

DIEGO DA SILVA DE MEDEIROS

**CANCELAMENTO DE SINAIS
INTERFERENTES USANDO ARRANJO DE
MICROFONES: APLICAÇÃO EM APARELHOS
AUDITIVOS**

FLORIANÓPOLIS

2013

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

**CANCELAMENTO DE SINAIS
INTERFERENTES USANDO ARRANJO DE
MICROFONES: APLICAÇÃO EM APARELHOS
AUDITIVOS**

Dissertação submetida à Universidade
Federal de Santa Catarina como parte
dos requisitos para a obtenção do grau
de Mestre em Engenharia Elétrica

DIEGO DA SILVA DE MEDEIROS

Florianópolis, Fevereiro de 2013

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

Medeiros, Diego da Silva de

Cancelamento de sinais interferentes usando arranjo de microfones : Aplicação em aparelhos auditivos / Diego da Silva de Medeiros ; orientador, Leonardo Silva Resende - Florianópolis, SC, 2013.

162 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia Elétrica.

Inclui referências

1. Engenharia Elétrica. 2. Cancelamento de interferências. 3. Arranjo de microfones. 4. Conformador de feixe. 5. Aparelhos auditivos. I. Resende, Leonardo Silva. II. Universidade Federal de Santa Catarina. Programa de Pós-graduação em Engenharia Elétrica. III. Título.

**CANCELAMENTO DE SINAIS INTERFERENTES
USANDO ARRANJO DE MICROFONES: APLICAÇÃO
EM APARELHOS AUDITIVOS**

DIEGO DA SILVA DE MEDEIROS

Esta Dissertação foi julgada adequada para obtenção do Título de Mestre em Engenharia Elétrica, Área de concentração Comunicações e Processamento de Sinais, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.

Leonardo Silva Resende, Dr.
Orientador

Patrick Kuo Peng, Dr.
Coordenador do Programa de Pós -Graduação em Engenharia Elétrica

Banca Examinadora:

Carlos Aurélio Faria da Rocha, Dr.
Presidente

Raimes Moraes, Ph.D.

Walter Pereira Carpes Junior, Dr.

AGRADECIMENTOS

Este trabalho só se tornou possível devido à ajuda de muitas pessoas.

Agradeço em primeiro lugar à minha família – meus pais, João e Marlene, e meu irmão Guilherme (Gui). Certamente, os meus valores existem graças a vocês. Além disso, agradeço pela paciência, pela cobrança saudável, pelo apoio, pela preocupação. Peço desculpas pela minha ausência, agradecendo a compreensão que vocês tiveram.

Minha namorada Larice (Ice) merece posição de destaque na lista de agradecimentos. Eu te agradeço pelo apoio quase infinito, pela compreensão nesses anos difíceis. Agradeço por toda a ajuda fornecida, seja na escrita do texto, seja no desenvolvimento do trabalho – “... será que o erro não tá na matriz?”, ou mesmo no *simples* fato de me acompanhar estudando durante as várias madrugadas.

Agradeço ao meu orientador, Leonardo (Leo), pelas reuniões esclarecedoras e pelas várias contribuições ao trabalho. Mas principalmente, agradeço pela compreensão, ao permitir que a minha dedicação ao mestrado fosse compartilhada com a docência no Instituto Federal de Santa Catarina (IFSC). Certamente, são raros os professores que tem esta postura, e serei eternamente grato por isso.

Pelas riquíssimas sugestões, agradeço aos membros da banca de defesa, Professores Aurélio, Raimes e Walter. Em especial, agradeço a ajuda do Professor Aurélio nessa fase final do trabalho, tornando possível a realização da defesa mesmo com o afastamento para Pós-Doc do Professor Leonardo.

Este trabalho não seria nada sem sinais de fala adequados à realização dos testes. Por isso, agradeço ao Professor Fábio Violaro da UNICAMP pela disposição em ajudar e agradeço ao Professor Carlos Alberto Ynoguti do INATEL pelas informações e disponibilização da sua base de áudios. Ambos os professores foram contatados por e-mail e se dispuseram prontamente a ajudar.

Agradeço também ao IFSC, por ter aliviado minha carga de trabalho mesmo sem um pedido formal de liberação. Da mesma forma, agradeço ao Programa de Pós-Graduação em Engenharia Elétrica da UFSC pela compreensão. Impossível não citar o “Wilson e o Marcelo da Pós”.

Também não pode faltar o CNPq pela bolsa concedida.

Por fim, agradeço aos meus colegas do GpqCom e amigos de outros laboratórios pelos ótimos anos de convivência. Em especial, Deise (Baiana), Renata (Baiana), Marcos (Caco, Cacobina, Baiano), Adriano (Baiano), Brena (Baiana), Julian (Rúlian) e Marcos (Chen).

Muito Obrigado a Todos!

Resumo da Dissertação apresentada à UFSC como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica

CANCELAMENTO DE SINAIS INTERFERENTES USANDO ARRANJO DE MICROFONES: APLICAÇÃO EM APARELHOS AUDITIVOS

Diego da Silva de Medeiros

Fevereiro de 2013

Orientador: Leonardo Silva Resende, Dr.

Área de Concentração: Comunicações e Processamento de Sinais.

Palavras-chave: Cancelamento de interferências, Arranjo de microfones, Conformador de feixe, Aparelhos auditivos.

À medida que o Brasil aumenta sua taxa de envelhecimento populacional, mais pessoas adquirem alguma forma de deficiência auditiva, dificultando o convívio social e reduzindo a qualidade de vida. Embora cada deficiência tenha uma causa e, portanto, uma solução específica, o aumento na relação sinal-ruído e sinal-interferência pode reduzir o seu impacto geral.

Este trabalho tem foco na redução de ruídos e interferências em sinais de áudio utilizando um arranjo de microfones, implementado adaptativamente através de uma estrutura conhecida como conformador de feixe. Através da redundância obtida pelos múltiplos microfones, esta estrutura tem a capacidade de filtrar sinais tanto em frequência quanto em ângulo de incidência, de forma a dar ênfase a sinais julgados de interesse e atenuar sinais de interferência. Por ser adaptativo, o arranjo pode ainda rastrear deslocamentos dos sinais interferentes de forma a efetuar a atenuação mesmo em ambientes ou sinais não estacionários.

Duas estruturas, uma clássica e outra com técnicas mais recentes de processamento, serão comparadas utilizando simulações. Estas

estruturas foram construídas através da variação de parâmetros que vão desde a construção do arranjo ao algoritmo de adaptação utilizado no problema de otimização. Neste ponto, serão confrontados o clássico algoritmo CLMS (*Constrained Least Mean Squares*) e o mais recente Multi-Split LMS (MS-LMS).

Como conclusão, a estrutura mais avançada obteve um desempenho substancialmente melhor que o da estrutura clássica, confirmando as expectativas.

Abstract of Dissertation presented to UFSC as a partial fulfillment of the requirements for the degree of Master in Electrical Engineering

CANCELLATION OF INTERFERING SIGNALS USING MICROPHONE ARRAY: APPLICATION IN HEARING AIDS

Diego da Silva de Medeiros

February 2013

Advisor: Leonardo Silva Resende, Dr.

Area of Concentration: Communications and Signal Processing.

Keyword: Interference cancellation, Microfone Array, Beamformer, Hearing aids.

As the Brazil increases its population aging rate, more people gain some form of hearing impairment, which complicates social interaction and reduces the quality of life. Although each deficiency has a cause and, therefore, a specific solution, the increasing of signal-to-noise and signal-to-interference ratio can reduce their general impact.

This work focuses on noise and interference reduction in audio signals using a microphone array, which is adaptively implemented via a structure known as beamformer. Through the obtained redundancy by the multiple microphones, this structure has the capacity to filter signals both in frequency and in incidence angle, in order to give emphasis to signals judged as of interest and to reduce interference signals. For being adaptive, the array can even track shifts of the interfering signals in order to perform the cancellation even in nonstationaries environments or signals.

Two structures, one classic and another with more recent processing techniques, will be compared using simulations. These structures were built through variation of parameters that ranging from

the array's construction to the adaptive algorithm used in the optimization problem. In this point, it will be confronted the classic CLMS (Constrained Least Mean Squares) algorithm and the more recent Multi-Split LMS (MS-LMS).

As conclusion, the more advanced structure has obtained a substantially better performance than the pioneer structure, confirming the expectations.

LISTA DE FIGURAS

Figura 1.1: Mascaramento em deficientes auditivos.....	25
Figura 1.2: Conformador de feixe.....	26
Figura 2.1: Diagrama de blocos do modelo anecoico do canal acústico.....	30
Figura 2.2: Diagrama de blocos do modelo reverberante do canal acústico.....	31
Figura 2.3: Exemplo de curva de energia de uma resposta ao impulso.....	33
Figura 2.4: Representação de um percurso de propagação de onda reflexivo por um percurso direto a partir de uma fonte sonora imagem.....	35
Figura 2.5: Distribuição das fontes sonoras no ambiente.....	36
Figura 2.6: Disposição dos coeficientes de reflexão em uma sala retangular.....	39
Figura 2.7: Resposta ao impulso acústica obtida através do método das imagens de Allen e Berkley.....	40
Figura 2.8: Resposta ao impulso acústica obtida através da expansão proposta por Peterson do método das imagens de Allen e Berkley.....	41
Figura 2.9: Comparação dos métodos de Allen e Berkley com o método de Peterson.	42
Figura 2.10: Respostas dos microfones de primeira ordem.....	44
Figura 2.11: Comparação do método de Peterson com o método proposto.....	46
Figura 2.12: Resposta ao impulso obtida com um microfone do tipo cardioide focalizado em diferentes regiões.....	48
Figura 3.1: Arranjo de sensores e ondas em far field.....	50
Figura 3.2: Distribuição dos microfones nas configurações broadside e endfire.....	51
Figura 3.3: Forma direta do beamformer.....	52
Figura 3.4: Recepção da onda pelo arranjo de sensores.....	54
Figura 3.5: Resposta de um conformador de feixe banda estreita.....	57
Figura 3.6: O conformador de feixe de banda larga na estrutura direta.....	58
Figura 3.7: Resposta de um conformador de feixe de banda larga.....	59

Figura 3.8: Estrutura do conformador de feixe utilizado por Frost.....	65
Figura 3.9: Arranjo de microfones rotacionado via adiço de blocos de atraso.....	66
Figura 3.10: Filtro de resposta desejada como a soma dos coeficientes dos filtros associados aos microfones.....	69
Figura 3.11: Diviso em blocos de uma rea espacial de ângulo e frequncia para definiço de restriçes pontuais mltiplas.....	74
Figura 3.12: Resposta do conformador obtida quando poucas restriçes pontuais so utilizadas.....	75
Figura 3.13: O conformador de feixe na forma do GSC.....	79
Figura 3.14: Filtragem adaptativa Split.....	85
Figura 3.15: Implementaço da filtragem Split via GSC.....	87
Figura 3.16: Simplificaço da implementaço da filtragem Split via GSC.....	88
Figura 3.17: A estrutura Multi-Split adaptativa.....	89
Figura 3.18: O Multi-Split na forma compacta, representado pela transformaço de Hadamard.....	91
Figura 3.19: Comparaço do MSE ao longo das iteraçes dos algoritmos RLS, MS- LMS, e LMS.....	95
Figura 3.20: Aplicaço do Multi-Split LMS no conformador de feixes adaptativo na forma do GSC.....	96
Figura 3.21: Conformador MS-GSC com a matriz de transformaço Multi-Split incorporada  matriz de bloqueio.....	97
Figura 4.1: Dimenses da sala e posicionamento do arranjo de microfones e das fontes sonoras.....	103
Figura 4.2: Energia das respostas das salas utilizadas nos testes.....	104
Figura 4.3: Resposta do filtro projetado para as restriçes de Frost.....	106
Figura 4.4: Potncia estimada de um sinal de voz considerando diferentes perodos de estacionariedade.....	110
Figura 4.5: Distribuiço dos Nveis de Fala Ativa da base de udios utilizada.....	113
Figura 4.6: Esquema utilizado para obter os sinais de saida do conformador de forma	

isolada.....	118
Figura 4.7: Resposta em frequência da configuração D do conformador.....	121
Figura 4.8: Resposta do conformador da estrutura “Frost” no cenário 1.....	122
Figura 4.9: Resposta em frequência da configuração C do conformador.....	124
Figura 4.10: Resposta do conformador da estrutura “Frost” no cenário 2.....	125
Figura 4.11: Resposta em frequência da configuração B do conformador.....	128
Figura 4.12: Resposta do conformador da estrutura “Frost” no cenário 3.....	129
Figura 4.13: Resposta em frequência da configuração A do conformador.....	131
Figura 4.14: Resposta do conformador da estrutura “Frost” no cenário 4.....	132
Figura 4.15: CSIs obtidos para a estrutura "Frost".....	133
Figura 4.16: Ganho obtido para a estrutura “Frost”.....	134
Figura 4.17: Resposta do conformador da estrutura “Multi-Split” no cenário 1.....	137
Figura 4.18: Resposta do conformador da estrutura “Multi-Split” no cenário 2.....	139
Figura 4.19: Resposta do conformador da estrutura “Multi-Split” no cenário 3.....	141
Figura 4.20: Resposta do conformador da estrutura “Multi-Split” no cenário 4.....	143
Figura 4.21: CSIs obtidos para a estrutura "Multi-Split".....	144
Figura 4.22: Ganho obtido para a estrutura “Multi-Split”.....	145
Figura 4.23: Comparação dos ganhos obtidos pelas duas estruturas.....	146
Figura 4.24: Cenário para o teste com três fontes sonoras.....	150

LISTA DE TABELAS

Tabela 1: Microfones de primeira ordem e componentes de pressão e gradiente.....	43
Tabela 2: Estruturas utilizadas na comparação de desempenho.....	99
Tabela 3: Parâmetros relacionados ao cenário dos testes realizados.....	101
Tabela 4: Parâmetros das duas salas utilizadas nos testes.....	102
Tabela 5: Exemplos de cenários obtidos na escolha do número de valores singulares mais significativos.....	108
Tabela 6: Número de microfones e coeficientes utilizados nos testes.....	109
Tabela 7: Frequências e pesos de inteligibilidade.....	116
Tabela 8: Níveis médios (dB) dos sinais utilizados com a estrutura "Frost" no cenário 1.....	120
Tabela 9: Resultados dos testes da estrutura "Frost" no cenário 1.....	120
Tabela 10: Níveis médios (dB) dos sinais utilizados com a estrutura "Frost" no cenário 2.....	123
Tabela 11: Resultados dos testes da estrutura "Frost" no cenário 2.....	123
Tabela 12: Níveis médios (dB) dos sinais utilizados com a estrutura "Frost" no cenário 3.....	126
Tabela 13: Resultados dos testes da estrutura "Frost" no cenário 3.....	127
Tabela 14: Níveis médios (dB) dos sinais utilizados com a estrutura "Frost" no cenário 4.....	130
Tabela 15: Resultados dos testes da estrutura "Frost" no cenário 4.....	130
Tabela 16: Níveis médios (dB) dos sinais utilizados com a estrutura "Multi-Split" no cenário 1.....	136
Tabela 17: Resultados dos testes da estrutura "Multi-Split" no cenário 1.....	136
Tabela 18: Níveis médios (dB) dos sinais utilizados com a estrutura "Multi-Split" no cenário 2.....	138
Tabela 19: Resultados dos testes da estrutura "Multi-Split" no cenário 2.....	138

XVIII

Tabela 20: Níveis médios (dB) dos sinais utilizados com a estrutura "Multi-Split" no cenário 3.....	140
Tabela 21: Resultados dos testes da estrutura “Multi-Split” no cenário 3.....	140
Tabela 22: Níveis médios (dB) dos sinais utilizados com a estrutura "Multi-Split" no cenário 4.....	142
Tabela 23: Resultados dos testes da estrutura “Multi-Split” no cenário 4.....	142
Tabela 24: Resultados dos testes finais com as duas estruturas.....	148
Tabela 25: Resultados dos testes finais para três fontes sonoras no ambiente.....	151

SUMÁRIO

1 INTRODUÇÃO.....	23
1.1 DEFICIÊNCIAS AUDITIVAS.....	23
1.2 OBJETIVOS DO TRABALHO.....	27
1.3 ORGANIZAÇÃO DO TRABALHO.....	28
2 O CANAL ACÚSTICO.....	29
2.1 MODELOS DE CANAL PARA USO EM ARRANJOS DE MICROFONE.....	29
2.1.1 O modelo anecoico.....	29
2.1.2 O modelo reverberante.....	30
2.2 A RESPOSTA AO IMPULSO ACÚSTICA.....	32
2.2.1 O tempo de reverberação.....	32
2.2.2 Resposta ao impulso acústica via medição.....	33
2.2.3 Resposta ao impulso acústica via simulação.....	34
2.2.4 Extensão do método das imagens para arranjos de microfones.....	40
2.2.5 Extensão do método das imagens para microfones direcionais.....	42
3 O CONFORMADOR DE FEIXES.....	49
3.1 DEFINIÇÃO DO CONFORMADOR DE FEIXES.....	50
3.1.1 Forma direta.....	51
3.1.2 O conformador de feixes banda larga.....	56
3.2 O PROBLEMA DE OTIMIZAÇÃO.....	62
3.2.1 O conformador LCMV adaptativo.....	63
3.3 RESTRIÇÕES LINEARES.....	64
3.3.1 As restrições de Frost.....	64
3.3.2 Restrições pontuais.....	70
3.3.3 Restrições autovetoriais.....	75
3.4 O CANCELADOR DE LÓBULO LATERAL GENERALIZADO.....	78

3.5 A FILTRAGEM MULTI-SPLIT ADAPTATIVA.....	83
3.5.1 A filtragem Split.....	84
3.5.2 A filtragem Multi-Split.....	88
3.5.3 O Multi-Split LMS.....	91
3.5.4 O conformador de feixe Multi-Split GSC.....	96
4 SIMULAÇÕES E RESULTADOS.....	99
4.1 PARÂMETROS DOS TESTES.....	99
4.1.1 Parâmetros relacionados ao cenário dos testes.....	100
4.1.1.1 <i>Número de fontes sonoras</i>	100
4.1.1.2 <i>Resposta do canal</i>	100
4.1.1.3 <i>Tipos de microfones</i>	103
4.1.1.4 <i>Configuração do arranjo de microfones</i>	104
4.1.1.5 <i>Potência do ruído gaussiano aditivo</i>	104
4.1.2 Parâmetros relacionados à filtragem adaptativa.....	105
4.1.2.1 <i>Restrições</i>	105
4.1.2.2 <i>Valores singulares das restrições autovetoriais</i>	106
4.1.2.3 <i>Número de microfones e coeficientes</i>	107
4.1.2.4 <i>Parâmetros dos algoritmos de otimização</i>	109
4.1.2.5 <i>Considerações finais sobre a simulação</i>	110
4.2 ÁUDIOS UTILIZADOS.....	112
4.3 MEDIDAS DE DESEMPENHO.....	113
4.3.1 Medidas objetivas de inteligibilidade.....	114
4.3.2 Medida de desempenho utilizada no trabalho.....	116
4.4 RESULTADOS.....	119
4.4.1 Resultados com a estrutura “Frost”.....	119
4.4.1.1 <i>Cenário 1: Ambiente pouco reverberante e microfones omnidirecionais</i>	119

4.4.1.2 <i>Cenário 2: Ambiente pouco reverberante e microfones cardioides</i>	122
4.4.1.3 <i>Cenário 3: Ambiente muito reverberante e microfones omnidirecionais</i>	125
4.4.1.4 <i>Cenário 4: Ambiente muito reverberante e microfones cardioides</i>	129
4.4.1.5 <i>Comparação dos resultados</i>	132
4.4.2 Resultados com a estrutura “Multi-Split”.....	135
4.4.2.1 <i>Cenário 1: Ambiente pouco reverberante e microfones omnidirecionais</i>	135
4.4.2.2 <i>Cenário 2: Ambiente pouco reverberante e microfones cardioides</i>	137
4.4.2.3 <i>Cenário 3: Ambiente muito reverberante e microfones omnidirecionais</i>	139
4.4.2.4 <i>Cenário 4: Ambiente muito reverberante e microfones cardioides</i>	141
4.4.2.5 <i>Comparação dos resultados</i>	143
4.4.3 Comparação das estruturas.....	145
4.4.3.1 <i>Teste adicional com os mesmos arquivos de áudio</i>	147
4.4.3.2 <i>Teste adicional com três fontes sonoras</i>	149
4.5 CONCLUSÕES DO CAPÍTULO.....	151
5 CONSIDERAÇÕES FINAIS.....	153
5.1 CONCLUSÕES.....	153
5.2 TRABALHOS FUTUROS.....	154

1 INTRODUÇÃO

1.1 DEFICIÊNCIAS AUDITIVAS

No Brasil, mais de 9,8 milhões de pessoas possuem alguma forma de deficiência auditiva, número que passa dos 5% do total da população [1]. Este número ainda tende a crescer, à medida que o Brasil aumenta a sua taxa de envelhecimento populacional.

São vários os fatores que podem causar deficiências auditivas. Dentre eles têm-se os pré-natais, os com origem no parto e os relacionados a exposição a sons muito altos. No caso dos pré-natais eles podem ser decorrentes da hereditariedade, doenças infectocontagiosas, drogas ou alcoolismo materno, carências alimentares e exposição à radiação. Os casos oriundos do parto normalmente se relacionam a infecções hospitalares. E, por fim, as deficiências auditivas causadas pela exposição contínua a sons altos são mais comum em jovens.

Das perdas auditivas, a forma mais comum, com mais de 90% das ocorrências, é a neurossensorial [2]. Este tipo de deficiência ocorre quando há uma impossibilidade de recepção do som, resultante de danos nas células sensoriais auditivas ou nos nervos auditivos. É causada tanto por fatores hereditários quanto por problemas no pré-natal, como contração de rubéola, toxoplasmose, herpes, alcoolismo e diabetes, além de traumas físicos, nascimento prematuro e trauma de parto.

Dentre os problemas enfrentados por pessoas com perdas neurossensoriais, os mais comuns são [3]:

Redução da faixa dinâmica: a intensidade entre o som mais fraco perceptível e o som mais forte suportado é reduzida. Com isso, há um grande desconforto na audição. Aparelhos auditivos amplificadores precisam ser projetados para amplificar de forma diferente sinais fracos e fortes, transformando a grande faixa dinâmica do ambiente numa faixa dinâmica menor.

Redução da seletividade espectral: uma das características psicoacústicas presente nos seres humanos é o mascaramento auditivo, que limita a percepção de um sinal sonoro devido à presença de outro sinal em frequência distinta. Deficientes auditivos possuem um maior mascaramento de frequências, dificultando a distinção de sons pelo ouvido e a posterior separação dos sons pelo cérebro.

O entendimento da fala pode ser ainda mais prejudicado quando uma componente espectral possui alta intensidade, já que a seletividade em frequência é ainda menor nestes casos. Como pessoas com perdas auditivas precisam amplificar os sons mais fracos devido à redução da sua faixa dinâmica, o mascaramento auditivo é conseqüentemente um fator significativo na perda da inteligibilidade [4].

A Figura 1.1 demonstra o mascaramento auditivo utilizando sinais acústicos com frequências de 500 Hz e 1000 Hz. A linha tracejada em azul demarca a banda espectral mascarada pelo sinal e a intensidade mínima que sinais dentro desta faixa de frequência precisam ter para serem percebidos. A figura superior, representa um indivíduo com audição normal, sendo que para a situação ambos os sinais serão percebidos. A figura inferior apresenta um maior mascaramento em virtude de deficiências auditivas. Neste caso, o sinal de 1000 Hz não será percebido pelo ouvinte, já que ele se encontra dentro da zona de mascaramento do sinal de 500 Hz.

Redução da seletividade temporal: pessoas sem deficiências auditivas conseguem extrair partes da informação de interesse durante momentos em que ruídos de fundo são interrompidos. Sendo rápida a flutuação dos ruídos ao longo do tempo, é preciso que a audição tenha resolução temporal suficiente para que a extração de informação seja bem sucedida. Pessoas com perdas auditivas possuem a seletividade temporal reduzida, perdendo a habilidade de escutar em intervalos de silêncio do ruído de fundo.

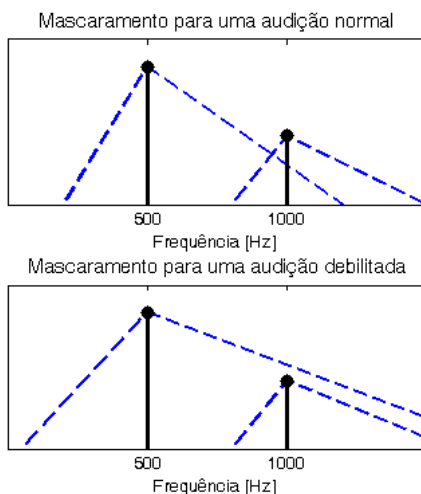


Figura 1.1: Mascaramento em deficientes auditivos. As figuras apresentam dois sinais em frequências de 500 Hz e 1000 Hz, assim como o mascaramento de sinal é representado pela linha tracejada. Na figura inferior, o sinal de 1000 Hz não será percebido pelo ouvinte, pois está mascarado pela componente de 500 Hz.

Adaptado de [75]

Neste cenário, este trabalho tem foco na redução de ruídos e interferências em aparelhos auditivos. Com a redução de interferências, uma minimização dos efeitos da baixa seletividade espectral e temporal é esperada, tendo em vista que com melhores relações sinal-ruído e sinal-interferência, tanto a probabilidade de mascaramento em frequência quanto a necessidade de uma alta resolução temporal são reduzidas. De fato, deficientes auditivos precisam de relações sinal-ruído e sinal interferência melhores para manter o mesmo nível de conversação que pessoas sem deficiências [3]. Estudos revelam que, em média, a diferença de relação sinal-ruído e sinal-interferência necessária para um reconhecimento de 50% das palavras entre pessoas sem deficiências e deficientes vai de 4 dB, para uma deficiência auditiva leve, a 18 dB para uma deficiência auditiva profunda [5] [6].

Para o processamento, será empregada uma estrutura conhecida como *conformador de feixe*, ou *beamformer*, que consiste em múltiplos microfones associados a filtros de resposta ao impulso finita (FIR - *Finite Impulse Response*) com coeficientes adaptáveis (Figura 1.2). Com a redundância temporal e espacial disponibilizada, é possível controlar a resposta do sistema tanto em nível de frequência quanto de ângulo de incidência. Desta forma, o sinal de interesse será mantido enquanto as interferências presentes no meio serão atenuadas, sendo os parâmetros para a distinção entre sinal de interesse e interferência, a faixa de frequência e a localização espacial dos sinais.

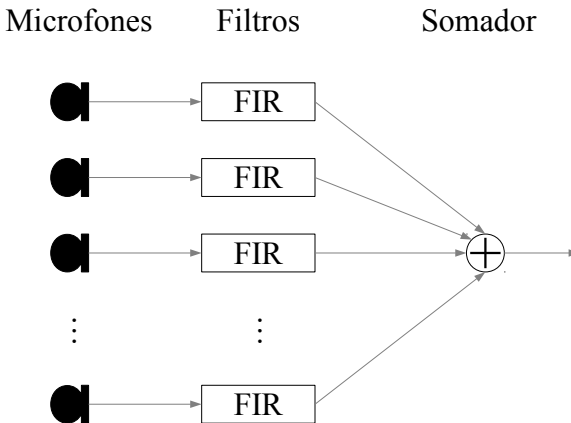


Figura 1.2: Conformador de feixe

Visto de uma outra forma, o conformador de feixe mantém uma resposta de magnitude unitária para uma determinada região de interesse, enquanto minimiza as regiões em que os sinais de interferência estão localizados. Neste caso, entende-se por região o par ângulo de incidência e frequência. Devido à sua natureza adaptativa, o sistema consegue acompanhar alterações das características dos sinais ao longo do tempo.

1.2 OBJETIVOS DO TRABALHO

A aplicação de conformador de feixe em aparelhos auditivos é bastante explorada na área acadêmica. Cientistas de todo o mundo aplicam as mais diversas técnicas de processamento de sinais em arranjo de microfones na redução de ruídos e interferências em aparelhos auditivos.

Um dos primeiros trabalhos nesta área, de Peterson [7], utilizou um conformador de feixe com restrições de Frost, que garante que sinais incidentes a zero graus são mantidos enquanto outros são minimizados pelo algoritmo adaptativo.

Desta forma, tendo como base o trabalho de Peterson, o objetivo deste trabalho é a investigação da aplicação de um esquema alternativo de implementação do conformador de feixes em aparelhos auditivos, buscando a melhora do desempenho e a redução da complexidade computacional.

Os objetivos específicos deste trabalho são:

- Estender os modelos de arranjo de microfones para a utilização de microfones direcionais;
- Propor uma figura de mérito para a comparação do desempenho dos conformadores de feixe investigados;
- Analisar o impacto da utilização de microfones direcionais no desempenho dos sistemas;
- Analisar o impacto da variação do número de microfones;
- Analisar o impacto do aumento da reverberação do ambiente;

1.3 ORGANIZAÇÃO DO TRABALHO

Com esta introdução, o presente trabalho possui cinco capítulos.

O capítulo dois aborda como o ambiente interfere em uma comunicação de fala. Duas modelagens do canal acústico amplamente utilizadas são apresentadas.

O capítulo três traz os fundamentos teóricos referentes às técnicas de filtragem empregadas no trabalho. O conformador de feixes é apresentado, juntamente com as ferramentas matemáticas utilizadas na sua operação. Os dois algoritmos adaptativos comparados neste trabalho também são apresentados nesse capítulo.

O capítulo quatro apresenta os tópicos relacionados aos testes realizados, assim como traz os resultados obtidos. Também é indicada a medida de desempenho utilizada como base na comparação dos algoritmos adaptativos.

Por fim, as considerações finais do trabalho e as sugestões para continuidade são apresentadas no capítulo cinco.

2 O CANAL ACÚSTICO

2.1 MODELOS DE CANAL PARA USO EM ARRANJOS DE MICROFONE

Um sinal $s(t)$ emitido por uma fonte sonora é afetado pelo meio antes de ser captado pelos microfones do arranjo. Dependendo do nível de realismo desejado, alguns fatores podem ser levados em consideração ou não na modelagem do canal acústico.

Serão apresentados, a seguir, dois modelos matemáticos utilizados na descrição do sinal recebido pelos microfones do arranjo: o modelo anecoico e o modelo reverberante.

2.1.1 O modelo anecoico

Um sinal captado pelos microfones do arranjo, utilizando um modelo anecoico, é uma versão atenuada e atrasada do sinal original. Em outras palavras, sendo um arranjo linear com K microfones, o sinal recebido pelo k -ésimo microfone em um determinado instante de tempo t é [12]:

$$y_k(t) = \alpha_k s[t - T_0 - F_k(\tau)] + v_k(t) \quad (2.1)$$

onde α_k com $k=0,1,\dots,K-1$ são os fatores de atenuação devido à propagação do sinal; $s(t)$ é o sinal emitido pela fonte sonora; T_0 é o tempo levado pelo sinal para se propagar da fonte sonora ao primeiro microfone do arranjo (microfone de referência); τ é o atraso sofrido pela onda para trafegar do microfone 0 ao 1; $F_k(\tau)$ é o atraso relativo entre os microfones 0 e k , de forma que $F_0(\tau)=0$ e $F_1(\tau)=\tau$; e por último, $v_k(t)$ é o ruído gaussiano de média zero presente no microfone k . O diagrama de blocos do modelo pode ser

visto na Figura 2.1, sendo que os blocos em cinza representam atrasos de sinal.

Por ser um modelo simplificado, o modelo anecoico não leva em consideração a reverberação inerente ao ambiente.

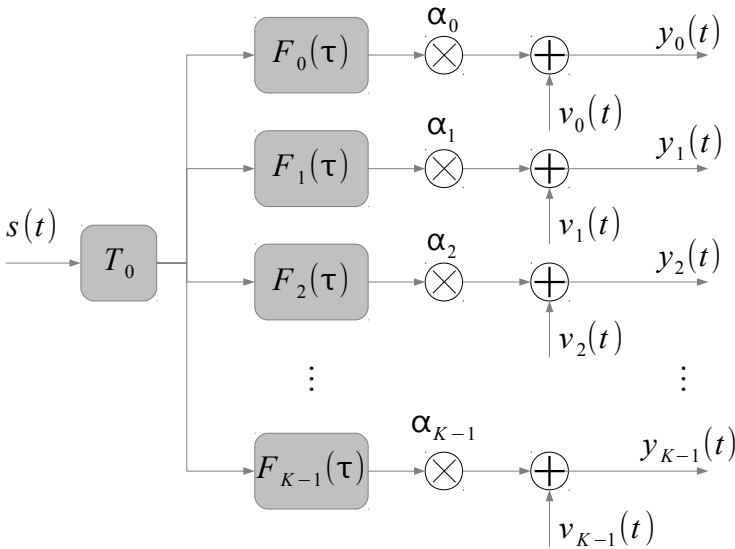


Figura 2.1: Diagrama de blocos do modelo anecoico do canal acústico

2.1.2 O modelo reverberante

Como o próprio nome já diz, o modelo reverberante inclui na descrição do sinal a reverberação presente no ambiente. Para um arranjo com K microfones, o sinal recebido pelo microfone k em um determinado tempo t é [12]:

$$y_k(t) = g_k * s(t) + v_k(t) \quad (2.2)$$

onde g_k é a resposta ao impulso acústica do ambiente, da fonte sonora ao microfone k , e o $*$ é o operador de convolução. O diagrama de blocos pode ser visto na Figura 2.2.

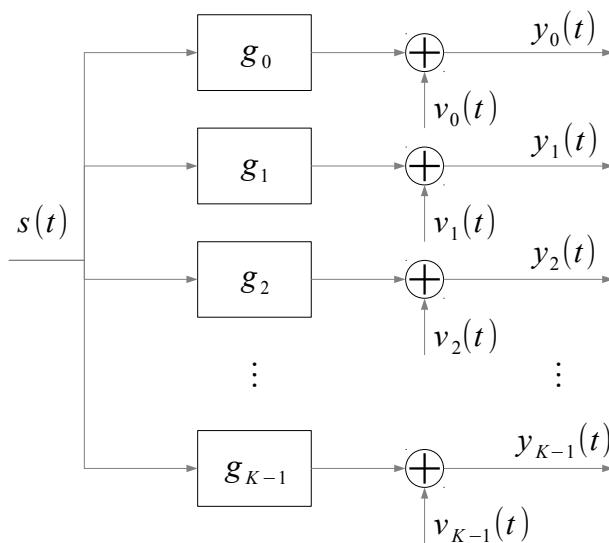


Figura 2.2: Diagrama de blocos do modelo reverberante do canal acústico

Através da resposta ao impulso acústica, toda a interação do sinal com o ambiente pode ser descrita. Além da reverberação, a resposta ao impulso acústica implementa o atraso de propagação da onda da fonte sonora até o primeiro microfone, o atraso de propagação da onda sonora entre os microfones e a atenuação devido à propagação. Ou seja, o modelo reverberante inclui, além da reverberação, todos os fatores implementados pelo modelo anecoico.

2.2 A RESPOSTA AO IMPULSO ACÚSTICA

Para utilizar o modelo reverberante, é necessário o conhecimento da resposta ao impulso do ambiente, com o “caminho” percorrido por uma onda acústica da fonte até cada um dos microfones. As amostras da resposta ao impulso podem ser obtidas através de medições no canal ou por meio de simulações, sendo a segunda forma a utilizada neste trabalho.

As seções a seguir apresentam mais detalhes sobre a resposta ao impulso de ambientes acústicos. O tempo de reverberação, uma importante característica no estudo de ambientes acústicos, será inicialmente definido. Posteriormente, os métodos de estimação da resposta ao impulso serão apresentados.

2.2.1 O tempo de reverberação

O tempo de reverberação, medida do quão reverberante é um ambiente acústico, é definido como o tempo necessário para a energia do som decair 60 dB. É comumente conhecido pela sigla T_{60} .

Por medir a reverberação do ambiente, o T_{60} é função dos coeficientes de reflexão e da área das paredes da sala, assim como também é função do volume total da sala. Outros fatores como temperatura do ambiente e a própria absorção do ar contribuem para a medida, embora sejam menos significativos.

Há diversas formas de estimar o tempo de reverberação, sendo as mais conhecidas a equação de Wallace Clement Sabine, proposta em seu estudo iniciado em 1900 [13], e a equação de Carl Ferdinand Eyring, de 1930 [14]. Embora antigas, ambas as equações são ainda utilizadas na estimação do tempo de reverberação em trabalhos atuais, como em [15] e [16].

Outra forma de estimar o tempo de reverberação é observar o decaimento da energia da resposta ao impulso do ambiente, normalizada

pelo seu valor máximo. O método de Schroeder [17] pode ser utilizado para obter a energia, através da seguinte equação:

$$E(t) = N \int_t^{\infty} h^2(l) dl \quad (2.3)$$

onde N é uma constante relacionada à densidade espectral de energia do sinal utilizado e $h(l)$ é a resposta ao impulso do ambiente. Um exemplo de curva de decaimento de energia de um ambiente pode ser visto na Figura 2.3, onde o tempo de reverberação é de, aproximadamente, 600 ms.

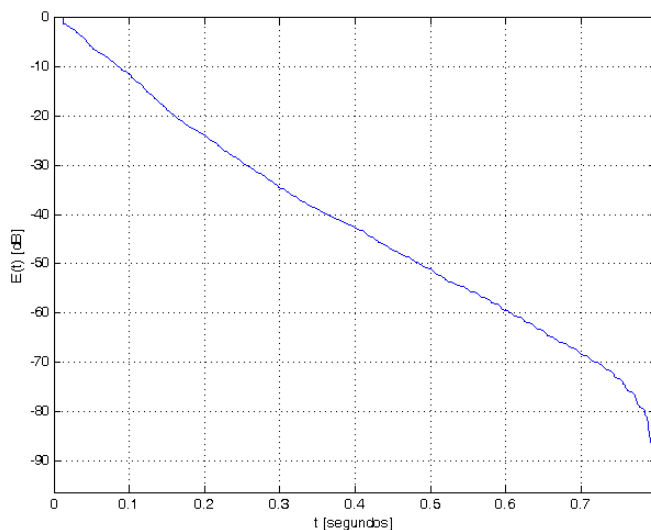


Figura 2.3: Exemplo de curva de energia de uma resposta ao impulso. É possível visualizar o tempo de reverberação de 600 ms

2.2.2 Resposta ao impulso acústica via medição

Uma das formas de se obter a resposta ao impulso do ambiente é através da medição [18]. Neste caso, um sinal é introduzido no canal e

captado pelos microfones. Sabendo que o sinal captado é a convolução do sinal transmitido pela resposta ao impulso do canal, é possível obter a resposta ao impulso através de um processo de deconvolução [19].

Os métodos de medição da resposta ao impulso acústica do canal variam em relação ao sinal introduzido no canal e no processo de deconvolução. Alguns dos métodos são encontrados em [20], [21], [22] e [23], e uma comparação deles pode ser encontrada em [19]. Na *internet* há alguns sites que disponibilizam respostas ao impulso medidas em uma variedade de ambientes, de pequenas salas a catedrais. Alguns exemplos são [24] e [25].

2.2.3 Resposta ao impulso acústica via simulação

Obter a resposta ao impulso do canal utilizando medições pode ser uma tarefa cara e demorada, tendo em vista que os equipamentos utilizados devem ser de qualidade e cada alteração no cenário de testes implica numa remontagem da estrutura.

Uma forma de se obter mais liberdade na utilização da resposta ao impulso do ambiente é através de simulação. O método das imagens proposto por Allen e Berkley [26] é o método de simulação mais utilizado quando o ambiente em questão possui geometria retangular [18].

O método das imagens baseia-se no fato de que um sinal refletido numa parede rígida pode ser visto como um sinal que “atravessa” a parede, transmitido por uma duplicata da fonte sonora localizada fora dos limites da sala (Figura 2.4) [18]. Este conceito é também estendido às duplicatas das duplicatas e assim por diante, incorporando os vários caminhos de propagação da onda sonora na sala. No caso de uma parede não rígida, a onda sonora refletida é ponderada pelo coeficiente de reflexão da parede. Em [26], os coeficientes de reflexão são assumidos independentes do ângulo de incidência da onda e da frequência do sinal.

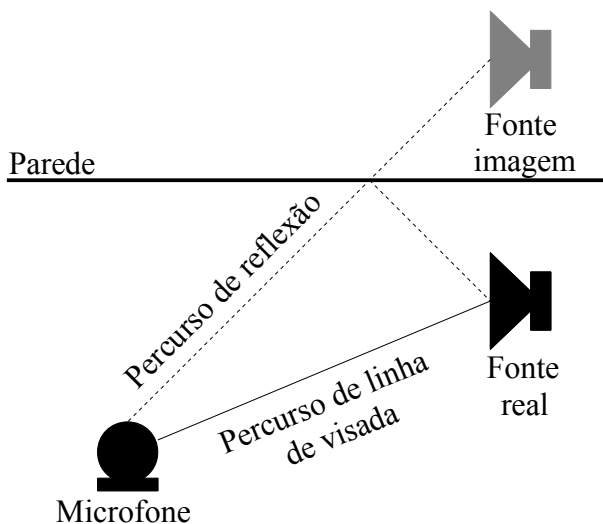


Figura 2.4: Representação de um percurso de propagação de onda reflexivo por um percurso direto a partir de uma fonte sonora imagem.

Para entender o método das imagens, é preciso ter em mente que cada fonte imagem levada em consideração irá contribuir com um impulso na resposta ao impulso final. Para que a simulação seja satisfatória, imagens até cT_{60} metros de distância do microfone devem ser incluídas, sendo c a velocidade de propagação da onda, aproximadamente 348 m/s. Estas imagens são utilizadas devido ao fato de elas, desconsiderando as perdas de reflexão, contribuir para a faixa audível da resposta ao impulso (0 dB a 60 dB).

As fontes sonoras imagem são distribuídas no modelo tendo como base a localização da fonte sonora real, a partir de sucessivos espelhamentos nas três dimensões da sala. Desta forma, estando a fonte real próxima a uma determinada parede, a primeira fonte imagem considerada nesta direção estará à mesma distância da parede, mas do lado oposto dela.

Para facilitar o entendimento do modelo, este será inicialmente observado em apenas uma dimensão. Sendo L_x o tamanho da sala e x_s a localização da fonte sonora real, a localização da primeira fonte imagem considerada será $2L_x - x_s$, já que esta sala imagem é um “espelho” da sala real. A próxima sala imagem considerada, um “espelho do espelho”, será idêntica a sala real, estando portanto a fonte imagem em $2L_x + x_s$. A próxima sala imagem, novamente espelhada em relação à sala real, terá a fonte sonora em $4L_x - x_s$. Este exemplo pode ser visto na Figura 2.5.

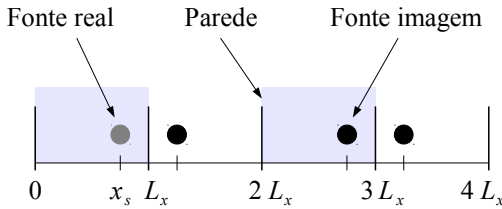


Figura 2.5: Distribuição das fontes sonoras no ambiente

Voltando para as três dimensões, a localização (x, y, z) de uma determinada fonte sonora, real ou imagem, dentro do modelo é:

$$\left(\pm x_s + 2q_x L_x, \quad \pm y_s + 2q_y L_y, \quad \pm z_s + 2q_z L_z \right) \quad (2.4)$$

onde q_x , q_y e q_z são inteiros que indicam o número da sala considerada em cada dimensão. Uma outra forma de localizar a imagem dentro do modelo é:

$$\left((-1)^{p_x} x_s + 2q_x L_x, \quad (-1)^{p_y} y_s + 2q_y L_y, \quad (-1)^{p_z} z_s + 2q_z L_z \right) \quad (2.5)$$

onde os inteiros p_x , p_y e p_z assumem o valor 1 quando a sala imagem é uma versão espelhada da sala real e 0 no outro caso. Para o restante desta apresentação, os inteiros p_i e q_i , $i=x, y, z$ utilizados serão agrupados nos vetores \mathbf{P} e \mathbf{Q} :

$$\mathbf{P} \stackrel{\text{def}}{=} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \quad \mathbf{Q} \stackrel{\text{def}}{=} \begin{bmatrix} q_x \\ q_y \\ q_z \end{bmatrix} \quad (2.6)$$

Para cada imagem incluída na simulação, devem ser estimados o atraso sofrido pela onda na propagação até o microfone, a atenuação devido aos efeitos de propagação, e a atenuação devido às reflexões do sinal, obtida através dos coeficientes de reflexão das paredes cruzadas no trajeto de propagação da onda. Estes três parâmetros são então combinados e estendidos a todas as imagens do modelo para se obter a resposta ao impulso do ambiente [18].

A seguinte equação sintetiza o método, calculando L amostras da resposta ao impulso $h(l)$:

$$h(l) = \sum_{p_x, p_y, p_z=0}^1 \sum_{q_x=-Q_x}^{Q_x} \sum_{q_y=-Q_y}^{Q_y} \sum_{q_z=-Q_z}^{Q_z} \left\{ \frac{\beta(\mathbf{P}, \mathbf{Q}) \times \delta[l - \tau(\mathbf{P}, \mathbf{Q})]}{d(\mathbf{P}, \mathbf{Q})} \right\} \quad (2.7)$$

$l = 1, 2, \dots, L$

onde o somatório em p_x , p_y e p_z considera as salas nas versões originais e espelhadas e o somatório em q_x , q_y e q_z varre todas as imagens até cT_{60} metros de distância do microfone, sendo os limites calculados com:

$$Q_i = \left\lceil \frac{cT_{60}}{2L_i} \right\rceil \quad i = \{x, y, z\} \quad (2.8)$$

onde $\lceil \cdot \rceil$ é o operador de arredondamento para o inteiro mais próximo. O parâmetro $\beta(\cdot)$ é a atenuação devido às reflexões do sinal, calculada por

$$\beta(\mathbf{P}, \mathbf{Q}) = \beta_{x_0}^{|q_x - p_x|} \beta_{x_1}^{|q_x|} \beta_{y_0}^{|q_y - p_y|} \beta_{y_1}^{|q_y|} \beta_{z_0}^{|q_z - p_z|} \beta_{z_1}^{|q_z|} \quad (2.9)$$

onde os termos β_{mn} são os coeficientes de reflexão das seis faces da sala retangular, sendo m o eixo perpendicular à face e n igual a zero para a face perpendicular à origem e um para a oposta (Figura 2.6). O parâmetro $d(\cdot)$ no denominador inclui no modelo a atenuação devido à propagação, sendo calculado por:

$$d(\mathbf{P}, \mathbf{Q}) = \|\mathbf{r} - \mathbf{r}_m\| \quad (2.10)$$

onde $\|\cdot\|$ é o operador de distância euclidiana, \mathbf{r} é o vetor tridimensional com as posições (x, y, z) da fonte sonora imagem:

$$\mathbf{r} = \begin{bmatrix} (-1)^{p_x} x_s + 2 q_x L_x \\ (-1)^{p_y} y_s + 2 q_y L_y \\ (-1)^{p_z} z_s + 2 q_z L_z \end{bmatrix} \quad (2.11)$$

e \mathbf{r}_m é o vetor tridimensional com as posições (x, y, z) do microfone:

$$\mathbf{r}_m = \begin{bmatrix} x_m \\ y_m \\ z_m \end{bmatrix} \quad (2.12)$$

Por fim, $\delta(l)$ é o *Delta de Dirac*, que inclui o atraso de propagação do som através de $\tau(\cdot)$, calculado a partir de:

$$\tau(\mathbf{P}, \mathbf{Q}) = \left\lfloor \frac{d(\mathbf{P}, \mathbf{Q})}{c t_s} \right\rfloor \quad (2.13)$$

onde t_s é o período de amostragem utilizado e o operador de arredondamento é utilizado para evitar uma quantidade fracionária de amostras.

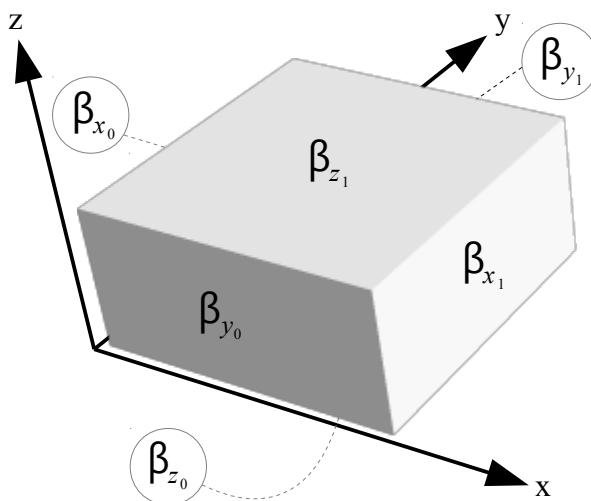


Figura 2.6: Disposição dos coeficientes de reflexão em uma sala retangular

Após a geração da resposta ao impulso, uma filtragem passa altas com uma frequência de corte em 1% da frequência de amostragem deve ser realizada, a fim de retirar a componente contínua obtida do modelo.

Na Figura 2.7 pode ser visto um exemplo de uma resposta ao impulso calculada com o método de Allen e Berkley [26]. Uma sala com dimensões $(3 \times 4 \times 3)$ metros foi utilizada, estando a fonte sonora na posição $(2 \times 2 \times 1,5)$ e o microfone na posição $(2 \times 1 \times 1,5)$. Foi utilizado como coeficiente de reflexão das paredes o valor 0,9, sendo usado para o chão e para o teto o valor de 0,7.

É importante salientar que o método das imagens considera apenas uma sala com quatro paredes, sem levar em conta janelas, portas, móveis, pessoas, e demais divisórias que possam existir no ambiente.

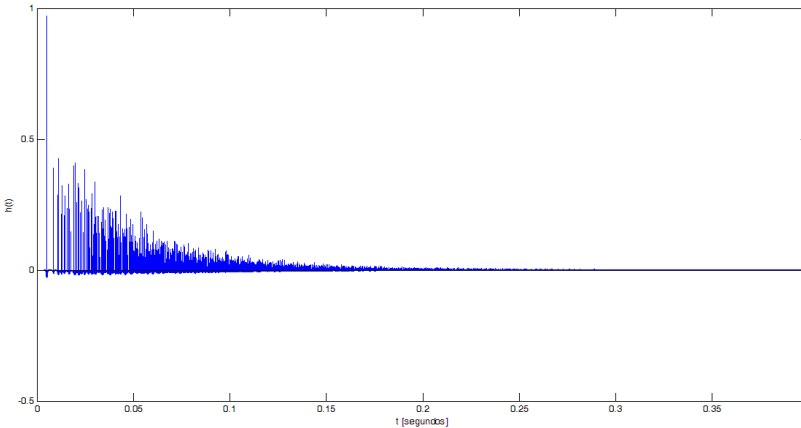


Figura 2.7: Resposta ao impulso acústica obtida através do método das imagens de Allen e Berkley.

2.2.4 Extensão do método das imagens para arranjos de microfones

No método das imagens, cada fonte imagem contribui para a resposta acústica com um impulso. A amostra da resposta ao impulso que vai abrigar a contribuição de uma determinada fonte imagem é obtida com a Equação (2.13). O operador de arredondamento da equação é utilizado para os casos onde a captação do sinal pelo microfone não ocorre em um instante de amostragem exato.

O erro de fase introduzido pelo operador de arredondamento é imperceptível quando a captação é feita por um único microfone, mas introduz erros consideráveis quando a captação é realizada por um arranjo de microfones [27].

Uma solução proposta por Peterson em [27] é que cada imagem contribua com um impulso filtrado em passa baixas, ocorrido no instante de tempo exato calculado com a Equação (2.13) sem o operador de arredondamento. No método proposto, o filtro escolhido é um FIR com janela de *Hanning*, ordem 40 e frequência de corte igual à $0,45 f_s$, sendo f_s a frequência de amostragem. A Figura 2.8 apresenta uma

resposta ao impulso calculada utilizando o método de Peterson, tendo como cenário a mesma configuração utilizada na Figura 2.7.

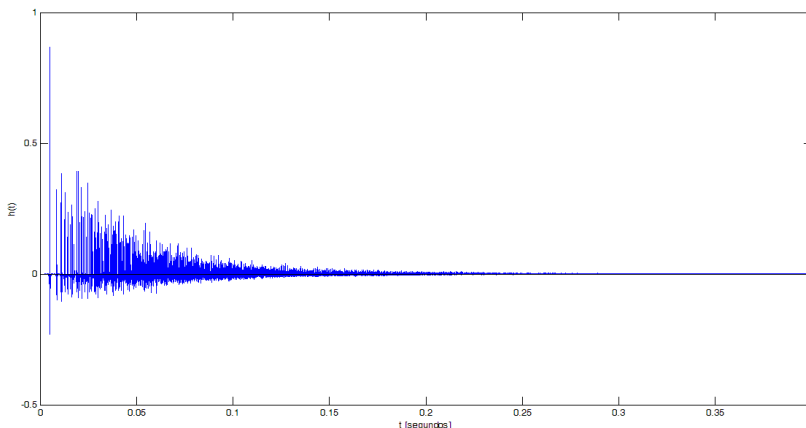


Figura 2.8: Resposta ao impulso acústica obtida através da expansão proposta por Peterson do método das imagens de Allen e Berkley.

A Figura 2.9 faz uma comparação da contribuição de uma única imagem para a resposta ao impulso, utilizando o método de Allen e Berkley e o método de Peterson. É possível notar que o instante exato de incidência da onda sonora no microfone ocorre entre os instantes de amostragem dezenove e vinte. Com o método de Allen e Berkley, o instante de incidência da onda é arredondado para o instante de amostragem mais próximo, no caso, o instante vinte. Com o método de Peterson, uma função *sinc* é usada na representação da onda incidente, preservando a fase da onda com o espalhamento da onda em algumas amostras.

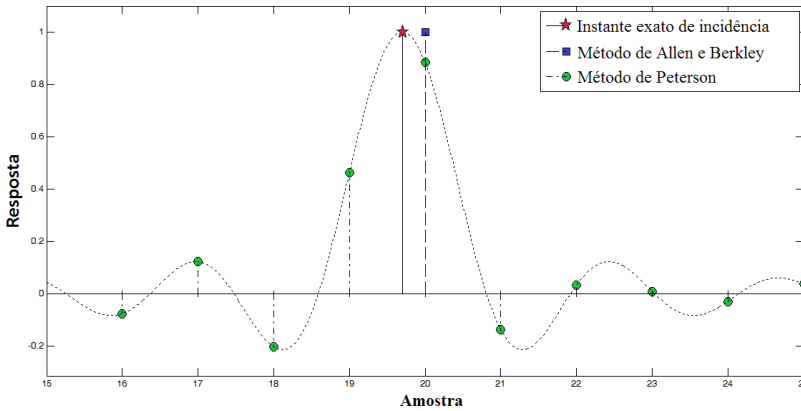


Figura 2.9: Comparação dos métodos de Allen e Berkley com o método de Peterson.

2.2.5 Extensão do método das imagens para microfones direcionais

Tanto o método de Allen e Berkley quanto a expansão proposta por Peterson não consideram na obtenção da resposta ao impulso a resposta angular do microfone, ou visto de uma outra forma, consideram microfones omnidirecionais. Isto limita o número de aplicações que podem utilizar o método como teste inicial. Pensando nesta limitação, uma expansão dos métodos apresentados pode ser feita, incluindo no cálculo da resposta ao impulso a resposta angular do microfone.

A maioria dos microfones direcionais usados atualmente são membros da família dos microfones de primeira ordem, que têm uma resposta polar dependente de um termo em cosseno com primeira potência [28]. A equação a seguir apresenta a resposta de um microfone de primeira ordem:

$$\mu(\theta, \phi) = |A + B \cos(\theta - \theta_f) \sin(\phi - \phi_f)| \quad (2.14)$$

onde θ e ϕ são as variáveis independentes da equação, sendo respectivamente o *azimute* e a *inclinação* (ou *elevação*) do sistema de coordenadas esférico, θ_f e ϕ_f são as coordenadas de foco do

microfone, e A e B são respectivamente as componentes de pressão e gradiente da resposta do microfone, sendo que $A+B=1$. A Tabela 1 apresenta os parâmetros A e B para os tipos mais comuns de microfones de primeira ordem. A Figura 2.10 apresenta a resposta dos tipos de microfones utilizados.

Tabela 1: Microfones de primeira ordem e componentes de pressão e gradiente

Tipo	A	B
Omnidirecional	1,00	0,00
Subcardioide	0,70	0,30
Cardioide	0,50	0,50
Supercardioide	0,37	0,63
Hipercardioide	0,25	0,75
Figure-8	0,00	1,00

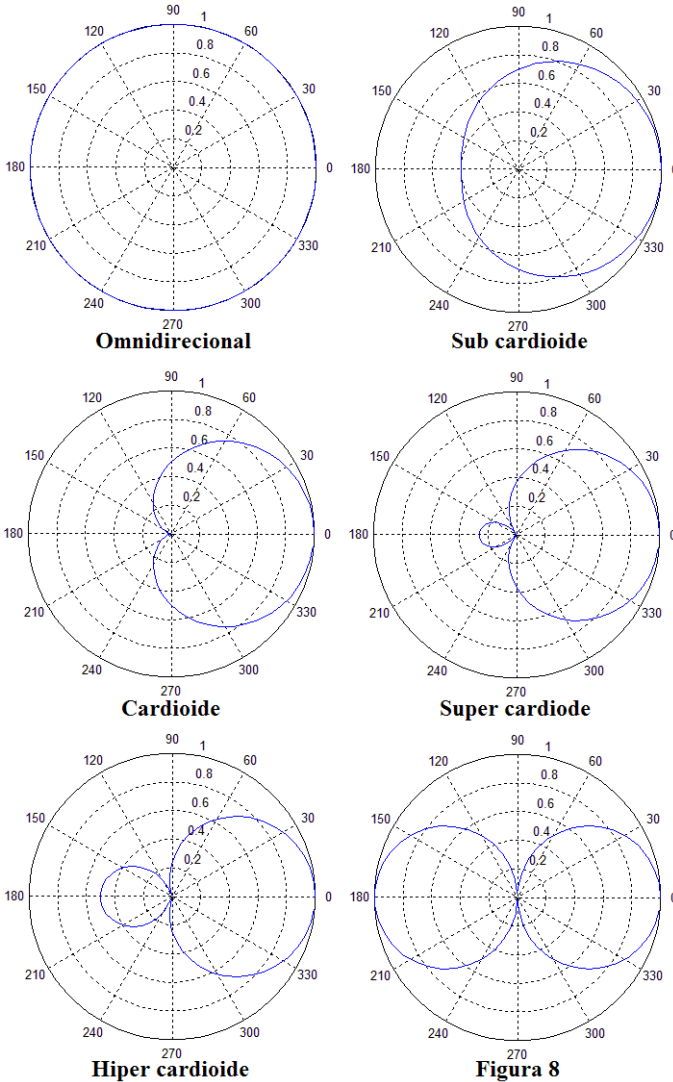


Figura 2.10: Respostas dos microfones de primeira ordem

Para considerar a resposta do microfone, cada fonte sonora (imagem e real) deve ser ponderada pela resposta do microfone no

ângulo de incidência da onda. Assim, é incluído um termo multiplicativo na Equação (2.7):

$$h(l) = \sum_{p_x, p_y, p_z=0}^1 \sum_{q_x=-Q_x}^{Q_x} \sum_{q_y=-Q_y}^{Q_y} \sum_{q_z=-Q_z}^{Q_z} \left\{ \frac{\beta(\mathbf{P}, \mathbf{Q}) \times \delta[l - \tau(\mathbf{P}, \mathbf{Q})]}{d(\mathbf{P}, \mathbf{Q})} \times \mu(\mathbf{P}, \mathbf{Q}) \right\} \quad (2.15)$$

$l=1, 2, \dots, L$

onde $\mu(\cdot)$ é a resposta do microfone para o ângulo de incidência (θ, ϕ) correspondente da onda, sendo este calculado a partir da localização do microfone na sala e dos vetores \mathbf{P} e \mathbf{Q} com:

$$\theta = \text{atan2}(r_2 - y_m, r_1 - x_m) \quad (2.16)$$

$$\phi = \arccos\left(\frac{r_3 - z_m}{d(\mathbf{P}, \mathbf{Q})}\right) \quad (2.17)$$

sendo r_1 , r_2 e r_3 elementos do vetor \mathbf{r} da Equação (2.11) e $\text{atan2}(y, x)$ a variação com dois argumentos da função $\arctan(y/x)$, definida como:

$$\text{atan2}(y, x) = \begin{cases} \arctan\left(\frac{y}{x}\right) & x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & y \geq 0, x < 0 \\ \arctan\left(\frac{y}{x}\right) - \pi & y < 0, x < 0 \\ +\frac{\pi}{2} & y > 0, x = 0 \\ -\frac{\pi}{2} & y < 0, x = 0 \\ \text{indefinida} & y = 0, x = 0 \end{cases} \quad (2.18)$$

A Figura 2.11 compara a resposta ao impulso obtida pelo método de Peterson (omnidirecional) com a resposta obtida levando a resposta do microfone em consideração, com um microfone do tipo cardioide

focalizado diretamente na fonte sonora. É possível notar que a componente de linha de visada, denotada com um “A”, não se altera, enquanto as demais componentes são atenuadas na versão com a resposta do microfone. Isto acontece porque a resposta do microfone para a componente de linha de visada é 1, e, como as demais componentes são relacionadas com componentes reflexivas, que incidem de várias direções, elas são ponderadas pela resposta do microfone nas demais direções, que é sempre menor que 1.

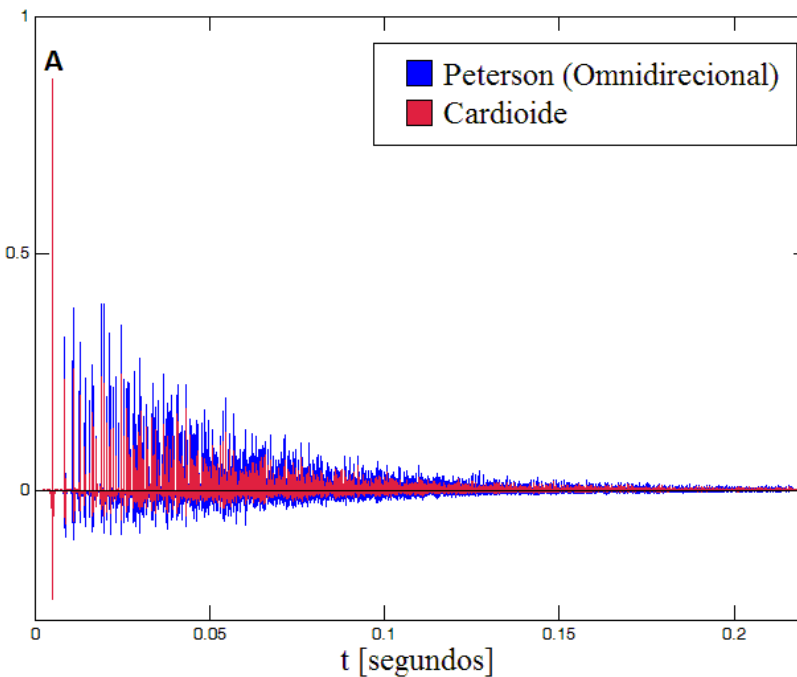


Figura 2.11: Comparação do método de Peterson com o método proposto, utilizando um microfone do tipo cardioide focalizado na fonte sonora.

A Figura 2.12 compara a resposta ao impulso obtida com um microfone do tipo cardioide focalizado em diferentes direções. É possível notar uma redução gradual da componente de linha de visada à

medida que o microfone desfocaliza a fonte sonora, que está fixada em 90° .

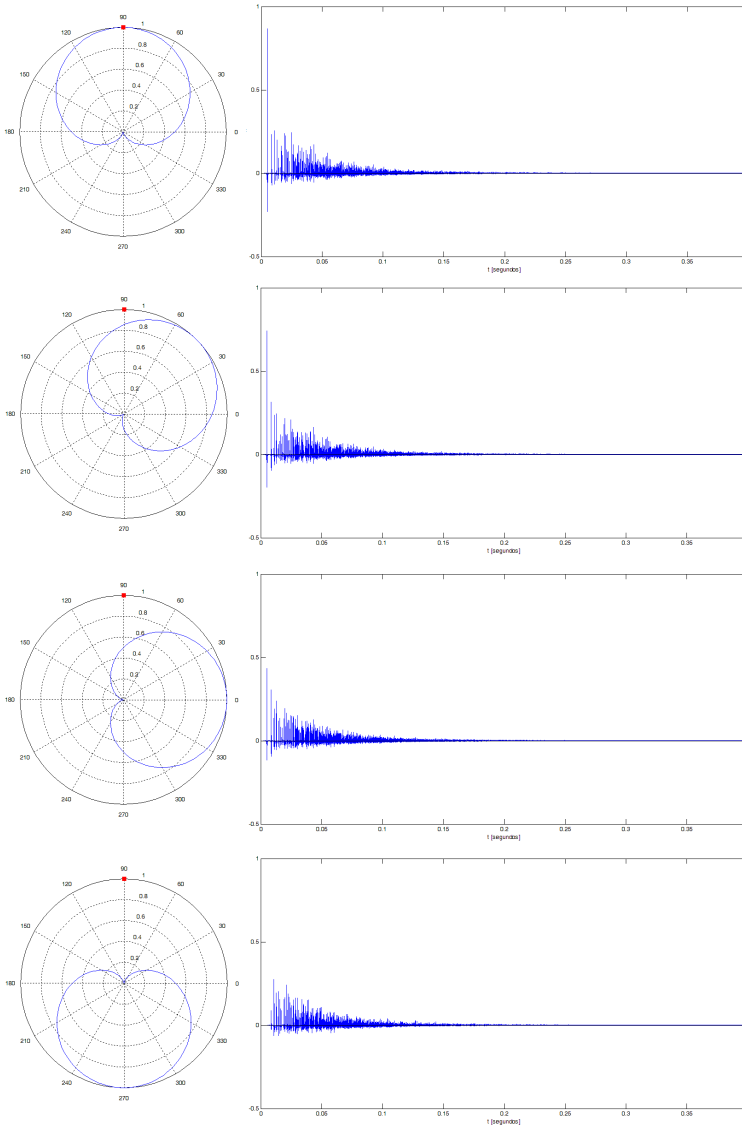


Figura 2.12: Resposta ao impulso obtida com um microfone do tipo cardioide focalizado em diferentes regiões. A fonte sonora está localizada em 90° .

3 O CONFORMADOR DE FEIXES

Os filtros digitais são em geral empregados para realizar uma *filtragem temporal* do sinal sendo processado. Porém, seu emprego em processamento espacial vem sendo intensificado nas duas últimas décadas, em função do desenvolvimento dos sistemas de comunicação sem fio e de processamento de sinais bioelétricos. Na *filtragem espacial*, um número determinado de sensores (comumente denominado *arranjo de sensores – sensor array*) é distribuído no ambiente para captar um sinal transmitido por uma fonte sonora. Com a diversidade de sensores, algumas propriedades dos sinais incidentes que não são passíveis na filtragem temporal são disponíveis, como o ângulo de incidência.

Nesta linha, a *conformação de feixe (beamforming)*, obtida pelo *conformador de feixe (beamformer)*, é o processo de distinção de características de sinais baseada em suas propriedades espaciais [29]. Qualquer configuração do arranjo pode ser utilizada, sendo a linear a mais comum.

São diversas as aplicações que usam o conceito de conformação de feixe, como radares, radioastronomia, sonares, comunicações sem fio e engenharia biomédica. Uma introdução ao assunto, com grande referencial para diversas aplicações, pode ser encontrada em [30].

Neste capítulo, o conceito de filtros espaciais será abordado, bem como técnicas avançadas de conformação de feixes utilizadas na literatura. Uma maior ênfase será dada ao conformador de feixes de mínima variância, linearmente restrito (*linearly constrained minimum variance beamformer- LCMV beamformer*), por ser base para diversos trabalhos de uso de conformadores em aparelhos auditivos.

3.1 DEFINIÇÃO DO CONFORMADOR DE FEIXES

Para apresentar as características do *beamformer*, o modelo anecoico de canal será utilizado. Como apresentado no Capítulo 2, esta abordagem simplifica os cálculos, mas desconsidera algumas características do canal, como a reverberação. Além disso, é considerado que a fonte transmissora da onda está distante o suficiente para que a frente de onda seja considerada plana sem grandes perdas (*far field* - campo distante), como demonstra a Figura 3.1.

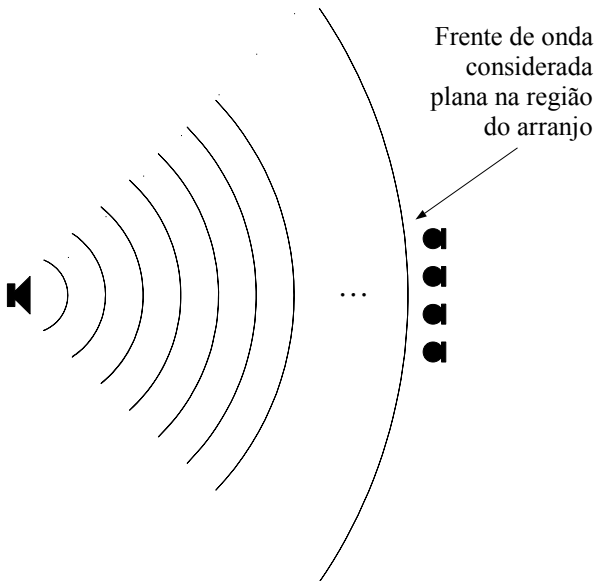


Figura 3.1: Arranjo de sensores e ondas em *far field*

Para as demonstrações, é sempre considerado que os microfones estão distribuídos no arranjo linear de maneira perpendicular à direção de propagação da frente de onda, configuração conhecida como

broadside. Uma outra opção bastante utilizada, em que os microfones do arranjo são dispostos ao longo da direção de propagação da frente de onda, é conhecida como *endfire*, e pode facilmente ser derivada a partir das equações da *broadside*. Ambas as configurações podem ser vistas na Figura 3.2.

Nesta seção, será apresentada a estrutura direta do conformador, sendo posteriormente apresentada uma forma alternativa conhecida como *cancelador de lóbulo lateral generalizado* (*generalized sidelobe canceller - GSC*).

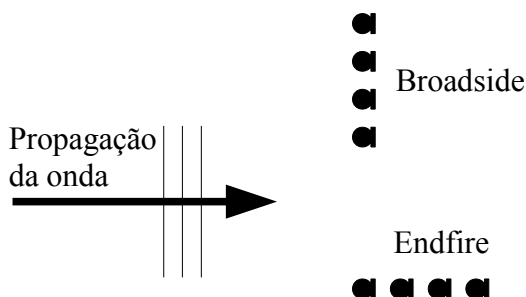


Figura 3.2: Distribuição dos microfones nas configurações *broadside* e *endfire*

3.1.1 Forma direta

A estrutura direta do *beamformer* pode ser vista na Figura 3.3, com K sensores. Um sinal captado pelos sensores do arranjo será ponderado pelos coeficientes dos sensores e o sinal de saída será a soma dos sinais ponderados. Embora qualquer configuração possa ser utilizada, neste trabalho será sempre assumida uma distribuição linear dos microfones, distanciados de d_{mic} metros.

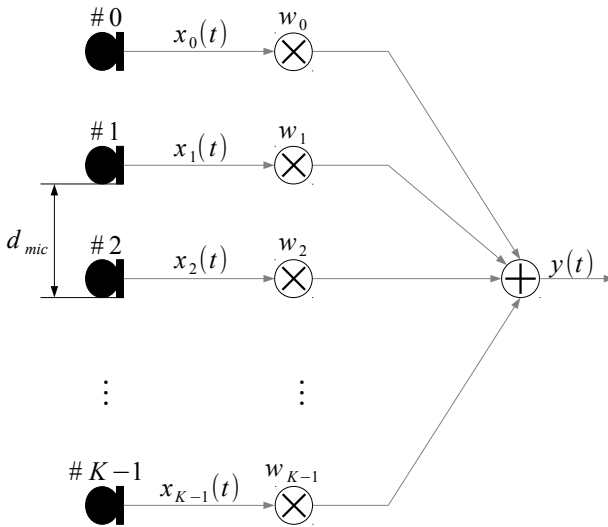


Figura 3.3: Forma direta do beamformer

O sinal de saída do *beamformer* é descrito como:

$$y(t) = \sum_{k=0}^{K-1} w_k x_k(t) \quad (3.1)$$

onde $x_k(t)$ é o sinal incidente no sensor k , t indica um instante de amostragem inteiro, e w_k é o coeficiente do sensor k . Na forma vetorial, a Equação (Erro: Origem da referência não encontrada) pode ser reescrita como:

$$y(t) = \mathbf{w}' \mathbf{x}(t) \quad (3.2)$$

onde \mathbf{w} e $\mathbf{x}(t)$ são respectivamente o *vetor de coeficientes* e o *vetor de entrada instantâneo (snapshot input vector)*, definidos como:

$$\mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{K-1} \end{bmatrix} \quad (3.3)$$

$$\mathbf{x}(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{K-1}(t) \end{bmatrix} \quad (3.4)$$

O valor dos coeficientes w_k depende de como o conformador de feixe é projetado, podendo ser fixo ou constantemente alterado por meio de um processamento adaptativo de sinais. É através desta ponderação que o conformador de feixe direciona a resposta espacial do arranjo.

A Figura 3.4 mostra o esquema de propagação de um sinal de banda estreita $s(t)$ transmitido por uma fonte sonora S , considerando os sensores linearmente espaçados a uma distância d_{mic} e a frente de onda plana (*far field*). Pela figura, é possível observar que o arranjo de sensores amostra a onda propagante.

Dado um ângulo de incidência θ , o atraso espacial sofrido pela frente de onda para trafegar do sensor #0 ao sensor #1 é:

$$d_{01} = d_{mic} \sin(\theta) \quad (3.5)$$

Da mesma forma, a fase acumulada entre dois sensores vizinhos é:

$$\phi_{\theta} = \frac{2\pi d_{mic}}{\lambda} \sin(\theta) \quad (3.6)$$

onde λ é o comprimento de onda do sinal $s(t)$, definido como $\lambda = c/f_0$, sendo f_0 a frequência da onda $s(t)$. A fase acumulada entre

dois sensores é também chamada de *ângulo elétrico da onda*, e pode ser reescrita como:

$$\Phi_\theta = 2\pi f_0 \tau(\theta) \quad (3.7)$$

onde $\tau(\theta)$ é o tempo de atraso de propagação entre dois sensores adjacentes, descrito como

$$\tau(\theta) = \frac{d_{\text{mic}}}{c} \sin(\theta) \quad (3.8)$$

Escrito desta forma, é possível observar que o ângulo elétrico da onda tem para o processamento espacial o mesmo papel que a frequência angular no processamento temporal, da mesma forma que o atraso de propagação entre dois sensores assume o papel do período de amostragem do processamento temporal.

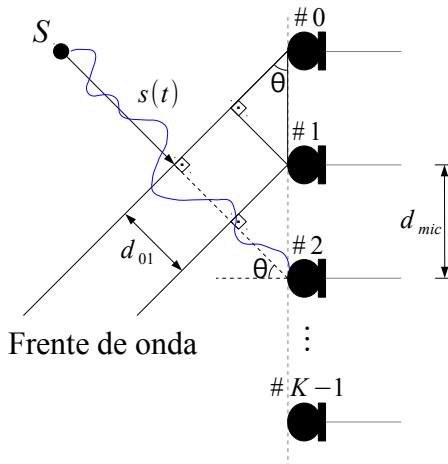


Figura 3.4: Recepção da onda pelo arranjo de sensores

Da mesma forma que em processamento temporal existe o critério de *Nyquist* para a amostragem de sinais, expresso em termos do período de amostragem temporal, em processamento espacial também há uma limitação semelhante. Tendo a relação entre o tempo de atraso de propagação $\tau(\theta)$ e período de amostragem, a limitação da amostragem espacial é:

$$\tau(\theta) \leq \frac{1}{2f_0} \quad (3.9)$$

Em termos da distância entre os sensores, o teorema da amostragem fica

$$d_{mic} \leq \frac{\lambda}{2 \text{sen}(\theta)} \quad (3.10)$$

e, por fim, tendo que uma medida mais conservadora é considerar θ igual a 90° , resultando o teorema em:

$$d_{mic} \leq \frac{\lambda}{2} \quad (3.11)$$

A resposta do arranjo é encontrada da mesma forma que em sistemas de processamento temporal. Sendo o sinal $s(t)$ uma exponencial complexa $e^{i2\pi f t}$ incidente num arranjo de sensores linear, e considerando que a onda incide primeiramente no sensor #0 (sensor de referência), o sinal incidente em um dado sensor k do arranjo é calculado como

$$x_k(t) = e^{i2\pi f [t - k \cdot \tau(\theta)]} \quad (3.12)$$

onde $\tau(\theta)$ é o tempo de atraso de propagação entre dois sensores adjacentes (Equação (3.8)) e i é o operador de número imaginário. Desta forma, o sinal de saída do conformador de feixe, dado pela Equação (3.5), é então:

$$\begin{aligned}
y(t) &= \sum_{k=0}^{K-1} w_k \cdot x_k(t) \\
&= \sum_{k=0}^{K-1} w_k \cdot e^{i2\pi f[t-k\cdot\tau(\theta)]} \\
&= e^{i2\pi f t} \sum_{k=0}^{K-1} w_k \cdot e^{-i2\pi f k \tau(\theta)} \\
&= e^{i2\pi f t} \cdot W(\theta)
\end{aligned} \tag{3.13}$$

Na equação anterior, o termo $W(\theta)$ é a *resposta espacial do arranjo*, definida por

$$W(\theta) = \sum_{k=0}^{K-1} w_k \cdot e^{-i2\pi f k \tau(\theta)} \tag{3.14}$$

E em notação matricial:

$$W(\theta) = \mathbf{w}^t \cdot \mathbf{e}(\theta) \tag{3.15}$$

onde \mathbf{w} é definido na Equação (3.7) e

$$\mathbf{e}(\theta) = \begin{bmatrix} 1 \\ e^{-i2\pi f \tau(\theta)} \\ e^{-i4\pi f \tau(\theta)} \\ \vdots \\ e^{-i2(K-1)\pi f \tau(\theta)} \end{bmatrix} \tag{3.16}$$

é o *vetor de direcionamento* do arranjo (*steering vector*) [31]. A magnitude de $W(\theta)$ define o diagrama de feixe do arranjo (*beam pattern*) [32].

3.1.2 O conformador de feixes banda larga

O conformador de feixe apresentado na seção anterior é comumente chamado de *conformador de feixe de banda estreita*

(*narrowband beamformer*) [33]. Este nome vem do fato de que o conformador não faz distinção da frequência do sinal de entrada. Ou seja, o conformador de feixe da seção anterior possui uma resposta plana em todo o eixo de frequências. Na Figura 3.5, há um exemplo de resposta de um conformador de feixe de banda estreita.

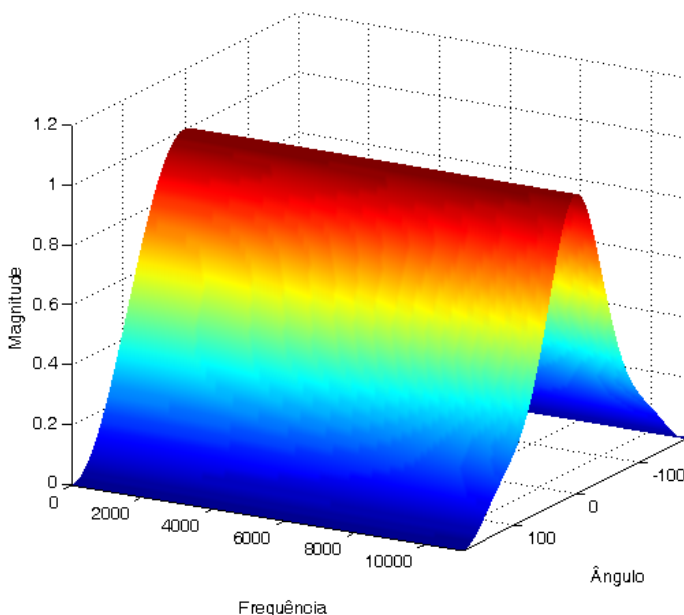


Figura 3.5: Resposta de um conformador de feixe banda estreita

Para melhor processar sinais de banda larga, o *conformador de feixe de banda larga (broadband beamformer)* pode ser utilizado. Nesta estrutura (Figura 3.6), um filtro transversal de J coeficientes é conectado à saída de cada um dos K sensores, em substituição ao coeficiente único da estrutura de banda estreita. Assim, é possível obter uma resposta em frequência diferente para cada ângulo de incidência (Figura 3.7).

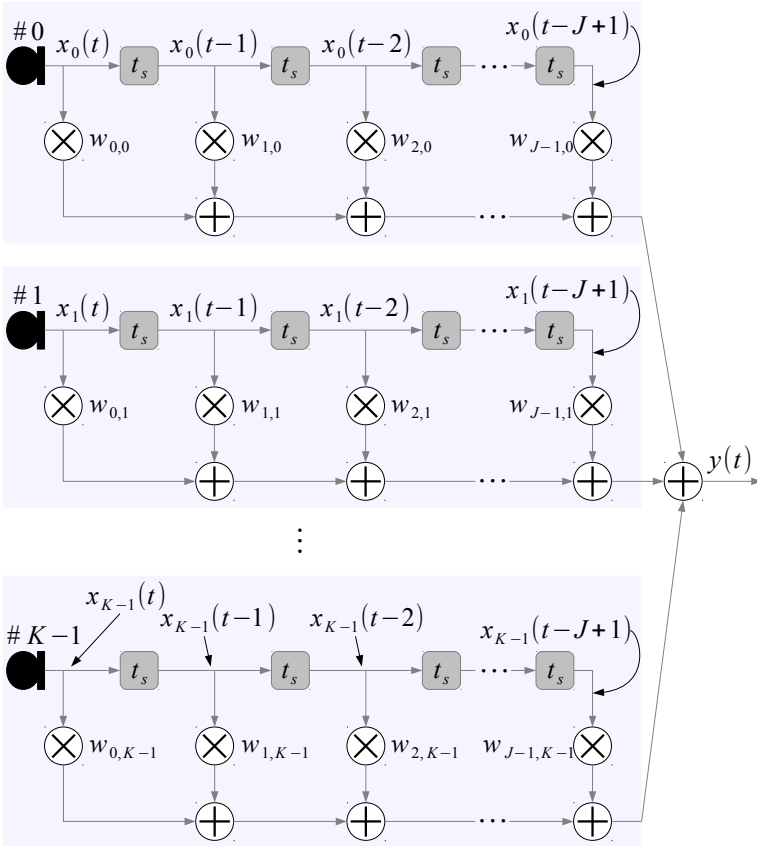


Figura 3.6: O conformador de feixe de banda larga na estrutura direta

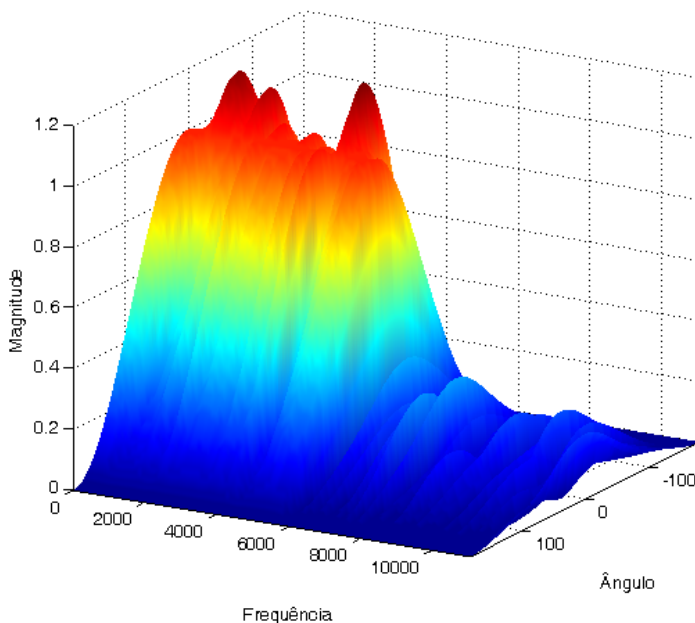


Figura 3.7: Resposta de um conformador de feixe de banda larga

Seguindo a mesma forma da Equação (3.5), o sinal de saída do conformador de feixe de banda larga é

$$y(t) = \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} \cdot x_k(t-j) \quad (3.17)$$

onde um determinado coeficiente $w_{j,k}$ indica o coeficiente j do filtro transversal associado ao sensor k , e uma determinada amostra $x_k(t-j)$ indica o sinal do sensor k atrasado de j amostras a partir do instante de tempo atual t . Na forma vetorial, a equação anterior fica como

$$y(t) = \mathbf{w}^t \mathbf{x}(t) \quad (3.18)$$

sendo o vetor \mathbf{w} definido como

$$\mathbf{w} = \begin{bmatrix} \mathbf{w}_0 \\ \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_{J-1} \end{bmatrix} \quad (3.19)$$

$$\mathbf{w}_j = \begin{bmatrix} w_{j,0} \\ w_{j,1} \\ \vdots \\ w_{j,K-1} \end{bmatrix} \quad (3.20)$$

e $\mathbf{x}(t)$:

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}(t-1) \\ \vdots \\ \mathbf{x}(t-J+1) \end{bmatrix} \quad (3.21)$$

$$\mathbf{x}(t-j) = \begin{bmatrix} x_0(t-j) \\ x_1(t-j) \\ \vdots \\ x_{K-1}(t-j) \end{bmatrix} \quad (3.22)$$

Para encontrar a resposta do conformador de banda larga, é utilizado o mesmo artifício do conformador de banda estreita. Considera-se que uma exponencial complexa incide no arranjo primeiramente no sensor #0. A amostra do sinal de entrada presente no *tap* j do filtro transversal associado ao sensor k é:

$$x_k(n-j) = e^{j2\pi f[n-j-k\tau(\theta)]} \quad (3.23)$$

E o sinal de saída é então

$$\begin{aligned}
y(t) &= \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} e^{i2\pi f [t-j-k\tau(\theta)]} \\
&= e^{i2\pi f t} \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} e^{-i2\pi f [j+k\tau(\theta)]} \\
&= e^{i2\pi f t} W(f, \theta)
\end{aligned} \tag{3.24}$$

onde $W(f, \theta)$ é a resposta do conformador, dada por:

$$W(f, \theta) = \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} e^{-i2\pi f [j+k\tau(\theta)]} \tag{3.25}$$

Na forma vetorial, a equação anterior fica

$$W(f, \theta) = \mathbf{w}^t \mathbf{e}(f, \theta) \tag{3.26}$$

onde \mathbf{w} é definido na Equação (Erro: Origem da referência não encontrada) e

$$\mathbf{e}(f, \theta) = \begin{bmatrix} e_0 \\ e_1 \\ \vdots \\ e_{J-1} \end{bmatrix} \tag{3.27}$$

$$\mathbf{e}_j = \begin{bmatrix} e^{-i2\pi f j} \\ e^{-i2\pi f [j+\tau(\theta)]} \\ e^{-i2\pi f [j+2\tau(\theta)]} \\ \vdots \\ e^{-i2\pi f [j+(K-1)\tau(\theta)]} \end{bmatrix} \tag{3.28}$$

3.2 O PROBLEMA DE OTIMIZAÇÃO

Os coeficientes do conformador indicam de que forma os sinais incidentes serão processados. O projetista define os coeficientes de forma a moldar a resposta do arranjo, para enfatizar ou atenuar um determinado par ângulo de incidência-frequência. Uma alternativa aos coeficientes fixos é usar *coeficientes adaptativos*, utilizando um *algoritmo de adaptação* para que uma determinada função seja executada pelo conformador. Este trabalho utiliza coeficientes adaptativos, de forma que o sistema consiga eliminar as interferências presentes no meio mesmo com a movimentação das suas fontes geradoras.

Os algoritmos adaptativos utilizados neste trabalho utilizam como critério a Mínima Variância Linearmente Restrita (LCMV - *Linearly Constrained Minimum Variance*), que tem como base a minimização da potência do sinal de saída do conformador enquanto algumas restrições são satisfeitas. É através das restrições que a resposta do conformador para um sinal de interesse é imposta. Não havendo restrições, a saída do conformador é zero.

Sabendo que o sinal de saída de um conformador é dado pela Equação (3.6), a sua potência de saída é:

$$\begin{aligned}
 P_y &= E[y^2(t)] \\
 &= E[\mathbf{w}^t \mathbf{x}(t) \mathbf{x}^t(t) \mathbf{w}] \\
 &= \mathbf{w}^t \mathbf{R}_{xx} \mathbf{w}
 \end{aligned}
 \tag{3.29}$$

onde \mathbf{R}_{xx} é a matriz de autocorrelação do sinal de entrada $\mathbf{x}(t)$. Assim, o problema de otimização para o algoritmo de Frost é um problema de pesquisa operacional, do tipo programação linear [34] [35], dado por:

$$\begin{aligned}
& \text{minimizar} && \mathbf{w}^t \mathbf{R}_{xx} \mathbf{w} \\
& && \mathbf{w} \\
& \text{sujeito a} && \mathbf{C}^t \mathbf{w} = \boldsymbol{\xi}
\end{aligned} \tag{3.30}$$

onde \mathbf{C} é chamada *matriz de restrição* e $\boldsymbol{\xi}$ é chamado *vetor de resposta às restrições*. Se G restrições são impostas a um conformador de J sensores e K coeficientes (JK coeficientes no total), \mathbf{C} é uma matriz de tamanho $JK \times G$ e $\boldsymbol{\xi}$ é um vetor de tamanho $G \times 1$.

O *valor ótimo* \mathbf{w}_{opt} é o valor de \mathbf{w} que minimiza a potência de saída, obedecendo as restrições do problema. Através do método dos multiplicadores de Lagrange [32] [29], o valor ótimo é:

$$\mathbf{w}_{\text{opt}} = \mathbf{R}_{xx}^{-1} \mathbf{C} [\mathbf{C}^t \mathbf{R}_{xx}^{-1} \mathbf{C}]^{-1} \boldsymbol{\xi} \tag{3.31}$$

3.2.1 O conformador LCMV adaptativo

A Equação (3.31) mostra que é necessário conhecer a matriz \mathbf{R}_{xx} para encontrar o valor ótimo dos coeficientes, o que nem sempre é possível. Uma solução para este problema é utilizar um algoritmo adaptativo, de forma que informações da matriz de autocorrelação sejam gradativamente aprendida pelo sistema.

A derivação do algoritmo inicia com a atribuição de um valor inicial que satisfaz às restrições ao vetor \mathbf{w} , ou:

$$\mathbf{w}(0) = \mathbf{C} [\mathbf{C}^t \mathbf{C}]^{-1} \boldsymbol{\xi} \tag{3.32}$$

A cada iteração, o vetor \mathbf{w} é movido na direção inversa do gradiente da potência de saída, limitado pelo plano das restrições.

Novamente, utilizando os multiplicadores de Lagrange, a versão final do algoritmo, chamada de LMS (*Least Mean Squares*) estocástico com restrições é encontrada:

$$\begin{aligned} \mathbf{w}(0) &= \mathbf{F} \\ \mathbf{w}(t+1) &= \mathbf{P}[\mathbf{w}(t) - \mu y(t)\mathbf{x}(t)] + \mathbf{F} \end{aligned} \quad (3.33)$$

onde $\mathbf{w}(0)$ é o valor inicial atribuído aos coeficientes do conformador, $\mathbf{w}(t+1)$ é o valor atribuído ao vetor de coeficientes após a iteração t , e os vetores \mathbf{F} e \mathbf{P} são respectivamente

$$\mathbf{F} = \mathbf{C}[\mathbf{C}'\mathbf{C}]^{-1}\boldsymbol{\xi} \quad (3.34)$$

$$\mathbf{P} = \mathbf{I} - \mathbf{C}[\mathbf{C}'\mathbf{C}]^{-1}\mathbf{C}' \quad (3.35)$$

3.3 RESTRIÇÕES LINEARES

Definido o critério LCMV dos conformadores de feixe, alguns tipos de restrições serão apresentadas a seguir.

3.3.1 As restrições de Frost

Em seu célebre artigo de 1972 [8], Frost propôs um algoritmo de adaptação dos coeficientes do conformador de feixe baseado no método do gradiente. O algoritmo de Frost, chamado de LMS restrito (*Constrained LMS - CLMS*), minimiza a potência do sinal de saída do conformador tendo como restrição o não cancelamento de sinais incidentes a zero graus no arranjo, sendo então um conformador LCMV. Assim, a técnica de minimização da potência age somente nos sinais com ângulo de incidência no arranjo diferente de zero graus.

A Figura 3.8 apresenta a estrutura do conformador de feixe utilizado por Frost. A estrutura superior mostra um conformador

comum, igual ao apresentado anteriormente, e a estrutura inferior apresenta um filtro transversal.

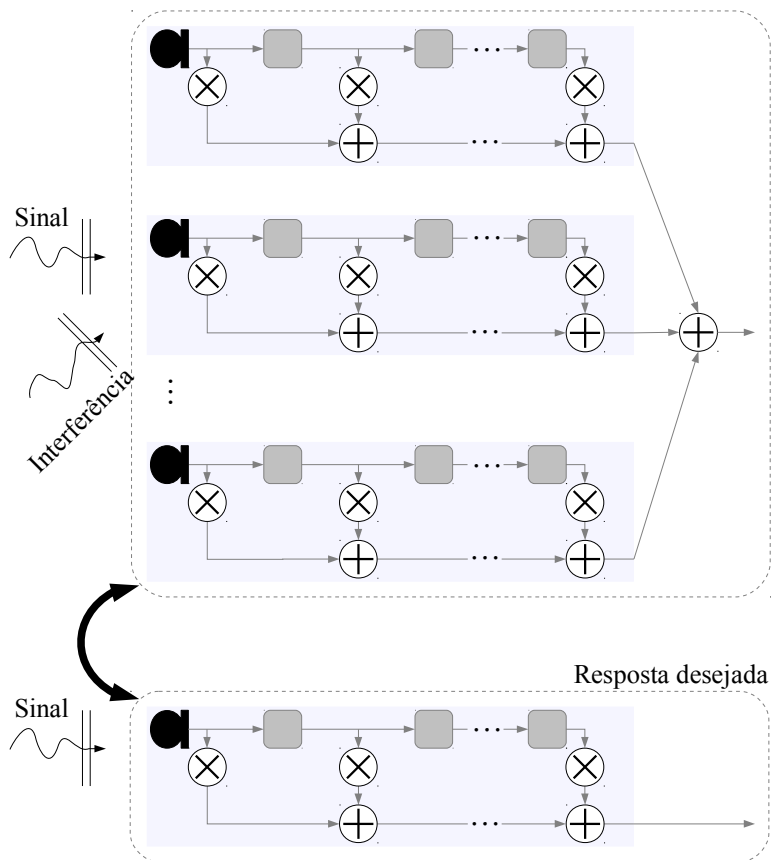


Figura 3.8: Estrutura do conformador de feixe utilizado por Frost

O algoritmo de Frost faz com que o conformador atue no sinal incidente a zero graus, denotado na figura como *senal*, de forma que a resposta resultante seja a resposta do filtro transversal, denominada *resposta desejada*. Enquanto isso, os outros sinais presentes no

ambiente, denotados na figura como *interferência*, são atenuados. A natureza adaptativa do conformador garante que, mesmo com a alteração dos seus ângulos de incidência e/ou faixa de frequência, os sinais de interferência serão atenuados.

Embora a linha de foco seja assumida como sendo a linha perpendicular à linha de propagação do sinal de interesse (arranjo em *broadside*), Frost previu também a rotação do arranjo, com a aplicação de blocos de atraso logo após os microfones. A função dos atrasos é fazer com que as amostras do sinal de interesse estejam sincronizadas na entrada dos filtros. A Figura 3.9 mostra um arranjo rotacionado em alguns graus. Os blocos de tamanho variável simbolizam blocos de atraso, de forma que blocos maiores indicam maiores atrasos temporais.

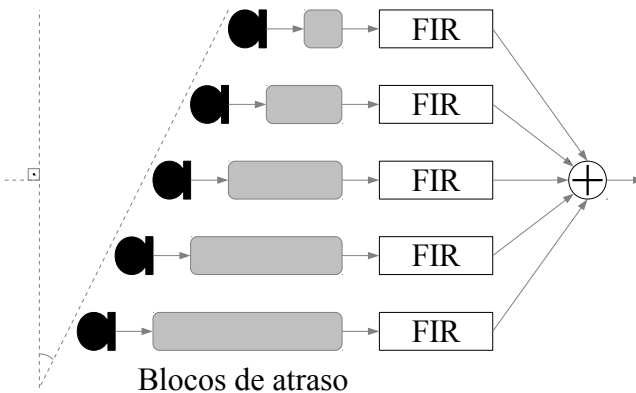


Figura 3.9: Arranjo de microfones rotacionado via adição de blocos de atraso

Tendo o algoritmo adaptativo da Equação (3.33), a matriz de restrição para as restrições de Frost é definida como

$$C = [c_1 \ c_2 \ \cdots \ c_J] \quad (3.36)$$

sendo c_j um vetor coluna de KJ elementos agrupados sob a forma de J grupos de K elementos, onde o grupo de número j tem os elementos iguais a um e os elementos dos outros $J-1$ grupos são iguais a zero:

$$\mathbf{c}_j = \left. \begin{array}{c} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{array} \right\} \begin{array}{l} K \\ \vdots \\ K \\ \text{grupo } j \text{ de } K \text{ elementos} \\ K \\ \vdots \\ K \end{array} \quad (3.37)$$

Uma forma simples de gerar a matriz \mathbf{C} é através de

$$\mathbf{C} = \mathbf{I}_J \circ \mathbf{1}_K \quad (3.38)$$

onde \mathbf{I}_J é uma matriz identidade de dimensão J , $\mathbf{1}_K$ é um vetor coluna com K elementos iguais a 1 e o símbolo \circ denota o produto de Kronecker [36].

O vetor de resposta às restrições, $\boldsymbol{\xi}$, é um vetor de J elementos com os coeficientes do filtro transversal que define a resposta desejada do conformador para sinais incidentes a zero graus (Figura 3.8):

$$\boldsymbol{\xi} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_J \end{bmatrix} \quad (3.39)$$

Uma forma de entender as restrições do algoritmo de Frost é resolver a equação vetorial das restrições. Cada linha da matriz \mathbf{C}^t (ou coluna da matriz \mathbf{C}), que multiplica o vetor de coeficientes \mathbf{w} do conformador, tem alguns elementos iguais a um e outros iguais a zero. Assim, está sendo feita uma seleção dos elementos de \mathbf{w} cuja soma resultará no valor de um coeficiente do filtro transversal. Para a linha j da matriz \mathbf{C}^t , dada a organização do vetor \mathbf{w} (Equações (3.30) e (3.31)), os elementos de \mathbf{w} selecionados para a soma são os coeficientes do *tap* j dos filtros transversais na saída de cada sensor, sendo o resultado da soma igual ao coeficiente do *tap* j do filtro transversal de resposta desejada (Figura 3.10). Por fim, a equação das restrições resultará em J equações com K variáveis cada.

Uma grande vantagem do algoritmo de Frost sobre outros algoritmos existentes na época da sua publicação é que o vetor de coeficientes sempre está de acordo com as restrições [8]. Devido a arredondamentos, quantização, entre outras características presentes em sistemas computacionais, outros algoritmos, como o de projeção do Gradiente [29], tinham o comportamento característico de perder as restrições ao longo das iterações [32].

Como limitação, o algoritmo de Frost tem o cancelamento do sinal de interesse quando este sofre um desvio da fonte de foco. Para que o algoritmo opere de forma satisfatória, é necessário que o ângulo de incidência do sinal desejado no arranjo seja de zero graus. Com um desvio de alguns graus, o sinal sai da faixa de sinal desejado e entra na faixa que o classifica como interferência, efetuando o cancelamento. Por último, as restrições de Frost são definidas unicamente para a faixa de interesse, não havendo controle da resposta do conformador para outros ângulos de incidência. Este controle é útil por exemplo no cancelamento

de uma interferência conhecida *a priori*, deixando a cargo do algoritmo adaptativo o cancelamento de outras interferências não conhecidas.

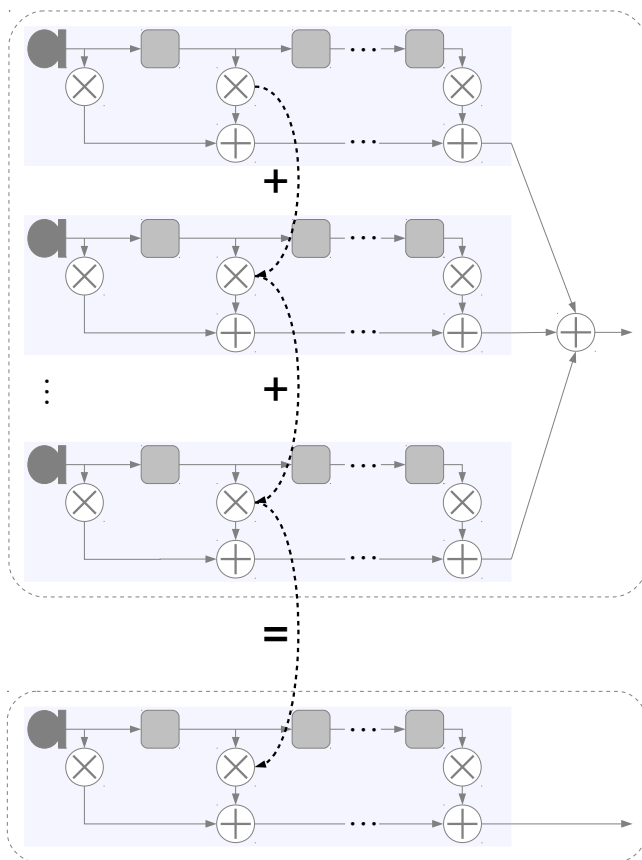


Figura 3.10: Filtro de resposta desejada como a soma dos coeficientes dos filtros associados aos microfones

3.3.2 Restrições pontuais

As restrições pontuais definem a resposta do conformador para um determinado par ângulo de incidência-frequência.

A partir da resposta do conformador de feixe da Equação (3.30), repetida aqui para comodidade:

$$W(f, \theta) = \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} e^{-i2\pi f[j+k\tau(\theta)]} \quad (3.40)$$

a resposta do conformador para um determinado ângulo de incidência θ_0 e uma determinada frequência f_0 é

$$\begin{aligned} W(f_0, \theta_0) &= \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} e^{-i2\pi f_0[j+k\tau(\theta_0)]} \\ &= \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} \cos(2\pi f_0[j+k\tau(\theta_0)]) + \\ &\quad -i \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} \operatorname{sen}(2\pi f_0[j+k\tau(\theta_0)]) \end{aligned} \quad (3.41)$$

Impondo na resposta do conformador no ângulo de incidência e na frequência indicada uma magnitude A e fase Ω :

$$\begin{aligned} W(f_0, \theta_0) &= \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} \cos(2\pi f_0[j+k\tau(\theta_0)]) + \\ &\quad -i \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} \operatorname{sen}(2\pi f_0[j+k\tau(\theta_0)]) \quad (3.42) \\ &= A \cos(\Omega) - i A \operatorname{sen}(\Omega) \end{aligned}$$

As equações de restrição podem então ser definidas como:

$$\begin{cases} \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} \cos(2\pi f_0 [j+k\tau(\theta_0)]) = A \cos(\Omega) \\ \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} w_{j,k} \text{sen}(2\pi f_0 [j+k\tau(\theta_0)]) = A \text{sen}(\Omega) \end{cases} \quad (3.43)$$

ou, na forma vetorial:

$$\begin{cases} \mathbf{c}'_{\cos} \mathbf{w} = f_{\cos} \\ \mathbf{c}'_{\text{sen}} \mathbf{w} = f_{\text{sen}} \end{cases} \quad (3.44)$$

onde

$$\mathbf{c}'_{\cos} = \begin{matrix} 1 \\ \cos(2\pi f_0) \\ \cos(4\pi f_0) \\ \cos(6\pi f_0) \\ \vdots \\ \cos(2\pi f_0 (J-1)) \\ \cdots \\ \cos(2\pi f_0 \tau(\theta_0)) \\ \cos(2\pi f_0 [1+\tau(\theta_0)]) \\ \cos(2\pi f_0 [2+\tau(\theta_0)]) \\ \vdots \\ \cos(2\pi f_0 [(J-1)+\tau(\theta_0)]) \\ \cdots \\ \cos(2\pi f_0 2\tau(\theta_0)) \\ \cos(2\pi f_0 [1+2\tau(\theta_0)]) \\ \cos(2\pi f_0 [2+2\tau(\theta_0)]) \\ \vdots \\ \cos(2\pi f_0 [(J-1)+2\tau(\theta_0)]) \\ \cdots \\ \cos(2\pi f_0 ((K-1)\theta_0)) \\ \cos(2\pi f_0 [1+(K-1)\tau(\theta_0)]) \\ \cos(2\pi f_0 [2+(K-1)\tau(\theta_0)]) \\ \vdots \\ \cos(2\pi f_0 [(J-1)+(K-1)\tau(\theta_0)]) \end{matrix} \quad \mathbf{e} \quad \mathbf{c}'_{\text{sen}} = \begin{matrix} 0 \\ \text{sen}(2\pi f_0) \\ \text{sen}(4\pi f_0) \\ \text{sen}(6\pi f_0) \\ \vdots \\ \text{sen}(2\pi f_0 (J-1)) \\ \cdots \\ \text{sen}(2\pi f_0 \tau(\theta_0)) \\ \text{sen}(2\pi f_0 [1+\tau(\theta_0)]) \\ \text{sen}(2\pi f_0 [2+\tau(\theta_0)]) \\ \vdots \\ \text{sen}(2\pi f_0 [(J-1)+\tau(\theta_0)]) \\ \cdots \\ \text{sen}(2\pi f_0 2\tau(\theta_0)) \\ \text{sen}(2\pi f_0 [1+2\tau(\theta_0)]) \\ \text{sen}(2\pi f_0 [2+2\tau(\theta_0)]) \\ \vdots \\ \text{sen}(2\pi f_0 [(J-1)+2\tau(\theta_0)]) \\ \cdots \\ \text{sen}(2\pi f_0 ((K-1)\theta_0)) \\ \text{sen}(2\pi f_0 [1+(K-1)\tau(\theta_0)]) \\ \text{sen}(2\pi f_0 [2+(K-1)\tau(\theta_0)]) \\ \vdots \\ \text{sen}(2\pi f_0 [(J-1)+(K-1)\tau(\theta_0)]) \end{matrix} \quad (3.45)$$

$$\begin{aligned} f_{\cos} &= A \cos(\Omega) \\ f_{\text{sen}} &= A \text{sen}(\Omega) \end{aligned} \quad (3.46)$$

A definição geral das restrições do algoritmo, presente na Equação (3.30), pode incluir as restrições pontuais. Isto é feito através da concatenação dos vetores da Equação (3.44). A definição das restrições torna-se portanto:

$$C' \mathbf{w} = \boldsymbol{\xi} \quad (3.47)$$

sendo:

$$C = [\mathbf{c}_{\cos} \ : \ \mathbf{c}_{\text{sen}}] \quad (3.48)$$

$$\boldsymbol{\xi} = \begin{bmatrix} f_{\cos} \\ f_{\text{sen}} \end{bmatrix} \quad (3.49)$$

Assim, com as restrições pontuais é possível incluir no problema informações sobre os sinais desejados e/ou interferentes. Por exemplo, se o ângulo de incidência do sinal desejado é conhecido, é possível definir a resposta do conformador para determinado ângulo de incidência. Da mesma forma, se é sabido que há um sinal de interferência forte numa frequência específica (sinal de banda estreita) incidindo de um certo ângulo, pode-se forçar a resposta do conformador a zero neste ponto.

Da mesma forma que nas restrições de Frost, o desempenho do conformador com uma restrição pontual é deteriorada por incertezas nos parâmetros. Se está definida uma resposta para um determinado par ângulo de incidência-frequência, sinais fora deste ponto são minimizados pelo algoritmo. Assim, qualquer desvio sofrido pelo sinal de interesse resulta na sua minimização. A vantagem das restrições pontuais sobre as restrições de Frost é que há a possibilidade de utilizar múltiplas restrições, cobrindo uma faixa de valores de ângulo e

frequência. Desta forma, uma margem de erro nos parâmetros é introduzida, garantindo um bom desempenho mesmo com pequenas variações na frequência e no ângulo de incidência.

A definição de múltiplas restrições pontuais se dá de forma análoga à Equação (3.47), sendo as matrizes \mathbf{C} e $\boldsymbol{\xi}$ definidas respectivamente como:

$$\mathbf{C} = [\mathbf{c}_1 \quad \mathbf{c}_2 \quad \cdots \quad \mathbf{c}_G \quad ; \quad s_1 \quad s_2 \quad \cdots \quad s_G] \quad (3.50)$$

onde \mathbf{c}_i é o vetor \mathbf{c}_{\cos} da equação (3.45) para um conjunto de parâmetros $\{f_i, \theta_i\}$, s_j é o vetor \mathbf{c}_{sen} da equação (3.45) para um conjunto de parâmetros $\{f_j, \theta_j\}$ e $2G$ passa a ser o número de restrições sendo impostas; e

$$\boldsymbol{\xi} = \begin{bmatrix} f_{c_1} \\ f_{c_2} \\ \vdots \\ f_{c_G} \\ \cdots \\ f_{s_1} \\ f_{s_2} \\ \vdots \\ f_{s_G} \end{bmatrix} \quad (3.51)$$

onde $f_{c_i} = A_i \cos(\Omega_i)$ é a resposta à restrição associada ao vetor de restrição \mathbf{c}_i e $f_{s_j} = A_j \text{sen}(\Omega_j)$ é a resposta à restrição associada ao vetor de restrição \mathbf{s}_j .

Para definir restrições para uma área espacial, deve-se dividir o espaço desejado em blocos de ângulo e frequência, e para cada bloco, definir uma restrição pontual (Figura 3.11). Porém, um cuidado adicional deve ser tomado na definição das restrições: caso muitas restrições sejam utilizadas, é possível que a precisão numérica utilizada não consiga representar duas restrições pontuais diferentes, resultando numa matriz de restrição mal condicionada devido à falta da

independência linear das suas colunas. O mau condicionamento da matriz pode tornar o sistema de Equações (3.47) incompatível. Da mesma forma, caso poucas restrições sejam utilizadas, uma restrição pontual será responsável por um bloco grande de frequências e/ou ângulos dentro da faixa total. Como consequência, é possível que haja minimização entre as restrições pontuais, resultando numa resposta distorcida em relação à resposta desejada do conformador, como pode ser visto na Figura 3.12. Na figura, as frequências entre as restrições pontuais, que estão em vermelho, foram minimizadas pelo algoritmo, obtendo uma resposta diferente da desejada.

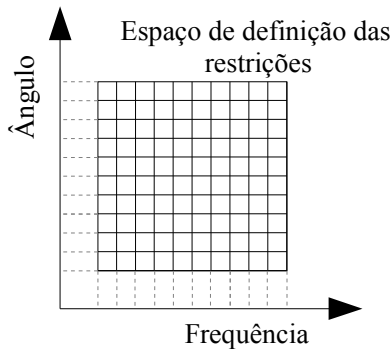


Figura 3.11: Divisão em blocos de uma área espacial de ângulo e frequência para definição de restrições pontuais múltiplas.

A cada bloco estará definida uma restrição pontual.

Por fim, a atuação das múltiplas restrições pontuais se dá às custas dos graus de liberdade utilizados na minimização das interferências. Quanto maior o número de restrições utilizadas, menor será o número de graus de liberdade disponíveis. Os graus de liberdade podem ser aumentados com o aumento da ordem dos filtros ou do número de sensores. Esta solução, porém, aumenta a complexidade do algoritmo adaptativo, podendo até impossibilitar o uso do conformador para a aplicação desejada.

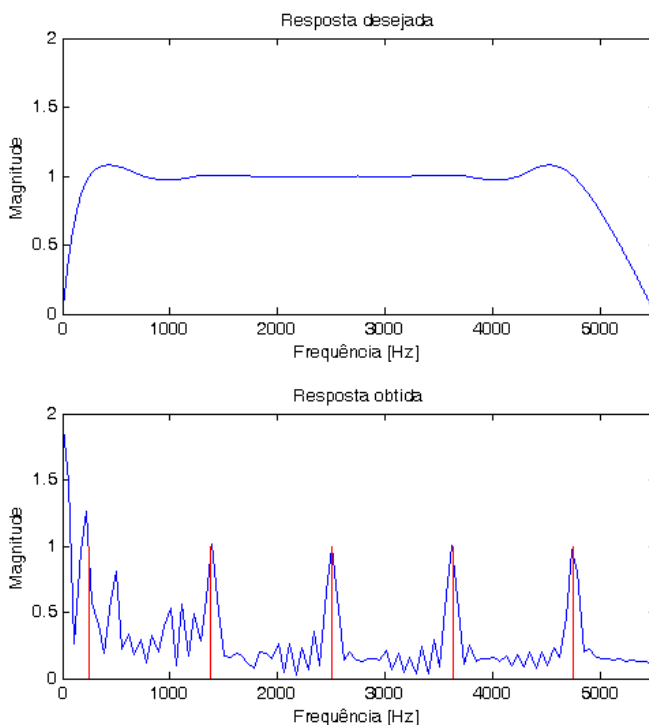


Figura 3.12: Resposta do conformador obtida quando poucas restrições pontuais são utilizadas.

3.3.3 Restrições autovetoriais

As restrições autovetoriais estendem as restrições pontuais de forma que a resposta resultante é a mais próxima possível da resposta desejada, usando menos restrições que as restrições pontuais. Foi proposta por Buckley em [37] e outras referências são [32] e [33].

A definição das restrições autovetoriais parte da matriz de restrições pontuais C . Utilizando a decomposição por valores singulares (*Singular Value Decomposition* - SVD), a matriz de restrições, com $2G$ restrições, pode ser descrita como:

$$C^t = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^t \quad (3.52)$$

onde

$$U = [u_1 \quad u_2 \quad \cdots \quad u_{2G}] \quad (3.53)$$

contém os vetores singulares esquerdos de C ,

$$V = [v_1 \quad v_2 \quad \cdots \quad v_{JK}] \quad (3.54)$$

contém os vetores singulares direitos de C e

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_P \end{bmatrix} \quad (3.55)$$

é a matriz diagonal formada pelos P valores singulares de C^t , ordenados de forma que $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots \geq \sigma_P > 0$. Selecionando os S valores singulares mais significativos, a Equação (3.52) se reduz a:

$$C^t \approx U_S \Sigma_S V_S^t \quad (3.56)$$

onde

$$U_S = [u_1 \quad u_2 \quad \cdots \quad u_S] \quad (3.57)$$

$$V_S = [v_1 \quad v_2 \quad \cdots \quad v_S] \quad (3.58)$$

$$\mathbf{\Sigma}_S = \begin{bmatrix} \sigma_1 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_S \end{bmatrix} \quad (3.59)$$

Utilizando esta representação da matriz \mathbf{C} na Equação (3.47), obtém-se:

$$\mathbf{U}_S \mathbf{\Sigma}_S \mathbf{V}_S^t \mathbf{w} = \mathbf{\xi} \quad (3.60)$$

que manipulado resulta em

$$\mathbf{V}_S^t \mathbf{w} = \mathbf{\Sigma}_S^{-1} \mathbf{U}_S^t \mathbf{\xi} \quad (3.61)$$

Comparando as duas Equações (3.47) e (3.61), é possível identificar a nova forma da matriz de restrições \mathbf{C}_{SVD} e do vetor de resposta às restrições $\mathbf{\xi}_{SVD}$. A equação a seguir sintetiza os resultados:

$$\begin{aligned} \mathbf{C}_{SVD} &= \mathbf{V}_S \\ \mathbf{\xi}_{SVD} &= \mathbf{\Sigma}_S^{-1} \mathbf{U}_S^t \mathbf{\xi} \end{aligned} \quad (3.62)$$

onde \mathbf{U}_S , $\mathbf{\Sigma}_S$ e \mathbf{V}_S são as matrizes \mathbf{U} , $\mathbf{\Sigma}$ e \mathbf{V} da decomposição em valores singulares da matriz de restrição \mathbf{C} , após sofrerem uma redução de posto de P para S .

Através das Equações (3.56) e (3.62), percebe-se que a redução do número de valores singulares S está diretamente ligada a uma redução no número de restrições. A redução do número de valores singulares reduz a matriz \mathbf{V} de JK colunas para a matriz \mathbf{V}_S de S colunas. Sendo \mathbf{C}_{SVD} igual a \mathbf{V}_S , e o seu número de colunas o número de restrições empregadas, se S for menor que $2G$, o número de restrições empregadas foi reduzido em $2G - S$ restrições.

A escolha do número de valores singulares S está diretamente ligada à qualidade da representação da matriz de restrições. Quanto mais próximo o valor S for do valor P , mais fiel é a representação da matriz C . Por outro lado, a redução do número de valores singulares implica no aumento do número de graus de liberdade disponíveis ao processo de minimização. Como os últimos valores singulares são os menos significativos, é possível reduzir o número de restrições sem perder muito na qualidade da representação da matriz de restrições. De fato, para um dado número de restrições, as restrições autovetoriais minimizam o erro quadrático entre a resposta desejada e a resposta obtida numa faixa de interesse [32].

Na prática, a matriz C utilizada é formada a partir de uma superamostragem da região de interesse, sem qualquer preocupação com relação ao seu condicionamento. Em geral, a matriz de restrições utilizada é mal condicionada, com posto JK , e o sistema de equações em (3.47) é sobredeterminado [32]. Com a utilização de apenas os valores singulares mais significativos, o mau condicionamento da matriz C é perdido na sua transformação para a matriz C_{SVD} , e a resposta obtida aproxima-se da resposta desejada.

3.4 O CANCELADOR DE LÓBULO LATERAL GENERALIZADO

O cancelador de lóbulo lateral generalizado (*Generalized Sidelobe Canceller - GSC*), proposto por Griffiths e Jim em 1982 [38], é uma forma alternativa de representar o conformador de feixe LCMV. Como vantagem, há a transformação do problema de otimização restrita em um problema de otimização clássico, permitindo o uso de algoritmos de adaptação clássicos, como o LMS. Além disso, a estrutura torna o processamento realizado pelo conformador de feixes mais simples de ser compreendido. Uma referência ao assunto pode ser encontrada em [32].

Tratando o método de forma simplificada, a estrutura direta vista na seção anterior é dividida em dois caminhos (Figura 3.13). O caminho superior, chamado de *conformador convencional*, é destinado à sinais de interesse; e o caminho inferior, chamado de *caminho cancelador de lóbulo lateral*, é destinado às interferências.

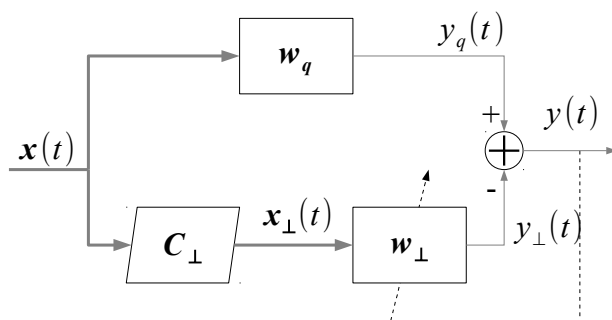


Figura 3.13: O conformador de feixe na forma do GSC

O vetor w_q , chamado de *vetor quiescente*, tem a função de implementar as restrições do problema. Por não ser adaptativo, ele independe dos sinais de entrada, de forma que o sinal de saída $y_q(t)$ é a aplicação da resposta desejada aos sinais dentro da faixa de interesse, sem nenhum cancelamento de interferências.

A matriz C_{\perp} , de tamanho $(JK) \times (JK - S)$ para as restrições autovetoriais, é o complemento ortogonal da matriz de restrições C_{SVD} no espaço R^{JK} , sendo obtida a partir da matriz de restrições com a utilização de qualquer método de ortogonalização disponível. A sua função é submeter aos filtros w_{\perp} somente os sinais de interferência, rejeitando o sinal de interesse. Por isso, a matriz C_{\perp} é chamada de *matriz de bloqueio de sinal* [32].

O sinal de saída do conformador $y(t)$ é igual ao sinal de saída do caminho superior $y_q(t)$ menos o sinal de saída do caminho inferior $y_{\perp}(t)$. Os filtros w_{\perp} implementam a parte adaptativa da estrutura,

tendo como função a ponderação das interferências para que, através da subtração dos sinais de saída dos caminhos superior e inferior, os sinais interferentes sejam cancelados.

Sabendo que a matriz C_{\perp} é o complemento ortogonal da matriz C_{SVD} no espaço $R^{J \times K}$, as colunas de C_{SVD} e C_{\perp} definem juntas o espaço $R^{J \times K}$. Desta forma, um vetor arbitrário w pode ser expresso em termos de suas componentes nos subespaços definidos pelas colunas de C_{SVD} e C_{\perp} da seguinte forma

$$w = C_{\text{SVD}} w_C - C_{\perp} w_{\perp} \quad (3.63)$$

onde w_C e w_{\perp} são componentes de w no subespaço das restrições e no subespaço ortogonal às restrições, respectivamente. Ao aplicar as restrições da equação (3.47), tem-se:

$$\begin{aligned} C_{\text{SVD}}^t w &= \xi_s \\ C_{\text{SVD}}^t [C_{\text{SVD}} w_C - C_{\perp} w_{\perp}] &= \xi_s \\ C_{\text{SVD}}^t C_{\text{SVD}} w_C - C_{\text{SVD}}^t C_{\perp} w_{\perp} &= \xi_s \\ C_{\text{SVD}}^t C_{\text{SVD}} w_C &= \xi_s \end{aligned} \quad (3.64)$$

onde o cancelamento de $C_{\text{SVD}}^t C_{\perp} w_{\perp}$ se deve ao fato de, por terem colunas ortogonais, $C_{\text{SVD}}^t C_{\perp} = 0$. Resolvendo:

$$w_C = (C_{\text{SVD}}^t C_{\text{SVD}})^{-1} \xi_s \quad (3.65)$$

e por fim, para que a Equação (3.63) fique de acordo com a Figura 3.13:

$$\begin{aligned} w_q &= C_{\text{SVD}} w_C \\ &= C_{\text{SVD}} (C_{\text{SVD}}^t C_{\text{SVD}})^{-1} \xi_s \end{aligned} \quad (3.66)$$

de forma que a Equação (3.63) torna-se:

$$w = w_q - C_{\perp} w_{\perp} \quad (3.67)$$

O sinal de saída do GSC é

$$\begin{aligned}
 y(t) &= \mathbf{w}^t \mathbf{x}(t) \\
 &= (\mathbf{w}_q - \mathbf{C}_\perp \mathbf{w}_\perp)^t \mathbf{x}(t) \\
 &= \mathbf{w}_q^t \mathbf{x}(t) - \mathbf{w}_\perp^t \mathbf{C}_\perp^t \mathbf{x}(t) \\
 &= y_q(t) - y_\perp(t)
 \end{aligned} \tag{3.68}$$

O problema de otimização da parte adaptativa do GSC implementa o LCMV através de uma minimização sem restrições. Como a matriz \mathbf{C}_\perp bloqueia o sinal de interesse, minimizar a potência de saída, constituída pela diferença dos dois caminhos, significa nesta estrutura minimizar somente a potência dos sinais interferentes. A potência do sinal de saída $y(t)$ do conformador é

$$\begin{aligned}
 P_y &= E[y^2(t)] \\
 &= E[(\mathbf{w}_q - \mathbf{C}_\perp \mathbf{w}_\perp)^t \mathbf{x}(t) \mathbf{x}^t(t) (\mathbf{w}_q - \mathbf{C}_\perp \mathbf{w}_\perp)] \\
 &= E[(\mathbf{w}_q - \mathbf{C}_\perp \mathbf{w}_\perp)^t \mathbf{R}_{xx} (\mathbf{w}_q - \mathbf{C}_\perp \mathbf{w}_\perp)]
 \end{aligned} \tag{3.69}$$

e o problema de otimização fica então

$$\underset{\mathbf{w}_\perp}{\text{minimizar}} \quad (\mathbf{w}_q - \mathbf{C}_\perp \mathbf{w}_\perp)^t \mathbf{R}_{xx} (\mathbf{w}_q - \mathbf{C}_\perp \mathbf{w}_\perp) \tag{3.70}$$

Como visto na Equação (3.70), o problema de otimização para a estrutura GSC é do tipo não restrita, e por consequência, algoritmos clássicos de otimização não restrita podem ser utilizados. O Quadro 3.1 ilustra a utilização do algoritmo LMS na estrutura GSC.

- Inicialização:
 - Se algum conhecimento *a priori* do vetor de coeficientes estiver disponível, usá-lo para selecionar um valor apropriado para $\mathbf{w}_\perp(0)$. Caso contrário, usar $\mathbf{w}_\perp(0)=\mathbf{0}$.
- Dado de entrada no instante t :
 - $\mathbf{x}(t)=\begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}(t-1) \\ \vdots \\ \mathbf{x}(t-J+1) \end{bmatrix}$, com $\mathbf{x}(t-j)=\begin{bmatrix} x_0(t-j) \\ x_1(t-j) \\ \vdots \\ x_{K-1}(t-j) \end{bmatrix}$
- Sinal de saída no instante t :
 - Caminho superior:
 - $y_q(t)=\mathbf{w}_q^t \mathbf{x}(t)$
 - Caminho inferior:
 - $\mathbf{x}_b(t)=\mathbf{C}_\perp^t \mathbf{x}(t)$
 - $y_\perp(t)=\mathbf{w}_\perp^t \mathbf{x}_b(t)$
 - Sinal de saída:
 - $y(t)=y_q(t)-y_\perp(t)$
- Adaptação dos coeficientes
 - $\mathbf{w}(t+1)=\mathbf{w}(t)+\mu \mathbf{x}_b(t) y(t)$

Quadro 3.1: Aplicação do algoritmo LMS na estrutura GSC

3.5 A FILTRAGEM MULTI-SPLIT ADAPTATIVA

Embora o LMS ainda seja um dos algoritmos mais utilizados em processamento de sinais, seu desempenho em termos de taxa de convergência e capacidade de rastreamento é dependente do espalhamento dos autovalores da matriz de autocorrelação do sinal de entrada. Isto limita seu espectro de aplicação quando se trata de processos não estacionários.

A filtragem Split adaptativa surgiu a partir das ideias propostas por Delsarte e Genin [39], que propuseram uma versão Split (segmentada) do algoritmo de Levinson para matrizes Toeplitz reais. Esta solução foi estendida à predição linear por Ho e Ching [40], que propuseram propondo um filtro adaptativo Split LMS. A filtragem Multi-Split LMS foi proposta por Ching e Wan [41], usando sucessivas transformações Split.

Resultados mostram que, sem aumento de complexidade, a filtragem Multi-Split LMS aplicada à estrutura GSC é uma opção mais eficiente do que o esquema GSC com o algoritmo LMS clássico, tanto em conformadores de feixe de banda estreita [42] quanto em conformadores de banda larga [31].

A justificativa para o bom desempenho do algoritmo foi apresentada em [43], devendo-se ao fato da operação de Multi-Split aumentar o fator de diagonalização da matriz de autocorrelação dos dados de entrada. Uma análise estatística do Multi-Split LMS pode ser encontrada em [44] e em [31].

Em [45], uma nova abordagem da filtragem Split é realizada, tendo como base a solução de um problema de otimização com restrições lineares aplicado à estrutura GSC. As restrições lineares aplicam as condições de simetria e antissimetria da filtragem Split na resposta ao impulso de dois filtros FIR conectados em paralelo. Esta abordagem será utilizada nesta seção como introdução à filtragem Multi-Split.

3.5.1 A filtragem Split

Qualquer sequência finita pode ser expressa pela soma de uma sequência conjugada simétrica e uma conjugada antissimétrica. A sequência finita pode ser, por exemplo, a resposta ao impulso de um filtro transversal. Em notação vetorial:

$$\mathbf{w} = \mathbf{w}_s + \mathbf{w}_a \quad (3.71)$$

sendo \mathbf{w} a resposta ao impulso de um filtro transversal com N_m coeficientes, \mathbf{w}_s a componente conjugada simétrica e \mathbf{w}_a a componente conjugada antissimétrica, respectivamente:

$$\mathbf{w} = [w_0 \quad w_1 \quad \cdots \quad w_{N_m-1}] \quad (3.72)$$

$$\mathbf{w}_s = \frac{1}{2} [\mathbf{w} + \mathbf{J} \mathbf{w}^*] \quad (3.73)$$

$$(3.74)$$

onde \mathbf{J} é chamada de *matriz de reflexão* ou *matriz de troca*, definida como uma matriz cujos elementos da diagonal secundária são unitários, sendo os outros elementos da matriz iguais a zero:

$$\mathbf{J} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \quad (3.75)$$

A Figura 3.14 mostra a divisão da resposta ao impulso de um filtro transversal nas componentes simétricas e antissimétricas.

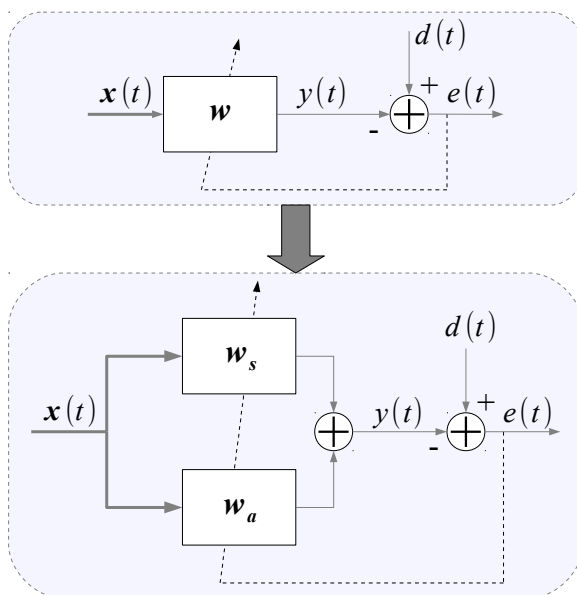


Figura 3.14: Filtragem adaptativa Split

Na abordagem apresentada em [45], a condição de simetria e antissimetria de w_s e w_a pode ser visualizada através de um conjunto de restrições lineares, definidas por:

$$C^t w(t) = f \quad (3.76)$$

onde C é a matriz de restrições, de dimensões $(N_m \times U)$, e f é um vetor de U elementos. As restrições utilizadas são:

$$C_s = \begin{bmatrix} I_U \\ \mathbf{0}_U^t \\ -J_U \end{bmatrix}, \quad C_a = \begin{bmatrix} I_U & \mathbf{0}_U \\ \mathbf{0}_U^t & \sqrt{2} \\ J_U & \mathbf{0}_U \end{bmatrix}, \quad f_s = \mathbf{0}_U \text{ e } f_a = \mathbf{0}_{U+1} \quad (3.77)$$

sendo:

$$U = \frac{N_m - 1}{2} \quad (3.78)$$

para N_m ímpar; ou

$$C_s = \begin{bmatrix} I_U \\ -J_U \end{bmatrix}, \quad C_a = \begin{bmatrix} I_U \\ J_U \end{bmatrix}, \quad f_s = f_a = \mathbf{0}_U \quad (3.79)$$

sendo:

$$U = \frac{N_m}{2} \quad (3.80)$$

para N_m par. As matrizes I_U , J_U e $\mathbf{0}_U$ usadas na definição das restrições são, respectivamente, a matriz identidade de ordem U , a matriz de reflexão da Equação (3.75) com ordem U e um vetor coluna com U elementos de valor zero. Esta dissertação assumirá deste ponto em diante que N_m é escolhido de forma a sempre ser um número par.

Baseado nas Equações (3.76), (3.77) e (3.79), as restrições impostas para estabelecer as condições de simetria e antissimetria de w_s e w_a são

$$C_s^t w_s(t) = f_s \quad (3.81)$$

e

$$C_a^t w_a(t) = f_a \quad (3.82)$$

Como visto na Seção 3.4, a estrutura GSC pode ser utilizada para transformar um problema de otimização restrita em um problema de otimização irrestrita. Este conceito é usado também na filtragem Split, substituindo os filtros w_s e w_a pelos seus equivalentes GSC. A Figura 3.15 ilustra a filtragem Split descrita através das estruturas GSC.

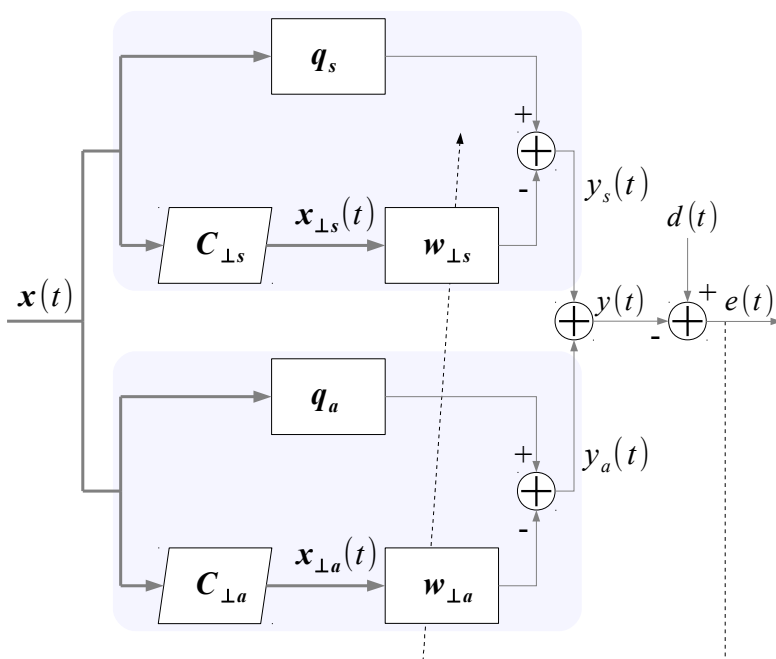


Figura 3.15: Implementação da filtragem Split via GSC

A estrutura Split GSC pode ser simplificada tendo em vista que os filtros q_s e q_a , definidos pela Equação (3.66) se anulam:

$$q_s = C_s (C_s^t C_s)^{-1} f_s = C_s (C_s^t C_s)^{-1} \mathbf{0} = \mathbf{0} \quad (3.83)$$

$$(3.84)$$

Outra simplificação a ser feita é que um bom candidato à matriz de bloqueio da parte simétrica $C_{\perp s}$ é a própria matriz de restrição da parte antissimétrica C_a , acontecendo o mesmo para a parte antissimétrica. A simplificação da estrutura de filtragem Split baseada na estrutura GSC resultante pode ser vista na Figura 3.16.

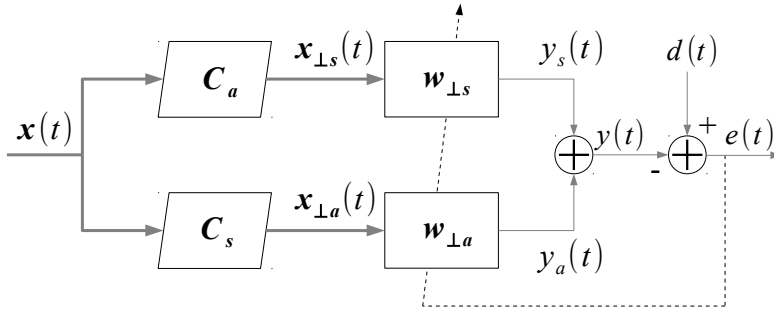


Figura 3.16: Simplificação da implementação da filtragem Split via GSC

3.5.2 A filtragem Multi-Split

Seguindo a mesma abordagem baseada na estrutura GSC, a estrutura Multi-Split é apresentada em [46]. Nela, N_m filtros de ordem zero (com apenas um coeficiente) são utilizados, obtidos através de sucessivas decomposições em componentes simétrica e antissimétrica dos filtros $w_{\perp s}$ e $w_{\perp a}$.

A Figura 3.17 ilustra o filtro Multi-Split obtido a partir de um filtro transversal com $N_m = 2^M$ coeficientes, sendo M um inteiro maior que um. A estrutura foi obtida após M passos com 2^{m-1} operações de segmentação ($m = 1, 2, \dots, M$) em cada passo.

O sinal de saída do filtro é obtido através da equação

$$y(t) = \sum_{i=0}^{N_m-1} x_{\perp i}(t) w_{\perp i}(t-1) \quad (3.85)$$

e, por fim, o sinal de erro de estimação é

$$e(t) = d(t) - y(t) \quad (3.86)$$

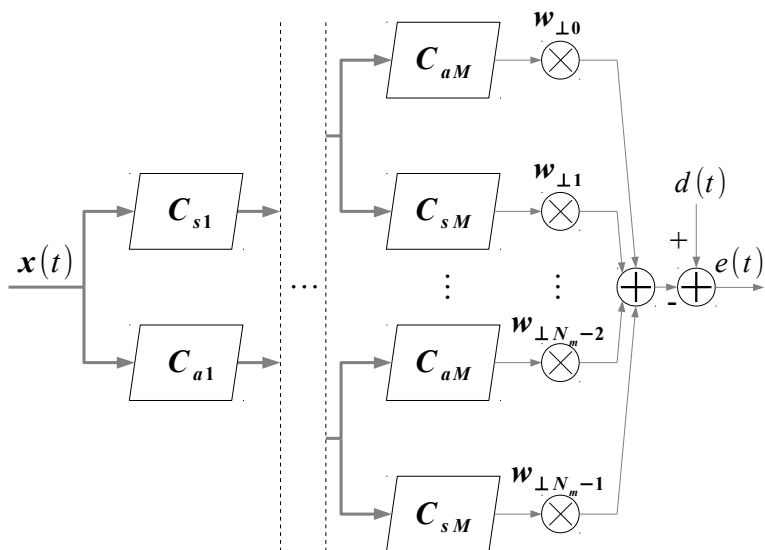


Figura 3.17: A estrutura Multi-Split adaptativa

O esquema Multi-Split pode ser visto como uma transformação linear de $x(t)$, denotada por:

$$\mathbf{x}_{\perp}(t) = \mathbf{T}^t \mathbf{x}(t) \quad (3.87)$$

onde

$$T = \begin{bmatrix} C_{aM}^t C_{a(M-1)}^t \cdots C_{a1}^t \\ C_{sM}^t C_{a(M-1)}^t \cdots C_{a1}^t \\ C_{aM}^t C_{s(M-1)}^t \cdots C_{a1}^t \\ C_{sM}^t C_{s(M-1)}^t \cdots C_{a1}^t \\ \vdots \\ C_{sM}^t C_{s(M-1)}^t \cdots C_{s1}^t \end{bmatrix}^t \quad (3.88)$$

Por último, permutações entre as colunas da matriz T equivalem a rearranjar os coeficientes $w_{\perp n}$ em diferentes sequências. Dentre as $N_m!$ permutações possíveis, uma delas transforma a matriz T numa matriz de Hadamard H_{N_m} de ordem N_m , de forma que o esquema Multi-Split passa a ser representado na forma compacta (Figura 3.18). A matriz de Hadamard de ordem N_m pode ser construída a partir de uma matriz de Hadamard de ordem $N_m \div 2$ com:

$$H_{N_m} = \begin{bmatrix} H_{N_m \div 2} & H_{N_m \div 2} \\ H_{N_m \div 2} & -H_{N_m \div 2} \end{bmatrix} \quad (3.89)$$

sendo que $H_1 = [1]$ por definição.

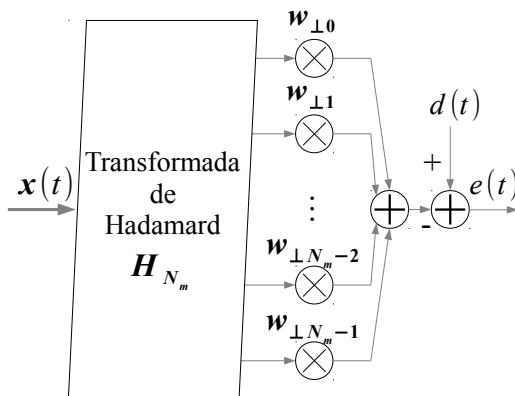


Figura 3.18: O Multi-Split na forma compacta, representado pela transformação de Hadamard

3.5.3 O Multi-Split LMS

Além da estrutura de pré-processamento Multi-Split apresentada até o presente, [46] propõe ainda um algoritmo de passo variável normalizado pelas estimativas da potência dos sinais de entrada dos filtros. O algoritmo é baseado nas técnicas de filtragem adaptativa de auto ortogonalização, conhecidas por melhorar a convergência do algoritmo LMS [29].

A cada iteração, a potência dos sinais de entrada é estimada, podendo um fator de esquecimento ser utilizado para atribuir um maior peso às amostras mais recentes do sinal. Isto se dá através da equação:

$$\tilde{\lambda}_i(t) = \frac{\sum_{m=1}^t \gamma^{t-m} x_{\perp i}^2(m)}{\sum_{m=1}^t \gamma^{t-m}} \quad (3.90)$$

onde $\tilde{\lambda}_i(t)$ é a estimativa da potência do sinal $x_{\perp_i}(t)$, com $i=0,1,2,\dots,N_m-1$, e γ é o fator de esquecimento utilizado, definido no intervalo $0 \leq \gamma \leq 1$. O uso do fator de esquecimento é útil quando os sinais envolvidos não são estacionários, caso dos sinais de voz, objeto desta dissertação. Usar $\gamma=1$ é aplicável a um ambiente estacionário no sentido amplo. A partir da equação anterior, uma forma recursiva para a estimação da potência pode ser encontrada:

$$\tilde{\lambda}_i(t) = \frac{x_{\perp_i}^2(t) + \tilde{\lambda}_i(t-1)[p(t)-1]}{p(t)} \quad (3.91)$$

onde:

$$p(t) = \gamma p(t-1) + 1 \quad (3.92)$$

$$p(0) = \gamma \quad (3.93)$$

A adaptação dos coeficientes é realizada através da equação:

$$w_{\perp_i}(t) = w_{\perp_i}(t-1) + \frac{\mu}{\tilde{\lambda}_i(t)} x_{\perp_i}(t) e(t) \quad (3.94)$$

onde μ é o passo constante de adaptação do algoritmo, que para garantir a convergência do algoritmo, deve se situar no intervalo [43]:

$$1 < \mu < \frac{1}{N_m \lambda_{\max}} \quad (3.95)$$

sendo λ_{\max} o maior autovalor da matriz de autocorrelação do sinal $x_{\perp_i}(t)$.

O algoritmo, chamado de Multi-Split LMS (MS-LMS), está sumariado no Quadro 3.2.

- Inicialização:
 - Para $i=0,1,\dots,N_m-1$, $w_{\perp_i}(0)=0$, $p(0)=\gamma$ e $\tilde{\lambda}_i(0)=0$.
- Dado de entrada no instante t :
 - Para $i=0,1,\dots,N_m-1$, $x_{\perp_i}(t)=\mathbf{t}_i^t \mathbf{x}(n)$ onde \mathbf{t}_i é a coluna $i+1$ da matriz de transformação \mathbf{T} e

$$\mathbf{x}(t)=[x(t) \quad x(t-1) \quad \dots \quad x(t-N_m+1)]^t$$
- Estimação da potência do sinal de entrada no instante t :
 - $p(t)=\gamma p(t-1)+1$
 - Para $i=0,1,\dots,N_m-1$, $\tilde{\lambda}_i(t)=\frac{x_{\perp_i}^2(t)+\tilde{\lambda}_i(t-1)[p(t)-1]}{p(t)}$
- Sinal de saída no instante t :
 - $y(t)=\sum_{i=0}^{N_m-1} x_{\perp_i}(t)w_{\perp_i}(t-1)$
- Sinal de erro no instante t :
 - $e(t)=d(t)-y(t)$
- Adaptação dos coeficientes
 - Para $i=0,1,\dots,N_m-1$, $w_{\perp_i}(t)=w_{\perp_i}(t-1)+\frac{\mu}{\tilde{\lambda}_i(t)}x_{\perp_i}(t)e(t)$

Quadro 3.2: O algoritmo Multi-Split LMS

Uma comparação em termos de taxa de convergência entre o algoritmo Multi-Split LMS (MS-LMS) e o LMS pode ser vista na Figura 3.19, retirada de [43], onde também pode ser vista a convergência do algoritmo RLS (*Recursive Least-Squares*). A figura foi obtida a partir da aplicação dos algoritmos em um equalizador

adaptativo, com oito coeficientes e controle sobre o espalhamento dos autovalores da matriz de autocorrelação do sinal de entrada. A curva superior mostra a convergência do MSE (*Mean Square Error*) para um baixo espalhamento dos autovalores; na curva inferior, um alto espalhamento dos autovalores foi utilizado.

É possível perceber na figura que, em se tratando de taxa de convergência, o Multi-Split LMS possui um melhor desempenho que o LMS clássico. Este é um bom resultado, tendo em vista que o Multi-Split apresenta uma complexidade computacional semelhante ao LMS clássico [31]. Além disso, o LMS clássico apresenta uma maior dependência do espalhamento dos autovalores, ficando mais lento à medida que os sinais são mais correlacionados [29].

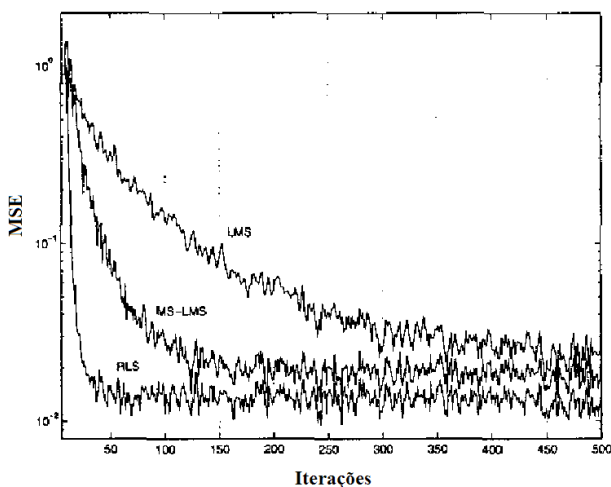
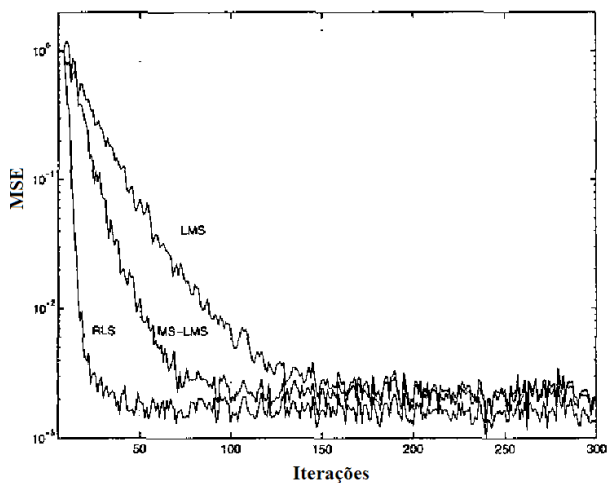


Figura 3.19: Comparação do MSE ao longo das iterações dos algoritmos RLS, MS-LMS, e LMS.

Adaptado de [43].

3.5.4 O conformador de feixe Multi-Split GSC

A Figura 3.20 ilustra a aplicação do Multi-Split no conformador de feixe adaptativo na forma GSC, denominado conformador MS-GSC [31]. Após o bloqueio do sinal desejado pela matriz de bloqueio \mathbf{C}_\perp , o pré-processamento Multi-Split com a matriz \mathbf{H}_{N_m} é realizado, e então o conjunto de filtros de ordem zero da estrutura é aplicado.

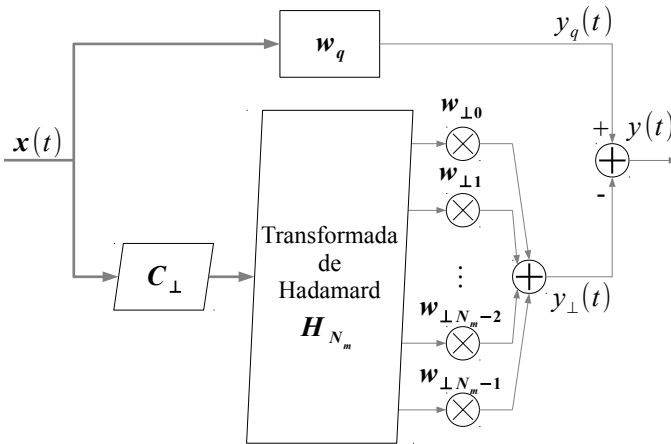


Figura 3.20: Aplicação do Multi-Split LMS no conformador de feixes adaptativo na forma do GSC

A fim de reduzir a carga computacional da estrutura, a matriz de bloqueio da estrutura GSC e a matriz quadrada de transformação Multi-Split podem ser agrupadas, de forma que o bloco resultante bloqueie o sinal desejado e realize a transformação Multi-Split. A Figura 3.21 apresenta a forma final do conformador MS-GSC, onde a matriz $\mathbf{C}_{\perp M}$ é obtida através da multiplicação dos dois blocos processadores:

$$\mathbf{C}_{\perp M} = \mathbf{C}_\perp \mathbf{T} \quad (3.96)$$

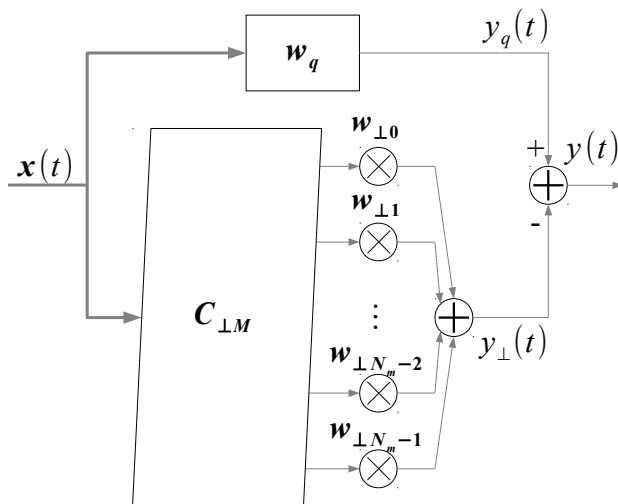


Figura 3.21: Conformador MS-GSC com a matriz de transformação Multi-Split incorporada à matriz de bloqueio

Cabe ressaltar que a Equação (3.96) é válida para qualquer matriz C_{\perp} , desde que a matriz T seja construída com matrizes que respeitem as Equações (3.77) e (3.79), no que tange ao número de linhas de entrada da estrutura Multi-Split. Assim, $N_m = JK - 2G$ para um problema com G restrições pontuais.

Da mesma forma, quando , com M sendo um inteiro maior que 0, a matriz T pode ser utilizada na forma da matriz de Hadamard H_N . O mesmo ocorre quando S restrições autovetoriais são utilizadas, caso em que o número de linhas de entrada da matriz de transformação Multi-Split é $JK - S$. Neste caso, o uso do Multi-Split na forma da matriz de Hadamard limita a escolha do número de valores singulares mais significativos S utilizados, de forma que $JK - S = 2^M$.

Por fim, o algoritmo Multi-Split LMS aplicado ao conformador de feixe GSC com S restrições autovetoriais está sumariado no Quadro 3.3.

- Inicialização:
 - Faça $N_m = JK - S$
 - Para $i=0, 1, \dots, N_m - 1$, $w_{\perp i}(0)=0$, $p(0)=\gamma$ e $\tilde{\lambda}_i(0)=0$.
- Dado de entrada no instante t :
 - $\mathbf{x}(t) = \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}(t-1) \\ \vdots \\ \mathbf{x}(t-J+1) \end{bmatrix}$, com $\mathbf{x}(t-j) = \begin{bmatrix} x_0(t-j) \\ x_1(t-j) \\ \vdots \\ x_{K-1}(t-j) \end{bmatrix}$
- Sinal de saída no instante t :
 - Caminho superior:
 - $y_q(t) = \mathbf{w}_q^t \mathbf{x}(t)$
 - Caminho inferior:
 - Para $i=0, 1, \dots, N_m - 1$, faça $x_{\perp i}(t) = \mathbf{c}_{\perp i}^t \mathbf{x}(t)$, onde $\mathbf{c}_{\perp i}$ é a coluna $i+1$ da matriz $\mathbf{C}_{\perp M}$
 - $y_{\perp}(t) = \sum_{i=0}^{N_m-1} x_{\perp i}(t) w_{\perp i}(t-1)$
 - Sinal de saída:
 - $y(t) = y_q(t) - y_{\perp}(t)$
- Estimação da potência:
 - $p(t) = \gamma p(t-1) + 1$
 - Para $i=0, 1, \dots, N_m - 1$, faça $\tilde{\lambda}_i(t) = \frac{x_{\perp i}^2(t) + \tilde{\lambda}_i(t-1)[p(t)-1]}{p(t)}$
- Adaptação dos coeficientes
 - Para $i=0, 1, \dots, N_m - 1$, faça $w_{\perp i}(t) = w_{\perp i}(t-1) + \frac{\mu}{\tilde{\lambda}_i(t)} x_{\perp i}(t) e(t)$

Quadro 3.3: O multisplit LMS aplicado ao conformador de feixe GSC

4 SIMULAÇÕES E RESULTADOS

A utilização de arranjo de microfones em aparelhos auditivos foi avaliada sob a forma de uma simulação. A simulação desenvolvida implementa os conceitos dos capítulos anteriores, e foi construída de forma a simplificar a sua utilização, para que novos pesquisadores possam utilizá-la como ferramenta de estudos.

O objetivo geral dos testes realizados é comparar o desempenho de duas estruturas diferentes de arranjo de microfones. A primeira estrutura, chamada por conveniência de “*Frost*”, utiliza um conformador de feixes de banda larga na forma direta, com o algoritmo CLMS aplicado a um problema de otimização com as restrições de Frost. A segunda estrutura, denominada “*Multi-Split*”, utiliza um conformador de feixes de banda larga na forma GSC, com a transformação Multi-Split acoplada, e o algoritmo MS-LMS aplicado a um problema de otimização com as restrições autovetoriais. A Tabela 2 resume as duas estruturas.

Tabela 2: Estruturas utilizadas na comparação de desempenho

	Estrutura	
	<i>Frost</i>	<i>Multi-Split</i>
Forma do conformador	Direta	Multi-Split GSC
Algoritmo	CLMS	MS-LMS
Restrições	Frost	Autovetoriais

4.1 PARÂMETROS DOS TESTES

Uma dificuldade presente neste trabalho diz respeito ao número de parâmetros presentes na simulação. O tamanho da sala reverberante,

os coeficientes de absorção das paredes, o número e a posição das fontes sonoras e dos microfones do arranjo, o formato das restrições utilizadas, refletem diretamente nos resultados e precisam ser escolhidos com parcimônia.

Desta forma, os parâmetros da simulação foram limitados a um conjunto de valores, a fim de padronizar os testes realizados. Visando uma simplicidade de apresentação, os parâmetros dos testes relacionados ao cenário e ao processamento adaptativo serão apresentados separadamente nas próximas seções.

4.1.1 Parâmetros relacionados ao cenário dos testes

A Tabela 3 sintetiza os parâmetros relacionados ao cenário dos testes realizados. As seções a seguir descrevem em detalhes os parâmetros.

4.1.1.1 Número de fontes sonoras

Duas fontes sonoras estão presentes no ambiente, sendo a referente ao sinal de interesse, localizada imediatamente à frente do arranjo de microfones (*broadside*), e a fonte de interferência, localizada a quarenta e cinco graus do foco do arranjo.

4.1.1.2 Resposta do canal

A resposta do canal é calculada seguindo os métodos do Capítulo 2, a partir dos parâmetros da sala. As dimensões da sala, os coeficientes de reflexão, o posicionamento das fontes sonoras e dos microfones estão relacionados à resposta do canal.

Tabela 3: Parâmetros relacionados ao cenário dos testes realizados

Parâmetros	Características
Fontes sonoras	Duas fontes sonoras, uma de interesse (0°) e outra de interferência (45°)
Resposta do canal	Ambiente com pouca reverberação ($T_{60}=150$ ms)
	Ambiente com muita reverberação ($T_{60}=600$ ms)
Tipo de microfone	Omnidirecional
	Cardioide
Configuração do arranjo de microfones	Broadside
Potência do ruído gaussiano aditivo	10 dB abaixo da potência média dos sinais de voz

Os parâmetros da resposta do canal foram escolhidos tendo como base o trabalho de Greenberg em [10]. Duas salas foram utilizadas, com diferentes tempos de reverberação T_{60} obtidos através dos coeficientes de absorção das superfícies, visando obter um cenário pouco reverberante ($T_{60}=150$ ms) e outro muito reverberante ($T_{60}=600$ ms). A posição das fontes sonoras e microfones é a mesma para ambas as salas. As fontes sonoras foram distribuídas de forma a respeitar os ângulos de incidência previstos. Os valores dos parâmetros utilizados na geração das respostas das salas podem ser vistos na Tabela 4.

Tabela 4: Parâmetros das duas salas utilizadas nos testes

Parâmetros	Sala 1	Sala 2
Classificação	Pouco reverberante	Muito reverberante
Tempo de reverberação T_{60}	150 ms	600 ms
Coefficiente de absorção das superfícies	0,6	0,2
Duração da resposta ao impulso calculada	0,25 segundos	0,8 segundos

Parâmetros comuns a ambas as salas

Tamanho da sala	$5,2 \times 3,4 \times 2,8$ metros
Localização do primeiro microfone do arranjo (microfone de referência)	$2,755 \times 1,38 \times 1,6$
Espaçamento entre os microfones	1 centímetro
Localização da fonte sonora 1 (sinal de interesse)	$2,755 \times 2,38 \times 1,6$
Localização da fonte sonora 2 (interferência)	$2,048 \times 2,087 \times 1,6$

Uma visão das dimensões da sala, bem como da localização das fontes sonoras e dos microfones pode ser obtida na Figura 4.1. As energias das respostas das salas utilizadas podem ser vistas na Figura 4.2, podendo o tempo de reverberação ser estimado a partir delas. Como já mencionado na Seção 2.2.3, as salas criadas possuem apenas paredes, teto e piso, sem aberturas, móveis ou outros objetos.

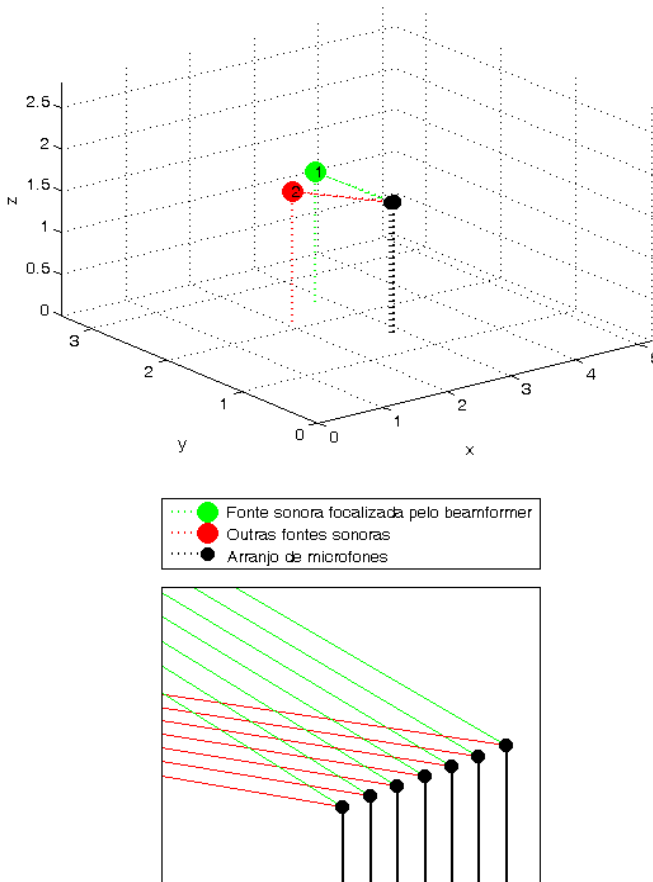


Figura 4.1: Dimensões da sala e posicionamento do arranjo de microfones e das fontes sonoras. No recorte, detalhe para os ângulos de incidência dos sinais de interesse e de interferência.

4.1.1.3 Tipos de microfones

Dentre os vários tipos de microfones de primeira ordem disponíveis, apresentados no Capítulo 2, os testes foram limitados a

utilizar microfones omnidirecionais ou cardioides, por serem mais comumente utilizados em aparelhos auditivos [47].

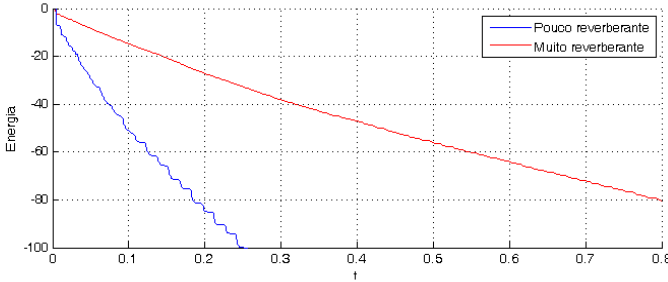


Figura 4.2: Energia das respostas das salas utilizadas nos testes

4.1.1.4 Configuração do arranjo de microfones

Como indicado no Item 3.1.1 da Seção 3.1, este trabalho assume uma distribuição linear dos microfones a espaços constantes de um centímetro, distância que respeita o teorema da amostragem espacial da Equação (3.11). Além disso, os microfones são distribuídos na linha perpendicular à linha de propagação da onda sonora de foco (*broadside*). Esta escolha se baseou no fato de a estrutura *Frost* ter sido desenvolvida para esta configuração de arranjo.

4.1.1.5 Potência do ruído gaussiano aditivo

A captação dos sinais pelos microfones do arranjo é concluída com a adição de um ruído branco gaussiano de média zero, que tem o objetivo de modelar as incertezas existentes, como o ruído térmico dos componentes eletrônicos. A relação sinal-ruído escolhida, de 10dB, baseia-se no livro de Loizou [48], que indica que ambientes silenciosos como hospitais, lojas e salas de aula apresentam uma relação sinal-ruído entre 5 e 15 dB.

A potência do ruído é obtida tendo como referência o Nível de Fala Ativa (ASL - *Active Speech Level*) dos sinais de voz presentes no ambiente, prática recomendada pela ITU-T (International Telecommunication Union - Telecommunication Standardization Sector) em [49]. O ASL descarta os períodos de silêncio da voz na estimação do nível sonoro através da estimação do envelope do sinal de voz, que é obtido através de uma filtragem IIR (*Infinite Impulse Response*) de segunda ordem do valor absoluto do sinal de voz [50]. As amostras do envelope são comparadas a um conjunto de limiares discretos, e uma contagem dos respectivos limiares excedidos pelo envelope é realizada. Através desta contagem, os instantes de atividade da voz são definidos, e os níveis do sinal são estimados. Neste trabalho, o nível de fala é calculado em relação ao limite inferior dos sons audíveis, $20 \mu\text{Pa}$, resultando na medida dB SPL (*Sound Pressure Level*) [51].

4.1.2 Parâmetros relacionados à filtragem adaptativa

4.1.2.1 Restrições

O projeto das estruturas “*Frost*” e “*Multi-Split*” deve ser iniciado com a definição das restrições utilizadas em cada problema de otimização. Estas foram definidas de forma a priorizar sinais na faixa de voz, presentes até aproximadamente 4000 Hz, deixando ainda uma banda de guarda de 1000 Hz. Desta forma, o objetivo dos algoritmos é minimizar as interferências enquanto uma filtragem passa-baixas com banda de passagem até 5000 Hz é aplicada ao sinal de interesse.

Para a estrutura “*Frost*”, deve ser definida a resposta ao impulso do filtro ao qual será submetido o sinal de interesse (resposta desejada) para compor o vetor de resposta às restrições ξ (Equação 3.39). Para isso, um filtro do tipo *equiripple* foi utilizado, com banda de passagem em 5000 Hz e banda de rejeição a partir de 6000 Hz. A ordem do filtro é, por definição, $J-1$. A resposta do filtro projetado pode ser vista na Figura 4.3.

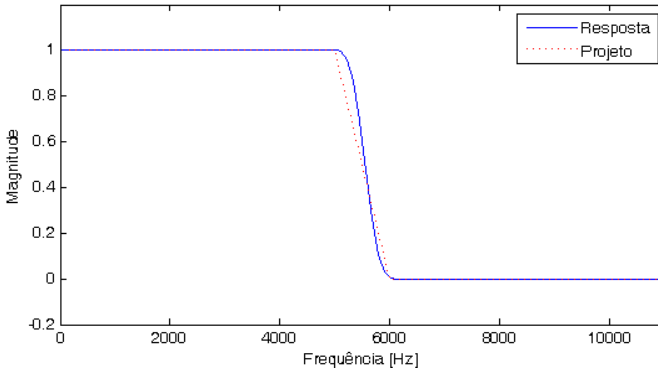


Figura 4.3: Resposta do filtro projetado para as restrições de Frost

No projeto das restrições para a estrutura “*Multi-Split*”, foram utilizados 50 pontos de frequência igualmente espaçados na faixa de frequência de 160 a 5000 Hz, para o ângulo de incidência de zero graus. Para esta faixa, uma resposta de magnitude unitária e fase linear foi adotado.

As restrições autovetoriais da estrutura “*Multi-Split*” são obtidas através da decomposição em valores singulares (SVD) da matriz de restrições pontuais criada. A redução de posto (Equação 3.56) é realizada tendo como base os valores singulares mais significativos S , selecionados manualmente a cada execução dos testes. A decomposição realizada é usada também na geração da matriz de bloqueio da estrutura GSC, através dos vetores singulares direitos $\mathbf{v}_{S+1}, \mathbf{v}_{S+2}, \dots, \mathbf{v}_{JK}$ (Equação 3.54).

4.1.2.2 Valores singulares das restrições autovetoriais

Como mencionado no Item 3.5.4 da Seção 3.5, o uso conjunto das restrições autovetoriais e da estrutura “*Multi-Split*”, quando esta é aplicada ao GSC através da matriz de Hadamard, limita a escolha dos valores singulares mais significativos. Foi também descrito que isto ocorre devido ao posto da matriz S ter de ser escolhido de forma que $JK - S = 2^M$, sendo M um inteiro maior que 0. Assim, dado que neste

trabalho optou-se pela aplicação da estrutura *Multi-Split* através da matriz de Hadamard, a escolha do número de valores singulares mais significativos foi feita a partir do conjunto de valores

$$S = \{JK - 2^M\} \quad (4.1)$$

com $M = 1, 2, \dots, \lfloor \log_2(JK) \rfloor$, sendo $\lfloor \cdot \rfloor$ o operador de arredondamento para o inteiro inferior (truncamento).

4.1.2.3 Número de microfones e coeficientes

A limitação dos valores singulares disponíveis inclui ainda uma outra dificuldade que se encontra relacionada ao número total de coeficientes do sistema JK . Dependendo da configuração de coeficientes utilizada, a escolha de S pode rejeitar valores singulares muito significativos, gerando uma matriz de restrições pouco similar à original. Também, há situações onde o S escolhido utiliza valores singulares em demasia, fazendo o número de restrições ser alto e, por consequência, reduzindo o número de graus de liberdade para os algoritmos adaptativos.

A Tabela 5 exhibe quatro configurações diferentes de coeficientes, em um cenário onde 500 restrições pontuais foram geradas no ângulo de zero graus e na faixa de frequência de 160 a 5000 Hz.

No cenário 1, a matriz de restrições para 7 sensores e 20 coeficientes possui originalmente posto 20. A partir da Equação (4.1), a escolha de S se limita aos valores 138, 136, 132, 124, 108, 76 e 12. A seleção de 12 valores descarta valores singulares importantes (por exemplo, 2,0124) e a matriz de restrições gerada não se aproxima da desejada. Fazendo $S=76$, valores singulares pouco significativos são utilizados.

No cenário 2, o menor valor de S possível de ser escolhido é o 44, já bastante próximo do posto da matriz de restrições, 48. Desta forma, todas as opções disponíveis para a escolha de S utilizam valores singulares desnecessários. O cenário 3 apresenta uma situação onde o

menor número de valores singulares disponível já é maior que o posto da matriz de restrições.

O cenário 4 apresenta uma situação considerada satisfatória. A escolha mais baixa para S , 22, representa uma redução aproximada de 27% no posto da matriz original. Desta forma, o valor singular menos significativo utilizado é 0,0039, sendo o valor do mais significativo descartado igual a 0,0006.

Tabela 5: Exemplos de cenários obtidos na escolha do número de valores singulares mais significativos

Cenário		1	2	3	4
Microfones (K)		7	5	8	5
Coeficientes (J)		20	60	10	30
Posto da matriz		20	48	10	30
Opção 1	Número de valores	12	44	16	22
	Valor singular	7,57E+000	8,84E-008	1,01E-014	3,92E-003
Opção 2	Número de valores	76	172	48	86
	Valor singular	1,51E-30	6,72E-44	4,34E-46	3,84E-30

Para a escolha do número de microfones e coeficientes utilizados, foram feitos testes de geração de restrições, em estruturas com um a dez microfones e dez a cem (a saltos de dez) coeficientes. As quatro configurações escolhidas podem ser vistas na Tabela 6. A denominação de cada teste, a saber: A, B, C e D, será utilizada no restante do trabalho. Embora a SVD seja utilizada apenas para as restrições autovetoriais aplicadas à estrutura “*Multi-Split*”, estas configurações foram também utilizados na estrutura “*Frost*”.

Tabela 6: Número de microfones e coeficientes utilizados nos testes

Configuração	Microfones	Coeficientes
A	2	90
B	4	40
C	5	30
D	7	80

O uso de um arranjo de microfones *broadside* com microfones distribuídos a distâncias de um centímetro, aliado ao número excessivo de microfones da configuração D da Tabela 6, inviabiliza a aplicação direta das técnicas deste trabalho em configurações convencionais de aparelhos auditivos. Alguns trabalhos endereçam este problema, sugerindo formas alternativas de distribuir os microfones do arranjo. Soede [52] indica a distribuição dos microfones em armações de óculos, possibilitando configurações tanto em *broadside* quanto em *endfire*. Widrow [53] sugere a utilização de um colar, onde os microfones do arranjo são distribuídos em forma de “V”. Mais informações podem ser obtidas em [54].

4.1.2.4 Parâmetros dos algoritmos de otimização

Além das configurações apresentadas, os algoritmos utilizados possuem ainda parâmetros específicos, que permitem controlar o seu funcionamento.

No algoritmo CLMS, o passo de adaptação μ controla a taxa de convergência. Após testes preliminares, o passo de adaptação foi fixado em 0,05.

No Multi-Split LMS, além do passo de adaptação, configurado com 0,0005, o fator de esquecimento da estimação da potência do sinal também é importante no comportamento do algoritmo. Para que a estimação da potência “lembre” as Q amostras mais recentes do sinal de fala, o fator de esquecimento γ foi escolhido a partir de:

$$\gamma = \frac{Q-1}{Q} \quad (4.2)$$

O valor de Q é escolhido tendo como base o intervalo de tempo em que sinais de fala podem ser considerados estacionários, dentro da faixa de 5 a 30 ms [55]. A Figura 4.4 apresenta a potência estimada de um sinal de voz utilizando três valores diferentes de γ , calculados para períodos de estacionariedade, T_e , de 10, 20 e 30 ms. Dada a pequena diferença entre os resultados, o sinal de voz foi considerado estacionário por 30 ms nas simulações.

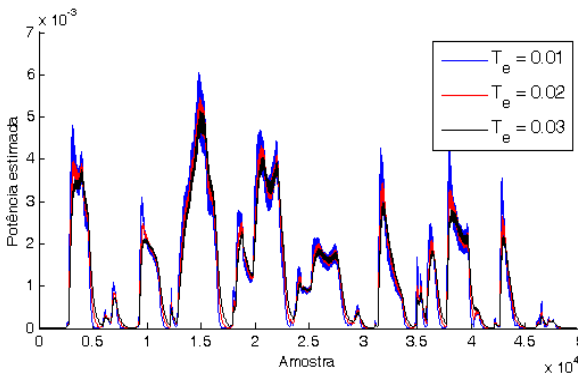


Figura 4.4: Potência estimada de um sinal de voz considerando diferentes períodos de estacionariedade

4.1.2.5 Considerações finais sobre a simulação

Por mais sofisticada que seja uma simulação, ela nunca poderá substituir um teste prático num cenário real. A prática envolve fatores que vão além do que qualquer modelo teórico considera. Por isso, ainda que o controle sobre os parâmetros e a rapidez de desenvolvimento atingidos pelo uso de simulações sejam imprescindíveis no desenvolvimento de qualquer projeto de engenharia, sua validação só poderá ser realizada por um teste prático.

Para a simulação desenvolvida neste trabalho, itens importantes e de influência conhecida foram desconsiderados.

Como analisado por Ivan Tashev [56], o desempenho de conformadores de feixe adaptativos é prejudicado devido a tolerâncias no processo de fabricação dos microfones utilizados no arranjo, principalmente no caso de variações da magnitude da resposta em frequência. Tashev [57] propõe um procedimento de calibração dos microfones do arranjo.

Além de variações nos microfones, o próprio arranjo pode apresentar imperfeições. Erros no alinhamento dos microfones do arranjo podem causar grande degradação do desempenho de conformadores [11].

Outro fator significativo desconsiderado neste trabalho é a influência da cabeça do usuário na resposta final do aparelho auditivo, fenômeno conhecido pelo termo em inglês *headshadow*. O efeito, causado por mecanismos de propagação de ondas como a difração, caracteriza-se por uma resposta variante em ângulo e frequência, como pode ser visto em [58].

Ainda que, como descrito anteriormente, seja indicado o uso de distribuições alternativas dos microfones, a influência do corpo do usuário na resposta deve ser levada em consideração para resultados mais realistas. Puder [59] endereça o problema e Hamacher *et al.* [60] apresenta diferenças na resposta do arranjo de microfones quando o *headshadow* é considerado, através de medições utilizando o manequim KEMAR¹. Hoffman [62] descreve a modelagem do efeito em simulações.

1 Knowles Electronics Manikin for Acoustic Research - Manequim amplamente utilizado em testes com sons. Para mais informações, ver [61]

4.2 ÁUDIOS UTILIZADOS

Para obter uma maior proximidade dos resultados de simulação com a prática, sinais de fala reais foram utilizados nos testes.

A base de sinais de fala utilizada, criada no trabalho de doutorado do Prof. Dr. Carlos Alberto Ynoguti [63], é composta pelas vinte listas de dez frases indicadas por Alcaim *et al.* [64]. As frases são foneticamente balanceadas, o que indica uma distribuição fonética similar à encontrada na língua espontânea. A utilização foi feita com permissão do Prof. Ynoguti.

Quarenta locutores adultos contribuíram com as gravações, sendo vinte homens e vinte mulheres. Cada locutor pronunciou quarenta frases, sendo cada frase repetida por oito locutores diferentes. As frases foram gravadas em arquivos de áudio separados, com frequência de amostragem de 11,025 kHz e resolução de dezesseis bits. Para este trabalho, a frequência de amostragem foi aumentada para 22,050 kHz via interpolação.

A potência dos sinais, estimada via Nível de Fala Ativa, situa-se na média em 70,31 dB SPL, tendo o mais fraco 55,59 dB SPL e o mais forte 79,71 dB SPL. O histograma da Figura 4.5 mostra como estão distribuídos os níveis dos sinais.

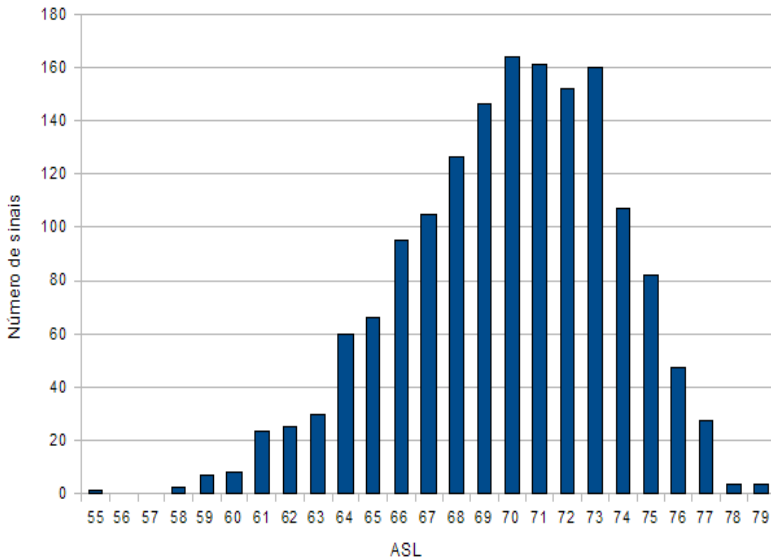


Figura 4.5: Distribuição dos Níveis de Fala Ativa da base de áudios utilizada

4.3 MEDIDAS DE DESEMPENHO

O desempenho de sistemas de processamento de sinais de fala é de difícil medição. A análise da qualidade da fala processada envolve fatores que são difíceis de serem ponderados, além de que o próprio conceito de qualidade varia entre aplicações e público-alvo [65].

Por conta da dificuldade de medição, avaliações subjetivas ainda são as mais indicadas na avaliação de desempenho. Estas são feitas submetendo ouvintes avaliadores a uma metodologia padronizada de testes, que deve ser extremamente organizada a fim de evitar erros de execução. Os procedimentos a serem adotados na realização de medidas subjetivas foram padronizados pela ITU (*International Telecommunication Union*) em [66], e um resumo sobre o assunto pode

ser encontrado em [65]. As práticas recomendadas pelo IEEE (*Institute of Electrical and Electronics Engineers*) são encontradas em [67].

Embora indicadas, as avaliações subjetivas são caras e trabalhosas de serem realizadas, seja pelo próprio procedimento ou pelo elevado número recomendado de avaliadores. Por conta disso, medidas objetivas são utilizadas com frequência em diversos trabalhos.

4.3.1 Medidas objetivas de inteligibilidade

A primeira medida objetiva utilizada em larga escala foi o Índice de Articulação (AI - *Articulation Index*), proposto por French e Steinberg [68] e, depois, sintetizado por Kryter [69]. O índice é calculado no domínio da frequência, e utiliza a contribuição de cada banda de frequência na inteligibilidade. O resultado é um número entre zero e um, onde números maiores indicam maior inteligibilidade. Um método que simplifica o cálculo do AI foi proposto por Mueller e Killion [70]. Uma padronização do cálculo pode ser encontrada na norma S3.5 do ANSI (*American National Standards Institute*) [71].

A partir do AI, uma nova medida foi desenvolvida, com várias extensões e correções. A medida criada, chamada de Índice de Inteligibilidade de Fala (SII - *Speech Intelligibility Index*), foi padronizada na versão de 1997 do padrão S3.5 do ANSI [72]. Uma introdução ao SII pode ser encontrada em [73].

Uma extensão do SII foi feita por Kates e Arehart [74], com a criação do Índice de Inteligibilidade de Fala baseado em Coerência (CSII - *Coherence-based Speech Intelligibility Index*). A principal diferença do CSII em relação ao SII é a utilização de medidas de coerência no cálculo da relação sinal-ruído, utilizada na computação dos índices. A coerência quadrática (MSC - *Magnitude-Squared Coherence*) é utilizada, entre outros fatores, em virtude da sua aplicabilidade a qualquer sinal de excitação, incluindo a fala. Assim como no AI e no SII, o CSII resulta num número entre zero e um, onde números maiores indicam maior inteligibilidade.

Para calcular o CSII, um sinal de referência $x(t)$ é utilizado para medir o ganho ou a perda de inteligibilidade causada pelo processamento, tendo como sinal resultante $y(t)$. O espectro dos sinais é obtido através da aplicação da Transformada Discreta de Fourier (*Discrete Fourier Transform* - DFT) em versões segmentadas dos sinais obtidas via janelamento.

A MSC é estimada usando

$$\text{MSC}(f) = \frac{\left| \sum_{m=0}^{M-1} X_m(f) Y_m^*(f) \right|^2}{\sum_{m=0}^{M-1} |X_m(f)|^2 \times \sum_{m=0}^{M-1} |Y_m(f)|^2} \quad (4.3)$$

onde $X_m(f)$ e $Y_m(f)$ são respectivamente o espectro do segmento m dos sinais $x(t)$ e $y(t)$ e f é o índice da DFT. Sendo $x(t)$ e $y(t)$ respectivamente a entrada e a saída de um sistema, a MSC representa o quanto da potência do sinal de saída é linearmente dependente da entrada [74], assim como $1 - \text{MSC}(f)$ representa a distorção e o ruído presentes.

Sendo $S_{yy}(f)$ a amostra f da densidade espectral de potência do sinal de saída, estimada via DFT, a relação sinal-ruído e interferência (SDR - *Signal-to-noise and Distortion Ratio*) pode ser estimada utilizando

$$\text{SDR}(b) = \frac{\sum_{f=0}^F I_b(f) \text{MSC}(f) S_{yy}(f)}{\sum_{f=0}^F I_b(f) [1 - \text{MSC}(f)] S_{yy}(f)} \quad (4.4)$$

onde o filtro $I_b(f)$ é um filtro que implementa o peso da banda de frequências b na inteligibilidade da fala, assim como efeitos como o mascaramento de frequências. A Tabela 7 apresenta as bandas de frequência utilizadas neste trabalho, com os respectivos pesos derivados da teoria da articulação [69].

Por fim, o CSII é calculado a partir de manipulações numéricas com a versão logarítmica da SDR calculada. Mais detalhes sobre o CSII podem ser encontrados em [74].

Tabela 7: Frequências e pesos de inteligibilidade

Frequência central	Peso
16	0,0083
20	0,0095
25	0,0150
31,5	0,0289
40	0,0440
50	0,0578
63	0,0653
80	0,0711
100	0,0818
125	0,0844
160	0,0882
200	0,0898
250	0,0868
315	0,0844
400	0,0771
500	0,0527

4.3.2 Medida de desempenho utilizada no trabalho

A medida de desempenho utilizada no presente trabalho é baseada no esquema de testes descrito em [9]. Tendo como referência os sinais de entrada do conformador, a inteligibilidade dos sinais de interesse e de

interferência é calculada utilizando o CSII. O ganho de inteligibilidade na utilização do conformador em teste será definido como:

$$G_I = \text{CSII}_s - \text{CSII}_i \quad (4.5)$$

onde CSII_s e CSII_i são os índices de inteligibilidade, calculados a partir dos sinais de interesse e interferência, respectivamente, na entrada e na saída do conformador. Desta forma, as seguintes situações podem ocorrer:

- $G_I = 1$: sinal de interesse mantido e de interferência totalmente cancelado;
- $0 < G_I < 1$: sinal de interesse possui maior inteligibilidade que o de interferência;
- $G_I = 0$: igual inteligibilidade dos sinais de interesse e interferência;
- $-1 < G_I < 0$: sinal de interesse possui menor inteligibilidade que o de interferência;
- $G_I = -1$: sinal de interesse totalmente cancelado e de interferência mantido.

O uso deste índice para o conformador tem como obstáculo a necessidade de acesso aos sinais de interesse e de interferência isolados. A solução vem novamente a partir da estrutura de testes utilizada em [9], ilustrada na Figura 4.6. O procedimento, baseado no princípio da superposição, utiliza três conformadores, nomeados de *Conformador independente*, *Conformador dependente 1* e *Conformador dependente 2*.

O *Conformador independente* assume o papel do conformador adaptativo, que realiza a adaptação dos coeficientes a fim de cancelar o sinal de interferência. Na sua saída haverá componentes referentes aos sinais de interesse e interferência.

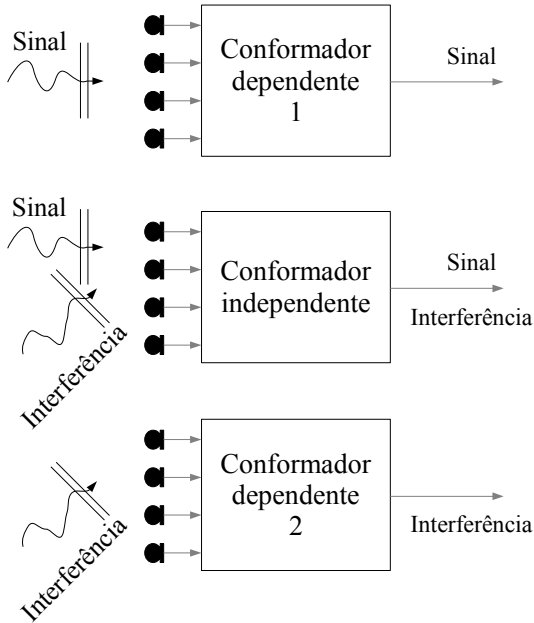


Figura 4.6: Esquema utilizado para obter os sinais de saída do conformador de forma isolada

Os conformadores dependentes são assim chamados por não realizarem adaptação dos seus coeficientes. A cada iteração, os coeficientes adaptados do *Conformador independente* são copiados pelos conformadores escravos, de forma que a submissão isolada dos sinais de entrada resulte na saída isolada das componentes referentes ao mesmo sinal. Pelo princípio da superposição, a soma dos sinais de saída dos conformadores escravos resulta no sinal de saída do *Conformador independente*.

Para a validação da medida de desempenho proposta, testes informais com áudios processados pelo sistema foram realizados, resultando numa boa correlação dos resultados da medida com resultados reais.

4.4 RESULTADOS

Esta seção apresenta os resultados obtidos com base nos parâmetros dos testes contidos na Seção 4.1.

As duas estruturas comparadas foram testadas em quatro cenários de testes, variando o tipo de ambiente e microfones utilizados. Em cada cenário, foram realizados quatro testes, cada um com uma das configurações de microfones e coeficientes apresentadas na Tabela 6.

Em cada teste, foram realizadas dez execuções, com arquivos de áudios escolhidos aleatoriamente (uniforme) dentro da base de áudios, e cada execução contou com dez realizações. O uso de diferentes realizações com os mesmos arquivos de áudio objetiva reduzir a influência de realizações específicas do ruído branco nos resultados obtidos. Execuções com diferentes arquivos de áudio retiram a dependência entre resultados, locutores e frases, além de simular a constante mudança na potência dos sinais presente num cenário real.

Visando uma boa organização do texto, os Itens 4.4.1 e 4.4.2 a seguir apresentarão os resultados das estruturas “*Frost*” e “*Multi-Split*” separadamente. Por sua vez, cada item apresentará separadamente os resultados dos quatro cenários, com as quatro configurações de microfones e coeficientes utilizados. Por fim, Item 4.4.3 compara as duas estruturas, apresentando resultados de testes adicionais realizados.

4.4.1 Resultados com a estrutura “*Frost*”

4.4.1.1 Cenário 1: *Ambiente pouco reverberante e microfones omnidirecionais*

A Tabela 8 apresenta a média dos níveis dos sinais de entrada do sistema utilizados nas dez execuções em cada uma das quatro configurações de microfones e coeficientes. Como pode ser visto, a única diferença significativa entre os níveis médios dos sinais de

interesse e de interferência ocorreu na configuração C, onde uma diferença de quase 5 dB foi obtida em favor do sinal de interferência.

Tabela 8: Níveis médios (dB) dos sinais utilizados com a estrutura "Frost" no cenário 1.

Configuração	Sinal de interesse		Sinal de interferência	
	Média	Desvio Padrão	Média	Desvio Padrão
A	70,74	3,16	70,47	4,56
B	70,88	3,95	71,42	3,60
C	69,81	3,43	74,43	2,08
D	70,59	3,12	70,57	3,25

Os resultados dos testes com as quatro configurações podem ser vistos na Tabela 9. Tanto os CSIIs quanto o G_I apresentados são valores médios calculados a partir das dez execuções com áudios diferentes.

Tabela 9: Resultados dos testes da estrutura "Frost" no cenário 1

		Configuração				
		A	B	C	D	
CSII	Sinal de Interesse	Média	1,000	1,000	1,000	0,999
		Desvio Padrão	0,000	0,000	0,001	0,002
	Sinal de Interferência	Média	1,000	0,946	0,789	0,570
		Desvio Padrão	0,000	0,040	0,111	0,163
G_I	Média	0,000	0,054	0,211	0,429	
	Desvio Padrão	0,000	0,040	0,111	0,162	

Como esperado, quanto maior o número de microfones e coeficientes utilizados, melhor o desempenho do sistema. Enquanto o sinal de interesse foi muito pouco distorcido, apresentando no pior caso um CSII de 0,999, houve um bom cancelamento do sinal de interferência, chegando no melhor caso a um CSII de 0,5699. O mesmo vale para o desvio padrão do ganho G_I , que em proporção ao valor médio, reduz à medida que o número de microfones é aumentado. Para as configurações A e B, um ganho praticamente nulo foi obtido.

A Figura 4.7 compara a resposta em frequência do conformador para o ângulo de zero graus com a resposta desejada da restrição de Frost. A resposta em frequência do conformador apresentada é a média das respostas das dez execuções realizadas com a configuração D, calculadas ao final do processamento dos sinais. Como pode ser visto, as respostas desejadas e obtidas ficaram bastante próximas.

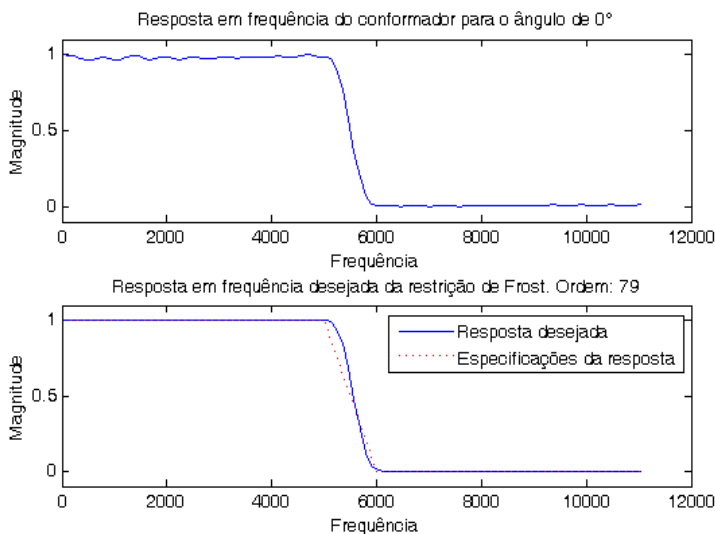


Figura 4.7: Resposta em frequência da configuração D do conformador

A Figura 4.8 apresenta um gráfico com a resposta em frequência do conformador em função do ângulo de incidência dos sinais, para as

quatro configurações do arranjo utilizadas. Cada resposta foi obtida através da média das respostas das dez execuções realizadas.

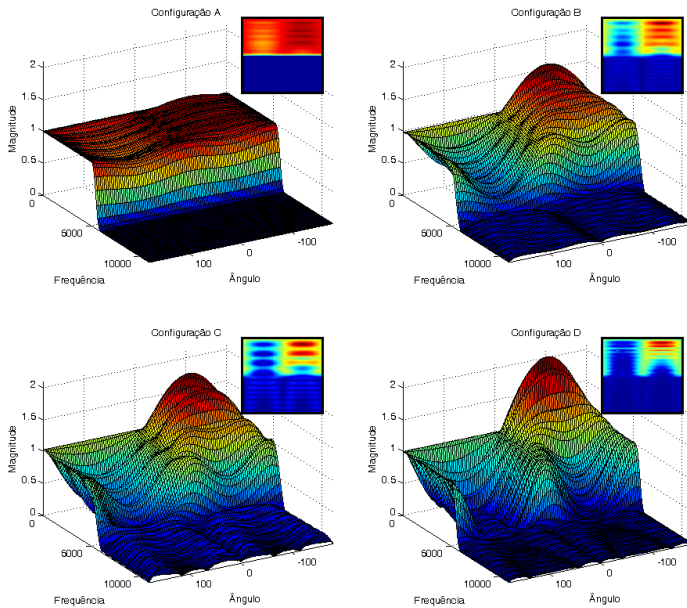


Figura 4.8: Resposta do conformador da estrutura “Frost” no cenário 1. No recorte, visão superior.

Como esperado, à medida que o número de microfones é aumentado, uma maior diretividade é obtida, o que pode ser visto na configuração D.

4.4.1.2 Cenário 2: Ambiente pouco reverberante e microfones cardioides

A Tabela 10 apresenta os resultados obtidos no cenário dois. Assim como no cenário um, os índices médios ficaram bastante próximos, beirando os 70 dB SPL .

Tabela 10: Níveis médios (dB) dos sinais utilizados com a estrutura "Frost" no cenário 2.

Configuração	Sinal de interesse		Sinal de interferência	
	Média	Desvio Padrão	Média	Desvio Padrão
A	69,82	4,74	70,14	5,11
B	69,15	3,02	70,58	3,68
C	68,97	3,37	72,29	3,17
D	70,11	2,84	70,75	3,49

Os resultados dos testes no cenário dois podem ser vistos na Tabela 11. Um aumento gradual, à medida que mais microfones são acrescentados, foi observado.

Tabela 11: Resultados dos testes da estrutura "Frost" no cenário 2

		Configuração				
		A	B	C	D	
CSII	Sinal de Interesse	Média	1,000	1,000	1,000	1,000
		Desvio Padrão	0,000	0,000	0,000	0,000
	Sinal de Interferência	Média	1,000	0,967	0,807	0,564
		Desvio Padrão	0,000	0,054	0,163	0,193
Gi	Média	0,000	0,034	0,193	0,435	
	Desvio Padrão	0,000	0,054	0,163	0,193	

A Figura 4.9 apresenta a resposta em frequência do conformador, comparando-a com a resposta desejada imposta pelas restrições de Frost. A resposta exibida foi obtida a partir da média das respostas obtidas nas dez execuções com a configuração C. As duas respostas ficaram

bastante próximas, mas a redução da ordem do filtro de resposta desejada introduziu uma pequena distorção na resposta em comparação à resposta apresentada na Figura 4.7, referente à configuração D.

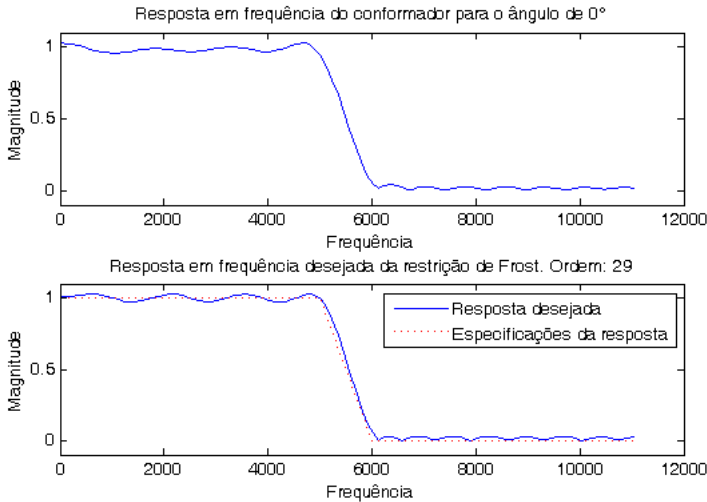


Figura 4.9: Resposta em frequência da configuração C do conformador

A resposta do conformador em função da frequência e do ângulo de incidência dos sinais, nas quatro configurações do arranjo utilizadas neste cenário, pode ser vista na Figura 4.10. Como pode ser visto, o microfone cardioide pondera a resposta do conformador, cancelando os lóbulos da resposta situados a 180° . A resposta do sistema fica muito mais diretiva utilizando microfones direcionais.

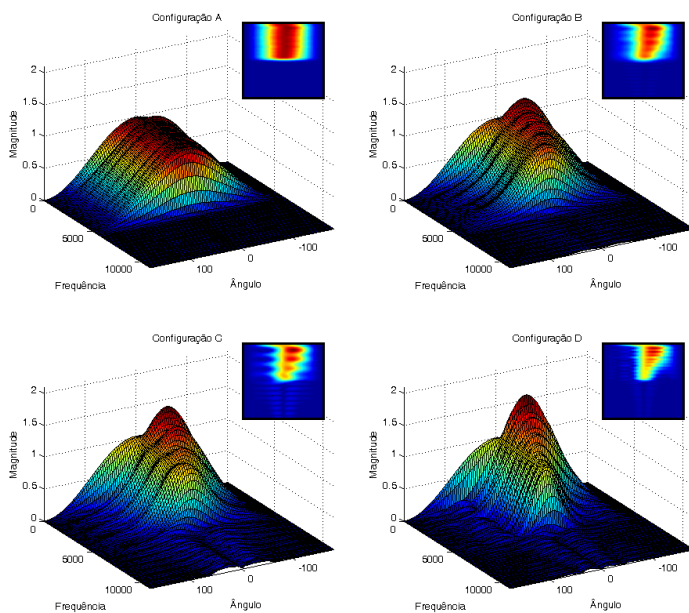


Figura 4.10: Resposta do conformador da estrutura “Frost” no cenário 2. No recorte, visão superior.

4.4.1.3 Cenário 3: Ambiente muito reverberante e microfones omnidirecionais

No terceiro cenário de testes, um arranjo com microfones omnidirecionais será aplicado em um ambiente com muita reverberação. Uma maior reverberação aumenta o número de componentes na resposta ao impulso do sinal, acarretando um aumento do comprimento significativo da resposta (T_{60}) e uma redução da relevância da componente de linha de visada. Assim, uma maior porção do sinal desejado incidirá no arranjo fora da linha de foco, assim como uma maior porção do sinal de interferência incidirá no arranjo a zero graus.

Na prática, uma maior correlação entre os sinais de interesse e de interferência é observada.

Os níveis médios dos sinais utilizados nos testes podem ser vistos na Tabela 12. Neste cenário, a maior diferença nos níveis médios foi de menos de 2dB na configuração D.

Tabela 12: Níveis médios (dB) dos sinais utilizados com a estrutura "Frost" no cenário 3.

Configuração	Sinal de interesse		Sinal de interferência	
	Média	Desvio Padrão	Média	Desvio Padrão
A	69,62	3,27	69,19	3,29
B	70,91	3,11	70,45	3,51
C	69,70	5,32	70,50	3,67
D	69,50	3,54	71,06	2,89

Os resultados obtidos podem ser vistos na Tabela 13. Como esperado, à medida que o número de microfones e coeficientes é elevado, um maior cancelamento no sinal de interferência é obtido. Este cancelamento, porém, também é observado no sinal de interesse.

Tabela 13: Resultados dos testes da estrutura “Frost” no cenário 3

			Configuração			
			A	B	C	D
CSII	Sinal de Interesse	Média	1,000	0,972	0,934	0,800
		Desvio Padrão	0,000	0,025	0,037	0,049
	Sinal de Interferência	Média	1,000	0,889	0,798	0,579
		Desvio Padrão	0,000	0,034	0,044	0,065
G_I	Média	0,000	0,083	0,136	0,221	
	Desvio Padrão	0,000	0,033	0,042	0,084	

A Figura 4.11 apresenta a resposta em frequência obtida para a configuração B do cenário três, e compara com a resposta desejada. Novamente, a figura foi obtida a partir da média das respostas das dez execuções realizadas. Assim como na configuração C (Figura 4.9), uma distorção da resposta com relação às especificações ocorreu devido à baixa ordem do filtro. De qualquer forma, a resposta obtida pelo conformador é muito semelhante à resposta desejada.

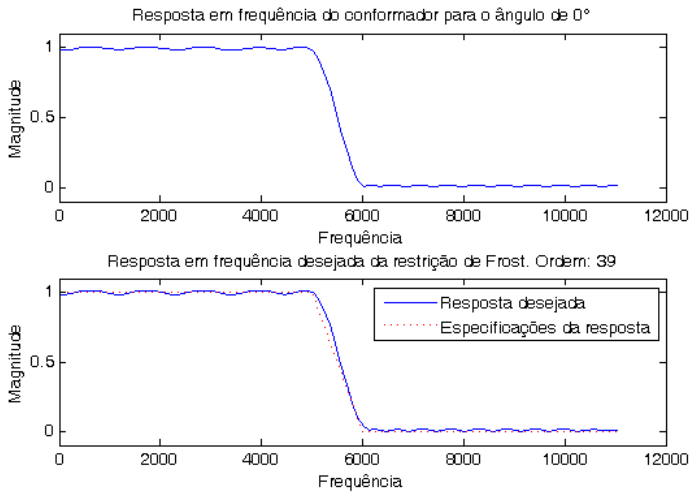


Figura 4.11: Resposta em frequência da configuração B do conformador

A resposta obtida para as quatro configurações pode ser vista na Figura 4.12.

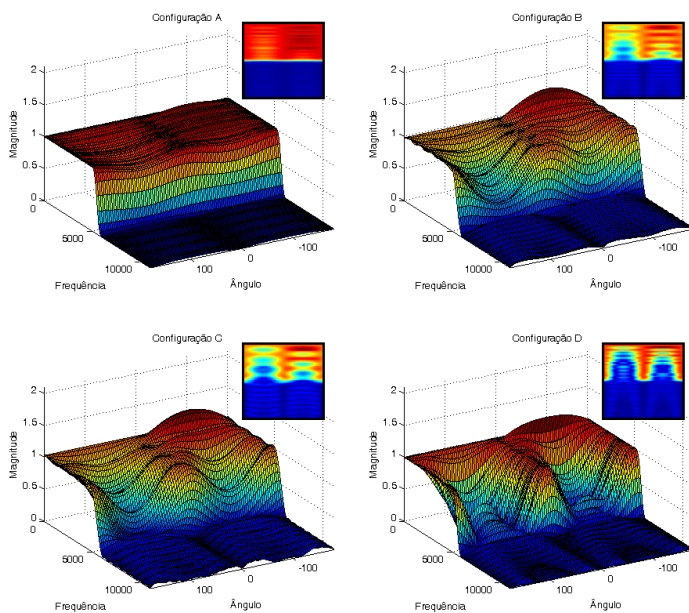


Figura 4.12: Resposta do conformador da estrutura “Frost” no cenário 3. No recorte, visão superior.

4.4.1.4 Cenário 4: Ambiente muito reverberante e microfones cardioides

Para o cenário quatro, os níveis dos sinais utilizados flutuaram ao redor dos 71 dB (Tabela 14).

Tabela 14: Níveis médios (dB) dos sinais utilizados com a estrutura "Frost" no cenário 4

Configuração	Sinal de interesse		Sinal de interferência	
	Média	Desvio Padrão	Média	Desvio Padrão
A	70,99	4,72	70,55	3,23
B	71,53	3,36	71,93	3,16
C	70,45	5,93	71,13	4,21
D	71,58	3,49	71,95	3,61

Os resultados dos testes realizados podem ser vistos na Tabela 15.

Tabela 15: Resultados dos testes da estrutura "Frost" no cenário 4

		Configuração				
		A	B	C	D	
CSII	Sinal de Interesse	Média	1,000	0,999	0,972	0,902
		Desvio Padrão	0,000	0,002	0,014	0,015
	Sinal de Interferência	Média	1,000	0,881	0,703	0,498
		Desvio Padrão	0,000	0,078	0,103	0,111
G_I	Média	0,000	0,118	0,270	0,404	
	Desvio Padrão	0,000	0,078	0,099	0,107	

A comparação das respostas em frequência obtida e desejada para a configuração A do cenário quatro pode ser vista na Figura 4.13. O resultado foi obtido a partir da média das dez execuções realizadas. Por possuir a maior ordem, a configuração A apresentou a melhor resposta, e foi também obtida corretamente após o processo de adaptação do conformador.

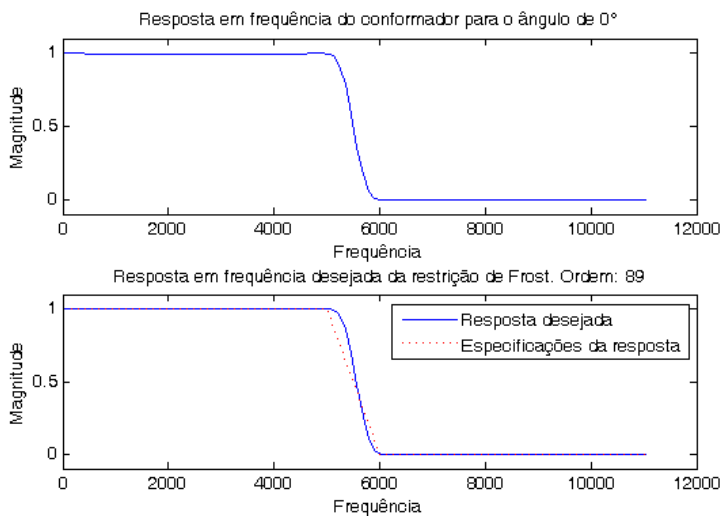


Figura 4.13: Resposta em frequência da configuração A do conformador

A Figura 4.14 apresenta a resposta do sistema nas quatro configurações testadas.

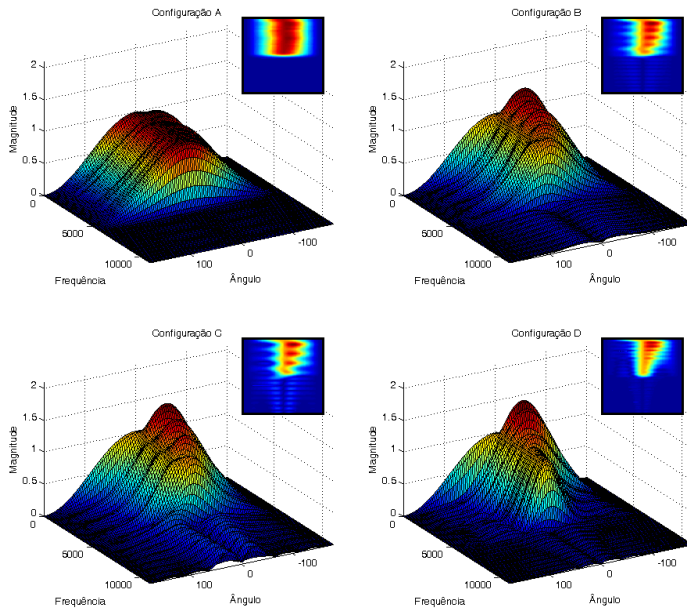


Figura 4.14: Resposta do conformador da estrutura “Frost” no cenário 4.
No recorte, visão superior.

4.4.1.5 Comparação dos resultados

A Figura 4.15 resume a redução de inteligibilidade dos sinais de interesse e interferência obtida pela aplicação das quatro configurações da estrutura “Frost” nos quatro cenários definidos. A Figura 4.16 apresenta os ganhos obtidos.

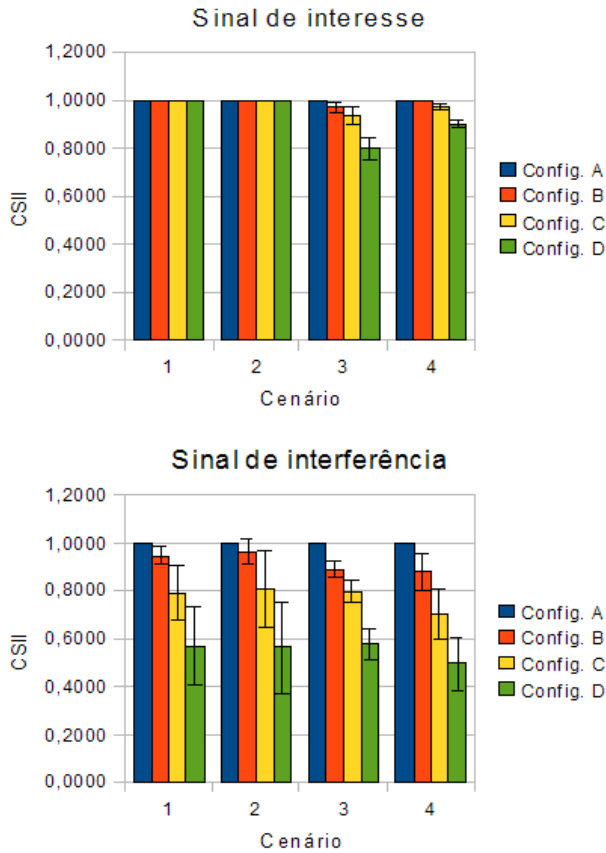


Figura 4.15: CSIs obtidos para a estrutura "Frost"

Como pode ser visto pelas figuras, a não ser por uma leve diferença na variação, resultados muito próximos foram obtidos para os cenários um e dois.

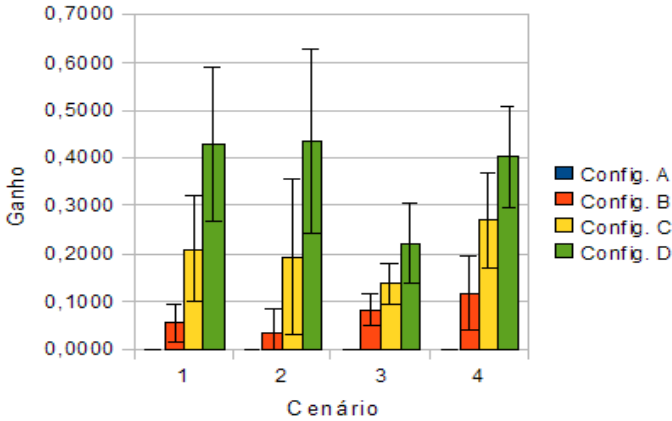


Figura 4.16: Ganho obtido para a estrutura “Frost”

A primeira causa possível é que o uso de microfones direcionais (cardioide) não é traduzido em um ganho significativo no sistema. Sendo o ambiente pouco reverberante, a contribuição da componente de linha de visada na resposta do ambiente é muito significativa em comparação com as demais componentes do sinal. Para o cenário de testes em questão, a atenuação da componente de linha de visada do sinal de interferência realizada pelo microfone direcional foi pequena.

A segunda causa possível para a semelhança nos resultados dos cenários um e dois é a não incorporação dos efeitos da utilização de microfones direcionais na medida de desempenho utilizada. Como o ganho G_I é função da inteligibilidade dos sinais, calculado a partir do CSII, uma pequena contribuição dos microfones direcionais na redução da inteligibilidade dos sinais de interferência pode ser a causa para o baixo ganho de desempenho obtido.

A comparação dos resultados dos cenários um e três mostra uma grande semelhança no cancelamento do sinal de interferência, mas um aumento expressivo do cancelamento do sinal de interesse no cenário três. Como esperado, devido à maior reverberação do ambiente, mais componentes do sinal de interesse são recebidas fora da linha de foco do conformador, assim como mais componentes do sinal de interferência

são recebidas dentro da linha de foco. Este fato causa um aumento da correlação entre os sinais considerados como de interesse e interferência, reduzindo o desempenho do sistema.

Comparando com o cenário dois, foi obtido no cenário quatro, um maior cancelamento dos sinais de interesse e interferência, com menor variação dos resultados.

O cenário quatro apresentou um menor cancelamento do sinal de interesse e um maior cancelamento do sinal de interferência que o cenário três, que possui reverberação semelhante, mas usa microfones omnidirecionais. Essa melhoria dos resultados é ainda acentuada à medida que mais microfones são utilizados no arranjo. Este comportamento se deve ao cancelamento de componentes reverberantes dos sinais pela resposta cardioide do microfone, reduzindo, conseqüentemente, a correlação entre os sinais de interesse e interferência. Uma causa para o mesmo comportamento não ser observado na comparação entre os cenários um e dois é a pequena reverberação do ambiente.

4.4.2 Resultados com a estrutura “*Multi-Split*”

4.4.2.1 Cenário 1: Ambiente pouco reverberante e microfones omnidirecionais

A Tabela 16 apresenta os níveis médios dos sinais utilizados nos testes no primeiro cenário com o algoritmo MS-LMS. Níveis médios próximos dos 70 dB SPL foram obtidos.

Tabela 16: Níveis médios (dB) dos sinais utilizados com a estrutura "Multi-Split" no cenário 1

Configuração	Sinal de interesse		Sinal de interferência	
	Média	Desvio Padrão	Média	Desvio Padrão
A	71,05	4,86	69,28	3,80
B	71,08	4,06	70,73	2,88
C	70,49	4,86	70,75	4,21
D	71,36	2,74	70,50	2,74

Os resultados dos testes podem ser vistos na Tabela 17. A partir do resultado com a primeira estrutura, muito próximo de zero, à medida que mais microfones são incluídos nos testes, melhores resultados podem ser observados. Para um cenário pouco reverberante, houve um grande cancelamento do sinal de interesse na configuração D, compensada pelo grande cancelamento do sinal de interferência.

Tabela 17: Resultados dos testes da estrutura "Multi-Split" no cenário 1

		Configuração				
		A	B	C	D	
CSII	Sinal de Interesse	Média	1,000	0,998	0,998	0,884
		Desvio Padrão	0,000	0,006	0,004	0,055
	Sinal de Interferência	Média	0,969	0,598	0,496	0,103
		Desvio Padrão	0,049	0,115	0,053	0,031
G_I	Média	0,031	0,400	0,503	0,781	
	Desvio Padrão	0,049	0,116	0,054	0,069	

A Figura 4.17 apresenta a resposta do conformador para as quatro configurações utilizadas. Pode ser visto na figura que as restrições designadas foram respeitadas. À medida que mais microfones são utilizados, uma maior seletividade é obtida.

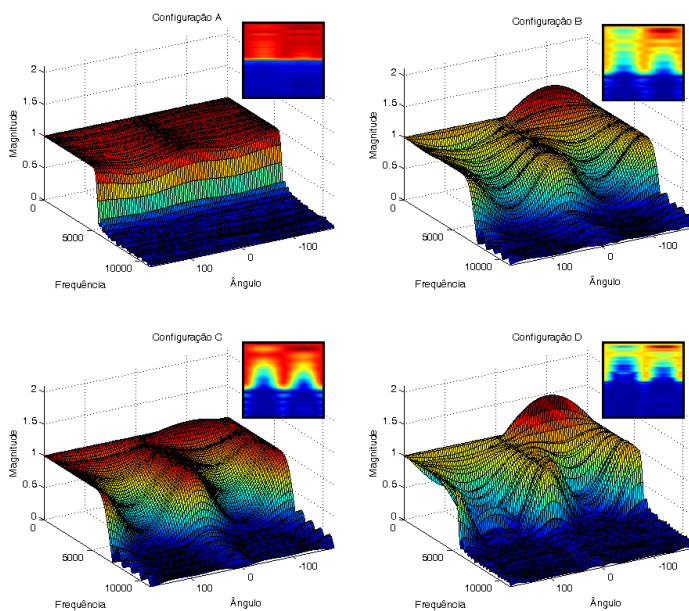


Figura 4.17: Resposta do conformador da estrutura “Multi-Split” no cenário 1. No recorte, visão superior.

4.4.2.2 Cenário 2: Ambiente pouco reverberante e microfones cardioides

Para o cenário dois, os sinais de áudio utilizados nos testes também flutuaram próximos dos 70 dB SPL (Tabela 18).

Tabela 18: Níveis médios (dB) dos sinais utilizados com a estrutura "Multi-Split" no cenário 2

Configuração	Sinal de interesse		Sinal de interferência	
	Média	Desvio Padrão	Média	Desvio Padrão
A	71,36	4,45	70,50	4,44
B	71,06	4,68	71,93	3,29
C	70,35	3,31	69,82	3,02
D	71,25	2,38	71,00	4,57

O resultado dos testes pode ser visto na Tabela 19. O bom ganho médio de 0,864 foi obtido na configuração D, a partir de um pequeno cancelamento do sinal de interesse e um grande cancelamento do sinal de interferência.

Tabela 19: Resultados dos testes da estrutura "Multi-Split" no cenário 2

		Configuração				
		A	B	C	D	
CSII	Sinal de Interesse	Média	1,000	1,000	0,998	0,971
		Desvio Padrão	0,000	0,000	0,006	0,069
	Sinal de Interferência	Média	0,909	0,587	0,495	0,108
		Desvio Padrão	0,096	0,166	0,120	0,058
G_I	Média	0,091	0,413	0,504	0,864	
	Desvio Padrão	0,096	0,166	0,120	0,105	

As respostas das quatro configuração de conformadores podem ser vistas na Figura 4.18. Assim como na estrutura "Frost", a utilização

de microfones direcionais pondera a resposta do conformador, atenuando sinais que incidem a um ângulo diferente de zero.

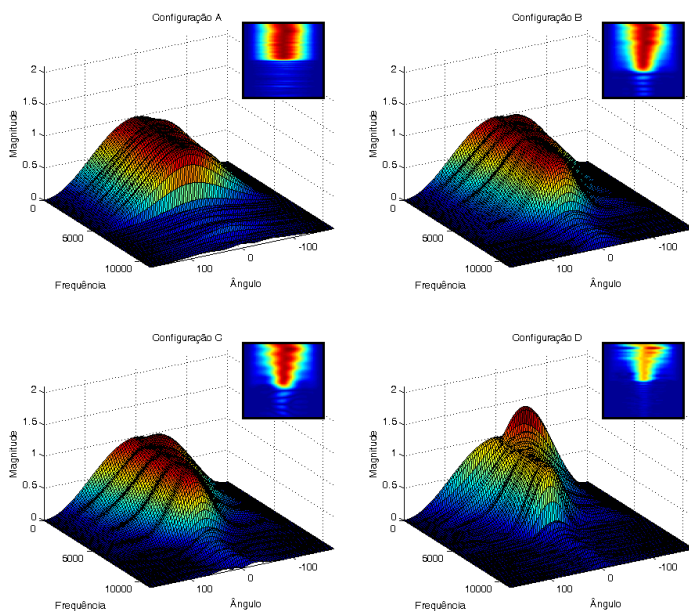


Figura 4.18: Resposta do conformador da estrutura “*Multi-Split*” no cenário 2. No recorte, visão superior.

4.4.2.3 Cenário 3: Ambiente muito reverberante e microfones omnidirecionais

A Tabela 20 apresenta os níveis médios dos sinais de áudio utilizados nos testes do terceiro cenário.

Tabela 20: Níveis médios (dB) dos sinais utilizados com a estrutura "Multi-Split" no cenário 3

Configuração	Sinal de interesse		Sinal de interferência	
	Média	Desvio Padrão	Média	Desvio Padrão
A	68,39	4,84	68,67	3,37
B	70,45	3,28	70,20	4,92
C	70,56	2,59	70,82	4,24
D	71,20	4,12	72,57	3,96

Os resultados podem ser vistos na Tabela 21. Como esperado, a maior reverberação do ambiente reduziu o desempenho do sistema, aumentando o cancelamento do sinal de interesse e reduzindo o cancelamento do sinal de interferência.

Tabela 21: Resultados dos testes da estrutura "Multi-Split" no cenário 3

		Configuração				
		A	B	C	D	
CSII	Sinal de Interesse	Média	0,994	0,898	0,808	0,520
		Desvio Padrão	0,009	0,057	0,059	0,100
	Sinal de Interferência	Média	0,918	0,682	0,617	0,246
		Desvio Padrão	0,066	0,065	0,036	0,061
G_I	Média	0,076	0,216	0,191	0,274	
	Desvio Padrão	0,066	0,065	0,085	0,101	

A Figura 4.19 apresenta a resposta das quatro configurações do sistema.

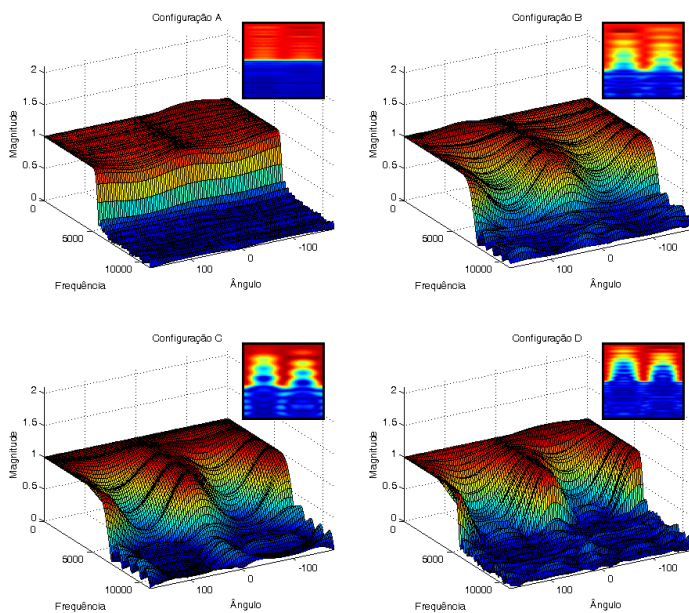


Figura 4.19: Resposta do conformador da estrutura “*Multi-Split*” no cenário 3. No recorte, visão superior.

4.4.2.4 Cenário 4: Ambiente muito reverberante e microfones cardioides

Os níveis médios áudios utilizados no último cenário de testes do algoritmo Multi-Split LMS estão exibidos na Tabela 22.

Tabela 22: Níveis médios (dB) dos sinais utilizados com a estrutura "Multi-Split" no cenário 4

Configuração	Sinal de interesse		Sinal de interferência	
	Média	Desvio Padrão	Média	Desvio Padrão
A	70,68	3,08	71,68	2,47
B	69,69	5,15	72,32	2,25
C	69,59	4,97	69,19	3,66
D	68,03	3,73	70,10	4,18

Os resultados obtidos podem ser vistos na Tabela 23.

Tabela 23: Resultados dos testes da estrutura "Multi-Split" no cenário 4

		Configuração				
		A	B	C	D	
CSII	Sinal de Interesse	Média	0,998	0,977	0,878	0,708
		Desvio Padrão	0,006	0,014	0,065	0,077
	Sinal de Interferência	Média	0,918	0,585	0,460	0,186
		Desvio Padrão	0,046	0,077	0,070	0,052
G_I	Média	0,080	0,392	0,418	0,523	
	Desvio Padrão	0,047	0,070	0,098	0,097	

A Figura 4.20 apresenta as respostas obtidas.

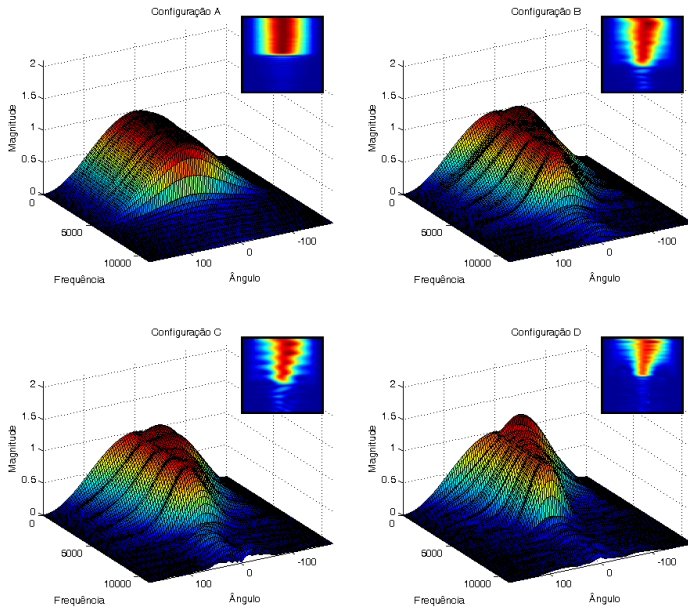


Figura 4.20: Resposta do conformador da estrutura “*Multi-Split*” no cenário 4. No recorte, visão superior.

4.4.2.5 Comparação dos resultados

Para a estrutura “*Multi-Split*”, a síntese dos resultados pode ser vista nas Figuras 4.21 e 4.22.

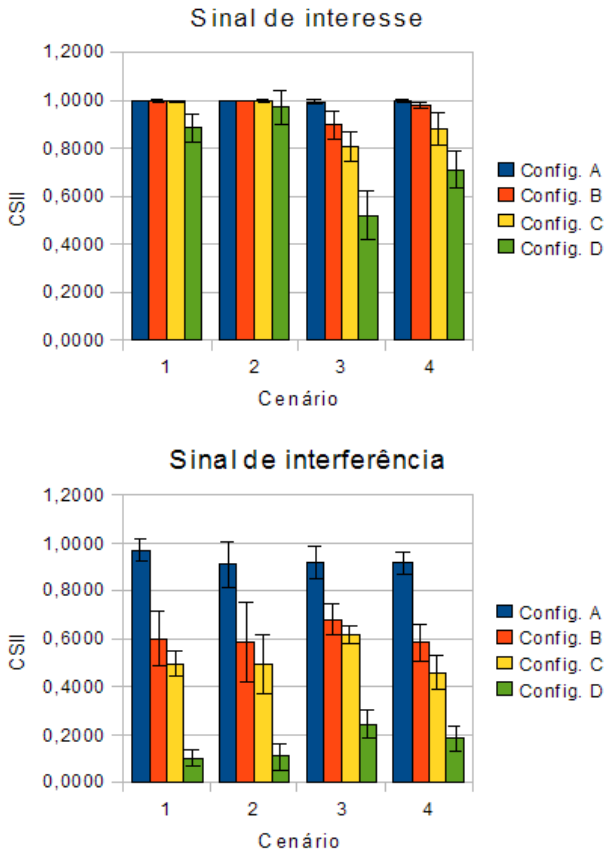


Figura 4.21: CSII's obtidos para a estrutura "Multi-Split"

Uma comparação dos resultados aponta para uma leve melhoria no desempenho do cenário dois em função do cenário um, principalmente na configuração A e D, onde o sinal de interferência foi mais cancelado e o sinal de interesse menos cancelado.

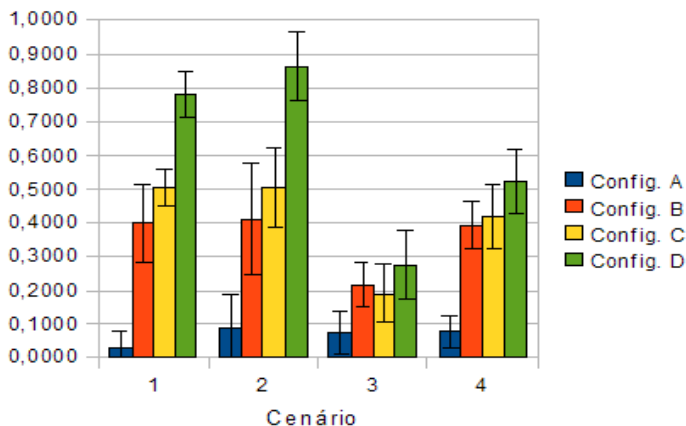


Figura 4.22: Ganho obtido para a estrutura “*Multi-Split*”

Comparado ao cenário um, o cenário três apresentou um desempenho pior, com grande cancelamento do sinal de interesse e menor cancelamento do sinal de interferência. Como esperado, a maior correlação entre o sinal de interesse e interferência reduziu o desempenho do sistema.

O cenário quatro ainda apresentou um grande cancelamento do sinal de interesse quando comparado ao cenário dois, mas este foi menor do que o obtido no cenário três. Isto indica que o aumento da correlação dos sinais causado pela maior reverberação do ambiente dificulta a ação do algoritmo Multi-Split LMS. A adição de microfones direcionais reduz a correlação entre os sinais de interesse e interferência, melhorando o desempenho. As mesmas conclusões são válidas na comparação do desempenho para o sinal de interferência.

4.4.3 Comparação das estruturas

Para comparar os resultados das duas estruturas, o ganho de inteligibilidade obtido na estrutura “*Multi-Split*” é subtraído pelo ganho da estrutura “*Frost*”. Desta forma, resultados positivos indicam um maior desempenho em favor da estrutura “*Multi-Split*”, assim como resultados negativos apontam para a estrutura “*Frost*”.

A Figura 4.23 apresenta a comparação dos resultados. Como pode ser visto, a estrutura “*Multi-Split*” obteve um melhor desempenho em todas as configurações e cenários, dado que todo o gráfico está voltado para a parte positiva do eixo.

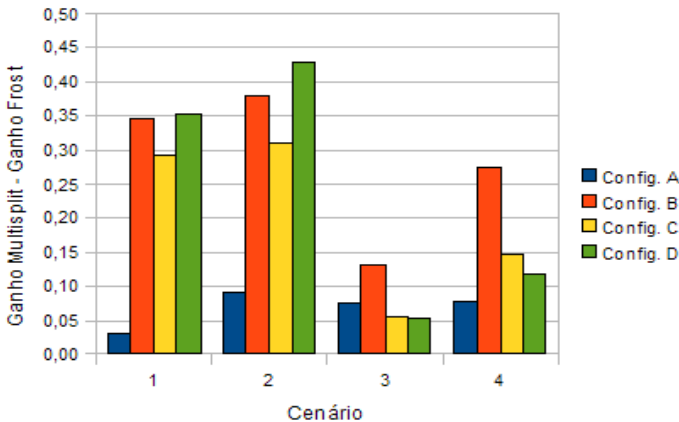


Figura 4.23: Comparação dos ganhos obtidos pelas duas estruturas

Comparando os resultados, pode ser visto que a estrutura “*Multi-Split*” (Figura 4.21) apresentou um maior cancelamento do sinal de interesse do que a estrutura “*Frost*” (Figura 4.15), comportamento acentuado à medida que mais microfones são adicionados na estrutura.

No cenário um, que utiliza microfones omnidirecionais e o ambiente apresenta pouca reverberação, o aumento do cancelamento do sinal de interesse pela estrutura “*Multi-Split*” em relação à “*Frost*” foi pouco significativo nas configurações A, B e C. Com sete microfones,

na configuração D, a redução do CSII foi de 11,55 pontos percentuais em relação à escala. O cenário dois apresenta resultados semelhantes ao cenário um, sendo a maior diferença entre os CSII no cenário D de 2,87 pontos percentuais.

Com o aumento da reverberação do ambiente no cenário três, o sinal de interesse foi ainda mais cancelado pelo Multi-Split LMS. A diferença entre os CSII do sinal de interesse nas estruturas “*Multi-Split*” e “*Frost*” é pouco significativa apenas na configuração A, e chega a pouco mais de 38 pontos percentuais na configuração D. Esta marca foi obtida a partir do CSII de 0,5195 do “*Multi-Split*”, o pior resultado obtido em todos os testes. O cenário quatro apresenta comportamento semelhante mas em menor magnitude, sendo a maior diferença entre os CSII de quase 19,38 pontos percentuais.

O aumento do cancelamento do sinal de interesse pela estrutura “*Multi-Split*” em relação à estrutura “*Frost*” foi, porém, compensado pelo aumento expressivo do cancelamento do sinal de interferência.

Com dois microfones, o aumento no cancelamento do sinal de interferência foi de até pouco mais de 9 pontos percentuais no cenário dois.

Para a configuração B uma maior redução do CSII do sinal de interferência foi observada, indo de 20,68 pontos percentuais no cenário três a 37,92 pontos percentuais no cenário dois. Os resultados da configuração C ficaram na média quatro pontos percentuais abaixo do cenário B.

A maior redução do CSII observada, como esperado, foi obtida para a configuração D, chegando a quase 47 pontos percentuais no cenário um.

4.4.3.1 Teste adicional com os mesmos arquivos de áudio

A fim de confirmar os resultados obtidos, um teste adicional foi realizado. Como os testes anteriores utilizaram arquivos de áudios aleatoriamente escolhidos dentro da base de áudios, é possível que os

sinais utilizados tenham alterado os resultados, mesmo considerando as múltiplas realizações. Além disso, os arquivos disponíveis são de pequena duração, inferior a cinco segundos, sendo interessante a realização de testes com áudios de maior duração.

O último teste, realizado apenas na configuração D e cenário dois, foi projetado tendo como entrada os mesmos arquivos de áudio para ambas as estruturas. Os arquivos de áudio utilizados foram criados a partir da concatenação de dez arquivos originais, escolhidos aleatoriamente dentro da pasta de cada locutor. O mesmo arquivo de áudio foi usado nos testes das estruturas “*Frost*” e “*Multi-Split*”. Os níveis de fala ativa resultante dos sinais de interesse e interferência foram, respectivamente, 67,81 dB SPL e 68,07 dB SPL .

A Tabela 24 apresenta os resultados obtidos. Neste ponto, cabe o apontamento que o desvio padrão obtido é resultante das múltiplas realizações do ruído gaussiano adicionado, já que este teste não foi realizado com diferentes distribuições de arquivos de áudio.

Pelos resultados, o cenário favorável à estrutura “*Multi-Split*” se confirma. Devido à natureza pouco reverberante do ambiente, um cancelamento mínimo do sinal de interesse foi observado. Para o sinal de interferência, o resultado final foi a total redução da inteligibilidade, com índice muito próximo de zero.

Tabela 24: Resultados dos testes finais com as duas estruturas

		Estrutura		
		<i>Frost</i>	<i>Multi-Split</i>	
CSII	Sinal de Interesse	Média	1,000	0,986
		Desvio Padrão	0,000	0,003
	Sinal de Interferência	Média	0,528	0,056
		Desvio Padrão	0,002	0,003
G_I	Média	0,472	0,929	
	Desvio Padrão	0,002	0,003	

4.4.3.2 *Teste adicional com três fontes sonoras*

Todos os resultados apontados até o presente utilizam como entrada um sinal de interesse e um de interferência. Nesta seção, resultados são apresentados a partir de um cenário onde três fontes sonoras competem a atenção do ouvinte (Figura 4.24). No caso, as fontes de interferência estão localizadas respectivamente a 45° e a 270° da localização da fonte e interesse, respectivamente nas coordenadas $(2,048 \times 2,087 \times 1,6)$ e $(3,75 \times 1,38 \times 1,6)$.

Os áudios utilizados foram gerados de acordo com a metodologia utilizada no teste anteriormente apresentado (Seção 4.4.3.1), a partir da concatenação de dez arquivos originais. Da mesma forma, os resultados compreendem apenas os valores médios obtidos a partir das múltiplas realizações de ruídos, já que somente uma distribuição de arquivos de áudio foi utilizada.

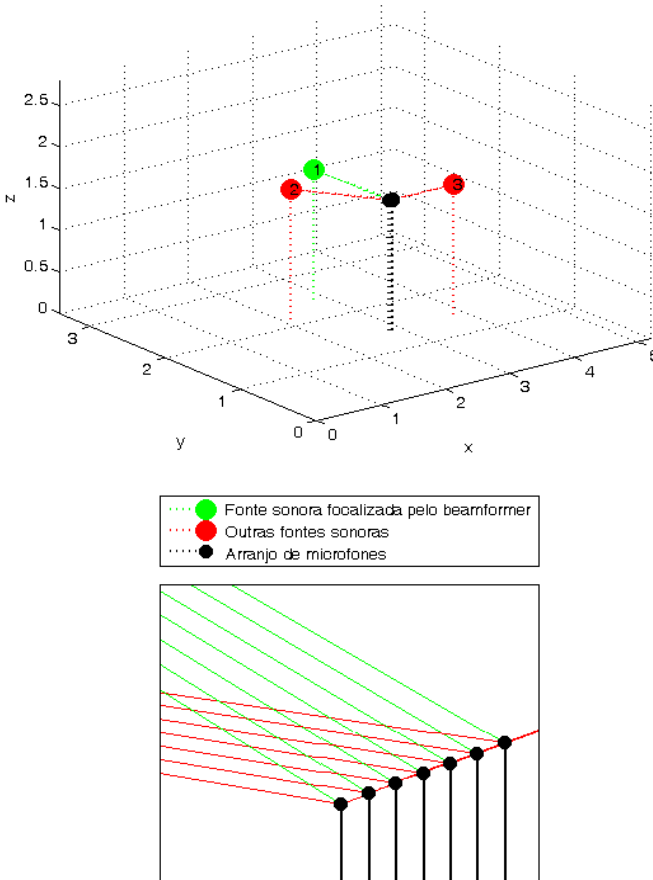


Figura 4.24: Cenário para o teste com três fontes sonoras. O recorte demonstra o ângulo de incidência das ondas geradas pelas fontes sonoras de interesse (verde) e interferências (vermelho).

Como pode ser visto na Tabela 25, a superioridade da estrutura “*Multi-Split*” se confirmou mesmo nesse cenário.

Tabela 25: Resultados dos testes finais para três fontes sonoras no ambiente

		Estrutura		
		<i>Frost</i>	<i>Multi-Split</i>	
CSII	Sinal de Interesse	Média	1,000	0,969
		Desvio Padrão	0,000	0,004
	Sinal de Interferência	Média	0,382	0,045
		Desvio Padrão	0,001	0,002
G_I		Média	0,619	0,924
		Desvio Padrão	0,001	0,004

4.5 CONCLUSÕES DO CAPÍTULO

Este capítulo apresentou os resultados dos testes realizados com as estruturas “*Frost*” e “*Multi-Split*”. Quatro configurações de arranjo foram utilizadas, variando o número de microfones e coeficientes dos filtros. Os testes foram realizados em dois ambientes distintos, através da variação da quantidade de reverberação presente, e dois tipos diferentes de microfones foram utilizados. A distribuição dos dois tipos de microfones nos dois ambientes deu origem aos quatro cenários utilizados nos testes.

Inicialmente, os parâmetros da simulação desenvolvida foram descritos, justificando sua utilização. Os cenários dos testes foram apresentados e os parâmetros dos algoritmos adaptativos foram definidos.

Em seguida, a base de áudios utilizada nos testes foi apresentada, sendo devidamente creditada ao professor Dr. Carlos Alberto Ynoguti. Considerações sobre os parâmetros dos arquivos de áudio foram apresentadas.

O próximo item apresentado foi a medida de desempenho utilizada na avaliação dos testes. Em virtude das dificuldades inerentes à utilização de medidas subjetivas de avaliação, foi escolhida uma medida objetiva como fonte de referência. A medida, calculada a partir CSII, indica o ganho de inteligibilidade efetuado na execução dos testes.

Por fim, os resultados foram apresentados, separados por estruturas e, posteriormente, separados por cenários.

Como resultado, a estrutura “*Multi-Split*” se mostrou superior à estrutura “*Frost*”, em todos os cenários e configurações testados. O desempenho de ambas as estruturas é prejudicado pelo aumento da reverberação do ambiente, e é ligeiramente melhorado com a utilização de microfones direcionais. O desempenho do sistema é proporcional ao aumento do número de microfones no arranjo de microfones.

5 CONSIDERAÇÕES FINAIS

5.1 CONCLUSÕES

Este trabalho apresentou a utilização de um arranjo de microfones visando a redução de ruídos e interferências em aparelhos auditivos. Reduzindo interferências, a capacidade de seletividade espectral e temporal da audição humana é aumentada, tornando a conversação mais agradável.

Os resultados apresentados foram obtidos através de simulações computacionais. Porém, a fim de obter a maior correlação possível dos resultados simulados com resultados obtidos em testes práticos, uma série de cuidados tiveram de ser tomados, como a utilização da resposta do canal acústico e dos microfones, bem como o uso de sinais de fala reais.

As técnicas adaptativas endereçadas neste trabalho baseiam-se em métodos de otimização numérica com restrições, amplamente utilizadas nas mais diversas áreas de estudo, das comunicações digitais à economia.

Como foco do trabalho, foi proposta a utilização do algoritmo MS-LMS na aplicação de interesse, em lugar ao algoritmo CLMS clássico, dada a maior eficiência apontada por estudos recentes do algoritmo. O algoritmo clássico foi ainda utilizado como referência para comparação dos resultados obtidos para a estrutura “*Multi-Split*”.

As diferenças entre os algoritmos foram ainda ampliadas pelo uso de estruturas (direta e GSC) e restrições (Frost e autovetoriais) diferentes. Desta forma, o objetivo do trabalho não foi a comparação de dois algoritmos, mas sim o confronto de dois sistemas, um clássico e um recente, e visualização dos ganhos obtidos com o avanço dos estudos da área.

De fato, resultados demonstram que o algoritmo MS-LMS aliado à estrutura GSC e restrições autovetoriais foi superior ao algoritmo LMS em todos os testes realizados, confirmando as expectativas.

Dos objetivos específicos, testes indicaram que a extensão do modelo de canal acústico para a utilização de microfones direcionais foi válida, reduzindo as componentes esperadas à medida que os microfones direcionais são utilizados.

Da mesma forma, os testes realizados a fim de validar a medida de desempenho proposta, confirmaram a correlação dos resultados da medida com os resultados práticos.

Este estudo também demonstrou os ganhos obtidos com o aumento do número de microfones no arranjo. Como esperado, o desempenho de ambos os sistemas é proporcional ao número de microfones utilizados, embora haja a limitação espacial dos aparelhos auditivos a um número de microfones elevado.

Os testes foram realizados em dois ambientes, alterando o tempo de reverberação. Os impactos da reverberação foram diretamente observados nos resultados, causando em geral um aumento do cancelamento do sinal de interesse e uma redução do cancelamento do sinal de interferência. Como uma alternativa para ambientes reverberantes, os microfones direcionais inibiram uma parte das componentes refletidas dos sinais, reduzindo o impacto desses ambientes nos resultados.

5.2 TRABALHOS FUTUROS

Em virtude da riqueza de detalhes que a aplicação envolve, uma série de desenvolvimentos podem ser propostos para a continuidade deste trabalho.

Visando uma maior correlação dos resultados obtidos com o observado no mundo real, uma implementação em dispositivos de

hardware como DSPs (*Digital Signal Processing*) e FPGAs (*Field-Programmable Gate Array*) pode ser feita. Para isto, várias medidas de otimização de código deverão ser tomadas, dadas as limitações de recursos desses dispositivos. Além da limitação de recursos, a implementação prática envolve parâmetros não considerados neste trabalho, como falhas na distribuição dos microfones do arranjo e tolerâncias do processo de fabricação de componentes. Uma avaliação do impacto desses fatores práticos pode ser realizada.

Embora amplamente utilizada, a modelagem do canal acústico utilizada neste trabalho é uma simplificação do que ocorre na prática, já que ambientes reais possuem objetos, aberturas, diferenças de material, entre vários outros fatores que podem contribuir significativamente para a resposta do ambiente. Desta forma, testes adicionais em ambientes reais podem ser feitos.

Um outro ponto que pode ser desenvolvido diz respeito à medida de desempenho proposta. Testes formais precisam ser realizados a fim de confirmar a sua validade. Outra opção é a utilização de outras medidas de desempenho, ou mesmo medidas subjetivas, para a avaliação das estruturas. Um teste prático envolvendo deficientes auditivos é de fundamental importância para a validação da aplicação das técnicas propostas na aplicação desejada.

Saindo do estudo prático e indo ao encontro da análise teórica, um estudo mais detalhado dos algoritmos aplicados a esse problema pode ser realizado. Uma análise estatística do comportamento dos algoritmos associada às características particulares dos sinais envolvidos é imprescindível para garantir a estabilidade do algoritmo e, por consequência, da aplicação. Ainda, mais detalhes sobre a influência da correlação dos sinais precisam ser investigados.

REFERÊNCIAS

- [1] Instituto Brasileiro de Geografia e Estatística - IBGE, *Censo demográfico 2010: Características gerais da população, religião e pessoas com deficiência*. Rio de Janeiro: , 2010.
- [2] Y. R. M. Maluenda, “Análise estatística do algoritmo LMS aplicado ao cancelamento de retroalimentação acústica em aparelhos auditivos”, Universidade Federal de Santa Catarina, Santa Catarina, 2009.
- [3] H. Dillon, *Hearing Aids*, 1º ed. Thieme, 2001.
- [4] B. R. Glasberg e B. C. Moore, “Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech”, *Scandinavian Audiology*, vol. Suppl 32, p. 25, 1988.
- [5] M. C. Killion, “SNR Loss: ‘I Can Hear What People Say but I Can’t Understand Them.’”, *The Hearing Review*, vol. 4, nº 12, 1997b.
- [6] M. C. Killion, “The SIN report: Circuits haven’t solved the hearing-in-noise problem”, *The hearing journal*, United States of America, 1997a.
- [7] P. M. Peterson, “Adaptive array processing for multiple microphone hearing aids”, 1989.
- [8] O. L. Frost, “An algorithm for linearly constrained adaptive array processing”, *Proc. IEEE*, vol. 60, nº 8, p. 926–935, 1972.
- [9] J. E. Greenberg, “A real-time adaptive-beamforming hearing aid”, Thesis, Massachusetts Institute of Technology, 1989.
- [10] J. E. Greenberg, “Improved design of microphone-array hearing aids”, Thesis, Massachusetts Institute of Technology, 1994.
- [11] M. W. Hoffman, T. D. Trine, K. M. Buckley, e D. J. Van Tasell, “Robust adaptive microphone array processing for hearing aids: realistic speech enhancement”, *J. Acoust. Soc. Am.*, vol. 96, nº 2 Pt 1, p. 759–770, ago. 1994.
- [12] J. Benesty, J. Chen, e Y. Huang, *Microphone array signal processing*. Springer, 2008.
- [13] W. C. Sabine, *Collected papers on acoustics*. Cambridge : Harvard University Press, 1922.
- [14] C. F. Eyring, “Reverberation Time in ‘Dead’ Rooms”, *The Journal*

- of the Acoustical Society of America*, vol. 1, n° 2A, p. 168, 1930.
- [15] L. L. Beranek, “Analysis of Sabine and Eyring equations and their application to concert hall audience and chair absorption”, *J. Acoust. Soc. Am.*, vol. 120, n° 3, p. 1399–1410, set. 2006.
- [16] A. Carvalho, “The use of the Sabine and Eyring Reverberation Equations to Churches”, 1995.
- [17] M. R. Schroeder, “New Method of Measuring Reverberation Time”, *The Journal of the Acoustical Society of America*, vol. 37, n° 6, p. 1187–1188, 1965.
- [18] Y. Huang, J. Benesty, e J. Chen, *Acoustic MIMO Signal Processing*, Softcover reprint of hardcover 1st ed. 2006. Springer, 2010.
- [19] S. Guy-Bart, E. Jean-Jacques, e D. Archambeau, “Comparison of different impulse response measurement techniques”, *Journal of the Audio Engineering Society*, vol. 50, n° 4, p. 249–262, 2002.
- [20] M. R. Schroeder, “Integrated-impulse method measuring sound decay without using impulses”, *The Journal of the Acoustical Society of America*, vol. 66, p. 497, 1979.
- [21] C. Dunn e M. Hawksford, “Distortion immunity of MLS-derived impulse response measurements”, *J. Aud. Eng. Soc.*, vol. 41, n° 5, p. 314–335, 1993.
- [22] N. Aoshima, “Computer-generated pulse signal applied for sound measurement”, *The Journal of the Acoustical Society of America*, vol. 69, p. 1484, 1981.
- [23] A. F. Dipartimento e A. Farina, “Simultaneous Measurement of Impulse Response and Distortion With a Swept-Sine Technique”, *108TH AES CONVENTION*, vol. 48, p. 18–22, 2000.
- [24] D. Murphy e S. Shelley, “OpenAIR: The Open Acoustic Impulse Response Library”. [Online]. Available: <http://www.openairlib.net/>. [Accessed: 12-out-2011].
- [25] Waves Studio, “Acoustics.Net.” [Online]. Available: <http://www.acoustics.net/>. [Accessed: 13-out-2011].
- [26] J. B. Allen e D. A. Berkley, “Image method for efficiently simulating small-room acoustics”, *J. Acoust. Soc. Am.*, vol. 65, n° 4, p. 943, 1979.
- [27] P. M. Peterson, “Simulating the response of multiple microphones to a single acoustic source in a reverberant room”, *The Journal of the Acoustical Society of America*, vol. 80, p. 1527, 1986.

- [28] J. Eargle, *The microphone book*. Focal Press, 2004.
- [29] S. Haykin, *Adaptive filter theory*, 4th ed. Upper Saddle River N.J.: Prentice Hall, 2002.
- [30] H. L. Van Trees, *Detection, estimation, and modulation theory. Part 4, Optimum array processing*. New York: Wiley-Interscience, 2002.
- [31] G. Marra, “Utilização do algoritmo multi-split LMS em conformadores adaptativos de feixe”, Universidade Federal de Santa Catarina, Florianópolis, 2006.
- [32] L. S. Resende, “Algoritmos recursivos de mínimos quadrados para processamento espacial/temporal com restrições lineares : aplicação em antenas adaptativas”, Universidade Estadual de Campinas, 1996.
- [33] D. Z. Filho, “Algoritmos de processamento espacial para enlaces de comunicação sem fio”, Electronic Thesis or Dissertation, Universidade Estadual de Campinas, Campinas, São Paulo, 2002.
- [34] F. S. Hillier e G. J. Lieberman, *Introduction to Operations Research*, 7^o ed. McGraw-Hill Science/Engineering/Math, 2002.
- [35] A. R. T. Góes, “Otimização na distribuição da carga horária de professores: Método exato, método heurístico, método misto e interface”, Universidade Federal do Paraná, Curitiba, Paraná, 2005.
- [36] W.-H. Steeb, *Problems and solutions in introductory and advanced matrix calculus*. Hackensack, NJ; London: World Scientific, 2006.
- [37] K. Buckley, “Spatial/Spectral filtering with linearly constrained minimum variance beamformers”, *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 35, n^o 3, p. 249–266, mar. 1987.
- [38] L. Griffiths e C. Jim, “An alternative approach to linearly constrained adaptive beamforming”, *IEEE Transactions on Antennas and Propagation*, vol. 30, n^o 1, p. 27–34, jan. 1982.
- [39] P. Delsarte e Y. Genin, “The split Levinson algorithm”, *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, n^o 3, p. 470 – 478, jun. 1986.
- [40] K. C. Ho e P. C. Ching, “Performance analysis of a split-path LMS adaptive filter for AR modeling”, *IEEE Transactions on Signal Processing*, vol. 40, n^o 6, p. 1375 –1382, jun. 1992.

- [41] P. C. Ching e K. F. Wan, “A unified approach to split structure adaptive filtering”, in , *1995 IEEE International Symposium on Circuits and Systems, 1995. ISCAS '95*, 1995, vol. 3, p. 1604 – 1607 vol.3.
- [42] L. S. Resende, R. D. Souza, e M. G. Bellanger, “Multisplit least-mean-square adaptive generalized sidelobe canceller for narrowband beamforming”, in *Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis, 2003. ISPA 2003*, 2003, vol. 2, p. 976 – 980 Vol.2.
- [43] R. D. Souza, L. S. Resende, e M. G. Bellanger, “A justification for the improved performance of the multi-split LMS algorithm”, in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03)*, 2003, vol. 6, p. VI – 89–92 vol.6.
- [44] L. S. Resende, C. A. F. Rocha, J. C. M. Bermudez, e M. G. Bellanger, “A statistical analysis of the multi-split LMS algorithm”, in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04)*, 2004, vol. 2, p. ii – 457–60 vol.2.
- [45] L. S. Resende, J. M. T. Romano, e M. G. Bellanger, “Adaptive split transversal filtering: a linearly-constrained approach”, in *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, 2000, p. 213 –217.
- [46] L. S. Resende, J. M. T. Romano, e M. Bellanger, “Multi-split adaptive filtering”, in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01)*, 2001, vol. 6, p. 3729 –3732 vol.6.
- [47] M. M. Tate, *Principles of Hearing Aid Audiology*. Singular Pub Group, 1994.
- [48] P. C. Loizou, *Speech Enhancement: Theory and Practice*, 1° ed. CRC Press, 2007.
- [49] International Telecommunication Union - ITU, *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*. 2001.
- [50] P. Kabal, “Measuring speech activity”, McGill University, Montreal, Canada, ago. 1997.

- [51] L. L. Beranek, *Acoustics*, Rev Sub. Amer Inst of Physics, 1986.
- [52] W. Soede, A. J. Berkhout, e F. A. Bilzen, “Development of a directional hearing instrument based on array technology”, *J. Acoust. Soc. Am.*, vol. 94, n° 2 Pt 1, p. 785–798, ago. 1993.
- [53] B. Widrow, “A microphone array for hearing aids”, in *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, 2000, p. 7–11.
- [54] B. Widrow e F.-L. Luo, “Microphone arrays for hearing aids: an overview”, *Speech Commun.*, vol. 39, n° 1–2, p. 139–146, jan. 2003.
- [55] R. da S. Maia, “Codificação CELP e Análise Espectral de Voz”, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2000.
- [56] I. Tashev, “Beamformer sensitivity to microphone manufacturing tolerances”. 29-abr-2010.
- [57] I. Tashev, “Gain self-calibration procedure for microphone arrays”, presented at the 2004 IEEE International Conference on Multimedia and Expo (ICME), 2004.
- [58] C. Oberzut e L. Olson, “Directionality and the head-shadow effect”, *The Hearing Journal*, vol. 56, n° 4, abr-2003.
- [59] H. Puder, “Adaptive signal processing for interference cancellation in hearing aids”, *Signal Process.*, vol. 86, n° 6, p. 1239–1253, jun. 2006.
- [60] V. Hamacher, J. Chalupper, J. Eggers, E. Fischer, U. Kornagel, H. Puder, e U. Rass, “Signal processing in high-end hearing aids: state of the art, challenges, and future trends”, *EURASIP J. Appl. Signal Process.*, vol. 2005, p. 2915–2929, jan. 2005.
- [61] M. D. Burkhard e R. M. Sachs, “Anthropometric manikin for acoustic research”, *The Journal of the Acoustical Society of America*, vol. 58, n° 1, p. 214–222, 1975.
- [62] M. W. Hoffman, “Robust microphone array processing for speech enhancement in hearing aids”, University of Minnesota, 1992.
- [63] C. A. Ynoguti, “Reconhecimento de fala contínua usando modelos ocultos de Markov”, Universidade Estadual de Campinas, Campinas, São Paulo, 1999.
- [64] A. Alcain, J. A. Solewicz, e J. A. de Moraes, “Frequência de ocorrência dos fones e listas de frases foneticamente balanceadas no português falado no Rio de Janeiro”, *Revista da Sociedade Brasileira de Telecomunicações*, vol. 7, n° 1, dez-1992.

- [65] J. G. A. Barbedo, *Avaliação objetiva de qualidade de codecs de voz na faixa de telefonia*. UNICAMP, 2001.
- [66] International Telecommunication Union - ITU, *Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems*, vol. 1. 1997.
- [67] Institute of Electrical and Electronics Engineers - IEEE, “IEEE Recommended Practice for Speech Quality Measurements”, *IEEE Transactions on Audio and Electroacoustics*, vol. 17, n° 3, p. 225–246, set. 1969.
- [68] N. R. French e J. C. Steinberg, “Factors Governing the Intelligibility of Speech Sounds”, *The Journal of the Acoustical Society of America*, vol. 19, n° 1, p. 90–119, 1947.
- [69] K. D. Kryter, “Methods for the Calculation and Use of the Articulation Index”, *J. Acoust. Soc. Am.*, vol. 34, n° 11, p. 1689–1697, nov. 1962.
- [70] H. G. Mueller e M. C. Killion, “An easy method for calculating the articulation index”, *The hearing journal*, p. 4, set-1990.
- [71] American National Standards Institute - ANSI, *American National Standards Methods for the Calculation of the Articulation Index*, vol. 5. 1969.
- [72] American National Standards Institute - ANSI, *American National Standards methods for the calculation of the speech intelligibility index*, vol. 5. 1997.
- [73] B. W. Y. Hornsby, “The Speech Intelligibility Index: What is it and what’s it good for?”, out-2004.
- [74] J. M. Kates e K. H. Arehart, “Coherence and the speech intelligibility index”, *J. Acoust. Soc. Am.*, vol. 117, n° 4 Pt 1, p. 2224–2237, abr. 2005.
- [75] H. Puder, “Hearing aids: an overview of the state-of-the-art, challenges, and future trends of an interesting audio signal processing application”, in *Proceedings of 6th International Symposium on Image and Signal Processing and Analysis, 2009. ISPA 2009*, 2009, p. 1–6.