Research Article

A Novel Improved Threshold Adaptive Forgetting Variable Step Size Blind Separation Model for Weak Signal Detection

Tianyi Yu [ID],[1] **Houming Wang,**[1] **Shunming Li,**[1] **Jiantao Lu** [ID],[1] **Siqi Gong,**[1] **Yanfeng Wang,**[2] **and Guangrong Teng**[2]

[1]*College of Energy and Power Engineering Nanjing University of Aeronautics and Stronautics, Nanjing, China*
[2]*Laboratory of AECC Sichuan Gas Turbine Establishment, Mianyang, China*

Correspondence should be addressed to Jiantao Lu; lujt@nuaa.edu.cn

The online blind source separation (BSS) is seriously disturbed by strong noise when extracting weak signals and has the defects that it cannot both give consideration to convergence speed and steady-state error. In order to solve the abovementioned problems, a novel improved threshold adaptive forgetting variable step size blind separation model (ITAFBS) for weak signal detection is proposed. Firstly, an improved lifting wavelet transform (ILWT) is proposed to reduce the noise of weak signals. In ILWT, a threshold function containing an adjustment factor is proposed to reduce the constant deviation so as to ensure a high signal-to-noise ratio and low distortion after denoising. Then, the separation index (SI) is constructed according to the convergence conditions of the BSS model. An adaptive variable step size blind separation model based on the SI is studied. At the initial stage of separation, the step size is increased to obtain a fast convergence rate, and at the end of separation and the step size is shortened to obtain a small steady-state error. Finally, the forgetting factor is introduced into the model to reduce the error accumulation in the early stage of the algorithm, and the Fourier norm is introduced to improve the convergence speed and separation accuracy of the model. The simulation and experimental results show that ITAFBS has a good performance in multi-frequency weak signal detection. Compared with other methods, the ITAFBS has a faster convergence speed and minimum steady-state error.

1. Introduction

Due to the complex and harsh load environments, the rotating parts (such as bearings and gears) in the rotating machinery (such as aerospace, wind equipment, automobiles, and ships) are very prone to fatigue and damage. In the early stage of the failure, the target signal is usually interfered with by strong noise, and the signal-to-noise ratio (SNR) is very low [1]. In the actual operation of the equipment, the energy of the target signal is continuously attenuated due to the complex transmission path and the interference noise. It becomes more difficult for fault signals to be detected. If the early faults of rotating parts are identified timely and accurately, the corresponding treatment will be made in time. There is a great significance to ensure the safe operation of mechanical equipment and reduce the frequency of production accidents [2].

Aiming at the characteristics of early weak fault signals, such as nonstationary and low SNR, a lot of research work have been carried out and fruitful results have been achieved by researchers. In the 1990s, Sweetens and Daubechies decomposed all wavelet transform into lifting process, and lifting wavelet threshold denoising method based on lifting wavelet transform (LWT) was proposed [3]. LWT realizes the simplification of the wavelet decomposition and reconstruction process and shortens the calculation time [4]. In addition, compared with the first generation wavelet, the lifting wavelet can also construct the wavelet adaptively according to different signal characteristics. The biggest advantages of LWT are simple structure, low distortion, reduced aliasing effects, and high computational efficiency [5]. After continuous improvement, LWT has been successfully applied in fault diagnosis [6], radar detection [7], spectral analysis [8], aerospace [9], and other fields.

The key point of LWT noise reduction is the selection of threshold function. Aiming at the problem that hard threshold function may produce oscillation point and soft threshold function has constant deviation, the improvement of threshold function has been studied deeply by scholars [10]. Li W. and Wang [11] proposed a modular square processing method of threshold functions. The method has good overall continuity but poor flexibility. The closer to the threshold, the greater the constant deviation. And there is a risk of loss of high-frequency components in noise reduction. Therefore, when the threshold function is improved in this paper, the problem of constant deviation will be fully considered so that the noise reduction effect is ideal and the distortion is low.

Blind source separation (BSS) is a kind of signal processing technology rapidly rising in the field of signal processing in the late 1990s. It cannot only obtain the target signal but also obtain every interference noise source, which is convenient for deeper noise source location and feature extraction. It is widely used in acoustic signal processing [12], wireless sensor signal transmission [13], biomedical engineering [14], and fault diagnosis [15]. BSS can be divided into online methods and offline methods [16]. In general, to achieve faster convergence, all data are used in offline methods during each iteration. However, the disadvantage of offline BSS is that a large storage capacity is required, so offline BSS is not suitable for time-varying situations. In contrast, in the online method, only a small amount of data storage is required. Online BBS can dynamically update the decomposition matrix in real-time based on current data, so it has been successfully applied to time-varying situations. At present, the most common online BSS are natural gradient algorithm [17] and the equivariant adaptive separation for independence (EASI) algorithm [18].

The convergence rate and steady-state error are used to evaluate the performance of online BSS. The faster the convergence speed, the stronger time-varying tracking ability. The smaller steady-state error, the higher convergence accuracy. In the process of signal separation, the two indicators are contradictory. But the convergence rate and steady-state error are related to the learning step size. Increasing the step size can improve the convergence speed and time-varying tracking ability of the separation model. Decreasing the step size can obtain higher convergence accuracy. Therefore, the balance ability of step size to convergence rate and steady-state error will be fully considered when designing the online BSS model.

In summary, considering the noise reduction ability of LWT and the separation ability of BSS for multi-frequency signals, a novel improved threshold adaptive forgetting variable step size blind separation model (ITAFBS) for weak signal detection is proposed. Firstly, the improved lifting wavelet transform (ILWT) is used to denoise the weak signal to reduce the interference of strong noise to the separation model. The AFBS is then applied to the preprocessed mixed signals. Fast and high precision separation can be achieved by adjusting step size adaptively and introducing the forgetting factor and Fourier norm. In simulation and experiment, the performance of ITAFBS and its application in

mechanical field and instrument field are verified by comparing it with other comparison algorithms.

The main innovations are as follows:

(1) A novel improved threshold adaptive forgetting variable step size blind separation model (ITAFBS) for weak signal detection is proposed. The ITAFBS has the ability of fast and accurate feature extraction for weak signals.

(2) In the ITAFBS, an improved threshold equation with adjustment factor and an optimal objective function with SNR and RMSE as indexes are proposed. The adjustment factor can handle the constant deviation problem flexibly. The objective function of the optimal value of adjustment factor can make the denoise signal with high *SNR* and low distortion at the same time.

(3) In the ITAFBS, an adaptively variable step size algorithm is proposed, which adaptively adjusts according to the separation index. At the initial stage of separation, the convergence rate of the model is improved by increasing the step size. When the separation is about to be completed, the steady-state error can be reduced to improve the separation accuracy by shortening the step size.

(4) In the ITAFBS, the forgetting factor and Fourier norm are introduced to further improve the overall convergence speed and separation accuracy of the model.

The organizational structure of the rest of this paper is as follows: Section 2 introduces the ILET and the improved threshold function; Section 3 introduces ITAFBS theory, formula derivation, and implementation steps in detail; Section 4, simulation signals are used to verify ITAFBS performance and compare it with other algorithms; Section 5, acceleration sensor signal and test circuit signal are used to verify ITAFBS performance; and Section 6 summarizes the full text.

2. Improved Lifting Wavelet Transform

2.1. Lifting Wavelet Transform. As one of the most popular frequency domain methods, wavelet transform (WT) has been widely studied in many aspects of signal processing [3]. Since many data sets have a strong correlation in time and frequency, wavelet can use the smallest coefficient to quickly extract the data set. LWT is called the second generation wavelet, it breaks through this limitation. The lifting procedure is a method for not only scheming wavelets but also doing wavelet transform.

The process of LWT includes two parts: forward transform decomposition and inverse transform reconstruction. The decomposition process of the positive transformation includes three stages of decomposition, prediction, and update. The three stages of inverse transform reconstruction are opposite to the former, including antiupdate, antiprediction, and merging [19]. The decomposition and reconstruction process of LWT is shown in Figure 1.
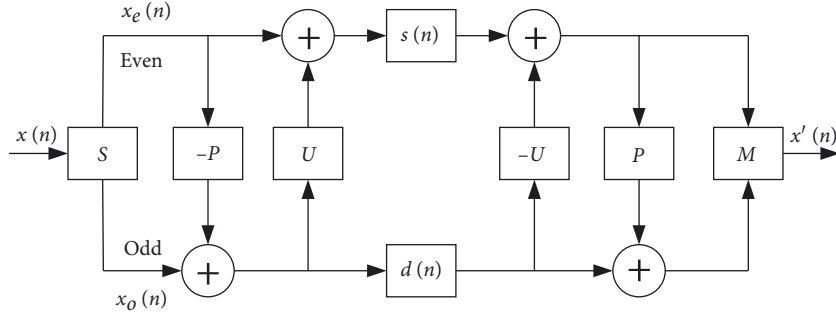
FIGURE 1: LWT process diagram.

In Figure 1, the original signal sequence $x(n)$ is decomposed into an even sequence $x_e(n)$ and an odd sequence $x_o(n)$ by the separation operator S. Because even sequences are related to odd sequences, even sequences are often used to predict odd sequences. The difference $d(n)$ between the odd sequence and the predicted value $P[x_e(n)]$ is a measure of approximation. $d(n)$ is the wavelet coefficient, which can be regarded as the high-frequency component of the original signal sequence. Prediction process: $d(n) = x_o(n) - P[x_e(n)]$. Where P is the prediction operator. The process of using $d(n)$ to adjust even sequences is called the updating process, which makes $s(n)$ contain only the low-frequency components of the original signal sequence. Update process: $s(n) = x_e(n) + U[d(n)]$. Where U is the update operator. On the other hand, the original signal sequence is reconstructed as $x'(n)$ through reverse update, reverse prediction, and merge [20].

There are many kinds of processing methods for the wavelet coefficients obtained by decomposition, such as correlation denoising, wavelet threshold denoising, modulus maximum denoising, and translation invariant denoising. Among them, the wavelet threshold denoising method has the least amount of computation, the highest efficiency, and the most extensive application fields. In the wavelet lifting scheme proposed in this paper, the high-frequency component of the signal to be measured can be decomposed first, and then the low-frequency component can be obtained by constructing the scaling function. The specific threshold denoising process is shown in Figure 2.

In Figure 2, "Input" represents the noisy signal, which is split into odd and even sequences by lazy wavelet transform. "ca1," "ca2", ... "ca$n$," respectively, decompose the low-frequency coefficients of the first layer, the second layer, and the nth layer. "cd1," "cd2", ... "cd$n$," respectively, decompose the high-frequency coefficients of the first layer, the second layer, and the nth layer. "cd1'," "cd2'", ... "cd$n$'," respectively, represent the high-frequency coefficients after threshold processing. "ca1'," "ca2'", ... "ca$n$-1'," respectively, represent the low-frequency coefficients of the reconstructed second layer, first layer, and nth layer. "Output" represents the reconstructed signal after threshold noise reduction processing.

2.2. Improved Lifting Wavelet Transform. Donoho and Johnstone proposed a wavelet threshold denoising method.

The basic idea is based on the distribution characteristics of noise and signal in the wavelet threshold. The classical threshold functions include hard and soft threshold functions [21].

2.2.1. Hard Threshold Function. When the absolute value of wavelet coefficients is less than the given threshold, it is zero. If the absolute value is greater than the threshold value, it will remain unchanged. The expression is as follows:

$$\widehat{w}_{j,k} = \begin{cases} w_{j,k}, & |w_{j,k}| \geq \lambda, \\ 0, & |w_{j,k}| < \lambda. \end{cases} \tag{1}$$

In (1), λ represents the threshold value, and $W_{j,k}$ represents the wavelet coefficients of the original signal decomposed in the first layer. According to (1), the continuity of the hard threshold function is relatively poor, and a jump phenomenon occurs at the threshold, which may cause the signal to oscillate during the inverse transformation process and make the reconstructed signal appear pseudo-Gibbs phenomenon.

2.2.2. Soft Threshold Function. The wavelet coefficients, whose absolute value is less than the threshold, are replaced by zero, and the wavelet coefficients, whose absolute value is greater than the threshold, are reduced by threshold. The expression is as follows:

$$\widehat{w}_{j,k} = \begin{cases} \text{sign}(w_{j,k}) \bullet (|w_{j,k}| - \lambda), & |w_{j,k}| \geq \lambda, \\ 0, & |w_{j,k}| < \lambda. \end{cases} \tag{2}$$

It can be seen from (2) that the soft threshold function has no discontinuity and has good continuity. However, for the wavelet coefficients, when the absolute value of the value is greater than the threshold, the signal will be compressed to a large extent. So that the deviation between $W_{j,k}$ and $W_{j,k}$ is relatively large, and the high-frequency component of the target signal will be lost to a certain extent.

2.2.3. Improved Threshold Function. Aiming at the problem that hard threshold function may produce an oscillation point and the soft threshold function has constant deviation, the classical threshold function is improved. The wavelet
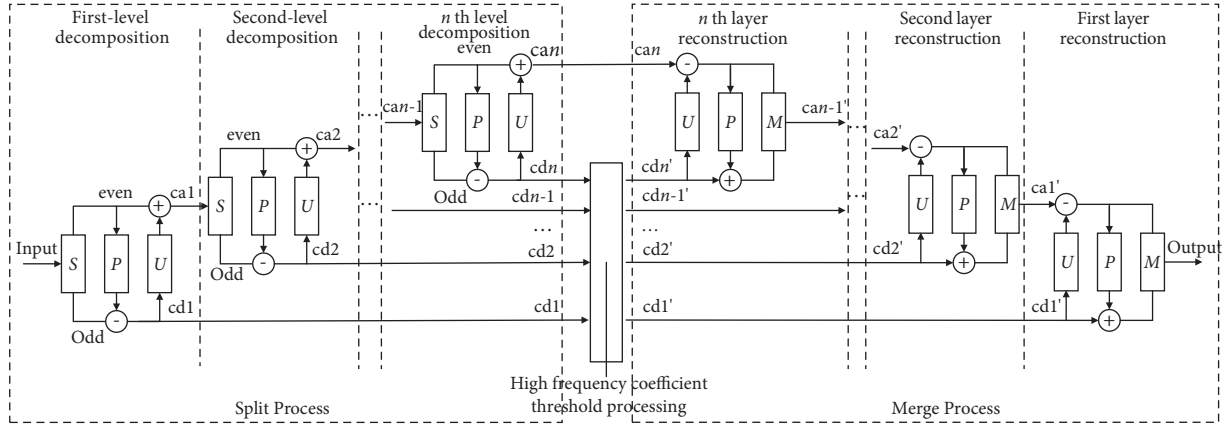
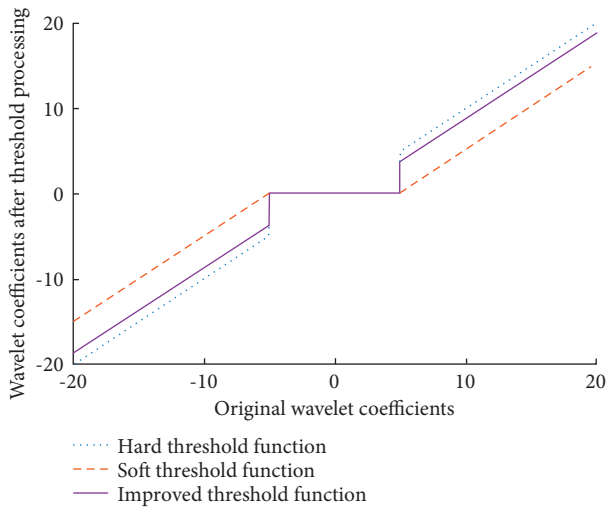FIGURE 2: Denoising process of lifting wavelet threshold.

FIGURE 3: Comparison of threshold function.

coefficients obtained by improved threshold processing are between hard and soft threshold functions, and then the wavelet coefficients closer to the real value are obtained. The expression of the improved threshold function is shown (3)as

$$\widehat{w}_{j,k} = \begin{cases} w_{j,k} - \text{sign}(w_{j,k}) \bullet \lambda (1 - \theta)^2, & |w_{j,k}| \geq \lambda, \\ 0, & |w_{j,k}| < \lambda. \end{cases} \quad (3)$$

In (3), the parameter θ is the adjustment factor, and its value range is [0, 1], and the constant deviation can be flexibly handled by adjusting its value. When the value of θ is close to 0, the improved threshold function is close to the soft threshold function. When the value of θ is close to 1, it is similar to the hard threshold function. The value of θ can be flexibly adjusted according to the actual signal characteristics, which has good adaptability.

By constantly adjusting the value of parameter θ, the improved threshold function can realize both the function of soft and hard threshold functions. The improved threshold function can better improve the shortcomings of hard and soft threshold functions, make the reconstructed signal

smoother, retain the high-frequency components of the target signal, reduce the loss of useful components, and achieve a better noise reduction effect. By setting the threshold $\lambda = 5$, and $\theta = 0.5$, the comparison diagram between the improved, the soft, and the hard threshold functions is obtained, as shown in Figure 3.

The noise reduction performance of three kinds of threshold functions is compared and analyzed by simulation signal. $y = \sin(2 \times \pi \times 20 \times t)$ is used as simulation signal. Gauss white noise with $SNR = -10$ dB is added to the simulation signal. db$N$ wavelet has good regularity. In this paper, considering the smoothness, localization ability, computation amount, and real-time performance, the db5 wavelet is selected as the basis function. The number of decomposition layers is set to three, and general threshold estimation is used to obtain the threshold. The SNR and RMSE of the reconstructed signal are used to evaluate the noise reduction performance of different threshold functions. The definitions of SNR and RMSE are shown in (4) and (5). In (5), $S(n)$ is the original signal and $\overline{S}(n)$ is the reconstructed signal after LWT.

$$SNR = 10 \log \frac{\text{Signal power}}{\text{Noise power}}, \quad (4)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^{N} [S(n) - \overline{S}(n)]^2}. \quad (5)$$

The SNR of the reconstructed signal is higher, and the RMSE is smaller, which indicates that the denoising effect is better and the distortion of the reconstructed signal is lower. By this characteristic, the objective function is constructed and the optimal value of adjustment factor is calculated. First, the adjustment factor is evaluated at 0.1 intervals within the range [0, 1]. Calculate SNR and $RMSE$ after denoising by each adjustment factor threshold. The functions $S(\theta)$ and $R(\theta)$ are obtained by fitting these points. The normalized functions $\widetilde{S}(\theta)$ and $\widetilde{R}(\theta)$ are obtained by using the maximum and minimum values of $S(\theta)$ and $R(\theta)$ to normalize. According to the abovementioned characteristics, the adjustment factor is the optimal value when the sum

of the distances from $\widetilde{S}(\theta)$ to max $\widetilde{S}$ and $\widetilde{R}(\theta)$ to min $\widetilde{R}$ is the smallest. The objective function is expressed as minimize$_{\theta}\{[\max\widetilde{S} - \widetilde{S}(\theta)] + [\min\widetilde{R} - \widetilde{R}(\theta)]\}$. Table 1 shows the comparison results of SNR and RMSE after denoising with different threshold functions.

According to Table 1, it can be seen that the improved threshold method cannot only obtain the highest SNR but also its RMSE is smallest. Therefore, the improved threshold function has a better denoising effect than soft and hard threshold functions, and the distortion of reconstructed signal is the lowest.

3. Improved Threshold Adaptive Forgetting Variable Step Size Blind Separation Model

3.1. Adaptive Forgetting Variable Step Size Blind Separation. Assuming that there are n signal sources, $s(t) = (s_1(t), s_2(t), \ldots, s_n(t))^T$ are the source signal vector and $x(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T$ are the observation signal vector. It is assumed that the time delay in the signal transmission process can be ignored, and the mixing method of each source signal collected by the sensor is linear. Based on these assumptions, a linear instantaneous mixture model of signal-to-noise is constructed. Its mathematical model can be expressed as

$$\mathbf{x}_i(t) = \sum_{j=1}^{n} a_{ij}\mathbf{s}_j(t) + \mathbf{K}_i(t), i = 1, 2, 3 \ldots n. \tag{6}$$

In (6), a_{ij} represents the mixing coefficient, and $\mathbf{K}_i(t)$ represents Gaussian white noise in a typical case. The matrix method is expressed as follows:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{K}(t), \tag{7}$$

where A is the full-rank $n \times n$ mixing matrix, each element in the matrix A represents the mixing coefficient a_{ij} of signal and noise. Generally, it is assumed that the source signals are statistically independent. Without any loss of generality, they are also assumed to have zero mean, and at most only one satisfies the Gaussian distribution. When the interference noise can be ignored or has been reduced to be negligible by noise reduction methods, the model can be simplified as shown in the following formula as follows:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t). \tag{8}$$

Therefore, the separation process of BSS is to find the optimal solution of the separation matrix B that satisfies the following function as

$$\mathbf{y}(t) = \mathbf{B}\mathbf{x}(t). \tag{9}$$

The separated signal vector $y(t)$ represents the estimated signal of the source signal $s(t)$. Figure 4 is the structure of the feedforward adaptive processor, which is used to solve the matrix B.

The purpose of online BSS is to adaptively adjust the separation matrix B to obtain an estimated signal vector $y(t)$, that is, similar to the source signal vector $s(t)$ as much as possible. Because the independent component analysis results have ranking uncertainties and amplitude

Table 1: Comparison of SNR and RMSE after denoising with different threshold functions.

Threshold functions	SNR (dB)	RMSE
Original signal	−10	2.256
Hard threshold method	−3.89	1.386
Soft threshold method	−2.192	1.012
Modular squared processing method	−2.851	1.175
Improved threshold method($\theta = 0.31$)	−0.897	0.784

uncertainties, the separation results are not unique. But it will not adversely affect the identification of the source signal. The whole process is to optimize the separation matrix B so that it can satisfy

$$\mathbf{BA} = \mathbf{P}\boldsymbol{\Lambda}. \tag{10}$$

In (10), P represents a generalized permutation matrix, and Λ represents a diagonal matrix.

In the process of BSS, the independence of separated signals is often regarded as the objective function of updating the separation matrix, and mutual information is often used as a criterion for evaluating the independence of separated signals. Based on the criterion of mutual information minimization, the natural gradient algorithm can be used as a learning algorithm for the separation matrix. The updated formula of the natural gradient algorithm to obtain the separation matrix B is shown as follows [17]:

$$\mathbf{B}_{t+1} = \mathbf{B}_t + \mu_t\left[\mathbf{I} - \boldsymbol{\psi}(\mathbf{y}_t)\mathbf{y}_t^T\right]\mathbf{B}_t. \tag{11}$$

In (11), μ_t represents the iteration step factor, B_t represents the tth iteration of the separation matrix, $\mathbf{I}$ represents the identity matrix, $\boldsymbol{\psi}(\cdot)$ is the activation function and the choice of $\boldsymbol{\psi}(\cdot)$ depends on the probability distribution of the source signal. In order to remove the second-order correlation between the original channel data and make the second-order statistical independence between the components, it is necessary to spheroidize the observation signal [22]. The online adjustment formula of the weight matrix of the spherical network can be derived by natural gradient method, as shown in the following:

$$\mathbf{V}_{t+1} = \mathbf{V}_t - \mu_t\left[\mathbf{z}_t\mathbf{z}_t^T - \mathbf{I}\right]\mathbf{V}_t. \tag{12}$$

In (12), $z(t)$ represents the observation signal after spheroidization. The separation process of blind source separation can be decomposed into two steps. The first step is adaptive updating of spheroidized matrix V. The second step is the adaptive update of the orthogonal normalization matrix U. So the separation matrix B can be expressed as the product of a spheroidized matrix V and an orthogonal normalized matrix U. The update equation of the orthogonal normalized matrix U is expressed as follows:

$$\begin{aligned}\mathbf{U}(t + 1) &= \mathbf{U}(t) - \mu_t\left[\boldsymbol{\psi}(\mathbf{y}_t)\mathbf{y}_t^T - \mathbf{y}_t\boldsymbol{\psi}^T(\mathbf{y}_t)\right]\mathbf{U}(t) \\ &= \left\{\mathbf{I} - \mu_t\left[\boldsymbol{\psi}(\mathbf{y}_t)\mathbf{y}_t^T - \mathbf{y}_t\boldsymbol{\psi}^T(\mathbf{y}_t)\right]\right\}\mathbf{U}(t).\end{aligned} \tag{13}$$

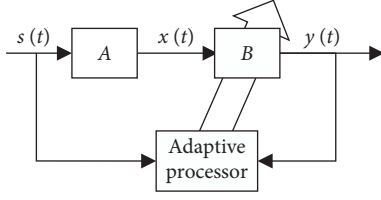Due to the separation matrix $B = UV$, combining the abovementioned two steps into one step, and (14) can get as

FIGURE 4: Structure diagram of feedforward adaptive processor.

$$\mathbf{B}(t+1) = \mathbf{U}(t+1)\mathbf{V}(t+1)$$
$$= \left\{\mathbf{I} - \mu_t\left[\mathbf{\psi}(\mathbf{y}_t)\mathbf{y}_t^T - \mathbf{y}_t\mathbf{\psi}^T(\mathbf{y}_t)\right]\right\}\mathbf{U}(t)\left[\mathbf{I} - \mu_t\left(\mathbf{z}_t\mathbf{z}_t^T - \mathbf{I}\right)\right]\mathbf{V}_t. \tag{14}$$

Ignore the μ_t^2 in the (14), the expression of the equivariant adaptive separation for independence (EASI) can be obtained [23, 24], which can be simplified as

$$\mathbf{B}_{t+1} = \mathbf{B}_t + \mu_t\left[\mathbf{I} - \mathbf{y}_t\mathbf{y}_t^T + \mathbf{y}_t\mathbf{\psi}^T(\mathbf{y}_t) - \mathbf{\psi}(\mathbf{y}_t)\mathbf{y}_t^T\right]\mathbf{B}_t. \tag{15}$$

When the EASI reaches a steady state, it must be satisfied.

$$E\left[\mathbf{I} - \mathbf{y}_t\mathbf{y}_t^T + \mathbf{y}_t\mathbf{\psi}^T(\mathbf{y}_t) - \mathbf{\psi}(\mathbf{y}_t)\mathbf{y}_t^T\right] = \mathbf{0}. \tag{16}$$

$E(\cdot)$ is taking the mean. This requirement can be divided into two parts.

$$D_1 = E\left[\mathbf{I} - \mathbf{y}_t\mathbf{y}_t^T\right] = \mathbf{0}, \tag{17}$$

$$D_2 = E\left[\mathbf{y}_t\mathbf{\psi}^T(\mathbf{y}_t) - \mathbf{\psi}(\mathbf{y}_t)\mathbf{y}_t^T\right] = \mathbf{0}. \tag{18}$$

The first condition (17) is used to ensure that the components are orthogonal, but they cannot guarantee mutual independence. The second condition (18) is used to ensure the independence of each component. Therefore, when the algorithm reaches convergence, it needs to be satisfied:

$$\|D_1\| = 0, \tag{19}$$

$$\|D_2\| = 0. \tag{20}$$

According to the steady-state condition of EASI, the separation index is defined as

$$\zeta = \max\left(\left\|E\left[\mathbf{I} - \mathbf{y}_t\mathbf{y}_t^T\right]\right\|, \left\|E\left[\mathbf{y}_t\mathbf{\psi}^T(\mathbf{y}_t) - \mathbf{\psi}(\mathbf{y}_t)\mathbf{y}_t^T\right]\right\|\right). \tag{21}$$

$E(\cdot)$ is taking the mean. ζ can be used as an index to measure the degree of signal separation. In the initial stage of separation, the value of ζ is relatively large. At this time, the degree of separation of the mixed signal is relatively low, and the value of the step size should be increased to improve the convergence speed. As the separation process continues, the estimated signal is more and more similar to the original signal. At this time, the value of ζ is relatively small, which indicates that a better signal separation effect is obtained at this time. The value of step size should be reduced to reduce the steady-state error of the algorithm.

In (19)–(21), $\|\cdot\|$ represents Frobenius norm (F-norm) of the matrix. F-norm refers to the sum of the squares of the elements in the matrix and then finding the square root. Compared with 1-norm and 2-norm, from the perspective of learning theory, F-norm can prevent the occurrence of over-fitting and improve generalization ability. In addition, from the perspective of optimization or numerical calculation, as a convex function, F-norm is convenient for derivative operation, so as to improve the stability and efficiency of solution operation and help to solve the problem of matrix inversion in some cases.

For online BSS, the step size needs to be updated online adaptively, so the method of obtaining the separation index ζ should be online updated. In order to obtain the online update of ζ, D_1 and D_2 should satisfy the online adaptive update. Let

$$\begin{cases} R_L = E\left[\mathbf{y}_t\mathbf{y}_t^T\right], \\ R_H = E\left[\mathbf{\psi}(\mathbf{y}_t)\mathbf{y}_t^T\right]. \end{cases} \tag{22}$$

According to (18), only the real-time online update of R_L and R_H is needed. In the initial stage of signal separation, the similarity between the estimated signal and the source signal is relatively low. With the continuous progress of the separation process, the estimated signal and the initial signal are more and more similar, and the steady-state error is also smaller and smaller. Errors will accumulate gradually during iteration. In order to reduce the accumulation of errors at the initial stage of separation, a real-time online update method with the forgetting factor is designed [25, 26]. According to (19), each element is updated by calculating the mean value. The update method is shown as

$$\begin{cases} \overline{R}_L(t) = \dfrac{\left[\mathbf{y}_t\mathbf{y}_t^T + \eta\Gamma_\eta(t-1)\overline{R}_L(t-1)\right]}{\Gamma_\eta(t)}, \\[3mm] \overline{R}_H(t) = \dfrac{\left[\mathbf{\psi}(\mathbf{y}_t)\mathbf{y}_t^T + \eta\Gamma_\eta(t-1)\overline{R}_H(t-1)\right]}{\Gamma_\eta(t)}. \end{cases} \tag{23}$$

In (23), $\overline{R}_L(t)$ represents the forgotten data of the tth update of R_L, and $\overline{R}_H(t)$ represents the forgotten data of the tth update of R_H. η represents the forgetting factor, and its value range is (0, 1). When η is close to 1, the forgetting speed is slow, and more attention is paid to historical data. When η approaches 0, the forgetting speed is faster, and the historical data has little influence on iteration. $\Gamma_\eta(t) = 1 + \eta + \ldots + \eta^{t-1}$. By (17), (18), and (22), D_1 and D_2 can be updated as

$$\begin{cases} \overline{R}_L(t) = \left[1 - (t-1)R_L(t-1)\right] + \mu(t-1)\overline{R}_L(t-1), \\ R_H(t) = \left[1 - \mu(t-1)\overline{R}_H(t-1)\right] + \mu(t-1)\overline{R}_L(t-1), \\ D_1(t) = I - R_L(t), \\ D_2(t) = R_H^T(t) - R_H(t). \end{cases} \tag{24}$$

In (24), $R_L(t)$ update is accomplished by the combination of the last updated data $R_L(t-1)$ and the forgotten data $\overline{R}_L(t-1)$, and $R_H(t)$ update is accomplished by the combination of the last updated data $R_H(t-1)$ and the forgotten

data $\overline{R}_H (t - 1)$. From the previous analysis, it can be seen that the step size plays a very important role in improving the steady-state error and convergence speed of the algorithm [24]. If the selected step size is relatively large, a faster convergence rate will be obtained, but at the same time, the steady-state error will be large, and the separated estimated signal will not be accurate enough. If the selected step size is relatively small, the steady-state error obtained will be relatively small but will reduce the convergence speed. Therefore, the choice of step size will directly determine the performance of the algorithm. In the early stage of signal separation, the similarity between the estimated signal and the source signal is low, so a larger step size should be selected. In the later stage of signal separation, the similarity between the estimated signal and the source signal increases, and the step length should be shortened to obtain a smaller steady-state error. Therefore, in order to have both a faster convergence speed and a smaller steady-state error, the step size should be adaptively adjusted online according to the separation index. A nonlinear mapping is constructed, and the step length can be dynamically adjusted according to it. The nonlinear monotonic increasing function is shown as

$$\overline{\mu}(t) = \begin{cases} \beta \tanh \{\alpha [\zeta(t) - \varepsilon]\} + \delta, \zeta(t) \geq 0.5, \\ \beta \alpha \zeta(t)^3, 0 \leq \zeta(t) < 0.5. \end{cases} \tag{25}$$

α and β are constants. Adjust the shape by α and adjust the scale by β. ε represents the position of the maximum variance rate of the separation index ζ. The ideal value of ε is half of the maximum value of the separation index ζ. In order to satisfy $\overline{\mu}(t) = 0$, while $\zeta(t) = 0$, δ can be calculated as

$$\delta = -\beta \tanh(-\alpha \cdot \varepsilon). \tag{26}$$

Since the selection of the step size will be affected by the previous step size, the update method of the step size is shown as [27]

$$\mu(t) = \eta \mu(t - 1) + (1 - \eta)\overline{\mu}(t). \tag{27}$$

3.2. The ITAFBS. The separation performance of BSS to mixed signals is seriously disturbed by strong noise. In order to solve this problem, a novel improved threshold adaptive forgetting variable step size blind separation model for weak signal detection (ITAFBS) is proposed. The framework of ITAFBS is shown in Figure 5. The specific steps of weak signal detection are as follows:

(1) Set the number of wavelet decomposition layers h and wavelet basis function. The wavelet coefficients $W_{j,k}$ are obtained by the forward lifting wavelet transform of mixed signal x with strong noise.

(2) The general threshold estimation method is used to calculate the threshold λ. The objective function was used to calculate the optimal adjustment factor θ. The improved threshold function (3) is used to quantify the wavelet coefficients $W_{j,k}$ of each layer, and the new wavelet coefficients $\mathrm{w}_{j,k}$ are obtained.

(3) The inverse transform of the lifting wavelet is performed on the new wavelet coefficients $\mathrm{w}_{j,k}$ to obtain the mixed signal x' after denoising.

(4) According to the Gaussian characteristics of each component in the mixed signal x, the activation function $\psi(\cdot)$ is set. Set initial step size $\mu(0)$, empirical parameter ε, constant α and β, iteration number m, forgetting factor η, $t = 1$, and initial separation matrix $\boldsymbol{B}_1$.

(5) Calculation. The separation signal $y(1)$ was calculated by (9). The separation index ζ_t was calculated in the order of (23), (24), and (21).

(6) Update. According to the separation index ζ_t, (25) and (27) are used to update the iteration step $\mu(t)$. According to the new iteration step $\mu(t)$, (15) is used to update the separation matrix B_{t+1}.

(7) Judge whether the preset number of iterations m has been reached. If no, $t = t + 1$, perform Step 4) and Step 5). If so, obtain the latest separation matrix $\boldsymbol{B}'$.

(8) Equation (9) and the separation matrix $\boldsymbol{B}'$ are used to calculate the estimated signal $\boldsymbol{y}'$.

4. Simulation

4.1. Parameter Selection. This section is proposed to study the influence of parameters α and β on the ITAFBS. Four source signals and four mixtures of them are used. The expression of the source signals are shown in (28), which are sinusoidal signal, pulse signal, amplitude-modulation signal, and white-noise signal, respectively.

$$\left. \begin{array}{l} s_1(t) = \sin(1300 \times \pi \times t), \\ s_2(t) = \mathrm{sign}(\sin(290) \times \pi \times t), \\ s_3(t) = \sin(18 \times \pi \times t) \times \sin(600 \times \pi \times t), \\ s_4(t) = \mathrm{ran\,dn}[1, N]. \end{array} \right\} \tag{28}$$

The sampling frequency and sampling length are set as 12000 Hz and 0.5 s. Furthermore, the mixing matrix $\boldsymbol{A}$, randomly generated for each trial, is subject to normal distribution $N(0, 1)$. And Gaussian white noise (SNR = −20 dB) is added to the mixed signal. The performance index (PI) [28] is set as a function of the global matrix $\boldsymbol{C} = \boldsymbol{BA}$, which is mainly used to test the performances of the proposed method.

$$P = \frac{1}{n(n-1)} \sum_i \left\{ \left(\sum_j \frac{|C_{ij}|}{\max_k |C_{ik}|} - 1 \right) \right.$$
$$\left. + \sum_j \left(\sum_i \frac{|C_{ij}|}{\max_k |C_{kj}|} - 1 \right). \right\} \tag{29}$$

In (29), C_{ij} is the (i, j) element of the global matrix $\boldsymbol{C}$. The actual BSS results can only make the global matrix $\boldsymbol{C}$ as close as possible to a generalized sorted matrix. Generally, the smaller the PI value, the better the separation performance of the algorithm.
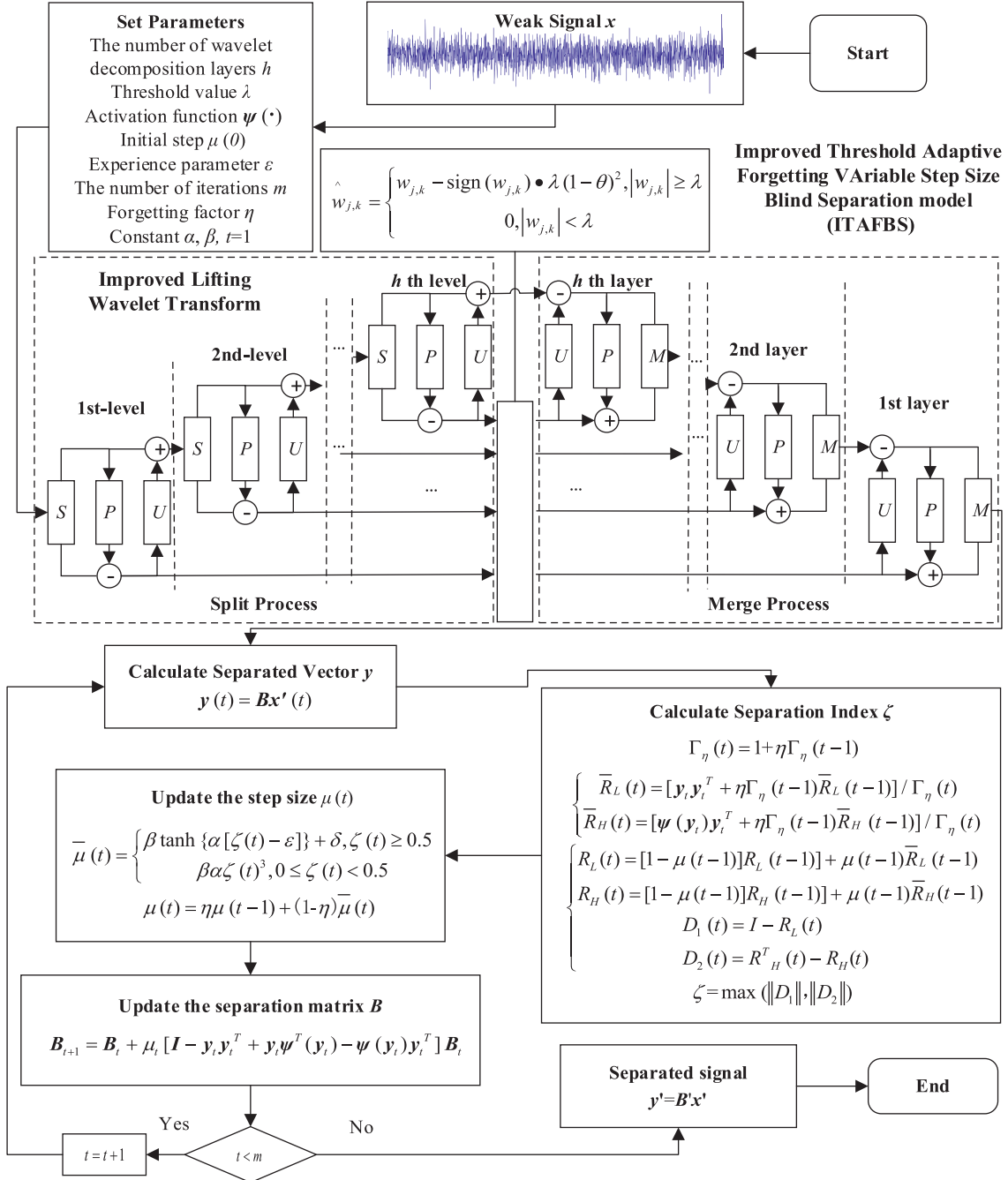
Figure 5: Framework of ITAFBS.

In the simulation, all the source signals are nonGaussian signals, so nonlinear function $\psi(y) = y^3$ is choice and $\varepsilon = 0.7$, $\mu(0) = 0.005$. PI curves of α and β with different values are shown in Figures 6 and 7.

In Figure 6, α has a great effect on the steady-state error, and with the decrease of α, the steady-state error decreases. In Figure 7, β has a great effect on the initial convergence rate, mainly because β controls the initial step size. The larger the parameter β is, the larger the initial step size is and the faster the initial convergence speed is. Considering the fast convergence speed and small steady-state error, set parameter $\alpha = 3$ and $\beta = 0.005$.

4.2. Contrast of Different Norm.
Aiming at the norm selection problem in (21), compare the separation performance of different norm corresponding algorithms according to the PI. Add the 1-norm, 2-norm, and F-norm of the matrix to the ITAFBS, and compare the PI, which is shown in Figure 8.

In Figure 8, norm has a great effect on the steady-state error and the initial convergence rate. Compared to 1-norm and 2-norm, the convergence speed of F-norm is close to 1.7 times that of 1-norm and 2-norm, indicating that its convergence speed is the fastest. In addition, the F-norm has the smallest final value of PI, indicating that its steady-state error is the smallest. It can be seen that the introduction of F-norm
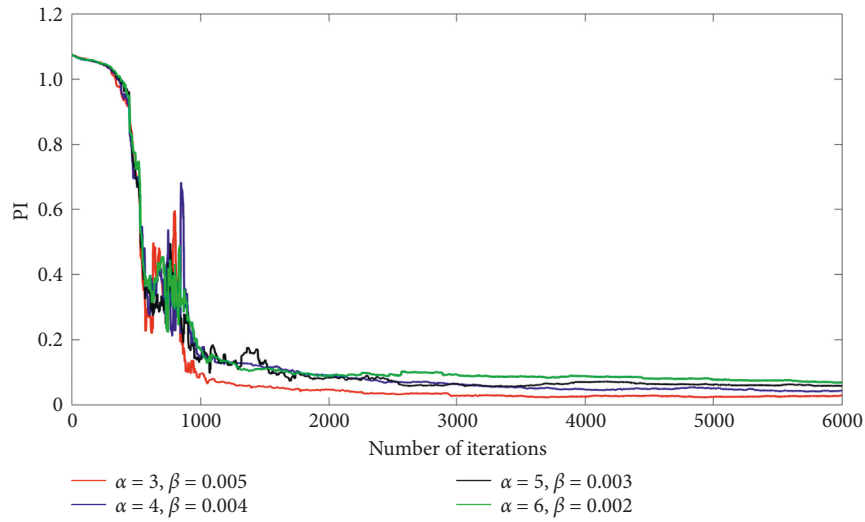
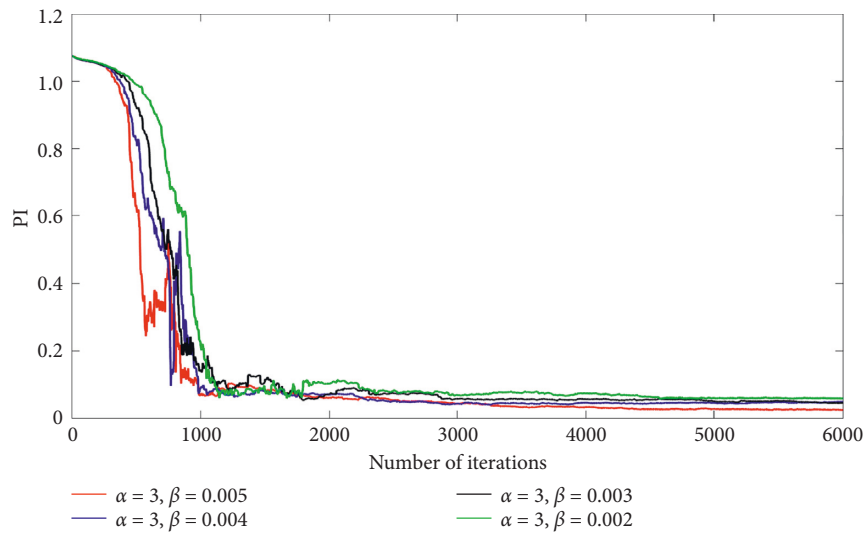Figure 6: Convergence performances with different parameters α.

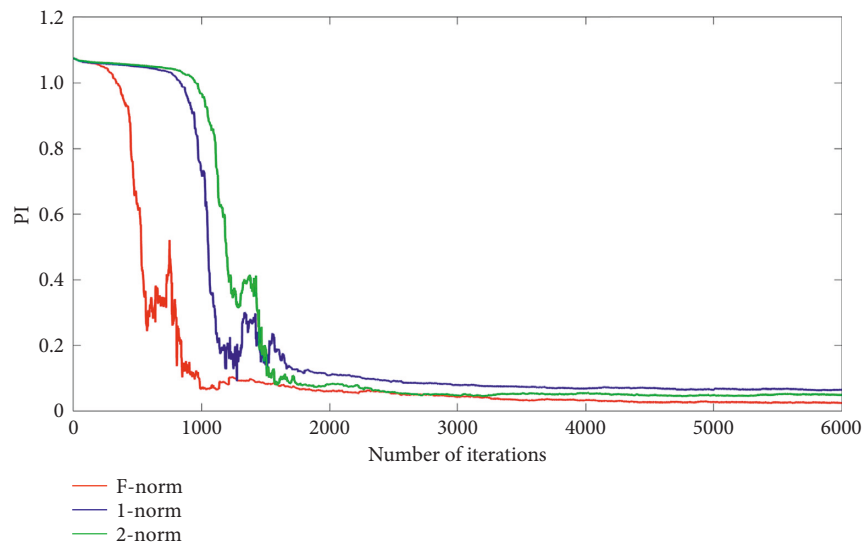
Figure 7: Convergence performances with different parameters β.

Figure 8: Convergence performances with different norms.
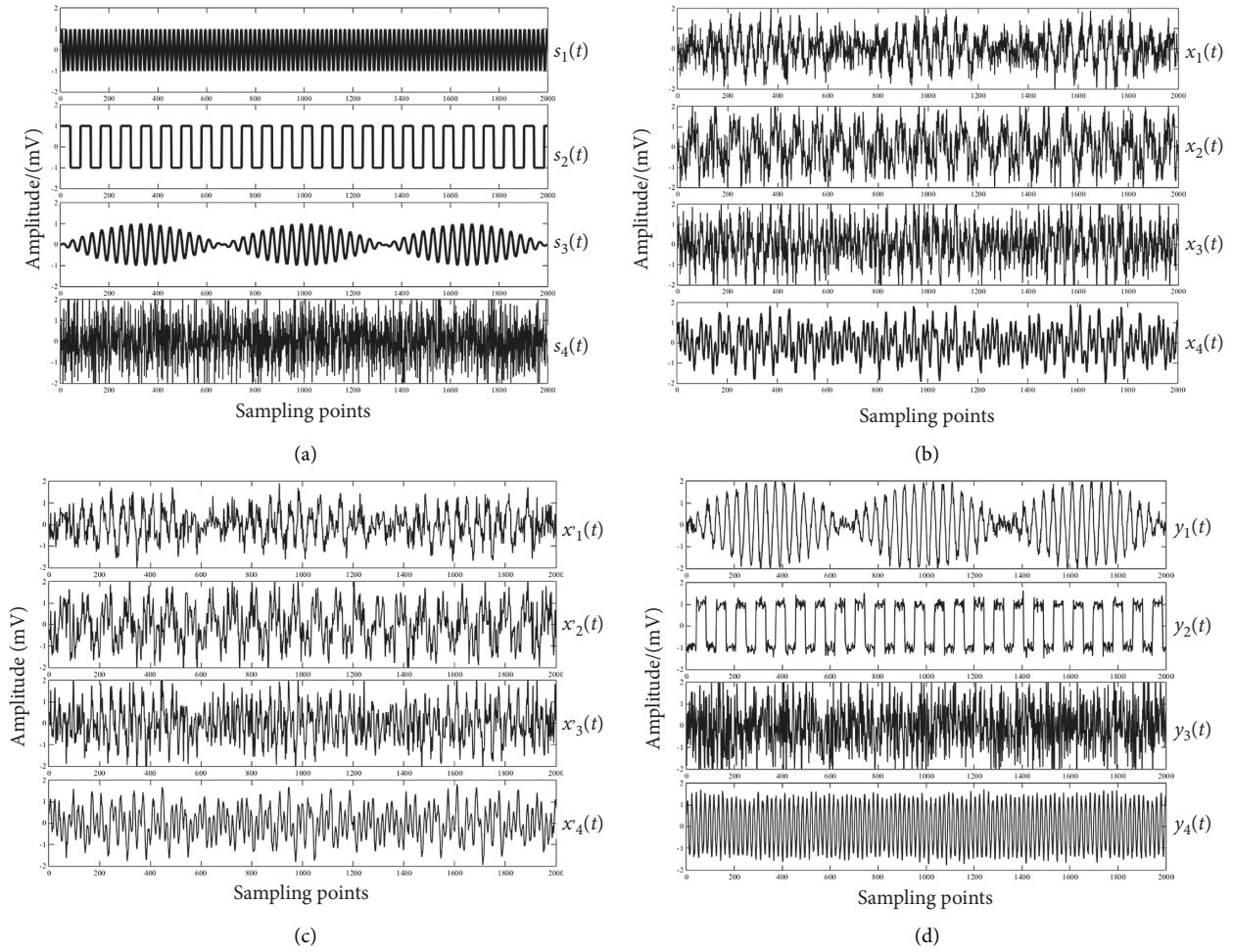
(a)

(b)

(c)

(d)

FIGURE 9: Simulation analysis of the ITAFBS. (a) Source signals, (b) mixed signals, (c) noise-reduced signals, and (d) separated signals.
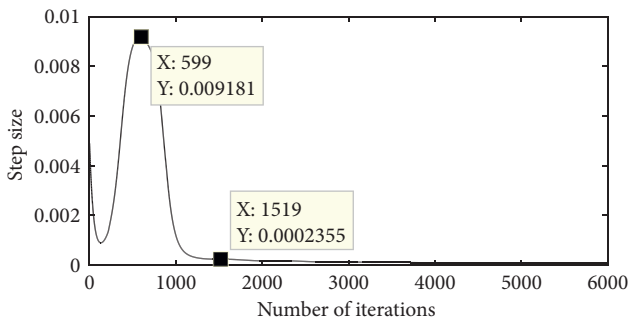

FIGURE 10: The step size of the ITAFBS.
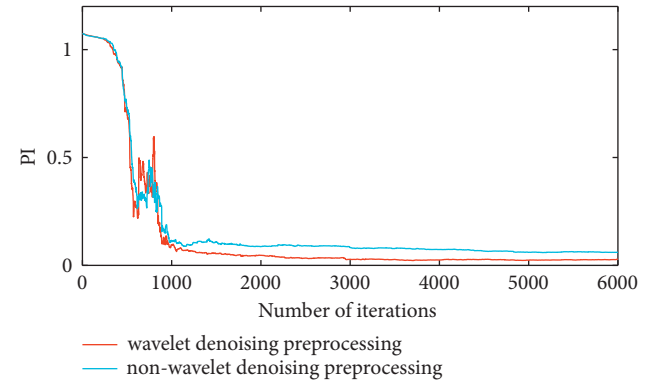

— wavelet denoising preprocessing
— non-wavelet denoising preprocessing

FIGURE 11: The comparison between wavelet denoising preprocessing and nonwavelet denoising preprocessing.

not only speeds up the convergence speed but also improves the stability of the learning algorithm.

4.3. Compare with Other Methods. The settings of the source signal are as described above. The proposed method is compared with the EASI method with fixed step size (FS-EASI) [24], the exponential-decay-step size method (EDS) [29], the adaptive step size method with weighted orthogonalization (AS-WO) [28] and variable step size algorithm with separation indicator (VS-SI) [15].

The initial parameters are set as follows: the step size of FS-EASI $\mu = 0.005$. The parameters of EDS are set as $\mu(0) = 0.005$, $K_0 = 1000$, $L_0 = 0.0012$, which are widely used in [19, 23]. The parameters of AS-WO are set as $\mu(0) = 0.005$, $\beta = 00.997$, $\rho = 0.003$. db$N$ wavelet has good regularity. In the ITAFBS, considering the smoothness, localization ability, computation amount, and real-time performance, db5 wavelet is selected as the wavelet basis function, and the
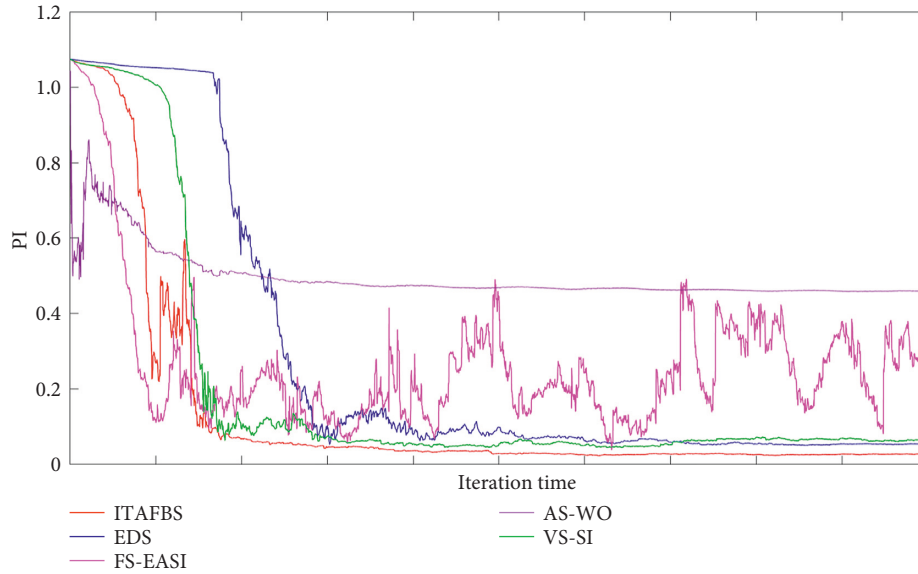
FIGURE 12: The performance comparison of the different BSS methods.

TABLE 2: Rolling bearing parameters.

Inner ring diameter (D_1/mm)	Outer ring diameter (d_1/mm)	Pressure angle (α/°)	Diameter of rolling element (d/mm)	Pitch diameter (D/mm)	Number of rolling elements (Z)
52	25	0	7.49004	39.0398	9

(a)

(b)
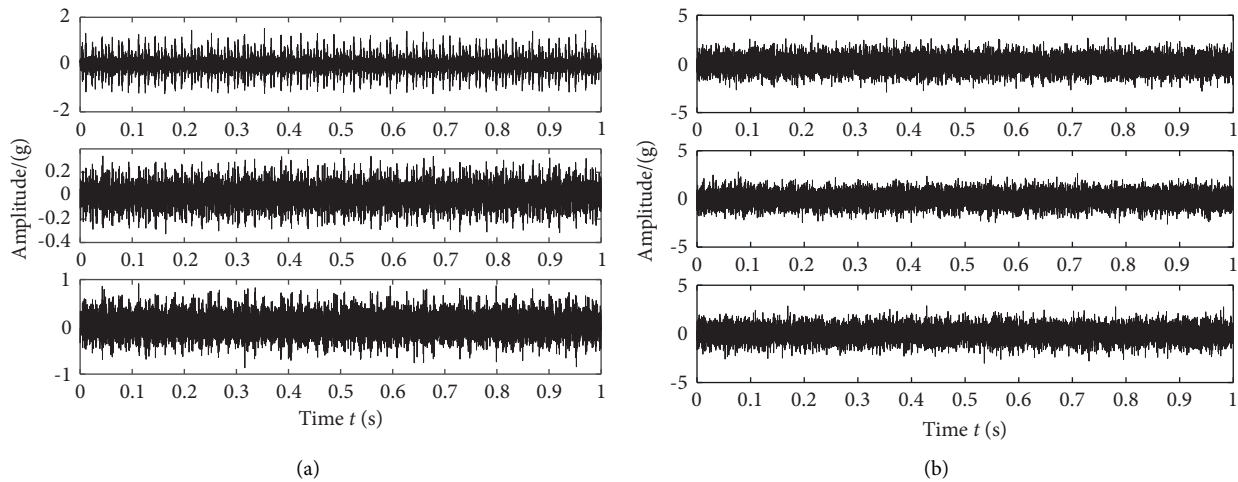
FIGURE 13: Time-domain waveform. (a) Original vibration signal and (b) vibration signal with noise.

number of decomposition layers $h = 3$, $\alpha = 3$, $\beta = 0.005$, $\varepsilon = 0.7$, $\eta = 0.98$.

Source signals and mixed signals of simulation analysis are shown in Figures 9(a) and 9(b). The noise-reduced signals and separated signals recovered by the ITAFBS are presented in Figures 9(c) and 9(d).

Compare Figures 9(b)–9(d), the ILWT can eliminate part of the noise, but cannot make the signal characteristics appear. Therefore, it is necessary to use a separation algorithm to further process the signal. Compare Figures 9(a) and 9(d), the recovery effect of the source signal is ideal except that the amplitude and sequence of the signal are

inconsistent, indicating the effectiveness of the ITAFBS. Figure 10 shows the trend diagram of the step size of the ITAFBS with the number of iterations.

It can be seen from Figure 10, that in the early stage of the separation process since the correlation between the signals is relatively strong, the degree of signal separation is relatively low. At this time, a larger step value should be selected to accelerate the algorithm convergence, the step size increases rapidly to near the maximum value of 0.0091. With the continuous separation process, the step size is reduced to about 0.00023. When the number of iterations reaches about 1500, the signal has been basically separated.
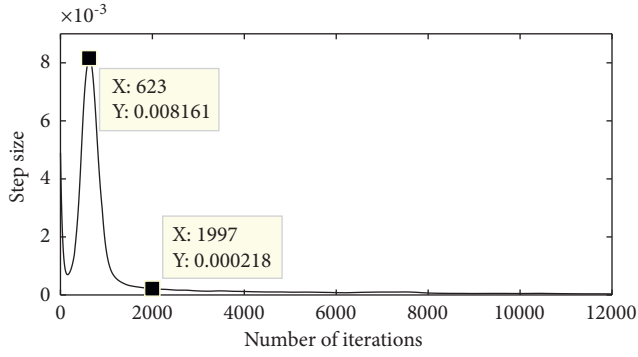
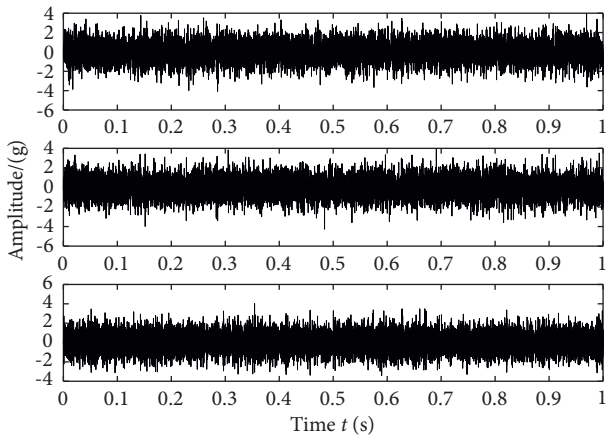FIGURE 14: The change of the step size of the ITAFBS.

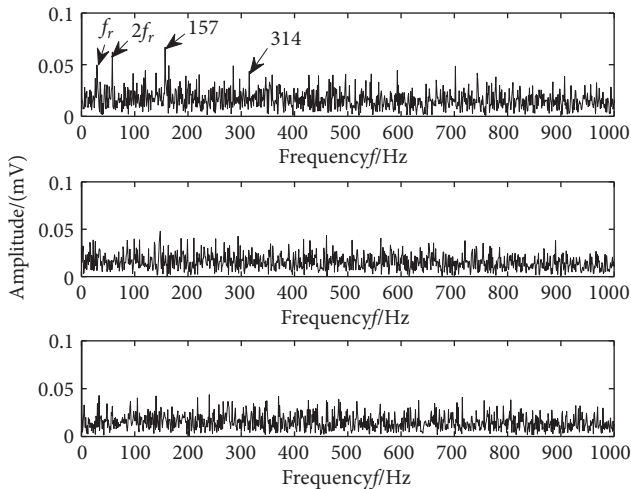FIGURE 15: Time-domain waveform after the ITAFBS.

FIGURE 16: Envelope spectrum of separated signal.

At this time, in order to obtain a smaller steady-state error ratio, the step size is smaller. Finally, in order to improve the accuracy and further improve the degree of signal separation, the step length keeps a slow speed and continues to decrease until its value fluctuates near a very small value. The changing trend of step size shows the correctness of the ITAFBS.

In order to illustrate the effect of wavelet denoising preprocessing, the comparison between wavelet denoising preprocessing and nonwavelet denoising preprocessing is performed, and the result is shown in Figure 11.

It can be seen from Figure 11, the two algorithms nearly converge after about 1000 iterations, which indicates that the convergence rates of the two algorithms are nearly the same. However, it can be seen that the algorithm with wavelet denoising preprocessing can obtain a smaller final value of the PI, which indicates that the algorithm has a smaller steady-state error. Therefore, the LWT can further improve the separation performance of the algorithm.

Figure 12 shows the performance comparison of the different BSS methods. It shows the PI obtained from EDS, FS-EASI, AS-WO, VS-SI, and the ITAFBS.

From Figure 12, the convergence speed of AS-WO is the fastest, and the steady-state error is the largest. The convergence speed of ITAFBS, VS-SI, and FS-EASI is faster than EDS. The steady-state error of FS-EASI is relatively large and has a large fluctuation due to the limitation of fixed step size. The ITAFBS has the smallest steady-state error compared with other methods. In conclusion, the ITAFBS has the best separation performance.

5. Experimentation

5.1. The Weak Signal of Acceleration Sensor. The experimental data used in this section are the open experimental data of Case Western Reserve University (CWRU). The faulty bearing is at the shaft of the motor. The rolling bearing model is SKF6205-2RS JEM deep groove ball bearing. The bearing fault is the pitting fault of the inner ring. The test speed is 1750 r/min, and the rotation frequency of the vibration signal is $f_r = 29.16$ Hz. The parameters of the rolling bearing at the drive end are shown in Table 2. In the process of collecting vibration signal data, a vibration acceleration sensor was placed on the drive end (DE), fan end (FE), and base of the motor (BA), and a 16-channel data logger was selected to save the data. The fault frequency of the outer ring can be calculated by (30).

$$f_i = \frac{Z}{2}\left(1 + \frac{d}{D}\cos\alpha\right)f_r. \tag{30}$$

The frequency of failure of the inner ring of the bearing is 157.94 Hz. The sampling frequency and sampling length are set as 12000 Hz and 1 s. The time-domain waveforms collected by sensors at three positions are shown in Figure 13(a). Observing the time-domain waveform diagram, it can be seen that the time-domain waveform in the drive end is ideal, and the impact characteristics are more prominent. Gaussian white noise (SNR = −10 dB) is added to the vibration signal of the rolling bearing. The time-domain waveform of the noise-added vibration signal is shown in Figure 13(b). It can be seen that the fault characteristics at this time have been completely overwhelmed by strong noise.
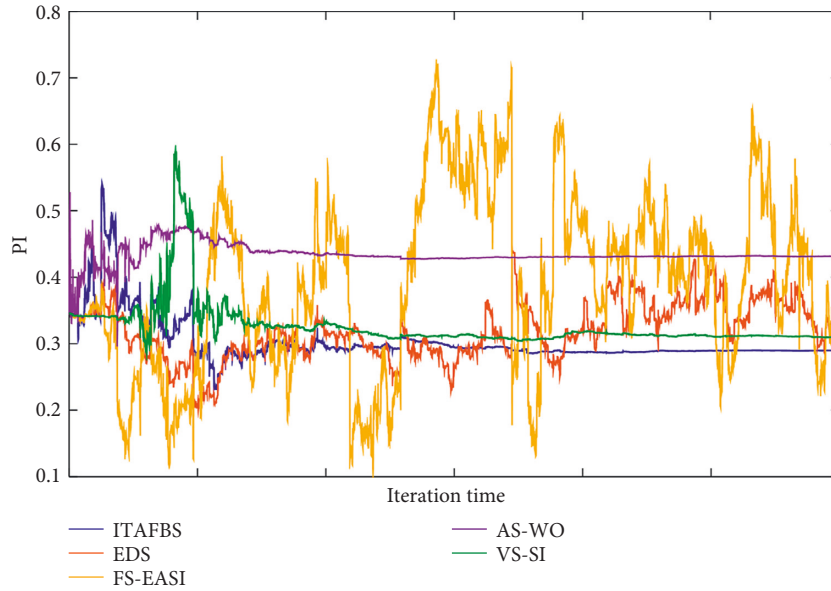
Figure 17: The performance comparison of the different BSS methods (CWRU data).
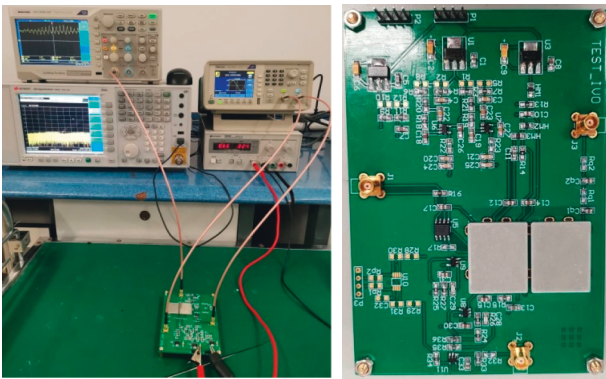

Figure 18: The test system and analog amplifier circuit board.

The parameter settings are consistent with Section 4.3. The change in the step size of the ITAFBS is shown in Figure 14.

It can be seen from Figure 14 that at the initial stage of separation (the number of iterations is 0 to 623), the signal separation degree is low, and the separation model needs to accelerate convergence by rapidly increasing the step size. When the number of iterations $t = 623$, the iteration step size reaches the maximum value ($\mu = 0.0082$), and the model reaches the maximum convergence speed. When the number of iterations is 623 to 2000, the separation index is close to the steady-state value, and the step size gradually decreases in order to obtain a small steady-state error. The number of iterations after 2000, iteration step size ($\mu = 0.0002$) is close to convergence, and the step size changes slowly, which indicates that the separation of mixed signals is basically completed. The change mode of step size conforms to the theory of Section 3.2.

The time-domain waveform of the signal after noise reduction processing and separated by the ITAFBS is shown in Figure 15.

It can be seen from Figure 15 that the noise is suppressed to a certain extent, but the shock characteristics of the vibration signal are still unable to be obtained from the time-domain waveform diagram. The Hilbert envelope spectrum is used to demodulate the three separated signals. The envelope spectrum is shown in Figure 16.

In Figure 16, the fault characteristic frequency 157 Hz and the second harmonic 314 Hz of the inner ring can be clearly found from the envelope spectrum of the first separated signal. It indicates that the bearing in drive end has the pitting fault of the inner ring.

Figure 17 shows the performance comparison of the different BSS methods. It shows the PI obtained from EDS, FS-EASI, AS-WO, VS-SI, and the ITAFBS.

From Figure 17, the convergence speed of ITAFBS is the fastest, and the steady-state error is the smallest. The ITAFBS has the best separation performance. To sum up, the ITAFBS can accurately detect the fault characteristics of bearing inner ring.

5.2. The Weak Signal of Test Circuit. This experiment aims to test the application effect of the ITAFBS in the field of instrument. The equipment of the test system includes signal generator, weak signal analog amplifier circuit board, signal acquisition instrument, shielding box, and computer. Target signal with a frequency of 79 Hz and a peak-to-peak value of 50 mVpp is generated by signal generator. At the same time, a clock signal with a frequency of 30 kHz and a duty cycle of 10% is generated. The target signal is attenuated by a large capacitor, and the amplification results are output by weak signal analog amplifier circuit board. Signal acquisition instrument is used to collect the output signal of the amplifying chip. Shielding box is used to shield outside interference. The test system and analog amplifier circuit board are shown in Figure 18.
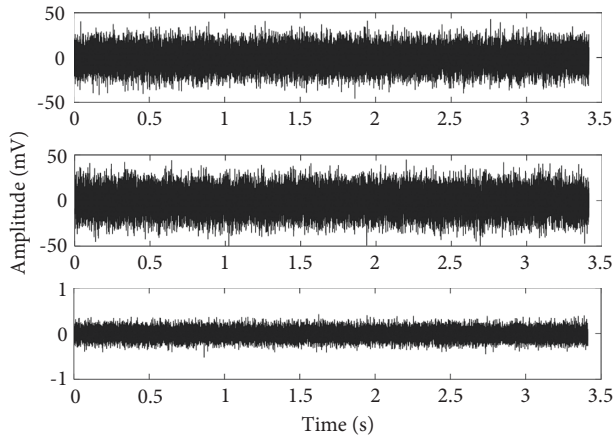
Figure 19: The time-domain waveform of observation signal (test system).
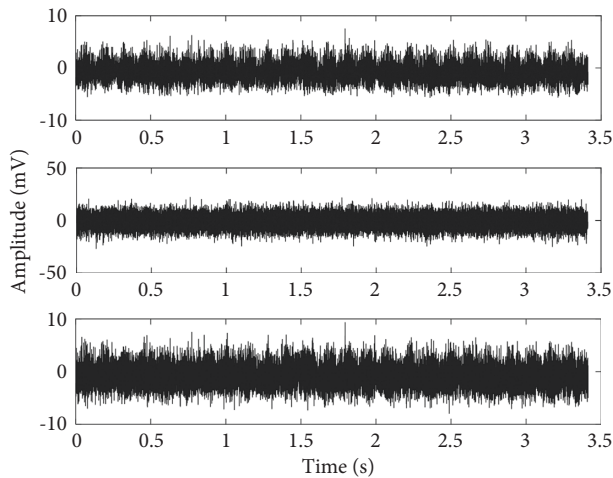

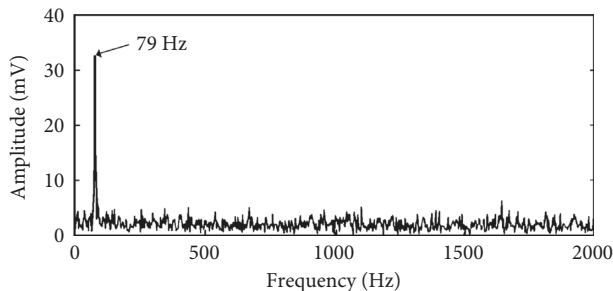Figure 20: The time-domain waveform of separated signal (test system).

Figure 21: The frequency domain of separated signal (test system).

The observation signal can be obtained by mixing the chip output signal and the two-channel output signal of the signal generator. The time-domain waveform is shown in Figure 19. The ITAFBS method is used to separate the observed signals, and the time-domain waveform of the separated signal is obtained as shown in Figure 20. The periodicity of the signal can be seen in the first figure in Figure 20. The signal is analyzed in the frequency domain

and the result is shown in Figure 21. The peak frequency of 79 Hz (the target frequency) can be clearly seen in Figure 21. Therefore, the ITAFBS can accurately detect the weak target signal generated by the test circuit.

6. Conclusion

A novel improved threshold adaptive forgetting variable step size blind separation model (ITAFBS) for weak signal detection is proposed. The ITAFBS has the ability of fast and accurate feature extraction for weak signals, which benefits from the ILWT to reduce the noise of weak signals, and the fast and accurate decomposition of AFBS to preprocessed mixed signals.

After simulation and experimental analysis, the following conclusions can be drawn:

(1) The proposed improved threshold function achieves better noise reduction results than other threshold functions

(2) The PI evaluation index analysis of five different algorithms shows that the ITAFBS has the fastest convergence speed, the smallest steady-state error, and the best separation performance

(3) The verification results of two experimental cases show that ITAFBS can be applied to mechanical and instrument fields, which is feasible and practical

In future work, we will further study the order indeterminacy and amplitude indeterminacy in the ITAFBS.

Data Availability

Data in the manuscript can be found in the link "https://engineering.case.edu/bearingdatacenter."

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

T.Y., H.W., and S.G. conceptualized the study; T.Y. and S.G. proposed the methodology; T.Y. provided software; T.Y., S.L., and J.L. validated the study; T.Y. and Y.W. performed formal analysis; T.Y. and G.T. investigated the study; T.Y. and S.L. were responsible for resources; T.Y. and H.W. performed data curation; T.Y. wrote the original draft; T.Y. reviewed and edited the manuscript; T.Y. and J.L. visualized the study; S.L. and J.L. supervised the study; Y.W.; was responsible for project administration; , G.T. was responsible for funding acquisition. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

References

[1] J. Z. Gao, *Detection of Weak Signals*, Tsinghua University Press, Beijing, China, 2004.

[2] Y. G. Lei, *Intelligent Fault Diagnosis and Remaining Useful Life Prediction of Rotating Machinery*, Xi'an Jiaotong University Press, Xian, China, 2017.

[3] W. Sweldens, "The lifting scheme: a construction of second generation Wavelets : a construction of second generation wavelets," *SIAM Journal on Mathematical Analysis*, vol. 29, no. 2, pp. 511–546, 1998.

[4] K. Loukhaoukha, J.-Y. Chouinard, and M. H. Taieb, "Optimal image watermarking algorithm based on LWT-SVD via multi-objective ant colony optimization," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, no. 4, pp. 303–319, 2011.

[5] Z. Prusa and P. Rajmic, "Real-time lifting wavelet transform algorithm," *Journal of Signal Processing*, vol. 2, no. 3, pp. 53–59, 2011.

[6] H. Zhang, J. Fang, and H. Liu, "Online current signal de-noising of magnetic bearing switching power amplifier based on lifting wavelet transform," *IET Electric Power Applications*, vol. 10, no. 8, pp. 799–806, 2016.

[7] X. Y. Cao, Z. J. Zhang, and J. J. Xiang, "Method of radar signal de-noising based on lifting wavelet improved threshold," *Computer Engineering and Applications*, vol. 48, no. 14, pp. 143–147, 2012.

[8] W. Ju, C. H. Lu, and Y. J. Zhang, "Open-Path fourier transform infrared spectrum de-noising based on improved threshold lifting wavelet transform and adaptive filter," *Spectroscopy and Spectral Analysis*, vol. 38, no. 06, pp. 1684–1690, 2018.

[9] Z. Li, Z. J. He, Y. Zi, and H. Jiang, "RotatingMachineryFault diagnosis using signal-adapted lifting scheme," *Mechanical Systems and Signal Processing*, vol. 22, no. 3, pp. 542–556, 2008.

[10] J. M. Wang, B. Li, and Q. Wang, "Measurement of gas/liquid two-phase flow velocity based on lifting wavelet time delay estimation," *Chinese Journal of Scientific Instrument*, vol. 38, no. 3, pp. 653–663, 2017.

[11] W. Li and X. Wang, "An improved wavelet threshold de-noising data processing method research in deformation monitoring," *Applied Mechanics and Materials*, vol. 90-93, pp. 2858–2863, 2011.

[12] P. He, M. Qi, W. Li, M. Tang, and Z. Zhao, "A general non-stationary and time-varying mixed signal blind source separation method based on online Gaussian process," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, no. 11, p. 2058015, 2020.

[13] J. He, W. Chen, and Y. Song, "Single channel blind source separation under deep recurrent neural network," *Wireless Personal Communications*, vol. 115, no. 2, pp. 1277–1289, 2020.

[14] K. Gajowniczek, I. Grzegorczyk, M. Gostkowski, and T. Zabkowski, "Blind source separation for the aggregation of machine learning algorithms: an arrhythmia classification case," *Electronics*, vol. 9, no. 3, p. 425, 2020.

[15] J. Lu, W. Cheng, and Y. Zi, "Online blind source separation method with adaptive step size in both time-invariant and time-varying cases," *Measurement Science and Technology*, vol. 31, no. 4, Article ID 045102, 2020.

[16] B. Iyer, S. L. Nalbalwar, and N. P. Pathak, "[Advances in intelligent systems and computing] computing, communication and signal processing volume 810," *Proceedings of ICCASP 2018) ‖ Comparative Analysis of ICA, PCA-Based EASI and Wavelet-Based Unsupervised Denoising for EEG Signals*, Springer, Singapore, pp. 749–759, 2019.

[17] M. Rattray, D. Saad, and S. I. Amari, "Natural gradient descent for on-line learning," *Physical Review Letters*, vol. 81, no. 24, pp. 5461–5464, 1998.

[18] C. Wang, H. Huang, Y. Zhang, and Y. Chen, "Variable learning rate EASI-based adaptive blind source separation in situation of nonstationary source and linear time-varying systems," *Journal of Vibroengineering*, vol. 21, no. 3, pp. 627–638, 2019.

[19] A. Mohsin Abdulazeez, D. Qader Zeebaree, D. M Hajy, and D. Asaad Zebari, "Robust watermarking scheme based LWT and SVD using artificial bee colony optimization," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 2, pp. 1218–1229, 2021.

[20] P. Singh and R. Chadha, "A survey of digital watermarking techniques, applications and attacks," *International Journal of Engineering and Innovative Technology*, vol. 2, no. 9, pp. 165–175, 2013.

[21] K. Loukhaoukha and J. Y. Chouinard, "Optimal image watermarking algorithm based on LWT-SVD via multi-objective ant colony optimization," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 4, 2011.

[22] A. Cichocki, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, John Wiley & Sons, NJ. USA, 2002.

[23] J. Ye, H. Jin, S. Lou, and K. You, "An optimized EASI algorithm," *Signal Processing*, vol. 89, no. 3, pp. 333–338, 2009.

[24] J. A. Chambers, M. G. Jafari, and S. Mclaughlin, "Variable step-size EASI algorithm for sequential blind source separation," *Electronics Letters*, vol. 40, no. 6, pp. 393-394, 2004.

[25] X. Zhang, X. Zhu, and Z. Bao, "Grading learning for blind source separation," *Science in China, Series A F*, vol. 46, no. 1, pp. 31–44, 2003.

[26] C. L. Fancourt and L. Parra, "The coherence function in blind source separation of convolutive mixtures of non-stationary signals," in *Proceedings of the Neural Networks for Signal Processing Xi, IEEE Signal Processing Society Workshop IEEE*, MA, USA, September 2001.

[27] T. P. von Hoff and A. G. Lindgren, "Adaptive step-size control in blind source separation," *Neurocomputing*, vol. 49, no. 1-4, pp. 119–138, 2002.

[28] X. Tang, X. Zhang, and J. Ye, "Adaptive step-size natural gradient ICA algorithm with weighted orthogonalization," *Circuits, Systems, and Signal Processing*, vol. 33, no. 1, pp. 211–221, 2014.

[29] H. H. Yang, "Serial updating rule for blind separation derived from the method of scoring," *IEEE Transactions on Signal Processing*, vol. 47, no. 8, pp. 2279–2285, 1999.