

Improved utility harmonic impedance measurement based on robust independent component analysis and bootstrap check

Feiyu Chen¹, Nanping Mao², Yang Wang¹, Ying Wang¹ ✉, Xianyong Xiao¹

¹College of Electrical Engineering, Sichuan University, Chengdu, People's Republic of China

²Electric Power Research Institute of State Grid Zhejiang Electric Power Company, Hangzhou, People's Republic of China

✉ E-mail: tuantuan1125@scu.edu.cn

ISSN 1751-8687

Received on 1st August 2019

Revised 25th September 2019

Accepted on 29th October 2019

E-First on 3rd February 2020

doi: 10.1049/iet-gtd.2019.1153

www.ietdl.org

Abstract: In this study, a new method for utility harmonic impedance measurement based on robust independent component analysis (RICA) and bootstrap check is proposed. The RICA approach which adopts exact line search to seek the optimal step-size in iteration has better performance for short data records. At the same time, a bootstrap check technique to remove the singular solutions of this sort of independent component analyses methods is presented as well. This verification technique is performed through judging the dispersion degree of the estimates of the bootstrap samples which is quantified by the coefficient of variation. Based on the results of computer simulation and field test, it is proved that the combination of the RICA approach with the bootstrap check technique can reduce the adverse impact of the background harmonic fluctuation and the singular solutions on impedance estimation effectively, and satisfactory measurement results can be acquired.

1 Introduction

With the widespread use of non-linear loads such as electronic power converters at each voltage level in electric power systems, harmonic pollution is becoming a more and more serious problem in the power quality area in recent years [1–3]. From the point of the utility side, the harmonic voltages in the grid are mainly caused by the harmonic currents injected by various non-linear loads from the customers. However, to the users, they normally require high-quality power supply by the utility. Therefore, the responsibility division of harmonic pollution at the point of common coupling (PCC) has become an urgent problem to be solved for both sides [4–6].

Norton equivalent circuit [7] is an often-used model to analyse the harmonic responsibility division problem. In this model, the key point of achieving reliable and valid partition results mainly lies in the accurate measurement of the utility harmonic impedance. The invasive methods [8, 9] for harmonic impedance estimation, in which harmonic currents are injected by specific equipment or branch circuits are switched, cannot be widely applied since their large adverse impacts on the networks. The non-invasive methods which calculate the harmonic impedance only by the harmonic voltage and current measured at the PCC during normal grid operation, in contrast, have been developed extensively due to their disturbance-free characteristics. The typical non-invasive methods for utility harmonic impedance measurement mainly comprise fluctuation method [10], linear regression method [11–14] and independent component analysis (ICA) method [15–18]. The fluctuation method and linear regression method are susceptible to the background harmonic variation inherently. The ICA method is a relatively new method used for harmonic impedance estimation based on the independence or weak correlation between the source signals of the utility and customer side, and the disturbance of the background harmonics is suppressed to a certain degree in this method. However, it needs long-term data to avoid falling into a local optimum which would lead to large-deviated singular solution, so it may lose efficacy for real-time estimation.

In this paper, a new method for utility harmonic impedance measurement based on robust ICA (RICA) and bootstrap check is proposed. The main research work and contribution of this paper are summarised as follows:

- Three published methods for harmonic impedance estimation are tested and compared by a large amount of field data in Sections 2.2, 2.3 and some shortcomings show up in them. Therefore, the RICA algorithm is introduced for signal demixing and impedance calculation in Sections 2.4 and 3.1.
- Although RICA can improve the performance of ICA method for finite samples, as the same as ICA method, some singular solutions still exist due to saddle points and spurious local extrema. In Section 3.2, a singular value elimination technique based on bootstrap is proposed. The bootstrap check technique removes the singular solutions by judging the dispersion degree of the estimates by the bootstrapping samples of the original data. This technique can eliminate the negative impact of singular solutions by RICA on impedance measurement effectively.
- Computer simulation and field test are conducted to verify the validity and applicability of the proposed method in Sections 4 and 5, respectively. The experimental results show that the estimated error can be limited within 4% by the improved measurement method even when the background harmonic fluctuation is much larger than the customer side. For short data size (300 points in our simulation), estimated error is confined <5%. It is proved that the proposed method can reduce the adverse effect of the background harmonics to an acceptable degree and can acquire satisfactory estimation results in small sample space.

2 Basic principles

2.1 Equivalent model for harmonic analysis

The Norton equivalent circuit model is adopted for harmonic analysis and impedance measurement in this paper as shown in Fig. 1.

In Fig. 1, I_s and I_c represent the harmonic current sources of the utility and customer side, respectively. Z_s and Z_c are the harmonic impedances of the utility and customer side, respectively. U_{pcc} and I_{pcc} represent the harmonic voltage and harmonic current at the PCC, respectively.

On the basis of the equivalent circuit above, (1) can be derived according to the Kirchhoff laws and superposition principle as follows:

$$\begin{cases} U_{pcc} = \frac{Z_s Z_c}{Z_s + Z_c} I_s + \frac{Z_s Z_c}{Z_s + Z_c} I_c \\ I_{pcc} = \frac{Z_s}{Z_s + Z_c} I_s - \frac{Z_c}{Z_s + Z_c} I_c \end{cases} \quad (1)$$

One can draw a conclusion by analysing the equation above that U_{pcc} and I_{pcc} are jointly contributed by I_s and I_c , and the contribution ratios are related to the harmonic impedances of both sides.

2.2 Review of published methods

In recent years, the non-invasive methods have been developed extensively due to their disturbance-free characteristics. In this section, three mainstream non-invasive methods for harmonic impedance measurement are reviewed, which are the fluctuation method, linear regression method and classical ICA method, respectively.

2.2.1 Fluctuation method: In reference [10], according to the Norton equivalent circuit, (2) can be deduced as below

$$\frac{\Delta U_{pcc}}{\Delta I_{pcc}} = \frac{\Delta I_c + \Delta I_s}{(1/Z_s)\Delta I_c + (1/Z_c)\Delta I_s} \quad (2)$$

where ΔU_{pcc} and ΔI_{pcc} denote the harmonic voltage fluctuation and the harmonic current fluctuation at the PCC, respectively. ΔI_s and ΔI_c denote the fluctuation of I_s and I_c , respectively. So according to (2), when the background harmonic is stable and ΔI_s is close to zero, namely ΔI_c dominates the fluctuation at the PCC, then the ratio of ΔU_{pcc} to ΔI_{pcc} is close to Z_s

$$\frac{\Delta U_{pcc}}{\Delta I_{pcc}} = Z_s \quad (3)$$

The approach of judging which side dominates the fluctuation at the PCC is by the symbol of the real part of Z_s . When the real part of Z_s is positive, the ratio of ΔU_{pcc} to ΔI_{pcc} can be believed as the valid estimation of Z_s . Conversely, if the real part of Z_s is negative, the estimated result is invalid and these samples should be discarded. In practical use of the fluctuation method, since the background harmonic fluctuation exists extensively, the adverse effect of ΔI_s cannot be neglected, so satisfactory results cannot be acquired in most practical cases.

2.2.2 Linear regression method: In the partial least square regression method in [12], Thevenin equivalent circuit is adopted and the complex-valued parameters of the equivalent equation are decomposed into the real parts and the imaginary parts. The decomposed equations are written as shown in (4), where subscripts x and y represent the real part and the imaginary part of the complex number, respectively. Obviously U_{pcc} and I_{pcc} are measurable, so the utility harmonic impedance can be calculated by solving the regression coefficients of (4). To perform linear regression algorithm, the variable V_s which represents the harmonic voltage source in the utility side is assumed to be constant in this method, so when the background harmonic varies

in practical cases, large errors would arise as the same as the fluctuation method

$$\begin{cases} U_{pccx}I_{pccy} - I_{pccx}U_{pccy} = V_{sx}I_{pccy} - V_{sy}I_{pccx} \\ \quad - Z_{sy}(I_{pccx}^2 + I_{pccy}^2) \\ U_{pccx}I_{pccx} + I_{pccy}U_{pccy} = V_{sx}I_{pccx} + V_{sy}I_{pccy} \\ \quad + Z_{sx}(I_{pccx}^2 + I_{pccy}^2) \end{cases} \quad (4)$$

To reduce the adverse impact of utility-side harmonic variation and sift out the estimated results approaching the actual impedance value, in reference [14], on basis of the complex-number least-square linear regression method, two techniques of data selection are proposed. The first selection technique is based on the idea that the calculated impedance samples may gather to the true value when the variance of I_{pcc} ($\text{Var}(I_{pcc})$) augments. If the variance of the 10% impedance samples with the largest $\text{Var}(I_{pcc})$ is <20% of the variance of all impedance samples, those 10% results are regarded as valid estimations. In the second technique, the authors propose a coefficient of determination (R^2) which represents the linear relationship between V_{pcc} and I_{pcc} . The stronger the linear relationship is, the closer the value of R^2 is to one so that the estimated impedance fits the actual impedance value better. To quantify the selecting criterion, a threshold of $R^2 = 0.9$ is used to select the usable impedance samples. Although the data selection techniques improve the estimating reliability of the linear regression method theoretically, it is found that no results can be provided by this method in most field cases due to the extensive existence of system variation.

2.2.3 ICA method: Up to the present day, fast ICA (FastICA) is the main ICA approach used for impedance measurement. In reference [16], the real-valued FastICA algorithm is adopted for mixed-signal decomposition based on the equation below

$$\begin{bmatrix} \Delta U_{pccx} \\ \Delta U_{pccy} \\ \Delta I_{pccx} \\ \Delta I_{pccy} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \cdot \begin{bmatrix} \Delta I_{cx} \\ \Delta I_{cy} \\ \Delta I_{sx} \\ \Delta I_{sy} \end{bmatrix} \quad (5)$$

As shown in (5), the Norton equivalent circuit equations are separated into the real and the imaginary parts, where symbol Δ means the fluctuating quantity. Subscripts x and y represent the real part and the imaginary part of the complex number, respectively. a_{ij} is the coefficient which is only related to the harmonic impedance. Equation (5) can be simplified as

$$X = A \cdot I \quad (6)$$

where X is the observed signal matrix which is the left part of the equal sign in (5). A is the coefficient matrix comprising of a_{ij} , and I is the source signal matrix denoting the rightest matrix in (5). Applying the real-valued FastICA algorithm to the observed matrix, the independent components S_m ($m=1,2,3,4$) can be obtained. The corresponding relation and scaling relation between S_m and the real source signals are unclear, so assume $S_m = c_m I_j$ (where $j=m$ and c_m is a non-zero real number), next the authors

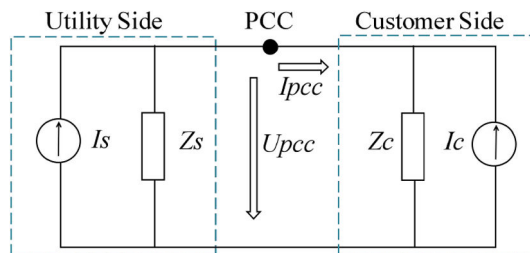


Fig. 1 Norton equivalent circuit diagram

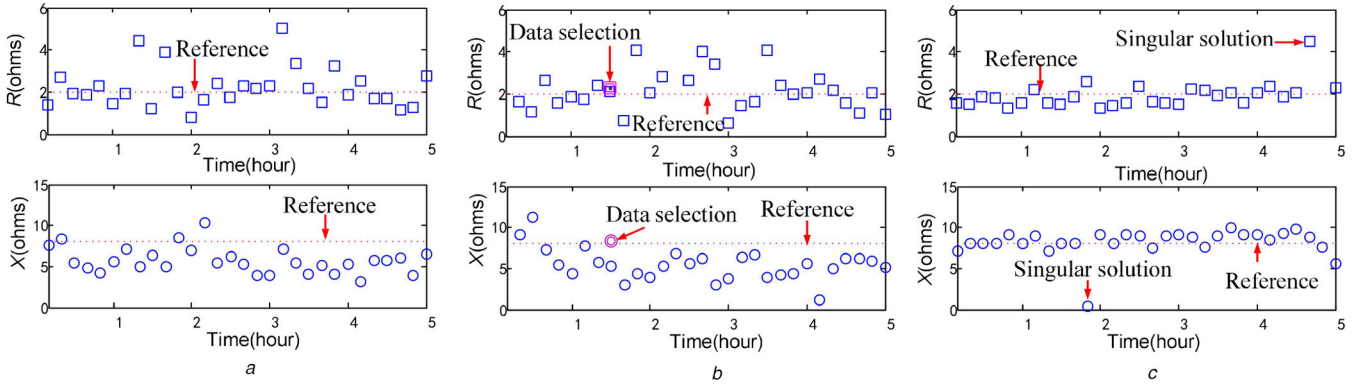


Fig. 2 Estimated results of the utility harmonic impedance of three published methods

(a) Estimated results of fluctuation method, (b) Estimated results of linear regression method, (c) Estimated results of ICA method

solve the mixing coefficients by means of the least square method which is expressed as

$$\mathbf{K}_m = (\mathbf{I}^T \mathbf{I})^{-1} \mathbf{I}^T \mathbf{X}_m \quad (7)$$

where $\mathbf{K}_m = [k_{1m}, k_{2m}, k_{3m}, k_{4m}]^T$ is the mixing coefficient matrix. From (5), the relation between \mathbf{K}_m and c_m can be determined as $a_{im} = K_{im} c_m$. By decomposing (1) into the real and the imaginary parts and comparing with the mixing coefficient matrix, the solution formulas of utility harmonic impedance are acquired as follows:

$$\begin{cases} Z_{\text{sreal}} = \frac{k_{1m}k_{3m} + k_{2m}k_{4m}}{k_{3m}^2 + k_{4m}^2} \\ Z_{\text{simag}} = \frac{k_{2m}k_{3m} - k_{1m}k_{4m}}{k_{3m}^2 + k_{4m}^2} \end{cases} \quad (8)$$

where Z_{sreal} represents the real part of the utility harmonic impedance, and Z_{simag} represents the imaginary part of the utility harmonic impedance.

In reference [17], another impedance calculation method based on complex-valued FastICA is presented. First the authors decompose the measured \mathbf{U}_{pcc} and \mathbf{I}_{pcc} into the fast-varying and slow-varying components by a linear filter. Then complex-valued FastICA algorithm is employed to the fast-varying components and the separating matrix and independent components are obtained. At last the harmonic impedance is estimated by solving an optimisation problem. No matter the real-valued or complex-valued FastICA method for impedance calculation, they both need long-term data to avoid falling into local optimum to guarantee estimation accuracy. Therefore, they are inapplicable for real-time estimation.

2.3 Remaining challenges faced by published methods

To verify and compare the performance of the three published methods, harmonic impedance estimations are conducted using a group of 5-hour field test data. This set of data was recorded at a 14.4 kV residential feeder of a 138 kV transformer substation in Alberta, Canada. During the measuring period the operation mode of the system keeps unchanged and the reference system harmonic impedance provided by the utility is $2 + 8j$. The data set is divided into 30 continuous segments (ten-minute data as one segment) and the reviewed methods are applied to these segments individually. The estimated results are shown in Fig. 2.

As shown in Figs. 2a and b, the estimated values of the utility harmonic impedance by fluctuation method and least square linear regression method vary largely around the reference value. This is an unsurprising circumstance because the two methods are inherently susceptible to the background harmonic variation which exists extensively in real power systems. The improved linear regression method based on data selection is expected to provide credible results. Nevertheless, due to the strict selection criteria, in this case, only one estimated value is provided and no more

effective results are available during a large portion of the time. That is because the system variation is not negligible and then few estimated results can meet the data selection requirement. This kind of situation of this method has been proved through extensive field data not only this case. Therefore, the data-selection-based linear regression method is still inapplicable in practice. The ICA method is a quite new method for harmonic impedance measurement which is deemed robust to the disturbance from the utility side. As seen from Fig. 2c, the estimated values are closer to the reference value and the varying trend is more stable compared to the other two methods. It is proved that the ICA method can reduce the adverse impact of the utility-side harmonic variation on impedance estimation to an acceptable degree and it can provide continuous estimated results with certain reliability. Therefore, the ICA method is a promising methodology to measure harmonic impedance for practical application.

However, some problems still exist in the classical ICA method. When the data size is short in real-time measurement, the ICA algorithm may fall into a local optimum or encounter saddle point which would bring about large error. As one can see in this field case, since impedance calculation is conducted every ten minutes and only 500 samples are involved in each computation, there are some singular solutions with large error in the estimated results. Therefore, how to remove these singular solutions of ICA method and improve its estimating accuracy with relatively small sample size in real-time estimation are primary problems that need to be solved in practical application.

2.4 Principle of ICA

2.4.1 ICA model: Assume there are m observed signals x_1, x_2, \dots, x_m which are linear combinations of another n random variables s_1, s_2, \dots, s_n . The relationship is expressed as follows

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \dots + a_{in}s_n \quad (9)$$

where a_{ij} is the coefficient. Rewrite (9) in the matrix form as below

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (10)$$

where t denotes the discrete time, $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$. To meet the requirements of ICA [19, 20], the source signals $\mathbf{s}(t)$ must be statistically independent and non-Gaussian distributions (at most one variable is Gaussian). $\mathbf{x}(t)$ is the matrix of the observed signals, $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$. $m \geq n$ is another necessary condition for ICA, and matrix \mathbf{A} is called the mixing matrix.

In short, what ICA solves is to restore the source signals under the circumstances that only the mixed signals are available but the source signals and mixing matrix are both unknown. In other words, it aims at seeking a separating matrix \mathbf{W} by which to achieve the optimal estimation of the real source signals $\mathbf{s}(t)$. The mathematical expression is shown as

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t) \quad (11)$$

where $\mathbf{y}(t)$ is the estimated source signals.

2.4.2 Robust independent analysis algorithm: Although FastICA is a very widely used ICA algorithm in the field of blind source separation, it is susceptible to sample size change and may get trapped in saddle areas or local extrema for short data records. In view of this, a new method for deflationary ICA named as RICA is proposed by Zarzoso and Comon in 2010 [21, 22]. In RICA algorithm, the kurtosis is selected as the contrast function and it is optimised by exact line search based on an optimal step-size (OS) technique. The OS along the search direction is found by fourth-degree polynomial rooting which can be performed algebraically at low cost at each iteration. Differing from the classical ICA algorithm, observed data prewhitening is needless for RICA, so the centring of observed variables is the only step of data preprocessing, and then the zero-mean variables are obtained by subtracting the mean value from each sample as shown below

$$x'(t_i) = x(t_i) - \frac{1}{n} \sum_{i=1}^n x(t_i) \quad (12)$$

Seek an extracting vector \mathbf{w} in the deflation approach to ICA to get the separated signal \mathbf{y} , its kurtosis $k(\mathbf{w})$ is defined as the normalised four-order marginal cumulant

$$k(\mathbf{w}) = \frac{E\{|\mathbf{y}|^4\} - 2E^2\{|\mathbf{y}|^2\} - |E\{\mathbf{y}^2\}|^2}{E^2\{|\mathbf{y}|^2\}} \quad (13)$$

where $E\{\cdot\}$ is the mathematical expectation. The search for OS is performed by the absolute kurtosis contrast function. The function is written as

$$\mu_{\text{opt}} = \arg_{\mu} \max |k(\mathbf{w} + \mu \mathbf{g})| \quad (14)$$

where $\arg_{\mu} \max$ means that μ_{opt} is the maximum output value of $|k(\mathbf{w} + \mu \mathbf{g})|$. \mathbf{g} is the search direction and it is typically the gradient expressed as follows

$$\mathbf{g} = \frac{4}{E^2\{|\mathbf{y}|^2\}} \left\{ E\left\{ \left| \mathbf{y} \right|^2 \mathbf{y}^* \mathbf{x} \right\} - E\{\mathbf{y} \mathbf{x}\} E\{\mathbf{y}^{*2}\} \right. \\ \left. - \frac{(E\{|\mathbf{y}|^4\} - |E\{\mathbf{y}^2\}|^2) E\{\mathbf{y}^* \mathbf{x}\}}{E\{|\mathbf{y}|^2\}} \right\} \quad (15)$$

where $*$ denotes adjoint matrix. Since exact line search is computationally complicated and has certain limitations, in addition for criteria like kurtosis can be expressed as the polynomial of μ , thus the globally OS can be established by finding the roots of a low-degree polynomial. Based on this idea, the RICA algorithm achieves optimisation of globally OS at each iteration by the following steps

(i) Calculate the OS polynomial coefficients. The OS polynomial of the kurtosis contrast function is given by

$$p(\mu) = \sum_{k=0}^4 a_k \mu^k \quad (16)$$

The coefficients $\{a_k\}_{k=0}^4$ are obtained by observed signal block and the current values of \mathbf{w} and \mathbf{g} at each iteration. Their detailed expressions can be found in reference [20].

(ii) Extract OS polynomial roots $\{\mu_k\}_{k=1}^4$

(iii) Select the root making the absolute objective function $|k(\mathbf{w} + \mu \mathbf{g})|$ maximise along the search direction

(iv) Update \mathbf{w}^* by the new-obtained μ_{opt} as below

$$\mathbf{w}^* = \mathbf{w} + \mu_{\text{opt}} \mathbf{g} \quad (17)$$

(v) Normalise \mathbf{w}^* by

$$\mathbf{w}^* \leftarrow \frac{\mathbf{w}^*}{\|\mathbf{w}^*\|} \quad (18)$$

Finally, the termination condition of iteration can be set as below

$$|1 - |\mathbf{w}^H \mathbf{w}^*|| < \varepsilon \quad (19)$$

where $(\cdot)^H$ represents the conjugate transpose operator, and ε is a significant small constant for convergence checking.

In conclusion, the procedure of RICA can be simply summarised as shown in Fig. 3.

RICA solves some problems existing in the classical ICA algorithm. To compare the two algorithms, the advantages and mechanism of RICA can be summarised as below

- Although pre-whitening in ICA method can decorrelate the observed signals and reduce the computational complexity, it imposes a bound on separation performance and brings about an estimation bias due to residual source correlations in short data records. The absence of data prewhitening in RICA makes it more tolerant to residual source correlations than the whitening-based ICA algorithm and the performance limitation can be avoided.
- For ICA, the step size is fixed to a constant value which leads to cubic convergence of the algorithm for infinite sample size. However, when the data length is short, convergence may slow down and even get trapped in saddle areas and local extrema. The methodology behind RICA is exact line search based on OS technique. Contrary to the classical line search techniques which can only perform local optimisation in the search direction, the OS technique computes the step size to globally optimise the kurtosis along the search direction at each extracting vector update, thus it is robust to the presence of saddle points and spurious local extrema and can improve the separation performance in short data circumstances.
- Benefiting from the generality of the kurtosis contrast function, RICA algorithm can process real-valued and complex-valued mixtures without any modification. Furthermore, it shows higher convergence speed and few iterations in the separation of signals of both types.

3 Improved harmonic impedance measurement

The harmonic impedance calculation method based on RICA is explained in this section first, and then a singular solution elimination technique based on bootstrap is proposed to further improve the estimation accuracy.

3.1 Utility harmonic impedance calculation based on RICA

In view of the superiority for short data sizes of RICA mentioned above, RICA algorithm is adopted to perform utility harmonic impedance estimation in this section. First, (1) can be rewritten in the matrix form as

$$\mathbf{X} = \mathbf{Z} \mathbf{I} \quad (20)$$

where

$$\begin{cases} \mathbf{X} = \begin{bmatrix} \mathbf{U}_{\text{pcc}} \\ \mathbf{I}_{\text{pcc}} \end{bmatrix} \\ \mathbf{Z} = \begin{bmatrix} \frac{Z_s Z_c}{Z_s + Z_c} & \frac{Z_s Z_c}{Z_s + Z_c} \\ \frac{Z_s}{Z_s + Z_c} & -\frac{Z_c}{Z_s + Z_c} \end{bmatrix} \\ \mathbf{I} = \begin{bmatrix} \mathbf{I}_s \\ \mathbf{I}_c \end{bmatrix} \end{cases} \quad (21)$$

As mentioned in the previous section, two conditions must be satisfied for ICA model, the dimension condition and the signal

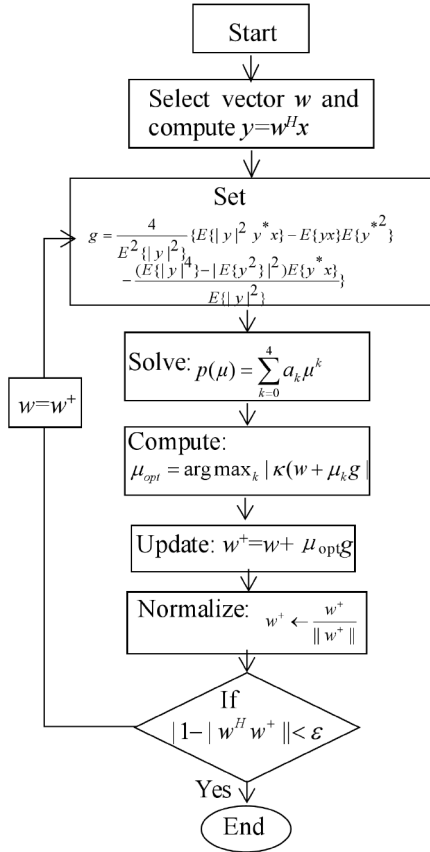


Fig. 3 Flow chart of RICA algorithm

property condition. Obviously, the number of observed signals and source signals are equal to two in the formulas above, so the first condition is satisfied. The second condition which requires the source signals to be statistically independent and non-Gaussian need to be further verified. In reference [15], the authors present that the harmonic voltage and current measured at the PCC can be divided into the fast-varying components and the slow-varying components by applying a linear filter. This operation is also equally applicable to the source signals, so based on this principle, (20) can naturally be decomposed into two parts as below

$$X_{\text{fast}} = Z I_{\text{fast}} \quad (22)$$

$$X_{\text{slow}} = Z I_{\text{slow}} \quad (23)$$

where X_{fast} and I_{fast} are the fast-varying components of matrices X and I , respectively. X_{slow} and I_{slow} are the slow-varying components of matrices X and I , respectively.

It is pointed out in reference [20] that the fast-varying components of harmonic load currents are generally independent and non-Gaussian. Since I_s and I_c are the combinations of the scattered harmonic load currents in the utility and customer side, it is rational to deduce that the fast-varying components of I_s and I_c are statistically independent and have non-Gaussian distributions. Therefore, the RICA algorithm can be applied to the fast-varying components of U_{pcc} and I_{pcc} , namely the matrix X_{fast} . By applying RICA algorithm, the separation matrix $W = [w_1, w_2]^T$ and separated independent components $I_{\text{fast-es}} = [I_{\text{fast1}}, I_{\text{fast2}}]^T$ are obtained, their relation can be expressed as

$$I_{\text{fast-es}} = W X_{\text{fast}} \quad (24)$$

There are two indeterminacies in the separated independent components, which are the ordering indeterminacy and the scaling indeterminacy. Assuming that $I_{\text{sfast-es}}$ and $I_{\text{cfast-es}}$ are the recovered signals belonging to I_s and I_c , respectively, then to solve the ordering indeterminacy is to find out the corresponding

relationship between $I_{\text{sfast-es}}$, $I_{\text{cfast-es}}$ and I_{fast1} , I_{fast2} . First the separating matrix W is utilised to recover the source signals including both the fast- and slow-varying components by the observed matrix X as follows

$$I_{\text{es}} = \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = W X = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} X \quad (25)$$

where $I_{\text{es}} = [I_1, I_2]^T$ is the recovered source signal matrix. In real power systems, the utility impedance is usually determined by the short-circuit capacity, so the magnitude of the utility impedance is normally much smaller than that of the customer side. In addition the utility- and customer-side impedances are inductive in most cases, which means the harmonic impedances of the two sides also conform to this situation, namely $|Z_s| \ll |Z_c|$. According to this assumption and (1), it can be easily deduced that $I_{\text{pcc}} \approx -I_c$. Evidently I_{pcc} is contributed mostly by I_c . In other words, there must be a strong linear correlation between I_{pcc} and I_c . On the contrary, the correlation between I_s and I_{pcc} is relatively weak. Therefore, the ordering indeterminacy can be solved by comparing the linear correlation between I_1 , I_2 and I_{pcc} . The correlative degree can be expressed by correlation coefficients as follows

$$\begin{aligned} \rho_{1\text{pcc}} &= \frac{\text{Cov}(I_1, I_{\text{pcc}})}{\sqrt{D(I_1)}\sqrt{D(I_{\text{pcc}})}} \\ \rho_{2\text{pcc}} &= \frac{\text{Cov}(I_2, I_{\text{pcc}})}{\sqrt{D(I_2)}\sqrt{D(I_{\text{pcc}})}} \end{aligned} \quad (26)$$

where $\rho_{1\text{pcc}}$ denotes the correlation coefficient between I_1 and I_{pcc} . $\rho_{2\text{pcc}}$ denotes the correlation coefficient between I_2 and I_{pcc} . $\text{Cov}(\cdot)$ represents the covariance of two variables, and $D(\cdot)$ represents the variance.

Next, rewrite (24) in a specific order as follows

$$\begin{bmatrix} I_{\text{sfast-es}} \\ I_{\text{cfast-es}} \end{bmatrix} = U X_{\text{fast}} \quad (27)$$

according to the calculated results in (26), compare the two variables $\rho_{1\text{pcc}}$ and $\rho_{2\text{pcc}}$, when $|\rho_{1\text{pcc}}| < |\rho_{2\text{pcc}}|$

$$I_{\text{sfast-es}} = I_{\text{fast1}}, I_{\text{cfast-es}} = I_{\text{fast2}}, U = W = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad (28)$$

when $|\rho_{1\text{pcc}}| > |\rho_{2\text{pcc}}|$

$$I_{\text{sfast-es}} = I_{\text{fast2}}, I_{\text{cfast-es}} = I_{\text{fast1}}, U = \begin{bmatrix} w_2 \\ w_1 \end{bmatrix} \quad (29)$$

Since there is scaling indeterminacy between the separated independent components and the real source signals, so assume that

$$\begin{aligned} I_{\text{sfast-es}} &= c_1 I_{\text{sfast}} \\ I_{\text{cfast-es}} &= c_2 I_{\text{cfast}} \end{aligned} \quad (30)$$

where c_1 and c_2 are non-zero real numbers representing the proportional relation between the separated signals and real source signals. Then (30) is plugged into (22) and written in full form as

$$X_{\text{fast}} = \begin{bmatrix} \frac{1}{c_1} \frac{Z_s Z_c}{Z_s + Z_c} & \frac{1}{c_2} \frac{Z_s Z_c}{Z_s + Z_c} \\ \frac{1}{c_1} \frac{Z_s}{Z_s + Z_c} & -\frac{1}{c_2} \frac{Z_c}{Z_s + Z_c} \end{bmatrix} \begin{bmatrix} I_{\text{sfast-es}} \\ I_{\text{cfast-es}} \end{bmatrix} \quad (31)$$

Next, by multiplying both sides of (27) by the inverse matrix of U , (32) can be obtained

$$\mathbf{X}_{\text{fast}} = \mathbf{U}^{-1} \cdot \begin{bmatrix} \mathbf{I}_{\text{sfast-es}} \\ \mathbf{I}_{\text{cfast-es}} \end{bmatrix} \quad (32)$$

At last, comparing (31) with (32), the utility harmonic impedance Z_s can be easily calculated as follows

$$Z_s = -\frac{\mathbf{U}^{-1}(1, 2)}{\mathbf{U}^{-1}(2, 2)} \quad (33)$$

where numbers in the round brackets indicate the element of the corresponding row and column in matrix \mathbf{U}^{-1} .

3.2 Bootstrap check for singular solution elimination

In regard to this sort of ICA methods, in some cases the separated results may have large deviation due to the presence of saddle points and spurious local extremum in iteration, thus some singular solutions of harmonic impedance may appear which makes the estimated results unreliable in these circumstances. To exhibit this situation, the proposed RICA method is applied to the set of data in Section 2.3. The estimated results are shown in Fig. 4.

From the results in Fig. 4, one can evidently see that the estimated results by RICA are close to the reference value. It is proved that as an improved ICA method, it can suppress the bad influence of background harmonics as the same as the classical ICA method. Even though satisfactory results are obtained by RICA method, singular values still exist as shown during data segment 1. In order to reduce the negative effect caused by these singular solutions, a validity verification technique of the estimated results based on bootstrap is proposed in this section. This technique is named 'bootstrap check' in this paper.

The bootstrap methodology [23] is a classical statistical approach to perform statistical interval estimation for finite sample sizes. It can raise the estimation performance of the concerned statistic without adding new observed samples only by re-sampling the original sample space with replacement. The group of re-sampling samples is called a bootstrap sample space. In this manner, 500 bootstrap sample spaces (500 sample points are involved in each space) are extracted from data segment 1 and data

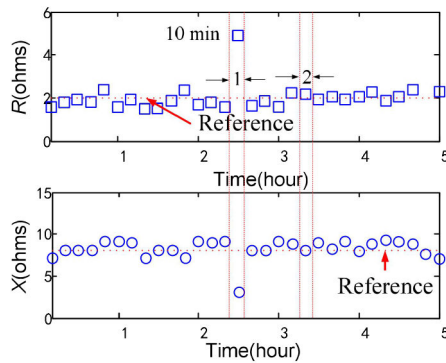


Fig. 4 Estimated results of Z_s by RICA method

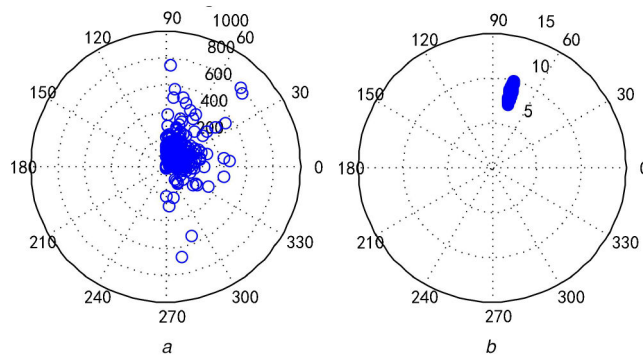


Fig. 5 Polar diagrams of the distributions of Z_s calculated by bootstrap samples

(a) Calculated Z_s by bootstrap samples of data segment 1, (b) Calculated Z_s by bootstrap samples of data segment 2

segment 2 in Fig. 4, respectively. Then Z_s is calculated by the bootstrap samples. The estimated values by the bootstrap samples of segment 1 and segment 2 are demonstrated in Fig. 5.

As shown in Fig. 5a, the distributions of calculated Z_s by bootstrap samples of data segment 1 in the polar diagram are dispersed widely. Conversely in Fig. 5b the counterpart values of data segment 2 are lumped within a small range. This is not an individual phenomenon. By vast filed data studies, it is found that when the sort of ICA methods encounters a singular solution of Z_s by a group of data, the counterpart results by its bootstrap samples are scattered. However, if the ICA methods reach a correct separation, the estimated results by bootstrap samples would gather within certain range. Based on this finding, the 'bootstrap check' technique is presented and its specific procedure can be summarised as the following steps:

- Re-sample n points from the original n samples which have been used to calculate Z_s with replacement, then one bootstrap sample space is acquired.
- Calculate Z_s based on the bootstrap sample space by the proposed RICA method.
- Repeat steps (a) and (b) for n times and then n Z_s samples are obtained.
- Analyse the dispersion of the n Z_s samples. To quantify the dispersion degree, the dimensionless parameter coefficient of variation (CV) is adopted in this paper which can be expressed as

$$CV = \frac{\sigma}{\mu} \quad (34)$$

where σ is the standard deviation of the n Z_s samples. μ is the mean value. If the calculated CV is >0.15 , it is considered that the originally estimated result of Z_s is a singular value and should be discarded.

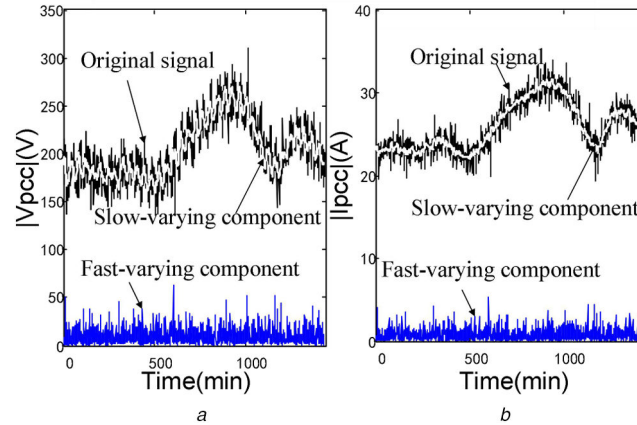
(e) If CV is <0.15 , the result is regarded as a valid estimation. Then apply the quantile method to the impedance samples to compute the confidence interval (CI) of Z_s . First, the impedance samples need to be sorted from smallest to largest and a vector is obtained ($Z_{s1}, Z_{s2}, \dots, Z_{sn}$). Then the CI can be calculated as below

$$\begin{aligned} k_1 &= [n \times \alpha/2] \\ k_2 &= [n \times (1 - \alpha/2)] \\ CI &= (Z_{sk_1}, Z_{sk_2}) \end{aligned} \quad (35)$$

where α is the confidence level. Generally, 95% confidence level is usually selected in most cases. Differing from CV which is used to verify the validity of the estimates, The index CI indicates a statistical range with a specific probability that the estimated value of Z_s lies within the range. Therefore, CI can be used as a further crosscheck of the estimated result.

Table 1 Parameters of simulation model

Parameters		Case 1	Case 2	Case 3
slow-varying component			typical load curves in [17]	
fast-varying component	I_s	Lap(0,0.2)	Lap(0,0.8)	Lap(0,0.5)
	I_c	Lap(0,0.8)	Lap(0,0.8)	Lap(0,0.2)
impedance, Ω	Z_s		$1 + j10$	
	Z_c		$10 + j80$ ($\pm 5\%$ sine variation)	

**Fig. 6** Waveform sampled at the PCC in case 1

(a) Magnitude of the harmonic voltage, (b) Magnitude of the harmonic current

Table 2 Estimated results of Z_s by two methods

Estimated results			ICA method	RICA method
case 1	Z_s		$0.9556 + j 9.6167$	$0.9709 + j10.2619$
	error(%)	R	4.44	2.91
		X	3.833	2.619
case 2	Z_s		$1.0721 + j10.5448$	$0.9647 + j9.6971$
	error(%)	R	7.21	3.53
		X	5.448	3.029
case 3	Z_s		$0.9227 + j10.556$	$1.0338 + j10.3516$
	error(%)	R	7.73	3.38
		X	5.56	3.516

4 Simulation case study

In this section, computer simulation is conducted based on the Norton equivalent circuit in Fig. 1 and the equivalent formulas in (1) by MATLAB R2014a software. The simulation program is written using MATLAB code to construct the circuit relations of the elements. To model the slow-varying components of the harmonic current sources of the utility and customer side, the typical load curves in reference [17] are adopted. The fast-varying components of the harmonic current sources are modelled with random variables obeying Laplace distribution which is denoted by $\text{Lap}(a,b)$, where a represents the expectation and b represents the variance. The Laplace distributed fast-varying components of both sides are statistically independent, thus they meet the requirements of ICA methods. The specific simulation parameters are shown in Table 1.

As one can see in Table 1, three cases are simulated to verify the effectiveness and validity of the proposed method in different background harmonic fluctuation circumstances. In case 1, the variance of the fast-varying component of I_c is much larger than that of I_s , which means the background harmonic fluctuation is relatively small and the variation of harmonic voltage at the PCC is mainly caused by the customer side. In case 2, the value of b of I_c is equal to that of I_s , denoting that the variation of harmonic voltage at the PCC is determined by both sides. In case 3, contrary to case 1, the value of b of I_c is much smaller than that of I_s , which means the background harmonic fluctuation is much larger than the

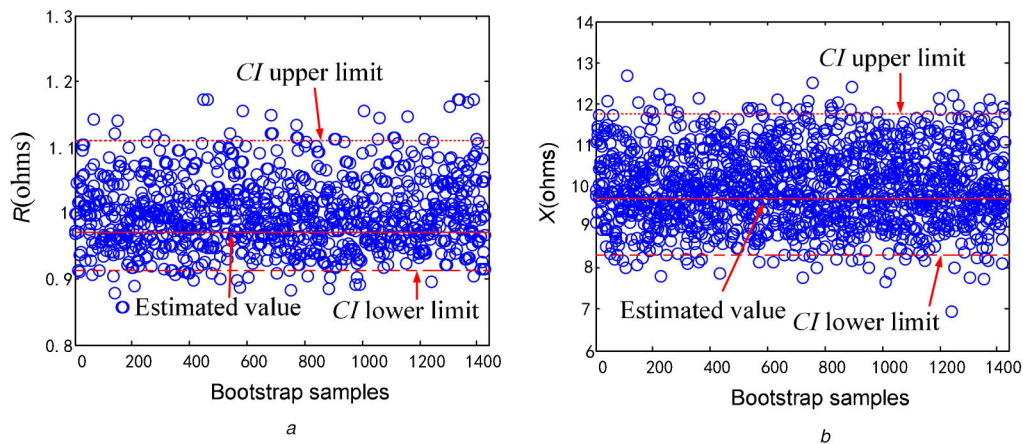
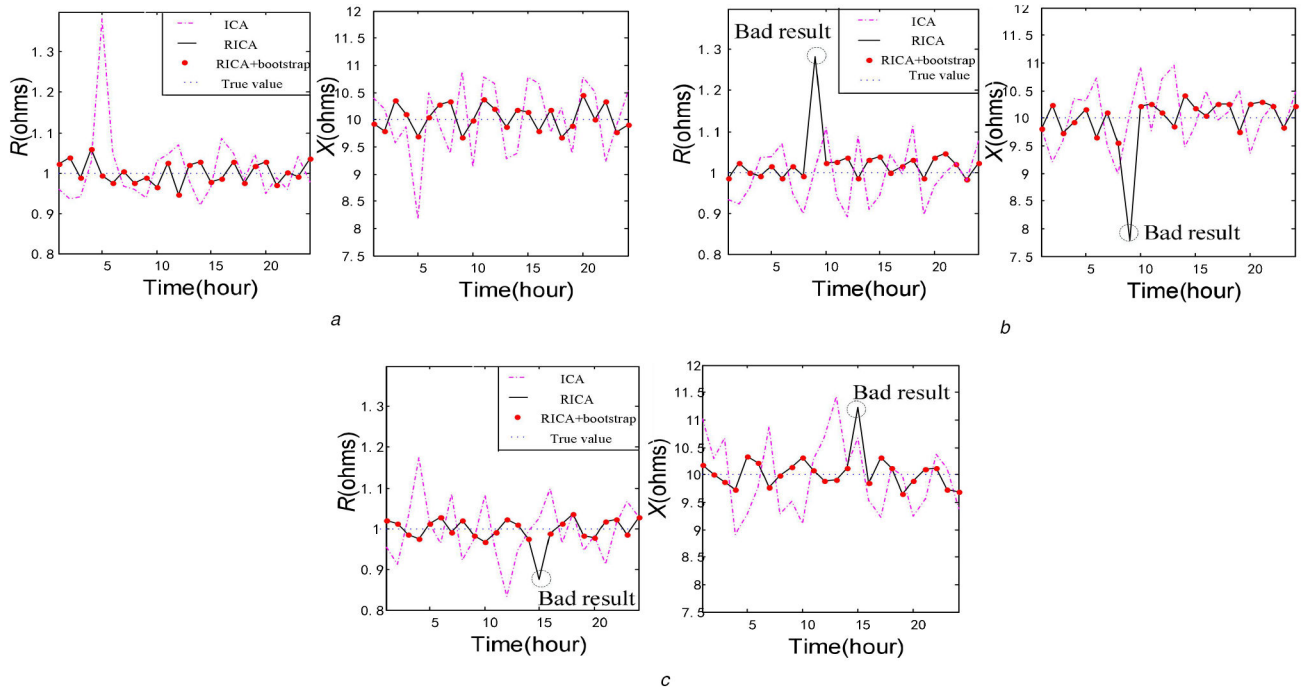
fluctuation from the customer side and the voltage variation at the PCC is dominated by the utility side. The value of Z_c is set much larger than that of Z_s , which fits most cases in real power systems, and $\pm 5\%$ sine fluctuation is added to both the real and imaginary parts of it. The length of the sampled data at the PCC is 1440 points for an entire day (one point per minute). The sampled harmonic voltage and harmonic current at the PCC are decomposed into the fast-varying components and the slow-varying components by using a 10-point moving average filter. Fig. 6 demonstrates the magnitude of harmonic voltage and harmonic current at the PCC and their decomposed components in case 1.

The method based on the classical ICA and the method proposed in this paper is applied to the sampled data at the PCC, respectively to perform utility harmonic impedance calculation. At the same time, the bootstrap samples are extracted from the fast-varying components of the measured data at the PCC to calculate CV and CI. The estimated results of Z_s and the calculated CV and CI by bootstrap samples for these cases are listed in Tables 2 and 3, respectively.

According to the results exhibited in Table 2, for three cases the estimating accuracy of the proposed RICA method is higher than the traditional ICA method. Besides the accuracy variation in three cases representing different fluctuation level of background harmonics is small which implies that the proposed RICA method can reduce the disturbances from the utility side effectively. In addition, from the results in Table 3, the three calculated CV are all < 0.15 , which means the results by RICA method are reliable and

Table 3 Calculated CV and CI by bootstrap samples

	CV	CI
case 1	0.0274	R: 0.9124–1.1108 X: 9.737–10.7661
case 2	0.0915	R: 0.8956–1.172 X: 9.2451–11.7315
case 3	0.0769	R: 0.8365–1.177 X: 8.6376–11.5024

**Fig. 7** Estimated results of Z_s by bootstrap samples(a) Estimated real part of Z_s , (b) Estimated imaginary part of Z_s **Fig. 8** Estimated results of Z_s per hour(a) Estimated results of Z_s in case 1, (b) Estimated results of Z_s in case 2, (c) Estimated results of Z_s in case 3

valid. At the same time, the CIs are calculated to crosscheck the estimated values of Z_s . Fig. 7 demonstrates the estimated results of the 1440 bootstrap sample spaces in case 1.

As one can see in Fig. 7 the estimated values by 1440 bootstrap sample spaces gather within a certain range, no matter the real parts or the imaginary parts. The calculated CI of bootstrap samples and the estimated results by original samples are marked in Fig. 7 as well. It is shown in Fig. 7 that the estimated value of Z_s by RICA method locates between the CI upper limit and the CI lower limit, thus one can conclude that the estimated value of Z_s is reliable.

To evaluate and compare the performance of the proposed RICA method and ICA method for short data sizes, and to verify the effectiveness of the bootstrap check technique, impedance calculation is performed every single hour with 300 sample points

(four points are added per minute by linear interpolation). Three methods are applied which are the ICA method, RICA method and method combining RICA with bootstrap. The estimated results of them are demonstrated in Fig. 8.

From Fig. 8, it can be observed that the estimates obtained by ICA method fluctuate widely while the ones obtained by RICA method are more stable during the entire day. It reflects that the proposed RICA method has better performance than ICA method for short data sizes. It is proved that RICA can acquire more satisfactory results in small sample spaces compared with ICA. In addition, for case 1, simply using RICA can achieve acceptable accuracy, thus the difference between 'RICA' and 'RICA+bootstrap' is small. However, there is a bad estimate in case 2 and case 3 as shown in Figs. 8b and 8c. The proposed bootstrap check successfully removes the bad points. Therefore, the effectiveness of the proposed bootstrap check technique to improve estimation



Fig. 9 Filed test picture
(a) Field connection diagram, (b) The measuring equipment

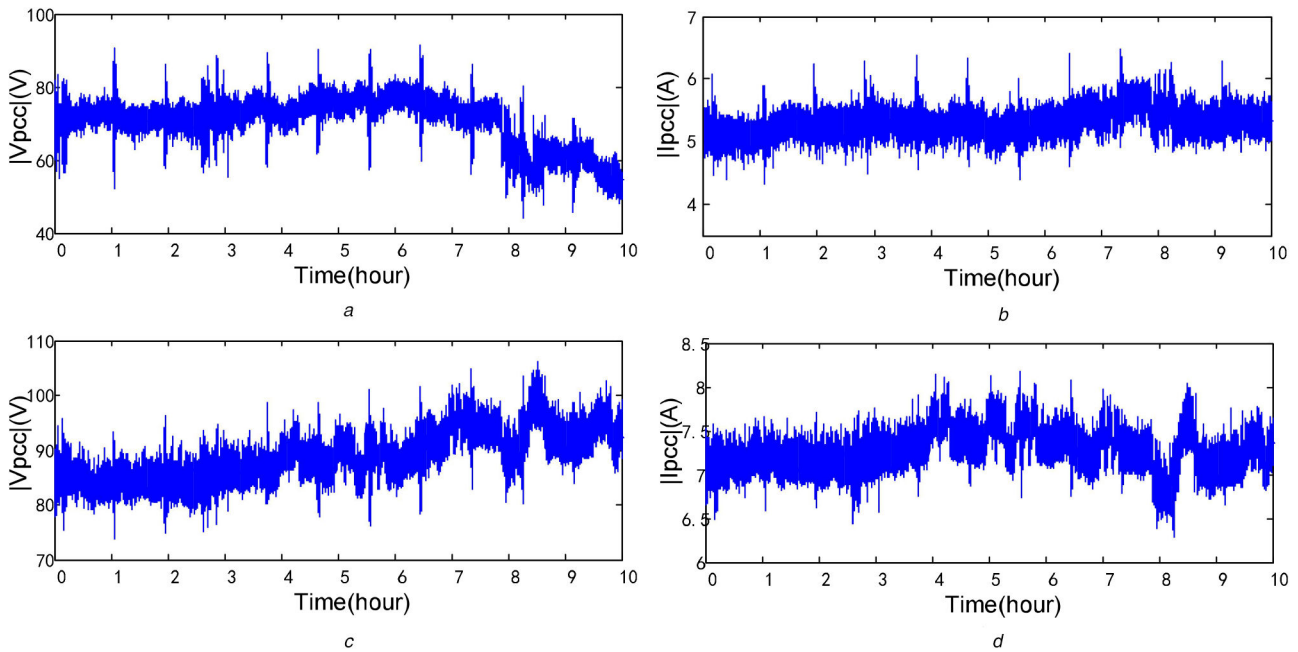


Fig. 10 Amplitude of the harmonic voltage and current at the PCC
(a) Amplitude of the fifth harmonic voltage, (b) Amplitude of the fifth harmonic current, (c) Amplitude of the seventh harmonic voltage, (d) Amplitude of the seventh harmonic current

accuracy is also verified. In conclusion, by our computer simulation, the proposed method combining RICA with bootstrap check can achieve the best harmonic impedance measurement.

5 Filed test

To verify the validity and applicability of the proposed method, field data is used to perform harmonic impedance calculation. This set of data is measured at a 14.4 kV residential feeder in a 138 kV transformer substation in Canada and recorded by a high-precision PQ analyser with the sampling frequency of 5 kHz for ten consecutive hours. The filed connection diagram and the picture of our measuring equipment are given by Fig. 9. It is programmed based on the FFT algorithm to compute the fifth and seventh harmonic components in accordance with IEC 61000-4-7. By FFT 90,000 points of harmonic samples are obtained as shown in Fig. 10, and then the three methods in the simulation are applied to them with 500 points as one segment (180 segments in total). The estimated results of the three methods are demonstrated in Fig. 11.

During the measuring period of this set of data, the operation mode of power system keeps unchanged, so it is reasonable to assume that the variation of utility harmonic impedance is small and the estimated values of Z_s would be close to the reference values supplied by the utility. The reference values of the fifth and seventh harmonic impedances are $1 + j10 \Omega$ and $1 + j11 \Omega$, respectively. According to the estimates of the fifth and seventh harmonic impedance given in Fig. 11, it is observable that the estimated values by classical ICA method fluctuate largely around the reference value due to insufficiency of samples. By contrast,

the fluctuation level of the estimates by RICA is much smaller than that of ICA. However, the results by RICA are still unsatisfactory because of the presence of singular values. From Fig. 10, it is shown that the 'RICA+bootstrap' method can provide reliable estimates closest to the reference value among three methods. Therefore, the effectiveness and applicability of the proposed method in real power systems is verified.

6 Conclusion

In this paper, an improved method for utility harmonic impedance measurement based on RICA and bootstrap check is proposed. Main conclusions of this paper are summarised as follows:

- (i) Although the classical ICA method overcomes the shortcoming of being susceptible to background harmonic fluctuations of the traditional methods, it requires long data records to ensure estimating accuracy and may encounter singular solutions because of saddle point and local extremum in the iteration. Therefore, the exact-line-research-based RICA algorithm which has better performance for short data records is utilised to perform impedance calculation.
- (ii) It is found that when ICA methods reach a singular solution for a group of data, the estimated results by its bootstrap samples are scattered widely. Conversely, the estimates by bootstrap samples gather within a small range if ICA methods achieve a valid separation. Based on this finding, a bootstrap check technique through judging the dispersion degree (quantified by CV) is

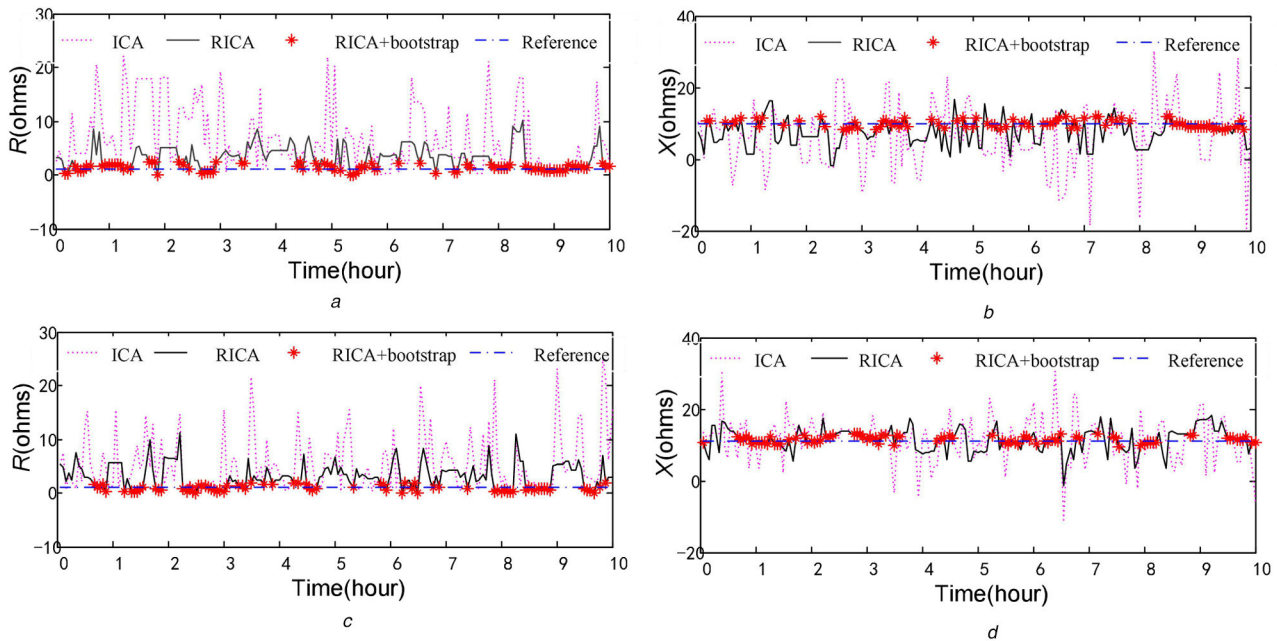


Fig. 11 Estimated results of the utility harmonic impedance

(a) Estimated real part of the fifth harmonic impedance, (b) Estimated imaginary part of the fifth harmonic impedance, (c) Estimated real part of the seventh harmonic impedance, (d) Estimated imaginary part of the seventh harmonic impedance

proposed to eliminate the adverse influence of the singular solutions in this sort of ICA methods.

(iii) Computer simulation and field test reveal that the proposed RICA method can acquire higher accuracy than ICA method, especially in the real-time estimation. Nevertheless, singular values still exist in its estimated results. The 'RICA + bootstrap' method shows the best performance among the three methods. Therefore, a combination of RICA with the bootstrap check is necessary to achieve satisfactory measurement in practical use.

The introduction of the bootstrap check technique solves the problem of singular solutions, but it increases the computing time at the same time. Since the time consumption of bootstrap check is related to the data segment length, how to find the optimal data segment size combining high reliability and desirable speed of computation is a task, which is worth of more research in the next step.

7 Acknowledgments

The authors acknowledge financial support from the Science and Technology Project of Electric Power Research Institute of State Grid Zhejiang Electric Power Company Ltd (5211DS180033).

8 References

- [1] Liu, Q., Li, Y., Hu, S., *et al.*: 'Power quality improvement using controllable inductive power filtering method for industrial DC supply system', *Control Eng. Pract.*, 2019, **83**, pp. 1–10
- [2] Tang, K., Chen, S.: 'Harmonic emission level assessment based on parameter identification analysis', *IET Gener. Transm. Distrib.*, 2019, **13**, (7), pp. 976–983
- [3] Liu, Q., Li, Y., Hu, S., *et al.*: 'A transformer integrated filtering system for power quality improvement of industrial DC supply system', *IEEE Trans. Ind. Electron.*, 2019, pp. 1–1
- [4] Srinivasan, K.: 'On separating customer and supply side harmonic contributions', *IEEE Trans. Power Deliv.*, 1996, **11**, (2), pp. 1003–1012
- [5] Xu, W., Liu, Y.: 'A method for determining customer and utility harmonic contribution at the point of common coupling', *IEEE Trans. Power Deliv.*, 2000, **15**, (2), pp. 804–811
- [6] Zang, T., He, Z., Fu, L., *et al.*: 'Adaptive method for harmonic contribution assessment based on hierarchical K-means clustering and Bayesian partial least squares regression', *IET Gener. Transm. Distrib.*, 2016, **10**, (13), pp. 3220–3227
- [7] Thunberg, E., Soder, L.: 'A norton approach to distribution network modeling for harmonic studies', *IEEE Trans. Power Deliv.*, 1999, **14**, (1), pp. 272–277
- [8] Monteiro, H.L.M., Duque, C.A., Silva, L.R.M., *et al.*: 'Harmonic impedance measurement based on short time current injections', *Electr. Power Syst. Res.*, 2017, **148**, pp. 108–116
- [9] Sumner, M., Palethorpe, B., Thomas, D.W.P., *et al.*: 'A technique for power supply harmonic impedance estimation using a controlled voltage disturbance', *IEEE Trans. Power Electron.*, 2002, **17**, (2), pp. 207–215
- [10] Huang, X., Nie, P., Gong, H.: 'A new assessment method of customer harmonic emission level'. Proc. Asia-Pacific Power and Energy Engineering Conf., Chengdu, People's Republic of China, March 2010, pp. 28–31
- [11] Xiao, X., Yang, H.: 'The method of estimating customer harmonic emission level based on bilinear regression'. Proc. 2nd Int. Conf. on Electric Utility Deregulation, Restructuring and Power Technologies, Hongkong, People's Republic of China, April 2004, pp. 05–08
- [12] Xu, Y., Huang, S., Liu, Y.: 'Partial least-squares regression based harmonic emission level assessing at the point of common coupling'. Proc. Int. Conf. on Power System Technology, Chongqing, People's Republic of China, October 2006, pp. 01–06
- [13] Mazin, H.E., Xu, W., Huang, B.: 'Determining the harmonic impacts of multiple harmonic-producing loads', *IEEE Trans. Power Deliv.*, 2011, **26**, (2), pp. 1187–1195
- [14] Hui, J., Freitas, W., Vieira, J.C.M., *et al.*: 'Utility harmonic impedance measurement based on data selection', *IEEE Trans. Power Deliv.*, 2012, **27**, (4), pp. 2193–2203
- [15] Karimzadeh, F., Esmaili, S., Hosseini, S.H.: 'A novel method for noninvasive estimation of utility harmonic impedance based on complex independent component analysis', *IEEE Trans. Power Deliv.*, 2016, **30**, (4), pp. 1843–1852
- [16] Zhao, X., Yang, H.: 'A new method to calculate the utility harmonic impedance based on FastICA', *IEEE Trans. Power Deliv.*, 2016, **31**, (1), pp. 381–388
- [17] Karimzadeh, F., Esmaili, S., Hosseini, S.H.: 'Method for determining utility and consumer harmonic contributions based on complex independent component analysis', *IET Gener. Transm. Distrib.*, 2016, **10**, (2), pp. 526–534
- [18] Liao, H.M., Niebur, D.: 'Load profile estimation in electric transmission networks using independent component analysis', *IEEE Trans. Power Syst.*, 2003, **18**, (2), pp. 707–715
- [19] Ferreira, D.D., Seixas, J.M.D., Cerqueira, A.S.: 'A method based on independent component analysis for single and multiple power quality disturbance classification', *Electr. Power Syst. Res.*, 2015, **119**, pp. 425–431
- [20] Gursoy, E., Niebur, D.: 'Harmonic load identification using complex independent component analysis', *IEEE Trans. Power Deliv.*, 2009, **24**, (1), pp. 285–292
- [21] Zarzoso, V., Comon, P.: 'Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size', *IEEE Trans. Neural Netw.*, 2010, **21**, (2), pp. 248–261
- [22] Zarzoso, V., Comon, P.: 'Comparative speed analysis of FastICA'. Proc. 7th Int. Conf. on Independent Component Analysis and Signal Separation, London, UK, September 2007, pp. 293
- [23] Efron, B.: 'Bootstrap methods: another look at the jackknife', *Ann. Stat.*, 1979, **7**, (1), pp. 1–26