Research Article

Decoupled Independent Vector Analysis Algorithm for Convolutive Blind Source Separation without Orthogonality Constraint on the Demixing Matrices

Pengfei Wang,[1,2] **Jiong Li** [ID],[3] **and Hang Zhang** [ID][1]

[1]*College of Communications Engineering, Army Engineering University of PLA, Nanjing, China*
[2]*PANDA Handa (Nanjing) Science and Technology Co., Ltd., Nanjing 221000, China*
[3]*School of Space Information, Space Engineering University of PLA, Beijing, China*

Correspondence should be addressed to Jiong Li; lij_2015@126.com and Hang Zhang; hangzh_2002@163.com

In this paper, we consider the problem of convolutive blind source separation in frequency domain and introduce a solution to the problem in an independent vector analysis (IVA) framework. IVA utilizes both the statistical independence of different sources in each frequency bin and the statistical dependence of the same source in different frequency bins. However, most of previous works impose orthogonality constraint on the rows of each separation matrix which may undermine the separation performance. In this work, we propose a nonorthogonal IVA algorithm based on decoupled relative Newton method. This proposed algorithm updates the separation matrices row by row, and unlike deflation separation algorithm, there is no separation error accumulation arising. Simulation results are provided to show the superior convergence behavior and separation performance of the proposed algorithm.

1. Introduction

Blind source separation (BSS) has been widely researched over the last decades since it is able to estimate the source signals from their mixtures (observed sensor signals) without knowing both the mixing process and the sources, such as crosstalk separation in telecommunications [1], speech enhancement [2], and biomedical signal processing [3]. From the mixing model point of view, the works on BSS can be divided into two categories, which are instantaneous mixing and convolutive mixing. For the instantaneous mixing BSS problem, the most popular and promising method to address this issue is independent component analysis (ICA) [4–6]. ICA assumes that the source signals are mutually statistically independent and at most one source is Gaussian distributed. Based on this assumption, some separation criteria are employed, such as non-Gaussianity maximization [7] and negentropy maximization [6]. However, in many real applications such as wireless communications, the instantaneous mixing model is invalid. In this case, the sources undergo propagation time delay and reverberation which

results in convolutive mixing. In the recent decades, various approaches have been proposed to deal with this case and they basically fall into two categories: time domain [8, 9] and frequency domain [10–12] algorithms. Time domain methods are mainly inspired by the existing blind deconvolution methods and the solutions usually require intensive computations due to the relationship of filter coefficients with each other. The computation load can be overcome by the frequency domain methods. By applying Fourier transform to the time domain convolutive observed signals, it can be converted to multiple independent linear mixing in the frequency domain. Subsequently, the separation methods for instantaneous mixing can be performed on every single frequency bin. However, there is a common problem for instantaneous BSS methods that the estimated signals are disordered. Therefore, further postprocessing is necessary to correct the permutation disorder at each frequency bin and then the separated signals in the time domain are reconstructed properly. Extensive works [13, 14] have been performed to solve the permutation problem, but most of these

permutation correction methods do not perform consistently well.

Independent vector analysis (IVA), an extension of ICA, can solve the frequency domain BSS problem efficiently and normally requires no bin-wise permutation correction post processing [10–12]. IVA algorithms extend the univariate function of ICA algorithms to multivariate function as the score function. In this way, the separation for different frequency bin data is no longer separate but joint, and the permutation problem is mitigated by exploiting the dependences of frequency bins. Various algorithms employ different multivariate prior distributions to preserve the interfrequency dependences for individual sources and corresponding nonlinear score functions are derived. However, most of these algorithms impose orthogonality constraint on the rows of each separation matrix which may undermine the separation performance. In addition, the most popular algorithms to address this issue are based on gradient descent method because of its simplicity [14, 15], but they suffer from slow convergence. Several Newton-based algorithms are proposed to speed up the convergence [12, 16]. However, performance of these algorithms, as shown in this paper, may be unsatisfactory in certain scenarios.

In this paper, we propose a nonorthogonal IVA algorithm based on decoupled relative Newton method. This proposed algorithm updates the separation matrices row by row, and, unlike deflation separation algorithm, there is no separation error accumulation arising. For real data, it has been shown that the decoupled IVA algorithm may converge faster than the vector gradient decent and Newton ones [17]. This paper is an extended version of the literature [17] to the complex domain. In order to avoid the occurrence of singular solutions, some approximate solutions are taken in the process of algorithm derivation, and the improved algorithm is applied to solve the problem of convolutive blind separation. Simulation results are demonstrated to confirm its superior performance.

The rest of this paper is organized as follows. The problem of IVA for frequency domain blind source separation is presented in Section 2. The proposed algorithm is introduced in Section 3. Finally, the performance of the proposed algorithm is evaluated in Section 4 by means of some simulations, and our conclusions are stated in Section 5.

2. Frequency Domain IVA

Let there be N sensors and N independent sources; the convolutedly mixture at the nth sensor is

$$\tilde{x}_n(m) = \sum_{j=1}^{N} \sum_{p=0}^{P-1} a_{nj}(p)\,\tilde{s}_j(m-p) \tag{1}$$

where $\tilde{s}_j(m)$ is jth source in the time domain, $a_{nj}(p)$ is the impulse response of the channel linking jth source and nth sensor, and P is the channel order. The signals are transformed to the frequency domain using the short time Fourier transform (STFT). A sliding window is used to perform discrete Fourier transform; then a time-frequency representation of the observation signal is created. Using

STFT, the sensor observation vector at time block t and frequency bin k becomes

$$\mathbf{x}_t^{[k]} = \mathbf{A}^{[k]}\mathbf{s}_t^{[k]} \quad 1 \le k \le K,\ 1 \le t \le L \tag{2}$$

where $\mathbf{s}_t^{[k]} = [s_{1,t}^{[k]}, s_{2,t}^{[k]}, \cdots, s_{N,t}^{[k]}]^T$ (superscript T denotes transpose operator) is a source vector and $\mathbf{A}^{[k]} \in \mathbb{C}^{N \times N}$ is the unknown nonsingular mixing matrix. The nth source component vector (SCV) can be written as $\mathbf{s}_{n,t}^T = [s_{n,t}^{[1]}, s_{n,t}^{[2]}, \cdots, s_{n,t}^{[K]}]$, which is independent of all the other SCVs. Then, the probability density function of the concatenated source vector can be written as $f(\mathbf{s}_{1,t}, \mathbf{s}_{2,t}, \cdots, \mathbf{s}_{N,t}) = \prod_{n=1}^{N} f(\mathbf{s}_{n,t})$. The IVA solution finds K demixing matrices and the corresponding source signal estimations for each frequency bin, denoted as $\mathbf{W}^{[k]}$ and $\mathbf{y}_t^{[k]} = \mathbf{W}^{[k]}\mathbf{x}_t^{[k]}$ for the kth frequency bin, respectively. nth estimated source of the kth frequency bin at time block t is $y_{n,t}^{[k]} = (\mathbf{w}_n^{[k]})^T \mathbf{x}_t^{[k]}$, where $(\mathbf{w}_n^{[k]})^T$ is nth row vector of $\mathbf{W}^{[k]}$.

3. Proposed Algorithm

3.1. Decoupled Relative Leaning. In the decoupled relative learning, the demixing matrix $\mathbf{W}^{[k]}$ is updated through repeated left multiplications with matrices in form [5]

$$\mathbf{W}_{new}^{[k]} = \left(\mathbf{I} + \mathbf{u}\mathbf{v}^H\right)\mathbf{W}_{old}^{[k]} \tag{3}$$

until convergence, where $\mathbf{I}$ is the $N \times N$ identity matrix and $\mathbf{u}$ and $\mathbf{v}$ are two $N \times 1$ vectors to be optimized. Noting that the eigenvalues of $\mathbf{I} + \mathbf{u}\mathbf{v}^H$ are 1 with multiplicity $N-1$ and $1+\mathbf{v}^H\mathbf{u}$, we can find that (3) is a rank-1 update as long as $1 + \mathbf{v}^H\mathbf{u} \ne 0$. Thus a nonsingular matrix can be transformed into another arbitrary nonsingular matrix within N steps in (3).

It is obvious that the decoupled relative learning transforms the matrix optimization problem into a series of vector optimization problems; thus we can design the algorithm more flexibly compared with the relative (natural) gradient algorithm. We only consider one special kind of relative rank-1 update, i.e., letting $\mathbf{u} = \mathbf{e}_n$, where $\mathbf{e}_n$ is a unit vector whose nth element is 1. Then, by letting $\mathbf{u} = \mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_N$, we can update $\mathbf{W}^{[k]}$ row by row via optimizing $\mathbf{v}$. In this way, as shown in the next subsection, the calculation of the Newton-based learning algorithm is simplified.

3.2. Proposed Algorithm. In this paper, we use the score function [18]

$$J_{IVA} \equiv \sum_{n=1}^{N} H[\mathbf{y}_{n,t}] - \sum_{k=1}^{K} \log\left|\det\left(\mathbf{W}^{[k]}\right)\right| - C_1 \tag{4}$$

where $H[\mathbf{y}_{n,t}] = E[-\log p(\mathbf{y}_{n,t})]$ is the entropy of $\mathbf{y}_{n,t}$ and C_1 is constant which is independent of demixing matrices $\mathbf{W}^{[k]}|_{k=1,2,\cdots,K}$. As processed in ICA, we select a nonlinear function $G : \mathbb{R}^{K \times 1} \longmapsto \mathbb{R}^{1 \times 1}$ to approximate the entropy of SCV; e.g.,

$$H[\mathbf{y}_{n,t}] = E\left\{G\left(\mathbf{y}_{n,t}^H \mathbf{y}_{n,t}\right)\right\}$$

$$G(\mathbf{y}_{n,t}) = \left(\mathbf{y}_{n,t}^H \mathbf{y}_{n,t}\right)^2 \tag{5}$$

Without loss of generality, according to (3), the optimization of the nth row of $\mathbf{W}^{[k]}$ can be written as

$$\Delta\mathbf{w}_n^{[k]} = \mathbf{w}_{n,new}^{[k]} - \mathbf{w}_{n,old}^{[k]} = \left(\Delta\mathbf{v}^H\mathbf{W}_{old}^{[k]}\right)^T \qquad (6)$$

where $\Delta\mathbf{v}$ is a small perturbation of $\mathbf{v}$ at $\mathbf{0}$. Correspondingly, the change of the nth estimated source within kth frequency bin is

$$\Delta y_{n,t}^{[k]} = \left(\Delta\mathbf{w}_n^{[k]}\right)^T\mathbf{x}_t^{[k]} = \Delta\mathbf{v}^H\mathbf{W}_{old}^{[k]}\mathbf{x}_t^{[k]} = \Delta\mathbf{v}^H\mathbf{y}_t^{[k]} \qquad (7)$$

Noting that $\det(\mathbf{W}_{new}^{[k]}) = (1 + \Delta\mathbf{v}^H\mathbf{e}_n)\det(\mathbf{W}_{old}^{[k]})$, the change of cost can be written as

$$\Delta J = E\left\{G\left(\mathbf{y}_{n,t} + \Delta y_{n,t}^{[k]}\mathbf{e}_k\right) - G\left(\mathbf{y}_{n,t}\right)\right\} \\ - \log\left|1 + \Delta\mathbf{v}^H\mathbf{e}_n\right| \qquad (8)$$

Then, using second-order Taylor series expansion to approximate $G(\mathbf{y}_{n,t} + \Delta y_{n,t}^{[k]}\mathbf{e}_k)$ and $\log|1 + \Delta\mathbf{v}^H\mathbf{e}_n|$, the change of cost ΔJ can be written as

$$\Delta J \approx \begin{bmatrix} \Delta\mathbf{v}^H & \Delta\mathbf{v}^T \end{bmatrix}\begin{bmatrix} \mathbf{f} \\ \mathbf{f}^* \end{bmatrix} \\ + \frac{1}{2}\begin{bmatrix} \Delta\mathbf{v}^H & \Delta\mathbf{v}^T \end{bmatrix}\begin{bmatrix} \mathbf{D}_0 & \mathbf{D}_1 \\ \mathbf{D}_1^* & \mathbf{D}_0^* \end{bmatrix}\begin{bmatrix} \Delta\mathbf{v} \\ \Delta\mathbf{v}^* \end{bmatrix} \qquad (9)$$

where $\mathbf{f} = E\{G'(\mathbf{y}_{n,t})|_{y_{n,t}^{[k]}}\mathbf{y}_t^{[k]}\} - \mathbf{e}_n$, $G'(\mathbf{y}_{n,t})|_{y_{n,t}^{[k]}}$, $G''(\mathbf{y}_{n,t})|_{y_{n,t}^{[k]}}$, $\mathbf{D}_0 = E\{G''(\mathbf{y}_{n,t})|_{y_{n,t}^{[k]}(y_{n,t}^{[k]})^*}\mathbf{y}_t^{[k]}(\mathbf{y}_t^{[k]})^H\}$, and $\mathbf{D}_1 = E\{G''(\mathbf{y}_{n,t})|_{y_{n,t}^{[k]}y_{n,t}^{[k]}}\mathbf{y}_t^{[k]}\mathbf{y}_t^{[k]}\} + \mathbf{e}_n\mathbf{e}_n^H$ are the first- and second-order derivatives of $G(\mathbf{y}_{n,t})$ with respect to $y_{n,t}^{[k]}$, respectively.

To design a Newton algorithm, the quadratic form in (9) is required to be positive definite. However, the quadratic form may not be positive definite when $y_{n,t}^{[k]}$ does not converge to the source. Therefore, we make an approximation of the quadratic form to ensure the positive definiteness. When $y_{n,t}^{[k]}$ converges to the source, $\mathbf{D}_0$ and $\mathbf{D}_1$ reduce to diagonal matrices respectively. Approximate $\mathbf{D}_0$ and $\mathbf{D}_1$ as

$$\widehat{\mathbf{D}}_0 = \mathrm{diag}\left\{E\left\{G''(\mathbf{y}_{n,t})\big|_{y_{n,t}^{[k]}(y_{n,t}^{[k]})^*}\mathbf{y}_t^{[k]}\odot\left(\mathbf{y}_t^{[k]}\right)^*\right\}\right\} \qquad (10)$$

$$\widehat{\mathbf{D}}_1 = \mathrm{diag}\left\{E\left\{G''(\mathbf{y}_{n,t})\big|_{y_{n,t}^{[k]}y_{n,t}^{[k]}}\mathbf{y}_t^{[k]}\odot\mathbf{y}_t^{[k]}\right\} + \mathbf{e}_n\right\} \qquad (11)$$

where $\mathrm{diag}\{\mathbf{d}\}$ generates a diagonal matrix with diagonal element vector $\mathbf{d}$ and $\odot$ denotes element-wise multiplication of two vectors. Then, the Hessian matrix for the ith element of $\Delta\mathbf{v}$ is $\begin{bmatrix} d_{0,i} & d_{1,i} \\ d_{1,i}^* & d_{0,i} \end{bmatrix}$, where $d_{0,i}$ and $d_{1,i}$ are the ith diagonal elements of $\widehat{\mathbf{D}}_0$ and $\widehat{\mathbf{D}}_1$, respectively. To ensure the positive definiteness of this Hessian matrix, a modification is made:

$$d_{0,i} = \max\left[d_{0,i}, (1+\alpha)\left|d_{1,i}\right|\right] \qquad (12)$$

where α is a small positive number. The Newton update rule for the ith element of $\Delta\mathbf{v}$ is derived as

$$\Delta v_i = -\frac{d_{0,i}f_i - d_{1,i}f_i^*}{d_{0,i}^2 - \left|d_{1,i}\right|^2} \qquad (13)$$

From the derivation process of the proposed algorithm, we can see that there is no orthogonality constraint on the demixing matrices and no separation error accumulation arises.

4. Simulation and Results

In the first simulation, we consider the separation of two convoluted speech signals generated by (1), and set the channel order $P = 3$. At each mixed signal, 1024-point FFT is done to time blocks. Figure 1 shows the separation result of two convolutedly mixed speech signals using the proposed algorithm. It indicates that the proposed algorithm is able to separate convolutive mixing successfully in frequency domain.

In order to numerically compare the separation performance of the proposed decoupled IVA algorithm with other IVA algorithms, multiple datasets separation for complex sources are simulated. The complex mixed signal is generated by (2). In addition, the dependence elements of SCV are generated as

$$\mathbf{s}_{n,t} = \sum_{i=0}^{\beta}\mathbf{M}_{n,i}\mathbf{z}_{n,t-i} \qquad (14)$$

where $\mathbf{M}_{n,i}$ is a real-valued $K \times K$ matrix and $\mathbf{z}_{n,t-i}$ is a zero-mean vector whose entities are subject to uniform distribution. In the following simulations, $\beta = 3$.

In this section, we compare the proposed decoupled IVA algorithm with other two algorithms, vector gradient, and Newton updates for IVA with multivariate Gaussian model [16], respectively. Performance is assessed using interference to source ratio (ISR)

$$ISR = \frac{1}{KN(N-1)}\sum_{k=1}^{K}\sum_{\substack{m,n=1 \\ m\neq n}}^{N}\left(C_{m,n}^{[k]}\right)^2 \qquad (15)$$

where $C_{m,n}^{[k]}$ is the (m,n) element of $\mathbf{C}^{[k]} = \mathbf{W}^{[k]}\mathbf{A}^{[k]}$; here we assume that no permutation exists. The smaller the value of ISR, the better the separation performance, specially, $ISR = 0$ implies ideal separation performance. In all the simulations, the results of 100 independent realizations are averaged.

Figure 2 shows the convergence behavior of the ISR. We observe that, among the three algorithms, the proposed decoupled relative Newton algorithm converges in the fewest iterations, and the vector gradient algorithm performs poorly compared with the other two algorithms. In addition, the decoupled relative Newton algorithm has the lowest ISR values when the algorithm converges, which implies that the decoupled relative Newton algorithm has the best separation performance.

Figure 3 depicts the separation performance versus the datasets number. We can observe that the separation performance of the proposed algorithm is the best among the three algorithms, and the ISR value of the proposed decoupled algorithm is lower than the other comparison algorithms around 15-20dB. Furthermore, the ISR of the three algorithms increases with the number of datasets. In addition, we also
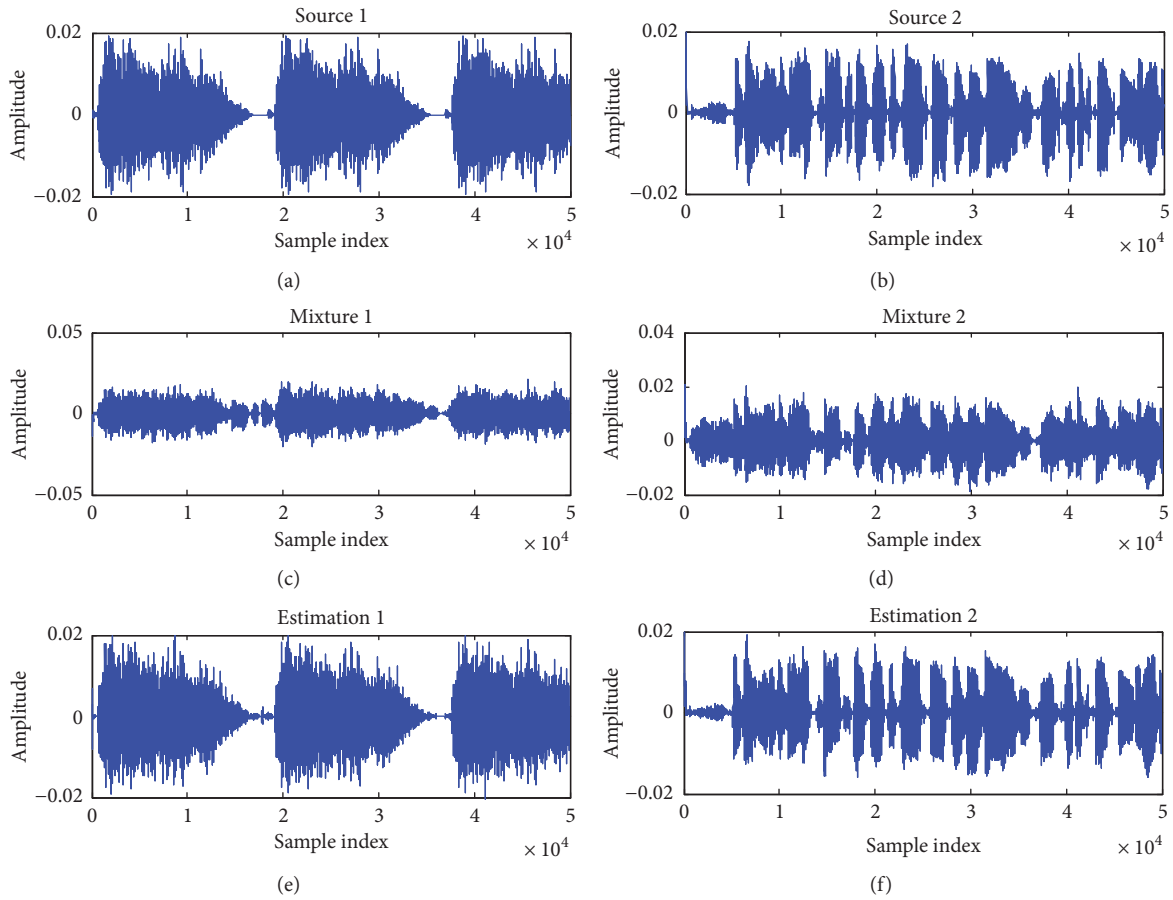
FIGURE 1: Separation of convoluted speech signals using the proposed algorithm.
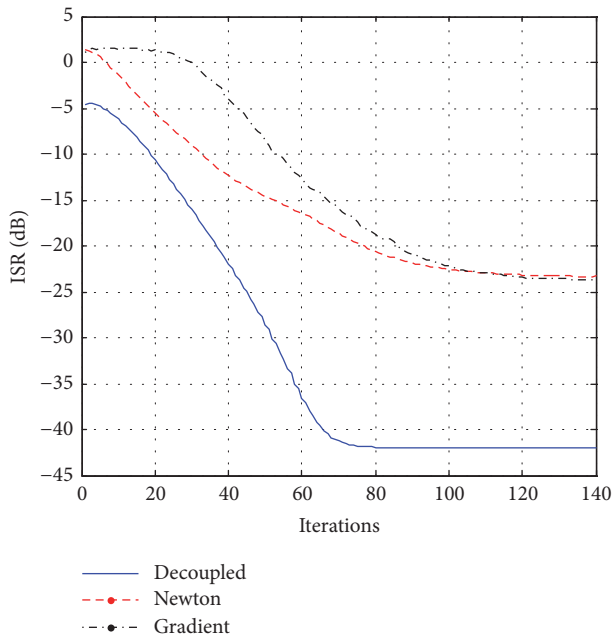

FIGURE 2: Mean ISR of 100 trails versus number of iterations for $K=4$, $N=4$, $T=3000$ using the vector gradient, Newton and proposed decoupled relative Newton optimization algorithms.
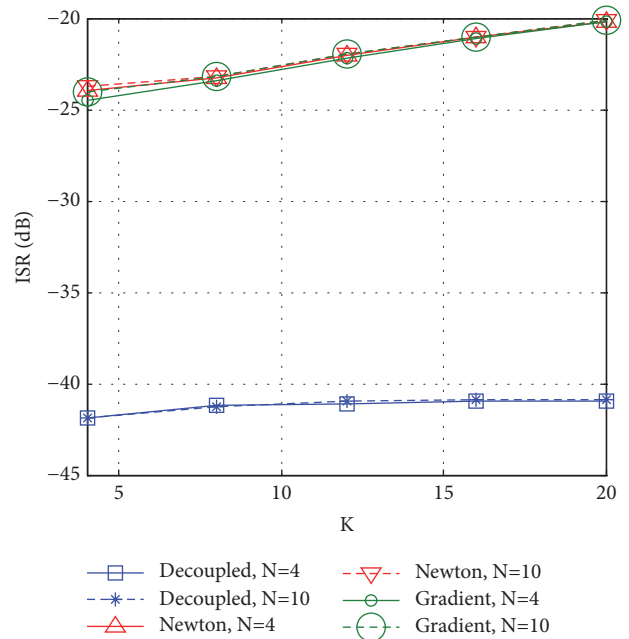
FIGURE 3: Mean ISR of 100 trails versus the number of datasets K for $N=4$, 10, $T=3000$ using the vector gradient, Newton, and proposed decoupled relative Newton optimization algorithms.
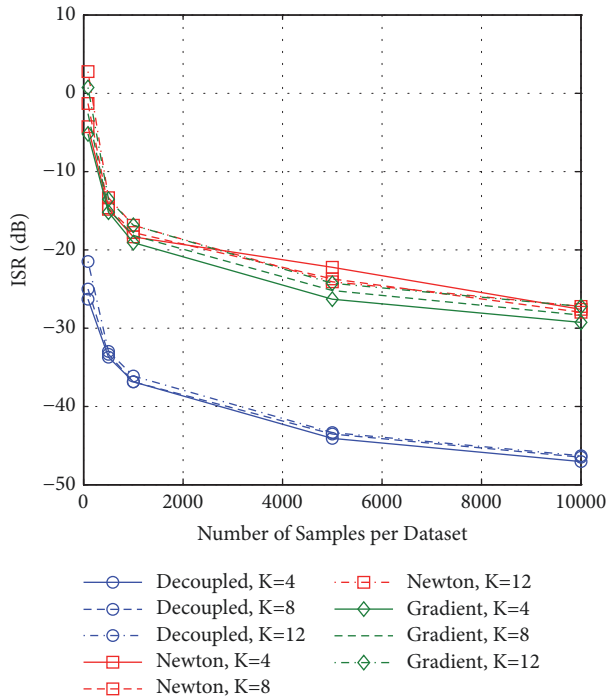
FIGURE 4: Mean ISR of 100 trails versus the number of samples per dataset for $K=4,8,12$, $N=4$ using the vector gradient, Newton, and proposed decoupled relative Newton optimization algorithms.

consider the influence of the number of sources. It shows that the number of sources has little effect on the separation results.

Figure 4 shows the ISR of the estimated sources as a function of the number of samples per dataset, where we can find that the ISR of the proposed algorithm is the lowest among the three algorithms and the ISR of the three algorithms increases as the number of datasets increases, with the largest rate of decrease occurring for the sample length smaller than 1000. When the sample length is smaller than 1000, vector gradient and Newton optimization algorithms have a similar separation performance, while, with the increase of sample length, the influence of the number of datasets becomes more apparent.

5. Conclusion

In this paper, we propose a new IVA algorithm for separating convolutedly mixed signals in frequency domain which is based on decoupled relative Newton approach. This algorithm decomposes the matrix optimization problem into a series of row vector optimization problems, and the separation matrices do not need to be orthogonal. Simulation results show that the proposed algorithm converges fast and its separation performance is superior to that of other two algorithms based on vector gradient and Newton methods, respectively.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] Y. Duan and H. Zhang, "Noisy blind signal-jamming separation algorithm based on VBICA," *Wireless Personal Communications*, vol. 74, no. 2, pp. 307–324, 2014.

[2] S. Gannot, E. Vincent, S. Markovich-Golan, and A. Ozerov, "A consolidated perspective on multimicrophone speech enhancement and source separation," *IEEE/ACM Transactions on Audio Speech and Language Processing*, vol. 25, no. 4, pp. 692–730, 2017.

[3] T. Adali, Y. Levin-Schwartz, and V. D. Calhoun, "Multimodal Data Fusion Using Source Separation: Application to Medical Imaging," *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1494–1506, 2015.

[4] E. Bingham and A. Hyvärinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals," *International Journal of Neural Systems*, vol. 10, no. 1, pp. 1–8, 2000.

[5] X.-L. Li and T. Adali, "Noncircular complex ICA by generalized Householder reflections," *IEEE Transactions on Signal Processing*, vol. 61, no. 24, pp. 6423–6430, 2013.

[6] M. Novey and T. Adali, "Complex ICA by negentropy maximization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 19, no. 4, pp. 596–609, 2008.

[7] M. Novey and T. Adalt, "Adaptable nonlinearity for complex maximization of nongaussianity and a fixed-point algorithm," in *Proceedings of the 16th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing (MLSP '06)*, pp. 79–84, September 2006.

[8] R. Dubroca, C. De Luigi, M. Castella, and E. Moreau, "A general algebraic algorithm for blind extraction of one source in a MIMO convolutive mixture," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2484–2493, 2010.

[9] S. Saito, K. Oishi, and T. Furukawa, "Convolutive blind source separation using an iterative least-squares algorithm for nonorthogonal approximate joint diagonalization," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 23, no. 12, pp. 2434–2448, 2015.

[10] Y. Liang, J. Harris, S. M. Naqvi, G. Chen, and J. A. Chambers, "Independent vector analysis with a generalized multivariate Gaussian source prior for frequency domain blind source separation," *Signal Processing*, vol. 105, pp. 175–184, 2014.

[11] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, "Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization," *IEEE/ACM Transactions on Audio Speech and Language Processing*, vol. 24, no. 9, pp. 1622–1637, 2016.

[12] H. Zhang, L. Li, and W. Li, "Independent vector analysis for convolutive blind noncircular source separation," *Signal Processing*, vol. 92, no. 9, pp. 2275–2283, 2012.

[13] H. Sawada, S. Araki, and S. Makino, "Measuring dependence of bin-wise separated signals for permutation alignment in frequency-domain BSS," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '07)*, pp. 3247–3250, May 2007.

[14] T. Kim, H. T. Attias, S.-Y. Lee, and T.-W. Lee, "Blind source separation exploiting higher-order frequency dependencies," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 15, no. 1, pp. 70–79, 2007.

[15] Y. F. Liang, S. M. Naqvi, and J. A. Chambers, "Adaptive step size independent vector analysis for blind source separation," in *Proceedings of the 17th International Conference on Digital Signal Processing (DSP '11)*, pp. 1–6, July 2011.

[16] M. Anderson, X.-L. Li, and T. Adal, "Complex-valued independent vector analysis: Application to multivariate Gaussian model," *Signal Processing*, vol. 92, no. 8, pp. 1821–1831, 2012.

[17] J. Li, H. Zhang, W. Ding, and P. Wang, "Fast joint blind source separation for multiple datasets using decoupled relative Newton approach," in *Proceedings of the 3rd IEEE International Conference on Computer and Communications (ICCC '17)*, pp. 2311–2314, December 2017.

[18] M. Anderson, T. Adali, and X.-L. Li, "Joint blind source separation with multivariate Gaussian model: algorithms and performance analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1672–1683, 2012.