Harmonic Separation Based on Independent Component Analysis Method

Yongle Ai
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo, China
Email: aiyongle@hpu.edu.cn

Haiyang Zhang
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo, China
Email: haiyang5256@126.com

Abstract—Harmonic separation plays an important role in power quality research due to the harmful impact to power system and electric apparatus. The accurate separation of harmonics is also the demand for filtering. This paper models the problem of harmonic separation as a blind source separation (BSS) task and the paper presents an algorithm called independent component analysis, which is widely used in the field of BSS. Meanwhile, the least square method is also used to solve the problem. During the harmonic separation, ICA is used to determine harmonic frequencies firstly. Then the least square method is used to determine amplitudes and phases of harmonic components. By MATLAB simulation, the result validates that harmonics can be separated effectively with a high precision by using the method proposed under the condition with low real-time demand.

Index Terms—harmonic separation, blind source separation, independent component analysis, least square method, Matlab simulation

I. INTRODUCTION

With the development of power electronic technology, problems relevant to harmonics are becoming more and more serious due to growing applications of non-linear load devices such as inverters, converters, switching power supply. The main effects of harmonics in power system are heating, overloading, accelerating the aging of equipments and increasing power losses. Therefore, harmonic separation becomes one of the key issues to improve the system’s power quality [1] and to make sure that the power system can work securely, economically and reliably.

There are a variety of harmonic separation methods [2-6], which are of their own characteristics. For example, the method of instantaneous reactive power theory has the properties of fast dynamic response speed, small time delay and strong real-time, but its topology structure is very complicated. Fourier transformation and its improved harmonic method which are high-precision and multi-functional have a large amount of calculation and exist spectrum leak and fencing effects. The method of neural network can apply neurons’ self-learning algorithm to the harmonic detection, but its convergence will be influenced by the size of learning rate.

Independent component analysis (ICA) [7-9] developed in 1990s is one of the many solutions to the blind source separation (BSS) problems. ICA looks for the components that are both statistically independent and non-Gaussian. The various ICA algorithms extract source signals based on the principle of information maximization, mutual information minimization, maximum likelihood estimation and maximizing nongaussianity. ICA has applied to many fields [9-14] such as in statistical signal processing, telecommunications, medical image processing, feature extraction and economic analysis.

This paper models the problem of harmonic separation as a blind source separation task and to solve it using a method combining Fast ICA [15-19] and Least Square method [20-21] together. Firstly, the proposed method uses Fast ICA to determine harmonic orders by constructing proper number of virtue observed channels. Then the least square method is used to determine amplitudes and phases of harmonics containing in the mixing signal. The result is shown that the method presented in this paper has a fast speed and high precision.

II. PRINCIPLE OF ICA

A. Mathematical Model of ICA

ICA is a new BSS technology, which is defined that multiple independent components are decomposed from observed signals by optimization algorithm to obtain estimation of the source signals according to principle of statistical independence. In other words, ICA is a way of finding a linear non-orthogonal co-ordinate system in any multivariate data. The directions of the axes of this co-ordinate system are determined by both the second and higher order statistics of the original data. The goal is to perform a linear transformation which makes the resulting variables as statistically independent from each other as possible.

Assume that we observe \( M \) linear mixtures \( x_1, x_2, \ldots, x_M \) of \( N \) independent sources
\[ x_i = a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{iN}s_N, \text{ for all } i. \] (1)

In (1), the time has been ignored. Instead, it is assumed that each mixture \( x_i \) as well as each independent component \( s_j \) is random variable.

Now rewrite the above equation by using vector-matrix notation. Denote by \( x \) the random vector whose elements are the mixtures \( x_1, x_2, \ldots, x_M \), and likewise by \( s \) the random vector with elements \( s_1, s_2, \ldots, s_N \). Denote by \( A \) the matrix with elements \( a_{ij} \). Therefore, the above mixing model (1) can be written as

\[ x = As. \] (2)

When the mixing signal is collected by \( K \) times, the model will be written as

\[ X = AS = \sum_{i=1}^{n} a_{i}s_j, \] (3)

where \( X, A, S \) are an \( M \times K \), \( M \times N \), \( N \times K \) matrix, respectively. The goal is to find a linear transformation matrix \( W \) without knowledge of the mixing matrix \( A \) and original sources \( S \), so that the estimation signals \( \hat{S} \) can approach \( S \) as much as possible, then the recovery model can be written as

\[ \hat{S} = WX. \] (4)

The following assumptions are adopted for ICA.

1) The components \( s_j \) are statistically independent;
2) At most one of the components \( s_j \) is Gaussian distributed;
3) The number of observed mixtures \( M \) is no less than that of independent components \( N \);
4) In many applications, it would be more realistic to assume that there is a noise in the measurement data, which would mean adding a noise term in the model. However, for simplicity, any noise terms are omitted.

However, there exist two ambiguities in the ICA model in (3).

1) We can not determine the variances of the independent components.

The reason is that, both \( s \) and \( A \) being unknown, we can freely change the order of the terms in the sum in (3), and call any of the independent components the first one. Formally, a permutation matrix \( P \) and its inverse \( P^{-1} \), which has only one nonzero element of value 1 in each row and column, can be substituted in the model to give \( X = AP^{-1}PS \), and it will not change the measurement matrix \( X \). The elements of \( PS \) are the original independent variables \( s_j \), but in another order. The matrix \( AP^{-1} \) is just a new unknown mixing matrix to be solved by the ICA algorithms.

**B. Fast ICA**

Fast ICA is an algorithm which is based on a fixed-point iteration scheme for finding a maximum of the non-Gaussian. It is the most popular algorithm used in various applications as it is simple, fast convergent and computationally less complex. It’s the first step to complete the task discussed in the paper.

Before applying Fast ICA algorithm on the data, it is usually very useful to do some preprocessing. The most basic and necessary preprocessing is to center \( x \), i.e. subtract its mean vector \( E(x) \) so as to make \( x \) a zero-mean variable

\[ x = x - E(x). \] (5)

After centering, another useful preprocessing strategy is to whiten the observed variables. The observed data \( x \) is transformed linearly so that a new vector \( z \) is uncorrelated and their variance equals unity, i.e. \( E(z^2) = I \). This whitening is done by using eigenvalue decomposition (EVD) of the covariance matrix. The new mixing matrix obtained using whitening is given by

\[ z = VD^{-1/2}x = Bx, \] (6)

where \( V \) is the matrix of orthogonal eigenvectors and \( D \) is a diagonal matrix with the corresponding eigenvalues, and \( D^{-1/2} = diag(d_1^{-1/2}, \ldots, d_n^{-1/2}) \). And whitening transforms the mixing matrix \( A \) into a new one \( B \).

The basic form of the Fast ICA algorithm is as follows,

1) Let \( i = 1 \);
2) Choose an initial weight vector \( w_i(0) \) (It satisfies \( \|w_i(0)\| = 1 \), and let \( k = 1 \);
3) Let \( w_i(k) = E\left[zw_i(k-1)^Tz\right] - 3w_i(k-1) \);
4) Normalize \( w_i(k) \), \( w_i(k) = w_i(k)/\|w_i(k)\| \);
5) If not converged, let \( k = k + 1 \) and go back to step 3;
6) If \( k < N \), let \( i = i + 1 \) and go back to step 2;

Then the resulting separated matrix \( W \) is

\[ W = [w(1), w(2), \ldots, w(M)]^T. \] (7)
And the mixing matrix \( A \) can be obtained by computing the inverse of \( W \).

III. APPLICATION IN HARMONIC SEPARATION BASED ON ICA

A. Construction of Virtue Observed Channels

The electric signals which can be directly observed are a cyclical non-sinusoidal signal, and denote it by \( x_1 \).

\[
x_1 = a_1 s_1 + a_2 s_2 + \cdots + a_N s_N = a_1 \sin(\omega t + \phi_1) + a_2 \sin(2\omega t + \phi_2) + \cdots + a_N \sin(N\omega t + \phi_N). \tag{8}
\]

When \( x_1 \) is decomposed by Fourier transformation, fundamental and a series of harmonic components can be obtained. Then \( a_{11}, a_{12}, \ldots, a_{1N} \) are amplitudes of the independent components \( s_1, s_2, \ldots, s_N \).

During the process of harmonic separation, there is only one actual observed channel, while it’s demanded that all of independent components \( N \) for ICA model. Therefore, it is needed that at least another \( M \) channels should be constructed, then it can be constructed by the following \((M - 1)\) expansion equations

\[
\begin{align*}
x_2 &= a_2 s_1 + a_{22} s_2 + \cdots + a_{2N} s_N = a_2 \sin(\omega t + \phi_1) + a_{22} \sin(2\omega t + \phi_2) + \cdots + a_{2N} \sin(N\omega t + \phi_N) \\
x_3 &= a_3 s_1 + a_{32} s_2 + \cdots + a_{3N} s_N = a_3 \sin(\omega t + \phi_1) + a_{32} \sin(2\omega t + \phi_2) + \cdots + a_{3N} \sin(N\omega t + \phi_N) \\
&\vdots \\
x_M &= a_M s_1 + a_{M2} s_2 + \cdots + a_{MN} s_N = a_M \sin(\omega t + \phi_1) + a_{M2} \sin(2\omega t + \phi_2) + \cdots + a_{MN} \sin(N\omega t + \phi_N)
\end{align*}
\]

B. Determination of Harmonic Frequencies

Firstly, a period of timed-sampling data is obtained from the signals \( x_1, x_2, \ldots, x_M \). In the paper, the number of timed-sampling data is chosen as \( K \), i.e. size \((x_1) = \text{size}(x_2) = \ldots = \text{size}(x_M) = (1 \times K)\). Then a matrix \( X \), i.e. size \((X) = (M \times K)\), is made up. Next, the Fast ICA algorithm presented in previous part is used to determine all the harmonic frequencies containing in the data. During this step, the separate matrix \( A \) and mixing matrix \( W \) are obtained. Finally, harmonic frequencies will be determined by \( W \) and \( A \).

C. Determination of Harmonic Amplitudes and Phases

From the previous part, it’s known that amplitudes and phases of the harmonic components need to be determined by the least square method after all the harmonic frequencies are determined.

In (9), \( s_N \) is expressed by the following equation

\[
s_N = C_N \sin(N\omega t + \phi_N). \tag{10}
\]

Normally, harmonic frequencies can be achieved by simple frequency analysis. However, harmonic amplitudes and phases are often unknown. Therefore, it needs to take the following transformation for signal \( s_N \)

\[
\begin{align*}
s_N &= C_N \sin(N\omega t + \phi_N) \\
&= C_N \sin(N\omega t) \cos(\phi_N) + C_N \cos(N\omega t) \sin(\phi_N) \\
&= A_N \sin(N\omega t) + B_N \cos(N\omega t) \\
&= s_{N1} + s_{N2}
\end{align*}
\]

Thus problems of estimation amplitude and phase of \( s_N \) become estimation of \( s_{N1} \) and \( s_{N2} \). Then (8) is rewritten in the following form of vector-matrix notation

\[
x_1 = [\sin(\omega t), \cos(\omega t), \ldots, \sin(N\omega t), \cos(N\omega t)]^T \cdot [a_1, a_2, \ldots, a_{2j-1}, a_{2j+1}, \ldots, a_{2N-1}, a_{2N}]^T.
\]

And the problems finally become to calculate values of \( a_1, a_2, \ldots, a_{2j-1}, a_{2j+1}, \ldots, a_{2N-1}, a_{2N} \).

When the least square method is used, firstly, a period of timed-sampling data is obtained from the signal \( x_1 \), then \( X \) is made up of by these discrete values, i.e. \( X = [x_{(1)}, x_{(2)}, \ldots, x_{(i)}, x_{(K)}]^T \), and let

\[
h = [\sin(\omega t), \cos(\omega t), \ldots, \sin(N\omega t), \cos(N\omega t)]^T, \tag{13}
\]

\[
H = [h_1, h_2, \ldots, h_i, \ldots, h_j]^T, \tag{14}
\]

\[
\theta = [a_1, a_2, \ldots, a_{2j-1}, a_{2j+1}, \ldots, a_{2N-1}, a_{2N}]^T. \tag{15}
\]

Then the following (16) is obtained

\[
X = H\theta. \tag{16}
\]

By taking a inverse operation to (16), the values of the amplitudes can be calculated out by (17)

\[
\theta = (H^T H)^{-1} H^T X. \tag{17}
\]

Then the phase of original source \( s_N \) can be also achieved by

\[
\phi_N = \arctan(a_{2N} / a_{2N-1}). \tag{18}
\]

IV. SIMULATION AND RESULTS ANALYSIS

A. Simulation

One observed signal consisting of fundamental, 3rd and 5th harmonics and 4 other virtual signals are simulated by MATLAB, and the fundamental frequency is assumed to be 50 Hz.

\[
\begin{align*}
x_1 &= 3 \sin(\omega t) + 0.6 \sin(3\omega t + \pi / 3) + 0.8 \sin(5\omega t + \pi / 5) \\
x_2 &= \sin(3\omega t) \\
x_3 &= \cos(3\omega t) \\
x_4 &= \sin(5\omega t) \\
x_5 &= \cos(5\omega t)
\end{align*}
\]

Ten cycles of time-sampling data are gathered during the simulation, and the five signals waves are shown in Fig. 1.
It’s assumed that the data is preprocessed by centering and whitening. Firstly, the harmonic frequencies are determined by using Fast ICA algorithm and the separated matrix $W$ and mixing matrix $A$ are as matrix (20) and (21), respectively.

$$ W = \begin{bmatrix}
-0.9998 & 0.0004 & 0.0007 & 0.0008 & 0.0006 \\
-0.0089 & -1.0000 & -0.0000 & 0.0000 & -0.0000 \\
0.0131 & 0.0000 & -0.0000 & -1.0000 & 0.0000 \\
0.0076 & -0.0000 & 0.0000 & -0.0000 & -1.0000 \\
\end{bmatrix} \quad (20) $$

$$ A = \begin{bmatrix}
-1.0002 & -0.0004 & -0.0007 & -0.0008 & -0.0006 \\
0.0089 & -1.0000 & 0.0000 & 0.0000 & 0.0000 \\
-0.0068 & -0.0000 & -1.0000 & -0.0000 & -0.0000 \\
-0.0131 & -0.0000 & 0.0000 & -1.0000 & 0.0000 \\
0.0076 & -0.0000 & 0.0000 & -0.0000 & -1.0000 \\
\end{bmatrix} \quad (21) $$

According to (20) and (21), it’s shown that the observed signal contains fundamental, 3rd and 5th harmonics. However, amplitudes and phases are not the real values of the containing harmonics. Then the least square method is used to determine the amplitudes and phases of harmonic components, and the resulting waves of separated signals are shown in Fig. 2.

And the resulting separated matrix $W$ and mixing matrix $A$ are as matrix (22) and (23), respectively.

$$ W = \begin{bmatrix}
0.3333 & -0.1000 & -0.1732 & -0.2157 & -0.1567 \\
0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \\
\end{bmatrix} \quad (22) $$

$$ A = \begin{bmatrix}
3.0000 & 0.3000 & 0.5196 & 0.6472 & 0.4702 \\
0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \\
\end{bmatrix} \quad (23) $$

### B. Error Analysis

Error analysis of harmonic separation is shown in Tab. I, and the values of amplitudes and phases are obtained by using the proposed in this paper.

From Tab. I, it is shown that the observed signals can be separated with a high precision by using the proposed algorithm.

<table>
<thead>
<tr>
<th>Harmonic orders</th>
<th>Fundamental</th>
<th>3rd</th>
<th>5th</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>sin(ωt)</td>
<td>cos(ωt)</td>
<td>sin(3ωt)</td>
</tr>
<tr>
<td>Amplitudes</td>
<td>3.0000</td>
<td>0.0000</td>
<td>0.3</td>
</tr>
<tr>
<td>Amplitude errors</td>
<td>0</td>
<td>0.0022%</td>
<td>0.00344%</td>
</tr>
<tr>
<td>Phases</td>
<td>0</td>
<td>$\pi/3$</td>
<td>$\pi/5$</td>
</tr>
<tr>
<td>Phase errors</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

TABLE I.
ERROR ANALYSIS OF HARMONIC SEPARATION
V. APPLICATION OF ICA IN REAL CASES

For testing the effectiveness of the combined algorithm, it is used to process data obtained from actual scenes. In this paper, original voltage and current time-sampling data are both achieved from a substation, and the measurement data are shown in Appendix A.

According to the principle of ICA, it firstly does the preprocessing of centering and whitening on the measurements before using the proposed algorithm. Then it takes two steps, that is, the first step is to determine harmonic frequencies by using Fast ICA and the second step is to determine harmonic amplitudes and phases by using least square method, to complete the separation task.

Firstly, harmonics containing in the voltage measurement data are separated by using the combined algorithm proposed above.

Waves of Fig. 3 show that the resulting separated voltage data, and Fig. 4 show that the separated voltage amplitudes of fundamental and all harmonics from 2nd to 18th.

Figure 3  Separated voltage data.

Figure 4  Amplitudes of the fundamental and all harmonics in voltage measurement data.

Figure 5  Separated current data.
Then, the algorithm is used to separate harmonics containing in the current measurement data. Waves of Fig. 5 show that the resulting separated current data, and Fig. 6 show that the separated current amplitudes of fundamental and other harmonics from 2nd to 18th.

The results of Tab. II and Tab. III show that the finally separated voltage and current amplitudes of fundamental and all harmonics in the voltage and current measurement data, respectively.

From Tab. II, it is shown that the harmonic containing rates of orders of 3rd, 5th, 7th, 9th, 11th, 15th and 17th are much higher than other harmonics in voltage measurement data. And from Tab. III, it is shown that the rates of orders of 2nd, 3rd, 5th, 7th, 15th and 17th are much higher in current data. In addition, they both hardly contain even-order harmonics in the voltage and current measurement data.

Similarly, errors are also analyzed in this real case. The voltage errors between real observed signals and separated signals are shown in Fig. 7, and the current errors between them are shown in Fig. 8.

From Fig. 7 and Fig. 8, it can be seen that the errors are quite small, and it’s suggested that the actual signal can also be separated with a high precision by using the proposed algorithm.

Then harmonic voltage distortion rate can be calculated by

\[
THD_v = \frac{\sqrt{U_2^2 + U_3^2 + \cdots + U_{17}^2}}{U_1} = 1.67\%. \quad (24)
\]

<table>
<thead>
<tr>
<th>Orders</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
<th>6th</th>
<th>7th</th>
<th>8th</th>
<th>9th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplitudes</td>
<td>8779.5</td>
<td>45.7</td>
<td>115.7</td>
<td>20.5</td>
<td>44.8</td>
<td>12.1</td>
<td>49.1</td>
<td>9.3</td>
<td>22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Orders</th>
<th>10th</th>
<th>11th</th>
<th>12th</th>
<th>13th</th>
<th>14th</th>
<th>15th</th>
<th>16th</th>
<th>17th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplitudes</td>
<td>7.2</td>
<td>18.3</td>
<td>6.9</td>
<td>6.4</td>
<td>5.9</td>
<td>8.3</td>
<td>5.2</td>
<td>8.3</td>
</tr>
</tbody>
</table>
TABLE III.
AMPLITUDES OF FUNDAMENTAL AND HARMONICS IN CURRENT MEASUREMENT DATA

<table>
<thead>
<tr>
<th>Orders</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
<th>6th</th>
<th>7th</th>
<th>8th</th>
<th>9th</th>
<th>10th</th>
<th>11th</th>
<th>12th</th>
<th>13th</th>
<th>14th</th>
<th>15th</th>
<th>16th</th>
<th>17th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplitudes</td>
<td>355.7015</td>
<td>2.1159</td>
<td>6.1954</td>
<td>0.7082</td>
<td>3.691</td>
<td>0.3771</td>
<td>5.2230</td>
<td>0.4455</td>
<td>0.1146</td>
<td>0.3280</td>
<td>0.3683</td>
<td>0.3023</td>
<td>0.4607</td>
<td>0.3406</td>
<td>2.9621</td>
<td>0.7385</td>
<td>3.7524</td>
</tr>
</tbody>
</table>

And harmonic current distortion rate is

\[
THD = \sqrt{\frac{I_2^2 + I_3^2 + \ldots + I_{17}^2}{I_1}} = 2.95%. \tag{25}
\]

In addition, the measurement data achieved from the substation are also analyzed by power quality analyzer WT3000. The results are that harmonic voltage and current distortion rate are 1.68% and 3.02%, respectively. Therefore, it validates that harmonics can be separated effectively by using the algorithm proposed in this paper.

VI. CONCLUSION

This paper models the problems of harmonic separation as blind source separation task, and presents an algorithm which combines the famous Fast ICA and least square method to carry on harmonic separation. During the separation, there is only one observed signal that can be used, virtual signal channels need to be constructed for ICA analysis. Besides, there exist two ambiguities in the ICA model. Therefore, it needs to take two steps to complete the separation task. The result is shown that fundamental and harmonics can be separated with a high precision and fast speed by using the algorithm presented in this paper.

APPENDIX A  ORIGINAL SAMPLING DATA

The original voltage and current time-sampling data achieved from the substation are shown in Fig. A-1 and Fig. A-2, respectively.

ACKNOWLEDGMENT

This work was supported by the Open Laboratory Foundation of Key disciplines of Control Engineering of Henan Provincial High Education Grant KG201101 and Doctor Foundation of Henan polytechnic University Grant 648207.

REFERENCES


Yongle Ai was born in Henan, China and received the B.S. degree in Power Electronics from Xi’an University of Technology, Xi’an, China, in 1987, received the M.S. degree in Electrical Engineering from Southwest Jiaotong University, Chengdu, China, in 1995, and the Ph.D. degree in Electrical Engineering from University of Stellenbosch, South Africa in 2006.

He joined the Henan Polytechnic University from 1987 to now. He is currently a professor at School of Electical Engineering and Automation, Henan Polytechnic University, China. His research interests include multiphase AC machine control and power system filter.

Prof. Ai is a member of Chinese Society for Electrical Engineering.

Haiyang Zhang was born in Henan, China and received the B.S. degree in Electrical Engineering from Henan Polytechnic University, Jiaozuo, Henan, China, in 2010.

He is currently pursuing the M.S. degree at School of Electrical Engineering and Automation, Henan Polytechnic University, China.