# 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の実験的評価*

○北村大地（香川高専），矢田部浩平（早稲田大）

## 1 はじめに

ブラインド音源分離（blind source separation: BSS）[1] は，混合系が未知の条件下で，複数の音源が混合した観測信号から混合前の音源信号を推定する技術である．BSS は音声認識精度向上や音楽信号解析等，様々な技術に応用されている．

音源数以上のマイクロホンで録音して得られる多チャネル観測信号を対象とする BSS（優決定条件 BSS）は，音源間の統計モデルを活用した 1994 年の独立成分分析（independent component analysis: ICA）[2] の登場以降盛んに研究されている．1998 年に周波数領域 ICA（frequency-domain ICA: FDICA）[3] が提案され，その後は FDICA で得られる周波数毎の分離信号の順番を適切に並び替えるパーミュテーション問題の解決が検討された [4, 5, 6]．2006 年には，音源の周波数構造を仮定した音源モデルを FDICA に導入することで，パーミュテーション問題を回避しながら分離信号を推定する独立ベクトル分析（independent vector analysis: IVA）[7, 8] が提案された．2011 年には，補助関数法 [9] 及び反復射影法（iterative projection: IP）[10] に基づく安定・高速な IVA（auxiliary-function-based IVA: AuxIVA）[11] が登場し，優決定条件 BSS におけるパーミュテーション問題のエレガントな回避法が，モデルと最適化の両観点から確立された．

パーミュテーション問題回避のために何らかの音源モデルを導入する IVA のアイデアは画期的であり，様々な音源モデルの導入へと発展した．例えば，非負値行列因子分解（nonnegative matrix factorization: NMF）[12] に基づく低ランク時間周波数構造を仮定する独立低ランク行列分析（independent low-rank matrix analysis: ILRMA）[13, 14]，学習データから音源モデルを深層学習する独立深層学習行列分析 [15] 等が提案されている．また，これらの音源モデルを plug-and-play で変更可能な最適化アルゴリズムを採用した BSS [16, 17] も提案されている．

一方，時間周波数領域におけるスペクトログラム無矛盾性 [18, 19] と呼ばれる性質を FDICA 及び IVA に導入した BSS が近年提案された [20]．スペクトログラムは通常，短時間フーリエ変換（short-time Fourier transform: STFT）における窓関数とそのオーバーラップの影響によって，近傍時間周波数グリッドに共起関係を持つ．スペクトログラムに対して BSS 等の信号処理を施した場合，この共起関係が崩されることにより，矛盾したスペクトログラムが生成される．このようなスペクトログラムの矛盾・無矛盾性に基づいて，最適化の過程で無矛盾性を担保しパーミュテーション問題をある程度回避する手法がスペクトログラム無矛盾性に基づく BSS である．著者らはこれを ILRMA に適用し，スペクトログラム無矛盾性に基づく ILRMA（consistent ILRMA） [21, 22] を提案している．

本稿では，consistent ILRMA の実験的評価として，最適化の毎反復における分離信号のスケール補正（プロジェクションバック法（back projection: BP）[23]）の有無が性能に与える影響を新たに明らかにする．なお，実験に用いた ILRMA 及び consistent ILRMA の MATLAB コードを GitHub（https://github.com/d-kitamura/ILRMA）にて公開しているため，文献 [22] と合わせて参照されたい．

## 2 Consistent ILRMA

### 2.1 定式化と優決定条件 BSS

離散時間信号の $l$ 番目のサンプルを $x[l]$ と表記し，$N$ 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，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．

信号 $\boldsymbol{z} = [\, z[1], \cdots, z[l], \cdots, z[L] \,]^{\mathrm{T}} \in \mathbb{R}^L$ の STFT を次式で表記する．

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

ここで，$\boldsymbol{\omega}$ 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 $\mathrm{ISTFT}_{\tilde{\omega}}(\cdot)$ 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．

$$\boldsymbol{z} = \mathrm{ISTFT}_{\tilde{\omega}}(\mathrm{STFT}_{\boldsymbol{\omega}}(\boldsymbol{z})) \qquad \forall \boldsymbol{z} \in \mathbb{R}^L \tag{5}$$

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．

$$\boldsymbol{s}_{ij} = \left[\, s_{ij1}, \cdots, s_{ijn}, \cdots s_{ijN} \,\right]^{\mathrm{T}} \quad \in \mathbb{C}^N \tag{6}$$

$$\boldsymbol{x}_{ij} = \left[\, x_{ij1}, \cdots, x_{ijm}, \cdots x_{ijM} \,\right]^{\mathrm{T}} \quad \in \mathbb{C}^M \tag{7}$$

$$\boldsymbol{y}_{ij} = \left[\, y_{ij1}, \cdots, y_{ijn}, \cdots y_{ijN} \,\right]^{\mathrm{T}} \quad \in \mathbb{C}^N \tag{8}$$

ここで，$i = 1, 2, \cdots, I$ 及び $j = 1, 2, \cdots, J$ はそれぞれ周波数ビン及び時間フレームのインデクスを表す．また，式 (8) の時間周波数行列を $\boldsymbol{Y}_n \in \mathbb{C}^{I \times J}$ とする．周波数領域 BSS は，次式の瞬時混合を仮定する．

$$\boldsymbol{x}_{ij} = \boldsymbol{A}_i \boldsymbol{s}_{ij} \tag{9}$$

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．

優決定条件 BSS では $M = N$ を仮定でき，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．

$$\boldsymbol{y}_{ij} = \boldsymbol{W}_i \boldsymbol{x}_{ij} \tag{10}$$

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．

---

*Experimental evaluation of consistent independent low-rank matrix analysis. By Daichi Kitamura (NIT Kagawa) and Kohei Yatabe (Waseda Univ.).

IVA や ILRMA 等の ICA に基づく BSS では，分離信号に次式の任意性が存在する．

$$\hat{\boldsymbol{y}}_{ij} = \hat{\boldsymbol{W}}_i \boldsymbol{x}_{ij} \quad (\hat{\boldsymbol{W}}_i = \boldsymbol{D}_i \boldsymbol{P}_i \boldsymbol{W}_i) \tag{11}$$

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 BP によって解析的に復元できるが，周波数毎の分離信号の順序の全周波数における整列（$\boldsymbol{P}_1 = \boldsymbol{P}_2 = \cdots = \boldsymbol{P}_I$ とする処理）はパーミュテーション問題と呼ばれる．

### 2.2 スペクトログラム無矛盾性

スペクトログラムは本来，文献 [22] の Fig. 1 のように，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

$$\mathcal{E}(\boldsymbol{Z}) = \boldsymbol{Z} - \mathrm{STFT}_{\boldsymbol{\omega}}(\mathrm{ISTFT}_{\widetilde{\boldsymbol{\omega}}}(\boldsymbol{Z})) \tag{12}$$

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．

パーミュテーション問題が生じた分離信号のスペクトログラムは，文献 [22] の Fig. 2 のように隣接周波数成分が不連続となるため，無矛盾性が大きく損なわれる．Consistent ILRMA [21, 22] では，ILRMA で仮定されていた NMF に基づく低ランク時間周波数音源モデルに加えて，分離信号 $\boldsymbol{Y}_n$ のスペクトログラム無矛盾性を最適化の過程で常に担保する．即ち，最適化の毎反復において，分離信号 $\boldsymbol{Y}_n$ を $\mathrm{STFT}_{\boldsymbol{\omega}}(\mathrm{ISTFT}_{\widetilde{\boldsymbol{\omega}}}(\boldsymbol{Y}_n))$ に更新する．その結果，周波数毎の音源成分が隣接周波数間で正しく整列された状態に誘導され，ILRMA よりも高精度にパーミュテーション問題を回避できる．

### 2.3 最適化アルゴリズム

Consistent ILRMA の最適化アルゴリズムは，従来の ILRMA における分離行列 $\boldsymbol{W}_i$ と NMF 音源モデル $\boldsymbol{T}_n \in \mathbb{R}^{I\times K}_{\geq 0}$ 及び $\boldsymbol{V}_n \in \mathbb{R}^{K\times J}_{\geq 0}$ の反復更新則に，分離信号 $\boldsymbol{Y}_n$ が無矛盾となる更新及び BP を加えたものとなる．ここで，$K$ は NMF の基底数を表す．この最適化アルゴリズムを Algorithm 1 に示す[1]．ここで，行列間の演算子 $\odot$ 及び分数はそれぞれ要素毎の積及び商，行列に対する $|\cdot|$ 及びドット付き指数はそれぞれ要素毎の絶対値及び指数乗，$[\cdot]_{r,c}$ は行列の $(r,c)$ 番目の要素を表す．また，$\boldsymbol{e}_n \in \{0,1\}^N$ は $n$ 番目の要素のみ 1 の単位ベクトル，$m_{\mathrm{ref}}$ は BP のためのリファレンスチャネルを表す．Algorithm 1 中の 3 行目が $\boldsymbol{Y}_n$ の無矛盾性を担保する更新，4 及び 5 行目が NMF 音源モデルの更新，6–8 行目が IP に基づく分離行列の更新，9–12 行目が $m_{\mathrm{ref}}$ 番目のチャネルへの BP を表す．従来の ILRMA は 4–8 及び 11 行目のみであり，BP は反復更新が収束した後に一度のみ適用される．

**Algorithm 1** Consistent ILRMA

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

1: Initialize $\{\boldsymbol{T}_n\}_{n=1}^{N}, \{\boldsymbol{V}_n\}_{n=1}^{N}, \{\boldsymbol{W}_i\}_{i=1}^{I}$
2: **for** iter $= 1, 2, \cdots,$ maxIter **do**
3: $\quad \boldsymbol{Y}_n \leftarrow \mathrm{STFT}_{\boldsymbol{\omega}}(\mathrm{ISTFT}_{\widetilde{\boldsymbol{\omega}}}(\boldsymbol{Y}_n))\ \forall n$
4: $\quad \boldsymbol{T}_n \leftarrow \boldsymbol{T}_n \odot \left\{ \frac{\left[|\boldsymbol{Y}_n|^{\cdot 2} \odot (\boldsymbol{T}_n\boldsymbol{V}_n)^{\cdot -2}\right]\boldsymbol{V}_n^{\mathrm{T}}}{(\boldsymbol{T}_n\boldsymbol{V}_n)^{\cdot -1}\boldsymbol{V}_n^{\mathrm{T}}} \right\}^{\cdot \frac{1}{2}}\ \forall n$
5: $\quad \boldsymbol{V}_n \leftarrow \boldsymbol{V}_n \odot \left\{ \frac{\boldsymbol{T}_n^{\mathrm{T}}\left[|\boldsymbol{Y}_n|^{\cdot 2} \odot (\boldsymbol{T}_n\boldsymbol{V}_n)^{\cdot -2}\right]}{\boldsymbol{T}_n^{\mathrm{T}}(\boldsymbol{T}_n\boldsymbol{V}_n)^{\cdot -1}} \right\}^{\cdot \frac{1}{2}}\ \forall n$
6: $\quad \boldsymbol{U}_{in} \leftarrow \frac{1}{J}\sum_j \frac{1}{[\boldsymbol{T}_n\boldsymbol{V}_n]_{i,j}} \boldsymbol{x}_{ij}\boldsymbol{x}_{ij}^{\mathrm{H}}\ \forall i, n$
7: $\quad \boldsymbol{w}_{in} \leftarrow (\boldsymbol{W}_i\boldsymbol{U}_{in})^{-1}\boldsymbol{e}_n\ \forall i, n$
8: $\quad \boldsymbol{w}_{in} \leftarrow \boldsymbol{w}_{in}\left(\boldsymbol{w}_{in}^{\mathrm{H}}\boldsymbol{U}_{in}\boldsymbol{w}_{in}\right)^{-\frac{1}{2}}\ \forall i, n$
9: $\quad \lambda_{in} \leftarrow \left[\boldsymbol{W}_i^{-1}\right]_{m_{\mathrm{ref}},n}\ \forall i, n$
10: $\quad \boldsymbol{w}_{in} \leftarrow \lambda_{in}\boldsymbol{w}_{in}\ \forall i, n$
11: $\quad y_{ijn} \leftarrow \boldsymbol{w}_{in}^{\mathrm{H}}\boldsymbol{x}_{ij}\ \forall i, j, n$
12: $\quad [\boldsymbol{T}_n]_{i,k} \leftarrow |\lambda_{in}|^2\,[\boldsymbol{T}_n]_{i,k}\ \forall i, k, n$
13: **end for**

## 3 実験的評価

### 3.1 実験条件

前述のように，ILRMA における分離行列の推定には周波数毎のスケールの任意性が生じる．この任意性は式 (11) の $\boldsymbol{D}_i$ となって現れるため，パーミュテーション問題を引き起こす $\boldsymbol{P}_i$ と同様に分離信号 $\boldsymbol{Y}_n$ のスペクトログラム無矛盾性を崩す要因となる．この影響を避けるために，consistent ILRMA では毎反復の無矛盾性を担保する処理の直前に BP を施してスケールを補正している．しかしながら，BP の有無で性能がどのように変化するかについては調査されていなかった．本稿では，毎反復の BP が consistent ILRMA の分離性能に与える影響を実験的に評価する．

本実験は音楽信号の BSS のみを対象とし，インパルス応答による畳み込み混合信号と実録音混合信号で評価する．比較手法には，AuxIVA（`IVA`），毎反復においてスペクトログラム無矛盾性を担保するが BP は適用しない AuxIVA（`Consis. IVA w/o BP`）及び BP を適用する AuxIVA（`Consis. IVA w/ BP`），従来の ILRMA（`ILRMA`），毎反復においてスペクトログラム無矛盾性を担保するが BP は適用しない ILRMA（`Consist. ILRMA w/o BP`）及び BP を適用する ILRMA（`Consist. ILRMA w/ BP`）の計 6 手法を用いた．STFT における窓関数は Hann 窓を用いた．その他の実験条件は文献 [22] を参照されたい．

### 3.2 インパルス応答による畳み込み混合の BSS

インパルス応答 E2A（$T_{60} = 300$ ms）を用いて，2 音源が混合した 2 チャネル観測信号を 10 曲分生成した．音源やインパルス応答の詳細は文献 [22] に記載の通りである．評価指標は総合的な分離尺度である source-to-distortion ratio（SDR）[24] の改善量 ΔSDR を用いた．10 曲の観測信号のそれぞれに対して，5 種類の乱数で初期化した計 50 回試行の ΔSDR の箱ひげ図を Fig. 1 に示す．ここで，STFT における窓長（window len.）とシフト長及び窓長の比率（shift len.）を変化させた際の結果をまとめて示している．

この結果を見ると，多くの実験条件においてスペクトログラム無矛盾性の担保が分離性能を向上させていることが確認できる．また，スペクトログラム無矛盾性を担保する直前に，BP を適用しスケール不定性を解消することで，より大きな性能改善が得られることが分かる．例えば，窓長 512 ms かつシフト長 1/2 の例では，consistent ILRMA は BP の有無によって中央値で 4 dB 弱もの改善があり，BP を適用する効果の大きさがうかがえる．この理由として，分離信号 $\boldsymbol{Y}_n$ のスケール不定性 $\boldsymbol{D}_i$ の影響を解消したうえでス

[1] 文献 [21] に記載の Algorithm 1 は 4 行目と 5 行目が逆（$\boldsymbol{W}_i$ と $\boldsymbol{T}_n$ 及び $\boldsymbol{V}_n$ の更新順が逆）の誤りがある．

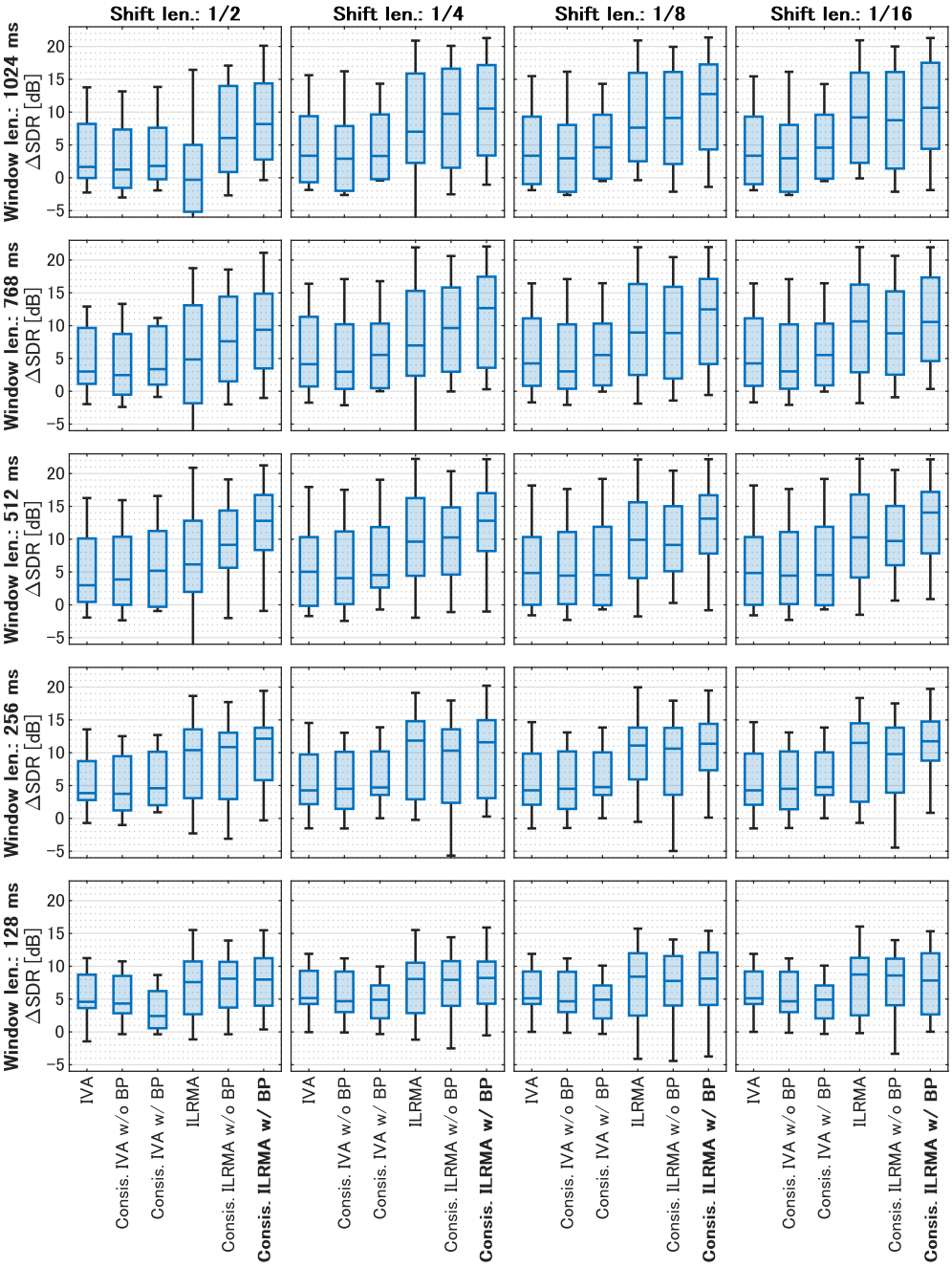


Fig. 1　SDR improvements for artificially mixed signals. The central lines of the box plots indicate the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively.

ペクトログラム無矛盾性を担保する更新を行う方が，$\boldsymbol{Y}_n$ 中のパーミュテーション問題 $\boldsymbol{P}_i$ の解消により強く作用することが考えられる．

### 3.3　実録音混合信号の BSS

残響時間 250 ms の環境で実際に録音された 2 音源混合の 2 チャネル観測信号を 12 曲分用意した．音源の詳細は文献 [22] に記載の通りである．また，ΔSDR に加え，source-to-interference ratio（SIR）[24] の改善量 ΔSIR 及び sources-to-artifact ratio（SAR）[24] も算出した．各曲に対して 5 種類の乱数で初期化した計 60 回試行の箱ひげ図を Fig. 2 に示す．但し，本実験では窓長を 512 ms，シフト長を 1/4 としている．

Fig. 2 より，実録音混合信号においても consistent ILRMA の優位性が確認できる．また，毎反復の BP は ILRMA と IVA の両手法に対して有効であり，特に IVA での性能改善（`Consis. IVA w/o BP` と `Consis. IVA w/ BP` の差）が顕著であった．SIR 及び SAR を見ると，毎反復の BP は SAR の性能改善に大きく寄与している．SAR は BSS で生じる人工歪みの少なさを表すことから，毎反復の BP でスケール不定性を解消した後にスペクトログラム無矛盾性を担保するという処理の流れが，より歪みの少ない分離信号 $\boldsymbol{Y}_n$ の推定に有効であることを示している．

## 4　おわりに

本稿では，consistent ILRMA において，毎反復で BP を適用する効果を実験的に評価した．実験結果より，consistent ILRMA における BP は分離性能の改善に寄与していることが確認された．また，この事実は ILRMA だけでなく IVA の分離実験でも確認され

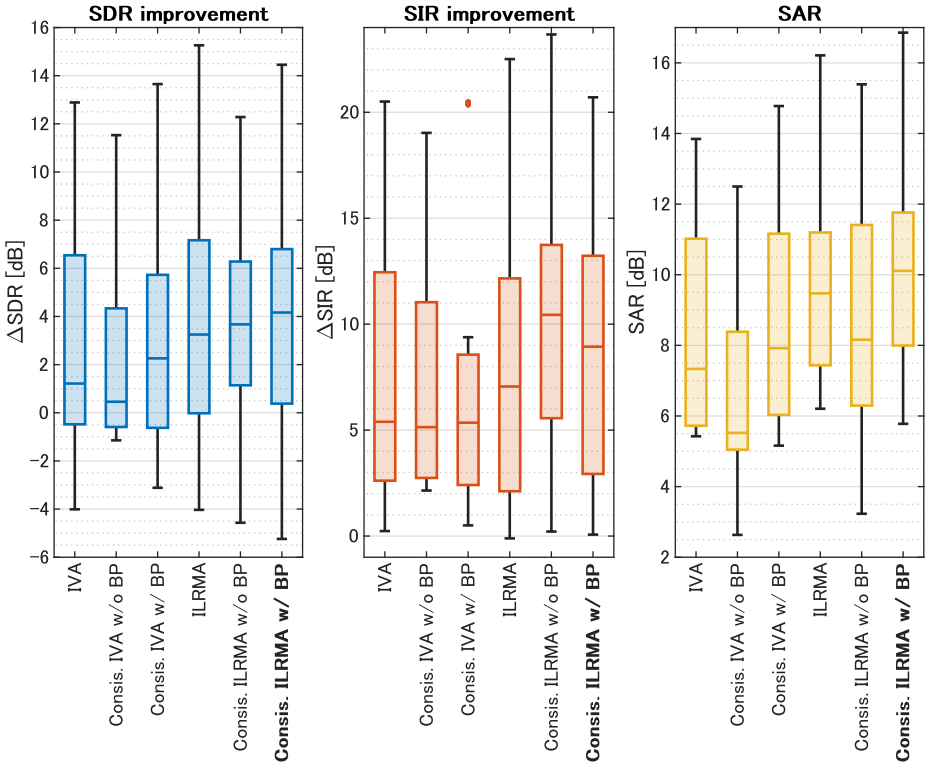


Fig. 2　Evaluation scores for live-recorded mixture signals. The left, center, and right figures show SDR improvements, SIR improvements, and SAR values, respectively.

たことから，あらゆる周波数領域 BSS においてスペクトログラム無矛盾性を担保する場合に BP を適用することが望ましい可能性が示唆された．

**謝辞**　本研究の一部は JSPS 科研費 19K20306 及び 19H01116 の助成を受けたものである．

## 参考文献

[1] H. Sawada, N. Ono, H. Kameoka, D. Kitamura, and H. Saruwatari, "A review of blind source separation methods: Two converging routes to ILRMA originating from ICA and NMF," APSIPA Trans. Signal and Info. Process., vol. 8, no. e12, pp. 1–14, 2019.

[2] P. Comon, "Independent component analysis, a new concept?," *Signal Process.*, vol. 36, no. 3, pp. 287–314, 1994.

[3] P. Smaragdis, "Blind separation of convolved mixtures in the frequency domain," *Neurocomputing*, vol. 22, pp. 21–34, 1998.

[4] N. Murata, S. Ikeda, and A. Ziehe, "An approach to blind source separation based on temporal structure of speech signals," *Neurocomputing*, vol. 41, no. 1–4, pp. 1–24, 2001.

[5] 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.

[6] 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.

[7] A. Hiroe, "Solution of permutation problem in frequency domain ICA using multivariate probability density functions," *Proc. ICA*, pp.601–608, 2006.

[8] 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.

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

[10] N. Ono and S. Miyabe, "Auxiliary-function-based independent component analysis for super-Gaussian sources," *Proc. LVA/ICA*, pp. 165–172, 2010.

[11] N. Ono, "Stable and fast update rules for independent vector analysis based on auxiliary function technique," *Proc. WASPAA*, pp. 189–192, 2011.

[12] 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.

[13] 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.

[14] 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.

[15] N. Makishima, S. Mogami, N. Takamune, D. Kitamura, H. Sumino, S. Takamichi, H. Saruwatari, and N. Ono, "Independent deeply learned matrix analysis for determined audio source separation," *IEEE/ACM Trans. ASLP*, vol. 27, no. 10, pp. 1601–1615, 2019.

[16] K. Yatabe and D. Kitamura, "Determined blind source separation via proximal splitting algorithm," *Proc. ICASSP*, pp. 776–780, 2018.

[17] K. Yatabe and D. Kitamura, "Time-frequency-masking-based determined BSS with application to sparse IVA," *Proc. ICASSP*, pp. 715–719, 2019.

[18] J. L. Roux, H. Kameoka, N. Ono, and S. Sagayama, "Fast signal reconstruction from magnitude STFT spectrogram based on spectrogram consistency," *Proc. DAFx*, 2010.

[19] 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.

[20] K. Yatabe, "Consistent ICA: Determined BSS meets spectrogram consistency," *IEEE Signal Process. Lett.*, vol. 27, pp. 870–874, 2020.

[21] 豊島直，北村大地，矢田部浩平，"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," **日本音響学会** *2020* **年秋季研究発表会講演論文集**, pp. 291–294, 2020.

[22] D. Kitamura and K. Yatabe, "Consistent independent low-rank matrix analysis for determined blind source separation," *EURASIP J. Adv. Signal Process.*, vol. 2020, no. 46, 35 pages, 2020.

[23] K. Matsuoka and S. Nakashima, "Minimal distortion principle for blind source separation," *Proc. ICA*, pp. 722–727, 2001.

[24] 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.