INCORPORATING SPATIAL INFORMATION IN BINAURAL BEAMFORMING FOR NOISE SUPPRESSION IN HEARING AIDS

Wei-Cheng Liao[*], *Mingyi Hong*[♮], *Ivo Merks*[†], *Tao Zhang*[†], *and Zhi-Quan Luo*[*]

[*]Department of Electrical and Computer Engineering
University of Minnesota, Minneapolis, MN, 55455, USA
[♮] Department of Industrial and Manufacturing Systems Engineering
Iowa State University, IA, 50011, USA
[†]Starkey Hearing Technologies, MN, 55344, USA

ABSTRACT

In this paper, we propose a beamforming algorithm for binaural hearing aids with enhanced noise suppression capability. The enhancement is based on incorporating *a priori* spatial information into the conventional multichannel Wiener filtering (MWF) approach for noise suppression. We develop a low complexity algorithm for the resulting quadratically constrained beamforming problem. Through numerical experiments, we demonstrate that the new algorithm can achieve better noise suppression performance than the existing beamforming algorithms under fairly realistic conditions. In addition, we propose two techniques to further reduce the algorithm's computational complexity and the communication overhead between two hearing aids without sacrificing the noise suppression performance.

***Index Terms*—** Binaural signal processing, dual decomposition, coordinate descent, spatial information

1. INTRODUCTION

Multichannel Wiener filtering (MWF) algorithm has been extensively studied for noise suppression in hearing aid design. The objective of MWF is to perform the minimum-mean-square-error (MMSE) estimation of a reference signal. The basic MWF design has been extended to situations involving binaural hearing aids by exploiting the extra degrees of freedom brought by the multiple microphones at both hearing aids (see [1] and references therein). While these algorithms can significantly improve the noise reduction performance of the binaural hearing aids, they inevitably cause undesirable speech distortions [2].

To mitigate speech distortions, speech distortion weighted MWF (SDW-MWF) has been proposed to balance these two design goals with a predetermined trade-off parameter [3]. Alternatively, it has been suggested to explicitly enforce speech distortion requirements using *a priori* acoustic transfer functions (ATFs). ATFs are used in the well-known Minimum Variance Distortionless Response (MVDR) [4] and Linearly Constrained Minimum Variance (LCMV) [5] filter designs which both require zero speech distortion. These filter designs have been used extensively in hearing aid design [6, 7]. However, the requirement of zero speech distortion reduces the amount of noise suppression dramatically. Therefore, it is sometimes preferred to allow a certain limited amount of speech distortion as for example proposed by the parameterized multichannel non-causal Wiener filter (PMWF) design [8]. Unfortunately, it's difficult to select a parameter to achieve the desired trade off between the two design factors in the SDW-MWF design, especially in the presence of multiple speech sources. Moreover, the optimal MMSE filter for these binaural MWF designs requires the signal correlation matrix to be accurately estimated, which is unrealistic in practice.

In this work, we revisit the binaural MWF hearing aid design problem. We incorporate the *a priori* knowledge of approximate ATFs for the signal sources into the design to improve the hearing aid performance. We formulate the design as a quadratically constrained quadratic program (QCQP) [9], explicitly striking a desirable balance between the two design factors. Since the constraints of the formulated QCQP do not depend on the correlation matrix of the signals, the resulting formulation is more robust to the speech nonstationarity. Moreover, we propose an iterative dual decomposition approach [10] to solve the proposed formulation. We also show how to significantly reduce the algorithm's computational complexity and the communication overhead between the hearing aids.

Notations: Boldfaced lowercase (resp. uppercase) letters are used to represent vectors (resp. matrices). The superscripts 'H' stands for the conjugate transpose. The set of all n-dimensional complex vectors are denoted by $\mathbb{C}^n$. We denote $\mathbf{x}_i \in \mathbb{C}$ as the ith element of $\mathbf{x} \in \mathbb{C}^n$, and $\mathbf{x}_{-i} \triangleq [\mathbf{x}_1^H, \ldots, \mathbf{x}_{i-1}^H, \mathbf{x}_{i+1}^H, \ldots, \mathbf{x}_n^H]^H$.

2. SYSTEM MODEL AND PROPOSED BINAURAL HEARING AID DESIGN

We consider a binaural hearing aid, which consists of two M-microphone arrays, one on each hearing aid. The signals are processed in the frequency-domain by transforming the received signals via short time fourier transform (STFT). The received signal for the ith time frame and on frequency band ω is denoted as $\boldsymbol{y}(i,\omega) \in \mathbb{C}^{2M}$, expressed below

$$\boldsymbol{y}(i,\omega) = \mathbf{x}(i,\omega) + \mathbf{v}(i,\omega) \in \mathbb{C}^{2M}. \tag{1}$$

Here $\mathbf{x}(i,\omega) = [(\mathbf{x}(i,\omega)^L)^H, (\mathbf{x}(i,\omega)^R)^H]^H$ and $\mathbf{v}(i,\omega) = [(\mathbf{v}(i,\omega)^L)^H, (\mathbf{v}(i,\omega)^R)^H]^H$ are, respectively, the speech component and the interference component; L (resp. R) labels the signal belonging to the left (resp. right) hearing aid. For notational simplicity, in the following, we will omit the time and frequency indices i and ω; the labels L and R will be used only when necessary.

In this paper, we will focus on the scenario with one desired target source s, multiple directional interference sources $n_\phi, \phi \in \Phi$, and one possibly non-directional interference n. We would like to note that it is possible to use the proposed method with more than

This work is supported by a research gift from Starkey Hearing Technologies.

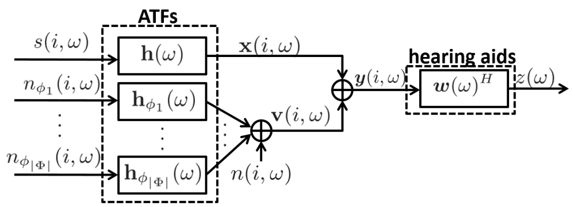

Fig. 1. The schematic of the considered system environment.

one target. The speech component $\mathbf{x}$ and the interference component $\mathbf{v}$ can be expressed as

$$\mathbf{x} = \mathbf{h}s, \quad \mathbf{v} = \sum_{\phi \in \Phi} \mathbf{h}_\phi n_\phi + n, \tag{2}$$

where $\mathbf{h}$ and $\mathbf{h}_\phi$ are the corresponding ATFs for s and n_ϕ, respectively. These notations are summarized in Fig. 1.

The hearing aids apply a receive beamformer $\boldsymbol{w}$ to linearly combine the received signal $\boldsymbol{y}$, i.e., the processed signal $z \in \mathbb{C}$ can be represented as

$$z = \boldsymbol{w}^H \boldsymbol{y} = \underbrace{\boldsymbol{w}^H \mathbf{x}}_{\text{desired speech}} + \underbrace{\boldsymbol{w}^H \mathbf{v}}_{\text{interferences}} . \tag{3}$$

The proposed binaural beamforming design tries to exploit the *a priori* knowledge of (approximate) ATFs and balances the following three design factors:

a) Reduce interference energy: The first requirement is that the variance of the interferences should be minimum, i.e., the filter should seek to solve

$$\min_{\boldsymbol{w}} \mathrm{E}_{\mathbf{v}}[|\boldsymbol{w}^H \mathbf{v}|^2] \equiv \min_{\boldsymbol{w}} \boldsymbol{w}^H \mathbf{R}_v \boldsymbol{w}, \tag{4}$$

where $\mathbf{R}_v \triangleq \mathrm{E}[\mathbf{v}\mathbf{v}^H]$ is the correlation matrix for both the directional and non-directional interference components.

b) Prevent speech distortion (SD) for s: the idea is to use the (approximate) ATFs to improve the SD performance. To this end, assume that a set of approximate ATFs for $\mathbf{h}$, denoted as $\{\mathbf{h}_\theta \in \mathbb{C}^{2M} | \ \theta \in \Theta\}$, is available. The following SD constraints are therefore enforced:

$$|\boldsymbol{w}^H \mathbf{h}_\theta - h_{\theta,\mathrm{ref}}|^2 \leq \epsilon_\theta |h_{\theta,\mathrm{ref}}|^2, \ \forall\, \theta \in \Theta, \tag{5}$$

where $h_{\theta,\mathrm{ref}} \in \mathbb{C}$ is the ATF of the reference microphone for $\mathbf{h}_\theta$, and ϵ_θ is the prespecified tolerable SD. Note that the set of ATFs may be obtained by direct measurement in an anechoic room with for different known directions of arrival (DOA).

c) Interference suppression: The last design criterion ensures that the strength amplification of the ϕ-th interference should not exceed the specified threshold ϵ_ϕ

$$|\boldsymbol{w}^H \mathbf{h}_\phi|^2 \leq \epsilon_\phi, \ \forall\, \phi \in \Phi. \tag{6}$$

In summary, the optimization-based hearing aid design considered in this paper can be formulated as follows:

$$\begin{aligned} \boldsymbol{w}^\star = \arg\min_{\boldsymbol{w}} \ & \boldsymbol{w}^H \mathbf{R}_v \boldsymbol{w} \\ \text{subject to } & (5) \text{ and } (6). \end{aligned} \tag{7}$$

Remark 1 *When constraint* (6) *is ignored, i.e., no constraints are put on the directional sources, and* $|\Theta| = 1$*, problem* (7) *reduces to the PMWF hearing aid design. Furthermore, if* $\epsilon_\theta = 0$*, problem* (7) *becomes the classical MVDR hearing aid design.*

Problem (7) is a QCQP problem, which can be solved using general purpose interior point solver [11]. However the associated computational complexity is too large for implementation in hearing aids. In the following section, we introduce an iterative algorithm for solving (7) whereby each update step is low-complexity and in closed-form. This algorithm can be further simplified by reducing the communication overhead between hearing aids and the computation complexity.

3. PROPOSED EFFICIENT DUAL DECOMPOSITION APPROACH ALGORITHM

In this section, the dual decomposition technique from convex optimization theory will be exploited to solve our design problem. First, we write the Lagrangian function for problem (7) as follows:

$$\begin{aligned} L(\boldsymbol{w}, \boldsymbol{\delta}) = \boldsymbol{w}^H \mathbf{R}_v \boldsymbol{w} + \sum_{\phi \in \Phi} \delta_\phi (|\boldsymbol{w}^H \mathbf{h}_\phi|^2 - \epsilon_\phi) \\ + \sum_{\theta \in \Theta} \delta_\theta (|\boldsymbol{w}^H \mathbf{h}_\theta - h_{\theta,\mathrm{ref}}|^2 - \epsilon_\theta |h_{\theta,\mathrm{ref}}|^2) \end{aligned} \tag{8}$$

where $\delta_\theta \geq 0$ and $\delta_\phi \geq 0$ are, respectively, the Lagrangian dual variables for the θ-th and ϕ-th constraint of (5) and (6); $\boldsymbol{\delta}$ is defined as $\boldsymbol{\delta} \triangleq \{\delta_\theta, \delta_\phi | \ \theta \in \Theta, \phi \in \Phi\}$.

Fixing the dual variables $\boldsymbol{\delta}$ and minimizing $L(\boldsymbol{w}, \boldsymbol{\delta})$ over $\boldsymbol{w}$, we obtain the following unconstrained optimization problem

$$\min_{\boldsymbol{w}} L(\boldsymbol{w}, \boldsymbol{\delta}), \tag{9}$$

whose optimal solution $\boldsymbol{w}$ can be obtained in closed-form as follows

$$\boldsymbol{w} = \left(\mathbf{R}_v + \sum_{\phi \in \Phi} \delta_\phi \mathbf{h}_\phi \mathbf{h}_\phi^H + \sum_{\theta \in \Theta} \delta_\theta \mathbf{h}_\theta \mathbf{h}_\theta^H \right)^{-1} \sum_{\theta \in \Theta} \delta_\theta \mathbf{h}_\theta \mathbf{h}_{\theta,\mathrm{ref}}^H . \tag{10}$$

Intuitively, the collection of dual variables $\boldsymbol{\delta}$ serve as the penalty coefficients of the violation for both the constraints. If δ_θ increases, the resulting constraint violation for (5), i.e., $(|\boldsymbol{w}^H \mathbf{h}_\theta - h_{\theta,\mathrm{ref}}|^2 - \epsilon_\theta |h_{\theta,\mathrm{ref}}|^2)$, decreases. The same argument applies to the resulting constraint violation of (6) for δ_ϕ as well. The following proposition follows from the standard convex optimization theory.

Proposition 1 *[12, Prop. 5.3.1]If problem* (7) *is feasible, there exists an optimal* $\boldsymbol{\delta}^\star$ *such that* $\boldsymbol{w}^\star = \arg\min_{\boldsymbol{w}} L(\boldsymbol{w}, \boldsymbol{\delta}^\star)$ *is the optimal receive beamformer for hearing aid design* (7). *Moreover,* $\boldsymbol{\delta}^\star = \arg\max_{\boldsymbol{\delta} \geq \mathbf{0}} L(\boldsymbol{w}^\star, \boldsymbol{\delta})$.

In the following, we describe an iterative gradient ascent algorithm for computing the optimal $\boldsymbol{\delta}^\star$. In particular, let t denote the iteration index, then the gradient directions for $\min_{\boldsymbol{w}} L(\boldsymbol{w}, \boldsymbol{\delta}^{(t)})$ with respect to $\delta_\theta^{(t)}$ and $\delta_\phi^{(t)}$ are, respectively, given by [9]

$$g_\theta^{(t)} = |\boldsymbol{w}^{(t)H} \mathbf{h}_\theta - h_{\theta,\mathrm{ref}}|^2 - \epsilon_\theta |h_{\theta,\mathrm{ref}}|^2, \tag{11a}$$

$$g_\phi^{(t)} = |\boldsymbol{w}^{(t)H} \mathbf{h}_\phi|^2 - \epsilon_\phi, \tag{11b}$$

where $\boldsymbol{w}^{(t)} = \arg\min_{\boldsymbol{w}} L(\boldsymbol{w}, \boldsymbol{\delta}^{(t)})$. Therefore, the dual variables $\boldsymbol{\delta}$ can be updated towards its ascent direction:

$$\delta_\theta^{(t+1)} = \left[\delta_\theta^{(t)} + \alpha^{(t)} g_\theta^{(t)} \right]^+, \ \delta_\phi^{(t+1)} = \left[\delta_\phi^{(t)} + \alpha^{(t)} g_\phi^{(t)} \right]^+, \tag{12}$$

where $[x]^+ \triangleq \max\{0, x\}$ and $\alpha^{(t)} > 0$ is the stepsize for updating the dual variables $\boldsymbol{\delta}$. The dual update procedure (12) also follows the intuition behind the Lagrangian formulation (8). Since the gradient direction $g_\theta^{(t)}$ and $g_\phi^{(t)}$ are the constraint violations for the current iteration, (12) increases the dual variables when the constraint violation is larger than zero, and decreases otherwise. Table 1 summarizes the proposed algorithm for hearing aid design (7). The convergence property of the algorithm is summarized in the result below.

Algorithm 1: Dual decomposition approach for (7):

1: **Initialize** $\boldsymbol{\delta}^{(0)} \geq \mathbf{0}$; set $t = 0$
2: **Repeat**
3: Update beamformer $\boldsymbol{w}^{(t)} = \arg\min_{\boldsymbol{w}} L(\boldsymbol{w}, \boldsymbol{\delta}^{(t)})$ by (10)
4: Update dual variables $\boldsymbol{\delta}^{(t+1)}$ by (12)
5: $t = t + 1$
6: **Until** Desired stopping criteria is met

Table 1. Summary of the proposed dual decomposition algorithm.

Proposition 2 *[12, Chapter 2.3] Assume problem* (7) *is feasible and the step size $\alpha^{(t)}$ satisfies $\alpha^{(t)} = c/t > 0$ where c is a constant. Then $\boldsymbol{w}^{(t)} \to \boldsymbol{w}^{\star}$ as $t \to \infty$.*

We note that later in our numerical experiments, we will focus on a simplified step size rule, i.e., we set $\alpha^{(t)} = c$, $\forall\ t$ for some constant $c > 0$. Since there is no explicit expression c, we have to pick it heuristically. Nevertheless the overall algorithm still achieves promising performance with small number of iterations.

Remark 2 *As a special case, consider $|\Phi| = 0$ and $|\Theta| = 1$ with $\mathbf{h}_\theta = \mathbf{h}$. The update procedure* (10) *reduces to SDW-MWF with parameter $\mu = P_s/\delta_\theta$ where $P_s = \mathrm{E}[|s|^2]$. Algorithm 1 can be viewed as an extension of SDW-MWF that iteratively updates the parameter. Moreover, Algorithm 1 achieves a desired perceptual performance in SD with the constraints* (5) *being explicitly satisfied.*

In the following subsections, we discuss modifications of Algorithm 1 to adapt to certain practical hardware constraints of the hearing aids.

3.1. Communication Overhead Reduction

The first practical constraint is the limited communication capability between two hearing aids. To address this issue, we first make the following practical assumptions: i) both hearing aids have full knowledge of ATFs in (7). ii) the left hearing aid contains the reference microphone; it computes $\boldsymbol{w}$ and forward the resulting $\boldsymbol{w}^R$ to the other side.

Based on these assumptions, the following three quantities need to be communicated between two sides in each frame: i) the computed beamformer $\boldsymbol{w}^R$ from left to right; ii) the processed signal z^L and z^R, exchanged between two sides; iii) the quantities needed to form the covariance $\mathbf{R}_v$, transmitted from right to left. Among these, the beamformer communication will be ignored since it happens less frequently. The processed signal exchange accounts for 2 data streams per frame. To compute $\mathbf{R}_v$, the right hearing aid should transmit $\boldsymbol{y}^R$ each noise-only frame to the left, which accounts for M data streams. In total, there are $2 + M$ data streams being exchanged during each noise-only frame, and 2 data streams during each speech-plus-noise frame.

On the other hand, the constraints (5) and (6), respectively, guarantee low SD for s and certain degree of interference suppression. It is therefore reasonable to obtain satisfactory beamformer design from (7) with only a rough estimate of $\mathbf{R}_v$. To obtain such estimate, we first approximate $\mathbf{R}_v$ to be a block diagonal matrix, i.e.,

$$\mathbf{R}_v = \begin{bmatrix} \mathbf{R}_v^{LL} & \mathbf{R}_v^{LR} \\ (\mathbf{R}_v^{LR})^H & \mathbf{R}_v^{RR} \end{bmatrix} \approx \begin{bmatrix} \mathbf{R}_v^{LL} & 0 \\ 0 & \mathbf{R}_v^{RR} \end{bmatrix}, \tag{13}$$

where $\mathbf{R}_v^{LL} = \mathrm{E}[\mathbf{v}^L(\mathbf{v}^L)^H]$, $\mathbf{R}_v^{RR} = \mathrm{E}[\mathbf{v}^R(\mathbf{v}^L)^R]$, and $\mathbf{R}_v^{LR} = \mathrm{E}[\mathbf{v}^L(\mathbf{v}^R)^H]$. Here $\mathbf{R}_v^{LL}$ and $\mathbf{R}_v^{RR}$ can be estimated locally at each hearing aid without exchanging the received signals, i.e., $\boldsymbol{y}^R$, during each noise-only frame. Instead, the right hearing aid can update the current $\mathbf{R}_v^{RR}$ to the left hearing aid less frequently, say every W noise-only frames. Hence, the communication overhead is decreased from $2 + M$ to $2 + M(M+1)/(2W)$.

Algorithm 2: Coordinate descent procedure for (10):

Initialize $\boldsymbol{w}^{(t,T)}$ with $T = 0$
For $T = 1 \sim T^{\max}$
$\boldsymbol{w}^{(t,T)} = \boldsymbol{w}^{(t,T-1)}$
For $i = 1 \sim 2M$
Update $\boldsymbol{w}_i^{(t,T)}$ by (14)

Table 2. Summary of the proposed coordinate descent procedure for updating $\boldsymbol{w}^{(t)}$.

3.2. Low Complexity Implementation

The second issue we address is the limited computational capability of each hearing aid. To this end, we will first analyze the computational complexity order of Algorithm 1. For the beamformer update (i.e., Step 3), the computational complexity is $O((2M)^3 + (2M)^2(|\Theta| + |\Phi| + 1))$, while that for the dual update (i.e., Step 4) is $O(2M(|\Theta| + |\Phi|))$. Therefore, per dual decomposition iteration, the total computational complexity is in the order of $(2M)^3 + (2M)^2(|\Theta| + |\Phi| + 1) + 2M(|\Theta| + |\Phi|)$.

In the following, we propose a low complexity procedure based on the so-called coordinate descent (CD) method [12, 13]. Specifically, the beamformer update will be approximated since it contains the main computational overhead (in the order of $(2M)^3$, incurred by the matrix inversion in (10)). First, we observe that by fixing $\boldsymbol{\delta}^{(t)}$, the objective function $L(\boldsymbol{w}, \boldsymbol{\delta}^{(t)})$ is convex, continuous, and differentiable with respect to $\boldsymbol{w}$. Secondly, for each $i = 1 \sim 2M$, the problem $\min_{\boldsymbol{w}_i} L(\boldsymbol{w}, \boldsymbol{\delta}^{(t)})$ has an unique optimal solution shown below:

$$\boldsymbol{w}_i = ([\mathbf{R}_v]_{ii} + \sum_{\theta\in\Theta} \delta_\theta |\mathbf{h}_{\theta,i}|^2 + \sum_{\phi\in\Phi} \delta_\phi |\mathbf{h}_{\phi,i}|^2)^{-1} \Bigg[\sum_{\theta\in\Theta} \delta_\theta \mathbf{h}_{\theta,i} \mathbf{h}_{\theta,\mathrm{ref}}^H - \left(\mathbf{R}_{v,(-i)i}^H + \sum_{\theta\in\Theta} \delta_\theta \mathbf{h}_{\theta,i} \mathbf{h}_{\theta,-i}^H + \sum_{\phi\in\Phi} \delta_\phi \mathbf{h}_{\phi,i} \mathbf{h}_{\phi,-i}^H \right) \boldsymbol{w}_{-i} \Bigg], \tag{14}$$

where $[\mathbf{R}_v]_{ii}$ is the ith diagonal element of $\mathbf{R}_v$ and $\mathbf{R}_{v,(-i)i} \triangleq \mathrm{E}[\mathbf{v}_{-i}\mathbf{v}_i^H]$. With these properties, we propose to apply the coordinate descent procedure that cyclically update each element of $\boldsymbol{w}$; see Table 2 for detailed description of the algorithm. The optimality of Algorithm 2 is shown via the following proposition:

Proposition 3 *[13, Thm. 4.1] For the convex function $f(\mathbf{x}) : \mathbb{C}^M \to \mathbb{R}$, assume the following condition holds: f is continuous and differentiable for $\mathbf{x}$. Then $\mathbf{x}$ obtained by the coordinate descent procedure with cyclic update rule converges to the optimal $\mathbf{x}^\star = \arg\min_{\mathbf{x}} f(\mathbf{x})$.*

As a result of Proposition 3, Algorithm 2 guarantees that $\boldsymbol{w}^{(t,T^{\max})} \to \boldsymbol{w}^{(t)}$ as $T^{\max} \to \infty$. Moreover, all terms in (14) except $\boldsymbol{w}_{-i}$ can be precomputed and the corresponding computational complexity is $O((|\Theta| + |\Phi|)(2M)^2)$. On the other hand, for updating $\boldsymbol{w}_i$, the computational complexity is $O(2M)$. Hence, the total computational complexity becomes $O((|\Theta| + |\Phi| + T^{\max})(2M)^2)$. In the numerical experiments in Sec. 4, $T^{\max}$ is set to be 1. Although this setting cannot guarantee the optimal convergence, we will see that little performance loss is incurred. Most importantly, it results in a great computational complexity reduction for beamformer update from $O((2M)^3 + (2M)^2(|\Theta| + |\Phi| + 1))$ to $O((|\Theta| + |\Phi| + 1)(2M)^2)$.

4. NUMERICAL EXPERIMENTS

In this section, the performance of the proposed algorithms is demonstrated and compared with that of the binaural hearing aid designs SDW-MWF (with the parameter $\mu = 1$) and the MVDR

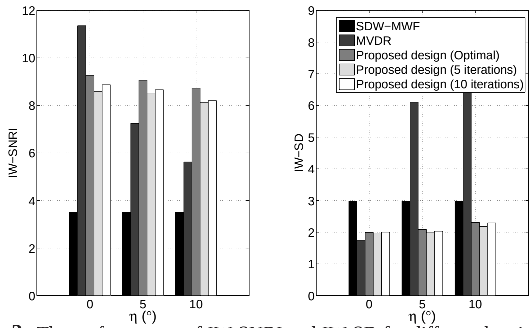

Fig. 2. The performances of IW-SNRI and IW-SD for different hearing aid designs with batch correlation matrix update.

that minimizes the output noise power, i.e., $\boldsymbol{w}^H \mathbf{R}_v \boldsymbol{w}$, with zero SD constraint.

We assume the person wearing hearing aids is located in the center of a 6m×6m×4m room. All the sources are 1m away from the hearing aids. The ATFs from sources to microphones are generated by the image method [14], and the head is modeled as a rigid sphere [15]. The desired speech comes from the direction 0° while 4 directional interference sources come respectively from 70°, 150°, 210°, and 290°, and no non-directional interference exists. The desired target source is a 23s signal with six different sentences, among which 3 sentences are from male speakers and 3 sentences are from female speakers. All of the sources are sampled at 16kHz, and they are taken from TIMIT database [16]. Additionally, there is a 0.5s silence period between each sentence. Each directional interference source consists of nonstop speech. Each hearing aid has 3 microphones ($M = 3$), and uses a 256-point FFT. The ATFs of anechoic room for the hearing aids are assumed to be known. In the sequel, we let $\Theta = \{\eta - 10^\circ, \eta - 5^\circ, \eta^\circ, \eta + 5^\circ, \eta + 10^\circ\}$ and $\Phi = \{80^\circ, 158^\circ, 202^\circ, 280^\circ\}$. Here η is the direction of ATF for MVDR. The proposed algorithms are tested under reverberant scenario with $T_{60} = 200$ms and the input SNR is -5dB. That means the *a priori* knowledge of ATFs are inaccurate due to both source direction estimation error and reverberation. The noise-only frames are used to estimate the correlation matrix of noise $\mathbf{R}_v$. The correlation matrix for the signal $\mathbf{R}_x = \mathbf{R}_y - \mathbf{R}_v + \epsilon\mathbf{I}$ can then be estimated from the signal-plus-noise interval, and $\epsilon \geq 0$ is chosen such that $\mathbf{R}_x$ is positive semi-definite. The parameters of problem (7) is chosen as $\epsilon_\Theta = \{(0.16, 0.09, 0.04, 0.09, 0.16)10^{0.5}\}$ and $\epsilon_\phi = 9 \times 10^{-3.5}$, $\forall\ \phi \in \Phi$, where SNR is the input signal and noise ratio (SNR). Since the coefficients of ATFs are small, The fixed step size $\alpha^{(t)}$, $\forall\ t$, of Algorithm 1 is set to as large as 5×10^6.

In Fig. 2, the two performance metrics, intelligibility-weighted SNR improvement (IW-SNRI) and intelligibility-weighted spectral distortion (IW-SD) [17], for different hearing aid designs are compared. For this set of experiments, batch correlation matrix estimation for $\mathbf{R}_v$ and $\mathbf{R}_x$ is used. We can easily observe that the proposed hearing aid design outperforms the other benchmark designs in terms of both metrics when η is inaccurately estimated. The advantage over MVDR hearing aid design in IW-SNRI comes from introducing tolerable amount of degradation in IW-SD, which can be controlled by parameter ϵ_Θ, and the extra approximated ATFs of the desired speech. Furthermore, the optimal solution of design (7) can be well approximated by running Algorithm 1 with only $5 \sim 10$ iterations.

Because SDW-MWF is sensitive to the VAD errors [3], Fig. 3 investigates the performance of the algorithms as function of miss rate with a constant false alarm rate of 30%. One can observe that the proposed design is not as sensitive as SDW-MWF. On the other hand, since MVDR also exploits the a priori spatial information, it is not affected by this error as much as SDW-MWF if η is appropriately approximated, i.e., Type 1 and 2. Note that as η increases, i.e., increasingly inaccurate DOA estimation of the desired speech, for MVDR, both performance metrics degrade dramatically.

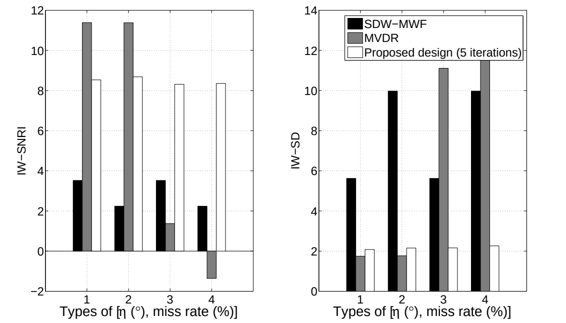

Fig. 3. The performances of IW-SNRI and IW-SD for different hearing aid designs with batch correlation matrix update and a constant false alarm rate of 30%. Here types 1∼4, respectively represent $(0^\circ, 10\%)$, $(0^\circ, 30\%)$, $(10^\circ, 0\%)$, and $(10^\circ, 30\%)$.

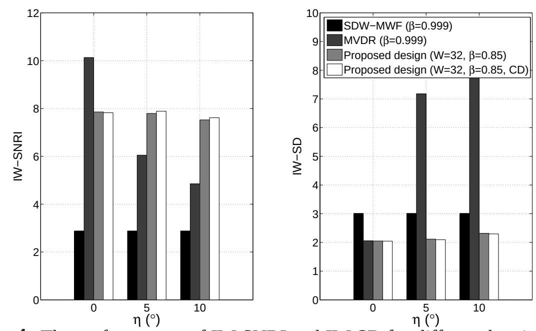

Fig. 4. The performances of IW-SNRI and IW-SD for different hearing aid designs with iterative correlation matrix update.

In the third set of experiments, the correlation matrix estimation for $\mathbf{R}_v$ and $\mathbf{R}_x$ is iterative updated instead. Specifically,

$$\mathbf{R}_v(i+1, \omega) = \beta \mathbf{R}_v(i, \omega) + (1-\beta)\mathbf{v}(i, \omega)\mathbf{v}(i, \omega)^H, \qquad (15)$$

and $\mathbf{R}_y(i, \omega)$ is obtained similarly. One iteration of Algorithm 1 is applied for every correlation matrix update. The two approximation techniques, i.e., low communication overhead scheme with $W = 36$ (Sec. 3.1) and low complexity Algorithm 2 ($T^{\max} = 1$) for updating beamformer $\boldsymbol{w}$ step in Algorithm 1, are also applied. This setting suggests that the communication overhead for the proposed approach is around 2.03 data streams per frame, which is reduced from the original 2.54 data streams per frame. Moreover, for the existing SDW-MWF and MVDR, extra communication overhead is needed for estimating $\mathbf{R}_y$ and $\mathbf{R}_v$. In Fig. 4, the performance of IW-SNRI and IW-SD for all hearing aid designs is compared. The relative performance of different algorithms is similar to that of the first experiment even with the fact that Algorithm 2 solves the hearing aid design (7) inexactly.

5. CONCLUSIONS

In summary, this paper proposes an extension of the MWF algorithm that exploits the *a priori* knowledge of spatial information. The algorithm uses the dual decomposition method to achieve more noise reduction with low distortion while being computational efficient and only requiring a low communication overhead. One future direction is extending the proposed noise suppression hearing aid design such that the binaural cue can also be preserved, e.g., [18].

6. REFERENCES

[1] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic beamforming for hearing aid applications," *Handbook on Array Processing and Sensor Networks*, pp. 269–302, 2009.

[2] J. Chen, J. Benesty, Y. Huang, and S. Doclo, "New insights into the noise reduction wiener filter," *IEEE Trans. Audio, Speech, and Language Process.*, vol. 14, no. 4, pp. 1218–1234, Jul. 2006.

[3] S. Doclo, M. Moonen, T. V. den Bogaert, and J. Wouters, "Reduced-bandwidth and distributed mwf-based noise reduction algorithms for binaural hearing aids," *IEEE Trans. Audio, Speech, and Language Process.*, vol. 17, no. 1, pp. 38–51, Jan. 2008.

[4] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.

[5] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1969.

[6] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, no. 1, pp. 27–34, Jan. 1982.

[7] S. Gannot and I. Cohen, *Adaptive beamforming and postfiltering*, Springer Handbook of Speech Processing, J. Benesty, Y. Huang, and M. M. Sondhi, Eds. New York: Springer-Verlag, 2007.

[8] M. Souden, J. Benesty, and S. Affes, "On optimal frequency-domain multichannel linear filtering for noise reduction," *IEEE Trans. Audio, Speech, and Language Process.*, vol. 18, no. 2, pp. 260–276, Feb. 2010.

[9] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[10] D. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.

[11] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," http://stanford.edu/∼boyd/cvx, June 2009.

[12] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1995.

[13] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *Journal of Optimization Theory and Applications*, vol. 109, no. 3, pp. 475–494, Jun. 2001.

[14] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating smallroom acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

[15] P. M. Morse, *Section 29 "Scattering of a sound" and Section 27 "Radiation from spheres" in Vibration and Sound*, Acoustical Society of America, 3rd edition, 1986.

[16] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, and N. L. Dahlgren, "Darpa timit acoustic phonetic continuous speech corpus," 1993.

[17] A. Spriet, M. Moonen, and J. Wouters, "Robustness analysis of multichannel wiener filtering and generalized sidelobe cancellation for multimicrophone noise reduction in hearing aid applications," *IEEE Trans. Speech and Audio Process.*, vol. 13, no. 4, pp. 487–503, Jul. 2005.

[18] D. Marquardt, E. Hadad, S. Gannot, and S. Doclo, "Optimal binaural LCMV beamformers for combined noise reduction and binaural cue preservation," in *Proc. of International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2014, pp. 289–293.