A SPEECH ENHANCEMENT SYSTEM USING BINAURAL HEARING AIDS AND AN EXTERNAL MICROPHONE

Dianna Yee[⋆†] , *Homayoun Kamkar-Parsi*[⋆], *Henning Puder*[⋆], *Rainer Martin*[†]

[⋆] Sivantos GmbH, Henri-Dunant Strasse 100, 91058 Erlangen, Germany,
dianna.yee@sivantos.com, homayoun.kamkarparsi@sivantos.com, henning.puder@sivantos.com
[†]Ruhr-Universität Bochum, Universitätsstrasse 150, 44801 Bochum, Germany, rainer.martin@rub.de

ABSTRACT

This paper presents a strategy for using an external microphone for enhancing noisy speech in single-microphone completely-in-canal (CIC) hearing aids. The external microphone is placed such that it benefits from the body shielding noise from the back hemisphere. The presented algorithm first enhances the external microphone signal without assuming an exact known location of the external microphone. The proposed algorithm then automatically incorporates the enhanced external microphone signal for post-processing enhancement of a conventional dual-channel binaural beamformer whenever the external microphone has a significant SNR advantage. The overall enhancement scheme avoids error-prone estimations of target voice activity detection and relative transfer functions between the microphones. Unlike single-channel post-processing filters which are limited to reducing stationary or diffuse noise, the proposed post-processing filter is able to reduce highly non-stationary directional noise from the back hemisphere. The resulting system provides enhancement even for noise arriving from the backward direction, which conventionally is difficult for CIC hearing aids where there exists a front-back ambiguity.

Index Terms— binaural beamforming, generalized sidelobe canceller, hearing aids, external microphone

1. INTRODUCTION

In hearing aid (HA) systems, speech enhancement algorithms are used to improve intelligibility of the desired signal while reducing surrounding noise. In practice, HA algorithms only combine signals of the HA for audio processing, combinations with external microphone (EMic) systems are not implemented. Multi-microphone enhancement algorithms have been proposed, such as multichannel Wiener filtering (MWF) [1, 2, 3] but this requires a reliable target voice activity detector (VAD) for estimating second order statistics of the target and noise signals. Another example is minimum-variance distortion-less response (MVDR) beamforming [4, 5], but this assumes that the relative transfer functions (RTF) of the target signal between all microphones are known. The RTF between the binaural HA can be estimated accurately but estimating the RTFs between the EMic and HAs is difficult as the EMic location is unknown and free to change. Furthermore, the sampling clocks of different devices are not synchronized in general. Such methods pose a difficult and possibly error-prone estimation problem.

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme under REA grant agreement [317521].

The advanced dual-channel beamformer (BF), which uses a version of a generalized sidelobe canceller (GSC) [6] tailored for binaural HAs, is robust and useful in practical HA applications [7]. It can attenuate lateral speech interferers but is limited in its ability to reduce interfering speech from 180°. To address this limitation, an additional EMic can be used to resolve the spatial ambiguity between 0° and 180°. In this work, the EMic is used to enhance an advanced dual-channel BF by computing a post-processing filter to further reduce noise from the back hemisphere. The phase mismatch between the HA and EMic signals does not allow for directly combining these signals and therefore, the proposed post-processing filter enhances the dual-channel BF using a frequency-dependent real-valued gain. This approach avoids accurate RTF or target VAD estimations. In addition, the proposed post-processing filter gain is advantageous to the conventional single-channel post-processing filter in attenuating back directional noise as the latter lacks ability in reducing highly non-stationary noise [8].

In this paper, we first present a placement strategy where the EMic can provide signal-to-noise-ratio (SNR) benefits for a binaural HA with a single microphone per device. Then, we outline the preliminary calibration steps required for incorporating an EMic with a HA system and propose a combined enhancement system which incorporates the EMic signal with a dual-channel BF. Lastly, we then evaluate the proposed combined system in comparison to the dual-channel BF. These topics are addressed in the following sections.

1.1. Notation and signal model

Assuming a linear time-invariant acoustic system, the m^{th} microphone signal at time t is denoted as $y_m(t)$. $y_m(t)$ is modeled as a convolution of the speech signal, $x(t)$, with the impulse response, $h_m(k)$, and additive noise $n_m(t)$,

$$y_m(t) = h_m(t) * s(t) + n_m(t). \tag{1}$$

Note that $n_m(t)$ can be a combination of directional interfering speakers and diffuse noise. m can take on values L, R and E to denote the left, right HA and EMic signals respectively. The short-time frequency domain representation of the signal model is given by

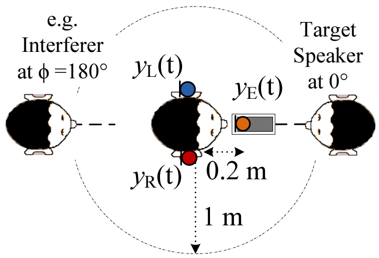
$$Y_m(k,n) = H_m(k)S(k,n) + N_m(k,n) \tag{2}$$

where k and n denote the sub-sampled frequency and time index respectively. K and N denotes the total number of subbands and time samples respectively.

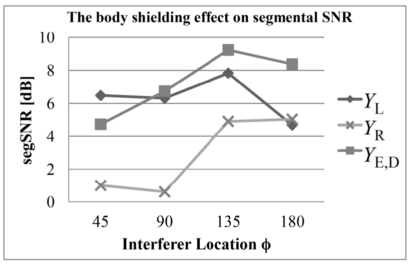
2. METHODS

2.1. Placement strategy for the body-shielding effect

We focused on the scenario shown in Figure 1a, where the EMic is centered and in front of the body at a distance of 20 cm. The HA user



(a) Microphones setup.



(b) Body shielding effect.

Fig. 1: SNR benefit from strategic placement of the EMic.

is wearing a single microphone CIC HA in each ear and the EMic is placed on a table at waist level. The target speaker DOA is 0° while the noise source location can vary along a 1 m radius circle around the HA user. The location of the noise source, ϕ, is varied in 45° increments. The front microphone of the iPhone 5s is used as the EMic and an ideal data transmission link exists between the EMic and HAs. A binaural link between the HAs is also assumed and the target equalization (EQ) filter weights (see Section 2.2) are known. Using this setup, the segmental SNR (segSNR)

$$\text{segSNR} = \frac{10}{KN}\sum_{k=0}^{K-1}\sum_{n=0}^{N-1}\log\left[\frac{|D_{\text{s}}(k,n)|^2}{|Y_i(k,n) - D_{\text{s}}(k,n)|^2}\right] \quad (3)$$

was calculated in the frequency domain using a block size of 30 ms and 50% overlap. $D_{\text{s}}(k,n)$ denotes the clean target signal in the i^{th} microphone signal, $Y_i(k,n)$. When calculating the segSNR for $Y_{\text{E,D}}(k,n)$, $Y_{\text{L}}(k,n)$ and $Y_{\text{R}}(k,n)$, $D_{\text{s}}(k,n)$ denotes the target signal component of $Y_{\text{E,D}}(k,n)$, $Y_{\text{L}}(k,n)$ and $Y_{\text{R}}(k,n)$ respectively. Figure 1b compares the segSNR of the EMic and CIC HAs when a speech interferer is present during target speech. The EMic signal has a significantly higher SNR than the raw HA signals when ϕ is in the back hemisphere. This characteristic is due to the effect of the body shielding back-directional noise. We refer to this effect as the body-shielding effect. For frontal interferers, the better ear, due to head-shadowing, has the better SNR. Note that head-shadowing is not present when ϕ = 180°. For this scenario, the body-shielding effect clearly offers an SNR advantage for the HA system.

We now focus on an enhancement scheme which takes advantage of the body-shielding effect in the EMic. Figure 2 shows the overall enhancement scheme. There are two enhancement subsystems - the top subsystem shows the scheme for enhancing the EMic while the lower shows scheme for the dual-channel BF. We first review the enhancement scheme of the dual-channel BF, then we elaborate on the enhancement of the EMic and finally, its incorporation into the HA system.

2.2. The dual-channel beamformer

The dual-channel BF, shown as the lower enhancement subsystem in Figure 2, has been an area of interest for many researchers [9, 10, 11]. The output of the fixed BF, $Y_{\text{HA,D}}(k,n)$, is derived from the average of the two target equalized signals, which predominately contains the target signal. The target EQ filter weights are calculated offline and applied during online processing. It is assumed that the target direction of arrival (DOA) is known. Note that DOA estimation is not within the scope of this paper. The noise estimate, $N_{\text{BM}}(k,n)$, is derived by taking the difference of the two target-equalized signals, which will contain mainly noise and ideally no target signal. The adaptive noise canceller (ANC) is implemented with a subband normalized-least-mean-squares (NLMS) algorithm where $\boldsymbol{N}_{\text{BM}}$, a vector of length L_{ANC} containing the current and $L_{\text{ANC}} - 1$ past values of N_{BM}, is used for adaptive noise reduction from $Y_{\text{HA,D}}(k,n)$ [12, 13]. A causality delay, d_{ANC}, is added to ensure a causal system. The filter coefficients, $\boldsymbol{H}_{\text{ANC}}(k,n)$, are adapted to minimize the variance of the output signal. Assuming uncorrelated target and noise signals, the result is an enhanced HA signal, $Y_{\text{HA,enh}}(k,n)$. Note that an ANC is also used for EMic enhancement. The output signals of the enhancement subsystems shown in Figure 2 are given as

$$Y_{i,\text{ANC}}(k,n) = Y_{i,\text{D}}(k,n) - \boldsymbol{H}_{\text{ANC}}(k,n)^H \boldsymbol{N}_{\text{BM}}(k,n) \quad (4)$$

where i is E or HA to denote the respective enhancement systems. The filter coefficient vector, $\boldsymbol{H}_{\text{ANC}}(k,n)$, is updated by

$$\boldsymbol{H}_{\text{ANC}}(k,n+1) = \boldsymbol{H}_{\text{ANC}}(k,n) \quad + \frac{\mu(k)\boldsymbol{N}_{\text{BM}}(k,n)Y^*_{i,\text{ANC}}(k,n)}{\boldsymbol{N}_{\text{BM}}(k,n)^H\boldsymbol{N}_{\text{BM}}(k,n) + \delta(k)} \quad (5)$$

where $\mu(k)$ is the NLMS step size and $\delta(k)$ is the regularization factor. Note that the dual-channel BF does not preserve spatial cues but rather strives for a maximum SNR.

2.3. External microphone hardware calibration

Calibration is required to account for the device group delay difference between the EMic and HA, the microphone characteristic difference, and the difference in location of the EMic to the HA. The first two factors are related to hardware (HW) calibration and for matching the microphone system characteristics between the EMic and HA. The HW calibrated signal is denoted as $Y_{\text{E,calib}}(k,n)$ in Figure 2. HW calibration can be calculated offline or estimated online [14]. The last factor, however, cannot be anticipated as the EMic location is unknown. This implies that the target signal received in the EMic will be different than that of the HA, and adjustment is needed to match the target signals in both systems.

2.4. External microphone target level adjustment

It is difficult to equalize the level and phase of the target signal between the different microphone systems since the RTF between the EMic and HA is unknown. A partial adjustment can be used instead where only the level of the target signal is adjusted by applying a gain, $\Gamma_i(k,n)$. Wittkop [15] suggests two approaches for estimating $\Gamma_i(k,n)$. The first approach calculates $\Gamma_i(k,n)$ in advance as a constant, using the mean of $\Gamma_i(k,n)$ over all time indices. The second approach calculates $\Gamma_i(k,n)$ when target speech presence probability is high. In our experiments, the first approach to estimating $\Gamma_i(k,n)$ is used. The target level adjusted signal, denoted by

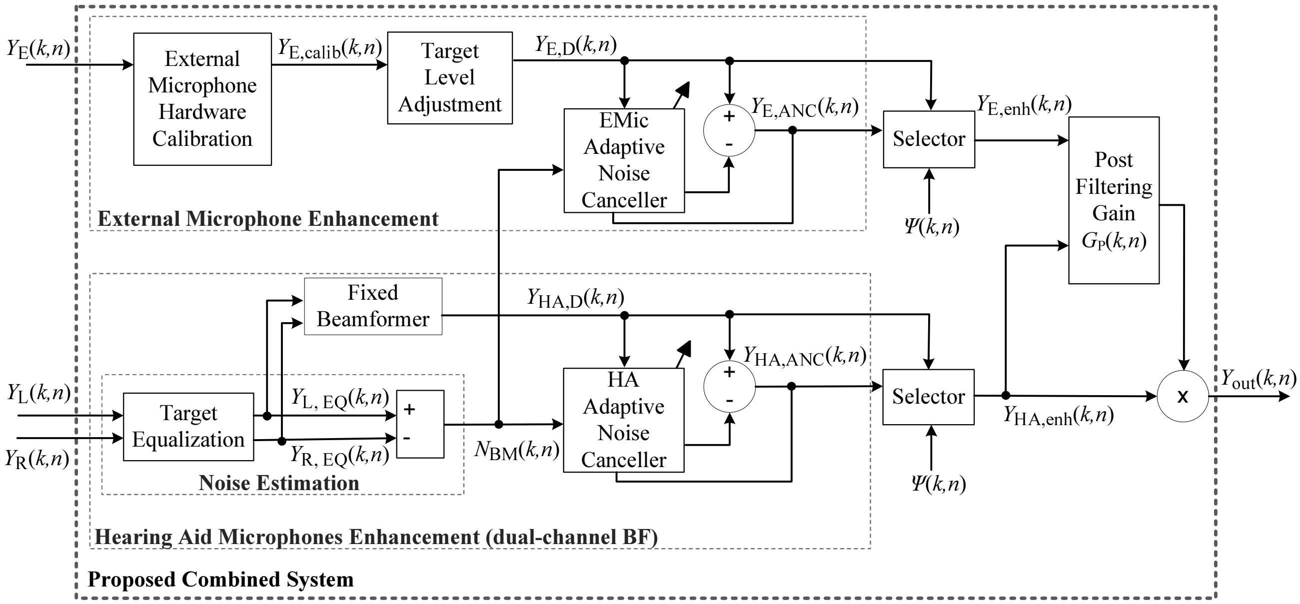


Fig. 2: The proposed enhancement scheme for processing input signals from a CIC hearing aid and an external microphone in the short-time Fourier domain.

$Y_{\mathrm{E,D}}(k,n)$, is derived using (6), i.e.

$$Y_{\mathrm{E,D}}(k,n) = \Gamma_i(k,n) Y_{\mathrm{E,calib}}(k,n). \tag{6}$$

2.5. External microphone enhancement

$Y_{\mathrm{E,D}}(k,n)$ is used as the reference signal for the ANC. The noise estimate, $N_{\mathrm{BM}}(k,n)$ is derived as discussed in Section 2.2 and the ANC is also implemented using the same approach using (4) and (5). The resulting enhanced EMic signal is denoted as $Y_{\mathrm{E,ANC}}(k,n)$.

2.6. Limitation of binaural noise estimate

Due to symmetry of the target location with ϕ = 180°, $N_{\mathrm{BM}}(k,n)$ essentially puts a notch at both 0° and 180°. $N_{\mathrm{BM}}(k,n)$ then contains only residual noise from 180° and target leakage at 0°. Both $Y_{\mathrm{E,ANC}}(k,n)$ and $Y_{\mathrm{HA,ANC}}(k,n)$ can suffer from SNR degradation due to the poor noise estimate. To avoid degradation of SNR, no enhancement is applied in this scenario by considering the following control signal,

$$\Psi(k,n) = \frac{\mathrm{E}\left\{|Y_{\mathrm{L,EQ}}(k,n) - Y_{\mathrm{R,EQ}}(k,n)|^2\right\}}{\mathrm{E}\left\{|Y_{\mathrm{L,EQ}}(k,n) + Y_{\mathrm{R,EQ}}(k,n)|^2\right\}} \leq \delta_{0,180}. \tag{7}$$

where $\delta_{0,180}$ is the threshold for determining when this symmetrical scenario exists. Due to symmetry of 0° and 180° in $Y_{\mathrm{L}}(k,n)$ and $Y_{\mathrm{R}}(k,n)$, $\Psi(k,n)$ is small compared to asymmetrical noise scenarios. When this scenario is detected, no enhancement is applied. In other words, the output signal of the i^{th} enhancement subsystem is defined as

$$Y_{i,\mathrm{enh}}(k,n) = \begin{cases} Y_{i,\mathrm{D}}(k,n), & \Psi(k,n) \leq \delta_{0,180} \\ Y_{i,\mathrm{ANC}}(k,n) & \text{otherwise.} \end{cases} \tag{8}$$

where i can take on values E or HA.

2.7. Incorporating the external microphone

$Y_{\mathrm{HA,enh}}$ and $Y_{\mathrm{E,enh}}$ have different advantages in different noise scenarios and it is useful to determine when to incorporate the EMic or not. When $Y_{\mathrm{E,enh}}$ benefits from a better SNR, it can be used for enhancing $Y_{HA,enh}$. Contrarily, when $Y_{\mathrm{E,enh}}$ has a worse SNR than $Y_{\mathrm{HA,enh}}$, it does not provide any benefit for enhancing $Y_{\mathrm{HA,enh}}$ and $Y_{\mathrm{HA,enh}}$ would be chosen as the final output of the system, denoted as Y_{out}. To determine which enhanced signal provides the better SNR, we consider the power of the enhanced signals. Assuming uncorrelated target and noise signals, the expected power of the i^{th} enhanced signal, denoted by $\sigma^2_{i,\mathrm{enh}}$, is given as

$$\sigma^2_{i,\mathrm{enh}}(k,n) = \mathrm{E}\left\{|H_{i,\mathrm{enh}}(k,n)S(k,n)|^2\right\} + \mathrm{E}\left\{|N_{i,\mathrm{enh}}(k,n)|^2\right\} \tag{9}$$

where i can be E or HA. Since $\Gamma_i(k,n)$ aims to match the level of the EMic to the reference HA and assuming no target distortion during the enhancement, it is assumed that

$$\mathrm{E}\left\{|H_i(k,n)S(k,n)|^2\right\} \approx \mathrm{E}\left\{|\Gamma_i(k,n)H_{\mathrm{E}}(k,n)S(k,n)|^2\right\} \tag{10}$$

where i refers to the reference HA. Under such assumptions, the enhanced signal with the lowest power would have the lower amount of residual noise and therefore has the better SNR. The signal with the better SNR can be chosen as Y_{out} or used in a post-processing step to further enhance $Y_{\mathrm{HA,D}}$ to yield Y_{out}. In applications where a target speech power estimate is required, directly switching to $Y_{\mathrm{E,enh}}$ would be more beneficial when $Y_{\mathrm{E,enh}}$ has an advantageous SNR. However, for listening applications, switching to $Y_{\mathrm{E,enh}}$ is not preferred due to audio artifacts arising from discrepancies between the microphone signals. Therefore, we incorporate the EMic in the HA system using the post-processing approach.

Although the phases of $Y_{\mathrm{HA,enh}}$ and $Y_{\mathrm{E,enh}}$ are not aligned, these signals can still be used to compute real-valued gains for post-processing the HA signals. Additionally, post-processing should only be applied when $Y_{\mathrm{E,enh}}$ has a better SNR than $Y_{\mathrm{HA,enh}}$ so that we can estimate the target signal in $Y_{\mathrm{HA,emh}}$ with $Y_{\mathrm{E,enh}}$ to derive the post-filtering gain. One realization of this concept is a Wiener type filter that can be applied to $Y_{\mathrm{HA,enh}}$ using the outputs from both enhancement subsystems. For a coherence-based definition, the

Wiener gain, $G_{\mathrm{p_o}}(k,n)$, is

$$G_{\mathrm{p_o}}(k,n) = \frac{\mathrm{E}\left\{|Y_{\mathrm{HA,enh}}(k,n)Y_{\mathrm{E,enh}}(k,n)^*|^2\right\}}{\mathrm{E}\left\{|Y_{\mathrm{HA,enh}}(k,n)|^2\right\}}. \tag{11}$$

$G_{\mathrm{p_o}}(k,n)$ preserves correlated components of $Y_{\mathrm{HA,enh}}$ and $Y_{\mathrm{E,enh}}$ while the uncorrelated components are attenuated. $G_{\mathrm{p_o}}(k,n)$ is applied when $10\log_{10}(\sigma^2_{\mathrm{HA,enh}}(k,n))$ - $10\log_{10}(\sigma^2_{\mathrm{E,enh}}(k,n)) > \delta_{\mathrm{G_p}}$ where $\delta_{\mathrm{G_p}}$ is a threshold which implicitly determines the SNR advantage at which $Y_{\mathrm{E,enh}}$ must provide before it is incorporated into the system. We found that $\delta_{\mathrm{G_p}}$ worked well between [-0.5, 0] dB. If $\delta_{\mathrm{G_p}}$ is too large, then the SNR benefit of the EMic is not used to its full advantage, while if $\delta_{\mathrm{G_p}}$ is too small, the EMic is incorrectly used when it does not provide an SNR advantage. The post-processing gain shown in Figure 2, $G_{\mathrm{p}}(k,n)$, is then defined as

$$G_{\mathrm{p}}(k,n) = \begin{cases} G_{\mathrm{p_o}}(k,n), & 10\log_{10}(\sigma^2_{\mathrm{HA,enh}}(k,n)) - \\ & \quad 10\log_{10}(\sigma^2_{\mathrm{E,enh}}(k,n)) > \delta_{\mathrm{G_p}} \\ 1 & \text{otherwise} \end{cases} \tag{12}$$

and is applied to $Y_{\mathrm{HA,enh}}(k,n)$ for further enhancement, i.e.

$$Y_{\mathrm{out}}(k,n) = G_{\mathrm{p}}(k,n)Y_{\mathrm{HA,enh}}(k,n). \tag{13}$$

3. RESULTS

The audio recordings used were obtained in a room with T_{60} =120 ms. The ANC was parametrized using d_{ANC} = 1, L_{ANC} = 4, $\mu(k)$ = 0.1 for all frequencies, and $\delta(k)$ was calculated in advance using $\delta(k) = \beta\sigma^2_{\mathrm{N_{BM}}}$ where β = 20 [16],[17]. The thresholds used are $\delta_{0,180}$ = -12 dB and $\delta_{\mathrm{G_p}}$ = -0.5 dB. The performance of the enhancement schemes is compared by the segSNR [18] and short time objective intelligibility (STOI) measures [19]. The segSNR was calculated using (3) where $D_{\mathrm{s}}(k,n)$ denotes the clean target signal. When calculating the segSNR for $Y_{\mathrm{E,enh}}(k,n)$, $D_{\mathrm{s}}(k,n)$ denotes the target signal component in $Y_{\mathrm{E,D}}(k,n)$. Likewise, when calculating the segSNR of $Y_{\mathrm{HA,enh}}(k,n)$ and $Y_{\mathrm{out}}(k,n)$, $D_{\mathrm{s}}(k,n)$ denotes the target signal component in $Y_{\mathrm{HA,D}}(k,n)$. The segSNR measures the average short-time power of target speech to residual noise, while the STOI predicts the objective speech intelligibility of the enhanced signal by using the temporal cross-correlation of the subband magnitude envelopes of the enhanced and target signal within one-third octave bands. The target signal used for calculating STOI follows the same definition as $D_{\mathrm{s}}(k,n)$ for calculating segSNR.

3.1. Performance for a single interferer

The performance of $Y_{\mathrm{HA,enh}}$, $Y_{\mathrm{E,enh}}$ and Y_{out} is compared when there is a single interferer at ϕ and a 0° target. Table 1 shows the segSNRs of enhanced signals along with $Y_{\mathrm{HA,D}}$ and $Y_{\mathrm{E,D}}$ for comparison. Although $Y_{\mathrm{E,enh}}$ gives the best SNR, post-filtering $Y_{\mathrm{HA,enh}}$ is opted for due to the mismatch of frequency responses between the EMic and HA microphone signals. When $\phi < 90°$, $Y_{\mathrm{HA,enh}}$ has a better segSNR and the selection strategy optimally chooses $Y_{\mathrm{HA,enh}}$ as Y_{out}. For $\phi \geq 90°$, $Y_{\mathrm{E,enh}}$ has a better segSNR and therefore post-filtering is applied. An overall segSNR and STOI improvement in $Y_{\mathrm{HA,enh}}$ results from the proposed incorporation of the EMic.

3.2. Advantages in adverse environments

The single-channel noise reference, N_{BM}, is sufficient when there exists a single interferer, however, with increasing numbers of interferers, the ANC provides less noise reduction. Nonetheless, the body-shielding effect inherent in the EMic signal can still be utilized as the SNR of $Y_{\mathrm{E,enh}}$ is still better than that of $Y_{\mathrm{HA,enh}}$. Table 2 shows various adverse noise scenarios where the pair of numbers indicates the locations of two interferers. In cases where $Y_{\mathrm{E,enh}}$ is beneficial for post-filtering of $Y_{\mathrm{HA,enh}}$, segSNR improves at least 1 dB and STOI is also slightly improved.

		Interferer Location ϕ			
		45	90	135	180
segSNR	$Y_{\mathrm{E,D}}$	0.1	1.7	0.8	3.1
	$Y_{\mathrm{HA,D}}$	-0.4	-0.3	-1.4	-0.1
	$Y_{\mathrm{E,enh}}$	1.8	3.5	1.9	3.1
	$Y_{\mathrm{HA,enh}}$	2.9	2.6	1.5	0.1
	Y_{out}	2.9	2.8	2.1	1.5
STOI	$Y_{\mathrm{E,enh}}$	0.73	0.77	0.69	0.69
	$Y_{\mathrm{HA,enh}}$	0.82	0.78	0.73	0.62
	Y_{out}	0.82	0.79	0.73	0.69

Table 1: Performance Measures for Single Interferer varying ϕ

		Adverse Noise Scenarios				
		45, 90	45, 135	45, 180	90, 180	135, 180
segSNR	$Y_{\mathrm{E,D}}$	-1.3	-1.5	-0.3	0.6	0.0
	$Y_{\mathrm{HA,D}}$	-2.0	-2.7	-1.8	-1.8	-2.7
	$Y_{\mathrm{E,enh}}$	0.2	-0.7	0.6	1.2	0.4
	$Y_{\mathrm{HA,enh}}$	1.6	0.8	-0.6	-0.6	-1.4
	Y_{out}	1.6	0.8	0.5	0.6	0.2
STOI	$Y_{\mathrm{E,enh}}$	0.68	0.63	0.66	0.68	0.65
	$Y_{\mathrm{HA,enh}}$	0.76	0.73	0.63	0.64	0.60
	Y_{out}	0.76	0.73	0.68	0.69	0.65

Table 2: Performance Measures in Adverse Noise Scenarios

3.3. Advantage to single-channel post-processing filter

Table 3 compares the SNR improvement resulting from the proposed post-processing filter and a single-input-single-output (SISO) post-processing Wiener filter. The latter is calculated using the noise power and apriori SNR of $Y_{\mathrm{HA,enh}}$, estimated using methods proposed in [20] and [21] respectively. As expected, the SISO filter is able to reduce diffuse cafeteria babble noise but does not perform well during interfering speech. By taking advantage of the body-shielding effect, the proposed post-processing filter can reduce highly non-stationary noise coming from the back direction.

		Noise Scenarios				
		babble	180	45, 180	90, 180	135, 180
segSNR	SISO	1.9	0.2	0.6	0.7	0.8
	proposed	0.3	1.3	1.0	1.2	1.6

Table 3: Post-Processing Filter segSNR Improvement

4. CONCLUSION

In this work, a new approach for combining a BF with an EMic device is introduced. The proposed system was shown to be a beneficial enhancement of the conventional dual-channel HA BF. The incorporation of the EMic improved SNR and STOI whenever there exists dominating back-directional noise. Benefits are noticed also in adverse noise environments with multiple interferers. The presented work strives for a better SNR performance but lacks binaural cues. Future work will also include the restoration of spatial cues to the enhanced signal.

5. ACKNOWLEDGMENTS

The authors would like to thank Eghart Fischer for participating in deep discussions related to this topic.

6. REFERENCES

[1] B. Cornelis and M. Moonen, "Speech intelligibility improvements with hearing aids using bilateral and binaural adaptive multichannel wiener filtering based noise reduction," *Journal Acoustical Society of America*, vol. 131, pp. 4743 – 4755, June. 2012.

[2] B. Cornelis, M. Moonen, and J. Wouters, "Reduced-bandwidth multi-channel wiener filter based binaural noise reduction and localization cue preservation in binaural hearing aids," *Elsevier-Signal Processing*, vol. 99, pp. 1 – 16, June. 2014.

[3] A. Bertrand and M. Moonen, "Robust distributed noise reduction in hearing aids with external acoustic sensor nodes," *EURASIP J. Adv. Signal Process*, vol. 2009, pp. 12:3–12:3, Jan. 2009.

[4] T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Robustness analysis of binaural hearing aid beamformer algorithms by means of objective perceptual quality measures," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 315 – 318, Oct. 2007.

[5] J. Riba, J. Goldberg, and D. Vazquez, "Robust beamforming for interference rejection in mobile communications," *IEEE Transactions on Signal Processing*, vol. 45, pp. 271–275, Jan. 1997.

[6] L. Griffiths and C. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Transactions on Antennas and Propagation*, vol. 30, pp. 27–34, July. 1982.

[7] S. Srinivasan, A. Pandharipande, and K. Janse, "Beamforming under quantization errors in wireless binaural hearing aids," *EURASIP Journal on Applied Signal Processing*, vol. 2008, June. 2008.

[8] I. Cohen, "Analysis of two-channel generalized sidelobe canceller with post-filtering," *IEEE Transactions on Speech and Audio Processing*, vol. 11, pp. 684–698, Nov. 2003.

[9] M. Kompis and N. Dillier, "Noise reduction for hearing aids: Combining directional microphones with an adaptive beamformer," *Journal Acoustical Society of America*, vol. 96, pp. 1910–1913, Sept. 1994.

[10] M. Kompis, P. Feuz, and M. Pelizzone, "A combined fixed/adaptive beamforming noise-reduction system for hearing aids," *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 20, pp. 3136–3139, Nov. 1998.

[11] L. Greenberg and P. Zurek, "Evaluation of an adaptive beamforming method for hearing aids," *Journal Acoustical Society of America*, vol. 91, pp. 1662–1676, Sept. 1992.

[12] S. Haykin, "Adaptive filter theory," *Prentice Hall*, pp. 432–437, 1996.

[13] A. Sayed, "Adaptive filters," *Wiley-Interscience*, 2008.

[14] M. Buck, T. Haulick, and H.J. Pfleiderer, "Microphone calibration for multi-channel signal processing," in *Speech and Audio Processing in Adverse Environments*, Eberhard Hänsler and Gerhard Schmidt, Eds., Signals and Communication Technology, pp. 417–467. Springer Berlin Heidelberg, 2008.

[15] T. Wittkop and V. Hohmann, "Strategy-selective noise reduction for binaural digital hearing aids," *Elsevier Science - Speech Communication*, vol. 39, pp. 111–138, Jan. 2003.

[16] J. Benesty and C. Paleologu, "On regularization in adaptive filtering," *IEEE Transactions on Audio, Speech, and Language Processing*, pp. 1734 – 1742, Aug. 2011.

[17] J. Benesty, M.M. Sondhi T. Gaensler, D. R. Morgan, and S. L. Gay, "Advances in network and acoustic echo cancellation," *Springer-Verlag*, 2001.

[18] P. Loizou, "Speech enhancement theory and practice," *CRC Press*, 2007.

[19] C. H. Taal, R. Heusdens R. C. Hendriks, and J. Jensen, "A short-time objective intelligibility measure for time-frequency weighted noisy speech," *IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 4214–4217, Mar. 2010.

[20] T. Gerkmann and R. C. Hendriks, "Unbiased mmse-based noise power estimation with low complexity and low tracking delay," *IEEE Transactions on Audio, Speech, Signal Processing*, pp. 1383–1393, May. 2012.

[21] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Transactions on Audio, Speech, Signal Processing*, pp. 1109–1121, Dec. 1984.