MEASUREMENT, ANALYSIS AND SIMULATION OF WIND NOISE SIGNALS FOR MOBILE COMMUNICATION DEVICES

Christoph Matthias Nelke, Peter Vary

Institute of Communication Systems and Data Processing (ind)
RWTH Aachen University, Germany

ABSTRACT

In this contribution, we study the characteristics of sound generated by wind and a signal model for the synthesis of wind noise signals is derived. An analysis of the statistics of wind noise recorded in a laboratory setup is carried out with respect to the spectral and temporal properties of the signals. In particular, an autoregresive model is developed for the spectral shape description and the temporal statistics are modeled by a Markov chain. These two components are combined in a model which synthesizes reproducible artificial wind noise signals. Furthermore, a database of measured wind noise signals is provided. The aim of this model and the measured audio data is to provide wind signals for the evaluation of speech enhancement and noise reduction systems.

Index Terms— Wind noise analysis, wind noise simulation, communication quality assessment, wind noise database

1. INTRODUCTION

For the development and evaluation of noise reduction algorithms it is always necessary having defined scenarios to rate the achieved performance. For the simulation of a certain acoustic scenario, a variety of databases and methods exist. Reverberation effects can be simulated, e.g., by the method proposed in [1] or by the impulse responses presented in [2]. Different background noise types are provided by several noise databases (e.g., [3], [4]). However, none of these databases includes measurements of wind noise. Wind noise signals clearly differ from other noise types in terms of spectral and temporal characteristics. Because of the non-stationary behavior several methods were especially developed for the reduction of wind noise (see, e.g., [5], [6], [7] and references therein). To evaluate and compare these class of algorithms audio samples of wind noise are required.

This paper starts with a description of the measurement setup in Sec. 2. A signal analysis with respect to the disturbance to speech signals is given in Sec. 3. From the temporal and spectral characteristics highlighted in Sec. 4, a model for the signal generation of wind noise is derived in Sec. 5. In Sec. 6 the provided audio data and model implemented in MATLAB are described[1].

2. MEASUREMENT SETUP

The aim of this contribution is to analyze and provide audio signals generated by a wind stream. Studies as [8], [9] and [10] measured the influence on microphones which are not attached to a certain device. In contrast to that, we investigate wind noise in a scenario simulating a mobile phone call. Therefore, a mock-up phone of the size 12 cm x 6 cm x 3 cm is used which is equipped with two omnidirectional Beyerdynamics MM1 microphones. There are two arrangements resulting in microphone distances of 2 cm and 10 cm. For the examinations in Sec. 3.1 speech samples were played by an artificial head (HEAD acoustics HMS II.3). For the speech recordings both the hand-held position (HHP) and the hands-free position (HFP) according to [11] were considered. The air stream for the wind was generated by a compressed air supply. The advantage of this wind stream is, that there are only marginal background sounds, which often occurs while generating an air stream (e.g., with a fan). The measurements were carried out in an acoustic booth with a low reverberation time (< 100 ms). A comparison with recordings from outdoor situations showed only marginal differences to the measurements in the audio booth. The drawbacks of outdoor recordings are that it is obviously not possible to create reproducible situations in terms of the wind characteristics and there are always background noise not directly generated by the wind. Therefore, in the following measured wind noise using the aforementioned setup was used.

3. EVALUATION OF MEASUREMENTS

Typically, acoustical noise is produced by sound sources in background of the desired sound source. In contrast to that wind noise is directly generated by turbulences in a boundary layer close to the microphones. This leads to a more severe situation because the captured noise might be inaudible to the near-end speaker. The setup described in Sec. 2 was used to measure wind noise and speech signals in a realistic scenario. In Sec. 3.1 objective measures are used to give an insight in the degradations to a speech signal disturbed by wind noise. Besides, measurements with two microphones were carried out and Sec. 3.2 shortly explains the spatial correlation properties of these measurements.

3.1. Influence on Speech Quality

Measurements using an artificial head to simulate the near-end speaker were carried out considering both the HHP and the HFP. The speech levels were chosen to 89.3 dB at the mouth reference point and to 65.3 dB at the HHP and the HFP, respectively, as defined in [11]. Speech samples of female and male speakers from [12] were randomly taken. The degree of degradation was measured in terms of the speech quality by the PESQ value [13], [14], [15] and the intelligibility given by the STOI [16]. The PESQ value ranges from 1 (poor quality) to 5 (no degradation) and the intelligibility coefficient estimated by STOI ranges from 0 to 1, where 1 indicates a perfect intelligibility. Besides the global SNR was calculated over the whole signal length. For the two positions three scenarios were investigated: a constant low wind stream (≈ 5 m/s), a constant

[1] Matlab code for the model and the wind noise database can be found at `http://www.ind.rwth-aachen.de/~bib/nelke14a`

high wind stream ($\approx$ 10 m/s) and a varying wind stream with wind speeds up to 10 m/s. The latter condition reflects a realistic scenario in which gusts of the wind leads to fast changes of the wind speed. The evaluation of all scenarios is given in Tab. 1. Clearly negative

		SNR [dB]	PESQ	STOI
low wind	HHP	6.08	1.38	0.93
($\approx$ 5 m/s)	HFP	-9.19	1.04	0.79
high wind	HHP	-5.41	1.09	0.87
($\approx$ 10 m/s)	HFP	-20.68	1.02	0.7
wind gusts	HHP	-2.95	1.09	0.78
(up to 10 m/s)	HFP	-18.22	1.06	0.52

Table 1. Measures from noisy speech

SNR values can be seen in all cases, except the low wind case in HHP. This extreme annoying noise impairs the speech quality as seen by the low PESQ values but has not such a high influence on the speech intelligibility given by the STOI estimates.

3.2. Coherence Analysis

For measurement setups with two microphones, many algorithms for speech enhancement exploit the correlation between the two microphone signals ([17], [18], [19]). As frequency dependent correlation measure the magnitude squared coherence (MSC) can be used

$$\mathrm{MSC}(f) = \left| \frac{\Phi_{xy}(f)}{\sqrt{\Phi_{xx}(f)\Phi_{yy}(f)}} \right|^2 \quad (1)$$

where Φ_{xy} and Φ_{xx}, Φ_{yy} are the cross- and auto PSDs of the microphone signals x and y. Different coherence models exist for several acoustic environments. E.g., for a single sound source with correlated microphone signals the coherence takes the value 1 and a diffuse noise field leads to a coherence given by a sinc function which is dependent on the distance between the microphones. In contrast to that the coherence of wind noise signals is close to 0 over the whole frequency range. This results from the fact that wind stream generates sound by turbulences close to the microphone which shows no or only low spatial correlation [20]. In Fig. 1 the MSC of two different microphone arrangements is presented. In the frequency range which is relevant for the wind noise (0-4000 Hz) the MSC for both microphone distances takes low values close to zero ($\leq$0.01). This coherence property can be used for the detection and reduction in multi microphone recordings (see [21] and references therein).

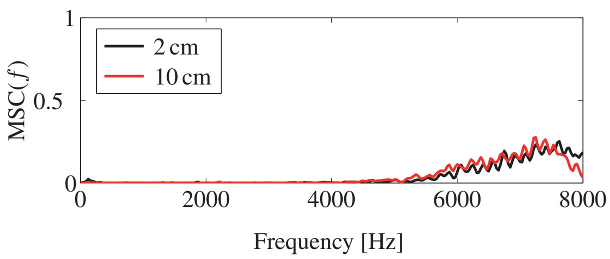

Fig. 1. Coherence of wind noise

4. SIGNAL ANALYSIS

In [22], it was shown that the noise field of a wind stream can be created by an arrangement of elementary emitters with monopole, dipole and quadrupole characters. The resulting noise is characterized by a low-frequency fast changing signal.

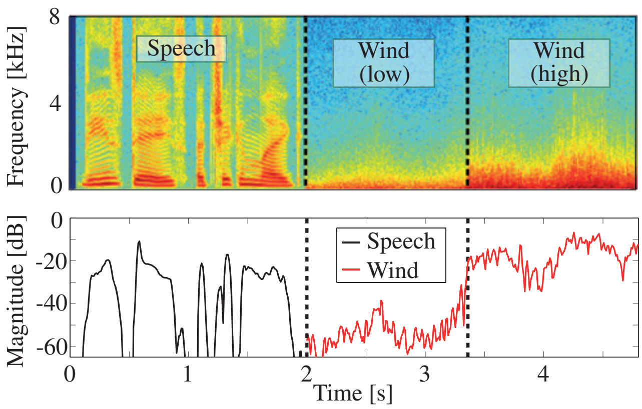

Fig. 2. Spectrogram of speech and wind (*top*) and corresponding frame energy (*bottom*)

4.1. Time Analysis

Examples of wind noise in the measured signals are shown in Fig. 2 in the second and third part of the spectrogram in the upper plot. It can clearly be seen, that wind noise results in a low-frequency signal with quickly changing temporal characteristics. The non-stationary characteristic is evident in the lower part of Fig. 2. This plot shows the progress of the short-term energy for wind noise over time. For a comparison a speech sample is also included. The frame size was chosen to 20 ms with 10 ms overlap as it is commonly used in real-time audio processing applications. The energy curve shows fast changing properties which are at least as high as the speech signal shown in the first two seconds. This nature makes it hardly possible to enhance speech with state-of-the-art approaches for noise reduction (e.g., [23] [24]) because they assume that noise signals varies slower than the speech signal. To quantify the non-stationary characteristic the variance of the short-term energy

$$\sigma_{\mathrm{ST}}(\lambda) = \frac{1}{L} \sum_{\kappa=\lambda-(L-1)/2}^{(L-1)/2} (E_x(\kappa) - \overline{E}_x)^2 \quad (2)$$

over L = 5 consecutive frames is measured. The frame index is given by λ, while κ is the position within the considered range of L frames which are represented by their energy $E_x(\kappa)$ and $\overline{E}_x$ giving the mean energy of the L frames. Fast changes of the signal within a short period result in a high variance of the short-term energy. The averaged values of $\sigma_{\mathrm{ST}}(\lambda)$ over 30 seconds audio samples resulting in $\overline{\sigma}_{ST}$ and are given in Tab. 2. For a comparison three typical noise types from [4] and 30 seconds of speech from [12] are also considered. The above mentioned temporal characteristics of wind noise can be read from the values of $\overline{\sigma}_{ST}$. Even the time varying Jackhammer noise shows a considerably lower variance compared to the wind signal. Only the speech signal shows a higher variance which is caused by the high energy differences between speech activity and speech pause segments.

	Car	Pub	Jackh.	Wind	Speech
$\overline{\sigma}_{ST}$ [dB]	4.47	6.4	7.58	13.67	18.28

Table 2. Averaged variance of short-term energy

4.2. Frequency Analysis

The spectral shape can be approximated by an $1/f$-shape with the frequency f. This behavior can be exploited for the estimation of the wind noise power spectral density (PSD) [5]. Further investigations were carried out by a linear prediction (LPC) analysis. Here, the measured wind noise signal is reproduced by an auto-regressive (AR) model resulting in an IIR filter described by the LPC coefficients. It turned out, that the LPC analysis provides a compact and accurate representation of the spectral shape of wind noise. In [25] a LPC analysis was applied to detect wind noise and speech segments in noisy signal. A sequential (sample-wise) approach is used for the determination of the LPC coefficients leading to the normalized least mean square (NLMS) computation rule (see, e.g., [26]). Considering the prediction error, it turned out that a prediction order of 5 is sufficient for modeling the spectral shape of wind noise signals. The aim of this investigation was to find out whether the LPC coefficients vary over time or with different energy levels of the wind noise. In Fig. 3 the progress of LPC coefficients estimation is depicted. A 10 second sample of measured wind noise was considered as shown in the spectrogram in the upper part. The middle part displays the short term energy and the lower plot reveals the temporal progress of the 5 LPC coefficients ($a_1, ..., a_5$). The 5th coefficient a_5 is close to zero over the whole segment, which proves that the chosen LPC order is high enough. Although the spectrogram might indicate a changing

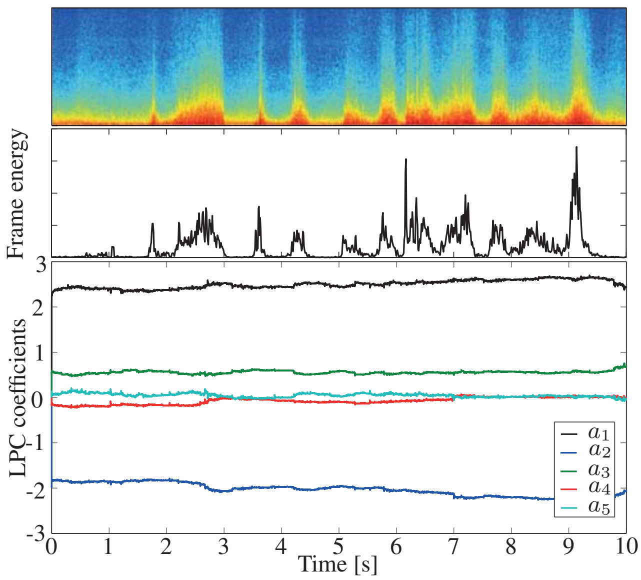

Fig. 3. Variation of LPC coefficients

spectral characteristic and thus variations of the LPC coefficients the progress of the estimated coefficients reveals a rather constant behavior. Hence, in the following the LPC coefficients describing the wind noise are assumed to be fix.

5. WIND NOISE MODEL

Based on the investigations from the previous section, a model is proposed which generates an artificial wind noise signal with predefined features. It should be mentioned that the derived model does not underlie the physics of wind noise generation. Primarily, the aim is to provide signals with similar statistics and spectral characteristics as recorded wind noise. A block diagram of this model is depicted in Fig. 4 where the containing blocks and parameters are described in Sec. 5.1 and 5.2.

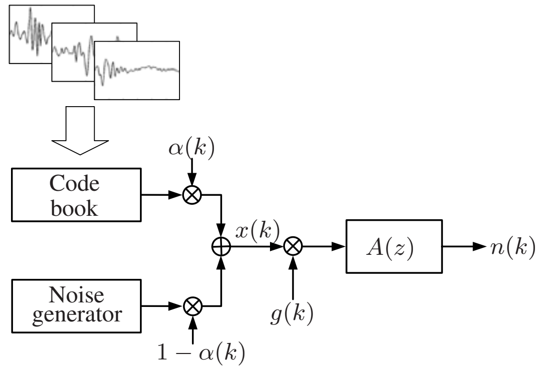

Fig. 4. Model for wind noise generation

5.1. Modeling the temporal characteristics

Regarding a wind stream in outdoor situations, the wind speed does not change abruptly. During a wind gust the wind speed is rising continuously to a high level and is then falling again. In contrast to that, the measured signal shows distinct changes. This is exemplified in Fig. 2, where the wind signal abruptly rises to a higher level at t=3.2 s. Thus, the temporal progress of measured wind noise signals can roughly be divided into two classes. In the first case the measured noise results from flow sound not directly next to the microphone (low wind). When wind gusts arise the sound level suddenly rises due to turbulences close to the microphone position (high wind). The two classes can be seen as two discrete states of a Markov model reflecting different wind speed conditions. Similar models were derived for the long time behavior of the wind speed ([27], [28]). Adding a third state when there is no air stream, the 3-state model depicted in Fig. 5 describes the temporal characteristic and is used in the following.

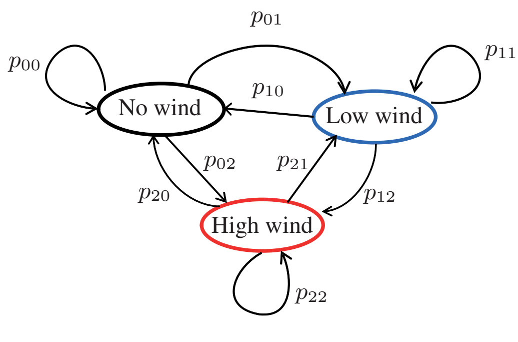

Fig. 5. 3-state Markov model

The transition probabilities p_{ij} from state i to state j determine the duration and occurrence rate of the corresponding wind condition. For the provided model the transition probabilities p_{02} and p_{20} were set to 0, because real measurements always show a transition phase of lower wind before reaching high wind segments from periods of still air and vice versa. The remaining transition probabilities can be trained by wind noise measurements. This is done by first labeling ranges of no, low and high wind speed in a given signal and compute the corresponding probabilities afterwards. The temporal characteristic in the proposed model is controlled by the gain $g(k)$ in Fig. 4. Here, the gain is given by a Gaussian distribution where the energy, i.e. the mean and variance, was set to values gained from wind measurements. These gains are recursively smoothed over time according to

$$g(k) = \alpha_g \cdot g(k-1) + (1-\alpha_g) \cdot \mathcal{X}_g(\mu_i, \sigma_i) \qquad (3)$$

with $\mathcal{X}_g$ as normal distributed variable with mean μ_i and variance σ_i of the state i of the Markov model.

5.2. Modelling the spectral characteristics

As mentioned in Sec. 4.1 an AR model is used for the generation of the synthesized wind signal. The LPC coefficients determine the LPC synthesis filter $A(z)$ in Fig. 4 and the gain $g(k)$ controls the energy of the synthesized signal over time. In this way $a_1 ... a_5$ define the spectral shape of the produced signal $n(k)$. It turned out that for the two aforementioned classes different input sequences $x(k)$ for the model must be considered. While for the low wind case a white noise signal is preferable (Fig. 4: Noise generator), the acoustics of the high wind case are better reproduced by taking short signal segments as input (Fig. 4: Code book). The code book was derived from the LPC analysis carried out in Sec. 4.1 by taking segment of 5-10 ms. In the following a code book size of 140 segments is considered from which excitation signal segments are randomly taken. The parameter $\alpha(k)$ determines the power ratio between the two excitation signals and must be adapted according to the current state (low wind / high wind). For the simulation of the fast changes in the wind signal α is toggled between 0 (low wind) and 0.5 (high wind).

5.3. Results

For the rectification of the model the averaged spectra of 10 seconds taken from measured and simulated wind noise are compared in Fig. 6. Both signals show a similar spectral shape of the low frequency noise. An example of a synthesized wind noise signal is

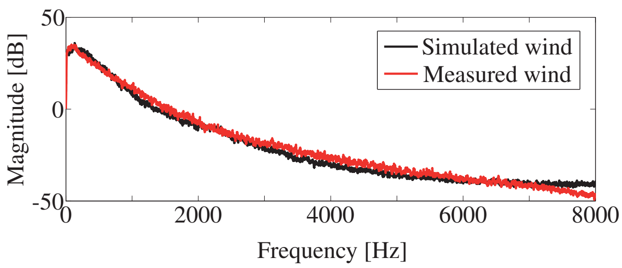

Fig. 6. Measured and simulated wind noise spectra

given in Fig. 7. The top spectrogram depicts a segment of measured wind noise comparison. The simulated wind signal in the bottom spectrogram shows a similar spectral and temporal characteristics. The averaged variance $\overline{\sigma}_{\text{ST}}$ of the simulated wind is 12.15 dB

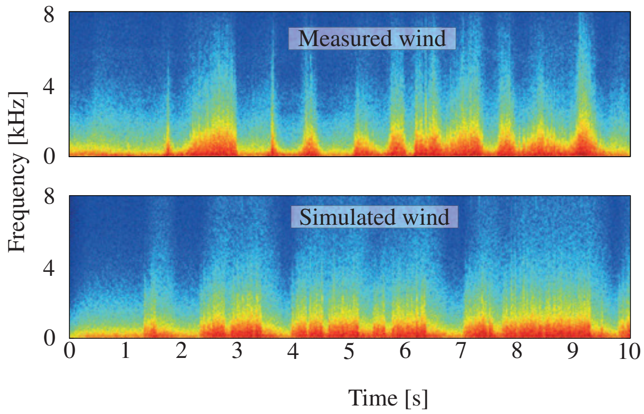

Fig. 7. Measured and simulated wind noise

which is close to the value of the measured wind signals (13.67 dB) which confirms a similar temporal behavior. Informal listening tests showed that the generated noise sounds very similar to measured wind noise.

6. PROVIDED DATA

Along with this contribution, a database of measured wind noise signals an MATLAB code for the wind noise model are provided. The database contains single microphone and multi microphone measurements. For both arrangements the setup was used as described in Sec. 2. The measurements includes conditions of constant wind stream and a more realistic scenario where the wind stream was varied during measurements. The latter signals contain the typical fluctuations resulting from wind gusts. Regarding the multi microphone measurements two setups were considered with microphone distances of 2 cm and 10 cm. Because especially for the multi microphone measurements at least one microphone contains only very little amount of wind noise in the HHP, all measurements provided are carried out in the HFP. Besides the MATLAB code for the wind noise simulator can be downloaded. The parameter p_{ij} representing the transition probabilities of the Markov model are pre-defined as trained from wind noise measurements. They can be changed to adjust the behavior of the generated wind noise signal. Further explanations for the execution are given in an included read-me file.

7. CONCLUSIONS

This contribution presents measured and artificially generated signals aiming for the quality assessment of noise reduction algorithms. Because in all commonly used noise databases wind noise signals are not included, it is necessary to provide those kind of signals for the evaluation of the special class of wind noise reduction methods. In a second step those signals were investigated with respect to their spectral and temporal properties. Then a simple model was to derived which generates artificial wind signals with similar characteristics as measured signals. The transition probabilities of the underlying Markov model can be adjusted to generate signals with desired properties in terms of the duration and occurrence rate of wind gusts.

8. REFERENCES

[1] J.B. Allen and D.A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

[2] M. Jeub, M. Schäfer, and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in *Digital Signal Processing, 2009 16th International Conference on*, Santorini, Greece, july 2009, pp. 1 –5.

[3] A. Varga and H. Steeneken, "Assessment for automatic speech recognition: Ii. noisex-92: A database and an experiment to study the effect of additive noise on speech recognition systems," *Speech Communication*, vol. 12, no. 3, pp. 247–251, 1993.

[4] ETSI EG 202 396-1, "Speech and multimedia transmission quality (STQ); Part 1: Background noise simulation technique and background noise database," March 2009.

[5] C.M. Nelke, N. Chatlani, C. Beaugeant, and P. Vary, "Single microphone wind noise PSD estimation using signal centroids," in *Proc. of IEEE Intern. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, Florence, Italy, May 2014.

[6] C.M. Nelke, N. Nawroth, M. Jeub, C. Beaugeant, and P. Vary, "Single microphone wind noise reduction using techniques of artificial bandwidth extension," in *Proc. of European Signal Processing Conf. (EUSIPCO)*, Bucharest, Romania, August 2012.

[7] C. Hofmann, T. Wolff, M. Buck, T. Haulick, and W. Kellermann, "A morphological approach to single-channel wind-noise suppression," in *Proc. of Intern. Workshop on Acoustic Echo and Noise Control (IWAENC)*, Aachen, Germany, Sept. 2012.

[8] J. Wuttke, "Microphones and wind," *Journal of the Audio Engineering Society*, vol. 40, no. 10, pp. 809–817, October 1992.

[9] E.B. Brixen and R. Hensen, "Wind generated noise in microphones - An overview - Part I," in *Audio Engineering Society Convention 120*, 5 2006.

[10] E.B. Brixen, "Wind generated noise in microphones - An overview - Part II," in *Audio Engineering Society Convention 121*, 10 2006.

[11] ETSI EG 201 377-2, "Speech processing, transmission and quality aspects (STQ); Specification and measurement of speech transmission quality; Part 2: Mouth-to-ear speech transmission quality including terminals," April 2004.

[12] P. Kabal, "TSP speech database," Tech. Rep., McGill University, Montreal, Canada, 2002.

[13] A.W. Rix, J.G. Beerends, M.P. Hollier, and A.P. Hekstra, "Perceptual evaluation of speech quality (PESQ) - A new method for speech quality assessment of telephone networks and codecs," in *Proc. of IEEE Intern. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, Salt Lake City, Utah, USA, 2001.

[14] ITU-T, "ITU-T Rec. P.862: Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs," February 2001.

[15] ITU-T, "ITU-T Rec. P.862.2: Wideband extension to recommendation P.862 for the assessment of wideband telephone networks and speech codecs," November 2007.

[16] C.H. Taal, R.C. Hendriks, R. Heusdens, and J. Jensen, "A short-time objective intelligibility measure for time-frequency weighted noisy speech," in *Proc. of IEEE Intern. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, Dallas, Texas, USA, march 2010, pp. 4214 –4217.

[17] Nima Yousefian, Philipos C. Loizou, and John H.L. Hansen, "A coherence-based noise reduction algorithm for binaural hearing aids," *Speech Communication*, , no. 0, pp. –, 2013.

[18] M. Jeub, C.M. Nelke, H. Krüger, C. Beaugeant, and P. Vary, "Robust dual-channel noise power spectral density estimation," in *Proc. of European Signal Processing Conf. (EUSIPCO)*, Barcelona, Spain, August 2011.

[19] C.M. Nelke, C. Beaugeant, and P. Vary, "Dual microphone noise PSD estimation for mobile phones in hands-free position exploiting the coherence and speech presence probability," in *Proc. of IEEE Intern. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, Vancouver, BC, Canada, May 2013.

[20] G. M. Corcos, "The structure of the turbulent pressure field in boundary-layer flows," *Journal of Fluid Mechanics*, vol. 18, pp. 353–378, 2 1964.

[21] C.M. Nelke and P. Vary, "Dual microphone wind noise reduction by exploiting the complex coherence," to appear in Proc. ITG-Fachtagung Sprachkommunikation, September 2014.

[22] Gerhard Müller and Michael Möser, *Handbook of engineering acoustics*, Springer, 2009.

[23] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 5, pp. 504–512, 2001.

[24] T. Gerkmann and R. Hendriks, "Noise power estimation based on the probability of speech presence," in *Proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, 2011.

[25] E. Nemer and W. Leblanc, "Single-microphone wind noise reduction by adaptive postfiltering," in *Proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, New York USA, 2009.

[26] P. Vary and R. Martin, *Digital Speech Transmission. Enhancement, Coding and Error Concealment*, Wiley-VCH Verlag, 2006.

[27] D.I. Jones and M.H. Lorenz, "An application of a markov chain noise model to wind generator simulation," *Mathematics and Computers in Simulation*, vol. 28, no. 5, pp. 391 – 402, 1986.

[28] A.D. Sahin and Z. Sen, "First-order markov chain approach to wind speed modelling," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 89, no. 34, pp. 263 – 269, 2001.