

# ON THE INDETERMINACIES OF CONVOLUTIVE BLIND SIGNAL SEPARATION BASED ON SECOND ORDER STATISTICS

*Daniël W.E. Schobben and Piet C.W. Sommen*

Eindhoven University of Technology,  
P.O. Box 513, 5600 MB Eindhoven, The Netherlands  
Email: D.W.E.Schobben@ele.tue.nl  
URL: <http://www.esp.ele.tue.nl/~daniels/>

## ABSTRACT

Recently, several blind signal separation algorithms have been developed which are based on second order statistics. Little has been published however on whether second order statistics are sufficient to obtain a unique solution. Especially for applications that involve convolutive mixing and unmixing of signals that are correlated in time, there is a lack of knowledge on why and in what cases second order statistics suffice. This paper investigates the indeterminacies that are introduced when second order statistics are used and presents a theorem for the unmixing system to be uniquely found using second order statistics.

## 1. INTRODUCTION

Blind Signal Separation deals with the problem of recovering independent sources from observed linear mixtures of them. These mixtures can be convolutive or instantaneous depending on the application. Second Order Statistics (SOS) approaches have the advantage that no a priori knowledge is required about the sources except that they must be independent and therefore uncorrelated. Also, SOS can be estimated more reliable than Higher Order Statistics (HOS).

BSS algorithms that use HOS include, among others, minimum mutual information and maximum likelihood approaches [1, 2, 3, 4, 5]. These algorithms contain non-linear elements which implicitly introduce HOS. These non-linear elements can be tuned to achieve a good performance for the particular source signals that are considered.

Recently, several algorithms have been introduced based on SOS [6, 7, 8]. Theoretically indeterminacies are introduced when SOS are used to solve the BSS problem. Therefore, some authors claim that SOS are insufficient [2, 9, 10]. This paper briefly gives an overview of the SOS indeterminacies known in literature. It proceeds with deriving a theorem for these indeterminacies in the more general case of convolutive mixing and unmixing of sources that are corre-

lated in time, i.e. sources that have non-zero autocorrelations over multiple time-lags. Also, remarks are made on implementation details and properties of real world signals which helps SOS based blind signal separation.

The notation is in accordance to Figure 1 which depicts the mixing/unmixing system. The independent sources  $s_1 \dots s_J$  are mixed by the mixing system  $H$  to obtain the sensor signals  $x_1 \dots x_J$ . Throughout, both the number of sources and the number of sensors are equal to  $J$ . Time indexes are not mentioned explicitly in all formulas. The expectation operator  $E\{\cdot\}$  gives the ensemble average of a variable. Superscripts denote the vector or matrix dimensions. A length  $L$  vector containing only zeros is denoted  $\underline{0}^L$ .

The unmixing system tries to recover the source signals blindly. I.e. the source signals and the mixing system are unknown so that BSS must be performed using only the sensor signals. In general, signal separation is achieved when

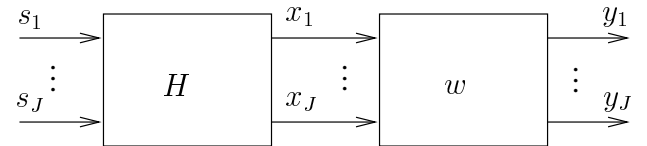


Figure 1: Cascaded mixing/unmixing system

the outputs of the unmixing system  $y_1 \dots y_J$  are independent. In this paper a SOS criterion is considered

$$\forall i \neq j : \rho_{y_i y_j}[\tau] = E\{y_i[n]y_j[n - \tau]\} = 0 \quad (1)$$

with  $n$  the time index and  $\tau = \tau_1 \dots \tau_P$  the correlation lags that are considered. The next additional constraint is used to prevent the BSS algorithm from producing zero outputs

$$\forall i : \rho_{y_i}[0] \neq 0 \quad (2)$$

The question that needs to be answered is therefore the following. When a SOS based BSS algorithm finds an unmixing system such that its outputs are uncorrelated for the set

of time-lags that are considered, under what conditions does this correspond to successful signal separation?

## 2. INDETERMINACIES

The separating criterion that is mentioned in the previous section does not incorporate any specific knowledge on the individual sources signals. The objective is merely to obtain separated signals. Consequently, instantaneous BSS algorithms can recover the sources up to a permutation and a scaling, i.e.

$$\underline{y}^J[n] = P^J D^J \underline{x}^J[n],$$

with  $P^J$  a permutation matrix which is obtained from the identity matrix by interchanging columns. Matrix  $D^J$  is a diagonal matrix which accounts for the scaling factors that cannot be determined. In practical applications, the scaling indeterminacy can be fixed by normalizing the recovered signals so that they have equal powers.

For convolutive BSS algorithms, the permutation indeterminacy is the same. The scaling indeterminacy however becomes a filtering indeterminacy. This means that the BSS algorithm has achieved its goal when filtered versions of the source signals are recovered.

Besides these basic indeterminacies, there are some specific indeterminacies that are related to the source signals, the mixing system and the information that is used. For instantaneous mixing and unmixing, these can be summarized as [11, 12]

- Mutually *uncorrelated* sources  $\underline{s}^J = (s_1 \dots s_J)^T$  can be recovered from an instantaneous mixture  $\underline{x}^J = H^J \underline{s}^J$  using second order statistics when the mixing matrix  $H^J$  is full rank and there is a  $\tau > 0$  such that for all  $i \neq j$

$$\frac{E\{s_i[n]s_i[n-\tau]\}}{E\{s_i^2\}} \neq \frac{E\{s_j[n]s_j[n-\tau]\}}{E\{s_j^2\}}$$

- Mutually *independent* sources can be recovered from an instantaneous mixture using higher order statistics when the mixing matrix is full rank and for all  $i \neq j$

$$\frac{E\{s_i^4\}}{E\{(s_i^2)^2\}} \neq \frac{E\{s_j^4\}}{E\{(s_j^2)^2\}}$$

Note that independent sources are uncorrelated. The opposite statement is not always true however. When sources that are not correlated in time are mixed instantaneously for example, SOS are insufficient to recover them from the mixture. When these signals have non-zero fourth order moments, HOS can be used. Sources with Gaussian PDFs have zero fourth order moments so that they cannot be recovered using HOS either in this case.

In [13] the uniqueness is studied of convolutive BSS of mixed stationary independent signals based on SOS. It was concluded that additional constraints are needed and that it is still not guaranteed that an adaptive algorithm converges to the desired solution. The cross power spectrum of the outputs can be written as

$$\Phi_{y_i y_j}(z) = \sum_{k=1}^J A_{ik}(z) A_{jk}^*(z) \Phi_{s_k}(z),$$

with  $A_{ik}(z)$  the  $z$ -transformed transfer function from source  $s_k$  to output  $y_i$  and  $\Phi_{s_k}(z)$  the power spectrum of source  $s_k$ . From this expression it follows that the true and the ‘permuted’ solutions are not the only ones. Setting the cross power spectrum to zero for the case of two sources and two sensors for example gives

$$A_{11}(z) A_{21}^*(z) \Phi_{s_1}(z) = -A_{12}(z) A_{22}^*(z) \Phi_{s_2}(z). \quad (3)$$

True signal separation is achieved when this equation holds for all spectra  $\Phi_{s_1}(z)$  and  $\Phi_{s_2}(z)$ . This is the case when crosstalk is eliminated, e.g.  $A_{21}(z) = A_{12}(z) = 0$ . This equation can also be fulfilled however for particular choices of the unmixing filters that depend on the spectra of the source signals without achieving signal separation. This is called the ‘phantom’ solution in [13]. It is mentioned that unmixing filters of infinite length are required for the phantom solution to exist. In the limiting case of very long unmixing filters it can be approximated however.

## 3. CONVOLUTIVE BSS INDETERMINACIES

In this section the uniqueness of the unmixing system is considered taking into account that the unmixing filters are of finite length. The transfer function of the cascaded mixing and unmixing system is also assumed to be of some finite length  $L$  and is expressed as

$$y_i[n] = \sum_{l=1}^J \sum_{m=0}^{L-1} a_{il}[m] s_l[n-m],$$

with  $a_{il}[m]$  the impulse response of the cascaded mixing and unmixing system from source  $s_l[n]$  to output  $y_i[n]$ . The cross-correlations  $\rho_{y_i y_j}[\tau] = E\{y_i[n] y_j[n-\tau]\}$  can be grouped into one equation for the correlation lags considered:

$$\begin{pmatrix} \rho_{y_i y_j}[\tau_1] \\ \vdots \\ \rho_{y_i y_j}[\tau_P] \end{pmatrix} = \Xi^{P, J(2L-1)} \begin{pmatrix} \widehat{a_{i1} a_{j1}}^{2L-1} \\ \vdots \\ \widehat{a_{iJ} a_{jJ}}^{2L-1} \end{pmatrix} \quad (4)$$

with the source correlation matrix  $\Xi^{P,J(2L-1)}$  given by

$$\begin{pmatrix} \rho_{s_1}[\tau_1-L] \dots \rho_{s_1}[\tau_1+L] \\ \vdots \\ \rho_{s_1}[\tau_P-L] \dots \rho_{s_1}[\tau_P+L] \end{pmatrix} \dots \begin{pmatrix} \rho_{s_J}[\tau_1-L] \dots \rho_{s_J}[\tau_1+L] \\ \vdots \\ \rho_{s_J}[\tau_P-L] \dots \rho_{s_J}[\tau_P+L] \end{pmatrix}$$

and with

$$\widehat{a_{il}a_{jl}}^{2L-1} = \begin{pmatrix} a_{il}[L-1]a_{jl}[0] \\ \sum_{n=L-2}^{L-1} a_{il}[n]a_{jl}[n-L+2] \\ \vdots \\ \sum_{n=1}^{L-1} a_{il}[n]a_{jl}[n-1] \\ \sum_{n=0}^{L-1} a_{il}[n]a_{jl}[n] \\ \sum_{n=1}^{L-2} a_{il}[n]a_{jl}[n+1] \\ \vdots \\ a_{il}[0]a_{jl}[L-1] \end{pmatrix}. \quad (5)$$

The  $\widehat{a_{il}a_{jl}}^{2L-1}$  is a correlation of  $\underline{a}_{il}^L$  and  $\underline{a}_{jl}^L$ , i.e. a convolution of  $\mathbf{J}^L \underline{a}_{il}^L$  and  $\underline{a}_{jl}^L$ , with  $\mathbf{J}^L$  the  $L \times L$  mirror matrix.

This interpretation reveals that  $\widehat{a_{il}a_{jl}}^{2L-1}$  equals zero iff  $\underline{a}_{il}^L$  or  $\underline{a}_{jl}^L$  equals zero. Uncorrelated outputs are sufficient to uniquely define the  $\widehat{a_{il}a_{jl}}^{2L-1}$  when the source correlation matrix  $\Xi^{P,J(2L-1)}$  is of rank  $J(2L-1)$ . In that case its nullspace is empty so that the only solution is  $\forall l, i \neq j : \widehat{a_{il}a_{jl}}^{2L-1} = \mathbf{0}^{2L-1}$ .

**Theorem 1** *An unmixing system that produces uncorrelated non-zero outputs at time lags  $\tau = \tau_1 \dots \tau_P$  achieves signal separation if the source correlation matrix  $\Xi^{P,J(2L-1)}$  is of rank  $J(2L-1)$  for this set of time lags.*

*Proof:*

The proof follows directly from the next two observations:

**Observation 1: uncorrelated outputs**

The separation system retrieves mutually uncorrelated outputs if the right hand side of (4) equals zero for  $i \neq j$ . This is the case iff  $\forall l : \underline{a}_{il}^L = \mathbf{0}^L$  or  $\underline{a}_{jl}^L = \mathbf{0}^L$ . When source  $l$  contributes to a certain output  $i$ , e.g.  $\underline{a}_{il}^L \neq \mathbf{0}^L$ , it cannot contribute to any other output as  $\underline{a}_{jl}^L = \mathbf{0}^L, j \neq i$ .

**Observation 2: non-zero outputs**

To prevent the system from setting the filters in the unmixing system to zero it is required that the autocorrelations of the outputs are non-zero,  $\forall i : \exists \tau : \rho_{y_i}[\tau] \neq 0$ . This requirement is met when  $\forall i : \exists l : \widehat{a_{il}a_{il}}^{2L-1} \neq \mathbf{0}^{2L-1}$  or similarly  $\forall i : \exists l : \underline{a}_{il}^L \neq \mathbf{0}^L$ . This means that there is at least one non-zero transfer function  $\underline{a}_{il}^L$  to every output  $i$ .

According to observation 2 there is a source that contributes to a particular output. According to observation 1 this source cannot contribute to any other outputs. Each source therefore uniquely corresponds to one output  $\square$ .

#### 4. RANK OF THE SOURCE CORRELATION MATRIX

According to Theorem 1 the source correlation matrix  $\Xi^{P,J(2L-1)}$  must be of rank  $J(2L-1)$  in order for BSS to succeed using SOS. The number of time lags that are considered  $P$  must therefore at least be equal to  $J(2L-1)$ . Whether the source correlation matrix  $\Xi^{P,J(2L-1)}$  is full rank then still depends on

- **Source similarity:** when the autocorrelations of two sources are similar, columns in the source correlation matrix can be eliminated and BSS is not possible in accordance with Theorem 1.
- **Correlation length:** the autocorrelation of at least one source needs to be of sufficient length. When the sources have autocorrelations equal to zero  $\rho_{s_i}[\tau] = 0$  for  $|\tau| > K$ , then the rank of the source correlation matrix is at most  $2L-1+2K$ .

The well known fact that white noise sources cannot be recovered using SOS is found from investigating the rank of the source correlation matrix. Another result is that sources with short autocorrelation sequences (e.g. moving average processes) can be recovered if one of the sources does have a long autocorrelation sequence. So if the correlation sequences are distinct, one source with a long correlation sequence can render the source correlation matrix full rank. The length of this correlation sequence must be such that  $2L-1+2K \geq J(2L-1)$  which gives

$$K \geq \frac{(J-1)}{2}(2L-1). \quad (6)$$

When there are many sources  $J$ , there must be a source with a long correlation sequence in order for SOS-based BSS to work. In audio applications this criterion is not met in general as the correlation length  $K$  for signals such as speech is shorter than the length of the cascaded mixing and unmixing system  $L$ . Despite of this, successful separation is achieved for several algorithms. This is due to the non-stationarity of the data and implementation details of such algorithms as discussed in the next section.

#### 5. REAL WORLD DATA CONSIDERATIONS

Source signals that have short autocorrelation sequences can be whitened when long unmixing filters are used. BSS is therefore not possible when using SOS, which is in correspondence with (6). However, typically constraints are incorporated into BSS algorithms to prevent them from whitening the data. For audio signals this is usually done so that the separated signals sound more natural. It has as a side effect however that the algorithms are forced to

separate as they are prevented from minimizing their cost-function by whitening the data.

Another important aspect is that most analysis - including the one in the previous section - assume stationarity of the source signals. Real world applications however typically involve the separation of signals that are not stationary. Consider for example speech signals which are recorded in a real room. Speech signals are stationary only over approximately 10-30 milliseconds. Thus, speech can be modeled by time varying innovation filters excited by white noise. These innovation filters change therefore every 10-30 ms. Typically, the convergence time of the SOS algorithm is slower than 10-30 ms so that the BSS algorithm fails to track the inverse of the innovation filters and only true separation results in uncorrelated recovered signals at all times. This is in correspondence with (3) in which for each choice of stationary source signals there exist a set of unmixing filters that renders the cross-correlations among the outputs equal to zero. This set is different for other stationary source signals but it always contains the desired solution that gives true signal separation. This true solution can therefore be found from the intersection of the solution sets. This is obtained automatically in an adaptive system when the stationarity of the source signals is over a shorter time than the convergence time of the adaptive system.

## 6. CONCLUSIONS

In this paper conditions are derived for second order statistics to be sufficient in convolutive blind signal separation which are based on mathematical proof and properties of real world signals. A theorem is given for second order statistics to be sufficient for the blind separation of signals that have long autocorrelation sequences compared to the length of the mixing and unmixing filters. For signals that do not have autocorrelations of such lengths it is argued that blind signals separation is still possible given that these signals are non-stationary.

## 7. ACKNOWLEDGMENTS

The authors would like to thank Jean-François Cardoso and Yves Grenier for the valuable discussions.

## 8. REFERENCES

- [1] T-W. Lee, A.J. Bell, and R. Lambert. Blind separation of delayed and convolved sources. *Advances in Neural Inf. Proc. Systems* 9, pages 758–764, 1997.
- [2] A.J. Bell and T.J. Sejnowski. An information maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159, 1995. MIT Press, Cambridge MA.
- [3] K. Torkkola. Blind separation of delayed sources based on information maximization. In *IEEE Workshop on Neural Networks for Signal Proc.*, Sept. 1996.
- [4] P. Smaragdis. Blind separation of convolved mixtures in the frequency domain. In *Int. Workshop on Independence & Artificial Neural Networks*, Febr. 1998.
- [5] R.H. Lambert. *Multichannel Blind Deconvolution: FIR Matrix Algebra and Separation of Multipath Mixtures*. University of Southern California, May 1996. Ph.D. Thesis.
- [6] D.W.E. Schobben and P.C.W. Sommen. A new convolutive blind signal separation algorithm based on second order statistics. In *Proceedings Int. Conf. on Signal and Image Proc.*, pages 564–569, Oct. 1998.
- [7] S. v. Gerven. *Adaptive Noise Cancellation and Signal Separation with Applications to Speech Enhancement*. Katholieke Universiteit Leuven, March 1996. Ph.D. Thesis.
- [8] D.C.B. Chan. *Blind Signal Separation*. Cambridge: Thesis University of Cambridge, 1997. Ph. D. Thesis.
- [9] C. Jutten and J. Herault. Independent component analysis (INCA) versus principal component analysis. In J.L. Lacoume et al., editor, *Signal Processing IV: Theories and Applications*, pages 643–646. Elsevier, 1988.
- [10] E. Weinstein, M. Feder, and A.V. Oppenheim. Multi-channel signal separation by decorrelation. *IEEE Transactions on Speech and Audio Processing*, 1(4):405–413, 1993.
- [11] L. Tong, R. Liu, V. Soon, and Y. Huang. Indeterminacy and identifiability of blind identification. *IEEE Transactions on Circuits and Systems*, 38(5):499–509, May 1991.
- [12] L. Molgedey and H.G. Schuster. Separation of independent signals using time-delayed correlations. *Physical Review Letters*, 72(23):3634–3637, June 6 1994.
- [13] S. Van Gerven and D. Van Compernelle. Signal separation by symmetric adaptive decorrelation: Stability, convergence, and uniqueness. *IEEE Transactions on Signal Processing*, 43(7):1602–1612, July 1995.