

観測信号の白色化による 線形ブラインド残響除去の高精度化

◎岡本 拓磨, 岩谷 幸雄, 鈴木 陽一

東北大学 電気通信研究所 / 大学院情報科学研究科

研究背景：音空間を構成する属性の分離・抽出

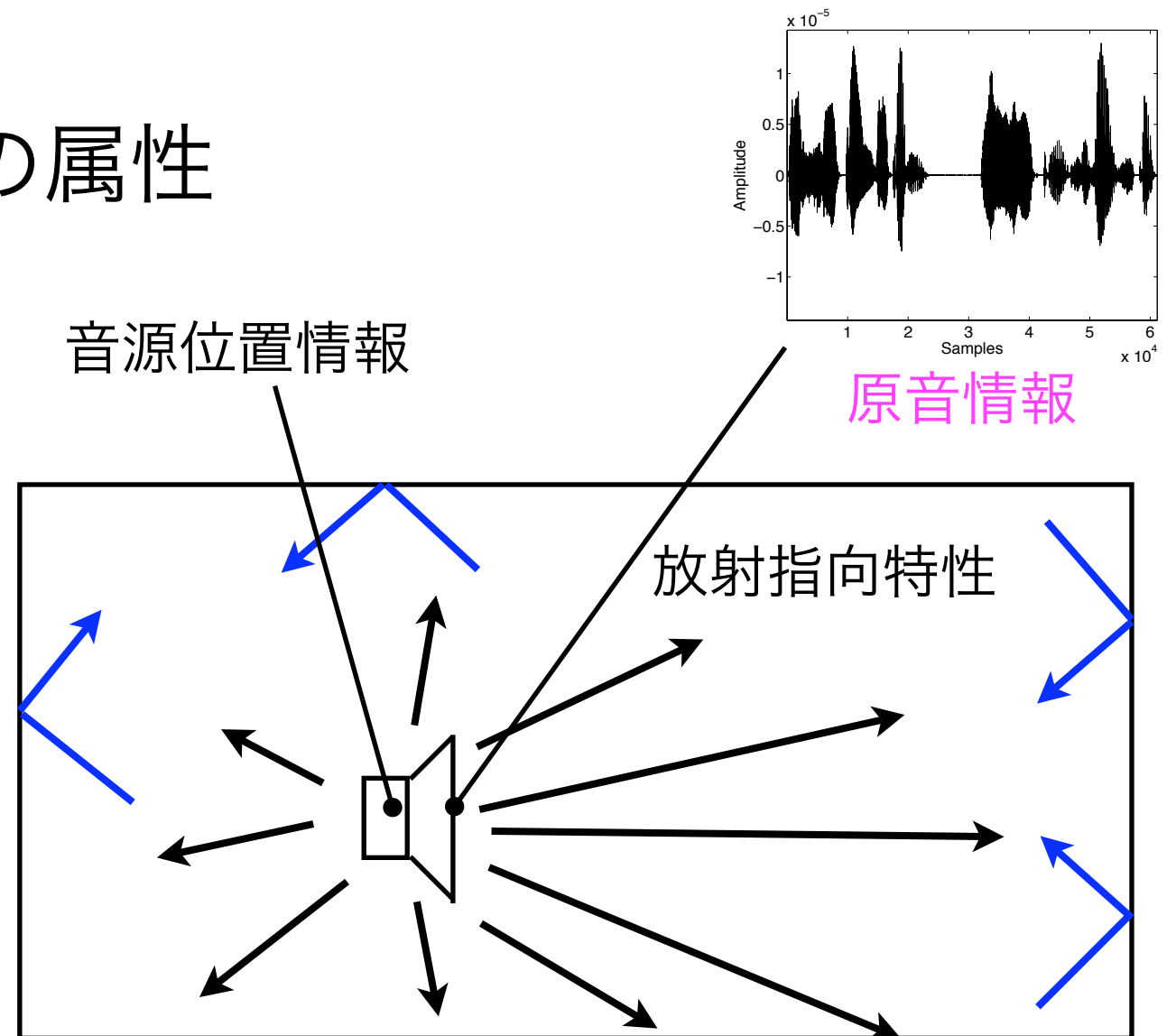
■ 室内音空間を構成する5つの属性

■ 音源の属性

1. 音源位置情報
2. 原音情報
3. 放射指向特性
…包囲による伝達特性の違い

■ 空間 (=室内)の属性

4. 反射音情報
5. 後部残響特性



➡ 各音源ずつ存在, 極めて複雑な室内音場

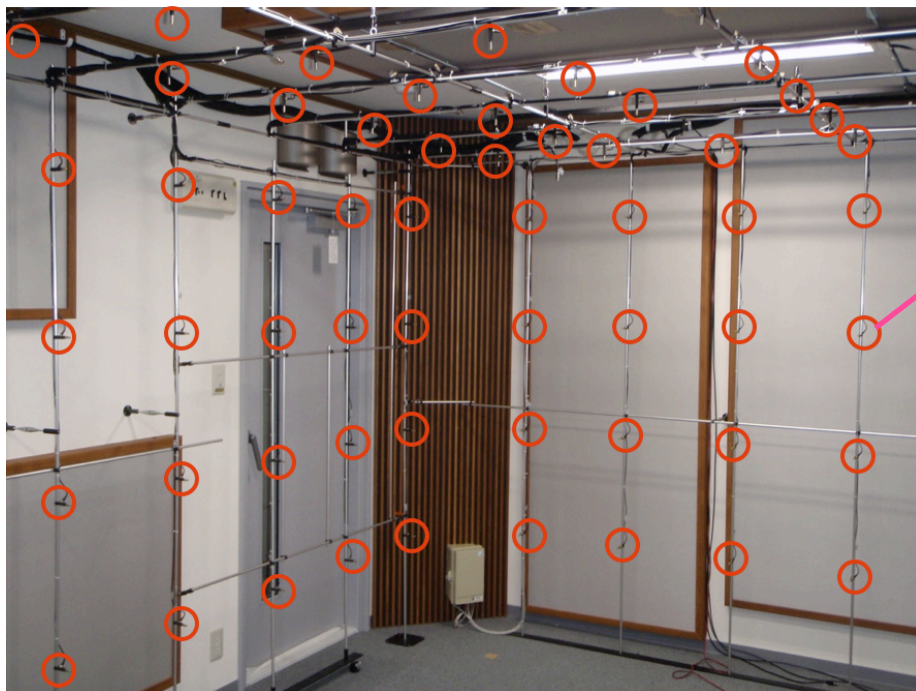
➡ 属性ごとの分離・抽出・解析 → 高臨場感再生などへの利用

室内残響下における原音情報の抽出に関する検討

包囲型マイクロホンアレイの構築

■ 包囲型マイクロホンアレイ

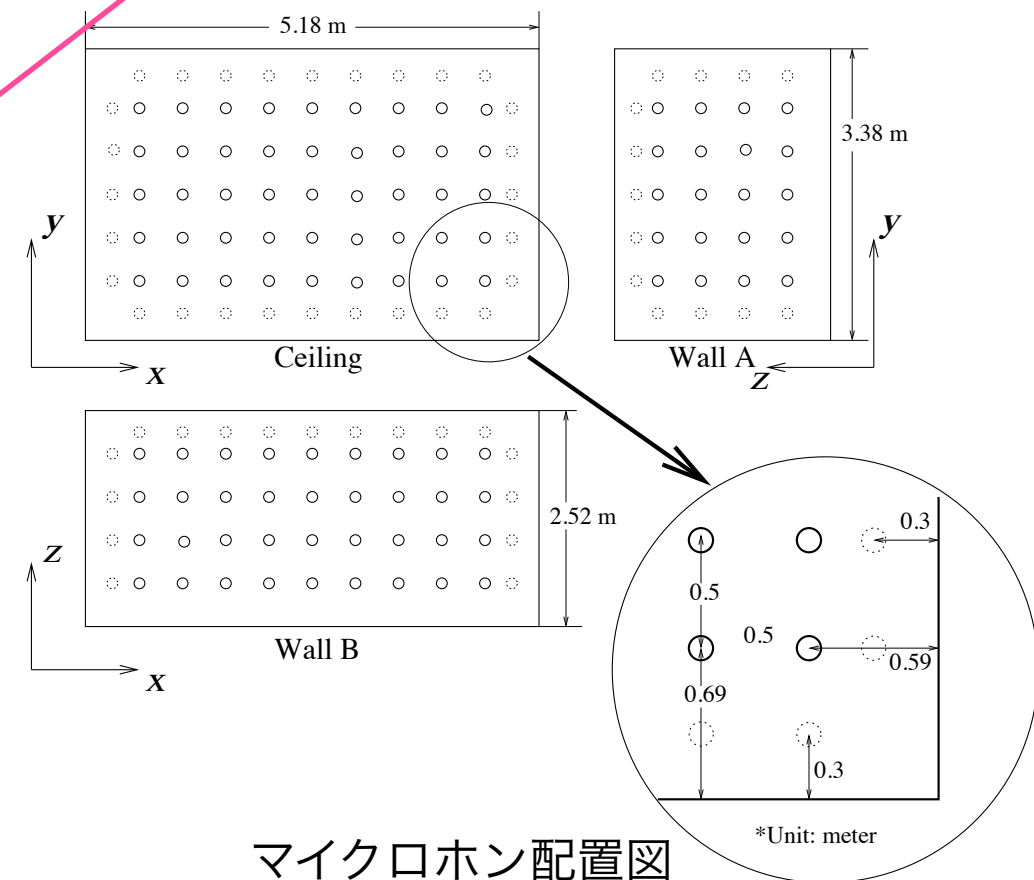
- 天井と壁4面にマイクロホンを格子状に配置
- 壁面から30 cm, 50 cm間隔にて格子状に配置
 - ※ マイクロホン：B & K Type 4951 × 157 ch
 - ※ マイクロホンアンプ：B & K Type 2694 × 10



包囲型マイクロホンアレイ@東北大学電気通信研究所



B & K Type 4951



マイクロホン配置図

- 48 kHzサンプリング157 chの同時収録を実現
- 音源特性を余すところなく收音可能

室内残響下における原音情報の高精度抽出

■ ブラインドシステム同定 → MINT型逆フィルタによる 原音抽出

MINT (Multiple input-output INverse theorem) M. Miyoshi *et al.* 1988

■ 観測信号のみから先験的知識なしで系のインパルス応答を推定

- ✱ インパルス応答長の知識が必要

- ✱ 室内残響は数千タップと非常に長い

→ 音響分野では困難な課題

■ ブラインド残響除去

■ 観測信号のみから系のMINT型逆フィルタを直接推定

- ✱ インパルス応答長が不明 → それより長い逆フィルタにより解決

- ✱ 音源が白色雑音を仮定 → 音声のような有色信号では出力結果が白色化

■ 線形予測を用いたブラインド残響除去

Linear-Predictive Multi-input Equalization (LIME) M. Delcroix *et al.* 2005

- ✱ 白色化問題を解決 → 良好な残響除去

目的：LIMEアルゴリズムの高精度化

■ 従来のブラインド残響除去のサンプリング周波数

- LIMEアルゴリズムをはじめとして、後段に音声認識を想定

- ✱ サンプリング周波数を8 kHzや16 kHzとしている

■ 本研究で想定しているブラインド残響除去

- 楽器や物音などの様々な音響事象の解析を想定

- ✱ 16 kHzよりもさらなる広帯域化が必要

LIMEアルゴリズムを単純に高帯域化



音声などの有色信号では精度が劣化

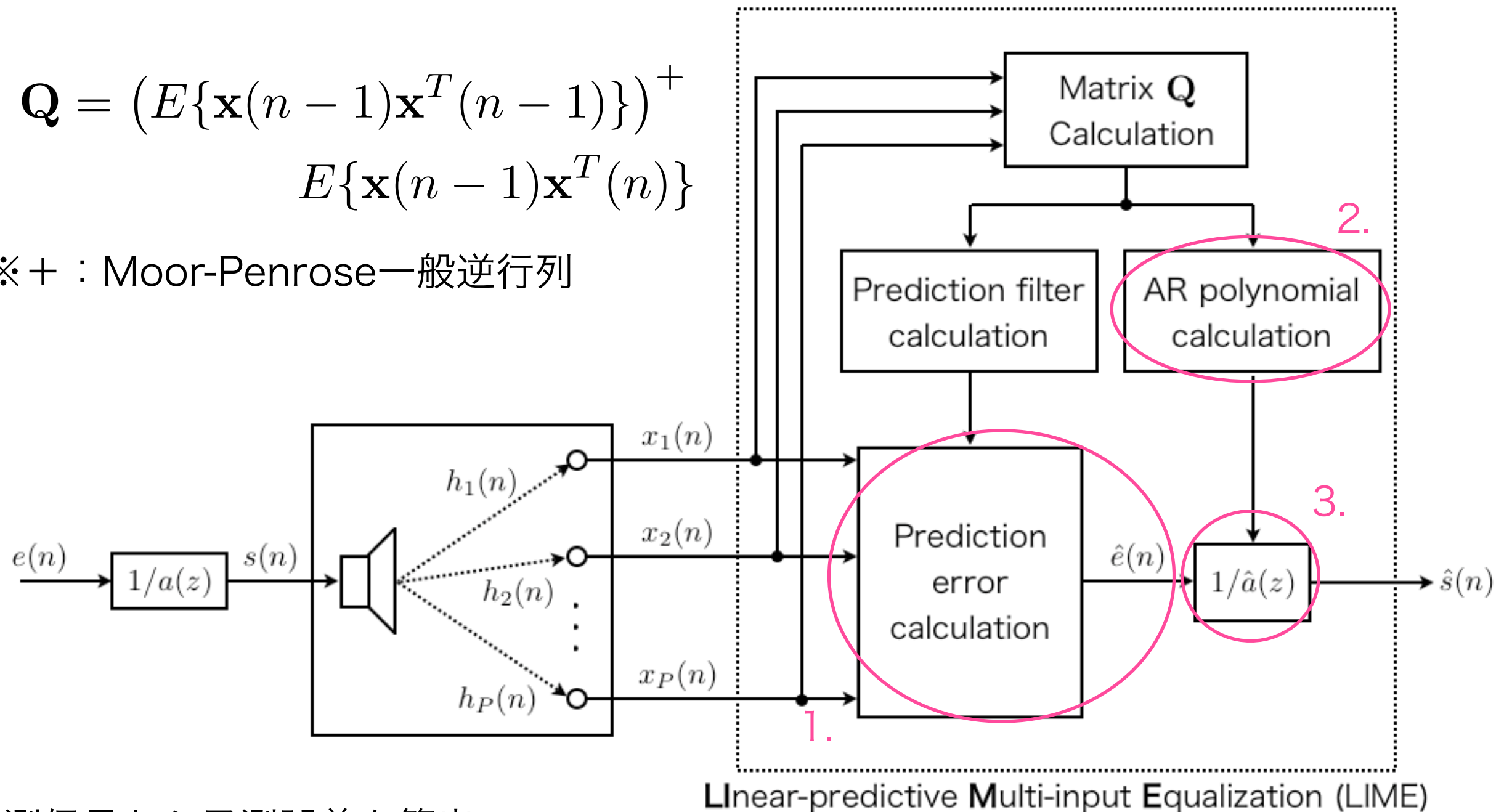


LIMEアルゴリズムの広帯域化を検討

LIMEアルゴリズム

$$\mathbf{Q} = \begin{pmatrix} E\{\mathbf{x}(n-1)\mathbf{x}^T(n-1)\} \\ E\{\mathbf{x}(n-1)\mathbf{x}^T(n)\} \end{pmatrix}^+$$

※+ : Moor-Penrose一般逆行列

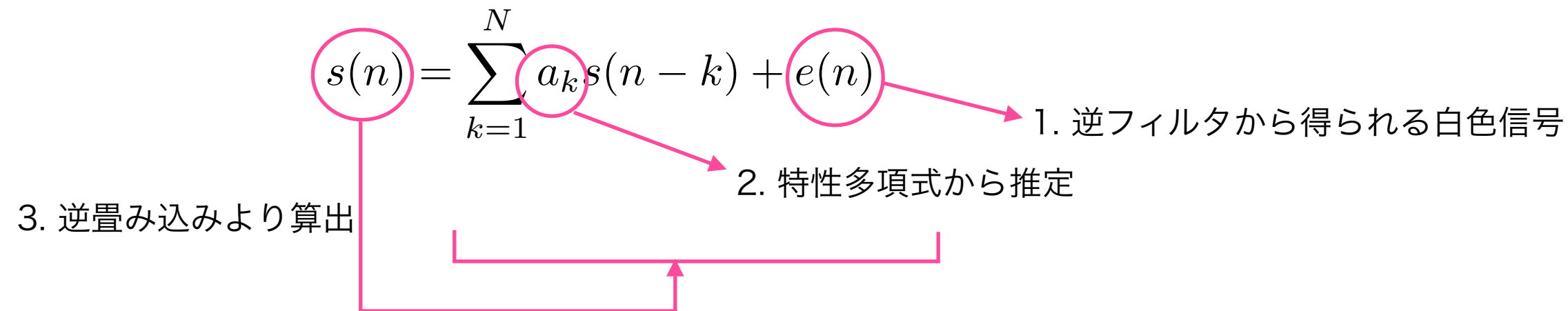


1. 観測信号から予測誤差を算出
2. 原音信号の平均スペクトル(=AR係数 $a(z)$)を特性多項式の係数として推定
3. 推定したAR係数を逆畳み込み → 逆フィルタによって白色化された信号 $\hat{e}(n)$ を回復

LIMEアルゴリズム (2)

■ LIMEアルゴリズムの特徴

- 入力信号の特性多項式から原信号のAR係数 $a(z)$ を別途推定

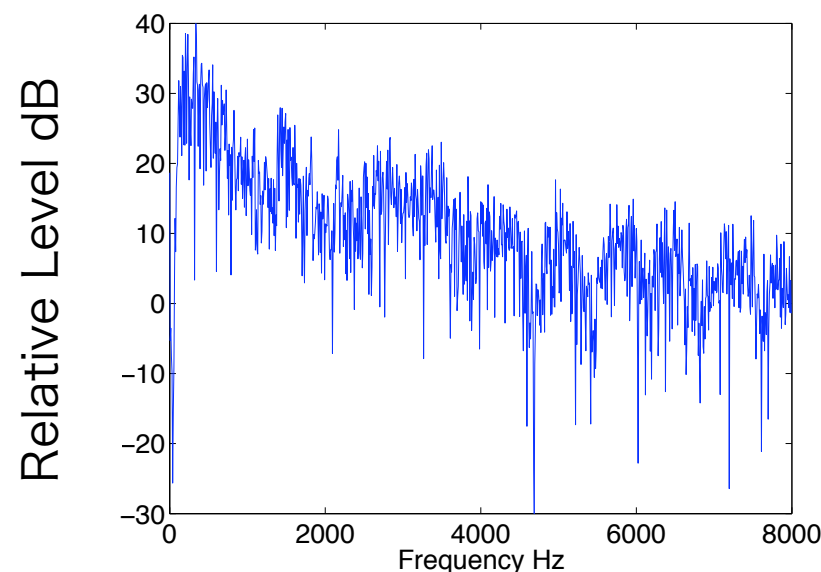


- ✳ 逆フィルタ長 L が室内インパルスより長く見積もっておけば, AR係数 $a(z)$ を求めた際, 長く見積もった分のAR係数は最後すべて0として計算されることによってインパルス応答長 M が未知である問題を解決
- ✳ AR係数 $a(z)$ の次数 $N = M + L - 1$

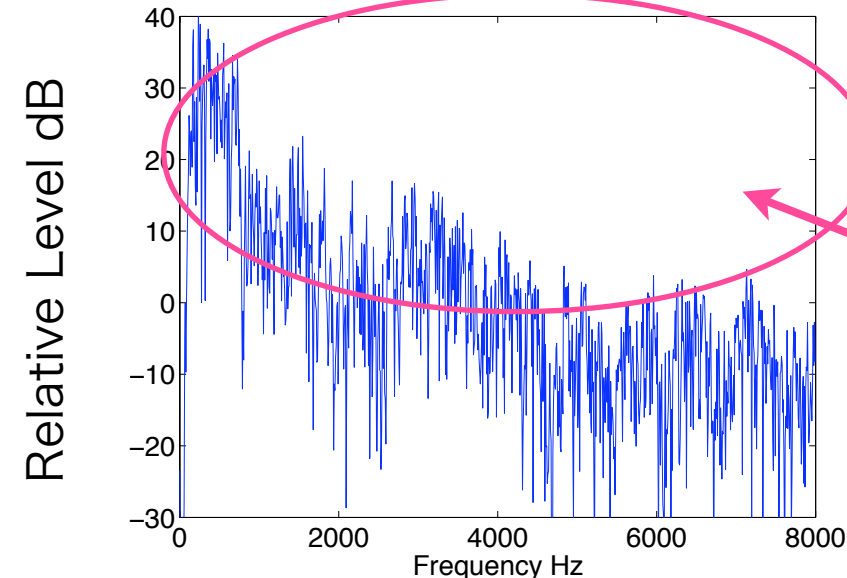
LIMEの高サンプリング時における不安定性

■ 観測信号の周波数領域におけるスパース性

- 実環境での信号：時間軸方向，周波数軸方向に対してパワーがまばらに存在



音声の振幅スペクトル



観測信号の振幅スペクトル

スパース性

■ LIMEアルゴリズムの高サンプリング時における不安定性

- $\mathbf{x}(n-1)\mathbf{x}^T(n-1)$ 観測信号の相関行列
 - サンプル周波数が高い場合，観測信号のスパース性が顕著となり，ランク数がAR係数 $a(z)$ の次数 N よりも低下

特性多項式の解が不安定 → AR係数の算出が不能

Pre-Whiteningフィルタを用いた LIMEアルゴリズムの高精度化

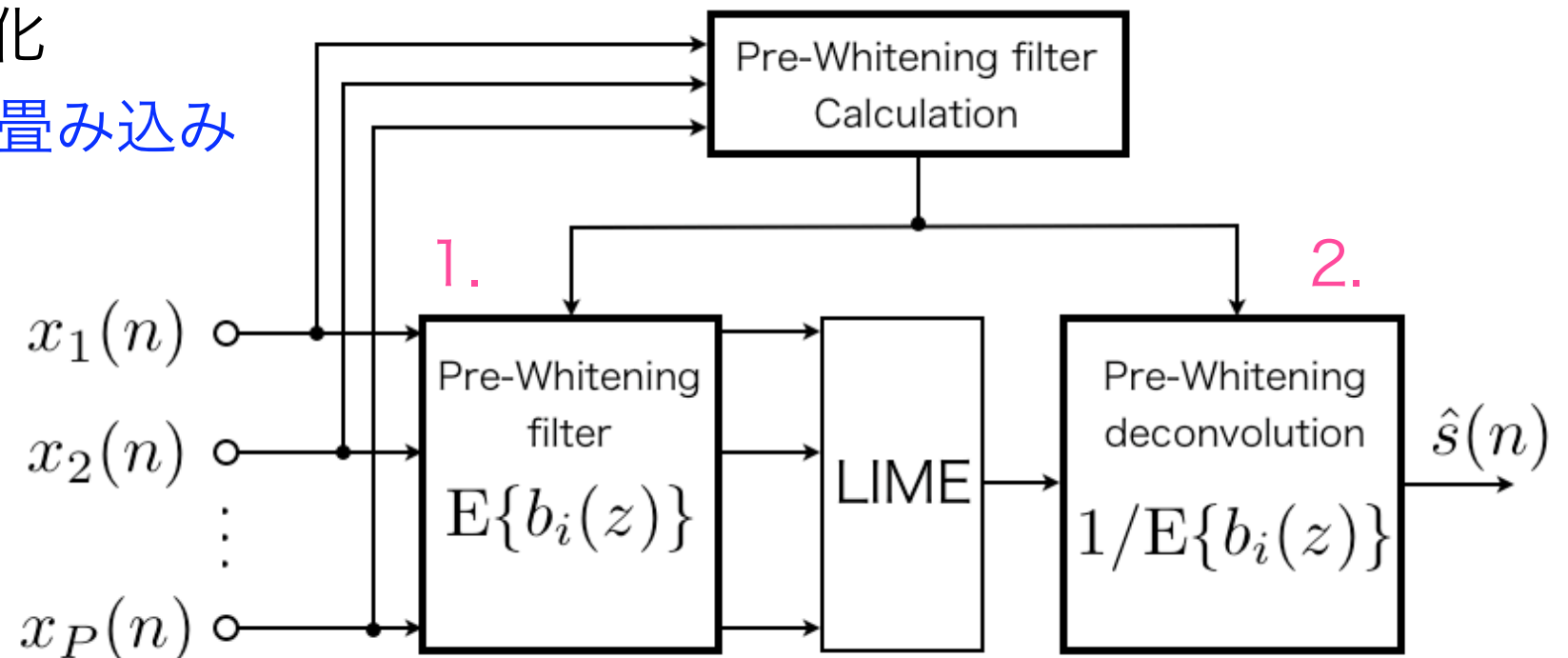
■ White-LIMEの提案

1. 全チャンネルに共通の白色化フィルタ

➡ 観測信号のスパース性を緩和 → 安定した $(E\{\mathbf{x}(n-1)\mathbf{x}^T(n-1)\})^+$ を算出

2. LIMEからの信号は白色化

➡ 白色化フィルタで逆畳み込み



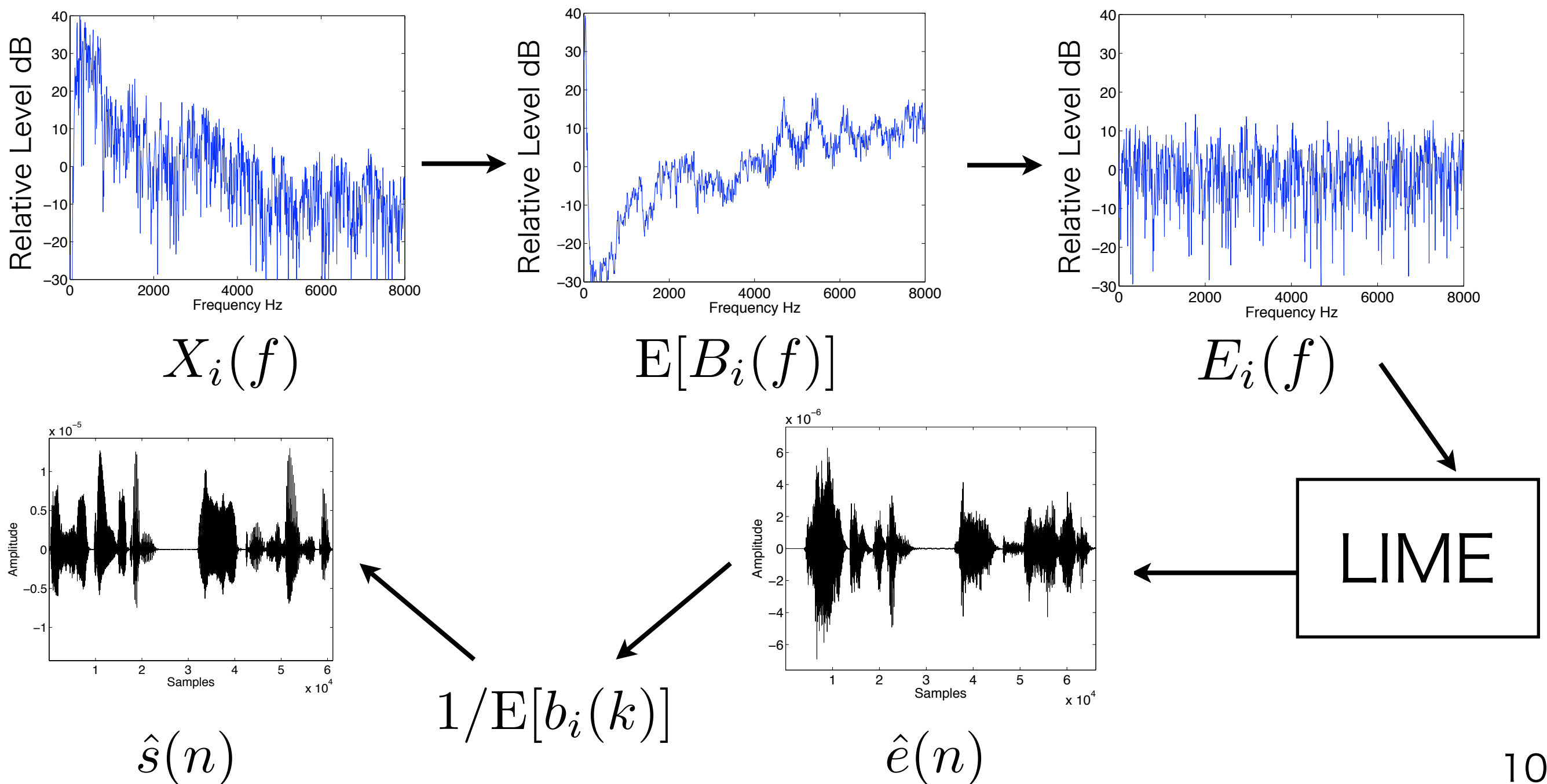
■ 使用するPre-Whiteningフィルタ

■ 各チャンネルをから求めたAR係数を平均化 N. D. Gaubitch *et. al* 2003

$$E\{b_i(z)\} = 1 - \{E\{b_{1i}\}z^{-1} + \dots + E\{b_{Ni}\}z^{-N}\} \quad x_i(n) = \sum_{k=1}^N b_{ki}x_i(n-k) + e_i(n)$$

White-LIMEの流れ

■ 高サンプリングの例 (24 kHz, Pre-Whiteningフィルタ長 1024 tap)

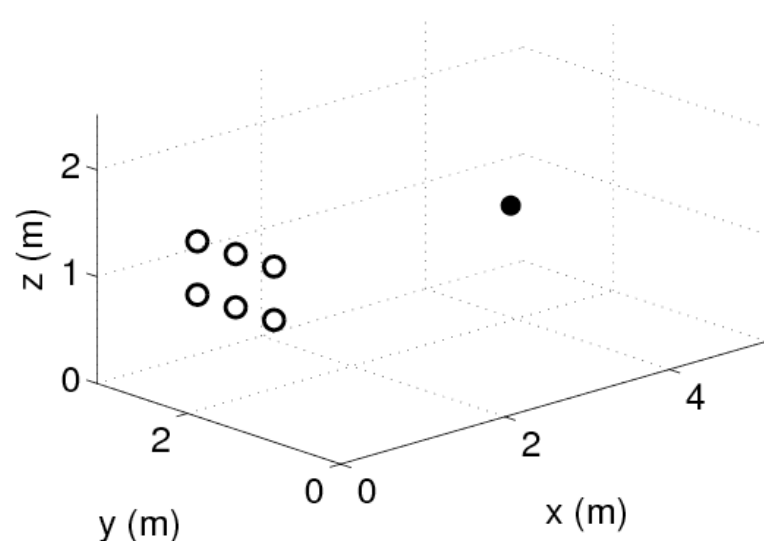


White-LIMEの有効性の検討

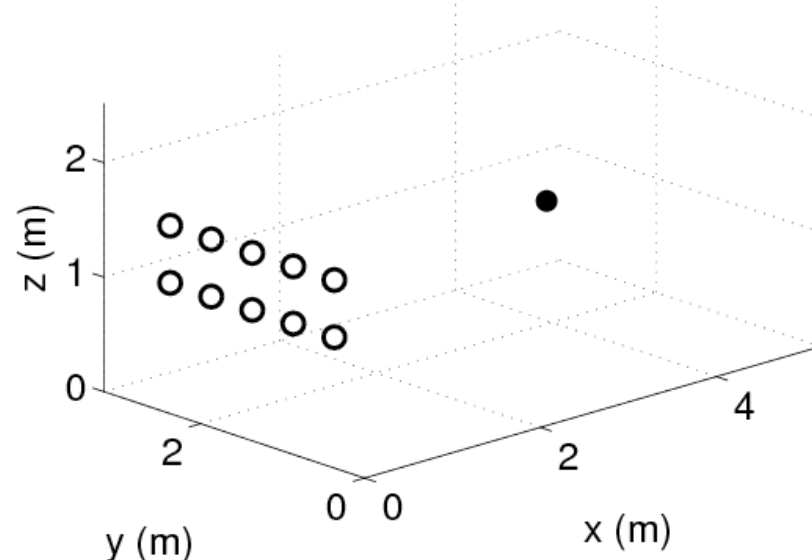
■ シミュレーションによる性能評価

■ シミュレーション条件

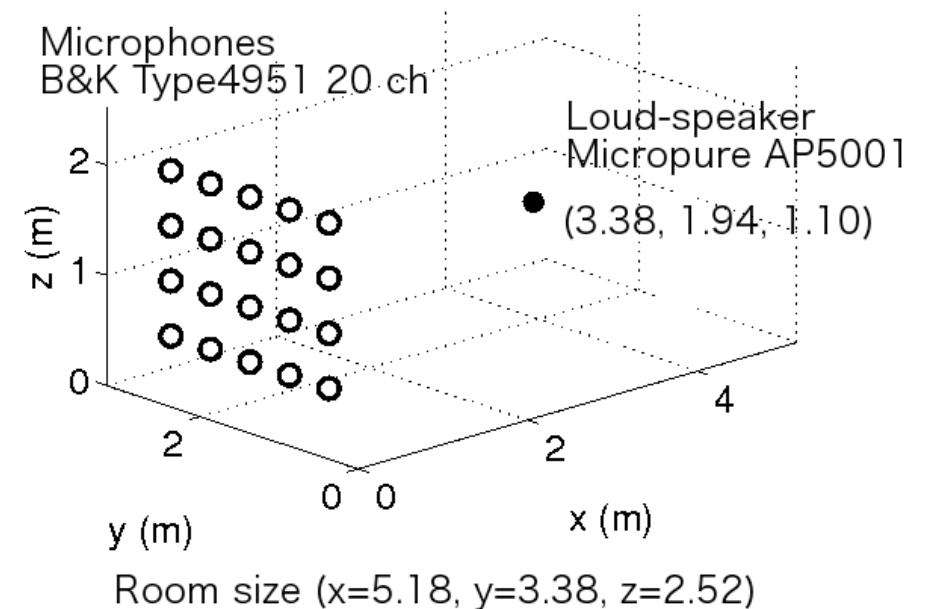
- ※ サンプル周波数：24k Hz → インパルス応答長：3600タップ
- ※ 残響時間：0.15 s
- ※ 原音：男性音声 2.6 s (NTT-AT広帯域音声データベース)
- ※ 白色化フィルタ長：256タップ
- ※ 観測信号：原音とインパルス応答の畳み込みにより作成 (=背景雑音なし)
- ※ チャンネル数：6, 10, 20 ch



6 ch



10 ch



20 ch

White-LIMEの性能評価

■ Pre-Whiteningフィルタによる残響除去精度の向上

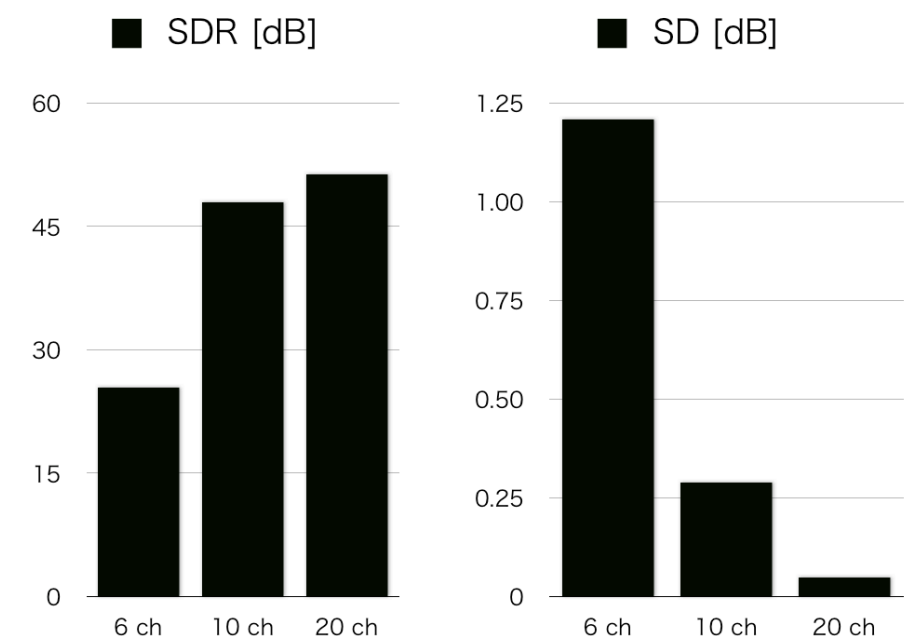
- 評価指標1 : Signal to Distortion Ratio (SDR) …値が大きいほど原音に近い

$$\text{SDR} = 10\log_{10} \left(\frac{\sum |s(n)|^2}{\sum |s(n) - \hat{s}(n)|^2} \right) \quad [\text{dB}]$$

- 評価指標2 : Spectral Distortion (SD) …値が小さいほど原音に近い

$$\text{SD} = \sqrt{\frac{1}{F} \sum_{f=0}^{F-1} \left(20\log|P(f)| - 20\log|\hat{P}(f)| \right)^2}. \quad [\text{dB}]$$

	order of $a(z)$	rank	SDR [dB]	SD [dB]
received signal $x_1(n)$			4.7	8.0
LIME (6 ch)	4328	4265	×	×
LIME (10 ch)	4008	3938	×	×
LIME (20 ch)	3798	3763	×	×
White-LIME (6 ch)	4328	4328	25.5	1.21
White-LIME (10 ch)	4008	4008	48.0	0.29
White-LIME (20 ch)	3798	3798	51.4	0.05



まとめと今後の課題

■ まとめ

- スパース性のある観測信号にPre-Whiteningフィルタを導入し、相関行列の条件数を緩和するWhite-LIMEを提案
 - ✱ スパース性のある観測信号に対しても高精度な残響除去が可能
 - ✱ 従来よりも広範囲な音信号に対する残響除去が可能

■ 今後の課題

- 背景雑音にロバストな高精度残響除去アルゴリズムの検討