Blind Instantaneous Noisy Mixture Separation with Best Interference-Plus-Noise Rejection*

Zbyněk Koldovský[1,2] and Petr Tichavský[1]

[1] Institute of Information Theory and Automation, Pod vodárenskou věží 4,
P.O. Box 18, 182 08 Praha 8, Czech Republic
zbynek.koldovsky@tul.cz

[2] Faculty of Mechatronic and Interdisciplinary Studies
Technical University of Liberec, Hálkova 6, 461 17 Liberec, Czech Republic

Abstract. In this paper, a variant of the well known algorithm FastICA is proposed to be used for blind source separation in off-line (block processing) setup and a noisy environment. The algorithm combines a symmetric FastICA with test of saddle points to achieve fast global convergence and a one-unit refinement to obtain high noise rejection ability. A novel test of saddle points is designed for separation of complex-valued signals. The bias of the proposed algorithm due to additive noise can be shown to be asymptotically proportional to σ^3 for small σ, where σ^2 is the variance of the additive noise. Since the bias of the other methods (namely the bias of all methods using the orthogonality constraint, and even of recently proposed algorithm EFICA) is asymptotically proportional to σ^2, the proposed method has usually a lower bias, and consequently it exhibits a lower symbol-error rate, when applied to blind separation of finite alphabet signals, typical for communication systems.

1 Introduction

The noisy model of Independent Component Analysis (ICA) considered in this paper, is

$$\mathbf{X} = \mathbf{AS} + \sigma\mathbf{N}, \tag{1}$$

where $\mathbf{S}$ denotes a vector of d independent random variables representing the original signals, $\mathbf{A}$ is an unknown regular $d \times d$ mixing matrix, and $\mathbf{X}$ represents the observed mixed signals. The noise $\mathbf{N}$ denotes a vector of independent variables having the covariance matrix $\mathbf{\Sigma}$. Without loss of generality, we will further assume that $\mathbf{\Sigma}$ equals to the identity matrix $\mathbf{I}$. Consequently, σ^2 is the variance of the added noise to the mixed signals. All signals considered here are i.i.d. sequences, i.e., they are assumed to be white in the analysis.

It is characteristic for most ICA methods that they were derived for the noiseless case, so to solve the task of estimating the mixing matrix $\mathbf{A}$ or its inversion

* This work was supported by Ministry of Education, Youth and Sports of the Czech Republic through the project 1M0572 and through the Grant 102/07/P384 of the Grant Agency of the Czech Republic.

M.E. Davies et al. (Eds.): ICA 2007, LNCS 4666, pp. 730–737, 2007.

$\mathbf{W} = \mathbf{A}^{-1}$. Then, abilities to separate noised data are studied experimentally, and the non-vanishing estimation error as $N \to +\infty$, N being length of data, is taken for a bias caused by the noise. To compensate such bias, several techniques were proposed [6]. Unfortunately, these methods have a drawback that the covariance structure of the noise needs to be known *a priori* [4].

In accord with [7], we suggest to measure the separation quality not through accuracy of estimation of the mixing mechanism but through the achieved interference + noise to signal ratio (INSR) or its inverse SINR. In separating the finite alphabet signals, the ultimate criterion should be the symbol error rate (SER). Computation of the INSR or the SER assumes that the permutation, scale, and sign or phase ambiguities were resolved by minimizing the INSR.

In the case $\mathbf{\Sigma} = \mathbf{I}$, the INSR of a k-th estimated signal can be computed as

$$\mathrm{INSR}_k = \frac{\sum_{i\neq k}^{d} (\mathbf{BA})_{ki}^2 + \sigma^2 \sum_{i=1}^{d} \mathbf{B}_{ki}^2}{(\mathbf{BA})_{kk}^2}, \tag{2}$$

where $\mathbf{B}$ is the separating transformation [7]. The solutions that minimize (2) are known to be given by the MMSE separating matrix, denoted by $\mathbf{W}^{\mathrm{MMSE}}$, that takes the form

$$\mathbf{W}^{\mathrm{MMSE}} = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H + \sigma^2\mathbf{I})^{-1} \tag{3}$$

where ^{H} denotes the conjugate (Hermitian) transpose. Signals given by $\mathbf{W}^{\mathrm{MMSE}}\mathbf{X}$ will be further called the *MMSE solution*. Note that these signals may not be necessarily normalized to have unit variance, unlike outcome of common blind separation methods, that produce normalized components. For exact comparisons, we introduce a matrix $\mathbf{W}^{\mathrm{NMMSE}}$ such that $\mathbf{W}^{\mathrm{NMMSE}}\mathbf{X}$ are the normalized MMSE signals.

The paper is organized as follows. In section 2, we briefly describe several variants of algorithm FastICA and the proposed method, including a novel test of saddle points for separating complex-valued signals. Section 3 presents analytic expressions for an asymptotic bias of solutions obtained by real domain FastICA variants [5,8] from the MMSE solution. Specifically, we study the biases of estimates of de-mixing matrix $\mathbf{W}$ from $\mathbf{W}^{\mathrm{NMMSE}}$, and the one-unit FastICA and the proposed algorithm is shown to be less biased than the other methods. Simulations in Section 4 demonstrate drawbacks of the unbiased algorithm [6] (further referred to as *unbiased FastICA*) following from required knowledge of $\mathbf{\Sigma}$ and/or σ. Conversely, the proposed algorithm with one-unit FastICA-like performance is shown to be the best blind MMSE estimator when separating noisy finite-alphabet signals.

2 FastICA and Its Variants

Common FastICA algorithms work with the decorrelated data $\mathbf{Z} = \mathbf{C}^{-1/2}\mathbf{X}$, where $\mathbf{C} = \mathrm{E}[\mathbf{X}\mathbf{X}^H]$ is the data covariance matrix. Only the unbiased FastICA [6] that aims at unbiased estimation of $\mathbf{A}^{-1}$ assuming that the noise has a known covariance matrix $\mathbf{\Sigma}$, uses the preprocessing $\mathbf{Z} = (\mathbf{C} - \mathbf{\Sigma})^{-1/2}\mathbf{X}$.

One-unit FastICA in real domain [5] estimates one de-mixing vector $\mathbf{w}_k^{1\mathrm{U}}$ iteratively via the recursion

$$\mathbf{w}_k^+ \leftarrow \mathrm{E}[\mathbf{Z}g({\mathbf{w}_k^{1\mathrm{U}}}^T\mathbf{Z})]\} - \mathbf{w}_k^{1\mathrm{U}}\mathrm{E}\{g'({\mathbf{w}_k^{1\mathrm{U}}}^T\mathbf{Z}))\}, \qquad \mathbf{w}_k^{1\mathrm{U}} \leftarrow \mathbf{w}_k^+/\|\mathbf{w}_k^+\| \quad (4)$$

until convergence is achieved. Here $g(\cdot)$ is a smooth nonlinear function that approximates/surrogates the score function corresponding to the distribution of the original signals [11]. The theoretical expectation values in (4) are, in practice, replaced by their sample-based counterparts.

Similar recursion was proposed for one-unit FastICA in the complex domain [1]. The symmetric (real or complex) variant performs the one-unit iterations in parallel for all d separating vectors, but the normalization in (4) is replaced by a symmetric orthogonalization.

The algorithm EFICA [8] combines the symmetric approach with the test of saddle points, an adaptive choice of nonlinearity $g_k(\cdot)$ for each signal separately, and it does the refinement step that relaxes the orthogonal constraint introduced by the symmetric approach and is designed towards asymptotic efficiency.

The unbiased FastICA [6] uses the recursion

$$\mathbf{w}_k^+ \leftarrow \mathrm{E}[\mathbf{Z}g({\mathbf{w}_k^{\mathrm{unb}}}^T\mathbf{Z})] - (\mathbf{I} + \widetilde{\boldsymbol{\Sigma}})\mathbf{w}_k^{\mathrm{unb}}\,\mathrm{E}[g'({\mathbf{w}_k^{\mathrm{unb}}}^T\mathbf{Z})],$$

where $\widetilde{\boldsymbol{\Sigma}} = (\mathbf{C} - \boldsymbol{\Sigma})^{-1/2}\boldsymbol{\Sigma}(\mathbf{C} - \boldsymbol{\Sigma})^{-1/2}$. Both approaches (one-unit and symmetric) can be considered; in simulations, we use the one-unit variant, and the resulting de-mixing matrix will be denoted by $\mathbf{W}^{\mathrm{UNB}}$. In order to compare performance of the unbiased FastICA by means of (2) with the other techniques fairly, it is necessary to consider a MMSE estimate derived from $\mathbf{W}^{\mathrm{UNB}}$, namely

$$\mathbf{W}^{\mathrm{MMSE\text{-}UNB}} = \boldsymbol{\Sigma}^{-1}(\mathbf{W}^{\mathrm{UNB}})^{-T} \times [(\mathbf{W}^{\mathrm{UNB}})^{-1}\boldsymbol{\Sigma}^{-1}(\mathbf{W}^{\mathrm{UNB}})^{-T} + \sigma^2\mathbf{I}]^{-1} \quad (5)$$

2.1 Proposed Algorithm

The proposed algorithm is a combination of symmetric FastICA, test of saddle points, and one-unit FastICA as a refinement. Usually, one unit FastICA is used in a deflation way, when the estimated components are subtracted from the mixture one by one. This is computationally effective method, but accuracy of the later separated components might be compromised. Therefore, we propose to initialize the algorithm using symmetric FastICA, that is known for having very good global convergence and allows equal separation precision for all components.

The test of saddle points was first proposed in [11] to improve probability of the symmetric FastICA to converge to the true global maximum of the cost function $[\mathrm{E}\{G(\mathbf{w}^T\mathbf{Z})\} - G_0]^2$ where $G(\cdot)$. is a primitive function of $g(\cdot)$ and $G_0 = \mathrm{E}\{G(\xi)\}$, where ξ is a standard Gaussian random variable.

In short, the test of saddle points consists in checking all pairs of the estimated components $(\mathbf{u}_k, \mathbf{u}_\ell)$, whether or not other pair of signals $(\mathbf{u}'_k, \mathbf{u}'_\ell)$ gives a higher value of the cost function

$$c(\mathbf{u}_k, \mathbf{u}_\ell) = [\mathrm{E}\{G(\mathbf{u}_k)\} - G_0]^2 + [\mathrm{E}\{G(\mathbf{u}_\ell)\} - G_0]^2, \quad (6)$$

where $\mathbf{u}'_k = (\mathbf{u}_k + \mathbf{u}_\ell)/\sqrt{2}$ and $\mathbf{u}'_\ell = (\mathbf{u}_k - \mathbf{u}_\ell)/\sqrt{2}$.

The motivation is that a random initialization of the algorithm may begin at a point of zero gradient of the cost function (a saddle point / an unstable point of the iteration) and terminate there, despite being not the desired stable solution. See [11] for details.

In the complex domain, the situation is a bit more tricky, because if $(\mathbf{u}_k, \mathbf{u}_\ell)$ is the pair of valid independent components in the mixture, not only their weighted sum and a difference represent a false (unstable) point of the iteration. *All* pairs $(\mathbf{u}'_k, \mathbf{u}'_\ell)$ of the form $\mathbf{u}'_k = (\mathbf{u}_k + e^{i\alpha}\mathbf{u}_\ell)/\sqrt{2}$ and $\mathbf{u}'_\ell = (\mathbf{u}_k - e^{i\alpha}\mathbf{u}_\ell)/\sqrt{2}$ are stationary but unstable for any phase factor $e^{i\alpha}$, $\alpha \in \mathcal{R}$.

Therefore we propose to do a phase shift of each separated component so that the real part and the imaginary part of the signal are as much independent each of other as possible before the test of the saddle points. This phase shift can be easily performed using a two-dimensional symmetric FastICA in the real domain applied to the real and imaginary part of the component. After this preprocessing, it is sufficient to perform the test of saddle points exactly as in the real-valued case, i.e. to check all pairs $(\mathbf{u}'_k, \mathbf{u}'_\ell)$ with $\mathbf{u}'_k = (\mathbf{u}_k + \mathbf{u}_\ell)/\sqrt{2}$ and $\mathbf{u}'_\ell = (\mathbf{u}_k - \mathbf{u}_\ell)/\sqrt{2}$, whether they give a higher value of the cost function (6) or not.

Validity of the above described complex domain test of the saddle points can be easily confirmed in simulations by starting the algorithm from the pairs $\mathbf{u}'_k = (\mathbf{u}_k + e^{i\alpha}\mathbf{u}_\ell)/\sqrt{2}$ and $\mathbf{u}'_\ell = (\mathbf{u}_k - e^{i\alpha}\mathbf{u}_\ell)/\sqrt{2}$ with an arbitrary $\alpha \in \mathcal{R}$ where $\mathbf{u}_k$ and $\mathbf{u}_\ell$ are the true independent sources. We have successfully tested this approach on separation of complex-valued finite alphabet sources known in communications (QAM, V27).

The resultant algorithm (symmetric FastICA + test of saddle points + one unit refinements) will be referred to as 1FICA.

3 Bias of the FastICA Variants

In this section, asymptotic expressions for bias of algorithms described in previous section working in the real domain will be presented. (The complex-domain FastICA exhibits a similar behavior in simulations.) For details of analysis, the reader is referred to [9] due to lack of space.

In brief, the theoretical analysis is done for "small" σ and infinite number of samples. Similarly to [11], for theoretical considerations, it is assumed that the analyzed method starts from the MMSE solution and stops after one iteration. This assumption is reasonable due to the following facts: (1) deviation of the global maximizer $\widehat{\mathbf{W}}$ of the FastICA cost function from $\mathbf{W}^{\mathrm{MMSE}}$ is of the order $O(\sigma^2)$, and (2) convergence of the algorithm is at least quadratic [10]. Therefore, after performing the one iteration, the deviation of the estimate from the global maximizer $\widehat{\mathbf{W}}$ is of the order $O(\sigma^4)$ and, hence, is negligible.

The bias of the algorithm will be studied in terms of the deviation of $\widehat{\mathbf{W}}(\mathbf{W}^{\mathrm{MMSE}})^{-1}$ from a diagonal matrix. More precisely, the bias is equal to the difference between $\mathrm{E}[\widehat{\mathbf{W}}](\mathbf{W}^{\mathrm{MMSE}})^{-1}$ and $\mathbf{D} = \mathbf{W}^{\mathrm{NMMSE}}[\mathbf{W}^{\mathrm{MMSE}}]^{-1}$, where $\mathbf{D}$ is the diagonal matrix that normalizes the MMSE signals $\mathbf{S}^{\mathrm{MMSE}} = \mathbf{W}^{\mathrm{MMSE}}\mathbf{X}$.

It holds that

$$\mathbf{D} = \mathbf{I} + \frac{1}{2}\sigma^2 \mathrm{diag}[\mathbf{V}_{11}, \ldots, \mathbf{V}_{dd}] + O(\sigma^3). \tag{7}$$

From here we use the notation $\mathbf{W} = \mathbf{A}^{-1}$ and $\mathbf{V} = \mathbf{W}\mathbf{W}^T$. Finally, for a matrix $\widehat{\mathbf{W}}$ that separates the data $\mathbf{S}^{\mathrm{MMSE}}$, the bias is $\mathrm{E}[\widehat{\mathbf{W}}] - \mathbf{D}$.

3.1 Bias of the One-Unit FastICA and 1FICA

It can be shown that the de-mixing vector $\mathbf{w}_k^{1\mathrm{U}}$ resulting from the one-unit FastICA (applied to the data $\mathbf{S}^{\mathrm{MMSE}}$), for $N \to +\infty$, is proportional to

$$\mathbf{w}_k^{1\mathrm{U}} = \tau_k \mathbf{e}_k + \frac{1}{2}\sigma^2 \mathbf{V}_{kk}(\tau_k + \delta_k)\mathbf{e}_k + O(\sigma^3) \tag{8}$$

where $\tau_k = \mathrm{E}[s_k g(s_k) - g'(s_k)]$, and δ_k is a scalar that depends on the distribution of s_k and on the nonlinear function g and its derivatives to the third order. Since (8) is a scalar multiple of $\mathbf{e}_k$ (the k-th column of the identity matrix), it follows that the asymptotic bias of the one-unit approach is $O(\sigma^3)$. Prospectively, the separating matrix $\mathbf{W}^{1\mathrm{F}}$ given by the proposed 1FICA has the same bias. Simulations confirm this expectation [9].

3.2 Bias of the Inversion Solution

It is interesting to compare the previous result with the solution that is given by exact inversion of the mixing matrix, i.e. $\mathbf{W}\mathbf{X} = \mathbf{S} + \sigma\mathbf{W}\mathbf{N}$; the signals will be called *the inversion solution*. From

$$\mathbf{W}(\mathbf{W}^{\mathrm{MMSE}})^{-1} = \mathbf{W}(\mathbf{A}\mathbf{A}^T + \sigma^2\mathbf{I})\mathbf{W}^T = \mathbf{I} + \sigma^2\mathbf{V}$$

it follows that the “bias” of the inversion solution is proportional to σ^2 and in general it is greater than that of 1FICA. In other words, **the algorithm 1FICA produces components that are asymptotically closer to the MMSE solution than to the inversion solution**.

3.3 Bias of Algorithms Using the Orthogonal Constraint

Large number of ICA algorithms (e.g. JADE [2], symmetric FastICA, etc.) use an orthogonal constraint, i.e., they enforce the separated components to have sample correlations equal to zero. Since the second-order statistics cannot be estimated perfectly, this constraint compromises the separation quality [3,11]. Here we show that the bias of all ICA algorithms that use the constraint has the asymptotic order $O(\sigma^2)$.

The orthogonality constraint can be written as

$$\mathrm{E}[\widehat{\mathbf{W}}\mathbf{X}(\widehat{\mathbf{W}}\mathbf{X})^T] = \widehat{\mathbf{W}}(\mathbf{A}\mathbf{A}^T + \sigma^2\mathbf{I})\widehat{\mathbf{W}}^T = \mathbf{I} \; . \tag{9}$$

It follows that the bias of all constrained algorithms is lower bounded by

$$\min_{\widehat{\mathbf{W}}(\mathbf{A}\mathbf{A}^T + \sigma^2\mathbf{I})\widehat{\mathbf{W}}^T = \mathbf{I}} \|\widehat{\mathbf{W}}(\mathbf{W}^{\mathrm{MMSE}})^{-1} - \mathbf{D}\|_F = O(\sigma^2) \tag{10}$$

where the minimization proceeds for $\widehat{\mathbf{W}}$. The matrix $\mathbf{D}$ in (10) is the same as in (7). For the minimizer $\widehat{\mathbf{W}}$ of (10) it holds that $\widehat{\mathbf{W}}(\mathbf{W}^{\mathrm{MMSE}})^{-1} = \mathbf{I}+\sigma^2\mathbf{\Gamma}+O(\sigma^3)$, where $\mathbf{\Gamma}$ is a nonzero matrix obeying $\mathbf{\Gamma}+\mathbf{\Gamma}^T = \mathbf{V}$; see [9] for details. This result can be interpreted in the way that the algorithms using the orthogonality constraint may prefer some of the separated components to give them a zero bias, but the total average bias for all components has the order $O(\sigma^2)$.

3.4 Bias of the Symmetric FastICA and EFICA

The biases of the algorithms can be expressed as

$$\mathrm{E}[\widehat{\mathbf{W}}](\mathbf{W}^{\mathrm{MMSE}})^{-1} - \mathbf{D} = \frac{1}{2}\sigma^2\mathbf{V} \odot (\mathbf{1}_{d\times d} - \mathbf{I} + \mathbf{H}) + O(\sigma^3), \tag{11}$$

where $\mathbf{H}_{k\ell} = \frac{|\tau_\ell|-|\tau_k|}{|\tau_k|+|\tau_\ell|}$ for the symmetric FastICA, and $\mathbf{H}_{k\ell} = \frac{c_{k\ell}|\tau_\ell|-|\tau_k|}{|\tau_k|+c_{k\ell}|\tau_\ell|}$ for EFICA, where $c_{k\ell} = \frac{|\tau_\ell|\gamma_k}{|\tau_k|(\gamma_\ell+\tau_\ell^2)}$ for $k \neq \ell$ and $c_{kk} = 1$. Here, $\gamma_k = \mathrm{E}[g_k^2(s_k)] - \mathrm{E}^2[s_k g_k(s_k)]$, and g_k is the nonlinear function chosen for the k-th signal.

It can be seen that the bias of both of the algorithms has the order $O(\sigma^2)$.

4 Simulations

In this section, we present results of two experiments to demonstrate and compare the performance of the proposed algorithm 1FICA with competing methods: The symmetric FastICA (marked by SYMM), the unbiased FastICA (unbiased FICA), EFICA, and JADE [2]. Results given by "oracle" MMSE solution and the inversion solution are included as well. Examples with complex signals are not included due to lack of space.

In the first example, we separate 10 randomly mixed [7] BPSK signals with added Gaussian noise, first, for various length of data N (Fig. 1(a)) and, second, for varying input signal-to-noise ratio (SNR) defined as $1/\sigma^2$ (Fig. 1(b)). The experiment encompasses several extremal conditions: In the first scenario, where SNR=5dB ($\sigma \doteq 0.56$), N goes from 100, which is quite low for the dimension $d = 10$. The second situation examines $N = 200$ and SNR going down to 0dB.

Note that the bias may be less important than the estimation variance when the data length N is low. Therefore, in simulations, we have included two slightly changed versions of 1FICA and EFICA algorithm, denoted by "1FICA-biga" and "EFICA-biga", respectively. The modifications consist in that the used nonlinear function g is equal to the score function of marginal pdfs of the signals to-be estimated (i.e., noisy BPSK that have **bi**modal **Ga**ussian distribution, therefore, "biga" in the acronym). Adopted from the noiseless case [11], better performance of the modified algorithms may be expected.

Figure 1 shows superior performance of the proposed algorithm 1FICA and of its modified version. The same performance is achieved by the modified EFICA for $N \leq 200$, but it is lower due to the bias when N is higher. The unbiased FastICA achieves the same accuracy for $N \geq 500$ but is unstable when N is low.

The average performance of an algorithm is often spoiled due to poorer stability, which occurs in high dimensions and low N cases, mainly. In this issue, we highlight positive effect of the test of saddle points that is included in the proposed 1FICA or in EFICA. For instance, the results achieved by the symmetric FastICA would be significantly improved if the test was included in it.

The second example demonstrates conditions when the covariance of the noise is not exactly known or varying. To this end, the noise level was changed randomly from trial to trial. Five BPSK signals of the length $N = 50000$ were mixed with a random matrix and disturbed by Gaussian noise with covariance $\sigma^2\mathbf{I}$, where σ was randomly taken from interval $[0, 1]$, and then blindly separated. The mean value of the noise covariance matrix, i.e. $\mathbf{I}/3$, was used as the input parameter of the unbiased FastICA. Note that INSR and BER of this method were computed for solutions given by $\mathbf{W}^{\text{MMSE-UNB}}$ defined in (5).

The following table shows the average INSR and bit error rate (BER) that were achieved in 1000 trials. The performance of the proposed 1FICA is almost the same like that of "oracle" MMSE separator, because, here, N is very high, and the estimation error is caused by the bias only. The unbiased FastICA significantly suffers from inaccurate information about the noise intensity.

algorithm	average INSR [dB]	BER [%]
1FICA	**-5,98**	**3,19**
Symmetric FastICA	-5,68	3,55
unbiased FastICA	*6,79*	*5,25*
EFICA	-5,79	3,41
MMSE solution	**-5,98**	**3,19**
inversion solution	-4,76	4,71
JADE	-5,68	3,55

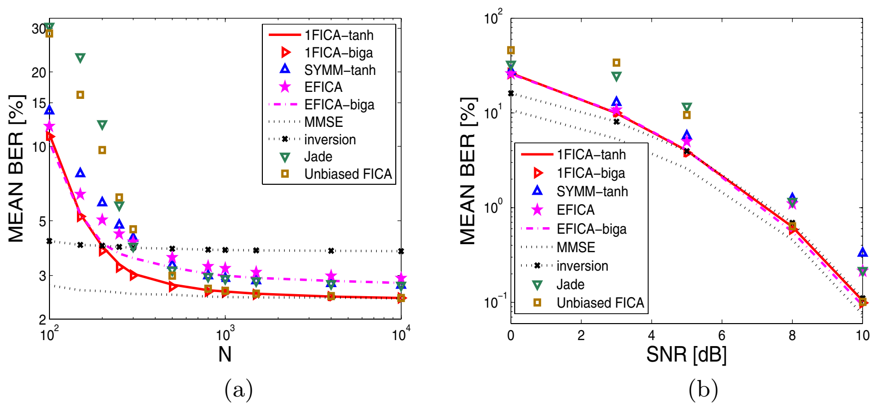

Fig. 1. Average BER of 10 separated BPSK signals when (a) SNR is fixed to 5dB and (b) a fixed number of data samples is $N = 200$. Averages are taken from 1000 independent trials for each settings.

5 Conclusions

This paper presents novel results from analysis of bias of several FastICA variants, whereby the one-unit FastICA was shown to be minimally biased from the MMSE solution, i.e., it achieves the best interference-plus-noise rejection rate for $N \to +\infty$.

By virtue of the theoretical results, a new variant of FastICA algorithm, called 1FICA, was derived to have the same global convergence as symmetric FastICA with the test of saddle points, and a noise rejection like the one-unit FastICA. Computer simulations show superior performance of the method when separating binary (BPSK) signals. Unlike the unbiased FastICA, it does not require prior knowledge of covariance of the noise to achieve the best MMSE separation.The Matlab codes for 1FICA in real and in complex domains can be downloaded from the first author's homepage, http://itakura.kes.tul.cz/zbynek/downloads.htm.

References

1. Bingham, E., Hyvärinen, A.: A fast fixed-point algorithm for independent component analysis of complex valued signals. Int. J. Neural Systems 10, 1–8 (2000)
2. Cardoso, J.-F., Souloumiac, A.: Blind Beamforming from non-Gaussian Signals. IEE Proc.-F 140, 362–370 (1993)
3. Cardoso, J.-F.: On the performance of orthogonal source separation algorithms. In: Proc. EUSIPCO, Edinburgh, pp. 776–779 (1994)
4. Davies, M.: Identifiability Issues in Noisy ICA. IEEE Signal Processing Letters 11, 470–473 (2005)
5. Hyvärinen, A., Oja, E.: A fast fixed-point algorithm for independent component analysis. Neural Computation 9, 1483–1492 (1997)
6. Hyvärinen, A.: Gaussian Moments for Noisy Independent Component Analysis. IEEE Signal Processing Letters 6, 145–147 (1999)
7. Koldovský, Z., Tichavský, P.: Methods of Fair Comparison of Performance of Linear ICA Techniques in Presence of Additive Noise. In: Proc. of ICASSP 2006, Toulouse, pp. 873–876 (2006)
8. Koldovský, Z., Tichavský, P., Oja, E.: Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound. IEEE Tr. Neural Networks 17, 1265–1277 (2006)
9. Koldovský, Z., Tichavský, P.: Asymptotic analysis of bias of FastICA-based algorithms in presence of additive noise, Technical report no 2181, ÚTIA, AV ČR (2007), Available at `http://itakura.kes.tul.cz/zbynek/publications.htm`
10. Oja, E., Yuan, Z.: The FastICA algorithm revisited: Convergence analysis. IEEE Trans. on Neural Networks 17, 1370–1381 (2006)
11. Tichavský, P., Koldovský, Z., Oja, E.: Performance Analysis of the FastICA Algorithm and Cramér-Rao Bounds for Linear Independent Component Analysis. IEEE Tr. on Signal Processing 54, 1189–1203 (2006)