NUMERICAL INVESTIGATION OF WEIGHT PARAMETERS FOR GEOMETRICALLY CONSTRAINED INDEPENDENT VECTOR ANALYSIS USING VECTORWISE COORDINATE DESCENT OR ITERATIVE SOURCE STEERING

Shinya Furunaga[1], *Kana Goto*[2], *Tetsuya Ueda*[1],
Li Li[3], *Takeshi Yamada*[2], *Shoji Makino*[1]

[1]Waseda University, 2-7 Hibikino, Wakamatsu-ku, Kita-Kyushu, Fukuoka 808-0135, Japan
[2]University of Tsukuba, Japan
[3]NTT Communication Science Laboratories, NTT, Japan
shin12621@toki.waseda.jp, goto.kana.ry@alumni.tsukuba.ac.jp, t.ueda@akane.waseda.jp,
lili-0805@ieee.org, takeshi@cs.tsukuba.ac.jp, s.makino@waseda.jp

ABSTRACT

In this paper, we experimentally investigate the effect of weight parameters in the geometrically constrained independent vector analysis (GC-IVA) based on the auxiliary function approach, where the algorithms derived using the vectorwise coordinate descent (VCD), and iterative source steering (ISS) are referred as to GC-AuxIVA-VCD and GC-AuxIVA-ISS, respectively. GC-IVA aims to achieve both high source separation performance by blind source separation and the capability of directional focusing by beamforming technique. Although the previous studies have shown that separation performance is highly dependent on the parameters that weigh the importance of each geometric constraint, the parameter space having been investigated is limited, where the weight parameters are assumed to be the same for the constraints for both the target and interference signals. Furthermore, the lack of guidance for the parameter tuning process makes applying these algorithms difficult. To improve separation performance, we separately investigate the effect of weight parameters for the constraints for the target and interference signals with numerical experiments. Moreover, we present several tips for weight parameter tuning based on the experimental results, which are necessary to bring GC-IVA one step closer to practical applications. Experimental results showed that separately considering the weights for the constraints effectively improved the source-to-distortions ratio (SDR) and source-to-interferences ratio (SIR).

***Index Terms*—** Multichannel source separation, independent vector analysis, geometric constraints, vectorwise coordinate descent , iterative source steering

1. INTRODUCTION

Speech processing applications have recently become widespread. However, the appearance of diffuse noise and directional interferences degrade the performance, leading to an increasing need for techniques to extract target signals from recorded sound mixtures.

Blind source separation (BSS) methods for the determined cases (where the number of sources is equal to that of microphones) estimate demixing matrices to obtain individual signals from the multichannel observations based on the assumption that source signals are statistically independent with each other [1–6]. However, the output order of the separated signals is arbitrary and those methods require postprocessing to select the channel outputting the target signal. To solve this selection problem simultaneously with separation problem, geometrically constrained BSS (GC-BSS), including a variety of geometrically constrained independent vector analysis (GC-IVA) methods [7–12], exploit spatial information to guide the demixing matrices estimation. Since GC-BSS usually performs signal separation using spatial nulls estimated based on the statistical independence of the source signals, it can be applied with a small number of microphones without any training samples and can provide reasonably satisfactory performance. Among them, GC-IVA with auxiliary function approach [5, 6] and vectorwise coordinate descent (VCD) [13], referred as to GC-AuxIVA-VCD, is an algorithm noteworthy in high performance, fast convergence, and no requirement of step-size parameter tuning [10, 11]. Currently, GC-IVA with auxiliary function approach and iterative source steering (ISS), GC-AuxIVA-ISS for short, has been proposed [12] to reduce the computational complexity in GC-AuxIVA-VCD and stabilize the numerical computation by replacing VCD [13] with ISS [14]. GC-AuxIVA-VCD and GC-AuxIVA-ISS have been shown in previous studies to achieve high separation performance with signals output in a predefined order [10–12].

In [10–12], three types of geometric constraints have been adopted to GC-AuxIVA-VCD and GC-AuxIVA-ISS: *unit response* constraint, *null* constraint, and *double* constraint. *unit* constraint aims to force the demixing filter to preserve the target signal by returning unit response in the target direction. On the other hand, *null* constraint aims to create a null in the direction, which is defined by the direction of arrival (DOA) of the interference signal, to suppress the interference signal. *Double* constraint is a combination of *unit* and *null* constraints, aiming to suppress the interference signals while keeping the target signal distortionless. Previous research has shown that separation performance highly depends on the weight parameters for each constraint and the optimal weights are different for *unit* and *null* constraints [12]. However, the parameter space having been investigated in [12] for *double* constraints is limited, where the weight parameters are assumed to be the same for the constraints for both the target and interference signals. Furthermore, there are still no clear rules for appropriately setting weight parameters, leading these two algorithms challenging to apply to practical applications.

In this paper, to improve separation performance of GC-AuxIVA-VCD and GC-AuxIVA-ISS, we separately investigate the effect of weight parameters for constraints for the target and

978-1-6654-6867-1/22/$31.00

interference signals. We conclude several tips based on the numerical experiment results for weight parameter tuning, which are paramount for making GC-IVA applicable for practical applications.

2. GEOMETRICALLY CONSTRAINED INDEPENDENT VECTOR ANALYSIS

2.1. GC-AuxIVA-VCD [10]

We consider a determined situation, where I microphones capture J sound sources. Let y_{ifn} and x_{ifn} denote the short-time Fourier transform (STFT) coefficients of the j-th estimated source and i-th microphone signals, respectively. Here, $f = 1, \ldots, F$ and $n = 1, \ldots, N$ are the indices of the frequency and frame, respectively. The vector representation of the observations and the estimated sources are shown as

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

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

where $(\cdot)^{\mathsf{T}}$ denotes the transpose. When $I = J$ and assuming a time-invariant instantaneous mixture model where the STFT window length is sufficiently longer than the impulse response between the sound source and microphones, the relationship between the observed signals and estimated sources can be expressed as

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

$\boldsymbol{W}_f = [\boldsymbol{w}_{1f}, \ldots, \boldsymbol{w}_{Jf}]^{\mathsf{H}}$ is an $I \times I$ demixing matrix containing demixing filters $\boldsymbol{w}_{jf} = [w_{1jf}, \ldots, w_{Ijf}]^{\mathsf{T}}$, and $(\cdot)^{\mathsf{H}}$ denotes the Hermitian transpose.

IVA uses the dependency between frequency components to solve frequency-domain permutation alignment by assuming the signal vectors of all frequencies to be spherical multivariate distributions. We can estimate the demixing matrices $\mathcal{W} = \{\boldsymbol{W}_f\}_f$ by minimizing the following negative log-likelihood function:

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

where $\boldsymbol{y}_{jn} = [y_{j1n}, \ldots, y_{jFn}]^{\mathsf{T}} \in \mathbb{C}^F$ is the source-wise vector representation and $\mathbb{E}[\cdot]$ denotes the expectation operator. Here, $G(\boldsymbol{y}_{jn})$ is the contrast function having the relationship $G(\boldsymbol{y}_{jn}) = -\log p(\boldsymbol{y}_{jn})$ and $p(\boldsymbol{y}_{jn})$ represents a multivariate probability density function of the j-th source at n-th frame. Since IVA assumes the signal vectors of all frequencies to be spherical multivariate distributions, we can use the following contrast function

$$G(\boldsymbol{y}_{jn}) = G_R(r_{jn}), \tag{5}$$

$$r_{jn} = ||\boldsymbol{y}_{jn}||_2 = \sqrt{\sum_f |y_{jfn}|^2}. \tag{6}$$

$G_R(r)$ 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 [6], the following equation is optimized instead of (4),

$$\begin{aligned}\mathcal{L}_{\text{IVA}}(\mathcal{W}) &\leq \mathcal{L}_{\text{AuxIVA}}(\Sigma, \mathcal{W}) \\ &= \frac{1}{2}\sum_{f=1}^{F}\sum_{j=1}^{J} \boldsymbol{w}_{jf}^{\mathsf{H}} \boldsymbol{\Sigma}_{jf} \boldsymbol{w}_{jf} - \sum_{f=1}^{F} \log|\det \boldsymbol{W}_f|, \end{aligned} \tag{7}$$

where, the weighted covariance $\boldsymbol{\Sigma}_{jf}$ is expressed as

$$\boldsymbol{\Sigma}_{jf} = \mathbb{E}\Big[\frac{G_R'(r_{jn})}{r_{jn}} \boldsymbol{x}_f \boldsymbol{x}_f^{\mathsf{H}}\Big]. \tag{8}$$

Here, $(\cdot)'$ denotes the derivative operator.

Considering geometric constraints [15] that restrict the far-field response of filters estimated by IVA in a set of directions Θ. The regularization term of the geometric constraint is expressed as

$$\mathcal{L}_{\text{GC}}(\mathcal{W}) = \sum_{j=1}^{J}\sum_{\theta\in\Theta} \lambda_{j\theta} \sum_{f=1}^{F} |\boldsymbol{w}_{jf}^{\mathsf{H}} \boldsymbol{d}_{f\theta} - c_{j\theta}|^2. \tag{9}$$

Here, $\Theta = \{\theta\}$ represents a set including all directions to be considered, $\boldsymbol{d}_{f\theta}$ is a steering vector pointing to the direction θ, $c_{j\theta}$ is a non-negative 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 $c_{j\theta} = 1$, it forces the spatial filter to form a conventional delay and sum beamformer that steers in the direction θ to preserve the target source. On the other hand, a small value of $c_{j\theta}$ essentially creates a spatial null toward direction θ, so multiple constraints of spatial nulls toward all interference directions can be used to suppress all interference. The objective function of GC-AuxIVA-VCD [10] is given as

$$\mathcal{L}(\Sigma, \mathcal{W}) = \mathcal{L}_{\text{AuxIVA}}(\Sigma, \mathcal{W}) + \mathcal{L}_{\text{GC}}(\mathcal{W}). \tag{10}$$

The update rule for Σ is derived by applying (6) into (8). The update rule for $\mathcal{W}$ is obtain by using the idea adopted in VCD [13] that organizing the terms in $\log|\det \boldsymbol{W}|$ by the properties of the cofactor expansion. The update rules are given as

$$\boldsymbol{D}_{jf} = \boldsymbol{\Sigma}_{jf} + \sum_{\theta\in\Theta} \lambda_{j\theta} \boldsymbol{d}_{f\theta} \boldsymbol{d}_{f\theta}^{\mathsf{H}}, \tag{11}$$

$$\boldsymbol{u}_{jf} = \boldsymbol{D}_{jf}^{-1} \boldsymbol{W}_f^{-1} \boldsymbol{e}_j, \tag{12}$$

$$\hat{\boldsymbol{u}}_{jf} = \boldsymbol{D}_{jf}^{-1} \sum_{\theta\in\Theta} \lambda_{j\theta} c_{j\theta} \boldsymbol{d}_{f\theta}, \tag{13}$$

$$h_{jf} = \boldsymbol{u}_{jf}^{\mathsf{H}} \boldsymbol{D}_{jf} \boldsymbol{u}_{jf}, \tag{14}$$

$$\hat{h}_{jf} = \boldsymbol{u}_{jf}^{\mathsf{H}} \boldsymbol{D}_{jf} \hat{\boldsymbol{u}}_{jf}, \tag{15}$$

$$\boldsymbol{w}_{jf} = \begin{cases} \frac{1}{\sqrt{h_{jf}}} \boldsymbol{u}_{jf} + \hat{\boldsymbol{u}}_{jf} \ \ (\text{if } \hat{h}_{jf} = 0), \\ \frac{\hat{h}_{jf}}{2h_{jf}} \Big[-1 + \sqrt{1 + \frac{4h_{jf}}{|\hat{h}_{jf}|^2}} \Big] \boldsymbol{u}_{jf} + \hat{\boldsymbol{u}}_{jf} \ \ (\text{o.w.}). \end{cases} \tag{16}$$

Here, $\boldsymbol{e}_j$ in (12) denotes the j-th column of the $I \times I$ identity matrix. The advantage of these update rules is that the algorithm converges fast and the objective function is monotonically decreasing. On the other hand, this method requires inverse matrix operations, which are computationally expensive and may cause the numerical computation unstable.

2.2. GC-AuxIVA-ISS [12]

GC-AuxIVA-ISS [12] is a method that aims to stabilize numerical computations and reduce computational costs of GC-AuxIVA-VCD by replacing VCD with ISS. ISS performs a rank-1 update for the whole demixing matrix as

$$\boldsymbol{W}_f \leftarrow \boldsymbol{W}_f - \boldsymbol{v}_{jf} \boldsymbol{w}_{jf}^{\mathsf{H}}. \tag{17}$$

By substituting (17) to (10), the objective function of GC-AuxIVA-ISS is given as

$$\mathcal{L}(\boldsymbol{v}_{jf}) = -\sum_{f=1}^{F} \log|\det \boldsymbol{W}_f - \boldsymbol{v}_{jf}\boldsymbol{w}_{jf}^{\mathsf{H}}| + \frac{1}{2}\sum_{f=1}^{F}\sum_{i=1}^{I}\Big\{(\boldsymbol{w}_{if} - v_{ijf}^{*}\boldsymbol{w}_{jf})^{\mathsf{H}}\boldsymbol{\Sigma}_{jf}(\boldsymbol{w}_{if} - v_{ijf}^{*}\boldsymbol{w}_{jf}) + \sum_{\theta\in\Theta}\lambda_{i\theta}|(\boldsymbol{w}_{if} - v_{ijf}^{*}\boldsymbol{w}_{jf})^{\mathsf{H}}\boldsymbol{d}_{f\theta} - c_{i\theta}|^2\Big\}. \quad (18)$$

From $\partial\mathcal{L}(\boldsymbol{v}_j)/\partial v_{ij}^{*} = 0$, the update rules for $\boldsymbol{v}_{ijf}$ can be derived [12]. In the following, we omit the index f for the notation simplicity. When $i \neq j$, the update rule for $\boldsymbol{v}_{ijf}$ is

$$v_{ij} = \frac{\sum_n \varphi(r_{in}) y_{in} y_{jn}^{*} + 2\sum_{\theta\in\Theta}\lambda_{i\theta} g_{j\theta}^{*}(g_{i\theta} - c_{i\theta})}{\sum_n \varphi(r_{in})|y_{jn}|^2 + 2\sum_{\theta\in\Theta}\lambda_{i\theta}|g_{j\theta}|^2}, \quad (19)$$

where $g_{j\theta} = \boldsymbol{w}_j^{\mathsf{H}}\boldsymbol{d}_\theta$, $\sum_n \varphi(r_{in}) y_{in} y_{jn}^{*} = \boldsymbol{w}_i^{\mathsf{H}}\boldsymbol{\Sigma}_i\boldsymbol{w}_j$, $\sum_n \varphi(r_{in})|y_{jn}|^2 = \boldsymbol{w}_j^{\mathsf{H}}\boldsymbol{\Sigma}_i\boldsymbol{w}_j$, and $\varphi(r_{in}) = G(r_{in})'/r_{in}$. When $i = j$, the update rule for $\boldsymbol{v}_{ijf}$ is

$$v_{jj} = \begin{cases} 1 - \frac{1}{\sqrt{\alpha_j}} \ (\beta_j = 0), \\ 1 - \beta_j^{*}\frac{|\beta_j| + \sqrt{|\beta_j|^2 + 4\alpha_j}}{2\alpha_j|\beta_j|} \ (\beta_j \neq 0), \end{cases} \quad (20)$$

where $\alpha_j = \sum_n \varphi(r_{jn})|y_{jn}|^2 + 2\sum_{\theta\in\Theta}\lambda_{j\theta}|g_{j\theta}|^2$ and $\beta_j = \sum_{\theta\in\Theta}\lambda_{j\theta}c_{j\theta}g_{j\theta}$. With the updated $\boldsymbol{v}_{ijf}$, we can obtain $g_{i\theta}$ and the output signal $\boldsymbol{y}_n$ by

$$g_{i\theta} \leftarrow g_{i\theta} - v_{ij}g_{i\theta}, \quad (21)$$

$$\boldsymbol{y}_n \leftarrow \boldsymbol{y}_n - \boldsymbol{v}_j y_{jn}. \quad (22)$$

These update rules do not require matrix inversion operations, which has been shown to provide numerical stability and reduce computation time [12].

3. PROBLEM AND APPROACH

We assume that DOAs of all sources are known and given as elements of Θ. We adopt *unit* constraint, *null* constraint, and *double* constraint by setting $\lambda_{j\theta}$ $(\forall j, \theta)$ as

- *unit*: $\lambda_{j\theta_{tgt}} > 0,\ \lambda_{j\theta_{inf}} = 0$ $(\forall j, \theta_{tgt}, \theta_{inf})$,
- *null*: $\lambda_{j\theta_{tgt}} = 0,\ \lambda_{j\theta_{inf}} > 0$ $(\forall j, \theta_{tgt}, \theta_{inf})$,
- *double*: $\lambda_{j\theta_{tgt}} > 0,\ \lambda_{j\theta_{inf}} > 0$ $(\forall j, \theta_{tgt}, \theta_{inf})$,

where θ_{tgt} and θ_{inf} denote the DOA for target and interference signals, respectively. Note that update rules of GC-AuxIVA-VCD and GC-AuxIVA-ISS with $\lambda_{j\theta} = 0$ $(\forall j, \theta)$ are equivalent to those employed in the AuxIVA [5]. In previous studies, experiments have only been conducted with the setting $\lambda_{j\theta_{tgt}} = \lambda_{j\theta_{inf}}$ $(\forall j)$ for the *double* constraint [12]. However, since the appropriate values for $\lambda_{j\theta_{tgt}}$ and $\lambda_{j\theta_{inf}}$ have been shown to be different for *unit* and *null* constraints, we cannot expect $\lambda_{j\theta_{tgt}} = \lambda_{j\theta_{inf}}$ $(\forall j)$ to be the optimal setting for the *double* constraint. Therefore, we propose to separately consider and investigate values for $\lambda_{j\theta_{tgt}}$ and $\lambda_{j\theta_{inf}}$ for the *double* constraint to improve the performance.

Furthermore, there are no rules to help parameter tuning in GC-AuxIVA-VCD and GC-AuxIVA-ISS, making it challenging to apply these two algorithms to practical applications. Since *double* constraints impose constraints on both the target and interference signals, the parameter tuning becomes more complicated. We conclude with several tips for parameter tuning based on numerical experiments to facilitate the difficulty.

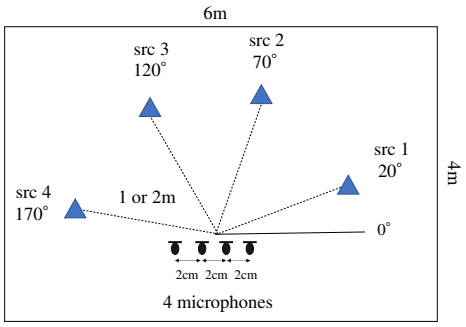

Fig. 1: Layout of speech sources and microphones.

Table 1: Avarage SDR and SIR [dB] achieved by each method with the optimal parameter setting, which was selected based on SDR and the accuracy of output signal order. The accuracy of output signal orders for all results in the table were 100%.

method	constraint	$\lambda_{j\theta}$	SDR [dB]	SIR [dB]
2 channels				
GC-AuxIVA-VCD	*unit*	100	8.26	9.70
	null	1	11.05	13.27
	double (previous) [12]	1	10.96	13.19
	double (proposed)	$\lambda_{j\theta_{tgt}} = 1$, $\lambda_{j\theta_{inf}} = 10$	**11.10**	**13.39**
GC-AuxIVA-ISS	*unit*	70	10.99	13.26
	null	100000	**11.08**	**13.34**
	double (previous) [12]	100	10.88	13.12
	double (proposed)	$\lambda_{j\theta_{tgt}} = 70$, $\lambda_{j\theta_{inf}} = 100000$	11.05	13.29
3 channels				
GC-AuxIVA-VCD	*unit*	10	6.90	8.55
	null	10	10.58	12.66
	double (previous) [12]	1	10.66	12.85
	double (proposed)	$\lambda_{j\theta_{tgt}} = 0.1$, $\lambda_{j\theta_{inf}} = 10$	**10.81**	**12.97**
GC-AuxIVA-ISS	*unit*	60	10.35	12.46
	null	100000	**10.66**	**12.80**
	double (previous) [12]	100	10.07	12.16
	double (proposed)	$\lambda_{j\theta_{tgt}} = 10$, $\lambda_{j\theta_{inf}} = 100000$	10.52	12.62
4 channels				
GC-AuxIVA-VCD	*unit*	10	4.36	5.96
	null	10	9.31	11.24
	double (previous) [12]	1	9.55	11.83
	double (proposed)	$\lambda_{j\theta_{tgt}} = 1$, $\lambda_{j\theta_{inf}} = 10$	**9.85**	**11.98**
GC-AuxIVA-ISS	*unit*	50	8.98	11.15
	null	20000	**9.28**	**11.48**
	double (previous) [12]	100	8.13	10.17
	double (proposed)	$\lambda_{j\theta_{tgt}} = 50$, $\lambda_{j\theta_{inf}} = 10000$	9.08	11.24

4. EXPERIMENT

4.1. Experimental conditions

To separately investigate the effect of $\lambda_{j\theta}$ $(\forall j, \theta)$ in separation performance of GC-AuxIVA-VCD and GC-AuxIVA-ISS, we conducted several speech separation experiments. We conducted the experiments for 2 to 4 sources and used the speech signals of 2 to 4 different speakers, which were randomly selected from 6 speakers (3 males and 3 females) in Set B of the ATR Japanese Speech Database [16]. We created a total of 48 patterns of observed signals

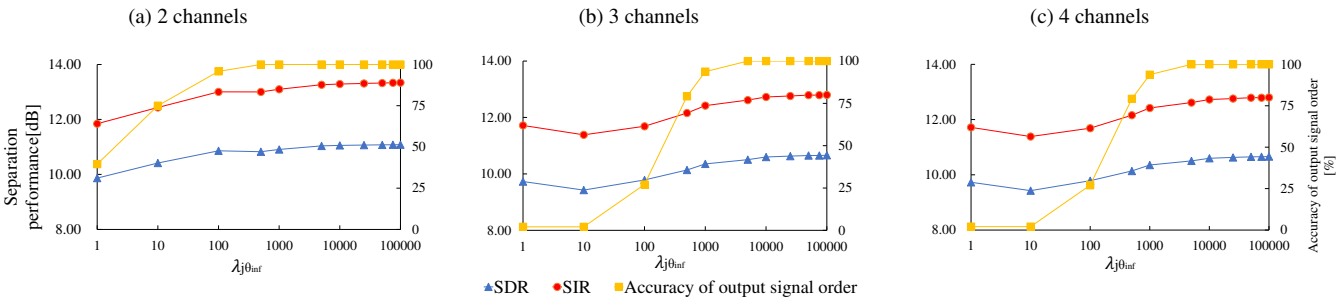

Fig. 2: Relation between $\lambda_{j\theta_{inf}}$ and average SDR, SIR [dB], and accuracy of output signal order [%] in GC-AuxIVA-ISS with *null* constraint, where $\lambda_{j\theta_{tgt}} = 0$.

for each number of speech sources by convoluting the speech with the room impulse responses (RIR) generated by the Python package `pyroomacoustics` [17]. The reverberation times (RT_{60}) were 100 and 300 ms. DOAs were set to 20° and 70° for the 2 sound sources, 20°, 70°, and 120° for the 3 sound sources, and 20°, 70°, 120° and 170° for 4 sound sources. The layout of sound sources and microphones is shown in Fig.1. The number of microphones was set equal to the number of sources, and the microphone spacing was set to 2 cm. In these experiments, the correct DOAs of the speakers were assumed to be known. Therefore, the direction set Θ was given as $\Theta = \{20°, 70°\}$, $\Theta = \{20°, 70°, 120°\}$, and $\Theta = \{20°, 70°, 120°, 170°\}$ for $J = 2$, $J = 3$, and $J = 4$, respectively.

We sampled all speech signals at 16 kHz. The STFT was computed using a Hanning window with a window length of 512 samples (32 ms) and a shift of 256 samples (16 ms). All methods were run for 50 iterations.

The following three objective metrics were used to evaluate the separation performance: source-to-distortions ratio (SDR), source-to-interferences ratio (SIR) [18], and the order of the output signals. For AuxIVA with no geometric constraints, the order of the output signals was determined as the one that achieved the highest SIR among all permutations.

4.2. Results

Table 1 shows the average SDR and SIR. The proposed *double* constraint in both GC-AuxIVA-VCD and GC-AuxIVA-ISS showed higher SDR and SIR scores than the previous *double* constraint. These results indicate that setting $\lambda_{j\theta_{tgt}}$ and $\lambda_{j\theta_{inf}}$ with different optimal values for *double* constraint is effective in improving source separation performance. In GC-AuxIVA-VCD, the separation performance of the proposed *double* constraint was higher than *unit* and *null* constraints as the number of sound sources increases. On the other hand, in GC-AuxIVA-ISS, *null* constraint has the best performance regardless of the number of sound sources. Therefore, we found that the effective constraints and parameters were widely different for each method.

Tables 2 shows the average SDR in GC-AuxIVA-VCD with *double* constraint applied for varying values of $\lambda_{j\theta_{tgt}}$ and $\lambda_{j\theta_{inf}}$. GC-AuxIVA-VCD achieved the highest SDR when $\lambda_{j\theta_{tgt}}$ was set below 1 for each $\lambda_{j\theta_{inf}}$. Therefore, in GC-AuxIVA-VCD with *double* constraint, by setting $\lambda_{j\theta_{tgt}}$ a low value, such as below 1, we can obtain high separation performance by tuning only $\lambda_{j\theta_{inf}}$.

Figure 2 shows the average SDR, SIR, and accuracy of output signal order in GC-AuxIVA-ISS with *null* constraint for various values of $\lambda_{j\theta_{inf}}$. GC-AuxIVA-ISS with *null* constraint achieved high separation performance in terms of SDR, SIR, and accuracy of output signal order when $\lambda_{j\theta_{inf}}$ was set at a high value, such as above 10,000. These results confirmed that setting $\lambda_{j\theta_{inf}}$ with high values could eliminate the need for parameter tuning.

Table 2: Avarage SDR [dB] of GC-AuxIVA-VCD with *double* constraint. Parentheses indicate scores with accuracy of output signal order less than 100%.

(a) 2 channels

$\lambda_{j\theta_{tgt}}$ \ $\lambda_{j\theta_{inf}}$	0.01	0.1	1	10	100	1000
0.01	(10.15)	**10.99**	11.06	11.02	10.53	9.40
0.1	(10.66)	(10.28)	**11.09**	11.05	10.63	9.53
1	(8.19)	(8.30)	10.96	**11.10**	**10.69**	**9.60**
10	**8.69**	8.91	9.92	9.84	9.91	9.27
100	8.43	8.86	9.68	9.89	9.93	8.72
1000	8.28	8.74	9.62	9.95	9.91	8.36

(b) 3 channels

$\lambda_{j\theta_{tgt}}$ \ $\lambda_{j\theta_{inf}}$	0.01	0.1	1	10	100	1000
0.01	(9.96)	(9.28)	10.64	10.64	9.33	6.65
0.1	(9.34)	(9.73)	**10.76**	**10.81**	9.70	6.94
1	(6.51)	(9.07)	10.66	10.81	**9.94**	**7.16**
10	**7.07**	**7.61**	8.65	8.63	8.47	6.65
100	6.18	6.54	7.19	7.31	7.09	5.71
1000	6.01	6.34	7.03	7.22	6.95	5.44

(c) 4 channels

$\lambda_{j\theta_{tgt}}$ \ $\lambda_{j\theta_{inf}}$	0.01	0.1	1	10	100	1000
0.01	(8.83)	(7.96)	9.36	9.50	7.99	5.05
0.1	(9.01)	(8.58)	9.51	9.80	8.47	5.47
1	(5.16)	**8.60**	**9.55**	**9.85**	**8.71**	**5.75**
10	**4.93**	5.51	6.76	6.64	6.44	5.04
100	3.91	4.32	5.15	5.48	4.87	3.29
1000	3.70	4.10	5.00	5.40	4.47	2.72

5. CONCLUSION

In this paper, we numerically investigated the effect of weight parameters for GC-AuxIVA-VCD and GC-AuxIVA-ISS. The experimental results showed that (1) the optimal weights for the constraints on the directions of the target and interference signals are different in these two algorithms; (2) setting different values of weights for the constraints was effective in improving separation performance; (3) the parameter tuning of GC-AuxIVA-VCD with *double* constraints could be simplified to only tuning weights for the interference constraints while setting those for the target constraints at values smaller than 1; (4) GC-AuxIVA-ISS with null constraints could achieve reasonably good and stable performance with the weight parameters for the interference constraints setting at a high value.

6. ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 19H04131.

7. REFERENCES

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

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

[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] N. Ono, "Fast stereo independent vector analysis and its implementation on mobile phone," in *Proc. IWAENC*, 2012, pp. 1–4.

[7] 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.

[8] H. Saruwatari, T. Kawamura, T. Nishikawa, A. Lee, and K. Shikano, "Blind source separation based on a fastconvergence algorithm combining ICA and beamforming," *IEEE Trans. ASLP*, vol. 14, no. 2, pp. 666–678, 2006.

[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, L. Li, R. Takahashi, S. Makino, and T. Yamada, "Study on geometrically constrained IVA with auxiliary function approach and VCD for in-car communication," in *Proc. APSIPA*, 2020, pp. 858–862.

[12] K. Goto, T. Ueda, L. Li, T. Yamada, and S. Makino (in press), "Geometrically constrained independent vector analysis with auxiliary function approach and iterative source steering," in *Proc. EUSIPICO*.

[13] 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.

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

[15] 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.

[16] 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.

[17] 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.

[18] 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.