時間領域低ランクスペクトログラム近似法による音源分離 *

☆鏡　英章 [1], 亀岡弘和 [2], △湯川正裕 [1]
([1] 慶應大・総合デザイン工, [2]NTT CS 研)

1　はじめに

混合音に含まれる各音源に対応する信号を分離抽出する問題を音源分離という。モノラル信号の音源分離を可能にするためには一般に何らかの仮定が必要である。混合音の各時刻における振幅（またはパワー）スペクトルが L 種類のスペクトルテンプレートの線形結合で表せるものとすると，混合音の振幅（あるいはパワー）スペクトログラムを二つの非負値行列の積で表現することができる。従って混合音の振幅スペクトログラムを二つの非負値行列の積で近似することにより，L 個のスペクトルテンプレートとともに各時刻でそれぞれのスペクトルテンプレートがどれくらい優勢であるかを推定することができる。この考え方に基づく音源分離手法を非負値行列因子分解 (Non-negative Matrix Factorization; NMF) 法と呼ぶ [1]。このアプローチは，振幅（またはパワー）スペクトルは加法的であるという必ずしも正確でない仮定に基づいている。この問題を解決するため，上述のアイディアを加法性が成り立つ複素スペクトログラムの表現に拡張した複素 NMF と呼ぶ枠組みが提案されている [2, 3]。複素 NMF のアイディアのポイントは，混合音の各時刻の複素スペクトルをスペクトルテンプレートに時変の位相スペクトルを乗じたものの和により表現する点にある。また，同様の動機から，混合音の複素スペクトログラムの生成モデルが提案されている [4, 5]。混合音の構成音の複素スペクトログラムの各要素が平均 0 の複素正規分布に従うと仮定できる場合，各構成音のパワースペクトログラムの最尤推定問題は板倉齋藤距離を規準として観測スペクトログラムを構成音のパワースペクトログラムの和でフィッティングする最適化問題と等価となる。従って，各構成音のパワースペクトログラムにランク 1 行列の構造を仮定することで，この問題は板倉斎藤擬距離規準の NMF に帰着する。この枠組を IS-NMF と呼ぶ。複素正規分布を複素 Cauchy 分布に置き換えた同様の生成モデルも提案されている [6]。

複素 NMF や IS-NMF では位相スペクトログラムの各要素が独立なパラメータ（もしくは潜在変数）として扱われるが，実際には互いに依存関係にある。これは，短時間 Fourier 変換 (Short-time Fourier Transform; STFT) やウェーブレット変換などにより得られる複素スペクトログラムは時間領域信号の冗長表現になっているからである。例えば STFT では信号の短時間フレームごとの Fourier 変換を連結して複素スペクトログラムを得る。フレーム間に重複区間がある場合は，それぞれの複素スペクトルは，重複区間における波形が無矛盾になるという条件を満たさなければならない [7]。複素 NMF や IS-NMF ではこのような制約が考慮されていなかったため，推定される複素スペクトログラムは時間領域信号に対応するとは限らず，時間領域信号に変換する際はできるだけ矛盾のない信号に近似する必要があった。

これまで我々は，複素 NMF のコンセプトを時間領域に拡張し，各構成音の時間領域信号をパラメータとした最適化問題として定式化した時間領域スペクトログラム近似法 (Time-domain Spectrogram Factorization; TSF) と呼ぶ枠組を提案している [9]。[9] では大きな行列の逆行列計算を含む最適化アルゴリズムが提案されているが，本稿では補助関数法と射影勾配法を用いてこれを回避する効率的なアルゴリズムを提案する。

2　非負値行列因子分解法

周波数と時刻に対応するインデックスを $k = 1, \ldots, K$, $m = 1, \ldots, M$ とし，観測信号の複素スペクトログラムを $Y_{k,m} \in \mathbb{C}$，$Y_{k,m}$ の絶対値またはその絶対値二乗を $Z_{k,m} = |Y_{k,m}|^{\alpha} \in \mathbb{R}_{+}$ $(\alpha = 1, 2)$ とする。NMF 法は観測振幅（またはパワー）スペクトログラム $\boldsymbol{Z} = (Z_{k,m})_{K\times M} \in \mathbb{R}_{+}^{K\times M}$ を非負値行列 $\boldsymbol{H} \in \mathbb{R}_{+}^{K\times L}$, $\boldsymbol{U} \in \mathbb{R}_{+}^{L\times M}$ の積 $\boldsymbol{Z}$ の各要素を

$$Z_{k,m} \simeq \sum_{l} H_{k,l} U_{l,m} \tag{1}$$

で近似することで，L 個のスペクトルテンプレート $H_{k,1}, \ldots, H_{k,L}$ と各時刻のアクティベーション $U_{1,m}, \ldots, U_{L,m}$ を同時推定する手法である [1]。式 (1) は各構成音の振幅スペクトログラムをランク 1 の行列 $(H_{k,l}U_{l,m})_{K\times M}$ でモデル化していることに相当する。式 (1) の分解表現が得られれば，例えば Wiener フィルタ

$$\hat{X}_{l,k,m} = Y_{k,m} \frac{(H_{k,l}U_{l,m})^{2/\alpha}}{(\sum_{l} H_{k,l}U_{l,m})^{2/\alpha}} \tag{2}$$

* Time-domain spectrogram factorization for sound source separation. by Hideaki Kagami (Keio University), Hirokazu Kameoka (NTT Communication Science Laboratories), Masahiro Yukawa (Keio University).

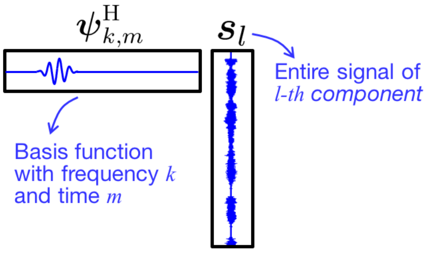

Fig. 1　信号の時間周波数表現

により各構成音の複素スペクトログラム $\hat{X}_{l,k,m}$ を得ることができる。

式 (1) の分解表現を得る問題は $\boldsymbol{Z}$ と $\boldsymbol{HU}$ の誤差の規準を用いた最適化問題として定式化することができる。誤差規準としては各行列要素の Euclid 距離（二乗誤差），I ダイバージェンス，板倉齋藤距離，β ダイバージェンスなどの和が用いられる。以上の目的関数を停留点へ収束させる $\boldsymbol{H}, \boldsymbol{U}$ は補助関数法に基づく反復アルゴリズムにより探索することができる [8, 10, 11]。

2.1　補助関数法

$\boldsymbol{\theta}$ に関して最小化したい目的関数を $F(\boldsymbol{\theta})$ とすると，

$$F(\boldsymbol{\theta}) = \min_{\bar{\boldsymbol{\theta}}} F^{+}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) \tag{3}$$

を満たす $F^{+}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}})$ を $F(\boldsymbol{\theta})$ の補助関数，$\bar{\boldsymbol{\theta}}$ を補助変数と呼ぶ。このとき，次の定理が成り立つ。

定理 1 $\bar{\boldsymbol{\theta}} \leftarrow \operatorname{argmin}_{\bar{\boldsymbol{\theta}}} F^{+}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}})$ と $\boldsymbol{\theta} \leftarrow \operatorname{argmin}_{\boldsymbol{\theta}} F^{+}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}})$ を交互に繰り返すことで目的関数 $F(\boldsymbol{\theta})$ は停留点に収束する。

3　複素 NMF

複素 NMF は，複素スペクトログラムが加法的である点に着目し，$Y_{k,m}$ をスペクトルテンプレート $H_{k,l}$，アクティベーション $U_{l,m}$，時変の位相スペクトル $c_{l,k,m}$ の積の和

$$Y_{k,m} \simeq \sum_{l} H_{k,l} U_{l,m} c_{l,k,m} \tag{4}$$

で近似する方法である [2, 3]。[2] では二乗誤差規準，[3] では KL ダイバージェンス規準の最適化アルゴリズムが補助関数法に基づきそれぞれ導出されている。

4　時間領域低ランクスペクトログラム近似

4.1　問題の定式化

観測信号を $\boldsymbol{y} = (y[1], \ldots, y[N])^{\mathsf{T}} \in \mathbb{R}^{N}$ とする。TSF は，各構成音 $\boldsymbol{s}_l$ の振幅スペクトログラムができるだけランク 1 構造に近くなるように $\boldsymbol{y}$ を $\boldsymbol{s}_l$ の和

$$\boldsymbol{y} = \sum_{l=1}^{L} \boldsymbol{s}_l, \tag{5}$$

に分解する方法である [9]。$\boldsymbol{\psi}_{k,m} \in \mathbb{C}^{N}$ を時間 m，周波数 k の時間周波数成分を得るための基底関数（図 1）とすると，$\boldsymbol{s}_l$ の振幅スペクトログラムは $|\boldsymbol{\psi}^{\mathsf{H}}_{k,m} \boldsymbol{s}_l|$ と表される。ただし，$(\cdot)^{\mathsf{H}}$ はベクトルの複素共役転置を表す。式 (5) を満たしながら $|\boldsymbol{\psi}^{\mathsf{H}}_{k,m} \boldsymbol{s}_l| \simeq H_{k,l} U_{l,m}$ となるように $\boldsymbol{S} = \{\boldsymbol{s}_l\}_l$, $\boldsymbol{H} = \{H_{k,l}\}_{k,l}$, $\boldsymbol{U} = \{U_{l,m}\}_{l,m}$ を求める問題は，

$$\text{minimize} \quad \mathcal{J}(\boldsymbol{\theta}) \tag{6}$$

$$\text{subject to} \quad \boldsymbol{y} = \sum_{l} \boldsymbol{s}_l, \tag{7}$$

$$\mathcal{J}(\boldsymbol{\theta}) := \sum_{l,k,m} \frac{(|\boldsymbol{\psi}^{\mathsf{H}}_{k,m} \boldsymbol{s}_l| - H_{k,l} U_{l,m})^2}{\beta_{l,k,m}} + R(\boldsymbol{U}) \tag{8}$$

のような最適化問題として定式化できる。ただし，$\boldsymbol{\theta} = \{\boldsymbol{H}, \boldsymbol{U}, \boldsymbol{S}, \boldsymbol{\beta}\}$, $\boldsymbol{\beta} = \{\beta_{l,k,m}\}$ であり，$\beta_{l,k,m}$ は $\sum_l \beta_{l,k,m} = 1$ を満たす正値変数とする。$\beta_{l,k,m}$ の重要性については後述する。$\mathcal{J}(\boldsymbol{\theta})$ の一つ目の項は振幅スペクトログラム $\boldsymbol{S}$ が完全にランク 1 構造を有したとき 0 になる。複素 NMF と同様，このモデルでは要素ごとに値を相殺することができるため $\boldsymbol{U}$ をスパースに誘導するための正則化項が必要になる。$R(\boldsymbol{U})$ はこのために導入した項であり，ここでは ℓ_p ノルム

$$R(\boldsymbol{U}) = 2\lambda \sum_{l,m} |U_{l,m}|^{p} \tag{9}$$

を用いる。ただし，$\lambda > 0$ は正則化項の寄与の大きさを決める正則化パラメータである。$\boldsymbol{U}$ が有界で $0 < p < 2$ のとき，$R(\boldsymbol{U})$ は $\boldsymbol{U}$ をスパースに誘導する働きをもつ項となる。$\boldsymbol{U}$ を有界にするため，

$$\sum_{k} H_{k,l}^{2} = 1 \tag{10}$$

とする。

4.2　$\mathcal{J}$ の補助関数

目的関数 $\mathcal{J}(\boldsymbol{\theta})$ は $|\boldsymbol{\psi}^{\mathsf{H}}_{k,m} \boldsymbol{s}_l|$ と $|U_{l,m}|^{p}$ $(0 < p < 1)$ に関して微分不可能な点を含むため $\mathcal{J}(\boldsymbol{\theta})$ を最小化する $\boldsymbol{s}_l$, $U_{l,m}$ は解析的に求まらないが，補助関数法により停留点への収束が保証された最適化アルゴリズムを導くことができる。以下に示す不等式 (14), (15) を用いることで $\mathcal{J}(\boldsymbol{\theta})$ の補助関数

$$\mathcal{J}^{+}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) = \sum_{l,k,m} \frac{1}{\beta_{l,k,m}} \left| \boldsymbol{\psi}^{\mathsf{H}}_{k,m} \boldsymbol{s}_l - H_{k,l} U_{l,m} c_{l,k,m} \right|^{2} + \lambda \sum_{l,m} \left\{ p |V_{l,m}|^{p-2} U_{l,m}^{2} + (2-p) |V_{l,m}|^{p} \right\}, \tag{11}$$

を得ることができる。ただし，$\bar{\boldsymbol{\theta}} = \{\boldsymbol{c}, \boldsymbol{V}\}$，$\boldsymbol{c} = \{c_{l,k,m}\}$，$\boldsymbol{V} = \{V_{l,m}\}$ である。補助変数 $\boldsymbol{C}$，$\boldsymbol{V}$ を

$$c_{l,k,m} = \boldsymbol{\psi}_{k,m}^{\mathsf{H}} \boldsymbol{s}_l / |\boldsymbol{\psi}_{k,m}^{\mathsf{H}} \boldsymbol{s}_l|, \tag{12}$$

$$V_{l,m} = U_{l,m}. \tag{13}$$

と置いたとき $\mathcal{J}^+(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) = \mathcal{J}(\boldsymbol{\theta})$ となる。

補題 1 任意の複素数 z，$|c| = 1$ を満たす複素数 c について次式が成り立つ。

$$-|z| \le -\mathrm{Re}[c^* z]. \tag{14}$$

等号は $c = z/|z|$ のとき成立する。

補題 2 $0 < p < 2$ のとき，任意の実数（または複素数）x について次式が成り立つ。

$$2|x|^2 \le p|v|^{p-2}|x|^2 + 2 - p|v|^p. \tag{15}$$

等号は $v = x$ のとき成立する。

5 射影勾配法を用いた TSF アルゴリズム

上記の補助関数を最小化する $\boldsymbol{s}_l$ は，Lagrange の未定乗数法を用いて

$$\boldsymbol{s}_l = \boldsymbol{\Psi}_l^{-1}(\boldsymbol{d}_l - \boldsymbol{\mu}), \tag{16}$$

となることが示せる。ただし，

$$\boldsymbol{\Psi}_l = \sum_{k,m} \frac{2\mathrm{Re}[\boldsymbol{\psi}_{k,m} \boldsymbol{\psi}_{k,m}^{\mathsf{H}}]}{\beta_{l,k,m}}, \tag{17}$$

$$\boldsymbol{d}_l = \sum_{k,m} \frac{2\mathrm{Re}[c_{l,k,m} \boldsymbol{\psi}_{k,m}] H_{k,l} U_{l,m}}{\beta_{l,k,m}}, \tag{18}$$

$$\boldsymbol{\mu} = \Big(\sum_l \boldsymbol{\Psi}_l\Big)^{-1} \Big(\sum_l \boldsymbol{\Psi}_l^{-1} \boldsymbol{d}_l - \boldsymbol{y}\Big), \tag{19}$$

である。式 (16) を $\boldsymbol{s}_l$ の更新式としたアルゴリズムは停留点への収束性が保証されるが，この更新式は大きな行列の逆行列計算を含む。[9] では $\beta_{1,k,m} = \cdots = \beta_{L,k,m}$ (i.e., $\forall l, \beta_{l,k,m} = 1/L$) という特定の条件で $\boldsymbol{\Psi}_l$ が単位行列になることを利用し，逆行列計算を回避する方法をとったが，β を固定することが音源分離性能に限界を与えていたことが検討の結果明らかになった。そこで，本章では射影勾配法を用いて当該逆行列計算を回避する効率的なアルゴリズムを提案する。

提案するアルゴリズムの $\boldsymbol{H}$ と $\boldsymbol{U}$ の更新式は [9] と同様のため，ここでは $\boldsymbol{s}_l$ の更新についてのみ考える。そのため，$\mathcal{J}^+$ を $\boldsymbol{S}$ の関数とし，$\mathcal{J}^+(\boldsymbol{S})$ と表記する。$\mathcal{J}^+(\boldsymbol{S})$ の $\boldsymbol{s}_l$ に関する偏微分は

$$\nabla_{\boldsymbol{s}_l} \mathcal{J}^+(\boldsymbol{S}) = \boldsymbol{\Psi}_l \boldsymbol{s}_l - \boldsymbol{d}_l \tag{20}$$

で与えられる。式 (20) の各項は効率的に計算可能である。例えば $\boldsymbol{\psi}_{k,m}$ が STFT を算出する基底関数（フレーム m の窓掛けされた複素正弦波）の場合，$\boldsymbol{\psi}_{k,m}^{\mathsf{H}} \boldsymbol{s}_l$ は $\boldsymbol{s}_l$ の STFT による複素スペクトログラムの (k, m) 成分である。また，式 (17) の $\sum_{k,m} \mathrm{Re}[\boldsymbol{\psi}_{k,m} \cdot]$ は逆 STFT に対応しており，式 (20) の初項は $2\boldsymbol{\psi}_{k,m}^{\mathsf{H}} \boldsymbol{s}_l / \beta_{l,k,m}$ の逆 STFT である。同様に式 (18) の $\boldsymbol{d}_l$ は $2H_{k,l} U_{l,m} c_{l,k,m} / \beta_{l,k,m}$ の逆 STFT である。よって $\boldsymbol{s}_l$ の勾配計算には高速 Fourier 変換を活用できるため，勾配法による更新則

$$\boldsymbol{s}_l \leftarrow \boldsymbol{s}_l - \gamma \nabla_{\boldsymbol{s}_l} \mathcal{J}^+(\boldsymbol{s}_l), \tag{21}$$

は効率的に計算することができる。ただし，$\gamma \in (0, 2/\kappa)$ はステップサイズである。ここで，$\kappa > 0$ は勾配 $\nabla_{\boldsymbol{s}_l} \mathcal{J}^+$ の Lipschitz 定数であり，行列 $\boldsymbol{\Psi}_l$ の最大固有値から得られるが，本稿では次の式で得られる γ をステップサイズとして用いる。

$$\gamma := \frac{2}{\rho} \in (0, 2/\kappa), \tag{22}$$

$$\rho := \sum_l \mathrm{tr}\,(\boldsymbol{\Psi}_l) = \sum_{l,n} \sum_{k,m} \frac{2\{w(n - \alpha_m)\}^2}{\beta_{l,k,m}}, \tag{23}$$

ただし，$w(n)$ は STFT の窓関数，α_m はフレームシフトの割合である。

式 (5) の制約を満たすように $\boldsymbol{s}_l$ を更新するため，本稿では射影勾配法を用いる。まず，

$$\tilde{\boldsymbol{s}} := \begin{bmatrix} \boldsymbol{s}_1 \\ \vdots \\ \boldsymbol{s}_L \end{bmatrix} \in \mathbb{R}^{LN}, \tag{24}$$

$$\boldsymbol{G} := [\boldsymbol{I}_N \ \cdots \ \boldsymbol{I}_N] \in \mathbb{R}^{N \times LN}, \tag{25}$$

を定義する。ここで，$\boldsymbol{I}_N$ は $N \times N$ の単位行列である。線形制約 $\boldsymbol{y} = \sum_l \boldsymbol{s}_l$ は $\boldsymbol{y} = \boldsymbol{G}\tilde{\boldsymbol{s}}$ と書ける。また

$$\mathcal{V} := \left\{\tilde{\boldsymbol{s}} \in \mathbb{R}^{LN} \mid \boldsymbol{G}\tilde{\boldsymbol{s}} = \boldsymbol{y}\right\}, \tag{26}$$

を定義する。このときアフィン空間 $\mathcal{V}$ への射影は

$$P_{\mathcal{V}}(\tilde{\boldsymbol{s}}) = \tilde{\boldsymbol{s}} - \boldsymbol{G}^{\mathsf{T}}(\boldsymbol{G}\boldsymbol{G}^{\mathsf{T}})^{-1}(\boldsymbol{G}\tilde{\boldsymbol{s}} - \boldsymbol{y}), \tag{27}$$

となる。さらに $\boldsymbol{G}\boldsymbol{G}^{\mathsf{T}} = L\boldsymbol{I}_N$，

$$\boldsymbol{G}^{\mathsf{T}}\boldsymbol{G} = \begin{bmatrix} \boldsymbol{I}_N \ \boldsymbol{I}_N \ \ldots \ \boldsymbol{I}_N \\ \boldsymbol{I}_N \ \boldsymbol{I}_N \ \ldots \ \boldsymbol{I}_N \\ \vdots \\ \boldsymbol{I}_N \ \boldsymbol{I}_N \ \ldots \ \boldsymbol{I}_N \end{bmatrix} \in \mathbb{R}^{LN \times LN}, \tag{28}$$

であることを用いると

$$[P_{\mathcal{V}}(\tilde{\boldsymbol{s}})]_l = \boldsymbol{s}_l - \frac{1}{L}\left(\sum_{l'=1}^{L} \boldsymbol{s}_{l'} - \boldsymbol{y}\right), \tag{29}$$

が得られる。実際は ρ の上界が κ であるため式 (22)(23) で得られる γ は $\gamma \in (0, 2/\kappa)$ を満たす。従って，式 (21), (29) を交互に繰り返すことにより，線形制約の下で $\mathcal{J}^+$ を最小化する値に収束することが保証される。

$\boldsymbol{H}$ と $\boldsymbol{U}$ の更新式は [9] と同様に得られる。まず，$\boldsymbol{U}$ に関して $\mathcal{J}^+$ の偏微分を 0 と置くことで更新式

$$U_{l,m} = \frac{\sum_k H_{k,l} |\boldsymbol{\psi}_{k,m}^{\mathsf{H}} \boldsymbol{s}_l| / \beta_{l,k,m}}{\sum_k H_{k,l}^2 / \beta_{l,k,m} + \lambda p |V_{l,m}|^{p-2}}, \tag{30}$$

が得られる。また $\boldsymbol{H}$，$\boldsymbol{\beta}$ についてはLagrangeの未定乗数法を用いることで

$$H_{k,l} = \frac{\sum_m U_{l,m}|\boldsymbol{\psi}^{\mathsf{H}}_{k,m}\boldsymbol{s}_l|/\beta_{l,k,m}}{\sqrt{\sum_k(\sum_m U_{l,m}|\boldsymbol{\psi}^{\mathsf{H}}_{k,m}\boldsymbol{s}_l|/\beta_{l,k,m})^2}}, \tag{31}$$

$$\beta_{l,k,m} = \frac{\big||\boldsymbol{\psi}^{\mathsf{H}}_{k,m}\boldsymbol{s}_l| - H_{k,l}U_{l,m}\big|}{\sum_l \big||\boldsymbol{\psi}^{\mathsf{H}}_{k,m}\boldsymbol{s}_l| - H_{k,l}U_{l,m}\big|}, \tag{32}$$

が得られる。全体の更新式は次のようになる。

1. Initialize $\boldsymbol{H}$, $\boldsymbol{U}$ and $\boldsymbol{S}$.
2. Update $\boldsymbol{c}$ and $\boldsymbol{V}$ using (12), (13).
3. Update $\boldsymbol{S}$ and γ using (20), (22)–(23), (29).
   for $t = 1$: iteration
   $$\boldsymbol{s}_l^{(t+1)} = P_{\mathcal{V}}\left(\boldsymbol{s}_l^{(t)} - \gamma\nabla_{\boldsymbol{s}_l}\mathcal{J}^+(\boldsymbol{s}_l^{(t)})\right)$$
   end
4. Update $\boldsymbol{U}$, $\boldsymbol{H}$ and $\boldsymbol{\beta}$ using (30)–(32).

6　複素NMFとの関係

複素NMFは $H_{k,l}$，　$U_{l,m}$，　$c_{l,k,m}$ に関して

$$\mathcal{I}(\boldsymbol{\theta}) = \sum_{k,m}\Big|Y_{k,m} - \sum_l H_{k,l}U_{l,m}c_{l,k,m}\Big|^2 + 2\lambda\sum_{l,m}|U_{l,m}|^p, \tag{33}$$

を最小化する最適化問題として定式化され，[2] では

$$\mathcal{I}^+(\boldsymbol{\theta},\bar{\boldsymbol{\theta}}) = \sum_{l,k,m}\frac{1}{\beta_{l,k,m}}\big|X_{l,k,m} - H_{k,l}U_{l,m}e^{j\phi_{l,k,m}}\big|^2 + \lambda\sum_{l,m}\left[p|V_{l,m}|^{p-2}U_{l,m}^2 + (2-p)|V_{l,m}|^p\right], \tag{34}$$

が $\mathcal{I}$ の補助関数として用いられている。ただし，$\bar{\boldsymbol{\theta}} = \{\boldsymbol{\beta}, \boldsymbol{X}\}$ が補助変数であり，$\boldsymbol{\beta} = \{\beta_{l,k,m}\}$，　$\boldsymbol{X} = \{X_{l,m}\}$ である。また，$\beta_{l,k,m}$ は $\sum_l \beta_{l,k,m} = 1$ を満たす正の数であり，$X_{l,k,m}$ は $Y_{k,m} = \sum_l X_{l,k,m}$ を満たす。$X_{l,k,m}$ は l 番目の構成音の複素スペクトログラムを表す。

(11) と (34) を比べると，複素NMFの補助関数における $X_{l,k,m}$ と $e^{j\phi_{l,k,m}}$ がTSFの補助関数における $\boldsymbol{s}_l$ と $c_{l,k,m}$ に対応していることがわかる。複素NMFの一つの問題は $X_{l,k,m}$ が複素スペクトログラムが時間領域信号に対応する条件が考慮されいない点にある。一方，提案手法では l 番目の複素スペクトログラム $\boldsymbol{\psi}^{\mathsf{H}}_{k,m}\boldsymbol{s}_l$ を手がかりに，適切な解空間の中で時間領域信号 $\boldsymbol{s}_l$ を推定することができる。

7　実験結果

提案手法の有効性を確認するため，SiSEC 2013 Database の”professionally produced music recordings”（5 音源からなるポピュラー音楽の音響信号）を用いて 3 分割交差検証による教師あり音源分離実験を行い，既存の TSF($\beta_{1,k,m} = \cdots = \beta_{L,k,m}$)，複素 NMF(CNMF-EUC)，I ダイバージェンス規準 NMF(NMF-KL)，二乗誤差規準 NMF(NMF-EUC)，板倉斎藤擬距離規準 NMF(NMF-IS) による各音源の Signal to Noise Ratio (SNR) 改善値を比較した。2 にその結果を示す。提案手法が他手法に比べて高い改善値を得られていることが確認できる。

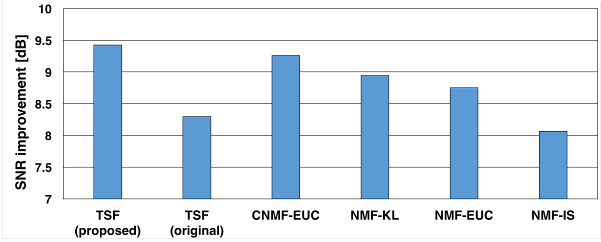

Fig. 2　SNR improvements.

8　おわりに

本稿では、時間領域で音源分離を行う効率的なTSFアルゴリズムを提案した。TSFの利点は、(i) 時間領域信号の加法性と (ii) 異なる時間周波数点における位相/振幅成分の相互依存性を利用することである。提案された反復アルゴリズムは、射影勾配法を利用した補助関数最小化に基づいて得られた。シミュレーション結果では、提案されたTSFアルゴリズムが既存法に比べ高い性能を比べて高い音源分離性能を得られることを確認した。

謝辞　本研究はJSPS科研費 15K06081, 15H02757, 26730100 の助成を受けて行われた。

参考文献

[1] Smaragdis et al., Proc. ICA, 414–421, 2007.

[2] Kameoka et al., Proc ICASSP, 3437–3440, 2009.

[3] 鏡他, 音講論 (秋), 1-P-5, 433–436, 2016.

[4] Parry & Essa, Proc. ICA, 536–543, 2007.

[5] Févotte et al., Neural Computation, 21(3), 793–830, 2009.

[6] Liutkus et al., Proc. WASPAA, 2015.

[7] Le Roux et al., Proc. DAFx, 397–403, 2010.

[8] Lee & Seung, Adv. NIPS, 556–562, 2000.

[9] Kameoka, Proc. ICASSP, 86–90, 2015.

[10] 亀岡他, 情処研報, 2006-MUS-66-13, 77–84, 2006

[11] M. Nakano et al., Proc. MLSP, 283–288, 2010.