Laplace Nonnegative Matrix Factorization with Application to Semi-supervised Audio Denoising

Hiroki Tanji*, Takahiro Murakami*†, and Hiroyuki Kamata*
* *Department of Electronics and Bioinformatics, School of Science and Technology,* Meiji University, Japan
† *Centre for Vision, Speech and Signal Processing (CVSSP),* University of Surrey, UK
{htanji, tmrkm, kamata}@meiji.ac.jp

***Abstract*—This paper proposes two statistical models for the nonnegative matrix factorization (NMF) based on heavy-tailed distributions. In the NMF for acoustic signals, previous works justify the additivity of an observed spectrogram using the reproductive property of a probability density function. However, the effectiveness of these properties is not clear. Consequently, to construct a model robust to noise, statistical models based on heavy-tailed distributions are recently growing up. In this paper, as heavy-tailed models for the NMF, we introduce statistical models based on the complex Laplace distributions, and call them Laplace-NMF. Moreover, we derive convergence-guaranteed optimization algorithms to estimate parameters. From our formulation, a statistical interpretation of the Itakura-Saito (IS) divergence-based NMF is newly revealed. We confirm the effectiveness of Laplace-NMF in semi-supervised audio denoising.**

***Index Terms*—complex Laplace distribution, nonnegative matrix factorization, majorization-minimization algorithm, source separation**

I. Introduction

The objective of the nonnegative matrix factorization (NMF) [1] is to decompose the nonnegative observed matrix $\boldsymbol{Y} \in \mathbb{R}_+^{M\times N}$ ($\mathbb{R}_+ = [0,\infty)$) into two nonnegative matrices $\boldsymbol{W} \in \mathbb{R}_+^{M\times K}$ and $\boldsymbol{H} \in \mathbb{R}_+^{K\times N}$ as $\boldsymbol{Y} \simeq \boldsymbol{W}\boldsymbol{H}$. In general, a value less than $\min(M,N)$ is used for K to obtain a low-rank approximation of $\boldsymbol{Y}$. We can obtain an approximation of $\boldsymbol{Y}$ by minimizing a divergence between $\boldsymbol{Y}$ and $\boldsymbol{W}\boldsymbol{H}$. In general, $\boldsymbol{W}$ and $\boldsymbol{H}$ are estimated using multiplicative update rules based on majorization-minimization (MM) algorithms [2].

In acoustic signal processing applications (e.g. music transcription [3], [4] and audio denoising [5]–[7]), an amplitude or a power spectrogram is used as the observed matrix. The observed spectrogram is decomposed into a set of spectra and time-varying activation weights corresponding to $\boldsymbol{W}$ and $\boldsymbol{H}$ [3].

To obtain $\boldsymbol{W}$ and $\boldsymbol{H}$ from the observed spectrogram, the generalized Kullback-Leibler (KL) divergence [1] is often used for the divergence in the NMF [7]–[9]. From a statistical perspective, we regard the NMF based on the generalized KL divergence (KL-NMF) as the statistical model that assumes each element of the observed matrix follows the Poisson distribution [10]. However, this assumption is not appropriate for audio signals because the elements of the observed spectrogram are continuous variables.

This work was partially supported by Grant-in-Aid for JSPS Fellows (KAKENHI 18J14238).

Towards appropriate statistical models for the NMF, recently, the NMF based on the α-stable distribution [11] (α-stable-NMF) has been investigated [12]. Because the α-stable distribution has the reproductive property in the sum of random variables, α-stable-NMF justifies the assumption on the additivity of fractional power spectra. α-stable-NMF includes the NMF based on the complex Gaussian (Gaussian-NMF [4]), the complex Cauchy (Cauchy-NMF [13]), and the Lévy (Lévy-NMF [14]) distributions for special cases. In particular, due to the heavy-tailed nature of the α-stable distribution, Cauchy-NMF and Lévy-NMF are effective in audio denoising when the observed spectrogram is corrupted by impulsive noise [13], [14]. However, it is not clear whether the reproductive property of the α-stable distribution for α-stable-NMF is effective in the denoising and the signal separation tasks.

As a statistical model based on a heavy-tailed distribution, the NMF based on the Student's t distribution (t-NMF) has been proposed [15]. The complex Student's t distribution is a generalization of the complex Cauchy and the complex Gaussian distributions. Thus, t-NMF is equivalent to Cauchy-NMF and Gaussian-NMF, when the number of degrees of freedom ν is $\nu = 1$ and $\nu \to \infty$, respectively. Although the reproductive property of the complex Student's t distribution does not hold for any ν, t-NMF provides promising results in signal separation [15].

As alternatives to the conventional models based on the heavy-tailed distributions, in this paper, we introduce Laplace-NMF, which utilizes the complex Laplace distributions. The complex Laplace distributions are derived by integrating the variance of the complex Gaussian distribution out using the exponential and the gamma distribution [16]–[18]. The probability density function (PDF) includes the modified Bessel function of the second kind when the exponential distribution is used [16], [17]. Also, the exponential function appears in the PDF when the gamma distribution is used [18]. The PDFs have heavier tails and sharper peaks than the complex Gaussian distribution. Nevertheless, the tails of the complex Laplace distributions are not as heavy as the α-stable distribution. This feature will bring to Laplace-NMF a characteristic that the reconstructed matrix fits to the observed spectrogram accurately compared with the conventional heavy-tailed models. The complex Laplace distributions cannot model the additivities of the amplitude and the power spectrograms. However, Laplace-NMF is appropriate as a statistical approach for modeling of

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

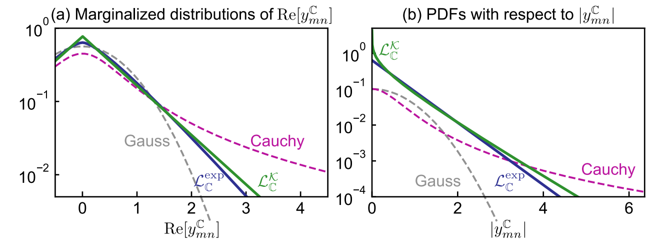

Fig. 1. Probability density functions used for the NMF.

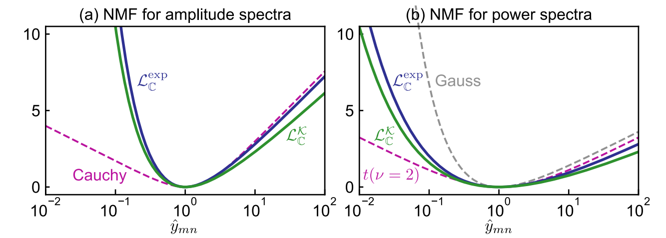

Fig. 2. Cost functions for $y_{mn} = 1$ in statistical models for the NMF.

acoustic signals.

In this study, we introduce two cost functions based on the complex Laplace distributions. Moreover, we discover a new statistical interpretation of the NMF based on the Itakura-Saito (IS) divergence (IS-NMF [4]). To estimate parameters of Laplace-NMF, using the MM algorithm, we derive convergence-guaranteed optimization algorithms. We evaluate the performances of Laplace-NMF in fitting synthetic datasets and semi-supervised audio denoising [5], [6].

II. Statistical models of Laplace-NMF

In this section, we formulate the NMF using the complex Laplace distributions.

A. Review of the complex Laplace distributions

The complex Laplace distributions are scale mixtures of Gaussians that have the form $p(x|\mu) = \int_{\mathbb{R}_+} \mathcal{N}_{\mathbb{C}}(x|\mu, \sigma^2)\, p_{\sigma^2}(\sigma^2) d\sigma^2$, where x is a complex-valued random variable, $\mathcal{N}_{\mathbb{C}}(x|\mu, \sigma^2)$ is the complex Gaussian distribution with mean μ and variance σ^2, and p_{σ^2} is any function. In [16], [17], the exponential distribution with scale λ^2 is used for $p_{\sigma^2}(\sigma^2)$. Then the complex Laplace distribution $p(x|\mu = 0) \triangleq \mathcal{L}^{\mathcal{K}}_{\mathbb{C}}(x; \lambda)$ is written as

$$\mathcal{L}^{\mathcal{K}}_{\mathbb{C}}(x; \lambda) = \frac{2}{\pi\lambda^2} \mathcal{K}_0\left(\frac{2|x|}{\lambda}\right), \tag{1}$$

where $\mathcal{K}_\nu(t)$ ($t \in \mathbb{R}_+$, $\nu \in \mathbb{R}$) is the modified Bessel function of the second kind defined as $\mathcal{K}_\nu(t) = \frac{1}{2}\int_{\mathbb{R}_+} u^{-\nu-1} \exp(-\frac{t}{2}(u + u^{-1}))du$.

Another form of the complex Laplace distribution is introduced in [18]. In this literature, the gamma distribution with shape $\frac{3}{2}$ and scale λ^2 is used for $p_{\sigma^2}(\sigma^2)$. The PDF, denoted by $\mathcal{L}^{\mathrm{exp}}_{\mathbb{C}}(x; \lambda)$, is written as

$$\mathcal{L}^{\mathrm{exp}}_{\mathbb{C}}(x; \lambda) = \frac{2}{\pi\lambda^2} \exp\left(-\frac{2|x|}{\lambda}\right). \tag{2}$$

By integrating $\mathrm{Im}[x]$ out from (1) and (2), we can obtain the marginalized distributions of $\mathrm{Re}[x]$ as $\int_{\mathbb{R}} \mathcal{L}^{\mathcal{K}}_{\mathbb{C}}(x; 1) d\mathrm{Im}[x] \propto \exp(-2|\mathrm{Re}[x]|)$ and $\int_{\mathbb{R}} \mathcal{L}^{\mathrm{exp}}_{\mathbb{C}}(x; 1) d\mathrm{Im}[x] \propto |\mathrm{Re}[x]|\mathcal{K}_1(2|\mathrm{Re}[x]|)$, respectively [16].

B. Cost functions of Laplace-NMF

The cost functions of Laplace-NMF are obtained from (1) and (2). Let $y^{\mathbb{C}}_{mn}$ be the observed complex spectrogram at the mth frequency bin and the nth frame. In this paper, we construct the observed nonnegative matrix $\boldsymbol{Y} = [y_{mn}]$ using $y_{mn} = |y^{\mathbb{C}}_{mn}|$. Then y^{δ}_{mn} is approximated using $\hat{y}_{mn} = \sum_{k=1}^{K} w_{mk} h_{kn}$, where $\delta > 0$. When $\delta = 1$ and $\delta = 2$, $[y^{\delta}_{mn}]$ are the amplitude and the power spectrogram, respectively. Also, $[\hat{y}^{1/\delta}_{mn}]$ can be regarded as the estimate of the amplitude spectrogram. The cost functions of Laplace-NMF are given by the negative log-likelihood functions with (1) and (2). By substituting y_{mn} and $\lambda \hat{y}^{1/\delta}_{mn}$ into $|x|$ and λ of (1) and (2), we obtain the cost functions based on the complex Laplace distributions as

$$f_{\mathcal{L}^{\mathcal{K}}_{\mathbb{C}}}(\boldsymbol{W}, \boldsymbol{H}) \overset{c}{=} \sum_{m,n} \left[\frac{2}{\delta} \log \hat{y}_{mn} - \log \mathcal{K}_0\left(\frac{2y_{mn}}{\lambda \hat{y}^{1/\delta}_{mn}}\right)\right] \tag{3}$$

$$f_{\mathcal{L}^{\mathrm{exp}}_{\mathbb{C}}}(\boldsymbol{W}, \boldsymbol{H}) \overset{c}{=} \sum_{m,n} \left[\frac{2}{\delta} \log \hat{y}_{mn} + \frac{2y_{mn}}{\lambda \hat{y}^{1/\delta}_{mn}}\right], \tag{4}$$

where $\boldsymbol{W} = [w_{mk}]$, $\boldsymbol{H} = [h_{kn}]$, and $\overset{c}{=}$ denotes equality up to constant terms. Equation (3) is minimized when $y_{mn} = \hat{y}^{1/\delta}_{mn}, \forall m, n$. The scale λ in (3) is therefore the solution of

$$\left.\frac{\partial f_{\mathcal{L}^{\mathcal{K}}_{\mathbb{C}}}(\boldsymbol{W}, \boldsymbol{H})}{\partial \hat{y}_{mn}}\right|_{\hat{y}^{1/\delta}_{mn} = y_{mn}} = 0. \tag{5}$$

Equation (5) is simplified as $\lambda \mathcal{K}_0(2\lambda^{-1}) = \mathcal{K}_1(2\lambda^{-1})$. Thus, we can obtain λ before evaluating the cost function. In a similar way, we have $\lambda = 1$ for (4). In this paper, we refer to the NMF with the cost functions (3) and (4) as $\mathcal{L}^{\mathcal{K}}_{\mathbb{C}}$-NMF and $\mathcal{L}^{\mathrm{exp}}_{\mathbb{C}}$-NMF, respectively.

When $\delta = 1$, that is, when the amplitude spectrogram is decomposed, (4) is equivalent to the cost function of IS-NMF. In the previous work [4], to decompose power spectra, IS-NMF is derived using the statistical model based on the complex Gaussian distribution. However, the statistical model of IS-NMF for amplitude spectra [19], [20] has not been revealed. $\mathcal{L}^{\mathrm{exp}}_{\mathbb{C}}$-NMF includes IS-NMF which decomposes not power spectra but amplitude spectra.

We show the marginalized distributions of $\mathrm{Re}[y^{\mathbb{C}}_{mn}]$ and the PDFs with respect to $|y^{\mathbb{C}}_{mn}|$ for the NMF in Fig. 1(a) and (b), respectively. The Cauchy distribution has the heaviest tail in the distributions in Fig. 1. Nevertheless, the PDFs for Laplace-NMF have heavier tails than the Gaussian distribution.

Moreover, in Fig. 2, we plot the divergences between y^{δ}_{mn} and $\hat{y}_{mn}$. We can obtain the divergences for Laplace-NMF by removing the summation signs in (3) and (4). In Fig. 2(a) and (b), δ is set to $\delta = 1$ and $\delta = 2$, respectively. When $\hat{y}_{mn}$ is smaller than y^{δ}_{mn}, Laplace-NMF gives larger penalties than t-NMF with $\nu = 2$ and Cauchy-NMF.

III. Optimization of the cost functions

In this section, we propose multiplicative update rules for Laplace-NMF. Our update rules are derived using the framework of the MM algorithm [21], which minimizes an upper bound of a cost function.

A. $\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}$-NMF

The derivation of an upper bound for $\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}$-NMF is not straightforward because (3) includes the special function of $\hat{y}_{mn}$. To obtain a tractable upper bound for (3), we apply the probabilistic form of the Jensen's inequality to $\log \mathcal{L}_{\mathbb{C}}^{\mathcal{K}}(y_{mn}^{\mathbb{C}}; \hat{y}_{mn})$. We can construct $\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}(y_{mn}^{\mathbb{C}};\ \hat{y}_{mn})$ as $\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}(y_{mn}^{\mathbb{C}}; \hat{y}_{mn}) = \int_{\mathbb{R}_+} p(y_{mn}^{\mathbb{C}}|z_{mn}) p(z_{mn}; \hat{y}_{mn}) dz_{mn}$, where $p(y_{mn}^{\mathbb{C}}|z_{mn})$ is the complex Gaussian distribution with variance z_{mn} and $p(z_{mn}; \hat{y}_{mn})$ is the exponential distribution with scale $\lambda^2 \hat{y}_{mn}^{2/\delta}$. Thus, using the probabilistic form of the Jensen's inequality, the upper bound for (3) is derived as follows:

$$
\begin{aligned}
f_{\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}}(\boldsymbol{W}, \boldsymbol{H}) &\stackrel{c}{=} -\sum_{m,n} \log \mathcal{L}_{\mathbb{C}}^{\mathcal{K}}(y_{mn}^{\mathbb{C}}; \hat{y}_{mn}) \\
&\leq -\sum_{m,n} \int_{\mathbb{R}_+} p(z_{mn}|y_{mn}^{\mathbb{C}}; \tilde{\hat{y}}_{mn}) \log \frac{p(y_{mn}^{\mathbb{C}}|z_{mn}) p(z_{mn}; \hat{y}_{mn})}{p(z_{mn}|y_{mn}^{\mathbb{C}}; \tilde{\hat{y}}_{mn})} dz_{mn} \\
&\stackrel{c}{=} \sum_{m,n} \left(\frac{2}{\delta} \log \hat{y}_{mn} + \frac{\mathbb{E}_{p(z_{mn}|y_{mn}^{\mathbb{C}}; \tilde{\hat{y}}_{mn})}[z_{mn}]}{\lambda^2 \hat{y}_{mn}^{2/\delta}} \right), \qquad (6)
\end{aligned}
$$

where $\tilde{\hat{y}}_{mn}$ is the latest estimate for $\hat{y}_{mn}$, z_{mn} is the latent variable, and $\mathbb{E}_{p(z_{mn}|y_{mn}^{\mathbb{C}}; \tilde{\hat{y}}_{mn})}[z_{mn}]$ is the expectation of z_{mn} with respect to the posterior distribution $p(z_{mn}|y_{mn}^{\mathbb{C}}; \tilde{\hat{y}}_{mn})$. Moreover, for $\log \hat{y}_{mn}$ and $\hat{y}_{mn}^{-2/\delta}$ in (6), we apply the first order Taylor expansion and the Jensen's inequality written as

$$
\log \hat{y}_{mn} \leq \frac{1}{\varphi_{mn}} (\hat{y}_{mn} - \varphi_{mn}) + \log \varphi_{mn} \qquad (7)
$$

$$
\hat{y}_{mn}^{-\gamma} \leq \sum_{k=1}^{K} \frac{\rho_{mnk}^{\gamma+1}}{(w_{mk} h_{kn})^{\gamma}}, \qquad (8)
$$

where $\gamma, \varphi_{mn}, \rho_{mnk} > 0, \forall m, n, k$, and $\sum_{k=1}^{K} \rho_{mnk} = 1$. The right hands of (7) and (8) are minimized with respect to φ_{mn} and ρ_{mnk} when $\varphi_{mn} = \hat{y}_{mn}$ and $\rho_{mnk} = w_{mk} h_{kn} / \hat{y}_{mn}$, respectively. By substituting (7) and (8) into (6), we obtain the upper bound $f_{\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}}^{+}(\boldsymbol{W}, \boldsymbol{H}, \boldsymbol{\varphi}, \boldsymbol{\rho})$ for $f_{\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}}(\boldsymbol{W}, \boldsymbol{H})$ as

$$
\begin{aligned}
f_{\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}}^{+}(\boldsymbol{W}, \boldsymbol{H}, \boldsymbol{\varphi}, \boldsymbol{\rho}) \stackrel{c}{=} \sum_{m,n} \Bigg[& \frac{2}{\delta} \left\{ \frac{1}{\varphi_{mn}} (\hat{y}_{mn} - \varphi_{mn}) + \log \varphi_{mn} \right\} \\
& + \frac{\zeta_{\delta mn}}{\lambda^2} \sum_{k=1}^{K} \frac{\rho_{mnk}^{2/\delta+1}}{(w_{mk} h_{kn})^{2/\delta}} \Bigg], \qquad (9)
\end{aligned}
$$

where $\boldsymbol{\varphi} = \{\varphi_{mn}\}$, $\boldsymbol{\rho} = \{\rho_{mnk}\}$, and $\zeta_{\delta mn} = \mathbb{E}_{p(z_{mn}|y_{mn}^{\mathbb{C}}; \tilde{\hat{y}}_{mn})}[z_{mn}]$. The minimization problems of $f_{\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}}^{+}(\boldsymbol{W}, \boldsymbol{H}, \boldsymbol{\varphi}, \boldsymbol{\rho})$ and $f_{\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}}(\boldsymbol{W}, \boldsymbol{H})$ are equivalent. By using the partial derivatives of $f_{\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}}^{+}(\boldsymbol{W}, \boldsymbol{H}, \boldsymbol{\varphi}, \boldsymbol{\rho})$ with respect to w_{mk} and h_{kn}, we obtain the update rules for $\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}$-NMF as

$$
w_{mk} \leftarrow w_{mk} \left(\frac{\sum_n \frac{\zeta_{\delta mn}}{\lambda^2 \hat{y}_{mn}^{2/\delta+1}} h_{kn}}{\sum_n h_{kn} / \hat{y}_{mn}} \right)^{\delta/(\delta+2)} \qquad (10)
$$

$$
h_{kn} \leftarrow h_{kn} \left(\frac{\sum_m \frac{\zeta_{\delta mn}}{\lambda^2 \hat{y}_{mn}^{2/\delta+1}} w_{mk}}{\sum_m w_{mk} / \hat{y}_{mn}} \right)^{\delta/(\delta+2)} \qquad (11)
$$

$$
\zeta_{\delta mn} = \lambda y_{mn} \hat{y}_{mn}^{\frac{1}{\delta}} \mathcal{K}_1 \left(\frac{2 y_{mn}}{\lambda \hat{y}_{mn}^{1/\delta}} \right) \mathcal{K}_0 \left(\frac{2 y_{mn}}{\lambda \hat{y}_{mn}^{1/\delta}} \right)^{-1}. \qquad (12)
$$

Equation (12) is derived from the first order moment of the generalized inverse Gaussian distribution [17].

B. $\mathcal{L}_{\mathbb{C}}^{\exp}$-NMF

Update rules for $\mathcal{L}_{\mathbb{C}}^{\exp}$-NMF are also derived using the MM algorithm. The upper bound $f_{\mathcal{L}_{\mathbb{C}}^{\exp}}^{+}(\boldsymbol{W}, \boldsymbol{H}, \boldsymbol{\varphi}, \boldsymbol{\rho})$ for (7) is given by substituting (7) and (8) into (4) as

$$
\begin{aligned}
f_{\mathcal{L}_{\mathbb{C}}^{\exp}}^{+}(\boldsymbol{W}, \boldsymbol{H}, \boldsymbol{\varphi}, \boldsymbol{\rho}) = \sum_{m,n} \Bigg[& \frac{2}{\delta} \left\{ \frac{1}{\varphi_{mn}} (\hat{y}_{mn} - \varphi_{mn}) + \log \varphi_{mn} \right\} \\
& + 2 y_{mn} \sum_{k=1}^{K} \frac{\rho_{mnk}^{1/\delta+1}}{(w_{mk} h_{kn})^{1/\delta}} \Bigg]. \qquad (13)
\end{aligned}
$$

The update rules for $\mathcal{L}_{\mathbb{C}}^{\exp}$-NMF are derived by using partial derivatives of $f_{\mathcal{L}_{\mathbb{C}}^{\exp}}^{+}(\boldsymbol{W}, \boldsymbol{H}, \boldsymbol{\varphi}, \boldsymbol{\rho})$ with respect to w_{mk} and h_{kn}, and they are written as

$$
w_{mk} \leftarrow w_{mk} \left(\frac{\sum_n \frac{y_{mn}}{\hat{y}_{mn}^{1/\delta+1}} h_{kn}}{\sum_n h_{kn} / \hat{y}_{mn}} \right)^{\delta/(\delta+1)} \qquad (14)
$$

$$
h_{kn} \leftarrow h_{kn} \left(\frac{\sum_m \frac{y_{mn}}{\hat{y}_{mn}^{1/\delta+1}} w_{mk}}{\sum_m w_{mk} / \hat{y}_{mn}} \right)^{\delta/(\delta+1)}. \qquad (15)
$$

When $\delta = 1$, (14) and (15) are equivalent to the update rules based on the MM algorithm for IS-NMF [22].

IV. Evaluation

To evaluate performances of Laplace-NMF, we applied Laplace-NMF to fitting random data and semi-supervised audio denoising [5], [6].

A. Fitting synthetic data

We performed the simulation to evaluate the data fitting abilities of Laplace-NMF.

We generated impulsive data and random data following the complex Laplace distributions. To generate impulsive data, we constructed w_{mk} and h_{kn} using 4th power of the Gaussian random noise. $y_{mn}^{\mathbb{C}}$ was then drawn from the complex Student's t distribution with degrees of freedom $\nu = 5$ and the scale $(\sum_{k=1}^{K} w_{mk} h_{kn})^{1/2}$. For Laplace distributed-data, w_{mk} was drawn from the gamma distribution with shape $\eta + 1$ and unit scale. Also, h_{kn} was drawn from the beta distribution with shapes $\alpha = \eta$ and $\beta = 1$. Using w_{mk} and h_{kn}, $y_{mn}^{\mathbb{C}}$ was generated from the complex Gaussian distribution with zero mean and the variance $\sum_{k=1}^{K} w_{mk} h_{kn}$. The shape parameter

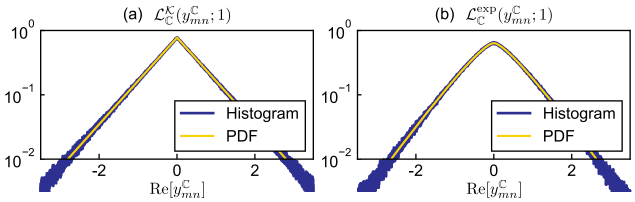

Fig. 3. Histograms of generated random data.

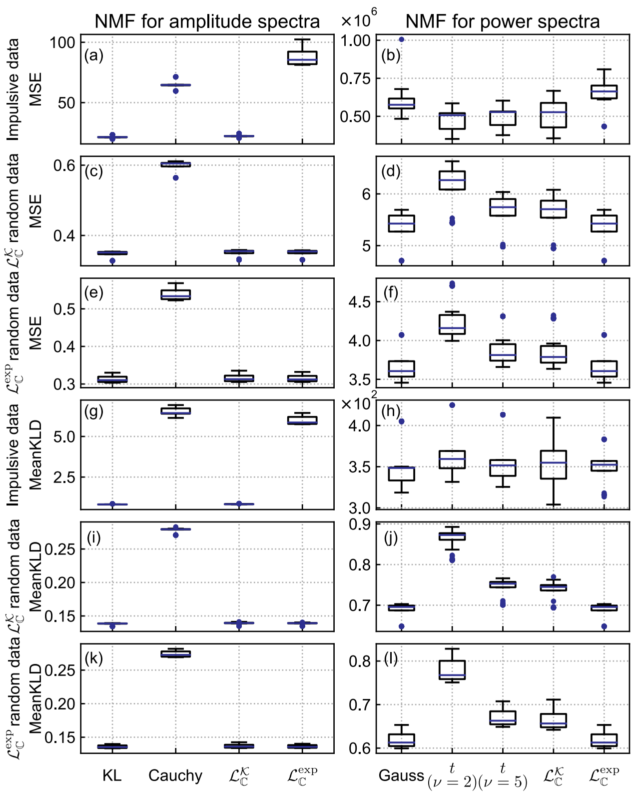

Fig. 4. The box plots of reconstruction qualities in fitting synthetic data. Lower is better.

was set to $\eta = 1/K$ and $\eta = 3/(2K)$ for $\mathcal{L}_{\mathbb{C}}^{\mathcal{K}}(y_{mn}^{\mathbb{C}};\lambda=1)$ and $\mathcal{L}_{\mathbb{C}}^{\exp}(y_{mn}^{\mathbb{C}};\lambda=1)$, respectively. We show the histograms of $\mathrm{Re}[y_{mn}^{\mathbb{C}}]$ of the generated data following the complex Laplace distributions in Fig. 3.

In this simulation, the size of the observed matrix was 400×500, and the number of bases K was set to 5. We constructed the observed matrix using $|y_{mn}^{\mathbb{C}}|^{\delta}$. When $\delta = 1$, Laplace-NMF was compared with KL-NMF [1] and Cauchy-NMF [13]. Moreover, when $\delta = 2$, Gaussian-NMF [4] and t-NMF [15] with $\nu = 2, 5$ were used to compare. $\boldsymbol{W}$ and $\boldsymbol{H}$ were initialized using random values. Then they were updated 500 times. We used the naive multiplicative update method [13] for Cauchy-NMF. For the others, we used the MM algorithms. Each algorithm performed 100 trials using 20 initial values and 5 observed matrices. To evaluate data fitting abilities, we used the mean squared error (MSE) defined as $\mathrm{MSE} = \frac{1}{MN}\sum_{m,n}(|y_{mn}^{\mathbb{C}}|^{\delta} - \hat{y}_{mn})^2$ and the mean of the generalized KL divergence.

In Fig. 4, we show the performance indices. In this figure, "MeanKLD" indicates the mean of the generalized KL divergence. When the random data follows the complex Laplace distribution, the fitting abilities of Laplace-NMF are better than those of Cauchy-NMF and t-NMF. However, the fitting ability of $\mathcal{L}_{\mathbb{C}}^{\exp}$-NMF deteriorates in fitting impulsive data. This is because $\mathcal{L}_{\mathbb{C}}^{\exp}$-NMF makes the approximations of outliers using larger values.

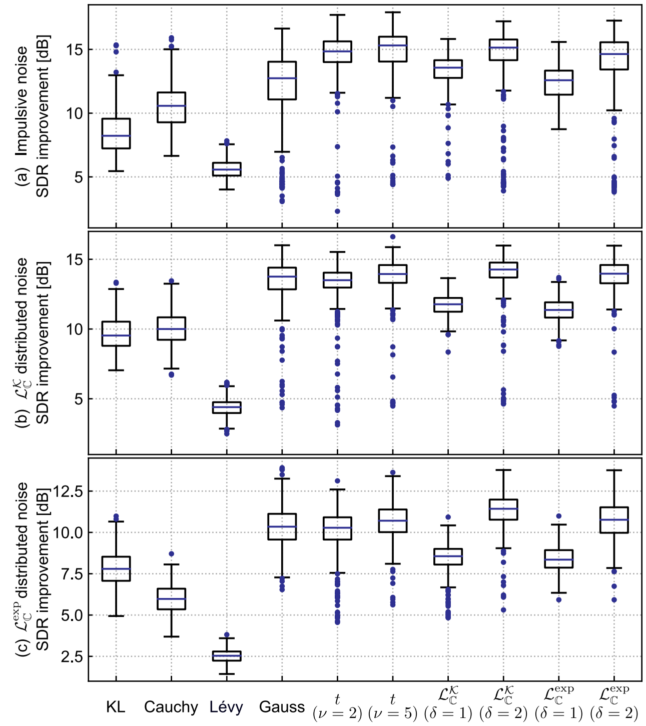

Fig. 5. Evaluation results in semi-supervised denoising. Higher is better.

B. Semi-supervised audio denoising

We evaluated performances of Laplace-NMF in semi-supervised audio denoising. In this simulation, the source signal was corrupted by the 0 [dB] of synthetic noise generated using the procedure in Sect. IV-A.

To reduce noise, we utilized the semi-supervised NMF [23]. Our semi-supervised denoising procedure consisted of two stages: the training and the denoising stages. In the training stage, the NMF was applied to the silent period of the source signal to obtain a basis matrix of noise $\boldsymbol{W}_{\mathrm{n}}$. Then, in the denoising stage, we constructed the basis matrix $\boldsymbol{W}$ with a basis matrix of source $\boldsymbol{W}_{\mathrm{s}}$ as $\boldsymbol{W} = [\boldsymbol{W}_{\mathrm{n}}, \boldsymbol{W}_{\mathrm{s}}]$. $\boldsymbol{W}$ and the weight matrix $\boldsymbol{H}$ were learned from the observed spectrogram. $\boldsymbol{W}_{\mathrm{s}}$ and $\boldsymbol{H}$ were updated using the update rules for the NMF while $\boldsymbol{W}_{\mathrm{n}}$ was fixed. Using the generalized Wiener filter [24], we estimated the source signal from $\boldsymbol{W}_{\mathrm{s}}$ and the corresponding rows of $\boldsymbol{H}$.

For the source signal, we used 6 recordings played on electric guitar in the IDMT-SMT-GUITAR database [25]. We extracted `AR_Lick[1-6]_KN.wav` from this database. We then resampled them to 11025 [Hz]. We obtained the complex spectrograms of the source signal using the short-time Fourier transform (STFT). For the STFT, the frame length was set to 512, the fast Fourier transform (FFT) length was 1024, and the hop size was 128. Before performing the FFT, the signal was Hamming-windowed. The Nyquist frequency and the DC components of the FFT spectra were excluded because they do not follow any complex distributions.

In this simulation, the MM algorithm for Lévy-NMF [14] was also used to compare. We performed each algorithm with 20 i.i.d. random initial values for $\boldsymbol{W}_{\mathrm{n}}$, $\boldsymbol{W}_{\mathrm{s}}$, and $\boldsymbol{H}$ and 5 i.i.d. noise signals. The numbers of bases were set to 5 and 10 for $\boldsymbol{W}_{\mathrm{n}}$ and $\boldsymbol{W}_{\mathrm{s}}$, respectively. The numbers of iterations were set to 500 and 700 for the training and the denoising stages, respectively. We evaluated the denoising performances using the improvement of the source-to-distortion ratio (SDR) [26]. The SDRs were calculated using "mir_eval" [27] including Python implementation of the BSS Eval Matlab toolbox [26]. The silent period used in the training stage is excluded for evaluation.

The evaluation results are shown in Fig. 5. The performances of NMF decomposing amplitude spectra are not outstanding. This suggests that using amplitude spectra is not effective for this task. In Fig. 5(a), t-NMF with $\nu = 5$ demonstrates the best performance in terms of the median of SDR improvement because the noise signal is impulsive and generated from the Student's t distribution. Nevertheless, as shown in Figs. 5(b) and (c), when the noise signal follows the complex Laplace distribution, Laplace-NMF with $\delta = 2$ shows favorable performance in terms of the median of SDR improvement.

V. Conclusions

In this paper, we have proposed two statistical models based on the complex Laplace distributions. Also, the statistical model of IS-NMF for amplitude spectra has been revealed. Moreover, we have derived the optimization algorithms for Laplace-NMF. Our experimental results have shown Laplace-NMF provides promising performances in semi-supervised audio denoising. Future works include investigating optimal values of δ for Laplace-NMF and the audio denoising task.

References

[1] D.D. Lee and H.S. Seung, "Learning the parts of objects with nonnegative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999.

[2] M. Nakano, H. Kameoka, J. Le Roux, Y. Kitano, N. Ono, and S. Sagayama, "Convergence-guaranteed multiplicative algorithms for nonnegative matrix factorization with β-divergence," in *Proc. 2010 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Aug. 2010, pp. 283–288.

[3] P. Smaragdis and J.C. Brown, "Non-negative matrix factorization for polyphonic music transcription," in *Proc. 2003 IEEE International Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct. 2003, pp. 177–180.

[4] C. Févotte, N. Bertin, and J.L. Durrieu, "Nonnegative matrix factorization with the Itakura-Saito divergence: with application to music analysis," *Neural Computation*, vol. 21, no. 3, pp. 793–830, Sep. 2008.

[5] N. Mohammadiha, P. Smaragdis, and A. Leijon, "Supervised and unsupervised speech enhancement using nonnegative matrix factorization," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 21, no. 10, pp. 2140-2151, Oct. 2013.

[6] N. Lyubimov and M. Kotov, "Non-negative matrix factorization with linear constraints for single-channel speech enhancement," *Proc. 14th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, Aug. 2013, pp. 446-450.

[7] K.W. Wilson, B. Raj, P. Smaragdis, and A. Divakaran, "Speech denoising using nonnegative matrix factorization with priors," in *Proc. 2008 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2008, pp. 4029–4032.

[8] P. Smaragdis, "Convolutive speech bases and their application to supervised speech separation," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 15, no. 1, pp. 1–12, Jan. 2007.

[9] S.A. Raczyński, N. Ono, and S. Sagayama, "Multipitch analysis with harmonic nonnegative matrix approximation," in *Proc. 8th International Conference on Music Information Retrieval (ISMIR)*, Sep. 2007, pp. 381–386.

[10] C. Févotte and A.T. Cemgil, "Nonnegative matrix factorisations as probabilistic inference in composite models," in *Proc. 17th European Signal Processing Conference (EUSIPCO)*, Aug. 2009, pp. 1913–1917.

[11] G.R. Arce, *Nonlinear signal processing: a statistical approach*, John Wiley & Sons, 2005.

[12] U. Şimşekli, A. Liutkus, and A.T. Cemgil, "Alpha-stable matrix factorization," *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2289–2293, Dec. 2015.

[13] A. Liutkus, D. FitzGerald, and R. Badeau, "Cauchy nonnegative matrix factorization," in *Proc. 2015 IEEE International Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct. 2015, pp. 1–5.

[14] P. Magron, R. Badeau, and A. Liutkus, "Lévy NMF for robust nonnegative source separation," in *Proc. 2017 IEEE International Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct. 2017, pp. 259–263.

[15] K. Yoshii, K. Itoyama, and M. Goto, "Student's t nonnegative matrix factorization and positive semidefinite tensor factorization for single-channel audio source separation," in *Proc. 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2016, pp. 51–55.

[16] S. Kotz, T.J. Kozubowski, and K. Podgórski, *The Laplace distribution and generalizations*, Birkhauser-Verlag, 2001.

[17] T. Eltoft, T. Kim, and T.W. Lee, "On the multivariate Laplace distribution," *IEEE Signal Processing Letters*, vol. 13, no. 5, pp. 300–303, May 2006.

[18] T. Kim, H.T. Attias, S.Y. Lee, and T.W. Lee, "Blind source separation exploiting higher-order frequency dependencies," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 15, no. 1, pp. 70–79, Jan. 2007.

[19] D. FitzGerald, M. Cranitch, and E. Coyle, "On the use of the beta divergence for musical source separation," in *Proc. IET Irish Signals and Systems Conference (ISSC)*, Jun. 2009, pp. 1–6.

[20] B. King, C. Févotte, and P. Smaragdis, "Optimal cost function and magnitude power for NMF-based speech separation and music interpolation," in *Proc. 2012 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Sep. 2012, pp. 1–6.

[21] D.R. Hunter and K. Lange, "A tutorial on MM algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, Feb. 2004.

[22] C. Févotte, "Majorization-minimization algorithm for smooth Itakura-Saito nonnegative matrix factorization," in *Proc. 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011, pp. 1980–1983.

[23] P. Smaragdis, B. Raj, and M. Shashanka, "Supervised and semi-supervised separation of sounds from single-channel mixtures," in *Proc. 2007 International Conference on Independent Component Analysis and Signal Separation (ICA)*, 2007, pp. 414–421.

[24] A. Liutkus and R. Badeau, "Generalized Wiener filtering with fractional power spectrograms," in *Proc. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2015, pp. 266–270.

[25] C. Kehling, J. Abeßer, C. Dittmar, and G. Schuller, "Automatic tablature transcription of electric guitar recordings by estimation of score- and instrument-related parameters," in *Proc. 17th International Conference on Digital Audio Effects (DAFx)*, Sep. 2014.

[26] E. Vincent, R. Gribonval, and C. Févotte, "Performance measurement in blind audio source separation," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 14, no. 4, pp. 1462–1469, Jul. 2006.

[27] C. Raffel, B. McFee, E.J. Humphrey, J. Salamon, O. Nieto, D. Liang, and D.P.W. Ellis, "Mir_eval: A transparent implementation of common MIR metrics," in *Proc. 15th International Society for Music Information Retrieval Conference (ISMIR)*, Oct. 2014, pp. 367–372.