

Chaotic signal denoising algorithm based on sparse decomposition

Jin-Wang Huang(黄锦旺), Shan-Xiang Lv(吕善翔), Zu-Sheng Zhang(张足生), Hua-Qiang Yuan(袁华强)

Citation: Chin. Phys. B . 2020, 29(6): 060505 . **doi:** 10.1088/1674-1056/ab8a3b
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Nonlinear continuous bi-inductance electrical line with dissipative elements: Dynamics of the low frequency modulated waves

S M Ngounou, F B Pelap

Chin. Phys. B . 2020, 29(4): 040502 . **doi:** 10.1088/1674-1056/ab75d6

Dynamical response of a neuron-astrocyte coupling system under electromagnetic induction and external stimulation

Zhi-Xuan Yuan(袁治轩), Pei-Hua Feng(冯沛华), Meng-Meng Du(独盟盟), Ying Wu(吴莹)

Chin. Phys. B . 2020, 29(3): 030504 . **doi:** 10.1088/1674-1056/ab7441

Dynamics of the plane and solitary waves in a Noguchi network: Effects of the nonlinear quadratic dispersion

S A T Fonkoua, M S Ngounou, G R Deffo, F B Pelap, S B Yamgoue, A Fomethe

Chin. Phys. B . 2020, 29(3): 030501 . **doi:** 10.1088/1674-1056/ab696a

Signal reconstruction in wireless sensor networks based on a cubature Kalman particle filter

Huang Jin-Wang, Feng Jiu-Chao

Chin. Phys. B . 2014, 23(7): 070504 . **doi:** 10.1088/1674-1056/23/7/070504

Gradient method for blind chaotic signal separation based on proliferation exponent

Lü Shan-Xiang, Wang Zhao-Shan, Hu Zhi-Hui, Feng Jiu-Chao

Chin. Phys. B . 2014, 23(1): 010506 . **doi:** 10.1088/1674-1056/23/1/010506

Chaotic signal denoising algorithm based on sparse decomposition*

Jin-Wang Huang(黄锦旺)[1], Shan-Xiang Lv(吕善翔)[2], Zu-Sheng Zhang(张足生)[1], and Hua-Qiang Yuan(袁华强)[1,†]

[1] *School of Cyberspace Science, Dongguan University of Technology, Dongguan* 523808, *China*
[2] *College of Cyber Security, Jinan University, Guangzhou* 510632, *China*

(Received 27 February 2020; revised manuscript received 10 April 2020; accepted manuscript online 17 April 2020)

Denoising of chaotic signal is a challenge work due to its wide-band and noise-like characteristics. The algorithm should make the denoised signal have a high signal to noise ratio and retain the chaotic characteristics. We propose a denoising method of chaotic signals based on sparse decomposition and K-singular value decomposition (K-SVD) optimization. The observed signal is divided into segments and decomposed sparsely. The over-complete atomic library is constructed according to the differential equation of chaotic signals. The orthogonal matching pursuit algorithm is used to search the optimal matching atom. The atoms and coefficients are further processed to obtain the globally optimal atoms and coefficients by K-SVD. The simulation results show that the denoised signals have a higher signal to noise ratio and better preserve the chaotic characteristics.

Keywords: sparse decomposition, denoising, K-SVD, chaotic signal

PACS: 05.45.–a, 05.40.Ca **DOI:** 10.1088/1674-1056/ab8a3b

1. Introduction

Chaos phenomenon widely exists in astronomy, meteorology, stock, and so on.[1,2] Some chaotic system can be described by definite equations, but its behavior is similar to stochastic and has the characteristics of broadband and similar to noise.[3,4] Because of the measurement accuracy and external influence, the actually measured chaotic signal is inevitably mixed with noise. Therefore, the characteristics of chaotic signals may be concealed, affecting the accuracy of parameter calculation and prediction of the signals, so the research of chaotic signal denoising algorithm is very important. And the non-linearity and broadband characteristics of chaotic signals makes the denoising more challenging.

The current denoising algorithms for chaotic signals can be classified into three categories. One is that the system model and parameters are known. The recursive expressions of chaotic signals are used as state equations, the adaptive filter, such as unscented Kalman filtering and cubature Kalman filter, is adopted to estimate the signal.[5–7] These methods can achieve online denoising, but the system model and parameters need to be known as *a prior*. The second is that the system model is known but the parameters are unknown. The phase space denoising method proposed in Ref. [8] belongs to this kind, which explicitly uses the source expression of the chaotic signals, but the recursive parameters are estimated by a least square method in the algorithm. Finally, the global error of the estimated signal is minimized. The third is that the system model and parameters are unknown. These algorithms belong to a completely blind. Wavelet threshold (WT) algorithm,[9,10] local curve fitting (LCF) algorithm,[11,12] and empirical mode decomposition (EMD) algorithm[13,14] are representative in all-blind methods. Essentially, they make use of the low-frequency property of chaotic flow signals, and carry out low-pass filtering while preserving the non-periodic property as far as possible in the filtering process. Their validity depends on the determination of the filtering boundary. The WT algorithm sets the coefficients to zero (hard threshold) under a certain threshold in high frequency band. The LCF algorithm performs hard projection under an optimal window length. The key of the EMD algorithm is to select the optimal number of eigenfunctions to reconstruct the source signal.

Based on these existing algorithms, we propose a noise suppression algorithm using sparse decomposition[15,16] and K-SVD.[17] Since the decomposition coefficient of noise in the dictionary is not sparse, the non-zero coefficients cover the whole transformation domain and generally are small compared with the coefficient of the signal. Moreover, the sparse decomposition coefficient of the signal in the dictionary is sparse, only K non-zero coefficients exist, the coefficient of the noisy signal has only K big coefficients. By decomposing the noisy signal in the dictionary $\boldsymbol{A}$ and selecting only the first K big coefficients, the linear combination of the atoms corresponding to these K large coefficients can approximate the clean signal. The approximation signal obtained contains most of the information of the useful signal and discards most of the noise information, which achieves the purpose of noise elimination.

The innovation of this algorithm lies in using noise inten-

*Project supported by the National Natural Science Foundation of China (Grant No. 61872083) and the Natural Science Foundation of Guangdong Province, China (Grant Nos. 2017A030310659 and 2019A1515011123).

†Corresponding author. E-mail: yuanhq@dgut.edu.cn

© 2020 Chinese Physical Society and IOP Publishing Ltd

sity to dynamically determine the sparsity of the reconstructed signals, giving a low-complexity dictionary design scheme and optimizing by the K-SVD algorithm. The organizational structure of this paper is as follows: The second part introduces the sparse decomposition denoising model, the third part introduces the algorithm and describes the principle of each step in detail, and the fourth part analyzes the algorithm. The fifth part compares the algorithm with other classical algorithms through numerical simulation, and finally comes to the conclusion.

2. Signal model

The observed noisy signal $y(n)$ from a noise-free signal $x(n)$ can be modeled as

$$y(n) = x(n) + v(n), \quad n = 1,2,3,...,N, \tag{1}$$

where $v(n)$ is the Gauss noise with zero mean and variance σ^2.

As the signal is divided into segments and processed, we use the window function $F(y(n), M')$ with window length $M' = 2M+1$ for signal segmentation, $y(n)$ is segmented with length M' and the moving step is M each time. The signal model can be rewrite in vector form $\boldsymbol{y}_t = \boldsymbol{x}_t + \boldsymbol{v}_t$. Here $\boldsymbol{y}_t = [y(t), y(t+1), \ldots, y(t+M'-1)]^{\mathrm{T}} \in \mathbb{R}^{M'}$, $\boldsymbol{x}_t$ and $\boldsymbol{v}_t$ have the similar expression. All the vectors $\boldsymbol{y}_t$ form a matrix $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_L]$, where $L = N/M$ and $\boldsymbol{Y} \in \mathbb{R}^{M' \times L}$.

For the noisy observed signal $\boldsymbol{y}_t$, the sparsest representation is then the solution to the optimization problem[15]

$$(P_{0,\delta}): \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \text{ subject to } \|\boldsymbol{y}_t - \boldsymbol{A}\boldsymbol{\alpha}\|_2 \le \delta, \tag{2}$$

where $\boldsymbol{A}$ is the dictionary and $\boldsymbol{\alpha}$ is the sparse decomposition coefficient vector. Note that $P_{0,0} \equiv P_0$, where P_0 is the optimization problem of noiseless signal sparse decomposition. If we apply this with $\delta > \varepsilon = \|\boldsymbol{y}_t - \boldsymbol{x}_t\|_2$, the problem has a sparse solution. In fact, the solution $\hat{\boldsymbol{\alpha}}_t$ obeys $\|\hat{\boldsymbol{\alpha}}_t\|_0 = val(P_{0,\delta}) \le val(P_0) = \|\boldsymbol{\alpha}_t\|_0$, where $\boldsymbol{\alpha}_t$ is the solution of noiseless optimization problem P_0, and $val(\cdot)$ is the solution of the optimization problem.

Usually the problem $(P_{0,\delta})$ demands exorbitant computational efforts in general and does not have a unique sparse solution. If $(P_{0,\delta})$ has the sparsest solution, the corresponding expression can be solved. In the case of noise, we can pursue a strategy of convexification, replacing the ℓ^0-norm in Eq. (2) by an ℓ^1-norm

$$(P_{1,\delta}): \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \text{ subject to } \|\boldsymbol{y}_t - \boldsymbol{A}\boldsymbol{\alpha}\|_2 \le \delta. \tag{3}$$

For the $(P_{1,\delta})$ problem, there are many algorithms to solve the convex quadratic program, such as iteratively-reweighted least squares (IRLS),[18] interior-point algorithms[19] and active-set methods. The solution of the $(P_{1,\delta})$ problem is closely related to the BPDN algorithm.[19] Both of those proposals amount to solving a corresponding convex optimization in Lagrangian form

$$(P'_{1,\lambda}): \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 + \|\boldsymbol{y}_t - \boldsymbol{A}\boldsymbol{\alpha}\|_2^2/\lambda, \tag{4}$$

when $\lambda = \lambda(\boldsymbol{y}_t, \delta)$, the problems $(P_{1,\delta})$ and $(P'_{1,\lambda(\boldsymbol{y}_t,\delta)})$ have the same solutions.

3. The proposed denoising algorithm

By decomposing the contaminated signal and selecting only the first K large coefficients, the linear combination of these K large coefficients and the corresponding atoms can approach the clean signal and the purpose of noise cancellation can be achieved. Since selecting the sparsity of each frame, setting up a suitable dictionary $\boldsymbol{A}$ and optimizing the atoms and coefficients are the key parts of this denoising algorithm, we propose a dynamic determination of sparsity, using the iterative equation expansion of the signal to get the dictionary and K-SVD algorithm to optimize the atoms and coefficients here. Figure 1 shows the flow chart of the sparse decomposition based denoising algorithm.

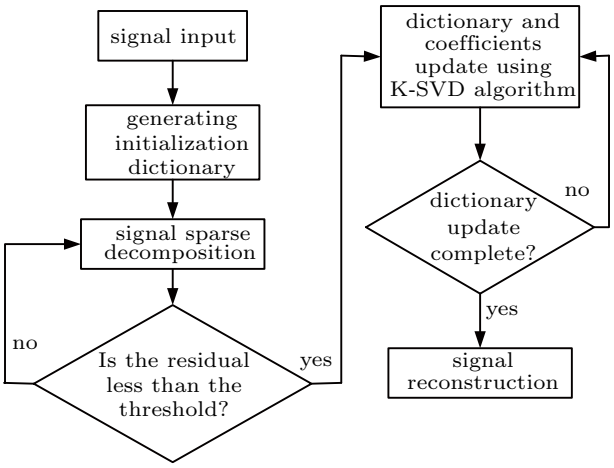


Fig. 1. Block diagram of sparse decomposition based denoising algorithm.

3.1. Sparsity determination

The energy of noise can be roughly estimated by means of least total variation or least-squares principle in the case of blind processing,[22,23] the sparsity in the denoising model is based on the parameter of noise intensity. Define the problem

$$(P_0^{\varepsilon}): \min_{\boldsymbol{A}\boldsymbol{\alpha}_t} \|\boldsymbol{y}_t - \boldsymbol{A}\boldsymbol{\alpha}_t\|_2 \text{ subject to } \|\boldsymbol{\alpha}_t\|_0 = f(\sigma_y^2, \sigma_v^2), \tag{5}$$

where $f(\sigma_y^2, \sigma_v^2) = \lfloor \frac{M}{\omega} \tanh(\frac{\sigma_y^2}{\sigma_v^2}) \rfloor + 1$ is the function of the input signal energy σ_y^2 and the estimated noise energy σ_v^2, ω is a constant, and $2M+1$ is the length of vector $\boldsymbol{y}_t$. The sparsity of the estimated signal reduces with the increase of noise, because the larger the noise is, the greater the probability that the atom (column) of the dictionary $\boldsymbol{A}$ obtained from zero norm tracking is misjudged. Therefore, only using atoms with high confidence probability to reconstruct $\boldsymbol{x}$ can achieve maximum noise suppression.

3.2. Dictionary construction

The chaotic flow signal can be described by differential equation $\dot{z} = F(z)$, and it is expressed recursively, so $x(\theta)$ generated from the system is $k+1$ order differentiable, the Taylor expansion formula of $x(n)$ at θ is

$$\begin{aligned} x(n) &= x(\theta) + x'(\theta)(n-\theta) + \cdots + \frac{x^{(k)}(\theta)}{k!}(n-\theta)^k \\ &\quad + o((n-\theta)^{k+1}). \end{aligned} \tag{6}$$

Each segment of the signal is regarded as centered on the origin, the above expression can be rewritten as follows:

$$x(n) = \sum_{k=0}^{\infty} \frac{x^{(k)}(0)}{k!} n^k. \tag{7}$$

Based on the above representation, each atom of the matrix $\boldsymbol{A}$ can be set as a power exponent corresponding to the sparse coefficient, namely, $\boldsymbol{A}(1:M', j) \triangleq \boldsymbol{a}_j = n^{j-1}$, where $1 \le j \le N'$, $n = -\lceil \frac{M'-1}{2} \rceil : 1 : \lceil \frac{M'-1}{2} \rceil$. Every atom $\boldsymbol{a}_j$ is normalized to $\boldsymbol{a}_j \leftarrow \boldsymbol{a}_j / \|\boldsymbol{a}_j\|_2$, and unit norm atom can simplify the steps of subsequent denoising algorithm.

3.3. Signal sparse decomposition

For the problem P_0^{ε}, using the ℓ_1 norm constraint instead of the ℓ_0 norm constraint will have a better recovery accuracy, but the convex relaxation algorithm corresponding to ℓ_1 norm is slower than the greedy algorithm corresponding to ℓ_0 norm with Iter$/K$, where Iter is the minimum number of iterations for the convex relaxation algorithm to reach the convergence state. In addition, the ℓ_1 norm tracking algorithm will introduce more atoms, resulting in poor performance of the denoising algorithm. The idea of the algorithm in this paper is to retain only the atoms with high confidence probability under the noise pollution, and the greedy algorithm is more suitable for solving the problem. We use the orthogonal matching pursuit (OMP) algorithm[20,21] to solve the problem of data reconstruction in each frame, because the length of the data to be processed in each segment is not large, the criterion of "single in and no out" for the OMP algorithm to select atoms is relatively simple. In contrast, the stage-wise orthogonal matching pursuit (StOMP) algorithm[24] and regularized orthogonal matching pursuit (ROMP) algorithm[25] use the criterion of "multi-in and multi-out", their superiority lies in dealing with big data recovery.

The initial residual $\boldsymbol{R}^1 = \boldsymbol{y}_t$, the orthogonal set of selected atoms is $\boldsymbol{D} = \boldsymbol{d}_p, p = 1, 2, \ldots, k$, and the atoms are selected according to certain principles in the iterative process to ensure the minimum residual between the sparse representation result and the original signal.

(1) Select the atom corresponding to the maximum result of residual and inner product in turn,

$$\boldsymbol{a}_{im} \text{ s.t.} |\langle \boldsymbol{R}^m, \boldsymbol{a}_{im} \rangle| = \max_{i \in \Gamma} |\langle \boldsymbol{R}^m, \boldsymbol{a}_i \rangle|, \tag{8}$$

where $m = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, n$.

(2) The above atoms are orthogonalized by Schmidt orthogonalization method,

$$\boldsymbol{d}_m = \boldsymbol{a}_{im} - \sum_{p=0}^{m-1} \frac{\langle \boldsymbol{a}_{im}, \boldsymbol{d}_p \rangle}{\|\boldsymbol{d}_p\|^2} \boldsymbol{d}_p. \tag{9}$$

(3) Normalize $\boldsymbol{d}_m$,

$$\boldsymbol{d}_m = \frac{\boldsymbol{d}_m}{\|\boldsymbol{d}_m\|^2}. \tag{10}$$

(4) Projecting the residual component $\boldsymbol{R}^m$ onto $\boldsymbol{d}_m$,

$$\boldsymbol{R}^m = \langle \boldsymbol{d}_m, \boldsymbol{R}^m \rangle \boldsymbol{d}_m + \boldsymbol{R}^{m+1}. \tag{11}$$

The residual component continues to be decomposed until it is less than a certain predetermined threshold. All the selected atoms constitute the atomic set $\boldsymbol{D}$, we can obtain a preliminary signal sparse representation using this atomic set. The signal is expressed as follow:

$$\boldsymbol{y}_t = \sum_{i=0}^{k-1} \langle \boldsymbol{d}_i, \boldsymbol{R}^i \rangle \boldsymbol{d}^i + \boldsymbol{R}^k = \boldsymbol{D}\boldsymbol{s} + \boldsymbol{R}^k, \tag{12}$$

where $\boldsymbol{s}$ is the representation coefficient.

3.4. Dictionary update

All the selected atoms are updated one by one using the K-SVD algorithm for global optimum. The atoms and the corresponding coefficients are updated simultaneously.

Choose a column $\boldsymbol{d}_k$ (i.e., the k-th atom) in the dictionary and the k-th row of the coefficient matrix $\boldsymbol{S}$ corresponding to $\boldsymbol{d}_k$, which is recorded as $\boldsymbol{s}_T^k$ (not the k-th vector $\boldsymbol{s}_k$ in $\boldsymbol{S}$). Equation $\|\boldsymbol{Y} - \boldsymbol{D}\boldsymbol{S}\|_F^2$ can be written as

$$\begin{aligned} \|\boldsymbol{Y} - \boldsymbol{D}\boldsymbol{S}\|_F^2 &= \left\| \boldsymbol{Y} - \sum_{j=1}^{K} \boldsymbol{d}_j \boldsymbol{s}_T^j \right\|_F^2 \\ &= \left\| \left(\boldsymbol{Y} - \sum_{j \ne k}^{K} \boldsymbol{d}_j \boldsymbol{s}_T^j \right) - \boldsymbol{d}_k \boldsymbol{s}_T^k \right\|_F^2 \\ &= \left\| \boldsymbol{E}_k - \boldsymbol{d}_k \boldsymbol{s}_T^k \right\|_F^2. \end{aligned} \tag{13}$$

Singular value decomposition is used to solve $\boldsymbol{d}_k$ and $\boldsymbol{s}_T^k$ in the above equation.

Define ω_k as the indicator set that points to the atom $\boldsymbol{d}_k$ in sample $\boldsymbol{y}_t$, namely, the non-zero position in $\boldsymbol{s}_T^k(i)$, $\omega_k = i | 1 \le i \le K, \boldsymbol{x}_T^k \ne 0$. Define matrix $\boldsymbol{\Omega}_k$ with dimension $N \times \omega_k$. The elements $(\omega_k(i), i)$ in the matrix are set to 1 and the others are set to 0. The row vector $\boldsymbol{s}_k^{\mathrm{T}}$ is shrunk with discarding the nonzero elements by multiplying $\boldsymbol{s}_k^{\mathrm{T}} \omega_k$. Similarly, the subset of the examples that are currently employing the $\boldsymbol{d}_k$ atom is denoted by $\boldsymbol{Y}\boldsymbol{\Omega}_k$.

Error matrix $\boldsymbol{E}_k^R = \boldsymbol{E}_k \boldsymbol{\Omega}_k$ denotes the error columns equivalent to the examples that employ the subset of $\boldsymbol{d}_k$, the minimization problem as mentioned before becomes

$$\|\boldsymbol{E}_k \boldsymbol{\Omega}_k - \boldsymbol{d}_k \boldsymbol{s}_T^k \boldsymbol{\Omega}_k\|_F^2 = \|\boldsymbol{E}_k^R - \boldsymbol{d}_k \boldsymbol{s}_R^k\|_F^2. \tag{14}$$

The error matrix $\boldsymbol{E}_k^R$ is decomposed into $\boldsymbol{E}_k^R = \boldsymbol{U} \triangle \boldsymbol{V}^{\mathrm{T}}$ using SVD algorithm. The atom $\boldsymbol{d}_k = \boldsymbol{u}_1$ and the coefficient $\boldsymbol{s}_R^k = \triangle(1,1) \cdot \boldsymbol{v}_1$ are both updated.

We update all the atoms in the atomic set $\boldsymbol{D}$ sequentially from the first column to the last column, and the optimized dictionary $\hat{\boldsymbol{D}}$ and the sparsity coefficient $\hat{s}$ are obtained. They are globally optimal and can represent the original signal better. The sparse representation is described as

$$\boldsymbol{y}_t' \approx \hat{\boldsymbol{D}}\hat{\boldsymbol{s}}. \tag{15}$$

3.5. Signal recovery

The sparse coefficients and optimized dictionary are obtained by sparse decomposition of $\boldsymbol{y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_L]$ after steps 3 and 4. We obtain the reconstructed signal $\boldsymbol{y}' = [\boldsymbol{y}_1', \boldsymbol{y}_2', \ldots, \boldsymbol{y}_L']$, each column is set as a signal segment, all signal segments are merged and mean processing is performed in the overlapping area. The obtained signal is the sparse denoised signal.

4. Algorithm analysis

Based on the matrix $\boldsymbol{A}$, the received signal can be rewritten as $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{s} + \boldsymbol{v}$ when the subscript is omitted. We discuss why the number of selected atoms in $\boldsymbol{A}$ is reduced when the noise increases. In OMP algorithm, the residual of the k-th step can be expressed as[20]

$$\begin{aligned} \boldsymbol{r}_k &= \boldsymbol{y} - \boldsymbol{p}_k \\ &= \boldsymbol{y} - \boldsymbol{A}_{\Gamma_k}\boldsymbol{A}_{\Gamma_k}'\boldsymbol{y} \\ &\cong (\boldsymbol{I} - \boldsymbol{W}_k)\boldsymbol{A}\boldsymbol{s} + (\boldsymbol{I} - \boldsymbol{W}_k)\boldsymbol{v} \cong \boldsymbol{q}_k + \boldsymbol{v}_k. \end{aligned} \tag{16}$$

According to the analysis of OMP algorithm in Ref. [20], if the real support set of pure data projected to $\boldsymbol{A}$ is recorded as $\boldsymbol{A}_\Gamma$, then the definition of maximum inner product coefficient can be derived from $\boldsymbol{q}_k$ and $\boldsymbol{v}_k$ in Eq. (16). Let the absolute value of the maximum inner product coefficient of $\boldsymbol{q}_k$ and the real support set be $Q_{t,1} \cong \max\limits_{\boldsymbol{a} \in \boldsymbol{A}_\Gamma} |\boldsymbol{a}^{\mathrm{T}}\boldsymbol{q}_k|$, the maximum inner product coefficient of the interference support set is $Q_{t,2} \cong \max\limits_{\boldsymbol{a} \in \boldsymbol{A} \backslash \boldsymbol{A}_\Gamma} |\boldsymbol{a}^{\mathrm{T}}\boldsymbol{q}_k|$, the inner product coefficient of the noise term and support set is $J_t \cong \max\limits_{\boldsymbol{a} \in \boldsymbol{A}} |\boldsymbol{a}^{\mathrm{T}}\boldsymbol{v}_k|$. Substituting $\boldsymbol{r}_k = \boldsymbol{q}_k + \boldsymbol{v}_k$ and using the absolute value inequality,[20] the sufficient condition for choosing the right supporting atom is

$$\max_{\boldsymbol{a} \in \boldsymbol{A}_\Gamma} \left|\boldsymbol{a}^{\mathrm{T}}\boldsymbol{r}_k\right| \geq Q_{t,1} - J_t > Q_{t,2} + J_t \geq \max_{\boldsymbol{a} \in \boldsymbol{A} \backslash \boldsymbol{A}_\Gamma} \left|\boldsymbol{a}^{\mathrm{T}}\boldsymbol{r}_k\right|. \tag{17}$$

It can be seen that the atoms that can be correctly selected must satisfy $Q_{t,1} - Q_{t,2} > 2J_t$. Without loss of generality, the modules of N' coefficients in $\boldsymbol{s}$ are recorded as $|s_1| \geq |s_2| \geq \cdots \geq |s_{N'}|$ according to the size relation. And only $|s_1|, \ldots, |s_l|$ satisfy $Q_{t,1} - Q_{t,2} > 2J_t$. Since $J_t = \sqrt{N'\sigma^2}\cos\theta$, it is obvious that the increase of σ^2 will reduce the number of correct atoms that can be selected. Introducing incorrect atoms will lead to performance deterioration. We reduce the number of selected atoms when the noise intensity increases, so as to ensure that all the selected atoms are "correct".

In assessing complexities, the algorithm is composed of iterating on three main stages: sparse coding (OMP process), dictionary update, and final averaging process. As mentioned in Ref. [29], both stages can be done efficiently in $\mathrm{O}(JM'kL_{nz})$ per segment, where J is the number of iterations in updating dictionary, M' is the size of the segment, k is the number of atoms in the dictionary, and L_{nz} is the maximal number of nonzero elements in each coefficient vector.

5. Simulation

The denoising of chaotic signals should not only be reflected in the improvement of SNR, but also in the maximum restoration of the characteristic parameters of the chaotic signals. In this section, the sparse decomposition based denoising (SDD) algorithm is compared with WT algorithm, LCF algorithm, and EMD algorithm from the aspect of SNR improvement, root mean square error (RMSE), and phase diagram recovery. The nonlinear time series data are generated from the following Lorenz system equations:[26]

$$\begin{cases} \mathrm{d}x/\mathrm{d}t = \alpha(y - x), \\ \mathrm{d}y/\mathrm{d}t = \beta x - xz - y, \\ \mathrm{d}z/\mathrm{d}t = -\gamma z + xy, \end{cases} \tag{18}$$

where $\alpha = 10$, $\beta = 28$, $\gamma = 8/3$. The equation is solved by the fourth-order Runge–Kutta method with step size of 0.01. In each experiment, the equation iterates from the random position of the chaotic area, and the state variable x is used as the chaotic signal, the sample length is $N = 8000$. The SNR and RMSE are defined as

$$\mathrm{SNR} = 10\log_{10} \frac{\sum\limits_{n=1}^{N} x^2(n)}{\sum\limits_{n=1}^{N} (\hat{x}(n) - x(n))^2}, \tag{19}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} [\hat{x}(n) - x(n)]^2}. \tag{20}$$

We add -5 dB to 25 dB noise to $x(n)$, and compare all these algorithms in SNR improvement. The parameter settings are: the moving length of window is $M = 20$, constant $\omega = 2.5$, the dimension of matrix $\boldsymbol{A}$ is 41×61; WT algorithm uses $J = 4$ level decomposition and hard threshold; LCF algorithm uses polynomial order $l = 3$ and const. $+x + x^2 + x^3$, the block length is $2\Delta + 1 = 17$; and EMD algorithm uses the number of iterations 9 and IMF layers 10.

Figure 2(a) shows the output SNR obtained by different algorithms when the input SNR is between -5 dB and 25 dB.

The SDD algorithm is about 2.35 dB higher than the other three in average, especially at lower SNR, because the SDD algorithm selects less atoms to avoid noise when the noise intensity increases. In essence, this SDD algorithm is an extension of the local curve fitting method, but it can use different atoms in different segments of the signal. Therefore, the consistency of the SDD algorithm is better than the local curve fitting algorithm. Figure 2(b) shows the RMSE of these algorithms, the SDD algorithm also shows better performance than the other three, especially when the SNR is lower. It is consistent with the SNR performance in Fig. 2(a).

In order to reflect these differences more intuitively, figure 3 shows the attractor flow patterns reconstructed by various methods when the input SNR is 15 dB. It can be seen that the reconstructed flow pattern becomes smoother around the attractor by SDD. The other three methods show a rough transfer in some areas.

The proliferation exponent (PE)[27] is used here to analyze the chaotic characteristics of the reconstructed signal. Although the Lyapunov exponent (LE)[28] can better reflect the chaotic properties of signals from the definition of chaos, its calculation method for continuous chaotic signals is based on the Jacobian matrix of the definition, which is not suitable for the calculation of a large number of sample data. The proliferation exponent can largely reflect the chaos characteristics of the original signal when the input sample data is large enough. The proliferation exponent is defined as

$$\mathrm{PE}_{\alpha} = \frac{D(V_{\mathrm{s}}(t))}{E^2(V_{\mathrm{s}}(t))}, \tag{21}$$

where $V_{\mathrm{s}}(t)$ is the phase space distance function,[28] and α is the control factor.

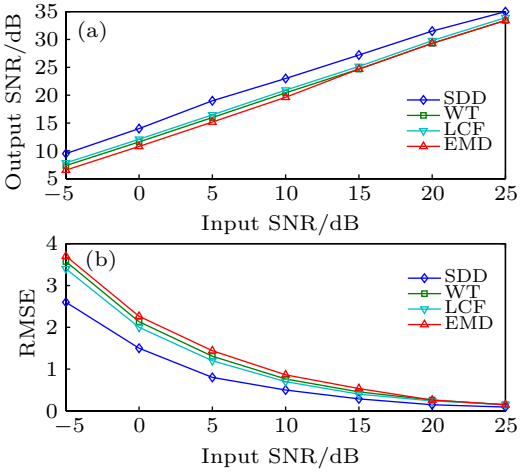


Fig. 2. (a) SNR improvement and (b) RMSE of different denoising algorithms under different noise intensities.

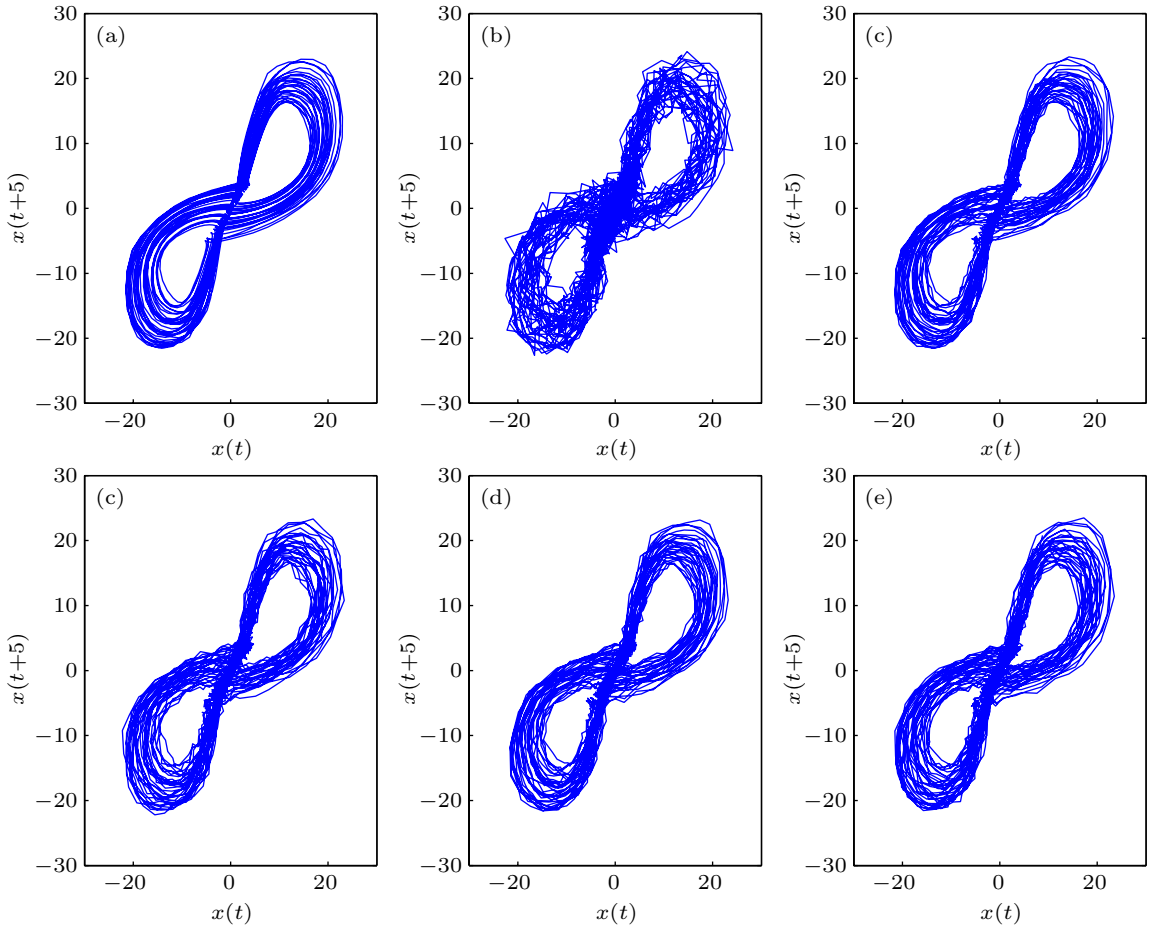


Fig. 3. Comparison of the reconstructed attractors by different denoising algorithms: (a) the original pure data, (b) the signal after noise pollution, and (c)–(f) the signal recovered by SDD, WT, LCF, and EMD, respectively.

Here we use a Lorenz signal with length of 8000 and sampling interval of 0.05, the same simulation parameter settings as the above. In terms of selecting the value of control factor α, it should be noted that since the selection of α takes the sampling step into account, that is, the optimal α value is set based on the sampling step to best reflect the chaotic characteristic. The simulation results with different α values may be slightly different. Here we set $\alpha = 6$ and calculate the PE of the restored signal by different algorithms. As shown in Fig. 4, the PE value by SDD algorithm is about 0.56 in the whole interval, and the PE value by the original data is 0.47, while the PE value obtained by WT, LCF, and EMD algorithms have greater deviation, which are 0.64, 0.62, and 0.65, respectively.

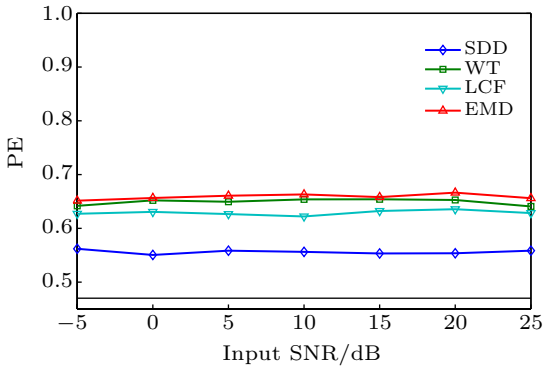


Fig. 4. Comparison of proliferation exponents of the restored signal by different denoising algorithms.

6. Conclusion

Based on the theory of sparse decomposition and K-SVD, we propose a chaotic signal denoising algorithm with dynamic sparsity determination, dictionary constructed by signal iteration equation and K-SVD optimization. The special point of this algorithm is that the constructed dictionary over complete projection matrix has sufficient frequency-domain components, so that it can better represent each segment of the signal. The K-SVD optimization process further improves the noise reduction. The simulation results show that the algorithm has better SNR improvement and RMSE than the other three denoising algorithms, especially in the case of low SNR. For the PE value calculation which characterizes the chaotic signal, the algorithm in this paper has better performance than the other three denoising algorithms.

References

[1] Alamodi A O A, Sun K H, Ai W, Chen C and Peng D 2019 *Chin. Phys. B* **28** 020503
[2] Bi H Y, Qi G Y, Hu J B and Wu Q L 2020 *Chin. Phys. B* **29** 020502
[3] Gao F, Li W Q, Tong H Q and Li Q L 2019 *Chin. Phys. B* **28** 090501
[4] Wang W B, Zhang X D, Chang Y C, Wang X L, Wang Z, Chen X and Zheng L 2015 *Chin. Phys. B* **25** 010202
[5] Yang H, Li Y A and Li G H 2015 *Acta Armamentarii* **36** 2330 (in Chinese)
[6] Liu K, Li H, Dai X C and Xu P X 2008 *JEIT* **30** 1849 (in Chinese)
[7] Wang M J, Zhou Z Q, Li Z J and Zeng Y C 2018 *Acta Phys. Sin.* **67** 060501 (in Chinese)
[8] Lv S X and Feng J C 2013 *Acta Phys. Sin.* **62** 230503 (in Chinese)
[9] Wang W B, Jin Y Y, Wang B, Li W G and Wang X L 2018 *Acta Electron. Sin.* **46** 1652 (in Chinese)
[10] Hai S, Gao H W and Ruan X J 2016 *Int. J. Signal Process., Image Process. and Pattern Recognit.* **9** 219
[11] Gao J B, Sultan H, Hu J and Tung W W 2010 *IEEE Signal Process. Lett.* **17** 237
[12] Tung W W, Gao J B, Hu J and Yang L 2011 *Phys. Rev. E* **83** 046210
[13] Wu Z H and Huang N E 2010 *Adv. Adap. Data Anal.* **2** 397
[14] Wang M J, Zhou Z Q, Li Z J and Zeng Y C 2019 *Circ. Syst. Signal Proces.* **38** 2471
[15] Donoho D L, Elad M and Temlyakov V N 2006 *IEEE Trans. Inf. Theory* **52** 6
[16] Tseng P 2009 *IEEE Trans. Inf. Theory* **55** 888
[17] Dumitrescu B and Irofti P 2017 *IEEE Signal Proces. Lett.* **24** 309
[18] Karlovitz L A 1970 *J. Appr. Theory* **3** 123
[19] Chen S S, Donoho D L and Saunders M A 2001 *SIAM Rev.* **43** 129
[20] Cai T T and Wang L 2011 *IEEE Trans. Inf. Theory* **57** 4680
[21] Karahanoglu N B and Erdogan H 2012 *Digital Signal Proces.: A Rev. J.* **22** 555
[22] Tian J and Chen L 2012 *IEEE Signal Proces. Lett.* **19** 395
[23] Chen P, Rong Y, Nordholm S and He Z Q 2017 *IEEE Trans. Vehic. Technol.* **66** 10567
[24] Lee D J 2016 *IEEE Commun. Lett.* **20** 2115
[25] Needell D and Vershynin R 2010 *IEEE J. Sel. Top. Signal Proces.* **4** 310
[26] Chowdhury M S H, Hashim I and Momani S 2009 *Chaos Soliton. Fract.* **40** 1929
[27] Lv S X, Wang Z S, Hu Z H and Feng J C 2014 *Chin. Phys. B* **23** 010506
[28] Holger K and Thomas S 2004 *Nonlinear Time Series Analysis* (Cambridge: Cambridge University Press) pp. 65–74
[29] Elad M and Aharon M 2006 *IEEE Trans. Image Proces.* **15** 3736