Performance Bound for Blind Extraction of Non-Gaussian Complex-Valued Vector Component from Gaussian Background

Václav Kautský, Zbyněk Koldovský, Petr Tichavský

Abstract

- We introduce Independent Vector Extraction (IVE), an approach for joint blind extraction of an independent vector component, the signal of interest (SOI), from K instantaneous mixtures.
- Similarly to Independent Component/Vector Analysis (ICA/IVA), the SOIs are assumed to be independent of the other signals in the mixture.
- The SOIs are assumed to be non-Gaussian or noncircular Gaussian, while the other signals are modeled as circular Gaussian.
- Cramér-Rao-Induced Bound (CRIB) for the achievable Interference-to-Signal Ratio (ISR) through IVE is derived and compared with similar bounds for ICA, IVA, and Independent Component Extraction (ICE).

Mixing model

- Linear mixture of d independent vector components which are formed from K scalar, possibly dependent but uncorrelated, sources

$$\mathbf{x}^k = \mathbf{A}^k\mathbf{u}^k, \tag{1}$$

for $k = 1, \dots, K$ and where $\mathbf{A}^k$ is a random mixing matrix.

- The mixing model could be written as

$$\mathbf{x}^k = \mathbf{A}^k\mathbf{u}^k = \mathbf{a}^k s^k + \mathbf{y}^k, \tag{2}$$

where s^k is the SOI in the kth mixture.

- The IVE mixing model is a generalization of the ICE model (When $K = 1$, ICE and IVE are the same model).
- Since $\mathbf{y}^k$ is not the object of extraction, we can assume $\mathbf{x}^k = \mathbf{A}^k_{\mathtt{ICE}}\mathbf{v}^k = [\mathbf{a}^k\mathbf{Q}^k]\mathbf{v}^k$, where $\mathbf{v}^k = [s^k; \mathbf{z}^k]$ and $\mathbf{Q}^k$ is such that $\mathbf{y}^k = \mathbf{Q}^k\mathbf{z}^k$, the choice of $\mathbf{Q}^k$ is based on the following steps.
- Parametrization of the demixing matrix for reduction of the ambiguity:

$$\mathbf{W}^k_{\mathtt{ICE}} = [\mathbf{w}^k; \mathbf{B}^k] \tag{3}$$

and denote $\mathbf{w}^k = [\beta^k; \mathbf{h}^k]$.

1. $\mathbf{B}^k$ is orthogonal to $\mathbf{a}^k = [\gamma^k; \mathbf{g}^k]$
   - straightforward selection is $\mathbf{B}^k = [\mathbf{g}^k \quad -\gamma^k\mathbf{I}_{d-1}]$
2. $\mathbf{W}^k_{\mathtt{ICE}}$ is the inverse of $\mathbf{A}^k_{\mathtt{ICE}}$
   - then $s^k = \mathbf{w}^{k^H}\mathbf{x}^k$
   - $\mathbf{A}^k_{\mathtt{ICE}} = [\mathbf{a}^k \quad \mathbf{Q}^k] = \begin{pmatrix} \gamma^k & \mathbf{h}^{k^H} \\ \mathbf{g}^k & \frac{1}{\gamma^k}\left(\mathbf{g}^k\mathbf{h}^{k^H} - \mathbf{I}_{d-1}\right) \end{pmatrix}$, where $\beta^k\gamma^k = 1 - \mathbf{h}^{k^H}\mathbf{g}^k$.

Signal model

- Random variables:
  - s^k (non-Gaussian), the target signal
  - $\mathbf{z}^k$ (multivariate Gaussian), background signals.
- The probability density function of $\mathbf{x}$ is

$$p(\mathbf{x}|\mathbf{a},\mathbf{w}) = p_{\mathbf{s}}(\{\mathbf{w}^{k^H}\mathbf{x}^k\}_{k=1}^K)p_{\mathbf{z}}(\{\mathbf{B}^k\mathbf{x}^k\}_{k=1}^K)\prod_{k=1}^K |\det \mathbf{W}^k_{\mathtt{ICE}}|^2 \tag{4}$$

where $\mathbf{w}^k$, $\mathbf{B}^k$ and $\mathbf{W}_{\mathtt{ICE}}{}^k$.

- Fix $\gamma^k = 1$ to avoid the scaling ambiguity and to reduce the number of parameters, then $|\det(\mathbf{W}_{\mathtt{ICE}}{}^k)| = 1$.
- The parameter vector is given by $[\mathbf{g}; \mathbf{h}]$, where $\mathbf{g} = [\mathbf{g}^1, \dots, \mathbf{g}^K]$ and $\mathbf{h} = [\mathbf{h}^1, \dots, \mathbf{h}^K]$.

Fisher Information Matrix

- Let $\boldsymbol{\theta}^k = [\mathbf{g}^k; \mathbf{h}^k]$ denote the parameter vector for the kth mixture, $\boldsymbol{\theta} = [\boldsymbol{\theta}^1; \dots; \boldsymbol{\theta}^K]$, and $\tilde{\boldsymbol{\theta}} = [\boldsymbol{\theta}; \boldsymbol{\theta}^*]$.
- For any unbiased estimator of $\tilde{\boldsymbol{\theta}}$, it holds that

$$\mathtt{cov}\left(\hat{\boldsymbol{\theta}}\right) \succeq \mathcal{J}^{-1}(\tilde{\boldsymbol{\theta}}) = \mathtt{CRLB}(\tilde{\boldsymbol{\theta}}), \tag{5}$$

where $\mathbf{C} \succeq \mathbf{D}$ means that $\mathbf{C} - \mathbf{D}$ is positive semi-definite, and $\mathcal{J}(\tilde{\boldsymbol{\theta}})$ is the Fisher information matrix (FIM) defined (in a block structure) as

$$\mathcal{J}(\tilde{\boldsymbol{\theta}}) = \begin{pmatrix} \mathbf{F} & \mathbf{P} \\ \mathbf{P}^* & \mathbf{F}^* \end{pmatrix} = \mathrm{E}\left[\frac{\partial\mathcal{L}}{\partial\tilde{\boldsymbol{\theta}}}\left(\frac{\partial\mathcal{L}}{\partial\tilde{\boldsymbol{\theta}}}\right)^H\right], \tag{6}$$

where $\mathcal{L} = \mathcal{L}(\tilde{\theta})$ is the log-likelihood function

$$\mathcal{L} = \log\left(p(\mathbf{x}|\mathbf{a},\mathbf{w})\right). \tag{7}$$

CRLB-Induced Bound for ISR

- Interference-to-Signal Ratio for the kth mixture in IVE is defined as

$$\mathtt{ISR}(\widehat{\mathbf{w}}^k) = \frac{(\widehat{\mathbf{w}}^k)^H\mathbf{C}^k_{\mathbf{y}}\widehat{\mathbf{w}}^k}{\sigma^2_{s^k}|(\widehat{\mathbf{w}}^k)^H\mathbf{a}^k|^2} = \frac{(\widehat{\mathbf{q}}^k_2)^H\mathbf{C}^k_{\mathbf{z}}\widehat{\mathbf{q}}^k_2}{|\widehat{q}^k_1|^2\sigma^2_{s^k}} \approx \frac{(\widehat{\mathbf{q}}^k_2)^H\mathbf{C}^k_{\mathbf{z}}\widehat{\mathbf{q}}^k_2}{\sigma^2_{s^k}}, \tag{8}$$

where $\sigma^2_{s^k}$ are the variances of the SOI, $\mathbf{C}^k_{\mathbf{y}} = \mathrm{E}[\mathbf{y}^k\mathbf{y}^k]^H$ and $(\widehat{\mathbf{q}}^k)^T = [\widehat{q}^k_1, (\widehat{\mathbf{q}}^k_2)^T] = (\widehat{\mathbf{w}}^k)^H\mathbf{A}^k_{\mathrm{ICE}} = \left[(\widehat{\mathbf{w}}^k)^H\mathbf{a}^k, (\widehat{\mathbf{w}}^k)^H\mathbf{Q}^k\right]$.

- Then, the mean ISR value reads

$$\mathrm{E}[\mathtt{ISR}(\widehat{\mathbf{w}}^k)] \approx \frac{\mathrm{E}\left[(\widehat{\mathbf{q}}^k_2)^H\mathbf{C}^k_{\mathbf{z}}\widehat{\mathbf{q}}^k_2\right]}{\sigma^2_{s^k}} = \frac{\mathtt{tr}\left(\mathbf{C}^k_{\mathbf{z}}\mathtt{cov}(\widehat{\mathbf{q}}^k_2)\right)}{\sigma^2_{s^k}}. \tag{9}$$

- Owing to the equivariance property of the BSE problem, we can consider the special case when $\mathbf{h} = \mathbf{0}$. Then, $\widehat{\mathbf{q}}^k_2 = \widehat{\mathbf{h}}^k$, and

$$\mathrm{E}[\mathtt{ISR}(\widehat{\mathbf{w}}^k)] \approx \frac{\mathtt{tr}\left(\mathbf{C}^k_{\mathbf{z}}\mathtt{cov}(\widehat{\mathbf{q}}^k_2)\right)}{\sigma^2_{s^k}} = \frac{\mathtt{tr}\left(\mathbf{C}^k_{\mathbf{z}}\mathtt{cov}(\widehat{\mathbf{h}}^k)\right)}{\sigma^2_{s^k}}, \tag{10}$$

then

$$\mathrm{E}[\mathtt{ISR}(\widehat{\mathbf{w}}^k)] \geq \sigma^{-2}_{s^k}\mathtt{tr}\left(\mathbf{C}^k_{\mathbf{z}}\mathtt{CRLB}(\mathbf{h}^k)\right). \tag{11}$$

- After computations and by considering N observations, the CRLB-induced bound for ISR for the kth mixture is

$$\mathrm{E}[\mathtt{ISR}_{\mathtt{IVE}}(\widehat{\mathbf{w}}^k)] \geq \frac{1}{N}\frac{d-1}{\kappa^k_{\mathrm{IVE}} - 1}, \tag{12}$$

where $\kappa^k_{\mathrm{IVE}} = \mathrm{E}\left[\left|\frac{\partial\log(p(\mathbf{s}))}{\partial s^k}\right|^2\right]$ where $p(\mathbf{s})$ is the joint pdf of $\mathbf{s} = s^1, \dots, s^K$ scaled to the unit variance.

Bounds for IVE, ICE, ICA, IVA

- Known bounds:

1. ICA (see [6, 5] for details):

$$\mathrm{E}[(\mathtt{ISR}_{\mathtt{ICA}})_{i,j}] \geq \frac{1}{N}\frac{\kappa_j}{\kappa_i\kappa_j - 1}, \tag{13}$$

where $\kappa_i = \mathrm{E}\left[\left|\frac{\partial\log(p_i(y_i))}{\partial y_i}\right|^2\right]$ where $p_i(y_i)$ is the pdf of the ith independent component scaled to the unit variance.

2. IVA (derived in [1]):

$$\mathrm{E}[(\mathtt{ISR}^k_{\mathtt{IVA}})_{i,j}] \geq \frac{1}{N}\frac{\kappa^k_j}{\kappa^k_i\kappa^k_j - 1}, \tag{14}$$

where $\kappa^k_i = \mathrm{E}\left[\left|\frac{\partial\log(p(\mathbf{y}_i))}{\partial y^k_i}\right|^2\right]$ where $p(\mathbf{y}_i)$ is the joint pdf of the ith vector component $\mathbf{y}_i = [y^1_i, \dots, y^K_i]$ scaled to the unit variance.

3. ICE (derived in [2]):

$$\mathrm{E}[\mathtt{ISR}_{\mathtt{ICE}}(\widehat{\mathbf{w}})] \geq \frac{1}{N}\frac{d-1}{\kappa_{\mathrm{ICE}} - 1}, \tag{15}$$

where $\kappa_{\mathrm{ICE}} = \mathrm{E}\left[\left|\frac{\partial\log(p(s))}{\partial s}\right|^2\right]$ and $p(s)$ is the pdf of the SOI s scaled to the unit variance.

- The following proposition shows that the dependence between signals from different mixtures can improve accuracy.

Proposition 1. *Let $p(s^1, \dots, s^K)$ denote the joint pdf of $s^1, \dots, s^K$, and $p_k(s^k)$ be the marginal pdf of s^k, $k = 1, \dots, K$. Then, $\kappa^k_{\mathrm{IVE}} \geq \kappa^k_{\mathrm{ICE}}$, and the equality when s^k is independent of the other random variables, or, equivalently, when $p(s^1, \dots, s^K) = p_k(s^k)p(s_1, \dots, s^{k-1}, s^{k+1}, \dots, s^K)$.*

- Comparison of CRIBs for $\mathrm{E}[\mathtt{ISR}(\widehat{\mathbf{w}}^k)]$:

	ICA	ICE	IVA	IVE
ICA	=	$\leq$	$\geq$	n/a
ICE	$\geq$	=	n/a	$\geq$
IVA	$\leq$	n/a	=	$\leq$
IVE	n/a	$\leq$	$\geq$	=

- The bound for ICA is lower than the bound for ICE (and IVA than IVE), since in ICA (and IVA) the background is not modeled as Gaussian.

Simulations

- Compare the bounds for ICE and IVE with empirical mean ISR achieved by the OGICE (Orthogonally Constrained ICE, see [4]) and by OGIVE, see [3], performing IVE.
- Both algorithms are properly initialized and the true score functions are used as the internal nonlinear function.
- For simplicity, only real-valued signals and mixing matrix are assumed.
- In one trial, $K = 3$ mixtures of $d = 5$ independent signals are generated: the background signals in mixtures are Gaussian with zero mean and unit variance, the SOIs (one SOI per mixture) are mutually dependent, drawn according to the joint pdf given by

$$p(s^1, \dots, s^K) \propto \exp\left(-\left(\lambda\sum_{i=1}^K |s^i|^2\right)^\alpha\right), \tag{16}$$

where $\lambda > 0$, and $\alpha \neq 1$ (for $\alpha = 1$, the pdf is Gaussian).

- All signals are mixed by a random mixing matrix.
- The following graph shows the comparison between IVE and ICE.

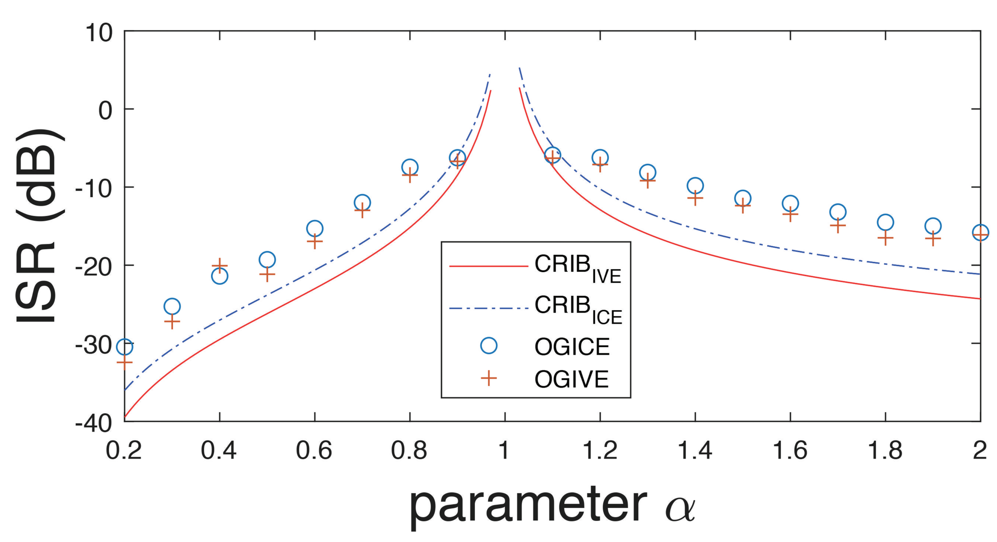

Fig. 1. CRIBs and average ISRs in 500 trials achieved by the compared algorithms for $d = 5$, $N = 5000$, $K = 3$.

Conclusions

- The CRIB on ISR achieved by IVE has shown that the structured (de-)mixing matrix model with a reduced number of parameters is not restrictive in terms of the achievable accuracy.
- The accuracy achievable by IVE is, in comparison to IVA, the same when the background is Gaussian.
- The dependence between the SOIs in the mixtures enable IVE to reach a better accuracy than ICE, which treats each mixture separately.

References

[1] M. Anderson, G. Fu, R. Phlypo, and T. Adalı. Independent vector analysis: Identification conditions and performance bounds. *IEEE Transactions on Signal Processing*, 62(17):4399–4410, Sept 2014.

[2] V. Kautský, Z. Koldovský, and P. Tichavský. Cramér-Rao-induced bound for interference-to-signal ratio achievable through non-gaussian independent component extraction. In *2017 IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 94–97, Dec 2017.

[3] Z. Koldovský and P. Tichavský. Gradient algorithms for complex non-gaussian independent component/vector extraction, question of convergence. *IEEE Transactions on Signal Processing*, 67(4):1050–1064, Feb 2019.

[4] Zbyněk Koldovský, Petr Tichavský, and Václav Kautský. Orthogonally constrained independent component extraction: Blind MPDR beamforming. In *Proceedings of European Signal Processing Conference*, pages 1195–1199, September 2017.

[5] Benedikt Loesch and Bin Yang. Cramér–Rao bound for circular and noncircular complex independent component analysis. In *IEEE Trans. Signal Processing*, volume 61, pages 365–379, Jan 2013.

[6] P. Tichavský, Z. Koldovský, and E. Oja. Performance analysis of the fastica algorithm and Cramér-Rao bounds for linear independent component analysis. *IEEE Transactions on Signal Processing*, 54(4):1189–1203, April 2006.

TECHNICAL UNIVERSITY OF LIBEREC www.tul.cz