Operational Rate-Constrained Beamforming in Binaural Hearing AIDS

Amini, Jamal; Hendriks, Richard C.; Heusdens, Richard; Guo, Meng; Jensen, Jesper Rindom

DOI
10.23919/EUSIPCO.2018.8553403

Publication date
2018

Document Version
Final published version

Published in
2018 26th European Signal Processing Conference (EUSIPCO)

Operational Rate-Constrained Beamforming in Binaural Hearing Aids

Jamal Amini[†], Richard C. Hendriks[†], Richard Heusdens[†], Meng Guo[⋆] and Jesper Jensen[⋆‡]
[†]Circuits and Systems (CAS) Group, Delft University of Technology, 2628 CD Delft, the Netherlands
[‡]Electronic Systems Department, Aalborg University, 9100 Aalborg, Denmark
[⋆]Oticon A/S, Kongebakken 9, 2765 Smørum, Denmark
E-mails: [†]{j.amini, r.c.hendriks, r.heusdens}@tudelft.nl, [‡⋆]{megu, jesj}@oticon.com

***Abstract*—Modern binaural hearing aids (HAs) can collaborate wirelessly with each other as well as with other assistive (wireless) devices. This enables multi-microphone noise reduction over small wireless acoustic sensor networks (WASNs) to increase the intelligibility under adverse conditions. In this work, we assume one of the HAs to serve as the fusion center (FC). The optimal beamforming strategy for processing the received data at the FC depends on the acoustic scene and physical constraints (e.g., the bit-rate for transmission to the FC), and might be frequency dependent. Selection of the optimal beamforming strategy, while satisfying rate constraints on the communication between the different devices is an important challenge in such setups. In this paper, we propose an operational rate-constrained beamforming system for optimal rate allocation and strategy selection across frequency. We show an example of the proposed framework, where both the algorithm selection as well as the required rates to transmit the necessary microphone signals are optimized using uniform quantizers, while minimizing the mean-square error (MSE) distortion measure. In contrast to a well-known (theoretically optimal) reference method based on remote source coding for two devices, the presented algorithm is practically implementable and only requires knowledge of joint signal statistics at the FC. Evaluations (based on simulation experiments) show clear improvement over other practically implementable strategies.**

***Index Terms*—Binaural hearing aids, multi-microphone noise reduction, operational rate-distortion tradeoff**

I. Introduction

Hearing aid (HA) devices are designed to increase the speech intelligibility. A typical way to improve the speech intelligibility is by means of multi-microphone noise reduction [1][2]. Modern HAs can collaborate through a wireless link to construct a binaural HA system. This considerably improves the potential of noise reduction, as effectively a larger microphone array can be used [3][4]. In addition, binaural HAs can collaborate with other assistive devices, and form a small wireless acoustic sensor network (WASN).

In such a small WASN, microphone recordings are received at the fusion center (FC), which estimates the target sources and suppresses the interferers. In this work, one of the two HAs is considered as an FC. A well-known binaural filter is the binaural multichannel Wiener filter (MWF) [5], which is based on constructing two monaural MWF beamformers. Each MWF tries to estimate the source of interest by linearly combining its locally recorded signals with those from the contralateral device such that the mean square error (MSE) between the target source and its estimate is minimized. Other binaural beamforming approaches, including [6] and [7], try to preserve some important spatial information of the target and interfering sources when minimizing the MSE.

To perform such binaural processing, the noisy observations need to be transmitted through wireless links to the FC. As the transmission capacities of such links are limited, the data must be quantized at a certain bit-rate [8]. This brings the notion of rate-constrained beamforming into the noise reduction problem. In [9] a binaural rate-constrained beamforming problem is introduced, assuming jointly Gaussian random sources, where an efficient trade-off between the transmission rate and the MSE between the target signal and its estimate is derived. However, this optimal framework is limited to only two processing nodes and is less practical due to the strong requirement that joint statistics are known at all processors and (infinitely) long-block vector quantizers are used. Transmission between binaural HAs and other assistive devices is thus not considered, nor how more practical implementations affect the performance. Different (sub-optimal) binaural rate-constrained approaches are proposed in [8] and [10], which provide more practical alternatives to the method in [9]. However, the performance of such methods depends heavily on the acoustic scene (e.g. target source location, spatial noise distributions, etc.) and it is typically far from optimal, even asymptotically, i.e., at sufficiently high rates.

In this work, the binaural HA problem is approached from a more general perspective. The general setup of a (small) WASN is considered here, where joint statistics are only assumed to be known at the FC, instead of at every node as in [9]. The binaural noise reduction problem is solved by minimizing a fidelity criterion, while satisfying a bit-rate constraint. To overcome the acoustic scene dependency, we consider a discrete set of processing candidates and a (discrete) set of operating resources (in this case bit-rates). This problem formulation of optimizing among a set of strategies under a rate constraint is related to operational rate-distortion optimization [11][12]. In [11] an elegant operational rate-distortion optimization method was proposed for rate allocation among an arbitrary set of quantizers. Most related approaches were inspired by the method in [11] for different applications such

This work was supported by the Oticon Foundation and the Dutch Technology Foundation STW.

ISBN 978-90-827970-1-5 © EURASIP 2018

as optimal time segmentation of speech [13] or finding optimal time-varying wavelet packet bases for signal expansion [14].

We propose a new operational rate-constrained beamforming algorithm based on both strategy selection and rate allocation in the frequency domain. The Lagrange multiplier (LM) based technique [11] is used to allocate the rates and select the best strategies over frequency, while minimizing the sum of estimation error power spectral densities (PSDs). Unlike the theoretical approaches [8][9], the proposed method allows an arbitrary range of operating rates in each frequency bin. Moreover, it enables forming the set of processing candidates from existing algorithms and optimally choosing between different strategies in different frequency bins. The proposed method is evaluated based on the output MSE gap between the monaural (i.e., no communication) setup and the (rate-constrained) generalized binaural setup. The results show significant improvements in comparison with naive strategy selection and equal rate allocation across frequencies.

II. Problem Statement

The generalized binaural HA system that we consider consists of two wireless collaborating HAs with M_1 and M_2 microphones, respectively, and M_A assistive processors, which can collaborate with the HAs. The total number of microphones is thus $M = M_1 + M_2 + M_A$. In general, each assistive device can be equipped with multiple microphones. However, in this work, it is assumed for simplicity that each assistive processor is equipped with a single microphone. The clocks of the devices are assumed to be synchronized. All microphones receive a filtered version of the target speech signal, which is indicated in the short-time frequency transform (STFT) domain by $S[k]$, with $[k]$ denoting the frequency bin index. Notice that the time-frame index is neglected for notational convenience. The target speech is degraded by interfering noise, which might originate from, e.g., interfering point sources, diffuse noise, and microphone self-noise. The interfering noise observed at a particular microphone is indicated by $N_i[k]$, with $i = 1, \dots, M$ the microphone index. The signals $S[k]$ and $N_i[k]$, for $i = 1, \dots, M$ are assumed additive and mutually uncorrelated. Altogether we then have

$$Y_i[k] = A_i[k]S[k] + N_i[k], \tag{1}$$

where A_i is the acoustic transfer function (ATF) between the target signal and the ith microphone. The signal model can be rewritten in vector notation by stacking all noisy microphone coefficients in a vector, as

$$\mathbf{y} = \mathbf{x} + \mathbf{n}, \tag{2}$$

where $\mathbf{x} = \mathbf{a}S$, $\mathbf{y} = [Y_1[k], ..., Y_M[k]]^{\mathrm{T}}$, and similarly for $\mathbf{n}$ and $\mathbf{a}$. Notice that we have left out the frequency bin index in (2) for notational convenience. The superscripts $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ denote transpose and conjugate transpose operators, respectively. The cross-power spectral density (CPSD) matrix $\boldsymbol{\Phi}_{\mathbf{y}}$ of the vector $\mathbf{y}$ is given by $\boldsymbol{\Phi}_{\mathbf{y}} = \boldsymbol{\Phi}_{\mathbf{x}} + \boldsymbol{\Phi}_{\mathbf{n}}$, where $\boldsymbol{\Phi}_{\mathbf{x}} = \Phi_S \mathbf{a}\mathbf{a}^{\mathrm{H}}$, $\boldsymbol{\Phi}_{\mathbf{n}} = E[\mathbf{n}\mathbf{n}^{\mathrm{H}}]$ with $\Phi_S = E[|S|^2]$ the PSD of the clean speech S, and with $E[\cdot]$ the expectation operator.

In this paper, our goal is to estimate the clean speech target signal at the FC. However, apart from the microphone signals acquired at the FC, the additional microphone signals are only available in quantized form. These signals are compressed at a certain operating rate, say R bits per sample (bps), which is considered as a (constrained) resource. Depending on this resource and the actual acoustic scene, different algorithm selections are optimal. Therefore, we address the problem of operational rate-constrained beamforming in order to find the optimal beamforming strategy, given a set of candidate algorithms, satisfying the bit-rate as a resource constraint.

III. Operational rate-constrained beamfomrming

Inspired by [11], in this section we propose operational rate-distortion optimization for beamforming based on both rate allocation and strategy selection across frequencies.

We are given a set $\mathcal{A} = \{A_1, A_2, \dots, A_{N_A}\}$ of strategy candidates (could be different microphone configurations, different beamforming algorithms, and/or different coding schemes on the microphone signals) with cardinality $|\mathcal{A}| = N_A$. The goal is to optimally select the candidates and allocate the resources (bit-rates) in order to minimize a distortion, in this case, the MSE between the remote-source S and its estimate $\hat{S}$ in the frequency domain, while satisfying the constraints on the total rate budget, say $R_{\max}$. The proposed optimization problem is given by

$$\begin{aligned} &\min_{\boldsymbol{\alpha} \in \mathcal{A}'} \ \min_{\mathbf{r} \in \mathcal{Q}} \quad D(\boldsymbol{\alpha}, \mathbf{r}) \\ &\text{subject to} \quad\ R(\mathbf{r}) \leq R_{\max}, \end{aligned} \tag{3}$$

where $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_{N_f}]^{\mathrm{T}}$ denotes a vector variable for possible choices of strategies for all N_f frequency bins. Similarly, $\boldsymbol{r} = [r_1, \dots, r_{N_f}]^{\mathrm{T}}$ indicates a vector variable for possible operating rates to be allocated to the frequency components. The set of all possible strategy choices is given by $\mathcal{A}' = \{\boldsymbol{\alpha} \mid \alpha_k \in \mathcal{A}\}$, for $k = 1, \dots, N_f$. The set $\mathcal{Q} = \{\mathbf{r} \mid r_k \in \mathcal{Q}_k\}$ consists of possible operating rates, where $\mathcal{Q}_k = \{p_k, \dots, q_k\}, q_k > p_k \geq 0$, with representative cardinality $N_r \triangleq \max\{|\mathcal{Q}_1|, \dots, |\mathcal{Q}_{N_f}|\}$, for all frequency bins. Note that p_k and q_k are the minimum and the maximum operating rates, respectively, for a particular frequency. $D(\boldsymbol{\alpha}, \mathbf{r})$ is the averaged PSD of the estimation error, given the algorithm choices and rate allocation across frequencies and is given by

$$D(\boldsymbol{\alpha}, \mathbf{r}) = \frac{1}{N_f} \sum_{k=1}^{N_f} d(\alpha_k, r_k), \tag{4}$$

where

$$d(\alpha_k, r_k) = E[|S[k] - \hat{S}[k]|^2 | \alpha_k, r_k], \tag{5}$$

which denotes the PSD of the estimation error in the kth discrete frequency bin, given the algorithm α_k and the quantization rate r_k. The cost function $R(\mathbf{r})$ is simply defined as the averaged rate over all bins and is given by

$$R(\mathbf{r}) = \frac{1}{N_f} \sum_{k=1}^{N_f} r_k. \tag{6}$$

ISBN 978-90-827970-1-5 © EURASIP 2018

The original problem in (3) is a (discrete) combinatorial optimization problem. Every possible solution is an operating point located in the 2-dimensional D-R coordinate system (D-R characteristics). Figure 1 illustrates an example D-R characteristic. The problem of finding the optimal operating point which satisfies the constraint in (3) is untractable. One way to make the search problem tractable is to approximate the convex hull of the set of all possible solutions and select a point on the convex hull which satisfies the constraints [12]. Using the LM technique [11], the original problem in (3) is reformulated to the following Lagrangian form as

$$\min_{\boldsymbol{\alpha}\in\mathcal{A}'} \min_{\mathbf{r}\in\mathcal{Q}} \quad D(\boldsymbol{\alpha},\mathbf{r}) + \lambda\, R(\mathbf{r}), \tag{7}$$

where λ is known as the Lagrange multiplier which satisfies $R(\mathbf{r}^*(\lambda)) \leq R_{\max}$. Substituting (4) and (6) into (7), we have

$$\min_{\boldsymbol{\alpha}\in\mathcal{A}'} \min_{\mathbf{r}\in\mathcal{Q}} \quad \frac{1}{N_f}\sum_{k=1}^{N_f} d(\alpha_k, r_k) + \lambda\, \frac{1}{N_f}\sum_{k=1}^{N_f} r_k. \tag{8}$$

As the optimization objective function is separable across frequency, the problem can be further simplified to

$$\frac{1}{N_f}\sum_{k=1}^{N_f}(\min_{r_k\in\mathcal{Q}_k} \min_{\alpha_k\in\mathcal{A}} (d(\alpha_k, r_k)) + \lambda\, r_k). \tag{9}$$

After optimizing over α_k the problem can finally be reformulated as

$$\frac{1}{N_f}\sum_{k=1}^{N_f}(\min_{r_k\in\mathcal{Q}_k} \quad d^\star(r_k) + \lambda\, r_k), \tag{10}$$

where $d^\star(r_k)$ is the minimum distortion per frequency with respect to the best strategy candidate choices, for a given rate r_k. Notice that for small N_A, $d^\star(r_k)$ can be found with exhaustive search. The final minimization problem can be solved by finding the operating point in the D-R curve which intersects first by the constant slope line $d_k + \lambda\, r_k = b$ with $b > 0$, for each frequency bin k [12]. This is illustrated in Figure 1. Alternatively, for small N_r, the best r_i values can be found by exhaustive search. The final step is to find a "good" λ satisfying the total rate budget constraint by iterating the same procedure in (10). For convex D-R relations, finding the optimal λ can be done using bisection algorithms [12][14]. However, as the D-R relations are not always convex, we use the method described in [11] (Variant 2) with a modified initialization formula, which is given by

$$\lambda^0 = \frac{1}{N_f}\sum_{k=1}^{N_f}[d^\star(\min(R_{\max}, q_k - 1)) - d^\star(\min(R_{\max}, q_k - 1) + 1)], \tag{11}$$

where λ^0 is the initial LM value, given a total rate budget $R_{\max}$ and q_k is the maximum operating rate at a particular frequency. More details about the method can be found in [11].

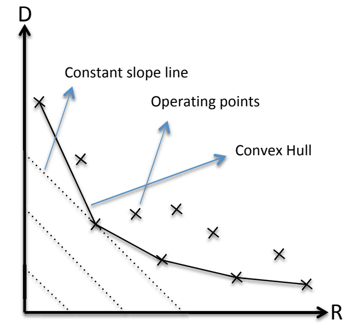

Fig. 1: Geometric interpretation of the problem in (10)

IV. Quantization aware MWF beamforming

In this section, we describe an application of the presented theory to rate-constrained MWF beamforming using uniform quantizers in a small WASN.

Let us assume the left side HA acts as an FC. The goal is to estimate the target signal S at the left reference microphone, given local (left-side) information and remote quantized signals from other microphones. The remote signals are quantized through uniform quantization as follows. The signal x is quantized, and the quantized signal is denoted by $\tilde{x}$. Therefore, under certain assumptions [15][16], the quantization error $e = x - \tilde{x}$ is uniformly distributed with variance $\sigma_e^2 = \frac{\Delta^2}{12}$, where $\Delta = \frac{2x_{\max}}{2^R}$ is a step size, which depends on the range of the signal (maximum value $x_{\max}$) and the quantization rate R.

Let $\tilde{\mathbf{y}}_{\text{rem}}$ denote the concatenation of the STFT coefficients obtained from the quantized and transmitted remote microphone signals. The vector $\tilde{\mathbf{y}}_{\text{rem}}$ is then combined with the local information $\mathbf{y}_{\text{loc}}$ to construct the total observation vector $\mathbf{y}_{\text{tot}} = [\mathbf{y}_{\text{loc}}^{\mathrm{T}} \tilde{\mathbf{y}}_{\text{rem}}^{\mathrm{T}}]^{\mathrm{T}}$. Finally, using the MWF beamformer, the estimated signal per frequency is given by $\hat{S} = \mathbf{w}^{\mathrm{H}}\mathbf{y}_{\text{tot}}$, where $\mathbf{w}$ denotes the vector of optimal Wiener filter coefficients. The PSD of the MWF estimation error (for a particular frequency bin) is then given by

$$d(S,\hat{S}) = E[|S-\hat{S}|^2] = \Phi_S - \mathbf{\Phi}_{S\mathbf{y}_{\text{tot}}}\mathbf{\Phi}_{\mathbf{y}_{\text{tot}}}^{-1}\mathbf{\Phi}_{\mathbf{y}_{\text{tot}}S}, \tag{12}$$

where $\mathbf{\Phi}_{S\mathbf{y}_{\text{tot}}} = \mathbf{\Phi}_{\mathbf{y}_{\text{tot}}S}^{\mathrm{H}}$ denotes the cross PSD vector between the target signal S and $\mathbf{y}_{\text{tot}}$, and $\mathbf{\Phi}_{\mathbf{y}_{\text{tot}}}$ denotes the cross PSD matrix of the vector $\mathbf{y}_{\text{tot}}$. The quantized signal vector $\tilde{\mathbf{y}}_{\text{rem}}$ is actually a function of the chosen strategy. Based on (12), distortions for different strategies and rates are computed. In this paper we consider a particular application of the presented theory, where the possible strategies consist of selection of local/remote signals and different bit-rate allocation schemes among these signals.

V. Experiments

In this section we apply the method proposed in Section III to an example acoustic scene and perform simulations to evaluate the performance.

ISBN 978-90-827970-1-5 © EURASIP 2018

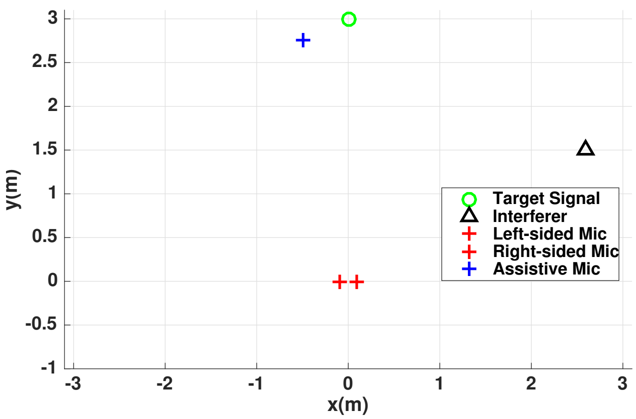

Fig. 2: Generalized binuaral HA setup

A. Setup

The experimental setup is shown in Figure 2. Two red "+" symbols denote two microphones (one microphone per HA) located along the horizontal x-axis at a distance 10 cm from the origin ($(x, y) = (0, 0)$). The target speech signal, shown by the green circle, is located in front of the binaural HA system (at zero degrees) with a distance of 3 m from the origin. In this paper, the angles are computed counter-clockwise and the straight looking direction corresponds to zero degrees. The blue "+" symbol shows the assistive wireless microphone located closer to the target speech signal, at $\theta = 10^\circ$ and a distance of 3 m from the origin, where $\theta = \arctan(\frac{y}{x}) - \frac{\pi}{2}$. The interfering signal, which is denoted by the black triangle, is located at -60° with a distance of 3 m from the origin. The point noise source (interfering signal) has a flat PSD $\Phi_{n_1}(\Omega)$ over the interval $\Omega \in [-\pi, \pi]$. Arround 10 s of the $F_s = 16$ kHz sampled speech of the "CMU-ARCTIC" [17] database are used for the PSD estimation (Φ_S) based on Welch's method. 512 discrete Fourier transform (DFT) coefficients, computed frame-by-frame from 50% overlapping speech frames, are used in the PSD estimation process. The cross PSD matrices are calculated using true ATFs [18] and corresponding estimated PSDs.

The reference microphone is chosen to be the microphone in the left-side HA (the FC). In addition to the target speech signal and the interferer, internal microphone noise is simulated and added, which is assumed to be uncorrelated between microphones. The signal-to-noise ratio (SNR) for the internal noise with respect to the target at the reference microphone is 40 dB. Similarly, the signal-to-interferer ratio (SIR) for the interferer is 0 dB.

B. Strategy Candidate Set for Simulations

Based on the acoustic scene shown in Figure 2, we design the following strategy candidate set:

- 2CH: Rate-constrained MWF beamforming with two microphone signals, i.e., the left side (FC) and the right side microphone signals.

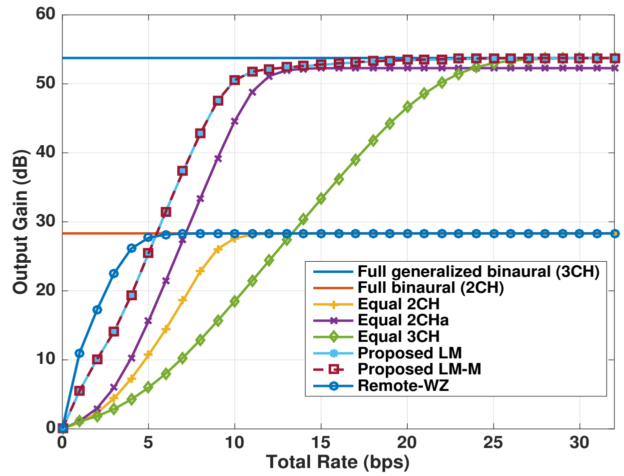

Fig. 3: Output Gain (dB) versus total rate (bit per sample)

- 2CHa: Rate-constrained MWF beamforming with two microphone signals, i.e., the left side and the assistive microphone signals.
- 3CH: Rate-constrained MWF beamforming with all three microphone signals. Note that opposed to the first two strategies, in this strategy multiple remote signals are selected. This implies that the total rate-budget now has to be allocated not only over frequency, but also over the two microphone signals.

When it happens that in one strategy (e.g., the candidate 3CH) there is more than one WASN node for which data needs to be quantized, then the candidate set is extended to cover all relevant rate allocations across microphones.

The number of all possible rate allocations across M microphones given N_r different operating rates ($0 \le r \le N_r - 1$) are computed as

$$|\mathcal{A}|^+ = \binom{M-1}{M-1} + \binom{M}{M-1} + \ldots + \binom{N_r + M - 2}{M-1}. \quad (13)$$

The final set will be the union of the initial strategy set and the set which consists of all combinations across microphones. For example, in the candidate 3CH two quantized signals are transmitted to the FC, i.e., $M = 2$ in (13). In the experiments the same rate range $0 \le r_k \le 32$ is chosen for all frequencies, i.e., $N_r = 33$. Therefore the total number of combinations (strategy choices) will be 561.

C. Evaluation

In this section, we compare variants of the proposed method with methods proposed in the literature. The following methods are compared:

- Full generalized binaural MWF: The MWF with all three microphone signals. This method serves as a performance bound assuming the signals are available at the infinite rate.
- Full binaural MWF: The MWF with both the left and right microphone signals. Similarly, this method serves as a performance bound for the binaural setup.

ISBN 978-90-827970-1-5 © EURASIP 2018

- Equal 2CH: The candidate 2CH. The rates are equally allocated over all frequencies.
- Equal 2CHa: The candidate 2CHa. The rates are equally allocated over all frequencies.
- Equal 3CH: The candidate 3CH. The rates are assumed to be equally allocated over all frequencies as well as across microphones.
- Proposed LM: The proposed method described in Section III. The distortions are computed based on (12), for different algorithm choices and rates. Note that this method optimally allocates the rates over all frequencies, but equally across microphones, when a strategy is selected that involves multiple microphones.
- Proposed LM-modified (LM-M): This method is based on the Proposed LM, and optimally allocates the rates over all frequencies and across microphones, using the extended strategy set described in Section V-B.
- Remote-Wyner-Ziv (WZ) [9]: The binaural rate-constrained beamforming presented in [9]. Note that only two HA microphones can be used in this method, joint statistics are needed at all processors (nodes) and long-block vector quantizers are impractical.

The performance measure is defined as the ratio of the MSE for the monaural configuration, i.e. when there is no communication with the FC, and the MSE achieved by the above-mentioned methods, and is given by

$$G = \frac{D(0)}{D(\boldsymbol{\alpha}, \mathbf{r})}. \tag{14}$$

The vectors $\boldsymbol{\alpha}$ and $\mathbf{r}$ are optimally chosen for the methods ”proposed LM” and ”proposed LM-M”. For the other (reference) methods, $\boldsymbol{\alpha}$ is fixed as no selection is possible. Figure 3 shows the output gains G in dB as a function of the total bit-rate budget ($R_{\max}$). The performance of the 2CH-based methods saturates to that of full binaural MWF, as expected. The performance of the remote-WZ method is computed based on the theoretical upper bound, described in [9]. As shown, the performance curve of the remote-WZ method saturates as the assistive microphone is not considered in this method.

The proposed methods select the best microphone configurations and find optimal rate allocations over frequency. The performance curves of the proposed methods LM and LM-M almost coincide, for this specific scenario, as the proposed optimization problem mostly chooses the 2CH-based candidates at low and middle rates. However, at middle and high rates the proposed methods tend to select the 3CH or 2CHa candidates, and the proposed LM-M method performs slightly better than the LM method, as unequal (efficient) rate allocations are chosen across the right-side and the assistive microphone signals.

VI. Conclusion

In this paper, we proposed an operational rate-distortion based optimization problem for both strategy selection and rate allocation over frequency in (small) WASNs. Unlike existing binaural beamforming algorithms, we considered a potential collaboration between the binaural HAs and some assistive wireless processors in a rate-distortion sense. The sensitivity of existing methods to the acoustic scene is addressed by introducing the strategy candidate set. The proposed framework was applied to the rate-constrained MWF beamforming problem. Assuming uniform quantizers the best microphone configurations and rate allocations were found. The proposed methods were evaluated based on the MSE performance gap between the monaural configuration and the rate-constrained generalized binaural setup. The efficiency of the proposed method is demonstrated in simulation experiments with an example acoustic scene.

References

[1] M. Brandstein and D. Ward (Eds.), *Microphone Arrays*, New York, NY, USA: Springer, 2001.

[2] P. Vary and R. Martin, *Digital Speech Transmission: Enhancement, Coding And Error Concealment*, John Wiley & Sons, 2006.

[3] K. Eneman *et al*, “Evaluation of signal enhancement algorithms for hearing instruments,” in *16th European Signal Processing Conference*, Aug 2008, pp. 1–5.

[4] R. Sockalingam, M. Holmberg, K. Eneroth, and M. Shulte, “Binaural hearing aid communication shown to improve sound quality and localization,” *The Hearing Journal*, vol. 62, no. 10, pp. 46–47, 2009.

[5] T. J. Klasen, T. Van den Bogaert, M. Moonen, and J. Wouters, “Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues,” *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1579–1585, April 2007.

[6] A. I. Koutrouvelis, R. C. Hendriks, R. Heusdens, and J. Jensen, “Relaxed binaural LCMV beamforming,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 1, pp. 137–152, Jan 2017.

[7] E. Hadad, S. Doclo, and S. Gannot, “The binaural LCMV beamformer and its performance analysis,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 3, pp. 543–558, March 2016.

[8] S. Srinivasan and A. den Brinker, “Rate-constrained beamforming in binaural hearing aids,” *EURASIP Journal on Advances in Signal Processing*, pp. 1–9, 2009.

[9] O. Roy and M. Vetterli, “Rate-constrained collaborative noise reduction for wireless hearing aids,” *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 645–657, Feb 2009.

[10] O. Roy and M. Vetterli, “Collaborating hearing aids,” *in Proceedings of MSRI Workshop on Mathematics of Relaying and Cooperation in Communication Networks*, April 2006.

[11] Y. Shoham and A. Gersho, “Efficient bit allocation for an arbitrary set of quantizers,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 9, pp. 1445–1453, Sep 1988.

[12] K. Ramchandran and M. Vetterli, “Best wavelet packet bases in a rate-distortion sense,” *IEEE Transactions on Image Processing*, vol. 2, no. 2, pp. 160–175, Apr 1993.

[13] P. Prandoni and M. Vetterli, “R/D optimal linear prediction,” *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 6, pp. 646–655, Nov 2000.

[14] Z. Xiong, K. Ramchandran, C. Herley, and M. T. Orchard, “Flexible tree-structured signal expansions using time-varying wavelet packets,” *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 333–345, Feb 1997.

[15] A. Sripad and D. Snyder, “A necessary and sufficient condition for quantization errors to be uniform and white,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, no. 5, pp. 442–448, Oct 1977.

[16] S. P. Lipshitz, R. A. Wannamaker, and J. Vanderkooy, “Quantization and dither: A theoretical survey,” *Audio Eng. Soc.*, vol. 40, pp. 355–375, 1992.

[17] J. Kominek, A. W. Black, and V. Ver, “CMU ARCTIC databases for speech synthesis,” Tech. Rep., 2003.

[18] E. A. P. Habets, “Room impulse response generator,” https://www.audiolabs-erlangen.de/fau/professor/habets/software/rir-generator/, 2010.

ISBN 978-90-827970-1-5 © EURASIP 2018