On the Separability of Nonlinear Mixtures of Temporally Correlated Sources

Shahram Hosseini and Christian Jutten, *Member, IEEE*

***Abstract*—It is well known that the source separation in nonlinear case is, in general, impossible: there exist many mappings with nondiagonal Jacobian matrices preserving the independence. In this letter, we suggest that using the time structure of the sources (if it exists), this indeterminacy may be reduced. In particular, we show that the classical examples used in the literature for demonstrating the nontrivial nonseparability of the nonlinear mixtures can be rejected by taking into account the temporal correlation of the sources.**

***Index Terms*—Independent component analysis (ICA), nonlinear mixtures, separability, source separation.**

I. Introduction

A NONLINEAR source separation problem can be formulated as follows. Consider a K-sensor array providing the signal $\mathbf{x}(t) = [x_1(t), x_2(t), \cdots, x_K(t)]$ which is a nonlinear instantaneous mixture of K unknown statistically independent sources $\mathbf{s}(t) = [s_1(t), s_2(t), \cdots, s_K(t)]$: $\mathbf{x}(t) = \mathcal{F}(\mathbf{s}(t))$, where the mapping $\mathcal{F}$ is an unknown differentiable bijective mapping. Is it possible, using only the statistical independence assumption, to recover the sources $\mathbf{s}$ from the nonlinear mixture? The answer is clearly no: if X and Y are two independent variables and f_1 and f_2 two arbitrary functions, $f_1(X)$ and $f_2(Y)$ are independent too. Hence, the sources may be recovered only up to any nonlinear distortion. This indeterminacy is characterized by the mappings with diagonal Jacobian and is called trivial indeterminacy. If it was the only indeterminacy, it could be tolerable because the sources would be separated. The problem is however more serious because X and Y can be mixed together and still be independent. There exist many separating mappings $\mathcal{G}$, providing a random vector $\mathbf{y}$ with independent components

$$\mathbf{y}(\mathbf{t}) = \mathcal{G}(\mathbf{x}(\mathbf{t})) = (\mathcal{G} \circ \mathcal{F})(\mathbf{s}(t)) = \mathcal{H}(\mathbf{s}(t)) \tag{1}$$

so that the overall mapping $\mathcal{H}$ has a nondiagonal Jacobian. A simple example derived from [1] is the following. Suppose $s_1(t) \in \mathcal{R}^+$ is a Rayleigh distributed variable with probability density function (pdf) $P_{s1}(s_1(t)) = s_1(t)\exp(-s_1^2(t)/2)$ and $s_2(t) \in [0, 2\pi)$ is uniform and independent of $s_1(t)$. Consider the nonlinear mapping

$$\begin{aligned}[y_1(t), y_2(t)] &= \mathcal{H}(s_1(t), s_2(t)) \\ &= [s_1(t)\cos(s_2(t)), s_1(t)\sin(s_2(t))]\end{aligned} \tag{2}$$

which has a nondiagonal Jacobian matrix

$$\mathbf{J} = \begin{pmatrix} \cos(s_2(t)) & -s_1(t)\sin(s_2(t)) \\ \sin(s_2(t)) & s_1(t)\cos(s_2(t)) \end{pmatrix}.$$

The joint pdf of $y_1(t)$ and $y_2(t)$ is

$$\begin{aligned}p_{y_1,y_2}(y_1(t), y_2(t)) &= \frac{p_{s_1,s_2}(s_1(t), s_2(t))}{|\mathbf{J}|} \\ &= \frac{1}{2\pi}\exp\left(\frac{-y_1^2(t) - y_2^2(t)}{2}\right) \\ &= \left(\frac{1}{\sqrt{2\pi}}\exp\frac{-y_1^2(t)}{2}\right) \\ &\quad \cdot \left(\frac{1}{\sqrt{2\pi}}\exp\frac{-y_2^2(t)}{2}\right).\end{aligned} \tag{3}$$

Relation (3) shows that the two random variables $y_1(t)$ and $y_2(t)$ are independent and Gaussian. We show now if the sources are temporally correlated, another test may allow to reject this mapping. It consists in verifying the equality

$$E[y_1(t+1)y_2(t)] = E[y_1(t+1)]E[y_2(t)] \tag{4}$$

on the recovered independent components. It is evident that if $y_1(t)$ and $y_2(t)$ are the actual sources (or a trivial mapping of them), the above equality is true. On the other hand, if the independent components are obtained from (2), the right side of (4) is equal to zero because y_1 and y_2 are zero mean Gaussian variables. The left side of (4) is equal to

$$\begin{aligned}&E[s_1(t+1)\cos(s_2(t+1))\,s_1(t)\sin(s_2(t))] \\ &\quad = E[s_1(t+1)s_1(t)]E[\cos(s_2(t+1))\sin(s_2(t))].\end{aligned} \tag{5}$$

If $s_1(t)$ and $s_2(t)$ are temporally correlated, it is highly probable that (5) is not zero (it depends evidently on the nature of the temporal correlation between two successive samples of the two sources) so that the equality (4) is false, and the solution can be rejected. In fact, the two *stochastic processes* $y_1(t)$ and $y_2(t)$ obtained from (2), are not statistically independent although their samples at each time instant (which are two *random variables*) are independent. This simple example shows how using the temporal correlation, we can distinguish the trivial and nontrivial mappings preserving the independence. Note that here we used only the first lag cross correlation of the signals. A more rigorous approach consists in testing the independence of $y_1(t+1)$

Manuscript received May 16, 2002; revised July 18, 2002. This work has supported in part by the European project BLInd Source Separation and applications (BLISS) under Contract IST 1999-14190. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Mounir Ghogho.

S. Hosseini was with LIS-INPG, Grenoble, France. He is now with Laboratoire d'Acoustique, Metrologie, Instrumentation (LAMI), Université Toulouse III (Paul Sabatier), 31062 Toulouse, France (e-mail: hosseini@cict.fr).

C. Jutten is with Laboratoire des Images et des Signaux (UMR CNRS 5083, INPG, UJF), 38031 Grenoble, France (e-mail: Christian.Jutten@inpg.fr).

Digital Object Identifier 10.1109/LSP.2002.807871

and $y_2(t)$, which may be even generalized to other time lags. In the following section, we study another classical example using this test.

II. Darmois Decomposition

Another example used in [1] and [2] for illustrating the nontrivial nonseparability of the nonlinear mixtures is the Darmois decomposition procedure. This constructive method permits to decompose a given vector into independent factors. For the case of two sources, it is enough to choose

$$\begin{aligned} y_1(t) &= g_1(x_1(t), x_2(t)) = F_{X_1}(x_1(t)) \\ y_2(t) &= g_2(x_1(t), x_2(t)) = F_{X_2|X_1}(x_1(t), x_2(t)) \end{aligned} \tag{6}$$

where F_{X_1} and $F_{X_2|X_1}$ are, respectively, the marginal and conditional cumulative distribution functions. It can be easily verified that the two random variables $y_1(t)$ and $y_2(t)$ are independent [2]. This construction also clearly shows that the decomposition in independent components is by no means unique. For example, we could first apply a linear transformation on the observation vector $\mathbf{x}$ to obtain another random vector $\mathbf{x}' = \mathbf{L}\mathbf{x}$, and then compute $\mathbf{y}' = \mathbf{g}'(\mathbf{x}')$ with $\mathbf{g}'$ being defined using the above procedure, where $\mathbf{x}$ is replaced by $\mathbf{x}'$. Thus, we obtain another decomposition of $\mathbf{x}$ into independent components. The resulting decomposition $\mathbf{y}' = \mathbf{g}'(\mathbf{L}\mathbf{x})$ is in general different from $\mathbf{y}$, and cannot be reduced to $\mathbf{y}$ by any simple transformation. A more rigorous justification of the nonuniqueness property has been given in [2].

We show now if the sources s_1 and s_2 are temporally correlated, the independent components y_1 end y_2 obtained from the above procedure do not generally satisfy the following equality where p_{Y_1} and p_{Y_2} are the marginal pdfs and p_{Y_1,Y_2} is the joint pdf:

$$p_{Y_1,Y_2}(y_1(t+1)y_2(t)) = p_{Y_1}(y_1(t+1))\, p_{Y_2}(y_2(t)) \tag{7}$$

while the trivial transformations of the real sources, in the forms of $y_1 = f_1(s_1)$ and $y_2 = f_2(s_2)$, satisfy obviously the above equality because of the independence of the two sources. Thus, the above test can be used to reject the nontrivial independent component analysis (ICA) solutions obtained from the Darmois decomposition. For the sake of clarity, we study only the case of two sources, two observations.

A. General Case

Supposing the Darmois transformation (6), we want to prove that (7) is not generally true if $p(s_1(t+1)s_1(t)) \neq p(s_1(t+1))p(s_1(t))$ and $p(s_2(t+1)s_2(t)) \neq p(s_2(t+1))p(s_2(t))$. In other words, we want to show that

$$\begin{aligned} &p(F_{X_1}(x_1(t+1)), F_{X_2|X_1}(x_1(t), x_2(t))) \\ &\qquad \neq p(F_{X_1}(x_1(t+1)))\, p\left(F_{X_2|X_1}(x_1(t), x_2(t))\right). \end{aligned} \tag{8}$$

The right side of the above relation is equal to

$$\begin{aligned} &p(F_{X_1}(x_1(t+1)))\, p\left(F_{X_2|X_1}(x_1(t), x_2(t))\right) \\ &\quad = \frac{p(x_1(t+1))}{\left|\frac{\partial F_{X_1}(x_1(t+1))}{\partial x_1(t+1)}\right|} p\left(F_{X_2|X_1}(x_1(t), x_2(t))\right) \\ &\quad = p\left(F_{X_2|X_1}(x_1(t), x_2(t))\right). \end{aligned} \tag{9}$$

Denoting $z = F_{X_2|X_1}(x_1(t), x_2(t))$ and using the auxiliary variable $w = x_1(t)$, we can write

$$\begin{aligned} p_{ZW}(z,w) &= \frac{p_{X_1,X_2}(x_1(t), x_2(t))}{\begin{vmatrix} \partial z/\partial x_1(t) & p_{X_2|X_1}(x_1(t), x_2(t)) \\ 1 & 0 \end{vmatrix}} \\ &= \frac{p_{X_1,X_2}(x_1(t), x_2(t))}{p_{X_2|X_1}(x_1(t), x_2(t))} = p_{X_1}(x_1(t)). \end{aligned} \tag{10}$$

Thus, we have

$$\begin{aligned} p(F_{X_2|X_1}(x_1(t), x_2(t))) &= p_Z(z) \\ &= \int_{-\infty}^{+\infty} p_{ZW}(z,w)dw = \int_{-\infty}^{+\infty} p_{X_1}(w)dw = 1. \end{aligned}$$

The right side of (8) is so equal to 1. To compute the left side of (8), denoting $w = F_{X_1}(x_1(t+1))$ and $v = F_{X_2|X_1}(x_1(t), x_2(t))$, and using the auxiliary variable $z = x_1(t)$, we can write

$$\begin{aligned} &p_{Z,W,V}(z,w,v) \\ &= \frac{p_{X_1,X_1,X_2}(x_1(t+1), x_1(t), x_2(t))}{\begin{vmatrix} 0 & 1 & 0 \\ p_{X_1}(x_1(t+1)) & 0 & 0 \\ 0 & \partial v/\partial x_1(t) & p_{X_2|X_1}(x_1(t), x_2(t)) \end{vmatrix}} \\ &= \frac{p_{X_1,X_1,X_2}(x_1(t+1), x_1(t), x_2(t))}{p_{X_1}\left(x_1(t+1)\right) p_{X_2|X_1}(x_1(t), x_2(t))}. \end{aligned} \tag{11}$$

Hence, the left side of (8) is equal to

$$\begin{aligned} &p(F_{X_1}(x_1(t+1)), F_{X_2|X_1}(x_1(t), x_2(t))) \\ &\qquad = p_{W,V}(w,v) = \int_{-\infty}^{+\infty} p_{Z,W,V}(z,w,v)dz. \end{aligned}$$

Denoting $a = x_1(t+1) = q(w)$, $b = x_1(t) = z$, and $c = x_2(t) = g(v,z)$ and using (11), the left side of (8) can be written as

$$\begin{aligned} &\int_{-\infty}^{+\infty} \frac{p_{A,B,C}(a,b,c)}{p_A(a)p_{C|B}(b,c)} dz \\ &= \int_{-\infty}^{+\infty} \frac{p_{A,B,C}(a,b,c)}{p_A(a)p_{B,C}(b,c)/p_B(b)} dz \\ &= \int_{-\infty}^{+\infty} \frac{1}{p_A(a)} p_{A|B,C}(a,b,c)p_B(b)dz \\ &= \int_{-\infty}^{+\infty} \frac{p_{A|B,C}(q(w), z, g(v,z))p_B(z)}{p_A(q(w))} dz \\ &= \frac{\int_{-\infty}^{+\infty} p_{A|B,C}(q(w), z, g(v,z))p_B(z)dz}{p_A(q(w))} \\ &= \frac{\int_{-\infty}^{+\infty} p_{A|B,C}(a,b,c)p_B(b)db}{p_A(a)}. \end{aligned} \tag{12}$$

In order that the two sides of (8) are not equal, the last expression must not be equal to one, or in the other words, the following inequality must be true:

$$\int_{-\infty}^{+\infty} p_{A|B,C}(a,b,c)p_B(b)db \neq p_A(a). \qquad (13)$$

In the general case where A, B, and C are mutually dependent (which is our case[1]), it is evident that the left side of (13) is, in general, a function of both a and c, and so different to the right side of (13). In Section II-B, we verify this fact in a special case of Gaussian sources with first-order autoregressive correlation structure.

B. Special Case of Gaussian Distribution

In this section, we verify (13) for a simple example. Suppose $x_1(t)$ and $x_2(t)$ are generated using the linear procedure[2] $x_1(t) = s_1(t)+s_2(t), x_2(t) = s_1(t)-s_2(t)$, where $s_i(t+1) = \rho_i s_i(t) + n_i(t), i = 1, 2$, and $n_i(t)$ are zero-mean independent and identically distributed Gaussian sequences with unit variance. Using the above model, we can write

$$\begin{aligned} s_1(t) &\sim \mathcal{N}(0, 1/(1-\rho_1^2)) \\ s_2(t) &\sim \mathcal{N}(0, (1/(1-\rho_2^2))) \\ n_1(t)+n_2(t) &\sim \mathcal{N}(0, 2) \\ s_1(t)+s_2(t) &\sim \mathcal{N}(0, (1/(1-\rho_1^2)) + (1/(1-\rho_2^2)) \end{aligned}$$

and

$$\begin{aligned} s_1(t) &= (x_1(t)+x_2(t))/2 \\ s_2(t) &= (x_1(t)-x_2(t))/2. \end{aligned}$$

We want at present to verify the following inequality:

$$\int_{-\infty}^{+\infty} p_{x_1(t+1)|x_1(t),x_2(t)}(a,b,c)p_{x_1(t)}(b)db \neq p_{x_1(t+1)}(a). \qquad (14)$$

The right side of (14) is equal to

$$\begin{aligned} R &= p_{s_1(t+1)+s_2(t+1)}(a) \\ &= \frac{1}{\sqrt{2\pi\left(\frac{1}{1-\rho_1^2}+\frac{1}{1-\rho_2^2}\right)}} \exp\left\{\frac{-a^2}{2\left(\frac{1}{1-\rho_1^2}+\frac{1}{1-\rho_2^2}\right)}\right\}. \qquad (15) \end{aligned}$$

To compute the left side, we write

$$\begin{aligned} &p_{x_1(t+1)|x_1(t),x_2(t)}(a,b,c) \\ &\quad = p_{s_1(t+1)+s_2(t+1)|s_1(t),s_2(t)}\left(a, \frac{b+c}{2}, \frac{b-c}{2}\right) \\ &\quad = p_{n_1(t+1)+n_2(t+1)}\left(a-\rho_1\frac{b+c}{2}-\rho_2\frac{b-c}{2}\right) \\ &\quad = \frac{1}{\sqrt{4\pi}} \exp\left\{\frac{-1}{4}\left(a-\rho_1\frac{b+c}{2}-\rho_2\frac{b-c}{2}\right)^2\right\}. \qquad (16) \end{aligned}$$

We know also that

$$\begin{aligned} p_{x_1(t)}(b) &= p_{s_1(t)+s_2(t)}(b) \\ &= 1/\sqrt{2\pi(1/1-\rho_1^2+1/1-\rho_2^2)} \\ &\quad \cdot \exp\{-b^2/2(1/1-\rho_1^2+1/1-\rho_2^2)\}. \end{aligned}$$

Using these relations, the left side of (14) can be written in the following manner:

$$\begin{aligned} L &= \int_{-\infty}^{+\infty} K_1 \exp\left\{-(K_2-K_3b)^2 - K_4b^2\right\} db \\ &= K_1 \exp\left(-K_2^2\right) \int_{-\infty}^{+\infty} \exp\left\{-\left(K_3^2+K_4\right)b^2 + 2K_2K_3b\right\} db \end{aligned} \qquad (17)$$

where

$$\begin{aligned} K_1 &= (1/\sqrt{4\pi})\, 1/\left(\sqrt{2\pi(1/1-\rho_1^2+1/1-\rho_2^2)}\right) \\ K_2 &= (1/2)(a-\rho_1 c/2+\rho_2 c/2) \\ K_3 &= (1/2)\left((\rho_1+\rho_2)/2\right) \\ K_4 &= (1/2)(1/1-\rho_1^2+1/1-\rho_2^2). \end{aligned}$$

Comparing with the Gaussian density formula

$$\exp\{-(b-m)^2/2\sigma^2\} = \exp\{-b^2/2\sigma^2 + mb/\sigma^2 - m^2/2\sigma^2\}$$

and denoting

$$\begin{aligned} \sigma^2 &= 1/2(K_3^2+K_4) \\ m/\sigma^2 &= 2K_2K_3 \\ m^2/2\sigma^2 &= (K_2^2K_3^2)/(K_3^2+K_4) \end{aligned}$$

(17) can be rewritten as

$$\begin{aligned} L &= K_1 \exp\left(-K_2^2\right) \int_{-\infty}^{+\infty} \exp\left\{\frac{-b^2}{2\sigma^2}+\frac{mb}{\sigma^2}-\frac{m^2}{2\sigma^2}+\frac{m^2}{2\sigma^2}\right\} db \\ &= K_1 \exp\left(-K_2^2+\frac{m^2}{2\sigma^2}\right)\sqrt{2\pi\sigma^2} \int_{-\infty}^{+\infty} \frac{\exp\left\{\frac{-(b-m)^2}{2\sigma^2}\right\}}{\sqrt{2\pi\sigma^2}} db \\ &= K_1\sqrt{2\pi\sigma^2}\exp\left\{-K_2^2+\frac{m^2}{2\sigma^2}\right\}. \end{aligned} \qquad (18)$$

Replacing the constants in this last expression, we obtain

$$\begin{aligned} L &= \frac{1}{\sqrt{4\pi\left(1+\frac{1}{8}(\rho_1+\rho_2)^2\left(\frac{1}{1-\rho_1^2}+\frac{1}{1-\rho_2^2}\right)\right)}} \\ &\quad \cdot \exp\left\{\frac{-1}{4}\left(a-\frac{\rho_1-\rho_2}{2}c\right)^2\right. \\ &\quad \left.\cdot \left(\frac{1}{1+\frac{1}{8}(\rho_1+\rho_2)^2\left(\frac{1}{1-\rho_1^2}+\frac{1}{1-\rho_2^2}\right)}\right)\right\}. \end{aligned} \qquad (19)$$

Thus, the left side of (14) is a function of c and cannot be equal to the right side of (14), unless $\rho_1 = \rho_2$. If $\rho_1 = \rho_2$, (19) becomes $L = 1/\sqrt{4\pi(1/1-\rho_1^2)}\exp\{-a^2/4(1/1-\rho_1^2)\}$, which is equal to (15) for $\rho_1 = \rho_2$. We may conclude that the test (14)

[1]In fact, we suppose also that the observations $x_1(t)$ and $x_2(t)$ are not independent, i.e., the mixture transformation $\mathcal{F}$ is not trivial.

[2]Note that the separation procedure using (6) stays, however, a nonlinear procedure.

excludes the nonlinear Darmois transformation as a separating function unless the two correlated sources have the same spectral densities. This result is analogous to those obtained in linear ICA for the Gaussian correlated sources [3].

III. Discussion and Perspectives

The results of preceding sections show that the two classical examples presented in the literature for illustrating that nonlinear mixtures are not separable up to a trivial indeterminacy can be rejected using the first-lag temporal dependence tests for the mixtures of temporally correlated sources. We must emphasize that this does not give any proof for the separability of these mixtures. There can be other forms of nontrivial mappings giving the independent factors. The question hence remains open until either a counterexample is presented to prove the nonseparability or a real proof of separability is proposed. A possible attempt to achieve such a proof is to consider that the two random variables $y_1(t) = g_1(s_1(t), s_2(t))$ and $y_2(t) = g_2(s_1(t), s_2(t))$, obtained by a nontrivial mapping of the independent and temporally correlated sources $s_1(t)$ and $s_2(t)$, are also independent, i.e.,

$$p(g_1(s_1(t), s_2(t)), g_2(s_1(t), s_2(t))) = p(g_1(s_1(t), s_2(t)))p(g_2(s_1(t), s_2(t)))$$

and to show that in the general case

$$p(g_1(s_1(t-1), s_2(t-1)), g_2(s_1(t), s_2(t))) \neq p(g_1(s_1(t-1), s_2(t-1)))p(g_2(s_1(t), s_2(t)))$$

at least for the known pdf and known temporal correlation structure of the sources. Nevertheless, the result encourages the use of criteria involving temporal structure of the sources in nonlinear mixtures. For improving the separation algorithms, a simplification using second-order statistics, could be sufficient, although this hypothesis requires further investigation. An attempt to use the temporal correlation for separating the nonlinear mixtures can be found in [4].

References

[1] A. Taleb and C. Jutten, “Source separation in post-nonlinear mixtures,” *IEEE Trans. Signal Processing*, vol. 47, pp. 2807–2820, Oct. 1999.

[2] A. Hyvärinen and P. Pajunen, “Nonlinear independent component analysis: Existence and uniqueness results,” *Neural Networks*, vol. 12, no. 3, pp. 429–439, 1999.

[3] L. Tong, R. W. Liu, V. C. Soon, and Y. F. Huang, “Indeterminacy and identifiability of blind identification,” *IEEE Trans. Circuits Syst.*, vol. 38, pp. 499–509, Apr. 1991.

[4] S. Harmeling, A. Ziehe, M. Kawanabe, and K.-R. Müller, “Kernel feature space and nonlinear blind source separation,” in *Advances in Neural Information Processing Systems*, T. G. Dietterich, S. Becker, and Z. Ghahramani, Eds. Cambridge, MA: MIT Press, 2002, vol. 14.