

# An improved complex ICA based method for wind farm harmonic emission levels evaluation

Jinshuai Zhao<sup>a</sup>, Honggeng Yang<sup>a,\*</sup>, Aiqiang Pan<sup>b</sup>, Fangwei Xu<sup>a</sup>

<sup>a</sup> The College of Electrical Engineering, Sichuan University, Chengdu, 610065, China

<sup>b</sup> State Grid Shanghai Electric Power Research Institute, Shanghai, 200437, China



## ARTICLE INFO

### Keywords:

Wind farm  
Harmonic emission level  
Harmonic impedance  
Improved complex independent component analysis  
Sparse component analysis  
Hybrid negentropy

## ABSTRACT

Wind farms, as typical renewable energy infrastructure, provide clean energy to the power grid but also inevitably generate harmonic pollution. In the case of a multi-wind farms system, evaluating the harmonic emissions for each farm is essential for the division of harmonic responsibility and the design of the harmonic mitigation scheme. However, this evaluation process faces two challenges. First, the background harmonics of a single wind farm might be unstable. In addition, the harmonic impedance of wind farm side is reduced by the installation of filters. As a result, traditional evaluation methods are no longer applicable. Complex independent component analysis is a well-known method for evaluating harmonic emissions. To make this method feasible in the above situations, we analyze its shortcomings. Furthermore, an improved method is proposed by introducing two screening mechanisms based on sparse component analysis and the maximization of hybrid negentropy. Additionally, the results of simulations and field case verification indicate that the traditional methods have large calculation errors when the background harmonics fluctuate greatly or the harmonic impedance of the wind farm side is not far larger than that of the utility side. In contrast, our novel evaluation method has high precision in the above cases.

## 1. Introduction

Grid-connected wind farms inject clean energy into power grids but also generate harmonic problems that degrade power quality [1,2]. The assessment of wind-farm harmonic emission properties is essential for studying the influence of these farms on their connected network. Yet, related studies are few [3,4]. In China, wind farms are usually connected to a 35-kV/110-kV network through long transmission lines, and the corresponding short-circuit capacity is low [5]. Owing to the weak connection between wind farms and the grid, voltage distortion is exacerbated [5,6]. In the case of a Chinese multi-wind farms system, for example, the maximum value for the total harmonic distortion voltage of the power pooling point is as high as 2.4%, which exceeds the Chinese standard limits [7]. Therefore, to mitigate harmonic pollution, it is important to evaluate the harmonic emission level for each wind farm at the grid connection point [8,9].

Evaluation methods can be classified as invasive and non-invasive methods. Invasive method is based on the variations of voltage and current which are caused by the injected disturbances [10]. Although the calculation accuracy of an invasive method is acceptable, the assessment results are not real time but only valid for the moment of

testing [11,12]. Additionally, the injected disturbances may be harmful for the operation of the power grid [13–15]. In contrast, noninvasive methods, which are based on natural harmonic variations, are free from the above problems, and are thus more widely adopted in the evaluation of harmonic emission levels.

Noninvasive methods can be further categorized as qualitative and quantitative methods. In contrast to quantitative methods, the qualitative methods only determine which side of the point of common coupling (PCC) dominates the harmonic and do not provide specific contribution levels [16–18]. Quantitative methods have thus received more attention.

A typical quantitative method has two steps: computing the utility and customer harmonic impedances from the harmonic data measured at the PCC; and then evaluating the harmonic emission levels of the both sides. Generally, the existing evaluating methods mainly rely on two assumptions: (1) the background harmonic is stable; and (2) the impedance of the customer side is large enough to ignore its influence on the evaluation, and then, only the impedance of the utility side ( $Z_u$ ) should be calculated in the evaluation [11,12,19–23].

Several studies have focused on the harmonic emissions evaluation. Among the methods adopted, the fluctuation method [22,23], which is

\* Corresponding author.

E-mail address: [PQlab99@126.com](mailto:PQlab99@126.com) (H. Yang).

a classical evaluation technology based on the aforementioned assumptions, can calculate  $Z_u$  approximately according to the ratios of the harmonic voltage fluctuations to current fluctuations at the PCC. However, this method is poor in terms of resisting the effects of background harmonics. According to the basic assumptions of the fluctuation method, many other technologies based on the statistical properties of harmonic data have been derived to reduce the effects of background harmonics on the evaluation process. For instance, on the basis of the linear clustering properties for harmonic voltages and currents at the PCC, a group of regression methods [19–21,24,25] can calculate  $Z_u$  using the slope of the clustering line. Although regression is one technology widely adopted in this field of research, its clustering effect excessively depends on the stability of the background harmonics. Additionally, in situations that harmonic filters are installed in the customer side, the harmonic impedance of the customer side may not be far larger than  $Z_u$  under a certain harmonic order. As a consequence, the slopes of the corresponding lines are affected, which decreases the accuracy of the calculated  $Z_u$ . To reduce the interference of background harmonic fluctuations, the covariance characteristic method, which is another statistics-based technology, has been proposed in Ref. [12]. This method reduces the effects of the background harmonics using the weak correlation between the background harmonic voltage and harmonic current measured at the PCC. However, under the conditions of the multi-wind farms system where harmonic filters are installed, it is difficult to establish the basic assumption of the covariance characteristic method that the harmonic impedance of the customer side is much larger than  $Z_u$ .

In summary, most existing methods are only suitable for situations where the both aforementioned assumptions hold. However, these two assumptions are valid for typical nonlinear customers but invalid for a single wind farm in a multi-wind farms system (Fig. 1). First, the utility side contains other wind farm harmonic sources, which potentially destabilize the background harmonics [8,9]. In addition, the harmonic impedance of the wind farm side ( $Z_w$ ) can be reduced by the installation of harmonic filters; thus,  $|Z_w| > |Z_u|$  might not hold [11,12]. Notably, the reduction of  $Z_w$  could also increase the difficulty of calculating  $Z_u$  [14]. Additionally, in this situation, not only  $Z_u$  but also  $Z_w$  should be calculated for the evaluation of harmonic emissions, which is a great challenge for existing methods.

Complex independent component analysis (CICA) is widely adopted in the field of blind source separation and promising for the evaluation of harmonic emission level [13,14,26]. According to the independence of the harmonic sources on the two sides of the PCC, CICA can calculate harmonic impedances more robustly, even when the background harmonic fluctuates. Yet, the calculation accuracy of CICA depends on a massive number of sample points, which can be ensured by extending the sampling time. However, the prolonged sampling time can cause impedances variation and consequently generate new errors [15].

The present paper analyzes the errors of CICA in terms of the signal separation capability and the optimization capability. These errors are

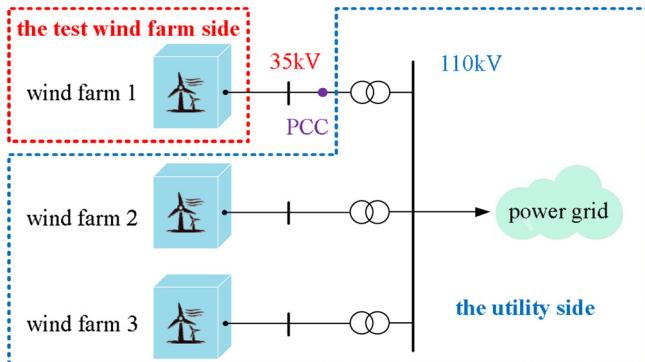


Fig. 1. Typical topological structure of the multi-wind farms system.

then greatly reduced by introducing a screening mechanism based on sparse component analysis (SCA) [27–29] and the maximization of hybrid negentropy. With the proposed screening mechanism, local signals, which are highly consistent with the true harmonic sources, can be found by searching nearby the signals separated from CICA. Finally, the harmonic impedances of both sides can be calculated accurately through these selected local signals, and a harmonic emission level is obtained. This improved CICA (ICICA) method is valid for situations that the background harmonic fluctuates greatly and the harmonic impedance amplitudes of both sides are close to each other. Results of simulation and a field study validate the proposed method.

## 2. Model of wind farm harmonic emissions

Fig. 1 shows the topology of a multi-wind farms system, where the grid connection point of a single wind farm is set as the PCC. The harmonic emission model can be represented by a Norton circuit [8,9]. Although both sides of the PCC have nonlinear components, the circuit can be linearized under the  $h$ th harmonic by treating the nonlinear parts as harmonic sources (Fig. 2) [11–22]. The subscript  $h$  is omitted hereafter for simplification. According to the principle of superposition, we have

$$\begin{cases} \dot{U}_{\text{pcc}} = \frac{Z_u Z_w}{Z_u + Z_w} (\dot{i}_u + \dot{i}_w) \\ \dot{i}_{\text{pcc}} = \frac{Z_w}{Z_u + Z_w} \dot{i}_w - \frac{Z_u}{Z_u + Z_w} \dot{i}_u \end{cases} \quad (1)$$

where  $\dot{U}_{\text{pcc}}$  and  $\dot{i}_{\text{pcc}}$  respectively represent the harmonic voltages and currents measured at the PCC, while  $\dot{i}_w$ ,  $\dot{i}_u$  and  $Z_w$ ,  $Z_u$  respectively represent the equivalent harmonic currents and impedances of the wind farm side and utility side.

Filters are installed in the wind farms, and  $|Z_w| > |Z_u|$  thus no longer holds for some harmonic orders. Therefore, the effects of both  $Z_w$  and  $Z_u$  should be considered in the evaluation of harmonic emissions. When  $\dot{i}_w$  and  $\dot{i}_u$  operate independently, the corresponding harmonic voltages at the PCC are [30]

$$\begin{cases} \dot{U}_{\text{pcc-w}} = \frac{Z_u Z_w}{Z_u + Z_w} \left( \frac{\dot{U}_{\text{pcc}}}{Z_w} + \dot{i}_{\text{pcc}} \right) \\ \dot{U}_{\text{pcc-u}} = \dot{U}_{\text{pcc}} - \dot{U}_{\text{pcc-w}} \end{cases} \quad (2)$$

where  $\dot{U}_{\text{pcc-w}}$  and  $\dot{U}_{\text{pcc-u}}$  are respectively the harmonic voltage emission levels of the wind farm and utility sides. It is clear that the calculation of  $Z_w$  and  $Z_u$  is the key to evaluating the harmonic emission levels for both sides of the PCC.

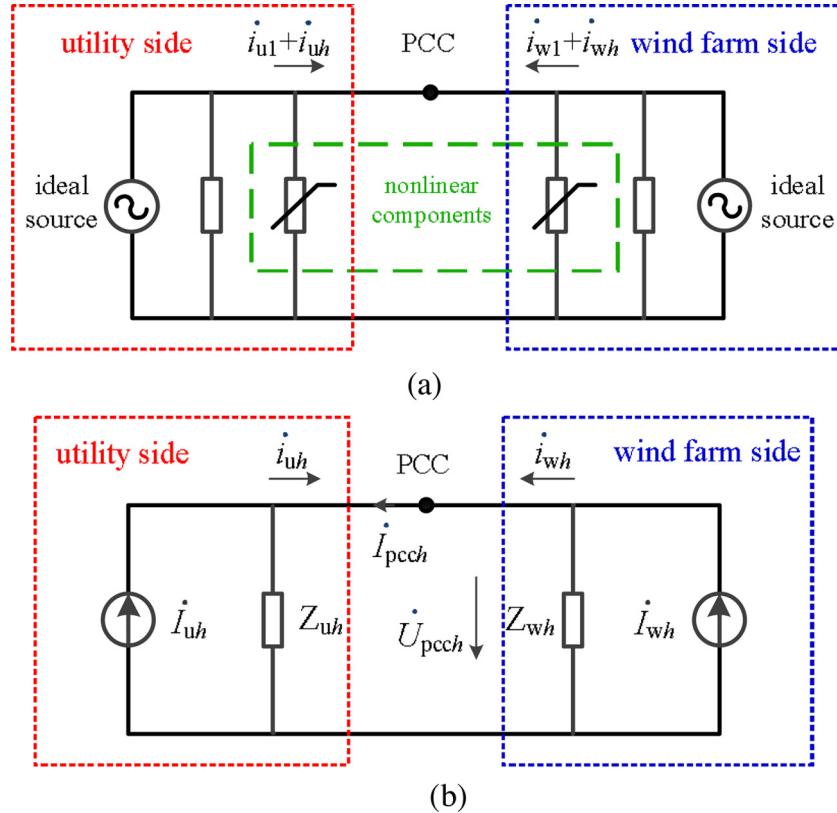
Existing calculation methods are based on the condition that the harmonic impedances are approximately constant. In practice,  $Z_u$  changes slowly and is usually stable within a short time segment [11,12,14]. However, fluctuations of  $Z_w$  are relatively frequent, and it is thus crucial to select the time interval during which the variations of  $Z_w$  are weak. It has been reported that, less variation of the fundamental current may lead to less variation of  $Z_w$  [14]. The fundamental current is thus an important factor in selecting appropriate time intervals in the present paper.

## 3. Improved complex ICA method

The fast-varying components of  $\dot{i}_w$  and  $\dot{i}_u$ , which can be extracted using an averaging filter, are approximately independent [13,14]. CICA can thus be adopted to separate the harmonic sources and calculate the harmonic impedances on both sides of the PCC.

### 3.1. Analysis of the traditional CICA algorithm

The central limit theorem indicates that the non-Gaussianities of



**Fig. 2.** Norton equivalent circuit. (a) Nonlinear circuit containing fundamental and  $h$ th harmonic frequencies. (b) Linearized circuit under the  $h$ th harmonic.

independent signals are weakened after linear combination. CICA thus reconstructs the source signals by maximizing the non-Gaussianities of the separated signals. The non-Gaussianity of a certain signal  $s$  can be measured by negentropy [31] as

$$J(s) \approx [E\{G(s)\} - E\{G(s_{\text{Gauss}})\}]^2 \quad (3)$$

where  $s$  and  $s_{\text{Gauss}}$  are signals with zero-mean and unit variance,  $s_{\text{Gauss}}$  obeys a Gaussian distribution,  $E\{\cdot\}$  denotes an expectation value, and  $G\{\cdot\}$  is a nonlinear function defined by Eq. (4). The non-Gaussianity of  $s$  strengthens as  $J(s)$  increases.

$$G\{s\} = -\exp(-s^2/2) \quad (4)$$

The blind source separation form of Eq. (1) is

$$\mathbf{X} = \mathbf{AS} \quad (5)$$

where  $\mathbf{X} = [\dot{U}_{pcc} \dot{I}_{pcc}]^T$  and  $\mathbf{S} = [\dot{I}_u \dot{I}_w]^T$  are respectively the fast-varying components of  $[\dot{U}_{pcc} \dot{I}_{pcc}]^T$  and  $[\dot{I}_u \dot{I}_w]^T$ . Matrix  $\mathbf{A}$  is unknown and mixed by  $Z_u$  and  $Z_w$ .

To solve Eq. (5),  $\mathbf{X}$  is first preprocessed by centering and whitening [13].  $\dot{I}_u^{\text{fast}}$  and  $\dot{I}_w^{\text{fast}}$  are then recovered through CICA. The calculation process has been described in the literature [13,14]. On the basis of the signals separated from CICA (denoted as  $\mathbf{Y}$ ), matrix  $\mathbf{A}$  can be calculated as

$$\hat{\mathbf{A}} = \mathbf{XY}^T (\mathbf{YY}^T)^{-1} \quad (6)$$

To solve the ordering indeterminacies of the results of CICA [13,14] (whether  $\mathbf{Y} = [\dot{Y}_u \dot{Y}_w]^T$  or  $\mathbf{Y} = [\dot{Y}_w \dot{Y}_u]^T$  is unknown), we define

$$\begin{cases} k_1 = \hat{\mathbf{A}}(1, 1)/\hat{\mathbf{A}}(2, 1) \\ k_2 = \hat{\mathbf{A}}(1, 2)/\hat{\mathbf{A}}(2, 2) \end{cases} \quad (7)$$

Due to the resistive part of impedance being positive, we have

$$k_i = \begin{cases} \hat{Z}_u, & \text{if } k_{i-x} > 0 \\ -\hat{Z}_w, & \text{if } k_{i-x} < 0 \end{cases} \quad (i = 1, 2) \quad (8)$$

where  $k_{i-x}$  represents the resistive part of  $k_i$ . The harmonic impedances of both sides are thus calculated through CICA.

In the process of evaluating harmonic emissions, under some harmonic orders, if  $|Z_w| > |Z_u|$  holds, we only need to calculate  $Z_u$ . Otherwise, if  $|Z_w|$  and  $|Z_u|$  are close to each other, it is necessary to calculate both  $Z_w$  and  $Z_u$ . With the criterion proposed in Ref. [14], we can judge whether  $|Z_w| > |Z_u|$  holds according to the degree of coincidence between  $\dot{Y}_w$  and  $\dot{I}_{pcc}^{\text{fast}}$ . If  $\dot{Y}_w$  is highly consistent with  $\dot{I}_{pcc}^{\text{fast}}$ ,  $|Z_w| > |Z_u|$  holds. Otherwise,  $|Z_w|$  and  $|Z_u|$  are close to each other.

The accuracy of CICA is strongly depends on the sample size. When the sample size is large, results are more accurate because the statistical characteristics of the data are obvious. However, the harmonic impedances are more likely to change in this situation, resulting in new errors [15]. Reducing the sample size ensures constant harmonic impedances, but the error of CICA still increases owing to the weak statistical properties of the data. To improve the CICA method, we analyze its errors as follows.

### 3.1.1. Signal separation capability

The calculated results have high accuracy when the signals separated from CICA coincide closely with the corresponding true source signals. Conversely, the error is large in the situation that the signal separation results are very different from the true source signal. Hence, strengthening the consistency between the separated signals and the true source signals is critical to improving CICA.

### 3.1.2. Optimization capability

On the basis of the central limit theorem, CICA recovers source signals by maximizing the non-Gaussianities of separated signals. Thus, when the non-Gaussianities of  $\mathbf{Y}$  are weaker than those of  $\mathbf{S}$ , (i.e., the optimization is inadequate), the accuracies of calculated impedances

are low. Premature convergence and local optimum may reduce the optimization capability of CICA and thereby lead to inadequate optimization [32].

To enhance the signal separation and optimization capability of CICA, we improve this method by introducing SCA and a hybrid negentropy screening mechanism into it.

### 3.2. SCA

SCA, a novel blind source separation technology, reconstructs source signals according to their sparsity. Sparse signals are nonzero only at a few sampling points and tend to zero at most times [27–29]. Thus, if source signals are sparse, the observed signals at each sampling point are at most generated by only one source signal. For instance,  $[s_1 \ s_2]^T$  and  $[x_1 \ x_2]^T$  are set as sources and observed signals, respectively:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \quad (9)$$

At sampling time  $t_1$ , assuming  $s_2(t_1) \approx 0$  and  $s_1(t_1) \neq 0$ , we have

$$\begin{cases} x_1(t_1) = a_{11}s_1(t_1) \\ x_2(t_1) = a_{21}s_1(t_1) \end{cases} \quad (10)$$

Thus,

$$x_1(t_1)/x_2(t_1) = a_{11}/a_{21} \quad (11)$$

Similarly, if  $s_1(t_2) \approx 0$  and  $s_2(t_2) \neq 0$  at sampling time  $t_2$ , we have

$$x_1(t_2)/x_2(t_2) = a_{12}/a_{22} \quad (12)$$

Applying the above analyses (i.e. Eqs. (9)–(12)) to Eq. (5), if  $S$  is sparse,  $X$  will cluster into two lines passing through the origin with slopes of  $A(1, 1)/A(2, 1)$  and  $A(1, 2)/A(2, 2)$ , respectively. According to Eqs. (7) and (8),  $Z_u$  and  $Z_w$  can be calculated from the slopes of these two lines.

Before adopting SCA, we should make source signals sparse using a sparse dictionary. Because most signals are sparse in the wavelet domain or time-frequency domain, the sparse dictionary  $D$  can be generated using a wavelet packet [27,28] or short-time Fourier transformation (STFT) [29]. The effect of the dictionary  $D$  is shown in Fig. 3. Like the source signals in Eq. (5), the signal  $s$  in Fig. 3 is set as a row vector, and  $s_{\text{sparse}}$  is its corresponding sparse form. There are two categories of sparse dictionaries, namely complete and over complete dictionaries. This paper adopts a complete dictionary because its inverse is valid. Thus, using an  $n \times n$  complete dictionary  $D$ , the  $1 \times n$  signal  $s$  can be transformed into its sparse form  $(D^{-1})^T$ .

Based on the sparse dictionary  $D$ , Eq. (5) can be mathematically transformed by post-multiplying by  $(D^{-1})^T$ , [28,29] as

$$X(D^{-1})^T = AS(D^{-1})^T \quad (13)$$

where  $S(D^{-1})^T$  denotes the sparse source signals and  $X(D^{-1})^T$  denotes the transformed observed signals, which are linearly clustered.

The sparse transformation does not change matrix  $A$ , and harmonic impedances of both sides can thus be calculated using the slopes of clustering lines. However, because  $i_u^{\text{fast}}$  and  $i_w^{\text{fast}}$  fluctuate randomly, they contain rich frequency information, which makes it hard for the STFT or wavelet packet transformation to convert the two source signals into sparse forms [33]. Thus, SCA cannot be used directly to calculate harmonic impedances for the aforementioned harmonic model.

Of note, our research finds that through their advantages, CICA and SCA exactly compensate for each other's weaknesses as shown in Fig. 4.

The rationale of combining the advantages of the two methods to calculate the harmonic impedances is as follow. In our proposed method, finding signals that coincide with the true source is one of the keys to calculating the harmonic impedances. This can be done using an SCA-based screening mechanism that will be exhaustively explained later. This procedure can be carried out more efficiently by narrowing the searching scope through focusing on signals that are similar to the true source signals. Notably, signals separated from CICA are similar to the true source signals, even though they do not completely coincide with them. Thus, the searching process can be optimized by searching around these separated signals. Meanwhile, because signals coinciding with the true sources can be precisely identified by the SCA based screening mechanism, the accuracy of signal separation for CICA is also improved. The specific process of combining SCA and CICA is as follows.

To reduce the search dimension, local signals are used instead of global signals. Symbols  $\dot{y}_{u,i}$  and  $\dot{y}_{w,i}$  are the  $i^{\text{th}}$  local signals obtained from screening locally around  $\dot{Y}_u$  and  $\dot{Y}_w$ .  $[i_{u,i}^{\text{fast}}, i_{w,i}^{\text{fast}}]^T$  and  $x_i$  respectively denote the local source signals and local observed signals corresponding to the sampling points of  $\dot{y}_{u,i}$  and  $\dot{y}_{w,i}$ . In the proposed ICICA method, the processes of calculating  $Z_u$  and  $Z_w$  are the same, and only the process of calculating  $Z_u$  is presented in the following analyses. As previously mentioned, the classic sparse technology in SCA (e.g. the STFT) is not suitable for many signals. Thus, in this paper, we calculate the sparse dictionary of  $\dot{y}_{u,i}$  using the sparse dictionary learning algorithm based on the K-means Singular Value Decomposition (KSVD) [34]. Unlike cases using STFT,  $D$  obtained from the KSVD can only make one of the sources sparse. When  $D$  corresponds to  $\dot{y}_{u,i}$ , it cannot transform  $\dot{y}_{w,i}$  into its sparse form. We therefore denote the sparse dictionary calculated from  $\dot{y}_{u,i}$  as  $D_{u,i}$ . Similarly, the dictionary calculated from  $\dot{y}_{w,i}$  is denoted as  $D_{w,i}$ . To measure the degree of similarity between  $\dot{y}_{u,i}$  and  $i_{u,i}^{\text{fast}}$ , we transform  $x_i$  into  $x_i(D_{u,i}^{-1})^T$ . Two possible situations exist as below:

- 1)  $\dot{y}_{u,i}$  is consistent with  $i_{u,i}^{\text{fast}}$ : In this case,  $D_{u,i}$  generated from  $\dot{y}_{u,i}$  can also make  $i_{u,i}^{\text{fast}}$  sparse. Thus,  $x_i(D_{u,i}^{-1})^T$  cluster into a line with the slope equaling  $Z_u$ .
- 2)  $\dot{y}_{u,i}$  is not consistent with  $i_{u,i}^{\text{fast}}$ : In this case,  $D_{u,i}$  cannot make  $i_{u,i}^{\text{fast}}$  sparse and  $x_i(D_{u,i}^{-1})^T$  usually does not undergo linear clustering unless [29].

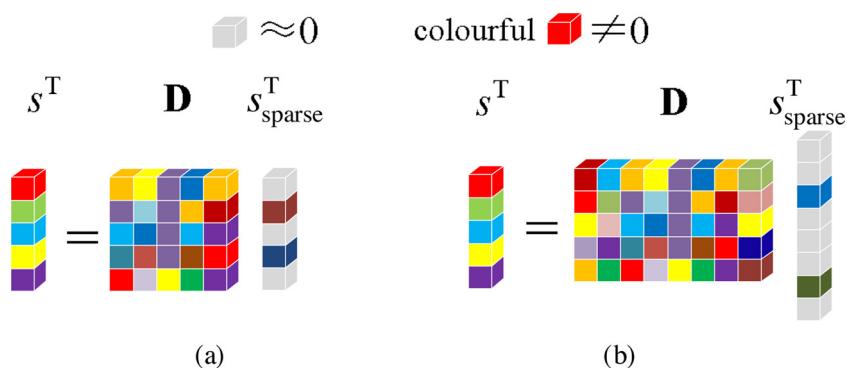


Fig. 3. Sparse dictionary matrix  $D$ . (a) complete. (b) over complete.

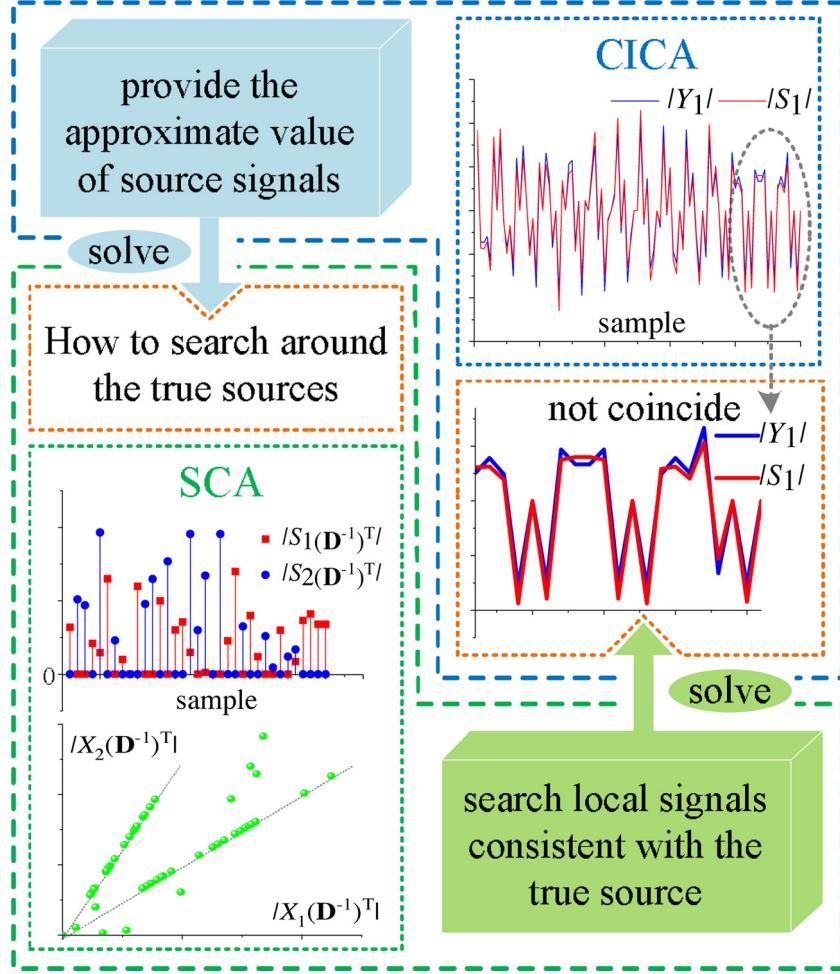


Fig. 4. The idea of combining SCA and CICA.

$$\frac{[\hat{i}_{u,i}^{\text{fast}}(\mathbf{D}_{u,i}^{-1})^T]_1}{[\hat{i}_{w,i}^{\text{fast}}(\mathbf{D}_{u,i}^{-1})^T]_1} = \frac{[\hat{i}_{u,i}^{\text{fast}}(\mathbf{D}_{u,i}^{-1})^T]_2}{[\hat{i}_{w,i}^{\text{fast}}(\mathbf{D}_{u,i}^{-1})^T]_2} = \dots = \frac{[\hat{i}_{u,i}^{\text{fast}}(\mathbf{D}_{u,i}^{-1})^T]_{n-T_0}}{[\hat{i}_{w,i}^{\text{fast}}(\mathbf{D}_{u,i}^{-1})^T]_{n-T_0}} \quad (14)$$

holds, where  $n$  is the length of local source signals  $\hat{i}_{u,i}^{\text{fast}}$  and  $\hat{i}_{w,i}^{\text{fast}}$ .  $T_0$  is the sparsity degree [33] of  $\hat{y}_{u,i}(\mathbf{D}_{u,i}^{-1})^T$ , i.e. the number of nonzero elements. The subscripts, 1, 2, ...,  $n - T_0$  represent the corresponding sampling points when  $\hat{y}_{u,i}(\mathbf{D}_{u,i}^{-1})^T$  is zero.

Assume the value obtained from Eq. (14) is equal to  $k_f$ . For signals meeting Eq. (14), it is deduced from Eq. (13) that

$$\frac{x_1(\mathbf{D}^{-1})^T}{x_2(\mathbf{D}^{-1})^T} = \frac{A(1, 1)k_f + A(1, 2)}{A(2, 1)k_f + A(2, 2)} \quad (15)$$

Thus, the transformed observed signals  $\mathbf{X}(\mathbf{D}^{-1})^T$  also cluster into a line passing through the origin but with a slope equaling  $\frac{A(1, 1)k_f + A(1, 2)}{A(2, 1)k_f + A(2, 2)}$  instead of  $Z_u$ . However, the previous studies showed that the probability that Eq. (14) holds is low [29].

Thus, using the linear clustering degree of  $x_i(\mathbf{D}_{u,i}^{-1})^T$ , we can judge whether  $\hat{y}_{u,i}$  is consistent with  $\hat{i}_{u,i}^{\text{fast}}$  through

$$\sigma_{\text{sparse}} = \max \left( \left| \frac{[x_i(\mathbf{D}_{u,i}^{-1})^T]_{1,j} - [x_i(\mathbf{D}_{u,i}^{-1})^T]_{1,k}}{[x_i(\mathbf{D}_{u,i}^{-1})^T]_{2,j} - [x_i(\mathbf{D}_{u,i}^{-1})^T]_{2,k}} \right| \right) < \varepsilon, \forall j \neq k \quad (16)$$

where  $\sigma_{\text{sparse}}$  and  $\varepsilon$  are respectively the sparse screening index and threshold. Subscripts 1, 2 represent the row numbers of  $x_i(\mathbf{D}_{u,i}^{-1})^T$  while subscripts  $j, k$  represent the corresponding sampling points when  $\hat{y}_{u,i}(\mathbf{D}_{u,i}^{-1})^T$  is zero. If Eq. (16) holds,  $\hat{y}_{u,i}$  is consistent with  $\hat{i}_{u,i}^{\text{fast}}$ , and  $\hat{Z}_{u,i}$  can thus be calculated from the slope of  $x_i(\mathbf{D}_{u,i}^{-1})^T$ .

To minimize the search dimension,  $n$  should be as small as possible.

However, Eq. (16) holds for all signals if  $n < 3$ . We thus set  $n = 3$  and the dimensions of  $\mathbf{D}_{u,i}$  are  $3 \times 3$ . To further reduce the search dimension, we adjust  $\dot{Y}_u$  and  $\dot{i}_u^{\text{fast}}$ :

$$\begin{cases} \dot{Y}_u = \dot{Y}_u / \dot{Y}_u(t_m) \\ \dot{i}_u^{\text{fast}} = \dot{i}_u^{\text{fast}} / \dot{i}_u^{\text{fast}}(t_m) \end{cases} \quad (17)$$

where  $t_m$  is the sampling point corresponding to the maximum amplitude of  $\dot{Y}_u$ .

Therefore, at point  $t_m$ , we have  $\dot{Y}_u(t_m) = \dot{i}_u^{\text{fast}}(t_m) = 1$ . The search dimension is thus further reduced to 2 by setting  $\dot{Y}_u(t_m)$  as a part of  $\dot{y}_{u,i}$ . The above process of dimension reduction is shown in Fig. 5. The length of the local signal is as low as 3. Furthermore, for each local signal and its corresponding local source, there is always a sample point with a signal value equaling 1. Thus, at this sample point,  $\dot{y}_{u,i}$  and  $\dot{i}_{u,i}^{\text{fast}}$  are naturally equal to each other, and then, only two sample points wait to be found. The search dimension is so low that enough local signals can always be found through sparse screening.

Furthermore, we have  $|\dot{Y}_u| \in (0, 1]$  after the transformation of Eq. (17). The amplitude of  $\dot{i}_u^{\text{fast}}$  is large at  $t_m$  because the general trend of  $\dot{Y}_u$  is usually similar to that of  $\dot{i}_u^{\text{fast}}$ . Therefore, using Eq. (17), we approximately have  $|\dot{i}_u^{\text{fast}}| \in (0, 1]$ . Eq. (17) thus narrows the search area by reducing the difference between  $\dot{Y}_u$  and  $\dot{i}_u^{\text{fast}}$ . The searching process can be accomplished by exhaustively searching around  $\dot{Y}_u$  or using an intelligent optimization algorithm (e.g. particle swarm optimization) with the objective function of minimizing  $\sigma_{\text{sparse}}$ . Through the same process, we can also search for  $\dot{y}_{w,i}$  and calculate  $\dot{Z}_{w,i}$ .

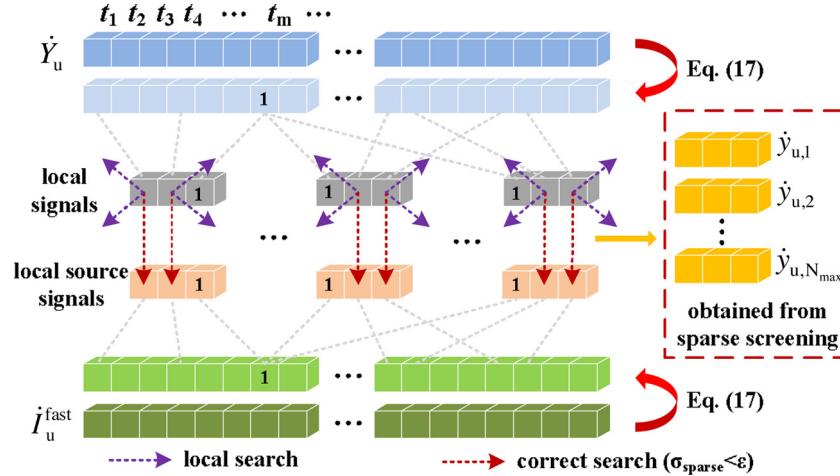


Fig. 5. Schematic diagram of the local search and sparse screening.

### 3.3. Hybrid negentropy screening mechanism

Although Eq. (14) usually does not hold, its occurrence can reduce the reliability of the sparse screening mechanism. Among the local signals searched out by Eq. (16), some are highly consistent with the true source signals but a few still have large errors. In this paper, we screen these local signals once again according to their non-Gaussianity.

In the mode of Eq. (1), when the complex signal separated from CICA has strong non-Gaussianity, its real and imaginary parts also have strong non-Gaussianity [15]. Since inadequate optimization can cause errors for CICA, when CICA has low precision, the non-Gaussianity for both real and imaginary parts of the separated signal are weak. By comparison, for true source signals, both the real and imaginary parts have strong non-Gaussianity. Thus, the non-Gaussianity for the real and imaginary parts of the calculated signals can be used as an index for secondary screening. The specific process is as follows.

The corresponding global signals of  $\hat{y}_{w,i}$  and  $\hat{y}_{u,i}$ , respectively denoted as  $\hat{y}_{w,i}^{\text{global}}$  and  $\hat{y}_{u,i}^{\text{global}}$ , are calculated as

$$\begin{bmatrix} \hat{y}_{w,i}^{\text{global}} \\ \hat{y}_{u,i}^{\text{global}} \end{bmatrix} = \begin{bmatrix} 1/\hat{Z}_{w,i} & 1 \\ 1/\hat{Z}_{u,i} & -1 \end{bmatrix} \begin{bmatrix} \hat{U}_{\text{pcc}}^{\text{fast}} \\ \hat{I}_{\text{pcc}}^{\text{fast}} \end{bmatrix} \quad (18)$$

where the lengths of  $\hat{y}_{w,i}^{\text{global}}$  and  $\hat{y}_{u,i}^{\text{global}}$  are the same as the length of  $I_u^{\text{fast}}$ .

The symbols  $y_{u-x,i}^{\text{global}}$  and  $y_{u-y,i}^{\text{global}}$  respectively denote the real and imaginary parts of  $\hat{y}_{u,i}^{\text{global}}$ . If  $\hat{y}_{u,i}^{\text{global}}$  is consistent with  $I_u^{\text{fast}}$ , both  $y_{u-x,i}^{\text{global}}$  and  $y_{u-y,i}^{\text{global}}$  have strong non-Gaussianity [15]. Otherwise, their respective non-Gaussianities are weak. Non-Gaussianity can be measured through negentropy as expressed by Eq. (3).

Through Eq. (19),  $y_{u-x,i}^{\text{global}}$  and  $y_{u-y,i}^{\text{global}}$  are transformed into signals with zero-mean and unit variance, which makes their negentropies comparable. The pending signal  $s$  follows

$$s \rightarrow s = s - \bar{s} \rightarrow s = s/\text{std}(s) \quad (19)$$

where  $\bar{s}$  and  $\text{std}(s)$  are respectively the mean value and standard deviation of  $s$ .

The negentropies of  $y_{u-x,i}^{\text{global}}$  and  $y_{u-y,i}^{\text{global}}$  are respectively denoted as  $J_{x,i}$  and  $J_{y,i}$ . For all signals obtained using Eq. (16), the corresponding  $J_{x,i}$  can be normalized to  $J_{x,i}$  through Eq. (20) and the corresponding  $J_{y,i}$  can be normalized to  $J_{y,i}$  in the same way.

$$J_{x,i} = (J_{x,i} - J_{x,\min})/(J_{x,\max} - J_{x,\min}) \quad (20)$$

where  $J_{x,\min}$  and  $J_{x,\max}$  are respectively the minimum and maximum values of all  $J_{x,i}$ .

The proposed hybrid negentropy of  $\hat{y}_{u,i}^{\text{global}}$  is defined as

$$\sigma_{HNE,i} = (J_{x,i} + J_{y,i})/2 \quad (21)$$

Therefore, among all  $\hat{y}_{u,i}$  obtained using Eq. (16), the one with the largest  $\sigma_{HNE,i}$  is the most optimal for calculating  $Z_u$ . Similarly,  $Z_w$  can be determined in the same way.

The probabilities that Eq. (14) holds are relatively low [29]. Therefore, theoretically, if sufficient local signals are obtained from sparse screening (where 50 local signals are sufficient on the basis of our simulation and practical experience), there are satisfactory results among them. Thereby, accurate results can always be found through hybrid negentropy screening.

As the complement relationship between SCA and CICA, the two proposed screening mechanisms also complement one another (Fig. 6). Hybrid negentropy screening can make up for the weakness of false screening resulting from Eq. (14). Meanwhile, because the index  $\sigma_{HNE,i}$  is based on the statistical properties provided by global signals, the search dimension will be terribly huge if hybrid negentropy screening is used directly. However, the search dimension is greatly reduced with the help of sparse screening and local searching.

On the basis of the two screening mechanisms, the ICICA method

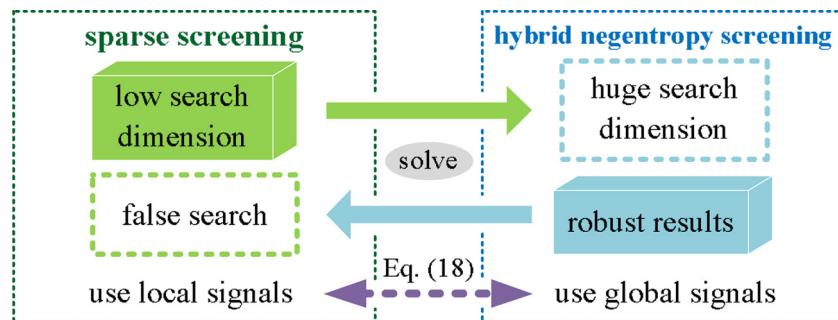


Fig. 6. Relationship of the two screening mechanisms.

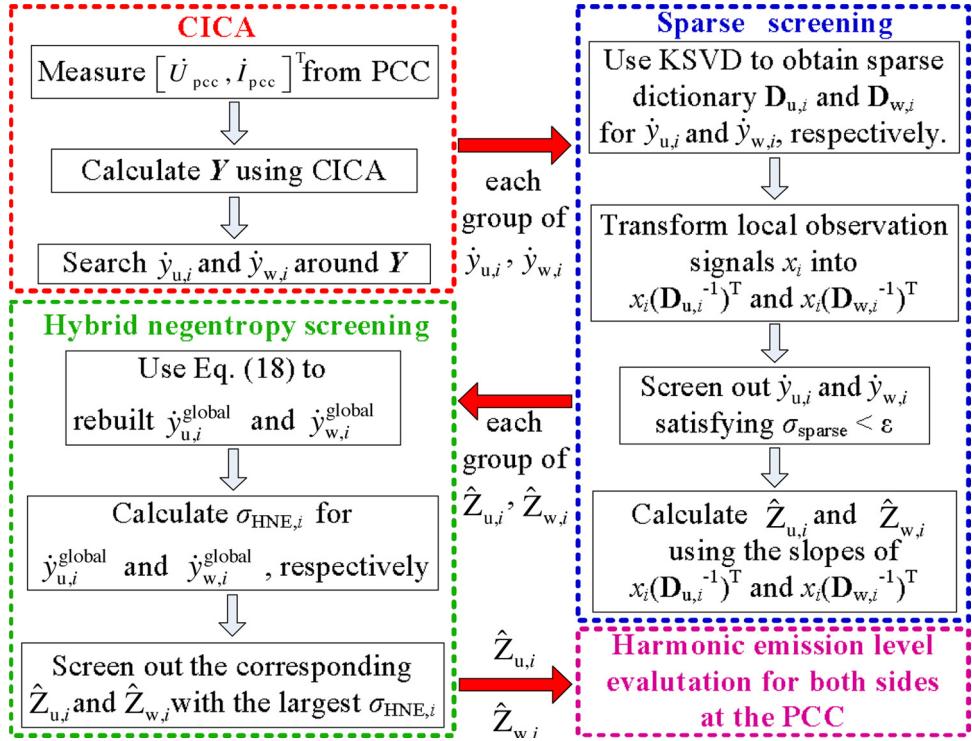


Fig. 7. Schematic diagram of the ICICA method.

evaluates harmonic emissions accurately. Fig. 7 is a schematic diagram of the ICICA.

The specific steps of ICICA in Fig. 7 are as follows.

- 1) Use CICA to separate  $\dot{Y}_w$  and  $\dot{Y}_u$  and adjust them using Eq. (17). Set  $i = 1$ .
- 2) Search around  $\dot{Y}_w$  and  $\dot{Y}_u$  locally to find  $\dot{y}_{w,i}$  and  $\dot{y}_{u,i}$ . Use the KSVD algorithm to calculate  $D_{u,i}$  and  $D_{w,i}$  for  $\dot{y}_{w,i}$  and  $\dot{y}_{u,i}$ , respectively. Transform  $x_i$  into  $x_i(D_{u,i}^{-1})^T$  and  $x_i(D_{w,i}^{-1})^T$ .
- 3) If  $x_i(D_{u,i}^{-1})^T$  or  $x_i(D_{w,i}^{-1})^T$  meets Eq. (16), calculate  $\hat{Z}_{u,i}$  or  $\hat{Z}_{w,i}$  using the slope of  $x_i(D_{u,i}^{-1})^T$  or  $x_i(D_{w,i}^{-1})^T$ , respectively.
- 4) Set  $i = i + 1$  and repeat steps 2) and 3) for the next  $\dot{y}_{w,i}$  and  $\dot{y}_{u,i}$  until  $i = N_{max}$ . In this paper, we set  $N_{max} = 50$ .
- 5) For all local signals  $\dot{y}_{w,i}$  and  $\dot{y}_{u,i}$ , use Eq. (18) to calculate the corresponding  $\dot{y}_{w,i}^{global}$  and  $\dot{y}_{u,i}^{global}$ , respectively. Use Eq. (21) to calculate  $\sigma_{HNE,i}$  of  $\dot{y}_{w,i}^{global}$  and  $\dot{y}_{u,i}^{global}$ .
- 6) Search out  $\dot{y}_{w,i}^{global}$  and  $\dot{y}_{u,i}^{global}$  with the largest hybrid negentropy. The corresponding  $\hat{Z}_{w,i}$  and  $\hat{Z}_{u,i}$  are the final harmonic impedances.
- 7) Evaluate the harmonic emission level using Eq. (2).

#### 4. Simulation cases and algorithm analysis

##### 4.1. Error analysis for each method

To compare the calculation error for each method, simulation data are set for the equivalent model shown in Fig. 2(b). PCC data are then generated with 15,000 sample points.

- 1) Harmonic current sources: The amplitude and phase angle of  $\dot{I}_w$  are respectively set as 100 A and  $-30^\circ$ . The amplitude of  $\dot{I}_u$  is  $k$  times  $\dot{I}_w$ . Considering the background harmonic may be unstable when other wind farms are connected to the power grid, we set  $k = 0.6, 0.8, \dots, 1.6$ . The phase angle of  $\dot{I}_u$  is set as  $-30^\circ$ . Sinusoidal fluctuations of  $\pm 10\%$  and random disturbances of  $\pm 5\%$  are added to the amplitude and phase angle of both  $\dot{I}_w$  and  $\dot{I}_u$ .
- 2) Harmonic impedances: We set  $Z_u = 5 + 15j\Omega$  and

$$Z_w = 8 + 23j\Omega \text{ and add } \pm 10\% \text{ sinusoidal fluctuations.}$$

The 15,000 data are divided into 100 segments for simulation analysis. Notably, most methods in this field are based on the statistical properties of the harmonic data. Among these properties, clustering, correlation, and independent properties have received widespread attentions [11–15,19–21,24–26]. We therefore choose typical technologies relating to these properties (i.e. the binary linear regression method, covariance characteristic method and CICA) for comparative analysis of methods. All these methods are used to calculate  $Z_u$ . Because most existing methods cannot calculate  $Z_w$ , the calculation errors of  $Z_w$  are only compared between CICA and ICICA. Average errors of each method are shown in Fig. 8.

It is obvious that the calculation errors of binary linear regression and covariance characteristic method are large. Additionally, these errors increase rapidly with the increasing of background harmonics. These unfavorable calculation results are explained as follows.

The basic functions of binary linear regression method are

$$\begin{cases} U_{pccx}I_{pccx} + U_{pccy}I_{pccy} = Z_{ux}|\dot{I}_{pcc}|^2 + \beta_1 \\ U_{pccy}I_{pccx} - U_{pccx}I_{pccy} = Z_{uy}|\dot{I}_{pcc}|^2 + \beta_2 \end{cases} \quad (22)$$

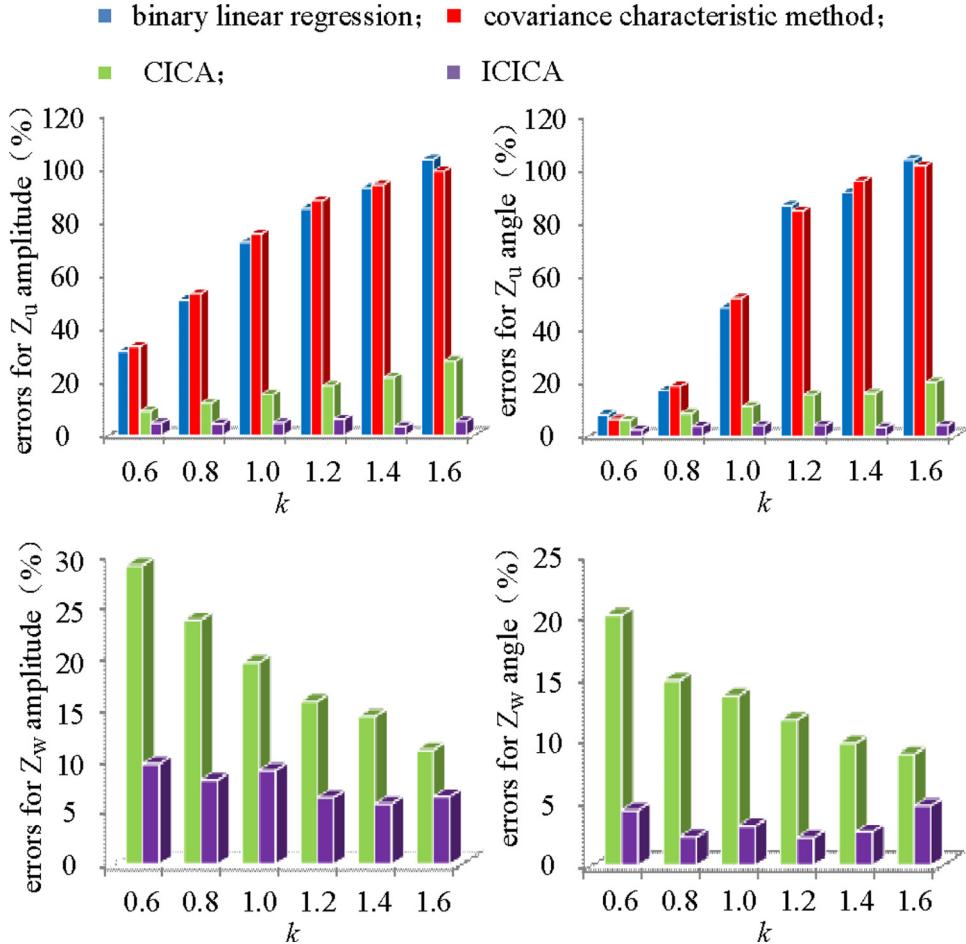
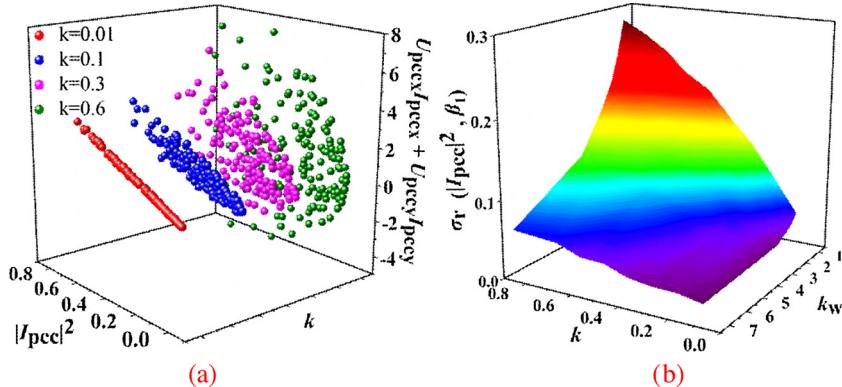
where,

$$\begin{cases} \beta_1 = U_{ux}I_{pccx} + U_{uy}I_{pccy} \\ \beta_2 = U_{uy}I_{pccx} - U_{ux}I_{pccy} \\ \dot{U}_u = \dot{I}_u Z_u \end{cases} \quad (23)$$

and the subscript x and y respectively represent the real and imaginary parts of a complex signal.

In Eq. (22),  $U_{pccx}I_{pccx} + U_{pccy}I_{pccy}$  and  $U_{pccy}I_{pccx} - U_{pccx}I_{pccy}$  are denoted as the dependent variables while  $|\dot{I}_{pcc}|^2$  is denoted as the independent variable.  $Z_{ux}$  and  $Z_{uy}$  are the slopes of the clustering lines while  $\beta_1$  and  $\beta_2$  are the vertical intercepts.

According to Eq. (23), when the background harmonics fluctuate, the vertical intercepts of the clustering lines become unstable, which makes it hard for the dependent and independent variables to cluster

Fig. 8. Errors for  $Z_u$  and  $Z_w$ .Fig. 9. Analysis of binary linear regression method. (a) Degree of linear clustering in each case. (b) Effects of the background harmonics and  $Z_w$ .

along a line. This phenomenon is also presented in Fig. 9(a); that is, with the strengthening background harmonics, the corresponding results are scattered instead of being linearly clustered. Additionally, according to the theory of regression, correlations between  $|I_{pcc}|^2$  and  $\beta_1$  (or  $\beta_2$ ) should be as weak as possible (This correlation can be quantified using the Pearson correlation coefficient in Eq. (24)). Otherwise, the slope of the clustering line will be affected, which decreases the calculation accuracy of  $Z_u$ . Fig. 9(b) shows the effects of  $k$  and  $k_w$  on  $\sigma_r(|I_{pcc}|^2, \beta_1)$  when resetting  $Z_w$  as  $k_w(8 + 23j)$ . When the background harmonics increase or  $Z_w$  is small, the corresponding  $\sigma_r(|I_{pcc}|^2, \beta_1)$  is large and the calculation accuracy of the binary linear regression method is low.

$$\sigma_r(S_1, S_2) = \frac{\sum_{i=1}^N (S_{1,i} - \bar{S}_1)(S_{2,i} - \bar{S}_2)}{\sqrt{\sum_{i=1}^N (S_{1,i} - \bar{S}_1)^2} \sqrt{\sum_{i=1}^N (S_{2,i} - \bar{S}_2)^2}} \quad (24)$$

where  $\bar{S}_1$  and  $\bar{S}_2$  are respectively the mean values of signals  $S_1$  and  $S_2$ .  $N$  is the sample size.

In the covariance characteristic method, the basic assumption that  $\dot{U}_u$  is weakly correlated with  $\dot{I}_{pcc}$  is also based on the premise that the background harmonics are relatively stable and  $Z_w$  is far larger than  $Z_u$ . Otherwise, as shown in Fig. 10, with strengthening background harmonics or decreasing  $Z_w$ , the covariance between  $\dot{U}_u$  and  $\dot{I}_{pcc}$  (denoted as  $C(\dot{U}_u, \dot{I}_{pcc})$ ) is enhanced, which leads to the failure of the covariance characteristic method. The covariance is

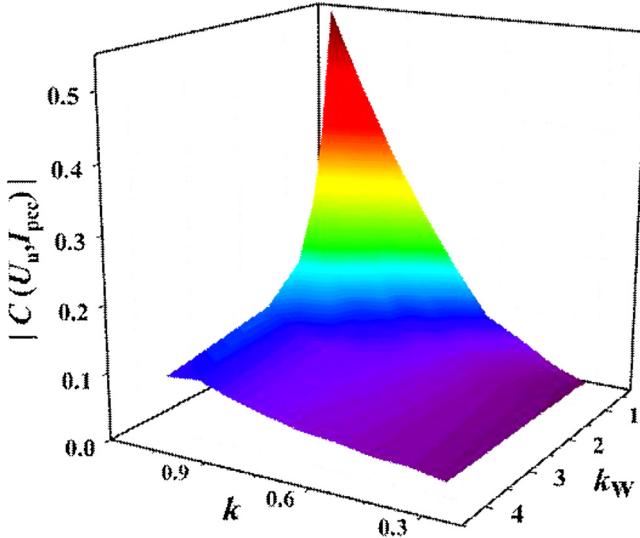


Fig. 10. Analysis of the covariance characteristic method.

$$C(\dot{U}_u, \dot{I}_{pcc}) = E\{[\dot{U}_u - E(\dot{U}_u)][\dot{I}_{pcc} - E(\dot{I}_{pcc})]^H\} \quad (25)$$

where  $E(\cdot)$  denotes a mean value, while the superscript H stands for the Hermitian transpose.

Compared with the results of binary linear regression and the covariance characteristic method, the results of CICA are better but not satisfactory, because the sample size for a single calculation is not large enough. It is thus hard to calculate both  $Z_w$  and  $Z_u$  accurately, especially when their amplitudes are close to each other. Unlike the three methods described, the proposed ICICA calculates the harmonic impedances of both sides accurately through the screening mechanism based on SCA and hybrid negentropy. The effectiveness of this screening mechanism is analyzed in sections IV.C and IV.D.

#### 4.2. Source of CICA errors

As previously mentioned, the signal separation capability and optimization capability may affect the accuracy for the results of CICA. The coefficient of correlation between the amplitudes of the separated signals and the true source signals can be used to assess the signal separation effects of CICA. The signal separation improves as this coefficient approaches 1. By running CICA repeatedly under  $k = 1$ , the harmonic impedance errors and correlation coefficients are obtained. Fig. 11 shows that when the calculated impedance has high accuracy, the coefficient is close to 1, which indicates that the separated signal is consistent with the true source signal. Additionally, when the impedance error is large, the correlation coefficient decreases

substantially. On the basis of the above discussion, the calculation accuracy of CICA is strongly related to the quality of signal separation.

The maximization of non-Gaussianity is set as the optimization objective of CICA and inadequate optimization thus introduces errors into the calculation of harmonic impedances. We decompose the real and imaginary parts of  $[\dot{Y}_u \dot{Y}_w]^T$  and  $[\dot{I}_u^{\text{fast}} \dot{I}_w^{\text{fast}}]^T$  to obtain the corresponding signals  $[Y_{u-x} Y_{u-y} Y_{w-x} Y_{w-y}]^T$  and  $[I_{u-x} I_{u-y} I_{w-x} I_{w-y}]^T$ . Through the transformation of Eq. (19), the negentropies of these signals, denoted as  $J(\cdot)$ , are comparable. The harmonic impedance errors and the comparison results of negentropies are obtained by running CICA repeatedly under  $k = 1$  (Fig. 12).

Fig. 12 shows that in most cases, the negentropies of  $[Y_{u-x} Y_{u-y} Y_{w-x} Y_{w-y}]^T$  are smaller than those of  $[I_{u-x} I_{u-y} I_{w-x} I_{w-y}]^T$ . The difference between their negentropies decreases as the calculation errors decrease. Thus, inadequate optimization is an important source of errors in CICA. Furthermore, in Fig. 12, the real or imaginary part of the separated signals has stronger negentropy than the true source signals at a few times. The reason for this phenomenon is that the central limit theorem is strictly established only when the number of sources is sufficiently large. However, in the modal of Eq. (1), there are only two sources, which, in a few situations, may lead to the non-Gaussianities of the source signals becoming even stronger after linear combination; that is over optimization.

To further analyze the optimization effects of CICA, the probability of each case occurring is presented in Fig. 13. It is indicated that, for CICA, the probability of inadequate optimization is large while over optimization seldom occurs. Because the central limit theorem is approximately valid for this model, it is scarcely possible that both the real and imaginary parts of a source signal are over optimized simultaneously.

To improve the signal separation and optimization capabilities of the algorithm, sparse and hybrid negentropy screening mechanisms are proposed.

#### 4.3. Sparse screening mechanism

In this study, sparse screening improves the signal separation ability of CICA. By quantifying the degree of similarity between the searched local signals and their corresponding true source signals, the effects of sparse screening index (i.e.  $\sigma_{\text{sparse}}$ ) can be examined. The similarity degree of two arbitrary signals can be assessed using Eq. (26). A larger  $\sigma_{\text{similar}}$  corresponds to a larger similarity degree.

$$\sigma_{\text{similar}} = \frac{1}{\|s_1/s_{1,\max} - s_2/s_{2,\max}\|_2} \quad (26)$$

where  $s_{1,\max}$  and  $s_{2,\max}$  are respectively the maximum amplitudes of  $s_1$  and  $s_2$ . The symbol  $\|\cdot\|$  represents the 2-norm.

Taking the calculation of  $Z_u$  under  $k = 1$  as an example, the indices  $\sigma_{\text{sparse}}$  and  $\sigma_{\text{similar}}$  for each local signal are shown in Fig. 14(a). The local

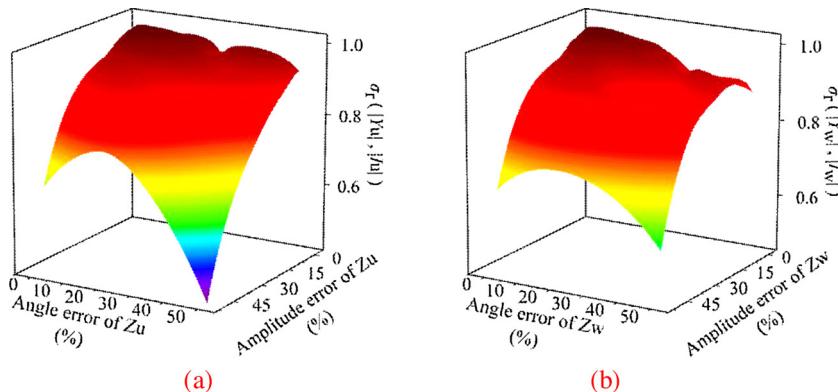


Fig. 11. Calculation errors and correlation coefficients. (a)  $Z_u$ . (b)  $Z_w$ .

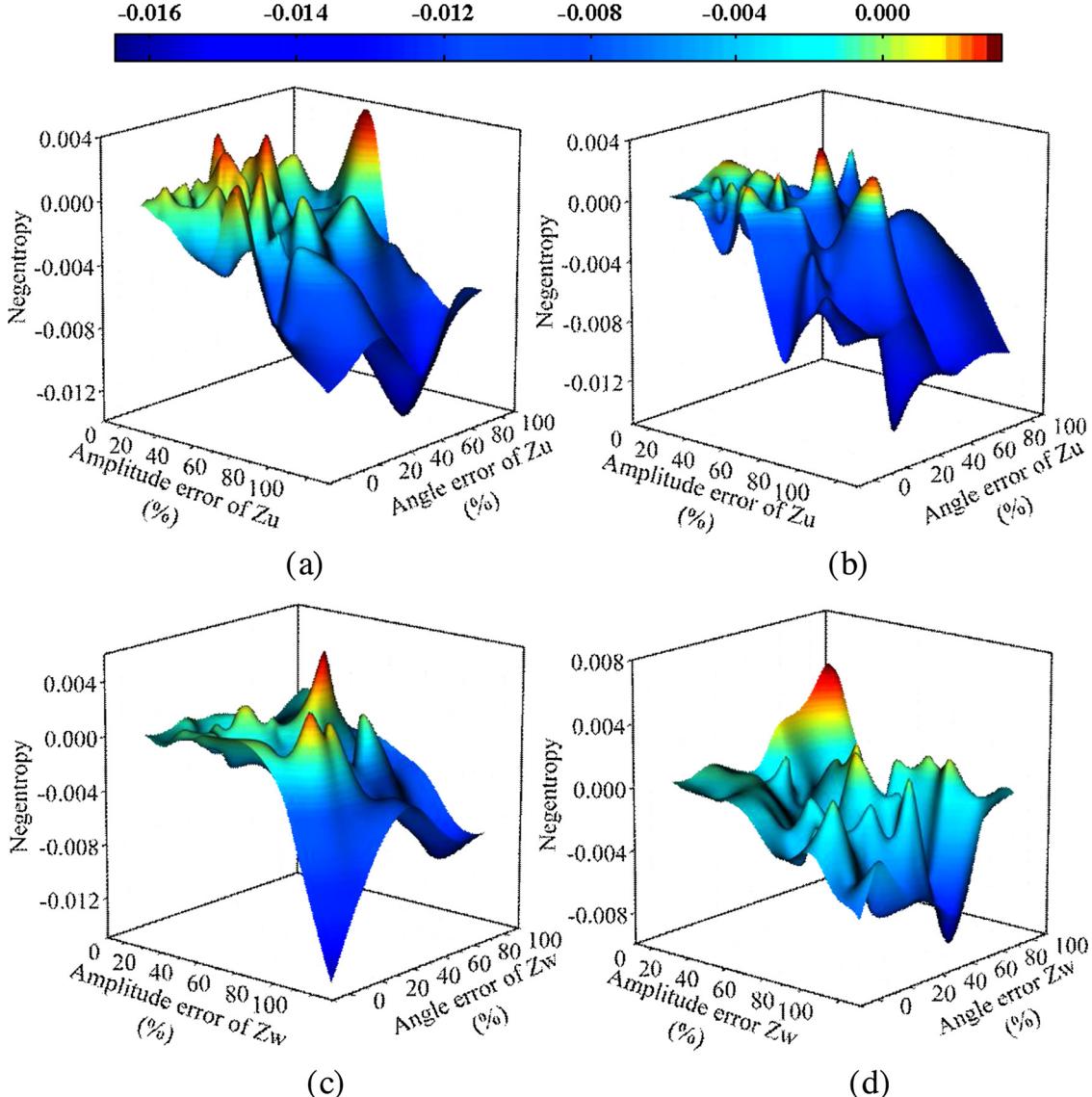


Fig. 12. Comparison of negentropies. (a)  $J(Y_{u-x}) - J(I_{u-x})$ . (b)  $J(Y_{u-y}) - J(I_{u-y})$ . (c)  $J(Y_{w-x}) - J(I_{w-x})$ . (d)  $J(Y_{w-y}) - J(I_{w-y})$ .

signals with high  $\sigma_{\text{similar}}$  clearly satisfy the sparse screening criterion, i.e.  $\sigma_{\text{sparse}} < \varepsilon$  ( $\varepsilon = 0.02$ ), while the value of  $\sigma_{\text{similar}}$  quickly drops with increasing  $\sigma_{\text{sparse}}$ . Therefore, through sparse screening we can find the local signals that are consistent with the true source signals.

However, among the local signals satisfying  $\sigma_{\text{sparse}} < \varepsilon$ , the index  $\sigma_{\text{similar}}$  is still low for those meeting Eq. (14) which leads to local signals being screened out even when they are not similar to the true source

signals. We use the index  $\sigma_{\text{wrong}}$ , defined by Eq. (27), to analyze this erroneous screening phenomenon. When  $\sigma_{\text{wrong}} \approx 0$ , Eq. (14) holds and the corresponding local signals are wrongly screened out. As shown in Fig. 14 (b), for local signals satisfying  $\sigma_{\text{sparse}} < \varepsilon$  and with high  $\sigma_{\text{similar}}$ , the corresponding index  $\sigma_{\text{wrong}}$  is obviously larger than zero. In contrast, for local signals satisfying  $\sigma_{\text{sparse}} < \varepsilon$  but with low  $\sigma_{\text{similar}}$ , the index  $\sigma_{\text{wrong}}$  is close to zero, which means these local signals are

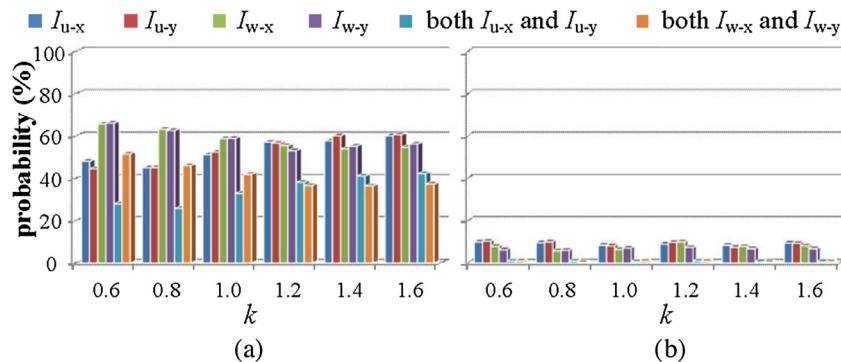


Fig. 13. Probability of each case occurring for CICA. (a) Inadequate optimization. (b) Over optimization.

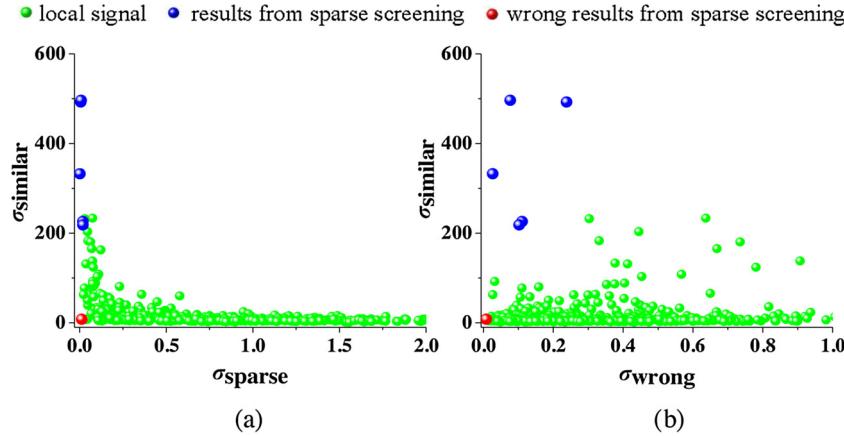


Fig. 14. Relationships of three indices. (a)  $\sigma_{\text{similar}}$  vs  $\sigma_{\text{sparse}}$  (b)  $\sigma_{\text{similar}}$  vs  $\sigma_{\text{wrong}}$ .

searched out mistakenly.

$$\sigma_{\text{wrong}} = \max \left( \left| \frac{[\hat{I}_{u,i}^{\text{fast}} (\mathbf{D}_{u,i}^{-1})^T]_{1,j}}{[\hat{I}_{c,i}^{\text{fast}} (\mathbf{D}_{u,i}^{-1})^T]_{2,j}} - \frac{[\hat{I}_{u,i}^{\text{fast}} (\mathbf{D}_{u,i}^{-1})^T]_{1,k}}{[\hat{I}_{c,i}^{\text{fast}} (\mathbf{D}_{u,i}^{-1})^T]_{2,k}} \right| \right), \forall j \neq k \quad (27)$$

where the meanings of the subscripts are the same as in Eq. (16). Consequently, through sparse screening, local signals that are consistent with the true source signals can be found, but a few wrong signals may also be obtained. We thus propose a new hybrid negentropy index for secondary screening.

#### 4.4. Hybrid negentropy screening mechanism

We take the calculation of  $Z_u$  as an example to illustrate the effects of hybrid negentropy screening. Fig. 15 (a) shows that after sparse screening, some results still have large errors. Notably, as  $\sigma_{\text{HNE},i}$  increases, the corresponding  $\hat{Z}_{u,i}$  becomes more accurate. The result with the highest accuracy can thus be selected adopting the proposed hybrid negentropy screening. In essence,  $y_{\text{global}}$  with large error is still the linear combination of source signals and its non-Gaussianity is weak according to the central limit theorem. Incorrect results can therefore be eliminated through hybrid negentropy screening. Fig. 15(b) shows that the negentropies of the final result and the true source signal are highly consistent. In addition, both the real and imaginary parts of the final result have large negentropies. The inadequate optimization of CICA is therefore overcome.

#### 5. Field verification

An actual multi-wind farms system comprising three wind farms (Fig. 16) is used to test each method. The 35-kV grid connection point of the test wind farm is set as the PCC. The maximum value of the total harmonic distortion voltage for the PCC is as high as 3.3%, which exceeds the Chinese standard limits. It is thus necessary to evaluate the harmonic emission level for wind farms.

The voltages and currents are measured from the PCC. A relatively high sampling frequency ensures the statistical characteristics of data. However, random noise is amplified with high a sampling frequency [35]. Our sampling frequency is empirically set at 10 kHz, which ensures the statistical properties of data and avoids the effect of noise at the same time. All harmonic data are obtained through fast Fourier transformation and there are five data points per second.

At each harmonic voltage at the PCC, the 5th harmonic ratio is relatively high and we thus take the 5th harmonic as an example for analysis. The 5th harmonic data for a period of 5 min are shown in Fig. 17. The corresponding fundamental currents fluctuate weakly, and the harmonic impedances are thus relatively stable to a certain degree.

However, the relatively long time interval still increases the risk of impedance variation. In this paper, to further keep the harmonic impedances constant within a single computation segment, PCC data are equally divided into 10 segments. It is thus considered that harmonic impedances are stable in each short period of 30 s.

The calculated 5th harmonic impedances for each method are shown in Fig. 18. The results of traditional methods clearly fluctuate greatly over the short interval of 5 min, which does not conform to

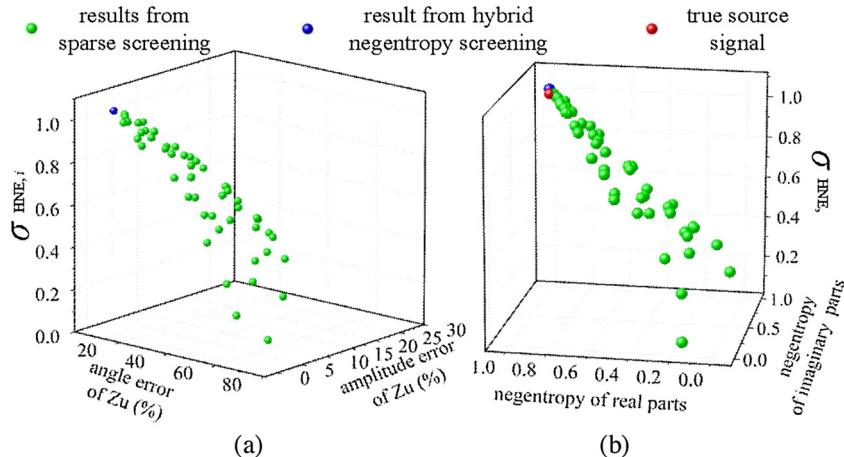


Fig. 15. Analysis of hybrid negentropy screening. (a) Errors analysis. (b) Negentropy analysis.

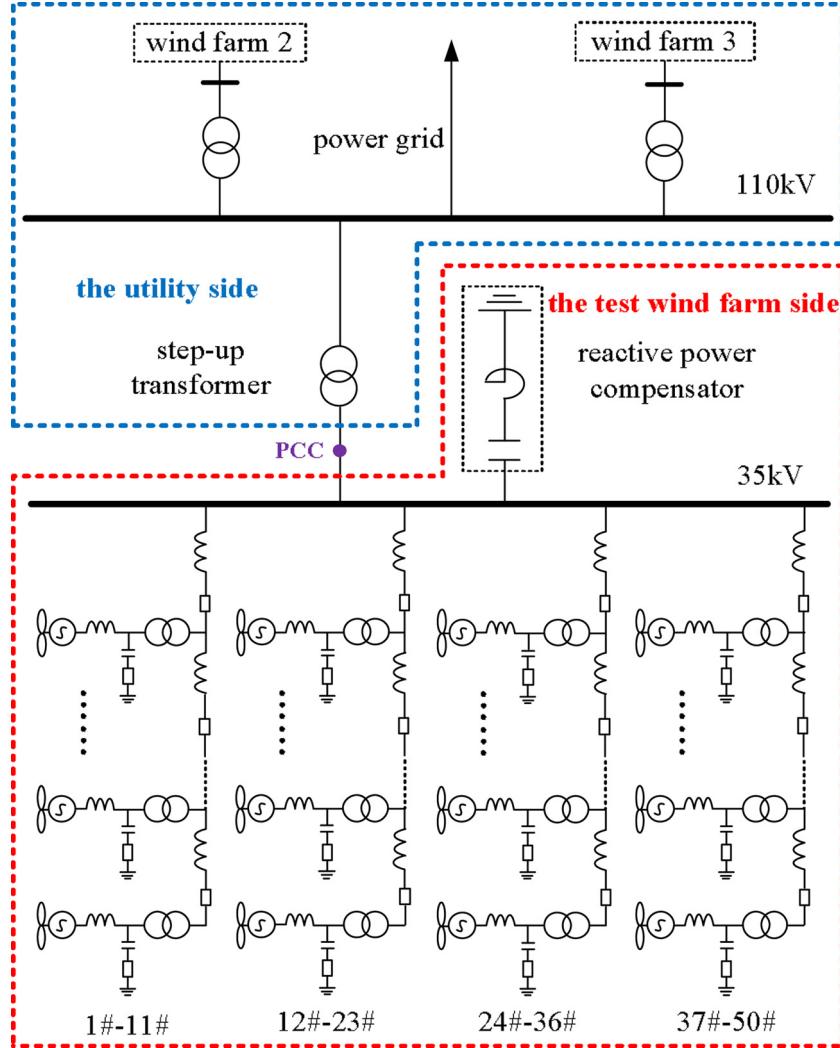


Fig. 16. Topological structure of a multi-wind farms system.

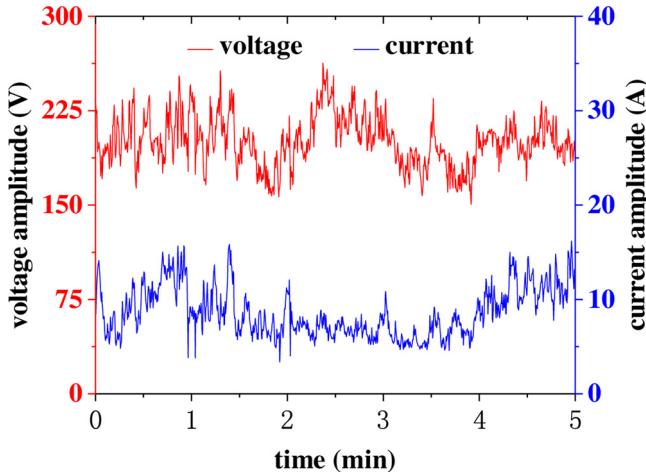


Fig. 17. The 5th harmonic voltages and currents at the PCC.

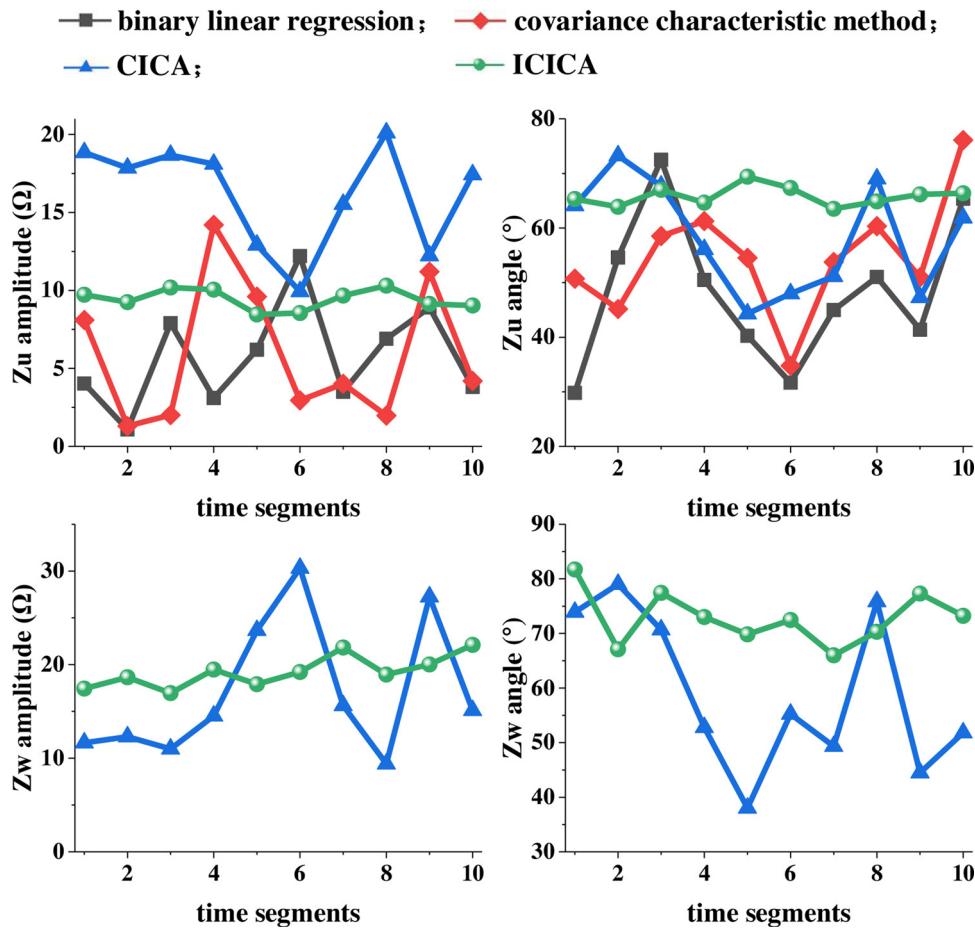
actual conditions. The reasons for these unfavorable results are that, in the multi-wind farms system, the necessary conditions for binary linear regression and covariance characteristic method do not hold. On the one hand, the utility side has two other wind farms, which potentially destabilize the background harmonics. On the other hand,  $|Z_w|$  is not much larger than  $|Z_u|$  because of the effect of filters. Furthermore, a

robust result of CICA relies heavily on thousands of samples for a single calculation. However, the sample size is limited to ensure the invariability of  $Z_w$  and  $Z_u$ . By comparison, the results of the proposed ICICA method are more stable, which is consistent with reality.

We further validate the harmonic impedances obtained using the proposed ICICA method by comparing average values (i.e.  $9.71 \Omega$  and  $19.58 \Omega$ ) with the corresponding reference values. The harmonic impedance of the step-up transformer ( $10.3 \Omega$ ) can be considered as the reference value of  $|Z_u|$ . This assumption is reasonable to some extent because  $Z_u$  is commonly dominated by the connection transformer [13,36]. Additionally, by building the model of the test wind farm in ETAP software and using the frequency scan program, we calculate the reference value of  $|Z_w|$  as  $18.52 \Omega$ . Obviously, both  $|Z_u|$  and  $|Z_w|$  calculated from ICICA are close to their reference values.

Theoretically, increasing the sample size for a single calculation enhances the statistical properties of the data and improve the calculation accuracy of CICA. However, a large sample size requires a long sampling time, which may lead to harmonic impedance fluctuation and new calculation errors. To explore the effect of a large sampling set, we use the whole 5 min of data for a single calculation. The calculated  $Z_u$  and  $Z_w$  are  $21.3 \Omega$  and  $13.6 \Omega$ , respectively, which still differ greatly from the reference values. By comparison, ICICA obtains accurate results without large samples, and thus avoids the errors generated from impedances changing.

Therefore, through  $Z_u$  and  $Z_w$  calculated in ICICA, the harmonic

Fig. 18.  $Z_u$  and  $Z_w$  calculated using each method.

emission levels of utility and the test wind farm side at the PCC are respectively obtained as 163.21 V and 119.88 V. Using the proposed method, the harmonic emission levels of other wind farms can also be assessed.

To further compare the effects of these methods, the calculation results for each typical order (i.e. the 5th, 7th, 11th, 13th, and 19th harmonics) are specifically analyzed (Tables 1 and 2) because their corresponding contents are relatively high. Notably, for the 5th and 11th harmonics, the reference amplitudes of  $Z_w$  are only a little larger than  $Z_u$ . Furthermore,  $Z_w$  is even smaller than  $Z_u$  for the 13th and 19th harmonics owing to the effects of harmonic filters at the wind farm. The errors of  $Z_u$  calculated by traditional methods are therefore large in these cases because  $|Z_w| > |Z_u|$  does not hold. Additionally, for the 7th harmonic,  $Z_w$  is obviously larger than  $Z_u$  and the calculation errors of  $Z_u$  are decreased for each method. However, the accuracies of the traditional methods are still unsatisfactory under the effects of the strong background harmonics.

**Table 1**  
Calculation results of  $Z_u$ .

	Harmonic order	5th	7th	11th	13th	19th
Reference value of $Z_u$ ( $\Omega$ )		10.30	14.42	22.64	26.75	39.08
Binary linear regression	$Z_u$ ( $\Omega$ )	6.04	10.19	32.18	39.49	18.19
	error (%)	-41.36	-29.33	42.14	47.63	-53.45
Covariance characteristic method	$Z_u$ ( $\Omega$ )	5.82	17.99	31.05	38.92	18.96
	error (%)	-43.50	24.76	37.15	45.50	-51.48
CICA	$Z_u$ ( $\Omega$ )	15.24	16.41	16.92	37.96	20.10
	error (%)	47.96	13.80	-25.27	41.91	-48.57
ICICA	$Z_u$ ( $\Omega$ )	9.71	15.34	24.17	24.44	43.08
	error (%)	-5.73	6.38	6.76	-8.64	10.24

**Table 2**  
Calculation results of  $Z_w$ .

Harmonic order	5th	7th	11th	13th	19th	
Reference value of $Z_w$ ( $\Omega$ )	18.52	47.06	41.46	19.36	15.22	
CICA	$Z_w$ ( $\Omega$ )	16.88	65.27	56.32	16.55	16.49
	error (%)	-8.86	38.70	35.84	-14.51	8.34
ICICA	$Z_w$ ( $\Omega$ )	19.58	41.79	45.15	20.87	14.24
	error (%)	5.72	-11.20	8.90	7.80	-6.44

Meanwhile, for CICA, the regularity for the calculation errors of  $Z_w$  is contrary to the regularity of  $Z_u$ . Specifically, for the 7th harmonic, where the reference amplitude of  $Z_w$  is obviously larger than  $Z_u$ , the error in  $Z_w$  is large, while the error in  $Z_u$  is relatively small. Additionally, for the 13<sup>th</sup> and 19<sup>th</sup> harmonics, because the reference amplitude of  $Z_w$  is obviously smaller than  $Z_u$ , the corresponding accuracy of  $Z_w$  are significantly increased, whereas the error in  $Z_u$  is large. Additionally, for situations where the amplitudes of  $Z_w$  are just a little larger than  $Z_u$  (i.e. the 5th and 11th harmonics) the calculation accuracies of both  $Z_u$  and  $Z_w$  are low. Of note, although the error in  $Z_w$  is just -8.86% for the 5th harmonic, the corresponding error of each segment in Fig. 18 is still large. These regularities for the calculation errors of CICA also coincide with the conclusions drawn from simulations in Ref. [14].

In contrast with existing methods, the calculation accuracy of ICICA is always satisfactory for both  $Z_u$  and  $Z_w$  at each harmonic order.

## 6. Conclusion

The present paper proposes an improved complex ICA algorithm to evaluate the harmonic emission levels of wind farm. The main

contributions of the paper are as follows.

- (1) The unstable background harmonics and the installation of harmonic filters at a wind farm increase the challenges for the existing methods for harmonic emissions evaluation.
- (2) To improve the traditional Complex ICA method such that it is suitable for wind farm systems, the sources of error in the algorithm are analyzed in terms of both the signal separation and optimization abilities.
- (3) The limitations of complex ICA are overcome using two screening mechanisms based on sparse component analysis and the maximization of hybrid negentropy. The sparse screening process finds local signals that strongly coincide with the true source signals while the hybrid negentropy screening process further improves the accuracy of the screening results.
- (4) Simulations and field cases demonstrate that the evaluation errors of traditional methods are large for wind farm scenes. In contrast, the proposed method is valid even when the background harmonics are unstable and/or the harmonic impedance of the wind farm side is not much larger than that of the utility side. In addition, compared with complex ICA, the novel method requires only a small sample size.

In future work, we will further study the proposed technique to enhance the efficiency of the local search.

#### Conflict of interest

None.

#### Acknowledgements

This paper is supported by the National Natural Science Foundation of China (#51477105), and the State Grid Corporation of China Research Program (SGSHDK00DWJS1800225).

#### References

- [1] S.T. Tentzerakis, S.A. Papathanassiou, An investigation of the harmonic emissions of wind turbines, *IEEE Trans. Energy Convers.* 22 (March (1)) (2007) 150–158.
- [2] S. Liang, Q. Hu, L. Wei-Jen, A survey of harmonic emissions of a commercially operated wind farm, *IEEE Trans. Ind. Appl.* 48 (June (3)) (2015) 1115–1123.
- [3] S.T. Tentzerakis, S.A. Papathanassiou, An investigation of the harmonic emissions of wind turbines, *IEEE Trans. Energy Convers.* 22 (March (1)) (2007) 150–158.
- [4] L. Sainz, J.J. Mesas, R. Teodorescu, P. Rodriguez, Deterministic and stochastic study of wind farm harmonic currents, *IEEE Trans. Energy Convers.* 25 (December (4)) (2010) 1071–1080.
- [5] L. Ye, L.Z. Lin, Study of superconducting fault current limiters for system integration of wind farms, *IEEE Trans. Appl. Supercond.* 20 (June (3)) (2010) 1233–1237.
- [6] M.H.J. Bollen, K. Yang, Harmonic aspects of wind power integration, *J. Modern Power Syst. Clean Energy* 1 (July (1)) (2013) 14–21.
- [7] Chinese Standards GB/T 14549-93, Quality of Electric Energy Supply Harmonic in Public Supply Network, Chinese Standard Press, 1993.
- [8] R. Vazquez, M.A. Muñoz, M. Alonso, H. Amaris, C. Alvarez, Background harmonic distortion measurement at power networks with wind farms, International Conference on Renewable Energies and Power Quality (ICREPQ'16), Madrid (Spain), 2016, pp. 4–6 vol. 1, May (14).
- [9] Fritz Santjer, Bernd Weise, Tina Pausch, Johannes Brombach, Aspects for improvement of measurement and assessment procedures of harmonic emission of wind power plants, DEWEK, May, 2015.
- [10] H. Saad, A. Schwob, Y. Vernay, Study of resonance issues between HVDC link and power system components using EMT simulations, 2018 Power Systems Computation Conference (PSCC), Dublin, Ireland, June, 2018.
- [11] J. Hui, W. Freitas, J.C.M. Vieira, H. Yang, Y. Liu, Utility harmonic impedance measurement based on data selection, *IEEE Trans. Power Deliv.* 27 (October (4)) (2012) 2193–2203.
- [12] J. Hui, H. Yang, S. Lin, M. Ye, Assessing utility harmonic impedance based on the covariance characteristic of random vectors, *IEEE Trans. Power Deliv.* 25 (July (3)) (2010) 1778–1786.
- [13] F. Karimzadeh, S. Esmaeili, S.H. Hosseiniyan, A novel method for noninvasive estimation of utility harmonic impedance based on complex independent component analysis, *IEEE Trans. Power Deliv.* 30 (August (4)) (2015) 1843–1852.
- [14] F. Karimzadeh, S. Esmaeili, S.H. Hosseiniyan, Method for determining utility and consumer harmonic contributions based on complex independent component analysis, *Gener. Transm. Distrib. IET* 10 (2) (2016) 526–534.
- [15] X. Zhao, H. Yang, A new method to calculate the utility harmonic impedance based on FastICA, *IEEE Trans. Power Deliv.* 31 (February (1)) (2016) 381–388.
- [16] F. Xu, H. Yang, J. Zhao, Z. Wang, Y. Liu, Study on constraints for harmonic source determination using active power direction, *IEEE Trans. Power Deliv.* 33 (December (6)) (2018) 2683–2692.
- [17] C. Li, W. Xu, T. Tayasanant, A “critical impedance”-based method for identifying harmonic sources, *IEEE Trans. Power Deliv.* 19 (April (2)) (2004) 671–678.
- [18] W. Xu, X. Liu, Y. Liu, An investigation on the validity of power-direction method for harmonic source determination, *IEEE Trans. Power Deliv.* 18 (January (1)) (2003) 214–219.
- [19] Zhang Yunyan, Cai Wei, Deng Chun, Xu Yonghai, A harmonic measurement mode based on partial least-squares regression method, *Electric Power Conference, IEEE*, 2009.
- [20] Jie Luo, Tianlei Zang, Ling Fu, A novel method to determine the harmonic responsibility for power distribution network including photovoltaic system, 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2), IEEE, 2017.
- [21] Xu Yonghai, Huang Shun, Liu Yingying, Partial Least-Squares Regression Based Harmonic Emission Level Assessing at the Point of Common Coupling, *International Conference on IEEE Power System Technology, 2006. PowerCon 2006* (2006).
- [22] H. Yang, P. Porotte, A. Robert, Assessing the harmonic emission level from one particular customer, *Proceedings of the 3rd International Conference on Power Quality (ICPQ)*, Amsterdam, Netherlands, 1994.
- [23] Yao Xiao, Jean-Claude Maun, Hedi Ben Mahmoud, Harmonic impedance measurement using voltage and current increments from disturbing loads, *Harmonics and Quality of Power, 2000. Proceedings. Ninth International Conference on*. Vol. 1. IEEE (2000).
- [24] Qi Fei, Li Jianwen, Li Yonggang, Sun Wei, Li Zhongliang, Research on the responsibility partition of harmonic pollution and harmonic impedance based on the total least-squares regression method, *International Conference on Power System Technology, IEEE*, 2014.
- [25] Y. Shujun, Y. Jing, W. Yan, L. Longhui, M. Luyao, The harmonic assessment method based on the generalized ridge regression, *Power & Energy Engineering Conference, IEEE*, 2012.
- [26] Y. Wang, W. Xu, J. Yong, K.-L. Chen, Estimating harmonic impact of individual loads using harmonic phasor data, *Int. Trans. Electr. Energy Syst.* 27 (October) (2017) e2384.
- [27] Y. Xianchuan, X. Jindong, H. Dan, X. Haihua, A new blind image source separation algorithm based on feedback sparse component analysis, *Signal Process.* 93 (1) (2013) 288–296.
- [28] Y. Li, S.-I. Amari, A. Cichocki, D.W.C. Ho, S. Xie, Underdetermined blind source separation based on sparse representation, *IEEE Trans. Signal Process.* 54 (February (2)) (2006) 423–437.
- [29] V.G. Reju, S.N. Koh, I.Y. Soon, An algorithm for mixing matrix estimation in instantaneous blind source separation, *Signal Process.* 89 (March (9)) (2009) 1762–1773.
- [30] X. Wilsun, Y. Liu, A method for determining customer and utility harmonic contributions at the point of common coupling, *IEEE Trans. Power Deliv.* 15 (April (2)) (2000) 804–811.
- [31] M. Novey, T. Adali, Complex ICA by negentropy maximization, *IEEE Trans. Neural Networks* 19 (April (4)) (2008) 596–609.
- [32] Na Li, Peng Du, Huijie Zhao, Independent component analysis based on improved quantum genetic algorithm: application in hyperspectral images, *IEEE International Geoscience and Remote Sensing Symposium* vol. 6, (2005), pp. 4323–4326.
- [33] A. Vahid, F. Saideh, S. Saeid, Blind separation of image sources via adaptive dictionary learning, *IEEE Trans. Image Process.* 21 (June (6)) (2012) 2921–2930.
- [34] M. Aharon, M. Elad, A. Bruckstein, K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation, *IEEE Trans. Signal Process.* 54 (November (11)) (2006) 4311–4322.
- [35] L. Xiong, F. Zhuo, F. Wang, X. Liu, M. Zhu, A fast orthogonal signal generation algorithm characterized by noise immunity and high accuracy for single-phase grid, *IEEE Trans. Power Electron.* 31 (March (3)) (2016) 1847–1851.
- [36] T. Pfajfar, B. Blazic, I. Papic, Harmonic contributions evaluation with the harmonic current vector method, *IEEE Trans. Power Deliv.* 23 (January (1)) (2008) 425–433.

**Jinshuai Zhao** was born in Yibin, Sichuan Province, China, in 1992. He received the B.S. degree from the College of Electrical Engineering and Information Technology, Sichuan University, Chengdu, China, in 2015, where he is currently pursuing the Ph.D. degree in power system and its automation. His main research interests are power system harmonic analysis and evaluation. ([jinshuai\\_zhao@scu.edu.cn](mailto:jinshuai_zhao@scu.edu.cn)).

**Honggeng Yang** was born in Chengdu, Sichuan, China, in 1949. He received the B.S. degree in electrical engineering from Chengdu University of S & T, China, and M.S. degree in electrical engineering from Harbin Institute of Technology, China, and PhD degree in electrical engineering from Liege University, Belgium, in 1982, 1985, and 1997, respectively. Since 2002, he has been a professor in the College of Electrical Engineering and Information Technology, Sichuan University, China. His research interests include power quality and reactive power/voltage control. ([PQlab99@126.com](mailto:PQlab99@126.com))

**Aiqiang Pan** was born in Shandong Province of China, in 1984. He received the Electrical Engineering M.S. degree in Shanghai Jiao Tong University. And now He is a senior engineer in State Grid Shanghai Electric Power Research Institute, Shanghai, China. His main research interests focus on Power Quality and harmonic analysis. ([aiqiang\\_pan@163.com](mailto:aiqiang_pan@163.com))

**Fangwei Xu** was born in Renshou, Sichuan Province of China, in 1978. She received the Electrical Engineering M.S. degree in Chongqing University in 2004. She received the Power System and its Automation PhD degree in Sichuan University in 2012. And now she is an associate professor in the College of Electrical Engineering and Information Technology, Sichuan University, Chengdu, Sichuan Province, China. Her main research interests focus on Power Quality characteristics extraction, harmonic source determination and harmonic emissions assessment, etc. ([xufangwei@scu.edu.cn](mailto:xufangwei@scu.edu.cn)).