Blind Nonparametric Determined and Underdetermined Signal Extraction Algorithm for Dependent Source Mixtures

Fasong Wang[1(✉)], Rui Li[2], Zhongyong Wang[1], and Xiangchuan Gao[1]

[1] School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China
fasongwang@126.com
[2] School of Sciences, Henan University of Technology, Zhengzhou, China

Abstract. Blind extraction or separation statistically independent source signals from linear mixtures have been well studied in the last two decades by searching for local extrema of certain objective functions, such as nonGaussianity (NG) measure. Blind source extraction (BSE) algorithm from underdetermined linear mixtures of the statistically dependent source signals is derived using nonparametric NG measure in this paper. After showing that maximization of the NG measure can also separate or extract the statistically weak dependent source signals, the nonparametric NG measure is defined by statistical distances between different source signals distributions based on the cumulative density function (CDF) instead of traditional probability density function (PDF), which can be estimated by the quantiles and order statistics using the L^2 norm efficiently. The nonparametric NG measure can be optimized by a deflation procedure to extract or separate the dependent source signals. Simulation results for synthesis and real world data show that the proposed nonparametric extraction algorithm can extract the dependent signals and yield ideal performance.

Keywords: Blind source separation (BSS) · Nongaussian measure · Independent component analysis (ICA) · Dependent component analysis (DCA) · Blind source extraction (BSE)

1 Introduction

Blind source separation (BSS) as an active topic in signal processing community aims to separate linearly or nonlinearly mixtures in both noise-free and noisy environments mixed latent source signals. It has become an important topic of research and development in many areas [1–3]. Various algorithms have been proposed in the last two decades to separate or extract source signals from their mixtures [1–7]. The propose of BSS is to separate the potential source signals from the mixtures obtained by the sensors without a priori about the source signals and the mixing process. This is realized by a variety of criteria, including the minimization of mutual information (MI), maximization of nonGaussianity (NG) and maximization of likelihood (ML) [8, 9]. A key factor in BSS is the assumption about the statistical properties of source signals

such as statistical independence among different ones. When the source signals are mutual independent, the BSS problem can be solved by using the so called independent component analysis (ICA) method which has attracted considerable attention in the signal processing fields and several efficient algorithms have been proposed [1–3].

Despite the success of using standard ICA for BSS in many applications, original ICA algorithms are in the sense that all sources are assumed to be statistically independent random variables, but the basic assumptions of ICA may not hold for some real-world situations. Several extented ICA models have been considered based on the basic ICA framework. This type of model can be called dependent component analysis (DCA) model. Multidimensional independent component analysis (MICA) model [11, 12] as the first DCA model of BSS problem, instead of assuming all the source signals to be mutually statistically independent, all the source signals are divided into several groups and the size of the groups can be different, where the signals between different groups are statistical independent and signals within the same group have dependencies, this DCA model can also be called independent subspace analysis (ISA). DCA related algorithm concern mostly the estimation of the entropy or of the MI. Various BSS algorithms have been developed in response to different DCA models [13–24].

Another extension of the original BSS task is the blind source extraction (BSE) problem. Unlike common BSS algorithms which consist of separating all the source signals simultaneously by means of the maximization of an independence measure between the output estimated signals, however, in some situations, it may be more appropriate to extract only a single source of interest based on a certain fundamental signal property, which is the task of BSE [2, 10]. One of the main advantages of BSE compared to traditional BSS is that it decreases computational cost since the degrees of freedom are reduced, the possibility to relax the need for preprocessing or postprocessing. Furthermore, this procedure has a great potential when the number of sensors and sources are not equal, even when unknown or underdetermined.

In addition, as compared with overdetermined or determined BSS problem, the underdetermined BSS one, where the number of available recorded mixtures is less than the underlying source signals, is more difficult to treat and attracts much attention in recent years. In this case, even if the mixing matrix is known or has been estimated, it is impossible to estimate the source signals directly. Therefore, in order to realize the source extraction from the mixed signals, some a *priori* knowledge about the whole system must to be exploited, such as independence or sparsity [25, 26].

The simultaneous assumption of the two extensions of BSS, that is, DCA combined with BSE in underdetermined situation seems to be a more realistic model than any of the two models alone. For example, at the biomedical signal processing, only small number of sources should be extracted which may be weak correlated in spatial. Cardoso showed that strong relationship existed among MI, correlation and NG of source estimates [27]. Inspired from this conclusion, we get that one can not resort minimization the MI, but on the other hand, according to maximization the NG, the dependent sources can be separated or extracted.

Here, we exploit some weaker conditions for extraction or separation source signals assuming that they have statistically dependent properties in the underdetermined situation. Based on the generalization of the central limit theorem (CLT) to special dependent variables, we will try to track the DCA model by maximization NG

measure. The proposed NG measure is defined in terms of cumulative distribution function (CDF) instead of the widespread probability density function (PDF) using the nonparametric estimation method, afterwards, the NG distance between the given CDF and the standard normal CDF is proposed which can be estimated by the order statistics using the L^2 norm efficiently. The NG distance based cost function will be optimized resorting to a deflation procedure by gradient iterative algorithm, whose local maximization performs the extraction of one dependent component.

2   Theory Fundamentals

The problem of linear instantaneous BSS problem can be formulated as Eq. (1) (see [1–3] for overview):

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{1}$$

where $\mathbf{s}(t) = (s_1(t), s_2(t), \cdots, s_N(t))^{\mathrm{T}}$ is an unknown source vector which contains N source signals. The M observed mixtures $\mathbf{x}(t) = (x_1(t), x_2(t), \cdots, x_M(t))^{\mathrm{T}}$ are sometimes called as sensor outputs. Matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{M \times N}$ is an unknown full column rank mixing matrix. $\mathbf{n}(t) = (n_1(t), n_2(t), \cdots, n_N(t))^{\mathrm{T}}$ is a vector of additive noise.

The task of BSS contains estimation of the mixing matrix $\mathbf{A}$ or its pseudoinverse separating (unmixing) matrix $\mathbf{W} = \mathbf{A}^{\dagger}$ in order to estimate the original source signals $\mathbf{s}(t)$, given only a finite number of observation data $\{\mathbf{x}(t),\ t = 1, \cdots, T\}$.

In order to simplify the problem, most of the algorithms of BSS problem contain a spatial decorrelation procedure $\mathbf{V}$ over noiseless $\mathbf{x}(t)$ to obtain the decorrelated signals $\mathbf{z}(t) = (z_1(t), z_2(t), \cdots, z_M(t))^{\mathrm{T}}$, that is,

$$\mathbf{z}(t) = \mathbf{V}\mathbf{x}(t) \tag{2}$$

So, the global mixture can be expressed as,

$$\mathbf{z}(t) = \mathbf{U}\mathbf{s}(t) \tag{3}$$

where $\mathbf{U} = \mathbf{V}\mathbf{A}$ is an unknown orthogonal matrix.

After some standard BSS methods (such as ICA) are used to the preprocessed decorrelated data $\mathbf{z}(t)$, one can obtain an unitary linear transformation $\mathbf{B}$ and the estimation of the source signals $\mathbf{y}(t)$:

$$\mathbf{y}(t) = \mathbf{B}\mathbf{z}(t) = \mathbf{B}\mathbf{V}\mathbf{x}(t) = \mathbf{B}\mathbf{V}\mathbf{A}\mathbf{s}(t) \tag{4}$$

Denote $\mathbf{W} = \mathbf{B}\mathbf{V}$, then one can get

$$\mathbf{y}(t) = \mathbf{WAs}(t) \tag{5}$$

Recall that two indeterminacies cannot be resolved in BSS without some priori knowledges: scaling and permutation ambiguities. Thus, if the estimate of the mixing matrix $\hat{\mathbf{A}}$ satisfies

$$\mathbf{P} = \mathbf{WA} = \hat{\mathbf{A}}\mathbf{A} = \mathbf{GD} \tag{6}$$

where $\mathbf{P}$ is a global transformation which combines the mixing and separating system, $\mathbf{G}$ is permutation matrix and $\mathbf{D}$ is some nonsingular scaling diagonal matrix, then $(\hat{\mathbf{A}}, \hat{\mathbf{s}})$ and $(\mathbf{A}, \mathbf{s})$ are said to be related by a waveform-preserving relation.

The propose of BSE is to design an extracting vector $\mathbf{w}$ to extract an expected source signal from the mixtures $\mathbf{x}(t)$,

$$y(t) = \mathbf{w}^{\mathrm{T}}\mathbf{x}(t) = \mathbf{w}^{\mathrm{T}}\mathbf{As}(t) \tag{7}$$

where $y(t)$ is an estimated of a source signal with scalar ambiguity.

In applications, the *priori* information of the expected source signals can be utilized to design proper extracted algorithms, so, any of the source signals could come out as the first one with particular property, such as absolute normalized kurtosis value [2], temporal structure [28–30], sparseness [31], morphological structure [32] and so on. In this paper, as shown in Sect. 4, the *priori* information of the desired source signal to be extracted is the maximum NG measure.

3   Generalized CLT and Nonparametric NG Measure

The Gaussian distribution has the maximum Shannon differential entropy (maximum uncertainty) over all the continuous distributions defined on the real space with the same variance. This fact makes the Gaussianity measure a very useful tool for the characterization of data. In recent years, a connection between NG and ICA has been suggested. It can also be explained by the CLT theory. Since CLT is not valid for any set of dependent variables, we must be aware that one may not always recover the original dependent source signals using maximum NG criteria. Caiafa et al. give a very special condition on sources, for which the linear combinations of dependent signals are not more Gaussian than the components and therefore the maximum NG criteria fails, but fortunately this is not the case in most of real world scenarios [33]. Moreover, the independence of source signals is not required when we solve blind deconvolution problem [34]. Additionally, based on minimum entropy, the dependent source signals can be recovered [35].

Conclusion 1. The maximum NG method can be described as exploring a linear transformation of the mixed signals in the unit-variance signals space, so that the transformed signals (source signal estimates) have maximum NG distributions.

According to this conclusion, if we choose a robust and efficient NG measure, the source signals can be extracted or separated properly.

A natural measure of NG based on the L^2 distance of an estimated PDF to the Gaussian PDF is introduced in [19, 33]. The NG measure is defined as Eq. (8):

$$d(y, g) = \left(\int \left[p_g(y) - p_y(y) \right]^2 dy \right)^{1/2} \tag{8}$$

where the integral is defined in Lebesgue sense and is taken on all the range of variable y, and $p_g(y)$ is the Gaussian PDF with the same variance of variable y whose PDF is $p_y(y)$. In this paper, we will build the NG measure using the concept of CDF instead of traditional PDF. Let us call F_y and F_g are the CDFs of random variable y to be analyzed and its equivalent Gaussian one, respectively, then, the NG measure based on CDF can defined as Eq. (9):

$$d(F_{y_i}, F_g) = \left(\int_{-\infty}^{\infty} \left[F_{y_i}(x) - F_g(x) \right]^2 dx \right)^{1/2} \tag{9}$$

The definition of NG measure $d(F_{y_i}, F_g)$ possesses the following property that the distance measure should have:

$$\begin{cases} d(F_{y_i}, F_g) = 0, & c = 2 \\ d(F_{y_i}, F_g) > 0, & c \neq 2 \end{cases} \tag{10}$$

where c is the shape parameter in the generalized Gaussian distribution (GGD).

The PDF of GGD can be described as: $p(y) = \frac{c}{2\gamma \Gamma(1/c)} \exp\left[-\left(|y - \mu_y|/\gamma \right)^c \right]$, where $\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt$ is Gamma function and $\gamma = \sqrt{\sigma^2 \Gamma(1/c)/\Gamma(3/c)}$ is the scale parameter. By changing the values $c(c > 0)$, a family of distributions with different sharpness will be given. From the relationship between the CDF based NG measure $d(F_{y_i}, F_g)$ and the shape parameter or Gaussian parameter c, we can conclude that the distance measure $d(F_{y_i}, F_g)$ can be used as the measure of NG since it offers a global minimum when the Gaussian parameter c various from 0^+ to infinite and it reaches its global minimum when $c = 2$. In other words, the measure of NG obtains its global minimum when the analyzed distribution is Gaussian.

As the definition of $d(F_{y_i}, F_g)$, we must estimate the CDF F_{y_i}. So, the next question is how to efficient get the estimation of CDF $\hat{F}_{y_i}$. As we know, it would need high computational cost through nonparametric histograms. Alternatively, an equivalent measure can be established in terms of inverse CDF, which is defined as:

$$Q_{y_i} = F_{y_i}^{-1}, \quad Q_g = F_g^{-1} \tag{11}$$

The relationship between CDF F and its inverse Q can be generalized to the NG distance, which is formulated as:

$$\begin{cases} D(Q_{y_i}, Q_g) = 0, & c = 2 \\ D(Q_{y_i}, Q_g) > 0, & c \neq 2 \end{cases} \tag{12}$$

Since the relationship between Q and F, they also present the monotone properties in the proper intervals of c, consequently, the distance $d(\cdot)$ and its correspondent $D(\cdot)$ preserve the same properties, as a result, one can conclude that

$$D(Q_{y_i}, Q_g) = \left(\int_0^1 \left[Q_{y_i}(x) - Q_g(x) \right]^2 dx \right)^{1/2} \tag{13}$$

is also a proper NG measure. In order to estimate the NG measure $D(Q_{y_i}, Q_g)$ from the discrete samples, we must estimate Q_{y_i} firstly. The estimation of Q_{y_i} can be performed robustly in a simple practical way by using the order statistics (OS) from a large set of discrete time samples. Then the quantiles of the CDF can be constructed using OS, which is a consistent estimator of the distribution [34]:

$$\hat{Q}_{y_i}\left(\frac{k}{T}\right) = y_{i(k)} \Leftrightarrow \hat{F}\left(y_{i(k)}\right) = \frac{k}{T} \tag{14}$$

As a result, the estimation of NG measure using the OS can be expressed as:

$$\hat{D}(Q_{y_i}, Q_g) = \frac{1}{T} \left(\sum_{k=1}^{T} \left[y_{i(k)} - Q_g\left(\frac{k}{T}\right) \right]^2 \right)^{1/2} \tag{15}$$

where $Q_g(k/T)$ is the k/T quantile of the equivalent Gaussian distribution.

4 Nonparametric NG Algorithm for Dependent Source Signals

Conclusion 2. The nonparametric NG measure $\hat{D}(Q_{y_i}, Q_g)$ will reach a local maximum at any output channel for each component if $\mathbf{b}_i$ is forced to be unitary [36].

To extract a different source signal at each output channel, a multistage deflation procedure must be applied to the separation system. The NG measure $\hat{D}(Q_{y_i}, Q_g)$ is maximized at each output channel successively under the constriction that the vector $\mathbf{b}_i$ has to be orthonormal to the previously obtained vectors, the separation matrix is composed of all the vectors $\mathbf{b}_i$. Taking into account Eq. (4) in vector form:

$$y_i(t) = \mathbf{b}_i^{\mathrm{T}} \mathbf{z}(t) \tag{16}$$

where $\mathbf{b}_i^{\mathrm{T}}$ is the i-row of the separation matrix $\mathbf{B}$, the goal is to update $\mathbf{b}_i$ at each stage by optimizing a cost function $J(\mathbf{b}_i)$. We take the objective function $J(\mathbf{b}_i)$ as:

$$J(\mathbf{b}_i) = D(Q_{y_i}, Q_g) \tag{17}$$

$J(\mathbf{b}_i)$ will be optimized by the stochastic gradient rule of the constrained optimization method with the constrains [36]:

$$\begin{cases} \mathbf{b}_i(k+1) = \mathbf{b}_i(k) + \mu\nabla J|_{\mathbf{b}_i(k)} \\ s.t.\ \mathbf{b}_i \text{ is orthonormal to } \{\mathbf{b}_1, \cdots, \mathbf{b}_{i-1}\} \end{cases} \tag{18}$$

Denote $Y_i(k) = y_{i(k)} - Q_g(k/T)$, then the gradient of $J(\mathbf{b}_i)$ in Eq. (15) is:

$$\nabla J|_{\mathbf{b}_i(k)} = \frac{1}{2T}\left(\sum_{t=1}^{T}[Y_i(t)]^2\right)^{-1/2} \frac{d\left(\sum_{t=1}^{T}[Y_i(t)]^2\right)}{d\mathbf{b}_i}\Bigg|_{\mathbf{b}_i(k)}$$

$$= \frac{1}{T}\left(\sum_{t=1}^{T}[Y_i(t)]^2\right)^{-1/2}\left(\sum_{t=1}^{T}Y_i(t)\right)\mathbf{z}\frac{dy_{i(t)}}{dy_i}\Bigg|_{\mathbf{b}_i(k)} \tag{19}$$

where $\frac{dy_{i(t)}}{dy_i}\Big|_{\mathbf{b}_i(k)} = \mathbf{e}_t = [0,0,\cdots,0,1,0,\cdots,0]^{\mathrm{T}}$ and $\mathbf{e}_t(l) = \begin{cases} 1 & \text{if } y_i(l) = y_{i(t)} \\ 0 & \text{else} \end{cases}\Bigg|_{t=1,\cdots,T}$.

After the ith source signal is extracted, $\mathbf{b}_i$ must be normalized and projected over the subspace orthonormal $\mathbf{C}_{i-1}$ to the vectors obtained at every previous stage. Let us quote that the $\mathbf{C}_{i-1}$ expression is

$$\mathbf{C}_{i-1} = \mathbf{I} - (\mathbf{B}_{i-1}\mathbf{B}_{i-1}^{\mathrm{T}})^{-1}\mathbf{B}_{i-1}^{\mathrm{T}} \tag{20}$$

where $\mathbf{B}_{i-1} = (\mathbf{b}_1, \cdots, \mathbf{b}_{i-1})$.

5   Computer Simulations

In order to show the performance and the validity of the proposed algorithm, simulations using Matlab are given below. The simulation results presented in this section are divided into four Examples. The statistical performance, or accuracy, was measured by the index signal-to-interference ratio (SIR) as,

$$\mathrm{SIR}(s_i, y_j) = 10\log\frac{\sum_{t=1}^{T}(s_i(t))^2}{\sum_{t=1}^{T}(|s_i(t)| - |y_j(t)|)^2}, \quad i,j = 1,\cdots,N$$

where $y_i(t)$ is the estimation of $s_i(t)$.

5.1 Simulations on Determined BSS Case

In this simulation, we use $N = 4$ source signals which are extracted from the real world photo. It should be noted that the source signals are extracted from different pixel columns of real world images and then stack these columns one by one to get a one dimensional source signals. By selecting different intervals between the columns of the image, we can control the level of dependence between the source signals. We choose the columns of the photo which are relatively far away, therefore they are mutually weak correlated. The source' correlation coefficients are shown in Table 1.

Table 1. The correlation coefficients between different source signals

	Source 1	Source 2	Source 3	Source 4
Source 1	1.0000	0.1396	−0.0305	−0.2182
Source 2	0.1396	1.0000	−0.0987	−0.0074
Source 3	−0.0305	−0.0987	1.0000	0.1967
Source 4	−0.2182	−0.0074	0.1967	1.0000

The input mixed signals of the algorithm are generated by mixing the four source signals with a 4×4 random mixing matrix in which the elements are distributed with $N(0, 1)$. After convergence of the proposed algorithm, the average results of the performance criteria evaluated by SIR over 10 experiments are shown in Table 2.

Table 2. Average SIR for different source signals using the proposed algorithm over 10 experiments

Performance index	Source signals			
	s_1	s_2	s_3	s_4
SIR	24.07	21.95	14.98	27.86

5.2 Simulations on Underdetermined BSE Case

In this example, we use 4 source signals which are extracted from the same real world photo as example 1. The 3*4 mixing matrix are generated by the randn function of Matlab. The source' correlation coefficients are shown in Table 3.

Table 3. The correlation coefficients between different source signals

	Source 1	Source 2	Source 3	Source 4
Source 1	1.0000	0.1923	0.0473	−0.1803
Source 2	0.1923	1.0000	0.1298	−0.1571
Source 3	0.0473	0.1298	1.0000	−0.1184
Source 4	−0.1803	−0.1571	−0.1184	1.0000

After convergence of the proposed algorithm, the average results of the performance criteria evaluated by SIR over 10 experiments are shown in Table 4. The correlation coefficients of source signals and their corresponding extracted signals using the proposed algorithm are shown in Table 5.

Table 4. Average SIR for different source signals using the proposed algorithm over 10 experiments

Performance index	Source signals		
	s_2	s_3	s_4
SIR	16.43	19.90	21.83

Table 5. The correlation coefficients of source signals and their corresponding extracted signals using the proposed algorithm

Extracted signals using the Proposed Algorithms	Source signals			
	s_1	s_2	s_3	s_4
y_1	0.1765	0.1232	0.0411	**−0.9967**
y_2	−0.2711	**−0.9154**	−0.0798	0.2098
y_3	−0.1476	0.2224	**0.9392**	−0.2493

5.3   Simulations on Effect of Strong Correlations

In order to verify the performance of the proposed algorithm for strong correlations between source signals, we choose four face images extracted from face databases of [37] as the source signals. The source' correlation coefficients are shown in Table 6. The input mixed signals of the algorithm are generated by mixing the four source signals with a 4×4 random mixing matrix in which the elements are distributed with $N(0, 1)$. After convergence of the proposed algorithm, the average results of the performance criteria evaluated by SIR over 10 experiments are shown in Table 7. The definition of SIR for images can be found in [38].

Table 6. The correlation coefficients between four different face images

	Source 1	Source 2	Source 3	Source 4
Source 1	1.0000	0.9106	0.8421	0.8388
Source 2	0.9106	1.0000	0.8379	0.8363
Source 3	0.8421	0.8379	1.0000	0.9279
Source 4	0.8388	0.8363	0.9279	1.0000

Table 7. Average SIR for different source signals using the proposed algorithm over 10 experiments

Performance index	Source signals			
	s_1	s_2	s_3	s_4
SIR	62.22	54.48	52.87	54.96

5.4   Simulations on Comparison with Other Algorithms

For comparison, we display the this simulation, at the same convergent conditions, the proposed algorithm was compared along two sets of criteria, statistical and computational with other popular BSS algorithms such as FastICA (the nonlinearity function to be chosen as y^3), COMBI, SOBI and JADEop algorithm [39]. The computational load was measured as CPU time needed for convergence (using Matlab R2010b and run in the CPU 3.0 GHz Pentium4 computer). The four source signals and mixed signals are the same as Sect. 5.1.

The statistical performance, or accuracy, was measured using an alternative popular BSS performance index called cross-talking error index E defined as [2]. The separation results of the 4 different sources are shown in Table 8 for various BSS algorithms (averaged over 100 Monte Carlo simulations).

From Table 8 we conclude that the proposed algorithm can make an ideal separation results for statistically dependent source signals, the BCA algorithm can also get ideal separation results for these source signals, the other four popular BSS algorithms can not all work well in this condition, but when the boundaries of the source signals are not satisfied properly, the proposed algorithm can also works well, that is to say our proposed method has a further wide field of applications. Then we will discuss the computational load for convergence, as a nonparametric method, the proposed algorithm requires more computation than all the other algorithms, but it has a better convergent performance, moreover the nonparametric estimation method has the robust properties, so from the convergent performance and computation load two aspects, the proposed algorithm can work in the DCA situation.

Table 8. The separation results of various BSS algorithms

Different Algorithms	FastICA	COMBI	SOBI	JADEop	BCA	Proposed
Performance index E	0.1999	0.1913	0.3972	0.3218	0.1671	0.1677
Computation time (s)	0.22	0.32	0.09	0.08	0.1374	0.96

6   Conclusions

Most of state-of-the-art algorithms for solving BSS or BSE problem rely on independence or at least second order statistical assumption of the source signals. In this paper, we developed a nonparametric BSE algorithm for statistical dependent source signals using the nonparametric NG measure. We show that maximization of the NG measure can not only separate the statistically independent but also dependent source signals, even in the underdetermined BSE situation. The NG measure is defined by statistical distances between distributions based on the CDF instead of traditional PDF which can be realized by the order statistics efficiently. Simulation results on both synthetic and real world data show that the proposed nonparametric algorithm is able to extract the dependent source signals and yield ideal performance. The next purpose of this study is utilizing the proposed method to extract ERP signals in the brain-computer interfaces (BCI) applications.

References

1. Comon, P., Jutten, C.: Handbook of Blind Source Separation: Independent Component Analysis and Applications. Elsevier, Oxford (2010)
2. Cichocki, A., Amari, S.: Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. Wiley, New York (2003)
3. Hyvarinen, A.: Independent component analysis: recent advances. Philos. Trans. R. Soc. A **371**, 20110534 (2013)
4. Cardoso, J.: Blind signal separation: statistical principles. Proc. IEEE **86**(10), 2009–2025 (1998)
5. Särelä, J., Valpola, H.: Denoising source separation. J. Mach. Learn. Res. **6**, 233–272 (2005)
6. Leong, W., Mandic, D.: Noisy component extraction (NoiCE). IEEE Trans. Circuits Syst. I. **57**(3), 664–671 (2010)
7. Deville, Y., Hosseini, S.: Recurrent networks for separating extractable-target nonlinear mixtures. Part I: Non-blind Configurations. Signal Process. **89**(4), 378–393 (2009)
8. Bell, A.J., Sejnowski, T.J.: An information-maximisation approach to blind separation and blind deconvolution. Neural Comput. **7**(6), 1129–1159 (1995)
9. Amari, S., Cichocki, A., Yang, H.: A new learning algorithm for blind signal separation. In: Advances in Neural Information Processing Systems, pp. 757−763. MIT Press, Cambridge (1996)
10. Bloemendal, B., Laar, J., Sommen, P.: A single stage approach to blind source extraction based on second order statistics. Signal Process. **93**(2), 432–444 (2013)
11. Cardoso, J.F.: Multidimensional independent component analysis. In: ICASSP 1998, Seattle, WA, USA, pp. 1941–1944. IEEE (1998)
12. Lahat, D., Cardoso, J.F., Messer, H.: Second-order multidimensional ica: performance analysis. IEEE Trans. Signal Process. **60**(9), 4598–4610 (2012)
13. Gutch, H.W., Theis, F.J.: Uniqueness of linear factorizations into independent subspaces. J. Multivar. Anal. **112**, 48–62 (2012)
14. Kawanabe, M., Muller, K.R.: Estimating functions for blind separation when sources have variance dependencies. J. Mach. Learn. Res. **6**, 453–482 (2005)
15. Hyvarinen, A., Hoyer, P.O., Inki, M.: Topographic independent component analysis. Neural Comput. **13**(7), 1527–1558 (2001)
16. Bach, F.R., Jordan, M.I.: Kernel independent component analysis. J. Mach. Learn. Res. **3**, 1–48 (2002)
17. Zhang, K., Chan, L.W.: An adaptive method for subband decomposition ICA. Neural Comput. **18**(1), 191–223 (2006)
18. Wang, F.S., Li, H., Li, R.: Novel nongaussianity measure based bss algorithm for dependent signals. In: Dong, G., Lin, X., Wang, W., Yang, Y., Yu, J.X. (eds.) APWeb/WAIM 2007. LNCS, vol. 4505, pp. 837–844. Springer, Heidelberg (2007)
19. Caiafa, C.: On the conditions for valid objective functions in blind separation of independent and dependent sources. EURASIP J. Adv. Signal Process. **2012**, 255 (2012)
20. Aghabozorgi, M.R., Doost-Hoseini, A.M.: Blind separation of jointly stationary correlated sources. Signal Process. **84**(2), 317–325 (2004)

21. Abrard, F., Deville, Y.: A time-frequency blind signal separation method applicable to underdetermined mixtures of dependent sources. Signal Process. **85**(7), 1389–1403 (2005)
22. Kopriva, I., Jeric, I., Brkljacic, L.: Nonlinear mixture-wise expansion approach to underdetermined blind separation of nonnegative dependent sources. J. Chemom. **27**(1), 189–197 (2013)
23. Cruces, S.: Bounded component analysis of linear mixtures: a criterion for minimum convex perimeter. IEEE Trans. Signal Process. **58**(4), 2141–2154 (2010)
24. Erdogan, A.T.: A class of bounded component analysis algorithms for the separation of both independent and dependent sources. IEEE Trans. Signal Process. **61**(22), 5730–5743 (2013)
25. Li, Y., Amari, S.I., Cichocki, A.: Underdetermined blind source separation based on sparse representation. IEEE Trans. Signal Process. **54**(2), 423–437 (2006)
26. Almeida, A., Luciani, X., Stegeman, A., Comon, P.: CONFAC decomposition approach to blind identification of underdetermined mixtures based on generating function derivatives. IEEE Trans. Signal Process. **60**(11), 5698–5713 (2012)
27. Cardoso, J.F.: Dependence, correlation and gaussianity in independent component analysis. J. Mach. Learn. Res. **4**, 1177–1203 (2003)
28. Cichocki, A., Thawonmas, R.: On-line algorithm for blind signal extraction of arbitrarily distributed, but temporally correlated sources using second order statistics. Neural Process. Lett. **12**(1), 91–98 (2000)
29. Barros, A.K., Cichocki, A.: Extraction of specific signals with temporal structure. Neural Comput. **13**(9), 1995–2003 (2001)
30. Anderson, M., Adali, T., Li, X.L.: Joint blind source separation with multivariate gaussian model: algorithms and performance analysis. IEEE Trans. Signal Process. **60**(4), 1672–1683 (2012)
31. Zibulevsky, M., Zeevi, Y.Y.: Extraction of a source from multichannel data using sparse decomposition. Neurocomputing **49**(1–4), 163–173 (2002)
32. Zhang, Z.L.: Morphologically constrained ICA for extracting weak temporally correlated signals. Neurocomputing **71**, 1669–1679 (2008)
33. Caiafa, C.F., Proto, A.N.: Separation of statistically dependent sources using an l^2 distance non-gaussianity measure. Signal Process. **86**(11), 3404–3420 (2006)
34. Hyvarinen, A.: Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Networks **10**(3), 626–634 (1999)
35. Friedman, J.H.: Exploratory projection pursuit. J. Am. Stat. Assoc. **82**(397), 249–266 (1987)
36. Blanco, Y., Zazo, S.: New gaussianity measures based on order statistics: application to ica. Neurocomputing. **51**, 303–320 (2003)
37. Psychology Department of University of Stirling. http://pics.psych.stir.ac.uk/
38. Vincent, E., Araki, S., Theis, F.J., Nolte, G., Bofill, P., et al.: The signal separation evaluation campaign (2007-2010): achievements and remaining challenges. Sig. Process. **92**, 1928–1936 (2012)
39. Cichocki, A., Amari, S., Siwek, K.: ICALAB toolboxes. http://www.bsp.brain.riken.jp/ICALAB