2023 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)

# Moving Interference Speaker removal using Geometrically Constrained Independent Vector Analysis

Shinya Furunaga*, Tetsuya Ueda*, and Shoji Makino*
* Waseda University, 2-7 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0135, Japan

*Abstract*—**In this paper, we propose Multiple Geometrically Constrained Auxiliary function-based Independent Vector Analysis with Iterative Source Steering (MGC-AuxIVA-ISS), an offline geometrically constrained source separation method that removes the moving interference signal. GC-AuxIVA is a method that combines AuxIVA, a Blind Source Separation method, with geometrical constraints so that it achieves high separation performance while obtaining the target speech at the desired output channel. Online processing for GC-AuxIVA has also been proposed and can handle the movement of sound sources. However, it requires the exact positions of the target and interference sound sources for each time frame to make an accurate geometrical constraint. On the other hand, it is relatively easier to obtain a range of the moving speaker. Thus, we propose overdetermined offline-GC-AuxIVA to remove the range of the moving speaker by providing multiple geometrical constraints. We conducted simulation experiments to confirm that the proposed method can remove the interference signal by giving geometrical constraints at equal intervals in the range of the movement.**

## I. Introduction

In recent years, speech processing applications have become widespread. However, there is a problem in that the speech quality is decreased by diffuse noises and directional interferences. Therefore, there is a need for techniques that can extract the target speech from the recorded speech mixture. One of the useful techniques to enhance the target speech is Blind Source Separation (BSS), which estimates individual source signals from microphone-observed signals without such prior information as training data [1]. BSS methods for the determined condition (where the number of sound sources is equal to the number of microphones) include Independent Component Analysis (ICA) [2], which estimates a demixing matrix that maximizes the independence between source signals, Independent Vector Analysis (IVA) [3], [4], which solves both the source separation problem and the frequency permutation problem by modeling the entire frequency components as multivariate variables following spherical multivariate distribution, and Auxiliary function-based IVA (AuxIVA) [5], which is more stable and fast update rule using the auxiliary function method. However, they do not uniquely determine the output order of the separated signals, which we call global permutation problem.

Geometrically Constrained ICA (GC-ICA) and its extensions have been proposed to solve the global permutation problem simultaneously and the source separation problem [6], [7], [8], [9], [10], [11]. They can guide the demixing matrix to obtain a signal from a desired direction by exploiting spatial information such as the Direction of Arrival (DOA) of the signals and the microphone positions. Among them, Geometrically Constrained Auxiliary function-based IVA with Iterative Source Steering (GC-AuxIVA-ISS) [11] solves the global permutation problem and achieves better stability and faster update through an update algorithm using Iterative Source Steering (ISS) [12] without requiring inverse matrix operations. Its extension to online processing has also been developed to separate signals in real-time (which we call online-GC-AuxIVA-ISS [13]).

This paper focuses on the problem when applying GC-AuxIVA-ISS to moving source separation. One way to remove the moving interference signal is to estimate a time-variant demixing matrix that removes the interference signals according to the interference sound sources' position. There are several online source separation methods [8], [14], [15], and online-GC-AuxIVA-ISS [13] can estimate time-variant demixing matrix recursively. However, when the interference source moves quickly, it is difficult to remove the interference signal in real time because we obtain less spatial information about the interference signal than when the interference is fixed. Also, in online-GC-AuxIVA-ISS, we require time-variant spatial information for geometric constraints because the position of the interference source is changing. However, it is difficult to know the exact position of the interference source in a real environment. Although previous research has proposed a method that estimates spatial information from the current demixing matrix, preliminary experiments have confirmed that its source separation performance is limited when the sound source moves quickly.

Although it is difficult to obtain the positional information of the moving sound source at each time, it is relatively easy to obtain information on its range. Therefore, we propose a source separation method that removes the signal in the range of the interference source's movement. In concrete, we assume offline processing and an overdetermined condition where the number of microphones is more than the number of sources. Then we add several geometric constraints to guide the demixing matrix to remove the range of the interference source movement. We call this method Multiple GCAuxIVA-ISS (MGC-AuxIVA-ISS) and evaluate its effectiveness using

979-8-3503-0067-3/23/$31.00 ©2023 IEEE

a simulation experiment. As a result, it can be expected to remove the interference signal without giving information on how the interference sound moves at each time.

## II. Problem formulation

We consider a determined or an overdetermined situation where $J$ sound source signals are mixed and captured by $I$ ($\geq J$) microphones. The Short-Time Fourier Transform (STFT) coefficients of the $j$-th estimated source and $i$-th microphone signals are denoted by

$$\boldsymbol{y}_{fn} = [y_{1fn}, \ldots, y_{Jfn}]^{\mathsf{T}} \in \mathbb{C}^J, \tag{1}$$
$$\boldsymbol{x}_{fn} = [x_{1fn}, \ldots, x_{Ifn}]^{\mathsf{T}} \in \mathbb{C}^I. \tag{2}$$

Here, $f = 1, \ldots, F$ and $n = 1, \ldots, N$ are the indices of the frequency and frame, respectively and $(\cdot)^{\mathsf{T}}$ denotes the transpose.

We consider an instantaneous mixture model in the time-frequency domain where the STFT window length is sufficiently longer than the impulse response between the sound source and microphones. Then, the relationship between the observed signals $\boldsymbol{x}_{fn}$ and estimated sources $\boldsymbol{y}_{fn}$ can be expressed as

$$\boldsymbol{y}_{fn} = \boldsymbol{W}_{fn}\boldsymbol{x}_{fn}, \tag{3}$$

where, $\boldsymbol{W}_{fn} = [\boldsymbol{w}_{1fn}, \ldots, \boldsymbol{w}_{Jfn}]^{\mathsf{H}}$ is a time-variant demixing matrix containing demixing filters $\boldsymbol{w}_{jfn} \in \mathbb{C}^I$, and $(\cdot)^{\mathsf{H}}$ denotes the Hermitian transpose. Our goal is to estimate the source signal $\boldsymbol{y}_{fn} = [y_{1fn}, \ldots, y_{Jfn}]^{\mathsf{T}}$ and solve the global permutation problem.

## III. Conventional method

In this section, we introduce two conventional methods: Offline-GC-AuxIVA-ISS and Online-GC-AuxIVA-ISS. Then, we describe the drawback of the conventional methods in Section III-C.

### A. *Offline-GC-AuxIVA-ISS [11]*

Offline-GC-AuxIVA estimates an time-invariant demixing matrix $\boldsymbol{W}_f = [\boldsymbol{w}_{1f}, \ldots, \boldsymbol{w}_{Jf}]^{\mathsf{H}}$. The demixing matrix $\mathcal{W} = \{\boldsymbol{W}_f\}_f$ can be estimated by minimizing the following negative log-likelihood function:

$$\mathcal{L}_{\text{IVA}}(\mathcal{W}) = \sum_{j=1}^{I} \mathbb{E}[G(\boldsymbol{y}_{jn})] - \sum_{f=1}^{F} \log|\det \boldsymbol{W}_f|, \tag{4}$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator and $\boldsymbol{y}_{jn}$ is the vector representation of the estimated sources:

$$\boldsymbol{y}_{jn} = [y_{j1n}, \ldots, y_{jFn}]^{\mathsf{T}} \in \mathbb{C}^F. \tag{5}$$

Here, $G(\boldsymbol{y}_{jn}) = -\log p(\boldsymbol{y}_{jn})$ is the contrast function where $p(\boldsymbol{y}_{jn})$ represents a multivariate probability density function at the $j$-th source. One typical choice of the contrast function is to use a spherical contract function [3], [4], [5], which is expressed as

$$G(\boldsymbol{y}_{jn}) = G_R(r_{jn}), \tag{6}$$
$$r_{jn} = ||\boldsymbol{y}_{jn}||_2 = \sqrt{\sum_f |y_{jfn}|^2} = \sqrt{\sum_f |\boldsymbol{w}_{jf}^{\mathsf{H}}\boldsymbol{x}_{fn}|^2}. \tag{7}$$

Here, $G_R(r)$ is a function of a real-valued scalar variable $r$, and $||\cdot||_2$ denotes the $L_2$ norm of a vector. Adopting the auxiliary function approach [5], the following equation is optimized instead of (4):

$$\begin{aligned}\mathcal{L}_{\text{IVA}}(\mathcal{W}) &\leq \mathcal{L}_{\text{AuxIVA}}(\Sigma, \mathcal{W}) \\ &= \frac{1}{2}\sum_{f=1}^{F}\sum_{j=1}^{J} \boldsymbol{w}_{jf}^{\mathsf{H}}\boldsymbol{\Sigma}_{jf}\boldsymbol{w}_{jf} - \sum_{f=1}^{F}\log|\det \boldsymbol{W}_f|,\end{aligned} \tag{8}$$

where $\Sigma = \{\boldsymbol{\Sigma}_{jf}\}_{jf}$ is the weighted spatial covariance matrix calculated as

$$\boldsymbol{\Sigma}_{jf} = \sum_n \varphi(r_{jn})\boldsymbol{x}_{fn}\boldsymbol{x}_{fn}^{\mathsf{H}}. \tag{9}$$

Here, $\varphi(r_{jn}) = G_R'(r_{jn})/r_{jn}$ and $(\cdot)'$ denotes the derivative operator.

To avoid the global permutation problem, GC-AuxIVA-ISS uses geometric constraints [6] that controls the $j$-th demixing filter to respond a value $c_{j\theta}$ toward $\theta$ direction. The regularization term of the geometric constraint is expressed as

$$\mathcal{L}_{\text{GC}}(\mathcal{W}) = \sum_{f=1}^{F}\sum_{j=1}^{J}\sum_{\theta\in\Theta} \lambda_{j\theta}|\boldsymbol{w}_{jf}^{\mathsf{H}}\boldsymbol{d}_{f\theta} - c_{j\theta}|^2. \tag{10}$$

Here, $\Theta = \{\theta\}$ represents a set including all directions to be considered, $\boldsymbol{d}_{f\theta}$ is a steering vector pointing to the direction $\theta$, $c_{j\theta}$ is a nonnegative value set for all frequency bins as constraints, and $\lambda_{j\theta} \geq 0$ is a parameter that weighs the importance of the constraint. When $c_{j\theta} = 1$, the $j$th demixing filter $\boldsymbol{w}_{jf}$ is guided to preserve the signals arriving from the $\theta$ direction. On the other hand, when the value of $c_{j\theta}$ is close to 0, the $j$th demixing filter $\boldsymbol{w}_{jf}$ is forced to create a spatial null to the $\theta$ direction. Thus, the objective function to be optimized in GC-AuxIVA is given by combining (8) and (10):

$$\mathcal{L}(\Sigma, \mathcal{W}) = \mathcal{L}_{\text{AuxIVA}}(\Sigma, \mathcal{W}) + \mathcal{L}_{\text{GC}}(\mathcal{W}). \tag{11}$$

Because there is no closed update rule to minimize the objective function in (11), GC-AuxIVA-ISS uses update an iterative update algorithm, ISS [11], [12]. In ISS, the demixing matrix is updated using the auxiliary variable $\boldsymbol{v}_{jf} = [v_{1jf}, \ldots, v_{Ijf}]^{\mathsf{T}} \in \mathbb{C}^I$ as follows:

$$\boldsymbol{W}_f \leftarrow \boldsymbol{W}_f - \boldsymbol{v}_{jf}\boldsymbol{w}_{jf}^{\mathsf{H}}. \tag{12}$$

Substituting (12) into the objective function in (11), we have a new objective function to be minimized. Then, we can obtain the update rule of the variable $\boldsymbol{v}_{jf}$ by calculating the partial derivative of the function and equating it to zero. Since the detailed update rule has been discussed in [11] and is not directly related to our main discussion, we skip explaining it.

### B. Online-GC-AuxIVA-ISS [13]

Online-GC-AuxIVA-ISS uses time-variant demixing matrices $\boldsymbol{W}_{fn}$ and estimates the matrices in each time frame $n$.

In offline-GC-AuxIVA-ISS, spatial covariance matrix $\boldsymbol{\Sigma}_{jf}$ is calculated using all time frames of the observed signal. On the other hand in online-GC-AuxIVA-ISS, the spatial covariance matrix is updated in an autoregressive manner:

$$\boldsymbol{\Sigma}_{jfn} = \alpha \boldsymbol{\Sigma}_{jf(n-1)} + (1-\alpha)\varphi(r_{jn})\boldsymbol{x}_{fn}\boldsymbol{x}_{fn}^{\mathsf{H}}. \tag{13}$$

Here, $0 < \alpha < 1$ denotes a forgetting factor. The ISS update can be straightforwardly applied to online-GC-AuxIVA by replacing the update rule of the covariance matrices with (13). The detailed update rule of online-GC-AuxIVA-ISS is also written in [13].

We also need to consider how to model the geometrical constraints in online processing. In this paper, we use the following time-variant GC term for online-GC-AuxIVA-ISS:

$$\mathcal{L}_{\mathrm{GC}}(\boldsymbol{W}_{fn}) = \sum_{f=1}^{F}\sum_{j=1}^{J}\sum_{\theta_n \in \Theta_n} \lambda_{j\theta_n} |\boldsymbol{w}_{jfn}^{\mathsf{H}}\boldsymbol{d}_{f\theta_n} - c_{j\theta_n}|^2. \tag{14}$$

As discussed in the previous research [13], we obtain the steering vector $\boldsymbol{d}_{f\theta_n}$ in each frame $n$ using a DOA estimation method, multiple signal classification (MUSIC) [16]. MUSIC is a subspace-based method that decomposes the spatial covariance matrix of the observed multichannel signals to obtain subspaces of signals and noise that are orthogonal to each other. The operation of MUSIC in this system is described in [13], [17]. We used [18] for implementing the MUSIC method.

### C. Drawback of the conventional method in moving source separation

Let us consider a situation of the moving source separation where one interference source is moving as shown in Fig. 1. In this case, offline-GC-AuxIVA-ISS described in Section III-A estimates a time-invariant demixing filter using the signals of all time frames, and thus cannot follow changes in the environment such as the movement of sound source.

On the other hand, online-GC-AuxIVA-ISS estimates a time-variant demixing matrix using only the observed signals up to the present frame. Therefore, it can handle the movement of sound sources. However, when the interference source moves quickly, it is difficult to remove the interference signal in real time because we obtain less spatial information about the interference signal than when the interference is fixed. Moreover, it is not necessarily easy to prepare the accurate direction $\theta_n$ in each time frame $n$ to provide geometrical constraints. Although there is a research to estimate the DOAs for online-GC-AuxIVA-ISS, the accuracy of the DOA estimation is limited because the estimation is conducted by using the observed mixture signals up to the present.

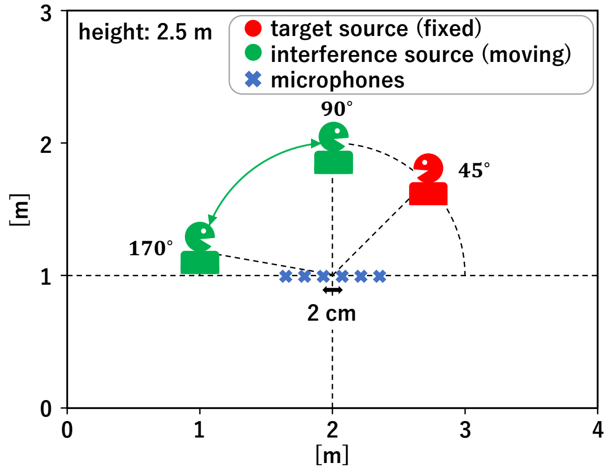



Fig. 1: Layout of a situation of the moving sourece separation. The target speaker is at $45°$. The interference speaker moves on an arc from $90°$ to $170°$ for the first 10 seconds and from $170°$ to $90°$ for the next 10 seconds.

## IV. Proposed method: MGC-AuxIVA-ISS

In this section, we propose a method to solve the issues described in Section III-C. Although it is difficult to get the positions of each source at each time, it is easier to obtain the range of the moving sound source. Thus, we propose to remove the range of the moving source by utilizing geometric constraints. In concrete, we extend offline-GC-AuxIVA-ISS to an overdetermined condition and provide null constraints in multiple directions to remove the moving interference signal. We call this proposed method Multiple GC-AuxIVA-ISS (MGC-AuxIVA-ISS). Here, we assume to obtain the range of the moving interference with a range $\phi$. For example in Fig. **??**, when the interference source move in the range in $90°$ to $170°$ ($\phi = \{\psi \mid 90° \leq \psi \leq 170°\}$) and we have 6 microphones, we set $\Theta = \{90, 110, \ldots, 170\}$, and apply the geometric constraint $c_{1\theta} = 0$ for all $\theta \in \Theta$. Thus, we expect that our MGC-AuxIVA-ISS can remove the entire moving range of the interference signal, regardless of the sound source's movement. This sets up $I-1$ null constraints equally spaced for $\phi$. In this study, we assume the direction range $\phi$ is known, and let its estimation be a future issue. The cost function for MGC-AuxIVA-ISS is equivalent to that of the Offline-GC-AuxIVA-ISS, and the parameter optimization algorithm is similar.

## V. Experimental Evaluation

### A. Experimental conditions

We conducted simulation experiments to evaluate the separation performance of MGC-AuxIVA-ISS in the situation of moving source separation. In this experiment, we used speech signals for 6 speakers (3 male and 3 female) selected randomly from a total of 503 sentences in Set B of the ATR Digital Japanese Speech Database [19]. To generate the mixture signals, we randomly selected speech signals from two different speakers and created a total of 50 patterns of mixture signals whose length is 20 seconds. We used `signal generator` [20] to create the mixture signals. We show the layout of sound sources and microphones using Fig. 1. The DOA of the target signal was fixed at $45°$ for all 20 seconds and the interference source moved on an arc from $90°$ to $170°$ for the first 10 seconds and moved on an arc from $170°$ to $90°$ for the next 10 seconds. We used 6 microphones and set the distance between adjacent microphones to 2 cm. The distance between sound sources and microphones was set to 1 m. We set the reverberation times ($RT_{60}$) to 200 ms. We conducted the convolution at a sampling frequency of 48 kHz and then resampled the convoluted signals at 16 kHz to separate the mixtures.

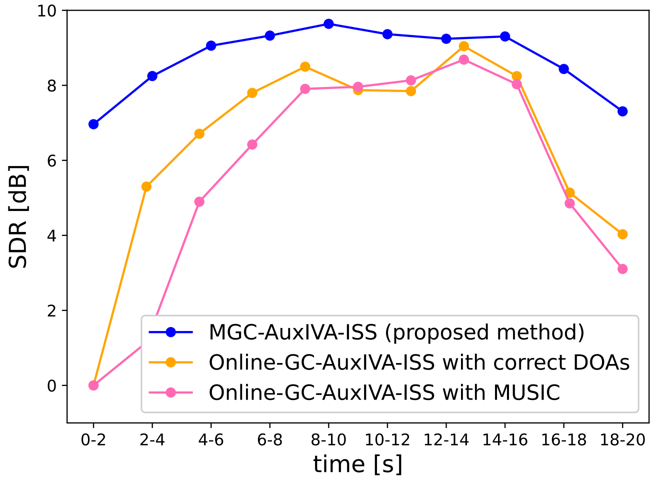
In this experiment, we used the following three methods, MGC-AuxIVA-ISS, online-GC-AuxIVA-ISS with correct DOAs and online-GC-AuxIVA-ISS with MUSIC. We used 6 microphones for MGC-AuxIVA-ISS, and 2 microphones for online-GC-AuxIVA-ISS. We gave information on the moving range of the interference source for MGC-AuxIVA-ISS. On the other hand, we gave the DOA of the interference signal at each time frame to online-GC-AuxIVA-ISS with correct DOAs. Online-GC-AuxIVA-ISS with MUSIC estimated the interference's DOA using MUSIC in each frame and applied the DOA into geometrical constraints. In this experiment, we used null constraints where $c_{j\theta} = 0$ and spatial nulls are formed in the direction of $\theta$. The STFT was computed using a Hanning window, whose length and shift were set at 2048 samples (128 ms) and 1024 samples (64 ms), respectively. The number of iterations is set to 50 for MGC-AuxIVA-ISS and 2 in each frame for online-GC-AuxIVA-ISS. In online-GC-AuxIVA-ISS, the forgetting factor $\alpha$ is set to 0.99, and the spatial covariance matrix $\mathbf{\Sigma}_{jf0}$ is initialized as an identity matrix.

The following two objective metrics were used to evaluate the separation performance: Source-to-Distortions Ratio (SDR) and Source-to-Interferences Ratio (SIR) [21]. Higher values of SDR and SIR indicate better separation performance. The values shown in the experimental results are obtained by tuning the parameter $\lambda_{j\theta}$ so that the accuracy of output signal order is 100% and SDR is the highest value.

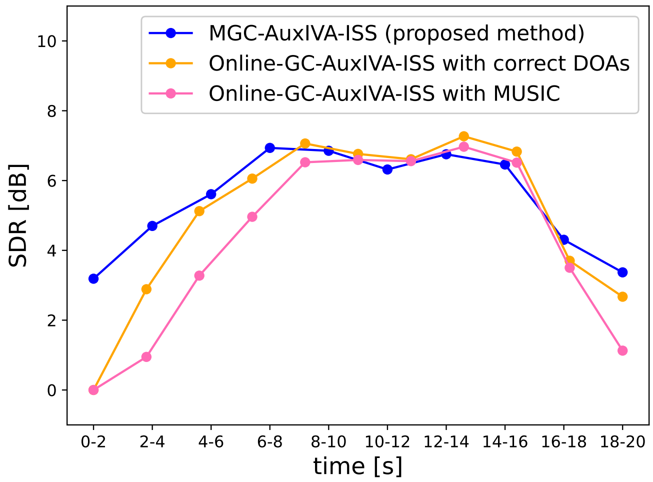
### *B. Results*

We compared the SDR and SIR improvement of MGC-AuxIVA-ISS and the conventional methods using Fig. 2 and Table I. Figure 2 shows the average SDR for the emphasized target signal every 2 seconds obtained by each method. Table I shows the average SDR and SIR of the target signal enhanced by each method in every 2 seconds. To evaluate the SDRs with and without background noise, we conducted two experiments, one without background noise and the other with a noise added by adjusting Source-to-Noise Ratio (SNR) = 25 dB.

We first show the result without background noise. In this experiment, we gave null constraints towards




(a) No background noise

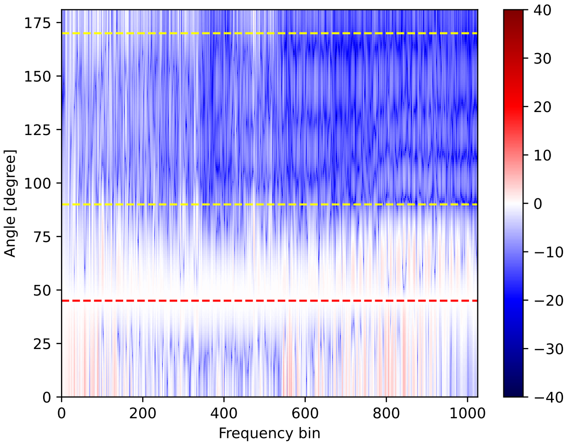



(b) SNR = 25 dB

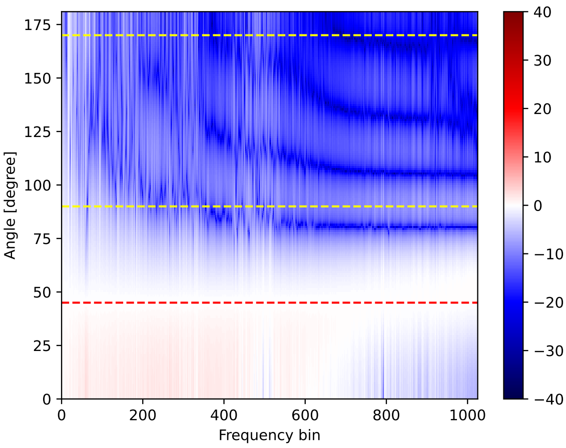
Fig. 2: Average SDR of the target signal enhanced by each method in every 2 s.

$\{90°, 110°, 130°, 150°, 170°\}$ and set $\lambda_{j\theta} = 100$ for MGC-AuxIVA-ISS. We set $\lambda_{j\theta} = 300$ for online-GC-AuxIVA-ISS with correct DOAs, and $\lambda_{j\theta} = 10$ for online-GC-AuxIVA-ISS with MUSIC. From Fig. 2 (a), MGC-AuxIVA-ISS shows the highest SDR overall times. Similarly, Table. I (a) shows that MGC-AuxIVA-ISS achieves the average SDR improvement of 1.64 dB and the average SIR improvement of 5.88 dB compared to online-GC-AuxIVA-ISS with correct DOAs. These results indicate that MGC-AuxIVA-ISS can remove the moving interference signal better than the conventional online-GC-AuxIVA-ISS.

Next, we show the result in a noisy environment. In this experiment, we gave null constraints towards $\{80°, 105°, 130°, 155°, 180°\}$ and set $\lambda_{j\theta} = 3590$ for MGC-AuxIVA-ISS. We set $\lambda_{j\theta} = 30$ for online-GC-AuxIVA-ISS with correct DOAs, and $\lambda_{j\theta} = 20$ for online-GC-AuxIVA-ISS




(a) No background noise




(b) SNR = 25 dB

Fig. 3: Examples of directivity patterns obtained by MGC-AuxIVA-ISS. Red line shows the position of the target source. Yellow lines show the range of the interference source.

with MUSIC. From Fig. 2 (b), MGC-AuxIVA-ISS shows the highest SDR before 4 seconds and after 16 seconds. However, the SDR of MGC-AuxIVA-ISS was close to that of online-GC-AuxIVA-ISS with correct DOAs between 4 to 16 seconds. Table I (b) also shows that MGC-AuxIVA-ISS performs equivalent SDR and higher SIR than online-GC-AuxIVA-ISS with correct DOAs. On the other hand, when compared to online-GC-AuxIVA-ISS with MUSIC, MGC-AuxIVA-ISS achieves the average SDR improvement of 0.77 dB. This result means that when we do not know the correct DOA but know the range of moving interference, using several null constraints toward the moving range is a good way, even in a noisy environment.

Finally, we show the directivity pattern using Fig. 3. The preferred result in Fig. 3 is that the demixing filter removes the steering vector arriving from the moving range of the

TABLE I: Average SDR [dB] and SIR [dB] of the target signal enhanced by each method. In these results, the accuracy of output order was 100%.

(a) No background noise

| method | SDR [dB] | SIR [dB] |
|---|---|---|
| MGC-AuxIVA-ISS (proposed method) | 8.69 | 16.01 |
| Online-GC-AuxIVA-ISS with correct DOAs | 7.05 | 10.13 |
| Online-GC-AuxIVA-ISS with MUSIC | 6.11 | 10.00 |

(b) SNR = 25 dB

| method | SDR [dB] | SIR [dB] |
|---|---|---|
| MGC-AuxIVA-ISS (proposed method) | 5.47 | 10.12 |
| Online-GC-AuxIVA-ISS with correct DOAs | 5.50 | 9.87 |
| Online-GC-AuxIVA-ISS with MUSIC | 4.70 | 9.69 |

interference source while remaining that from $45°$. So, we normalized the response in each frequency by the response from $45°$. This result shows that MGC-AuxIVA-ISS removes the moving range between $90°$ and $170°$.

## VI. Conclusions

In this paper, we proposed MGC-AixIVA-ISS to remove the moving interference speaker. Online-GC-AuxIVA-ISS, which is effective for the movement of the interference source, requires accurate DOAs of the interference signal for each time frame. On the other hand, MGC-AuxIVA-ISS extends offline-GC-AuxIVA-ISS to the overdetermined condition and provides equally spaced null constraints within the moving range of the interference source. To investigate the effectiveness of MGC-AuxIVA-ISS, we conducted speech enhancement experiments to remove the moving interference signal. As a result, we confirmed that the separation performance of MGC-AuxIVA-ISS is more than that of online-GC-AuxIVA-ISS.

## Acknowledgment


This work was supported by JSPS KAKENHI Grant Number 23H03423.


## References


[1] S. Makino, T.-W. Lee, and H. Sewada, *Blind Speech Separation*, Springer, 2007.

[2] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4-5, pp. 411–430, 2000.

[3] T. Kim, T. Eltoft, and T.-W. Lee, "Independent vector analysis: An extension of ICA to multivariate components," in *Proc. ICA*, 2006, pp. 165–172.

[4] A. Hiroe, "Solution of permutation problem in frequency domain ICA using multivariate probability density functions," in *Proc. ICA*, 2006, pp. 601–608.

[5] N. Ono, "Stable and fast update rules for independent vector analysis based on auxiliary function technique," in *Proc. WASPAA*, 2011, pp. 189–192.

[6] L. C. Parra and C. V. Alvino, "Geometric source separation: Merging convolutive source separation with geometric beamforming," *IEEE Trans. SAP*, vol. 10, no. 6, pp. 352–362, 2002.

[7] Y. Mitsui, N. Takamune, D. Kitamura, H. Saruwatari, Y. Takahashi, and K. Kondo, "Vectorwise coordinate descent algorithm for spatially regularized independent low-rank matrix analysis," in *Proc. ICASSP*, 2018, pp. 746–750.

[8] A. H. Khan, M. Taseska, and E. A. P. Habets, "A geometrically constrained independent vector analysis algorithm for online source extraction," in *Proc. LVA/ICA*, 2015, pp. 396–403.

[9] M. Knaak, S. Araki, and S. Makino, "Geometrically constrained independent component analysis," *IEEE Trans. ASLP*, vol. 15, no. 2, pp. 715–726, 2007.

[10] L. Li and K. Koishida, "Geometrically constrained independent vector analysis for directional speech enhancement," in *Proc. ICASSP*, 2020, pp. 846–850.

[11] K. Goto, T. Ueda, L. Li, T. Yamada, and S. Makino, "Geometrically constrained independent vector analysis with auxiliary function approach and iterative source steering," in *Proc. EUSIPICO*, 2022, pp. 757–761.

[12] R. Scheibler and N. Ono, "Fast and stable blind source separation with rank-1 updates," in *Proc. ICASSP*, 2020, pp. 236–240.

[13] K. Goto, T. Ueda, L. Li, T. Yamada, and S. Makino, "Accelerating online algorithm using geometrically constrained independent vector analysis with iterative source steering," in *Proc. APSIPA*, 2022, pp. 755–760.

[14] 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 *Proc. HSCMA*, 2014, pp. 107–111.

[15] L. Li, K. Koishida, and S. makino, "Online directional speech enhancement using geometrically constrained independent vector analysis," in *Proc. Interspeech*, 2020, pp. 61–65.

[16] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.

[17] K. Goto, T. Ueda, L. Li, T. Yamada, and S. Makino, "Investigation of partial update and source localization in geometrically constrained independent vector analysis with auxiliary function approach for moving source," in *Proc. NCSP*, 2023, pp. 426–429.

[18] R. Scheibler, E. Bezzam, and I. Dokmanic, "Pyroomacoustics: A python package for audio room simulations and array processing algorithms," in *Proc. ICASSP*, 2018, pp. 351–355.

[19] A. Kurematsu, K. Takeda, Y. Sagisaka, S. Katagiri, H. Kuwabara, and K. Shikano, "ATR Japanese speech database as a tool of speech recognition and synthesis," *Speech communication*, vol. 9, no. 4, pp. 357–363, 1990.

[20] "Signal generator," *https://www.audiolabs-erlangen.de/fau/professor/ habets/software/signalgene*.

[21] E. Vincent, R. Gribonval, and C. Févotte, "Performance measurement in blind audio source separation," *IEEE/ACM Trans. ASLP*, vol. 14, no. 4, pp. 1462–1469, 2006.