BLIND EXTRACTION OF MOVING SOURCES VIA INDEPENDENT COMPONENT AND VECTOR ANALYSIS: EXAMPLES

N. Amor[1]*, J. Čmejla*[1]*, V. Kautský*[2]*, Z. Koldovský*[1]*, and T. Kounovský*[1]

[1]Acoustic Signal Analysis and Processing Group, Technical University of Liberec, Czech Republic
[2] Faculty of Nuclear Sciences and Physical Engineering, CTU in Prague, Czech Republic

ABSTRACT

This paper is devoted to the recently proposed mixing model with constant separating vector (CSV) for Blind Source Extraction of moving sources using the FastDIVA algorithm, which is an extension of the famous FastICA and FastIVA for static mixtures. The benefits due to the CSV model and FastDIVA are demonstrated in three new applications. First, the extraction of a moving speaker in a noisy reverberant environment using a dense array of 48 MEMS microphones is considered. Second, a case study on the blind extraction of moving brain activity from visually evoked potentials in electroencephalogram is reported. Third, a simulation of block-by-block online extraction of a moving source is demonstrated. In these examples, the CSV and FastDIVA show their new potential and good performance in handling the blind moving source extraction problem.

Index Terms— Blind Source Extraction, Independent Component Analysis, Independent Vector Analysis, Speech Separation, Electroencephalogram, Moving Sources

1. INTRODUCTION

Blind Source Extraction and Separation (BSE/BSS) aim at the extraction of a desired signal or of all signals, respectively, from their observed mixture without using prior information. In the vast majority of methods, the static linear mixing model is considered, given by

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{1}$$

where $\mathbf{x}$ is a $d \times 1$ vector representing the mixed signals; $\mathbf{A}$ is a $d \times d$ non-singular time-invariant mixing matrix; and $\mathbf{s}$ is a $d \times 1$ vector representing the original signals. In BSE, only some columns of $\mathbf{A}$ and the corresponding elements of $\mathbf{s}$ are to be estimated. The determined case, i.e. where $\mathbf{A}$ is square and the dimensions of $\mathbf{x}$ and $\mathbf{s}$ are the same, is often preferred, because then the goal can be formulated as to estimate a square de-mixing matrix $\mathbf{W}$ (or some of its row in BSE). In Independent Component Analysis (ICA), where the main assumption is that the components in $\mathbf{s}$ are mutually independent, $\mathbf{W}$ is sought such that the signals $\mathbf{W}\mathbf{x}$ are as independent as possible [1].

Problems of time-varying mixtures have rarely been addressed by considering other mixing models than the fully parametrized time-variant version of (1); a linearly time-varying BSS mixing model was considered in [2]; see also [3, 4]. However, since many real-world problems are dynamic, there are various works considering sequential deployments of static mixture-based methods [5–9]. Recently, we have introduced a novel mixing model that allows for BSE of a moving source through the classical statistical estimation approach. The model, referred to as CSV (Constant Separating Vector) assumes time-varying mixing and time-invariant separating parameters in the determined model [10]. Some advantageous theoretical properties as well as practical usefulness of CSV have been reported in [11–13]. However, the practical potential of CSV has still not been fully explored.

This paper reports about new applications of CSV. In Sections 2 and 3, a brief summary of the model and of the recently announced one-unit FastDIVA algorithm is provided. FastDIVA is the successor of the famous FastICA/FastIVA [14, 15] allowing for the CSV mixing model [13]. Sections 4 is devoted to three examples of use: 1) BSE of a moving speaker using a dense microphone array, 2) BSE of a moving activity in visually evoked potentials, and 3) sequential BSE employing CSV. In these examples, CSV together with FastDIVA show promising results.

Notation: Upper index $\cdot^T$, $\cdot^H$, or $\cdot^*$ denotes, respectively, transposition, conjugate transpose, or complex conjugate. The Matlab convention for matrix/vector concatenation will be used, e.g., $[1;\, \mathbf{g}] = [1,\, \mathbf{g}^T]^T$.

2. CSV MIXING MODEL

In [10] it was shown that, for the static BSE problem, it is sufficient to parameterize $\mathbf{A}$ in (1) with two parameter vectors associated with the source of interest (SOI): the mixing and separating vector. When extracting the SOI jointly from a set of $K > 1$ mixtures via Independent Vector Extraction (IVE) [10, 16], the kth mixing matrix, $k = 1, \ldots, K$, is parameterized similarly by two k-indexed parameter vectors.

Let N samples of signals be available in K mixtures, and let the samples be divided into $T \geq 1$ time-intervals called blocks, for simplicity, of the same length N_b; hence $N = T \cdot N_b$. Consider a time-varying (block-wise) determined instantaneous mixing model given by

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

where $k = 1, \ldots, K$ is the mixture index; $t = 1, \ldots, T$ is the block index; $\mathbf{A}_{k,t}$, $\mathbf{s}_{k,t}$ and $\mathbf{x}_{k,t}$ represent the mixing matrix, the original signals and the observed mixed signals in the kth mixture and tth block, respectively. CSV is a special case of (2) where the mixing matrices are parameterized as

$$\mathbf{A}_{k,t} = \begin{pmatrix} \mathbf{a}_{k,t} & \mathbf{Q}_{k,t} \end{pmatrix} = \begin{pmatrix} \gamma_{k,t} & \mathbf{h}_k^H \\ \mathbf{g}_{k,t} & \frac{1}{\gamma_{k,t}}(\mathbf{g}_{k,t}\mathbf{h}_k^H - \mathbf{I}_{d-1}) \end{pmatrix}, \tag{3}$$

This work was supported by the United States Department of the Navy, Office of Naval Research Global, through Project No. N62909-19-1-2105, by The Czech Science Foundation through Project No. 20-17720S and by the Student Grant Competition of the Technical University of Liberec under the project No. SGS-2019-3060.

where $\mathbf{a}_{k,t} = [\gamma_{k,t}; \mathbf{g}_{k,t}]$ is the *mixing vector* and $\mathbf{w}_k = [\beta_k; \mathbf{h}_k]$ is the separating vector corresponding to the SOI in the kth mixture and tth block; the vectors are linked through $\mathbf{a}_{k,t}^H \mathbf{w}_k = 1$ for all t and k. The inverse matrix of $\mathbf{A}_{k,t}$ is equal to

$$\mathbf{W}_{k,t} = \begin{pmatrix} \mathbf{w}_k^H \\ \mathbf{B}_{k,t} \end{pmatrix} = \begin{pmatrix} \beta_k^* & \mathbf{h}_k^H \\ \mathbf{g}_{k,t} & -\gamma_{k,t}\mathbf{I}_{d-1} \end{pmatrix}, \tag{4}$$

where $\mathbf{I}_d$ is the $d \times d$ identity matrix. The structure of $\mathbf{s}_{k,t}$ is assumed to be $\mathbf{s}_{k,t} = [s_{k,t}; \mathbf{z}_{k,t}]$ where $s_{k,t}$ stands for the SOI while $\mathbf{z}_{k,t}$ represents the other background signals that are not subject of separation into independent signals like in ICA or IVA. Note that, for $T = 1$, CSV coincides with the static IVE model parametrized with two parameter vectors per each mixture [10].

While the mixing vectors, which are referring mainly to the position of the source, depend on t, the key property of CSV is that the separating vectors do not [13]. The idea is that each separating vector is estimated to extract the SOI from the entire area of its motion, which is delimited by the mixing vectors. Although the time-invariant separator need not always exist, it appears that such assumption can be reasonable in short/middle-term intervals [12]. The reduced parameterization gives the CSV an advantage over the models with time-variant separating vectors. To justify, it has been shown in [11] that the achievable residual interference-to-signal ratio (ISR) after the extraction based on CSV (for one data-set only, i.e., $K = 1$), is lower-bounded in the sense that

$$\mathrm{E}[\mathtt{ISR}] \geq \frac{1}{N} \frac{d-1}{\kappa_s - 1} \tag{5}$$

provided that the SOI has the same pdf (and variance) in all blocks and the background signals are circular Gaussian; $\kappa_s = \mathrm{E}[|\psi|^2]$ where ψ is the score function corresponding to the normalized pdf of the SOI. The bound in (5) has order $\mathcal{O}(N^{-1})$. This is in contrast to the conventional static mixture-based BSE where data are independently processed in each block, which ensures time-variant demixing on one hand, but on the other hand yields ISR whose lower bound is obviously $\mathcal{O}(N_b^{-1})$ [11]. Moreover, there is the ambiguity problem that a different source other than the SOI can be extracted in each block when they are processed independently. By contrast, in CSV, it is guaranteed that the same source is extracted in each block.

3. ONE-UNIT FASTDIVA

FastDIVA (Fast Dynamic IVA) is an algorithm that has been recently derived in [13]; this section provides its brief description. The one-unit FastDIVA performs BSE based on the CSV model. Starting from an initial guess, the algorithm performs updates of the separating vectors, $k = 1, \ldots, K$, according to

$$\mathbf{w}_k^{\mathrm{new}} = \mathbf{w}_k - \left\langle \left(\frac{\hat{\nu}_{k,t} - \hat{\rho}_{k,t}}{\hat{\nu}_{k,t}} \right)^* \frac{\widehat{\mathbf{C}}_{k,t}}{\hat{\sigma}_{k,t}^2} \right\rangle_t^{-1} \nabla^k. \tag{6}$$

where

$$\nabla^k = \left\langle \mathbf{a}_{k,t} - \frac{1}{\hat{\nu}_{k,t}} \hat{\mathrm{E}} \left[\phi_k \left(\left\{ \frac{\hat{s}_{k,t}}{\hat{\sigma}_{k,t}} \right\}_k \right) \frac{\mathbf{x}_{k,t}}{\hat{\sigma}_{k,t}} \right] \right\rangle_t, \tag{7}$$

$\hat{s}_{k,t} = \mathbf{w}_k^H \mathbf{x}_{k,t}$ is the current estimate of the SOI; $\hat{\sigma}_{k,t}^2$, $\hat{\nu}_{k,t}$, and $\hat{\rho}_{k,t}$ are equal, respectively, to the sample-based estimates of the variance, $\mathrm{E}[\frac{s_{k,t}}{\sigma_{k,t}} \phi_k]$, and $\mathrm{E}[\frac{\partial}{\partial s_k} \phi_k]$; ϕ_k is the score function corresponding to a selected model density of the SOI (the true density is not known in BSE); $\widehat{\mathbf{C}}_{k,t}$ is the sample-based covariance matrix of $\mathbf{x}_{k,t}$; $\mathbf{a}_{k,t}$ is linked to $\mathbf{w}_k$ through the orthogonal constraint, i.e., $\mathbf{a}_{k,t} = \widehat{\mathbf{C}}_{k,t} \mathbf{w}_k (\mathbf{w}_k^H \widehat{\mathbf{C}}_{k,t} \mathbf{w}_k)^{-1}$; $\langle \cdot \rangle_t$ denotes the average of the argument over $t = 1, \ldots, T$; $\{\cdot\}_k$ means that the argument involves all the variables dependent on k for $k = 1, \ldots, K$. After each update, $\mathbf{w}_k^{\mathrm{new}}$ is normalized so that the output scale of the extracted SOI is one. The algorithm stops when the change in $\mathbf{w}_k^{\mathrm{new}}$ is below a threshold for all k.

For $T = 1$ and $K = 1$, FastDIVA coincides with FastICA [14] in the real-valued case and is closely related to the complex-valued FastICA [17] and FastIVA [15] in the complex-valued case and $K > 1$, respectively. FastDIVA currently appears, to our best knowledge, to be the fastest method for the dynamic case $T > 1$. Performance analysis of the algorithm provided in [13] has shown that the asymptotic mean ISR achieved within the kth data set is equal to the Cramér-Rao-induced lower bound from [11] provided that the true score function of the SOI is used.

4. CASE STUDIES

4.1. Moving speaker extraction with dense microphone array

The usefulness of CSV in blind extraction of a moving speaker has already been proven in [12]. In the example here, we consider a special dense array consisting of 48 MEMS microphones arranged in an 8×6 vertical planar grid, see Fig. 1. The device captures sound at the sampling frequency of 25 kHz, which is down-sampled to 16kHz. All further processing is then done in the time-frequency domain obtained by a short-time Fourier transform with a frame length of 256 samples and 128 samples overlap.

Two speakers were recorded in a noisy office with the array. The SOI is a female speaker, which is performing a circular motion from 0° to 90° over the duration of 10 s at a 1 m distance from the array. The second speaker is a static male interferer standing at -90° at ~ 1.5 m distance. Both speakers were simulated using loudspeakers. White Gaussian noise was added so that the Interference-to-Noise ratio (INR) is -10 dB, and this was mixed with the SOI's signal at a Signal-to-Interference-plus-Noise ratio (SINR) of 5 dB. We chose this low INR level to verify the usefulness of the dense array (and algorithm) in strong isotropic noise, which cannot be attenuated in any way other then increasing the number of microphones.

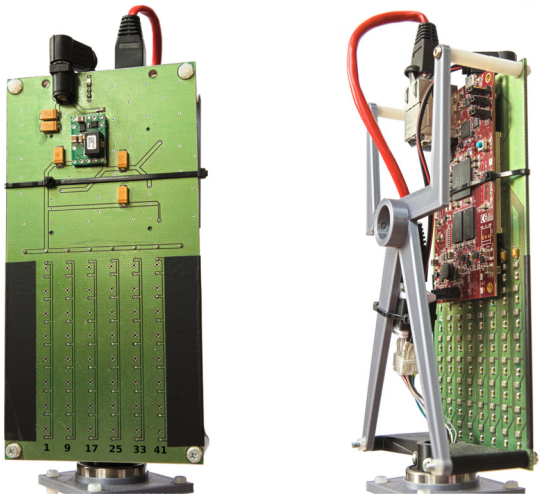


Fig. 1. Recording device with 48 working MEMS microphones arranged in a 5×7cm regular plane array.

One-unit FastDIVA was used to extract a component from the mixture. In the first extraction case, the entire signal was processed as a static mixture, i.e., with $T = 1$. In the second case, $T = 7$ blocks in the CSV model were considered.

Table 1 shows the results in terms of SINR improvement, SOI attenuation and Signal-to-Distortion ratio (SDR) when some number of microphones is used. The SDR is defined as

$$\mathrm{SDR} = \frac{\widehat{\mathrm{E}}[\tilde{s}^2]}{\min_\alpha \widehat{\mathrm{E}}[(\tilde{s} - \alpha s)^2]}, \tag{8}$$

where $\widehat{\mathrm{E}}$ is the sample-mean operator, $\tilde{s}$ is the contribution of SOI in the extracted signal $\widehat{s}$, and s is the true SOI signal. For the case $d = 2$, microphones 1 and 41 were used; for $d = 6$, the bottom-most row of microphones was used; for $d = 24$, every second row of microphones, starting at the bottom, was used. Increasing the number of available microphones provides a steady increase in SINR improvement. SDR, on the other hand, steadily decreases, since the methods using a lower number of microphones do not perform any significant signal filtering. In all cases, the approach using multiple blocks yields better extraction performance than the one with the static mixing model.

Table 1. SINR improvement, SDR and SOI attenuation [dB] for a different number of microphones.

#mics (d)	2		6		24		48	
#CSV blocks (T)	1	7	1	7	1	7	1	7
SINR$_{\mathrm{imp}}$	-0.30	1.68	3.06	4.90	3.32	6.28	7.59	8.25
SDR	14.86	15.84	11.77	14.13	7.84	11.47	7.22	11.91
SOI atn	-2.50	-0.94	-2.74	-1.19	-4.53	-1.47	-1.26	-0.52

Fig. 2 compares the resulting signals when 48 microphones were used. FastDIVA was successful in attenuating the interfering signal in both settings $T = 1$ and $T = 7$. With the static mixing model, however, it yielded a so-called "half-way" solution where the first half of the extracted SOI is severely attenuated. This is caused by the source movement. Using multiple blocks efficiently solves this issue together with an effective employment of all microphones to achieve the improved SINR.

4.2. Moving activity in visual evoked potentials

Visual Evoked Potentials (VEPs) are common measures of brain activity in the cognitive neuroscience community. VEPs are measured from Electroencephalograms (EEGs) that correspond to the electrical activity in the brain that occurs in response to a stimulus [18–20]. Using visual stimulation, various regions of cerebral cortex are activated, called *brain sources* [21]. Extracting these brain sources helps to accurately understand and analyze brain diseases.

ICA is a powerful tool for analyzing VEPs [22–24] as it is able to extract multiple functionally distinct sources. Extraction of these sources greatly enhances the awareness of VEPs by providing a cleaner and less ambiguous measure of source activity. So far, ICA does not handle moving EEG sources because the mixing model assumes that the independent components come from fixed positions. However, advanced neurological research has shown that brain sources are moving depending on various internal and external stimuli [25, 26]. Here, we show that One-unit FastDIVA is able to, at least, partly capture the moving activity in VEPs.

In the example here, we used the data containing motion-onset VEPs due to slowly moving objects with high contrast structures studied in [27]. These data were recorded using 8 channels (F3, F4, C3, C4, Pz, Oz, Ol, and Or); the sampling frequency was 250 Hz. The grand average of data of length 512 samples are subject to the BSS/BSE analysis.

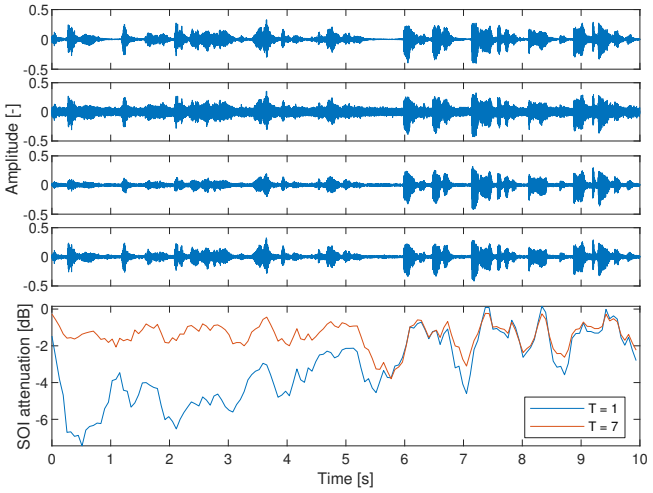


Fig. 2. Dynamic source extraction using one-unit FastDIVA. From the top to bottom: original SOI signal; mixture on the first microphone; extracted signal with $T = 1$; extracted signal using $T = 7$; attenuation of the SOI.

We first perform the classical ICA analysis using EFICA (Efficient FastICA) from [28]. This gives us components, which are in correspondence to those obtained in [27]. There are four independent components, denoted $s_1, \ldots, s_4$, shown in Fig.3(a) which contain the brain activity due to the stimuli; cf. Fig. 2 in [27]. These activities are localized, respectively, in the parieto-occipital region at 124 ms, occipito-parietal region at 164 ms, frontal region at 224 ms, and occipito-parietal region at 596 ms [27].

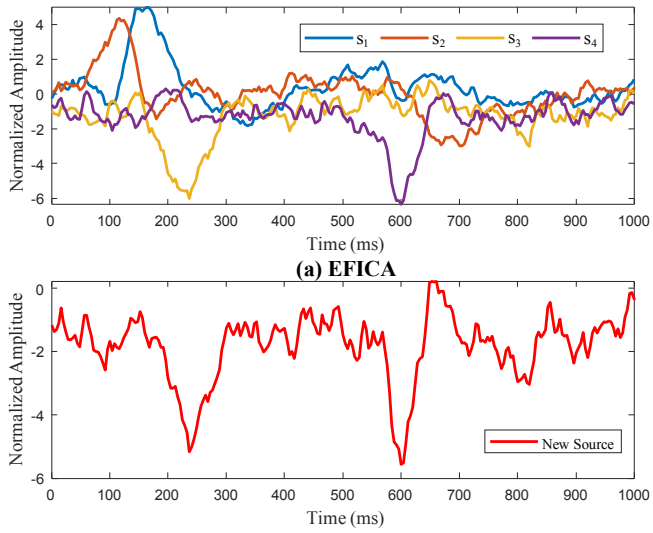
Since BSE algorithms (one-unit FastDIVA) can have different stationary points compared to BSS methods (EFICA), we verify whether the active components by EFICA are also the stationary points of one-unit FastDIVA. To ensure this "stability", we define new input data that correspond to the (already separated) active components $s_1, \ldots, s_4$. When one-unit FastDIVA is restarted from these solutions with $T = 1$, it stops after less than 10 iterations yielding the same components up to a small statistical deviation. In other words, the estimated mixing vectors corresponding to the extracted components are approximately equal to $\mathbf{e}_1, \ldots, \mathbf{e}_4$, up to scaling factors; $\mathbf{e}_i$ denotes the ith column of the 4×4 identity matrix. This confirms that both algorithms estimate the same components when their mixing models coincide.

An important result is obtained when one-unit FastDIVA is similarly applied with $T = 4$, which means that the data are divided into four blocks each having $N_b = 128$ samples in length. Fig.3(b) displays a new active source which is found by the algorithm when it is initialized in s_4. This source has two negative peaks at 224 ms and 596 ms, respectively, which corresponds with the activations of s_3 and s_4, respectively. The estimated mixing vectors of the source found on block 3 and 4 are equal to

$$\mathbf{a}_3 = \begin{pmatrix} -0.0331 \\ -0.1547 \\ 1.3695 \\ 0.0605 \end{pmatrix} \qquad \mathbf{a}_4 = \begin{pmatrix} 0.1189 \\ -0.0410 \\ 0.1661 \\ 1.1660 \end{pmatrix}, \tag{9}$$

which is, up to the scale, approximately equal to $\mathbf{e}_3$ and $\mathbf{e}_4$, respectively. It means that One-unit FastDIVA with $T = 4$ extracted the source that moved from the frontal region in block 3 to the occipito-parietal region in block 4.

A natural question is why the activity in s_1 and s_2 is not involved in the new component. We conjecture that this might be caused by the insufficient time resolution. The activations in s_1 and s_2 appear within the same block as that of s_3 and come from different regions; the movement is too fast. Unfortunately, increasing the time-resolution by dividing the data into more blocks does not help because the number of samples within one block becomes too small.



(a) EFICA

(b) One-unit FastDIVA

Fig. 3. (a) Active brain sources separated using EFICA. (b) A moving active brain source extracted using One-unit FastDIVA.

4.3. Simulated block-by-block online extraction

This example demonstrates the usability of CSV and FastDIVA in online BSE where data are processed sequentially block-by-block. Such processing is always a question of a correctly set length of blocks. On one hand, over-shortening the block length usually causes problems with convergence as the accuracy of the local parameter estimates is low. On the other hand, setting the block length too long can lead to an insufficient representation of highly dynamic mixtures involving fast-moving sources. The benefit of using CSV in online processing is that it allows for dynamics within the block by setting $T > 1$. This enables us to increase the block length without losing the time-resolution of the online BSE. We verify this feature in a simulated example where online BSE based on the static mixing model and CSV are compared.

In one trial, an instantaneous real-valued mixture ($K = 1$) of dimension $d = 10$ involving one Laplacean moving SOI, eight Laplacean interferers and additional white Gaussian noise is generated. The length of the mixture is $N = 60000$ samples; the INR is set to 50 dB and the input SINR is set to 5 dB over all sensors. The mixing vector of the SOI is continuously changing in a linear manner between $\mathbf{a}_1$ and $\mathbf{a}_N$, i.e., the mixing vector at the nth sample, $n = 1, \ldots, N$ is

$$\mathbf{a}_n = \left(1 - \frac{n-1}{N-1}\right)\mathbf{a}_1 + \left(\frac{n-1}{N-1}\right)\mathbf{a}_N. \qquad (10)$$

The interferers are static. The movement speed of the SOI is controled through the angular distance between $\mathbf{a}_1$ and $\mathbf{a}_N$. The separating vector is initialized by the LCMP beamformer steered in the directions given by $\mathbf{a}_1$ and $\mathbf{a}_N$. The data are then processed block-by-block (with overlap) by performing one FastDIVA iteration per block, initialized by the separating vector from the previous block.

We performed 750 repetitions of the experiment for various speeds of movement, block lengths and values of the CSV parameter T. The shift between neighbouring blocks is set to the block length divided by T (the sub-block length within CSV). Table 2 shows the resulting extraction accuracy in terms of average output SINR and SDR over all samples.

Table 2. Results of the simulated online BSE in terms of output SINR and SDR [dB].

Angle $\angle(\mathbf{a}_1, \mathbf{a}_N)$			0°		3°		10°		30°	
Block length	**Block shift**	**#CSV blocks** (T)	**SINR**	**SDR**	**SINR**	**SDR**	**SINR**	**SDR**	**SINR**	**SDR**
500	100	1	13.8	25.6	13.4	25.0	12.2	21.4	5.8	9.0
2500	500	1	21.7	32.9	18.6	27.2	9.0	12.0	8.7	4.4
5000	1000	1	24.8	36.2	14.9	19.6	10.0	6.3	14.0	2.4
500	100	5	13.6	25.5	13.2	24.9	13.1	24.1	13.1	19.9
2500	500	5	21.7	32.9	21.1	31.0	21.0	25.0	20.2	13.9
5000	1000	5	24.8	36.2	24.1	32.1	23.3	23.4	23.0	15.7

In the static case (angle 0°), both SINR and SDR obviously grow with the block length. Here, the number of CSV sub-blocks has no influence, i.e., CSV with $T > 1$ brings no advantage. In other cases, where the SOI is moving, the optimum block length depends on the movement speed, and a trade-off between the output SINR and SDR has to be sought. Here, the online processing with $T = 5$ brings significantly better SINR as well as SDR in all settings compared to $T = 1$, which confirms the advantage brought by CSV.

5. CONCLUSIONS

We presented examples where the CSV mixing model together with one-unit FastDIVA show their potential in blind extraction of a moving source. In particular, the approach solves the problem of "halfway" solutions, which appears when conventional static ICA or IVA is applied to mixtures involving moving sources. While ICA/IVA tends to split moving sources into several components, FastDIVA can find it as one independent component. We showed that the approach is useful in the multichannel speech enhancement problem where the speaker moves and is helpful in online block-by-block processing regime. It can also serve as a novel tool for exploratory data analysis, here demonstrated on BSE of moving brain activity in VEPs. This, however, needs further development and experimental validation. Also, an extensive experimental comparison of the CSV-based methods with the conventional sequential methods has to be done in future. The Matlab implementation of FastDIVA and of the examples presented in this paper are available on our website[1].

6. ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to professor Jan Kremláček from the Department of Medical Biophysics at the Faculty of Medicine in Hradec Králové, Charles University, Czech Republic, for providing us the VEPs data.

[1] https://asap.ite.tul.cz/downloads/ice/

7. REFERENCES

[1] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Independent Component Analysis and Applications Series, Elsevier Science, 2010.

[2] A. Yeredor, "TV-SOBI: An expansion of sobi for linearly time-varying mixtures," in *Proc. of The 4th International Symposium on Independent Component Analysis and Blind Source Separation (ICA 2003)*, April 2003.

[3] Y. Deville and A. Deville, "Blind quantum source separation and process tomography systems for time-varying coupling," in *2017 IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics (ECMSM)*, pp. 1–6, 2017.

[4] Y. Deville, A. Deville, S. Rebeyrol, and A. Mansour, "Analytical performance analysis for blind quantum source separation with time-varying coupling," in *2017 23rd Asia-Pacific Conference on Communications (APCC)*, pp. 1–6, 2017.

[5] F. Nesta, A. Brutti, and L. Cristoforetti, "Real-time prototype for multiple source tracking through generalized state coherence transform and particle filtering," in *2011 Joint Workshop on Hands-free Speech Communication and Microphone Arrays*, pp. 161–162, 2011.

[6] J. Chien and H. Hsieh, "Nonstationary source separation using sequential and variational bayesian learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 5, pp. 681–694, 2013.

[7] T. Taniguchi, N. Ono, A. Kawamura, and S. Sagayama, "An auxiliary-function approach to online independent vector analysis for real-time blind source separation," in *HSCMA 2014*, pp. 107–111, May 2014.

[8] A. H. Khan, M. Taseska, and E. A. P. Habets, *A Geometrically Constrained Independent Vector Analysis Algorithm for Online Source Extraction*, pp. 396–403. Cham: Springer International Publishing, 2015.

[9] S. H. Hsu, T. R. Mullen, T. P. Jung, and G. Cauwenberghs, "Real-time adaptive eeg source separation using online recursive independent component analysis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 3, pp. 309–319, 2016.

[10] 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, pp. 1050–1064, Feb 2019.

[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] J. Janský, Z. Koldovský, J. Málek, T. Kounovský, and J. Čmejla, "Auxiliary function-based algorithm for blind extraction of a moving speaker," 2020, arXiv: 2002.12619.

[13] Z. Koldovský, V. Kautský, and P. Tichavský, "Dynamic independent component/vector analysis: Time-variant linear mixtures separable by time-invariant beamformers," 2020, arXiv: 2007.11241.

[14] A. Hyvärinen, "Fast and robust fixed-point algorithm for independent component analysis," *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

[15] I. Lee, T. Kim, and T.-W. Lee, "Fast fixed-point independent vector analysis algorithms for convolutive blind source separation," *Signal Processing*, vol. 87, no. 8, pp. 1859–1871, 2007.

[16] T. Kim, H. T. Attias, S.-Y. Lee, and T.-W. Lee, "Blind source separation exploiting higher-order frequency dependencies," *IEEE Transactions on Audio, Speech, and Language Processing*, pp. 70–79, Jan. 2007.

[17] 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, Feb. 2000.

[18] M. Kuba, J. Kremláček, Z. Kubová, J. Szanyi, J. Langrová, and F. Vít, "Diagnostic substantiation and current possibilities of veps examination," *Clinical Neurophysiology*, vol. 129, no. 4, p. e8, 2018.

[19] X. Zhang, G. Xu, X. Mou, A. Ravi, M. Li, Y. Wang, and N. Jiang, "A convolutional neural network for the detection of asynchronous steady state motion visual evoked potential," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 6, pp. 1303–1311, 2019.

[20] Y. Lee, W. Lin, F. Cherng, and L. Ko, "A visual attention monitor based on steady-state visual evoked potential," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 3, pp. 399–408, 2016.

[21] G. Cruz, M. Miyakoshi, S. Makeig, K. Kilborn, and J. Evans, "Erps and their brain sources in perceptual and conceptual prospective memory tasks: Commonalities and differences between the two tasks," *Neuropsychologia*, vol. 91, pp. 173 – 185, 2016.

[22] L. Dong, X. Liu, L. Zhao, Y. Lai, D. Gong, T. Liu, and D. Yao, "A comparative study of different eeg reference choices for event-related potentials extracted by independent component analysis," *Frontiers in Neuroscience*, vol. 13, p. 1068, 2019.

[23] S. Hsu, L. Pion-Tonachini, J. Palmer, M. Miyakoshi, and S. Makeig, "Modeling brain dynamic state changes with adaptive mixture independent component analysis," *NeuroImage*, vol. 183, pp. 47–61, Feb 2018.

[24] P. Sheela and S. D. Puthankattil, "A hybrid method for artifact removal of visual evoked EEG," *Journal of Neuroscience Methods*, vol. 336, p. 108638, 2020.

[25] B. Ebinger, N. Bouaynaya, P. Georgieva, and L. Mihaylova, "Eeg dynamic source localization using marginalized particle filtering," in *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 454–457, 2015.

[26] P. Georgieva, N. Bouaynaya, L. Mihaylova, and F. Silva, "Bayesian approach for reconstruction of moving brain dipoles," in *Image Analysis and Recognition* (M. Kamel and A. Campilho, eds.), (Berlin, Heidelberg), pp. 565–572, Springer Berlin Heidelberg, 2013.

[27] J. Kremláček and M. Kuba, "Global brain dynamics of transient visual evoked potentials," *Physiological research*, vol. 48, no. 4, p. 303 — 308, 1999.

[28] Z. Koldovský, P. Tichavský, and E. Oja, "Efficient variant of algorithm FastICA for independent component analysis attaining the Cramér-Rao lower bound," *IEEE Transactions on Neural Networks*, vol. 17, pp. 1265–1277, Sept 2006.