PARETO OPTIMAL BINAURAL MVDR BEAMFORMER WITH CONTROLLABLE INTERFERENCE SUPPRESSION

Elior Hadad[⋆] *Simon Doclo*[†] *Sven Nordholm*[‡] *Sharon Gannot*[⋆]

[⋆]Bar-Ilan University, Israel; email: {elior.hadad,sharon.gannot}@biu.ac.il
[†]University of Oldenburg, Germany; email: simon.doclo@uni-oldenburg.de
[‡]Curtin University, Australia; email: s.nordholm@curtin.edu.au

2022 International Workshop on Acoustic Signal Enhancement (IWAENC) | 978-1-6654-6867-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/IWAENC53105.2022.9914759

ABSTRACT

The objective of binaural multi-microphone speech enhancement algorithms can be viewed as a multi-criteria design problem as there are several requirements to be met. When applying distortionless beamforming, it is necessary to suppress interfering sources and ambient background noise, and to extract an undistorted replica of the target source. In the binaural versions, it is also important to preserve the binaural cues of the target and the interference sources. In this paper, we propose a unified Pareto optimization framework for binaural distortionless beamformers, which is achieved by defining a multi-objective problem (MOP) to control the amount of interference suppression and noise reduction simultaneously. The derivation is given for the multi-interference case by introducing separate mean squared error (MSE) cost functions for each of the respective interference sources and the background noise. A Pareto optimal set of solutions is provided for any set of parameters. The performance of the proposed method in a noisy and reverberant environment is presented, demonstrating the impact of the trade-off parameters using real-signal recordings.

Index Terms— MVDR beamforming, Pareto optimization, binaural cues, noise reduction, hearing aids.

1. INTRODUCTION

Binaural hearing aid devices consisting of a hearing aid mounted on each ear of a hearing-impaired person, are known to outperform their monaural counterparts in terms of noise reduction performance and their capability to preserve the binaural cues and, consequently, the spatial impression of the acoustical scene [1]. For directional sources, preservation of the interaural level difference (ILD) and the interaural time difference (ITD) cues can be achieved by preserving the so-called relative transfer function (RTF), which is defined as the ratio of the acoustical transfer functions relating the source and the two ears.

In the last decade, several binaural speech enhancement algorithms that aim to preserve the binaural cues have been developed [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. If the target needs to be processed without distortion, the binaural minimum variance distortionless response (BMVDR) beamformer can be applied [13, 5]. However, a major drawback of the BMVDR beamformer is that the binaural cues of the noise are not preserved. To control both the suppression and the binaural cue preservation of directional interfering sources, the binaural linearly constrained minimum variance (BLCMV) beamformer was proposed in [4, 10]. In the BLCMV criterion, a hard constraint controlling the amount of interference reduction was added to the BMVDR cost function. It was shown that the BLCMV beamformer is able to preserve the binaural cues of both the target and interfering sources. Several extensions of the BLCMV were proposed in [5, 7, 11, 12], for controlling the binaural cue preservation of both the target and interfering sources. The objective of the speech enhancement algorithm can be viewed as a multi-criteria design problem as there are several requirements to be met: extraction of the target speaker without distortion and suppression of interfering sources and ambient background noise, while preserving the binaural cues of both the target and the undesired sound sources.

In this paper, we propose a unified Pareto optimization framework for the aforementioned binaural distortionless beamformers, which is achieved by defining a MOP to simultaneously control the amount of interference suppression and noise reduction. The derivation is given for the multi-interference case by introducing a separate MSE cost function for each of the respective interference sources and the background noise. Under this new formulation, a set of Pareto optimal solutions is given, instead of a single solution that optimizes a specific objective. In addition, we provide two types of trade-off parameters, namely, scaling and weighting parameters. The scaling parameter determines the respective MOP to be optimized, while the weighting parameters are used to select a preferred solution from a set of Pareto optimal solutions in a, so called, decision-making procedure.

The proposed MOP takes into account interference reduction (IR) and noise reduction (NR) cost functions, such that the optimization does not consider binaural cue preservation as an explicit requirement. The cue preservation is determined by the controlling parameter. Specific parameter setting can lead to binaural cue preservation as a by-product, e.g., the setting of an identical scaling parameters for both the left and right filters, and the preferred solution on the optimal Pareto frontier can be selected in the aforementioned decision-making procedure. The Pareto formulation of the binaural multichannel Wiener filter (MWF), in the case of a single interference source, is introduced in [14].

2. PROBLEM FORMULATION

We consider an acoustic scenario comprising a single target speaker and N_u competing interference speakers in a noisy and reverberant environment. The binaural hearing device, consists of two hearing aids equipped with $M = M_L + M_R$ microphones (i.e., M_L microphones on the left hearing aid and M_R microphones on the right hearing aid). All microphone signals can be stacked in the M-dimensional vector $\boldsymbol{y}(\omega)$ in the frequency domain as

$$\boldsymbol{y}(\omega) = \boldsymbol{x}(\omega) + \sum_{i=1}^{N_u} \boldsymbol{u}_i(\omega) + \boldsymbol{n}(\omega) = \boldsymbol{x}(\omega) + \boldsymbol{v}(\omega), \quad (1)$$

with

$$\boldsymbol{y}(\omega)=\left[y_1^L(\omega),\ldots,y_{M_L}^L(\omega),\ y_1^R(\omega),\ldots,y_{M_R}^R(\omega)\right]^T,\quad (2)$$

$\boldsymbol{x}(\omega)$ the target source component, $\boldsymbol{u}_i(\omega)$ the ith interfering source component, $\boldsymbol{n}(\omega)$ the additional background noise, and $\boldsymbol{v}(\omega)=\sum_{i=1}^{N_u}\boldsymbol{u}_i(\omega)+\boldsymbol{n}(\omega)$ the overall noise component. $\boldsymbol{x}(\omega)$, $\boldsymbol{u}_i(\omega)$, $\boldsymbol{n}(\omega)$, and $\boldsymbol{v}(\omega)$ are defined similarly to $\boldsymbol{y}(\omega)$. The variable ω is henceforth omitted for brevity. We can further write $\boldsymbol{x}=S_{\mathrm{x}}\boldsymbol{a}$ and $\boldsymbol{u}_i=S_{\mathrm{u}}^i\boldsymbol{b}_i$, where S_{x} and S_{u}^i are the target and the ith interfering source signals and $\boldsymbol{a}$ and $\boldsymbol{b}_i$ are the acoustic transfer functions (ATFs) relating the target and the ith interfering source positions and the microphones, respectively. Assuming the directional sources and the noise are uncorrelated, the spatial correlation matrix of the noisy microphone signals can be written as

$$\boldsymbol{R}_Y=\boldsymbol{R}_X+\sum_{i=1}^{N_u}\boldsymbol{R}_U^i+\boldsymbol{R}_N,\quad (3)$$

where $\boldsymbol{R}_X=\mathcal{E}\left\{\boldsymbol{x}\boldsymbol{x}^H\right\}=P_X\boldsymbol{a}\boldsymbol{a}^H$, $\boldsymbol{R}_U^i=\mathcal{E}\left\{\boldsymbol{u}_i\boldsymbol{u}_i^H\right\}=P_U^i\boldsymbol{b}_i\boldsymbol{b}_i^H$, and $\boldsymbol{R}_N=\mathcal{E}\left\{\boldsymbol{n}\boldsymbol{n}^H\right\}$ are the target source, the ith interfering source, and the noise correlation matrices, respectively. $\mathcal{E}\{\cdot\}$ denotes the expectation operator and $P_X=\mathcal{E}\left\{|S_{\mathrm{x}}|^2\right\}$ and $P_U^i=\mathcal{E}\left\{|S_{\mathrm{u}}^i|^2\right\}$ denote the power spectral densities (PSDs) of the target source and the ith interfering source, respectively. Without loss of generality, the first microphone on the left and the right hearing aid are selected as the reference microphones, i.e.,

$$z_L(t,k)=\boldsymbol{e}_L^H\boldsymbol{y}(t,k),\quad z_R(t,k)=\boldsymbol{e}_R^H\boldsymbol{y}(t,k),\quad (4)$$

where $\boldsymbol{e}_L$ and $\boldsymbol{e}_R$ are M-dimensional vectors with '1' in the left and right reference microphones, and '0' elsewhere. The output signals at the left and the right hearing devices are given by $z_L=\boldsymbol{w}_L^H\boldsymbol{y}$ and $z_R=\boldsymbol{w}_R^H\boldsymbol{y}$, respectively, where $\boldsymbol{w}_L$ and $\boldsymbol{w}_R$ denote M-dimensional complex-valued weight vectors. Furthermore, we define the $2M$-dimensional stacked weight vector as $\boldsymbol{w}=\left[\begin{array}{cc}\boldsymbol{w}_L & \boldsymbol{w}_R\end{array}\right]^T$.

The input RTFs of the target and the ith interfering source between the reference microphones on the left and the right hearing aid are defined as the ratio of the ATFs, i.e.,

$$\mathrm{RTF}_{X,\mathrm{IN}}=\frac{a_L}{a_R},\qquad \mathrm{RTF}_{U,\mathrm{IN}}=\frac{b_{i,L}}{b_{i,R}}.\quad (5)$$

The output RTFs of the target source and the ith interfering source are defined as the ratio of the filtered components on the left and the right hearing aid, i.e.,

$$\mathrm{RTF}_{X,\mathrm{OUT}}=\frac{\boldsymbol{w}_L^H\boldsymbol{a}}{\boldsymbol{w}_R^H\boldsymbol{a}},\quad \mathrm{RTF}_{U,\mathrm{OUT}}=\frac{\boldsymbol{w}_L^H\boldsymbol{b}_i}{\boldsymbol{w}_R^H\boldsymbol{b}_i}.\quad (6)$$

The binaural ILD and ITD cues can be computed from the complex-valued frequency-dependent RTF, i.e., [15]

$$\mathrm{ILD}=20\log_{10}(|\mathrm{RTF}|),\qquad \mathrm{ITD}=\frac{\angle(\mathrm{RTF})}{\omega},\quad (7)$$

with $\angle$ denoting the phase.

3. THE MULTI OPTIMIZATION PROBLEM

In this section, the multi objective problem is described. First, the mathematical foundations of the MOP are derived for the distortionless family of binaural beamformers. Then, the filter decomposition is derived, constituting a Pareto optimal set of distortionless filters solving the MOP.

3.1. Pareto Binaural MVDR MOP

In this section, we will introduce a family of binaural distortionless beamformers that are able to extract the target source without distortion such that they preserve the binaural cues of the target source. These beamformers must satisfy the following constraint set for the target source

$$\{\boldsymbol{w}\in\mathbb{C}^M:\boldsymbol{w}_L^H\boldsymbol{a}=a_L,\boldsymbol{w}_R^H\boldsymbol{a}=a_R\}.\quad (8)$$

The traditional BMVDR reproduces the target source component without distortion, while minimizing the overall noise power, i.e.,

$$\min_{\boldsymbol{w}}\mathcal{E}\left\{\left\|\left[\begin{array}{c}\boldsymbol{w}_L^H\boldsymbol{v}\\ \boldsymbol{w}_R^H\boldsymbol{v}\end{array}\right]\right\|^2\right\}\ \text{s.t. }\boldsymbol{w}_L^H\boldsymbol{a}=a_L,\boldsymbol{w}_R^H\boldsymbol{a}=a_R.\quad (9)$$

This single-objective problem (SOP) leads to a unique optimal filter. In the BMVDR all sound sources are perceived as arriving from the target direction. Therefore, the RTF of the interfering source is typically distorted, which is clearly an undesired phenomenon, since the spatial impression of the acoustic scene is altered. To suppress the interfering source, while preserving its RTF, several SOP extensions of the BMVDR were proposed in [5, 10, 11].

In the current study, we take a different perspective of the problem, in which multiple cost functions are simultaneously minimized, i.e., the interference suppression for each of the interfering sources and the noise reduction are separately controlled. The MOP can be formulated as

$$\underset{\boldsymbol{w}}{\text{minPareto}}\ \mathcal{C}_{\text{Pareto}}(\boldsymbol{w})\quad (10)$$

with

$$\begin{aligned}\mathcal{C}_{\text{Pareto}}(\boldsymbol{w})&=[J_1(\boldsymbol{w}),\cdots,J_I(\boldsymbol{w})]\\ &\text{s.t. }\boldsymbol{w}_L^H\boldsymbol{a}=a_L,\boldsymbol{w}_R^H\boldsymbol{a}=a_R,\end{aligned}\quad (11)$$

where the IR term $J_i(\boldsymbol{w})$, for the ith interfering source is defined as

$$J_i(\boldsymbol{w})=\mathcal{E}\left\{\left\|\left[\begin{array}{c}\eta_i u_{i,L}-\boldsymbol{w}_L^H\boldsymbol{u}_i\\ \eta_i u_{i,R}-\boldsymbol{w}_R^H\boldsymbol{u}_i\end{array}\right]\right\|^2\right\},\quad (12)$$

for $i=1,\ldots,N_u$, $N_u=I-1$ and η_i is defined as the interference scaling parameter for each interference source. In general, we limit $0\le\eta_i\le1$, such that the parameter controls the amount of IR for each respective interfering source. Note that, in order to preserve the binaural cues of the interference sources, in this study the same value was set for both the left and the right filters. Nevertheless, different values can be set [16]. The last element in $\mathcal{C}(\boldsymbol{w})$ is defined as the NR term $J_I(\boldsymbol{w})=J_n(\boldsymbol{w})$ where

$$J_n(\boldsymbol{w})=\mathcal{E}\left\{\left\|\left[\begin{array}{c}\boldsymbol{w}_L^H\boldsymbol{n}\\ \boldsymbol{w}_R^H\boldsymbol{n}\end{array}\right]\right\|^2\right\}.\quad (13)$$

The Pareto cost function $\mathcal{C}(\boldsymbol{w})$, defined in (11), is an $I=N_u+1$ vector-valued global objective function such that it represents a set of criteria. Each element in $\mathcal{C}(\boldsymbol{w})$ is associated with a different local single cost function.

Note that, in an ideal situation, where there is a unique filter that minimizes all criteria, a single solution is obtained, as in the SOP case. The notion of Pareto optimality becomes of paramount importance in cases where a unique solution is not feasible. In these cases, a set of Pareto-optimal solutions that comprises all solutions that individually minimize each cost function as well as the solutions that trade-off these cost functions is obtained. The Pareto optimality can be formulated as follows.

Definition 1: A filter solution $\boldsymbol{w}^* \in \mathcal{W}$ is Pareto optimal solution iff there does not exist another filter solution $\boldsymbol{w} \in \mathcal{W}$, such that $J_i(\boldsymbol{w}) \leq J_i(\boldsymbol{w}^*)$ for all $i = 1, 2, \ldots, I$ and $J_j(\boldsymbol{w}) < J_j(\boldsymbol{w}^*)$ for at least one index j (cost function).

Definition 2: All the Pareto optimal filter solutions solve the MOP and lie on the boundary of the feasible criterion space [17]. The set of Pareto solutions constitutes the Pareto *frontier*.

A filter is a Pareto optimal solution if no other filter exists that improves at least one cost function without leading to a degradation in another cost function. We note that there may be an infinite number of optimal solutions in the Pareto optimal set.

3.2. Pareto filter decomposition

The Pareto set of solutions can be obtained by various means. We use the so-called *scalarization* method, as described in [18], to compute the Pareto optimal set. For this technique, the set of optimal Pareto solutions is calculated by optimizing a single-objective, generalized cost function, defined as a weighted sum of the local cost functions, i.e.,

$$J(\boldsymbol{w}) = \lambda_1 J_1(\boldsymbol{w}) + \ldots + \lambda_I J_I(\boldsymbol{w})$$
$$\text{s.t.} \sum_{i=1}^{I} \lambda_i = 1, \;\; \boldsymbol{w}_L^H \boldsymbol{a} = a_L, \boldsymbol{w}_R^H \boldsymbol{a} = a_R, \tag{14}$$

where $0 < \lambda_i, \; i = 1, 2, \ldots, I$ are defined as the *weighting parameters* that provide a trade-off between the local cost function terms. The minimization of the single-objective generalized cost function $J(\boldsymbol{w})$ is sufficient for finding a Pareto optimal solution if $J(\boldsymbol{w})$ increases monotonically with respect to each cost function [19], such that any filter that solves the generalized cost function lies on the Pareto frontier.

Now, having a set of Pareto optimal solutions, a subsequent *decision-making* procedure is needed to obtain a selected filter from the Pareto optimal set [20]. This can be achieved based on considerations that are independent of the MOP by setting the weighting parameters. It is attractive to collect all the optimal solutions in the Pareto set since it makes the decision making process less complex. Otherwise, we would get only solutions for a single point that cannot be easily compared with other optimal solutions.

The filters minimizing the cost function in (14) can be computed as:[1]

$$\boldsymbol{w}_L = \boldsymbol{R}_\lambda^{-1}\boldsymbol{r}_\text{L} + \boldsymbol{R}_\lambda^{-1}\boldsymbol{C}\left[\boldsymbol{C}^H\boldsymbol{R}_\lambda^{-1}\boldsymbol{C}\right]^{-1}\left(\boldsymbol{g}_\text{L} - \boldsymbol{C}^H\boldsymbol{R}_\lambda^{-1}\boldsymbol{r}_\text{L}\right)$$
$$\boldsymbol{w}_R = \boldsymbol{R}_\lambda^{-1}\boldsymbol{r}_\text{R} + \boldsymbol{R}_\lambda^{-1}\boldsymbol{C}\left[\boldsymbol{C}^H\boldsymbol{R}_\lambda^{-1}\boldsymbol{C}\right]^{-1}\left(\boldsymbol{g}_\text{R} - \boldsymbol{C}^H\boldsymbol{R}_\lambda^{-1}\boldsymbol{r}_\text{R}\right) \tag{15}$$

with

$$\boldsymbol{R}_{\boldsymbol{\lambda}} = \sum_{u=1}^{N_u} \lambda_u \boldsymbol{R}_U^u + \lambda_n \boldsymbol{R}_N, \; \boldsymbol{R}_{\boldsymbol{\eta}} = \sum_{u=1}^{N_u} \lambda_u \eta_u \boldsymbol{R}_U^u \tag{16}$$

with $\lambda_n = \lambda_I$ and $\boldsymbol{r}_\text{L} = \boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{e}_L$ and $\boldsymbol{r}_\text{R} = \boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{e}_R$. The left and right constraint sets are given by

$$\boldsymbol{C}^H\boldsymbol{w}_L = \boldsymbol{g}_\text{L}, \quad \boldsymbol{C}^H\boldsymbol{w}_R = \boldsymbol{g}_\text{R}, \tag{17}$$

where $\boldsymbol{C}$ denotes the constraint matrix, and $\boldsymbol{g}_\text{L}$ and $\boldsymbol{g}_\text{R}$ denote the left and right desired vectors, respectively. For the single distortionless target source beamfomer, the constraint matrix $\boldsymbol{C}$ and the left and the right desired vectors $\boldsymbol{g}_\text{L}$ and $\boldsymbol{g}_\text{R}$ are given by

$$\boldsymbol{C} = \boldsymbol{a}, \quad \boldsymbol{g}_\text{L} = a_L^*, \quad \boldsymbol{g}_\text{R} = a_R^*. \tag{18}$$

The set of filters are referred to as the Pareto-BMVDR.

There are two types of trade-off parameters, the η-based scaling parameters and the λ-based weighting parameters. The scaling parameters, which control the respective ith interference suppression, define a family of MOPs such that these parameters determine the respective required IR while the weighting parameters are responsible for selecting the preferred filter from a set of Pareto solutions.

4. SIMULATIONS WITH NOISY SPEECH SIGNALS

In this section, experimental performance evaluation is provided considering one target source and two interference sources in a noisy and reverberant environment. The evaluation demonstrates the impact of the trade-off parameters on various performance measures, i.e., in terms of the binaural signal-to-interference-and-noise ratio (SINR), signal-to-noise ratio (SNR), and signal-to-interference ratio (SIR) improvements and the binaural cue preservation capabilities.

All experiments were carried out using Oldenburg database [21]. Each of the hearing aids is equipped with two microphones. Binaural Behind-the-Ear Impulse Responses (BTE-IRs) measured on an artificial head in a cafeteria were used to generate the signal components. The target and the first and second interfering sources were located at 135° and a distance of $129\,\text{cm}$, -90° and a distance of $52\,\text{cm}$, and -45° and a distance of $117.5\,\text{cm}$, respectively. Background ambient noise recorded in the Oldenburg cafeteria was added to the speech components. The sampling frequency was $16\,\text{kHz}$. The signals were processed in the STFT domain with 1024 points and 50% overlap. The input SIR and SNR were set to 10 dB and 6 dB, respectively. For the estimation procedure, training segments were used. The correlation matrices of the target and interference sources were estimated during non-concurrent activity of the target and interference speakers. The correlation matrix of the noise was estimated during a segment in which none of the speech sources was active. The target and the ith interfering source RTF vectors were estimated by picking the major generalized-eigenvectors of the generalized eigenvalue decomposition (GEVD) of the target and the respective ith interfering source correlation metrics and the noise correlation matrix.

We define the input and output narrow-band SINR as the ratio of the average input and respectively, output PSDs of the target source and the overall noise components in the left and the right hearing aids, i.e., [5]

$$\text{SINR}^\text{in} = \frac{\boldsymbol{e}_L^H\boldsymbol{R}_X\boldsymbol{e}_L + \boldsymbol{e}_R^H\boldsymbol{R}_X\boldsymbol{e}_R}{\boldsymbol{e}_L^H\boldsymbol{R}_V\boldsymbol{e}_L + \boldsymbol{e}_R^H\boldsymbol{R}_V\boldsymbol{e}_R}, \tag{19}$$

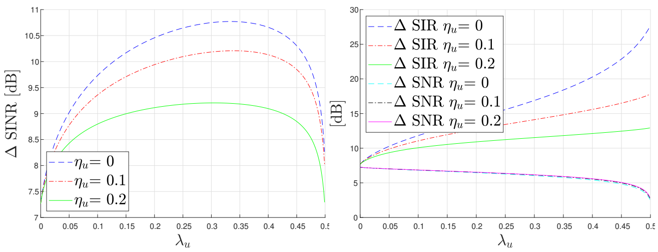
$$\text{SINR}^\text{out} = \frac{\boldsymbol{w}_L^H\boldsymbol{R}_X\boldsymbol{w}_L + \boldsymbol{w}_R^H\boldsymbol{R}_X\boldsymbol{w}_R}{\boldsymbol{w}_L^H\boldsymbol{R}_V\boldsymbol{w}_L + \boldsymbol{w}_R^H\boldsymbol{R}_V\boldsymbol{w}_R}. \tag{20}$$

The binaural input and output SINR are defined as the average of the narrow-band binaural input and, respectively, the output SINR in dB over all frequencies.The SINR improvement is calculated as the difference between the binaural output SINR and binaural input SINR. The binaural SIR and SNR improvements are similarly defined by substituting the overall noise signal with the interference/background noise signal, respectively. In terms of binaural

[1] The derivation can be extended to any number of constrained sources, e.g., for multi-target sources, and constrained multi-interference sources, by extending the constraint sets accordingly.

cues, we define the narrow-band ILD and ITD errors as the absolute values of the difference between the input ILD and ITD and the output ILD and ITD (cf. (7)). The ILD and ITD errors are defined as the narrow-band ILD and ITD errors averaged over all frequencies.

The same interference scaling parameter $\eta_u = \eta_1 = \eta_2$ was used for both interfering sources. The performance measures are examined as functions of the interference weighting parameter $\lambda_u = \lambda_1 = \lambda_2$ with $\lambda_n = 1 - 2\lambda_u$ for various interference scaling parameter values (i.e., η_u equal to [0, 0.1, 0.2]).



(a) Binaural SINR improvement (b) SNR and SIR improvement

Fig. 1: Dependence of binaural SINR improvement (a) and binaural SIR and SNR improvement (b) on interference weighting parameter λ_u.

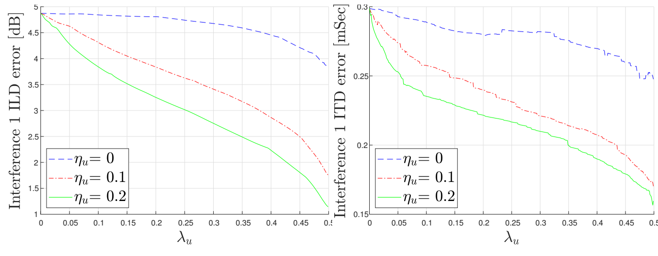
Figure 1 depicts the binaural SINR improvement, the binaural SIR improvement, and the binaural SNR improvement performance measures as functions of the interference weighting parameter λ_u. As the value of the interference weighting parameter λ_u increases, the relative importance of the IR term becomes larger such that, on the one hand, the binaural SNR improvement decreases and, on the other hand, the SIR improvement increases. The SINR improvement peaks at approximately $\lambda_u = 0.35$. This result is expected since the SINR improvement trades off interference reduction and noise reduction. As the interference scaling parameter η_u increases, both the SIR improvement and SINR improvement decreases since the MOP condition can be more easily met if a lower IR is required.

Figure 2 depicts the interference binaural cues as a function of the interference weighting parameter λ_u. As the value of the interference weighting parameter λ_u increases, the interference binaural cue errors for both sources decrease. As the interference scaling parameter η_u increases, the interference binaural cue errors for both sources decrease. It is evident that for η_u equal to zero, the impact of the weighting parameter λ_u on the interference binaural cue errors is negligible.

5. DISCUSSION AND CONCLUSION

In this paper, we propose a unified Pareto optimization framework for binaural distortionless beamformers, in the presence of multi-interference signals in a noisy reverberant environment. To suppress the interfering source while preserving its RTF, several SOP extensions of the BMVDR were proposed in [5, 10, 11]. In the current study, the problem is considered from a different perspective. The approach here is to minimize simultaneously multiple cost functions, i.e., to separately control the interference suppression for each of the interfering sources and the noise reduction.

The proposed framework provides a different perspective for solving the binaural problem. We provide a procedure with two sub-tasks. First, a set of Pareto optimal solutions is provided, rather than an SOP solution, as obtained when one optimizes a single cost function. All solutions in the obtained set are equally optimal. Then,



(a) Interference 1 ILD error (b) Interference 1 ITD error

(c) Interference 2 ILD error (d) Interference 2 ITD error

Fig. 2: Dependence of interference sources' binaural cues on interference weighting parameter λ_u.

in the second sub-task, the most preferred solution out of the optimal Pareto set is selected by setting the controlling parameters in a decision-making procedure. The trade-off parameters can be set in accordance with various considerations, e.g., those that are based on a required interference/noise reduction and those that are based on interference binaural cue preservation. In this study, the interference sources' binaural cue preservation is not an explicit part of the MOP. However, it is impacted by the parameters in the application of the algorithm.

The novel approach can be easily extended by introducing additional cost functions to the respective Pareto MOP (e.g., by adding multi-target and explicit RTF cost functions to the MOP).

6. REFERENCES

[1] J. Wouters, S. Doclo, R. Koning, and T. Francart, "Sound processing for better coding of monaural and binaural cues in auditory prostheses," *Proc. IEEE*, vol. 101, no. 9, pp. 1986–1997, Sept. 2013.

[2] Y. Suzuki, S. Tsukui, F. Asano, and R. Nishimura, "New design method of a binaural microphone array using multiple constraints," *IEICE Tran. on Fundamentals of Elect., Comm. and Comp. Sci.*, vol. 82, no. 4, pp. 588–596, 1999.

[3] T. Lotter and P. Vary, "Dual-channel speech enhancement by superdirective beamforming," *EURASIP J. Appl. Sig. Proc.*, Jan. 2006.

[4] E. Hadad, S. Gannot, and S. Doclo, "Binaural linearly constrained minimum variance beamformer for hearing aid applications," in *Proc. Int. Workshop Acoust. Sig. Enhancement (IWAENC)*, Aachen, Germany, Sept. 2012, pp. 117–120.

[5] E. Hadad, D. Marquardt, S. Doclo, and S. Gannot, "Theoretical analysis of binaural transfer function MVDR beamformers with interference cue preservation constraints," *IEEE/ACM Tran. Audio, Speech, Lang. Proc.*, vol. 23, no. 12, pp. 2449–2464, Dec. 2015.

[6] B. Cornelis, S. Doclo, T. Van dan Bogaert, M. Moonen, and J. Wouters, "Theoretical analysis of binaural multimicrophone noise reduction techniques," *IEEE Tran. Audio, Speech, Lang. Proc.*, vol. 18, no. 2, pp. 342–355, Feb. 2010.

[7] D. Marquardt, E. Hadad, S. Gannot, and S. Doclo, "Theoretical analysis of linearly constrained multi-channel Wiener filtering algorithms for combined noise reduction and binaural cue preservation in binaural hearing aids," *IEEE/ACM Tran. Audio, Speech, Lang. Proc.*, vol. 23, no. 12, pp. 2384–2397, Dec. 2015.

[8] W.C. Liao, Z.Q. Luo, I. Merks, and T. Zhang, "An effective low complexity binaural beamforming algorithm for hearing aids," in *Proc. IEEE Workshop Applicat. Sig. Proc. Audio Acoust. (WASPAA)*, New Paltz, NY, USA, 2015, pp. 1–5.

[9] E. Hadad, D. Marquardt, S. Doclo, and S. Gannot, "Extensions of the binaural MWF with interference reduction preserving the binaural cues of the interfering source," in *Proc. IEEE Int. Conf. Acoust., Speech, Sig. Proc. (ICASSP)*, Shanghai, China, Mar. 2016.

[10] E. Hadad, S. Doclo, and S. Gannot, "The binaural LCMV beamformer and its performance analysis," *IEEE/ACM Tran. Audio, Speech, Lang. Proc.*, vol. 24, no. 3, pp. 543–558, Mar. 2016.

[11] Andreas I Koutrouvelis, Richard Christian Hendriks, Richard Heusdens, and Jesper Jensen, "Relaxed binaural LCMV beamforming," *IEEE/ACM Tran. Audio, Speech, Lang. Proc.*, vol. 25, no. 1, pp. 137–152, 2017.

[12] Hala As'ad, Martin Bouchard, and Homayoun Kamkar-Parsi, "A robust target linearly constrained minimum variance beamformer with spatial cues preservation for binaural hearing aids," *IEEE/ACM Tran. Audio, Speech, Lang. Proc.*, vol. 27, no. 10, pp. 1549–1563, 2019.

[13] 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, 2010.

[14] E. Hadad, S. Doclo, S. Nordholm, and S. Gannot, "A class of pareto optimal binaural beamformers," *IEEE/ACM Tran. Audio, Speech, Lang. Proc.*, 2022.

[15] B. Cornelis, S. Doclo, T. Van dan Bogaert, M. Moonen, and J. Wouters, "Theoretical analysis of binaural multimicrophone noise reduction techniques," *IEEE Tran. Audio, Speech, Lang. Proc.*, vol. 18, no. 2, pp. 342–355, Feb. 2010.

[16] D. Marquardt, E. Hadad, S. Gannot, and S. Doclo, "Optimal binaural LCMV beamformers for combined noise reduction and binaural cue preservation," in *Proc. Int. Workshop Acoust. Sig. Enhancement (IWAENC)*, Antibes-Juan les Pins, France, Sept. 2014.

[17] T. W. Athan and P. Y. Papalambros, "A note on weighted criteria methods for compromise solutions in multi-objective optimization," *Engineering Optimization*, vol. 27, no. 2, pp. 155–176, 1996.

[18] Massimiliano Caramia and Paolo Dell'Olmo, "Multi-objective optimization," *Multi-objective Management in Freight Logistics: Increasing Capacity, Service Level and Safety with Optimization Algorithms*, pp. 11–36, 2008.

[19] T. Marler and S. Arora, "Survey of multi-objective optimization methods for engineering," *Structural and Multidisciplinary Optimization*, vol. 26, no. 6, pp. 369–395, 2004.

[20] KM Miettinen, "Nonlinear multiobjective optimization," 1998.

[21] H. Kayser, S. Ewert, J. Annemüller, T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Database of multichannel In-Ear and Behind-The-Ear Head-Related and Binaural Room Impulse Responses," *Eurasip Journal on Advances in Signal Processing*, vol. 2009, pp. 10 pages, 2009.