1-P-7

ディリクレ非負値行列因子分解のハウリング抑圧への応用

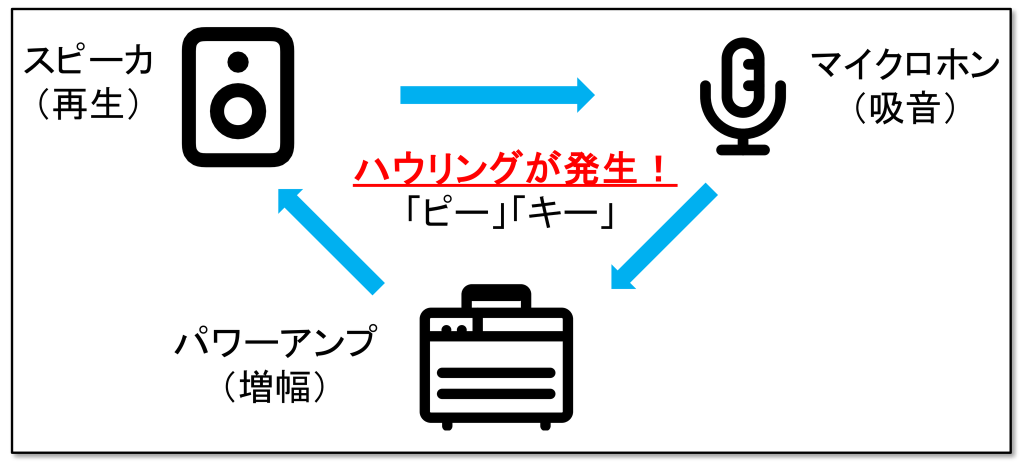
◎小川遼，北村大地（香川高専），綾野翔馬（科学大）

本発表の概要

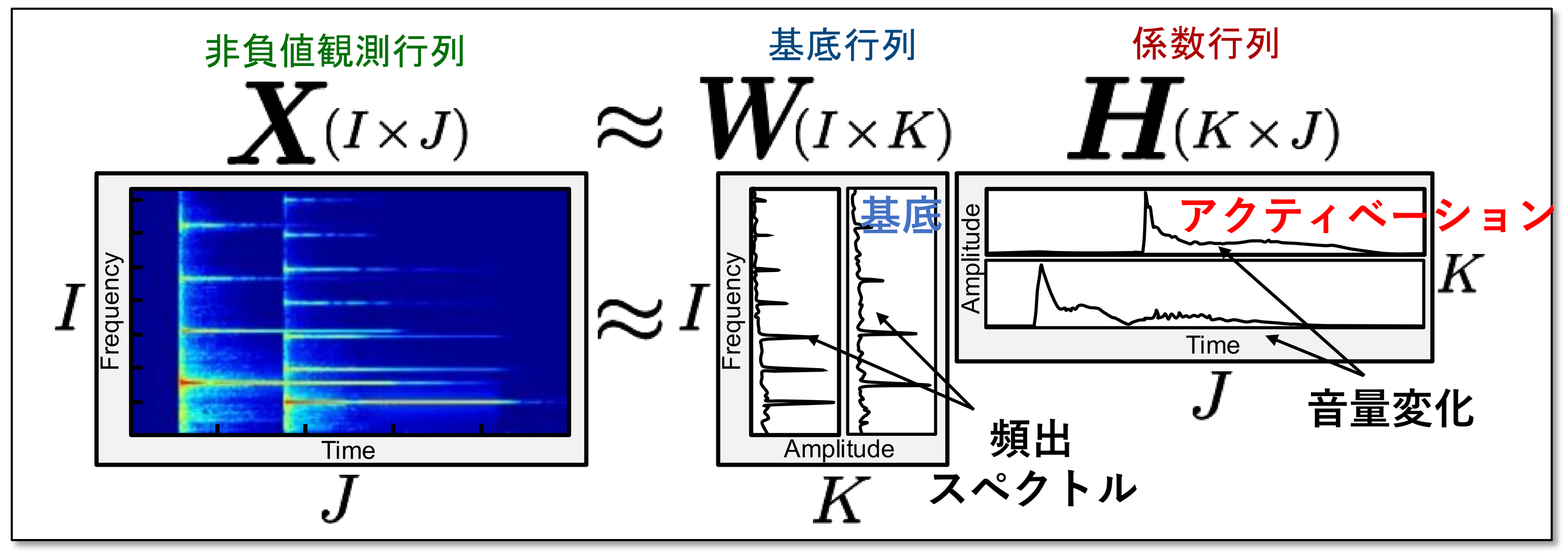
- ハウリングはマイク・アンプ・スピーカによるフィードバックにより生じる発振現象
  - 音質劣化の原因となる重要な問題
- スペクトログラム分解によりハウリング成分のみを抑圧する新しいアプローチを提案
  - ➡ ディリクレ事前分布を導入した非負値行列因子分解（ディリクレNMF）
- さらにディリクレ分布の母数を負の領域まで拡張した拡張ディリクレNMFを提案
- 音声信号および音楽信号を用いた実験により従来のスパースNMF（L_1-NMF）より高いハウリング抑圧性能を確認

1. 研究背景

- **ハウリング（howling）**
  - 音が循環し特定の周波数が増幅される現象
  - 発生すると，大きな不快音が持続し，音質が劣化する
  - 通常は，適応ノッチフィルタなどで抑圧することが一般的



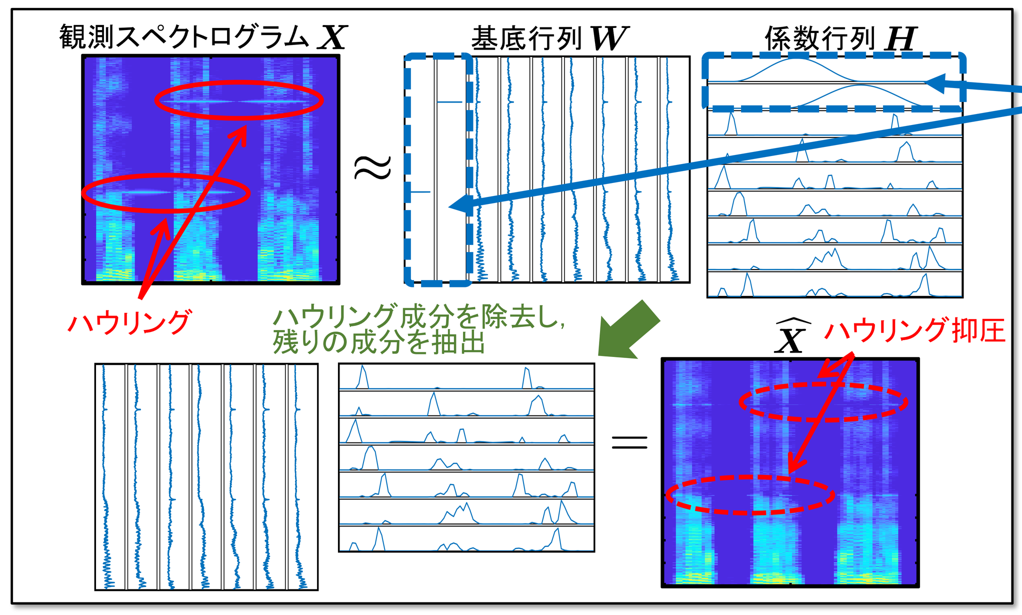
- **非負値行列因子分解（nonnegative matrix factorization: NMF）**[Lee+, 1999]
  - 観測非負行列を別の2つの非負行列の行列積で低ランク近似する手法
  - 音響信号では時間周波数のモデリングに頻繁に用いられる



新規性 適応ノッチフィルタ以外の新しいハウリング抑圧手法を検討
ディリクレNMFでハウリングの基底を抽出・除去する手法を提案

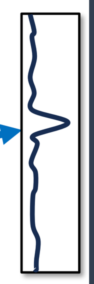
2. 提案手法

- **NMFを用いたハウリング抑圧**



問題点① 乱数により度の基底にハウリング成分が出るか不明

問題点② ハウリング成分だけの基底を作れない場合もある



特定の基底ベクトルがone-hotになるように誘導（提案手法）

- ハウリング成分を特定のベクトルに誘導可能
- ハウリング基底に他の成分が混ざることを防ぐ

- **ディリクレ分布**
  - 非負制約とノルム制約を満たす確率変数ベクトル（標準単体上のベクトル）の分布
  - 標準単体の定義

$$\mathcal{S}_{I-1} = \left\{ \mathbf{x} \in \mathbb{R}^I \mid \underbrace{\mathrm{x}_i \geq 0\ \forall i}_{\text{非負制約}},\ \underbrace{\sum_i \mathrm{x}_i = 1}_{\text{ノルム制約}} \right\}$$

  - ディリクレ分布の確率密度関数

$$\mathbf{x} \sim p(\mathbf{x};\boldsymbol{\alpha}) = \frac{1}{B(\boldsymbol{\alpha})}\prod_i \mathrm{x}_i^{\alpha_i - 1}$$

確率変数ベクトル　パラメータベクトル　多変量ベータ関数

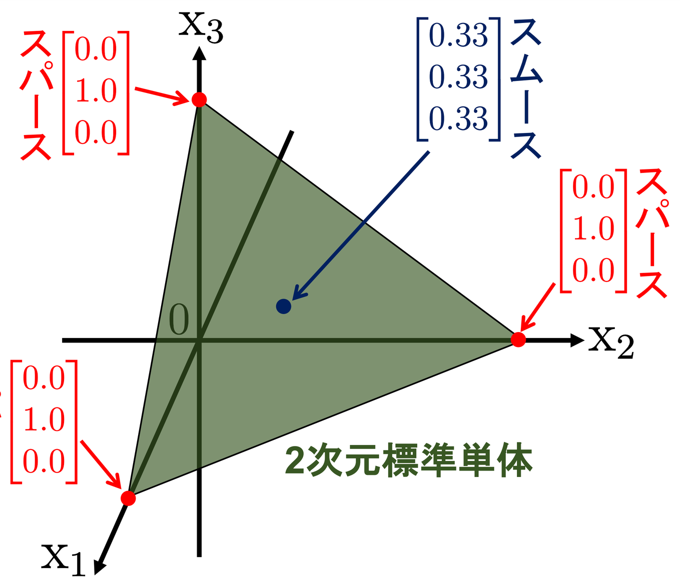
$$B(\boldsymbol{\alpha}) = \frac{\prod_i \Gamma(\alpha_i)}{\Gamma\left(\sum_i \alpha_i\right)}$$

- **3次元確率変数ベクトルの場合**
  - 2次元標準単体上の確率変数ベクトル（3変数だがノルム制約より自由度は2）

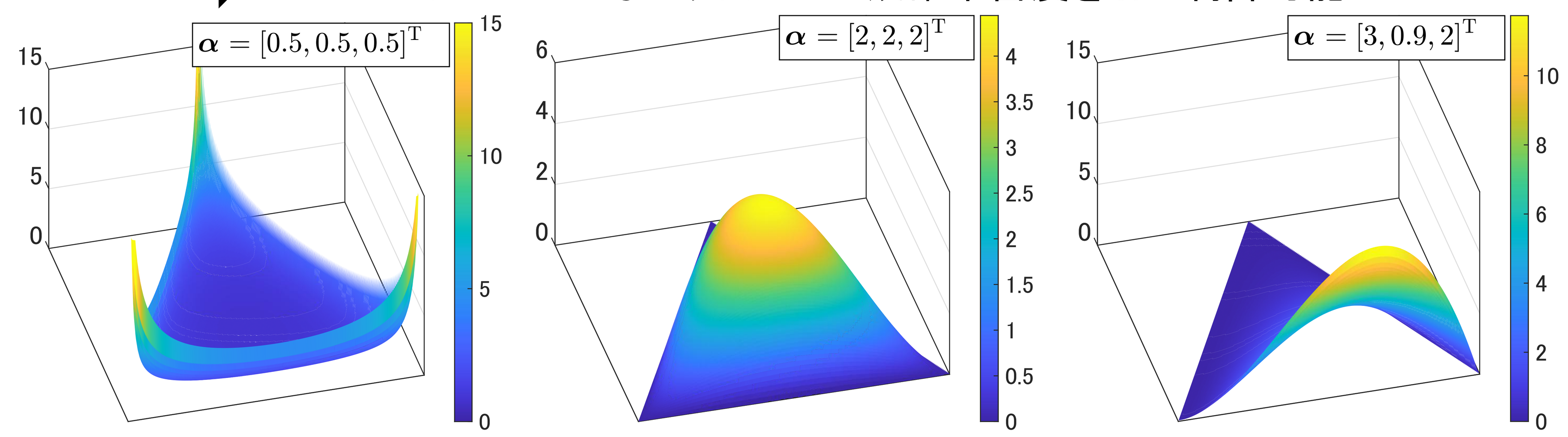
3次元確率変数ベクトル
$$\mathbf{x} = \begin{bmatrix}\mathrm{x}_1\\ \mathrm{x}_2\\ \mathrm{x}_3\end{bmatrix}$$

ノルム制約: $\mathrm{x}_1 + \mathrm{x}_2 + \mathrm{x}_3 = 1$
非負制約: $\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3 \geq 0$

例えば・・・ $\mathbf{x} = \begin{bmatrix}0.2\\0.3\\0.5\end{bmatrix}, \mathbf{x} = \begin{bmatrix}0.1\\0.1\\0.8\end{bmatrix}, \mathbf{x} = \begin{bmatrix}0.0\\0.9\\0.1\end{bmatrix}$ 等



  - パラメータベクトル $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \alpha_3]^{\mathrm{T}}$ は確率密度の各頂点への集中度に対応
  - ➡ スパースやスムースなベクトルの生成確率密度を α で制御可能



- **ディリクレNMFの目的関数と最適化の方針**
  - ディリクレ分布から基底行列の各列（基底ベクトル）が独立に生成されると仮定

$$\underset{\boldsymbol{W},\boldsymbol{H}}{\text{Minimize}}\ \mathcal{D}(\boldsymbol{X}|\boldsymbol{W}\boldsymbol{H}) + \sum_{i,k}(\alpha_{ik}-1)\log w_{ik}^{-1} \ \text{s.t.}\ w_{ik}, h_{kj} \geq 0\ \forall i,j,k,\ \sum_i w_{ik} = 1\ \forall k$$

  - NMFでは常套手段の**補助関数法**を用いるが，ディリクレ分布由来の**ノルム制約**を陽に扱って解く必要がある ➡ **補助関数法の上限関数最小化に制約条件を付与**

- **補助関数の非負制約・ノルム制約条件付き最適化問題**

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

  - 補助関数の最小化問題に対して**非負制約**と**ノルム制約**の両方を考慮
  - **一般化KL擬距離に対する補助関数**と**ディリクレ分布由来の正則化項（そのまま）**

$$\mathcal{J}^+(\boldsymbol{W},\boldsymbol{H},\boldsymbol{\Delta}) = \sum_{i,j}\left[x_{i,j}\log x_{i,j} - x_{i,j}\sum_k \delta_{i,j,k}\log\frac{w_{i,k}h_{k,j}}{\delta_{i,j,k}} - x_{i,j} + \sum_k w_{i,k}h_{k,j}\right] + \sum_{i,k}(\alpha_{i,k}-1)\log w_{i,k}^{-1}$$

- **ディリクレNMFの反復更新式**

$$w_{ik} = \begin{cases}\hat{w}_{ik} & (\text{if}\ \ \hat{w}_{ik} \geq 0)\\ 0 & (\text{otherwise})\end{cases}$$

$$\hat{w}_{ik} = \frac{w_{ik}\sum_j \frac{x_{ij}}{\sum_{k'} w_{ik'}h_{k'j}}h_{kj} + \alpha_{ik} - 1}{\sum_{i\in\mathcal{I}_k}\left(w_{ik}\sum_j \frac{x_{ij}}{\sum_{k'} w_{ik'}h_{k'j}}h_{kj} + \alpha_{ik} - 1\right)}$$

補助関数の最適化問題に対するラグランジアン $\mathcal{L}$ を考えたときに，$\partial\mathcal{L}/\partial w_{ik} = 0$ を満たす w_{ik} が非負となるインデクス i の集合（k 毎にこの集合を定義している）

- **拡張ディリクレNMF**

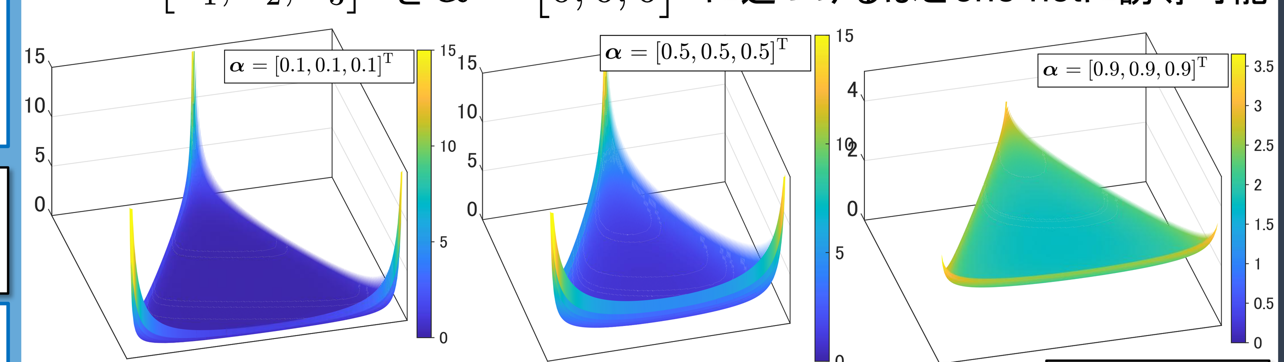
ディリクレ分布の母数の制約（$\alpha_{ik} > 0$）によりone-hot性への誘導が頭打ちに．．．

$$\sum_{i,k}(\alpha_{ik}-1)\log w_{ik}^{-1}$$

$\alpha_{ik} = 0.1, 0.01, \cdots$ として も係数部分が-0.9，-0.99・・・となり正則化の効きがほぼ変わらない

α_{ik} に負の値を許すことでより強力なone-hot性への誘導を実現（拡張ディリクレNMF）

$\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \alpha_3]^{\mathrm{T}}$ を $\boldsymbol{\alpha} = [0, 0, 0]^{\mathrm{T}}$ に近づけるほどone-hotに誘導可能



拡張ディリクレNMFの問題点

1. 母数 $\alpha_{ik} \leq 0$ ではディリクレ分布にならない（単純な正則化付きNMFとなりMAP推定ではなくなる）
2. 基底ベクトルがゼロベクトルとなりNMFが**破綻する**可能性がある

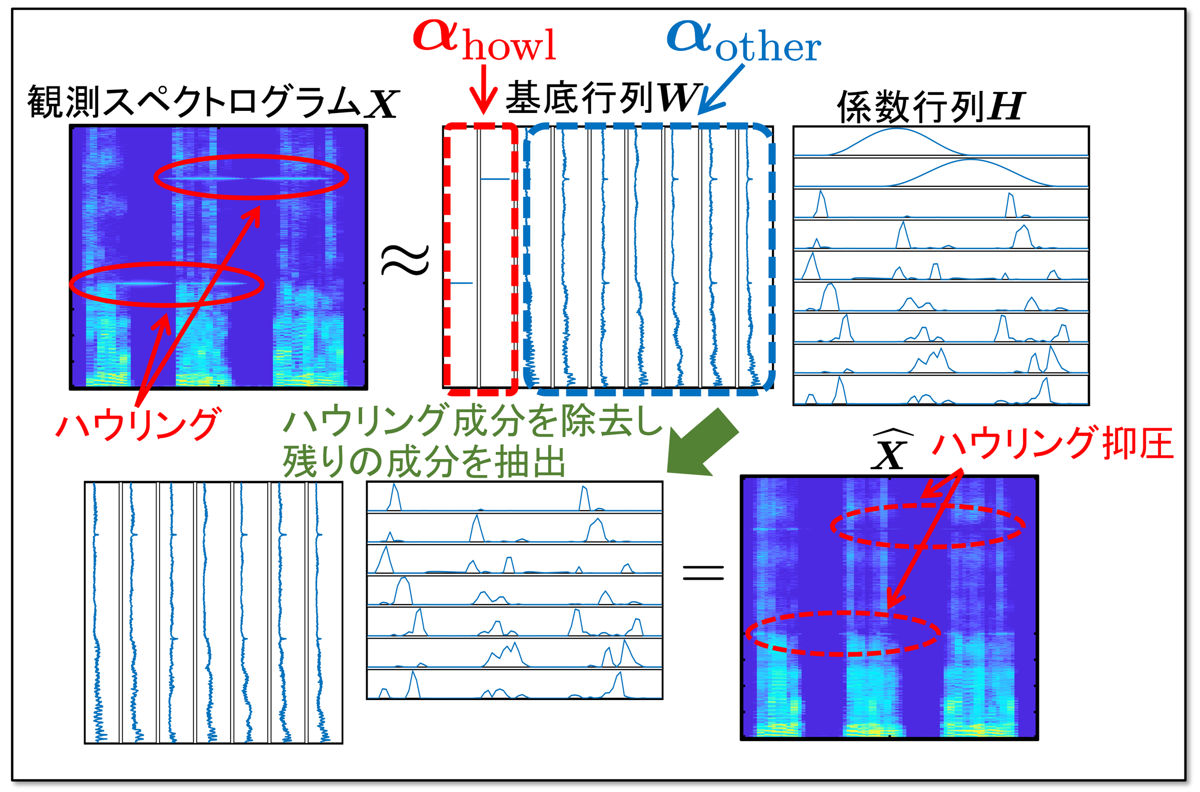
いずれかの基底が**ゼロベクトル**に

$$w_k = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

3. ハウリング抑圧実験

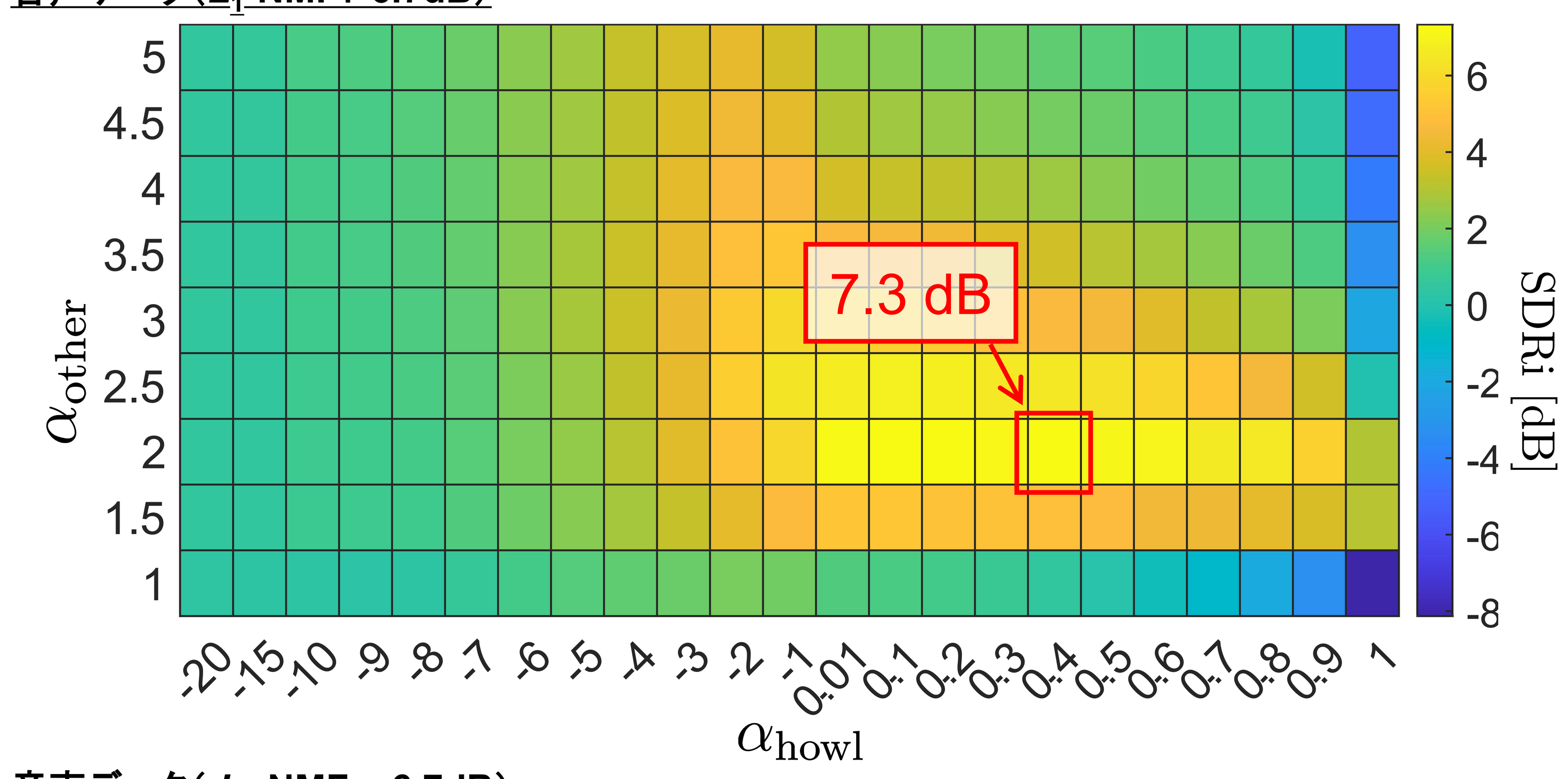
- **実験目的**
  - ディリクレNMFおよび拡張ディリクレNMFを用いたハウリング抑圧
- **実験条件**
  - 人工的にハウリングを2本付与した音声信号及び音楽信号を使用
  - ハウリングを誘導するパラメータ α_{howl}，その他の成分を誘導する α_{other} を設定

条件	詳細
データセット（音声）	JVS corpus
データセット（音楽）	DSD100
信号長（音声）	10 s（先頭区間）
信号長（音楽）	30–40 s
サンプリング周波数	16 kHz
窓関数	ハン窓
窓長（音声）	64 ms
窓長（音楽）	128 ms
試行回数	100回
評価指標	SDRi

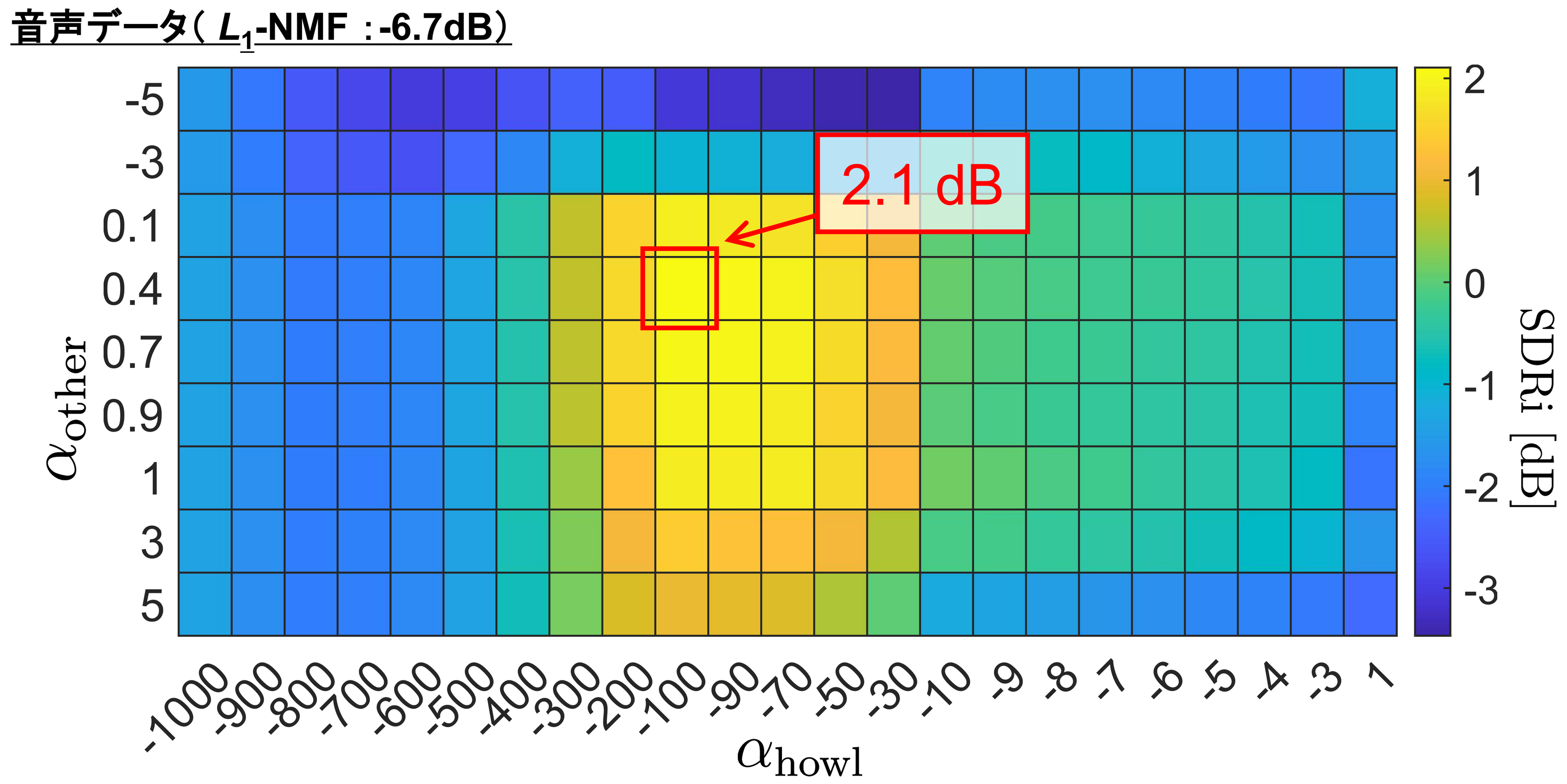


- **実験結果**

音声データ（L_1-NMF：-6.7dB）



音声データ（ L_1-NMF ：-6.7dB）



破綻率