

Received March 19, 2017, accepted May 8, 2017, date of publication May 11, 2017, date of current version June 28, 2017. *Digital Object Identifier 10.1109/ACCESS.2017.2703840*

Stationary Points of a Kurtosis Maximization Criterion for Noisy Blind Source Extraction

SHANSHAN LU, WEI WANG, AND GUOYU WANG

College of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China

Corresponding author: Shanshan Lu (sslu2004@sina.com)

This work was supported by the National Natural Science Foundation of China under Grant 61571451 and Grant 61501487.

ABSTRACT Blind source extraction (BSE) is often posed as the maximization of a statistical criterion under a unitary constraint. This paper addresses the convergence problem of a kurtosis-based criterion in the presence of noise. We present the stationary points of such criterion, and show that these extrema are simplified as the minimum mean square error (MMSE) solutions with some approximations. Moreover, we introduce a robust preprocessing approach, which allows one to find the MMSE separation matrix up to an orthogonal factor. The excellent performance of the BSE algorithm based on this preprocessing approach shows that the analysis of the stationary point is reliable.

INDEX TERMS Blind source extraction (BSE), blind source separation (BSS), kurtosis maximization, Gaussian noise.

I. INTRODUCTION

Blind source separation (BSS) has received considerable attention because of its applicability to various fields, such as communication signal analysis, image processing, texture modelling, and others [1]–[4]. Blind source extraction (BSE) is a special class of BSS methods, which extracts one source or a selected number of the sources at a time [5]. A large number of BSE algorithms have been proposed over the past decades that are mostly derived from search methods aiming to maximize statistical criteria that typically reflect some known structural properties of the transmitted signals [6], [7].

The most widely used statistical criteria in the BSE are the kurtosis-based ones. Then, an important question is whether the stationary points of such a criterion will yield an ''approximate'' extraction. The existing studies of stationary points are usually proposed under the assumption of no noise. In such cases, a convergence proof was provided by Ding and Nguyen in [8], which shows that all extrema will extract a single source while rejecting all the interference. However, the bias caused by the noise, which may be present in many applications, will affect the convergence behavior of the systems.

The BSE algorithms based on the kurtosis-type criteria are known to provide good results. However, the stationary points in the presence of noise are not as clearly understood as the noiseless alternates. This paper addresses this difficult problem. We introduce the Lagrange multiplier approach to present the stationary points of such criterion. Moreover, we show that these extrema are simplified as the minimum mean square error (MMSE) solutions with some approximations.

Many of the algorithms that have been proposed for the noisy data produce the zero-forcing solution [9], [10]. However, such solution does not correspond to the MMSE criterion. Inspired by the above study of the stationary points, we employ a robust preprocessing approach in [11], which allows one to find the MMSE separation matrix up to an orthogonal or unitary factor. The simulations show that the algorithm based on the robust preprocessing step provides excellent performance. This result supports our analysis of the stationary points.

II. PROBLEM FORMULATION

Consider the typical model where the observed mixtures are expressed as a linear mixing of independent sources. Denoting the observed mixtures by $x(k) = [x_1(k)]$, $x_2(k), \ldots, x_m(k)$ ^T, then, the signal model is described by

$$
x(k) = As(k) + n(k)
$$
 (1)

where $s(k) = [s_1(k), s_2(k), \dots, s_n(k)]^T$ is the source vector, $A=[a_1, a_2, \ldots, a_n]$ is an unknown mixing matrix, and $n(k)$ is a Gaussian noise vector. Moreover, without loss of generality, we assume that the sources have a zero mean and unit variance.

To retrieve one of the source signals, we apply an extraction operation given by the demixing vector *w* to the

mixtures $x[k]$, which yields the recovered signal $y(k)$, given by

$$
y(k) = w^H x(k) = \sum_{i=1}^n g_i s_i(k) + w^H n(k)
$$
 (2)

where $g_i = w^H a_i$ is a coefficient of the source s_i . Assuming that the sources are second order circular signal, then the kurtosis of a random signal *y*[*k*] is defined as [12]

$$
kt(y(k)) = E\{|y(k)|^4\} - 2E^2\{|y(k)|^2\}
$$
 (3)

where $E\{\cdot\}$ denotes the statistical expectation. In addition, it is noted that the kurtosis-based BSE methods cannot extract the source signal with zero kurtosis. Therefore, we assume that the sources of interest always have non-zero kurtosis.

III. STATIONARY POINTS OF A KURTOSIS MAXIMIZATION CRITERION

In the BSE, the kurtosis-based technology is one of the most common methods. The kurtosis-type criterion usually requests that the output variance is bounded. Considering this constraint, we obtain the optimization problem described by

max
$$
J_1(w) = \text{kt}(y(k))
$$
,
subject to $J_2(w) = E\{|y(k)|^2\} = 1$. (4)

The kurtosis of the signal $y(k)$, can be expressed as the sum of the individual kurtosis by considering that the kurtosis of a Gaussian signal is zero and the properties of kurtosis for linear mixing systems, and can be expressed as

$$
J_1(w) = \text{kt}(y(k)) = \sum_{i=1}^n |g_i|^4 \text{kt}(s_i(k))
$$
 (5)

where $kt(s_i(k))$ denotes the kurtosis of $s_i(k)$. Otherwise, if the objective source has a negative kurtosis, we simply change the sign of J_1 . Therefore, without loss of generality, we assume that the objective source has a positive kurtosis.

Our goal is to obtain the stationary points of the constrained optimization problem described in (4). To this end, we first introduce the Lagrange multiplier method to transform it into an unconstrained optimization problem

$$
J(\mathbf{w}, \lambda) = \sum_{i=1}^{n} |g_i|^4 \text{kt}(s_i(k)) + \lambda (1 - \mathbf{w}^H \mathbf{R}_{\mathbf{x}} \mathbf{w}) \qquad (6)
$$

where \mathbf{R}_x is the covariance matrix of $\mathbf{x}(k)$, and λ is a Lagrange multiplier. It is well known that a stationary point is located where the gradient of the function equals to zero

$$
\nabla_{\mathbf{w}} J(\mathbf{w}, \lambda) = \sum_{i=1}^{n} 2\text{kt}(s_i(k)) |g_i|^2 \mathbf{a}_i \mathbf{a}_i^H \mathbf{w} - \lambda \mathbf{R}_{\mathbf{x}} \mathbf{w} = 0. \quad (7)
$$

To obtain the solution of this equation, we take the variable $2kt(s_h(k))|g_h|^2 a_h a_h^H w$ on the one side of the equal sign, and the rest on the other side, Then, we obtain the solution that makes the equation tenable, and which is given by

$$
\mathbf{w}_{hopt} = b_h(\mathbf{R}_{\mathbf{x}} - \sum_{i=1, i \neq h}^{n} \alpha_i \mathbf{a}_i \mathbf{a}_i^H)^{-1} \mathbf{a}_h \tag{8}
$$

where b_h ($h = 1, 2, ..., n$) is a non-zero coefficient that will not influence our conclusion, and $\alpha_i = 2kt(s_i(k))|g_i|^2/\lambda$ is a coefficient depending on the vector *w*. The subscript *h* indicates that this stationary point extracts the *h*-th source signal.

The existing study have shown that when a signal is received in the presence of noises with zero kurtosis (such as Gaussian noise), the kurtosis maximum algorithm (KMA) will converge to the capturing of the signal with MMSE setting [8]. From the solution in (8), we have that our result supports such conclusion, but has a more general significance.

In order to further analyze the stationary points of the model, we expand the expression of the solution in (8), and simplify it with some suitable approximations. According to the property of matrix inversion, we have formula as follows

$$
(\mathbf{R}_{\mathbf{x}} - \sum_{i=1, i \neq h}^{n} \alpha_{i} \mathbf{a}_{i} \mathbf{a}_{i}^{H})^{-1} \mathbf{a}_{h}
$$

= $\mathbf{R}_{x}^{-1} \mathbf{a}_{h} + \sum_{i=1, i \neq h}^{n} \sum_{j=1, j \neq h}^{n} c_{ij} d_{jh} \mathbf{R}_{x}^{-1} \mathbf{a}_{i}$ (9)

where $d_{jh} = a_j^H R_x^{-1} a_h$ and c_{ij} are scalar coefficients. The scalar d_{jh} denoting the interference comes from the *h*-th source signal when extracting the *j*-th source under the MMSE framework is typically a relatively small quantity. The coefficients c_{ij} often contains the factors d_{ij} ($i \neq j$), which makes the second part on the RHS of (9) contain some second order of the small quantity. Neglecting this small quantity, we obtain the pre-simplification of the stationary point in (8), given by

$$
(\boldsymbol{R}_{\boldsymbol{x}} - \sum_{i=1, i \neq h}^{n} \alpha_i \boldsymbol{a}_i \boldsymbol{a}_i^H)^{-1} \boldsymbol{a}_h \approx \boldsymbol{R}_{\boldsymbol{x}}^{-1} \boldsymbol{a}_h + \sum_{i=1, i \neq h}^{n} \alpha_i \beta_i d_{ih} \boldsymbol{R}_{\boldsymbol{x}}^{-1} \boldsymbol{a}_i
$$
\n(10)

where $\beta_i = 1/(1 - \alpha_i d_{ii})$ is a scalar parameter. In order to intuitively illustrate the above approximation process, we employ a simple case as the example, in which the source's number *n* is set as 3. Then, we have the following equation

$$
(\mathbf{R}_{\mathbf{x}} - \alpha_i \mathbf{a}_i \mathbf{a}_i^H - \alpha_j \mathbf{a}_j \mathbf{a}_j^H)^{-1} \mathbf{a}_h
$$

\n
$$
= \mathbf{R}_{\mathbf{x}}^{-1} (\mathbf{a}_h + \alpha_i \beta_i d_{ih} \mathbf{a}_i + \alpha_i \alpha_j \beta_i \beta_j d_{ij} d_{jh} \mathbf{a}_i + \alpha_i^2 \alpha_j \beta_i^2 \beta_{ij} d_{ij} d_{ih} \mathbf{a}_i + \alpha_i \alpha_j \beta_i \beta_j d_{ij} d_{jh} \mathbf{a}_i + \alpha_i \alpha_j^2 \beta_j^2 \beta_{ji} d_{ij} d_{ij} \mathbf{a}_j + \alpha_i \alpha_j \beta_i \beta_j d_{ji} d_{ih} \mathbf{a}_j)
$$
(11)

where $\beta_{ij} = 1/(1 - \alpha_j d_{jj} - \alpha_i \alpha_j \beta_i d_{ij} d_{ji})$ is a scalar coefficient. We find that the result in (11) is consistent with the above analysis, and that can be approximated as the expression in (10). It should be noted that a similar example is achieved for more source signals but is not presented here.

Substituting the pre-simplification described in (10) into α_i and neglecting the above 3-order terms of d_{ii} ($i \neq j$), we find that α_i is the second order of d_{ij} ($i \neq j$), which makes the second part on the RHS of (10) a third order of the small quantity

 d_{ii} ($i \neq j$) that can be neglected. Therefore, the stationary point given by (8) can be further simplified as

$$
w_{hopt} \approx R_x^{-1} a_h \tag{12}
$$

which is the MMSE solution of the system. Therefore, it can be concluded that, if the mutual interference is small enough under the MMSE framework, the stationary points of the constrained optimization problem described in (4) would be approximately equals to the MMSE solutions.

IV. ROBUST-FASTICA ALGORITHM

Some algorithms for blind extraction require prewhitening of mixed signals. After prewhitening, the BSE problems usually become somewhat easier and well posed. The additive noise needs to be taken into account in many practices. And then the quasi-whitening process is developed in the literatures. Before presenting this preprocessing, we first review the eigenvalue decomposition of the covariance matrix R_x

$$
\mathbf{R}_{x} = E\{\mathbf{x}(k)\mathbf{x}^{H}(k)\} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{H} \tag{13}
$$

where $\Lambda = \text{diag}\{\lambda_1 > \lambda_2...\ > \lambda_m\}$ contains the *m* eigenvalues, and $U = [u_1, u_2, \dots, u_m]$ is the eigenvectors matrix. It is well known that the matrix U can be divided into $U_s = [u_1, u_2, \dots, u_n]$ contains the eigenvectors that correspond to the *n* principal eigenvalues of Λ_s = $diag{\lambda_1 > \lambda_2 \ldots > \lambda_n}$, and $U_n = [u_{n+1}, \ldots, u_m]$ contains the *m*-*n* eigenvectors. Then, the preprocessing matrix that is applied in the quasi-whitening processing is given by

$$
\boldsymbol{P}_{qua} = (\boldsymbol{\Lambda}_s - \sigma^2 \boldsymbol{I})^{-1/2} \boldsymbol{U}_s^H
$$
 (14)

where σ^2 is the noise variance. This preprocessing allows one to find the zero-forcing separation solution up to an orthogonal factor. It well known that such solution is able to suppress the mutual interference of source signals.

The analysis of the stationary points shows that the kurtosis- based BSE algorithms approximately converge to the MMSE separation solutions. Thus, the quasi-whitening preprocessing, which is proposed for the zero-forcing separation solution, is not suitable for such a condition. In this section, we employ the robust preprocessing proposed in [11], which can overcome the above problem. The preprocessing matrix that is applied in the robust preprocessing is given by

$$
\boldsymbol{P}_{rob} = (\boldsymbol{\Lambda}_s - \sigma^2 \boldsymbol{I})^{1/2} \boldsymbol{\Lambda}_s^{-1} \boldsymbol{U}_s^H. \tag{15}
$$

The mixed signal after robust preprocessing is $z(k)$ = $P_{rob}x(k)$. From [11], we have that the only difference between the MMSE separation solution and the robust preprocessing matrix is an orthogonal transformation. This condition means that one can find the MMSE solution up to an orthogonal factor. The MMSE solution can not only suppress the mutual interference of source signals, but also reduce the effect of noise.

The following BSE method is to design a demixing vector *w* such that $y(k)=w^Hz(k)$ can recover one of the source signals. Our starting point is a Newton's approach based

on the Lagrangian function. The second term on the RHS of (3) is ignored owing to the unity-variance constraint. Then, the Lagrange function for the optimization problem in (4) is given by

$$
J(\mathbf{w}, \lambda) = J_3(\mathbf{w}) + \lambda (1 - J_2(\mathbf{w})) \tag{16}
$$

where $J_3 = E\{|y(k)|^4\}$ is the simplified contrast function, J_2 is the constraint in (4), and λ is a Lagrange multiplier. The Newton update to this Lagrangian is given by [13]

$$
\Delta w = -(H_w J_3 + \lambda R_z)^{-1} (\nabla_w J_3 + \lambda R_z w) \tag{17}
$$

where $\nabla_w J_3$ and $\mathbf{H}_w J_3$ are the gradient and the Hessian of the function J_3 , respectively. Using the result of [12], we obtain that the gradient is $\nabla_w J_3 = 2E\{|y(k)|^2 y(k)^* z(k)\}\$, and the Hessian is $\mathbf{H}_w J_3 = 4E\{|y(k)|^2\}R_z$. Substituting these results into (17) and simplifying, we obtain the fixed-point update as

$$
w^{+} = 2E\{|y(k)|^{2}\}w + R_{z}^{-1}E\{|y(k)|^{2}y(k)^{*}z(k)\}\newline w^{+} = w^{+}/((w^{+})^{H}R_{z}w^{+})
$$
\n(18)

Note that if the preprocessing is the whitening (that is $R_z = I$), this update is exactly the method proposed in [12]. This update can be extended to the estimation of all the source signals. To prevent different vectors from converging to the same ones, the demixing vector *w* should be decorrelated after every update. One way to accomplish this is given by [12]

$$
\mathbf{w}_{p+1} = \mathbf{w}_{p+1} - \sum_{j=1}^{p} \mathbf{w}_j \mathbf{w}_j^H \mathbf{w}_{p+1}
$$
 (19)

After this deflation scheme, we obtain the orthogonal demixing matrix $W = [w_1, w_2, \dots, w_n]$. Combining the robust preprocessing and this fast fixed-point algorithm (FastICA), we then obtain the robust-FastICA method used in this paper.

V. SIMULATIONS

In this section, some simulation examples are carried out to illustrate the validity of the stationary point analysis. We take the communication signal processing under consideration. An eight-element uniform linear array is employed. The distance is set to half of the carrier wavelength of the sources. Each of the sensor signals is polluted by an additive noise.

In the first example, the source signals are assumed as four QPSK modulated signals. These independent sources have a zero mean and a unit variance with the length of 8000 samples each. The directions of arrival of the sources were set $\hat{\theta}_1 = 0^\circ$, $\theta_2 = 10^\circ$, $\theta_3 = 20^\circ$, and $\theta_4 = 30^\circ$. Fig. 1 shows the beam patterns of the MMSE extractor and our robust-FastICA algorithm when the input SNR is fixed at 5 dB.

From Fig. 1, we find that the beam pattern of the proposed algorithm is similar to that of the MMSE extractor. This result is consistent with our analytical result that the stationary points of the optimization problem described in (4) are approximately equal to the MMSE extraction solutions. It should be noted that only the beam pattern for the third

FIGURE 1. The beam pattern of the MMSE extractor and the robust-FastICA algorithm.

source signal is presented in this example. Similar performance is achieved for extracting the other sources but is not presented here.

In the second example, we compared the performance of the robust-FastICA algorithm and the KMG algorithm based on the quasi-whitening proposed in [14]. To evaluate the performance of the algorithms, we take the signalto-interference-plus-noise ratio (SINR*i*) as the performance index defined as

$$
\text{SINR}i = |g_{ik}|^2 / (\sum_{j=1}^n |g_{ij}|^2 + \sigma^2 \sum_{j=1}^m |s_{ij}|^2 - |g_{ik}|^2) \tag{20}
$$

where g_{ij} is the *ij*-th element of the matrix $G=WPA$, s_{ij} is the *ij*-th element of the matrix $S=WP$, and $g_{ik} = \max\{g_{i1},\}$ g_{i2}, \ldots, g_{in} . This criterion indicates the degree of inhibition of the objective signal to interference signal plus noise, and is suggested as the most meaningful performance criterion.

This simulation example employs the same parameters as those used in example 1, except that the source signals arrive at the sensors from the angles $\theta_1 = -6^\circ$, $\theta_2 = 0^\circ$, $\theta_3 = 9^\circ$, and $\theta_4 = 16^\circ$. Fig. 2 presents the SINR*i* of the separated signals after the convergence of the algorithms. Note that the MMSE-bound in Fig. 2 shows the upper limit of the performance corresponding to the MMSE separation matrix.

From Fig. 2, we find that the separation performance of the robust-FastICA algorithm approaches the MMSE-bound for all the four output signals. However, the performance of the KMG method only approaches the MMSE-bound for the first output signals. Note that the criterion employed in the KMG algorithm is also the kurtosis maximization ones. For the rest of the output signals, the performance degradation of the KMG algorithm is due to the bias from the quasiwhitening step. The excellent performance of the robust-FastICA algorithm shows that the analysis of the stationary point is reliable.

To further illustrate the performance of the algorithms, the input SNR is set as 0-15 dB, and 16 experiments are conducted under different SNRs, where 100 Monte Carlo

FIGURE 2. SINRi of 4 separated signals using the MMSE extractor, the robust-FastICA algorithm, and the KMG algorithm.

FIGURE 3. Average SINR over 100 independent runs for the MMSE extractor, the robust-FastICA algorithm, and the KMG algorithm.

experiments are performed for each SNR case. The averages, SINR (the mean of SINR*i*), are presented in Fig. 3.

From Fig. 3, we find that the robust-FastICA algorithm has an improved separation performance than the contrast method, and closely approaches the MMSE-bound under all considered SNRs. This simulation demonstrates the superior performance of the robust-FastICA algorithm. This result also shows the reliability of the stationary point analysis.

VI. CONCLUSION

In this paper, we have provided a convergence analysis for the kurtosis maximization criterion in the presence of Gaussian noise. We have shown that if the mutual interference is small enough under the MMSE framework, the stationary points of such criterion can be simplified as the MMSE solutions with some suitable approximations. Moreover, we have proposed a FastICA method based on robust preprocessing to illustrate the correctness of the convergence analysis. The simulation results support our theoretical analysis.

REFERENCES

[1] M. Novey and T. Adali, "Complex ICA by negentropy maximization,'' *IEEE Trans. Neural Netw.*, vol. 19, no. 4, pp. 596–609, Apr. 2008.

- [2] Z. Shi and C. Zhang, ''Blind source extraction using generalized autocorrelations,'' *IEEE Trans. Neural Netw.*, vol. 18, no. 5, pp. 1516–1524, Sep. 2007.
- [3] T. Adalı, H. Li, M. Novey, and J. Cardoso, ''Complex ICA using nonlinear functions,'' *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4536–4544, Sep. 2008.
- [4] A. T. Erdogan, ''On the convergence of ICA algorithms with symmetric orthogonalization,'' *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2209–2221, Jun. 2009.
- [5] W. Liu, D. P. Mandic, and A. Cichocki, ''Blind second-order source extraction of instantaneous noisy mixtures,'' *IEEE Trans. Circuits Syst. II, Express Briefs*, vol. 53, no. 9, pp. 931–935, Sep. 2006.
- [6] C. B. Papadias, "Globally convergent blind source separation based on a multiuser kurtosis maximization criterion,'' *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3508–3519, Dec. 2000.
- [7] M. Castella and E. Moreau, ''New kurtosis optimization schemes for MISO equalization,'' *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1319–1330, Mar. 2012.
- [8] Z. Ding and T. Nguyen, ''Stationary points of a kurtosis maximization algorithm for blind signal separation and antenna beamforming,'' *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1587–1596, Jun. 2000.
- [9] W. Liu and D. P. Mandic, ''A normalised kurtosis-based algorithm for blind source extraction from noisy measurements,'' *Signal Process.*, vol. 86, no. 7, pp. 1580–1585, Jul. 2006.
- [10] S. C. Douglas, "Combined subspace tracking, prewhitening, and contrast optimization for noisy blind signal separation,'' in *Proc. ICA*, Helsinki, Finland, 2000, pp. 585–590.
- [11] S.-S. Lu, W. Wang, Z.-Y. Gu, and G.-Y. Wang, "A robust preprocessing method for noisy blind source separation,'' in *Proc. PIERS*, Shanghai, China, Aug. 2016, pp. 1887–1893.
- [12] E. Bingham and A. Hyvärinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals,'' *Int. J. Neural Syst.*, vol. 10, no. 1, pp. 1–8, Feb. 2000.
- [13] M. Novey and T. Adali, "On extending the complex FastICA algorithm to noncircular sources,'' *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 2148–2154, May 2008.
- [14] H. Li and T. Adali, "A class of complex ICA algorithms based on the kurtosis cost function,'' *IEEE Trans. Neural Netw.*, vol. 19, no. 3, pp. 408–420, Mar. 2008.

SHANSHAN LU received the M.S. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2010, where he is currently pursuing the Ph.D. degree. His fields of interest include radar signal processing, the applications to independent component analysis, and electronic countermeasures.

WEI WANG received the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2003. He is currently a Professor with the National University of Defense Technology. His fields of interest include radar signal processing, and electromagnetic environment effects.

GUOYU WANG received the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 1999. He is currently a Professor with the National University of Defense Technology. His fields of interest include radar system simulation and electromagnetic environment effects.

 \sim \sim \sim