

本発表の概要

- NMFにスパース正則化を付与した**SNMF**について掘り下げる
- 本発表では**SNMFの単調非増加性を保証した乗法更新式**を提案する
- 実験では単調非増加性を維持したまま**従来手法と同等以上の性能**を示した

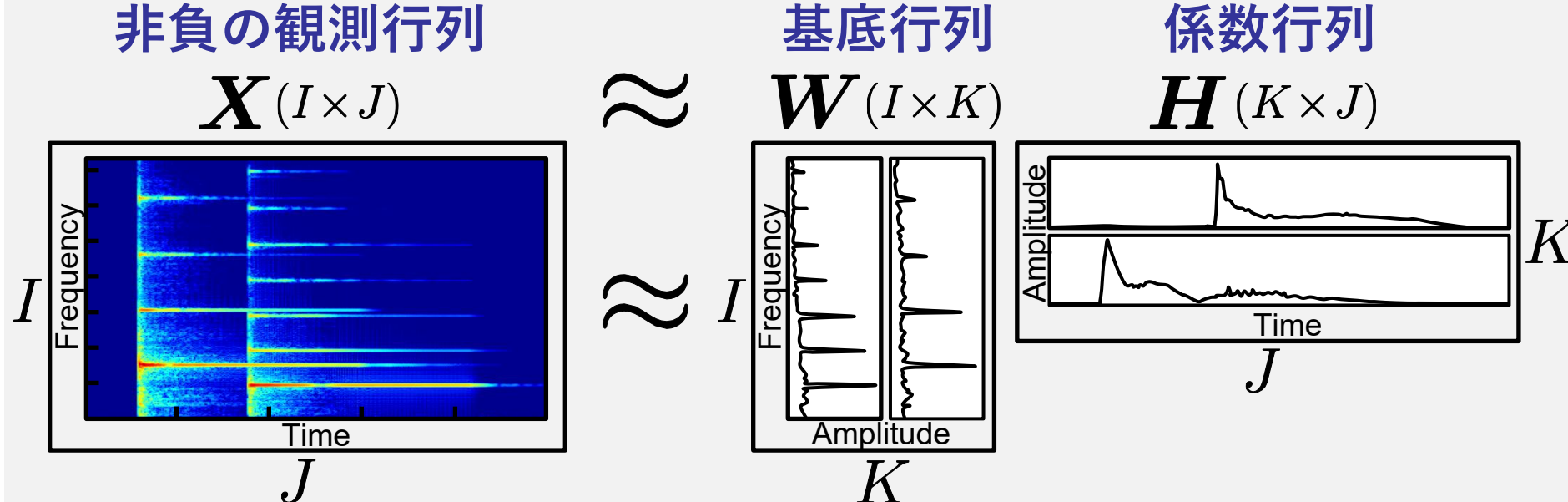
1. 研究背景

- 非負値行列因子分解(Nonnegative matrix factorization: NMF)

低ランク近似手法

NMFの最適化問題

$$\text{Minimize } \mathcal{D}(X|WH) \\ \text{s.t. } w_{ik}, h_{kj} \geq 0 \quad \forall i, j, k$$



- スパースNMF (sparse NMF: SNMF) とスケール任意性問題

$$\text{Minimize } \mathcal{D}(X|WH) + \mu \|H\|_1 \\ \text{s.t. } w_{ik}, h_{kj} \geq 0 \quad \forall i, j, k$$

- 損失関数 $\mathcal{D}$
- L_1 ノルム $\|H\|_1$ の同時最小化

スパースな H が推定される

SNMFのスケール任意性問題

WH のスケール任意性で
正則化項が**無意味**となる

$$\mathcal{D}(X|WH) + \mu \|H\|_1 \quad \text{スケール任意性} \\ WH = (\alpha W)(H/\alpha) \\ \text{一定 } \mathcal{D}(X|(\alpha W)(H/\alpha)) + \mu \|(H/\alpha)\|_1 \text{ 最小化}$$

- 既存SNMFの最適化手法と特徴

射影勾配法

勾配降下後に非負制約へ射影

- ステップサイズが必要

$$\theta_i^{(t+1)} = \max \left(0, \theta_i^{(t)} - \eta_t \frac{\partial \mathcal{J}(\theta^{(t)})}{\partial \theta_i} \right)$$

Heuristic update rule [Févotte+, 2011]

勾配の正項と負項から乗法更新式を構成

$$\theta_i^{(t+1)} = \theta_i^{(t)} \cdot \frac{\left[\frac{\partial \mathcal{J}(\theta^{(t)})}{\partial \theta_i} \right]_-}{\left[\frac{\partial \mathcal{J}(\theta^{(t)})}{\partial \theta_i} \right]_+}, \quad \frac{\partial \mathcal{J}(\theta^{(t)})}{\partial \theta_i} = \left[\frac{\partial \mathcal{J}(\theta^{(t)})}{\partial \theta_i} \right]_+ - \left[\frac{\partial \mathcal{J}(\theta^{(t)})}{\partial \theta_i} \right]_-$$

補助関数法 / 上界最小化 (majorization-minimization: MM) アルゴリズム [Hunter+, 2000]

目的関数の上界となる補助関数の設計と最小化を反復

- 反復最小化における**目的関数値の単調非増加性を保証**

$\mathcal{J}(\theta) = \text{Minimize}_{\theta} \mathcal{J}^+(\theta, \tilde{\theta})$

$\mathcal{J}^+$: 補助関数 $\tilde{\theta}$: 補助変数

文献	最適化問題	最適化手法	単調非増加性
[Hoyer, 2002]	Minimize $\mathcal{D}_{\text{Eu}}(X WH)$ s.t. $w_{ik} \geq 0, h_{kj} \geq 0, s(h_k) = 1 \quad \forall i, j, k$	射影勾配法	無し
[Hoyer, 2004]	Minimize $\mathcal{D}_{\text{KL}}(X WH) + \mu \ H\ _1$ s.t. $w_{ik} \geq 0, h_{kj} \geq 0, \ w_k\ _1 = 1 \quad \forall i, j, k$	射影勾配法	無し
[Liu+, 2004]	Minimize $\mathcal{D}_{\text{KL}}(X WH) + \mu \ H\ _1$ s.t. $w_{ik} \geq 0, h_{kj} \geq 0, \ w_k\ _1 = 1 \quad \forall i, j, k$	補助関数法かつ 反復毎の正規化	無し
[Eggert+, 2004] [Le Roux+, 2015]	Minimize $\mathcal{D}_{\beta}(X WH) + \mu \ H\ _1$ s.t. $w_{ik} \geq 0, h_{kj} \geq 0 \quad \forall i, j, k$	Heuristic update rule	無し
[Leplat+, 2021]	Minimize $\mathcal{D}_{\beta}(X WH) + \mu \ H\ _1$ s.t. $w_{ik} \geq 0, h_{kj} \geq 0, \ w_k\ _2^2 = 1 \quad \forall i, j, k$	ラグランジュの未定乗数法 + ニュートン・ラプソン法	保証有り
[Marmin+, 2023]	Minimize $\mathcal{D}_{\beta}(X WH) + \mu \ WH\ _1$ s.t. $w_{ik} \geq 0, h_{kj} \geq 0 \quad \forall i, j, k$	補助関数法	保証有り

- 単調非増加性を保証した乗法更新式は限られる ([Marmin+, 2023] だけ)

本研究
の目的

W か H だけをスパース化

+

補助関数法に基づく乗法更新式

両立するSNMFの提案

位置付け

本研究はノルム制約を L_1 , $\beta = 1$ としたときの [Leplat+, 2021] と一致

単調非増加性を保証する乗法更新式を, [Filstroff+, 2021, Appendix C] の解法を流用し導出

ポスター完成直前
(9/4) に発見...orz

2. 提案手法

- ノルム制約付きKL基準SNMFの定式化

$$\text{Minimize}_{W, H} \underbrace{\mathcal{D}_{\text{KL}}(X|WH)}_{\text{一般化KL擬距離 (非凸)}} + \mu \sum_{k,j} h_{kj} \quad \text{s.t. } w_{ik}, h_{kj} \geq 0, \sum_i w_{ik} = 1 \quad \forall i, j, k$$

L_1 ノルム制約

- 目的関数: $\mathcal{J} = \mathcal{D}_{\text{KL}}(X|WH) + \mu \sum_{k,j} h_{kj}$
- W に対するスパース正則化も一般性を失うことなく議論可能

- 補助関数法に基づく乗法更新式の導出

- 補助変数を用いて非凸な損失関数の上界補助関数を設計

$$\mathcal{D}_{\text{KL}}(X|WH) \stackrel{c}{=} \sum_{i,j} \left(-x_{ij} \log \sum_k w_{ik} h_{kj} + \sum_k w_{ik} h_{kj} \right) \quad \text{Jensenの不等式}$$

$$\leq \sum_{i,j} \left(-x_{ij} \sum_k \delta_{ijk} \log \frac{w_{ik} h_{kj}}{\delta_{ijk}} + \sum_k w_{ik} h_{kj} \right) \quad \delta_{ijk}: \text{補助変数} \quad \sum_k \delta_{ijk} = 1$$

- 補助関数 $\mathcal{J}^+$ に**ノルム制約を引き継いだ**最小化問題 [Filstroff+, 2021, Appx. C]

$$\text{Minimize}_{W, H, \Delta} \mathcal{J}^+ \quad \text{s.t. } w_{ik}, h_{kj} \geq 0, \sum_i w_{ik} = 1 \quad \forall i, j, k$$

$$\mathcal{J}^+ \stackrel{c}{=} \sum_{i,j} \left(-x_{ij} \sum_k \delta_{ijk} \log \frac{w_{ik} h_{kj}}{\delta_{ijk}} + \sum_k w_{ik} h_{kj} \right) + \mu \sum_{k,j} h_{kj}$$

$\mathcal{J}^+ \geq \mathcal{J}$ の等式条件

$$\delta_{ijk} = \frac{w_{ik} h_{kj}}{\sum_{k'} w_{ik'} h_{k'j}}$$

- この最小化問題のラグランジアンとKKT条件

$$\mathcal{L} = \mathcal{J}^+ - \sum_k \lambda_k \left(\sum_i w_{ik} - 1 \right)$$

$$\text{原稿訂正: p.3 式(12)} \\ \mathcal{L} = \mathcal{J}^+ + \sum_k \lambda_k \left(\sum_i w_{ik} - 1 \right) \quad (12)$$

KKT条件

$$w_{ik} \frac{\partial \mathcal{L}}{\partial w_{ik}} = 0 \quad \forall i, k \quad (\text{i})$$

$$\frac{\partial \mathcal{L}}{\partial w_{ik}} \geq 0 \quad \forall i, k \quad (\text{ii})$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_k} = 0 \quad \forall k \quad (\text{iii})$$

- (i) より, $w_{ik} = 0$ または $\partial \mathcal{L} / \partial w_{ik} = 0$ を得る

$$\frac{\partial \mathcal{L}}{\partial w_{ik}} = \sum_j \left(-x_{ij} \frac{\delta_{ijk}}{w_{ik}} + h_{kj} \right) - \lambda_k = 0 \quad \text{変形} \quad w_{ik} = \frac{\sum_j x_{ij} \delta_{ijk}}{\sum_j h_{kj} - \lambda_k}$$

- (iii) 及びノルム制約より $\sum_i w_{ik} = 1$ を用いて λ_k を解く

$$\sum_i w_{ik} = \sum_i \frac{\sum_j x_{ij} \delta_{ijk}}{\sum_j h_{kj} - \lambda_k} = 1 \quad \text{分母は} i \text{ に依らない} \quad \lambda_k = \sum_j h_{kj} - \sum_{i,j} x_{ij} \delta_{ijk}$$

$$w_{ik} \text{ から } \lambda_k \text{ を消去} \quad w_{ik} = \frac{\sum_j x_{ij} \delta_{ijk}}{\sum_j h_{kj} - \lambda_k} = \frac{\sum_j x_{ij} \delta_{ijk}}{\sum_{i,j} x_{ij} \delta_{ijk}}$$

- 同様に, KKT条件を用いて $\partial \mathcal{L} / \partial h_{kj} = 0$ より h_{kj} について解く

$$\frac{\partial \mathcal{L}}{\partial h_{kj}} = \sum_j \left(-x_{ij} \frac{\delta_{ijk}}{h_{kj}} + w_{ik} \right) + \mu = 0 \quad \text{変形} \quad h_{kj} = \frac{\sum_j x_{ij} \delta_{ijk}}{w_{ik} + \mu}$$

- 反復更新式

- 等式条件を w_{ik} と h_{kj} のそれぞれに代入して反復更新式を得る

$$w_{ik} = w_{ik} \frac{\sum_j \frac{x_{ij} h_{kj}}{\sum_{k'} w_{ik'} h_{k'j}}}{\sum_i w_{ik} \sum_j \frac{x_{ij} h_{kj}}{\sum_{k'} w_{ik'} h_{k'j}}}, \quad h_{kj} = h_{kj} \frac{\sum_i \frac{x_{ij} w_{ik}}{\sum_{k'} w_{ik'} h_{k'j}}}{\sum_i w_{ik} + \mu}$$

- 収束パラメータや正規化処理を必要としない**乗法更新式**

- 補助関数法とKKT条件を厳密に満たし, **理論的に単調非増加性を保証**

補足

$$\mathbf{1}_I \in \{1\}^I$$

β ダイバージェンスの $\beta \leq 1$ に拡張したときの反復更新式 [Marmin+, 2023]

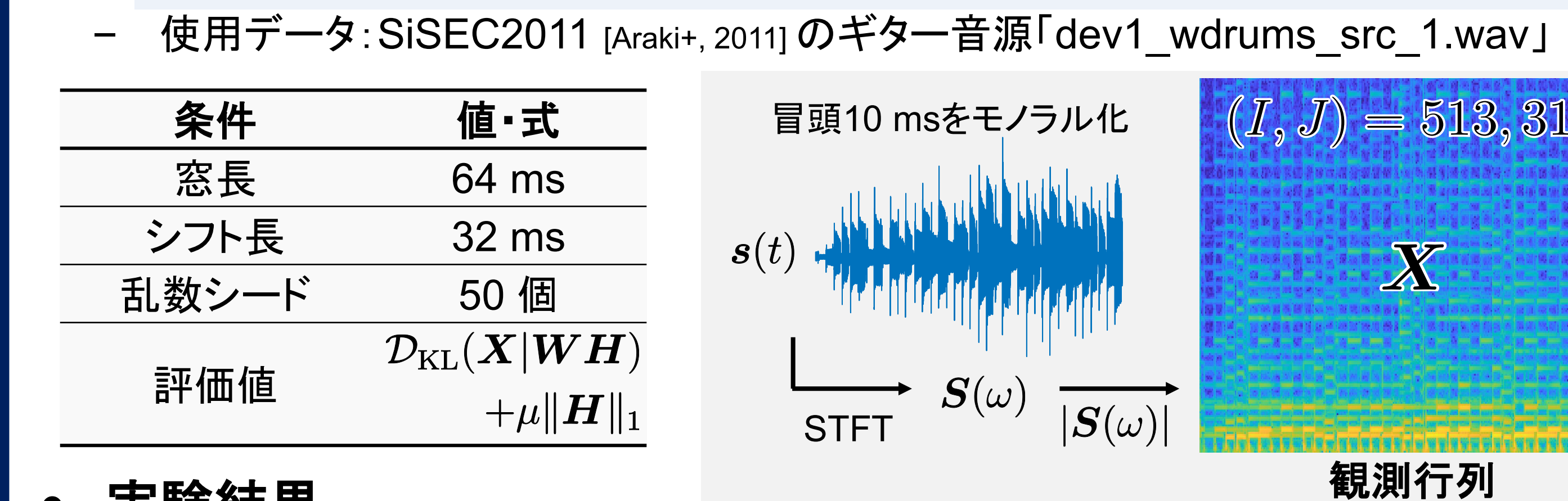
$$W = W \odot \left(\frac{X \odot (WH)^{(\beta-2)}}{(WH)^{(\beta-1)} H^T - \mathbf{1}_I \lambda^T} \right)^{\frac{1}{2-\beta}} \quad W \text{ がラグランジュ乗数 } \lambda \text{ に依存する} \\ (\lambda \text{ を解析的に導けない})$$

3. 実験

- 音源データを用いた振幅スペクトログラムの近似実験

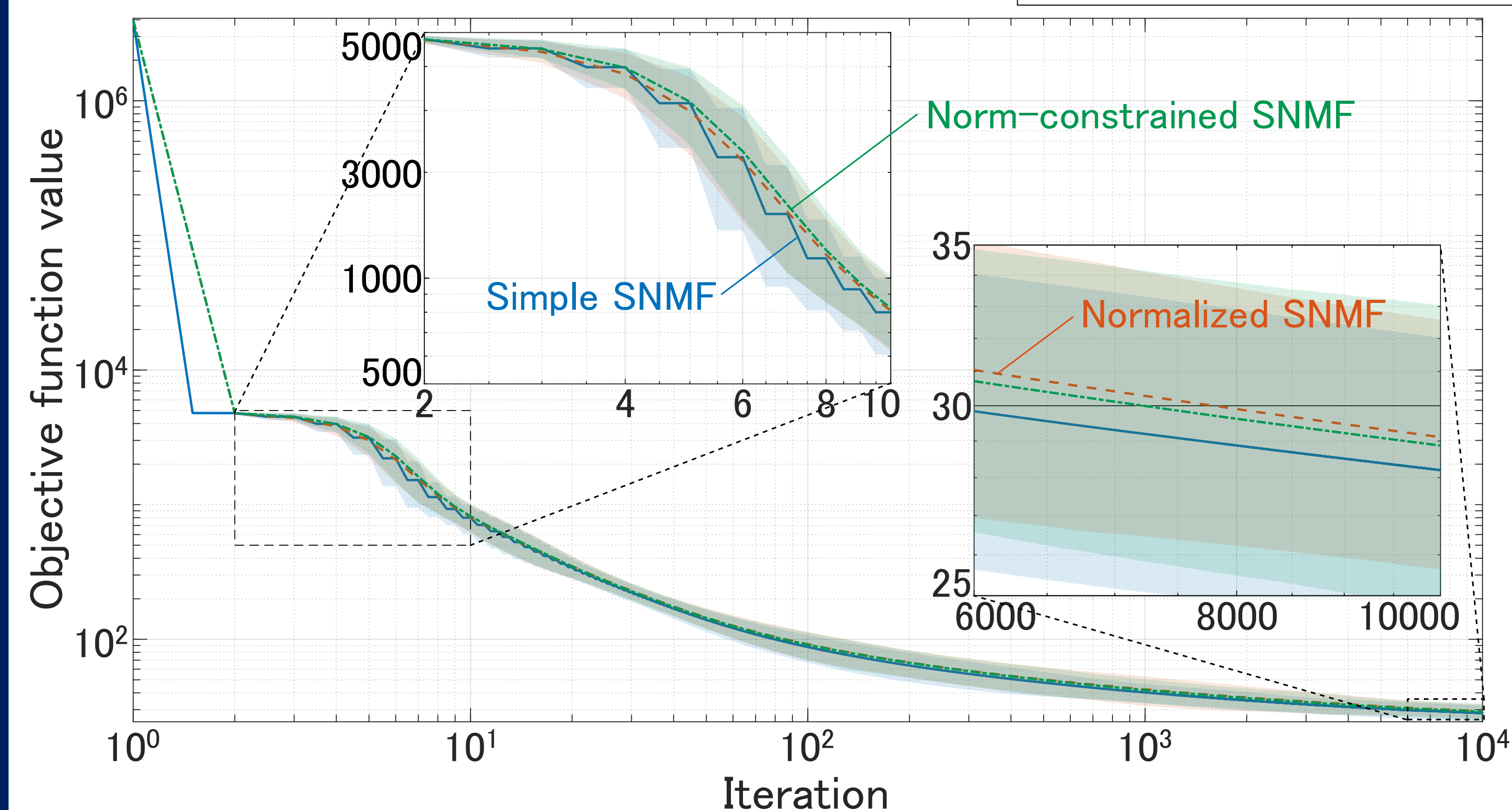
- 目的: 振幅スペクトログラムの近似実験により**単調非増加性を確認**
- 評価方法: 目的関数が一致する3手法の収束挙動を比較

- Simple SNMF [Liu+, 2004]: 更新毎に正規化するため前後の目的関数値計算
- Normalized SNMF [Le Roux+, 2015]: $\beta = 1$ における反復更新式を利用
- Norm-constrained SNMF: 提案手法

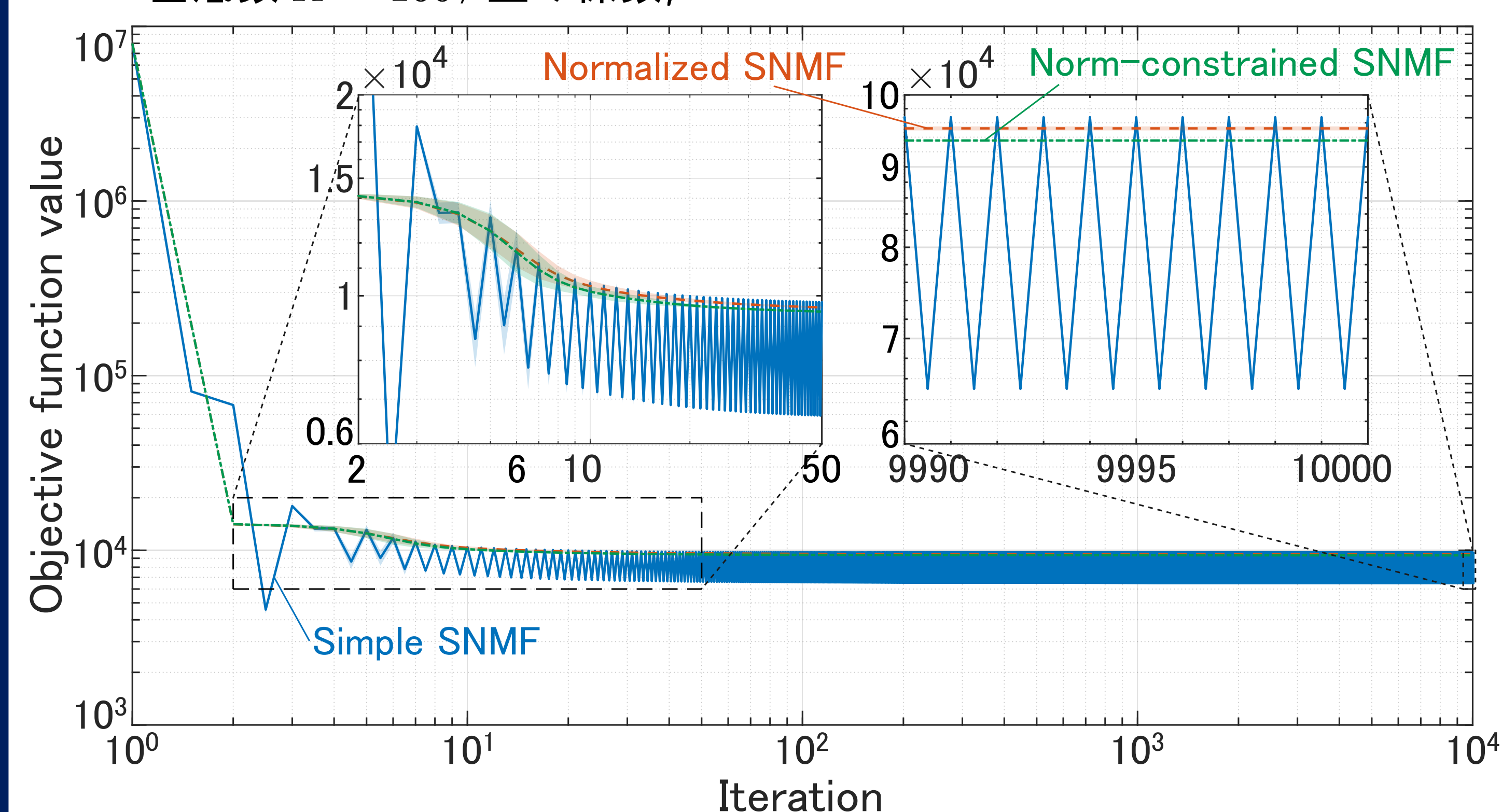


- 実験結果

- 基底数 $K = 100$, 重み係数 $\mu = 10^{-6}$



- 基底数 $K = 100$, 重み係数 $\mu = 1$



実験結果

基底数 K 及び重み係数 μ の様々な値に対して...

- 収束挙動が安定しており, **常に単調非増加性を維持**する
- 従来手法と**同等かより優れた目的関数値に収束**する