# Robust Trade-off Parameter for the Speech Distortion Weighted Multichannel Wiener Filter

Ina Kodrasi*, Daniel Marquardt, and Simon Doclo

*Abstract*—The use of an optimal trade-off parameter in the speech distortion weighted multichannel Wiener filter (MWF) is of crucial importance, since it controls the resulting speech distortion and noise reduction.

Aiming at achieving a high noise reduction performance at a low speech distortion, in this letter we propose computing the time-frequency dependent trade-off parameter as the maximum curvature of the parametric plot of noise power versus speech distortion. Within the proposed framework speech distortion and noise power can be weighted in advance, e.g. based on perceptually motivated criteria. Theoretical and experimental results show that the proposed parameter results in a more robust filter to SNR estimation errors than the traditional MWF, significantly improving the performance.

*Index Terms*—SDW-MWF, trade-off parameter, curvature

## I. INTRODUCTION

Multichannel Wiener filtering (MWF) is a commonly used technique aiming at noise reduction by minimizing the mean-square error between the output signal and the speech component in the reference microphone [1]. The error signal typically consists of a speech distortion term and a noise power term. While the traditional MWF assigns equal importance to both terms, the speech distortion weighted multichannel Wiener filter (SDW-MWF) incorporates a trade-off parameter to trade-off between speech distortion and noise reduction [2], [3]. An empirically selected and fixed trade-off parameter has been generally used in the SDW-MWF [3], [4], which can be suboptimal since it does not reflect the typically changing speech and noise powers in different time-frequency bins. Hence, in [5]–[7] it has been proposed to use a soft output voice activity detector [8] to weight the speech distortion term by the probability that speech is present and the noise power term by the probability that speech is absent. This principle has been further extended in [9] where again an empirical strategy for the selection of the trade-off parameter based on the instantaneous masking threshold [10] has been proposed. In this paper a systematic method for determining a time-frequency dependent trade-off parameter is established. Aiming at keeping both speech distortion and noise power low, it is proposed to use the parameter that maximizes the curvature of the parametric plot of noise power versus speech distortion. Within the proposed framework, speech distortion and noise

power can be weighted in advance based on what is more important for the speech communication application under consideration or based on perceptually motivated criteria. Theoretical and experimental results show that the proposed parameter yields a more robust filter to commonly occurring signal-to-noise ratio (SNR) estimation errors than the traditional MWF.

## II. THE SDW-MWF

Consider an $M$-channel acoustic system, where the $m$-th microphone signal $Y_m(k,l)$ at frequency index $k$ and time index $l$ consists of a speech component $X_m(k,l)$ and a noise component $V_m(k,l)$, i.e.,

$$Y_m(k,l) = X_m(k,l) + V_m(k,l). \qquad (1)$$

For conciseness the frequency and time indices $k$ and $l$ will be omitted in the remainder of the letter. In vector notation, the $M$-dimensional vector $\mathbf{y}$ of the received microphone signals can be written as $\mathbf{y} = \mathbf{x} + \mathbf{v}$, with $\mathbf{y} = [Y_1 \ \ldots \ Y_M]^T$, and the speech and noise vectors $\mathbf{x}$ and $\mathbf{v}$ similarly defined. Defining the vector of filter coefficients $\mathbf{w}$ similarly as $\mathbf{y}$, the output signal $Z$ is given by

$$Z = \mathbf{w}^H \mathbf{y} = \mathbf{w}^H \mathbf{x} + \mathbf{w}^H \mathbf{v}. \qquad (2)$$

The MWF aims at noise reduction by minimizing the mean-square error between the output signal and the received speech component in the $m$-th microphone, i.e., reference microphone [1]. In the SDW-MWF, a trade-off parameter $\mu$ has been incorporated which allows to trade-off between noise reduction and speech distortion [2], [3]. Assuming that the speech and noise components are uncorrelated, the SDW-MWF cost function can be written as

$$\min_{\mathbf{w}} \ \underbrace{\mathcal{E}\{|\mathbf{w}^H\mathbf{x} - \mathbf{e}_m^T\mathbf{x}|^2\}}_{\psi_\mathbf{x}} + \mu \underbrace{\mathcal{E}\{|\mathbf{w}^H\mathbf{v}|^2\}}_{\psi_\mathbf{v}}, \qquad (3)$$

with $\mathcal{E}$ the expected value operator, $\mathbf{e}_m$ the $M$-dimensional selector vector, i.e., a vector whose $m$-th element is equal to 1 and all other elements are equal to 0, $\psi_\mathbf{x}$ the speech distortion, and $\psi_\mathbf{v}$ the noise power. The filter minimizing the cost function in (3) is given by

$$\mathbf{w} = [\mathbf{R}_\mathbf{x} + \mu\mathbf{R}_\mathbf{v}]^{-1} \mathbf{R}_\mathbf{x}\mathbf{e}_m, \qquad (4)$$

with $\mathbf{R}_\mathbf{x}$ and $\mathbf{R}_\mathbf{v}$ the speech and noise correlation matrices, defined as

$$\mathbf{R}_\mathbf{x} = \mathcal{E}\{\mathbf{x}\mathbf{x}^H\} = P_s\mathbf{a}\mathbf{a}^H \ \text{ and } \ \mathbf{R}_\mathbf{v} = \mathcal{E}\{\mathbf{v}\mathbf{v}^H\}, \qquad (5)$$

where $P_s = \mathcal{E}\{|S|^2\}$ is the power spectral density of the speech source and $\mathbf{a} = [A_1 \ \ldots \ A_M]^T$ is the vector of the acoustic transfer functions. The SDW-MWF in (4) can be decomposed into a Minimum Variance Distortionless Response (MVDR) beamformer $\mathbf{w}_{\text{MVDR}}$ and a single channel Wiener postfilter $G$ applied to the MVDR output [11], i.e.,

$$\mathbf{w} = A_m^* \underbrace{\frac{\mathbf{R_v}^{-1}\mathbf{a}}{\mathbf{a}^H \mathbf{R_v}^{-1}\mathbf{a}}}_{\mathbf{w}_{\text{MVDR}}} \underbrace{\frac{\rho}{\mu + \rho}}_{G}, \qquad (6)$$

with $A_m = \mathbf{e}_m^T \mathbf{a}$ and $\rho$ the SNR at the MVDR output, i.e.,

$$\rho = P_s \mathbf{a}^H \mathbf{R_v}^{-1}\mathbf{a}. \qquad (7)$$

By setting $\mu = 0$ in (6) the SDW-MWF results in the MVDR beamformer, which reduces the noise while keeping the speech component in the reference microphone undistorted, i.e., $\mathbf{w}_{\text{MVDR}}^H \mathbf{a} = A_m$. Using $\mu > 0$, the residual noise at the MVDR output can be further suppressed at the cost of introducing speech distortion. Setting $\mu = 1$, the SDW-MWF results in the MWF. If $\mu > 1$, the noise power is further reduced at the expense of increased speech distortion. Hence the selection of the trade-off parameter in the SDW-MWF is of crucial importance.

## III. TRADE-OFF PARAMETER SELECTION

In the following, the L-curve method [12] used in regularized acoustic multichannel equalization [13] is exploited to derive an analytical expression for a time-frequency dependent trade-off parameter. Furthermore, insights on the resulting postfilter and its robustness to SNR estimation errors will be provided.

### A. Parameter derivation

Using (3), (5), (6), and (7), speech distortion and noise power can be expressed as

$$\psi_{\mathbf{x}} = P_s |A_m|^2 \frac{\mu^2}{(\mu + \rho)^2} \quad \text{and} \quad \psi_{\mathbf{v}} = P_s |A_m|^2 \frac{\rho}{(\mu + \rho)^2}. \quad (8)$$

Fig. 1(a) depicts a typical parametric plot of $\psi_{\mathbf{v}}$ versus $\psi_{\mathbf{x}}$ for 50 trade-off parameters linearly spaced between $10^{-4}$ and 5, with the marked points showing the value of $\mu$ at the given positions. Due to the nature of the arising trade-off between $\psi_{\mathbf{v}}$ and $\psi_{\mathbf{x}}$ (cf. (8)), this parametric plot has an L-shape, with the corner, i.e., point of maximum curvature, located where the filter $\mathbf{w}$ changes in nature from being dominated by large noise power to being dominated by large speech distortion. At the point of maximum curvature, i.e., $\mu = 1$ in the depicted example, speech distortion and noise power are simultaneously minimized. Hence we propose to select the trade-off parameter in the SDW-MWF as the point of maximum curvature of the parametric plot of $\psi_{\mathbf{v}}$ versus $\psi_{\mathbf{x}}$.

Using such a parameter inherently implies that the performance in terms of speech distortion and noise reduction is equally valuable in all frequency bands and for all speech communication systems. However, depending on the system as well as on the frequency band under consideration, one performance measure might be more important than the other.
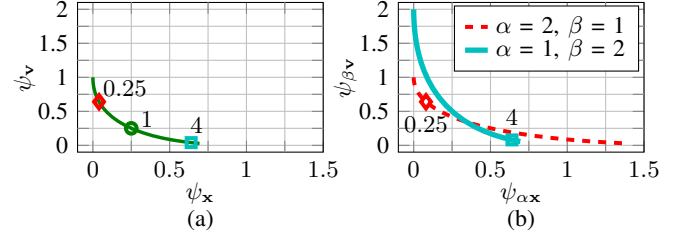


Fig. 1: Typical parametric plot of (a) noise power versus speech distortion and (b) (weighted) noise power versus (weighted) speech distortion

To account for this difference, we propose introducing weighting functions, i.e.,

$$\psi_{\alpha\mathbf{x}} = \alpha \psi_{\mathbf{x}} \quad \text{and} \quad \psi_{\beta\mathbf{v}} = \beta \psi_{\mathbf{v}}, \qquad (9)$$

with $\alpha$ and $\beta$ being the speech distortion and noise power weighting functions, defined e.g. based on psychoacoustically motivated measures such as average masking threshold [10] or speech intelligibility weighting [14] (cf. Section IV). Introducing a weighting function changes the point of maximum curvature as shown in Fig. 1(b), which depicts the parametric plots of $\psi_{\beta\mathbf{v}}$ versus $\psi_{\alpha\mathbf{x}}$ for different weights. Weighting the speech distortion term more yields a lower trade-off parameter, i.e., the point of maximum curvature is at $\mu = 0.25$. On the other hand weihgting the noise power more yields a higher trade-off parameter, i.e., $\mu = 4$. These points are also marked in the original plot in Fig. 1(a), showing the impact of introducing weighting functions on the resulting trade-off. The curvature $\kappa$ of $\psi_{\beta\mathbf{v}}(k)$ versus $\psi_{\alpha\mathbf{x}}(k)$ is defined as [15]

$$\kappa = \frac{\psi'_{\alpha\mathbf{x}}\psi''_{\beta\mathbf{v}} - \psi''_{\alpha\mathbf{x}}\psi'_{\beta\mathbf{v}}}{\{(\psi'_{\alpha\mathbf{x}})^2 + (\psi'_{\beta\mathbf{v}})^2\}^{\frac{3}{2}}}, \qquad (10)$$

where $\{\cdot\}'$ and $\{\cdot\}''$ denote the first and second derivative with respect to $\mu$. The computation of the derivatives yields

$$\psi'_{\alpha\mathbf{x}} = \frac{2\alpha P_s |A_m|^2 \mu\rho}{(\mu + \rho)^3}, \quad \psi''_{\alpha\mathbf{x}} = \frac{2\alpha P_s |A_m|^2 \rho(-2\mu + \rho)}{(\mu + \rho)^4}, \quad (11)$$

$$\psi'_{\beta\mathbf{v}} = \frac{-2\beta P_s |A_m|^2 \rho}{(\mu + \rho)^3}, \quad \psi''_{\beta\mathbf{v}} = \frac{6\beta P_s |A_m|^2 \rho}{(\mu + \rho)^4}. \quad (12)$$

Using the computed derivatives, $\kappa$ can be written as

$$\kappa = \frac{\alpha\beta(\mu + \rho)^3}{2P_s |A_m|^2 \rho(\alpha^2\mu^2 + \beta^2)^{\frac{3}{2}}}. \qquad (13)$$

To compute the point of maximum curvature, (13) is maximized by setting its derivative to 0, i.e.,

$$\kappa' = \frac{3\alpha\beta(\mu + \rho)^2(\beta^2 - \alpha^2\mu\rho)}{2P_s |A_m|^2 \rho(\alpha^2\mu + \beta^2)^{\frac{5}{2}}} = 0. \qquad (14)$$

The solution to (14) yields

$$\mu = \frac{\beta^2}{\alpha^2\rho}. \qquad (15)$$

The point of maximum curvature in (15) only depends on the user-defined weighting functions $\alpha$, $\beta$, and on the SNR at the MVDR output, which can be estimated by employing an SNR

estimator [16]–[20].

In summary, three choices of the trade-off parameter will be analyzed, i.e., the point of maximum curvature when no additional weights are applied $\mu_\mathrm{o} = \frac{1}{\rho}$, when a higher weight is put on the speech distortion term $\mu_\mathrm{sd} = \frac{1}{\alpha^2 \rho}$, and when a higher weight is put on the noise reduction term $\mu_\mathrm{nr} = \frac{\beta^2}{\rho}$. The weighting functions used in this paper are discussed in Section IV.

### B. Robustness to SNR estimation errors

Different choices of the trade-off parameter clearly yield different postfilters. Defining a postfilter to be the optimal one is not feasible, since different postfilters might be optimal for different applications. However, for a given desired postfilter, its robustness to SNR estimation errors can be evaluated. The robustness of a postfilter to SNR estimation errors can be evaluated by analyzing its derivative with respect to the SNR $\rho$ [21], [22]. If the derivative is large, misestimating the SNR will yield a large deviation between the desired and the applied postfilter. On the contrary, a low derivative value will not cause a large fluctuation in the postfilter.

Table I summarizes the postfilters and their derivatives obtained by using $\mu = 1$ and the proposed trade-off parameters. These different postfilters and the respective derivatives are also depicted in Figs. 2(a) and 2(b) for $-10$ dB $\leq \rho \leq 10$ dB and for weighting parameters $\alpha = 2$ and $\beta = 2$. Fig. 2(a) shows that for $\rho < 0$ dB, using $\mu_\mathrm{o}$ yields a more aggressive postfilter than using $\mu = 1$, i.e., a higher noise reduction performance as well as a higher speech distortion. For $\rho > 0$ dB the proposed postfilter is less aggressive, yielding a lower speech distortion as well as a lower noise reduction. Weighting the speech distortion or the noise power more shifts the postfilter to the left or right respectively. The derivatives depicted in Fig. 2(b) show that using $\mu_\mathrm{o}$ is more robust to SNR estimation errors than using $\mu = 1$ for $\rho < -4.6$ dB and $\rho > 4.6$ dB. Furthermore, using $\mu_\mathrm{sd}$ yields a generally more sensitive postfilter, whereas using $\mu_\mathrm{nr}$ yields a generally more robust postfilter. As the SNR increases, it can be seen that all postfilters exhibit a low sensitivity to SNR estimation errors.

## IV. EXPERIMENTAL RESULTS

In this section, the MWF is compared to the SDW-MWF using the different proposed trade-off parameters. The weights $\alpha$ and $\beta$ are determined based on the speech intelligibility index [14], where each frequency band is assigned an importance weight to reflect how much a performance improvement

TABLE I: Postfilter and derivative for different choices of the trade-off parameter

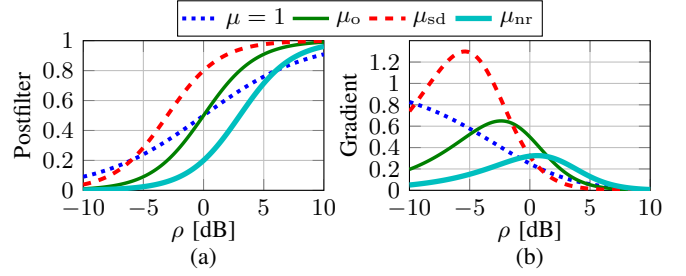| Trade-off parameter | Postfilter | Derivative |
|---|---|---|
| MWF with $\mu = 1$ | $\frac{\rho}{1+\rho}$ | $\frac{1}{(1+\rho)^2}$ |
| SDW-MWF with $\mu_\mathrm{o} = \frac{1}{\rho}$ | $\frac{\rho^2}{1+\rho^2}$ | $\frac{2\rho}{(1+\rho^2)^2}$ |
| SDW-MWF with $\mu_\mathrm{sd} = \frac{1}{\alpha^2\rho}$ | $\frac{\rho^2}{\alpha^{-2}+\rho^2}$ | $\frac{2\alpha^{-2}\rho}{(\alpha^{-2}+\rho^2)^2}$ |
| SDW-MWF with $\mu_\mathrm{nr} = \frac{\beta^2}{\rho}$ | $\frac{\rho^2}{\beta^2+\rho^2}$ | $\frac{2\beta^2\rho}{(\beta^2+\rho^2)^2}$ |



Fig. 2: (a) Postfilter and (b) postfilter derivative as a function of the SNR $\rho$ for different choices of the trade-off parameter

in that band contributes to the overall speech intelligibility improvement. In this work, the intelligibility indices are arbitrarily scaled between 1 and 10. By setting $\alpha$ or $\beta$ to the scaled intelligibility indices, speech distortion or noise power are weighted more in more important frequency bands.

### A. Setup and performance measures

We have considered a scenario with $M = 2$ microphones placed 5 cm apart and a speech source located at $0°$. The speech components were generated using measured room impulse responses with reverberation time $T_{60} \approx 450$ ms [23]. The noise components consisted of spatially diffuse nonstationary babble speech generated using [24]. Several broadband input SNRs ranging from $-10$ dB to 10 dB have been evaluated. The MVDR filter coefficients have been computed using anechoic steering vectors assuming knowledge of the direction-of-arrival of the speech source and a theoretically diffuse noise correlation matrix. The signals were processed at a sampling frequency $f_s = 16$ kHz using a weighted overlap-add framework with a block size of 512 samples and an overlap of $50\%$ between successive blocks. The SNR at the MVDR output has been estimated using the minimum statistics-based estimator [16] combined with the cepstral smoothing approach [20]. The true SNR has also been computed using the periodograms of the speech and noise component. The minimum gain for all postfilters has been set to $-10$ dB.

The performance is evaluated in terms of segmental noise reduction ($\mathrm{NR}_\mathrm{seg}$) [25], segmental speech SNR ($\mathrm{SSNR}_\mathrm{seg}$) reflecting the speech preservation performance [25], and the segmental SNR improvement ($\Delta\mathrm{SNR}_\mathrm{seg}$) reflecting the overall improvement in terms of both noise reduction and speech preservation. To evaluate the robustness of the postfilters to SNR estimation errors, the mismatch between the postfilter $G$ computed based on the true SNR and the postfilter $\hat{G}$ computed based on the estimated SNR has been calculated as

$$M_G = 10 \log_{10} \sum_{l=1}^{L} \sum_{k=1}^{K} |G(k,l) - \hat{G}(k,l)|^2,$$

with $L$ and $K$ being the total number of time and frequency bins respectively.

### B. Results

The performance for different trade-off parameters when the SNR at the MVDR output is known is depicted in Fig. 3. As
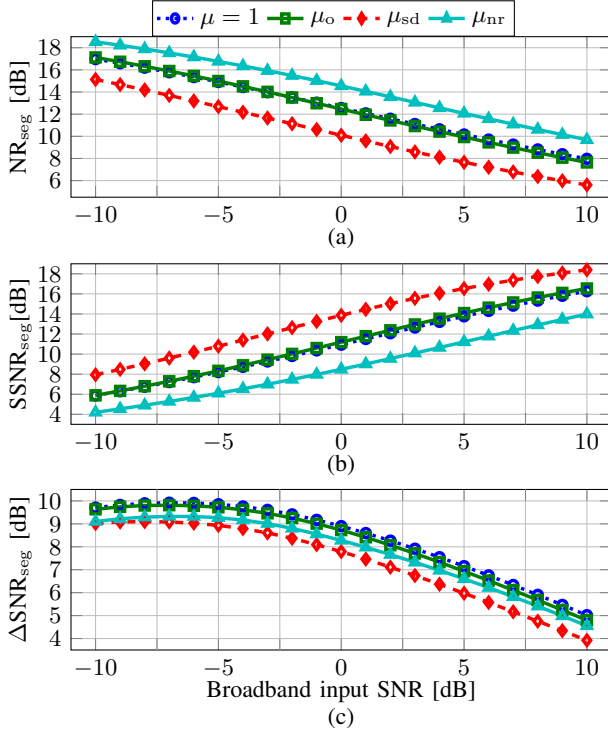
Fig. 3: (a) Segmental noise reduction, (b) segmental speech SNR, and (c) segmental SNR improvement based on the true SNR for different trade-off parameters
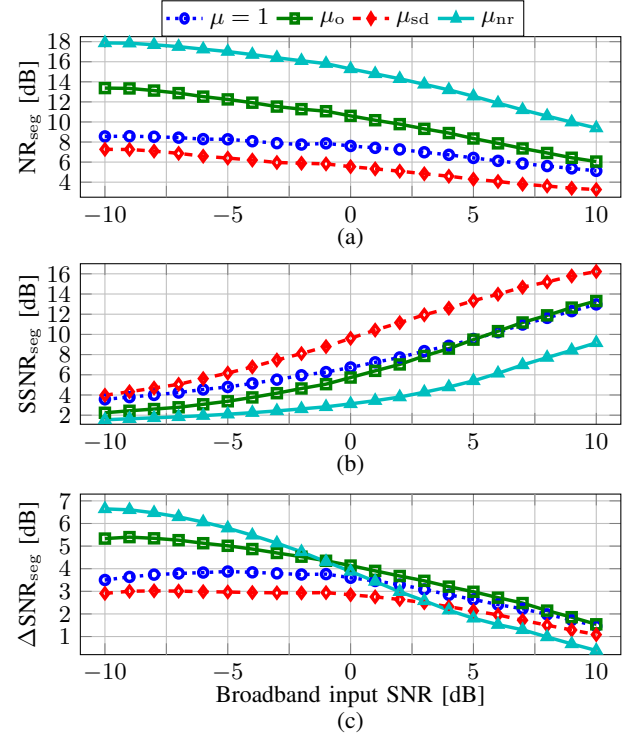


Fig. 4: (a) Segmental noise reduction, (b) segmental speech SNR, and (c) segmental SNR improvement based on the estimated SNR for different trade-off parameters

TABLE II: Postfilter mismatch for different trade-off parameters averaged over all input SNRs

| Trade-off parameter | $\mu = 1$ | $\mu_{\mathrm{o}}$ | $\mu_{\mathrm{sd}}$ | $\mu_{\mathrm{nr}}$ |
|---|---|---|---|---|
| Postfilter mismatch $M_G$ [dB] | 6.77 | 6.39 | 8.32 | 4.84 |

illustrated in Figs. 3(a) and 3(b), using $\mu = 1$ or $\mu_{\mathrm{o}}$ yields a very similar performance in terms of noise reduction and speech SNR. As a result, the overall $\Delta\mathrm{SNR}_{\mathrm{seg}}$ depicted in Fig. 3(c) achieved by both the MWF and the SDW-MWF is very similar. Furthermore, using $\mu_{\mathrm{sd}}$ yields as expected a loss in noise reduction and a gain in speech SNR, whereas using $\mu_{\mathrm{nr}}$ increases the noise reduction performance at the cost of a decrease in speech SNR. Finally, Fig. 3(c) shows that when the true SNR is known, weighting the speech distortion or noise power more yields a lower $\Delta\mathrm{SNR}_{\mathrm{seg}}$ than when no weights are applied.

Fig. 4 depicts the performance in the realistic scenario of using an SNR estimator for computing the postfilters. As illustrated, using $\mu_{\mathrm{o}}$ instead of $\mu = 1$ yields a significantly higher noise reduction performance in particular for low input SNRs without significantly worsening the speech SNR. This performance increase is reflected in achieving a higher $\Delta\mathrm{SNR}_{\mathrm{seg}}$ over all input SNRs as depicted in Fig. 4(c). Since the proposed trade-off parameter outperforms $\mu = 1$ only when the SNR is estimated, i.e., in the likely presence of SNR estimation errors, this performance gain can be explained by the higher robustness of the resulting postfilter.

Weighting the speech distortion or noise power terms yields as expected similar results as before. Since at low input SNRs by using $\mu_{\mathrm{nr}}$ the increase in noise reduction does not cause a significant decrease in speech SNR, using this trade-off parameter yields the highest $\Delta\mathrm{SNR}_{\mathrm{seg}}$ in this input SNR region. However, as the input SNR increases also the loss in speech SNR increases, leading to $\mu_{\mathrm{nr}}$ yielding the lowest

$\Delta\mathrm{SNR}_{\mathrm{seg}}$. The latter can also be explained by the previous robustness analysis suggesting that at higher input SNRs, the postfilter derived using $\mu_{\mathrm{nr}}$ is the most sensitive one to SNR estimation errors.

Finally, Table II presents the mismatch between the desired and estimated postfilters averaged over all input SNRs. As expected from the theoritcal robustness analysis, using $\mu_{\mathrm{nr}}$ yields the lowest mismatch value whereas using $\mu_{\mathrm{sd}}$ reults in the highest mismatch value. Furthermore, the proposed trade-off parameter $\mu_{\mathrm{o}}$ results in a more robust postfilter than $\mu = 1$.

## V. CONCLUSION

In this letter it has been proposed to compute the trade-off parameter in the SDW-MWF as the one that maximizes the curvature of the parametric plot of noise power versus speech distortion. Based on what is more important for the speech communication application or on perceptually motivated criteria, speech distortion and noise power can also be weighted in advance. Theoretical and experimental results have shown that the proposed parameter results in a more robust filter to SNR estimation errors than the traditional MWF, significantly improving the performance.

## REFERENCES

[1] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic beamforming for hearing aid applications," in *Handbook on Array Processing and Sensor Networks*, S. Haykin and K. J. R. Liu, Eds. USA: Wiley, 2010.

[2] A. Spriet, M. Moonen, and J. Wouters, "Spatially pre-processed speech distortion weighted multi-channel Wiener filtering for noise reduction," *Signal Processing*, vol. 84, no. 12, pp. 2367–2387, Dec. 2004.

[3] S. Doclo, A. Spriet, J. Wouters, and M. Moonen, "Frequency-domain criterion for the speech distortion weighted multichannel Wiener filter for robust noise reduction," in *Speech Communication, special issue on Speech Enhancement*, vol. 49, Jul.-Aug. 2007, pp. 636–656.

[4] B. Cornelis, "Design and evaluation of noise reduction techniques for binaural hearing aids," Ph.D. dissertation, Katholieke Universiteit Leuven, 2011.

[5] K. Ngo, A. Spriet, M. Moonen, J. Wouters, and S. H. Jensen, "Variable speech distortion weighted multichannel Wiener filter based on soft output voice activity detection for noise reduction in hearing aids," in *Proc. International Workshop on Acoustic Echo and Noise Control*, Seattle, Washington USA, Sep. 2008.

[6] K. Ngo, M. Moonen, S. H. Jensen, and J. Wouters, "A flexible speech distortion weighted multi-channel Wiener filter for noise reduction in hearing aids," in *Proc. International Conference on Acoustics, Speech and Signal Processing*, Prague, Czech Republic, May 2011, pp. 2528–2531.

[7] K. Ngo, A. Spriet, M. Moonen, J. Wouters, and S. H. Jensen, "Incorporating the conditional speech presence probability in multi-channel Wiener filter based noise reduction in hearing aids," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, Jul. 2009.

[8] S. Gazor and W. Zhang, "A soft voice activity detector based on a Laplacian-Gaussian model," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 5, pp. 498–505, Sep. 2003.

[9] B. Defraene, K. Ngo, T. van Waterschoot, M. Diehl, and M. Moonen, "A psychoacoustically motivated speech distortion weighted multi-channel Wiener filter for noise reduction," in *Proc. International Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, Mar. 2012, pp. 4637–4640.

[10] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proc. of the IEEE*, vol. 88, no. 4, pp. 451–515, Apr. 2000.

[11] K. U. Simmer, J. Bitzer, and C. Marro, "Post-filtering techniques," in *Microphone Arrays*, M. Brandstein and D. Ward, Eds. Berlin, Germany: Springer, 2001.

[12] P. Hansen and D. O'Leary, "The use of the L-Curve in the regularization of discrete ill-posed problems," *SIAM Journal on Scientific Computing*, vol. 14, no. 6, pp. 1487–1503, Jan. 1993.

[13] I. Kodrasi, S. Goetze, and S. Doclo, "Regularization for partial multi-channel equalization for speech dereverberation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 9, pp. 1879–1890, Sep. 2013.

[14] *American National Standard Methods for Calculation of the Speech Intelligibility Index*, Acoustical Society of America, ANSI S3.5-1997, Jun. 1997.

[15] S. Sternberg, *Curvature in Mathematics and Physics*. USA: Dover Pubications, 2012.

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

[17] I. Cohen, "Noise spectrum estimation in adverse environments: improved minima controlled recursive averaging," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 5, pp. 466–475, Sep. 2003.

[18] T. Gerkmann and R. Hendriks, "Unbiased MMSE-based noise power estimation with low complexity and low tracking delay," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 4, pp. 1383–1393, May 2012.

[19] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 32, no. 6, pp. 1109–1121, Dec. 1984.

[20] C. Breithaupt, T. Gerkmann, and R. Martin, "A novel a priori SNR estimation approach based on selective cepstro-temporal smoothing," in *Proc. International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, USA, Mar. 2008, pp. 4897–4900.

[21] C. Demiroglu, "Multisensor segmentation-based noise suppression for intelligibility improvement in MELP coders," Ph.D. dissertation, Georgia Institute of Technology, 2006.

[22] P. S. Whitehead and D. V. Anderson, "Robust Bayesian analysis applied to Wiener filtering of speech," in *Proc. International Conference on Acoustics, Speech and Signal Processing*, Prague, Czech Republic, May 2011, pp. 5080–5083.

[23] J. Wen, N. D. Gaubitch, E. A. P. Habets, T. Myatt, and P. A. Naylor, "Evaluation of speech dereverberation algorithms using the MARDY database," in *Proc. International Workshop on Acoustic Echo and Noise Control*, Paris, France, Sep. 2006.

[24] E. A. P. Habets, I. Cohen, and S. Gannot, "Generating nonstationary multisensor signals under a spatial coherence constraint," *Journal of the Acoustical Society of America*, vol. 124, no. 5, pp. 2911–2917, Nov. 2008.

[25] T. Lotter and P. Vary, "Speech enhancement by MAP spectral amplitude estimation using a super-Gaussian speech model," *EURASIP Journal on Applied Signal Processing*, vol. 2005, pp. 1110–1126, Jan. 2005.