SOBM - A BINARY MASK FOR NOISY SPEECH THAT OPTIMISES AN OBJECTIVE INTELLIGIBILITY METRIC

Leo Lightburn and Mike Brookes

Dept. of Electrical and Electronic Engineering, Imperial College London, UK

ABSTRACT

It is known that the intelligibility of noisy speech can be improved by applying a binary-valued gain mask to a time-frequency representation of the speech. We present the SOBM, an oracle binary mask that maximises STOI, an objective speech intelligibility metric. We show how to determine the SOBM for a deterministic noise signal and also for a stochastic noise signal with a known power spectrum. We demonstrate that applying the SOBM to noisy speech results in a higher predicted intelligibility than is obtained with other masks and show that the stochastic version is robust to mismatch errors in SNR and noise spectrum.

Index Terms— Speech enhancement, noise reduction, speech intelligibility, binary mask, intelligibility metric

1. INTRODUCTION

At Signal-to-Noise Ratios (SNRs) below about $0\,\mathrm{dB}$ the intelligibility of noisy speech is significantly reduced and conventional speech enhancement techniques are normally unable to improve intelligibility even though they may give substantial improvements in SNR [1, 2]. A number of studies [3, 4] have shown that the intelligibility of noisy speech can be improved by applying a binary-valued gain mask in the Time-Frequency (TF) domain. The mask is set to 1 in TF regions dominated by speech energy and to a low value, often 0, in TF regions dominated by noise. These studies have inspired the development of enhancement algorithms that determine a binary mask by classifying the TF cells of the degraded speech as speech-dominated or noise-dominated and then synthesise the enhanced speech from the masked TF representation of the noisy speech [5, 6]. These algorithms typically use features extracted from the noisy speech as the input to a classifier. The internal parameters of the classifier are found during training by applying noisy speech samples together with a target output consisting of an oracle mask, i.e. a mask that is obtained with knowledge of the clean speech.

The most widely used oracle mask is the so-called Ideal Binary Mask (IBM) introduced in [7], which is a function of the instantaneous SNR in the corresponding TF cell. The mask is given by

$$B_{IBM}(k,m) = \begin{cases} 1 & |X(k,\,m)|^2 > \beta\,|N(k,\,m)|^2 \\ 0 & \text{otherwise} \end{cases}$$

where $X(k,m)$ and $N(k,m)$ are the complex Short Time Fourier Transform (STFT) coefficients of the speech and noise respectively in frequency bin k of frame m. The Local Criterion (LC), β, determines the SNR threshold above which the mask will equal 1. The observation that speech at an arbitrarily low SNR could be made fully intelligible by setting β approximately equal to the average SNR was explained in [8] whose authors suggested that the masked speech provides two independent speech cues, a noisy speech signal and a vocoded noise signal, and that it is the vocoded component that is responsible for improving the intelligibility. In [9] the vocoded signal component is created by the Target Binary Mask (TBM) in which the speech energy in each TF cell is compared with $\overline{X}(k)$, the average speech energy in that frequency bin. The TBM is given by

$$B_{\mathrm{TBM}}(k,m) = \begin{cases} 1 & X(k,\,m) > \beta'\overline{X}(k) \\ 0 & \text{otherwise} \end{cases} \qquad (1)$$

where β', the Relative Criterion (RC), typically lies in the range $\pm 5\,\mathrm{dB}$. The Universal Target Binary Mask (UTBM) [5] eliminates the speaker-dependence of the TBM by replacing $\overline{X}(k)$ in (1) by $\alpha\overline{\overline{X}}(k)$ where α is the average speech power and $\overline{\overline{X}}(k)$ is a speaker-independent power-normalised Long Term Average Speech Spectrum (LTASS) [10].

There is evidence that the intelligibility of speech depends not only on the instantaneous spectrum but also on its temporal modulation [11, 12]. The intelligibility of the mask-processed speech will not therefore be maximised if the classifier training target uses a mask such as the IBM, TBM or UTBM that depends only on the instantaneous spectrum. In this paper we propose an alternative oracle binary mask, the STOI-optimal Binary Mask (SOBM). The SOBM explicitly maximises an intelligibility metric, the Short-Time Objective Intelligibility Measure (STOI), that takes account of spectral modulation.

2. OBJECTIVE INTELLIGIBILITY MEASURE

The work of [13] led to the Articulation Index (AI) [14] as a standardised method of objectively estimating the intelligibility of speech. The AI and its successors, the SII and STI [15, 16], are computed from the SNRs in a set of frequency bands and have been extensively validated for speech

degraded by additive stationary noise. It has been found, however, that these SNR-based metrics are unable to model the effects of speech enhancement algorithms operating in the TF domain such as [17]. A number of more recent metrics are based on the correlation of the spectral amplitude modulation of the clean and degraded speech signals in each frequency band (see [18]). The most successful of these is STOI [19] which has been found to correlate well with the subjective intelligibility of both unenhanced and enhanced noisy speech signals [20, 21, 22]. Accordingly, in this paper, we advocate an oracle mask that optimises STOI.

We present here a brief overview of the STOI metric; readers are referred to [19] for a more detailed description. The clean speech is first converted into the STFT domain using 50%-overlapping Hanning analysis windows of length $25.6\,\mathrm{ms}$. The resultant complex-valued STFT coefficients, $X(k,m)$, are then combined into J third-octave bands by computing the TF cell amplitudes

$$X_j(m) = \sqrt{\sum_{k=K_j}^{K_{j+1}-1} |X(k,m)|^2} \quad \text{for } j = 1,\, \ldots,\, J \qquad (2)$$

where K_j is the lowest STFT frequency bin within frequency band j. The correlation between clean and degraded speech is performed on vectors of duration $(25.6 \times 30)/2 = 384\,\mathrm{ms}$. For each m, we therefore define the modulation vector

$$\mathbf{x}_{j,m} = [X_j(m-M+1), X_j(m-M+2), \ldots, X_j(m)]^T \qquad (3)$$

comprising $M = 30$ consecutive TF cells within frequency band j. The corresponding quantities for the degraded speech are $Y(k,m)$, $Y_j(m)$ and $\mathbf{y}_{j,m}$. Before computing the correlation, the degraded speech is clipped to limit the impact of frames containing low speech energy. The clipped TF cell amplitudes, denoted by a tilde superscript, are determined as

$$\widetilde{Y}_j(m) = \min\left(Y_j(m),\, \lambda \frac{\|\mathbf{y}_{j,m}\|}{\|\mathbf{x}_{j,m}\|} X_j(m)\right) \qquad (4)$$

where $\lambda = 6.623$ and $\|\ \|$ is the Euclidean norm. The corresponding modulation vectors are $\tilde{\mathbf{y}}_{j,m}$. The STOI contribution of the TF cell $(j,\, m)$ is then given by

$$d\left(\mathbf{x}_{j,m},\, \tilde{\mathbf{y}}_{j,m}\right) \quad \triangleq \quad \frac{\left(\mathbf{x}_{j,m} - \bar{x}_{j,m}\right)^T \tilde{\mathbf{y}}_{j,m}}{\|\mathbf{x}_{j,m} - \bar{x}_{j,m}\|\, \|\tilde{\mathbf{y}}_{j,m} - \bar{\tilde{y}}_{j,m}\|} \qquad (5)$$

where $\bar{x}_{j,m}$ denotes the mean of vector $\mathbf{x}_{j,m}$. The overall STOI metric is found by averaging the contributions of TF cells over all bands, j, and all frames, m.

3. STOI-OPTIMAL BINARY MASK

We derive the SOBM, the binary mask that maximises STOI for two cases: for a deterministic noise signal (DSOBM) and for stochastic noise with a known power spectrum (SSOBM).

3.1. SOBM for Deterministic noise (DSOBM)

We apply a binary mask, $B_j(m) \in \{0,\, 1\}$, by forming the masked signal $Z_j(m) = B_j(m)Y_j(m)$ and thence, analogous to (4), (3), the clipped masked vector $\tilde{\mathbf{z}}_{j,m}$. We optimise the mask separately in each band, j, by computing

$$B_j(m) = \underset{\{B_j(m):m=1,\ldots,T\}}{\arg\max} \left(\sum_{m=1}^{T} d\left(\mathbf{x}_{j,m}, \tilde{\mathbf{z}}_{j,m}\right)\right). \qquad (6)$$

We can compute this efficiently using a dynamic programming approach in which the active states at frame m are a subset of the 2^M possible values of $\mathbf{b}_{j,m}$. Associated with each active state is the STOI sum, $\sum_{s=1}^{m} d\left(\mathbf{x}_{j,s}, \tilde{\mathbf{z}}_{j,s}\right)$, corresponding to the best sequence $\{B_j(i) : i = 1, \ldots, m\}$ whose final M values match the entries of the corresponding $\mathbf{b}_{j,m}$ vector. At each iteration of the dynamic programming, we first form a list of potential active states at frame $m+1$ by appending $B_j(m+1) = 0$ and $B_j(m+1) = 1$ to each of the active states at frame m; this doubles the number of active states and may result in some duplicated states. For each of these potential active states, the STOI sum is updated to frame $m+1$ and the D distinct states that have the highest STOI sums are retained as the active states at frame $m+1$. The dynamic programming is initialised by taking $\mathbf{b}_{j,0}$ to be an all-zero vector. For the tests in Sec. 4, we used $D = 20000$.

3.2. SOBM for Stochastic noise (SSOBM)

For the stochastic case, we wish to determine the mask that maximises the expected value of STOI when $X(k,m)$ is known and the noise, $N(k,m) = Y(k,m) - X(k,m)$, is a stationary zero-mean complex Gaussian random variable with variance

$$\langle N(k,m)N^*(k,m)\rangle = \sigma_j^2 \qquad (7)$$

where $\langle\ \rangle$ denotes the expected value and σ_j^2 is assumed to have the same value for all k in frequency band j. We now wish to maximise the expected value of the sum given in (6). To make the analysis tractable, we assume that clipping is very rare in the stochastic noise case, so that $\widetilde{Y}_j(m) \approx Y_j(m)$ in (4). It follows from (7) that $2\sigma_j^{-2}\,|Y(k,m)|^2$ has a non-central χ^2 distribution with 2 degrees of freedom and non-centrality parameter $R(k,m) = 2\sigma_j^{-2}\,|X(k,m)|^2$. From (2), therefore, $2\sigma_j^{-2}Y_j^2(m)$ has a non-central χ^2 distribution with $\nu_j = 2\,(K_{j+1} - K_j)$ degrees of freedom and non-centrality parameter

$$R_j(m) = 2\sigma_j^{-2} \sum_{k=K_j}^{K_{j+1}-1} |X(k,m)|^2 .$$

Thus $\sqrt{2}\sigma_j^{-1}Y_j(m)$ has a non-central χ distribution with mean [23, 24] given by

$$\left\langle \sqrt{2}\sigma_j^{-1}Y_j(m) \right\rangle = \sqrt{\frac{\pi}{4}}\sigma_j L_{0.5}^{(0.5\nu_j-1)}\left(-0.5R_j(m)\right)$$

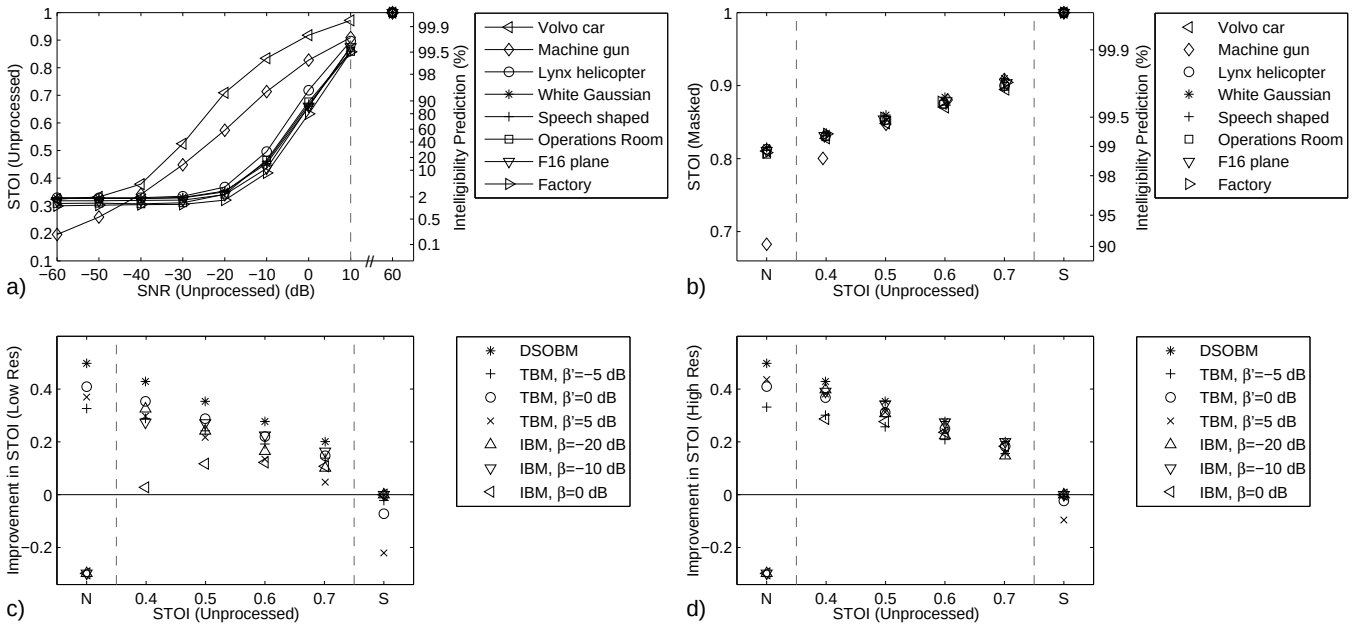

Fig. 1: a) STOI against SNR for the 8 tested noise types. b) Average STOI of masked speech against STOI before processing for the deterministic algorithm, DSOBM, applied to speech containing different noise types. Average improvement in STOI across all noise types against STOI before processing. The TBMs and IBMs have c) third-octave band resolution and d) full STFT resolution. "N" and "S" denote "noise-only" and "clean speech" input signals, respectively.

and second moment $\left\langle 2\sigma_j^{-2} Y_j^2(m) \right\rangle = \nu_j + R_j(m)$ where $L_n^{(\alpha)}(z)$ is a generalised Laguerre polynomial [25]. Defining the non-centrality vector, $\mathbf{r}_{j,m}$, analogous to (3), we can write

$$\langle \mathbf{z}_{j,m} \rangle = \sqrt{\frac{\pi}{4}} \sigma_j \mathbf{b}_{j,m} \circ L_{0.5}^{(0.5\nu_j - 1)} \left(-0.5 \mathbf{r}_{j,m} \right) \qquad (8)$$

where $\circ$ denotes elementwise multiplication and $L_n^{(\alpha)}(\ )$ acts elementwise on a vector argument. If we assume $Y_j(m)$ and $Y_j(n)$ are independent for $m \neq n$, we have

$$\begin{aligned} \left\langle \left\| \mathbf{z}_{j,m} - \bar{z}_{j,m} \right\|^2 \right\rangle &= \left\langle \left\| \mathbf{z}_{j,m} \right\|^2 \right\rangle - M \left\langle \bar{z}_{j,m}^2 \right\rangle \\ &= 0.5\sigma_j^2 \frac{M-1}{M} \mathbf{b}_{j,m}^T \left(\nu_j + \mathbf{r}_{j,m} \right) \\ &\quad - \frac{\pi \sigma_j^2}{4M} \left(\mathbf{b}_{j,m}^T L_{0.5}^{(0.5\nu_j - 1)} \left(-0.5 \mathbf{r}_{j,m} \right) \right)^2 \\ &\quad + \frac{\pi \sigma_j^2}{4M} \left\| \mathbf{b}_{j,m}^T \circ L_{0.5}^{(0.5\nu_j - 1)} \left(-0.5 \mathbf{r}_{j,m} \right) \right\|^2 . \end{aligned} \qquad (9)$$

Finally, combining (5), (8) and (9), we can calculate

$$\left\langle d\left(\mathbf{x}_{j,m},\, \mathbf{z}_{j,m} \right) \right\rangle \approx \frac{\left(\mathbf{x}_{j,m} - \bar{x}_{j,m} \right)^T \left\langle \mathbf{z}_{j,m} \right\rangle}{\left\| \mathbf{x}_{j,m} - \bar{x}_{j,m} \right\| \sqrt{\left\langle \left\| \mathbf{z}_{j,m} - \bar{z}_{j,m} \right\|^2 \right\rangle}}.$$

4. EVALUATION

The SOBM was evaluated using a subset of TIMIT [26] and seven noise types from the NOISEX-92 corpus [27]. Fig. 1a shows the average STOI plotted against SNR for speech degraded with each noise type. Most noise types give similar curves, with the exceptions of 'Volvo', which is predominately low frequency, and 'machine gun', which is highly non-stationary. The right hand axis gives the predicted intelligibility from [19] for previously unheard sentences.

Fig. 1b plots the average STOI of the masked speech against the STOI before processing, for the DSOBM applied to speech degraded with different noise types. The symbols "N" and "S" on the horizontal axis denote "noise-only" and "clean speech" input signals, respectively. The DSOBM resulted in a large improvement in STOI for all noise types, at all noise levels except for "S"; in the latter case, STOI was unchanged from a unprocessed value of 1. With the exception of machine gun noise at very poor SNRs, the DSOBM resulted in an improvement in STOI that was largely independent of noise type and in an average STOI above 0.8 for every noise level including "N" (corresponding to >98% intelligibility).

Fig. 1c shows the average improvement in STOI across all noise types against the STOI before processing, for the DSOBM, and selected IBMs and TBMs, where the masks all use identical third-octave band frequency resolutions. The DSOBM outperformed all of the tested TBMs and IBMs at all input noise levels excluding "S". After the DSOBM, the best performing mask was the TBM with β'=0 dB. The TBMs gave consistently good results for noisy speech, but degraded the intelligibility of clean speech. The IBMs preserved the intelligibility of clean speech, but performed worse than the

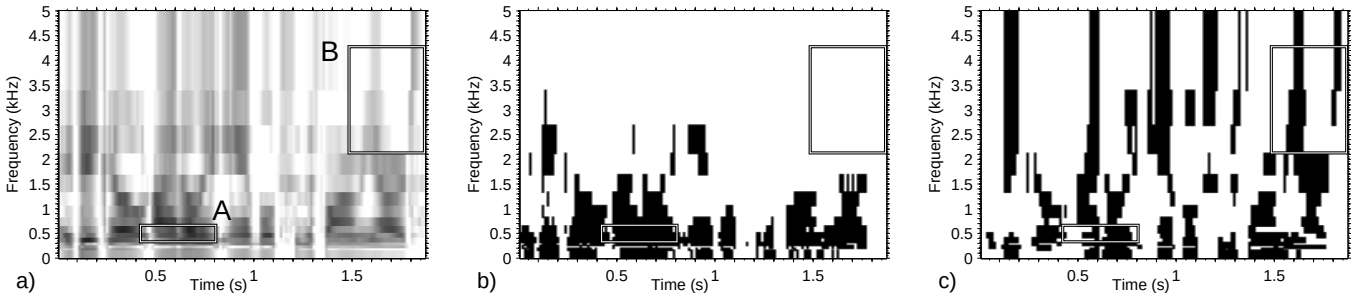

Fig. 2: Third-octave band resolution spectrogram of a) clean speech, and b) an IBM, computed by mixing the speech with WGN at -20 dB SNR, with β=-20 dB. c) The SSOBM, optimised for the same noise type and SNR. High energy (A) and low energy (B) regions of the plots are highlighted for comparison.

TBMs with very noisy speech.

In Fig. 1d the IBMs and TBMs used the full STFT resolution, much higher than that of the DSOBM. For test samples with unprocessed STOIs below 0.6, the DSOBM still gave the greatest improvement in STOI of all tested masks. For unprocessed STOIs of 0.6 and above, the improvement in STOI given by the DSOBM and the IBM with β=-10 dB was approximately equal.

Fig. 3 plots the improvement in STOI for different SSOBMs relative to the DSOBM averaged over all noises except machine gun noise, which is plotted separately. The SSOBM gives about 0.02 less STOI improvement than the DSOBM at all noise levels except for "S". To assess the effect of mismatch, we determined the SSOBMs for white-noise at SNRs of -60 and $-10\,\mathrm{dB}$ and applied these masks to all test signals ($\triangleright$, $\triangleleft$ in Fig. 3). We see that, except for "S", the STOI improvement is almost equal to that of the SSOBM that used a matched noise spectrum and SNR. This demonstrates that it is possible to use the SSOBM for $-60\,\mathrm{dB}$ white noise as a noise-independent and SNR-independent mask with little loss in intelligibility compared to the optimum DSOBM. The highly non-stationary machine gun noise is plotted separately in Fig. 3; its intermittent nature means that the SSOBM performs significantly worse than the DSOBM.

Fig. 2 shows a third-octave resolution spectrogram of speech, alongside an IBM with matching resolution and β=-20 dB, and the SSOBM, both computed for speech with white noise at -20 dB SNR. In both the high energy (A) and low energy (B) highlighted regions of the spectrogram the SOBM has captured the temporal modulations in the speech spectrum more successfully than the IBM. The average STOI contributions, (5), in regions A and B respectively are 0.52 and -0.18 for the IBM versus 0.82 and 0.85 for the SSOBM.

Fig. 4 shows the distribution of the difference in TF cell STOI contributions, (5), between the SSOBM and the IBM for the example of Fig. 2. In 76% of TF cells, (5) from the SSOBM was higher than from the IBM and in a significant number of cells it was much higher.

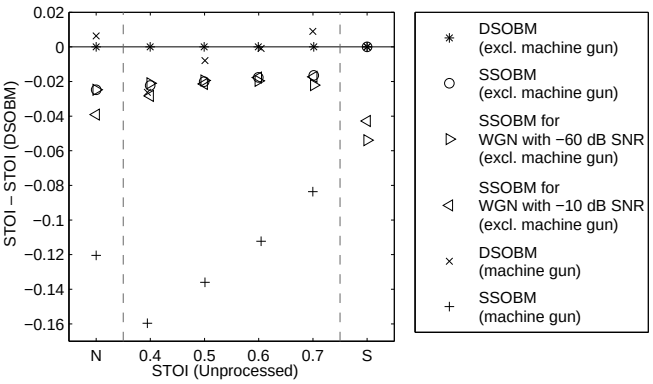

Fig. 3: Improvement in STOI for different masks relative to the DSOBM averaged over all noises other than machine gun noise, which is plotted separately.

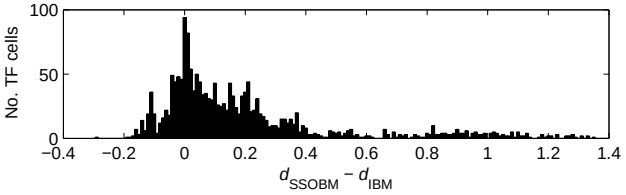

Fig. 4: Distribution of the difference between (5) computed on corresponding pairs of modulation vectors in SSOBM-processed and IBM-processed speech.

5. CONCLUSION

We have presented a new oracle mask, the SOBM, that explicitly maximises an objective intelligibility metric and is suitable for training a mask-based speech enhancer. For deterministic additive noise, the DSOBM always results in a higher predicted intelligibility than other oracle masks. When we assume a stochastic noise signal, the SSOBM achieves a performance close to the DSOBM for a wide range of SNRs and noise types, even when the noises used for mask optimisation and testing are mismatched.

6. REFERENCES

[1] Yi Hu and Philipos C. Loizou, "A comparative intelligibility study of single-microphone noise reduction algorithms," *J. Acoust. Soc. Am.*, vol. 122, pp. 1777–1786, 2007.

[2] Gaston Hilkhuysen, Nikolay Gaubitch, Michael Brookes, and Mark Huckvale, "Effects of noise suppression on intelligibility: dependency on signal-to-noise ratios," *J. Acoust. Soc. Am.*, vol. 131, no. 1, pp. 531–539, 2012.

[3] Ning Li and Philipos C. Loizou, "Factors influencing intelligibility of ideal binary-masked speech: Implications for noise reduction," *J. Acoust. Soc. Am.*, vol. 123, no. 3, pp. 1673–1682, Mar. 2008.

[4] Douglas S. Brungart, Peter S. Chang, Brian D. Simpson, and DeLiang Wang, "Isolating the energetic component of speech-on-speech masking with ideal time-frequency segregation," *J. Acoust. Soc. Am.*, vol. 120, pp. 4007–4018, 2006.

[5] Sira Gonzalez and Mike Brookes, "Mask-based enhancement for very low quality speech," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, May 2014.

[6] A. A. Kressner, D. V. Anderson, and Rozell C. J., "Causal binary mask estimation for speech enhancement using sparsity constraints," in *Proc Intl Congress on Acoustics*, Montreal, June 2013.

[7] DeLiang Wang, "On ideal binary mask as the computational goal of auditory scene analysis," in *Speech Separation by Humans and Machines*, P. Divenyi, Ed., pp. 181–197. Kluwer Academic, 2005.

[8] Ulrik Kjems, Michael S. Pedersen, Jesper B. Boldt, Thomas Lunner, and DeLiang Wang, "Speech intelligibility of ideal binary masked mixtures," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Aalborg, Denmark, Aug. 2010, pp. 1909–1913.

[9] Ulrik Kjems, Jesper B. Boldt, Michael S. Pedersen, Thomas Lunner, and DeLiang Wang, "Role of mask pattern in intelligibility of ideal binary-masked noisy speech," *J. Acoust. Soc. Am.*, vol. 126, no. 3, pp. 1415–1426, Sept. 2009.

[10] D. Byrne, H. Dillon, K. Tran, S. Arlinger, K. Wilbraham, R. Cox, B. Hayerman, R. Hetu, J. Kei, C. Lui, J. Kiessling, M. N. Kotby, N. H. A. Nasser, W. A. H. El Kholy, Y. Nakanishi, H. Oyer, R. Powell, D. Stephens, T. Sirimanna, G. Tavartkiladze, G. I. Frolenkov, S. Westerman, and C. Ludvigsen, "An international comparison of long-term average speech spectra," *J. Acoust. Soc. Am.*, vol. 96, no. 4, pp. 2108–2120, Oct. 1994.

[11] Les Atlas and Shihab A Shamma, "Joint acoustic and modulation frequency," *EURASIP Journal on Applied Signal Processing*, vol. 7, pp. 668–675, 2003.

[12] Rob Drullman, Joost M Festen, and Reinier Plomp, "Effect of temporal envelope smearing on speech reception," *J. Acoust. Soc. Am.*, vol. 95, pp. 1053, 1994.

[13] N. R. French and J. C. Steinberg, "Factors governing the intelligibility of speech sounds," *J. Acoust. Soc. Am.*, vol. 19, no. 1, pp. 90–119, 1947.

[14] ANSI, "Methods for the calculation of the articulation index," ANSI Standard ANSI S3.5–1969, American National Standards Institute, New York, 1969.

[15] ANSI, "Methods for the calculation of the speech intelligibility index," ANSI Standard S3.5–1997 (R2007), American National Standards Institute, 1997.

[16] IEC, "Objective rating of speech intelligibility by speech transmission index," EU Standard EN60268-16, International Electrotechnical Commission, May 2003.

[17] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 2, pp. 443–445, 1985.

[18] Cees H. Taal, Richard C Hendriks, Richard Heusdens, and Jesper Jensen, "An evaluation of objective measures for intelligibility prediction of time-frequency weighted noisy speech," *J. Acoust. Soc. Am.*, vol. 130, no. 5, pp. 3013–3027, 2011.

[19] C. H. Taal, R. C. Hendriks, R. Heusdens, and J. Jensen, "An algorithm for intelligibility prediction of time-frequency weighted noisy speech," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 7, pp. 2125–2136, Sept. 2011.

[20] Gaston Hilkhuysen, Nickolay Gaubitch, Michael Brookes, and Mark Huckvale, "Effects of noise suppression on intelligibility. II: An attempt to validate physical metrics," *J. Acoust. Soc. Am.*, vol. 135, no. 1, pp. 439–450, Jan. 2014.

[21] Angel M. Gomez, Belinda Schwerin, and Kuldip Paliwal, "Objective intelligibility prediction of speech by combining correlation and distortion based techniques," in *Proc. Interspeech Conf.*, 2011.

[22] Belinda Schwerin and Kuldip Paliwal, "An improved speech transmission index for intelligibility prediction," *Speech Communication*, 2014.

[23] J. H. Park, "Moments of the generalized Rayleigh distribution," *Quarterly of Applied Mathematics*, vol. 19, pp. 45–49, 1961.

[24] A. B. Olde Daalhuis, "Confluent hypergeometric functions," In Olver et al. [28], chapter 13, pp. 321–349.

[25] T. H. Koornwinder, R. Wong, R. Koekoek, and R. F. Swarttouw, "Orthogonal polynomials," In Olver et al. [28], chapter 18, pp. 436–484.

[26] John S. Garofolo, Lori F. Lamel, William M. Fisher, Jonathan G. Fiscus, David S. Pallett, Nancy L. Dahlgren, and Victor Zue, "TIMIT acoustic-phonetic continuous speech corpus," Corpus LDC93S1, Linguistic Data Consortium, Philadelphia, 1993.

[27] A. Varga and H. J. M. 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. 3, no. 3, pp. 247–251, July 1993.

[28] Frank W. J. Olver, Danel W. Lozier, Ronald F. Boisvert, and Charles W. Clark, Eds., *NIST Handbook of Mathematical Functions*, CUP, 2010.