スペクトログラム無矛盾性を用いた独立低ランク行列分析 *

☆豊島直, 北村大地 (香川高専), 矢田部浩平 (早稲田大)

1 はじめに

音源分離とは，複数の音源が混合した観測信号から，混合前の音源を推定する技術である．特に，音源位置やマイクロホン位置等が未知の条件で音源分離を達成する技術はブラインド音源分離（blind source separation: BSS）と呼ばれる．観測信号のチャネル数（マイクロホン数）と混合されている音源数が等しい条件下では，独立成分分析（independent component analysis: ICA）[1] に基づく BSS として，周波数領域 ICA（frequency-domain ICA: FDICA）[2]，独立ベクトル分析（independent vector analysis: IVA）[3, 4]，及び独立低ランク行列分析（independent low-rank matrix analysis: ILRMA）[5, 6] 等が提案されてきた．これらはいずれも，FDICA におけるパーミュテーション問題 [7]（周波数毎に得られる分離信号の順序を整列する問題）を解決するための FDICA の拡張である．IVA や ILRMA は，その分離性能の高さや初期値に対する頑健性等の利点から，数多くの改良手法や一般化手法へと発展している（例えば [8] 等）．

一方，多くの音響信号処理で用いられる短時間フーリエ変換（short-time Fourier transform: STFT）では，窓関数とそのオーバーラップにより近傍時間周波数グリッドに共起性が生じる（文献 [9] の Fig. 1 参照）．これは時間周波数表現された信号が根本的に有している性質であり，スペクトログラム無矛盾性と呼ばれる [10, 11]．近年，このスペクトログラム無矛盾性を FDICA や IVA に導入した BSS が提案され，分離行列最適化時のパーミュテーション問題の解決をアシストする新しい基準となることが示された [9]．その原理は，分離信号のパーミュテーションが近傍周波数で異なる場合に，スペクトログラム無矛盾性（隣接周波数の共起関係等）が大きく損なわれるという性質に基づいている（文献 [9] の Fig. 2 参照）．

本稿では，文献 [9] で示された結果に基づき，より高精度な BSS 達成を目的として，スペクトログラム無矛盾性を ILRMA に導入したアルゴリズムを提案する．また，音楽信号と音声信号の BSS 実験を通して，ILRMA においてもスペクトログラム無矛盾性が分離性能の向上に貢献することを示す．

2 BSS とスペクトログラム無矛盾性

2.1 STFT 及び BSS の定式化

離散時間信号の l 番目のサンプルを $x[l]$ のように表記し，N 個の音源信号が M 個のマイクロホンで観測される状況を考える．多チャネルの音源信号，観測信号，及び分離信号をそれぞれ次式で表す．

$$\boldsymbol{s}[l] = \left[\, s_1[l], \cdots, s_n[l], \cdots s_N[l] \,\right]^{\mathrm{T}} \quad \in \mathbb{R}^N \tag{1}$$

$$\boldsymbol{x}[l] = \left[\, x_1[l], \cdots, x_m[l], \cdots x_M[l] \,\right]^{\mathrm{T}} \quad \in \mathbb{R}^M \tag{2}$$

$$\boldsymbol{y}[l] = \left[\, y_1[l], \cdots, y_n[l], \cdots y_N[l] \,\right]^{\mathrm{T}} \quad \in \mathbb{R}^N \tag{3}$$

ここで，$n = 1, 2, \cdots, N$，$m = 1, 2, \cdots, M$，及び $l = 1, 2, \cdots, L$ はそれぞれ音源，マイクロホン（チャネル），及び離散時間のインデクスであり，$\cdot^{\mathrm{T}}$ は転置を表す．BSS では，音源信号 $\boldsymbol{s}$ に近い分離信号 $\boldsymbol{y}$ を，観測信号 $\boldsymbol{x}$ から推定することが目的となる．

FDICA，IVA，及び ILRMA 等の周波数領域 BSS では，信号を時間周波数領域で取り扱う．STFT における窓長とシフト長をそれぞれ Q 及び τ とおくと，信号 $z[l]$ の j 番目の短時間セグメントは次式となる．

$$\begin{aligned}\boldsymbol{z}^{[j]} &= \left[\, z[(j-1)\tau+1], z[(j-1)\tau+2], \right. \\ &\qquad\qquad \left. \cdots, z[(j-1)\tau+Q] \,\right]^{\mathrm{T}} \\ &= \left[\, z^{[j]}[1], \cdots, z^{[j]}[q], \cdots, z^{[j]}[Q] \,\right]^{\mathrm{T}} \in \mathbb{R}^Q\end{aligned}$$

ここで，$j = 1, 2, \cdots, J$ 及び $q = 1, 2, \cdots, Q$ はそれぞれ短時間セグメント及びセグメント内サンプルのインデクスを表す．セグメント数は（必要に応じて信号に零詰めを行ったうえで）$J = L/\tau$ となる．信号 $\boldsymbol{z} = [\, z[1], \cdots, z[l], \cdots, z[L] \,]^{\mathrm{T}} \in \mathbb{R}^L$ の STFT を

$$\boldsymbol{Z} = \mathrm{STFT}_{\boldsymbol{\omega}}(\boldsymbol{z}) \quad \in \mathbb{C}^{I \times J} \tag{4}$$

と表し，$\boldsymbol{Z}$ の (i, j) 要素は次式で与えられる．

$$z_{ij} = \sum_{q=1}^{Q} \omega[q]\, z^{[j]}[q]\, \mathrm{e}^{-\imath 2\pi (q-1)(i-1)/F} \tag{5}$$

ここで，$i = 1, 2, \cdots, I$ は周波数ビンのインデクス，F は $\lfloor F/2 \rfloor + 1 = I$ を満たす整数，$\lfloor \cdot \rfloor$ は床関数，$\imath$ は虚数単位，及び $\boldsymbol{\omega}$ は解析時の窓関数を表す．合成時の窓関数を $\tilde{\boldsymbol{\omega}}$ とおくとき，逆 STFT を $\mathrm{ISTFT}_{\tilde{\boldsymbol{\omega}}}(\cdot)$ と表記する．本稿では，$\boldsymbol{\omega}$ と $\tilde{\boldsymbol{\omega}}$ のペアが次式の完全再構成条件を満たすことを仮定する．

$$\boldsymbol{z} = \mathrm{ISTFT}_{\tilde{\boldsymbol{\omega}}}(\mathrm{STFT}_{\boldsymbol{\omega}}(\boldsymbol{z})) \qquad \forall \boldsymbol{z} \in \mathbb{R}^L \tag{6}$$

各チャネルに STFT を適用して得られる音源信号，観測信号，及び分離信号のスペクトログラムの (i, j) 番目の要素をそれぞれ次式で表す．

$$\boldsymbol{s}_{ij} = [\, s_{ij1}, \cdots, s_{ijn}, \cdots s_{ijN} \,]^{\mathrm{T}} \quad \in \mathbb{C}^N \tag{7}$$

$$\boldsymbol{x}_{ij} = [\, x_{ij1}, \cdots, x_{ijm}, \cdots x_{ijM} \,]^{\mathrm{T}} \quad \in \mathbb{C}^M \tag{8}$$

$$\boldsymbol{y}_{ij} = [\, y_{ij1}, \cdots, y_{ijn}, \cdots y_{ijN} \,]^{\mathrm{T}} \quad \in \mathbb{C}^N \tag{9}$$

また，式 (7)–(9) の時間周波数行列をそれぞれ $\boldsymbol{S}_n \in \mathbb{C}^{I \times J}$，$\boldsymbol{X}_m \in \mathbb{C}^{I \times J}$，及び $\boldsymbol{Y}_n \in \mathbb{C}^{I \times J}$ と定義する．周波数領域 BSS では，次式の瞬時混合を仮定する．

$$\boldsymbol{x}_{ij} = \boldsymbol{A}_i \boldsymbol{s}_{ij} \tag{10}$$

ここで，$\boldsymbol{A}_i \in \mathbb{C}^{M \times N}$ は周波数毎の混合行列である．式 (10) は，混合系の残響時間が窓長よりも十分短い場合に近似的に成立する．

以後，決定的な系（$M = N$）を考える．この場合，BSS は $\boldsymbol{A}_i$ の逆行列を推定する問題となる．この逆行列を $\boldsymbol{W}_i \approx \boldsymbol{A}_i^{-1}$ とすると，分離信号は次式となる．

$$\boldsymbol{y}_{ij} = \boldsymbol{W}_i \boldsymbol{x}_{ij} \tag{11}$$

*Independent low-rank matrix analysis based on spectrogram consistency. By Nao TOSHIMA, Daichi KITAMURA (NIT Kagawa), and Kohei YATABE (Waseda Univ.).

ここで，$\boldsymbol{W}_i = [\boldsymbol{w}_{i1}, \cdots, \boldsymbol{w}_{in}, \cdots, \boldsymbol{w}_{iN}]^{\mathrm{H}} \in \mathbb{C}^{N\times M}$ は分離行列と呼ばれ，$\cdot^{\mathrm{H}}$ はエルミート転置を表す．

ICA に基づく BSS では，分離音源のスケールや順序が不定であるため，次式の任意性が存在する．

$$\hat{\boldsymbol{y}}_{ij} = \hat{\boldsymbol{W}}_i \boldsymbol{x}_{ij} \quad (\hat{\boldsymbol{W}}_i = \boldsymbol{D}_i \boldsymbol{P}_i \boldsymbol{W}_i) \tag{12}$$

ここで，$\boldsymbol{D}_i \in \mathbb{C}^{N\times N}$ 及び $\boldsymbol{P}_i \in \{0,1\}^{N\times N}$ はそれぞれ任意の対角行列及びパーミュテーション行列である．周波数毎のスケールの任意性についてはプロジェクションバック法 [13] によって解析的に復元可能であるが，周波数毎の分離信号の順序の整列はパーミュテーション問題と呼ばれ，大きな課題である．

2.2 スペクトログラム無矛盾性に基づくパーミュテーション解決 [9]

文献 [9] では，スペクトログラム無矛盾性と呼ばれる STFT の性質を用いてパーミュテーション問題が解決できる可能性が示されている．スペクトログラムは本来，時間周波数表現における不確定性原理より，近傍の時間周波数成分が共起関係を持つ．スペクトログラム $\boldsymbol{Z}$ において，この共起関係が崩れている状態を矛盾と呼び，逆 STFT はその矛盾した成分を無矛盾な状態に復元する．即ち，$\boldsymbol{Z}$ の無矛盾性は

$$\mathcal{E}(\boldsymbol{Z}) = \boldsymbol{Z} - \mathrm{STFT}_{\boldsymbol{\omega}}(\mathrm{ISTFT}_{\tilde{\boldsymbol{\omega}}}(\boldsymbol{Z})) \tag{13}$$

のノルム $\|\mathcal{E}(\boldsymbol{Z})\|$ によって特徴付けられ，それが 0 となるスペクトログラムを無矛盾と呼ぶ．大雑把に言えば，逆 STFT はスペクトログラムの時間及び周波数方向への滲みがない成分に対して，スムージングをかけるような処理となる（文献 [9] の Fig. 1 参照）．

文献 [9] の Fig. 2 のように，パーミュテーション問題が生じている分離信号のスペクトログラムは，隣接周波数の不連続性に起因して無矛盾性が大きく損なわれている．従って，FDICA や IVA における $\boldsymbol{W}_i$ の反復最適化の過程で $\mathrm{STFT}_{\boldsymbol{\omega}}(\mathrm{ISTFT}_{\tilde{\boldsymbol{\omega}}}(\boldsymbol{Y}_n))$ を作用させることで，周波数毎の音源成分が正しく整列された（パーミュテーション問題が解決された）状態に少しずつ誘導できることが示されている．本稿では，より高い分離性能を発揮できる ILRMA においても，上記のスペクトログラム無矛盾性を反復毎に担保するアルゴリズムが性能向上に貢献することを実証する．

3 提案手法

3.1 スペクトログラム無矛盾性に基づく ILRMA

ILRMA は次式の目的関数を最小化する [6].

$$\mathcal{L} = -2J \sum_i |\det \boldsymbol{W}_i| + \sum_{i,j,n} \left(\frac{|\boldsymbol{w}_{in}^{\mathrm{H}} \boldsymbol{x}_{ij}|^2}{\sum_k t_{ikn} v_{kjn}} + \log \sum_k t_{ikn} v_{kjn} \right) \tag{14}$$

ここで，t_{ikn} 及び v_{kjn} は非負行列 $\boldsymbol{T}_n \in \mathbb{R}_{\geq 0}^{I\times K}$ 及び $\boldsymbol{V}_n \in \mathbb{R}_{\geq 0}^{K\times J}$ の要素であり，$k = 1, 2, \cdots, K$ は $\boldsymbol{T}_n$ の列ベクトルのインデクスである．式 (14) の最小化は，ICA に基づく分離信号間の独立性の最大化と非負値行列因子分解（nonnegative matrix factorization: NMF）[12] に基づく分離信号の低ランク近似 $|\boldsymbol{Y}_n|^2 \approx \boldsymbol{T}_n \boldsymbol{V}_n$ を同時に行うことに対応する．但し，行列に対する $|\cdot|^2$ は要素毎の絶対値二乗を表す．これは即ち，混合前の音源信号 $\boldsymbol{S}_n$ が低ランク時間周波数構造を有することを仮定し，分離信号 $\boldsymbol{Y}_n$ を NMF でモデル化しながら分離行列 $\boldsymbol{W}_i$ を最適化することで，パーミュテーション問題が生じない $\boldsymbol{W}_i$ を推定している．

提案手法では，スペクトログラム無矛盾性を担保する処理を ILRMA の最適化アルゴリズムに組み込む．従来の ILRMA の分離行列の反復最適化計算 [4]

$$\boldsymbol{U}_{in} \leftarrow \frac{1}{J} \sum_j \frac{1}{\sum_k t_{ikn} v_{kjn}} \boldsymbol{x}_{ij} \boldsymbol{x}_{ij}^{\mathrm{H}} \tag{15}$$

$$\boldsymbol{w}_{in} \leftarrow (\boldsymbol{W}_i \boldsymbol{U}_{in})^{-1} \boldsymbol{e}_n \tag{16}$$

$$\boldsymbol{w}_{in} \leftarrow \boldsymbol{w}_{in} \left(\boldsymbol{w}_{in}^{\mathrm{H}} \boldsymbol{U}_{in} \boldsymbol{w}_{in} \right)^{-\frac{1}{2}} \tag{17}$$

$$y_{ijn} \leftarrow \boldsymbol{w}_{in}^{\mathrm{H}} \boldsymbol{x}_{ij} \tag{18}$$

及び NMF 音源モデルの反復最適化計算

$$t_{ikn} \leftarrow t_{ikn} \sqrt{\frac{\sum_j |y_{ijn}|^2 \left(\sum_{k'} t_{ik'n} v_{k'jn}\right)^{-2} v_{kjn}}{\sum_j \left(\sum_{k'} t_{ik'n} v_{k'jn}\right)^{-1} v_{kjn}}} \tag{19}$$

$$v_{kjn} \leftarrow v_{kjn} \sqrt{\frac{\sum_i |y_{ijn}|^2 \left(\sum_{k'} t_{ik'n} v_{k'jn}\right)^{-2} t_{ikn}}{\sum_i \left(\sum_{k'} t_{ik'n} v_{k'jn}\right)^{-1} t_{ikn}}} \tag{20}$$

において，

$$\boldsymbol{Y}_n \leftarrow \mathrm{STFT}_{\boldsymbol{\omega}}(\mathrm{ISTFT}_{\tilde{\boldsymbol{\omega}}}(\boldsymbol{Y}_n)) \tag{21}$$

なる演算を挿入することで，毎回の反復においてスペクトログラム無矛盾性を担保する．ここで，$\boldsymbol{e}_n \in \{0,1\}^N$ は n 番目の要素のみが 1 の単位ベクトルである．式 (21) は，分離信号のスペクトログラム $\boldsymbol{Y}_n$ を無矛盾スペクトログラムの集合へと射影していることに対応する．そのため，もし $\boldsymbol{Y}_n$ が無矛盾であれば式 (21) は何もしておらず，$\boldsymbol{Y}_n$ に矛盾があれば時間及び周波数の両方向にスムージングがかかる．

3.2 反復毎のプロジェクションバック法の適用

ILRMA の推定は FDICA や IVA と同様に，式 (12) に示す任意性がある．とくに，任意の対角行列 $\boldsymbol{D}_i$ に起因する周波数毎のスケール不定性は，それ自身がスペクトログラム無矛盾性を崩してしまう．提案手法では，$\boldsymbol{D}_i$ に起因する矛盾成分の影響を最小限に抑えるため，毎回の反復計算において式 (21) を行う直前に，次式のプロジェクションバック法 [13] を適用する．

$$\tilde{\boldsymbol{y}}_{ijn} = \boldsymbol{W}_i^{-1} (\boldsymbol{e}_n \circ \boldsymbol{y}_{ij}) = y_{ijn} \boldsymbol{\lambda}_{in}, \tag{22}$$

ここで，$\tilde{\boldsymbol{y}}_{ijn} = [\tilde{y}_{ijn1}, \cdots, \tilde{y}_{ijnm}, \cdots, \tilde{y}_{ijnM}]^{\mathrm{T}} \in \mathbb{C}^M$ はスケール補正後の分離信号の (i,j) 番目の成分，$\boldsymbol{\lambda}_{in} = [\lambda_{in1}, \cdots, \lambda_{inm}, \cdots, \lambda_{inM}]^{\mathrm{T}} \in \mathbb{C}^M$ はスケール補正係数，及び $\circ$ は要素毎の積を表す．また，式 (22) によって目的関数 (14) の値が変動することを防ぐために，他の変数も次のように補正する．

$$\boldsymbol{w}_{in} \leftarrow \boldsymbol{w}_{in} \lambda_{inm_{\mathrm{ref}}} \tag{23}$$

$$y_{ijn} \leftarrow \boldsymbol{w}_{in}^{\mathrm{H}} \boldsymbol{x}_{ij} \tag{24}$$

$$t_{ikn} \leftarrow t_{ikn} |\lambda_{inm_{\mathrm{ref}}}|^2 \tag{25}$$

ここで，m_{ref} はプロジェクションバック法で用いるリファレンスチャネルのインデクスである．

これらをまとめると，提案手法は Algorithm 1 に示すアルゴリズムによって実現される．

Algorithm 1 Consistent ILRMA

Input: $\{\boldsymbol{x}_{ij}\}_{i=1,j=1}^{I,J}$, maxIter

Output: $\{\boldsymbol{y}_{ij}\}_{i=1,j=1}^{I,J}$

Initialize $\{\boldsymbol{T}_n\}_{n=1}^{N}, \{\boldsymbol{V}_n\}_{n=1}^{N}, \{\boldsymbol{W}_i\}_{i=1}^{I}$
for iter $= 1, 2, \cdots,$ maxIter **do**
  Ensure consistency by (21) $\forall n$
  Update $\boldsymbol{W}_i$ by (15)–(18) $\forall i, j, n$
  Update $\boldsymbol{T}_n$ and $\boldsymbol{V}_n$ by (19), (20) $\forall i, j, k, n$
  Apply back projection by (22) $\forall i, j, n$
  Update parameters by (23)–(25) $\forall i, j, k, n$
end for

Table 1　Experimental conditions

Window function	Hann window
Window length	64, 128, 256, 512, 768, 1024 ms
Window shift length	1/16 or 1/2 of window length
Number of bases K for each source in ILRMA	10 for music signals and 2 for speech signals
Initialization of parameters	$\boldsymbol{W}_i$: identity matrix $\boldsymbol{T}_n$ and $\boldsymbol{V}_n$: random matrices
Number of iterations	100
Number of trials	5 with different random seeds
Reference channel m_{ref}	1

4　BSS 実験

4.1　実験条件

提案手法の有効性を確認するために，BSS 性能を従来の ILRMA と比較した．ここで，従来手法は提案手法から式 (21) の演算のみを省いたものとし，その他の条件は全て提案手法と統一している．実験には，文献 [14] と同じ音楽信号 4 種と音声信号 4 種及び混合系（インパルス応答）を用いた．本稿ではインパルス応答 JR2（$T_{60} = 470$ ms）の結果のみを示すが，文献 [15] には網羅的な実験結果が示されているので参照されたい．評価値は音源対歪み比（source-to-distortion ratio: SDR）[16] の改善量を用いた．その他の実験条件は Table 1 に示す通りである．

4.2　実験結果

Fig. 1 は，提案手法の各反復における矛盾成分のエネルギー値の例である．ここで，縦軸 $\|\mathcal{E}(\mathsf{Y})\|_2^2/\|\mathsf{X}\|_2^2$ 中の X 及び Y はそれぞれ $\mathsf{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2]$ 及び $\mathsf{Y} = [\boldsymbol{Y}_1, \boldsymbol{Y}_2]$ を表し，$\mathcal{E}$ の定義は式 (13) に示されている．矛盾成分のエネルギー値は反復初期において大きく増加するものの，反復後半では一定の値に収束している．このことから，提案手法の最適化ができるだけ無矛盾な解へと誘導する様子を確認できる．

Fig. 2 は，各窓長及びシフト長での平均 SDR 改善量の比較を示している．図中のラベル Conv. 及び Prop. はそれぞれ従来手法と提案手法に対応しており，異なる初期値に対する試行や 4 種類の信号に対する結果を全てまとめて箱ひげ図で示している．この結果より，式 (10) が成立しやすい長い窓長において ILRMA は高い性能を発揮することが確認できる．また提案手法は，従来手法の分離性能が比較的高い条件において性能が向上する傾向が確認でき，場合によっては 2〜3 dB 程度の改善が見られた．これは，音源分離が成功するほど，分離信号 $\boldsymbol{Y}_n$ は本来の音源信号 $\boldsymbol{S}_n$ に近づき，提案手法においてスペクトログラム無矛盾性を担保する効果が高くなるためと推測される．

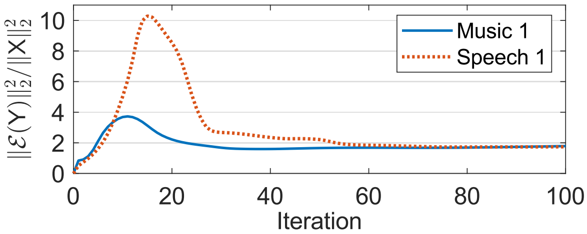

Fig. 1　Normalized energy of the inconsistent components of proposed method, where window and shift lengths were 256 ms and 32 ms, respectively.

5　おわりに

本稿では，反復最適化の中でスペクトログラム無矛盾性を担保する ILRMA を提案し，その有効性を実験的に確認した．今後の課題として，スペクトログラム無矛盾性を陽に加味した新しい音源モデルの導入や目的関数の設計等が挙げられる．

謝辞　本研究の一部は JSPS 科研費 19K20306 の助成を受けたものである．

参考文献

[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, pp. 21–34, 1998.
[3] 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.
[4] N. Ono, "Stable and fast update rules for independent vector analysis based on auxiliary function technique," *Proc. WASPAA*, pp. 189–192, 2011.
[5] 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.
[6] 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., pp. 125–155. Springer, Cham, 2018.
[7] 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. SAP*, vol. 12, no. 5, pp. 530–538, Sep. 2004.
[8] K. Yatabe and D. Kitamura, "Time-frequency-masking-based determined BSS with application to sparse IVA," *Proc. ICASSP*, pp. 715–719, 2019.
[9] K. Yatabe, "Consistent ICA: Determined BSS meets spectrogram consistency," *IEEE Signal Process. Lett.*, vol. 27, pp. 870–874, 2020.
[10] J. Le Roux and E. Vincent, "Consistent Wiener filtering for audio source separation," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 217–220, 2013.
[11] K. Yatabe, Y. Masuyama, T. Kusano, and Y. Oikawa, "Representation of complex spectrogram via phase conversion," *Acoust. Sci. & Tech.*, vol. 40, no. 3, pp. 170–177, 2019.
[12] 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, 2009.
[13] K. Matsuoka and S. Nakashima, "Minimal distortion principle for blind source separation," *Proc. ICA*, pp. 722–727, 2001.
[14] D. Kitamura, N. Ono, and H. Saruwatari, "Experimental analysis of optimal window length for independent low-rank matrix analysis," *Proc. EUSIPCO*, pp. 1210–1214, 2017.
[15] D. Kitamura and K. Yatabe, "Consistent independent low-rank matrix analysis for determined blind source separation," *arXiv:2007.00274*, 2020.
[16] 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.

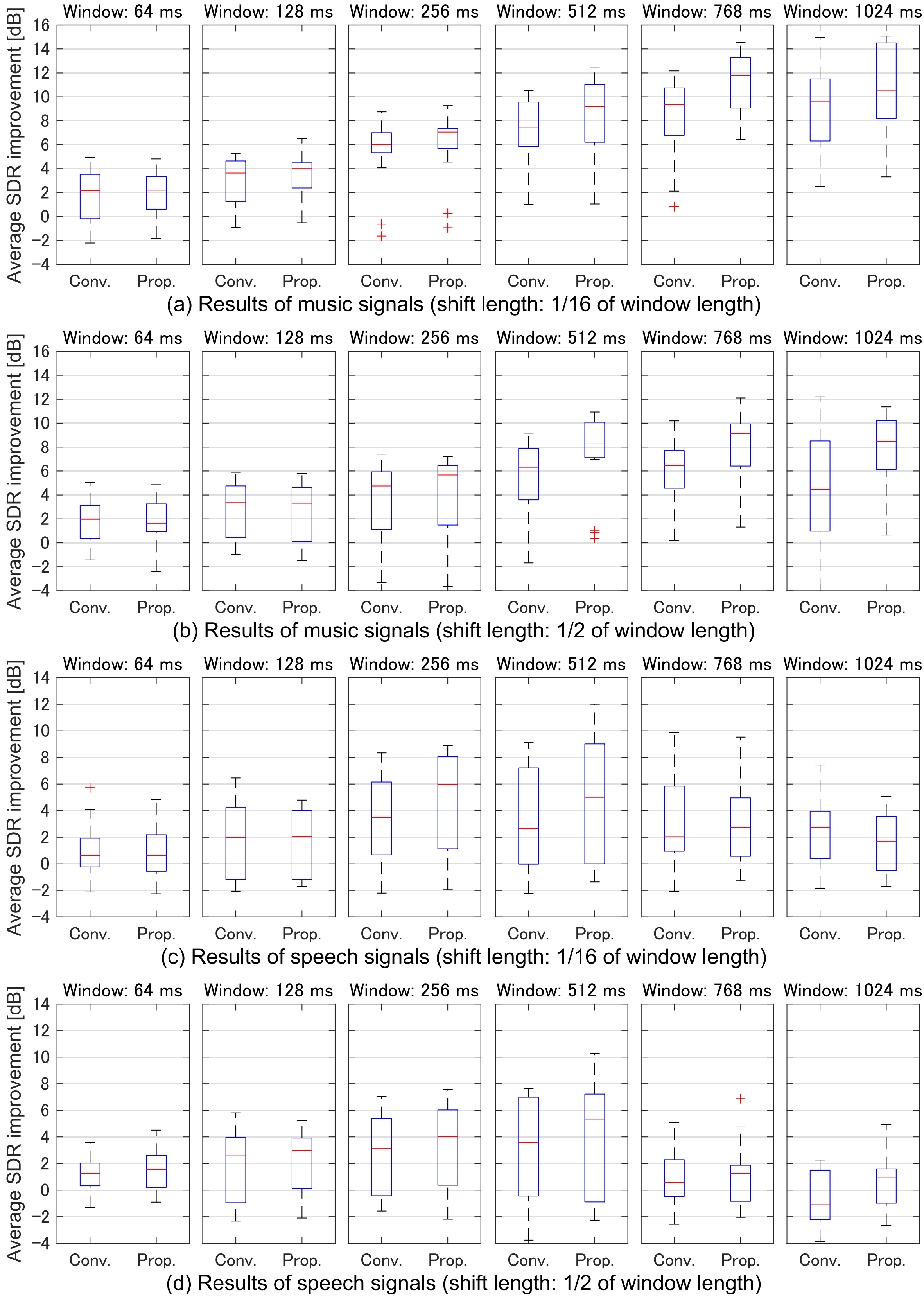

Fig. 2 SDR improvements for (a) music with 1/16 window shifting, (b) music with 1/2 window shifting, (c) speech with 1/16 window shifting, and (d) speech with 1/16 window shifting, where “Conv.” and “Prop.” respectively indicate conventional and proposed methods.