A PREDICTION MATRIX APPROACH TO CONVOLUTIVE ICA

Lars Kai Hansen and Mads Dyrholm
Informatics and Mathematical Modelling
Technical University of Denmark B321
DK-2800 Lyngby, DENMARK
lkh@imm.dtu.dk mad@imm.dtu.dk

Abstract. *A linear prediction approach reduces convolutive independent component analysis (ICA) to the following three steps: Solution of a set of multivariate linear prediction problems, a linear multivariate deconvolution problem with known matrix coefficients, and finally solution of a conventional instantaneous mixing ICA problem.*

CONVOLUTIVE MIXING

Independent component analysis (ICA) of convolutive mixtures is a problem of considerable current interest in neural computation, say for modeling speech processing and furthermore has numerous applications in signal processing, see, e.g., [1, 2, 3, 4, 5].

Convolutive ICA in its simplest form concerns reconstruction of the $L + 1$ mixing matrices and the T source signal vectors from a D-dimensional convolutive mixture,

$$\mathbf{x}_t = \sum_{\tau=0}^{L} \mathbf{A}_\tau \mathbf{s}_{t-\tau}, \quad t = 1, ..., T. \tag{1}$$

The K-dimensional source signal vectors are assumed temporally independent: $p(\{\mathbf{s}_t\}) = \prod_{t=1}^{T} p(\mathbf{s}_t)$. We will assume L is so large that the correlations in the process $\mathbf{x}$ can be explained by the mixing matrices. In fact, as noted by [1], possible auto-correlations of the source signals can not be identified without additional a priori information. In order to see this, note that in the frequency domain the convolution becomes a product of Fourier transforms

$$\mathbf{x}_\omega = \mathbf{A}_\omega \mathbf{s}_\omega, \tag{2}$$

hence, any non-zero 'filter' $h(\omega)$ can be multiplied on a given source if $1/h(\omega)$ is applied to the corresponding column of $\mathbf{A}_\omega$. Another observation is that

for stationary Gaussian white noise sources, the sufficient statistic $\langle \mathbf{x}_t \mathbf{x}_{t+\delta}^\top \rangle$ does not allow full recovery of the mixing matrices since

$$\langle \mathbf{x}_t \mathbf{x}_{t+\delta}^\top \rangle = \sum_\tau \mathbf{A}_\tau \mathbf{A}_{\tau+\delta}^\top, \tag{3}$$

which is invariant to common rotation $\mathbf{A}_\tau \to \mathbf{A}_\tau \mathbf{U}$ of all mixing matrices.

Most earlier approaches to convolutive ICA are based on frequency domain estimation using (2). This leads to a set of conventional 'instantaneous' ICA problems, one for each frequency, and is hampered by a massive permutation problem which can be tamed by adding a prior source 'smoothness' information or other more elaborate schemes [2, 3, 4, 5, 6]. Another line of work is based on optimization of certain 'independency measures', information maximization or other heuristics, see e.g., [7, 8, 9, 10, 11, 12, 13].

Statistically motivated maximum likelihood schemes have been proposed, typically leading to high-dimensional optimizations w.r.t. to all elements of all mixing matrices, see e.g. [14, 1, 15]. The aim in this paper is to invoke a few simple approximations and use these and straightforward linear algebra to reduce the problem to a conventional ICA problem. We will avoid the frequency domain representation all together, hence, we will not further address the frequency component permutation problem.

TEMPORAL UN-MIXING

We will present a temporal un-mixing procedure in which the key new ingredient is the use of prediction matrices, hence, this step is first illustrated on the well-understood problem of *instantaneous* ICA of temporally correlated sources.

Consider the instantaneous mixing system

$$\mathbf{x}_t = \mathbf{A}\mathbf{s}_t. \tag{4}$$

For simplicity we will consider square mixing so that $D = K$, i.e., $\mathbf{A}$ is a $K \times K$ matrix with real elements. Let the prediction matrix $\mathbf{W}_\tau$ be the best linear predictor of the series $\mathbf{x}$

$$\mathbf{x}_{t+\tau} = \mathbf{W}_\tau \mathbf{x}_t + \boldsymbol{\epsilon}_{t+\tau}. \tag{5}$$

Now right multiply (5) by the transposed source vector $\mathbf{s}_t^\top$ and average w.r.t. the source distribution. If we assume $\langle \boldsymbol{\epsilon}_{t+\tau} \mathbf{s}_t^\top \rangle \approx \mathbf{0}$, we obtain

$$\begin{aligned} \mathbf{A}\langle \mathbf{s}_{t+\tau} \mathbf{s}_t^\top \rangle &= \mathbf{W}_\tau \mathbf{A} \langle \mathbf{s}_t \mathbf{s}_t^\top \rangle, \\ \mathbf{A}\mathbf{C}_\tau &= \mathbf{W}_\tau \mathbf{A}\mathbf{C}_0, \end{aligned} \tag{6}$$

where the matrices $\mathbf{C}_0$ and $\mathbf{C}_\tau$ are diagonal because the sources are independent and constant in time by stationarity. From (6) we learn that the mixing matrix $\mathbf{A}$ is the matrix formed by the eigenvectors of the prediction matrix.

The eigenvalues in the diagonal of the matrix $\mathbf{C}_\tau\mathbf{C}_0^{-1}$ are normalized auto-correlation values of the given source at the lag τ. Equations (5-6) form an alternative route to the so-called Molgedey-Schuster algorithm, see [16, 17]. This algorithm is a quick (closed form) ICA approach, for mixing problems with time-correlated sources and where there are values of τ for which the sources have different normalized auto-correlations, see e.g., [18] for a more detailed discussion and multi-media applications.

Next we will show how the prediction matrix method can be use to simplify the convolutive mixing problem. The linear prediction approach is first generalized to a multi-lag linear predictor of the form,

$$\mathbf{x}_{t+\tau} = \sum_{\lambda=0}^{M} \mathbf{W}_{\tau,\lambda}\mathbf{x}_{t-\lambda} + \boldsymbol{\epsilon}_{t+\tau}. \tag{7}$$

Substituting the convolutive process (1) we find

$$\sum_{\tau'=0}^{L} \mathbf{A}_{\tau'}\mathbf{s}_{t+\tau-\tau'} = \sum_{\lambda=0}^{M} \mathbf{W}_{\tau,\lambda} \sum_{\tau'=0}^{L} \mathbf{A}_{\tau'}\mathbf{s}_{t-\tau'-\lambda}. \tag{8}$$

As above we multiply (8) by $\mathbf{s}_t^\top$ and average w.r.t. the source distribution now assuming, as discussed above, that the sources are *temporally uncorrelated*: $\langle \mathbf{s}_{t+\tau}\mathbf{s}_t^\top \rangle = \mathbf{C}_0\delta_{\tau,0}$, to get

$$\mathbf{A}_\tau\mathbf{C}_0 = \mathbf{W}_{\tau,0}\mathbf{A}_0\mathbf{C}_0, . \tag{9}$$

Furthermore, assuming that all sources have non-vanishing variance we can divide by the diagonal source covariance matrix $\mathbf{C}_0$ to arrive at the result,

$$\mathbf{A}_\tau = \mathbf{W}_{\tau,0}\mathbf{A}_0. \tag{10}$$

Hence, the existence of the linear predictor (7) implies that the delayed mixing matrices are generated from the 'zero lag' mixing matrix by the prediction matrices.

We estimate the prediction matrices $\mathbf{W}_{\tau,\lambda}$ by least squares. For each value of τ separately we obtain a coupled set of equations,

$$\langle \mathbf{x}_{t+\tau}\mathbf{x}_{t-\delta}^\top \rangle = \sum_{\lambda=0}^{M} \widehat{\mathbf{W}}_{\tau,\lambda}\langle \mathbf{x}_{t-\lambda}\mathbf{x}_{t-\delta}^\top \rangle, \tag{11}$$

with the expectations estimated from the measured time series $\mathbf{x}_t$ by $\langle ... \rangle \approx \frac{1}{T}\sum_t(...)$. The linear equations in (11) are easily solved for $\widehat{\mathbf{W}}_{\tau,\lambda}$ by matrix inversion. For each value of τ we will eventually need the set of $L+1$ matrices $\mathbf{W}_{\tau,0}$, c.f., (10). Note that the coupling to the other prediction matrices (for a given τ) in (11) makes $\mathbf{W}_{\tau,0}$ different from the matrix obtained by making a linear prediction in (7) with $M = 0$.

The generator property (10) is next used to simplify the convolutive mixing problem. First rewrite (1)

$$\mathbf{x}_t = \sum_{\tau=0}^{L} \mathbf{W}_{\tau,0} \mathbf{A}_0 \mathbf{s}_{t-\tau} \equiv \sum_{\tau=0}^{L} \mathbf{W}_{\tau,0} \mathbf{u}_{t-\tau}, . \tag{12}$$

The signals $\mathbf{u}_t = \mathbf{A}_0 \mathbf{s}_t$ form an uncorrelated series as they are proportional to the source series $\mathbf{s}_t$.

We have already estimated the prediction matrices from measured data, hence, (12) is a standard linear MIMO system with *known* matrix coefficients $\mathbf{W}_{\tau,0}$, and can be solved by a variety of methods producing an estimate of the time series $\widehat{\mathbf{u}}_t$, $t = 1, ..., T$. In this work we use the simple recursive filter

$$\widehat{\mathbf{u}}_t = \mathbf{W}_{0,0}^{-1} \mathbf{x}_t - \sum_{\tau=1}^{L} \mathbf{W}_{0,0}^{-1} \mathbf{W}_{\tau,0} \widehat{\mathbf{u}}_{t-\tau}. \tag{13}$$

This filter may become unstable, in such case a more robust *regularized* estimator can be invoked, e.g., substituting

$$\mathbf{W}_{0,0}^{-1} \rightarrow \left(\kappa \mathbf{1} + \mathbf{W}_{0,0}^{\top} \mathbf{W}_{0,0} \right)^{-1} \mathbf{W}_{0,0}^{\top}, \tag{14}$$

in (13). The remaining problem is to estimate $\mathbf{A}_0$ and the source signals $\mathbf{s}_t$ from the series

$$\widehat{\mathbf{u}}_t = \mathbf{A}_0 \mathbf{s}_t. \tag{15}$$

This is a conventional ICA problem with temporally independent source signals and can be solved by any of the standard approaches. If the distribution of the source signals have positive kurtosis, as appropriate for, e.g., speech signals we can use the Infomax approach of Bell and Sejnowski [19].

Solving the problem (15) we obtain $\widehat{\mathbf{A}}_0$ and $\widehat{\mathbf{s}}_t$, using (10) we can then generate the matrices $\widehat{\mathbf{A}}_\tau$ using the $\widehat{\mathbf{W}}_{\tau,0}$'s, hence concluding our recipe for solving the convolutive mixing problem (1).

SIMULATION EXAMPLE

We illustrate the viability of the prediction based approach by a small simulation example.

A $D = K = 2$ convolutive mixture was created by first designing a set of 2×2 mixing matrices ($L = 30$). These were next applied as in (1) to a i.i.d. random source signal ($T = 30000$). The distribution of the source signals was made non-gaussian, with positive kurtosis, by the transformation $s = \text{sign}(u) * |u|^2$ where $u \sim \mathcal{N}(0,1)$. The source and the mixed signals can be seen in Figure 1, while the mixing matrices are shown for reference in Figure 7.

In Figures 2-3 we first illustrate the excellent quality of the linear model in (7).
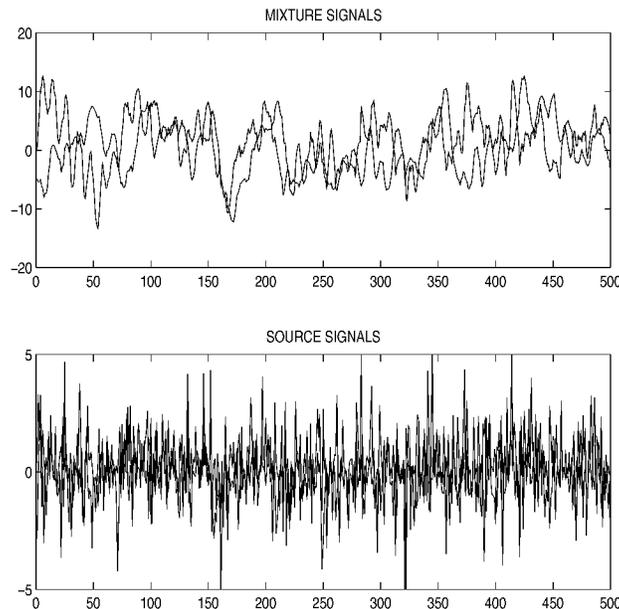
Figure 1: Simulation experiment involving $L = 30$ square mixing matrices ($D = 2$) convolved with i.d.d. long-tailed source signals. In the upper panel we show a short segment of the two convolved signals ($\mathbf{x}_t$), in the lower panel we show the corresponding segments of the two sources ($\mathbf{s}_t$). The mixing matrices are shown in Figure 7.

Figure 2 shows scatter plots of the prediction error ($\epsilon_{t+\tau}$, $\tau = 3$) vs. the source signal ($\mathbf{s}_t$). It is important for the generator relation (10) that these time series are roughly uncorrelated. In Figure 3 we have further quantified this relation as function of the prediction horizon (τ). As expected, the predictions become more and more noisy as we increase τ, i.e., the relative power in ϵ_t increases, however, more important is it that the correlation between the source signal and the error remains limited, supporting relation (10).

Next we investigate the quality of the prediction matrix estimates. The ratios $\mathbf{A}_\tau \mathbf{A}_0^{-1}$ were computed with the 'true' matrices used in the simulation. In Figure 4 we compare these matrices with the matrices estimated from data, the match is good and the other four channels are of similar quality (data not shown).

The MIMO problem is solved using (13). The relative reconstruction error was small ($\langle(\mathbf{x}_t - \widehat{\mathbf{x}}_t)^2\rangle / \langle(\mathbf{x}_t)^2\rangle < 10^{-6}$). Using our in-house implementation of the Bell and Sejnowski algorithm[1], instantaneous ICA was applied to the resulting time series $\widehat{\mathbf{u}}_t$. The estimated sources are compared with the 'true' sources in Figure 5 and the consistency is remarkable.

Using the reconstructed $\mathbf{A}_0$ we estimated the remaining matrices $\mathbf{A}_\tau$ us-
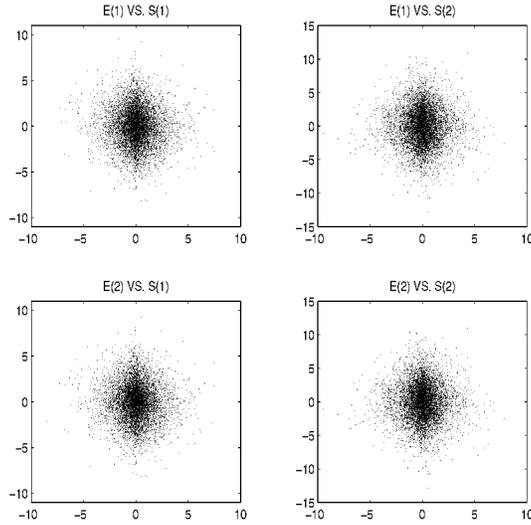
---

[1] MatLab toolbox available from www.imm.dtu.dk/cisp

Figure 2: Simulation as in Figure 1. The scatter plots illustrate the dependency between $\epsilon_{t+\tau}$ ($\tau = 3$) and $\mathbf{s}_t$, c.f. (7).
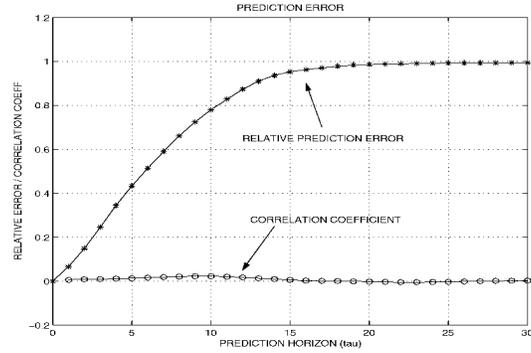

Figure 3: Simulation experiment as in Figure 1. We show the relative prediction error (the mean square error normalized by the signal variance) as function of the prediction horizon (τ), and the correlation coefficients between $\epsilon_{1,t+\tau}$ and $\mathbf{s}_{2,t}$. While the predictions become increasingly random, the correlation coefficients stay in the range $-0.05 - 0.05$, ensuring that the error in (10) is bounded.

ing (10). The matrix elements $\mathbf{A}_{1,2,\tau}$ are compared in Figure 6 with the corresponding element of matrices found by generation using the true $\mathbf{A}_0$. Apart from the absolute amplitude, these elements are in good agreement, indicating that the approach has quite successfully solved the convolutive mixing problem.
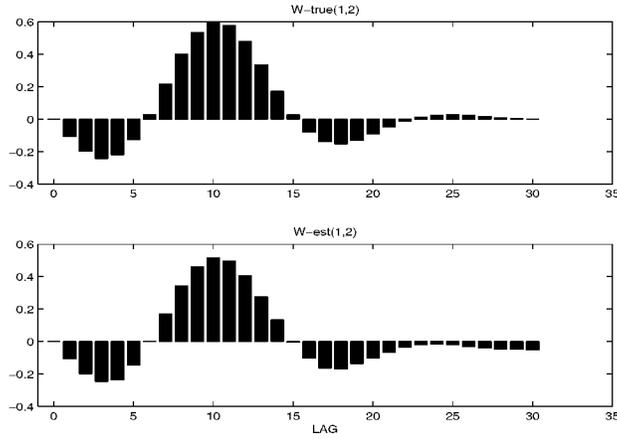
Figure 4: Simulation experiment as in Figure 1. The estimated W-matrices (using (13)) compared favorably with the 'true' matrices $\mathbf{W}_{\tau,0} \equiv \mathbf{A}_\tau \mathbf{A}_0^{-1}$.
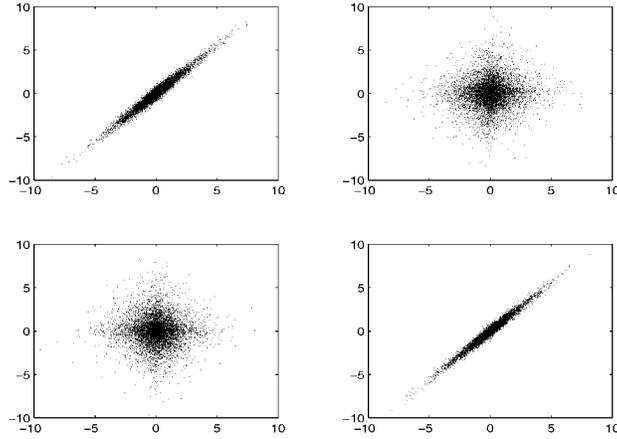

Figure 5: Simulation experiment as in Figure 1. We plot the true sources of the simulation experiment vs. the reconstructed sources. The sign and the ordering of the reconstructed sources have been modified for clarity. The reconstructed sources are well aligned with the true sources, this is highly non-trivial for convolutive mixtures.

CONCLUSION

We have proposed a linear prediction approach to the convolutive ICA problem. Within a linear prediction assumption and linear algebra, the problem is reduced to the following three steps: Solving a set of multivariate linear prediction problems, solving a linear multivariate deconvolution problem with known matrix coefficients, and finally solving a conventional instantaneous mixing ICA problem. A small simulation example showed that the
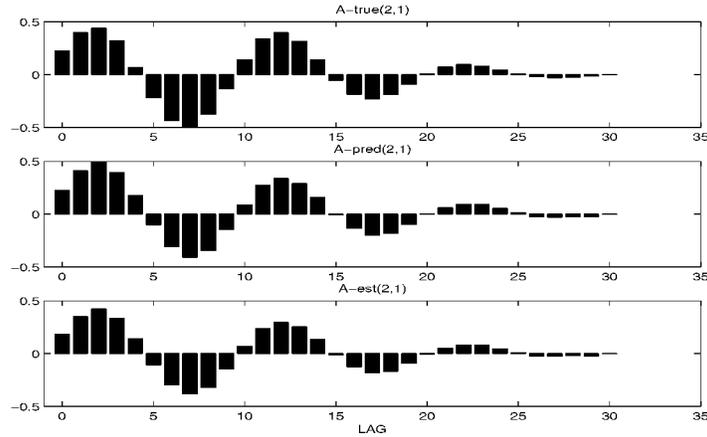
Figure 6: Simulation experiment as in Figure 1. The recovered matrix elements $\mathbf{A}_{2,1,\tau}$ (bottom row) are compared with true matrices (upper row) and the matrices obtained by multiplying the prediction matrices $\widehat{\mathbf{W}}_{\tau,0}$ on the true $\mathbf{A}_0$-matrix (middle row).

approach is able to accurately estimate the mixing matrices and the source signals. We are currently trying to identify proper conditions for the linear prediction assumption and also to invoke more robust schemes for solving the MIMO problem.

Acknowledgments

References

[1] H. Attias and C. E. Schreiner, "Blind source separation and deconvolution: the dynamic component analysis algorithm," *Neural Computation*, vol. 10, no. 6, pp. 1373–1424, 1998.

[2] L. Parra, C. Spence, and B. D. Vries, "Convolutive blind source separation based on multiple decorrelation," in *IEEE Workshop on Neural Networks and Signal Processing, Cambridge, UK, September 1998*, pp. 23–32, 1998.

[3] L. Parra and C. Spence, "Convolutive blind source separation of nonstationary sources," *IEEE Trans. on Speech and Audio Processing*, vol. 8,
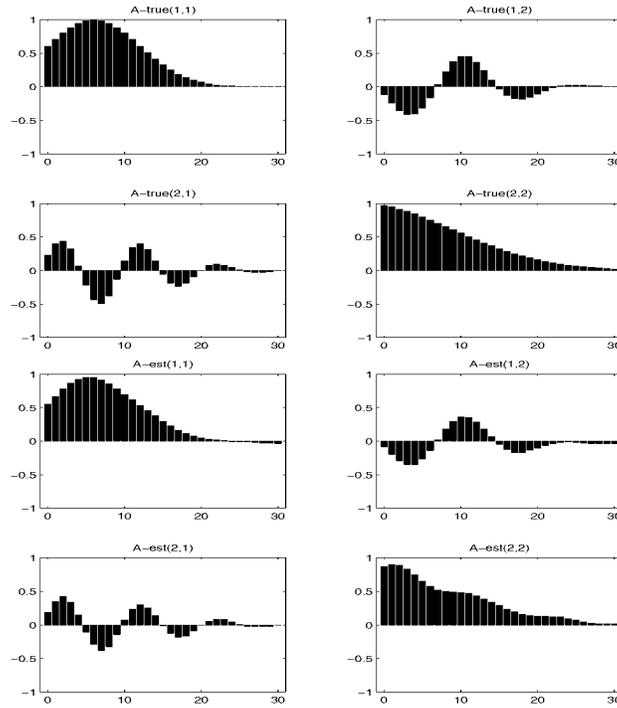
Figure 7: Simulation experiment as in Figure 1. The complete set of recovered matrix elements $\widehat{\mathbf{A}}_\tau$ (bottom four panels) are compared with the true matrix elements $\mathbf{A}_\tau$ (upper four panels).

pp. 320–327, 2000.

[4] K. Rahbar and J. Reilly, "Blind source separation of convolved sources by joint approximate diagonalization of cross-spectral density matrices," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 2745–2748, 2001.

[5] K. Rahbar, J. P. Reilly, and J. H. Manton, "A frequency domain approach to blind identification of mimo fir systems driven by quasi-stationary signals," in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 1717–1720, 2002.

[6] W. Baumann, B.-U. Kohler, D. Kolossa, and R. Orglmeister, "Real time separation of convolutive mixtures," in *3'rd International Conference on ICA. Eds. T-W. Lee et al.*, pp. 65–69, San Diego December 2001 2001.

[7] K. Torkkola, "Blind separation of convolved sources based on information maximization," in *IEEE Workshop on Neural Networks for Signal Processing, Kyoto, Japan*, pp. 423–432, September 4-6 1996.

[8] A. Belouchrani, K. A. Meraim, J.-F. Cardoso, and Éric Moulines, "A blind source separation technique based on second order statistics," *IEEE Trans. on Signal Processing*, vol. 42, pp. 434–444, 1997.

[9] S. Choi and A. Cichocki, "Blind signal deconvolution by spatio- temporal decorrelation and demixing," in *Neural Networks for Signal Processing, Proc. of the 1997 IEEE Workshop (NNSP-97), IEEE Press, N.Y. 1997*, pp. 426–435, 1997.

[10] S. Douglas, A. Cichocki, and S. Amari, "Self-whitening algorithms for adaptive equalization and deconvolution," *IEEE Trans. on Signal Processing*, vol. 47, pp. 1161–1165, 1999.

[11] R. Cristescu, T. Ristaniemi, J. Joutsensalo, and J. Karhunen, "Blind separation of convolved mixtures for cdma systems," in *Proc. of the X European Signal Processing Conference (EUSIPCO 2000), Tampere, Finland*, pp. 619–622, 2000.

[12] P. Comon, E. Moreau, and L. Rota, "Blind separation of convolutive mixtures: A constrast based joint diagonalization appraoch," in *3'rd International Conference on ICA. Eds. T-W. Lee et al.*, pp. 686–691, San Diego December 2001 2001.

[13] X. Sun and S. Douglas, "A natural gradient convolutive blind source separation algorithms for speech mixtures," in *3'rd International Conference on ICA. Eds. T-W. Lee et al.*, pp. 59–64, San Diego December 2001 2001.

[14] E. Moulines, J.-F. Cardoso, and E. Gassiat, "Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models," in *Proc. ICASSP'97 Munich*, pp. 3617–3620, 1997.

[15] S. Deligne and R. Gopinath, "An em algorithm for convolutive independent component analysis," p. To appear, 2002.

[16] L. Molgedey and H. Schuster, "Separation of a mixture of independent signals using time delayed correlation," *Physical Review Letters*, vol. 72, pp. 3634–3637, 1994.

[17] A. S. Lukic, M. N. Wernick, L. K. Hansen, and S. C. Strother, "An ica algorithm for analyzing multiple data sets," in *IEEE 2002 Int. Conf. on Image Processing (ICIP-2002)* (M. T. et al., ed.), IEEE, 2002.

[18] L. K. Hansen, J. Larsen, and T. Kolenda, "On independent component analysis for multimedia signals," in *Multimedia Image and Video Processing, CRC Press* (S. K. L. Guan and J. Larsen, eds.), pp. 175–199, 2000.

[19] T. Bell and T. Sejnowski, "An information maximisation approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, pp. 1129–1159, 1995.