

Shifted Independent Component Analysis

Morten Mørup, Kristoffer H. Madsen, and Lars K. Hansen

Technical University of Denmark
Informatics and Mathematical Modelling
Richard Petersens Plads, Building 321
DK-2800 Kgs. Lyngby, Denmark
`{mm,khm,lkh}@imm.dtu.dk`

Abstract. Delayed mixing is a problem of theoretical interest and practical importance, e.g., in speech processing, bio-medical signal analysis and financial data modelling. Most previous analyses have been based on models with integer shifts, i.e., shifts by a number of samples, and have often been carried out using time-domain representation. Here, we explore the fact that a shift τ in the time domain corresponds to a multiplication of $e^{-i\omega\tau}$ in the frequency domain. Using this property an algorithm in the case of sources \leq sensors allowing arbitrary mixing and delays is developed. The algorithm is based on the following steps: 1) Find a subspace of shifted sources. 2) Resolve shift and rotation ambiguity by information maximization in the complex domain. The algorithm is proven to correctly identify the components of synthetic data. However, the problem is prone to local minima and difficulties arise especially in the presence of large delays and high frequency sources. A Matlab implementation can be downloaded from [1].

1 Introduction

Factor analysis is widely used to reconstruct latent effects from mixtures of multiple effects based on the model

$$\mathbf{X}_{n,m} = \sum_d \mathbf{A}_{n,d} \mathbf{S}_{d,m} + \mathbf{E}_{n,m}, \quad (1)$$

where $\mathbf{E}_{n,m}$ is additive noise. However, this decomposition is not unique since $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{Q}$ and $\tilde{\mathbf{S}} = \mathbf{Q}^{-1}\mathbf{S}$ yields same approximation as \mathbf{A}, \mathbf{S} . Consequently, constraints have been imposed such as Varimax rotation for Principal Component Analysis (PCA) [2], statistical independence of the sources \mathbf{S} as in Independent Component Analysis (ICA)[3, 4]. A related strategy is sparse coding where the objective of minimizing the error is combined with a term penalizing the non-sparsity of \mathbf{S} [5].

Factor analysis in the setting of ICA is often illustrated by the so-called cocktail party problem. Here mixtures of several speakers are recorded in several microphones forming the measured signal \mathbf{X} . The task is to identify the sources

\mathbf{S} of each original speaker. However, even in an anechoic environment the mixing model is typically not accurate because of different delays in the microphones. Consider two microphones placed at distance L and $L + h$ from a given speaker. Under normal atmospheric conditions, the speed of sound is approximately $c = 344$ m/s while a typical sampling rate is $f_s = 22$ kHz. Then the delay in samples between the two microphones is given by: $\#samples = \frac{f_s h}{c}$ such that the delay increases linearly with the difference in distance. Consequently, a distance of 1 cm gives a delay of 0.6395 samples while $h = 1$ m leads to a delay of 63.95 samples. Harshman and Hong [6] proposed a generalization of the factor models in which the underlying sources have specific delays when they reach the sensors. The model is called shifted factor analysis (SFA), and reads

$$\mathbf{X}_{n,m} = \sum_d \mathbf{A}_{n,d} \mathbf{S}_{d,m-\tilde{\tau}_{n,d}} + \mathbf{E}_{n,m}. \quad (2)$$

In real acoustic environments we expect echoes due to paths that are created by reflection off surfaces. To account for general delay mixing effects, the ICA model has been generalized to convolutive mixtures, see e.g., [7–9]

$$\mathbf{X}_{n,m} = \sum_{\tau,d} \mathbf{A}_{n,d}^{\tau} \mathbf{S}_{d,m-\tau} + \mathbf{E}_{n,m}. \quad (3)$$

Here \mathbf{A}^{τ} is a filter that accounts for the presence of each source in the sensors at time delay τ . The shifted factor model, thus is a special case of the convolutive model where the filter coefficients $\mathbf{A}_{n,d}^{\tau} = \mathbf{A}_{n,d}$ if $\tilde{\tau}_{n,d} = \tau$ else $\mathbf{A}_{n,d}^{\tau} = 0$.

In fact shifted mixtures are also seen in many other contexts. For instance, astronomy where star motion Doppler effects induce frequency red shifts that can be modelled using SFA. Here we will focus on the delayed source model. In [6] strong support was found for the conjecture that the incorporation of shifts can strengthen the model enough to make the parameters identifiable up to scaling and permutation (essential uniqueness). We will demonstrate that this conjecture is not correct when allowing for arbitrary shifts. Indeed, the model is, as for regular factor analysis, ambiguous. In [10] an algorithm was proposed to estimate the model. However, the algorithm has the following drawbacks.

1. All potential shifts have to be specified in the model.
2. Exhaustive integer search for the delays is expensive.
3. The model only accounts for shifts by whole samples.
4. The model is in general not unique.

Prior to the work of [6, 10] Bell and Sejnowski [4] sketched how to handle time delays in networks based on a model similar to equation 2. This was further explored in [11]. Although their algorithms derive gradients to search for the delays (alleviating the first two drawbacks above) the models are still based on pure integer delays. In [12] a different model based on equally mixed sources, i.e. $\mathbf{A} = \mathbf{1}$, formed by moving averages incorporated non-integer delays by signal interpolation. Yeredor [13] solved the SFA model by joint diagonalization of

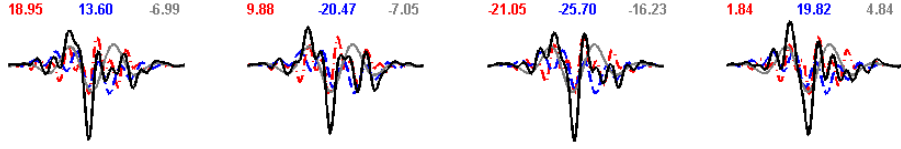


Fig. 1. Example of activities obtained (black graph) when summing three components (gray, blue dashed and red dash-dotted graphs) each shifted to various degrees (given in samples by the colored numbers). Clearly, the resulting activities are heavily impacted by the shifts such that a regular instantaneous ICA analysis would be inadequate.

the source cross spectra based on the AC-DC algorithm with non-integer shifts for the 2×2 system. This approach was extended to complex signals in [14]. The algorithm is least squares optimal for equal number of sensors and sources. More sensors than sources is not a problem for conventional ICA; we simply reduce dimension by variance decomposition, this procedure is exact for noiseless mixing. Due to the delays projection based dimensional reduction will not reproduce the simple single delay structure, but rather lead to a more general convolutive mixture. We will therefore aim at an algorithm for finding a shift invariant subspace. Hence, solve equation 2 by use of the fact that a shift τ in the time domain can be approximated by multiplication by the complex coefficients $e^{-i\omega\tau}$ in the frequency domain. This alleviates the first three drawbacks of the SFA algorithm. We will denote this algorithm a Shift Invariant Subspace Analysis (SISA). To further deal with shift and rotation ambiguities, we impose independence in the complex domain based on information-maximization (IM) [4]. Hence, we form an algorithm for ICA with shifted sources (SICA). Notice, that algorithms for ICA in the complex domain without shifts have previously been derived, see for instance [9, 15] and references therein.

2 Method and Results

In the following \mathbf{U} will denote a matrix in the time domain, while $\tilde{\mathbf{U}}$ denotes the corresponding matrix in the frequency domain. \mathcal{U} and $\tilde{\mathcal{U}}$ denotes 3-way arrays in the time and frequency domains respectively. Furthermore, $\mathbf{U} \bullet \mathbf{V}$ denotes the direct product, i.e. element-wise multiplication. Also, $\omega = 2\pi \frac{f-1}{M}$ such that $\tilde{\mathbf{U}}(f) = \mathbf{U} \bullet e^{-i2\pi \frac{f-1}{M} \tau}$. Finally, the i^{th} row of a matrix will be denoted $\mathbf{U}_{i,:}$.

2.1 Shift Invariant Subspace Analysis (SISA)

In the following we will devise an algorithm to find a shift invariant subspace based on the SFA model. Consider the SFA model and its frequency transformed

$$\mathbf{X}_{n,m} = \sum_d \mathbf{A}_{n,d} \mathbf{S}_{d,m-\tau_{n,d}} + \mathbf{E}_{n,m}, \quad \tilde{\mathbf{X}}_{n,f} = \sum_d \mathbf{A}_{n,d} \tilde{\mathbf{S}}_{d,f} e^{-i2\pi \frac{f-1}{M} \tau_{n,d}} + \tilde{\mathbf{E}}_{n,f}. \quad (4)$$

In matrix notation this can be stated as

$$\tilde{\mathbf{X}}_f = \tilde{\mathbf{A}}^{(f)} \tilde{\mathbf{S}}_f + \tilde{\mathbf{E}}_f. \quad (5)$$

Due to Parseval's identity the following holds

$$C_{ls} = \sum_{n,m} \|\mathbf{E}_{n,m}\|_F^2 = \frac{1}{M} \sum_{n,f} \|\tilde{\mathbf{E}}_{n,f}\|_F^2. \quad (6)$$

Thus, minimizing the least square error in the time and frequency domain is equivalent. The algorithm will be based on alternatingly solving for \mathbf{A} , \mathbf{S} and $\boldsymbol{\tau}$.

S update: According to equation 5, \mathbf{S}_f can be estimated as

$$\tilde{\mathbf{S}}_f = \tilde{\mathbf{A}}^{(f)\dagger} \tilde{\mathbf{X}}_f. \quad (7)$$

Although, \mathbf{S} is updated in the frequency domain the updated version has to remain real when taking the inverse FFT. For \mathbf{S} to be real valued the following has to hold

$$\tilde{\mathbf{S}}_{M-f+1} = \tilde{\mathbf{S}}_f^*, \quad (8)$$

where * denotes complex conjugate. This constraint is enforced by updating the first $\lfloor M/2 \rfloor + 1$ elements, i.e. up to the Nyquist frequency, while setting the remaining elements according to equation 8.

A update: Let $\tilde{\mathbf{S}}_{d,f}^{(n)}$ denote the delayed version of the source signal $\tilde{\mathbf{S}}_{d,f}$ to the n^{th} channel, i.e. $\tilde{\mathbf{S}}_{d,f}^{(n)} = \tilde{\mathbf{S}}_{d,f} e^{-i2\pi \frac{f-1}{M} \tau_{n,d}}$. Then equation 2 can be restated as

$$\mathbf{X}_{n,:} = \mathbf{A}_{n,:} \mathbf{S}^{(n)} + \mathbf{E}_{n,:}, \quad (9)$$

This is the regular factor analysis problem giving the update

$$\mathbf{A}_{n,:} = \mathbf{X}_{n,:} \mathbf{S}^{(n)\dagger}. \quad (10)$$

$\boldsymbol{\tau}$ update: The least square error for the model stated in equation 5, is given by

$$C_{ls} = \frac{1}{M} \sum_f (\tilde{\mathbf{X}}_f - \tilde{\mathbf{A}}^{(f)} \tilde{\mathbf{S}}_f)^H (\tilde{\mathbf{X}}_f - \tilde{\mathbf{A}}^{(f)} \tilde{\mathbf{S}}_f), \quad (11)$$

where H denotes the conjugate transpose. Define $\mathbf{T}^{ND \times 1} = \text{vec}(\boldsymbol{\tau})$, i.e. the vectorized version of the matrix $\boldsymbol{\tau}$ such that $\mathbf{T}_{n+(d-1)N} = \tau_{n,d}$. Let further

$$\tilde{\mathbf{Q}}_{n,d,f} = \tilde{\mathbf{A}}_{n,d}^{(f)} \tilde{\mathbf{S}}_{d,f}, \quad \tilde{\mathbf{E}}_f = \tilde{\mathbf{X}}_f - \tilde{\mathbf{A}}^{(f)} \tilde{\mathbf{S}}_f. \quad (12)$$

Then the gradient of C_{ls} with respect to $\tau_{n,d}$ is given as

$$\mathbf{g}_{n+(d-1)N} = \frac{\partial C_{ls}}{\partial \mathbf{T}_{n+(d-1)N}} = \frac{\partial C_{ls}}{\partial \tau_{n,d}} = \frac{-1}{M} \sum_f 2\omega \Im[\tilde{\mathbf{Q}}_{n,d,f} \tilde{\mathbf{E}}_{n,f}^*] \quad (13)$$

The Hessian has the following structure

$$\mathbf{H}_{n+(d-1)N, n'+(d'-1)N} = \begin{cases} \frac{-2}{M} \sum_f \omega^2 \Re[\tilde{\mathcal{Q}}_{n,d,f} \tilde{\mathcal{Q}}_{n',d',f}^*] & \text{if } n \neq n' \wedge d \neq d' \\ \frac{-2}{M} \sum_f \omega^2 \Re[\tilde{\mathcal{Q}}_{n,d,f} (\tilde{\mathcal{Q}}_{n',d',f}^* + \tilde{\mathbf{E}}_{n',f}^*)] & \text{if } n = n' \wedge d = d' \end{cases} \quad (14)$$

As a result, $\boldsymbol{\tau}$ can be estimated using the Newton-Raphson method

$$\mathbf{T} \leftarrow \mathbf{T} - \eta \mathbf{H}^{-1} \mathbf{g}, \quad (15)$$

where η is a step size parameter that is tuned to keep decreasing the cost function.

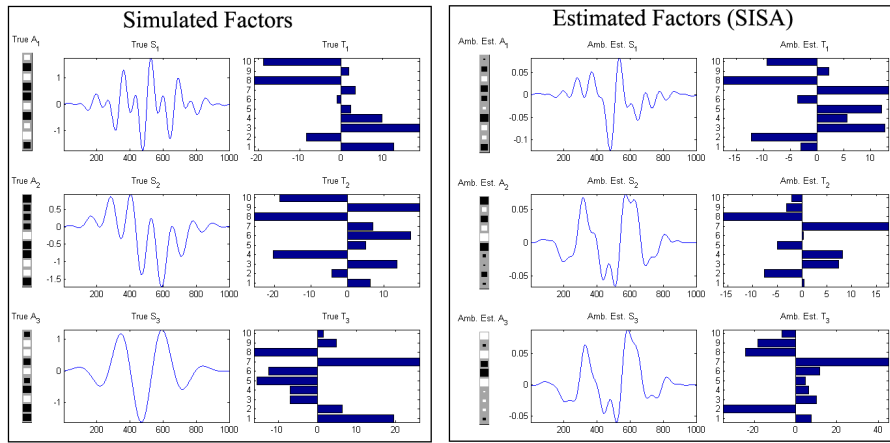


Fig. 2. Results obtained by a shift invariant subspace analysis (SISA). Left panel: the true factors forming a synthetic data set. To the left, the strength of the mixing \mathbf{A} of each source is indicated in gray color scale. In the middle, the three sources are shown and to the right is given the time delays of each source to each channel. Right panel: The estimated factors from the SISA analysis. Although, all the variance is explained the decomposition has not identified the true underlying components but an ambiguous mix. Clearly, as for regular factor analysis the SISA is not unique.

2.2 SISA is not unique

According to equation 5, the reconstructed signal in the complex domain is given as $\tilde{\mathbf{X}}_f \approx \tilde{\mathbf{A}}(f) \tilde{\mathbf{S}}_f = \tilde{\mathbf{A}}(f) \tilde{\mathbf{W}}(f) \tilde{\mathbf{W}}(f)^{-1} \tilde{\mathbf{S}}_f$. Such that $\tilde{\mathbf{W}}(f) = \mathbf{W} \bullet e^{-i2\pi \frac{f-1}{M} \hat{\boldsymbol{\tau}}}$ is a rotation, scaling and shift matrix. Assume the inverse of $\tilde{\mathbf{W}}(f)$ is also a rotation, scaling and shift matrix, i.e. $\tilde{\mathbf{W}}(f)^{-1} = \mathbf{V} \bullet e^{-i2\pi \frac{f-1}{M} \hat{\boldsymbol{\tau}}}$. Since $\tilde{\mathbf{W}}(f) \tilde{\mathbf{W}}(f)^{-1} = \mathbf{I}$, we find

$$\sum_{d''} \mathbf{W}_{d,d''} \mathbf{V}_{d',d''} e^{-i2\pi \frac{f-1}{M} (\hat{\boldsymbol{\tau}}_{d,d''} + \hat{\boldsymbol{\tau}}_{d',d''})} = \begin{cases} 0 & \text{for } d \neq d' \forall f \\ 1 & \text{for } d = d' \forall f \end{cases} \quad (16)$$

From $f = 1$ we obtain the relation $\mathbf{V} = \mathbf{W}^{-1}$. For the remaining frequencies this expression can only be valid if $\hat{\tau}_{dd''} + \check{\tau}_{d''d} = 0$ (diagonal elements) and $\hat{\tau}_{dd''} + \check{\tau}_{d''d} = k_{dd'}$ (off diagonal elements) where $k_{dd'}$ denotes an arbitrary constant. The first relation gives the constraint that $\hat{\tau} = -\check{\tau}^T$. The second relation further constraints all the elements of the columns of $\hat{\tau}$ to be equal. Thus the ambiguity is given by $\widetilde{\mathbf{W}}^{(f)} = [\mathbf{W} \text{diag}(e^{-i2\pi \frac{f-1}{M} \hat{\tau}})]$. Where $\hat{\tau}$ is a vector describing the shift ambiguity.

2.3 Shifted Independent Component Analysis (SICA)

A common approach to ICA is the maximum likelihood (ML) method [16] which corresponds to the approach of maximizing information proposed in [4]. In the framework of ML a non-gaussian distribution on the sources is assumed such that ambiguity can be resolved up to the trivial ambiguities of scale, permutation and source shifting relative to the time delays.

Define, $\widetilde{\mathbf{U}}_f = \widetilde{\mathbf{W}}^{(f)} \widetilde{\mathbf{S}}_f$, i.e. the sources at frequency f when transformed according to the rotation and shift ambiguity described in the previous section. The ambiguity can be resolved by maximizing the log-likelihood assuming the (non-gaussian) Laplace distribution $p(\widetilde{\mathbf{U}}_f) \propto e^{-|\widetilde{\mathbf{U}}_{d,f}|}$, i.e.

$$p(\widetilde{\mathbf{S}}_f | \mathbf{W}, \hat{\tau}) = \prod_f p(\widetilde{\mathbf{S}}_f | \mathbf{W}, \hat{\tau}) = \prod_f |\det(\widetilde{\mathbf{W}}^{(f)})| p(\widetilde{\mathbf{W}}^{(f)} \widetilde{\mathbf{S}}_f) \quad (17)$$

Such that the log-likelihood as a function of \mathbf{W} and $\hat{\tau}$ becomes

$$\mathcal{L}(\mathbf{W}, \hat{\tau}) = \sum_f \ln |\det(\widetilde{\mathbf{W}}^{(f)})| - \sum_d |\widetilde{\mathbf{W}}^{(f)} \widetilde{\mathbf{S}}_f|_d \quad (18)$$

By maximizing $\mathcal{L}(\mathbf{W}, \hat{\tau})$ \mathbf{W} and $\hat{\tau}$ is estimated and a new unambiguous \mathbf{S} solution found by $\widetilde{\mathbf{S}}_f = \widetilde{\mathbf{W}}^{(f)} \widetilde{\mathbf{S}}_f$. The corresponding mixing and delays can be estimated alternating between the \mathbf{A} and τ update. We initialized \mathbf{A} as $\mathbf{A} = \mathbf{A}\mathbf{W}^{-1}$ and $\tau_{i,d}$ by the maximum cross-correlation between $\mathbf{X}_{i,:}$ and $\mathbf{S}_{d,:}$.

3 Discussion

Traditionally, ICA analysis is based on subspace analysis often using singular value decomposition (SVD). The sources are then found by rotating the vectors spanning the subspace according to a measure of independence. Similarly, we derived the SISA algorithm to find a shift invariant subspace by alternating least squares. Shift and rotation ambiguities were solved by imposing independence on the amplitudes of the frequency transform of the sources. While SVD has a closed form solution the SISA algorithm is non-convex. Estimating \mathbf{A} and \mathbf{S} has a closed form solution for fixed values of τ , \mathbf{S} and \mathbf{A} . However, τ is estimated using an iterative method potentially leading to many local minima. Furthermore, the problem becomes increasingly difficult for high frequency sources and large shifts

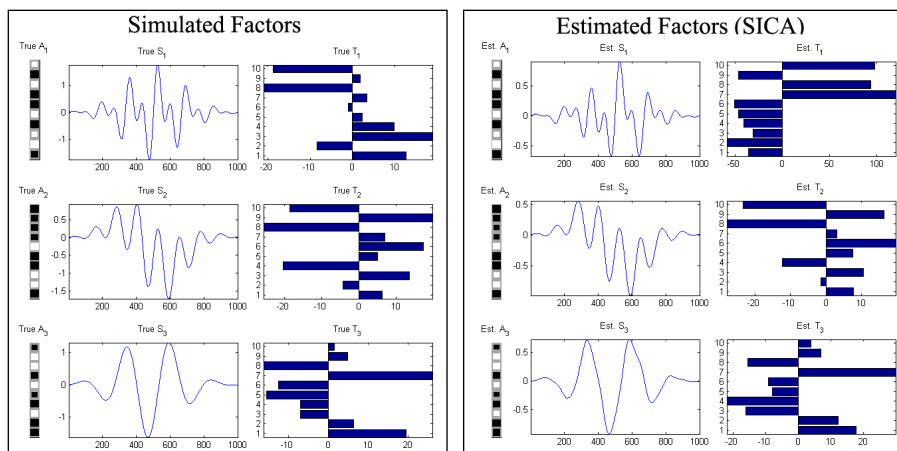


Fig. 3. Result obtained using the SICA on the decomposition found using SISA. By imposing independence, e.g., requiring the amplitudes in the frequency domain to be sparse, the rotation and shift ambiguity inherited in the model is resolved. Clearly the true underlying components and their respective mixing are correctly identified. However, a local minimum has been found, resulting in errors in the estimation of the delays for the first component.

due to additional local minima. In an example we saw this happen: The SICA algorithm failed in correctly identifying the delays of the first component; the component with the highest frequencies. A multistart strategy was invoked, we choose the best of ten random initializations to obtain a good initial solution for the estimation of the shift invariant subspace. While our algorithm was based on likelihood maximization, Yeredor [13] developed an algorithm based on joint diagonalization. The present SISA is potentially useful as a preprocessing step for this latter algorithm when estimating less sources than sensors. Current work comprises performance comparisons with this algorithm (A. Yeredor, personal communication).

Previous work based on integer shifts conjectured the decomposition to be unique [6]. When using integer shifts some shifts might perform better than others due to a better integer rounding error. Hence, this might be why the integer shifts formed seemingly unique solutions. However, as demonstrated in figure 2 the shifted factor analysis model is not in general unique. But, by imposing independence unique solutions can be obtained up to trivial permutation, scaling and specific onset relative to the delays of the sources as demonstrated in figure 3. The shift/delay model may prove useful for a wide range of data where ICA already has been employed. Furthermore, the extra information of delays can be useful for spatial source localization when combined with information of position of the sensors. Future work will focus on implementing additional constraints such as non-negativity and attempt to improve the identifiability in the presence of many local minima. The current algorithm can be downloaded from [1].

References

1. Mørup, M., Madsen, K.H.: www2.imm.dtu.dk/pubdb/views/publication_details.php?id=5206. (2007)
2. Kaiser, H.F.: The varimax criterion for analytic rotation in factor analysis. *Psychometrika* **23** (1958) 187–200
3. Comon, P.: Independent component analysis, a new concept? *Signal Processing* **36** (1994) 287–314
4. Bell, A.J., Sejnowski, T.J.: An information maximization approach to blind source separation and blind deconvolution. *Neural Computation* **7** (1995) 1129–1159
5. Olshausen, B. A., F.D.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381** (1996) 607–609
6. Harshman, R., Hong, S., Lundy, M.: Shifted factor analysis – part i: Models and properties. *Journal of Chemometrics* **17** (2003) 363–378
7. Attias, H., Schreiner, C.: Blind source separation and deconvolution: the dynamic component analysis algorithm. *Neural Computation* **10**(6) (1998) 1373–1424
8. Parra, L., Spence, C., Vries, B.: Convolutional blind source separation based on multiple decorrelation. *IEEE Workshop on Neural Networks and Signal Processing* (1998) 23–32
9. Anemuller, J., Sejnowski, T.J., Makeig, S.: Complex independent component analysis of frequency-domain electroencephalographic data. *Neural Networks* **16**(9) (2003) 1311–1323
10. Harshman, R., Hong, S., Lundy, M.: Shifted factor analysis – part ii: Algorithms. *Journal of Chemometrics* **17** (2003) 379–388
11. Torkkola, K.: Blind separation of delayed sources based on information maximization. *Acoustics, Speech, and Signal Processing, ICASSP-96* **6** (1996) 3509–3512
12. Emile, B., Comon, P.: Estimation of time delays between unknown colored signals. *Signal Processing* **68**(1) (1998) 93–100
13. Yeredor, A.: Time-delay estimation in mixtures. *ICASSP* **5** (2003) 237–240
14. Yeredor, A.: Blind source separation in the presence of doppler frequency shifts. *ICASSP* **5** (2005) 277–280
15. Cardoso, J.F., Tulay, A.: The maximum likelihood approach to complex ica. *ICASSP* (2006) 673–676
16. Hyvarinen, A., Karhunen, J., Oja, E.: *Independent Component Analysis*. John Wiley and Sons. (2001)