TIME DOMAIN OPTIMIZATION TECHNIQUES FOR BLIND SEPARATION OF NON-STATIONARY CONVOLUTIVE MIXED SIGNALS

Iain Russell, Alfred Mertins, Jiangtao Xi

Telecommunications and Information Technology Research Institution,
University of Wollongong,
Wollongong, N.S.W 2522, AUS
iain@titr.uow.edu.au, mertins@elec.uow.edu.au, jiangtao@uow.edu.au

ABSTRACT

This paper aims to solve the problem of Blind Signal Separation (BSS) in a convolutive environment based on output correlation matrix diagonalization. Firstly an extension of the closed form gradient and Newton methods used by Joho and Rahbar [1] is developed which encapsulates the more difficult convolutive mixing case. This extension is completely in the time domain and thus avoids the inherent permutation problem associated with frequency domain approaches. We also compare the performance of three commonly used algorithms including Gradient, Newton and global optimization algorithms in terms of their convergence behavior and separation performance in the instantaneous case and then the convolutive case.

Keywords - *Blind source separation, Global optimization, Joint diagonalization, multivariate optimization, Newton method, Steepest gradient descent.*

1. INTRODUCTION

Blind Signal Separation (BSS) has been a topic which attracted many researchers in recent years. BSS is a challenging problem in that neither the signal sources nor the mixing system are known. The only information which can be used is the assumption that the unknown signal sources are statically independent. Suppose there are N statistically independent sources, $\mathbf{s}(t) = [s_1(t), ..., s_N(t)]^T$. These sources are mixed in a medium providing M sensor or observed signals, $\mathbf{x}(t) = [x_1(t), ..., x_M(t)]^T$, given by:

$$\mathbf{x}(t) = \mathbf{H}(\mathbf{t}) * \mathbf{s}(t) \tag{1}$$

where $\mathbf{H}(t)$ is a $M \times N$ mixing matrix with its element $\mathbf{h}_{ij}(t)$ being the impulse response from jth source signal to ith measurement. $*$ defines the convolution of corresponding elements of $\mathbf{H}(t)$ and $\mathbf{s}(t)$ following the same rules for matrix multiplication.

Assuming that the mixing channels can be modelled as FIR filters with length P, Equation (1) can be rewritten as:

$$\mathbf{x}(t) = \sum_{\tau=0}^{P-1} \mathbf{H}(\tau)\mathbf{s}(t-\tau). \tag{2}$$

The M observed signals $\mathbf{x}(t)$ are coupled to the N reconstructed signals $\hat{\mathbf{s}}(t)$ via the de-mixing system. The demixing system has a similar structure to the mixing system. It contains $N \times M$ FIR filters of length Q,where $Q \geq P$. The de-mixing system can also be expressed as a $N \times M$ Matrix $\mathbf{W}(t)$, with its element $\mathbf{w}_{ij}(t)$ being the impulse response from jth measurement to ith output. The reconstructed signal can be obtained as:

$$\hat{\mathbf{s}}(t) = \sum_{\tau=0}^{Q-1} \mathbf{W}(\tau)\mathbf{x}(t-\tau) \tag{3}$$

where $\hat{\mathbf{s}}(t) = [\hat{s}_1(t), ..., \hat{s}_N(t)]^T$. A straight forward approach for BSS is to identify the unkown system first and then to apply the inverse of the identified system to the measurement signals in order to restore the signal sources. However this approach can lead to problems of instability. Therefore it is desired that the de-mixing system be estimated based on the observations of mixed signals.

The simplest case is the instantaneous mixing in which matrix $\mathbf{H}(t)$ is a constant matrix with all elements being scalar values. Many effective algorithms have been developed for this case during the past years. However, in many practical applications mixing is convolutive, in which situation BSS is much more difficult due to the expansion of complexity associated with the mixing system. The frequency domain approaches are considered to be effective to separate signal sources in convolutive cases, but another difficult issue, the inherent permutation and scaling ambiguity in each individual frequency bin, arise which makes the reconstruction of signal sources almost impossible [2]. Therefore it is worthwhile to develop an effective approach in the time domain.

Joho and Mathis [1] proposed a BSS approach based on joint diagonalization of the output signal correlation matrix using gradient and Newton optimization methods. However the approaches in [1] are limited to the instantaneous mixing cases whilst in the time domain.

This paper aims to extend the approaches in [1] to the convolutive mixing cases. In Section 2 the approaches in [1] are briefly reviewed. The extended approach in convolutive mixing cases is given in Section 3. Section 4 presents computer simulation results which gives a performance comparison of three optimization methods: Gradient, Newton and Global optimization.

The following notations are used in this paper. Vectors and matrices are printed in boldface with matrices being in capitals. Matrix and vector transpose, complex conjugation, and Hermitian transpose are denoted by $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H \triangleq ((\cdot)^*)^T$, respectively. † is the pseudo-inverse while $E(\cdot)$ means the expectation operation. $\|\cdot\|_F$ is the Frobenius norm of a matrix. With $\mathbf{a} = diag(\mathbf{A})$ we obtain a vector whose elements are the diagonal elements of $\mathbf{A}$ and $diag(\mathbf{a})$ is a square diagonal matrix which contains the elements of $\mathbf{a}$. $ddiag(\mathbf{A})$ is a diagonal matrix where its diagonal elements are the same as the diagonal elements of $\mathbf{A}$ and

$$off(\mathbf{A}) \triangleq \mathbf{A} - ddiag(\mathbf{A}). \tag{4}$$

2. OPTIMIZATION OF INSTANTANEOUS BSS

This section gives a brief review of the algorithms proposed in [1]. Assuming that the sources are statistically independent and non-stationary, observing the signals over K different time slots, we define the following noise free instantaneous BSS problem. In the instantaneous mixing cases both the mixing and de-mixing matrices are constant, that is, $\mathbf{H}(t) = \mathbf{H}$ and $\mathbf{W}(t) = \mathbf{W}$. In this case the reconstructed signal vector can be expressed as:

$$\hat{\mathbf{s}}(t) = \mathbf{W}\mathbf{x}(t). \tag{5}$$

The instantaneous correlation matrix of $\hat{\mathbf{s}}(t)$ at timeframe k can be obtained as:

$$\mathbf{R}_{\hat{s}\hat{s},k} = \mathbf{W}\mathbf{R}_{xx,k}\mathbf{W}^H \tag{6}$$

where $\mathbf{R}_{xx,k}$ is defined as:

$$\mathbf{R}_{xx,k} = \mathbf{x}(k)\mathbf{x}^H(k). \tag{7}$$

For a given set of K observed instantaneous correlation matrices, $\{\mathbf{R}_{xx,k}\}_{k=1}^K$, the aim is to find a matrix $\mathbf{W}$ that minimizes the following cost function:

$$\mathcal{J}_1 \triangleq \sum_{k=1}^{K} \beta_k \|off(\mathbf{W}\mathbf{R}_{xx,k}\mathbf{W}^H)\|_F^2 \tag{8}$$

where $\{\beta_k\}$ are positive weighting *normilization* factors such that the cost function is independent of the absolute norms and are given as:

$$\beta_k = (\sum_{k=1}^{K} \|\mathbf{R}_{xx,k}\|_F^2)^{-1}. \tag{9}$$

Perfect joint diagonalization is possible under the condition that $\{\mathbf{R}_{xx,k}\} = \{\mathbf{H}\mathbf{\Lambda}_{ss,k}\mathbf{H}^H\}$ where $\{\mathbf{\Lambda}_{ss,k}\}$ are diagonal matrices due to the assumption of the mutually independent unknown sources. This means that full diagonalization is possible, and when this is achieved, the cost function is zero at its global minimum. This constrained non-linear multivariate optimization problem can be solved using various techniques including gradient-based steepest descent, Newton and global optimization routines. However, the performance of the first two techniques depends on the initial guess of the global minimum, which in turn relies heavily on an initialization of the unknown system that is near the global trough. If this is not the case then the solution may be sub-optimal as the algorithm gets trapped in one of the local multi-minima points. Global optimization routines such as those that utilize tunnelling, simulated annealing and a combination of first and second order methods allow a more robust convergence of the cost function to the global minimum.

To prevent a trivial solution where $\mathbf{W} = \mathbf{0}$ would minimize Equation (8), some constraints need to be placed on the unknown system $\mathbf{W}$ to prevent this. One possible constraint is that $\mathbf{W}$ is unitary. This can be implemented as a penalty term such as given below:

$$\mathcal{J}_2 \triangleq \|\mathbf{W}\mathbf{W}^H - \mathbf{I}\|_F^2 \tag{10}$$

or as a hard constraint that is incorporated into the adaptation step in the optimization routine. For problems where the unknown system is constrained to be unitary, Manton presented a routine for computing the Newton step on the manifold of unitary matrices referred to as the *complex Stiefel manifold*. For further information on derivation and implementation of this hard constraint refer to [1] and references therein.

The closed form analytical expressions for first and second order information used for gradient and Hessian expressions in optimization routines are taken from Joho and Rahbar and will be referred to when generating results for convergence. Both the Steepest gradient descent (SGD) and Newton methods are implemented following the same frameworks used by Joho and Rahbar. The primary weakness of these optimization methods is that although they do converge relatively quickly there is no guarantee for convergence to a global minimum which provides the only true solution. This is exceptionally noticeable when judging the audible separation of speech signals. As one contribution of this paper we provide a comparative analysis of an existing global optimization algorithm which solves the problem of converging to local minima without a good initialization of the unknown system and note the differences in convergence and quality of separation with existing SGD and Newton methods.

3. EXTENDED CONVOLUTIVE BSS ALGORITHM IN THE TIME DOMAIN

As mentioned previously and as with most BSS algorithms that assume convolutive mixing, solving many BSS problems in the frequency domain for individual frequency bins can exploit the same algorithm derivation as the instantaneous BSS algorithms in the time domain. However the inherent *frequency permutation problem* remains a problem and will always need to be addressed. The tradeoff is that by formulating algorithms in the frequency domain we can perform less computations and processing time falls but we still must fix the permutations for individual frequency bins so that they are all aligned correctly. The main contribution in this paper is to provide a way to utilize the existing algorithm developed for instantaneous BSS but avoid the permutation problem.

Now we extend the above approach to the convolutive cases. We still assume that the de-mixing systems are defined by Equation (3),which consists of $N \times M$ FIR filters with length Q. We want to get a similar expression to those in the instantaneous cases. It can be shown that Equation (3) can be written as the following matrix form:

$$\hat{\mathbf{s}}(n) = \mathcal{W}\mathcal{X}(n) \tag{11}$$

where $\mathcal{W}$ is a $(N \times QM)$ matrix given by:

$$\mathcal{W} = [\mathbf{W}(0), \mathbf{W}(1), ..., \mathbf{W}(Q-1)] \tag{12}$$

and $\mathcal{X}(n)$ is a $(QM \times 1)$vector defined as:

$$\mathcal{X}(n) = \begin{bmatrix} \mathbf{x}(n) \\ \mathbf{x}(n-1) \\ \vdots \\ \mathbf{x}(n-(Q-1)) \end{bmatrix}. \tag{13}$$

Then the output correlation matrix at time k can be derived as:

$$\mathbf{R}_{\hat{s}\hat{s},k} = \mathcal{W}\mathbf{R}_{\mathcal{X}\mathcal{X},k}\mathcal{W}^H \tag{14}$$

where,

$$\mathbf{R}_{\mathcal{X}\mathcal{X},k} = \mathcal{X}(k)\mathcal{X}^H(k). \tag{15}$$

Correlation matrices for the recovered sources for all necessary time lags can also be obtained as:

$$\mathbf{R}_{\hat{s}\hat{s},k}(\tau) = \mathcal{W}E\{\mathcal{X}(k)\mathcal{X}^H(k+\tau)\}\mathcal{W}^H = \mathcal{W}\mathbf{R}_{\mathcal{X}\mathcal{X},k}(\tau)\mathcal{W}^H \tag{16}$$

If the inner order of $\mathcal{W}$ does not match the inner order of $\mathbf{R}_{\mathcal{X}\mathcal{X},k}(\tau)$ due to the fact that Q is chosen arbitrarily then a subset of $\mathbf{R}_{\mathcal{X}\mathcal{X},k}(\tau)$ or a zero padded version of it is used depending on whether $Q < \tau$ or $Q > \tau$, respectively.

Using the joint-diagonalization criterion in [1] for the instantaneous modelling of the BSS problem we can formulate a similar expression for convolutive mixing in the time domain. Consider the correlation matrices with all different time lags we should have the following cost function:

$$\mathcal{J}_3 \triangleq \sum_{\tau=-\tau_{min}}^{\tau_{max}} \sum_{k=1}^{K} \beta_k \|off(\mathcal{W}\mathbf{R}_{\mathcal{X}\mathcal{X},k}(\tau)\mathcal{W}^H)\|_F^2. \tag{17}$$

The only difference between $\mathcal{J}_1$ and $\mathcal{J}_3$ is that we now take into account all the different time lags τ for the correlation matrices for each respective time epoch k where the SOS are changing. Also β_k is now defined as,

$$\beta_k = (\sum_{\tau=-\tau_{min}}^{\tau_{max}} \sum_{k=1}^{K} \|\mathbf{R}_{\mathcal{X}\mathcal{X},k}(\tau)\|_F^2)^{-1}, \tag{18}$$

and we note the new structure of $\mathcal{W}$. In the ideal case where we know the exact system $\mathcal{W}_{ideal}$, all off-diagonal elements would equal zero and the value of the objective function would reach its global minimum where $\mathcal{J}_3 = 0$. Each value of k represents a different time window frame where the Second Order Statistics (SOS) are considered stationary over that particular time frame. In adjacent non-overlapping time frames k, the SOS are changing due to the non-stationarity assumption. As this is a non-linear constrained optimization problem with NQM unknown parameters we can rewrite it as,

$$\mathcal{W}_{opt} = \arg\min_{\mathcal{W}} \mathcal{J}_3(\mathcal{W}) \quad s/t \ \|ddiag(\mathcal{W}\mathcal{W}^H - \mathbf{I})\|_F^2 = 0 \tag{19}$$

Due to the structure of the matrices and with the technique of matrix multiplication to perform convolution in the time domain, optimization algorithms similar to those performed in the instantaneous climate can be utilized. Notice also that in the instantaneous version the constraint used to prevent the trivial solution $\mathbf{W} = 0$ was a unitary one. In the convolutive case a different constraint is used where the row vectors of $\mathcal{W}$ are normalized to have length one. Again referring to the SGD and Newton algorithms closed form analytical expressions of the gradient and Hessian deduced by Joho and Rahbar are extended slightly to accommodate the time domain convolutive climate of the new algorithm. These expressions are shown in Table 1. $\mathbf{R}_{\mathcal{X}\mathcal{X},k}(\tau)$ will be denoted as $\mathbf{R}^{\tau}_{\mathcal{X}\mathcal{X},k}$. With these expressions the steepest-descent algorithm and the Newton method are summarized in the Tables 2 and 3 respectively and are mainly similar to the method proposed by Joho and Rahbar, however, they

Table 1: Closed Form Analytical Expressions for the gradient and Hessian of the cost function and constraints

Cost function - $\mathcal{J}_{\mathcal{W}}$
$\mathcal{J}_{\mathcal{W}} \triangleq \sum_{\tau=-\tau_{min}}^{\tau_{max}} \sum_{k=1}^{K} \left\| off(\mathcal{W}\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}\mathcal{W}^H) \right\|_F^2$
Gradient - $\mathbf{G}_{\mathcal{W}}$
$\mathbf{G}_{\mathcal{W}} = 2\sum_{\tau=-\tau_{min}}^{\tau_{max}} \sum_{k=1}^{K} \{ off(\mathcal{W}\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}\mathcal{W}^H)\mathcal{W}{\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^H$ $+off(\mathcal{W}{\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^H\mathcal{W}^H)\mathcal{W}\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}\}$
Hessian - $\mathbf{H}_{\mathcal{W}}$
$\mathbf{H}_{\mathcal{W}} = 2\sum_{\tau=-\tau_{min}}^{\tau_{max}} \sum_{k=1}^{K} \{ ({\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^* \otimes off(\mathcal{W}\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}\mathcal{W}^H))$ $+({\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^T \otimes off(\mathcal{W}{\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^H\mathcal{W}^H))$ $+({\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^T\mathcal{W}^T \otimes \mathbf{I})\mathbf{P}_{off}(\mathcal{W}^*{\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^* \otimes \mathbf{I})$ $+({\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^*\mathcal{W}^T \otimes \mathbf{I})\mathbf{P}_{off}(\mathcal{W}^*{\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^T \otimes \mathbf{I})$ $+(\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}\mathcal{W}^H \otimes \mathbf{I})\mathbf{P}_{vec}\mathbf{P}_{off}(\mathcal{W}^*{\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^* \otimes \mathbf{I})$ $+({\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^H\mathcal{W}^H \otimes \mathbf{I})\mathbf{P}_{off}\mathbf{P}_{vec}(\mathcal{W}^*{\mathbf{R}_{\mathcal{X}\mathcal{X},k}^{\tau}}^T \otimes \mathbf{I})\}$
Row-normailzed Constraint
$\mathcal{J}_4 = \left\| ddiag(\mathcal{W}\mathcal{W}^H - \mathbf{I}) \right\|_F^2$
Constraint Gradient
$\mathbf{G}_4 = 4ddiag(\mathcal{W}\mathcal{W}^H - \mathbf{I})\mathcal{W}$

Table 2: Steepest-descent algorithm for the joint-diagonalization task with a weighted constraint

Initialization $(k = 0) : \mathcal{W}_0$

For $k = 1, 2, ...$

$\triangle\mathcal{W}_k = -\mu(\mathbf{G}_{\mathcal{W}} + \alpha\mathbf{G}_4)$
$\mathcal{W}_{k+1} = \mathcal{W}_k + \triangle\mathcal{W}_k$

work for convolutive mixing. A software package called TOMLAB [3] was used to solve the constrained global optimization problem. The application serves as an interface between the user defined problem which in our case has been outlined as the convolutive BSS algorithm in the time domain, and numerous global optimization solver routines that are included. For a more detailed explanation of the global solver routines refer to [3] and references therein.

Table 3: Newton algorithm for the joint-diagonalization task with a weighted constraint

Initialization $(k = 0) : \mathcal{W}_0$

For $k = 1, 2, ...$

$\triangle\mathcal{W}_k = (\mathbf{H}_{\mathcal{W}} + \alpha\mathbf{H}_4)^{-1}(\mathbf{G}_{\mathcal{W}} + \alpha\mathbf{G}_4)$
$\mathcal{W}_{k+1} = \mathcal{W}_k + \triangle\mathcal{W}_k$

4. SIMULATION

To demonstrate the performance of the extended convolutive BSS algorithm in the time domain we firstly investigated the instantaneous BSS algorithm using a variety of optimization techniques. A set of $K = 15$ real diagonal square matrices $\{\Lambda_k\}$ were randomly chosen representing the unknown source input uncorrelated matrices. The diagonal assumption is crucial to all BSS problems as it reflects that the sources are mutually independent allowing separation. Following the assumption that the unknown separating system $\mathbf{W}$ is unitary, preventing the trivial solution, the observed correlation matrices can be constructed where $\{\mathbf{R}_{xx,k}\} = \{\mathbf{H}\mathbf{\Lambda}_k\mathbf{H}^H\}$ and $\mathbf{H}$ is chosen as a two by two unitary mixing matrix,

$$\mathbf{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}. \tag{20}$$

This simulation compares the different optimization methods used in [1] with the global optimization solver routine *glcCluster* available in the TOMLAB software environment. Equation (8) forms the objective to be optimized while Equation (10) forms the constraint preventing a trivial solution of the unknown separating system $\mathbf{W}$. Figure 1 shows the comparison of the convergence rates of each optimization algorithm for ten independent runs with ten distinct sets of correlation matrices. It is evident that with the second order information available, convergence of the Newton algorithm is much quicker. The optimization for this particular instantaneous BSS problem where the system is assumed to be unitary is performed on the Stiefel manifold. The step size $\mu = 0.2$ was used and the various slopes of the different convergence curves of the gradient method depends entirely on the ten different sets of randomly generated diagonal input matrices.

With the SGD and Newton methods, convergence to the global minimum depends entirely on a good initial starting point $\mathbf{W}_0$. The starting point selected in this simulation was

$$\mathbf{W}_0 = \begin{bmatrix} cos(1) & -sin(1) \\ sin(1) & cos(1) \end{bmatrix}. \tag{21}$$

Although taking slightly longer to converge than the Newton method, the benefit of using global optimization solvers such as *glcCluster* to solve the constrained optimization problem is that no information on the initial starting point is necessary for the global minimum to be found. This is an important point as the goal of BSS problems is to separate observed signals with the least amount of information and assumptions as possible. Although techniques like geometric beamforming [4] are good procedures to allow feasible starting points for optimization, they require additional information and assumptions on the problem space.

To investigate the performance of the extended instantaneous BSS algorithm to the convolutive case in the time domain the SGD and Newton algorithm implementation were slightly changed. As the constraint no longer requires the unknown system $\mathcal{W}$ to be unitary the constraint was changed to that given in Equation (19). The technique of weighted penalty functions was used to ensure the constraints preventing the trivial solution were met. No longer performing the optimization on the Stiefel manifold the SGD and Newton algorithms were changed and the steps for each algorithm are shown in Tables 2 and 3 respectively. A TITO two tap FIR known mixing system was chosen and is given below in the z domain.

$$\mathbf{H}(z) = \begin{bmatrix} 1+z^{-1} & -1+z^{-1} \\ -1+z^{-1} & 1+z^{-1} \end{bmatrix}. \qquad (22)$$

The corresponding known unmixing system which would separate mixed signals which are produced by convolving the source signals with the TITO mixing system $\mathbf{H}(z)$ given above is

$$\mathbf{W}_{ideal}(z) = \begin{bmatrix} 1+z^{-1} & 1-z^{-1} \\ 1-z^{-1} & 1+z^{-1} \end{bmatrix}. \qquad (23)$$

The convolution of these two systems in cascade would ensure the global system $\mathbf{G}(z) = \mathbf{W}_{ideal}(z)\mathbf{H}(z)$ would be a delayed version of the identity, i.e. $z^{-1}\mathbf{I}$. Using matrix multiplication to perform convolution in the time domain, Equation (12) can be used to represent the equivalent structure of Equation (23),

$$\mathcal{W}_{ideal} = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix}. \qquad (24)$$

Setting the parameters $\mu = 0.6$ and $\alpha = 0.2$ we solve the constrained optimization problem given in Equation (19) using the SGD, Newton method and global optimization solver(s) from TOMLAB. With the Newton method it should be noted that in Table 3 a closed form expression for the Hessian of the constraint has not been given. For this part numerical differentiation using finite differences was used to obtain $\mathbf{H}_4$. A set of $K = 15$ real diagonal square uncorrelated matrices for the unknown source input signals were randomly generated. Using convolution in the time domain a corresponding set of correlation matrices $\mathbf{R}^{\tau}_{\mathcal{X}\mathcal{X},k}$ for each respective time instant $k = 1, ..., 15$ at multiple time lags τ were generated for the observed signals. Each optimization algorithm was run ten independent times and convergence graphs similar to the instantaneous case were observed and are shown in Figure 2. The main difference between Figure 1 and Figure 2 is that the algorithms take longer to converge. This is due to the increased dimension of unknown variables. Again the initialization of the SGD and Newton algorithms plays an important role in the convergence to either a local or global minimum. Initial values for the estimated unmixing system $\mathcal{W}$ were randomly generated by adding Gaussian random variables with standard

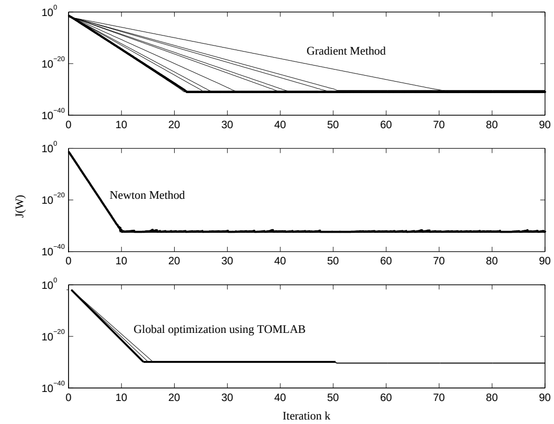

Figure 1: Convergence of differing optimization methods for instantaneous BSS.

deviation $\sigma = 0.1$ to the coefficients of the true system. For the global optimization routine *glcCluster* in TOMLAB no initial value for the unknown system is needed. This particular solver uses a global search to approximately obtain the set of all global solutions and then uses a local search method which utilizes the derivative expressions to obtain more accuracy on each global solution.

After convergence of the objective function to an order of magnitude approximately equal to 10^{-16} the unknown de-mixing FIR filter system $\mathcal{W}$ in cascade with the known mixing system $\mathbf{H}(z)$ resulted in a global system which was equivalent to a scaled and permuted version of the true global system $z^{-1}\mathbf{I}$ as can be seen by the following example,

$$\mathbf{G}(0) = \begin{bmatrix} 0.008 & 0.003 \\ 0.008 & 0.003 \end{bmatrix}, \quad \mathbf{G}(1) = \begin{bmatrix} 0.007 & -0.98 \\ -0.87 & 0.006 \end{bmatrix}, \quad \mathbf{G}(2) = \begin{bmatrix} 0.013 & -0.003 \\ -0.013 & 0.003 \end{bmatrix}. \qquad (25)$$

5. CONCLUSION

A new method for convolutive BSS in the time domain using an existing instantaneous BSS framework has been presented. This method avoids the inherent permutation problem when dealing with solving the convolutive BSS problem in the frequency domain. Optimization algorithms including SGD and Newton methods have been compared to a global optimization routine *glcCluster* available in TOMLAB, a robust software package for solving global opti-

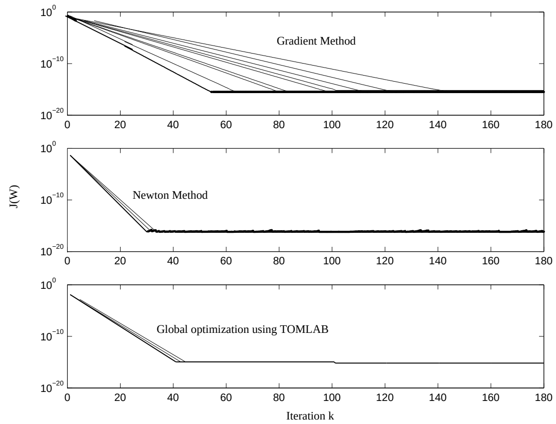

Figure 2: Convergence of differing optimization methods for convolutive BSS.

mization problems, for both the instantaneous and convolutive mixing environments. Future work will be directed at implementing the simulations with recorded data such as speech and audio mixed in a real reverberant environment.

6. REFERENCES

[1] M. Joho and H. Mathis, "Joint diagonalization of correlation matrices by using Newton methods with application to blind signal separation," (Rossyln, VA), pp. 403–407, SAM, Aug 4-6 2002.

[2] M. Ikram and D. Morgan, "Exploring permutation inconsistency in blind separation of speech signals in a reverberant environment," (Instanbul, Turkey), pp. 1041–1044, ICASSP, June 2000.

[3] K. Holmström, "User's Guide for TOMLAB v4.0." URL: http://tomlab.biz/docs/tomlabv4.pdf, Sept. 2 2002.

[4] L. Parra and C. Alvino, "Geometric source separation: Merging convolutive source separation with geometric beamforming," *IEEE Trans. on Speech and Audio Process.*, vol. 10, pp. 352–362, September 2002.