NEW INSIGHTS INTO NON-CAUSAL MULTICHANNEL LINEAR FILTERING FOR NOISE REDUCTION

Mehrez Souden, Jacob Benesty, and Sofiène Affes
INRS-ÉMT, 800, de la Gauchetière Ouest, Suite 6900, Montréal, H5A 1K6, Qc, Canada.
{souden,benesty,affes}@emt.inrs.ca

ABSTRACT

We investigate a general framework for noise reduction which consists in controlling the level of signal distortion while reducing the level of noise. A parameterized non-causal filter that allows for tuning the signal distortion and noise reduction inversely is obtained and is referred to as parameterized multichannel non-causal Wiener filter (PMWF) herein. The same optimization problem leads to the minimum variance distortionless response (MVDR) as a particular case of the PMWF. In contrast to earlier works, the proposed expressions of the PMWF and MVDR are simplified and require the knowledge of the speech and noise statistics only. To rigorously quantify the gains and losses when using these filters, we establish simplified closed-form expressions for three measures, namely, the signal distortion index, the noise reduction factor, and the output signal-to-noise ratio (SNR), and highlight the tradeoff between noise reduction and speech distortion in the multichannel case.

Index Terms— Multichannel noise reduction, Wiener filter, minimum variance distortionless response, speech distortion.

1. INTRODUCTION

Speech signals perceived by communication devices are generally corrupted by background noise or interference from other competing sources. To cope with this issue, several noise reduction approaches have been developed so far including [1]-[12]. In contrast to the single-channel based techniques, microphone-array based processing is promising since it takes advantage of the spatial aperture in addition to the classical frequency and time dimensions.

When compared to their time-domain counterparts, frequency-domain approaches for noise reduction are generally preferred because each frequency bin can be processed apart from the others. This allows for easier calculations and interesting relationships can be found. The well known multichannel Wiener filter is optimal in the mean-square error sense. However, it can introduce undesirable distortions to the speech [1]. Parameterized multichannel filtering allows for tuning the signal distortion and noise reduction [1, 2, 3] while forcing a distortionless response when reducing the noise power leads to the MVDR filter [2, 4, 5]. In [6, 7], the parametrization of the adaptive noise canceler in the standard Griffith and Jim generalized sidelobe canceler [8] has been shown to reduce the speech distortions by controlling the signal leakage due to the system model errors (microphones mismatch, spatial aliasing, reverberation, etc). In [9], a general cost function combining system model prior and estimated model terms was considered. However, system-model-based prior is known to deteriorate the performance of the filters in the presence of system model errors.

In this paper, we focus on a general framework that does not require any preprocessing and consists in minimizing the power of the noise captured by the microphones and filtered by the filter of interest while controlling the desired signals distortion which is defined as the dissimilarity between one noise-free reference microphone signal and the overall filtered noise-free microphone signals. This approach, albeit essentially equivalent to the traditional way of reducing the signal distortion subject to some constraint on the output noise, is more intuitive in the sense that it allows to see the connection with the MVDR. By doing so, we develop a new simplified expression for the PMWF that depends on the noise and speech statistics only. The second contribution of this work consists in analytically investigating the tradeoff between noise reduction and speech distortion in the multichannel case and studying the effect of some key parameters, namely, the input SNR and number of microphones on the performance of these filters.

1.1. Data Model

We consider the following frequency-domain representation of the data model [2]:

$$Y_n(j\omega) = G_n(j\omega)S(j\omega) + V_n(j\omega) = X_n(j\omega) + V_n(j\omega), \quad (1)$$

where $Y_n(j\omega)$, $G_n(j\omega)$, $S(j\omega)$, and $V_n(j\omega)$ are the discrete-time Fourier transforms (DTFT's) of the nth microphone output, the channel impulse response between the source and the nth microphone, the desired speech signal, and the additive noise, respectively.

Our aim is to reduce the noise and recover one of the signal components, say $X_{n_0}(j\omega)$, $n_0 \in \{1, ..., N\}$, the best way we can (along some criteria to be defined later) by applying a linear filter $\mathbf{h}_{n_0}(j\omega)$ to the overall observation vector $\mathbf{y}(j\omega) = [Y_1(j\omega)\ Y_2(j\omega)\ \cdots\ Y_N(j\omega)]^T$. The output of this filter is:

$$Z(j\omega) = \mathbf{h}_{n_0}^H(j\omega)\mathbf{y}(j\omega) = \underbrace{\mathbf{h}_{n_0}^H(j\omega)\mathbf{x}(j\omega)}_{D_{n_0}(j\omega)} + \underbrace{\mathbf{h}_{n_0}^H(j\omega)\mathbf{v}(j\omega)}_{\nu_{n_0}(j\omega)}, \quad (2)$$

where $\mathbf{x}(j\omega)$ and $\mathbf{v}(j\omega)$ are defined like $\mathbf{y}(j\omega)$. $D_{n_0}(j\omega)$ and $\nu_{n_0}(j\omega)$ are the speech and noise components at the output of $\mathbf{h}_{n_0}(j\omega)$, respectively. We also define $\mathbf{g}(j\omega) = [G_1(j\omega)\ G_2(j\omega)\ \cdots\ G_N(j\omega)]^T$.

1.2. Definitions

We use the definitions given in [2]. For completeness, we specify some of them here. First, we define the power spectrum density (PSD) matrix of a vector $\mathbf{a}(j\omega)$ as $\mathbf{\Phi}_{aa}(j\omega) = E\left\{\mathbf{a}(j\omega)\mathbf{a}^H(j\omega)\right\}$. Since we are taking the n_0th noise-free microphone signal as a reference, we define the local input SNR as $\mathrm{SNR}(\omega) = \frac{\phi_{x_{n_0}x_{n_0}}(\omega)}{\phi_{v_{n_0}v_{n_0}}(\omega)}$, where $\phi_{aa}(\omega) = E\left\{|A(j\omega)|^2\right\}$ is the PSD of $a(t)$ [having $A(j\omega)$ as DTFT]. Recall that our aim is to have an optimal estimate of $X_{n_0}(j\omega)$ at the output of the linear filter $\mathbf{h}_{n_0}(j\omega)$. Hence, we define the error signals $\mathcal{E}_{x,n_0}(j\omega) = X_{n_0}(j\omega) - D_{n_0}(j\omega)$ and $\mathcal{E}_{v,n_0}(j\omega) = \nu_{n_0}(j\omega)$. We obtain:

$$\mathcal{E}_{x,n_0}(j\omega) = [\mathbf{u}_{n_0} - \mathbf{h}_{n_0}(j\omega)]^H \mathbf{x}(j\omega), \quad (3)$$

$$\mathcal{E}_{v,n_0}(j\omega) = \mathbf{h}_{n_0}^H(j\omega)\mathbf{v}(j\omega), \quad (4)$$

where $\mathbf{u}_{n_0}$ is an N-dimensional vector with entries being all zeroes except the n_0th one which equals 1. We use the definitions of the local signal distortion index, $\upsilon_{\rm sd}\left[\mathbf{h}_{n_0}(j\omega)\right]$, and local noise reduction factor, $\xi_{\rm nr}\left[\mathbf{h}_{n_0}(j\omega)\right]$, initially proposed in [2] which are directly derived from (3) and (4) as:

$$\upsilon_{\rm sd}\left[\mathbf{h}_{n_0}(j\omega)\right] = \frac{\left[\mathbf{u}_{n_0}-\mathbf{h}_{n_0}(j\omega)\right]^H \mathbf{\Phi}_{xx}(j\omega)\left[\mathbf{u}_{n_0}-\mathbf{h}_{n_0}(j\omega)\right]}{\phi_{x_{n_0}x_{n_0}}(\omega)}, \tag{5}$$

$$\xi_{\rm nr}\left[\mathbf{h}_{n_0}(j\omega)\right] = \frac{\phi_{v_{n_0}v_{n_0}}(\omega)}{\mathbf{h}_{n_0}^H(j\omega)\mathbf{\Phi}_{vv}(j\omega)\mathbf{h}_{n_0}(j\omega)}. \tag{6}$$

Finally, we define the local output SNR as [2]:

$$\mathrm{SNR}_{\rm o}\left[\mathbf{h}_{n_0}(j\omega)\right] = \frac{\mathbf{h}_{n_0}^H(j\omega)\mathbf{\Phi}_{xx}(j\omega)\mathbf{h}_{n_0}(j\omega)}{\mathbf{h}_{n_0}^H(j\omega)\mathbf{\Phi}_{vv}(j\omega)\mathbf{h}_{n_0}(j\omega)}. \tag{7}$$

Similar definitions were first proposed in [10] to study the tradeoff between signal distortion and noise reduction with the time-domain single channel Wiener filter. Herein, we use the above definitions to study this tradeoff achieved by the PMWF. As a rule of thumb, noise reduction comes at the price of speech distortion. This fact is well known in the single-channel case where *any* noise reduction leads to speech distortion [2, 10, 11]. In the multichannel case, however, noise reduction can be achieved with no speech distortion in theory [2, 5]. The effect of preserving the speech signal on the noise reduction is quantified herein. In what follows, we start by presenting the general framework for noise reduction under low speech distortion constraint and develop new expressions for the PMWF and MVDR. Then, we study the tradeoff between noise reduction and speech distortion in the multichannel case. Note that we assume that the noise is stationary enough [12] so that the noise PSD matrix $\mathbf{\Phi}_{vv}(j\omega)$ can be estimated during the periods of silence of the target speech. Then, we can also obtain $\mathbf{\Phi}_{xx}(j\omega)=\mathbf{\Phi}_{yy}(j\omega)-\mathbf{\Phi}_{vv}(j\omega)$.

2. GENERAL FRAMEWORK FOR NOISE REDUCTION

2.1. Parameterized Multichannel Wiener Filter

In contrast to the classical multichannel non-causal Wiener filter, the PMWF is able to achieve a tradeoff between noise reduction and signal distortion. Traditionally, parameterized filters allowing the tuning of the levels of residual noise and signal distortion are derived by minimizing the signal distortion under the constraint of an upper bound on the residual noise power [1, 2, 12]. Herein, we derive our parameterized filter by minimizing the output noise energy under some constraint on the level of the output signal distortion. Although both approaches are equivalent, this formulation will better show us the link with the MVDR that we will investigate next. Specifically, we consider this optimization problem:

$$\min_{\mathbf{h}_{n_0}(j\omega)} E\left\{|\mathcal{E}_{v,n_0}(j\omega)|^2\right\}, \text{ sub. to } E\left\{|\mathcal{E}_{x,n_0}(j\omega)|^2\right\} \leq \sigma^2(\omega), \tag{8}$$

where $\sigma^2(\omega)$ represents the maximum allowable local signal distortion. Setting the derivative of the Lagrangian associated with (8) with respect to $\mathbf{h}_{n_0}^H(j\omega)$ to zero, we obtain the PMWF:

$$\mathbf{h}_{{\rm W}\beta,n_0}(j\omega)=\left[\mathbf{\Phi}_{xx}(j\omega)+\beta\mathbf{\Phi}_{vv}(j\omega)\right]^{-1}\mathbf{\Phi}_{xx}(j\omega)\mathbf{u}_{n_0}, \tag{9}$$

where $\beta=\frac{1}{\gamma}$ is a positive-valued factor that allows for tuning the signal distortion and noise reduction at the output of $\mathbf{h}_{{\rm W}\beta,n_0}(j\omega)$ and γ is the Lagrange multiplier associated with (8). The relationship between β and $\sigma(\omega)$ will be discussed in Section 3. Note also that (9) can be found in earlier works such as [2]. The time-domain equivalent of this expression can be found in [6, 12]. Unfortunately, this expression does not offer enough flexibility to investigate the tradeoff between noise reduction and signal distortion in the multichannel case and to establish the link with other optimal filters such as the MVDR. Herein, we propose a more simplified form.

***Result 1*:** For $\beta\neq 0$, the PMWF can be written as :

$$\mathbf{h}_{{\rm W}\beta,n_0}(j\omega) = \frac{\mathbf{\Phi}_{vv}^{-1}(j\omega)\mathbf{\Phi}_{xx}(j\omega)}{\beta+\lambda(\omega)}\mathbf{u}_{n_0}, \tag{10}$$

where $\lambda(\omega)=\mathrm{tr}\left\{\mathbf{\Phi}_{vv}^{-1}(j\omega)\mathbf{\Phi}_{xx}(j\omega)\right\}$.

***Proof*:** see Appendix I.

This parameterized filter is denoted as PMWF-β in the sequel. Its expression is notable since it will allow us to show that the MVDR is a particular case of (10). In addition, new simplified expressions of the performance measures will be derived next using (10). These measures will give us good insights into the behavior of the PMWF-β in terms of signal distortion and noise reduction. The classical Multichannel Wiener is obtained when $\beta=1$ (PMWF-1). Next, we show that the case $\beta=0$ corresponds to the MVDR. This generalization is not straightforward and an explicit formulation of the problem is required to establish its link with (8) and (10).

2.2. Minimum Variance Distortionless Response

The MVDR consists in reducing the noise under the constraint of no distortion of $X_{n_0}(j\omega)$. This can be formulated as [2, 4]:

$$\min_{\mathbf{h}_{n_0}(j\omega)} E\left\{|\mathcal{E}_{v,n_0}(j\omega)|^2\right\}, \text{ sub. to } E\left\{|\mathcal{E}_{x,n_0}(j\omega)|^2\right\}=0. \tag{11}$$

Clearly, this optimization problem is a particular case of (8) with $\sigma(\omega)=0$. Rewriting the constraint in (11) as a function of $\mathbf{h}_{n_0}(j\omega)$ and $\mathbf{g}(j\omega)$ only and setting the derivative of the Lagrangian associated with (11) with respect to $\mathbf{h}_{n_0}^H(j\omega)$ to zero lead to [2, 4]:

$$\mathbf{h}_{{\rm MVDR},n_0}(j\omega)=G_{n_0}^*(j\omega)\frac{\mathbf{\Phi}_{vv}^{-1}(j\omega)\mathbf{g}(j\omega)}{\mathbf{g}^H(j\omega)\mathbf{\Phi}_{vv}^{-1}(j\omega)\mathbf{g}(j\omega)}. \tag{12}$$

Multiplying and dividing the second term in (12) by $\phi_{ss}(\omega)$ and knowing that $\mathbf{g}^H(j\omega)\mathbf{\Phi}_{vv}^{-1}(j\omega)\mathbf{g}(j\omega) = \mathrm{tr}\left[\mathbf{\Phi}_{vv}^{-1}(j\omega)\mathbf{g}(j\omega)\mathbf{g}^H(j\omega)\right]$, we get rid of the explicit dependence of (12) on the channel transfer functions and obtain [2]:

$$\mathbf{h}_{{\rm MVDR},n_0}(j\omega)=\frac{\mathbf{\Phi}_{vv}^{-1}(j\omega)\mathbf{\Phi}_{xx}(j\omega)}{\lambda(\omega)}\mathbf{u}_{n_0}. \tag{13}$$

Since we are literally solving a particular case of the general problem defined in (8), we see from (10) and (13) that the MVDR is nothing but the PMWF with $\beta=0$ (i.e., PMWF-0). More importantly, note that in (10) and (13), there is no need to know neither the channel transfer functions nor their ratios in contrast to [4, 9]. Indeed, these new expressions are explicitly dependent on the speech and noise statistics only.

3. PERFORMANCE ANALYSIS

Our analysis is based on the performance measures defined in Subsection 1.2. For completeness, we investigate the general case of the PMWF-β. Plugging (10) into (5), (6), and (7), we obtain:

***Result 2*:**

$$\upsilon_{\rm sd}\left[\mathbf{h}_{{\rm W}\beta,n_0}(j\omega)\right] = \frac{\beta^2}{\left[\beta+\lambda(\omega)\right]^2}, \tag{14}$$

$$\xi_{\rm nr}\left[\mathbf{h}_{{\rm W}\beta,n_0}(j\omega)\right] = \frac{\left[\beta+\lambda(\omega)\right]^2}{\mathrm{SNR}(\omega)\lambda(\omega)}, \tag{15}$$

$$\mathrm{SNR}_{\rm o}\left[\mathbf{h}_{{\rm W}\beta,n_0}(j\omega)\right] = \lambda(\omega). \tag{16}$$

***Proof*:** see Appendix II .

Clearly, $\upsilon_{\rm sd}\left[\mathbf{h}_{{\rm W}\beta,n_0}(j\omega)\right]$ and $\xi_{\rm nr}\left[\mathbf{h}_{{\rm W}\beta,n_0}(j\omega)\right]$ are increasing with respect to β. In Fig. 1 (a), we plot the theoretical variations of these performance measures with respect to β in the case of a white noise. We notice that the tradeoff between noise reduction and signal distortion has to be made in the multichannel case too. However, noise reduction can be achieved [i.e., $\xi_{\rm nr}\left[\mathbf{h}_{{\rm W}\beta,n_0}(j\omega)\right]$ can be higher than 1] even when there is no signal distortion [i.e., $\upsilon_{\rm sd}\left[\mathbf{h}_{{\rm W}\beta,n_0}(j\omega)\right]=0$] which is not possible in the single-channel case [2, 10]. This is observed with the MVDR filter which preserves the speech and reduces the noise. Thanks to the new simplified expression (14) and our new formulation (8), one can also find the relationship between $\sigma(\omega)$ defined in (8) and β. By considering (5), the constraint in (8) is equivalent to $\upsilon_{\rm sd}(\omega)\leq\tilde{\sigma}^2(\omega)$ with $\tilde{\sigma}^2(\omega)=\frac{\sigma^2(\omega)}{\phi_{x_{n_0}x_{n_0}}(\omega)}$. Using (14) and this result, we obtain the relationship:

$$\beta\leq\lambda(\omega)\frac{\tilde{\sigma}(\omega)}{1-\tilde{\sigma}(\omega)}. \tag{17}$$

For a given $\tilde{\sigma}(\omega)<1$, one can choose β and vice versa. Now, to better understand the gains in terms of signal distortion and noise reduction when using multiple microphones, we investigate in the sequel the particular case of spatially incoherent[1] noise components (identically distributed).

Particular Case-Spatially Incoherent Noise: in this case, it can be easily shown that:

$$\lambda(\omega)={\rm SNR}(\omega)\left[1+R_{n_0}(\omega)\right], \tag{18}$$

where $R_{n_0}(\omega)=\sum_{n=1,n\neq n_0}^{N}\frac{|G_n(j\omega)|^2}{|G_{n_0}(j\omega)|^2}$.

Plugging (18) into (14)-(16), we draw out four important conclusions: *(i)* For an invariant environment, increasing the number of microphones amounts to adding more diversity (other propagation paths), increasing, thereby, $R_{n_0}(\omega)$. Hence, when the number of microphones increases, the performance of the PMWF-β is enhanced (decreasing signal distortion and increasing noise reduction and output SNR). A similar improvement is observed when the input SNR is increased. *(ii)* As the input SNR or the number of microphones increases, the PMWF-β ($\beta\neq0$) tends to have closer performance to the MVDR in terms of noise reduction and signal distortion. For example, at a given frequency ω, the Wiener and MVDR filters are related up to a scalar coefficient: $\mathbf{h}_{{\rm MVDR},n_0}(j\omega)=\frac{1+\lambda(\omega)}{\lambda(\omega)}\mathbf{h}_{{\rm W}1,n_0}(j\omega)$. Fig. 1 (b) depicts the theoretical variations of the scalar coefficient relating both filters at a given frequency with respect to the input SNR and the number of microphones. Clearly, both filters seem to have similar effects on the input signals when the number of microphones and/or the SNR is sufficiently high. The major differences between both filters can be noticed at low SNR and small N. *(iii)* Since ${\rm SNR}(\omega)=\frac{\phi_{ss}(\omega)}{\phi_{vv}(\omega)}|G_{n_0}(j\omega)|^2$, choosing the signal microphone experiencing the highest input SNR leads to the best performances. *(iv)* The same performance measures corresponding to the non-causal single-channel Wiener filter have been derived in [2]. Those results correspond to the particular case above $N=1$. Thus, the multichannel case theoretically provides better performances than the single-channel processing.

Finally, it is important to mention that in [3], we provided an analytical proof of the SNR improvement at the output of the PMWF-β confirming, thereby, its theoretical effectiveness in speech enhancement regardless of the choice of β (even for the MVDR).

[1]The case of spatially coherent noise was omitted herein due to lack of space. See [3] for more details.

4. NUMERICAL EXAMPLES

To show the tradeoff between signal distortion and noise reduction in the multichannel case, we investigate the performance of the filters PMWF-1 (i.e., Wiener), PMWF-10 ($\beta=10$), and PMWF-0 (i.e., the MVDR). Without loss of generality, we will take the first microphone $n_0=1$ as a reference. The results are presented in terms of the signal distortion index and the output SNR (other performance measures were also tested and similar conclusions were reached). In the investigated scenarios, the speaker is located in a reverberant room in addition to a uniform linear array of N (varied between 2 and 10) microphones. The microphones spacing is $\Delta=0.2$ m. The source is a 2-minutes long female speech sampled at 8 kHz and located at 1.3 m away from the first microphone (taken as a reference herein). The image method [13] was used to generate the impulse responses (with a reverberation time $T_{60}\approx270$ ms) which are convolved with the speech signal before adding a computer generated white Gaussian noise with a long-term input SNR $=0$ dB. The signals are cut into 75% overlapping frames of duration 256 ms each. We are interested in assessing the performance of the filters developed above and the different tradeoffs. Hence, we put aside the problem of noise statistics estimation and suppose that the noise samples are known for any processed data frame as in [1]. For further details on noise estimation, we refer the readers to [11]. All statistics are estimated in a batch mode using the Welch's modified periodogram.

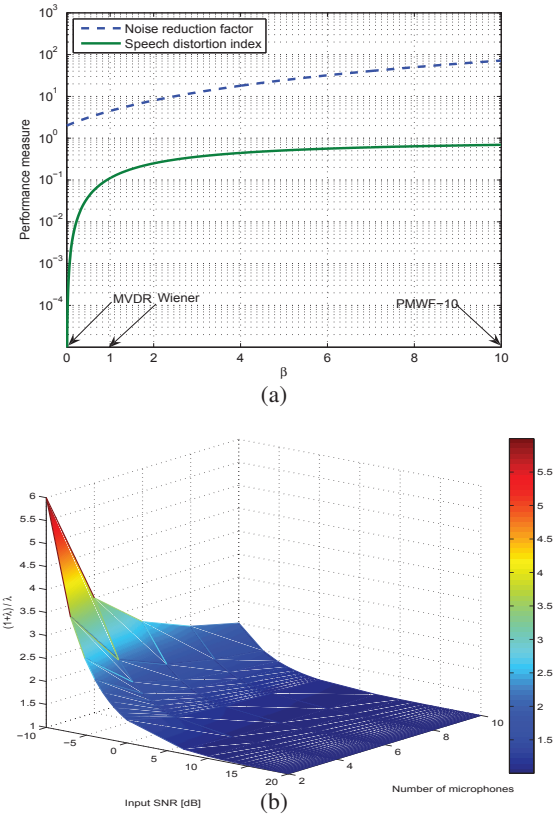

Fig. 1. Theoretical analysis: (a) output signal distortion index and noise reduction factor vs. β; $N=2$ and input ${\rm SNR}=0$ dB; (b) scalar coefficient relating the Wiener and MVDR filters vs. input SNR and N; anechoic environment and white noise.

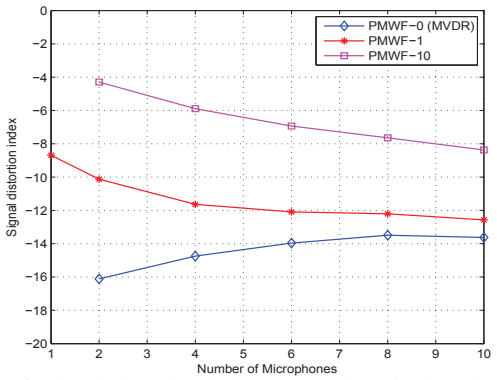

Fig. 2. Signal distortion index vs. number of microphones.

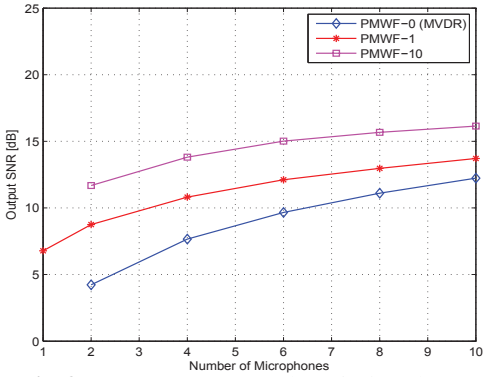

Fig. 3. Output SNR vs. number of microphones.

Fig. 2 presents the variations of the signal distortion measures with respect to the number of microphones (from 2 to 10). Note that we also included the performance of the single-channel non-causal Wiener filter (PMWF-1 with $N = 1$). We first see that using PMWF-1 with multiple microphones is more beneficial in terms of signal distortion than the single-channel case. In addition, the highest signal distortions are observed with the PMWF-10 and PMWF-1 while lower signal distortions are seen with the MVDR. This confirms the effect of the choice of the tuning parameter β that we expected in Section 3. Increasing the number of microphones reduces the signal distortion for both filters PMWF-1 and PMWF-10. The signal distortion achieved by the MVDR unexpectedly increases with the number of microphones. To explain this fact, recall that in theory the MVDR leads to zero signal distortion regardless of the number of microphones. This is not the case in practice because of the numerical inaccuracies in estimating the PSD matrices. Increasing the number of microphones leads to more estimation errors since the required PSD matrices become of larger sizes and more auto- and cross-PSD terms are estimated, thereby, increasing the overall estimation errors and increasing the signal distortion. In all cases, we observe that the resulting signal distortion measure complies with the theoretical effect of the tuning parameter β and the number of microphones N. In Fig. 3, we see that increasing the number of microphones leads, as expected, to more output SNR gains. Again, the effect of the choice of the parameter β complies with the theoretical findings of Section 3. Indeed, the highest output SNR values are achieved by the PMWF-10 while the lowest are achieved by the MVDR. This proves the tradeoff of signal distortion vs. noise reduction in the multichannel case. The MVDR filter is desired because of its low speech distortion. However, this comes at the price of low output SNR especially when few microphones are used. Note also that when the number of microphones increases, the PMWF-1 and the MVDR tend to have comparable effects in terms of output SNR and speech distortion, confirming, again, our theoretical results.

5. CONCLUSIONS

In this paper, a general framework for the design of non-causal noise reduction filters for microphone arrays is investigated leading to the PMWF. The Wiener filter and MVDR filters are essentially derived from the same optimization problem and are particular cases of the PMWF whose expression is shown to depend on the signal and noise statistics only. We investigated the theoretical performance of the PMWF and found interesting relationships between the input SNR, noise reduction, signal distortion, and the output SNR. Indeed, we highlighted the tradeoff between signal distortion and noise reduction in the multichannel case.

APPENDIX I: PROOF OF RESULT 1

The matrix $\mathbf{\Phi}_{xx}(j\omega) = \phi_{ss}(\omega)\mathbf{g}(j\omega)\mathbf{g}^H(j\omega)$ is of rank one and so is $\mathbf{\Phi}_{vv}^{-1}(j\omega)\mathbf{\Phi}_{xx}(j\omega)$ whose unique positive eigenvalue, $\lambda(\omega)$, is given by:

$$\lambda(\omega) = \text{tr}\left\{\mathbf{\Phi}_{vv}^{-1}(j\omega)\mathbf{\Phi}_{xx}(j\omega)\right\}. \quad (19)$$

For $\beta \neq 0$, we have the Woodbury's identity:

$$[\mathbf{\Phi}_{xx}(j\omega) + \beta\mathbf{\Phi}_{vv}(j\omega)]^{-1} = \frac{1}{\beta}[\mathbf{\Phi}_{vv}^{-1}(j\omega) - \frac{\mathbf{\Phi}_{vv}^{-1}(j\omega)\mathbf{\Phi}_{xx}(j\omega)\mathbf{\Phi}_{vv}^{-1}(j\omega)}{\beta + \lambda(\omega)}]. \quad (20)$$

Plugging (20) into (9), we obtain (10).

APPENDIX II: PROOF OF RESULT 2

We use the simplified expression (10) with (19) to obtain:

$$\mathbf{h}_{\text{W}\beta,n_0}^H(j\omega)\mathbf{\Phi}_{xx}(j\omega)\mathbf{h}_{\text{W}\beta,n_0}(j\omega) = \frac{\phi_{x_{n_0}x_{n_0}}(\omega)\lambda^2(\omega)}{[\beta + \lambda(\omega)]^2}, \quad (21)$$

$$\mathbf{u}_{n_0}^T\mathbf{\Phi}_{xx}(j\omega)\mathbf{h}_{\text{W}\beta,n_0}(j\omega) = \frac{\lambda(\omega)\phi_{x_{n_0}x_{n_0}}(\omega)}{\beta + \lambda(\omega)}, \quad (22)$$

and

$$\mathbf{h}_{\text{W}\beta,n_0}^H(j\omega)\mathbf{\Phi}_{vv}(j\omega)\mathbf{h}_{\text{W}\beta,n_0}(j\omega) = \frac{\lambda(\omega)\phi_{x_{n_0}x_{n_0}}(\omega)}{[\beta + \lambda(\omega)]^2}. \quad (23)$$

Plugging (21) and (22) in (5) leads to (14). Plugging (23) in (6) leads to (15). Finally, plugging (23) and (21) in (7) leads to (16).

6. REFERENCES

[1] Y. Huang, J. Benesty, J. Chen, "Analysis and comparison of multichannel noise reduction methods in a common framework," *IEEE Trans. Audio, Speech, Language Process.*, vol. 16, no. 5, pp. 957–968, Jul. 2008.

[2] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Springer-Verlag, Berlin, Germany, 2008.

[3] M. Souden, J. Benesty, and S. Affes, "On Optimal Frequency-Domain Multichannel Linear Filtering for Noise Reduction," Submitted to *IEEE Trans. Audio, Speech, Language Process.*, May 2008.

[4] S. Gannot, D. Burshtein, and E. Weinstein, "Signal enhancement using beamforming and nonstationarity with applications to speech," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1614-1626, Aug. 2001.

[5] S. Gannot and I. Cohen, "Adaptive beamforming and postfitering," in *Springer Handbook of Speech Processing*, editors-in-chief, J. Benesty, Y. Huang, and M.M. Sondhi, Springer, Nov. 2007.

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

[7] S. Doclo, A. Spriet, M. Moonen, J. Wouters, " Frequency-domain criterion for the speech distortion weighted multichannel wiener filter for robust noise reduction," *Speech Communication*, vol. 49, no. 7-8, pp. 636-656, Aug. 2007.

[8] L. J. Griffiths and C. W. Jim,"An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.* vol. AP-30, pp. 27–34, Jan. 1982.

[9] A. Spriet , S. Doclo, M. Moonen, J. Wouters, " A unification of adaptive multi-microphone noise reduction systems," in *Proc. IWAENC*, 2006, pp. 1–4.

[10] J. Chen, J. Benesty, Y. Huang, and S. Doclo, "New insights into the noise reduction Wiener filter," *IEEE Trans. Audio, Speech, Language Process.*, vol. 14, pp. 1218-1234, Jul. 2006.

[11] P. C. Loizou, *Speech enhancement: Theory and Practice*. CRC Press, New York, USA, 2007.

[12] S. Doclo and M. Moonen,"GSVD-based optimal filtering for single and multimicrophone speech enhancement," *IEEE Trans. Signal Process.*, vol. 50, pp. 2230–2244, Sept. 2002.

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