Analytical solution of the blind source separation problem using derivatives

Sebastien Lagrange[1,2], Luc Jaulin[2], Vincent Vigneron[1], and Christian Jutten[1]

[1] Laboratoire Images et Signaux, Institut National Polytechnique de Grenoble, 46 avenue Félix Viallet, 38031 Grenoble, France.
`Vincent.Vigneron@lis.inpg.fr , Christian.Jutten@inpg.fr`
`http://www.lis.inpg.fr`

[2] Laboratoire LISA, Institut des Sciences et Techniques de l'Ingénieur, 62 avenue notre dame du lac, 49000 Angers.
`{sebastien.lagrange, luc.jaulin}@isita.univ-angers.fr`
`http://www.istia.univ-angers.fr`

Abstract. In this paper, we consider independence property between a random process and its first derivative. Then, for linear mixtures, we show that cross-correlations between mixtures and their derivatives provide a sufficient number of equations for analytically computing the unknown mixing matrix. In addition to its simplicity, the method is able to separate Gaussian sources, since it only requires second order statistics. For two mixtures of two sources, the analytical solution is given, and a few experiments show the efficiency of the method for the blind separation of two Gaussian sources.

1 Introduction

Blind source separation (BSS) consists in finding unknown sources $s_i(t)$, $i = 1, ..., n$ supposed statistically independent, knowing a mixture of these sources, called observed signals $x_j(t)$, $j = 1, ..., p$. In the literature, various mixtures have been studied : linear instantaneous [1–3] or convolutive mixtures [4–6], nonlinear and especially post-nonlinear mixtures [7, 8]. In this paper, we assume (i) the number of sources and observations are equal, $n = p$, (ii) the observed signals are linear instantaneous mixtures of the sources, *i.e,*

$$x_j(t) = \sum_{i=1}^{n} a_{ij} s_i(t), \quad j = 1, \ldots n. \tag{1}$$

In vector form, denoting the source vector $\mathbf{s}(t) = [s_1(t), ..., s_n(t)]^T \in \mathbb{R}^n$, and the observation vector $\mathbf{x}(t) = [x_1(t), ..., x_n(t)]^T \in \mathbb{R}^n$, the observed signal is

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \tag{2}$$

where $\mathbf{A} = [a_{ij}]$ is the $n \times n$ mixing matrix, assumed regular.

Without prior knowledge, the BSS problem can be solved by using independent component analysis (ICA) [9], which involves higher (than 2) order statistics, and requires that at most one source is Gaussian. With weak priors, like source coloration [10–12] or non-stationarity [13, 14], it is well known that BSS can be solved by jointly diagonalizing variance-covariance matrices, *i.e.* using only second order statistics, and thus allowing separation of Gaussian sources.

For square ($n \times n$) mixtures, the unknown sources can be indirectly estimated by estimating a separating matrix denoted $\mathbf{B}$, which provides a signal $\mathbf{y}(t) = \mathbf{B}\mathbf{x}(t)$ with independent components. However, it is well known that independence of the components of $\mathbf{y}(t)$ is not sufficient for estimating exactly $\mathbf{B} = \mathbf{A}^{-1}$, but only $\mathbf{BA} = \mathbf{DP}$, pointing out a scale (diagonal matrix $\mathbf{D}$) and permutation (permutation matrix $\mathbf{DP}$) indeterminacies [9]. It means that source power cannot be estimated. Thus, in the following, we will assumed unit power sources.

In this paper, we propose a new method based on second order statistics between the signals and their first derivatives. In Section 2, a few properties concerning statistical independence are derived. The main result is presented in Section 3, with the proof in Section 4, and a few experiments in Section 5, before the conclusion.

2 Statistical independence

In this section, we will introduce the main properties used below. For p random variables $x_1, \dots x_p$, a simple definition of independence is based on the factorisation of the joint density as the product of the marginal densities:

$$p_{x_1,\dots,x_p}(u_1, \dots, u_p) = \prod_{i=1}^{p} p_{x_1}(u_i). \tag{3}$$

We can also define the independence of random processes.

Definition 1. *Two random processes $x_1(t)$ and $x_2(t)$ are independent if and only if any random vectors, $x_1(t_1), ..., x_1(t_1 + k_1)$ and $x_2(t_2), ..., x_2(t_2 + k_2)$, $\forall t_i$, and k_j, $(i, j = 1, 2)$, extracted from them, are independent.*

Consequently, if two random signals (processes) $x_1(t)$ and $x_2(t)$ are statistically independent, then $\forall t_1, t_2$, $x_1(t_1)$ and $x_2(t_2)$ are statistically independent random variables, too [16].

Notation 1 *In the following, the independence between two random signals $x_1(t)$ and $x_2(t)$ will be denoted $x_1(t) \mathbb{I} x_2(t)$.*

Proposition 1. *Let $x_1(t), x_2(t), ..., x_n(t)$ and $u(t)$ be random signals. We have*

$$\begin{array}{ccc} x_1(t)\mathbb{I}u(t) & \Longrightarrow & u(t)\mathbb{I}x_1(t) \\ x_1(t)\mathbb{I}u(t),\dots,x_n(t)\mathbb{I}u(t) & \Longrightarrow & (x_1(t) + \dots + x_n(t))\,\mathbb{I}u(t)\,. \\ \forall\alpha\in\mathbb{R}, x_1(t)\mathbb{I}u(t) & \Longrightarrow & \alpha x_1(t)\mathbb{I}u(t) \end{array} \tag{4}$$

We now consider independence properties involving signals and their derivatives.

Lemma 1. *Let $x_1(t)$ and $x_2(t)$ be differentiable (with respect to t) signals. Then,*

$$x_1(t) \mathbb{I} x_2(t) \Longrightarrow \begin{cases} x_1(t) \ \mathbb{I} \ \dot{x}_2(t) \\ \dot{x}_1(t) \ \mathbb{I} \ x_2(t) \\ \dot{x}_1(t) \ \mathbb{I} \ \dot{x}_2(t) \end{cases} . \tag{5}$$

As a direct consequence, if $x_1(t)$ and $x_2(t)$ are sufficiently differentiable, for all $m, n \in \mathbb{N}$,

$$x_1(t) \mathbb{I} x_2(t) \Longrightarrow x_1^{(n)}(t) \mathbb{I} x_2^{(m)}(t). \tag{6}$$

Proof. If $x_1(t)$ and $x_2(t)$ are statistically independent then

$$x_1(t) \mathbb{I} x_2(t) \Longrightarrow \forall t_1, \forall t_2,\ x_1(t_1) \mathbb{I} x_2(t_2). \tag{7}$$

According to (4), $\forall t_1, \forall t_2$,

$$\left.\begin{array}{l} x_1(t_1) \mathbb{I} x_2(t_2) \\ x_1(t_1) \mathbb{I} x_2(t_2+\tau) \end{array}\right\} \Longrightarrow x_1(t_1) \mathbb{I} \frac{x_2(t_2) - x_2(t_2+\tau)}{\tau}. \tag{8}$$

Hence, since x_2 is differentiable with respect to t

$$\lim_{\tau \to 0} \frac{x_2(t_2) - x_2(t_2+\tau)}{\tau} < \infty, \tag{9}$$

and we have $\forall t_1, t_2,\ x_1(t_1) \mathbb{I} \dot{x}_2(t_2)$ where $\dot{x}(t)$ denotes the derivative of $x(t)$ with respect to t. Similar proof can be done for showing $\forall t_1, t_2, \dot{x}_1(t_1) \mathbb{I} x_2(t_2)$, and more generally $\forall t_1 \in \mathbb{R}, t_2 \in \mathbb{R},\ x_1^{(n)}(t_1) \mathbb{I} x_2^{(m)}(t_2)$.

Lemma 2. *Let $x(t)$ be a differentiable signal with the auto-correlation function $\gamma_{xx}(\tau) = E(x(t)x(t-\tau))$, then $E(x\dot{x}) = 0$.*

Proof. Since $x(t)$ is derivable, its autocorrelation function is derivable in zero:

$$\begin{aligned} \dot{\gamma}_{xx}(0) &= \lim_{\tau \to 0} \frac{\gamma_{xx}(0) - \gamma_{xx}(\tau)}{-\tau} && (10) \\ &= \lim_{\tau \to 0} \frac{E(x(t)x(t)) - E(x(t)x(t-\tau))}{-\tau} && (11) \\ &= \lim_{\tau \to 0} E(x(t)(\frac{x(t) - x(t-\tau)}{-\tau})) && (12) \\ &= E(x(t). \lim_{\tau \to 0}(\frac{x(t) - x(t-\tau)}{-\tau})) = -E(x\dot{x}). && (13) \end{aligned}$$

Finally, since γ_{xx} is even, $\dot{\gamma}_{xx}(0) = 0$, and consequently $E(x\dot{x}) = 0$.

Lemma 3. *If $\mathbf{x} = \mathbf{A}\mathbf{s}$, where component s_i of $\mathbf{s}$ are mutually independent, then $E(\mathbf{x}\dot{\mathbf{x}}^T) = 0$.*

Proof. Since $\mathbf{x} = \mathbf{A}\mathbf{s}$, we have $E(\mathbf{x}\dot{\mathbf{x}}^T) = \mathbf{A}E(\mathbf{s}\dot{\mathbf{s}}^T)\mathbf{A}^T$

Using Lemmas 2 and 1, one has $E(s_i\dot{s}_i) = 0$ and $E(s_i\dot{s}_j) = 0$, respectively. Consequently, $E(\mathbf{x}\dot{\mathbf{x}}^T) = 0$.

3 Theorem

In this section, we present the main result of the paper. The proof will be shown in the next section (4).

First, let us define the set $\mathcal{T}$ of trivial linear mixings, *i.e.* linear mappings which preserve independence for any distributions. One can show that $\mathcal{T}$ is set of square regular matrices which are the product of one diagonal regular matrix and one permutation matrix. In other words, $\mathbf{B}$ is a separating matrix if $\mathbf{BA} = \mathbf{DP} \in \mathcal{T}$.

Theorem 1. *Let $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$, be an unknown regular mixture of sources $\mathbf{s}(t)$, whose components $s_i(t)$ are ergodic, stationary, derivable and mutually independent signals, the separating matrices $\mathbf{B}$, such that $\mathbf{y}(t) = \mathbf{B}\mathbf{x}(t)$ has mutually independent components, are the solutions of the equation set:*

$$\begin{aligned}\mathbf{B}E(\mathbf{x}\mathbf{x}^T)\mathbf{B}^T &= E(\mathbf{y}\mathbf{y}^T)\\ \mathbf{B}E(\dot{\mathbf{x}}\dot{\mathbf{x}}^T)\mathbf{B}^T &= E(\dot{\mathbf{y}}\dot{\mathbf{y}}^T)\end{aligned}$$

where $E(\mathbf{y}\mathbf{y}^T)$ and $E(\dot{\mathbf{y}}\dot{\mathbf{y}}^T)$ are diagonal matrices.

4 Proof of theorem

The proof is given for 2 mixtures of 2 sources. It will be admitted in the general case.

The estimated sources are $\mathbf{y} = \mathbf{B}\mathbf{x}$ where $\mathbf{B}$ is a separating matrix of $\mathbf{A}$. After derivation, one has a second equation: $\dot{\mathbf{y}} = \mathbf{B}\dot{\mathbf{x}}$.

The independence assumption of the estimated sources $\mathbf{y}$ implies that the following matrix is diagonal:

$$E\left(\mathbf{y}\mathbf{y}^T\right) = E\left(\mathbf{B}\mathbf{x}\mathbf{x}^T\mathbf{B}^T\right) = \mathbf{B}E\left(\mathbf{x}\mathbf{x}^T\right)\mathbf{B}^T. \tag{14}$$

Moreover, using Lemma 1 and 2, the following matrix is diagonal, too:

$$E\left(\dot{\mathbf{y}}\dot{\mathbf{y}}^T\right) = E\left(\mathbf{B}\dot{\mathbf{x}}\dot{\mathbf{x}}^T\mathbf{B}^T\right) = \mathbf{B}E\left(\dot{\mathbf{x}}\dot{\mathbf{x}}^T\right)\mathbf{B}^T. \tag{15}$$

By developing the vectorial equations (14) and (15), one gets four scalar equations:

$$\begin{cases} b_{11}^2 E\left(x_1^2\right) + 2b_{11}b_{12}E\left(x_1x_2\right) + b_{12}^2 E\left(x_2^2\right) = E\left(y_1^2\right)\\ b_{21}^2 E\left(x_1^2\right) + 2b_{21}b_{22}E\left(x_1x_2\right) + b_{22}^2 E\left(x_2^2\right) = E\left(y_2^2\right)\\ b_{11}b_{21}E\left(x_1^2\right) + b_{11}b_{22}E\left(x_1x_2\right) + b_{12}b_{21}E\left(x_1x_2\right) + b_{12}b_{22}E\left(x_2^2\right) = 0\\ b_{11}b_{21}E\left(\dot{x}_1^2\right) + b_{11}b_{22}E\left(\dot{x}_1\dot{x}_2\right) + b_{12}b_{21}E\left(\dot{x}_1\dot{x}_2\right) + b_{12}b_{22}E\left(\dot{x}_2^2\right) = 0. \end{cases} \tag{16}$$

This system is a set of polynomials with respect to the b_{ij}. It has six unknowns for only four equations. In fact, the two unknowns, $E(y_1^2)$ and $E(y_2^2)$ are not relevant, due to the scale indeterminacies of source separation. Since the source power cannot be estimated, we can then consider this unknows as parameters, or even constraint these parameter to be equal to a constant (e.g. 1). Here, we parameterize the set of solutions (two dimensional manifold) by two real parameters λ_1 and λ_2 such that $|\lambda_1|=E\left(y_1^2\right)$ and $|\lambda_2|=E\left(y_2^2\right)$ in Eq. (16) .

Groëbner Basis decomposition [15] give the solutions

$$\mathbf{B} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \phi_1 & \phi_1\eta_1 \\ \phi_2 & \phi_2\eta_2 \end{pmatrix} \quad \text{or} \quad \mathbf{B} = \begin{pmatrix} 0 & \lambda_1 \\ \lambda_2 & 0 \end{pmatrix} \begin{pmatrix} \phi_1 & \phi_1\eta_1 \\ \phi_2 & \phi_2\eta_2 \end{pmatrix}, \tag{17}$$

where

$$\phi_1 = \left(E(x_1^2) + 2\eta_1 E(x_1x_2) + \eta_1^2 E(x_2^2)\right)^{-\frac{1}{2}}, \tag{18}$$

$$\phi_2 = \left(E(x_1^2) + 2\eta_2 E(x_1x_2) + \eta_2^2 E(x_2^2)\right)^{-\frac{1}{2}}, \tag{19}$$

$$\eta_1 = -\beta\left(1 + \sqrt{1 - \frac{\alpha}{\beta^2}}\right), \tag{20}$$

$$\eta_2 = -\beta\left(1 - \sqrt{1 - \frac{\alpha}{\beta^2}}\right), \tag{21}$$

$$\alpha = \frac{E(x_1^2)E(\dot{x}_1\dot{x}_2) - E(x_1x_2)E(\dot{x}_1^2)}{E(x_1x_2)E(\dot{x}_2^2) - E(x_2^2)E(\dot{x}_1\dot{x}_2)}, \tag{22}$$

$$\beta = \frac{1}{2}\frac{E(x_1^2)E(\dot{x}_2^2) - E(x_2^2)E(\dot{x}_1^2)}{E(x_1x_2)E(\dot{x}_2^2) - E(x_2^2)E(\dot{x}_1\dot{x}_2).} \tag{23}$$

Then, let

$$\tilde{\mathbf{B}} = \begin{pmatrix} \phi_1 & \phi_1\eta_1 \\ \phi_2 & \phi_2\eta_2 \end{pmatrix}, \tag{24}$$

any matrix $\mathbf{T}\tilde{\mathbf{B}}$ where $\mathbf{T} \in \mathcal{T}$, is still a solution of (16). Especially, it exists a particular matrix $\tilde{\mathbf{T}} \in \mathcal{T}$ with $\lambda_1 = E(y_1^2)$ and $\lambda_2 = E(y_2^2)$ such that:

$$\mathbf{A}^{-1} = \tilde{\mathbf{T}}\tilde{\mathbf{B}}. \tag{25}$$

Thus, all the possible separating matrices are solutions of (16), and the Theorem 1 is proved.

5 Experiments

Consider two independent Gaussian sources obtained by Gaussian white noises filtered by low-pass second-order filters. Filtering ensures to obtain differentiable sources by preserving the Gaussian distribution of the sources.

The two sources are depicted on Figures 1 and 2 ($90,000$ samples, sampling period $Te = 0.1$). The corresponding joint distribution is represented on Figure 3.

The derivative joint distribution (Fig. 4) shows the two signals $\dot{s}_1(t)$ and $\dot{s}_2(t)$ are independent (as predicted by Lemma 2).

The mixing matrix

$$\mathbf{A} = \begin{pmatrix} 1.5 \; 0.8 \\ 0.6 \; 1.2 \end{pmatrix}, \tag{26}$$

provides the observation signals (mixtures) $x_1(t)$ and $x_2(t)$. The joint distribution of the mixtures (Fig. 4) and of their derivatives (Fig. 5) points out the statistical dependence of $x_1(t)$ and $x_2(t)$ as well as $\dot{x}_1(t)$ and $\dot{x}_2(t)$.

From the Theorem 1, the separation matrix $\tilde{\mathbf{B}}$ is analytically computed:

$$\tilde{\mathbf{B}} = \begin{pmatrix} 0.456 & -.606 \\ -0.454 & 1.136 \end{pmatrix}. \tag{27}$$

The estimated sources $\tilde{\mathbf{B}}\mathbf{x}(t)$ are independent and unit power. We also can check that there exists $\tilde{\mathbf{T}} \in \mathcal{T}$ such as $\tilde{\mathbf{T}}\tilde{\mathbf{B}}\mathbf{x}(t) = \mathbf{s}(t)$, *i.e.* $\tilde{\mathbf{T}}\tilde{\mathbf{B}} = \mathbf{A}^{-1}$. In this example, let

$$\tilde{\mathbf{T}} = \begin{pmatrix} 2 \; 0 \\ 0 \; 2 \end{pmatrix} \in \mathcal{T}, \; (\lambda_1 = \lambda_2 = 2), \tag{28}$$

we have

$$\tilde{\mathbf{T}}\tilde{\mathbf{B}} = \begin{pmatrix} 0.912 & -.613 \\ -0.453 & 1.135 \end{pmatrix}, \tag{29}$$

and

$$\mathbf{A}^{-1} = \begin{pmatrix} 0.909 & -.606 \\ -0.454 & 1.136 \end{pmatrix}. \tag{30}$$

In order to study the robustness of the solution for Gaussian mixtures, Fig. 7 shows the separation performance (using the index: $E(norm(\mathbf{s} - \hat{\mathbf{s}}))$) versus the sample number. Over 2600 samples, the analytical solution provides good performance, with an error less than -20 dB.

6 Conclusion

In this article, we proposed a new source separation criterion based on variance-covariance matrices of observations and of their derivatives. Since the method only uses second-order statistics, source separation of Gaussians remains possible. The main (and weak) assumption is the differentiability of the unknown sources.

We derived the analytical solution for two mixtures of two sources. In the general case (n mixtures of n sources), the analytical solution seems tricky to compute. Moreover, for an ill-conditioned set of equations (16), the analytical solution can be very sensitive to the statistical moment estimations. For overcoming this problem, we could estimate the solution, by using approximate joint diagonalization algorithm of $E(\mathbf{x}\mathbf{x}^T)$ and $E(\dot{\mathbf{x}}\dot{\mathbf{x}}^T)$. Moreover, other variance-covariance matrices, based on higher-order derivatives or using different delays (assuming sources are colored) or on different temporal windows (assuming sources are non stationary), could be used for estimating the solution by joint diagonalization.

Further investigations include implementation of joint diagonalization, and extension to more complex signals mixtures, *e.g.* based on state variable models.

References

1. Jutten, C., Hérault, J.: Blind Separation of Sources, Part I: an Adaptive Algorithm Based on a Neuromimetic Architecture. Signal Processing, Vol. 24 (1991) 1–10
2. Cardoso, J.-F., Souloumiac, A.: Blind beamforming for non Gaussian signals. IEE Proceedings-F, Vol. 140 (1993) 362–370
3. Bell, T., Sejnowski, T.: An Information-Maximization Approach to Blind Separation and Blind Deconvolution. Neural Comutation, Vol. 7 (1995) 1004-1034
4. Yellin, D., Weinstein, E.: Criteria for multichannel signal separation. IEEE Trans. on Signal Processing, Vol. 42 (1994) 2158–2168
5. Nguyen Thi, H. L., Jutten, C.: Blind Sources Separation For Convolutive mixtures. Signal Processing, Vol. 45 (1995) 209–229
6. Babaie-Zadeh, M.: On blind source separation in convolutive and nonlinear mixtures. PhD thesis, INP Grenoble (2002)
7. Pajunen, P., Hyvärinen, A., Karhunen, J.: Non linear source separation by self-organizing maps. Proceedings of ICONIP 96, Hong-Kong (1996)
8. Taleb, A., Jutten, C.: Source separation in post nonlinear mixtures. IEEE Transactions on Signal Processing, Vol. 47 (1999) 2807–2820
9. Comon, P.: Independent Component Analysis, a new concept? Signal Processing, Vol. 36 (1994) 287–314
10. Tong, L., Soon, V., Liu, R., Huang, Y.: AMUSE: a new blind identification algorithm. Proceedings ISCAS, New Orleans, USA (1990)
11. Belouchrani, A., Abed Meraim, K., Cardoso, J.-F., Moulines, E.: A blind source separation technique based on second order statistics. IEEE Trans. on Signal Processing, Vol. 45 (1997) 434–444
12. Molgedey, L., Schuster, H. G.: Separation of a mixture of independent signals using time delayed correlation. Physical Review Letters, Vol. 72 (1994) 3634–3636
13. Matsuoka, K., Ohya, M., Kawamoto, M.: A neural net for blind separation of nonstationary signals. Neural Networks, Vol. 8 (1995) 411–419
14. Pham, D. T., Cardoso, J.-F.: Blind Separation of Instantaneous Mixtures of Non Stationary Sources. IEEE Transaction on Signal Processing, Vol. 49 (2001) 1837–1848
15. Adams, W. W., Loustaunau, P.: An Introduction to Grbner Bases. Providence, RI: Amer. Math. Soc., 1994
16. Papoulis, A.: Probability, Random Variables, and Stochastic Processes. Third Edition, McGraw-Hill, International Editions. Electronical and Electronic Engineering Series.

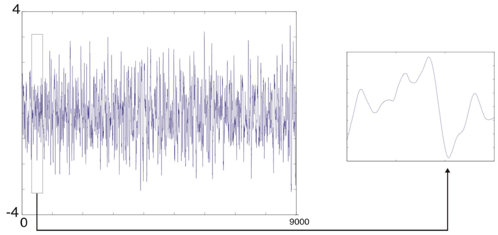

Fig. 1. Source $s_1(t)$

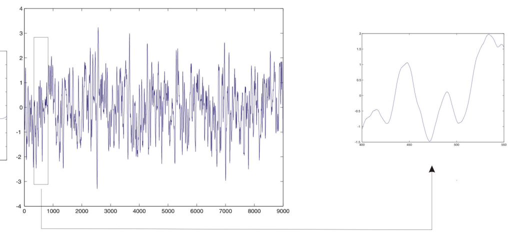

Fig. 2. Source $s_2(t)$

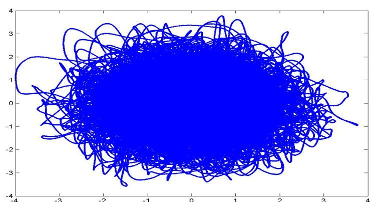

Fig. 3. Distribution of the sources $(s_1(t), s_2(t))$

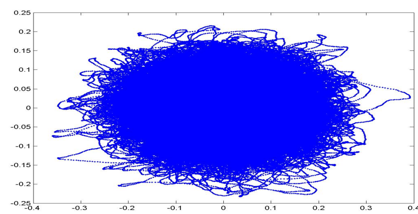

Fig. 4. Distribution of the signals $(\dot{s}_2(t), \dot{s}_1(t))$

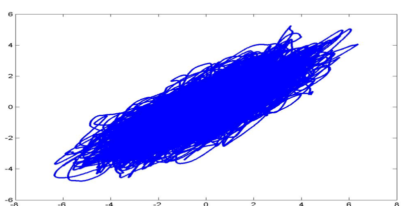

Fig. 5. Joint distribution of the mixtures $(x_1(t), x_2(t))$

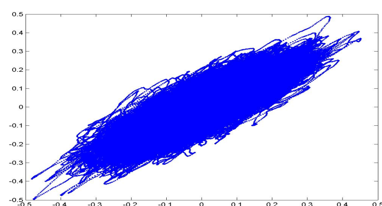

Fig. 6. Joint distribution of the mixture derivatives $(\dot{x}_1(t), \dot{x}_2(t))$

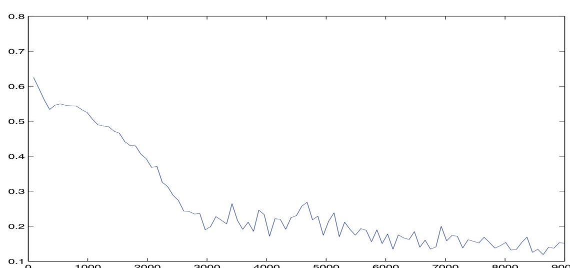

Fig. 7. The error estimation of the sources according to the number of samples