DIRECTION PRESERVING WIND NOISE REDUCTION OF B-FORMAT SIGNALS

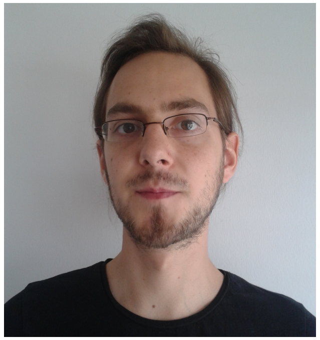
ICASSP 2021

Adrian Herzog, Daniele Mirabilii, and Emanuël A.P. Habets

FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Fraunhofer IIS

DIRECTION PRESERVING WIND NOISE REDUCTION OF B-FORMAT SIGNALS



A. Herzog

D. Mirabilii

Outline

- Introduction
- Problem Formulation
- Noise Reduction Methods
- Proposed Method
- Performance Evaluation

Outline

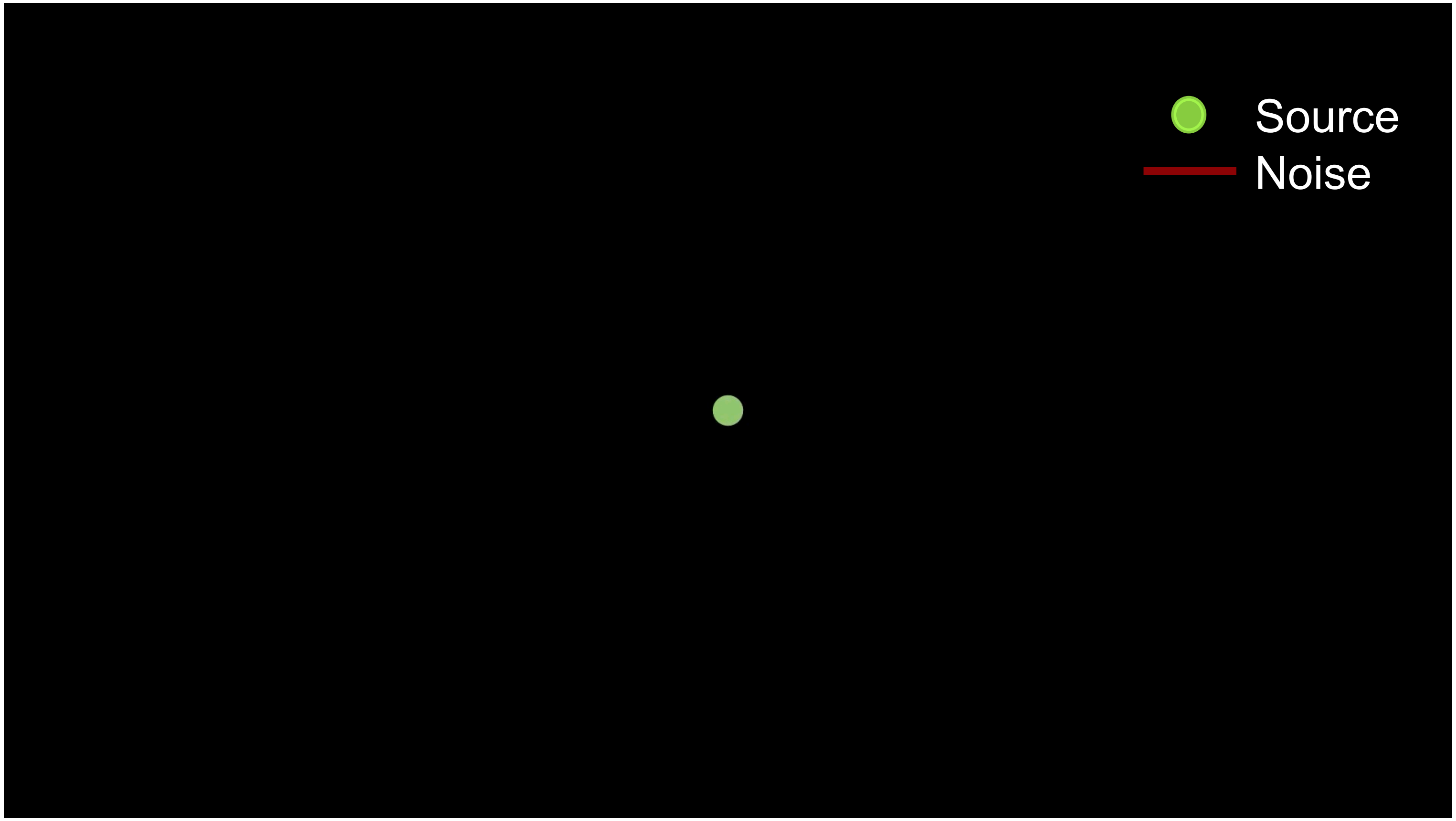
- **Introduction**
- Problem Formulation
- Noise Reduction Methods
- Proposed Method
- Performance Evaluation

Introduction
Motivation

- Wind noise is an adverse acoustic condition in outdoor recordings
- Degradation of sound quality and speech intelligibility
- Denoising in B-format signals
  - Suppress noise
  - Preserve the original spatial properties of the sound field

Introduction
Motivation

Binauralized sample: headphones recommended!



Introduction
Main idea

- Multi-channel wind noise reduction [Mirabilii and Habets, 2019]
  - Use of a Parametric Multi-channel Wiener Filter (PMWF)
  - Power Ratio (PR): noise reduction/speech distortion trade-off

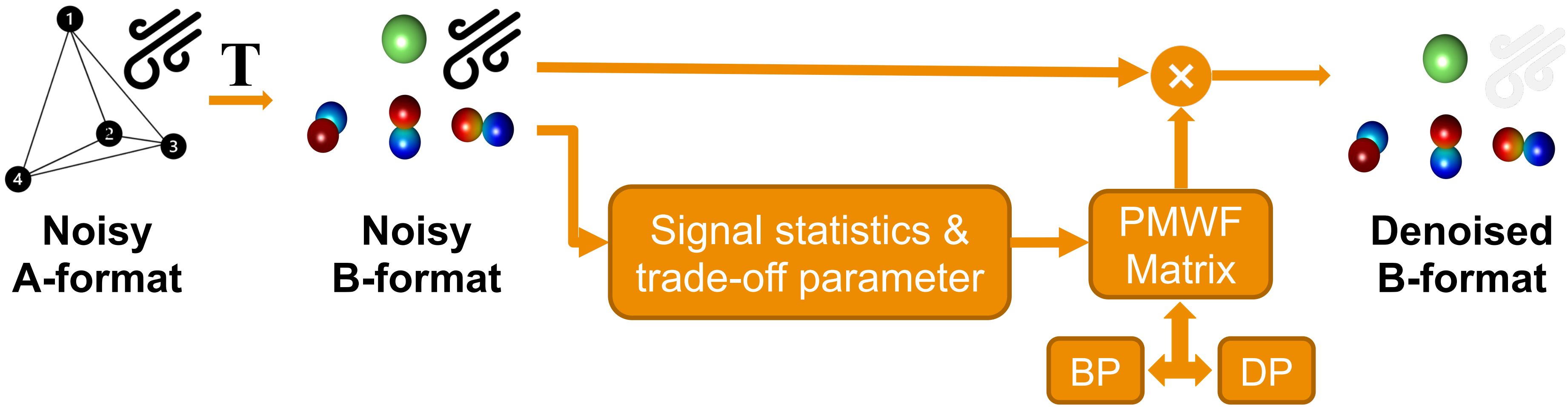
- Direction-preserving noise reduction for Ambisonics [Herzog and Habets 2019]
  - PMWF matrix
    - Beamform-and-project (BP)
    - Direction-preserving (DP)

Mirabilii and Habets, “Multi-channel wind noise reduction using the Corcos model,” ICASSP, UK, 2019.

Herzog and Habets, “Direction-preserving Wiener matrix filtering for Ambisonic input-output systems,” ICASSP, UK, 2019.

Introduction
Main idea

Direction-preserving wind noise reduction of B-format signals

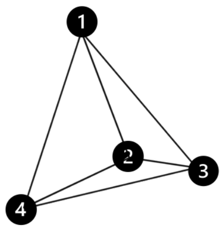


Outline

- Introduction
- **Problem Formulation**
- Noise Reduction Methods
- Proposed Method
- Performance Evaluation

Problem Formulation
A-format and B-format Signals

A-format

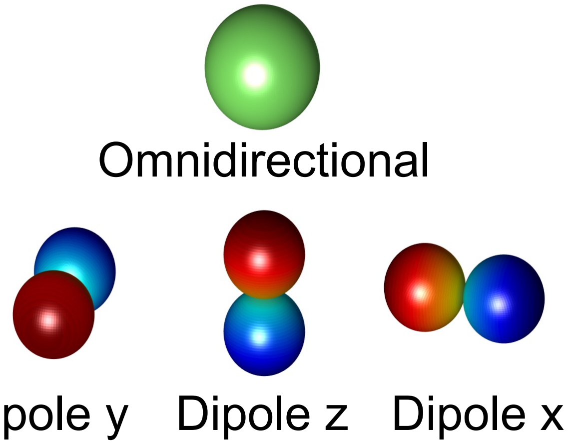


Soundfield microphone

- 4 Cardioid microphones
- Tetrahedron arrangement

$\mathbf{T}$

B-format



$$\mathbf{x}_A = \begin{bmatrix} X_{A1} \\ X_{A2} \\ X_{A3} \\ X_{A4} \end{bmatrix}$$

$$\mathbf{x}_B = \mathbf{T}\mathbf{x}_A = \begin{bmatrix} X_{Bo} \\ X_{Bx} \\ X_{By} \\ X_{Bz} \end{bmatrix}$$

Problem Formulation

Signal Model

$$\mathbf{x}_B = \underbrace{S\,\mathbf{b}(\Omega)}_{\text{Desired}} + \underbrace{\mathbf{v}_B}_{\text{Undesired}}$$

- Speech source S
- B-format steering vector $\mathbf{b}(\Omega)$ (plane wave)
- Direction Ω
- Wind noise $\mathbf{v}_B$

Tasks

- Estimate desired component
- Reduce undesired component
- Preserve the spatial information of the original soundfield

Outline

- Introduction
- Problem Formulation
- **Noise Reduction Methods**
- Proposed Method
- Performance Evaluation

Noise Reduction Methods

Noise Reduction Using a Filter Matrix

Beamforming

$$Z = \mathbf{w}^H \mathbf{x}_B$$

- Single-channel output Z
- Beamformer $\mathbf{w}$
- Z is an estimate of S

Multichannel-to-single-channel signal enhancement

No spatial information after beamforming

Matrix Spatial Filtering

$$\mathbf{z}_B = \mathbf{W} \mathbf{x}_B$$

- Multichannel output $\mathbf{z}_B$
- Filter matrix $\mathbf{W}$
- $\mathbf{z}_B$ is an estimate of $S\,\mathbf{b}(\Omega)$

Multichannel-to-multi-channel signal enhancement

Spatial information can be preserved to some degree in output

Noise Reduction Methods

Matrix PMWF [Doclo and Moonen, 2002]

PMWF beamformer

Beamform-and-project method $\mathbf{W}_{\mathrm{BP}} = \mathbf{b}(\Omega)\mathbf{w}_{\mathrm{PMWF}}^{H}$

- $\mathbf{w}_{\mathrm{PMWF}}$ contains trade-off parameter μ
  - Small μ: less speech distortion, less noise reduction
  - Large μ: high noise reduction but high speech distortion
- **Disadvantage:** residual noise projected to direction Ω

Partial noise reduction [Klasen et al., 2007]

- Mixing between filtered and unprocessed signal, $\mathbf{W}_{\mathrm{BP+PM}}$
- Partially preserves spatial distribution of noise

Doclo and Moonen, "GSVD-based optimal filtering for single and multimicrophone speech enhancement," *IEEE Trans. Signal Processing*, 2002.

Klasen et al., “Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues,” *IEEE Trans. Signal Process.,* 2007.

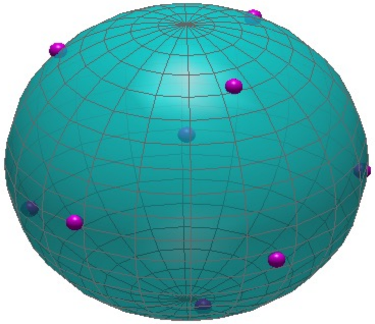
Noise Reduction Methods

Direction-Preserving Method [Herzog and Habets, 2019]

- Optimally preserves directions of plane waves
- Direction-preserving filter matrix $\mathbf{W}_{\mathrm{DP}} = \mathbf{W}_{\mathrm{DP}}(\alpha_1, \alpha_2, ..., \alpha_Q)$

  $\alpha_1, \alpha_2, ..., \alpha_Q$: directional gains for Q virtual sampling directions

- Insert $\mathbf{W}_{\mathrm{DP}}$ in MPMWF cost function
- Optimal solution for $\alpha_1, \alpha_2, ..., \alpha_Q$
- Lower bound $\alpha_{\min}$


Q=9 virtual sampling directions

In this work

Derived B-format expressions for $\mathbf{W}_{\mathrm{DP}}$ and $\alpha_1, \alpha_2, ..., \alpha_Q$

Herzog and Habets, "Direction-preserving Wiener matrix filtering for Ambisonic input-output systems", ICASSP, UK, 2019.

Outline

- Introduction
- Problem Formulation
- Noise Reduction Methods
- **Proposed Method**
- Performance Evaluation

Proposed Method
Motivation

- **In [Herzog and Habets, 2019]**
  - Fixed trade-off parameter μ for Ambisonic-to-Ambisonic noise reduction

Problem: Wind-noise is highly time-varying

➡ μ which adapts to current signal statistics is desired

- **In [Mirabilii and Habets, 2019]**
  - Difference-to-sum power ratio used for μ in the context of beamforming and wind-noise reduction

In this work 1. Derivation of the dipole-to-omnidirectional power ratio

2. Use as trade-off parameter for noise reduction

Mirabilii and Habets, "Multi-channel wind noise reduction using the Corcos Model," *ICASSP, UK,* 2019.

Herzog and Habets, "Direction-preserving Wiener matrix filtering for Ambisonic input-output systems", ICASSP, UK, 2019.

Proposed Method
B-format Power Ratio

Dipole-to-omnidirectional power ratio $\mathrm{PR} = \frac{g_o^2}{g_d^2} \frac{\phi_{\mathrm{dip}}}{\phi_{\mathrm{omni}}}$

g_o, g_d: arbitrary omnidirectional and dipole gains from B-format encoding

$\phi_{\mathrm{omni}} = \mathcal{E}\left\{|X_{Bo}|^2\right\}$ omnidirectional power

$\phi_{\mathrm{dip}} = \mathcal{E}\left\{|X_{Bx}|^2 + |X_{By}|^2 + |X_{Bz}|^2\right\}$ dipole power

Statistical expectation

- For plane wave $\mathrm{PR} = 1$
- For wind noise $\mathrm{PR} = 9$ (spatially white assumption)
- "Windiness" parameter $\widetilde{\mathrm{PR}} = \min\{\max\{(\mathrm{PR}-1)/8, 0\}, 1\}$
- Proposed trade-off parameter $\mu = 1 + \rho\widetilde{\mathrm{PR}}$

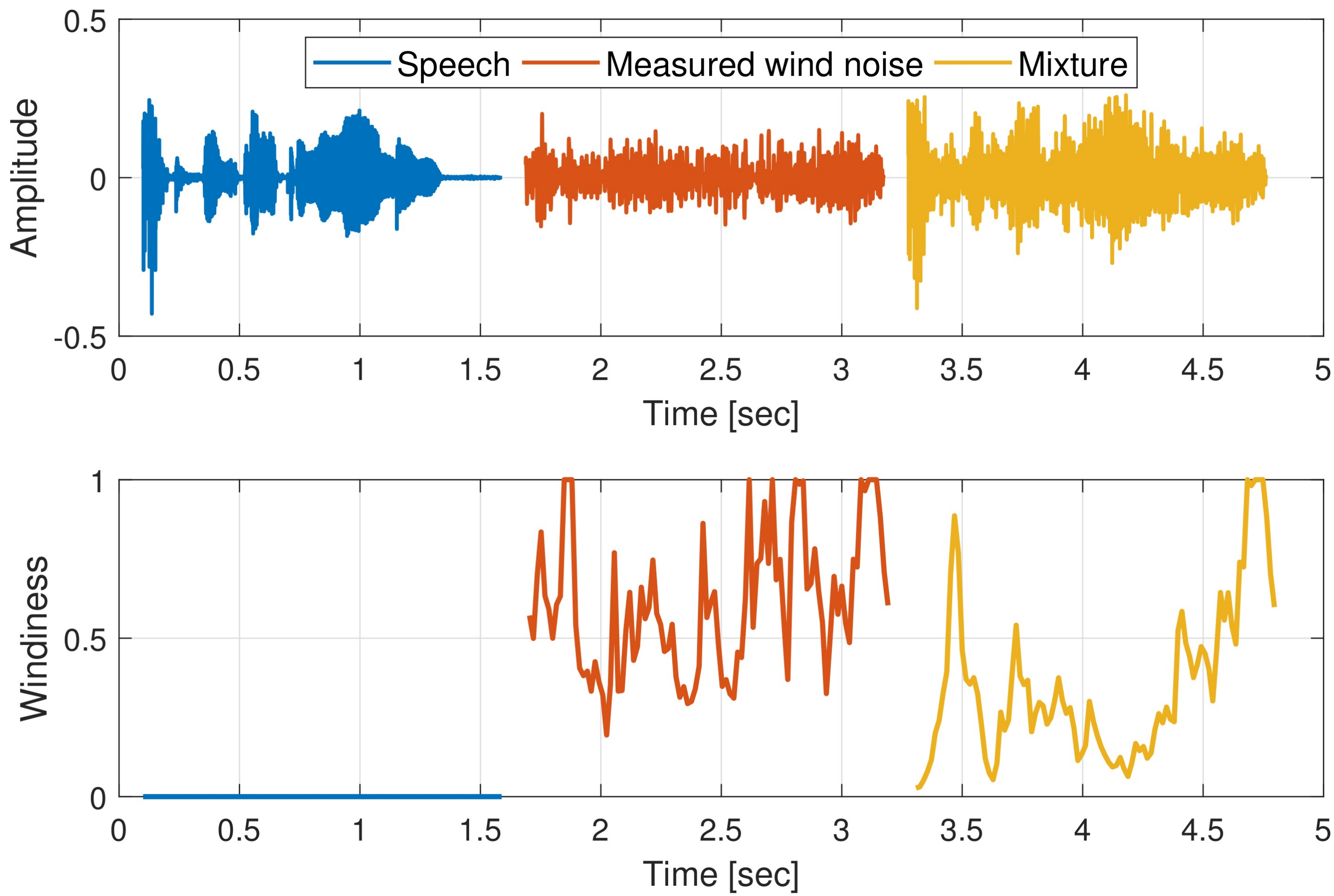
Adjustable scaling parameter

Proposed Method

Windiness

$$\mu = 1 + \rho\,\widetilde{\mathrm{PR}}$$

Windiness



Outline

- Introduction
- Problem Formulation
- Noise Reduction Methods
- Proposed Method
- **Performance Evaluation**

Performance Evaluation

Experimental setup

- Wind noise recorded with an AMBEO VR mic (10 samples)

×

- 1 male and 1 female speaker

×

- 4 different directions Ω_i

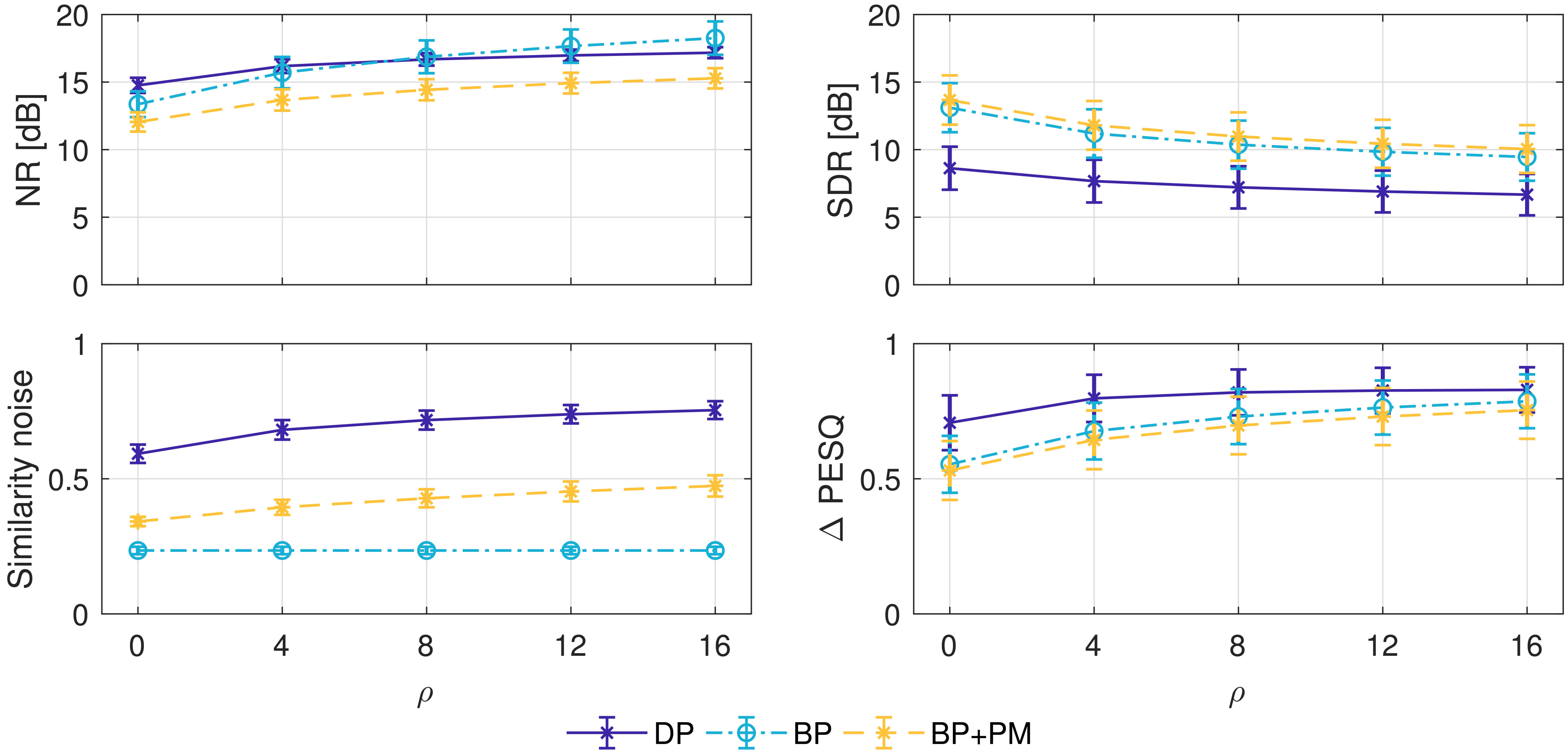
➡ 80 test samples (0 dB of input SNR)

Performance Evaluation
Processing and Performance Measures

- Methods under test
  - $\mathbf{W}_{\mathrm{DP}}, \mathbf{W}_{\mathrm{BP}}, \mathbf{W}_{\mathrm{BP+PM}}$ for different scaling parameters ρ
- Processing
  - Short-time Fourier transform domain
  - Recursive estimation of signal statistics
  - Oracle speech direction
- Objective evaluation
  - Noise Reduction (NR)
  - Signal-to-Distortion Ratio (SDR)
  - Perceptual Evaluation of Speech Quality (PESQ)
  - Noise Angular Similarity

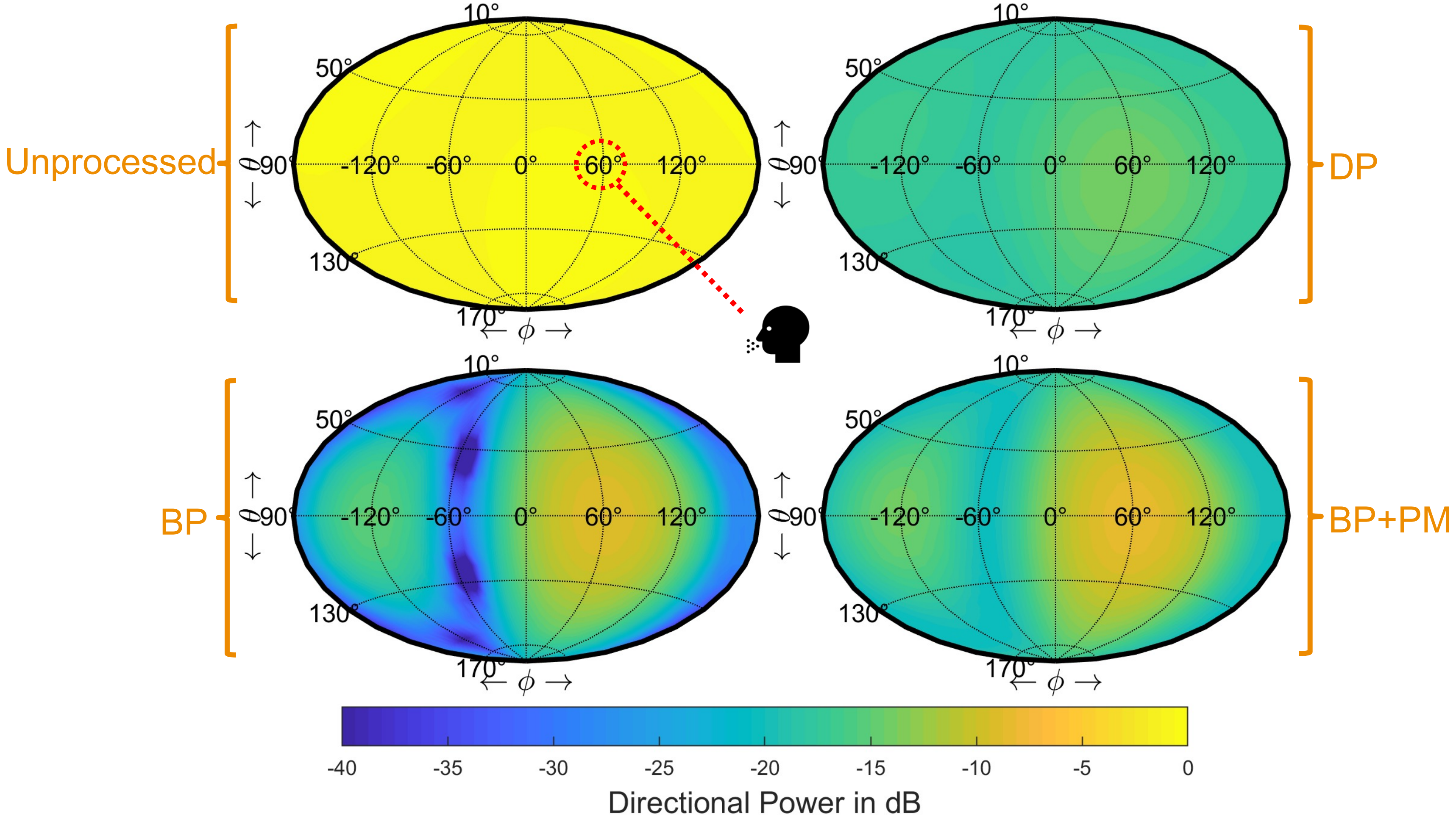
Performance Evaluation

Results



Performance Evaluation

Directional Distribution of Residual Noise

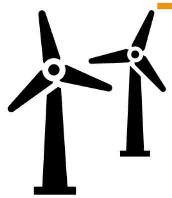


Audio samples: https://www.audiolabs-erlangen.de/resources/2021-ICASSP-BWNR

Summary

- Trade-off parameter of the PMWF matrix
  - Based on the dipole-to-omnidirectional power ratio
  - Adapts to fast-changing statistics of wind noise
  - Overall improvement of noise reduction, spatial preservation of the noise and speech quality compared to a fixed parameter

- Methods to reduce wind noise and preserve its spatial distribution
  - **BP**: + noise reduction and signal-to-distortion ratio
    - spatial preservation of the noise and speech quality
  - **BP+PM**: + spatial preservation of the noise
  - **DP**: + spatial preservation of the noise and speech quality
    - signal-to-distortion ratio



Thank you!