

Article

Accelerated Conjugate Gradient for Second-Order Blind Signal Separation

Hai Huyen Dam ^{*,†} and Sven Nordholm ^{*,†} 

School of Electrical Engineering, Computing and Mathematical Sciences (EECMS), Curtin University, Perth, WA 6845, Australia

* Correspondence: h.dam@curtin.edu.au (H.H.D.); s.nordholm@curtin.edu.au (S.N.)

† These authors contributed equally to this work.

Abstract: This paper proposes a new adaptive algorithm for the second-order blind signal separation (BSS) problem with convolutive mixtures by utilising a combination of an accelerated gradient and a conjugate gradient method. For each iteration of the adaptive algorithm, the search point and the search direction are obtained based on the current and the previous iterations. The algorithm efficiently calculates the step size for the accelerated conjugate gradient algorithm in each iteration. Simulation results show that the proposed accelerated conjugate gradient algorithm with optimal step size converges faster than the accelerated descent algorithm and the steepest descent algorithm with optimal step size while having lower computational complexity. In particular, the number of iterations required for convergence of the accelerated conjugate gradient algorithm is significantly lower than the accelerated descent algorithm and the steepest descent algorithm. In addition, the proposed system achieves improvement in terms of the signal to interference ratio and signal to noise ratio for the dominant speech outputs.

Keywords: accelerated conjugate gradient; second-order blind signal separation; optimal step size



Citation: Dam, H.H.; Nordholm, S. Accelerated Conjugate Gradient for Second-Order Blind Signal Separation. *Acoustics* **2022**, *4*, 948–957. <https://doi.org/10.3390/acoustics4040058>

Academic Editor: Michel Darmon

Received: 5 July 2022

Revised: 21 September 2022

Accepted: 24 October 2022

Published: 11 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Blind signal separation (BSS) has been of great interest for various practical applications, including image enhancement, speech enhancement, and biomedical signal processing [1–8]. This is attributed to the fact that BSS has the ability to separate sources from the observed mixtures without requiring source localization knowledge or array geometry information. Such flexibility has made BSS a popular technique in many applications.

In general, BSS techniques can be classified into the classes of higher order-based BSS and second order-based BSS. Higher order-based BSS requires assumption of the source density function. Second order-based BSS, on the other hand, requires assumption of the second order statistics of the source, such as non-stationarity or non-whiteness. In this paper, we concentrate on second order-based BSS.

The second order-based BSS for convolutive mixtures with fixed step size has previously been investigated in [4]. In [9], a steepest descent algorithm with an optimum step size procedure was proposed for the second order gradient-based BSS to improve the convergence. In [7,10], a conjugate gradient algorithm was developed for the BSS problem in which a transformation is employed to convert the constrained optimization problem with complex unmixing matrices into an unconstrained optimization problem. An accelerated gradient method with optimal step size was proposed for solving the problem in [8]. The algorithm was combined with the kurtosis measure and single speech enhancement technique to further reduce the background noise in the speech-dominant outputs.

In this paper, we propose improving the convergence rate of the accelerated gradient BSS algorithm further by incorporating the conjugate gradient descent into the accelerated

algorithm. The search direction for each iteration is obtained by utilizing both the accelerated gradient descent direction at the current iteration and the direction obtained from the previous iteration. The idea of an accelerated conjugate gradient algorithm has previously been applied for MR Imaging reconstruction in [11,12]. The image processing problem uses sparsity in the cost function according to [11]. Here, the audio separation is applied to short-time Fourier transform (STFT) signals over different time lags, and correlation matrices are formed for each frequency and time lag. Then, we use the non-stationarity of the speech sources as a criterion by diagonalizing multiple correlation matrices over a range of different time lags. As such, the key observation is the non-stationarity of the speech sources. Thus, the problem formulation for the speech separation problem is very different from the image processing problem; however, we have found that the accelerated conjugate gradient update can be used for both problems. In addition, a linear search method is employed to obtain the optimal step size for each iteration. This results in an accelerated conjugate gradient algorithm with optimal step size.

In addition, we calculate the optimal step size for each iteration in order to improve the convergence of the accelerated conjugate gradient algorithm. Simulation results based on data from a simulated room environment and real data from a recording environment show that the proposed algorithm converges faster than the accelerated algorithm and the steepest descent algorithm while requiring approximately the same computational complexity per iteration. This results in a lower total computational complexity for convergence of the accelerated conjugate gradient algorithm, as the proposed algorithm requires fewer iterations for convergence. In addition, the proposed system can simultaneously separate sources and suppress noise in reverberant environments, resulting in improved signal-to-interference and signal-to-noise ratios compared to existing methods.

The rest of this paper is organized as follows. The second order BSS is formulated in Section 2. In Section 3, the accelerated conjugate gradient algorithm is developed. Simulation results are provided in Section 4, and Section 5 contains the conclusions.

2. Second Order Blind Signal Separation

Consider the model used in [4] for a convolutive mixture of N sources in which the received signal vector with L microphones is $\mathbf{x}(t) = [x_1(t), \dots, x_L(t)]^T$ and where t is the sampled time index. The received signal vector is passed through $N \times L$ unmixing matrices $\{\mathbf{w}(q), 0 \leq q \leq Q-1\}$, where Q is the length of the unmixing filters, producing an output signal vector $\mathbf{y}(t)$:

$$\mathbf{y}(t) = \sum_{q=0}^{Q-1} \mathbf{w}(q)\mathbf{x}(t-q). \quad (1)$$

The objective is to estimate the unmixing matrices \mathbf{w} to recover the sources up to an arbitrary scaling and permutation [4]. The exact number of sources is usually unknown, and is assumed to be the same as the number of microphones, i.e., $N = L$. The received signal is processed in block BSS in the frequency domain. The signal is divided into M blocks, and the correlation matrix $\mathbf{R}_x(\omega, m), 0 \leq m \leq M-1$ for each frequency $\omega \in \Omega$ is calculated [9], where Ω denotes the frequency set.

We denote by \mathbf{W} the set consisting of all the frequency domain unmixing matrices $\mathbf{W} = \{\mathbf{W}(\omega), \omega \in \Omega\}$. The problem is to estimate \mathbf{W} that jointly decorrelates M correlation matrices $\mathbf{R}_x(\omega, m)$. This problem can be formulated as minimizing the following weighted cost over the frequencies ω ,

$$\begin{aligned} f(\mathbf{W}) &= \sum_{\omega \in \Omega} f(\mathbf{W}, \omega) \\ &= \sum_{\omega \in \Omega} \gamma(\omega) \sum_{m=0}^{M-1} \|\text{offdiag}\{\mathbf{R}_y(\omega, m)\}\|_F^2 \end{aligned} \quad (2)$$

where

$$\mathbf{R}_y(\omega, m) = \mathbf{W}(\omega)\mathbf{R}_x(\omega, m)\mathbf{W}^H(\omega) \tag{3}$$

and $\|\cdot\|_F^2$ is the Frobenius norm. Here, $\text{offdiag}\{\cdot\}$ denotes the matrix which is the same as $\{\cdot\}$ for elements outside the diagonal and zeros for elements along the diagonal. The weighting function $\gamma(\omega)$ is obtained from the correlation matrix $\mathbf{R}_x(\omega, m)$, $0 \leq m \leq M - 1$ as

$$\gamma(\omega) = \frac{1}{\sum_{m=0}^{M-1} \|\mathbf{R}_x(\omega, m)\|_F^2}. \tag{4}$$

Additional constraints on the unmixing matrices are included to avoid a zero solution, where the diagonal elements of the matrix $\mathbf{W}(k)$ are restricted to being one for all k , i.e., $\mathbf{W}_{l,l}(\omega) = 1$ for all l and ω . The minimization of the weighted cost function in (2) can lead to an arbitrary permutation of the frequency bins. A method to overcome this problem is to constrain the corresponding time domain unmixing weight \mathbf{w}_{l_1,l_2} to length D . This effectively requires zero coefficients for time domain elements greater than D , which restricts the solutions to being continuous or smooth in the frequency domain [4]. As such, the optimization problem to obtain the unmixing weighting matrix can be formulated as follows:

$$\begin{aligned} & \min_{\mathbf{W}} f(\mathbf{W}) \\ & \text{subject to } \mathbf{W}_{l,l}(\omega) = 1, \text{ for all } l \text{ and } \omega \text{ and } \mathbf{w}_{l_1,l_2} \text{ has length } D. \end{aligned} \tag{5}$$

3. Accelerated Conjugate Gradient Algorithm

In [8], the accelerated gradient algorithm was developed to solve problem (5). Here, we propose the combination of the accelerated gradient algorithm and the conjugate gradient algorithm to further improve the convergence of the optimization problem in (5). This results in an accelerated conjugate gradient algorithm with a step size optimized for each iteration. As problem (5) restricts the elements along the diagonal of matrix $\mathbf{W}(w)$ for all w to 1, we only need to update the coefficients for the off-diagonal elements. For each iteration $k > 0$, the descent search direction is calculated as

$$\begin{aligned} \Delta\mathbf{W}^{(k)}(\omega) &= -\text{offdiag}\left[\nabla f(\mathbf{W}^{(k)}, \omega)\right] \\ &= -2\text{offdiag}\left[\gamma(\omega) \sum_{m=0}^{M-1} \text{offdiag}[\mathbf{R}_y^{(k)}(\omega, m)]\mathbf{W}^{(k)}(\omega)\mathbf{R}_x(\omega, m)\right] \end{aligned} \tag{6}$$

where ∇ denotes the gradient. This direction is then transformed into the time domain and the coefficients are truncated to a length D . The constrained time domain search direction is transformed back into the frequency domain to obtain the direction $\hat{\Delta}\mathbf{W}^{(k)}$.

For each iteration k of the accelerated gradient algorithm [8], instead of using the descent direction at $\mathbf{W}^{(k)}$ we use the descent direction at the point $\mathbf{V}^{(k)}$, which is a combination of the unmixing matrices $\mathbf{W}^{(k)}$ and $\mathbf{W}^{(k-1)}$ at the k and $k - 1$ iterations, respectively:

$$\mathbf{V}^{(k)} = \mathbf{W}^{(k)} + \frac{k-1}{k+2}(\mathbf{W}^{(k)} - \mathbf{W}^{(k-1)}). \tag{7}$$

This provides an improvement in the convergence of the accelerated algorithm, as it takes into consideration the unmixing matrix at the previous iterations [13–15]. The descent direction at $\mathbf{V}^{(k)}$ is obtained, resulting in $\Delta\mathbf{V}^{(k)}$. Similar to $\Delta\mathbf{W}^{(k)}$, this direction is then transformed into the time domain and the coefficients are truncated to a length D . The constrained time domain search direction is transformed back into the frequency domain, resulting in the direction $\hat{\Delta}\mathbf{V}^{(k)}$.

With the accelerated conjugate gradient algorithm [11], instead of using the gradient direction $\hat{\Delta}\mathbf{V}^{(k)}$ from $\mathbf{V}^{(k)}$, the conjugate gradient direction is set based on the gradient direction and the previous searched direction. We denote by $\mathbf{s}^{(k-1)}$ the search direction in

the $(k - 1)^{\text{th}}$ direction. The search direction $\mathbf{s}^{(k)}$ at the k^{th} iteration is obtained from $\hat{\Delta}\mathbf{V}^{(k)}$ and $\mathbf{s}^{(k-1)}$ as

$$\mathbf{s}^{(k)} = \hat{\Delta}\mathbf{V}^{(k)} + \beta^{(k)}\mathbf{s}^{(k-1)} \tag{8}$$

where $\beta^{(k)}$ is the step size, which is defined as in [16,17]. One possible choice for $\beta^{(k)}$ is

$$\beta^{(k)} = \frac{\|\mathbf{s}^{(k)}\|^2}{\|\mathbf{s}^{(k-1)}\|^2} \tag{9}$$

where $\|\cdot\|$ denotes the l_2 norm. The combination of the current search direction and the previous direction allows the algorithm to converge faster to the optimal solution compared to the accelerated algorithm.

For each iteration, the step size $\mu^{(k)}$ is searched in the direction $\mathbf{s}^{(k)}$ as

$$\mu^{(k)} = \arg \min_{\mu} f(\mathbf{V}^{(k)} + \mu\mathbf{s}^{(k)}). \tag{10}$$

The unmixing weight matrices are updated according to

$$\mathbf{W}^{(k+1)} = \mathbf{V}^{(k)} + \mu^{(k)}\mathbf{s}^{(k)}. \tag{11}$$

The proposed algorithm can be summarized as follows:

Procedure 1: Accelerated conjugate gradient algorithm for BSS

- Step 1: Initialize the iteration $k = 1$ and the time domain unmixing weight matrices $\mathbf{w}^{(1)}(\tau)$ as

$$\mathbf{w}^{(1)}(\tau) = \begin{cases} \mathbf{I}_{L \times L}, & \text{if } \tau = 0 \\ \mathbf{0}_{L \times L}, & \tau \geq 1 \end{cases} \tag{12}$$

where $\mathbf{I}_{L \times L}$ is a $L \times L$ identity matrix and τ is the time domain index. Transform $\mathbf{w}^{(1)}(\tau)$ into the frequency domain to obtain $\mathbf{W}^{(1)}$.

- Step 2: Obtain $\mathbf{V}^{(k)}$ as in (7). As $\mathbf{V}^{(k)}$ can have a higher objective function than $\mathbf{W}^{(k)}$, we can compare the cost function of $\mathbf{V}^{(k)}$ and $\mathbf{W}^{(k)}$. If $f(\mathbf{V}^{(k)}) > \rho f(\mathbf{W}^{(k)})$ where ρ is a constant greater than 1, then we can reset $\mathbf{V}^{(k)} = \mathbf{W}^{(k)}$. The accelerated gradient algorithm “slides” slightly further than the gradient descent direction obtained from $\mathbf{W}^{(k)}$. The projected search direction $\hat{\Delta}\mathbf{V}^{(k)}$ is obtained from $\Delta\mathbf{V}^{(k)}$ by transforming the direction $\Delta\mathbf{V}^{(k)}$ to the time domain, truncating to length D and transforming it back to the frequency domain. Next, we obtain the conjugate gradient descent direction $\mathbf{s}^{(k)}$ as in (8).
- Step 3: Obtain the step size for the k^{th} iteration as in (10) and update the coefficient matrices as in (11).
- Step 4: Calculate the cost function $f^{(k+1)}(\mathbf{W})$ at the $k + 1$ iteration. If

$$|10 \log_{10} f^{(k+1)}(\mathbf{W}) - 10 \log_{10} f^{(k)}(\mathbf{W})| < \epsilon \tag{13}$$

where ϵ is a small tolerance and $|\cdot|$ denotes the absolute value. Then, proceed to Step 5. Otherwise, set $k = k + 1$ and return to the beginning of Step 2.

- Step 5: Stop the procedure. The optimum unmixing matrix is $\mathbf{W}^{(k+1)}$.

For each iteration k , the problem in (10) becomes finding the optimal step size $\mu^{(k)}$ that minimizes the objective function $f(\mathbf{V}^{(k)} + \mu\mathbf{s}^{(k)})$. We now employ an efficient procedure for searching the optimum step size for each iteration k . The search for $\mu^{(k)}$ can be viewed as a one-dimensional optimization problem with respect to μ . Thus, a line search using a golden search algorithm is employed to obtain the optimal step size. The search for the optimum step size can be described as follows.

Procedure 2: Search for an optimum step size $\mu^{(k)}$ that minimizes the cost function (10).

- Step 1: Initialize a step size $\mu > 0$, a constant $c > 1$, and an accuracy level ϵ_1 . Set $\mu_0 = 0$, $s = \mu$, and $\mu_1 = s$.
- Step 2: Obtain the cost functions $f(\mu_0)$ and $f(\mu_1)$. If $f(\mu_0) \leq f(\mu_1)$, then reduce the initial step size by setting $s/c \rightarrow s$. Let $\mu_1 = s$ and proceed to the beginning of Step 2. Otherwise, $f(\mu_0) > f(\mu_1)$ and continue to Step 3.
- Step 3: Increase s by setting $cs \rightarrow s$ and let $\mu_2 = \mu_1 + s$. Calculate the cost function $f(\mu_2)$. If $f(\mu_1) > f(\mu_2)$, then set $\mu_0 = \mu_1$, $\mu_1 = \mu_2$ and return to the beginning of Step 3. Otherwise, proceed to Step 4.
- Step 4: We now have three points, μ_0 , μ_1 , and μ_2 , satisfying

$$f(\mu_0) > f(\mu_1), f(\mu_1) \leq f(\mu_2), \text{ and } \mu_0 < \mu_1 < \mu_2. \quad (14)$$

Thus, there exists a local minimum in the interval $[\mu_0, \mu_2]$. As such, a parabola is fitted among three points $[\mu_0, f(\mu_0)]$, $[\mu_1, f(\mu_1)]$, and $[\mu_2, f(\mu_2)]$. The minimum a of this parabola in the interval $[\mu_0, \mu_2]$ can be calculated using the parabola approximation [7]. Update μ_0 , μ_1 , μ_2 ; stop the procedure if $|\mu_2 - \mu_1|$ is small enough and output the optimal step size.

The advantage of Procedure 2 is that it is relatively simple. In addition, by combining parabolic interpolation with a step size search, the optimum step size $\mu^{(k)}$ can be found with just a few calculations of the cost function.

4. Design Examples

The proposed speech enhancement scheme is evaluated in a simulated room environment with a fast-ISM room simulator and in a real car recording environment. The performance measures for the source outputs are obtained by assuming the access of the source, interference, and noise separately. Here, fourth-order kurtosis is employed to determine the two outputs with dominant speech levels [10]. A high value of kurtosis indicates that the distribution tends towards supergaussian. Because speech signals follow a Laplacian distribution, they belong to the supergaussian case. For a case with speech sources, interference, and noise, the dominant speech outputs are determined as the ones corresponding to the highest kurtosis.

As such, fourth-order kurtosis is employed to determine the two most dominant speech outputs. For each output, the source with the highest power is viewed as the main source, while the other is viewed as the interference. Hence, for each output we denote \hat{P}_s , \hat{P}_{int} and \hat{P}_n as the power spectral density (PSD) of the source, interference, and noise, respectively. The performance measure of each output is quantified in terms of a signal to interference ratio and signal to noise ratio, defined as

$$\text{SIR} = 10 \log_{10} \frac{\int_{-\pi}^{\pi} \hat{P}_s(\omega) d\omega}{\int_{-\pi}^{\pi} \hat{P}_{int}(\omega) d\omega} \text{ [dB]} \quad (15)$$

$$\text{SNR} = 10 \log_{10} \frac{\int_{-\pi}^{\pi} \hat{P}_s(\omega) d\omega}{\int_{-\pi}^{\pi} \hat{P}_n(\omega) d\omega} \text{ [dB]}. \quad (16)$$

We first consider the case with a simulated room environment.

4.1. Case 1: Simulated Room Environment

We consider the case of a simulated room environment and a linear array; the room has dimensions $4 \times 4 \times 2.5 \text{ m}^3$ with a sampling rate $f_s = 16,000 \text{ Hz}$ (see Figure 1). The inter-element distance for the microphone array is 0.04 m. We have three microphones at the $[x, y, z]$ positions $[1.96, 0.5, 1]$, $[2.0, 0.5, 1]$, and $[2.04, 0.5, 1]$, respectively. The speech signal is from the TIMIT library and the noise is from the NOISEX-92 library. The signal-to-interference ratio (SIR) is 0 dB, while the signal-to-noise ratio (SNR) is 0 dB and 10 dB. The length of the data is 8 seconds. The outputs are obtained using a fast-ISM room simulator [18] with reverberation time $T_{60} = 0.15 \text{ s}$.

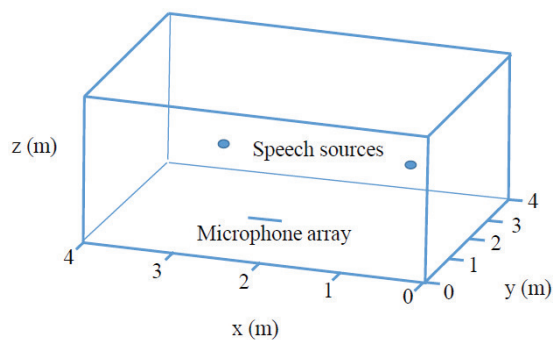


Figure 1. Setup of acoustic room and layout of microphone array for the indoor BSS.

Figures 2 and 3 show the convergence for (i) the steepest descent, (ii) the accelerated gradient, and (iii) the accelerated conjugate gradient algorithms for the case with 0 dB and 10 dB SNR, respectively. All algorithms have optimal step size in each iteration and converge with $\epsilon = 2 \times 10^{-3} = 0.002$ in (13); moreover, we set the constant ρ as $\rho = 1.1$ and allow the unmixing weight matrices $\mathbf{V}^{(k)}$ to have a cost function up to 10% larger than the cost of the unmixing weight matrices $\mathbf{W}^{(k)}$.

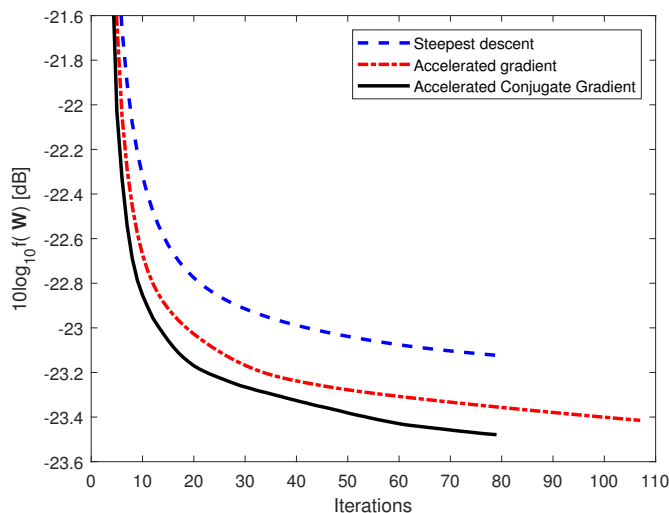


Figure 2. Convergence for simulated room recording data with three microphones and 0 dB SNR.

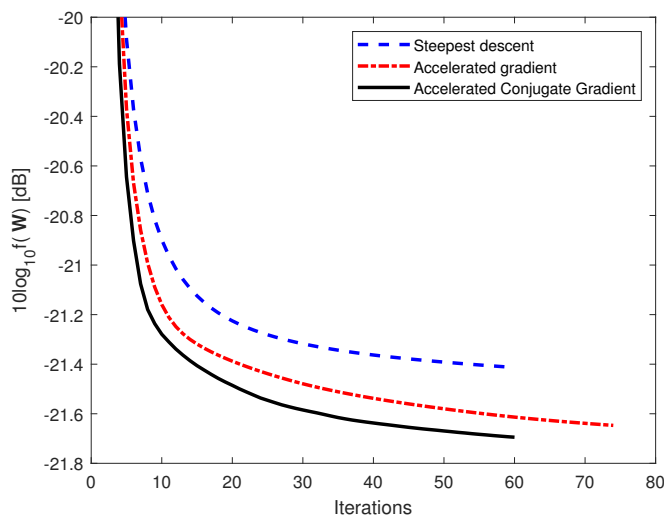


Figure 3. Convergence for simulated room recording data with three microphones and 10 dB SNR.

It can be seen from the figures that the accelerated conjugate gradient converges faster than the accelerated gradient and steepest descent algorithms for both SNR = 0 dB and 10 dB with a smaller objective function. In particular, for the case with SNR = 0 dB, the accelerated conjugate gradient requires 79 iterations for convergence with a smaller objective function, while the accelerated gradient requires 109 iterations for convergence.

Tables 1 and 2 show the SIR and SNR for the two dominant speech outputs for the three methods with the SNR = 0 dB and 10 dB, respectively. For each output, one speech is viewed as the source while the other is viewed as the interference. It can be seen from the tables that the accelerated conjugate gradient has improved performance over the other two methods.

Table 1. Second order BSS in a simulated room environment with two speech sources, noise, and three microphones for the dominant speech output detected by the kurtosis. The SNR is 0 dB.

Methods	Output 1		Output 2	
	SIR [dB]	SNR [dB]	SIR [dB]	SNR [dB]
Steepest descent with optimum step size	7.1859 dB	6.8365 dB	7.1438 dB	6.2152 dB
Accelerated gradient with optimum step size	10.2324 dB	9.4395 dB	12.3647 dB	10.3396 dB
Accelerated conjugate gradient with optimum step size	10.9354 dB	9.6536 dB	12.4842 dB	11.0761 dB

Table 2. Second order BSS in a simulated room environment with two speech sources, noise, and three microphones for the dominant speech output detected by the kurtosis. The SNR is 10 dB.

Methods	Output 1		Output 2	
	SIR [dB]	SNR [dB]	SIR [dB]	SNR [dB]
Steepest descent with optimum step size	7.8648 dB	16.8737 dB	8.0902 dB	16.9402 dB
Accelerated gradient with optimum step size	10.2414 dB	18.9972 dB	12.1539 dB	20.0236 dB
Accelerated conjugate gradient with optimum step size	10.9882 dB	19.0562 dB	13.7579 dB	20.8592 dB

For the case with 10 dB SNR, the accelerated conjugate gradient method has an improvement of 3.12 dB SIR over the steepest descent method and 0.75 dB SIR over the accelerated gradient method for the first output. For the second output, the accelerated conjugate gradient method has an improvement of 5.67 dB SIR over the steepest descent method and 1.60 dB SIR over the accelerated gradient method. Similar improvements of the accelerated conjugate gradient method over other existing methods can be seen with the SNR measure.

4.2. Case 2: Real Car Recording Data

We now consider a second case with recording data from a real car. Evaluations are performed for a double-talk situation in a real car hands-free environment for a Toyota Landcruiser 4WD driven on sealed roads at a speed of 60 km/h. A linear array with 40 mm spacing was mounted on the dashboard in front of the passenger seat. Data were gathered on a multichannel DAT-recorder with a sampling rate of 16 kHz. The main speech source is a male speaker in the front seat, while the interference is the female speaker in another seat. The male and female speech sources have the same power and operate at the same time, resulting in an SIR of 0 dB. The values of ϵ and ρ are the same as in Case 1.

First, we focus on the convergence comparison between the accelerated conjugate gradient, the accelerated gradient method, and the steepest descent method. Figures 4 and 5

compare the convergence rates for the case with three microphones and SNR = 0 dB and 10 dB, respectively. It can be seen from both figures that the accelerated conjugate gradient with an optimal step size converges faster than the steepest descent and the accelerated gradient with optimal step size, with a lower objective function.

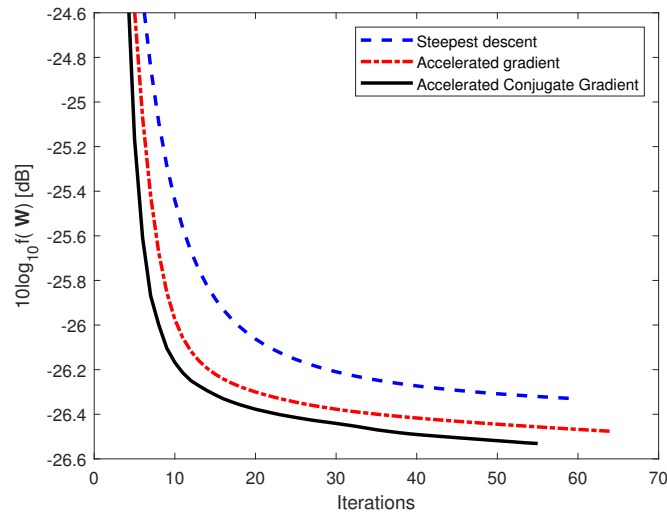


Figure 4. Convergence for real car recording data with three microphones and 0 dB SNR.

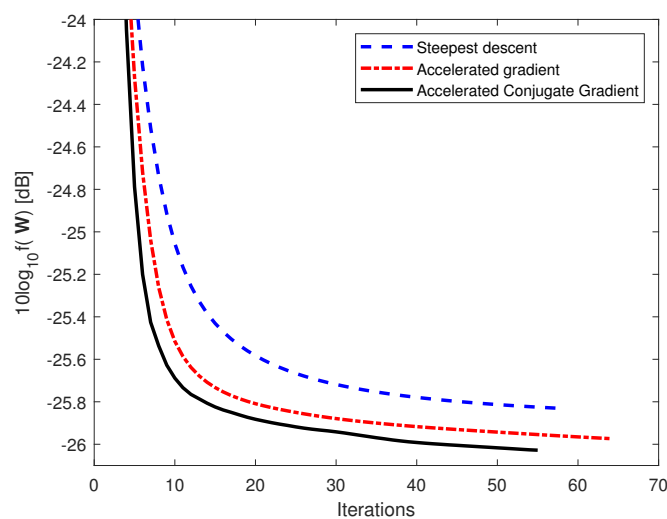


Figure 5. Convergence for real car recording data with three microphones and 10 dB SNR.

Tables 3 and 4 show the SIR and SNR measures for the two speech dominant outputs with SNR = 0 dB and 10 dB, respectively. The outputs from the accelerated conjugate gradient method have higher SIR and SNR when compared with the other two methods. In particular, for 0 dB input SNR, the first output for the accelerated conjugate gradient method achieves an improvement of 1.50 dB for SIR and 2.00 dB for SNR over the accelerated gradient method. In addition, the method achieves an improvement of 2.73 dB for SIR and 3.12 dB for SNR over the steepest descent method. For the second output, the accelerated conjugate gradient method achieves a 2.52 dB improvement for SIR and 4.08 dB for SNR over the accelerated gradient method. In addition, the method achieves an improvement of 2.66 dB for SIR and 5.36 dB for SNR over the steepest descent method.

Table 3. Second order BSS in a real car recording environment with two speech sources, noise, and three microphones for the dominant speech output detected by the kurtosis. The SNR is 0 dB.

Methods	Output 1		Output 2	
	SIR [dB]	SNR [dB]	SIR [dB]	SNR [dB]
Steepest descent with optimum step size	7.6236 dB	7.5242 dB	0.8961 dB	4.1718 dB
Accelerated gradient with optimum step size	8.8505 dB	8.6473 dB	2.4538 dB	6.8649 dB
Accelerated conjugate gradient with optimum step size	10.3518 dB	10.6519 dB	4.9757 dB	9.5280 dB

Table 4. Second order BSS in a real car recording environment with two speech sources, noise, and three microphones for the dominant speech output detected by the kurtosis. The SNR is 10 dB.

Methods	Output 1		Output 2	
	SIR [dB]	SNR [dB]	SIR [dB]	SNR [dB]
Steepest descent with optimum step size	7.1142 dB	16.1042 dB	0.8555 dB	14.0435 dB
Accelerated gradient with optimum step size	8.3185 dB	16.8156 dB	2.1784 dB	16.6476 dB
Accelerated conjugate gradient with optimum step size	9.8636 dB	18.5544 dB	4.5940 dB	19.5139 dB

5. Conclusions

This paper proposes a new algorithm based on an accelerated conjugate gradient method for solving the second order gradient-based blind signal separation (BSS) with convolutive mixtures. Simulation results on a simulated room environment and real recording environment from a car show that the proposed algorithm converges faster than the steepest descent algorithm and accelerated algorithm while having lower computational complexity. In addition, the results show that the proposed system can separate speech signals while significantly reducing background noise.

Author Contributions: Formal analysis, H.H.D.; Investigation, H.H.D. and S.N.; Writing—original draft, H.H.D.; Writing—review & editing, S.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kasak, P.; Jarina, R. Evaluation of Blind Source Separation Algorithms Applied to Music Signals. In Proceedings of the ELEKTRO, Krakow, Poland, 23–26 May 2022 ; pp. 1–8.
2. Zhou, R.; Han, J.; Li, T.; Guo, Z. Fast Independent Component Analysis Denoising for Magnetotelluric Data Based on a Correlation Coefficient and Fast Iterative Shrinkage Threshold Algorithm. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–15. [[CrossRef](#)]
3. Ezilarasan, M.R.; Pari, J.B.; Aanandhasaravanan, K.; Prasanna, N.V.; Balaji, D. Performance analysis of ICA algorithm for Blind Source Separation. In Proceedings of the International Conference on Smart Structures and Systems (ICSSS), Chennai, India, 23–24 July 2020.
4. Parra, L.; Spence, C. Convolutive blind separation for non-stationary sources. *IEEE Trans. Speech Audio Process.* **2000**, *6*, 320–327. [[CrossRef](#)]

5. Yong, P.C.; Nordholm, S.; Dam, H.H. Noise Estimation Based on Soft Decisions and Conditional Smoothing for Speech Enhancement. In Proceedings of the International Workshop on Acoustic Signal Enhancement (IWAENE 12), Aachen, Germany, 4–6 September 2012; pp. 4640–4643.
6. Yong, P.C.; Nordholm, S.; Dam, H.H. Optimization and evaluation of sigmoid function with a priori SNR estimate for real-time speech enhancement. *Speech Commun.* **2013**, *55*, 358–376. [[CrossRef](#)]
7. Dam, H.H.; Rimantho, D.; Nordholm, S. Second-Order Blind Signal Separation with Optimal Step Size. *Speech Commun.* **2013**, *55*, 535–543. [[CrossRef](#)]
8. Dam, H.H.; Nordholm, S. Accelerated gradient with optimal step size for second-order blind signal separation. *Multidimens. Syst. Signal Process.* **2018**, *29*, 903–919. [[CrossRef](#)]
9. Dam, H.H.; Nordholm, S.; Low, S.Y.; Cantoni, A. Blind Signal Separation using Steepest Descent Method. *IEEE Trans. Signal Process.* **2007**, *55*, 4198–4207.
10. Dam, H.H.; Cantoni, A.; Nordholm, S.; Teo, K.L. Second-Order Blind Signal Separation for Convolutional Mixtures Using Conjugate Gradient. *IEEE Signal Process. Lett.* **2008**, *15*, 79–82.
11. Li, X.; Wang, W.; Xiang, S.Z.W.; Wu, X. Generalized Nesterov Accelerated Conjugate Gradient Algorithm for a Compressively Sampled MR Imaging Reconstruction. *IEEE Access.* **2020**, 157130–157139. [[CrossRef](#)]
12. Yi, D.; Ji, S.; Bu, S. An Enhanced Optimization Scheme based on Gradient Descent Method for Machine Learning. *Symmetry* **2019**, *11*, 942. [[CrossRef](#)]
13. Beck, A.; Teboulle, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problem. *SIAM J. Imag. Sci.* **2009**, *2*, 183–203. [[CrossRef](#)]
14. Tseng, P. On accelerated proximal gradient methods for convex concave optimization. *Siam J. Optim.* 2008, submitted for publication. Available online: <http://www.math.washington.edu/tseng/papers/apgm.pdf> (accessed on 21 September 2022).
15. Giselsson, P.; Doanb, M.D.; Keviczky, T.; Schutter, B.D.; Rantzer, A. Accelerated gradient methods and dual decomposition in distributed model predictive control. *Automatica* **2013**, *49*, 829–833. [[CrossRef](#)]
16. McCormick, G.P. *Nonlinear Programming. Theory, Algorithms, and Applications*; John Wiley & Sons: New York, NY, USA, 1983.
17. Fletcher, R.; Reeves, C.M. Functions minimization by conjugate gradient. *Comput. J.* **1964**, *7*, 149–154. [[CrossRef](#)]
18. Lehmann, E.; Johansson, A. Prediction of energy decay in room impulse responses simulated with an image-source model. *J. Acoust. Soc. Am.* **2008**, *124*, 269–277. [[CrossRef](#)] [[PubMed](#)]