



COPPE/UFRJ

NOVOS ALGORITMOS EM SUBBANDAS PARA FILTRAGEM ADAPTATIVA
E SEPARAÇÃO CEGA DE FONTES

Paulo Bulkool Batalheiro

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia Elétrica.

Orientadora: Mariane Rembold Petraglia

Rio de Janeiro

Julho de 2009

NOVOS ALGORITMOS EM SUBBANDAS PARA FILTRAGEM ADAPTATIVA
E SEPARAÇÃO CEGA DE FONTES

Paulo Bulkool Batalheiro

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Aprovada por:

Prof^ª. Mariane Rembold Petraglia, Ph.D.

Prof. João Marcos Travassos Romano , D.Sc.

Prof. José Carlos Moreira Bermudez, Ph.D.

Prof. José Gabriel Rodriguez Carneiro Gomes, Ph.D.

Prof. Roberto Aizik Tenenbaum, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

JULHO DE 2009

Batalheiro, Paulo Bulkool

Novos Algoritmos em Subbandas para Filtragem Adaptativa e Separação Cega de Fontes/Paulo Bulkool Batalheiro. – Rio de Janeiro: UFRJ/COPPE, 2009.

XX, 138 p.: il.; 29, 7cm.

Orientadora: Mariane Rembold Petraglia

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2009.

Referências Bibliográficas: p. 130 – 138.

1. Filtragem Adaptativa. 2. Análise de Convergência.
3. Erro Quadrático Médio. 4. Processamento Multitaxas.
5. Banco de Filtros Uniforme. 6. Banco de Filtros Não-uniforme. 7. Separação cega de Fontes. 8. Separação cega de Fontes em Subbandas. 9. Misturas Convolutivas. 10. Sinais de Áudio. I. Petraglia, Mariane Rembold. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*“Se eu falar as línguas dos
homens, e dos anjos, e não tiver
caridade, sou como metal que
soa, ou como o sino que tine. E
se eu tiver o dom da profecia, e
conhecer todos os mistérios, e
quando se pode saber: e se tiver
toda a fé, até o ponto de
transportar montes, e não tiver
caridade, não sou nada. A
caridade nunca jamais há de
acabar, ou deixem de ter lugar as
profecias, ou cessem as línguas,
ou seja abolida a ciência.
Quando eu era menino, falava
como menino, julgava como
menino, discorria como menino.
Mas depois que eu cheguei a ser
homem feito, dei de mão as
coisas que eram de menino. Nós
agora vemos Deus como por um
espelho em enigmas, mas então
face a face. Agora conheço-o em
parte, mas então hei de
conhecê-lo, como eu mesmo sou
também dele conhecido. Agora
pois permanecem a fé, a
esperança, a caridade, estas três
virtudes; porém, a maior delas é
a caridade.” 1ª Epístola de
São Paulo aos Coríntios 13,
1.2.9.10-13.*

*À minha esposa Lilian, aos meus
filhos Pedro e Marcos, ao meu
Pai José Augusto (in
memoriam), à minha querida
Mãe Marina, aos meus sogros
Quirino e Benita (in
memoriam). Sem o amor de
vocês eu nada seria...*

Agradecimentos

Gostaria de agradecer a todos que fizeram e que fazem parte desta jornada que culminará, se Deus quiser, com o título de Doutor em Ciências em Engenharia Elétrica. Durante este período em que as pressões são grandes, tive a felicidade de encontrar e contar com pessoas maravilhosas que nunca mais irei esquecer. Particularmente, mesmo correndo o risco de esquecer algumas pessoas, posso citar:

- Minha família e meus familiares, de quem recebi todo apoio e incentivo, nos momentos alegres e nas horas mais difíceis;

- Minha orientadora, professora Mariane, pela sua paciência, tranquilidade, dedicação, capacidade profissional, e principalmente, por ter acreditado em mim;

- A Chefia e os colegas do departamento de Engenharia Eletrônica e Telecomunicações da Faculdade de Engenharia da UERJ, instituição de ensino superior onde trabalho;

- Todos os professores da COPPE responsáveis pelo meu crescimento intelectual e profissional; e todos os funcionários pelo suporte extra classe durante esta jornada;

- O amigo Diego Barreto Haddad, que tem sido um colaborador importante nesta caminhada.

A todos, muito obrigado! E estejam certos de que poderão sempre contar comigo.

O autor e a orientadora gostariam, também, de agradecer o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico da República Federativa do Brasil - CNPq.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

NOVOS ALGORITMOS EM SUBBANDAS PARA FILTRAGEM ADAPTATIVA E SEPARAÇÃO CEGA DE FONTES

Paulo Bulkool Batalheiro

Julho/2009

Orientadora: Mariane Rembold Petraglia

Programa: Engenharia Elétrica

Neste trabalho são investigadas novas estruturas e algoritmos em subbandas para filtragem adaptativa e para separação cega de fontes. Primeiramente é derivada uma estrutura adaptativa não-uniforme maximamente decimada e apresentado um novo algoritmo adaptativo, a partir do qual são realizadas as análises de convergência, do erro quadrático médio em regime estacionário e dos coeficientes ótimos. Em seguida, são propostas duas estruturas e seus respectivos algoritmos em subbandas para o problema de separação cega de fontes no domínio do tempo e envolvendo misturas lineares e convolutivas: uma decompõe as misturas em bandas não-uniformes e realiza a separação de forma *off-line*; a outra utiliza decomposição em bandas uniformes e usa um algoritmo *on-line* para promover a separação. Para tais estruturas são empregados bancos de filtros uniformes e não-uniformes com coeficientes reais e filtros extras para cancelamento da sobreposição do espectro entre subbandas adjacentes, proporcionando redução do erro quadrático médio e aceleração da taxa de convergência dos algoritmos adaptativos na identificação de sistemas em presença de sinal de entrada colorido, e na melhoria da razão sinal-interferência na separação cega de fontes utilizando sinais de áudio.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

NEW SUBBAND ALGORITHMS FOR ADAPTIVE FILTERING
AND BLIND SOURCE SEPARATION

Paulo Bulkool Batalheiro

July/2009

Advisor: Mariane Rembold Petraglia

Department: Electrical Engineering

In this work new structures and subband algorithms for adaptive filtering and blind source separation are investigated. First, a nonuniform adaptive subband structure is derived and a new adaptive algorithm is presented, from which the analysis of the convergence rate, steady-state mean-square error and optimal coefficients are accomplished. Afterwards, two structures and their respective subband algorithms are proposed with the purpose of solving the blind source separation problem in the time-domain for linear and convolutive mixtures: one of them decomposes the mixtures in nonuniform bands and employs an off-line separation method; the other one uses decomposition in uniform bands and an on-line algorithm to perform the separation. For these structures, real-coefficient uniform and non-uniform filter banks with critical sampling, including extra filters to cancel aliasing between adjacent channels, are employed to reduce the mean-square error and improve the convergence behavior of the adaptive subband algorithms for colored input signals in system identification applications, and to improve the signal to noise interference ratio in blind source separation applications with audio signals.

Sumário

Lista de Figuras	xii
Lista de Tabelas	xvi
Lista de Símbolos	xviii
Lista de Abreviaturas	xix
1 Introdução	1
2 Algoritmos Adaptativos	8
2.1 Introdução	8
2.2 Algoritmos para Filtragem Adaptativa	8
2.2.1 Método do Gradiente	9
2.2.2 Algoritmo <i>Least Mean-Square</i> (LMS)	11
2.2.3 Algoritmo LMS Normalizado (NLMS)	13
3 Banco de Filtros e Sistemas Multitaxas	15
3.1 Introdução	15
3.2 Banco de Filtros Uniforme Maximamente Decimado	15
3.2.1 Banco de Filtros Não Maximamente Decimado	17
3.3 Banco de Filtros Estruturado em Árvore	17
3.4 Banco de Filtros Não-Uniforme Maximamente Decimado	19
3.5 Representação Polifásica	21
3.5.1 Banco de Filtros Uniforme	21
3.5.2 Banco de Filtros Não-Uniforme	24
3.6 Banco de Filtros Modulados por Cosseno	25

3.6.1	Erros de Distorção e de <i>Aliasing</i>	26
3.6.2	Protótipos para Banco de Filtros Modulados por Cosseno	27
3.6.2.1	Protótipos do Tipo PQMF	27
3.6.2.2	Protótipos com Reconstrução Quase Perfeita	29
3.6.2.3	Protótipos com Reconstrução Perfeita e Maxima- mente Decimados	30
3.6.2.4	Protótipos com Reconstrução Perfeita sem De- cimação Máxima	32
4	Filtragem Adaptativa em Subbandas Não-Uniformes	33
4.1	Introdução	33
4.2	Estrutura Adaptativa em Subbandas sem Decimação	35
4.3	Estrutura em Subbandas Não-Uniforme Criticamente Decimada	38
4.4	Coefficientes Ótimos	41
4.5	Algoritmo de Adaptação	46
4.6	Análise da Convergência	48
4.7	Erro Quadrático Médio em Regime Estacionário	51
4.8	Complexidade Computacional	54
4.9	Resultados Experimentais	56
4.9.1	Experimento 1	56
4.9.2	Experimento 2	59
4.9.3	Experimento 3	60
4.9.4	Experimento 4	61
4.10	Conclusões	64
5	Separação Cega de Sinais de Áudio	66
5.1	Introdução	66
5.1.1	Separação Cega de Fontes para Misturas Instantâneas e De- terminadas	70
5.1.2	Análise de Componentes Independentes	72
5.2	Separação Cega de Fontes para Misturas Convolutivas	77
5.2.1	Separação Cega de Fontes no Domínio do Tempo	79
5.2.1.1	Função Objetivo e Algoritmo de Atualização	82

5.2.1.2	Complexidade Computacional	85
5.2.2	Separação Cega de Fontes no Domínio da Frequência	86
5.2.2.1	Algoritmo Adaptativo	88
5.2.2.2	Permutação e Escalamento	88
5.3	Medidas de Desempenho	90
5.3.1	Medidas de Desempenho Globais	91
6	Separação Cega de Sinais em Subbandas	94
6.1	Introdução	94
6.2	BSS em Subbandas no Domínio do Tempo <i>Off-line</i>	95
6.2.1	Complexidade Computacional	98
6.2.2	Resultados Experimentais	98
6.2.2.1	Experimento 1	100
6.2.2.2	Experimento 2	101
6.3	BSS em Subbandas no Domínio do Tempo <i>On-line</i>	104
6.3.1	Complexidade Computacional	110
6.4	Resultados Experimentais	111
6.4.1	Experimento 1	111
6.4.2	Experimento 2	115
6.4.3	Experimento 3	116
6.4.4	Experimento 4	118
6.4.5	Experimento 5	120
6.5	Conclusões	126
7	Conclusões e Sugestões	128
7.1	Conclusões	128
7.2	Sugestões	129
	Referências Bibliográficas	130

Lista de Figuras

1.1	Estrutura geral de um filtro adaptativo.	2
1.2	Estrutura geral para filtragem adaptativa em subbandas.	4
1.3	Estrutura geral para separação cega de fontes.	6
2.1	Estrutura geral para identificação de sistemas.	9
3.1	Banco de filtros uniforme e resposta típica dos filtros de análise. . . .	16
3.2	Exemplo de banco de filtros não maximamente decimado.	18
3.3	(a) Exemplo de banco de filtros em árvore maximamente decimado e (b) representação equivalente.	19
3.4	(a) Banco de filtros não-uniforme em árvore maximamente decimado e (b) representação equivalente.	20
3.5	Resposta em frequência de um banco de filtros em oitavas com $M = 4$.	21
3.6	Estrutura multitaxa geral com a representação polifásica dos bancos de análise e síntese.	22
3.7	Estrutura polifásica simplificada	23
3.8	(a) k -ésima subbanda de um banco de filtros não-uniforme com PR; (b) expansão da k -ésima subbanda em l_k subbandas.	24
3.9	Expansão da quarta subbanda de um banco de filtros em árvore com $M = 4$ bandas	25
3.10	(a) Resposta em frequência do filtro protótipo, (b) Resposta em frequência dos filtros de análise.	28
4.1	Estrutura adaptativa utilizando um banco de filtros de análise e sub- filtros adaptativos esparsos.	36

4.2	Estrutura em subbandas para implementação do sistema FIR com função de transferência $S(z)z^{-\Delta}$	38
4.3	Estrutura adaptativa em subbandas não-uniforme sem decimação. . .	39
4.4	k -ésima subbanda após a inclusão do banco de filtros criticamente amostrado.	40
4.5	k -ésima subbanda com filtros adaptativos trabalhando em taxas menores.	41
4.6	Estrutura simplificada com $M = 4$ subbandas aplicada na identificação de sistemas.	42
4.7	(a) k -ésima subbanda da estrutura da Fig. 4.3; (b) k -ésima subbanda com a representação estendida de $G_k(z^{L_k})$	43
4.8	Estrutura adaptativa da Fig. 4.3 com a representação polifásica do banco de filtros não-uniforme.	44
4.9	Erro da k -ésima subbanda.	47
4.10	Resposta em frequência dos filtros de análise da Fig. 4.6 ($H_{k,k}(z)$: linha cheia e $H_{k,k+1}(z)$: linha tracejada).	57
4.11	Evolução do MSE da estrutura em subbandas não-uniforme para o Experimento 1.	59
4.12	Evolução do MSE para diferentes estruturas em subbandas com $M = 4$ para o Experimento 1.	60
4.13	Evolução do MSE da estrutura não-uniforme para o Experimento 2. . .	61
4.14	Evolução do MSE das estruturas uniforme e não-uniforme para o Experimento 3.	62
4.15	Evolução do MSE e valores teóricos do excesso de MSE para o Experimento 4, considerando sinal de entrada branco.	63
4.16	Evolução do MSE e valores teóricos do MSE em excesso para o Experimento 4, considerando sinal de entrada colorido.	63
4.17	Resposta em frequência dos filtros protótipos (linha cheia) e os respectivos MSEs experimentais (linha tracejada), considerando sinal de entrada colorido.	64
5.1	Mistura instantânea com $Q = 2$ fontes e $P = 2$ misturas.	67
5.2	Mistura convolutiva com $Q = 2$ fontes e $P = 2$ misturas.	68

5.3	Esquema geral de um sistema linear MIMO para separação cega de fontes.	70
5.4	Sistema MIMO de P entradas e Q saídas para misturas instantâneas.	71
5.5	Exemplificação do teorema do limite central.	74
5.6	Representação gráfica da informação mútua.	75
5.7	Sistema linear MIMO de P entradas e P saídas para mistura convolutiva.	79
5.8	Esquema geral para separação cega de fontes no domínio da frequência.	87
5.9	Esquema para mensurar SIR_1 em um sistema TITO.	91
5.10	Decomposição da estimativa da q -ésima fonte para cálculo das medidas de desempenho.	92
6.1	Configuração TITO linear para BSS em subbandas.	96
6.2	Cenário virtual usado nos experimentos.	99
6.3	Evolução do SIR para os dois esquemas de normalização: (a) NN e (b) NA.	100
6.4	Resposta em frequência dos filtros de análise $H_{k,k}(z)$ para $M = 4$. .	102
6.5	Evolução do SIR para os algoritmos em banda cheia e em subbandas com filtros de misturas de diferentes comprimentos: $U=256, 512$ e 1024 (de cima para baixo, respectivamente).	103
6.6	Espectro da fonte masculina original , do sinal separado em banda cheia e em subbandas, com filtros de mistura de diferentes comprimentos: $U=256, 512$ e 1024	105
6.7	k -ésimo canal da configuração linear TITO para BSS em subbandas. .	106
6.8	Evolução do SIR para os algoritmos em banda cheia (linha tracejada) e em subbandas (linha cheia) com filtros de mistura de diferentes comprimentos: (a) $U=256$, (b) $U=512$ e (c) $U=1024$	113
6.9	Ocorrência de permutação para os algoritmos em banda cheia ($M=1$) e em subbandas: (b) $M=2$, (c) $M=4$, (d) $M=8$ e (e) $M=16$	114
6.10	Espectrogramas das fontes originais (linha tracejada) e de suas estimativas (linha cheia).	115

6.11	Evolução do SIR bloco a bloco para o algoritmo com $M = 8$ sub- bandas e diferentes sobreposições: $J = \frac{D}{2}$ (linha cheia) e D (linha tracejada).	121
6.12	Evolução dos últimos 250 ms dos sinais dos microfones e das saídas dos algoritmos participantes da competição e do algoritmo proposto para as misturas da Sala 1.	124
6.13	Evolução dos últimos 250 ms dos sinais dos microfones e das saídas dos algoritmos participantes da competição e do algoritmo proposto para as misturas da Sala 2.	125
6.14	Evolução dos últimos 250 ms dos sinais dos microfones e das saídas dos algoritmos participantes da competição e do algoritmo proposto para as misturas da Sala 3.	126

Lista de Tabelas

4.1	Complexidade computacional dos algoritmos das estruturas não-uniforme e uniforme maximamente decimadas.	55
4.2	Parâmetros da estrutura não-uniforme para o Experimento 1.	58
4.3	Razão entre o maior e menor autovalor da matriz $\mathbf{\Gamma}$ para o sinal colorido do Experimento 1.	58
4.4	Relação de autovalores da matriz $\mathbf{\Gamma}$ para o Experimento 2.	60
4.5	MSEs teóricos e experimentais (em dB) para diferentes bancos de filtros do Experimento 4.	62
5.1	Curtose para os exemplos mostrados na Fig. 5.5.	74
6.1	Tempo de processamento em minutos.	101
6.2	Parâmetros da estrutura não-uniforme para $M = 4$ e $U = S = 1024$	101
6.3	SIR Máximo (em dB).	104
6.4	Tempo de processamento (TP) em minutos e número de multiplicações por bloco (NMPB) requerida pelos algoritmos.	104
6.5	Parâmetros para os algoritmos em banda cheia e em subbandas.	112
6.6	SIR final (em dB).	112
6.7	Número de multiplicações por bloco.	112
6.8	Medidas de desempenho globais obtidas pela estrutura em subbandas com $M = 8$ e diferentes filtros protótipos, considerando diversas misturas.	116
6.9	SIR global médio obtido pela estrutura proposta com $M = 4, 8$ e 16 subbandas e subfiltros de separação de mesmo comprimento em todas as bandas.	118

6.10	SIR global médio obtido pela estrutura proposta com $M = 4, 8$ e 16 subbandas e subfiltros de separação de comprimento reduzido somente na banda de mais alta frequência.	118
6.11	SIR global médio obtido pela estrutura proposta com $M = 4, 8$ e 16 subbandas e subfiltros de separação de comprimento reduzido em $\frac{1}{4}$ das bandas de mais alta frequência.	119
6.12	Comprimentos dos subfiltros de separação utilizados nas simulações apresentados na Tabela 6.9.	119
6.13	Comprimentos reduzidos dos subfiltros de separação utilizados nas simulações apresentados na Tabela 6.10.	120
6.14	Comprimentos reduzidos dos subfiltros de separação utilizados nas simulações apresentados na Tabela 6.11.	120
6.15	Parâmetros usados nas simulações das salas 1, 2 e 3.	123

Lista de Símbolos

D	número e atrasos no tempo levados em conta no cálculo da correlação, p. 80
K	número de coeficientes dos filtros adaptativos, p. 38
L_k	fator de decimação na k -ésima subbanda, p. 16
M	número de subbandas, p. 16
N	tamanho do bloco de sinal de saída, p. 80
N_S	ordem do sistema desconhecido, p. 45
N_d	comprimento do sistema desconhecido, p. 38
N_p	comprimento do filtro protótipo, p. 25
P	número de sensores, p. 5
Q	número de fontes, p. 5
S	comprimento dos filtros FIR de separação, p. 69
U	comprimento dos filtros FIR de mistura, p. 69
β	fator de lembrança, p. 48
λ	fator de esquecimento, p. 110
μ	passo de adaptação, p. 10

Lista de Abreviaturas

BSS	<i>Blind Source separation</i> , p. 5
CMFB	<i>Cosine Modulated Filter Bank</i> , p. 25
DCT	<i>Discrete Cosine Transform</i> , p. 25
DSP	<i>Digital Signal Processor</i> , p. 6, 69
FFT	<i>Fast Fourier Transform</i> , p. 56
FIR	<i>Finite Impulse Response</i> , p. 2
HOS	<i>Higher Order Statistics</i> , p. 78
ICA	<i>Independent Component Analysis</i> , p. 67
IDFT	<i>Inverse Discrete Fourier Transform</i> , p. 88
IIR	<i>Infinite Impulse Response</i> , p. 2
IVA	<i>Independent Vector Analysis</i> , p. 122
LMS	<i>Least-Mean-Square</i> , p. 3
LS	<i>Least-Squares</i> , p. 123
MIMO	<i>Multiple Input Multiple Output</i> , p. 5
ML	<i>Maximum Likelihood</i> , p. 74
MMC	Mínimo Múltiplo Comum, p. 24
MSE	<i>Mean-Square-Error</i> , p. 4
ND	<i>Nonstationary Decorrelation</i> , p. 78

NG	<i>Natural Gradient</i> , p. 77
NLMS	<i>Normalized Least Mean-Square</i> , p. 8
NMPA	número de multiplicações reais por amostra do sinal de entrada, p. 54
NMPB	número de multiplicações reais por bloco do sinal de entrada, p. 85
NPR	<i>Near Perfect Reconstruction</i> , p. 26
PQMF	<i>Pseudo Quadrature Mirror Filter</i> , p. 26
PR	<i>Perfect Reconstruction</i> , p. 17
QMF	<i>Quadrature Mirror Filter</i> , p. 17
RLS	<i>Recursive-Least-Square</i> , p. 3
SAR	<i>Signal-to-Artifact Ratio</i> , p. 92
SCA	<i>Sparse Component Analysis</i> , p. 67
SDR	<i>Signal-to-Distortion Ratio</i> , p. 92
SIR	<i>Signal to Interference Ratio</i> , p. 90
SOS	<i>Second Order Statistics</i> , p. 78
STFT	<i>Short-Time Fourier Transform</i> , p. 87
TDD	<i>Time-Delayed Decorrelation</i> , p. 78
TDOA	<i>Time Difference Of Arrival</i> , p. 122
TITO	<i>Two Input Two Output</i> , p. 84
pdf	<i>probability density function</i> , p. 73

Capítulo 1

Introdução

Nas últimas décadas têm havido significantes contribuições na área de processamento digital de sinais devido ao grande avanço tecnológico obtido no desenvolvimento de circuitos digitais. Como consequência, estes sistemas têm se tornado muito atrativos devido à sua precisão, pequena dimensão e flexibilidade. Dentre as várias ferramentas de processamento digital de sinais estão a filtragem adaptativa e a separação cega de fontes.

Filtragem é um processamento que tem por objetivo manipular a informação contida em um determinado sinal; ou seja, um filtro digital é um dispositivo que mapeia o sinal de entrada em outro sinal (normalmente denominado sinal de saída) possibilitando a extração da informação desejada, contida naquele sinal de entrada. Nos filtros invariantes no tempo, os parâmetros internos e a estrutura do filtro são fixos, e se o filtro for linear, cada amostra do sinal de saída será uma combinação linear (possivelmente infinita) de amostras do sinal de entrada.

O filtro adaptativo é um filtro variante no tempo, uma vez que seus parâmetros são ajustados continuamente por um algoritmo de adaptação, o qual almeja otimizar alguma métrica de desempenho. Normalmente, a definição do critério de desempenho requer um sinal de referência, o que usualmente é desconsiderado no projeto de filtros fixos. No projeto de filtros fixos é necessária uma caracterização completa da relação entre os sinais de entrada e saída para projetar o filtro mais apropriado que satisfaça o desempenho desejado. Infelizmente, na prática, isto nem sempre é possível porque as condições dos sistemas e sinais envolvidos não são bem definidos, tampouco bem comportados.

A estrutura geral para filtragem adaptativa supervisionada é apresentada na Fig. 1.1, onde $x(n)$ é o sinal de entrada do filtro adaptativo, $y(n)$ é o sinal de saída, $d(n)$ é o sinal de referência ou sinal desejado, e $e(n)$ é o sinal de erro, que é obtido pela diferença entre os sinais $d(n)$ e $y(n)$, isto é, $e(n) = d(n) - y(n)$.

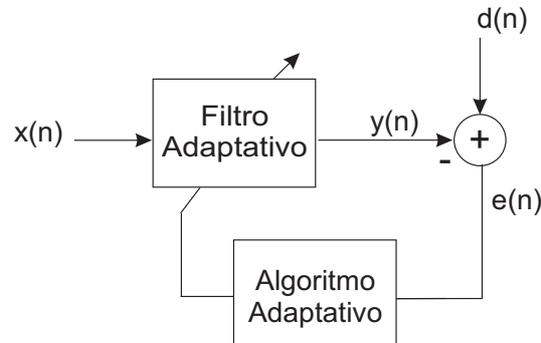


Figura 1.1: Estrutura geral de um filtro adaptativo.

O sinal de erro é utilizado na composição da função objetivo (também denominada função de desempenho ou função custo) que é responsável pela atualização dos coeficientes do filtro adaptativo. A otimização da função objetivo é realizada através da atualização dos coeficientes do filtro adaptativo; com isso, a resposta do filtro se aproxima da resposta desejada, o que implica uma forte correlação entre os sinais $y(n)$ e $d(n)$.

O filtro adaptativo pode ser implementado de maneiras distintas e em estruturas diferentes, que podem influenciar na complexidade computacional (número de operações matemáticas por iteração) e no número de iterações necessárias para alcançar o desempenho desejado. Basicamente existem dois tipos de realizações para os filtros digitais, que se distinguem pela duração das suas respostas ao impulso:

- ↪ Resposta ao Impulso Finita (FIR, *Finite Impulse Response*) ;
- ↪ Resposta ao Impulso Infinita (IIR, *Infinite Impulse Response*) .

Os filtros FIR são normalmente implementados com estruturas não recursivas, enquanto os filtros IIR utilizam implementações recursivas. Devido à existência de pólos não-nulos, os filtros IIR possuem melhores características de transição da faixa de passagem para faixa de rejeição com ordens menores do que os filtros FIR. Por outro lado, os filtros IIR podem se tornar instáveis se os pólos se deslocarem para fora do círculo unitário. Esta característica é particularmente preocupante

em estruturas adaptativas. Ao contrário dos filtros IIR, os filtros FIR adaptativos em geral apresentam funções objetivo com um único ponto de mínimo (mínimo global), facilitando a busca por um melhor desempenho quando a função objetivo é quadrática. Por estas razões a maior parte dos trabalhos, inclusive este, são dedicados a filtros adaptativos FIR.

Diversos algoritmos para atualização dos coeficientes do filtro adaptativo mostrado na Fig. 1.1 foram desenvolvidos nos últimos anos. Podemos citar alguns: o algoritmo LMS (*Least-Mean-Square*) convencional, que, embora possua baixa complexidade computacional, apresenta comportamento da convergência sensível às características do sinal de entrada, acarretando uma convergência lenta para sinais de entrada com forte correlação entre suas amostras (“sinais coloridos”); o algoritmo LMS normalizado, que torna a velocidade de convergência independente da potência do sinal de entrada e possui maior complexidade computacional do que o convencional; o algoritmo LMS no domínio da frequência, que para filtros adaptativos de ordem alta apresenta reduzida complexidade computacional e acelerada convergência para sinais coloridos em relação ao LMS convencional; e o RLS (*Recursive-Least-Square*), que possui alta velocidade de convergência, porém elevada complexidade computacional e, eventualmente, instabilidade numérica.

Recentemente, estruturas adaptativas em subbandas foram propostas com duplo propósito: redução da complexidade computacional e aumento na velocidade de convergência do algoritmo adaptativo para sinais de entrada correlacionados. A redução no esforço computacional é alcançada com a utilização de filtros adaptativos de ordens menores que o filtro adaptativo em banda cheia, os quais operam numa taxa de amostragem menor que a taxa de amostragem do sinal de entrada. Os sinais decompostos pelo banco de análise são menos correlacionadas em cada subbanda que o sinal de entrada em banda cheia, podendo promover melhores taxas de convergência. A Fig. 1.2 mostra a estrutura de um filtro adaptativo em subbandas.

Essencialmente, os sinais de entrada e desejado são separados em subbandas adjacentes por um banco de filtros de análise; em seguida, o sinal de cada subbanda é subamostrado e o algoritmo LMS é aplicado, utilizando o erro de cada subbanda

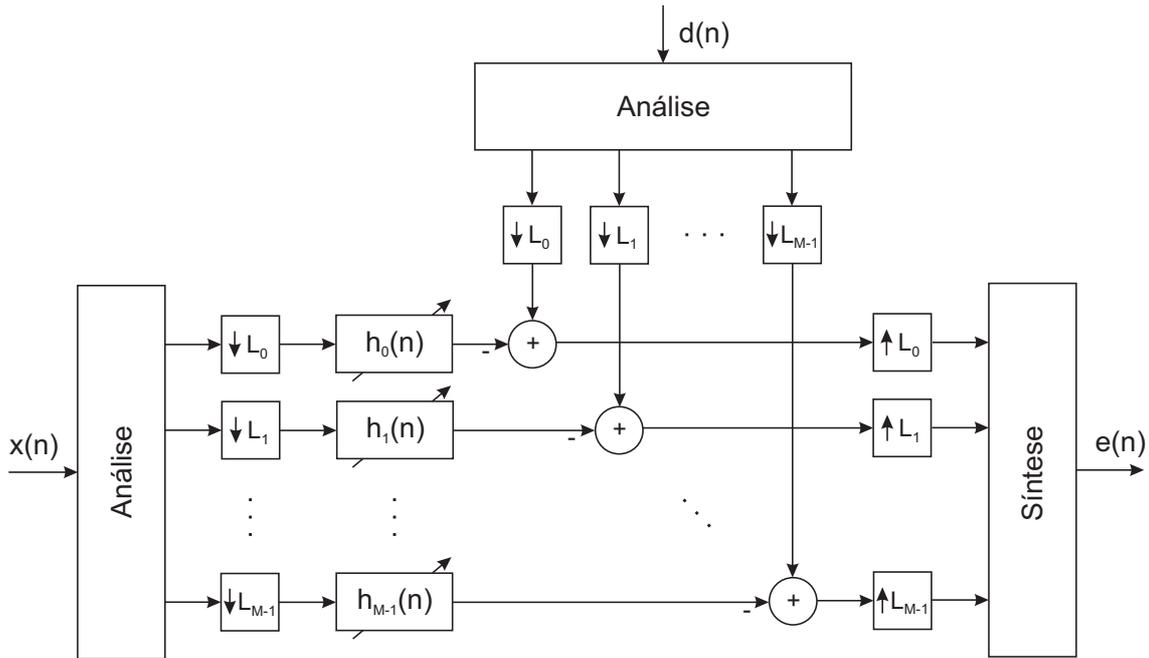


Figura 1.2: Estrutura geral para filtragem adaptativa em subbandas.

para adaptar o respectivo filtro adaptativo¹. A partir daí, os sinais são interpolados e combinados pelo banco de síntese, gerando o sinal de erro da estrutura ($e(n)$), com o qual podemos calcular o MSE (*Mean-Square-Error*) da estrutura. Neste caso, a resposta impulsional identificada será uma composição não-trivial das respostas impulsivas dos subfiltros. A subamostragem reduz a complexidade computacional e o processamento em subbandas aumenta a velocidade de convergência do algoritmo LMS, já que a dispersão entre os valores máximo e mínimo da densidade espectral de potência dos sinais nas diversas subbandas é reduzido, implicando a diminuição do espalhamento dos autovalores da matriz de correlação dos sinais em subbandas [1] ou melhorando o condicionamento da matriz de correlação. Outro fator que contribui para a aceleração da taxa de convergência é a normalização do passo de adaptação de cada subbanda pela energia do sinal de entrada do respectivo filtro adaptativo.

Na primeira parte deste trabalho derivaremos um algoritmo de adaptação em subbandas do tipo gradiente, que emprega um banco de filtros não-uniforme para decomposição do sinal de entrada e passos de adaptação normalizados. Estudaremos

¹Outra possibilidade seria decompor o erro final da estrutura em subbandas e utilizar o erro de cada canal para adaptar os filtros adaptativos.

a convergência do algoritmo resultante e seu erro quadrático médio em excesso. A motivação pela busca de uma nova estrutura não-uniformes deve-se a resultados recentes, os quais demonstraram que estruturas não-uniformes podem ter desempenho superior a estrutura uniformes [2].

Na segunda parte, estudaremos o problema de separação cega de fontes ilustrado na Fig. 1.3, cuja solução tem sido bastante investigada nas últimas décadas. Os sinais $x_p(n)$ capturados pelos P sensores (microfones) são resultantes de misturas lineares e convolutivas entre as Q fontes sonoras $s_q(n)$. Como os sinais de saída são estimativas dos sinais das fontes, obtidas sem nenhum conhecimento das suas posições, de seus conteúdos espectrais e/ou do sistema de mistura, este problema é conhecido como separação cega de fontes (BSS, *Blind Source separation*). São utilizadas hipóteses mínimas que se aplicam a uma grande variedade de configurações, entre elas de que as fontes são independentes entre si e apresentam distribuições não-gaussianas ² (frequentemente supergaussianas). Este sistema é modelado como um sistema MIMO (*Multiple Input Multiple Output*) de Q entradas e P saídas, e cada sinal $x_p(n)$ corresponde ao somatório das fontes convoluídas com as respostas impulsivas definidas desde cada uma das Q fontes até o p -ésimo microfone. Este sistema de mistura é modelado como filtros FIR cujas ordens (tipicamente alguns milhares) dependem do tempo de reverberação do ambiente. O algoritmo adaptativo de aprendizagem pode ser do tipo gradiente descendente [1], porém a função objetivo a ser minimizada leva em conta a informação mútua entre os sinais $x_p(n)$, de forma que os coeficientes do sistema de separação sejam adaptados para que os sinais $y_i(n)$ se tornem mutuamente independentes e correspondam às estimativas das fontes (a menos das ambiguidades de escalamento e permutação).

Recentemente, diversos trabalhos foram publicados tentando resolver este problema no domínio do tempo e/ou no da frequência usando diversas medidas estatísticas. Soluções no domínio do tempo têm um custo computacional muito grande devido às convoluções. Esta complexidade pode ser reduzida no domínio da frequência substituindo as convoluções por produtos de funções de transferência no domínio da transformada. Porém, tais procedimentos acarretam sérios problemas de permutação entre as raias dos diferentes filtros de separação. Para minimizar este

²Admite-se que no máximo uma das fontes apresente distribuição gaussiana

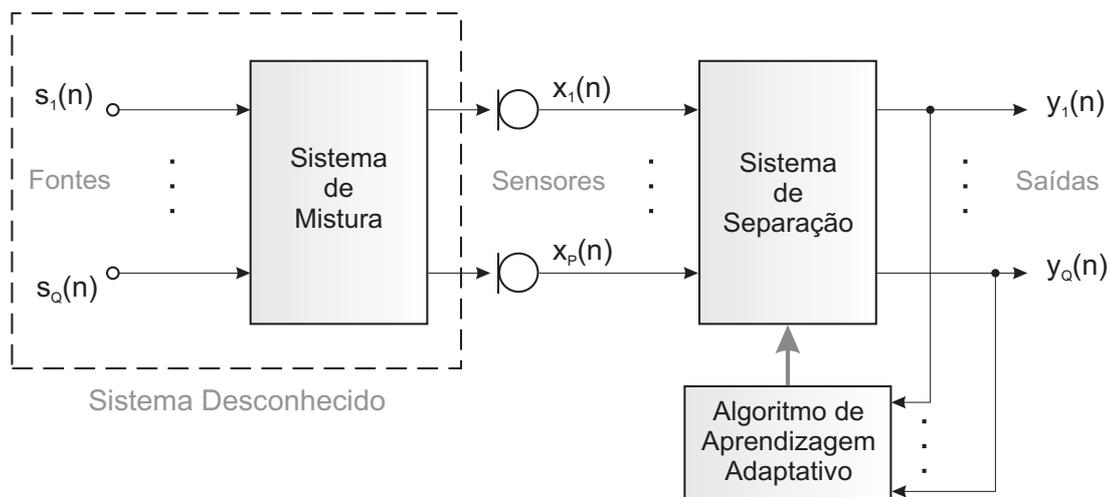


Figura 1.3: Estrutura geral para separação cega de fontes.

problema existem propostas que utilizam algum tipo de informação sobre a localização das fontes. Neste trabalho propomos uma solução intermediária (no domínio do tempo) decompondo os sinais dos microfones em subbandas através de um banco de filtros maximamente decimado com coeficientes reais, possibilitando a redução do número de coeficientes nos sistemas de separação e conseqüentemente o esforço computacional. A idéia é estender o trabalho proposto em [3] para um tratamento em subbandas não-uniformes e uniformes, utilizando estatísticas de segunda ordem e diferentes critérios nos diversos canais.

Nossa proposta é tratar o problema de misturas convolutivas e determinadas no domínio do tempo, decompondo os sinais em subbandas utilizando banco de filtros maximamente decimados com coeficientes reais. Através da decomposição e da decimação dos sinais observados, esperamos reduzir a complexidade computacional da estrutura; outra vantagem é a utilização de filtros com coeficientes reais que são atrativos em implementações em dispositivos para processamento de sinais (DSP, *Digital Signal Processor*). Um desafio será resolver o problema de permutação entre os sinais das diversas subbandas para recompor o sinal estimado das fontes corretamente e sem degradação de desempenho.

Para atingir os objetivos descritos acima, a organização deste trabalho se dará da seguinte forma: no Capítulo 2 será feita uma introdução a algoritmos adaptativos

e a apresentação de alguns algoritmos adaptativos baseados no método do gradiente descendente. No Capítulo 3 serão apresentados conceitos básicos de banco de filtros e sistemas multitaxas, tais como: bancos de filtros uniformes e não-uniformes, maximamente decimados e não maximamente decimados, representação polifásica de sistemas multitaxas, formas de implementação de banco de filtros e erros de reconstrução, além de diferentes métodos de desenvolvimento de filtros protótipos para banco de filtros modulados por cosseno com e sem decimação máxima. No Capítulo 4 será apresentada uma estrutura adaptativa em subbandas não-uniformes sem decimação, a partir da qual será proposta a estrutura não-uniforme maximamente decimada. Para esta última será derivado seu algoritmo de adaptação e analisados sua taxa de convergência e o erro quadrático em excesso em regime permanente, e apresentados os resultados experimentais. No Capítulo 5 serão apresentadas uma introdução para separação cega de fontes, o modelamento deste sistema para os casos de misturas instantâneas e convolutivas, a análise de componentes independentes e algoritmos adaptativos para a BSS. No Capítulo 6 será derivada nossa proposta de implementação de separação de fontes em subbandas, abordando algoritmos *off-line* e *on-line*, e serão mostradas algumas simulações computacionais. Finalmente, no Capítulo 7 serão apresentadas as conclusões e propostas para continuação deste trabalho.

Capítulo 2

Algoritmos Adaptativos

2.1 Introdução

Os algoritmos adaptativos têm-se tornado uma ferramenta muito valiosa devido à sua capacidade de modificar o comportamento de um sistema ao longo do tempo, baseado em algum critério de desempenho. Neste capítulo faremos uma revisão do método de adaptação baseado na técnica do gradiente descendente, que será usado para derivar a estrutura adaptativa não-uniforme apresentada no Capítulo 4 e, também, para obter os algoritmos para separação cega de fontes no Capítulo 5. Em seguida, vamos apresentar os algoritmos *Least Mean-Square* e *Normalized Least Mean-Square* (NLMS) que são algoritmos estocásticos baseados no gradiente [1], e que serão usados para comparação de desempenho com o algoritmo adaptativo proposto no Capítulo 4.

2.2 Algoritmos para Filtragem Adaptativa

Os filtros adaptativos vêm sendo bastante estudados nas últimas décadas devido à sua grande versatilidade e aplicabilidade. Dentre as diversas aplicações, podemos citar por exemplo: identificação de sistemas, realce de sinais, equalização de canais e predição de sinais.

Em aplicações de identificação de sistemas o sinal desejado $d(n)$ é a saída do sistema desconhecido excitado por um sinal de entrada $x(n)$ de banda larga, geralmente considerado um ruído branco [4]; este mesmo sinal é utilizado na entrada

do filtro adaptativo, como mostrado na Fig. 2.1. Quando o erro quadrático médio da estrutura é minimizado, o filtro adaptativo representa o modelo do sistema desconhecido. Considerando que o filtro adaptativo da Fig. 2.1 possui N coeficientes,

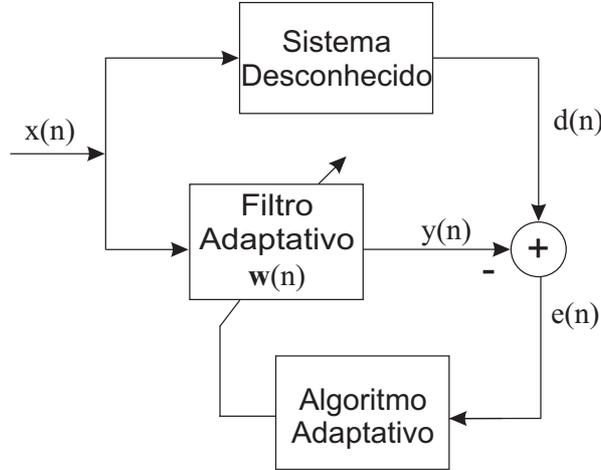


Figura 2.1: Estrutura geral para identificação de sistemas.

podemos definir seus vetores de entrada e de coeficientes, respectivamente, como:

$$\mathbf{x}(n) = [x(n) \quad \dots \quad x(n - (N - 1))]^T \quad (2.1)$$

$$\mathbf{w}(n) = [w_0(n) \quad \dots \quad w_{N-1}(n)]^T \quad (2.2)$$

e, como consequência, o sinal na saída deste filtro será:

$$y(n) = \mathbf{w}^T(n)\mathbf{x}(n). \quad (2.3)$$

2.2.1 Método do Gradiente

Para modificar os coeficientes do filtro adaptativo $\mathbf{w}(n)$, um método numérico amplamente utilizado, devido à sua simplicidade, é o método do gradiente. Cada iteração é baseada na minimização de uma determinada função objetivo $J(n)$. Para filtragem adaptativa supervisionada, esta função é baseada no erro entre um sinal de referência $d(n)$ e o sinal de saída $y(n)$ (ver Fig. 2.1). Para separação cega de fontes esta função objetivo é baseada, por exemplo, na informação mútua na saída do sistema e não utiliza um sinal de referência. O procedimento para encontrar o valor mínimo J_{min} da função objetivo escolhida é o seguinte:

1. arbitrar valores iniciais para os coeficientes do filtro $\mathbf{w}(n)$ contidos no vetor $\mathbf{w}(0)$;
2. usando os valores iniciais, computar o vetor gradiente, formado pelas derivadas parciais de $J(n)$ em relação a cada um dos coeficientes de $\mathbf{w}(n)$;
3. atualizar os valores dos coeficientes do filtro adaptativo na direção contrária ao vetor gradiente;
4. retornar ao passo 2 e repetir o procedimento.

A equação recursiva de atualização dos coeficientes é dada por [1]

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu [-\nabla(J(n))], \quad (2.4)$$

sendo μ uma constante real positiva denominada passo de adaptação, e o vetor gradiente da função objetivo $\nabla(J(n))$ dado por

$$\nabla(J(n)) = \begin{bmatrix} \frac{\partial J(n)}{\partial w_0(n)} \\ \frac{\partial J(n)}{\partial w_1(n)} \\ \vdots \\ \frac{\partial J_n}{\partial w_{M-1}(n)} \end{bmatrix}, \quad (2.5)$$

onde $\partial J(n)/\partial w_i(n)$ é a derivada parcial da função objetivo em relação ao i -ésimo coeficiente do filtro $\mathbf{w}(n)$.

Uma das funções objetivo mais utilizadas em filtragem adaptativa é o erro médio quadrático (*mean-square error* - MSE) definido como

$$J_n = \xi(n) = E[e^2(n)] = E[d^2(n) - 2d(n)y(n) + y^2(n)]. \quad (2.6)$$

Substituindo a Eq. (2.3) na função objetivo acima, temos:

$$\begin{aligned} \xi(n) &= E[d^2(n) - 2d(n)\mathbf{w}^T(n)\mathbf{x}(n) + \mathbf{w}^T(n)\mathbf{x}(n)\mathbf{x}(n)^T\mathbf{w}(n)] \\ &= E[d^2(n)] - 2E[d(n)\mathbf{w}^T(n)\mathbf{x}(n)] + E[\mathbf{w}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}(n)] \\ &= E[d^2(n)] - 2\mathbf{w}(n)^T\mathbf{p} + \mathbf{w}(n)^T\mathbf{R}\mathbf{w}(n), \end{aligned} \quad (2.7)$$

sendo $\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}^T(n)]$ a matriz de autocorrelação do sinal de entrada e $\mathbf{p} = E[d(n)\mathbf{x}(n)]$ o vetor de correlação cruzada entre o sinal desejado e o vetor de entrada.

O vetor gradiente para a função objetivo acima é dado por:

$$\nabla(J(n)) = \frac{\partial \xi(n)}{\partial \mathbf{w}(n)} = -2\mathbf{p} + 2\mathbf{R}\mathbf{w}(n). \quad (2.8)$$

Substituindo este resultado na Eq. (2.4) tem-se a equação recursiva baseada no método do gradiente descendente expressa por:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu [\mathbf{p} - \mathbf{R}\mathbf{w}(n)]. \quad (2.9)$$

Assumindo um filtro com coeficientes fixos, a solução ótima para a Eq. (2.7) é obtida fazendo $\partial\xi(n)/\partial\mathbf{w}(n) = \mathbf{0}$. Esta solução, denominada solução de Wiener-Hopf [1], é dada por:

$$\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p}. \quad (2.10)$$

Porém, na prática, estimativas precisas para \mathbf{R} e \mathbf{p} não estão disponíveis e o vetor gradiente precisa ser estimado a partir dos dados disponíveis. Nesse sentido, serão apresentados, a seguir, dois algoritmos baseados no gradiente estocástico: *Least-Mean-Square* (LMS) e LMS normalizado (NLMS).

2.2.2 Algoritmo *Least Mean-Square* (LMS)

O algoritmo LMS é um dos mais amplamente utilizados em filtragem adaptativa por várias razões, tais como: baixa complexidade computacional, comprovada convergência para meios estacionários, convergência na média para a solução de Wiener e comportamento estável quando implementado com aritmética de precisão finita. Porém, sua velocidade de convergência depende da dispersão dos autovalores da matriz de correlação do vetor de entrada.

No impossibilidade de termos toda a informação necessária para computar precisamente o gradiente (Eq. (2.8)), o LMS busca estocasticamente com os dados de entrada disponíveis uma solução que se aproxime na média da solução ótima, como mostrado a seguir:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu\hat{\mathbf{g}}\mathbf{w}(n), \quad (2.11)$$

onde $\hat{\mathbf{g}}\mathbf{w}(n)$ representa a estimativa do vetor gradiente de (ξ) em relação a $\mathbf{w}(n)$ para um determinado instante. Esta estimativa é obtida derivando-se o valor instantâneo do erro ao quadrado em relação aos elementos de $\mathbf{w}(n)$, ou seja:

$$\begin{aligned} \hat{\mathbf{g}}\mathbf{w}(n) &= -2d(n)\mathbf{x}(n) + 2\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}(n) \\ &= 2\mathbf{x}(n)[-d(n) + \mathbf{x}^T(n)\mathbf{w}(n)] \\ &= -2e(n)\mathbf{x}(n), \end{aligned} \quad (2.12)$$

onde $\mathbf{x}(n)\mathbf{x}^T(n)$ e $d(n)\mathbf{x}(n)$ representam, respectivamente, as estimativas da matriz \mathbf{R} e do vetor \mathbf{p} . Substituindo este resultado na Eq. (2.11) temos:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu e(n)\mathbf{x}(n), \quad (2.13)$$

sendo μ o passo de adaptação que deve ser escolhido de forma a garantir a convergência do algoritmo em ambientes estacionários [4]. A faixa de valores para μ para convergência do algoritmo na média é dada por:

$$0 < \mu < \frac{1}{\lambda_{\max}}, \quad (2.14)$$

onde λ_{\max} é o maior autovalor da matriz \mathbf{R} .

Outras propriedades relativas a algoritmos adaptativos são:

↪ **MSE mínimo** - corresponde ao MSE obtido com a solução ótima \mathbf{w}_o , decorrente da presença de ruído adicional não correlacionado com o sinal de entrada $x(n)$, ou então, devido ao fato do filtro adaptativo não ser um modelo exato do sistema desconhecido;

↪ **Excesso de MSE** - os coeficientes do filtro adaptativo convergem na média para a solução ótima (\mathbf{w}_o), ao passo que instantaneamente ocorre um desvio $\Delta\mathbf{w}(n) = \mathbf{w}(n) - \mathbf{w}_o$ acarretando um aumento (excesso) do MSE que, no caso do LMS, é dado por [4]:

$$\xi = \lim_{n \rightarrow \infty} \Delta\xi(n) \approx \frac{\mu\sigma_v^2 \text{tr}[\mathbf{R}]}{1 - \mu \text{tr}[\mathbf{R}]}. \quad (2.15)$$

Para $\mu \text{tr}[\mathbf{R}] \ll 1$, o excesso de MSE pode ser aproximado para:

$$\xi \approx \mu\sigma_v^2 \text{tr}[\mathbf{R}] = \mu N \sigma_v^2 \sigma_x^2, \quad (2.16)$$

sendo N o comprimento do filtro adaptativo, e σ_x^2 e σ_v^2 , respectivamente, as variâncias do sinal de entrada e do ruído aditivo.

↪ **Desajuste** - é um parâmetro muito utilizado para comparar diferentes algoritmos adaptativos, já que representa a relação entre o excesso de MSE e o MSE mínimo [4]. Para o LMS, é dado por:

$$\mathbf{M}^* \triangleq \frac{\xi_{exc}}{\xi_{min}} = \frac{\mu \text{tr}[\mathbf{R}]}{1 - \mu \text{tr}[\mathbf{R}]}. \quad (2.17)$$

2.2.3 Algoritmo LMS Normalizado (NLMS)

No LMS a correção $e(n)\mathbf{x}(n)$ aplicada ao vetor de coeficientes adaptativas $\mathbf{w}(n)$ é diretamente proporcional ao tamanho de vetor de entrada $\mathbf{x}(n)$. Quando estes vetores são muito grandes, o LMS sofre o problema de amplificação ruído do gradiente [1]. Para superar esta dificuldade foi proposto o algoritmo LMS com passo de adaptação normalizado¹. O algoritmo resultante, o LMS normalizado, utiliza um passo de adaptação variável (μ_n).

Existem estratégias diferentes para derivação do NLMS. Em [4] foi considerada a minimização do erro quadrático instantâneo (EQI) dado por

$$e^2(n) = d^2(n) + \mathbf{w}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}(n) - 2d(n)\mathbf{w}^T(n)\mathbf{x}(n), \quad (2.18)$$

o qual foi considerado como uma forma bem simples de estimar o MSE.

Fazendo uma mudança no vetor de coeficientes do filtro adaptativo do tipo $\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu_n e(n)\mathbf{x}(n) = \mathbf{w}(n) + \Delta\mathbf{w}'(n)$, o erro quadrático correspondente pode ser expresso por

$$\begin{aligned} e'^2(n) &= e^2(n) + 2\Delta\mathbf{w}'^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}(n) + \Delta\mathbf{w}'^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\Delta\mathbf{w}'(n) \\ &\quad - 2d(n)\Delta\mathbf{w}'^T(n)\mathbf{x}(n). \end{aligned} \quad (2.19)$$

Desta forma, pode-se calcular a variação de $e^2(n)$:

$$\begin{aligned} \Delta e^2(n) &\triangleq e'^2(n) - e^2(n) \\ &\quad - 2\Delta\mathbf{w}'^T(n)\mathbf{x}(n)e(n) + \Delta\mathbf{w}'^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\Delta\mathbf{w}'(n). \end{aligned} \quad (2.20)$$

Substituindo $\Delta\mathbf{w}'(n)$ por $2\mu_n e(n)\mathbf{x}(n)$ na equação acima, tem-se:

$$\Delta e^2(n) = -4\mu_n e^2(n)\mathbf{x}^T(n)\mathbf{x}(n) + 4\mu_n^2 e^2(n)[\mathbf{x}^T(n)\mathbf{x}(n)]^2. \quad (2.21)$$

Com o objetivo encontrar o passo μ_n apropriado que maximize a redução do erro quadrático, faz-se $\frac{\partial \Delta e^2(n)}{\partial \mu_n} = 0$, e obtém-se:

$$\mu_n = \frac{1}{2\mathbf{x}^T(n)\mathbf{x}(n)}. \quad (2.22)$$

¹O passo de adaptação normalizado também é chamado de passo de adaptação variável, uma vez que a normalização pode variar a cada iteração.

Como $\frac{\partial^2 \Delta e^2(n)}{\partial^2 \mu_n} > 0$, o valor de μ_n na Eq. (2.22) corresponde a um ponto de mínimo de $\Delta e^2(n)$. Com isso, a fórmula de atualização dos coeficientes do filtro adaptativo é a seguinte:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{e(n)\mathbf{x}(n)}{\mathbf{x}^T(n)\mathbf{x}(n)}. \quad (2.23)$$

Na prática, um passo de adaptação μ_{NLMS} pode ser incluído na equação de atualização para controlar o desajuste, já que todas as derivações feitas são baseadas no erro quadrático instantâneo e não no MSE. Um parâmetro γ também deve ser incluído para controlar o passo de adaptação variável quando o produto $\mathbf{x}^T(n)\mathbf{x}(n)$ torna-se muito pequeno. Desta forma, pode-se reescrever a Eq. (2.23) da seguinte forma:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu_{NLMS}}{\gamma + \mathbf{x}^T(n)\mathbf{x}(n)} e(n)\mathbf{x}(n). \quad (2.24)$$

Para a convergência na média do algoritmo NLMS, o passo de adaptação constante deve satisfazer a seguinte condição [1], [4]:

$$0 < \mu_{NLMS} < 2. \quad (2.25)$$

Capítulo 3

Banco de Filtros e Sistemas Multitaxas

3.1 Introdução

Os bancos de filtros e os sistemas multitaxas são utilizados em diversas aplicações; como exemplo podemos citar: sistemas de áudio digital, codificação e compressão em sinais de voz e imagem, filtragem adaptativa em subbandas, conversão A/D, redução de taxa de amostragem, e multiplexação de sinais em telefonia digital [5].

Neste capítulo apresentaremos alguns conceitos básicos do estudo de bancos de filtros e de sistemas multitaxas, que incluem bancos de filtros uniformes e não-uniformes, maximamente decimados e não maximamente decimados, decomposição polifásica, implementação eficiente de bancos de filtros e projeto de filtros protótipos. Estes tópicos serão importantes para o entendimento dos próximos capítulos e serão apresentados por facilidade de notação no domínio da transformada z^1 .

3.2 Banco de Filtros Uniforme Maximamente Decimado

Como mostrado na Fig. 3.1(a), os conjuntos de filtros $\{H_0(z) \dots H_{M-1}(z)\}$ e $\{F_0(z) \dots F_{M-1}(z)\}$ são denominados de banco de análise e de síntese, respectiva-

¹Transformada z é uma generalização da transformada de Fourier para sinais discretos no tempo muito útil na caracterização sistemas no tempo discreto [6].

mente. Neste sistema o sinal $x(n)$ é decomposto em M sinais através dos filtros de

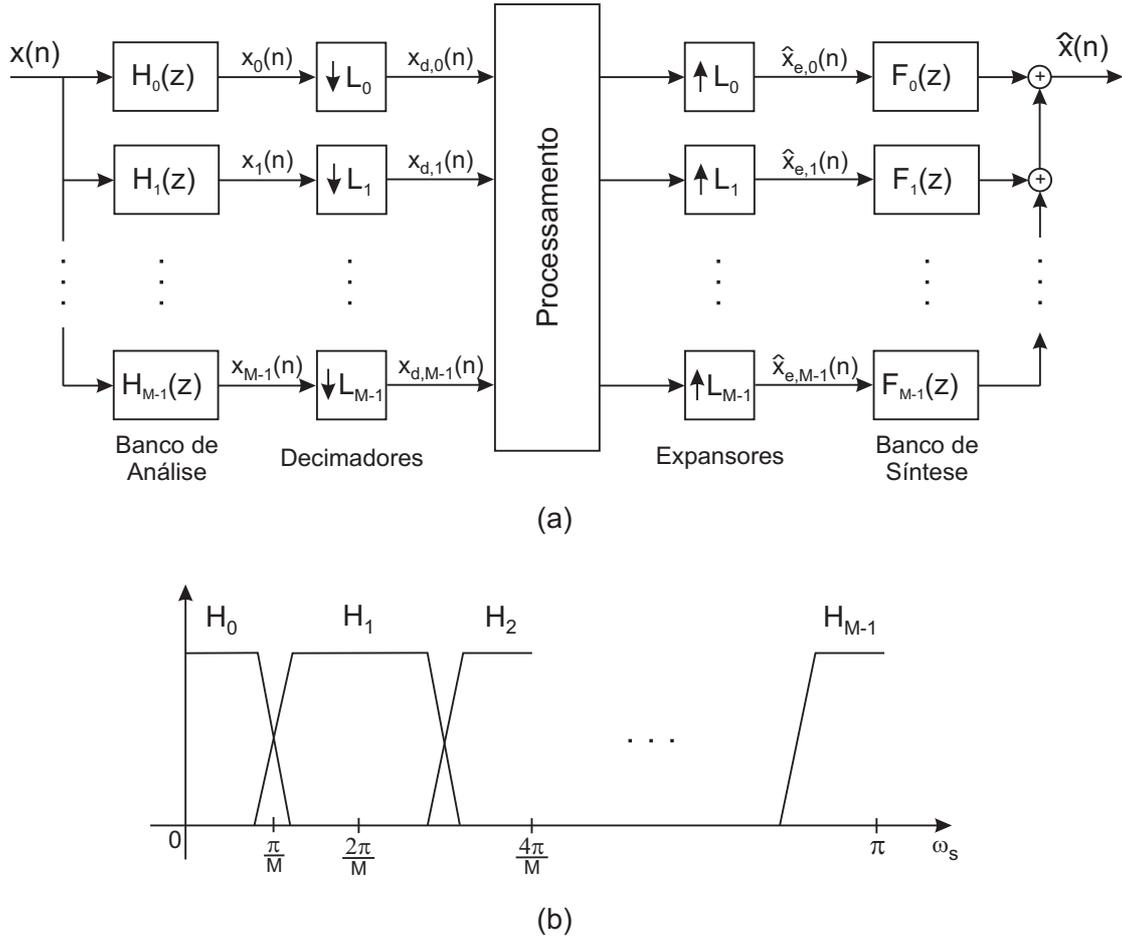


Figura 3.1: Banco de filtros uniforme e resposta típica dos filtros de análise.

análise $H_k(z)$. Estes sinais $x_k(n)$ são em seguida decimados por fatores inteiros L_k , resultando nos sinais decimados $x_{d,k}(n)$. Estes sinais, que agora estão em uma taxa de amostragem L_k vezes menor que a taxa do sinal de entrada, sofrem algum tipo de processamento e em seguida são interpolados, também por um fator L_k , originando os sinais $\hat{x}_{e,k}(n)$. Estes sinais estão novamente na mesma taxa de amostragem do sinal de entrada $x(n)$ e são recombinados pelos filtros de síntese $F_k(z)$. Os filtros de análise $H_k(z)$ têm por finalidade decompor o sinal $x(n)$ em M bandas distintas de frequência, enquanto os filtros de síntese $F_k(z)$ combinam os sinais das M subbandas em um único sinal $\hat{x}(n)$. O banco de filtros é classificado como um banco uniforme quando todos os filtros $H_k(z)$ e $F_k(z)$ têm a mesma largura de faixa de passagem. A Fig. 3.1(b) mostra um esboço da resposta em frequência típica dos filtros de análise de um banco de filtros uniforme. Um banco de filtros é denominado maximamente

ou criticamente decimado quando

$$\sum_{k=0}^{M-1} \frac{1}{L_k} = 1. \quad (3.1)$$

Na reconstrução do sinal na saída do banco de filtros, além da modificação introduzida pelo processamento existe ainda o erro de reconstrução dos bancos de filtros que pode ter características distintas: distorção de amplitude, distorção de fase e sobreposição do espectro (*aliasing*) entre as respostas em frequência dos filtros de análise e de síntese. Felizmente, estes erros podem ser evitados durante o projeto do banco de filtros, levando à reconstrução perfeita do sinal de entrada [5], [7]. Normalmente, o sinal reconstruído $\hat{x}(n)$ é uma versão atrasada do sinal de entrada, ou seja, $\hat{x}(n) = \frac{1}{K}x(n - D)$, onde D é o atraso introduzido pelo banco de filtros e o fator $\frac{1}{K}$ é decorrente do processo de decimação [5]; porém, este último pode ser compensado durante a obtenção dos filtros de análise e síntese.

Historicamente, podemos dizer que a teoria de reconstrução perfeita (PR, *Perfect Reconstruction*) teve início com banco de filtros tipo QMF (*Quadrature Mirror Filter*) [8], mas somente em [9] e [10] demonstrou-se como eliminar os três tipos de distorção para um banco de filtros de dois canais ($M = 2$).

3.2.1 Banco de Filtros Não Maximamente Decimado

Considerando M o número de bandas e L_k o fator de decimação em cada subbanda, dizemos que um banco de filtros é não maximamente decimado quando $\sum_{k=0}^{M-1} \frac{1}{L_k} > 1$. A Fig. 3.2 mostra um banco subamostrado de 4 canais decimado e expandido por um fator $L_k = 2$ para $k = 0, 1, 2$ e 3 (neste caso, $\sum_{k=0}^3 \frac{1}{L_k} = 2$).

Este tipo de banco de filtros tem como característica uma menor sobreposição de espectros na saída dos filtros de análise, além de que suas transições da faixa de passagem para a faixa de rejeição podem ser mais suaves [7].

3.3 Banco de Filtros Estruturado em Árvore

Neste tipo de banco de filtros o sinal de entrada é dividido em duas subbandas por um banco de filtros de análise maximamente decimado de dois canais; em seguida os sinais de cada subbanda são, novamente, divididos e decimados. Cada etapa

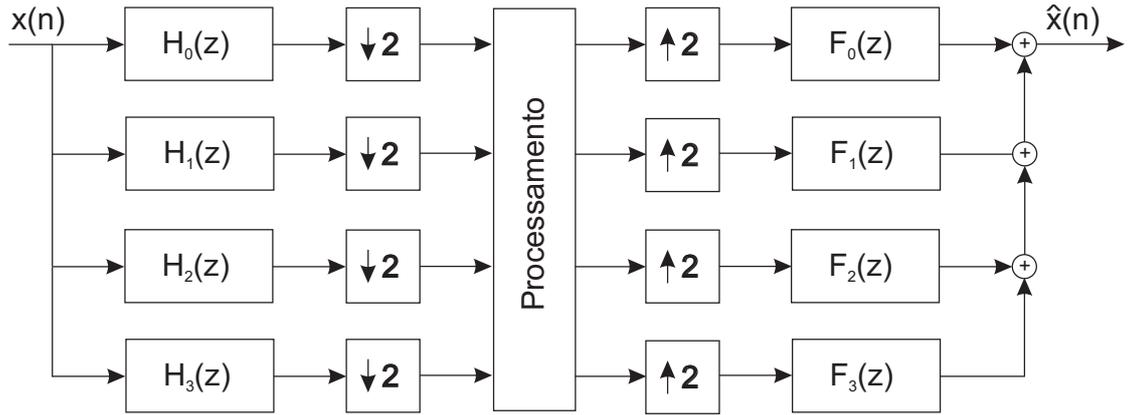


Figura 3.2: Exemplo de banco de filtros não maximamente decimado.

de divisão e decimação é considerada um nível de decomposição e quanto maior o número de níveis maior o número de subbandas do banco de filtros resultante. Estas subbandas são então re combinadas duas a duas através de um banco de síntese de dois canais. Este sistema, maximamente decimado e uniforme, com dois níveis de decomposição é mostrado na Fig. 3.3(a), conhecido também como banco de filtros binário estruturado em árvore [5]. Nesta figura os filtros $[H^{0,0}(z), H^{1,0}(z)]$ e $[F^{0,0}(z)$ e $F^{1,0}(z)]$ compõem os bancos de análise e síntese de um banco de filtro com $M = 2$ canais, respectivamente, onde o primeiro índice sobrescrito igual a 0 representa um subfiltro passa-baixas e igual a 1 um subfiltro passa-altas, enquanto que o segundo índice sobrescrito indica o nível de decomposição da estrutura em árvore. Usando as identidades nobres [5] podemos mover os decimadores do i -ésimo nível do banco de análise para a direita dos filtros $H^{0,i+1}(z)$ e $H^{1,i+1}(z)$ do próximo nível. Análogamente, podemos mover os expansores do i -ésimo nível do banco de síntese para a esquerda dos filtros $F^{0,i+1}(z)$ e $F^{1,i+1}(z)$ do nível seguinte. Com isso, podemos obter a representação equivalente para o banco de filtros binário mostrada na Fig. 3.3(b), onde os $M = 4$ filtros de análise $H_k(z)$ são dados por

$$\begin{aligned} H_0(z) &= H^{0,0}(z)H^{0,1}(z^2), & H_1(z) &= H^{0,0}(z)H^{1,1}(z^2), \\ H_2(z) &= H^{1,0}(z)H^{0,1}(z^2), & H_3(z) &= H^{1,0}(z)H^{1,1}(z^2), \end{aligned} \quad (3.2)$$

e os $F_k(z)$ filtros de síntese são dados por

$$\begin{aligned} F_0(z) &= F^{0,0}(z)F^{0,1}(z^2), & F_1(z) &= F^{0,0}(z)F^{1,1}(z^2), \\ F_2(z) &= F^{1,0}(z)F^{0,1}(z^2), & F_3(z) &= F^{1,0}(z)F^{1,1}(z^2). \end{aligned} \quad (3.3)$$

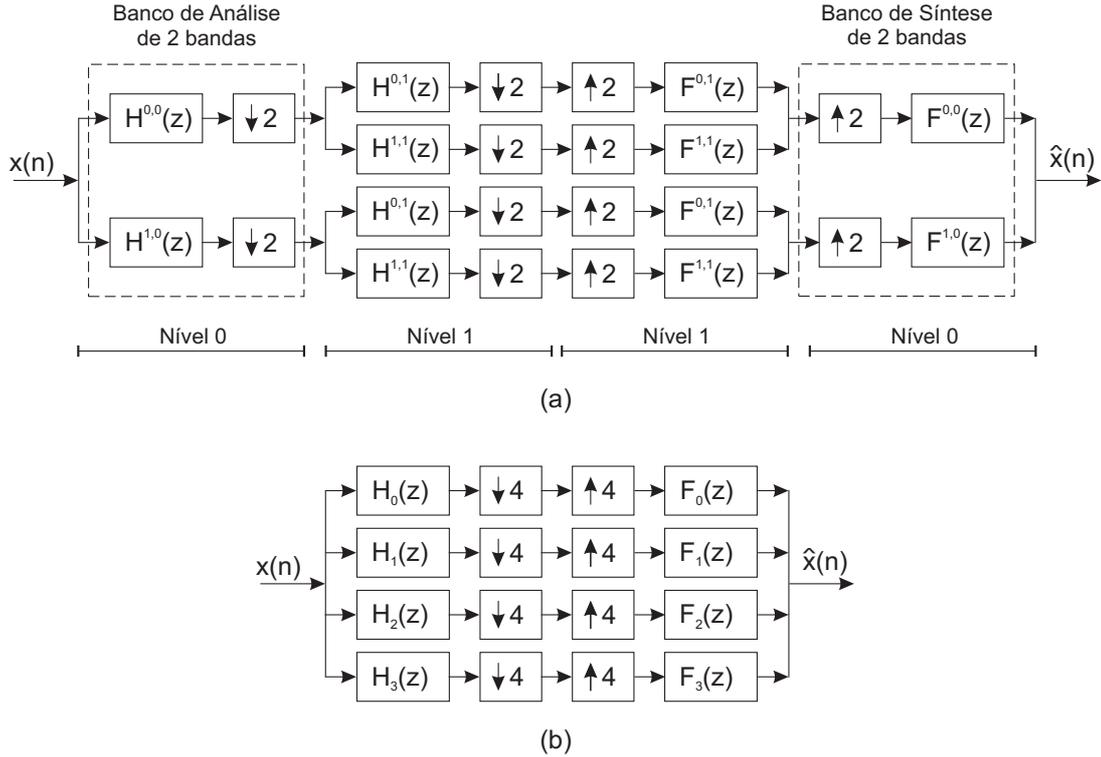


Figura 3.3: (a) Exemplo de banco de filtros em árvore maximamente decimado e (b) representação equivalente.

3.4 Banco de Filtros Não-Uniforme Maximamente Decimado

Um banco de filtros que divide o sinal de entrada em M subbandas que possuem diferentes larguras de bandas de passagem é denominado de banco de filtros não-uniforme. Este tipo de banco de filtros pode ser obtido de diversas maneiras, como por exemplo, a partir de um banco uniforme em que os sinais de suas subbandas são combinados de forma a gerar novas bandas não-uniformes. Uma outra forma é baseada na decomposição binária vista na Seção 3.3.

A Fig. 3.4(a) mostra um banco de filtros não-uniforme binário com $M = 4$ subbandas obtido pela decomposição em árvore de um banco de filtros uniforme e maximamente decimado de 2 canais. Neste caso, são necessários três níveis de decomposição, sendo que cada nível subsequente é obtido pela decomposição, apenas, do sinal decimado na saída do filtro passa-baixas do estágio anterior ($H^{0,i}(z)$). A

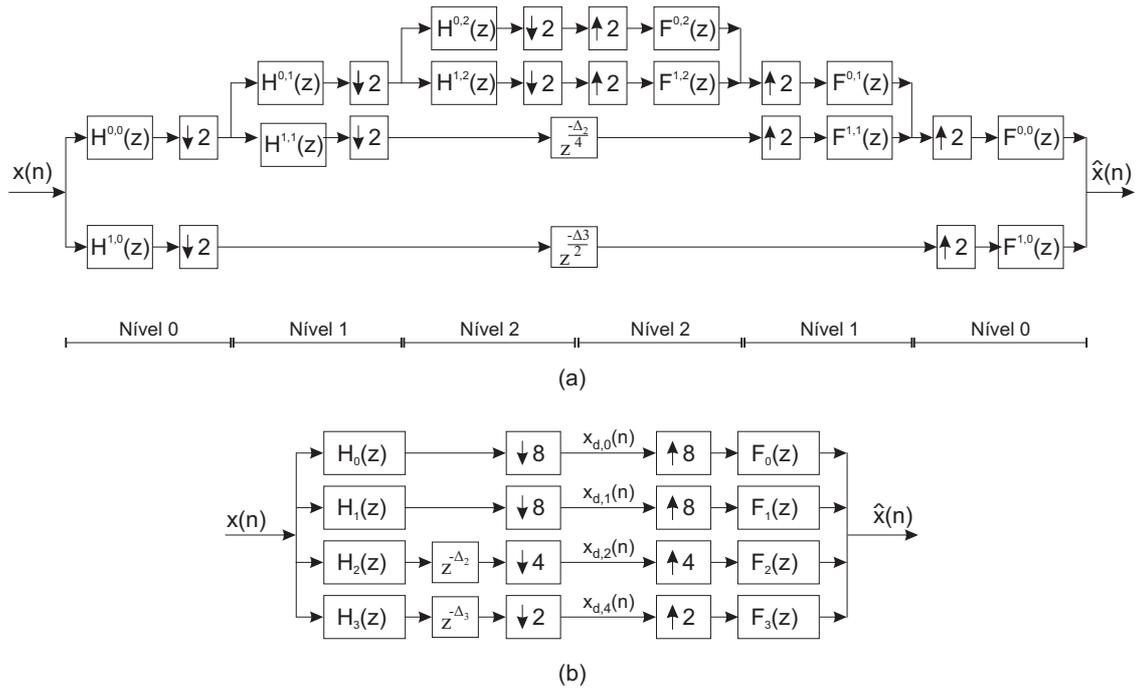


Figura 3.4: (a) Banco de filtros não-uniforme em árvore maximamente decimado e (b) representação equivalente.

representação equivalente da estrutura em árvore (após usadas as identidades nobres) é vista na Fig. 3.4(b), onde os sinais $x_{d,k}(n)$ operam em diferentes taxas de amostragem. De acordo com a Eq. (3.1) o sistema resultante também é criticamente amostrado.

Os M filtros de análise $H_k(z)$ da representação equivalente são dados por

$$H_0(z) = \prod_{j=0}^{M-2} H^{0,j}(z^{2^j}),$$

$$H_k(z) = H^{1,M-1-k}(z^{2^{M-1-k}}) \prod_{j=0}^{M-k-2} H^{0,j}(z^{2^j}), \quad \text{para } k = 1, \dots, M-1, \quad (3.4)$$

e os filtros de síntese $F_k(z)$ são dados por

$$F_0(z) = \prod_{j=0}^{M-2} F^{0,j}(z^{2^j}),$$

$$F_k(z) = F^{1,M-1-k}(z^{2^{M-1-k}}) \prod_{j=0}^{M-k-2} F^{0,j}(z^{2^j}), \quad \text{para } k = 1, \dots, M-1. \quad (3.5)$$

Os atrasos Δ_k são necessários para compensar a diferença de comprimento entre os filtros de análise², e os fatores de decimação e expansão da representação equivalente

²A expressão para Δ_k está descrita no Capítulo 4 (ver Eq. (4.12)).

são obtidos da seguinte forma:

$$\begin{cases} L_0 = 2^{M-1}, \\ L_k = 2^{M-k}, \text{ para } 1 \leq k < M - 1. \end{cases} \quad (3.6)$$

A Fig. 3.5 mostra a resposta em frequência típica dos filtros de análise e síntese do banco de filtros não-uniforme da Fig. 3.4(b). Neste tipo de estrutura cada in-

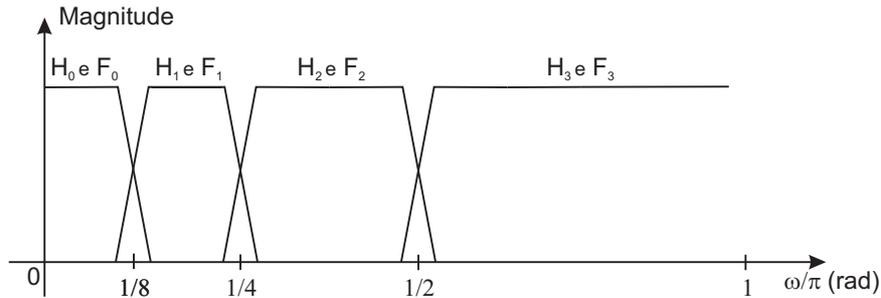


Figura 3.5: Resposta em frequência de um banco de filtros em oitavas com $M = 4$.

cremento de um nível de decomposição implica a redução da largura da banda de passagem à metade da largura correspondente do nível anterior. Por isso, este tipo de estrutura também é conhecida como banco de filtros não-uniforme com decomposição de frequência em oitavas, ou simplesmente, banco de filtros em oitavas.

3.5 Representação Polifásica

Um grande avanço no processamento de sinais multitaxas se deve à representação polifásica dos bancos de análise e síntese. Com isto, foi possível uma grande simplificação da análise teórica dos sistemas multitaxas, e uma maior eficiência computacional tanto na implementação de filtros decimadores e expansores, como na de banco de filtros [5], [7].

3.5.1 Banco de Filtros Uniforme

A Fig. 3.6 mostra a estrutura geral de um sistema multitaxas com bancos de filtros uniformes, usando a representação polifásica tanto do banco de análise como do de síntese, supondo que os fatores de decimação e expansão são todos iguais a L . Decompondo os filtros de análise $H_i(z)$ da seguinte forma:

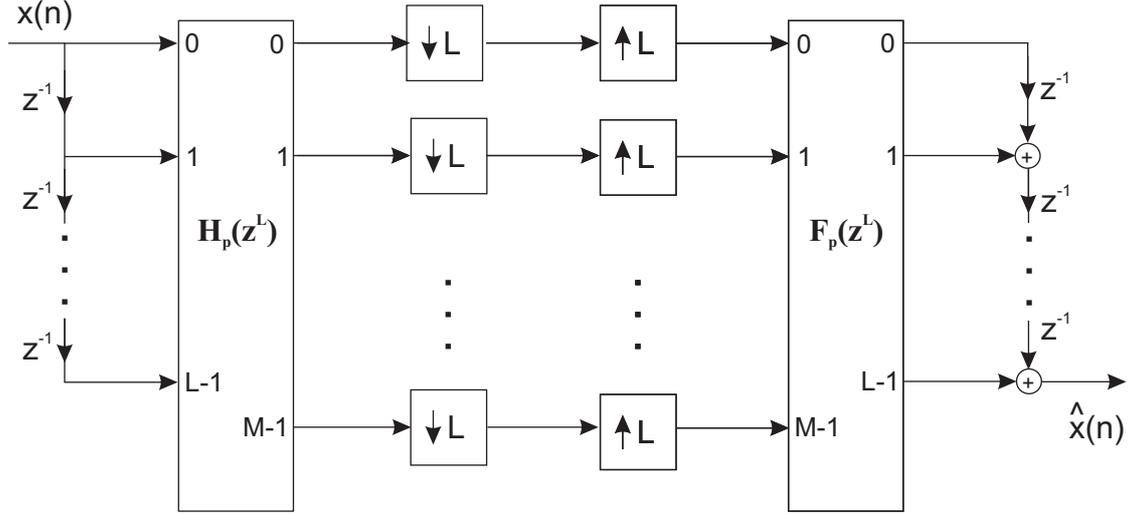


Figura 3.6: Estrutura multitaxa geral com a representação polifásica dos bancos de análise e síntese.

$$H_i(z) = \sum_{j=0}^{L-1} z^{-j} H_{i,j}(z^L), \quad (3.7)$$

onde $H_{i,j}(z)$ são os componentes polifásicos do tipo 1 do i -ésimo filtro de análise³ $h_i(n)$, isto é,

$$H_{i,j}(z) = \sum_{l=-\infty}^{\infty} h_i(lL + j) z^{-l}, \quad (3.8)$$

a matriz polifásica de análise $\mathbf{H}_p(z^L)$ de dimensão $M \times L$ é definida como

$$\mathbf{H}_p(z^L) = \begin{bmatrix} H_{0,0}(z^L) & H_{0,1}(z^L) & \dots & H_{0,L-1}(z^L) \\ H_{1,0}(z^L) & H_{1,1}(z^L) & \dots & H_{1,L-1}(z^L) \\ \vdots & \vdots & \ddots & \vdots \\ H_{M-1,0}(z^L) & H_{M-1,1}(z^L) & \dots & H_{M-1,L-1}(z^L) \end{bmatrix}. \quad (3.9)$$

A representação utilizada para o banco de síntese é a decomposição polifásica do tipo 2 [5], sendo os filtros $F_i(z)$ de síntese dados por:

$$F_i(z) = \sum_{j=0}^{L-1} z^{-(L-1-j)} F_{i,j}(z^L), \quad (3.10)$$

onde $F_{i,j}(z)$ são os componentes polifásicos do i -ésimo filtro de síntese $f_i(n)$, isto é,

$$F_{i,j}(z) = \sum_{l=-\infty}^{\infty} f_i(L(l+1) - j - 1) z^{-l}. \quad (3.11)$$

³Existem dois tipos de decomposição polifásica: tipo 1 e tipo 2. A primeira está associada ao banco de análise e a segunda ao banco de síntese [5], [7].

A matriz polifásica de síntese $\mathbf{F}_p(z^L)$ de dimensão $L \times M$ é definida como

$$\mathbf{F}_p(z^L) = \begin{bmatrix} F_{0,0}(z^L) & F_{1,0}(z^L) & \dots & F_{M-1,0}(z^L) \\ F_{0,1}(z^L) & F_{1,1}(z^L) & \dots & F_{M-1,1}(z^L) \\ \vdots & \vdots & \ddots & \vdots \\ F_{0,L-1}(z^L) & F_{1,L-1}(z^L) & \dots & F_{M-1,L-1}(z^L) \end{bmatrix}. \quad (3.12)$$

Utilizando as identidades nobres [7] podemos simplificar a estrutura da Fig. 3.6 como mostra a Fig. 3.7, onde os filtros de análise e síntese operam numa taxa L vezes menor que o sinal de entrada $x(n)$.

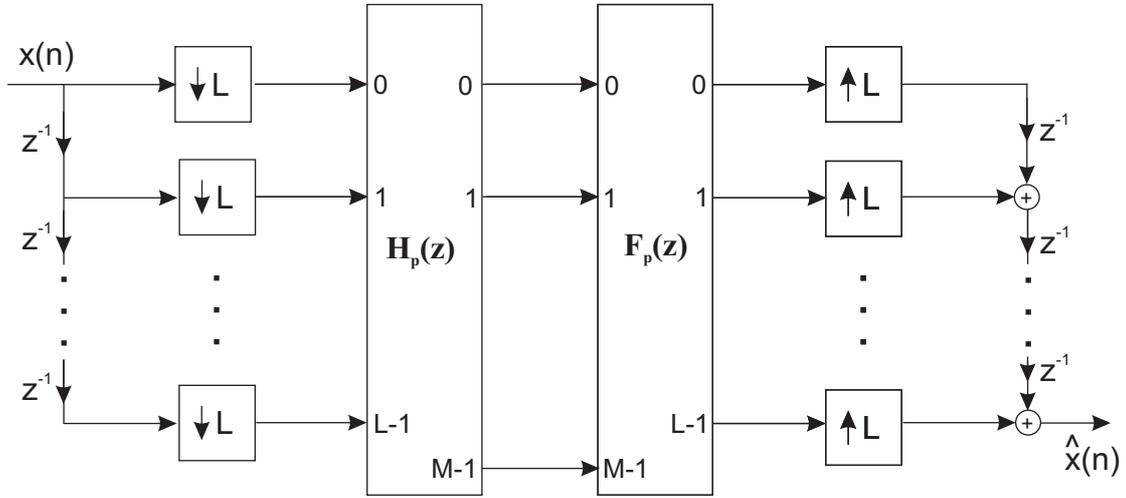


Figura 3.7: Estrutura polifásica simplificada

Para um sistema multitaxa de M bandas, as condições para se obter reconstrução perfeita foram desenvolvidas em [11] e [12] para banco de filtros ortogonais⁴, utilizando a propriedade da paraunitariedade das matrizes polifásicas [5], ou seja:

$$\mathbf{F}_p(z)\mathbf{H}_p(z) = z^{-D_0}\mathbf{I}, \quad (3.13)$$

onde $\mathbf{F}_p(z) = \mathbf{H}_p^T(z^{-1})$, e D_0 é o atraso introduzido no domínio polifásico, sendo o atraso total do sistema $D = L - 1 + D_0L$.

Uma vez que o projeto separado dos filtros de análise e síntese para M bandas ainda se mostrava muito complexo, uma solução alternativa foi obtê-los a partir de

⁴Um banco de filtros é dito ortogonal quando suas matrizes polifásicas são paraunitárias. No domínio do tempo a condição de paraunitariedade é satisfeita quando os coeficientes dos filtros de síntese são reversos no tempo em relação aos coeficientes dos filtros de análise [7].

um único filtro como foi desenvolvido em [13], [14] e [15]. Neste caso, os filtros de análise e síntese são versões moduladas de um único filtro protótipo.

3.5.2 Banco de Filtros Não-Uniforme

A idéia básica é expandir o banco de filtros não-uniforme com M subbandas e diferentes fatores de decimação L_k em um banco de filtros com L_0 bandas e um único fator de decimação N para todos os canais, dado pelo mínimo múltiplo comum (MMC) entre os diferentes fatores L_k . Para o banco de filtros em oitava, mostrado na Seção 3.4, os fatores de decimação e expansão para todas as subbandas são relacionados por

$$N = \text{MMC}(L_0, L_1, \dots, L_{M-1}) = L_0. \quad (3.14)$$

Para atingir este objetivo, a k -ésima banda de análise/síntese da estrutura em árvore não-uniforme com PR mostrada na Fig. 3.8 (a) é, então, estendida em $l_k = L_0/L_k$ subbandas como mostra a Fig. 3.8(b).

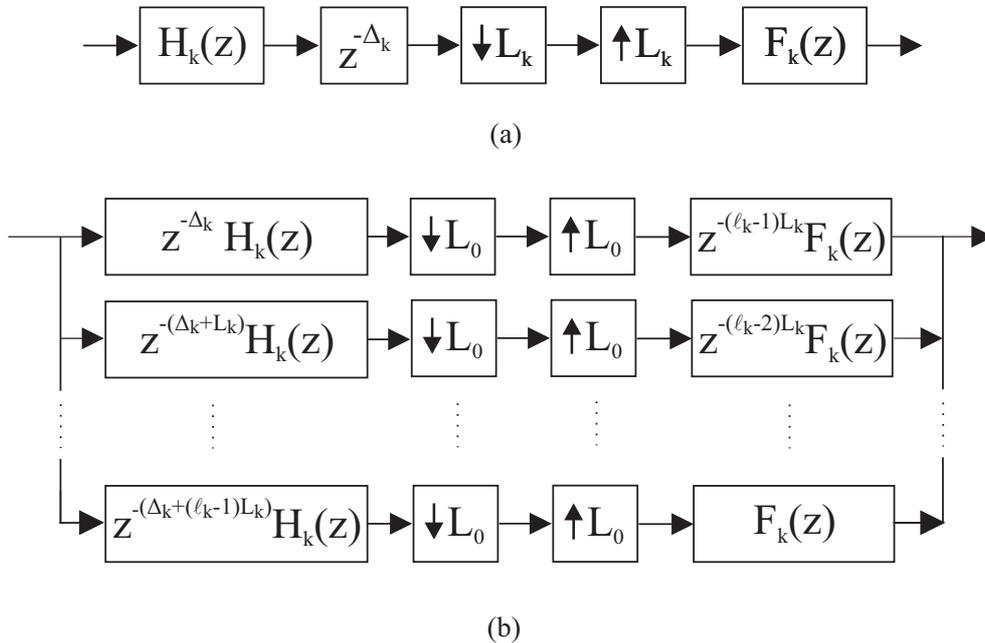


Figura 3.8: (a) k -ésima subbanda de um banco de filtros não-uniforme com PR; (b) expansão da k -ésima subbanda em l_k subbandas.

Como exemplo, é vista na Fig. 3.9 a extensão da quarta subbanda ($k = 3$) do banco de filtros não-uniforme em árvore mostrado na Fig. 3.4(b) com $L_0/L_3 = 4$ novas bandas. Desta forma, um banco de filtros não-uniforme de M canais pode ser

estendido em um banco de L_0 bandas e ser representado na forma geral da Fig. 3.6. A partir desta expansão, a representação polifásica do banco de filtros em oitavas é

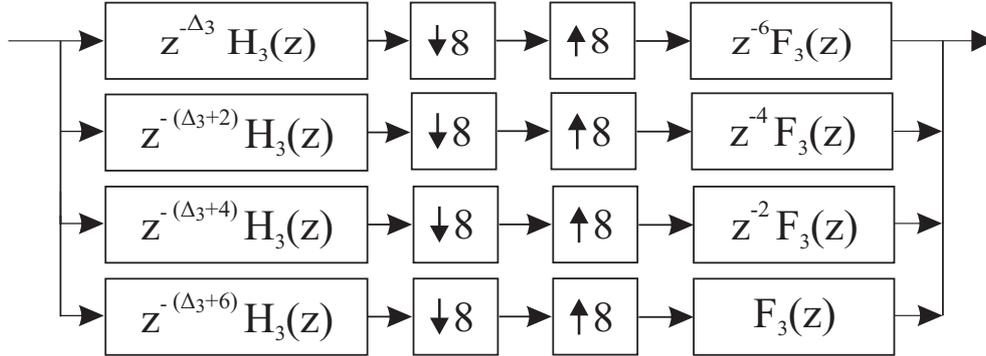


Figura 3.9: Expansão da quarta subbanda de um banco de filtros em árvore com $M = 4$ bandas .

análoga à do banco de filtros uniforme apresentada na Seção 3.5.1.

3.6 Banco de Filtros Modulados por Cosseno

Na literatura são encontradas diversas formas de projeto e implementação de bancos de filtros [5], [7]. Nesta seção vamos descrever apenas o banco de filtros modulado por cosseno (CMFB, *Cosine Modulated Filter Bank*) que será utilizado neste trabalho.

Nestes sistemas, todos os M filtros de análise e os M filtros de síntese são obtidos pela modulação por cosseno de um único filtro protótipo. Algumas extraordinárias vantagens deste sistema são:

- a) O custo computacional para implementação dos bancos de análise e de síntese é igual ao custo computacional de um único filtro e de uma transformada DCT (*Discrete Cosine Transform*);
- b) Durante a fase do projeto do banco de filtros, o número de parâmetros a serem otimizados é reduzido porque somente o filtro protótipo tem que ser otimizado;
- c) Se o protótipo possuir coeficientes reais, todos os filtros dos bancos de análise e de síntese também terão coeficientes reais.

Supondo que o filtro protótipo tenha comprimento N_p e resposta ao impulso $p(n)$, os filtros que irão compor os bancos de análise e síntese são obtidos da seguinte

forma:

$$h_k(n) = 2p(n) \cos \left[\frac{\pi}{L}(k + 0,5) \left(n - \frac{D}{2} \right) + \theta_k \right] \quad (3.15)$$

$$f_k(n) = 2p(n) \cos \left[\frac{\pi}{L}(k + 0,5) \left(n - \frac{D}{2} \right) - \theta_k \right] \quad (3.16)$$

onde $D = N_p - 1$ e $\theta_k = (-1)^k \frac{\pi}{4}$, para $0 \leq k \leq M - 1$ e $0 \leq n \leq N_p - 1$.

Nos últimos anos vêm sendo estudadas algumas classes de CMFB; em ordem cronológica, podemos citar:

i) Banco de filtros pseudo-QMF (PQMF, *Pseudo Quadrature Mirror Filter*), que caracteriza-se por cancelar apenas a sobreposição do espectro entre bandas adjacentes;

ii) Banco de filtros com reconstrução quase perfeita (NPR, *Near Perfect Reconstruction*), que caracteriza-se por cancelar os erros de distorção e de sobreposição do espectro entre bandas adjacentes;

iii) Banco de filtros com reconstrução perfeita (PR), que caracteriza-se por não possuir erros de distorção e de sobreposição do espectro.

O que diferencia estes três tipos de bancos de filtros é o filtro protótipo utilizado, mais precisamente as restrições que a este são impostas durante o seu projeto. Veremos este assunto com mais detalhes na Seção 3.6.2.

3.6.1 Erros de Distorção e de *Aliasing*

Para um banco de filtros de M bandas com fatores de decimação L , a relação entre o sinal de entrada e de saída é [7]:

$$\hat{X}(z) = \sum_{l=0}^{L-1} T_l(z) X(zW_L^l) \quad (3.17)$$

onde

$$T_l(z) = \frac{1}{L} \sum_{k=0}^{M-1} H_k(zW_L^l) F_k(z), \quad (3.18)$$

sendo $W_L = e^{-j\frac{2\pi}{L}}$.

A função de transferência $T_0(z)$ multiplica o espectro original do sinal de entrada e é conhecida como função de distorção. As funções de transferência $T_1(z)$, ..., $T_{L-1}(z)$ multiplicam versões deslocadas do espectro do sinal de entrada e são

conhecidas como funções de transferência de sobreposição de espectro. O objetivo é encontrar um conjunto de filtros ($H_k(z)$ e $F_k(z)$) que garantam a reconstrução perfeita do sinal de entrada; para tal é necessário que:

$$\begin{cases} T_0(z) = z^{-D} \\ T_l(z) = 0, \quad \text{para } 1 \leq l \leq L-1 \end{cases} \quad (3.19)$$

sendo que a sobreposição do espectro entre bandas adjacentes é representado por $T_1(z)$ e entre bandas não adjacentes representado por $T_2(z), \dots, T_{L-1}(z)$.

3.6.2 Protótipos para Banco de Filtros Modulados por Cosseno

Nesta seção serão apresentados quatro tipos de filtros protótipos: Pseudo-QMF (PQMF), com reconstrução quase perfeita (NPR), maximamente decimados com reconstrução perfeita (PR) e sem decimação máxima com PR, a partir dos quais serão implementados os bancos de análise e de síntese modulados por cosseno (Eqs. (3.15) e (3.16)).

Como o cosseno é a soma de duas exponenciais complexas ($\cos(w) = 0,5[e^{jw} + e^{-jw}]$), a resposta em frequência do filtro protótipo é deslocada em duas direções, e se este for um filtro FIR de coeficientes reais e fase linear, todos os filtros do banco de análise e de síntese também terão coeficientes reais e fase linear. A Fig. 3.10 mostra a resposta em frequência do filtro protótipo e dos filtros do banco de análise.

3.6.2.1 Protótipos do Tipo PQMF

Inicialmente foram desenvolvidas técnicas para desenvolvimento de sistemas com reconstrução aproximada. Estes sistemas foram chamados de *Pseudo Quadrature Mirror Filter* (PQMF), inicialmente estudados em [16].

Nestes sistemas, os filtros de análise $H_k(z)$ e de síntese $F_k(z)$ são escolhidos de forma que somente a sobreposição do espectro entre bandas adjacentes é cancelada, e a função de distorção $T_0(z)$ é aproximadamente um atraso. Estes sistemas são aceitáveis em algumas aplicações práticas [5].

Para que o banco de filtros tenha resposta suficientemente plana durante toda a

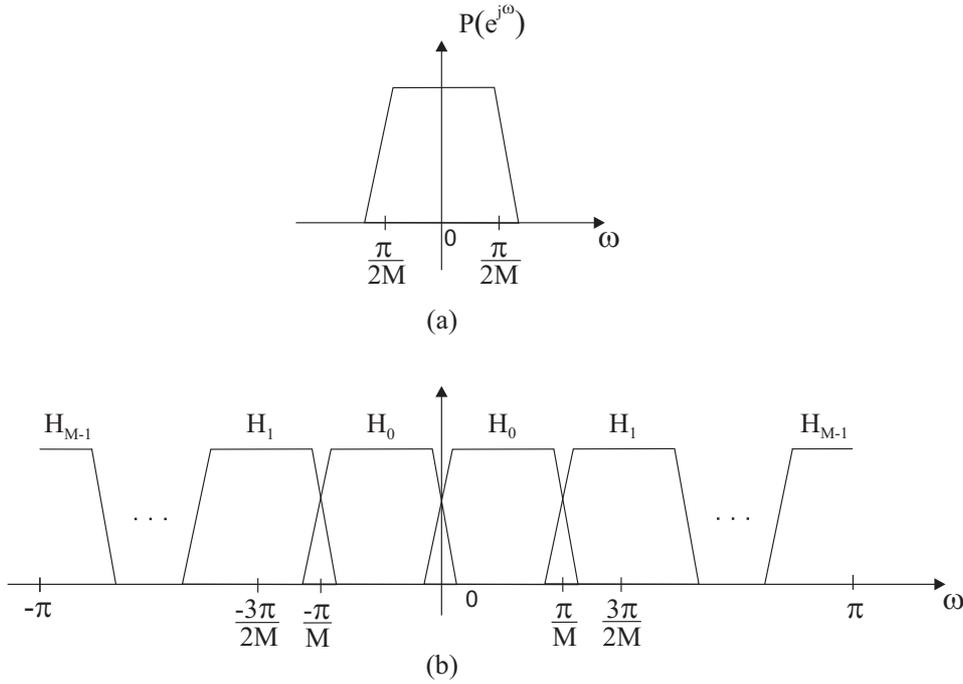


Figura 3.10: (a) Resposta em frequência do filtro protótipo, (b) Resposta em frequência dos filtros de análise.

faixa de frequências digitais $[0, \pi]$, é necessário forçar que o protótipo seja complementar em potência através da função objetivo abaixo:

$$\xi_1 = \sum_{w=0}^{\frac{\pi}{M}} \left(|P(e^{jw})|^2 + |P(e^{j(w-\frac{\pi}{M})})|^2 - 1 \right)^2. \quad (3.20)$$

Para que o banco de filtros tenha boa seletividade e para reduzir a sobreposição do espectro entre bandas não adjacentes é importante que o filtro protótipo tenha grande atenuação na faixa de rejeição. Podemos então minimizar a sua energia na faixa de rejeição através da seguinte função objetivo:

$$\xi_2 = \sum_{w=\frac{\pi}{M}+\gamma}^{\pi} (|P(e^{jw})|)^2, \quad (3.21)$$

onde γ controla a seletividade do protótipo.

Finalmente, podemos otimizar os coeficientes $p(n)$ de $P(z)$ compondo uma função objetivo que combina as funções acima da seguinte forma:

$$\xi = \alpha_1 \xi_1 + \alpha_2 \xi_2, \quad (3.22)$$

onde α_1 e α_2 controlam a importância de cada parcela durante a otimização.

O protótipo $P(z)$ é um filtro FIR de fase linear e coeficientes reais com frequência de corte $\pi/2M$ (Fig. 3.10). Os erros de fase e de sobreposição do espectro entre canais adjacentes são cancelados fazendo $f_k(n) = h_k(Np - n)$ e $\theta_k = (-1)^k \frac{\pi}{4}$ na Eq. (3.15) [5].

3.6.2.2 Protótipos com Reconstrução Quase Perfeita

Um filtro protótipo $P(z)$ de comprimento $N = 2(mM + m_1)$, para m inteiro e $0 \leq m_1 \leq M - 1$, que resulta em reconstrução quase perfeita do sinal de entrada de um banco de filtros modulado por cosseno pode ser obtido através da decomposição espectral de um filtro de $2M$ bandas com fase linear [7]. Os erros de reconstrução são cancelados e o erro de sobreposição do espectro na saída do banco de filtros é comparável à atenuação deste protótipo na faixa de rejeição. Utilizando a notação acima, o filtro de $2M$ bandas é definido da seguinte forma:

$$G(z) = \sum_{n=0}^{4mM+4m_1-2} g(n)z^{-n} \quad (3.23)$$

onde

$$g(n) = \begin{cases} 0, & \text{para } n = 2(mM + m_1) - 1 - 2lM \text{ e } \begin{cases} 1 \leq l \leq m - 1, & m_1 = 0 \\ 1 \leq l \leq m, & m_1 \neq 0 \end{cases} \\ \frac{1}{2M}, & \text{para } n = 2(mM + m_1) - 1. \end{cases} \quad (3.24)$$

Em [17] foram obtidas as restrições quadráticas que garantem que o filtro protótipo corresponde à decomposição espectral do filtro $G(z)$ de $2M$ bandas, tornando possível a obtenção de protótipos com reconstrução quase perfeita através de uma otimização com restrições quadráticas.

Considerando apenas o caso em que M é par, devem ser impostas as seguintes restrições ao protótipo:

$$\begin{cases} \mathbf{p}^T \mathbf{D}_n \mathbf{p} = 0, \text{ para } \begin{cases} \lfloor \frac{m+1}{2} \rfloor \leq l \leq (m-1), & m_1 = 0 \\ \lfloor \frac{m+1}{2} \rfloor \leq l \leq m, & m_1 \neq 0 \end{cases} \\ \mathbf{p}^T (\mathbf{D}_n + \mathbf{J} \mathbf{D}_{n-mM-m_1} + \mathbf{D}_{n-mM-m_1} \mathbf{J}) \mathbf{p} = 0, \text{ para } 1 \leq l \leq \lfloor \frac{m+1}{2} \rfloor - 1 \\ \mathbf{p}^T (\mathbf{J} \mathbf{D}_{mM+m_1-1} + \mathbf{D}_{mM+m_1-1} \mathbf{J}) \mathbf{p} = \frac{1}{2M} \end{cases} \quad (3.25)$$

onde

$$\mathbf{J} = \begin{pmatrix} 0 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 0 \end{pmatrix} \quad (3.26)$$

$$[\mathbf{D}_n]_{i,j} = \begin{cases} 1 & , \text{ para } n = i + j \\ 0 & , \text{ para qualquer outro n.} \end{cases} \quad (3.27)$$

sendo $(mM + m_1) \times (mM + m_1)$ as dimensões das matrizes \mathbf{J} e \mathbf{D}_n , $n = 2M(m - l) + 2m_1 - 1$, e \mathbf{p} o vetor que consiste dos primeiros $mM + m_1$ coeficientes de $p(n)$, ou seja,

$$\mathbf{p} = \left[p(0) \quad p(1) \quad \dots \quad p(mM + m_1 - 1) \right]^T. \quad (3.28)$$

Para obter o filtro protótipo é preciso otimizar sua energia na faixa de rejeição, de acordo com a seguinte função objetivo, atendendo às m restrições impostas na Eq. (3.25):

$$\xi = \sum_{w=w_s}^{\pi} |P(e^{jw})|^2 \quad (3.29)$$

onde $w_s = \frac{\pi}{M}$ (frequência correspondente ao início da faixa de rejeição do filtro protótipo).

3.6.2.3 Protótipos com Reconstrução Perfeita e Maximamente Decimados

Foram obtidas em [18] as condições necessárias e suficientes para os $2M$ componentes polifásicos de um filtro protótipo de comprimento $N = 2mM$ ($m \geq 1$) e fase linear, tal que a matriz dos componentes polifásicos de um banco de filtros modulado por cosseno seja paraunitária [7]. Nesta referência encontramos o seguinte Lema:

Lema: Sendo $\mathbf{h}(z)$ a matriz que representa o banco de filtros de análise obtido a partir de um filtro protótipo $P(z)$ de comprimento $N = 2mM$, fase linear e coeficientes reais, então, a matriz dos componentes polifásicos de $\mathbf{h}(z)$ é paraunitária (sem perdas) se e somente se:

a) Para M par:

$$G_k(z^{-1})G_k(z) + G_{M+k}(z^{-1})G_{M+k}(z) = \frac{1}{2M} \quad (3.30)$$

onde $G_k(z)$ são os componentes polifásicos do tipo 1 de $P(z)$ para $0 \leq k \leq \frac{M}{2} - 1$.

b) Para M ímpar:

$$G_k(z^{-1})G_k(z) + G_{M+k}(z^{-1})G_{M+k}(z) = \frac{1}{2M} \quad (3.31)$$

e

$$2G_{\frac{M-1}{2}}(z^{-1})G_{\frac{M-1}{2}}(z) = \frac{1}{2M} \quad (3.32)$$

para $0 \leq k \leq \lfloor \frac{M}{2} \rfloor - 1$, onde $\lfloor x \rfloor$ representa o maior inteiro menor que x .

Em [19] as restrições acima (Eqs. (3.30), (3.31) e (3.32)) foram reescritas em função dos coeficientes do filtro protótipo ($P(z)$), tornando possível a obtenção de protótipos com reconstrução perfeita (PR) através de uma otimização com restrições quadráticas.

Considerando apenas o caso em que M é par, pode-se reescrever a Eq. (3.30) da seguinte forma:

$$\mathbf{p}^T [\mathbf{V}_k \mathbf{J} \mathbf{D}_n \mathbf{V}_k^T + \mathbf{V}_{M+k} \mathbf{J} \mathbf{D}_n \mathbf{V}_{M+k}^T] \mathbf{p} = \begin{cases} 0 & , 0 \leq n \leq m-2 \\ \frac{1}{2M} & , n = m-1 \end{cases} \quad (3.33)$$

sendo

$$\mathbf{p} = [p(0) \ p(1) \ \cdots \ p(mM-1)]^T \quad (3.34)$$

e

$$[\mathbf{V}_k]_{i,j} = \begin{cases} 1, & \begin{cases} i = k + 2jM, & k + 2jM < mM \\ i = 2M(m-j) - 1 - k, & k + 2jM \geq mM \end{cases} \\ 0, & \text{para qualquer outro } i \end{cases} \quad (3.35)$$

onde $0 \leq k \leq \frac{M}{2} - 1$, $0 \leq n \leq m-1$, e as matrizes \mathbf{J} e \mathbf{D}_n são obtidas de acordo com as Eqs. (3.26) e (3.27), respectivamente.

Levando em conta a simetria do filtro protótipo, as $\frac{M}{2}$ condições de reconstrução perfeita na Eq. (3.30) são reescritas como $\frac{mM}{2}$ restrições quadráticas na Eq. (3.33). Note que as dimensões de \mathbf{p} , \mathbf{V}_k , \mathbf{J} e \mathbf{D}_n são $(mM \times 1)$, $(mM \times m)$, $(m \times m)$ e $(m \times m)$, respectivamente.

Para obter o filtro protótipo é necessário otimizar sua energia na faixa de rejeição, de acordo com a função objetivo mostrada na Eq. (3.29), atendendo às $\frac{mM}{2}$ restrições impostas na Eq. (3.33).

3.6.2.4 Protótipos com Reconstrução Perfeita sem Decimação Máxima

Foram desenvolvidas em [20] as condições necessárias para reconstrução perfeita em bancos de filtros modulados por cosseno sem decimação máxima, para um atraso D arbitrário.

Considerando $P(z)$ o filtro protótipo de comprimento $N = 2mM$ e apenas o caso em que M é par, pode-se escrever as condições PR da seguinte forma:

$$\sum_{l=0}^{2R-1} P_{k+lL}(z)P_{2M-1-k-lL}(z) = \frac{z^{-D_1}}{2L}, \text{ para } k = 0, \dots, \left\lceil \frac{L}{2} \right\rceil - 1 \quad (3.36)$$

onde R é a razão entre o número de bandas M e o fator de decimação L , $P_j(z)$ são os componentes polifásicos do tipo 1 do filtro protótipo, D_1 é o atraso introduzido pelo sistema análise-síntese e $\lceil x \rceil$ representa o menor inteiro maior que x . O atraso no domínio polifásico D_0 é definido como:

$$D_0 = 2RD_1 + 2R - 1 \quad (3.37)$$

com D_1 podendo variar entre 0 e $2m - 2$; porém, o protótipo terá fase linear se $D_1 = m - 1$. Neste caso, o atraso entrada-saída introduzido pelos bancos de análise e de síntese é dado por $D = L - 1 + D_0L$ (Seção 3.5.1).

Para obter o filtro protótipo PR, novamente será necessário otimizar sua energia na faixa de rejeição de acordo com a função objetivo mostrada na Eq. (3.29), atendendo às $2R$ restrições impostas na Eq. (3.36).

Capítulo 4

Filtragem Adaptativa em Subbandas Não-Uniformes

4.1 Introdução

Nos últimos anos, algumas estruturas em subbandas para filtragem adaptativa foram apresentadas com o objetivo de acelerar a convergência para sinais de entrada correlacionados no tempo (sinais coloridos) e, em alguns casos, de reduzir a complexidade computacional, promovendo a adaptação dos coeficientes dos filtros adaptativos em uma taxa de amostragem mais baixa do que a taxa de amostragem do sinal de entrada da estrutura. Em contrapartida, estas estruturas apresentam um atraso entrada-saída e um acréscimo de MSE inerentes a utilização do banco de filtros para decomposição do sinal de entrada, sobreposições do espectro nas diversas subbandas que devem ser contornadas antes de promover a adaptação dos subfiltros, a programação que deve ser implementada para cada canal, entre outras desvantagens.

Em [21] foi proposta uma estrutura com decimação máxima capaz de fazer um modelamento quase exato de sistemas FIR, através da inserção de filtros cruzados e considerando que só existia sobreposição do espectro entre bandas adjacentes. Neste caso, tanto o sinal de entrada quanto o sinal desejado eram decompostos em subbandas, e o erro gerado em cada subbanda era utilizado para atualizar os respectivos filtros adaptativos (direto e cruzados) relacionados à subbanda.

Em [22] foram propostas duas estruturas sem decimação máxima ($L < M$). Com isto os filtros adaptativos eram adaptados numa taxa L vezes menor que o sinal de

entrada; porém, como o efeito da sobreposição do espectro é diretamente proporcional ao fator de decimação, quanto menor o valor de L menor o MSE mínimo da estrutura. Para valores fixos de M e L , pode-se obter um banco de filtros ótimo que minimiza o MSE final da estrutura. A diferença entre as duas estruturas propostas é que em uma o sinal desejado é decomposto em subbandas, enquanto que na outra é o erro final da estrutura que é decomposto.

Duas novas estruturas em subbandas foram propostas em [23], [24]. A primeira utiliza um banco de análise sem decimação, seguido por filtros adaptativos esparsos; já a segunda, que é derivada a partir da primeira, utiliza um banco de filtros maximamente decimado com reconstrução perfeita e os filtros adaptativos operam na taxa reduzida.

A maioria das estruturas adaptativas em subbandas emprega bancos de filtros uniformes. Trabalhos recentes [25], [26] porém, mostraram que estruturas adaptativas em subbandas não-uniformes podem apresentar desempenho superior ao das estruturas uniformes, em termos de velocidade de convergência e/ou erro de modelagem, devido à sua maior flexibilidade. Através da alocação dinâmica das subbandas, que evita componentes de sinal com alta energia nas proximidades dos limites das subbandas [27], ou selecionando as larguras das bandas e os fatores de decimação que minimizam a potência dos componentes polifásicos (de ordens diferentes de zero) do sistema desconhecido [2], a utilização de filtros adaptativos em subbandas não-uniformes pode resultar em uma substancial redução do erro de modelagem quando comparada à utilização de subbandas uniformes. Por outro lado, com algum conhecimento das características espectrais do sinal de entrada, a taxa de convergência pode ser significativamente melhorada usando algoritmos adaptativos em subbandas não-uniformes com passo de adaptação normalizado [26], [25]. As estruturas não-uniformes apresentadas em [25]-[2] empregam sinais em subbandas sem decimação crítica. Em [26] foi derivada uma estrutura não-uniforme criticamente decimada com três subbandas.

Neste trabalho estenderemos os resultados de [26] para um número arbitrário de subbandas, e derivaremos um algoritmo de adaptação do tipo gradiente, que trabalha na menor taxa de amostragem dentre as subbandas envolvidas e emprega um passo de adaptação normalizado, resultando na melhoria da taxa de convergência para

sinais coloridos quando comparado ao algoritmo LMS convencional em banda cheia e duas estruturas em subbandas: uma uniforme e criticamente decimada¹ e outra não-uniforme subamostrada. Como consequência da decomposição não-uniforme em frequência do sinal de entrada, filtros adaptativos distintos trabalham em diferentes taxas, levando a algumas particularidades no algoritmo de adaptação, como o uso de diferentes números de amostras de erro e fatores de convergência com diferentes normalizações na equação de atualização dos coeficientes [28].

4.2 Estrutura Adaptativa em Subbandas sem Decimação

A estrutura adaptativa em subbandas mostrada na Fig. 4.1, que utiliza um banco de filtros de análise e subfiltros esparsos adaptativos, era tida como capaz de modelar somente uma classe particular de sistemas FIR, devido ao comprimento dos filtros de análise ser maior do que o número de coeficientes adaptativos [29]. Porém, é mostrado em [30] que escolhendo apropriadamente o banco de filtros e o número de coeficientes dos subfiltros adaptativos esparsos, a estrutura da Fig. 4.1 é capaz de modelar qualquer sistema FIR com a introdução de um atraso.

Considerando a representação polifásica do banco de análise da estrutura da Fig. 4.1, a matriz polifásica de dimensão $M \times M$ é definida como [7]

$$\mathbf{H}_p(z) = \begin{bmatrix} H_{0,0}(z) & H_{0,1}(z) & \dots & H_{0,M-1}(z) \\ H_{1,0}(z) & H_{1,1}(z) & \dots & H_{1,M-1}(z) \\ \vdots & \vdots & \ddots & \vdots \\ H_{M-1,0}(z) & H_{M-1,1}(z) & \dots & H_{M-1,M-1}(z) \end{bmatrix} \quad (4.1)$$

onde $H_{i,j}(z)$ são os componentes polifásicos do tipo 1 do i -ésimo filtro de análise $H_i(z) = \sum_{n=0}^{N_h-1} h_i(n)z^{-n}$, dadas por

$$H_{i,j}(z) = \sum_{n=0}^{\lceil (N_h-1)/M \rceil - 1} h_i(nM + j)z^{-n}. \quad (4.2)$$

¹A título de esclarecimento as expressões “criticamente decimada” e “maximamente decimada” têm o mesmo significado, assim como as expressões “subamostrada” e “sem decimação crítica”.

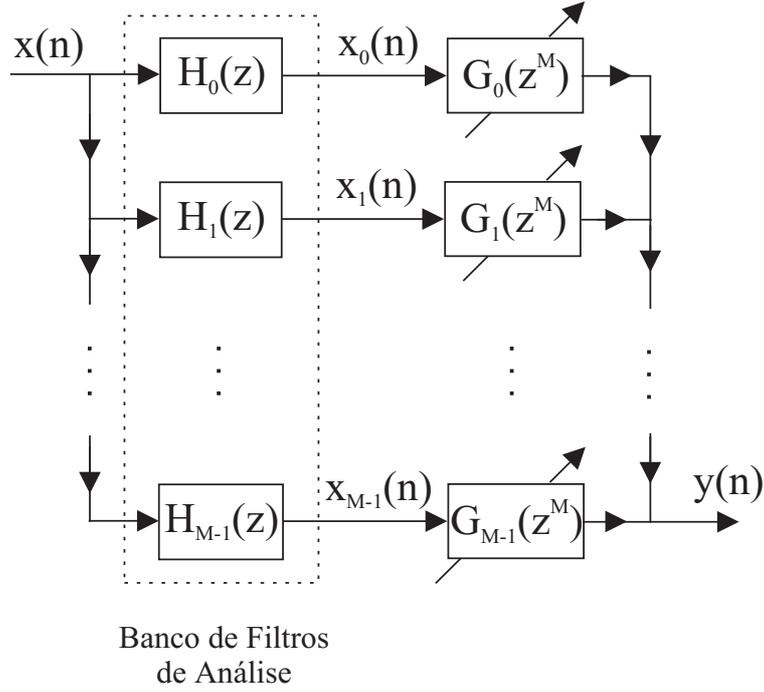


Figura 4.1: Estrutura adaptativa utilizando um banco de filtros de análise e subfiltros adaptativos esparsos.

Portanto, a função de transferência implementada pela estrutura da Fig. 4.1 é

$$H(z) = \begin{bmatrix} G_0(z^M) & G_1(z^M) & \cdots & G_{M-1}(z^M) \end{bmatrix} \mathbf{H}_p(z^M) \begin{bmatrix} 1 \\ z^{-1} \\ \vdots \\ z^{-(M-1)} \end{bmatrix}. \quad (4.3)$$

Em aplicações de identificação de sistemas, os coeficientes dos subfiltros esparsos $G_i(z^M)$ são adaptados para modelar um sistema FIR desconhecido, o qual será denominado $S(z)$. A decomposição polifásica do tipo 1 da função de transferência do sistema desconhecido é dada por

$$S(z) = \begin{bmatrix} S_0(z^M) & S_1(z^M) & \cdots & S_{M-1}(z^M) \end{bmatrix} \begin{bmatrix} 1 \\ z^{-1} \\ \vdots \\ z^{-(M-1)} \end{bmatrix}. \quad (4.4)$$

A partir das Eqs. (4.3) e (4.4), podemos observar que a estrutura em subbandas

modelará exatamente um sistema FIR desconhecido quando

$$\begin{bmatrix} G_0(z^M) & G_1(z^M) & \cdots & G_{M-1}(z^M) \end{bmatrix} \mathbf{H}_p(z^M) = \begin{bmatrix} S_0(z^M) & S_1(z^M) & \cdots & S_{M-1}(z^M) \end{bmatrix}. \quad (4.5)$$

Observando a Eq. (4.5) vemos que a igualdade não pode ser alcançada, porque para subfiltros adaptativos esparsos de comprimento MK e filtros de análise de comprimento N_h , o produto $G_i(z^M)H_{i,j}(z^M)$ tem comprimento $M\lceil N_h/M \rceil + MK - 1$, o qual é maior que o número de coeficientes MK que estão sendo adaptados. Porém, se

$$\begin{bmatrix} G_0(z^M) & G_1(z^M) & \cdots & G_{M-1}(z^M) \end{bmatrix} = \begin{bmatrix} S_0(z^M) & S_1(z^M) & \cdots & S_{M-1}(z^M) \end{bmatrix} \mathbf{F}_p(z^M) \quad (4.6)$$

tal que $\mathbf{F}_p(z^M)\mathbf{H}_p(z^M) = z^{-\Delta}\mathbf{I}$, onde \mathbf{I} é a matriz identidade de dimensão $M \times M$, a função de transferência implementada pela estrutura da Fig. 4.1 será

$$H(z) = S(z)z^{-\Delta}. \quad (4.7)$$

As matrizes $\mathbf{H}_p(z)$ e $\mathbf{F}_p(z)$ que satisfazem as condições acima correspondem, respectivamente, às matrizes polifásicas de análise e de síntese de um banco de filtros com reconstrução perfeita. A matriz polifásica do banco de síntese é definida como

$$\mathbf{F}_p(z) = \begin{bmatrix} F_{0,0}(z) & F_{1,0}(z) & \cdots & F_{M-1,0}(z) \\ F_{0,1}(z) & F_{1,1}(z) & \cdots & F_{M-1,1}(z) \\ \vdots & \vdots & \ddots & \vdots \\ F_{0,M-1}(z) & F_{1,M-1}(z) & \cdots & F_{M-1,M-1}(z) \end{bmatrix} \quad (4.8)$$

onde $F_{i,j}(z)$ são os componentes polifásicos do tipo 2 do i -ésimo filtro de síntese $F_i(z) = \sum_{n=0}^{N_f-1} f_i(n)z^{-n}$ dadas por

$$F_{i,j}(z) = \sum_{n=0}^{\lceil (N_f-1)/M \rceil} f_i(nM - j + M - 1)z^{-n}. \quad (4.9)$$

Então, utilizando um banco de filtros de análise que permita reconstrução perfeita e subfiltros adaptativos esparsos com ordem suficiente para satisfazer a Eq.

(4.6), a estrutura da Fig. 4.1, redesenhada na Fig. 4.2, implementa exatamente o sistema FIR com função de transferência dada na Eq. (4.7). Porém, deve ser enfatizado que o atraso introduzido pelo banco de filtros precisa ser considerado no algoritmo de adaptação dos coeficientes dos subfiltros esparsos.

Para filtros de análise e síntese de fase linear e comprimentos $N_h = N_f = N_p$, o atraso é dado por $\Delta = N_p - M$. O número de coeficientes adaptativos K não nulos dos subfiltros esparsos $G_i(z^M)$ deve ser no mínimo:

$$K = \lceil (N_d + N_p)/M \rceil - 1, \quad (4.10)$$

onde N_d é o comprimento do sistema desconhecido.

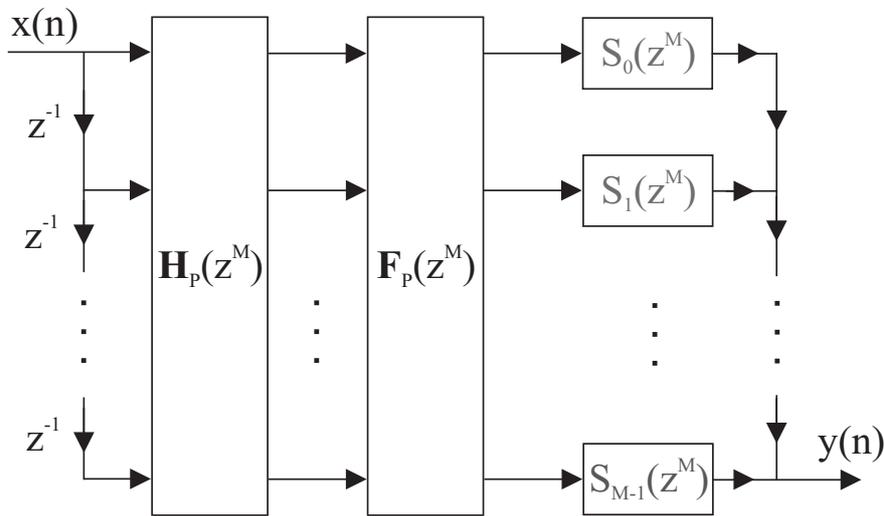


Figura 4.2: Estrutura em subbandas para implementação do sistema FIR com função de transferência $S(z)z^{-\Delta}$.

4.3 Estrutura em Subbandas Não-Uniforme Criticamente Decimada

A estrutura adaptativa não-uniforme em subbandas é derivada a partir da estrutura mostrada na Fig. 4.1, porém empregando um banco de filtros de análise não-uniforme. Esta estrutura é ilustrada na Fig. 4.3, onde $x(n)$ é o sinal de entrada, $H_k(z)$ são os subfiltros de um banco de análise não-uniforme de M canais, $G_k(z^{L_k})$

são os subfiltros adaptativos esparsos, $d(n)$ é o sinal desejado, e $e(n)$ é o sinal de erro utilizado no algoritmo de adaptação.

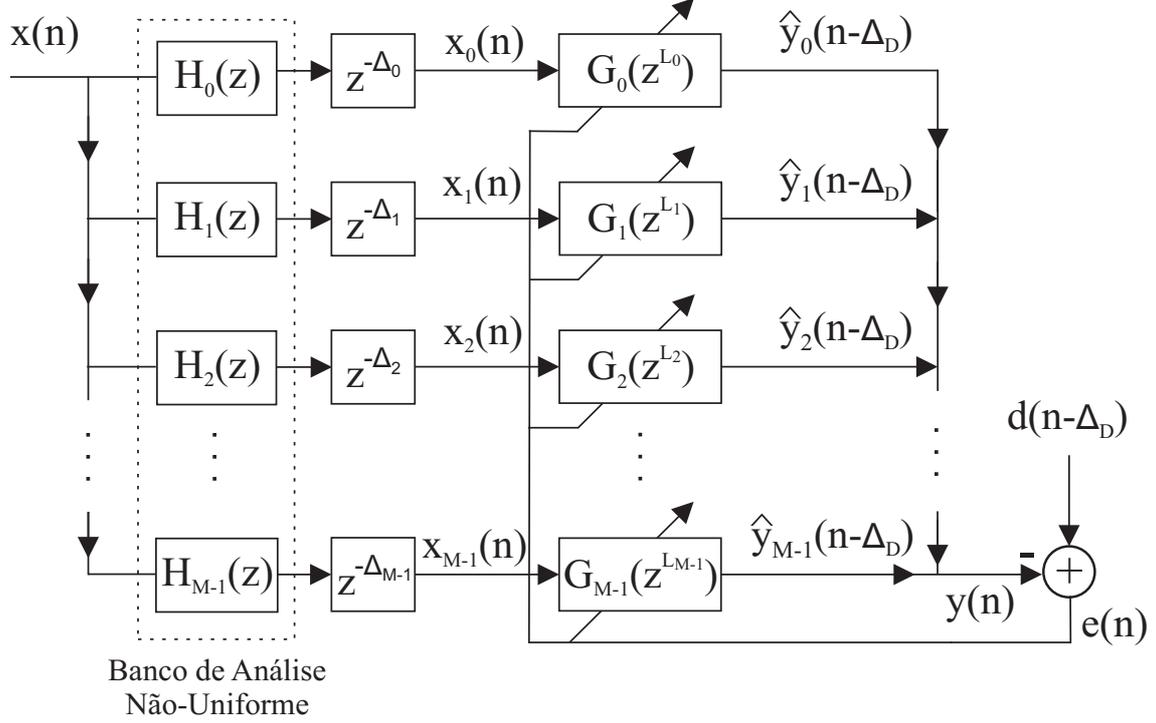


Figura 4.3: Estrutura adaptativa em subbandas não-uniforme sem decimação.

O desenvolvimento da estrutura em subbandas com amostragem crítica proposta e do algoritmo adaptativo correspondente serão feitos para bancos de filtros que utilizam decomposição em frequência por oitavas, os quais são frequentemente empregados na análise de sinais de áudio. Esta escolha foi feita devido às características passa-baixas comumente encontradas nos sistemas a serem identificados em aplicações de cancelamento de ecos acústicos e de equalização de canais. Contudo, os resultados podem ser facilmente estendidos para outros tipos de bancos de filtros não-uniformes.

Para um banco de filtros em oitavas, os filtros de análise equivalentes de um banco de filtros PR de M-canal são obtidos como mostrado na Seção 3.4 (ver Eq. (3.4)). Estes filtros de análise têm ordens

$$\begin{aligned}
 N_{H_0} &= \sum_{j=0}^{M-2} 2^j N_{H^{0,j}}, \\
 N_{H_k} &= \sum_{j=0}^{M-k-2} 2^j N_{H^{0,j}} + 2^{M-k-1} N_{H^{1,j}},
 \end{aligned} \tag{4.11}$$

sendo $N_H^{0,j}$ e $N_H^{1,j}$ as ordens de $H^{0,j}(z)$ e $H^{1,j}(z)$, respectivamente. Os fatores de esparsidade L_k são dados pela Eq. (3.6) e os atrasos Δ_k na Fig. 4.3, que são necessários para compensar a introdução de atrasos diferentes nas subbandas devido às diferenças de comprimento dos filtros que compõem o banco de análise [5], são dados por

$$\Delta_k = N_{H_0} - N_{H_k}. \quad (4.12)$$

O atraso entrada-saída total introduzido pela estrutura é

$$\Delta_D = N_{H_0}. \quad (4.13)$$

Para derivar a estrutura criticamente decimada, um sistema multitaxas não-uniforme com reconstrução perfeita com filtros de análise $H_k(z)$ e filtros de síntese $F_k(z)$ é incluído após cada sub-filtro adaptativo da Fig. 4.3. A Fig. 4.4 mostra a k -ésima subbanda da estrutura resultante, a qual possibilitará, como veremos a seguir, que os subfiltros operem numa menor taxa de amostragem.

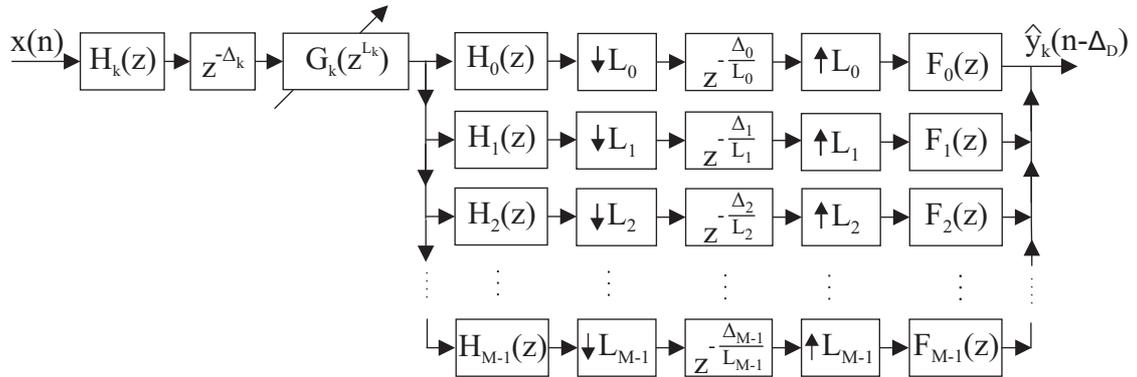


Figura 4.4: k -ésima subbanda após a inclusão do banco de filtros criticamente amostrado.

Para se obter uma estrutura com menor complexidade consideramos que os filtros de análise são suficientemente seletivos para assumir que existe sobreposição do espectro somente entre as respostas em frequência das subbandas adjacentes, a k -ésima banda ilustrada na Fig. 4.4 pode ser simplificada como mostrado na Fig. 4.5, onde $H_{i,j}(z) = H_i(z)H_j(z)$ e os subfiltros esparsos $G_k(z^{L_k})$ foram movidos para a direita dos decimadores L_k , de acordo com uma das identidades nobres [5]. Observando a k -ésima banda da estrutura simplificada mostrada na Fig. 4.5, percebe-se

que os filtros adaptativos trabalham numa taxa de amostragem que é L_k ou L_{k+1} vezes menor que a taxa do sinal de entrada.

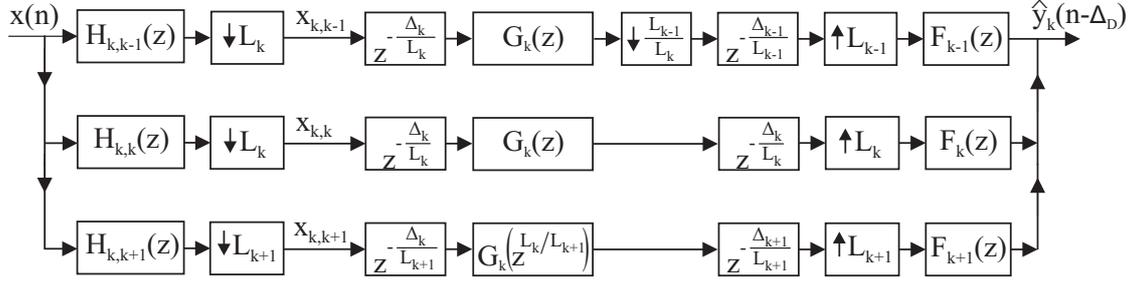


Figura 4.5: k -ésima subbanda com filtros adaptativos trabalhando em taxas menores.

A estrutura poder ser ainda mais simplificada observando que $H_{i,j}(z) = H_{j,i}(z)$ e combinando os sinais das bandas adjacentes que são filtrados pelo mesmo filtro de síntese $F_k(z)$. A Fig. 4.6 mostra a estrutura completa para $M = 4$ subbandas aplicada à identificação de sistemas, onde $v(n)$ é um ruído de medição presente no sinal desejado $d(n)$. Esta figura ilustra os erros de cada subbanda $e_k(m)$, que serão usados na Seção 4.5 para derivação do algoritmo de adaptação.

4.4 Coeficientes Ótimos

De acordo com a seção anterior, assumindo que são utilizados filtros de análise bem seletivos (tal que exista sobreposição de espectro somente entre subbandas adjacentes), os coeficientes da estrutura com decimação crítica da Fig. 4.5 são idênticos àqueles da estrutura esparsa da Fig. 4.3. A partir desta hipótese, nesta seção derivaremos uma expressão para os coeficientes ótimos da estrutura não-uniforme criticamente decimada supondo o caso de modelagem de um sistema FIR arbitrário.

Os filtros adaptativos $G_k(z^{L_k})$ de cada subbanda da Fig. 4.3 podem ser representados em termos de versões atrasadas de seus $l_k = L_0/L_k$ componentes polifásicos [5], como ilustrado na Fig. 4.7, resultando em l_k filtros de esparsidade L_0 .

Definindo como $\mathbf{G}_k(z)$ o vetor que contém os l_k componentes polifásicos do k -ésimo subfiltro, dado por

$$\mathbf{G}_k(z) = \begin{bmatrix} G_{k,0}(z) & G_{k,1}(z) & \cdots & G_{k,l_k-1}(z) \end{bmatrix}^T, \quad (4.14)$$

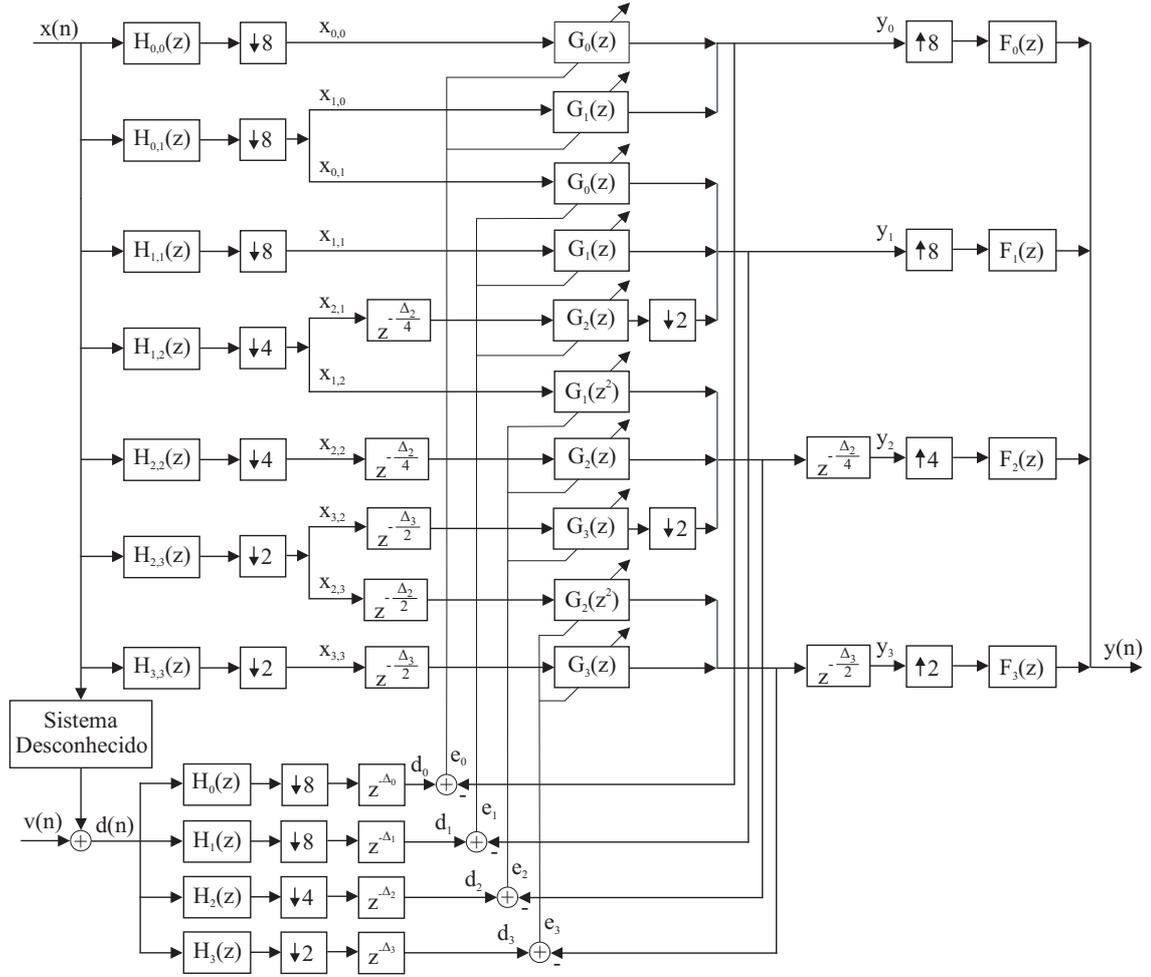


Figura 4.6: Estrutura simplificada com $M = 4$ subbandas aplicada na identificação de sistemas.

e $\widehat{\mathbf{G}}(z)$ o vetor formado pelos vetores polifásicos dos M subfiltros $G_k(z)$, dado por

$$\begin{aligned} \widehat{\mathbf{G}}(z) &= \left[\widehat{G}_0(z) \quad \widehat{G}_1(z) \quad \cdots \quad \widehat{G}_{L_0-1}(z) \right]^T \\ &= \left[\mathbf{G}_0^T(z) \quad \mathbf{G}_1^T(z) \quad \cdots \quad \mathbf{G}_{M-1}^T(z) \right]^T, \end{aligned} \quad (4.15)$$

a estrutura em subbandas da Fig. 4.3 pode ser redesenhada em termos dos L_0 componentes de $\widehat{\mathbf{G}}(z)$, como mostrado na Fig. 4.8. Estes componentes são relacionadas com os componentes polifásicos dos subfiltros $G_k(z)$ pelas seguintes relações: $\widehat{G}_0(z) = G_{0,0}(z)$ e $\widehat{G}_{l_k+i}(z) = G_{k,i}(z)$, para $k = 1, \dots, M-1$ e $i = 0, \dots, l_k-1$. Por exemplo, considerando a estrutura não-uniforme em subbandas com $M = 4$ (ilustrada na Fig. 4.6), teremos: $\widehat{G}_0(z) = G_{0,0}(z)$, $\widehat{G}_1(z) = G_{1,0}(z)$, $\widehat{G}_2(z) = G_{2,0}(z)$, $\widehat{G}_3(z) = G_{2,1}(z)$, $\widehat{G}_4(z) = G_{3,0}(z)$, $\widehat{G}_5(z) = G_{3,1}(z)$, $\widehat{G}_6(z) = G_{3,2}(z)$, e $\widehat{G}_7(z) = G_{3,3}(z)$.

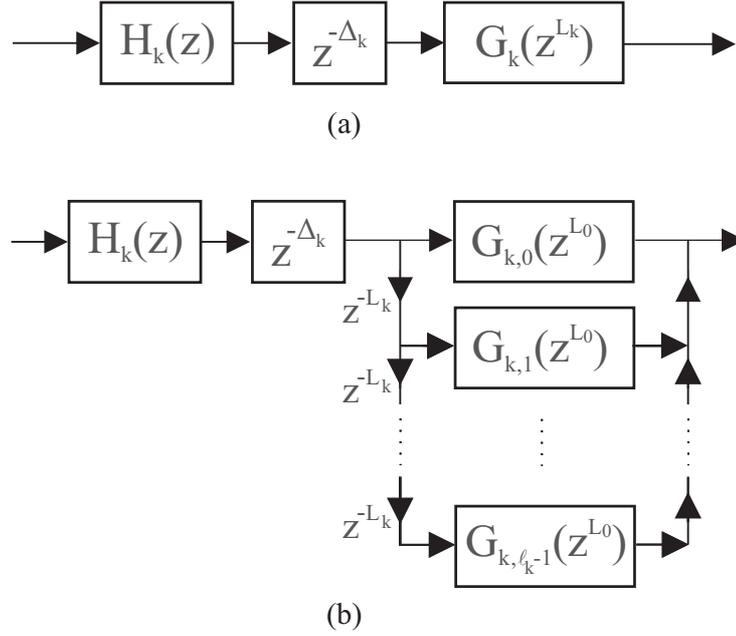


Figura 4.7: (a) k -ésima subbanda da estrutura da Fig. 4.3; (b) k -ésima subbanda com a representação estendida de $G_k(z^{L_k})$.

Na Fig. 4.8, $\mathbf{H}_p(z)$ é a matriz de dimensão $L_0 \times L_0$ que contém os componentes polifásicos do tipo 1 dos filtros de análise, dada por

$$\mathbf{H}_p(z) = \begin{bmatrix} \mathbf{H}_0^T(z) & \mathbf{H}_1^T(z) & \cdots & \mathbf{H}_{M-1}^T(z) \end{bmatrix}^T \quad (4.16)$$

onde $\mathbf{H}_k(z)$ é a matriz $l_k \times L_0$ com a i -ésima linha ($i = 0, \dots, l_k - 1$) formada pelos L_0 componentes polifásicos de $z^{-(\Delta_k + iL_k)} H_k(z)$ (ver Fig. 4.7).

Observando a Fig. 4.8, a função de transferência implementada pela estrutura adaptativa não-uniforme da Fig. 4.3 é dada por

$$T(z) = \widehat{\mathbf{G}}^T(z^{L_0}) \mathbf{H}_p(z^{L_0}) \begin{bmatrix} 1 & z^{-1} & \cdots & z^{-(L_0-1)} \end{bmatrix}^T. \quad (4.17)$$

Em aplicações de identificação de sistemas, os coeficientes dos subfiltros $G_k(z^{L_k})$ são adaptados para modelar um sistema FIR desconhecido. A função de transferência do sistema desconhecido, denotada por $S(z)$, pode ser expressa em termos de seus L_0 componentes polifásicos como

$$S(z) = \begin{bmatrix} S_0(z^{L_0}) & S_1(z^{L_0}) & \cdots & S_{L_0-1}(z^{L_0}) \end{bmatrix} \begin{bmatrix} 1 & z^{-1} & \cdots & z^{-(L_0-1)} \end{bmatrix}^T. \quad (4.18)$$

A partir das Eqs. (4.17) e (4.18), a estrutura em subbandas modela exatamente um

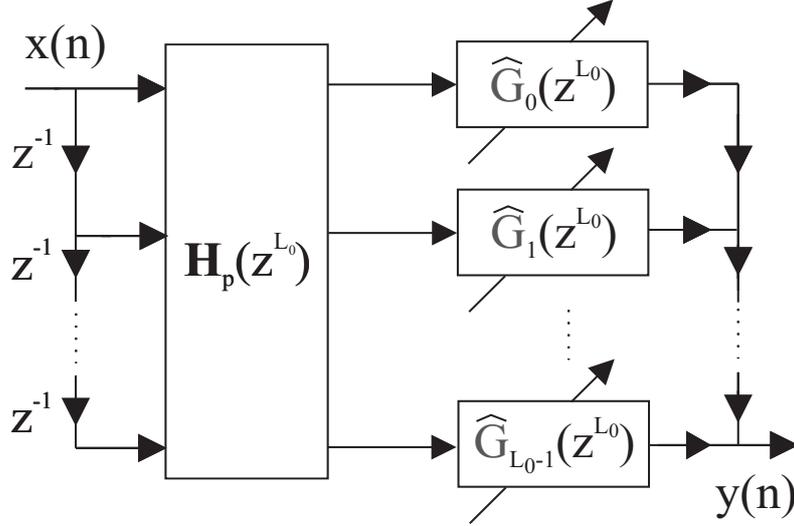


Figura 4.8: Estrutura adaptativa da Fig. 4.3 com a representação polifásica do banco de filtros não-uniforme.

sistema FIR $S(z)$ quando

$$\widehat{\mathbf{G}}^T(z)\mathbf{H}_p(z) = \begin{bmatrix} S_0(z) & S_1(z) & \cdots & S_{L_0-1}(z) \end{bmatrix}. \quad (4.19)$$

Pós-multiplicando ambos os lados da Eq. (4.19) pela matriz $\mathbf{F}_p(z)$, tal que

$$\mathbf{H}_p(z)\mathbf{F}_p(z) = z^{-\Delta p}\mathbf{I}, \quad (4.20)$$

sendo \mathbf{I} a matriz identidade de dimensão $L_0 \times L_0$ e Δp um inteiro positivo, obtemos a seguinte relação entre os coeficientes dos subfiltros esparsos e os coeficientes do sistema desconhecido:

$$\widehat{\mathbf{G}}^T(z)z^{-\Delta p} = \begin{bmatrix} S_0(z) & S_1(z) & \cdots & S_{L_0-1}(z) \end{bmatrix} \mathbf{F}_p(z). \quad (4.21)$$

A matriz $\mathbf{F}_p(z)$ que satisfaz a Eq. (4.20) corresponde à matriz polifásica dos filtros de síntese que resulta num sistema com reconstrução perfeita [5].

Na Fig. 4.7 (a) é mostrada a k -ésima banda de um sistema análise/síntese não-uniforme com PR e na Fig. 4.7 (b) é ilustrada sua expansão em l_k componentes. A partir desta figura, $\mathbf{F}_p(z)$ é a matriz de dimensão $L_0 \times L_0$ que contém os componentes polifásicos do tipo 2 dos filtros de síntese expandidos, dada por

$$\mathbf{F}_p(z) = \begin{bmatrix} \mathbf{F}_0(z) & \mathbf{F}_1(z) & \cdots & \mathbf{F}_{M-1}(z) \end{bmatrix}, \quad (4.22)$$

onde $\mathbf{F}_k(z)$ é a matriz $L_0 \times l_k$ com a i -ésima coluna ($i = 0, \dots, l_k - 1$) formada pelos L_0 componentes polifásicos do tipo 2 de $z^{-(l_k-i-1)L_k} F_k(z)$.

Os coeficientes ótimos do filtro da k -ésima subbanda $G_k^*(z)$, assumindo a hipótese de existência de sobreposição do espectro apenas entre bandas adajacentes, são dados por

$$G_k^*(z) = \sum_{i=0}^{l_k-1} z^{-i} G_{k,i}(z^{l_k}) \quad (4.23)$$

sendo os filtros $G_{k,i}(z)$ relacionados com $\widehat{G}_k(z)$ através das Eqs. (4.14) e (4.15).

De acordo com as Eqs. (4.21) e (4.23), para uma estrutura não-uniforme de M -bandas com filtros de síntese de ordem N_{F_k} , o número de coeficientes adaptativos dos subfiltros $G_k(z)$ deve ser no mínimo

$$N_{G_k} = \left\lceil \frac{N_S + N_{F_k}}{L_k} \right\rceil + 1 \quad (4.24)$$

sendo N_S a ordem do sistema desconhecido.

Então, utilizando um banco de filtros não-uniforme que permite reconstrução perfeita e subfiltros adaptativos esparsos com ordens suficientes para satisfazer a Eq. (4.24), a estrutura da Fig. 4.5 implementa exatamente qualquer sistema FIR. Porém, deve ser enfatizado que o atraso introduzido pelo banco de filtros deve ser considerado no algoritmo de adaptação.

Para exemplificar os resultados obtidos nessa seção vamos considerar a identificação de sistema desconhecido com a estrutura proposta supondo uma decomposição em $M = 3$ bandas. De acordo com as Eqs. (4.14) e (4.15), podemos escrever a relação entre os coeficientes dos subfiltros e os coeficientes do sistema desconhecida (ver Eq. (4.21) supondo $\Delta_p = 0$) da seguinte forma:

$$\begin{bmatrix} G_{0,0}(z) \\ G_{1,0}(z) \\ G_{2,0}(z) \\ G_{2,1}(z) \end{bmatrix}^T = \begin{bmatrix} S_0(z) \\ S_1(z) \\ S_2(z) \\ S_3(z) \end{bmatrix}^T \begin{bmatrix} F_{0,0}(z) & F_{1,0}(z) & F_{2,0}(z) & F_{3,0}(z) \\ F_{0,1}(z) & F_{1,1}(z) & F_{2,1}(z) & F_{3,1}(z) \\ F_{0,2}(z) & F_{1,2}(z) & F_{2,2}(z) & F_{3,1}(z) \\ F_{0,3}(z) & F_{1,3}(z) & F_{2,3}(z) & F_{3,1}(z) \end{bmatrix}. \quad (4.25)$$

Em seguida, a partir da Eq. (4.25), podemos encontrar os coeficientes ótimos dados pela Eq. (4.23), ou seja:

$$\begin{aligned} G_0^*(z) &= G_{0,0}(z) \\ G_1^*(z) &= G_{1,0}(z) \\ G_2^*(z) &= G_{2,0}(z^2) + z^{-1}G_{2,1}(z^2), \end{aligned} \quad (4.26)$$

sendo $L_0 = L_1 = 4$, $L_2 = 2$, $l_0 = l_1 = 1$ e $l_2 = 2$.

4.5 Algoritmo de Adaptação

Nesta seção derivaremos um método baseado no algoritmo LMS para atualização dos coeficientes dos subfiltros da estrutura não-uniforme em subbandas com amostragem crítica proposta. O algoritmo adaptativo resultante opera na menor taxa de amostragem entre as subbandas e emprega um passo de adaptação normalizado pelas potências dos sinais responsáveis pelos erros de cada subbanda. [31].

Definimos como $x_{i,j}(m)$ o sinal decimado na saída dos filtros $H_{i,j}(z)$ (ver Fig. 4.5) na iteração m , como $\mathbf{x}_{i,j}(m)$ o vetor que contém as N_{G_i} amostras mais recentes de $x_{i,j}(m)$ que foi decimado por um fator $L_{i,j} = \min(L_i, L_j)$ em relação ao sinal de entrada $x(n)$, como $\mathbf{g}_i(m)$ o vetor contendo os coeficientes do subfiltro $G_i(z)$, e consideramos como função objetivo o somatório das médias dos erros quadráticos² de cada subbanda, isto é,

$$J(m) = \sum_{k=0}^{M-1} \frac{1}{l_k} \sum_{q=0}^{l_k-1} e_k^2(m'_k) \quad (4.27)$$

sendo $m'_k = l_k m - q$ e $q = 0, \dots, l_k - 1$. A equação de atualização do tipo gradiente descendente [1] para os coeficientes do k -ésimo subfiltro é dada por

$$\mathbf{g}_k(m+1) = \mathbf{g}_k(m) - \frac{\mu}{2} \frac{\partial J(m)}{\partial \mathbf{g}_k}, \quad (4.28)$$

sendo μ o passo de adaptação. Os erros de cada subbanda, mostrados na Fig. 4.9, são dados por

$$\begin{aligned} e_k(m'_k) &= d_k(m'_k - \Delta_k) - \mathbf{x}_{k,k}^T \left(m'_k - \frac{\Delta_k}{L_k} \right) \mathbf{g}_k(m) \\ &- \mathbf{x}_{k-1,k}^{\downarrow T} \left(m'_k - \frac{\Delta_{k-1}}{L_k} \right) \mathbf{g}_{k-1}(m) \\ &- \mathbf{x}_{k+1,k}^T \left(\frac{L_k}{L_{k+1}} m'_k - \frac{\Delta_{k+1}}{L_k} \right) \mathbf{g}_{k+1}(m), \end{aligned} \quad (4.29)$$

sendo $\mathbf{x}_{k,k+1}^{\downarrow}$ o vetor que contém as N_{G_k} amostras do sinal $x_{k,k+1}(m)$ considerando somente as (L_k/L_{k+1}) -ésimas amostras, por causa do fator de esparsidade L_k/L_{k+1} do subfiltro adaptativo $G_k(z^{L_k/L_{k+1}})$ mostrado na Fig. 4.5.

²Os filtros adaptativos trabalham em taxas de amostragens diferentes, por isso consideramos, nas subbandas que operam em taxas mais altas, a média dos erros das amostras nas saídas dos respectivos subfiltros.

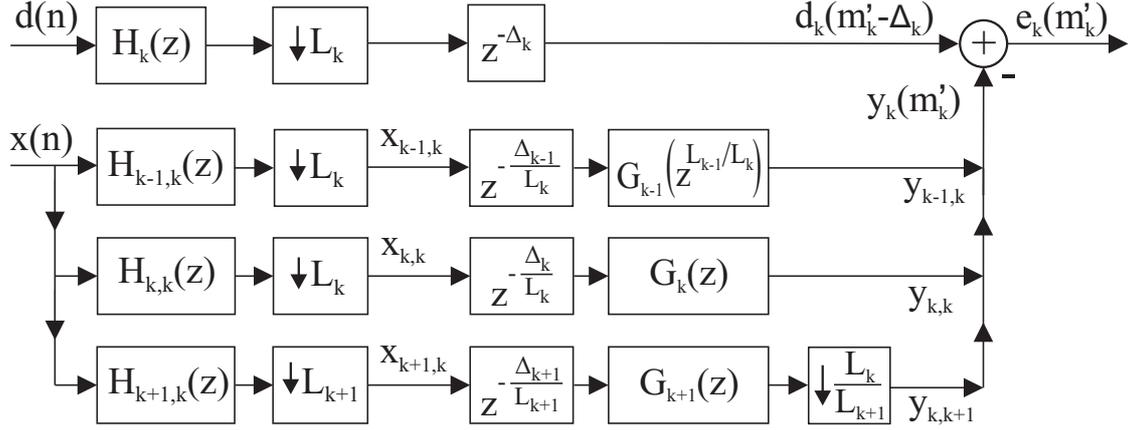


Figura 4.9: Erro da k -ésima subbanda.

Substituindo a Eq. (4.29) na Eq. (4.27), temos que

$$\begin{aligned}
\frac{\partial J(m)}{\partial \mathbf{g}_k} &= \frac{2}{l_k} \sum_{q=0}^{l_k-1} e_k(m'_k) \frac{\partial e_k(m'_k)}{\partial \mathbf{g}_k} \\
&+ \frac{2}{l_{k-1}} \sum_{q=0}^{l_{k-1}-1} e_{k-1}(m'_{k-1}) \frac{\partial e_{k-1}(m'_{k-1})}{\partial \mathbf{g}_k} \\
&+ \frac{2}{l_{k+1}} \sum_{q=0}^{l_{k+1}-1} e_{k+1}(m'_{k+1}) \frac{\partial e_{k+1}(m'_{k+1})}{\partial \mathbf{g}_k}, \tag{4.30}
\end{aligned}$$

sendo

$$\begin{aligned}
\frac{\partial e_k(m'_k)}{\partial \mathbf{g}_k} &= -\mathbf{x}_{k,k} \left(m'_k - \frac{\Delta_k}{L_k} \right), \\
\frac{\partial e_{k-1}(m'_{k-1})}{\partial \mathbf{g}_k} &= -\mathbf{x}_{k,k-1} \left(\frac{L_{k-1}}{L_k} m'_{k-1} - \frac{\Delta_k}{L_{k-1}} \right), \\
\frac{\partial e_{k+1}(m'_{k+1})}{\partial \mathbf{g}_k} &= -\mathbf{x}_{k,k+1}^\downarrow \left(m'_{k+1} - \frac{\Delta_k}{L_{k+1}} \right). \tag{4.31}
\end{aligned}$$

Substituindo, agora, as Eqs. (4.30) e (4.31) na Eq. (4.28) e considerando diferentes passos de adaptação para cada termo, obtemos

$$\begin{aligned}
\mathbf{g}_k(m+1) &= \mathbf{g}_k(m) + \frac{\mu_{k,k}}{l_k} \sum_{q=0}^{l_k-1} \mathbf{x}_{k,k} \left(m'_k - \frac{\Delta_k}{L_k} \right) e_k(m'_k) \\
&+ \frac{\mu_{k-1,k}}{l_{k-1}} \sum_{q=0}^{l_{k-1}-1} \mathbf{x}_{k,k-1} \left(\frac{L_{k-1}}{L_k} m'_{k-1} - \frac{\Delta_k}{L_{k-1}} \right) e_{k-1}(m'_{k-1}) \\
&+ \frac{\mu_{k,k+1}}{l_{k+1}} \sum_{q=0}^{l_{k+1}-1} \mathbf{x}_{k,k+1}^\downarrow \left(m'_{k+1} - \frac{\Delta_k}{L_{k+1}} \right) e_{k+1}(m'_{k+1}). \tag{4.32}
\end{aligned}$$

Para melhorar a robustez e consequentemente a taxa de convergência do algoritmo adaptativo quando sinais de entrada coloridos são considerados, cada passo

de adaptação da Eq. (4.32) é feito inversamente proporcional à soma das potências dos sinais responsáveis pelo erro correspondente [31], ou seja,

$$\begin{aligned}\mu_{k-1,k} &= \frac{\tilde{\mu}}{P_{k-2,k-1} + P_{k-1,k-1} + P_{k-1,k}}, \\ \mu_{k,k} &= \frac{\tilde{\mu}}{P_{k-1,k} + P_{k,k} + P_{k,k+1}}, \\ \mu_{k,k+1} &= \frac{\tilde{\mu}}{P_{k,k+1} + P_{k+1,k+1} + P_{k+1,k+2}},\end{aligned}\tag{4.33}$$

sendo $\tilde{\mu}$ um fator constante e $P_{r,s}$ a potência do sinal $x_{r,s}(m)$ que pode ser estimada recursivamente por

$$P_{r,s}(m+1) = \beta P_{r,s}(m) + (1-\beta)x_{r,s}^2(m)\tag{4.34}$$

com $0 \leq \beta < 1$.

Portanto, na adaptação recursiva da Eq. (4.32), cada termo de erro é ponderado por um passo normalizado diferente, tendo maior peso o erro gerado pelo sinal decimado de menor potência.

4.6 Análise da Convergência

Para analisar o comportamento da convergência do algoritmo adaptativo da Seção 4.5, estudamos a evolução da média dos vetores de erro dos coeficientes, assumindo sinais estacionários e usando a teoria da independência [1]. O vetor de erro dos coeficientes da k -ésima subbanda é definido como

$$\tilde{\mathbf{g}}_k(m) = \mathbf{g}_k(m) - \mathbf{g}_k^*,\tag{4.35}$$

sendo \mathbf{g}_k^* o vetor de coeficientes ótimos correspondente. Considerando que existe sobreposição de espectro dos filtros de análise somente entre subbandas adjacentes (ver Seção 4.3) e que o sistema a ser modelado é FIR, o sinal desejado em cada subbanda d_k pode ser expresso, em termos do vetor de coeficientes ótimos, como

$$\begin{aligned}d_k(m'_k - \Delta_k) &= \mathbf{x}_{k,k}^T \left(m'_k - \frac{\Delta_k}{L_k} \right) \mathbf{g}_k^* \\ &+ \mathbf{x}_{k-1,k}^{\downarrow T} \left(m'_k - \frac{\Delta_{k-1}}{L_k} \right) \mathbf{g}_{k-1}^* \\ &+ \mathbf{x}_{k+1,k}^T \left(\frac{L_k}{L_{k+1}} m'_k - \frac{\Delta_{k+1}}{L_{k+1}} \right) \mathbf{g}_{k+1}^* \\ &+ e_k^*(m'_k),\end{aligned}\tag{4.36}$$

sendo e_k^* o erro de modelagem residual (erro obtido utilizando-se os coeficientes ótimos) da k -ésima subbanda.

Com o propósito de simplificar a notação, definimos os vetores de dimensão $N_{G_i} \times 1$ que contêm os sinais das subbandas com os respectivos atrasos como

$$\widehat{\mathbf{x}}_{i,j}(m) = \mathbf{x}_{i,j} \left(m - \frac{\Delta_i}{L_{i,j}} \right) \quad (4.37)$$

e

$$\widehat{\mathbf{x}}_{i,j}^\downarrow(m) = \mathbf{x}_{i,j}^\downarrow \left(m - \frac{\Delta_i}{L_{i,j}} \right). \quad (4.38)$$

sendo $L_{i,j} = \min(L_i, L_j)$. Portanto, substituindo as Eqs. (4.29), (4.35) e (4.36) na Eq. (4.32), tirando o valor esperado de ambos os lados da equação resultante, e assumindo que o erro $e_k^*(m)$ tem média zero, temos que

$$\begin{bmatrix} E[\widetilde{\mathbf{g}}_0(m+1)] \\ \vdots \\ E[\widetilde{\mathbf{g}}_{M-1}(m+1)] \end{bmatrix} = (\mathbf{I} - \mathbf{\Gamma}) \begin{bmatrix} E[\widetilde{\mathbf{g}}_0(m)] \\ \vdots \\ E[\widetilde{\mathbf{g}}_{M-1}(m)] \end{bmatrix}, \quad (4.39)$$

sendo \mathbf{I} a matriz identidade de ordem $\sum_{k=0}^{M-1} N_{G_k}$ e

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{C}_0 & \mathbf{D}_0 & \mathbf{E}_0 & \mathbf{0} & \dots & \dots & \dots & \mathbf{0} \\ \mathbf{B}_1 & \mathbf{C}_1 & \mathbf{D}_1 & \mathbf{E}_1 & \mathbf{0} & \ddots & \dots & \vdots \\ \mathbf{A}_2 & \mathbf{B}_2 & \mathbf{C}_2 & \mathbf{D}_2 & \mathbf{E}_2 & \mathbf{0} & \ddots & \vdots \\ \mathbf{0} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \vdots & \ddots & \mathbf{0} & \mathbf{A}_{M-3} & \mathbf{B}_{M-3} & \mathbf{C}_{M-3} & \mathbf{D}_{M-3} & \mathbf{E}_{M-3} \\ \vdots & \dots & \ddots & \mathbf{0} & \mathbf{A}_{M-2} & \mathbf{B}_{M-2} & \mathbf{C}_{M-2} & \mathbf{D}_{M-2} \\ \mathbf{0} & \dots & \dots & \dots & \mathbf{0} & \mathbf{A}_{M-1} & \mathbf{B}_{M-1} & \mathbf{C}_{M-1} \end{bmatrix} \quad (4.40)$$

com

$$\begin{aligned}
\mathbf{A}_k &= \mu_{k-1,k} E \left[\widehat{\mathbf{x}}_{k,k-1}(m'_{k-1}) \widehat{\mathbf{x}}_{k-2,k-1}^{\downarrow T}(m'_{k-1}) \right], \\
\mathbf{B}_k &= \mu_{k-1,k} E \left[\widehat{\mathbf{x}}_{k,k-1}(m'_{k-1}) \widehat{\mathbf{x}}_{k-1,k-1}^T(m'_{k-1}) \right] \\
&\quad + \mu_{k,k} E \left[\widehat{\mathbf{x}}_{k,k}(m'_k) \widehat{\mathbf{x}}_{k-1,k}^{\downarrow T}(m'_k) \right], \\
\mathbf{C}_k &= \mu_{k-1,k} E \left[\widehat{\mathbf{x}}_{k,k-1}(m'_{k-1}) \widehat{\mathbf{x}}_{k,k-1}^T(m'_{k-1}) \right] \\
&\quad + \mu_{k,k} E \left[\widehat{\mathbf{x}}_{k,k}(m'_k) \widehat{\mathbf{x}}_{k,k}^T(m'_k) \right] \\
&\quad + \mu_{k,k+1} E \left[\widehat{\mathbf{x}}_{k,k+1}^{\downarrow}(m'_{k+1}) \widehat{\mathbf{x}}_{k,k+1}^{\downarrow T}(m'_{k+1}) \right], \\
\mathbf{D}_k &= \mu_{k,k} E \left[\widehat{\mathbf{x}}_{k,k}(m'_k) \widehat{\mathbf{x}}_{k+1,k}^T(m'_k) \right] \\
&\quad + \mu_{k,k+1} E \left[\widehat{\mathbf{x}}_{k,k+1}^{\downarrow}(m'_{k+1}) \widehat{\mathbf{x}}_{k+1,k+1}^T(m'_{k+1}) \right], \\
\mathbf{E}_k &= \mu_{k,k+1} E \left[\widehat{\mathbf{x}}_{k,k+1}^{\downarrow}(m'_{k+1}) \widehat{\mathbf{x}}_{k+2,k+1}^T(m'_{k+1}) \right],
\end{aligned} \tag{4.41}$$

onde as dimensões das matrizes \mathbf{A}_k , \mathbf{B}_k , \mathbf{C}_k , \mathbf{D}_k e \mathbf{E}_k são $N_{G_k} \times N_{G_{k-2}}$, $N_{G_k} \times N_{G_{k-1}}$, $N_{G_k} \times N_{G_k}$, $N_{G_k} \times N_{G_{k+1}}$ e $N_{G_k} \times N_{G_{k+2}}$, respectivamente.

As matrizes acima podem ser reescritas em termos dos coeficientes dos filtros de análise e da função de autocorrelação do sinal de entrada, uma vez que

$$E \left[\widehat{\mathbf{x}}_{i,j}(m'_j) \widehat{\mathbf{x}}_{r,s}^T(m'_s) \right] = \widehat{\mathbf{H}}_{i,j} \mathbf{R}_{xx} \widehat{\mathbf{H}}_{r,s}^T, \tag{4.42}$$

e

$$E \left[\widehat{\mathbf{x}}_{i,j}(m'_j) \widehat{\mathbf{x}}_{r,s}^{\downarrow T}(m'_s) \right] = \widehat{\mathbf{H}}_{i,j} \mathbf{R}_{xx} \widehat{\mathbf{H}}_{r,s}^{\downarrow T}. \tag{4.43}$$

Nas equações acima, \mathbf{R}_{xx} é a matriz de autocorrelação do sinal de entrada de dimensão $N_T \times N_T$, sendo $N_T = \max[(L_j N_{G_i} + N_{H_{i,j}} + 1 + \Delta_j), (L_s N_{G_r} + N_{H_{r,s}} + 1 + \Delta_r)]$; $\widehat{\mathbf{H}}_{i,j}$ é a matriz de dimensão $N_{G_j} \times N_T$, com a primeira linha de elementos diferentes de zero na posição $\Delta_i + 1$ e contendo os coeficientes de $H_{i,j}(z)$, com as linhas seguintes sendo formadas deslocando-se os elementos da linha anterior para a direita de L_j posições; e $\widehat{\mathbf{H}}_{i,j}^{\downarrow}$ é similar a $\widehat{\mathbf{H}}_{i,j}$ porém com deslocamentos de L_i posições de uma linha para a próxima.

Observando a Eq. (4.39), podemos concluir que a taxa de convergência do algoritmo adaptativo em subbandas, dado pela Eq. (4.32), é governada pela relação entre os autovalores da matriz $\mathbf{\Gamma}$. Este resultado não só nos permite prever o comportamento do algoritmo para um determinado banco de análise, mas também pode ser usado, com algum conhecimento das estatísticas do sinal de entrada, para selecionar o número ótimo de subbandas em relação à velocidade de convergência do algoritmo adaptativo.

Na Seção 4.9 será mostrado que a relação de autovalores da matriz $\mathbf{\Gamma}$ é significativamente reduzida para sinais de entrada coloridos quando aumenta-se o número de subbandas, resultando numa taxa de convergência melhor que a obtida pelo algoritmo LMS em banda cheia.

4.7 Erro Quadrático Médio em Regime Estacionário

Nesta seção vamos apresentar a análise do erro quadrático médio (MSE) em regime permanente para o algoritmo da estrutura não-uniforme em subbandas, que leva em conta apenas o erro causado pela suposição de que não havia sobreposição de espectro dos filtros de análise entre subbandas não-adjacentes durante a derivação do algoritmo adaptativo (Seção 4.5). Outros erros podem ser modelados pela análise convencional do LMS.

Analisando o caso geral (ou seja, considerando que existe sobreposição de espectro entre as subbandas não-adjacentes), o sinal desejado $d_k(m'_k - \Delta_k)$ para a k -ésima subbanda pode ser expresso em termos dos coeficientes adaptativos ótimos \mathbf{g}_k^* na iteração m como

$$\begin{aligned} d_k(m'_k - \Delta_k) &= \sum_{i=0}^{k-1} \mathbf{x}_{i,k}^{\downarrow T} (m'_k - \frac{\Delta_i}{L_k}) \mathbf{g}_i^* \\ &+ \sum_{i=k}^{M-1} \mathbf{x}_{i,k}^T (\frac{L_k}{L_i} m'_k - \frac{\Delta_i}{L_k}) \mathbf{g}_i^* + v_k(m'_k) \end{aligned} \quad (4.44)$$

sendo $v_k(m)$ o ruído de medição da k -ésima subbanda.

Definindo $\mathbf{d}(m) = [d_0(l_0 m) \quad d_1(l_1 m) \quad \cdots \quad d_{M-1}(l_{M-1} m)]^T$ como o vetor formado pelos sinais desejados das M subbandas no tempo $l_k m$, este pode ser escrito na forma matricial como

$$\mathbf{d}(m) = \mathbf{X}(m) \mathbf{g}^* + \mathbf{v}(m) \quad (4.45)$$

sendo $\mathbf{g}^* = [\mathbf{g}_0^{*T} \quad \mathbf{g}_1^{*T} \quad \cdots \quad \mathbf{g}_{M-1}^{*T}]^T$ o vetor que contém os coeficientes ótimos de todos os subfiltros adaptativos, $\mathbf{v}(m)$ o vetor formado pelos erros residuais de modelagem de todas as M subbandas no tempo $l_k m$, e $\mathbf{X}(m)$ a matriz de dimensão $M \times (\sum_{k=0}^{M-1} N_{G_k})$ com os sinais de entrada das subbandas, cuja k -ésima linha é

dada por

$$[\mathbf{X}(m)]_k = \begin{bmatrix} \hat{\mathbf{x}}_{0,k}^{\downarrow T}(l_k m) & \cdots & \hat{\mathbf{x}}_{k-1,k}^{\downarrow T}(l_k m) & \hat{\mathbf{x}}_{k,k}^T(l_k m) \\ \hat{\mathbf{x}}_{k+1,k}^T(2l_k m) & \cdots & \hat{\mathbf{x}}_{M-1,k}^T(2^{M-1-k}l_k m) & \end{bmatrix} \quad (4.46)$$

Considerando, agora, somente a sobreposição de espectro entre as subbandas adjacentes, o vetor $\mathbf{y}(m) = [y_0(l_0 m) \ y_1(l_1 m) \ \cdots \ y_{M-1}(l_{M-1} m)]^T$ formado pelos sinais de saída das M subbandas da estrutura não-uniforme no tempo $l_k m$ (ver Fig. 4.6) pode ser escrito como

$$\mathbf{y}(m) = \hat{\mathbf{X}}(m)\mathbf{g}(m) \quad (4.47)$$

sendo $\hat{\mathbf{X}}(m)$, a matriz formada pelos vetores de sinais na entrada dos filtros adaptativos, dada por

$$[\hat{\mathbf{X}}(m)]_k = \begin{bmatrix} \mathbf{0}_0^T & \cdots & \mathbf{0}_{k-2}^T & \hat{\mathbf{x}}_{k-1,k}^{\downarrow T}(l_k m) & \hat{\mathbf{x}}_{k,k}^T(l_k m) \\ \hat{\mathbf{x}}_{k+1,k}^T(2l_k m) & \mathbf{0}_{k+2}^T & \cdots & \mathbf{0}_{M-1}^T & \end{bmatrix} \quad (4.48)$$

e $\mathbf{g}(m) = [\mathbf{g}_0^T(m) \ \mathbf{g}_1^T(m) \ \cdots \ \mathbf{g}_{M-1}^T(m)]^T$ é o vetor que contém todos os coeficientes adaptativos, com o vetor $\mathbf{0}_i^T$ de dimensão $1 \times N_{G_i}$ sendo formado por zeros em substituição ao vetor de sinal na entrada do i -ésimo filtro adaptativo de uma banda não-adjacente.

Portanto, o vetor formado pelos sinais de erro das subbandas resultantes da hipótese de que não havia sobreposição de espectro entre os filtros de análise não-adjacentes na derivação da estrutura com amostragem crítica é dado por

$$\begin{aligned} \mathbf{e}(m) &= \mathbf{d}(m) - \mathbf{y}(m) \\ &= \mathbf{X}(m)\mathbf{g}^* - \hat{\mathbf{X}}(m)\mathbf{g}(m) + \mathbf{v}(m). \end{aligned} \quad (4.49)$$

sendo $\mathbf{v}(m)$ o vetor que contém os erros de modelagem residual de cada subbanda.

Um banco de filtros é considerado sem perdas quando a potência do erro é igual à soma das potências dos erros de cada subbanda [32]. Portanto, para estes bancos de filtros, o MSE total é dado por

$$\begin{aligned} \xi(m) &= \sum_{k=0}^{M-1} \frac{1}{l_k} \sum_{q=0}^{l_k-1} E[e_k^2(m'_k)] \\ &\approx \sum_{k=0}^{M-1} E[e_k^2(l_k m)] = E[\mathbf{e}(m)^T \mathbf{e}(m)]. \end{aligned} \quad (4.50)$$

Substituindo as Eqs. (4.45), (4.47) e (4.49) na equação anterior, assumindo que os erros residuais têm média zero e que, após a convergência, $E[\mathbf{g}(\infty)] \approx \mathbf{g}^*$, temos que

$$\xi_{\text{SS}} = \xi(\infty) \approx \mathbf{g}^{*T} \mathbf{\Phi} \mathbf{g}^* + \sigma_v^2 \quad (4.51)$$

sendo

$$\mathbf{\Phi} = E \left\{ \left[\mathbf{X}(m) - \widehat{\mathbf{X}}(m) \right] \left[\mathbf{X}(m) - \widehat{\mathbf{X}}(m) \right]^T \right\} \quad (4.52)$$

e σ_v^2 a variância do ruído de medição. A k -ésima linha da matriz $\left[\mathbf{X}(m) - \widehat{\mathbf{X}}(m) \right]$, dada por

$$\begin{aligned} [\mathbf{X}(m) - \widehat{\mathbf{X}}(m)]_k &= [\widehat{\mathbf{x}}_{0,k}^{\downarrow T}(l_k m) \quad \cdots \quad \widehat{\mathbf{x}}_{k-2,k}^{\downarrow T}(l_k m) \\ &\mathbf{0}_{k-1}^T \quad \mathbf{0}_k^T \quad \mathbf{0}_{k+1}^T \quad \widehat{\mathbf{x}}_{k+2,k}^T(2l_k m) \quad \cdots \quad \widehat{\mathbf{x}}_{M-1,k}^T(2^{M-1-k} l_k m)], \end{aligned} \quad (4.53)$$

contém os componentes do sinal de entrada da k -ésima subbanda nas bandas de frequência não adjacentes a esta subbanda. Desta forma, a matriz $\mathbf{\Phi}$ é formada pelas matrizes de correlação dos sinais nas subbandas não-adjacentes.

Para filtros de análise seletivos, os coeficientes adaptativos convergirão aproximadamente para os coeficientes ótimos da estrutura esparsa. Conseqüentemente, os elementos do vetor \mathbf{g}^* serão aproximadamente iguais aos coeficientes dos filtros $G_k^*(z)$ dados pela Eq. (4.23).

Observando a Eq. (4.51), o MSE da estrutura em subbandas não-uniforme em regime permanente será, em geral, maior que a variância σ_v^2 do ruído de medição porque a sobreposição residual do espectro não é cancelada na estrutura simplificada. O correspondente acréscimo no MSE está relacionado com a atenuação na faixa de rejeição dos filtros de análise e pode ser estimado de forma simplificada pela expressão a seguir:

$$\xi_{\text{SS}} \approx \lambda_{\text{mean}}(\mathbf{\Phi}) |\mathbf{g}^*|^2 + \sigma_v^2 \quad (4.54)$$

sendo $\lambda_{\text{mean}}(\mathbf{\Phi})$ o autovalor médio da matriz $\mathbf{\Phi}$. Esta forma simplificada de estimar o excesso de MSE não requer nenhum conhecimento dos coeficientes ótimos, mas somente a norma ao quadrado do vetor de coeficientes \mathbf{g}^* correspondente, a qual pode ser estimada pela razão entre as potências dos sinais na saída e na entrada destes subfiltros. Esta expressão simplificada também é função dos coeficientes dos filtros de análise e da autocorrelação da sequência de entrada.

4.8 Complexidade Computacional

Nesta seção analisaremos a complexidade computacional dos algoritmos, considerando o número de multiplicações reais por amostra do sinal de entrada (NMPA) necessárias para atualizar os coeficientes do filtro adaptativo utilizado na identificação de sistemas FIR e para obter o sinal de saída.

Para o algoritmo LMS convencional, o número de multiplicações necessárias é dado por

$$\text{NMPA} = 2N_S, \quad (4.55)$$

sendo N_S o comprimento do sistema desconhecido em banda cheia.

Para a estrutura não-uniforme maximamente decimada mostrada na Fig 4.6 com M canais, o número de multiplicações necessárias (por amostra do sinal de entrada) é dado por

$$\begin{aligned} \text{NMPA} = & \sum_{k=0}^{M-1} \frac{2N_{H_k} + 1}{L_k} + \sum_{k=0}^{M-2} \frac{N_{H_k} + N_{H_{k+1}} + 1}{L_{k+1}} + 2 \sum_{k=0}^{M-1} \frac{N_{H_k} + 1}{L_k} \\ & + \sum_{k=1}^{M-1} \frac{N_{G_{k-1}}}{L_k} + \sum_{k=0}^{M-1} \frac{N_{G_k}}{L_k} + \sum_{k=0}^{M-2} \frac{N_{G_{k+1}}}{L_k} \\ & + \sum_{k=1}^{M-1} \frac{l_{k-1} N_{G_{k-1}}}{L_0} + \sum_{k=0}^{M-1} \frac{l_k N_{G_k}}{L_0} + \sum_{k=0}^{M-2} \frac{l_{k+1} N_{G_{k+1}}}{L_0}. \end{aligned} \quad (4.56)$$

Os dois primeiros termos correspondem à decomposição do sinal de entrada pelos filtros diretos $H_{k,k}(z)$ e extras $H_{k,k+1}(z)$; o terceiro termo corresponde à decomposição pelos filtros $H_k(z)$ e $F_k(z)$ (considerando que eles têm a mesma ordem, ou seja, $N_{H_k} = N_{F_k}$), sendo o fator L_k no denominador devido à decimação; os três termos seguintes correspondem à filtragem dos sinais decompostos pelos subfiltros $G_k(z)$, que é realizada numa taxa de amostragem L_k vezes menor que a taxa do sinal de entrada; finalmente, os últimos três termos dizem respeito à atualização dos coeficientes adaptativos realizada na taxa L_0 . Substituindo os valores de N_{G_k} (Eq.

(4.24)) e de l_k na equação acima, temos:

$$\begin{aligned}
\text{NMPA} &= \sum_{k=0}^{M-1} \frac{2N_{H_k}(2L_k + 1) + 3L_k}{L_k^2} \\
&+ \sum_{k=1}^{M-1} \left(\frac{NH_{k-1} + NH_k + 1}{L_k} \right) + \left(\frac{NH_{k-1}}{L_{k-1}} + \frac{NH_k}{L_k} \right) \left(\frac{1}{L_k} + \frac{1}{L_{k+1}} \right) \\
&+ N_S \left(\frac{2}{L_0^2} + \sum_{k=1}^{M-1} \left(\frac{2}{L_k^2} + \frac{1}{L_k} + \frac{1}{L_{k+1}} \right)^2 \right). \tag{4.57}
\end{aligned}$$

Considerando apenas o caso em que o comprimento do sistema desconhecido é muito maior que o comprimento dos filtros protótipos, a Tabela 4.1 mostra, comparativamente, a complexidade computacional para a estrutura em subbandas não-uniforme apresentada neste capítulo e para a estrutura em subbandas uniforme mostrada em [33], variando-se o número de bandas e o comprimento N_p dos filtros protótipos utilizados na obtenção do banco de filtros não-uniforme, e considerando um sistema desconhecido de comprimento $N_d = 4096$ (que implica uma complexidade de 8192 multiplicações por amostra para o algoritmo LMS convencional). Para a estrutura proposta, consideramos também, a utilização de filtros protótipos cujos números de coeficientes são reduzidos pela metade à medida em que o número de níveis utilizados na estrutura em árvore vai aumentando (N_p variável).

Tabela 4.1: Complexidade computacional dos algoritmos das estruturas não-uniforme e uniforme maximamente decimadas.

		Estrutura Não-uniforme		Estrutura Uniforme		
M	N_p	Complexidade		M	N_p	Complexidade
		N_p fixo	N_p variável			
2	32 / 64	8410 / 8634	8410 / 8634	2	32 / 64	8410 / 8634
3	32 / 64	6849 / 7311	6790 / 7252	4	64 / 128	5545 / 5977
4	32 / 64	6622 / 7311	6454 / 7143	8	128 / 256	3633 / 4457
5	32 / 64	6728 / 7643	6399 / 7314	16	256 / 512	3061 / 4657

Podemos notar que nas estruturas maximamente decimadas a complexidade em todos os casos é menor que a complexidade do LMS convencional, exceto para $M = 2$ bandas. Algoritmos rápidos para implementação das convoluções e correlações

utilizando FFTs (*Fast Fourier Transform*) podem ser usados em implementações eficientes destas estruturas, reduzindo as complexidades descritas na Tabela 4.1.

4.9 Resultados Experimentais

Simulações computacionais são apresentadas para ilustrar o comportamento da convergência do algoritmo da estrutura não-uniforme descrito na Seção 4.5, e para comparar o desempenho desta estrutura com os desempenhos da estrutura adaptativa uniforme com decimação crítica apresentada em [34] e da estrutura adaptativa não-uniforme sem decimação crítica investigada em [2].

4.9.1 Experimento 1

Nesse experimento consideramos a identificação de um sistema FIR de ordem $N_S = 879$. Os coeficientes do sistema desconhecido são obtidos aleatoriamente (sinal branco³ com distribuição uniforme). O sinal de entrada foi gerado passando um sinal branco gaussiano por um filtro IIR de primeira ordem com pólo em $z = 0,9$. Um ruído branco com variância $\sigma_n^2 = 10^{-6}$ foi somado ao sinal desejado.

A estrutura adaptativa em subbandas não-uniforme foi implementada utilizando um banco de filtros com reconstrução perfeita e decomposição em oitavas com $M=3$, 4 e 5 subbandas, utilizando filtros protótipos projetados de acordo com [19]. A Tabela 6.5 contém os fatores de decimação L_k , os fatores $l_k = L_0/L_k$, as ordens dos filtros de análise $H_k(z)$, os atrasos Δ_k e as ordens dos filtros adaptativos $G_k(z)$ usados nas simulações com $M = 4$ subbandas. A Fig. 4.10 mostra as respostas em frequência dos filtros de análise correspondentes. O passo de adaptação da Eq. (4.33) que resultou na melhor convergência, obtido por tentativa e erro, foi $\tilde{\mu} = 1/N_{G_0}$.

A Fig. 4.11 mostra a evolução do erro quadrático médio da estrutura em subbandas não-uniforme proposta para diferentes números de subbandas. A Tabela 4.3 contém as razões dos maiores e menores autovalores das matrizes $\mathbf{\Gamma}$ (Eq. (4.40)) para as simulações da Fig. 4.11.

Quando analisamos os resultados teóricos, devemos considerar que os coeficientes adaptativos são atualizados numa taxa de amostragem que é $L_0 = 2^{M-1}$ vezes menor

³Sinais brancos são sinais aleatórios decorrelacionados no tempo.

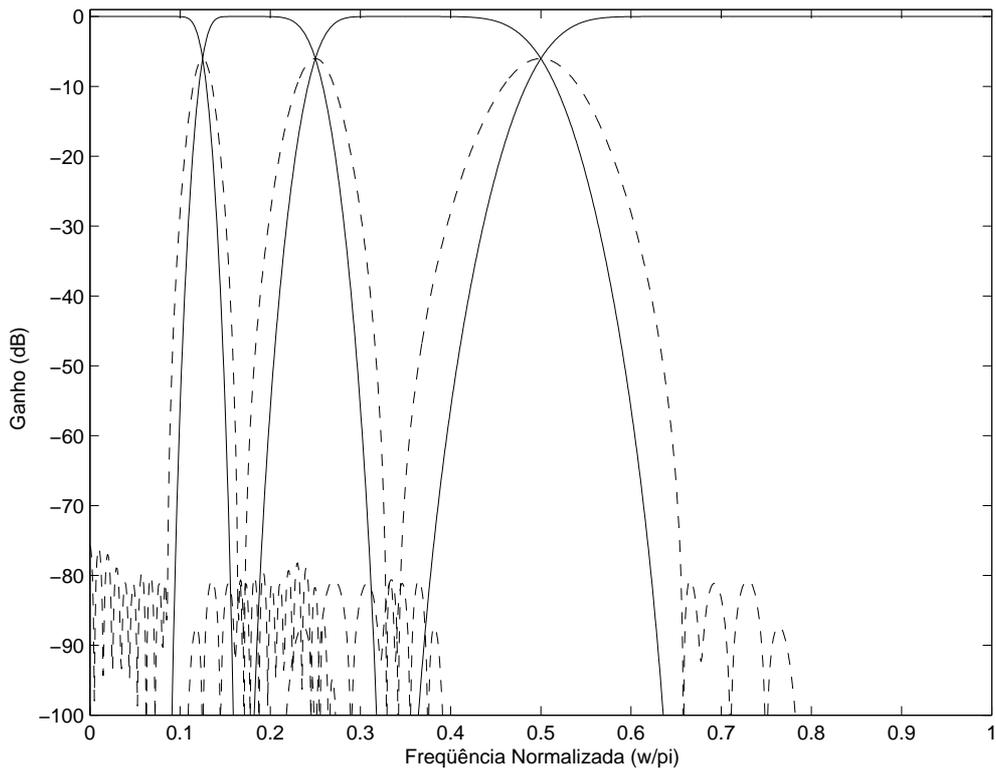


Figura 4.10: Resposta em frequência dos filtros de análise da Fig. 4.6 ($H_{k,k}(z)$: linha cheia e $H_{k,k+1}(z)$: linha tracejada).

que a taxa do sinal de entrada. Então, as relações de autovalores efetivas são aquelas da Tabela 4.3 multiplicadas por L_0 . Com esta observação em mente, podemos verificar que os resultados teóricos da Tabela 4.3 estão de acordo com os resultados experimentais da Fig. 4.11, indicando que uma melhoria significativa na taxa de convergência do algoritmo proposto pode ser obtida, para sinais de entrada coloridos, com o acréscimo do número de canais no algoritmo em subbandas. Para o sinal de entrada colorido aplicado nesse experimento, a utilização de $M = 4$ subbandas é suficiente, na prática, para descorrelacionar suas amostras. O pequeno ganho na velocidade de convergência quando acrescentamos mais um nível de decomposição ($M = 5$ subbandas), que era esperado pelos resultados teóricos da Tabela 4.3, não é verificado nos resultados experimentais (Fig. 4.11) porque o número total de coeficientes adaptativos da estrutura em subbandas torna-se maior. O efeito deste aumento no número de coeficientes adaptativos não será perceptível na modelagem de sistemas de ordens muito elevadas, tal como o caminho do eco em salas e/ou auditórios.

Tabela 4.2: Parâmetros da estrutura não-uniforme para o Experimento 1.

k	L_k	l_k	N_{H_k}	Δ_k	N_{G_k}
0	8	1	361	0	156
1	8	1	361	0	156
2	4	2	173	188	264
3	2	4	63	298	472

Tabela 4.3: Razão entre o maior e menor autovalor da matriz $\mathbf{\Gamma}$ para o sinal colorido do Experimento 1.

M	$\lambda_{\max}/\lambda_{\min}$
1	360,60
3	32,23
4	9,06
5	3,10

A seguir, compararemos a evolução do MSE considerando diferentes estruturas adaptativas em subbandas com $M = 4$ bandas. A Fig. 4.12 apresenta a evolução do MSE para a estrutura não-uniforme com decimação crítica proposta, para a estrutura uniforme criticamente decimada de [33] e para a estrutura não-uniforme subamostrada de [2]. A partir desta figura, podemos observar que a estrutura não-uniforme proposta tem a mais rápida velocidade de convergência. Os problemas de convergência da estrutura subamostrada foram investigados em [35], e podem ser reduzidos usando filtros de análise com largura de banda estendida [36]. A melhoria na taxa de convergência obtida com a estrutura não-uniforme, quando comparada à estrutura uniforme, é devida à decomposição do sinal de entrada em bandas mais estreitas nas baixas frequências, resultando em uma menor razão entre os valores máximo e mínimo da densidade espectral de potência dos sinais em subbandas [1]. Deve ser observado que a melhoria na convergência obtida com a estrutura proposta empregando banco de filtros com decomposição em oitavas foi observada para sinais de entrada coloridos com características passa-baixas. Para sinais de entrada com outros tipos de características, diferentes tipos de banco de filtros podem ser mais apropriados. Por exemplo, para sinais de entrada coloridos do tipo passa-altas, a

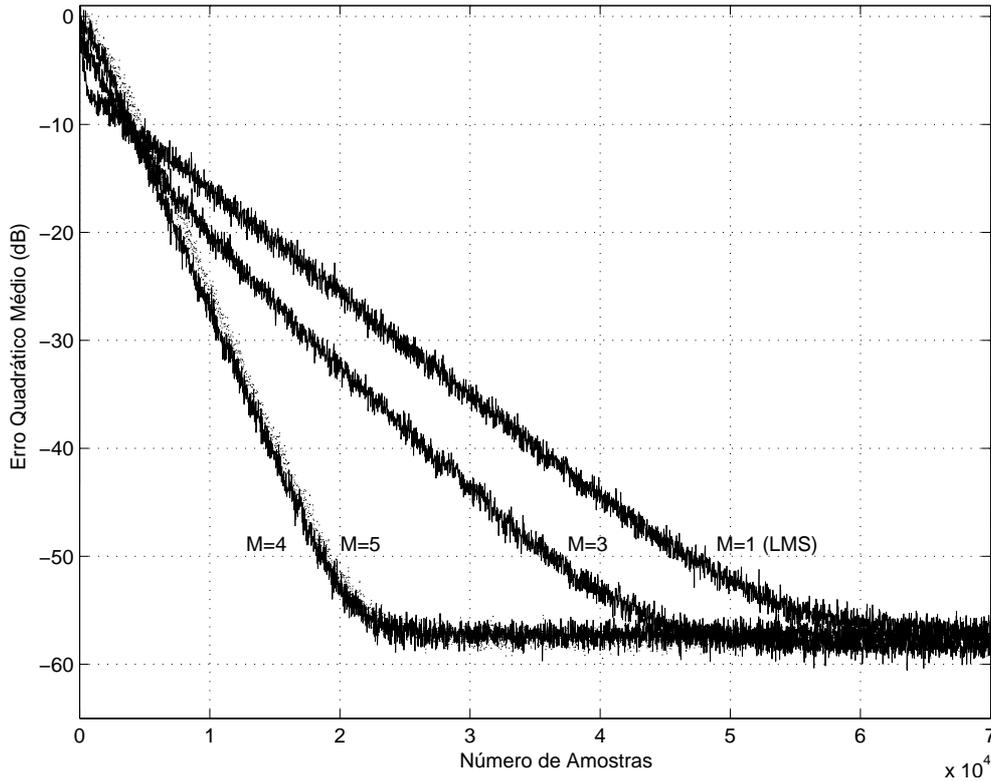


Figura 4.11: Evolução do MSE da estrutura em subbandas não-uniforme para o Experimento 1.

decomposição das bandas de altas frequências em mais níveis que as bandas de baixas frequências pode resultar numa melhor convergência.

4.9.2 Experimento 2

Esse experimento é similar à identificação do sistema FIR de ordem $N_S = 879$ considerado no Experimento 1, porém com sinal de entrada gerado passando-se um ruído branco gaussiano por um filtro IIR de primeira ordem com pólo em $z = 0,95$.

A Fig. 4.13 mostra a evolução do MSE da estrutura não-uniforme proposta com $M = 4$ e 5 subbandas. A Tabela 4.4 contém a relação dos autovalores da matriz Γ dada pela Eq. (4.40) para as simulações da Fig. 4.13. Podemos observar pela Fig. 4.13 e pela Tabela 4.4 que, para o sinal colorido usado nesse experimento, a estrutura proposta com $M = 4$ subbandas não é capaz de decorrelacioná-lo completamente; porém, uma significativa melhoria na taxa de convergência pode ser obtida aumentando-se o número de subbandas para $M = 5$.

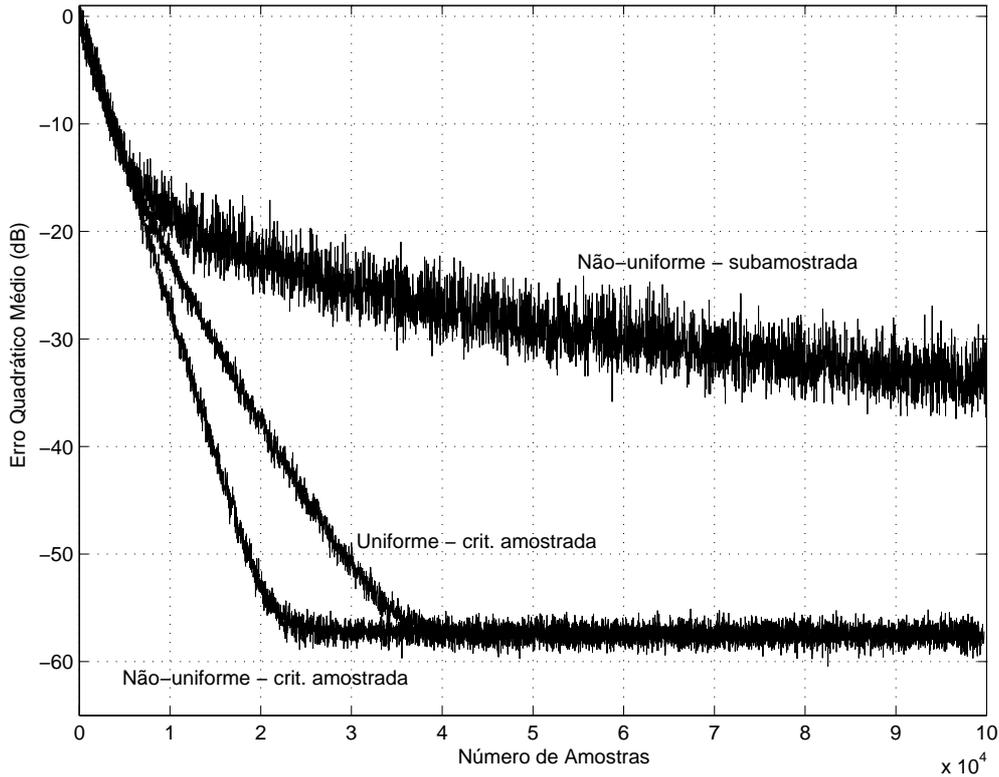


Figura 4.12: Evolução do MSE para diferentes estruturas em subbandas com $M = 4$ para o Experimento 1.

Tabela 4.4: Relação de autovalores da matriz $\mathbf{\Gamma}$ para o Experimento 2.

M	$\lambda_{\max}/\lambda_{\min}$
4	33,38
5	9,47

4.9.3 Experimento 3

Esse experimento avalia a capacidade da estrutura adaptativa em subbandas não-uniforme em responder rapidamente a mudanças no sistema desconhecido. Inicialmente consideramos a identificação de um sistema FIR passa-baixas de ordem $N_S = 299$ com faixa de passagem de 0 a $0,4\pi$. Depois de 15.000 iterações, o sistema desconhecido é alterado para um sistema FIR de mesmo comprimento, também passa-baixas, porém com faixa de passagem de 0 a $0,2\pi$. O sinal de entrada da estrutura foi colorido exatamente como no Experimento 1, e ao sinal desejado foi somado um ruído branco com variância $\sigma_v^2 = 10^{-5}$. A estrutura em subbandas foi implementada utilizando o mesmo banco de filtros em oitavas e com reconstrução perfeita

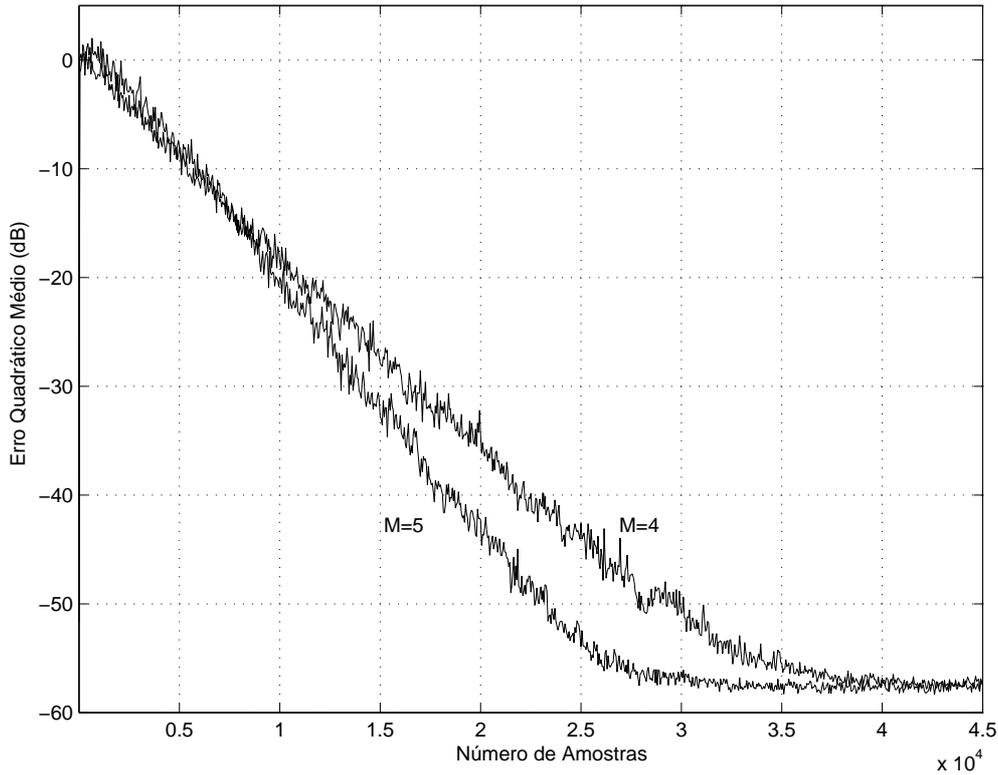


Figura 4.13: Evolução do MSE da estrutura não-uniforme para o Experimento 2.

do Experimento 1 com $M = 4$. A Fig. 4.14 mostra a evolução do erro quadrático médio (MSE) da estrutura criticamente decimada. Observando esta figura podemos perceber que o algoritmo da estrutura proposta é capaz de rastrear variações rápidas do sistema desconhecido e tem um desempenho similar ao da estrutura uniforme.

4.9.4 Experimento 4

Nesse experimento vamos comparar os MSEs em regime permanente teóricos e experimentais (após a convergência) obtidos na identificação de um sistema FIR de ordem $N_S = 127$, com coeficientes gerados novamente de forma aleatória. As simulações serão realizadas para sinais de entrada brancos e coloridos sem adição de ruído de medição ao sinal desejado. O sinal de entrada colorido foi produzido exatamente como no Experimento 1. A estrutura em subbandas não-uniforme proposta na Seção 4.3 foi simulada com $M = 4$ subbandas, empregando um banco de filtros em árvore com 3 níveis, implementado a partir de protótipos com reconstrução perfeita de ordens $N_{H_3} = 15, 31$ e 63 [7]. As estimativas teóricas do MSE (ξ_{SS} dadas pelas Eqs. (4.51) e (4.54)), assim como os MSEs experimentais (ξ_{exp}) obtidos com

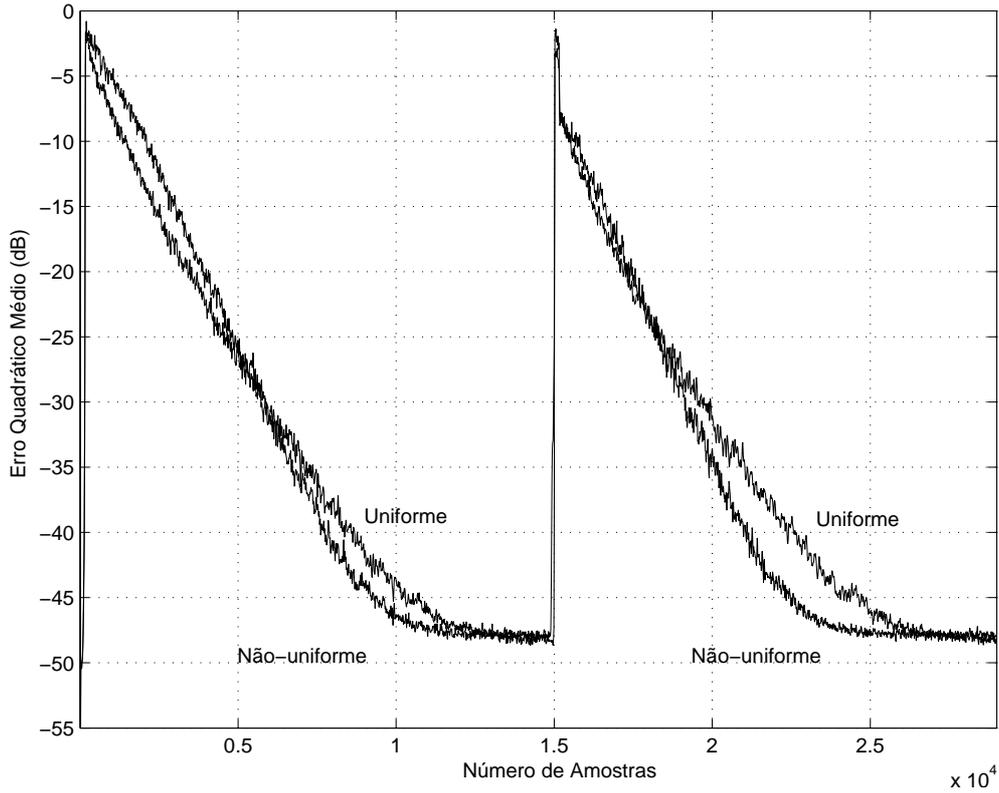


Figura 4.14: Evolução do MSE das estruturas uniforme e não-uniforme para o Experimento 3.

os três filtros protótipos são dados na Tabela 4.5.

Tabela 4.5: MSEs teóricos e experimentais (em dB) para diferentes bancos de filtros do Experimento 4.

entrada	ruído branco			sinal colorido		
N_{H_3}	15	31	63	15	31	63
ξ_{ss} (Eq. (4.51))	-38,0	-60,0	-87,2	-37,1	-56,5	-81,3
ξ_{ss} (Eq. (4.54))	-36,8	-59,0	-86,5	-36,1	-56,2	-81,3
ξ_{exp}	-37,1	-61,0	-88,6	-32,1	-54,0	-80,2

As Figs. 4.15 e 4.16 apresentam a evolução do erro quadrático médio e os valores teóricos correspondentes para a estrutura em subbandas proposta com filtros protótipos de diferentes ordens, para sinais de entrada brancos e coloridos, respectivamente.

A Fig. 4.17 mostra a resposta em frequência dos três filtros protótipos empregados e seus respectivos MSEs experimentais (ver Tabela 4.5) para sinal colorido na

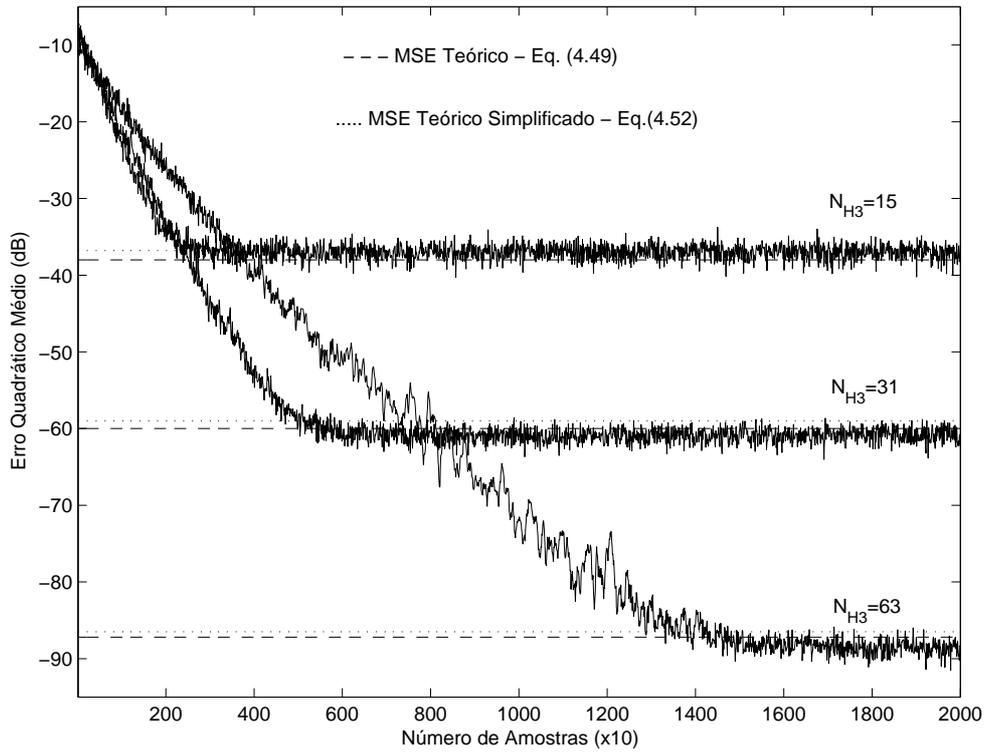


Figura 4.15: Evolução do MSE e valores teóricos do excesso de MSE para o Experimento 4, considerando sinal de entrada branco.

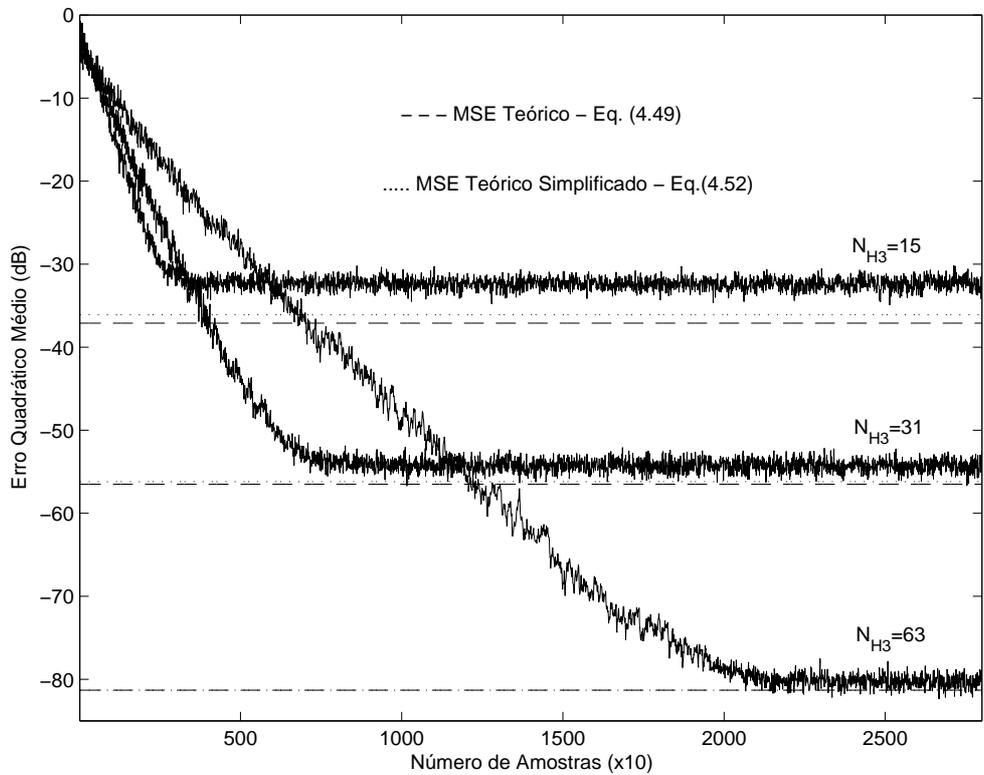


Figura 4.16: Evolução do MSE e valores teóricos do MSE em excesso para o Experimento 4, considerando sinal de entrada colorido.

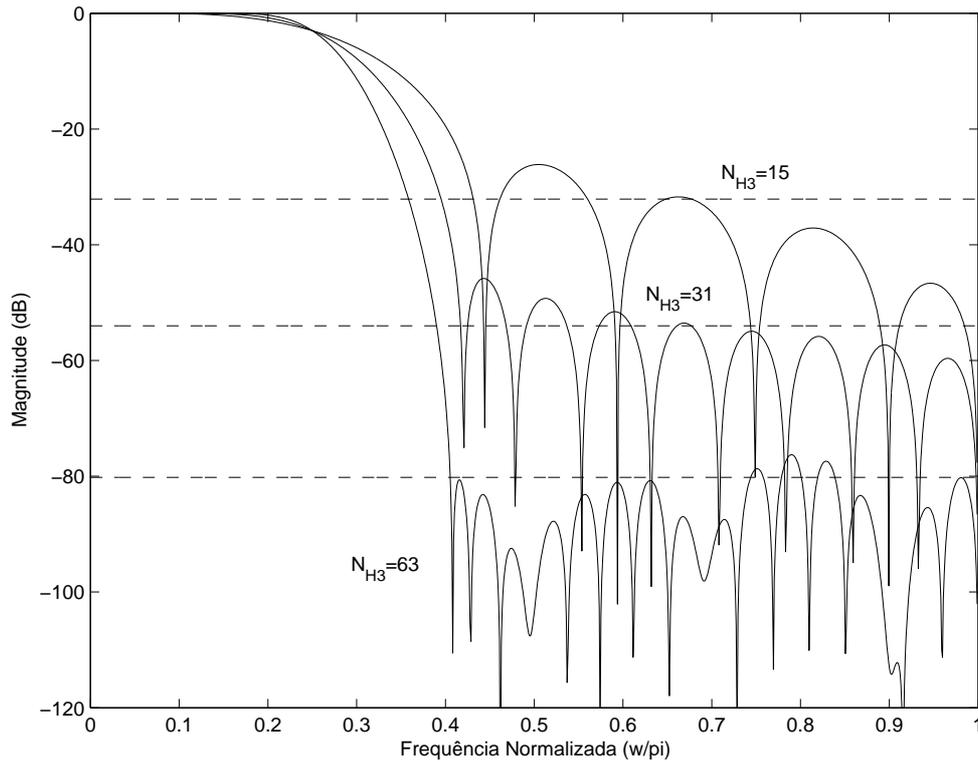


Figura 4.17: Resposta em frequência dos filtros protótipos (linha cheia) e os respectivos MSEs experimentais (linha tracejada), considerando sinal de entrada colorido.

entrada. Observando as Figs. 4.15, 4.16 e 4.17, e a Tabela 4.5, percebemos que o erro quadrático mínimo fica limitado pela atenuação finita na faixa de rejeição dos filtros protótipos usados na implementação dos bancos de análise do sinal de entrada e do sinal desejado.

Observando as Figs. 4.15 e 4.16 e a Tabela 4.5, podemos perceber que os valores teóricos do excesso de MSE estão de acordo com os resultados experimentais. Portanto, as expressões das Eqs. (4.51) e (4.54) podem ser úteis para selecionar o banco de filtros apropriado para um MSE desejado em regime permanente.

4.10 Conclusões

Neste capítulo derivamos uma estrutura adaptativa criticamente decimada que emprega um banco de filtros não-uniforme em árvore para decomposição do sinal de entrada. A adaptação é feita na menor taxa de amostragem dentre as empregadas nas várias subbandas, utilizando um algoritmo do tipo LMS com passo de adaptação normalizado. Apresentamos uma análise de convergência do algoritmo adaptativo

proposto, a partir da qual podemos estimar a taxa de convergência e o MSE teórico; e analisamos a complexidade computacional da estrutura. Simulações computacionais foram realizadas mostrando que uma significativa melhoria na taxa de convergência pode ser obtida para sinais de entrada coloridos com o algoritmo proposto quando comparado ao algoritmo LMS convencional em banda cheia e a outros algoritmos em subbandas.

Capítulo 5

Separação Cega de Sinais de Áudio

5.1 Introdução

Atualmente a tecnologia envolvida no reconhecimento de fala é de fundamental importância para a comunicação entre os seres humanos e as máquinas. Porém, a taxa de reconhecimento se degrada rapidamente quando várias pessoas falam ao mesmo tempo ou quando existe ruído de fundo (por exemplo, ar condicionado, barulhos de motores de máquinas, burburinho, etc.). O grande desafio é prover às máquinas a capacidade humana de distinguir uma fala desejada dentre diversas conversas (ou ruídos) paralelas. Este problema é denominado *cocktail-party* (festa-coquetel), onde se busca detectar uma fala individual dentre as misturas capturadas pelos microfones.

As técnicas de separação cega de fontes (BSS, *Blind Source Separation*) têm sido bastante estudadas na última década e apresentam inúmeras aplicações, entre as quais podemos destacar: sistemas de áudio [37], sensoriamento remoto [38], análise de sinais sísmicos [39], realce de imagens [40] e comunicações digitais [41]. Estas técnicas emergentes possibilitam a extração do sinal de uma fonte desejada a partir de uma mistura de sinais de várias fontes observadas. O termo “cega” se refere ao fato de que o sistema de misturas é desconhecido e de que há ausência de informações a respeito das fontes (posições, estruturas espectrais, etc.). De forma geral, podemos classificar os tipos de misturas em duas grandes classes: lineares e não-lineares. O grupo das misturas lineares é o contemplado neste trabalho, podendo ser subdividido em dois casos: misturas instantâneas e misturas convolutivas. Cada um destes casos

pode ainda ter três classificações: misturas sobre-determinadas, determinadas e sub-determinadas, nas quais o número P de misturas é superior, se iguala e é inferior ao número Q de fontes, respectivamente.

Para atingir este objetivo (estimar as fontes originais), geralmente é utilizada uma técnica computacional e estatística baseada na análise de componentes independentes (ICA, *Independent Component Analysis*), que possibilita a extração do sinal de cada uma das fontes a partir do sinal capturado pelos sensores, e que deve ser aplicada em sinais que possam ser modelados como misturas lineares de componentes estatisticamente independentes e não-gaussianos¹. As configurações sub-determinadas normalmente requerem o emprego de técnicas de análise de componentes esparsos (SCA, *Sparse Component Analysis*), que não serão abordadas neste trabalho.

O problema de BSS está muito bem resolvido para o caso particular de misturas lineares instantâneas e determinadas (a partir deste momento chamada simplesmente de misturas instantâneas), mostrada na Fig. 5.1, que envolvem apenas atenuações. Um grande desafio no momento é obter uma boa solução para o problema de misturas

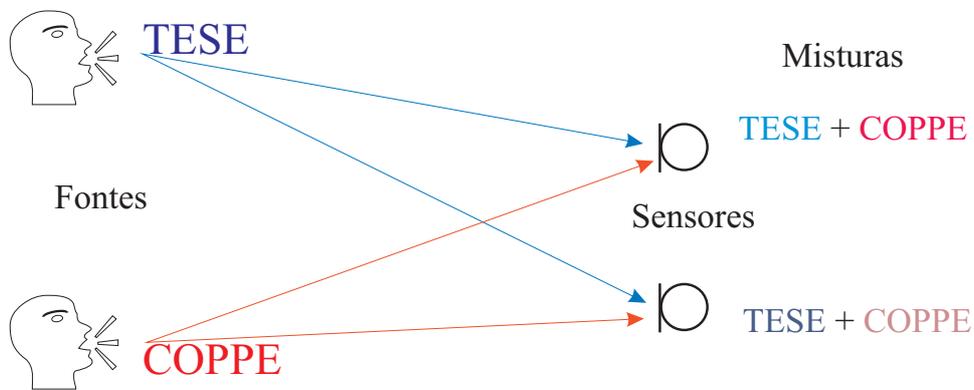


Figura 5.1: Mistura instantânea com $Q = 2$ fontes e $P = 2$ misturas.

lineares convolutivas determinadas (doravante denominadas misturas convolutivas) que além de atenuações envolvem atrasos e reverberação, como mostra a Fig. 5.2. Neste trabalho trataremos o caso de misturas convolutivas envolvendo sinais da fala humana e de música, permitindo assim modelar-se a reverberação sempre presente em ambientes ecóicos. Recentemente, muita pesquisa tem sido feita nesta área e

¹No caso de misturas instantâneas e determinadas admite-se no máximo uma fonte gaussiana.

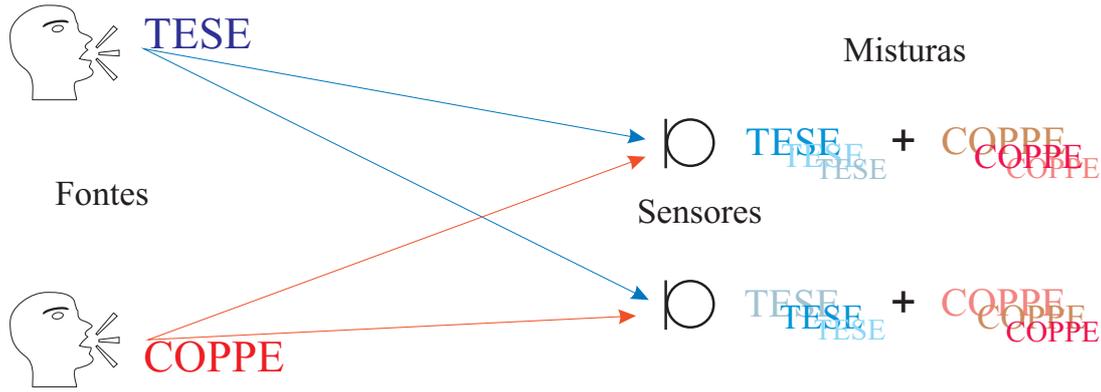


Figura 5.2: Mistura convolutiva com $Q = 2$ fontes e $P = 2$ misturas.

basicamente existem três maneiras de tratar o problema de misturas convolutivas: no domínio do tempo [42, 43], no domínio da frequência [44, 45] e uma solução híbrida [46]. O grande problema das soluções no domínio do tempo é a demanda computacional devido ao tempo de reverberação de uma sala, que sendo de tamanho médio², é da ordem de 0,25s. Para sinais amostrados em uma frequência de 8kHz, por exemplo, teríamos que modelar filtros de separação da ordem de 2.000 coeficientes³. Tentando amenizar este problema, existem propostas de desfazer a mistura no domínio da frequência, uma vez que a convolução no domínio do tempo é transformada em produto e pode ser vista como um caso de mistura instantânea em cada raia de frequência. As desvantagens são os problemas de escalamento e permutação entre as raias, a equivalência convolução-produto que só é válida no domínio discreto quando a convolução é circular e não linear como ocorre na prática, além do uso de blocos grandes de amostras dos sinais, necessários para estimar filtros longos. Neste caso, devido à não-estacionariedade do ambiente acústico (sistema de misturas) e dos sinais de áudio, não há amostras suficientes no domínio da frequência para uma correta estimativa das estatísticas em cada raia. Estas desvantagens podem limitar severamente o desempenho destes algoritmos. Existem, ainda, soluções mistas que combinam as soluções no domínio da frequência e no domínio do tempo para melhorar o desempenho da BSS com menor complexidade computacional. Neste cenário, algumas soluções em subbandas têm sido propostas, principalmente pela

²Uma sala de tamanho médio tem dimensões de aproximadamente 3,5 m de largura por 4,5 m de profundidade por 2,5 m de altura.

³Supondo que os filtros de separação sejam da mesma ordem que os filtros de mistura.

possibilidade de trabalhar com filtros de separação independentes e de ordem reduzida em cada canal, além da redução da taxa de amostragem. Estes métodos utilizam, normalmente, banco de filtros uniformes subamostrados com coeficientes complexos [47].

Nossa proposta é tratar o problema de misturas convolutivas e determinadas no domínio do tempo, decompondo os sinais em subbandas utilizando banco de filtros maximamente decimados com coeficientes reais. Através da decomposição e da decimação dos sinais observados, esperamos reduzir a complexidade computacional da estrutura; outra vantagem é a utilização de filtros com coeficientes reais que são atrativos em implementações em dispositivos para processamento de sinais (DSP, *Digital Signal Processor*). Um desafio será resolver o problema de permutação entre os sinais das diversas subbandas para recompor o sinal estimado das fontes corretamente e sem degradação de desempenho.

No caso de separação de fontes de áudio, alguns microfones (sensores) são colocados em diferentes posições de forma que cada um capture uma mistura das fontes originais ligeiramente atrasadas e com diferentes níveis de potência (diversidade espacial). Para situações do mundo real, podemos considerar sinais de áudio como sendo as fontes originais e uma sala ou um auditório como o sistema que realiza a mistura. Neste caso, os sinais capturados pelo microfones são afetados pela reverberação do ambiente. A Fig. 5.3 mostra o esquema geral de um sistema MIMO de separação cega de fontes. Considerando que o sistema desconhecido de misturas pode ser modelado por um conjunto de filtros FIR de comprimento U (mistura convolutiva), os sinais das Q fontes capturados pelos P microfones podem ser escritos como

$$x_p(n) = \sum_{q=1}^Q \sum_{k=0}^{U-1} g_{pq}(k) s_q(n-k) \quad \text{para } p = 1, \dots, P, \quad (5.1)$$

onde g_{pq} é o filtro que modela a função de transferência acústica do eco entre a q -ésima fonte e o p -ésimo sensor.

Para desfazer a mistura a partir dos sinais observados pelos microfones são estimados, através de um algoritmo de aprendizagem adaptativo, os coeficientes dos filtros FIR de separação w_{qp} de comprimento S de forma que os sinais de saída $y_q(n)$ tornem-se maximamente independentes. Portanto, o sinal estimado pela q -ésima

Existem, porém, alguns problemas inerentes à BSS conhecidos como permutação e escalamento. Estes fenômenos aparecem quando $\mathbf{C} = \mathbf{W}\mathbf{G} \neq \mathbf{I}$, como mostra o exemplo a seguir, supondo um sistema MIMO 5×5 (P sensores \times Q fontes):

$$\mathbf{C} = \begin{bmatrix} -4.3 & 0 & 0 & 0 & 0 \\ 0 & 5.7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2.2 \\ 0 & 0 & 0 & 1.5 & 0 \\ 0 & 0 & -3.7 & 0 & 0 \end{bmatrix}.$$

Portanto, o algoritmo de aprendizagem cego, ou seja, desprovido de qualquer conhecimento a respeito das fontes e do sistema de misturas, garante apenas que

$$y_q(n) = c_q s_p(n), \quad (5.8)$$

sendo c_q um valor constante⁴ que, diferente da unidade, indica escalamento da fonte, e $q \neq p$ indica a existência de permutação. Estes efeitos são facilmente contornados para misturas instantâneas, porém podem limitar o desempenho do algoritmo para misturas convolutivas como será visto na Seção 5.2.

5.1.2 Análise de Componentes Independentes

Análise de componentes independentes (ICA) é um método estatístico que foi introduzido originalmente no contexto de redes neurais [48],[49], [50], onde a matriz de separação \mathbf{W} era encarada como uma rede linear. Estudadas intensivamente desde meados da década de 90, estas técnicas têm sido largamente empregadas para promover a BSS de sinais de áudio que são capturados por alguns microfones, de ondas cerebrais gravadas a partir de múltiplos sensores, de sobreposição de imagens e de interferência entre sinais de rádio que chegam em uma estação móvel [51]. Alguns autores usam os termos ICA e BSS de forma intercambiável. Devido à diversidade de aplicações, a ICA tem se tornado uma boa ferramenta em diversas aplicações como processamento de sinais, comunicações sem fio, teoria da informação, redes neurais, etc.

No caso de separação cega de fontes, assumindo que as fontes originais são mutuamente independentes, a ICA consiste em estimar o sistema \mathbf{W} que desfaz a mistura

⁴Para que todas as fontes sejam recuperadas é necessário $c_q \neq 0$ para todo q .

e as fontes $s_l(n)$, a partir dos sinais observados $x_k(n)$.

O conceito de independência, mencionado acima, é mais forte que o conceito de decorrelação porque, enquanto este último envolve estatísticas de segunda ordem, o primeiro envolve as estatísticas de ordens maiores que dois, as quais exploram informações que não estão contidas na matriz de correlação. Supondo um problema de decorrelação não-linear de um sistema MIMO 2×2 sem nenhuma perda de generalidade, se a matriz de separação é do tipo $\mathbf{C} = \mathbf{W}\mathbf{G} = \mathbf{I}$, $y_1(n)$ e $y_2(n)$ são independentes e têm média zero, e a função não-linear $\Phi(\cdot)$ é uma função ímpar tal que $\Phi(y_1(n))$ também tem média zero, então

$$E[\Phi(y_1(n))y_2(n)] = E[\Phi(y_1(n))]E[y_2(n)] = 0. \quad (5.9)$$

A escolha da função não-linear é baseada em algumas teorias desenvolvidas para ICA, como por exemplo, minimização da informação mútua, maximização da não-gaussianidade e maximização da verossimilhança. A informação mútua é uma medida de independência estatística baseada na teoria da informação, que é sempre não-negativa. A minimização da informação mútua (ou maximização da entropia) pode ser interpretada como a maximização da independência entre as componentes. A segunda aproximação é baseada na maximização da não-gaussianidade (ou maximização da curtose). O teorema do limite central da teoria da probabilidade diz que a distribuição da soma de variáveis aleatórias não-gaussianas independentes tende a se tornar uma distribuição gaussiana, como mostra a Fig. 5.5. A Fig. 5.5 (a) apresenta a distribuição de um sinal de fala ($Q = 1$), a (b) exhibe a distribuição da soma de dois sinais de voz ($Q = 4$), a (c) mostra o histograma do somatório de $Q = 4$ fontes e a (d) exhibe a distribuição quase gaussiana da soma de 8 fontes. Portanto, as componentes independentes podem ser encontradas nas direções em que os dados são maximamente não-gaussianos. Na teoria da ICA, as variáveis aleatórias têm distribuição não-gaussiana, e esta hipótese é válida para sinais de áudio que possuem distribuições tipicamente supergaussianas (ver Fig. 5.5 (a)), ou seja, sua função densidade de probabilidade (pdf, *probability density function*) é relativamente grande em torno de zero se comparada à distribuição gaussiana (ver Fig. 5.5 (d)). Uma maneira tradicional de se avaliar a distribuição de uma variável aleatória é através de sua curtose, definida como [52]:

$$K(x) = E\{x^4\} - 3(E\{x^2\})^2. \quad (5.10)$$

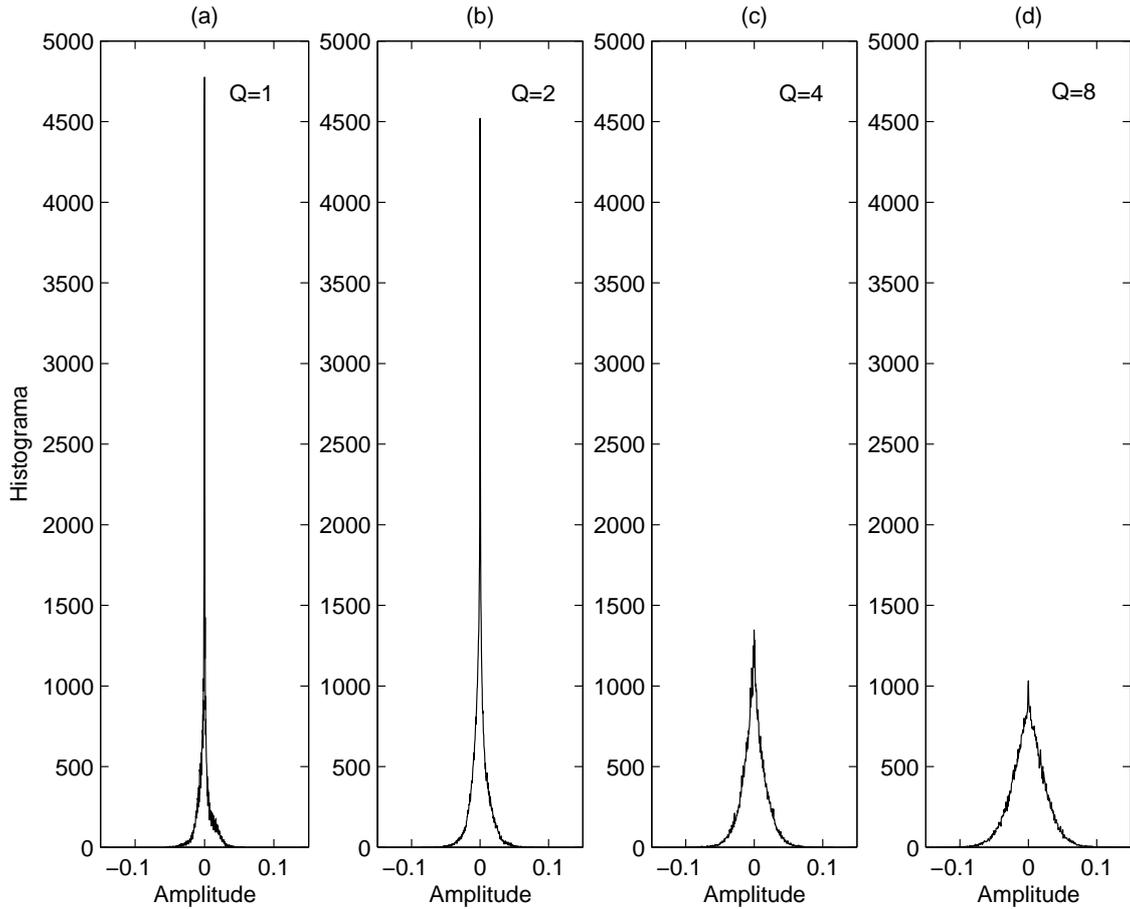


Figura 5.5: Exemplificação do teorema do limite central.

Para distribuições super-gaussianas $K(x) > 0$, para gaussianas $K(x) = 0$ e para sub-gaussianas $K(x) < 0$. A tabela 5.1 apresenta o valor da curtose para os exemplos mostrados na Fig. 5.5.

Tabela 5.1: Curtose para os exemplos mostrados na Fig. 5.5.

Q	1	2	4	8	∞
$K(x)$	3,96	3,24	1,57	0,94	0

A terceira aproximação é baseada na maximização da verossimilhança. A estimação da máxima verossimilhança (ML, *Maximum Likelihood*) é a técnica mais popular utilizada para ICA, e pode ser entendida como a estimativa de parâmetros que proporcionam a maior probabilidade para as observações. A ML está intimamente relacionada ao princípio da maximização da informação em redes neurais. Este princípio é baseado na maximização da entropia de saída ou do fluxo de informação de uma rede neural de saídas não-lineares. A maximização da informação

mútua é equivalente à maximização da entropia da saída, que por sua vez é equivalente à maximização da verossimilhança. É interessante perceber que, na verdade, as três aproximações descritas acima são equivalentes [43], [53] e [54]. Omitindo momentaneamente o índice (n) para facilitar a notação, a informação mútua $I(y_1, y_2)$ entre as saídas $y_1(n)$ e $y_2(n)$ é expressa por

$$I(y_1, y_2) = H(y_1) + H(y_2) - H(y_1, y_2) \quad (5.11)$$

sendo $H(y_k)$ a entropia marginal e $H(y_1, y_2)$ a entropia conjunta da saída. Uma representação gráfica da informação mútua pode ser vista na Fig. 5.6. Observando

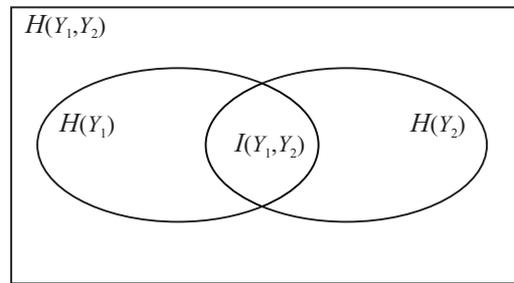


Figura 5.6: Representação gráfica da informação mútua.

a Eq. (5.11) percebe-se que a informação mútua $I(y_1, y_2)$ é minimizada quando os dois primeiros termos são minimizados ou quando o terceiro termo é maximizado. Sinais gaussianos maximizam os dois primeiros termos, portanto, maximizar a não-gaussianidade nos leva à minimização da informação mútua. Por outro lado, a maximização da entropia conjunta da saída maximiza o último termo da Eq. (5.11).

Para encontrar a matriz \mathbf{W} na Eq. (5.6) que minimize a informação mútua, maximize a não-gaussianidade ou maximize a verossimilhança da saída, podemos utilizar o método do gradiente [1]. Um algoritmo muito simples foi derivado em [55], o qual maximiza a entropia dada por

$$H(y) = -E[\ln(p(y))] = - \int_{-\infty}^{\infty} p(y) \ln(p(y)) dy, \quad (5.12)$$

sendo $(p(y))$ a pdf de y .

Inicialmente, assumindo um sistema do tipo $y = wx$ (com uma entrada e uma saída), no qual uma transformação não-linear aplicada à saída y ($\phi(y)$) é monotonicamente crescente ou decrescente [55] (ou seja, tem uma única inversa) o pdf da

saída pode ser escrita como uma função da pdf da entrada ($p(x)$):

$$p(y) = \frac{p(x)}{|\partial y / \partial x|}. \quad (5.13)$$

Substituindo a Eq. (5.13) na Eq. (5.12) obtém-se

$$H(y) = E \left[\ln \left| \frac{\partial y}{\partial x} \right| \right] - E [\ln(p(x))]. \quad (5.14)$$

Como o segundo termo não é afetado pelo parâmetro w , para maximizar a entropia de y é necessário maximizar apenas o primeiro termo de Eq. (5.14). Isto pode ser feito calculando o gradiente da entropia em relação ao coeficiente w , ou seja,

$$\nabla_w H \approx \frac{\partial H}{\partial w} = \frac{\partial}{\partial w} \left(\ln \left| \frac{\partial y}{\partial x} \right| \right) = \left(\frac{\partial y}{\partial x} \right)^{-1} \frac{\partial}{\partial w} \left(\frac{\partial y}{\partial x} \right). \quad (5.15)$$

Para sinais de áudio, uma função não-linear bastante utilizada é $y = \phi(y) = \tanh(y)$, para qual temos

$$\begin{aligned} \frac{\partial \phi(y)}{\partial x} &= \frac{\partial \left(\frac{e^y - e^{-y}}{e^y + e^{-y}} \right)}{\partial x} = \frac{\partial}{\partial x} \left(\frac{e^{wx}}{e^{wx} + e^{-wx}} \right) - \frac{\partial y}{\partial x} \left(\frac{e^{-wx}}{e^{wx} + e^{-wx}} \right) \\ &= \frac{w(e^{wx} + e^{-wx})^2}{(e^{wx} + e^{-wx})^2} - \frac{w(e^{wx} - e^{-wx})^2}{(e^{wx} + e^{-wx})^2} = w(1 - \phi^2(y)) \end{aligned} \quad (5.16)$$

e

$$\begin{aligned} \frac{\partial}{\partial w} \left(\frac{\partial \phi(y)}{\partial x} \right) &= 1 - \frac{\partial}{\partial w} (w\phi^2(y)) \\ &= 1 - \phi^2(y) - 2w\phi(y) \frac{\partial \phi(y)}{\partial w} \\ &= 1 - \phi^2(y) - 2x\phi(y)w(1 - \phi^2(y)) \\ &= (1 - \phi^2(y)) - 2x\phi(y)w(1 - \phi^2(y)). \end{aligned} \quad (5.17)$$

Substituindo estes resultados na Eq. (5.18), o gradiente da entropia é dado por

$$\nabla_w H \approx \frac{1}{w} - 2x\phi(y) \quad (5.18)$$

Generalizando para um sistema de P entradas e Q saídas, teremos o gradiente estocástico da função objetivo (entropia) dado por

$$\nabla_{\mathbf{W}} H \approx [\mathbf{W}^H]^{-1} - 2E [\Phi(\mathbf{y})\mathbf{x}^H]. \quad (5.19)$$

Para aumentar a estabilidade e a velocidade de convergência deste algoritmo e evitar problemas de mal-condicionamento da matriz de mistura, foi proposto em

[49] o método do gradiente natural, que pode ser visto como um caso especial de decorrelação não-linear. O gradiente natural (NG, *Natural Gradient*) é obtido a partir do gradiente tradicional, através da relação

$$\nabla_{\mathbf{W}}^{NG} H = \nabla_{\mathbf{W}} H \mathbf{W} \mathbf{W}^H \mathbf{W}. \quad (5.20)$$

Substituindo a Eq. (5.19) na Eq. (5.20), obtemos

$$\nabla_{\mathbf{W}}^{NG} H = \{ \mathbf{I} - 2E [\Phi(\mathbf{y})\mathbf{y}^H] \} \mathbf{W}. \quad (5.21)$$

A equação de atualização dos coeficientes da matriz de separação que desfaz a mistura, baseada no método do gradiente natural, é dada por

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu \{ \mathbf{I} - 2E [\Phi(\mathbf{y}(n))\mathbf{y}^H(n)] \} \mathbf{W}(n). \quad (5.22)$$

5.2 Separação Cega de Fontes para Misturas Convolutivas

Para misturas convolutivas de fontes, os algoritmos descritos na seção anterior não apresentam um bom desempenho devido à reverberação do ambiente. Uma maneira simples de melhorar a separação das fontes é diagonalizar a matriz de correlação das saídas \mathbf{R}_{YY} , a qual para um sistema linear MIMO $P \times Q$ com $P = Q$ é dada por

$$\mathbf{R}_{YY} = \begin{bmatrix} \langle \Phi(Y_1)Y_1^H \rangle & \langle \Phi(Y_1)Y_2^H \rangle & \dots & \langle \Phi(Y_1)Y_P^H \rangle \\ \langle \Phi(Y_2)Y_1^H \rangle & \langle \Phi(Y_2)Y_2^H \rangle & \dots & \langle \Phi(Y_2)Y_P^H \rangle \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \langle \Phi(Y_P)Y_1^H \rangle & \langle \Phi(Y_P)Y_2^H \rangle & \dots & \langle \Phi(Y_P)Y_P^H \rangle \end{bmatrix}, \quad (5.23)$$

sendo $\langle \cdot \rangle$ o operador de média estatística. Os coeficientes $w_{qp}(n)$ dos filtros que desfazem a mistura devem convergir para valores que minimizam a informação mútua entre as saídas, os quais correspondem aos elementos que estão fora da diagonal principal da matriz de correlação, ou seja,

$$\langle \Phi(Y_i)Y_j^H \rangle = 0 \quad \text{para } i \neq j. \quad (5.24)$$

Já os elementos da diagonal principal, que controlam o escalamento das saídas, devem ser restritos a constantes apropriadas c_i , ou seja:

$$\langle \Phi(Y_i)Y_i^H \rangle = c_i. \quad (5.25)$$

A equação iterativa de atualização dos coeficientes dos filtros de separação baseada no método do gradiente é dada por

$$\mathbf{W}_{i+1} = \mathbf{W}_i + \mu \Delta \mathbf{W}_i, \quad (5.26)$$

sendo

$$\Delta \mathbf{W}_i = \begin{bmatrix} c_1 - \langle \Phi(Y_1)Y_1^H \rangle & \langle \Phi(Y_1)Y_2^H \rangle & \dots & \langle \Phi(Y_1)Y_P^H \rangle \\ \langle \Phi(Y_2)Y_1^H \rangle & c_2 - \langle \Phi(Y_2)Y_2^H \rangle & \dots & \langle \Phi(Y_2)Y_P^H \rangle \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \langle \Phi(Y_P)Y_1^H \rangle & \langle \Phi(Y_P)Y_2^H \rangle & \dots & c_P - \langle \Phi(Y_P)Y_P^H \rangle \end{bmatrix}. \quad (5.27)$$

No procedimento descrito acima, se $\Phi(Y_i) = Y_i$ temos um caso de decorrelação simples que não é suficiente para garantir a independência dos sinais de saída da estrutura BSS. Contudo, se as fontes são não-estacionárias, podemos usar estatísticas de segunda ordem (SOS, *Second Order Statistics*) considerando vários blocos de amostras dos sinais de saída. Este método é conhecido como decorrelação não-estacionária (ND, *Nonstationary Decorrelation*) [56]. Outro método, para fontes coloridas, que também usa SOS considera a decorrelação de blocos de sinais atrasados no tempo (TDD, *Time-Delayed Decorrelation*), ou seja,

$$\langle \Phi(Y_i)Y_j^H \rangle = \langle Y_i(m)Y_j(m + \tau_i)^H \rangle = 0, \quad (5.28)$$

que também pode resolver o problema da BSS. Usando estes tipos de decorrelação (ND e TDD) tem-se informação extra suficiente para estimação dos filtros de separação, não sendo necessárias informações de estatísticas de ordens mais elevadas para garantir a independências entre as amostras das estimativas das fontes [43].

Por outro lado, quando consideramos $\Phi(Y_i) = \tanh(Y_i)$ temos

$$\langle \Phi(Y_i)Y_j^H \rangle = \langle \tanh(Y_i)Y_j^H \rangle = 0, \quad (5.29)$$

que pode ser visto como um caso de decorrelação não-linear, como mostra a expansão de Taylor de $\tanh(Y_i)$, dada por

$$\left\langle \left(Y_i - \frac{Y_i^3}{3} + \frac{2Y_i^5}{15} - \frac{17Y_i^7}{315} + \dots \right) Y_2^H \right\rangle = 0. \quad (5.30)$$

Este método utiliza, portanto, estatísticas de mais alta ordem (HOS, *Higher Order Statistics*) ou decorrelação não-linear, e também pode resolver o problema de separação cega de misturas convolutivas.

5.2.1 Separação Cega de Fontes no Domínio do Tempo

A separação cega de fontes é baseada na hipótese fundamental de que as diferentes fontes de sinais são estatisticamente mutuamente independentes. No mundo real, devido à reverberação do ambiente, os sinais originais das fontes são filtrados por um sistema MIMO linear antes de serem capturados pelos microfones. A partir deste momento, vamos assumir que o número de fontes é igual ao número de microfones como mostrado na Fig. 5.7 e apresentar a solução baseada em estatísticas de segunda ordem proposta em [3].

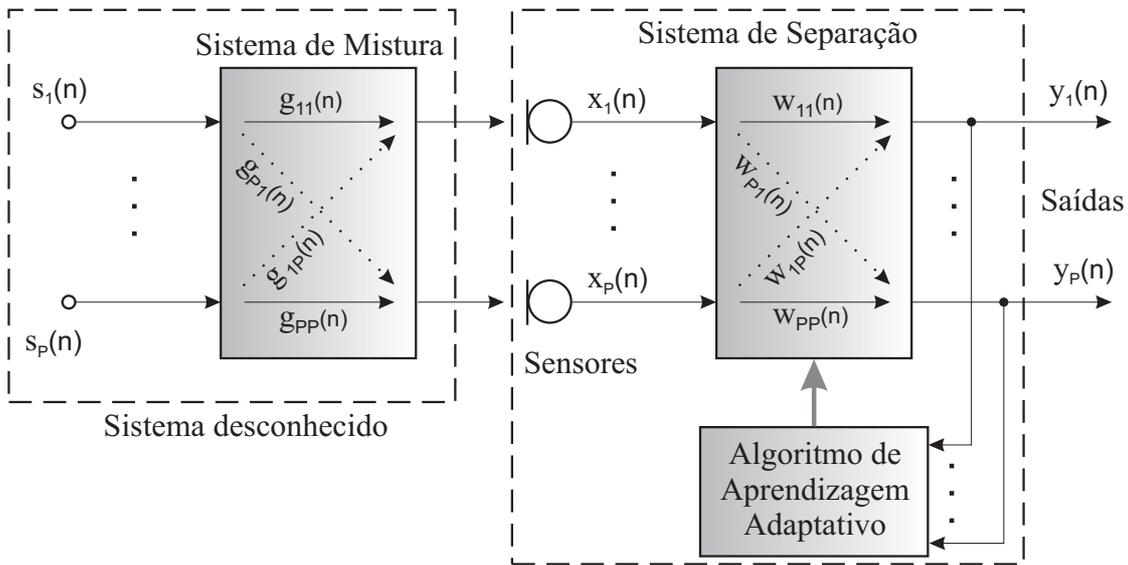


Figura 5.7: Sistema linear MIMO de P entradas e P saídas para mistura convolutiva.

No problema de BSS, estamos interessados no sistema que desfaz a mistura que, neste caso, é descrito por

$$y_q(n) = \sum_{p=1}^P \sum_{k=0}^{S-1} \tilde{w}_{pq}(k) x_p(n-k). \quad (5.31)$$

Por conveniência, mantendo a notação utilizada em [3], o desenvolvimento a seguir utilizará coeficientes $\tilde{w}_{pq}(k)$, onde $\tilde{w}_{pq}(k) = w_{qp}(k)$.

Para sinais coloridos e não-estacionários, como sinais de áudio, vimos na Subseção 5.2 que o problema pode ser resolvido diagonalizando a matriz de correlação de saída considerando múltiplos blocos em diferentes instantes de tempo. Estendendo a formulação dos sinais de saída para uma forma matricial, podemos

descrever o sinal da q -ésima saída no instante n como

$$y_q(n) = \sum_{p=1}^P \mathbf{x}_p^T(n) \tilde{\mathbf{w}}_{pq}, \quad (5.32)$$

sendo $\mathbf{x}_p(n) = [x_p(n), x_p(n-1), \dots, x_p(n-S+1)]^T$ o vetor que contém as últimas S amostras capturadas pelo p -ésimo microfone e $\tilde{\mathbf{w}}_{pq}(n) = [\tilde{w}_{pq}(0), \tilde{w}_{pq}(1), \dots, \tilde{w}_{pq}(S-1)]^T$ o vetor que contém os S coeficientes do filtro FIR que modela o caminho do p -ésimo sensor à q -ésima saída.

Dois novos parâmetros necessários para generalização da formulação são o número D de atrasos no tempo levados em conta no cálculo da correlação ($1 \leq D \leq S$) e o tamanho N do bloco de sinal de saída.

Da Eq. (5.32) podemos descrever o vetor que contém um bloco de amostras de tamanho N da q -ésima saída no instante m como

$$\begin{aligned} \mathbf{y}_q(m) &= \sum_{p=1}^P \hat{\mathbf{X}}_p^T(m) \tilde{\mathbf{w}}_{pq} \\ &= [y_q(mS) \dots y_q(mS + N - 1)]^T, \end{aligned} \quad (5.33)$$

sendo

$$\begin{aligned} \hat{\mathbf{X}}_p(m) &= [\mathbf{x}_p(mS) \dots \mathbf{x}_p(mS + N - 1)]^T \\ &= \begin{bmatrix} x_p(mS) & \dots & x_p(mS - S + 1) \\ x_p(mS + 1) & \ddots & x_p(mS - S + 2) \\ \vdots & \ddots & \vdots \\ x_p(mS + N - 1) & \dots & x_p(mS - S + N) \end{bmatrix} \end{aligned} \quad (5.34)$$

a matriz Toeplitz de dimensão $S \times N$ que contém os S blocos com as versões atrasadas das amostras do sinal capturado pelo p -ésimo sensor.

A seguir, a Eq. (5.33) pode ser estendida para incorporar as amostras dos D blocos atrasados no tempo. Assim sendo, a matriz com a informação da q -ésima saída, de dimensão $N \times D$, é dada por

$$\mathbf{Y}_q(m) = \sum_{p=1}^P \mathbf{X}_p^T(m) \tilde{\mathbf{W}}_{pq} \quad (5.35)$$

$$= \begin{bmatrix} y_q(mS) & \dots & y_q(mS - D + 1) \\ y_q(mS + 1) & \ddots & y_q(mS - D + 2) \\ \vdots & \ddots & \vdots \\ y_q(mS + N - 1) & \dots & y_q(mS - D + N) \end{bmatrix}. \quad (5.36)$$

Para assegurar a convolução linear de todos os elementos de $\mathbf{Y}_q(m)$ até o número máximo de atrasos de tempo $D = S$ [42], são necessários dois blocos de sinais de entrada $\hat{\mathbf{X}}_p^T$. Com isso, as dimensões de $\mathbf{X}_p(m)$ e $\tilde{\mathbf{W}}_{pq}$ são $N \times 2S$ e $2S \times D$, respectivamente. As matrizes $\mathbf{X}_p(m)$ são obtidas dobrando-se o tamanho da matriz $\hat{\mathbf{X}}_p$, ou seja,

$$\mathbf{X}_p(m) = \left[\hat{\mathbf{X}}_p^T(m), \hat{\mathbf{X}}_p^T(m-1) \right]. \quad (5.37)$$

A matriz $\hat{\mathbf{X}}_p^T(m-1)$ também é uma matriz Toeplitz, de forma que a primeira linha da matriz $\mathbf{X}_p(m)$ contém $2S$ amostras do p -ésimo sinal de entrada e cada linha subsequente é obtida pelo deslocamento da linha anterior para a direita de uma amostra, contendo uma nova amostra por linha. A matriz $\tilde{\mathbf{W}}_{pq}$ é do tipo *Sylvester* de dimensão $2S \times D$, definida como

$$\tilde{\mathbf{W}}_{pq} = \begin{bmatrix} \tilde{w}_{pq}(0) & 0 & \dots & 0 \\ \tilde{w}_{pq}(1) & \tilde{w}_{pq}(0) & \ddots & \vdots \\ \vdots & \tilde{w}_{pq}(1) & \ddots & 0 \\ \tilde{w}_{pq}(S-1) & \vdots & \ddots & \tilde{w}_{pq}(0) \\ 0 & \tilde{w}_{pq}(S-1) & \ddots & \tilde{w}_{pq}(1) \\ \vdots & 0 & \ddots & \vdots \\ 0 & \dots & 0 & \tilde{w}_{pq}(S-1) \\ 0 & \dots & 0 & 0 \\ \vdots & \dots & \dots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix}, \quad (5.38)$$

que tem as últimas $S - D + 1$ linhas formadas por zeros para assegurar a compatibilidade com $\mathbf{X}_p(m)$ no caso geral de $1 \leq D \leq S$. Para permitir uma notação conveniente do algoritmo de atualização, que será vista a seguir, podemos reescrever a Eq. (5.35) de uma forma mais compacta, ou seja,

$$\mathbf{Y}(m) = \mathbf{X}(m)\tilde{\mathbf{W}}, \quad (5.39)$$

sendo

$$\mathbf{Y}(m) = [\mathbf{Y}_1(m) \dots \mathbf{Y}_P(m)] \quad (5.40)$$

a matriz de dimensão $N \times PD$ que contém os blocos de sinais de saída de todos os canais,

$$\mathbf{X}(m) = [\mathbf{X}_1(m) \dots \mathbf{X}_P(m)] \quad (5.41)$$

a matriz de ordem $N \times 2SP$ contendo todos os blocos atrasados no tempo de todos os sensores, e

$$\tilde{\mathbf{W}} = \begin{bmatrix} \tilde{\mathbf{W}}_{11} & \dots & \tilde{\mathbf{W}}_{1P} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{W}}_{P1} & \dots & \tilde{\mathbf{W}}_{PP} \end{bmatrix} \quad (5.42)$$

a matriz de dimensão $2SP \times DP$ que contém todos os coeficientes de todos os filtros de separação.

5.2.1.1 Função Objetivo e Algoritmo de Atualização

Novamente, por conveniência, mantendo a notação utilizada em [3], o desenvolvimento a seguir utilizará coeficientes $\tilde{g}_{qp}(k)$, onde $\tilde{g}_{qp}(k) = g_{pq}(k)$.

Analogamente ao sistema de separação descrito pela Eq. (5.39), o sistema de mistura pode ser modelado por $\mathbf{X}(m) = \mathbf{S}(m)\tilde{\mathbf{G}}$, sendo $\mathbf{S}(m)$ a matriz de dimensão $N \times P(U+S-1)$ que contém as versões atrasadas dos sinais das fontes e $\tilde{\mathbf{G}}$ a matriz de mistura do tipo *Sylvester* de ordem $P(U+S-1) \times 2PS$ contendo os coeficientes da resposta ao impulso de todos os filtros $\tilde{g}_{qp}(n)$. Estas dimensões resultam, novamente, da condição de linearidade das convoluções realizadas. Portanto, é possível obter uma matriz diagonal por blocos $\mathbf{C} = \tilde{\mathbf{G}}\tilde{\mathbf{W}}$, tal que $\mathbf{C} - \text{bdiag } \mathbf{C} = \text{boff } \mathbf{C} = \mathbf{0}$. O operador *bdiag* opera sobre uma matriz formada por submatrizes, zerando todas as submatrizes que não pertencem a diagonal principal. Analogamente, o operador *boff* zera todas as submatrizes localizadas na diagonal principal.

Para alcançar este objetivo, usando estatísticas de segunda ordem, os sinais estocásticos são descritos pela sequência de matrizes reduzidas de correlação temporal com atrasos baseada na formulação matricial apresentada anteriormente. Estas matrizes são definidas por

$$\mathbf{R}_{xx}(m) = \mathbf{X}^H(m)\mathbf{X}(m) \quad (5.43)$$

e

$$\mathbf{R}_{yy}(m) = \mathbf{Y}^H(m)\mathbf{Y}(m) \quad (5.44)$$

tendo dimensões $2PS \times 2PS$ e $PD \times PD$, respectivamente. Para que a Eq. (5.44) tenha posto completo, é necessário que o comprimento do bloco do vetor de saída seja $N \geq D$.

A função objetivo, baseada na sequência de matrizes de correlação reduzidas e atrasadas, é dada por

$$\mathfrak{S}(m) = \sum_{i=0}^{\infty} \beta(i, m) \{ \log [\det (\text{bdiag}(\mathbf{Y}^H(i)\mathbf{Y}(i)))] - \log [\det(\mathbf{Y}^H(i)\mathbf{Y}(i))] \} \quad (5.45)$$

cujo ponto de equilíbrio corresponde exatamente à solução da BSS desejada, ou seja boff $\mathbf{C} = \mathbf{0}$. O parâmetro β é um valor constante normalizado de acordo com $\sum_{i=0}^{\infty} \beta(i, m) = 1$. Usando a formulação matricial da Eq. (5.39) para calcular as matrizes de correlação temporal reduzidas da Eq. (5.44), a função objetivo inclui todos os D atrasos no tempo de todas as autocorrelações e correlações cruzadas dos sinais de saída do sistema BSS.

Considerando um algoritmo baseado no método do gradiente, a equação recursiva para atualização dos coeficientes dos filtros que desfazem a mistura é dada por

$$\tilde{\mathbf{W}}(m+1) = \tilde{\mathbf{W}}(m) - \mu \nabla_{\tilde{\mathbf{W}}} \mathfrak{S}(m). \quad (5.46)$$

Recorrendo à formulação do gradiente natural [49] que é mais robusta e menos pesada computacionalmente⁵, obtém-se a seguinte equação recursiva para atualização dos coeficientes:

$$\tilde{\mathbf{W}}(m+1) = \tilde{\mathbf{W}}(m) - \mu \nabla_{\tilde{\mathbf{W}}}^{NG} \mathfrak{S}(m), \quad (5.47)$$

sendo o gradiente natural da função objetivo (Eq. (5.45)) dado por

$$\begin{aligned} \nabla_{\tilde{\mathbf{W}}}^{NG} \mathfrak{S}(m) &= \tilde{\mathbf{W}} \tilde{\mathbf{W}}^H \nabla_{\tilde{\mathbf{W}}} \mathfrak{S}(m) \\ &= 2 \sum_{i=0}^{\infty} \beta(i, m) \tilde{\mathbf{W}} \{ \mathbf{R}_{yy}(i) - \text{bdiag} \mathbf{R}_{yy}(i) \} \text{bdiag}^{-1} \mathbf{R}_{yy}(i) \end{aligned} \quad (5.48)$$

e μ o passo de adaptação do algoritmo.

O operador $\text{bdiag}(\cdot)$ interpreta a matriz ao qual é aplicado como uma composição de submatrizes, zerando todas as submatrizes que não pertencem à sua diagonal principal. Para exemplificar este operador, vamos supor um sistema com 3 fontes.

⁵A modificação proposta no gradiente remove os efeitos de mau-condicionamento da matriz de misturas \mathbf{G} e promovendo um desempenho uniforme devido à propriedade da equivariância obtida no gradiente natural, ou seja, a convergência do algoritmo de separação fica independente das fontes [42].

Neste caso, a matriz $\mathbf{R}_{\mathbf{y}\mathbf{y}}(m)$ pode ser descrita como:

$$\mathbf{R}_{\mathbf{y}\mathbf{y}}(m) = \begin{bmatrix} \mathbf{Y}_1^H(m)\mathbf{Y}_1(m) & \mathbf{Y}_1^H(m)\mathbf{Y}_2(m) & \mathbf{Y}_1^H(m)\mathbf{Y}_3(m) \\ \mathbf{Y}_2^H(m)\mathbf{Y}_1(m) & \mathbf{Y}_2^H(m)\mathbf{Y}_2(m) & \mathbf{Y}_2^H(m)\mathbf{Y}_3(m) \\ \mathbf{Y}_3^H(m)\mathbf{Y}_1(m) & \mathbf{Y}_3^H(m)\mathbf{Y}_2(m) & \mathbf{Y}_3^H(m)\mathbf{Y}_3(m) \end{bmatrix}. \quad (5.50)$$

As matrizes $\mathbf{Y}_i^H(m)\mathbf{Y}_i(m)$ (com $i = 1, 2$ e 3) são as matrizes de autocorrelação da i -ésima saída, enquanto que as matrizes $\mathbf{Y}_i^H(m)\mathbf{Y}_j(m)$, com $i \neq j$, são as matrizes de correlação cruzada entre a i -ésima e a j -ésima saída. É natural subdividir a matriz $\mathbf{R}_{\mathbf{y}\mathbf{y}}(m)$ em submatrizes, sendo as submatrizes de autocorrelação as pertencentes à diagonal principal da matriz de submatrizes. Assim, $\text{bdiag } \mathbf{R}_{\mathbf{y}\mathbf{y}}(m)$ produz o seguinte resultado:

$$\text{bdiag } \mathbf{R}_{\mathbf{y}\mathbf{y}}(m) = \begin{bmatrix} \mathbf{Y}_1^H(m)\mathbf{Y}_1(m) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}_2^H(m)\mathbf{Y}_2(m) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Y}_3^H(m)\mathbf{Y}_3(m) \end{bmatrix}, \quad (5.51)$$

onde $\mathbf{0}$ é uma matriz nula, com as mesmas dimensões que $\mathbf{Y}_i^H(m)\mathbf{Y}_j(m) = \mathbf{R}_{\mathbf{y}_i\mathbf{y}_j}(m)$.

Durante o processo de atualização dos coeficientes é necessário garantir a estrutura de matriz *Sylvester* $\tilde{\mathbf{W}}(m+1)$. O uso indiscriminado de um gradiente que atue sobre toda a matriz pode destruir esta característica, retirando a redundância que permite uma relação biunívoca entre as matrizes $\tilde{\mathbf{W}}_{pq}$ e os filtros correspondentes ($\tilde{\mathbf{w}}_{pq}$). Isto é facilmente imposto selecionando uma das colunas da matrizes $\tilde{\mathbf{W}}_{pq}$ que contenha todos os coeficientes dos filtros $\tilde{w}_{pq}(k)$ (para $k = 0, \dots, S-1$) e gerar $\nabla_{\tilde{\mathbf{W}}}^{NG} \mathfrak{S}(m)$ de acordo com a Eq. (5.38). Em [3] é mostrado que a escolha dos S primeiros elementos da primeira coluna de $\tilde{\mathbf{W}}_{pq}$ é a melhor opção para propósitos de otimização.

O algoritmo *off-line* tipo batelada⁶ (*batch*) para ajustar os coeficientes dos filtros de separação, considerando um sistema com duas fontes e dois sensores (*TITO*, *Two Input Two Output*), é dado por:

$$\tilde{\mathbf{W}}(i) = \tilde{\mathbf{W}}(i-1) - \frac{2\mu}{b} \sum_{m=1}^b \begin{bmatrix} \tilde{\mathbf{W}}_{12}\mathbf{R}_{y_2y_1}\mathbf{R}_{y_1y_1}^{-1} & \tilde{\mathbf{W}}_{11}\mathbf{R}_{y_1y_2}\mathbf{R}_{y_2y_2}^{-1} \\ \tilde{\mathbf{W}}_{22}\mathbf{R}_{y_2y_1}\mathbf{R}_{y_1y_1}^{-1} & \tilde{\mathbf{W}}_{21}\mathbf{R}_{y_1y_2}\mathbf{R}_{y_2y_2}^{-1} \end{bmatrix} \quad (5.52)$$

⁶O algoritmo utilizado é do tipo batelada, ou seja, todo o sinal é processado para gerar uma iteração *off-line*.

sendo \mathbf{R}_{pq} , de dimensão $D \times D$, uma submatriz de \mathbf{R}_{yy} (Eq. (5.44)), i o número de iterações (*off-line*) e μ o passo de adaptação do algoritmo.

Para reduzir a complexidade computacional do algoritmo, os fatores de normalização $\mathbf{R}_{qq}^{-1}(m)$ podem ser simplificados considerando $\mathbf{R}_{qq}(m)$ uma matriz diagonal [57], ou seja,

$$\mathbf{R}_{qq}(m) \approx \text{diag} \{ \mathbf{R}_{qq}(m) \}. \quad (5.53)$$

Desta maneira, \mathbf{R}_{qq}^{-1} (Eq. (5.52)) pode ser obtida invertendo os elementos de sua diagonal, que corresponde ao inverso da potência dos blocos atrasados no tempo do sinal da q -ésima saída (ver Eq. (5.33)). Com o objetivo de reduzir ainda mais o custo computacional, foi proposto em [58] reduzir o fator de normalização a um escalar. Neste caso,

$$\mathbf{R}_{qq}(m) \approx (\mathbf{y}_q(m))^T \mathbf{y}_q(m) \mathbf{I}, \quad (5.54)$$

com $\mathbf{y}_q^k(m)$ dado pela Eq. (5.33), sendo sua inversa reduzida ao inverso da potência do bloco mais recente de amostras da saída $y_q(n)$.

5.2.1.2 Complexidade Computacional

Conforme descrito na Seção 4.8, vamos considerar como complexidade computacional o número de multiplicações reais por bloco do sinal de entrada (NMPB) necessárias para atualizar os coeficientes dos filtros FIR de separação, considerando implementações na forma direta.

O NMPB total requerido pela estrutura em banda cheia, considerando um sistema linear MIMO $P \times P$ (P sensores e P fontes), é dado por

$$\text{NMPB}_{\text{FB}} = 2P^2NSD + P^2(N + DN) + P^2 \left(\left(\sum_{j=1}^S j \right) + S \right). \quad (5.55)$$

A primeira parcela diz respeito ao custo para gerar todos os blocos atrasados no tempo de todas as saídas (Eq. (5.39)). O segundo fator representa o custo da obtenção das submatrizes $\mathbf{R}_{y_q y_p}(m)$ da matriz \mathbf{R}_{yy} (ver Eq. (5.52)) para o cálculo do gradiente. O terceiro termo indica o custo da adaptação considerando apenas a atualização das S primeiras linhas da primeira coluna da matriz $\tilde{\mathbf{W}}_{pq}(m)$ (Eq. (5.38)).

Para filtros de separação de ordens altas, a equação acima pode ser simplificada considerando somente os termos dominantes, resultando em

$$\text{NMPB}_{\text{FB}} \approx 4P^2S^3. \quad (5.56)$$

As expressões acima foram obtidas considerando o fator de normalização simplificado da Eq. (5.54), $D = S$ e $N = 2D$.

5.2.2 Separação Cega de Fontes no Domínio da Frequência

O método de separação cega de fontes no domínio da frequência tem por objetivo transformar a mistura convolutiva no domínio do tempo em um problema de mistura instantânea no domínio da frequência, aplicar a solução mostrada na Eq. (5.22) (adaptada para valores complexos) em cada raia de frequência e resolver a BSS [3], [44], [45], [46]. Portanto, os sinais capturados pelos microfones no domínio da transformada podem ser considerados misturas lineares, ou seja:

$$\mathbf{X}(\omega) = \mathbf{G}(\omega)\mathbf{S}(\omega), \quad (5.57)$$

sendo

$$\mathbf{X}(\omega) = [X_1(\omega) \ \dots \ X_P(\omega)]^T \quad (5.58)$$

o vetor que contém a transformada de *Fourier* dos P sinais observados,

$$\mathbf{S}(\omega) = [S_1(\omega) \ \dots \ S_Q(\omega)]^T \quad (5.59)$$

o vetor que contém a transformada de *Fourier* dos sinais das Q fontes e

$$\mathbf{G}(\omega) = \begin{bmatrix} G_{11}(\omega) & \dots & G_{1Q}(\omega) \\ \vdots & \ddots & \vdots \\ G_{P1}(\omega) & \dots & G_{PQ}(\omega) \end{bmatrix} \quad (5.60)$$

a matriz contendo a resposta na frequência ω dos filtros FIR $g_{pq}(n)$ que misturam os sinais de todas as fontes $s_q(n)$.

A Fig. 5.8 mostra o esquema geral da BSS no domínio da frequência, onde os sinais observados $x_p(n)$ no domínio do tempo são “janelados” e convertidos para os

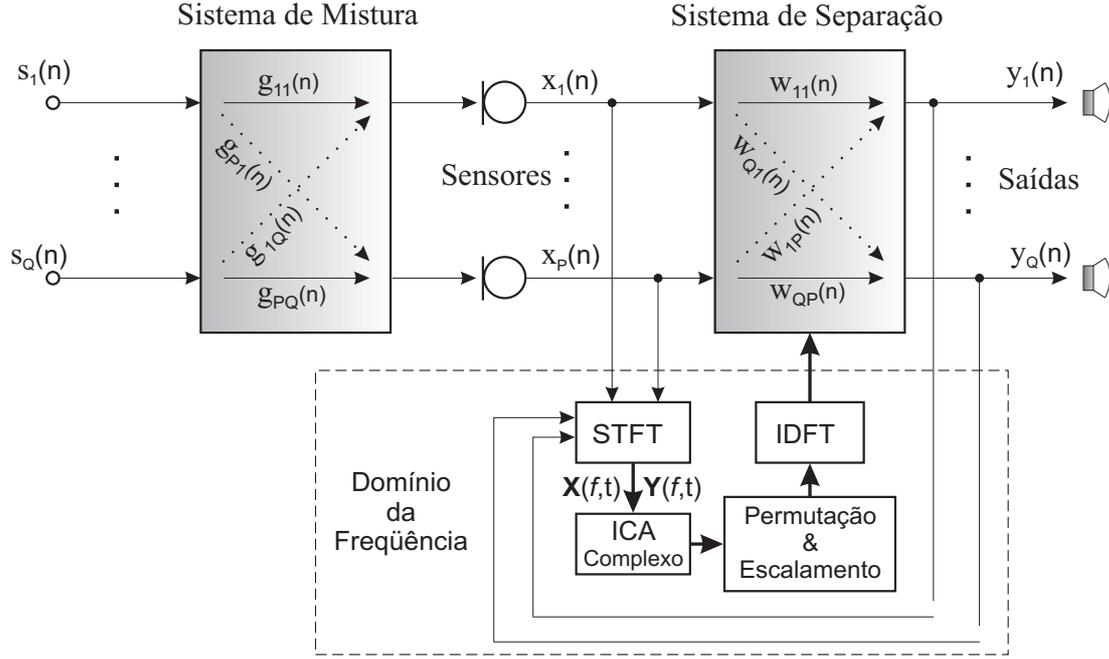


Figura 5.8: Esquema geral para separação cega de fontes no domínio da frequência.

sinais $X_P(\omega, k)$ no domínio da frequência através de uma transformada curta de *Fourier* de S -pontos (STFT, *Short-Time Fourier Transform*):

$$X_P(\omega, k) = \sum_{r=-\frac{S}{2}}^{\frac{S}{2}-1} x_p(k+r) \text{win}(r) e^{-j\Omega r}, \quad (5.61)$$

sendo $\Omega \in \{0, \frac{1}{S}2\pi, \dots, \frac{S-1}{S}2\pi\}$ a frequência normalizada, $\text{win}(r)$ a função janela com as extremidades suavizadas em torno de zero (normalmente a janela de *Hanning*) e k o índice da amostra do sinal janelado. Plotando os valores espectrais em raias de frequência de cada sensor, podemos considerá-los como uma série temporal designada por

$$\mathbf{X}(f, t) = [X_1(f, t) \dots X_P(f, t)]^T, \quad (5.62)$$

onde f representa uma determinada raia⁷ (para $0 \leq f \leq S-1$) da t -ésima janela aplicada ao sinal $x_p(n)$.

Seguindo a mesma notação, podemos expressar cada raia das Q séries temporais de saída como

$$\mathbf{Y}(f, t) = \mathbf{W}(f)\mathbf{X}(f, t), \quad (5.63)$$

⁷A relação entre a frequência ω e uma determinada raia f pode ser expressa da seguinte forma: $\omega = \frac{2\pi f}{S}$.

sendo

$$\mathbf{Y}(f, t) = [Y_1(f, t) \ \dots \ Y_Q(f, t)]^T \quad (5.64)$$

e

$$\mathbf{W}(f) = \begin{bmatrix} W_{11}(f) & \dots & W_{1P}(f) \\ \vdots & \ddots & \vdots \\ W_{Q1}(f) & \dots & W_{QP}(f) \end{bmatrix}. \quad (5.65)$$

De acordo com a equação anterior, a mistura convolutiva na Eq. (5.1) pode ser vista como uma mistura instantânea em cada raia de frequência.

Este procedimento deve ser aplicado a todas as raias de frequência. Finalmente, podemos aplicar a transformada inversa de *Fourier* (IDFT, *Inverse Discrete Fourier Transform*) para obter os coeficientes dos filtros $w_{qp}(n)$ e obter os sinais resultantes nas Q saídas.

5.2.2.1 Algoritmo Adaptativo

O algoritmo de adaptação obtido para misturas instantâneas reais no domínio do tempo da Eq. (5.22) pode ser utilizado para a separação de misturas convolutivas no domínio da frequência, tomando o cuidado de adaptá-lo para misturas instantâneas com valores complexos. A equação de atualização baseada no gradiente natural para cada raia é dada por

$$\mathbf{W}_{i+1}(f) = \mathbf{W}_i(f) - \mu [\mathbf{I} - 2\langle \Phi(\mathbf{Y}(f, t))\mathbf{Y}^H(f, t) \rangle_t] \mathbf{W}_i(f), \quad (5.66)$$

sendo o operador $\langle \cdot \rangle_t$ a média temporal, o índice i referente à i -ésima iteração, μ o fator de convergência,

$$\Phi(\mathbf{Y}(f, t)) = [\Phi(Y_1(f, t)), \dots, \Phi(Y_Q(f, t))]^T \quad (5.67)$$

o vetor que contém as funções não-lineares das saídas e

$$\Phi(Y_q(f, t)) \equiv \tanh(\text{Re}(Y_q(f, t))) + \tanh(\text{Im}(Y_q(f, t))). \quad (5.68)$$

5.2.2.2 Permutação e Escalamento

Na separação cega de fontes baseada na análise de componentes independentes os filtros de separação são atualizados para que as saídas tornem-se mutuamente independentes. Neste tipo de solução pode haver permutação entre as saídas e estas

ainda podem estar escaladas por um fator constante. A solução no domínio da frequência é realizada diretamente nas raias da DFT dos filtros que desfazem a mistura. Neste caso, a permutação e o escalamento entre raias de diferentes filtros levam à degradação de desempenho do algoritmo adaptativo, sendo o maior desafio a ser superado.

Algumas propostas para reduzir a ocorrência de permutações são encontradas na literatura: restrição do comprimento do filtro no domínio do tempo [59], minimização de correlação [60], estimativa de direção de chegada de cada fonte [61],[62] e soluções mistas destas duas últimas técnicas [63].

Para atenuar a ambiguidade de escalamento foi proposto em [45] forçar que todas as matrizes de dimensão $Q \times Q$ que desfazem a mistura tenham determinante unitário, ou seja,

$$\mathbf{W}(f)^{norm} = \mathbf{W}(f)^{orig} |\mathbf{W}(f)^{orig}|^{-\frac{1}{P}}, \quad (5.69)$$

sendo $\mathbf{W}(f)^{norm}$ a matriz de separação a ser utilizada (de determinante unitário) e $\mathbf{W}(f)^{orig}$ a matriz original. Isto assegura a conservação do volume de todas as matrizes de separação [45], mantendo o conteúdo espectral inalterado. Outra maneira de contornar o problema de escalamento é mediante o princípio da distorção mínima [64], que modifica a matriz de separação $\mathbf{W}(f)$ após a convergência do método, da seguinte forma:

$$\mathbf{W}(f) \leftarrow \text{diag} [\mathbf{W}^{-1}(f)] \mathbf{W}(f), \quad (5.70)$$

onde o operador $\text{diag}(\cdot)$ anula todos os coeficientes de uma matriz que não estejam em sua diagonal principal. Supondo que a separação seja razoável e que não há permutação (sem perda de generalidade), podemos aproximar a matriz $\mathbf{W}(f)$ por:

$$\mathbf{W}(f) \approx \mathbf{D}(f) \mathbf{G}^{-1}(f), \quad (5.71)$$

onde $\mathbf{D}(f)$ é uma matriz diagonal que contém os coeficientes de escalamento. Esta aproximação nos permite verificar que $\text{diag} [\mathbf{W}^{-1}(f)] \mathbf{W}(f) \approx \text{diag} [\mathbf{G}(f)] \mathbf{G}^{-1}(f)$. Esta aproximação implica que a transformação em (5.70) gera uma matriz $\mathbf{W}(f)$ com um escalamento razoável, e não com um escalamento completamente arbitrário.

Outra dificuldade diz respeito à equivalência entre a convolução linear, implementada no domínio do tempo, e a convolução circular, implementada pela técnica

no domínio da frequência. Para contornar este problema, é feita uma restrição no número de coeficientes não-nulos dos filtros de separação no domínio do tempo [59].

5.3 Medidas de Desempenho

Existem várias formas de avaliação de desempenho dos métodos de separação de fontes. Uma medida quantitativa bastante utilizada é a relação sinal interferência (SIR, *Signal to Interference Ratio*), definida para misturas instantâneas e determinadas como

$$\text{SIR} = 10 \log_{10} \left(\frac{1}{Q} \sum_{q=1}^Q \text{SIR}_q \right) \quad (5.72)$$

com $\text{SIR}_q = \max [\text{SIR}_{q1}, \dots, \text{SIR}_{qQ}]$ e

$$\text{SIR}_{qp} = \frac{\sum_n |c_{pq}s_q(n)|^2}{\sum_{l=1, l \neq q}^Q \sum_n |c_{pl}s_l(n)|^2}, \quad (5.73)$$

sendo SIR_{qp} a relação entre a potência do sinal presente na saída $y_p(n)$ quando somente a fonte $s_q(n)$ está ativa e a potência do sinal presente na saída $y_p(n)$ quando somente a fonte $s_q(n)$ está inativa⁸, e c_{pl} o elemento da p -ésima linha e l -ésima coluna da matriz $\mathbf{C} = \mathbf{WH}$.

Existem formas qualitativas de avaliar o desempenho da separação que consistem na visualização dos sinais ao longo do tempo seguida de uma comparação entre os sinais das fontes, misturas e saídas obtidas (estimativas das fontes); ou ainda, na análise da magnitude espectral ao longo do tempo, ou seja, comparando o espectrograma das fontes originais com os espectrogramas dos sinais observados e dos sinais separados podemos perceber a qualidade da separação. Uma forma subjetiva de avaliação seria escutar os sinais observados e separados e verificar o nível de separação.

Para o caso de misturas convolutivas, é necessário introduzir uma modificação na medida de avaliação SIR apresentada anteriormente. Neste caso, a Eq. (5.73) deve ser generalizada da seguinte forma:

$$\text{SIR}_{ij} = \frac{\sum_n |c_{ji}(n) * s_i(n)|^2}{\sum_{l=1, l \neq i}^N \sum_n |c_{jl}(n) * s_l(n)|^2}, \quad (5.74)$$

⁸Quando $\max [\text{SIR}_{qp}]$ ocorre em $q \neq p$, existe permutação entre a q -ésima fonte e a p -ésima saída.

sendo $c_{jl}(n)$ a soma das convoluções dos filtros da j -ésima linha da matriz \mathbf{W} com os filtros da l -ésima coluna da matriz \mathbf{H} , ou seja, $\mathbf{C} = \mathbf{W} * \mathbf{H}$. O operador “ $*$ ” representa a operação de convolução.

Exemplificando, para um sistema linear TITO, a Fig. 5.9 (a) mostra a estimativa da saída $y_1(n)$ quando somente a fonte $s_1(n)$ está ativa e a Fig. 5.9 (b) apresenta a estimativa da saída $y_1(n)$ quando somente a fonte $s_2(n)$ está ativa. Neste caso,

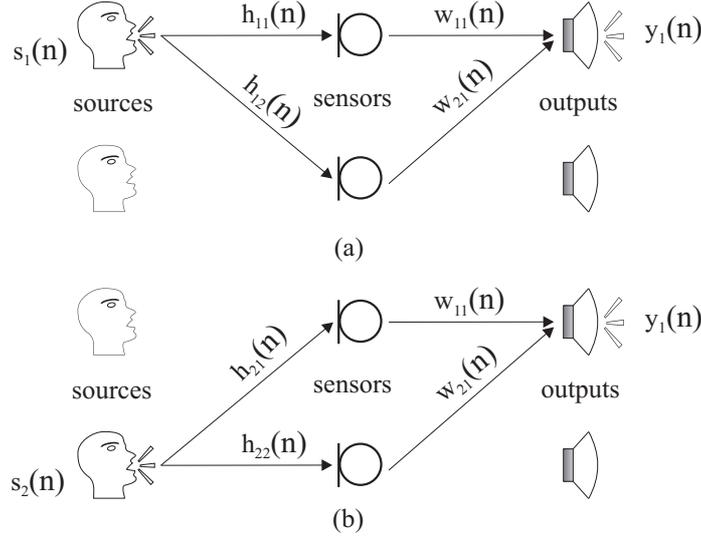


Figura 5.9: Esquema para mensurar SIR_1 em um sistema TITO.

assumindo que não há problema de permutação, podemos definir:

$$SIR_1 = \frac{(s_1(n) * [h_{11}(n) * w_{11}(n) + h_{12}(n) * w_{21}(n)])^2}{(s_2(n) * [h_{21}(n) * w_{11}(n) + h_{22}(n) * w_{21}(n)])^2}, \quad (5.75)$$

$$SIR_2 = \frac{(s_2(n) * [h_{22}(n) * w_{22}(n) + h_{21}(n) * w_{12}(n)])^2}{(s_1(n) * [h_{12}(n) * w_{22}(n) + h_{11}(n) * w_{12}(n)])^2}. \quad (5.76)$$

5.3.1 Medidas de Desempenho Globais

Uma forma alternativa para avaliação do desempenho global dos métodos BSS foi proposta em [65], na qual não é necessário o conhecimento do sistema de mistura, apenas das fontes originais e de suas estimativas.

Supondo fontes \mathbf{s}_q ($q = 1, \dots, Q$) descorrelacionadas e ausência de ruído de medição, pode-se decompor $\hat{\mathbf{s}}_q$ (a estimativa da q -ésima fonte) como:

$$\hat{\mathbf{s}}_q = \mathbf{s}_q^{\text{target}} + \mathbf{e}_q^{\text{interf}} + \mathbf{e}_q^{\text{artef}}, \quad (5.77)$$

sendo

$$\mathbf{s}_q^{\text{target}} = \frac{\langle \hat{\mathbf{s}}_q, \mathbf{s}_q \rangle \mathbf{s}_q}{\|\mathbf{s}_q\|^2} \quad (5.78)$$

uma deformação aceitável na q -ésima fonte,

$$\mathbf{e}_q^{\text{interf}} = \sum_{q' \neq q} \frac{\langle \hat{\mathbf{s}}_q, \hat{\mathbf{s}}_{q'} \rangle \mathbf{s}_{q'}}{\|\mathbf{s}_{q'}\|^2} \quad (5.79)$$

uma deformação na q -ésima fonte provocada pela interferência das outras $Q - 1$ fontes e

$$\mathbf{e}_q^{\text{artef}} = \hat{\mathbf{s}}_q - \mathbf{s}_q^{\text{target}} - \mathbf{e}_q^{\text{interf}} \quad (5.80)$$

os artefatos introduzidos pelos algoritmos de separação.

Após esta decomposição, ilustrada na Fig. 5.10, pode-se definir três medidas globais de desempenho [65], ou seja, a razão sinal-interferência global (SIR_G , *Global Signal to Interference Ratio*), a razão sinal-artefato global (SAR_G , *Global Signal-to-Artifact Ratio*) e a razão sinal-distorção global (SDR_G , *Global Signal-to-Distortion Ratio*), dadas por:

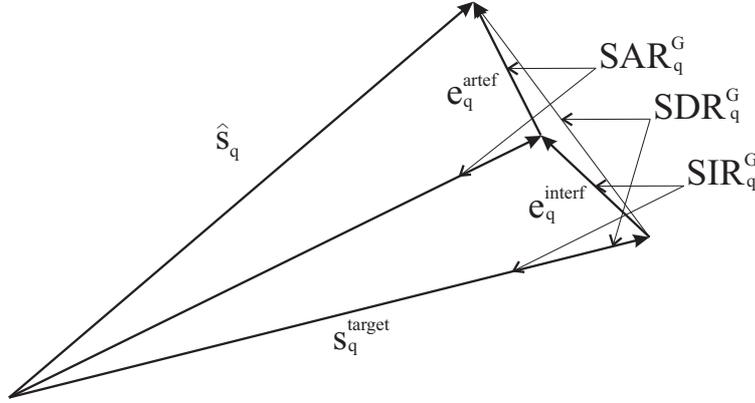


Figura 5.10: Decomposição da estimativa da q -ésima fonte para cálculo das medidas de desempenho.

$$\text{SIR}_q^G = 10 \log_{10} \frac{\|\mathbf{s}_q^{\text{target}}\|^2}{\|\mathbf{e}_q^{\text{interf}}\|^2}, \quad (5.81)$$

$$\text{SAR}_q^G = 10 \log_{10} \frac{\|\mathbf{s}_q^{\text{target}} + \mathbf{e}_q^{\text{interf}}\|^2}{\|\mathbf{e}_q^{\text{artef}}\|^2}, \quad (5.82)$$

$$\text{SDR}_q^G = 10 \log_{10} \frac{\|\mathbf{s}_q^{\text{target}}\|^2}{\|\mathbf{e}_q^{\text{interf}} + \mathbf{e}_q^{\text{artef}}\|^2}. \quad (5.83)$$

As medidas de desempenho globais médias podem ser obtidas como descrito a seguir:

$$\text{SIR}^G = 10\log_{10} \left(\frac{1}{Q} \sum_{q=1}^Q \text{SIR}_q^G \right), \quad (5.84)$$

$$\text{SAR}^G = 10\log_{10} \left(\frac{1}{Q} \sum_{q=1}^Q \text{SAR}_q^G \right) \quad (5.85)$$

e

$$\text{SDR}^G = 10\log_{10} \left(\frac{1}{Q} \sum_{q=1}^Q \text{SDR}_q^G \right). \quad (5.86)$$

Capítulo 6

Separação Cega de Sinais em Subbandas

6.1 Introdução

Métodos BSS em subbandas têm sido propostos recentemente com o objetivo de reduzir a complexidade computacional e melhorar a taxa de convergência da adaptação, porém negligenciando os efeitos da sobreposição do espectro e mantendo um número suficiente de amostras para estimar as estatísticas dos sinais em subbandas [47], [66] e [67]. Usualmente estes resultados são alcançados usando banco de filtros DFT subamostrados, porém este trabalho descartará a utilização coeficientes complexos.

Neste capítulo apresentaremos dois métodos para separação cega de misturas convolutivas em subbandas, que utilizam bancos de filtros com coeficientes reais: um algoritmo *off-line*, o qual utiliza um banco de filtros não-uniforme subamostrado com bandas de frequência decompostas em oitavas; e um algoritmo *on-line*, o qual emprega uma estrutura em subbandas uniformes com amostragem crítica e filtros extras que cancelam a sobreposição do espectro entre canais adjacentes. Devido aos filtros de separação nas subbandas trabalharem em taxas de amostragem reduzidas e terem ordens menores, os métodos propostos apresentam menor complexidade computacional, maior taxa de convergência e maior relação sinal-interferência em regime estacionário quando comparados aos algoritmos correspondentes em banda cheia. Os coeficientes dos filtros que desfazem as misturas em cada subbanda são

ajustados independentemente por um algoritmo no domínio do tempo, que emprega estatísticas de segunda ordem e um esquema de normalização que resulta numa maior taxa de convergência e reduzida complexidade quando comparado ao esquema de normalização original. Outra vantagem dos métodos propostos consiste no uso de filtros com coeficientes reais, que são atraentes para implementação em dispositivos para processamento de sinais (DSPs).

6.2 BSS em Subbandas no Domínio do Tempo

Off-line

Nesta seção investigaremos o uso de uma estrutura em subbandas que emprega processamento multitasas com taxas de amostragem menores nas baixas frequências, onde o sinal de fala concentra menor energia, em conjunto com o algoritmo BSS em blocos no domínio do tempo apresentado na Seção 5.2.1, o qual explora duas características dos sinais das fontes simultaneamente: correlação e não-estacionariedade.

A idéia é explorar as características de melhoria da taxa de convergência e de redução da complexidade computacional inerentes às estruturas em subbandas. Uma estrutura que emprega banco de filtros uniforme com coeficientes complexos aplicada à BSS foi usada em [47]. Nossa proposta é utilizar um banco de filtros não-uniforme em oitavas com coeficientes reais e sem decimação crítica. A Fig. 6.1 mostra um sistema BSS em subbandas de duas entrada e duas saídas considerando um banco de filtros não-uniforme de M canais. Esta estrutura é uma versão modificada, em relação à estrutura apresentada em [28], onde os sinais presentes nas entradas dos filtros de separação de cada subbanda $\tilde{w}_{pq}^k(n)$ são decimados pela metade do fator de decimação crítico com o objetivo de reduzir o efeito da sobreposição de espectro durante o processo de adaptação dos coeficientes. As decimações dos sinais nas saídas dos filtros de separação de cada subbanda por um fator igual a 2 restauram a taxa de amostragem crítica da estrutura antes da etapa de reconstrução do sinal de saída.

Assumindo que $h_p(n)$ é a resposta ao impulso do filtro protótipo de comprimento N_P que permite reconstrução perfeita em um banco de filtros uniforme modulado por cosseno de 2 bandas [19], os filtros de análise e síntese são dados, respectivamente,

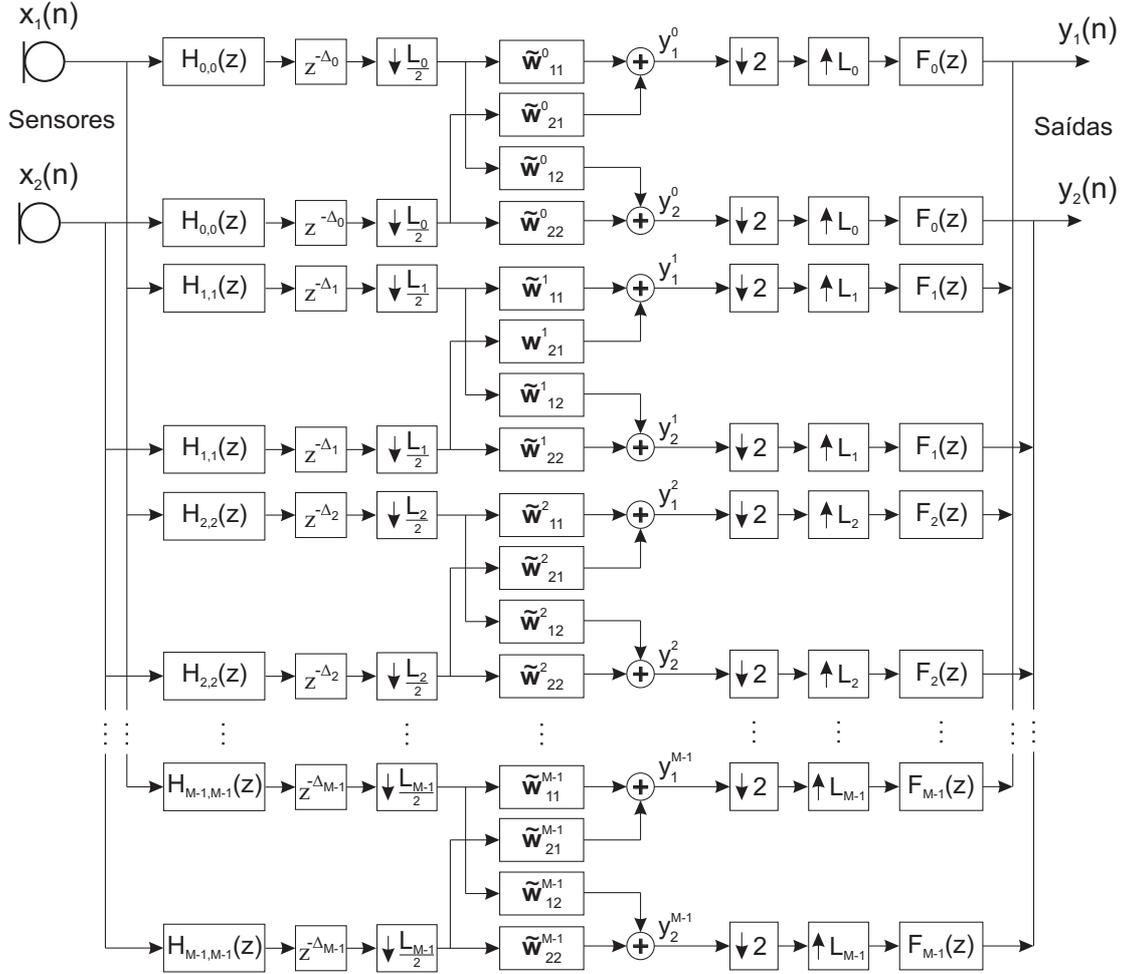


Figura 6.1: Configuração TITO linear para BSS em subbandas.

por [5]:

$$h^l(n) = 2h_p(n) \cos \left[\frac{\pi}{2} (l + 0, 5) \left(n - \frac{N_P - 1}{2} \right) + \theta_l \right], \quad (6.1)$$

$$f^l(n) = 2h_p(n) \cos \left[\frac{\pi}{2} (l + 0, 5) \left(n - \frac{N_P - 1}{2} \right) - \theta_l \right], \quad (6.2)$$

sendo $\theta_l = (-1)^l \frac{\pi}{4}$ para $l = 0, 1$ e $0 \leq n \leq N_P - 1$.

Para um banco de filtros com bandas de frequência em oitavas, implementado a partir de banco de filtros uniforme de 2 canais (Eqs. (6.1) e (6.2)), os filtros de análise ($H_k(z)$) e síntese ($F_k(z)$) equivalentes de um banco de filtros em árvore não-uniforme de M-canais são obtidos conforme apresentado no Capítulo 3 (ver Seção 3.4). O número de coeficientes de cada subfiltro de separação na k -ésima subbanda

$w_{pq}^k(n)$ deve ser no mínimo [28]

$$S_k = 2 \left\lceil \frac{S - 1 + N_{F_k}}{L_k} \right\rceil, \quad (6.3)$$

sendo N_{F_k} a ordem do k -ésimo filtro do banco de síntese correspondente.

Para ajustar os coeficientes de cada subfiltro de separação empregamos o algoritmo *off-line* em banda cheia (Eq. (5.47)) adaptado para subbandas, sendo a equação de atualização dos coeficientes dos filtros da k -ésima banda dada por:

$$\tilde{\mathbf{W}}^k(i) = \tilde{\mathbf{W}}^k(i-1) - \frac{2}{b_k} \left\{ \sum_{m=1}^{b_k} \begin{bmatrix} \tilde{\mathbf{W}}_{12}^k \mathbf{R}_{21}^k \mathbf{R}_{11}^{k-1} & \tilde{\mathbf{W}}_{11}^k \mathbf{R}_{12}^k \mathbf{R}_{22}^{k-1} \\ \tilde{\mathbf{W}}_{22}^k \mathbf{R}_{21}^k \mathbf{R}_{11}^{k-1} & \tilde{\mathbf{W}}_{21}^k \mathbf{R}_{12}^k \mathbf{R}_{22}^{k-1} \end{bmatrix} \right\} \begin{bmatrix} \mu_1^k \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mu_2^k \mathbf{I} \end{bmatrix} \quad (6.4)$$

sendo

$$\mathbf{R}_{pq}^k(m) = [\mathbf{Y}_p^k(m)]^H \mathbf{Y}_q^k(m) \quad (6.5)$$

e

$$\mathbf{Y}_q^k(m) = \begin{bmatrix} y_q^k(mS_k) & \dots & y_q^k(mS_k - D_k + 1) \\ y_q^k(mS_k + 1) & \dots & y_q^k(mS_k - D_k + 2) \\ \vdots & \dots & \vdots \\ y_q^k(mS_k + N_k - 1) & \dots & y_q^k(mS_k - D_k + N_k) \end{bmatrix}. \quad (6.6)$$

As matrizes acima têm dimensões $D_k \times D_k$ (com $1 \leq D_k \leq S_k$) e $N_k \times D_k$ (com $N_k \geq D_k$), respectivamente, b_k é o número de blocos, N_k é o tamanho de cada bloco, μ_q^k é o q -ésimo passo de adaptação da k -ésima subbanda, i é o número de iterações (*off-line*) para todas as subbandas e y_q^k é o sinal da q -ésima saída na k -ésima subbanda (ver Figura 6.1).

Para reduzir a complexidade computacional do algoritmo, o fator de normalização $\mathbf{R}_{qq}^{k-1}(m)$ pode ser simplificado para um escalar [42]. Neste caso,

$$\mathbf{R}_{qq}^k(m) \approx [\mathbf{y}_q^k(m)]^T \mathbf{y}_q^k(m) \mathbf{I}, \quad (6.7)$$

com $\mathbf{y}_q^k(m)$ correspondendo à primeira coluna da matriz de blocos atrasados no tempo (Eq. (6.6)), sendo sua inversa reduzida ao inverso da potência do bloco mais recente de amostras da saída $y_q^k(n)$. Esta simplificação foi proposta sem sabermos que já tinha sido apresentado em [42].

A implementação em subbandas proposta apresenta maior flexibilidade que a versão em banda cheia, podendo trabalhar, por exemplo, com filtros de separação de diferentes comprimentos e passos de adaptação independentes nos diversos canais.

6.2.1 Complexidade Computacional

O número de multiplicações reais por bloco do sinal de entrada (NMPB), necessárias para atualizar os coeficientes dos filtros FIR de separação nas diversas subbandas da estrutura não-uniforme de M canais e para obter um bloco dos sinais de saída em banda cheia, considerando implementação na forma direta e um sistema linear MIMO $P \times P$, é dado por

$$\begin{aligned} \text{NMPB}_{\text{SB}} = & P \sum_{k=1}^M \frac{NH_{kk}N_k}{\frac{L_k}{2}} + P^2 \sum_{k=1}^M \frac{2S_k D_k N_k}{\frac{L_k}{2}} + PS \sum_{k=1}^M \frac{N_{F_k}}{L_k} \\ & + P^2 \sum_{k=1}^M \frac{N_k + D_k N_k}{\frac{L_k}{2}} + P^2 \sum_{k=1}^M \frac{\left(\sum_{j=1}^{S_k} j\right) + S_k}{\frac{L_k}{2}} \end{aligned} \quad (6.8)$$

As duas primeiras parcelas correspondem ao custo para gerar todos os blocos atrasados no tempo de todas as saídas nas M subbandas. A terceira parcela mostra o custo do processamento pelo banco de síntese para gerar os blocos das saídas em banda cheia. A quarta parcela representa o custo da obtenção das submatrizes de correlação $\mathbf{R}_{pq}^k(m)$ (ver Eq. (6.5)) para o cálculo do gradiente. O quinto termo indica o custo da adaptação considerando apenas a atualização dos S_k coeficientes dos filtros de separação nas diversas subbandas. O fator no denominador de todas as parcelas indica que as operações estão sendo realizadas numa taxa de amostragem $\frac{L_k}{2}$ vezes menor que a taxa em banda cheia. Para filtros de separação de altas ordens, a equação acima pode ser simplificada, considerando somente os termos dominantes, e expressa por

$$\text{NMPB}_{\text{SB}} \approx P^2 \sum_{k=1}^M \left(\frac{8S_k^3}{L_k} \right). \quad (6.9)$$

Os resultados acima foram obtidos considerando o fator de normalização simplificado (Eq. (6.7)), $D_k = S_k$ e $N_k = 2D_k$ para o algoritmo em subbandas.

6.2.2 Resultados Experimentais

Todos os experimentos foram realizados usando dois sinais de fala com 10 segundos de duração e frequência de amostragem $f_s = 8\text{kHz}$: uma voz feminina em inglês e uma voz masculina em português. Esses sinais foram convoluídos com as respostas

impulsivas sintéticas obtidas para o cenário virtual ¹ mostrado na Fig. 6.2, com tempo de reverberação de 250 ms, espaçamento entre os microfones de 5cm e dimensões 3,55 m × 4,55 m × 2,5 m. As fontes sonoras foram posicionadas a 1 m de distância do ponto médio entre os sensores em duas direções diferentes: -50° e 45° .

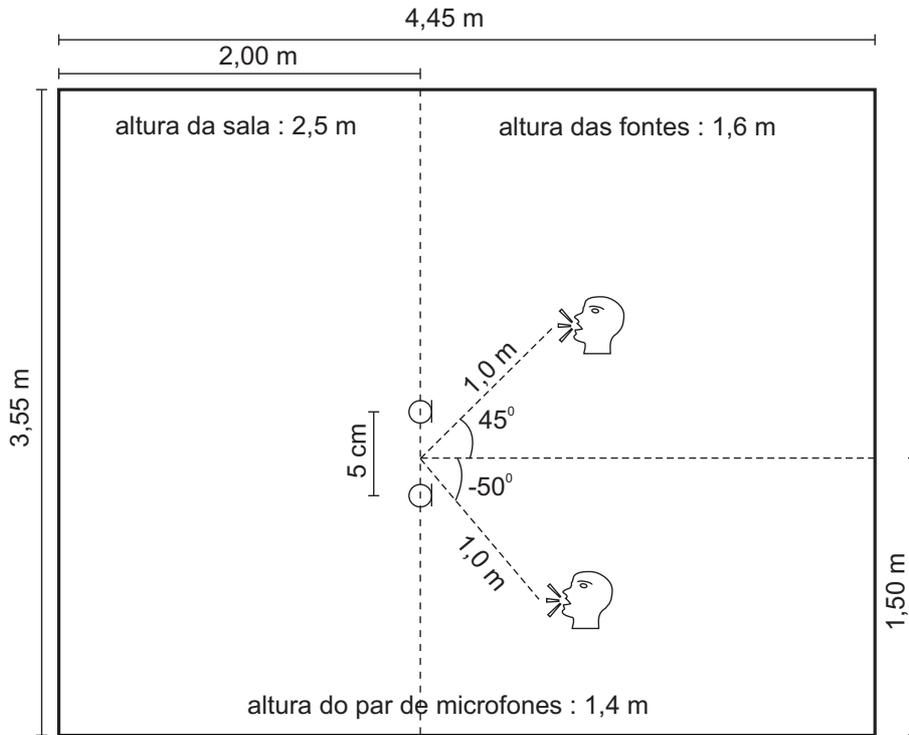


Figura 6.2: Cenário virtual usado nos experimentos.

Em nossos experimentos as misturas foram realizadas considerando diferentes condições de reverberação, reamostrando as respostas impulsivas obtidas em [69] para $f_s = 8\text{kHz}$ e truncando-as para o tamanho desejado (ou seja, considerando somente suas S primeiras amostras). O comprimento dos filtros de separação foi fixado, em todas as simulações, igual ao dos filtros de mistura ($U = S$). Os algoritmos foram executados em um processador Intel Core 2 Duo 2GHz. Adotamos, para avaliação de desempenho, a relação sinal interferência (SIR) definida na Seção 5.3.

¹Este cenário foi proposto para competição promovida em 2006 pela comunidade de separação de sinais *Signal Separation Evaluation Campaign (SiSEC 2006)*, com os resultados apresentados [68] em um painel de discussões durante a *7th International Conference on Independent Component Analysis and Signal Separation (ICA 2007)*. Originalmente o cenário tinha $Q = 4$ fontes distintas em diferentes direções, entretanto, nos nossos experimentos para reduzir a complexidade do procedimento BSS utilizamos somente duas destas fontes.

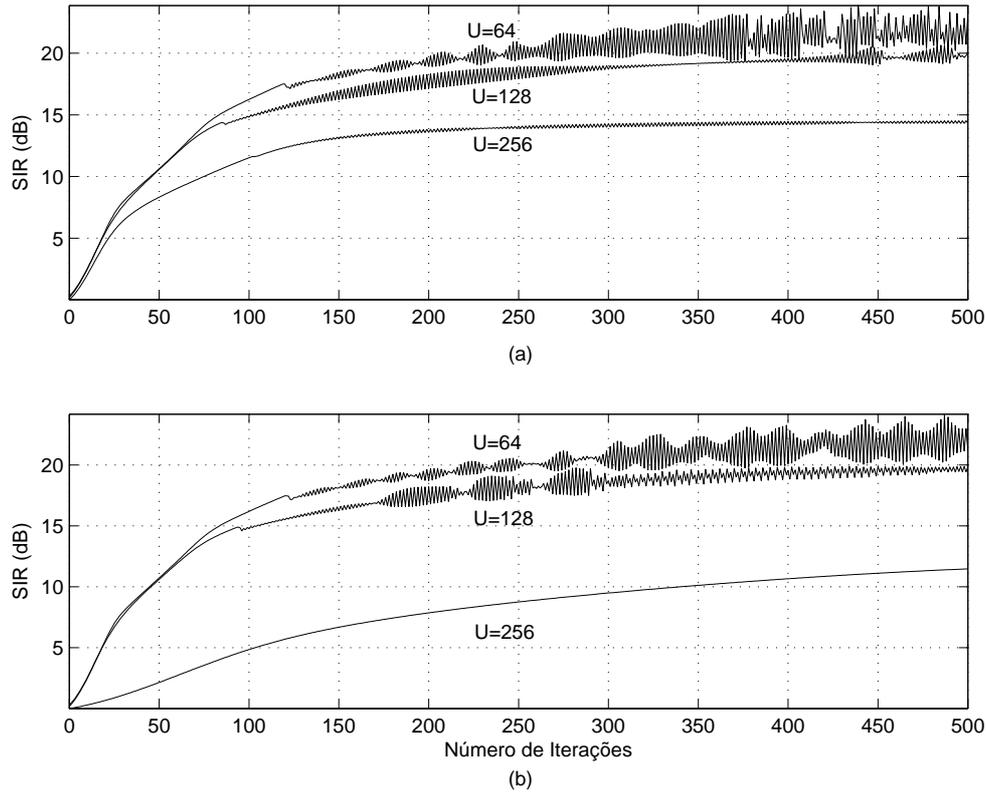


Figura 6.3: Evolução do SIR para os dois esquemas de normalização: (a) NN e (b) NA.

Os coeficientes dos filtros de separação de cada subbanda $\tilde{w}_{pq}^k(n)$ foram inicializados sempre iguais a zero, exceto para $p = q$ e $n = 0$, os quais foram feitos iguais a 1. O mesmo procedimento foi adotado para o algoritmo em banda cheia.

6.2.2.1 Experimento 1

Nesse experimento comparamos o desempenho do algoritmo em banda cheia usando dois diferentes esquemas de normalização: (a) normalização nova (NN - Eq. (5.54)) e (b) normalização antiga (NA - Eq. (5.53)). A Fig. 6.3 mostra a evolução do SIR considerando filtros de mistura de diferentes comprimentos $U = 64, 128$ e 256 . O passo de adaptação em todos os casos foi $\mu = 5 \times 10^{-3}$, exceto para a nova normalização com $U = 256$, onde foi usado $\mu = 1 \times 10^{-3}$ por razões de convergência. A Tabela 6.1 mostra o tempo de processamento para as simulações exibidas na Fig. 6.3. Analisando os resultados da Tabela 6.1 e da Fig. 6.3 podemos observar que para filtros de misturas de comprimentos maiores (correspondendo a ambientes mais reverberantes), o novo esquema de normalização reduz significativamente o tempo de processamento

Tabela 6.1: Tempo de processamento em minutos.

S=U	NN	NA
64	13	22
128	23	55
256	57	181

Tabela 6.2: Parâmetros da estrutura não-uniforme para $M = 4$ e $U = S = 1024$.

k	L_k	N_{H_k}	Δ_k	S_k	μ_1^k	μ_2^k
0	8	441	0	366	8.8×10^{-3}	8×10^{-3}
1	8	441	0	366	17.6×10^{-3}	16×10^{-3}
2	4	189	252	606	35.2×10^{-3}	32×10^{-3}
3	2	63	378	575	70.4×10^{-3}	64×10^{-3}

e melhora sensivelmente a velocidade de convergência do algoritmo em banda cheia.

6.2.2.2 Experimento 2

Nesse experimento, descartando o uso de banco de filtros com coeficientes complexos, confrontamos apenas as estruturas em banda cheia e em subbandas, ambas com os métodos aproximados de normalização de acordo com as Eqs. (5.54) e (6.7). A estrutura em subbandas não-uniforme foi implementada utilizando um banco de filtros com bandas em oitavas e com reconstrução perfeita para $M = 4$ subbandas. A Tabela 6.2 apresenta os fatores de decimação L_k , as ordens dos filtros de análise N_{H_k} (que são iguais às ordens dos filtros de síntese N_{F_k}), os atrasos Δ_k , as ordens S_k dos filtros de separação w_{pq}^k e os passos de adaptação, μ_1^k e μ_2^k , usados nas simulações em subbandas com $U = S = 1024$. O passo de adaptação para o algoritmo em banda cheia foi o mesmo usado no Experimento 1, exceto para $U = 1024$ onde foi usado $\mu = 3 \times 10^{-3}$. Estes passos resultaram na melhor convergência e foram obtidos experimentalmente por tentativa e erro.

A Fig. 6.4 mostra as respostas em frequência dos filtros de análise $H_{k,k}(z)$ correspondentes (ver Fig. 6.1) para $M = 4$ subbandas não-uniformes.

Para reduzir o custo computacional sem degradação significativa no processo de separação, a ordem dos filtros de separação na banda de maior frequência ($k =$

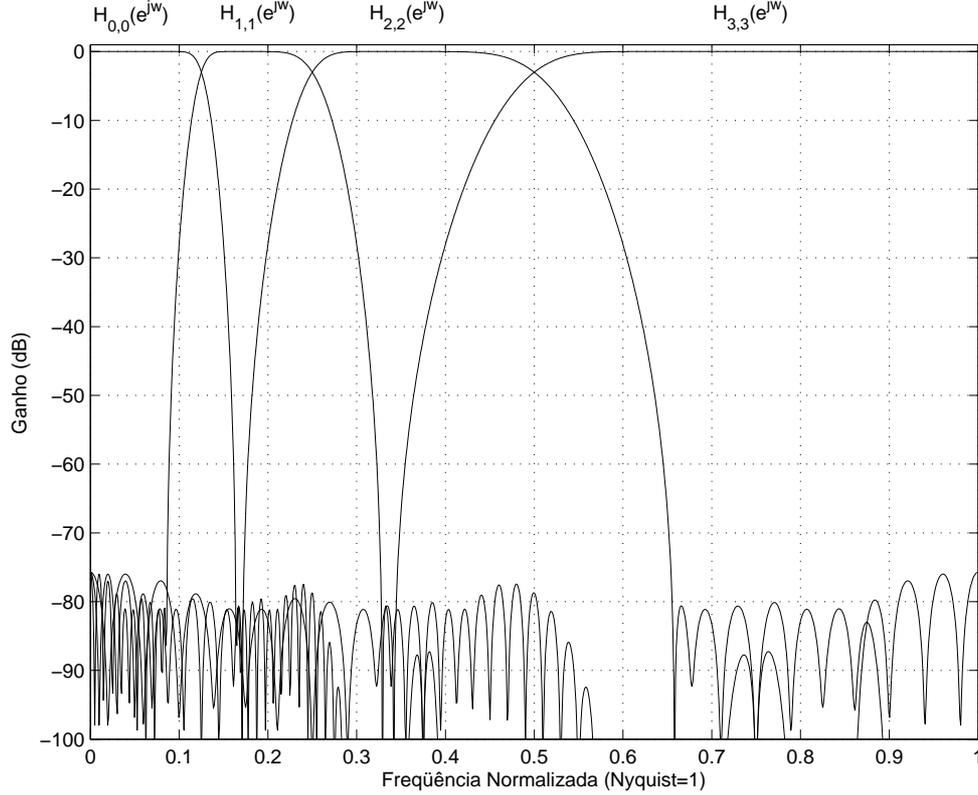


Figura 6.4: Resposta em frequência dos filtros de análise $H_{k,k}(z)$ para $M = 4$

3), que são adaptados na mesma taxa dos sinais capturados pelos microfones, foi reduzida em relação à da Eq. (6.3) para

$$S_3 = 2 \left\lceil \frac{\frac{S-1}{2} + N_{F_k}}{L_k} \right\rceil. \quad (6.10)$$

O resultado acima foi obtido por tentativa e erro, ou seja, filtro de separação na banda 3 de ordem menor que a dada pela Eq. (6.10) pode comprometer o desempenho do algoritmo em subbandas. Esta redução foi possível devido às características de reverberação e à pequena concentração de energia dos sinais de fala em altas frequências.

A Fig. 6.5 mostra a evolução do SIR para os algoritmos em banda cheia (Eq. (5.52)) e em subbandas (Eq. (6.27)), com filtros de misturas de comprimentos diferentes ($U = 256, 512$ e 1024).

A Tabela 6.3 exhibe o SIR máximo em banda cheia e em cada uma das 4 bandas da estrutura não-uniforme. A Tabela 6.4 mostra o tempo de processamento para as simulações da Fig. 6.5 e a complexidade computacional dos algoritmos de acordo com as Eqs. (5.56) e (6.9). Analisando estas tabelas e os resultados da

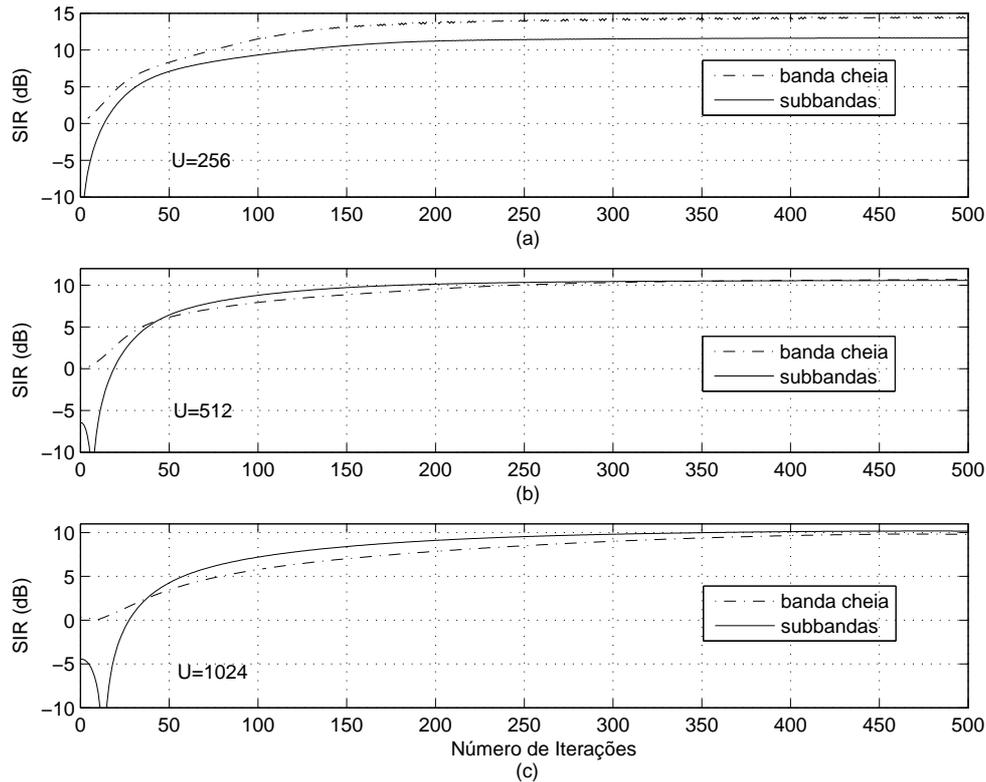


Figura 6.5: Evolução do SIR para os algoritmos em banda cheia e em subbandas com filtros de misturas de diferentes comprimentos: $U=256$, 512 e 1024 (de cima para baixo, respectivamente).

Fig. 6.5, podemos observar que na medida em que aumenta a ordem do sistema de mistura (ambientes mais reverberantes), as vantagens da estrutura em subbandas sobre a estrutura em banda cheia tornam-se mais evidentes, resultando numa taxa de convergência mais rápida, numa menor complexidade computacional e num menor tempo de processamento, sendo que este pode ser ainda reduzido usando processamento paralelo para implementações em subbandas.

O desempenho da BSS na banda de frequência mais alta ($k = 3$) foi sempre pior que nas demais subbandas (devido ao uso de filtros de separação de comprimento reduzido nesta banda), causando uma redução no SIR final. Contudo, nesta faixa de frequência a potência dos sinais de fala é reduzida e auditivamente os resultados ficaram significativamente melhores que em banda cheia.

A Fig. 6.6 mostra o espectro de potência da fonte original e de suas estimativas em banda cheia e em subbandas, para filtros de mistura de diferentes comprimentos. Estes resultados mostram a robustez dos algoritmos para o problema de

Tabela 6.3: SIR Máximo (em dB).

Filtros de Mistura	SIR em subbandas				SIR
	$k = 0$	$k = 1$	$k = 2$	$k = 3$	
U=S					--
256	21.72	11.06	12.57	6.63	14.56
512	13.81	9.41	10.37	6.31	10.70
1024	12.40	8.73	9.39	6.44	9.92

Tabela 6.4: Tempo de processamento (TP) em minutos e número de multiplicações por bloco (NMPB) requerida pelos algoritmos.

$S = U$	Banda cheia		Subbandas	
	TP	NMPB	TP	NMPB
256	45	2.68×10^8	57	4.90×10^8
512	127	2.15×10^9	83	1.96×10^9
1024	377	1.72×10^{10}	167	1.05×10^{10}

branqueamento das fontes, comum quando se aplicam técnicas de deconvolução cega multicanal para separar as fontes, e também para o escalamento dos sinais de saída.

Durante nossos experimentos monitoramos a correlação entre as estimativas das fontes nas diversas subbandas para evitar problemas de permutação, mas os algoritmos mostraram-se muito robustos, não demandando qualquer correção.

6.3 BSS em Subbandas no Domínio do Tempo

On-line

Nesta seção generalizaremos o algoritmo *on-line* BSS proposto em [3] para operar em multitaxas. A estrutura BSS em subbandas emprega um banco de filtros uniforme criticamente decimado ² com coeficientes reais e subfiltros FIR de separação de ordens reduzidas. Filtros extras são usados na decomposição dos sinais observados para cancelar a sobreposição do espectro entre as bandas adjacentes. A idéia é

²Até o presente momento, o autor desconhece na literatura estruturas criticamente decimadas aplicadas ao problema de BSS. Optamos por utilizar banco de filtros uniforme por razões de simplicidade. As propostas apresentadas podem ser estendidas para estruturas não-uniformes.

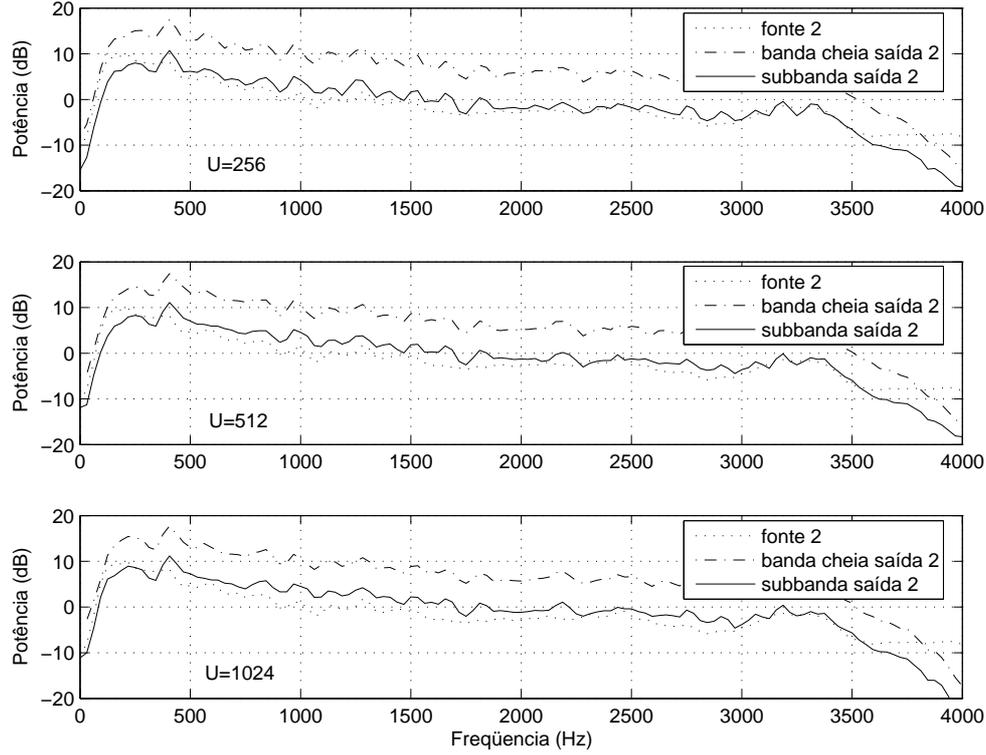


Figura 6.6: Espectro da fonte masculina original, do sinal separado em banda cheia e em subbandas, com filtros de mistura de diferentes comprimentos: $U=256$, 512 e 1024 .

explorar as características de melhor taxa de convergência e reduzida complexidade computacional do algoritmo adaptativo em subbandas apresentado em [70]. Da mesma forma que na Seção 6.2, os filtros que desfazem as misturas nas diferentes bandas são ajustados independentemente pelo algoritmo adaptativo no domínio da tempo, que utiliza estatísticas de segunda ordem, apresentado em [3].

A Fig. 6.7 mostra a k -ésima banda da configuração linear TITO para BSS em subbandas de M -canais proposta. Nesta estrutura cada sinal observado ($x_q(n)$) é decomposto pelos filtros diretos ($h_{k,k}(n)$) e pelos filtros extras ($h_{k,k-1}(n)$ e $h_{k,k+1}(n)$). Os sinais resultantes ($x_q^{k,k}(n)$, $x_q^{k,k-1}(n)$ e $x_q^{k,k+1}(n)$) são decimados pelo fator crítico (M) e em seguida passam pelos filtros de separação ($w_{qp}^k(n)$, $w_{qp}^{k-1}(n)$ e $w_{qp}^{k+1}(n)$, respectivamente). Os sinais de saída correspondentes têm suas taxas de amostragem restauradas e são recombinados pelos filtros de síntese ($f_k(n)$) para formar os sinais de saída em banda cheia (estimativas das fontes, $y_q(n)$). Considerando banco de filtros com reconstrução perfeita (PR), esta estrutura é capaz de modelar exatamente qualquer sistema FIR que desfaça as misturas [70].

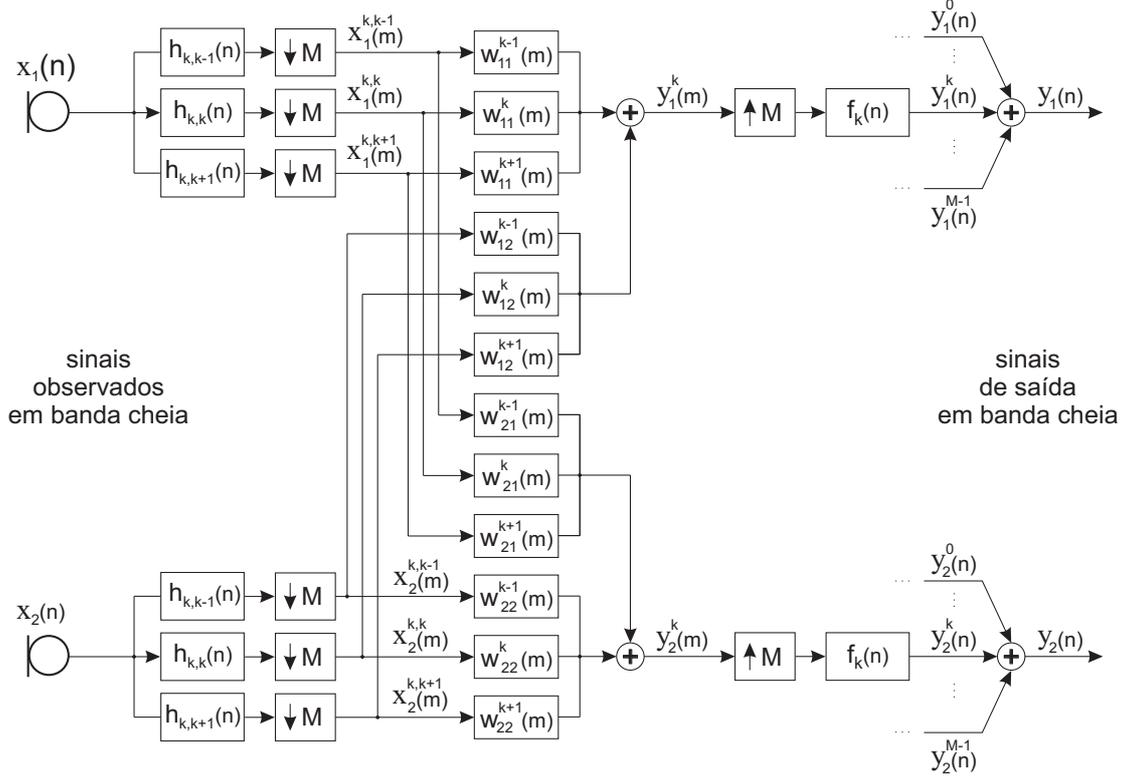


Figura 6.7: k -ésimo canal da configuração linear TITO para BSS em subbandas.

Assumindo que $h_p(n)$ é a resposta ao impulso de um filtro protótipo de comprimento N_P que permite reconstrução perfeita em um sistema multitasas modulado por cosseno de M bandas [19], os coeficientes dos filtros de análise e de síntese são dados, respectivamente, pelas Eqs. (3.15) e (3.16).

Os filtros $h_{k,i}(n)$ da Fig. 6.7, que decompõem os sinais observados $x_q(n)$, têm respostas ao impulso dadas por $h_{k,i}(n) = h_k(n) * h_i(n)$. O número de coeficientes de cada subfiltro de separação na k -ésima subbanda (\mathbf{w}_{pq}^k) deve ser no mínimo [70]:

$$K = \left\lceil \frac{S + N_P}{M} \right\rceil. \quad (6.11)$$

Os sinais observados $x_p^{k,i}(m)$ mostrados na Fig. 6.7 podem ser expressos como

$$x_p^{k,i}(m) = \mathbf{x}_p^T(m) \mathbf{h}_{k,i}, \quad (6.12)$$

sendo $\mathbf{x}_p(m) = [x_p(mM), x_p(mM - 1), \dots, x_p(mM - N_H + 1)]^T$ o vetor que contém as últimas $N_H = 2N_p - 1$ amostras do sinal do p -ésimo sensor e $\mathbf{h}_{k,i} = [h_{k,i}(0), h_{k,i}(1), \dots, h_{k,i}(N_H - 1)]^T$ o vetor que contém os N_H coeficientes do filtro de análise $h_{k,i}(n)$.

Para sinais coloridos e não-estacionários, como os sinais de áudio, o problema de BSS pode ser resolvido diagonalizando a matriz de correlação das saídas, considerando múltiplos blocos de amostras das saídas em diferentes instantes de tempo (TDD). Nesta seção vamos estender a solução *on-line* do tipo banda larga proposta em [3] para o domínio em subbandas, empregando processamento multitaxas. O método derivado em [3] é baseado em estatísticas de segunda ordem e considera duas características dos sinais das fontes simultaneamente [43]:

1. o não-branqueamento, explorando a diagonalização simultânea das matrizes de correlação das saídas considerando múltiplos atrasos;
2. a não-estacionariedade, explorando a diagonalização simultânea das matrizes reduzidas de correlação temporal das saídas em diferentes intervalos de tempo.

A partir da Fig. (6.7), considerando que não existe sobreposição entre as respostas de frequência dos filtros não-adjacentes $h_{k,i}(n)$, o sinal da q -ésima saída na k -ésima subbanda é descrito por

$$y_q^k(m) = \sum_{p=1}^P \sum_{i=k-1}^{k+1} [\mathbf{w}_{qp}^i]^T \mathbf{x}_p^{k,i}(m), \quad (6.13)$$

sendo $\mathbf{x}_p^{k,i}(m)$ o vetor que contém as últimas K amostras do sinal decomposto em subbandas do p -ésimo sensor $x_p^{k,i}(m)$ e $\mathbf{w}_{qp}^i = [w_{qp}^i(0), w_{qp}^i(1), \dots, w_{qp}^i(K-1)]^T$ o vetor que contém os K coeficientes do sub-filtro de separação da i -ésima subbanda $w_{qp}^i(m)$. O vetor $\mathbf{x}_p^{k,i}(m)$ pode ser expresso como

$$\mathbf{x}_p^{k,i}(m) = \mathbf{X}_p(m) \mathbf{h}_{k,i} \quad (6.14)$$

sendo $\mathbf{X}_p(m)$ a matriz de dimensão $K \times N_H$ dada por

$$\mathbf{X}_p(m) = \begin{bmatrix} x_p(mM) & x_p(mM-1) & \cdots & x_p(mM-N_H+1) \\ x_p((m-1)M) & x_p((m-1)M-1) & \cdots & x_p((m-1)M-N_H+1) \\ \vdots & \vdots & \ddots & \vdots \\ x_p((m-K+1)M) & x_p((m-K+1)M-1) & \cdots & x_p((m-K+1)M-N_H+1) \end{bmatrix}. \quad (6.15)$$

No algoritmo BSS em subbandas genérico no domínio do tempo, definindo N como o tamanho do bloco ($N \geq D$) e D como o número de blocos que serão utilizados

na estimativa das correlações ($1 \leq D \leq K$), um bloco da q -ésima saída na k -ésima banda indexado por ℓ pode se escrito como

$$\begin{aligned} [\mathbf{y}_q^k(\ell)]^T &= \sum_{p=1}^P \sum_{i=k-1}^{k+1} [\mathbf{w}_{qp}^i]^T \hat{\mathbf{X}}_p^{k,i}(\ell) \\ &= \left[y_q^k(\ell K), y_q^k(\ell K + 1), \dots, y_q^k(\ell K + N - 1) \right] \end{aligned} \quad (6.16)$$

sendo a matriz $\hat{\mathbf{X}}_p^{k,i}(\ell)$ de dimensão $K \times N$ expressa como

$$\begin{aligned} \hat{\mathbf{X}}_p^{k,i}(\ell) &= \begin{bmatrix} x_p^{k,i}(\ell K) & x_p^{k,i}(\ell K + 1) & \dots & x_p^{k,i}(\ell K + N - 1) \\ x_p^{k,i}(\ell K - 1) & x_p^{k,i}(\ell K) & \dots & x_p^{k,i}(\ell K + N - 2) \\ \vdots & \vdots & \ddots & \vdots \\ x_p^{k,i}((\ell - 1)K + 1) & x_p^{k,i}((\ell - 1)K + 2) & \dots & x_p^{k,i}((\ell - 1)K + N) \end{bmatrix} \\ &= \left[\mathbf{X}_p(\ell K), \mathbf{X}_p(\ell K + 1), \dots, \mathbf{X}_p(\ell K + N - 1) \right] \mathbf{H}_{k,i}, \end{aligned} \quad (6.17)$$

onde a matriz $\mathbf{H}_{k,i}$ de dimensão $N_H N \times N$ tem a primeira coluna formada pelos coeficientes do filtros de análise $h_{k,i}(n)$ seguidos de $(N - 1)N_H$ zeros e as colunas seguintes são versões deslocadas circularmente de N_H posições em relação às colunas anteriores. As matrizes $\mathbf{Y}_q^k(\ell)$, de dimensão $D \times N$, formadas pelos D subsequentes vetores de saída, podem ser expressas como

$$\begin{aligned} \mathbf{Y}_q^k(\ell) &= \sum_{p=1}^P \sum_{i=k-1}^{k+1} \mathbf{W}_{qp}^i(\ell) \mathbf{X}_p^{k,i}(\ell) \\ &= \begin{bmatrix} y_q^k(\ell K) & y_q^k(\ell K + 1) & \dots & y_q^k(\ell K + N - 1) \\ \vdots & \vdots & \ddots & \vdots \\ y_q^k(\ell K - D + 1) & y_q^k(\ell K - D + 2) & \dots & y_q^k(\ell K - D + N) \end{bmatrix}, \end{aligned} \quad (6.18)$$

sendo $\mathbf{X}_p^{k,i}(\ell)$ a matriz de dimensão $2K \times N$ dada por

$$\mathbf{X}_p^{k,i}(\ell) = \left[\hat{\mathbf{X}}_p^{k,i}(\ell), \hat{\mathbf{X}}_p^{k,i}(\ell - 1) \right]^T \quad (6.19)$$

e $\mathbf{W}_{pq}^i(\ell)$ a matriz tipo Sylvester de dimensão $D \times 2K$ expressa por

$$\mathbf{W}_{qp}^i(\ell) = \begin{bmatrix} w_{pq}^i(0) & w_{pq}^i(1) & \dots & w_{pq}^i(K-1) & 0 & \dots & 0 & 0 \\ 0 & w_{pq}^i(0) & w_{pq}^i(1) & \dots & w_{pq}^i(K-1) & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \dots & 0 & w_{pq}^i(0) & w_{pq}^i(1) & \dots & w_{pq}^i(K-1) & 0 \end{bmatrix}. \quad (6.20)$$

Combinando as P saídas em cada subbanda, a Eq. (6.18) pode ser expressa concisamente como

$$\begin{aligned}\mathbf{Y}^k(\ell) &= \left[\mathbf{Y}_1^k(\ell), \quad \dots, \quad \mathbf{Y}_P^k(\ell) \right]^T \\ &= \sum_{i=k-1}^{k+1} \mathbf{W}^i(\ell) \mathbf{X}^{k,i}(\ell),\end{aligned}\quad (6.21)$$

sendo

$$\mathbf{X}^{k,i}(\ell) = \left[\mathbf{X}_1^{k,i}(\ell), \quad \dots, \quad \mathbf{X}_P^{k,i}(\ell) \right]^T \quad (6.22)$$

e

$$\mathbf{W}^i(\ell) = \begin{bmatrix} \mathbf{W}_{11}^i(\ell) & \dots & \mathbf{W}_{1P}^i(\ell) \\ \vdots & \ddots & \vdots \\ \mathbf{W}_{P1}^i(\ell) & \dots & \mathbf{W}_{PP}^i(\ell) \end{bmatrix}. \quad (6.23)$$

As matrizes acima têm dimensões $PD \times N$, $2KP \times N$ e $PD \times 2KP$, respectivamente.

Numa formulação matricial, a função objetivo para BSS em subbandas é dada por

$$\mathfrak{S}^k(\ell) = \log(\det(\text{bdiag}(\mathbf{Y}^k(\ell)[\mathbf{Y}^k(\ell)]^T))) - \log(\det(\mathbf{Y}^k(\ell)[\mathbf{Y}^k(\ell)]^T)). \quad (6.24)$$

Aplicando o método do gradiente à função objetivo da Eq. (6.24), teremos

$$\nabla_{\mathbf{W}_k}^{GN} \mathfrak{S}(\ell) = 2\mathbf{W}^k(\ell) \{ \mathbf{R}_{yy}^k(\ell) - \text{bdiag}(\mathbf{R}_{yy}^k(\ell)) \} \{ \text{bdiag}(\mathbf{R}_{yy}^k(\ell)) \}^{-1}, \quad (6.25)$$

sendo

$$\mathbf{R}_{yy}^k(\ell) = \mathbf{Y}^k(\ell)[\mathbf{Y}^k(\ell)]^T. \quad (6.26)$$

a matriz de dimensão $PD \times PD$.

O algoritmo *on-line* para ajustar os coeficientes dos subfiltros de separação, considerando um sistema TITO, é dado por (omitindo o índice ℓ das matrizes de correlação para facilitar a notação)

$$\mathbf{W}^k(\ell) = \mathbf{W}^k(\ell - 1) - \mu \left[\lambda \nabla_{\mathbf{W}_k}^{GN} \mathfrak{S}(\ell - 1) + (1 - \lambda) \mu \nabla_{\mathbf{W}_k}^{GN} \mathfrak{S}(\ell) \right] \quad (6.27)$$

com

$$\nabla_{\mathbf{W}_k}^{GN} \mathfrak{S}(\ell) = 2 \begin{bmatrix} [\mathbf{R}_{y_1 y_1}^k]^{-1} [\mathbf{R}_{y_2 y_1}^k]^T \mathbf{W}_{21}^k & [\mathbf{R}_{y_1 y_1}^k]^{-1} [\mathbf{R}_{y_2 y_1}^k]^T \mathbf{W}_{22}^k \\ [\mathbf{R}_{y_2 y_2}^k]^{-1} [\mathbf{R}_{y_1 y_2}^k]^T \mathbf{W}_{11}^k & [\mathbf{R}_{y_2 y_2}^k]^{-1} [\mathbf{R}_{y_1 y_2}^k]^T \mathbf{W}_{12}^k \end{bmatrix}, \quad (6.28)$$

sendo $\mathbf{R}_{y_q y_p}^k(\ell)$ a submatriz de dimensão $D \times D$ da matriz $\mathbf{R}_{yy}^k(\ell)$, μ o passo de adaptação do algoritmo e λ o fator de esquecimento para que as amostras mais recentes dos filtros de separação tenham maior influência, permitindo o mapeamento de ambientes acústicos variantes no tempo [1].

Devido às redundâncias na matriz $\mathbf{W}_{qp}^k(\ell)$ (ver Eq. (6.20)) e por motivos de convergência [3], somente os primeiros K elementos da primeira linha desta matriz são atualizados a cada iteração. Para reduzir a complexidade computacional do algoritmo, o fator de normalização $[\mathbf{R}_{y_q y_q}^k(\ell)]^{-1}$ na Eq. (6.28) é simplificado considerando apenas o bloco de amostras mais recentes [57], ou seja ,

$$\mathbf{R}_{y_q y_q}^k(\ell) \approx [\mathbf{y}_q^k(\ell)]^T \mathbf{y}_q^k(\ell) \mathbf{I} \quad (6.29)$$

com $\mathbf{y}_q^k(\ell)$ dado pela Eq. (6.16).

6.3.1 Complexidade Computacional

O número total de multiplicações reais por bloco do sinal de entrada (NMPB) requerido pela estrutura uniforme em subbandas de M canais, considerando um sistema linear MIMO $P \times P$ e implementações na forma direta, é dado por

$$\begin{aligned} \text{NMPB}_{\text{SB}} &= \frac{PN(2M-1)(2N_p-1)}{M} + \frac{2P^2(3M-2)NKD}{M} \\ &+ \frac{PMN_p S}{M} + \frac{P^2M(N+ND)}{M} + \frac{P^2M((\sum_{j=1}^K j) + K)}{M} \end{aligned} \quad (6.30)$$

A primeira parcela corresponde ao custo para gerar a primeira linha da matriz $\mathbf{X}_p^{k,i}(\ell)$ (Eq. (6.17)), a única com informação nova. A segunda parcela representa o custo para gerar todos os blocos atrasados no tempo das saídas em todas as subbandas (Eq. (6.21)). A terceira parcela mostra o custo do processamento pelo banco de síntese para gerar os blocos das saídas em banda cheia. O quarto termo representa o custo da obtenção das matrizes $\mathbf{R}_{y_q y_p}^k(\ell)$ (ver Eq. (6.28)) para o cálculo do gradiente. Finalmente, o quinto termo indica o custo da adaptação considerando apenas a atualização das K primeiras colunas da primeira linha da matriz $\mathbf{W}_{qp}^i(\ell)$ (Eq. (6.20)). O fator no denominador de todas as parcelas indica que as operações estão sendo realizadas numa taxa de amostragem M vezes menor que a taxa em banda cheia. Para filtros de separação de altas ordens, a Eq. (6.30) pode ser

simplificada considerando somente os termos dominantes, resultando em

$$\text{NMPB}_{\text{SB}} \approx \frac{P^2(12MK^3 - 8K^3)}{M}. \quad (6.31)$$

Os resultados acima foram obtidos considerando o fator de normalização simplificado (Eq. (6.29)), $D = K$ e $N = 2K$ para o algoritmo em subbandas.

6.4 Resultados Experimentais

Em todos os nossos experimentos utilizaremos sinais amostrados em $f_s = 16 \text{ kHz}$. Nas simulações a seguir, em todos os casos, o número de blocos atrasados no tempo usados nos cálculos das correlações (Eq. (6.26)) foi $D = K$, o comprimento dos blocos dos sinais de saída foi $N = 2K$ e o comprimento dos filtros de separação foi sempre igual ao dos filtros de mistura ($U = S$). A inicialização dos filtros de separação foi feita exatamente como apresentado na Seção 6.2.2.

6.4.1 Experimento 1

Nos experimentos a seguir foram utilizados dois sinais de fala com duração variando de 60 a 75 segundos: uma voz feminina e uma voz masculina, ambas no idioma inglês. Esses sinais foram convoluídos com respostas impulsivas artificiais, obtidas em [69] e considerando o cenário virtual (proposto para competição de separação de sinais realizada durante ICA 2007) mostrado na Fig. 6.2. As misturas foram realizadas considerando diferentes condições de reverberação, utilizando as respostas impulsivas obtidas em [69], truncando-as com o tamanho desejado (ou seja, considerando somente suas S primeiras amostras). O comprimento dos filtros de separação foram fixados, em todas as simulações, igual ao dos filtros de mistura ($U = S$). Adotamos, para avaliação de desempenho, a relação sinal interferência (SIR) definida na Seção 5.3. Nesses experimentos comparamos o desempenho do algoritmo em banda cheia apresentado em [3] com o algoritmo *on-line* em subbandas proposto na Seção 6.3, considerando filtros de mistura de diferentes comprimentos: $U = 256, 512$ e 1024 .

A estrutura uniforme em subbandas foi implementada usando bancos de filtros modulados por cosseno maximamente decimados com $M = 2, 4, 8$ e 16 bandas, com filtros protótipos de comprimento $N_p = 16M$ que proporcionam reconstrução

Tabela 6.5: Parâmetros para os algoritmos em banda cheia e em subbandas.

M	Comprimento K dos subfiltros de separação			Passo
	$U = 256$	$U = 512$	$U = 1024$	μ
1	256	512	1024	2.5×10^{-4}
2	144	272	528	5×10^{-4}
4	80	144	272	1×10^{-3}
8	48	80	144	2×10^{-3}
16	32	48	80	3×10^{-3}

Tabela 6.6: SIR final (em dB).

$S = U$	SIR Final				
	$M = 1$	$M = 2$	$M = 4$	$M = 8$	$M = 16$
256	13.03	13.71	13.92	13.27	15.99
512	9.74	9.03	9.97	10.00	11.94
1024	7.25	6.59	7.45	7.51	8.84

perfeita. A Tabela 6.5 apresenta o comprimento K dos subfiltros de separação $w_{pq}^i(m)$ e os passos de adaptação utilizados nas simulações em banda cheia ($M=1$) e em subbandas para diferentes misturas. Estes valores dos passos foram escolhidos de tal maneira que os algoritmos apresentassem taxas de convergência similares.

A Fig. 6.8 mostra a evolução do SIR (para $M = 1$ e $M = 16$), a Tabela 6.6 contém o SIR final (correspondendo ao valor médio das últimas 10 iterações) para os algoritmos em banda cheia e em subbandas, e a Tabela 6.7 apresenta o número de multiplicações por bloco de acordo com as Eqs. (6.31) e (5.56).

Analisando as Tabelas 6.6 e 6.7 e a Fig. 6.8 podemos observar que com o aumento da ordem do sistema de mistura (correspondendo à maior reverberação no ambiente

Tabela 6.7: Número de multiplicações por bloco.

$S = U$	$M = 1$	$M = 2$	$M = 4$	$M = 8$	$M = 16$
256	2.68×10^8	9.56×10^7	2.05×10^7	4.87×10^6	1.51×10^6
512	2.15×10^9	6.44×10^8	1.19×10^8	2.25×10^7	5.09×10^6
1024	1.72×10^{10}	4.71×10^9	8.05×10^8	1.31×10^8	2.35×10^7

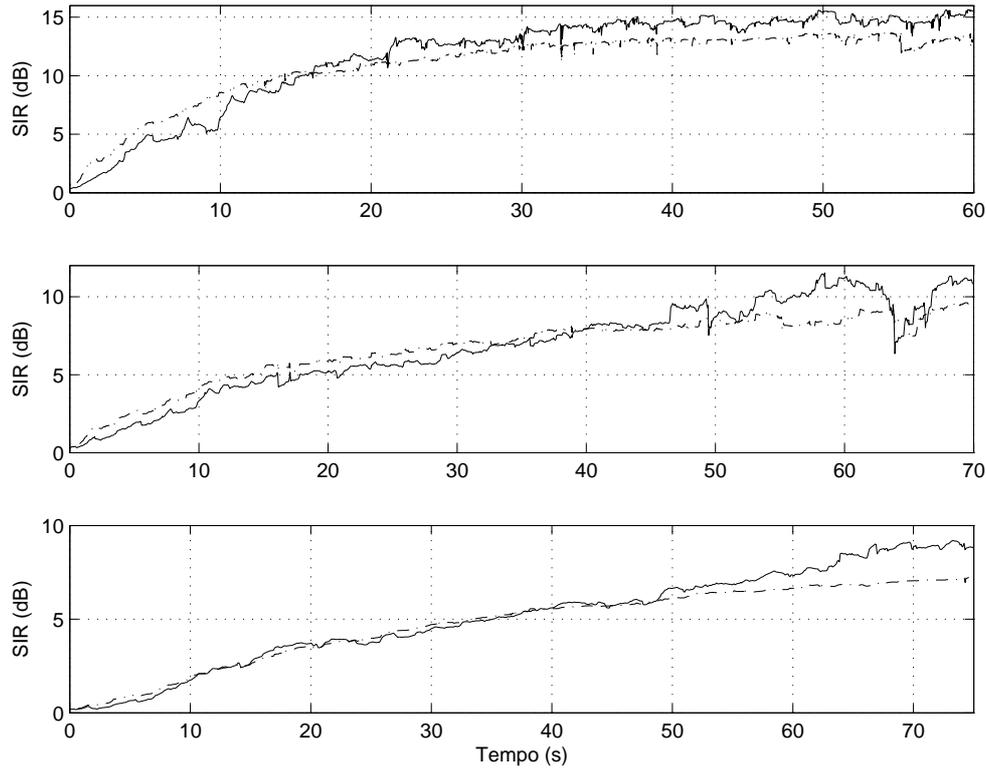


Figura 6.8: Evolução do SIR para os algoritmos em banda cheia (linha tracejada) e em subbandas (linha cheia) com filtros de mistura de diferentes comprimentos: (a) $U=256$, (b) $U=512$ e (c) $U=1024$.

acústico), as vantagens da estrutura em subbandas sobre a estrutura em banda cheia tornam-se mais evidentes, resultando num SIR final significativamente maior e num tempo de processamento menor quando comparado ao algoritmo em banda cheia.

Durante nossos experimentos monitoramos a correlação entre as estimativas das fontes nas diversas subbandas para evitar problemas de permutação, mas os algoritmos mostraram-se muito robustos, não demandando qualquer correção. A Fig. 6.9 mostra a ocorrência de permutação (amplitude igual a 1 indica ocorrência de permutação e igual a zero aponta ausência de permutação) ao longo das iterações *on-line* para as simulações da Fig. 6.8 (a).

A Fig. 6.10 apresenta os espectrogramas das fontes originais e das saídas estimadas para as simulações da Fig. 6.8 (a), mostrando a robustez dos algoritmos ao escalamento e ao branqueamento das estimativas das fontes. Na Figs. 6.10 (a) e (b) temos as fontes 1 e 2, respectivamente, e suas estimativas para o algoritmo em banda cheia; nas Figs. 6.10 (c) e (d) temos as fontes 1 e 2, respectivamente, e suas

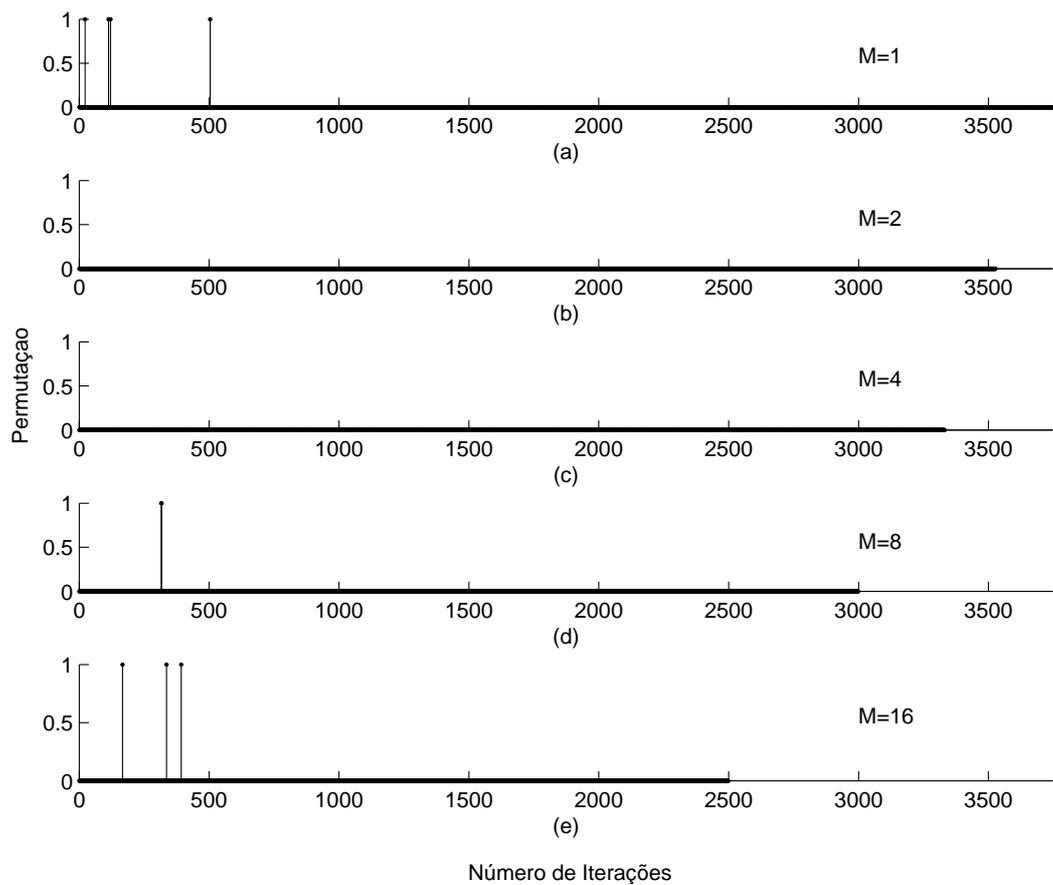


Figura 6.9: Ocorrência de permutação para os algoritmos em banda cheia ($M=1$) e em subbandas: (b) $M=2$, (c) $M=4$, (d) $M=8$ e (e) $M=16$.

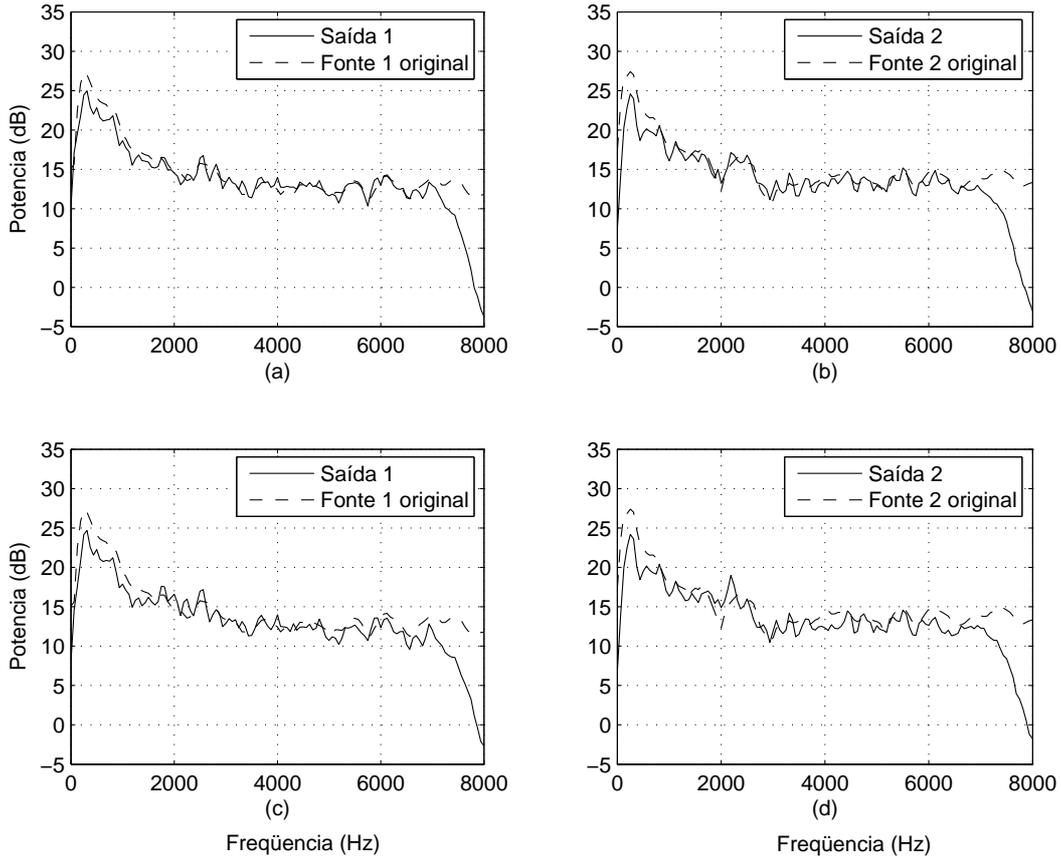


Figura 6.10: Espectrogramas das fontes originais (linha tracejada) e de suas estimativas (linha cheia).

estimativas para o algoritmo em subbandas com $M = 4$.

6.4.2 Experimento 2

Nesse experimento vamos avaliar o impacto do banco de filtros (através da utilização de diferentes filtros protótipos) nas medidas globais de desempenho apresentadas na Seção 5.3.1. Vamos utilizar o SIR^G , o SAR^G e o SDR^G para estabelecer a melhor relação entre o número de canais (M) e o comprimento do protótipo (N_p) para aplicações em BSS.

A estrutura uniforme em subbandas foi implementada usando banco de filtros modulado por cosseno maximamente decimado com $M = 8$ bandas e filtros protótipos com comprimentos $N_p = 4M, 8M$ e $16M$, com reconstrução perfeita. Os passos de adaptação utilizados foram os mesmos do experimento anterior. A Tabela 6.8 apresenta as medidas globais médias do SIR^G (Eq. (5.84)), do SAR^G (Eq.

(5.85)) e do SDR^G (Eq. (5.86)), considerando diferentes condições de reverberação: $U = 256$ e $U = 512$. Estas medidas foram obtidas através do *toolbox* para medidas de desempenho em BSS disponível em [71].

Tabela 6.8: Medidas de desempenho globais obtidas pela estrutura em subbandas com $M = 8$ e diferentes filtros protótipos, considerando diversas misturas.

	$U = 128$			$U = 256$			$U = 512$		
N_p	SIR^G	SAR^G	SDR^G	SIR^G	SAR^G	SDR^G	SIR^G	SAR^G	SDR^G
32	15,5	20,4	14,0	12,3	19,7	11,6	9,9	9,9	6,6
64	15,3	21,2	14,3	12,8	20,2	12,0	10,5	10,2	7,0
128	12,3	0,5	-1,2	12,4	0,3	-0,2	9,9	0,6	-0,6

A estrutura uniforme utilizada, que cancela a sobreposição do espectro entre subbandas adjacentes, foi derivada inicialmente para aplicações de identificação de sistemas, nas quais existia um sinal de referência utilizado na minimização do erro quadrático. Em aplicações de BSS, a proposta é minimizar a correlação entre as saídas considerando diferentes blocos de sinais em diferentes instantes de tempo, sem qualquer sinal de referência.

Analisando a Tabela 6.8 percebe-se que o uso de filtros protótipos muito seletivos ($N_p = 16M$), que reforçariam a hipótese de sobreposição do espectro somente entre bandas adjacentes, as vezes não garante um desempenho esperado em termos de SIR^G , SAR^G e SDR^G , podendo inserir distorções nas estimativas das fontes. Por outro lado, filtros pouco seletivos ($N_p = 4M$) podem comprometer a separação, porque aumentam o erro mínimo da estrutura [70]. Portanto, de acordo com a Tabela 6.8, a escolha que apresenta as melhores medidas de desempenho é a que utiliza protótipos com $N_p = 8M$. Esta relação entre N_p e M será adotada como padrão nas próximas simulações.

6.4.3 Experimento 3

As características de reverberação de salas e/ou auditórios têm despertado o interesse de vários pesquisadores. Destas, o tempo de reverberação tem relevante importância devido à qualidade dos sinais que serão captados pelos microfones. Em [72] foi apresentado o resultado de 50 ensaios realizados em salas de diferentes di-

mensões, demonstrando que o tempo de reverberação médio diminui à medida em que a frequência dos sinais de teste aumenta, principalmente pela perda na propagação devido à viscosidade do ar para frequências acima de 1 kHz [73]. Baseado nestes resultados, vamos explorar a flexibilidade da estrutura BSS em subbandas proposta na Seção 6.3 e reduzir o comprimento dos filtros de separação nas bandas de alta frequência, comparando os resultados obtido com o algoritmo *on-line* em banda cheia ($M = 1$) apresentado em [58], considerando filtros de mistura de diferentes comprimento: $U = 256, 512$ e 1024 .

Nesse experimento foram utilizados 4 sinais de voz: duas falas femininas e duas masculinas, sendo um par (feminino e masculino) em idioma inglês e outro em idioma português. Os sinais foram misturados em pares (dois a dois), e utilizando o cenário virtual mostrado na Fig. 6.2. Assim sendo, trabalhamos com 6 diferentes combinações de misturas. A estrutura uniforme em subbandas foi implementada usando bancos de filtros modulados por cosseno maximamente decimados com $M = 4, 8$ e 16 bandas, e filtros protótipos de tamanho $N_p = 8M$ com reconstrução perfeita.

A Tabela 6.9 mostra o SIR global médio, calculado a partir do SIR^G (Eq. (5.84)), obtido para cada uma das 6 possibilidades de mistura, sem redução no comprimento dos subfiltros de separação na estrutura em subbandas. A Tabela 6.10 apresenta o SIR global médio com redução do subfiltro de separação somente na banda de mais alta frequência. Esta redução foi feita considerando o comprimento de cada filtro de separação igual à metade do valor efetivamente usado para realizar as misturas, ou seja, $K_i = \left\lceil \frac{\frac{S}{2} + N_P}{M} \right\rceil$. Para as demais subbandas o valor de K foi obtido conforme a Eq. (6.11). A Tabela 6.11 exhibe o SIR global médio alcançado com a redução dos subfiltros de separação em um quarto das subbandas de mais alta frequência. As Tabelas 6.12, 6.13 e 6.14 apresentam os comprimentos dos filtros de separação utilizados nas simulações exibidas nas Tabelas 6.9, 6.10 e 6.11, respectivamente.

Analisando as tabelas acima, percebe-se que a redução no comprimento dos subfiltros de separação nas bandas de alta frequência não implica perda de desempenho na estimativa das fontes. Isto se deve ao fato do tempo de reverberação das componentes de alta frequência ser menor, além da característica dos sinais de fala de concentrar maior energia em baixas frequências.

Tabela 6.9: SIR global médio obtido pela estrutura proposta com $M = 4, 8$ e 16 subbandas e subfiltros de separação de mesmo comprimento em todas as bandas.

	SIR global médio		
M	$U = 256$	$U = 512$	$U = 1024$
1	12, 2	10, 1	7, 4
4	13, 1	9, 9	7, 2
8	12, 5	10, 7	7, 2
16	14, 2	12, 4	8, 2

Tabela 6.10: SIR global médio obtido pela estrutura proposta com $M = 4, 8$ e 16 subbandas e subfiltros de separação de comprimento reduzido somente na banda de mais alta frequência.

	SIR global médio		
M	$U = 256$	$U = 512$	$U = 1024$
1	12, 2	10, 1	7, 4
4	13, 5	9, 9	7, 1
8	13, 1	10, 7	7, 2
16	14, 2	12, 4	8, 2

6.4.4 Experimento 4

Nesse experimento vamos variar o fator de sobreposição dos blocos de amostras das saídas para avaliar o impacto desta modificação no desempenho da estrutura em subbandas uniformes aplicada à BSS. Para tanto é necessário uma pequena modificação na Eq. (6.18), substituindo o índice K pelo parâmetro J para possibilitar uma sobreposição de amostras entre blocos adjacentes diferente dos 50% que vínhamos utilizando ³. Neste caso, as matrizes $\mathbf{Y}_q^k(\ell)$ de dimensão $D \times N$, formadas pelos D

³A sobreposição de 50% é alcançada quando $D = K$.

Tabela 6.11: SIR global médio obtido pela estrutura proposta com $M = 4, 8$ e 16 subbandas e subfiltros de separação de comprimento reduzido em $\frac{1}{4}$ das bandas de mais alta frequência.

	SIR global médio		
M	$U = 256$	$U = 512$	$U = 1024$
1	12, 2	10, 1	7, 4
4	13, 0	9, 9	7, 1
8	12, 4	10, 6	7, 2
16	14, 2	12, 4	8, 2

Tabela 6.12: Comprimentos dos subfiltros de separação utilizados nas simulações apresentados na Tabela 6.9.

	Comprimento dos subfiltros $w_{pq}^i(\ell)$		
M	$U = 256$	$U = 512$	$U = 1024$
1	$S = 256$	$S = 512$	$S = 1024$
4	$K = 71$	$K = 135$	$K = 263$
8	$K = 39$	$K = 71$	$K = 135$
16	$K = 23$	$K = 39$	$K = 71$

subsequentes vetores de saída, podem ser expressas como

$$\begin{aligned}
 \mathbf{Y}_q^k(\ell) &= \sum_{p=1}^P \sum_{i=k-1}^{k+1} \mathbf{W}_{qp}^i(\ell) \mathbf{X}_p^{k,i}(\ell) \\
 &= \begin{bmatrix} y_q^k(\ell J) & y_q^k(\ell J + 1) & \dots & y_q^k(\ell J + N - 1) \\ \vdots & \vdots & \ddots & \vdots \\ y_q^k(\ell J - D + 1) & y_q^k(\ell J - D + 2) & \dots & y_q^k(\ell J - D + N) \end{bmatrix} \quad (6.32)
 \end{aligned}$$

com $1 \leq J \leq D$.

As simulações foram realizadas considerando o algoritmo em subbandas com $M = 4$ e 8 , filtros de mistura com $U = 128, 256$ e 512 coeficientes, e $J = \frac{D}{2}$ e D , considerando sobreposição de 75% e 50%, respectivamente. A Fig. 6.11 mostra a evolução do SIR médio bloco a bloco (SIR_{bb}) obtido para o algoritmo em subbandas com $M = 8$, filtros de mistura com $U = 512$ coeficientes e sinais com duração de 50 segundos.

Tabela 6.13: Comprimentos reduzidos dos subfiltros de separação utilizados nas simulações apresentados na Tabela 6.10.

	Comprimento dos subfiltros $w_{pq}^i(\ell)$		
M	$U = 256$	$U = 512$	$U = 1024$
4	$K_3 = 39$	$K_3 = 71$	$K_3 = 135$
8	$K_7 = 23$	$K_7 = 39$	$K_7 = 71$
16	$K_{15} = 15$	$K_{15} = 23$	$K_{15} = 39$

Tabela 6.14: Comprimentos reduzidos dos subfiltros de separação utilizados nas simulações apresentados na Tabela 6.11.

	Comprimento dos subfiltros $w_{pq}^i(\ell)$		
M	$U = 256$	$U = 512$	$U = 1024$
8	$K_6 = K_7 = 23$	$K_6 = K_7 = 39$	$K_6 = K_7 = 71$
16	$K_{12}=K_{13}=K_{14}=K_{15}=15$	$K_{12}=K_{13}=K_{14}=K_{15}=23$	$K_{12}=K_{13}=K_{14}=K_{15}=39$

Assumindo um sistema *TITO* e ausência de permutação, o SIR_{bb} foi calculado para cada bloco ℓ de tamanho D das estimativas das fontes da seguinte forma:

$$\text{SIR}(\ell) = \beta \text{SIR}(\ell - 1) + (1 - \beta) 10 \log_{10} \left(\frac{1}{2} \sum_{q=1}^2 \text{SIR}_q(\ell) \right), \quad (6.33)$$

sendo β um fator de lembrança e $\text{SIR}_q(\ell)$ a relação sinal-interferência da q -ésima saída (ver Seção 5.3). O fator de lembrança utilizado nestes experimentos foi $\beta = 0,99$.

Analisando a Fig. 6.11 percebe-se que o aumento na sobreposição de amostras na composição da matriz $\mathbf{Y}_q^k(\ell)$ resulta na melhoria significativa da taxa de crescimento do SIR_{bb} ; porém o número de iterações (ℓ) também aumenta, implicando um maior custo computacional. Uma alternativa seria rodar o algoritmo inicialmente com uma grande sobreposição entre blocos, reduzindo-a ao longo do tempo.

6.4.5 Experimento 5

Nesse experimento vamos comparar o desempenho do algoritmo *on-line* proposto com os algoritmos que participaram da competição promovida em 2008 pela comunidade de separação de sinais, *Signal Separation Evaluation Campaign (SiSEC*

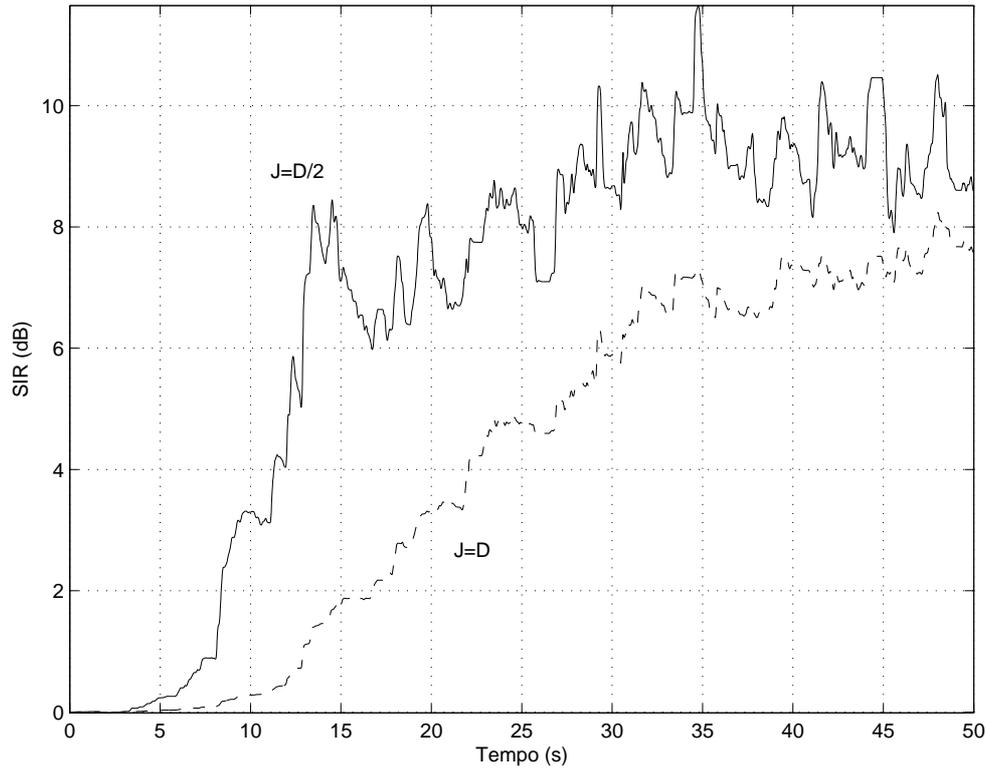


Figura 6.11: Evolução do SIR bloco a bloco para o algoritmo com $M = 8$ subbandas e diferentes sobreposições: $J = \frac{D}{2}$ (linha cheia) e D (linha tracejada).

2008), com os resultados apresentados em [74] e em um painel de discussões durante a *8th International Conference on Independent component Analysis and Signal Separation (ICA 2009)*. Os cenários acústicos reais propostos para a tarefa de estimação das fontes originais a partir das misturas de sinais de música e de fala são descritos a seguir:

Sala 1 - Câmara com paredes almofadadas de dimensões $1,5 \text{ m} \times 2,0 \text{ m} \times 2,5 \text{ m}$ (largura \times profundidade \times altura). As fontes foram colocadas aleatoriamente na sala, no chão ou numa mesa. Os microfones também foram posicionados aleatoriamente a aproximadamente 50 cm das paredes em diferentes alturas;

Sala 2 - Sala de conferência de tamanho médio de dimensões $10 \text{ m} \times 8 \text{ m} \times 3 \text{ m}$. As fontes foram colocadas randomicamente na sala, no chão, em uma mesa ou em um banco, com distância de 2 metros entre as fontes. Os microfones foram colocados ao longo da parede mais próxima às fontes, a aproximadamente 1 metro da parede, em alturas diferentes e uniformemente espaçados de aproximadamente 1 metro;

Sala 3 - Escritório de tamanho médio de dimensões $3,0 \text{ m} \times 3,0 \text{ m} \times 2,5 \text{ m}$.

As fontes foram colocadas ao acaso na sala, no chão, em uma mesa ou em um banco, com distância de 1,5 metros entre as fontes. Os microfones foram posicionados em alturas diferentes e uniformemente espaçados de aproximadamente 1 metro.

Para todas as salas foram disponibilizados aos concorrentes 4 canais de misturas com duração de 10 segundos, amostrados em $f_s = 16 \text{ KHz}$, capturados por 4 microfones e variando o número de fontes presentes, ou seja, $P = 2, 3$ e 4 . Portanto, foram explorados casos de misturas determinadas (sistemas MIMO 4×4) e sobre-determinadas (sistemas MIMO 4×2 e 4×3). As fontes originais utilizadas nas três salas foram sempre as mesmas e nenhuma delas foi disponibilizada para evitar qualquer tipo de manipulação durante o processo de separação.

Ao todo, seis algoritmos foram selecionados com os resultados tabelados e disponibilizados em [75], a saber:

Algoritmo 1 (A1) - A BSS é realizada no domínio da frequência, através de filtragem linear baseada em ICA [76]. O problema de permutação é resolvido combinando a TDOA (*Time Difference Of Arrival*) com a estimação espectral recursiva;

Algoritmo 2 (A2) - A BSS é feita aplicando ICA no domínio da frequência com a correção da permutação utilizando uma função objetivo baseada na verossimilhança [77];

Algoritmo 3 (A3) - Método BSS no domínio da frequência que utiliza IVA (*Independent Vector Analysis*) como uma extensão do ICA para funções multivariáveis [78].

Algoritmo 4 (A4) - Esse algoritmo foi proposto por Itahashi Takashi do *Department of Brain Science and Engineering* do *Kyushu Institute of Technology*. Como nenhuma informação adicional foi fornecida, esse algoritmo não será considerado neste trabalho.

Algoritmo 5 (A5) - Realiza a BSS no domínio do tempo, sendo uma versão modificado do conhecido algoritmo baseado no gradiente natural (*Infomax*), que

minimiza os problemas de divergência devido à inicialização dos filtros de separação e à magnitude dos sinais de entrada, impondo restrições ao gradiente [79].

Algoritmo 6 (A6) - Efetua a BSS no domínio do Tempo e consiste numa extensão do conhecido algoritmo *FastICA*, que combina um branqueamento das misturas através de predição linear LS (*Least-Squares*) com um procedimento adaptativo para impor restrições que levem a filtros de separação paraunitários [80].

Nossas simulações foram realizadas considerando o algoritmo em subbandas com $M = 8$ e fator de esquecimento $\lambda = 0,2$. O passo de adaptação (μ), o tamanho da sobreposição entre os blocos (J) e o comprimento dos subfiltros de separação (K) são mostrados na Tabela 6.15. Assumimos o sistema como determinado (2×2) trabalhando com os sinais de apenas 2 microfones, embora houvesse 4 canais de misturas disponíveis.

Tabela 6.15: Parâmetros usados nas simulações das salas 1, 2 e 3.

Salas	μ	J	K
1	8×10^{-4}	16	135
2	1×10^{-3}	16	263
3	1×10^{-3}	8	135

Para visualizar e avaliar subjetivamente a qualidade da separação dos sinais de música apresentamos a evolução no tempo dos últimos 250 ms das misturas e das estimativas das fontes presentes nas Salas 1, 2 e 3.

A Fig. 6.12 mostra a evolução dos sinais na Sala 1 da seguinte forma: em (a) e (b) 2 canais de misturas capturadas pelos microfones, em (c) e (d) as saídas obtidas pelo algoritmo A1, em (e) e (f) pelo A2, em (g) e (h) pelo A3, em (i) e (j) pelo A5, em (l) e (m) pelo A6 e em (n) e (o) pelo algoritmo em subbandas proposto (A_{sb}).

Na falta das fontes originais não foi possível calcular o SIR para mensurar nossos resultados e comparar com os resultados obtidos pelos competidores. Alternativamente, escutamos atentamente as estimativas obtidas para entender o que foi separado. Numa das saídas percebemos um trecho de *jazz* acompanhado de flauta,

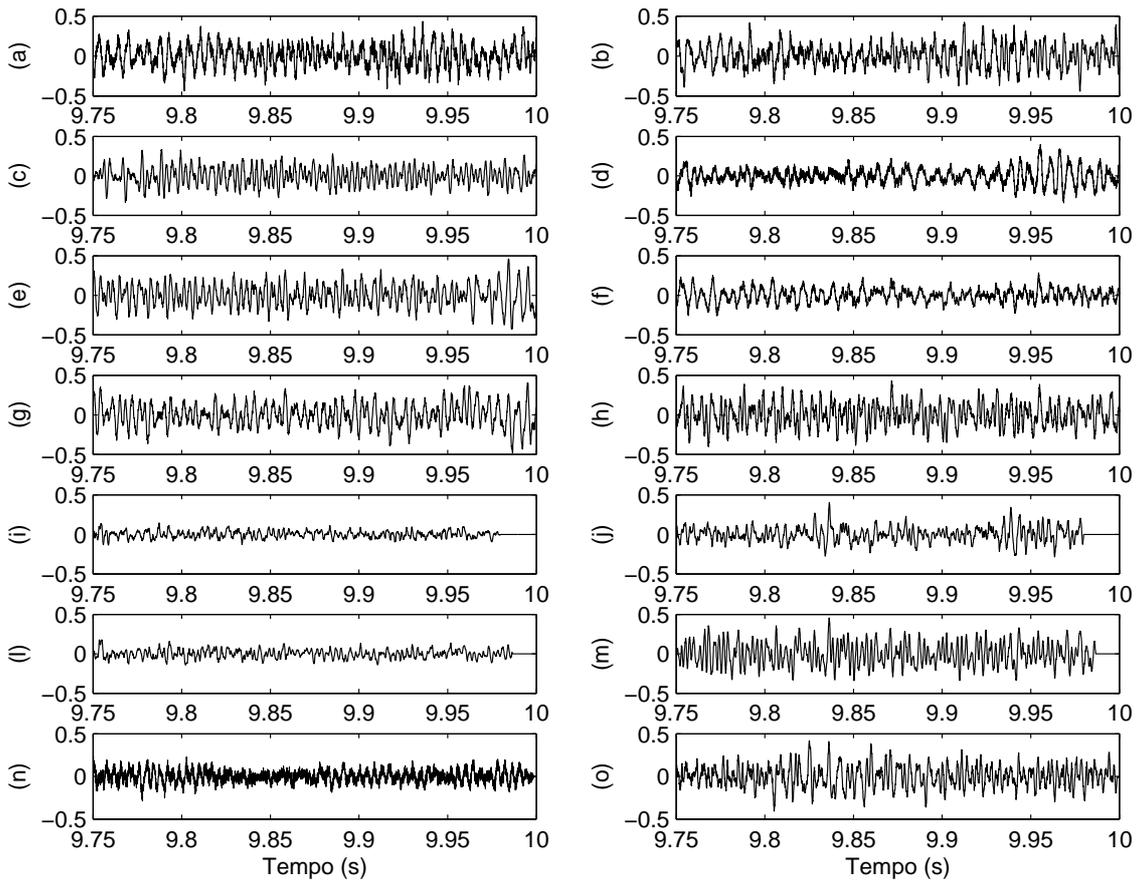


Figura 6.12: Evolução dos últimos 250 ms dos sinais dos microfones e das saídas dos algoritmos participantes da competição e do algoritmo proposto para as misturas da Sala 1.

baixo e bateria, a qual denominamos fonte A, na outra, a qual nomeamos de fonte B, temos um trecho diferente de *jazz* acompanhado de saxofone, de baixo e de um instrumento harmônico que parece ser um tipo de xilofone.

Analisando a Fig. 6.12, auxiliado pela audição das misturas e estimativas das saídas dos algoritmos de BSS aplicados a Sala 1, que estão disponíveis para *download* em [81], podemos concluir que os algoritmos A_{sb} proposto e A1 tiveram um desempenho superior aos demais. O algoritmo A_{sb} apresentou na saída 1 a estimativa da fonte A, que tem uma característica mais “ruidosa” devido à presença da bateria, e na saída 2 a estimativa da fonte B, que possui um aspecto mais “harmônico” devido ao xilofone. O algoritmo A1 apresentou na saída 1 a estimativa da fonte B e na saída 2 a estimativa da fonte A.

As Figs. 6.13 e 6.14 mostram a evolução dos sinais nas Salas 2 e 3, respectivamente, da seguinte forma: em (a) e (b) 2 canais de misturas capturadas pelos microfones, em (c) e (d) as saídas obtidas pelo algoritmo A1, em (e) e (f) pelo A3, em (g) e (h) pelo A5, em (i) e (j) pelo A6 e em (l) e (m) pelo algoritmo A_{sb} .

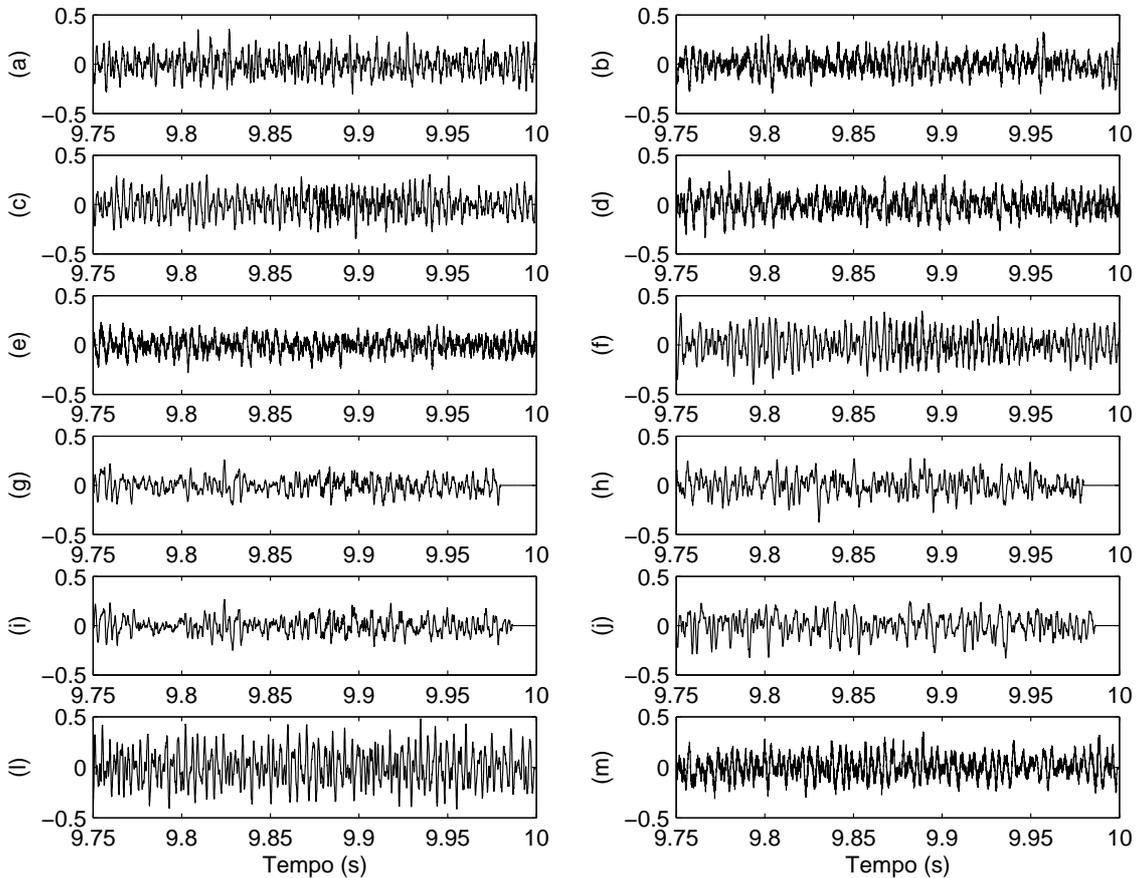


Figura 6.13: Evolução dos últimos 250 ms dos sinais dos microfones e das saídas dos algoritmos participantes da competição e do algoritmo proposto para as misturas da Sala 2.

Analisando as Fig. 6.13 e as Fig. 6.14, auxiliado pela audição das misturas e estimativas das saídas dos algoritmos de BSS aplicados as Salas 2 e 3, que estão disponíveis para *download* em [81], podemos dizer que em ambientes mais reverberantes o desempenho do algoritmo A_{sb} proposto foi bem superior a todos os outros na tarefa de separação de sinais de música, apresentando na saída 1 a estimativa da fonte B e na saída 2 a estimativa da fonte A.

Analisando os resultados divulgados em [75], pudemos observar que nenhum algoritmo foi eficiente para todos os tipos de misturas, sendo alguns mais eficazes

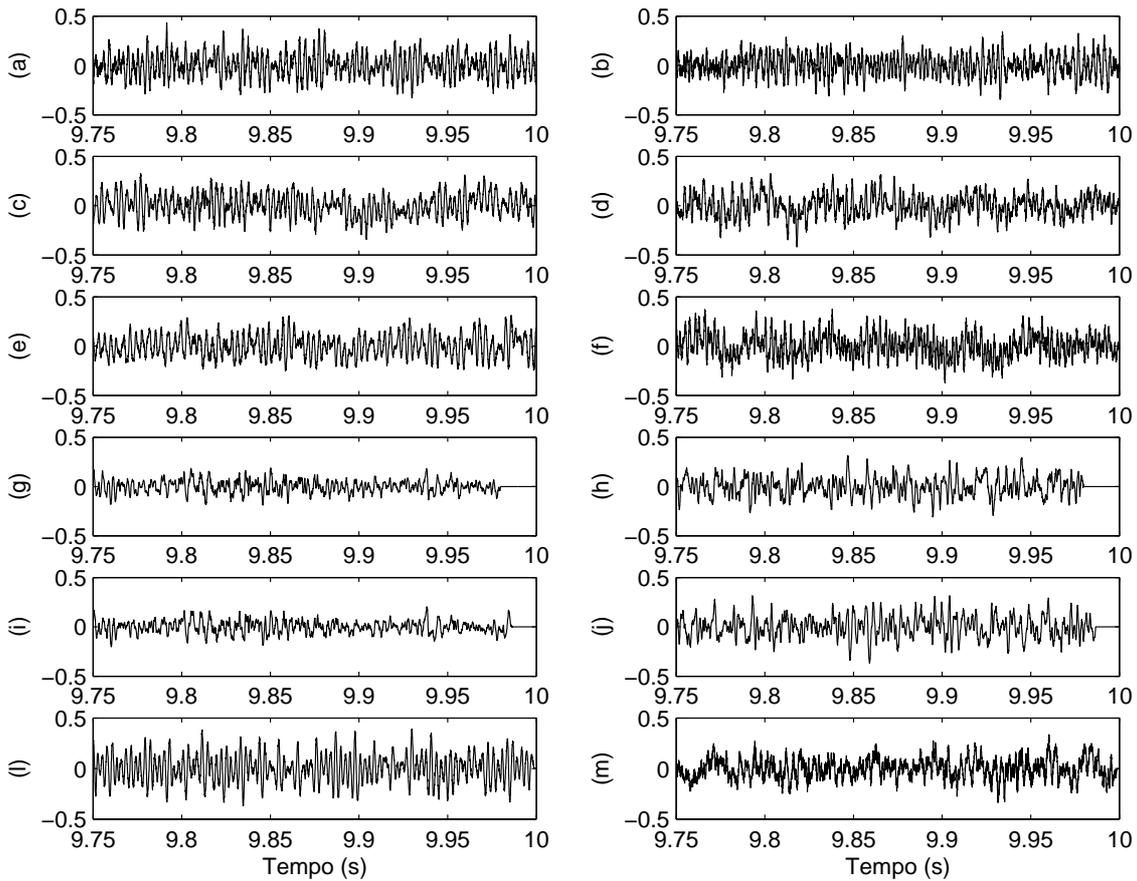


Figura 6.14: Evolução dos últimos 250 ms dos sinais dos microfones e das saídas dos algoritmos participantes da competição e do algoritmo proposto para as misturas da Sala 3.

em separação de fala e outros em separação de música. O algoritmo proposto teve um desempenho muito bom para sinais de música; entretanto, para sinais de fala, suas estimativas ainda apresentavam-se um pouco misturadas. Por isso não apresentamos os resultados para as demais Salas 4, 5, *C* e *O* (propostas para competição) que envolviam sinais de voz.

Não foram apresentados os resultados dos algoritmos A2 nas Salas 2 e 3, e do A4 nas Salas 1, 2 e 3, por não estarem disponíveis em [75] (provavelmente porque os autores não devem ter submetido resultados para estas tarefas).

6.5 Conclusões

Nesse capítulo propusemos dois novos algoritmos em subbandas para separação cega de fontes, que empregam bancos de filtros não-uniforme subamostrado e uniforme

criticamente decimado na decomposição dos sinais observados pelos sensores. No primeiro algoritmo, os filtros de separação têm diferentes comprimentos e são adaptados de forma *off-line* em diferentes taxas, ou seja, as bandas de frequências mais baixas são mais estreitas e operam em menores taxas de amostragem. No segundo algoritmo, são utilizados filtros extras para cancelamento da sobreposição do espectro entre bandas adjacentes e os filtros que desfazem as misturas são adaptados de forma *on-line* sempre na taxa de amostragem crítica.

A adaptação é realizada por algoritmos baseados no gradiente natural com esquemas de normalizações simplificados, que resultam na redução do custo computacional. Simulações computacionais foram realizadas com sinais de fala, mostrando as vantagens das estruturas em subbandas sobre a estrutura em banda cheia em relação ao tempo de processamento, taxa de convergência e razão sinal interferência. Outros resultados experimentais foram apresentados para o algoritmo *on-line* com sinais de música e ambientes acústicos diversos, obtendo-se desempenho superior a outros dois algoritmos implementados no domínio do tempo em banda cheia e três algoritmos que realizam a separação no domínio da frequência.

Capítulo 7

Conclusões e Sugestões

Neste capítulo serão apresentados as conclusões finais deste trabalho, assim como algumas sugestões que poderão ser investigadas oportunamente. Os principais resultados obtidos foram a obtenção de uma estrutura não-uniforme criticamente decimada de filtragem adaptativa em subbandas, capaz de promover melhores taxas de convergências, e a análise de convergência do respectivo algoritmo adaptativo; derivação de uma estrutura uniforme em subbandas com decimação crítica para separação cega de fontes e comparação do seu desempenho com os de outras técnicas no domínio do tempo e da frequência.

7.1 Conclusões

O grande destaque que o processamento adaptativo multitaxas vem alcançando, principalmente em modelagem de sistemas, cancelamento de ecos acústicos em sistemas de teleconferência, equalização de sinais e separação cega de fontes, motivou a busca de novos algoritmos para estruturas em subbandas.

Nos Capítulos 2 e 3, foi feita uma revisão sobre tópicos importantes para o desenvolvimento deste trabalho, dos quais podemos citar: algoritmos adaptativos em banda cheia, banco de filtros e sistemas multitaxas, implementação e desenvolvimento de protótipos para banco de filtros modulado por cosseno. No Capítulo 4 foi apresentada a estrutura adaptativa não-uniforme derivada neste trabalho, seu algoritmo adaptativo e as análises de convergência, do erro quadrático médio e coeficientes ótimos, e a análise da complexidade computacional. Os resultados experimentais mostraram melhoria na taxa de convergência do algoritmo adaptativo para sinais de entrada coloridos e validaram o resultado da análise do MSE teórico.

No Capítulo 5 foi apresentado o problema de separação cega de fontes, as formas de abordagem nos domínios do tempo e da frequência e mostradas formas de avaliações de desempenho. No Capítulo 6 propusemos duas estruturas adaptativas em subbandas para recuperação de fontes: uma não-uniforme subamostrada e a outra uniforme maximamente decimada com filtros extras para cancelamento de *aliasing* entre bandas adjacentes. Para estas estruturas, a primeira implementada de forma *off-line* e a segunda de forma *on-line*, foram descritos os algoritmos adaptativos baseados no método do gradiente natural com diferentes procedimentos de normalização, e calculadas suas complexidades computacionais. Os resultados experimentais envolvendo sinais de áudio mostraram a melhoria na relação sinal-interferência e a redução do custo computacional obtidos com as estruturas em subbandas, principalmente envolvendo misturas de música.

7.2 Sugestões

Dando seguimento a este trabalho, podemos sugerir para pesquisas futuras os seguintes itens:

1. Explorar formas eficientes de implementação para banco de filtros modulados por cosseno;
2. Desenvolver protótipos que minimizem o MSE final da estrutura e/ou acelerem a taxa de convergência;
3. Avaliar o comportamento das estruturas em subbandas em cancelamento de ecos acústicos, utilizando os protótipos dedicados;
4. Derivar a estrutura não-uniforme com amostragem crítica para aplicações em separação cega de fontes;
5. Investigar o uso de estatísticas de ordens superiores na função objetivo usada em BSS ;
6. Implementar outros métodos BSS em subbandas;
7. Testar o uso de funções *wavelets* convencionais na estrutura não-uniforme;

Referências Bibliográficas

- [1] HAYKIN, S., *Adaptive Filter Theory - 4th ed.*. Prentice-Hall: Englewood Cliffs, NJ, 2001.
- [2] GRIESBACH, J. D., LIGHTNER, M. R., ETTER, D. M., “Constituent subband allocation for system modeling Nonuniform subband adaptive filters”, *IEEE Trans. Signal Process.*, v. 53, n. 2, pp. 539–549, Feb. 2005.
- [3] BUCHNER, H., AICHNER, R., KELLERMANN, W., “A Generalization of Blind Source separation Algorithms for Convolutional Mixtures Based on Second-Order Statistics”, *IEEE Trans. on Speech and Audio Process.*, v. 13, n. 1, pp. 120–134, Jan. 2005.
- [4] DINIZ, P. S. R., *Adaptive Filtering: Algorithms and Practical Implementation*. Kluwer Academic Publishers: Boston, MA, 1997.
- [5] VAIDYANATHAN, P. P., *Multirate Systems and Filter Banks*. Prentice-Hall: Englewood Cliffs, NJ, 1993.
- [6] DINIZ, P. S. R., DA SILVA, E. A. B., NETTO, S. L., *Processamento Digital de Sinais - Projeto e Análise de Sistemas*. Bookman, 2004.
- [7] STRANG, G., NGUYEN, T., *Wavelets and Filter Banks*. Wellesley - Cambridge Press: Wellesley, MA, 1996.
- [8] CROISER, A., ESTEBAN, D., GALAND, C., “Perfect channel splitting by use of interpolation, decimation, tree decomposition techniques”. In: *Int. Conf. on Info. Sciences / Systems*, pp. 443–446, Patras, Greece, Aug. 1976.

- [9] SMITH, M. J. T., BARNWELL-III, T. P., “A procedure for designing exact reconstruction filter banks for tree structure subband coders”. In: *Proc. IEEE Int. Conf. Acoust. Speech, and Signal Processing*, pp. 2711–2714, San Diego, CA, Mar. 1984.
- [10] MINTZER, F., “Filters for distortion-free two-band multirate filter banks”, *IEEE Trans. on Acoust., Speech, and Signal Processing*, v. ASSP-33, pp. 626–630, Jun. 1985.
- [11] VAIDYANATHAN, P. P., “Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having perfect reconstruction property”, *IEEE Trans. on Acoust., Speech, and Signal Processing*, v. ASSP-35, pp. 476–492, Apr. 1987.
- [12] VAIDYANATHAN, P. P., “Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques”, *IEEE ASSP magazine*, v. 4, pp. 4–20, Jul. 1987.
- [13] MALVAR, H. S., “Modulated QMF filter banks with perfect reconstruction”, *Electronics Letters*, v. 26, pp. 906–907, Jun. 1990.
- [14] KOILPILLAI, R. D., VAIDYANATHAN, P. P., “Cosine-modulated filter banks satisfying perfect reconstruction”, *IEEE Trans. on Signal Processing*, v. SP-40, pp. 770–783, Apr. 1992.
- [15] RAMSTAD, T. A., “Cosine modulated analysis-synthesis filter bank with critical sampling and perfect reconstruction”. In: *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, pp. 1789–1792, Toronto, Canada, May 1991.
- [16] NUSSBAUMER, J. H., “Pseudo QMF filter banks”. In: *IBM Tech. disclosure Bulletin*, v. 24, pp. 3081–3087, Nov. 1981.
- [17] NGUYEN, T. Q., “Near Perfect Reconstruction pseudo-QMF banks”, *IEEE Trans. on Signal Processing*, v. 42, pp. 65–76, Jan. 1994.

- [18] KOILPILLAI, R. D., VAIDYANATHAN, P. P., “Cosine-modulated FIR filter banks satisfying perfect reconstruction”, *IEEE Trans. Signal Processing*, v. 40, pp. 770–783, Apr. 1992.
- [19] NGUYEN, T. Q., “Digital filter banks design - quadratic constrained formulation”, *IEEE Trans. on Signal Processing*, v. 43, pp. 2103–2108, Sep. 1995.
- [20] KLIEWER, J., MERTINS, A., “Oversampled cosine-modulated filter banks with arbitrary system delay”, *IEEE Trans. on Signal Processing*, v. 46, pp. 941–955, Apr. 1998.
- [21] GILLOIRE, A., VETTERLI, M., “Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation”, *IEEE Trans. on Signal Processing*, v. 40, pp. 1862–1875, Aug. 1992.
- [22] PETRAGLIA, M. R., MITRA, S. K., “Performance analysis of adaptive structures based on subband decomposition”. In: *Proc. IEEE Int. Symp. on Circuits and Systems*, pp. 60–63, Chicago, IL, May 1993.
- [23] PETRAGLIA, M. R., MITRA, S. K., “Adaptive FIR filter structure based on the generalized subband decomposition of FIR filters”, *IEEE Trans. on Circuits and Systems-II: Analog and Digital Processing*, v. 40, n. 9, pp. 354–362, Jun. 1993.
- [24] PETRAGLIA, M. R., ALVES, R. G., “A new adaptive subband structure with critical sampling”. In: *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 1861–1864, Phoenix, AR, Apr. 1999.
- [25] MCCLOUD, M. L., ETTER, D. M., “Subband adaptive filtering with time-varying nonuniform filter banks”. In: *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 1953–1956, 1997.
- [26] VASCONCELLOS, R. T. B., PETRAGLIA, M. R., ALVES, R. G., “A New Critically Sampled Non-Uniform Subband Adaptive Structure”. In: *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 3713–3716, 2001.

- [27] OAKMAN, A., NAYLOR, P. A., “Dynamic structures for non-uniform adaptive filters”. In: *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 3717–3720, 2001.
- [28] PETRAGLIA, M. R., BATALHEIRO, P. B., “Nonuniform Subband Adaptive Filtering With Critical Sampling”, *IEEE Trans. on Signal Processing.*, v. 56, pp. 565–575, Feb. 2008.
- [29] USEVITCH, B. E., ORCHARD, M. T., “Adaptive filtering using filter banks”, *IEEE trans. on Circuits and Systems-II: Analog and Digital Signal Processing*, v. 43, n. 3, pp. 255–265, Mar. 1996.
- [30] PETRAGLIA, M. R., TORRES, J. C. B., “Performance Analysis of Adaptive Filter Structure Employing Wavelet and Sparse Subfilters”, *IEE Proc. - Vis. Image Signal Process.*, v. 149, n. 2, pp. 115–119, Apr. 2002.
- [31] PAPOULIS, E. V., STATHAKI, T., “A Transmultiplexer-Based Adaptive Filtering Structure using a New Adaptation Scheme”. In: *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 53–56, 2002.
- [32] VAIDYANATHAN, P. P., “Quadrature mirror filter banks, M-band extensions, and perfect-reconstruction technique”, *IEEE ASSP Mag.*, v. 4, n. 3, pp. 4–20, Jul. 1987.
- [33] PETRAGLIA, M. R., BATALHEIRO, P. B., “Filter Bank Design for an Adaptive Subband Structure with Critical Sampling using a New Adaptation Scheme”. In: *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 114–117, 2005.
- [34] PETRAGLIA, M. R., ALVES, R. G., DINIZ, P. S. R., “New structures for adaptive filtering in subbands with critical sampling”, *IEEE Trans. Signal Process.*, v. 48, n. 12, pp. 3316–3327, Dec. 2000.
- [35] MORGAN, D. R., “Slow asymptotic convergence of LMS acoustic echo cancellers”, *IEEE Trans. Speech Audio Process.*, v. 3, pp. 126–136, Mar. 1995.
- [36] DE LEON II, P. L., ETTER, D. M., “Experimental results with increased bandwidth analysis filters in oversampled subband acoustic echo cancellers”, *IEEE Signal Process. Lett.*, v. 2, n. 1, pp. 1–3, Jan. 1995.

- [37] SAWADA, H., MUKAI, R., ARAKI, S., et al., “A Robust and Precise Method for Solving the Permutation Problem of Frequency-Domain Blind Source Separation”, *IEEE Transactions on Speech And Audio Processing*, v. 12, n. 5, pp. 530–538, 2004.
- [38] NACEUR, M. S., LOGHMARI, M. A., BOUSSEMA, M. R., “The Contribution of the Sources Separation Method in the Decomposition of Mixed Pixels”, *IEEE Transactions on Geoscience and Remote Sensing*, v. 42, n. 11, pp. 2642–2653, 2004.
- [39] ACERNESE, F., CIARAMELLA, A., MARTINO, S. D., et al., “Neural Networks for Blind-Source Separation of Stromboli Explosion Quakes”, *IEEE Transactions on Neural Networks*, v. 14, n. 1, pp. 167–175, 2003.
- [40] ICHIR, M. M., MOHAMMAD-DJAFARI, A., “Hidden Markov Models for Wavelet-Based Blind Source Separation”, *IEEE Transactions on Image Processing*, v. 15, n. 7, pp. 1887–1899, 2006.
- [41] ERDOGAN, A. T., “Globally Convergent Deflationary Instantaneous Blind Source Separation Algorithm for Digital Communication Signals”, *IEEE Transactions on Signal Processing*, v. 55, n. 5, pp. 2182–2192, 2007.
- [42] BUCHNER, H., AICHNER, R., KELLERMANN, W., “A Generalization of a Class of Blind Source Separation Algorithms for convolutive mixtures”. In: *Proc. Int. Symp. Independent Component Analysis Blind Signal Separation*, pp. 945–950, Apr. 2003.
- [43] HYVÄRINEN, A., KARHUNEN, H., OJA, E., *Independent Component Analysis*. John Wiley & Sons, 2001.
- [44] SARUWATARI, H., KAWAMURA, T., NISHIKAWA, T., et al., “Blind Source separation Based on a Fast-Convergence Algorithm Combining ICA and Beamforming”, *IEEE Trans. Audio, Speech, and Language Process.*, v. 14, n. 2, pp. 666–678, Mar. 2006.
- [45] SMARAGDIS, P., “Blind Separation of convolved mixtures in the Frequency Domain”, *Neurocomputing*, v. 22, pp. 21–34, 1998.

- [46] NISHIKAWA, T., SARUWATARI, H., SHIKANO, K., “Comparison of Time-Domain ICA, Frequency-Domain ICA and Multistage ICA for Blind Source Separation”. In: *Proc. Eur. Signal Processing Conf.*, v. 2, pp. 15–18, 2002.
- [47] ARAKI, S., MAKINO, S., R.AICHNER, et al., “Subband-Based Blind Separation for Convolutional Mixtures of speech”, *IEICE Trans. Fundamentals*, v. E88-A, pp. 3593–3603, 2005.
- [48] COMON, P., “Independent component analysis: A new concept?” *Signal Process.*, Elsevier, v. 36, n. 3, pp. 287–314, Apr. 1994.
- [49] AMARI, S., CICHOCKI, A., YANG, H., “A New Learning Algorithm for Blind Signal Separation”. In: *Advances in Neural Information Process. Syst.*, v. 8, pp. 757–763, The MIT Press, 1996.
- [50] CHOI, S., CICHOCKI, A., PARK, H., et al., “Blind source separation and independent component analysis: a review”, *Neural Information Process. - Lett. and Rev.*, v. 06, n. 1, pp. 1–57, Jan. 2005.
- [51] CICHOCKI, A., ICHI AMARI, S., *Adaptive Blind Signal and Image Processing*. John Wiley & Sons, 2002.
- [52] PAPOULIS, A., *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 2001.
- [53] LEE, T. W., GIROLAM, M., BELL, A. J., et al., “A Unifying Information-Theoretic Framework for Independent Component Analysis”, *Computers and Mathematics with Applications*, v. 31, n. 11, pp. 1–12, Mar. 2000.
- [54] LEE, T. W., *Independent Component Analysis - Theory and Applications*. Kluwer, 1998.
- [55] BELL, A. J., SEJNOWSKI, T. J., “An Information-Maximization Approach to Blind Separation and Blind Deconvolution”, *Neural Computation*, v. 07, n. 06, pp. 1129–1159, 1995.

- [56] MATSUOKA, K., OHYA, M., KAWAMOTO, M., “A Neural Net for Blind Separation of Nonstationary Signals”, *Neural Networks*, v. 08, n. 3, pp. 411–419, 1995.
- [57] AICHNER, R., BUCHNER, H., KELLERMANN, W., “Exploiting Narrowband Efficiency for Broadband Convolutional Blind Source Separation”, *EURASIP Journal on Applied Signal*, v. 2007, pp. 1–9, Sep. 2006.
- [58] AICHNER, R., BUCHNER, H., ARAKI, S., et al., “On-Line Time-Domain Blind Source Separation of Nonstationary Convolved Signals”. In: *Proc. Eur. Signal Processing Conf.*, pp. 987–992, Apr. 2003.
- [59] PARRA, L., SPENCE, C., “Convolutional Blind Separation of Non-Stationary Sources”, *IEEE Transactions on Speech and Audio Processing*, v. 8, n. 3, pp. 320–327, 2000.
- [60] MURATA, N., IKEDA, S., ZIEHE, A., “An approach to blind source separation based on temporal structure of speech signals”, *Neurocomputing*, v. 41, pp. 1–24, 2001.
- [61] IKRAM, M. Z., MORGAN, D. R., “A beamforming approach to permutation alignment for multichannel frequency-domain blind speech separation”, *Proc. ICASSP*, pp. 881–884, 2002.
- [62] KURITA, S., SARUWATARI, H., KAJITA, S., et al., “Evaluation of Blind Signal Separation Method Using Directivity Pattern Under Reverberant Conditions”. In: *Proc. Eur. Signal Processing Conf.*, pp. 3140–3143, Jun. 2000.
- [63] SAWADA, H., MUKAI, R., ARAKI, S., et al., “A Robust and Precise Method for Solving the Permutation Problem of Frequency-Domain Blind Source Separation”, *IEEE Transactions on Speech And Audio Processing*, v. 12, n. 5, pp. 530–538, 2004.
- [64] MATSUOKA, K., NAKASHIMA, S., “Minimal distortion principle for blind source separation”, *Proc. ICA*, pp. 722–727, 2001.

- [65] VINCENT, E., GRIBONVAL, R., FÉVOTTE, C., “Performance Measurement in Blind Audio Source Separation”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 14, n. 4, pp. 1462–1469, 2006.
- [66] RUSSELL, I., XI, J., MERTINS, A., et al., “Blind Source Separation of Nonstationary Convolutively Mixed Signals in the Subband Domain”, *Proc. ICASSP*, pp. 481–484, 2004.
- [67] PARK, H., OH, S., LEE, S., “A Uniform Oversampled Filter Bank Approach to Independent Component Analysis”, *Proc. ICASSP*, pp. 249–252, 2003.
- [68] VINCENT, E., SAWADA, H., BOFILL, P., et al., “First stereo audio source separation evaluation campaign: data, algorithms and results”. In: *Proc. Int. Conf. on Independent Component Analysis and Signal Separation*, pp. 1–8, Sep. 2007.
- [69] Disponível em: <http://sassec.gforge.inria.fr/>. Acesso em: 29 jun. 2009, 10:00:00.
- [70] PETRAGLIA, M. R., BATALHEIRO, P. B., “Filter Bank Design for a Subband Adaptive Filtering Structure With Critical Sampling”, *IEEE Trans. on Circuits And Systems I - Fundamental Theory And Applications.*, v. 51, pp. 1194–1202, Jun. 2004.
- [71] FÉVOTTE, C., GRIBONVAL, R., VINCENT, E., “BSS_EVAL toolbox user guide”, *IRISA Technical Report 1706*, Apr. 2005. Disponível em: http://www.irisa.fr/metiss/bss_eval/. Acesso em: 29 jun. 2009, 10:01:00.
- [72] EVEREST, F. A., POHLMANN, K., *Master HandBook of Acoustics*. McGraw-Hill, 2001.
- [73] BERANEK, L., *Music, Acoustics and Architecture*. John Wiley & Sons, 1962.
- [74] VINCENT, E., BOFILL, S. A. P., “The 2008 Signal Separation Evaluation Campaign: A Community-Based Approach to Large-Scale Evaluation”. In: *Proc. Int. Conf. on Independent Component Analysis and Signal Separation*, pp. 734–741, Mar. 2009.

- [75] Disponível em: <http://sisec.wiki.irisa.fr/tiki-index.php>. Acesso em: 29 jun. 2009, 10:02:00.
- [76] NESTA, F., OMOLOGO, M., SVAIZER, P., “Multiple TDOA estimation by using a state coherence transform for solving the permutation problem in frequency-domain BSS”. In: *Proc. Machine Learning for Signal Processing*, pp. 43–48, Oct. 2008.
- [77] LEE, I., “Permutation correction in blind source separation using sliding sub-band likelihood function”. In: *Proc. Int. Conf. on Independent Component Analysis and Signal Separation*, pp. 767–774, Mar. 2009.
- [78] LEE, I., KIM, T., LEE, T.-W., *Blind speech separation*. Springer Netherlands, 2007.
- [79] DOUGLAS, S. C., MALAY GUPTA, “Scaled Natural Gradient Algorithms for Instantaneous and Convolutive Blind Source Separation”. In: *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 637–640, Apr. 2007.
- [80] DOUGLAS, S. C., GUPTA, M., SAWADA, H., et al., “Spatio-Temporal FastICA Algorithms for the Blind Separation of Convolutive Mixtures”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 15, n. 5, pp. 1511–1520, 2007.
- [81] Disponível em: <http://www.eng.uerj.br/deptos/mostra.php?id=206&tipo=trabalhos>. Acesso em: 29 jun. 2009, 10:03:00.
- [82] LUENBERGER, D. G., *Introduction to Linear and Nonlinear Programming*. Addison Wesley: Reading, MA, 1973.