

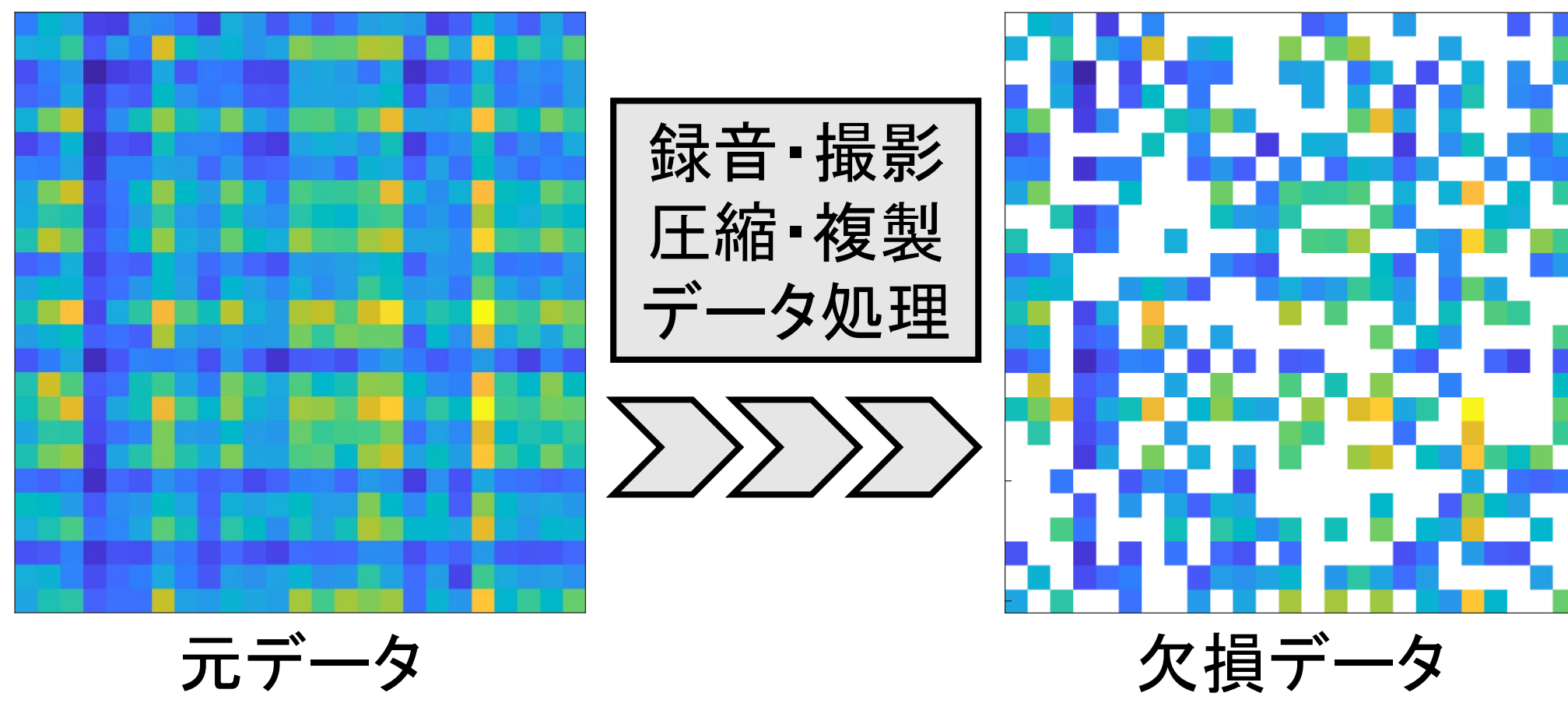
上界最小化アルゴリズムに基づくスパースNMFを用いた欠損値補完と音源分離への応用

☆和気佑弥, 北村大地(香川高専)

1. 研究背景

欠損値

- データには諸要因によって欠損値が発生する
 - 画像の白飛び/黒潰れ
 - アンケートの未回答
 - マイクの音割れ



欠損値を補完することでデータ品質を保ちたい

非負値行列因子分解 (nonnegative matrix factorization: NMF) [Lee+, 1999]

- 非負制約付きの任意基底数 K による低ランク近似

$$X(I \times J) \approx W(I \times K) H(K \times J) = \hat{X}(I \times J)$$

非負値観測行列 X (サイズ $I \times J$) は基底行列 W (サイズ $I \times K$) と係数行列 H (サイズ $K \times J$) の積として近似される。ここで W は基底 (Basis) と呼ばれ、 H はアクティベーション (Activation) と呼ばれる。

2. NMFによる行列補完

NMFを用いたナイーブな行列補完 [Kitamura+, 2014]

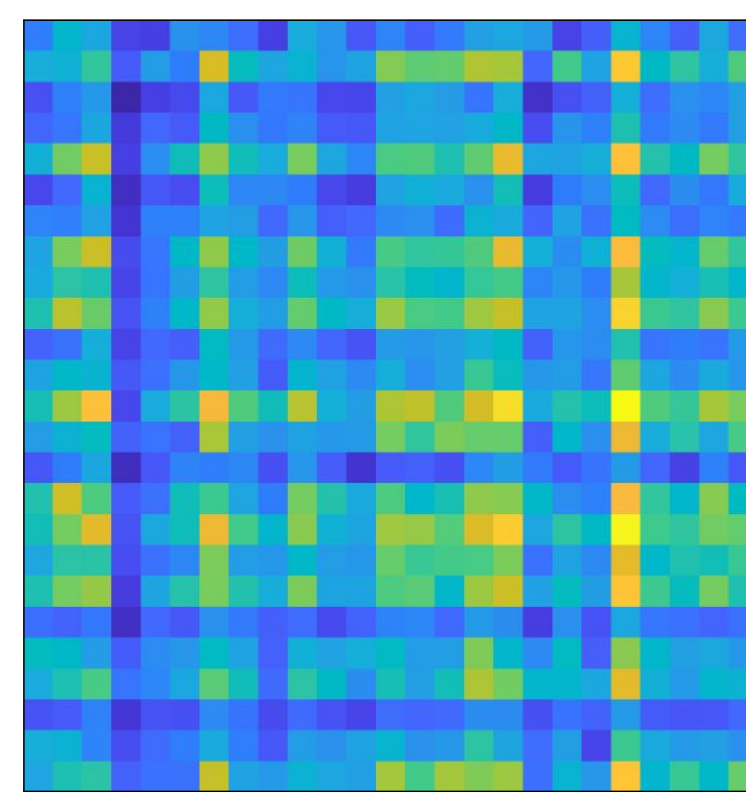
- 欠損要素を無視してNMFのコスト関数値を最小化

$$\text{Minimize}_{W, H} \sum_{i,j} m_{ij} D(x_{ij} | \sum_k w_{ik} h_{kj})$$

s.t. $w_{ik}, h_{kj} \geq 0 \quad \forall i, j, k$

- 低ランクかつ元データの基底数が既知の場合において高精度に補完可能

実際は... ノイズによりフルランク化する欠損データから基底数の予測は困難



上記欠損データのNMFに基づく補完

正則化付きNMFを用いた行列補完 [和気ら, 2023]

- 正則化により補完精度が向上する
- L_1 ノルムを用いたスパース正則化付きNMFにおいて性能向上

3. 提案手法

スパース正則化付きNMF (sparse NMF: SNMF)

- SNMFに基づく行列補完の最適化問題

$$\text{Minimize}_{W, H} \sum_{i,j} m_{ij} D(x_{ij} | \sum_k w_{ik} h_{kj}) + \mu \sum_{k,j} h_{kj}$$

s.t. $w_{ik}, h_{kj} \geq 0 \quad \forall i, j, k$

重み係数 μ

- 重み係数 μ で正則化項が最適化に与える影響の強さを調整する
- 様々な手法により更新式が導かれる

- ヒューリスティック手法によるSNMF (Heuristic SNMF) の更新式

$$W \leftarrow W \otimes \frac{M \otimes X H^T}{W H^T}, \quad H \leftarrow H \otimes \frac{W^T M \otimes X}{W^T M + \mu}$$

- 更新毎に変数 W 及び H の正規化及びリスケールが必要

$$W \leftarrow W \text{diag}(c), \quad H \leftarrow \text{diag}(c)^{-1} H, \quad c_k = \sum_{i=1}^I w_{ik}, \quad \forall k$$

正規化 (Normalization) リスケール (Rescaling)

- H のリスケールを回避したSNMF (Well-done SNMF) の更新式

- H のリスケールによる正則化項の変動を抑制可能 [Le Roux+, 2015]
- H の更新式はHeuristic SNMFと同様かつリスケールが不要 $E \in \{1\}^{I \times J}$

$$W \leftarrow W \otimes \frac{(M \otimes X) H^T + W \otimes (E E^T (W \otimes (M H^T)))}{M H^T + W \otimes (E E^T (W \otimes (M \otimes X H^T)))}$$

- 上界最小化アルゴリズムに基づくSNMF (MM-based SNMF) の

- 行列補完の最適化問題 [Marmin+, 2023]

$$\text{Minimize}_{W, H} \sum_{i,j} m_{ij} D(x_{ij} | \sum_k w_{ik} h_{kj}) + \mu \sum_{i,k,j} w_{ik} h_{kj}$$

s.t. $w_{ik}, h_{kj} \geq 0 \quad \forall i, j, k$

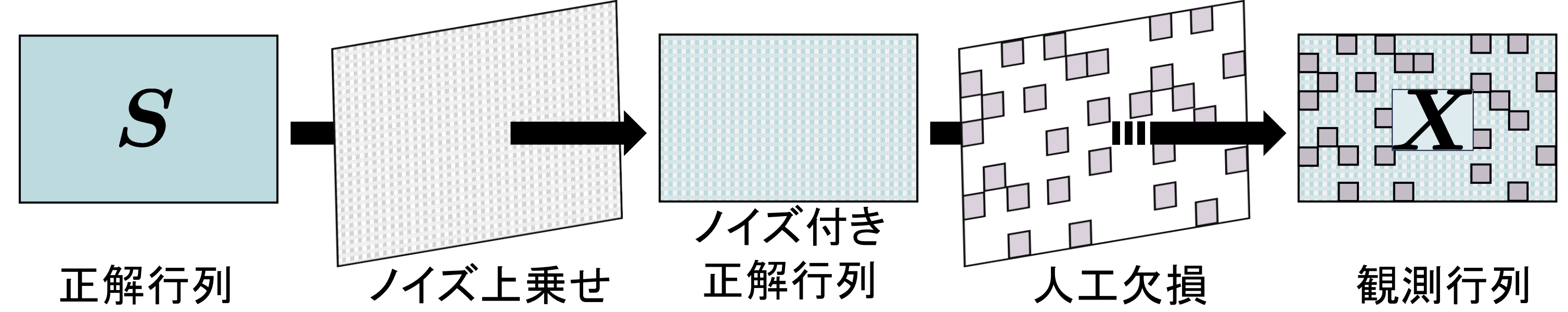
- MM-based SNMFの更新式

- 正規化及びリスケールが不要で、更新による単調収束が保証される

$$W \leftarrow W \otimes \frac{M \otimes X H^T}{M H^T + \mu W^T \mathbf{1}}, \quad H \leftarrow H \otimes \frac{W^T M \otimes X}{W^T M + \mu \mathbf{1} H^T}$$

4. 人工欠損行列の補完実験

乱数生成した人工欠損行列の補完実験

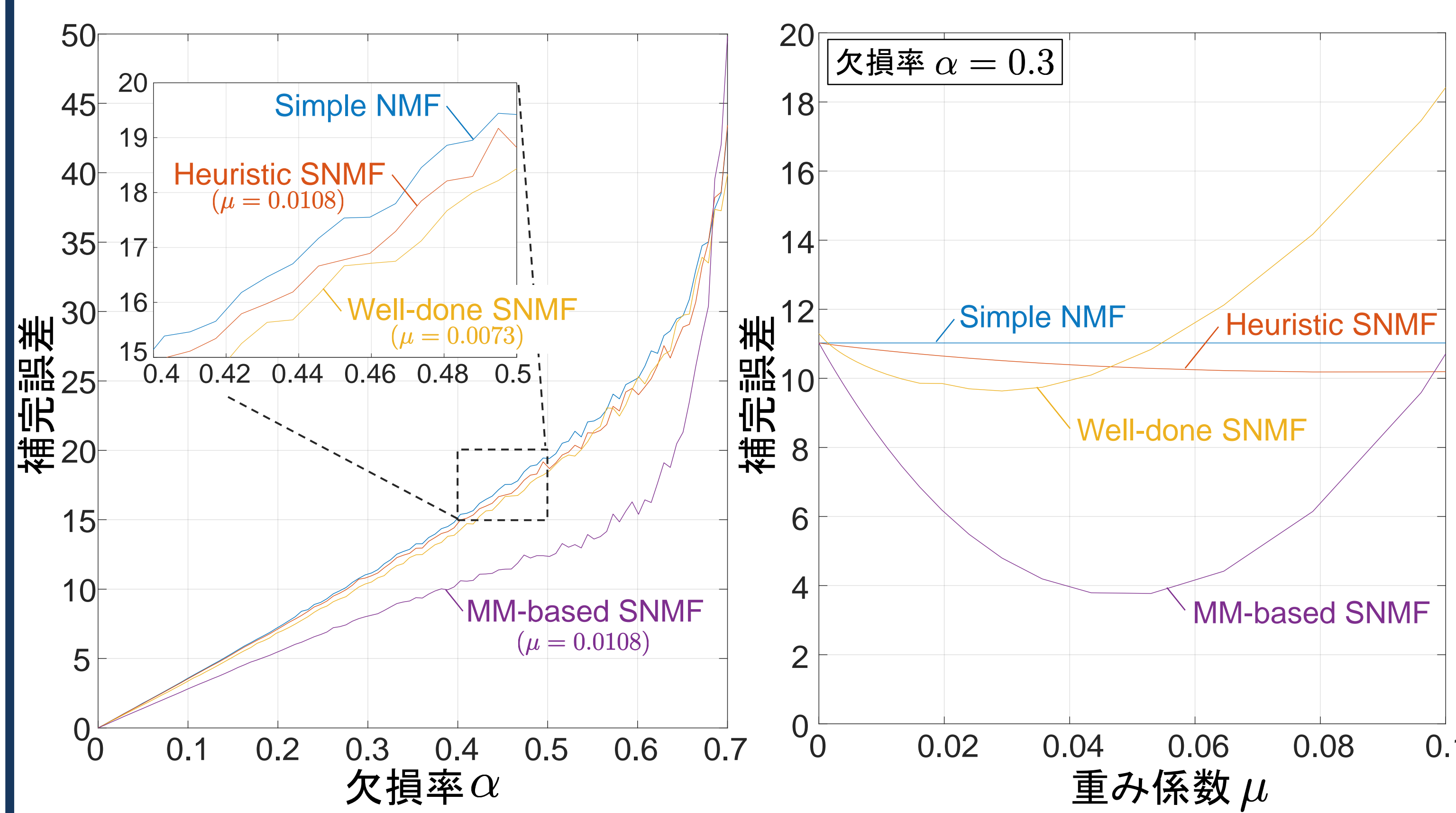


- 一様乱数を用いたサイズ 100×200 , ランク10の正解行列 S の生成
- 所望のSNR [dB]と欠損率 α に沿ったノイズと欠損値を生成
- 観測行列を行列補完し補完精度を各手法で比較 → 観測行列 X

- 比較手法: 正則化無しNMF (Simple NMF) 及び各種SNMF
- 補完誤差: 観測行列の欠損要素の真値と補完値のKL擬距離
- 欠損率 α , 重み係数 μ の複数条件を検証 (SNR:20dB)

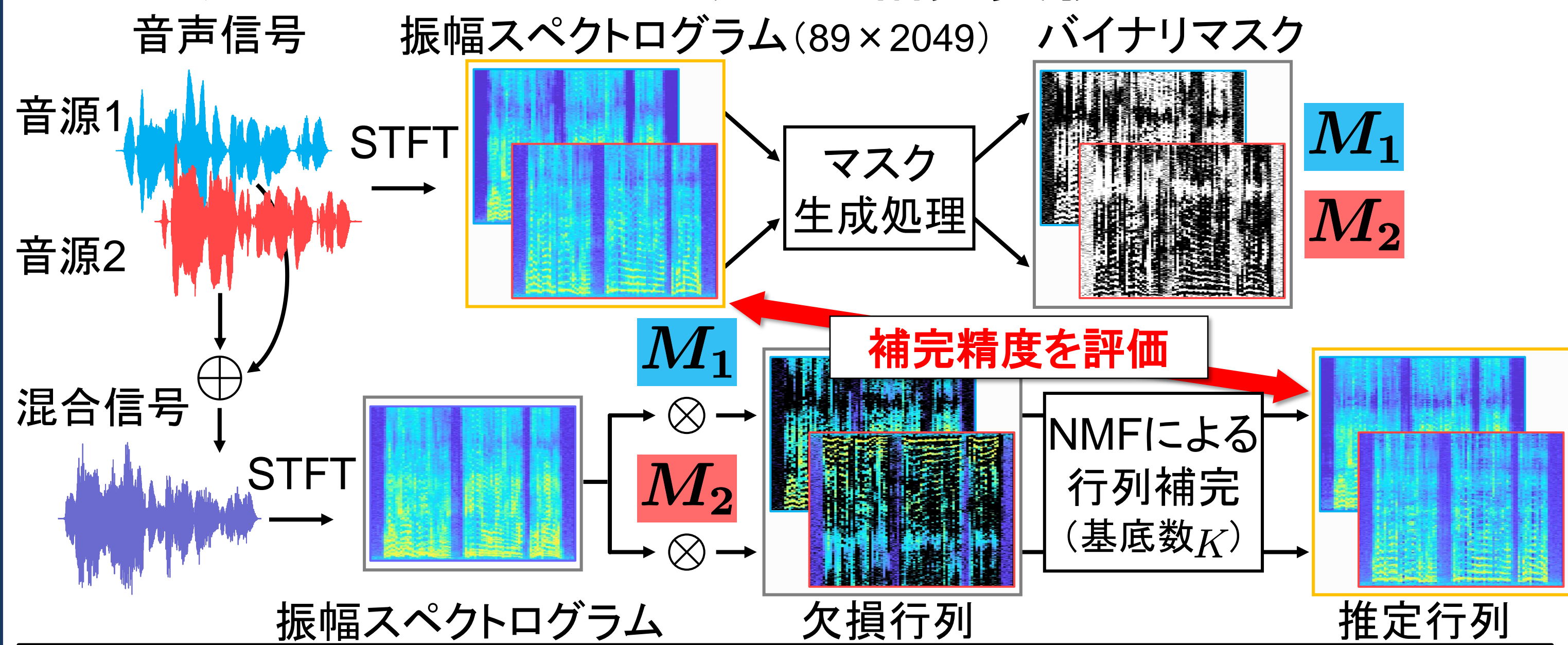
実験結果 (左図: 各手法最良の μ で α を推移, 右図: $\alpha = 0.3$ において μ を推移)

- 多くの条件でMM-based SNMFが最も補完精度が高い



5. スペクトログラムの補完実験

バイナリマスクスペクトログラムの補完実験



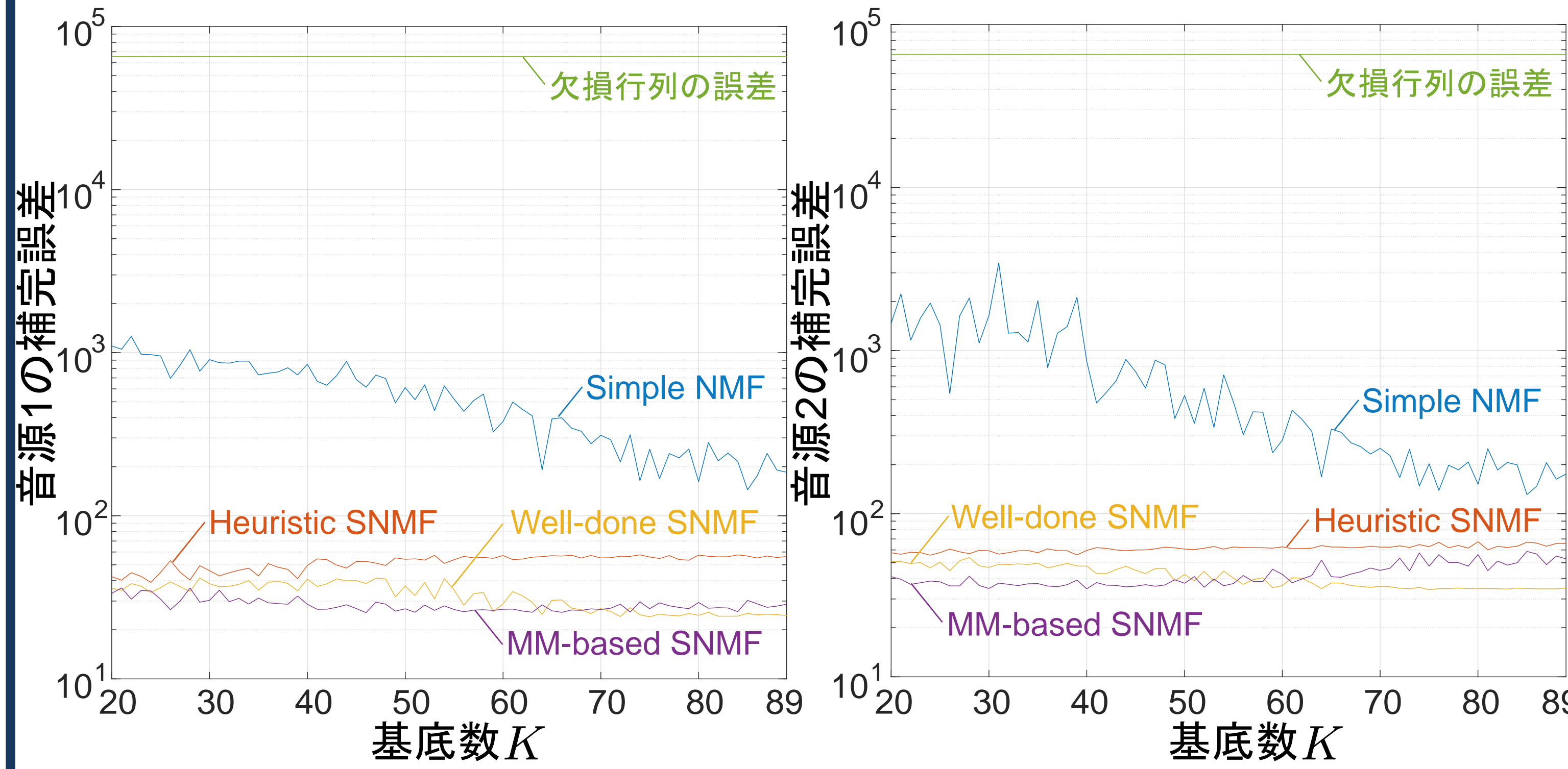
- 2音源を混合し時間周波数領域へ変換 (STFT)
- 各音源をもとに生成した理想バイナリマスクを掛け欠損行列を生成
- 欠損行列を行列補完し無補完時と補完誤差を比較

- 音源にはJVS corpus [Takamichi+, 2019]の"parallel"を使用
- 音源1: "jvs001/parallel100/wav24kHz16bit/VOICEACTRESS100_018.wav"
- 音源2: "jvs002/parallel100/wav24kHz16bit/VOICEACTRESS100_018.wav"

- 評価値: 欠損行列の欠損要素の真値と補完値のKL擬距離

実験結果 (最良の重み係数 μ における基底数 K を変化させたときの補完誤差の推移)

- K が小さいときMM-based SNMFが最も補完精度が高い
- K が大きいときWell-done SNMFも同等の補完精度



Heuristic SNMF: $\mu = 0.3162$, Well-done SNMF: $\mu = 0.0268$, MM-based SNMF: $\mu = 0.0268$