Closed-form solution to the multichannel Wiener filter with interaural level difference preservation

Diego Marques Do Carmo, Ricardo Augusto Borsoi, Márcio Holsbach Costa

This article has been accepted for publication in IEEE/ACM Transactions on Audio, Speech and Language Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TASLP.2023.3291530

Closed-Form Solution to the Multichannel Wiener Filter with Interaural Level Difference Preservation

Diego Marques do Carmo, Ricardo Borsoi and Márcio Holsbach Costa

***Abstract*— This paper presents a multichannel Wiener filter (MWF) based noise reduction method with preservation of the interaural level difference (ILD). It minimizes the MWF cost function subject to two constraints for ILD preservation. Under this approach, the weighting coefficient that establishes the trade-off between noise reduction and binaural cue preservation takes a physical interpretation, facilitating its design. The proposed approach results in a convex optimization problem that admits a computationally efficient semi-analytical closed-form solution. Simulation experiments in hearing aid applications were performed considering practical acoustic scenarios. Considering an appropriate set of control parameters, the average performance of the proposed method preserves the ILD of a single interfering source in the same way as the conventional MWF-ILD, keeping the same level of noise reduction as the classic MWF, with approximately the same amount of speech ILD distortion. The proposed method is particularly interesting for implementing real-time noise reduction methods in binaural hearing aids.**

***Index Terms*— Wiener filter, interaural level difference, binaural hearing aids, noise reduction, speech enhancement.**

I. Introduction

Binaural hearing is the ability of the human auditory system to combine and compare information from acoustic signals captured in both ears. It increases speech loudness compared to the case in which only one ear is stimulated [1]. By using signals at both ears, the auditory system creates sound objects facilitating the localization, separation, and identification of sound sources [2].

An example of the binaural hearing advantage occurs when a person listens to the desired speaker while another (undesired speaker) speaks simultaneously. If the spatial positions of both the desired and undesired speakers are close enough, an effect known as "spatial masking" occurs. In such a situation, information from the desired speaker may be missed due to its proximity to the undesired source. Consequently, there may be a decrease in perception and understanding of the desired information content. However, if the undesired speaker moves away, the spatial masking decreases, leading to an increase in intelligibility and a reduction in the necessary cognitive effort. This phenomenon is called *spatial release from masking* [3] [4] [5].

Noise reduction methods are an essential part of modern hearing aids (HA) since they may improve the quality and intelligibility of noisy speech [6] [7]. Strict-sense binaural noise reduction methods combine noise reduction, which aims to recover the desired speech from noisy observations [8] [9], with binaural cue preservation, whose goal is to avoid significant distortions of the spatial information associated with the sound sources [9]. Binaural HA constitutes the most advanced noise reduction device for hearing-impaired people. These gadgets employ a wireless communication link between the left and right ears, by which signals and control parameters are exchanged, thus increasing the diversity of information.

Assuming a sound field generated by a pointwise source, the acoustic power in each ear is generally different due to physical phenomena, such as reverberation, reflection, and diffraction [2]. This power dissimilarity allows the human auditory system to identify different sound objects (e.g., speech) and localize their corresponding physical sources (e.g., the person speaking) [2]. The primary spatial information used by the auditory system to localize sound sources is the interaural time difference (ITD). This binaural information is predominant for frequencies below 1500 Hz [2] and is based on time differences between the signals that reach the ears. Complementary spatial information is provided by the acoustic power difference in both ears, known as the interaural level difference (ILD). It is mathematically defined as the difference (in logarithmic scale) between the left and right acoustic powers from the waves reaching the ears [2]. The proportion between the human head's width and the wavelength of audible sounds makes the ILD the primary binaural cue for frequencies above 1500 Hz [2]. This binaural cue is robust to coherence variations [10] and carries enough information to create complex acoustic scenarios, such as in amplitude stereo panning techniques, even with headphones [11].

Techniques based on the multichannel Wiener filter (MWF)

This work was supported by CNPq under grants 315020/2018-0 and 302492/2021-6. D. Carmo is with the Graduate Program in Electrical Engineering, Federal University of Santa Catarina, Florianópolis-SC, Brazil, R. A. Borsoi is with the Centre de Recherche en Automatique de Nancy (CRAN), Université de Lorraine, CNRS, Vandoeuvre-lès-Nancy, France, and M. H. Costa is with the Department of Electrical and Electronic Engineering, Federal University of Santa Catarina, Florianópolis-SC, Brazil. E-mails: diego.m.carmoh@gmail.com, raborsoi@gmail.com, and costa@eel.ufsc.br.

This article has been accepted for publication in IEEE/ACM Transactions on Audio, Speech and Language Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TASLP.2023.3291530

are among the most studied noise reduction methods for binaural HA. In binaural systems based on the MWF, the binaural cues of the processed speech are unchanged [12]. However, the binaural cues of the filtered noise are distorted, inheriting the same spatial characteristics as the speech [12]. Due to this behavior, HA users may not take advantage of psychoacoustic mechanisms such as the *best ear advantage* [13] or the *spatial release from masking* [3].

Many MWF-based techniques have been proposed to preserve the binaural cues. The MWF with partial noise estimation (MWF-PNE) was proposed in [14]. Its noise reduction filters have a simple closed form, being defined as the convex combination of the MWF solution and an all-pass filter. This solution may preserve the ILD and IPD of the undesired additive component, by adjusting the amount of unprocessed noise at the output of the gadget, but results in significant loss in noise reduction performance.

An important class of MWF-based techniques achieves (strict-sense) binaural cue preservation of the interference signal through designing and optimizing cost functions incorporating regularization terms that directly penalize distortions in the ILD and IPD. This is the case of the MWF with interaural transfer function preservation (MWF-ITF) [15]. This technique preserves the interfering source's ILD and IPD at the cost of losing noise reduction performance and/or distorting the binaural cues of the speech source. Despite their advantages, these techniques do not have closed-form solutions, which may prevent their real-time implementation.

MWF-based methods with ILD preservation (MWF-ILD) generally employ cost functions comprised of two terms: i) the conventional MWF cost function, which aims to minimize the power of the overall noise (restraining the speech distortion); and ii) an ILD penalty term, which penalizes solutions that deviate from the original ILD of the input noise [16] [17] [18]. A weighting parameter establishes the trade-off between the optimization effort for each term in the cost function.

The first proposed MWF-ILD technique defined the ILD penalty term as the mean squared difference between the input and output ILDs [16]. The work in [17] proposed a variation of the ILD penalty term based on an approximation of the logarithm function, resulting in a cost function that equally penalizes positive and negative ILD errors. In both works, the design of the binaural noise reduction filters was based on an unconstrained minimization of the MWF cost function plus the weighted ILD penalty term. However, despite their notable performance in offline experiments, these techniques are inappropriate for online implementation in embedded systems with severe computational limitations, such as in HA applications. Lately, an adaptive filter implementation of the method presented in [17] was proposed in [18] to deal with this problem. This algorithm allows for the practical implementation of the MWF-ILD. However, the slow convergence rate of the adaptive algorithm may lead to suboptimal solutions, compromising the maximum attainable noise reduction and spatial preservation. As a result, previous MWF-ILD methods have some considerable drawbacks, which may be summarized as: i) the design of the binaural filters is based on the unconstrained minimization of a nonconvex and highly nonlinear cost function, which may result in long optimization times unsuitable for HA applications; ii) there is no guarantee of global optimality and algorithm convergence; iii) the weighting parameter employed for setting the trade-off between noise reduction and ILD preservation does not have a direct relationship with physical performance measures, making its design difficult.

Considering the presented facts, this work proposes a new MWF-ILD-based noise reduction method surpassing the abovementioned limitations. The contributions of this paper are the following: Firstly, the original (unconstrained) MWF-ILD optimization problem is changed to a constrained (CO) form (CO-MWF-ILD) and then reformulated to an equivalent nonconvex quadratically constrained quadratic program (QCQP) with two quadratic constraints (QC-MWF-ILD). This makes the parameter design intuitive and directly linked to physical measures. Secondly, a convex semidefinite program (SDP) relaxation of the QC-MWF-ILD is derived (SD-MWF-ILD), which is guaranteed to have a globally optimal solution. Using recent results from nonconvex optimization literature, we show that the SD-MWF-ILD achieves the same solution as the QC-MWF-ILD, which provides a means of computing the globally optimal solution to the CO-MWF-ILD problem. Thirdly, we derive a reformulation of the SD-MWF-ILD, called constrained binaural MWF-ILD (CB-MWF-ILD), as a nonlinear optimization problem with linear constraints by considering the case of single pointwise speech and interfering sources. A semi-analytical closed-form solution requiring only simple algebraic operations is then derived. Thus, we obtain the optimal solution to the original non-convex CO-MWF-ILD cost function at very low computational complexity and without requiring any iterative optimization procedure. This achievement may contribute to the availability of low-cost commercial binaural hearing aid gadgets. Fourthly, the relation between the proposed solution with other techniques described in the literature is presented. Fifthly, computer simulations are provided for an acoustic scenario comprised of one pointwise speech source and one pointwise interfering source, corroborating the effectiveness of the proposed method. Results indicate that the proposed method leads to a computational complexity similar to the MWF algorithm, reducing the computational time to estimate the noise reduction filters and presenting noise reduction and spatial preservation performance equivalent to the conventional (unconstrained) MWF-ILD implementation.

The remainder of this paper is structured as follows. Section II defines the employed nomenclature. In Section III, the binaural HA application is introduced. Section IV describes the conventional MWF and the conventional MWF-ILD methods. Section V presents the CO-MWF-ILD and the QC-MWF-ILD. The CO-MWF-ILD equivalent dual problem and its semidefinite relaxation form (SD-MWF-ILD) are also shown. In Section VI, a semi-analytical closed-form solution (CB-MWF-ILD) is derived. Section VII describes the computational simulation setup, while in Section VIII, results are presented and discussed. Finally, Section 0 presents the conclusions of this work.

II. Definitions and Notation

Throughout this text, lowercase italic symbols represent scalars, while lowercase and uppercase bold symbols denote vectors and matrices, respectively. Subscripts in lowercase italic letters $\{\cdot\}_l$ denote indexes, while uppercase letters $\{\cdot\}_{\mathrm{L}}$ denote literals. Terms 'minimum', 'argument that minimizes,' 'argument that maximizes,' and 'subject to' are abbreviated to 'min.,' 'arg.min.,' 'arg.max.,' and 's.t.,' respectively. The matrix inequality representation $\mathbf{A} \succ 0$ ($\mathbf{A} \succeq 0$) means $\mathbf{A}$ is positive (semi-) definite. The vector inequality representation $\mathbf{a} > 0$ ($\mathbf{a} \geq 0$) means all elements of $\mathbf{a}$ are greater (or equal) than 0. Equivalent meaning is considered for $\mathbf{a} < 0$ ($\mathbf{a} \leq 0$). Zero matrices of dimension $m \times n$ are represented as $\mathbf{0}_{m\times n}$. The identity matrix is represented as $\mathbf{I}$, and its order is defined according to the context. The sets of real, positive real, and complex numbers are represented as $\mathbb{R}$, $\mathbb{R}_{++}$, and $\mathbb{C}$, respectively. The operator diag(·) creates a diagonal matrix, handling different kinds of input arguments, e.g., for the scalar a and square matrix $\mathbf{A}$ of order m, diag(a,$\mathbf{A}$) leads to:

$$\mathrm{diag}(a, \mathbf{A}) = \begin{bmatrix} a & \mathbf{0}_{1\times m} \\ \mathbf{0}_{m\times 1} & \mathbf{A} \end{bmatrix}. \tag{1}$$

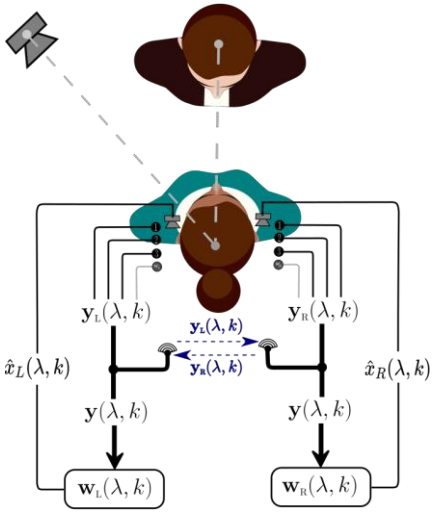

Fig. 1. Binaural hearing aid system.

III. Signal and Systems

Fig. 1 depicts the investigated acoustic scenario constituted by a HA user, a desired pointwise source (speaker), an interfering pointwise source (loudspeaker), and background noise. The binaural system comprises two HA: one on the left (L) side with M_{L} microphones and another on the right (R) side with M_{R} microphones. The total number of microphones is $M = M_{\mathrm{L}}+M_{\mathrm{R}}$.

The time-frequency representation of the noisy input signal, at the frequency bin index k, time frame index λ, microphone $m \in \{1, 2, \ldots, M_\ell\}$, and side $\ell \in \{\mathrm{L}, \mathrm{R}\}$ is

$$y_{m,\ell}(\lambda,k) = x_{m,\ell}(\lambda,k) + v_{m,\ell}(\lambda,k), \tag{2}$$

in which $x_{m,\ell}(\lambda,k)$ is the desired component (speech); and $v_{m,\ell}(\lambda,k) = u_{m,\ell}(\lambda,k)+n_{m,\ell}(\lambda,k)$ is the overall noise component, which is the sum of the interfering component $u_{m,\ell}(\lambda,k)$ and background noise $n_{m,\ell}(\lambda,k)$.

The noisy-speech input vector $\mathbf{y}_\ell \in \mathbb{C}^{M_\ell}$ at side ℓ is defined as

$$\mathbf{y}_\ell(\lambda,k) = [\, y_{\ell,1}(\lambda,k)\ y_{\ell,2}(\lambda,k) \ldots y_{\ell,M_\ell}(\lambda,k)\,]^{\mathrm{T}}, \tag{3}$$

in which $\{\cdot\}^{\mathrm{T}}$ is the transpose operation. The binaural noisy vector, $\mathbf{y}(\lambda,k) \in \mathbb{C}^M$, is defined as

$$\mathbf{y}(\lambda,k) = [\, \mathbf{y}_{\mathrm{L}}^{\mathrm{T}}(\lambda,k) \quad \mathbf{y}_{\mathrm{R}}^{\mathrm{T}}(\lambda,k)\,]^{\mathrm{T}}. \tag{4}$$

This vector is available on both HAs due to a full-duplex communication link and can be decomposed as

$$\mathbf{y}(\lambda,k) = \mathbf{x}(\lambda,k) + \mathbf{v}(\lambda,k) = \mathbf{x}(\lambda,k) + \mathbf{u}(\lambda,k) + \mathbf{n}(\lambda,k), \tag{5}$$

in which $\mathbf{x}(\lambda,k)$, $\mathbf{u}(\lambda,k)$, $\mathbf{n}(\lambda,k)$, and $\mathbf{v}(\lambda,k)$ are, respectively, the speech, interference, background noise, and overall noise vectors, all similarly defined as $\mathbf{y}(\lambda,k)$ in (3) and (4).

The pointwise speech $\mathbf{x}(\lambda,k)$ and interfering $\mathbf{u}(\lambda,k)$ vectors are modeled as

$$\mathbf{x}(\lambda,k) = s_{\mathrm{x}}(\lambda,k)\mathbf{a}(\lambda,k), \tag{6}$$

$$\mathbf{u}(\lambda,k) = s_{\mathrm{u}}(\lambda,k)\mathbf{b}(\lambda,k), \tag{7}$$

in which $s_{\mathrm{x}}(\lambda,k)$ and $s_{\mathrm{u}}(\lambda,k)$ are the clean speech and interfering signals; and vectors $\mathbf{a}(\lambda,k) = [\, a_{\mathrm{L},1}(\lambda,k)\ a_{\mathrm{L},2}(\lambda,k)\ \ldots\ a_{\mathrm{L},M_{\mathrm{L}}}(\lambda,k)\ a_{\mathrm{R},1}(\lambda,k)\ \ldots\ a_{\mathrm{R},M_{\mathrm{R}}}(\lambda,k)\,]^{\mathrm{T}}$ and $\mathbf{b}(\lambda,k) = [\, b_{\mathrm{L},1}(\lambda,k)\ b_{\mathrm{L},2}(\lambda,k)\ \ldots\ b_{\mathrm{L},M_{\mathrm{L}}}(\lambda,k)\ b_{\mathrm{R},1}(\lambda,k)\ \ldots\ b_{\mathrm{R},M_{\mathrm{R}}}(\lambda,k)\,]^{\mathrm{T}}$ are the complex M-dimensional acoustic transfer function (ATF) vectors related to the speech and interfering sources, which carry information about the environment and the head and torso of the HA user.

In general, a microphone at each HA is defined as the reference; whose associated signal is defined as:

$$\mathbf{q}_\ell^{\mathrm{T}}\mathbf{y}(\lambda,k) = x_\ell(\lambda,k) + u_\ell(\lambda,k) + n_\ell(\lambda,k), \tag{8}$$

in which $\mathbf{q}_\ell$ is a microphone selection vector with entries equal to 1 in the position of the reference microphone at the side ℓ and zero elsewhere; $x_\ell(\lambda,k)$, $u_\ell(\lambda,k)$, and $n_\ell(\lambda,k)$ are the speech, interference, and background noise components at the reference microphones.

The output signals on the left and right HAs are defined as

$$z_\ell(\lambda,k) = \mathbf{w}_\ell^{\mathrm{H}}(\lambda,k)\mathbf{y}(\lambda,k), \tag{9}$$

in which $\mathbf{w}_\ell \in \mathbb{C}^M$ are the binaural noise reduction filters on the HA at side ℓ.

IV. The MWF-ILD Noise Reduction Method

The conventional binaural MWF noise reduction method is based on the mean squared error (MSE) criterion. It defines the best linear estimators for the speech in the reference microphones, i.e., $\hat{x}_{\mathrm{L}}(\lambda,k)$ and $\hat{x}_{\mathrm{R}}(\lambda,k)$, through the minimization of the following cost function [12]:

$$J_{\mathrm{MWF}}(\lambda,k) = \mathbb{E}\left\{ \left\| \begin{bmatrix} x_{\mathrm{L}}(\lambda,k) - \mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\mathbf{y}(\lambda,k) \\ x_{\mathrm{R}}(\lambda,k) - \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\mathbf{y}(\lambda,k) \end{bmatrix} \right\|^2 \right\}, \tag{10}$$

in which $\|\cdot\|$ is the Euclidean norm, and $\mathbb{E}\{\cdot\}$ is the expected value operator. Considering that speech, interference and noise are uncorrelated ($\mathbb{E}\{\mathbf{x}(\lambda,k)\mathbf{v}^{\mathrm{H}}(\lambda,k)\} = \mathbb{E}\{\mathbf{v}(\lambda,k)\mathbf{x}^{\mathrm{H}}(\lambda,k)\} = \mathbf{0}_{M\times M}$) (10) turns to [12]

$$J_{\mathrm{MWF}}(\lambda,k) = \mathbf{w}^{\mathrm{H}}(\lambda,k)\boldsymbol{\Phi}_{\mathrm{yy}}(\lambda,k)\mathbf{w}(\lambda,k) - \mathbf{w}^{\mathrm{H}}(\lambda,k)\mathbf{p}_{\mathrm{xx}}(\lambda,k) - \mathbf{p}_{\mathrm{xx}}^{\mathrm{H}}(\lambda,k)\mathbf{w}(\lambda,k) + p_{\mathrm{xx}}(\lambda,k), \quad (11)$$

in which

$$\mathbf{w}(\lambda,k) = [\mathbf{w}_{\mathrm{L}}^{\mathrm{T}}(\lambda,k) \quad \mathbf{w}_{\mathrm{R}}^{\mathrm{T}}(\lambda,k)]^{\mathrm{T}}, \quad (12)$$

$$\boldsymbol{\Phi}_{\mathrm{yy}}(\lambda,k) = \mathrm{diag}(\boldsymbol{\Phi}_{\mathrm{y}}(\lambda,k), \boldsymbol{\Phi}_{\mathrm{y}}(\lambda,k)), \quad (13)$$

$$\mathbf{p}_{\mathrm{xx}}(\lambda,k) = \left[\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathrm{x}}^{\mathrm{T}}(\lambda,k) \quad \mathbf{q}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathrm{x}}^{\mathrm{T}}(\lambda,k)\right]^{\mathrm{T}}, \quad (14)$$

$$p_{\mathrm{xx}}(\lambda,k) = \mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathrm{x}}(\lambda,k)\mathbf{q}_{\mathrm{L}} + \mathbf{q}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathrm{x}}(\lambda,k)\mathbf{q}_{\mathrm{R}}. \quad (15)$$

The coherence matrices are defined as

$$\boldsymbol{\Phi}_{\mathrm{y}}(\lambda,k) = \mathbb{E}\{\mathbf{y}(\lambda,k)\mathbf{y}^{\mathrm{H}}(\lambda,k)\} = \boldsymbol{\Phi}_{\mathrm{x}}(\lambda,k) + \boldsymbol{\Phi}_{\mathrm{v}}(\lambda,k), \quad (16)$$

$$\boldsymbol{\Phi}_{\mathrm{v}}(\lambda,k) = \mathbb{E}\{\mathbf{v}(\lambda,k)\mathbf{v}^{\mathrm{H}}(\lambda,k)\} = \boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k) + \boldsymbol{\Phi}_{\mathrm{n}}(\lambda,k), \quad (17)$$

$$\boldsymbol{\Phi}_{\mathrm{x}}(\lambda,k) = \mathbb{E}\{\mathbf{x}(\lambda,k)\mathbf{x}^{\mathrm{H}}(\lambda,k)\} = p_{\mathrm{sx}}(\lambda,k)\mathbf{a}(\lambda,k)\mathbf{a}^{\mathrm{H}}(\lambda,k), \quad (18)$$

$$\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k) = \mathbb{E}\{\mathbf{u}(\lambda,k)\mathbf{u}^{\mathrm{H}}(\lambda,k)\} = p_{\mathrm{su}}(\lambda,k)\mathbf{b}(\lambda,k)\mathbf{b}^{\mathrm{H}}(\lambda,k), \quad (19)$$

in which $\boldsymbol{\Phi}_{\mathrm{n}}(\lambda,k) = \mathbb{E}\{\mathbf{n}(\lambda,k)\mathbf{n}^{\mathrm{H}}(\lambda,k)\}$; and $p_{\mathrm{sx}}(\lambda,k) = \mathbb{E}\{|s_{\mathrm{x}}(\lambda,k)|^2\}$ and $p_{\mathrm{su}}(\lambda,k) = \mathbb{E}\{|s_{\mathrm{u}}(\lambda,k)|^2\}$ are the power spectrum densities (PSD) of the speech and interference, respectively.

A closed-form solution to the minimum point of the cost function defined in (11) can be obtained by taking the gradient of $J_{\mathrm{MWF}}(\lambda,k)$ with relation to $\mathbf{w}(\lambda,k)$ and equating it to zero. Assuming that $\boldsymbol{\Phi}_{\mathrm{yy}}(\lambda,k) \succeq 0$, it results in the generalized MWF (G-MWF) filter, given by [19]

$$\mathbf{w}_{\text{G-MWF}}(\lambda,k) = \boldsymbol{\Phi}_{\mathrm{yy}}^{\dagger}(\lambda,k)\boldsymbol{\Phi}_{\mathrm{xx}}(\lambda,k)\mathbf{q}, \quad (20)$$

in which $\{\cdot\}^{\dagger}$ is the Moore-Penrose inverse (pseudo-inverse); $\boldsymbol{\Phi}_{\mathrm{xx}}(\lambda,k) = \mathrm{diag}(\boldsymbol{\Phi}_{\mathrm{x}}(\lambda,k), \boldsymbol{\Phi}_{\mathrm{x}}(\lambda,k))$; and $\mathbf{q} = [\mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \; \mathbf{q}_{\mathrm{R}}^{\mathrm{T}}]^{\mathrm{T}}$. In the particular case in which $\boldsymbol{\Phi}_{\mathrm{yy}}(\lambda,k) \succ 0$, then $\boldsymbol{\Phi}^{\dagger}{}_{\mathrm{yy}}(\lambda,k) = \boldsymbol{\Phi}^{-1}{}_{\mathrm{yy}}(\lambda,k)$ and (20) becomes the conventional binaural MWF [12]. In such a situation, considering the pointwise speech source case, it is theoretically proven that the binaural cues of both processed speech and residual noise at the output are equal to the binaural cues of the input speech [12]. As a result, both signals are psychoacoustically perceived as arriving from the speech direction [18].

A. MWF-ILD

To achieve the correct spatial perception of the interfering source at the HA output, some works extend the MWF cost function with additional terms for penalizing solutions $\mathbf{w}(\lambda,k)$ that distort the original binaural cues of the interference. The ILD has been an interesting alternative for achieving this goal. The MWF-ILD was originally defined as [16] [17]

$$J_{\text{MWF-ILD}}(\lambda,k) = J_{\mathrm{MWF}}(\lambda,k) + \alpha(\lambda,k)J_{\mathrm{ILD}}^{\mathrm{u}}(\lambda,k), \quad (21)$$

in which $\alpha(\lambda,k) \in \mathbb{R}_{++}$ is a frequency-dependent weighting parameter, which impacts the tradeoff between noise reduction and ILD preservation, and $J^{\mathrm{u}}{}_{\mathrm{ILD}}$ is the interference ILD penalty term, generically defined as [16] [17] [20]

$$J_{\mathrm{ILD}}^{\mathrm{u}}(\lambda,k) = \left[g(\mathrm{ILD}_{\mathrm{ou}}^{\mathrm{u}}(\lambda,k)) - g(\mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}}(\lambda,k))\right]^2, \quad (22)$$

in which $g(\cdot) = 10\cdot\log_{10}(\cdot)$, and $\mathrm{ILD}^{\mathrm{u}}{}_{\mathrm{in}}(\lambda,k)$ and $\mathrm{ILD}^{\mathrm{u}}{}_{\mathrm{ou}}(\lambda,k)$ are the input and output ILDs of the interfering source, respectively defined as [16][17]

$$\mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}}(\lambda,k) = \frac{\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k)\mathbf{q}_{\mathrm{L}}}{\mathbf{q}_{\mathrm{R}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k)\mathbf{q}_{\mathrm{R}}}, \quad (23)$$

and

$$\mathrm{ILD}_{\mathrm{ou}}^{\mathrm{u}}(\lambda,k) = \frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k)\mathbf{w}_{\mathrm{L}}(\lambda,k)}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k)\mathbf{w}_{\mathrm{R}}(\lambda,k)}, \quad (24)$$

in which $\mathbf{q}_{\mathrm{R}}{}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k)\mathbf{q}_{\mathrm{R}} > 0$ and $\mathbf{w}_{\mathrm{R}}{}^{\mathrm{H}}(\lambda,k)\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k)\mathbf{w}_{\mathrm{R}}(\lambda,k) > 0$.

The first ILD penalty term was proposed in [16], corresponding to a first-order Taylor series approximation for $g(\cdot)$, i.e., $g(x) \cong x-1$. A more accurate approximation for (22) was proposed in [17] based on the inverse hyperbolic tangent function approximation of the logarithm, i.e., $g(x) \cong (x-1)/(x+1)$.

The optimum noise reduction filter that minimizes (21) can be obtained by solving the conventional (unconstrained) MWF-ILD problem:

$$\mathbf{w}_{\text{MWF-ILD}}(\lambda,k) = \underset{\mathbf{w}(\lambda,k)}{\arg.\min.} \; J_{\text{MWF-ILD}}(\lambda,k). \quad (25)$$

It has been shown that the optimal filter obtained from (25) provides adequate noise reduction as well as psychoacoustic spatial preservation for $g(x) \cong x-1$ and $g(x) \cong (x-1)/(x+1)$ approximations [16] [17] [18]. However, compared to the closed-form solution of the conventional MWF in (20) the formulated minimization problem in (25) has some inconveniences. Firstly, there is no direct relation between the parameter $\alpha(\lambda,k)$ and the maximum tolerated input-output ILD error, which imposes difficulties in its design. Secondly, the cost function $J^{\mathrm{u}}{}_{\mathrm{ILD}}(\lambda,k)$ is nonconvex and highly nonlinear. Consequently, usual approaches to solve (25) may result in considerable additional computational burden compared to (20), as general-purpose optimization solvers have to be applied. Moreover, these approaches are not guaranteed to find the optimal global solution and may get trapped at local minima. These disadvantages motivate a reformulation of the conventional MWF-ILD method.

The following sections omit the time-frame λ and the frequency index k for space-saving and clarity.

V. Convex Formulation of the MWF-ILD

In this section, an equivalent convex formulation for the conventional (unconstrained) MWF-ILD problem presented in (25) is proposed. It aims to provide tractable means for computing the globally optimal solution to the original problem.

A. Constrained MWF-ILD

An alternative ILD penalty term can be defined as:

$$J_{\text{ILD-REL}}^{\mathrm{u}} = J_{\mathrm{ILD}}^{\mathrm{u}} \cdot g(\mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}})^{-2}, \quad (26)$$

for $g(\mathrm{ILD}^{\mathrm{u}}{}_{\mathrm{in}}) \neq 0$, i.e., the interference source is not positioned directly in front of the HA user. The $\mathrm{ILD}^{\mathrm{u}}{}_{\mathrm{in}} = 1$ scenario is not considered in this work because, in this case, the classic MWF will impose no binaural cue distortions.

Unlike the $J^{\mathrm{u}}{}_{\mathrm{ILD}}$ penalty term presented in (22), $J^{\mathrm{u}}{}_{\text{ILD-REL}}$ incorporates a scale factor that weights the input-output ILD error according to the input ILD. Thus, it measures relative ILD deviations.

Constraining (26) in a given range, limited by $\delta \in \mathbb{R}_{++}$ (the ILD relative error tolerance), the constrained MWF-ILD (CO-MWF-ILD) noise reduction method can be defined as [21]:

$$\mathbf{w}_{\text{CO-MWF-ILD}} = \underset{\mathbf{w}}{\arg.\min.}\ J_{\text{MWF}} \quad \text{s.t. } J^{\text{u}}_{\text{ILD-REL}} \le \delta^2 , \tag{27}$$

in which $\mathbf{w}_{\text{CO-MWF-ILD}}$ is the optimal solution. By using this approach, we change the original problem, which consists of determining the value of α in (21) (which is hard to design because it does not have a physical interpretation) by defining δ (which is an easier task because it is directly related to the maximum tolerable amount of ILD relative error).

To write the constraint in (27) in a mathematically tractable form, we apply the antilogarithm and rearrange the ratio, which results in two cases with two quadratic inequalities:

i) If $\text{ILD}^{\text{u}}_{\text{in}} > 1$ then

$$\mathbf{w}_{\text{L}}^{\text{H}}\mathbf{\Phi}_{\text{u}}\mathbf{w}_{\text{L}} - \delta_1\mathbf{w}_{\text{R}}^{\text{H}}\mathbf{\Phi}_{\text{u}}\mathbf{w}_{\text{R}} \ge 0 , \tag{28}$$

$$\mathbf{w}_{\text{L}}^{\text{H}}\mathbf{\Phi}_{\text{u}}\mathbf{w}_{\text{L}} - \delta_2\mathbf{w}_{\text{R}}^{\text{H}}\mathbf{\Phi}_{\text{u}}\mathbf{w}_{\text{R}} \le 0 , \tag{29}$$

ii) If $0 < \text{ILD}^{\text{u}}_{\text{in}} < 1$ then

$$\mathbf{w}_{\text{L}}^{\text{H}}\mathbf{\Phi}_{\text{u}}\mathbf{w}_{\text{L}} - \delta_1\mathbf{w}_{\text{R}}^{\text{H}}\mathbf{\Phi}_{\text{u}}\mathbf{w}_{\text{R}} \le 0 , \tag{30}$$

$$\mathbf{w}_{\text{L}}^{\text{H}}\mathbf{\Phi}_{\text{u}}\mathbf{w}_{\text{L}} - \delta_2\mathbf{w}_{\text{R}}^{\text{H}}\mathbf{\Phi}_{\text{u}}\mathbf{w}_{\text{R}} \ge 0 , \tag{31}$$

in which

$$\delta_1 = (\text{ILD}^{\text{u}}_{\text{in}})^{1-\delta} \ \text{ and } \ \delta_2 = (\text{ILD}^{\text{u}}_{\text{in}})^{1+\delta} . \tag{32}$$

We can generically represent (28)-(31) using (12), leading to:

$$(-1)^{t-1}\left[-\mathbf{w}^{\text{H}}\mathbf{\Phi}_{\text{cc}}(\delta_1)\mathbf{w} \quad \mathbf{w}^{\text{H}}\mathbf{\Phi}_{\text{cc}}(\delta_2)\mathbf{w}\right]^{\text{T}} \le \mathbf{0}_{2\times 1} , \tag{33}$$

in which

$$\mathbf{\Phi}_{\text{cc}}(\kappa) = \text{diag}(\mathbf{\Phi}_{\text{u}}, -\kappa\mathbf{\Phi}_{\text{u}}) , \tag{34}$$

for $\kappa \in \{\delta_1, \delta_2\}$ and t defined as

$$t = \begin{cases} 1 , & \text{ILD}^{\text{u}}_{\text{in}} > 1 \\ 2 , & 0 < \text{ILD}^{\text{u}}_{\text{in}} < 1 \end{cases} . \tag{35}$$

Using (33), the constrained problem in (27) is equivalently represented as

$$\mathbf{w}_{\text{QC-MWF-ILD}} = \underset{\mathbf{w}}{\arg.\min.}\ J_{\text{MWF}} \quad \text{s.t. } \mathbf{c}(\mathbf{w}) \le \mathbf{0}_{2\times 1} . \tag{36}$$

where $\mathbf{c}(\mathbf{w})$ is defined as the left-hand side of (33).

Both optimization problems in (27) and (36) are equivalent, resulting in $\mathbf{w}_{\text{CO-MWF-ILD}} = \mathbf{w}_{\text{QC-MWF-ILD}}$.

The reformulation of the unconstrained (conventional) MWF-ILD problem (defined in (25)) into the MWF-ILD with two quadratic constraints (QC-MWF-ILD) in (36) yields a straightforward form to control the ILD distortion. The problem defined in (36) is known as a quadratically constrained quadratic problem (QCQP), and is widely studied in the optimization literature [22] [23] [24].

B. Convex Reformulation of the QC-MWF-ILD

The QC-MWF-ILD problem in (36) is hard to solve because it is non-convex. This section defines a convex reformulation of the QC-MWF-ILD with theoretical guarantees to achieve the globally optimal solution.

Computing the Lagrangian of (36) results in

$$J_{\mathcal{L}}(\mathbf{w}, \boldsymbol{\tau}) = J_{\text{MWF}}(\mathbf{w}) + \boldsymbol{\tau}^{\text{T}}\mathbf{c}(\mathbf{w}) , \tag{37}$$

in which $\boldsymbol{\tau} = [\,\tau_1 \ \tau_2\,]^{\text{T}}$ are the Lagrange multipliers. Substituting (11) and (33) in (37) leads to:

$$J_{\mathcal{L}}(\mathbf{w}, \boldsymbol{\tau}) = \mathbf{w}^{\text{H}}\mathbf{\Phi}_{\text{yc}}(\boldsymbol{\tau})\mathbf{w} - \mathbf{w}^{\text{H}}\mathbf{p}_{\text{xx}} - \mathbf{p}_{\text{xx}}^{\text{H}}\mathbf{w} + p_{\text{xx}} , \tag{38}$$

in which

$$\mathbf{\Phi}_{\text{yc}}(\boldsymbol{\tau}) = \mathbf{\Phi}_{\text{yy}} + (-1)^{t-1}\sum_{j=1}^{2}(-1)^j\tau_j\mathbf{\Phi}_{\text{cc}}(\delta_j) . \tag{39}$$

Using (13), (16), and (34) in (39) results in

$$\mathbf{\Phi}_{\text{yc}}(\boldsymbol{\tau}) = \text{diag}(\mathbf{\Phi}_{\text{yu}}(\beta_1(\boldsymbol{\tau})), \mathbf{\Phi}_{\text{yu}}(\beta_2(\boldsymbol{\tau}))) , \tag{40}$$

in which:

$$\mathbf{\Phi}_{\text{yu}}(\beta_r(\boldsymbol{\tau})) = \mathbf{\Phi}_{\text{y}} + \beta_r(\boldsymbol{\tau})\mathbf{\Phi}_{\text{u}} , \tag{41}$$

$$\beta_r(\boldsymbol{\tau}) = (-1)^{t-1}\boldsymbol{\delta}_r^{\text{T}}\boldsymbol{\tau} , \tag{42}$$

for $r \in \{1, 2\}$, $\boldsymbol{\delta}_1 = [-1\ 1]^{\text{T}}$, and $\boldsymbol{\delta}_2 = [\delta_1 \ -\delta_2]^{\text{T}}$.

From (38), the Lagrangian dual problem of the QC-MWF-ILD is defined as [22]:

$$J_{\mathcal{L}}(\mathbf{w}^{\text{opt}}, \boldsymbol{\tau}^{\text{opt}}) = \max_{\boldsymbol{\tau}\ge 0}.\ \min_{\mathbf{w}}.\ J_{\mathcal{L}}(\mathbf{w}, \boldsymbol{\tau}) . \tag{43}$$

in which $\mathbf{w}^{\text{opt}}$ is the optimal solution, and $\boldsymbol{\tau}^{\text{opt}}$ denotes the optimal Lagrange multipliers. The dual problem, defined in (43), can be formulated for any constrained optimization problem and has two essential properties [22]: i) it is always a concave problem, whatever the form of the primal problem; and ii) for any $\boldsymbol{\tau} \ge 0$ it defines a lower bound for the QC-MWF-ILD objective function, i.e.,

$$J_{\text{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}}) \ge J_{\mathcal{L}}(\mathbf{w}^{\text{opt}}, \boldsymbol{\tau}) . \tag{44}$$

The inequality in (44) is not necessarily tight. However, considering that the QC-MWF-ILD in (36) is a QCQP (in complex variables) with two quadratic constraints, and assuming that it is strictly feasible[1], then the optimal solution of both the primal problem $J_{\text{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}})$ and the dual problem $J_{\mathcal{L}}(\mathbf{w}^{\text{opt}}, \boldsymbol{\tau}^{\text{opt}})$ are the same [23]. Thus, the inequality in (44) holds with equality, and $\mathbf{w}^{\text{opt}} = \mathbf{w}_{\text{QC-MWF-ILD}}$ [22] [23], i.e.,

$$J_{\text{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}}) = J_{\mathcal{L}}(\mathbf{w}_{\text{QC-MWF-ILD}}, \boldsymbol{\tau}^{\text{opt}}) . \tag{45}$$

The result in (45) is known as the *strong duality property* [22]. As a consequence, the solution to the QC-MWF-ILD problem in (36) can be computed precisely by (43), i.e.,

$$J_{\text{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}}) = \max_{\boldsymbol{\tau}\ge 0}.\ \min_{\mathbf{w}}.\ J_{\mathcal{L}}(\mathbf{w}, \boldsymbol{\tau}) . \tag{46}$$

Since the Lagrangian is a quadratic function over $\mathbf{w}$, (46) can be reduced to a single maximization problem. Firstly, the inner minimization problem must be solved, which can be done by equating the gradient of (38) with respect to $\mathbf{w}$ to zero, which leads to the following linear system:

$$\mathbf{\Phi}_{\text{yc}}(\boldsymbol{\tau})\mathbf{w} = \mathbf{p}_{\text{xx}} . \tag{47}$$

Assuming $\mathbf{\Phi}_{\text{yc}}(\boldsymbol{\tau}) \succeq 0$, without loss of generality, an analytical solution to (47) is given by

$$\mathbf{w}^{\text{opt}}(\boldsymbol{\tau}) = \mathbf{\Phi}_{\text{yc}}^{\dagger}(\boldsymbol{\tau})\mathbf{p}_{\text{xx}} . \tag{48}$$

Substituting (48) in (38) leads to

$$J_{\mathcal{L}}(\mathbf{w}^{\text{opt}}(\boldsymbol{\tau}), \boldsymbol{\tau}) = \min_{\mathbf{w}}.\ J_{\mathcal{L}}(\mathbf{w}, \boldsymbol{\tau}) = -\mathbf{p}_{\text{xx}}^{\text{H}}\mathbf{\Phi}_{\text{yc}}^{\dagger}(\boldsymbol{\tau})\mathbf{p}_{\text{xx}} + p_{\text{xx}} . \tag{49}$$

[1] The QC-MWF-ILD is strictly feasible if there exists $\mathbf{w}$ that satisfies the constraints $\mathbf{c}(\mathbf{w})$, such that $\mathbf{c}(\mathbf{w}) < 0$.

Without the optimal $\boldsymbol{\tau}$, the solution $\mathbf{w}^{\text{opt}}(\boldsymbol{\tau})$ in (48) is not necessarily equal to $\mathbf{w}_{\text{QC-MWF-ILD}}$, it only determines a lower bound solution. To find $\mathbf{w}_{\text{QC-MWF-ILD}}$, it is necessary to find $\boldsymbol{\tau}$ that maximizes (49). This is obtained by introducing constraints $\boldsymbol{\tau} \geq 0$ and $\boldsymbol{\Phi}_{\text{yc}}(\boldsymbol{\tau}) \succeq 0$ (which are required for the existence of a finite solution to the inner optimization problem in (46)) explicitly in the maximization problem, i.e.,

$$\begin{aligned} \boldsymbol{\tau}^{\text{opt}} = \ & \underset{\boldsymbol{\tau}}{\text{arg.max.}} \ -\mathbf{p}_{\text{xx}}^{\text{H}} \boldsymbol{\Phi}_{\text{yc}}^{\dagger}(\boldsymbol{\tau}) \mathbf{p}_{\text{xx}} + p_{\text{xx}} \\ & \text{s.t.} \quad \boldsymbol{\tau} \geq 0, \quad \boldsymbol{\Phi}_{\text{yc}}(\boldsymbol{\tau}) \succeq 0 \end{aligned} . \tag{50}$$

This problem can be written as a convex semidefinite (SD) programming (SDP) of the form (see [22] and [23] for details)

$$\phi^{\text{opt}}, \boldsymbol{\tau}^{\text{opt}} = \arg \min_{\phi, \boldsymbol{\tau}} \{ -\phi \ \text{ s.t. } \ \boldsymbol{\tau} \succeq 0; \boldsymbol{\Phi}_{\text{g}}(\boldsymbol{\tau}) \succeq 0 \} \ , \tag{51}$$

in which

$$\boldsymbol{\Phi}_{\text{g}}(\boldsymbol{\tau}) = \begin{bmatrix} \boldsymbol{\Phi}_{\text{yc}}(\boldsymbol{\tau}) & \mathbf{p}_{\text{xx}} \\ \mathbf{p}_{\text{xx}}^{\text{H}} & p_{\text{xx}} - \phi \end{bmatrix}. \tag{52}$$

Finally, replacing $\boldsymbol{\tau}^{\text{opt}}$ in (48) leads to:

$$\mathbf{w}_{\text{SD-MWF-ILD}}(\boldsymbol{\tau}^{\text{opt}}) = \boldsymbol{\Phi}_{\text{yc}}^{\dagger}(\boldsymbol{\tau}^{\text{opt}}) \mathbf{p}_{\text{xx}} . \tag{53}$$

As a consequence of the strong duality of the QCQP in (36), it is proven in [23] that $\mathbf{w}_{\text{SD-MWF-ILD}} = \mathbf{w}_{\text{QC-MWF-ILD}}$. Therefore, one approach to determine $\mathbf{w}_{\text{QC-MWF-ILD}}$ is to find $\boldsymbol{\tau}^{\text{opt}}$ solving the convex SDP in (51) (which is performed in polynomial time [22]) and replace its optimal solution in (53).

VI. Semi-Analytical Closed-Form Solution

This section presents an efficient semi-analytical closed-form procedure to the original CO-MWF-ILD optimization problem. The procedure is based on solving the concave dual problem presented in (50). Firstly, the bidimensional dual problem in (50) is transformed into two univariate problems; then, the semidefinite constraint is converted into a set of linear constraints.

A. Reducing the Dimensionality of the Maximization Problem

To reduce the dimensionality of the bidimensional maximization problem in (50), (37) is substituted in (45), leading to the following form for the strong duality property:

$$\begin{aligned} J_{\text{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}}) = J_{\text{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}}) \\ + (\boldsymbol{\tau}^{\text{opt}})^{\text{T}} \mathbf{c}(\mathbf{w}_{\text{QC-MWF-ILD}}) \end{aligned} , \tag{54}$$

which necessarily implies that:

$$(\boldsymbol{\tau}^{\text{opt}})^{\text{T}} \mathbf{c}(\mathbf{w}_{\text{QC-MWF-ILD}}) = 0 . \tag{55}$$

Due to the non-negativity of the elements $\boldsymbol{\tau}^{\text{opt}}$ and non-positivity of the entries in $\mathbf{c}(\mathbf{w}_{\text{QC-MWF-ILD}})$, the identity in (55) can be represented as:

$$\tau_j^{\text{opt}} (-1)^{t+j-2} \mathbf{w}_{\text{QC-MWF-ILD}}{}^{\text{H}} \boldsymbol{\Phi}_{\text{cc}}(\delta_j) \mathbf{w}_{\text{QC-MWF-ILD}} = 0 , \tag{56}$$

for $j \in \{1, 2\}$. The identity in (56) implies that (55) is true. This occurs in the following situations: i) $\tau_j^{\text{opt}} = 0$; ii) $\mathbf{w}_{\text{QC-MWF-ILD}}{}^{\text{H}} \boldsymbol{\Phi}_{\text{cc}}(\delta_j) \mathbf{w}_{\text{QC-MWF-ILD}} = 0$; or iii) $\tau_j^{\text{opt}} = 0$ and $\mathbf{w}_{\text{QC-MWF-ILD}}{}^{\text{H}} \boldsymbol{\Phi}_{\text{cc}}(\delta_j) \mathbf{w}_{\text{QC-MWF-ILD}} = 0$, for $j \in \{1, 2\}$.

Considering that $\mathbf{w}_{\text{QC-MWF-ILD}}$ must be a feasible solution to the problem in (36) and since $\delta_j > 0$, the constraint in (36) implies that the feasible solution must satisfy one of the following three cases:

i) only the first constraint is active:

$$-(-1)^t \mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}} \boldsymbol{\Phi}_{\text{cc}}(\delta_1) \mathbf{w}_{\text{QC-MWF-ILD}} = 0 , \tag{57}$$

$$(-1)^t \mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}} \boldsymbol{\Phi}_{\text{cc}}(\delta_2) \mathbf{w}_{\text{QC-MWF-ILD}} < 0 . \tag{58}$$

ii) only the second constraint is active:

$$-(-1)^t \mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}} \boldsymbol{\Phi}_{\text{cc}}(\delta_1) \mathbf{w}_{\text{QC-MWF-ILD}} < 0 , \tag{59}$$

$$(-1)^t \mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}} \boldsymbol{\Phi}_{\text{cc}}(\delta_2) \mathbf{w}_{\text{QC-MWF-ILD}} = 0 . \tag{60}$$

iii) both constraints are inactive:

$$-(-1)^t \mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}} \boldsymbol{\Phi}_{\text{cc}}(\delta_1) \mathbf{w}_{\text{QC-MWF-ILD}} < 0 , \tag{61}$$

$$(-1)^t \mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}} \boldsymbol{\Phi}_{\text{cc}}(\delta_2) \mathbf{w}_{\text{QC-MWF-ILD}} < 0 . \tag{62}$$

From (57) to (62), it is possible to verify that any optimal solution $\mathbf{w}_{\text{QC-MWF-ILD}}$ will lead to at least one inactive constraint, which implies that at least one optimal Lagrange multiplier must be zero. Therefore, the optimal Lagrange multiplier $\boldsymbol{\tau}^{\text{opt}}$ will be one out of two kinds: $\boldsymbol{\tau}_1 = [\ \tau_1 \ 0\ ]^{\text{T}}$, associated with (57)-(58); or $\boldsymbol{\tau}_2 = [\ 0 \ \tau_2\ ]^{\text{T}}$, associated with (59)-(60). The Lagrange multiplier associated with (61)-(62) is a particular case of $\boldsymbol{\tau}_1$ or $\boldsymbol{\tau}_2$. Consequently, we can find the optimal solution to the problem in (50) by equivalently solving two univariate optimization problems corresponding to each of the situations above. These problems can be formulated as:

$$\begin{aligned} \tau_j^{\text{opt}} = \ & \underset{\tau_j}{\text{arg.max.}} \ -\mathbf{p}_{\text{xx}}^{\text{H}} \boldsymbol{\Phi}_{\text{yc}}^{\dagger}(\tau_j) \mathbf{p}_{\text{xx}} + p_{\text{xx}} \\ & \text{s.t.} \ \ \tau_j \geq 0, \quad \boldsymbol{\Phi}_{\text{yc}}(\tau_j) \succeq 0 \end{aligned} \tag{63}$$

for $j \in \{1, 2\}$, in which (40), (41), and (42) turn into

$$\boldsymbol{\Phi}_{\text{yc}}(\tau_j) = \text{diag}(\boldsymbol{\Phi}_{\text{yu}}(\beta_1(\tau_j)), \boldsymbol{\Phi}_{\text{yu}}(\beta_2(\tau_j))) , \tag{64}$$

$$\boldsymbol{\Phi}_{\text{yu}}(\beta_r(\tau_j)) = \boldsymbol{\Phi}_{\text{y}} + \beta_r(\tau_j) \boldsymbol{\Phi}_{\text{u}} , \tag{65}$$

$$\beta_r(\tau_j) = (-1)^{t+j+r} (\delta_j)^{r-1} \tau_j , \tag{66}$$

for $r \in \{1, 2\}$. Solving (63) for $j = 1$ and $j = 2$ leads to $\boldsymbol{\tau}_1^{\text{opt}}$ and $\boldsymbol{\tau}_2^{\text{opt}}$, respectively. The optimal solution to (50) is the one that maximizes the dual cost function, i.e.,

$$\boldsymbol{\tau}^{\text{opt}} = \text{arg.max.}\{ \ J_{\mathcal{L}}(\boldsymbol{\tau}_1^{\text{opt}}), \ J_{\mathcal{L}}(\boldsymbol{\tau}_2^{\text{opt}}) \ \} . \tag{67}$$

Finally, the solution $\mathbf{w}_{\text{QC-MWF-ILD}}$ is obtained by substituting the resulting $\boldsymbol{\tau}^{\text{opt}}$ in (53).

B. Simplification of the Constraint

The positive semidefinite constraint of $\boldsymbol{\Phi}_{\text{yc}}(\tau_j)$ makes the solution to the (63) computationally intensive. Here, we show that it is possible to write the semidefinite constraint equivalently in the form of a set of linear constraints.

Defining $\boldsymbol{\chi}(\mathbf{A})$ as the vector containing the eigenvalues of a generic diagonalizable matrix $\mathbf{A}$ and considering the block diagonal structure of $\boldsymbol{\Phi}_{\text{yc}}(\tau_j)$, then $\boldsymbol{\chi}(\boldsymbol{\Phi}_{\text{yc}}(\tau_j)) = [\ \boldsymbol{\chi}^{\text{T}}(\boldsymbol{\Phi}_{\text{yu}}(\beta_1(\tau_j))) \ \boldsymbol{\chi}^{\text{T}}(\boldsymbol{\Phi}_{\text{yu}}(\beta_2(\tau_j)))\ ]^{\text{T}}$ comprises the eigenvalues of the diagonal blocks [25]. Thus, matrix $\boldsymbol{\Phi}_{\text{yc}}(\tau_j)$ is positive semi-definite if and only if $\boldsymbol{\chi}(\boldsymbol{\Phi}_{\text{yu}}(\beta_1(\tau_j))) \geq 0$ and $\boldsymbol{\chi}(\boldsymbol{\Phi}_{\text{yu}}(\beta_2(\tau_j))) \geq 0$, i.e.,

$$\mathbf{\Phi}_{\text{yc}}(\tau_j) \succeq 0 \;\; \Leftrightarrow \;\; \boldsymbol{\chi}(\mathbf{\Phi}_{\text{yu}}(\beta_r(\tau_j))) \geq 0\,, \tag{68}$$

for $r \in \{1, 2\}$. Therefore, we need to determine conditions in which all eigenvalues of $\mathbf{\Phi}_{\text{yu}}(\beta_r(\tau_j))$ are nonnegative. The following theorem provides this:

Theorem I. Considering that $\mathbf{\Phi}_{\text{yu}}(\beta_r(\tau_j)) = \mathbf{\Phi}_{\text{y}}+\beta_r(\tau_j)\mathbf{\Phi}_{\text{u}}$, with $\mathbf{\Phi}_{\text{y}}$ defined in (16), $\mathbf{\Phi}_{\text{u}}$ defined in (19) and $\beta_r(\tau_j)$ defined in (66); if $\mathbf{\Phi}_{\text{y}}$ is a symmetric positive semidefinite matrix and $\mathbf{\Phi}_{\text{u}}$ is a symmetric rank-1 matrix, then the eigenvalues of $\mathbf{\Phi}_{\text{yu}}(\beta_r(\tau_j)))$ are non-negative if and only if

$$(-1)^{t+j+r}\,\tau_j \geq -\eta_{\text{b}}^{-1}(\delta_j)^{1-r}\,, \tag{69}$$

for $r \in \{1, 2\}$ and $\eta_{\text{b}} = p_{\text{su}}\mathbf{b}^{\text{H}}\mathbf{\Phi}_{\text{y}}^{\dagger}\mathbf{b} = \text{trace}(\mathbf{\Phi}_{\text{y}}^{\dagger}\mathbf{\Phi}_{\text{u}})$.

Proof: See Appendix A.

Considering (69) and the constraint $\tau_j \geq 0$, the feasible region of (63) can be written as:

$$\{\tau_j \geq 0\} \cap \{(-1)^{t+j+r}\,\tau_j \geq -\eta_{\text{b}}^{-1}(\delta_j)^{1-r}\}\,, \tag{70}$$

for $r \in \{1, 2\}$. Considering the cases for $r = 1$ and $r = 2$, the interval in (70) is given by

$$\{\tau_j \geq 0\} \cap \left\{(-1)^{t+j}\,\tau_j \leq \eta_{\text{b}}^{-1}\right\} \cap \left\{(-1)^{t+j}\,\tau_j \geq -\eta_{\text{b}}^{-1}\delta_j^{-1}\right\}. \tag{71}$$

Considering (71), the feasible regions of (63) for $t = 1$ and $t = 2$ are defined, respectively, as

$$0 \leq \tau_j \leq \eta_{\text{b}}^{-1}(\delta_j)^{1-j}\,, \tag{72}$$

$$0 \leq \tau_j \leq \eta_{\text{b}}^{-1}(\delta_j)^{j-2}. \tag{73}$$

The intervals in (72) and (73) can be generically represented by

$$0 \leq \tau_j \leq \eta_{\text{b}}^{-1}(\delta_j)^{f} \tag{74}$$

in which $f = (-1)^t\,(j - t)$. Using (74) in (63), leads to the following optimization problem:

$$\begin{aligned} \tau_j^{\text{opt}} = \;& \underset{\tau_j}{\text{arg.max.}} \; -\mathbf{p}_{\text{xx}}^{\text{H}}\mathbf{\Phi}_{\text{yc}}^{\dagger}(\tau_j)\mathbf{p}_{\text{xx}} + p_{\text{xx}} \\ & \text{s.t. } 0 \leq \tau_j \leq \eta_{\text{b}}^{-1}(\delta_j)^{f} \end{aligned}. \tag{75}$$

For each $j \in \{1, 2\}$, the optimization problem described in (75) encompasses a concave cost function with a linear interval constraint. Therefore, its stationary point will be either in the interior of this interval, or it is one of the limits of the feasible region.

C. Rewriting the Cost Function of the Dual Problem

Considering the expressions for the pseudo-inverse of a block diagonal matrix such as $\mathbf{\Phi}_{\text{yc}}$, and of the sum of a positive semidefinite and a rank-1 matrix such as $\mathbf{\Phi}_{\text{yu}}(\beta_r(\tau_j))$, given respectively in [26] and [27], the cost function of (75) is defined as

$$J_{\mathcal{L}}(\tau_j) = c_1\left[-\frac{(-1)^{j+t}\,\text{ILD}_{\text{in}}^{\text{x}}\,\tau_j}{1-(-1)^{j+t}\,\eta_{\text{b}}\,\tau_j} + \frac{(-1)^{j+t}\,\delta_j\,\tau_j}{1+(-1)^{j+t}\,\delta_j\eta_{\text{b}}\,\tau_j}\right] + c_2\,, \tag{76}$$

in which $c_1 = p_{\text{xR}}|\eta_{\text{ab}}|^2$; $c_2 = (\text{ILD}_{\text{in}}{}^{\text{x}}+1)p_{\text{xR}}|\eta_{\text{a}}|^2$; $\eta_{\text{a}} = p_{\text{sx}}\mathbf{a}^{\text{H}}\mathbf{\Phi}_{\text{y}}^{\dagger}\mathbf{a} = \text{trace}(\mathbf{\Phi}_{\text{y}}^{\dagger}\mathbf{\Phi}_{\text{x}})$; $\eta_{\text{b}} = p_{\text{su}}\mathbf{b}^{\text{H}}\mathbf{\Phi}_{\text{y}}^{\dagger}\mathbf{b} = \text{trace}(\mathbf{\Phi}_{\text{y}}^{\dagger}\mathbf{\Phi}_{\text{u}})$; $|\eta_{\text{ab}}|^2 = p_{\text{sx}}\,p_{\text{su}}\,|\mathbf{a}^{\text{H}}\mathbf{\Phi}_{\text{y}}^{\dagger}\mathbf{b}|^2 = \text{trace}(\mathbf{\Phi}_{\text{y}}^{\dagger}\mathbf{\Phi}_{\text{x}}\mathbf{\Phi}_{\text{y}}^{\dagger}\mathbf{\Phi}_{\text{u}})$; $p_{\text{sx}} = \mathbb{E}\{|s_{\text{x}}|^2\}$; $p_{\text{su}} = \mathbb{E}\{|s_{\text{u}}|^2\}$; $p_{\text{xR}} = \mathbf{q}_{\text{R}}{}^{\text{T}}\mathbf{\Phi}_{\text{x}}\mathbf{q}_{\text{R}}$; and $\text{ILD}_{\text{x}}{}^{\text{in}} = \mathbf{q}_{\text{L}}{}^{\text{T}}\mathbf{\Phi}_{\text{x}}\mathbf{q}_{\text{L}}(\mathbf{q}_{\text{R}}{}^{\text{T}}\mathbf{\Phi}_{\text{x}}\,\mathbf{q}_{\text{R}})^{-1}$.

Therefore, the dual problem in (75) can be represented as

$$\tau_j^{\text{opt}} = \underset{\tau_j}{\text{arg.max.}}\; J_{\mathcal{L}}(\tau_j) \text{ s.t. } 0 \leq \tau_j \leq \eta_{\text{b}}^{-1}(\delta_j)^{f}\,, \tag{77}$$

in which $J_{\mathcal{L}}(\tau_j)$ is defined in (76). There are two possibilities for τ_j^{opt} in (77). Either it belongs to the interior of the feasible region ($0 < \tau_j^{\text{opt}} < \eta_{\text{b}}{}^{-1}\delta_j{}^{f}$), in which case it is a stationary point of (76), or it is one of the limits of the feasible interval, i.e., $\tau_j^{\text{opt}} = 0$ or $\tau_j^{\text{opt}} = \eta_{\text{b}}{}^{-1}\delta_j{}^{f}$.

Thus, (77) can be solved by computing the set of possible solutions and finding the one that maximizes the objective function.

D. The Stationary Point Inside the Feasible Region

Since the dual problem is always concave, and considering the Hessian of $J_{\mathcal{L}}(\tau_j)$ is not singular, there is at most one stationary point inside the feasible region, in which case it is the optimal Lagrange multiplier. To obtain it, we compute the gradient of (76), leading to:

$$\frac{\partial J_{\mathcal{L}}(\tau_j)}{\partial \tau_j} = -\frac{c_1\,(-1)^{t+j}\,\text{ILD}_{\text{in}}^{\text{x}}}{((-1)^{t+j} - \eta_{\text{b}}\,\tau_j)^2} + \frac{c_1\,(-1)^{t+j}\,\delta_j}{((-1)^{t+j} + \eta_{\text{b}}\delta_j\,\tau_j)^2}\,. \tag{78}$$

Equating (78) to zero results in two possible solutions:

$$\frac{(-1)^{t+j} + \eta_{\text{b}}\delta_j\,\tau_j}{(-1)^{t+j} - \eta_{\text{b}}\,\tau_j} = (-1)^{i-1}\psi_j\,, \tag{79}$$

in which $i \in \{1, 2\}$; and

$$\psi_j = \delta_j^{1/2}(ILD_{\text{in}}^{\text{x}})^{-1/2}\,. \tag{80}$$

Rearranging (79) leads to:

$$(-1)^{t+j+1}(1+(-1)^i\psi_j) = \eta_{\text{b}}(\delta_j - (-1)^i\psi_j)\,\tau_j\,. \tag{81}$$

Therefore, the generic representation for the stationary points (sp) of the dual problem in (75) is given by:

$$\tau_{j,i}^{\text{sp}} = \frac{(-1)^{t+j+1}}{\eta_{\text{b}}}\,\frac{1+(-1)^i\psi_j}{\delta_j - (-1)^i\psi_j}\,. \tag{82}$$

TABLE I.

CONDITIONS FOR THE STATIONARY POINTS IN (82) TO BE CONTAINED IN THE FEASIBLE REGION OF (77), FOR $t \in \{1,2\}$, $j \in \{1,2\}$, AND $i \in \{1,2\}$.

		$i = 1$	$i = 2$
$t = 1$	$j = 1$	$0 \leq \frac{1}{\eta_{\text{b}}}\frac{\psi_1 - 1}{\psi_1 + \delta_1} \leq \frac{1}{\eta_{\text{b}}}$	$0 \leq \frac{1}{\eta_{\text{b}}}\frac{1+\psi_1}{\psi_1 - \delta_1} \leq \frac{1}{\eta_{\text{b}}}$
	$j = 2$	$0 \leq \frac{1}{\eta_{\text{b}}}\frac{1-\psi_2}{\delta_2 + \psi_2} \leq \frac{1}{\eta_{\text{b}}\delta_2}$	$0 \leq \frac{1}{\eta_{\text{b}}}\frac{1+\psi_2}{\delta_2 - \psi_2} \leq \frac{1}{\eta_{\text{b}}\delta_2}$
$t = 2$	$j = 1$	$0 \leq \frac{1}{\eta_{\text{b}}}\frac{1-\psi_1}{\delta_1 + \psi_1} \leq \frac{1}{\eta_{\text{b}}\delta_1}$	$0 \leq \frac{1}{\eta_{\text{b}}}\frac{1+\psi_1}{\delta_1 - \psi_1} \leq \frac{1}{\eta_{\text{b}}\delta_1}$
	$j = 2$	$0 \leq \frac{1}{\eta_{\text{b}}}\frac{\psi_2 - 1}{\delta_2 + \psi_2} \leq \frac{1}{\eta_{\text{b}}}$	$0 \leq \frac{1}{\eta_{\text{b}}}\frac{1+\psi_2}{\psi_2 - \delta_2} \leq \frac{1}{\eta_{\text{b}}}$

Note that (82) represents two solutions, one for $i = 1$ and another for $i = 2$. However, one can be eliminated since at most one of the stationary points will be a feasible solution to the dual problem. This can be verified by replacing (82) in the feasible region of (77) considering the values for t, j, and i that lead to

This article has been accepted for publication in IEEE/ACM Transactions on Audio, Speech and Language Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TASLP.2023.3291530

the conditions for the stationary points to be feasible solutions, as shown in TABLE I. Manipulating the inequalities for each entry of TABLE I leads to the equivalent interval representations shown in TABLE II.

TABLE II.
EQUIVALENT REPRESENTATION OF THE FEASIBLE REGIONS IN TABLE I.

		$i = 1$	$i = 2$
$t = 1$	$j = 1$	$\{\psi_1 \geq 1\} \cap \{\delta_1 \geq -1\}$	$\{\psi_1 \geq \delta_1\} \cap \{\delta_1 \leq -1\}$
	$j = 2$	$\{\psi_2 \leq 1\} \cap \{\delta_2 \geq -1\}$	$\{\psi_2 \leq \delta_2\} \cap \{\delta_2 \leq -1\}$
$t = 2$	$j = 1$	$\{\psi_1 \leq 1\} \cap \{\delta_1 \geq -1\}$	$\{\psi_1 \leq \delta_1\} \cap \{\delta_1 \leq -1\}$
	$j = 2$	$\{\psi_2 \geq 1\} \cap \{\delta_2 \geq -1\}$	$\{\psi_2 \geq \delta_2\} \cap \{\delta_2 \leq -1\}$

TABLE III
PSEUDOCODE FOR THE CB-MWF-ILD PROPOSED METHOD.

Input: $\mathbf{\Phi}_y$, $\mathbf{\Phi}_u$, $\mathbf{\Phi}_x$, δ

Output: $\mathbf{w}_{\text{CB-MWF-ILD}}$

If $\text{ILD}^u_{in} \neq 1$

  Compute δ_1 and δ_2 using (32) and δ
  Compute ILD_{in}^x using $\mathbf{\Phi}_x$ in (23)
  Compute ψ_1 and ψ_2 using (80) and ILD_{in}^x
  Initialize $\tau_1^{opt} = 0$ and $\tau_2^{opt} = 0$
If $t = 1$ **and** $\psi_1 \geq 1$ **then**
  Compute $\tau_1^{opt} = \tau_{1,1}^{ifr}$ using (83), for $t = 1$ and $j = 1$
end If
If $t = 1$ **and** $\psi_2 \leq 1$ **then**
  Compute $\tau_2^{opt} = \tau_{2,1}^{ifr}$ using (83), for $t = 1$ and $j = 2$
end If
If $t = 2$ **and** $\psi_1 \leq 1$ **then**
  Compute $\tau_1^{opt} = \tau_{1,2}^{ifr}$ using (83), for $t = 2$ and $j = 1$
end If
If $t = 2$ **and** $\psi_2 \geq 1$ **then**
  Compute $\tau_2^{opt} = \tau_{2,2}^{opt}$ using (83), for $t = 2$ and $j = 2$
end If

  If $J_{\mathcal{L}}(\tau_1^{opt}) \geq J_{\mathcal{L}}(\tau_2^{opt})$ **then**
    $\boldsymbol{\tau}^{opt} = [\ \tau_1^{opt}\ \ 0\ ]^T$
  else
    $\boldsymbol{\tau}^{opt} = [\ 0\ \ \tau_2^{opt}\ ]^T$
  end If

  $\mathbf{w}_{\text{CB-MWF-ILD}} = \mathbf{\Phi}^{\dagger}_{yc}(\boldsymbol{\tau}^{opt})\mathbf{p}_{xx}$
else
  $\mathbf{w}_{\text{MWF-ILD}} = \mathbf{\Phi}^{\dagger}_{yy}\mathbf{p}_{xx}$
end If

Since $\delta_j > 0$ (see (32)) and $\psi_j > 0$ (see (80)), the stationary points associated with column $i = 2$ in TABLE II. will never be inside of the feasible region (since this would require that $\delta_j < 0$). Therefore, the stationary points inside the feasible region (ifr) can only be those for which $i = 1$, i.e.,

$$\tau_{j,t}^{\text{ifr}} = \frac{(-1)^{t+j+1}}{\eta_b}\frac{1-\psi_j}{\delta_j+\psi_j}. \tag{83}$$

E. The Stationary Points Outside the Feasible Region

If the stationary points are outside the feasible region, the maximum value of the dual problem is achieved at one of the limits of the feasible region. Considering the problem in (77), $\tau_j = 0$ always leads to $J_{\mathcal{L}}(\tau_j) = c_2$. Also, as τ_j approaches $\eta_b^{-1}\delta_j{}^f$ the value of $J_{\mathcal{L}}(\tau_j)$ approaches $-\infty$. Therefore, the upper limit of the feasible interval is not a solution to the problem for any combination of $f \in \{1,2\}$ and $j \in \{1,2\}$, since $J_{\mathcal{L}}(0) > J_{\mathcal{L}}(\eta_b^{-1}\delta_j{}^f)$. Therefore, the optimal Lagrange multiplier is zero ($\tau_j^{opt} = 0$) when the stationary point of the cost function is outside the feasible region.

F. Implementation

TABLE III presents the pseudocode to implement the proposed semi-analytical closed-form solution to the constrained binaural MWF-ILD named CB-MWF-ILD noise reduction method. It comprises equations (53), (83) and the reasoning presented in Section VI.E. It is worth noting that for $\text{ILD}^u_{in} = 1$, the conventional MWF is the solution to the problem; its closed formula may be directly applied to speed up the calculation.

G. Relationship of the CB-MWF-ILD with other techniques

Assuming $\mathbf{\Phi}_y \succ 0$, the CB-MWF-ILD filters are given by:

$$\mathbf{w}_{\text{CB-MWF-ILD}}(\boldsymbol{\tau}^{opt}) = \mathbf{\Phi}_{yc}^{-1}(\boldsymbol{\tau}^{opt})\mathbf{p}_{xx}. \tag{84}$$

The left and right filters in (84) can be written as:

$$\mathbf{w}_{\text{CB-MWF-ILD},\ell}(\boldsymbol{\tau}^{opt}) = \mathbf{\Phi}_{yu}^{-1}(\beta_r(\boldsymbol{\tau}^{opt}))\mathbf{\Phi}_{xx}\mathbf{q}_\ell, \tag{85}$$

in which $r = 1$ when $\ell = \text{L}$ and $r = 2$ when $\ell = \text{R}$. Using (19) in (41), replacing the result in (85), and finally using the Sherman-Morrison formula [25], the left and right CB-MWF-ILD solutions can be decomposed as:

$$\mathbf{w}_{\text{CB-MWF-ILD},\ell} = \mathbf{w}^x_{\text{MWF},\ell} - \frac{\beta_1(\boldsymbol{\tau})\varphi_{ba}}{1+\beta_1(\boldsymbol{\tau})\varphi_b}\frac{a_L^*}{p_{su}b_L^*}\mathbf{w}^u_{\text{MWF},\ell}, \tag{86}$$

in which $\varphi_{ba} = \mathbf{b}^H\mathbf{\Phi}_y^{-1}\mathbf{a}$; $\varphi_b = \mathbf{b}^H\mathbf{\Phi}_y^{-1}\mathbf{b}$; $\mathbf{w}^x_{\text{MWF},\ell} = \mathbf{\Phi}_y^{-1}\mathbf{\Phi}_x\mathbf{q}_\ell$ is the conventional binaural MWF, while $\mathbf{w}^u_{\text{MWF},\ell} = \mathbf{\Phi}_y^{-1}\mathbf{\Phi}_u\mathbf{q}_\ell$ is its counterpart related to the interfering source. In the light of this decomposition, the CB-MWF-ILD can be related to the MVDR beamformer, leading to [12]:

$$\mathbf{w}_{\text{CB-MWF-ILD},\ell} = \upsilon_a\mathbf{w}^x_{\text{MVDR},\ell} - \frac{\beta_1(\boldsymbol{\tau})(\varphi_{ba}a_L^*\upsilon_b)}{(1+\beta_1(\boldsymbol{\tau})\varphi_b)p_{su}b_L^*}\mathbf{w}^u_{\text{MVDR},\ell}, \tag{87}$$

in which $\upsilon_a = p_{sx}/(1+p_{sx}\,\varphi_a)$; $\varphi_a = \mathbf{a}^H\mathbf{\Phi}_y^{-1}\mathbf{a}$; $\upsilon_b = p_{sx}/(1+p_{sx}\,\varphi_b)$; $\varphi_b = \mathbf{b}^H\mathbf{\Phi}_y^{-1}\mathbf{b}$; $\mathbf{w}^x_{\text{MVDR},\ell} = a^*_\ell\mathbf{\Phi}_y^{-1}\mathbf{a}/\varphi_a$ is the conventional MVDR, while $\mathbf{w}^u_{\text{MVDR},\ell} = b^*_\ell\mathbf{\Phi}_y^{-1}\mathbf{b}/\varphi_b$ is its counterpart related to the interfering source.

Equation (87) also relates the CB-MWF-ILD with the parametric unconstrained beamformer (see Section IV in [28]).

VII. COMPUTER SIMULATIONS

In this section, computer simulations are presented to assess the performance of the CB-MWF-ILD semi-analytical closed-form solution provided in TABLE III. Comparisons were performed with four techniques: i) the conventional MWF described in (20); ii) the conventional MWF-ILD described in

This article has been accepted for publication in IEEE/ACM Transactions on Audio, Speech and Language Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TASLP.2023.3291530

(25); iii) the MWF-ITF ($J_{\text{MWF-ITF}} = J_{\text{MWF}} + \rho \cdot J^{\text{u}}_{\text{ITF}}$) whose cost function is given by [12] [15]:

$$J^{\text{u}}_{\text{ITF}}(\mathbf{w}) = \frac{\mathbf{w}^{\text{H}}\boldsymbol{\Phi}_{\text{u1}}\mathbf{w}}{\mathbf{w}^{\text{H}}\boldsymbol{\Phi}_{\text{u2}}\mathbf{w}}, \tag{88}$$

in which:

$$\boldsymbol{\Phi}_{\text{u1}} = \begin{bmatrix} \boldsymbol{\Phi}_{\text{u}} & -(\text{ITF}^{\text{u}}_{\text{in}})^{*}\boldsymbol{\Phi}_{\text{u}} \\ -\text{ITF}^{\text{u}}_{\text{in}}\boldsymbol{\Phi}_{\text{u}} & |\text{ITF}^{\text{u}}_{\text{in}}|^{2}\,\boldsymbol{\Phi}_{\text{u}} \end{bmatrix},\ \boldsymbol{\Phi}_{\text{u2}} = \begin{bmatrix} \mathbf{0}_{M\times M} & \mathbf{0}_{M\times M} \\ \mathbf{0}_{M\times M} & \boldsymbol{\Phi}_{\text{u}} \end{bmatrix}, \tag{89}$$

$\text{ITF}^{\text{u}}_{\text{in}} = (\mathbf{q}_{\text{L}}^{\text{T}}\boldsymbol{\Phi}_{\text{u}}\mathbf{q}_{\text{L}})/(\mathbf{q}_{\text{R}}^{\text{T}}\boldsymbol{\Phi}_{\text{u}}\mathbf{q}_{\text{R}})$, and $(\cdot)^{*}$ is the complex conjugate operation; and iv) the MWF-PNE [12], whose solution is given by:

$$\mathbf{w}_{\text{MWF-PNE}} = (1-\kappa)\mathbf{w}_{\text{MWF}} + \kappa\mathbf{q}, \tag{90}$$

in which κ is the scaling parameter, and $\mathbf{q}$ is the selection vector.

A. Acoustic Scenario

The acoustic scenario comprises two pointwise sources (the desired speech and the interfering source) and additive (environmental and electrical) noise. Pointwise sources were simulated by convolving signals with the corresponding acoustic transfer functions (ATFs). The ATFs characterize the acoustic path between the source location and the acquisition microphones in a pair of behind-the-ear HA mounted on an artificial head and torso mannequin. The room reverberation time is $T_{60} \cong 300$ ms [29]. Thirty different speech signals were selected from the database presented in [30]. Half of them were considered speech of interest, while the other half were interference. Signals are two seconds long on average and uttered by a single speaker.

The total number of microphones is $M = 6$ (3 on each side). The desired source was emulated at 0° azimuth (in front of the HA), while the azimuth of the interfering source (θ_{u}) was varied from −90° (left side) to 90° (right side) in steps of 10°, resulting in 19 different azimuths. Both pointwise sources' elevation angle and radial distance were set at 0° and 3 m, respectively.

The environmental noise was assumed cylindrical and isotropic. It was generated according to the procedure defined in [31] and [32].

The long-term signal-to-interference ratio (SIR) and signal-to-noise ratio (SNR) were both set to 0 dB, resulting in a signal-to-interference-plus-noise ratio (SINR) of −3 dB.

For each noise reduction technique, 285 simulations (15 sets of signals and 19 azimuths for the interfering source) were performed.

B. Time-Frequency Representation

The input signals were sampled at 16 kHz and processed in frames of 256 samples with 50% overlap. Frames were weighted by an analysis window and transformed to the frequency domain using the Short-time Fourier transform with $K = 512$ points, using zero padding. After filtering by the noise reduction method, the processed signals were transformed back to the time domain using the inverse Fourier transform. A synthesis window weighted adjacent frames, and an overlap-and-add algorithm was applied to restore the filtered signal to the time domain. The square root of the Hann window was used for both analysis and synthesis.

For obtaining an upper bound performance, a batch procedure for estimating the required coherence matrices was implemented, resulting in a unique coefficient vector $\mathbf{w}$ per bin, which was calculated a priori and then employed for filtering the whole noisy speech [32] [33] [34].

The methodology used to estimate the coherence matrices was the same used in [35], [36], and [37]. We assume the use of a (own) voice activity detector [38] for segmenting background noise, interference plus noise, and speech plus noise epochs to estimate $\boldsymbol{\Phi}_{\text{n}}$; $\boldsymbol{\Phi}_{\text{un}} = \boldsymbol{\Phi}_{\text{u}} + \boldsymbol{\Phi}_{\text{n}}$ and $\boldsymbol{\Phi}_{\text{xn}} = \boldsymbol{\Phi}_{\text{x}} + \boldsymbol{\Phi}_{\text{n}}$ using:

$$\hat{\boldsymbol{\Phi}}_{\text{d}}(k) = \frac{1}{P_{\text{d}}}\sum_{\lambda=1}^{P_{\text{d}}}\mathbf{d}(\lambda,k)\mathbf{d}^{\text{H}}(\lambda,k), \tag{91}$$

in which $d \in \{\text{n, un, xn}\}$; and P_{d} is the number of frames used to estimate the coherence matrices.

Estimation of $\boldsymbol{\Phi}_{\text{x}}$ and $\boldsymbol{\Phi}_{\text{u}}$ was performed using the covariance whitening method [35] [36] [37] [39] [40]. Firstly, a squared-root decomposition of $\hat{\boldsymbol{\Phi}}_{\text{n}}$ is performed, resulting in:

$$\hat{\boldsymbol{\Phi}}_{\text{n}}(k) = \hat{\boldsymbol{\Phi}}_{\text{n}}^{1/2}(k)\hat{\boldsymbol{\Phi}}_{\text{n}}^{\text{H}/2}(k). \tag{92}$$

Then, matrices $\hat{\boldsymbol{\Phi}}_{\text{xn}}$ and $\hat{\boldsymbol{\Phi}}_{\text{un}}$ are prewhitened using (92), i.e.,

$$\hat{\boldsymbol{\Phi}}^{\text{w}}_{\text{xn}}(k) = \hat{\boldsymbol{\Phi}}_{\text{n}}^{-1/2}(k)\hat{\boldsymbol{\Phi}}_{\text{xn}}(k)\hat{\boldsymbol{\Phi}}_{\text{n}}^{-\text{H}/2}(k), \tag{93}$$

$$\hat{\boldsymbol{\Phi}}^{\text{w}}_{\text{un}} = \hat{\boldsymbol{\Phi}}_{\text{n}}^{-1/2}(k)\hat{\boldsymbol{\Phi}}_{\text{un}}(k)\hat{\boldsymbol{\Phi}}_{\text{n}}^{-\text{H}/2}(k). \tag{94}$$

Finally, $\boldsymbol{\Phi}_{\text{x}}$ and $\boldsymbol{\Phi}_{\text{u}}$ were estimated as rank-1 matrices according to:

$$\hat{\boldsymbol{\Phi}}_{\text{x}}(k) = \hat{p}_{\text{sx}}(k)\hat{\mathbf{a}}(k)\hat{\mathbf{a}}^{\text{H}}(k), \tag{95}$$

$$\hat{\boldsymbol{\Phi}}_{\text{u}}(k) = \hat{p}_{\text{su}}(k)\hat{\mathbf{b}}(k)\hat{\mathbf{b}}^{\text{H}}(k), \tag{96}$$

in which

$$\hat{p}_{\text{sx}}(k) = \nu\{\hat{\boldsymbol{\Phi}}^{\text{w}}_{\text{xn}}\},\ \hat{p}_{\text{su}}(k) = \nu\{\hat{\boldsymbol{\Phi}}^{\text{w}}_{\text{xu}}\}, \tag{97}$$

$$\hat{\mathbf{a}}(k) = \hat{\boldsymbol{\Phi}}_{\text{n}}\mathbf{f}\{\hat{\boldsymbol{\Phi}}^{\text{w}}_{\text{xn}}\},\ \hat{\mathbf{b}}(k) = \hat{\boldsymbol{\Phi}}_{\text{n}}\mathbf{f}\{\hat{\boldsymbol{\Phi}}^{\text{w}}_{\text{un}}\}, \tag{98}$$

where $\nu\{\cdot\}$ is the largest eigenvalue of the matrix in its argument; and $\mathbf{f}\{\cdot\}$ is its associated eigenvector. This method is widely used for speech ($\boldsymbol{\Phi}_{\text{x}}$) and interference ($\boldsymbol{\Phi}_{\text{u}}$) coherence matrix estimation in low-SNR complex acoustic scenarios [41].

C. Parameters and Optimization Algorithms

The '`fmincon`' routine from Matlab was employed to obtain the optimal solution for the conventional MWF-ILD method defined in (25) [42] and the MWF-ITF with cost function presented in (88) [15]. Both gradient and Hessian (see Appendix B) for equations (21) and (88) were provided to the solver. Parameters $\alpha(\lambda,k)$ and $\rho(\lambda,k)$ were defined through the iterative algorithm presented in Appendix C and kept fixed for all λ. Vector $\boldsymbol{\alpha} \in \mathbb{R}^{\text{G}}_{++}$ was defined by $G = 500$ elements linearly distributed in the logarithm scale between 10^{-10} and 1. This iterative algorithm is similar to the algorithm presented in [32]. The ILD tolerance $\delta(\lambda,k)$ was set to 10^{-6}.

D. Objective Performance Measures

Seven objective measures were used to assess the noise reduction and binaural cue preservation performances of the investigated methods.

Speech quality for the left and right ears was measured by the

wideband perceptual evaluation of the speech quality (PESQ) [43], and intelligibility was estimated with the short-time objective indelibility (STOI) [44].

The signal-to-interfering-plus-noise ratio variation (ΔSINR), the signal-to-interfering ratio variation (ΔSIR), and the signal-to-noise ratio variation (ΔSNR) are defined as [35]

$$\Delta\text{SINR} = \frac{10}{K}\sum_{k=1}^{K}\log_{10}\left(\frac{p_{zx_{LR}}(k)\,p_{v_{LR}}(k)}{p_{zv_{LR}}(k)\,p_{x_{LR}}(k)}\right), \tag{99}$$

$$\Delta\text{SIR} = \frac{10}{K}\sum_{k=1}^{K}\log_{10}\left(\frac{p_{zx_{LR}}(k)\,p_{u_{LR}}(k)}{p_{zu_{LR}}(k)\,p_{x_{LR}}(k)}\right), \tag{100}$$

$$\Delta\text{SNR} = \frac{10}{K}\sum_{k=1}^{K}\log_{10}\left(\frac{p_{zx_{LR}}(k)\,p_{n_{LR}}(k)}{p_{zn_{LR}}(k)\,p_{x_{LR}}(k)}\right), \tag{101}$$

in which $p_{zd_{LR}} = \mathbf{w}_L{}^H\mathbf{\Phi}_d\mathbf{w}_R$; $p_{d_{LR}} = \mathbf{q}_L{}^T\mathbf{\Phi}_d\mathbf{q}_R$; and d ∈ { n,u,x }.

The ILD and IPD variations, which measure the input-output binaural cue preservation, are defined respectively as [17] [32]

$$\Delta\text{ILD}_d = \frac{10}{K}\sum_{k=1}^{K}\left|\log_{10}\left(\frac{p_{zd_L}(k)\,p_{d_R}(k)}{p_{zd_R}(k)\,p_{d_L}(k)}\right)\right|, \tag{102}$$

$$\Delta\text{IPD}_d = \frac{1}{K\pi}\sum_{k=1}^{K}\left|\angle p_{zd_{LR}}(k) - \angle p_{d_{LR}}(k)\right|, \tag{103}$$

in which $p_{zd_\ell} = \mathbf{w}_\ell{}^H\mathbf{\Phi}_d\mathbf{w}_\ell$; $p_{d_\ell} = \mathbf{q}_\ell{}^T\mathbf{\Phi}_d\mathbf{q}_\ell$; and d ∈ { n,u,x }. The global performance for each noise reduction technique was calculated for all signals and interference azimuths (285 values).

VIII. Results and Discussion

This section presents performance results for the proposed CB-MWF-ILD method and compares it with the conventional MWF, MWF-ILD, MWF-ITF, and MWF-PNE.

The dissimilarity between the employed objective criteria for the left and right sides results from the input/output ILD definitions presented in (23) and (24), as described in [17].

Fig. 2 and Fig. 3 depict input-output ILD variation for the interference and speech signals, respectively. As expected, the conventional MWF severely distorts the interference ILD compared to the MWF-ITF and MWF-ILD techniques. This observation is corroborated by the ΔILD_u (for all θ_u) presented in TABLE IV, in which: ΔILD_u = 8.8 dB, ΔILD_u = 3.2 dB, ΔILD_u = 2.4 dB, and ΔILD_u = 2.4 dB for the MWF, MWF-ITF, MWF-ILD, and CB-MWF-ILD, respectively. The proposed method and the MWF-ILD result in similar ILD preservation for interference. Both MWF-ILD techniques present a difference of 0.8 dB on the median ΔILD_u compared to the MWF-ITF. The CB-MWF-ILD results in the same speech ILD distortion as both the MWF-ILD and the MWF-ITF and approximately the same as the conventional MWF.

Fig. 4 and Fig. 5 show ΔIPD_u and ΔIPD_x. The speech ΔIPD of the proposed method is very small and is in the same range as the MWF, MWF-ILD, and MWF-ITF. The interference ΔIPD is not negligible but in the same range as the MWF and the MWF-ILD techniques. This side effect may influence the localization of low-frequency interfering acoustic sources by the HA user due to incoherent ILD and IPD binaural cues. However, this situation was not observed in the processed signals[2]. The MWF-ITF presents lower interference ΔIPD values than the MWF and MWF-ILD techniques at the cost of a higher speech ΔIPD for azimuths near 0°. This behavior is expected since the ITF cost function accounts for interference IPD and ILD preservation. TABLE IV also indicates that despite some variations, ΔIPD_u and ΔIPD_x for all interference azimuths and methods are in the same range.

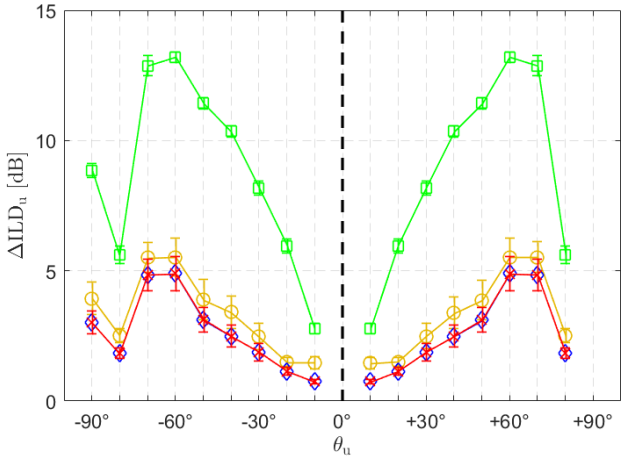

Fig. 2. Interference ΔILD calculated from 15 input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◊), CB-MWF-ILD in red (×), and MWF-ITF in yellow (○).

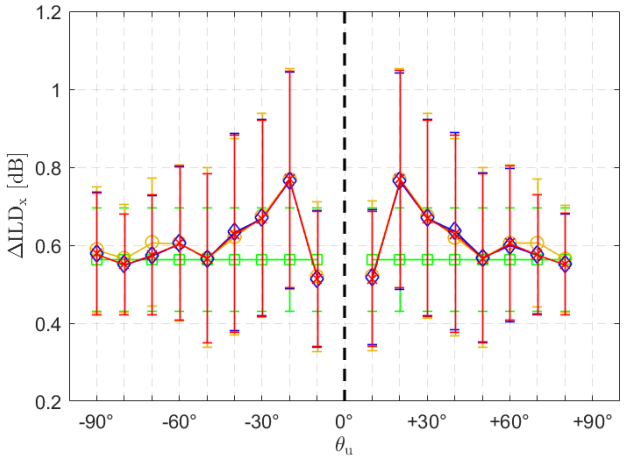

Fig. 3. Speech ΔILD calculated from 15 input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◊), CB-MWF-ILD in red (×), and MWF-ITF in yellow (○).

[2] Audio files with some examples obtained from the investigated acoustic scenarios can be found in https://github.com/diego-carmo/Closed-MWF-ILD/blob/main/README.md

This article has been accepted for publication in IEEE/ACM Transactions on Audio, Speech and Language Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TASLP.2023.3291530

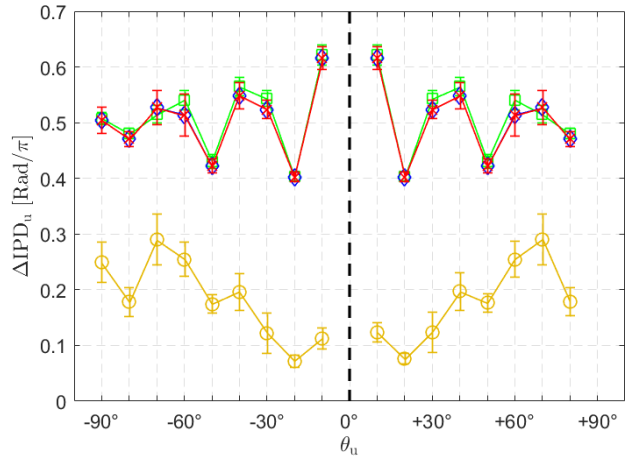

Fig. 4. Interference ΔIPD calculated from 15 input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◊), CB-MWF-ILD in red (×), and MWF-ITF in yellow (○).

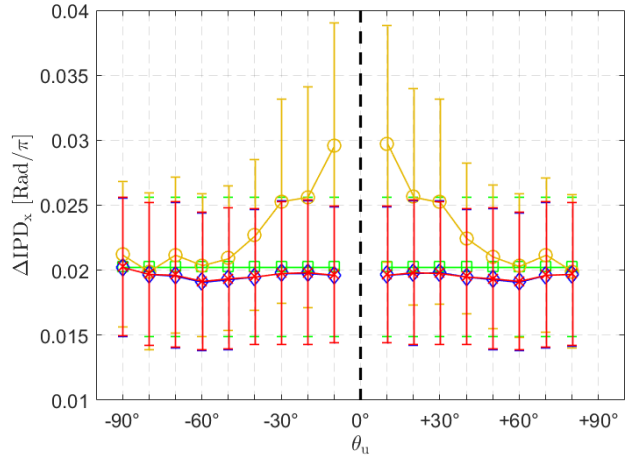

Fig. 5. Speech ΔIPD calculated from 15 input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◊), CB-MWF-ILD in red (×), and MWF-ITF in yellow (○).

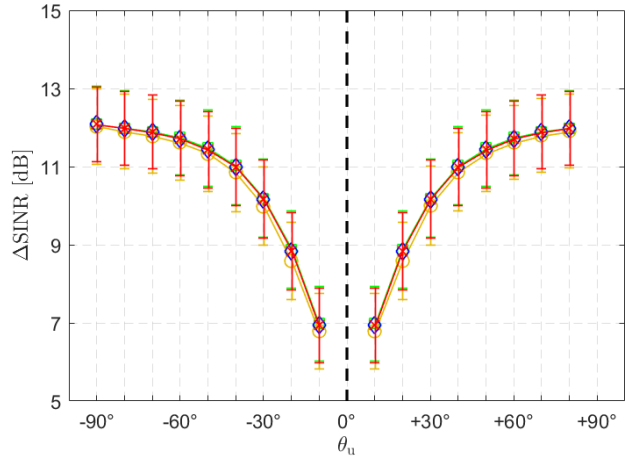

Fig. 6. ΔSINR calculated from 15 different input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◊), CB-MWF-ILD in red (×), and MWF-ITF in yellow (○).

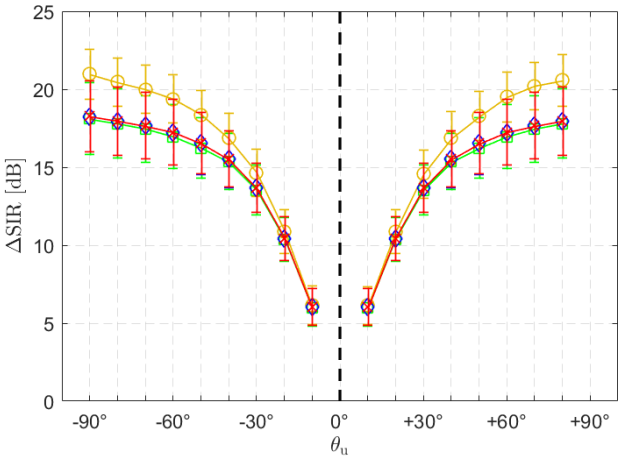

Fig. 7. ΔSIR calculated from 15 different input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◊), CB-MWF-ILD in red (×), and MWF-ITF in yellow (○).

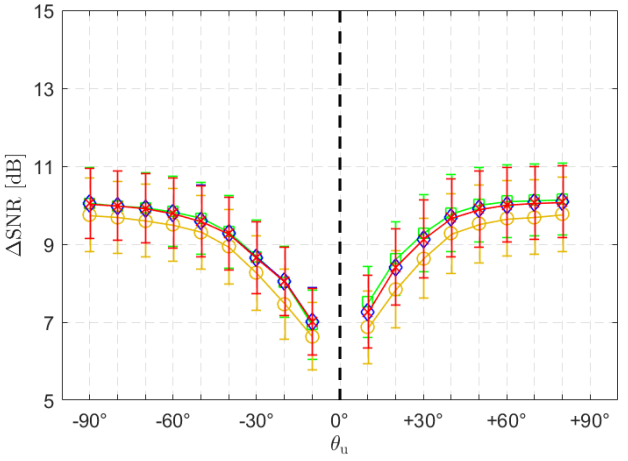

Fig. 8. ΔSNR calculated from 15 different input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◊), CB-MWF-ILD in red (×), and MWF-ITF in yellow (○).

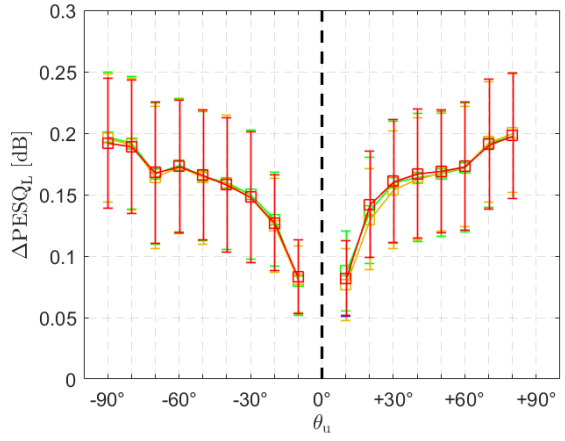

Fig. 9. $\Delta PESQ_L$ calculated from 15 different input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◊), CB-MWF-ILD in red (×), and MWF-ITF in yellow (○).

This article has been accepted for publication in IEEE/ACM Transactions on Audio, Speech and Language Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TASLP.2023.3291530

TABLE IV
GLOBAL MEDIAN AND STANDARD DEVIATION ($M \pm \Sigma$) FOR SPEECH (X) AND INTERFERING (U) ΔILD AND ΔIPD.

	ΔILD_u [dB]	ΔILD_x [dB]	ΔIPD_u [rad/π]	ΔIPD_x [rad/π]
MWF	8.8 ± 3.3	0.5 ± 0.1	$0.5 \pm 6{\cdot}10^{-2}$	$2{\cdot}10^{-2} \pm 5{\cdot}10^{-3}$
MWF-ITF	3.2 ± 1.5	0.6 ± 0.2	$0.1 \pm 7{\cdot}10^{-2}$	$2{\cdot}10^{-2} \pm 7{\cdot}10^{-3}$
MWF-ILD	2.4 ± 1.4	0.6 ± 0.2	$0.5 \pm 6{\cdot}10^{-2}$	$2{\cdot}10^{-2} \pm 5{\cdot}10^{-3}$
CB-MWF-ILD	2.4 ± 1.4	0.6 ± 0.2	$0.5 \pm 6{\cdot}10^{-2}$	$2{\cdot}10^{-2} \pm 5{\cdot}10^{-3}$

TABLE V
GLOBAL MEDIAN AND STANDARD DEVIATION ($M \pm \Sigma$) FOR ΔSINR, ΔSIR, AND ΔSNR.

	ΔSINR [dB]	ΔSIR [dB]	ΔSNR [dB]
MWF	11.1 ± 1.8	15.1 ± 3.9	9.3 ± 1.3
MWF-ITF	11.0 ± 1.9	17.3 ± 4.7	8.9 ± 1.3
MWF-ILD	11.1 ± 1.8	15.3 ± 4.0	9.3 ± 1.3
CB-MWF-ILD	11.1 ± 1.8	15.3 ± 4.0	9.3 ± 1.3

TABLE VI
AVERAGE TIME FOR CALCULATING THE OPTIMAL FILTERS ($\mathbf{w}_L$ AND $\mathbf{w}_R$) IN THE ASSESSED ACOUSTIC SCENARIOS.

	MWF	MWF-ITF	MWF-ILD	CB-MWF-ILD
Time [s]	$7{\cdot}10^{-5}$	3.2	2.6	$2{\cdot}10^{-4}$

Fig. 6 to Fig. 8 show ΔSINR, ΔSIR, and ΔSNR for distinct $\theta_u \neq \theta_x = 0°$, respectively. The proposed method presents a ΔSINR reduction performance equivalent to the conventional MWF and MWF-ILD for all interference azimuths (TABLE V). Considering the ΔSIR, the MWF-ITF presents a performance of 2 dB higher than the MWF and MWF-ILD techniques. On the other hand, considering the ΔSNR, the MWF-ITF presents a reduction of 0.5 dB compared to the other techniques.

Fig. 9 to Fig. 12 show $\Delta PESQ_L$, $\Delta PESQ_R$, $\Delta STOI_L$, and $\Delta STOI_R$, respectively. Differences between PESQ median values are lower than 0.2 units indicating no relevant perceptual variation in speech quality [45]. Similar results are observed for intelligibility, since STOI differences, for both left and right sides, are lower than 0.1 [46].

Fig. 13 and Fig. 14 present comparisons of the CB-MWF-ILD performance with the MWF-ITF and MWF-PNE methods for $\theta_u = -60°$. ΔILD_u and ΔSINR criteria were evaluated for the following set of parameters: δ (CB-MWF-ILD), ρ (MWF-ITF) and, κ (MWF-PNE) $\in \{ 10^{-6}\ 0.1,\ 0.2,\ \ldots,\ 1 \}$. The leftmost point in each plot is $\delta = \rho = \kappa = 10^{-6}$. Note that these parameters have different interpretations, despite setting the trade-off between noise reduction and spatial preservation of the interference source. Results presented in both Fig. 13 and Fig. 14 indicate that CB-MWF-ILD and MWF-ITF are barely influenced by changes of δ and ρ in the range analyzed. On the contrary, the choice of κ strongly impacts the MWF-PNE performance. Furthermore, comparing CB-MWF-ILD and MWF-PNE for the same ΔILD_u setpoint (ΔILD_u = 4.85 dB), the CB-MWF-ILD presents 3.3 dB higher ΔSINR than the MWF-PNE. Considering the case where both methods have the same ΔSINR setpoint (ΔSINR = 11.75 dB), the CB-MWF-ILD results in 7.7 dB lower ΔILD_u than the MWF-PNE. Therefore, considering the same setpoint for a given metric, the CB-MWF-ILD performs better than the MWF-PNE.

Massive simulations performed on a desktop personal computer with an Intel® Core i7-3770 processor, running at 3.40 GHz, and Matlab® indicated that the numerical process required for obtaining the optimal filters ($\mathbf{w}_L$ and $\mathbf{w}_R$) for the CB-MWF-ILD is 13,000 and 16,000 times faster than for the conventional (unconstrained) MWF-ILD and the MWF-ITF methods, respectively; but only 0.35 times slower than the conventional MWF, on average (see TABLE VI).

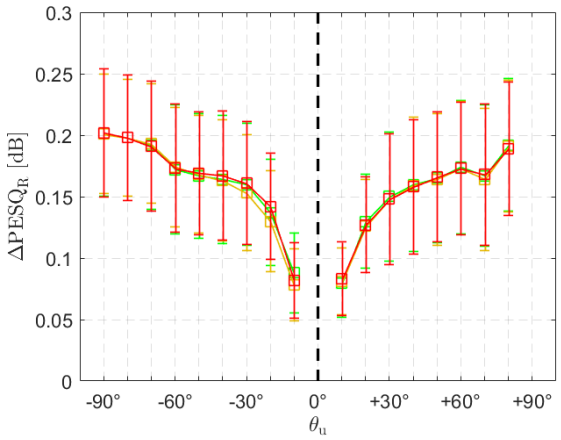

Fig. 10. $\Delta PESQ_R$ calculated from 15 different input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◊), CB-MWF-ILD in red (×), and MWF-ITF in yellow (○).

This article has been accepted for publication in IEEE/ACM Transactions on Audio, Speech and Language Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TASLP.2023.3291530

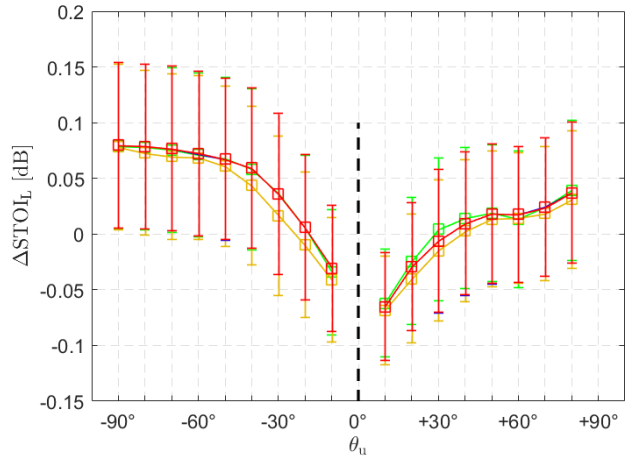

Fig. 11. $\Delta\text{STOI}_\text{L}$ calculated from 15 different input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◊), CB-MWF-ILD in red (×), and MWF-ITF in yellow (○).

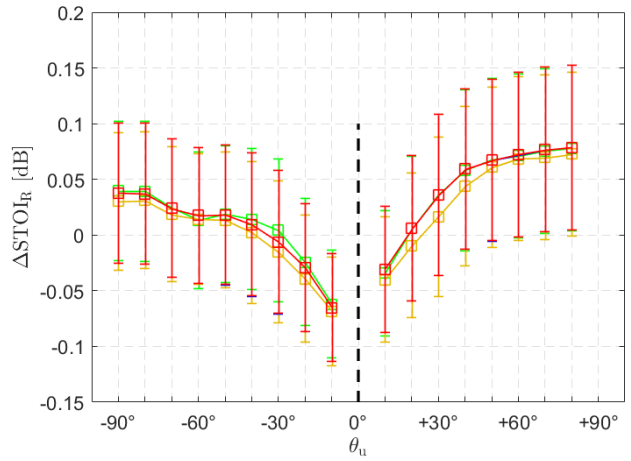

Fig. 12. $\Delta\text{STOI}_\text{R}$ calculated from 15 different input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◊), CB-MWF-ILD in red (×), and MWF-ITF in yellow (○).

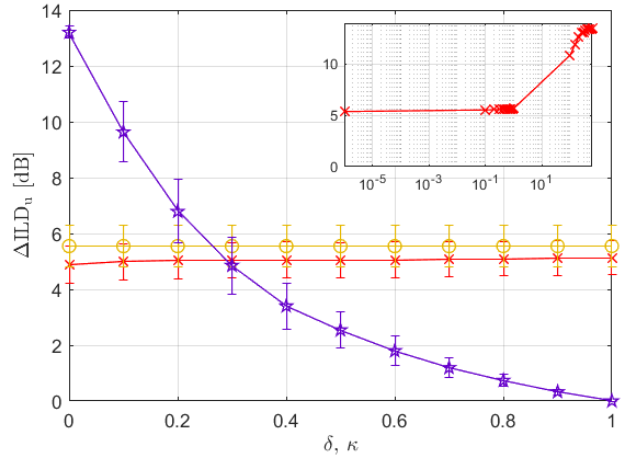

Fig. 13. Interference ΔILD for CB-MWF-ILD in red (×), MWF-ITF in yellow (○), and MWF-PNE in purple (☆). The abscissa (concomitantly) indicates parameters δ and κ, respectively required by the CB-MWF-ILD and MWF-PNE. $\theta_u = -60°$. The inset shows the CB-MWF-ILD performance for $10^{-6} \leq \delta \leq 6{\cdot}10^{2}$.

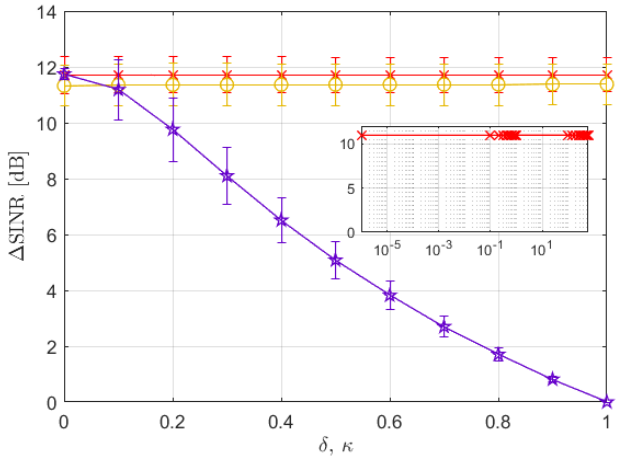

Fig. 14. ΔSINR for CB-MWF-ILD in red (×), MWF-ITF in yellow (○), and MWF-PNE in purple (☆). The abscissa (concomitantly) indicates parameters δ and κ, respectively required by the CB-MWF-ILD and MWF-PNE. $\theta_u = -60°$. The inset shows the CB-MWF-ILD performance for $10^{-6} \leq \delta \leq 6{\cdot}10^{2}$.

The CB-MWF-ILD was designed for acoustic scenarios with a single interference source. In the case of multiple sources, the rank-1 assumption is violated, and the performance of the semi-analytical closed-form solution is not guaranteed to be the same as the conventional MWF-ILD. However, note that the conventional forms of the MWF-ILD and MWF-ITF, presented in (21) and (88), are inappropriate for multiple interferences. In this case, other procedures are required to estimate the coherence matrices, and additional cost functions must be included. Methods such as the parametric unconstrained beamforming [28] or the relaxed binaural LCMV [47] can be alternatively used. It is also expected that an increase in the reverberation time would lead to a performance decrease, but keeping the correspondence with the conventional MWF-ILD. In such case, spatial preservation of the late reverberation would require control over the interaural coherence, and the ILD, ITD, and ITF cost functions are no longer adequate [48].

IX. Conclusion

This paper presented a new multichannel Wiener filter (MWF) based noise reduction method with interaural level difference (ILD) preservation. It minimizes the MWF cost function subject to two constraints (with physical meaning) for ILD preservation. Theoretical analysis shows that the resulting convex optimization problem leads to the same optimal solution as the conventional nonconvex MWF-ILD. However, differently from it, the proposed technique has a semi-analytical closed-form solution, which is a very interesting and desired characteristic for real-time applications. Simulation results indicate that the proposed method results in approximately the same signal-to-interference-plus-noise ratio performance presented by the conventional MWF and ILD preservation comparable to the conventional MWF-ILD. The proposed semi-analytical closed form leads to a considerable decrease (up to 13,000 times less) in the optimization time compared to the conventional MWF-

This article has been accepted for publication in IEEE/ACM Transactions on Audio, Speech and Language Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TASLP.2023.3291530

ILD method, which makes this method very attractive for binaural hearing aid applications.

APPENDIX A
PROOF OF THEOREM I

Decompose $\mathbf{\Phi}_y$ as

$$\mathbf{\Phi}_y = \mathbf{Q}_y \mathbf{\Lambda}_y \mathbf{Q}_y^H, \tag{104}$$

in which $\mathbf{\Lambda}_y$ is the diagonal matrix of eigenvalues ordered from the largest to the smallest; and $\mathbf{Q}_y$ is an orthonormal matrix with the corresponding eigenvectors. Assuming $\mathbf{\Phi}_y \succeq 0$ then (104) can be described as:

$$\mathbf{\Phi}_y = \begin{bmatrix}\mathbf{C}_y & \mathbf{N}_y\end{bmatrix} \mathrm{diag}(\mathbf{\Sigma}_y, \mathbf{0}_{(M-T)\times(M-T)}) \begin{bmatrix}\mathbf{C}_y & \mathbf{N}_y\end{bmatrix}^H, \tag{105}$$

in which $\mathbf{\Sigma}_y$ is a diagonal matrix of order $T \le M$, whose diagonal entries are the non-zero eigenvalues of $\mathbf{\Lambda}_y$; and matrices $\mathbf{C}_y$ (with dimension $M\times T$) and $\mathbf{N}_y$ (with dimension $M\times(M-T)$) are formed by the eigenvectors corresponding to the non-zero eigenvalues and the zero eigenvalues of $\mathbf{\Lambda}_y$, respectively. From (16), (17), and (19) it can be verified that vector $\mathbf{b}$ is in the column space of $\mathbf{\Phi}_y$. Using the eigenvectors of $\mathbf{\Phi}_y$ as a basis, vector $\mathbf{b}$ can be written as

$$\mathbf{b} = \begin{bmatrix}\mathbf{C}_y & \mathbf{N}_y\end{bmatrix} \begin{bmatrix}\mathbf{b}_y^T & \mathbf{0}_{1\times(M-T)}\end{bmatrix}^T, \tag{106}$$

in which $\mathbf{b}_y$ is the component of $\mathbf{b}$ in the column space of $\mathbf{\Phi}_y$. Using (105) and (106), (65) can be defined as:

$$\begin{aligned}&\mathbf{\Phi}_{yu}(\beta_r(\tau_j)) \\ &\quad = \mathbf{Q}_y \mathrm{diag}(\mathbf{\Sigma}_y + \beta_r(\tau_j) p_{su} \mathbf{b}_y \mathbf{b}_y^H, \mathbf{0}_{(M-T)\times(M-T)}) \mathbf{Q}_y^H.\end{aligned} \tag{107}$$

Because $\mathbf{Q}_y$ is invertible, pre- and post-multiplication of (107) by $\mathbf{Q}_y$ and its conjugate transpose preserves its positive definiteness characteristic [25]. Thus, the block diagonal structure of the inner matrix in (107) implies that $\mathbf{\Phi}_{yu}(\beta_r(\tau_j))$ is positive semidefinite if and only if

$$\mathbf{\Sigma}_y + p_{su}\beta_r(\tau_j)\mathbf{b}_y\mathbf{b}_y^H \succeq 0, \tag{108}$$

which can be expressed as:

$$\mathbf{\Sigma}_y^{1/2}(\mathbf{I} + \beta_r(\tau_j) p_{su} \mathbf{\Sigma}_y^{-1/2}\mathbf{b}_y\mathbf{b}_y^H\mathbf{\Sigma}_y^{-1/2})\mathbf{\Sigma}_y^{1/2} \succeq 0. \tag{109}$$

Because the columns of $\mathbf{C}_y$ in (106) consist of unity norm, orthogonal, and linearly independent vectors, the following relation between $\mathbf{b}_y$ and $\mathbf{b}$ can be established:

$$\mathbf{b}_y = \mathbf{C}_y^H \mathbf{b}. \tag{110}$$

Pre-multiplying both sides of (110) by $\mathbf{\Sigma}_y^{-1/2}$ leads to:

$$\mathbf{\Sigma}_y^{-1/2}\mathbf{b}_y = \mathbf{\Sigma}_y^{-1/2}\mathbf{C}_y^H\mathbf{b}. \tag{111}$$

Using (111) in (109), leads to:

$$\mathbf{\Sigma}_y^{1/2}(\mathbf{I} + \beta_r(\tau_j)\mathbf{\Sigma}_y^{-1/2}\mathbf{C}_y^H\mathbf{b}\mathbf{b}^H\mathbf{C}_y\mathbf{\Sigma}_y^{-1/2})\mathbf{\Sigma}_y^{1/2} \succeq 0 \tag{112}$$

The matrix inside the parenthesis in (112) is an identity matrix plus a rank-1 matrix. Thus, its eigenvalues are 1 ($T-1$ times) and $1+\beta_r(\tau_j)p_{su}\mathbf{b}_y{}^H\mathbf{C}_y\mathbf{\Sigma}_y{}^{-1}\mathbf{C}_y{}^H\mathbf{b}$ [49]. Therefore, since (112) consists in a congruence relation, $\mathbf{\Phi}_{yu}(\beta_r(\tau_j))$ is positive semidefinite if and only if:

$$\begin{aligned}0 &\le 1 + \beta_r(\tau_j) p_{su} \mathbf{b}^H \mathbf{C}_y \mathbf{\Sigma}_y^{-1} \mathbf{C}_y^H \mathbf{b} \\ &= 1 + \beta_r(\tau_j) p_{su} \mathbf{b}^H \mathbf{\Phi}_y^\dagger \mathbf{b} \\ &= 1 + \beta_r(\tau_j)\eta_b\end{aligned} \tag{113}$$

in which $\eta_b = p_{su}\mathbf{b}^H\mathbf{\Phi}_y{}^\dagger\mathbf{b} = \mathrm{trace}(\mathbf{\Phi}_y{}^\dagger\mathbf{\Phi}_u)$, and $\mathbf{\Phi}^\dagger{}_y$ is the Moore-Penrose inverse of matrix $\mathbf{\Phi}_y$, defined as $\mathbf{\Phi}^\dagger{}_y = \mathbf{Q}_y\mathrm{diag}(\mathbf{\Sigma}_y{}^{-1}, \mathbf{0}_{(M-t)\times(M-t)})\mathbf{Q}_y{}^H$. Considering $\eta_b > 0$, equation (113) results in

$$\beta_r(\tau_j) \ge -\eta_b^{-1}. \tag{114}$$

Replacing (66) in (114) completes the proof.

APPENDIX B
GRADIENT AND HESSIAN OF J_{MWF}, J_{ILD}, AND J_{ITF}

Since $\mathbf{\Phi}_{yy} = \mathbf{\Phi}_{yy}{}^H$, the J_{MWF} cost function defined in (11) can be written as:

$$J_{MWF}(\mathbf{w}) = \Re\{\mathbf{w}^H\mathbf{\Phi}_{yy}\mathbf{w}\} - 2\Re\{\mathbf{w}^H\mathbf{p}_{xx}\} + p_{xx}. \tag{115}$$

Defining

$$\breve{\mathbf{w}} = [\Re\{\mathbf{w}^T\} \quad \Im\{\mathbf{w}^T\}]^T = [\mathbf{w}_\Re^T \quad \mathbf{w}_\Im^T]^T \tag{116}$$

and using it in (115), leads to

$$J_{MWF}(\breve{\mathbf{w}}) = \breve{\mathbf{w}}^T\breve{\mathbf{\Phi}}_{yy}\breve{\mathbf{w}} - 2\breve{\mathbf{w}}^T\breve{\mathbf{p}}_{xx} + p_{xx}, \tag{117}$$

in which

$$\breve{\mathbf{\Phi}}_{yy} = \begin{bmatrix}\Re\{\mathbf{\Phi}_{yy}\} & -\Im\{\mathbf{\Phi}_{yy}\} \\ \Im\{\mathbf{\Phi}_{yy}\} & \Re\{\mathbf{\Phi}_{yy}\}\end{bmatrix} \text{ and } \breve{\mathbf{p}}_{xx} = \begin{bmatrix}\Re\{\mathbf{p}_{xx}\} \\ \Im\{\mathbf{p}_{xx}\}\end{bmatrix}. \tag{118}$$

Calculating the gradient of J_{MWF} with respect to $\breve{\mathbf{w}}$ leads to:

$$\nabla_{\breve{\mathbf{w}}} J_{MWF}(\breve{\mathbf{w}}) = 2\breve{\mathbf{\Phi}}_{yy}\breve{\mathbf{w}} - 2\breve{\mathbf{p}}_{xx}, \tag{119}$$

while its Hessian matrix is given by:

$$\nabla^2_{\breve{\mathbf{w}}} J_{MWF}(\breve{\mathbf{w}}) = 2\breve{\mathbf{\Phi}}_{yy}. \tag{120}$$

The gradient of J^u_{ILD} in (22) can also be defined as a function of $\breve{\mathbf{w}}$:

$$\nabla_{\breve{\mathbf{w}}} J^u_{ILD}(\breve{\mathbf{w}}) = \frac{40}{\log_e(10)} e_{ILD}(\breve{\mathbf{w}})\mathbf{p}(\breve{\mathbf{w}}), \tag{121}$$

in which

$$e_{ILD}(\breve{\mathbf{w}}) = 10\log_{10}(\mathrm{ILD}^u_{ou}(\breve{\mathbf{w}})) - 10\log_{10}(\mathrm{ILD}^u_{ou}), \tag{122}$$

$$\mathbf{p}(\breve{\mathbf{w}}) = \frac{\breve{\mathbf{\Phi}}_1\breve{\mathbf{w}}}{\breve{\mathbf{w}}^T\breve{\mathbf{\Phi}}_1\breve{\mathbf{w}}} - \frac{\breve{\mathbf{\Phi}}_2\breve{\mathbf{w}}}{\breve{\mathbf{w}}^T\breve{\mathbf{\Phi}}_2\breve{\mathbf{w}}}, \tag{123}$$

and

$$\breve{\mathbf{\Phi}}_1 = \mathrm{diag}(\Re\{\mathbf{\Phi}_u\}, \mathbf{0}_{M\times M}), \tag{124}$$

$$\breve{\mathbf{\Phi}}_2 = \mathrm{diag}(\mathbf{0}_{M\times M}, \Re\{\mathbf{\Phi}_u\}). \tag{125}$$

The Hessian matrix of the $J_{ILD}(\breve{\mathbf{w}})$ is given by

$$\begin{aligned}\nabla^2_{\breve{\mathbf{w}}} J^u_{ILD}(\breve{\mathbf{w}}) &= \frac{40}{\log_e(10)} e_{ILD}(\breve{\mathbf{w}})\nabla_{\breve{\mathbf{w}}}\mathbf{p}(\breve{\mathbf{w}}) \\ &\quad - \frac{40}{\log_e(10)}\nabla_{\breve{\mathbf{w}}} e_{ILD}(\breve{\mathbf{w}})\mathbf{p}^T(\breve{\mathbf{w}})\end{aligned} \tag{126}$$

The cost function in (88) can be written with respect to $\breve{\mathbf{w}}$ leading to:

$$J^{\mathrm{u}}_{\mathrm{ITF}}(\tilde{\mathbf{w}}) = \frac{\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_3\tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_4\tilde{\mathbf{w}}}, \tag{127}$$

in which:

$$\tilde{\mathbf{\Phi}}_3 = \begin{bmatrix} \Re\{\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\} & -\Im\{\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\} \\ \Im\{\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\} & \Re\{\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\} \end{bmatrix}, \tag{128}$$

$$\tilde{\mathbf{\Phi}}_4 = \begin{bmatrix} \mathbf{0}_{M\times M} & \mathbf{0}_{M\times M} \\ \mathbf{0}_{M\times M} & \Re\{\tilde{\mathbf{\Phi}}_{\mathrm{u2}}\} \end{bmatrix}. \tag{129}$$

The gradient of (127) is given by:

$$\nabla_{\tilde{\mathbf{w}}} J^{\mathrm{u}}_{\mathrm{ITF}}(\tilde{\mathbf{w}}) = 2\frac{\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\tilde{\mathbf{w}}\cdot\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u2}}\tilde{\mathbf{w}} - \tilde{\mathbf{\Phi}}_{\mathrm{u2}}\tilde{\mathbf{w}}\cdot\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\tilde{\mathbf{w}}}{(\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u2}}\tilde{\mathbf{w}})^2} \tag{130}$$

Also, the Hessian matrix of (127) is given by:

$$\begin{aligned}\nabla^2_{\tilde{\mathbf{w}}} J^{\mathrm{u}}_{\mathrm{ITF}}(\tilde{\mathbf{w}}) = {} & 4\frac{\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\cdot\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\tilde{\mathbf{w}} - \tilde{\mathbf{\Phi}}_{\mathrm{u1}}\tilde{\mathbf{w}}\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u1}}}{(\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u2}}\tilde{\mathbf{w}})^2} \\ & - 4\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\frac{\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\tilde{\mathbf{w}}}{(\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u2}}\tilde{\mathbf{w}})^2} - 2\frac{\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\tilde{\mathbf{w}}\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u2}}}{(\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u2}}\tilde{\mathbf{w}})^2} \\ & + 4\frac{\tilde{\mathbf{\Phi}}_{\mathrm{u2}}\tilde{\mathbf{w}}\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u2}}}{(\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u2}}\tilde{\mathbf{w}})^3}\cdot\tilde{\mathbf{w}}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_{\mathrm{u1}}\tilde{\mathbf{w}}\end{aligned} \tag{131}$$

APPENDIX C

ITERATIVE ALGORITHM FOR OBTAINING OPTIMUM α/ρ AND $\mathbf{w}$ FOR THE UNCONSTRAINED MWF-ILD OR MWF-ITF.

Algorithm: Iterative search for finding α and $\mathbf{w}$ in each discrete frequency $k = 1, \ldots, N$.

Input: δ^2, $\mathrm{ILD}_{\mathrm{in}}$, $\mathrm{ITF}_{\mathrm{in}}$, $\boldsymbol{\alpha}$, $\mathbf{\Phi}_{\mathrm{u}}$, $\mathbf{\Phi}_{\mathrm{y}}$, $\mathbf{\Phi}_{\mathrm{x}}$

Output: $\mathbf{w}$

Choose $J(\lambda,k) \in \{ J_{\mathrm{ILD}}(\lambda,k), J_{\mathrm{ITF}}(\lambda,k) \}$
Set G as the length vector $\boldsymbol{\alpha}$
Set $e^2_{\mathrm{ILD}} = \infty$
Set $i = 0$
Set $\mathbf{w}$ with random initialization
While $e^2_{\mathrm{ILD}} > \delta^2$ **and** $i < G$
  $i = i + 1$
  Considering $\alpha_i = \boldsymbol{\alpha}(i)$, find the binaural filters solving:
  $\mathbf{w}(\alpha_i) = \arg\min_{\mathbf{w}} J_{\mathrm{MWF}}(\lambda,k) + \alpha_i J(\lambda,k)$
  $\mathbf{w}_{\mathrm{UNC}} = \mathbf{w}(\alpha_i)$
  $e^2_{\mathrm{ILD}} = J^{\mathrm{u}}_{\mathrm{ILD}}(\mathbf{w}(\alpha_i))$ using (22)
End While
Return $\mathbf{w} = \mathbf{w}_{\mathrm{UNC}}$

REFERENCES

[1] S. A. Gelfand, *Hearing: An introduction to psychological and physiological acoustics*, CRC Press, 2018.

[2] J. Blauert, *Spatial hearing – The psychoacoustics of human sound localization*, MIT Press 1997.

[3] T. Y. C. Ching, E. Wanrooy, H. Dillon, and L. Carter, "Spatial release from masking in normal-hearing children and children who use hearing aids," *J. Acoust. Soc. Am.*, vol. 129, no. 1, pp. 368-375, 2011.

[4] W. A. Yost, "Spatial release from masking based on binaural processing for up to six maskers," *J. Acoust. Soc. Am.*, vol. 141, no. 3, pp. 2093-2106, 2017.

[5] L. Wagner, L. Geiling, C. Hauth, T. Hocke, S. Plonke, and T. Rahne, "Improved binaural speech reception thresholds through small symmetrical separation of speech and noise," *PloS One*, vol. 15, no. 8, pp. 1-10, 2020.

[6] N. Li and P. C. Loizou "Factors influencing intelligibility of ideal binary-masked speech: Implications for noise reduction," *J. Acoust. Soc. Am.*, vol. 123, no. 3, pp. 1673-1682, 2008.

[7] Y. Hu and P. C. Loizou, "Subjective comparison and evaluation of speech enhancement algorithms," *Speech Commun.*, vol. 49, no. 7-8, pp. 588-601, 2007.

[8] J. Benesty, J. Chen, Y. Huang, and I. Cohen, *Noise reduction in speech processing*, Springer, 2009.

[9] S. Doclo, S. Gannot, D. Marquart, and E. Hadad, "Binaural speech processing with application to hearing devices," in *Audio source separation and speech enhancement*, John Wiley & Sons, 2018. ch. 18, pp. 413-442.

[10] W. M. Hartmann and Z. A. Constan, "Interaural level differences and the level-meter model," *J. Acoust. Soc. Am.*, vol. 112, no. 3, pp. 1037-1045, 2002.

[11] K. Blum, G.-J. Rooyen, and H. A. Engelbrecht, "Spatial audio to assist speaker identification in telephony," in *Proc. IWSSIP*, 2010, pp. 1-4.

[12] B. Cornelis, S. Doclo, T. Bogaert, M. Moonen, and J. Wouters, "Theoretical analysis of binaural multimicrophone noise reduction techniques," *IEEE Trans. Audio Speech Lang. Process.*, vol. 18, no. 2, pp. 342-355, 2010.

[13] M. L. Hawley, R. Y. Litovsky, and J. F. Culling, "The benefit of binaural hearing in a cocktail party: effect of location and type of interferer," *J. Acoust. Soc. Am.*, vol. 15, no 2, pp. 833-843, 2004.

[14] T. Bogaert, S. Doclo, J. Wouters, and M. Moonen, "The effect of multimicrophone noise reduction systems on sound source localization by users of binaural hearing aids," *J. Acoustic. Soc. Am.*, vol. 124, pp. 484-497, 2008.

[15] T. Bogaert, J. Wouters, S. Doclo, and M. Moonen, "Binaural cue preservation for hearing aids using an interaural transfer function multichannel Wiener filter," in *Proc. ICASSP*, 2007, pp. 565-568.

[16] S. Doclo, R. Dong, T. J. Klasen, J. Wouters, S. Haykin, and M. Moonen, "Extension of the multi-channel Wiener filter with localization cues for noise reduction in binaural hearing aids," in *Proc. IWAENC*, 2005, pp. 221-224.

[17] M. H. Costa and P. A. Naylor, "ILD preservation in the multichannel Wiener filter for binaural hearing aid applications," in *Proc. EUSIPCO*, 2014, pp. 636-640.

[18] D. M. Carmo and M. H. Costa, "Online approximation of the multichannel Wiener filter with preservation of interaural level difference for binaural hearing aids," *Comput. Biol. Med.*, vol. 95, pp. 188-197, 2018.

[19] E. Chaumette, J. Vila-Valls, and F. Vincent, "On the general conditions of existence for linear MMSE filters: Wiener and Kalman," *Signal Process.*, vol. 184, pp. 1-5, 2021.

[20] M. Raspaud, H. Viste, and G. Evangelista, "Binaural source localization by joint estimation of ILD and ITD," *IEEE Trans. Audio Speech Lang. Process.*, vol. 18, no. 1, pp. 68-77, 2009.

[21] J. Werner and M. H. Costa, "A Noise-reduction method with coherence enhancement for binaural hearing aids," *J. Commun. Inf. Syst.*, vol. 35, no. 1, pp. 338-348, 2020.

[22] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.

[23] A. Beck and Y. C. Eldar, "Strong duality in nonconvex quadratic optimization with two quadratic constraints," *SIAM J. Optimiz.*, vol. 17, no. 3, pp. 844-860, 2006.

[24] X. Lin, Y. Huang, and W.-K. Ma, "Robust downlink transmit optimization under quantized channel feedback via the strong duality for QCQP," *IEEE Signal Proc. Let.*, vol. 28, pp. 1-5, 2021.

[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 2012.

[26] N. Castro-González, M. F. Martínez-Serrano, and J. Robles, "Expressions for the Moore–Penrose inverse of block matrices involving the Schur complement," *Linear Algebra Appl.*, vol. 471, pp. 353-368, 2015.

[27] K. Kohno, Y. Inouye, and M, Kawamoto, "A matrix pseudo-inver-

sion lemma for positive semidefinite hermitian matrices and its application to adaptive blind deconvolution of MIMO systems," *IEEE Trans. Circ. Sys. I*, vol. 55, no. 1, pp. 424-435, 2008.

[28] J. Zhang and G. Zhang, "A parametric unconstrained beamformer based binaural noise reduction for assistive hearing," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 30, pp. 292-304, 2021.

[29] H. Kayser, S. D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Database of multichannel in-ear and behind-the-ear head related and binaural room impulse responses," *EURASIP J. Adv. Sig Pr.*, vol. 2009, pp. 1-10, 2009.

[30] G. E. Henter, T. Merritt, M. Shannon, C. Mayo, and S. King, "Repeated Harvard sentence prompts corpus version 0.5" [dataset]. The University of Edinburgh, Centre for Speech Technology Research, Cambridge University Engineering Department, 2014. Available at https://doi.org/10.7488/ds/39.

[31] E. A. P. Habets, I. Cohen, and S. Gannot, "Generating nonstationary multisensor signals under a spatial coherence constraint," *J. Acoust. Soc. Am.*, vol. 124, no. 5, pp. 2911-2917, 2008.

[32] D. Marquardt, V. Hohmann, and S. Doclo, "Coherence preservation in multi-channel Wiener filtering based noise reduction for binaural hearing aids," in *Proc. ICASSP*, 2013, pp. 8648-8652.

[33] B. Cornelis, M. Moonen, and J. Wouters, "Performance analysis of multichannel Wiener filter-based noise reduction in hearing aids under second order statistics estimation errors," *IEEE Trans. Audio Speech Lang. Process.*, vol. 19, no. 5, pp. 1368-1381, 2011.

[34] J. Zhang and C. Li, "Quantization-aware binaural MWF based noise reduction incorporating external wireless devices," *IEEE/ACM Trans. on Audio Speech Lang. Process.*, vol. 29, pp. 3118-3131, 2021.

[35] E. Hadad, S. Doclo, and S. Gannot, "The binaural LCMV beamformer and its performance analysis," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 24, no. 3, pp. 543-558, 2016.

[36] N. Gößling, E. Hadad, S. Gannot, and S. Doclo, "Binaural LCMV beamforming with partial noise estimation," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 28, pp. 2942-2955, 2020.

[37] N. Gößling, D. Marquardt, I. Merks, T. Zhang, and S. Doclo, "Optimal binaural LCMV beamforming in complex acoustic scenarios: Theoretical and practical insights," in *Proc. IWAENC*, 2018, pp. 381-385.

[38] P. Pertila, E. Fagerlund, A. Huttunen, and V. Myllyla, "Online own voice detection for a multi-channel multi-sensor in-ear device," *IEEE Sens. J.*, vol. 21, no. 24, pp. 27686-27697.

[39] S. Markovich, S. Gannot, and I. Cohen, "Multichannel eigenspace beamforming in a reverberant environment with multiple interfering speech signals," *IEEE Trans. Audio Speech Lang. Process.*, vol. 17, no. 6, pp. 1071-1086, 2009.

[40] R. Serizel, M. Moonen, B. Van Dijk, and J. Wouters, "Low-rank approximation based multichannel wiener filter algorithms for noise reduction with application in cochlear implants," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 22, no. 4, pp. 785-799, 2014.

[41] S. Markovich-Golan, S. Ganot, and W. Kellerman, "Performance comparison of the covariance-whitening and the covariance-subtraction methods for estimating the relative transfer function," in *Proc. EUSIPCO*, 2018, pp. 2513-2517.

[42] Mathworks, "fmincon - find a minimum of a constrained nonlinear multivariable function." Available at https://www.mathworks.com /help/optim/ug/fmincon.html.

[43] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra, "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs," in *Proc. ICASSP*, 2001, pp. 749-752.

[44] C. H. Taal, R. C. Hendriks, R. Heusdens, and J. Jensen, "A short-time objective intelligibility measure for time-frequency weighted noisy speech," in *Proc. ICASSP*, 2010, pp. 4214-4217.

[45] A. Servetti, and J. C. D. Martin, "802.11 MAC protocol with selective error detection for speech transmission," in Quality of Service in Multiservice IP Networks. Springer, 2005, pp. 509-519.

[46] P. C. Loizou, Speech Enhancement: Theory and Practice. CRC Press, 2013.

[47] A. I. Koutrouvelis, R. C. Hendriks, R. Heusdens, and J. Jensen, "Relaxed binaural LCMV beamforming," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 25, no. 1, pp. 137-152, 2017.

[48] J. Werner and M. H. Costa, "Improved spatialization performance for joint speech dereverberation and noise reduction in binaural hearing aids," *Biomed. Signal Process. Control*, vol. 68, pp. 1-15, 2021.

[49] J. Ding, and A. Zhou, "Eigenvalues of rank-one updated matrices with some applications," *Appl. Math. Lett.*, vol. 20, no. 12, pp. 1223-1226, 2007.