forum acusticum 2026
8-12 September · Graz, Austria

IMPROVING INDEPENDENT COMPONENT ANALYSIS-BASED ACOUSTIC BLIND SOURCE SEPARATION WITH BEAMFORMING

Péter Fiala[1], Máté Tóth[1]

[1] *Dept. of Networked Systems and Services, Budapest University of Technology and Economics, Hungary*

This paper is a revised version of the authors' submitted manuscript that was peer-reviewed by at least two anonymous reviewers.

Preliminary document

Abstract

Blind Source Separation (BSS) is a long-standing signal processing problem where individual signals cannot be observed separately, only as noisy mixtures produced by an unknown mixing system. The objective is to recover the original sources as efficiently as possible. This problem is highly relevant in telecommunications, medical signal processing, and acoustics, and has been widely studied in the literature.

A related problem in acoustic and radio signal processing is beamforming, which aims to determine the spatial orientation of sound sources using a microphone array of known geometry and to separate them by focusing on their directions. Given the conceptual similarity between BSS and beamforming, this paper investigates how classical BSS techniques can be adapted for acoustic beamforming.

We explore frequency-domain independent component analysis (FD-ICA), a widely used BSS method for convolutive mixtures in real acoustic environments. The paper reviews Complex-ICA methods underlying FD-ICA and examines how microphone array properties — known sensor positions and a larger number of microphones than sources — provide additional information that can mitigate limitations of classical BSS approaches. The applicability of traditional beamforming techniques within the FD-ICA framework is also analyzed.

Based on these insights, we propose a new FD-ICA-based source separation algorithm tailored to microphone arrays. The method is evaluated on real measurements, with performance compared to classical beamforming techniques and the impact of using multiple microphones assessed.

Corresponding author: *fiala@hit.bme.hu.*
Copyright: *©2026 Péter Fiala et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.*

Keywords: *independent component analysis, blind source separation, frequency domain complex ICA, beamforming, microphone Arrays*

1. Introduction

Blind Source Separation (BSS) is a common problem in signal processing, where individual signals cannot be directly observed, only their mixtures produced by an unknown mixing system, often corrupted by noise. The goal is to separate the original sources from these mixtures. A well-known example is the cocktail party problem, where humans can focus on a single speaker in noisy environments. Algorithmic solutions to this task have significant applications in telecommunications, medical signal processing, and acoustic processing.

A widely accepted solution to the BSS problem is Independent Component Analysis (ICA) [1]. Although ICA is well established and supported by efficient implementations, classical ICA is primarily applicable to instantaneous linear mixtures, which are rarely encountered in practical acoustic scenarios.

Another related problem is beamforming, widely used in acoustic and radio signal processing. Acoustic beamforming uses a microphone array with known geometry to estimate source directions and separate signals by spatial focusing. While beamforming can be interpreted as a special case of BSS, traditional beamforming methods rely on shaping the array directivity rather than exploiting statistical independence, as in ICA.

This paper investigates the feasibility and advantages of ICA-based beamforming. Since practical acoustic mixtures are typically convolutive, frequency-domain ICA (FD-ICA) [2] is adopted. The paper reviews classical ICA, complex ICA methods, and FD-ICA, and proposes improvements exploiting microphone array properties such as known sensor positions and a large number of microphones.

Based on these results, a novel beamforming-oriented FD-ICA (BF-FD-ICA) architecture is proposed. The method is evaluated using real microphone array measurements and compared with classical beamforming and existing FD-ICA approaches.

2. Independent Component Analysis for Convolutive Mixtures

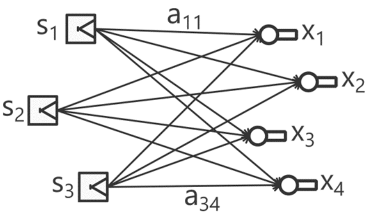

Figure 1: The mixing model

2.1. Fundamentals of ICA

Independent Component Analysis (ICA) is one of the most widely used approaches for solving the blind source separation problem [1]. It should be emphasized that the number of independent sources is apriory known. ICA is a statistical method that estimates original discrete-time signals from their instantaneous linear mixtures. Formally, let $S_1, S_2, \ldots, S_N$ denote N independent random variables representing the sources, and let $X_1, X_2, \ldots, X_N$ denote the observed mixtures. The ICA model assumes the linear relationship $\vec{X} = \mathbf{A}\vec{S}$, where $\mathbf{A}$ is an unknown mixing matrix. The objective is to estimate a demixing matrix $\mathbf{W}$ such that $\vec{Y} = \mathbf{W}\vec{X} \approx \vec{S}$. ICA methods solve this problem by finding a transformation that maximizes statistical independence between recovered signals. Perfect independence guarantees recovery up to permutation and scaling [3]. Common independence measures include non-Gaussianity, mutual information, tensor-based approaches, and maximum likelihood estimation. This work focuses on non-Gaussianity-based methods.

Most ICA algorithms assume whitened input signals. Whitening ensures zero mean, unit variance, and decorrelation, and is typically performed using Principal Component Analysis (PCA): $\vec{Z} = \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{Q}^{\mathrm{T}}\hat{\vec{X}}$, where $\hat{\vec{X}} = \vec{X} - \mathbb{E}[\vec{X}]$, is the centered input and $\mathbf{\Lambda}$ and $\mathbf{Q}$ are the eigenvalues and eigenvectors of the input covariance $\mathbf{\Sigma} = \mathbb{E}[\hat{\vec{X}}\hat{\vec{X}}^{\mathrm{T}}]$, respectively.

Non-Gaussianity-based ICA relies on the Central Limit Theorem, which states that sums of independent variables tend toward Gaussian distributions. Therefore, independent components can be estimated by maximizing non-Gaussianity. The general procedure is subsequent estimation of orthogonal mixing vectors $\vec{w}$ (rows of $\mathbf{W}$) by maximizing non-Gaussianity:

$$\vec{w} = \underset{\|\vec{w}\|=1}{\operatorname{argmax}}\, G(\vec{w}^{\mathrm{T}}\vec{Z}), \tag{1}$$

where G is a suitable function measuring non-Gaussianity. Finally, the sources are recovered as $\vec{Y} = \mathbf{W}\vec{Z}$.

In this work, we use negentropy-based ICA. Negentropy is based on differential entropy:

$$H(X) = -\int p_X(\eta)\log p_X(\eta)\,\mathrm{d}\eta, \tag{2}$$

and is defined as $J(X) = H(X_{\mathrm{gauss}}) - H(X)$, where $H(X_{\mathrm{gauss}})$ is the differential entropy of a Gaussian random variable with the same variance as X.

ICA suffers from inherent ambiguities. Let $\mathbf{D}$ be a diagonal scaling matrix and $\mathbf{P}$ a permutation matrix. Then $\mathbf{A}\vec{\tilde{S}} = \mathbf{APD}\vec{S}$. Since independence is preserved, ICA can only recover sources up to: (1) unknown scaling factors and (2) unknown permutation of sources.

2.2. Frequency-Domain ICA Theory

Classical ICA assumes instantaneous linear mixtures, which is rarely valid in real acoustic environments. In microphone recordings, signals propagate with delays, reflections, and frequency-dependent attenuation, resulting in a convolutive mixing model.

Frequency-domain ICA (FD-ICA) provides an efficient framework by transforming the convolutive problem into a set of frequency-dependent instantaneous mixtures. In practice, Short-Time Fourier Transform (STFT) is applied to obtain time–frequency representations. For each frequency bin f, the model becomes $\mathbf{X_f} = \mathbf{A_f}\mathbf{S_f}$, where $\mathbf{A_f}$ is the frequency-dependent mixing matrix. This formulation enables applying ICA independently for each frequency bin.

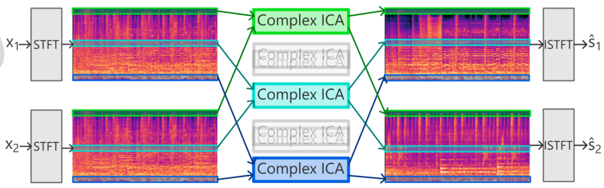

Figure 2: Schematic workflow of frequency domain ICA

The FD-ICA procedure is as follows (as shown in Fig. 2): (1) Apply STFT to microphone signals, (2) perform complex ICA independently for each frequency bin: $\hat{\mathbf{S}}_\mathbf{f} = \mathbf{W_f}\mathbf{X_f}$, (3) reconstruct separated signals in the frequency domain, and (4) Apply inverse STFT to obtain time-domain signals.

2.3. Complex ICA

Frequency-domain ICA requires ICA algorithms capable of handling complex-valued signals. Complex ICA has extensive applications in telecommunications, radar processing, and computer vision. Early methods such as FOBI [4] relied on fourth-order statistics, while JADE [5] improved general performance but scales poorly with increasing source numbers. Complex FastICA [6] offers fast convergence and practical implementations but assumes circular sources, which is not guaranteed in FD-ICA. In this work, we applied Adaptive Complex Maximization of Non-Gaussianity (ACMN) [7], a negentropy-based fixed-point method

Preliminary document

supporting both circular and noncircular sources, offering strong separation performance at higher computational cost.

The ACMN method assumes a whitened complex ICA model: $\vec{Z} = \mathbf{VA}\vec{S}$. The algorithm maximizes complex negentropy:

$$J(\vec{w}) = \mathbb{E}\left[\left|G(\vec{w}^{\mathrm{H}}\vec{Z})\right|^2\right], \tag{3}$$

leading to a fixed-point iterative algorithm for finding the optimal demixing vectors $\vec{w}$. Multiple sources are extracted by enforcing orthogonality of the demixing vectors using Gram–Schmidt projection.

2.4. Problems of Naive FD-ICA

Although FD-ICA combined with complex ICA can be implemented directly, naive FD-ICA typically yields poor separation performance. This is due to the inherent ICA ambiguities discussed earlier: scaling and permutation indeterminacies.

In FD-ICA, these ambiguities appear independently in each frequency bin. Scaling ambiguity introduces frequency-dependent amplitude and phase distortions, while permutation ambiguity causes frequency bins of different sources to be randomly reordered. As a result, inverse STFT recombines the separated spectra into mixed signals, therefore resolving scaling and permutation ambiguities is essential.

2.4.1. Scaling Ambiguity Resolution

A widely used solution is projection back to the observation space [2]. Assuming the demixing matrix $\mathbf{W} = (\hat{\mathbf{A}}\mathbf{\Lambda})^{-1}$, the estimated sources become $\vec{\hat{S}} \approx \mathbf{\Lambda}^{-1}\vec{S}$. The scaling can then be removed using the diagonal of $\mathbf{W}^{-1}$:

$$\hat{S}_{\mathrm{obs}} = \mathbf{W}_D\vec{\hat{S}} \approx \hat{\mathbf{A}}_D\vec{S}. \tag{4}$$

This projects the separated signals back to the microphone domain, yielding scaled but physically meaningful source estimates corresponding to microphone observations.

2.4.2. Permutation Ambiguity Resolution

Permutation ambiguity is typically solved using source or channel modeling. In this work, the Reduced Likelihood Ratio Jump (RLRJ) method [8] is applied, based on time-frequency source modeling.

Assuming complex Laplace-distributed STFT coefficients, the RLRJ method assigns a time varying energy estimate to each recovered source, and reorders their frequency bins iteratively such that the permutation likelihood is maximized. To reduce complexity, RLRJ restricts permutations to pairwise swaps, significantly lowering computational cost while preserving convergence. This approach provides an efficient and robust solution to permutation ambiguity in FD-ICA.

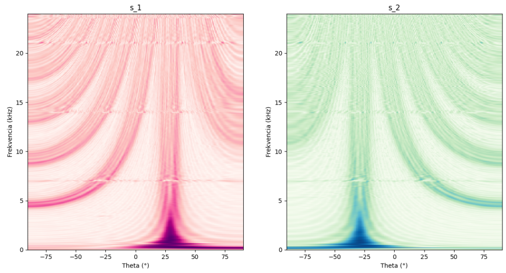

Figure 3: Magnitude of FD-ICA directivity patterns

3. Beamforming FD ICA

In this section we investigate how the performance of frequency domain ICA can be improved when two additional conditions, typical to microphone array applications, are met: (1) The number of microphones M is much higher than the number of independent sources S, and (2) the (relative) location of microphones is known.

3.1. Exploiting Multiple Microphones

When $M > S$, direct FD-ICA may degrade performance due to whitening of noise subspace components. This can be addressed using PCA-based dimensionality reduction.
For $\mathbf{X}_f \in \mathbb{C}^{M\times T}$, compute the covariance

$$\mathbf{\Sigma} = \mathbf{X_f}\mathbf{X_f}^H. \tag{5}$$

Selecting the S dominant eigenvectors yields the reduced whitening matrix

$$\mathbf{V^D} = \mathbf{\Lambda^S}^{\circ -\frac{1}{2}}\mathbf{Q^S}. \tag{6}$$

This projects signals onto the source subspace, improving SNR and separation performance. Scaling ambiguity can still be resolved using the pseudo-inverse of the demixing matrix.

3.2. Exlpoiting Known Microphone Positions

Known microphone geometry allows for acoustic beamforming and provides additional information for permutation resolution. In this section it will be shown that beamforming-based initialization improves convergence of RLRJ.
For a linear array and assuming far field (planar) source signals, the beamformer response is

$$F(f, \vec{w^f}, \theta) = \sum_{m=1}^{M} \tilde{w}_m^f \exp\left(\mathrm{i}\omega(m-1)d\frac{\sin(\theta)}{c}\right), \tag{7}$$

where d is microphone spacing and θ is the direction of arrival (DOA). The beamformer responses, as shown in Fig. 3 are used to assign a dominant frequency-based DOA to each mixing vector. The dominant DOA define an initial permutation of sources so that the frequency dependent variation of the direction of arrival of a source is minimized.

Preliminary document

3.3. BF-FD-ICA Architecture

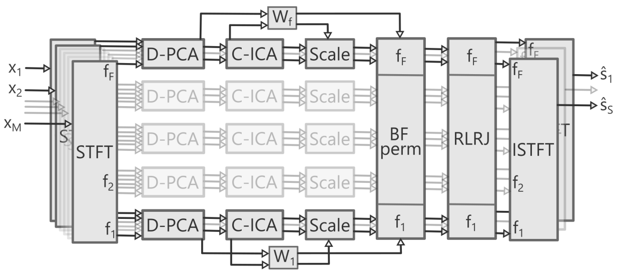

Figure 4: Bock diagram of BF-FD-ICA

By exploiting microphone array properties, the final beamforming-oriented FD-ICA architecture (BF-FD-ICA) is obtained by extending the method presented in Section 2.2. The high-level structure is shown in Fig. 4.
The BF-FD-ICA algorithm operates as follows:

1. Apply STFT to the input microphone signals $\vec{x}_m \in \mathbb{R}^N$ to obtain $\mathbf{X}_m$.
2. Construct frequency-bin matrices $\mathbf{X}_f \in \mathbb{C}^{M\times T}$.
3. Perform dimensionality-reduction PCA (D-PCA) to obtain whitened signals $\mathbf{Z}_f \in \mathbb{C}^{S\times T}$ and whitening matrices $\mathbf{V_f^D}$. The number of sources S can be estimated manually or using MUSIC.
4. Apply complex ICA (e.g., ACMN) to obtain separated signals $\hat{\mathbf{S}}_\mathbf{f}$ and demixing matrices $\mathbf{W_f^{white}}$.
5. Resolve scaling ambiguity using projection back to the observation space with $\mathbf{W_f} = \mathbf{W_f^{white}} \mathbf{V_f^D}$.
6. Estimate initial permutations using beamforming-based direction-of-arrival information, producing $\hat{\mathbf{S}}_\mathbf{f}^\mathbf{sb}$.
7. Refine permutations using the RLRJ algorithm, yielding $\hat{\mathbf{S}}_\mathbf{f}^\mathbf{sbp}$.
8. Reconstruct source STFT matrices $\hat{\mathbf{S}}_\mathbf{s}$.
9. Apply inverse STFT to obtain time-domain separated signals $\vec{\hat{s}}_s$.

4. Evaluation by Measurements

To obtain realistic performance estimates, the proposed method was evaluated using real measurements. Measurements were performed using a Gfai mcdRec system with a 24-channel linear microphone array. The microphones were placed on a line with spacing $d = 5\,\text{cm}$, resulting in an array length of $L = 1.15\,\text{m}$. The Nyquist aliasing limit of this array is $f_{\text{Nyquist}} = 3400\,\text{Hz}$. Regarding its spatial resolution, above $f > 2000\,\text{Hz}$ the $-3\,\text{dB}$ main lobe width is $6°$.

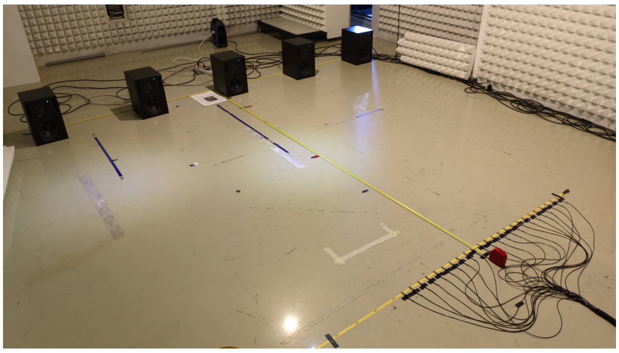

Figure 5: Measurement setup

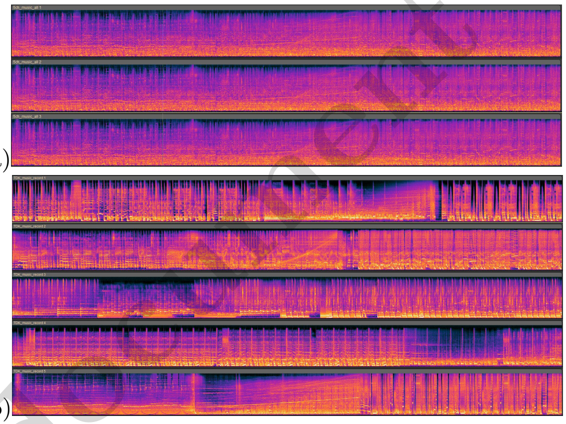
(a)
(b)

Figure 6: Spectrograms of (a) mixed signals and (b) the reference signals

Five Genelec 1030A loudspeakers were used as sources, arranged symmetrically in front of the array at distance $D = 0.7\,\text{m}$. The measurement was conducted in a semi-anechoic room with minimal reflections, as shown in Fig. 5. The resulting DOA vector was approximately $\vec{\theta} \approx [-30°, -16.1°, 0°, 16.1°, 30°]$.
A second measurement was performed in a reflective environment using a wooden panel placed behind the array, producing a strong reflection with delay $\tau \approx 0.012\,\text{s}$.
Complex electronic music excerpts were used as source signals. Reference signals were recorded by playing each source individually and recording it with a central microphone. Fig. 6 shows the spectrograms of the separate sources and their mixtures.
Subjective evaluation is possible by listening or inspecting spectrograms, but quantitative evaluation requires comparison with reference signals. Since time-domain error alone is insufficient, an STFT-based time–frequency SNR metric is proposed. The separated and reference signals were normalized using RMS values in a selected frequency band. The signals were synchronized using cross-correlation. The frequency domain RMS error was then computed as

$$err_i[f] = \sqrt{\sum_t S_i^{err}[f,t]^2} \tag{8}$$

Preliminary document

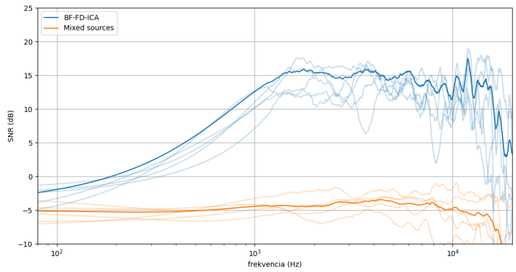

Figure 7: Separation performance of BF-FD-ICA

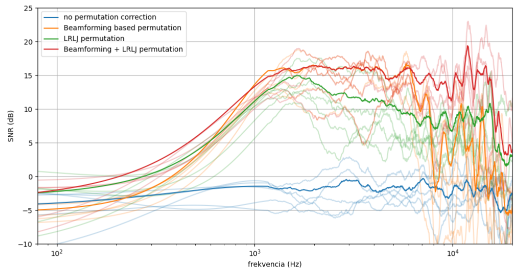

Figure 9: The effect of permutation correction algorithm on the separation performance

where the error spectrogram is defined as $\mathbf{S_i^{err}} = |\mathbf{S_i^{ref}} - \mathbf{S_i}|$ Finally, the frequeny domain SNR is defined as $SNR_i[f] = sig_i[f]/err_i[f]$, where

$$sig_i[f] = \sqrt{\sum_t S_i^{sig}[f,t]^2}. \tag{9}$$

This metric captures both time- and frequency-domain separation performance on real measurements.

4.1. Performance Evaluation of BF-FD-ICA

Unless otherwise stated, all 24 microphones were used in the non-reflective setup, with the number of sources manually set to $S = 5$. The STFT window length was $F = 4096$, and the ACMN algorithm was used for complex ICA. Separation performance was evaluated using the SNR metric defined in Section 4. Frequency-dependent SNR curves were smoothed using a Savitzky–Golay filter and plotted for individual sources and their average.

Using the parameters above, the BF-FD-ICA achieved approximately 15 dB average separation performance in the 1000–15000 Hz band (Fig. 7). This corresponds to roughly 20 dB improvement compared to the mixed signals (-4 to -6 dB SNR). Performance degrades at low frequencies due to small phase differences and near-singular mixing matrices, but remains acceptable (about 6 dB at 500 Hz).

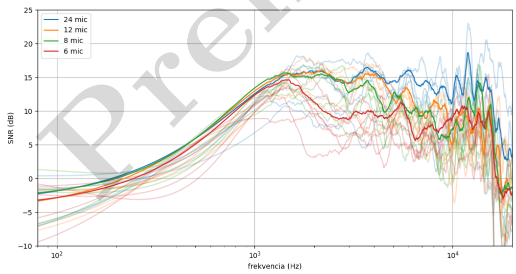

Figure 8: The effect of the number of microphones on separation performance

Fig. 8 shows performance degradation when reducing microphone count (24→12→8→6). The degradation is more significant at higher frequencies, likely due to reduced beamforming accuracy and weaker

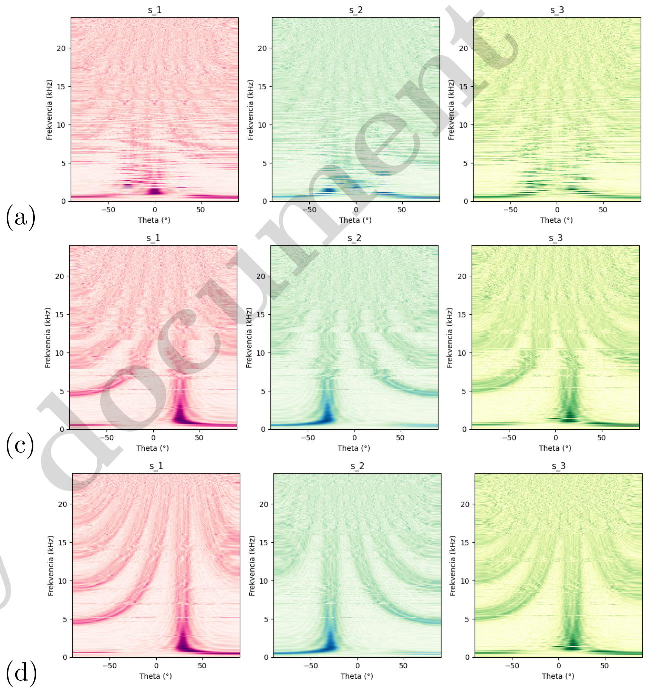

Figure 10: Recovered directivities (a) without permutation correction, (c) using BF-based permutation correction, (d) using RLRJ and BF permutation correction

dimensionality-reduction benefits.

Fig. 9 compares permutation correction methods. Without correction, separation fails. Pure beamforming-based permutation performs well at low frequencies but degrades at higher frequencies due to spatial aliasing. RLRJ performs consistently but slightly worse on average. Combining both methods provides the best overall performance. Directional patterns for each method are shown in Fig. 10.

Fig. 11 compares BF-FD-ICA (24 microphones, dimensionality reduction, beamforming+RLRJ) with classical FD-ICA (5 microphones, RLRJ). BF-FD-ICA improves performance by approximately 6–8 dB above 1500 Hz, confirming the effectiveness of the proposed extensions.

Performance was also evaluated in the reflective setup. Fig. 12 shows approximately 5–10 dB degradation due to longer impulse responses. Despite this, BF-FD-ICA

Preliminary document

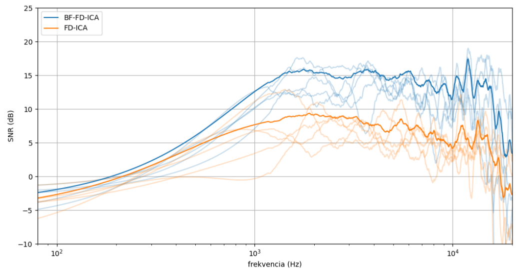

Figure 11: Comparing the performance of classical FD-ICA and BF-FD-ICA

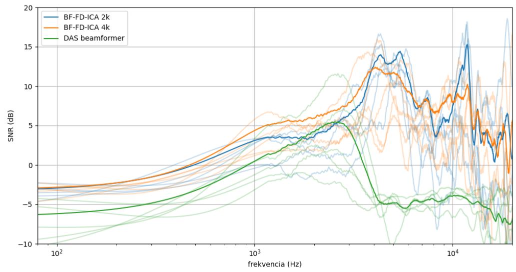

Figure 12: Comparing the performance of DAS and BF-FD-ICA for reflective environments

still achieved 10–15 dB separation above 1 kHz.
Finally, comparison with a classical Delay-And-Sum beamformers (Fig. 12) shows that DAS performance degrades with increasing frequency, while BF-FD-ICA remains superior.

5. Conclusion

This paper reviewed the theoretical background of ICA, FD-ICA, and complex ICA, and analyzed the limitations of naive FD-ICA implementations along with solutions proposed in the literature. Based on these, a practically applicable FD-ICA architecture was developed.
The additional information available in microphone array beamforming – namely the larger number of microphones than sources and the known microphone geometry – was then exploited. Two novel extensions were introduced: dimensionality-reduction FD-ICA and beamforming-based permutation correction. These were combined into the proposed BF-FD-ICA architecture optimized for microphone array environments.
Experimental results showed that the proposed method achieved approximately 20 dB separation performance in a wide 1500–15000 Hz frequency range, while maintaining acceptable performance even at lower frequencies (around 10 dB at 500 Hz). Compared to classical FD-ICA, the proposed BF-FD-ICA significantly improved performance above approximately 1000 Hz.
The BF-FD-ICA method was also compared with a classical beamforming algorithms (Delay-and-Sum). The proposed method significantly outperformed DAS in the measured environment.
Finally, performance was evaluated in a reflective environment. Although separation performance decreased compared to the near-anechoic case, BF-FD-ICA still outperformed conventional beamforming methods.

6. Acknowledgments

This work has been supported by the Hungarian National Research, Development and Innovation Office under contract No. K–143436.

References

[1] P. Comon, "Independent component analysis, a new concept?" *Signal Processing*, vol. 36, no. 3, pp. 287–314, 1994, Higher Order Statistics, ISSN: 0165-1684. DOI: `https://doi.org/10.1016/0165-1684(94)90029-9`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/0165168494900299`.

[2] N. Mitianoudis, "Audio source separation using independent component analysis," Ph.D. dissertation, University of London, 2004.

[3] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Jun. 2001, vol. 26, ISBN: 9780471405405. DOI: `10.1002/0471221317`.

[4] J.-F. Cardoso, "Source separation using higher order moments," in *International Conference on Acoustics, Speech, and Signal Processing,*, IEEE, 1989, pp. 2109–2112.

[5] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non-gaussian signals," in *IEE proceedings F (radar and signal processing)*, IET, vol. 140, 1993, pp. 362–370.

[6] E. Bingham and A. Hyvärinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals," *International journal of neural systems*, vol. 10, pp. 1–8, Mar. 2000. DOI: `10.1142/S0129065700000028`.

[7] M. Novey and T. Adali, "Adaptable nonlinearity for complex maximization of nongaussianity and a fixed-point algorithm," in *2006 16th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing*, IEEE, 2006, pp. 79–84.

[8] D. Mallis, T. Sgouros, and N. Mitianoudis, "Convolutive audio source separation using robust ica and an intelligent evolving permutation ambiguity solution," *Evolving Systems*, vol. 9, pp. 315–329, 2018.

Preliminary document