M Lei and G Meng, The influence of noise on nonlinear time series detection based on Volterra-Wiener-Korenberg model. *Chaos, Solitons and Fractals*, 2008. 36: 512-516.
- [17] E N Lorenz, Deterministic nonperiodic flow. *Journal of The Atmospheric Sciences*, 1963. 20: 130-141.
- [18] T Y Li and J A Yorke, Period three implies chaos. *Amer Math Monthly*, 1975. 82: 985-992.
- [19] R M May, Simple mathematical models with very complicated dynamics. *Nature*, 1976. 261: 459-467.
- [20] G U Yule., On a method of investigating periodicities in disturbed series with special reference to wolfer's sunspot numbers. *Phil. Trans. R. Soc. London A*, 1927. 226(636-646): 267-298
- [21] F Takens, Detecting Strange Attractors in Turbulence. *Dynamical Systems and Turbulence*, 1980. 898: p. 366-381.
- [22] N H Packard, J P Crutchfield, J D Farmer, et al., Geometry from a time series. *Phys. Rev. Lett.*, 1980. 45: 712-716.
- [23] K Vibe and J M Vesin, On chaos detection methods. *International Journal of Bifurcation and Chaos*, 1996. 6(3): 529-543.

- [24] A M Fraser, Information and entropy in strange attractors. *IEEE Trans. Inf. Theory*, 1989. 35(2): 245-262.
- [25] B Cheng and H Tong, Orthogonal Projection, Embedding Dimension and Sample size in chaotic time series from a statistical perspective. *Nonlinear time series and chaos*, 1994. 2(15): 1-30.
- [26] H S Kim, R Eykholt, and J D Salas, Nonlinear dynamics, delay times, and embedding windows. *Physica D*, 1999. 127: 48-60.
- [27] D Kugiumtzis, State space reconstruction parameters in the analysis of chaotic time series — the role of the time window length. *Physica D*, 1996. 95(1): 13–28.
- [28] M T Rosenstein, J J Collins, and C J D Luca, Reconstruction expansion as a geometry-based framework for choosing proper delay times. *Physica D*, 1994. 73: 82-98.
- [29] J M Martinerie, A M Albano, A I Mees, et al., Mutual information, strange attractors, and the optimal estimation of dimension. *Phys. Rev. A* 1992. 45: 7058-7064.
- [30] A M Fraser and H L Swinney, Independent coordinates for strange attractors from mutual information. *Phys. Rev. A*, 1986. 33(2): 1134-1140.
- [31] D S Broomhead and G P King, Extracting qualitative dynamics from experimental data. *Physica D*, 1986. 20(2-3): 217–236.
- [32] M Knennel, R Brown, and H Abarbanel., Determining embedding dimension for phase — space reconstruction using a geometrical construction. *Phys. Rev. A*, 1992. 45: 3403–3411.
- [33] P Grassberger and I Procaccia, Measuring the strangeness of strange attractors. *Physica D*, 1983. 9(1-2): 189-208.
- [34] A Galka, T Maaß, and P G., Estimating the dimension of high-dimensional attractors: a comparison between two algorithms. *Physica D*, 1998. 121: 237-251.
- [35] K Judd and A Mees, Embedding as a modeling problem. *Physica D*, 1998. 120: 273-286.
- [36] L Cao, Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D*, 1997. 110(1-2): 43–50.
- [37] M Palus and I Dvorak, Singular-value decomposition in attractor reconstruction: pitfalls and precautions. *Physica D*, 1992. 55: 221-234.
- [38] K Judd, An improved estimator of dimension and some comments on providing confidence intervals. *Physica D*, 1992. 56: 216-228.
- [39] P Grassberger, Generalized dimensions of strange attractors. *Phys. Lett. A*, 1983. 97: 227-234.
- [40] P Grassberger and I Procaccia, Characterization of strange attractors. *Phys. Rev. Lett.*, 1983. 50: 346–349.
- [41] M Ding, C Grebogi, E Ott, et al., Plateau onset for correlation dimension: when does it occur? *Phys. Rev. Lett.*, 1993. 70: 3872–3875.
- [42] M Small and K Judd, Detecting nonlinearity in experimental data. *International Journal of Bifurcation and Chaos*, 1998. 8(6): 1231-1244.
- [43] A Wolf, J Swift, H Swinney, et al., Determining Lyapunov exponents from a time series. *Physica D*, 1985. 16: 285-317.
- [44] M Sano and Y Sawada, Measurement of the Lyapunov spectrum from a chaotic time series. *Phys. Rev. Lett.*, 1985. 55(10): 1082-1085.

- [45] M T Rosenstein, J J Collins, and C J D Luca, A practical method for calculating largest lyapunov exponents from small data sets. *Physica D*, 1993. 65: 117-134.
- [46] A I Mees, P E Rapp, and L S Jennings, Singular-value decomposition and embedding dimension. *Phys. Rev. A*, 1987. 36: 340-346.
- [47] A M Fraser., Reconstructing attractors from scalar time series : a comparison of singular system and redundancy criteria. *Physica D*, 1989. 34: 391-404.
- [48] A Brandstater, H L Swinney, and G I Chapman, Characterizing turbulent channel flow. in: *Dimensions and entropies in chaotic systems*, G. Mayer-Kree(Ed.), Springer, Berlin 1986: 150.
- [49] M Lei, Z Z Wang, and Z J Feng, A method of embedding dimension estimation based on symplectic geometry. *Phys. Lett. A*, 2002. 303(2-3): 179-189.
- [50] X Lu and R Schmid, Symplectic integration of Sine-Gordon type systems. *Mathematics and Computers in Simulation*, 1999. 50: 255-263.
- [51] Z G Ying and W Q Zhu, Exact stationary solutions of stochastically excited and dissipated gyroscopic systems. *International Journal of Non-Linear Mechanics*, 2000. 35: 837-848.
- [52] A C J Luo and R P S Han, The resonance theory for stochastic layers in nonlinear dynamic systems. *Chaos, Solitons and Fractals*, 2001. 12: 2493-2508.
- [53] K Feng, *Proceeding of the 1984 Beijing Symposium on Differential Geometry and Differential Equations: Computation of partial differential equations*, ed. F. Kang. 1985: Science Press, 42.
- [54] C V Loan, A Symplectic Method for Approximating All the Elgenvalues of a Hamiltonian Matrix. *Linear Algebra and Its Applications*, 1984. 61: 233-251.
- [55] M Lei, Z Z Wang, and Z J Feng, The application of symplectic geometry on nonlinear dynamics analysis of the experimental data. *14th International Conference on Digital Signal Processing Proceeding*, 2002. 1-2: 1137-1140.
- [56] M Lei and G Meng, Detecting nonlinearity of sunspot number. *International Journal of Nonlinear Sciences and Numerical Simulation*, 2004. 5(4): 321-326.
- [57] X Niu, F Qu, and N Wang, Evaluating Sprinters' Surface EMG Signals Based on EMD and Symplectic Geometry. *J of Ocean Univ. of Qingdao*, 2005. 35(1): 125-129.(in Chinese).
- [58] H Xie, Z Z Wang, and H Huang, Identification determinism in time series based on symplectic geometry spectra. *Phys. Lett. A*, 2005. 342(1-2): 156-161.
- [59] M Lei and G Meng, Symplectic principal component analysis: a new method for time series analysis. *Mathematical Problems in Engineering*, 2011. 2011: Article ID 793429, 14pages.
- [60] M Lei and G Meng, A noise reduction method for continuous chaotic systems based on symplectic geometry. *The Seventh International Conference on Vibration Engineering and Technology of Machinery*, 2011. 1: Article ID 173, 6pages.
- [61] Y Gong and J Xu, Chaos signal and noise. *Signal Processing*, 1997. 13(2): 112-118 (in Chinese)
- [62] J Yuan and X C Xiao, Higher-order singular-spectrum analysis of nonlinear time series. *Acta Physica Sinica*, 1998. 47(6): 897-905 (in Chinese).

- [63] B B Mandelbrot, Self-affine fractals and fractal dimension. *Physica Scripta*, 1985. 32: 257-260.
- [64] B B Mandelbrot, *The fractal geometry of Nature*. 1982: San Francisco: Freeman.
- [65] K Falconer, *Fractal Geometry: Mathematical Foundations and Applications*. 2003: John Wiley & Sons Ltd. 337.
- [66] B D Malamud and D L Turcotte, Self-affine time series: measures of weak and strong persistence. *Journal of Statistical Planning and Inference*, 1999. 80: 173-196.
- [67] A N Pavlov and V S Anishchenko, Multifractal analysis of complex signals. *Physics-Uspekhi*, 2007. 50(8): 819-834.
- [68] R H Riedi, M S Crouse, V J Ribeiro, et al., A multifractal wavelet model with application to network traffic. *IEEE Trans. On Information Theory*, 1999. 45(3): 992-1018.
- [69] A Z R Langi, K Soemintapura, and W Kinsner. Multifractal processing of speech signals. in *International Conference on Information, Communications and Signal Processing*. 1997.
- [70] Z Xu and S Xiao. Fractal dimension of surface EMG and its determinants. in *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 1997.
- [71] O Adeyemi and G F Boudreaux-Bartels. Improved accuracy in the singularity spectrum of multifractal chaotic time series. in *IEEE International Conference on Acoustics, Speech, and Signal Processing*. 1997.
- [72] C J Aumuth, Fractals dimension of electromyographic signals recorded with surface electrodes during isometric contractions with muscle activation. *Muscle & Nerve*, 1994. 17: 953-954.
- [73] J Huang, *The practical modeling method of static and dynamic mathematical model*. 1988, Beijing: Mechanical Industry Press (in Chinese).