# A BOUNDED COMPONENT ANALYSIS APPROACH FOR THE SEPARATION OF CONVOLUTIVE MIXTURES OF DEPENDENT AND INDEPENDENT SOURCES

*Huseyin A. Inan, and Alper T. Erdogan*

Electrical and Electronics Engineering Department,
Koc University,
Istanbul, Turkey

**ABSTRACT**

Bounded Component Analysis is a new framework for Blind Source Separation problem. It allows separation of both dependent and independent sources under the assumption about the magnitude boundedness of sources. This article proposes a novel Bounded Component Analysis optimization setting for the separation of the convolutive mixtures of sources as an extension of a recent geometric framework introduced for the instantaneous mixing problem. It is shown that the global maximizers of this setting are perfect separators. The article also provides the iterative algorithm corresponding to this setting and the numerical examples to illustrate its performance especially for separating convolutive mixtures of sources that are correlated in both space and time dimensions.

***Index Terms***— Convolutive Blind Source Separation, Bounded Component Analysis, Independent Component Analysis, Dependent Component Analysis.

## 1. INTRODUCTION

Blind Source Separation (BSS) is one of the basic problems in signal processing and machine learning with a diverse set of applications [1]. Independent Component Analysis approach, which assumes and exploits the mutual independence among sources, has been the most popular approach in efforts to obtain solution to this problem [1,2].

Following the works exploiting the boundedness of sources within the ICA framework (e.g., [3–6]), Bounded Component Analysis (BCA) has been recently introduced as a new framework which allows separation of both dependent and independent sources for bounded sources [7]. It exploits the knowledge about the boundedness of sources to replace the mutual independence assumption with a more generic domain separability assumption. Therefore, the requirement one the pdf separability in terms of products of marginals is abandoned. Potential benefits of this new approach is mainly twofold:

- Under the standing source boundedness assumption, it provides a more general framework for the separation of both independent and dependent (even correlated) sources.
- Moreover, even when the sources are independent, short data records used for adaptation may not reflect this behavior. Therefore, abandoning of the independence assumption can provide performance improvement for finite data records even for mutually independent case.

In [8], a geometric framework for the construction of BCA algorithms was introduced. The approach introduced in this framework makes use of two geometric objects related to separator output samples, namely principal hyper-ellipsoid and bounding hyper-rectangle, and pose BCA problem as optimization of the relative sizes of these objects. The framework introduced in [8] was limited to memoryless, or instantaneous, mixtures. In this article, we extend this framework to more general case of convolutive mixing, where the sources can be mixed in both source and sample (space and time) dimensions. The proposed convolutive BCA approach allows separation of sources which can be potentially correlated in both source and sample dimensions.

The organization of the article is as follows. In Section 2 we introduce the convolutive BCA setup that we use throughout the article. In Section 3 the proposed convolutive BCA approach is provided. The corresponding iterative algorithm is provided in Section 4. Numerical examples to illustrate the separation performance for especially the convolutive mixtures of space-time correlated sources are given in Section 5.

## 2. CONVOLUTIVE BCA SETUP

The components of the convolutive BCA setup that we consider throughout the article are as follows:

- We assume a setup with $p$ real sources. The sources are represented by a vector of zero mean (without loss of generality) wide sense stationary process $\{\boldsymbol{s}(k) \in$

This work is supported in part by TUBITAK 112E057 project.

$\Re^p ; k \in \mathbb{Z}\}$ which has the Power Spectral Density (PSD) $\{\mathbf{P_s}(f) ; f \in [-\frac{1}{2}, \frac{1}{2})\}$.

We further assume that the sources have bounded ranges, i.e. $s_i(k) \in [\alpha_i, \beta_i]$ where $\alpha_i, \beta_i \in \Re, \beta_i > \alpha_i$ for $i = 1, ..., p$ and $k \in \mathbb{Z}$. We also define $\gamma_i = \mathcal{R}(s_i) = \beta_i - \alpha_i$ as the range of $s_i$. where $\mathcal{R}(\cdot)$ is the range operator which returns the support length for the pdf of its argument. We decompose the sources as

$$\boldsymbol{s}(k) \stackrel{\triangle}{=} \Upsilon \underline{\boldsymbol{s}}(k), \quad k \in \mathbb{Z}, \tag{1}$$

where $\Upsilon = diag(\gamma_1, \gamma_2, \ldots, \gamma_p)$ is the range matrix of $\boldsymbol{s}$ and $\{\underline{\boldsymbol{s}}(k) \in \Re^p ; k \in \mathbb{Z}\}$ is the normalized source process whose components have unit ranges.

Unlike ICA, we do not assume that sources are independent, or uncorrelated, which implies that $\mathbf{P_s}(f)$ is allowed to be non-diagonal. We assume sources satisfy BCA's domain separability assumption [7], which is a weaker assumption than independence and which can be stated as follows:

– (**A1**) The (convex hulls of the) domain of the sources can be written as the cartesian product of (the convex hulls of the) the individual source domains.

Furthermore, we do not assume independent identically distributed (i.i.d.) samples for sources. The samples can, in fact, be correlated, i.e., $\mathbf{P_s}(f)$ can vary with frequency.

- The source signals are mixed by a MIMO system with a $q$ x $p$ transfer matrix $\mathbb{H}(f)$, whose output is denoted by $\{\boldsymbol{y}(k) \in \Re^q ; k \in \mathbb{Z}\}$. We have

$$\boldsymbol{Y}(f) = \mathbb{H}(f)\mathbf{S}(f), \tag{2}$$

where $\boldsymbol{Y}(f)$ is the Discrete Time Fourier Transform (DTFT) of $\{\boldsymbol{y}(k) \in \Re^q ; k \in \mathbb{Z}\}$, $\mathbf{S}(f)$ is the DTFT of $\{\boldsymbol{s}(k) \in \Re^p ; k \in \mathbb{Z}\}$ and $\mathbb{H}(f) = \sum_{l=0}^{L-1} \boldsymbol{H}(l)e^{-j2\pi fl}$. We assume that $\mathbb{H}(f)$ is an equalizable transfer function of order $L-1$ [9].

- Separator $\mathbb{W}(f)$ is a $q$ x $p$ FIR transfer matrix of order $M-1$ and the separator output sequence is denoted by $\{\boldsymbol{o}(k) \in \Re^p ; k \in \mathbb{Z}\}$ whose DTFT can be written as

$$\boldsymbol{O}(f) = \mathbb{W}(f)\boldsymbol{Y}(f), \tag{3}$$

where $\mathbb{W}(f) = \sum_{l=0}^{M-1} \boldsymbol{W}(l)e^{-j2\pi fl}$. We also define $\tilde{\boldsymbol{W}} = \left[\begin{array}{cccc} \boldsymbol{W}(0) & \boldsymbol{W}(1) & \ldots & \boldsymbol{W}(M-1) \end{array}\right]$ as the separator coefficient matrix.

- The overall system function is defined as

$$\mathbb{G}(f) = \mathbb{W}(f)\mathbb{H}(f) = \sum_{l=0}^{P-1} \boldsymbol{G}(l)e^{-j2\pi fl}, \tag{4}$$

where $P-1$ is the order of overall system. Therefore, the sources $\{\boldsymbol{s}(k) \in \Re^p ; k \in \mathbb{Z}\}$ and the separator outputs $\{\boldsymbol{o}(k) \in \Re^p ; k \in \mathbb{Z}\}$ are related by

$$\boldsymbol{o}(k) = \sum_{l=0}^{P-1} \boldsymbol{G}(l)\boldsymbol{s}(k-l), \quad k \in \mathbb{Z}. \tag{5}$$

Defining $\tilde{\boldsymbol{G}} = \left[\begin{array}{cccc} \boldsymbol{G}(0) & \boldsymbol{G}(1) & \ldots & \boldsymbol{G}(P-1) \end{array}\right]$ and $\tilde{\boldsymbol{s}}(k) = \left[\begin{array}{cccc} \boldsymbol{s}(k) & \boldsymbol{s}(k-1) & \ldots & \boldsymbol{s}(k-P+1) \end{array}\right]^T$, we have $\boldsymbol{o}(k) = \tilde{\boldsymbol{G}}\tilde{\boldsymbol{s}}(k)$, for $k \in \mathbb{Z}$. We obtain the range matrix of $\tilde{\boldsymbol{s}}$ as $\tilde{\Upsilon} = I \otimes \Upsilon$.

## 3. A CONVOLUTIVE BCA APPROACH

In this section, we will extend the instantaneous BCA approach introduced in [8] to the convolutive BSS separation problem. The first objective function in [8], can be modified as

$$J(\boldsymbol{W}) = \frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} log(det(\mathbf{P_o}(f)))df - log\left(\prod_{l=1}^{p} \mathcal{R}(o_i)\right), \tag{6}$$

where $\mathbf{P_o}(f)$ is the PSD of the separator output sequence. In the proposed objective function, the only change is in the log volume of the principal hyper-ellipse, i.e., the first term. The principal hyper-ellipse volume definition is extended from sample based correlation information to process based correlation information, capturing inter-sample correlations.

The following theorem shows that the proposed objective is useful for achieving separation of convolutive mixtures whose setup is outlined in Section 2.

**Theorem:** Assuming the setup in Section 2 and $\boldsymbol{H}(f)$ is equalizable by an FIR separator matrix of order $M-1$, the set of global maxima for $J$ in (6) is equal to a set of perfect separator matrices.

**Proof:** Using the fact that

$$\mathbf{P_o}(f) = \mathbb{G}(f)\Upsilon\mathbf{P}_{\underline{\mathbf{s}}}(f)\Upsilon^T\mathbb{G}(f)^H,$$

we obtain

$$\begin{aligned}
&\int_{-\frac{1}{2}}^{\frac{1}{2}} log(det(\mathbf{P_o}(f)))df = \\
&\int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(|det(\mathbb{G}(f)\Upsilon)|^2 det(\mathbf{P}_{\underline{\mathbf{s}}}(f))\right) df = \\
&\int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(|det(\mathbb{G}(f)\Upsilon)|^2\right) df + \int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(det(\mathbf{P}_{\underline{\mathbf{s}}}(f))\right) df.
\end{aligned} \tag{7}$$

Using the Hadamard inequality [10] yields

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(|det(\mathbb{G}(f)\Upsilon)|^2\right) df \leq$$

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(\prod_{m=1}^{p} ||\left(\mathbb{G}(f)\Upsilon\right)_{m,:}||_2^2\right) df =$$

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{m=1}^{p} log\left(||\left(\mathbb{G}(f)\Upsilon\right)_{m,:}||_2^2\right) df =$$

$$\sum_{m=1}^{p} \int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(||\left(\mathbb{G}(f)\Upsilon\right)_{m,:}||_2^2\right) df, \quad (8)$$

where $\left(\mathbb{G}(f)\Upsilon\right)_{m,:}$ is the $m^{th}$ row of $\mathbb{G}(f)\Upsilon$. From Jensen's inequality [11], for $m = 1, ..., p$, we have

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(||\left(\mathbb{G}(f)\Upsilon\right)_{m,:}||_2^2\right) df \leq$$

$$log\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} ||\left(\mathbb{G}(f)\Upsilon\right)_{m,:}||_2^2 df\right). \quad (9)$$

The use of Parseval's theorem yields

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} ||\left(\mathbb{G}(f)\Upsilon\right)_{m,:}||_2^2 df = ||(\tilde{\boldsymbol{G}}\tilde{\Upsilon})_{m,:}||_2^2. \quad (10)$$

Thus, from (8-10), we obtain

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(|det(\mathbb{G}(f)\Upsilon)|^2\right) df \leq \sum_{m=1}^{p} log\left(||(\tilde{\boldsymbol{G}}\tilde{\Upsilon})_{m,:}||_2^2\right), \quad (11)$$

which further implies,

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} log(det(\mathbf{P_o}(f)))df \leq \sum_{m=1}^{p} log\left(||(\tilde{\boldsymbol{G}}\tilde{\Upsilon})_{m,:}||_2^2\right) + \int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(det(\mathbf{P}_{\underline{\mathbf{s}}}(f))\right) df. \quad (12)$$

As a result,

$$J(\boldsymbol{W}) \leq \frac{1}{2}\sum_{m=1}^{p} log\left(||(\tilde{\boldsymbol{G}}\tilde{\Upsilon})_{m,:}||_2^2\right) - log\left(\prod_{l=1}^{p} \mathcal{R}(o_i)\right) + \frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(det(\mathbf{P}_{\underline{\mathbf{s}}}(f))\right) df. \quad (13)$$

Under the BCA's domain separability assumption (A1) stated in Section 2, we can write the range of $i^{th}$ component of $\boldsymbol{o}$ as $\mathcal{R}(o_i) = ||\tilde{\boldsymbol{G}}_{i,:}\tilde{\Upsilon}||_1$. We can further define $\boldsymbol{Q} \triangleq \tilde{\boldsymbol{G}}\tilde{\Upsilon}$, the range vector for the separator outputs can be rewritten as

$$\mathcal{R}(\boldsymbol{o}) = \begin{bmatrix} ||\boldsymbol{Q}_{1,:}||_1 & ||\boldsymbol{Q}_{2,:}||_1 & \dots & ||\boldsymbol{Q}_{m,:}||_1 \end{bmatrix}. \quad (14)$$

If we rewrite the inequality (13) in terms of $\boldsymbol{Q}$ we obtain

$$\begin{aligned} J(\boldsymbol{W}) &\leq \sum_{m=1}^{p} log\left(||\boldsymbol{Q}_{m,:}||_2\right) - log\left(\prod_{m=1}^{p} ||\boldsymbol{Q}_{m,:}||_1\right), \\ &+ \frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(det(\mathbf{P}_{\underline{\mathbf{s}}}(f))\right) df \\ &= \sum_{m=1}^{p} log\left(||\boldsymbol{Q}_{m,:}||_2\right) - \sum_{m=1}^{p} log\left(||\boldsymbol{Q}_{m,:}||_1\right) \\ &+ \frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(det(\mathbf{P}_{\underline{\mathbf{s}}}(f))\right) df. \end{aligned} \quad (15)$$

Note that,

$$\sum_{m=1}^{p} log\left(||(\boldsymbol{Q}_{m,:})||_2\right) \leq \sum_{m=1}^{p} log\left(||\boldsymbol{Q}_{m,:}||_1\right), \quad (16)$$

due to the ordering $||\boldsymbol{q}||_1 \geq ||\boldsymbol{q}||_2$ for any $\boldsymbol{q}$. Therefore,

$$J(\boldsymbol{W}) \leq \frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} log\left(det(\mathbf{P}_{\underline{\mathbf{s}}}(f))\right) df. \quad (17)$$

The inequality in (16) is achieved if and only if each row of $\boldsymbol{Q}$ has only one non-zero element which results in each row of $\tilde{\boldsymbol{G}}$ has only one non-zero element. The inequality in (8) is achieved if and only if the rows of $\mathbb{G}(f)$ are perpendicular to each other which yields that the rows of $\tilde{\boldsymbol{G}}$ should be perpendicular to each other and the non-zero elements should not be positioned in the same indexes with respect to mod $p$.

As a result, the inequality in (17) is achieved if and only if $\tilde{\boldsymbol{G}}$ corresponds to perfect separator transfer matrix in the form

$$\mathbb{G}(z) = diag(\alpha_1 z^{-d_1}, \alpha_2 z^{-d_2}, \dots, \alpha_p z^{-d_p})\boldsymbol{P}$$

where $\alpha_k$'s are non-zero real scalings, and $d_k$'s are non-negative integer delays. The FIR equalizability of the mixing system implies the existence of such parameters.

## 4. ADAPTIVE IMPLEMENTATION

In this section, we provide the adaptive algorithm corresponding to the optimization setting in (6). In the adaptive implementation, we assume a set of finite observations of mixtures $\{\boldsymbol{y}(0), \boldsymbol{y}(1), ..., \boldsymbol{y}(N-1)\}$ and modify the objective as

$$J(\mathbf{W}) = \sum_{l=-\nu}^{\nu} \frac{1}{2\eta} log(det(\hat{\mathbf{P}}_{\mathbf{o}}(l))) - log\left(\prod_{l=1}^{p} \hat{\mathcal{R}}(o_i)\right), \quad (18)$$

where $\nu = N + M - 1, \eta = 2\nu + 1$ is the DFT size and we use the PSD estimate for the separator outputs given by

$$\hat{\mathbf{P}}_{\mathbf{o}}(l) = \sum_{k=-\nu}^{\nu} \hat{\boldsymbol{R}}_{\mathbf{o}}(k) e^{-j2\pi lk/\eta}, \quad (19)$$

for $l \in \{-\nu, ..., \nu\}$, where $N$ is the number of samples and $\hat{\boldsymbol{R}}_{\mathbf{o}}$ is the output sample autocovariance function, defined as

$$\hat{\boldsymbol{R}}_{\mathbf{o}}(k) = \frac{1}{\nu + 1 - |k|} \sum_{q=max(0,-k)}^{min(\nu,\nu-k)} \boldsymbol{o}(q)\boldsymbol{o}^T(q+k), \quad (20)$$

for $k = -\nu, ..., \nu$. We point out that we use $\hat{\mathcal{R}}(\boldsymbol{o})$ for the range vector of the sample outputs for which we have

$$\hat{\mathcal{R}}(o_i) = \max_{k \in \{1,2,...,N\}} o_i(k) - \min_{k \in \{1,2,...,N\}} o_i(k), \quad (21)$$

for $i = 1, 2, ..., p$. Note that the derivative of the first part of $J(\boldsymbol{W})$ with respect to $\boldsymbol{W}(n)$ is

$$\frac{1}{2\eta} \frac{\partial \sum_{l=-\nu}^{\nu} log(det(\hat{\mathbf{P}}_{\mathbf{o}}(l)))}{\partial \mathbf{W}(n)} = \frac{1}{\eta} \sum_{l=-\nu}^{\nu} \Re e \left\{ \hat{\mathbf{P}}_{\mathbf{o}}(l)^{-1} \hat{\mathbb{W}}(l) \hat{\mathbf{P}}_{\mathbf{y}}(l) e^{j2\pi nl/\eta} \right\}, \quad (22)$$

where

$$\hat{\mathbb{W}}(l) = \sum_{k=-\nu}^{\nu} \boldsymbol{W}(k) e^{-j2\pi lk/\eta}, \quad (23)$$

and

$$\hat{\mathbf{P}}_{\mathbf{y}}(l) = \sum_{k=-\nu}^{\nu} \hat{\boldsymbol{R}}_{\mathbf{y}}(k) e^{-j2\pi lk/\eta}. \quad (24)$$

Therefore, the iterative update equation is as

$$\boldsymbol{W}^{(i+1)}(n) = \boldsymbol{W}^{(i)}(n) + \mu^{(i)} \left( \frac{1}{\eta} \sum_{l=-\nu}^{\nu} \Re e \left\{ \hat{\mathbf{P}}_{\mathbf{o}}(l)^{-1} \hat{\mathbb{W}}(l) \hat{\mathbf{P}}_{\mathbf{y}}(l) e^{j2\pi nl/\eta} \right\} - \sum_{m=1}^{p} \frac{1}{\mathbf{e}_m^T \hat{\mathcal{R}}(\boldsymbol{o}_{\mathbf{W}^{(i)}})} \mathbf{e}_m \left( \boldsymbol{y}(l_m^{max(i)}) - \boldsymbol{y}(l_m^{min(i)}) \right) \right), \quad (25)$$

where $\mu^{(i)}$ is the step-size at the $i^{th}$ iteration and $l_m^{max(i)}$ ($l_m^{min(i)}$) is the sample index for which the maximum (minimum) value for the $m^{th}$ separator output is achieved at the $i^{th}$ iteration.

## 5. NUMERICAL EXAMPLES AND CONCLUSION

We consider the following scenario to illustrate the separation capability of the proposed algorithm for the convolutive mixtures of space-time correlated sources: In order to generate space-time correlated sources, we first generate a samples of a $\tau p$ size vector, $\mathbf{d}$, with zero-mean adjusted Copula-t distribution, a perfect tool for generating vectors with controlled correlation, with 4 degrees of freedom whose correlation matrix parameter is given by $\mathbf{R} = \mathbf{R}_t \otimes \mathbf{R}_s$ where $\mathbf{R}_t$ ($\mathbf{R}_s$) is a Toeplitz matrix whose first row is $\begin{bmatrix} 1 & \rho_t & \dots & \rho_t^{\tau-1} \end{bmatrix}$ ( $\begin{bmatrix} 1 & \rho_s & \dots & \rho_s^{p-1} \end{bmatrix}$). Each sample of $\mathbf{d}$ is partitioned to produce source vectors, $\mathbf{d}(k) = \begin{bmatrix} \mathbf{s}(k\tau) & \mathbf{s}(k\tau+1) & \dots & \mathbf{s}((k+1)\tau-1) \end{bmatrix}$. Therefore, we obtain the source vectors as samples of a wide-sense cyclostationary[1] process whose correlation structure in time direction and space directions are governed by the parameters $\rho_t$ and $\rho_s$, respectively.

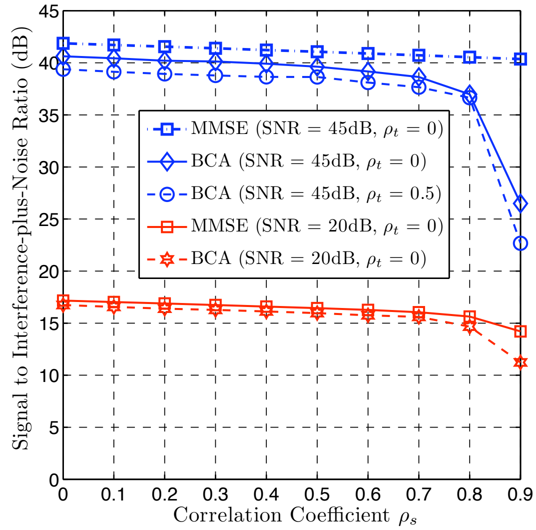


**Fig. 1**. Dependent convolutive mixtures separation performance results

In the simulations, we considered a scenario with 3 sources and 5 mixtures, an i.i.d. Gaussian convolutive mixing system with order 3 and a separator of order 5. At each run, we generate 50000 source vectors where $\tau$ is set as 5. Figure 1 shows the output total Signal energy to total Interference+Noise energy (over all outputs) Ratio (SINR) obtained for the proposed approach for two different SNRs (45dB and 20dB), and various space and time correlation parameters. SINR performance of Minimum Mean Square Error (MMSE) filter of the same order, which uses full information about mixing system and source/noise statistics, is also shown to evaluate the relative success of the proposed approach. These results demonstrate that the proposed algorithm's performance closely follows its MMSE counterpart for a wide range of correlation values. Therefore, we obtain a convolutive extension of the BCA approach introduced in [8], which is capable of separating convolutive mixtures of space-time correlated sources.

[1] This actually violates the stationarity assumption on sources when $\rho_t \neq 0$. However, we still use this as a convenient method to generate space-time correlated sources

## 6. REFERENCES


[1] Pierre Comon and Christian Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, 2010.

[2] Aapo Hyvärinen, Juha Karhunen, and Erkki Oja, *Independent Component Analysis*, John Wiley and Sons Inc., 2001.

[3] Dinh-Tuan Pham, "Blind separation of instantenaous mixtrures of sources based on order statistics," *IEEE Trans. on Signal Processing*, vol. 48, pp. 363–375, February 2000.

[4] S. Cruces and I. Duran, "The minimum support criterion for blind source extraction: a limiting case of the strenthened young's inequality," *Lecture Notes in Computer Science*, pp. 57–64, September 2004.

[5] F. Vrins, M. Verleysen, and C. Jutten, "SWM: A class of convex contrasts for source separation," *Proceedings of IEEE ICASSP 2005*, vol. V, pp. 161–164, March 2005.

[6] Alper T. Erdogan, "A simple geometric blind source separation method for bounded magnitude sources," *IEEE Trans. on Signal Processing*, vol. 54, pp. 438–449, February 2006.

[7] S. Cruces, "Bounded component analysis of linear mixtures: A criterion for minimum convex perimeter," *IEEE Trans. on Signal Process.*, vol. 58, no. 4, pp. 2141–2154, April 2010.

[8] Alper T. Erdogan, "A family of bounded component analysis algorithms," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1881–1884, March 2012.

[9] Yuijro Inouye and Ruey-Wen Liu, "A system-theoretic foundation for blind equalization of an FIR MIMO channel system," *IEEE Trans. on CAS-I: Fund. Th. and Apps.*, vol. 49, no. 4, pp. 425–436, April 2002.

[10] D.J.H. Garling, *Inequalities: a journey into linear analysis*, Cambridge University Press, 2007.

[11] W. Rudin, *Real and Complex Analysis*, McGraw-Hill Book Co., 1987.