

*Faculty of Informatics*

*Faculty of Informatics - Papers*

---

*University of Wollongong*

*Year 2004*

---

De-cumulant based approaches for  
convolutive blind source separation

T. Mei\*                      J. Xi†  
J. F. Chicharo‡              F. Yin\*\*

\*University of Wollongong, tiemin@uow.edu.au

†University of Wollongong, jiangtao@uow.edu.au

‡University of Wollongong, chicharo@uow.edu.au

\*\*Dalian University of Technology, China

This paper originally appeared as: Mei, T, Xi, J, Chicharo, JF and Yin, F, De-cumulant based approaches for convolutive blind source separation, Proceedings of 2004 International Intelligent Multimedia, Video and Speech Processing, 20-22 October 2004, 471-474. Copyright IEEE 2004.

This paper is posted at Research Online.

<http://ro.uow.edu.au/infopapers/191>

## DE-CUMULANT BASED APPROACHES FOR CONVOLUTIVE BLIND SOURCE SEPARATION

Tiemin Mei<sup>1,2</sup>, Jiangtao Xi<sup>1</sup>, Joe Chicharo<sup>1</sup> and Fuliang Yin<sup>2</sup>

<sup>1</sup>School of Electrical, Computer and Telecommunications Engineering,  
The University of Wollongong, NSW 2522, Australia, email: Jiangtao@uow.edu.au

<sup>2</sup>School of Electronic and Information Engineering  
Dalian University of Technology, Dalian, Liaoning, 116023 P. R. China

### ABSTRACT

*Abstract:* This paper studies the blind separation of signal sources (BSS) based on the approaches of de-cumulant. It considers the cases where independent signal sources are mixed through convolutive mixing system with unity autochannel frequency responses and causal cross-channel FIR filters. Firstly, it tries to show that de-cumulant is sufficient for separation. Secondly, novel algorithms are developed based on zero-forcing of cross-cumulant pairs. These algorithms are developed in time-domain and so there is not the frequency permutation ambiguity problem usually suffered by most of the frequency-domain algorithms. Simulation results are presented to support the validity of the proposed algorithms.

### 1. INTRODUCTION

Many High Order Statistics (HOS) based approaches [1] have been considered as an effective solution for the challenging problem of blind source separation [2]. However, for the cases of convolutive environment, the effectiveness of HOS based techniques has not been studied thoroughly yet. In this paper we consider the application of de-cumulant to the convolutive mixing model has been studied in [3][4], with the purpose to show that de-cumulants are sufficient for BSS.

### 2. THE MIXING AND SEPARATING MODELS

In this paper we consider the two-by-two cases. That is, there are two signal sources, two observation signals and two separated signals as well, denoted by  $S(t)$ ,  $X(t)$  and  $Y(t)$  respectively, where  $S(t) = [s_1(t), s_2(t)]^T$ ,  $X(t) = [x_1(t), x_2(t)]^T$  and  $Y(t) = [y_1(t), y_2(t)]^T$ . Note that  $s_1(t)$  and  $s_2(t)$  are real-valued, of zero-mean and statistically mutually independent.

The cross-cumulant of real valued stochastic processes  $x(t_1)$  and  $y(t_2)$  with order of  $(r_1 + r_2)$  is denoted as:

$$c_{xy}^{r_1 r_2}(t_1, t_2) \equiv \text{cum}(x^{r_1}(t_1), y^{r_2}(t_2)) \\ \equiv \text{cum}(\underbrace{x(t_1), \dots, x(t_1)}_{r_1}, \underbrace{y(t_2), \dots, y(t_2)}_{r_2}) \quad (1)$$

Also the  $r$ th-order auto-cumulant of  $x(t)$  is denoted by  $c_x^r(t_1, \dots, t_r) = \text{cum}(x(t_1), \dots, x(t_r))$ . As usual, we assume that signal sources are independent, that is  $\text{cum}(s_1(t_{11}), \dots, s_1(t_{1p}), s_2(t_{21}), \dots, s_2(t_{2q})) = 0$  (2) for  $\forall t_{1i}, t_{2j} (i = 1, 2, \dots, p; j = 1, 2, \dots, q)$  and  $\forall p, q \geq 1$ .

### 2.1 The Convolutive Mixing System Model

The mixing system considered in this paper is described as follows:

$$x_1(t) = s_1(t) + \sum_{i=0}^N a_{12}(i)s_2(t-i) \quad (3)$$

$$x_2(t) = s_2(t) + \sum_{i=0}^N a_{21}(i)s_1(t-i) \quad (4)$$

where  $a_{12}(i)$  and  $a_{21}(i) (i=0, 1, 2, \dots, N)$  are the impulse responses of the two coupling filters. We assume that the two FIR filters are of the same length  $N$ . Note that the above model is equivalent to the four mixing filter model [5], but much easier to cope with.

### 2.2 The Separation System Model

There are two types of separation systems that can be employed: the forward separation system and the backward separation system. The two systems are equivalent in the sense that both can achieve source separation. However, the backward system is better in that it directly yields the source signals themselves rather than the filtered version of the sources as those by the forward system.

*The forward separation neural network*

The forward separation neural network is described by:

$$y_1(t) = x_1(t) - \sum_{i=0}^N h_{12}(i)x_2(t-i) \quad (5)$$

$$y_2(t) = x_2(t) - \sum_{i=0}^N h_{21}(i)x_1(t-i) \quad (6)$$

Combining Equations (3), (4), (5) and (6), we have:

$$y_1(t) = v_{11}(t) + v_{12}(t) \quad (7)$$

$$y_2(t) = v_{21}(t) + v_{22}(t) \quad (8)$$

where

$$v_{11}(t) = \sum_{k=0}^{2N} w_{11}(k) s_1(t-k), \quad w_{11}(k) = \delta(k) - \sum_{i=0}^N a_{12}(i) h_{21}(k-i) \quad (9)$$

$$k=0, 1, 2, \dots, 2N.$$

$$v_{12}(t) = \sum_{k=0}^N w_{12}(k) s_2(t-k), \quad w_{12}(k) = a_{12}(k) - h_{12}(k) \quad (10)$$

$$k=0, 1, 2, \dots, N.$$

$$v_{21}(t) = \sum_{k=0}^N w_{21}(k) s_1(t-k), \quad w_{21}(k) = a_{21}(k) - h_{21}(k) \quad (11)$$

$$k=0, 1, 2, \dots, N$$

$$v_{22}(t) = \sum_{k=0}^{2N} w_{22}(k) s_2(t-k), \quad w_{22}(k) = \delta(k) - \sum_{i=0}^N a_{12}(i) h_{21}(k-i) \quad (12)$$

$$k=0, 1, 2, \dots, 2N$$

The backward separation neural network

The backward network is described as follows.

$$y_1(t) = x_1(t) - \sum_{i=0}^N c(i) y_2(t-i) \quad (13)$$

$$y_2(t) = x_2(t) - \sum_{i=0}^N d(i) y_1(t-i) \quad (14)$$

where  $c(i)$  and  $d(i)$  are the impulse response of the two feedback filters.

It is easy to find that the end-to-end relationship between the sources and the separated output for backward separation system is that for forward separation system cascaded by a filter  $H(z) = [1 - H_{12}(z)H_{21}(z)]^{-1}$ . Thus, the separation conditions for the forward separation system are also valid for the backward separation system.

### 3. THE SEPARATION OF CONVOLUTIVE MIXTURES BASED ON DECUMULANTS

In this section, we try to show that de-cumulants are sufficient for the separation of convolutely mixed signals. The  $(r_1 + r_2)$  order cross-cumulant of  $y_1(t)$  and  $y_2(t - \tau)$  is as follows

$$c_{y_1 y_2}^{r_1 r_2}(t, t - \tau) = c_{v_{11} v_{21}}^{r_1 r_2}(t, t - \tau) + c_{v_{12} v_{22}}^{r_1 r_2}(t, t - \tau) \quad (15)$$

where

$$c_{v_{11} v_{21}}^{r_1 r_2}(t, t - \tau) = \sum_{k_1=0}^{2N} \dots \sum_{k_{r_1-1}=0}^{2N} \dots \sum_{l_2=0}^N \left( \prod_{i=1}^{r_1} w_{11}(k_i) \prod_{j=1}^{r_2} w_{12}(l_j) \right) \times \quad (16)$$

$$\times c_{s_1}^{r_1+r_2}(t-k_1, \dots, t-k_{r_1}, t-l_1-\tau, \dots, t-l_{r_2}-\tau)$$

$$c_{v_{12} v_{22}}^{r_1 r_2}(t, t - \tau) = \sum_{k_1=0}^N \dots \sum_{k_{r_1-1}=0}^N \dots \sum_{l_2=0}^{2N} \left( \prod_{i=1}^{r_1} w_{21}(k_i) \prod_{j=1}^{r_2} w_{22}(l_j) \right) \times \quad (17)$$

$$\times c_{s_2}^{r_1+r_2}(t-k_1, \dots, t-k_{r_1}, t-l_1-\tau, \dots, t-l_{r_2}-\tau)$$

The derivation of equation (15) has utilized the following assumption, that is

$$\text{cum}(s_1(t-\tau_{11}), \dots, s_1(t-\tau_{1p}), s_2(t-\tau_{21}), \dots, s_2(t-\tau_{2q})) = 0 \quad (18)$$

$$p, q > 0; \quad p+q = r_1 + r_2$$

Clearly the above condition less restrict when compared to independence.

Note that  $c_{v_{11} v_{21}}^{r_1 r_2}(t, t - \tau)$  and  $c_{v_{12} v_{22}}^{r_1 r_2}(t, t - \tau)$  have very different changing trends with respect to  $\tau$ . Hence even the two independent sources have similar or the same statistics and the mixing system is symmetric, it is still impossible that  $c_{v_{11} v_{21}}^{r_1 r_2}(t, t - \tau) = -c_{v_{12} v_{22}}^{r_1 r_2}(t, t - \tau)$  for all time delay  $\tau$ . Thus, let  $c_{y_1 y_2}^{r_1 r_2}(t, t - \tau) = 0$  ( $\forall t, \tau$  and  $r_1, r_2 \geq 1$ ), we must have

$$c_{v_{11} v_{21}}^{r_1 r_2}(t, t - \tau) = 0, \quad \forall t, \tau \quad (19)$$

$$c_{v_{12} v_{22}}^{r_1 r_2}(t, t - \tau) = 0, \quad \forall t, \tau \quad (20)$$

Considering that  $v_{11}(t)$  and  $v_{21}(t)$  are originated from the same source signal  $s_1(t)$ , the condition of (19) implies that  $v_{11}(t) = 0$  or  $v_{21}(t) = 0$ . However, because  $a_{12}(t)$  and  $h_{21}(t)$  are causal FIR filters, it is not possible that  $a_{12}(t)h_{21}(t) = \delta(t)$ , which imply that  $v_{11}(t) \neq 0$ . Thus  $v_{21}(t) = 0$  is the unique solution that makes (19) holds. Similarly we conclude that  $v_{12}(t) = 0$  is the unique solution that makes (20) holds.

From the above discussion we can conclude that

$$c_{y_1 y_2}^{r_1 r_2}(t, t - \tau) = 0 \quad (\forall t, \tau \text{ and } r_1, r_2 \geq 1) \quad (21)$$

is a sufficient condition for the separation of convolutely mixed sources. Equation (21) is the decumulant criteria for convolutive mixture separation.

Equation (18) also gives the separable conditions for mixed signal sources. The smaller  $r_1$  and  $r_2$  are, the lower order the decumulant approaches can be used. A special case is that  $r_1, r_2 = 1$ , the two signal sources are uncorrelated, in which case decorrelation is sufficient for

achieving source separation.

#### 4. ALGORITHMS DEVELOPMENT BASED ON DECUMULANTS

In this section, we present a family of algorithms for the separation of convolutely mixed sources under the decumulant criteria.

Consider the backward separation network which contains strict causal channels, that is,  $c(0) = 0$  and  $d(0) = 0$  [6][7]. Let  $r_1 = 1$ ,  $r_2 = r$  and  $\tau = 1, 2, \dots, N$  in (21), we obtain the following equation groups

$$c_{y_1 y_2}^{1r}(t, t - \tau) = 0, \tau = 1, 2, \dots, N \quad (22)$$

Substituting (5) and (6) into (22) and expressing the result in matrix form, we obtain

$$c_{x_1 y_2}^{1r} - C_{y_2 y_2}^{1r} W_1 = c_{y_1 y_2}^{1r} = 0 \quad (23)$$

Substituting (5) and (6) into (21) and also letting  $r_1 = r$ ,  $r_2 = 1$  and  $\tau = 1, 2, \dots, N$  gives

$$c_{x_2 y_1}^{1r} - C_{y_1 y_1}^{1r} W_2 = c_{y_2 y_1}^{1r} = 0 \quad (24)$$

where

$$W_1 = [c(1), \dots, c(N)]^T \quad (25)$$

$$W_2 = [d(1), \dots, d(N)]^T \quad (26)$$

$$Y_i = [y_i(t-1), \dots, y_i(t-N)]^T \quad i = 1, 2 \quad (27)$$

$$c_{x_k y_l}^{1r} = [\text{cum}(x_k(t), y_l^r(t-1)), \dots, \text{cum}(x_k(t), y_l^r(t-N))]^T \quad (28)$$

$$k, l \in [1, 2], k \neq l$$

$$c_{y_k y_l}^{1r} = [\text{cum}(y_k(t), y_l^r(t-1)), \dots, \text{cum}(y_k(t), y_l^r(t-N))]^T \quad (29)$$

$$k, l \in [1, 2], k \neq l$$

$$C_{y_k y_k}^{1r} = \text{cum}[Y_k^r Y_k^{rT}] \quad (30)$$

$$= [\text{cum}(y_k^r(t-i), y_k^r(t-j))]_{i,j=1,2,\dots,N} \quad k \in [1, 2]$$

where in (30),  $C_{y_k y_k}^{1r}$  ( $k=1, 2$ ) are two square matrices,  $\text{cum}(y_k^r(t-i), y_k^r(t-j))$  is their  $i$ th row  $j$ th column entry.

The equations (23) and (24) are two groups of nonlinear equations. Applying Robbins-Monro first-order stochastic approximation methods [8], (23) and (24) give rise to the following adaptive algorithm

$$W_1(t+1) = W_1(t) + \mu(t)c_{y_1 y_2}^{1r} \quad (31)$$

$$W_2(t+1) = W_2(t) + \mu(t)c_{y_2 y_1}^{1r} \quad (32)$$

We now discuss two special cases of the algorithm family (31) and (32).

*Special case 1:* let  $r=1$  in (23) and (24) and substitute instantaneous values for their ensemble values, we obtain a decorrelation-based algorithm:

$$W_1(t+1) = W_1(t) + \mu_1 y_1(t) Y_2(t) \quad (33)$$

$$W_2(t+1) = W_2(t) + \mu_2 y_2(t) Y_1(t) \quad (34)$$

*Special case 2:* let  $r=3$  in (23) and (24), and replace the  $N$  by 1 vectors  $c_{y_1 y_2}^{13}$  and  $c_{y_2 y_1}^{13}$  with their instantaneous value, we obtain a 4<sup>th</sup>-order de-cumulant algorithm:

$$W_1(t+1) = W_1(t) + \mu_1 y_1(t) \text{diag}\{Y_2^2(t) - 3E[Y_2^2(t)]\} Y_2(t) \quad (35)$$

$$W_2(t+1) = W_2(t) + \mu_2 y_2(t) \text{diag}\{Y_1^2(t) - 3E[Y_1^2(t)]\} Y_1(t) \quad (36)$$

where  $Y_i^r(t) \equiv [y_i^r(t-1), \dots, y_i^r(t-N)]^T$  ( $r=2$ ) ( $i=1, 2$ ). Operator  $\text{diag}\{\cdot\}$  rearranges a vector into a diagonal matrix whose diagonal entries are the corresponding elements of the vector;  $E[\cdot]$  represents the ensemble average operation.

Let  $\bar{\mu}_i(t) = \mu_i \text{diag}\{Y_i^2(t) - 3E[Y_i^2(t)]\}$ , the algorithms (35) and (36) can be considered as decorrelation algorithm with time-changing learning rates. On the other hand, if we let  $F(Y_i(t)) = \text{diag}\{Y_i^2(t) - 3E[Y_i^2(t)]\} Y_i(t)$ , (35) and (36) can be rewritten as

$$W_1(t+1) = W_1(t) + \mu_1 y_1(t) F(Y_2(t)) \quad (37)$$

$$W_2(t+1) = W_2(t) + \mu_2 y_2(t) F(Y_1(t)) \quad (38)$$

Equation (37) and (38) can be interpreted as an algorithm based on nonlinear neural network. So we may also select other nonlinear functions such as  $F(Y_i(t)) = \alpha Y_i(t) + \beta Y_i^3(t)$  in (37) and (38), as did in [9]. Note that the parameters  $\alpha$  and  $\beta$  should be selected carefully.

We can even replace  $y_i(t)$  ( $i=1, 2$ ) with  $f(y_i(t))$ , where  $f(\cdot)$  is a nonlinear function such as  $f(y) = \alpha y + \tanh(\gamma y)$ , in (33) and (34), resulting in the following algorithm:

$$W_1(t+1) = W_1(t) + \mu_1 f(y_1(t)) Y_2(t) \quad (39)$$

$$W_2(t+1) = W_2(t) + \mu_2 f(y_2(t)) Y_1(t) \quad (40)$$

Combining the second- and fourth-order de-cumulant algorithms (33), (34) and (35), (36), we obtain the

following algorithm.

$$W_1(t+1) = W_1(t) + \mu_{11}y_1(t)Y_2(t) + \mu_{12}y_1(t)\text{diag}\{Y_2^2(t) - 3E[Y_2^2(t)]\}Y_2(t) \quad (41)$$

$$W_2(t+1) = W_2(t) + \mu_{21}y_2(t)Y_1(t) + \mu_{22}y_2(t)\text{diag}\{Y_1^2(t) - 3E[Y_1^2(t)]\}Y_1(t) \quad (42)$$

### 5. NUMERICAL EXPERIMENTS

Computer simulations have been performed for verify the algorithms. For simplicity, we denote (33) and (34) as algorithm 1, (35) and (36) as algorithm 2, (37) and (38) as algorithm 3, (39) and (40) as algorithm 4, (41) and (42) as algorithm 5.

Two speech signals are used in the experiments. The length of the signals are 30000 samples, sampling frequency is 16kHz. The mixing filters  $a_{12}(t)$  and  $a_{21}(t)$  unchanged as follows:

$$a_{12}(t) = [0;0;0;0.2;0.4;0.05;-0.3;0.15;-0.1;0;0.05;0.08;0.09;0;0.03;0]$$

$$a_{21}(t) = [0;0;0;-0.1;0.4;0.1;0.01;-0.4;0.15;0;0.05;0;0.06;0;0.04;0]$$

In algorithms 2 and 5, the  $E[\text{diag}\{Y_i^2(t)\}]$  ( $i=1, 2$ ) is estimated by

$$E[\text{diag}\{Y_i^2(t)\}] = \alpha E[\text{diag}\{Y_i^2(t-1)\}] + (1-\alpha)\text{diag}\{Y_i^2(t)\}$$

and where  $\alpha=0.999999$ .

The parameters for  $F(Y_i(t)) = \alpha Y_i(t) + \beta Y_i^3(t)$  in algorithm 3 are that  $\alpha=0.6$  and  $\beta=3.5$ ; The parameters for  $f(y) = \alpha y + \tanh(\gamma y)$  in algorithm 4 are that  $\alpha=0.5; \gamma=2.5$ ; The learning rates are listed in Table I.

The algorithm 1 through 5 is investigated intensively. We use Signal-to-interference ratio (SIR) to measure separation performance of the algorithms. Table I lists the separation results of algorithms 1 through 5 based on 10 times of operation on the same data. From Table I, we can see that all algorithms give quite good separation results. Algorithm 5 has much better performance than others, but it is more complex as well. Algorithm 3 is almost as good as algorithm 5 is much simpler. Hence we may say that Algorithm 3 is the best option if both performance and simplicity are all considered.

TABLE I PERFORMANCE COMPARISON

Algorithms	$\mu_A$	$\mu_B$	SIR <sub>1</sub> (dB)	SIR <sub>2</sub> (dB)
1	0.003	0.000	22.54	21.85
2	0.000	0.010	23.95	19.23
3	0.003	0.000	24.92	22.54
4	0.003	0.000	23.26	22.11
5	0.003	0.010	24.69	23.01

Note: (1) The SIRs of the mixtures: SIR<sub>1</sub>=8.83 dB; SIR<sub>2</sub>=6.22 dB  
 (2) In algorithms 1,3 and 4:  $\mu_1=\mu_2=\mu_A$ ; in algorithm 2:  $\mu_1=\mu_2=\mu_B$ ; in algorithm 5:  $\mu_{11}=\mu_{12}=\mu_A$  and  $\mu_{21}=\mu_{22}=\mu_B$ ;

### 6. CONCLUSIONS

For the equivalent convolutive mixing system model with unity auto-channel frequency responses and causal cross-channel FIR filters, we prove that the zero-forcing of cross-cumulants of the outputs of separation system is sufficient for the separation of convolutive mixtures of statistically mutually independent sources. Experiment results show that the combination of SOS and HOS is more effective than SOS or HOS only. Those nonlinear function-based algorithms are not only simple in computation but also of better performance than the SOS based algorithm.

### REFERENCES

- [1] J. F. Cardoso, "Blind signal separation: Statistical Principles," Proc. of the IEEE, vol.86, no.10, pp.2009-2025, 1998.
- [2] C. Jutten and J. Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture," Signal Processing, vol.24, pp.1-10, 1991.
- [3] D. Yellin and E. Weinstein, "Multichannel signal separation: Methods and analysis," IEEE Trans. Signal Processing, vol.44, no.1, pp.106-118, 1996.
- [4] D. Yellin and E. Weinstein, "Criteria for multichannel signal separation," IEEE Trans. Signal Processing, vol.42, no.8, pp.2156-2168, 1994.
- [5] M. Kawamoto and K. Matsuoka et al, "A method of blind separation for convolved non-stationary signals," Neurocomputing, vol.22, pp.157-171, 1998.
- [6] S. Van Gerven and D. Van Compernelle, "Signal separation by symmetric adaptive decorrelation: stability, convergence and uniqueness," IEEE Trans. Signal Processing, vol.43, no.7, pp.1602-1612, 1995.
- [7] G. Mirchandani, R. L. Zinser, et al, "A new adaptive noise cancellation scheme in the presence of crosstalk," IEEE Trans. Signal Processing, vol. 39, no. 10, pp. 681-694, 1992.
- [8] P. Comon, "Blind separation of sources, part II: problems statement," Signal Processing, vol. 24, pp.11-20, 1991.
- [9] S. Amari and A. Cichocki, "Adaptive blind signal processing-Neural network approaches," Proc. of the IEEE, vol. 86, no.10, pp.2026-2048, 1998.