NATURAL GRADIENT MULTICHANNEL BLIND DECONVOLUTION AND SOURCE SEPARATION USING CAUSAL FIR FILTERS

Scott C. Douglas

Department of Electrical Engineering
Southern Methodist University
Dallas, Texas 75275 USA

Hiroshi Sawada and Shoji Makino

NTT Communication Science Laboratories
NTT Corporation
Kyoto 619-0327 Japan

ABSTRACT

Practical gradient-based adaptive algorithms for multichannel blind deconvolution and convolutive blind source separation typically employ FIR filters for the separation system. Inadequate use of signal truncation within these algorithms can introduce steady-state biases into their converged solutions that lead to degraded separation and deconvolution performances. In this paper, we derive a natural gradient multichannel blind deconvolution and source separation algorithm that mitigates these effects for estimating causal FIR solutions to these tasks. Numerical experiments verify the robust convergence performance of the new method both in multichannel blind deconvolution tasks for i.i.d. sources and in convolutive BSS tasks for acoustic sources, even for extremely-short separation filters.

1. INTRODUCTION

Blind source separation (BSS) is a field that has received much recent attention in several research fields, including acoustics, bioinformatics, communications, control, data mining, and signal processing. Most formulations to the BSS task assume that a set of sensor signals contain linear mixtures of several source signals of interest. The goal is to process the sensor signals to extract versions of each of the source signals without any crosstalk and without precise knowledge of the source signals, the mixing conditions, or any training information. This paper focuses on the convolutive BSS task, in which a vector sequence $\mathbf{s}(k) = [s_1(k) \; \cdots \; s_m(k)]^T$ of m sources $\{s_i(k)\}$ is mixed by a causal time-dispersive multichannel system as

$$\mathbf{x}(k) = \sum_{l=0}^{\infty} \mathbf{A}_l \mathbf{s}(k - l), \tag{1}$$

where $\mathbf{x}(k) = [x_1(k) \; \cdots \; x_m(k)]^T$ contains the m sensor signals and $\mathbf{A}_l$ is the $(m \times m)$ coefficient mixing matrix at lag l with elements $\{a_{ijl}\}$, $1 \leq \{i, j\} \leq m$. In practice, the separation system has the multichannel finite-impulse-response (FIR) form

$$\mathbf{y}(k) = \sum_{l=0}^{L} \mathbf{W}_l(k) \mathbf{x}(k - l), \tag{2}$$

where $\mathbf{W}_l(k)$, $0 \leq l \leq L$ are the $(L + 1)$ matrix coefficients of size $(m \times m)$ for the separation system at time k. The separation system is made time-varying with the assumption that the $\{\mathbf{W}_l(k)\}$ can be iteratively adapted to achieve the separation task.

In [1, 2], natural gradient multichannel blind deconvolution and source separation procedures are derived that have a number of useful features: (1) They are based on a sound minimum mutual information criterion [3]; (2) They only assume that the sources have non-Gaussian amplitude statistics; (3) They have a simple FIR-based multiply/add computational structure that is amenable to a sample-by-sample or a block frequency-domain implementation; and (4) They appear to provide some separating capability for real-world mixtures such as acoustic recordings of speech [2, 4]. The time-domain coefficient updates are [1]

$$\mathbf{W}_l(k+1) = \mathbf{W}_l(k) + \mu[\mathbf{W}_l(k) - \mathbf{f}(\mathbf{y}(k-L))\mathbf{u}^T(k-l)] \tag{3}$$

$$\mathbf{u}(k) = \sum_{q=0}^{L} \mathbf{W}_{L-q}^T(k)\mathbf{y}(k - q), \tag{4}$$

$\mathbf{f}(\mathbf{y}) = [f_1(y_1) \; \cdots \; f_m(y_m)]^T$, and $f_i(y)$ is a nonlinear function that is related to the amplitude statistics of the source signal $s_j(k)$ extracted at the ith output. Several researchers have used (2)–(4) as the basis for modified approaches that overcome some of the limitations of the original procedures for acoustic source separation, such as the spectral flattening of the extracted sources [5]–[9].

We have identified a performance limitation of the algorithm in (2)–(4), as demonstrated in Figs. 1 and 2. Complete details of this example are provided in Section 4. Fig. 1 shows the impulse responses $\{a_{ijp}\}$ for a two-loudspeaker, two-microphone laboratory measurement setup at a sampling rate of 8kHz. Fig. 2 shows the impulse responses $\{c_{ijp}(k)\}$ of the combined system

$$\mathbf{C}_p(k) = \sum_{l=0}^{\infty} \mathbf{W}_{p-l}(k)\mathbf{A}_l \tag{5}$$

obtained after applying (2)–(4) to i.i.d. binary-$\{\pm 1\}$-distributed sources that are mixed by an acoustic channel. Notice that the edges of the impulse responses of the various sub-filters within the combined system exhibit "spikes" at either end of their active temporal windows. These spikes create pre- and post-echoes that harm the overall separation performance of the scheme. Although illustrated using binary i.i.d. sources, these artifacts appear with most signal sets after an extended number of algorithm iterations, and they cause a significant decrease in overall separation performance. Several researchers have developed extensions of these algorithms in an attempt to mitigate these effects, which generally increases the complexity of the approach [5, 6, 8]. Clearly, a structural change to the algorithm is required to improve its steady-state separation and deconvolution performance for a wider class of mixing systems, including those that are non-minimum phase or that have non-finite-length inverses.

In this paper, we study the effect that signal windowing and filter truncation play in the natural gradient method for blind de-
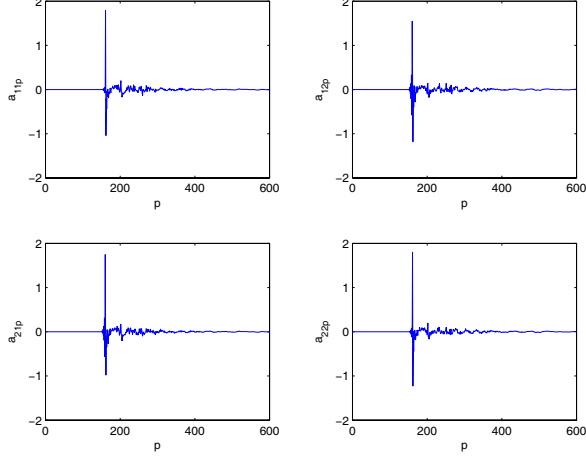
Fig. 1: Impulse responses for the acoustic channel.

Fig. 2: Combined system responses for the algorithm in (2)-(4).

convolution and source separation. We show that windowing approximations used in the derivation of (2)–(4) result in additional input signal terms within the natural gradient updates that are not part of the standard gradient. Simulation evidence suggests that these terms cause a bias in the steady-state solution obtained by the scheme. We then introduce a modified natural gradient method for multichannel blind deconvolution that does not include these additional terms. Simulations show that the proposed algorithm performs multichannel blind deconvolution on i.i.d. sources without pre- or post-echo in the combined impulse responses. Application to acoustical mixtures shows that much smaller filter lengths can often be chosen for the new method as compared to the original algorithm without sacrificing performance.

2. SIGNAL WINDOWING IN EXISTING NATURAL GRADIENT ALGORITHMS

In this section, we compare the structure of the coefficient updates in (2)–(4) with the gradient of the cost function on which this procedure is based. This cost function is

$$
\mathcal{J}(\mathcal{W}_k(z)) \\
= -\sum_{i=1}^{m} E\{\log \widehat{p}_i(y_i(k))\} - \frac{1}{2\pi j} \oint \log|\det \mathcal{W}_k(z)|z^{-1}dz \quad (6)
$$

where $\mathcal{W}_k(z) = \sum_{l=0}^{L} \mathbf{W}_l(k)z^{-l}$ is the z-transform of $\mathbf{W}_l(k)$, $E\{\cdot\}$ denotes statistical expectation, and $\widehat{p}_i(y)$ is a model of the p.d.f. of the ith source to be extracted. It can be shown that (6) is, up to a constant independent of $\mathcal{W}_k(z)$, proportional to the mutual information of the output signal sequences $\{y_i(k)\}$ when $\widehat{p}_i(y)$ is the p.d.f. of the ith extracted source sequence [3]. Minimizing this measure results in a set of sequences $y_i(k)$ that are most independent from sample to sample and from channel to channel. When $\mathbf{x}(k)$ fits the model in (1) and each $s_i(k)$ is independent of every $s_j(l)$ for $i \neq j$ and any k and l, minimizing (6) results in the multichannel deconvolution of the signal mixture.

Standard gradient minimization of (6) takes on the form

$$
\mathbf{W}_l(k+1) = \mathbf{W}_l(k) - \mu \frac{\partial \mathcal{J}(\mathcal{W}_k(z))}{\partial \mathbf{W}_l(k)}, \quad (7)
$$

where μ is the algorithm step size. For the moment, we shall focus on the first term within the cost function in (6) whose gradient with
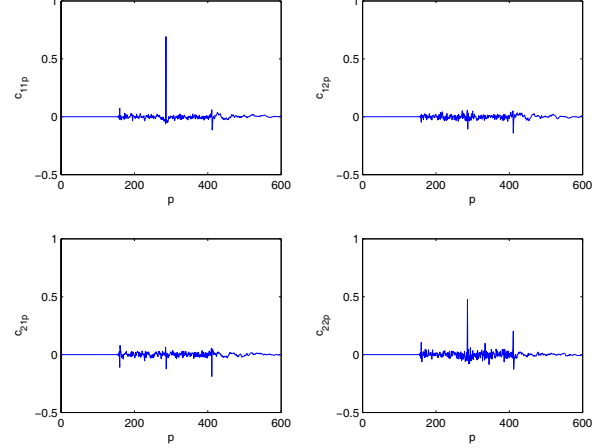
respect $\mathbf{W}_l(k)$ can be shown to be

$$
\frac{\partial}{\partial \mathbf{W}_l(k)}\left(-\sum_{i=1}^{m} E\{\log \widehat{p}_i(y_i(k))\}\right) = E\{\mathbf{f}(\mathbf{y}(k))\mathbf{x}^T(k-l)\} \quad (8)
$$

where $f_i(y) = -\partial \log \widehat{p}_i(y)/\partial y$ in $\mathbf{f}(\mathbf{y})$. Thus, the coefficient updates for the standard gradient method in (7) depend only on the most-recent output signal vector $\mathbf{y}(k)$ and the $(L+1)$ most-recent input signal vectors $\{\mathbf{x}(k-l)\}$, $0 \leq l \leq L$. If an L-sample delayed coefficient update is used instead, the coefficient updates depend on signal terms of the form

$$
\frac{\partial}{\partial \mathbf{W}_l(k-L)}\left(-\sum_{i=1}^{m} E\{\log \widehat{p}_i(y_i(k-L))\}\right) \\
= E\{\mathbf{f}(\mathbf{y}(k-L))\mathbf{x}^T(k-L-l)\}, \quad (9)
$$

in which only $\{\mathbf{x}(k-l)\}$, $L \leq l \leq 2L$ appear in the updates.

The natural gradient procedure in (2)–(4) is a filtered-gradient stochastic update rule in which an L-sample delay is introduced within the signal-dependent update terms of (8) to maintain causality. Substituting the expressions for $\mathbf{y}(k)$ in (2) and $\mathbf{u}(k)$ in (4) into the right-hand side of (3), we obtain

$$
\mathbf{W}_l(k+1) = (1+\mu)\mathbf{W}_l(k) \\
- \mu \mathbf{f}(\mathbf{y}(k-L))\sum_{p=0}^{L}\sum_{q=0}^{L}\{\mathbf{W}_{L-q}^T(k-l) \\
\times \mathbf{W}_p(k-q-l)\mathbf{x}(k-p-q-l)\}. \quad (10)
$$

The second update term on the right-hand side of (10) includes input signal samples $\{x_j(k-l)\}$ for $0 \leq l \leq 3L$. These terms fall outside the range of lags $L \leq l \leq 2L$ in the delayed gradient calculation in (9). Hence, the update in (3) employs input signal samples that are *not part of the gradient of the cost function for the chosen FIR filter structure.*

How these additional terms affect the overall convergence performance of the system depends strongly on the source and mixture characteristics. In the unlikely situation where a finite-length equalizer is adequate to both separate and deconvolve the source mixtures (*e.g.*, the mixing system is multichannel autoregressive of order $P \leq L$) and the sources are temporally-independent, the additional terms within the coefficient updates are likely to improve

V - 478

convergence performance, because they provide additional signal-dependent constraints that fit the truncated separation model, such that $E\{\mathbf{f}(\mathbf{y}(k-L))\mathbf{x}^T(k-l)\} = \mathbf{0}$ outside the range $L \le l \le 2L$. In more-common situations where a finite-length equalizer cannot separate and deconvolve the sources, or in situations where the sources are temporally-dependent such as in speech separation, the additional terms are likely to harm convergence performance because they are not consistent with the truncated separation model, *i.e.* $E\{\mathbf{f}(\mathbf{y}(k-L))\mathbf{x}^T(k-l)\} \ne \mathbf{0}$ outside the range $L \le l \le 2L$.

3. CAUSAL NATURAL GRADIENT ALGORITHM FOR MULTICHANNEL SYSTEMS

In this section, we develop a procedure that largely mitigates the detrimental coefficient bias observed in Fig. 2. The modification is straightforward: Remove the terms in (3) that do not depend on $\{\mathbf{x}(k-l)\}, L \le l \le 2L$. If we also assume that adaptation is slow such that $\mathbf{W}_p(k-q-l)$ and $\mathbf{W}_{L-q}^T(k-l)$ can be replaced by $\mathbf{W}_p(k-L)$ and $\mathbf{W}_{L-q}^T(k-L)$, respectively, we obtain

$$\mathbf{W}_l(k+1) = (1+\mu)\mathbf{W}_l(k) - \mu\mathbf{f}(\mathbf{y}(k-L))$$
$$\times \sum_{p=0}^{L-l} \sum_{q=\max\{L-l-p,0\}}^{\min\{L,2L-l-p\}} \{\mathbf{W}_{L-q}^T(k-L)$$
$$\times \mathbf{W}_p(k-L)\mathbf{x}(k-p-q-l)\}. \quad (11)$$

To see the inherent structure of this new algorithm, define

$$\underline{\mathbf{x}}(k) = [\mathbf{x}^T(k) \ \mathbf{x}^T(k-1) \ \cdots \ \mathbf{x}^T(k-L)]^T \quad (12)$$
$$\underline{\mathbf{W}}(k) = [\mathbf{W}_0(k) \ \mathbf{W}_1(k) \ \cdots \ \mathbf{W}_L(k)]. \quad (13)$$

Then, we can represent

$$\mathbf{y}(k) = \sum_{l=0}^{L} \mathbf{W}_l(k)\mathbf{x}(k-l) = \underline{\mathbf{W}}(k)\underline{\mathbf{x}}(k). \quad (14)$$

Define the m-element vector $\mathbf{z}_l(k)$ and $(m \times m)$ matrix $\mathbf{R}_p(k)$ as

$$\mathbf{z}_l(k) = \sum_{j=0}^{L} \mathbf{R}_{j-l}(k)\mathbf{x}(k-j) \quad (15)$$

$$\mathbf{R}_l(k) = \begin{cases} \sum_{p=0}^{L-l} \mathbf{W}_p^T(k)\mathbf{W}_{p+l}(k) & \text{if } 0 \le l \le L \\ \mathbf{R}_{-l}^T(k) & \text{if } -L \le l \le -1 \end{cases} \quad (16)$$

Define the $(m(L+1) \times m(L+1))$ matrix $\underline{\mathbf{R}}(k)$ and the $m(L+1)$-element vector $\underline{\mathbf{z}}(k)$ as

$$\underline{\mathbf{R}}(k) = \begin{bmatrix} \mathbf{R}_0(k) & \mathbf{R}_1(k) & \cdots & \mathbf{R}_L(k) \\ \mathbf{R}_{-1}(k) & \mathbf{R}_0(k) & & \mathbf{R}_{L-1}(k) \\ \vdots & & & \vdots \\ \mathbf{R}_{-L}(k) & \cdots & \cdots & \mathbf{R}_0(k) \end{bmatrix} \quad (17)$$

$$\underline{\mathbf{z}}(k) = [\mathbf{z}_0^T(k) \ \cdots \ \mathbf{z}_L^T(k)]^T = \underline{\mathbf{R}}(k)\underline{\mathbf{x}}(k), \quad (18)$$

respectively. Then, it can be shown that (11) is equivalent to

$$\underline{\mathbf{W}}(k+1) = \underline{\mathbf{W}}(k) + \mu[\underline{\mathbf{W}}(k) - \mathbf{f}(\mathbf{y}(k-L))\underline{\mathbf{z}}^T(k-L)], (19)$$

where $\mathbf{f}(\mathbf{y}(k)) = [f_1(y_1(k)) \ \cdots \ f_m(y_m(k))]^T$. Moreover, since $\{\mathbf{x}(k-l)\}, 0 \le l < L$ no longer appear in the updates, we can remove the signal delay introduced into (19) to obtain

$$\underline{\mathbf{W}}(k+1) = \underline{\mathbf{W}}(k) + \mu[\underline{\mathbf{W}}(k) - \mathbf{f}(\mathbf{y}(k))\underline{\mathbf{z}}^T(k)]. \quad (20)$$
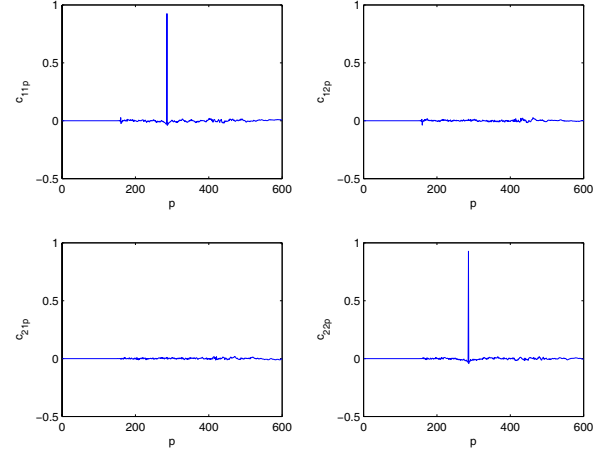

Fig. 3: Combined system response for the proposed scheme.

The calculation of $\underline{\mathbf{z}}(k)$ in (18) involves $m^2(L+1)^2$ multiply/adds, making it prohibitive for large filter lengths. As such, we propose a simplified approach to approximating the value of $\underline{\mathbf{z}}(k)$. This approach make use of the block Toeplitz structure of $\underline{\mathbf{R}}(k)$ to recursively update the vector

$$\underline{\widehat{\mathbf{t}}}(k) = [\widehat{\mathbf{t}}_0^T(k) \ \widehat{\mathbf{t}}_1^T(k) \ \cdots \ \widehat{\mathbf{t}}_{2L-1}^T(k) \ \widehat{\mathbf{t}}_{2L}^T(k)]^T \quad (21)$$

whose last $m(L+1)$ entries contained in the vector $\underline{\widehat{\mathbf{t}}}(k)$ are equal to $\underline{\mathbf{z}}(k)$ if $\underline{\mathbf{R}}(k)$ does not change. The recursive update for the vector elements of $\underline{\widehat{\mathbf{t}}}(k)$ is given by

$$\widehat{\mathbf{z}}_p(k) = \widehat{\mathbf{t}}_{L+p}(k) \quad (22)$$

$$\widehat{\mathbf{t}}_p(k) = \begin{cases} \mathbf{R}_L^T(k)\mathbf{x}(k) & \text{if } p = 0 \\ \widehat{\mathbf{t}}_{p-1}(k-1) + \mathbf{R}_{L-p}^T(k)\mathbf{x}(k) & \text{if } 1 \le p \le L \\ \widehat{\mathbf{t}}_{p-1}(k-1) + \mathbf{R}_{p-L}(k)\mathbf{x}(k) \\ \quad -\mathbf{R}_{2L+1-p}^T(k-L-1)\mathbf{x}(k-L-1) & \text{if } L+1 \le p \le 2L \end{cases} \quad (23)$$

Each $\{\mathbf{R}_l(k)\}, 0 \le l \le L$ is only updated at the time instants $k = n(L+1)$, where n is an integer. The coefficient updates become

$$\underline{\mathbf{W}}(k+1) = \underline{\mathbf{W}}(k) + \mu[\underline{\mathbf{W}}(k) - \mathbf{f}(\mathbf{y}(k))\widehat{\underline{\mathbf{z}}}^T(k)]. \quad (24)$$

Equations (14) and (22)—(24) define the proposed multichannel blind deconvolution and source separation procedure employing causal filters. The complexity of the algorithm is approximately 6.5 multiply/adds per adaptive filter coefficient per time instant including the periodic calculation of $\{\mathbf{R}_l(k)\}$, making it approximately 63% more complex than (2)–(4).

4. NUMERICAL SIMULATIONS

We now explore the behavior of the proposed procedure via numerical simulations. All of the examples in this section use the same two-input, two-output impulse response in Fig. 1. This impulse response was generated from an acoustic laboratory setup consisting of a pair of omnidirectional lapel microphones spaced 4cm apart and mounted in a V-configuration approximately 1.5m from the floor in the center of a 4.45m-by-3.55m-by-2.50m room. The reverberation time of this room is 130ms. A pair of loudspeakers located 1.2m away from the microphones at -40 degrees and $+30$ degrees from the on-axis direction were used as the acoustic sources. Bandlimited white noise played through these loudspeakers was then used to characterize the individual impulse responses
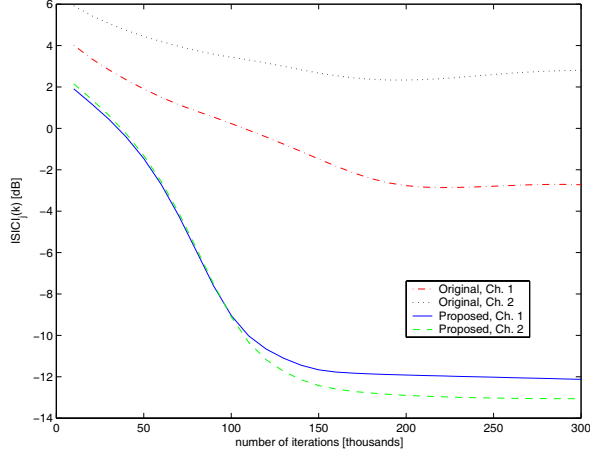
Fig. 4: Per-channel ISICIs for the two algorithms.

of the loudspeaker-to-microphone acoustic paths using linear estimation techniques. We then generated the source signal mixtures by filtering recorded signals using these impulse responses digitally. While not exactly identical to real-world acoustic mixtures, these signals allow us to accurately observe and characterize the impulse response distortions due to FIR filter and signal truncation through the combined system impulse responses. These distortions are generally *unobservable* when looking at the separation system coefficients alone. All signals are sampled at 8kHz.

Our first experiment tests the ability of the algorithms to both separate and deconvolve binary–$\{\pm 1\}$ i.i.d. source mixtures. Both the original natural gradient procedure in (3) and the proposed procedure in (24) were applied to this data, where $L = 250$, $\mathbf{W}_l(0) = 0.5\mathbf{I}\delta_{l-125}$, $f_i(y) = y^3$, and $\mu = 0.0001$. Shown in Fig. 2 is the combined impulse responses $\mathbf{C}_p(k)$ delayed by 175 samples at iteration $k = 300000$ for the original procedure. As noted in the introduction, the "spikes" present near $p = 175$ and $p = 425$ are troublesome and prevent equalization. Shown in Fig. 3 are the combined impulse responses for the proposed method at iteration $k = 300000$ on the same data. As can be seen, the separation and deconvolution performance is nearly perfect, with ideal delta-function responses in $\{c_{11p}\}$ and $\{c_{22p}\}$ and nearly-zero responses in $\{c_{12p}\}$ and $\{c_{21p}\}$, respectively. Plots of the separation system impulse responses $\mathbf{W}_l(k)$ for the two schemes show a slight difference in the coefficient values, although it is not clear which impulse response set yields the better result.

Shown in Fig. 4 are the per-channel combined inter-symbol and inter-channel interferences (ISICIs), computed as

$$ISICI_j(k) = \left(\sum_{l=0}^{2L} \{c_{j1l}^2(k) + c_{j2l}^2(k)\} \Big/ \max_{0 \le j \le 2L} c_{jjl}^2(k) \right) - 1 \quad (25)$$

for $j \in \{1, 2\}$ respectively, during their respective convergence periods. As can be seen, the original natural gradient approach fails to accurately deconvolve the acoustic channel, whereas the proposed method is quite effective at reducing the ISICI.

We now turn to examples involving acoustic sources. In this case, we replace the two random uniform sources with two 7-second isolated recordings of a single male speaker from a radio newscast. These signals were repeated six times before being filtered by the acoustic channel to create a 300000-sample pair of signal mixtures. Both the original and proposed algorithms were applied to these signals, where $\mathbf{W}_l(0) = 10\mathbf{I}\delta_{l-L/2}$ and

Table 1: Signal-to-Interference Ratios for Speech Separation

Channel	Original		Proposed		
	$L = 100$	$L = 50$	$L = 100$	$L = 50$	$L = 24$
1	0.7 dB	2.2 dB	8.6dB	13.9dB	13.3dB
2	-1.0 dB	3.1 dB	11.0dB	13.5dB	12.8dB

$f_i(y) = \text{sgn}(y)$. The initial step size values were chosen to be $\mu = 0.00005$ and $\mu = 0.00015$ for the original and proposed methods, respectively, and these values were lowered to 0.000025 and 0.000075, respectively, at iteration $k = 150000$.

Table 1 lists the signal-to-interference ratios (SIRs) obtained by the original and proposed methods for various filter lengths L. The data indicate that the original natural gradient algorithm fails to provide any degree of separation with these parameter choices, and performance is marginally better for shorter filters. Large "spikes" can be seen in the $\{c_{ijp}\}$ sequences (not shown). In contrast, the proposed method provides reasonable performances for filter lengths as small as $L = 24$ taps. Such short-filter systems can only be used when the room reverberation time is short and the microphone array has closely-spaced sensors. These results are extremely promising for practical acoustic source separation in offices and other small-room environments.

5. CONCLUSIONS

In this paper, we have demonstrated through argument and simulation that an existing natural gradient multichannel blind deconvolution and source separation procedure [1] can achieve a biased result when the channel to be inverted is not minimum phase and/or the separation filter length is too short. We then propose a new natural gradient procedure that avoids these difficulties. The complexity of the new algorithm, while somewhat greater than the original approach, is still proportional to the number of parameters in the separation system, and it uses only multiplies and adds in its operation. We have demonstrated its capabilities both in multichannel blind deconvolution tasks involving synthetic signals as well as convolutive BSS tasks involving speech signals. Moreover, the algorithm functions in a reasonable manner even when the filter lengths chosen are much shorter than would be required for an accurate channel inverse.

6. REFERENCES

[1] S. Amari, S.C. Douglas, A. Cichocki, and H.H. Yang, "Multichannel blind deconvolution and equalization using the natural gradient," *Proc. 1st IEEE Workshop Signal Process. Adv. Wireless Comm.,* Paris, France, pp. 101–104, Apr. 1997.

[2] R.H. Lambert and A.J. Bell, "Blind separation of multiple speakers in a multipath environment," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing,* Munich, Germany, vol. 1, pp. 423-426, Apr. 1997.

[3] D.-T. Pham, "Mutual information approach to blind separation of stationary sources," *Proc. 1st Workshop Indep. Compon. Anal. Signal Separation,* Aussois, France, pp. 215-220, Jan. 1999.

[4] S. Amari, S.C. Douglas, A. Cichocki, and H.H. Yang, "Novel on-line adaptive learning algorithms for blind deconvolution using the natural gradient approach," *Proc. 11th IFAC Symp. Syst. Ident.,* Kitakyushu City, Japan, vol. 3, pp. 1057-1062, July 1997.

[5] T.-W. Lee, A. Ziehe, R. Orglmeister and T. J. Sejnowski, "Combining time-delayed decorrelation and ICA: towards solving the cocktail party problem," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing,* Seattle, WA., pp. 1089–1092, May 1998.

[6] K. Matsuoka and S. Nakashima, "Minimal distortion principle for blind source separation," *Proc. 3rd Int. Workshop Indep. Compon. Anal. Signal Separation,* San Diego, CA, pp. 722–727, Dec. 2001.

[7] S.C. Douglas, "Blind separation of acoustic signals," in *Microphone Arrays: Techniques and Applications,* M. Brandstein and D. Ward, eds. (New York: Springer-Verlag, 2001), pp. 355-380.

[8] S.C. Douglas and X. Sun, "Convolutive blind separation of speech mixtures using the natural gradient," *Speech Communication,* vol. 39, pp. 65-78, Dec. 2002.

[9] T. Takatani, T. Nishikawa, H. Saruwatari, and K. Shikano, "SIMO-model-based independent component analysis for high-fidelity blind separation of acoustic signals," *Proc. 4th Int. Symp. Indep. Compon. Anal. Signal Separation,* Nara, Japan, pp. 993–998, Apr. 2003.