Independent Low-Rank Matrix Analysis Based on Time-Variant Sub-Gaussian Source Model

Shinichi Mogami*, Norihiro Takamune*, Daichi Kitamura†, Hiroshi Saruwatari*,
Yu Takahashi‡, Kazunobu Kondo‡, Hiroaki Nakajima‡, and Nobutaka Ono§
* The University of Tokyo, Tokyo, Japan
† National Institute of Technology, Kagawa College, Kagawa, Japan
‡ Yamaha Corporation, Shizuoka, Japan
§ Tokyo Metropolitan University, Tokyo, Japan

***Abstract*—Independent low-rank matrix analysis (ILRMA) is a fast and stable method for blind audio source separation. Conventional ILRMAs assume time-variant (super-)Gaussian source models, which can only represent signals that follow a super-Gaussian distribution. In this paper, we focus on ILRMA based on a generalized Gaussian distribution (GGD-ILRMA) and propose a new type of GGD-ILRMA that adopts a time-variant sub-Gaussian distribution for the source model. By using a new update scheme called generalized iterative projection for homogeneous source models, we obtain a convergence-guaranteed update rule for demixing spatial parameters. In the experimental evaluation, we show the versatility of the proposed method, i.e., the proposed time-variant sub-Gaussian source model can be applied to various types of source signal.**

I. Introduction

Blind source separation (BSS) [1]–[8] is a technique for extracting specific sources from an observed multichannel mixture signal without knowing a priori information about the mixing system. The most commonly used algorithm for BSS in the (over)determined case (number of microphones $\geq$ number of sources) is independent component analysis (ICA) [1]. As a state-of-the-art ICA-based BSS method, Kitamura et al. proposed *independent low-rank matrix analysis (ILRMA)* [9], [10], which is a unification of independent vector analysis (IVA) [5], [6] and nonnegative matrix factorization (NMF) [11]. ILRMA assumes both statistical independence between sources and a low-rank time-frequency structure for each source, and the frequency-wise demixing systems are estimated without encountering the permutation problem [3], [4]. ILRMA is a faster and more stable algorithm than multichannel NMF (MNMF) [12]–[14], which is an algorithm for BSS that estimates the mixing system on the basis of spatial covariance matrices.

The original ILRMA based on Itakura–Saito (IS) divergence assumes a time-variant isotropic complex Gaussian distribution for the source generative model. Hereafter, we refer to the original ILRMA as *IS-ILRMA*. Recently, various types of source generative model have been proposed in ILRMA for robust BSS. In particular, t-ILRMA [15] and GGD-ILRMA [16], [17] have been proposed as generalizations of IS-ILRMA with a complex Student's t distribution and a complex generalized Gaussian distribution (GGD), respectively. In t-ILRMA and GGD-ILRMA, the kurtosis of the generative models' distributions can be parametrically changed along with the degree-of-freedom parameter in Student's t distribution and the shape parameter in the GGD. By changing the kurtosis of the distributions, we can control how often the source signal outputs outliers or its expected sparsity. In particular, in sub-Gaussian models, i.e., models that follow distributions with a *platykurtic* shape, the source signal rarely outputs outliers. Therefore, the sub-Gaussian modeling of sources is expected to accurately estimate the source spectrogram without ignoring its important spectral peaks. Furthermore, many audio sources follow platykurtic distributions; it is known that musical instrument signals obey sub-Gaussian distributions [18].

However, neither conventional t-ILRMA nor GGD-ILRMA assumes that the source model follows a sub-Gaussian distribution. Both t-ILRMA and GGD-ILRMA can adopt only a super-Gaussian (or Gaussian) model, i.e., models that follow distributions with a *leptokurtic* shape, for the source model. In t-ILRMA, this is because the complex Student's t distribution becomes only super-Gaussian for any degree-of-freedom parameter. In GGD-ILRMA, on the other hand, it is because the estimation algorithm for the demixing matrix has not yet been derived for a sub-Gaussian case, although the GGD itself can represent a sub-Gaussian distribution depending on its shape parameter. More specifically, the conventional *iterative projection (IP)* [7], which is an algorithm that updates the demixing matrix mainly used in IVA and ILRMA, cannot be applied to sub-Gaussian-based GGD-ILRMA owing to mathematical difficulties.

In this paper, we propose a new type of ILRMA that assumes time-variant sub-Gaussian source models. This paper includes three novelties. First, we construct a new update scheme for the demixing matrix called *generalized IP for homogeneous source models (GIP-HSM)*. Second, we derive a convergence-guaranteed update rule for the demixing matrix in GGD-ILRMA with a shape parameter of four. To the best of our knowledge, this is the world's first attempt to model the source signal with a time-variant sub-Gaussian distribution, and we derive the update rule by applying the above-mentioned scheme. Third, we show the validity of the proposed sub-Gaussian GGD-ILRMA via BSS experiments on music and speech signals. We confirm that the proposed method is a versatile approach to source modeling, i.e., the proposed time-variant sub-Gaussian model can represent super-Gaussian or Gaussian signals as well as sub-Gaussian signals owing to its time-variant nature, whereas the conventional models can only represent super-Gaussian or Gaussian signals.

II. Problem Formulation

A. Formulation of Demixing Model

Let N and M be the numbers of sources and channels, respectively. The short-time Fourier transforms (STFTs) of the multichannel source, observed, and estimated signals are defined as

$$\boldsymbol{s}_{ij} = (s_{ij1}, \ldots, s_{ijN})^\top \in \mathbb{C}^N, \tag{1}$$
$$\boldsymbol{x}_{ij} = (x_{ij1}, \ldots, x_{ijM})^\top \in \mathbb{C}^M, \tag{2}$$
$$\boldsymbol{y}_{ij} = (y_{ij1}, \ldots, y_{ijN})^\top \in \mathbb{C}^N, \tag{3}$$

where $i = 1, \ldots, I$; $j = 1, \ldots, J$; $n = 1, \ldots, N$; and $m = 1 \ldots, M$ are the integral indices of the frequency bins, time frames, sources, and channels, respectively, and $^\top$ denotes the transpose. We assume the mixing system

$$\boldsymbol{x}_{ij} = \boldsymbol{A}_i \boldsymbol{s}_{ij}, \tag{4}$$

where $\boldsymbol{A}_i = (\boldsymbol{a}_{i1}, \ldots, \boldsymbol{a}_{iN}) \in \mathbb{C}^{M \times N}$ is a frequency-wise mixing matrix and $\boldsymbol{a}_{in}$ is the steering vector for the nth source. When $M = N$ and $\boldsymbol{A}_i$ is not a singular matrix, the estimated signal $\boldsymbol{y}_{ij}$ can be expressed as

$$\boldsymbol{y}_{ij} = \boldsymbol{W}_i \boldsymbol{x}_{ij}, \tag{5}$$

where $\boldsymbol{W}_i = \boldsymbol{A}_i^{-1} = (\boldsymbol{w}_{i1}, \ldots, \boldsymbol{w}_{iN})^{\mathsf{H}}$ is the demixing matrix, $\boldsymbol{w}_{in}$ is the demixing filter for the nth source, and $^{\mathsf{H}}$ denotes the Hermitian transpose. ILRMA estimates both $\boldsymbol{W}_i$ and $\boldsymbol{y}_{ij}$ from only the observation $\boldsymbol{x}_{ij}$ assuming statistical independence between s_{ijn} and $s_{ijn'}$, where $n \neq n'$.

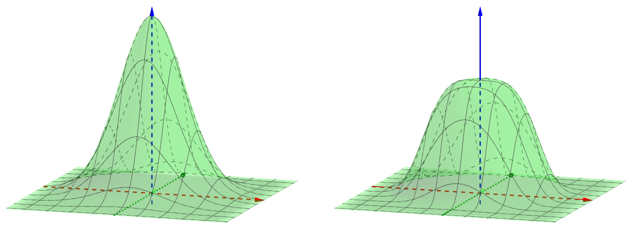
B. Generative Model and Cost Function in GGD-ILRMA

GGD-ILRMA utilizes the isotropic complex GGD. The probability density function of the GGD is

$$p(z) = \frac{\beta}{2\pi r^2 \Gamma(2/\beta)} \exp\left(-\frac{|z|^\beta}{r^\beta}\right), \tag{6}$$

where β is the shape parameter, r is the scale parameter, and $\Gamma(\cdot)$ is the gamma function. Fig. 1 shows the shapes of the GGD with $\beta = 2$ and $\beta = 4$. When $\beta = 2$, (6) corresponds to the probability density function of the complex Gaussian distribution with a *mesokurtic* shape. In the case of $0 < \beta < 2$, the distribution becomes super-Gaussian with a leptokurtic shape. In the case of $\beta > 2$, the distribution becomes sub-Gaussian with a platykurtic shape.

In GGD-ILRMA, we assume the time-variant isotropic complex GGD as the source generative model, which is independently defined in each time-frequency slot as follows:



(a) GGD with $\beta = 2$. (b) GGD with $\beta = 4$.

Fig. 1. Examples of shapes of complex GGD. (a) When $\beta = 2$, shape of GGD corresponds to that of Gaussian distribution. (b) When $\beta = 4$, GGD is platykurtic.

$$\begin{aligned} p(\boldsymbol{Y}_n) &= \prod_{i,j} p(y_{ijn}) \\ &= \prod_{i,j} \frac{\beta}{2\pi {r_{ijn}}^2 \Gamma(2/\beta)} \exp\left(-\frac{|y_{ijn}|^\beta}{{r_{ijn}}^\beta}\right), \end{aligned} \tag{7}$$

$${r_{ijn}}^p = \sum_k t_{ikn} v_{kjn}, \tag{8}$$

where the local distribution $p(y_{ijn})$ is defined as a circularly symmetric complex Gaussian distribution, i.e., the probability of $p(y_{ijn})$ only depends on the power of the complex value y_{ijn}. r_{ijn} is the time-frequency-varying scale parameter and p is the domain parameter in NMF modeling. Moreover, the variables t_{ikn} and v_{kjn} are the elements of the basis matrix $\boldsymbol{T}_n \in \mathbb{R}_{\geq 0}^{I \times K}$ and the activation matrix $\boldsymbol{V}_n \in \mathbb{R}_{\geq 0}^{K \times J}$, respectively, where $\mathbb{R}_{\geq 0}$ denotes the set of nonnegative real numbers. $k = 1, \cdots, K$ is the integral index, and K is set to a much smaller value than I and J, which leads to the low-rank approximation. From (7), the negative log-likelihood function $\mathcal{L}_{\mathrm{GGD}}$ of the observed signal $\boldsymbol{x}_{ij}$ can be obtained as follows by assuming independence between sources:

$$\begin{aligned} \mathcal{L}_{\mathrm{GGD}} = &-2J \sum_i \log |\det \boldsymbol{W}_i| \\ &+ \sum_{i,j,n} \left(\frac{|y_{ijn}|^\beta}{{r_{ijn}}^\beta} + 2 \log r_{ijn} \right), \end{aligned} \tag{9}$$

where $y_{ijn} = \boldsymbol{w}_{in} \boldsymbol{x}_{ij}$ and we used the transformation of random variables from $\boldsymbol{x}_{ij}$ to $\boldsymbol{y}_{ij}$. The cost function of GGD-ILRMA (9) coincides with that of IS-ILRMA when $\beta = p = 2$. By minimizing (9) w.r.t. $\boldsymbol{W}_i$ and r_{ijn} under the limitation (8), we estimate the demixing system that maximizes the independence between sources.

Fig. 2 shows a conceptual model of GGD-ILRMA. When each of the original sources has a low-rank spectrogram, the spectrogram of their mixture should be more complicated, where the rank of the mixture spectrogram will be greater than that of the source spectrogram. On the basis of this assumption, in GGD-ILRMA, the low-rank constraint for each estimated spectrogram is introduced by employing NMF. The demixing matrix $\boldsymbol{W}_i$ is estimated so that the spectrogram of

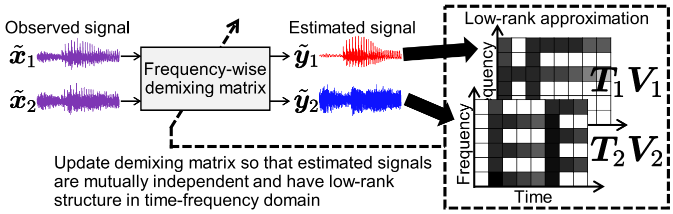


Fig. 2. Principle of source separation in GGD-ILRMA, where $\tilde{\boldsymbol{x}}_m$ and $\tilde{\boldsymbol{y}}_n$ are time-domain signals of x_{ijn} and y_{ijn}, respectively.

the estimated signal becomes a low-rank matrix modeled by $\boldsymbol{T}_n \boldsymbol{V}_n$, whose rank is at most K. The estimation of $\boldsymbol{W}_i, \boldsymbol{T}_n$, and $\boldsymbol{V}_n$ can consistently be carried out by minimizing (9) in a fully blind manner.

III. Conventional Method

A. Update Rule for Demixing Matrix

In IS-ILRMA, the demixing matrix $\boldsymbol{W}_i$ can be efficiently updated by IP, which can be applied only when the cost function is the sum of $-\log|\det \boldsymbol{W}_i|$ and the quadratic form of $\boldsymbol{w}_{in}$ (this corresponds to GGD-ILRMA with $\beta = 2$). In GGD-ILRMA, the update rule of $\boldsymbol{W}_i$ is also derived using the majorization-minimization (MM) algorithm [19]. When $0 < \beta \leq 2$, we can use the following inequality of weighted arithmetic and geometric means to design the majorization function:

$$|y_{ijn}|^\beta \leq \frac{\beta}{2}\frac{|y_{ijn}|^2}{\alpha_{ijn}{}^{2-\beta}} + \left(1 - \frac{\beta}{2}\right)\alpha_{ijn}{}^\beta, \tag{10}$$

where α_{ijn} is an auxiliary variable and the equality of (10) holds if and only if $\alpha_{ijn} = |y_{ijn}|$. By applying (10) to (9), the majorization function of (9) can be designed as

$$\begin{aligned}\mathcal{L}_{\mathrm{GGD}} \leq & -2J\sum_i \log|\det \boldsymbol{W}_i| \\ & + J\sum_{i,n} \boldsymbol{w}_{in}^{\mathsf{H}} \boldsymbol{F}_{in} \boldsymbol{w}_{in} + \mathrm{const.},\end{aligned} \tag{11}$$

$$\boldsymbol{F}_{in} = \frac{\beta}{2J}\sum_j \frac{1}{\alpha_{ijn}{}^{2-\beta}(\sum_k t_{ikn}v_{kjn})^{\frac{\beta}{p}}} \boldsymbol{x}_{ij}\boldsymbol{x}_{ij}^{\mathsf{H}}, \tag{12}$$

where the constant term is independent of $\boldsymbol{w}_{in}$. By applying IP to (11) and substituting the equality condition $\alpha_{ijn} = |y_{ijn}|$ into (12), the update rule for $\boldsymbol{W}_i$ is derived as

$$\boldsymbol{F}_{in} = \frac{\beta}{2J}\sum_j \frac{1}{|y_{ijn}|^{2-\beta}\left(\sum_k t_{ikn}v_{kjn}\right)^{\frac{\beta}{p}}} \boldsymbol{x}_{ij}\boldsymbol{x}_{ij}^{\mathsf{H}}, \tag{13}$$

$$\boldsymbol{w}_{in} \leftarrow \boldsymbol{F}_{in}^{-1}\boldsymbol{W}_i^{-1}\boldsymbol{e}_n, \tag{14}$$

$$\boldsymbol{w}_{in} \leftarrow \boldsymbol{w}_{in}\sqrt{1/(\boldsymbol{w}_{in}^{\mathsf{H}}\boldsymbol{F}_{in}\boldsymbol{w}_{in})}. \tag{15}$$

When $\beta = p = 2$, these update rules (13)–(15) coincide with those in IS-ILRMA.

Note that the update rules (13)–(15) are valid only when $0 < \beta \leq 2$, which is equivalent to the condition that the inequality (10) holds. In fact, when $\beta > 2$, it is thought to be impossible to design a majorization function to which we can apply IP because no quadratic function w.r.t. x can majorize x^β.

Conventional GGD-ILRMA achieves various types of source generative model: when $\beta = 2$, the entry of the source spectrogram follows the complex Gaussian distribution (the same model as that of IS-ILRMA), and when $\beta < 2$, the entry of the source spectrogram follows the complex leptokurtic distribution. However, a source generative model that follows a platykurtic complex GGD is yet to be achieved. Since the marginal distribution of the time-variant super-Gaussian or Gaussian model w.r.t. the time frame becomes only super-Gaussian, any signals that follow sub-Gaussian distributions, such as music signals, cannot be appropriately dealt with by the conventional GGD-ILRMA.

B. Update Rule for Low-Rank Source Model

The update rules for $\boldsymbol{T}_n$ and $\boldsymbol{V}_n$ in IS-ILRMA and GGD-ILRMA can be derived by the MM algorithm, which is a popular approach for NMF. We obtain the following update rules:

$$t_{ikn} \leftarrow t_{ikn}\left(\frac{\beta\sum_j \dfrac{|y_{ijn}|^\beta}{(\sum_{k'} t_{ik'n}v_{k'jn})^{\frac{\beta}{p}+1}}v_{kjn}}{2\sum_j \dfrac{1}{\sum_{k'} t_{ik'n}v_{k'jn}}v_{kjn}}\right)^{\frac{p}{\beta+p}}, \tag{16}$$

$$v_{kjn} \leftarrow v_{kjn}\left(\frac{\beta\sum_j \dfrac{|y_{ijn}|^\beta}{(\sum_{k'} t_{ik'n}v_{k'jn})^{\frac{\beta}{p}+1}}t_{ikn}}{2\sum_j \dfrac{1}{\sum_{k'} t_{ik'n}v_{k'jn}}t_{ikn}}\right)^{\frac{p}{\beta+p}}. \tag{17}$$

See Appendix A for their detailed derivation.

In GGD-ILRMA, the cost function (9) is minimized by alternately repeating the update of the demixing matrix $\boldsymbol{W}_i$ using (13)–(15) and the update of the low-rank source models $\boldsymbol{T}_n$ and $\boldsymbol{V}_n$ using (16) and (17), respectively. A monotonic decrease in the cost is guaranteed over these update rules.

IV. Proposed Method

A. Motivation

The conventional methods [9], [16] have a limitation that the source signal cannot be appropriately represented when the signal follows a sub-Gaussian distribution. In this paper, we propose an MM-algorithm-based update rule for GGD-ILRMA to maximize the likelihood based on the sub-Gaussian source model. To derive the update rule, we also extend the problem of demixing matrix estimation into a more generalized form and propose its optimization scheme based on GIP-HSM.

In contrast to the time-variant super-Gaussian or Gaussian model, the marginal distribution of the time-variant sub-Gaussian model w.r.t. the time frame can be sub-Gaussian as well as Gaussian or super-Gaussian, depending on its time variance of the scale parameter r_{ijn}. For example, the time-variant sub-Gaussian model is platykurtic when r_{ijn} is constant w.r.t. the time frame, whereas it becomes mesokurtic or leptokurtic when r_{ijn} fluctuates appropriately. This shows that the proposed time-variant sub-Gaussian model covers distributions with a wider range between platykurtic and leptokurtic shapes than other conventional source models. Therefore, the proposed GGD-ILRMA is expected to have a robust performance against the variation of the target signals.

B. Derivation of GIP-HSM

The cost functions of IS-ILRMA and GGD-ILRMA are generalized as

$$\mathcal{L} = \sum_{i=1}^{I} \left[-2\log|\det \boldsymbol{W}_i| + \sum_{n=1}^{N} f_{in}(\boldsymbol{w}_{in}) \right] + \text{const.}, \quad (18)$$

where the constant term is independent of $\boldsymbol{w}_{in}$ and $f_{in} : \mathbb{C}^N \to \mathbb{R}$ is a real-valued function that satisfies the following three conditions:

1) $f_{in}(\boldsymbol{w})$ is differentiable w.r.t. $\boldsymbol{w}$ at an arbitrary point.
2) $^\forall c > 0$, $\left\{\boldsymbol{w} \in \mathbb{C}^N \mid f_{in}(\boldsymbol{w}) \le c\right\}$ is convex (naturally satisfied when $f_{in}(\boldsymbol{w})$ is convex).
3) $^\forall \eta$, $f_{in}(\eta \boldsymbol{w}) = \eta^d f_{in}(\boldsymbol{w})$, namely, f_{in} is a homogeneous function of degree d.

The term $f_{in}(\boldsymbol{w}_{in})$ is determined by the distribution of the source generative model, e.g., $f_{in}(\boldsymbol{w}_{in}) = (1/J)\sum_j \left(|\boldsymbol{w}_{in}^{\mathsf{H}}\boldsymbol{x}_{ij}|^\beta / {r_{ijn}}^\beta\right)$ in GGD-ILRMA.

Here we show that the optimization of (18) w.r.t. $\boldsymbol{w}_{in}$ is composed of "direction optimization" and "scale optimization" for each frequency bin. Let $\boldsymbol{u}_{in}$ be an N-dimensional vector that satisfies $f_{in}(\boldsymbol{u}_{in}) = 1$. Then, $\boldsymbol{w}_{in}$ can be uniquely represented as $\boldsymbol{w}_{in} = \eta_{in}\boldsymbol{u}_{in}$, where η_{in} is a positive real value. By regarding $f_{in}(\boldsymbol{u})$ as the norm of $\boldsymbol{u}$, we can interpret $\boldsymbol{u}_{in}$ as a unit vector w.r.t. the f_{in}-norm. Substituting $\boldsymbol{w}_{in} = \eta_{in}\boldsymbol{u}_{in}$ into (18), the cost function is represented as

$$\begin{aligned}\mathcal{L} &= \sum_{i=1}^{I} \Bigg[-2\log\left|\det \begin{bmatrix} \eta_1 \boldsymbol{u}_1 & \cdots & \eta_N \boldsymbol{u}_N \end{bmatrix}^{\mathsf{H}}\right| \\ &\qquad + \sum_{n=1}^{N} f_{in}(\eta_{in}\boldsymbol{u}_{in}) \Bigg] \\ &= \sum_{i=1}^{I} \left[-2\log\left(\prod_n \eta_{in} \cdot |\det \boldsymbol{U}_i|\right) + \sum_{n=1}^{N} {\eta_{in}}^d f_{in}(\boldsymbol{u}_{in}) \right] \\ &= \sum_{i=1}^{I} \left[-2\log|\det \boldsymbol{U}_i| + \sum_{n=1}^{N} \left[-2\log\eta_{in} + {\eta_{in}}^d\right] \right], \quad (19)\end{aligned}$$

where $\boldsymbol{U}_i = \begin{bmatrix} \boldsymbol{u}_{i1} & \cdots & \boldsymbol{u}_{iN} \end{bmatrix}^{\mathsf{H}}$. Therefore, the minimization of the cost function can be interpreted as the minimization of $-\log|\det \boldsymbol{U}_i|$ for each frequency bin and the minimization of $-2\log\eta_{in} + {\eta_{in}}^d$ for each source and frequency bin. These direction optimization and scale optimization problems are independent of each other. The optimal η_{in} can be calculated by a closed form because the derivative of the cost function w.r.t. η_{in} can be written as

$$\frac{d}{d\eta_{in}}(-2\log\eta_{in} + {\eta_{in}}^d) = -\frac{2}{\eta_{in}} + d{\eta_{in}}^{d-1}. \quad (20)$$

Hence, letting the right side of (20) be zero, we can obtain the optimal η_{in} as

$$\eta_{in} = \sqrt[d]{2/d}. \quad (21)$$

The actual difficulty in the optimization of the demixing matrix is the direction optimization, i.e., the minimization of $-2\log|\det \boldsymbol{U}_i|$. Since minimizing $-\log x^2$ is equivalent to maximizing x^2, we can reformulate this problem as

$$\text{maximize } |\det \boldsymbol{U}_i|^2 \quad \text{s.t. } f_{in}(\boldsymbol{u}_{in}) = 1. \quad (22)$$

Since it is generally difficult to solve this problem by a closed form, we apply an approach called vectorwise coordinate descent. In this algorithm, we focus on $\boldsymbol{u}_{in}$, namely, the Hermitian transpose of a particular row vector of $\boldsymbol{U}_i$. By cofactor expansion, we can deform the problem (22) as

$$\text{maximize } \left|\boldsymbol{b}_{in}^{\mathsf{H}}\boldsymbol{u}_{in}\right|^2 \quad \text{s.t. } f_{in}(\boldsymbol{u}_{in}) = 1, \quad (23)$$

where $\boldsymbol{b}_{in}$ is a column vector of the adjugate matrix $\boldsymbol{B}_i = \begin{bmatrix} \boldsymbol{b}_{i1} & \cdots & \boldsymbol{b}_{iN} \end{bmatrix}^{\mathsf{H}}$ of $\boldsymbol{U}_i$. Since $\boldsymbol{b}_{in}$ only depends on $\boldsymbol{u}_{in'}$ $(n' \neq n)$ and is independent of $\boldsymbol{u}_{in}$, (23) can be regarded as a function of $\boldsymbol{u}_{in}$ by fixing the other row vectors of $\boldsymbol{U}_i$. Using the method of Lagrange multipliers, the stationary condition is

$$\boldsymbol{b}_{in}(\boldsymbol{b}_{in}^{\mathsf{H}}\boldsymbol{u}_{in}) + \lambda\frac{\partial f_{in}}{\partial \boldsymbol{u}_{in}^{\mathsf{H}}}(\boldsymbol{u}_{in}) = 0, \quad (24)$$

where λ is a Lagrange multiplier. Since $(\boldsymbol{b}_{in}^{\mathsf{H}}\boldsymbol{u}_{in})$ is a scalar, the stationary condition can be rewritten as

$$\frac{\partial f_{in}}{\partial \boldsymbol{u}_{in}^{\mathsf{H}}}(\boldsymbol{u}_{in}) \parallel \boldsymbol{b}_{in}, \quad (25)$$

where the binary relation "$\boldsymbol{x} \parallel \boldsymbol{y}$" means that $\boldsymbol{x}$ is parallel to $\boldsymbol{y}$. In (25), $\boldsymbol{b}_{in}$ is represented in terms of $\boldsymbol{W}_{in}$ as

$$\begin{aligned}\boldsymbol{b}_{in} &= (\det \boldsymbol{U}_i)\boldsymbol{U}_i^{-1}\boldsymbol{e}_n \\ &= (\det \boldsymbol{U}_i)(\mathrm{diag}(\eta_{i1}^{-1}, \ldots, \eta_{iN}^{-1})\boldsymbol{W}_i)^{-1}\boldsymbol{e}_n \\ &= (\det \boldsymbol{U}_i)\boldsymbol{W}_i^{-1}\mathrm{diag}(\eta_{i1}, \ldots, \eta_{iN})\boldsymbol{e}_n \\ &= (\eta_{in}\det \boldsymbol{U}_i)\boldsymbol{W}_i^{-1}\boldsymbol{e}_n \\ &\parallel \boldsymbol{W}_i^{-1}\boldsymbol{e}_n, \end{aligned} \quad (26)$$

where $\mathrm{diag}(c_1, \ldots, c_N)$ denotes the $N \times N$ diagonal matrix whose (n, n)th element is c_n, and $\boldsymbol{e}_n$ is an N-dimensional vector whose nth element is one and whose other elements are zero. Since f_{in} is convex, the stationary point of the objective function (23) must also be the optimal point. Therefore, the cost function (19) that includes (22) monotonically decreases with each update of the direction $\boldsymbol{u}_{in}$.

In conclusion, to minimize the cost function (18), we update the vector $\boldsymbol{w}_{in}$ by the following two steps in GIP-HSM. (a) Find a vector $\boldsymbol{w}'_{in}$ that satisfies

$$\frac{\partial f_{in}}{\partial \boldsymbol{w}_{in}^{\mathsf{H}}}(\boldsymbol{w}'_{in}) \parallel \boldsymbol{W}_i^{-1}\boldsymbol{e}_n. \quad (27)$$

(b) Update $\boldsymbol{w}_{in}$ as

$$\boldsymbol{w}_{in} \leftarrow \boldsymbol{w}'_{in}\sqrt[d]{2/(d \cdot f_{in}(\boldsymbol{w}'_{in}))}. \quad (28)$$

The first step (a) and second step (b) correspond to the direction and scale optimizations, respectively. Note that $\boldsymbol{w}'_{in}$ calculated in the first step does not need to satisfy $f_{in}(\boldsymbol{w}'_{in}) = 1$ because the scale is automatically adjusted in the second step. In fact, if $\boldsymbol{w}'_{in}$ is represented as $\boldsymbol{w}'_{in} = \eta'_{in}\boldsymbol{u}_{in}$, the second step results in

$$\boldsymbol{w}_{in} \leftarrow \eta'_{in}\boldsymbol{u}_{in} \cdot \sqrt[d]{2/(d \cdot {\eta'_{in}}^{d})} = \boldsymbol{u}_{in}\sqrt[d]{2/d}, \tag{29}$$

at which point both the direction and the scale are optimized.

C. Sub-Gaussian ILRMA Based on GIP-HSM

Using GIP-HSM, we propose a new update rule in GGD-ILRMA whose shape parameter β is set to four (time-variant sub-Gaussian model). The cost function of GGD-ILRMA with $\beta = 4$ is written as

$$\mathcal{J} = -2J\sum_i \log|\det \boldsymbol{W}_i| + \sum_{i,j,n}\frac{|\boldsymbol{w}_{in}^{\mathsf{H}}\boldsymbol{x}_{ij}|^4}{{r_{ijn}}^4} + \text{const.}, \tag{30}$$

where the constant term is independent of $\boldsymbol{w}_{in}$. It seems possible to apply GIP-HSM by letting $f_{in}(\boldsymbol{w}_{in}) = (1/J)\sum_j\left(|\boldsymbol{w}_{in}^{\mathsf{H}}\boldsymbol{x}_{ij}|^4/{r_{ijn}}^4\right)$. In this case, however, it is difficult to solve (27), which is reduced to a cubic vector equation w.r.t. $\boldsymbol{w}'_{in}$. Instead, we apply an MM algorithm to derive an update rule that does not contain any cubic vector equations. Hereafter, we prove the following theorem, and then design a new type of majorization function of (30) using the theorem.

Theorem 1: Let $f(\boldsymbol{w}) = (1/J)\sum_{j=1}^{J}(|\boldsymbol{w}^{\mathsf{H}}\boldsymbol{x}_j|^4/{r_j}^4)$ and $g(\boldsymbol{w}) = (\boldsymbol{w}^{\mathsf{H}}\boldsymbol{G}\boldsymbol{w})^2$, where $\boldsymbol{G}$ is defined in terms of a vector $\tilde{\boldsymbol{w}}$ as

$$\boldsymbol{H} = \begin{bmatrix}\frac{1}{r_1}\boldsymbol{x}_1 & \cdots & \frac{1}{r_J}\boldsymbol{x}_J\end{bmatrix}, \tag{31}$$

$$\tilde{\boldsymbol{q}} = \begin{bmatrix}\tilde{q}_1 & \cdots & \tilde{q}_J\end{bmatrix}^{\top} = \boldsymbol{H}^{\mathsf{H}}\tilde{\boldsymbol{w}}, \tag{32}$$

$$\tilde{\boldsymbol{Q}} = \begin{bmatrix}\|\tilde{\boldsymbol{q}}\|^2 & -\tilde{q}_1\tilde{q}_2^* & \cdots & -\tilde{q}_1\tilde{q}_J^* \\ -\tilde{q}_2\tilde{q}_1^* & \|\tilde{\boldsymbol{q}}\|^2 & \cdots & -\tilde{q}_2\tilde{q}_J^* \\ \vdots & \vdots & \ddots & \vdots \\ -\tilde{q}_J\tilde{q}_1^* & -\tilde{q}_J\tilde{q}_2^* & \cdots & \|\tilde{\boldsymbol{q}}\|^2\end{bmatrix}, \tag{33}$$

$$\boldsymbol{G} = \frac{1}{\sqrt{J\sum_j|\tilde{q}_j|^4}}\boldsymbol{H}\tilde{\boldsymbol{Q}}\boldsymbol{H}^{\mathsf{H}}. \tag{34}$$

Then, $g(\boldsymbol{w})$ satisfies $f(\boldsymbol{w}) \leq g(\boldsymbol{w})$ for arbitrary $\boldsymbol{w}$ and the equality holds when $\boldsymbol{w} = \tilde{\boldsymbol{w}}$.

Proof: Let $\boldsymbol{q} = \begin{bmatrix}q_1 & \cdots & q_J\end{bmatrix}^{\top} = \boldsymbol{H}^{\mathsf{H}}\boldsymbol{w}$. $f(\boldsymbol{w})$ and $g(\boldsymbol{w})$ can be written as

$$f(\boldsymbol{w}) = \frac{1}{J}\sum_j|q_j|^4, \tag{35}$$

$$g(\boldsymbol{w}) = \frac{1}{J\sum_j|\tilde{q}_j|^4}(\boldsymbol{q}^{\mathsf{H}}\tilde{\boldsymbol{Q}}\boldsymbol{q})^2. \tag{36}$$

Then, the objective inequality $f(\boldsymbol{w}) \leq g(\boldsymbol{w})$ holds if and only if

$$\left(\sum_j|q_j|^4\right)\left(\sum_j|\tilde{q}_j|^4\right) \leq (\boldsymbol{q}^{\mathsf{H}}\tilde{\boldsymbol{Q}}\boldsymbol{q})^2. \tag{37}$$

The quadratic form of the right side in (37) can be deformed as

$$\begin{aligned}\boldsymbol{q}^{\mathsf{H}}\tilde{\boldsymbol{Q}}\boldsymbol{q} &= \mathrm{tr}(\tilde{\boldsymbol{Q}}\boldsymbol{q}\boldsymbol{q}^{\mathsf{H}}) \\ &= \left(\sum_j|q_j|^2\right)\|\tilde{\boldsymbol{q}}\| - \sum_{i\neq j}q_iq_j^*\tilde{q}_i^*\tilde{q}_j \\ &= \left(\sum_j|q_j|^2\right)\left(\sum_j|\tilde{q}_j|^2\right) - \sum_{i\neq j}q_iq_j^*\tilde{q}_i^*\tilde{q}_j.\end{aligned} \tag{38}$$

Hence, we prove the following inequality hereafter:

$$\begin{aligned}&\left(\sum_j|q_j|^4\right)\left(\sum_j|\tilde{q}_j|^4\right) \\ &\leq \left(\left(\sum_j|q_j|^2\right)\left(\sum_j|\tilde{q}_j|^2\right) - \sum_{i\neq j}q_iq_j^*\tilde{q}_i^*\tilde{q}_j\right)^2.\end{aligned} \tag{39}$$

Let

$$x_1 = \sum_j|q_j|^2, \tag{40}$$

$$x_2 = \sqrt{\sum_{i\neq j}|q_i|^2|q_j|^2}, \tag{41}$$

$$y_1 = \sum_j|\tilde{q}_j|^2, \tag{42}$$

$$y_2 = \sqrt{\sum_{i\neq j}|\tilde{q}_i|^2|\tilde{q}_j|^2}. \tag{43}$$

Since

$${x_1}^2 - {x_2}^2 = \sum_j|q_j|^4 \geq 0, \tag{44}$$

$${y_1}^2 - {y_2}^2 = \sum_j|\tilde{q}_j|^4 \geq 0, \tag{45}$$

and $x_1, x_2, y_1, y_2 \geq 0$, it is obvious that

$$x_1y_1 - x_2y_2 \geq 0. \tag{46}$$

Furthermore, we obtain the following inequality by applying the Cauchy–Schwarz inequality:

$$\begin{aligned}x_2y_2 &= \sqrt{\left(\sum_{i\neq j}|q_i|^2|q_j|^2\right)\left(\sum_{i\neq j}|\tilde{q}_i|^2|\tilde{q}_j|^2\right)} \\ &\geq \left|\sum_{i\neq j}q_iq_j^*\tilde{q}_i^*\tilde{q}_j\right| \geq \sum_{i\neq j}q_iq_j^*\tilde{q}_i^*\tilde{q}_j,\end{aligned} \tag{47}$$

where we used the fact that

$$\begin{aligned}\sum_{i\neq j}q_iq_j^*\tilde{q}_i^*\tilde{q}_j &= \sum_{i<j}\left(q_iq_j^*\tilde{q}_i^*\tilde{q}_j + q_i^*q_j\tilde{q}_i\tilde{q}_j^*\right) \\ &= 2\sum_{i<j}\mathrm{Re}[q_iq_j^*\tilde{q}_i^*\tilde{q}_j] \in \mathbb{R}.\end{aligned} \tag{48}$$

From (46) and (47),

$$(x_1y_1 - x_2y_2)^2 \le \left(x_1y_1 - \sum_{i\neq j} q_i q_j^* \tilde{q}_i^* \tilde{q}_j\right)^2. \quad (49)$$

Therefore, (39) can be proven by using the Brahmagupta identity, (44), (45), and (49), as follows:

$$\begin{aligned}
&\left(\sum_j |q_j|^4\right)\left(\sum_j |\tilde{q}_j|^4\right)\\
&= ({x_1}^2 - {x_2}^2)({y_1}^2 - {y_2}^2)\\
&= (x_1y_1 - x_2y_2)^2 - (x_1y_2 - x_2y_1)^2\\
&\le (x_1y_1 - x_2y_2)^2\\
&\le \left(x_1y_1 - \sum_{i\neq j} q_i q_j^* \tilde{q}_i^* \tilde{q}_j\right)^2\\
&= \left(\left(\sum_j |q_j|^2\right)\left(\sum_j |\tilde{q}_j|^2\right) - \sum_{i\neq j} q_i q_j^* \tilde{q}_i^* \tilde{q}_j\right)^2. \quad (50)
\end{aligned}$$

It is easy to prove that the equality of (39) holds if $\boldsymbol{w} = \tilde{\boldsymbol{w}}$ because then $\boldsymbol{q} = \tilde{\boldsymbol{q}}$ holds. ■

Applying Theorem 1, we can design a majorization function of (30) as

$$\begin{aligned}
\mathcal{J} &\le -2J\sum_i \log|\det \boldsymbol{W}_i| + J\sum_{i,n}(\boldsymbol{w}_{in}^{\mathsf{H}}\boldsymbol{G}_{in}\boldsymbol{w}_{in})^2 + \text{const.}\\
&=: \mathcal{J}^+, \quad (51)
\end{aligned}$$

where

$$\boldsymbol{H}_{in} = \left[\tfrac{1}{r_{i1n}}\boldsymbol{x}_{i1} \ \cdots \ \tfrac{1}{r_{iJn}}\boldsymbol{x}_{iJ}\right], \quad (52)$$

$$\tilde{\boldsymbol{q}}_{in} = \left[\tilde{q}_{i1n} \ \cdots \ \tilde{q}_{iJn}\right]^\top = \boldsymbol{H}_{in}^{\mathsf{H}}\tilde{\boldsymbol{w}}_{in}, \quad (53)$$

$$\tilde{\boldsymbol{Q}}_{in} = \begin{bmatrix} \|\tilde{\boldsymbol{q}}_{in}\|^2 & -\tilde{q}_{i1n}\tilde{q}_{i2n}^* & \cdots & -\tilde{q}_{i1n}\tilde{q}_{iJn}^* \\ -\tilde{q}_{i2n}\tilde{q}_{i1n}^* & \|\tilde{\boldsymbol{q}}_{in}\|^2 & \cdots & -\tilde{q}_{i2n}\tilde{q}_{iJn}^* \\ \vdots & \vdots & \ddots & \vdots \\ -\tilde{q}_{iJn}\tilde{q}_{i1n}^* & -\tilde{q}_{iJn}\tilde{q}_{i2n}^* & \cdots & \|\tilde{\boldsymbol{q}}_{in}\|^2 \end{bmatrix}, \quad (54)$$

$$\boldsymbol{G}_{in} = \frac{1}{\sqrt{J\sum_j |\tilde{q}_{ijn}|^4}}\boldsymbol{H}_{in}\tilde{\boldsymbol{Q}}_{in}\boldsymbol{H}_{in}^{\mathsf{H}}, \quad (55)$$

where $\tilde{\boldsymbol{w}}_{in}$ is an auxiliary variable and the equality of (51) holds when $\boldsymbol{w}_{in} = \tilde{\boldsymbol{w}}_{in}$. Since $g_{in}(\boldsymbol{w}_{in}) = (\boldsymbol{w}_{in}^{\mathsf{H}}\boldsymbol{G}_{in}\boldsymbol{w}_{in})^2$ is a differentiable, convex, and homogeneous function of $\boldsymbol{w}_{in}$, we can apply GIP-HSM to minimize $\mathcal{J}^+$. The optimal condition for the direction of $\boldsymbol{w}_{in}$ is determined as

$$\frac{\partial g_{in}}{\partial \boldsymbol{w}_{in}^{\mathsf{H}}}(\boldsymbol{w}'_{in}) = ({\boldsymbol{w}'_{in}}^{\mathsf{H}}\boldsymbol{G}_{in}\boldsymbol{w}'_{in})\boldsymbol{G}_{in}\boldsymbol{w}'_{in} \parallel \boldsymbol{W}_i^{-1}\boldsymbol{e}_n. \quad (56)$$

Since $({\boldsymbol{w}'_{in}}^{\mathsf{H}}\boldsymbol{G}_{in}\boldsymbol{w}'_{in})$ is a scalar, one of the solutions of (56) is

$$\boldsymbol{w}'_{in} = \boldsymbol{G}_{in}^{-1}\boldsymbol{W}_i^{-1}\boldsymbol{e}_n. \quad (57)$$

Substituting $\tilde{\boldsymbol{w}}_{in} = \boldsymbol{w}_{in}$ into (53)–(55), we obtain the following update rule for optimizing the direction of $\boldsymbol{w}_{in}$:

$$\boldsymbol{H}_{in} = \left[\tfrac{1}{r_{i1n}}\boldsymbol{x}_{i1} \ \cdots \ \tfrac{1}{r_{iJn}}\boldsymbol{x}_{iJ}\right], \quad (58)$$

$$\boldsymbol{q}_{in} = \left[q_{i1n} \ \cdots \ q_{iJn}\right]^\top = \boldsymbol{H}_{in}^{\mathsf{H}}\boldsymbol{w}_{in}, \quad (59)$$

$$\boldsymbol{Q}_{in} = \begin{bmatrix} \|\boldsymbol{q}_{in}\|^2 & -q_{i1n}q_{i2n}^* & \cdots & -q_{i1n}q_{iJn}^* \\ -q_{i2n}q_{i1n}^* & \|\boldsymbol{q}_{in}\|^2 & \cdots & -q_{i2n}q_{iJn}^* \\ \vdots & \vdots & \ddots & \vdots \\ -q_{iJn}q_{i1n}^* & -q_{iJn}q_{i2n}^* & \cdots & \|\boldsymbol{q}_{in}\|^2 \end{bmatrix}, \quad (60)$$

$$\boldsymbol{G}_{in} = \frac{1}{\sqrt{J\sum_j |q_{ijn}|^4}}\boldsymbol{H}_{in}\boldsymbol{Q}_{in}\boldsymbol{H}_{in}^{\mathsf{H}}, \quad (61)$$

$$\boldsymbol{w}_{in} \leftarrow \boldsymbol{G}_{in}^{-1}\boldsymbol{W}_i^{-1}\boldsymbol{e}_n. \quad (62)$$

Finally, we operate the following scale optimization by applying (28):

$$\boldsymbol{q}_{in} = \left[q_{i1n} \ \cdots \ q_{iJn}\right]^\top = \boldsymbol{H}_{in}^{\mathsf{H}}\boldsymbol{w}_{in}, \quad (63)$$

$$\boldsymbol{w}_{in} \leftarrow \boldsymbol{w}_{in}\sqrt[4]{J/(2\textstyle\sum_j |q_{ijn}|^4)}, \quad (64)$$

which is the scale optimization w.r.t. the f_{in}-norm.

In GGD-ILRMA with $\beta = 4$, the demixing matrix $\boldsymbol{W}_i$ is updated by (58)–(64), and the low-rank models $\boldsymbol{T}_n$ and $\boldsymbol{V}_n$ are updated by (16) and (17), respectively. These update rules are derived using the MM algorithm and GIP-HSM, thus guaranteeing a monotonic decrease of the cost function (30).

V. Experimental Evaluation

A. BSS Experiment on Music Signals

We compared the separation performance of the proposed sub-Gaussian GGD-ILRMA ($\beta = 4$) with those of conventional IS-ILRMA [9] and GGD-ILRMA ($\beta < 2$) [16]. We artificially produced monaural dry music sources of four melody parts (melody 1: main melody, melody 2: counter melody, midrange, and bass) using Microsoft GS Wavetable Synth, where several musical instruments were chosen to play these melody parts [20], [21]. Six combinations of sources, Music 1–Music 6, were constructed by selecting typical combinations of instruments with different melody parts. The combinations of dry sources used in this experiment are shown in Table I. To simulate a reverberant mixture, the observed signals were produced by convoluting the impulse response E2A, which was obtained from the RWCP database [22], shown in Fig. 3. As the evaluation score, we used the improvement of the signal-to-distortion ratio (SDR) [23], which indicates the overall separation quality. An STFT was performed using a 128-ms-long Hamming window with a 64-ms-long shift. The shape parameter β of the GGD was set to 1, 1.99, 2 in conventional GGD-ILRMA and 4 in the proposed method, where $\beta = 1.99$ is the best parameter according to [16]. The other conditions are shown in Table II.

Fig. 4 shows the average SDR improvements for Music 1–Music 6. The proposed sub-Gaussian GGD-ILRMA on average outperforms the other conventional ILRMAs. This

TABLE I
COMBINATIONS OF DRY SOURCES

Index	Source 1	Source 2
Music 1	Fg. (bass)	Ob. (melody 1)
Music 2	Fg. (bass)	Tp. (melody 1)
Music 3	Ob. (melody 1)	Fl. (melody 2)
Music 4	Pf. (midrange)	Ob. (melody 1)
Music 5	Pf. (midrange)	Tp. (melody 1)
Music 6	Tp. (melody 1)	Fl. (melody 2)

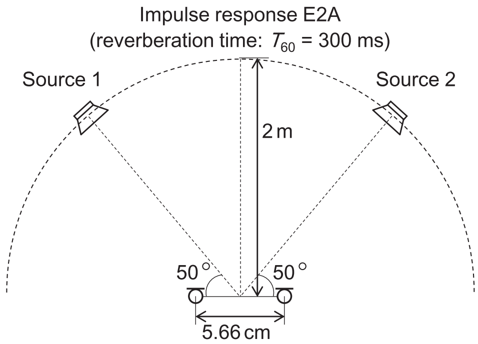


Fig. 3. Recording conditions of impulse response E2A ($T_{60} = 300\,\mathrm{ms}$) obtained from RWCP database [22].

result confirms that the proposed sub-Gaussian source model is more appropriate for dealing with music signals than other conventional (super-)Gaussian source models.

B. BSS Experiment on Speech Signals

We also confirmed the separation performance of the proposed sub-Gaussian GGD-ILRMA for speech signals, which are less likely to follow a sub-Gaussian distribution than music signals. We used the monaural dry speech sources from the source separation task in SiSEC2011 [24], Speech 1–Speech 4. An STFT was performed using a 256-ms-long Hamming window with a 128-ms-long shift. The other conditions were the same as those of the music source separation experiment.

Fig. 5 shows the average SDR improvements for Speech 1–Speech 4. The proposed GGD-ILRMA on average outperforms the other conventional ILRMAs even for speech signals, which are expected to be sparse and follow super-Gaussian distributions. This shows that the proposed time-variant sub-Gaussian model can appropriately model super-Gaussian signals as well as sub-Gaussian signals owing to its time-variant property, as described in Sect. IV-A.

VI. CONCLUSION

We proposed a new type of ILRMA, which assumes that the source signal follows the time-variant isotropic complex sub-Gaussian GGD. By using a new update scheme called GIP-HSM, we obtained a convergence-guaranteed update rule for the demixing matrix. Furthermore, in the experimental evaluation, we revealed the versatility of the proposed method, i.e., the proposed time-variant sub-Gaussian source model can deal with various types of source signal, ranging from sub-Gaussian music signals to super-Gaussian speech signals.

TABLE II
EXPERIMENTAL CONDITIONS FOR MUSIC AND SPEECH SOURCE SEPARATION

Sampling frequency	16 kHz
Number of iterations	1000
Number of bases	20
Number of trials	10
Domain parameter	$p = 0.5$
Initial demixing matrix $\boldsymbol{W}_i$	identity matrix
Entries of initial source model matrices $\boldsymbol{T}_n$ and $\boldsymbol{V}_n$	uniformly distributed random values

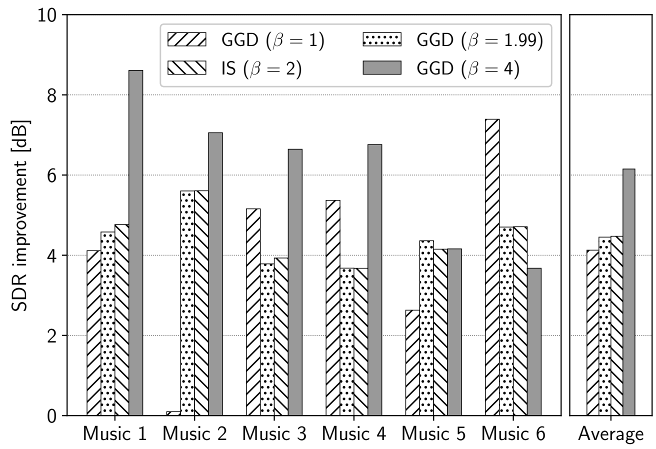


Fig. 4. Average SDR improvement of GGD-ILRMA ($\beta = 1$), GGD-ILRMA ($\beta = 1.99$), IS-ILRMA, and proposed GGD-ILRMA ($\beta = 4$) for six music signals. "Average" bar graph on right side shows average SDR improvement among six music signals for each method.

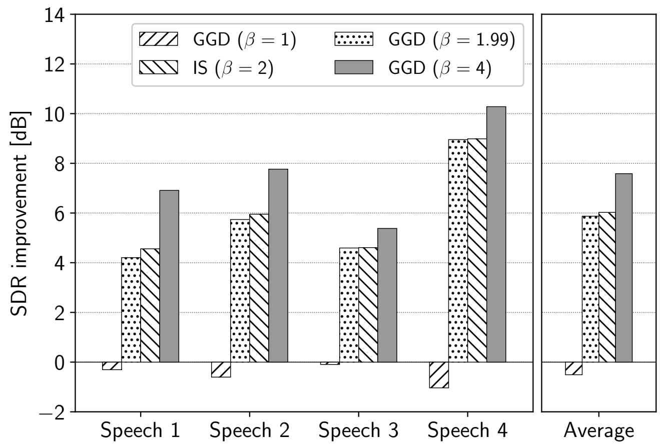


Fig. 5. Average SDR improvement of GGD-ILRMA ($\beta = 1$), GGD-ILRMA ($\beta = 1.99$), IS-ILRMA, and proposed GGD-ILRMA ($\beta = 4$) for four speech signals. "Average" bar graph on right side shows average SDR improvement among four speech signals for each method.

ACKNOWLEDGMENT

This work was partly supported by SECOM Science and Technology Foundation and JSPS KAKENHI Grant Numbers JP17H06572 and JP16H01735.

REFERENCES

[1] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, 1994.
[2] P. Smaragdis, "Blind separation of convolved mixtures in the frequency domain," *Neurocomputing*, vol. 22, no. 1, pp. 21–34, 1998.
[3] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," *IEEE Trans. ASLP*, vol. 12, no. 5, pp. 530–538, 2004.
[4] H. Saruwatari, T. Kawamura, T. Nishikawa, A. Lee, and K. Shikano, "Blind source separation based on a fast-convergence algorithm combining ICA and beamforming," *IEEE Trans. ASLP*, vol. 14, no. 2, pp. 666–678, 2006.
[5] A. Hiroe, "Solution of permutation problem in frequency domain ICA using multivariate probability density functions," in *Proc. ICA*, 2006, pp. 601–608.
[6] T. Kim, H. T. Attias, S.-Y. Lee, and T.-W. Lee, "Blind source separation exploiting higher-order frequency dependencies," *IEEE Trans. ASLP*, vol. 15, no. 1, pp. 70–79, 2007.
[7] N. Ono, "Stable and fast update rules for independent vector analysis based on auxiliary function technique," in *Proc. WASPAA*, 2011, pp. 189–192.
[8] K. Yatabe and D. Kitamura, "Determined blind source separation via proximal splitting algorithm," in *Proc. ICASSP*, 2018, pp. 776–780.
[9] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, "Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization," *IEEE/ACM Trans. ASLP*, vol. 24, no. 9, pp. 1626–1641, 2016.
[10] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, "Determined blind source separation with independent low-rank matrix analysis," in *Audio Source Separation*, S. Makino, Ed. Springer, Cham, March 2018, ch. 6, pp. 125–155.
[11] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
[12] A. Ozerov and C. Févotte, "Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation," *IEEE Trans. ASLP*, vol. 18, no. 3, pp. 550–563, 2010.
[13] A. Ozerov, C. Févotte, R. Blouet, and J. L. Durrieu, "Multichannel nonnegative tensor factorization with structured constraints for user-guided audio source separation," in *Proc. ICASSP*, 2011, pp. 257–260.
[14] H. Sawada, H. Kameoka, S. Araki, and N. Ueda, "Multichannel extensions of non-negative matrix factorization with complex-valued data," *IEEE Trans. ASLP*, vol. 21, no. 5, pp. 971–982, 2013.
[15] S. Mogami, D. Kitamura, Y. Mitsui, N. Takamune, H. Saruwatari, and N. Ono, "Independent low-rank matrix analysis based on complex Student's t-distribution for blind audio source separation," in *Proc. MLSP*, 2017.
[16] D. Kitamura, S. Mogami, Y. Mitsui, N. Takamune, H. Saruwatari, N. Ono, Y. Takahashi, and K. Kondo, "Generalized independent low-rank matrix analysis using heavy-tailed distributions for blind source separation," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 28, pp. 1–25, 2018.
[17] R. Ikeshita and Y. Kawaguchi, "Independent low-rank matrix analysis based on multivariate complex exponential power distribution," in *Proc. ICASSP*, 2018, pp. 741–745.
[18] G. R. Naik and W. Wang, "Audio analysis of statistically instantaneous signals with mixed Gaussian probability distributions," *International Journal of Electronics*, vol. 99, no. 10, pp. 1333–1350, 2012.
[19] D. R. Hunter and K. Lange, "Quantile regression via an MM algorithm," *J. Comput. Graph. Stat.*, vol. 9, no. 1, pp. 60–77, 2000.
[20] D. Kitamura, H. Saruwatari, H. Kameoka, Y. Takahashi, K. Kondo, and S. Nakamura, "Multichannel signal separation combining directional clustering and nonnegative matrix factorization with spectrogram restoration," *IEEE Trans. ASLP*, vol. 23, no. 4, pp. 654–669, 2015.
[21] D. Kitamura, "Open dataset: songkitamura," http://d-kitamura.net/en/dataset_en.htm. Accessed 27 May 2018.
[22] S. Nakamura, K. Hiyane, F. Asano, T. Nishiura, and T. Yamada, "Acoustical sound database in real environments for sound scene understanding and hands-free speech recognition," in *Proc. LREC*, 2000, pp. 965–968.
[23] E. Vincent, R. Gribonval, and C. Févotte, "Performance measurement in blind audio source separation," *IEEE Trans. ASLP*, vol. 14, no. 4, pp. 1462–1469, 2006.
[24] S. Araki, F. Nesta, E. Vincent, Z. Koldovský, G. Nolte, A. Ziehe, and A. Benichoux, "The 2011 signal separation evaluation campaign (SiSEC2011): - audio source separation -," in *Proc. LVA/ICA*, 2012, pp. 414–422.

APPENDIX

A. Derivation of Update Rule for Low-Rank Source Model

The update rules for $\boldsymbol{T}_n$ and $\boldsymbol{V}_n$ in GGD-ILRMA can be derived by the MM algorithm. In the MM algorithm, we minimize the majorization function instead of the original cost function. To derive the majorization function in GGD-ILRMA, we introduce Jensen's inequality

$$\left(\sum_k t_{ikn}v_{kjn}\right)^{-\frac{\beta}{p}} \leq \sum_k \phi_{ijnk}\left(\frac{t_{ikn}v_{kjn}}{\phi_{ijnk}}\right)^{-\frac{\beta}{p}} \tag{65}$$

and the tangent-line inequality

$$\log\sum_k t_{ikn}v_{kjn} \leq \frac{1}{\psi_{ijn}}\left(\sum_k t_{ikn}v_{kjn} - 1\right) + \log\psi_{ijn}, \tag{66}$$

where $\phi_{ijnk} > 0$ and $\psi_{ijn} > 0$ are auxiliary variables and ϕ_{ijnk} satisfies $\sum_k \phi_{ijnk} = 1$. The equalities of (65) and (66) hold if and only if

$$\phi_{ijnk} = \frac{t_{ikn}v_{kjn}}{\sum_{k'} t_{ik'n}v_{k'jn}}, \tag{67}$$

$$\psi_{ijn} = \sum_k t_{ikn}v_{kjn}, \tag{68}$$

respectively. By substituting (8) into (9) and applying (65) and (66) to (9), the majorization function of (9) can be designed as

$$\mathcal{L}_{\mathrm{GGD}} \leq \sum_{i,j,n,k}\left(\frac{\phi_{ijnk}^{\frac{\beta}{p}+1}|y_{ijn}|^{\beta}}{(t_{ikn}v_{kjn})^{\frac{\beta}{p}}} + \frac{2t_{ikn}v_{kjn}}{p\psi_{ijn}}\right) + \mathrm{const.}, \tag{69}$$

where the constant term is independent of t_{ikn} and v_{kjn}. By setting the partial derivatives of (69) w.r.t. t_{ikn} and v_{kjn} to zero, we obtain the update rules (16) and (17), respectively.