AN LMS VIEWPOINT ON THE LOCAL STABILITY OF SECOND ORDER BLIND SOURCE SEPARATION

Julien Bourgeois

DaimlerChrysler Research and Technology
P.O. Box 23 60 – 89013 Ulm – Germany
julien.bourgeois@daimlerchrysler.com

ABSTRACT

The local stability of blind source separation (BSS) algorithms has been mostly investigated in the vicinity of the mixing system inverse, i.e. assuming not only separation but also deconvolution of the sources. In the case of convolutive mixtures, other equilibrium points exist that do not deconvolve the sources. In this paper, we examine the local stability at such an equilibrium point for a second-order BSS algorithm based on the source non-stationarity. Considering explicitly source silences (for example with speech sources) greatly simplifies the obtained local stability conditions and allows a comparison with the classical LMS noise canceller. This analysis reveals a discrepancy between causal and double-sided (acausal) mixing channels.

1. INTRODUCTION

Blind source separation (BSS) is an emerging technique that tackles the problem of recovering several source signals from a set of observed mixtures. A famous example is the so-called cocktail party problem. Consider the scenario in which several speakers talk simultaneously and are recorded by a microphone array. BSS is an appropriate technique to separate the individual speech flows, in particular when no prior information on the microphone arrangement or on the source position is available. BSS relies solely on the assumption that the source signals are mutually independent, which makes it unnecessary to model the source-sensor propagation. Nevertheless, in many practical setups, information on the relative source-microphone position is available. Then, adaptive beamforming is another relevant method.

These two approaches are closely related. As pointed out by Araki *et. al.*, there is an equivalence between BSS and beamforming, since both achieve the separation by spatially filtering the observed signals [1]. However, this *functional* equivalence does not account for *algorithmic* differences that are significant from a practical point of view.

The noise canceller (the adaptive component of a beamformer) is typically driven by the Normalized Least-Mean Square algorithm (NLMS), which minimizes a quadratic power criterion. The *global* convergence of this algorithm is well-known [10]. By contrast, the parameter update in BSS algorithms usually involves non-linear terms. This makes an analysis of the global convergence hardly tractable [8, 5]. In this paper, we will compare the *local* behavior of NLMS and BSS in the vicinity of their equilibrium points, which is described in particular by the local stability.

Stability conditions of the NLMS algorithm are very simple. They depend only on the adaptation step-size. On the other hand, the local stability of BSS algorithms has been investigated in the vicinity of the inverse of the mixing system. There, the sources are not only separated but also deconvolved, possibly using a feedback filter architecture [6, 11]. However, in many practical cases, only the separation of the sources is desired (or achievable). Therefore, we will study the stability around an equilibrium point that separates but does not deconvolve the sources.

This paper is organized as follows: Section 2 defines the mixture/separation model and the notations. In Section 3, we specify which BSS algorithm is under consideration. The stability analysis is carried out in Section 4. The results are discussed in Section 6. Section 7 concludes.

2. MODEL AND NOTATIONS

A common framework for BSS and beamforming is easily obtained when we consider the problem of separating the two-source two-sensor convolutive mixture depicted in Fig. 1 (left). The sources $\mathbf{s}(t) = (s_1(t), s_2(t))^T$ propagate over unknown multipath channels to the observations $\mathbf{x}(t) = (x_1(t), x_2(t))^T$. In general, the source-observation relationship can be expressed in the z-domain as

$$\mathbf{X}(z) = \begin{pmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{pmatrix} \mathbf{S}(z) \quad (1)$$

where $H_{ij}(z) = \sum_{p\in\mathbb{Z}} h_{ij}(p) z^{-p}$. h_{ij} is the impulse response from source j to sensor i. Let us assume that the

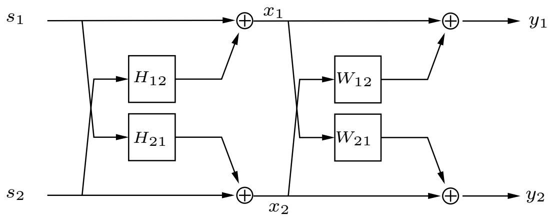

Fig. 1. Left: the mixing channels – right: the separation filters

direct channels $H_{ii}(z)$ are invertible. Then, the mixing can be equivalently rewritten as

$$\begin{aligned} \mathbf{X}(z) &= \mathbf{H}(z)\mathbf{S}(z), & (2) \\ \text{with } \mathbf{H}(z) &= \begin{pmatrix} 1 & H_{12}(z) \\ H_{21}(z) & 1 \end{pmatrix}, & (3) \end{aligned}$$

keeping in mind that $H_{12}(z)$ and $H_{21}(z)$ represent ratio of acoustic transfer functions. The channels $H_{12}(z)$ and $H_{21}(z)$ characterize the amount of cross-talk in the microphones. The observation-output relationship is written in the z-domain as

$$\begin{aligned} \mathbf{Y}(z) &= \mathbf{W}(z)\mathbf{X}(z) & (4) \\ \mathbf{W}(z) &= \begin{pmatrix} 1 & W_{12}(z) \\ W_{21}(z) & 1 \end{pmatrix}, & (5) \end{aligned}$$

where $W_{12}(z)$ and $W_{12}(z)$ are finite-impulse-response filters (FIR). The separation matrix $\mathbf{W}(z)$ must inverse $\mathbf{H}(z)$ in the sense that the off-diagonal terms of the global system $\mathbf{W}(z)\mathbf{H}(z)$ vanish, whereas the diagonal terms are non-zero. Under the constraint that diag $\mathbf{W}(z) = \mathbf{I}_2$ ($\mathbf{I}_N$ denotes the $N \times N$ identity matrix, the operator diag $\mathbf{A}$ sets the off-diagonal elements of $\mathbf{A}$ to zero), this is achieved uniquely for

$$\mathbf{W}(z) = \begin{pmatrix} 1 & -H_{12}(z) \\ -H_{21}(z) & 1 \end{pmatrix}. \quad (6)$$

With this mixture-separation model, beamforming and BSS tackle the same *identification* problem. (In practice, adaptive beamforming performs a noise cancellation task, i.e. the optimal filters depend on the source signals.)

At the equilibrium (6), the source-output transfer function is

$$W_{\text{eq}}(z) = 1 - H_{12}(z)H_{21}(z). \quad (7)$$

Since $W_{\text{eq}}(z)$ is not a unit response, the source are not deconvolved at the equilibrium point. For later reference, we denote its time domain counterpart by $w_{\text{eq}} = \delta - h_{12} * h_{21}$.

In the next sections, we will consider causal (resp. double-sided) mixing channels and separation filters h_{ij}, w_{ij} with support $\mathcal{S}_c = [0, L-1]$ (resp. $\mathcal{S}_d = [-L+1, L-1]$). The case of double-sided filters is presented using positive and negative filters taps $w_{ij}(k), -L+1 \leq k \leq L-1$. We define the truncation operator on the z-tranform $C(z)$, denoted by $[C(z)]_{\mathcal{S}}$, as

$$[C(z)]_{\mathcal{S}} = \sum_{k \in \mathcal{S}} c(k) z^{-k}. \quad (8)$$

$\mathcal{S}$ refers to the filter support: $\mathcal{S} = \mathcal{S}_c$ in the causal case, $\mathcal{S} = \mathcal{S}_d$ for double-sided filters.

Concerning the sources, we consider blockwise stationary white source signals with powers $\sigma_{1,i}^2$ and $\sigma_{2,i}^2$, where $i = 1, \ldots, I$ denotes the time block index. We will pay particular attention to the case where these sources have periods of silence:

$$\exists (i_0, j_0) \text{ with } i_0 \neq j_0 \text{ and } \sigma_{1,i_0}^2 = \sigma_{1,j_0}^2 = 0.$$

Let us define the output vector in the time domain $\mathbf{y}(t) = (y_1(t), y_2(t))^T$ and the output correlation matrix at the i-th data block $\mathbf{R}_{\mathbf{yy}}^{(i)}(k) = \mathbf{E}_i\{\mathbf{y}(t)\mathbf{y}^T(t-k)\}$. In practice, $\mathbf{R}_{\mathbf{yy}}^{(i)}(k)$ is estimated by averaging $\mathbf{y}(t)\mathbf{y}^T(t-k)$ over the i-th data block. This estimation is denoted by $\hat{\mathbf{R}}_{\mathbf{yy}}^{(i)}(k)$. The sequence $\hat{\mathbf{R}}_{\mathbf{yy}}^{(i)}(k), k \in \mathbb{Z}$ has the z-transform

$$\hat{\mathbf{\Phi}}_{\mathbf{yy}}^{(i)}(z) = \sum_{k \in \mathbb{Z}} \hat{\mathbf{R}}_{\mathbf{yy}}^{(i)}(k) z^{-k}. \quad (9)$$

Similarly, we denote by $\hat{\mathbf{R}}_{\mathbf{yx}}^{(i)}(k)$ (resp. $\hat{\mathbf{\Phi}}_{\mathbf{yx}}^{(i)}(z)$) the input-output cross-correlation in the time domain (resp. z-domain).

The adaptation of the filters $\mathbf{W}(z)$ is performed iteratively according to

$$\mathbf{W}_{n+1}(z) = \mathbf{W}_n(z) - \mu \Delta \mathbf{W}_n(z). \quad (10)$$

In the following, we will consider batch algorithms, which means that $\Delta \mathbf{W}_n(z)$ is computed using the whole available data set.

3. SECOND-ORDER STATISTICS ALGORITHMS

Applied separately for each source, the beamforming approach consists of two adaptive interference cancellers, i.e. two decoupled identification problems, as illustrated in Fig. 2 (top). For example, the cancellation of $s_2(t)$ in $y_1(t)$ requires the identification of h_{12}. This approach applies best when no contribution of the desired source $s_1(t)$ is present at the sensor $x_2(t)$: w_{12} must be adapted only if $h_{21} * s_1(t)$ is weak compared to $s_2(t)$. Otherwise, this leads to the well-known power inversion effect associated with least-square methods [10]. In speech applications, the adaptation occurs if the desired speaker is silent. The adaptation of two interference cancellers minimizes the output

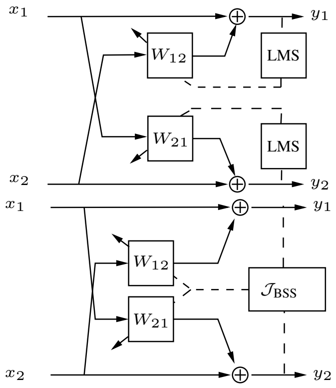

Fig. 2. Top: separation with two decoupled interference cancellers (11) – bottom: separation with BSS (13)

powers when the one or the other source is silent. With the NLMS algorithm, the parameter update can be written as:

$$\Delta \mathbf{W}_n(z) = \sum_{i=1}^{I} \mathbf{Detect}^{(i)} \mathbf{N}_{NLMS}^{(i)} \left[\text{off} \left(\hat{\mathbf{\Phi}}_{\mathbf{yx}}^{(i)}(z) \right) \right]_{\mathcal{S}}. \quad (11)$$

The normalization is given by

$$\hat{\mathbf{N}}_{NLMS}^{(i)} = \begin{pmatrix} \hat{R}_{x_2x_2}^{(i)}(0) & 0 \\ 0 & \hat{R}_{x_1x_1}^{(i)}(0) \end{pmatrix}^{-1}.$$

The silence detection is represented by the 2×2 matrix $\mathbf{Detect}^{(i)}$ that contains only zeros, except the (p, p) element which is 1 if $\sigma_{p,i}^2 = 0$, i.e. if $s_p(t) = 0$. This models a perfect silence detection.

BSS techniques are robust in the sense that even if $s_1(t)$ has significant power, they are able to cancel the interferer $s_2(t)$ in the output $y_1(t)$. The separation filters w_{12} and w_{21} *jointly* minimize $\mathcal{J}_{\text{BSS}}$, a dependence measure that is built on an assumed source probability distribution, as depicted in Fig. 2 (bottom). Fundamentally, BSS techniques require no adaptation control.

Let us consider the second-order BSS algorithm that is proposed in [2]. Assuming blockwise stationary white Gaussian sources, the output mutual information is given by [4]

$$\mathcal{J}_{\text{BSS}} = \sum_{i=1}^{I} \log\det \text{diag}\, \hat{\mathbf{R}}_{\mathbf{yy}}^{(i)}(0) - \log\det \hat{\mathbf{R}}_{\mathbf{yy}}^{(i)}(0). \quad (12)$$

Therefore, $\mathcal{J}_{\text{BSS}}$ is a measure of the output independence. In fact, $\mathcal{J}_{\text{BSS}}$ can be seen as the measure of diagonality for the matrices $\hat{\mathbf{R}}_{\mathbf{yy}}^{(i)}(0)$: Minimizing $\mathcal{J}_{\text{BSS}}$ results in the simultaneous diagonalization of these correlation matrices.

To give the learning rule $\Delta \mathbf{W}_n(z)$ for BSS, we need to define the normalization matrix with $\hat{\mathbf{N}}_{BSS}^{(i)} = (\text{diag}\, \hat{\mathbf{R}}_{\mathbf{yy}}^{(i)}(0))^{-1}$. The minimization of $\mathcal{J}_{\text{BSS}}$ is performed with the natural gradient under the constraint that $\text{diag}\, \mathbf{W}_n(z) = \mathbf{I}_2$ using

$$\Delta \mathbf{W}_n(z) = \sum_{i=1}^{I} \hat{\mathbf{N}}_{BSS}^{(i)} \left[\text{off}\, \hat{\mathbf{\Phi}}_{\mathbf{yy}}^{(i)}(z) \right]_{\mathcal{S}}, \quad (13)$$

where $\text{off}\, \mathbf{A} = \mathbf{A} - \text{diag}\, \mathbf{A}$.

4. LOCAL STABILITY ANALYSIS IN THE MEAN

Let us define the mean error

$$\mathbf{V}_n(z) = \mathbf{E}\{\text{off}\, (\mathbf{W}_n(z) + \mathbf{H}(z))\} \quad (14)$$

at the n-th iteration. According to (13), the evolution of the mean error is expressed in the time domain with $v_{ij,n}(k), k \in \mathcal{S}$ as

$$\begin{cases} v_{12,n+1}(k) = v_{12,n}(k) - \mu \sum_{i=1}^{I} R_{y_1y_2}^{(i)}(k) / \mathbf{E}_i\{y_1^2(t)\}, \\ v_{21,n+1}(k) = v_{21,n}(k) - \mu \sum_{i=1}^{I} R_{y_2y_1}^{(i)}(k) / \mathbf{E}_i\{y_2^2(t)\}. \end{cases} \quad (15)$$

For an analysis in the mean, all estimated quantities in (13) are replaced by their true values. The output correlations are quadratic functions of $v_{12,n}, v_{21,n}$. Therefore, (15) is a non-linear recursion. We linearize (15) in the vicinity of $\mathbf{V}(z) = \mathbf{0}$ and the step-size is assumed to be sufficiently small to ensure that permutation of the sources does not occur. The linearization of, e.g., $\mathbf{R}_{y_1y_2}^{(i)}(k) = \mathbf{E}_i\{y_1(t)y_2(t-k)\}$ for v_{ij} close to zero and decorrelated white source signals is given by[1]

$$\begin{aligned} \mathbf{R}_{y_1y_2}^{(i)}(k) \quad &= \quad \sigma_{1,i}^2 \sum_{v \in \mathcal{S}} v_{21}(v) w_{\text{eq}}(k+v) \quad (16) \\ &+ \sigma_{2,i}^2 \sum_{v \in \mathcal{S}} w_{\text{eq}}(v) v_{12}(k+v) + \mathcal{O}(v_{12}, v_{21}) \end{aligned}$$

where $\mathcal{O}(v_{12}, v_{21})$ contains the quadratic dependencies in (v_{12}, v_{21}). Combining (15) and (16), we can write the first-order evolution of $v_{12}(k)$ as

$$\begin{aligned} \Delta v_{12}(k) \quad &= \quad \frac{1}{\|w_{eq}\|_2^2} \sum_{i=1}^{I} \left(\sum_{v \in \mathcal{S}} v_{21}(v) w_{\text{eq}}(k+v) \right. \\ &+ \left. \frac{\sigma_{2,i}^2}{\sigma_{1,i}^2} \sum_{v \in \mathcal{S}} w_{\text{eq}}(v) v_{12}(k+v) \right), \quad (17) \end{aligned}$$

where $\|w_{eq}\|_2^2 = \sum_k w_{eq}^2(k)$. The source signals act upon the BSS algorithm through the two following variables:

$$\Theta = \sum_{i=1}^{I} \frac{\sigma_{2,i}^2}{\sigma_{1,i}^2}, \text{ and } \Theta' = \sum_{i=1}^{I} \frac{\sigma_{1,i}^2}{\sigma_{2,i}^2}.$$

[1] The iteration index n is omitted for notational convenience.

Θ and Θ' quantify the relative nonstationarity of the sources. It can be shown that the (Θ, Θ')-space is spanned by $I = 2$ and

$$(\sigma_{1,1}^2, \sigma_{1,2}^2, \sigma_{2,1}^2, \sigma_{2,2}^2) = (1, \sigma_{1,2}^2, \sigma_{2,1}^2, 1),\ \forall\, \sigma_{1,2}^2, \sigma_{2,1}^2 \geq 0.$$

All possible runs of the algorithm with respect to the source distribution are covered with two time blocks. Without loss of generality, we set $I = 2, \sigma_{1,1}^2 = \sigma_{2,2}^2 = 1$ in the following.

With locally silent sources, the NLMS algorithm converges to $\mathbf{V}_n(z) = \mathbf{0}$. In this case, i.e. when $\sigma_{1,2}^2 = \sigma_{2,1}^2 = 0$, we have

$$R_{y_i x_j}^{(j)}(k) = \mathbf{E}_j\{x_j^2(t)\} v_{ij}(k) \text{ for } i, j = 1, 2, i \neq j. \quad (18)$$

The evolution of the error (v_{12}, v_{21}) for the NLMS adaptation (11) can be written as

$$v_{ij,n+1}(k) = v_{ij,n}(k)(1 - \mu), \quad (19)$$

which means that all error terms converge to zero with the same speed controlled by $1 - \mu$. Therefore, the NLMS algorithm is stable in the mean for $0 < \mu < 2$, independently of the mixing channels.

4.1. Double-sided systems

Similarly to (17), we can linearize $v_{21,n+1}(k)$ in (15). For $\mathcal{S} = \mathcal{S}_d$, the linearized system can be written in matrix form as:

$$\mathbf{v}_{n+1} = \left(\mathbf{I}_{4L-2} - \mu \frac{1}{\|w_{eq}\|_2^2} \mathbf{D}\right) \mathbf{v}_n, \quad (20)$$

$$\mathbf{D} = \begin{pmatrix} \Theta \mathbf{H} & f(\mathbf{H}) \\ f(\mathbf{H}) & \Theta' \mathbf{H} \end{pmatrix}, \quad (21)$$

$$\mathbf{H} = \begin{pmatrix} w_{\text{eq}}(0) & \cdots & w_{\text{eq}}(2L-2) \\ \vdots & \ddots & \vdots \\ w_{\text{eq}}(-2L+2) & \cdots & w_{\text{eq}}(0) \end{pmatrix}, \quad (22)$$

$$\mathbf{v}_n = (\mathbf{v}_{12,n}^T, \mathbf{v}_{21,n}^T), \quad (23)$$

$$\mathbf{v}_{ij,n} = (v_{ij,n}(-L+1), \ldots, v_{ij,n}(L+1))^T. \quad (24)$$

The function $f(\mathbf{A})$ returns the $M \times N$ matrix $\mathbf{A}$ with columns preserved and rows flipped in the up/down direction: $f(\mathbf{A})_{m,n} = \mathbf{A}_{(M-m+1),n}$, $m = 1, \ldots, M$. If $\Theta = \Theta' = 1$, which corresponds to stationary sources, $\mathbf{D}$ is singular and the algorithm is unstable. Another remark about (20) is that $\mathbf{D}$ is no correlation matrix (it is not positive definite). The positiveness of its eigenvalues, which is required for local stability, is not guaranteed. Although $\mathbf{H}$ has a Toeplitz structure, $\mathbf{D}$ is only banded by block. Therefore, the Fourier tools cannot be applied to evaluate the eigenvalues of $\mathbf{D}$. Moreover, $\mathbf{D}$ does not present a favorable form (e.g., triangular) that would provide a closed formula of its eigenvalues.

However, the particular case of locally silent sources opens a possibility for further study. Let us set $\sigma_{1,2}^2 = \sigma_{2,1}^2 = \sigma^2$ and study the limit $\sigma^2 \to 0$. A key observation is that

$$\lim_{\sigma^2 \to 0} \Theta = \lim_{\sigma^2 \to 0} \Theta' = +\infty,$$

and subsequently the diagonal submatrices of $\mathbf{D}$ become dominant. The local stability relates to the *direction* given by the update, not to the norm. Therefore, we can choose an arbitrarily small step-size μ. In order to balance the normalization, we set the step-size to $\mu = \mu_0 \|w_{eq}\|^2 \sigma^2$. Then, the limit $\lim_{\sigma^2 \to 0} \mu \mathbf{D}$ exists and we can neglect the off-diagonal submatrices of $\mathbf{D}$. The eigenvalues of $\mathbf{D}$ are those of $\mathbf{H}$, since $\det(\mathbf{I}_{4L-2} - \lambda \mathbf{D}) = \det^2(\mathbf{I}_{2L-1} - \lambda \mathbf{H})$ for any λ. Therefore, a *necessary and sufficient* local stability condition for locally silent sources is that the eigenvalues $\lambda_k^{\mathbf{H}}$ of $\mathbf{H}$ satisfy

$$\forall k \;\; \Re\{\lambda_k^{\mathbf{H}}\} > 0 \text{ where } \Re\{z\} \text{ is the real part } z \in \mathbb{C}. \quad (25)$$

This is not very enlightening. Let us reformulate the condition (25) with the Discrete Fourier Transform (DFT). Let $\mathbf{H}_c$ be the square circulant matrix defined with its first column $\mathbf{c}_1 = (w_{\text{eq}}(0), \ldots, w_{\text{eq}}(-2L+2), w_{\text{eq}}(2L-2), \ldots, w_{\text{eq}}(1))^T$. Its $(2L-1) \times (2L-1)$ upper left block is $\mathbf{H}$. $\mathbf{H}_c$ is diagonalized by the $(4L-3) \times (4L-3)$ Fourier matrix $\mathbf{F}$: $\mathbf{H}_c = \mathbf{F} \boldsymbol{\Delta} \mathbf{F}^H$, where $\boldsymbol{\Delta}$ is diagonal and contains the values of the DFT of $\mathbf{c}_1$. We denote the $M \times N$ zero matrix by $\mathbf{0}_{M \times N}$ ($\mathbf{0}_M$ if $M = N$) and define

$$\boldsymbol{\Pi} = \begin{pmatrix} \mathbf{I}_{2L-1} & \mathbf{0}_{(2L-1)\times(2L-2)} \\ \mathbf{0}_{(2L-2)\times(2L-1)} & \mathbf{0}_{(2L-2)\times(2L-2)} \end{pmatrix},$$

$$\boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Pi} & \mathbf{0}_{(4L-3)\times(4L-3)} \\ \mathbf{0}_{(4L-3)\times(4L-3)} & \boldsymbol{\Pi} \end{pmatrix}.$$

Then (20) becomes

$$\underbrace{\begin{pmatrix} \mathbf{v}_{12,n+1} \\ \mathbf{0}_{(2L-2)\times 1} \\ \mathbf{v}_{21,n+1} \\ \mathbf{0}_{(2L-2)\times 1} \end{pmatrix}}_{\tilde{\mathbf{v}}_{n+1}} = \boldsymbol{\Omega} \left(\mathbf{I}_{8L-6} - \mu_0 \underbrace{\begin{pmatrix} \mathbf{H}_c & \mathbf{0}_{4L-3} \\ \mathbf{0}_{4L-3} & \mathbf{H}_c \end{pmatrix}}_{\tilde{\mathbf{D}}} \right) \tilde{\mathbf{v}}_n. \quad (26)$$

As we constructed $\tilde{\mathbf{D}}$ from $\mathbf{D}$, we construct $\tilde{\mathbf{F}}$ and $\tilde{\boldsymbol{\Delta}}$ by doubling the sizes of $\mathbf{F}$ and $\boldsymbol{\Delta}$. Then, we have

$$\begin{aligned} \tilde{\mathbf{v}}_{n+1} &= \boldsymbol{\Omega}(\mathbf{I}_{8L-6} - \mu_0 \tilde{\mathbf{F}} \tilde{\boldsymbol{\Delta}} \tilde{\mathbf{F}}^H) \tilde{\mathbf{v}}_n \\ \tilde{\mathbf{v}}_{n+1} &= \boldsymbol{\Omega} \tilde{\mathbf{F}} (\mathbf{I}_{8L-6} - \mu_0 \tilde{\boldsymbol{\Delta}}) \tilde{\mathbf{F}}^H \boldsymbol{\Omega} \tilde{\mathbf{v}}_n \\ \tilde{\mathbf{F}}_0^H \tilde{\mathbf{v}}_{n+1} &= (\boldsymbol{\Omega} - \mu_0 \tilde{\boldsymbol{\Delta}}_0) \tilde{\mathbf{F}}_0^H \tilde{\mathbf{v}}_n \end{aligned}$$

with $\tilde{\mathbf{F}}_0 = \boldsymbol{\Omega} \tilde{\mathbf{F}}$ and $\tilde{\boldsymbol{\Delta}}_0 = \boldsymbol{\Omega} \tilde{\boldsymbol{\Delta}}$. The local stability of the rotated error vector $\mathbf{v}_n^r = \tilde{\mathbf{F}}_0^H \tilde{\mathbf{v}}_n$ is governed by the

DFT of $\mathbf{c}_1$. We denote the DFT of $\mathbf{c}_1$ for the frequency bins $k = 0, \ldots, 4L-4$ by

$$W_{eq}[k] = \sum_{n=-2L+2}^{2L-2} w_{\text{eq}}(n) e^{2i\pi kn/(4L-3)}. \quad (27)$$

A *sufficient* local stability condition for double-sided filters with periods of silence is that for all $k \in \{0, \ldots, 4L-4\}$, $\Re\{W_{eq}[k]\} > 0$. Since $w_{\text{eq}} = \delta - h_{12} * h_{21}$, we have $W_{eq}[k] = 1 - H_{12}[k]H_{21}[k]$, where $H_{12}[k]$ and $H_{21}[k]$ are defined similarly to (27). Therefore, the real part of $W_{eq}[k]$ is positive if $|H_{12}[k]H_{21}[k]|$ is strictly smaller than one. This condition is satisfied if

$$\sum_{n \in \mathcal{S}_d} |h_{ij}(n)| < 1 \text{ for } i, j = 1, 2\, i \neq j. \quad (28)$$

4.2. Causal systems

Linearizing (15) for causal systems gives a similar form to (20)

$$\mathbf{v}_{n+1} = \left(\mathbf{I}_{2L} - \mu \frac{1}{\|w_{eq}\|_2^2} \underbrace{\begin{pmatrix} \Theta\mathbf{B} & \mathbf{A} \\ \mathbf{A} & \Theta'\mathbf{B} \end{pmatrix}}_{\mathbf{D}_{\text{causal}}} \right) \mathbf{v}_n \quad (29)$$

with

$$\mathbf{A} = \begin{pmatrix} w_{\text{eq}}(0) & \cdots & w_{\text{eq}}(L-1) \\ \vdots & \ddots & \vdots \\ w_{\text{eq}}(L-1) & \cdots & w_{\text{eq}}(2L-2) \end{pmatrix},$$

$$\mathbf{B} = \begin{pmatrix} w_{\text{eq}}(0) & \cdots & w_{\text{eq}}(L-2) & w_{\text{eq}}(L-1) \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & w_{\text{eq}}(1) \\ 0 & \cdots & 0 & w_{\text{eq}}(0) \end{pmatrix},$$

$$\mathbf{v}_n = (\mathbf{v}_{12,n}^T, \mathbf{v}_{21,n}^T),$$

where $\mathbf{v}_{ij,n} = (v_{ij,n}(0), \ldots, v_{ij,n}(L+1))^T$. Like in the double-sided case, $\mathbf{D}_{\text{causal}}$ is not positive definite in general. The local stability is not guaranteed.

For locally silent sources, the extraction of the eigenvalues of $\mathbf{D}_{\text{causal}}$ is possible. Letting σ^2 tend to zero as in the previous section, we neglect the off-diagonal blocks in $\mathbf{D}_{\text{causal}}$. Then, $\mathbf{D}_{\text{causal}}$ becomes an upper-triangular matrix. This reveals that, in the vicinity of the equilibrium $v_{ij} = 0$, $v_{ij}(k)$ depends at the first order only on $v_{ij}(l), l \geq k$. The eigenvalues of $\mathbf{D}_{\text{causal}}$ are the diagonal terms $w_{\text{eq}}(0) = 1 - h_{12}(0)h_{21}(0)$. Then, the local stability condition is

$$h_{12}(0)h_{21}(0) < 1. \quad (30)$$

This is similar to the asymptotic stability conditions derived in [6, 7].

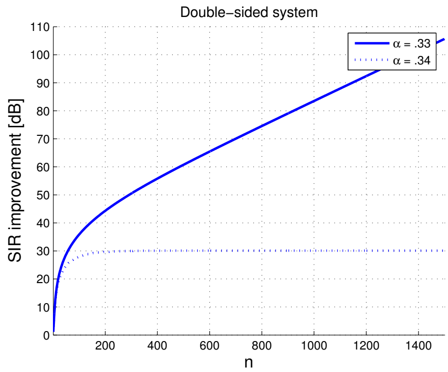

Fig. 3. SIR improvement for a stable ($\alpha = 0.33$) and an unstable ($\alpha = 0.34$) double-sided system.

5. NUMERICAL EXPERIMENTS

To illustrate the effect of the local stability on the global convergence, we carry out a numerical experiment. For this, the signal statistics are *not* estimated on finite length signals but computed *exactly* using known source statistics and mixing channels. The local stability is a property that concerns the vicinity of the separation point $-\mathbf{H}(z)$: if an algorithm is unstable, it will not be able to reach that point. However it may still provide signal-to-interference ratio (SIR) improvement to a certain extent. We consider the following double-sided mixing:

$$H_{12}^{(\alpha)}(z) = H_{21}^{(\alpha)}(z) = \alpha z + .5 + \alpha z^{-1}. \quad (31)$$

The source variance profile is

$$(\sigma_{1,1}^2, \sigma_{1,2}^2, \sigma_{2,1}^2, \sigma_{2,2}^2) = (1, .5, .5, 1).$$

For $\alpha = 0.33$, the eigenvalues of the matrix $\mathbf{D}$ in (20) have positive real-parts, whereas it is not the case for $\alpha = 0.34$. As we can see in Fig. 3, this does not lead to divergence but rather to an early saturation.

6. DISCUSSION

It is difficult to draw any conclusion on the separation performance of BSS from local stability conditions. In practice, the separation performance is mainly affected by the global convergence. Nevertheless, the influence of the causality on experimental results has been pointed out [3, 9]. It is important to keep in mind the physical interpretation of the stability conditions (28) and (30). Condition (30) is satisfied if one of the mixing filter is strictly causal, i.e. if $h_{12}(0)h_{21}(0) = 0$. This is the case when the sources are placed apart from the microphones median plane, as depicted in Fig. 4 (left). A scenario involving acausal mixing

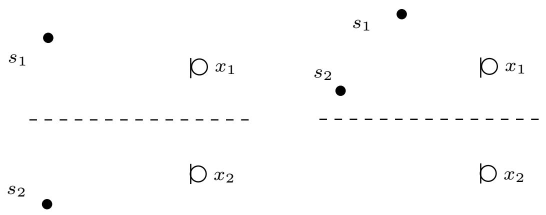

Fig. 4. Left: The sources are apart from the microphones median plan and causal separation filters are sufficient. – Right: The sources are in the same half-plane and acausal separation filters are necessary.

channels is depicted in Fig. 4 (right). In this situation, condition (28) sets an upper bound on the amount of crosstalk and is clearly more conservative than (30). Therefore, one can expect BSS to be less robust when acausal separation filters are involved [3]. Note that $H_{12}(z)$ and $H_{21}(z)$ represent ratio of acoustic transfer functions. Therefore, acausal separation filters may also be useful in the scenario depicted in Fig. 4 (left), depending on $H_{ii}(z), i = 1, 2$.

It is often proposed to apply BSS algorithms in the frequency domain, particularly in the context of acoustical mixtures. Thereby, the mixing (2) is processed at each frequency $z = e^{i\omega}$ independently. Such a formulation of the mixing system hides its causality. Nevertheless, the performance of frequency-domain BSS algorithm is also affected by the causality of the mixing/separation system [9].

7. CONCLUSION

We examined the local stability of a second-order BSS algorithm. Because of the existence of periods of silence, e.g. with speech sources, BSS is compared to a set of two interference cancellers. Taking these silences explicitly into account also simplifies the local stability analysis and shows a discrepancy between causal and double-sided mixing channels. In addition to the well-known permutation ambiguity, double-sided systems suffer from more restrictive local stability conditions than causal ones. These differences causal/double-sided do not exist for the NLMS adaptation of the interference cancellers, whose stability depends solely on the step-size (for a perfect activity detection). This analysis is restricted to white sources and its value is rather illustrative. If $H_{ii} \neq 1$, the sensors receive colored signals.

Acknowledgments. The author acknowledges the financial support provided through the European Community's Human Potential Programme under contract HPRN-CT-2002-00276 HOARSE. We are also grateful to R. Aichner, H. Buchner and Prof. Dr.-Ing. W. Kellermann for their helpful comments.

8. REFERENCES

[1] S. Araki, S. Makino, R. Mukai, Y. Hinamoto, T. Nishikawa, and H. Saruwatari. Equivalence between frequency domain blind source separation and frequency domain adaptive beamforming. *Proceedings of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Student Forum, 2002.

[2] R. Aichner, H. Buchner, S. Araki, and S. Makino. On-line Time-Domain Blind Source Separation of Nonstationary Convolved Signals. *Proc. Int. Symposium on Independent Component Analysis and Blind Signal Separation (ICA)*, pp. 987–992, Nara, Japan, 2003.

[3] R. Aichner, H. Buchner and W. Kellermann. On the causality problem in time-domain blind source separation and deconvolution algorithms. *Proceedings IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Philadelphia, PA, USA, March 2005.

[4] H. Buchner, R. Aichner, W. Kellermann. A generalization of blind source separation algorithms for convolutive mixtures based on second order statistics. *IEEE Trans. on Speech and Audio Processing*, Vol. 13, Num. 1, pp. 120–134, 2005.

[5] J.F. Cardoso. On the stability of source separation algorithms. *Journal of VLSI Signal Processing Systems*, pp. 7–14, Vol. 26, no 1/2, 2000.

[6] N. Charkani and Y. Deville. Self-adaptive separation of convolutively mixed signals with a recursive structure. Part I: Stability analysis and optimization of asymptotic behaviour. *Signal Processing*, no. 73, pp. 225–254, 1999.

[7] S. Cruces, L. Castedo. Stability analysis of adaptative algorithms for blind source separation of convolutive mixtures. *Signal Processing*, no. 78, pp. 265–275, 1999.

[8] S.C. Douglas and S. Amari. Natural Gradient Adaptation. *in Unsupervised Adaptive Filtering*, S. Haykin, ed. Wiley, New-York, 1999.

[9] E. Robledo-Arnuncio and B.H. Juang. Issues in frequency-domain blind source separation - a critical revisit. *Proceedings IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Philadelphia, PA, USA, March 2005.

[10] B. Widrow, S. D. Stearn. Adaptive Signal Processing. Prentice-Hall Signal Processing Series, Englewood Cliffs, 1985.

[11] L. Zhang, A. Cichocki and S. Amari. Semiparametric approach for blind deconvolution of nonminimum phase systems, *NIPS'99, Advance in Neural Information Processing*, pp. 363–369. S.A. Solla, T.K. Leen and K.-R. Muller, eds, MIT Press, 2000.