

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Citation: J. Čmejla, Z. Koldovský, V. Kautský and T. Adali, "Dynamic Independent Component Extraction with Blending Mixing Vector: Lower Bound on Mean Interference-to-Signal Ratio," ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023, pp. 1-5, doi: 10.1109/ICASSP49357.2023.10096924.

DOI: <https://doi.org/10.1109/ICASSP49357.2023.10096924>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

DYNAMIC INDEPENDENT COMPONENT EXTRACTION WITH BLENDING MIXING VECTOR: LOWER BOUND ON MEAN INTERFERENCE-TO-SIGNAL RATIO

Jaroslav Čmejla¹, Zbyněk Koldovský¹, Václav Kautský¹, and Tülay Adalı²

¹Faculty of Mechatronics, Informatics, and Interdisciplinary Studies
Technical University of Liberec, Czech Republic

²Department of Electrical and Computer Engineering
University of Maryland, Baltimore County, Baltimore, MD, 21250, USA

ABSTRACT

This paper deals with dynamic Blind Source Extraction (BSE) from where the mixing parameters characterizing the position of a source of interest (SOI) are allowed to vary over time. We present a new source extraction model called CvxCSV which is a parameter-reduced modification of the recent Constant Separation Vector (CSV) mixing model. In CvxCSV, the mixing vector evolves as a convex combination of its initial and final values. We derive a lower bound on the achievable mean interference-to-signal ratio (ISR) based on the Cramér-Rao theory. The bound reveals advantageous properties of CvxCSV compared with CSV and compared with a sequential BSE based on independent component extraction (ICE). In particular, the achievable ISR by CvxCSV is lower than that by the previous approaches. Moreover, the model requires significantly weaker conditions for identifiability, even when the SOI is Gaussian.

Index Terms— Blind source extraction, independent component extraction, nonstationary mixing, dynamic models, moving sources

1. INTRODUCTION

Blind Source Separation/Extraction (BSS/BSE) deals with separating a mixture of signals or extracting a signal-of-interest (SOI) from the mixture without prior information or training. It is successfully realized using Independent Component Analysis/Extraction (ICA/ICE) [1]. For example, in audio source separation [2], blind techniques provide attractive alternatives to deep learning methods through direct interpretability [3]. This often happens in problems where the mixture variability is too high so that not all situations can be represented with training examples.

Even though limiting, so far, research has been focused almost exclusively on static mixing models. Recently, there has been a growing interest in dynamic models because real situations involve moving sources. This is typically solved by deploying methods in online processing mode [4]. A recent dynamic (semi-time-variant) mixing model for ICE and Independent Vector Extraction (IVE), called Constant Separating Vector (CSV) addresses this gap [5]. Here, the goal is to estimate a time-invariant beamformer that extracts the source from multiple locations covering the entire motion space [6]. This way, the parameters are interconnected, and their number is reduced. CSV makes it possible not only to achieve

higher accuracy than a fully time-varying model (time-variant mixing and de-mixing parameters) but also to avoid problems caused by the permutation uncertainty of the SOI [7]. The fully time-varying model can choose an arbitrary source as the SOI at each movement location, and thus the extracted signal can consist of concatenated intervals from multiple sources.

This paper proposes an effective modification of CSV with a significantly reduced number of parameters, which moderates the demands on the amount of data—lacking in online mode. In the original CSV, the source movement is captured by a time-varying mixing vector whose value is allowed to change from interval to interval. The proposed model enables a continuous or interval-to-interval change of the mixing vector where its current value is a convex combination of the initial and final values. Similar to CSV, the separating vector is assumed to be time-invariant, and hence, the proposed model is called Convex-CSV (CvxCSV). To our best knowledge, this powerful mixing model has not yet been considered in the context of BSE. A similar model was considered by Yeredor et al. in [8, 9]; however, it was in the context of BSS, where the entire mixing matrix changes over time, which generally excludes the existence of a constant separating vector.

In this paper, we also derive the Cramér-Rao-induced lower bound (CRIB) for the Interference-to-Signal Ratio (ISR). The bound reveals important features of the model: the minimum ISR achievable by any algorithm based on the model and the identification conditions. The result points to several advantageous properties over the original CSV. First, it is the effective use of the non-stationarity of the mixing model and of the non-stationarity of the SOI (so-called double non-stationarity [10]) causing the achievable ISR to be lower than the original CSV [11]. Second, it is the fact that the model is identifiable even when the SOI is circular Gaussian, which is not possible with CSV [11].

2. MODEL

2.1. Determined mixture with blending mixing vector

We consider a complex-valued mixture comprising the SOI and additional noise observed by d sensors. Let N observed samples be divided into T time intervals of the same length N_b , called blocks; in the case that $T = N$ and $N_b = 1$, the blocks correspond to the samples. The n th observed sample, $n = 1, \dots, N$ is described by

$$\mathbf{x}(n) = \mathbf{a}_t s(n) + \mathbf{y}(n), \quad (1)$$

where $s(n)$ and $\mathbf{y}(n)$ stand, respectively, for the SOI and noise; \mathbf{a}_t stands for the *mixing vector* of the SOI, and $t = \lceil \frac{nT}{N} \rceil$; the depen-

Part of this work was supported by The Czech Science Foundation through Project No. 20-17720S.

dence of t on n will not be explicitly written, for simplicity. Let \mathbf{a}_1 and \mathbf{a}_T be free parameter vectors. The key assumption in CvxCsv is that the SOI movement is a smooth blending of two positions corresponding to \mathbf{a}_1 and \mathbf{a}_T so that

$$\mathbf{a}_t = \lambda_t \mathbf{a}_1 + (1 - \lambda_t) \mathbf{a}_T, \quad (2)$$

where $\lambda_t \in [0, 1]$, $\lambda_1 = 1$, $\lambda_T = 0$. The case that $\lambda_t = \frac{t-1}{T-1}$ will be referred to as *linear*.

We now introduce constraints that enable us to write (1) in the form of a determined instantaneous mixture; the steps are similar to those in [7, 11, 12].

Let $\mathbf{B}_t = (\mathbf{g}_t - \gamma_t \mathbf{I}_{d-1})$, where γ_t and \mathbf{g}_t are the upper and lower parts of \mathbf{a}_t , i.e., $\mathbf{a}_t = \begin{pmatrix} \gamma_t \\ \mathbf{g}_t \end{pmatrix}$; \mathbf{I}_d is the $d \times d$ identity matrix. \mathbf{B}_t is a $(d-1) \times d$ blocking matrix since $\mathbf{B}_t \mathbf{a}_t = \mathbf{0}$. We can introduce $d-1$ background sources $\mathbf{z}(n) = \mathbf{B}_t \mathbf{x}(n) = \mathbf{B}_t \mathbf{y}(n)$ since they do not contain the SOI. By adopting the assumption of CSV that a *time-invariant* separating vector \mathbf{w} exists such that $s(n) = \mathbf{w}^H \mathbf{x}(n)$ for $n = 1, \dots, N$, \cdot^H stands for the conjugate transposition, the de-mixing model with the $d \times d$ de-mixing matrix \mathbf{W}_t reads

$$\begin{pmatrix} s(n) \\ \mathbf{z}(n) \end{pmatrix} = \mathbf{W}_t \mathbf{x}(n) = \begin{pmatrix} \mathbf{w}^H \\ \mathbf{B}_t \end{pmatrix} \mathbf{x}(n). \quad (3)$$

By inversion, the mixing model is obtained as

$$\mathbf{x}(n) = \mathbf{A}_t \begin{pmatrix} s(n) \\ \mathbf{z}(n) \end{pmatrix}, \quad (4)$$

where $\mathbf{A}_t = \mathbf{W}_t^{-1} = (\mathbf{a}_t \ \mathbf{Q}_t)$,

$$\mathbf{Q}_t = \begin{pmatrix} \mathbf{h}^H \\ \frac{1}{\gamma_t} (\mathbf{g}_t \mathbf{h}^H - \mathbf{I}_{d-1}) \end{pmatrix}, \quad (5)$$

and it must be satisfied that $\mathbf{w}^H \mathbf{a}_t = 1$ for every $t = 1, \dots, T$. By (2), this gives two conditions called *distortionless constraints*:

$$\mathbf{w}^H \mathbf{a}_1 = 1, \quad \mathbf{w}^H \mathbf{a}_T = 1. \quad (6)$$

The other constraint for the equivalence of (1) and (4) is that $\mathbf{y}(n) = \mathbf{Q}_t \mathbf{z}(n) = \mathbf{Q}_t \mathbf{B}_t \mathbf{x}(n) = (\mathbf{I}_d - \mathbf{a}_t \mathbf{w}^H) \mathbf{x}(n)$, which means that the noise subspace has dimension $d-1$. This restriction pays off for mathematical tractability, and its importance decreases as the number of sensors d increases.

2.2. Complexity of CvxCsv vs CSV

The free parameters in CvxCsv are the $d \times 1$ parameter vectors \mathbf{a}_1 , \mathbf{a}_T , and \mathbf{w} , which are linked through (6). Hence, CvxCsv involves $3d-2$ free parameters.

In CSV [7], each mixing vector \mathbf{a}_t , $t = 1, \dots, T$, is a free parameter vector linked with \mathbf{w} through $\mathbf{w}^H \mathbf{a}_t = 1$. Hence, CSV has $T(d-1) + d$ free parameters. CvxCsv and CSV coincide when $T = 2$, and both models correspond to the conventional (static) ICE model [12] when $T = 1$.

2.3. Statistical model

We consider a similar statistical model of signals as in [7, 10]. In particular, $s(n)$ and $\mathbf{z}(n)$ are assumed to be mutually independent; the samples of $s(n)$ as well as of $\mathbf{z}(n)$ are assumed to be independently distributed; let their pdfs depend on t and be denoted p_{s_t} and $p_{\mathbf{z}_t}$, respectively; let s_t and \mathbf{z}_t symbolize random variables having

the corresponding pdfs. By (3), the pdf of $\mathbf{x}(n)$, $n = 1, \dots, N$, is given by

$$p_{\mathbf{x}_t}(\mathbf{x}(n) | \mathbf{a}_1, \mathbf{a}_T, \mathbf{w}) = p_{s_t}(s(n)) p_{\mathbf{z}_t}(\mathbf{z}(n)) |\det(\mathbf{W}_t)|^2, \quad (7)$$

where $\det(\mathbf{W}_t) = (-1)^{d-1} \gamma_t^{d-2}$; see Eq. (15) in [12]. Hence, the log-likelihood function involving the distortionless constraint (6) reads

$$\begin{aligned} \mathcal{L}(\mathbf{x} | \mathbf{g}_1, \mathbf{g}_T, \mathbf{h}) = \sum_{n=1}^N \Big\{ & \log p_{s_t}(\mathbf{w}^H \mathbf{x}(n)) + \\ & + \log p_{\mathbf{z}_t}(\mathbf{B}_t \mathbf{x}(n)) + \\ & + (d-2) \log |1 - \mathbf{h}^H \mathbf{g}_t|^2 \Big\}, \end{aligned} \quad (8)$$

where \mathbf{x} symbolizes the entire observation $\mathbf{x}(1), \dots, \mathbf{x}(N)$.

3. CRAMÉR-RAO-INDUCED BOUND ON ISR

The Cramér-Rao lower bound (CRLB) determines a minimum variance achievable by unbiased estimators of deterministic parameters and corresponds to the asymptotic mean square error of maximum likelihood estimators. Similarly to [11], we use the CRLB to compute the lowest achievable value of the Interference-to-Signal Ratio (ISR) measured at the output of the BSE model proposed in this paper.

3.1. ISR

Let \mathbf{w} , $\hat{\mathbf{w}}$, and $\boldsymbol{\epsilon}$ be, respectively, the true separating vector, its estimate, and an error vector of the asymptotic order $o(1)$. We assume that $\hat{\mathbf{w}} = \mathbf{w} + \boldsymbol{\epsilon}$, thus, $\hat{\mathbf{w}}^H \mathbf{A}_t = (\mathbf{e}_1 + \mathbf{A}_t^H \boldsymbol{\epsilon})^H$, where \mathbf{e}_1 is the first column of \mathbf{I}_d . The estimated SOI then equals

$$\begin{aligned} \hat{s}(n) &= \hat{\mathbf{w}}^H \mathbf{x}(n) = \hat{\mathbf{w}}^H \mathbf{A}_t \begin{pmatrix} s(n) \\ \mathbf{z}(n) \end{pmatrix} = \\ &= (1 + \boldsymbol{\epsilon}) s(n) + \mathbf{q}^H \mathbf{z}(n), \end{aligned} \quad (9)$$

where $\mathbf{A}_t^H \boldsymbol{\epsilon} = [\boldsymbol{\epsilon}; \mathbf{q}]$. The ISR measured on the estimated SOI evaluated over all N samples is given by

$$\text{ISR} = \frac{\langle \mathbf{E} [|\mathbf{q}^H \mathbf{z}_t|^2] \rangle_t}{|1 + \boldsymbol{\epsilon}|^2 \langle \mathbf{E} [|s_t|^2] \rangle_t} = \frac{\text{tr} [\langle \mathbf{C}_{\mathbf{z}_t} \rangle_t \mathbf{q} \mathbf{q}^H]}{\langle \sigma_t^2 \rangle_t} + o(1), \quad (10)$$

where $\langle \cdot \rangle_t$ is the average value of the argument over $t = 1, \dots, T$, $\mathbf{C}_{\mathbf{z}_t} = \mathbf{E} [\mathbf{z}_t \mathbf{z}_t^H]$, and $\text{tr}[\cdot]$ denotes the trace of the argument. Hence, the approximate lower bound on the mean value of ISR is given by

$$\mathbf{E} [\text{ISR}] \geq \frac{\text{tr} [\langle \mathbf{C}_{\mathbf{z}_t} \rangle_t \text{cov}[\mathbf{q}]]}{\langle \sigma_t^2 \rangle_t}, \quad (11)$$

where $\text{cov}[\mathbf{q}] = \mathbf{E} [\mathbf{q} \mathbf{q}^H]$, and the Cramér-Rao-induced lower bound (CRIB) on ISR is obtained when $\text{cov}[\mathbf{q}]$ is replaced by the CRLB for the parameter vector \mathbf{q} . The important fact is that \mathbf{q} corresponds to the lower part of $\hat{\mathbf{w}}$, which is $\hat{\mathbf{h}}$ when the true separating and mixing vectors are all unit vectors, that is, $\mathbf{w} = \mathbf{a}_t = \mathbf{e}_1$. This has been referred to as *equivariance*, which significantly simplifies the whole analysis.

3.2. FIM

In the CRLB analysis for complex-valued parameters [13], for any unbiased estimator of a parameter vector $\boldsymbol{\theta}$, it holds that $\text{cov}[\boldsymbol{\theta}] \succeq \mathcal{J}^{-1}(\boldsymbol{\theta}) = \text{CRLB}(\boldsymbol{\theta})$, where $\mathcal{J}(\boldsymbol{\theta})$ is the Fisher Information Matrix (FIM), and $\mathbf{C} \succeq \mathbf{D}$ means that $\mathbf{C} - \mathbf{D}$ is a positive semi-definite matrix. The FIM can be partitioned as $\mathcal{J}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{F} & \mathbf{P} \\ \mathbf{P}^* & \mathbf{F}^* \end{pmatrix}$, where \cdot^* denotes the conjugate value,

$$\mathbf{F} = \mathbb{E} \left[\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}^*} \left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}^*} \right)^H \right], \quad \mathbf{P} = \mathbb{E} \left[\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}^*} \left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}^*} \right)^T \right], \quad (12)$$

where the derivatives are defined according to the Wirtinger calculus.

In our estimation problem, the only free parameters are the lower parts of \mathbf{a}_1 , \mathbf{a}_T , and \mathbf{w} , i.e., \mathbf{g}_1 , \mathbf{g}_T , and \mathbf{h} because of the constraints (6). The β in \mathbf{w} can be fixed to $\beta = 1$ because this removes a redundant parameter due to the scaling ambiguity in (1) (\mathbf{a}_t and $s(n)$ can be replaced by $a\mathbf{a}_t$ and $s(n)/a$ for any $a \neq 0$) and because the ISR is invariant to the scaling. Therefore, we define the whole parameter vector as $\boldsymbol{\theta} = [\mathbf{g}_1^T, \mathbf{g}_T^T, \mathbf{h}^T]^T$, and hence, \mathbf{F} and \mathbf{P} are square matrices of dimension $3(d-1)$.

3.3. Circular Gaussian background

From now on, we assume that the background signals $\mathbf{z}(n)$ are circular Gaussian with the covariance matrix $\mathbf{C}_{\mathbf{z}_t}$, so $\log p_{\mathbf{z}_t}(\mathbf{z}(n)) = -\mathbf{z}(n)^H \mathbf{C}_{\mathbf{z}_t}^{-1} \mathbf{z}(n) + \text{const}$. This special case allows a deeper analytical insight and gives an approximate idea of situations where the background is not Gaussian [14, 11]. By putting into (8), the log-likelihood function is equal to

$$\mathcal{L}(\mathbf{x}|\boldsymbol{\theta}) = \sum_{n=1}^N \log p_s(\mathbf{w}^H \mathbf{x}(n)) + \mathbf{z}(n)^H \mathbf{C}_{\mathbf{z}_t}^{-1} \mathbf{z}(n) + (d-2) \log |1 - \mathbf{h}^H \mathbf{g}_t|^2. \quad (13)$$

The next step is to compute the Wirtinger derivatives (12) and to evaluate them for $\mathbf{h} = \mathbf{g}_t = \mathbf{0}$.

$$\nabla^{\mathbf{h}} = \left. \frac{\partial \mathcal{L}}{\partial \mathbf{h}^*} \right|_{\substack{\mathbf{h}=\mathbf{0} \\ \mathbf{g}_t=\mathbf{0}}} = \sum_{n=1}^N \phi_t(s(n)) \mathbf{z}(n), \quad (14)$$

$$\nabla^{\mathbf{g}_1} = \left. \frac{\partial \mathcal{L}}{\partial \mathbf{g}_1^*} \right|_{\substack{\mathbf{h}=\mathbf{0} \\ \mathbf{g}_t=\mathbf{0}}} = - \sum_{n=1}^N (1 - \lambda_t) s(n)^* \mathbf{C}_{\mathbf{z}_t}^{-1} \mathbf{z}(n), \quad (15)$$

$$\nabla^{\mathbf{g}_T} = \left. \frac{\partial \mathcal{L}}{\partial \mathbf{g}_T^*} \right|_{\substack{\mathbf{h}=\mathbf{0} \\ \mathbf{g}_t=\mathbf{0}}} = - \sum_{n=1}^N \lambda_t s(n)^* \mathbf{C}_{\mathbf{z}_t}^{-1} \mathbf{z}(n), \quad (16)$$

where $\phi_t(s) = -\frac{\partial \log p_{s_t}(s, s^*)}{\partial s}$ is the score function of the SOI's pdf on the t th block. Since the samples are assumed to be independently distributed, the FIM has the structure $\mathcal{J}(\boldsymbol{\theta}) = N_b \sum_{t=1}^T \mathcal{J}_t(\boldsymbol{\theta})$ where

$$\mathcal{J}_t(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{F}_t & \mathbf{P}_t \\ \mathbf{P}_t^* & \mathbf{F}_t^* \end{pmatrix} \quad (17)$$

is the FIM corresponding to any sample within the t th block where the data are i.i.d. By using (14)-(16) for any n and $t = \lceil \frac{nT}{N} \rceil$,

$$\mathbf{F}_t = \begin{pmatrix} (1 - \lambda_t)^2 \mathbf{R}_t & (1 - \lambda_t) \lambda_t \mathbf{R}_t & -(1 - \lambda_t) \mathbf{I}_{d-1} \\ (1 - \lambda_t) \lambda_t \mathbf{R}_t & \lambda_t^2 \mathbf{R}_t & -\lambda_t \mathbf{I}_{d-1} \\ -(1 - \lambda_t) \mathbf{I}_{d-1} & -\lambda_t \mathbf{I}_{d-1} & \kappa_{s_t} \mathbf{C}_{\mathbf{z}_t} \end{pmatrix}, \quad (18)$$

where $\kappa_{s_t} = \mathbb{E}[|\phi_t(s(n))|^2]$, and $\mathbf{R}_t = \sigma_t^2 \mathbf{C}_{\mathbf{z}_t}^{-1}$. Owing to the assumed circularity of $\mathbf{z}(n)$, $\mathbf{P}_t = \mathbf{0}$, so $\mathcal{J}(\boldsymbol{\theta})$ is block-diagonal, and we can only focus on the computation of \mathbf{F}^{-1} to derive the CRLB. The structure of \mathbf{F} is

$$\mathbf{F} = N_b \sum_{t=1}^T \mathbf{F}_t = N_b \left(\begin{array}{c|c} \frac{A}{C} & \frac{B}{D} \end{array} \right) = N_b \left(\begin{array}{cc|c} \sum_{t=1}^T (1 - \lambda_t)^2 \mathbf{R}_t & \sum_{t=1}^T (1 - \lambda_t) \lambda_t \mathbf{R}_t & -\frac{T}{2} \mathbf{I}_{d-1} \\ \sum_{t=1}^T (1 - \lambda_t) \lambda_t \mathbf{R}_t & \sum_{t=1}^T \lambda_t^2 \mathbf{R}_t & -\frac{T}{2} \mathbf{I}_{d-1} \\ \hline -\frac{T}{2} \mathbf{I}_{d-1} & -\frac{T}{2} \mathbf{I}_{d-1} & \sum_{t=1}^T \kappa_{s_t} \mathbf{C}_{\mathbf{z}_t} \end{array} \right). \quad (19)$$

By the block-matrix inversion lemma,

$$\mathbf{F}^{-1} = \left[N_b \begin{pmatrix} A & B \\ C & D \end{pmatrix} \right]^{-1} = \frac{1}{N_b} \begin{pmatrix} I & J \\ K & L \end{pmatrix}, \quad (20)$$

and the CRLB for the lower part of \mathbf{w} , which is \mathbf{h} , corresponds to the lower corner of \mathbf{F}^{-1} , which is

$$\text{CRLB}(\mathbf{h})|_{\substack{\mathbf{h}_t=\mathbf{0} \\ \mathbf{g}_t=\mathbf{0}}} = \frac{1}{N_b} L = \frac{1}{N_b} (D - CA^{-1}B)^{-1}. \quad (21)$$

By putting (19) into (20) and then into (11), the desired lower bound on ISR is obtained. However, the closed-form expression cannot be derived in general. We now briefly discuss two special cases that enable deeper analytical insights.

First, as discussed in Section 2.2, the CvxCSV model coincides with CSV for $T = 2$ where the corresponding bounds must be the same. This is easily seen since when $\lambda_1 = 1$ and $\lambda_2 = 0$, the block A in (19) becomes block-diagonal, and the FIM corresponds with Equation (67) in [11].

Second, the closed-form bound on ISR can be expressed for the special case when \mathbf{R}_t is independent of t and when we have the linear blending case, i.e., $\lambda_t = \frac{t-1}{T-1}$. The computations are straightforward, however, we do not show them here due to limited space. The resulting bound for such a special case says that

$$\mathbb{E}[\text{ISR}] \gtrsim \frac{1}{N} \frac{d-1}{\langle \bar{\kappa}_{s_t} \rangle_t - 1}, \quad (22)$$

where $\bar{\kappa}_{s_t} = \kappa_{s_t} \sigma_{s_t}^2$ and it holds that $\bar{\kappa}_{s_t} \geq 1$ and $\bar{\kappa}_{s_t} = 1$ if and only if the SOI is circular Gaussian within the t th block. This result is similar to Equation (78) in [11], which corresponds to CSV when the SOI has the same distribution on all blocks. The result points to the important strength of both dynamic models: The lower bound for ISR is proportional to N^{-1} and not to N_b^{-1} . For CSV, the SOI must *not* be circular Gaussian in this special case, otherwise, the model is not identifiable (the bound is infinite). For CvxCSV, this condition is relaxed: for example, it suffices that the SOI is non-Gaussian within, at least, one block.

When the SOI is non-stationary, the identifiability conditions under CvxCSV are even more relaxed than under CSV, allowing the SOI to be Gaussian throughout the observation period. We demonstrate this fact by a numerical comparison of the bounds in the following section.

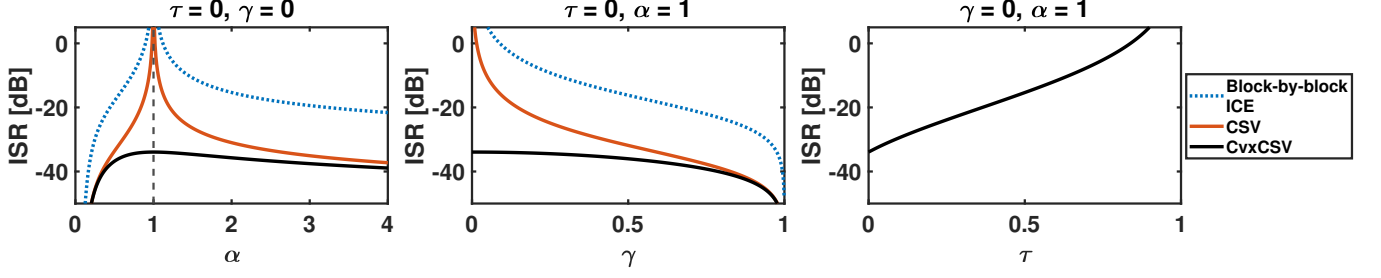


Fig. 1. The charts show the theoretically achievable ISR (CRIB) in terms of SOI properties (α , γ and τ parameters) and compare the output for block-by-block processing, the CSV model, and the linear variant of the proposed CvxCSV model.

4. NUMERICAL EVALUATION

In this section, we evaluate the CRIB for specific scenarios involving (non-)stationary, (non-)Gaussian, and (non-)circular SOI to analyze the influence of these signal features on the achievable extraction accuracy. We compare the results with similar CRIBs for the previous dynamic BSE models. Namely, we consider a naive blind approach where the conventional ICE is applied block-by-block to adapt to dynamic mixing; the approach is denoted as BICE in [11]. Next, the comparison includes the original CSV model from [7].

The CRIBs are evaluated for cases where the number of sensors is $d = 5$ and the number of observed samples is $N = 5000$. The data are assumed to be divided into $T = 10$ non-overlapping blocks of the same length, hence, $N_b = 500$. The SOI is assumed to be distributed according to the complex-valued Generalized Gaussian distribution (GGD) [15]. The pdf corresponding to a normalized GGD random variable (zero mean and unit variance) is given by [16]

$$p(s, s^*) = \frac{\alpha \rho \exp(-[\frac{\rho/2}{\gamma^2-1}(\gamma s^2 + \gamma(s^*)^2 - 2ss^*)])}{\pi \Gamma(1/\alpha)(1 - \gamma^2)^{\frac{1}{2}}}, \quad (23)$$

where $\rho = \frac{\Gamma(2/\alpha)}{\Gamma(1/\alpha)}$; $\Gamma(\cdot)$ denotes the Gamma function; $\alpha \in [0, \infty]$ is a shape parameter; $\gamma \in [0, 1]$ controls the level of (non-)circularity. Specifically, the distribution is super-Gaussian for $\alpha < 1$, Gaussian for $\alpha = 1$, and sub-gaussian for $\alpha > 1$; it is circular when $\gamma = 0$ and non-circular for $\gamma > 0$; it is degenerate when $\gamma = 1$. The corresponding score function is equal to

$$\phi(s, s^*) = \frac{2\alpha(\rho/2)^\alpha}{(\gamma^2 - 1)^\alpha} (\gamma s^2 + \gamma(s^*)^2 - 2ss^*)^{\alpha-1} (\gamma s - s^*). \quad (24)$$

By definition, for a SOI whose pdf on the t th block is GGD and its variance is σ_t^2 , it holds that

$$\bar{\kappa}_{st} = \kappa_{st} \sigma_t^2 = \frac{\alpha^2 \Gamma(2/\alpha)}{(1 - \gamma^2) \Gamma^2(1/\alpha)}. \quad (25)$$

To simulate the (non-)stationarity of the SOI, we consider its variance on the t th block to be given by

$$\sigma_{st} = \tau + (1 - \tau) \sin\left(\frac{\pi t}{2T}\right), \quad (26)$$

where $\tau \in [0, 1]$ controls the level of non-stationarity. For $\tau = 1$, the SOI is stationary (i.i.d.) throughout the observation period; otherwise, it is non-stationary. The background sources \mathbf{z}_t are assumed to be circular Gaussian, i.e., $\gamma = 0$, $\alpha = 1$, with identity covariance, i.e., $\mathbf{C}_{\mathbf{z}_t} = \mathbf{I}_{d-1}$.

The CRIBs for all three compared models are shown in Fig. 1. The left-hand chart of Fig. 1 shows the CRIB for circular non-stationary SOI ($\gamma = 0$, $\tau = 0$) in terms of varying α . It can be noticed that the SOI's non-stationarity ensures that the CvxCSV model is identifiable even in the case of circular Gaussian SOI ($\alpha = 1$). In contrast, the other models are unidentified (the CRIB grows to infinity). The second chart of Fig. 1 depicts the results for non-stationary Gaussian SOI ($\tau = 0$, $\alpha = 1$) in the case of varying circularity parameter. The CRIBs of all three models decrease as the SOI becomes more non-circular ($\gamma \rightarrow 1$). Finally, the right-hand chart in Fig. 1 shows the results for circular Gaussian SOI ($\gamma = 0$, $\alpha = 1$) when its level of (non-)stationarity is changing. In this case, the CSV and the block-by-block ICE models are not identifiable; therefore, the only displayed result is for CvxCSV. The corresponding CRIB decreases with increasing non-stationarity of the SOI ($\tau \rightarrow 0$), and vice versa. For $\tau = 1$, that is, when the SOI is stationary and circular Gaussian, the model is not identifiable.

5. CONCLUSION

We have proposed a new BSE model for extracting a SOI whose position given by the mixing vector changes in time. The change is approximated by a convex combination of the initial and final values of the vector, thereby reducing the number of free parameters. The analysis of the achievable accuracy in terms of mean ISR shows that with CvxCSV a higher extraction accuracy can be achieved than with the original CSV model or the naive sequential ICE. At the same time, the model enables the identification of a wider set of signals, especially Gaussian signals.

Future research will be focused on the development of efficient algorithms that achieve the optimum accuracy given by the bound, as has been achieved with the other BSE models [7, 11]. Algorithms based on the CvxCSV model have potential applications in situations where very little data is available. For example, this occurs when the SOI moves rapidly so that the change in its position must be estimated from short observations obtained through a correspondingly short time interval.

6. REFERENCES

- [1] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, ser. Independent Component Analysis and Applications Series. Elsevier Science, 2010.
- [2] E. Vincent, T. Virtanen, and S. Gannot, *Audio Source Separation and Speech Enhancement*, 1st ed. Wiley Publishing, 2018.

- [3] C. Boeddeker, F. Rautenberg, and R. Haeb-Umbach, "A comparison and combination of unsupervised blind source separation techniques," 2021. [Online]. Available: <https://arxiv.org/abs/2106.05627>
- [4] T. Nakashima and N. Ono, "Inverse-free online independent vector analysis with flexible iterative source steering," 2022. [Online]. Available: <https://arxiv.org/abs/2209.00937>
- [5] Z. Koldovský, J. Málek, and J. Janský, "Extraction of independent vector component from underdetermined mixtures through block-wise determined modeling," vol. 7903–7907, May 2019.
- [6] J. Janský, Z. Koldovský, J. Málek, T. Kounovský, and J. Čmejla, "Auxiliary function-based algorithm for blind extraction of a moving speaker," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2022, no. 1, p. 1, Jan 2022.
- [7] Z. Koldovský, V. Kautský, P. Tichavský, J. Čmejla, and J. Málek, "Dynamic independent component/vector analysis: Time-variant linear mixtures separable by time-invariant beamformers," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2158–2173, 2021.
- [8] A. Yeredor, "TV-SOBI: An expansion of sobi for linearly time-varying mixtures," in *Proceedings of The 4th International Symposium on Independent Component Analysis and Blind Source Separation (ICA2003)*, April 2003.
- [9] T. Weisman and A. Yeredor, "Separation of periodically time-varying mixtures using second-order statistics," in *Independent Component Analysis and Blind Signal Separation*, J. Rosca, D. Erdogmus, J. C. Principe, and S. Haykin, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 278–285.
- [10] Z. Koldovský, V. Kautský, and P. Tichavský, "Double non-stationarity: Blind extraction of independent nonstationary vector/component from nonstationary mixtures—Algorithms," *IEEE Transactions on Signal Processing*, vol. 70, pp. 5102–5116, 2022.
- [11] V. Kautský, Z. Koldovský, P. Tichavský, and V. Zarzoso, "Cramér–Rao bounds for complex-valued independent component extraction: Determined and piecewise determined mixing models," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5230–5243, 2020.
- [12] Z. Koldovský and P. Tichavský, "Gradient algorithms for complex non-Gaussian independent component/vector extraction, question of convergence," *IEEE Transactions on Signal Processing*, vol. 67, no. 4, pp. 1050–1064, Feb 2019.
- [13] T. Menni, E. Chaumette, P. Larzabal, and J. P. Barbot, "New results on deterministic Cramér–Rao bounds for real and complex parameters," in *IEEE Trans. Signal Processing*, vol. 60, March 2012, pp. 1032–1049.
- [14] T. Adalı and P. J. Schreier, "Optimization and estimation of complex-valued signals: Theory and applications in filtering and blind source separation," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 112–128, 2014.
- [15] M. Novey, T. Adalı, and A. Roy, "A complex generalized Gaussian distribution—characterization, generation, and estimation," *IEEE Trans. Signal Processing*, vol. 58, pp. 1427–1433, March 2010.
- [16] B. Loesch and B. Yang, "Cramér–Rao bound for circular and noncircular complex independent component analysis," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 365–379, 2013.