Wind Noise Reduction with Orthogonal Polynomial Expansion

Li Du*, Chao Pan*, Lijun Zhang*
* Center of Intelligent Acoustics and Immersive Communications,
School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an, China
E-mail: dulia@mail.nwpu.edu.cn, {cpan, zhanglj7385}@nwpu.edu.cn

***Abstract*—This paper is dedicated to mitigating unwanted ambient wind noise in outdoor multi-channel recordings. Different from conventional methodologies such as the multi-channel Wiener filter (MWF) and its extensions, we propose a wind noise reduction method from a beamforming perspective based on the spatial characteristics of wind noise captured by closely spaced microphones. Specifically, the noisy observation is decomposed into a parameterization space concerning frequency through orthogonal polynomial expansion. Subsequently, the estimated target source signals are reconstructed by selecting coefficients corresponding to desired components within the polynomial basis. Moreover, we introduce the Alternating Direction Method of Multipliers (ADMM) framework to attain a sparse solution for the coefficients during the decomposition phase. In comparison to existing multi-channel wind noise reduction methods, simulation results demonstrate the superior performance of the proposed method, particularly in low signal-to-noise ratio (SNR) scenarios.**

I. INTRODUCTION

Wind noise is a serious impediment to the intelligibility and quality of speech communication in outdoor scenarios, presenting a challenging task compared to other ambient noise [1], [2], [3], [4], [5], [6], [7], [8]. The inherent difficulties stem from the highly dynamic and nonstationary nature of wind-induced disturbances, compounded by a low signal-to-noise ratio (SNR). Traditional approaches for wind noise reduction can be divided into single-channel methods [2], [6], [9], [10] and multi-channel methods [7], [11], [12], [13]. In the multi-channel processings, common methods are based on the assumption that wind noise is uncorrelated. However, recent investigations showed that the spatial coherence of wind noise closely depends on the direction and the speed of free-field airflow, particularly apparent in the presence of strong winds [14], [15]. The non-zero spatial coherence of wind noise captured by closely spaced microphones is elucidated and can be modeled by a fluid-dynamics model in [16].

Utilizing the spatial coherence feature of wind noise, it is natural to design a multi-channel Wiener filter to effectively reduce wind noise. This approach relies on the accurate estimation of the power spectral densities (PSDs) for both wind noise and target signal, e.g., speech, as detailed in the works of [7], [17]. And extensions to this work include the estimation of a trade-off parameter for the wiener filter, denoted as the parametric multi-channel Wiener filter (PMWF) [17], [18] and improved PSD estimation methods [19].

A critical component of multi-channel noise reduction is beamforming, which exploits the spatio-temporal information to extract the target source signal from noisy observations [20]. The beamformers possess the ability to form a main beam directed towards the desired signal based on knowledge of the directions of the intended signal and interference signal, instead of the statistical properties. Among the diverse beamforming approaches, the differential beamformers provide the potential in speech communication systems due to their frequency-invariant beampatterns and small apertures [21], [22]. Inspired by the design of a differential beamformer with orthogonal polynomials [22], we present a novel method for wind noise reduction through orthogonal polynomial expansion.

To deal with the wind noise in microphone arrays, we first present the signal model and some important definitions. Then we investigate the spatial coherence between the speech signal and wind noise, clarifying their distinctions. An orthogonal polynomial expansion is performed to redefine the noisy observations in a frequency dependent parameterization space. Specifically, by analyzing the characteristics of speech source signals and wind noise on each order of orthogonal polynomials, we identify the orders in which speech source signals are dominant. Thus, the corresponding orders are selected to reconstruct an enhanced version of the desired source signal. An additional contribution is the introduction of the Alternating Direction Method of Multipliers (ADMM) scheme to optimize the decomposition coefficients. Simulations are conducted to illustrate the performance of the proposed method in terms of SNR, speech quality and speech intelligibility.

II. SIGNAL MODEL

Consider an M-element uniform linear microphone array in the outdoor environment where the wind noise dominates the ambient noise. The distance between two successive sensors is δ_0 and the source signal is from the direction of θ. The reverberant effect can be omitted. We formulate the problem and devise the approach in the short-time Fourier transform (STFT) domain [23]. Hence, the observation of the mth sensor can be modeled as

$$Y_m(f,t) = X_m(f,t) + V_m(f,t), m \in \{0, \cdots, M-1\}, \quad (1)$$

where f is the frequency bin index, t is the time frame index, $X_m(f,t)$ is the clean speech signal captured by the

mth sensor, and $V_m(f,t)$ is the additive ambient noise. In this work, $V_m(f,t)$ mainly consists of wind noise.

For convenience, the signal model in (1) is concatenated into the vector form

$$\begin{aligned}\mathbf{y}(f,t) &\triangleq \left[\begin{array}{cccc} Y_0(f,t) & Y_1(f,t) & \cdots & Y_{M-1}(f,t) \end{array}\right]^T \\ &= \mathbf{x}(f,t) + \mathbf{v}(f,t) \\ &= \mathbf{d}(f,\theta)X_0(f,t) + \mathbf{v}(f,t), \end{aligned} \tag{2}$$

where $\mathbf{x}(f,t)$ and $\mathbf{v}(f,t)$ are defined in the same way as $\mathbf{y}(f,t)$, and

$$\mathbf{d}(f,\theta) \triangleq \left[\begin{array}{c} 1 \\ e^{-j\frac{2\pi f}{c}\delta_0\cos(\theta)} \\ \vdots \\ e^{-j\frac{2\pi f}{c}(M-1)\delta_0\cos(\theta)} \end{array}\right] \tag{3}$$

is the array manifold vector, where $j=\sqrt{-1}$ is the imaginary unit and c is the speed of sound in air. It should be pointed out that, for general array geometries, the array manifold vector is a function of source direction and sensor positions, which is not as simple as expressed in (3). Then, the correlation matrix of $\mathbf{y}(f,t)$ is

$$\begin{aligned}\mathbf{\Phi}_{\mathbf{y}}(f,t) &= \mathbf{\Phi}_{\mathbf{x}}(f,t) + \mathbf{\Phi}_{\mathbf{v}}(f,t) \\ &= \phi_{X_0}(f,t)\mathbf{d}(f,\theta)\mathbf{d}^H(f,\theta) + \mathbf{\Phi}_{\mathbf{v}}(f,t), \end{aligned} \tag{4}$$

where $\mathbf{\Phi}_{\mathbf{x}}(f,t)$ and $\mathbf{\Phi}_{\mathbf{v}}(f,t)$ are the correlation matrices of $\mathbf{x}(f,t)$ and $\mathbf{v}(f,t)$, respectively, and $\phi_{X_0}(f,t) \triangleq \mathbb{E}\left[|X_0(f,t)|^2\right]$ is the variance of $X_0(f,t)$.

By applying a filter $\mathbf{h}(f,t)$ to the array observation $\mathbf{y}(f,t)$, the array output can be expressed as

$$\begin{aligned} Z(f,t) &= \mathbf{h}^H(f,t)\mathbf{y}(f,t) \\ &= X_{\text{fd}}(f,t) + V_{\text{rn}}(f,t), \end{aligned} \tag{5}$$

where

$$X_{\text{fd}}(f,t) = \mathbf{h}^H(f,t)\mathbf{x}(f,t), \tag{6}$$

$$V_{\text{rn}}(f,t) = \mathbf{h}^H(f,t)\mathbf{v}(f,t) \tag{7}$$

are the filtered desired source signal and the residual noise, respectively. The variance of the array output $Z(f,t)$ can be expressed as

$$\begin{aligned} \phi_Z(f,t) &= \mathbf{h}^H(f,t)\mathbf{\Phi}_{\mathbf{y}}(f,t)\mathbf{h}(f,t) \\ &= \phi_{X_{\text{fd}}}(f,t) + \phi_{V_{\text{rn}}}(f,t). \end{aligned} \tag{8}$$

Similarly, the variance/power of the desired source signal $\phi_{X_{\text{fd}}}(f,t)$ and the variance/power of the residual noise $\phi_{V_{\text{rn}}}(f,t)$ are

$$\phi_{X_{\text{fd}}}(f,t) = \phi_{X_0}(f,t)|\mathbf{h}^H(f,t)\mathbf{d}(f,\theta)|^2, \tag{9}$$

$$\phi_{V_{\text{rn}}}(f,t) = \mathbf{h}^H(f,t)\mathbf{\Phi}_{\mathbf{v}}(f,t)\mathbf{h}(f,t). \tag{10}$$

The output SNR can be expressed as

$$\text{oSNR}(f,t) \triangleq \frac{\phi_{X_0}(f,t)|\mathbf{h}^H(f,t)\mathbf{d}(f,\theta)|^2}{\mathbf{h}^H(f,t)\mathbf{\Phi}_{\mathbf{v}}(f,t)\mathbf{h}(f,t)}. \tag{11}$$

Various optimal beamformers are built for different design targets and optimization criteria, where adaptive beamformers like the minimum variance distortionless response (MVDR) beamformer require the good estimations of the covariance matrices of both desired source signal and additive noise. In this work, we introduce considerations of the spatial coherence of wind noise in a samll aperture array and design a beamformer that fully exploits the difference between the correlation matrices $\mathbf{\Phi}_{\mathbf{x}}(f,t)$ and $\mathbf{\Phi}_{\mathbf{v}}(f,t)$.

III. Covariance Matrix and Orthogonal Expansion

A. Covariance Matrix of Wind Noise

The covariance matrix of wind noise is dependent on the far-field wind velocity U_c and the wind stream direction ϑ, as stated in [14] and [16], with ϑ being defined identically to θ. Covariance matrix of the noise can be expressed as

$$\mathbf{\Phi}_{\mathbf{v}}(f,t) = \phi_V(f,t)\mathbf{\Upsilon}(f,\vartheta) \odot \left[\mathbf{a}(f,\vartheta)\mathbf{a}^H(f,\vartheta)\right], \tag{12}$$

where $\odot$ stands for Hadamard product, $\phi_V(f,t)$ is the variance of the noise, the (i,j)th element of the spatial coherence matrix $\mathbf{\Upsilon}(f,\vartheta)$ modeled by the Corcos model can be expressed as

$$\left[\mathbf{\Upsilon}(f,\vartheta)\right]_{i,j} = e^{-\frac{2\pi f}{U_c}\delta_0|i-j|g(\vartheta)} \tag{13}$$

with

$$g(\vartheta) = \alpha_1\left|\cos(\vartheta)\right| + \alpha_2\left|\sin(\vartheta)\right| \tag{14}$$

denotes a coherence decay parameter, α_1, α_2 are the longitudinal and the lateral coherence decay rates respectively, experimentally provided in [15]. For a wind stream with constant direction and speed, (13) is assumed to be time-invariant. And the vector $\mathbf{a}(f,\vartheta)$ is defined as

$$\mathbf{a}(f,\vartheta) \triangleq \left[\begin{array}{c} 1 \\ e^{j\frac{2\pi f}{U_c}\delta_0\cos(\vartheta)} \\ \vdots \\ e^{j\frac{2\pi f}{U_c}(M-1)\delta_0\cos(\vartheta)} \end{array}\right] \tag{15}$$

By comparing (3) and (15), the vector $\mathbf{a}(f,\vartheta)$ shares a similar form with the array manifold vector. Since U_c is often far smaller than the speed of sound in air, it is very likely that $\mathbf{d}(f,\theta)$ and $\mathbf{a}(f,\vartheta)$ belong to different subspaces, even if the intended source signal and wind noise are coming from the same direction, as do the $\mathbf{\Phi}_{\mathbf{x}}(f,t)$ and $\mathbf{\Phi}_{\mathbf{v}}(f,t)$.

B. The Orthogonal Expansion and the Decomposition

Following the work presented in [22], we express the array manifold vector as

$$\begin{aligned}\mathbf{d}(f,\theta) &= \left[\begin{array}{cccc} 1 & e^{-\jmath\varpi_1\cos\theta} & \cdots & e^{-\jmath\varpi_m\cos\theta} \end{array}\right]^T \\ &= \sum_{n=0}^{N-1}\mathbf{c}_n(f)\mathcal{P}_n(\cos\theta), \end{aligned} \tag{16}$$

where $\varpi_m = 2\pi(m-1)f\delta_0/c$, and

$$\mathbf{c}_n(f) \triangleq \left[\begin{array}{cccc} C_{n,0}(f) & C_{n,1}(f) & \cdots & C_{n,M-1}(f) \end{array}\right]^T \tag{17}$$

is a coefficient vector of length M, $n = 0, 1, 2, ..., N-1$ is the polynomial order, $\mathcal{P}_n(\cos\theta)$ is the orthogonal polynomial to approximate the exponential function of the array manifold vector. Several alternatives are available for expanding the array manifold, such as the MacLaurin's series, Jacobi polynomials [22], [24] and so on. Therefore, $\mathbf{c}_n(f)$ is determined by the geometric structure of the microphone array, the frequency band, the type of polynomial expansion and other factors.

We normalize the vector $\mathbf{c}_n(f)$'s according to

$$\mathbf{c}_n(f) \leftarrow \frac{1}{\|\mathbf{c}_n(f)\|}\mathbf{c}_n(f), \tag{18}$$

and then construct a matrix $\mathbf{C}(f)$ as

$$\mathbf{C}(f) \triangleq \left[\begin{array}{cccc} \mathbf{c}_0(f) & \mathbf{c}_1(f) & \cdots & \mathbf{c}_{N-1}(f) \end{array} \right], \tag{19}$$

which is clearly a matrix of size $M \times N$. Without loss of generality, we assume that $N \geq M$.

Given an array observation $\mathbf{y}(f,t)$, which is a function of the array manifold vector, we can get the optimal approximation according to (16) and divide it into

$$\begin{aligned} \mathbf{y}(f,t) &= \sum_{n=0}^{N-1} \beta_{y,n}(f,t)\mathbf{c}_n(f) \\ &= \mathbf{C}(f)\boldsymbol{\beta}_y(f,t), \end{aligned} \tag{20}$$

where

$$\boldsymbol{\beta}_y(f,t) \triangleq \left[\begin{array}{cccc} \beta_{y,0}(f,t) & \beta_{y,1}(f,t) & \cdots & \beta_{y,N-1}(f,t) \end{array} \right]^T. \tag{21}$$

Considering the speech signal in the STFT domain presents a large dynamic range, the normalization of the array observation $\mathbf{y}(f,t)$ is applied in practice

$$\mathbf{y}(f,t) \leftarrow \frac{1}{\|\mathbf{y}(f,t)\|_2}\mathbf{y}(f,t). \tag{22}$$

Similarly, we can define $\mathbf{x}(f,t) = \mathbf{C}(f)\boldsymbol{\beta}_x(f,t)$ and $\mathbf{v}(f,t) = \mathbf{C}(f)\boldsymbol{\beta}_v(f,t)$. It is evidently observed that a direct matrix inversion for solving the coefficient $\boldsymbol{\beta}_y(f,t)$ is prone to an ill-posed problem. Then we reformulate the matrix inversion as an optimization problem, introducing a regularization term into the cost function [25]. From a mathematical standpoint, the introduction of the ℓ_1 norm of $\boldsymbol{\beta}(f,t)$ is beneficial to the robustness of the solution. Furthermore, by the sparsity of the speech signal distribution exhibited in the frequency band, it is reasonable to assume that the $\boldsymbol{\beta}(f,t)$ is sparse in the STFT domain. The unconstrained optimization problem is shown in the following equation

$$\min \frac{1}{2}\|\mathbf{C}(f)\boldsymbol{\beta}_y(f,t) - \mathbf{y}(f,t)\|_2^2 + \lambda\|\boldsymbol{\beta}_y(f,t)\|_1. \tag{23}$$

Equation (23) is then solved by the ADMM framework [26]. Note that the variables in (23) are complex ones instead of real ones, so the optimization problem should be extended to the complex domain. A detailed discussion on the initialization and selection of parameters in the ADMM will be provided in the next section.

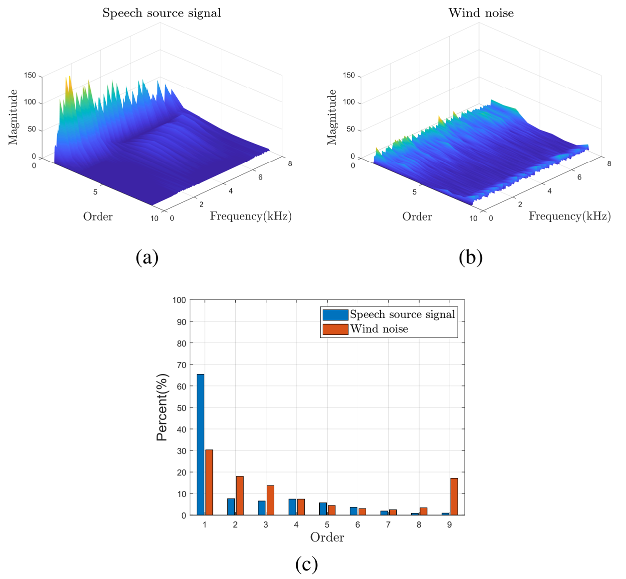

Fig. 1. $\boldsymbol{\beta}(f)$ coefficient distributions of (a) speech signal and (b)wind noise. (c) A percentage of distribution for the $\boldsymbol{\beta}(f)$ over each order.

Fig.1a and Fig.1b illustrate the values of $\boldsymbol{\beta}_x(f,t)$ and $\boldsymbol{\beta}_v(f,t)$, which correspond to the speech signal and wind noise respectively. It can be clearly seen that $\boldsymbol{\beta}_x(f,t)$ predominantly concentrates on the first and second orders at the lower frequencies, while $\boldsymbol{\beta}_v(f,t)$ is more evenly distributed across multiple orders. Fig.1c shows a bar chart depicting the magnitude of $\boldsymbol{\beta}_x(f,t)$, $\boldsymbol{\beta}_v(f,t)$ at each order as a percentage of the total. This above discussion implies that the low order components of $\boldsymbol{\beta}_y(f,t)$ contain the majority of speech signal components. As a consequence, the filter is constrained to these analytical findings in the reconstruction phase.

C. Reconstruction and Beamforming

According to (20), we have an estimation of the speech source signal $\mathbf{x}(f,t)$ as

$$\begin{aligned} \hat{\mathbf{x}}(f,t) &= \mathbf{C}(f)\hat{\boldsymbol{\beta}}_x(f,t) \\ &= \mathbf{C}(f)\mathbf{W}_L\boldsymbol{\beta}_y(f,t), \end{aligned} \tag{24}$$

where $\mathbf{W}_L$ is a constant matrix, which is defined as

$$\mathbf{W}_L \triangleq \left[\begin{array}{cc} \mathbf{I}_{L\times L} & \mathbf{0}_{L\times(N-L)} \\ \mathbf{0}_{(N-L)\times L} & \mathbf{0}_{(N-L)\times(N-L)} \end{array} \right], \tag{25}$$

and L is the order index at which the speech signal dominates, with $L \leq N$. Finally, the $\hat{\mathbf{x}}(f,t)$ is inverse normalized by

$$\mathbf{z}(f,t) \leftarrow \|\mathbf{y}(f,t)\|_2\hat{\mathbf{x}}(f,t). \tag{26}$$

In the case that we have an estimation of the array manifold vector toward the desired source, i.e., $\hat{\mathbf{d}}(f,t)$, we can calculate the array output according to

$$Z(f,t) = \frac{1}{\|\hat{\mathbf{d}}(f,t)\|^2}\hat{\mathbf{d}}^H(f,t)\mathbf{z}(f,t)$$
$$= \frac{1}{\|\hat{\mathbf{d}}(f,t)\|^2}\|\mathbf{y}(f,t)\|_2\hat{\mathbf{d}}^H(f,t)\mathbf{C}(f)\mathbf{W}_L\boldsymbol{\beta}_y(f,t) \quad (27)$$

In this paper, the direction of the source θ is known as a priori, and the $\mathbf{d}(f,\theta)$ is taken as the estimate of the array manifold vector, i.e., $\hat{\mathbf{d}}(f,t)$.

IV. Experiments and Analyses

A. Experimental Setup

Several experiments were conducted to validate and analyze the proposed method, with wind noise reduction methods developed in [17] serving as baselines. Sentences were randomly selected from the LibriSpeech dataset [22] and sampled at 16kHz. Wind noise was generated using an artificial wind noise generator [16] with temporal, spectral, and spatial characteristics matching measured multi-channel wind noise observations. The slow air stream velocity was set to $U_c = 1.8$ m/s and the fast velocity to 5 m/s in this work. Target speech signals were corrupted at three SNR values: -10dB, -5dB and 0dB. We utilized a closely spaced uniform linear array with eight microphones, each having an interelement spacing of 4 mm. The speech was kept in the endfire position, as well as wind noise in the direction $\vartheta = [0, 45°, 90°]$. We apply the STFT domain where the frame length is 32ms with half of the overlap between consecutive frames.

B. Results and Discussion

The parameter initialization and selection of the ADMM in the decomposition phase influence the performance of the proposed method. At each time-frequency frame, the dual variable λ in (23) is set to $\lambda = 0.1\mathbf{C}(f)^H\mathbf{y}(f,t)$. Fig.2 plots the spectrograms of the clean speech signal, wind noise, noisy signal at iSNR = 0dB, and the enhanced speech signal. Note that wind noise is predominantly focused in the low frequency band, which is the difference with other ambient noise characteristics. Fig.2d illustrates that the proposed method effectively addresses wind noise in the low frequency domain.

Considering the high fluctuation possessed by wind-induced disturbances, we mainly measure the fwSNR metrics of the proposed algorithm under various scenarios such as three wind directions and two wind speeds, with iSNR = -10dB. Averaged over all wind directions, the results in Table I show the effectiveness of the proposed method in complex wind turbulence situations.

In addition, a comparative analysis was conducted with the PMWF from [17] where the spatial complex coherence matrix of wind noise is also modeled as the classic Corcos model, and the baseline method MWF [7] and MVDR with the spatial coherence being an identity matrix. The results in Table II showcase the outstanding fwSNR, PESQ, and STOI improvements achieved by the proposed method at every iSNR level. *Highlighted values correspond to the best results.*

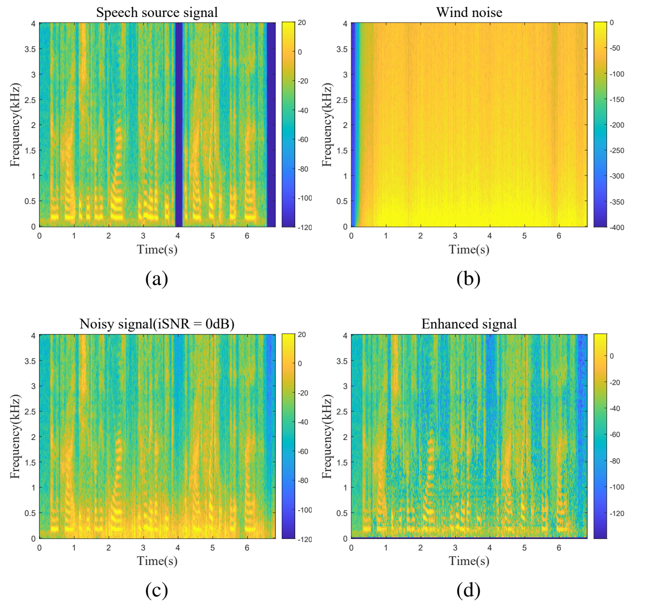

Fig. 2. The spectrograms of (a) clean signal, (b) wind noise, (c) noisy signal, (d) enhanced signal.

TABLE I
fwSNR improvements using the proposed method with the three wind directions and two wind speeds (iSNR = -10dB)

Speeds \ Directions	0	45°	90°	Average
1.8 m/s	14.61	18.14	17.02	16.59
5 m/s	15.10	17.93	15.84	16.29

TABLE II
fwSNR, PESQ and STOI Results With the proposed method and Baseline Methods

iSNR	Algorithm \ Metrics	ΔfwSNR	ΔPESQ	ΔSTOI
-10 dB	MVDR	4.16	0.45	0.0067
	MWF	12.12	0.93	0.0261
	PMWF	13.44	1.03	**0.1061**
	Proposed	**16.45**	**1.07**	**0.1061**
-5 dB	MVDR	4.03	0.44	0.0083
	MWF	7.44	1.04	0.0826
	PMWF	9.59	1.07	0.0827
	Proposed	**13.37**	**1.21**	**0.0926**
0 dB	MVDR	3.01	0.39	0.0156
	MWF	7.53	1.08	0.0383
	PMWF	7.71	1.27	0.0418
	Proposed	**10.21**	**1.48**	**0.0683**

Furthermore, the proposed method is more efficient and robust in the case of lower iSNR.

V. Conclusions

This paper proposes a multi-channel wind noise reduction method for outdoor recordings. Through analyzing the spatial characteristics of wind noise captured by closely spaced microphones and modeled by the Corcos model, a comprehensive analysis of the covariance matrix of target speech signals and ambient wind noise is conducted. Leveraging this analysis, we decompose the corrupted signals using the orthogonal polynomial expansion, highlighting the different coefficient distributions on the polynomial basis for speech source signals and wind noise. Then the estimated source signals are reconstructed by choosing low order coefficients corresponding to speech components. An additional contribution is the application of the ADMM framework, offering a more sparse and robust solution in the decomposition phase. The experiment results demonstrated that the proposed method outperforms baseline solutions in terms of fwSNR, PESQ, and STOI.

ACKNOWLEDGMENT

This work was supported in part by the NSFC (Nos. 62171373).

REFERENCES

[1] S. Kuroiwa, Y. Mori, S. Tsuge, M. Takashina, and F. Ren, "Wind noise reduction method for speech recording using multiple noise templates and observed spectrum fine structure," in *2006 International Conference on Communication Technology*. IEEE.

[2] M. N. Schmidt, J. Larsen, and F.-T. Hsiao, "Wind noise reduction using non-negative sparse coding," in *2007 IEEE Workshop on Machine Learning for Signal Processing*. IEEE.

[3] B. King and L. Atlas, "Coherent modulation comb filtering for enhancing speech in wind noise," in *International Workshop on Acoustice Echo and Noise Control*, 2008, pp. 14–17.

[4] E. Nemer and W. Leblanc, "Single-microphone wind noise reduction by adaptive postfiltering," in *2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. IEEE.

[5] J. A. Zakis, "Wind noise at microphones within and across hearing aids at wind speeds below and above microphone saturation," vol. 129, no. 6, pp. 3897–3907.

[6] C. Hofmann, T. Wolff, M. Buck, T. Haulick, and W. Kellermann, "A morphological approach to single-channel wind-noise suppression," in *IWAENC 2012; International Workshop on Acoustic Signal Enhancement*. VDE, 2012, pp. 1–4.

[7] P. Thuene and G. Enzner, "Maximum-likelihood approach to adaptive multichannel-wiener postfiltering for wind-noise reduction," in *Speech Communication; 12. ITG Symposium*. VDE, 2016, pp. 1–5.

[8] C. M. Nelke, *Wind noise reduction: signal processing concepts*. Wissenschaftsverlag Mainz, 2016.

[9] C. M. Nelke, N. Chatlani, C. Beaugeant, and P. Vary, "Single microphone wind noise psd estimation using signal centroids," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 7063–7067.

[10] C. M. Nelke and P. Vary, "Wind noise short term power spectrum estimation using pitch adaptive inverse binary masks," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 5068–5072.

[11] ——, "Dual microphone wind noise reduction by exploiting the complex coherence," in *Speech Communication; 11. ITG Symposium*. VDE, 2014, pp. 1–4.

[12] J. Park, J. Park, S. Lee, J. Kim, and M. Hahn, "Coherence-based dual microphone wind noise reduction by wiener filtering," in *Proceedings of the 8th International Conference on Signal Processing Systems*, 2016, pp. 170–172.

[13] G. W. Elko, J. M. Meyer, and T. F. Gaensler, "Noise-reducing directional microphone array," Mar. 29 2016, uS Patent 9,301,049.

[14] G. M. Corcos, "The structure of the turbulent pressure field in boundary-layer flows," vol. 18, no. 03, p. 353.

[15] R. H. Mellen, "On modeling convective turbulence," vol. 88, no. 6, pp. 2891–2893.

[16] D. Mirabilii and E. A. Habets, "Simulating multi-channel wind noise based on the corcos model," in *2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*. IEEE, sep 2018.

[17] ——, "Multi-channel wind noise reduction using the corcos model," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, may 2019.

[18] A. Herzog, D. Mirabilii, and E. A. P. Habets, "Direction preserving wind noise reduction of b-format signals," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE.

[19] D. Mirabilii and E. A. P. Habets, "Spatial coherence-aware multi-channel wind noise reduction," vol. 28, pp. 1974–1987.

[20] J. Benesty, M. M. Sondhi, and Y. A. Huang, Eds., *Springer Handbook of Speech Processing*. Springer Berlin Heidelberg.

[21] J. Benesty and C. Jingdong, *Study and design of differential microphone arrays*. Springer Science & Business Media, 2012, vol. 6.

[22] C. Pan, J. Benesty, and J. Chen, "Design of robust differential microphone arrays with orthogonal polynomials," *The Journal of the Acoustical Society of America*, vol. 138, no. 2, pp. 1079–1089, aug 2015.

[23] J. Benesty, I. Cohen, and J. Chen, *Fundamentals of signal enhancement and array signal processing*. John Wiley & Sons, 2017.

[24] L. Zhao, J. Benesty, and J. Chen, "Design of robust differential microphone arrays with the jacobi–anger expansion," *Applied Acoustics*, vol. 110, pp. 194–206, sep 2016.

[25] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer-Verlag New York Inc., Aug. 2006.

[26] S. Boyd, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.