Short communication

Can back-projection fully resolve polarity indeterminacy of independent component analysis in study of event-related potential?

Fengyu Cong*, Igor Kalyakin, Tapani Ristaniemi

Department of Mathematical Information Technology, PL 35 (Agora), University of Jyväskylä, 40014 Jyväskylä, Finland

**Article Info**

Article history:
Received 4 January 2010
Received in revised form 19 May 2010
Accepted 19 May 2010
Available online 15 September 2010

Keywords:
ICA
Global optimization
Local optimization
Parallel
Sequential
Projection
Polarity
ERP

**Abstract**

In the study of event-related potentials (ERPs) using independent component analysis (ICA), it is a traditional way to project the extracted ERP component back to electrodes for correcting its scaling (magnitude and polarity) indeterminacy. However, ICA tends to be locally optimized in practice, and then, the back-projection of a component estimated by the ICA can possibly not fully correct its polarity at every electrode. We demonstrate this phenomenon from the view of the theoretical analysis and numerical simulations, and suggest checking and modifying the abnormal polarity of the projected component in the electrode field before further analysis. Moreover, when several components are to be projected, instead of the parallel projection of those components simultaneously, the sequential projection of component by component permits the correction of the abnormal polarity of a certain projected component at a certain electrode, which can improve the accuracy of the back-projection. Furthermore, after one extracted component by the ICA is projected back to electrodes under the global optimization, we cannot achieve the real source yet, but the determined scaled source, i.e., the multiplication between the real source and the mapping coefficient from the source to the point at the scalp.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Independent component analysis (ICA) [1] has been extensively used in the study of event-related potentials (ERPs) [2]. It is known that the independent components (ICs) estimated by ICA tend to possess the magnitude and polarity (positive or negative) indeterminacy [1]. Because the peak amplitude is one of the most important parameters to describe an ERP [3], the correction of such ambiguity of an IC is necessary to the study of ERPs. Thus, a back-projection of the desired ERP component to the electrode field often follows the ICA decomposition [4]. For example, as Makeig and colleagues noted, the projection of the ith IC onto the original data channels is given by the outer product of the ith row of the component activation matrix with the ith column of the inverse unmixing matrix, and is in the original units (e.g., microvolts) [4]. However, despite of the vital influence of the back-projection to the further study of ERPs, the performance of the back-projection has never been deeply analyzed except by our previous report [5].

This study attempts to analyze the mathematical composition of the back-projection of ICs under the global and local optimization of the ICA decomposition [1,6,7], respectively. Particularly, the polarity is important to identify an ERP, for example, N1 is the negative peak and P3 is the positive peak [3] under certain reference. This study mainly discusses the polarity of the back-projection of an estimated ERP-like IC in the electrode field.

2. Abnormal polarity of projected components at electrodes

2.1. General solution of ICA on EEG

The classic ICA decomposition on EEG may be illustrated as

\[ x(t) = Ax(t), \]

\[ y(t) = Wx(t), \]

where \( x(t) = [x_1(t), \ldots, x_N(t)]^T \) denotes available multichannel EEG recordings at the scalp, \( s(t) = [s_1(t), \ldots, s_M(t)]^T \) represents unknown sources of the brain, \( A \) with full rank contains the unknown mapping coefficients from any source to any electrode, \( W \) exhibits the unmixing matrix, and \( y(t) = [y_1(t), \ldots, y_M(t)]^T \) manifests the extracted components by the ICA. The ICA decomposition is to find the unmixing matrix through minimizing or maximizing some cost functions [1,7]. Without loss of generality, we assume \( N \) sources and \( N \) sensors in this study.

Usually, after the unmixing matrix is obtained, its inverse is often computed to recover of original units of recordings. The projection matrix is simply obtained by

\[ B = W^{-1}. \]
This is because every column of the projection matrix includes the relative projection strengths of the corresponding component onto all of the scalp sensors [4]. Furthermore, the accuracy of such projection depends on the performance of the ICA decomposition.

2.2. Global matrix and projection matrix

One method to evaluate the performance of ICA is to investigate its global matrix \( C \) [7]. Then, the ICA solution in Section 2.1 can be interpreted as below

\[
C = WA. \tag{4}
\]

\[
y(t) = Cs(t). \tag{5}
\]

\[
B = AC^{-1}. \tag{6}
\]

When only one nonzero element exists in each row and each column of the global matrix \( C \), the performance of the ICA is globally optimized; otherwise it is locally optimized.

2.2.1. Global optimization and projection matrix

Under the global optimization, \( C \) can be decomposed into a permutation matrix \( P \) and a diagonal matrix \( D \), i.e.,

\[
C = PD. \tag{7}
\]

Thus, the estimated components turn to be

\[
y(t) = PDs(t). \tag{8}
\]

Explicitly, Eq. (8) means that every estimated component is scaled version of the corresponding source and does not contain the information of other sources.

Furthermore, the projection matrix becomes

\[
B = AD^{-1}P^{-1}. \tag{9}
\]

Because \( D^{-1} \) and \( P^{-1} \) are also the diagonal and permutation matrices respectively, every column of the projection matrix \( B \) corresponds to the permuted and scaled version of the relevant column of the mapping matrix \( A \).

2.2.2. Local optimization and projection matrix

Under the local optimization, the global matrix cannot be factorized into the multiplication of the permutation matrix \( P \) and the diagonal matrix \( D \), hence, Eqs. (7)–(9) become inequations, i.e.,

\[
C \neq PD, \quad y(t) \neq PDs(t), \quad B \neq AD^{-1}P^{-1}. \tag{10}
\]

Consequently, every estimated component is still the mixture of some sources, and every column of the projection matrix \( B \) is the mixture of randomly scaled relative projection strengths of some components onto all of the scalp sensors. In this case, the estimated components and the projection matrix can only be obtained from Eqs. (2) and (3). Since such a case may appear with very high probability using ICA to extract ERPs [5,6], we are interested in studying the projection under the local optimization.

2.3. Project one component back to electrodes

The projection of the \( k \)th IC at the \( i \)th electrode can be described as

\[
e_k(t) = b_{ik}y_k(t), \tag{10}
\]

where \( b_{ik} \) is the \( i \)th element of the \( k \)th column of \( B \) and \( y_k(t) \) is the \( k \)th component of \( y(t) \).

2.3.1. Projection under global optimization

Under the global optimization, based on Eq. (8), the estimated component can be described as

\[
y_k(t) = p_{km}d_{mm}s_m(t) = d_{mm}s_m(t), \tag{11}
\]

where \( p_{km} \) with value ‘1’ in a permutation matrix is the nonzero element of \( P \) at the \( k \)th row and the \( m \)th column, \( d_{mm} \) is the \( m \)th diagonal element of \( D \), and \( s_m(t) \) is the \( m \)th element of \( s(t) \). Based on Eq. (9), the projection coefficients at the \( k \)th column of \( B \) and the \( i \)th electrode can be interpreted as

\[
b_{ik} = a_{im} \frac{1}{d_{mm}} p_{mk} = \frac{a_{im}}{d_{mm}}. \tag{12}
\]

Substituting Eqs. (11) and (12) into Eq. (10), we obtain the projection of the \( k \)th component at the \( i \)th electrode as below,

\[
e_k(t) = a_{im}s_m(t), \tag{13}
\]

where, to discriminate the projection under the local optimization, \( e_k(t) \) is used for the notation instead of \( e_k(t) \) for the projection under the global optimization.

2.3.2. Projection under local optimization

Under the local optimization, the estimated component derived from Eq. (5) can be illustrated as

\[
y_k(t) = \sum_{j=1}^{N} c_{kj}s_j(t), \tag{14}
\]

where some \( c_{kj} \) is the element of \( \{1, \ldots, N\} \) may be zero and this depends on the performance of the ICA. Substituting Eq. (14) into (10), we achieve the projection of the \( k \)th component at the \( i \)th electrode under the local optimization as the following

\[
e_k(t) = b_{ik} \sum_{j=1}^{N} c_{kj}s_j(t) = b_{ik} c_{km}s_m(t) + b_{ik} \sum_{j=1, j \neq m}^{N} c_{kj}s_j(t). \tag{15}
\]

The right part of Eq. (15) is separated into two items because \( b_{ik} c_{km} \) can dominate \( e_k(t) \) under the satisfactory ICA performance, i.e., in the \( k \)th row of the global matrix \( C \),

\[
|c_{km}| > |c_{kj}|, \tag{16}
\]

where \(|\cdot|\) denotes the absolute value of a scalar.

Under the assumption of Eq. (16), to a certain source, \( b_{ik} c_{km} \) can determine the polarity and can almost determine magnitude of the projected component in the electrode field. Under the global optimization, \( b_{ik} c_{km} \) in Eq. (15) is equal to the mapping coefficient \( a_{im} \) in Eq. (13); however, under the local optimization, both \( c_{km} \) and \( b_{ik} \) do not obey any essential relationship as illustrated in Section 2.2.2, moreover, they may have random polarities, respectively, hence, the sign of \( b_{ik} c_{km} \) may be different with that of \( a_{im} \). Then, the polarity of the projected component \( e_k(t) \) may be opposite to that of the mapping component \( e_k(t) \). In such a case, we define that the abnormal polarity happens [5].

To correct the abnormal polarity, the projected component \( e_k(t) \) can be multiplied by ‘−1’. Indeed, such correction does not change the sign of the ‘actual’ brain sources. Under the local optimization, Eq. (15) is the approximation of Eq. (13) under the assumption mentioned above. This means there are errors between \( q_k(t) \) and \( e_k(t) \). Such errors are just originated from the separation matrix \( W \) in Eq. (2). Under the global optimization, as illustrated by Eqs. (8) and (11), the separation matrix \( W \) can guarantee that only one estimated component corresponds to one source. However, under the local optimization, as illustrated by Eq. (14), the separation matrix \( W \) cannot guarantee that only one estimated component contains the information of one source. Furthermore, since the projection matrix \( B \) is the inverse of the separation matrix, it contains errors too. The correction made in this study is indeed the post-processing after the ICA, and it can also be regarded as the further processing on certain element of the projection matrix \( B \). So, the correction does not change the real source in the brain.
### 2.4. Project some components back to electrodes

In some cases, projection of several components simultaneously back to electrodes is used [2]. Two components are assumed to be two desired sources in this study. For example, the projection of two components is usually implemented as

\[
\mathbf{e}(t) = [\mathbf{b}_1, \mathbf{b}_2] [y_{k_1}(t), y_{k_2}(t)]^T = \mathbf{b}_{k_1} y_{k_1}(t) + \mathbf{b}_{k_2} y_{k_2}(t) = \mathbf{e}_{k_1}(t) + \mathbf{e}_{k_2}(t). \tag{17}
\]

From Eq. (17), we can observe that the parallel projection of two components is equal to the sum of the sequential projection of each component no matter under the global or local optimization. Since the polarity indeterminacy cannot be fully modified under the local optimization, the polarity indeterminacy at some electrodes may happen on either \(e_{k_1}(t)\) or \(e_{k_2}(t)\) or both of them. Moreover, the parallel projection in Eq. (17) does not permit correcting the abnormal polarities of a certain component at some electrodes. However, if the projection is implemented as component by component, the correction becomes available for the abnormal polarity of a certain component at a certain electrode. Such sequential projection should become closer to mapping process than the parallel projection does.

Next three steps are the paradigm to project \(Q (1 < Q < N)\) components:

1. Project each selected component back to electrodes respectively according to Eq. (15) to gain the \(\mathbf{e}_{k_q}(t) (q = 1, \ldots, Q)\). \(k_q\) is the element of \(\{1, \ldots, N\}\).
2. Correct the abnormal polarity at some electrodes for each projected components to acquire the \(\hat{\mathbf{e}}_{k_q}(t)\). If there is no abnormal polarity, \(\hat{\mathbf{e}}_{k_q}(t) = \mathbf{e}_{k_q}(t)\).
3. Sum the modified projection of each component to obtain the final projection of all selected components as below

\[
\mathbf{e}(t) = \sum_{q=1}^{Q} \hat{\mathbf{e}}_{k_q}(t). \tag{18}
\]

It should be noted that the mixture \(\mathbf{x}(t)\) is produced if all components are projected back to electrodes. Obviously, this is no use for further study.

### 3. Simulation

This part demonstrates how the polarity problem affects the projection as illustrated by Eq. (15) under the local optimization of the ICA decomposition. During the simulation, the global matrix \(\mathbf{C}\) is generated directly instead of any ICA solution, and then \(\mathbf{W}\) is computed from Eq. (4). By doing this way, the possibility of errors taken by any ICA algorithm is avoided. Since the ICA algorithm seeks the unmixing matrix, this treatment may not take any negative effect for the further analysis.

---

**Fig. 1.** Waveforms of sources, mixtures, and estimated components.
Six brain sources with unit variance and zero mean are obtained from the data set of an auditory ERP [8]. The sampling frequency was 200 Hz, and the ERPs were elicited by an uninterrupted sound under the oddball paradigm [8]. The mixing matrix with full rank is randomly generated. According to Eq. (16), let $C$ have only one element with a comparatively large absolute value in each row and each column. This covers the locally optimized ICA decomposition. To quantify the simulation, we define the signal-to-interference ratio (SIR) for the ICA to the mixtures and estimated components as

$$
SIRA = 10 \log_{10} \left\{ \frac{\sum_{t=1}^{T} [a_{km}(t)]^2}{\sum_{t=1}^{T} \left[ \sum_{j=1, j \neq m}^{n} q_{j}(t) \right]^2} \right\},
$$

$$
SIRC = 10 \log_{10} \left\{ \frac{\sum_{t=1}^{T} [c_{km}(t)]^2}{\sum_{t=1}^{T} \left[ \sum_{j=1, j \neq m}^{n} c_{j}(t) \right]^2} \right\}.
$$

The signal-to-difference ratio improvement (SDRI) for the corrected projection of the abnormal polarity can be expressed as

$$
SDRI = 10 \log_{10} \left\{ \frac{\text{var} [e_{1}(t) - q_{1}(t)]}{\text{var} [-e_{1}(t) - q_{1}(t)]} \right\},
$$

where var($\bullet$) means the variance of a vector. The SDRI interprets the improvement of the projection through the correction of the abnormal polarity.

Fig. 1 shows the estimated ICs, mixtures, and sources. The second and third sources are chosen as the target activations. From the visual inspection of the waveform, we can observe that the separation is satisfactory. Moreover, $SIRA = -8.0$ dB, $SIRC = 32.6$ dB.

On the projection stage, firstly, the two desired sources and the corresponding estimated components are mapped and projected back to electrodes respectively according to Eqs. (13) and (15). Fig. 2 shows that the projection of the second component has the polarity reversal at the second channel. After modification, the corrected projection becomes almost identical to the mapping, i.e., the theoretical expectation. Secondly, the two sources and the two desired components are respectively mapped and projected to electrodes simultaneously. Fig. 2 evidently exhibits that the parallel projection of two components at the second channel severely violates the true mapping. With the sequential projection and the abnormal polarity correction through Eq. (18), the corrected projection becomes closer to the mapping. This demonstrates the effectiveness of the sequential projection.

To reflect the validation of the recommended projection procedure, the same simulation paradigms are run 1000 times with the same sources. The SIRA and SIRC are variable across 1000 runs. Table 1 demonstrates the number of abnormal polarities and other averaged parameters for the performance of the ICA and the benefit of the abnormal polarities correction to each source.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Times</th>
<th>Mean SIRA (dB)</th>
<th>Mean SIRC (dB)</th>
<th>Mean SDRI (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source-#2</td>
<td>135</td>
<td>-27.9</td>
<td>36.2</td>
<td>8.6</td>
</tr>
<tr>
<td>Source-#3</td>
<td>149</td>
<td>-5.3</td>
<td>15.1</td>
<td>7.7</td>
</tr>
</tbody>
</table>
4. Conclusion

Since ICA decomposition tends to be locally optimized in real EEG data, we find that the back-projection of the extracted ERP component may possibly not fully correct the polarity ambiguity of the estimation in the electrode field when using ICA to study ERPs. Because the polarity of an ERP is very important, it is necessary to correct this indeterminacy. As the polarity has only two factors, i.e., positive and negative, the projected component in the electrode field can either keep consistent to the truth or conflicts to the truth, and the correction is possible. In the case of the projection of several components, the parallel projection of those components does not permit the modification of the abnormal polarities of a certain component in the electrode field. Fortunately, the parallel projection is equal to the sum of the sequential projection of each component. The simulations have shown that the recommended sequential projection with the abnormal polarity correction outperforms the parallel one. Also, the simulations have demonstrated that the abnormal polarity may appear with a higher probability at the electrodes where ERPs are smaller.

Without a question, to achieve more accurate estimation heavily depends on the better performance of ICA. However, under the certain extent of the ICA decomposition, the correction of the abnormal polarity of the projected component is helpful. Hence, in the study of ERPs when using ICA or other linear model related methods, we suggest to check correspondence of the polarity of the projected component in the electrode field in contrast to the theoretical expectations. If the checked polarity is reversed, the correction of the abnormal polarity may improve the accuracy of the back-projection.

Moreover, Eq. (13) also explicitly interprets that, under the global optimization, after projection of one component extracted by the ICA back to electrode, we cannot achieve the real source yet, but the determined scaled source, i.e., the multiplication between the real source and the mapping coefficient from the source in the brain to the point at the scalp.

Furthermore, the model of the ICA and the simulation in this study are very simple. The real EEG recordings should be much more complicated. However, as long as the EEG recordings conform to the linear transformation of latent variables, the conclusions of this study are plausible. We indeed acknowledge the fact that, without a proper evaluation based on real EEG recordings, it is impossible to know the actual occurrence of the ‘random polarities’, that is, how common they are. This will be our research topic in the near future.

Acknowledgments

Cong thanks Professor Heikki Lyytinen, Dr. Piia Astikainen (University of Jyväskylä, Finland) and Mr. Zhilin Zhang (University of California, San Diego) for invaluable discussion.

References