UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Diego Marques do Carmo

Binaural Multichannel Wiener Filter Methods for Noise Reduction and Preservation of Binaural Sound Impression

Florianópolis
2023

Diego Marques do Carmo

Binaural Multichannel Wiener Filter Methods for Noise Reduction and Preservation of Binaural Sound Impression

Tese submetida ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina para a obtenção do título de doutor em Engenharia Elétrica.
Orientador: Prof. Márcio Holsbach Costa, Dr.

Florianópolis
2023

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Carmo, Diego Marques do
Binaural Multichannel Wiener Filter Methods for Noise Reduction and Preservation of Binaural Sound Impression / Diego Marques do Carmo ; orientador, Márcio Holsbach Costa, 2023.
119 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Elétrica, Florianópolis, 2023.

Inclui referências.

1. Engenharia Elétrica. 2. Aparelhos auditivos biauriculares. 3. Informações de localização biauriculares . 4. Filtro multicanal de Wiener. 5. Redução de ruído. I. Costa, Márcio Holsbach. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Elétrica. III. Título.

Diego Marques do Carmo

Binaural Multichannel Wiener Filter Methods for Noise Reduction and Preservation of Binaural Sound Impression

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Vitor Heloiz Nascimento, Ph.D.
Universidade de São Paulo (USP)

Prof. Sérgio José de Melo Almeida, Dr.
Universidade Católica de Pelotas (UCPel)

Prof. Bruno Catarino Bispo, Ph.D.
Universidade Federal de Santa Catarina (UFSC)

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de doutor em Engenharia Elétrica.

Documento assinado digitalmente
Telles Brunelli Lazzarin
Data: 13/07/2023 23:02:32-0300
CPF: ***.400.719-**
Verifique as assinaturas em https://v.ufsc.br

Prof. Telles Brunelli Lazzarin, Dr.
Coordenador do Programa

Documento assinado digitalmente
Marcio Holsbach Costa
Data: 13/07/2023 11:09:19-0300
CPF: ***.170.680-**
Verifique as assinaturas em https://v.ufsc.br

Prof. Márcio Holsbach Costa, Dr.
Orientador

Florianópolis, 31 de maio de 2023.

Este trabalho é dedicado a Helino e Roseth, meus pais.

ACKNOWLEDGEMENTS

A tese de doutorado é a materialização de vários anos de pesquisa e apresenta uma ideia original e inédita do autor, mas ela é incompleta. Ela é incompleta não porque o leitor (entendido do tema) não conseguirá compreender a ideias contidas entre a primeira e a última página, mas, sim, porque a história da produção deste texto existe para além desses marcos de início e fim. O texto técnico foi gerado a partir de vários outros textos que foram publicadas em congressos e revistas, e de textos que foram melhorados e outros que foram abandonados. As ideias que vingaram e que permanecerão registradas para sempre são filhas de outras que duraram alguns minutos e outras que duraram alguns meses antes de morrerem, deixando apenas a inspiração. A autoria do texto é atribuída uma pessoa, porém várias outras contribuíram de diferentes maneiras para que o texto se tornasse realidade. É para resolver esse problema da incompletude de qualquer obra que existe os agradecimentos. Abaixo, agradeço a todos aqueles que contribuíram para que a história dessa tese fosse construída.

Roseth e Helino, meus pais, me deram as condições básicas para que eu pudesse sonhar o caminho que eu desejasse. Sem o apoio incondicional deles a minha capacidade de sonhar seria limitada por tantas dificuldades que aparecem no caminho de uma pessoa pobre do norte do Brasil. Por isso, a eles todo meu amor é infinito agradecimento por serem meu alicerce.

Se as condições básicas foram fornecidas pelos meus pais, as condições técnicas para produção desta tese só foram possíveis graças ao Professor Márcio Costa, meu orientador durante nesses anos de UFSC. Márcio merece meus agradecimentos por me transmitir sua maneira de conduzir e pensar uma pesquisa científica, mas também por me aconselhar e me dar apoio durante vários momentos difíceis desses anos de UFSC. Muito obrigado, Professor.

Agradeço também ao Ricardo Borsoi, meu amigo, uma pessoa generosa que é coautor nos trabalhos que compõem esse texto. Alguém com o qual eu compartilhei vários dias e noites de laboratório com ótimos papos de temas diversos.

As pessoas mais próximas sofrem e se alegram conosco durante todo o trajeto. Por isso, eu agradeço a Maria Negri, minha namorada, por estar comigo durante esse caminho, me dando suporte em momentos difíceis e mais ainda por ser minha fonte de amor, diversão e alegria.

São com nossos amigos e colegas de laboratório que compartilhamos nossas angústias quanto ao andamento do doutorado, mas é com eles que dividimos momentos prazerosos de bate-papo com café. Por isso, agradeço a voĉes esse momentos que lembrarei para sempre:Celso, Karoline, Luciana Menezes, Paulo Goubert, Rivael e Vitor.

A vida de pós-graduação envolve lidar com preenchimento de papéis e cumprimento de protocolos e prazo. E os servidores do PPGEEL facilitam todo esse processo. Por isso,

gostaria de agradecer ao Marcelo, Wilson e Lis pelo seu trabalho sempre dedicado e preocupado com os alunos da pós-graduação.

Por fim, o financiamento de ideias é o que possibilita a concretude de trabalhos como este, por isso agradeço à Capes pelo financiamento desta tese.

"You have to make some noise if you want to be heard" (Enola Holmes)

ABSTRACT

Binaural hearing aids are the most advanced hearing compensation device for people with mild to moderate hearing loss in both ears. It consists of a pair of hearing aids with a communication channel that is used to share signals and control parameters. Due to this characteristic, these devices can use binaural noise reduction techniques that jointly increase the power of the emitted signals, which improves the quality and intelligibility of speech; and preserve information associated with the perception of acoustic sources, which allows the correct spatial location of these sound sources. The binaural multichannel Wiener filter (MWF) is one of the most studied approaches for binaural noise reduction. The MWF design is based on minimization of the mean squared error cost function and generates filters that reduce additive noise. To maintain the correct spatial perception of the location of acoustic sound sources, terms that penalize solutions that distort the spatial perception of sound sources are added to the cost function that characterize the MWF. In this context, this thesis presents three contributions to improve different aspects of techniques based on the MWF. The first contribution is a method for the automatic adjustment of techniques based on the MWF cost function, which aims to maintain the desired noise reduction performance and preserve the spatial perception of sound sources in situations where there is a variation in the power of the signals captured by the hearing aid microphones. The results obtained from simulations of acoustic scenarios and psychoacoustic experiments carried out with human beings indicate that the proposed method maintained the desired trade-off between noise reduction and perception of the spatial position of sound sources. Competing techniques, for example, cannot maintain the performance setup in situations where there is variation in the power of the input signals, a problem that does not occur in the proposed technique. The second contribution is a closed-form solution for the MWF with preservation of the interaural level difference (ILD). Simulations were performed to compare the proposed closed-form solution with the traditionally used MWF technique with preservation of ILD, which uses numerical optimization methods. The results show that the proposed technique considerably reduces the calculation time of the noise reduction filters while maintaining the same noise reduction performance and preservation of spatial location cues compared to the technique traditionally used. Finally, the third contribution is a set of three techniques that can be used to preserve binaural information from point sources and to preserve this information in diffuse sources. The first technique is a new cost function used in unconstrained optimization, the second technique is based on optimization with quadratic constraints, and the third technique is based on semidefinite optimization. The results obtained show that the new cost function and the technique based on optimization with quadratic constraints preserve both the perception of point sound sources and that of diffuse sound sources, a characteristic that is not presented by any state-of-the-art technique based on MWF. The third technique reduces the time required to calculate noise reduction filters. In addition, it preserves binaural information from point sources, distorts the perception of diffuse sources, and has a noise reduction performance equivalent to that presented by the MWF.

Keywords: binaural hearing aids, noise reduction, binaural cues, Multichannel Wiener Filter.

RESUMO

Aparelhos auditivos biauriculares são os aparelhos de compensação auditiva mais avançados para pessoas com perdas auditivas leve à moderada em ambas as orelhas. O dispositivo consiste em um par de aparelhos auditivos com um canal de comunicação usado para compartilhar sinais e parâmetros de controle. Devido a essa característica, esses dispositivos podem empregar técnicas de redução biauricular de ruído que, conjuntamente, reduzem a potência de sinais indesejados, o que melhora a qualidade e inteligibilidade da fala; e preservam informações associadas à percepção de fontes acústicas, o que permite a correta localização espacial dessas fontes sonoras. O filtro biauricular multicanal de Wiener (*multichanel Wiener filter* – MWF) é uma das abordagens mais estudadas para redução biauricular de ruído. O projeto do MWF é feito a partir da minimização da função custo do erro quadrático médio e gera filtros que reduzem o ruído aditivo. Para se manter a correta percepção espacial da localização de fontes sonoras acústicas, termos que penalizam soluções que distorcem a percepção espacial das fontes sonoras são adicionados à função custo que caracteriza o MWF. Neste contexto, esta tese apresenta três contribuições para melhorar diferentes aspectos das técnicas baseadas no MWF. A primeira contribuição é um método para o ajuste automático de técnicas biauriculares baseadas no MWF que visa a manter o desempenho de redução de ruído desejado e preservar a percepção espacial de fontes sonoras em situações em que há variação de potência dos sinais captados pelos microfones dos aparelhos auditivos. Resultados obtidos a partir de simulações de cenários acústicos e de experimentos psicoacústicos realizados com seres humanos indicam que o método proposto manteve o compromisso desejado entre redução de ruído e percepção da posição espacial de fontes sonoras. As técnicas concorrentes, por exemplo, não conseguem manter o desempenho projetado em situações em que há variação de potência dos sinais captados, problema que não ocorre na técnica proposta. A segunda contribuição é uma solução de forma fechada para o MWF com preservação da diferença de nível interauricular (*interaural level difference* – ILD). Simulações foram realizadas para comparar a proposta de solução de forma fechada com a técnica tradicionalmente utilizada do MWF com preservação de ILD, que utiliza métodos números de otimização. Os resultados mostram que a técnica proposta reduz consideravelmente o tempo de cálculo dos filtros ótimos de redução de ruído enquanto mantém o mesmo desempenho de redução de ruído e preservação das pistas de localização espacial comparado à técnica tradicionalmente utilizada. Por fim, a terceira contribuição é um conjunto de três técnicas que podem ser utilizadas tanto para preservação das informações biauriculares de fontes pontuais quanto para preservação dessas informações em fontes difusas. A primeira técnica é uma nova função custo utilizada em otimização sem restrição, a segunda técnica é baseada em otimização com restrições quadráticas, e a terceira técnica é baseada em otimização semi-definida. Os resultados obtidos mostram que a nova função de custo e a técnica baseada em otimização com restrições quadráticas preservam tanto a percepção de fontes sonoras pontuais quanto a de fontes sonoras difusas, característica que ultimamente nenhuma técnica baseada em MWF apresenta. A terceira técnica apresenta um tempo menor para o cálculo dos filtros de redução de ruído. Além disso, ela preserva as informações biauriculares de fontes pontuais, distorce a percepção de fontes difusas, e possui um desempenho de redução de ruído equivalente ao apresentado pelo MWF.

Palavras-chave: Aparelhos auditivos biauriculares, informações biauriculares, filtro multicanal de Wiener.

RESUMO ESTENDIDO

Aparelhos auditivos biauriculares são os dispositivos mais avançados para a compensação de perdas auditivas leves à moderadas. Consistem em um par de aparelhos auditivos, posicionados um em cada orelha, que se comunicam através de um canal de comunicação sem fio, de forma a compartilhar sinais e parâmetros de controle. Devido a essa característica, esses dispositivos podem empregar estratégias de redução de ruído para, conjuntamente, reduzir a potência de sinais indesejados (o que pode melhorar a qualidade e a inteligibilidade da fala) e preservar informações associadas às pistas acústicas espaciais (o que permite manter a percepção espacial do cenário acústico). O filtro biauricular multicanal de Wiener (multichanel Wiener filter – MWF) é uma das abordagens mais estudadas para redução biauricular de ruído. O projeto do MWF é feito a partir da minimização do erro quadrático médio entre o sinal desejado e o sinal contaminado processado. No caso biauricular, termos que penalizam soluções que distorcem as pistas biauriculares são adicionados à função custo original. Neste contexto, esta tese apresenta três contribuições para melhorar diferentes aspectos de técnicas baseadas no MWF.

A primeira contribuição é um método para projetar o parâmetro de ponderação da função custo aumentada de métodos de redução de ruído baseados no MWF biauricular. O parâmetro em questão estabelece o compromisso desejado entre redução de ruído e preservação das pistas biauriculares. A estratégia proposta foi derivada especialmente para a preservação da diferença interauricular de nível (interaural level difference – ILD), da diferença interauricular de fase (interaural time difference – IPD) e da coerência interauricular (interaural coherence – IC). A estratégia proposta é definida em função da potência média do ruído de entrada nos microfones, conferindo robustez contra a influência de mudanças conjuntas da potência do ruído e da fala (efeito Lombard), bem como variações da razão sinal-ruído. Um arcabouço teórico baseado na definição matemática do grau de homogeneidade é apresentado e aplicado a uma função custo genérica aumentada baseada no MWF. Os insights teóricos obtidos são suportados por simulações computacionais e experimentos psicoacústicos usando o MWF com preservação da função de transferência interauricular do ruído (MWF-ITF), como um estudo de caso. A análise estatística indica que, em comparação com o uso de um parâmetro de ponderação fixo (usualmente utilizado na literatura), a estrutura e o método de projeto propostos oferecem robustez significativa contra alterações nas pistas biauriculares do ruído, ao custo de uma pequena diminuição do desempenho de redução de ruído e uma modesta distorção das pistas biauriculares da fala.

A segunda contribuição é um método de redução de ruído baseado no MWF com preservação da ILD. A proposta minimiza a função custo do MWF sujeita a duas restrições quadráticas para preservação da ILD. Nesta abordagem, o coeficiente de ponderação que estabelece o compromisso entre redução de ruído e preservação da pista biauricular apre-

senta uma interpretação física, o que facilita o seu projeto. A abordagem proposta resulta em um problema de otimização convexa que admite uma solução semi-analítica de forma fechada computacionalmente eficiente. Experimentos de simulação em aplicações com aparelhos auditivos foram realizados considerando cenários acústicos quotidianos. Para um conjunto adequado de parâmetros de controle, o desempenho médio do método proposto preserva a ILD da fonte interferente da mesma forma que o MWF-ILD convencional e mantém o mesmo nível de redução de ruído do MWF clássico, com aproximadamente a mesma quantidade de distorção da ILD da fala. O método proposto é particularmente interessante para implementações computacionalmente eficientes de métodos de redução de ruído em tempo real em aparelhos auditivos biauriculares.

Por fim, a terceira contribuição é um conjunto de três técnicas que podem ser utilizadas tanto para preservação das informações biauriculares de fontes pontuais quanto de campos sonoros difusos. A primeira técnica é uma nova função custo composta por um termo para preservação da ILD e outro para IC. Inicialmente, o projeto dos filtros de redução de ruído é feito a partir de um processo de otimização sem restrição. Apesar dos resultados bem sucedidos da técnica proposta na redução de ruído e preservação de pistas espaciais, uma importante limitação é o tempo necessário para o cálculo dos filtros. A segunda técnica proposta reformula o problema original do MWF com preservação de ILD e IC e coloca os termos da nova função custo como restrições de um problema de otimização, que são posteriormente transformadas em funções quadráticas dos coeficientes do filtro. Como consequência da reformulação realizada, o controle dos erros nas pistas biauriculares pode ser feito de forma mais eficiente, o que diminui o tempo de cálculo. A terceira técnica é uma reformulação da técnica com restrições quadráticas que utiliza uma restrição semi-definida e foi desenvolvida para reduzir ainda mais o tempo de cálculo dos coeficientes. A avaliação das técnicas propostas foi realizada a partir da simulação de cenários acústicos relevantes para a aplicação em questão. Os resultados indicam que as técnicas baseadas no processo de otimização sem restrição e a com restrições quadráticas conseguem preservar as pistas biauriculares de fontes pontuais e de campos sonoros difusos, resultado inexistente no estado da arte. Por outro lado, a técnica que utiliza otimização com restrição semi-definida, embora diminua substancialmente o tempo de cálculo dos coeficientes e resulte em níveis de redução de ruído compatíveis com o MWF, não preserva efetivamente as pistas espaciais de campos acústicos difusos.

Palavras-chave: Aparelhos auditivos biauriculares, informações biauriculares, filtro multicanal de Wiener.

LIST OF FIGURES

LIST OF SYMBOLS

Sets

$\mathbb{C}$ Set of complex numbers

$\mathbb{R}$ Set of real numbers

$\mathbb{R}_+$ Set of non-negative real numbers

Scalars

α Parameter that weights the cost function used for preservation of binaural cues

ΔILD_d Variation between input and output ILD measured from a generic signal d

ΔIPD_d Variation between input and output IPD measured from a generic signal d

ΔMSC_d Variation between input and output MSC measured from a generic signal d

$\Delta\text{SNR}_{\ell,\text{in}}$ Variation between the overall bin input and output SNR in the reference microphone of hearing aid ℓ

γ Constant that defines the setpoint associated with the desired trade-off between noise reduction and binaural cue preservation

λ Discrete time index

L Left side label

R Right side label

$\overline{\text{SNR}_{\text{in}}}$ overall microphone input SNR

$\overline{g}$ Average Lombard gain

τ Elevation angle

$\text{IC}_{\text{in}}^{\text{d}}$ input IC of the signal d

$\text{IC}_{\text{ou}}^{\text{d}}$ output IC of the signal d

$\text{ILD}_{\text{in}}^{\text{d}}$ input ILD of the signal d

$\text{ILD}_{\text{ou}}^{\text{d}}$ output ILD of the signal d

$\text{IPD}_{\text{in}}^{\text{d}}$ input IPD of the signal d

$\text{IPD}_{\text{ou}}^{\text{d}}$ output IPD of the signal d

$\text{ITF}_{\text{in}}^{\text{d}}$ input ITF of the signal d

$\mathrm{ITF}_{\mathrm{ou}}^{\mathrm{d}}$ output ITF of the signal d

$\mathrm{MSC}_{\mathrm{in}}^{\mathrm{d}}$ input MSC of the signal d

$\mathrm{MSC}_{\mathrm{ou}}^{\mathrm{d}}$ output MSC of the signal d

$\mathrm{SNR}_{\ell,\mathrm{in}}$ input SNR in the reference microphone of hearing aid ℓ

$\mathrm{SNR}_{\ell,\mathrm{ou}}$ output SNR in the reference microphone of hearing aid ℓ

$\mathrm{SNR}_{\mathrm{WORST}}(k)$ The worst predicted input SNR condition

θ Azimuth angle

$a_{\ell,m}$ Acoustic transfer function from the desired source to the m-th microphone of side ℓ

$b_{\ell,m}$ Acoustic transfer function from the interfering source to the m-th microphone of side ℓ

d Generic received signal, i.e., $\{x, n, v, u\}$

d_ℓ Generic signal at hearing aid of side ℓ

f_s Sampling frequency

g Lombard gain

$i\mathrm{SNR}_{\ell,\mathrm{in}}$ Overall bin input SNR in the reference microphone of hearing aid ℓ

J_{T} Total cost function comprised of different terms

k Discrete frequency index

M Number of microphones in a single hearing aid

m Microphone index

M_ℓ Number of microphones in the hearing aid of side ℓ

$o\mathrm{SNR}_{\ell,\mathrm{in}}$ Overall bin output SNR in the reference microphone of hearing aid ℓ

$p_{d\ell}$ Generic processed set of samples at hearing aid of side ℓ

r Distance between a source and the center of binaural coordinate system

s_u Interfering signal generated by the source

s_x Speech signal generated by the source

$u_{\ell,m}$ Interfering signal in the microphone m of side ℓ

$v_{\ell,m}$ Received total noise

$x_{\ell,m}$ Received speech signal

$x_{\ell,m}$ Speech signal in the microphone m of side ℓ

$y_{\ell,m}$ Noise in the microphone m of side ℓ

$y_{\ell,m}$ Noisy signal in the microphone m of side ℓ

$y_{\ell,m}$ Noisy signal

$y_{\ell,m}$ Overall noise in the microphone m of side ℓ

z_{ℓ} Processed signal at side ℓ

Matrices

$\mathbf{\Phi}_{(\bullet)}$ Coherence matrix

Operators

$(\bullet)^{-1}$ Inverse element

$\angle\big(\bullet\big)$ Unwrapped phase of a complex number

$\mathbb{E}\{\bullet\}$ Expected value

$\nabla_{\mathbf{w}}\{\bullet\}$ Gradient vector with respect to vector $\mathbf{w}$

$\mathrm{Tr}\{\bullet\}$ Trace of a square matrix

$\{\bullet\}^{\mathrm{H}}$ Hermitian transpose

$\{\bullet\}^{\top}$ Transposition

$\{\cdot\}^{\dagger}$ Moore-Penrose inverse (pseudo-inverse)

Vectors

$\overline{\mathbf{v}}$ Overall noise vector normalized by the square root of the noise power

$\overline{\mathbf{x}}$ Speech vector normalized by the square root of the speech power

$\mathbf{a}_{\ell,m}$ Vector of acoustic transfer function from the desired source to the microphones of the hearing aids

$\mathbf{b}_{\ell,m}$ Vector of acoustic transfer function from the interfering source to the microphones of the hearing aids

$\mathbf{d}$ Generic vector of samples from both hearing aids

$\mathbf{n}$ Binaural interfering signal vector

$\mathbf{n}$ Binaural noise vector

$\mathbf{q}$ Binaural selection vector

$\mathbf{q}_\ell$ Selection vector of the reference microphone of the haring aid at side ℓ

$\mathbf{v}$ Binaural overall noise vector

$\mathbf{v}_\ell$ Vector of overall noise samples in the hearing aid ℓ

$\mathbf{w}$ Binaural noise reduction filter

$\mathbf{w}_\ell$ Noise reduction vector of the hearing aid ℓ

$\mathbf{x}$ Binaural speech signal vector

$\mathbf{x}_\ell$ Vector of speech samples in the hearing aid ℓ

$\mathbf{x}_{\text{SNR}}$ Speech Speech vector normalized by the square root of the speech power and multiplied by the input SNR

$\mathbf{y}$ Binaural noisy vector

$\mathbf{y}_\ell$ Vector of noisy samples in the hearing aid ℓ

CONTENTS

1 INTRODUCTION

Human hearing plays a fundamental role in the communication process. Consequently, an untreated loss in hearing capacity can affect personal and professional relationships, negatively affecting quality of life, can put people at risk [1] and is correlated with other health problems [2]. Noise reduction methods embedded in assistive hearing devices, such as hearing aids (HAs), reduce noise while avoiding speech distortion [3], and improve quality [4] and speech intelligibility [5]. Therefore, these methods directly contribute to the social well-being and living conditions of hearing-impaired people. However, in addition to reducing noise, noise reduction methods may affect the sound impression of the acoustic scenario, impairing the user's ability to take advantage of visual information, for instance. To mitigate this problem, this thesis mainly focuses on the development of methods that jointly reduce noise and preserve the acoustic impression of the sound source.

1.1 JUSTIFICATION

The last Brazilian census, carried out in 2010, estimated the number of people deaf or hearing impaired to be about 5.2% [6]. This represented a total of 9.8 million inhabitants[1]. Since 2004, this group is under the national hearing care policy, which distributes hearing aids and performs cochlear implant operations through the unified health system, among other actions.

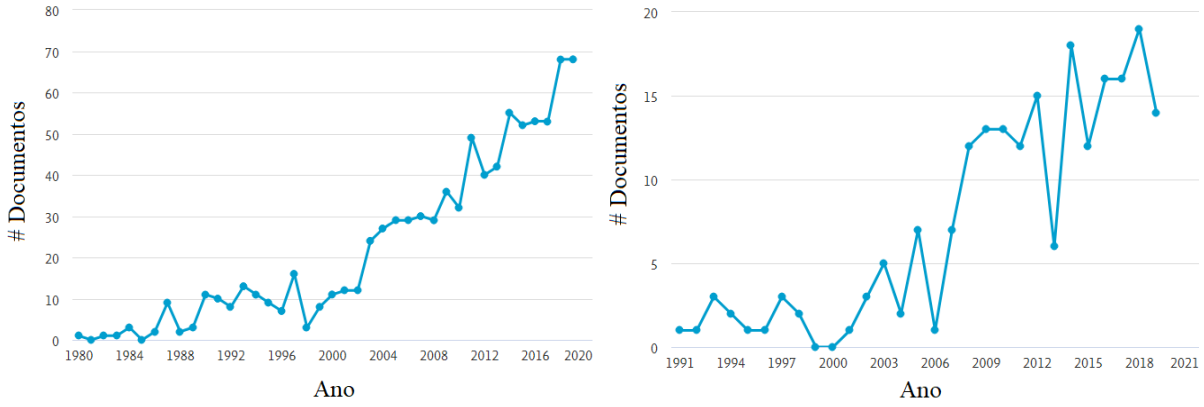
Until 2018, estimates by the World Health Organization (WHO) indicated that 6.1% of the world population, or 466 million people, had some type of hearing loss [7]. WHO also highlights that untreated hearing loss has an annual impact of $760 billion for the global economy. Due to aging and growth of the world population, if nothing is done, projections indicate that the number of people with disabling hearing loss may reach 630 million people in 2030 and 900 million in 2050. In this context, hearing aids have the benefit of increasing quality of life, as well as improving cognition and communication skills [8].

Despite the benefits mentioned, the main complaint of hearing aid users is acoustic noise. Figure 1 presents the results of a keyword search carried out based on the academic database *Scopus*. Figure 1 (a) shows the number of occurrences over time for the research: "*'hearing aids' AND ('noise reduction' OR 'speech enhancement')*". Figure 1 (b) presents the results obtained from the keywords of the previous search plus the term *'binaural'*. In both, there is an increase in interest in research for both traditional HAs and binaural HAs (HAs that share information).

Restoring hearing capacity to normal levels involves not only reducing noise but

[1] This group is divided into approximately 344,200, who cannot hear (deaf); 1 million and 800 thousand with great difficulty in hearing, and 7 million and 600 thousand who have some difficulty, even using hearing aids.

Figure 1 – Search results by keywords in `scopus.com`. Number of articles published in two topic groups per year: (a) "*'hearing aids' AND ('noise reduction' OR 'speech enhancement')*", e (b) "*'hearing aids' AND ('noise reduction' OR 'speech enhancement')* AND 'binaural' ".



also preserving a number of other abilities performed by the human auditory system, including binaural hearing. Therefore, methods that address this problem are of great importance.

1.2 HEARING AIDS AND NOISE REDUCTION

Hearing aids are electronic devices that can compensate for mild to moderate hearing loss [9]. The first written record of a device used to help a person with hearing loss dates from the 17th century in a text by Francis Bacon [10]. Such early devices were bugle-shaped and made of animal horns, metal, or wood. The sound amplification was a direct consequence of the acoustical properties of the material's shape and composition. Throughout the 20th and 21st centuries, HAs have followed the development of technology in different areas, resulting in changes in its shape, size, and usage. They changed from purely analog devices to gadgets with digital technology.

Communication in noisy environments is a common everyday situation. The presence of noise decreases the intelligibility and quality of the speech, especially for hearing-impaired persons. Therefore, noise reduction methods are an important part of HAs since they reduce the power of the undesired signals, such as those that originate from reverberation, impulsive noise, or wind [11]. Experiments carried out with HAs users under real-life usage indicated that noise reduction methods decrease auditory effort, improve sound quality, and facilitate the localization of sound sources [11]. In general, users prefer to enable noise reduction methods in their HA devices [12]. Therefore, noise reduction methods are an essential part of modern HAs since they improve the quality and intelligibility of noisy speech [13, 14].

1.3 BINAURAL HEARING AND SOUND SOURCE LOCALIZATION

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[2] compared to the case in which only one ear is stimulated [15]. By using signals at both ears, the auditory system creates sound objects that facilitate the localization, separation, and identification of sound sources [16].

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 content of the information. However, if the undesired speaker moves away, 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 [17–19].

The sound signal that reaches the ears carries information about the spatial attributes of the sound source. Therefore, the ability of the auditory system to localize, segment and track a sound source depends on *interaural dissimilarities* [16], that is, on the difference between the signals that reach the ears. The spatial localization of a point source in the horizontal plane by the auditory system is mainly made by two binaural cues: i) the interaural time difference (ITD), which originates from the difference between the arrival times; and ii) the interaural level difference (ILD), which originates from the difference between the intensities at both ears.

Another important binaural cue is interaural coherence (IC), which affects the perception of ILD and ITD in scenarios where the signals in both ears do not have perfect correlation with each other. When the number of point sources in the environment increases, the directionality of the resulting sound field is lost. In the limit, when this number becomes very large, there is no perception of direction in the acoustic field, which is then called a *diffuse sound field*. In this scenario, the IC is the main binaural cue, since it is related to the apparent width of the acoustic field [20, 21].

With normal hearing, the auditory system naturally preserves the binaural localization cues, i.e., information used to localize sound sources. However, this ability is severely reduced for hearing aid users, particularly for bilateral HAs, in which the signal in each ear is processed independently of the other. Due to the distortions in the ILD, people using bilateral HAs can perceive the sound of a single source as coming from two different spatial locations [22]. Due to this, bilateral HA users locate a sound source more easily when they do not use their devices [23]. As a consequence, the sound provided by the HAs

[2] Loudness is the perceived sound volume

to the user may not have adequate information to locate a sound source in the acoustic environment. This reinforces the need to include the preservation of binaural cues in the design of noise reduction methods.

1.4 BINAURAL NOISE REDUCTION METHODS

To restore binaural hearing capacity, binaural HAs are the most advanced option on the market for hearing-impaired people. These gadgets employ a wireless communication link between the two HAs, by which signals and control parameters are exchanged, thus increasing the diversity of information. The presence of this link allows the development of noise reduction methods with significantly better performance compared to the methods available for bilateral HAs. Moreover, it also makes it possible to preserve the binaural cues of the signals, thus preserving the original location of the sound source.

Binaural noise reduction methods must be designed considering three main objectives [24, 25]: i) decrease the power of undesired signals; ii) limit distortions in speech; and iii) reduce distortions in the binaural cues of the received signals. The first two objectives are associated with the traditional noise reduction problem and influence the quality and intelligibility of the speech signal. The last objective characterizes the binaural methods and aims to explore the advantages of binaural hearing to restore auditory spatial awareness of the environment, providing the user with an auditory notion of the spatial arrangement of sound sources.

Since binaural cues are a function of signals in both ears, binaural noise reduction methods need at least two microphones (one in each ear) to function properly. Therefore, they consist of multichannel noise reduction methods which can exploit the spectrotemporal information and the spatial information captured by the set of two or more microphones.

The binaural multichannel Wiener filter (MWF) is one of the most used methods for binaural noise reduction. It determines a linear estimator that is computed by minimizing a cost function based on the mean squared error between the desired signal and its estimated version. Its implementation depends exclusively on second-order statistics. In principle, the location of the desired source is not needed for its design, which makes the MWF robust to changes in the location of the source. In its parameterized form, the MWF can establish a compromise between speech distortion and noise reduction. Another important feature of the MWF is that it can be decomposed into two filters: one that performs spatial filtering, and another that can be interpreted as a spectral mask [26]. Regarding the preservation of sound sources, the MWF does not severely affect the ILD, ITD and IC of the desired speech source [26], leaving its perceived location intact.

However, the MWF has important disadvantages. First, it is sensitive to errors in the estimation of second-order statistics, resulting in a decrease in noise reduction performance, which affects speech quality and intelligibility [26]. In addition, such errors may degrade

the binaural cues of the desired speech signal and cause, from the perceptual point of view, a shift in the perceived position of the desired source towards the interfering source. Second, the binaural cues of the interfering source and the perception of the ambient noise may be distorted, changing their location towards the desired source. Therefore, the MWF may not achieve the objective of preserving spatial awareness of the acoustic scenario in the signals captured by the HAs.

1.5 PRESERVATION OF BINAURAL CUES

Different methods have recently been proposed to reduce the distortions introduced by the MWF in the interfering signal by calculating a spatial filter resulting from the minimization of an augmented cost function, that is, a cost function comprised of the MWF cost function and an additional term related to preservation of a binaural cue. The importance of each term in the augmented cost function is determined by a parameter that, in principle, should be adjusted by the system designer. This parameter generally depends on the statistical characteristics of the input signals and on the acoustic scenario, and must be carefully selected so that both the noise reduction and spatial preservation objectives are met.

The first cost function for preserving the binaural characteristics of an interfering source was proposed in [27]. The main idea for the design of this cost function is to penalize distortions in the ILD and ITD of a point source by minimizing the mean square error between the ILD and ITD of the input signals and of the signals processed by the hearing aids.

Furthermore, considering the case of an interfering point source, in [28] a cost function based on the interaural transfer function (ITF) was proposed. It intrinsically carries information about both ITD and ILD [29]. This approach uses a single cost function based on the mean squared error between the average reference ITF and the instantaneous ITF of the interfering processed signal.

Considering the situation of a desired source immersed in a diffuse sound field, in [30] an extension of the MWF with an additional cost function was proposed to minimize errors between the IC of the input and processed signals. Its objective was to preserve the perception of the signal coming from a diffuse sound field.

1.6 CHALLENGES WITH THE EXISTING METHODS

Despite the important advances in the development of methods for binaural cue preservation in noise reduction, existing methods face fundamental challenges when used in real-life scenarios, which prevent them from delivering satisfactory results. We discuss three major challenges below.

1.6.1 The effect of the power variation in the augmented MWF cost function

One situation that can affect the performance of binaural noise reduction methods is a variation in the power of the input signal. This can occur, for example, in situations in which the power of the speech signal varies according to the power of the interfering signal, a phenomenon known as *Lombard effect* [31]. This phenomenon may affect the adjustment of the weight parameter between the cost functions, significantly altering the performance of the noise reduction filters. Therefore, a robust design of this parameter is of fundamental importance for binaural noise reduction methods based on the addition of terms to the MWF cost function.

1.6.2 Efficient methods for MWF-based noise reduction

The design of binaural noise reduction filters using MWF-based methods is performed by solving nonconvex unconstrained optimization problems. This approach presents notable performance in offline experiments; however, it is inappropriate for online implementation in embedded systems with severe computational limitations, such as in HA applications, since the time taken to compute the coefficients is too high. Therefore, MWF-based methods that admit simple strategies to compute the filter coefficients are of paramount importance, since they can be implemented in real life HA hardware. This is a still open research field for augmented MWF methods.

1.6.3 A multichannel Wiener filter approach for preservation of the main binaural cues

Several binaural noise reduction methods have been proposed to preserve some of the binaural cues important for the localization of sound sources. However, none of these proposals promotes the joint preservation of the three main important binaural cues: ILD, ITD and IC. This means that each of the binaural noise reduction methods available in the literature is restricted to some specific acoustic scenario, and, in the eventual change of the acoustic scene, a classifier is needed to identify the change and choose the most appropriate cost function for the situation. Therefore, a technique that can preserve the principal binaural cues is important, as it can be used in different acoustic scenarios.

1.7 OBJECTIVES

From the context of the binaural noise reduction methods presented above, the general and specific objectives of this work are defined below.

The general objective of this work is to propose MWF-based binaural noise reduction methods for applications in binaural hearing aids, considering acoustic scenarios composed of a desired source and a generic noise field.

The particular objectives of this work are:

a) Develop new methods that can deal with changes in the acoustic environment and in the power of the received signals.

b) Propose time-efficient methods based on the multichannel Wiener filter for the preservation of binaural cues.

c) Validate the propositions through extensive experiments.

1.8 THE MAIN CONTRIBUTIONS OF THIS THESIS

The main contributions of this thesis include new, theoretically motivated methods that address the challenges aforementioned and extensive experimental validation.

The first contribution, presented in Chapter 3, is a method for designing the parameter that sets the trade-off between the two important parts of the MWF-based technique. The proposed method is designed for acoustic scenarios consisting of one speech source with either an point noise source [27, 28] or a diffuse noise field [30]. The strategy automatically adjusts the trade-off parameter to maintain the desired noise reduction performance without changing the perception of the location of an acoustic source. The approach can be applied directly to a variety of noise reduction methods for binaural hearing aids, such as [27, 28, 32–35] and [36].

The second contribution, presented in Chapter 4, is a closed-form solution for the MWF with the ILD preservation problem (MWF-ILD problem). The solution has parameters that are easier to design, since they are intuitive and directly linked to physical measures, compared to the parameters of the traditional approach to solve the MWF-ILD problem. The solution is guaranteed to be the globally optimal solution (exact solution) of the original problem. Finally, the derived solution severely reduces the time it takes to find the noise reduction filters, compared to the traditional approach.

The third contribution, presented in Chapter 5, is a new cost function design to preserve the main important binaural cues used to provide a realistic impression of the position and diffuseness of an acoustic source. To this end, the proposed cost function is used to preserve the binaural impression of signals generated by point sources or diffuse fields.

1.9 ORGANIZATION

This text is divided as described below. Chapter 2 presents the concepts and definitions necessary to understand the research topic. Chapter 3 discusses the problem of lack of homogeneity of terms that comprise MWF-based binaural noise reduction. From the analysis of this problem, a strategy for designing the parameter that sets the trade-off between the two important parts of the augmented MWF-based technique is proposed. This approach prevents that variations in the power of the input signal affect the

performance of the methods. Chapter 4 presents some important mathematical limitations of the MWF with ILD preservation in the form of an optimization problem. From these limitations, we propose a different reformulation of the original optimization problem, which results in a closed-form solution to the MWF-ILD technique. Chapter 5 deals with the problem of how to preserve the ILD, IPD, and IC using a single binaural cost function. The first proposition is a new cost function composed of terms that preserve both the ILD and IC of the noise source. The second proposition is a technique that approximates the solutions of the new cost function based on constrained optimization. A simulated acoustic scenario was used to assess the performance of the proposed techniques. The results show that the noise reduction coefficients designed using such approaches can be used to preserve binaural cues of point sources and diffuse fields with approximately the same noise reduction performance.

2 BACKGROUND

This chapter presents concepts and definitions related to anatomy, psychoacoustic, mathematics, and HA technology that will help the reader throughout this thesis.

2.1 NOTATION

Throughout this text, lowercase italic symbols represent scalars (e.g., a), while lowercase and uppercase bold symbols denote, respectively, vectors (e.g., $\mathbf{a}$) and matrices (e.g., $\mathbf{A}$). The subscripts in lowercase italic letters $\{\bullet\}_\ell$ denote indexes, while literals are denoted by uppercase letters $\{\bullet\}_\mathrm{L}$. The terms "minimum", "argument that minimizes" and "subject to" are abbreviated to "min.", "arg.min.", and "s.t.", respectively. The matrix inequality representation $\mathbf{A} \succ 0$ ($\mathbf{A} \succeq 0$) means A is positive (semi-)definite. The vector inequality representation $\mathbf{a} > 0$ ($\mathbf{a} \geqslant 0$) means that all elements of $\mathbf{a}$ are greater (or equal) than 0. The equivalent meaning is considered for $\mathbf{a} < 0$ ($\mathbf{a} \leqslant 0$). The zero matrix of dimension $m \times n$ is represented as $\mathbf{0}_{m\times n}$. The identity matrix is represented as $\mathbf{I}$, and its dimension is defined by the context. The sets of real, positive real, and complex numbers are represented as $\mathbb{R}$, $\mathbb{R}_{++}$, and $\mathbb{C}$, respectively. The operator $\mathrm{diag}(\bullet)$ creates a diagonal matrix with its arguments and handles different sizes of the input arguments, e.g., for the scalar a and a square matrix $\mathbf{A}$ of order m, $\mathrm{diag}(a, \mathbf{A})$ leads to the following.

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

The whole mathematical notation is defined in the list of symbols.

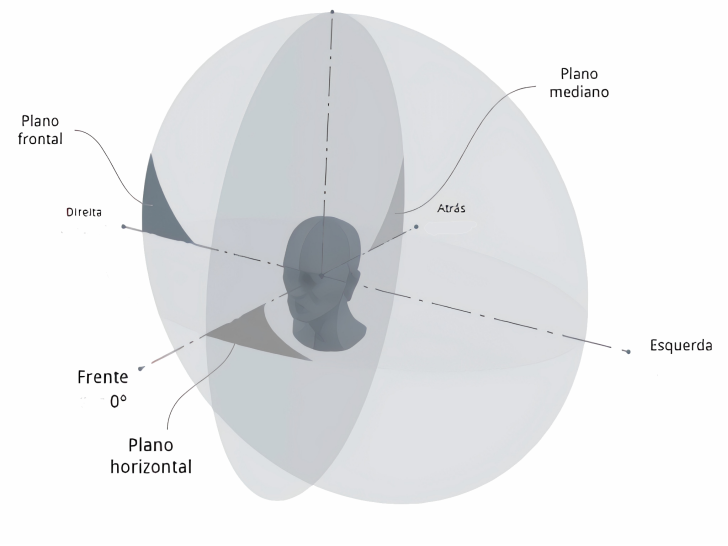
2.2 INTERAURAL COORDINATE SYSTEM

Figure 2 depicts the interaural coordinate system, i.e., each point in space is specified relative to the orientation of the head. The origin of the coordinate system is the midpoint of the imaginary line that joins the entrances of the ear canals. Mathematically, the coordinate system is a spherical coordinate system and the points are defined by three parameters: azimuth angle (θ), which quantifies horizontal angular variations; elevation angle (τ), which quantifies vertical angular variations; and distance (r), which quantifies the space between the acoustic source and the origin of the coordinate system. In this thesis, it is assumed that the sound sources are located on the horizontal plane, i.e., the plane defined for all points with zero degree of elevation ($\tau = 0$).

2.3 HUMAN HEARING

A simple anatomy of the human ear and the information used by the auditory system for sound source localization will be presented in the following sections.

Figure 2 – Interaural coordinate system adopted in this thesis: the reference point is the midpoint of the line segment joining the entrances of the ear canals. A point in space is specified by three parameters: azimuth angle (θ), elevation angle (τ), and the distance (r).



Reference: Modified from [37].

2.3.1 Anatomy of the human ear

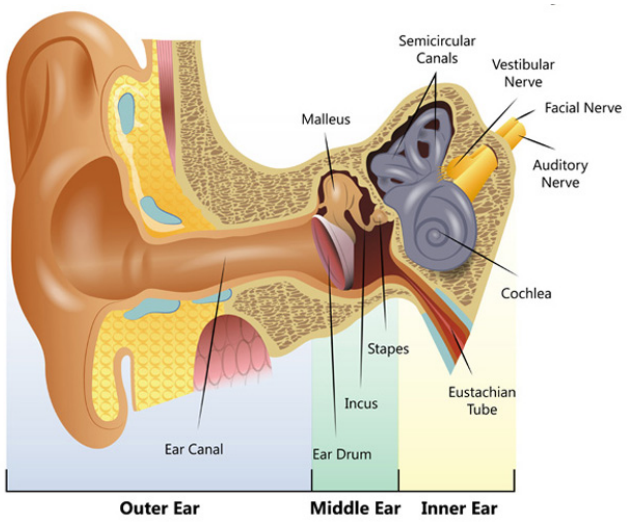
The human ear is the sensory part of the auditory system that allows auditory perception (hearing) and sense of balance. It is divided into three parts: the external ear, the middle ear, and the internal ear, which are illustrated in Figure 3.

The external ear is the initial portion of the auditory system and is made up of the pinna and the ear canal. The eardrum is located at the end of the ear canal and determines the boundary between the outer ear and the middle ear. The tympanic membrane transmits sound stimuli to a chain of three ossicles called hammer, anvil, and stirrup, whose function is to efficiently transfer the mechanical stimuli that arrive from the eardrum. In the inner ear, the auditory stimulus travels through the cochlea, a spiral tube filled with fluid (endolymph). In the cochlea, the vibrations in the endolymph are translated into an electrical pulse sequence that is sent through the auditory nerve to the central auditory pathways.

2.3.2 Spatial hearing

The concept of spatial hearing involves the study of the relationship between sound events (such as a person speaking or an alarm sounding) and auditory events, that is, the interpretation of this event by the human auditory system [16]. Thus, the localization of a sound source is the relationship between its physical attributes (position, distance, and spatial extent), which are associated with sound events; and the interpretation of these attributes by the auditory system.

Figure 3 – Human ear anatomy.



Reference: [38].

2.3.3 Binaural hearing and binaural localization cues

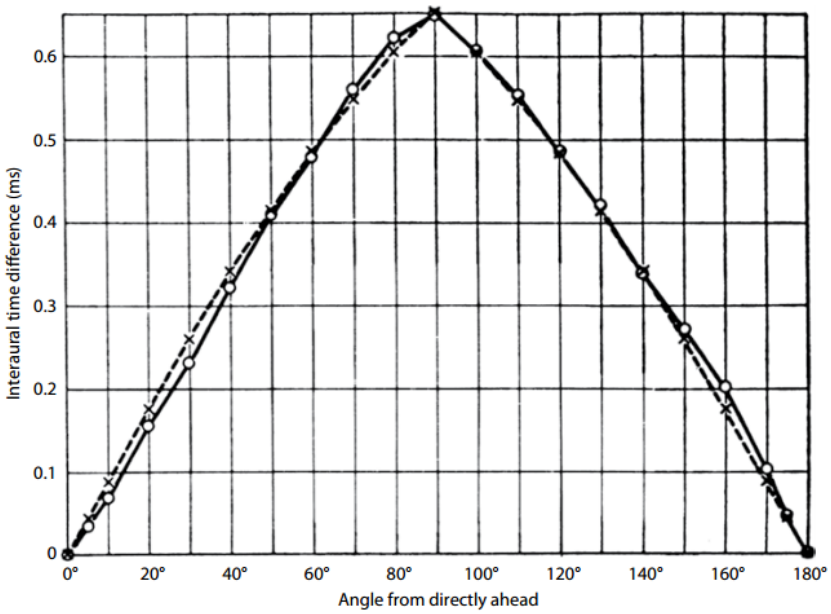
The spatial location of sound sources in the horizontal plane occurs because of differences between the sound signals that reach the ears. These differences carry information about the localization of the source of the sound event as well as the environment, and are called binaural cues.

In environments with low reverberation, the spatial location of point sources in the horizontal plane is performed by the human auditory system from ITD and ILD. Interaural time difference quantifies the difference between the arrival times of the same wavefront in ears, and is the main binaural cue at frequencies below 1500 Hz [16], since the wavelength in this region is greater than the average size of the human head.

Figure 4 presents an illustration of the ITD estimates with respect to the azimuth of the sound source [39]. In the azimuth 0°, the source is located immediately in front of the person. As the sound source moves sideways, the ITD increases, reaching its maximum value when the sound source is located immediately to the right, in azimuth $+90^\circ$. When the sound source moves towards the back of the person, the ITD decreases again, becoming minimal right behind the person, in the azimuth 180°. Since the ITDs of the azimuths 0° and 180° are the same, the limitation of this acoustic cue in differentiating when a sound source is located in front of or behind the individual is evident. This problem is known as front-back confusion and can be resolved by slightly rotating the head. In terms of sensitivity, the just noticeable difference (JND) of the ITD is around 80μs in the side region, between 30° and 90° azimuth; and around 30μs in the front region, between 0° and 30°.

The acoustic power in each ear is generally different due to physical phenomena, such as reverberation, reflection, and diffraction [16]. This power dissimilarity allows the

Figure 4 – The anatomical range of the ITD.



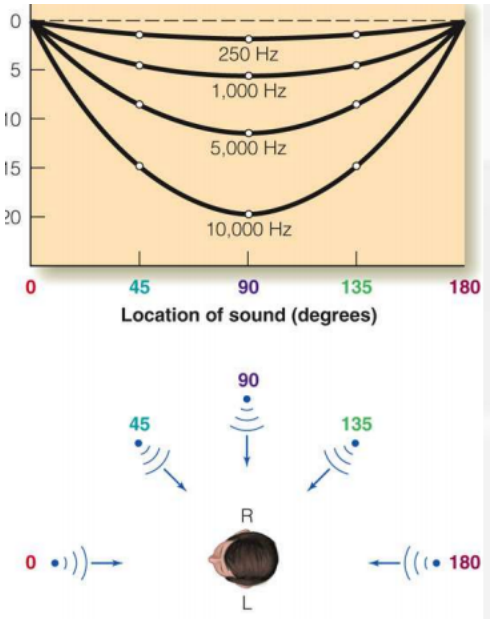
Reference: Modified representation of the figure in Feddersen's paper [39, Fig. 1] taken from [15].

human auditory system to identify different sound objects (e.g., speech) and localize their corresponding physical sources (e.g., the person speaking) [16]. The spatial information provided by the difference in acoustic power in both ears is known as the interaural level difference (ILD). It is mathematically defined as the difference (in logarithmic scale) between the left and right acoustic powers of the waves reaching the ears [16]. The proportion between the width of the human head and the wavelength of audible sounds makes the ILD the primary binaural cue for frequencies above 1500 Hz [16]. This binaural cue is robust to coherence variations [40] and contains enough information to create complex acoustic scenarios, such as in amplitude stereo panning techniques, even with headphones [41].

Figure 5 presents an example of the behavior of the ILD considering the frequency of the sound component and the azimuth of the source. For low frequencies, ILD provides little capacity for discrimination between different azimuths. As the frequency increases, ILD increases the capacity to differentiate different spatial positions. In anechoic environments, the minimum perceptible variation of ILD (JND-ILD) is around 0.6 dB [40]. In reverberant environments, JND-ILD can be up to twice as large [42], i.e. it can be up to 1.2 dB.

Interaural coherence is defined as the normalized cross-correlation between signals in the ears and determines the reliability of ILD and ITD [30], being associated with the perceived size of the sound source [42]. When the sound field is diffuse, IC is the relevant binaural information [30].

Figure 5 – ILD variation per frequency as a function of the source azimuth.



Reference: [43].

2.4 BINAURAL SYSTEM

Figure 6 depicts the investigated acoustic scenario, which is made up of a hearing aid user, a point desired source (speaker), and a point interfering source, both contaminated by noise. The binaural system consists of two hearing aids: 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 in a frequency bin index k, time frame index λ, microphone $m = \{1, 2, \cdots, M_\ell\}$ and on the side $\ell \in \{\mathrm{L}, \mathrm{R}\}$, is given by

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

in which $x_{\ell,m}(\lambda, k)$ is the desired component (speech); and $v_{\ell,m}(\lambda, k)$ is the overall noise component, which is represented as a sum of an interfering component $u_{\ell,m}(\lambda, k)$ and a remaining (coherent or incoherent) noise component $n_{\ell,m}(\lambda, k)$:

$$v_{\ell,m}(\lambda, k) = u_{\ell,m}(\lambda, k) + n_{\ell,m}(\lambda, k). \tag{3}$$

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

$$\mathbf{y}_\ell(\lambda, k) = \left[\; y_{\ell,1}(\lambda, k) \;\; y_{\ell,2}(\lambda, k) \;\; y_{\ell,3}(\lambda, k) \;\; \cdots \;\; y_{\ell,M_\ell}(\lambda, k) \right]^\top, \tag{4}$$

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

$$\mathbf{y}(\lambda, k) = [\; \mathbf{y}_{\mathrm{L}}^\top(\lambda, k) \quad \mathbf{y}_{\mathrm{R}}^\top(\lambda, k) \;]^\top. \tag{5}$$

Vector $\mathbf{y}(\lambda, k)$ is available on both HAs due to a full duplex communication link and can be decomposed as

$$\mathbf{y}(\lambda, k) = \mathbf{x}(\lambda, k) + \mathbf{v}(\lambda, k), \tag{6}$$

in which $\mathbf{x}(\lambda, k)$ is the speech vector and $\mathbf{v}(\lambda, k)$ is the overall noise vector, whose entries are defined similarly as $\mathbf{y}(\lambda, k)$ in (5). Due to (3), vector $\mathbf{v}(\lambda, k)$ can be decomposed as:

$$\mathbf{v}(\lambda, k) = \mathbf{u}(\lambda, k) + \mathbf{n}(\lambda, k), \tag{7}$$

in which $\mathbf{u}(\lambda, k)$ and $\mathbf{n}(\lambda, k)$ are, respectively, the interfering and noise vectors.

For point speech and interfering sources, vectors $\mathbf{x}(\lambda, k)$ and $\mathbf{u}(\lambda, k)$ can be modeled as

$$\mathbf{x}(\lambda, k) = s_{\mathrm{x}}(\lambda, k)\mathbf{a}(\lambda, k), \tag{8}$$

$$\mathbf{u}(\lambda, k) = s_{\mathrm{u}}(\lambda, k)\mathbf{b}(\lambda, k), \tag{9}$$

in which $s_{\mathrm{x}}(\lambda, k)$ and $s_{\mathrm{u}}(\lambda, k)$ are clean speech and interfering signals; and vectors

$$\mathbf{a}(\lambda, k) = [\, a_{\mathrm{L},1}(\lambda, k)\, a_{\mathrm{L},2}(\lambda, k)\, \cdots\, a_{\mathrm{L},M_\ell}(\lambda, k)\, a_{\mathrm{R},1}(\lambda, k) \cdots\, a_{\mathrm{R},M_{\mathrm{R}}}(\lambda, k)]^\top, \tag{10}$$

and $$\mathbf{b}(\lambda, k) = [\, b_{\mathrm{L},1}(\lambda, k)\, b_{\mathrm{L},2}(\lambda, k)\, \cdots\, b_{\mathrm{L},M_{\mathrm{L}}}(\lambda, k)\, b_{\mathrm{R},1}(\lambda, k)\, \cdots\, b_{\mathrm{R},M_{\mathrm{R}}}(\lambda, k)\,]^\top, \tag{11}$$

are 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 this work, it is assumed that the ATFs are time-invariant and that the narrow band model holds [25, 44], leading to $\mathbf{a}(\lambda, k) = \mathbf{a}(k)$.

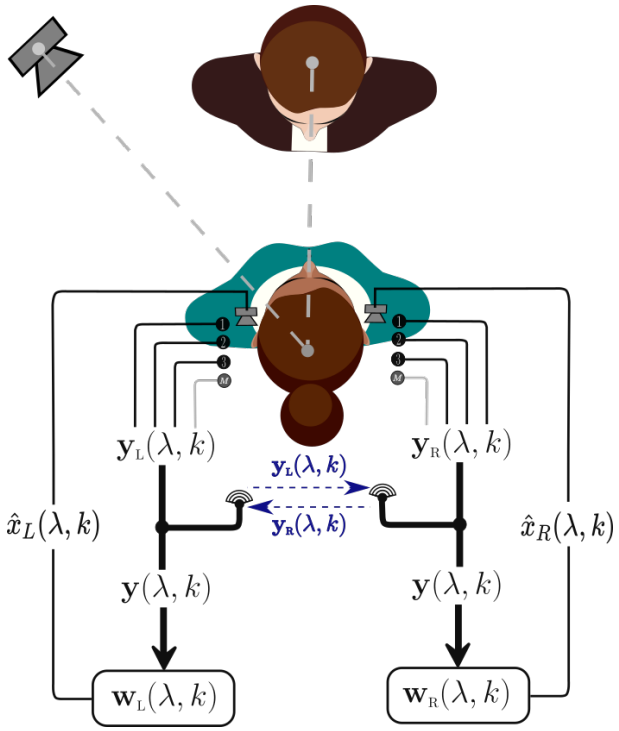
In general, a microphone on each hearing aid is defined as a reference microphone, whose associated signal is defined as

$$\begin{aligned}
\mathbf{q}_\ell^\top \mathbf{y}(\lambda, k) &= y_\ell(\lambda, k) && \text{(12a)}\\
&= x_\ell(\lambda, k) + v_\ell(\lambda, k) && \text{(12b)}\\
&= x_\ell(\lambda, k) + u_\ell(\lambda, k) + n_\ell(\lambda, k) && \text{(12c)}\\
&= s_{\mathrm{x}}(\lambda, k)a_\ell(\lambda, k) + s_{\mathrm{u}}(\lambda, k)b_\ell(\lambda, k) + n_\ell(\lambda, k), && \text{(12d)}
\end{aligned}$$

in which $\mathbf{q}_\ell$ is a microphone selection vector with entry equal to 1 in the position of the reference microphone on the side ℓ and zero elsewhere; $x_\ell(\lambda, k)$, $u_\ell(\lambda, k)$, $n_\ell(\lambda, k)$ and $v_\ell(\lambda, k)$ are the speech, interfering, noise and overall-noise components in the reference microphone; and $a_\ell(\lambda, k)$, and $b_\ell(\lambda, k)$ are the ATFs associated with $x_\ell(\lambda, k)$ and $u_\ell(\lambda, k)$, respectively.

The noise reduction problem consists of estimating the desired speech for each HA, in each time and frequency bin, using the noisy samples of the $2M$ microphones. The

Figure 6 – Diagram of a binaural hearing aid system with a full-duplex communication link.



estimated speech components are defined in the frequency domain as:

$$z_\ell(\lambda,k) = \mathbf{w}_\ell^{\mathrm{H}}(\lambda,k)\mathbf{y}(\lambda,k) \tag{13a}$$
$$= \mathbf{w}_\ell^{\mathrm{H}}(\lambda,k)\mathbf{x}(\lambda,k) + \mathbf{w}_\ell^{\mathrm{H}}(\lambda,k)\mathbf{v}(\lambda,k) \tag{13b}$$
$$= \mathbf{w}_\ell^{\mathrm{H}}(\lambda,k)\mathbf{x}(\lambda,k) + \mathbf{w}_\ell^{\mathrm{H}}(\lambda,k)\mathbf{u}(\lambda,k) + \mathbf{w}_\ell^{\mathrm{H}}(\lambda,k)\mathbf{n}(\lambda,k), \tag{13c}$$

in which $\mathbf{w}_\ell(\lambda,k) \in \mathbb{C}^{2M}$ is the noise reduction filter on the side ℓ. The binaural selection vector $\mathbf{q}$ is defined, respectively, as:

$$\mathbf{q} = \left[\ \mathbf{q}_{\mathrm{L}}^\top \ \ \mathbf{q}_{\mathrm{R}}^\top\ \right]^\top. \tag{14}$$

The coherence matrix of a generic received component d is defined as

$$\mathbf{\Phi}_{\mathrm{d}}(\lambda,k) = \mathbb{E}\{\mathbf{d}(\lambda,k)\mathbf{d}^{\mathrm{H}}(\lambda,k)\}, \tag{15}$$

in which $\mathbb{E}\{\bullet\}$ is the expected value operator; $\mathbf{d}(\lambda,k)$ is a generic vector, representing the vectors in the set $\{\mathbf{x}(\lambda,k), \mathbf{u}(\lambda,k), \mathbf{n}(\lambda,k), \mathbf{v}(\lambda,k)\}$; and d is a generic label, representing elements in the set $\{\mathrm{x}, \mathrm{u}, \mathrm{n}, \mathrm{v}\}$. Substituting the definitions in (8) and (9) into the generic definition in (15), leads to the definition of the coherence matrices of the speech, and interference, which are, respectively, given by:

$$\mathbf{\Phi}_{\mathrm{x}}(\lambda,k) = p_{sx}(\lambda,k)\mathbf{a}(\lambda,k)\mathbf{a}^{\mathrm{H}}(\lambda,k) \tag{16}$$
$$\mathbf{\Phi}_{\mathrm{u}}(\lambda,k) = p_{su}(\lambda,k)\mathbf{b}(\lambda,k)\mathbf{b}^{\mathrm{H}}(\lambda,k), \tag{17}$$

in which $p_{sx}(\lambda,k)=\mathbb{E}\{|s_{\mathrm{x}}(\lambda,k)|^2\}$ and $p_{su}(\lambda,k)=\mathbb{E}\{|s_{\mathrm{u}}(\lambda,k)|^2\}$ are the power spectrum density (PSD) of clean speech and interfering signals, respectively, and $|\bullet|$ is the absolute value operator.

Considering $\mathbf{d}(\lambda,k)=\mathbf{y}(\lambda,k)$ in (15) and assuming that the speech component is uncorrelated with the overall noise component, i.e., $\mathbb{E}\{\mathbf{x}\mathbf{v}^{\mathrm{H}}\}=\mathbb{E}\{\mathbf{x}\mathbf{v}^{\mathrm{H}}\}=\mathbf{0}_{M\times M}$, and that $\mathbb{E}\{\mathbf{x}\}=\mathbb{E}\{\mathbf{v}\}=\mathbf{0}$ [28, 32], the noisy input coherence matrix ($\mathbf{\Phi}_{\mathrm{y}}$) is defined from (5) as:

$$\mathbf{\Phi}_{\mathrm{y}}(\lambda,k)=\mathbf{\Phi}_{\mathrm{x}}(\lambda,k)+\mathbf{\Phi}_{\mathrm{v}}(\lambda,k) \tag{18}$$

in which $\mathbf{\Phi}_{\mathrm{v}}=\mathbb{E}\{\mathbf{v}(\lambda,k)\mathbf{v}^{\mathrm{H}}(\lambda,k)\}$ is the coherence matrix of the overall noise, given by:

$$\mathbf{\Phi}_{\mathrm{v}}(\lambda,k)=\mathbf{\Phi}_{\mathrm{u}}(\lambda,k)+\mathbf{\Phi}_{\mathrm{n}}(\lambda,k), \tag{19}$$

in which $\mathbf{\Phi}_{\mathrm{n}}(\lambda,k)$ is the coherence matrix of the noise component.

2.5 BINAURAL CUES

In this section, generic expressions for binaural cues are presented. To avoid divergent definitions, generic signals are used in the definition. The signal $d_\ell(\lambda,k)$ is the received component in the reference microphone, e.g., $\mathbf{q}_\ell^\top\mathbf{x}(\lambda,k)$ or $\mathbf{q}_\ell^\top\mathbf{u}(\lambda,k)$. The signal $p_{\mathrm{d}\ell}$ represents a generic processed component, e.g., $\mathbf{w}_\ell^\top\mathbf{x}$ or $\mathbf{w}_\ell^\top\mathbf{v}$; and $\mathbf{\Phi}_{\mathrm{d}}$ is defined in (15). The input binaural cues are functions of d_ℓ, while the output binaural cues are functions of $p_{\mathrm{d}\ell}$. Note that both d_ℓ and $p_{\mathrm{d}\ell}$ depend on $\mathbf{d}$; for this reason, all binaural cue definitions carry the superscript $d\in\{x,\,v\}$. The operation $(\bullet)^*$ below is the complex conjugate.

The input and output interaural transfer functions (ITFs) are defined respectively as [45]:

$$\mathrm{ITF}_{\mathrm{in}}^{\mathrm{d}}=\frac{\mathbb{E}\{d_{\mathrm{L}}\,d_{\mathrm{R}}^*\}}{\mathbb{E}\{|d_{\mathrm{R}}|^2\}}=\frac{\mathbf{q}_{\mathrm{L}}^\top\mathbf{\Phi}_{\mathrm{d}}\mathbf{q}_{\mathrm{R}}}{\mathbf{q}_{\mathrm{R}}^\top\mathbf{\Phi}_{\mathrm{d}}\mathbf{q}_{\mathrm{R}}}, \tag{20}$$

$$\mathrm{ITF}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w})=\frac{\mathbb{E}\{p_{d\mathrm{L}}p_{d\mathrm{R}}^*\}}{\mathbb{E}\{|p_{d\mathrm{R}}|^2\}}=\frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}}. \tag{21}$$

The input and output interaural level differences (ILDs) are defined, respectively, as [45]:

$$\mathrm{ILD}_{\mathrm{in}}^{\mathrm{d}}=\frac{\mathbb{E}\{|d_{\mathrm{L}}|^2\}}{\mathbb{E}\{|d_{\mathrm{R}}|^2\}}=\frac{\mathbf{q}_{\mathrm{L}}^\top\mathbf{\Phi}_{\mathrm{d}}\mathbf{q}_{\mathrm{L}}}{\mathbf{q}_{\mathrm{R}}^\top\mathbf{\Phi}_{\mathrm{d}}\mathbf{q}_{\mathrm{R}}}, \tag{22}$$

$$\mathrm{ILD}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w})=\frac{\mathbb{E}\{|p_{d\mathrm{L}}|^2\}}{\mathbb{E}\{|p_{d\mathrm{R}}|^2\}}=\frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}}. \tag{23}$$

The input and output interaural phase differences (IPDs) are defined, respectively, as [45]:

$$\mathrm{IPD}_{\mathrm{in}}^{\mathrm{d}}=\angle(\mathbb{E}\{d_{\mathrm{L}}\,d_{\mathrm{R}}^*\})=\angle(\mathbf{q}_{\mathrm{L}}^\top\mathbf{\Phi}_{\mathrm{d}}\mathbf{q}_{\mathrm{R}}), \tag{24}$$

$$\mathrm{IPD}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w})=\angle(\mathbb{E}\{p_{d\mathrm{L}}p_{d\mathrm{R}}^*\})=\angle(\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}). \tag{25}$$

with $\angle(\bullet)$ denoting the phase of its argument (complex number).

The input and output interaural coherences (ICs) are, respectively, defined as [30]:

$$\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}} = \frac{\mathbb{E}\{d_{\mathrm{L}}\, d_{\mathrm{R}}^{*}\}}{\sqrt{\mathbb{E}\{|d_{\mathrm{L}}|^2\} \cdot \mathbb{E}\{|d_{\mathrm{R}}|^2\}}} = \frac{\mathbf{q}_{\mathrm{L}}^{\top} \mathbf{\Phi}_{\mathrm{d}} \mathbf{q}_{\mathrm{R}}}{\sqrt{\mathbf{q}_{\mathrm{L}}^{\top} \mathbf{\Phi}_{\mathrm{d}} \mathbf{q}_{\mathrm{L}} \cdot \mathbf{q}_{\mathrm{R}}^{\top} \mathbf{\Phi}_{\mathrm{d}} \mathbf{q}_{\mathrm{R}}}}, \tag{26}$$

$$\mathrm{IC}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w}) = \frac{\mathbb{E}\{p_{d\mathrm{L}} p_{d\mathrm{R}}^{*}\}}{\sqrt{\mathbb{E}\{|p_{d\mathrm{L}}|^2\} \cdot \mathbb{E}\{|p_{d\mathrm{R}}|^2\}}} = \frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}} \mathbf{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}}}{\sqrt{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}} \mathbf{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{L}} \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \mathbf{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}}}}. \tag{27}$$

The input and output mean square coherences (MSCs) are defined, respectively, as [46]:

$$\mathrm{MSC}_{\mathrm{in}}^{\mathrm{d}} = |\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}|^2, \tag{28}$$

$$\mathrm{MSC}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w}) = |\mathrm{IC}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w})|^2, \tag{29}$$

From (20), (24) and (26), and (21), (25) and (27) the following relations can be obtained [30]:

$$\mathrm{IPD}_{\mathrm{in}}^{\mathrm{d}} = \angle(\mathrm{ITF}_{\mathrm{in}}^{\mathrm{d}}) = \angle(\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}), \tag{30}$$

$$\mathrm{IPD}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w}) = \angle(\mathrm{ITF}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w})) = \angle(\mathrm{IC}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w})). \tag{31}$$

From (20), (22), (24), (26), and (28) we can define the following relation for the input ITF [30]:

$$\mathrm{ITF}_{\mathrm{in}}^{\mathrm{d}} = [\mathrm{ILD}_{\mathrm{in}}^{\mathrm{d}}]^{1/2} \cdot \mathrm{IC}_{\mathrm{in}}^{\mathrm{d}} \tag{32a}$$

$$= [\mathrm{ILD}_{\mathrm{in}}^{d} \cdot \mathrm{MSC}_{\mathrm{in}}^{d}]^{1/2} \cdot \exp\left(j\mathrm{IPD}_{\mathrm{in}}^{\mathrm{d}}\right). \tag{32b}$$

In the same way, from (21), (23), (25), (27), and (29) we obtain the following generic relation for the output binaural cues of a generic sound field [30]:

$$\mathrm{ITF}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w}) = [\mathrm{ILD}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w})]^{1/2} \cdot \mathrm{IC}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w}) \tag{33a}$$

$$= [\mathrm{ILD}_{\mathrm{ou}}^{\mathrm{d}} \cdot \mathrm{MSC}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w})]^{1/2} \cdot \exp\left(j\mathrm{IPD}_{\mathrm{ou}}^{\mathrm{d}}(\mathbf{w})\right). \tag{33b}$$

2.6 THE MULTICHANNEL WIENER FILTER (MWF)

The classical binaural MWF noise reduction method is based on the mean squared error (MSE) criterion. It defines the best linear estimators for speech in reference microphones ($x_{\mathrm{L}}(\lambda,k) = \mathbf{q}_{\mathrm{L}}^{\top}\mathbf{x}(\lambda,k)$ and $x_{\mathrm{R}}(\lambda,k) = \mathbf{q}_{\mathrm{R}}^{\top}\mathbf{x}(\lambda,k)$) by minimizing the following cost function [26]:

$$J_{\mathrm{MWF}}(\lambda,k) = \mathbb{E}\left\{\left\|\begin{bmatrix}\mathbf{q}_{\mathrm{L}}^{\top}\mathbf{x}(\lambda,k) - \mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\mathbf{y}(\lambda,k) \\ \mathbf{q}_{\mathrm{R}}^{\top}\mathbf{x}(\lambda,k) - \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\mathbf{y}(\lambda,k)\end{bmatrix}\right\|^2\right\}, \tag{34}$$

in which $||\bullet||$ is the Euclidean norm.

From (7) the following expected value can be defined:

$$\mathbb{E}\{\mathbf{x}(\lambda,k)\mathbf{v}^{\mathrm{H}}(\lambda,k)\} = \mathbb{E}\{\mathbf{x}(\lambda,k)\mathbf{u}^{\mathrm{H}}(\lambda,k)\} + \mathbb{E}\{\mathbf{x}(\lambda,k)\mathbf{n}^{\mathrm{H}}(\lambda,k)\}, \tag{35}$$

and considering that the real and imaginary parts of the random variables have zero mean [32] and that speech is uncorrelated with the interference and noise signals [29, 33, 44, 47], or precisely that

$$\mathbb{E}\{\mathbf{x}(\lambda,k)\mathbf{u}^{\mathrm{H}}(\lambda,k)\} = \mathbf{0}_{M\times M}, \tag{36}$$

$$\mathbb{E}\{\mathbf{x}(\lambda,k)\mathbf{v}^{\mathrm{H}}(\lambda,k)\} = \mathbf{0}_{M\times M}, \tag{37}$$

the cost function in (34) turns to [26]:

$$\begin{aligned} J_{\mathrm{MWF}}(\lambda,k) =& \mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\mathbf{\Phi}_{\mathrm{y}}(\lambda,k)\mathbf{w}_{\mathrm{L}}(\lambda,k) - \mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\mathbf{\Phi}_{\mathrm{x}}(\lambda,k)\mathbf{q}_{\mathrm{L}} - \mathbf{q}_{\mathrm{L}}^{\top}(\lambda,k)\mathbf{\Phi}_{\mathrm{x}}(\lambda,k)\mathbf{w}_{\mathrm{L}}(\lambda,k) \\ &+\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\mathbf{\Phi}_{\mathrm{y}}(\lambda,k)\mathbf{w}_{\mathrm{R}}(\lambda,k) - \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\mathbf{\Phi}_{\mathrm{x}}(\lambda,k)\mathbf{q}_{\mathrm{R}} - \mathbf{q}_{\mathrm{R}}^{\top}(\lambda,k)\mathbf{\Phi}_{\mathrm{x}}(\lambda,k)\mathbf{w}_{\mathrm{R}}(\lambda,k) \\ &+\mathbf{q}_{\mathrm{L}}^{\top}\mathbf{\Phi}_{\mathrm{x}}(\lambda,k)\mathbf{q}_{\mathrm{L}} + \mathbf{q}_{\mathrm{R}}^{\top}\mathbf{\Phi}_{\mathrm{x}}(\lambda,k)\mathbf{q}_{\mathrm{R}}. \end{aligned} \tag{38}$$

Representing (38) in a simplified form leads to:

$$\begin{aligned} J_{\mathrm{MWF}}(\lambda,k) = \mathbf{w}^{\mathrm{H}}(\lambda,k)\mathbf{\Phi}_{\mathrm{yy}}(\lambda,k)\mathbf{w}(\lambda,k) - \mathbf{w}^{\mathrm{H}}(\lambda,k)\mathbf{p}_{\mathrm{xx}}(\lambda,k) \\ - \mathbf{p}_{\mathrm{xx}}^{\mathrm{H}}(\lambda,k)\mathbf{w}(\lambda,k) + p_{\mathrm{xx}}(\lambda,k), \end{aligned} \tag{39}$$

in which:

$$\mathbf{w}(\lambda,k) = \left[\begin{array}{cc} \mathbf{w}_{\mathrm{L}}^{\top}(\lambda,k) & \mathbf{w}_{\mathrm{R}}^{\top}(\lambda,k) \end{array}\right]^{\top}, \tag{40}$$

$$\mathbf{p}_{\mathrm{xx}}(\lambda,k) = [\,\mathbf{q}_{\mathrm{L}}^{\top}\mathbf{\Phi}_{\mathrm{x}}^{\top}(\lambda,k)\;\mathbf{q}_{\mathrm{R}}^{\top}\mathbf{\Phi}_{\mathrm{x}}^{\top}(\lambda,k)\,]^{\top}, \tag{41}$$

$$p_{\mathrm{xx}}(\lambda,k) = \mathbf{q}_{\mathrm{L}}^{\top}\mathbf{\Phi}_{\mathrm{x}}(\lambda,k)\mathbf{q}_{\mathrm{L}} + \mathbf{q}_{\mathrm{R}}^{\top}\mathbf{\Phi}_{\mathrm{x}}(\lambda,k)\mathbf{q}_{\mathrm{R}}, \tag{42}$$

$$\mathbf{\Phi}_{\mathrm{yy}}(\lambda,k) = \mathrm{diag}(\mathbf{\Phi}_{\mathrm{y}}(\lambda,k), \mathbf{\Phi}_{\mathrm{y}}(\lambda,k)). \tag{43}$$

The optimization problem to find the MWF coefficients is defined as:

$$\mathbf{w}_{\mathrm{MWF}}(\lambda,k) = \underset{\mathbf{w}(\lambda,k)}{\arg.\min.}\; J_{\mathrm{MWF}}(\lambda,k). \tag{44}$$

To find a closed-form solution to $\mathbf{w}_{\mathrm{MWF}}(\lambda,k)$, we first take the gradient of $J_{\mathrm{MWF}}(\lambda,k)$ with relation to $\mathbf{w}(\lambda,k)$, which leads to:

$$\nabla_{\mathbf{w}} J_{\mathrm{MWF}}(\lambda,k) = \mathbf{w}^{\mathrm{H}}(\lambda,k)\mathbf{\Phi}_{\mathrm{yy}}(\lambda,k) - \mathbf{p}_{\mathrm{xx}}(\lambda,k) \tag{45}$$

Then, we equate (45) to zero, assuming $\mathbf{\Phi}_{\mathrm{yy}}(\lambda,k) \succeq 0$. This process results in the closed-form solution of the generalized MWF (GMWF) filter, given by [48]:

$$\mathbf{w}_{\mathrm{GMWF}} = \mathbf{\Phi}_{\mathrm{yy}}^{\dagger}(\lambda,k)\mathbf{p}_{\mathrm{xx}}(\lambda,k), \tag{46}$$

in which $\{\cdot\}^{\dagger}$ is the Moore-Penrose inverse (pseudo-inverse) [49], and $\mathbf{\Phi}_{\mathrm{xx}}(\lambda,k)$ is a block diagonal matrix given by:

$$\mathbf{\Phi}_{\mathrm{xx}}(\lambda,k) = \begin{bmatrix} \mathbf{\Phi}_{\mathrm{x}}(\lambda,k) & \mathbf{0}_{M\times M} \\ \mathbf{0}_{M\times M} & \mathbf{\Phi}_{\mathrm{x}}(\lambda,k) \end{bmatrix}. \tag{47}$$

In the particular case in which $\mathbf{\Phi}_{\mathrm{yy}}(\lambda,k) \succ 0$, then $\mathbf{\Phi}^{\dagger}_{\mathrm{yy}}(\lambda,k) = \mathbf{\Phi}^{-1}_{\mathrm{yy}}(\lambda,k)$, and the solution in (46) becomes the classical binaural MWF [26]:

$$\mathbf{w}_{\mathrm{MWF}} = \mathbf{\Phi}^{-1}_{\mathrm{yy}}(\lambda,k)\mathbf{p}_{\mathrm{xx}}(\lambda,k). \tag{48}$$

In this situation, considering the case of the point speech source, it is theoretically proven that the binaural cues of processed speech and the residual noise at the output are equal to the binaural cues of the input speech [26]. As a result, both signals are psychoacoustically perceived to be coming from the speech direction [50].

2.7 THE MULTICHANNEL WIENER FILTER WITH ADDITIONAL TERM FOR PRESERVATION OF BINAURAL CUES

The MWF cost function can be complemented by additional penalty terms that aim to prevent solutions that do not preserve the original spatial characteristics [27, 30, 32]. This leads to a minimization problem of an MWF-based augmented cost function in the following form:

$$\mathbf{w}_{\text{MWF-BM}}(\lambda,k) = \underset{\mathbf{w}(\lambda,k)}{\arg.\min.}\ J_{\mathrm{T}}(\lambda,k) \tag{49a}$$

$$= \underset{\mathbf{w}(\lambda,k)}{\arg.\min.}\ J_{\mathrm{MWF}}(\lambda,k) + \sum_{i=1}^{I} \alpha_i(\lambda,k) J_i(\lambda,k) \tag{49b}$$

where $\alpha_i(\lambda,k) \in \mathbb{R}_+$ are the weighting (trade-off) parameters that control the balance (minimization effort) among different terms; and I is the number of penalty terms. $J_i(\lambda,k)$ represents different terms related to the preservation of the same binaural cue – ITD, IPD, ILD, and IC. Some preservation terms are presented in the following subsections.

2.7.1 Preservation of Interaural Level Difference (ILD) and Interaural Phase Difference (IPD)/Interaural Time Difference (ITD)

Considering an acoustic scenario composed of one speech and one single interfering sources, the authors of [27] first presented a penalty term for the preservation of both the ILD and the IPD of the interfering component, given by:

$$\sum_{i=1}^{I} \alpha_i(\lambda,k) J_i(\lambda,k) = \alpha_1(\lambda,k) J_1(\lambda,k) + \alpha_2(\lambda,k) J_2(\lambda,k) \tag{50}$$

$$= \alpha_{\mathrm{ILD}}(\lambda,k) J^{\mathrm{d}}_{\mathrm{ILD}}(\lambda,k) + \alpha_{\mathrm{IPD}}(\lambda,k) J^{\mathrm{d}}_{\mathrm{IPD}}(\lambda,k) \tag{51}$$

in which $\alpha_1(\lambda,k) = \alpha_{\mathrm{ILD}}(\lambda,k)$, $\alpha_2(\lambda,k) = \alpha_{\mathrm{IPD}}(\lambda,k)$, $J_1(\lambda,k) = J^{\mathrm{d}}_{\mathrm{ILD}}(\lambda,k)$ and $J_2(\lambda,k) = J^{\mathrm{d}}_{\mathrm{IPD}}(\lambda,k)$. The terms $J^{\mathrm{d}}_{\mathrm{ILD}}(\lambda,k)$ and $J^{\mathrm{d}}_{\mathrm{IPD}}(\lambda,k)$ are defined respectively as:

$$J^{\mathrm{d}}_{\mathrm{ILD}}(\lambda,k) = [\mathrm{ILD}^{\mathrm{d}}_{\mathrm{ou}}(\lambda,k) - \mathrm{ILD}^{\mathrm{d}}_{\mathrm{in}}(\lambda,k)]^2, \tag{52}$$

$$J^{\mathrm{d}}_{\mathrm{IPD}}(\lambda,k) = 1 - \cos(\mathrm{IPD}^{\mathrm{d}}_{\mathrm{ou}}(\lambda,k) - \mathrm{IPD}^{\mathrm{d}}_{\mathrm{in}}(\lambda,k)). \tag{53}$$

2.7.2 Preservation of the Interaural Transfer Function (ITF)

Considering the same acoustic scenario in [27], the authors in [51] presented a cost function to preserve the interaural transfer function (ITF) of the noise component. The ITF carries intrinsic information about IPD and ILD [29]. As a result, the approach in [51] makes use of a single auxiliary cost function defined as:

$$J_{\mathrm{ITF}}^{\mathrm{d}}(\lambda,k)=\frac{E\{|\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\mathbf{d}(\lambda,k)-\mathrm{ITF}_{\mathrm{in}}^{\mathrm{v}}(\lambda,k)\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\mathbf{d}(\lambda,k)|^2\}}{E\{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\mathbf{d}(\lambda,k)\mathbf{d}^{\mathrm{H}}(\lambda,k)\mathbf{w}_{\mathrm{R}}(\lambda,k)\}}\,. \tag{54}$$

2.7.3 Preservation of the Interaural Coherence (IC)

In [30], a cost function was presented to preserve the original perception of a diffuse noise field. It is based on the error between the output and input interaural coherence (IC) of the noise component, i.e.,

$$J_{\mathrm{IC}}^{\mathrm{d}}(\lambda,k)=|\mathrm{IC}_{\mathrm{ou}}^{\mathrm{d}}(\lambda,k)-\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}(\lambda,k)|^2. \tag{55}$$

It was demonstrated in [32] that (55) is also effective in preserving the ITD.

3 ROBUST PARAMETER STRATEGY FOR WIENER-BASED BINAURAL NOISE REDUCTION METHODS IN HEARING AIDS

A large and widely employed subclass of MWF-based binaural noise reduction methods aims to preserve both speech and noise localization cues. It is based on spatial filters calculated as the solution to an optimization problem that minimizes the weighted sum of two specific cost functions. The first is the MWF cost function, which aims to promote noise reduction at a limited speech distortion. The second is a set of penalization terms related to binaural cue preservation, which aims to retain the auditory impression of the acoustic sources in the processed signal. The relative weight between both terms of this augmented cost function is determined by an adjustable parameter, whose setpoint depends on the statistical characteristics of the received signals and acoustic scenario. This parameter must be carefully selected to establish the desired trade-off between noise reduction and binaural cue preservation.

In [27], an augmented MWF-based cost function was proposed considering two additive penalty terms to preserve both the ITD and the ILD of the processed noise, which are known to be the prevalent binaural cues for localization of point sound sources in the median plane of the head. In [28], the preservation of both the ITD and ILD was obtained using a single penalty term, which considered the interaural transfer function (ITF). The ITF carries information from both ITD and ILD binaural cues [29]. However, this method was shown to be incapable of preserving the spatiality of diffuse noise fields. To address this issue, the IC cost function was proposed in [30] for the preservation of environmental diffuse noise. In [32], it was shown that the IC cost function can also be used to preserve the ITD of point sources with improved lateralization performance in comparison to the approach presented in [27].

A systematic review of the literature performed in two of the most important scientific databases indicated that two strategies have been applied for designing the weight parameters employed by augmented MWF-based noise reduction techniques, which are: (i) empirical/arbitrary (fixed) parameter setting [32, 34, 51]; and (ii) psychoacoustical parameter optimization [30, 36]. Fixed weight parameters achieve adequate performance in acoustic scenarios under stationary and time-invariant conditions. In [30], a psychoacoustically motivated method for obtaining constraint boundaries for the MWF-IC weight parameter was presented. However, this method is limited to diffuse noise fields and requires an exhaustive/iterative search procedure that is computationally expensive, because it has to be applied to each time-frequency bin.

Regardless of the nature of the acoustic field, the desired trade-off between noise reduction and spatial preservation should be maintained constant under speech/noise power variations. These variations are intrinsic to spoken communication, in which unpredictable changes in the environmental noise may compel the speaker to adjust the speech level to maintain intelligible communication [52]. This phenomenon is known as the Lombard

effect [31, 53, 54] and affects a number of speech features, such as voice intensity, spectral slope of glottal waveforms, formant locations and bandwidths, and energy ratios in voiced/unvoiced phonemes [55, 56]. Although this is a well-studied problem in the speech recognition area, it has not received sufficient attention in the binaural hearing aid field.

A review of MWF-based binaural techniques shows that most binaural-cue penalty terms presented in literature are invariant to input power variations [27–30, 32]. However, this is not the case for the MWF cost function. As a result, the optimal trade-off between noise reduction and preservation of noise binaural cues, provided by a fixed weight parameter (applied into the augmented cost function), may be significantly affected by the absolute power of the received signals (e.g., Lombard effect), requiring a conservative setting. Such observations impose the need for a normalized weight parameter to achieve the optimal performance under power variations in the input signals. This approach would alleviate the need for repetitive manual adjustments of the binaural hearing-aid control-parameters to accommodate variations of the acoustic scenario [57].

This chapter presents a method for the automatic adjustment of MWF-based binaural techniques to maintain the desired noise reduction performance, as well to preserve the original perception of the acoustic scenario. To the best of the author's knowledge, this is the first approach in this area, whose results can be directly applied to a variety of noise reduction methods for binaural hearing aids such as [27, 28, 32] and [33–36].

The proposed method is designed for acoustic scenarios comprising of one speech source with either an interfering point noise source [27, 28] or a diffuse noise field [30]. A theoretical analysis is presented to highlight the mechanisms by which input signal power variations affect the performance of binaural MWF-based methods. Subsequently, we present the mathematical properties that should be satisfied by an augmented MWF cost function to present setpoint invariance to input signal power variations. These observations motivated the design of a normalized weight parameter to achieve robust setpoint invariance. Computational simulations with objective criteria, as well as lateralization psychoacoustic experiments with headphones, illustrate the performance of the proposed method. The main contributions of this study are as follows: (a) a structure and a method for designing the weighting parameter of Wiener-based binaural noise reduction methods for hearing aid applications; (b) a theoretical analysis to elucidate its operating mechanism and support its expected performance; and (c) computational simulations with objective measures and real psychoacoustic results to corroborate the effectiveness of the proposed method.

The remainder of this chapter is structured as follows. Section 3.1 introduces the problem formulation, and Section 3.2 presents the proposed structure and method. The experimental setup is described in Section 3.3, and the results and discussion are presented in Sections 3.4 and 3.5, respectively. Finally, concluding remarks are presented in Section 3.6.

3.1 PROBLEM FORMULATION

An extensive review of the literature indicated that, in general, the previously proposed augmented MWF techniques employ time-invariant trade-off parameters (i.e., $\alpha_i(\lambda, k) = \beta_i(k)$) in the cost function presented in (49) [26, 30, 32, 58]. Thus, the performance of the binaural noise reduction method may be influenced by the power variations in the received signals [26]. To highlight this characteristic, speech and noise signals are represented in a normalized form as follows:

$$\mathbf{v}(\lambda, k) = g(\lambda, k)\widetilde{\mathbf{v}}(\lambda, k), \tag{56}$$

$$\mathbf{x}(\lambda, k) = g(\lambda, k)\mathrm{SNR}_{\mathrm{in}}^{1/2}(\lambda, k)\widetilde{\mathbf{x}}(\lambda, k), \tag{57}$$

in which $g^2(\lambda, k) \in \mathbb{R}_+$ is the mean noise power at the input microphones of the hearing aids given by:

$$g^2(\lambda, k) = \mathbb{E}\{||\mathbf{v}(\lambda, k)||^2\}, \tag{58}$$

vectors $\widetilde{\mathbf{v}}(\lambda, k)$ and $\widetilde{\mathbf{x}}(\lambda, k)$ are the unit variance noise and speech vectors given, respectively, by:

$$\widetilde{\mathbf{v}}(\lambda, k) = \frac{\mathbf{v}(\lambda, k)}{\mathbb{E}\{\|\mathbf{v}(\lambda, k)\|^2\}^{1/2}}, \tag{59}$$

$$\widetilde{\mathbf{x}}(\lambda, k) = \frac{\mathbf{x}(\lambda, k)}{\mathbb{E}\{\|\mathbf{x}(\lambda, k)\|^2\}^{1/2}}, \tag{60}$$

and the $\mathrm{SNR}_{\mathrm{in}}(\lambda, k)$ is the overall input SNR given by:

$$\mathrm{SNR}_{\mathrm{in}}(\lambda, k) = \frac{\mathbb{E}\{||\mathbf{x}(\lambda, k)||^2\}}{\mathbb{E}\{||\mathbf{v}(\lambda, k)||^2\}} = \frac{\mathbb{E}\{||\mathbf{x}(\lambda, k)||^2\}}{g^2(\lambda, k)}. \tag{61}$$

Using (57) and (56) in (2) results in:

$$\mathbf{y}(\lambda, k) = g(\lambda, k)\left[\,\mathrm{SNR}_{\mathrm{in}}^{1/2}(\lambda, k)\widetilde{\mathbf{x}}(\lambda, k) + \widetilde{\mathbf{v}}(\lambda, k)\,\right] \tag{62a}$$

$$= g(\lambda, k)\left[\,\widetilde{\mathbf{x}}_{\mathrm{SNR}} + \widetilde{\mathbf{v}}(\lambda, k)\,\right], \tag{62b}$$

for

$$\widetilde{\mathbf{x}}_{\mathrm{SNR}} = \mathrm{SNR}_{\mathrm{in}}^{1/2}(\lambda, k)\widetilde{\mathbf{x}}(\lambda, k). \tag{63}$$

Equation (62) models the noisy input vector using a mutual gain and a weighted sum of two normalized signals. Because $\mathbb{E}\{||\widetilde{\mathbf{x}}(\lambda, k)||^2\} = \mathbb{E}\{||\widetilde{\mathbf{v}}(\lambda, k)||^2\} = 1$, $\mathrm{SNR}_{\mathrm{in}}(\lambda, k)$ establishes the SNR of the noisy input vector $\mathbf{y}(\lambda, k)$. The parameter $g(\lambda, k)$ is named here as the Lombard gain [59], because it is associated with the speech effort ($\mathbb{E}\{||\widetilde{\mathbf{x}}(\lambda, k)||^2\} = \mathrm{SNR}_{\mathrm{in}}(\lambda, k)g^2(\lambda, k)$) required to maintain the same communication condition (i.e., a constant input SNR). Assuming a fixed SNR, equation (58) establishes a simplified model to

represent the Lombard effect, because it does not consider other spectral modifications [55, 56]. Despite its simplicity, equation (62) constitutes a meaningful method of investigating the effect of input power variations on the performance of MWF-based binaural noise reduction methods. This can be achieved by considering the joint changes of both noise and speech absolute power (constant SNR), as well as modifications of the input SNR. To do this, let us represent the J_{MWF} presented in (34) by a different form given by:

$$J_{\mathrm{MWF}}(\lambda,k)=\mathbb{E}\left\{\left\|\begin{bmatrix}(\mathbf{q}_{\mathrm{L}}^{\top}-\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k))\mathbf{x}(\lambda,k)-\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\mathbf{v}(\lambda,k)\\(\mathbf{q}_{\mathrm{R}}^{\top}-\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k))\mathbf{x}(\lambda,k)-\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\mathbf{v}(\lambda,k)\end{bmatrix}\right\|^{2}\right\}, \tag{64}$$

which is equivalent to:

$$\begin{aligned}J_{\mathrm{MWF}}(\lambda,k)&=\mathbb{E}\{||(\mathbf{q}_{\mathrm{L}}^{\top}-\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k))\mathbf{x}(\lambda,k)-\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\mathbf{v}(\lambda,k)||^{2}\}\\&\quad+\mathbb{E}\{||(\mathbf{q}_{\mathrm{R}}^{\top}-\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k))\mathbf{x}(\lambda,k)-\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\mathbf{v}(\lambda,k)||^{2}\}.\end{aligned} \tag{65}$$

Using (56) and (63) in (65) results in:

$$\begin{aligned}J_{\mathrm{MWF}}(\lambda,k)&=\mathbb{E}\{||g(\lambda,k)((\mathbf{q}_{\mathrm{L}}^{\top}-\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k))\widetilde{\mathbf{x}}_{\mathrm{SNR}}(\lambda,k)-g(\lambda,k)\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\widetilde{\mathbf{v}}(\lambda,k)||^{2}\}\\&\quad+\mathbb{E}\{||g(\lambda,k)(\mathbf{q}_{\mathrm{R}}^{\top}-\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k))\widetilde{\mathbf{x}}_{\mathrm{SNR}}(\lambda,k)-g(\lambda,k)\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\widetilde{\mathbf{v}}(\lambda,k)||^{2}\}.\end{aligned} \tag{66}$$

Replacing (56) in (66), and isolating $g(\lambda,k)$, since it is a real deterministic scalar value, equation (66) results in:

$$\begin{aligned}J_{\mathrm{MWF}}(\lambda,k)&=g^{2}(\lambda,k)\cdot\big[\mathbb{E}\{||(\mathrm{SNR}_{\mathrm{in}}(\lambda,k))^{1/2}\cdot(\mathbf{q}_{\mathrm{L}}^{\top}-\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k))\widetilde{\mathbf{x}}(\lambda,k)-\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\widetilde{\mathbf{v}}(\lambda,k)||^{2}\}\\&\quad+\mathbb{E}\{||(\mathrm{SNR}_{\mathrm{in}}(\lambda,k))^{1/2}\cdot(\mathbf{q}_{\mathrm{R}}^{\top}-\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k))\widetilde{\mathbf{x}}(\lambda,k)-\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\widetilde{\mathbf{v}}(\lambda,k)||^{2}\}\big].\end{aligned} \tag{67}$$

Manipulating (67) considering speech and noise both as zero-mean random variables [32] and uncorrelated with each other [47, 60], results in:

$$\begin{aligned}J_{\mathrm{MWF}}(\lambda,k)&=g^{2}(\lambda,k)\cdot\big[\mathrm{SNR}_{\mathrm{in}}(\lambda,k)\cdot(\mathbf{q}_{\mathrm{L}}^{\top}-\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k))\boldsymbol{\Phi}_{\widetilde{\mathrm{x}}}(\mathbf{q}_{\mathrm{L}}-\mathbf{w}_{\mathrm{L}}(\lambda,k))-\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\boldsymbol{\Phi}_{\widetilde{\mathrm{v}}}\mathbf{w}_{\mathrm{L}}(\lambda,k)\\&\quad+\mathrm{SNR}_{\mathrm{in}}(\lambda,k)\cdot(\mathbf{q}_{\mathrm{R}}^{\top}-\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k))\boldsymbol{\Phi}_{\widetilde{\mathrm{x}}}(\mathbf{q}_{\mathrm{R}}-\mathbf{w}_{\mathrm{R}}(\lambda,k))-\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\boldsymbol{\Phi}_{\widetilde{\mathrm{x}}}\mathbf{w}_{\mathrm{R}}(\lambda,k)\big].\end{aligned} \tag{68}$$

in which $\boldsymbol{\Phi}_{\widetilde{\mathrm{x}}}(\lambda,k)=\boldsymbol{\Phi}_{\mathrm{x}}(\lambda,k)/\mathbb{E}\{||\mathbf{x}(\lambda,k)||^{2}$ and $\boldsymbol{\Phi}_{\widetilde{\mathrm{v}}}(\lambda,k)=\boldsymbol{\Phi}_{\mathrm{v}}(\lambda,k)/\mathbb{E}\{||\mathbf{v}(\lambda,k)||^{2}$. Representing equation (68) using (14) and (40) leads to:

$$J_{\mathrm{MWF}}(\lambda,k)=g^{2}(\lambda,k)\big[\mathrm{SNR}_{\mathrm{in}}(\lambda,k)(\mathbf{q}^{\top}-\mathbf{w}^{\mathrm{H}}(\lambda,k))\boldsymbol{\Phi}_{\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}}(\mathbf{q}-\mathbf{w}(\lambda,k))-\mathbf{w}^{\mathrm{H}}(\lambda,k)\boldsymbol{\Phi}_{\widetilde{\mathbf{v}}\widetilde{\mathbf{v}}}\mathbf{w}(\lambda,k)\big], \tag{69}$$

in which $\boldsymbol{\Phi}_{\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}}(\lambda,k)=\mathrm{diag}\big(\boldsymbol{\Phi}_{\widetilde{\mathrm{x}}}(\lambda,k),\boldsymbol{\Phi}_{\widetilde{\mathrm{x}}}(\lambda,k)\big)$ and $\boldsymbol{\Phi}_{\widetilde{\mathbf{v}}\widetilde{\mathbf{v}}}(\lambda,k)=\mathrm{diag}\big(\boldsymbol{\Phi}_{\widetilde{\mathrm{v}}}(\lambda,k),\boldsymbol{\Phi}_{\widetilde{\mathrm{v}}}(\lambda,k)\big)$.

Finally, equation (69) can be simplified as:

$$J_{\mathrm{MWF}}(\lambda,k)=g^{2}(\lambda,k)\left[\mathrm{SNR}_{\mathrm{in}}(\lambda,k)\left\|\mathbf{q}-\mathbf{w}(\lambda,k)\right\|^{2}_{\boldsymbol{\Phi}^{1/2}_{\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}}}+\left\|\mathbf{w}(\lambda,k)\right\|^{2}_{\boldsymbol{\Phi}^{1/2}_{\widetilde{\mathbf{v}}\widetilde{\mathbf{v}}}}\right], \tag{70}$$

in which $||\mathbf{d}||^{2}_{\mathbf{A}^{1/2}}=||\mathbf{A}^{1/2}\mathbf{d}||^{2}$ for $\mathbf{A}^{1/2}\in\mathbb{C}^{M\times M}$.

In the following, the time frame λ and frequency index k are omitted in the equations for clarity, whenever possible.

3.1.1 Homogeneity Degree

A useful concept for obtaining theoretical insights into the interplay between the Lombard gain ($g(\lambda, k)$) and the intrinsic trade-off (speech-distortion/noise-reduction versus speech/noise binaural cue preservation) associated with the augmented MWF cost function is the "homogeneity degree" [61] of a function, which can be defined as follows:

Definition 1: A function $f : \mathbb{C}^M \times \mathbb{C}^M \rightarrow \mathbb{R}$ is said to be positively homogeneous of degree P (or P-homogeneous), with respect to (w.r.t.) both $\mathbf{a}$, $\mathbf{b} \in \mathbb{C}^M$, for $P \in \mathbb{R}$, if $f(c\mathbf{a}, c\mathbf{b}) = c^P f(\mathbf{a}, \mathbf{b})$ for all nonzero $c \in \mathbb{R}_+$.

Because the minimum of a function is invariant to positive scaling, the minimum of any P-homogeneous function is invariant to joint positive changes in $\mathbf{a}$ and $\mathbf{b}$, i.e., min. $f(c\mathbf{a}, c\mathbf{b}) =$ min. $c^P f(\mathbf{a}, \mathbf{b})$, for $c \in \mathbb{R}_+$. This property can be used to understand the influence of the Lombard effect (speech and noise joint power variation) on the performance of MWF-based binaural noise reduction methods, as well as to formulate a robust choice for the weighting parameters $\alpha_i(\lambda, k)$ of (49).

Applying Definition 1 into (70), it can be concluded that J_{MWF} is a 2-homogenous function w.r.t. both $\widetilde{\mathbf{x}}_{\text{SNR}}$ and $\widetilde{\mathbf{v}}$, because:

$$J_{\text{MWF}}(g\widetilde{\mathbf{x}}_{\text{SNR}}, g\widetilde{\mathbf{v}}) = g^2 J_{\text{MWF}}(\widetilde{\mathbf{x}}_{\text{SNR}}, \widetilde{\mathbf{v}}). \tag{71}$$

On the other hand, it can be verified that the binaural preservation terms J_i, defined in (52)-(55), and the ratio form defined in (20)-(27) are 0-homogenous functions w.r.t. $\widetilde{\mathbf{x}}_{\text{SNR}}$, $\widetilde{\mathbf{x}}$ and $\widetilde{\mathbf{v}}$ because (for noise)

$$J_i(g\widetilde{\mathbf{x}}_{\text{SNR}}, g\widetilde{\mathbf{v}}) = J_i(g\widetilde{\mathbf{v}}) = J_i(\widetilde{\mathbf{v}}) \tag{72}$$

and (for speech)

$$J_i(g\widetilde{\mathbf{x}}_{\text{SNR}}, g\widetilde{\mathbf{v}}) = J_i(g\widetilde{\mathbf{x}}_{\text{SNR}}) = J_i(\widetilde{\mathbf{x}}_{\text{SNR}}) = J_i(\widetilde{\mathbf{x}}). \tag{73}$$

Using (71), (72) and (73) in (49), and considering a time-invariant weight parameter $\alpha(\lambda, k) = \beta(k)$, it is easy to verify that the augmented cost function J_{T} in (49) is not homogeneous, because J_{MWF} and J_i have different degrees of homogeneity. As a result, the minimum of J_{T} is affected by both $g(\lambda, k)$ and $\text{SNR}_{\text{in}}(\lambda, k)$, which means that the desired trade-off between noise reduction and preservation of the binaural cues may change under different acoustic conditions.

3.2 PROPOSED METHOD

In this section, we propose a design method for the weighting parameters $\alpha_i(\lambda, k)$, with the aim of obtaining a robust trade-off between noise-reduction/speech-distortion and binaural-cue preservation, against variations in g and SNR_{in}. As demonstrated in Section 3.1.1, time-invariant weighting parameters $\alpha_i(\lambda, k) = \beta_i(k)$ lead to inhomogeneous J_{T},

which is affected by input power variations. Therefore, we propose to employ a normalized $\alpha_i(\lambda, k)$ defined as the multiplication of a time-invariant weight by the noise power (squared Lombard gain), resulting in:

$$\alpha_i(\lambda, k) = \beta_i(k) g^2(\lambda, k), \tag{74}$$

in which $\beta_i(k)$ is a time-invariant (for each k and i) that defines the setpoint associated with the desired trade-off between noise reduction and binaural cue preservation. Using (71)–(74) in (49) yields

$$J_{\mathrm{T}}(g\mathbf{x}_{\mathrm{SNR}}, g\widetilde{\mathbf{v}}) = g^2(\lambda, k)\big[J_{\mathrm{MWF}}(\widetilde{\mathbf{x}}_{\mathrm{SNR}}, \widetilde{\mathbf{v}}) + \beta_i(k) \sum_{i=1}^{I} J_i(\widetilde{\mathbf{v}})\big] \tag{75a}$$

$$= g^2(\lambda, k) J_{\mathrm{T}}(\widetilde{\mathbf{x}}_{\mathrm{SNR}}, \widetilde{\mathbf{v}}), \tag{75b}$$

in which I is the number of penalty terms associated with the noise binaural cues.

Equation (75) shows that the optimal solution ($\mathbf{w}_{\mathrm{L}}$ and $\mathbf{w}_{\mathrm{R}}$) to the resulting augmented MWF-based cost function is invariant to the Lombard gain. This occurs, because in (75) J_{T} is 2-homogeneous w.r.t. $\widetilde{\mathbf{x}}_{\mathrm{SNR}}$, $\widetilde{\mathbf{x}}$ and $\widetilde{\mathbf{v}}$. This means that joint variations of speech and noise power should not affect the designed performance of the binaural noise reduction method.

3.2.1 Robustness to SNR variations

Although (75) ensures robustness against the Lombard effect, the resulting augmented cost function is still inhomogeneous w.r.t. $\mathrm{SNR}_{\mathrm{in}}$ variations and, as a result, there is no guarantee of setpoint invariance under such conditions. To investigate the impact of $\mathrm{SNR}_{\mathrm{in}}$ variations on the performance of the proposed method, we employed a heuristic reasoning analysis. By using (70) in (75), the minimization problem to obtain the optimal (left and right) coefficient-vectors can be written as:

$$\begin{aligned}\text{min. } J_{\mathrm{T}}(\lambda, k) = g^2(\lambda, k) \text{ min. } \Big[& \overbrace{\mathrm{SNR}_{\mathrm{in}}(\lambda, k)\big\|\mathbf{q} - \mathbf{w}(\lambda, k)\big\|^2_{\mathbf{\Phi}^{1/2}_{\mathbf{xx}}} + \big\|\mathbf{w}(\lambda, k)\big\|^2_{\mathbf{\Phi}^{1/2}_{\mathbf{vv}}}}^{J_{\mathrm{MWF}}(\widetilde{\mathbf{x}}_{\mathrm{SNR}}, \widetilde{\mathbf{v}})} \\ & + \beta_i(k) \sum_{i=1}^{I} J_i(\lambda, k)\Big].\end{aligned} \tag{76}$$

From (76), it can be noted that increasing $\mathrm{SNR}_{\mathrm{in}}$ boosts the significance of $\big\|\mathbf{q} - \mathbf{w}(\lambda, k)\big\|^2_{\mathbf{\Phi}^{1/2}_{\bar{x}\bar{x}}}$, which can be interpreted as a weighted similarity measure between $\mathbf{q}$ and $\mathbf{w}$. Thus, as $\mathrm{SNR}_{\mathrm{in}}$ increases, the strength of the first term in (76) becomes bigger than that of the other two, which are not directly influenced by $\mathrm{SNR}_{\mathrm{in}}$. This leads to solutions to the minimization problem $\mathbf{w}_{\mathrm{L}}$ and $\mathbf{w}_{\mathrm{R}}$ to approximate $\mathbf{q}_{\mathrm{L}}$ and $\mathbf{q}_{\mathrm{R}}$, respectively. In this case, increases in the input SNR lead to decreases in the magnitude and phase distortions of the noise in both ears, restoring its original binaural-cues. As a result, the parameters

$\beta_i(k)$ that satisfy the noise binaural cue preservation requirements at a given input SNR, also satisfy the same requirements at a higher input SNR condition. A similar behavior was previously demonstrated in [62] with the conventional Wiener filter, whose optimal solution corresponds to an all-pass filter for very large input SNRs.

3.2.2 Weight Parameter Design

Considering the previous reasoning, devising a parameter-design strategy now becomes straightforward. The proposed approach is based on the structure presented in (74), which should provide setpoint invariance to the Lombard gain. Then, considering the worst predicted $\mathrm{SNR}_{\mathrm{in}}$ condition, the weight parameters β_i are chosen as the smallest values such that the minimum required preservation of the speech/noise binaural-cues are obtained. In this way, higher input SNRs naturally reinforce binaural-cue preservation. A pseudocode for this procedure is detailed in Algorithm 1. Although it is an iterative procedure, this only needs to be done once (a-priori) for each frequency bin.

Algorithm 1 Pseudocode for the parameter design strategy

Input: $\mathrm{SNR}_{\mathrm{WORST}}(k)$
Output: $\beta_i(k)$

1: Generate an acoustic scenario using the worst possible SNR ($\mathrm{SNR}_{\mathrm{WORST}}$) as $\mathrm{SNR}_{\mathrm{in}}$.
2: For all k
3: Set $\beta(k) = 0$;
4: Compute the solution of the MWF-based method in (76) for $\beta_i(k)$;
5: Assess the binaural cue preservation of the noise source at the output using objective criteria;
6: **if** the binaural cue preservation is not satisfactory **then**
7: Increase $\beta_i(k)$ by a small amount and return to step 4;
8: **Otherwise**
9: Finish;
10: **Return** $\beta_i(k)$

3.3 EXPERIMENTAL SETUP

The main goal of the presented case study is to illustrate the improved robustness of the proposed method against variations in both Lombard gain and input SNR, better preserving the original perception of the noise source. In this case study, the considered acoustic scenario was comprised of a target speech source and a single point noise source. Subscripts x and v will be used to associat a method, a parameter or a performance measure to the speech and noise component, respectively.

3.3.1 Processing methods

To illustrate the performance of the proposed method we present numerical simulations and psychoacoustic experiments performed using the MWF defined by the minimization of (38), and the MWF with the noise interaural-transfer-function preservation (MWF-ITF$_{\mathrm{v}}$) approach, implemented by $J_i = J_{\mathrm{ITF}}$ defined in (54). Comparisons were conducted among unprocessed noise (RAW), noise processed by the binaural MWF method, by the MWF-ITF$_{\mathrm{v}}$ using a fixed weight parameter, and by the proposed method applied to the MWF-ITF$_{\mathrm{v}}$; here called the robust multichannel Wiener filter with noise interaural-transfer-function preservation (R-MWF-ITF$_{\mathrm{v}}$). This last method comprises Equations (58) and (74), and the method described in Algorithm 1 applied to the MWF-ITF$_{\mathrm{v}}$.

3.3.2 Acoustic scenario simulation

The acoustic scenario was simulated using head-related impulse responses (HRIRs) obtained from an anechoic chamber using a manikin Bruel & Kjær type 4128-C wearing a pair of behind-the-ear hearing aids [63]. The anechoic scenario was chosen to avoid the influence of extraneous binaural cue distortions due to acoustic reflections in the presented experiments; however, this is not a precondition. Each hearing aid has $M_{\mathrm{L}} = M_{\mathrm{R}} = 3$ microphones, resulting in $M = M_{\mathrm{L}} + M_{\mathrm{R}} = 6$ microphones. The speech source was assumed to be located at azimuth $\theta_{\mathrm{x}} = 0^{\circ}$ and distance $r_{\mathrm{x}} = 0.8$ m from the manikin, whereas the noise source was at $\theta_{\mathrm{v}} = -60^{\circ}$ (left side) and $r_{\mathrm{v}} = 3.0$ m. Both sources have 0° of elevation angle.

3.3.3 Speech and noise Signals

A set of 32 speech audio files from the repeated Harvard database [64] was selected. The sentence "the fruit peel was cut in thick slices" was uttered by the same female speaker. Contamination noise was obtained from a set of 32 non-overlapping segments of ICRA noise type I [65], which is a speech-like noise with spectral and temporal properties similar to those of the human voice. Figure 7(a) and Figure 7(b) illustrate the speech and noise spectrograms, respectively. Figure 8 shows the estimated power spectrum densities (PSD) of both speech and noise. The sampling frequency was $f_s = 16$ kHz.

3.3.4 Objective performance measures

Simulation results were assessed according to the following objective performance measures.

The overall input SNR of hearing aid of side ℓ is defined as [29]:

$$i\mathrm{SNR}_{\ell} = \sum_{k=1}^{K} 10\log_{10}\left(\frac{\mathbf{q}_{\ell}^{\top}\mathbf{\Phi}_{\mathrm{x}}(k)\mathbf{q}_{\ell}}{\mathbf{q}_{\ell}^{\top}\mathbf{\Phi}_{\mathrm{v}}(k)\mathbf{q}_{\ell}}\right), \tag{77}$$

Figure 7 – Spectrograms of: (a) speech sentence: "the fruit peel was cut in thick slices"; and (b) one segment of the ICRA noise type I.

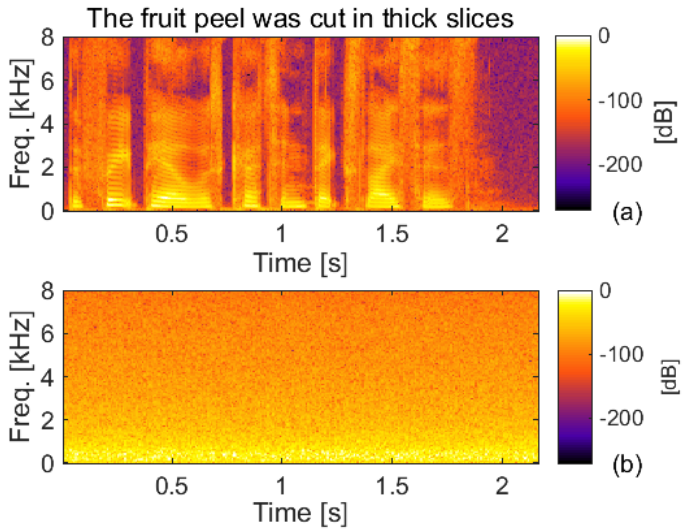
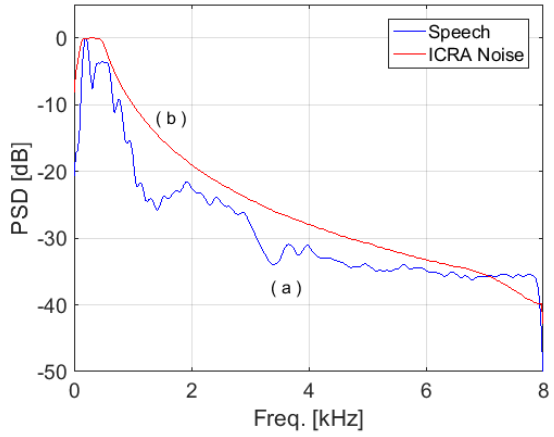


Figure 8 – Estimated power spectral densities (PSD): (a) speech (blue); and (b) ICRA type I noise (red).



in which K is the number of bins of frequency.

The variation between the overall input and output SNR in the HA ℓ, which measures the noise reduction performance of the techniques and is given by:

$$\Delta\text{SNR}_\ell = \sum_{k=1}^{K} 10\log_{10}\left(\frac{\mathbf{w}_\ell^{\text{H}}(k)\mathbf{\Phi}_{\text{x}}(k)\mathbf{w}_\ell(k)}{\mathbf{w}_\ell^{\text{H}}(k)\mathbf{\Phi}_{\text{v}}(k)\mathbf{w}_\ell(k)}\right) - \sum_{k=1}^{K} 10\log_{10}\left(\frac{\mathbf{q}_\ell^{\top}\mathbf{\Phi}_{\text{x}}(k)\mathbf{q}_\ell}{\mathbf{q}_\ell^{\top}\mathbf{\Phi}_{\text{v}}(k)\mathbf{q}_\ell}\right). \tag{78}$$

The ILD and IPD variation, which measure the input-output binaural cue preservation, are defined respectively as [29, 35]:

$$\Delta\text{IPD}_{\text{d}} = \frac{10}{K}\sum_{k=1}^{K}\left|\angle\mathbf{w}_{\text{L}}^{\text{H}}\mathbf{\Phi}_{\text{d}}(k)\mathbf{w}_{\text{R}} - \angle\mathbf{q}_{\text{L}}^{\top}\mathbf{\Phi}_{\text{d}}(k)\mathbf{q}_{\text{R}}\right|, \tag{79}$$

$$\Delta\text{ILD}_{\text{d}} = \frac{10}{K}\sum_{k=1}^{K}\left|\log_{10}\left(\frac{\mathbf{w}_{\text{L}}^{\text{H}}(k)\mathbf{\Phi}_{\text{d}}(k)\mathbf{w}_{\text{L}}(k)}{\mathbf{w}_{\text{R}}^{\text{H}}(k)\mathbf{\Phi}_{\text{d}}(k)\mathbf{w}_{\text{R}}(k)}\right) - \log_{10}\left(\frac{\mathbf{q}_{\text{L}}^{\top}(k)\mathbf{\Phi}_{\text{d}}(k)\mathbf{q}_{\text{L}}(k)}{\mathbf{q}_{\text{R}}^{\top}(k)\mathbf{\Phi}_{\text{d}}(k)\mathbf{q}_{\text{R}}(k)}\right)\right|. \tag{80}$$

The overall microphone SNR over all bins is defined as:

$$\overline{\mathrm{SNR}}_{\mathrm{in}} = \sum_{k=1}^{K} 10\log_{10}\left(\frac{\mathrm{Tr}\{\mathbf{\Phi}_{\mathrm{x}}(k)\}}{\mathrm{Tr}\{\mathbf{\Phi}_{\mathrm{v}}(k)\}}\right), \tag{81}$$

in which Tr(•) is the trace of a matrix [66].

The average squared Lombard gain defined in dB as:

$$\overline{g}^2 = \sum_{k=1}^{K} 10\log_{10}\Big(\mathrm{Tr}\left\{\mathbf{\Phi}_{\mathrm{v}}(k)\right\}\Big). \tag{82}$$

The coherence matrices $\mathbf{\Phi}_{\mathrm{x}}(k)$ and $\mathbf{\Phi}_{\mathrm{v}}(k)$ were calculated directly from the input speech-only (x) and noise-only signals (v).

3.3.5 Psychoacoustic experiments

Psychoacoustic experiments were conducted to assess the performance of the volunteers during a lateralization task. A headphone Sennheiser HD 202 was connected to a laptop, using a Realtek® high-definition audio onboard sound card, under the Windows 10 operating system.

Noise-only signals processed by the assessed methods were employed to quantify the perceived azimuth of arrival. A total of 16 volunteers participated in the experiment, and were divided into two groups. In the first group, employed for Lombard gain analysis, there were two females and six males, with an average age of 25.6 years and a standard deviation of 5 years. In the second group, which was employed to analyze SNR variations, three females and five males performed the experiment. The average age of this group was 23.2 years and the standard deviation was 4 years. The experiments were approved by the Ethics Committee on Human Research, under certificate 90899518.7.0000.0121 CEP-UFSC. All volunteers read and signed a written informed consent form.

After an initial period necessary for adjusting the volume to comfortable levels and providing instructions, the experiment procedure was into three stages: (a) learning: in which the volunteers listened to audios associated with visual information about their respective azimuths, (from –90° to 90°, spaced by 15°, including the edge points); (b) training: in which volunteers were instructed to associate seven audio clips with a set of seven azimuths: {–90°, –60°, –30°, 0°, 30°, 60°, 90°} – during this stage volunteers were allowed to listen to the audio clips and to change their choices as many times as they wanted before finishing the experiment; and (c) testing: in which the volunteers were instructed to listen to a set of randomly selected audios and to blindly assign them to an azimuth. The volunteers were allowed to listen to the audio clips as many times as they wanted before assigning them a value.

Unprocessed noise (RAW), as well as, MWF, MWF-ITF$_{\mathrm{v}}$, and R-MWF-ITF$_{\mathrm{v}}$ processed noise were obtained according to the acoustic scenario described in Section

3.3.2, for four different noise runs, and either 0 dB $\leqslant \bar{g}^2 \leqslant$ 30 dB or –5 dB $\leqslant \mathrm{SNR}_{\mathrm{in}} \leqslant$ 25 dB in steps of 5 dB [51, 52]. Additionally, five unprocessed noise-only audio clips related to the azimuths {–90°, –30°, 0°, 30°, 90°} were also employed, totaling 117 audio clips for each volunteer in each experiment.

3.4 RESULTS

In this section, the performance of the proposed method is assessed by both objective quality metrics and psychoacoustic experiments, considering either Lombard-gain or SNR variations. The acoustic scenario was created according to Sections 3.3.2 and 3.3.3. TABLE 1 lists the input-signal quality measures. The proposed method assumed the following arbitrary design requirements: $\mathrm{SNR}_{\mathrm{WORST}} = -5$ dB [51, 52] (which is estimated according to (81), but considering the worst expected SNR condition during operation); and a binaural cue preservation requirement for the noise source defined by $\Delta\mathrm{ILD}_{\mathrm{v}} < 2$ dB dB and $\Delta\mathrm{IPD}_{\mathrm{v}} < 2 \cdot 10^{-4}$, which resulted (approximately) in $\beta(\lambda, k) = 1$ for all k values. All volunteers were able to assign each noise signal to their correct azimuth in stage 2 (training) of the psychoacoustic experiment, showing adequate individual lateralization capacity.

Table 1 – INPUT QUALITY MEASURES

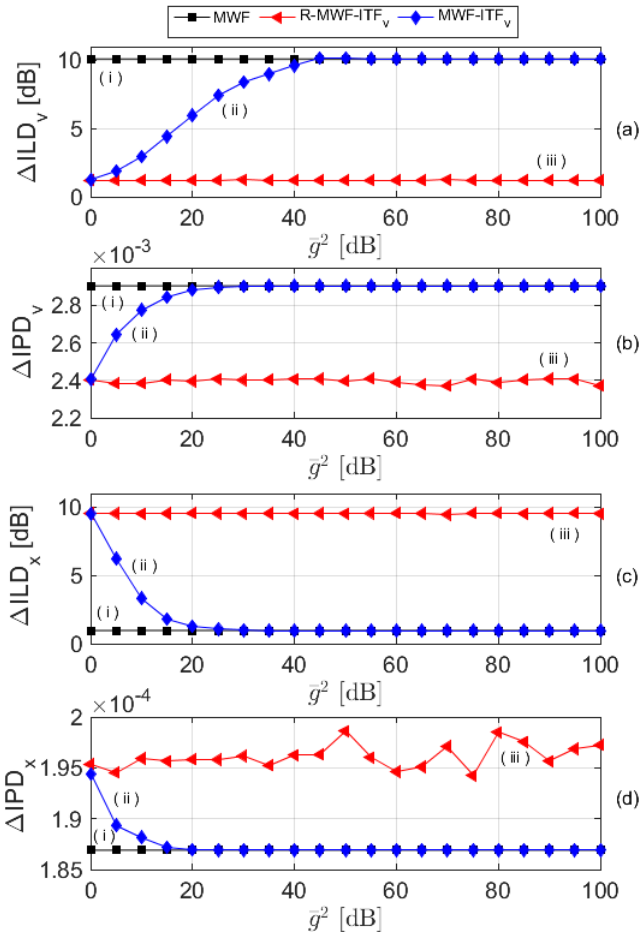
Measure	Average Value ± Standard Deviation
$\mathrm{PESQ}_{\mathrm{L}}$	1.05 ± 0.01
$\mathrm{PESQ}_{\mathrm{R}}$	1.07 ± 0.01
$\mathrm{SNR}_{\mathrm{L}}$ [dB]	-6.71 ± 0.01
$\mathrm{SNR}_{\mathrm{R}}$ [dB]	-2.60 ± 0.02
SNR [dB]	-5.00 ± 0.40

3.4.1 Robustness to the Lombard gain: numerical simulations

Figure 9 shows the speech and noise $\Delta\mathrm{ILD}_{\mathrm{v}}$ and $\Delta\mathrm{IPD}_{\mathrm{v}}$ values for MWF, MWF-$\mathrm{ITF}_{\mathrm{v}}$, and R-MWF-$\mathrm{ITF}_{\mathrm{v}}$ processed signals as a function of the average squared Lombard gain (see (82)). Noticeably, the performance of the conventional MWF-$\mathrm{ITF}_{\mathrm{v}}$ was significantly affected by variations in the Lombard gain. Figure 9 also shows that both MWF and R-MWF-$\mathrm{ITF}_{\mathrm{v}}$ were approximately invariant to changes in $\bar{g}^2$. Considering $\Delta\mathrm{ILD}_{\mathrm{v}}$ and $\Delta\mathrm{IPD}_{\mathrm{v}}$, respectively, as shown in Figure 9(a) and Figure 9(b), the MWF presented the highest noise distortion, whereas the R-MWF-$\mathrm{ITF}_{\mathrm{v}}$ presented the lowest. On the other hand, according to Figure 9(c) and Figure 9(d), the MWF resulted in the lowest speech binaural cue distortion, whereas the R-MWF-$\mathrm{ITF}_{\mathrm{v}}$ resulted in the highest (this is not surprising, because there are no penalties for spatial preservation of speech). Note, however, that the R-MWF-$\mathrm{ITF}_{\mathrm{v}}$ was not affected by $\bar{g}^2$. A similar behavior was verified for $\Delta\mathrm{SNR}_{\mathrm{L}}$ and $\Delta\mathrm{SNR}_{\mathrm{R}}$ presented, respectively, in Figure 10(a) and Figure 10(b), in

which SNR increases (up to 2.7 dB and 0.9 dB for the left and right ears, respectively) were obtained for the proposed method (in comparison with the MWF-ITF$_{\mathrm{v}}$) for large Lombard gains, resulting in the same MWF noise reduction performance.

Figure 9 – Averaged binaural-cue errors with respect to the averaged squared Lombard gain ($\overline{g}^2$) in decibels (dB) for: (i) MWF (black square), (ii) MWF-ITF$_{\mathrm{v}}$ (blue diamond), and (iii) R-MWF-ITF$_{\mathrm{v}}$ (red triangle). (a) ΔILD$_{\mathrm{v}}$ [dB]; (b) ΔIPD$_{\mathrm{v}}$; (c) ΔILD$_{\mathrm{x}}$ [dB]; (d) ΔIPD$_{\mathrm{x}}$.

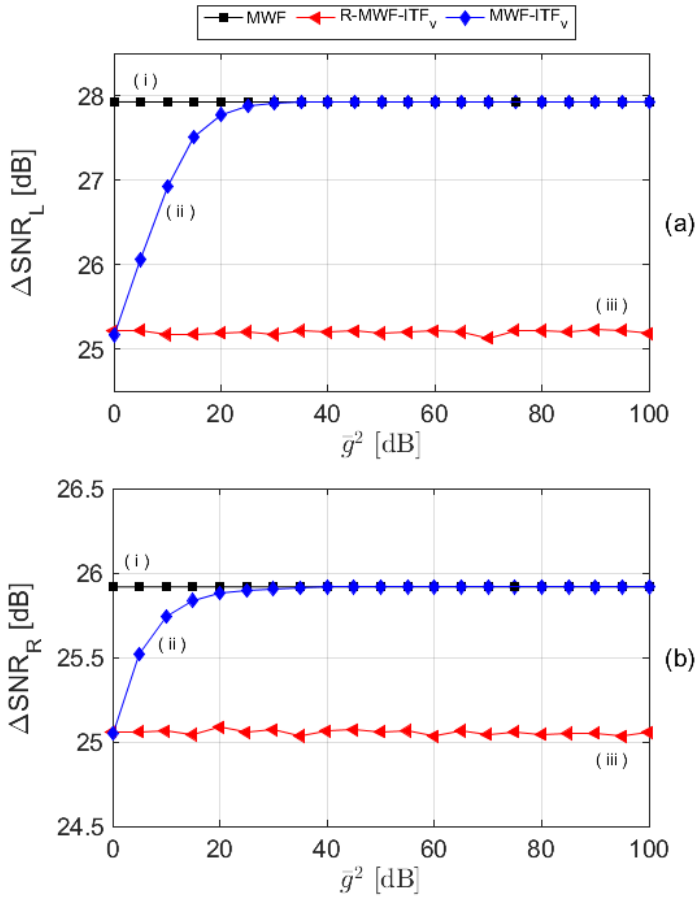


3.4.2 Robustness to the Lombard gain: psychoacoustic experiments

Figure 11(a) and Figure 11(b) present the average and median perceived azimuth of arrival for each noise reduction method and unprocessed noise, respectively, obtained in the psychoacoustic experiments with volunteers. Notably, the average/median perceived azimuth obtained from signals processed by the MWF-ITF$_{\mathrm{v}}$ was significantly affected by the Lombard gain.

Figure 11(c) presents box-and-whisker diagrams including all Lombard gains. The average perceived azimuth for the unprocessed (RAW) noise was –71° (median of –75°). In this sense, we verified that volunteers overestimated the true azimuth of the noise source located at –60° (left). As expected for signals processed by the MWF, the perceived noise

Figure 10 – Averaged input-output SNR differences with respect to the average squared Lombard gain ($\overline{g}^2$) in decibels (dB) for: (i) MWF (black square), (ii) MWF-ITF$_{\mathrm{v}}$ (blue diamond), and (iii) R-MWF-ITF$_{\mathrm{v}}$ (red triangle). (a) $\Delta\mathrm{SNR_L}$ [dB]; and (b) $\Delta\mathrm{SNR_R}$ [dB].

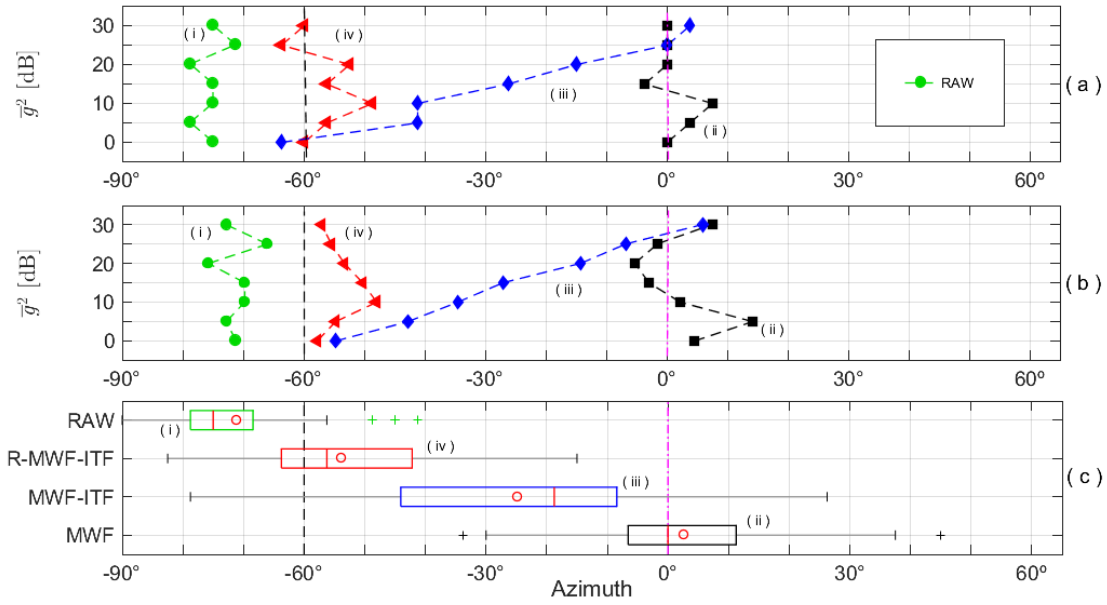


azimuth changed toward the speech source at azimuth of 0° (average of 3° and median of 0°). The MWF-ITF$_{\mathrm{v}}$ presented an average perceived azimuth of –25° (median of –19°), indicating a displacement toward the speech azimuth. This was especially true for large Lombard gains, when the cost function was largely unbalanced. The proposed R-MWF-ITF$_{\mathrm{v}}$ resulted in accurate estimates of the original noise azimuth with an average of –54° (median of –56°). Considering the average response of the volunteers for each audio, and that each $\overline{g}^2$ indicates a different condition in which the azimuth was measured, a one-way repeated measures ANOVA test was applied to analyze the three methods with respect to the perceived azimuth.

First, the Shapiro-Wilk (SW) test was applied to the data of each method to verify the hypothesis of Gaussianity. Its null hypothesis was defined as "H$_0$: the distributions are Gaussian" at the level of significance $p > 0.05$. H$_0$ was not reject for all groups. Mauchly's test was used to analyze the sphericity assumption of the ANOVA test. The null hypothesis of this test was "H$_0$: differences between all possible pairs are equal," and the level of significance was $p > 0.05$. The sphericity assumption was rejected, and the Greenhouse-

Geisser correction was applied. Because no extreme outliers were identified, the one-way repeated ANOVA, with the null hypothesis defined as "H_0: all distributions are the same, at the level of significance $p > 0.05$" was applied. As a result, the null hypothesis was rejected ($F(2,68) = 66.5$, $p < 0.0001$, $\eta^2 = 0.56$). Then, the Dunn-Bonferroni post-hoc test was applied to verify which pairs of distributions were different from each other. The null hypothesis was "H_0: the pairs of groups have the same distributions, at a level of significance $p > 0.05$". Finally, we concluded that all pairwise differences, between the methods were statistically different ($p \leqslant 0.05$).

Figure 11 – Psychoacoustic experiments for different average squared Lombard gains: (i) RAW (unprocessed) (green circle), (ii) MWF (black square), (iii) MWF-ITF$_v$ (blue diamond), (iv) R-MWF-ITF$_v$ (red triangle), true noise azimuth (black dashed line), and true speech azimuth (pink dotted-dashed line). (a) Average and (b) median perceived azimuths. (c) Box-and-whisker diagrams for all perceived azimuths grouped by technique (circles represent the average values).



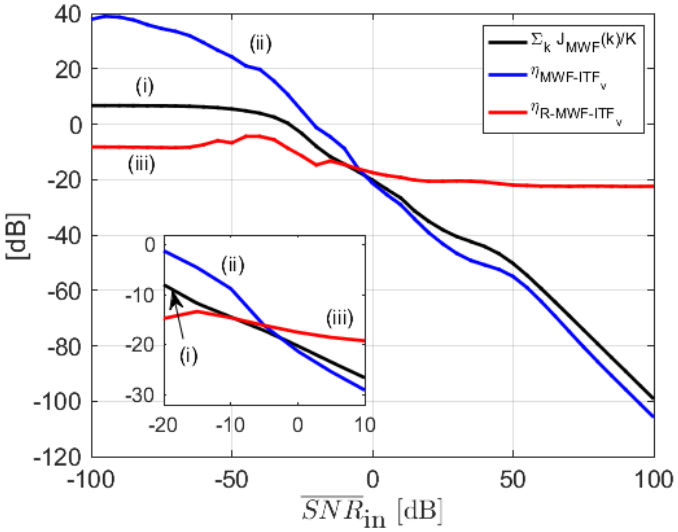
3.4.3 Robustness to SNR variations: numerical simulations

A preliminary computational simulation was performed to corroborate the theoretical hypothesis presented in Section 3.2.1, with respect to the robustness of equation (76) against $\mathrm{SNR}_{\mathrm{in}}$ variations.

Figure 12 highlights the influence of J_{MWF} in (76). It shows a comparison between $\sum_k J_{\mathrm{MWF}}(k)/K$ and the quantity $\eta = \sum_k r(k)/K$, as a function of $\mathrm{SNR}_{\mathrm{in}}$, where $r(k)$ is defined as the ratio of the following terms in (76):

$$r(k) = \frac{J_{\mathrm{MWF}}(k)}{\beta_1(k) J_1(k)} \tag{83}$$

Figure 12 – Simulation results for equation (83): (i)$\sum_k J_{\text{MWF}}(\lambda,k)/K$ (black); (ii) $\eta_{\text{MWF-ITF}_\text{v}}$ (blue); and (iii) $\eta_{\text{R-MWF-ITF}_\text{v}}$ (red). See the description before (83) for an explanation about η.



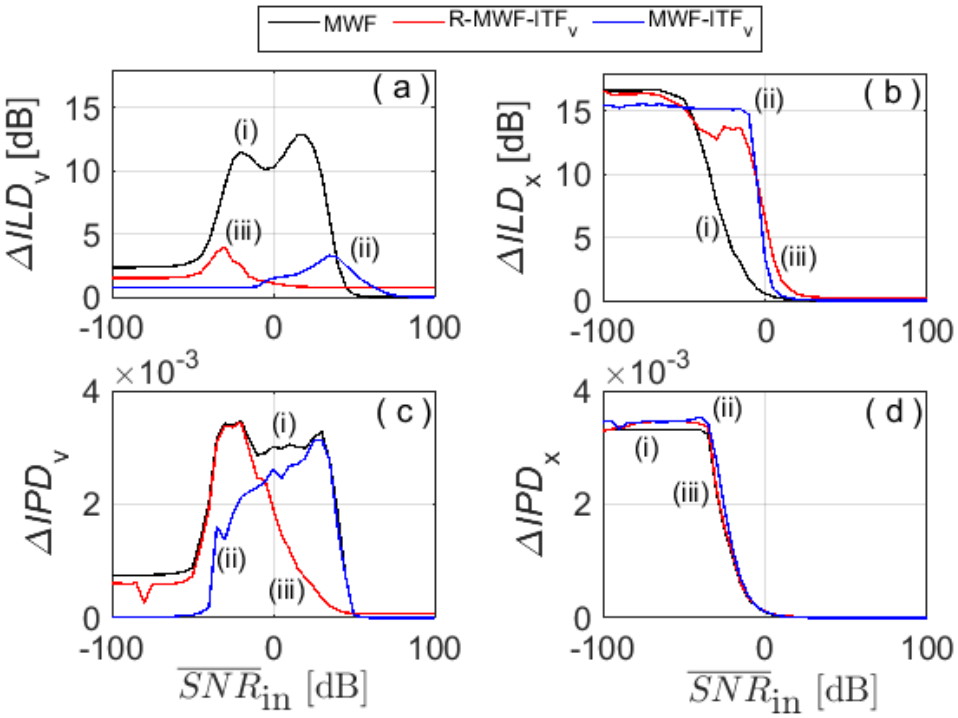
for $\beta_i \sum_{i=1}^{I} J_i(\mathbf{\Phi}_{\tilde{\mathbf{v}}})/K$ defined by $I = 1$, $\beta_1 \in \{g^{-2}, 1\}$. The quantities $\eta_{\text{MWF-ITF}_\text{v}}$ ($\beta_1 = g^{-2}$) and $\eta_{\text{R-MWF-ITF}_\text{v}}$ ($\beta_1 = 1$) are associated with the MWF-ITF and R-MWF-ITF, respectively. From Figure 12, it can be verified that for $\beta_1 = g^{-2}$ (MWF-ITF$_\text{v}$), η substantially decreases with increasing SNR$_{\text{in}}$. Specifically, for SNR$_{\text{in}} > 0$ dB, η decays logarithmically with SNR$_{\text{in}}$. On the other hand, for $\beta_1 = 1$ (R-MWF-ITF$_\text{v}$), the result of (83) is approximately constant for all ranges of SNR$_{\text{in}}$. From these observations it can be inferred that ensuring the preservation of the binaural noise cues for a given SNR condition ensures that they will also be preserved for higher SNRs.

Figure 13 shows the speech and noise ΔILD and ΔIPD as a function of SNR$_{\text{in}}$. Both the MWF-ITF$_\text{v}$ and R-MWF-ITF$_\text{v}$ methods have the same performance at SNR$_{\text{in}} = -5$ dB. Speech binaural cues for both MWF-ITF$_\text{v}$ and R-MWF-ITF$_\text{v}$ were similar and not very distinct from MWF results, indicating that the azimuth of the target source was approximately preserved.

3.4.4 Robustness to SNR variations: psychoacoustic experiments

Figure 14(a) and Figure 14(b) present the average and median azimuth perceived by volunteers for each noise reduction method and for unprocessed noise as a function of $\overline{\text{SNR}}_{\text{in}}$, respectively. Notably, the azimuth perceptions from signals processed by both MWF and MWF-ITF$_\text{v}$ were more affected by the $\overline{\text{SNR}}_{\text{in}}$ than the R-MWF-ITF$_\text{v}$. Figure 14(c) presents box-and-whisker diagrams of all perceived azimuths grouped by the processing method. The average (and median) perceived azimuth for the unprocessed noise (RAW) was –68°. In this sense, volunteers overestimated the real azimuth of the noise source, as verified in the psychoacoustic experiments in Section 3.3.3. As expected, the MWF-

Figure 13 – Influence of the SNR in the binaural noise cue preservation: (i) MWF (black); (ii) MWF-ITF$_v$ (blue); and (iii) R-MWF-ITF$_v$ (red). (a) ΔILD$_v$, (b) ΔILD$_x$, (c) ΔIPD$_v$, and (d) ΔIPD$_x$.

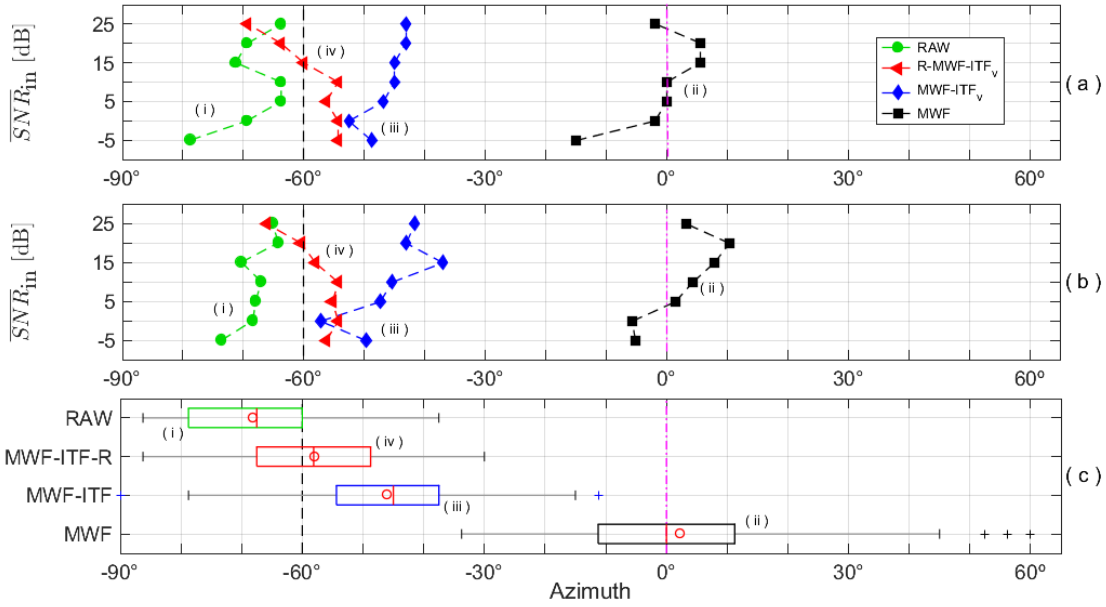


perceived azimuths were severely displaced to the 0° azimuth (average of 0° and median of 2°). The MWF-ITF$_v$ resulted in a perceived average azimuth of –46° (median of –45°), whereas R-MWF-ITF$_v$ provided an average (and median) of –58°. Unprocessed and processed noise by the MWF, MWF-ITF$_v$ and R-MWF-ITF$_v$ methods were analyzed with respect to their perceived azimuth performance, constituting four groups. No extreme outliers were identified in any group. Gaussianity was not rejected, but the assumption of sphericity was rejected. Thus, the Greenhouse-Geisser transformation was applied. The one-way repeated ANOVA test was applied, in which the null hypothesis was "H$_0$: all distributions are the same," at the level of significance $p > 0.05$. The null hypothesis was rejected. All pair comparisons rejected the null hypothesis using the Dunn-Bonferroni post-hoc test. Finally, we concluded that all pairwise differences between the methods were statistically different ($p \leqslant 0.05$).

3.5 DISCUSSION

Numerical simulations and psychoacoustic experiments with volunteers using fixed weighting parameters indicated that the performance of the MWF-ITF$_v$ binaural noise reduction method can be severely affected by speech/noise power variations. The proposed normalized weighting parameter and its design method aim to provide a robust operating setpoint, for the desired noise reduction effort and preservation of the original acoustic

Figure 14 – Psychoacoustic experiments for different $\overline{\mathrm{SNR}}_{\mathrm{in}}$ in: RAW (unprocessed) (green circle), MWF (black square), MWF-ITF$_{\mathrm{v}}$ (blue diamond), and R-MWF-ITF$_{\mathrm{v}}$ (red triangle), true noise azimuth (black dashed line), and true speech azimuth (pink dotted-dashed line). (a) Average and (b) median perceived azimuths. (c) Box-and-whisker diagrams for all perceived azimuths grouped by technique (circles represent the average values).



scenario.

In Section 3.1, the influence of both the Lombard effect and the SNR condition on augmented MWF-based cost functions was theoretically described. Numerical simulations with objective criteria indicate that the proposed method presented in Section 3.2 is robust to both the Lombard effect (Figure 9 and Figure 10) and input SNR variations (Figure 13).

Figure 9 and Figure 10 show that the performance of the MWF-ITF$_{\mathrm{v}}$ was significantly affected by the Lombard gain. Notably, increasing the Lombard gain increases both $\Delta\mathrm{SNR}_{\mathrm{L}}$ and $\Delta\mathrm{SNR}_{\mathrm{R}}$, as well as the preservation of the speech binaural cues. However, both $\Delta\mathrm{ILD}_{\mathrm{v}}$ and $\Delta\mathrm{IPD}_{\mathrm{v}}$ indicate progressive distortion. For $\bar{g}^2 > 45$ dB the MWF-ITF$_{\mathrm{v}}$ performance was equivalent to that of the conventional MWF, and the spatial preservation of the noise source was lost. As a result, the desired trade-off between noise reduction and preservation of binaural cues was no longer achieved. On the other hand, the R-MWF-ITF$_{\mathrm{v}}$ performance was approximately constant, maintaining the same binaural-cue preservation, independently of the whole range of assessed Lombard gains. The $\Delta\mathrm{ILD}_{\mathrm{x}}$ and $\Delta\mathrm{IPD}_{\mathrm{x}}$ increased for the R-MWF-ITF$_{\mathrm{v}}$, which was a side effect of not applying auxiliary penalty terms for the speech spatial preservation. The consequences associated with this arbitrary case study are alleviated in practical applications considering that, in the most common listening situations, the speaker of interest is in front of the hearing aid user ($\theta_{\mathrm{x}} = 0^{\circ}$)

[52, 67] (to allow for lip reading). In this situation, limited distortions in the perceived azimuth are not considered to be a major problem. Speech spatial preservation may also be included in both MWF-ITF and R-MWF-ITF approaches if required.

Figure 11 shows psychoacoustic experiments with volunteers, corroborating the observations presented in Figure 9(a) and Figure 9(b). Noticeably, changes in the Lombard gain significantly modify the perception of the noise azimuth in signals processed by the MWF-ITF$_{\mathrm{v}}$. The perceived azimuth ranges varied from –64° to 4° (with an average of –25°) for 0 dB $\leqslant \bar{g}^2 \leqslant$ 30 dB. The R-MWF-ITF$_{\mathrm{v}}$ provides the most accurate noise azimuth perception (varying from –64° to –49°, with an average of –54°), whereas the unprocessed-noise azimuth was overestimated by the volunteers (varying from –79° to –71°, resulting in an average of –71°).

Figure 12 corroborates the theoretical hypothesis that the proposed weighting parameter provides robustness to input SNR variations, showing bounded limits to the balance of both parts (noise reduction and spatial preservation) of the MWF-ITF$_{\mathrm{v}}$ technique.

Figure 13 indicates that both the MWF-ITF$_{\mathrm{v}}$ and the R-MWF-ITF$_{\mathrm{v}}$ present similar binaural cue variation along the analyzed range of $\overline{\mathrm{SNR}}_{\mathrm{in}}$. However, the largest amounts for each method occurred in different ranges. The MWF-ITF$_{\mathrm{v}}$ shows large binaural cue distortions for both $\Delta\mathrm{ILD}_{\mathrm{v}}$ and $\Delta\mathrm{IPD}_{\mathrm{v}}$ in the 5 dB $< \overline{\mathrm{SNR}}_{\mathrm{in}} <$ 55 dB range, whereas the R-MWF-ITF$_{\mathrm{v}}$ was only significantly affected in the –40 dB $< \overline{\mathrm{SNR}}_{\mathrm{in}} <$ –15 dB range. This observation agrees with the theoretical assumption that the robustness of the proposed method increases with $\overline{\mathrm{SNR}}_{\mathrm{in}} \rightarrow \infty$. Notably, according to [52] [68], the most common listening situations in which adults with hearing losses are exposed to, are in the 2 dB $< \overline{\mathrm{SNR}}_{\mathrm{in}} <$ 14 dB range. Thus, the proposed method provides better spatial preservation than the conventional MWF-ITF$_{\mathrm{v}}$ under practical conditions.

Finally, Figure 14 also corroborates both theoretical assumptions and numerical simulations with regard to robustness to input SNR variations. In the same way as in Figure 6, the unprocessed noise azimuth was overestimated by the volunteers (varying, in average, from –79° to –64°, for -5 dB $< \Delta\overline{\mathrm{SNR}}_{\mathrm{in}} <$ 30 dB), resulting in a global average/median of –68°. The MWF resulted in an azimuth range from –15° to 6° and a global average of 2° (median of 0°, which is the speech azimuth), while the MWF-ITF$_{\mathrm{v}}$ resulted in a range from –53° to –43°, and a global average of –46° (median of –45°). The R-MWF-ITF$_{\mathrm{v}}$ method provided the most accurate azimuth perception (varying from –69° to –54°, with a global average/median of –58°). The variation in the perceived azimuth in the psychoacoustic experiments was expected because its average error increases for more lateral angles [16, 69].

As expected for the fixed parameter strategy [34, 51], it was verified that the desired trade-off between noise reduction and binaural preservation was not preserved for different power levels and, as a result, new setpoints must be calculated. This behavior was observed

for the MWF-ITF$_{\mathrm{v}}$, but it is also valid to other MWF-based noise reduction methods, according to the theory developed in Section 3.1. Previously proposed psychoacoustical parameter optimization strategies for the MWF-IC under diffuse noise fields [30] [36] were based on exhaustive/iterative search procedures, which are computationally very intensive. Thus, the proposed method is the first one that is valid for both diffuse noise fields and point noise sources, while also being the first computationally efficient solution to this problem. Assuming that speech and noise second-order statistic estimation are accurate and the fixed part of the weight parameter was calculated according to Algorithm 1, then the proposed method ensures adequate robustness in the -5 dB $< \overline{\mathrm{SNR}}_{\mathrm{in}} <$ 30 dB range. Acoustic scenarios associated to high degrees of nonstationarity and large reverberation time must take into consideration an adequate (fast and accurate) choice of second-order statistic estimators, as well as MWF variations including dereverberation (such as in [34]) to avoid significant performance loss. Further studies must be made to assess the real-time performance of the proposed method in open-programming application-specific instruction-set processors.

3.6 CONCLUSION

This chapter proposed a normalized weighting parameter for MWF-based binaural noise reduction methods in hearing-aid applications. It aims to provide a robust preservation of spatial localization cues under speech and noise power variations. A design method was presented to guarantee that the desired operational setpoint, which establishes a trade-off between noise reduction and binaural-cue preservation, is maintained. An application example using the MWF-ITF$_{\mathrm{v}}$ noise reduction method (which can be easily expanded to the ILD, IPD, and IC cases) demonstrated that the proposed technique is robust to both the Lombard effect and input SNR variations. Simulation results with objective criteria as well as psychoacoustic experiments with volunteers corroborated the theoretical arguments, indicating that the proposed method maintained the desired trade-off between noise reduction and binaural-cue preservation, even under challenging conditions.

4 CLOSED-FORM SOLUTION TO THE MULTICHANNEL WIENER FILTER WITH INTERAURAL LEVEL DIFFERENCE PRESERVATION

Methods based on the Multichannel Wiener Filter with ILD preservation (MWF-ILD) generally employ cost functions comprised of two terms: i) the classical 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 [27, 35, 50]. 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 [27]. The work in [35] 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 adaptative filter implementation of the method presented in [35] was proposed in [50] 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.

In light of the presented facts, this work proposes a new MWF-ILD-based noise reduction method surpassing the abovementioned limitations. The contributions of the proposal in this chapter are the following: Firstly, the original (unconstrained) MWF-ILD optimization problem is changed to a constrained 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 a point speech source. 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 a very low computational complexity and without requiring any iterative optimization procedure. Fourthly, computer simulations are provided for an acoustic scenario comprised of one point speech source and one point interfering source, corroborating the effectiveness of the proposed method.

Results indicate that the proposed method achieves comparable or better performance than the conventional (unconstrained) MWF-ILD implementation, with a complexity similar to that of the classical MWF algorithm.

The remainder of this chapter is structured as follows. Section 4.1 describes the conventional MWF-ILD method, i.e., the unconstrained version of the problem. Section 4.2 presents the quadratic non-convex (QC-MWF-ILD) and a convex (CO-MWF-ILD) versions of the MWF-ILD problem. The CO-MWF-ILD equivalent dual problem and its semidefinite relaxation form (SD-MWF-ILD) are also shown in this section. In Section 4.3, a semi-analytical closed-form solution (CB-MWF-ILD) is derived. Section 4.4 describes the computational simulation setup, while in Section 4.5, results are presented and discussed. Finally, Section 4.6 presents the conclusions of this work.

4.1 MULTICHANNEL WIENER FILTER WITH INTERAURAL LEVEL DIFFERENCE PRESERVATION (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 [27, 35]

$$J_{\text{MWF-ILD}}(\lambda, k) = J_{\text{MWF}}(\lambda, k) + \alpha(\lambda, k) J^{\text{u}}_{\text{ILD}}(\lambda, k), \tag{84}$$

in which $\alpha(\lambda, k) \in \mathbb{R}_{++}$ is a time-frequency weighting parameter, which impacts the tradeoff between noise reduction and ILD preservation, and $J^{\text{u}}_{\text{ILD}}$ is the ILD penalty term, generically defined as [27, 35, 69]

$$J^{\text{u}}_{\text{ILD}}(\lambda, k) = [\varrho(\text{ILD}^{\text{u}}_{\text{ou}}(\lambda, k)) - \varrho(\text{ILD}^{\text{u}}_{\text{in}}(\lambda, k))]^2, \tag{85}$$

in which $\varrho(\bullet) = 10 \cdot \log_{10}(\bullet)$, and $\text{ILD}^{\text{u}}_{\text{in}}(\lambda, k)$ and $\text{ILD}^{\text{u}}_{\text{ou}}(\lambda, k)$ are the input and output ILDs of the interfering source, respectively defined by (22) and (23) making $d = u$, leading

to [27, 35]:

$$\mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}}(\lambda,k)=\frac{\mathbf{q}_{\mathrm{L}}^{\top}\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k)\mathbf{q}_{\mathrm{L}}}{\mathbf{q}_{\mathrm{R}}^{\top}\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k)\mathbf{q}_{\mathrm{R}}}, \tag{86}$$

and

$$\mathrm{ILD}_{\mathrm{ou}}^{\mathrm{u}}(\lambda,k)=\frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}(\lambda,k)\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k)\mathbf{w}_{\mathrm{L}}(\lambda,k)}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}(\lambda,k)\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k)\mathbf{w}_{\mathrm{R}}(\lambda,k)}, \tag{87}$$

in which $\mathbf{q}_{\mathrm{R}}^{\top}\boldsymbol{\Phi}_{\mathrm{u}}(\lambda,k)\mathbf{q}_{\mathrm{R}}>0$ 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 [27], corresponding to a first-order Taylor series approximation for $\varrho(\bullet)$, i.e., $\varrho(x)\cong x-1$. A more accurate approximation for (85) was proposed in [35] based on the inverse hyperbolic tangent function approximation of the logarithm, i.e., $\varrho(x)\cong(x-1)/(x+1)$. The optimum noise reduction filter that minimizes (84) can be obtained by solving the conventional (unconstrained) MWF-ILD problem:

$$\mathbf{w}_{\text{MWF-ILD}}(\lambda,k)=\underset{\mathbf{w}(\lambda,k)}{\arg.\min.}\ J_{\text{MWF-ILD}}(\lambda,k). \tag{88}$$

It has been shown that the optimal filter obtained from (88) provides adequate noise reduction as well as psychoacoustic spatial preservation for $\varrho(x)\cong x-1$ and $\varrho(x)\cong(x-1)/(x+1)$ approximations [27, 35, 50]. However, compared to the closed-form solution of the classical MWF in (46), the formulated minimization problem in (88) 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 $J_{\mathrm{ILD}}^{\mathrm{u}}(\lambda,k)$ is nonconvex and highly nonlinear. Consequently, usual approaches to solve (88) may result in considerable additional computational burden compared to (46), 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.

4.2 CONVEX FORMULATION OF THE MWF-ILD

In this section, an equivalent convex formulation for the conventional (unconstrained) MWF-ILD problem presented in (88) is proposed. It aims to provide tractable means for computing the globally optimal solution to the original problem.

4.2.1 Constrained MWF-ILD

An alternative ILD penalty term can be defined as:

$$J_{\text{ILD-REL}}(\mathbf{w})=\frac{J_{\mathrm{ILD}}^{\mathrm{u}}(\mathbf{w})}{(\varrho(\mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}}))^{2}}, \tag{89}$$

for $\varrho(\mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}}) \neq 0$, i.e., the interference source is not positioned directly in front of the HA user. The $\mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}} = 1$ scenario is not considered in this work because, in this case, the classic MWF will impose no binaural cue distortions. Unlike the $J_{\mathrm{ILD}}^{\mathrm{u}}$ penalty term presented in (85), $J_{\text{ILD-REL}}$ incorporates a scale factor that weights the input-output ILD error according to the input ILD. Thus, it measures relative ILD deviations.

Constraining (89) 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:

$$\mathbf{w}_{\text{CO-MWF-ILD}} = \arg.\min_{\mathbf{w}}.\ J_{\mathrm{MWF}}(\mathbf{w}) \tag{90a}$$

$$\text{s.t. } J_{\text{ILD-REL}}(\mathbf{w}) < \delta^2, \tag{90b}$$

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 (84) (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 (90b) in a mathematically tractable form, we apply the antilogarithm and rearrange the ratio, which results in two cases with two quadratic inequalities:

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

$$\mathbf{w}_{\mathrm{L}}^{\top}\mathbf{\Phi}_{\mathrm{u}}\mathbf{w}_{\mathrm{L}} - \delta_1 \cdot \mathbf{w}_{\mathrm{R}}^{\top}\mathbf{\Phi}_{\mathrm{u}}\mathbf{w}_{\mathrm{R}} \geqslant 0, \tag{91}$$

$$\mathbf{w}_{\mathrm{L}}^{\top}\mathbf{\Phi}_{\mathrm{u}}\mathbf{w}_{\mathrm{L}} - \delta_2 \cdot \mathbf{w}_{\mathrm{R}}^{\top}\mathbf{\Phi}_{\mathrm{u}}\mathbf{w}_{\mathrm{R}} \leqslant 0, \tag{92}$$

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

$$\mathbf{w}_{\mathrm{L}}^{\top}\mathbf{\Phi}_{\mathrm{u}}\mathbf{w}_{\mathrm{L}} - \delta_1 \cdot \mathbf{w}_{\mathrm{R}}^{\top}\mathbf{\Phi}_{\mathrm{u}}\mathbf{w}_{\mathrm{R}} \leqslant 0, \tag{93}$$

$$\mathbf{w}_{\mathrm{L}}^{\top}\mathbf{\Phi}_{\mathrm{u}}\mathbf{w}_{\mathrm{L}} - \delta_2 \cdot \mathbf{w}_{\mathrm{R}}^{\top}\mathbf{\Phi}_{\mathrm{u}}\mathbf{w}_{\mathrm{R}} \geqslant 0, \tag{94}$$

  in which

$$\delta_1 = (\mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}})^{1-\delta}, \text{ and } \delta_2 = (\mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}})^{1+\delta}. \tag{95}$$

We can generically represent (91)-(94) using (40), leading to:

$$(-1)^{t-1}\begin{bmatrix} -\mathbf{w}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{cc}}(\delta_1)\mathbf{w} \\ \mathbf{w}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{cc}}(\delta_2)\mathbf{w} \end{bmatrix} \leqslant \mathbf{0}_{2\times 1}, \tag{96}$$

in which

$$\mathbf{\Phi}_{\mathrm{cc}}(\kappa) = \begin{bmatrix} \mathbf{\Phi}_{\mathrm{u}} & \mathbf{0}_{M\times M} \\ \mathbf{0}_{M\times M} & -\kappa\mathbf{\Phi}_{\mathrm{u}} \end{bmatrix}, \tag{97}$$

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

$$t = \begin{cases} 1, & \mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}} > 1 \\ 2, & 0 < \mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}} < 1 \end{cases}. \tag{98}$$

Using (96), the constrained problem in (90) is equivalently represented as

$$\mathbf{w}_{\text{QC-MWF-ILD}} = \arg.\min_{\mathbf{w}}. \; J_{\mathrm{MWF}}(\mathbf{w}) \text{ s.t. } \mathbf{c}(\mathbf{w}) \leqslant \mathbf{0}_{2\times 1}, \tag{99}$$

where $\mathbf{c}(\mathbf{w})$ is defined as the left-hand side of (96), i.e.,

$$\mathbf{c}(\mathbf{w}) = (-1)^{t-1} \begin{bmatrix} -\mathbf{w}^{\mathrm{H}} \mathbf{\Phi}_{\mathrm{cc}}(\delta_1)\mathbf{w} \\ \mathbf{w}^{\mathrm{H}} \mathbf{\Phi}_{\mathrm{cc}}(\delta_2)\mathbf{w} \end{bmatrix}. \tag{100}$$

Both optimization problems in (90) and (99) 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 (88)) into the MWF-ILD with two quadratic constraints (QC-MWF-ILD) in (99) yields a straightforward form to control the ILD distortion. The problem defined in (99) is known as a quadratically constrained quadratic problem (QCQP), and is widely studied in the optimization literature [70–72].

4.2.2 Convex reformulation of the QC-MWF-ILD

The QC-MWF-ILD problem in (99) 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 (99) results in

$$J_{\mathcal{L}}(\mathbf{w}, \boldsymbol{\tau}) = J_{\mathrm{MWF}}(\mathbf{w}) + \boldsymbol{\tau}^{\top} \mathbf{c}(\mathbf{w}), \tag{101}$$

in which $\boldsymbol{\tau} = [\, \tau_1 \;\; \tau_2 \,]^{\top}$ are the the Lagrange multipliers. Substituting (39) and (100) in (101) leads to:

$$J_{\mathcal{L}}(\mathbf{w}, \boldsymbol{\tau}) = \mathbf{w}^{\mathrm{H}} \mathbf{\Phi}_{\mathrm{yc}}(\boldsymbol{\tau})\mathbf{w} - \mathbf{w}^{\mathrm{H}} \mathbf{p}_{\mathrm{xx}} - \mathbf{p}^{\mathrm{H}} \mathbf{w} + p_{\mathrm{xx}}, \tag{102}$$

in which $\mathbf{p}_{\mathrm{xx}}$ and p_{xx} are defined in (41) and (42), respectively, and the $\mathbf{\Phi}_{\mathrm{yc}}(\boldsymbol{\tau})$ matrix defined as:

$$\mathbf{\Phi}_{\mathrm{yc}}(\boldsymbol{\tau}) = \mathbf{\Phi}_{\mathrm{yy}} + (-1)^{t-1} \sum_{j=1}^{2} (-1)^{j} \tau_j \mathbf{\Phi}_{\mathrm{cc}}(\delta_j), \tag{103}$$

Using (18), (43) and (97) in (103) results in

$$\mathbf{\Phi}_{\mathrm{yc}}(\boldsymbol{\tau}) = \begin{bmatrix} \mathbf{\Phi}_{\mathrm{yu}}(\beta_1(\boldsymbol{\tau})) & \mathbf{0}_{M\times M} \\ \mathbf{0}_{M\times M} & \mathbf{\Phi}_{\mathrm{yu}}(\beta_2(\boldsymbol{\tau})) \end{bmatrix}, \tag{104}$$

in which:

$$\boldsymbol{\Phi}_{\mathrm{yu}}(\beta_r(\boldsymbol{\tau})) = \boldsymbol{\Phi}_{\mathrm{y}} + \beta_r(\boldsymbol{\tau})\boldsymbol{\Phi}_{\mathrm{u}}, \tag{105}$$

$$\beta_r(\boldsymbol{\tau}) = (-1)^{t-1}\boldsymbol{\delta}_r^{\top}\boldsymbol{\tau}, \tag{106}$$

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

From (102), the Lagrangian dual problem of the QC-MWFILD is defined as [70]:

$$J_{\mathcal{L}}(\mathbf{w}^{\mathrm{opt}}, \boldsymbol{\tau}^{\mathrm{opt}}) = \max_{\boldsymbol{\tau} \succcurlyeq 0} . \min_{\mathbf{w}} .\ J_{\mathcal{L}}(\mathbf{w}, \boldsymbol{\tau}), \tag{107}$$

in which $\mathbf{w}^{\mathrm{opt}}$ is the optimal solution and $\boldsymbol{\tau}^{\mathrm{opt}}$ denotes the optimal Lagrange multipliers. The dual problem, defined in (107), can be formulated for any constrained optimization problem and has two essential properties [70]: i) it is always a concave problem, whatever the form of the primal problem; and ii) for any $\boldsymbol{\tau} \succcurlyeq 0$ it defines a lower bound for the QC-MWF-ILD objective function, i.e.,

$$J_{\mathrm{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}}) \geqslant J_{\mathcal{L}}(\mathbf{w}^{\mathrm{opt}}, \boldsymbol{\tau}). \tag{108}$$

The inequality in (108) is not necessarily tight. However, considering that the QC-MWF-ILD in (99) is a QCQP (in complex variables) with two quadratic constraints, and assuming that it is strictly feasible then the optimal solution of both the primal problem $J_{\mathrm{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}})$ and the dual problem $J_{\mathcal{L}}(\mathbf{w}^{\mathrm{opt}}, \boldsymbol{\tau}^{\mathrm{opt}})$ are the same [71]. Thus, the inequality in (108) holds with equality, and $\mathbf{w}^{\mathrm{opt}} = \mathbf{w}_{\text{QC-MWF-ILD}}$ [70, 71], i.e.,

$$J_{\mathrm{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}}) = J_{\mathcal{L}}(\mathbf{w}_{\text{QC-MWF-ILD}}, \boldsymbol{\tau}^{\mathrm{opt}}). \tag{109}$$

The result in (109) is known as the strong duality property [70]. As a consequence, the solution to the QC-MWF-ILD problem in (99) can be computed precisely by (107), i.e.,

$$J_{\mathrm{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}}) = \max_{\boldsymbol{\tau} \succcurlyeq 0} . \min_{\mathbf{w}} .\ J_{\mathcal{L}}(\mathbf{w}, \boldsymbol{\tau}). \tag{110}$$

Since the Lagrangian is a quadratic function over $\mathbf{w}$, (110) 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 (102) with respect to $\mathbf{w}$ to zero, which leads to the following linear system:

$$\boldsymbol{\Phi}_{\mathrm{yc}}(\boldsymbol{\tau})\mathbf{w} = \mathbf{p}_{\mathrm{xx}}. \tag{111}$$

Assuming $\boldsymbol{\Phi}_{\mathrm{yc}}(\boldsymbol{\tau}) \succeq 0$, without loss of generality, an analytical solution to (111) is given by

$$\mathbf{w} = \boldsymbol{\Phi}_{\mathrm{yc}}^{\dagger}(\boldsymbol{\tau})\mathbf{p}_{\mathrm{xx}}. \tag{112}$$

Substituting (112) in (102) leads to

$$J_{\mathcal{L}}(\mathbf{w}^{\text{opt}}(\boldsymbol{\tau}), \boldsymbol{\tau}) = \min_{\mathbf{w}} . \; J_{\mathcal{L}}(\mathbf{w}, \boldsymbol{\tau}) \tag{113a}$$

$$= -\mathbf{p}_{\text{xx}}^{\text{H}} \boldsymbol{\Phi}_{\text{yc}}^{\dagger}(\boldsymbol{\tau}) \mathbf{p}_{\text{xx}} + p_{\text{xx}}. \tag{113b}$$

Without the optimal $\boldsymbol{\tau}$, the solution $\mathbf{w}^{\text{opt}}(\boldsymbol{\tau})$ in (112) 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 (113). This is obtained by introducing constraints $\boldsymbol{\tau} \geqslant 0$ 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 (110)) explicitly in the maximization problem, i.e.,

$$\boldsymbol{\tau}^{\text{opt}} = \arg.\max_{\boldsymbol{\tau}} . \; -\mathbf{p}_{\text{xx}}^{\text{H}} \boldsymbol{\Phi}_{\text{yc}}^{\dagger}(\boldsymbol{\tau}) \mathbf{p}_{\text{xx}} + p_{\text{xx}} \tag{114a}$$

$$\text{s.t. } \boldsymbol{\tau} \geqslant 0, \; \boldsymbol{\Phi}_{\text{yc}}(\boldsymbol{\tau}) \succeq 0. \tag{114b}$$

This problem can be written as a convex semidefinite (SD) programming (SDP) of the form (see [70] and [71] for details)

$$\boldsymbol{\tau}^{\text{opt}} = \arg.\min_{\boldsymbol{\tau}, \phi} . \; \phi \tag{115a}$$

$$\text{s.t. } \boldsymbol{\tau} \geqslant 0, \; \boldsymbol{\Phi}_{g}(\boldsymbol{\tau}) \succeq 0. \tag{115b}$$

in which ϕ is a slack variable, and matrix $\boldsymbol{\Phi}_{\text{g}}(\boldsymbol{\tau})$ is defined as:

$$\boldsymbol{\Phi}_{\text{g}}(\boldsymbol{\tau}) = \begin{bmatrix} \boldsymbol{\Phi}_{\text{yc}}(\kappa) & \mathbf{p}_{\text{xx}} \\ \mathbf{p}_{\text{xx}}^{\text{H}} & p_{\text{xx}} - \phi \end{bmatrix}. \tag{116}$$

Finally, replacing $\boldsymbol{\tau}^{\text{opt}}$ in (112) leads to:

$$\mathbf{w}_{\text{SD-MWF-ILD}}(\boldsymbol{\tau}^{\text{opt}}) = \boldsymbol{\Phi}_{\text{yc}}^{\dagger}(\boldsymbol{\tau}^{\text{opt}}) \mathbf{p}_{\text{xx}}. \tag{117}$$

As a consequence of the strong duality of the QCQP in (99), it is proven in [71] 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 (115a) (which is performed in polynomial time [70]) and replace its optimal solution in (117).

4.3 THE PROPOSAL: 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 (114). Firstly, the bidimensional dual problem in (114) is transformed into two univariate problems; then, the semidefinite constraint is converted into a set of linear constraints.

4.3.1 Reducing the dimensionality of the maximization problem

To reduce the dimensionality of the bidimensional maximization problem in (114), (101) is substituted in (109), leading to the following form for the strong duality property:

$$J_{\text{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}}) = J_{\text{MWF}}(\mathbf{w}_{\text{QC-MWF-ILD}}) + (\boldsymbol{\tau}^{\text{opt}})^{\top}\mathbf{c}(\mathbf{w}_{\text{QC-MWF-ILD}}), \tag{118}$$

which necessarily implies that:

$$(\boldsymbol{\tau}^{\text{opt}})^{\top}\mathbf{c}(\mathbf{w}_{\text{QC-MWF-ILD}}) = 0. \tag{119}$$

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 (119) can be represented as:

$$\tau_j^{\text{opt}}(-1)^{t+j-2}\mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}}\boldsymbol{\Phi}_{\text{cc}}(\delta_j)\mathbf{w}_{\text{QC-MWF-ILD}} = 0, \tag{120}$$

for $j \in \{1, 2\}$. The identity in (120) implies that (119) 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 (99) and since $\delta_j > 0$, the constraint in (99) implies that the feasible solution must satisfy one of the following three cases:

i) only the first constraint is active:

$$-(-1)^{t-1}\mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}}\boldsymbol{\Phi}_{\text{cc}}(\delta_1)\mathbf{w}_{\text{QC-MWF-ILD}} = 0, \tag{121}$$
$$(-1)^{t-1}\mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}}\boldsymbol{\Phi}_{\text{cc}}(\delta_2)\mathbf{w}_{\text{QC-MWF-ILD}} < 0, \tag{122}$$

ii) only the second constraint is active:

$$-(-1)^{t-1}\mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}}\boldsymbol{\Phi}_{\text{cc}}(\delta_1)\mathbf{w}_{\text{QC-MWF-ILD}} < 0, \tag{123}$$
$$-(-1)^{t-1}\mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}}\boldsymbol{\Phi}_{\text{cc}}(\delta_2)\mathbf{w}_{\text{QC-MWF-ILD}} = 0, \tag{124}$$

iii) both constraints are inactive:

$$(-1)^{t-1}\mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}}\boldsymbol{\Phi}_{\text{cc}}(\delta_1)\mathbf{w}_{\text{QC-MWF-ILD}} < 0, \tag{125}$$
$$(-1)^{t-1}\mathbf{w}_{\text{QC-MWF-ILD}}^{\text{H}}\boldsymbol{\Phi}_{\text{cc}}(\delta_2)\mathbf{w}_{\text{QC-MWF-ILD}} < 0, \tag{126}$$

From (121) to (126), 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: $\boldsymbol{\tau}_1 = [\,\tau_1\ 0\,]^{\top}$, associated with (121)-(122); or $\boldsymbol{\tau}_2 = [\,0\ \tau_2]^{\top}$, associated with (123)-(124). The Lagrange multiplier associated with (125)-(126) is a particular case of $\boldsymbol{\tau}_1$ or $\boldsymbol{\tau}_2$. Consequently, we can find the optimal solution to the problem

in (114) by equivalently solving two univariate optimization problems corresponding to each of the situations above. These problems can be formulated as:

$$\tau_j^{\text{opt}} = \underset{\tau_j}{\arg.\max.} \; -\mathbf{p}_{\text{xx}}^{\text{H}} \boldsymbol{\Phi}_{\text{yc}}^{\dagger}(\tau_j)\mathbf{p}_{\text{xx}} + p_{\text{xx}} \tag{127a}$$

$$\text{s.t. } \tau_j \geqslant 0, \;\; \boldsymbol{\Phi}_{\text{yc}}(\tau_j) \succeq 0. \tag{127b}$$

for $j \in \{1, 2\}$, in which (104), (105), and (106) turn into

$$\boldsymbol{\Phi}_{\text{yc}}(\tau_j) = \begin{bmatrix} \boldsymbol{\Phi}_{\text{yu}}(\beta_1(\tau_j)) & \mathbf{0}_{M\times M} \\ \mathbf{0}_{M\times M} & \boldsymbol{\Phi}_{\text{yu}}(\beta_2(\tau_j)) \end{bmatrix}, \tag{128}$$

in which:

$$\boldsymbol{\Phi}_{\text{yu}}(\beta_r(\tau_j)) = \boldsymbol{\Phi}_{\text{y}} + \beta_r(\tau_j)\boldsymbol{\Phi}_{\text{u}}, \tag{129}$$

$$\beta_r(\tau_j) = (-1)^{t+j+r}\delta_r^{r-1}\tau_j, \tag{130}$$

for $r \in \{1, 2\}$. Solving (127) for $j = 1$ and $j = 2$ leads to $\boldsymbol{\tau}_1^{\text{opt}}$ and $\boldsymbol{\tau}_2^{\text{opt}}$, respectively. The optimal solution to (114) is the one that maximizes the dual cost function, i.e.,

$$\boldsymbol{\tau}^{\text{opt}} = \arg.\max.\{\, J_{\mathcal{L}}(\boldsymbol{\tau}_1^{\text{opt}}),\; J_{\mathcal{L}}(\boldsymbol{\tau}_2^{\text{opt}})\}. \tag{131}$$

Finally, the solution $\mathbf{w}_{\text{QC-MWF-ILD}}$ is obtained by substituting the resulting $\boldsymbol{\tau}^{\text{opt}}$ in (117).

4.3.2 Simplification of the constraint

The positive semidefinite constraint of $\boldsymbol{\Phi}_{\text{yc}}(\tau_j)$ makes the solution to (127) 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 $\boldsymbol{\Phi}_{\text{yc}}(\delta_j)$, then $\boldsymbol{\chi}(\boldsymbol{\Phi}_{\text{yc}}(\delta_j)) = [\,\boldsymbol{\chi}^{\top}(\boldsymbol{\Phi}_{\text{yu}}(\beta_1(\tau_j)))\;\boldsymbol{\chi}^{\top}(\boldsymbol{\Phi}_{\text{yc}}(\beta_2(\tau_j)))\,]^{\top}$ comprises the eigenvalues of the diagonal blocks [66]. Thus, matrix $\boldsymbol{\Phi}_{\text{yc}}(\delta_j)$ is positive semidefinite if and only if $\boldsymbol{\chi}(\boldsymbol{\Phi}_{\text{yu}}(\beta_1(\tau_j))) \geqslant 0$ and $\boldsymbol{\chi}(\boldsymbol{\Phi}_{\text{yu}}(\beta_2(\tau_j))) \geqslant 0$, i.e.,

$$\boldsymbol{\Phi}_{\text{yc}}(\delta_j) \Longleftrightarrow \boldsymbol{\chi}(\boldsymbol{\Phi}_{\text{yu}}(\beta_r(\tau_j))) \geqslant 0, \tag{132}$$

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

Theorem I. Considering that $\boldsymbol{\Phi}_{\text{yu}}(\beta_r(\tau_j)) = \boldsymbol{\Phi}_{\text{y}} + \beta_r(\tau_j)\boldsymbol{\Phi}_{\text{u}}$, with $\boldsymbol{\Phi}_{\text{y}}$ defined in (18), $\boldsymbol{\Phi}_{\text{u}}$ defined in (17) and $\beta_r(\tau_j)$ defined in (130); if $\boldsymbol{\Phi}_{\text{y}}$ is a symmetric positive semidefinite matrix and $\boldsymbol{\Phi}_{\text{u}}$ is a symmetric rank-1 matrix, then the eigenvalues of $\boldsymbol{\Phi}_{\text{yu}}(\beta_r(\tau_j))$ are non-negative if and only if

$$(-1)^{t+j+r}\tau_j \geqslant -\eta_b^{-1}(\delta_j)^{1-r}, \tag{133}$$

for $r \in \{1,2\}$ and $\eta_{\mathrm{b}} = p_{su}\mathbf{b}^{\mathrm{H}}\boldsymbol{\Phi}_y^{\dagger}\mathbf{b} = \mathrm{trace}\{\boldsymbol{\Phi}_y^{\dagger}\boldsymbol{\Phi}_u\}$. Proof: See Appendix A.

Considering (133) and the constraint $\tau_j \geqslant 0$, the feasible region of (127) can be written as:

$$\{\tau_j \geqslant 0\} \cap \{(-1)^{t+j+r}\tau_j \geqslant -\eta_b^{-1}(\delta_j)^{1-r}\}, \tag{134}$$

for $r \in \{1,2\}$. Considering the cases for $r=1$ and $r=2$, the interval in (134) is given by

$$\{\tau_j\} \cap \{(-1)^{t+j}\tau_j \leqslant \eta_b^{-1}\} \cap \{(-1)^{t+j}\tau_j \geqslant -\eta_b^{-1}(\delta_j)^{-1}\}, \tag{135}$$

Considering (135), the feasible regions of (127) for $t=1$ and $t=2$ are defined, respectively, as

$$0 \leqslant \tau_j \leqslant \eta_b^{-1}(\delta_j)^{1-j}, \tag{136}$$

$$0 \leqslant \tau_j \leqslant \eta_b^{-1}(\delta_j)^{j-2}. \tag{137}$$

The intervals in (136) and (137) can be generically represented by

$$0 \leqslant \tau_j \leqslant \eta_b^{-1}(\delta_j)^{f}, \tag{138}$$

in which $f = (-1)^t(j-t)$. Using (138) in (127), leads to the following optimization problem:

$$\tau_j^{\mathrm{opt}} = \underset{\tau_j}{\arg.\max.} \; -\mathbf{p}_{\mathrm{xx}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{yc}}^{\dagger}(\tau_j)\mathbf{p}_{\mathrm{xx}} + p_{\mathrm{xx}} \tag{139a}$$

$$\text{s.t. } 0 \leqslant \tau_j \leqslant \eta_b^{-1}(\delta_j)^{f}. \tag{139b}$$

For each $j \in \{1,2\}$, the optimization problem described in (139) 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.

4.3.3 Rewriting the cost function of the dual problem

Considering the expressions for the pseudo-inverse of a block diagonal matrix such as $\boldsymbol{\Phi}_{\mathrm{yc}}$, and the sum of a positive semidefinite and a rank-1 matrix such as $\boldsymbol{\Phi}_{\mathrm{yu}}(\beta_r(\tau_r))$, given respectively in [73] and [74], the cost function of (139) is defined as

$$J_{\mathcal{L}}(\tau_j) = c_1\left[-\frac{(-1)^{j+t}\mathrm{ILD}_{\mathrm{in}}^{\mathrm{x}}\tau_j}{1-(-1)^{j+t}\eta_{\mathrm{b}}\tau_j} + \frac{(-1)^{j+t}\delta_j\tau_j}{1+(-1)^{j+t}\eta_{\mathrm{b}}\delta_j\tau_j}\right] + c_2, \tag{140}$$

in which $c_1 = p_{xR}|\eta_{ab}|^2$; $c_2 = (\mathrm{ILD}_{\mathrm{in}}^{\mathrm{x}})p_{xR}|\eta_a|^2$; $\eta_a = p_{su}\mathbf{b}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{y}}^{\dagger}\mathbf{b} = \mathrm{trace}(\boldsymbol{\Phi}_{\mathrm{y}}^{\dagger}\boldsymbol{\Phi}_{\mathrm{u}})$; $\eta_{ab} = p_{sx}p_{su}|\mathbf{a}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{y}}^{\dagger}\mathbf{b}| = \mathrm{trace}(\boldsymbol{\Phi}_{\mathrm{y}}^{\dagger}\boldsymbol{\Phi}_{\mathrm{x}}\boldsymbol{\Phi}_{\mathrm{y}}^{\dagger}\boldsymbol{\Phi}_{\mathrm{u}})$; $p_{xR} = \mathbf{q}_{\mathrm{R}}^{\top}\boldsymbol{\Phi}_{\mathrm{x}}\mathbf{q}_{\mathrm{R}}$. Therefore, the dual problem in (139) can be represented as

$$\tau_j^{\mathrm{opt}} = J_{\mathcal{L}}(\tau_j) \;\; \text{s.t.} \;\; 0 \leqslant \tau_j \leqslant \eta_b^{-1}(\delta_j)^{f}, \tag{141}$$

in which $J_{\mathcal{L}}(\tau_j)$ is defined in (140). There are two possibilities for τ_j in (141). Either it belongs to the interior of the feasible region $0 \leqslant \tau_j \leqslant \eta_b^{-1}(\delta_j)^f$, in which case it is a stationary point of (140), or it is one of the limits of the feasible interval, i.e., $\tau_j = 0$ or $\tau_{\text{opt}} = \eta_b^{-1}\delta_j^f$. Thus, (141) can be solved by computing the set of possible solutions and finding the one that maximizes the objective function.

4.3.4 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 (140), leading to:

$$\frac{\nabla J_{\mathcal{L}}(\tau_j)}{\nabla \tau_j} = -\frac{c_1(-1)^{j+t}\text{ILD}_{\text{in}}^{\text{x}}}{((-1)^{j+t} - \eta_{\text{b}}\tau_j)^2} + \frac{c_1(-1)^{j+t}\delta_j}{((-1)^{j+t} + \eta_{\text{b}}\delta_j\tau_j)^2}. \tag{142}$$

Equating (142) to zero results in two possible solutions:

$$\frac{(-1)^{j+t} + \eta_{\text{b}}\delta_j\tau_j}{(-1)^{j+t} - \eta_{\text{b}}\tau_j} = (-1)^{i-1}\psi_j, \tag{143}$$

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

$$\psi_j = \frac{\delta_j}{\text{ILD}_{in}^{\text{x}}}. \tag{144}$$

Therefore, the generic representation for the stationary points (sp) of the dual problem in (139) is given by:

$$\tau_{j,i}^{\text{sp}} = \frac{(-1)^{t+j+1}}{\eta_{\text{b}}}\frac{1 - (-1)^i\psi_j}{\delta_j - (-1)^i\psi_j}. \tag{145}$$

Note that (145) 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 (145) in the feasible region of (141) considering the values for t, j, and i that lead to the conditions for the stationary points to be feasible solutions, as shown in Table 2. Manipulating the inequalities for each entry of Table 2 leads to the equivalent interval representations shown in Table 3.

Since $\delta_j > 0$ (see (95)) and $\psi_j > 0$ (see (144)), the stationary points associated with column $i = 2$ in Table 3 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.,

$$\tau_{j,t}^{\text{ifr}} = \frac{(-1)^{t+j+1}}{\eta_{\text{b}}}\frac{1 - \psi_j}{\delta_j + \psi_j}. \tag{146}$$

Table 2 – Conditions for the stationary points in (145) to be contained in the feasible region of (141), for $t \in \{1,2\}$, $j \in \{1,2\}$, and $i \in \{1,2\}$.

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

Table 3 – Equivalent representation of the feasible regions in Table 2.

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

4.3.5 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 (141), $\tau_j = 0$ always leads to $J_{\mathcal{L}}(\tau_j) = c_2$. Also, as τ_j approaches $\eta_\mathrm{b}^{-1}\delta_j^f$ 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 $J_{\mathcal{L}}(0) > J_{\mathcal{L}}(\eta_\mathrm{b}^{-1}\delta_j^f)$. Therefore, the optimal Lagrange multiplier is zero ($\tau_j^{\mathrm{opt}} = 0$) when the stationary point of the cost function is outside the feasible region.

4.3.6 Implementation

Table 4 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 (117) and (146) and the reasoning presented in Section VI.E.

4.4 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 4. Comparisons with the classical MWF (in (48)) and the conventional MWF-ILD (in (88)) methods are also performed.

Table 4 – Pseudocode for the CB-MWF-ILD proposed method.

Input: $\mathbf{\Phi}_{\mathrm{y}}$, $\mathbf{\Phi}_{\mathrm{u}}$, $\mathbf{\Phi}_{\mathrm{x}}$, δ
Output: $\mathbf{w}_{\text{CB-MWF-ILD}}$

If $\mathrm{ILD}_{\mathrm{in}}^{\mathrm{u}} \neq 1$ **then**
Compute δ_1 and δ_2 using (95) and δ
Compute $ILD_{\mathrm{in}}^{\mathrm{x}}$ using $\mathbf{\Phi}_{\mathrm{x}}$ in (22)
Compute ψ_1 and ψ_2 using (144) and $ILD_{\mathrm{in}}^{\mathrm{x}}$
Initialize $\tau_1^{\mathrm{opt}} = 0$ and $\tau_2^{\mathrm{opt}} = 0$
If $t = 1$ **and** $\psi_1 \geqslant 1$ **then**
Compute $\tau_1^{\mathrm{opt}} = \tau_{1,1}^{\mathrm{ifr}}$ using (146) , for $t = 1$ and $j = 1$
end IF
If $t = 1$ **and** $\psi_2 \leqslant 1$ **then**
Compute $\tau_2^{\mathrm{opt}} = \tau_{2,1}^{\mathrm{ifr}}$ using (146) , for $t = 1$ and $j = 2$
end IF
If $t = 2$ **and** $\psi_1 \leqslant 1$ **then**
Compute $\tau_1^{\mathrm{opt}} = \tau_{1,2}^{\mathrm{ifr}}$ using (146) , for $t = 2$ and $j = 1$
end IF
If $t = 2$ **and** $\psi_2 \geqslant 1$ **then**
Compute $\tau_2^{\mathrm{opt}} = \tau_{2,2}^{\mathrm{ifr}}$ using (146) , for $t = 2$ and $j = 2$
end IF
If $J_{\mathcal{L}}(\tau_1^{\mathrm{opt}}) \geqslant J_{\mathcal{L}}(\tau_2^{\mathrm{opt}})$ **then**
$\boldsymbol{\tau} = [\ \tau_1^{\mathrm{opt}}\ \ 0\ ]^{\top}$
else
$\boldsymbol{\tau} = [\ 0\ \tau_2^{\mathrm{opt}}\ ]^{\top}$
end If
$\mathbf{w}_{\text{CB-MWF-ILD}} = \mathbf{\Phi}_{\mathrm{yc}}^{\dagger}(\tau^{\mathrm{opt}})\mathbf{p}_{\mathrm{xx}}$
else
$\mathbf{w}_{\mathrm{MWF}} = \mathbf{\Phi}_{\mathrm{yy}}^{\dagger}(\tau^{\mathrm{opt}})\mathbf{p}_{\mathrm{xx}}$
end If

4.4.1 Acoustic scenario

The acoustic scenario comprises two point sources (the desired speech and the interfering source) and additive (environmental and electrical) noise. Point 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 $\mathrm{T}_{60} \cong 300$[1] ms [63]. Thirty different speech signals were selected from the database presented in [64]. 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.

[1] T_{60}: is a measure of the reverberation time, i.e., the time required for the acoustic impulse response of the environment decay by 60 dB.

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. The elevation angle and the radial distance of both point sources 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 [75] and [46].

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.

4.4.2 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 [76]. 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 [46, 45, 77].

The methodology used to estimate the coherence matrices was the same used in [60], [78], and [79]. We assume the use of a (own) voice activity detector [80] for segmenting background noise, interference plus noise, and speech plus noise epochs to estimate $\mathbf{\Phi}_n$; $\mathbf{\Phi}_\mathrm{un} = \mathbf{\Phi}_\mathrm{u} + \mathbf{\Phi}_n$ and $\mathbf{\Phi}_\mathrm{xn} = \mathbf{\Phi}_\mathrm{x} + \mathbf{\Phi}_n$ using:

$$\hat{\mathbf{\Phi}}_\mathrm{d}(k) = \frac{1}{P_d} \sum_{\lambda=1}^{P_d} \mathbf{d}(\lambda, k)\mathbf{d}^\mathrm{H}(\lambda, k), \tag{147}$$

in which $d \in \{n, un, xn\}$; and P_d is the number of frames used in the estimative of the coherence matrices.

Estimation of $\mathbf{\Phi}_\mathrm{x}$ and $\mathbf{\Phi}_\mathrm{u}$ was performed using the covariance whitening method [60, 78, 79, 81, 82]. Firstly, a squared-root decomposition of $\hat{\mathbf{\Phi}}_n$ is performed, resulting in

$$\hat{\mathbf{\Phi}}_\mathrm{n} = \hat{\mathbf{\Phi}}_\mathrm{n}^{1/2}(k)\hat{\mathbf{\Phi}}_\mathrm{n}^{\mathrm{H}/2}(k). \tag{148}$$

Then, matrices $\mathbf{\Phi}_{\mathrm{xn}}$ and $\hat{\mathbf{\Phi}}_{\mathrm{un}}$ are prewhitened using (148) , i.e.,

$$\hat{\mathbf{\Phi}}^{\mathrm{w}}_{\mathrm{xn}}(k) = \hat{\mathbf{\Phi}}^{-1/2}_{\mathrm{n}}(k)\hat{\mathbf{\Phi}}_{\mathrm{xn}}(k)\hat{\mathbf{\Phi}}^{-\mathrm{H}/2}_{\mathrm{n}}(k), \tag{149}$$

$$\hat{\mathbf{\Phi}}^{\mathrm{w}}_{\mathrm{un}}(k) = \hat{\mathbf{\Phi}}^{-1/2}_{\mathrm{n}}(k)\hat{\mathbf{\Phi}}_{\mathrm{un}}(k)\hat{\mathbf{\Phi}}^{-\mathrm{H}/2}_{\mathrm{n}}(k). \tag{150}$$

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

$$\hat{\mathbf{\Phi}}_{\mathrm{x}}(k) = \hat{p}_{sx}(k)\hat{\mathbf{a}}(k)\hat{\mathbf{a}}^{\mathrm{H}}(k), \tag{151}$$

$$\hat{\mathbf{\Phi}}_{\mathrm{u}}(k) = \hat{p}_{sx}(k)\hat{\mathbf{b}}(k)\hat{\mathbf{b}}^{\mathrm{H}}(k), \tag{152}$$

in which

$$\hat{p}_{sx}(k) = \nu\{\hat{\mathbf{\Phi}}^{\mathrm{w}}_{\mathrm{xn}}(k)\}, \tag{153}$$

$$\hat{p}_{su}(k) = \nu\{\hat{\mathbf{\Phi}}^{\mathrm{w}}_{\mathrm{un}}(k)\}, \tag{154}$$

$$\hat{\mathbf{a}}(k) = \hat{\mathbf{\Phi}}_{\mathrm{n}}(k)\mathbf{f}\{\hat{\mathbf{\Phi}}^{\mathrm{w}}_{\mathrm{xn}}(k)\}, \tag{155}$$

$$\hat{\mathbf{b}}(k) = \hat{\mathbf{\Phi}}_{\mathrm{n}}(k)\mathbf{f}\{\hat{\mathbf{\Phi}}^{\mathrm{w}}_{\mathrm{un}}(k)\}, \tag{156}$$

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 ($\mathbf{\Phi}_{\mathrm{x}}$) and interference ($\mathbf{\Phi}_{\mathrm{u}}$) coherence matrix estimation in low-SNR complex acoustic scenarios [83].

4.4.3 Parameters and optimization algorithms

The 'fmincon' routine from Matlab [84] was employed to obtain the optimal solution for the conventional MWF-ILD method defined in (88). Both gradient and Hessian (see Appendix B) for equation (84) were provided to the solver. Parameters $\alpha(\lambda, k)$ were defined through the iterative algorithm presented in Appendix C and kept fixed for all λ. Vector $\boldsymbol{\alpha} \in \mathbb{R}^{G}_{++}$ 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 [46]. The ILD tolerance $\delta(\lambda, k)$ was set to 10^{-6}.

4.4.4 Objective performance measures

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

The signal-to-interference plus noise ratio variation (ΔSINR), the signal-to-interference

ratio variation (ΔSIR), and the signal-to-noise ratio variation (ΔSNR) are defined as [60]

$$\Delta\text{SINR} = \frac{10}{K}\sum_{k=1}^{K}\log_{10}\left(\frac{p_{\text{zx}_{\text{LR}}}(k)p_{\text{v}_{\text{LR}}}(k)}{p_{\text{zv}_{\text{LR}}}(k)p_{\text{x}_{\text{LR}}}(k)}\right), \tag{157}$$

$$\Delta\text{SIR} = \frac{10}{K}\sum_{k=1}^{K}\log_{10}\left(\frac{p_{\text{zx}_{\text{LR}}}(k)p_{\text{u}_{\text{LR}}}(k)}{p_{zu_{\text{LR}}}(k)p_{\text{x}_{\text{LR}}}(k)}\right), \tag{158}$$

$$\Delta\text{SNR} = \frac{10}{K}\sum_{k=1}^{K}\log_{10}\left(\frac{p_{\text{zx}_{\text{LR}}}(k)p_{\text{n}_{\text{LR}}}(k)}{p_{\text{zn}_{\text{LR}}}(k)p_{\text{x}_{\text{LR}}}(k)}\right), \tag{159}$$

in which $p_{\text{zd}_{\text{LR}}}(k) = \mathbf{w}_{\text{L}}^{\text{H}}(k)\mathbf{\Phi}_{\text{d}}(k)\mathbf{w}_{\text{R}}(k)$; $p_{\text{d}_{\text{LR}}}(k) = \mathbf{q}_{\text{L}}^{\top}(k)\mathbf{\Phi}_{\text{d}}(k)\mathbf{q}_{\text{R}}(k)$; and $\text{d} \in \{\text{n, u, x}\}$. The ILD and IPD variations, which measure the input-output binaural cue preservation, are defined respectively as [35, 46]

$$\Delta\text{ILD}_{\text{d}} = \frac{10}{K}\sum_{k=1}^{K}\left|\log_{10}\left(\frac{p_{\text{zd}_{\text{L}}}(k)p_{d_{\text{R}}}(k)}{p_{\text{zd}_{\text{R}}}(k)p_{d_{\text{L}}}(k)}\right)\right|, \tag{160}$$

$$\Delta\text{IPD}_{\text{d}} = \frac{1}{K\pi}\sum_{k=1}^{K}\left|\angle p_{\text{z}d_{\text{LR}}}(k) - \angle p_{d_{\text{LR}}}(k)\right|, \tag{161}$$

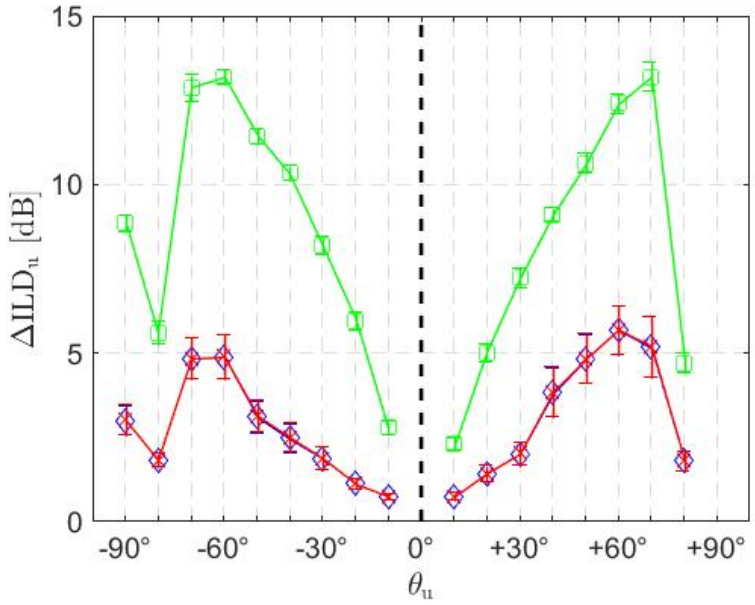
in which $p_{\text{zd}_\ell}(k) = \mathbf{w}_\ell^{\text{H}}\mathbf{\Phi}_{\text{d}}\mathbf{w}_\ell$; $p_{\text{d}_\ell}(k) = \mathbf{q}_\ell^{\top}\mathbf{\Phi}_{\text{d}}\mathbf{q}_\ell$. The global performance for each noise reduction technique was calculated for all signals and interference azimuths (285 values).

4.5 RESULTS AND DISCUSSION

This section presents performance results for the proposed CB-MWF-ILD method and compares it with the classical MWF and the conventional MWF-ILD. Figure 15 and Figure 16 depict input-output ILD variation for the interference and speech signals, respectively. As expected, the classical MWF severely distorts the ILD of the interfering source compared to the MWF-ILD techniques. This observation is corroborated by the $\Delta\text{ILD}_{\text{u}}$ (for all θ_{u}) presented in Table 5 , in which: $\Delta\text{ILD}_{\text{u}} = 8.4$ dB, $\Delta\text{ILD}_{\text{u}} = 2.5$ dB, and $\Delta\text{ILD}_{\text{u}} = 2.5$ dB for the MWF, MWF-ILD, and CB-MWF-ILD, respectively. The proposed method and the MWF-ILD result in similar ILD preservation for interference. The CB-MWF-ILD results in the same speech ILD distortion as the MWF-ILD and approximately the same as the classical MWF.

Figure 17 and Figure 18 show $\Delta\text{IPD}_{\text{u}}$ and $\Delta\text{IPD}_{\text{x}}$. The speech ΔIPD of the proposed method is very small and is in the same range as the MWF and MWF-ILD methods. The ΔIPD of the interfering source is not negligible, despite being 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 because of uncoherent ILD and IPD binaural cues. However, this situation was not observed in the processed signals. Table 5 also indicates that despite some variations, $\Delta\text{IPD}_{\text{u}}$ and $\Delta\text{IPD}_{\text{x}}$

Figure 15 – Interference ΔILD calculated from 15 input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◇), and CB-MWF-ILD in red (×).



for all interference azimuths and methods are in the same range. Figure 19, Figure 20, and Figure (21) show ΔSINR, ΔSIR, and ΔSNR for distinct $\theta_u \neq \theta_x = 0°$. The proposed method presents a noise reduction performance equivalent to the classic MWF and the conventional MWF-ILD for all interference azimuths (Table 6).

The dissimilarity between noise reduction performances for both left and right sides is a consequence of the input/output ILD definitions presented in (22) and (23), as described in [35].

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 12,000 times faster than for the conventional (unconstrained) MWF-ILD and only 0.35 times slower than the classical MWF, on average (see Table 7).

Table 5 – 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	8.4 ± 3.4	0.5 ± 0.1	$0.5 \pm 5 \times 10^{-2}$	$2 \times 10^{-2} \pm 5 \times 10^{-3}$
MWF-ILD	2.5 ± 1.6	0.5 ± 0.1	$0.5 \pm 5 \times 10^{-2}$	$2 \times 10^{-2} \pm 5 \times 10^{-3}$
CB-MWF-ILD	2.5 ± 1.6	0.5 ± 0.1	$0.5 \pm 5 \times 10^{-2}$	$2 \times 10^{-2} \pm 5 \times 10^{-3}$

Figure 16 – Speech ΔILD calculated from 15 input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◇), and CB-MWF-ILD in red (×).

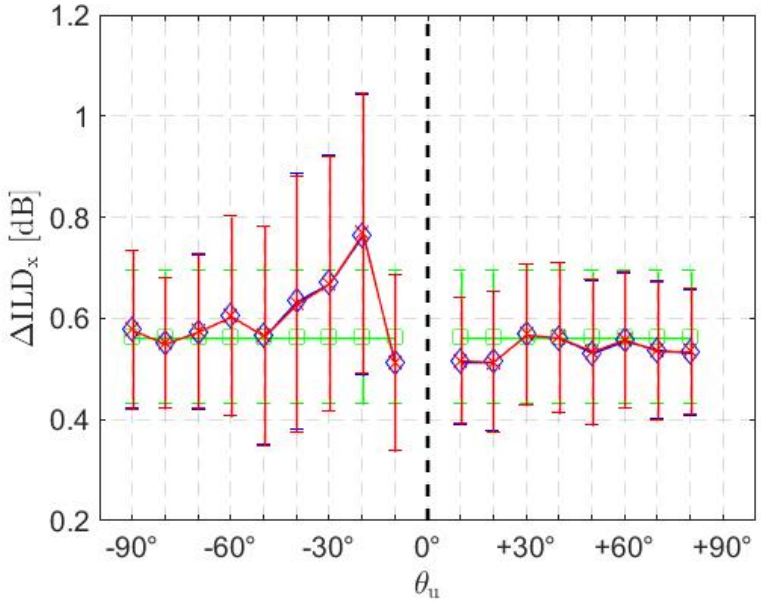


Figure 17 – ΔIPD of the interfering source calculated from 15 input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◇), and CB-MWF-ILD in red (×).

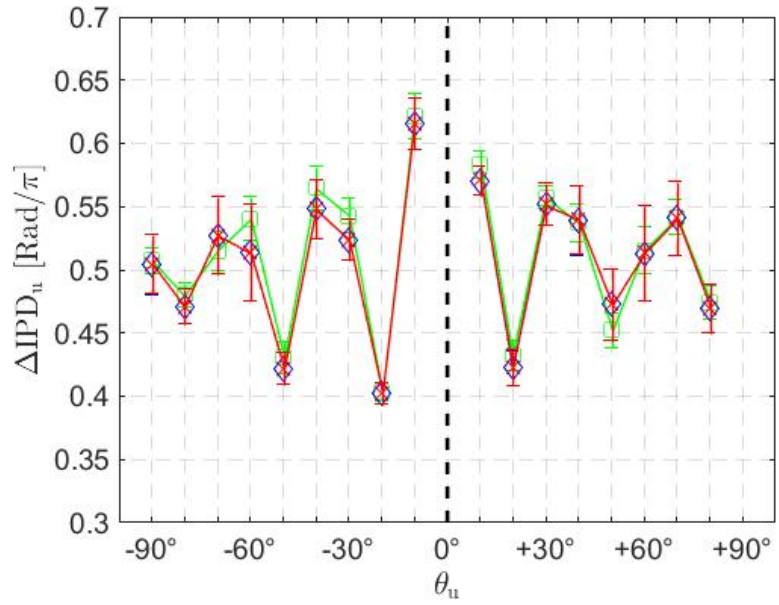


Figure 18 – Speech ΔIPD calculated from 15 input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◇), and CB-MWF-ILD in red (×).

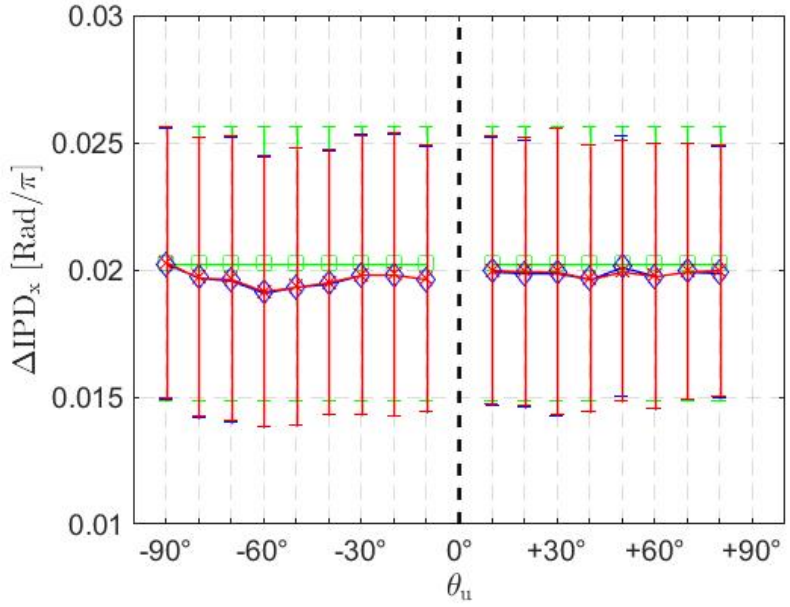


Figure 19 – ΔSINR calculated from 15 input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◇), and CB-MWF-ILD in red (×).

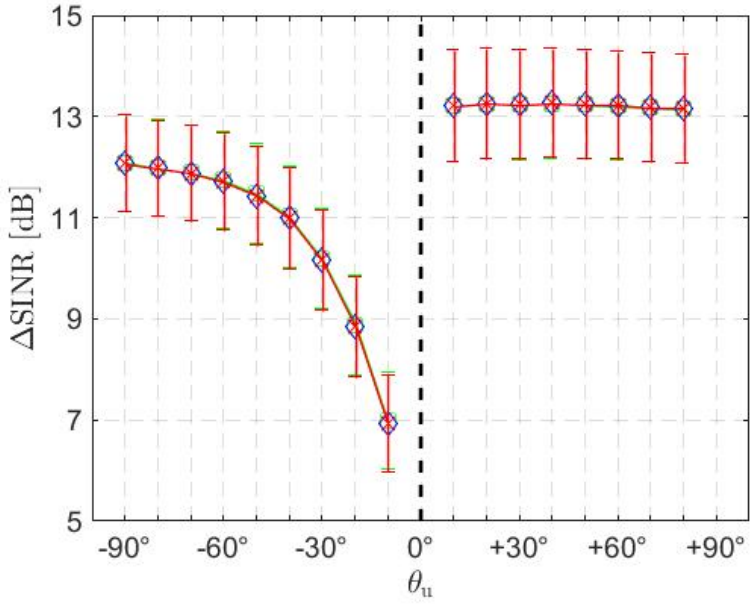


Figure 20 – ΔSIR calculated from 15 input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◇), and CB-MWF-ILD in red (×).

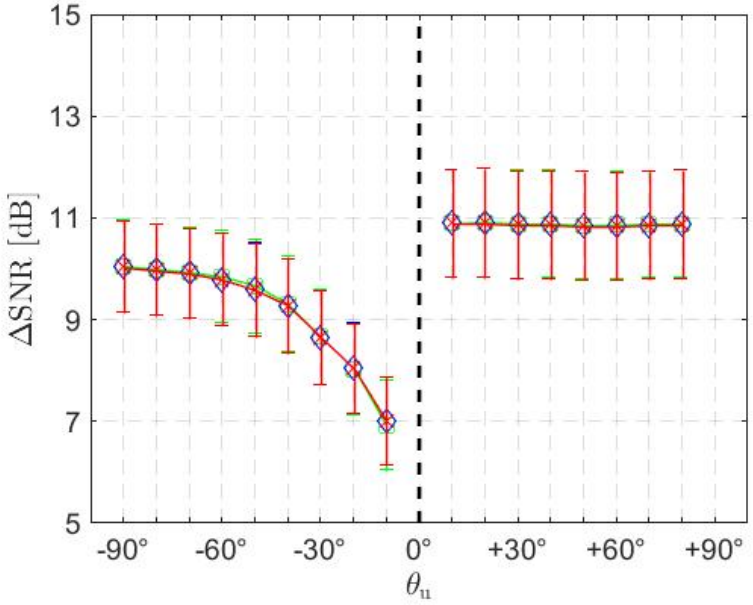


Figure 21 – ΔSNR calculated from 15 input signals for each $\theta_u \neq \theta_x = 0°$. MWF in green (□), MWF-ILD in blue (◇), and CB-MWF-ILD in red (×).

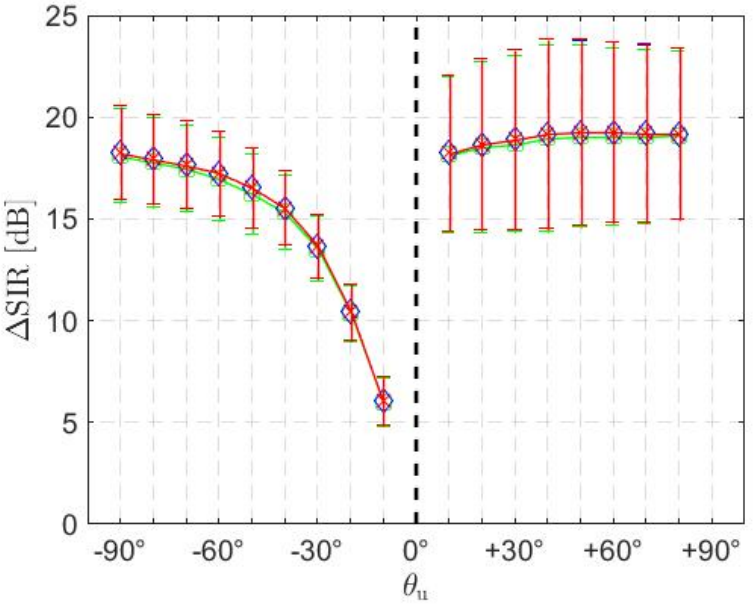


Table 6 – Global median and standard deviation ($M \pm \Sigma$) for ΔSINR, ΔSIR and ΔSNR.

	ΔSINR [dB]	ΔSIR [dB]	ΔSNR [dB]
MWF	12.2 ± 2.0	16.8 ± 5.0	10.1 ± 1.4
MWF-ILD	12.2 ± 2.0	16.8 ± 5.0	10.1 ± 1.4
CB-MWF-ILD	12.2 ± 2.0	16.8 ± 5.0	10.1 ± 1.4

Table 7 – Average time for calculating the optimal filters ($\mathbf{w}_{\mathrm{L}}$ and $\mathbf{w}_{\mathrm{R}}$) in the assessed acoustic scenario.

	MWF	MWF-ILD	CB-MWF-ILD
Time [s]	7×10^{-5}	2.4	2×10^{-4}

4.6 CONCLUSIONS

This chapter 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 classical MWF and an ILD preservation comparable to the conventional MWF-ILD. The proposed semi-analytical closed-form leads to a considerable decrease (up to 12,000 times less) in the optimization time compared to the conventional MWF-ILD method, which makes this method very attractive for binaural hearing aid applications.

5 MULTICHANNEL WIENER FILTER TECHNIQUES FOR PRESERVATION OF POINTWISE SOURCES AND DIFFUSE SOUND FIELDS

This chapter discusses the limitations of current binaural noise reduction (NR) techniques and proposes a new approach to address these limitations.

The preservation of binaural sound impression is associated with two types of signals: signals generated by a pointwise source or signals generated by diffuse fields. Current state-of-the-art MWF-based binaural noise reduction (NR) techniques are available to preserve the binaural impression of one of those signals.

The reason for this limitation is that the binaural preservation terms used in these techniques are typically associated with the preservation of one or two main binaural cues, i.e., the interaural level difference (ILD), interaural phase difference (IPD), and interaural coherence (IC). As a result, MWF-based binaural NR techniques are effective only in preserving the binaural impression of signals generated from either a pointwise source or a diffuse field, but not both.

In order to tackle this issue with MWF-based techniques, this chapter presents three techniques for preservation of a pointwise source or a diffuse field. The goal is to jointly preserve ILD, ITD/IPD, and IC using a single approach.

This chapter is divided into two parts. The first deals with the proposition of a new cost function for binaural hearing aids, comprehending Sections 5.1 to 5.4; while the second describes an efficient form to implement it. In Section 5.1, we investigate the abilities of binaural cost functions to preserve the three main binaural cues. Section 5.2 presents a cost function that aims to preserve the ILD, ITD/IPD, and IC. In Section 5.3, we investigate the noise reduction and binaural preservation performances of the proposed cost function. In Section 5.4, we present and discuss the results of our experiments. Sections 5.5 and 5.6 present the derivation of the proposed implementation algorithm. Section 5.7 presents the experimental setup to assess the performance of constrained optimization techniques. Section 5.8 presents the results and discussion. Finally, in Section 5.9, we present our conclusion for this chapter.

5.1 AN INVESTIGATION OF THE COST FUNCTIONS FOR PRESERVATION OF THE BINAURAL CUES

This section focuses on challenges associated with preserving binaural cues in signals generated from pointwise sources and diffuse fields. Despite a significant amount of interest devoted to designing perceptually relevant binaural preservation terms, currently there is no unique preservation term that can effectively maintain the binaural cues of such signals. Specifically, typical binaural cost functions used to preserve interaural level differences (ILD), interaural phase differences (IPD), and interaural coherence (IC) are only used for signals originating from pointwise sources and diffuse fields, not both. In the

following paragraphs, we will describe the specific set of binaural cues preserved by the binaural preservation terms.

Firstly, let us analyze the $J^{\mathrm{d}}_{\mathrm{ILD-ITD}}(\mathbf{w})$ cost function presented in (50). This term is added to the MWF cost function (J_{MWF}) in order to preserve ILD and ITD/IPD in the case of pointwise sources. The auxiliary cost function is a weighted sum of two separate terms: one for the ILD and another for the ITD/IPD preservation. However, it has been shown in [32] that the term associated with the preservation of ITD / IPD in (50) has an unsatisfactory psychoacoustic performance for pointwise noise sources.

Secondly, let us analyze the ITF cost function, $J^{\mathrm{d}}_{\mathrm{ITF}}(\mathbf{w})$ defined in (54). The purpose of this cost function is to preserve the ITF of a pointwise source. Consequently, it also ensures that the ILD and IPD of signals produced by such source are preserved.

To provide a clear understanding of this claim, we can manipulate both the denominator and the numerator of (54). By doing so, we arrive at the following result:

$$J^{\mathrm{d}}_{\mathrm{ITF}}(\mathbf{w}) = \frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}} - (\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}})^{*}\cdot\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}} - (\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}})\cdot\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}} + |\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}}|^{2}\cdot\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}} \tag{162}$$

$$= \frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}} - (\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}})^{*}\cdot\frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}} - (\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}})\cdot\frac{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}} + |\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}}|^{2}\cdot\frac{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}}. \tag{163}$$

By substituting $\mathrm{ILD}^{\mathrm{d}}_{\mathrm{ou}}$, which is defined in (23), and $\mathrm{ITF}^{\mathrm{d}}_{\mathrm{ou}}$, which is defined in (21), into Equation (162), we obtain:

$$J^{\mathrm{v}}_{\mathrm{ITF}}(\mathbf{w}) = \mathrm{ILD}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w}) + |\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}}|^{2} - 2\cdot\Re\{(\mathrm{ITF}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w}))^{*}\cdot\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}}\} \tag{164}$$

Equation (164) can be equivalent represented as:

$$J^{\mathrm{v}}_{\mathrm{ITF}}(\mathbf{w}) = |\mathrm{ITF}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w})|^{2} + |\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}}|^{2} - 2\cdot\Re\{(\mathrm{ITF}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w}))^{*}\cdot\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}}\} + \mathrm{ILD}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w}) - |\mathrm{ITF}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w})|^{2} \tag{165}$$

Because (33), $|\mathrm{ITF}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w})|^{2} = \mathrm{ILD}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w})|\mathrm{IC}_{\mathrm{ou}}(\mathbf{w})|^{2}$, and (165) can be represented as:

$$J^{\mathrm{v}}_{\mathrm{ITF}}(\mathbf{w}) = |\mathrm{ITF}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w})|^{2} + |\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}}|^{2} - 2\cdot\Re\{(\mathrm{ITF}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w}))^{*}\cdot\mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}}\} + \mathrm{ILD}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w})\cdot(1 - |\mathrm{IC}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w})|^{2}) \tag{166}$$

From (166), it is clear to see that the $J^{\mathrm{v}}_{\mathrm{ITF}}$ can be represented as the sum of two terms:

$$J^{\mathrm{d}}_{\mathrm{ITF}}(\mathbf{w}) = J^{\mathrm{d}}_{1}(\mathbf{w}) + J^{\mathrm{d}}_{2}(\mathbf{w}), \tag{167}$$

in which:

$$J^{\mathrm{d}}_{1}(\mathbf{w}) = |\mathrm{ITF}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w}) - \mathrm{ITF}^{\mathrm{d}}_{\mathrm{in}}|^{2}, \tag{168}$$

$$J^{\mathrm{d}}_{2}(\mathbf{w}) = \mathrm{ILD}^{\mathrm{d}}_{\mathrm{ou}}(\mathbf{w})\cdot(1 - |\mathrm{IC}_{\mathrm{ou}}(\mathbf{w})|^{2}). \tag{169}$$

From an investigation of Equations (168) and (169), we can see that the term $J_1^{\mathrm{d}}(\mathbf{w})$ of (169) is directly associated with the preservation of ITF. On the other hand, the term $J_2^{\mathrm{d}}(\mathbf{w})$ from (169) maximizes $|\mathrm{IC}_{\mathrm{ou}}^{\mathrm{d}}|^2$, which is restricted to a maximum value of 1 [30]. This explains why the $J_{\mathrm{ITF}}^{\mathrm{d}}(\mathbf{w})$ cost function is highly effective in preserving the binaural cues of a signal generated by a point acoustic source. However, this characteristic is fundamentally incompatible with the preservation of the IC of a signal generated by a diffuse field. This is due to the fact that the magnitude of $\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}(\mathbf{w})$ can vary between 0 and 1 depending on the frequency [30]. This is the first theoretical explanation about the performance of $J_{\mathrm{ITF}}^{\mathrm{d}}(\mathbf{w})$ without any simplifications in the original cost function.

Finally, let us analyze $J_{\mathrm{IC}}^{\mathrm{d}}(\lambda, k)$ presented in (55). This cost function can be used for the preservation of the IPD of a pointwise source [32], as well as the IC of a signal generated by a diffuse field, as shown in [46]. However, since it is invariant to changes in the relative magnitudes of $\mathbf{w}_{\mathrm{L}}$ and $\mathbf{w}_{\mathrm{R}}$, it cannot ensure the preservation of ILD.

5.2 COST FUNCTION FOR PRESERVATION OF POINTWISE SOURCES AND DIFFUSE SOUND FIELDS

Preserving the interaural level difference (ILD) and interaural coherence (IC) are crucial for effective preservation of the binaural impression of an acoustic source. However, to date, no augmented MWF cost function has been proposed to jointly preserve these cues. To address this gap, the following auxiliary cost function is proposed:

$$J_{\mathrm{ILD–IC}}^{\mathrm{d}}(\mathbf{w}) = \alpha_{\mathrm{ILD}} J_{\mathrm{ILD}}^{\mathrm{d}}(\mathbf{w}) + \alpha_{\mathrm{IC}} J_{\mathrm{IC}}^{\mathrm{d}}(\mathbf{w}) \tag{170}$$

in which $J_{\mathrm{ILD}}^{\mathrm{d}}$ and $J_{\mathrm{IC}}^{\mathrm{d}}$ are defined in (52) and (55), respectively. Using (170) in the optimization problem defined in (49), it is possible to obtain the desired trade-off between noise reduction and spatial preservation of the binaural sound impression for either a diffuse field or a signal generated by a pointwise source.

Both $J_{\mathrm{ILD}}^{\mathrm{d}}$ and $J_{\mathrm{IC}}^{\mathrm{d}}$ are nonnegative functions, minimization of the cost function $J_{\mathrm{ILD}}^{\mathrm{d}}$ naturally results in $\mathrm{ILD}_{\mathrm{ou}}^{\mathrm{d}} = \mathrm{ILD}_{\mathrm{in}}^{\mathrm{d}}$, while minimization of $J_{\mathrm{IC}}^{\mathrm{d}}$, results in $\left|\mathrm{IC}_{\mathrm{ou}}^{\mathrm{d}}\right| = \left|\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\right|$ e $\mathrm{IPD}_{\mathrm{ou}}^{\mathrm{d}} = \mathrm{IPD}_{\mathrm{in}}^{\mathrm{d}}$.

5.3 EXPERIMENTAL SETUP

This section presents the experimental setup to assess the performance of the proposed binaural noise reduction technique defined by the cost function in (170).

5.3.1 Binaural noise reduction techniques

The proposed method is named MWF-ILD-IC in the experiments, and the design of the binaural noise reduction filters $\mathbf{w}_{\mathrm{L}}$ and $\mathbf{w}_{\mathrm{R}}$ is made in each frequency bins by an unconstrained minimization of the J_{MWF} plus the $J_{\mathrm{ILD–IC}}^{\mathrm{d}}(\mathbf{w})$, defined in (170). The

performance of MWF-ILD-IC is compared with three other techniques: the MWF defined by the minimization of (38); the MWF with noise ITF preservation, named MWF-ITF, whose NR filters are designed by an unconstrained minimization of the cost function composed by the J_{MWF} plus the J_{ITF}, which is defined in (54); and the MWF with noise IC preservation, whose cost function is defined by the J_{MWF} plus the J_{IC} (defined in (55)). All unconstrained minimization problems were solved using the 'fmincon' routine from Matlab® [84].

5.3.2 Simulated acoustic scenarios

Two acoustic scenarios called AS_1 and AS_2 were simulated for the experiments. In all simulations, the speech signal (x) and the overall additive noise (v) were used. In the scenario AS_1, the overall noise is dominated by an point interfering source, which is approximately to $v = u$. In the scenario AS_2, the noise is dominated by its diffuse component, which is approximately to $v = n$.

5.3.3 Acoustic scenario 1 (AS_1): speech and interfering generated by pointwise sources

The signal generated by the speech source was simulated using 3 seconds of a low-pitched speech signal acquired from a database described in [85]. The clean interfering signal was simulated using a spectral-temporal signal with characteristics similar to speech and developed for use in clinical tests with hearing aids [65].

The acoustic paths between the pointwise sources and the hearing aid microphones were simulated using an impulse response database described in [63]. The database provided head-related impulse responses (HRIRs) associated with a pair of behind the ear (BTE) hearing aids, each having three microphones ($M = 3$). The HRIRs carry information about the impulse response of the hearing aids microphones, the ears, and other parts of the human anatomy, and the direct path between the pointwise source and the microphones. HRIRs were measured in an anechoic environment.

The source of speech is located at azimuth $\theta_{\text{x}} = 0^\circ$ and 80 cm from the hearing aid user, and the interference is located at azimuth $\theta_{\text{u}} = -60^\circ$ (left hemisphere). The elevantion angle was 0°.

The speech and interfering components captured by the microphones were generated by filtering the original audios with the impulse responses associated with the positions of the respective sources.

5.3.4 Acoustic scenario 2 (AS_2): speech plus noise generated by a diffuse sound field

In this scenario, the speech signal was generated in the same way as in scenario AS_1. On the other hand, the undesired component is generated by a homogeneous cylindrical diffuse field, which was simulated using the method described in [75].

5.3.5 Noisy signal and time-frequency representation

In the simulated acoustic scenarios, speech and overall noise were combined, so that the SNR in the left ear resulted in 0 dB, i.e., $\text{iSNR}_\text{L} = 0$ dB. This leads to $\text{iSNR}_\text{R} = 3.1$ dB in the right ear – see Equation (77) for the definition of iSNR_L and iSNR_R.

The time-frequency representation of the received signals was estimated using the method described in [76] which is summarized as follows. A set of 128 noisy input samples was weighted by an analysis window and transformed to the frequency domain using the Fourier transform containing $K = 256$ points, filled with zeros. After being processed, the input signal was transformed to the time domain using the inverse Fourier transform, and a synthesis window was applied over the result. The analysis and synthesis windows are the square root of the Hanning window. The processed output signal was reconstructed in the time domain by adding the adjacent set of samples with 50% overlap between them. The sampling frequency of the system is $f_s = 16$ kHz.

5.3.6 Estimation of coherence matrices

The estimations of the matrices $\mathbf{\Phi}_\text{y}$ and $\mathbf{\Phi}_\text{v}$ were calculated from the complete set of samples of the contaminated signal (y) and noise (v), respectively. In both scenarios, the coherence matrix of the speech signal was approximated by subtracting the noise coherence matrix from the noisy input signal coherence matrix, as follows: $\mathbf{\Phi}_\text{x} = \mathbf{\Phi}_\text{y} - \mathbf{\Phi}_\text{v}$. The effect that errors in the $\mathbf{\Phi}_\text{v}$ matrix have on the performance of the investigated binaural noise reduction techniques was not analyzed in this work.

5.3.7 Objective performance measures

The measures used to evaluate the performance of the proposed technique are described in the following.

Variation between input and output binaural cues: i) ΔILD, defined in (80); ii) ΔIPD, defined in (79); and iii) ΔMSC defined as:

$$\Delta\text{MSC}_\text{d} = \frac{10}{K}\sum_{k=1}^{K}\left|\frac{\mathbf{w}_\text{L}^\text{H}(k)\mathbf{\Phi}_\text{d}(k)\mathbf{w}_\text{R}(k)}{\left[\mathbf{w}_\text{L}^\text{H}(k)\mathbf{\Phi}_\text{d}(k)\mathbf{w}_\text{L}\cdot\mathbf{w}_\text{R}^\text{H}\mathbf{\Phi}_\text{d}(k)\mathbf{w}_\text{R}\right]^{1/2}} - \frac{\mathbf{q}_\text{L}^\top\mathbf{\Phi}_\text{d}(k)\mathbf{q}_\text{R}}{\left[\mathbf{q}_\text{L}^\top\mathbf{\Phi}_\text{d}\mathbf{q}_\text{L}\cdot\mathbf{q}_\text{R}^\top\mathbf{\Phi}_\text{d}\mathbf{q}_\text{R}\right]^{1/2}}\right|^2. \tag{171}$$

in which d can be associated with speech (x) or one of the parts of overall noise (u and n).

The variation between the input and output SNR (ΔSNR) was defined in (78).

5.4 RESULTS AND DISCUSSION

This section presents the results and discussion of the experiments for the two acoustic scenarios described in Sections 5.3.3 and 5.3.4.

5.4.1 Acoustic scenario 1 (AS_1): speech and interfering generated by pointwise sources

Figures 22, 22 and 22 present the ΔILD, ΔIPD and ΔMSC of the intefering and speech sources.

The performance of the MWF is represented by a black dashed line. The initial flat region in the graphs (where $\alpha \to 0$) corresponds to the region where all techniques have equivalent performance. In this region, the perceived position of the point interfering source is the same as that of the speech source. This is a known behavior of the MWF and was investigated by theoretical [28] and psychoacoustic experiments [50].

As the parameter α grows, the filters generated by the MWF-IC distort the original ILD of the interfering source – note that the $\Delta\text{ILD}_\text{u}$ for the MWF-IC is greater than or equal to that produced by the MWF. On the other hand, the $\Delta\text{IPD}_\text{u}$ and the $\Delta\text{MSC}_\text{u}$ tend to zero as α grows. This ability of the IC to preserve IPD was presented in [32]. Therefore, for MWF-IC, the increase of α makes the interfering source perceived as having two components: one at its original position (at –60°), since the IPD is preserved, and the other at the speech position (at 0°). For the speech source, the results presented by the MWF-IC can be classified as closer to the MWF, suggesting that the location of the speech source is perceived closer to the correct location. Taking into account the noise reduction performance, Figure 23 shows that the MWF-IC has a performance similar to the MWF, for a wide range of α; however, they differ for large α. This shows the trade-off between noise reduction and binaural cue preservation.

As presented in Section 5.1 the MWF-ITF cost function has an additional term associated with the preservation of the ITF. In the acoustic scenario AS_1, this term is associated with the preservation of the interfering source, which is a pointwise source. In this way, MSC tends to 1. Figures 22(a), 22(b) and 22(c) show a decrease in $\Delta\text{ILD}_\text{u}$, $\Delta\text{IPD}_\text{u}$ and $\Delta\text{MSC}_\text{u}$ for the MWF-ITF when the parameter α is around 10^{-6}. Note that the decrease in $\Delta\text{MSC}_\text{u}$ is probably due to the term that keeps the value of $\text{MSC}_\text{u}^\text{ou}$ close to 1 (see Equation (169)). This probably indicates that as α increases, the perceived position of the interfering source is moving closer to the correct position of the source. However, the decreases in these noise measures are accompanied by increases in $\text{ILD}_\text{x}^\text{ou}$,

$\text{IPD}_{\text{x}}^{\text{ou}}$, and $\text{MSC}_{\text{x}}^{\text{ou}}$, especially for values of α greater than 10^{-3}. This shows a trade-off between preservation of the binaural cues of the interfering and the speech source. Figure 23 indicates that MWF-ITF and MWF-IC have a similar performance noise reduction: for values of α greater than 10^{-3} the binaural NR performance of MWF-ITF decreases compared to the MWF. However, the values of $\Delta\text{SNR}_{\text{L}}$ and $\Delta\text{SNR}_{\text{R}}$ are the lowest among the techniques in the experiment. All results in this paragraph point towards the existence of a trade-off between two competing objectives. On the one hand, there is the need to reduce the noise while preserving the binaural characteristics of the speech source. On the other hand, preservation of the binaural cues of the interfering source also needs to be taken into account. These competing objectives lead to a complex trade-off that must be carefully balanced. These results indicate that attempts to increase noise reduction and binaural preservation of the speech source may compromise preservation of the binaural cues of the interfering source. Therefore, this trade-off should be carefully considered in any application of the MWF for binaural noise reduction.

The proposed MWF-ILD-IC technique shows a decrease in the values of $\Delta\text{ILD}_{\text{u}}$ and $\Delta\text{IPD}_{\text{u}}$ with an increase of α. This finding is similar to that of the MWF-ITF, although the MWF-ILD-IC achieves smaller values of $\Delta\text{MSC}_{\text{u}}$ compared to the MWF-ITF. It should be noted that the $\Delta\text{MSC}_{\text{u}}$ for the MWF-ITF seems to decrease more slowly than that of the proposed technique. A significant difference between the MWF-ILD-IC and the MWF-ITF is the preservation of the binaural cue of the speech source. The proposed technique has significantly lower values of $\Delta\text{ILD}_{\text{x}}$, $\Delta\text{IPD}_{\text{x}}$, and $\Delta\text{MSC}_{\text{x}}$ for $\alpha \geqslant 10^{-3}$ compared to MWF-ITF. This suggests that the proposed technique better preserves the perceived position of the speech. Considering the binaural noise reduction performance of MWF-ILD-IC presented in Figure 23, the proposed technique has better $\Delta\text{SNR}_{\text{L}}$ performance than MWF-ITF and MWF-IC for $\alpha \geqslant 10^{-3}$ and a $\Delta\text{SNR}_{\text{R}}$ performance similar to the of MWF-IC. Therefore, in the region where binaural preservation of the interfering source can be achieved, the proposed technique is the most effective in reducing the power of noise in the ear closer to the interfering source, compared to MWF-ITF and MWF-IC.

5.4.2 Acoustic scenario 2 (AS_2): speech plus a noise generated by a diffuse sound field

Figures 24(a), 24(b) and 24(c) respectively present $\Delta\text{ILD}_{\text{n}}$, $\Delta\text{IPD}_{\text{n}}$ and $\Delta\text{MSC}_{\text{n}}$.

The performance of the MWF is represented by a black dashed line in Figures 24 and 25. Similar to the previous Section, it can be observed that all techniques present similar performance for α near 0.

Considering the results in Figures 24(a), 24(b) and 24(c) for the MWF-ITF, it can be seen that $\Delta\text{ILD}_{\text{n}}$ increases for $\alpha \geqslant 10^{-4}$, and $\Delta\text{MSC}_{\text{n}}$ decreases until $\alpha \cong 10^{-2}$. This behavior can be explained by Equations (32) and (33), which reveal that different values of the product between $\text{ILD}_{\text{n}}^{\text{ou}}$ and $\text{MSC}_{\text{n}}^{\text{ou}}$ can lead to the same value for $|\text{ITF}_{\text{n}}^{\text{ou}}|$. For example,

Figure 22 – Errors in the binaural cues measured in the acoustic scenario with point noise: (a) $\Delta\text{ILD}_\text{u}$, (b) ΔILDx, (c)$\Delta\text{IPD}_\text{u}$, (d)$\Delta\text{IPDx}$, (e) $\Delta\text{MSC}_\text{u}$ e (f) ΔMSCx.

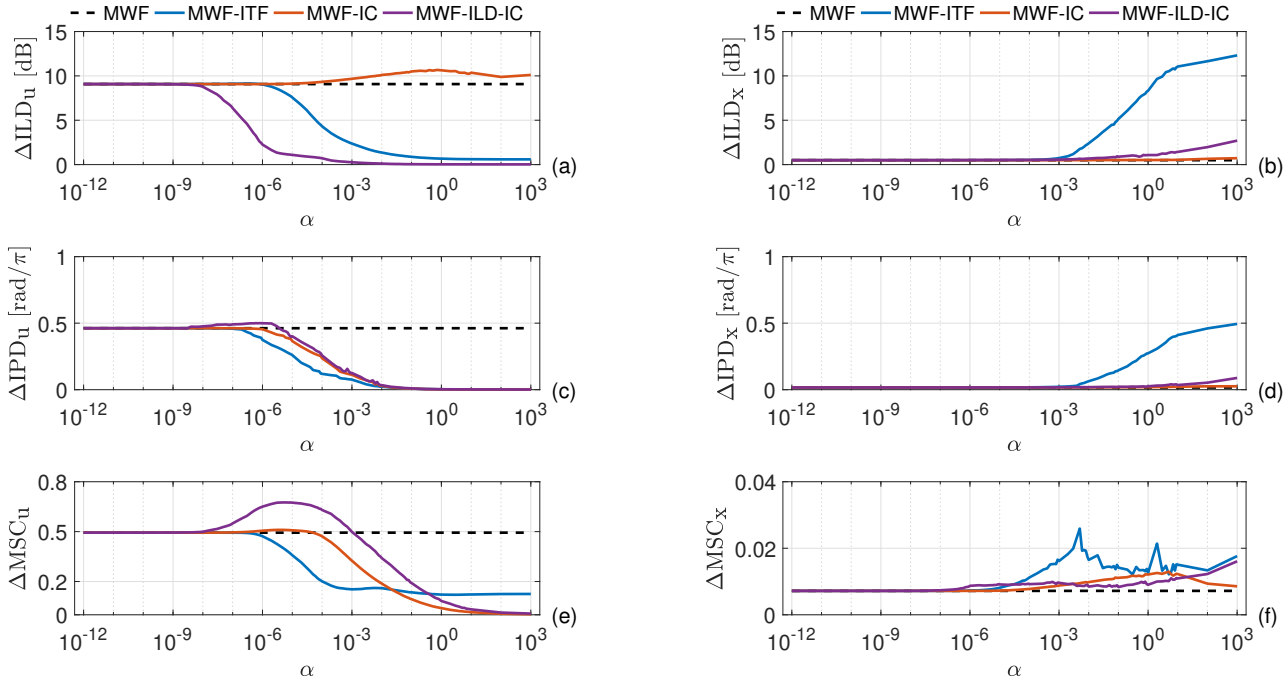
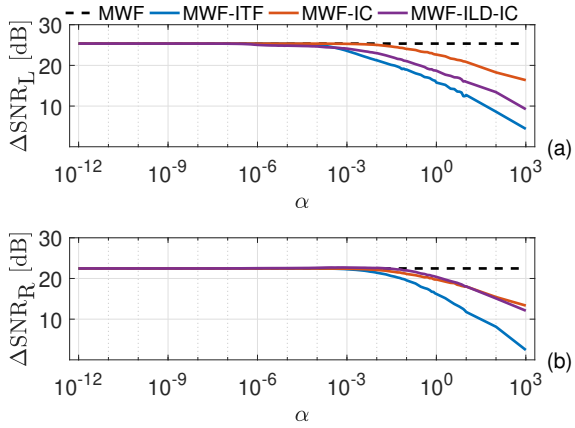


Figure 23 – Noise reduction variation (ΔSNR) in the left (L) and right (R) hearing aids in the acoustic scenario with point noise: (a) $\Delta\text{SNR}_\text{L}$ e (b) $\Delta\text{SNR}_\text{R}$.



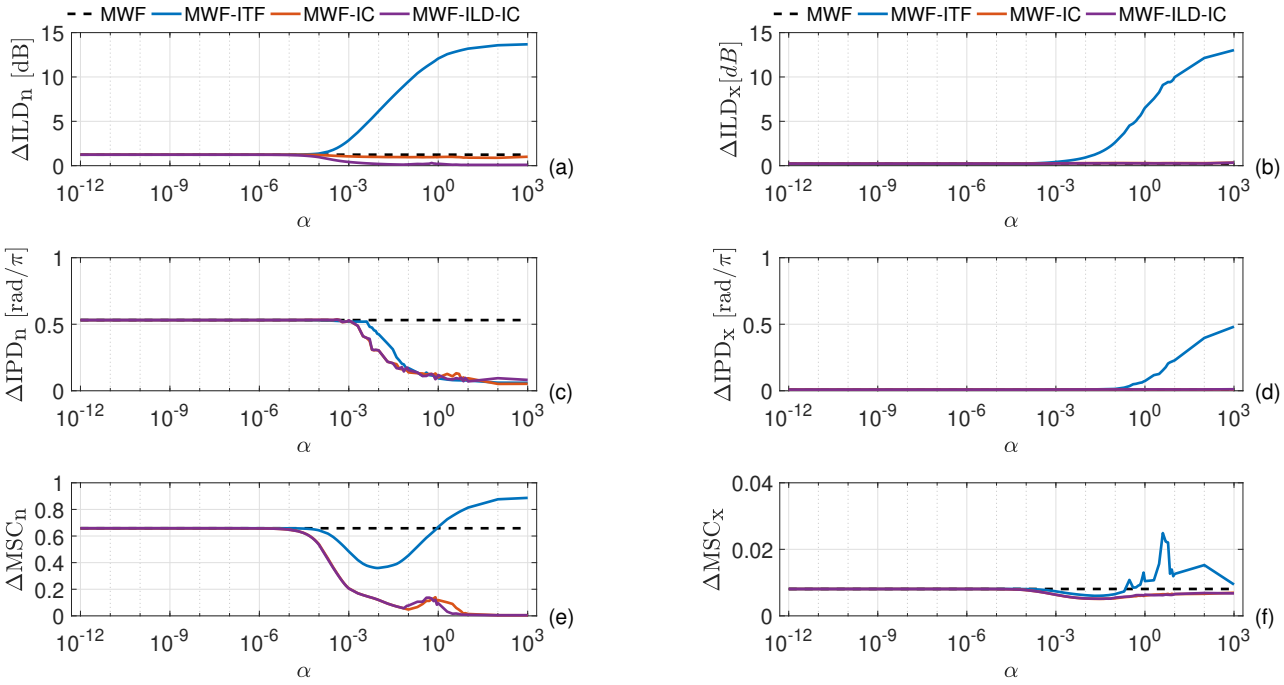
if $\text{ILD}_\text{n}^\text{in} = 2$ and $\text{MSC}_\text{n}^\text{in} = 0.25$ (which leads to $|\text{ITF}_\text{n}^\text{in}| = (0.5)^{1/2}$), then $\text{ILD}_\text{n}^\text{ou} = 4$ and $\text{MSC}_\text{n}^\text{in} = 0.125$, or $\text{ILD}_\text{n}^\text{ou} = 8$ and $\text{MSC}_\text{n}^\text{in} = 0.0625$ lead both to $|\text{ITF}_\text{n}^\text{in}| = (0.5)^{1/2}$. Furthermore, $\Delta\text{IPD}_\text{n}^\text{in}$ decreases for $\alpha \geqslant 10^{-3}$, indicating preservation of the original IPD noise. Figure 25 shows that MWF-ITF has equal or lower values of SNR_R than MWF-IC and MWF-ILD-IC for $\alpha \geqslant 10^{-3}$. The results for SNR_R are similar to those presented by SNR_L.

Figures 24(a), 24(b) and 24(c) show a decrease of $\Delta\text{IPD}_\text{ou}^\text{n}$ and $\Delta\text{MSC}_\text{ou}^\text{n}$ for MWF-IC technique for $\alpha \geqslant 10^{-3}$, while keeping the values of $\Delta\text{ILD}_\text{ou}^\text{n}$ closer to the values presented by MWF. This suggests that the MWF-IC preserves the perception of the diffuseness of the noise. Furthermore, the ability of the MWF-IC to maintain the binaural

cues of speech signals is similar to that of the MWF. This indicates that the proposed technique preserves the perception of the original location of the speech. The MWF-IC presents an average reduced performance in $\Delta\text{SNR}_\text{L}$ and $\Delta\text{SNR}_\text{R}$ compared to the MWF.

Finally, considering the presented results in Figures 24(a), 24(b) and 24(c), it can be concluded that the proposed MWF-ILD-IC technique reduces errors in noise IPD and MSC, showing a performance very similar to MWF-IC. However, unlike MWF-IC, the proposed technique also leads to a numerical reduction in the values of $\Delta\text{ILD}^\text{n}_\text{ou}$. This indicates that the MWF-ILD-IC also preserves the diffuseness perception of the noise signal. In addition, the proposed technique presents a noise reduction performance similar to that of the MWF-IC.

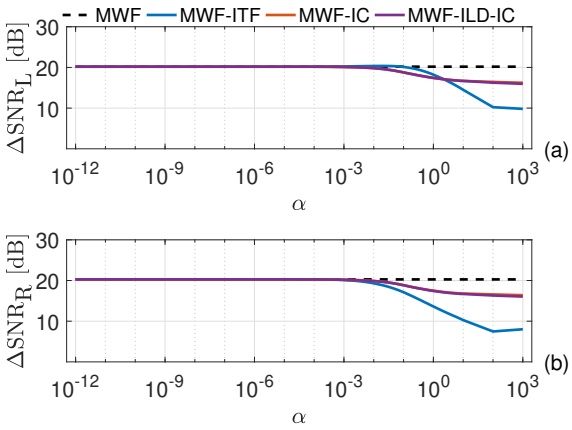
Figure 24 – Errors in the binaural cues measured in the acoustic scenario with noise generated by a diffuse field: (a) $\Delta\text{ILD}_\text{n}$, (b) ΔILDx, (c)$\Delta\text{IPD}_\text{n}$, (d)$\Delta\text{IPDx}$, (e) $\Delta\text{MSC}_\text{n}$ e (f) ΔMSCx.



5.5 A QUADRATIC REFORMULATION FOR THE MWF WITH ILD/IC PRESERVATION

The design of noise reduction filters through the minimization problem in (49b) has some disadvantages. One of them is how to design the parameters α_i a priori (i.e., before executing the optimization procedure) to achieve a guaranteed performance improvement in some performance measure (e.g., SNR or binaural cue variation measures). The difficulty in this task lies in the fact that it is not clear how α_i is associated with some measure of noise reduction or variation of binaural signals. Furthermore, the complexity in determining α_i increases with the number of terms in J_T. Despite the possibility of setting up these

Figure 25 – Noise reduction variation (ΔSNR) in the left (L) and right (R) hearing aids in the acoustic scenario with undesired diffuse field (noise): (a) $\Delta\text{SNR}_\text{L}$ e (b) $\Delta\text{SNR}_\text{R}$.



parameters a posteriori (using methods such as [30]: iterative search algorithms, grid search algorithms, or SNR maximization), this procedure results in increased computational time due to the number of times that the optimization problem in (49b) has to be solved. To overcome this issue, we propose to first rewrite the unconstrained problem in (170) as a constrained problem in the following form:

$$\mathbf{w}_{\text{MWF-LC}} = \arg.\min.\ J_{\text{MWF}}(\mathbf{w}) \tag{172a}$$

$$\text{s.t. } J_{\text{ILD}}(\mathbf{w}) \leqslant \delta_{\text{ILD}}^2, \tag{172b}$$

$$J_{\text{IC}}(\mathbf{w}) \leqslant \delta_{\text{IC}}^2, \tag{172c}$$

in which δ_{ILD}^2 and δ_{IC}^2 the parameters directly associated with the variations of ILD and IC. We can reformulate the ILD constraint in (172b) as a set of quadratic constraints, leading to the constraints in (96), rewritten here as:

$$(-1)^{t-1} \cdot \mathbf{w}^{\text{H}} \mathbf{\Phi}_{\text{cc}}(\delta_1)\mathbf{w} \geqslant 0, \tag{173}$$

$$(-1)^{t-1} \cdot \mathbf{w}^{\text{H}} \mathbf{\Phi}_{\text{cc}}(\delta_2)\mathbf{w} \leqslant 0, \tag{174}$$

in which t is defined in (98).

Now, to rewrite the IC constraint in (172c) as a set of quadratic constraints, we considered the relation between the modulus of a complex number and its real and imaginary parts. Then, the following equivalent representation for the $J_{\text{IC}}(\mathbf{w})$ constraint in (172c) can be derived:

$$J_{\text{IC}}(\mathbf{w}) = \Re^2\{e_{\text{IC}}(\mathbf{w})\} + \Im^2\{e_{\text{IC}}(\mathbf{w})\} \leqslant \delta_{\text{IC}}^2, \tag{175}$$

in which:

$$e_{\text{IC}}(\mathbf{w}) = \text{IC}_{\text{ou}}^{\text{d}}(\mathbf{w}) - \text{IC}_{\text{in}}^{\text{d}}. \tag{176}$$

Using (27) in (176) leads to:

$$e_{\mathrm{IC}}(\mathbf{w}) = \frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}}{(\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}} \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}})^{1/2}} - \mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}. \tag{177}$$

Using auxiliary variables to separate the real and imaginary parts of the constraint in (175) in two other constraints leads to the following:

$$J_{\mathrm{IC}}(\mathbf{w}) = \varepsilon_{\Re}^2 + \varepsilon_{\Im}^2 \leqslant \delta_{\mathrm{IC}}^2 \tag{178a}$$

$$\Re^2\{e_{\mathrm{IC}}(\mathbf{w})\} \leqslant \varepsilon_{\Re}^2 \tag{178b}$$

$$\Im^2\{e_{\mathrm{IC}}(\mathbf{w})\} \leqslant \varepsilon_{\Im}^2 \tag{178c}$$

in which $\varepsilon_{\Re}$ and $\varepsilon_{\Im}$ are real auxiliary variables (slack variables), which must be optimized together with $\mathbf{w}_{\mathrm{L}}$ and $\mathbf{w}_{\mathrm{R}}$.

Using (177) in both (178b) and (178c), and manipulating the result leads to the following:

$$r(\mathbf{w}) \cdot (\Re\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} - \varepsilon_{\Re}) \leqslant \Re\{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}\} \leqslant r(\mathbf{w}) \cdot (\Re\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} + \varepsilon_{\Re}), \tag{179}$$

$$r(\mathbf{w}) \cdot (\Im\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} - \varepsilon_{\Im}) \leqslant \Im\{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}\} \leqslant r(\mathbf{w}) \cdot (\Im\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} + \varepsilon_{\Im}) \tag{180}$$

in which:

$$r(\mathbf{w}) = (\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}} \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}})^{1/2}\ . \tag{181}$$

Using the definition of the real and imaginary parts of a complex number in (179) and (180) lead to:

$$2 \cdot r(\mathbf{w}) \cdot (\Re\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} - \varepsilon_{\Re}) \leqslant \mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}} + \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}} \leqslant 2 \cdot r(\mathbf{w}) \cdot (\Re\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} + \varepsilon_{\Re}), \tag{182}$$

$$2j \cdot r(\mathbf{w}) \cdot (\Im\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} - \varepsilon_{\Im}) \leqslant \mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}} - \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}} \leqslant 2j \cdot r(\mathbf{w}) \cdot (\Im\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} + \varepsilon_{\Im}). \tag{183}$$

Rewrite the constraints in (182) and (183) as a set of quadratic forms with respect to (w.r.t.) $\mathbf{w}_{\mathrm{L}}$ and $\mathbf{w}_{\mathrm{R}}$ (and also w.r.t. $\mathbf{w}$), it is necessary to write $r(\mathbf{w})$ as a square function of these variables. In this way, $r(\mathbf{w})$ can be written as:

$$(\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}} \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}})^{1/2} = \left(\frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}} \cdot (\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}})^2\right)^{1/2} \tag{184a}$$

$$= \left(\frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}}}\right)^{1/2} \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}} \tag{184b}$$

Replacing (23) in (184), results in:

$$(\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{L}} \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}})^{1/2} = (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{ou}}(\mathbf{w}))^{1/2} \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{d}}\mathbf{w}_{\mathrm{R}} \tag{185}$$

If we select a value δ_{ILD} in (172b) small enough to ensure negligible errors between the output and input ILDs, it is possible to assume for (186) that

$$\mathrm{ILD}_{\mathrm{ou}}(\mathbf{w}) \approx \mathrm{ILD}_{\mathrm{in}}. \tag{186}$$

Then, using (186) on the right-hand side of (185), it can be approximated as follows:

$$r(\mathbf{w}) = (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{ou}}(\mathbf{w}))^{1/2} \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} \approx (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{in}})^{1/2} \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} = \tilde{r}(\mathbf{w}). \tag{187}$$

Therefore, replacing $r(\mathbf{w})$ by its approximation defined in (187) in both equations (182) and (183) leads to:

$$2 \cdot \tilde{r}(\mathbf{w}) \cdot (\Re\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} - \varepsilon_{\Re}) \leqslant \mathbf{w}_{\mathrm{L}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} + \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{L}} \leqslant 2 \cdot (\Re\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} + \varepsilon_{\Re}) \cdot \tilde{r}(\mathbf{w}), \tag{188}$$

$$2j \cdot \tilde{r}(\mathbf{w}) \cdot (\Im\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} - \varepsilon_{\Im}) \leqslant \mathbf{w}_{\mathrm{L}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} - \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{L}} \leqslant 2j \cdot (\Im\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}}\} + \varepsilon_{\Im}) \cdot \tilde{r}(\mathbf{w}) \tag{189}$$

Replacing the definition of $\tilde{r}(\mathbf{w})$, splitting the interval constraint in (188) into one constraint for the lower bound and another for the upper bound, and isolating the variables on the left-hand side leads to:

$$\mathbf{w}_{\mathrm{L}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} + \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{L}} - 2 \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} \cdot (\Re\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}} \cdot (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{in}})^{1/2}\} - \varepsilon_{\Re} \cdot (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{in}})^{1/2}) \geqslant 0, \tag{190}$$

$$\mathbf{w}_{\mathrm{L}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} + \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{L}} - 2 \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} \cdot (\Re\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}} \cdot (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{in}})^{1/2}\} + \varepsilon_{\Re} \cdot (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{in}})^{1/2}) \leqslant 0. \tag{191}$$

Proceeding in the same way for the constraint in (189), results in:

$$\mathbf{w}_{\mathrm{L}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} - \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{L}} - 2j \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} \cdot (\Im\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}} \cdot (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{in}})^{1/2}\} - \varepsilon_{\Im} \cdot (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{in}})^{1/2}) \geqslant 0, \tag{192}$$

$$\mathbf{w}_{\mathrm{L}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} - \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{L}} - 2j \cdot \mathbf{w}_{\mathrm{R}}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{d}} \mathbf{w}_{\mathrm{R}} \cdot (\Im\{\mathrm{IC}_{\mathrm{in}}^{\mathrm{d}} \cdot (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{in}})^{1/2}\} + \varepsilon_{\Im} \cdot (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{in}})^{1/2}) \leqslant 0, \tag{193}$$

Replacing the equality $\mathrm{ITF}_{\mathrm{in}}^{\mathrm{d}} = \mathrm{IC}_{\mathrm{in}}^{\mathrm{d}} \cdot (\mathrm{ILD}_{\mathrm{d}}^{\mathrm{in}})^{1/2}$ defined in (21) in equations (190)-(193), and rewriting the result using the vector $\mathbf{w}$ defined in (40), the four constraints can be represented as the following set of quadratic constraints:

$$\mathbf{w}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{rr}}(-\varepsilon_{\Re}) \mathbf{w} \geqslant 0, \tag{194}$$

$$\mathbf{w}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{rr}}(+\varepsilon_{\Re}) \mathbf{w} \leqslant 0, \tag{195}$$

$$\mathbf{w}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{ii}}(-\varepsilon_{\Im}) \mathbf{w} \geqslant 0, \tag{196}$$

$$\mathbf{w}^{\mathrm{H}} \boldsymbol{\Phi}_{\mathrm{ii}}(+\varepsilon_{\Im}) \mathbf{w} \leqslant 0, \tag{197}$$

in which $\boldsymbol{\Phi}_{\mathrm{rr}}(\varepsilon_{\mathrm{rr}})$ and $\boldsymbol{\Phi}_{\mathrm{ii}}(\varepsilon_{\mathrm{ii}})$ are defined, respectively, as:

$$\boldsymbol{\Phi}_{\mathrm{rr}}(\varepsilon_{\mathrm{rr}}) = \begin{bmatrix} \mathbf{0}_{M \times M} & \boldsymbol{\Phi}_{\mathrm{v}} \\ \boldsymbol{\Phi}_{\mathrm{v}} & -2(\Re\{\mathrm{ITF}_{\mathrm{v}}^{\mathrm{in}}\} + \varepsilon_{\mathrm{rr}} \cdot (\mathrm{ILD}_{\mathrm{v}}^{\mathrm{in}})^{1/2}) \cdot \boldsymbol{\Phi}_{\mathrm{v}} \end{bmatrix}, \tag{198}$$

$$\boldsymbol{\Phi}_{\mathrm{ii}}(\varepsilon_{\mathrm{ii}}) = \begin{bmatrix} \mathbf{0}_{M \times M} & \boldsymbol{\Phi}_{\mathrm{v}} \\ -\boldsymbol{\Phi}_{\mathrm{v}} & -2j(\Im\{\mathrm{ITF}_{\mathrm{v}}^{\mathrm{in}}\} + \varepsilon_{\mathrm{ii}} \cdot (\mathrm{ILD}_{\mathrm{v}}^{\mathrm{in}})^{1/2}) \cdot \boldsymbol{\Phi}_{\mathrm{v}} \end{bmatrix}, \tag{199}$$

for $\varepsilon_{\mathrm{rr}} \in \{-\varepsilon_{\Re}, +\varepsilon_{\Re}\}$ and $\varepsilon_{\mathrm{ii}} \in \{-\varepsilon_{\Im}, +\varepsilon_{\Im}\}$.

Using the constraints in (173), (174) and (194)-(197) it is possible to obtain an approximate solution for the problem in (172) solving the following optimization problem:

$$\mathbf{w}_{\text{QCQP-MWF-LC–1}} = \underset{\mathbf{w}, \varepsilon_{\Re}, \varepsilon_{\Im}}{\arg.\min.} \; J_{\mathrm{MWF}}(\mathbf{w}) \tag{200a}$$

$$\text{s.t. } \varepsilon_{\Re}^2 + \varepsilon_{\Im}^2 \leqslant \delta_{IC}^2, \tag{200b}$$

$$\mathbf{w}^{\mathrm{H}} \boldsymbol{\Phi}_q \mathbf{w} \leqslant 0, \; q = 1, 2, \cdots, 6, \tag{200c}$$

in which:

$$\mathbf{\Phi}_1 = -(-1)^{t-1}\mathbf{\Phi}_{\mathrm{cc}}(\delta_1), \tag{201}$$

$$\mathbf{\Phi}_2 = (-1)^{t-1}\mathbf{\Phi}_{\mathrm{cc}}(\delta_2), \tag{202}$$

$$\mathbf{\Phi}_3 = -\mathbf{\Phi}_{\mathrm{rr}}(-\varepsilon_{\Re}), \tag{203}$$

$$\mathbf{\Phi}_4 = \mathbf{\Phi}_{\mathrm{rr}}(+\varepsilon_{\Re}), \tag{204}$$

$$\mathbf{\Phi}_5 = -\mathbf{\Phi}_{\mathrm{ii}}(-\varepsilon_{\Im}), \tag{205}$$

$$\mathbf{\Phi}_6 = \mathbf{\Phi}_{\mathrm{ii}}(-\varepsilon_{\Im}) \tag{206}$$

To reduce the number of variables to optimize, it is proposed to set $\varepsilon^2_{\Re}$ and $\varepsilon^2_{\Im}$ a priori in the following form:

$$\varepsilon^2_{\Re} = \varepsilon^2_{\Im} = \delta^2_{\mathrm{IC}}. \tag{207}$$

Using (207) in (200) leads to:

$$\mathbf{w}_{\text{QCQP-MWF-LC–2}} = \arg.\min_{\mathbf{w}}.\ J_{\mathrm{MWF}}(\mathbf{w}) \tag{208a}$$

$$\text{s.t. } \mathbf{w}^{\mathrm{H}}\mathbf{\Phi}_q\mathbf{w} \leqslant 0,\ q = 1, 2, \cdots, 6. \tag{208b}$$

5.6 A CONVEX FORM FOR THE MWF WITH PRESERVATION OF ILD AND IC

The quadratic reformulation of the optimization problem in (208) presented in the previous section demonstrates that the original problem in (172) can be approximated by a quadratically constrained quadratic problem (QCQP). However, the new approximated form for the MWF-ILD-IC in (208) is a non-convex problem, which in practical terms means that it is computationally difficult to find a solution. QCQPs on complex variables with more than two constraints are NP-hard problems in the general case [86], i.e., it is not possible to find the optimal solution in polynomial time [86]. This is a considerable limitation for hearing aid applications, which require computationally efficient algorithms. Considering these issues, in this section a convex relaxation for the QCQP in (208) is proposed. The method finds suboptimal solutions that can be calculated in less time compared to the QCQP proposed in (208).

To find the convex method, we compute the Lagrangian of (208), given by [70, pg. 215]:

$$\mathcal{L}(\mathbf{w}, \boldsymbol{\mu}) = \mathbf{w}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{yq}}(\boldsymbol{\mu})\mathbf{w} - \mathbf{w}^{\mathrm{H}}\mathbf{p}_{\mathrm{xx}} - \mathbf{p}^{\mathrm{H}}_{\mathrm{xx}}\mathbf{w} + p_{\mathrm{xx}}\,, \tag{209}$$

in which $\mathbf{p}_{\mathrm{xx}}$ and p_{xx} are defined in (41) and (42), respectively; $\boldsymbol{\mu} \in \mathbb{R}^6$ is a vector whose entries are the Lagrange multipliers, and matrix $\mathbf{\Phi}_{\mathrm{yq}}(\boldsymbol{\mu})$ is given by:

$$\mathbf{\Phi}_{\mathrm{yq}}(\boldsymbol{\mu}) = \mathbf{\Phi}_{\mathrm{yy}} + \mu_{qq}\sum_{qq=1}^{6}\mathbf{\Phi}_{qq}\,. \tag{210}$$

The *dual relaxation* of the problem (200) is given by [70, pg. 223]:

$$\max_{\boldsymbol{\mu}\geqslant 0} . \min_{\mathbf{w}} . \ \mathcal{L}(\mathbf{w}, \boldsymbol{\mu}) . \tag{211}$$

As also presented in Section 4.2.2, the solution of this problem gives a lower bound solution to the problem in (208), and it can always be transformed into a convex problem, even if the original problem is not convex [70, pg. 255], which is the case of problem in (208).

Note that if the parameters $\boldsymbol{\mu}$ are fixed, then it is possible to find an analytical solution for $\mathbf{w}$ in (209). Taking the gradient of $\mathcal{L}(\mathbf{w}, \boldsymbol{\mu})$ with respect to $\mathbf{w}$ leads to:

$$\nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \boldsymbol{\mu}) = \boldsymbol{\Phi}_{\mathrm{yq}}(\boldsymbol{\mu})\mathbf{w} - \boldsymbol{\Phi}_{\mathrm{xx}}\mathbf{q} . \tag{212}$$

Equating (212) to zero and solving for $\mathbf{w}$, leads to the minimum:

$$\mathbf{w} = \boldsymbol{\Phi}_{\mathrm{yq}}(\boldsymbol{\mu})^{-1}\mathbf{p} , \tag{213}$$

in which we assume $\boldsymbol{\Phi}_{\mathrm{yq}}(\boldsymbol{\mu}) \succeq 0$.

To find the parameter $\boldsymbol{\mu}$, the argument that maximizes problem (211) need to be find. Due to this, let us replace replacing $\mathbf{w}$, defined in (213), in the Lagrangian $\mathcal{L}(\mathbf{w}, \boldsymbol{\mu})$, defined in (209), which leads to:

$$\boldsymbol{\mu}^{\mathrm{opt}} = \arg . \max_{\boldsymbol{\mu}\geqslant 0} . \ -\mathbf{p}^{\mathrm{H}}\boldsymbol{\Phi}_{\mathrm{yq}}(\boldsymbol{\mu})\mathbf{p} , \quad \text{s.t.} \quad \boldsymbol{\Phi}_{\mathrm{yq}}(\boldsymbol{\mu}) \succeq 0 . \tag{214}$$

The problem presented in (214) can be represented in equivalent form using an auxiliary variable γ^{opt} and the Schur complement [70, pg. 650], resulting in:

$$\begin{aligned} \gamma^{\mathrm{opt}}, \boldsymbol{\mu}^{\mathrm{opt}} = \arg . \max_{\gamma, \boldsymbol{\mu}} . \quad & \gamma \\ \text{s.t.} \quad & \boldsymbol{\mu} \geqslant 0 , \\ & \begin{bmatrix} \boldsymbol{\Phi}_{\mathrm{yq}}(\boldsymbol{\mu}) & \mathbf{p}_{\mathrm{xx}} \\ \mathbf{p}_{\mathrm{xx}}^{\mathrm{H}} & p_{\mathrm{xx}} - \gamma \end{bmatrix} \succeq 0 . \end{aligned} \tag{215}$$

Therefore, the solution of the proposed technique is defined to the one obtained by replacing the value of $\boldsymbol{\mu}^{\mathrm{opt}}$, found by the problem in (215), in Equation (213), leading to:

$$\mathbf{w}_{\text{SDP-MWF-LC}} = \boldsymbol{\Phi}_{\mathrm{yq}}(\boldsymbol{\mu}^{\mathrm{opt}})^{-1}\mathbf{p} , \tag{216}$$

The problem presented in (215) is a semidefinite programming (SDP) problem with variables γ e $\boldsymbol{\mu}$, which has suboptimal solutions that can be obtained efficiently, that is, in less time than the algorithm in (208).

5.7 EXPERIMENTAL SETUP

In this section computational simulations are presented to evaluate the performance of the proposed technique.

5.7.1 Noise reduction techniques

Three noise reduction techniques were compared in the experiments: i) the conventional MWF defined in (48); ii) the proposed non-convex QCQP-MWF-LC[1] presented in (208); and iii) the convex SDP-MWF-LC technique presented in (216). For the last two techniques, the following values were used: $\delta_{\text{ILD}} = 10^{-4}$ e $\delta_{\text{IC}} = 10^{-4}$. These values were chosen to ensure small errors in the binaural cues, necessary condition to the approximation in (187).

5.7.2 Simulated acoustic scenario

Two simulated acoustic scenarios, called AS_1 and AS_2, were used in the experiments, both of which contained a point speech source. In the acoustic scenario AS_1, the noise component of the noisy signal y comprises only a diffuse acoustic field, i.e., $v = n$. In the acoustic scenario AS_2, in addition to n, the noise also contains a component from an point interfering source (signal u), then, $v = u + n$. In both acoustic scenarios, the speech source is located at the right azimuth $\theta_{\text{x}} = 15^{\circ}$ (right hemisphere) and is 3 meters from the hearing aid user. In the AS_2, the point interfering source is also 3 meters from the user, but its azimuth (θ_{u}) has been varied from $\theta_{\text{u}} = -90^{\circ}$ (left) to $\theta_{\text{u}} = +90^{\circ}$ (right) in steps of 15°. The diffuse acoustic field was assumed to be cylindrical and was generated according to the algorithm in [75]. The head-related impulse response database (HRIR), which determines the acoustic characteristics carried out by the received signals, is described in [63]. Acoustic signals were transmitted in a reverberating environment ($\text{T}_{60} \cong 300$ ms) and acquired by a pair of hearing aids coupled to a body and head simulator. Each hearing aid has $M_{\text{L}} = M_{\text{R}} = 3$ microphones for a total of $M = 6$.

5.7.3 Received signals

The clean speech and interference signals generated at the pointwise sources were simulated using approximately 2 seconds long audio [64]. The sampling frequency of the signals is 16 kHz. The speech and interference components received in each microphone were generated by individually convolving the clean signal with the HRIRs associated with the acoustic path between the acoustic source and each of the hearing aid microphones. The noisy input signal was generated in the microphones by adding the speech component and the respective noise component of each acoustic scenario. In the acoustic scenario AS_1, the signal-to-noise ratio (SNR) between speech and diffuse noise is 0 dB. In the AS_2 acoustic scenario, the SNR and the signal-to-interference ratio (SIR) are both 0 dB, resulting in a signal-to-noise-plus-interference ratio (SINR) of –3 dB.

[1] L and C are used to indicate the preservation of the interaural level difference and the interaural coherence, respectvaly.

5.7.4 Time-frequency representation and processing

The noisy input signal was processed using windows of 256 samples (16 ms), which were weighted by an analysis window and transformed to the frequency domain using the 512-point Fourier transform. The overlap between the adjacent set of samples is 50%. After filtering the input signal in the frequency domain, the inverse Fourier transform was used to transform the processed signal back to the time domain. The output signal was obtained using the overlap-and-add technique, using a synthesis window. Both the analysis and the synthesis windows are the square root of the Hanning window. A detailed description of the algorithm to represent signals in the time-frequency domain is presented in [76].

The coherence matrices $\mathbf{\Phi}_{\mathrm{n}}$ and $\mathbf{\Phi}_{\mathrm{y}}$ were calculated directly from the diffuse noise component and the noisy input signal. In AS_1, the coherence matrix $\mathbf{\Phi}_{\mathrm{x}}$ was estimated by performing the eigendecomposition between $\mathbf{\Phi}_{\mathrm{x}}$ and $\mathbf{\Phi}_{\mathrm{n}}$, similar to the one presented in 4.4.2. In AS_2, the coherence matrices $\mathbf{\Phi}_{\mathrm{x}}$ and $\mathbf{\Phi}_{\mathrm{u}}$ were determined using $\mathbf{\Phi}_{\mathrm{n}}$ and performing the eigendecomposition exactly as presented in Section 4.4.2.

The coefficients $\mathbf{w}_{\mathrm{L}}$ and $\mathbf{w}_{\mathrm{R}}$ of the QCQP-MWF-LC technique were obtained using the 'fmincon' function of Matlab [84]. The coefficients of the SDP-MWF-LC technique were obtained using the SeDuMi solver [87]. Modeling of both optimization programs was done using the Yalmip toolbox [88].

5.7.5 Objective performance measure

The binaural-cue preservation performance of the proposed technique was assessed by the ΔILD, ΔIPD, and ΔMSC defined in (80), (79), and (171), respectively. Noise reduction performance was assessed using ΔSINR, defined in (157), and ΔSNR defined in (157).

5.8 RESULTS AND DISCUSSION

Table 8 presents the $\Delta\mathrm{ILD}_{\mathrm{n}}$, $\Delta\mathrm{IPD}_{\mathrm{n}}$,and $\Delta\mathrm{MSC}_{\mathrm{n}}$ measured for the acoustic scenario AS_1. All these measures were calculated using the diffuse component, i.e., signal n.

The proposed QCQP-MWF-LC presents the lowest values for the $\Delta\mathrm{ILD}_{\mathrm{n}}$, $\Delta\mathrm{IPD}_{\mathrm{n}}$ and $\Delta\mathrm{MSC}_{\mathrm{n}}$. This result suggests an adequate preservation of the diffuseness perception of the noise component. However, QCQP-MWF-LC has a significant performance loss in terms of SNR compared to those of MWF and SDP-MWF-LC.

The proposed SDP-MWF-LC has a performance similar to that of the MWF. This suggests that this technique changes the perception of the diffuse noise component to a pointwise source whose direction is the same as the speech source.

Figure 26 presents ILD errors (Figure 26(a)) and IPD (Figure 26(b)) of the interfering source as a function of θ_{u}, as well as ΔSINR (Figure 26(c)), for the acoustic

Table 8 – Acoustic Scenario AS_1 – Variation of the binaural cues of the diffuse noise component when processed by the MWF, QCQP-MWF-LC, SDP-MWF-LC.

	MWF	QCQP-MWF-LC	SDP-MWF-LC
ΔILD_n [dB]	0,895	0,114	0,818
ΔIPD_n [rad/π]	0,628	0,040	0,628
ΔMSC_n	0,980	0,004	0,980
ΔSNR [dB]	11,47	7,42	11,77

scenario AS_2. The MWF presents the highest errors in the binaural cues of the interfering source in all analyzed situations. This result corroborates already known theoretical and practical results and indicates the apparent displacement of the interfering source to the position of the speech source, i.e., $\theta_u = \theta_x = 0^\circ$. SDP-MWF-LC significantly reduces ILD and IPD errors compared to MWF. At azimuth $\theta_u = -60^\circ$, the difference of ΔILD_u between MWF and SDP-MWF-LC and between MWF and QCQP-MWF-LC is 9.48 and 7.68, respectively. Compared to QCQP-MWF-LC, SDP-MWF-LC has a greater reduction in ILD and IPD errors. Analyzing the $\Delta SINR$, the SDP-MWF-LC presents a performance approximately equal to that of the MWF, while the QCQP-MWF-LC reaches, on average, 1.5 dB less compared to the SDP-MWF-LC- IC. ΔSNR plots have been omitted as they are similar to the SINR plots, and therefore the conclusions are the same.

The simulations presented were performed on a desktop computer, with an Intel Core i7-3770 processor (*clock* 3.40 GHz) and using Matlab™. Table 9 presents the processing times of the three techniques analyzed. We found that the SPD-MWF-LC presents a significant reduction in processing time compared to the QCQP-MWF-LC.

Table 9 – Optimization time for the techniques MWF, QCQP-MWF-LC e SDP-MWF-LC.

	MWF	QCQP-MWF-LC	SDP-MWF-LC
time [s]	9×10^{-5}	1,319	0,074

5.9 CONCLUSION

This chapter focuses on presenting various MWF-based techniques with the aim of preserving the ILD and IC of undesired sources. The first technique discussed involves a cost function for noise reduction, which employs a penalty term to maintain the dominant binaural cues of signals generated by both pointwise sources and diffuse fields. Previous cost functions proposed in the literature have only been effective in preserving the acoustic characteristics of one of these scenarios. In contrast, the proposed cost function achieves preservation of relevant acoustic features using a single penalty term, thus combining the objectives of prior works. The results demonstrate that this new cost function preserves the acoustic characteristics of both pointwise sources and diffuse sound fields and achieves superior quantitative performance compared to state-of-the-art cost functions used to

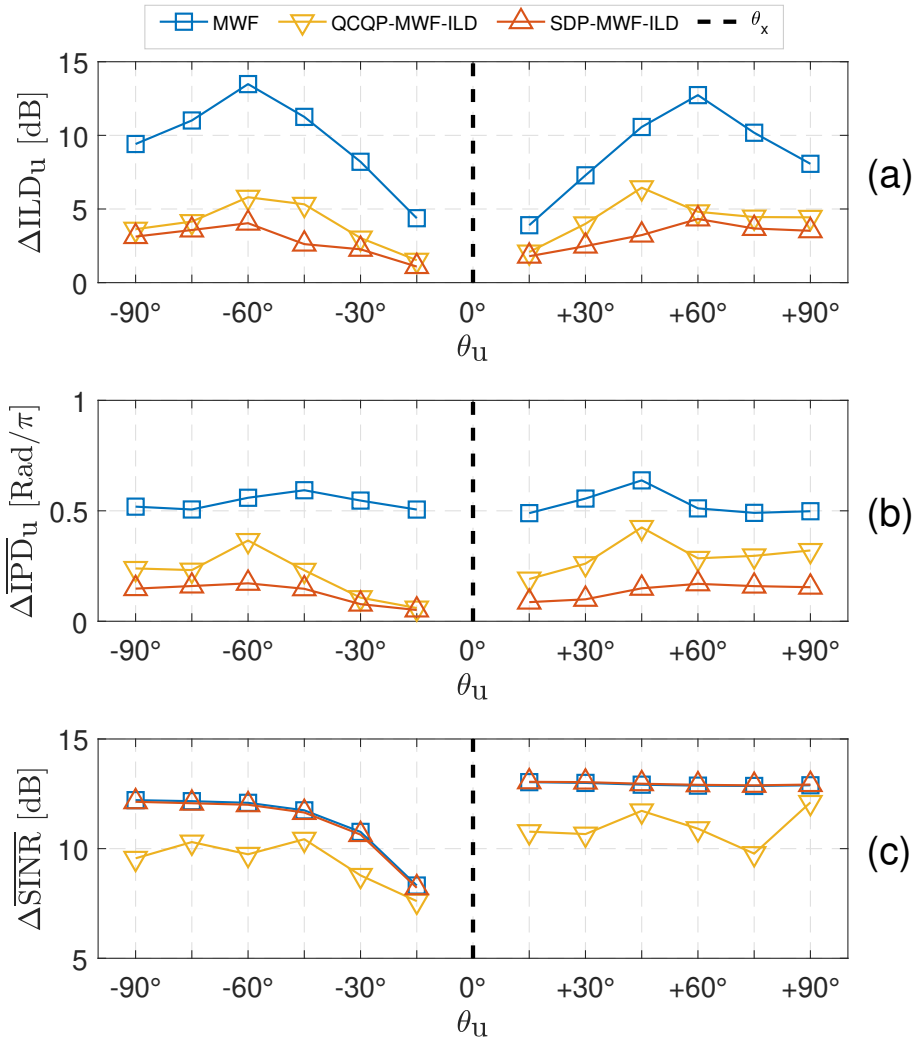


Figure 26 – Acoustic Scenario AS_2 – Measures for the interfering source: (a) ΔILD_u, (b) ΔIPD_u, (c) ΔSINR. Processing techniques: (i) MWF (–□–), (ii) QCQP-MWF-LC (– ▽ –), (iii) SDP-MWF-LC (– △ –).

preserve each of these scenarios individually. The second proposition consists of two parts: first, a reformulation of the MWF-ILD-IC as a quadratically constrained quadratic problem (QCQP-MWF-LC); and second, a formulation of the problem as a semidefinite programming (SDP-MWF-LC). The QCQP-MWF-LC facilitates the setup of the desired ILD and IC errors, while the SDP-MWF-LC is also a convex and computationally efficient approach to calculate an approximate solution to the MWF-ILD-IC technique. Results obtained from simulated acoustic scenarios indicate the ability of the QCQP-MWF-LC in preserving the binaural cues of a diffuse field. Additionally, the QCQP-MWF-LC technique reduces the errors in the ILD and IPD compared to the errors presented by the MWF. Finally, the SDP-MWF-LC also reduces the errors in the ILD and IPD of a pointwise source, however it distorts the binaural cues of a diffuse field, presenting a performance equivalent to the MWF.

6 CONCLUSION

This thesis presents new contributions to noise reduction with binaural cue preservation using the multichannel Wiener filter (MWF) for applications in binaural hearing aids. The first contribution is related to techniques based on the augmented MWF cost function, with terms related to the preservation of binaural cues. It proposed a strategy for adjusting the parameters that is robust to power variations in the input signals. The second proposal is a closed-form solution to the MWF with preservation of the interaural level difference (ILD). The last proposal is a set of techniques that can be used for the preservation of the ILD and the interaural coherence (IC). Following, the main contributions of this thesis are presented.

6.1 ROBUST PARAMETER STRATEGY FOR WIENER-BASED BINAURAL NOISE REDUCTION METHODS

A strategy for adjusting the parameter that weights the terms of the MWF augmented cost function is presented in Chapter 3. The strategy compensates for joint variations in the speech and noise power (the Lombard effect) and for signal-to-noise (SNR) variations. Objective and psychoacoustic performance measures show that the original MWF-ITF changes its setpoint with variations of speech and noise powers. As a practical effect of this behavior, the perceived position of the point noise source may change toward the speech position. This behavior is perceived in listening experiments conducted by volunteers. The proposed technique guarantees the preservation of the setpoint for variations in the power of the input signals. As a practical consequence, the perceived position of the noise source is maintained.

6.2 CLOSED-FORM SOLUTION TO THE MULTICHANNEL CHANNEL WIENER FILTER WITH INTERAURAL LEVEL DIFFERENCE PRESERVATION

A closed-form solution to MWF with preservation of ILD is presented in Chapter 4. The experiments considered practical and challenging acoustic scenarios with a speech source, point interfering source, and background diffuse noise. The environment also includes a low level of reverberation. Due to the estimation errors, the proposed technique presented non-null errors in the ΔILD; however, it is significantly lower than the errors presented by the MWF. The proposed closed-form solution to the MWF-ILD presented noise reduction performance similar to the MWF. Therefore, the proposed technique preserves the ILD binaural cue, with no practical reduction in noise reduction performance.

6.3 MULTICHANNEL WIENER FILTER TECHNIQUES FOR PRESERVATION OF POINT SOURCE AND DIFFUSE SOUND FIELDS

Three binaural MWF techniques are presented for the preservation of the ILD and IC, leading to the preservation of binaural cues for signals generated by point sources or diffuse fields. The first proposal is a term for binaural preservation for the preservation of the ILD and IC. The results show that the proposal has equivalent performance compared to the MWF with ITF preservation, which is the state-of-the-art technique in acoustic scenarios with a single point interfering/noise source. In acoustic scenarios in which the prevalent acoustic noise is generated by a diffused field, the proposed technique presents performance equivalent to the MWF with IC preservation, which is the state-of-the-art MWF technique for preservation of the perception of a diffuse component. The second proposal is an approximation to the MWF-ILD-IC technique reformulated as a constrained optimization problem. This technique presented an easier way, compared to the original MWF-ILD-IC, to set its parameters and avoid distortions in the binaural cues. The results show that the proposed technique presents good performance in preserving the ILD and IC of the two investigated acoustic scenarios. Finally, the third technique is based on a relaxation of the constrained problem. The goal of this new algorithm is to reduce the computational time required to compute the noise reduction filters. The results show that the reformulation presents a better performance for the preservation of ILD and IC compared with the constrained formulation of the MWF-ILD-IC for signals generated by a point source. However, in the case of a diffuse field, the proposed algorithm presents poor performance, since it presents ILD and IC errors similar to the MWF.

6.4 PUBLISHED PAPERS

This section presents submitted and published works during the period of this thesis.

6.4.1 Journals

- **DO CARMO, D.M.**; BORSOI, R.A.; COSTA, M.H. Robust parameter strategy for Wiener-based binaural noise reduction methods in hearing aids. Biomedical Signal Processing and Control, v. 74, p. 103461, 2022.
- **DO CARMO, D.M.**; BORSOI, R.A.; COSTA, M.H. Closed-Form Solution to the Multichannel Wiener Filter with Interaural Level Difference Preservation. (submitted)

6.4.2 Conference proceedings

- **DO CARMO, D.M.**; BORSOI, R.A.; COSTA, M.H. Uma Abordagem Convexa para o Filtro de Wiener Multicanal com Preservação de Pistas Biauriculares em Aparelhos Auditivos. XXVII Congresso Brasileiro de Engenharia Biomédica (CBEB) e IX Congresso Brasileiro de Engenharia Biomédica, 2022.
- **DO CARMO, D.M.**; BORSOI, R.A.; COSTA, M.H. Filtro multicanal de Wiener com restrição quadrática para preservaçao de pistas biauriculares em aparelhos auditivos. Simp. Bras. Telecom. Proc. Sinais, p. 1-5, 2021.
- **DO CARMO, D.M.**; BORSOI, R.A.; COSTA, M.H. Proposta de um parâmetro robusto para redução de ruído e preservação espacial em aparelhos auditivos biauriculares. Simp. Bras. Telecom. Proc. Sinais, p. 1-5, 2021.
- **DO CARMO, D.M.**; BORSOI, R.A.; COSTA, M.H. Proposta de uma função custo para a preservação espacial de fontes sonoras em aparelhos auditivos biauriculares. Simp. Bras. Telecom. Proc. Sinais, p. 1-5, 2020.
- REYS, A.D.; **DO CARMO, D.M.**; COSTA, M.H. et al. Implementação em tempo real de um sistema de redução de ruído binaural com preservação da função de transferência interaural. Simp. Bras. Telecom. Proc. Sinais, p. 1-5, 2019.

6.4.3 Awards

- Cândido Pinto de Melo Award (3º place) - XXV Congresso Brasileiro de Engenharia Biomédica, Sociedade Brasileira de Engenharia Biomédica, 2016. *Uma Abordagem Convexa para o Filtro de Wiener Multicanal com Preservação de Pistas Biauriculares em Aparelhos Auditivos.*

REFERENCES

1 LIN, Frank R; FERRUCCI, Luigi. Hearing loss and falls among older adults in the United States. **Archives of internal medicine**, American Medical Association, v. 172, n. 4, p. 369–371, 2012.

2 RUTHERFORD, Bret R; BREWSTER, Katharine; GOLUB, Justin S; KIM, Ana H; ROOSE, Steven P. Sensation and psychiatry: linking age-related hearing loss to late-life depression and cognitive decline. **American Journal of Psychiatry**, Am Psychiatric Assoc, v. 175, n. 3, p. 215–224, 2018.

3 CHIEA, Rafael Attili; COSTA, Márcio Holsbach; BARRAULT, Guillaume. New insights on the optimality of parameterized Wiener filters for speech enhancement applications. **Speech Communication**, Elsevier, v. 109, p. 46–54, 2019.

4 HU, Yi; LOIZOU, Philipos C. Evaluation of objective quality measures for speech enhancement. **IEEE Transactions on audio, speech, and language processing**, IEEE, v. 16, n. 1, p. 229–238, 2007.

5 JAZI, Nima Yousefian; LOIZOU, Philipos C. **Method and system for enhancing the intelligibility of sounds relative to background noise**. [S.l.]: Google Patents, July 2014. US Patent App. 13/990,942.

6 IBGE. **Censo demográfico 2010: características gerais da população, religião e pessoas com deficiência**. [S.l.: s.n.], 2012.

7 OMS et al. Addressing the rising prevalence of hearing loss, 2018.

8 EHIMA. **Hearing Aids improve Hearing - and a LOT more, Trends derived from the EuroTrak databases 2009 - 2020**. [S.l.: s.n.], June 2020.

9 POPELKA, G.R.; MOORE, B.C.J.; FAY, R.R.; POPPER, A.N. **Hearing aids**. [S.l.: s.n.], 2016.

10 BAUMAN, N. The hearing aids of yesteryear-a brief history of hearing aids from then to now. **Signal, the journal of the Association of Hearing Instrument Practitioners of Ontario**, v. 100, p. 22–28, 2014.

11 WU, Y.; STANGL, E.; CHIPARA, O.; HASAN, S.S.; DEVRIES, S.; OLESON, J. Efficacy and effectiveness of advanced hearing aid directional and noise reduction technologies for older adults with mild to moderate hearing loss. **Ear and hearing**, v. 40, n. 4, p. 805–822, 2019.

12 WONG, L.L.; CHEN, Y.; WANG, Q.; KUEHNEL, V. Efficacy of a hearing aid noise reduction function. **Trends in hearing**, v. 22, p. 2331216518782839, 2018.

13 LI, Ning; LOIZOU, Philipos C. Factors influencing intelligibility of ideal binary-masked speech: Implications for noise reduction. **The Journal of the Acoustical Society of America**, v. 123, n. 3, p. 1673–1682, 2008.

14 HU, Yi; LOIZOU, Philipos C. Subjective comparison and evaluation of speech enhancement algorithms. **Speech communication**, v. 49, n. 7-8, p. 588–601, 2007.

15 GELFAND, S.A. **Hearing: An Introduction to Psychological and Physiological Acoustics, Sixth Edition**. Sixth edition. [S.l.]: CRC Press, 2018.

16 BLAUERT, Jens. **Spatial hearing: the psychophysics of human sound localization**. [S.l.: s.n.], 1997.

17 CHING, T.Y.C.; VAN WANROOY, E.; DILLON, H.; CARTER, L. Spatial release from masking in normal-hearing children and children who use hearing aids. **The Journal of the Acoustical Society of America**, v. 129, n. 1, p. 368–375, 2011.

18 YOST, W.A. Spatial release from masking based on binaural processing for up to six maskers. **The Journal of the Acoustical Society of America**, v. 141, n. 3, p. 2093–2106, 2017.

19 WAGNER, Luise; GEILING, Lukas; HAUTH, Christopher; HOCKE, Thomas; PLONTKE, Stefan; RAHNE, Torsten. Improved binaural speech reception thresholds through small symmetrical separation of speech and noise. **PloS one**, v. 15, n. 8, p. 1–10, 2020.

20 AVNI, Amir; RAFAELY, Boaz. Interaural cross correlation and spatial correlation in a sound field represented by spherical harmonics. In: 7. FIRST International Symposium on Ambisonics and Spherical Acoustics (Ambisonics 2009). Graz, Austria. [S.l.: s.n.], 2009.

21 KÄSBACH, Johannes; MARSCHALL, Marton; EPP, Bastian; DAU, Torsten. The relation between perceived apparent source width and interaural cross-correlation in sound reproduction spaces with low reverberation. In: 39TH German Annual Conference on Acoustics. [S.l.: s.n.], 2013.

22 GORODENSKY, J.H.; ALEMU, R.Z.; GILL, S.S.; SANDOR, M.T.; PAPSIN, B.C.; CUSHING, S.L.; GORDON, K.A. Binaural hearing is impaired in children with hearing loss who use bilateral hearing aids. **The Journal of the Acoustical Society of America**, v. 146, n. 6, p. 4352–4362, Dec. 2019.

23 VAN DEN BOGAERT, Tim; KLASEN, Thomas J; MOONEN, Marc; DEUN, Van; WOUTERS, Jan. Horizontal localization with bilateral hearing aids: Without is better than with. **Citation: The Journal of the Acoustical Society of America**, v. 119, p. 515, 2006.

24 BENESTY, J.; CHEN, J.; HUANG, Y.; COHEN, I. **Noise reduction in speech processing**. [S.l.: s.n.], 2009. v. 2.

25 DOCLO, S.; GANNOT, S.; MARQUARDT, D.; HADAD, E. Binaural Speech Processing with Application to Hearing Devices. **Audio Source Separation and Speech Enhancement**, p. 413–442, 2018.

26 CORNELIS, B.; DOCLO, S.; VAN DEN BOGAERT, T.; MOONEN, M.; WOUTERS, J. Theoretical analysis of binaural multimicrophone noise reduction techniques. **IEEE Transactions on Audio, Speech, and Language Processing**, v. 18, n. 2, p. 342–355, 2009.

27 DOCLO, S.; DONG, Rong; KLASEN, Thomas J; WOUTERS, Jan; HAYKIN, Simon; MOONEN, Marc. Extension of the multi-channel Wiener filter with localisation cues for noise reduction in binaural hearing aids. In: PROC. IWAENC. [S.l.: s.n.], 2005. P. 221–224.

28 DOCLO, S.; KLASEN, Thomas J; VAN DEN BOGAERT, Tim; WOUTERS, Jan; MOONEN, Marc. Theoretical analysis of binaural cue preservation using multi-channel Wiener filtering and interaural transfer functions. In: PROC. Int. Workshop Acoust. Echo Noise Control (IWAENC). [S.l.: s.n.], 2006. P. 1–4.

29 MARQUARDT, D.; HADAD, E.; GANNOT, S.; DOCLO, S. Theoretical analysis of linearly constrained multi-channel Wiener filtering algorithms for combined noise reduction and binaural cue preservation in binaural hearing aids. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, v. 23, n. 12, p. 2384–2397, 2015.

30 MARQUARDT, D.; HOHMANN, V.; DOCLO, S. Interaural coherence preservation in multi-channel Wiener filtering-based noise reduction for binaural hearing aids. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, v. 23, n. 12, p. 2162–2176, 2015.

31 LOMBARD, Etienne. Le signe de l'elevation de la voix. **Ann. Mal. de L'Oreille et du Larynx**, p. 101–119, 1911.

32 ITTURRIET, F.P.; COSTA, M.H. Perceptually relevant preservation of interaural time differences in binaural hearing aids. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, v. 27, n. 4, p. 753–764, 2019.

33 MARQUARDT, Daniel; DOCLO, Simon. Interaural coherence preservation for binaural noise reduction using partial noise estimation and spectral postfiltering. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, IEEE, v. 26, n. 7, p. 1261–1274, 2018.

34 WERNER, Johnny; COSTA, Marcio Holsbach. A Noise-Reduction Method With Coherence Enhancement for Binaural Hearing Aids. **Journal of Communication and Information Systems**, v. 35, n. 1, p. 338–348, 2020.

35 COSTA, M.H.; NAYLOR, P.A. ILD preservation in the multichannel Wiener filter for binaural hearing aid applications. In: IEEE. 2014 22nd European Signal Processing Conference (EUSIPCO). [S.l.: s.n.], 2014. P. 636–640.

36 BRAUN, S.; TORCOLI, M.; MARQUARDT, D.; HABETS, E.A.P.; DOCLO, S. Multichannel dereverberation for hearing aids with interaural coherence preservation. In: 2014 14th International Workshop on Acoustic Signal Enhancement (IWAENC). [S.l.: s.n.], 2014. P. 124–128.

37 CORDIOLI, J.A. **Implementação de um sistema para ensaios subjetivos de percepção sonora com separação espacial voltada para candidatos e usuários de próteses auditivas**. 2019. MA thesis – Universidade Federal de Santa Catarina (UFSC).

38 HEARING LINK. **How the ear works**. [S.l.: s.n.], 2020. Acessado em junho de 2020. Available from: `https://www.hearinglink.org/your-hearing/about-hearing/how-the-ear-works/`.

39 FEDDERSEN, WE; SANDEL, TT; TEAS, DC; JEFFRESS, LA. Localization of high-frequency tones. **the Journal of the Acoustical Society of America**, v. 29, n. 9, p. 988–991, 1957.

40 HARTMANN, William M; CONSTAN, Zachary A. Interaural level differences and the level-meter model. **The Journal of the Acoustical Society of America**, v. 112, n. 3, p. 1037–1045, 2002.

41 BLUM, Konrad; ROOYEN, Gert-Jan van; ENGELBRECHT, H. Spatial audio to assist speaker identification in telephony. **Proc. IWSSIP**, p. 1–4, 2010.

42 KLOCKGETHER, S.; VAN DE PAR, S. Just noticeable differences of spatial cues in echoic and anechoic acoustical environments. **The Journal of the Acoustical Society of America**, v. 140, n. 4, el352–el357, 2016.

43 SCHARFF, A. **Chapter 12: Sound Localization and the Auditory Scene**. [S.l.: s.n.], 2020. Pagina 20 do material online. Acessado em setembro de 2020. Available from: `https://courses.washington.edu/psy333/lecture_pdfs/Week9_Day2.pdf`.

44 GANNOT, Sharon; VINCENT, Emmanuel; MARKOVICH-GOLAN, Shmulik; OZEROV, Alexey. A consolidated perspective on multimicrophone speech enhancement and source separation. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, IEEE, v. 25, n. 4, p. 692–730, 2017.

45 CORNELIS, B.; MOONEN, M.; WOUTERS, J. Performance analysis of multichannel Wiener filter-based noise reduction in hearing aids under second order statistics estimation errors. **IEEE Transactions on Audio, Speech, and Language Processing**, v. 19, n. 5, p. 1368–1381, 2010.

46 MARQUARDT, Daniel; HOHMANN, Volker; DOCLO, Simon. Coherence preservation in multi-channel Wiener filtering based noise reduction for binaural hearing aids. In: IEEE. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. [S.l.: s.n.], 2013. P. 8648–8652.

47 DOCLO, S.; GANNOT, S.; MARQUARDT, D.; HADAD, E. Binaural Speech Processing with Application to Hearing Devices. In: AUDIO Source Separation and Speech Enhancement. [S.l.]: John Wiley & Sons, Ltd, 2018. chap. 18, p. 413–442.

48 CHAUMETTE, Eric; VILÀ-VALLS, Jordi; VINCENT, François. On the general conditions of existence for linear MMSE filters: Wiener and Kalman. **Signal Processing**, Elsevier, v. 184, p. 108052, 2021.

49 PENROSE, Roger. A generalized inverse for matrices. In: CAMBRIDGE UNIVERSITY PRESS, 3. MATHEMATICAL proceedings of the Cambridge philosophical society. [S.l.: s.n.], 1955. P. 406–413.

50 CARMO, D.M. do; COSTA, M.H. Online approximation of the multichannel Wiener filter with preservation of interaural level difference for binaural hearing-aids. **Computers in biology and medicine**, v. 95, p. 188–197, 2018.

51 VAN DEN BOGAERT, Tim; WOUTERS, Jan; DOCLO, Simon; MOONEN, Marc. Binaural cue preservation for hearing aids using an interaural transfer function multichannel Wiener filter. In: 2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07. [S.l.: s.n.], 2007. P. iv–565.

52 WU, Y.; STANGL, E.; CHIPARA, O.; HASAN, S. S.; WELHAVEN, A.; OLESON, J. Characteristics of real-world signal-to-noise ratios and speech listening situations of older adults with mild-to-moderate hearing loss. **Ear and Hearing**, v. 39, n. 2, p. 293, 2018.

53 ZOLLINGER, S.A.; BRUMM, H. The evolution of the Lombard effect: 100 years of psychoacoustic research. **Behaviour**, v. 148, n. 11-13, p. 1173–1198, 2011.

54 LUO, J.; HAGE, S.R.; MOSS, C.F. The Lombard effect: From acoustics to neural mechanisms. **Trends in neurosciences**, v. 41, n. 12, p. 938–949, 2018.

55 JUNQUA, Jean-Claude. The Lombard reflex and its role on human listeners and automatic speech recognizers. **The Journal of the Acoustical Society of America**, v. 93, n. 1, p. 510–524, 1993.

56 HANSEN, John HL. Analysis and compensation of speech under stress and noise for environmental robustness in speech recognition. **Speech communication**, v. 20, n. 1-2, p. 151–173, 1996.

57 KIM, J.Y.; NAM, K.W.; LEE, J.C.; HWANG, J.H.; JANG, D.P.; KIM, I.Y. Scalp tapping-based protocol for adjusting the parameters of binaural hearing aids. **Biomedical Signal Processing and Control**, v. 45, p. 91–97, 2018.

58 CHISAKI, Yoshifumi; MATSUO, Kotaro; HAGIWARA, Katsumori; NAKASHIMA, Hidetoshi; USAGAWA, Tsuyoshi. Real-time processing using the frequency domain binaural model. **Applied Acoustics**, v. 68, n. 8, p. 923–938, 2007.

59 MARXER, R.; BARKER, J.; ALGHAMDI, N.; MADDOCK, S. The impact of the Lombard effect on audio and visual speech recognition systems. **Speech communication**, v. 100, p. 58–68, 2018.

60 HADAD, E.; DOCLO, S.; GANNOT, S. The binaural LCMV beamformer and its performance analysis. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, v. 24, n. 3, p. 543–558, 2016.

61 BRICKELL, F. A theorem on homogeneous functions. **Journal of the London Mathematical Society**, v. 1, n. 1, p. 325–329, 1967.

62 CHEN, J.; BENESTY, J.; HUANG, Y.; DOCLO, S. New insights into the noise reduction Wiener filter. **IEEE Transactions on audio, speech, and language processing**, v. 14, n. 4, p. 1218–1234, 2006.

63 KAYSER, H.; EWERT, S. D.; ANEMÜLLER, J.; ROHDENBURG, T.; HOHMANN, V.; KOLLMEIER, B. Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. **Eurasip Journal on Advances in Signal Processing**, v. 2009, 2009.

64 HENTER, G.E.; MERRITT, T.; SHANNON, M.; MAYO, C.; KING, S., et al. Repeated Harvard Sentence Prompts corpus version 0.5. University of Edinburgh, Centre for Speech Technology Research; Cambridge, 2014.

65 DRESCHLER, Wouter A; VERSCHUURE, Hans; LUDVIGSEN, Carl; WESTERMANN, Søren. ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment: Ruidos ICRA: Señates de ruido artificial con espectro similar al habla y propiedades temporales para pruebas de instrumentos auditivos. **Audiology**, v. 40, n. 3, p. 148–157, 2001.

66 HORN, Roger A; JOHNSON, Charles R. **Matrix analysis**. [S.l.]: Cambridge university press, 2012.

67 WALDEN, Brian E; SURR, Rauna K; CORD, Mary T; DYRLUND, Ole. Predicting hearing aid microphone preference in everyday listening. **Journal of the American Academy of Audiology**, American Academy of Audiology, v. 15, n. 05, p. 365–396, 2004.

68 SMEDS, Karolina; WOLTERS, Florian; RUNG, Martin. Estimation of signal-to-noise ratios in realistic sound scenarios. **Journal of the American Academy of Audiology**, v. 26, n. 2, p. 183–196, 2015.

69 RASPAUD, Martin; VISTE, Harald; EVANGELISTA, Gianpaolo. Binaural source localization by joint estimation of ILD and ITD. **Ieee transactions on audio, speech, and language processing**, v. 18, n. 1, p. 68–77, 2010.

70 BOYD, Stephen; BOYD, Stephen P; VANDENBERGHE, Lieven. **Convex optimization**. [S.l.]: Cambridge university press, 2004.

71 BECK, Amir; ELDAR, Yonina C. Strong duality in nonconvex quadratic optimization with two quadratic constraints. **SIAM Journal on optimization**, v. 17, n. 3, p. 844–860, 2006.

72 LIN, Xianming; HUANG, Yongwei; MA, Wing-Kin. Robust Downlink Transmit Optimization Under Quantized Channel Feedback via the Strong Duality for QCQP. **IEEE Signal Processing Letters**, IEEE, v. 28, p. 1–5, 2020.

73 CASTRO-GONZÁLEZ, N; MARTÍNEZ-SERRANO, MF; ROBLES, J. Expressions for the Moore–Penrose inverse of block matrices involving the Schur complement. **Linear Algebra and its Applications**, Elsevier, v. 471, p. 353–368, 2015.

74 KOHNO, Kiyotaka; INOUYE, Yujiro; KAWAMOTO, Mitsuru. A matrix pseudo-inversion lemma for positive semidefinite hermitian matrices and its application to adaptive blind deconvolution of MIMO systems. **IEEE Transactions on Circuits and Systems I: Regular Papers**, IEEE, v. 55, n. 1, p. 424–435, 2008.

75 HABETS, E.A.P.; COHEN, I.; GANNOT, S. Generating nonstationary multisensor signals under a spatial coherence constraint. **The Journal of the Acoustical Society of America**, v. 124, n. 5, p. 2911–2917, 2008.

76 CROCHIERE, R. A weighted overlap-add method of short-time Fourier analysis/synthesis. **IEEE Transactions on Acoustics, Speech, and Signal Processing**, v. 28, n. 1, p. 99–102, 1980.

77 ZHANG, Jie; LI, Changheng. Quantization-Aware Binaural MWF Based Noise Reduction Incorporating External Wireless Devices. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, IEEE, v. 29, p. 3118–3131, 2021.

78 GÖSSLING, Nico; HADAD, Elior; GANNOT, Sharon; DOCLO, Simon. Binaural LCMV beamforming with partial noise estimation. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, IEEE, v. 28, p. 2942–2955, 2020.

79 GÖSSLING, Nico; MARQUARDT, Daniel; MERKS, Ivo; ZHANG, Tao; DOCLO, Simon. Optimal binaural LCMV beamforming in complex acoustic scenarios: Theoretical and practical insights. In: IEEE. 2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC). [S.l.: s.n.], 2018. P. 381–385.

80 PERTILÄ, Pasi; FAGERLUND, Eemi; HUTTUNEN, Anu; MYLLYLÄ, Ville. Online Own Voice Detection for a Multi-Channel Multi-Sensor In-Ear Device. **IEEE Sensors Journal**, IEEE, v. 21, n. 24, p. 27686–27697, 2021.

81 MARKOVICH, Shmulik; GANNOT, Sharon; COHEN, Israel. Multichannel eigenspace beamforming in a reverberant noisy environment with multiple interfering speech signals. **IEEE Transactions on Audio, Speech, and Language Processing**, IEEE, v. 17, n. 6, p. 1071–1086, 2009.

82 SERIZEL, Romain; MOONEN, Marc; VAN DIJK, Bas; WOUTERS, Jan. Low-rank approximation based multichannel Wiener filter algorithms for noise reduction with application in cochlear implants. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, IEEE, v. 22, n. 4, p. 785–799, 2014.

83 MARKOVICH-GOLAN, Shmulik; GANNOT, Sharon; KELLERMANN, Walter. Performance analysis of the covariance-whitening and the covariance-subtraction methods for estimating the relative transfer function. In: IEEE. 2018 26th European Signal Processing Conference (EUSIPCO). [S.l.: s.n.], 2018. P. 2499–2503.

84 MATHWORKS. **fmincon - find a minimum of a constrained nonlinear multivariable function.** [S.l.: s.n.]. Available from: `https://www.mathworks.com/help/optim/ug/fmincon.html`.

85 ITU-T. **Series p: Telephone transmission quality, telephone installations, local line networks-subjective video quality assessment methods for multimedia applications - Test Vectors Associated to Rec. ITU-T P.50 Appendix I**. [S.l.: s.n.], 1998. Available from: `https://www.itu.int/net/itu-t/sigdb/genaudio/AudioForm-g.aspx?val=1000050`.

86 LINDEROTH, Jeff. A simplicial branch-and-bound algorithm for solving quadratically constrained quadratic programs. **Mathematical programming**, Springer, v. 103, p. 251–282, 2005.

87 STURM, Jos F. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. **Optimization methods and software**, Taylor & Francis, v. 11, n. 1-4, p. 625–653, 1999.

88 LOFBERG, Johan. YALMIP: A toolbox for modeling and optimization in MATLAB. In: IEEE. 2004 IEEE international conference on robotics and automation (IEEE Cat. No. 04CH37508). [S.l.: s.n.], 2004. P. 284–289.

Appendix

APPENDIX A – PROOF OF THEOREM I

Decompose $\mathbf{\Phi}_y$ as

$$\mathbf{\Phi}_y = \mathbf{Q}_y \mathbf{\Lambda}_y \mathbf{Q}_y^{\mathrm{H}} \tag{217}$$

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 $\mathbf{\Phi}_y \geq 0$ then (217)

$$\mathbf{\Phi}_y = \begin{bmatrix} \mathbf{C}_y \\ \mathbf{N}_y \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_y & \mathbf{0} \\ \mathbf{0} & \mathbf{N}_y \end{bmatrix} \begin{bmatrix} \mathbf{C}_y^{\mathrm{H}} & \mathbf{N}_y^{\mathrm{H}} \end{bmatrix} \tag{218}$$

in which $\mathbf{\Sigma}_y$ is a diagonal matrix of order $T \leqslant M$, 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 (17), (18), 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

$$\mathbf{b} = \begin{bmatrix} \mathbf{C}_y & \mathbf{N}_y \end{bmatrix} \begin{bmatrix} \mathbf{b}_y \\ \mathbf{0}_{(M-T)\times 1} \end{bmatrix} \tag{219}$$

in which $\mathbf{b}_y$ is the component of $\mathbf{b}$ in the column space of y. Using (218) and (219), (129) can be defined as:

$$\mathbf{\Phi}_{yu}(\beta_r(\tau_j)) = \mathbf{Q}_y \begin{bmatrix} \mathbf{\Sigma}_y + \beta_r(\tau_j) p_{su} \mathbf{b}_y \mathbf{b}_y^{\mathrm{H}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{(M-T)\times(M-T)} \end{bmatrix} \mathbf{Q}_y^{\mathrm{H}} \tag{220}$$

Because $\mathbf{Q}_y$ is invertible, pre- and post-multiplication of (220) by $\mathbf{Q}_y$ and its conjugate transpose preserves its positive definiteness characteristic [66]. Thus, the block diagonal structure of the inner matrix in (220) implies that $\mathbf{\Phi}_{yu}(\beta_r(\tau_j))$ is positive semi-definite if and only if

$$\mathbf{\Sigma}_y + \beta_r(\tau_j) p_{su} \mathbf{b}_y \mathbf{b}_y^{\mathrm{H}} \geq 0, \tag{221}$$

which can be expressed as:

$$\mathbf{\Sigma}_y^{1/2} (\mathbf{I} + \beta_r(\tau_j) p_{su} \mathbf{\Sigma}_y^{-1/2} \mathbf{b}_y \mathbf{b}_y^{\mathrm{H}} \mathbf{\Sigma}_y^{-1/2}) \mathbf{\Sigma}_y^{1/2} \geq 0. \tag{222}$$

Because the columns of $\mathbf{C}_y$ in (219) have unity norm, and are orthogonal and linearly independent vectors, the following relation between $\mathbf{b}_y$ and $\mathbf{b}$ can be established:

$$\mathbf{b}_y = \mathbf{C}_y^{\mathrm{H}} \mathbf{b}. \tag{223}$$

Pre-multiplying both sides of (223) by $\mathbf{\Sigma}_y^{-1/2}$ leads to:

$$\mathbf{\Sigma}_y^{-1/2} \mathbf{b}_y = \mathbf{\Sigma}_y^{-1/2} \mathbf{C}_y^{\mathrm{H}} \mathbf{b}. \tag{224}$$

Using (224) in (222), leads to:

$$\boldsymbol{\Sigma}_y^{1/2}(\mathbf{I} + \beta_r(\tau_j)p_{su}\boldsymbol{\Sigma}_y^{-1/2}\mathbf{C}_y^{\mathrm{H}}\mathbf{b}\mathbf{b}^{\mathrm{H}}\mathbf{C}_y\boldsymbol{\Sigma}_y^{-1/2})\boldsymbol{\Sigma}_y^{1/2} \succeq 0. \tag{225}$$

The matrix inside the parenthesis in (222) is an identity matrix plus a rank-one matrix. Thus, its eigenvalues are 1 ($T-1$ times) and $1 + \beta_r(\tau_j)p_{su}\mathbf{b}^{\mathrm{H}}\mathbf{C}_y\boldsymbol{\Sigma}_y^{-1}\mathbf{C}_y^{\mathrm{H}}\mathbf{b}$, which can be represented as:

$$\beta_r(\tau_j)p_{su}\mathbf{b}^{\mathrm{H}}\mathbf{C}_y\boldsymbol{\Sigma}_y^{-1}\mathbf{C}_y^{\mathrm{H}}\mathbf{b} = \beta_r(\tau_j)p_{su}\mathbf{b}^{\mathrm{H}}\boldsymbol{\Phi}_y^{\dagger}\mathbf{b} \tag{226a}$$

$$= \beta_r(\tau_j)\eta_b, \tag{226b}$$

in which $\eta_b = p_{su}\mathbf{b}^{\mathrm{H}}\boldsymbol{\Phi}_y^{\dagger}\mathbf{b} = \mathrm{trace}\{\boldsymbol{\Phi}_y^{\dagger}\boldsymbol{\Phi}_u\}$, and the $\boldsymbol{\Phi}_y^{\dagger}$ is the Moore-Penrose inverse of matrix $\boldsymbol{\Phi}_y$, defined as:

$$\boldsymbol{\Phi}_y^{\dagger} = \mathbf{Q}_y \begin{bmatrix} \boldsymbol{\Sigma}_y^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{(M-t)\times(M-t)} \end{bmatrix} \mathbf{Q}_y^{\mathrm{H}}. \tag{227}$$

Therefore, since (222) consists in a congruence relation, matrix $\boldsymbol{\Phi}_{yu}(\beta_r(\tau_j))$ is positive semidefinite if and only if:

$$1 + \beta_r(\tau_j)\eta_b \geqslant 0, \tag{228}$$

considering $\eta_b > 0$, equation (226b) results in

$$\beta_r(\tau_j) \geqslant -\frac{1}{\eta_b}. \tag{229}$$

Replacing (130) in (229) completes the proof.

APPENDIX B – GRADIENT OF $J_{\mathbf{MWF\text{–}ILD}}$

Considering $\mathbf{\Phi}_{yy} = \mathbf{\Phi}_{yy}^{\mathrm{H}}$, the J_{MWF} cost function defined in (38) can be written as:

$$J_{\mathrm{MWF}}(\mathbf{w}) = \Re\{\mathbf{w}^{\mathrm{H}}\mathbf{\Phi}_{yy}\mathbf{w}\} - 2\Re\{\mathbf{w}^{\mathrm{H}}\mathbf{p}_{xx}\} + p_{xx}. \tag{230}$$

Defining

$$\breve{\mathbf{w}} = \begin{bmatrix}\Re\{\mathbf{w}\}\\ \Im\{\mathbf{w}\}\end{bmatrix} = \begin{bmatrix}\mathbf{w}_{\Re}\\ \mathbf{w}_{\Im}\end{bmatrix}, \tag{231}$$

and using (231) in (230), leads to

$$J_{\mathrm{MWF}}(\breve{\mathbf{w}}) = \breve{\mathbf{w}}^{\top}\breve{\mathbf{\Phi}}_{yy}\breve{\mathbf{w}} - 2\breve{\mathbf{w}}^{\mathrm{H}}\breve{\mathbf{p}}_{xx} + p_{xx}. \tag{232}$$

in which

$$\breve{\mathbf{\Phi}}_{yy} = \begin{bmatrix}\Re\{\mathbf{\Phi}_{yy}\} & -\Im\{\mathbf{\Phi}_{yy}\}\\ \Im\{\mathbf{\Phi}_{yy}\} & \Re\{\mathbf{\Phi}_{yy}\}\end{bmatrix}, \tag{233}$$

and

$$\breve{\mathbf{p}}_{xx} = \begin{bmatrix}\Re\{\mathbf{p}_{xx}\}\\ \Im\{\mathbf{p}_{xx}\}\end{bmatrix}. \tag{234}$$

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

$$\nabla_{\breve{\mathbf{w}}} J_{\mathrm{MWF}}(\breve{\mathbf{w}}) = 2\breve{\mathbf{\Phi}}_{yy}\breve{\mathbf{w}} - 2\breve{\mathbf{p}}_{xx}. \tag{235}$$

while its Hessian matrix is given by:

$$\nabla^2_{\breve{\mathbf{w}}} J_{\mathrm{MWF}}(\breve{\mathbf{w}}) = 2\breve{\mathbf{\Phi}}_{yy}. \tag{236}$$

The gradient of J^u_{ILD} in (85) can also be defined as a function of $\breve{\mathbf{w}}$:

$$\nabla_{\breve{\mathbf{w}}} J^u_{\mathrm{ILD}}(\breve{\mathbf{w}}) = \frac{40}{\log_e(10)}\, e_{\mathrm{ILD}}(\breve{\mathbf{w}})\mathbf{p}(\breve{\mathbf{w}}). \tag{237}$$

in which:

$$e_{\mathrm{ILD}}(\breve{\mathbf{w}}) = 10\log_{10}(ILD^u_{\mathrm{ou}}) - 10\log_{10}(ILD^u_{\mathrm{in}}), \tag{238}$$

$$\mathbf{p}(\breve{\mathbf{w}}) = \frac{\breve{\mathbf{\Phi}}_1\breve{\mathbf{w}}}{\breve{\mathbf{w}}^{\top}\breve{\mathbf{\Phi}}_1\breve{\mathbf{w}}} - \frac{\breve{\mathbf{\Phi}}_2\breve{\mathbf{w}}}{\breve{\mathbf{w}}^{\top}\breve{\mathbf{\Phi}}_2\breve{\mathbf{w}}}. \tag{239}$$

for

$$\breve{\mathbf{\Phi}}_1 = \begin{bmatrix}\Re\{\mathbf{\Phi}_u\} & \mathbf{0}_{M\times M}\\ \mathbf{0}_{M\times M} & \mathbf{0}_{M\times M}\end{bmatrix}, \tag{240}$$

and

$$\breve{\mathbf{\Phi}}_2 = \begin{bmatrix}\mathbf{0}_{M\times M} & \mathbf{0}_{M\times M}\\ \mathbf{0}_{M\times M} & \Re\{\mathbf{\Phi}_u\}\end{bmatrix}, \tag{241}$$

The Hessian matrix of the $J^u_{\mathrm{ILD}}(\breve{\mathbf{w}})$ is given by:

$$\nabla^2_{\breve{\mathbf{w}}} J^u_{\mathrm{ILD}}(\breve{\mathbf{w}}) = \frac{40}{\log_e(10)}\left[e_{\mathrm{ILD}}(\breve{\mathbf{w}})\nabla_{\breve{\mathbf{w}}}\mathbf{p}(\breve{\mathbf{w}}) - \nabla_{\breve{\mathbf{w}}} e_{\mathrm{ILD}}(\breve{\mathbf{w}})\mathbf{p}^{\top}(\breve{\mathbf{w}})\right]. \tag{242}$$

APPENDIX C – ITERATIVE ALGORITHM FOR OBTAINING OPTIMUM α AND $\mathbf{w}$ FOR THE UNCONSTRAINED MWF-ILD.

Table 10 – Iterative Search for Finding α and $\mathbf{w}$ in Each Discrete Frequency $k = 1, \cdots, N$

Input: δ^2, $\mathrm{ILD}_{\mathrm{in}}$, $\mathbf{\Phi}_{\mathrm{u}}$, $\mathbf{\Phi}_{\mathrm{y}}$, $\mathbf{\Phi}_{\mathrm{x}}$
Output: $\mathbf{w}$
Set G as the length of vector α
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 $\alpha_i = \alpha(i)$, find the binaural filters solving:
$\mathbf{w}(\alpha_i) = \arg\min \; J_{\mathrm{MWF}}(\lambda, k) + \alpha_i J^{\mathrm{u}}_{\mathrm{ILD}}(\lambda, k)$
$\mathbf{w}_{\mathrm{UNC}} = \mathbf{w}(\alpha_i)$
$e^2_{\mathrm{ILD}} = J^{\mathrm{u}}_{\mathrm{ILD}}(\mathbf{w}(\alpha_i))$

Annex

ANNEX A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado voluntário

As informações contidas nesse termo foram fornecidas por Márcio Holsbach Costa, professor do Departamento de Engenharia Elétrica e Eletrônica da Universidade Federal de Santa Catarina. O objetivo desse documento é informar sobre o trabalho realizado pelo referido pesquisador, para obter uma autorização por escrito, referente à vossa participação espontânea na pesquisa que está sendo realizada. O título desse trabalho é "Desenvolvimento de Técnicas e Algoritmos para Redução de Ruído em Aparelhos Auditivos e Implantes Cocleares". Objetivo: Este trabalho tem como objetivo o estudo, desenvolvimento e análise de desempenho de métodos e algoritmos de processamento de sinais para aparelhos auditivos e implantes cocleares. Metodologia: Sua participação é voluntária e consiste em ouvir sons, emitidos por alto-falantes, fones de ouvido ou moldes auriculares e, em sequência, responder a perguntas sobre características percebidas, como qualidade, inteligibilidade, conforto acústico e espacialidade (direção de chegada). O tempo total do procedimento é de cerca de cinquenta (50) minutos. Dependendo do tipo de experimento poderá ser solicitada a realização de um procedimento de audiometria, realizado por profissionais capacitados, para avaliação da sua condição auditiva. Benefícios: Os benefícios esperados envolvem a produção de conhecimento científico relacionado ao aprimoramento e desenvolvimento de novos métodos de processamento de sinais para aparelhos auditivos e implantes cocleares. Desconfortos e riscos: Os possíveis riscos ou desconfortos decorrentes de sua participação na pesquisa são: o desconforto acústico (devido a volume sonoro elevado), o desconforto mecânico (devido ao uso de moldes auriculares ou fones de ouvido), a fadiga (devido ao tempo de realização do procedimento) e a claustrofobia (quando o procedimento for realizado em ambiente fechado). O desconforto acústico é evitado através do controle (diminuição) de intensidade sonora que poderá ser realizado a qualquer momento. O desconforto mecânico é evitado pelo reposicionamento ou retirada, a qualquer momento, dos moldes auriculares ou fones de ouvido. A fadiga é minimizada pela limitação do tempo total do procedimento e pela possibilidade de pausa a qualquer momento ou de interrupção precoce. O sentimento de claustrofobia é evitado pela saída do ambiente fechado (quando utilizado) a qualquer momento. Outras informações: Não haverá ônus ou bônus financeiro em nenhum momento decorrente da participação nessa pesquisa. O Sr(a) receberá respostas e esclarecimentos a todas as suas perguntas e dúvidas sobre os procedimentos realizados e assuntos relacionados a essa pesquisa por meio do contato com o pesquisador, que assume o compromisso de proporcionar informações atualizadas sobre o estudo. O pesquisador declara que cumprirá as exigências contidas na Resolução CNS 466/2012, que o sigilo e a privacidade dos participantes serão garantidos durantes todas as etapas da pesquisa, inclusive na divulgação dos resultados e que os participantes terão direito ao ressarcimento de eventuais despesas e indenização diante de eventuais

danos produzidos por essa pesquisa. Em caso de necessidade de contato, o endereço do pesquisador responsável é: Departamento de Engenharia Elétrica e Eletrônica, Centro Tecnológico, Bloco E, terceiro andar, Universidade Federal de Santa Catarina, Campus

Universitário João David Ferreira Lima, bairro Trindade, CEP 88.040-900, Florianópolis-SC, telefone (48) 3721-2260, e-mail: costa@eel.ufsc.br. O endereço do Comitê de Ética em Pesquisa da Universidade Federal de Santa Catarina é: Prédio Reitoria II, Rua Desembargador Vitor Lima, número 222, sala 401, Trindade, Florianópolis-SC, CEP 88.040-400, telefone (48) 3721-6094, e-mail: cep.propesq@contato.ufsc.br. CONSENTIMENTO PÓS-INFORMADO

Eu,, portador do RG e CPF , concordo em participar do trabalho "Desenvolvimento de Técnicas e Algoritmos para Redução de Ruído em Aparelhos Auditivos e Implantes Cocleares", desde que seja mantido o sigilo da minha identificação, conforme as normas do Comitê de Ética em Pesquisa com Seres Humanos desta Universidade. A minha participação é voluntária podendo ser cancelada a qualquer momento. Pelo presente consentimento, declaro que fui esclarecido(a) sobre a pesquisa a ser realizada, de forma detalhada, livre de qualquer constrangimento e obrigação, e que recebi uma cópia desse termo, assinada pelos pesquisadores.

Florianópolis, de de .

Assinatura do participante

Assinatura do Pesquisador Principal/ Responsável

Márcio Holsbach Costa RG: 1031584426 SJTC/RS

Elaborado com base na Resolução 466/2012 do CNS.