



COMPRESSÃO DE SINAIS USANDO DECOMPOSIÇÕES ATÔMICAS COM
BASE EM DICIONÁRIOS REDUNDANTES

Michel Pompeu Tcheou

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.

Orientadores: Eduardo Antônio Barros da
Silva
Luiz Wagner Pereira Biscainho

Rio de Janeiro
Março de 2011

COMPRESSÃO DE SINAIS USANDO DECOMPOSIÇÕES ATÔMICAS COM
BASE EM DICIONÁRIOS REDUNDANTES

Michel Pompeu Tcheou

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.

Examinada por:

Prof. Eduardo Antônio Barros da Silva, Ph.D.

Prof. Luiz Wagner Pereira Biscainho, D.Sc.

Prof. Moisés Vidal Ribeiro, D.Sc.

Prof. Marcello Luiz Rodrigues de Campos, Ph.D.

Prof. Lisandro Lovisolo, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2011

Tcheou, Michel Pompeu

Compressão de Sinais Usando Decomposições Atômicas
Com Base em Dicionários Redundantes/Michel Pompeu
Tcheou. – Rio de Janeiro: UFRJ/COPPE, 2011.

XVI, 167 p.: il.; 29, 7cm.

Orientadores: Eduardo Antônio Barros da Silva

Luiz Wagner Pereira Biscainho

Tese (doutorado) – UFRJ/COPPE/Programa de
Engenharia Elétrica, 2011.

Referências Bibliográficas: p. 154 – 167.

1. Processamento de sinais. 2. Compressão. 3.
Decomposições atômicas. I. da Silva, Eduardo Antônio
Barros *et al.* II. Universidade Federal do Rio de Janeiro,
COPPE, Programa de Engenharia Elétrica. III. Título.

Agradecimentos

Agradeço a Deus pela saúde e disposição, fundamentais para a realização e conclusão deste trabalho. Aos meus orientadores, Eduardo e Luiz Wagner, pela excelente orientação e pelo entusiasmo com que me ajudaram a conduzir a tese. Ao professor Lisandro Lovisolo que acompanhou de perto o desenvolvimento da tese, proporcionando importantes contribuições sempre através do caos. Ao professor Moisés Ribeiro por aceitar fazer parte da banca examinadora. Aos pesquisadores Marco Antonio Macciola e André Miranda por disponibilizarem a ferramenta de análise automática de faltas em linhas de transmissão. Aos professores Paulo Diniz e Marcello Campos, responsáveis por minha iniciação científica. Ao CNPq pelo suporte financeiro no início da pesquisa.

À minha querida esposa Thais, pelo apoio, carinho e compreensão durante o período da tese. Aos meus pais, Tcheou Cheng Ho e Chou Sin Whei Tcheou, pelo apoio e dedicação, sempre prezando pela minha formação acadêmica e profissional. À minha irmã, Cíntia, aos meus tios, Atuim, Inja e Didiu, e aos meus sogros, Adelia e Takumi, pelo incentivo e suporte durante esses anos.

Aos meus amigos de doutorado, Filipe Diniz, Mauricio Quêlhas e Tadeu Ferreira que me acompanharam nesta jornada árdua de trabalho. Aos meus amigos do LPS por permitirem que as horas gastas neste laboratório fossem sempre as mais agradáveis possíveis. Em especial: João Baptista, Fábio Freeland, Amaro Lima, Rodrigo Torres, José Fernando Leite, Nuno Rodrigues, Bernardo Costa, Alexandre Leizor, Alan Tygel, Leonardo Nunes, Wallace Martins, Markus Lima, Guilherme Pinto, Camila Gussen, Ana Fernanda, Flávio Ávila, Alessandro Dutra, Gabriel Matos, João Dias, Fabiano Castoldi, Arnaldo Gunzi, Rafael Amado, Andreas Ellmauthaler, Ana Luisa Santos, Marcello Artimos, Carlo Marcello, Leonardo Baltar, Iúri Kothe, José Márcio Faier, Jürgen Wittmann.

À minha chefe Maria Elvira, pelo incentivo e apoio para o término da tese; e aos meus amigos do CEPTEL, em especial: Welington de Oliveira, Renato Cabral, Saulo Aires, Vitor Duarte, Débora Jardim, André Diniz, Roberto Pinto, Ana Lúcia e Carlos Henrique Sabóia, Tiago Norbiato, Igor Raupp, Manuel Fragoso, Leonardo Moraes, Francisco Faria, Diego Oneto, Michel Ennes, Cesar Vilasboa, Cesar Moreira Jr. e Thiago Baptista.

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

COMPRESSÃO DE SINAIS USANDO DECOMPOSIÇÕES ATÔMICAS COM BASE EM DICIONÁRIOS REDUNDANTES

Michel Pompeu Tcheou

Março/2011

Orientadores: Eduardo Antônio Barros da Silva
Luiz Wagner Pereira Biscainho

Programa: Engenharia Elétrica

Este trabalho tem como objetivo propor um arcabouço de compressão de sinais usando decomposições atômicas com base em dicionários redundantes. Em geral, utiliza-se um dicionário fixo na etapa de decomposição e codificam-se os coeficientes e os índices dos átomos. Nesta tese, utiliza-se um novo conceito de compressão de sinais usando decomposições atômicas com base em múltiplos dicionários. Além de transmitir os coeficientes e os índices dos átomos, o codificador precisa enviar ao decodificador informação adicional a respeito do dicionário utilizado.

Considerando este paradigma de compressão, desenvolvem-se dois sistemas de compressão, um aplicado a sinais elétricos e outro aplicado a sinais de áudio. O sistema de compressão aplicado a sinais elétricos usa um método de decomposição com base em um dicionário de senóides amortecidas, denominado MPSA (*Matching Pursuit* por Senóides Amortecidas); e utiliza uma técnica de otimização taxa distorção através de curvas taxa-distorção operacionais. Avalia-se o sistema com o auxílio de uma ferramenta de análise automática de faltas.

O sistema de compressão de sinais de áudio é desenvolvido com base no sistema aplicado a sinais elétricos, realizando-se algumas modificações para adequá-lo a sinais de áudio. Na decomposição, usa-se um dicionário de senóides amortecidas, porém, com frequências da escala musical cromática. Na alocação de bits, utiliza-se uma técnica de otimização taxa-distorção por faixa de energia atômica junto a um codificador aritmético de modelo de probabilidade fixo. Avalia-se o sistema através de uma ferramenta de medição objetiva de qualidade de 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.)

SIGNAL COMPRESSION FOR ATOMIC DECOMPOSITIONS BASED ON REDUNDANT DICTIONARIES

Michel Pompeu Tcheou

March/2011

Advisors: Eduardo Antônio Barros da Silva

Luiz Wagner Pereira Biscainho

Department: Electrical Engineering

In this work, we present a framework for signal compression using atomic decompositions based on redundant dictionaries. In general, a fixed dictionary is used in the decomposition stage, then one encodes the coefficients and indices of the atoms. In this thesis, we use a concept of signal compression using atomic decompositions based on multiple dictionaries. Besides transmitting the coefficients and indices of the atoms, the encoder needs to send additional information with respect to the dictionary used.

Given this compression paradigm, we develop two compression systems, one applied to electric signals and another applied to audio signals. The compression system for electric signals uses an atomic decomposition method based on a dictionary of damped sinusoids, called DSMP (Matching Pursuit using Damped Sinusoids) and uses a rate-distortion optimization technique based on operational rate-distortion curves. This compression system is evaluated by using an automatic tool for fault analysis.

The compression system for audio signals is developed based on the compression system for electric signals. As decomposition, one uses a dictionary of damped sinusoids but with frequencies of the chromatic musical scale. We use a rate-distortion optimization technique that performs bit allocation inside each energy range. The symbols are encoded by an arithmetic coder with fixed model. This compression system is evaluated using an objective tool for measuring the quality of audio signals.

Sumário

Lista de Figuras	x
Lista de Tabelas	xvi
1 Introdução	1
1.1 Análise e Síntese de Sinais	1
1.2 Representações Compactas	2
1.3 Decomposições Adaptativas Redundantes	3
1.4 Proposta da Tese	4
1.5 Organização	6
2 Fundamentos Teóricos	8
2.1 Decomposições Atômicas	8
2.1.1 <i>Matching Pursuit</i>	11
2.1.2 Dicionários Redundantes Parametrizados	13
2.1.3 Decomposições Coerentes	15
2.2 Compressão usando Decomposições Atômicas	17
2.2.1 Compressão usando Decomposições Atômicas com Múltiplos Dicionários	17
2.2.2 Compressão usando Decomposições Atômicas com Dicionários Parametrizados	18
2.2.3 Seleção do Dicionário de Compromisso Taxa-Distorção Ótimo	20
3 Decomposição Atômica de Sinais Elétricos	25
3.1 Modelo de Sinais Elétricos	25
3.2 Algoritmo de Decomposição – MPSA	28
3.2.1 Dicionário Parametrizado de Senoides Amortecidas	28
3.2.2 Aproximação Preliminar do Resíduo	29
3.2.3 Reconhecimento de Fenômenos Localizados	31
3.2.4 Critério de Parada	34
3.3 Aplicações na Modelagem de Sinais Elétricos	39
3.3.1 Modelagem Coerente de Sinais	39

3.3.2	Supressão de Ruído por Síntese	41
3.3.3	Extração da Componente Fundamental e Separação de Transitórios	41
3.3.4	Filtragem da componente “DC”	42
4	Compressão de Sinais Elétricos Usando Decomposições Atômicas	44
4.1	Motivação	44
4.2	Sistema de Compressão	47
4.3	Quantização do Livro de Estruturas	48
4.4	Otimização Taxa-Distorção	49
4.4.1	Resultados Experimentais	50
4.5	Avaliação de Desempenho do Sistema de Compressão	52
4.5.1	Análise de Faltas	53
4.5.2	Sistema de Análise Automática de Faltas	54
4.5.3	Avaliação do Sistema de Compressão	58
5	Decomposição Atômica de Sinais de Áudio	65
5.1	<i>Matching Pursuit</i> Aplicado a Sinais de Áudio	65
5.2	MPSA de Escala Musical Cromática - MPSACr	66
5.2.1	Segmentação do Sinal	67
5.2.2	Dicionário de Escala Musical Cromática	68
5.2.3	Estratégia de Redução de Complexidade Computacional	71
5.2.4	Heurísticas de Aproximação	77
5.2.5	Heurística de Continuidade de Átomos entre Quadros	79
5.2.6	Critério de Parada	86
5.3	Simulações Computacionais	87
6	Compressão de Áudio usando Decomposições Atômicas	96
6.1	Introdução	96
6.2	Sistema de Compressão	98
6.2.1	Quantização e Otimização Taxa-Distorção através de Curvas Operacionais	100
6.2.2	Codificação	104
6.2.3	Simulações Computacionais	106
6.3	Otimização Taxa-Distorção por Faixa de Energia Atômica	110
6.3.1	Motivação	110
6.3.2	Problema de Otimização Taxa-Distorção	112
6.3.3	Alocação Ótima de Bits	115
6.3.4	Função-Custo Taxa-Distorção	117
6.4	Aperfeiçoamento do Sistema de Compressão	122

6.4.1	Quantização e Otimização Taxa-Distorção por Faixa de Energia Atômica	123
6.4.2	Distribuição Gaussiana Generalizada	131
6.4.3	Codificação	133
6.4.4	Simulações Computacionais	137
7	Conclusão	150
7.1	Contribuições	152
7.2	Trabalhos Futuros	153
	Referências Bibliográficas	154

Lista de Figuras

1.1	Diagrama de blocos de um sistema de análise e síntese para modelagem de sinais.	2
1.2	Graus de liberdade de uma senoide, de uma <i>wavelet</i> e de um átomo exponencial. Essas figuras foram cedidas pelo autor de [1].	4
2.1	Análise e síntese de um sinal utilizando decomposições atômicas com base em um dicionário D	9
2.2	Representação gráfica da projeção ortogonal no primeiro passo da decomposição de \mathbf{x}	11
2.3	Representação em tempo-frequência do átomo g_γ	14
2.4	Representações no plano tempo-frequência do sinal (a) utilizando a STFT (b), a distribuição de Choi-Williams (c), e a distribuição de Wigner-Ville (d). Estas figuras foram cedidas pelo autor de [2].	15
2.5	(a) sinal decomposto e (b) taxa de aproximação nos 300 primeiros passos da decomposição utilizando um dicionário de exponenciais.	16
2.6	Compressão por decomposições atômicas usando um único dicionário redundante D	17
2.7	Compressão por decomposições atômicas usando um conjunto \mathcal{D} de dicionário redundantes.	18
2.8	Problema da compressão de decomposições atômicas com base em dicionários parametrizados.	19
2.9	Compressão de decomposições atômicas usando um conjunto de dicionários parametrizados.	21
2.10	Interpretação gráfica da otimização Lagrangeana.	22
2.11	Fecho convexo contendo os pontos ótimos em termos de taxa-distorção.	23
2.12	Traçando o fecho convexo. Neste caso, $\theta_{\min} = \theta_1$	24
3.1	Exemplo de forma de onda gerada a partir do modelo de sinais de perturbações elétricas.	26

3.2	Diagrama de blocos do algoritmo de decomposição atômica de sinais. Na primeira iteração, a chave está na posição 1, e nas demais iterações permanece na posição 0.	28
3.3	Exemplos de átomos do dicionário de senoides amortecidas.	29
3.4	Resultado da otimização dos parâmetros do átomo.	30
3.5	Falha do MP na tentativa de encontrar estruturas coerentes com os padrões do sinal.	31
3.6	De cima para baixo, os resultados referentes à busca no dicionário discreto de exponenciais, à maximização da aproximação do resíduo através da otimização dos parâmetros e à busca por melhor suporte temporal.	32
3.7	De cima para baixo, os resultados referentes à busca no dicionário discreto de exponenciais, à maximização da aproximação do resíduo através da otimização dos parâmetros e à busca por melhor suporte temporal.	33
3.8	De cima para baixo, os resultados da busca por melhor suporte temporal do átomo, seguida pela quantização da frequência e pela nova busca por melhor suporte.	34
3.9	De cima para baixo, os resultados da busca por melhor suporte temporal do átomo, seguida pela quantização da frequência e pela busca por melhor suporte novamente.	35
3.10	De cima para baixo, os resultados da quantização da frequência seguida pela busca por suporte, e da identificação da senoide pura.	36
3.11	De cima para baixo, os resultados da quantização da frequência seguida pela busca por suporte, e da identificação da senoide pura.	37
3.12	Comportamento da taxa de aproximação para ruídos de diferentes comprimentos.	37
3.13	Buscando estruturas coerentes em sinais de distúrbios elétricos reais provenientes de [3].	38
3.14	Geração de um modelo coerente de sinais sujeito a diferentes razões sinal-ruído.	40
3.15	Desempenho da supressão de ruído por síntese.	41
3.16	Extração da fundamental e separação de transitórios de sinais elétricos.	43
3.17	Filtro de Fourier aplicado após a filtragem da componente “DC” do sinal sintético.	43
4.1	Esquema simplificado de um sistema elétrico.	46
4.2	Diagrama de blocos do processo de compressão de sinais de perturbações elétricas.	47

4.3	Sinais elétricos e suas versões reconstruídas com taxas de compressão de 0,95 e 0,578 bit/amostra respectivamente, com seus erros de reconstrução.	51
4.4	Curvas operacionais obtidas utilizando-se os conjuntos completo e reduzido de quantizadores para os sinais R_1 e R_2 . ($[*]= [4]$)	52
4.5	Um sinal de tensão com módulo e ângulo calculados usando filtro de Fourier.	55
4.6	Comportamento do módulo da corrente (acima) e seu índice de variabilidade correspondente (abaixo).	56
4.7	Detecção do tempo de início de falta. Acima: erro absoluto médio versus intervalos de taxas de bits; Abaixo: FDA dos erros absolutos na detecção do início da falta abaixo de 1 bit/amostra. A linha tracejada indica 3,5 ms.	60
4.8	Cálculo de módulo e ângulo dos fasores de corrente e tensão nos períodos de pré-falta e falta: (a) RMS médio dos erros percentuais do módulo versus intervalos de taxas de bits, (b) e RMS médio dos erros absolutos de ângulo versus intervalos de taxas de bits.	61
4.9	Classificação de falta. Acima: indica quais taxas de bits resultam em erro de classificação; Abaixo: mostra a probabilidade de erro de classificação da falta em diferentes intervalos de taxas de bits.	63
4.10	Localização de faltas: (a) erros absoluto versus taxa de bits, (b) erro absoluto médio versus intervalos de taxas de bits.	64
4.11	FDA dos erros absolutos no intervalo 1 a 2 bits/amostra em relação à localização de falta.	64
5.1	Diagrama de blocos do método de decomposição MPSACr aplicado a sinais de áudio. Na primeira iteração, a chave está na posição 1, e nas demais iterações permanece na posição 0.	68
5.2	Diagrama de blocos do MPSA. Na primeira iteração, a chave está na posição 1, e nas demais iterações permanece na posição 0.	68
5.3	JNDF (<i>just noticeable difference in frequency</i>) em função da frequência para uma modulação na frequência em 4 Hz. Adaptado de [5]	70
5.4	Descontinuidade de um átomo presente em dois quadros consecutivos devido à segmentação do sinal (a). A linha vertical pontilhada indica o local da separação entre os dois blocos e (b) apresenta a ampliação do trecho circulado em (a).	79
5.5	Decomposição de sinais usando heurística de continuidade de átomos entre quadros consecutivos.	80

5.6	Extensão de uma átomo do quadro $i - 1$ para quadro i : (a) o átomo do quadro $i - 1$ a ser estendido, (b) a envoltória deste átomo, (c) a extensão da envoltória ao quadro i e (d) o resultado da extensão do átomo do quadro $i - 1$ ao quadro i	81
5.7	Eliminação de (a) descontinuidade do átomo apresentado na Figura 5.4. A linha vertical pontilhada indica o local da separação entre os dois quadros e (b) apresenta a ampliação do trecho circulado em (a).	83
5.8	Sinal com variação de decaimento entre quadros: (a) extensão de uma átomo para um quadro consecutivo e (b) o resultado da obtenção do decaimento ótimo. A linha vertical pontilhada indica o local da separação entre os dois quadros.	85
5.9	Sinal com mudança radical de padrão entre quadros: (a) extensão de uma átomo para um quadro consecutivo e (b) o resultado da obtenção do decaimento ótimo. A linha vertical pontilhada indica o local da separação entre os dois quadros.	85
5.10	Comportamento da taxa de aproximação para ruídos de diferentes comprimentos.	87
5.11	Forma de onda do sinal <i>pianoA3</i>	89
5.12	Número de átomos versus índice do quadro para o sinal <i>pianoA3</i>	89
5.13	Forma de onda do sinal <i>fluteA4</i>	90
5.14	Número de átomos versus índice do quadro para o sinal <i>fluteA4</i>	90
5.15	Forma de onda do sinal <i>celloA4</i>	91
5.16	Número de átomos versus índice do quadro para o sinal <i>celloA4</i>	91
5.17	Forma de onda do sinal <i>tirofinal</i>	92
5.18	Número de átomos versus índice do quadro para o sinal <i>tirofinal</i>	92
5.19	Aproximações relativas às seis primeiras iterações da decomposição do quadro 10 do sinal <i>pianoA3</i>	94
5.20	Átomos continuados que formam a representação do sinal <i>pianoA3</i> . As linhas verticais tracejadas indicam os limites entre quadros.	95
6.1	Esquema geral de compressão de sinais.	99
6.2	Compressão de sinais de áudio realizando a decomposição do sinal via MPSACr e a otimização taxa-distorção através de curvas operacionais.	99
6.3	Curva operacional contendo os pontos ótimos em termos de taxa-distorção.	103
6.4	Otimização taxa-distorção conjunta de três quadros de quantização consecutivos.	104
6.5	Organização dos símbolos dentro do feixe de bits.	105

6.6	Curva operacional de taxa-distorção do sinal <i>pianoA3</i> e pares taxa-distorção calculados variando-se as taxas entre o coeficiente e os parâmetros do átomo. O círculo indica os pares taxa-distorção pertencentes à curva operacional.	107
6.7	Comportamento taxa-distorção marginal em função da taxa aplicada à amplitude. O eixo das abcissas corresponde à taxa total de bits por amostra e o das ordenadas corresponde ao erro quadrático. Cada símbolo está associado a uma taxa de bits aplicada à amplitude e dentro de cada grupo de símbolos variam-se as taxas do decaimento e da fase.	107
6.8	Comportamento taxa-distorção marginal em função da taxa aplicada ao decaimento com a taxa de amplitude fixada em 8 bits. O eixo das abcissas corresponde à taxa total de bits por amostra e o das ordenadas corresponde ao erro quadrático. Cada símbolo está associado a uma taxa de bits aplicada ao decaimento e dentro de cada grupo de símbolos varia-se a taxa da fase.	108
6.9	Comportamento taxa-distorção marginal em função da taxa aplicada à fase com a taxa de amplitude fixada em 8 bits. O eixo das abcissas corresponde a taxa total de bits por amostra e o das ordenadas corresponde ao erro quadrático. Cada símbolo está associado a uma taxa de bits aplicada à fase e dentro de cada grupo de símbolos varia-se a taxa do decaimento.	109
6.10	Curvas operacionais de taxa-distorção de diferentes sinais de áudio.	110
6.11	Esquema de otimização taxa-distorção considerando a relevância dos átomos na representação do sinal em termos de energia, inserida em um contexto de codificação.	116
6.12	Construção de faixas de energia com espaçamento não-uniforme e seus representantes médios.	119
6.13	Compressão de sinais de áudio realizando a decomposição do sinal via MPSACr, a otimização taxa-distorção por faixa de energia atômica seguida de codificação aritmética.	122
6.14	Distribuição gaussiana generalizada para diferentes parâmetros de forma p mas considerando a mesma variância	132
6.15	Organização dos símbolos dentro do feixe de bits.	133

6.16	Codificação de entropia dos símbolos usando codificador aritmético com modelo de probabilidade fixo. Os símbolos são codificados e decodificados na seguinte ordem de parâmetros: coeficientes, decaimentos, frequências, fases, amostras iniciais do suporte e tamanhos do suporte. As funções de probabilidade dos coeficientes e dos parâmetros do átomo exponencial são representadas por $p_\alpha, p_\rho, p_\xi, p_\phi, p_{n^s}$ e p_{n^L}	135
6.17	Codificação da mensagem <i>abba!</i>	137
6.18	Decodificação da mensagem <i>abba!</i>	138
6.19	Superfícies de distorção média dos átomos exponenciais por faixa de energia para diferentes combinações de passos de quantização do decaimento (q_ρ) e da fase (q_ϕ). Ilustra-se o caso de quatro faixas de energia, sendo a primeira faixa a de menor energia, e a quarta faixa a de maior energia.	140
6.20	Histogramas da amplitude normalizada, do decaimento e do tamanho do suporte do átomo ajustados através de distribuição gaussiana generalizada.	141
6.21	Histogramas dos parâmetros de fase, amostra inicial do suporte temporal do átomo e frequência.	142
6.22	Curvas de taxa-distorção obtidas através da codificação de quatro sinais realizando a otimização taxa-distorção através de curvas operacionais (<i>curvaOper</i>), por faixa de energia atômica modelando o coeficiente e os parâmetros do átomo como distribuição uniforme (<i>faixaEnergia</i>) e por faixa de energia atômica ajustando as distribuições do coeficiente e dos parâmetros do átomo através de gaussiana generalizada (<i>faixaEnergiaGG</i>).	144
6.23	O sinal <i>berimbau</i> original (à esquerda) e o resíduo (à direita) com relação à representação obtida removendo-se os átomos cujos coeficientes se tornariam nulos com a quantização a uma taxa de 320 kbps, porém sem quantizar os coeficientes e os parâmetros dos átomos restantes.	148
6.24	Versão ampliada de um trecho do resíduo apresentado na Figura 6.23. As linhas verticais tracejadas indicam os limites entre os quadros de decomposição.	149

Lista de Tabelas

3.1	Tabela de valores do critério de parada obtidos a partir da decomposição de sinais de ruído de diferentes comprimentos utilizando MP com base no dicionário de exponenciais com parâmetros contínuos.	38
4.1	Taxa de bits e razão de compressão para distintos erros de localização.	64
5.1	Tabela de valores do critério de parada obtidos a partir da decomposição de sinais de ruído de diferentes comprimentos utilizando MPSACr com base no dicionário discreto de senoides amortecidas de frequências da escala musical cromática.	86
5.2	Média de átomos por quadro para diferentes sinais de áudio.	93
6.1	Parâmetros de um átomo continuado em dois quadros de 4096 amostras e a representação deste átomo agrupando estes dois quadros em um quadro de 8192 amostras.	100
6.2	Distorção total do sinal e a soma das distorções individuais de cada átomo para diferentes taxas de compressão e sinais distintos.	111
6.3	Codificação da mensagem <i>abba!</i>	137
6.4	Avaliação através de PEAQ da representação obtida via MPSACr de diferentes sinais de áudio com razão sinal-ruído de 60 dB.	146
6.5	Avaliação do desempenho do sistema de compressão proposto através do PEAQ. A coluna <i>Nota-SQ</i> se refere às notas associadas às representações dos sinais sem a quantização dos coeficientes e dos parâmetros e com a remoção dos átomos cujos coeficientes se tornariam nulos com a quantização associada à taxa de compressão desejada. A coluna <i>Nota-Q</i> se refere às notas dadas aos sinais codificados considerando a quantização dos coeficientes e dos parâmetros dos átomos.	147

Capítulo 1

Introdução

A modelagem de sinais consiste em descrever matematicamente e de forma suficientemente acurada seus fenômenos intrínsecos. A representação matemática desses fenômenos pode ser obtida e processada através de um esquema de análise e síntese. A análise corresponde ao processo de ajuste de um modelo a um sinal particular, e a síntese é responsável por reconstruir o sinal utilizando o modelo e a informação obtida pela análise. Neste trabalho, utiliza-se um modelo aditivo que representa os sinais como uma soma ponderada de componentes mais simples.

1.1 Análise e Síntese de Sinais

Um sistema generalizado de análise e síntese de sinais é apresentado na Figura 1.1. O bloco de análise retorna, de acordo com um modelo matemático, uma representação do sinal que é usada pelo bloco de síntese para produzir uma estimativa do sinal. Se essa estimativa não for perfeita, a diferença entre o sinal original e o reconstruído é diferente de zero, resultando em um resíduo. Quando esse resíduo é considerado pelo bloco de análise, denomina-se o sistema como de análise-por-síntese [6, 7]. De um modo geral, os dados de análise podem ser processados para fins de supressão de ruído [8], filtragem [9], compressão [10–12], realce (*enhancement*) ou modificação de sinais [6, 13].

Uma abordagem comum para a modelagem matemática de sinais envolve a utilização de blocos de análise e síntese formados por bancos de filtros [14, 15]. Alternativamente, a análise dos sinais é realizada através de técnicas como a transformada de Fourier de tempo curto, a transformada *wavelet* discreta, a transformada cosseno discreta e a transformada ortogonal com sobreposição [16, 17]. Neste trabalho, propõe-se um sistema de análise-por-síntese cujo bloco de análise realiza aproximações sucessivas de sinais através de um algoritmo denominado *Matching Pursuit* [16, 18]. É comum como alternativa à representação de sinais de voz e áudio, por exemplo, recorrer-se à modelagem física que busca mimetizar o trato vocal

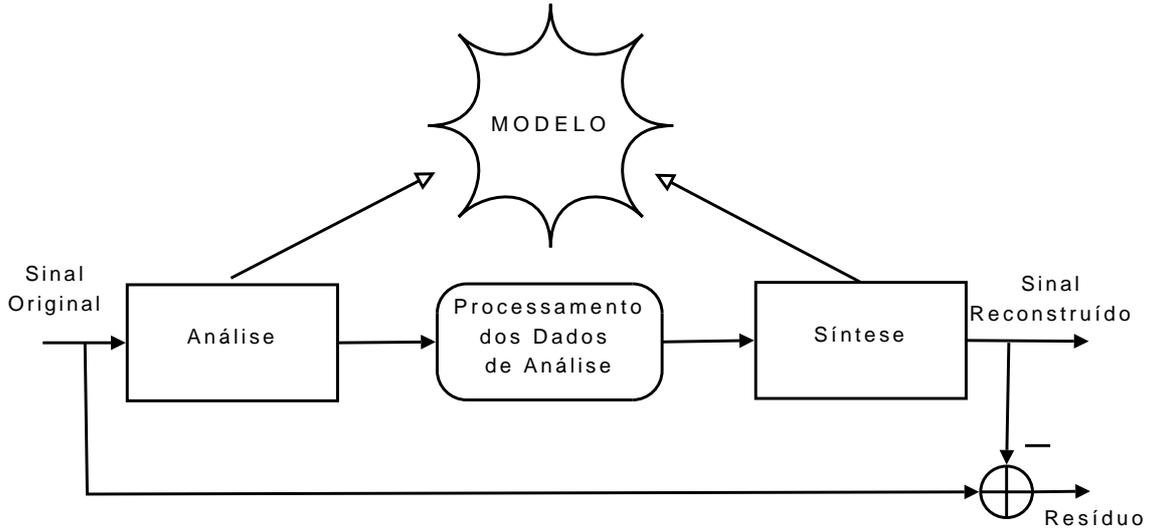


Figura 1.1: Diagrama de blocos de um sistema de análise e síntese para modelagem de sinais.

e o instrumento musical [19–21]. Neste caso, o modelo precisa de uma excitação adequada de forma tal que sua saída corresponda a uma boa aproximação do sinal.

O foco deste trabalho é em modelagem matemática de sinais com base em modelos aditivos, também conhecidos como atômicos, da forma

$$\mathbf{x} = \sum_{i=1}^M \alpha_i \mathbf{g}_{\gamma(i)}, \quad (1.1)$$

em que o sinal \mathbf{x} é representado como uma soma ponderada de componentes básicas $\mathbf{g}_{\gamma(i)}$; tais modelos são conhecidos também como expansões ou decomposições atômicas [11, 16]. É importante ressaltar que os modelos atômicos permitem o refinamento sucessivo de sinais. Os algoritmos de modelagem atômica podem ser desenvolvidos de tal forma que a aproximação do sinal é aperfeiçoada progressivamente à medida que o número de elementos na decomposição aumenta. Esses aumentos de qualidade na representação do sinal podem ser medidos através de uma métrica como o erro quadrático médio, por exemplo.

1.2 Representações Compactas

Um modelo deve ser capaz de fornecer uma descrição do sinal que seja útil à aplicação em vista. Geralmente, é difícil avaliar se um modelo é adequado à aplicação sem um conhecimento *a priori* das características do sinal. Contudo, considerando um modelo acurado e que seja apropriado à aplicação, é possível avaliá-lo com base em uma métrica associada à compactidade da representação que o modelo fornece. Se a representação for acurada e compacta simultaneamente, pode-se concluir que a re-

apresentação captura as estruturas intrínsecas mais significativas do comportamento do sinal. Quanto mais compacta é a representação, mais padrões coerentes com o sinal ela consegue extrair [18, 22]. A coerência corresponde ao grau de similaridade entre os elementos da representação e os fenômenos representados no sinal. A obtenção de representações compactas e coerentes com os sinais é uma das principais questões tratadas neste trabalho.

1.3 Decomposições Adaptativas Redundantes

A grande variedade de padrões presentes em sinais complexos motiva o uso de técnicas de decomposição adaptativa redundante, as quais realizam uma aproximação de uma dada função de forma adaptativa com base em um conjunto de funções pertencentes a um dicionário redundante ou sobrecompleto [16, 18]. A decomposição é considerada adaptativa porque as funções selecionadas no dicionário para representar o sinal dependem deste, e redundante porque o dicionário possui mais funções do que as necessárias para se estabelecer uma base. O resultado da decomposição não é único, no que difere das decomposições usando bases. É interessante destacar que um aumento da redundância do dicionário implica aumento da adaptabilidade permitida e maior compacidade na representação do sinal obtida pela decomposição.

Em contraposição, as decomposições em uma base, como ocorre em Fourier ou em *wavelets*, não são muito flexíveis [18]. As funções-base de Fourier não possibilitam uma boa representação de sinais bem localizados no tempo. O mesmo acontece ao representar-se os sinais cujo suporte no eixo tempo-frequência seja diferente do padrão estabelecido pelas funções-base da *wavelet*. Em ambos os casos, não é possível proceder à identificação correta de padrões a partir dos coeficientes da transformada, porque a informação fica diluída por todas as funções-base.

Ao representar componentes de sinais cuja localização no tempo e na frequência varia bastante, as decomposições adaptativas redundantes, apesar de serem não-lineares, trazem maior flexibilidade em relação às lineares que usam uma base fixa. Nesse caso, o sinal deve ser decomposto em funções cujas características, tanto no tempo quanto na frequência, estejam de acordo com os padrões ou fenômenos nele existentes. Tais funções podem ser denominadas átomos ou estruturas do dicionário. A Figura 1.2 ilustra o número de graus de liberdade de uma senoide, de uma *wavelet* e de um átomo modulado por uma janela exponencial. Observa-se que a senoide pode ser ajustada através de dois parâmetros: a amplitude e frequência; já a *wavelet* é definida por três: a amplitude, a escala e deslocamento; e por fim, tem-se o átomo exponencial, o qual apresenta cinco graus de liberdade: a amplitude, a frequência, o fator de decaimento, e os tempos de início e fim do suporte temporal.

Portanto, os impulsos obtêm uma melhor representação quando decompostos em

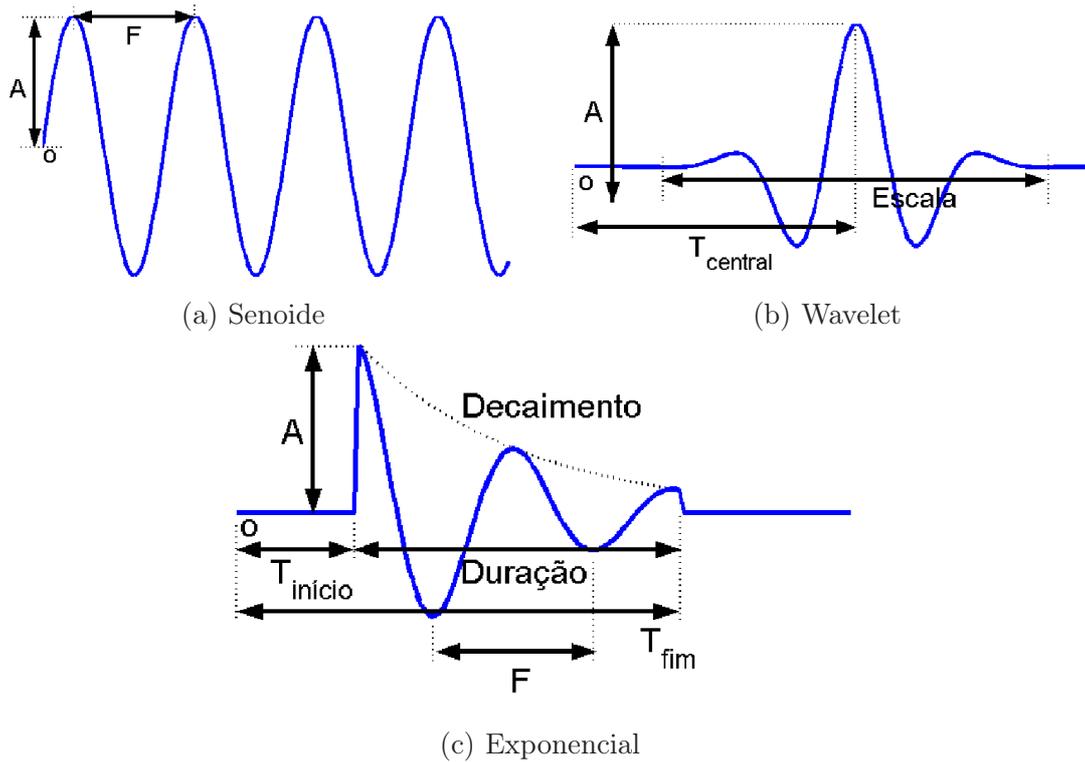


Figura 1.2: Graus de liberdade de uma senoide, de uma *wavelet* e de um átomo exponencial. Essas figuras foram cedidas pelo autor de [1].

funções bem localizadas no tempo, enquanto que as senoides são melhor representadas por funções de banda estreita na frequência. Já os sinais que possuem ambos os elementos são melhor representados por funções de um dicionário redundante. Para isso é necessário definir procedimentos para a seleção dessas funções. Um método bastante popular de decomposição adaptativa redundante, adotado neste trabalho, é o *Matching Pursuit* [18].

1.4 Proposta da Tese

Este trabalho tem como objetivo propor um arcabouço de compressão de sinais usando decomposições atômicas com base em dicionários redundantes. Em geral, utiliza-se um dicionário fixo na etapa de decomposição e codificam-se os coeficientes e os índices dos átomos [23, 24]. Nesta tese, utiliza-se um novo conceito de compressão de sinais usando decomposições atômicas com base em múltiplos dicionários. Além de transmitir os coeficientes e os índices dos átomos, o codificador precisa enviar ao decodificador informação adicional a respeito do dicionário utilizado. Uma forma prática de comprimir com inúmeros dicionários sem tornar o processo computacionalmente inviável consiste em realizar a decomposição atômica através do algoritmo de *Matching Pursuit* (MP) usando dicionários parametrizados redundan-

tes cujos átomos são definidos por parâmetros contínuos; e quantizar, em seguida, os coeficientes e parâmetros dos átomos. O MP realiza aproximações sucessivas de sinais, buscando em um dicionário redundante a melhor aproximação a cada iteração. Dentro deste arcabouço de compressão, desenvolve-se uma técnica de otimização taxa-distorção capaz de selecionar o dicionário de compromisso taxa-distorção ótimo, que se baseia na obtenção de curvas de taxa-distorção operacionais.

A partir do método de decomposição desenvolvido em [25] (MPSA—*Matching Pursuit* por Senoides Amortecidas), constrói-se um sistema de compressão aplicado a sinais elétricos, introduzindo-se o esquema de otimização taxa-distorção através de curvas operacionais, que fornece o dicionário correspondente à menor distorção dada uma taxa desejada. O desempenho do método de compressão de sinais elétricos é avaliado através procedimentos de análise de faltas, utilizando-se uma ferramenta de análise automática de faltas no lugar de figuras ou métricas normalmente adotadas, como o erro quadrático médio. Vale destacar que esta ferramenta foi desenvolvida pelo CEPEL (Centro de Pesquisas de Energia Elétrica) [26, 27].

Desenvolve-se um método de decomposição, inspirado no MPSA, aplicado a sinais de áudio, denominado MPSACr, que se baseia em um dicionário de senoides amortecidas com frequências da escala musical cromática. Como o MPSACr provê a representação do sinal no domínio de tempo, o sinal é segmentado em quadros sem sobreposição, produzindo descontinuidades em átomos presentes em quadros consecutivos. Para evitar estes artefatos, propõe-se e desenvolve-se uma heurística de continuidade dos átomos presentes em quadros consecutivos. Além disso, o MPSACr traz consigo uma nova estratégia de redução de complexidade computacional no cálculo dos produtos internos, de maneira que não há restrição para o espaçamento frequencial no projeto do dicionário. Isto se adequa perfeitamente ao dicionário de escala musical cromática, que apresenta frequências geometricamente espaçadas.

A partir do MPSACr, constrói-se um sistema de compressão aplicado a sinais de áudio, utilizando-se o esquema de otimização taxa-distorção através de curvas operacionais. Nesse caso, as taxas ótimas de bits são aplicadas a todos os coeficientes e parâmetros dos átomos pertencentes à representação do sinal. Dessa maneira, os parâmetros de átomos de baixa energia são quantizados com a mesma intensidade que os parâmetros de átomos de alta energia. Contudo, a influência de um átomo de baixa energia na distorção total é menor do que a influência de um átomo de alta energia. Com base nesta constatação, introduzem-se a fundamentação teórica e a descrição da técnica de otimização taxa-distorção por faixa de energia atômica. Através deste método é possível alocar bits de modo a privilegiar a representação de átomos de maior energia em detrimento dos de menor energia.

Aprimora-se o sistema de compressão aplicado a sinais de áudio, introduzindo um novo esquema de alocação ótima de bits por faixa de energia em substituição

à técnica que se baseia em curvas operacionais. Além disso, utiliza-se um codificador aritmético de modelo fixo, modelando o coeficiente e os parâmetros do átomo exponencial através de gaussianas generalizadas. Por fim, avalia-se o desempenho do sistema de compressão através do PEAQ (*Perceptual Evaluation of Audio Quality*) [28–31], que é um método de medição objetiva de qualidade de sinais de áudio padronizado pela ITU (*International Telecommunication Union*).

1.5 Organização

No Capítulo 2, são abordados inicialmente os aspectos gerais a respeito de decomposições atômicas: fundamentos, exemplos de aplicações, vantagens de utilizá-las junto a dicionários redundantes em vez de decomposições usando base. Em seguida, são apresentados os fundamentos da compressão de sinais usando decomposições atômicas.

No Capítulo 3, apresenta-se o método de decomposição aplicado a sinais elétricos que fornece representações de sinais com base no modelo de senoides amortecidas e é denominado *Matching Pursuit* por Senoides Amortecidas (MPSA). São apresentadas diferentes aplicações da decomposição de sinais elétricos através de MPSA.

No Capítulo 4, apresenta-se o sistema de compressão aplicado a sinais elétricos proposto e desenvolvido nesta tese. Este sistema se baseia no método de decomposição atômica com base no dicionário de senoides amortecidas (MPSA) e realiza a alocação de bits através de curvas operacionais. Avalia-se o desempenho desse sistema com o auxílio de um sistema de análise automática de faltas.

No Capítulo 5, apresenta-se um novo método de decomposição atômica aplicado a sinais de áudio, o MPSACr, fundamentado no MPSA. Ele se baseia em um dicionário composto por senoides amortecidas com as frequências da escala musical cromática. Nesse método, introduz-se uma nova estratégia de redução de complexidade computacional de maneira que não haja restrição de espaçamento frequencial uniforme no projeto do dicionário. Além disso, apresenta-se também uma heurística de continuidade de átomos entre quadros consecutivos. Por fim, são apresentadas simulações computacionais com o método de decomposição proposto.

No Capítulo 6, apresenta-se o sistema de compressão aplicado a sinais de áudio proposto e desenvolvido neste trabalho. Este sistema utiliza o método de decomposição MPSACr e um novo paradigma de alocação ótima de bits por faixa de energia atômica. Além disso, utiliza-se codificação aritmética e os parâmetros associados aos átomos são modelados através de gaussianas generalizadas. Por fim, avalia-se o desempenho do sistema de compressão através do PEAQ, que é um método de medição objetiva de qualidade de sinais de áudio padronizado pela ITU (*International Telecommunication Union*).

No Capítulo 7 apresentam-se as conclusões desta tese e as indicações para os trabalhos futuros.

Capítulo 2

Fundamentos Teóricos

Na primeira parte deste capítulo, são abordados inicialmente os aspectos gerais a respeito de decomposições atômicas: fundamentos, exemplos de aplicações, vantagens de utilizá-las junto a dicionários redundantes em vez de decomposições em base. Em seguida, apresenta-se o método de decomposição atômica denominado de *Matching Pursuit* (MP), que realiza aproximações sucessivas de sinais iterativamente com base em dicionários redundantes. Esta técnica corresponde ao núcleo dos métodos de decomposição desenvolvidos neste trabalho. Descreve-se uma forma de se gerar dicionários redundantes parametrizados a partir de funções-protótipo. Por fim, apresentam-se os conceitos relacionados a decomposições coerentes e à métrica de taxa de aproximação que indica quanto do sinal está sendo aproximado a cada iteração do *Matching Pursuit*.

Na segunda parte, são apresentados os fundamentos da compressão de sinais usando decomposições atômicas. Primeiramente, introduz-se a forma mais usual, que consiste na utilização de um dicionário redundante de parâmetros fixos e pré-definidos, quantizando-se somente os coeficientes. Neste caso, o codificador transmite os coeficientes e os índices dos átomos ao decodificador. Em seguida, introduz-se o novo conceito de compressão usando decomposições atômicas com múltiplos dicionários. Além de transmitir os coeficientes e os índices dos átomos, o codificador precisa enviar ao decodificador informação adicional a respeito do dicionário utilizado. Apresenta-se uma forma prática de lidar com inúmeros dicionários sem tornar a decomposição inviável computacionalmente, utilizando-se MP com dicionários parametrizados. Por fim, descreve-se um modo de selecionar o dicionário de compromisso taxa-distorção ótimo dentro dessa nova abordagem.

2.1 Decomposições Atômicas

As decomposições atômicas têm como objetivo selecionar um subconjunto de elementos, denominados átomos ou estruturas, a partir de um dicionário de formas

de onda pré-definidas, a fim de aproximar um sinal como combinação linear desses elementos. Considerando o dicionário D e o átomo \mathbf{g}_γ , um sinal \mathbf{x} é aproximado na forma

$$\mathbf{x} \approx \hat{\mathbf{x}} = \sum_{i=1}^M \alpha_i \mathbf{g}_{\gamma(i)}, \quad \mathbf{g}_{\gamma(i)} \in D. \quad (2.1)$$

Os átomos $\mathbf{g}_{\gamma(i)}$ são indexados por $\gamma(i)$, que é definido como $\gamma : \mathbb{Z}^+ \rightarrow \{1, \dots, \#_D\}$; $\#_D$ é a cardinalidade—o número de elementos—do dicionário D , portanto $\gamma(i) \in \{1, \dots, \#_D\}$. O parâmetro α_i é o coeficiente que pondera $\mathbf{g}_{\gamma(i)}$ e M corresponde ao número de átomos selecionados para representar \mathbf{x} . A Figura 2.1 ilustra um esquema de análise e síntese utilizando decomposições atômicas. Na análise do sinal, obtêm-se os coeficientes e os índices dos átomos com base em um dicionário D , enquanto que na síntese o sinal é reconstruído através da equação (2.1) tendo em vista o mesmo dicionário D .

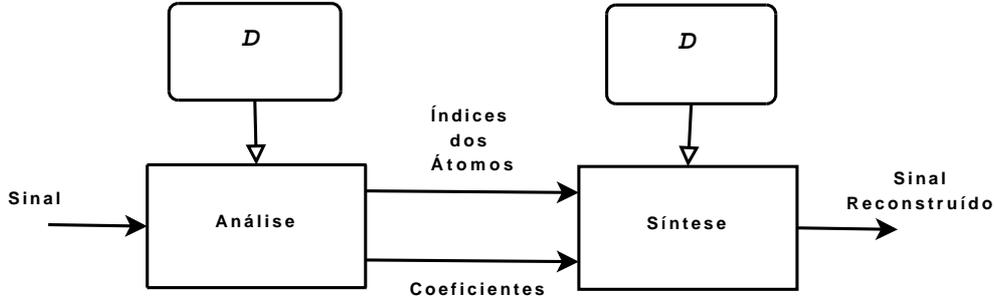


Figura 2.1: Análise e síntese de um sinal utilizando decomposições atômicas com base em um dicionário D .

Existem diversas aplicações relacionadas às decomposições atômicas. Estas têm sido empregadas na filtragem de sinais e supressão de ruído [8, 16], análise de fenômenos físicos intrínsecos dos sinais em conjunto com reconhecimento padrões e modelagem matemática de sinais [7, 14, 16, 32–35], análise tempo-frequência [16, 18] e análise harmônica [7, 36]. Elas atuam também como ferramentas de compressão de sinais [23, 36–39], dado que são capazes de gerar representações compactas, característica fundamental para o bom desempenho de um codificador. Além disso, as decomposições atômicas podem ser utilizadas na separação de realizações de diferentes processos gaussianos [40].

A distorção entre a aproximação de M termos (equação (2.1)) e o sinal \mathbf{x} é descrita por

$$d(\mathbf{x}, M, D) = \|\mathbf{x} - \hat{\mathbf{x}}\| = \left\| \mathbf{x} - \sum_{i=1}^M \alpha_i \mathbf{g}_{\gamma(i)} \right\|. \quad (2.2)$$

Esta distorção depende do (i) número de elementos M usados para representar $\hat{\mathbf{x}}$; (ii) dos átomos $\mathbf{g}_{\gamma(i)}$ escolhidos para exprimir $\hat{\mathbf{x}}$; (iii) e dos coeficientes α_i dos átomos. A representação mais compacta de \mathbf{x} é aquela que utiliza o menor número

de átomos com distorção nula [14]. Contudo, na prática, busca-se a representação mais eficiente, ou seja, aquela que utiliza o menor número de átomos que aproxime \mathbf{x} com a menor distorção aceitável de acordo com a aplicação [4, 41].

Em decomposições atômicas, permite-se o uso de dicionários compostos por elementos linearmente dependentes, o que as distingue de metodologias clássicas baseadas em transformadas, nas quais se utiliza uma base que provê uma representação única do sinal. Como exemplos de bases ortonormais, têm-se a base de *Fourier* e as *wavelets* discretas [16]. A dependência linear do dicionário está associada à existência de mais elementos no dicionário do que o necessário para se gerar um espaço vetorial, e portanto, há mais de uma representação possível do sinal. Nesse caso, o dicionário é denominado redundante ou sobrecompleto. Para que um dicionário D seja capaz de representar qualquer sinal $\mathbf{x} \in \mathbb{X}$ com uma dada distorção $d(\mathbf{x}, M, D)$, D deve ser completo em \mathbb{X} . Isto quer dizer que existe pelo menos uma combinação linear de elementos de D que resulta em $\mathbf{x} = \hat{\mathbf{x}}$, para todo $\mathbf{x} \in \mathbb{X}$, ou seja, D gera \mathbb{X} . Os dicionários redundantes podem ser formados por uma união de bases ortogonais, por transformadas redundantes (como *wavelet packet* e *frames* de Gabor) [16] ou por formas de onda parametrizadas [4, 14, 41, 42].

A redundância resulta em poder representar o sinal de diferentes maneiras utilizando os mesmos elementos do dicionário. Logo, o dicionário redundante corresponde a uma base linearmente dependente. A utilização de dicionários altamente redundantes possibilita a extração direta de uma variedade maior de padrões e fenômenos presentes em sinais, resultando em representações mais compactas e eficientes. Entretanto, a compacidade da representação de sinais não depende somente do nível de redundância, mas também da acurácia do modelo matemático no qual se baseia o projeto do dicionário em representar de forma coerente os fenômenos intrínsecos de uma determinada classe de sinais: os átomos do dicionário devem apresentar um alto grau de similaridade com os padrões existentes da classe. Em processos físicos, considera-se o sinal observado como uma mistura de componentes \mathbf{p}_i , que representam os fenômenos físicos, dada por

$$\mathbf{x} = \sum_i \beta_i \mathbf{p}_i + \mathbf{n}, \quad (2.3)$$

em que \mathbf{n} é o ruído inerente à observação. Quanto mais parecidos forem os átomos $\mathbf{g}_{\gamma(i)}$ e respectivos coeficientes α_i utilizados na representação de um sinal \mathbf{x} com as componentes \mathbf{p}_i e respectivos coeficientes β_i efetivamente presentes em \mathbf{x} , melhor será a representação obtida para modelagem do sinal e reconhecimento de padrões. Portanto, conclui-se que a redundância e a coerência são fatores fundamentais para se atingir um maior nível de compacidade na representação de sinais.

De maneira geral, nota-se que através das decomposições atômicas é possível mo-

delar sinais de forma matematicamente acurada, compacta e coerente, com baixa distorção. Um algoritmo bastante popular utilizado na obtenção dessas decomposições é o *Matching Pursuit* (MP) [16, 18, 43].

2.1.1 *Matching Pursuit*

O algoritmo de *Matching Pursuit* (MP) realiza aproximações sucessivas de sinais, de forma adaptativa, buscando em um dicionário redundante a melhor aproximação a cada iteração [2, 4, 18].

Considere um sinal \mathbf{x} pertencente ao espaço de Hilbert \mathbb{H} de dimensão finita. Define-se um dicionário redundante $D = (\mathbf{g}_\gamma)_{\gamma \in \Gamma}$, tal que $\|\mathbf{g}_\gamma\| = 1$, em que γ é o índice que define \mathbf{g} e Γ é o conjunto de todos os γ 's possíveis. O MP gera uma expansão linear de \mathbf{x} em função de um subconjunto de elementos de D . Isso é feito através de aproximações sucessivas de \mathbf{x} com projeções ortogonais sobre os elementos do dicionário. O sinal \mathbf{x} pode ser decomposto da seguinte maneira

$$\mathbf{x} = \langle \mathbf{x}, \mathbf{g}_{\gamma_0} \rangle \mathbf{g}_{\gamma_0} + \mathbf{r}_x, \quad (2.4)$$

em que \mathbf{r}_x é o resíduo da projeção de \mathbf{x} na direção de \mathbf{g}_{γ_0} e o operador \langle, \rangle representa o produto interno. Na Figura 2.2, observa-se que \mathbf{g}_{γ_0} é ortogonal a \mathbf{r}_x , logo tem-se:

$$\|\mathbf{x}\|^2 = |\langle \mathbf{x}, \mathbf{g}_{\gamma_0} \rangle|^2 + \|\mathbf{r}_x\|^2. \quad (2.5)$$

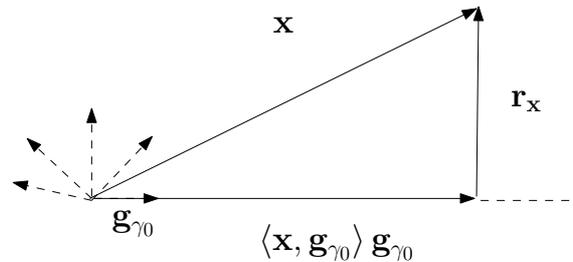


Figura 2.2: Representação gráfica da projeção ortogonal no primeiro passo da decomposição de \mathbf{x} .

A cada passo da decomposição, com o intuito de minimizar $\|\mathbf{r}_x\|$, escolhe-se o elemento \mathbf{g}_{γ_0} do dicionário tal que $|\langle \mathbf{x}, \mathbf{g}_{\gamma_0} \rangle|$ seja máximo. Visto que Γ é finito, obtém-se um \mathbf{g}_{γ_0} ótimo no seguinte sentido:

$$|\langle \mathbf{x}, \mathbf{g}_{\gamma_0} \rangle| \geq \max_{\gamma \in \Gamma, \gamma \neq \gamma_0} |\langle \mathbf{x}, \mathbf{g}_\gamma \rangle|. \quad (2.6)$$

Em seguida, o resíduo \mathbf{r}_x é decomposto pela sua projeção no elemento do dicionário que melhor o representa, como foi feito anteriormente. Essa projeção irá gerar

um segundo resíduo, \mathbf{r}_x^2 , o qual sofrerá o mesmo procedimento, gerando um terceiro resíduo, e assim por diante. Dessa forma, obtém-se a seguinte formulação para se calcular o resíduo de ordem n :

$$\mathbf{r}_x^n = \langle \mathbf{r}_x^n, \mathbf{g}_{\gamma_n} \rangle \mathbf{g}_{\gamma_n} + \mathbf{r}_x^{n+1}, \quad (2.7)$$

em que $\mathbf{r}_x^0 = \mathbf{x}$ e define-se o resíduo de ordem $n + 1$, \mathbf{r}_x^{n+1} .

Ainda que seja não-linear, o MP é caracterizado pela conservação de energia, característica intrínseca de decomposições lineares e ortogonais. Para fins de demonstração, dado que \mathbf{r}_x^{n+1} é ortogonal a \mathbf{g}_{γ_n} , tem-se

$$\|\mathbf{r}_x^n\|^2 = |\langle \mathbf{r}_x^n, \mathbf{g}_{\gamma_n} \rangle|^2 + \|\mathbf{r}_x^{n+1}\|^2. \quad (2.8)$$

Em seguida, escreve-se $\|\mathbf{x}\|^2$ da seguinte forma:

$$\|\mathbf{x}\|^2 = \sum_{n=0}^{m-1} (\|\mathbf{r}_x^n\|^2 - \|\mathbf{r}_x^{n+1}\|^2) + \|\mathbf{r}_x^m\|^2. \quad (2.9)$$

Combinando as equações (2.7) e (2.9), obtém-se a equação de conservação de energia

$$\|\mathbf{x}\|^2 = \sum_{n=0}^{m-1} |\langle \mathbf{r}_x^n, \mathbf{g}_{\gamma_n} \rangle|^2 + \|\mathbf{r}_x^m\|^2. \quad (2.10)$$

Decompondo \mathbf{x} até o resíduo de ordem m obtém-se o seguinte somatório:

$$\mathbf{x} = \sum_{n=0}^{m-1} (\mathbf{r}_x^n - \mathbf{r}_x^{n+1}) + \mathbf{r}_x^m. \quad (2.11)$$

Com base na representação da equação (2.7), tem-se:

$$\mathbf{x} = \sum_{n=0}^{m-1} \langle \mathbf{r}_x^n, \mathbf{g}_{\gamma_n} \rangle \mathbf{g}_{\gamma_n} + \mathbf{r}_x^m. \quad (2.12)$$

Em [18], apresenta-se um teorema que garante a convergência deste método de decomposição. Se o dicionário for completo, $\lim_{n \rightarrow \infty} \|\mathbf{r}_x^n\| = 0$. Na prática, esse processo iterativo termina quando o resíduo $\|\mathbf{r}_x^m\|$ for suficientemente reduzido. Ao fim da decomposição, obtém-se uma representação para \mathbf{x} como uma soma dos elementos do dicionário que melhor representam os resíduos; logo, tem-se que

$$\mathbf{x} \approx \sum_{n=0}^{M-1} \alpha_n \mathbf{g}_{\gamma_n}, \quad (2.13)$$

em que M corresponde ao número de passos do MP e, também, ao número de átomos do dicionário D cuja combinação linear aproxima \mathbf{x} , e $\alpha_n = \langle \mathbf{r}_{\mathbf{x}}^n, \mathbf{g}_{\gamma_n} \rangle$. Portanto, \mathbf{x} pode ser aproximado pela sequência $(\alpha_n, \gamma_n)_{n \in N}$, a qual pode ser denominada livro de estruturas (*structure book* [18]).

2.1.2 Dicionários Redundantes Parametrizados

Os átomos de um dicionário redundante são gerados escalonando-se, trasladando-se e modulando-se uma função $h(t)$. Supondo $h(t)$ real e centralizada em 0, o átomo complexo $g_\gamma(t)$ é definido por [16, 18]

$$g_\gamma(t) = \frac{1}{\sqrt{s}} h\left(\frac{t-u}{s}\right) e^{i\xi t}, \quad (2.14)$$

em que o fator $\frac{1}{\sqrt{s}}$ torna unitária a energia de g_γ . O índice γ é composto pelos parâmetros de escalonamento s , frequência de modulação ξ e deslocamento no tempo u , ou seja, $\gamma = (s, u, \xi)$, onde $\gamma \in \Gamma = R^+ \times R^2$ e Γ é o conjunto de todos os parâmetros do dicionário. A função $g_\gamma(t)$ possui sua energia concentrada em torno do seu centro u . A transformada de Fourier de $\mathbf{g}_\gamma(t)$ é dada por

$$\widehat{g}_\gamma(\omega) = \sqrt{s} \widehat{h}(s(\omega - \xi)) e^{i(\omega - \xi)u}, \quad (2.15)$$

dado que $|\widehat{h}(\omega)|$ é par, $|\widehat{g}_\gamma(\omega)|$ possui sua energia concentrada em torno da frequência central $\omega = \xi$. A redundância deste tipo dicionário é comprovada pelo fato de que quando $\gamma = (s_0, u, \xi)$, $g_\gamma(t)$ se reduz às funções-base da STFT (*Short-Time Fourier Transform*) [16]. Da mesma forma, quando $\gamma = (s, u, \frac{\xi_0}{s})$ obtêm-se as funções-base das *Wavelets* [16].

No caso de sinais reais, utiliza-se um dicionário de átomos reais gerados da seguinte forma:

$$g_{(\gamma, \phi)}(t) = \frac{K_{(\gamma, \phi)}}{\sqrt{s}} h\left(\frac{t-u}{s}\right) \cos(\xi t + \phi), \quad (2.16)$$

em que $K_{(\gamma, \phi)}$ é um fator de normalização para que $\|g_{(\gamma, \phi)}\| = 1$, e a fase ϕ pertence ao intervalo $[0, 2\pi)$.

Em [16, 18], mostra-se que a localização no tempo de $g_\gamma(t)$ é mensurada através da variância de seu módulo ao quadrado, em torno de u , e sua localização na frequência, através da variância da transformada de Fourier de seu módulo ao quadrado, em torno de ξ . Supondo-se $\|h(t)\| = \|\widehat{h}(\omega)\| = 1$, tem-se que

$$\sigma_{t,s}^2 = \int_{-\infty}^{+\infty} (t-u)^2 |g_\gamma(t)|^2 dt = s^2 \int_{-\infty}^{+\infty} t^2 |h(t)|^2 dt \quad (2.17)$$

e

$$\sigma_{\omega,s}^2 = \int_{-\infty}^{+\infty} (\omega - \xi)^2 |\hat{g}_\gamma(\omega)|^2 d\omega = \frac{1}{s^2} \int_{-\infty}^{+\infty} \omega^2 |\hat{h}(\omega)|^2 d\omega. \quad (2.18)$$

A Figura 2.3 ilustra que a resolução tempo-frequência de g_γ é representada no plano tempo-frequência por uma caixa de Heisenberg centralizada em (u, ξ) cujas larguras no tempo e na frequência são $\sigma_{t,s}$ e $\sigma_{\omega,s}$, respectivamente [16]. Pelo princípio da incerteza de Heisenberg, prova-se que a área do retângulo é no mínimo $\frac{1}{2}$:

$$\sigma_{t,s}^2 \sigma_{\omega,s}^2 \geq \frac{1}{4} \Rightarrow \sigma_{t,s} \sigma_{\omega,s} \geq \frac{1}{2}. \quad (2.19)$$

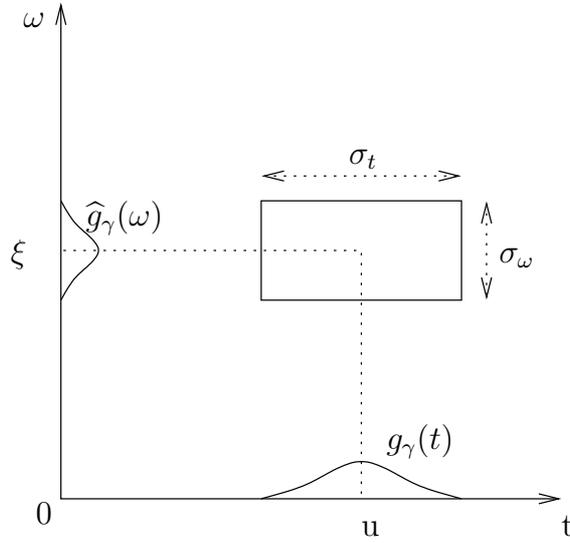


Figura 2.3: Representação em tempo-frequência do átomo g_γ .

Esse valor mínimo é alcançado se e somente se $h(t)$ é uma gaussiana [16], e, nesse caso, pode-se denominá-lo uma função de Gabor. Em [18], define-se uma distribuição de energia no plano tempo-frequência de um dado sinal a partir da distribuição de Wigner-Ville [16] dos elementos do dicionário nos quais ele foi decomposto pelo algoritmo de MP. Logo, a densidade de energia é dada por

$$Ef(t, \omega) = \sum_{n=0}^{\infty} | \langle r_x^n, g_{\gamma_n} \rangle |^2 W_{g_{\gamma_n}}(t, \omega), \quad (2.20)$$

em que $W_{g_{\gamma_n}}(t, \omega)$ corresponde à distribuição de Wigner-Ville do átomo de índice $\gamma_n = (s_n, u_n, \xi_n)$ do dicionário. Na Figura 2.4, observam-se as representações no plano tempo-frequência dos átomos de Gabor, os quais representam o sinal em (a), para diferentes tipos de distribuição. Em (b), realizou-se a transformada de Fourier de tempo curto dos átomos, resultando em um espectrograma. Em (c), apresenta-se a distribuição de Choi-Williams [16]. Por fim, em (d), utilizou-se a distribuição de

Wigner-Ville.

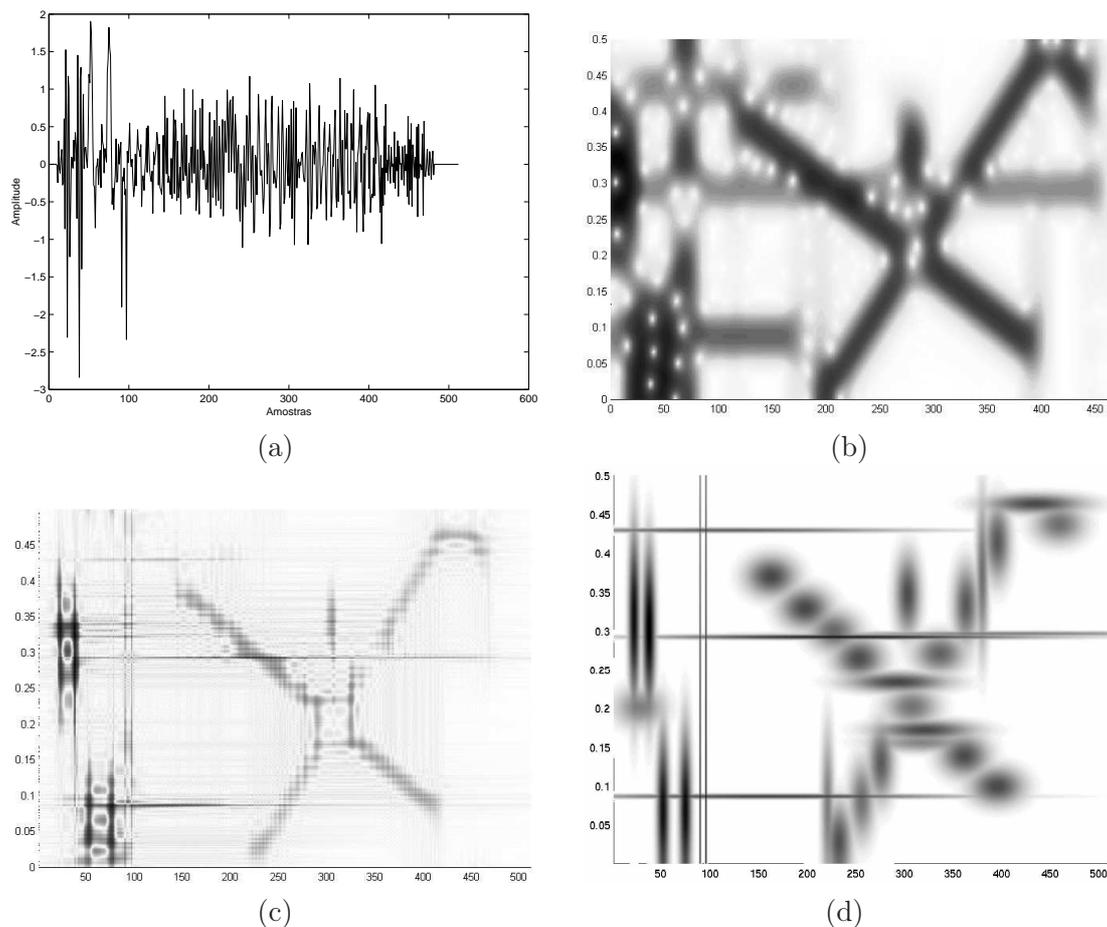


Figura 2.4: Representações no plano tempo-frequência do sinal (a) utilizando a STFT (b), a distribuição de Choi-Williams (c), e a distribuição de Wigner-Ville (d). Estas figuras foram cedidas pelo autor de [2].

Obtém-se uma boa estimativa do conteúdo tempo-frequência do sinal ao se utilizar a distribuição de Wigner-Ville dos átomos de Gabor resultantes da decomposição do mesmo, pois o átomo gaussiano possui a melhor concentração no plano tempo-frequência, e a densidade de energia é calculada individualmente para cada átomo [18]. Contudo, isso não significa que a decomposição com base em um dicionário de Gabor resulte na representação mais coerente possível de quaisquer sinais. Para isso, deve-se utilizar um dicionário formado por elementos que sejam coerentes com os fenômenos ou padrões existentes no sinal a ser decomposto.

2.1.3 Decomposições Coerentes

Uma decomposição coerente pressupõe alto grau de similaridade do sinal com relação ao dicionário, ou seja, este deve ser formado por elementos representativos dos fenômenos ou padrões existentes naquele. Se isto ocorrer até um certo número de passos do MP, obtêm-se os átomos que são coerentes com as estruturas do sinal. Contudo,

não se sabe a partir de qual passo se perde esta coerência, ou seja, o resíduo passa a corresponder a ruído. Portanto, faz-se necessário utilizar alguma métrica com o propósito de discernir os elementos coerentes daqueles correspondentes a processos ruidosos, obtidos durante a decomposição. Uma forma de se medir a distância entre o resíduo e o elemento escolhido a cada passo se dá através do cálculo da taxa de aproximação, ou seja, o cosseno do ângulo entre eles [18]. Se r_x^n é o resíduo no passo n e g_{γ_n} , o átomo escolhido, define-se a taxa de aproximação da seguinte maneira:

$$\lambda(r_x^n) = \frac{\| \langle r_x^n, g_{\gamma_n} \rangle \|}{\| r_x^n \|}. \quad (2.21)$$

Em [18], os Lemas 1 e 2 provam que o resíduo decresce exponencialmente em função dos passos de decomposição em uma espaço de dimensão finita. A Figura 2.5 apresenta o resultado de um experimento numérico que mostra o comportamento exponencial do resíduo. Observa-se que nos primeiros passos de decomposição, onde se obtêm os elementos coerentes, a taxa de decaimento é alta. Essa taxa de decaimento diminui à medida que a decomposição evolui, até que o resíduo converge para um valor médio estacionário. A partir desse ponto, considera-se que o resíduo é igualmente correlacionado com todos os elementos do dicionário, comportando-se como um ruído branco estacionário [22]. Os resíduos aproximados nesse contexto correspondem ao ruído do dicionário.

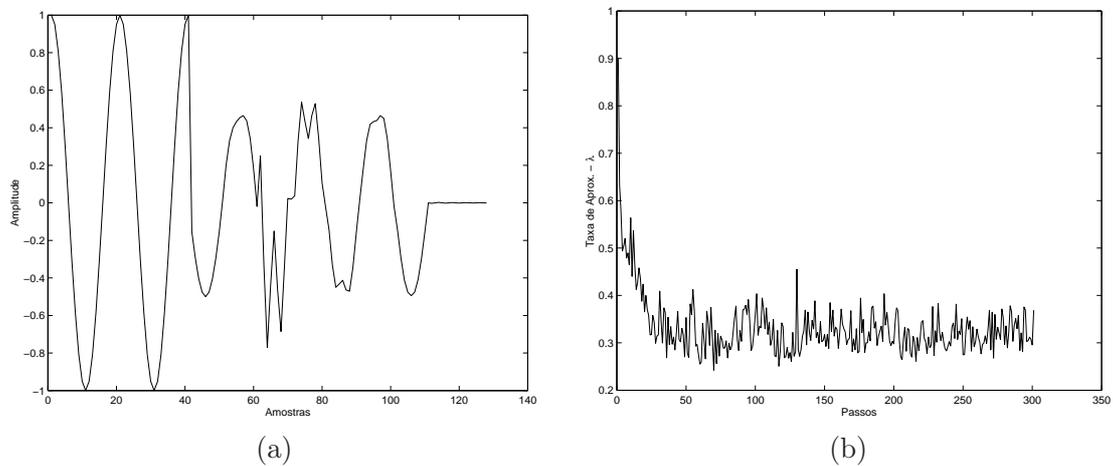


Figura 2.5: (a) sinal decomposto e (b) taxa de aproximação nos 300 primeiros passos da decomposição utilizando um dicionário de exponenciais.

É importante destacar que o ruído depende do dicionário em questão [22]. Dessa maneira, um resíduo caracterizado como ruído para um determinado dicionário pode ainda conter elementos coerentes com relação a outro dicionário. Por exemplo, um sinal senoidal não possui padrões coerentes com um dicionário de impulsos de Dirac; no entanto, é bastante coerente com um dicionário de exponenciais complexas.

A escolha do dicionário é muito importante para se alcançar um bom desempenho

no MP. Para isso, o dicionário deve ser constituído de elementos coerentes com o sinal, permitindo, assim, uma análise fidedigna dos fenômenos presentes.

2.2 Compressão usando Decomposições Atômicas

Em geral, a compressão por decomposições atômicas consiste em se obter, primeiramente, a aproximação do sinal \mathbf{x} como descrito na equação (2.1). Em seguida, os coeficientes e índices dos átomos são codificados e transmitidos ou armazenados. No decodificador, a fim de reconstruir o sinal, emprega-se o mesmo dicionário do codificador. Tal esquema está ilustrado na Figura 2.6. Nota-se que não existe a necessidade de se codificar qualquer informação referente ao dicionário utilizado, dado que o decodificador já o conhece de antemão. Neste caso, o melhor custo-benefício na relação taxa-distorção é obtido com base no compromisso entre o número de elementos na representação do sinal e o nível de quantização dos coeficientes [24].

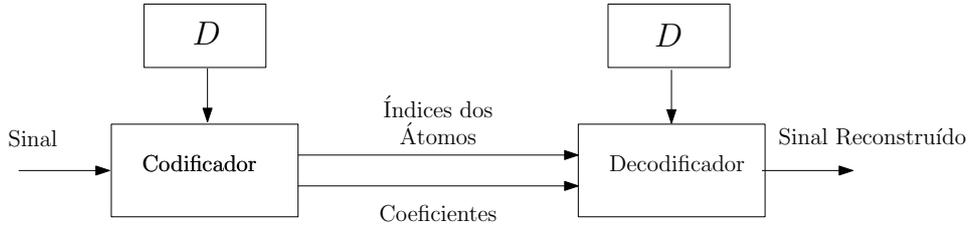


Figura 2.6: Compressão por decomposições atômicas usando um único dicionário redundante D .

2.2.1 Compressão usando Decomposições Atômicas com Múltiplos Dicionários

Agora, suponha que em vez de um único dicionário, seja considerado um conjunto de dicionários redundantes descrito por $\mathcal{D} = \{D_i\}_{i=1,\dots,I}$, em que I é o número de dicionários pertencentes a \mathcal{D} . Esta abordagem é ilustrada na Figura 2.7. Seja cada dicionário $D_i = \{\mathbf{g}_{\gamma(k)}^{D_i}\}_{k=1,\dots,\#D_i}$, onde $\mathbf{g}_{\gamma(k)}^{D_i}$ representa o átomo $\in D_i$ e $\#D_i$, a cardinalidade de D_i . Neste esquema, o codificador escolhe o dicionário a ser utilizado na decomposição atômica. Considerando D_i como o dicionário escolhido, a decomposição atômica fornece a aproximação do sinal

$$\hat{\mathbf{x}}_{D_i} = \sum_{n=1}^{M_{D_i}} \alpha_n^{D_i} \mathbf{g}_{\gamma(n)}^{D_i}. \quad (2.22)$$

em que M_{D_i} é o número de átomos selecionados a partir do dicionário D_i para compor a representação do sinal $\hat{\mathbf{x}}_{D_i}$. Cada dicionário resulta em diferentes representações do sinal. Logo, codificam-se os coeficientes, os índices dos átomos e a

informação adicional especificando o dicionário utilizado para a decomposição do sinal. O decodificador, com base na informação adicional, seleciona em \mathcal{D} o dicionário usado pelo codificador e reconstrói do sinal. O desempenho ótimo em termos de taxa-distorção é alcançado quando obtém-se o compromisso entre o número de bits gastos em informação adicional, em índices dos átomos e coeficientes, que implica distorção mínima. Observa-se que a otimização deste problema apresenta alta demanda computacional e é de difícil resolução, posto que é necessário lidar com um enorme conjunto de dicionários, impossibilitando o uso desta abordagem de decomposição. Esta alta demanda computacional pode ser reduzida através do uso de dicionários parametrizados, como será visto a seguir.

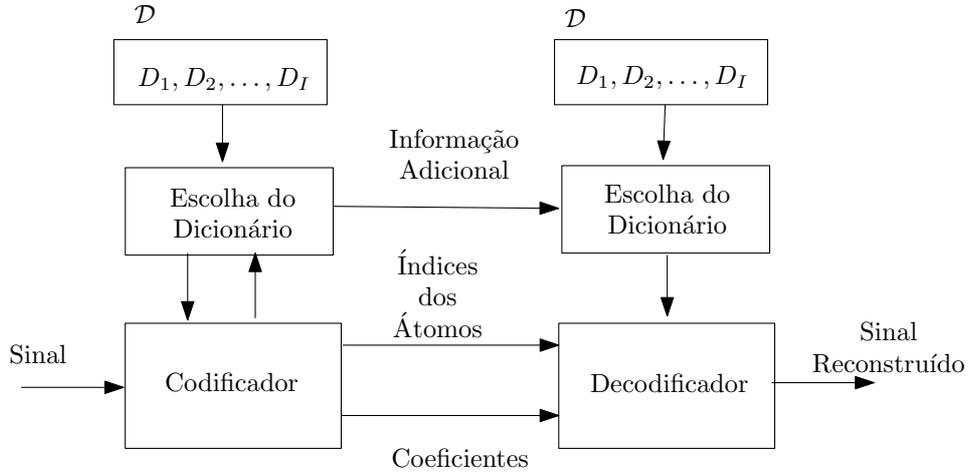


Figura 2.7: Compressão por decomposições atômicas usando um conjunto \mathcal{D} de dicionário redundantes.

2.2.2 Compressão usando Decomposições Atômicas com Dicionários Parametrizados

Os dicionários parametrizados são definidos como um conjunto $D = \{\mathbf{g}_{\gamma(k)}\}_{\gamma(k) \in \Gamma}$, tal que $\|\mathbf{g}_{\gamma(k)}\| = 1$, em que Γ é o espaço de parâmetros—o conjunto dos $\gamma(k)$ permitidos, e o índice $\gamma(k)$ é definido como $\gamma : \mathbb{Z}^+ \rightarrow \mathbb{R}^P$. O conjunto de parâmetros $\gamma(k) = (\nu_p(k))_{p=1, \dots, P}$, onde $\nu_p(k)$ representa o valor do parâmetro e P é a quantidade de parâmetros que definem $\mathbf{g}_{\gamma(k)}$. Como visto na seção 2.1.2, estes dicionários são geralmente formados por átomos reais gerados através do escalonamento, translação e modulação de uma única função janela $g(t)$, resultando em [18]:

$$g_{\gamma}(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) \cos(\xi t + \phi), \quad (2.23)$$

onde s é a escala, ξ é a frequência de modulação, u é a translação no tempo e ϕ é a fase; logo, $\gamma(k) = (s_k, \xi_k, u_k)$.

A princípio, a compressão de decomposições atômicas através de diversos dicionários redundantes parametrizados aparenta ser uma árdua tarefa, visto que o espaço de parâmetros Γ é infinito. Dessa forma, a compressão somente é possível caso seja selecionado um subconjunto finito e contável $\Gamma_{D_i} \in \Gamma$ para codificar o sinal. Cada subconjunto de parâmetros define um dicionário distinto (ver Figura 2.8).

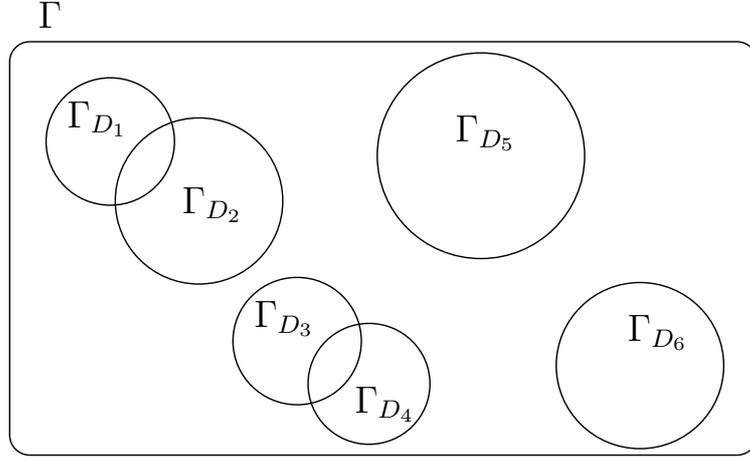


Figura 2.8: Problema da compressão de decomposições atômicas com base em dicionários parametrizados.

Uma abordagem prática para tratar este problema consiste em realizar a decomposição atômica através do *Matching Pursuit* com base em um dicionário redundante parametrizado, cujos átomos são definidos por parâmetros contínuos ou densamente quantizados [4, 41, 42]. Os parâmetros são restringidos somente pela sua representação numérica. Esta abordagem será vista mais adiante com mais detalhes.

A indexação dos diferentes dicionários é realizada através da quantização dos parâmetros dos átomos, conforme ilustra o diagrama de blocos da Figura 2.9. Para que isto seja viável, a cada passo do processo iterativo do MP, busca-se, inicialmente, o átomo mais correlacionado com o resíduo em um dicionário com o espaço de parâmetros amostrado. Em seguida, otimizam-se os parâmetros γ com o propósito de encontrar aqueles que maximizam a aproximação entre o átomo e o resíduo [4]. Ao fim da decomposição, obtém-se o livro de estruturas composto pelos pares $(\alpha_n, \gamma(n))$, $n = 1, \dots, M$. Com o intuito de comprimir o sinal, quantiza-se o livro de estruturas; uma das maneiras de fazer isto é quantizar α_n e cada parâmetro de $\gamma(n)$ utilizando um quantizador escalar uniforme [4, 41] definido por [44]

$$x_q = I_x \times \Delta_{q(x)}, \quad \text{com } I_x = \left\lfloor \frac{x + \frac{\Delta_{q(x)}}{2}}{\Delta_{q(x)}} \right\rfloor, \quad (2.24)$$

em que x é qualquer parâmetro, x_q representa sua versão quantizada, $\Delta_{q(x)}$ é o passo de quantização e I_x corresponde ao índice associado a x . Além disso, define-se $\lfloor z \rfloor = \text{máx}\{m \in \mathbb{Z} | m \leq z\}$. Os parâmetros são quantizados de acordo com um

intervalo dinâmico definido por seus respectivos valores máximo e mínimo dentre todos os elementos do livro de estruturas, ou seja [44],

$$\Delta_{q(x)} = \begin{cases} \frac{x_{\max} - x_{\min}}{2^{b_x} - 1}, & \text{se o quantizador for } \textit{mid-rise} \\ \frac{x_{\max} - x_{\min}}{2^{b_x} - 2}, & \text{se o quantizador for } \textit{mid-thread}, \end{cases}$$

em que b_x é o número de bits alocados para x . Os intervalos dinâmicos e o número de bits alocados aos parâmetros são codificados como informação adicional [4, 41]. Outros tipos de quantizadores podem ser utilizados e avaliados, além do quantizador uniforme. A quantização dos parâmetros e dos coeficientes resulta na seguinte aproximação do sinal

$$\hat{\mathbf{x}} = \sum_{n=1}^M Q^\alpha[\alpha_n] \mathbf{g}_{Q^i[\gamma(n)]}, \quad (2.25)$$

em que $Q^\alpha[\cdot]$ é a função de quantização dos coeficientes e $Q^i[\cdot]$ representa a função de quantização associada ao dicionário D_i . Dependendo, da quantidade de bits alocados entre os parâmetros, diferentes átomos irão compor a aproximação do sinal, ou seja, utilizar-se-ão dicionários distintos. Mais precisamente, o dicionário D_i usado é definido pela função $Q^i[\gamma(n)] : \mathbb{R}^T \rightarrow \mathbb{R}^T$. Dessa maneira, o problema de otimização taxa-distorção consiste em buscar os quantizadores $Q^\alpha[\cdot]$ e $Q^i[\cdot]$ que levam à distorção mínima para uma dada taxa.

Deve-se destacar que o esquema de compressão apresentado difere de outros sistemas de compressão que utilizam o MP. Normalmente, esses sistemas realizam a compressão de sinais utilizando um dicionário fixo, quantizando somente os coeficientes [23, 24, 38]. Aqui, a compressão é realizada quantizando-se os parâmetros do livro de estruturas, e os átomos usados para reconstruir o sinal são diferentes dos átomos obtidos na decomposição. Como foi discutido, este esquema de compressão é equivalente a usar múltiplos dicionários seguidos da escolha de um deles para a codificação do sinal.

2.2.3 Seleção do Dicionário de Compromisso Taxa-Distorção Ótimo

O objetivo da otimização taxa-distorção é obter a melhor reprodução do sinal para uma dada taxa de compressão alvo [45]. No paradigma de compressão em questão, busca-se o compromisso entre o número de átomos da representação do sinal, a quantização de cada coeficiente e a escolha do dicionário $D_i \in \mathcal{D}$ que é definido pelos quantizadores dos parâmetros do átomo. Primeiramente, define-se o número

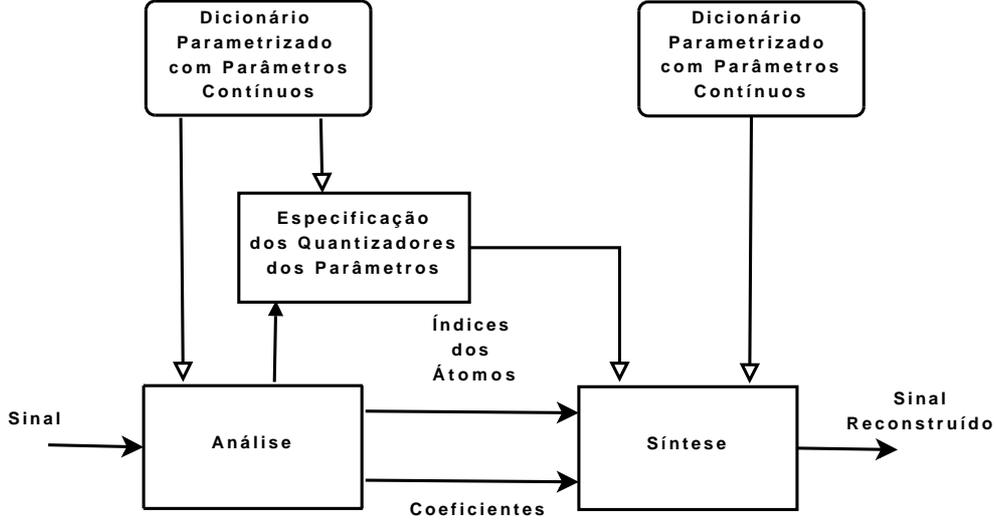


Figura 2.9: Compressão de decomposições atômicas usando um conjunto de dicionários parametrizados.

de bits de um átomo como

$$r = r_\alpha + r_{\nu_1} + \dots + r_{\nu_t} + \dots + r_{\nu_T}, \quad (2.26)$$

em que r_α representa o número de bits alocados aos coeficientes α e r_{ν_t} é a quantidade de bits alocados ao parâmetro ν_t , tal que $\gamma = \{\nu_t\}_{t=1,\dots,T}$, e T é o número parâmetros do átomo. Deve-se também determinar o número M de termos da decomposição. O número total de bits gastos será rM . A distorção total é expressa em função das taxas de bits dos coeficientes e dos parâmetros, resultando em

$$d_S = f(r_\alpha, r_{\nu_1}, \dots, r_{\nu_T}). \quad (2.27)$$

Considere o quantizador uniforme definido por (2.24) e os comprimentos de bits de cada parâmetro na $(T+1)$ -upla $\mathbf{b}_k = (r_\alpha, r_{\nu_1}, \dots, r_{\nu_T}) \in \mathcal{B}$, onde \mathcal{B} representa o conjunto de todas as possíveis combinações permitidas de taxas de bits dentro do intervalo definido por cada elemento \mathbf{b}_k , $k = [1, 2, \dots, K_{\mathcal{B}}]$, sendo $K_{\mathcal{B}}$ o número de elementos em \mathcal{B} . Note que cada \mathbf{b}_k define um quantizador de coeficiente e os quantizadores para os parâmetros do átomo. Os quantizadores dos parâmetros correspondem à escolha de um dicionário D_i (ver equação (2.22)). A fim de obter o melhor compromisso taxa-distorção, deve-se buscar o \mathbf{b}_k que minimiza a distorção total inserida no processo de codificação, dada uma quantidade de bits disponíveis r_{alvo} . A solução é obtida através da resolução do seguinte problema de otimização [45]:

$$\min_{\mathbf{b}_k \in \mathcal{B}} d_S \quad \text{sujeito a } Mr = r_{\text{alvo}}. \quad (2.28)$$

A solução clássica para este problema se baseia em otimização Lagrangeana [45–47], que corresponde a minimizar a função-custo

$$J = d_S + \lambda (M r - r_{\text{alvo}}), \quad (2.29)$$

onde λ é multiplicador Lagrangeano. A idéia básica desta técnica está ilustrada na Figura 2.10. Para um dado λ , encontra-se o par $(d_S^{\text{opt}}, r_{\text{alvo}}^{\text{opt}})$ onde J é mínimo. Os diversos pontos ótimos relativos a diferentes λ 's formam a curva taxa-distorção operacional. A solução do problema de minimização irrestrita da função dada em (2.29) é obtida através da resolução do seguinte sistema de equações:

$$\frac{\partial J}{\partial r_\alpha} = 0 \quad (2.30)$$

$$\frac{\partial J}{\partial r_{\nu_1}} = 0 \quad (2.31)$$

$$\vdots \quad (2.32)$$

$$\frac{\partial J}{\partial r_{\nu_T}} = 0 \quad (2.33)$$

$$\frac{\partial J}{\partial \lambda} = 0. \quad (2.34)$$

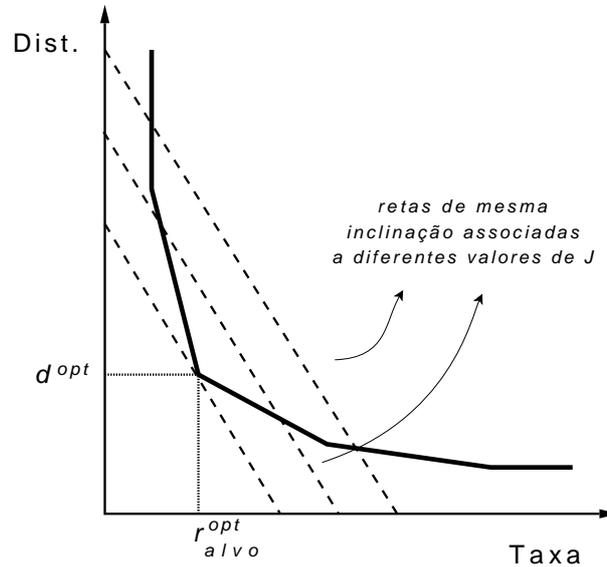


Figura 2.10: Interpretação gráfica da otimização Lagrangeana.

Quando não existir forma fechada para d_S em função das taxas $(r_\alpha, r_{\nu_1}, \dots, r_{\nu_T})$, é possível adotar uma abordagem empírica para obter as curvas operacionais [45]. Para cada $\mathbf{b}_k = (r_\alpha, r_{\nu_1}, \dots, r_{\nu_T})$, e para um dado sinal, calcula-se o par taxa-distorção (r_k, d_k) , resultando em um gráfico taxa-distorção (T-D) como apresentado na Figura 2.11. Nota-se, por exemplo, que o ponto C é certamente pior do que o A ,

porque apresenta a mesma distorção, porém maior taxa. Da mesma forma, o ponto B é também pior do que A , porque apresenta maior distorção para uma mesma taxa. Portanto, seleciona-se o ponto que pertence ao fecho convexo que para uma desejada taxa de compressão fornece o dicionário e o quantizador de coeficientes que levam à distorção mínima.

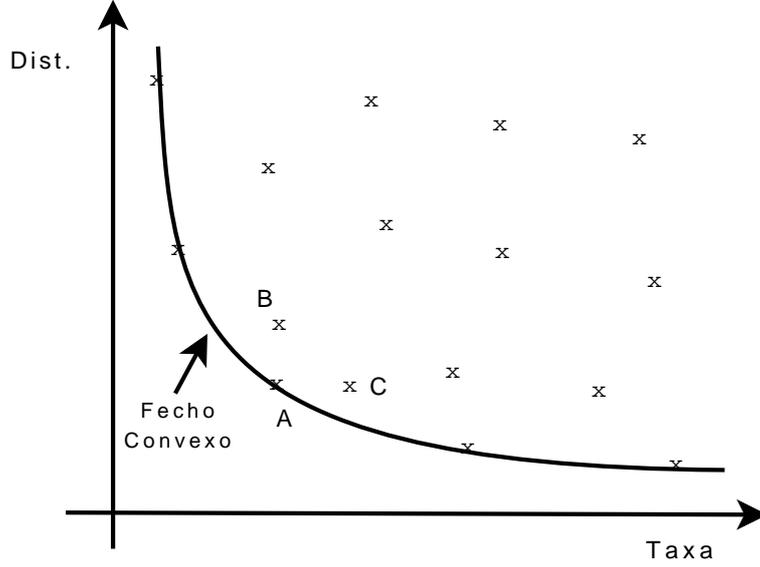


Figura 2.11: Fecho convexo contendo os pontos ótimos em termos de taxa-distorção.

A curva operacional é obtida conectando-se os pontos que pertencem ao fecho convexo da região definida pelos pares T-D gerados para cada $\mathbf{b}_k \in \mathcal{B}$, onde $k \in \mathcal{K} = \{1, 2, \dots, K_B\}$ e K_B é o número de elementos em \mathcal{B} . O procedimento para se obter o fecho convexo, cujo pseudocódigo é descrito no Algoritmo 2.1, transcorre da seguinte forma:

1. Busca-se $\mathbf{b}_{k_o} = \arg \min_{\mathbf{b}_k \in \mathcal{B}} r_k$;
2. Atribui-se a $\mathbf{b}_{k_{atual}} = \mathbf{b}_{k_o}$, portanto $[r_{k_{atual}} ; d_{k_{atual}}] = [r_{k_o} ; d_{k_o}]$;
3. Traça-se uma reta do ponto $[r_{k_{atual}} ; d_{k_{atual}}]$ a todos os outros pontos $[r_k ; d_k]$, onde $r_k > r_{k_{atual}}$, como ilustra a Figura 2.12. Cada reta possui um ângulo θ_k com a horizontal (correspondente à inclinação 0°), calculada da seguinte maneira $\theta_k = \arctg\left(\frac{d_k - d_{k_{atual}}}{r_k - r_{k_{atual}}}\right)$;
4. Obtém-se $\mathbf{b}_{k_{próximo}}$, cujo par correspondente $[r_{k_{próximo}} ; d_{k_{próximo}}]$ possui o menor ângulo θ_k , ou seja, $\theta_{\min} = \min_k \theta_k$;
5. Se $\theta_{\min} \leq 0$, inclui-se $\mathbf{b}_{k_{atual}}$ na curva operacional e atualiza-se $\mathbf{b}_{k_{atual}} = \mathbf{b}_{k_{próximo}}$;
6. Se $\theta_{\min} > 0$, interrompe-se o procedimento;

7. Repetem-se os procedimentos de 3 a 6 até alcançar-se o par de r_k máximo;
8. Ao fim, os pontos pertencentes à curva operacional correspondem aos \mathbf{b}_k (dicionário) ótimos do sinal.

Após o cálculo da curva operacional, ilustrada na Figura 2.11 , realiza-se a alocação de bits definida por \mathbf{b}_k cuja taxa é a mais próxima da taxa alvo r_{alvo} .

Algoritmo 2.1 Cálculo das curvas taxa-distorção operacionais

```

 $\mathbf{b}_{k_o} = \arg \min_{\mathbf{b}_k \in \mathcal{B}} r_k$ 
 $\mathbf{b}_{k_{\text{atual}}} = \mathbf{b}_{k_o}$ 
 $[r_{k_{\text{atual}}}; d_{k_{\text{atual}}}] = [r_{k_o}; d_{k_o}]$ 
enquanto  $r_{k_{\text{atual}}} < \max_{k \in \mathcal{K}}(r_k)$  faça
  para todo  $\mathcal{K}' = \{k \in \mathcal{K} | r_k > r_{k_{\text{atual}}}\}$  faça
     $\theta_k = \arctg\left(\frac{d_k - d_{k_{\text{atual}}}}{r_k - r_{k_{\text{atual}}}}\right)$ 
  fim para
   $k_{\text{próximo}} = \arg \min_{k \in \mathcal{K}'} \theta_k$ 
   $\theta_{\text{mín}} = \min_{k \in \mathcal{K}'} \theta_k$ 
  se  $\theta_{\text{mín}} \leq 0$  então
     $\mathbf{b}_{k_{\text{atual}}}$  pertence à curva operacional
     $\mathbf{b}_{k_{\text{atual}}} = \mathbf{b}_{k_{\text{próximo}}}$ 
     $[r_{k_{\text{atual}}}; d_{k_{\text{atual}}}] = [r_{k_{\text{próximo}}}; d_{k_{\text{próximo}}}]$ 
  senão
    FIM
  fim se
fim enquanto

```

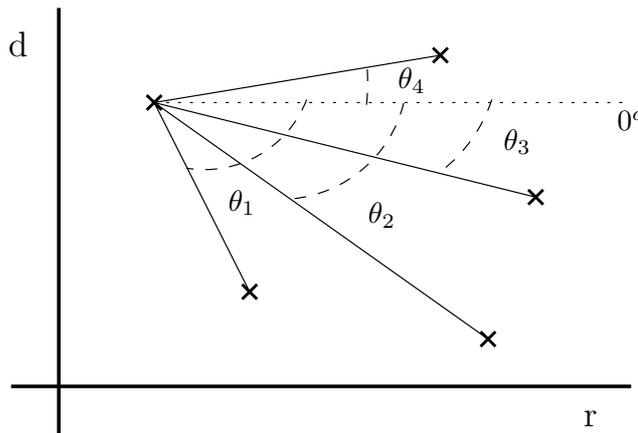


Figura 2.12: Traçando o fecho convexo. Neste caso, $\theta_{\text{mín}} = \theta_1$.

Capítulo 3

Decomposição Atômica de Sinais Elétricos

Neste capítulo, apresenta-se inicialmente um modelo de sinais elétricos definido como um somatório de senoides amortecidas. Em seguida, descreve-se o método de decomposição aplicado a sinais elétricos que tem como engrenagem fundamental o algoritmo de *Matching Pursuit*, desenvolvido em [25]. Este método fornece representações de sinais com base no modelo de senoides amortecidas e é denominado *Matching Pursuit* por Senoides Amortecidas (MPSA). No laço de decomposição, são introduzidas heurísticas que resultam em maior coerência entre átomo e resíduo. No fim do capítulo, apresentam-se diferentes aplicações da decomposição de sinais elétricos através de MPSA.

3.1 Modelo de Sinais Elétricos

Para se obter um método eficiente de decomposição de sinais de perturbações elétricas registrados em arquivos de oscilografia, seria desejável criar um modelo que conseguisse prever todos os fenômenos presentes no sistema elétrico. Todavia, devido às não-linearidades do sistema e às variações de comportamento dos sinais elétricos no tempo, é extremamente complexo gerar um modelo capaz de prever um conjunto abrangente das respostas do sistema elétrico aos variados tipos de perturbações existentes [2]. Algumas das técnicas mais usualmente empregadas na modelagem e análise de sinais de perturbações elétricas são filtragem de Fourier [48, 49]; análise de Prony [50, 51]; modelo auto-regressivo com média-móvel (ARMA) e métodos de rastreamento do espaço de estados [52]; e *wavelets* [50, 53–59]. Em alguns casos esses métodos são utilizados em conjunto com estratégias de inteligência artificial [51, 60–62].

Com base em um modelo simplificado, pode-se dizer que o sistema elétrico é

formado, basicamente, por fontes, cargas e linhas de transmissão, resultando, assim, em circuitos RLC. Portanto, as correntes e tensões são calculadas através de equações diferenciais. Além disso, devem-se considerar também os harmônicos gerados por transformadores e dispositivos de eletrônica de potência e as descontinuidades decorrentes de chaveamentos de circuito causadas pelo sistema de proteção ou por manobras.

A fim de atender a essas premissas, neste trabalho o sinal elétrico foi modelado como um conjunto de senoides moduladas por uma janela exponencial real, concatenadas e sobrepostas, cada uma delas com um suporte no tempo bem definido. Esse modelo é dado por [1, 2, 63]:

$$f(t) = \sum_{q=0}^{Q-1} A_q \cos(2\pi k_q F t + \phi_q) e^{-\rho_q(t-t_{i_q})} [u(t-t_{i_q}) - u(t-t_{f_q})], \quad (3.1)$$

em que Q é o número de componentes, F é a frequência fundamental (50/60Hz) e cada componente é representada pela sêxtupla $(A_q, k_q, \rho_q, \phi_q, t_{i_q}, t_{f_q})$, na qual A_q é a amplitude, k_q é o múltiplo inteiro do harmônico, ρ_q é o fator de decaimento, ϕ_q é a fase, t_{i_q} e t_{f_q} são respectivamente os tempos de início e fim do suporte. A Figura 3.1 ilustra uma forma de onda gerada a partir desse modelo.

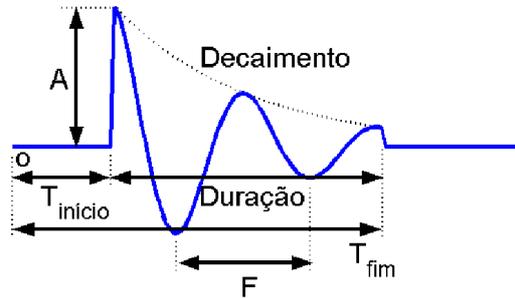


Figura 3.1: Exemplo de forma de onda gerada a partir do modelo de sinais de perturbações elétricas.

Ainda que esse modelo não seja capaz de representar todos os fenômenos do sistema elétrico, ele contempla os seus distúrbios mais comuns e importantes [64]:

1. Harmônicos: senoides com frequência múltipla da frequência fundamental (50 ou 60Hz).
2. Transitórios: impulsos ou oscilações de alta frequência sobrepostos aos sinais de tensão e corrente da frequência fundamental. Normalmente, são causados por raios ou chaveamentos de equipamentos elétricos.
3. Subtensão (*sag*): queda momentânea de tensão de meio ciclo a 1 minuto,

aproximadamente. É ocasionada por energização de cargas pesadas de grandes motores.

4. Sobretensão (*swell*): corresponde à elevação momentânea de tensão, durante meio ciclo a 1 minuto, aproximadamente. Normalmente, está associada a elevações momentâneas nas fases não envolvidas durante uma falta do tipo fase-terra. São bem menos comuns do que as subtensões.

Há um modelo similar proveniente do método de Prony [50, 51] bastante conhecido e utilizado como ferramenta de análise de sinais elétricos, que representa o sinal da seguinte forma:

$$f(t) = \sum_{q=0}^{Q-1} A_q \cos(2\pi f_q t + \phi_q) e^{-\rho_q(t)}. \quad (3.2)$$

Contudo, comparado ao modelo segundo (3.1), este modelo não considera as descontinuidades do sinal devido aos chaveamentos, nem que os fenômenos referentes às perturbações podem surgir em instantes de tempo diferentes no oscilograma [4]. Deve-se notar que o modelo de senoides amortecidas permite a geração de um dicionário completo, pois engloba as bases de impulsos de Dirac no tempo e de exponenciais complexas, isto é, impulsos na frequência.

Na literatura de processamento de sinais, a modelagem através de senoides amortecidas está presente em diversas aplicações. Por exemplo, em [65–67] tais componentes são utilizadas na análise e detecção de transitórios. As inúmeras aplicações referentes às senoides amortecidas são motivadas pelo fato de que estas são soluções de equações diferenciais ordinárias intrínsecas a diversos modelos de sistemas físicos [14, 33, 68, 69]. Há um bom tempo, pesquisadores têm projetado sistemas e algoritmos a fim de estimar os parâmetros destas componentes, encontradas em diversos tipos de sinais [70–77]. Em [78, 79], por exemplo, os sinais elétricos são modelados utilizando-se a componente fundamental com um conjunto de harmônicos, e após subtrair-se do sinal estas componentes, o sinal resultante é decomposto através da transformada *wavelet*. O modelo de sinais em (3.1) difere daqueles em [78, 79], por não restringir a componente fundamental e os harmônicos a possuírem amplitudes e suportes temporais constantes.

Apresenta-se a seguir um método de decomposição que envolve o uso do MP com base em um dicionário de senoides amortecidas, além de um conjunto de heurísticas dentro do laço do MP, com o objetivo de obter uma representação compacta e coerente para sinais de perturbações elétricas.

3.2 Algoritmo de Decomposição – MPSA

Nesta seção, apresenta-se o algoritmo de decomposição atômica que obtém representações de sinais de acordo com o modelo matemático da equação (3.1), o MPSA. A Figura 3.2 apresenta o diagrama de blocos do algoritmo de decomposição. O algoritmo tem como base o *Matching Pursuit* e usa um dicionário parametrizado de senoides amortecidas com parâmetros contínuos. O simples uso do MP com um dicionário de senoides amortecidas não garante uma boa representação do sinal. A fim de aprimorar a modelagem do sinal, um conjunto de heurísticas é introduzido no laço de decomposição do MP tradicional de forma a selecionar átomos com maior grau de similaridade com os padrões do sinal. Esses procedimentos são descritos a seguir e derivam daqueles em [4].

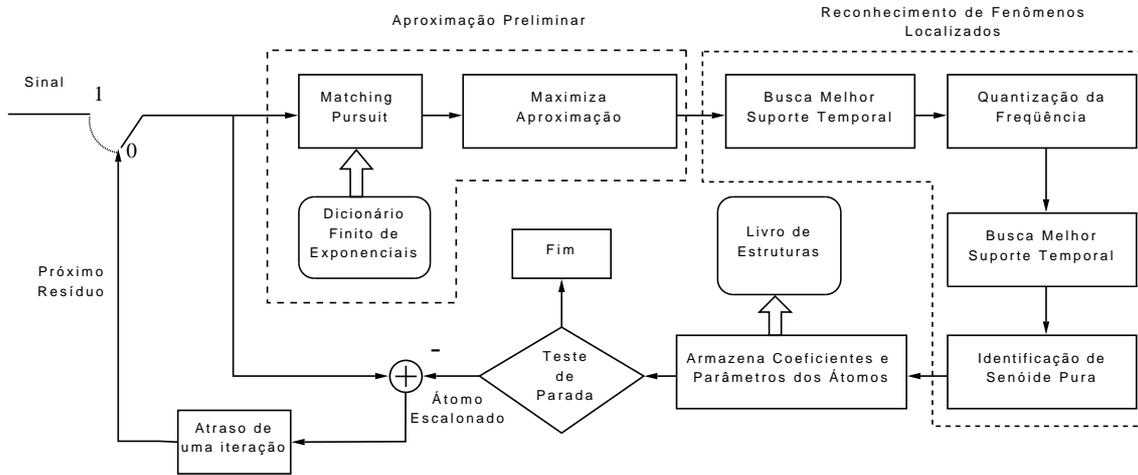


Figura 3.2: Diagrama de blocos do algoritmo de decomposição atômica de sinais. Na primeira iteração, a chave está na posição 1, e nas demais iterações permanece na posição 0.

3.2.1 Dicionário Parametrizado de Senoides Amortecidas

Os elementos \mathbf{g}_γ do dicionário parametrizado de senoides amortecidas são definidos como

$$g_\gamma(n) = K_\gamma g(n) \cos(\xi n + \phi) [u(n - n^s) - u(n - n^e)], \quad (3.3)$$

$$g(k) = \begin{cases} 1, & \text{se } \rho = 0 \quad (\text{Senóide pura (se } \xi \neq 0), \\ & \text{DC ou impulso unitário (se } \xi = 0)) \\ e^{-\rho(n-n^s)}, & \text{se } \rho > 0 \quad (\text{Exponencial decrescente}) \\ e^{\rho(n^e-n)}, & \text{se } \rho < 0 \quad (\text{Exponencial crescente}), \end{cases}$$

em que $K_\gamma \in \mathbb{R}$ é tal que $\|\mathbf{g}_\gamma\| = 1$ e N é o comprimento do sinal. Ademais, o átomo na equação (3.3) é definido pela quintupla $\gamma = (\rho, \xi, \phi, n^s, n^e)$, em que ρ representa o fator de decaimento, ξ é a frequência, ϕ é a fase, e n^s e n^e são as amostras inicial e final, respectivamente. A fase do átomo é calculada de forma analítica em função dos demais parâmetros, sendo ótima no sentido de prover o máximo produto interno entre o átomo e o resíduo a cada iteração [1, 4, 80]. Observa-se que este dicionário de senoides amortecidas é composto tanto por exponenciais crescentes e decrescentes moduladas quanto por impulsos de Dirac, níveis DC e harmônicos, sendo sobrecompleto. É importante destacar que as exponenciais crescentes são obtidas a partir da reversão temporal das decrescentes. A Figura 3.3 mostra alguns elementos desse dicionário.

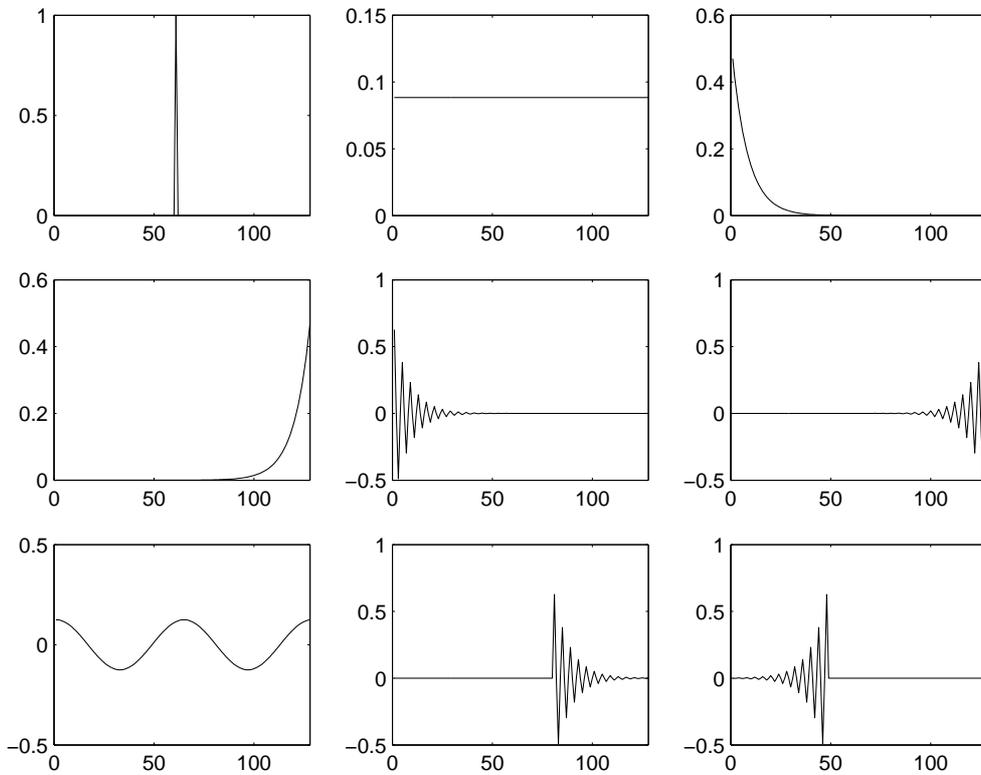


Figura 3.3: Exemplos de átomos do dicionário de senoides amortecidas.

3.2.2 Aproximação Preliminar do Resíduo

Inicialmente, o algoritmo procura o átomo de maior correlação com o resíduo em um dicionário finito de exponenciais com uma discretização pré-definida do espaço de parâmetros. Com base na discretização dos parâmetros do átomo gaussiano realizada em [1, 2, 4, 18], amostra-se γ de tal forma a obter $\gamma_d = (m_f 2^{-j}, k\pi 2^{1-j}, \phi, n^s, n^e)$, onde $\rho = m_f 2^{-j}$; $\xi = k\pi 2^{1-j}$; $n^s = p2^j$ e $n^e = N - 1$, se $\rho \geq 0$; $n^s = 0$ e $n^e = p2^j$, se

$\rho < 0$; N é a dimensão do dicionário e $j, p, k \in \mathbb{Z}$. Dessa maneira, o átomo exponencial discreto é definido somente pelos parâmetros (m_f, j, k, ϕ, p) , onde $m_f = \{-1, 1\}$ e j se referem ao fator de decaimento ou elevação do átomo, p ao seu deslocamento e k à sua modulação. A fase ϕ é calculada em função dos demais parâmetros [25, 80]. Observa-se que os parâmetros discretizados de γ foram ρ , ξ , n^s ou n^e . No dicionário discreto, o deslocamento de átomos exponencialmente decrescentes ocorre variando-se a amostra inicial do suporte e o deslocamento de átomos exponencialmente crescentes ocorre variando-se a amostra final. Os elementos deste dicionário discreto são descritos como

$$g_{\gamma_d}(n) = g_j(n) \cos(nk\pi 2^{1-j} + \phi) \cdot [u(n - n^s) - u(n - n^e)], \quad (3.4)$$

$$n = \{0, \dots, N - 1\}$$

$$g_j(n) = \begin{cases} \delta(j), & j = 0 \\ K_{\gamma_d} e^{-m_f 2^{-j}(n-p2^j)}, & m_f = 1 \text{ e } j \in [1, L] \\ K_{\gamma_d} e^{-m_f 2^{-j}(p2^j-n)}, & m_f = -1 \text{ e } j \in [1, L] \\ \frac{1}{\sqrt{N}}, & j = L \end{cases}.$$

O dicionário possui $L = \log_2(N)$ escalas e os intervalos dos parâmetros que definem os elementos são $j \in [0, L]$, $p \in [0, N2^{-j}]$ e $k \in [0, 2^j]$.

Em seguida, os parâmetros de γ_d encontrados são otimizados com o propósito de obter o γ do conjunto de parâmetros contínuos que maximiza a correlação entre o átomo e o resíduo atual através de um método de otimização denominado pseudo-Newton [1, 4, 25]. A Figura 3.4 ilustra o resultado dessa otimização.

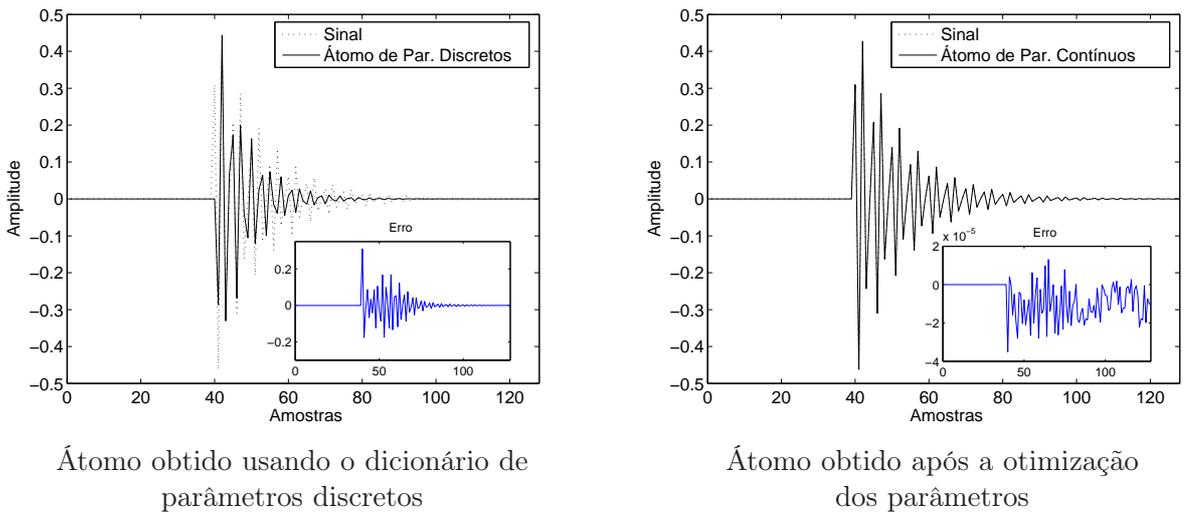


Figura 3.4: Resultado da otimização dos parâmetros do átomo.

A simples decomposição utilizando o MP com um dicionário de senoides amorte-

cidas não garante a obtenção de uma representação coerente do sinal, ou seja, uma representação com interpretação física em relação aos fenômenos presentes em sinais de perturbações elétricas. Na Figura 3.5, apresenta-se um exemplo do que ocorre quando um sinal de falta é decomposto através do MP com base em um dicionário de senoides amortecidas, considerando toda a região de suporte do sinal. A falta ocorre após a amostra 200 do sinal. Observe que o átomo de maior produto interno com sinal não representa bem o sinal de falta em questão, ou seja, este resultado não é satisfatório em termos de coerência, pois não houve casamento do átomo com nenhum dos fenômenos localizados presentes no sinal. Em [81], menciona-se que esse é um problema comum de algoritmos vorazes¹ como o *Matching Pursuit*. Visto que o produto interno não é uma métrica quanto à forma, e sim de correlação, o fato de o produto interno entre o átomo e o sinal possuir um valor significativo não quer dizer que eles sejam coerentes.

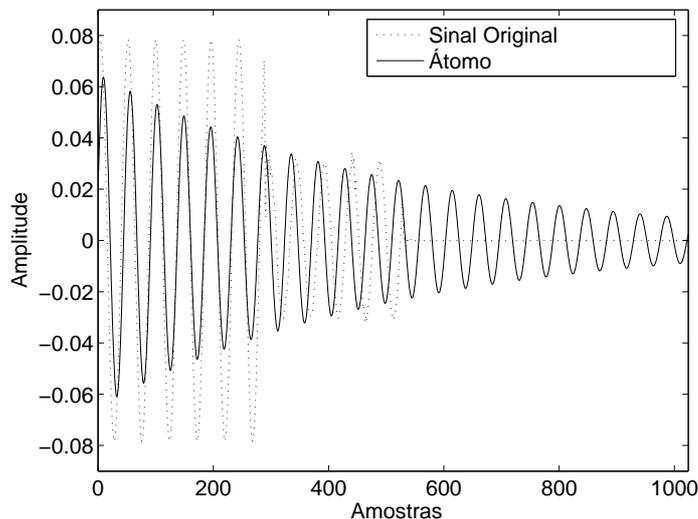


Figura 3.5: Falha do MP na tentativa de encontrar estruturas coerentes com os padrões do sinal.

3.2.3 Reconhecimento de Fenômenos Localizados

Com o objetivo de obter uma decomposição coerente, depois de selecionar o átomo de parâmetros contínuos que melhor aproxima o resíduo, realizam-se procedimentos para o reconhecimento de fenômenos intrínsecos do sinal. Primeiramente, buscase o melhor suporte temporal do átomo determinado pelos parâmetros n^s e n^e . A região de suporte do átomo é reduzida amostra a amostra, verificando-se se um novo suporte temporal é capaz de produzir um melhor casamento entre o átomo e o resíduo

¹Algoritmos vorazes são aqueles que buscam a melhor aproximação do sinal a cada iteração (ótimos locais) com o propósito de alcançar o ótimo global.

atual. É importante mencionar que a norma do átomo é mantida unitária a cada alteração do suporte através de sua renormalização. A Figura 3.6 ilustra o resultado da busca por melhor suporte temporal, aplicada após a aproximação inicial de um sinal elétrico sintético [25]. Nota-se que foram eliminadas as amostras de valor nulo, contudo o átomo ainda não é capaz de realizar um bom casamento com os fenômenos localizados do resíduo. Nesse caso, confundem-se duas senoides de mesma frequência, mas de amplitudes diferentes, com uma única exponencial decrescente modulada. O mesmo acontece com o sinal real [25] ilustrado na Figura 3.7.

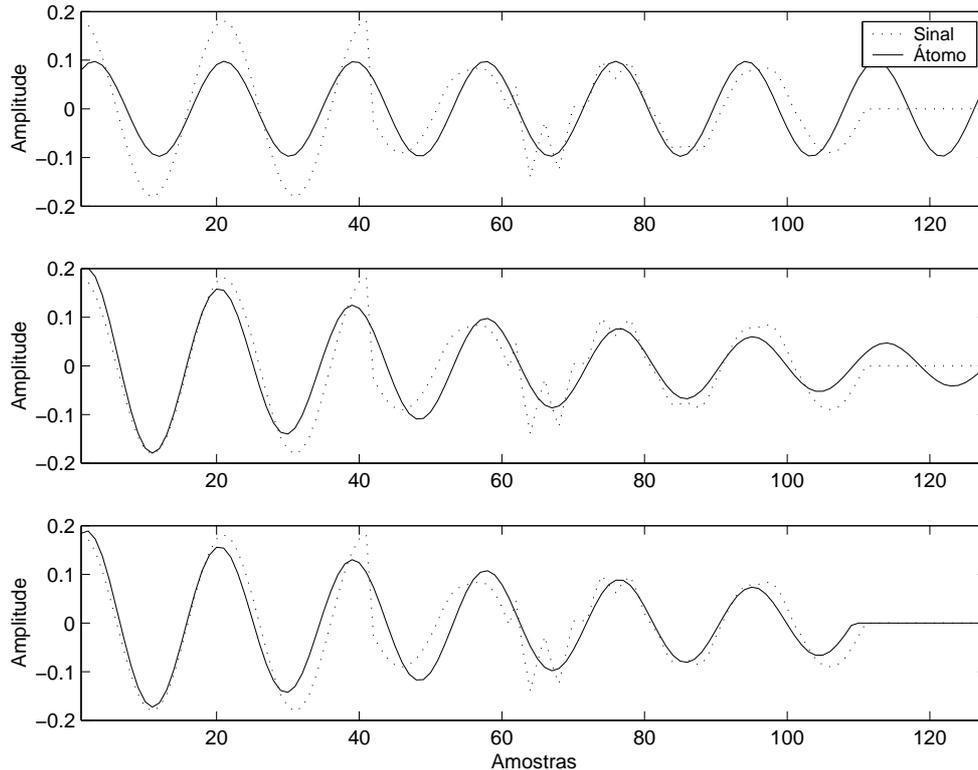


Figura 3.6: De cima para baixo, os resultados referentes à busca no dicionário discreto de exponenciais, à maximização da aproximação do resíduo através da otimização dos parâmetros e à busca por melhor suporte temporal.

O próximo passo do algoritmo de decomposição é quantizar a frequência para um múltiplo da frequência fundamental, tendo em vista o modelo da equação (3.1), e repetir a busca por melhor suporte temporal para a nova frequência quantizada [4, 63]. A quantização da frequência permite separar duas senoides contíguas que estejam defasadas devido à falta ocorrida, mesmo que possuam diferentes amplitudes, como ilustra a Figura 3.8. Observa-se que ao realizar-se a busca do suporte após a aproximação inicial do sinal, o átomo resultante possui uma frequência diferente da do sinal. Ao quantizar-se a frequência, o átomo passa a estar em fase com a senoide pura de maior energia referente à pré-falta. Por fim, quando novamente busca-se um melhor suporte, ocorre o casamento entre o átomo e a senoide.

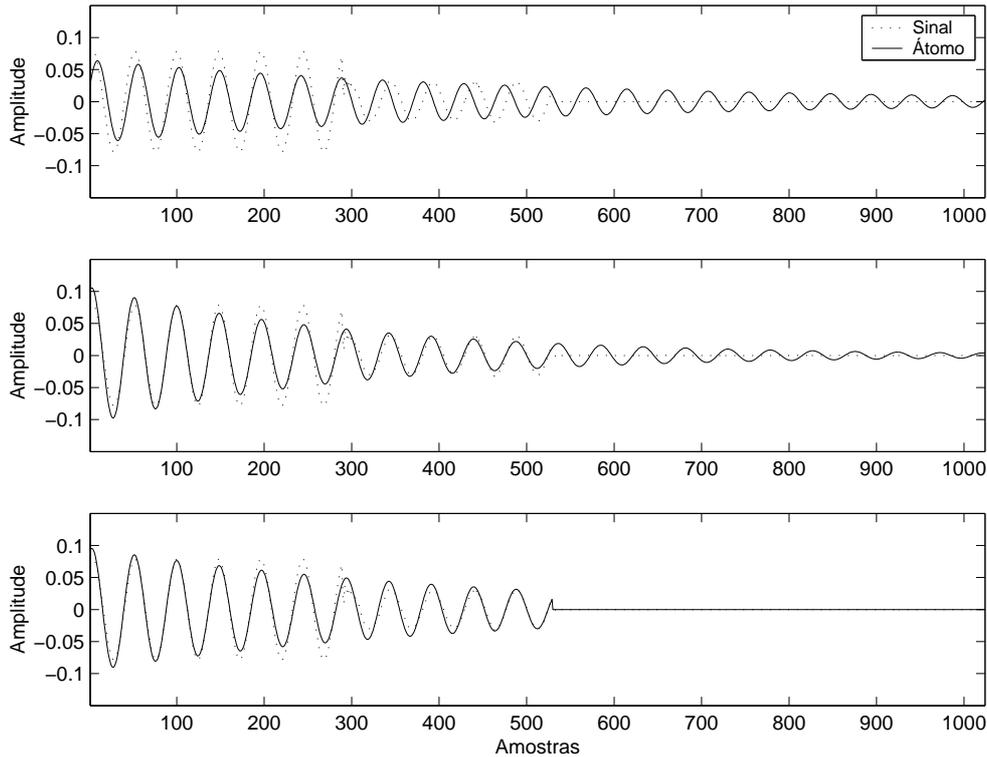


Figura 3.7: De cima para baixo, os resultados referentes à busca no dicionário discreto de exponenciais, à maximização da aproximação do resíduo através da otimização dos parâmetros e à busca por melhor suporte temporal.

No caso em que, por exemplo, existem dois trechos de senoides puras em fase, a quantização de frequência seguida da busca por suporte não é capaz de distinguir uma da outra. Tal situação é ilustrada na Figura 3.9; nesse caso, observa-se que as senoides de um sinal sintético [25] nos trechos inicial e final estão em fase. Dessa forma, o átomo obtido após a quantização da frequência encontra-se em fase com essas duas senoides, impossibilitando que, logo em seguida, a busca por suporte aproxime o átomo de uma delas.

Com o propósito de aproximar o átomo de senoides puras localizadas, em casos nos quais não se obtém êxito após a quantização da frequência, realiza-se uma heurística baseada em uma métrica de similaridade quanto à forma. Escolhe-se, entre o átomo exponencial e uma senoide pura, aquele que melhor atender um compromisso definido entre o erro por amostra e o produto interno em relação ao resíduo dentro da região de suporte [1, 4]. As Figuras 3.10 e 3.11 ilustram bons resultados do procedimento de identificação de senoide pura, onde a quantização da frequência não obtivera sucesso.

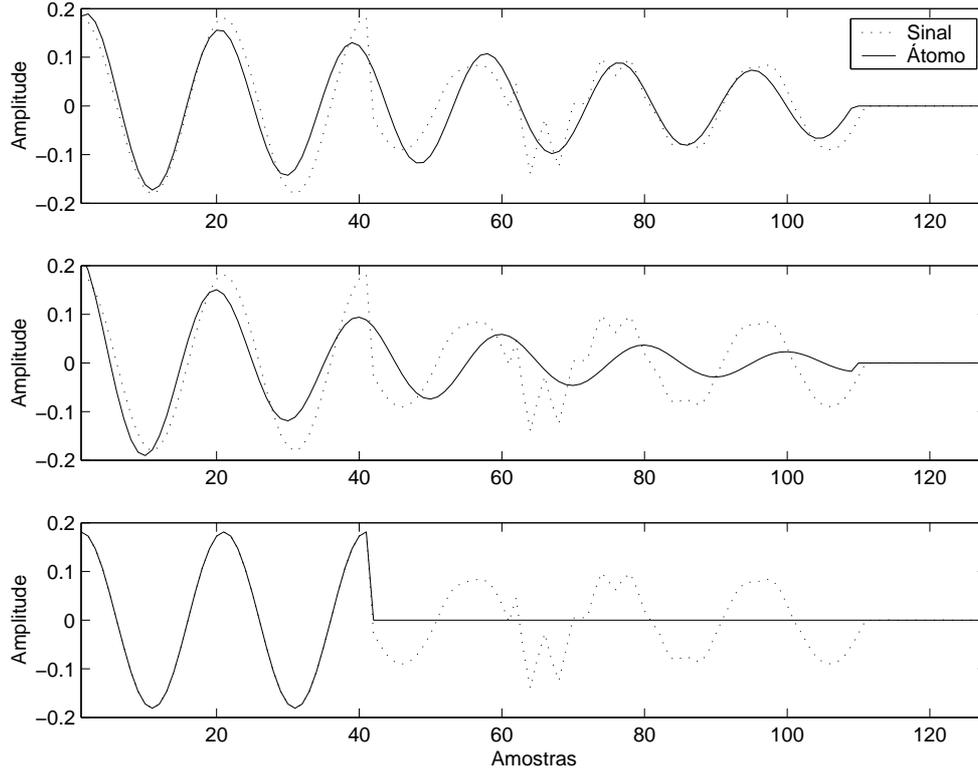


Figura 3.8: De cima para baixo, os resultados da busca por melhor suporte temporal do átomo, seguida pela quantização da frequência e pela nova busca por melhor suporte.

3.2.4 Critério de Parada

O algoritmo de decomposição é interrompido quando a aproximação do sinal for suficientemente boa, significando que os átomos escolhidos para a aproximação são coerentes, ou seja, altamente correlacionados com as características intrínsecas do sinal. Caso o critério de parada não seja satisfeito, o átomo é ponderado pelo seu coeficiente e subtraído do resíduo, produzindo um novo resíduo a ser aproximado na próxima iteração.

Normalmente, utiliza-se a norma do resíduo ou um número máximo de iterações como critério de parada [14, 18, 82]. A norma do resíduo é uma medida do erro de aproximação, contudo com base apenas nela não é possível averiguar se o resíduo ainda é significativamente correlacionado com algum átomo do dicionário [42]. Para isso, opta-se pelo uso da taxa de aproximação [18, 22], introduzida na Seção 2.1.3, definida como

$$\eta(i) = \frac{|\langle \mathbf{r}_x^{i-1}, \mathbf{g}_{\gamma(i)} \rangle|}{\|\mathbf{r}_x^{i-1}\|}, \quad (3.5)$$

que mede o quanto do resíduo \mathbf{r}_x^{i-1} é aproximado na iteração i . É importante lembrar que a taxa de aproximação possui um comportamento exponencialmente decrescente,

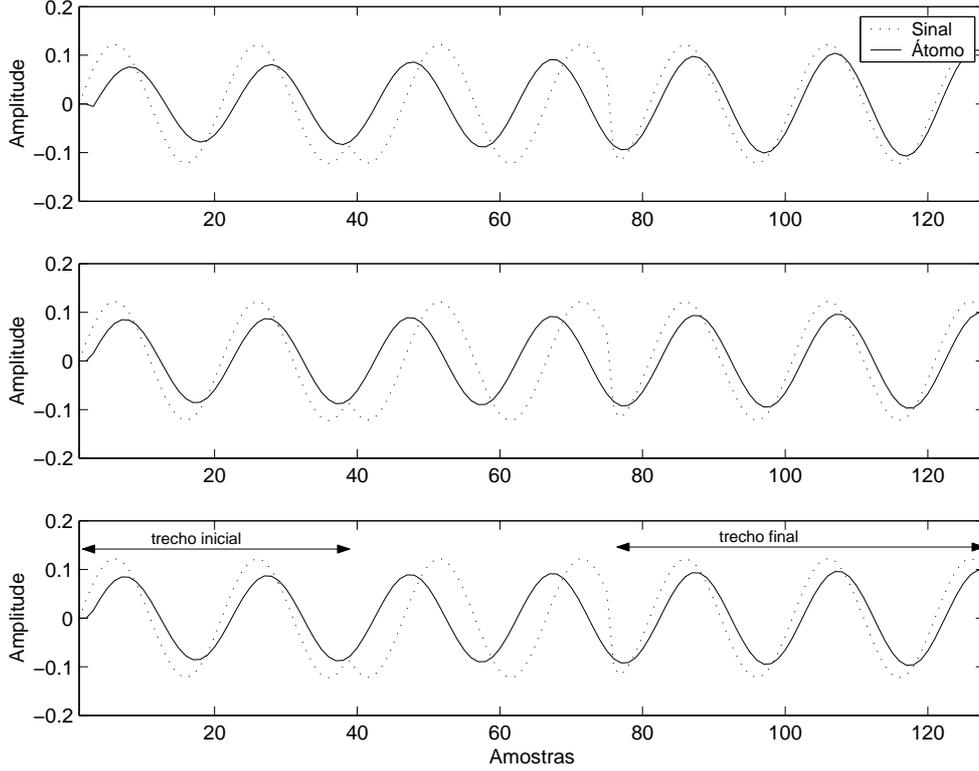


Figura 3.9: De cima para baixo, os resultados da busca por melhor suporte temporal do átomo, seguida pela quantização da frequência e pela busca por melhor suporte novamente.

como é demonstrado em [18], e a derivada da envoltória dessa exponencial diminui à medida que o número de iterações aumenta até que $\eta(i)$ convirja para um valor médio estacionário. Quando a taxa de aproximação atinge esse valor, poder-se dizer que não há mais elementos do dicionário que sejam coerentes com o resíduo, portanto a decomposição pode ser interrompida. Na prática, para um dado passo M , calcula-se uma média móvel considerando os próximos L passos

$$\bar{\eta}(\mathbf{r}_x^M) = \frac{1}{L} \sum_{n=M}^{M+L-1} \frac{\|\langle \mathbf{r}_x^n, \mathbf{g}_{\gamma_n} \rangle\|}{\|\mathbf{r}_x^n\|}, \quad (3.6)$$

onde $L = \log_2 N$, sendo N o comprimento do sinal. Nesse caso, escolheu-se L de forma empírica com o intuito de criar uma relação entre a média e o comprimento do sinal. Quando média móvel se aproxima do valor médio estacionário η_o significa que o resíduo passa a ser incoerente com o dicionário. Dessa forma, a decomposição é realizada enquanto a seguinte desigualdade for satisfeita:

$$\bar{\eta}(\mathbf{r}_x^M) > \eta_o + \epsilon, \quad (3.7)$$

onde ϵ é uma constante de confiabilidade. Dessa forma, obtém-se uma representação

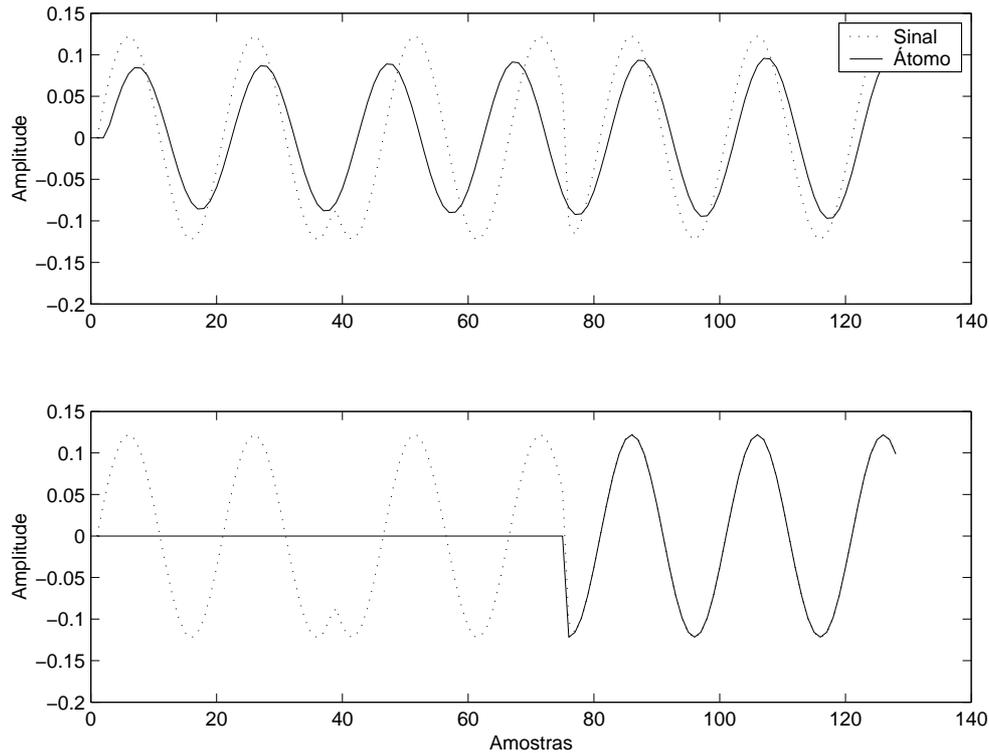


Figura 3.10: De cima para baixo, os resultados da quantização da frequência seguida pela busca por suporte, e da identificação da senoide pura.

do sinal decomposto constituída somente de átomos coerentes, separando-os do ruído [4, 42, 83].

É importante mencionar que η_o depende do dicionário e de sua dimensão, ou seja, do comprimento do sinal. Visto que o dicionário discreto de exponenciais só possui dimensões que são potências de 2, obteve-se η_o para dicionários de dimensões $N = 2^j$, onde $j = \{6, 7, 8, 9, 10, 11\}$, da seguinte forma:

1. Decompõe-se um sinal de ruído branco gaussiano de comprimento N e variância unitária gerado sinteticamente, apenas aproximando o resíduo utilizando MP com base no dicionário de exponenciais de parâmetros contínuos;
2. Calcula-se a taxa de aproximação para um grande número de passos;
3. Ao fim, obtém-se η_o calculando-se a média das taxas de aproximação de todos os passos.

A Figura 3.12 mostra o comportamento da taxa de aproximação em mil passos para ruídos gaussianos de todos os comprimentos relativos ao índice j (64, 128, 256, 512, 1024 e 2048) [25]. Observa-se que para cada comprimento, a taxa oscila em torno de um valor médio desde o início da decomposição. Tendo em vista esses

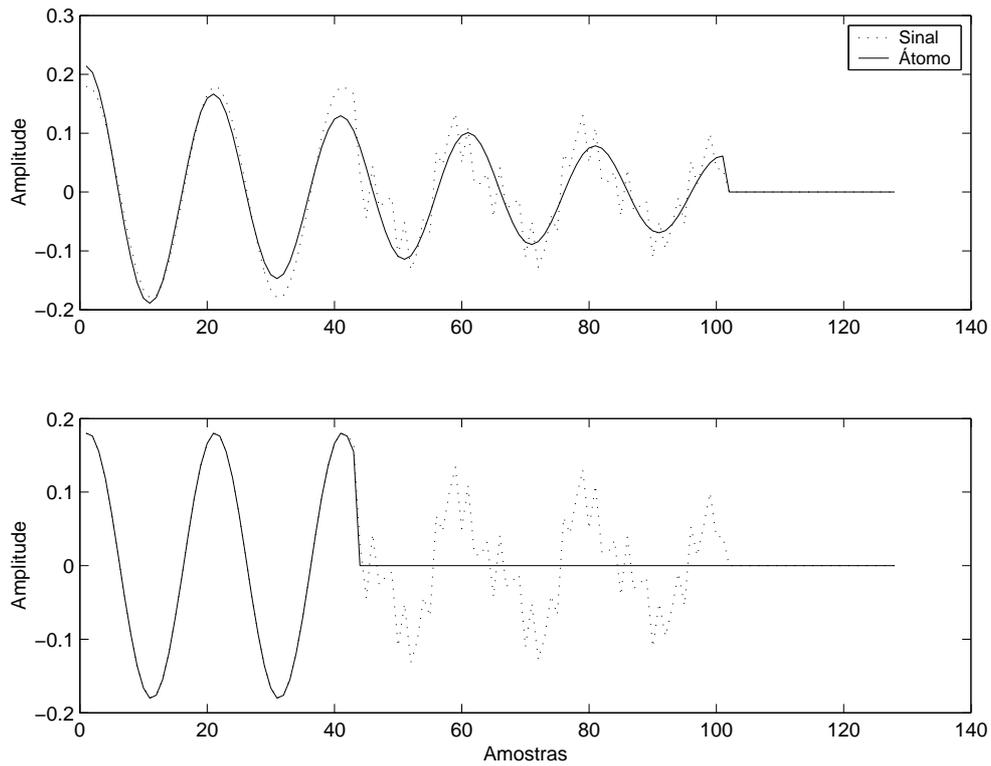


Figura 3.11: De cima para baixo, os resultados da quantização da frequência seguida pela busca por suporte, e da identificação da senoide pura.

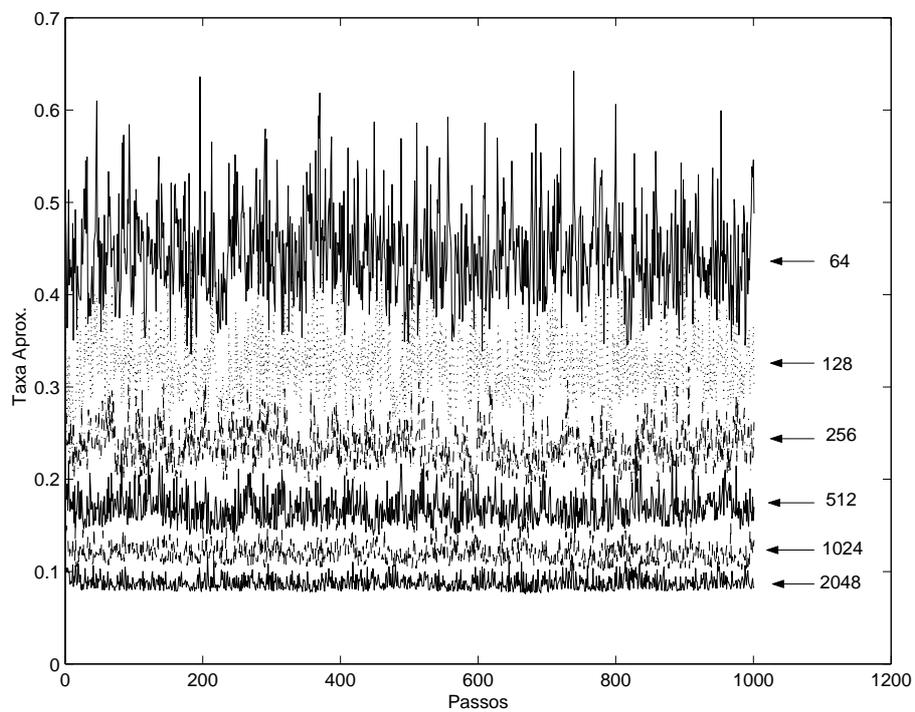


Figura 3.12: Comportamento da taxa de aproximação para ruídos de diferentes comprimentos.

Tabela 3.1: Tabela de valores do critério de parada obtidos a partir da decomposição de sinais de ruído de diferentes comprimentos utilizando MP com base no dicionário de exponenciais com parâmetros contínuos.

Comprimento do Sinal	64	128	256	512	1024	2048
η_o	0,4439	0,3256	0,2361	0,17	0,1203	0,0881

resultados, obtiveram-se os valores de η_o a serem utilizados no critério de parada da decomposição, mostrados na Tabela 3.1.

Ao fim da decomposição, obtém-se a aproximação do sinal representado pela sequência de pares $(\alpha_m, \gamma(m))$, $m = 0, \dots, M - 1$, onde $\gamma(m) = (\rho_m, \xi_m, \phi_m, n_m^s, n_m^e)$ (ver Equações (3.1) e (3.3)). Note que o algoritmo retorna os valores discretos dos parâmetros do átomo n_m^s , n_m^e e ξ_m , enquanto os parâmetros restantes ρ_m , ϕ_m e α_m são contínuos.

A Figura 3.13 mostra como o algoritmo de decomposição se comporta nas primeiras quatro iterações quando aplicado a um sinal elétrico real [3]. O primeiro resíduo é o próprio sinal. Note que os átomos encontrados a cada iteração do algoritmo casam coerentemente com os resíduos correspondentes.

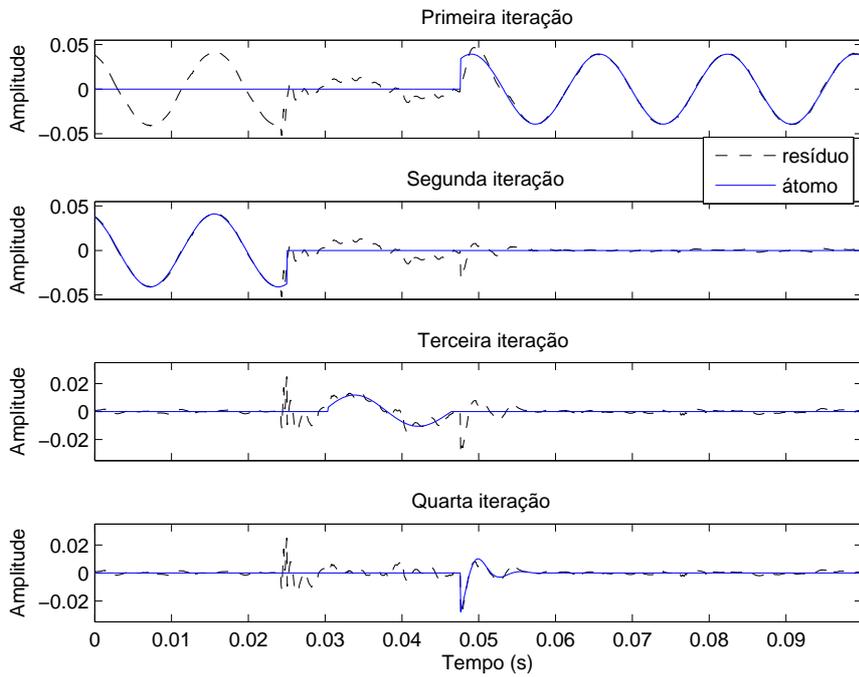


Figura 3.13: Buscando estruturas coerentes em sinais de distúrbios elétricos reais provenientes de [3].

3.3 Aplicações na Modelagem de Sinais Elétricos

Nesta seção apresentam-se algumas aplicações da modelagem de sinais elétricos através do algoritmo de decomposição atômica proposto.

3.3.1 Modelagem Coerente de Sinais

Geralmente, os sinais são adquiridos em ambiente severamente ruidosos, principalmente quando consideramos sistemas de baixa tensão. Idealmente, deseja-se identificar as componentes que caracterizam o sinal em meio ao ruído que possa ser adicionado ao sinal. Contudo, caso o ruído possua uma energia comparável à energia das componentes do sinal, torna-se difícil distinguí-las. Com o uso do método de decomposição apresentado é possível detectar essas componentes mesmo que o sinal esteja corrompido por ruído.

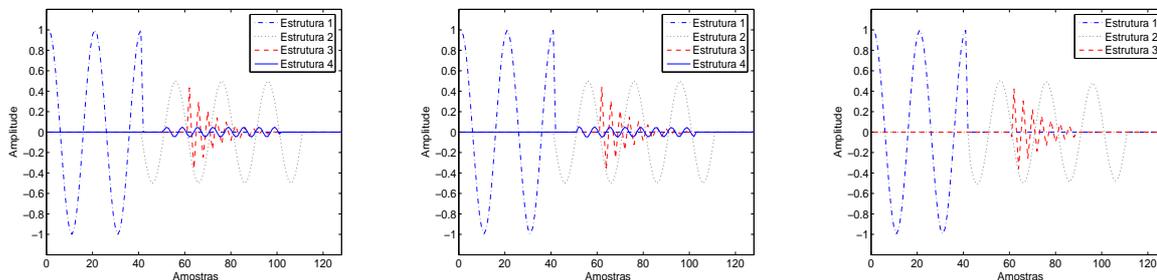
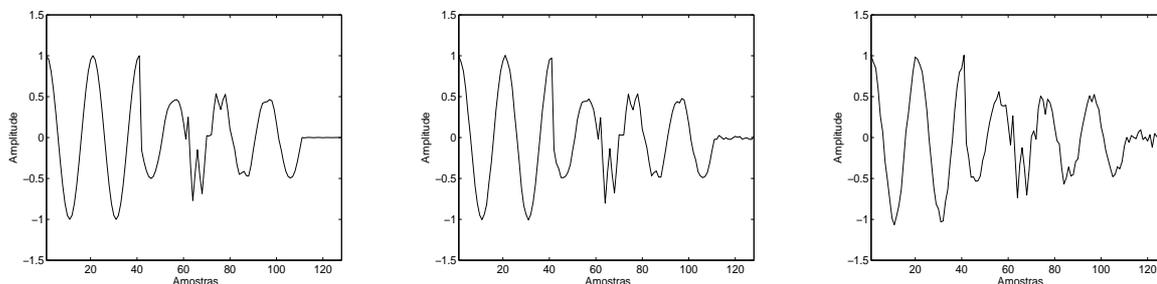
Define-se o sinal com ruído aditivo como

$$\mathbf{x}_{\text{noise}} = \mathbf{x} + \mathbf{n}, \quad (3.8)$$

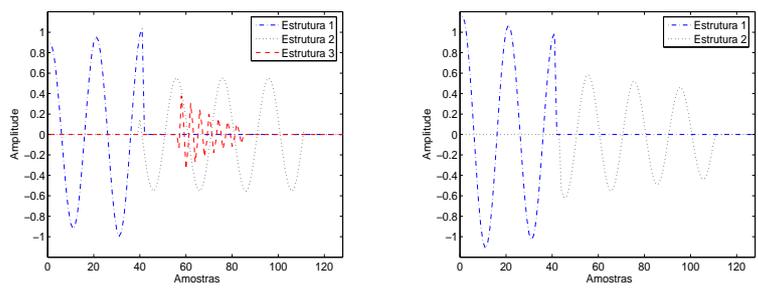
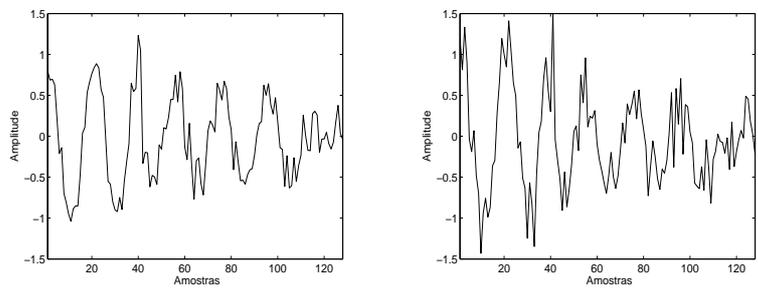
em que \mathbf{n} é o sinal de ruído. A partir desta definição, pode-se calcular

$$\text{SNR}_C = 10 \log_{10} \left(\frac{\|\mathbf{x}\|^2}{\|\mathbf{n}\|^2} \right) = 10 \log_{10} \left(\frac{\|\mathbf{x}\|^2}{\|\mathbf{x} - \mathbf{x}_{\text{noise}}\|^2} \right) \quad [\text{dB}] \quad (3.9)$$

para medir o quanto \mathbf{x} é corrompido pelo ruído. A Figura 3.14 mostra as componentes identificadas em um dado sinal corrompido por sinais de ruído com diferentes níveis de SNR_C , obtidas através do algoritmo de decomposição apresentado na seção anterior. O sinal sintético original (sem adição de ruído) é apresentado no gráfico superior da Figura 3.14.a) e as componentes utilizadas na sua geração estão no gráfico inferior da Figura 3.14.a). A Figura 3.14.b) mostra o sinal da Figura 3.14.a) corrompido pelo ruído tal que $\text{SNR}_C = 30$ dB e as estruturas encontradas pela decomposição. Note que elas são similares às utilizadas na geração do sinal sintético. A Figura 3.14.c) e a Figura 3.14.d) mostram o mesmo sinal corrompido por ruído tal que $\text{SNR}_C = 20$ dB e $\text{SNR}_C = 10$ dB, respectivamente. É possível observar que nestes casos, as três estruturas de maior energia são identificadas, porém a quarta estrutura não é identificada. Quando o ruído adicionado ao sinal é tal que $\text{SNR}_C = 5$ dB, ver Figura 3.14.e), somente as duas estruturas de maior energia são identificadas, embora não tão bem quanto nos casos anteriores. Neste caso, o ruído possui maior energia que a terceira e a quarta estruturas.



a) Original. b) Sujeito a $SNR_C = 30$ dB. c) Sujeito a $SNR_C = 20$ dB.



d) Sujeito a $SNR_C = 10$ dB. e) Sujeito a $SNR_C = 5$ dB.

Figura 3.14: Geração de um modelo coerente de sinais sujeito a diferentes razões sinal-ruído.

3.3.2 Supressão de Ruído por Síntese

O método de decomposição MPSA é capaz de identificar as componentes de sinais, mesmo quando sujeitos a altos níveis de ruído. Dessa forma, esse método também pode ser utilizado com o propósito de remover o ruído presente no sinal. A fim de avaliar a capacidade desta estratégia de suprimir o ruído, gerou-se um conjunto de versões de sinais corrompidos por ruído com diferentes valores de SNR_C (ver equação (3.9)). Dessa maneira, decompõe-se cada versão de sinal corrompido e calcula-se a razão sinal-ruído da reconstrução

$$\text{SNR}_R = 10 \log_{10} \left(\frac{\|\mathbf{x}\|^2}{\|\mathbf{x} - \hat{\mathbf{x}}\|^2} \right) \quad [\text{dB}], \quad (3.10)$$

em que $\hat{\mathbf{x}}$ é o sinal sintetizado (ver equação (2.1)) para diferentes versões corrompidas de \mathbf{x} .

A Figura 3.15 apresenta SNR_R em função de SNR_C para cada um dos sinais na primeira linha da Figura 3.13, provenientes de [3]. É possível notar que SNR_R é sempre maior que SNR_C , especialmente para valores reduzidos de SNR_C , mostrando que esta abordagem de decomposição análise-por-síntese é eficiente na remoção de ruído de sinais.

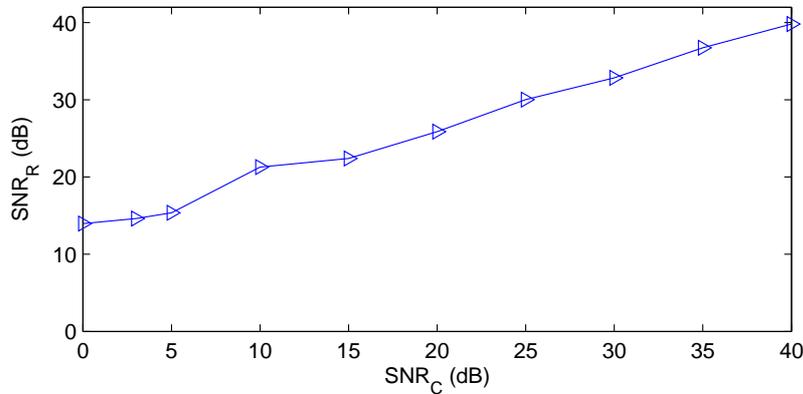


Figura 3.15: Desempenho da supressão de ruído por síntese.

3.3.3 Extração da Componente Fundamental e Separação de Transitórios

Diversos trabalhos têm proposto métodos de análise que extraem inicialmente a componente fundamental dos sinais e em seguida utilizam o sinal residual para a análise do distúrbio, bem como a sua classificação [84–86]. Dado que o método de decomposição extrai automaticamente a componente fundamental do sinal quando esta possui forte presença, pode-se subtraí-la do sinal a fim de separá-la das com-

ponentes associadas a transitórios. As Figuras 3.16.a) e 3.16.b) mostram essa separação de transitórios. Por exemplo, na Figura 3.16.b), observa-se que o método proposto detecta a presença de componentes “DC” com um transitório (evento de potência) ocorrendo a 0,015 segundo. Essa componente está associada a fenômenos de frequência bem reduzida ou nula.

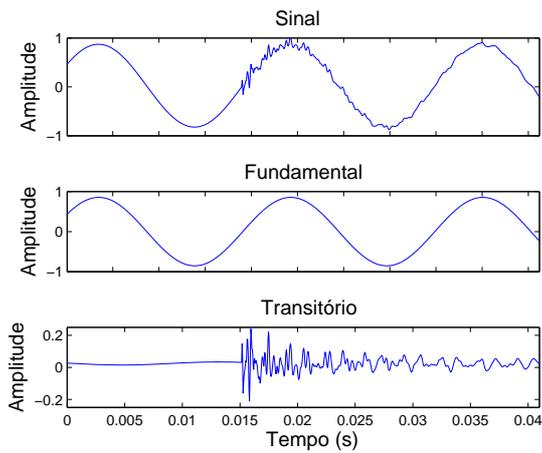
3.3.4 Filtragem da componente “DC”

Às vezes, estão presentes em sinais elétricos componentes “DC” no instante em que o distúrbio ocorre [48]. O sinal corrompido pela componente “DC” (decaimento exponencial) pode ser modelado como

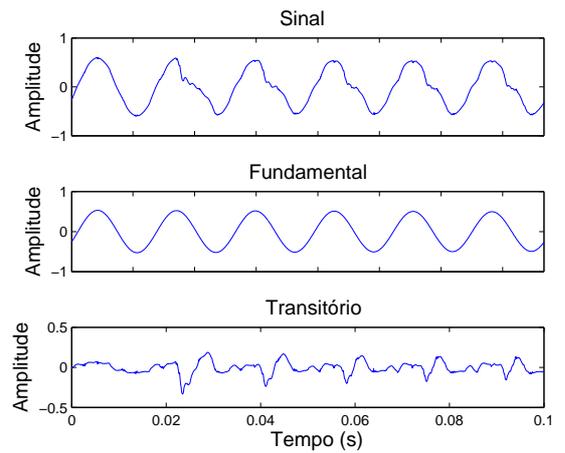
$$Ae^{-\rho t}[u(t - t^s) - u(t - t^e)] + B\text{sen}(2\pi Ft + \phi), \quad (3.11)$$

onde t^s e t^e são os instantes de tempo inicial e final da componente “DC” e ρ expressa a constante de decaimento exponencial. Visto que a equação (3.11) é um caso particular da equação 3.1, o algoritmo de decomposição é capaz de extrair/identificar a componente “DC”. Uma vez que o sinal é decomposto, a componente “DC” pode ser filtrada na síntese do sinal. Esta filtragem é obtida ignorando-se na síntese do sinal todas as estruturas de frequência bem baixa ou nula e que não possuem natureza impulsiva (suporte temporal que não seja menor que 10% do período da frequência fundamental) [4, 9] obtidas através da análise do sinal.

De forma geral, a análise de sinais de distúrbios elétricos se baseia na comparação de valores de grandezas de corrente e tensão, usualmente através da representação fasorial. Para isso, o sinal é filtrado a fim de se obter somente a contribuição da frequência fundamental, usando, por exemplo, o filtro de Fourier [48, 49]. Utiliza-se essa medida para avaliar a eficiência do método proposto em filtrar a componente “DC”. A Figura 3.17 mostra um exemplo da filtragem dessa componente “DC” presente em um sinal sintético gerado com base no modelo da equação (3.1). As componentes do sinal original são duas senoides de 60Hz com amplitudes 1,0 e 2,0 e fases 0° and 90° , que vão das amostras 0 a 50, e 50 a 100, respectivamente. Uma componente “DC” é adicionada ao sinal iniciando-se na amostra 50 e terminando na amostra 100. Seu fator de decaimento é 0,05 e sua amplitude é 3,0. Na Figura 3.17, observa-se que após a filtragem a componente “DC” é quase totalmente eliminada. Além disso, os fasores de tensão e corrente são bem próximos dos fasores calculados para o sinal sintético sem a adição da componente “DC”. Essa filtragem se mostra bastante efetiva quando aplicada a sinais sintéticos e naturais, bem como a sinais simulados em ATP-EMTP [87]. Outro exemplo deste processo de filtragem pode ser encontrado em [4].



a) Sinal elétrico proveniente de [3].



b) Sinal elétrico proveniente de [4].

Figura 3.16: Extração da fundamental e separação de transitórios de sinais elétricos.

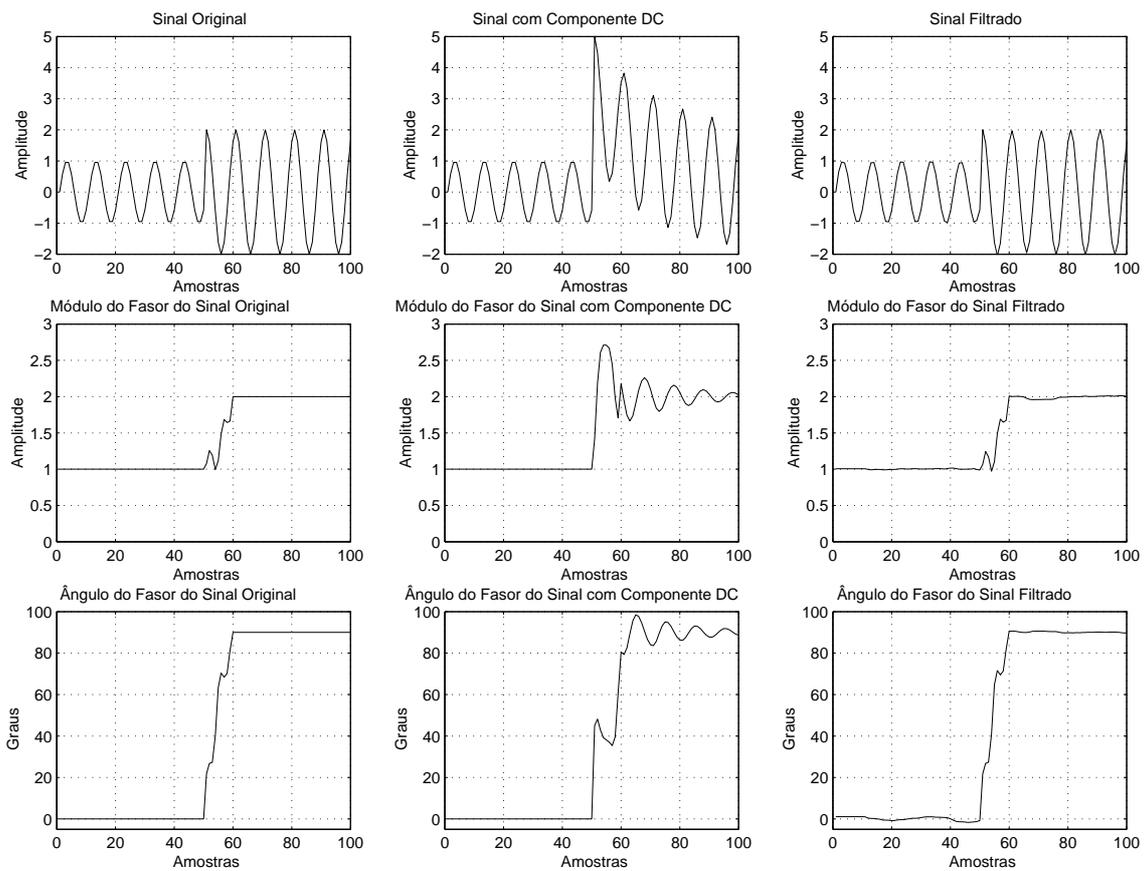


Figura 3.17: Filtro de Fourier aplicado após a filtragem da componente “DC” do sinal sintético.

Capítulo 4

Compressão de Sinais Elétricos Usando Decomposições Atômicas

Neste capítulo, apresenta-se inicialmente a motivação para a compressão de registros de perturbações elétricas. Em seguida, apresenta-se o sistema de compressão aplicado a sinais elétricos proposto e desenvolvido nesta tese. Este sistema se baseia no método de decomposição atômica com base no dicionário de senoides amortecidas (MPSA), apresentado no Capítulo 3, e no procedimento de seleção do dicionário com compromisso taxa-distorção ótimo (ver Seção 2.2—página 17). Descrevem-se os detalhes dos processos de quantização do livro de estruturas e alocação ótima de bits através de curvas de taxa-distorção operacionais, que fazem parte deste arcabouço de compressão. Por fim, avalia-se o desempenho desse sistema, com o propósito de determinar o nível de compressão atingível de forma que não ocorra uma discrepância significativa entre os resultados obtidos na análise de faltas do arquivo comprimido com relação ao não comprimido. Esta avaliação é realizada com o auxílio de um sistema de análise automática de faltas em linhas de transmissão.

4.1 Motivação

A análise integrada de perturbações no sistema elétrico de potência brasileiro teve início nos anos 70 pela atuação do hoje extinto GCOI - Grupo Coordenador para Operação Interligada [88]. Até então, distúrbios em sistemas elétricos eram tratados de forma empírica, sem uma análise criteriosa das causas e efeitos, principalmente pela falta de conhecimentos ou experiência em análise desses tipos de perturbações no país.

Através da análise sistemática de perturbações, realizada principalmente em empresas concessionárias de serviços de energia elétrica, houve condição para a aquisição de conhecimentos e consolidação de técnicas de análise e diagnóstico. A grande

dificuldade, na época, tanto para as concessionárias quanto para a indústria, era a falta de ferramentas adequadas para a obtenção dos dados.

Já nos anos 50, havia surgido a oscilografia de falhas em sistemas elétricos, que registrava as perturbações através de aparelhos eletromecânicos a tinta. A partir dos anos 70, surgiram os oscilógrafos analógicos que utilizavam papel fotográfico, os quais, todavia, eram utilizados apenas em sistemas de extra-alta tensão, por serem bastante onerosos. Além disso, com o tempo, o papel fotográfico se degradava e perdiam-se os dados.

Os registradores digitais de perturbações apareceram na segunda metade da década de 80 do século XX e se consolidaram durante a década seguinte. Com a evolução tecnológica, o barateamento do *hardware* e a ampliação do mercado, os recursos de oscilografia foram incorporados em grande parte dos relés digitais utilizados em sistemas de proteção.

Atualmente, o sistema elétrico brasileiro consiste de uma vasta malha interligada composta de diversos agentes que possuem diferentes funções, como geração, transmissão e distribuição de energia, como exemplificado na Figura 4.1. A geração de energia é realizada por usinas hidráulicas, térmicas ou fontes não convencionais (ex. eólica e solar). Os agentes de transmissão têm o propósito de transportar a energia elétrica das usinas geradoras aos centros de consumo, sendo seus principais ativos as linhas de transmissão e as subestações. Os agentes de distribuição têm a finalidade de transportar e distribuir a energia elétrica aos consumidores finais, através de linhas de subtransmissão, subestações (transformadoras abaixadoras e seccionadoras), redes primárias, transformadores de distribuição e redes secundárias [88].

A análise de perturbações é mais utilizada pelos agentes de transmissão e de distribuição, visto que a maioria dos distúrbios elétricos são registrados a partir do monitoramento das linhas de transmissão. É importante destacar que a oscilografia tem o objetivo de viabilizar a análise pós-evento dos distúrbios, diferentemente dos sistemas de proteção, que devem atuar em tempo real em resposta aos distúrbios. Na realidade, a oscilografia é um instrumento complementar aos sistemas de proteção, pois permite que o especialista em análise de perturbações verifique a correção dos ajustes da proteção, assim como eventuais defeitos que venham a surgir.

Um cálculo bastante útil realizado por especialistas a partir dos sinais de oscilografia é o da distância em que ocorreu o distúrbio [26, 89]. Nesse caso, o especialista informa à equipe de manutenção em que região da linha de transmissão ela deve atuar a fim de reparar os danos causados pela perturbação, tornando o seu trabalho mais rápido e eficiente. Além disso, o especialista realiza outros procedimentos, tais como a análise fasorial, para verificar o desbalanceamento entre as fases, e a análise harmônica, para observar a intensidade dos harmônicos presentes no sinal.

A digitalização dos sinais de oscilografia motivou o crescimento do número de

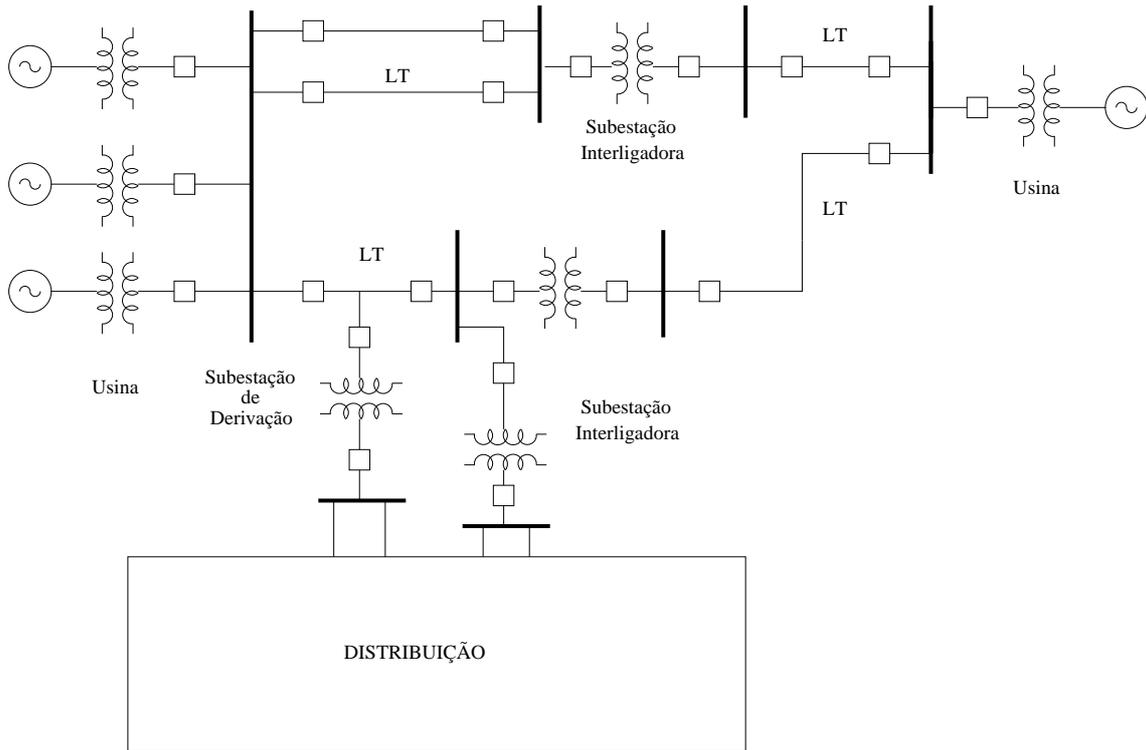


Figura 4.1: Esquema simplificado de um sistema elétrico.

ferramentas computacionais desenvolvidas para o auxílio da análise de perturbações, permitindo também o desenvolvimento de ferramentas sofisticadas de processamento de sinais, e sistemas de processamento inteligente [26].

Além disso, em face da disseminação de oscilógrafos digitais, tem-se observado um aumento da quantidade de eventos registrados, o que implica um crescimento progressivo do número de arquivos de oscilografia. Dessa maneira, surge a necessidade de se estudar e desenvolver métodos de compressão com o propósito de reduzir o espaço necessário para o armazenamento desses arquivos e realizar um melhor uso dos recursos de transmissão. Na literatura, é possível encontrar diversos métodos de compressão aplicados a sinais elétricos; em geral, se baseiam em representação de sinais através de *wavelets* [79, 90–93]. As técnicas de análise através de *wavelets* são caracterizadas por proporcionarem representações adequadas aos fenômenos transitórios. Contudo, elas provêm uma representação mais limitada para fenômenos de banda estreita bem localizados na frequência, como componentes fundamentais e harmônicas. O método de decomposição usado no sistema de compressão proposto neste trabalho apresenta maior flexibilidade na representação tanto de fenômenos transitórios quanto de fenômenos bem localizados na frequência.

4.2 Sistema de Compressão

O sistema de compressão proposto se baseia no método de decomposição atômica, o MPSA, e em técnicas de quantização e de alocação ótima de bits, também desenvolvidas neste trabalho. A Figura 4.2 apresenta o diagrama de blocos do sistema de compressão. Inicialmente, cada sinal do oscilograma é decomposto em uma combinação linear de senóides amortecidas utilizando o algoritmo de decomposição MPSA, de acordo com o modelo matemático da equação (3.1). Como visto anteriormente, o algoritmo seleciona em um dicionário de senóides amortecidas parametrizadas um subconjunto de átomos, que são os mais correlacionados com os padrões existentes no sinal. Definindo-se \mathbf{x} como o sinal, obtém-se a seguinte aproximação de M termos do sinal, apresentada novamente por conveniência:

$$\mathbf{x} \approx \hat{\mathbf{x}} = \sum_{m=1}^M \alpha_m \mathbf{g}_{\gamma(m)}, \quad (4.1)$$

em que cada componente é caracterizada por seu coeficiente α_m correspondente e por um conjunto de parâmetros $\gamma(m)$, definindo a senoide amortecida $\mathbf{g}_{\gamma(m)}$.

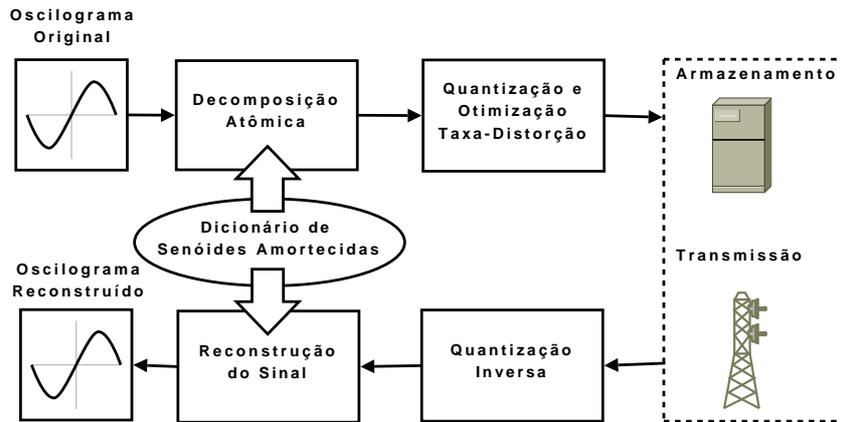


Figura 4.2: Diagrama de blocos do processo de compressão de sinais de perturbações elétricas.

Ao fim da decomposição, tem-se a sequência dos pares $(\alpha_m, \gamma(m))$, $m = 1, 2, \dots, M$, que formam o livro de estruturas, cujo $\gamma(m) = (\rho_m, \xi_m, \phi_m, n_m^s, n_m^e)$ (ver equações (3.1) e (3.3)). Com o objetivo de comprimir, os coeficientes e os parâmetros do livro de estruturas são quantizados, produzindo uma sequência de bits a ser transmitida ou armazenada. A alocação de bits entre os parâmetros é realizada através de um esquema de otimização taxa-distorção. Após a decodificação, posto que há perda de informação em decorrência da quantização, obtém-se uma aproximação do livro de estruturas codificado através de quantização inversa. Observa-se que a quantização é realizada sobre os parâmetros, e não diretamente em formas de onda. Finalmente, o sinal é reconstruído através da soma ponderada dos elementos

do livro de estruturas reconstruído, como na equação (4.1). Note que quanto maior o erro produzido pela quantização do livro de estruturas, maior é a distorção no sinal reconstruído.

4.3 Quantização do Livro de Estruturas

O coeficiente α_m e cada parâmetro de $\gamma(m)$ associados aos elementos do livro de estruturas são quantizados utilizando-se um quantizador escalar uniforme definido como [44]

$$x_q = I_x \times \Delta_{q(x)}, \quad \text{no qual } I_x = \left\lfloor \frac{x + \frac{\Delta_{q(x)}}{2}}{\Delta_{q(x)}} \right\rfloor, \quad (4.2)$$

em que x é um parâmetro qualquer, x_q representa sua versão quantizada, $\Delta_{q(x)}$ é o passo de quantização associado a x e I_x corresponde ao índice associado a x . Os quantizadores de cada parâmetro são definidos por um intervalo dinâmico e um número de níveis, restritos a potências de 2, de forma a serem indexados usando um número inteiro de bits. As amostras inicial e final da região de suporte do átomo, n_m^s e n_m^e , são quantizadas com $\log_2(N_s)$ bits, em que N_s é o tamanho do sinal. A frequência ξ_m é quantizada com $\log_2\left(\frac{F_s/2}{F}\right)$ bits, onde F_s é a frequência de amostragem e F representa a frequência fundamental em que o sistema elétrico opera. Os coeficientes α_m , os fatores de decaimento ρ_m e as fases ϕ_m são quantizados de acordo com um intervalo dinâmico definido por seus respectivos valores máximo e mínimo dentre todos os elementos do livro de estruturas, ou seja,

$$\Delta_{q(x)} = \frac{x_{\text{máx}} - x_{\text{mín}}}{2^{b_x} - 1}, \quad (4.3)$$

onde b_x é o número de bits alocados a x . Para o quantizador do coeficiente faz-se $\alpha_{\text{mín}} = 0$. Os intervalos dinâmicos e o número de bits alocados aos parâmetros α_m , ρ_m e ϕ_m , o tamanho do sinal e as frequências de amostragem e fundamental são codificados como informação adicional. Neste caso, opta-se por utilizar um quantizador *mid-rise*.

É importante lembrar que o método de decomposição MPSA utiliza um dicionário de senoides amortecidas, cujos átomos são definidos por parâmetros contínuos ou densamente quantizados [4, 41, 42]. Isso significa dizer que os parâmetros são restringidos somente pela representação numérica de ponto flutuante. Neste caso, o dicionário possui uma infinidade de elementos devido às inúmeras possibilidades de valores que os parâmetros do átomo podem assumir. Considerando o espaço de parâmetros infinito Γ associado ao dicionário de parâmetros contínuos, a quantização dos parâmetros do livro de estruturas equivale a selecionar subconjuntos $\Gamma_{D_i} \in \Gamma$ para

codificar o sinal. Cada subconjunto de parâmetros define um dicionário distinto. Portanto, a indexação dos diferentes dicionários é realizada através da quantização dos parâmetros dos átomos e de seus coeficientes.

A quantização dos parâmetros e dos coeficientes resulta na seguinte aproximação do sinal:

$$\tilde{\mathbf{x}} = \sum_{m=1}^M Q^\alpha[\alpha_m] \mathbf{g}_{Q^i[\gamma(m)]}, \quad (4.4)$$

em que $Q^\alpha[\cdot]$ é a função de quantização dos coeficientes e $Q^i[\cdot]$ representa a função de quantização associada ao dicionário D_i . Dependendo da quantidade de bits alocados entre os parâmetros, diferentes átomos irão compor a aproximação do sinal, ou seja, utilizar-se-á um dicionário distinto. Mais precisamente, o dicionário usado é definido pela função $Q^i[\gamma(m)]$, e $\tilde{\mathbf{x}}$ corresponde à soma ponderada dos elementos selecionados do dicionário. Dessa maneira, o problema de otimização taxa-distorção consiste em buscar os quantizadores $Q^\alpha[\cdot]$ e $Q^i[\cdot]$ que levam à mínima distorção para uma dada taxa de compressão.

4.4 Otimização Taxa-Distorção

No paradigma de compressão em questão, busca-se o compromisso entre o número de átomos da representação do sinal, a quantização de cada coeficiente e a escolha do dicionário que é definido pelos quantizadores dos parâmetros do átomo (ver Seção 2.2). Neste caso, o número de bits de um átomo senoidal amortecido é definido como

$$r = r_\alpha + r_\xi + r_\rho + r_\phi + r_{n^s} + r_{n^e}, \quad (4.5)$$

em que r_α , r_ξ , r_ρ , r_ϕ , r_{n^s} e r_{n^e} representam os números de bits alocados a α , ξ , ρ , ϕ , n^s e n^e , respectivamente. O total de bits gastos na compressão do sinal, com exceção da informação adicional, é de $r \times M$, sendo M o número de termos da reconstrução do sinal (ver equação (4.4)). As grandezas r_{n^s} e r_{n^e} são definidas pelo tamanho do sinal, enquanto r_ξ é definido pelas frequências de amostragem e fundamental, i.e., elas não variam em função da escolha do dicionário. Dessa forma, a distorção total é expressa somente em função dos números de bits gastos em coeficientes, fatores de decaimento e fases, resultando em

$$d = \frac{1}{N_s} \|\mathbf{x} - \tilde{\mathbf{x}}\|^2 = f(r_\alpha, r_\rho, r_\phi), \quad (4.6)$$

em que N_s é comprimento do sinal. Note que d também representa erro quadrático médio (MSE).

Considerando-se o quantizador definido por (4.2) e os comprimentos de bits de

cada parâmetro na sêxtupla

$$\mathbf{b}_k = (r_\alpha, r_\xi, r_\rho, r_\phi, r_{n^s}, r_{n^e}) \in \mathcal{B}, \quad (4.7)$$

em que \mathcal{B} é definido pelos intervalos de r_α , r_ρ e r_ϕ (r_ξ , r_{n^s} e r_{n^e} são fixos e dependem somente do sinal original), cada \mathbf{b}_k define um quantizador de coeficientes e os quantizadores para os parâmetros do átomo. Os quantizadores dos parâmetros correspondem à escolha de um dicionário (ver equação (4.4)). A fim de se obter o melhor compromisso taxa-distorção, deve-se buscar o \mathbf{b}_k que minimiza a distorção total inserida no processo de codificação, dada uma quantidade de bits desejada. Para isso, utiliza-se o método de alocação ótima de bits apresentado na Seção 2.2.3 (página 20) correspondente ao Algoritmo 2.1, responsável por obter as curvas operacionais de taxa-distorção. É importante destacar que lança-se mão deste algoritmo, porque d não possui forma fechada em função de $(r_\alpha, r_\xi, r_\rho, r_\phi, r_{n^s}, r_{n^e})$.

Observe que os conjuntos de pares T-D diferem entre sinais, ou seja, um \mathbf{b}_k ótimo para um sinal pode não ser ótimo para outros sinais. Por isso, propõe-se uma estratégia com o intuito de reduzir o tamanho de \mathcal{B} , utilizando-se somente os \mathbf{b}_k que fornecem um dicionário e um quantizador de coeficientes ótimos em compromisso T-D para pelo menos um sinal. Constroem-se curvas operacionais para um conjunto de sinais como treinamento e a partir destas curvas obtém-se um conjunto reduzido formado por \mathbf{b}_k associados aos pontos ótimos das curvas. Esta estratégia é equivalente à redução do número de dicionários que podem ser escolhidos (ver Figuras 2.7 e 2.9, Seção 2.2—páginas 18 e 21) para decompor os sinais. A estratégia proposta objetiva reduzir o custo computacional da otimização taxa-distorção, diminuindo o número de total de pares T-D a serem calculados. A partir desses pares, obtém-se para cada sinal o fecho convexo.

4.4.1 Resultados Experimentais

Nesta seção, avalia-se o desempenho do método proposto para otimização taxa-distorção (T-D). O conjunto reduzido de \mathbf{b}_k utilizado neste esquema de otimização é obtido a partir de um conjunto \mathcal{B} definido variando-se r_α de 3 a 16 bits, e r_ρ e r_ϕ de 1 a 12 bits. Dessa forma, \mathcal{B} possui 2016 elementos. Usando \mathcal{B} , obtêm-se as curvas taxa-distorção operacionais para um conjunto de treinamento de 29 sinais de distúrbios elétricos adquiridos no sistema elétrico brasileiro. Esse conjunto de treinamento é caracterizado pelos fenômenos mais comuns existentes em sinais elétricos [64], descritos na Seção 3.1 (página 3.1). A quantidade de sinais deste conjunto é considerada adequada para gerar o conjunto reduzido de \mathbf{b}_k capaz de fornecer um desempenho em termos de taxa-distorção quase ótimo, como verificado

a seguir. Foi observado que de um conjunto completo de 2016 possíveis \mathbf{b}_k utilizados para obter as curvas operacionais do conjunto de treinamento, somente 292 eram de fato ótimos para pelo menos um sinal. Portanto, o uso deste conjunto reduzido resulta em uma redução de 7 vezes na complexidade de codificação.

Os valores de r_α , r_ρ e r_ϕ , as frequências fundamental e de amostragem e o comprimento do sinal, bem como os intervalos dinâmicos dos parâmetros e dos coeficientes são codificados como um cabeçalho de 149 bits.

A Figura 4.3 mostra dois sinais de perturbações elétricas R_1 e R_2 que foram utilizados em [4]. Eles são originalmente representados com 16 bits por amostra e não pertencem ao conjunto de treinamento empregado na seleção dos quantizadores utilizados na otimização T-D. O sinal R_1 possui uma taxa de amostragem de 1000 amostras por segundo e o sinal R_2 , uma taxa de amostragem de 24960 amostras por segundo. Em [4], os quantizadores foram escolhidos sem a preocupação de se obter o melhor compromisso taxa-distorção. Por exemplo, para o sinal R_1 em [4] utilizando uma taxa de 1,035 bit/amostra, o sinal foi reconstruído com uma razão sinal-ruído (SNR = $10 \log_{10} \left(\frac{\|\mathbf{x}\|^2}{\|\mathbf{x}-\tilde{\mathbf{x}}\|^2} \right)$ dB) de 28,08 dB, enquanto a otimização taxa-distorção forneceu um quantizador que codifica R_1 a uma taxa de 0,95 bit/amostra com 31,65 dB de SNR. Para o sinal R_2 , um quantizador definido por $r_\alpha = 6$, $r_\rho = 6$ e $r_\phi = 6$ foi empregado em [4], resultando em uma SNR de 31,12 dB a 0,584 bit/amostra. A otimização T-D proposta encontrou melhores quantizadores para R_2 a uma taxa de bits similar, este quantizador é dado por $r_\alpha = 6$, $r_\rho = 3$ e $r_\phi = 4$ resultando em uma taxa de 0,578 bit/amostra associada a uma SNR de 31,39 dB (ver Figura 4.4). A Figura 4.3 mostra também as versões reconstruídas dos sinais R_1 e R_2 comprimidos a 0,950 e 0,578 bit/amostra respectivamente, e seus erros de reconstrução.

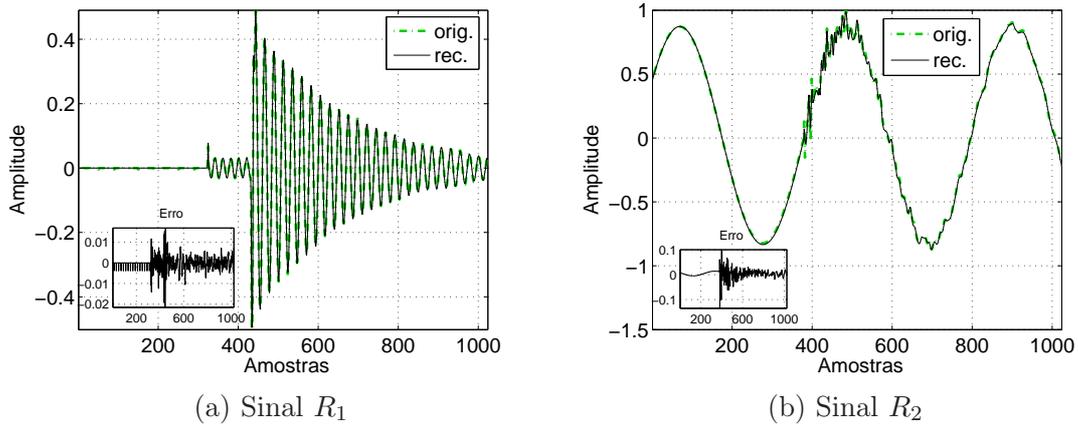


Figura 4.3: Sinais elétricos e suas versões reconstruídas com taxas de compressão de 0,95 e 0,578 bit/amostra respectivamente, com seus erros de reconstrução.

É possível verificar através da Figura 4.4 a eficácia do conjunto reduzido de quantizadores utilizado na otimização T-D. Esta figura apresenta as curvas operacionais

de R_1 e R_2 usando ambos os conjuntos de quantizadores \mathbf{b}_k , o reduzido e o completo. A taxa é medida em bits/amostra e a distorção corresponde ao erro médio quadrático, dado pela equação (4.6). Note que, para ambos os sinais, as curvas operacionais através do uso do conjunto reduzido apresentam boa aproximação com relação às curvas obtidas usando o conjunto completo. Isso demonstra a efetividade do estágio de treinamento, que possibilitou que se encontrasse um conjunto de dicionários (especificados por \mathbf{b}_k) com bom desempenho quando aplicado a sinais elétricos, permitindo a implementação de um esquema de compressão com baixa complexidade computacional que também é capaz de alcançar um desempenho quase ótimo em termos de taxa-distorção.

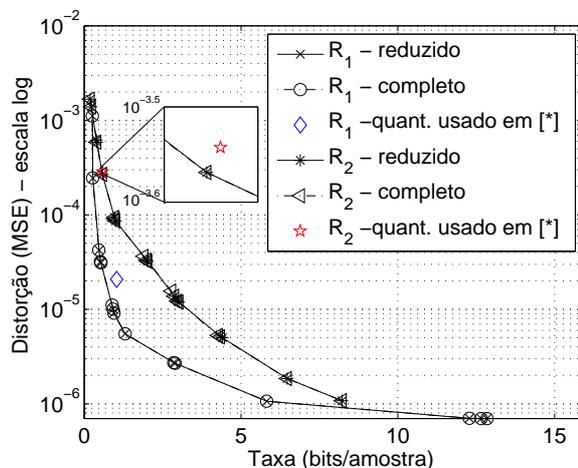


Figura 4.4: Curvas operacionais obtidas utilizando-se os conjuntos completo e reduzido de quantizadores para os sinais R_1 e R_2 . ($[*] = [4]$)

4.5 Avaliação de Desempenho do Sistema de Compressão

Os registradores digitais de perturbações (RDP) têm sido amplamente usados em sistemas elétricos modernos a fim de melhor compreender as características das faltas ocorridas, bem como para avaliação da qualidade de energia. Dessa maneira, quando formas de onda de perturbações elétricas são comprimidas, é de fundamental importância verificar se a perda de informação introduzida pela compressão resulta em erro de interpretação por parte de engenheiros ou sistemas especialistas em análise de faltas. Vale destacar que, considerando esta questão, os requisitos de avaliação para a compressão de sinais elétricos apresentam características bem diferentes dos utilizados na compressão de imagens ou sinais sonoros. Quando se avalia o desempenho de métodos de compressão aplicados a imagens e áudio, interessa saber a quantidade de informação que pode ser eliminada do sinal original de tal forma que

o olho/cérebro ou ouvido/cérebro não perceba a informação removida [28, 94]. Os oscilogramas requerem uma abordagem diferente, posto que eles são normalmente processados por ferramentas que auxiliam o engenheiro especialista na análise das causas dos distúrbios elétricos.

Neste trabalho, avalia-se o desempenho do método de compressão proposto utilizando-se um sistema de análise automática desenvolvido no CEPTEL (Centro de Pesquisas de Energia Elétrica) com ênfase na análise de faltas em linhas de transmissão de alta tensão [26, 27, 95]. Sua implementação tem como base a experiência de vários engenheiros especialistas em análise de distúrbios elétricos correspondentes à atuação de sistemas de proteção. Em geral, essa análise envolve a detecção do início da falta, a classificação da falta, a localização da falta, a análise de eventos e a análise da operação de sistemas de proteção e de disjuntores. O sistema de análise automática gera diversos resultados que caracterizam as faltas; contudo, visando a maior concisão e clareza, somente algumas figuras de mérito foram utilizadas: o momento de início da falta, os fasores calculados a partir dos sinais de tensão e corrente, o tipo de falta e a distância a que a falta ocorreu.

4.5.1 Análise de Faltas

A análise de faltas é uma tarefa imprescindível em empresas de transmissão e distribuição de energia para se manter a confiabilidade de seus sistemas. Através dela é possível obter informações a respeito do comportamento do sistema elétricos e dos equipamentos antes e durante a falta. Isto possibilita apontar inúmeros problemas e assim poder corrigi-los a tempo, de forma correta e adequada. Embora os registros de RDPs não sejam a única fonte de informação usada na análise de faltas, eles são de grande importância, pois permitem [26]:

- a compreensão da falta, incluindo o comportamento dos equipamentos do sistema elétrico de potência, especialmente disjuntores e relés de proteção;
- a localização da falta em linhas de transmissão;
- a medição das condições de pré-falta do sistema, que pode auxiliar no entendimento das causas da falta;
- a compreensão do comportamento pós-falta do sistema, que revela a resposta dos equipamentos do sistema com relação à falta;
- a detecção de problemas incipientes, como degeneração de transformadores de corrente (TC), que podem levar a um mau funcionamento da proteção;

- a detecção de problemas de montagem e configuração do sistemas, tais como contatos folgados, resultando em canais desconectados ou ruidosos, ou escalonamento errado (entre os níveis primário e secundário);
- o levantamento de estatísticas de faltas e taxa de falhas de atuação dos equipamentos.

4.5.2 Sistema de Análise Automática de Faltas

O grande número de instalações e de RDPs distribuídos pela complexa malha interconectada que caracteriza o sistema elétrico brasileiro tem tornado árdua a tarefa de análise de faltas, visto que é impossível analisar a enorme quantidade de dados oscilográficos produzidos simultaneamente. Por essa razão, a pesquisa e desenvolvimento de ferramentas de análise automática de faltas tem se intensificado [96]. O objetivo é explorar a capacidade de processamento de computadores modernos, a fim de facilitar a tarefa dos especialistas em análise de faltas. Desta forma, eles podem reduzir o tempo que o especialista gasta realizando tarefas repetitivas. Neste trabalho, empregam-se tais procedimentos automáticos de análise para extrair características de faltas com o propósito de avaliar o método de compressão proposto. É importante destacar que os algoritmos empregados são ferramentas tradicionais, vastamente utilizadas por especialistas em análise de sistemas elétricos [48, 97–99]. Além disso, enfatiza-se a análise de sinais provenientes do monitoramento de linhas de transmissão. A seguir, são descritos os algoritmos usados no cálculo das figuras de mérito usadas na avaliação de desempenho do sistema de compressão.

Cálculo dos Fasores de Tensão e Corrente

As tensões e correntes em regime permanente de sistemas AC são normalmente representadas através de fasores. O fasor representa uma forma de onda senoidal de frequência ω_0 , módulo $|v(t)|$ e ângulo $\phi_v(t)$ variantes no tempo; no caso particular de

$$v(t) = A \cos(\omega_0 t + \varphi_0), \quad (4.8)$$

têm-se $|v(t)| = A$ e $\phi_v(t) = \varphi_0$. Os fasores são geralmente medidos na frequência fundamental, e são essenciais para a análise do comportamento de sistemas de controle e proteção [100].

As perturbações elétricas podem introduzir oscilações dinâmicas no sistema, harmônicos, transitórios. Dessa maneira, as formas de onda de tensão e corrente não podem ser adequadamente modeladas pela equação (4.8), ou seja, a amplitude A e o ângulo φ_0 podem variar no tempo. Existem diversas técnicas que objetivam obter, em tais casos, uma representação aceitável dos sinais de tensão e corrente na

forma fasorial [97]. A técnica mais amplamente utilizada se baseia em filtro de Fourier, que filtra a componente fundamental antes do cálculo do fasor [48]. Utilizam-se janelas retangulares deslizantes de cossenos e senos para calcular os valores do fasor a cada amostra k do sinal, como descrito abaixo para a implementação do filtro de Fourier de um ciclo

$$\begin{aligned}
 Y_C(k) &= \frac{2}{N_{sc}} \sum_{l=0}^{N_{sc}-1} v(k+l-N_{sc}+1) \cdot \cos\left(\frac{2\pi}{N_{sc}}l\right), \\
 Y_S(k) &= \frac{2}{N_{sc}} \sum_{l=0}^{N_{sc}-1} v(k+l-N_{sc}+1) \cdot \sen\left(\frac{2\pi}{N_{sc}}l\right), \\
 |v(k)| &= \sqrt{Y_C^2(k) + Y_S^2(k)}, \quad \phi_v(k) = -\arctg\frac{Y_S(k)}{Y_C(k)},
 \end{aligned} \tag{4.9}$$

onde N_{sc} representa o número de amostras por ciclo da frequência fundamental. N_{sc} pode ser escrito como a razão $N_{sc} = \frac{F_s}{F}$, em que F_s e F são as frequências de amostragem e fundamental, respectivamente. A Figura 4.5 mostra um sinal de tensão e seus correspondentes módulo e ângulo calculados utilizando-se o filtro de Fourier. Note que para o cálculo do fasor, o ângulo é mantido constante, compensando-se os incrementos de fase devido ao deslocamento da janela. Esta compensação é realizada multiplicando-se $Y_C(k)$ e $Y_S(k)$ na equação (4.9) por $e^{j\frac{2\pi}{N_{sc}}k}$. Detalhes sobre as implementações de meio ciclo e dois ciclos do filtro podem ser encontrados em [48].

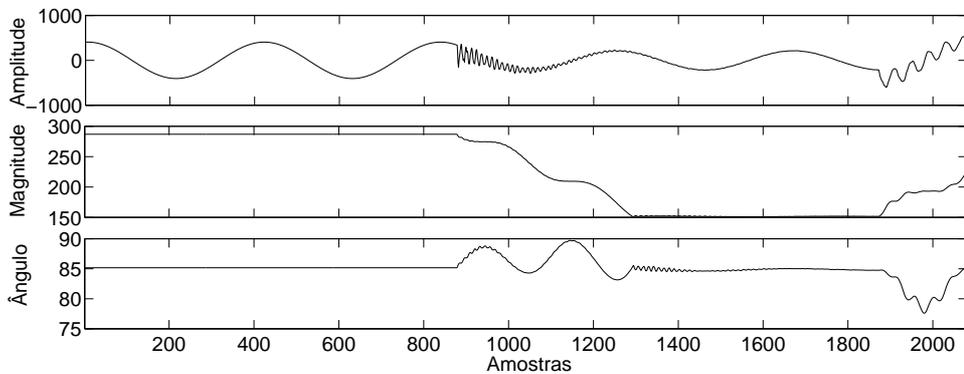


Figura 4.5: Um sinal de tensão com módulo e ângulo calculados usando filtro de Fourier.

Outros fenômenos também podem aparecer em sinais de tensão e corrente tais como exponenciais decrescentes sobrepostas, componentes sub-síncronas, ruído de chaveamento, desvio na frequência fundamental do sistema de potência e saturação do TC. Estes fenômenos acarretam o cálculo impreciso do fasor, mesmo utilizando-se o filtro de Fourier. Foi visto uma aplicação do método de decomposição MPSA discutido anteriormente na filtragem das componentes exponenciais, reduzindo a falta de precisão no cálculo do fasor [4, 42]. Entretanto, o objetivo deste trabalho

é avaliar o sistema de compressão baseado em MPSA. Dada a vasta utilização da abordagem de cálculo fasorial com filtro de Fourier, considera-se adequado usá-la para avaliar o desempenho do sistema de compressão proposto.

Tempo de Início da Falta

O algoritmo do sistema de análise automática para a identificação do tempo de início da falta se baseia em detectar uma variação no módulo do fasor durante a pré-falta e no período de falta. Em alguns momentos, um ruído impulsivo pode produzir pequenas variações de módulo, resultando em uma identificação precoce do início da falta. Para minimizar o efeito deste ruído, considera-se um índice de variabilidade que é calculado utilizando-se duas janelas deslizantes: as janelas de referência e de dados, como ilustra a Figura 4.6. Dessa forma, o ruído impulsivo e outros fenômenos rápidos são filtrados. A transição ao estado faltoso é considerada somente se os valores de módulo do fasor forem suficientemente distintos nessas janelas. O índice de variabilidade é definido como

$$I_{var} = \frac{|\bar{v}_p - \bar{v}|}{\bar{v}}, \quad (4.10)$$

onde \bar{v}_p e \bar{v} são os módulos médios calculados utilizando-se a janela de dados e a janela de referência, respectivamente. O tempo de início da falta é detectado quando o índice de variabilidade é maior que um determinado limiar. Os comprimentos das janelas são parâmetros de entrada do algoritmo. A Figura 4.6 ilustra o comportamento do módulo do fasor e seu índice de variabilidade correspondente.

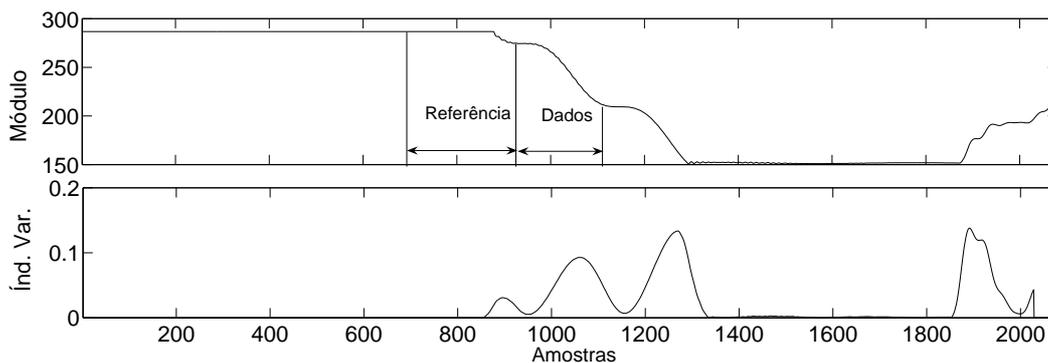


Figura 4.6: Comportamento do módulo da corrente (acima) e seu índice de variabilidade correspondente (abaixo).

Classificação da Falta

A classificação de falta é uma questão importante no processo de compreensão da falta. Também é peça-chave quando a localização de falta em linhas de transmissão

é realizada com base na informação de somente um terminal, como explicado mais à frente.

Para a classificação de falta, os valores fasoriais no período de pré-falta são comparados aos do período de falta, a fim de determinar se a fase está ou não envolvida na falta. Os valores de pré-falta são obtidos a partir do segmento de pré-falta do oscilograma usando um filtro de *Fourier* de um ciclo, terminando meio ciclo antes do início da falta. Os valores de falta são calculados da mesma forma, contudo utilizando a informação do segmento de falta, posicionado a 1,75 ciclo depois do ponto de início da falta [95].

O algoritmo de classificação de falta é descrito da seguinte forma [95]:

1. Se qualquer fasor de corrente possui módulo maior que seu valor em regime permanente por uma constante (>1), então a fase é considerada em falta. A constante é determinada empiricamente, e o valor 1,4 fornece bons resultados.
2. Se mais que uma fase é considerada em falta pelo critério acima, a razão entre os módulos de corrente de distintas fases é analisada. Se a razão é maior que um certo valor, então somente a maior fase é considerada faltosa. O valor 2,5 fornece bons resultados.
3. Caso contrário, se nenhuma fase é considerada em falta pelos teste anteriores, os módulos dos fasores de tensão são analisados. Se o módulo de qualquer fase de tensão é menor que 90% do seu módulo em regime permanente, então ela é considerada uma fase faltosa.
4. Para determinar se a falta envolve terra, as componentes simétricas das correntes são calculadas. Se a razão entre a sequência zero e a sequência positiva das correntes na falta é maior que a razão entre a sequência zero e a sequência positiva das correntes na pré-falta, então considera-se que o terra está envolvido.

Os valores das constantes acima foram definidos empiricamente através de testes massivos aplicando-se este algoritmo em um conjunto grande de oscilogramas reais provenientes do sistema elétrico brasileiro [95].

Localização da Falta

Para a confiabilidade da operação do sistema elétrico de potência, é de suma importância a estimação correta da localização de falta. Em casos de falta permanente, ela permite que a equipe de manutenção seja rapidamente solicitada a se dirigir à seção da linha onde ocorreu a falta para restabelecer o sistema. Para faltas não permanentes, a localização de falta pode ajudar a encontrar a seção da linha onde

um problema intermitente ocorre, devido, por exemplo, a árvores ou fogo abaixo da linha de transmissão.

O algoritmo de localização de falta adotado tem como base as referências [98] e [99]. Os fasores dos segmentos de pré-falta e falta são determinados na etapa de classificação de falta, como visto na Seção 4.5.2. Com isso, distância da falta é calculada como descrito a seguir

- Faltas fase-terra:

$$d = \frac{\text{Im}(V_F \times I_F^*)}{\text{Im} \left[\sum_{k=A,B,C} (Z_{Fk} \times I_k^*) \times I_F^* \right]}, \quad (4.11)$$

onde $F \in \{A, B, C\}$ é a fase faltosa, $\text{Im}(x)$ é a parte imaginária de x , V_F e I_F são os fasores de tensão e corrente, respectivamente, e Z_{Fk} é a impedância entre as fases F e k .

- Faltas fase-fase e faltas trifásicas:

$$d = \frac{\text{Im} \{ (V_{F1} - V_{F2}) \times (I_{F1} - I_{F2})^* \}}{\text{Im} \left[\sum_{k=A,B,C} (Z_{F1k} - Z_{F2k}) \times I_k \times (I_{F1} - I_{F2})^* \right]}, \quad (4.12)$$

onde para faltas fase-fase $F1, F2 \in \{A, B, C\}$, e $F1 \neq F2$ são as fases em falta; e para faltas trifásicas, $F1$ e $F2$ podem ser quaisquer fases, desde que $F1 \neq F2$.

Nestas equações, Z_{Fi} corresponde a cada elemento da matriz de impedância da linha de transmissão, onde $i \in \{1, 2\}$; e I_k representa a corrente de pré-falta da fase $k \in \{A, B, C\}$. Note que este algoritmo é altamente sensível ao cálculo de fasores e erros na classificação do tipo de falta.

Imprecisões no cálculo fasorial e na medição dos parâmetros de linha, transposição não uniforme das fases, clima e condições do solo são fatores que podem acarretar uma imprecisa localização de falta através deste algoritmo. No entanto, o propósito deste trabalho não é avaliar o desempenho do algoritmo de localização de falta. O objetivo é utilizar esta ferramenta tradicional para avaliar o método de compressão aplicado a sinais elétricos proposto neste trabalho.

4.5.3 Avaliação do Sistema de Compressão

O sistema de compressão proposto, descrito na Seção 4.2 (página 47), é avaliado utilizando-se as ferramentas do sistema de análise automática anteriormente descritas. Neste processo de avaliação, foram utilizados 18 oscilogramas que registram faltas em linhas de transmissão de uma empresa de distribuição de energia, cada um

contendo de 6 a 8 canais analógicos (os sinais de tensões e correntes em circuitos trifásicos mais os canais neutros). As formas de onda são armazenadas em formato IEEE COMTRADE [101]. O conjunto de oscilogramas é composto por distúrbios em situações distintas, contendo fenômenos tais como subtensões, sobrecorrentes, harmônicos, transitórios e chaveamentos de circuitos. As amostras dos sinais são originalmente representadas usando 16 bits, e a frequência fundamental do sistema elétrico em questão é de 60 Hz. As taxas de amostragem dos arquivos deste conjunto são 1200, 2000, 2040 e 3840 Hz, e o número de amostras varia de 378 a 1991. O sistema de análise automática é aplicado a esses oscilogramas, bem como a suas versões comprimidas a diferentes taxas de compressão. Reunindo todas as possíveis versões codificadas desses 18 arquivos, obtém-se um conjunto completo de 321 arquivos. Os resultados de análise dos arquivos comprimidos são comparados com os obtidos na análise dos arquivos não comprimidos.

Influência Sobre a Detecção do Início de Falta

Inicialmente, avalia-se o efeito da compressão na detecção do tempo de início da falta, o ponto inicial do processo de análise de falta. Uma correta identificação do início da falta é importante para a sua classificação de tipo e, por consequência, para a sua localização na linha de transmissão. Calcula-se o erro absoluto, representado em milissegundos, entre os instantes de tempo para os oscilogramas comprimidos e não comprimidos; os resultados são apresentados na Figura 4.7. A Figura 4.7.(a) exhibe a média dos erros absolutos de todos os oscilogramas para diferentes intervalos de taxas de compressão. Pode-se observar que o erros absolutos médios não ultrapassam o valor de 3,5 ms, que é considerado pequeno quando comparado ao tempo de duração dos eventos de faltas tais como operação de relés de proteção ou abertura de disjuntores. Com o propósito de obter uma análise mais detalhada dos intervalos de taxas de bits com os maiores valores de erro absoluto médio, calcula-se a função distribuição acumulada (FDA: $F_X(x) = P(X \leq x)$) dos erros absolutos dos oscilogramas comprimidos com 0 a 1 bit/amostra na Figura 4.7.(a), visualizada graficamente através da Figura 4.7.(b). Observa-se que embora o máximo erro absoluto alcance o valor de 80 ms, aproximadamente 97% dos erros absolutos são menores que 3,5 ms (indicado pela linha tracejada). Esta alta porcentagem de erro de pequena magnitude demonstra que o sistema de compressão proposto, mesmo operando a baixas taxas de bits, preserva a detecção de início de falta. Uma análise detalhada mostra que para taxas acima de 0,7 bit por amostra, o erro absoluto na identificação do início da falta não excede 3,5 ms; logo, para evitar erros significativos de detecção de tempo de início de falta, recomenda-se comprimir os oscilogramas a taxas acima de 0,7 bit/amostra.

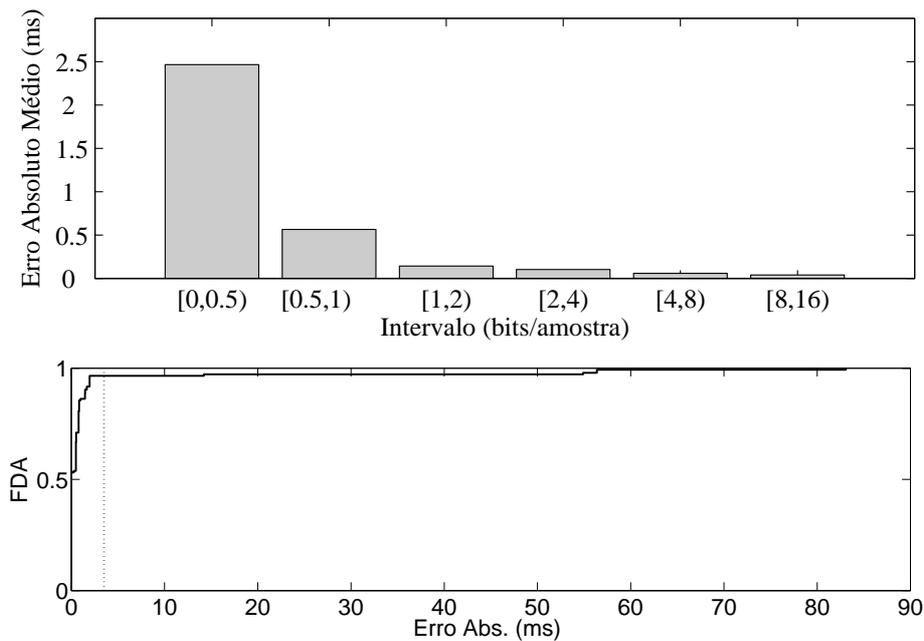


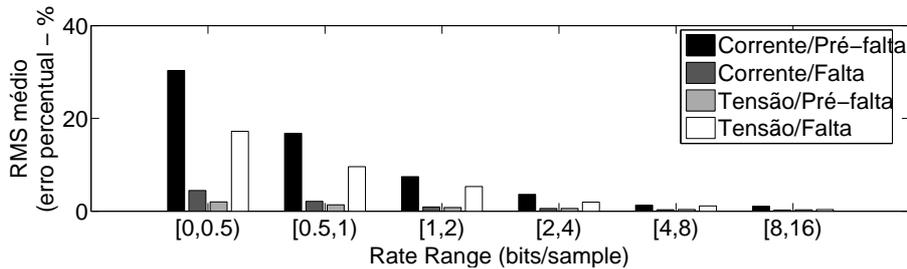
Figura 4.7: Detecção do tempo de início de falta. Acima: erro absoluto médio versus intervalos de taxas de bits; Abaixo: FDA dos erros absolutos na detecção do início da falta abaixo de 1 bit/amostra. A linha tracejada indica 3,5 ms.

Influência Sobre o Cálculo dos Fasores

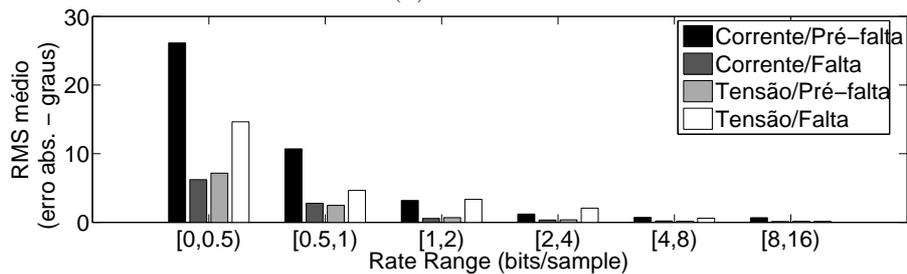
Como previamente discutido, o cálculo correto do módulo e do ângulo dos fasores de tensão e corrente é de grande importância. Dessa maneira, o efeito da distorção introduzida pelo sistema de compressão sobre o cálculo dos fasores deve ser investigado. Os módulos e ângulos são calculados alguns ciclos da frequência fundamental antes e depois do momento de ocorrência da falta, e são usados como parâmetros de entrada para a classificação da falta e o cálculo da distância da falta. Dado que estas grandezas dependem da detecção correta do instante de início da falta, os oscilogramas comprimidos com erros absolutos inaceitáveis (valores maiores que 3,5 ms) não são considerados.

Primeiramente, calculam-se os erros percentuais dos módulos dos fasores de tensão e corrente entre os arquivos não comprimidos e comprimidos a diferentes taxas de compressão. Em seguida, calculam-se as médias, empregando diferentes faixas de taxas, dos valores médios quadráticos (RMS: $x_{\text{rms}} = \sqrt{\frac{1}{n} \sum_i^n x_i^2}$) destes erros percentuais, como ilustrado na Figura 4.8.(a). A Figura 4.8.(b) apresenta o RMS médio, para os diferentes oscilogramas do conjunto analisado, dos erros absolutos entre os ângulos dos arquivos comprimidos e não comprimidos, em função dos intervalos de taxa de compressão. Ambas as figuras mostram, separadamente, os valores RMS dos erros para os sinais de corrente e de tensão sendo que para cada tipo de sinal considera-se a divisão entre os períodos de pré-falta e falta. A partir destes

resultados, nota-se que para sinais de corrente a distorção introduzida na pré-falta é maior que na falta, e o oposto ocorre para os sinais de tensão, particularmente a baixas taxas. Isso ocorre porque, em geral, as faltas produzem sobrecorrente e subtensão. Dessa forma, para sinais de corrente a energia é mais concentrada no período de falta, enquanto que para sinais de tensão, a energia se concentra na pré-falta. À medida que o nível de compressão se eleva, para cada oscilograma do sinal, as componentes de baixa energia são descartadas antes das de maior energia. Portanto, o sistema de compressão proposto, baseado em MPSA, tende a apresentar um efeito mais severo em sinais de corrente no período de pré-falta e em sinais de tensão no período de falta.



(a) Módulo



(b) Ângulo

Figura 4.8: Cálculo de módulo e ângulo dos fasores de corrente e tensão nos períodos de pré-falta e falta: (a) RMS médio dos erros percentuais do módulo versus intervalos de taxas de bits, (b) e RMS médio dos erros absolutos de ângulo versus intervalos de taxas de bits.

A análise do efeito da distorção, introduzida pela codificação, no cálculo do fasor não é suficiente para se determinar a intensidade de compressão tolerável, posto que para isso seria necessário saber quanto de distorção no cálculo do módulo e ângulo do fasor é aceitável. Para se obter esse nível de compressão tolerável sem que haja uma interpretação errada dos resultados por parte dos especialistas, outras medições também são realizadas [26, 27]. Por exemplo, os fasores calculados alguns ciclos antes e depois do início da falta são utilizados na classificação e localização da falta. Por isso, decidiu-se avaliar este método de compressão tendo em vista a sua capacidade de não comprometer os resultados da análise de faltas dos sinais comprimidos. O seguinte procedimento é empregado: primeiro, a falta é classificada usando o algoritmo descrito na Seção 4.5.2; uma vez classificada corretamente, calcula-se a

distância da falta com base no algoritmo da Seção 4.5.2.

Influência Sobre a Classificação de Falta

Compara-se o casamento entre o tipo de falta obtido na classificação de oscilogramas comprimidos e não comprimidos. A Figura 4.9.(a) apresenta os resultados desta avaliação, indicando quais taxas de compressão estão associadas à classificação errônea (0 - correto; 1 - incorreto) A Figura 4.9.(b) mostra a probabilidade de erro de classificação da falta em diferentes intervalos de taxas de bits. Para taxas acima de 1,13 bit por amostra não há erro de classificação. Portanto, arquivos de RDPs devem ser comprimidos acima desta taxa ao se utilizar este codificador. Na Seção 4.5.3, verificou-se que uma taxa maior que 0,7 bit/amostra é necessária para obter pequenos erros de detecção de início de falta. Dessa maneira, para o conjunto de sinais considerado aqui, uma taxa de codificação maior que 1,13 bit/amostra preserva o resultado de classificação da falta com relação ao arquivo não comprimido e produz erros desprezíveis na identificação do tempo de início de falta.

Influência Sobre a Localização de Falta

Como discutido anteriormente, uma importante função dos oscilogramas é seu uso na localização de falta. Nesta seção, se estabelece o nível de compressão que pode ser aplicado aos oscilogramas usando este sistema de compressão sem comprometer significativamente o resultado da localização de falta.

É importante notar que a acurácia na classificação de falta é crucial para o cálculo da distância de falta. Portanto, quando o resultado da classificação de falta do sinal comprimido é diferente do resultado da classificação do sinal não comprimido, torna-se inútil comparar os resultados da localização de falta entre eles. Por isso, o desempenho do codificador MPSA é avaliado considerando a localização de falta, dado que o tipo de falta identificado no arquivo comprimido é o mesmo identificado no arquivo não comprimido. A Figura 4.10.(a) apresenta os erros absolutos entre as distâncias calculadas para os sinais comprimidos e não comprimidos. As distâncias são calculadas em percentual do comprimento da linha de transmissão. Com base nestes resultados observa-se, por exemplo, que se erros absolutos menores que 4% sejam aceitáveis, os sinais podem ser comprimidos usando taxas acima de 2,5 bits/amostra, como indicado pela linha tracejada na Figura 4.10.(a).

A Figura 4.10.(b) exhibe o erro absoluto médio no cálculo da distância da falta para diferentes intervalos de taxas de compressão, e a Tabela 4.1 mostra os erros absolutos máximos para cada intervalo, suas respectivas taxas de bits e razões de compressão (a taxa do arquivo original—16 bits/amostra—dividida pela taxa de bits do arquivo comprimido). Comparando os dados da Figura 4.10.(b) com os dados

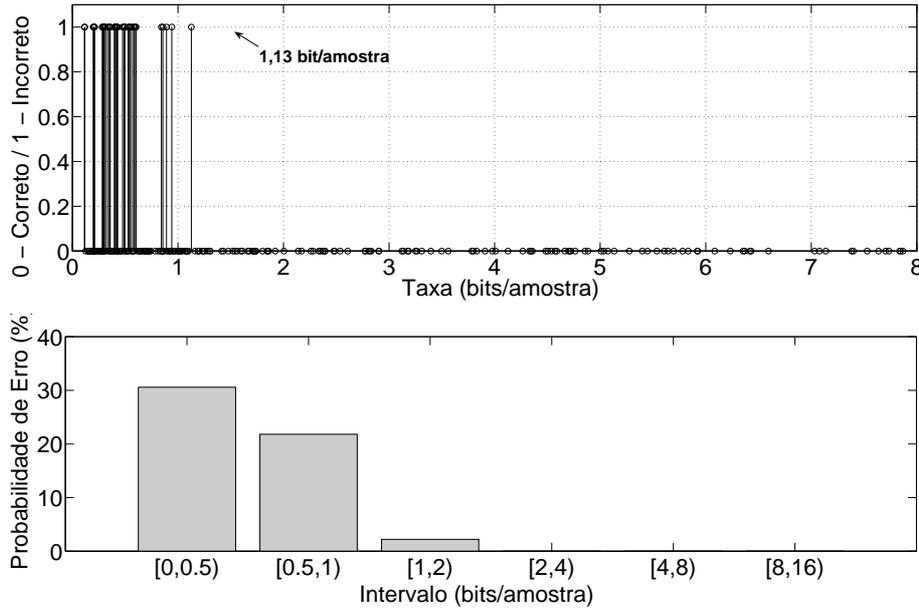
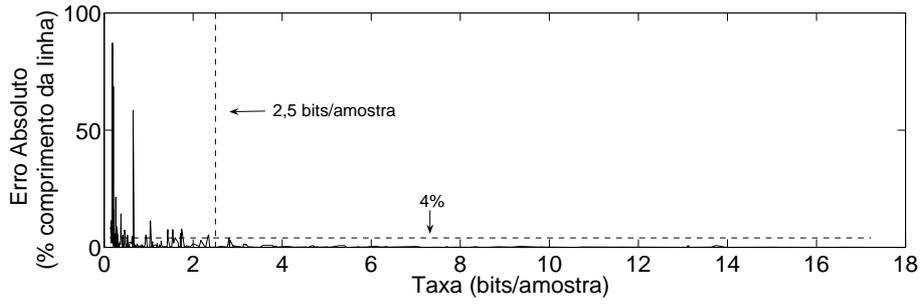


Figura 4.9: Classificação de falta. Acima: indica quais taxas de bits resultam em erro de classificação; Abaixo: mostra a probabilidade de erro de classificação da falta em diferentes intervalos de taxas de bits.

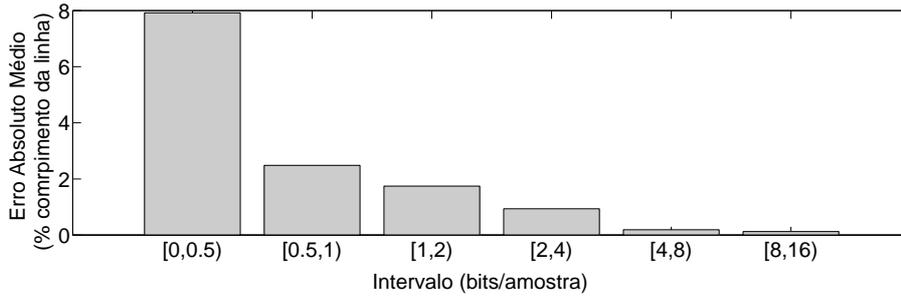
da Tabela 4.1, nota-se que o erro absoluto médio é consideravelmente menor que o erro absoluto máximo, indicando que erros próximos aos máximos não são frequentes, especialmente para intervalos de baixas taxas de bits. Deve-se destacar que o máximo erro absoluto de 11,3% no intervalo de 1 a 2 bits/amostra ainda é bastante significativo. Contudo, observando-se a FDA dos erros absolutos correspondentes a este intervalo, apresentada na Figura 4.11, verifica-se que aproximadamente 80% dos erros são menores que 4%. Dessa maneira, neste intervalo o sistema de compressão ainda pode ser considerado confiável. É importante mencionar que o erro de localização da falta aceitável depende do comprimento da linha de transmissão; para linhas curtas podem-se aceitar erros maiores que 4%.

Considerando os resultados de forma geral, observa-se que o nível de compressão que pode ser empregado ao utilizar o sistema de compressão proposto neste trabalho é determinado pelo procedimento de localização de falta. Em outras palavras, a taxa de compressão necessária para se obter um erro tolerável em localização de falta é mais do que suficiente para assegurar tanto erros desprezíveis na identificação do início da falta quanto uma classificação de falta bem acurada sem erros. No entanto, convém lembrar que a quantidade de erro tolerável na localização da falta depende do comprimento da linha de transmissão.

Concluindo, para o conjunto de oscilogramas de teste usado neste trabalho e considerando o erro aceitável de 4% no cálculo da distância da falta, uma taxa de 2,5 bits/amostra pode ser empregada. Isso indica que o sistema de compressão desenvolvido aqui é capaz de prover alta razão de compressão (6,4:1) com consi-



(a)



(b)

Figura 4.10: Localização de faltas: (a) erros absoluto versus taxa de bits, (b) erro absoluto médio versus intervalos de taxas de bits.

Tabela 4.1: Taxa de bits e razão de compressão para distintos erros de localização.

Intervalo(b/a)	Erro máximo (%)	Taxa (b/a)	Razão de compressão
[0; 0,5)	87,21	0,19	84,21
[0,5; 1)	58,50	0,65	24,61
[1; 2)	11,30	1,04	15,38
[2; 4)	5,34	2,35	6,80
[4; 8)	0,74	4,69	3,40
[8; 16)	0,77	13,11	1,20

derável confiabilidade e robustez, sem comprometer o processo de análise falta dos oscilogramas comprimidos.

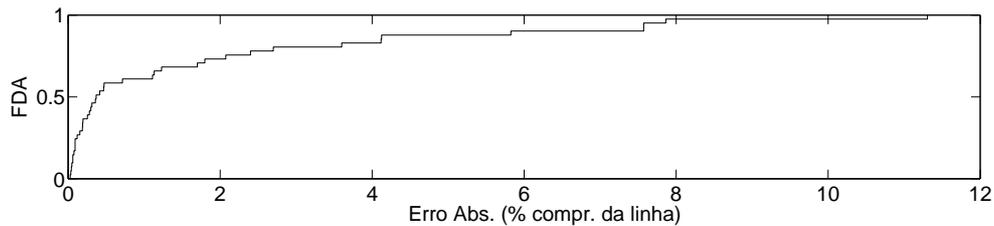


Figura 4.11: FDA dos erros absolutos no intervalo 1 a 2 bits/amostra em relação à localização de falta.

Capítulo 5

Decomposição Atômica de Sinais de Áudio

Neste capítulo, descrevem-se inicialmente trabalhos encontrados na literatura que utilizam o *Matching Pursuit* para a decomposição atômica de sinais de áudio. Em seguida, apresenta-se um novo método de decomposição atômica aplicado a sinais de áudio, o MPSACr, fundamentado no MPSA (ver Capítulo 3). Ele se baseia em um dicionário composto por senoides amortecidas de frequências na escala musical cromática. Nesse método introduz-se uma nova estratégia de redução de complexidade computacional de maneira a não restringir o espaçamento frequencial no projeto do dicionário. Além disso, apresenta-se também uma heurística de continuidade de átomos entre quadros consecutivos. Por fim, são apresentadas simulações computacionais com o método de decomposição proposto.

5.1 *Matching Pursuit* Aplicado a Sinais de Áudio

Na literatura, existem diversos trabalhos em que se utiliza o algoritmo de *Matching Pursuit* para a análise e representação de sinais de áudio para fins de transcrição [33, 82], codificação [14, 32, 102–108], codificação de áudio espacial [109], separação de fontes [110], reconhecimento de áudio (*audio retrieval*) [111] e modelagem psicoacústica [34, 112].

Em geral, as abordagens propostas têm em comum a utilização do MP como mecanismo principal para a obtenção da representação atômica de sinais de áudio, e o que as diferencia é o dicionário usado para modelar os fenômenos neles presentes. É possível encontrar uma grande variedade de dicionários, todos criteriosamente escolhidos e projetados com o intuito de obter representações compactas e coerentes.

O projeto do dicionário deve levar em consideração as características intrínsecas de um sinal de áudio, como seu comportamento variante no tempo e conteúdo

espectral. Para sinais musicais, é interessante que o dicionário ao menos apresente componentes com frequências associadas às notas musicais de um escala cromática igualmente temperada (utilizada no Ocidente) que são espaçadas geometricamente com razão $2^{\frac{1}{12}}$. Além disso, existem os harmônicos (múltiplos inteiros) associados a cada frequência fundamental. Na literatura, verifica-se o uso de dicionários compostos por componentes cujas frequências são espaçadas geometricamente de acordo com a escala musical cromática [82], ou por componentes cujas frequências são espaçadas linearmente com o objetivo de representar os harmônicos [33, 113]. Observa-se que um dicionário com espaçamento frequencial linear não possui elementos para representar corretamente componentes de frequência da escala cromática, bem como um dicionário com espaçamento frequencial geométrico não apresenta elementos para representar corretamente componentes harmônicas não presentes na escala cromática.

Outra característica do projeto de dicionários está associada à escolha da janela utilizada. Ao percorrer os trabalhos, verifica-se o uso da janela gaussiana (Gabor) [18, 33], da janela de Hanning [82] e da janela exponencial [14, 114], entre outras [105, 106]. Outra abordagem que também proporciona flexibilidade na representação consiste em compor o dicionário de *wavelets* [32].

Em trabalhos recentes [107, 115, 116], propõe-se o uso de moléculas ou subespaços vetoriais em vez de átomos para a representação de sinais de áudio. A idéia geral é rastrear uma parcial pertencente a uma estrutura harmônica e, em seguida, ajustar uma molécula composta por componentes harmônicas em torno da parcial detectada.

5.2 MPSA de Escala Musical Cromática - MPSACr

Neste trabalho, desenvolve-se um método de decomposição de sinais de áudio, o MPSACr, que agrega características do método apresentado em [82] ao algoritmo de decomposição MPSA, que, a princípio, foi desenvolvido e ajustado para a representação de sinais elétricos. Em [82], verifica-se que o uso de um dicionário composto por átomos de frequências da escala musical cromática possibilita boa resolução para análise tempo-frequência de sinais musicais. A fim de obter uma representação mais compacta e coerente com os fenômenos presentes em sinais de áudio, em especial, os sinais musicais, são realizadas modificações no MPSA relacionadas ao projeto do dicionário e às heurísticas existentes no laço de decomposição. É importante destacar o uso de heurísticas para se obter representações mais coerentes com os fenômenos localizados dentro do suporte do sinal como sendo uma característica peculiar do MPSA em relação a métodos de decomposição baseados em MP encontrados na literatura, como visto na Seção 5.1, em que normalmente se utiliza o MP como o algoritmo-núcleo, preocupando-se somente com a modelagem do dicio-

nário [14, 33, 82, 105, 107]. Fenômenos localizados são muito frequentes em sinais variantes no tempo como os sinais elétricos e de áudio.

Na Figura 5.1, é apresentado o diagrama de blocos do MPSACr. Inicialmente, o sinal é segmentado em quadros de iguais comprimentos. À exceção do primeiro quadro, realiza-se a seleção, no quadro anterior, dos átomos fortemente correlacionados com os padrões do quadro corrente para compor seu livro de estruturas. Os parâmetros de amplitude e fase são ajustados para que haja continuidade dos átomos entre quadros consecutivos. Além disso, atualiza-se o resíduo à medida que os átomos continuados são selecionados.

Em seguida, a cada iteração da decomposição, busca-se a melhor aproximação do resíduo em um dicionário finito de senoides amortecidas com frequências pertencentes à escala musical cromática. Na sequência, realiza-se a busca por um melhor suporte temporal e maximiza-se a aproximação entre o átomo e o resíduo, otimizando-se o fator de decaimento através de um algoritmo pseudo-Newton [1, 4].

Por fim, armazenam-se o coeficiente e os parâmetros do átomo no livro de estruturas e verifica-se se é satisfeito o critério de parada, que tem como base a taxa de aproximação. Em caso contrário, o átomo é escalonado e subtraído do resíduo, gerando um novo resíduo para a próxima iteração.

Com o propósito de facilitar a identificação das diferenças entre o MPSACr e o MPSA, o diagrama de blocos do MPSA é novamente ilustrado através da Figura 5.2. A seguir, descrevem-se os detalhes de cada etapa do MPSACr, ressaltando-se as principais modificações que transformam o MPSA em MPSACr.

5.2.1 Segmentação do Sinal

Dado que os sinais de áudio contêm tipicamente um grande número de amostras, é computacionalmente inviável decompor o sinal inteiro utilizando um dicionário de mesma dimensão. Dessa forma, segmenta-se o sinal em quadros de tamanhos menores, analisando-os/decompondo-os um por vez. No MPSACr, o sinal é segmentado utilizando-se uma janela retangular com saltos do mesmo tamanho da janela, ou seja, não se utiliza nenhuma janela de suavização como a de Hanning, o que levaria à necessidade de sobreposição entre quadros consecutivos e ao consequente aumento do número total de átomos da representação. A fim de preservar a identidade dos segmentos a representar no tempo, aqui optou-se pela janela retangular sem sobreposição entre quadros.

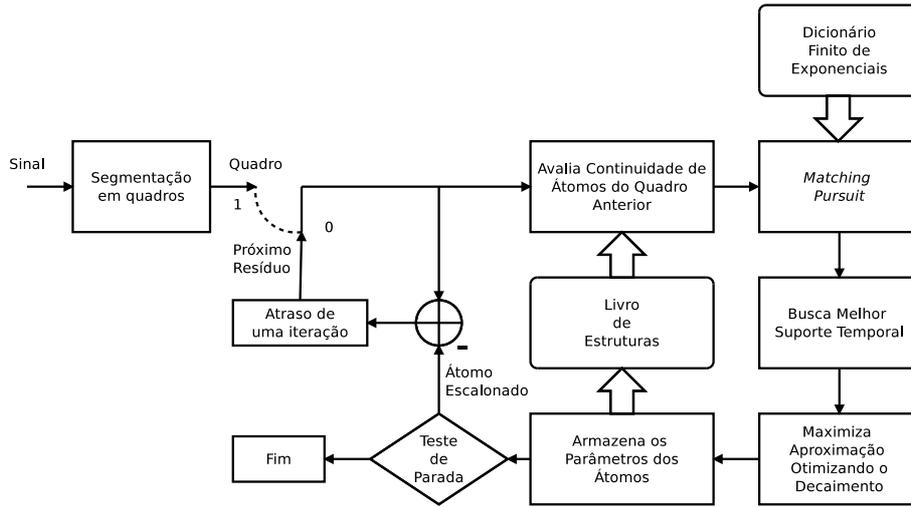


Figura 5.1: Diagrama de blocos do método de decomposição MPSACr aplicado a sinais de áudio. Na primeira iteração, a chave está na posição 1, e nas demais iterações permanece na posição 0.

5.2.2 Dicionário de Escala Musical Cromática

O dicionário usado no MPSACr é o mesmo usado no MPSA, sendo formado por senoides amortecidas. Um átomo g_γ deste dicionário é definido como

$$g_\gamma(n) = K_\gamma g(n) \cos(\xi n + \phi) [u(n - n^s) - u(n - n^e)], \quad (5.1)$$

$$g(k) = \begin{cases} 1, & \text{se } \rho = 0 \quad (\text{Senoide pura (se } \xi \neq 0), \\ & \text{DC ou impulso unitário (se } \xi = 0)) \\ e^{-\rho(n-n^s)}, & \text{se } \rho > 0 \quad (\text{Exponencial decrescente}) \\ e^{\rho(n^e-n)}, & \text{se } \rho < 0 \quad (\text{Exponencial crescente}), \end{cases}$$

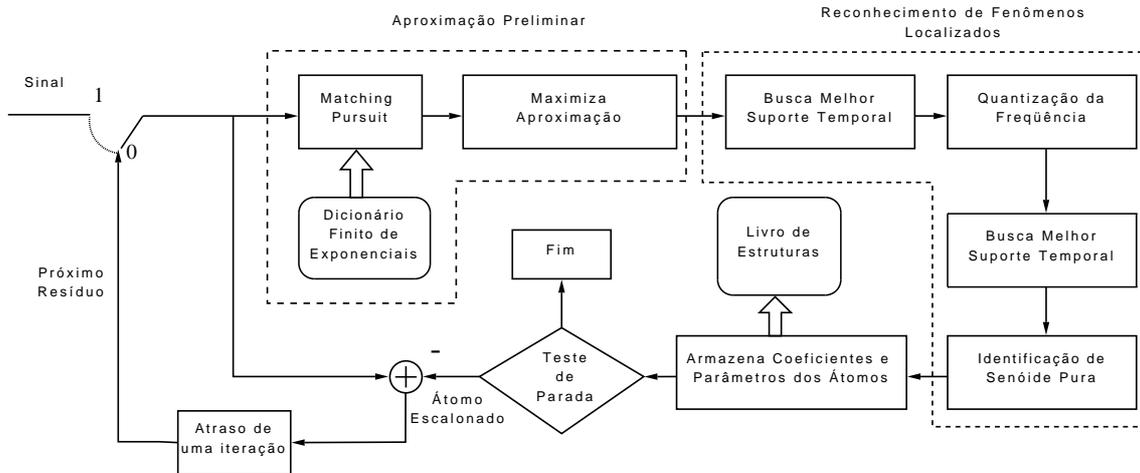


Figura 5.2: Diagrama de blocos do MPSA. Na primeira iteração, a chave está na posição 1, e nas demais iterações permanece na posição 0.

em que $K_\gamma \in \mathbb{R}$ é tal que $\|\mathbf{g}_\gamma\| = 1$ e N é o comprimento do sinal. Ademais, o átomo na equação (5.1) é definido pela quintupla $\gamma = (\rho, \xi, \phi, n^s, n^e)$, em que ρ representa o fator de decaimento, ξ é a frequência, ϕ é a fase, e n^s e n^e são as amostras inicial e final, respectivamente. A fase do átomo é calculada de forma analítica em função dos demais parâmetros, sendo ótima no sentido de prover o máximo produto interno entre o átomo e o resíduo a cada iteração [1, 4, 80].

Contudo, neste caso, as frequências não são espaçadas linearmente, e sim geometricamente como em [82]. A frequência é discretizada da seguinte forma:

$$\xi = \xi_0 2^{\frac{k}{24}}, \quad k \in \{0, \dots, k_{\text{máx}}\}, \quad (5.2)$$

em que k é o índice da frequência correspondente a um quarto de tom da escala musical cromática entre uma frequência mínima ξ_0 e uma frequência máxima associada a $k_{\text{máx}}$. Portanto, as frequências são separadas por intervalos de um quarto de tom, que corresponde à metade da menor divisão formal, o semitom. Desta forma, é possível representar de maneira mais precisa sinais provenientes de instrumentos de afinação imperfeita. O semitom corresponde a um espaçamento geométrico de $2^{\frac{1}{12}}$ na escala musical ocidental. As frequências variam de 7,28 a 15804,26 Hz. O limite inferior se situa à 23 semitons (2 oitavas) abaixo de 27,5 Hz, frequência associada à primeira nota do piano (A0); e o limite superior se situa à 23 semitons (2 oitavas) acima de 4186,01 Hz, frequência associada à última nota do piano (C8).

Em [5], verifica-se que os indivíduos têm um limite natural para discernir a diferença de *pitch*¹ entre dois tons puros (senoides puras) de mesma intensidade quando emitidos um logo após o outro: quando essa diferença ultrapassa a chamada JNDF (*just-noticeable difference in frequency*), ela passa a ser detectada pelo ouvido humano. A JNDF é medida utilizando-se modulação na frequência senoidal. O desvio na frequência, Δf , é definido como a diferença entre a frequência não-modulada ou de portadora, f , e a máxima/mínima frequência. Usando esta definição, a frequência varia entre $f - \Delta f$ e $f + \Delta f$, logo, a variação total é dada por $2\Delta f$, que corresponde à JNDF, apresentada na Figura 5.3 como uma função da frequência de portadora. Essa curva foi obtida considerando a frequência de modulação de 4 Hz e o nível de audibilidade de 60 *phons*², calculando-se valores médios para diferentes indivíduos [5]. Observe que, em baixas frequências, a JNDF é aproximadamente constante e possui um valor em torno de 3,6 Hz. Para frequências acima de 500 Hz, a JNDF aumenta proporcionalmente à frequência de portadora f , valendo aproximadamente $0,007f$. Isto significa que uma variação na frequência de 0,7% é percebida nesta faixa de frequência. Em frequências baixas, a variação relativa de JNDF aumenta

¹O *pitch* é a frequência fundamental percebida pelo sistema auditivo humano [5].

²Unidade de nível de audibilidade [5, 15]

chegando a um quarto de tom (2,9%) na frequência de 120 Hz e um semitom (3,6%) em 50 Hz.

Portanto, para componentes de frequência maiores que 120 Hz seria necessário adotar maior resolução/discretização para a frequência no dicionário. Contudo, existe uma restrição para o aumento da resolução, pois a decomposição via MPSACr se torna mais custosa computacionalmente à medida que a discretização da frequência passa a ser mais refinada, deixando o dicionário mais populoso. Opta-se por uma discretização mais grosseira em troca de menor complexidade computacional.

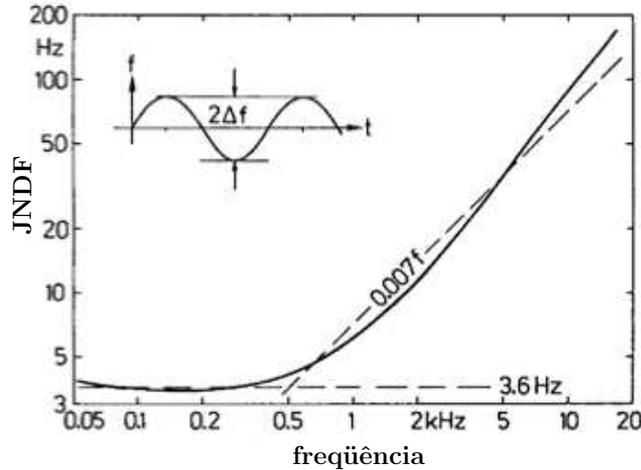


Figura 5.3: JNDF (*just noticeable difference in frequency*) em função da frequência para uma modulação na frequência em 4 Hz. Adaptado de [5]

Deve-se observar que, dependendo da escala ($\frac{1}{\rho}$), não é possível analisar de maneira eficaz componentes de frequências que possuem menos de 1 ciclo dentro da escala. Portanto, para impedir a inserção de um número de elementos maior que o necessário no dicionário, seleciona-se, para uma dada escala, ξ_0 de modo que as frequências apresentem mais que 1 ciclo.

No MPSA, o dicionário discreto é obtido através da discretização dos parâmetros da seguinte forma: $\rho = 2^{-j}$, $j \in [0, \log_2 N]$; $n^s = p2^j$, $p \in [0, N2^{-j})$ e $\xi = k\pi 2^{1-j}$, $k \in [0, 2^{-j}]$, onde N é a dimensão do dicionário. No MPSACr, o decaimento continua sendo discretizado como $s = 2^{-j}$, sendo $j = 0, 1, \dots, \log_2(N)$. Quanto ao deslocamento no tempo, para todos os pares (escala, frequência) faz-se sempre o espaçamento entre os deslocamentos $\Delta u = 1$, ou seja $u = i$, $i = 0, 1, \dots, N - 1$. Dessa forma, o número de átomos do dicionário aumenta significativamente, elevando, assim, o custo computacional; entretanto, com o uso da abordagem proposta para o cálculo rápido dos produtos internos na busca no dicionário finito de exponenciais apresentado a seguir na Seção 5.2.3, reduz-se o impacto deste aumento. Na nova abordagem proposta neste trabalho, para escala e frequência dadas, calculam-se N produtos internos associados aos deslocamentos temporais dos átomos com uma

complexidade $\mathcal{O}(12N \log_2 2N)$ enquanto que na abordagem tradicional verifica-se uma complexidade $\mathcal{O}(N^2)$.

5.2.3 Estratégia de Redução de Complexidade Computacional

Tradicionalmente, o *Matching Pursuit* (MP) é considerado um algoritmo de alto custo computacional para aplicação em problemas reais envolvendo sinais de grande dimensão. A implementação numérica típica deste algoritmo apresenta uma complexidade de $\mathcal{O}(N^2)$ (N é o comprimento do sinal), que leva a tempos de execução muito elevados [117]. Esta seção tem o propósito de discutir e apresentar estratégias para a redução da complexidade computacional do MP, em especial na etapa de cálculo dos produtos internos. É importante ressaltar que existem estratégias na literatura que se aplicam ao MP com base em um dicionário fixo [18, 117], o que não se aplica aqui, dado que existem heurísticas do laço de decomposição e o dicionário é de parâmetros contínuos.

Neste trabalho, desenvolve-se uma forma mais rápida de calcular os produtos internos sem restrição de espaçamento frequencial. Esta estratégia se ajusta bem ao dicionário de escala musical cromática, visto que suas frequências não são linearmente espaçadas.

Algoritmo Básico

O MP realiza aproximações sucessivas de um sinal \mathbf{x} de forma iterativa. Primeiramente, inicia-se o resíduo $\mathbf{r}_x^0 = \mathbf{x}$. A cada iteração i :

1. Calculam-se as correlações entre o resíduo \mathbf{r}_x^{i-1} e cada átomo do dicionário redundante D , utilizando-se produtos internos:

$$\forall \mathbf{g}_\gamma \in D, \langle \mathbf{r}_x^{i-1}, \mathbf{g}_\gamma \rangle; \quad (5.3)$$

2. Busca-se o átomo mais correlacionado com o resíduo, procurando-se o produto interno máximo:

$$\hat{\mathbf{g}}_{\gamma(i)} = \operatorname{argm\acute{a}x}_{\mathbf{g}_\gamma \in D} \left| \langle \mathbf{r}_x^{i-1}, \mathbf{g}_\gamma \rangle \right|; \quad (5.4)$$

3. Subtrai-se o átomo ponderado $\alpha_i \mathbf{g}_{\gamma(i)}$ do resíduo \mathbf{r}_x^{i-1} :

$$\mathbf{r}_x^i = \mathbf{r}_x^{i-1} - \alpha_i \mathbf{g}_{\gamma(i)}; \quad (5.5)$$

4. Se o critério de parada for satisfeito, termina-se a decomposição; em caso contrário, faz-se $i = i + 1$ e retorna-se ao passo 1.

Ao fim da decomposição, obtém-se a seguinte aproximação em M termos do sinal \mathbf{x} :

$$\hat{\mathbf{x}} = \sum_{i=1}^M \alpha_i \mathbf{g}_{\gamma(i)}. \quad (5.6)$$

Complexidade Computacional

Para um dicionário de dimensão N , observa-se que cada produto interno $\langle \mathbf{r}_{\mathbf{x}}^{i-1}, \mathbf{g}_{\gamma} \rangle$ requer N multiplicações e $N - 1$ adições, e em cada iteração calculam-se $\#_D \geq N$ ($\#_D$ corresponde à cardinalidade do dicionário) produtos internos. Portanto, o custo do cálculo dos produtos internos no passo 1 atinge $\mathcal{O}(N^2)$ a cada iteração, tornando o uso do MP inviável para sinais de dimensão grande.

No passo 2, realiza-se, normalmente, uma varredura por todas as correlações obtidas, o que corresponde a uma complexidade computacional de $\mathcal{O}(\#_D)$ por iteração. Dado que são realizadas M iterações, o custo computacional total $\mathcal{O}(M\#_D)$ de implementações clássicas de MP ultrapassa $\mathcal{O}(N^2)$, tornando o uso do MP ainda mais impraticável computacionalmente.

Produto Interno entre o Resíduo e o Átomo Real

Em [80], é apresentada uma proposição para o cálculo do produto interno entre o sinal e um átomo real do dicionário apresentando fase ótima. Como visto na Seção 2.1.2, os elementos de um dicionário redundante podem ser gerados escalonando-se, trasladando-se e modulando-se uma função-janela $g(t)$ contínua no tempo. Aqui, em vez da função contínua $g(t)$, considera-se sua versão discreta $g(n)$, que corresponde à amplitude de $g(t)$ a cada instante nT , onde T representa o período de amostragem. No caso de um dicionário de átomos reais e discretos, cada átomo $g_{\gamma,\phi}(n)$ é definido como

$$g_{\gamma,\phi}(n) = \frac{g\left(\frac{n-\tau}{s}\right) \cos(\xi n + \phi)}{\|g\left(\frac{n-\tau}{s}\right) \cos(\xi n + \phi)\|}, \quad (5.7)$$

em que $\gamma = (s, \tau, \xi)$, s é a escala, τ é o deslocamento, ξ é a frequência e ϕ representa a fase. Em seguida, definem-se os vetores

$$P_{\gamma}(n) = g\left(\frac{n-\tau}{s}\right) \cos(\xi n), \quad Q_{\gamma}(n) = g\left(\frac{n-\tau}{s}\right) \sin(\xi n). \quad (5.8)$$

Então, tem-se

$$\mathbf{g}_{\gamma,\phi} = \frac{\mathbf{P}_{\gamma} \cos \phi - \mathbf{Q}_{\gamma} \sin \phi}{\|\mathbf{P}_{\gamma} \cos \phi - \mathbf{Q}_{\gamma} \sin \phi\|} \quad (5.9)$$

e, portanto, o produto interno é escrito como

$$\langle \mathbf{x}, \mathbf{g}_{\gamma, \phi} \rangle = \frac{\langle \mathbf{x}, \mathbf{P}_{\gamma} \rangle \cos \phi - \langle \mathbf{x}, \mathbf{Q}_{\gamma} \rangle \operatorname{sen} \phi}{\|\mathbf{P}_{\gamma} \cos \phi - \mathbf{Q}_{\gamma} \operatorname{sen} \phi\|}. \quad (5.10)$$

Sejam \mathbf{x} um vetor de sinal, $\gamma = (s, \tau, \xi)$, $\mathbf{P} = \mathbf{P}_{\gamma}$, $\mathbf{Q} = \mathbf{Q}_{\gamma}$ e

$$\begin{aligned} a_1 &= \langle \mathbf{x}, \mathbf{P} \rangle \|\mathbf{Q}\|^2 - \langle \mathbf{x}, \mathbf{Q} \rangle \langle \mathbf{P}, \mathbf{Q} \rangle, \\ b_1 &= \langle \mathbf{x}, \mathbf{Q} \rangle \|\mathbf{P}\|^2 - \langle \mathbf{x}, \mathbf{P} \rangle \langle \mathbf{P}, \mathbf{Q} \rangle. \end{aligned} \quad (5.11)$$

Então, a fase ótima ϕ_0 , para a qual $\max_{\phi \in [0, 2\pi]} |\langle \mathbf{x}, \mathbf{g}_{\gamma, \phi} \rangle|$, é dada por uma das soluções abaixo

1. Se $\xi = 0$, então

$$\phi_0 = 0 \text{ e } \langle \mathbf{x}, \mathbf{g}_{\gamma, \phi} \rangle = \frac{\langle \mathbf{x}, \mathbf{P} \rangle}{\|\mathbf{P}\|}. \quad (5.12)$$

2. Se $\xi \neq 0$ e $a_1 = 0$, então

$$\phi_0 = \frac{\pi}{2} \text{ e } \langle \mathbf{x}, \mathbf{g}_{\gamma, \phi} \rangle = -\frac{\langle \mathbf{x}, \mathbf{Q} \rangle}{\|\mathbf{Q}\|}. \quad (5.13)$$

3. Se $\xi \neq 0$ e $a_1 \neq 0$, então

$$\phi_0 = \operatorname{arctg} \left(-\frac{b_1}{a_1} \right) \text{ e } \langle \mathbf{x}, \mathbf{g}_{\gamma, \phi} \rangle = \frac{\langle \mathbf{x}, \mathbf{P} \rangle a_1 + \langle \mathbf{x}, \mathbf{Q} \rangle b_1}{\|\mathbf{P} a_1 + \mathbf{Q} b_1\|}. \quad (5.14)$$

Observa-se, portanto, que para se obter a projeção $\langle \mathbf{x}, \mathbf{g}_{\gamma, \phi_0} \rangle$ do vetor sinal sobre o átomo de fase ótima são necessários cinco fatores:

$$\langle \mathbf{x}, \mathbf{P} \rangle = \sum_{n=0}^{N-1} x(n) g \left(\frac{n - \tau}{s} \right) \cos(\xi n), \quad (5.15)$$

$$\langle \mathbf{x}, \mathbf{Q} \rangle = \sum_{n=0}^{N-1} x(n) g \left(\frac{n - \tau}{s} \right) \operatorname{sen}(\xi n), \quad (5.16)$$

$$\langle \mathbf{P}, \mathbf{P} \rangle = \sum_{n=0}^{N-1} g^2 \left(\frac{n - \tau}{s} \right) \cos^2(\xi n), \quad (5.17)$$

$$\langle \mathbf{Q}, \mathbf{Q} \rangle = \sum_{n=0}^{N-1} g^2 \left(\frac{n - \tau}{s} \right) \operatorname{sen}^2(\xi n), \quad (5.18)$$

$$\langle \mathbf{Q}, \mathbf{P} \rangle = \sum_{n=0}^{N-1} g^2 \left(\frac{n - \tau}{s} \right) \operatorname{sen}(\xi n) \cos(\xi n). \quad (5.19)$$

Redução de Custo Computacional Proposto em [80] com Restrição de Espaçamento Freqüencial

Podem-se calcular N produtos internos através da transformada rápida de Fourier (FFT - *Fast Fourier Transform*) para N valores de frequência ξ , tornando, assim, o passo 1 do algoritmo de MP mais célere. Para que se aproveite ao máximo o potencial computacional da FFT é necessário que o comprimento N do sinal seja potência de 2, i.e., $N = 2^J$. Desta maneira, os únicos valores permitidos de escala são $s = 2^j$, $j = 1, 2, \dots, J$. O parâmetro de frequência é discretizado da seguinte forma: $\xi = \frac{2\pi k}{N}$, $k = 0, 1, \dots, N - 1$. Os valores de deslocamento τ podem ser definidos livremente, podendo-se utilizar a mesma convenção da implementação original [16, 18]: $\tau = p\frac{s}{2}$, onde $0 \leq p \leq 2N/s$. A metodologia de cálculo rápido dos produtos internos utilizando FFT, proposta em [80], é apresentada a seguir.

Inicialmente, para uma dada escala s e deslocamento τ , definem-se

$$P_k(n) = g\left(\frac{n-\tau}{s}\right) \cos\left(\frac{2\pi nk}{N}\right), \quad (5.20)$$

$$Q_k(n) = g\left(\frac{n-\tau}{s}\right) \text{sen}\left(\frac{2\pi nk}{N}\right). \quad (5.21)$$

Deve-se notar que essas equações são iguais a (5.8), fazendo-se $\gamma = (2^j, \tau, 2\pi k/N)$. Para cada valor de $k = 0, 1, \dots, N - 1$, utiliza-se a FFT para calcular as grandezas

$$\langle \mathbf{x}, \mathbf{P}_k \rangle, \langle \mathbf{x}, \mathbf{Q}_k \rangle, \langle \mathbf{P}_k, \mathbf{P}_k \rangle, \langle \mathbf{Q}_k, \mathbf{P}_k \rangle \text{ e } \langle \mathbf{Q}_k, \mathbf{Q}_k \rangle, \quad (5.22)$$

que são usadas no cálculo da projeção $\langle \mathbf{x}, \mathbf{g}_{\gamma, \phi} \rangle$.

Seja a transformada discreta de Fourier (DFT) $Y(k)$ de uma seqüência $y(j)$ de comprimento N

$$Y(k) = \sum_{j=0}^N y(j) e^{-i\frac{2\pi jk}{N}} = \sum_{j=0}^N y(j) \cos\frac{2\pi jk}{N} - i \sum_{j=0}^N y(j) \text{sen}\frac{2\pi jk}{N}. \quad (5.23)$$

Tendo em mente as equações (5.15)-(5.19), ao se realizar a transformada de Fourier da seqüência $y_1(j) = x(j)g\left(\frac{j-\tau}{s}\right)$, obtêm-se para $k = 0, 1, \dots, N - 1$ os produtos internos

$$\langle \mathbf{x}, \mathbf{P}_k \rangle = \text{Re}\{Y_1(k)\}, \quad (5.24)$$

$$\langle \mathbf{x}, \mathbf{Q}_k \rangle = -\text{Im}\{Y_1(k)\}. \quad (5.25)$$

Agora, considerando como entrada da transformada de Fourier a seqüência

$y_2(j) = g^2 \left(\frac{j-\tau}{s} \right)$, obtêm-se

$$\langle \mathbf{P}_k, \mathbf{P}_k \rangle = \frac{1}{2} (\text{Re} \{Y_2(0)\} + \text{Re} \{Y_2(2k)\}), \quad (5.26)$$

$$\langle \mathbf{Q}_k, \mathbf{Q}_k \rangle = \frac{1}{2} (\text{Re} \{Y_2(0)\} - \text{Re} \{Y_2(2k)\}), \quad (5.27)$$

$$\langle \mathbf{Q}_k, \mathbf{P}_k \rangle = -\frac{1}{2} \text{Im} \{Y_2(2k)\}, \quad (5.28)$$

para $k = 0, 1, \dots, \frac{N}{2} - 1$, e

$$\langle \mathbf{P}_k, \mathbf{P}_k \rangle = \left\langle \mathbf{P}_{k-\frac{N}{2}}, \mathbf{P}_{k-\frac{N}{2}} \right\rangle, \quad (5.29)$$

$$\langle \mathbf{Q}_k, \mathbf{Q}_k \rangle = \left\langle \mathbf{Q}_{k-\frac{N}{2}}, \mathbf{Q}_{k-\frac{N}{2}} \right\rangle, \quad (5.30)$$

$$\langle \mathbf{Q}_k, \mathbf{P}_k \rangle = \left\langle \mathbf{Q}_{k-\frac{N}{2}}, \mathbf{P}_{k-\frac{N}{2}} \right\rangle, \quad (5.31)$$

para $k = \frac{N}{2}, \dots, N - 1$.

O Algoritmo 5.1 descreve, em pseudocódigo, a implementação do MP utilizando esta estratégia proposta em [80]. Esse algoritmo apresenta uma complexidade $\mathcal{O}(2N \log_2 N)$. A grande vantagem desta estratégia é obter os produtos internos entre o resíduo e N átomos através do uso de duas FFTs. Contudo, sua desvantagem é que o espaçamento entre as frequências do dicionário, nesse caso, é obrigatoriamente linear.

Redução de Custo Computacional sem Restrição de Espaçamento Frequencial

Com base na estratégia de [80], propõe-se uma versão rápida do MP tendo a frequência ξ como parâmetro de livre escolha, em vez do deslocamento temporal τ . Nesta implementação, utiliza-se também a transformada rápida de Fourier. Com o intuito de aproveitar ao máximo o potencial computacional dessa transformada, os sinais devem ter como comprimento uma potência de 2. Dessa maneira, os valores de escala s continuam sendo potências de 2, i.e., $s = 2^j$, onde $j = 1, 2, \dots, \log_2 N$.

Primeiramente, para dadas escala s e frequência ξ , definem-se

$$P_\tau(n) = g \left(\frac{n-\tau}{s} \right) \cos(\xi n), \quad (5.32)$$

$$Q_\tau(n) = g \left(\frac{n-\tau}{s} \right) \text{sen}(\xi n). \quad (5.33)$$

A fim de calcular uma série de projeções do sinal sobre versões deslocadas no tempo de um átomo de forma mais rápida, valemo-nos de uma propriedade específica da DFT referente à convolução circular no tempo [17]. Essa propriedade enuncia

Algoritmo 5.1 Implementação de *Ferrando et al* [80] do *Matching Pursuit*

```
s = N
coef = 0
enquanto s > 1 faça
  dτ =  $\frac{s}{2}$ 
  para τ = 0 ; τ < N; τ++ = dτ faça
    para j = 0; j < N; j++ faça
      y1(j) = x(j)g( $\frac{j-\tau}{s}$ )
      y2(j) = g2( $\frac{j-\tau}{s}$ )
    fim para
    Y1 = FFT(y1)
    Y2 = FFT(y2)
    para k = 0; k < N; k++ faça
      calcula as grandezas de (5.22)
      calcula a1, b1 de acordo com (5.11)
      calcula prodint = < x, gγ,φ >
      se |prodint| > coef então
        coef = prodint
        g = gγ,φ
      fim se
    fim para
  fim para
  s =  $\frac{s}{2}$ 
fim enquanto
```

que o produto escalar entre as transformadas para o domínio da frequência de dois vetores equivale à convolução circular entre eles no domínio do tempo, ou seja

$$\sum_{l=0}^{N-1} z(l)w(n-l) = \sum_{l=0}^{N-1} z(n-l)w(l) \leftrightarrow Z(k)W(k). \quad (5.34)$$

Dessa forma, é possível calcular a convolução circular de duas sequências $z(n)$ e $w(n)$, de comprimentos N_z e N_w , utilizando a DFT inversa do produto de suas DFTs. Contudo, o que se deseja é obter a convolução linear entre essas duas sequências utilizando-se este artifício. Para isso, deve-se aumentar os seus comprimentos para $N_z + N_w - 1$, preenchendo as amostras adicionais com zero.

Sejam $z(n) = x(n)e^{i\xi n}$ e $w(n) = g(\frac{N-1-n}{s})$ duas sequências de comprimento N , e $z_1(n)$ e $w_1(n)$ suas respectivas sequências de comprimento aumentado $2N - 1$ após o preenchimento com zeros. A convolução linear entre $z(n)$ e $w(n)$ é calculada da seguinte forma

$$[z * w](n) = \mathcal{F}^{-1}\{Z_1(k) \cdot W_1(k)\}, \quad (5.35)$$

em que $\mathcal{F}^{-1}\{\cdot\}$ é o operador de DFT inversa. Observa-se que $w(n)$ corresponde à reversão temporal de $g(\frac{n}{s})$, pois, dessa forma, através da convolução linear entre

$z(n)$ e $w(n)$ obtêm-se a correlação temporal entre as sequências $x(n)e^{i\xi n}$ e $g(\frac{n}{s})$, que é o que de fato se almeja.

Tendo em mente as equações (5.15) e (5.16) e em mãos a sequência $[z * w](n)$, obtêm-se

$$\langle \mathbf{x}, \mathbf{P}_\tau \rangle = \text{Re} \{ [z * w](\tau + N - 1) \}, \quad (5.36)$$

$$\langle \mathbf{x}, \mathbf{Q}_\tau \rangle = \text{Im} \{ [z * w](\tau + N - 1) \}. \quad (5.37)$$

Agora, sejam $z(n) = e^{i2\xi n}$ e $w(n) = g^2(\frac{N-1-n}{s})$, e $z_1(n)$ e $w_1(n)$ suas respectivas sequências de comprimento aumentado $2N - 1$ com preenchimento de zeros. Obtêm-se a convolução linear $[z * w](n)$ como descrito na equação (5.35). Com base nesse resultado e nas equações (5.17)-(5.19), obtêm-se

$$\langle \mathbf{P}_\tau, \mathbf{P}_\tau \rangle = \frac{1}{2} \{ \text{Re} \{ [z_{\xi=0} * w](\tau + N - 1) \} + \text{Re} \{ [z * w](\tau + N - 1) \} \}, \quad (5.38)$$

$$\langle \mathbf{Q}_\tau, \mathbf{Q}_\tau \rangle = \frac{1}{2} \{ \text{Re} \{ [z_{\xi=0} * w](\tau + N - 1) \} - \text{Re} \{ [z * w](\tau + N - 1) \} \}, \quad (5.39)$$

$$\langle \mathbf{Q}_\tau, \mathbf{P}_\tau \rangle = \frac{1}{2} \text{Im} \{ [z * w](\tau + N - 1) \}, \quad (5.40)$$

em que $z_{\xi=0}(n) = 1$.

O Algoritmo 5.2 descreve, em pseudocódigo, a implementação do MP utilizando esta estratégia. Esse algoritmo apresenta uma complexidade $\mathcal{O}(12N \log_2 2N)$. Deve-se observar que através de duas FFTs e uma FFT inversa, todas de comprimentos $2N - 1$, é possível calcular N produtos internos associados aos deslocamentos temporais dos átomos. Portanto, embora esta estratégia apresente maior custo computacional que o método proposto por Ferrando *et al.* em [80], não exige que o espaçamento entre as frequências seja linear.

5.2.4 Heurísticas de Aproximação

Originalmente, no MPSA busca-se, a cada iteração no dicionário discreto de senoides amortecidas, o átomo de maior correlação com o resíduo e, em seguida, maximiza-se a aproximação entre o átomo e o resíduo, otimizando-se os parâmetros de decaimento, frequência e deslocamento no tempo. Em sequência, realizam-se os procedimentos de busca por melhor suporte temporal, quantização da frequência e identificação de senoide pura. Na primeira etapa obtêm-se a melhor aproximação do sinal em termos de correlação (produto interno) com o resíduo, considerando o suporte completo do sinal/resíduo. Em seguida, na segunda etapa, são realizados procedimentos com o objetivo de maximizar a coerência entre o átomos e estes fenômenos localizados.

No MPSACr, simplifica-se o método de decomposição, procurando-se, desde o

Algoritmo 5.2 Implementação de *Ferrando et al.* [80] modificada do *Matching Pursuit*

```

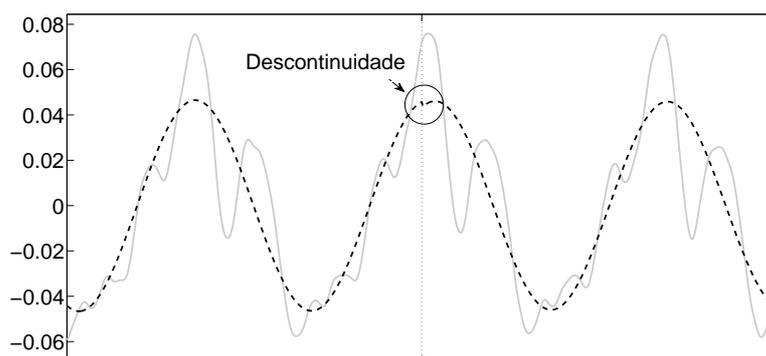
s = N
coef = 0
enquanto s > 1 faça
  para  $\xi = 0; \xi < 2\pi; \xi(+ = ou* =)d\xi$  faça
    para  $n = 0; n < 2N - 1; n ++$  faça
       $z_1(n) = 0; w_1(n) = 0; z_2(n) = 0; w_2(n) = 0$ 
    fim para
    para  $n = 0; n < N - 1; n ++$  faça
       $z_1(n) = x(n)e^{i\xi n}$ 
       $w_1(n) = g \left( \frac{N-1-n}{s} \right)$ 
       $z_2(n) = e^{i2\xi n}$ 
       $w_2(n) = g^2 \left( \frac{N-1-n}{s} \right)$ 
    fim para
    conv_zw1 = IFFT(FFT( $\mathbf{z}_1$ ) · FFT( $\mathbf{w}_1$ ))
    conv_zw2 = IFFT(FFT( $\mathbf{z}_2$ ) · FFT( $\mathbf{w}_2$ ))
    se  $\xi = 0$  então
      conv_zxi0_w2 = conv_zw2
    fim se
    para  $\tau = -N + 1; \tau < N; \tau + = d\tau$  faça
      calcula as grandezas de (5.36)-(5.40)
      calcula  $a_1, b_1$  de acordo com (5.11)
      calcula prodint =  $\langle \mathbf{x}, \mathbf{g}_{\gamma, \phi} \rangle$ 
      se  $|\text{prodint}| > \text{coef}$  então
        coef = prodint;
         $\mathbf{g} = \mathbf{g}_{\gamma, \phi}$ 
      fim se
    fim para
  fim para
   $s = \frac{s}{2}$ 
fim enquanto

```

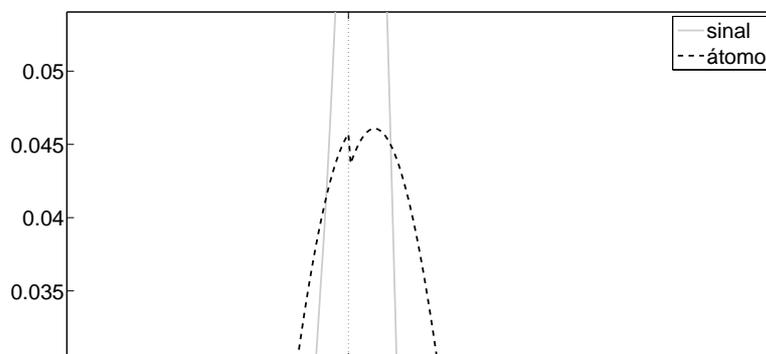
início do processo, maximizar simultaneamente a correlação e a coerência. Dessa forma, após a busca em um dicionário discreto o deslocamento no tempo e a frequência são ótimos em termos de correlação e coerência, posto que o dicionário discreto representa as frequências da escala cromática e todos os deslocamentos temporais. Em seguida, há um refinamento da região de suporte do átomo. Por fim, otimiza-se o decaimento de forma a maximizar a aproximação entre o átomo e o resíduo também em termos de correlação e coerência. É importante destacar que o procedimento de identificação de trechos de senoide pura no sinal não está presente no MPSACr, porque ele aceita uma correlação mais baixa entre o átomo e o resíduo em detrimento de maior similaridade quanto à forma. Como o objetivo é aplicar este método em um sistema de compressão, decide-se buscar sempre o átomo de maior produto interno com o resíduo.

5.2.5 Heurística de Continuidade de Átomos entre Quadros

O processo de decomposição, quando realizado quadro a quadro, resulta na descontinuidade de átomos que estão presentes em quadros consecutivos, como ilustrado na Figura 5.4. Isto ocorre quando a representação de cada quadro é obtida independentemente dos demais quadros adjacentes. É verdade que estas descontinuidades vão sendo corrigidas à proporção que as aproximações sucessivas do quadro são realizadas; no entanto, o número de átomos necessários para se representar o sinal se torna cada vez maior.



(a)



(b)

Figura 5.4: Descontinuidade de um átomo presente em dois quadros consecutivos devido à segmentação do sinal (a). A linha vertical pontilhada indica o local da separação entre os dois blocos e (b) apresenta a ampliação do trecho circulado em (a).

Com o intuito de evitar a ocorrência deste fenômeno, propõe-se o uso de uma heurística de continuidade de átomos entre quadros consecutivos, cujo procedimento é ilustrado na Figura 5.5. A idéia geral desta heurística consiste em, a cada quadro: buscar os átomos no livro de estruturas proveniente do quadro anterior que são altamente correlacionados com o sinal no quadro corrente; em seguida, adicionar ao livro de estruturas do quadro corrente átomos que correspondem exatamente a

versões continuadas dos átomos do quadro anterior.

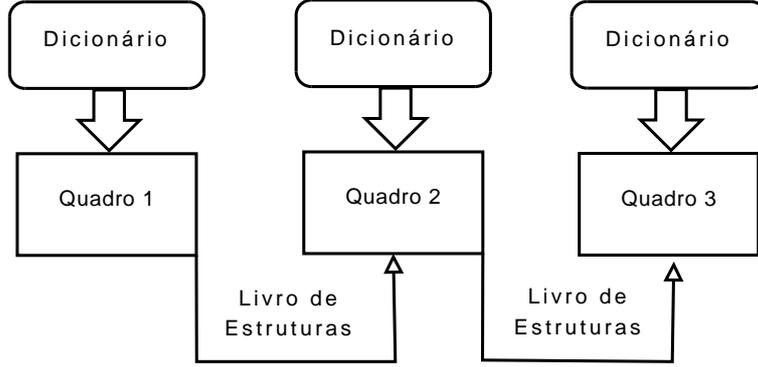


Figura 5.5: Decomposição de sinais usando heurística de continuidade de átomos entre quadros consecutivos.

Os detalhes desse procedimento são descritos a seguir:

1. Estendem-se, para o quadro analisado, os átomos do quadro anterior cuja amostra final (n^e) corresponde à última amostra do quadro;
2. Busca-se o fator de decaimento (ρ_o) que minimiza o erro quadrático entre os átomos continuados/estendidos;
3. Se o coeficiente do átomo for maior que ϵ_α (10^{-8} , neste caso) e o valor absoluto do decaimento for menor que ϵ_ρ (4, neste caso), ele é selecionado; em caso contrário, é descartado.

O valor de ϵ_α foi escolhido de modo que se o coeficiente for menor que ϵ_α passa a ser desprezível, ainda mais se lhe for aplicada alguma quantização em uma etapa posterior. O valor $\epsilon_\rho = 4$ foi escolhido porque para decaimentos maiores que 4, os átomos se aproximam do impulso unitário.

A Figura 5.6 mostra a extensão de um átomo do quadro $i - 1$ para o quadro i . Primeiramente, calcula-se a envoltória do átomo $\mathbf{g}_{\gamma(i-1)}$ e estende-se esta envoltória até o fim do quadro i . Depois, com base no trecho da envoltória do quadro i , calculam-se os parâmetros do átomo $\mathbf{g}_{\gamma(i)}$ que corresponde à extensão de $\mathbf{g}_{\gamma(i-1)}$ no quadro i .

Sejam $\alpha(i - 1)$ e $\gamma(i - 1) = \{\rho(i - 1), \xi(i - 1), \phi(i - 1), n^s(i - 1), n^e(i - 1)\}$ o coeficiente e os parâmetros relativos a $\mathbf{g}_{\gamma(i-1)}$; $\alpha_0(i - 1)$ e $\gamma_0(i - 1) = \{\rho_0(i - 1), \xi_0(i - 1), \phi_0(i - 1), n_0^s(i - 1), n_0^e(i - 1)\}$ o coeficiente e os parâmetros relativos à envoltória $\mathbf{g}_{\gamma_0(i-1)}$; e L_Q o tamanho dos quadros. Atribuem-se aos parâmetros de $\gamma_0(i - 1)$ os

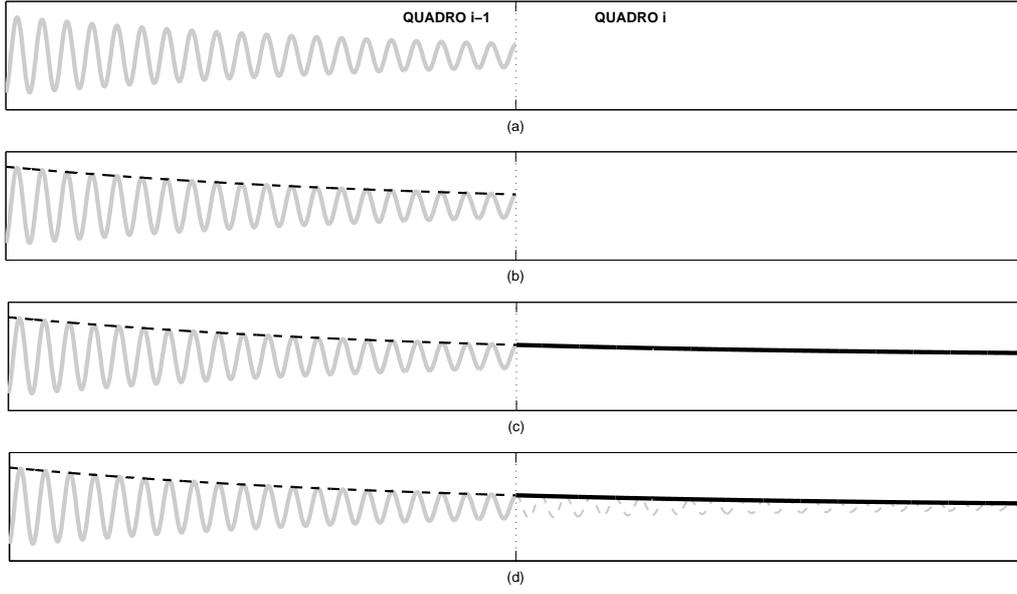


Figura 5.6: Extensão de uma átomo do quadro $i - 1$ para quadro i : (a) o átomo do quadro $i - 1$ a ser estendido, (b) a envoltória deste átomo, (c) a extensão da envoltória ao quadro i e (d) o resultado da extensão do átomo do quadro $i - 1$ ao quadro i .

valores

$$\rho_0(i - 1) = \rho(i - 1), \quad (5.41)$$

$$\xi_0(i - 1) = 0, \quad (5.42)$$

$$\phi_0(i - 1) = 0,$$

$$n_0^s(i - 1) = n^s(i - 1),$$

$$n_0^e(i - 1) = (2 * L_Q) - 1.$$

Em seguida, calcula-se o coeficiente

$$\alpha_0(i - 1) = \alpha(i - 1) \frac{\|\mathbf{g}_{\gamma_0(i-1)}\|}{\|\mathbf{g}_{\gamma(i-1)}\|} \quad (5.43)$$

e calcula-se o vetor

$$\mathbf{x} = \alpha_0(i - 1) \mathbf{g}_{\gamma_0(i-1)}^{[L_Q:(2*L_Q)-1]}, \quad (5.44)$$

em que $\mathbf{g}_{\gamma_0(i-1)}^{[L_Q:(2*L_Q)-1]}$ contém os L_Q últimos elementos do vetor $\mathbf{g}_{\gamma_0(i-1)}$.

Sejam $\alpha_0(i)$ o coeficiente e $\gamma_0(i) = \{\rho_0(i), \xi_0(i), \phi_0(i), n_0^s(i), n_0^e(i)\}$ os parâmetros de $\mathbf{g}_{\gamma_0(i)}$, átomo associado ao trecho da envoltória no quadro i . Atribuem-se aos

parâmetros de $\gamma_0(i)$ os valores

$$\begin{aligned}\rho_0(i) &= \rho_0(i-1), \\ \xi_0(i) &= \xi_0(i-1), \\ \phi_0(i) &= \phi_0(i-1), \\ n_0^s(i) &= 0, \\ n_0^e(i) &= L_Q - 1\end{aligned}\tag{5.45}$$

e calcula-se

$$\alpha_0(i) = \mathbf{x}^T \mathbf{g}_{\gamma_0(i)}.\tag{5.46}$$

Sejam $\alpha(i)$ e $\gamma(i) = \{\rho(i), \xi(i), \phi(i), n^s(i), n^e(i)\}$ o coeficiente e os parâmetros de $\mathbf{g}_{\gamma(i)}$ que corresponde a extensão $\tilde{\mathbf{g}}_{\gamma(i-1)}$ no quadro i . Atribuem-se aos parâmetros de $\gamma(i)$ os valores

$$\begin{aligned}\rho(i) &= \rho(i-1), \\ \xi(i) &= \xi(i-1), \\ n^s(i) &= 0, \\ n^e(i) &= L_Q - 1\end{aligned}\tag{5.47}$$

e ajusta-se a fase $\phi(i)$

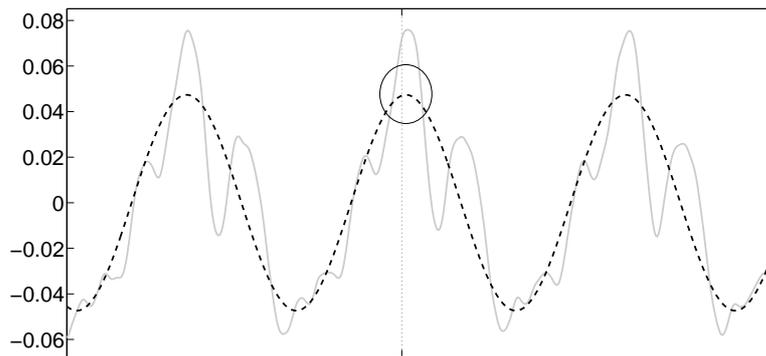
$$\phi(i) = (L_Q \xi(i-1) + \phi(i-1)) - \left\lfloor \frac{L_Q \xi(i-1) + \phi(i-1)}{2\pi} \right\rfloor 2\pi.\tag{5.48}$$

A fase $\phi(i)$ é calculada de modo que seu ponto de referência passe a ser a amostra inicial do quadro i . Finalmente, calcula-se

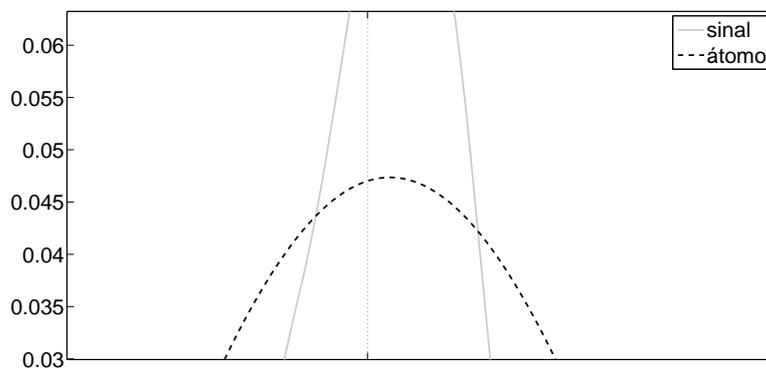
$$\alpha(i) = \alpha_0(i) \frac{\|\mathbf{g}_{\gamma(i)}\|}{\|\mathbf{g}_{\gamma_0(i)}\|}.\tag{5.49}$$

A Figura 5.7 mostra que a descontinuidade apresentada na Figura 5.4 desaparece ao se utilizar a extensão do átomo no quadro anterior para o quadro corrente. Em [68, 118], apresentam-se técnicas de extrapolação de sinais de áudio para restauração e aumento da resolução espectral em métodos baseados em transformada rápida de Fourier de tempo curto.

Calculados $\alpha(i)$ e $\gamma(i)$, busca-se o decaimento que minimiza o erro quadrático



(a)



(b)

Figura 5.7: Eliminação de (a) descontinuidade do átomo apresentado na Figura 5.4. A linha vertical pontilhada indica o local da separação entre os dois quadros e (b) apresenta a ampliação do trecho circulado em (a).

entre $\mathbf{g}_{\gamma(i)}$ e o resíduo \mathbf{r} a ser aproximado no quadro i , ou seja,

$$\rho(i) = \arg \min_{\rho} \|\mathbf{r} - \alpha(i)\mathbf{g}_{\gamma(i)}\|^2. \quad (5.50)$$

Para isso, utiliza-se o mesmo algoritmo pseudo-Newton usado na etapa de aproximação inicial da decomposição [1, 4, 25]. Dessa maneira, é possível avaliar se a extensão do átomo do quadro anterior possui ou não correlação com os padrões do sinal presente no quadro analisado. Além disso, podem-se detectar também variações na envoltória do átomo.

A cada variação do decaimento durante a otimização, deve-se recalculer o coeficiente correspondente. Seja $\gamma_0'(i) = \{\rho_0'(i), \xi_0(i), \phi_0(i), n_0^s(i), n_0^e(i)\}$, onde $\rho_0'(i)$ corresponde a uma variação de $\rho_0(i)$. Tem-se que

$$\alpha_0'(i) = \alpha_0(i) \frac{\mathbf{g}_{\gamma_0'(i)}^{[1]}}{\mathbf{g}_{\gamma_0(i)}^{[1]}}, \quad (5.51)$$

$$\alpha(i) = \alpha_0'(i) \frac{\|\mathbf{g}_{\gamma(i)}\|}{\|\mathbf{g}_{\gamma_0'(i)}\|}, \quad (5.52)$$

em que $\mathbf{g}_{\gamma_0'(i)}^{[1]}$ é o primeiro elemento do vetor $\mathbf{g}_{\gamma_0'(i)}$ e $\mathbf{g}_{\gamma_0(i)}^{[1]}$ é o primeiro elemento do vetor $\mathbf{g}_{\gamma_0(i)}$.

Na Figura 5.8, gerou-se um sinal sintético com 8192 amostras, dividido em dois quadros de 4096 amostras. Este sinal apresenta uma variação de envoltória no limite entre os quadros. Observe que, no segundo quadro, a extensão do átomo do primeiro quadro possui baixa correlação com o sinal do segundo quadro, devido à variação de envoltória. Ao buscar-se o decaimento que minimiza o erro quadrático entre o átomo continuado e o sinal do segundo quadro, é possível melhorar o casamento entre eles, como pode ser visto na Figura 5.8(b).

No caso em que o padrão do sinal muda completamente de um quadro para outro, a otimização do decaimento trata de buscar um fator de decaimento tão alto que torna a influência do átomo continuado desprezível no quadro corrente, como ilustrado na Figura 5.9. O sinal sintético é formado por uma senoide amortecida decaindo suavemente no primeiro quadro e por um sinal ruidoso de mais baixa potência no segundo. Ao otimizar-se o decaimento, o átomo continuado adquire um fator de decaimento bastante expressivo (nesse caso, ρ passou de 0,00009999 para 52,63) e sua forma de onda passa a ter um valor próximo de zero por todo o segundo quadro.

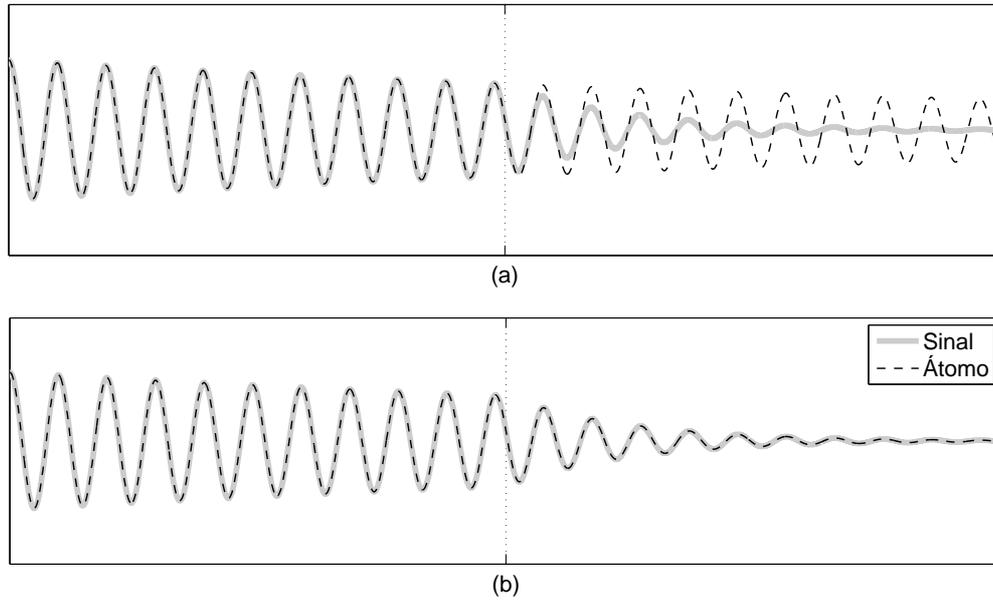


Figura 5.8: Sinal com variação de decaimento entre quadros: (a) extensão de uma átomo para um quadro consecutivo e (b) o resultado da obtenção do decaimento ótimo. A linha vertical pontilhada indica o local da separação entre os dois quadros.

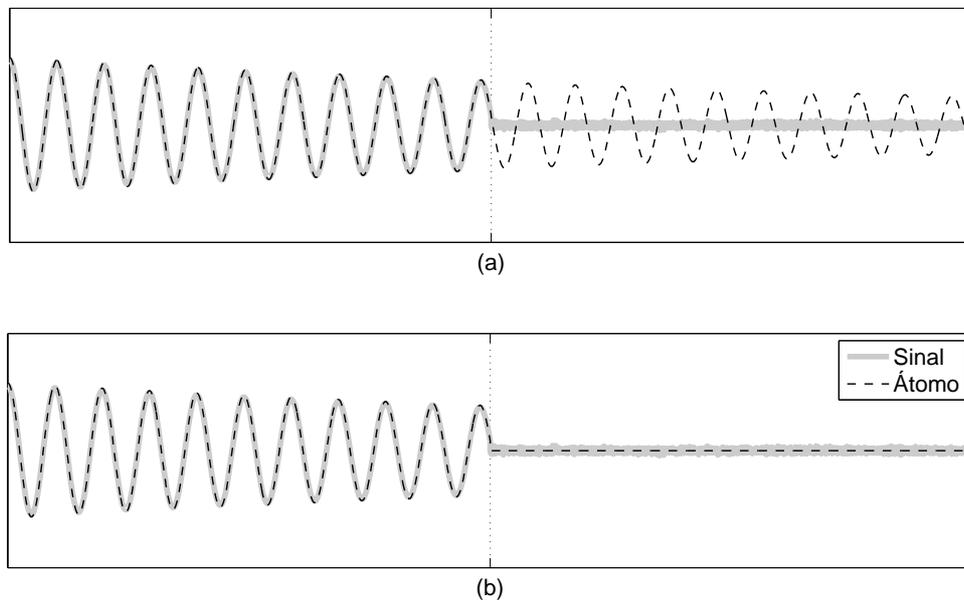


Figura 5.9: Sinal com mudança radical de padrão entre quadros: (a) extensão de uma átomo para um quadro consecutivo e (b) o resultado da obtenção do decaimento ótimo. A linha vertical pontilhada indica o local da separação entre os dois quadros.

5.2.6 Critério de Parada

O critério de parada do MPSACr é o critério utilizado no MPSA (ver Seção 3.2.4) em conjunto com uma medida de razão entre as energias do sinal e do ruído, a SNR (*signal-to-noise ratio*). Para um dado passo M , calcula-se uma média móvel considerando os próximos L passos

$$\bar{\eta}(\mathbf{r}_x^M) = \frac{1}{L} \sum_{n=M}^{M+L-1} \frac{\|\langle \mathbf{r}_x^n, \mathbf{g}_{\gamma_n} \rangle\|}{\|\mathbf{r}_x^n\|}, \quad (5.53)$$

em que $L = \log_2 N$, sendo N o comprimento do sinal. Vale recordar que L é escolhido de forma empírica com o intuito de criar uma relação entre a média e o comprimento do sinal. Quando a média móvel se aproxima do valor médio estacionário η_o significa que o resíduo passa a estar incoerente com relação ao dicionário. Dessa forma, a decomposição é realizada enquanto a seguinte desigualdade for satisfeita:

$$\bar{\eta}(\mathbf{r}_x^M) > \eta_o + \epsilon, \quad (5.54)$$

onde $\epsilon \in \mathbb{R}$ é uma constante de confiabilidade. Obtém-se, assim, uma representação do sinal constituída somente de átomos coerentes, separando-os do ruído [4, 42, 83].

É importante lembrar que o valor de η_o depende do dicionário e de sua dimensão, ou seja, do comprimento do sinal. Considerando o dicionário discreto usado no MPSACr, apresentado na Seção 5.2.2, calcula-se η_o para as dimensões 512, 1024, 2048 e 4096, pois estes são os valores mais usuais quando se analisam sinais de áudio amostrados de 44,1 kHz a 48 kHz [15, 119].

A Figura 5.10 mostra o comportamento da taxa de aproximação para mil realizações de um processo de ruído gaussiano branco de média zero para diferentes comprimentos de sinal. Observa-se que para cada comprimento, a taxa oscila em torno de um valor médio. Os valores de η_o e o desvio padrão a serem utilizados no critério de parada da decomposição são mostrados na Tabela 5.1.

Tabela 5.1: Tabela de valores do critério de parada obtidos a partir da decomposição de sinais de ruído de diferentes comprimentos utilizando MPSACr com base no dicionário discreto de senoides amortecidas de frequências da escala musical cromática.

Comprimento do Sinal	512	1024	2048	4096
η_o	0,2013	0,1488	0,1095	0,0801
desvio padrão	0,0141	0,0095	0,0065	0,0043

Portanto, para que a decomposição termine, deve-se satisfazer o critério $\bar{\eta}(\mathbf{r}_x^M) < \eta_o + \epsilon$ e simultaneamente uma SNR alvo dever ser atingida. Dessa forma, garante-se que ao final da decomposição o resíduo consiste de um ruído branco gaussiano e

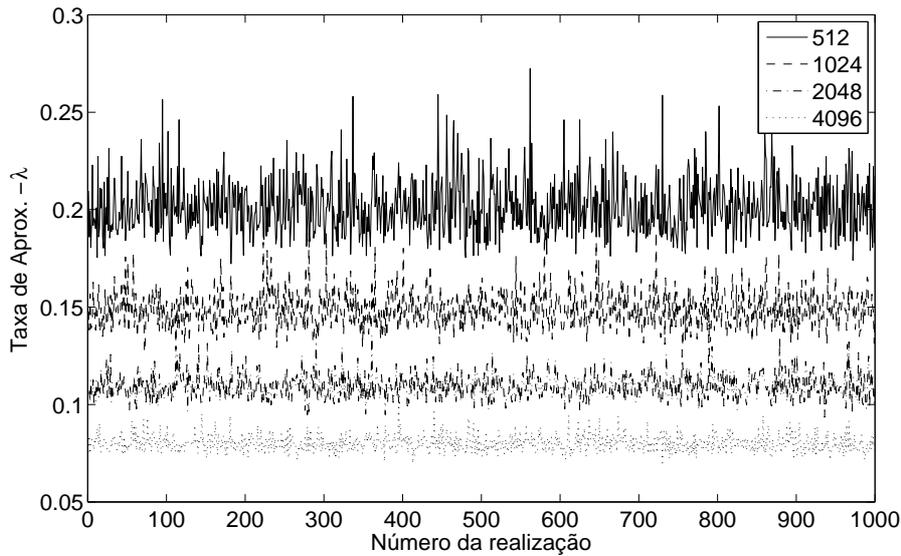


Figura 5.10: Comportamento da taxa de aproximação para ruídos de diferentes comprimentos.

que a representação do sinal apresenta potencialmente uma boa qualidade sonora ao ouvido humano.

5.3 Simulações Computacionais

O objetivo das simulações computacionais apresentadas nesta seção é realizar uma avaliação do desempenho e de potencialidades do método de decomposição proposto, o MPSACr, quando aplicado a sinais de áudio. Inicialmente, avalia-se a capacidade do MPSACr de obter representações compactas e coerentes para diferentes tipos de sinais. Para este experimento, utilizam-se quatro sinais de áudio:

- *pianoA3*: sinal de piano associado à nota musical A3 proveniente de uma banco de dados de instrumentos musicais, o MUMS (*McGill University Master Samples*) [120]. Possui duração de 1 segundo e taxa de amostragem de 44,1 kHz, portanto são 44100 amostras.
- *fluteA4*: sinal de flauta, com duração de 1 segundo, associado à nota musical A4 proveniente de uma banco de dados de sinais musicais, o RWC (*Real World Computing*) [121]. Possui duração de 1 segundo e taxa de amostragem de 44,1 kHz, portanto são 44100 amostras.
- *celloA4*: sinal de violoncelo, com duração de 1 segundo, associado à nota musical A4 proveniente de uma banco de dados de sinais musicais, o RWC (*Real*

World Computing) [121]. Possui duração de 1 segundo e taxa de amostragem de 44,1 kHz, portanto são 44100 amostras.

- *tirofinal*: sinal de um tiro de revólver com duração de 1,5 segundo e taxa de amostragem de 44,1 kHz, portanto são 65687 amostras.

Ao serem decompostos através do MPSACr, os sinais são subdivididos em quadros de 2048 amostras, o que corresponde a 46,4 ms a uma taxa de amostragem de 44,1 kHz, com saltos de 2048 amostras. Dessa forma, não há sobreposição entre quadros consecutivos. Além disso, não se utiliza nenhuma janela de suavização.

O tamanho do quadro foi escolhido de modo que a decomposição fosse executada em tempo computacional viável, posto que o tamanho do bloco influencia diretamente na dimensão do dicionário. Além disso, pode-se assumir quase-estacionariedade em quadros de 46,4 ms [119]. Dessa maneira, ao utilizar-se a heurística de continuidade, espera-se que o decaimento de um átomo continuado não varie muito abruptamente de um quadro para o outro.

Para o critério de parada, adota-se o valor de 60 dB para a razão sinal-ruído alvo (SNR_{alvo}) e o valor de 0,1095 para taxa de aproximação média desejada, como descrito na Tabela 5.1.

As Figuras 5.11 e 5.12 mostram a forma de onda e o número de átomos obtidos na decomposição em função do quadro, respectivamente, para o sinal *pianoA3*, as Figuras 5.13 e 5.14 mostram para o sinal *fluteA4*, as Figuras 5.15, 5.16 mostram para o sinal *celloA4* e as Figuras 5.17 e 5.18 mostram para o sinal *tirofinal*.

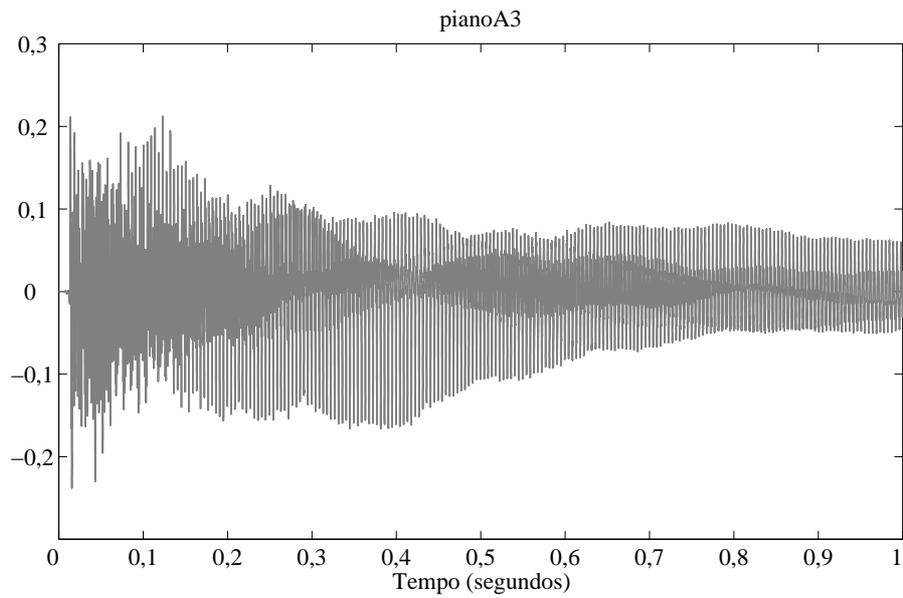


Figura 5.11: Forma de onda do sinal *pianoA3*.

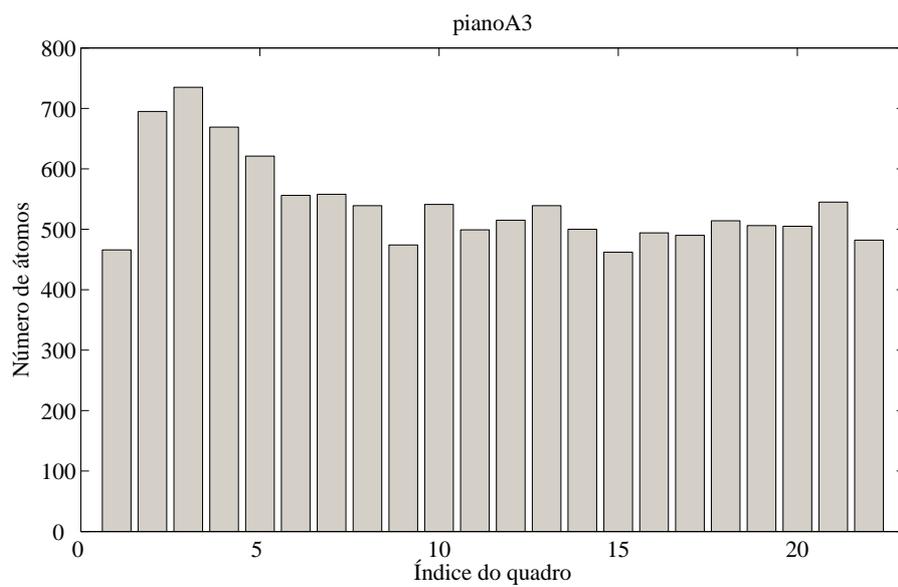


Figura 5.12: Número de átomos versus índice do quadro para o sinal *pianoA3*.

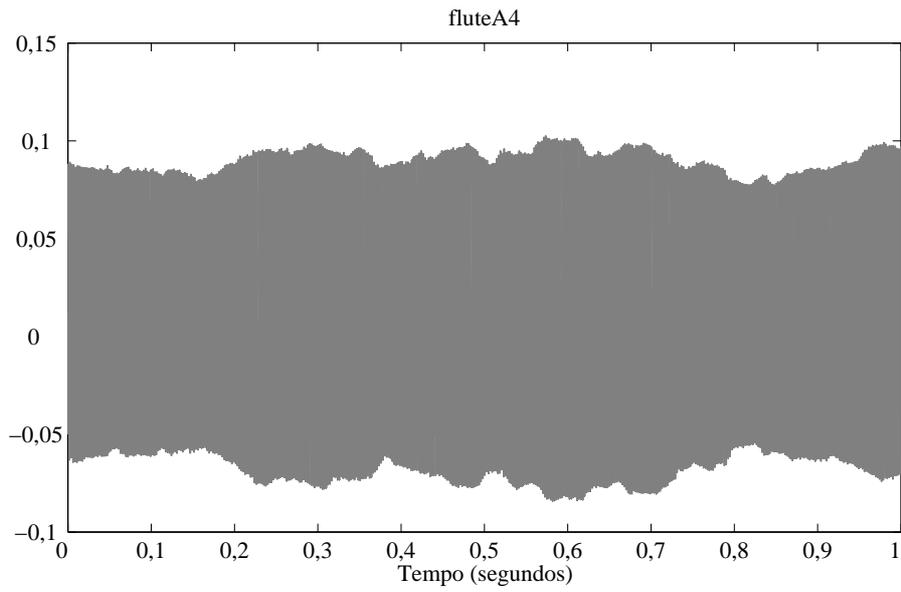


Figura 5.13: Forma de onda do sinal *fluteA4*.

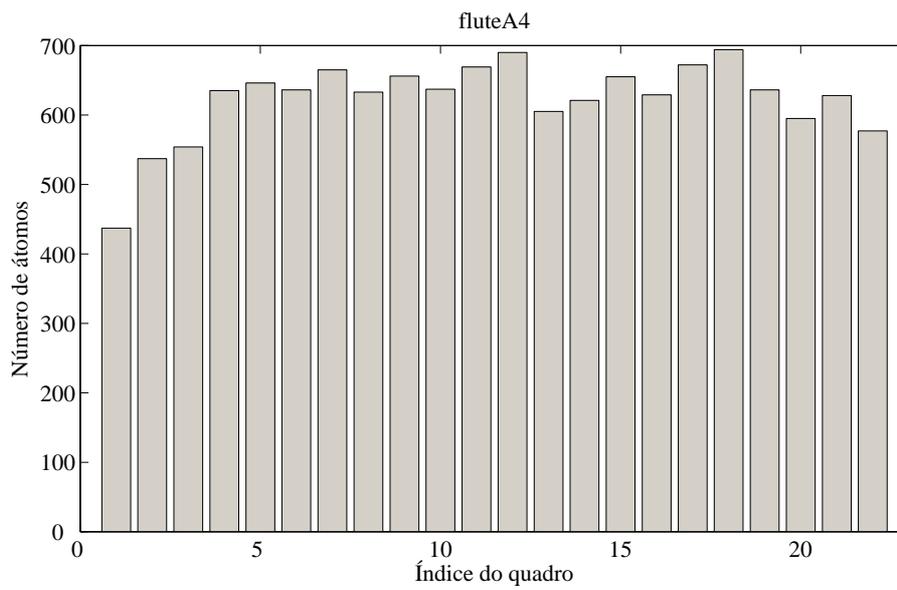


Figura 5.14: Número de átomos versus índice do quadro para o sinal *fluteA4*.

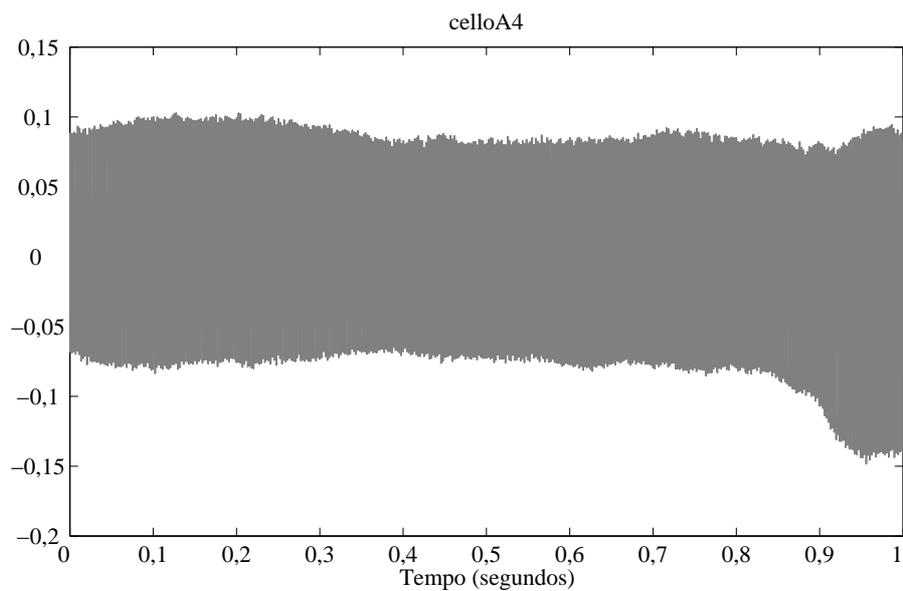


Figura 5.15: Forma de onda do sinal *celloA4*.

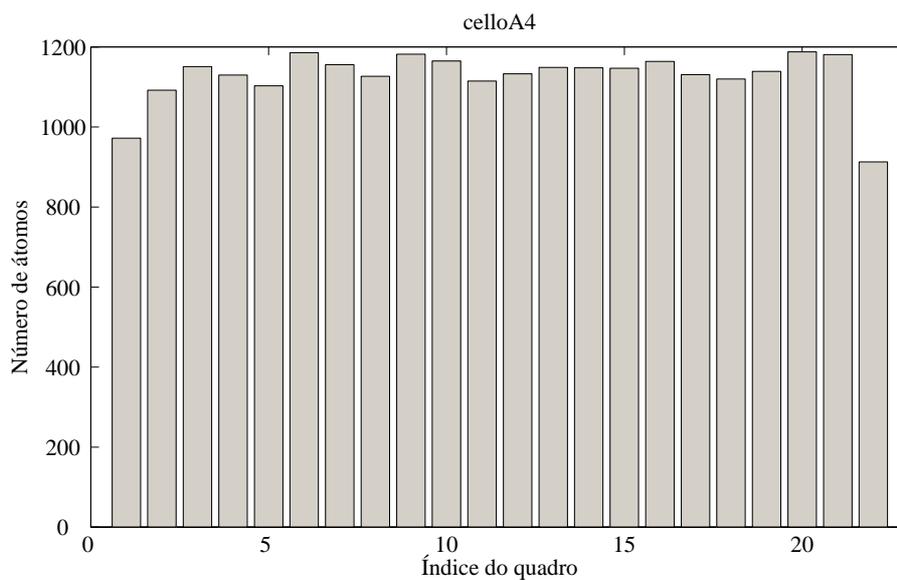


Figura 5.16: Número de átomos versus índice do quadro para o sinal *celloA4*.

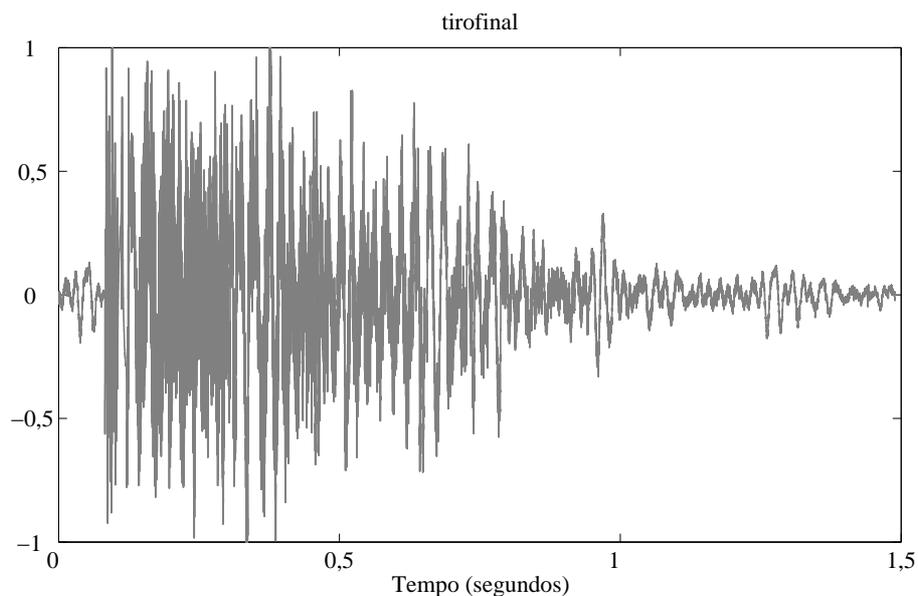


Figura 5.17: Forma de onda do sinal *tirofinal*.

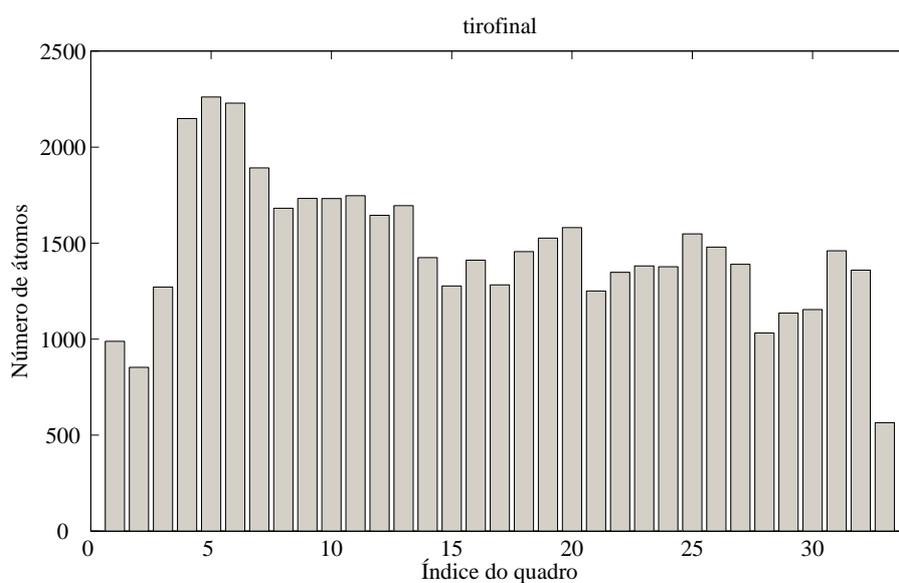


Figura 5.18: Número de átomos versus índice do quadro para o sinal *tirofinal*.

Considerando os sinais na ordem *pianoA3*, *fluteA4*, *celloA4*, e *tirofinal*, é possível verificar que o número de átomos necessários para as representações destes sinais atingirem um mesmo nível de razão sinal-ruído (SNR=60 dB) aumenta. A Tabela 5.2 apresenta a média de átomos por quadro para os diferentes sinais de áudio e corrobora esta constatação. Isto se deve ao nível de ruído inerente ao sinal, que não é bem correlacionado com o dicionário de senoides amortecidas, fazendo com que seja preciso realizar um número maior de iterações de decomposição para alcançar

uma SNR desejada. Ao ouvir estes sinais, verifica-se que o sinal *fluteA4* apresenta um ruído, relativo ao sopro de flautista, de maior intensidade que o ruído inerente ao sinal de *pianoA3*. O sinal *celloA4*, por sua vez, apresenta um ruído relativo à raspagem do arco na corda do violoncelo, de maior intensidade que o ruído inerente ao sinal *fluteA4*. Finalmente, o sinal *tirofinal* é o que apresenta maior nível de ruído não correlacionado com o dicionário de senoides amortecidas, por isso demanda o maior número de átomos para representação, dentre estes sinais.

Tabela 5.2: Média de átomos por quadro para diferentes sinais de áudio.

Sinal	Média de átomos por quadro
<i>pianoA3</i>	541
<i>fluteA4</i>	623
<i>celloA4</i>	1126
<i>tirofinal</i>	1464

Na Figura 5.19, apresentam-se as aproximações relativas às seis primeiras iterações da decomposição do quadro 10 do sinal *pianoA3*. Neste caso, a heurística de continuidade não foi utilizada. Observa-se que no decorrer das iterações da decomposição surgem componentes senoidais puras, exponenciais decrescentes e exponenciais crescentes. É interessante ressaltar que os padrões apresentados nessa figura, em que durante o processo iterativo de aproximações sucessivas exponenciais decrescentes aparecem seguidas de exponenciais crescentes, se repete na decomposição dos demais blocos do sinal. Isso ocorre devido à presença de oscilações de baixa frequência, podendo ser representativas de aspectos físicos do piano. Os pianos têm de uma a três cordas por nota e uma pequena desafinação entre as cordas pode produzir essas oscilações [21]. Pode existir também o acoplamento entre diferentes modos de vibração ao longo do tempo. Alguns modos começam com mais energia, que é então transferida para outros modos, e assim por diante. Esta oscilação entre os modos mais energéticos ao longo do tempo pode causar o surgimento de exponenciais crescentes e decrescentes na modelagem [20].

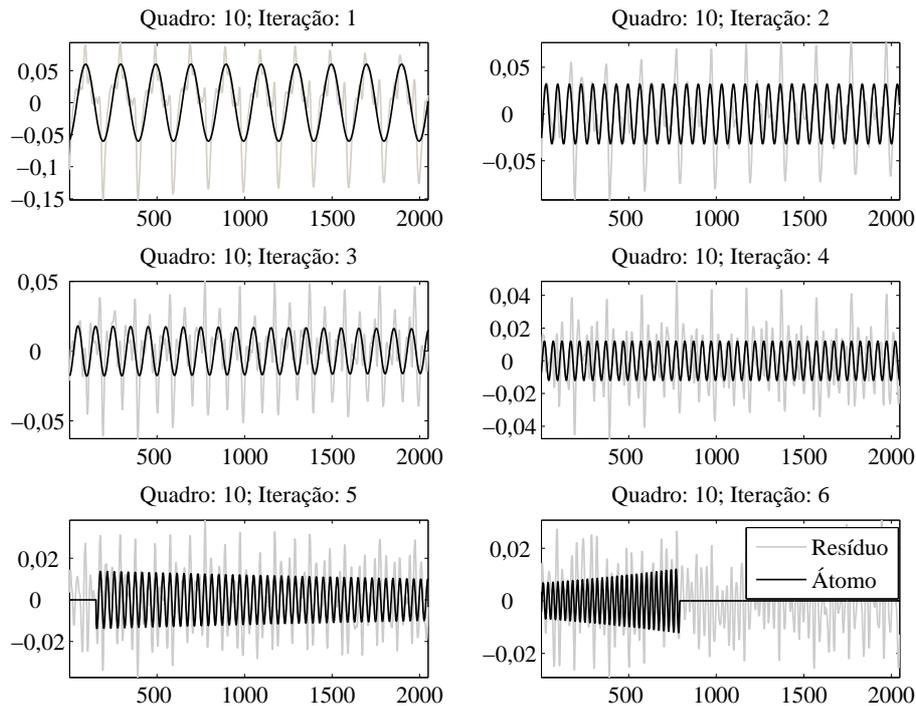


Figura 5.19: Aproximações relativas às seis primeiras iterações da decomposição do quadro 10 do sinal *pianoA3*.

O efeito da heurística de continuidade sobre o resultado da representação do sinal está ilustrado na Figura 5.20. Observam-se três átomos continuados que aproximam o sinal *pianoA3* obtidos pelo MPSACr usando a heurística de continuidade. Nos três casos, nota-se que o primeiro quadro em que surge o átomo corresponde a uma exponencial crescente, nos quadros seguintes o decaimento pode variar positiva e negativamente, e no último quadro termina-se com uma exponencial decrescente. Verifica-se, portanto, que a heurística de continuidade obtém êxito ao eliminar descontinuidades em um mesmo átomo devido à segmentação do sinal em quadros, bem como acompanhar a envoltória de padrões existentes no sinal.

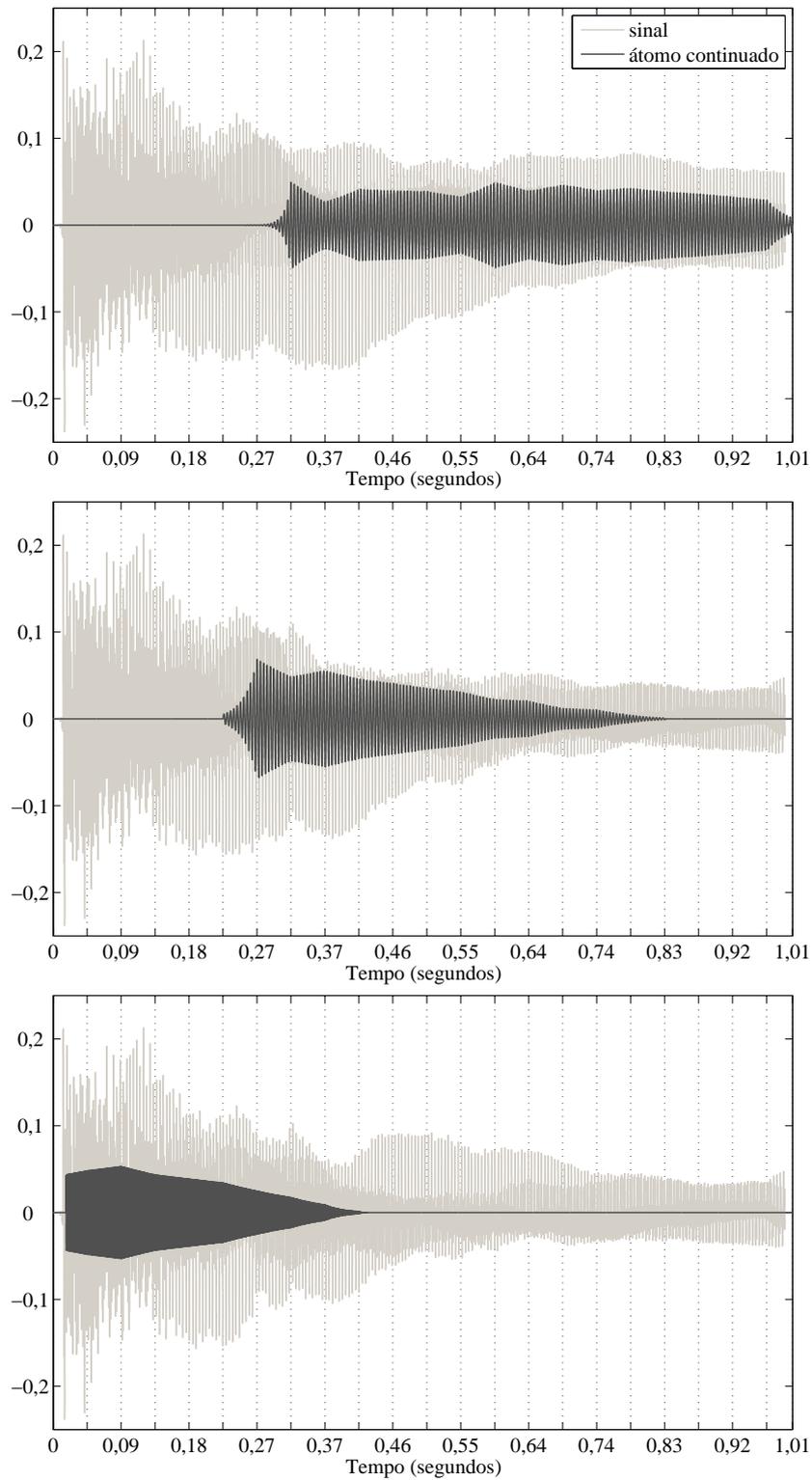


Figura 5.20: Átomos continuados que formam a representação do sinal *pianoA3*. As linhas verticais tracejadas indicam os limites entre quadros.

Capítulo 6

Compressão de Áudio usando Decomposições Atômicas

Na primeira parte deste capítulo, apresenta-se o sistema de compressão aplicado a sinais de áudio proposto e desenvolvido neste trabalho. Este sistema utiliza o método de decomposição com base em um dicionário de senoides amortecidas de frequências na escala musical cromática, o MPSACr, e realiza a alocação ótima de bits através das curvas de taxa-distorção operacionais. Por fim, são apresentadas simulações computacionais desse arcabouço de compressão.

Na segunda parte, apresentam-se a fundamentação teórica e a descrição de um método de otimização taxa-distorção por faixa de energia atômica. Através deste método é possível alocar bits de modo a privilegiar a representação de átomos de maior energia em detrimento dos de menor energia.

Na terceira parte, este novo paradigma de alocação ótima de bits é introduzido no sistema de compressão, em substituição ao método que se baseia em curvas operacionais. Além disso, tendo em vista o uso de codificação aritmética, alguns dos parâmetros associados aos átomos são modelados através de gaussianas generalizadas, promovendo uma atuação mais eficiente do codificador aritmético. Por fim, são apresentadas simulações computacionais deste sistema de compressão aperfeiçoado. Avalia-se o desempenho do sistema de compressão através do PEAQ (*Perceptual Evaluation of Audio Quality*) [29–31], que é um método de medição objetiva de qualidade de sinais de áudio padronizado pela ITU (*International Telecommunication Union*) [28].

6.1 Introdução

A compressão ou codificação de áudio é utilizada para obter representações digitais compactas de sinais de áudio com o intuito de alcançar maior eficiência na trans-

missão ou armazenamento. O objetivo principal é representar o sinal com o número mínimo de bits. No caso de este ser um requisito, deve-se garantir que o codec seja transparente, ou seja, que o sinal reproduzido não seja perceptivamente distinguível do sinal original.

Com a introdução do CD (*compact disc*) no início dos anos 80, a mídia digital de áudio substituiu rapidamente a mídia analógica. O sinal armazenado no CD é amostrado a uma taxa de 44,1 kHz usando PCM (*pulse code modulation*), com representação de 16 bits por amostra. Isto resulta em uma razão sinal-ruído de 90 dB e uma taxa de 706,6 kbps (quilobits por segundo) para um sinal mono. Um sinal estéreo requer uma taxa de 1,41 Mbps (megabit por segundo). Em geral, a qualidade de reprodução do CD é considerada como referência no desenvolvimento de métodos de compressão aplicados a áudio.

Com o desenvolvimento e a proliferação de sistemas de comunicação sem-fio e o aumento do fluxo de mídia (*streaming*) através da Internet, surgiu a demanda por transmissão de áudio digital a taxas baixas [119] devido às restrições de largura de banda, capacidade de armazenamento e custo. Como consequência, ao longo dos últimos anos, diversos algoritmos têm sido desenvolvidos para a codificação perceptivamente transparente de sinais com a qualidade de CD. Em [15, 119, 122–124], encontram-se as descrições de diversas metodologias e atividades de padronização relacionadas à codificação de áudio, que pode ser classificada em dois tipos: sem perdas e com perdas. A codificação sem perdas permite uma razão de compressão de até 3:1, enquanto que através da codificação com perdas é possível atingir uma razão de compressão de 12:1 ou mais [122]. Em codificação com perdas, quanto maior a razão de compressão, maior é o nível de degradação do sinal; no entanto, quando existe uma exigência de taxa de compressão que não pode ser atendida através de técnicas sem perdas, a única alternativa é utilizar técnicas de codificação com perdas. Em [15, 119, 124], é possível encontrar atividades de pesquisa relacionadas ao desenvolvimento de diferentes padrões e métodos de codificação de áudio perceptiva com perdas. Geralmente, a codificação perceptiva é caracterizada por remover redundância estatística e informação perceptivamente irrelevante de sinais de áudio de modo a reduzir a taxa total [125, 126]. A redundância estatística é reduzida usando uma representação esparsa, tal que a energia do sinal seja concentrada em poucos coeficientes ou parâmetros. A informação perceptivamente irrelevante é explorada através de quantização mais grosseira ou até remoção de componentes que não são perceptíveis ao ouvido humano.

Na literatura, existem diversos trabalhos em que se utiliza o algoritmo de *Matching Pursuit* (MP) para a análise e representação de sinais de áudio para fins de codificação [14, 32, 34, 102–108, 112, 126]. De maneira geral, as abordagens utilizam o MP como engrenagem principal na obtenção da representação atômica do sinal de

áudio, e o que as diferencia é o dicionário e a forma como a análise psicoacústica é realizada. Os dicionários podem ser compostos por bases de MDCT (*Modified Discrete Cosine Transform*) [126], senoides amortecidas [14], senoides puras [34, 102, 112], *wavelets* [32], senoides moduladas por amplitude [105, 106]. Com relação à análise psicoacústica, limiares de mascaramento podem ser obtidos em função do sinal e com base nesses limiares os átomos inaudíveis são removidos da representação do sinal [105, 106, 126]. Por outro lado, a cada iteração do MP é possível calcular os limiares de mascaramento em função do sinal reconstruído utilizando os átomos obtidos até a iteração corrente [34, 102, 112]. Neste caso, quando o resíduo passa a ser inaudível, o processo iterativo de decomposição termina. Vale ressaltar que nos trabalhos encontrados na literatura, o dicionário utilizado é fixo e somente os coeficientes e os índices dos átomos são codificados.

O foco deste trabalho está em desenvolver um sistema de compressão de áudio com perdas com base no arcabouço de compressão de sinais usando múltiplos dicionários, segundo o qual além de transmitir os coeficientes e os índices dos átomos, o codificador precisa enviar informação adicional sobre o dicionário utilizado. O sistema de compressão desenvolvido explora a redundância estatística usando um método de decomposição com base em um dicionário de senoides amortecidas de frequências da escala musical cromática, o MPSACr. A informação perceptivamente irrelevante não é explorada pelo sistema de compressão neste estágio de desenvolvimento, contudo é de grande importância para alcançar menor taxas de compressão reproduzindo o sinal de forma transparente, e será tópico de investigação em trabalhos futuros.

6.2 Sistema de Compressão

Em geral, um método de compressão de sinais pode ser dividido em três partes [44], como ilustrado na Figura 6.1:

1. **Transformação:** obtém-se uma representação compacta do sinal, resultando em um menor número de coeficientes;
2. **Quantização:** mapeiam-se os coeficientes da transformação em um conjunto finito de símbolos;
3. **Codificação:** mapeiam-se os símbolos em bits.

Normalmente, na etapa de quantização realiza-se a otimização taxa-distorção que permite a alocação ótima de bits na etapa de codificação.

O sistema de compressão de áudio proposto neste trabalho é apresentado na Figura 6.2. Ele faz uso, na etapa de transformação, do método de decomposição

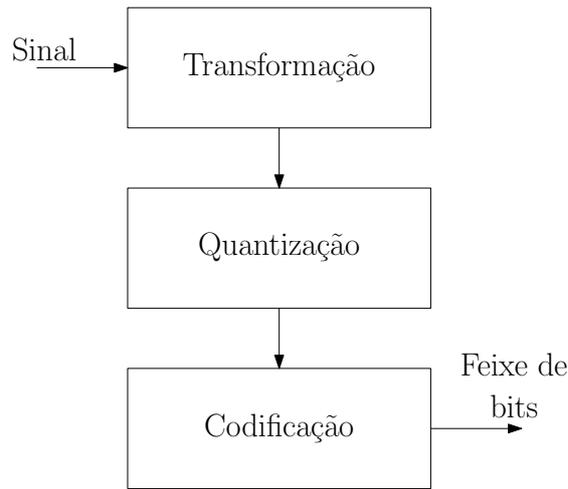


Figura 6.1: Esquema geral de compressão de sinais.

atômica descrito no Capítulo 5, o MPSACr. Portanto, o modelo de sinais em questão é composto por senóides amortecidas com frequências da escala musical cromática. Ao final da decomposição o MPSACr provê o livro de estruturas contendo os coeficientes e parâmetros dos átomos. Em seguida, realiza-se a otimização taxa-distorção através de curvas operacionais, técnica esta apresentada na Seção 2.2 (página 17) e cuja aplicação em compressão de sinais elétricos foi descrita no Capítulo 4. Definida a alocação ótima de bits entre os coeficientes e parâmetros dos átomos, quantiza-se o livro de estruturas, produzindo símbolos que são codificados e transmitidos ao decodificador. No decodificador, o feixe de bits é decodificado, gerando os símbolos. Estes, por sua vez, sofrem o processo de quantização inversa, produzindo o livro de estruturas quantizado. Enfim, com base neste livro de estruturas, reconstrói-se o sinal.

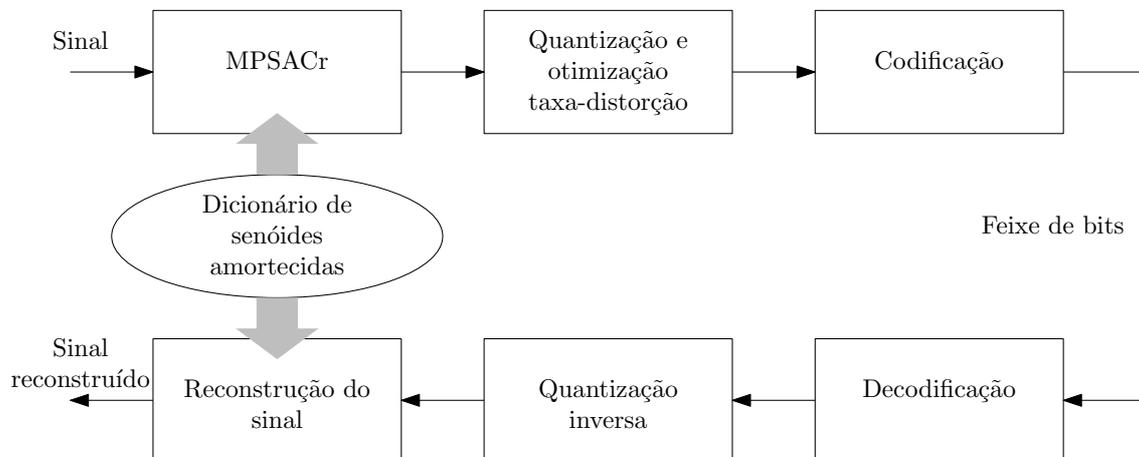


Figura 6.2: Compressão de sinais de áudio realizando a decomposição do sinal via MPSACr e a otimização taxa-distorção através de curvas operacionais.

6.2.1 Quantização e Otimização Taxa-Distorção através de Curvas Operacionais

Os processos de quantização do livro de estruturas e da otimização taxa-distorção são realizados a cada quadro da mesma forma que a decomposição. Entretanto, o tamanho do átomo pode ser maior do que o tamanho do quadro usado na decomposição. É interessante que isto ocorra quando se utiliza a heurística de continuidade de átomos entre quadros consecutivos. Dessa forma, agrupam-se diferentes trechos de um átomo continuado, representando-os através de um único átomo, em que cada trecho se refere a um quadro. Para que trechos consecutivos sejam agrupados em um único átomo é preciso que o fator de decaimento não varie significativamente entre eles. Na prática, tolera-se uma variação de 1 % de um quadro para o outro. Por exemplo, na Tabela 6.1 apresentam-se a representação de um átomo continuado em dois quadros consecutivos de 4096 amostras (α_1, γ_1) e (α_2, γ_2) e a representação em um quadro de 8192 amostras ao agruparem-se estes dois quadros, $(\alpha_{\text{grp}}, \gamma_{\text{grp}})$. Para obter γ_{grp} , fazem-se

$$\begin{aligned}
 \rho_{\text{grp}} &= \rho_1 \\
 \xi_{\text{grp}} &= \xi_1 \\
 \phi_{\text{grp}} &= \phi_1 \\
 n_{\text{grp}}^s &= n_1^s \\
 n_{\text{grp}}^e &= n_1^e + 4096,
 \end{aligned} \tag{6.1}$$

em que 4096 é o tamanho do quadro de decomposição. Para obter α_{grp} calcula-se

$$\alpha_{\text{grp}} = [\alpha_1 \mathbf{g}_1^T \quad \alpha_2 \mathbf{g}_2^T] \mathbf{g}_{\text{grp}}, \tag{6.2}$$

em que \mathbf{g}_1 , \mathbf{g}_2 e \mathbf{g}_{grp} correspondem aos vetores (verticais) contendo as amostras dos átomos relativos a γ_1 , γ_2 e γ_{grp} , respectivamente.

Tabela 6.1: Parâmetros de um átomo continuado em dois quadros de 4096 amostras e a representação deste átomo agrupando estes dois quadros em um quadro de 8192 amostras.

	α	ρ	ξ	ϕ	n^s	n^e
(α_1, γ_1)	0,1989	0,0003	0,0313	3,46	500	4095
(α_2, γ_2)	0,069	0,0003	0,0313	6,19	0	4095
$(\alpha_{\text{grp}}, \gamma_{\text{grp}})$	0,2106	0,0003	0,0313	3,46	500	8191

Considera-se que após a decomposição e o agrupamento de quadros da decomposição, obtém-se a sequência dos pares $(\alpha_m, \gamma(m))$,

$m = 1, 2, \dots, M$, que formam o livro de estruturas, no qual $\gamma(m) = (\rho_m, \xi_m, \phi_m, n_m^s, n_m^e)$ e M é o número total de elementos do livro de estruturas. De maneira equivalente, o suporte temporal do átomo pode ser definido através de n_m^s e o comprimento do suporte $n_m^L = n_m^e - n_m^s$. Logo, tem-se que $\gamma(m) = (\rho_m, \xi_m, \phi_m, n_m^s, n_m^L)$. O resultado do agrupamento de quadros da decomposição é referenciado como quadro de quantização. Em geral, como existem átomos continuados obtidos na decomposição do sinal recomenda-se que a quantização e a otimização taxa-distorção sejam realizadas utilizando-se quadros de quantização de comprimento maior que quadros de decomposição.

O coeficiente α_m e cada parâmetro de $\gamma(m)$ associados aos elementos do livro de estruturas são quantizados utilizando-se um quantizador escalar uniforme definido como [44]

$$x_q = I_x \times q(x), \quad \text{no qual } I_x = \left\lfloor \frac{x + \frac{q(x)}{2}}{q(x)} \right\rfloor, \quad (6.3)$$

em que x é qualquer parâmetro, x_q representa sua versão quantizada, $q(x)$ é o passo de quantização e I_x corresponde ao símbolo associado a x .

Os quantizadores de cada parâmetro são definidos por um intervalo dinâmico e um número de níveis, restritos a potências de 2 de forma a serem indexados usando um número inteiro de bits. A amostra inicial da região de suporte do átomo n_m^s é quantizada com $\log_2(N_d)$ bits, em que N_d é o tamanho do quadro de decomposição. Isto é possível porque o processo de codificação, a ser detalhado mais adiante, pode ser realizado em quadros do tamanho do quadro de decomposição. O comprimento da região de suporte n_m^L é quantizado com $\lceil \log_2(L_{\text{máx}}) \rceil$ bits, em que $L_{\text{máx}}$ é o comprimento máximo encontrado no livro de estruturas. O parâmetro $L_{\text{máx}}$ deve ser transmitido como informação lateral para o decodificador. Decidiu-se codificar o tamanho do suporte em vez da amostra final do suporte, porque dentro do livro de estruturas a amostra final do suporte é uniformemente distribuída, enquanto que a probabilidade de existirem átomos de suporte pequeno é maior do que átomos de suporte grande. Isto permite a utilização de um codificador de entropia de modo a alocar menos bits a tamanhos de suporte pequeno e mais bits a tamanhos de suporte grande, resultando em menor taxa total. Veremos o efeito disto mais adiante na Seção 6.4. A frequência ξ_m é quantizada de acordo com a discretização da frequência do dicionário de senoides amortecidas da escala musical cromática. O número de bits alocados a ξ_m consiste de $\lceil \log_2(N_{\text{freq}}) \rceil$ bits, em que N_{freq} é número de frequências do dicionário. Com relação a estes três parâmetros n_m^s , n_m^L e ξ_m não há perda de informação com a quantização.

Os coeficientes α_m , os fatores de decaimento ρ_m e as fases ϕ_m são quantizados de acordo com um intervalo dinâmico definido por seus respectivos valores máximo

e mínimo dentre todos os elementos do livro de estruturas, ou seja,

$$q(x) = \frac{x_{\text{máx}} - x_{\text{mín}}}{2^{b_x} - 1}, \quad (6.4)$$

em que b_x é o número de bits alocados a x . Note que, neste caso, opta-se por utilizar um quantizador *mid-rise*. Para o quantizador do coeficiente fazem-se $\alpha_{\text{mín}} = 0$ e $\alpha_{\text{máx}} = \text{máx} \{ \alpha_m \}$, para o quantizador do decaimento fazem-se $\rho_{\text{mín}} = 0$ e $\rho_{\text{máx}} = \text{máx} \{ |\rho_m| \}$ e para o quantizador da fase fazem-se $\phi_{\text{mín}} = 0$ e $\phi_{\text{máx}} = 2\pi$. Dessa forma, é necessário codificar como informação adicional $\alpha_{\text{máx}}$ e $\rho_{\text{máx}}$, que dependem do conteúdo do livro de estruturas, do número de bits alocados aos parâmetros α_m , ρ_m e ϕ_m , da taxa de amostragem e dos tamanhos do sinal e dos quadros de decomposição e quantização.

No paradigma de compressão em questão, busca-se o compromisso entre o número de átomos da representação do sinal, a quantização de cada coeficiente e a escolha do dicionário (definido pelos quantizadores dos parâmetros do átomo, ver Seção 2.2.3—página 20), que resulta na menor distorção dada uma taxa desejada. Neste caso, o número de bits de um átomo senoidal amortecido é definido como

$$r = r_\alpha + r_\xi + r_\rho + r_\phi + r_{n^s} + r_{n^L}, \quad (6.5)$$

em que r_α , r_ξ , r_ρ , r_ϕ , r_{n^s} e r_{n^L} representam os números de bits alocados a α , ξ , ρ , ϕ , n^s e n^L , respectivamente. O total de bits gastos na compressão do sinal, com exceção da informação adicional, é de $r \times M$, em que M é o número de termos da reconstrução do sinal. Como as grandezas r_{n^s} , r_{n^L} e r_ξ são fixas, não variando em função da escolha do dicionário, a distorção total é expressa somente em função dos números de bits gastos em coeficientes, fatores de decaimento e fases, resultando em

$$d = \|\mathbf{x} - \tilde{\mathbf{x}}\|^2 = f(r_\alpha, r_\rho, r_\phi), \quad (6.6)$$

em que \mathbf{x} e $\tilde{\mathbf{x}}$ são respectivamente os sinais original e reconstruído; d é, neste caso, o erro quadrático.

A fim de se obter o melhor compromisso taxa-distorção, deve-se buscar as taxas $(r_\alpha, r_\rho, r_\phi)$ que minimizam a distorção total inserida pelo processo de codificação, dada uma quantidade de bits desejada. Para isso, utiliza-se o método de alocação ótima de bits apresentado na Seção 2.2 correspondente ao algoritmo 2.1, que minimiza a seguinte função-custo taxa-distorção:

$$J = d + \lambda M r, \quad (6.7)$$

em que M é o número de elementos do livro de estruturas. Obtêm-se as curvas

operacionais de taxa-distorção, calculando-se a distorção e a taxa para diferentes combinações de $(r_\alpha, r_\rho, r_\phi)$. Estas combinações são obtidas variando-se r_α de 1 a 16 bits, r_ρ de 1 a 22 bits e r_ϕ de 1 a 18 bits. Isto resulta em um total de 6336 pares taxa-distorção para a obtenção do fecho convexo. Os limites superiores foram determinados de acordo com as faixas dinâmicas dos parâmetros α , ρ e ϕ de modo a poder representá-los com erros suficientemente pequenos. A Figura 6.3 ilustra a curva operacional obtida através deste algoritmo de otimização taxa-distorção. Observa-se que cada ponto está associado a um multiplicador de Lagrange λ_i obtido em função do ângulo θ_i através do seguinte cálculo: $\lambda_i = \arctg(\theta_i)$.

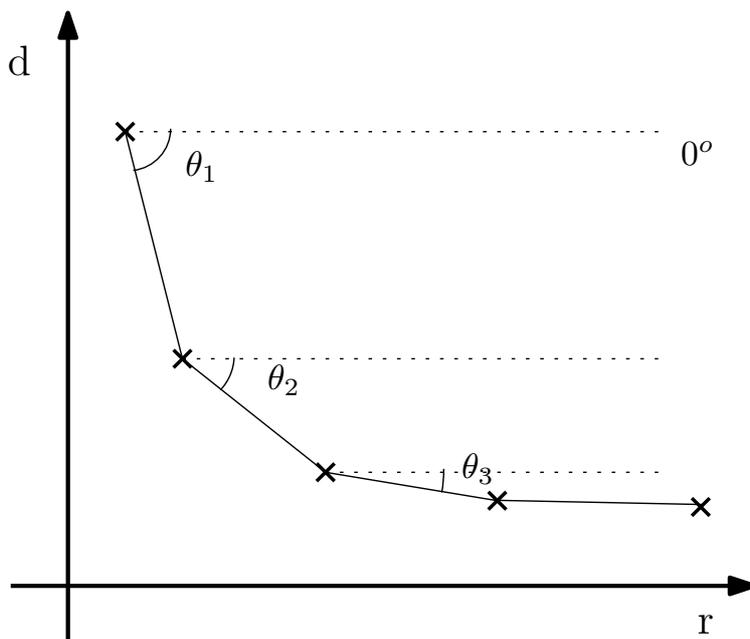


Figura 6.3: Curva operacional contendo os pontos ótimos em termos de taxa-distorção.

Geralmente, a alocação de bits se inicia definindo-se uma taxa desejada para o sinal, que é igualmente distribuída entre os quadros de quantização. A otimização taxa-distorção é realizada em cada quadro de forma independente tendo em vista a taxa alvo. Dessa maneira, a alocação de bits é localmente ótima a cada quadro, não sendo globalmente ótima em relação a todo o sinal. Para que a alocação seja globalmente ótima, o quadro de quantização deve possuir o mesmo comprimento do sinal. No entanto, como os sinais de áudio normalmente possuem longa duração, portanto inúmeras amostras, o processo de otimização taxa-distorção nessa situação torna-se impraticável em termos computacionais.

O que pode ser feito é obter as curvas operacionais de taxa-distorção para um conjunto de quadros de quantização e encontrar um multiplicador de Lagrange $\hat{\lambda}$, associado a um ângulo $\hat{\theta}$, que resulte em uma taxa \hat{r} próxima da taxa desejada, como ilustrado na Figura 6.4. No exemplo ilustrado, consultam-se através das curvas

operacionais dos quadros $q1$, $q2$ e $q3$ as taxas r_{q1} , r_{q2} e r_{q3} associadas a $\hat{\theta}$ que somadas resultam em \hat{r} . A consulta é feita da seguinte forma: se $\theta_{i-1} < \hat{\theta} \leq \theta_i$, retorna-se a taxa correspondente a θ_i . O $\hat{\theta}$ pode ser ajustado de forma iterativa até que \hat{r} atinja a taxa desejada. Com o advento e a venda em larga escala de processadores com múltiplos núcleos, bem como a evolução da programação paralela [127], o tratamento de quadros de quantização longos passa a ser viável, visto que a obtenção das curvas operacionais é realizada de forma independente em cada quadro. É importante ressaltar que o propósito desta análise é verificar a viabilidade de se implementar a otimização taxa distorção através de programação paralela em trabalhos futuros, apesar de este não ser o foco desta tese.

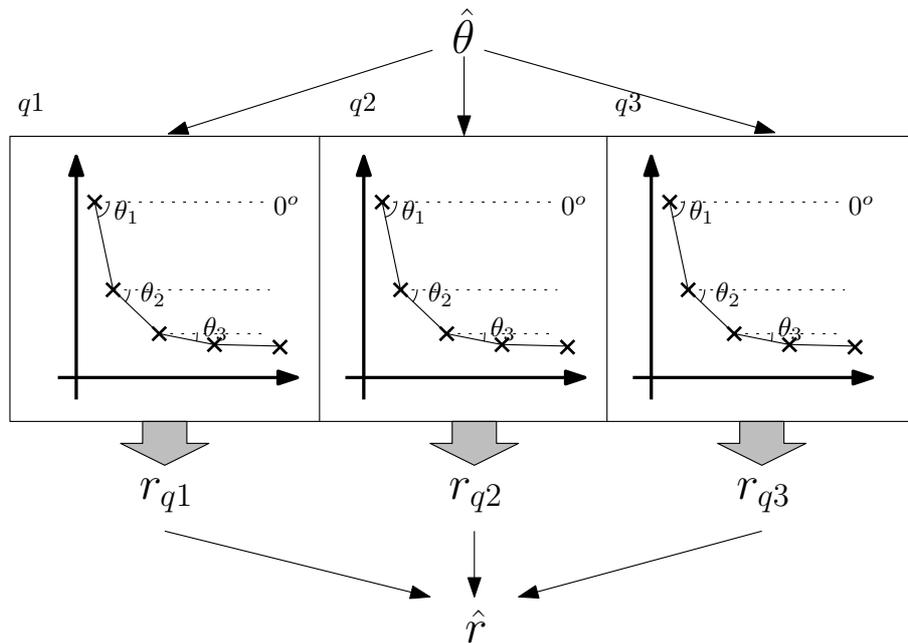


Figura 6.4: Otimização taxa-distorção conjunta de três quadros de quantização consecutivos.

6.2.2 Codificação

Na fase de codificação, forma-se o feixe de bits, ilustrado na Figura 6.5, contendo os códigos binários dos cabeçalhos do sinal e do quadro de quantização, bem como os símbolos dos coeficientes e parâmetros dos átomos. Os símbolos são representados por $(I_\alpha(i), I_\xi(i), I_\rho(i), I_\phi(i), I_{n^s}(i), I_{n^L}(i))$, onde i é o índice do átomo codificado e varia de 1 a M (número de elementos do livro de estruturas quantizado). De acordo com a Figura 6.5, a informação do feixe de bits é codificada e decodificada da esquerda para direita, de cima para baixo. Observe que os símbolos são separados por quadros de decomposição, sendo que a amostra inicial do suporte do átomo é responsável por indicar o quadro ao qual ele pertence. Os símbolos são ordenados em função do parâmetro, ou seja, primeiro enviam-se os símbolos do coeficiente,

seguidos pelos símbolos do decaimento, frequência, fase, amostra inicial do suporte e tamanho da região de suporte, nesta ordem. Utiliza-se o símbolo 0 do coeficiente para identificar o fim da lista de coeficientes e, desta forma, identificar implicitamente o número de átomos N_{qd}^k codificados no quadro de decomposição k . Isto é possível porque este símbolo não é usado na representação dos coeficientes, dado que átomos de coeficiente nulo não são codificados. Ao agregarem-se átomos em quadros de decomposição no lugar de quadros de quantização economizam-se bits na representação da amostra inicial do átomo, visto que em geral o tamanho do quadro de decomposição é menor que o do quadro de quantização.

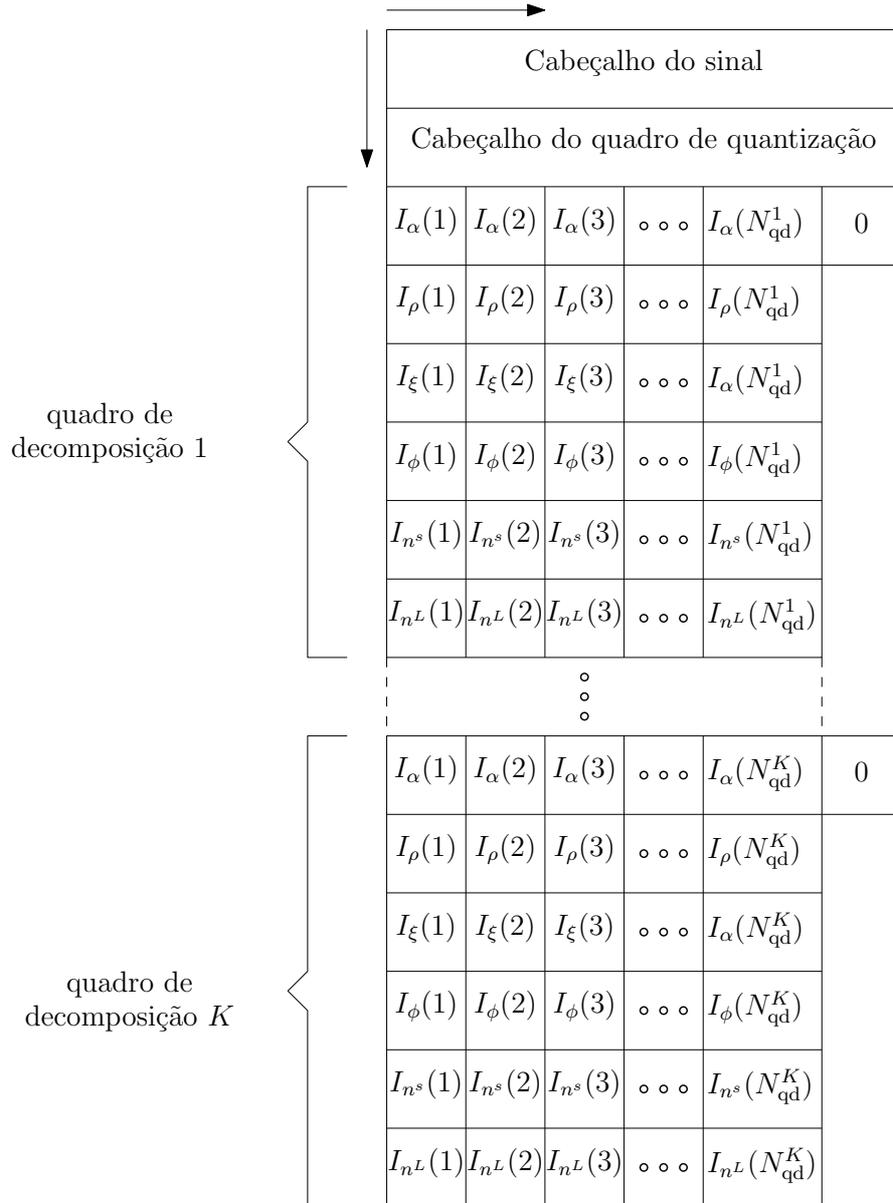


Figura 6.5: Organização dos símbolos dentro do feixe de bits.

O cabeçalho do sinal é composto pelas seguintes informações:

- fator de normalização do sinal;

- tamanho do sinal;
- taxa de amostragem;
- tamanho do quadro de decomposição;
- tamanho do quadro de quantização.

O cabeçalho do quadro de quantização é composto pelas seguintes informações:

- valor máximo da amplitude $\alpha_{\text{máx}}$;
- valor máximo do decaimento $\rho_{\text{máx}}$;
- valor máximo do tamanho do suporte $n_{\text{máx}}^L$;
- taxa de bits da amplitude α ;
- taxa de bits do decaimento ρ ;
- taxa de bits da fase ϕ .

Os valores mínimos da amplitude e do decaimento são iguais a zero e os valores mínimo e máximo da fase são 0 e 2π , respectivamente. Portanto, como estas informações já são conhecidas tanto no codificador quanto no decodificador, não é necessário enviá-las num cabeçalho.

6.2.3 Simulações Computacionais

Nesta seção, apresentam-se resultados de simulações computacionais a respeito do método de compressão de sinais áudio desenvolvido e proposto neste trabalho. Os sinais avaliados aqui são os mesmos utilizados na Seção 5.3 (página 87).

Ao serem decompostos através do MPSACr, os sinais são divididos em quadros de 2048 amostras, o que corresponde a 46,4 ms a uma taxa de amostragem de 44,1 kHz ou 42,7 ms a uma taxa de amostragem de 48 kHz, com saltos de 2048 amostras. Dessa forma, os quadros são ponderados por uma janela retangular e não há sobreposição entre quadros consecutivos. Para o critério de parada, adotam-se o valor de 60 dB para a razão sinal-ruído alvo (SNR_{alvo}) e o valor de 0,1095 para taxa de aproximação média desejada.

Inicialmente, analisa-se o comportamento das curvas operacionais que permitem a alocação ótima de bits entre os coeficientes e os parâmetros dos átomos. A Figura 6.6 apresenta a curva operacional do sinal *pianoA3* junto aos pares taxa-distorção calculados para diferentes combinações de taxas aplicadas ao coeficiente, ao decaimento e à fase. Em seguida, avalia-se o comportamento marginal da nuvem de pares taxa-distorção em função de cada parâmetro separadamente.

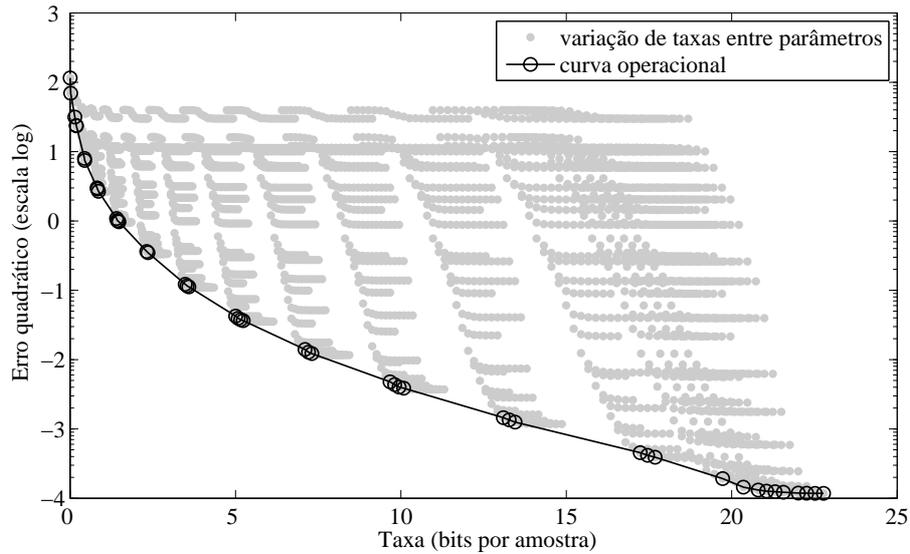


Figura 6.6: Curva operacional de taxa-distorção do sinal *pianoA3* e pares taxa-distorção calculados variando-se as taxas entre o coeficiente e os parâmetros do átomo. O círculo indica os pares taxa-distorção pertencentes à curva operacional.

A Figura 6.7 mostra os pares taxa-distorção agrupados de acordo com o número de bits gastos na amplitude. Observe que à medida que a taxa da amplitude aumenta, a nuvem de pontos se desloca para a direita e se torna mais esparsa, alcançando menores níveis de distorção.

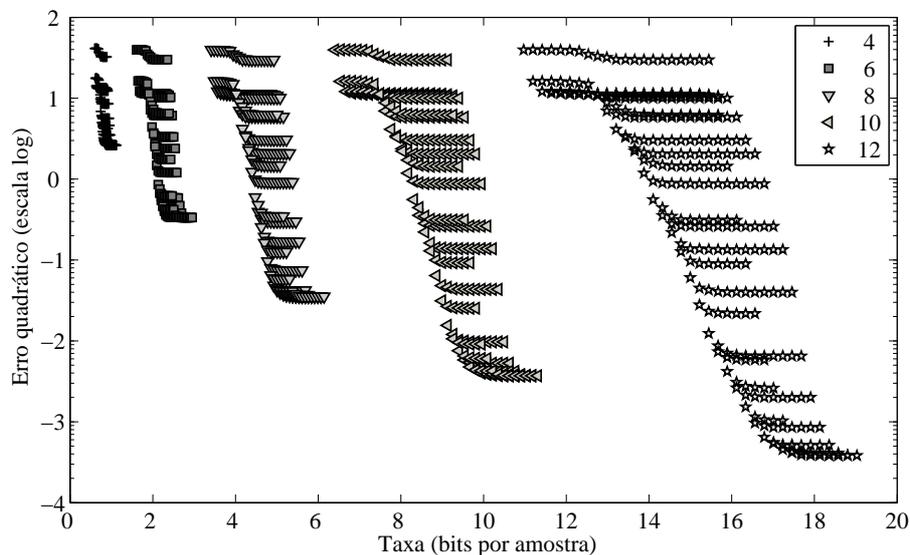


Figura 6.7: Comportamento taxa-distorção marginal em função da taxa aplicada à amplitude. O eixo das abcissas corresponde à taxa total de bits por amostra e o das ordenadas corresponde ao erro quadrático. Cada símbolo está associado a uma taxa de bits aplicada à amplitude e dentro de cada grupo de símbolos variam-se as taxas do decaimento e da fase.

A Figura 6.8 ilustra o comportamento taxa-distorção marginal em função da

taxa aplicada ao decaimento com a taxa de amplitude fixada em 8 bits. Note que para uma dada taxa alocada ao decaimento, a distorção decai em função da taxa total saturando em um patamar inferior. À proporção que a taxa do decaimento aumenta, a curva desloca-se para a direita e o patamar inferior decai, alcançando menores níveis de distorção.

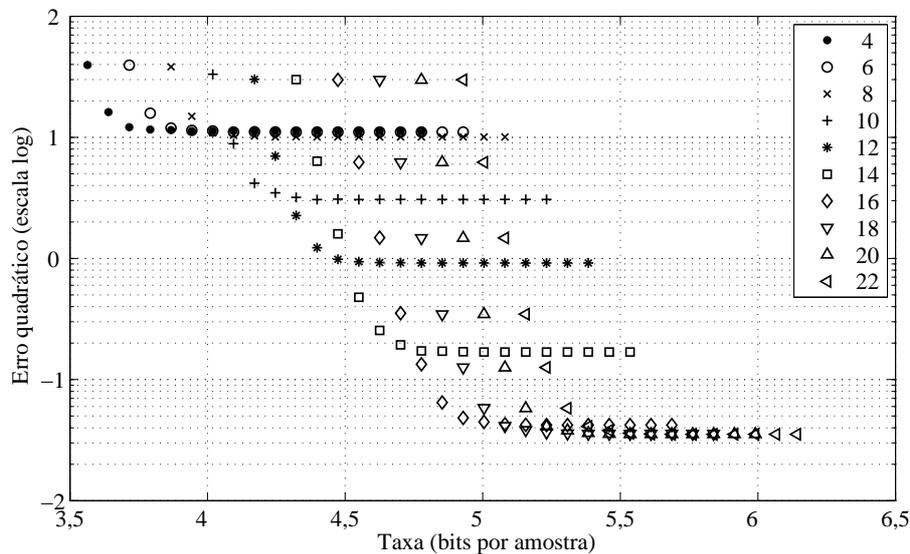


Figura 6.8: Comportamento taxa-distorção marginal em função da taxa aplicada ao decaimento com a taxa de amplitude fixada em 8 bits. O eixo das abcissas corresponde à taxa total de bits por amostra e o das ordenadas corresponde ao erro quadrático. Cada símbolo está associado a uma taxa de bits aplicada ao decaimento e dentro de cada grupo de símbolos varia-se a taxa da fase.

A Figura 6.9 ilustra o comportamento taxa-distorção marginal em função da taxa aplicada à fase com a taxa de amplitude fixada em 8 bits. Note que para uma dada taxa alocada à fase, à medida que a taxa total aumenta, a distorção inicia em um patamar superior e, em algum momento, passa a decair até saturar, enfim, em um patamar inferior. Por exemplo, para uma taxa de 6 bits no decaimento, a distorção inicia em um patamar de 10, a partir de 4,2 bits/amostra ela passa a decair e satura inferiormente em 0,1 a partir de 5,1 bits/amostra. Ao passo que a taxa do decaimento aumenta, a curva desloca-se para a direita e o patamar inferior decai alcançando menores níveis de distorção. Os comportamentos marginais destas nuvens de pares taxa-distorção apresentados nestes gráficos se repetem para os demais sinais: *fluteA4*, *celloA4* e *tirofinal*.

Na Figura 6.10, apresentam-se as curvas operacionais de taxa-distorção dos sinais *pianoA3*, *fluteA4*, *celloA4* e *tirofinal*. Para que estas curvas sejam comparáveis, o erro quadrático é normalizado em função da norma ao quadrado de cada sinal, por isso a taxa zero está associada à distorção igual a 1. Na Seção 5.3 (página 93), foi observado que quanto maior o número de padrões incoerentes com o dicionário

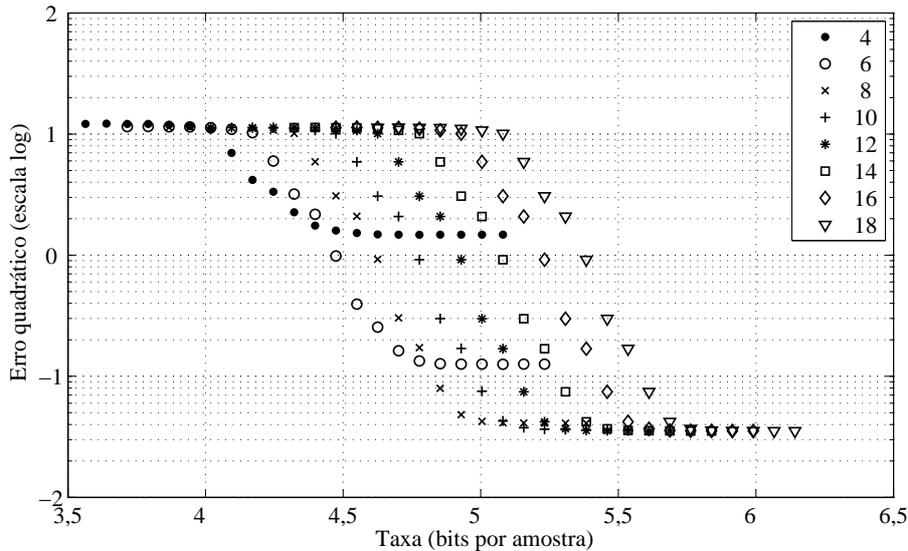


Figura 6.9: Comportamento taxa-distorção marginal em função da taxa aplicada à fase com a taxa de amplitude fixada em 8 bits. O eixo das abcissas corresponde a taxa total de bits por amostra e o das ordenadas corresponde ao erro quadrático. Cada símbolo está associado a uma taxa de bits aplicada à fase e dentro de cada grupo de símbolos varia-se a taxa do decaimento.

adotado (neste caso o dicionário de senoides amortecidas de escala musical cromática), maior é o número de átomos necessários para representar o sinal com um determinado nível de qualidade desejado. Isto se reflete no resultado da compressão. Ao compararem-se as curvas operacionais dos quatro sinais, é possível notar que a quantidade de bits para se atingir um mesmo nível de distorção aumenta na seguinte ordem: *pianoA3*, *fluteA4*, *celloA4* e *tirofinal*, que corresponde à mesma ordem em termos do número de átomos na representação para se atingir um mesmo nível de qualidade. Como observado na Seção 5.3 (página 87), isto ocorre devido à quantidade e à intensidade de padrões inerentes ao sinal que não são bem correlacionados com o dicionário de senoides amortecidas, fazendo com que seja necessário representá-lo com uma maior quantidade de bits para alcançar uma SNR desejada. O sinal *fluteA4* apresenta um som descorrelacionado com o dicionário, relativo ao sopro de flautista, de maior intensidade que o som inerente ao sinal *pianoA3*. O sinal *celloA4*, por sua vez, apresenta um som descorrelacionado com o dicionário, relativo à raspagem do arco na corda do violoncelo, ainda de maior intensidade que o som inerente ao sinal *flute A4*. Finalmente, o sinal *tirofinal* é o que apresenta maior intensidade de som não correlacionado com o dicionário de senoides amortecidas, por isso demanda a maior taxa de bits dentre estes sinais.

A Tabela 6.2 apresenta a distorção total do sinal e a soma das distorções individuais de cada átomo para diferentes taxas de compressão e sinais distintos. Através desta tabela, constata-se que a soma das distorções dos átomos se aproxima bem

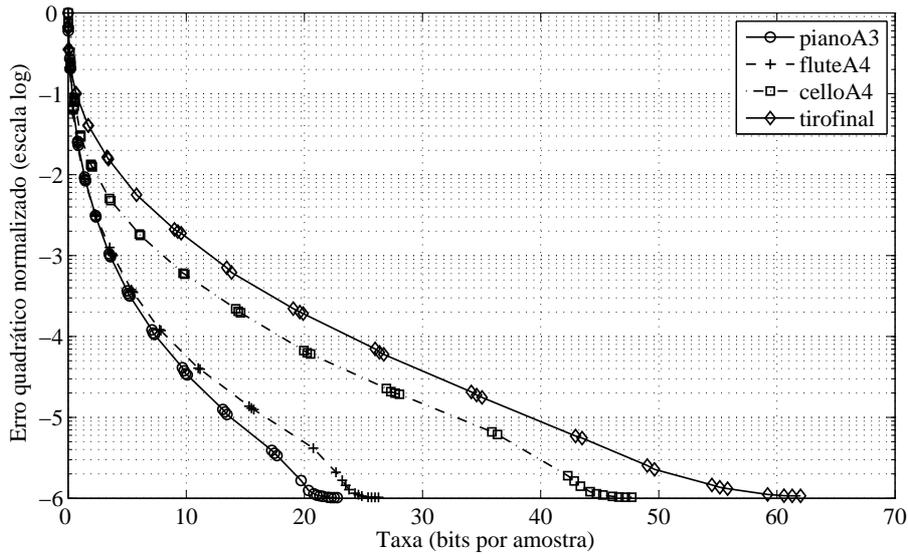


Figura 6.10: Curvas operacionais de taxa-distorção de diferentes sinais de áudio.

da distorção total do sinal, especialmente para taxas mais altas. Observa-se que somente a taxas de bits por amostra baixas obtém-se um erro percentual de 14 %; a taxas mais altas esse percentual não ultrapassa 5 %. Esta constatação é crucial para a validade da formulação matemática do problema de otimização taxa-distorção que será apresentada na seção a seguir.

6.3 Otimização Taxa-Distorção por Faixa de Energia Atômica

Nesta seção apresenta-se um esquema de otimização taxa-distorção, proposto e desenvolvido nesta tese, que realiza a alocação ótima de bits por entre os parâmetros dos átomos levando em consideração a relevância dos átomos na representação do sinal em termos de energia.

6.3.1 Motivação

Na Seção 2.2.3 (página 20), foi visto um método de otimização que busca o dicionário de ótimo compromisso taxa-distorção. Neste aspecto, o dicionário ótimo é definido pelas taxas de bits alocadas para o coeficiente e os parâmetros do átomo e estas taxas são aplicadas a todos os átomos do livro de estruturas. A aplicação deste método na compressão de sinais de áudio foi apresentada na Seção 6.2.1 (página 100).

Dessa maneira, os parâmetros de átomos de baixa energia são quantizados com a mesma intensidade que os parâmetros de átomos de alta energia. Contudo a influência de um átomo de baixa energia na distorção total é menor do que a influência de

Tabela 6.2: Distorção total do sinal e a soma das distorções individuais de cada átomo para diferentes taxas de compressão e sinais distintos.

Sinal	Taxa (b/a)	Distorção total	Soma das distorções dos átomos	Erro percentual (%)
<i>pianoA3</i>	0,44	7,456479	8,118704	8,88
	0,85	2,665812	2,840021	6,53
	2,31	0,362504	0,358818	1,01
	3,48	0,121762	0,120248	1,24
	5,00	0,042402	0,043187	1,85
	7,10	0,013980	0,013726	1,81
	9,67	0,004766	0,004759	0,14
	10,08	0,003868	0,003825	1,11
<i>fluteA4</i>	0,41	4,969812	5,681759	14,32
	0,79	2,042206	2,040408	0,08
	2,29	0,287508	0,296970	3,29
	3,52	0,110783	0,112769	1,79
	5,35	0,033262	0,033373	0,33
	7,74	0,010660	0,010663	0,02
	7,85	0,010378	0,010353	0,24
	11,02	0,003562	0,003511	1,43
<i>celloA4</i>	0,53	5,904323	6,107211	3,43
	1,04	2,034987	2,120965	4,22
	1,98	0,850147	0,849915	0,02
	3,50	0,333874	0,336071	0,65
	6,12	0,116904	0,116786	0,10
	9,73	0,039799	0,040665	2,17
	9,87	0,038789	0,039500	1,83
	14,60	0,012939	0,012938	0,01
<i>tirofinal</i>	0,66	326,323380	332,708899	1,95
	3,3	52,911736	52,459096	0,85
	9,57	6,030283	5,980756	0,82
	13,83	1,969808	2,001819	1,62
	19,89	0,613050	0,614965	0,31
	26,71	0,191488	0,190803	0,35

um átomo de alta energia.

Tendo em vista esta constatação, é possível obter um melhor desempenho taxa-distorção do codificador ao se utilizar uma ferramenta capaz de alocar os bits de maneira ótima por entre os parâmetros dos átomos, discriminando os átomos através de sua energia. Dessa forma, os bits são distribuídos por diferentes faixas de amplitude, podendo-se gastar mais bits na representação de parâmetros de átomos de maior energia em detrimento dos átomos de mais baixa energia.

A contrapartida desta abordagem consiste em um aumento na sobrecarga de

informação lateral a fim de indicar as taxas usadas em cada faixa de amplitude. Entretanto, no esquema de otimização taxa-distorção desenvolvido neste trabalho, este fardo é evitado através de um aumento na complexidade no decodificador, como será visto mais adiante.

6.3.2 Problema de Otimização Taxa-Distorção

Antes de descrever detalhes a respeito do procedimento de alocação ótima de bits, é necessário conhecer o problema a ser resolvido. A seguir, apresenta-se a formulação matemática do problema de otimização taxa-distorção agregando os átomos em diferentes faixas de amplitude.

Considere i o índice da faixa de amplitude variando em ordem crescente de energia, onde $i = \{1, 2, \dots, N_{\text{amp}}\}$ sendo N_{amp} o número total de faixas de amplitude. Dessa forma, o índice $i = 1$ está associado à faixa de menor energia e $i = N_{\text{amp}}$ está associado à de maior energia. Além disso, cada faixa de amplitude possui M_i átomos.

Define-se a taxa de bits de um átomo da faixa de amplitude i como

$$r(i) = r_\alpha + r_{\nu_1}(i) + \dots + r_{\nu_t}(i) + \dots + r_{\nu_T}(i), \quad (6.8)$$

em que r_α representa a taxa de bits alocada aos coeficientes α e $r_{\nu_t}(i)$ é a quantidade de bits alocados ao parâmetro ν_t na faixa de amplitude i , tal que $\gamma = \{\nu_t\}_{t=1, \dots, T}$, e T é o número de parâmetros que definem o átomo. Observe que r_α não varia em função da faixa, sendo a mesma para todo o livro de estruturas.

A distorção do átomo j na faixa de amplitude i é expressa como função das taxas de bits dos coeficientes e dos parâmetros, resultando em

$$d_{i,j} = f_j(r_\alpha, r_\gamma(i)), \quad (6.9)$$

em que $r_\gamma(i) = \{r_{\nu_1}(i), r_{\nu_2}(i), \dots, r_{\nu_T}(i)\}$.

Com o propósito de obter o melhor compromisso taxa-distorção, deve-se buscar as taxas $\{r_\alpha, \{r_\gamma(i)\}_{i=1,2,\dots,N_{\text{amp}}}\}$ que minimizam a distorção total inserida no processo de codificação para uma dada taxa alvo r_{alvo} . A solução é obtida através da resolução do seguinte problema de otimização [45]:

$$\begin{aligned} \underset{r_\alpha, \{r_\gamma(i)\}_{i=1,2,\dots,N_{\text{amp}}}}{\text{mín}} \quad & \sum_{i=1}^{N_{\text{amp}}} \sum_{j=1}^{M_i} d_{i,j} \\ \text{s. a.} \quad & \sum_{i=1}^{N_{\text{amp}}} M_i r(i) = r_{\text{alvo}}, \end{aligned} \quad (6.10)$$

em que M_i é o número de átomos na faixa de amplitude i .

O problema de minimização irrestrita correspondente ao problema em (6.10) é dado por [45–47]

$$\min_{\lambda, r_\alpha, \{r_\gamma(i)\}_{i=1,2,\dots,N_{\text{amp}}}} J = \sum_{i=1}^{N_{\text{amp}}} \sum_{j=1}^{M_i} d_{i,j} + \lambda \left(\left(\sum_{i=1}^{N_{\text{amp}}} M_i r(i) \right) - r_{\text{alvo}} \right), \quad (6.11)$$

λ é o multiplicador de Lagrange. É importante observar que esta formulação matemática possui uma condição *sine qua non*:

$$d_{\text{tot}} = \sum_{i=1}^{N_{\text{amp}}} \sum_{j=1}^{M_i} d_{i,j}, \quad (6.12)$$

ou seja, a distorção total do sinal d_{tot} deve ser igual à soma das distorções de cada átomo. Esta condição é considerada atendida para a validade do esquema de otimização taxa-distorção desenvolvido neste trabalho e apresentado a seguir. Na Seção 6.2.3 (página 106), sua validade foi comprovada de maneira empírica.

A distorção de um átomo j na faixa de amplitude i é expressa por

$$d_{i,j} = \int_{t_s}^{t_e} [\alpha(i,j) g_{\gamma(i,j)}(t) - \alpha_q(i,j) g_{\gamma_q(i,j)}(t)]^2 dt, \quad (6.13)$$

t_s e t_e são os tempos inicial e final, respectivamente, que definem o suporte temporal do átomo; $\alpha(i,j)$ e $g_{\gamma(i,j)}(t)$ são o coeficiente e a função geradora do átomo, respectivamente; $\alpha_q(i,j)$ e $g_{\gamma_q(i,j)}(t)$ são o coeficiente quantizado e a função geradora do átomo com os parâmetros quantizados. Para maior clareza, simplifica-se a notação retirando os índices i e j considerando-os implicitamente. Logo, tem-se que

$$\begin{aligned} d &= \int_{t_s}^{t_e} [\alpha g_\gamma(t) - \alpha g_{\gamma_q}(t) + \alpha g_{\gamma_q}(t) - \alpha_q g_{\gamma_q}(t)]^2 dt \\ &= \int_{t_s}^{t_e} [(\alpha - \alpha_q) g_{\gamma_q}(t) + \alpha (g_\gamma(t) - g_{\gamma_q}(t))]^2 dt \\ &= \int_{t_s}^{t_e} [(\alpha - \alpha_q)^2 (g_{\gamma_q}(t))^2 + \alpha^2 (g_\gamma(t) - g_{\gamma_q}(t))^2 + \\ &\quad 2(\alpha - \alpha_q) g_{\gamma_q}(t) \alpha (g_\gamma(t) - g_{\gamma_q}(t))] dt \\ &= (\alpha - \alpha_q)^2 \int_{t_s}^{t_e} (g_{\gamma_q}(t))^2 dt + \alpha^2 \int_{t_s}^{t_e} (g_\gamma(t) - g_{\gamma_q}(t))^2 dt + \\ &\quad 2\alpha(\alpha - \alpha_q) \int_{t_s}^{t_e} g_{\gamma_q}(t) (g_\gamma(t) - g_{\gamma_q}(t)) dt. \end{aligned} \quad (6.14)$$

Vale lembrar que os átomos possuem norma unitária, portanto

$$\begin{aligned} \int_{t_s}^{t_e} (g_\gamma(t))^2 dt &= 1 \\ \int_{t_s}^{t_e} (g_{\gamma_q}(t))^2 dt &= 1. \end{aligned} \quad (6.15)$$

Além disso, assume-se que o erro de quantização do átomo $(g_\gamma(t) - g_{\gamma_q}(t))$ é decorrelacionado no tempo com $g_{\gamma_q}(t)$, logo

$$\int_{t_s}^{t_e} g_{\gamma_q}(t) (g_\gamma(t) - g_{\gamma_q}(t)) dt \approx 0. \quad (6.16)$$

Substituindo-se

$$\begin{cases} \int_{t_s}^{t_e} (g_\gamma(t) - g_{\gamma_q}(t))^2 dt &= d_g \\ \alpha - \alpha_q &= \Delta\alpha \end{cases}$$

em (6.14), tem-se que

$$\begin{aligned} d &= (\Delta\alpha)^2 + \alpha^2 d_g + 2\alpha(\Delta\alpha) - 2\alpha(\Delta\alpha) \\ &= d_\alpha + \alpha^2 d_g. \end{aligned} \quad (6.17)$$

Logo, a distorção de um átomo equivale à soma entre o quadrado do erro de quantização da amplitude e a distorção do átomo ponderada pelo quadrado da amplitude. O problema (6.11) pode ser rescrito da seguinte forma

$$\begin{aligned} \min_{\lambda, r_\alpha, \{r_\gamma(i)\}_{i=1,2,\dots,N_{\text{amp}}}} J &= \sum_{n=1}^M d_\alpha(n) + \sum_{i=1}^{N_{\text{amp}}} \sum_{j=1}^{M_i} \alpha^2(i, j) d_g(i, j) + \\ &+ \lambda \left(\left(\sum_{i=1}^{N_{\text{amp}}} M_i r(i) \right) - r_{\text{alvo}} \right), \end{aligned} \quad (6.18)$$

em que M é o número total de elementos no livro de estruturas. Considerando que α^2 e d_g são decorrelacionados, tem-se que

$$\begin{aligned} \min_{\lambda, r_\alpha, \{r_\gamma(i)\}_{i=1,2,\dots,N_{\text{amp}}}} J &= \sum_{n=1}^M d_\alpha(n) + \sum_{i=1}^{N_{\text{amp}}} M_i \overline{\alpha^2}(i) \overline{d_g}(i) + \\ &+ \lambda \left(\left(\sum_{i=1}^{N_{\text{amp}}} M_i r(i) \right) - r_{\text{alvo}} \right), \end{aligned} \quad (6.19)$$

em que $\overline{\alpha^2}(i)$ e $\overline{d_g}(i)$ são a amplitude quadrática média e a distorção média dos

átomos da faixa de amplitude i , respectivamente. Por fim, tem-se que

$$\begin{aligned} \min_{\lambda, r_\alpha, \{r_\gamma(i)\}_{i=1,2,\dots,N_{\text{amp}}}} J = & \sum_{n=1}^M d_\alpha(n) + \lambda M r_\alpha + \\ & + \sum_{i=1}^{N_{\text{amp}}} M_i \overline{\alpha^2}(i) \overline{d}_g(i) + \lambda \left(\sum_{i=1}^{N_{\text{amp}}} M_i r_g(i) \right) - \lambda r_{\text{alvo}}, \end{aligned} \quad (6.20)$$

onde $r_g(i) = r_{\nu_1}(i) + \dots + r_{\nu_t}(i) + \dots + r_{\nu_T}(i)$.

6.3.3 Alocação Ótima de Bits

Posto o problema de otimização taxa-distorção baseado na agregação dos átomos em faixas de energia, propõe-se uma maneira prática de resolvê-lo. Na Figura 6.11, observa-se o diagrama de blocos do esquema de otimização taxa-distorção, que se baseia na relevância energética dos átomos, dentro de um contexto de codificação.

Considere o livro de estruturas $\{A \cup \Gamma\}$, no qual $A = \{\alpha(n)\}_{n=1}^M$ é o conjunto dos coeficientes e $\Gamma = \{\gamma(n)\}_{n=1}^M$ o conjunto dos parâmetros dos átomos, sendo M o número de átomos do livro de estruturas. Tendo o livro de estruturas em mãos, o processo de codificação se inicia com a quantização dos coeficientes dos átomos de A , obtendo-se A_q que é composto pelos coeficientes quantizados não nulos. Logo, os átomos com coeficientes nulos passam a ser desconsiderados. Com base nos coeficientes não nulos, os átomos são agregados em diferentes faixas de energia, ou subconjuntos $\{\Gamma_{\text{amp},i}\}_{i=1}^{N_{\text{amp}}}$, N_{amp} é o número de faixas de energia. A quantidade de faixas é associada ao nível de quantização aplicado aos coeficientes. Portanto, dado um nível de quantização dos coeficientes, realiza-se a alocação ótima de bits entre os parâmetros dos átomos em cada faixa de energia separadamente. Por fim, calcula-se o valor do custo taxa-distorção. Este processo se repete para diferentes níveis desejados de quantização do coeficiente.

Em seguida, selecionam-se as taxas que resultaram no menor valor de custo taxa-distorção. Calcula-se a taxa total e verifica-se o atendimento da taxa alvo. Caso positivo, as taxas ótimas obtidas são utilizadas na quantização dos parâmetros dos átomos. Os símbolos associados ao coeficiente e aos parâmetros quantizados são tratados por um codificador por entropia [44, 128, 129]. Finalmente, o feixe de bits é formatado e enviado ao decodificador. Inserido neste feixe de bits, é enviado como informação lateral o nível de quantização dos coeficientes. Se a taxa alvo não for atendida, reinicia-se o processo alterando-se o multiplicador de Langrange que compõe a função custo taxa-distorção.

No decodificador, os coeficientes dos átomos são decodificados primeiramente utilizando a informação lateral de nível de quantização. Com base nos coeficientes, os átomos são agregados em diferentes faixas de energia. Em seguida, realiza-se a

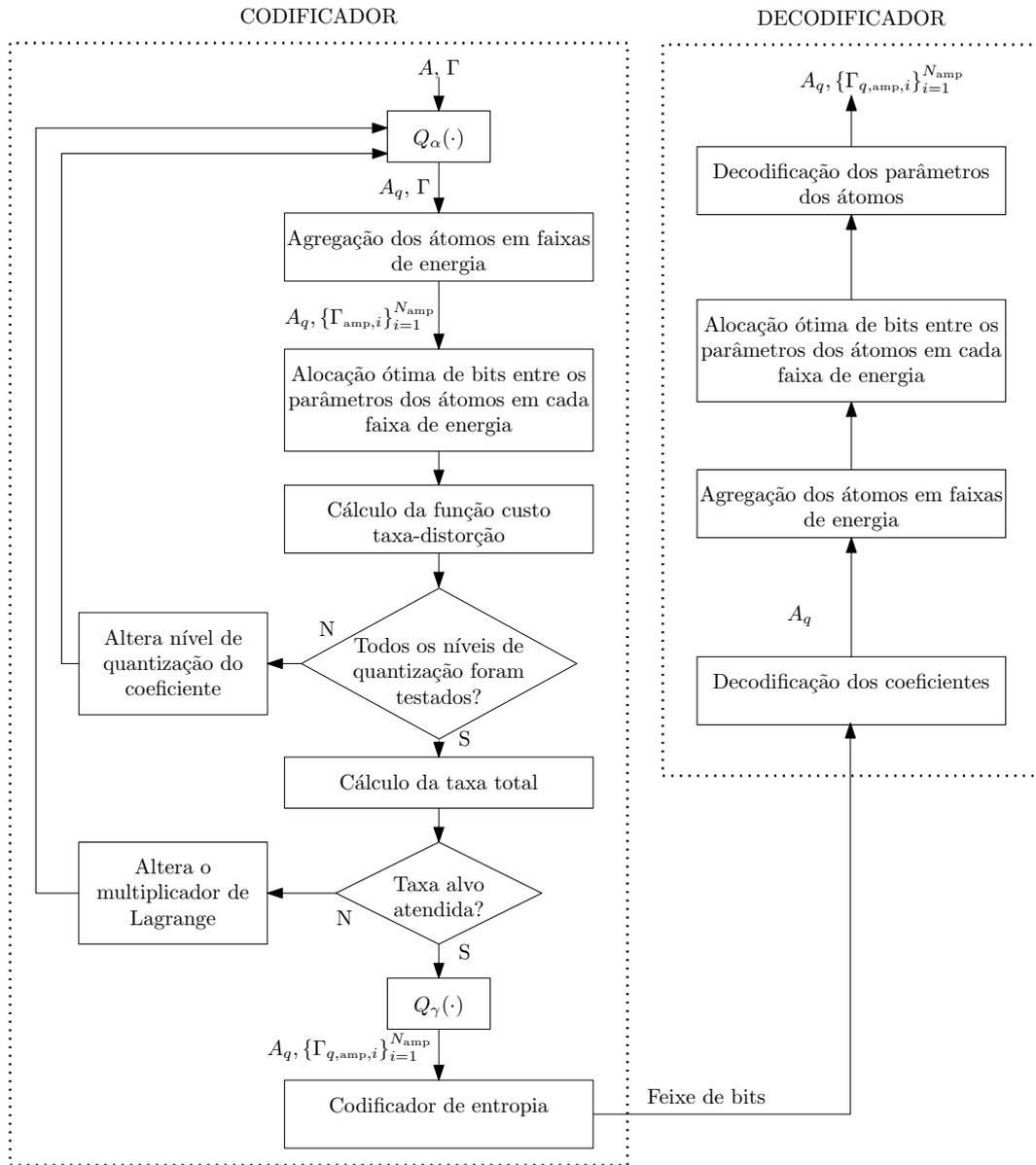


Figura 6.11: Esquema de otimização taxa-distorção considerando a relevância dos átomos na representação do sinal em termos de energia, inserida em um contexto de codificação.

alocação ótima de bits entre os parâmetros dos átomos em cada faixa de energia. Por fim, decodificam-se os parâmetros dos átomos.

É importante destacar que o procedimento de alocação ótima de bits entre os parâmetros dos átomos é realizado tanto no codificador quanto no decodificador com o propósito de evitar a sobrecarga de informação lateral com a indicação das taxas ótimas definidas para cada faixa de energia.

6.3.4 Função-Custo Taxa-Distorção

A função-custo na qual o método de otimização taxa-distorção em questão se baseia é definida por

$$J = \sum_{n=1}^M d_{\alpha}(n) + \lambda M h_{\alpha} + \sum_{i=1}^{N_{\text{amp}}} M_i \overline{\alpha^2}(i) \overline{d}_g(i) + \lambda \sum_{i=1}^{N_{\text{amp}}} M_i h_g(i). \quad (6.21)$$

Deve-se notar que existem algumas diferenças entre esta função e aquela que consta no problema descrito em (6.20). Em vez de taxas r_{α} e $r_g(i)$, utilizam-se a entropia da amplitude h_{α} e a soma das entropias dos parâmetros do átomo, $h_g(i) = h_{\nu_1}(i) + \dots + h_{\nu_t}(i) + \dots + h_{\nu_T}(i)$, em que $h_{\nu_t}(i)$ é a entropia do parâmetro ν_t na faixa de amplitude i . Dessa maneira, é possível realizar no decodificador a alocação de bits entre os parâmetros dos átomos por faixa de amplitude sem a necessidade de se enviar as taxas como informação lateral. A entropia é calculada somente em função das funções de probabilidade dos parâmetros, que são conhecidas tanto pelo codificador quanto pelo decodificador.

Além disso, a taxa alvo é desconsiderada na função custo taxa-distorção. Na prática, usa-se o multiplicador de Lagrange λ , ajustando-o de modo que a taxa final de codificação se aproxime da taxa alvo, como ilustrado no diagrama de blocos da Figura 6.11. Recomenda-se realizar o ajuste de λ através de um método de busca por bisseção [130], por exemplo.

Quantização dos Coeficientes e Definição das Faixas de Energia

Os coeficientes dos átomos são quantizados usando-se um quantizador escalar uniforme descrito por

$$\alpha_q = I_{\alpha} q_{\alpha}, \quad \text{com } I_{\alpha} = \left\lfloor \frac{\alpha + \frac{q_{\alpha}}{2}}{q_{\alpha}} \right\rfloor, \quad (6.22)$$

em que α é o coeficiente, α_q representa sua versão quantizada, q_{α} é o passo de quantização e I_{α} corresponde ao símbolo associado a α_q . O coeficiente é quantizado de acordo com um intervalo dinâmico definido por seus respectivos valores máximo

e mínimo dentre todos os elementos do livro de estruturas, ou seja,

$$q_\alpha = \frac{L_\alpha}{N_{I_\alpha} - 1}, \quad (6.23)$$

em que $L_\alpha = (\alpha_{\text{máx}} - \alpha_{\text{mín}})$ e N_{I_α} é o número de símbolos. Neste caso, opta-se por um quantizador *mid-rise*, dado que é possível garantir que os coeficientes sejam não negativos. Nos casos em que os coeficientes são negativos e positivos, além de ser fundamental que o valor nulo seja representado precisamente, recomenda-se a utilização de um quantizador *mid-thread* [44].

A distorção de um coeficiente é dada por

$$d_\alpha(n) = (\alpha(n) - \alpha_q(n))^2, \text{ onde } n = 1, 2, \dots, M. \quad (6.24)$$

Ademais, é necessário fornecer a função densidade de probabilidade do coeficiente para o cálculo da entropia. Neste caso, a entropia do coeficiente é definida como

$$h_\alpha = - \sum_{k=1}^{N_{I_\alpha}} p_\alpha(I_\alpha(k)) \log_2(p_\alpha(I_\alpha(k))) \text{ (bits/símbolo)}, \quad (6.25)$$

em que $p_\alpha(\cdot)$ corresponde à função densidade de probabilidade do coeficiente α . Tendo realizado a quantização dos coeficientes e o cálculo da entropia, é possível calcular a parcela da função custo taxa-distorção da equação (6.21) relativa aos coeficientes dos átomos

$$J_\alpha = \sum_{n=1}^M d_\alpha(n) + \lambda M h_\alpha. \quad (6.26)$$

A quantização e a codificação dos demais parâmetros dos átomos dependem da quantização dos coeficientes, pois as faixas de energia são definidas de acordo com o número de símbolos empregados na codificação dos coeficientes dos átomos. O número de faixas é dado por $N_{amp} = \log_2(N_{I_\alpha})$; sendo assim, para que N_{amp} seja um número inteiro, é preciso que N_{I_α} seja potência de 2. Em princípio, as faixas poderiam ser uniformemente espaçadas; no entanto, opta-se por um espaçamento não-uniforme, de modo que as faixas mais largas sejam as de maior energia. Dessa forma, a distribuição dos átomos entre as faixas de amplitude fica mais balanceada, posto que, em geral, a decomposição atômica resulta em uma representação do sinal formada por poucos átomos de alta energia e muitos de baixa energia. Os limites

das faixas seguem a seguinte regra de formação:

$$l_\alpha(i) = \begin{cases} 0, & \text{para } i = 0 \\ \frac{L_\alpha}{2^{N_{\text{amp}}-i}}, & \text{para } i = 1, 2, \dots, N_{\text{amp}}. \end{cases} \quad (6.27)$$

Além disso, por simplificação, opta-se por utilizar a média aritmética entre os limites de cada faixa de energia como representante da média dos coeficientes, ou seja, para cada faixa i tem-se

$$\bar{\alpha}(i) = \frac{l(i) + l(i-1)}{2}, \text{ onde } i = 1, 2, \dots, N_{\text{amp}}. \quad (6.28)$$

A Figura 6.12 ilustra a construção das faixas de energia com espaçamento não-uniforme e os representantes médios.

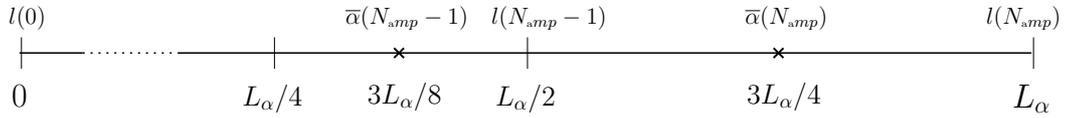


Figura 6.12: Construção de faixas de energia com espaçamento não-uniforme e seus representantes médios.

Alocação Ótima de Bits por Faixa de Energia entre os Parâmetros dos Átomos

A parcela da função custo taxa-distorção da equação (6.21) relativa aos parâmetros dos átomos é expressa por

$$J_g = \sum_{i=1}^{N_{\text{amp}}} M_i \bar{\alpha}^2(i) \bar{d}_g(i) + \lambda \sum_{i=1}^{N_{\text{amp}}} M_i h_g(i), \quad (6.29)$$

$h_g(i) = h_{\nu_1}(i) + \dots + h_{\nu_t}(i) + \dots + h_{\nu_T}(i)$ e $\bar{\alpha}^2(i) = \bar{\alpha}^2(i)$, com $\bar{\alpha}(i)$ definido na equação (6.28). As entropias dos parâmetros do átomo são calculadas da seguinte

forma

$$\begin{aligned}
h_{\nu_1}(i) &= - \sum_{k=1}^{N_{I_{\nu_1}}} p_{\nu_1}(I_{\nu_1}(k)) \log_2(p_{\nu_1}(I_{\nu_1}(k))) \text{ (bits/símbolo),} \\
h_{\nu_2}(i) &= - \sum_{k=1}^{N_{I_{\nu_2}}} p_{\nu_2}(I_{\nu_2}(k)) \log_2(p_{\nu_2}(I_{\nu_2}(k))) \text{ (bits/símbolo),} \\
&\vdots \\
h_{\nu_t}(i) &= - \sum_{k=1}^{N_{I_{\nu_t}}} p_{\nu_t}(I_{\nu_t}(k)) \log_2(p_{\nu_t}(I_{\nu_t}(k))) \text{ (bits/símbolo),} \\
&\vdots \\
h_{\nu_T}(i) &= - \sum_{k=1}^{N_{I_{\nu_T}}} p_{\nu_T}(I_{\nu_T}(k)) \log_2(p_{\nu_T}(I_{\nu_T}(k))) \text{ (bits/símbolo),}
\end{aligned}$$

em que $I_{\nu_t}(k)$ é o k -ésimo símbolo associado ao t -ésimo parâmetro, $N_{I_{\nu_t}}$ é o número de símbolos e $p_{\nu_t}(\cdot)$ é a função densidade de probabilidade de ν_t . Por sua vez, a distorção média por átomo \bar{d}_g na faixa de energia i é descrita por

$$\begin{aligned}
\bar{d}_g(i) &= \int \int \dots \int \dots \int \\
&\quad d_g(\nu_1, \nu_2, \dots, \nu_t, \dots, \nu_T, q_{\nu_1}, q_{\nu_2}, \dots, q_{\nu_t}, \dots, q_{\nu_T}) \\
&\quad p_\gamma(\nu_1, \nu_2, \dots, \nu_t, \dots, \nu_T) d\nu_1 d\nu_2 \dots d\nu_t \dots d\nu_T, \quad (6.30)
\end{aligned}$$

em que q_{ν_t} é o passo de quantização de ν_t e $p_\gamma(\cdot)$ é a função de probabilidade conjunta dos parâmetros do átomo.

Por simplificação, desenvolve-se a equação (6.30) somente para dois parâmetros ν_1 e ν_2 . Além disso, considera-se a utilização de quantizadores uniformes. Dessa forma, tem-se que

$$\begin{aligned}
\bar{d}_g(i) &= \int \int d_g(\nu_1, \nu_2, q_{\nu_1}, q_{\nu_2}) p_\gamma(\nu_1, \nu_2) d\nu_1 d\nu_2 \\
&= \sum_{k=0}^{N_{\nu_1}-1} \int_{kq_{\nu_1}}^{(k+1)q_{\nu_1}} \sum_{r=0}^{N_{\nu_2}-1} \int_{r q_{\nu_2}}^{(r+1)q_{\nu_2}} d_g(\nu_1, \nu_2, q_{\nu_1}, q_{\nu_2}) p_\gamma(\nu_1, \nu_2) d\nu_1 d\nu_2 \quad (6.31)
\end{aligned}$$

em que N_{ν_1} e N_{ν_2} são os números de níveis de quantização e q_{ν_1} e q_{ν_2} correspondem aos passos de quantização de ν_1 e ν_2 , respectivamente. Substituindo-se

$$\begin{cases} \nu_1 &= \Delta\nu_1 + \left(k + \frac{1}{2}\right) q_{\nu_1} \\ \nu_2 &= \Delta\nu_2 + \left(r + \frac{1}{2}\right) q_{\nu_2}, \end{cases}$$

sendo $\Delta\nu_1$ e $\Delta\nu_2$ os erros de quantização de ν_1 e ν_2 , tem-se

$$\begin{aligned}
\overline{d}_g(i) &= \sum_{k=0}^{N_{\nu_1}-1} \int_{-\frac{q_{\nu_1}}{2}}^{\frac{q_{\nu_1}}{2}} \sum_{r=0}^{N_{\nu_2}-1} \int_{-\frac{q_{\nu_2}}{2}}^{\frac{q_{\nu_2}}{2}} \\
&\quad d_g \left(\left\{ \Delta\nu_1 + \left(k + \frac{1}{2} \right) q_{\nu_1} \right\}, \left\{ \Delta\nu_2 + \left(r + \frac{1}{2} \right) q_{\nu_2} \right\}, \Delta\nu_1, \Delta\nu_2 \right) \\
&\quad p_\gamma \left(\left\{ \Delta\nu_1 + \left(k + \frac{1}{2} \right) q_{\nu_1} \right\}, \left\{ \Delta\nu_2 + \left(r + \frac{1}{2} \right) q_{\nu_2} \right\} \right) d(\Delta\nu_1) d(\Delta\nu_2), \\
&= \sum_{k=0}^{N_{\nu_1}-1} \sum_{r=0}^{N_{\nu_2}-1} \int_{-\frac{q_{\nu_1}}{2}}^{\frac{q_{\nu_1}}{2}} \int_{-\frac{q_{\nu_2}}{2}}^{\frac{q_{\nu_2}}{2}} \\
&\quad d_g \left(\left\{ \Delta\nu_1 + \left(k + \frac{1}{2} \right) q_{\nu_1} \right\}, \left\{ \Delta\nu_2 + \left(r + \frac{1}{2} \right) q_{\nu_2} \right\}, \Delta\nu_1, \Delta\nu_2 \right) \\
&\quad p_\gamma \left(\left\{ \Delta\nu_1 + \left(k + \frac{1}{2} \right) q_{\nu_1} \right\}, \left\{ \Delta\nu_2 + \left(r + \frac{1}{2} \right) q_{\nu_2} \right\} \right) d(\Delta\nu_1) d(\Delta\nu_2).
\end{aligned} \tag{6.32}$$

É importante observar que a distorção média dos átomos por faixa de energia $\overline{d}_g(i)$ pode ser escrita em função dos passos de quantização dos parâmetros, bem como a entropia $h_g(i)$, pois os passos de quantização influenciam diretamente no número de símbolos. O número de símbolos do t -ésimo parâmetro ν_t é dado por

$$N_{I_{\nu_t}} = \left\lceil \frac{\hat{\nu}_t - \underline{\nu}_t}{q_{\nu_t}} \right\rceil, \tag{6.33}$$

em que $\underline{\nu}_t$ e $\hat{\nu}_t$ são valores mínimo e máximo de ν_t em todo o livro de estruturas, respectivamente, e

$$h_{\nu_t}(i) = - \sum_{k=1}^{N_{I_{\nu_t}}} p_{\nu_t}(I_{\nu_t}(k)) \log_2(p_{\nu_t}(I_{\nu_t}(k))) \text{ (bits/símbolo)}, \tag{6.34}$$

com

$$I_{\nu_t} = \left\lfloor \frac{\nu_t + \frac{q_{\nu_t}}{2}}{q_{\nu_t}} \right\rfloor. \tag{6.35}$$

Portanto, pode-se rescrever o custo taxa-distorção do átomo da equação (6.29) como

$$J_g(\lambda, q_\gamma(i)) = \sum_{i=1}^{N_{\text{amp}}} M_i \overline{\alpha}^2(i) \overline{d}_g(i, q_\gamma(i)) + \lambda \sum_{i=1}^{N_{\text{amp}}} M_i h_g(i, q_\gamma(i)), \tag{6.36}$$

em que $q_\gamma(i) = \{q_{\nu_1}(i), \dots, q_{\nu_T}(i)\}$. Os passos de quantização dos parâmetros do átomo variam em função da faixa de energia. Em cada faixa de energia, calcula-se $J_g(\lambda, q_\gamma(i))$ para diferentes combinações de $q_\gamma(i)$ e seleciona-se o conjunto $\{q_\gamma(i)\}_{i=1}^{N_{\text{amp}}}$

ótimo que minimiza $J_g(\lambda, q_\gamma(i))$.

A seguir, será vista a aplicação deste esquema de otimização taxa-distorção dentro de um contexto de codificação de sinais de áudio.

6.4 Aperfeiçoamento do Sistema de Compressão

Nesta seção, descreve-se o método de compressão de áudio resultante do aperfeiçoamento do sistema de compressão apresentado na Seção 6.2 (página 98). Na etapa de quantização, a otimização taxa-distorção passa a ser realizada por faixa de energia atômica, como apresentado na Seção 6.3, e na etapa de codificação introduz-se um codificador por entropia, mais especificamente o codificador aritmético.

A Figura 6.13 ilustra o diagrama de blocos do método de compressão proposto. Inicialmente, obtém-se a representação do sinal através do MPSACr. Ao fim da decomposição, o MPSACr provê o livro de estruturas contendo os coeficientes e parâmetros dos átomos. Em seguida, realiza-se a otimização taxa-distorção por faixa de energia atômica utilizando uma tabela de distorção média dos átomos pré-calculada para cada faixa de energia, dado um nível de quantização da amplitude. Definida a alocação ótima de bits entre os coeficientes e parâmetros dos átomos, quantiza-se o livro de estruturas produzindo símbolos que são tratados por um codificador aritmético e transmitidos ao decodificador. No decodificador, o feixe de bits é decodificado, gerando os símbolos. Estes, por sua vez, sofrem o processo de quantização inversa, produzindo o livro de estruturas quantizado. Enfim, com base neste livro de estruturas, reconstrói-se o sinal.

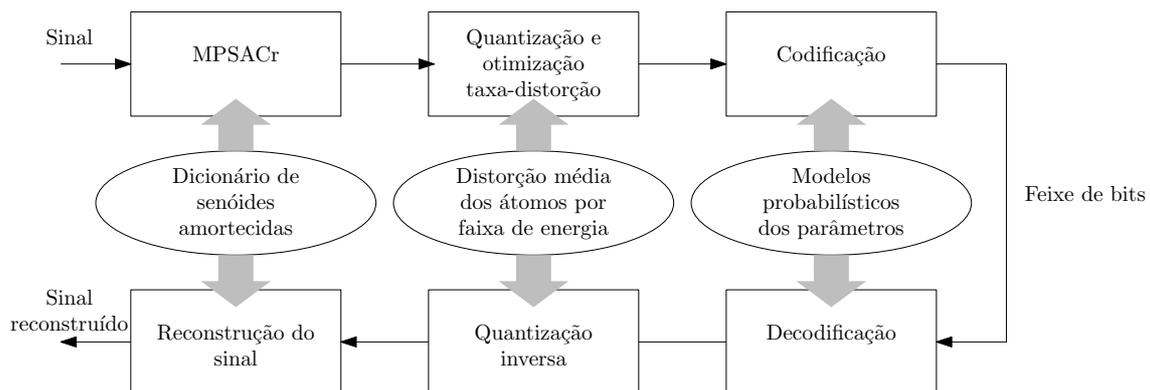


Figura 6.13: Compressão de sinais de áudio realizando a decomposição do sinal via MPSACr, a otimização taxa-distorção por faixa de energia atômica seguida de codificação aritmética.

Através deste aperfeiçoamento, alcançam-se menores taxas de bits/amostra para um mesmo nível de distorção. Além disso, a nova estratégia de otimização taxa-distorção permite uma alocação ótima de bits por faixa de energia sem sobrecarga

de informação lateral. Se a otimização taxa-distorção por faixa de energia fosse realizada através de curvas operacionais, haveria a necessidade de se transmitir ao decodificador as taxas ótimas usadas em cada faixa de energia como informação lateral, o que tenderia a diminuir muito a eficiência de compressão do método. Além disso, a mudança de estratégia de otimização taxa-distorção proporciona redução do custo computacional. A obtenção das curvas operacionais apresenta alto custo computacional, pois pressupõe o cálculo de pares taxa-distorção para diferentes taxas aplicadas aos parâmetros do átomo. No método de compressão aperfeiçoado, utiliza-se uma tabela de distorção média dos átomos pré-calculada que agiliza significativamente o processo de otimização taxa-distorção.

6.4.1 Quantização e Otimização Taxa-Distorção por Faixa de Energia Atômica

Considera-se aqui o livro de estruturas obtido através do MPSACr $\{A \cup \Gamma\}$, onde $A = \{\alpha(m)\}_{m=1}^M$ é o conjunto dos coeficientes e $\Gamma = \{\gamma(m)\}_{m=1}^M$ o conjunto dos parâmetros dos átomos, sendo M o número de átomos do livro de estruturas. Cada elemento do livro de estruturas é definido pelo par $(\alpha(m), \gamma(m))$, onde $\gamma(m) = (\rho_m, \xi_m, \phi_m, n_m^s, n_m^L)$ e $m = 1, 2, \dots, M$. Vale lembrar que o dicionário é composto por exponenciais moduladas e que ρ_m corresponde ao decaimento, ξ_m corresponde à frequência, ϕ_m denota a fase, n_m^s é amostra inicial do suporte temporal e n_m^L é o tamanho do suporte.

O Algoritmo 6.1 ilustra o pseudocódigo do procedimento de otimização taxa-distorção por faixa de energia no contexto de codificação via MPSACr. Primeiramente, definem-se o multiplicador de Lagrange λ e o conjunto de números de faixas de energia atômica Υ , onde cada elemento define o nível de quantização dos coeficientes. Neste caso, adota-se $\Upsilon = \{1, 2, \dots, 16\}$. Dado um número de faixas de energia N_{amp} :

1. define-se o número de símbolos de coeficiente;
2. define-se o coeficiente máximo dentro do livro de estruturas;
3. calcula-se o passo de quantização do coeficiente;
4. quantizam-se os coeficientes através de um quantizador escalar uniforme.
5. calculam-se as distorções dos coeficientes;
6. calcula-se a entropia do coeficiente;
7. calcula-se a parcela da função custo taxa-distorção referente ao coeficiente;

8. definem-se as faixas de energia não-uniformemente espaçadas;
9. agregam-se os átomos em faixas de energia;
10. calcula-se soma dos custos taxa-distorção mínimos referentes aos parâmetros dos átomos por faixa de energia;
11. calcula-se o custo taxa-distorção total;

Após repetir este procedimento para todos os números de faixas de energias N_{amp} pré-definidos, obtêm-se o número ótimo de faixas N_{amp}^* e os passos de quantização ótimos $q_\alpha^*, \{q_\rho^*(i), q_\phi^*(i)\}_{i=1}^{N_{\text{amp}}^*}$. Vale notar que q_ξ, q_{n^s} e q_{n^L} são pré-definidos e fixados, não variando em função da faixa de energia, por isso não se buscam passos de quantização ótimos para ξ, n^s e n^L .

Algoritmo 6.1 Otimização taxa-distorção por faixa de energia atômica.

Define-se λ

Define-se Υ

para $N_{\text{amp}} \in \Upsilon$ **faça**

$$N_{I_\alpha} = 2^{N_{\text{amp}}}$$

$$L_\alpha = \max_{\alpha \in A} \alpha$$

$$q_\alpha = \frac{L_\alpha}{N_{I_\alpha} - 1} \text{ (ver equação (6.23))}$$

para $n = 1 ; n \leq M ; n = n + 1$ **faça**

$$\alpha_q(n) = \left\lfloor \frac{\alpha(n) + \frac{q_\alpha}{2}}{q_\alpha} \right\rfloor q_\alpha \text{ (ver equação (6.22))}$$

$$d_\alpha(n) = (\alpha(n) - \alpha_q(n))^2 \text{ (ver equação (6.24))}$$

fim para

$$h_\alpha = - \sum_{k=1}^{N_{I_\alpha}} p_\alpha(I_\alpha(k)) \log_2(p_\alpha(I_\alpha(k))) \text{ (ver equação (6.25))}$$

$$J_\alpha = \sum_{n=1}^M d_\alpha(n) + \lambda M h_\alpha. \text{ (ver equação (6.26))}$$

$$l_\alpha(0) = 0$$

para $i = 1 ; i \leq N_{\text{amp}} ; i = i + 1$ **faça**

$$l_\alpha(i) = \frac{L_\alpha}{2^{N_{\text{amp}} - i}} \text{ (ver equação (6.27) e Figura 6.12)}$$

$$\Gamma_{\text{amp}}(i) = \{\gamma | l_\alpha(i-1) < \alpha_q \leq l_\alpha(i)\}$$

fim para

$$J_{g,\text{mín}}^{\text{tot}} = \sum_{i=1}^{N_{\text{amp}}} \min_{q_\rho(i), q_\phi(i)} J_g(\lambda, q_\rho(i), q_\phi(i), q_\xi, q_{n^s}, q_{n^L})$$

$$J(N_{\text{amp}}) = J_\alpha + J_{g,\text{mín}}^{\text{tot}}$$

fim para

$$N_{\text{amp}}^* = \arg \min_{N_{\text{amp}}} J$$

Obtêm-se os passos de quantização ótimos $q_\alpha^*, \{q_\rho^*(i), q_\phi^*(i)\}_{i=1}^{N_{\text{amp}}^*}$.

A seguir, é detalhada a maneira como se calcula a parcela do custo taxa-distorção relativa aos parâmetros dos átomos por faixa de energia. Esta parcela é definida

como

$$J_g(\lambda, q_\rho(i), q_\phi(i), q_\xi, q_{n^s}, q_{n^L}) = \sum_{i=1}^{N_{\text{amp}}} M_i \bar{\alpha}^2(i) \bar{d}_g(i, q_\rho(i), q_\phi(i), q_\xi, q_{n^s}, q_{n^L}) + \lambda \sum_{i=1}^{N_{\text{amp}}} M_i h_g(i, q_\rho(i), q_\phi(i), q_\xi, q_{n^s}, q_{n^L}), \quad (6.37)$$

em que

$$h_g(i, q_\rho(i), q_\phi(i), q_\xi, q_{n^s}, q_{n^L}) = h_\rho(i) + h_\phi(i) + h_\xi + h_{n^s} + h_{n^L}, \quad (6.38)$$

sendo que

$$\begin{aligned} h_\rho(i) &= - \sum_{k=1}^{N_{I_\rho(i)}} p_\rho(I_\rho(k)) \log_2(p_\rho(I_\rho(k))) \text{ (bits/símbolo)}, \\ h_\phi(i) &= - \sum_{k=1}^{N_{I_\phi(i)}} p_\phi(I_{\nu_2}(k)) \log_2(p_\phi(I_{\nu_2}(k))) \text{ (bits/símbolo)}, \\ h_\xi &= - \sum_{k=1}^{N_{I_\xi}} p_\xi(I_\xi(k)) \log_2(p_{\nu_i}(I_\xi(k))) \text{ (bits/símbolo)}, \\ h_{n^s} &= - \sum_{k=1}^{N_{I_{n^s}}} p_{n^s}(I_{n^s}(k)) \log_2(p_{n^s}(I_{n^s}(k))) \text{ (bits/símbolo)}, \\ h_{n^L} &= - \sum_{k=1}^{N_{I_{n^L}}} p_{n^L}(I_{n^L}(k)) \log_2(p_{n^L}(I_{n^L}(k))) \text{ (bits/símbolo)}, \end{aligned}$$

Deve-se observar que como somente os passos de quantização q_ρ e q_ϕ variam em função da faixa de amplitude, o mesmo ocorre com as entropias. Dessa maneira, rescreve-se a equação (6.37) da seguinte forma

$$\begin{aligned} J_g(\lambda, q_\rho(i), q_\phi(i)) &= \sum_{i=1}^{N_{\text{amp}}} M_i \bar{\alpha}^2(i) \bar{d}_g(i, q_\rho(i), q_\phi(i)) + \\ &\quad \lambda \sum_{i=1}^{N_{\text{amp}}} M_i (h_\rho(i) + h_\phi(i)) + \\ &\quad \lambda M (h_\xi + h_{n^s} + h_{n^L}), \end{aligned} \quad (6.39)$$

em que M_i é o número de átomos da faixa i , M é o número total de átomos e $\bar{\alpha}^2(i) = \bar{\alpha}^2(i)$ ($\bar{\alpha}(i)$ corresponde à média aritmética entre os limites da faixa i , definida na equação (6.28)).

Para calcular a distorção média dos átomos por faixa de energia, resolve-se para

cada faixa a seguinte integral

$$\begin{aligned} \overline{d}_g(q_\rho, q_\phi) = \int \int \int \int \int d_g(\rho, \phi, \omega, n^L, q_\rho, q_\phi) \\ p_\gamma(\alpha, \rho, \phi, \omega, n^L) d(\alpha)d(\rho)d(\phi)d(\omega)d(n^L), \end{aligned} \quad (6.40)$$

em que $p_\gamma(\cdot)$ é a função de probabilidade conjunta entre os parâmetros e $d_g(\cdot)$ é a função de distorção do átomo. Deve-se notar que d_g não depende de n^s pois seu cálculo exige uma integral definida de n^s a $n^s + n^L$, que corresponde a integrar de 0 a n^L . Detalhes sobre o cálculo de d_g serão vistos mais adiante.

Infelizmente, não é possível obter uma forma fechada para \overline{d}_g em função de q_ρ e q_ϕ , posto que não se tem de antemão a probabilidade conjunta entre os parâmetros $p_\gamma(\cdot)$. Neste caso, opta-se por calcular previamente uma tabela contendo a distorção média dos átomos por faixa de energia através de um conjunto de sinais de treinamento. Este procedimento é descrito através de pseudocódigo do Algoritmo 6.2 e detalhado a seguir:

1. define-se o conjunto de número de faixas de amplitude Υ ;
2. define-se o conjunto de passos de quantização de decaimento \mathcal{Q}_ρ ;
3. define-se o conjunto de passos de quantização de fase \mathcal{Q}_ϕ ;
4. para cada número de faixas de energia $N_{\text{amp}} \in \Upsilon$:
 - (a) definem-se as faixas de energia não-uniformemente espaçadas através dos limites l_α ;
 - (b) agregam-se os átomos em faixas de energia $\Gamma_{\text{amp}}(i)$;
 - (c) para cada faixa de energia i , com base em $\Gamma_{\text{amp}}(i)$, calcula-se a distorção média \overline{d}_g variando-se os passos de quantização q_ρ e q_ϕ ;

Ao final deste processo, obtêm-se as tabelas de distorção média de átomos em função dos passos de quantização $q_\rho \in \mathcal{Q}_\rho$ e $q_\phi \in \mathcal{Q}_\phi$. Cada tabela está associada a um número de faixas de amplitude $N_{\text{amp}} \in \Upsilon$ e uma faixa de amplitude.

Na prática, os conjuntos são definidos da seguinte maneira

$$\left\{ \begin{array}{l} \Upsilon = \{k \in \mathbb{N} | 1 \leq k \leq 16\} \\ \mathcal{Q}_\rho = \{L_\rho / (2^b - 1) \in \mathbb{R} | b = 1 + k \Delta b \leq b_{\text{máx}}, k \in \mathbb{N}, \\ \quad \text{com } L_\rho = 4, b_{\text{máx}} = 22, \Delta b = 0,25\} \\ \mathcal{Q}_\phi = \{L_\phi / (2^b - 1) \in \mathbb{R} | b = 1 + k \Delta b \leq b_{\text{máx}}, k \in \mathbb{N}, \\ \quad \text{com } L_\phi = 2\pi, b_{\text{máx}} = 18, \Delta b = 0,25\} \end{array} \right. \quad (6.41)$$

Algoritmo 6.2 Cálculo da distorção média dos átomos por faixa de energia para diferentes níveis de quantização de amplitude.

Define-se Υ
Define-se \mathcal{Q}_ρ
Define-se \mathcal{Q}_ϕ
para $N_{\text{amp}} \in \Upsilon$ **faça**
 $L_\alpha = \underset{\alpha \in A}{\text{máx}} \alpha$
para $i = 1 ; i \leq N_{\text{amp}} ; i = i + 1$ **faça**
 $l_\alpha(i) = \frac{L_\alpha}{2^{N_{\text{amp}}-i}}$
 $\Gamma_{\text{amp}}(i) = \{\gamma | l_\alpha(i-1) < \alpha_q \leq l_\alpha(i)\}$
fim para
para $i = 1 ; i \leq N_{\text{amp}}$ **faça**
para $q_\rho \in \mathcal{Q}_\rho$ **faça**
para $q_\phi \in \mathcal{Q}_\phi$ **faça**
calcula-se a distorção média $\overline{d}_g(N_{\text{amp}}, i, q_\rho, q_\phi)$ tal que $\gamma \in \Gamma_{\text{amp}}(i)$.
fim para
fim para
fim para
fim para

Os valores de L_ρ e L_ϕ foram definidos sabendo-se que durante a decomposição do sinal MPSACr garante-se que $|\rho| \leq 4$ e $0 \leq \phi \leq 2\pi$. Em relação ao decaimento, dependendo do sinal, é possível que o máximo valor de $|\rho|$ seja menor que 4. Neste caso, o conjunto \mathcal{Q}_ρ de passos de quantização de decaimento fica mais restrito, ou seja,

$$\begin{aligned} \mathcal{Q}_\rho &= \{L_\rho / (2^b - 1) \in \mathbb{R} | b = 1 + k \Delta b \leq b_{\text{máx}}, k \in \mathbb{N}, \\ &\quad \text{com } L_\rho = 4, b_{\text{máx}} = 22, \Delta b = 0,25 \text{ e} \\ &\quad L_\rho / (2^b - 1) < \rho_{\text{máx}}, \text{ onde } \rho_{\text{máx}} = \underset{\rho}{\text{máx}} |\rho|\}. \end{aligned} \quad (6.42)$$

É importante destacar que os passos de quantização estão associados a taxas de bits de valores reais, e isto é possível tendo em vista o uso de um codificador de entropia. As taxas máximas de 22 e 18 bits proveem intervalos considerados suficientes para otimização taxa-distorção e alocação ótima de bits entre os parâmetros ρ e ϕ , de forma a não restringir o codificador a baixas taxas de bits. O passo de taxa Δb foi escolhido de modo a fornecer uma boa resolução ao processo de otimização taxa-distorção sem inviabilizar seu tempo de execução.

A fim de reduzir o custo computacional na obtenção destas tabelas, desenvolve-se uma expressão analítica para o cálculo da distorção do átomo d_g . O átomo exponencial possui norma unitária e é definido por

$$g_\gamma(t) = \frac{1}{\kappa_\gamma} e^{-\rho t} \cos(\xi t + \phi) [u(t - n^s) - u(t - n^s - n^L)], \quad (6.43)$$

em que $u(\cdot)$ é a função degrau unitário, $\gamma = (\rho, \xi, \phi, n^s, n^L)$ e κ_γ é definido tal que $\|\mathbf{g}_\gamma\| = 1$ da seguinte maneira

$$(\kappa_\gamma)^2 = \int_{n^s}^{n^L} [e^{-\rho t} \cos(\xi t + \phi)]^2 dt. \quad (6.44)$$

Seguindo a dedução matemática de κ_γ^2 , se $\xi \neq 0$ e $\rho \neq 0$, então

$$\begin{aligned} \kappa_\gamma^2 &= \int_{n^s}^{n^L} [e^{-\rho t} \cos(\xi t + \phi)]^2 dt \\ &= \left[-\frac{e^{-2\rho t} \{\rho^2 (\cos(2\xi t + 2\phi) + 1) + \xi^2 - \rho\xi \operatorname{sen}(2\xi t + 2\phi)\}}{4\rho^3 + 4\rho\xi^2} \right]_{n^s}^{n^L} \\ &= F_\kappa(n^L) - F_\kappa(n^s), \end{aligned} \quad (6.45)$$

em que

$$F_\kappa(t) = -\frac{e^{-2\rho t} \{\rho^2 (\cos(2\xi t + 2\phi) + 1) + \xi^2 - \rho\xi \operatorname{sen}(2\xi t + 2\phi)\}}{4\rho^3 + 4\rho\xi^2}. \quad (6.46)$$

Se $\xi \neq 0$ e $\rho = 0$, então

$$\begin{aligned} \kappa_\gamma^2 &= \int_{n^s}^{n^L} [\cos(\xi t + \phi)]^2 dt \\ &= \left[\frac{t}{2} + \frac{\operatorname{sen}(2\xi t + 2\phi)}{4\xi} \right]_{n^s}^{n^L} \\ &= F_\kappa(n^L) - F_\kappa(n^s), \end{aligned} \quad (6.47)$$

em que

$$F_\kappa(t) = \frac{t}{2} + \frac{\operatorname{sen}(2\xi t + 2\phi)}{4\xi}. \quad (6.48)$$

Se $\xi = 0$ e $\rho \neq 0$, tem-se

$$\begin{aligned} \kappa_\gamma^2 &= \int_{n^s}^{n^L} [e^{-\rho t} \cos(\phi)]^2 dt \\ &= F_\kappa(n^L) - F_\kappa(n^s), \end{aligned} \quad (6.49)$$

em que

$$F_\kappa(t) = -\frac{\cos^2(\phi)}{(2\rho)} e^{-2\rho t}. \quad (6.50)$$

Se $\xi = 0$ e $\rho = 0$, tem-se

$$\begin{aligned}\kappa_\gamma^2 &= \int_{n^s}^{n^L} [\cos(\phi)]^2 dt \\ &= F_\kappa(n^L) - F_\kappa(n^s),\end{aligned}\tag{6.51}$$

em que

$$F_\kappa(t) = t (\cos(\phi))^2.\tag{6.52}$$

A distorção do átomo consiste no erro quadrático entre o átomo gerado com os parâmetros originais e o gerado com os parâmetros quantizados. No caso do átomo exponencial, a distorção é definida como

$$d_g = \int_{n^s}^{n^L} \left(\frac{1}{\kappa_\gamma} e^{-\rho t} \cos(\xi t + \phi) - \frac{1}{\kappa_{\gamma_q}} e^{-\rho_q t} \cos(\xi t + \phi_q) \right)^2 dt.\tag{6.53}$$

Vale lembrar que os parâmetros quantizados com perdas são somente o decaimento e a fase; ρ_q e ϕ_q são as versões quantizadas de ρ e ϕ ; e κ_γ e κ_{γ_q} são tais que $\|\mathbf{g}_\gamma\| = 1$ e $\|\mathbf{g}_{\gamma_q}\| = 1$. Continuando a dedução de d_g , tem-se que

$$\begin{aligned}d_g &= \int_{n^s}^{n^L} \frac{1}{\kappa_\gamma^2} e^{-2\rho t} \cos^2(\xi t + \phi) dt + \\ &\quad \int_{n^s}^{n^L} \frac{1}{\kappa_{\gamma_q}^2} e^{-2\rho_q t} \cos^2(\xi t + \phi_q) dt - \\ &\quad \int_{n^s}^{n^L} 2 \frac{1}{\kappa_\gamma \kappa_{\gamma_q}} e^{-(\rho + \rho_q)t} \cos(\xi t + \phi) \cos(\xi t + \phi_q) dt.\end{aligned}\tag{6.54}$$

Dado que

$$2 \cos(\xi t + \phi) \cos(\xi t + \phi_q) = \cos(2\xi t + \phi + \phi_q) + \cos(\phi - \phi_q),\tag{6.55}$$

então

$$\begin{aligned}
d_g &= \frac{1}{\kappa_\gamma^2} \int_{n^s}^{n^L} e^{-2\rho t} \cos^2(\xi t + \phi) dt + \frac{1}{\kappa_{\gamma_q}^2} \int_{n^s}^{n^L} e^{-2\rho_q t} \cos^2(\xi t + \phi_q) dt - \\
&\quad \frac{1}{\kappa_\gamma \kappa_{\gamma_q}} \int_{n^s}^{n^L} e^{-(\rho+\rho_q)t} \cos(2\xi t + \phi + \phi_q) dt - \frac{1}{\kappa_\gamma \kappa_{\gamma_q}} \int_{n^s}^{n^L} e^{-(\rho+\rho_q)t} \cos(\phi - \phi_q) dt \\
&= \frac{1}{\kappa_\gamma^2} (F_1(\rho, \phi, n^L) - F_1(\rho, \phi, n^s)) + \frac{1}{\kappa_{\gamma_q}^2} (F_1(\rho_q, \phi_q, n^L) - F_1(\rho_q, \phi_q, n^s)) - \\
&\quad \frac{1}{\kappa_\gamma \kappa_{\gamma_q}} (F_2(\rho, \phi, \rho_q, \phi_q, n^L) - F_2(\rho, \phi, \rho_q, \phi_q, n^s)) - \\
&\quad \frac{1}{\kappa_\gamma \kappa_{\gamma_q}} (F_3(\rho, \phi, \rho_q, \phi_q, n^L) - F_3(\rho, \phi, \rho_q, \phi_q, n^s)), \tag{6.56}
\end{aligned}$$

em que

$$\begin{cases}
F_1(\rho, \phi, t) &= \int e^{-2\rho t} \cos^2(\xi t + \phi) dt \\
F_2(\rho, \phi, \rho_q, \phi_q, t) &= \int e^{-(\rho+\rho_q)t} \cos(2\xi t + \phi + \phi_q) dt \\
F_3(\rho, \phi, \rho_q, \phi_q, t) &= \int e^{-(\rho+\rho_q)t} \cos(\phi - \phi_q) dt.
\end{cases}$$

As expressões analíticas destas integrais indefinidas são dadas por

$$F_1(\rho, \phi, t) = \begin{cases} -\frac{e^{-2\rho t}}{4\rho^3+4\rho\xi^2} (\rho^2 (\cos(2\xi t + 2\phi) + 1) + \xi^2 - \rho\xi \text{sen}(2\xi t + 2\phi)), \\ \quad \text{se } \xi \neq 0 \text{ e } \rho \neq 0 \\ \frac{t}{2} + \frac{\text{sen}(2\xi t + 2\phi)}{4\xi}, \text{ se } \xi \neq 0 \text{ e } \rho = 0 \\ -\cos^2(\phi) \frac{e^{-2\rho t}}{2\rho}, \text{ se } \xi = 0 \text{ e } \rho \neq 0 \\ t \cos^2(\phi), \text{ se } \xi = 0 \text{ e } \rho = 0; \end{cases} \tag{6.57}$$

$$F_2(\rho, \phi, \rho_q, \phi_q, t) = \begin{cases} -e^{-(\rho+\rho_q)t} \frac{(\rho \cos(2\xi t + \phi + \phi_q) + \rho_q \cos(2\xi t + \phi + \phi_q) - 2\xi \text{sen}(2\xi t + \phi + \phi_q))}{\rho^2 + 2\rho\rho_q + \rho_q^2 + 4\xi^2}, \\ \quad \text{se } \xi \neq 0 \text{ e } (\rho + \rho_q) \neq 0 \\ \frac{\text{sen}(2\xi t + \phi + \phi_q)}{2\xi}, \text{ se } \xi \neq 0 \text{ e } (\rho + \rho_q) = 0 \\ -e^{-(\rho+\rho_q)t} \frac{\cos(\phi + \phi_q)}{\rho + \rho_q}, \text{ se } \xi = 0 \text{ e } (\rho + \rho_q) \neq 0 \\ t \cos(\phi + \phi_q), \text{ se } \xi = 0 \text{ e } (\rho + \rho_q) = 0; \end{cases} \tag{6.58}$$

$$F_3(\rho, \phi, \rho_q, \phi_q, t) = \begin{cases} -e^{-(\rho+\rho_q)t} \frac{\cos(\phi - \phi_q)}{\rho + \rho_q} & \text{se } (\rho + \rho_q) \neq 0 \\ t \cos(\phi - \phi_q) & \text{se } (\rho + \rho_q) = 0. \end{cases} \tag{6.59}$$

Um detalhe importante diz respeito ao átomo exponencial crescente: este, na realidade consiste na reversão temporal de um átomo decrescente. Portanto, ao se

calcular a distorção de um átomo crescente através da fórmula analítica é necessário utilizar os parâmetros referentes do átomo decrescente correspondente. Por exemplo, dado um $\gamma = (\rho, \xi, \phi, n^s, n^L)$ que define um átomo crescente, os parâmetros do átomo decrescente correspondente são $\gamma' = (-\rho, \xi, -\phi, N_Q - 1 - n^L, N_Q - 1 - n^s)$, em que N_Q é o tamanho do quadro.

6.4.2 Distribuição Gaussiana Generalizada

As entropias do coeficiente e dos parâmetros do átomo dependem diretamente de suas funções de densidade de probabilidade. Lança-se mão de representá-las através da distribuição gaussiana generalizada [131, 132], que é definida como

$$f(x) = \frac{p}{2a\Gamma(1/p)} e^{-\left(\frac{|x-\mu|}{a}\right)^p}, \quad (6.60)$$

em que μ é a média, a é o fator de escalonamento e p é o parâmetro de forma da distribuição. A função $\Gamma(\cdot)$ é definida por

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt, \quad x > 0. \quad (6.61)$$

A implementação numérica desta função pode ser encontrada em [130]. O fator de escalonamento a é definido como

$$a = \sqrt{\frac{\sigma^2 \Gamma(1/p)}{\Gamma(3/p)}}, \quad (6.62)$$

em que σ^2 corresponde à variância. A Figura 6.14 mostra a distribuição gaussiana generalizada para diferentes parâmetros de forma p , mas considerando a mesma variância. Note que $p = 2$ resulta em uma função de densidade de uma distribuição normal (gaussiana) e $p = 1$ resulta na função de densidade de uma laplaciana. Valores mais baixos de p resultam em distribuições mais pontiagudas.

A função de densidade de probabilidade (fdp) da gaussiana generalizada é obtida a partir de três parâmetros: a média (μ), a variância (σ^2) e a forma (p). Em [131, 132], indica-se o caminho a se percorrer para a estimação destes parâmetros. A média e a variância são estimadas através das seguintes relações

$$\begin{aligned} \hat{\mu} &= \frac{1}{N} \sum_{i=1}^N x_i \\ \hat{\sigma}^2 &= \frac{1}{N} \sum_{i=1}^N (x_i - \hat{\mu})^2, \end{aligned} \quad (6.63)$$

em que N é o número de elementos da amostra de parâmetros. A estimativa de

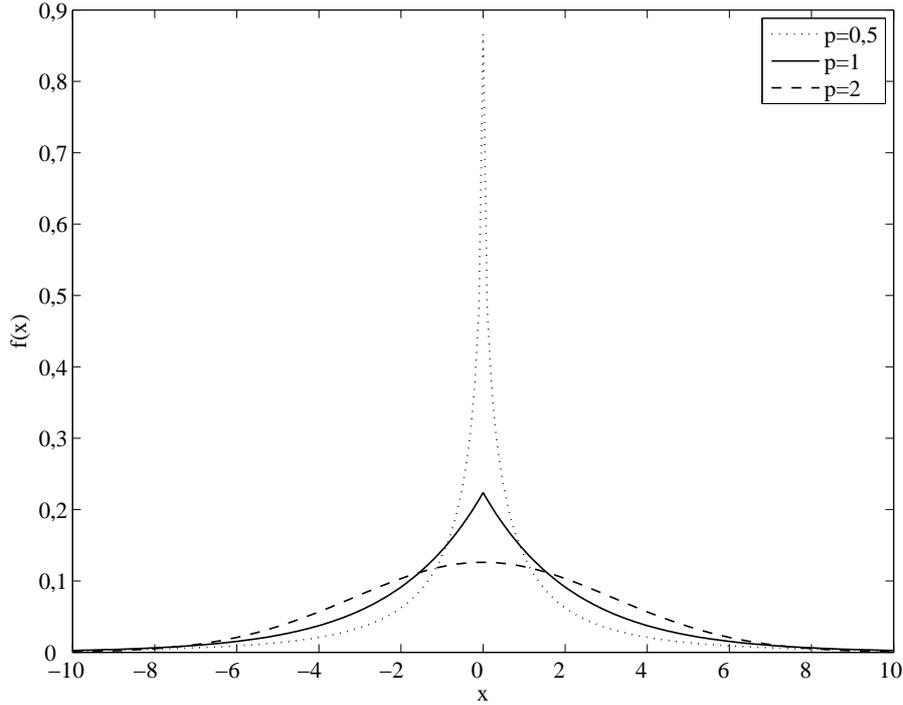


Figura 6.14: Distribuição gaussiana generalizada para diferentes parâmetros de forma p mas considerando a mesma variância

p é obtida com base em uma relação entre a variância (σ^2) e a média dos valores absolutos $E(|X|)$ de uma fdp gaussiana generalizada de média zero

$$R(p) = \frac{E^2(|X|)}{\sigma^2} = \frac{\Gamma^2(2/p)}{\Gamma(1/p)\Gamma(3/p)}, \quad (6.64)$$

em que $R(p)$ é denominada como função de razão gaussiana generalizada [132]. Para estimar a forma p adotam-se os seguintes procedimentos:

1. calculam-se $\hat{\mu}$ e $\hat{\sigma}^2$;
2. calcula-se uma estimativa da média dos valores absolutos

$$\hat{E}(|X|) = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{\mu}|; \quad (6.65)$$

3. calcula-se a razão

$$\eta = \frac{\hat{E}^2(|X|)}{\hat{\sigma}^2}; \quad (6.66)$$

4. consulta-se em uma tabela (*look-up table*), que contém valores pré-calculados de $R(p)$ em função de diferentes valores de p , o \hat{p} que resulta no $\hat{R}(\hat{p})$ mais próximo de η .

6.4.3 Codificação

A organização dos símbolos dentro do feixe de bits é a mesma descrita na Seção 6.2.2 (página 104). Por conveniência, repete-se na Figura 6.15 a ilustração da formatação do feixe de bits. A informação do feixe de bits é codificada e decodificada da esquerda para direita, de cima para baixo. Os símbolos são separados por quadros de decomposição, sendo que a amostra inicial do suporte do átomo é responsável por indicar o quadro ao qual ele pertence. O símbolo 0 do coeficiente identifica o fim de lista de coeficientes, indicando implicitamente o número de átomos N_{qd}^k codificados no quadro de decomposição k .

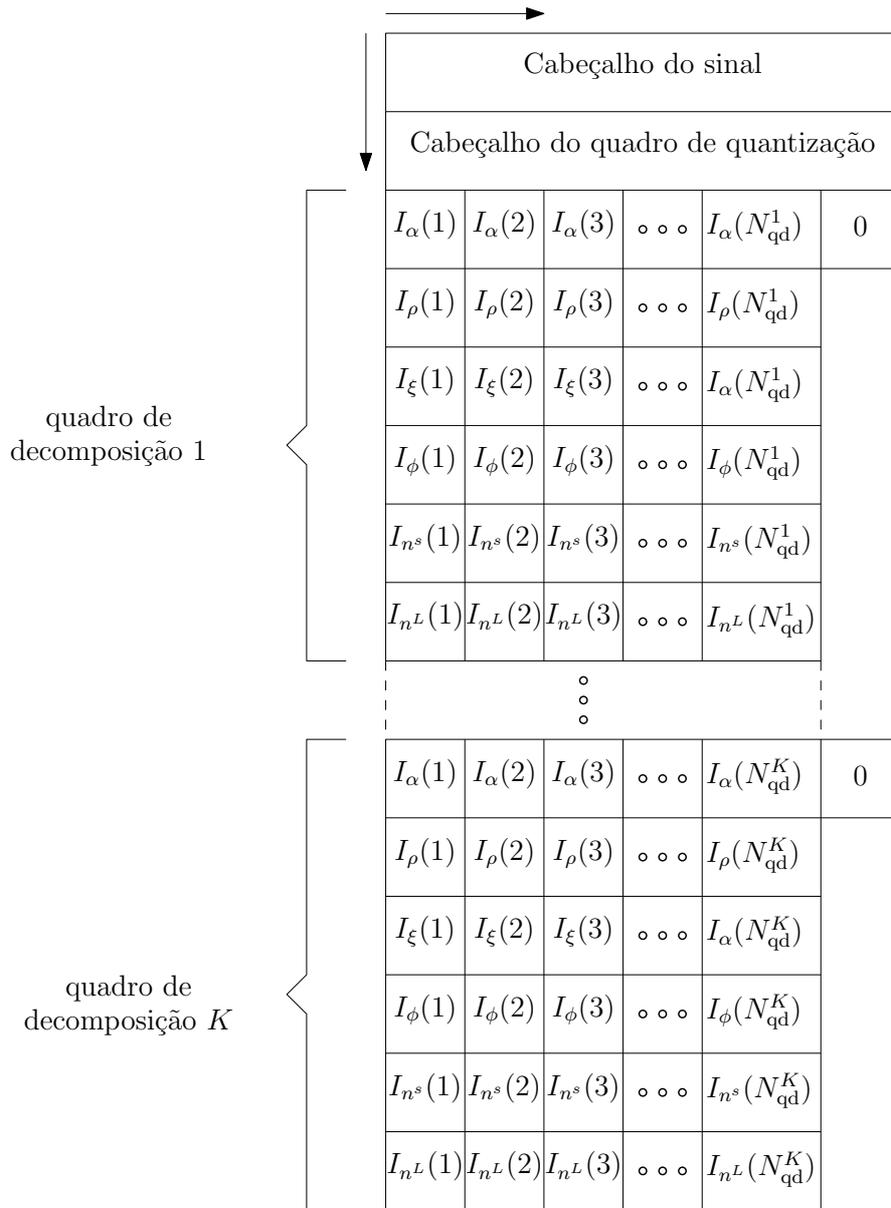


Figura 6.15: Organização dos símbolos dentro do feixe de bits.

O cabeçalho do sinal é composto pelas seguintes informações:

- fator de normalização do sinal;
- tamanho do sinal;
- taxa de amostragem;
- tamanho do quadro decomposição;
- tamanho do quadro de quantização.

O cabeçalho do quadro de quantização é composto pelas seguintes informações:

- valor máximo da amplitude $\alpha_{\text{máx}}$;
- valor máximo do decaimento $\rho_{\text{máx}}$;
- valor máximo do tamanho do suporte $n_{\text{máx}}^L$;
- número de faixas de energia;
- multiplicador de Lagrange λ .

Os valores mínimos da amplitude e do decaimento são iguais a zero, e os valores mínimo e máximo da fase são 0 e 2π , respectivamente. Como estas informações já são conhecidas tanto no codificador quanto no decodificador, não é necessário enviá-las através de cabeçalho.

Pode-se observar que ocorreram modificações no cabeçalho de quantização: em vez das taxas aplicadas aos coeficientes e aos parâmetros dos átomos, são enviados ao decodificador o número de faixas de energia, que também indica o nível de quantização dos coeficientes, e o multiplicador de Lagrange λ obtido na otimização taxa-distorção. Através do λ é possível obter a alocação ótima de bits entre os parâmetros do átomo por faixa de energia. Dessa maneira, evita-se a sobrecarga do cabeçalho com o envio de informações a respeito das taxas ótimas dos parâmetros do átomo por faixa de energia.

Com o propósito de alcançar uma representação mais compacta, aplica-se um codificador de entropia aos símbolos, como ilustrado na Figura 6.16. Opta-se por utilizar o codificador aritmético com modelo de probabilidade fixo [44, 128, 129]. A implementação eficiente do codificador aritmético adotada neste trabalho está descrita em [129]. Note que os símbolos são codificados na mesma ordem em que são organizados no feixe de bits (ver Figura 6.15) e cada parâmetro possui seu próprio modelo, que é conhecido tanto no codificador quanto no decodificador.

As probabilidades dos símbolos do coeficiente e dos parâmetros do átomo são calculadas com base na função densidade de probabilidade de cada parâmetro. Considere o parâmetro ν , cujos símbolos são representados por $\{I_\nu(k)\}_{k=1}^{N_{I_\nu}}$, sendo N_{I_ν}

o número de símbolos; o passo de quantização q_ν ; e a função densidade de probabilidade de ν , $f_\nu(\nu(k))$, onde $\nu(k) = I_\nu(k) q_\nu$. A probabilidade do símbolo $I_\nu(k)$ é definida por

$$p_\nu(k) = \frac{f_\nu(\nu(k))q_\nu}{\sum_{k=1}^{N_{I_\nu}} f_\nu(\nu(k))q_\nu} = \frac{f_\nu(\nu(k))}{\sum_{k=1}^{N_{I_\nu}} f_\nu(\nu(k))}, \quad (6.67)$$

onde o somatório no denominador garante que $\sum_{k=1}^{N_{I_\nu}} p_\nu(k) = 1$. As funções de densidade de probabilidade associadas ao coeficiente, ao decaimento e ao tamanho do suporte correspondem a distribuições gaussianas generalizadas. Os parâmetros de frequência, fase e amostra inicial do suporte são representados por distribuição uniforme. Isto foi definido de acordo com os resultados das estatísticas dos parâmetros, que serão vistos mais adiante, na seção de Simulações Computacionais.

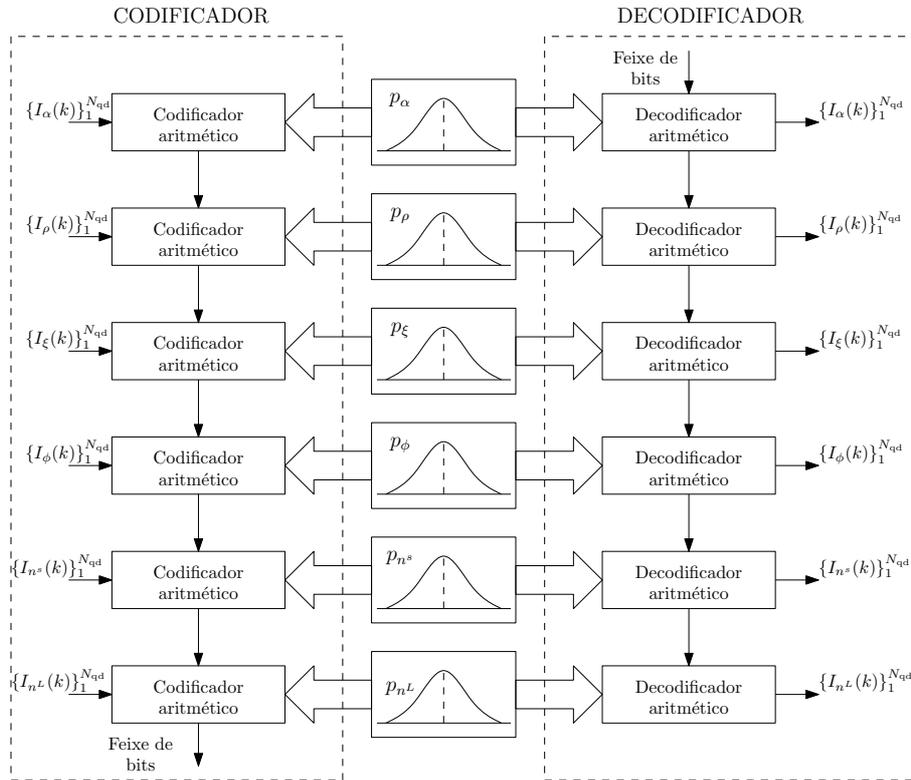


Figura 6.16: Codificação de entropia dos símbolos usando codificador aritmético com modelo de probabilidade fixo. Os símbolos são codificados e decodificados na seguinte ordem de parâmetros: coeficientes, decaimentos, frequências, fases, amostras iniciais do suporte e tamanhos do suporte. As funções de probabilidade dos coeficientes e dos parâmetros do átomo exponencial são representadas por p_α , p_ρ , p_ξ , p_ϕ , p_{n^s} e p_{n^L} .

Em codificação aritmética, uma mensagem é representada por um intervalo de números reais entre 0 e 1 [128]. À proporção que a mensagem se torna mais longa,

menor é o intervalo necessário para representá-la e o número de bits que especifica o intervalo aumenta. Ao codificarmos sucessivos símbolos de uma mensagem, reduz-se o tamanho do intervalo de acordo a probabilidade do símbolo produzido pelo modelo. Os símbolos mais prováveis reduzem menos o intervalo do que os símbolos menos prováveis, portanto adicionam menos bits à mensagem.

Sendo mais específico, considere a mensagem $S = \beta_1\beta_2 \dots \beta_N$, cujos símbolos são extraídos de um alfabeto $\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_M\}$. Cada símbolo α_i do alfabeto está associado a um intervalo real $[a_i, b_i)$ contido em $[0; 1)$, definido de acordo com a sua probabilidade. A mensagem S é codificada como um intervalo $I^S = [L, H)$ contido em $[0; 1)$. Para isso:

1. define-se um intervalo corrente inicial com $[0; 1)$;
2. para cada símbolo da mensagem:
 - (a) particiona-se o intervalo corrente em subintervalos, um para cada símbolo do alfabeto, que são definidos de acordo com as probabilidades dos símbolos;
 - (b) o subintervalo correspondente ao símbolo codificado passa a ser o novo intervalo corrente;
3. codifica-se a mensagem através do número real com a representação binária mais compacta dentro do último intervalo calculado, que corresponde a I^S .

A fim de exemplificar este processo, considere o alfabeto $\mathcal{A} = \{a, b, !\}$ onde os intervalos associados aos símbolos são definidos da seguinte forma

$$\begin{aligned}
 a &\rightarrow [0; 0,5) \\
 b &\rightarrow [0,5; 0,8) \\
 ! &\rightarrow [0,8; 1)
 \end{aligned}
 \tag{6.68}$$

e o símbolo ! indica o fim da mensagem. A codificação da mensagem $S = abba!$ é descrita na Tabela 6.3 e ilustrada através da Figura 6.17. Com base nesta tabela, a mensagem $abba!$ é codificada como o intervalo $[0,343; 0,3475)$. Seleciona-se o número real dentro deste intervalo com a mais compacta representação binária que é enviada ao decodificador. A este número real dá-se o nome de valor-código (*code value*) [129].

O decodificador, por sua vez, realiza o processo inverso, como ilustrado na Figura 6.18. Considere que o valor-código seja o limite inferior do intervalo associado à mensagem $abba!$, ou seja, 0,343. A cada símbolo decodificado, ocorre o mapeamento dos subintervalos dentro do intervalo do símbolo anterior e identifica-se o subintervalo ao qual pertence o valor-código. Este procedimento se repete até que o

Tabela 6.3: Codificação da mensagem *abba!*.

mensagem codificada	símbolo adicionado	intervalo do símbolo adicionado	intervalo anterior	próximo intervalo
a	a	[0; 0,5)	[0; 1)	[0; 0,5)
ab	b	[0,5; 0,8)	[0; 0,5)	[0,25; 0,4)
abb	b	[0,5; 0,8)	[0,25; 0,4)	[0,325; 0,37)
abba	a	[0; 0,5)	[0,325; 0,37)	[0,325; 0,3475)
abba!	!	[0,8; 1)	[0,325; 0,3475)	[0,343; 0,3475)

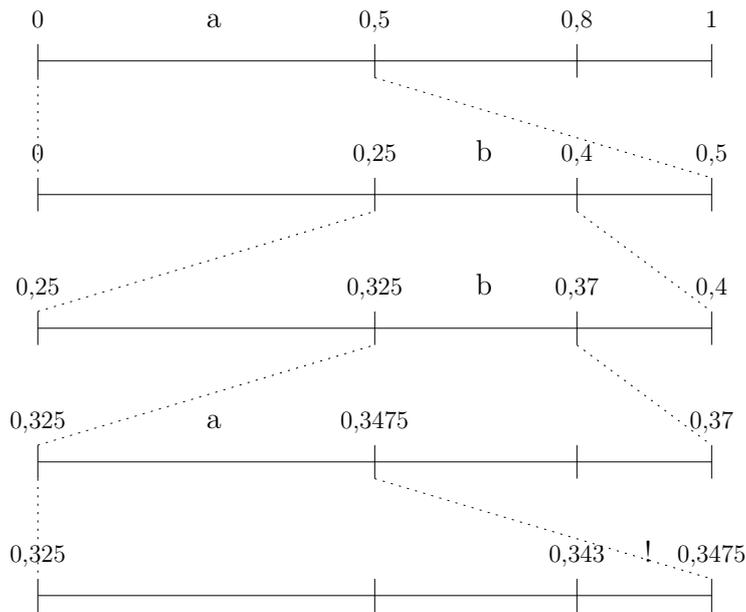


Figura 6.17: Codificação da mensagem *abba!*.

símbolo de fim de mensagem (!) seja decodificado. O símbolo de fim de mensagem é fundamental para que o processo de decodificação não se perca entrando em laço infinito.

6.4.4 Simulações Computacionais

Nesta seção, apresentam-se resultados de simulações computacionais relacionadas ao método de compressão proposto, que realiza a otimização taxa-distorção por faixa de energia atômica e utiliza codificação aritmética.

Na Seção 6.3.4, foi visto que a otimização taxa-distorção por faixa de energia atômica requer o pré-cálculo de uma tabela contendo as distorções médias dos átomos exponenciais por faixa de energia, dado um número de faixas. A distorção média do átomo é calculada em função dos passos de quantização do decaimento e da fase. Essa tabela é obtida a partir de um conjunto de sinais de treinamento constituído de diferentes tipos de sinais de áudio:

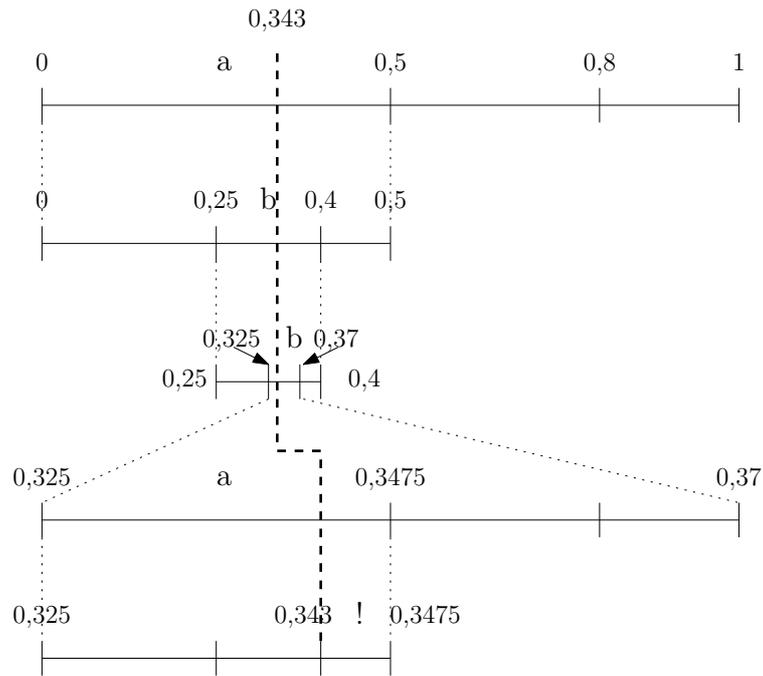


Figura 6.18: Decodificação da mensagem *abba!*.

- notas avulsas dos instrumentos musicais: piano, flauta, violoncelo e fagote;
- sequências de notas de piano e violão;
- trechos musicais polifônicos que misturam diferentes instrumentos;
- trecho musical com uma sequência de acordes de violão;
- solo de um instrumento de percussão;
- trecho de voz falada;
- trecho musical contendo voz e trio instrumental;
- solo de harmônica de vidro;
- solo de órgão de tubos;
- som de tiro de revólver.

Obtém-se uma representação para cada sinal deste conjunto através do método de decomposição MPSACr. Garante-se que essa representação está associada a uma razão sinal-ruído de 60 dB. Após a decomposição, agrupam-se os átomos continuados de forma que a variação do decaimento de um quadro ao outro não ultrapasse 1 %. O número de faixas de amplitude varia de 1 a 16. Dado um número de faixas, separam-se os átomos por faixa de amplitude. Para cada faixa, calcula-se a distorção média dos átomos usando a expressão analítica em (6.56) (página 130) e armazena-se o

número de átomos. Finalmente, calculam-se as distorções médias dos átomos de todos os sinais do conjunto de treinamento, classificando-as por número de faixas, por faixa de energia e pelos passos de quantização do decaimento e da fase.

A Figura 6.19 ilustra as superfícies de distorção média dos átomos exponenciais por faixa de energia em função dos passos de quantização do decaimento (q_ρ) e da fase (q_ϕ) para o caso de quatro faixas de energia. A primeira faixa corresponde à faixa de menor energia, e a quarta faixa corresponde à faixa de maior energia. Os passos de quantização do decaimento e da fase variam de 0 a 4 e de 0 a 2π , respectivamente, com espaçamento não-uniforme, sendo mais refinado para passos de quantização menores. Como esperado, à medida que são reduzidos os passos de quantização do decaimento e da fase, reduz-se também a distorção média dos átomos. As reduções mais significativas de distorção média (para aproximadamente -80 dB) ocorrem quando os passos de quantização são bem pequenos, da ordem de 10^{-6} . É importante ressaltar que estes gráficos mostram as superfícies médias considerando diferentes sinais com características distintas entre si, o que possivelmente explica sua convexidade. As superfícies de distorção média dos átomos para os demais números de faixas de amplitude apresentam comportamento similar às superfícies ilustradas neste exemplo da Figura 6.19 relativo a quatro faixas.

O método de compressão proposto neste trabalho faz uso da codificação aritmética e abrange o cálculo de entropia, ambos diretamente dependentes das funções de densidade de probabilidade da amplitude e dos parâmetros do átomo. A Figura 6.20 apresenta os histogramas da amplitude normalizada, do decaimento e do tamanho do suporte do átomo. Dado que a amplitude máxima dentro do livro de estruturas varia de acordo com o sinal, opta-se por ajustar a distribuição da amplitude normalizada (amplitude dividida pela amplitude máxima). As distribuições destes três parâmetros foram ajustadas através de gaussiana generalizada por apresentarem um comportamento exponencial. Observa-se que a distribuição do tamanho do suporte é bimodal e se concentra em dois pontos: 0 e 2047. Dessa maneira, para ajustar a distribuição do tamanho do suporte foi necessário utilizar a soma de três gaussianas generalizadas.

A função de probabilidade gaussiana generalizada é definida como

$$f(x; \mu; a; p) = \frac{p}{2a\Gamma(1/p)} e^{-\left(\frac{|x-\mu|}{a}\right)^p}, \quad (6.69)$$

onde μ é a média, a é o fator de escalonamento, p é o parâmetro de forma da distribuição. Adotando-se o procedimento de estimação dos parâmetros da gaussiana generalizada, descrito na Seção 6.4.2 (página 131), tem-se que a função de

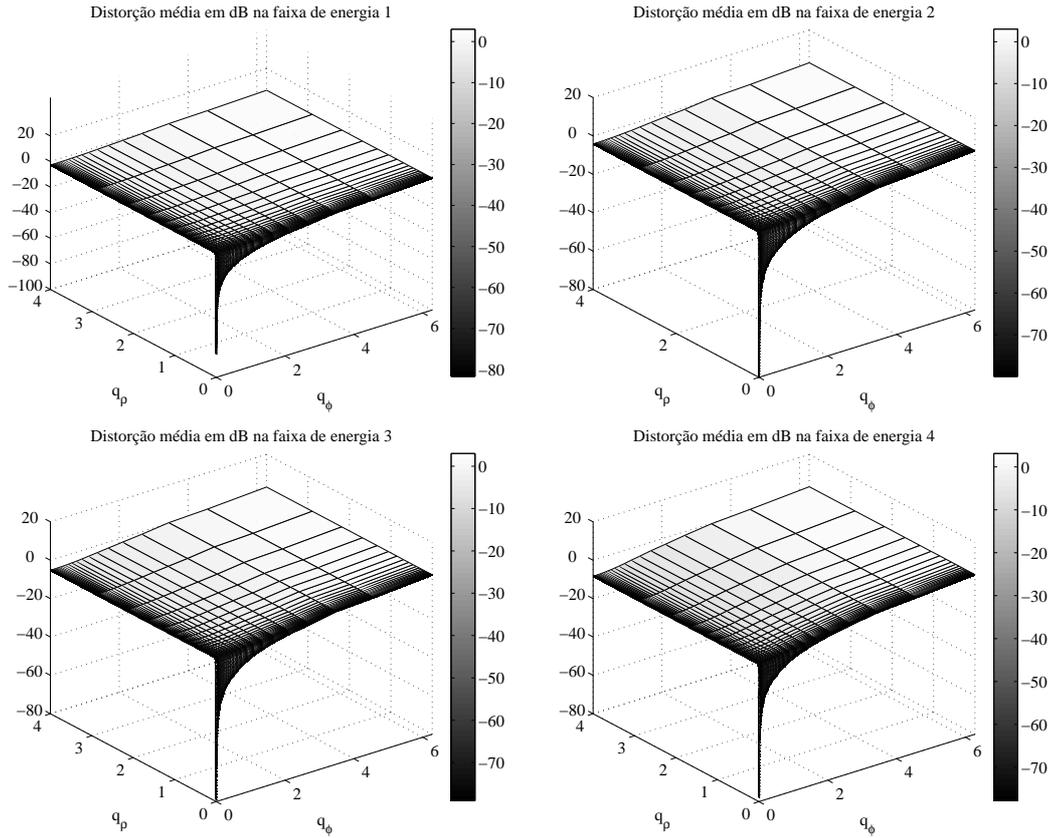


Figura 6.19: Superfícies de distorção média dos átomos exponenciais por faixa de energia para diferentes combinações de passos de quantização do decaimento (q_p) e da fase (q_ϕ). Ilustra-se o caso de quatro faixas de energia, sendo a primeira faixa a de menor energia, e a quarta faixa a de maior energia.

probabilidade da amplitude normalizada é descrita por

$$f_\alpha(x) = f(x; 0; 0,0000562; 0,379), \quad (6.70)$$

e a função de probabilidade do decaimento é descrita por

$$f_\rho(x) = f(x; 0; 0,0235; 0,49). \quad (6.71)$$

Os parâmetros da função de probabilidade do tamanho do suporte foram obtidas ajustando-se sua distribuição de forma visual. Isso é permitido porque o codificador aritmético adotado é de modelo fixo. Esta função é descrita da seguinte forma

$$f_{n^L}(x) = 0,5 f(x; 0; 10; 1) + f(x; 2047; 10; 1) + 1,2 f(x; 1023; 1024; 100). \quad (6.72)$$

Os demais parâmetros do átomo são modelados como distribuições uniformes. A Figura 6.21 ilustra os histogramas dos parâmetros de fase, amostra inicial do suporte temporal do átomo e frequência. Observe que a fase e a amostra inicial do suporte

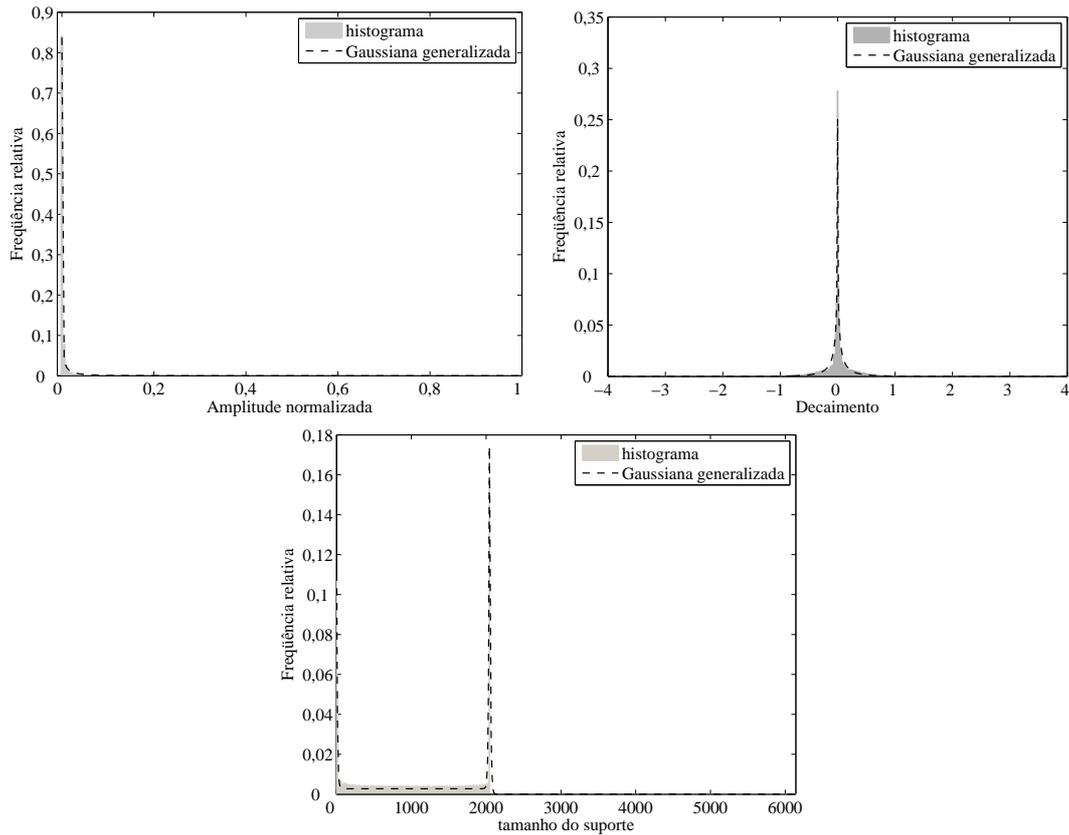


Figura 6.20: Histogramas da amplitude normalizada, do decaimento e do tamanho do suporte do átomo ajustados através de distribuição gaussiana generalizada.

são bem representadas por uma distribuição uniforme. Com relação à frequência, nota-se que sua distribuição não é exatamente uniforme. Contudo, como o espectro de frequências é característico de cada sinal, não é interessante adotar uma função de probabilidade que privilegie algumas frequências em detrimento de outras. Nesse caso, a distribuição uniforme é o modelo mais adequado para as frequências dos átomos.

Agora, avaliamos o desempenho do sistema de compressão com base em otimização taxa-distorção por faixa de energia comparado ao sistema de compressão baseado em otimização taxa-distorção por curvas operacionais, apresentado na Seção 6.2 (página 98). Além do aprimoramento relacionado à otimização taxa-distorção, o codificador proposto utiliza codificação aritmética, modelando as distribuições dos parâmetros através de gaussiana generalizada. Comparam-se as curvas de taxa-distorção de três arcabouços de codificação:

- *curvaOper*: codificação através de otimização taxa-distorção por curvas operacionais;
- *faixaEnergia*: codificação através de otimização taxa-distorção por faixa de energia atômica com codificação aritmética e modelagem das distribuições do

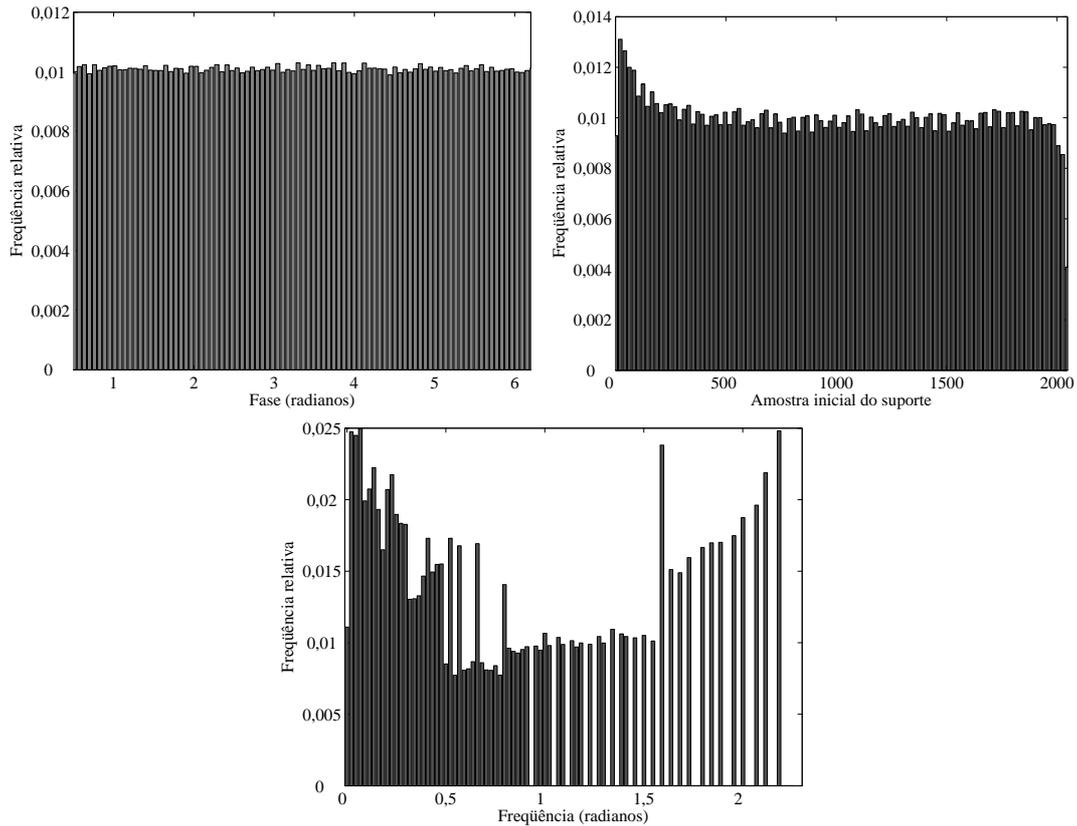


Figura 6.21: Histogramas dos parâmetros de fase, amostra inicial do suporte temporal do átomo e frequência.

coeficiente e dos parâmetros do átomo por função de probabilidade uniforme;

- *faixaEnergiaGG*: codificação através de otimização taxa-distorção por faixa de energia atômica com codificação aritmética e modelagem das distribuições do coeficiente e dos parâmetros do átomo por função de probabilidade gaussiana generalizada.

Ao utilizar o sistema de compressão na configuração *faixaEnergia*, onde os parâmetros são modelados através de distribuição uniforme, avalia-se o efeito isolado da alocação ótima de bits entre as faixas de amplitude. Na configuração *faixaEnergiaGG*, verifica-se o efeito conjunto da nova abordagem de alocação ótima de bits e uma modelagem mais adequada das distribuições dos parâmetros através de gaussiana generalizada.

A parametrização do MPSACr é a mesma adotada originalmente (ver Seção 6.2.3, página 106), ou seja, os sinais são divididos em quadros de 2048 amostras com saltos de 2048 amostras. Para o critério de parada, adota-se o valor de 60 dB para a razão sinal-ruído alvo (SNR_{alvo}) e o valor de 0,1095 para taxa de aproximação média desejada.

Na Figura 6.22, mostram-se as curvas de taxa-distorção obtidas através da co-

dificação de quatro sinais, quando se realiza a otimização taxa-distorção através de curvas operacionais (*curvaOper*), por faixa de energia atômica modelando o coeficiente e os parâmetros do átomo como distribuição uniforme (*faixaEnergia*) e por faixa de energia atômica ajustando as distribuições do coeficiente e dos parâmetros do átomo através de gaussiana generalizada (*faixaEnergiaGG*). Os sinais são os mesmos descritos na Seção 5.3 (página 87) e usados nas simulações computacionais relativas ao codificador na configuração *curvaOper* (ver Seção 6.2.3—página 106): *pianoA3*, *fluteA4*, *celloA4* e *tirofinal*. Ressalta-se que estes sinais fazem parte do conjunto de treinamento usado para a obtenção da tabela contendo as distorções médias dos átomos exponenciais por faixa de energia. Portanto, ao codificar-se cada sinal, utiliza-se a tabela de distorções médias por faixa de energia calculada com base nos demais sinais do conjunto. Dessa forma, evita-se qualquer tipo de viés nos resultados de codificação quando é utilizada a abordagem de otimização taxa-distorção por faixa de energia atômica.

Deve-se observar que, para os quatro sinais analisados, a otimização taxa-distorção por faixa de energia implica melhor desempenho do sistema de compressão, posto que para um dado nível de distorção, reduz-se a taxa. Esta redução é ainda mais acentuada quando as distribuições dos parâmetros são modeladas por gaussianas generalizadas. Além disso, o ganho no desempenho taxa-distorção é mais nítido para taxas mais altas. Possivelmente, isto ocorre porque mais átomos passam a compor a representação do sinal, fazendo com que as estatísticas dos parâmetros se aproximem mais dos modelos de probabilidade ajustados previamente. Dessa maneira, o codificador aritmético consegue atuar de forma mais eficiente.

Até o momento, verificou-se que os aperfeiçoamentos realizados no sistema de codificação propostos e desenvolvidos nesta tese resultaram em melhor desempenho taxa-distorção, ou seja, dado um nível de distorção, reduziu-se a taxa de compressão. A etapa seguinte consiste em avaliar o quanto se pode reduzir a taxa durante a codificação de modo que a distorção introduzida no sinal seja imperceptível ao ouvido humano. A forma mais confiável de se avaliar o desempenho obtido após o processamento de um sinal de áudio, qualquer que tenha sido o processamento realizado, consiste na avaliação subjetiva. Para que isto seja feito sistematicamente, é preciso recorrer a um grande número de pessoas com boa sensibilidade auditiva para ouvirem o áudio processado e avaliarem sua qualidade. Devido às dificuldades inerentes a esta categoria de testes, há bastante tempo realizam-se pesquisas com o objetivo de desenvolver métodos para avaliação objetiva de qualidade, isto é, que não necessitem de pessoas [28–30, 133, 134]. Dentre os diversos métodos de avaliação objetiva de qualidade da literatura, destaca-se o PEAQ (*Perceptual Evaluation of Audio Quality*) [28–30]. O PEAQ é um método de medição objetiva da qualidade de sinais de áudio padronizado pela ITU (*International Telecommunication Union*),

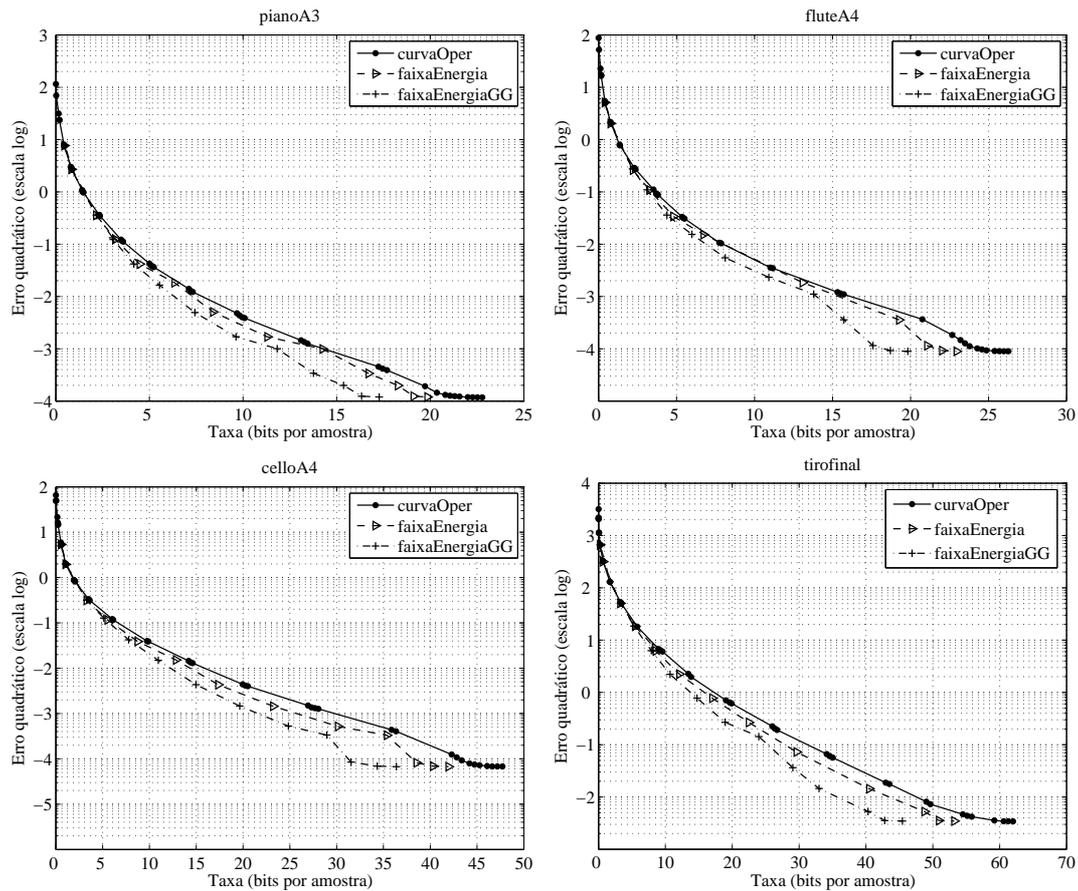


Figura 6.22: Curvas de taxa-distorção obtidas através da codificação de quatro sinais realizando a otimização taxa-distorção através de curvas operacionais (*curvaOper*), por faixa de energia atômica modelando o coeficiente e os parâmetros do átomo como distribuição uniforme (*faixaEnergia*) e por faixa de energia atômica ajustando as distribuições do coeficiente e dos parâmetros do átomo através de gaussiana generalizada (*faixaEnergiaGG*).

originalmente destinado à avaliação de codecs com perdas. Ele utiliza diversas medidas de base psicoacústica que, combinadas, resultam em uma medida de qualidade relativa entre duas instâncias de um sinal—um sinal de referência (original) e um sinal teste (reconstruído) [28]. Essa medida de qualidade é classificada em diferentes níveis de degradação do sinal de áudio:

- **0:** imperceptível;
- **-1:** perceptível, mas não incômoda;
- **-2:** pouco incômoda;
- **-3:** incômoda;
- **-4:** muito incômoda.

Avalia-se aqui o desempenho do sistema de compressão utilizando-se uma implementação do PEAQ elaborada por Peter Kabal [30]. É importante ressaltar que esta implementação permite atribuir notas de valores positivos, porém próximos de zero. Na recomendação BS.1387 [28], elaborada pela ITU, destaca-se que os sinais de teste devem possuir uma duração de 10 a 20 segundos. Os sinais utilizados em testes subjetivos normalmente apresentam duração dentro deste intervalo. Se os sinais forem muito longos, o ouvinte avaliador acaba “esquecendo” o início do sinal, focando somente nas degradações que ocorrem no seu final. Se o sinal for muito curto, o ouvinte avaliador não consegue perceber a degradação. Outros cuidados ao se utilizar o PEAQ dizem respeito ao alinhamento e ao ganho entre os sinais de referência e de teste [30]. Esses parâmetros devem ser ajustados antes de se realizar a avaliação objetiva de qualidade do sinal. Além disso, os sinais devem apresentar uma taxa de amostragem de 48 kHz; do contrário, devem ser reamostrados a 48 kHz.

O experimento consiste em avaliar a qualidade dos sinais codificados através do PEAQ. Os sinais de teste utilizados são representados originalmente em 16 bits no formato PCM (*Pulse Code Modulation*), gravados em mono e na taxa de amostragem de 44,1 kHz, com exceção do sinal *blues*, que possui taxa de 48kHz. Eles são listados e descritos a seguir:

- *berimbau*: solo de violão, em acordes, que imita a sonoridade de um berimbau. Possui duração de 12 segundos.
- *blues*: trecho musical de um quarteto de *jazz* contendo vibrafone, contrabaixo, piano e bateria. Possui duração de 11 segundos.
- *largoschulz*: solo de harmônica de vidro, de característica tonal com forte inarmonicidade. Possui duração de 12,5 segundos.
- *violaofaixa2*: solo de violão predominantemente monofônico. Possui duração de 9,8 segundos.
- *pianomidinota2*: sequência sintética de notas de piano de altura e duração variadas sequencialmente. Possui duração de 9,9 segundos.
- *hagenteaqui*: trecho musical contendo voz e trio instrumental. Possui duração de 6 segundos.

Inicialmente, avalia-se a qualidade da representação do sinal obtida através do MP-SACr antes de se introduzir perdas através de quantização. Vale lembrar que o MPSACr foi parametrizado de forma a obter representações de sinais com razão sinal-ruído (SNR) de 60 dB. A Tabela 6.4 apresenta os resultados de avaliação de

qualidade de áudio através do PEAQ, aplicado às representações dos sinais de áudio obtidas via MPSACr, tomando-se todos os cuidados para seu uso adequado. Observa-se que todos os sinais obtiveram o grau de diferença como imperceptível, com exceção do *pianomidinota2*, que obteve um grau de perceptível, mas não incômodo. Apesar disso, através de testes informais de audição, não foi possível perceber diferença entre os sinais de referência e os sinais reconstruídos a eles associados.

Tabela 6.4: Avaliação através de PEAQ da representação obtida via MPSACr de diferentes sinais de áudio com razão sinal-ruído de 60 dB.

Sinal	Nota
<i>berimbau</i>	-0,134
<i>blues</i>	0,133
<i>largoschulz</i>	0,068
<i>pianomidinota2</i>	-1,18
<i>violaofaixa2</i>	-0,31
<i>hagenteaqui</i>	0,142

Em seguida, será avaliado, com auxílio do PEAQ a qualidade perceptiva dos sinais codificados pelo sistema de compressão proposto, em quatro taxas específicas: 128, 256, 320 e 512 kbps (quilobits por segundo). A taxa correspondente em bits por amostra pode ser obtida dividindo-se a taxa em bits por segundo pela frequência de amostragem. A Tabela 6.5 mostra os resultados de PEAQ quando aplicado aos sinais codificados. A coluna *Nota-SQ* se refere às notas dadas às reconstruções dos sinais sem a quantização dos coeficientes e dos parâmetros e com a retirada dos átomos cujos coeficientes se tornariam nulos com a quantização associada à taxa de compressão desejada. Com isso, objetiva-se verificar em que condição inicial o codificador, parte antes da inserção de erros nos coeficientes e parâmetros dos átomos. Nota-se que considerando este conjunto de sinais, o PEAQ indica uma degradação perceptível, mas não incômoda, a uma taxa de 512 kbps, com exceção do sinal *berimbau*, cuja degradação é graduada como incômoda. Portanto, as representações providas pelo MPSACr não são suficientemente compactas, de modo que a remoção de átomos de baixa energia sem a quantização do coeficiente e dos parâmetros dos átomos restantes já implica distorções perceptíveis pelo ouvido humano, mesmo a uma taxa alta como 512 kbps.

A coluna *Nota-Q* se refere às notas dadas aos sinais codificados considerando a quantização dos coeficientes e dos parâmetros dos átomos. Como esperado, as notas dadas pelo PEAQ tendem a indicar maior nível de degradação audível, visto que, além de haver a remoção de átomos de mais baixa energia, os coeficientes e os

parâmetros dos átomos são quantizados, introduzindo ainda mais distorção no sinal.

Tabela 6.5: Avaliação do desempenho do sistema de compressão proposto através do PEAQ. A coluna *Nota-SQ* se refere às notas associadas às representações dos sinais sem a quantização dos coeficientes e dos parâmetros e com a remoção dos átomos cujos coeficientes se tornariam nulos com a quantização associada à taxa de compressão desejada. A coluna *Nota-Q* se refere às notas dadas aos sinais codificados considerando a quantização dos coeficientes e dos parâmetros dos átomos.

Sinal	Taxa (kbps)	Nota-SQ	Nota-Q
<i>berimbau</i>	128	-3,837	-3,827
	256	-3,750	-3,806
	320	-3,437	-3,762
	512	-2,568	-3,594
<i>blues</i>	128	-3,537	-3,419
	256	-3,202	-3,542
	320	-2,068	-3,076
	512	-0,898	-1,566
<i>largoschulz</i>	128	-3,694	-3,643
	256	-3,730	-3,713
	320	-3,550	-3,632
	512	-1,306	-1,957
<i>pianomidinota2</i>	128	-3,913	-3,913
	256	-2,860	-3,871
	320	-2,308	-3,687
	512	-1,170	-2,656
<i>violaofaixa2</i>	128	-3,763	-3,727
	256	-3,658	-3,790
	320	-3,164	-3,627
	512	-0,803	-1,630
<i>hagenteaqui</i>	128	-3,740	-3,748
	256	-3,464	-3,520
	320	-2,480	-2,869
	512	-1,229	-1,418

A Figura 6.23 ilustra o sinal *berimbau* e seu resíduo com sua representação obtida removendo-se os átomos cujos coeficientes se tornariam nulos com a quantização a uma taxa de 320 kbps, porém sem quantizar os coeficientes e os parâmetros dos átomos restantes. No decorrer do resíduo, observam-se picos que estão associados a amostras próximas dos limites entre quadros de decomposição. A Figura 6.24 mostra um trecho ampliado do resíduo da Figura 6.23 para melhor visualização desses picos. Isso demonstra que o método de decomposição ainda é ineficiente para representar a informação em torno das bordas dos quadros. Em termos perceptivos, estes picos de diferença entre os sinais original e reconstruído se traduzem em cliques, que normalmente são incômodos ao ouvido humano. A representação das bordas

entre quadros no processo de decomposição merece investigações mais aprofundadas em trabalhos futuros. Isto deverá proporcionar melhor desempenho na compressão. Outro fator que possivelmente está prejudicando o codificador está relacionado ao projeto do dicionário. As frequências dos dicionário estão limitadas a um valor correspondente a aproximadamente duas oitavas acima da última nota do piano (C8), ou seja, 15804,26 Hz. Dado que os sinais originais possuem uma taxa de amostragem 44,1 kHz, existem componentes de frequência entre 15804,26 e 22050 Hz (metade da taxa de amostragem) que não constam no dicionário. Isto reduz sensivelmente a capacidade do método de decomposição em obter representações compactas; prejudicando, por consequência, o desempenho do codificador.

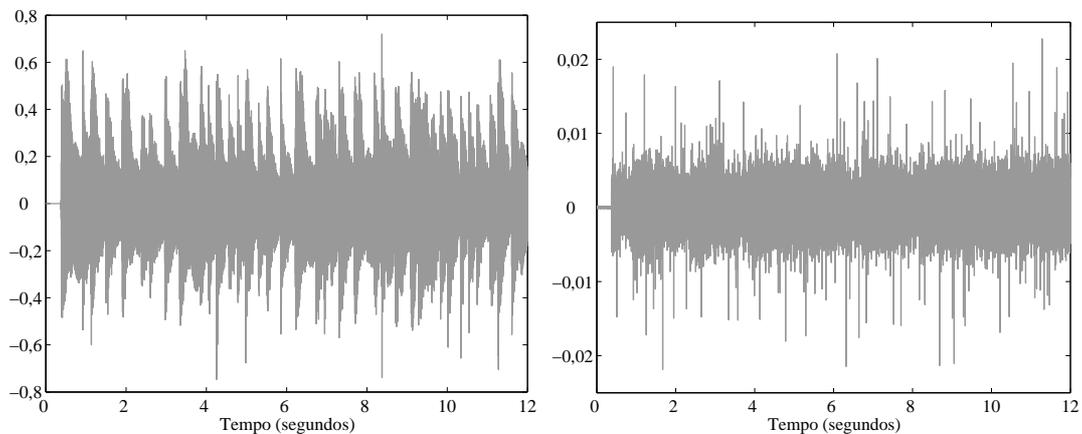


Figura 6.23: O sinal *berimbau* original (à esquerda) e o resíduo (à direita) com relação à representação obtida removendo-se os átomos cujos coeficientes se tornariam nulos com a quantização a uma taxa de 320 kbps, porém sem quantizar os coeficientes e os parâmetros dos átomos restantes.

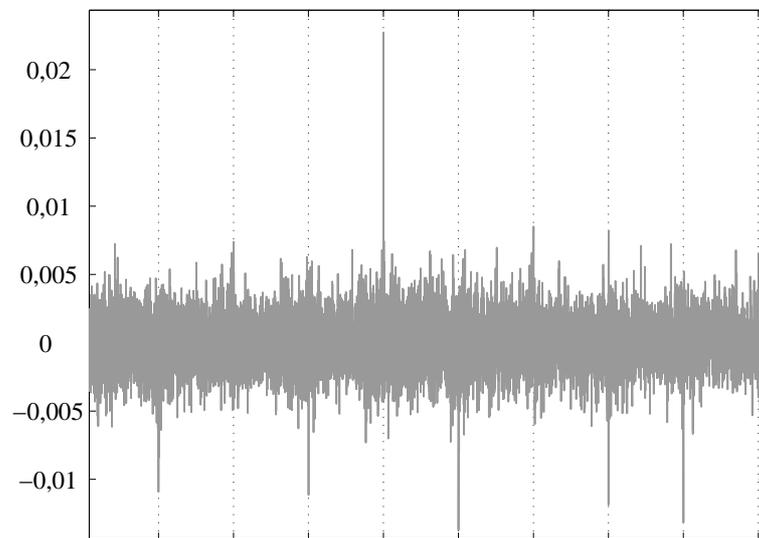


Figura 6.24: Versão ampliada de um trecho do resíduo apresentado na Figura 6.23. As linhas verticais tracejadas indicam os limites entre os quadros de decomposição.

Capítulo 7

Conclusão

Nesta tese, foi proposto um arcabouço de compressão de sinais usando decomposições atômicas com base em dicionários redundantes. Em geral, utiliza-se um dicionário fixo na etapa de decomposição e codificam-se os coeficientes e os índices dos átomos. Em contrapartida, neste trabalho, introduziu-se um novo conceito de compressão de sinais usando decomposições atômicas com base em múltiplos dicionários. Além de transmitir os coeficientes e os índices dos átomos, o codificador precisa enviar ao decodificador informação adicional a respeito do dicionário utilizado. Este paradigma foi aplicado tanto a sinais elétricos quanto a sinais de áudio.

Inicialmente, foram abordados os aspectos gerais a respeito de decomposições atômicas: fundamentos, exemplos de aplicações, vantagens de utilizá-las junto a dicionários redundantes em vez de decomposições em base. Além disso, foram apresentados os fundamentos do uso de decomposições atômicas em compressão de sinais, introduzindo-se o conceito de múltiplos dicionários.

Em seguida, apresentou-se o sistema de compressão aplicado a sinais elétricos proposto e desenvolvido nesta tese. Este sistema se baseia no método de decomposição atômica com base no dicionário de senoides amortecidas (MPSA) e no procedimento de seleção do dicionário com compromisso taxa-distorção ótimo. O desempenho deste sistema foi avaliado com o auxílio de uma ferramenta de análise automática de faltas. As figuras de mérito utilizadas foram o instante de início da falta, os fasores calculados a partir dos sinais de tensão e corrente, o tipo de falta e a distância a que a falta ocorreu. Com base nos resultados, verificou-se que a localização de faltas é o procedimento determinante da taxa de compressão, em bits por amostra, abaixo da qual não se recomenda comprimir registros de perturbações elétricas. Para o conjunto teste de oscilogramas utilizado e considerando o erro de 4% no cálculo da distância, pode-se empregar uma taxa de 2,5 bits/amostra sem comprometer qualquer das características de falta analisadas. Isso indica que o sistema de compressão proposto aqui é capaz de prover alta razão de compressão (6,4:1) com considerável confiabilidade e robustez, sem comprometer o processo de

análise de faltas dos oscilogramas comprimidos.

Fundamentado no MPSA, desenvolveu-se um novo método de decomposição atômica aplicado a sinais de áudio, o MPSACr, que se baseia em um dicionário composto por senoides amortecidas de frequências na escala musical cromática. Nesse método traz consigo uma nova estratégia de redução de complexidade computacional de maneira a não restringir o projeto do dicionário a um espaçamento frequencial uniforme. Além disso, esse método é composto por uma heurística de continuidade de átomos entre quadros consecutivos.

Com base no MPSACr, desenvolveu-se um sistema de compressão aplicado a sinais de áudio. Nesse sistema, realiza-se uma otimização taxa-distorção por faixa de energia atômica. Dessa maneira, foi possível alocar os bits por entre os parâmetros dos átomos, discriminando os átomos através de sua energia. Os bits são distribuídos por diferentes faixas de amplitude, podendo-se gastar mais bits na representação de parâmetros de átomos de maior energia em detrimento dos átomos de mais baixa energia. Além dessa estratégia de alocação de bits, utiliza-se a codificação aritmética, modelando-se o coeficiente e os parâmetros do átomo através de gaussiana generalizada.

O desempenho do sistema de compressão foi avaliado através do PEAQ, que é um método de medição objetiva de qualidade de sinais de áudio padronizado pela ITU (*International Telecommunication Union*). Através dos resultados do PEAQ, observou-se a necessidade de se utilizar taxas altas compressão de modo que a distorção inserida pela compressão não seja perceptível ao ouvido humano. Verificou-se que o método de decomposição ainda é ineficiente para representar a informação em torno das bordas do quadros, pois analisar o resíduo entre o sinal original e o comprimido aparecem picos próximos às bordas. Estes picos se traduzem em cliques, que normalmente são incômodos. A representação das bordas entre quadros no processo de decomposição merece investigações mais aprofundadas em trabalhos futuros, posto que deverá proporcionar melhor desempenho na compressão. Outro fator que possivelmente está prejudicando o codificador está relacionado ao projeto do dicionário. As frequências dos dicionário estão limitadas a um valor correspondente à aproximadamente duas oitavas acima da última nota do piano (C8), ou seja, 15804,26 Hz. Dado que os sinais originais possuem uma taxa de amostragem 44,1 kHz, existem componentes de frequência entre 15804,26 e 22050 Hz (metade da taxa de amostragem) que não constam no dicionário. Isto reduz sensivelmente a capacidade do método de decomposição em obter representações compactas, prejudicando, por consequência, o desempenho do codificador.

7.1 Contribuições

A seguir, enumeram-se os aprimoramentos e as inovações apresentados nesta tese de doutorado:

- Introduziu-se um conceito de compressão de sinais usando decomposições atômicas com base em múltiplos dicionários;
- Desenvolveu-se uma maneira computacionalmente viável de lidar com inúmeros dicionários através do uso de dicionários parametrizados, cujos átomos são definidos por parâmetros contínuos;
- Desenvolveu-se um método de otimização taxa-distorção capaz de selecionar o dicionário de compromisso taxa-distorção ótimo dentro do paradigma de compressão com múltiplos dicionários. Este método se baseia na obtenção das curvas de taxa-distorção operacionais;
- A partir do método de decomposição desenvolvido em [25] (MPSA), construiu-se um sistema de compressão aplicado a sinais elétricos, introduzindo-se o esquema de otimização taxa-distorção através de curvas operacionais, que fornece o dicionário correspondente a menor distorção dada uma taxa desejada;
- Avaliou-se o desempenho do método de compressão de sinais elétricos por procedimentos de análise de faltas, utilizando-se uma ferramenta de análise automática de faltas no lugar de figuras ou métricas normalmente adotadas, como o erro quadrático médio;
- Inspirado no MPSA, desenvolveu-se um método de decomposição aplicado a sinais de áudio: o MPSACr, que se baseia em um dicionário de senoides amortecidas com frequências da escala musical cromática. Como o MPSACr provê a representação do sinal no domínio de tempo, o sinal é segmentado em quadros sem sobreposição, produzindo descontinuidades em átomos presentes em quadros consecutivos;
- Dessa maneira, lança-se mão de uma heurística de continuidade dos átomos proposta e desenvolvida neste trabalho;
- O MPSACr traz consigo uma nova estratégia de redução de complexidade computacional no cálculo dos produtos internos, de maneira que não há restrição para o espaçamento frequencial no projeto do dicionário. Isto se ajusta perfeitamente ao dicionário de escala musical cromática, que possui frequências geometricamente espaçadas;

- Desenvolveu-se um sistema de compressão aplicado a sinais de áudio, que utiliza o MPSACr e o esquema de otimização taxa-distorção através de curvas operacionais;
- Introduziram-se a fundamentação teórica e a descrição da técnica de otimização taxa-distorção por faixa de energia atômica. Através deste método é possível alocar bits de modo a privilegiar a representação de átomos de maior energia em detrimento dos de menor energia;
- Aprimoraram-se o sistema de compressão aplicado a sinais de áudio, introduzindo este novo paradigma de alocação ótima de bits por faixa de energia em substituição à técnica que se baseia em curvas operacionais. Além disso, utiliza-se um codificador aritmético de modelo fixo, modelando o coeficiente e os parâmetros do átomo exponencial através de gaussiana generalizada. Por fim, avaliou-se o desempenho do sistema de compressão através do PEAQ, que é um método de medição objetiva de qualidade de sinais de áudio padronizado pela ITU (*International Telecommunication Union*).

7.2 Trabalhos Futuros

Propõem-se para trabalhos futuros os seguintes tópicos:

- Realizar ajustes no dicionário de senoides amortecidas com frequências da escala musical cromática, estendendo as frequências até a metade da taxa de amostragem do sinal. Dessa forma, espera-se obter representações mais compactas dos sinais;
- Aprimorar o método de decomposição MPSACr na representação da informação nas bordas dos quadros para eliminar os picos que surgem no resíduo entre o sinal original e sua versão codificada;
- Introduzir no sistema de compressão de sinais de áudio um modelo psicoacústico a fim de remover os átomos não perceptíveis ao ouvido humano;
- Introduzir no sistema de compressão de sinais elétricos o esquema de otimização taxa-distorção por faixa de energia atômica;
- Tratar, na etapa de codificação, os átomos continuados através de uma estrutura/molécula, mesmo com variações significativas de decaimento no decorrer dos quadros. Dessa forma, espera-se alcançar um melhor desempenho de compressão.

Referências Bibliográficas

- [1] LOVISOLO, L. *Representações de Sinais Usando Dicionário Redundantes*. Tese de Doutorado, PEE/COPPE,UFRJ, Rio de Janeiro, Brasil, 2006.
- [2] RODRIGUES, M. A. M. *Decomposições Eficientes para Codificação de Sinais*. Tese de Doutorado, PEE/COPPE,UFRJ, Rio de Janeiro, Brasil, 1999.
- [3] IEEE PES WORKING GROUP P1433. “Power Quality Definitions”. . <http://grouper.ieee.org/groups/1433/>, Acessado em maio de 2007.
- [4] LOVISOLO, L., DA SILVA, E. A. B., RODRIGUES, M. A. M., et al. “Efficient Coherent Representations of Power Systems Signals Using Damped Sinusoids”, *IEEE Transactions on Signal Processing*, v. 53, n. 10, pp. 3831–3846, Out. 2005.
- [5] ZWICKER, E., FASTL, H. *Psychoacoustics: Facts and Models*. 3 ed. New York, USA, Springer-Verlag, 1990.
- [6] GOODWIN, M. *Adaptive Signal Models: Theory, Algorithms, and Audio Applications*. Ph.D. Thesis, University of California at Berkeley, Berkeley, USA, 1997.
- [7] GOODWIN, M. M. *Adaptive Signal Models: Theory, Algorithms, and Audio Applications*. Kluwer Inter. Series in Engineering and Computer Science. 1 ed. New York, USA, Kluwer, 1998.
- [8] KRIM, H., TUCKER, D., MALLAT, S., et al. “On Denoising and Best Signal Representations”, *IEEE Transactions on Information Theory*, v. 45, n. 7, pp. 2225–2238, Nov. 1999.
- [9] LOVISOLO, L., RODRIGUES, M. A. M., DA SILVA, E. A. B., et al. “Eliminação da Resposta Transitória em Oscilogramas utilizando Matching Pursuits”. In: *VII Seminário Técnico de Proteção e Controle*, pp. 22–27, Rio de Janeiro, Brasil, Jun. 2003.

- [10] COIFMAN, R., WICKERHAUSER, M. “Entropy-based Algorithms for Best Basis Selection”, *IEEE Transactions on Information Theory*, v. 38, n. 2, pp. 713–718, Mar. 1992.
- [11] CHEN, S. S., DONOHO, D. L., MICHAEL, et al. “Atomic Decomposition by Basis Pursuit”, *SIAM Journal on Scientific Computing*, v. 20, n. 1, pp. 33–61, Ago. 1998.
- [12] ETEMOGLU, C., CUPERMAN, V. “Matching Pursuits Sinusoidal Speech Coding”, *IEEE Transactions on Speech and Audio Processing*, v. 11, n. 5, pp. 413–424, Set. 2003.
- [13] JANG, H., PARK, J. S. “Multiresolution Sinusoidal Model with Dynamic Segmentation for Timescale Modification of Polyphonic Audio Signals”, *IEEE Transactions on Speech and Audio Processing*, v. 13, n. 2, pp. 254–262, Mar. 2005.
- [14] GOODWIN, M. M., VETTERLI, M. “Matching Pursuits and Atomic Signal Models Based on Recursive Filters Banks”, *IEEE Transactions on Signal Processing*, v. 47, n. 7, pp. 1890–1902, Jul. 1999.
- [15] BOSI, M., GOLDBERG, R. E. *Introduction to Digital Audio Coding and Standards*. 1 ed. Dordrecht, The Netherlands, Kluwer, 2003.
- [16] MALLAT, S. *A Wavelet Tour of Signal Processing*. 2 ed. California, USA, Academic Press, 1998.
- [17] DINIZ, P. S. R., DA SILVA, E. A. B., NETTO, S. L. *Processamento Digital de Sinais*. 1 ed. Rio de Janeiro, Brasil, Artmed, 2004.
- [18] MALLAT, S., ZHANG, Z. “Matching Pursuits With Time-Frequency Dictionaries”, *IEEE Transactions on Signal Processing*, v. 41, n. 12, pp. 3397–3415, Dez. 1993.
- [19] MULLEN, J., HOWARD, D., MURPHY, D. “Waveguide Physical Modeling of Vocal Tract Acoustics: Flexible Formant Bandwidth Control from Increased Model Dimensionality”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 14, n. 3, pp. 964–971, Mai. 2006.
- [20] LEGGE, K. A., FLETCHER, N. H. “Nonlinear Generation of Missing Modes on a Vibrating String”, *Journal of Acoustical Society of America*, v. 76, pp. 5–12, Jul. 1984.
- [21] FLETCHER, N. H. “The Nonlinear Physics of Musical Instruments”, *Reports of Progress in Physics*, v. 62, pp. 723–764, Mai. 1999.

- [22] DAVIS, G. *Adaptive Nonlinear Approximations*. Ph.D. Thesis, Department of Mathematics, New York University, New York, USA, 1994.
- [23] NEFF, R., ZAKHOR, A. “Modulus Quantization for Matching-Pursuit Video Coding”, *IEEE Transactions on Circuits and Systems for Video Technology*, v. 10, n. 6, pp. 895–912, Set. 2000.
- [24] FROSSARD, P., VANDERGHEYNST, P., FIGUERAS I VENTURA, R. M., et al. “A Posteriori Quantization of Progressive Matching Pursuit Streams”, *IEEE Transactions on Signal Processing*, v. 52, n. 2, pp. 525–535, Fev. 2004.
- [25] TCHEOU, M. P. *Análise e Representação de Sinais de Oscilografia Usando Decomposições Adaptativas Redundantes*. Dissertação de Mestrado, COPPE/UFRJ, Rio de Janeiro, Brasil, 2005.
- [26] RODRIGUES, M. A. M., FIGUEIREDO, M. V. F., MIRANDA, A. L. L., et al. “Oscillography for power system operational planning”. In: *Proceedings of the VII SEPOPE – Symposium of Specialists in Electric Operational and Expansion Planning*, Mai. 2000.
- [27] RODRIGUES, M. A. M., OLIVEIRA, J. C. C., MIRANDA, A. L. L., et al. “Evaluation of protection system performance using DFR and relay com-trade files”. In: *Proceedings of the Cigre Session 42*, Paris, France, 2008.
- [28] ITU-R. *Rec. BS.1387: Method for Objective Measurements of Perceived Audio Quality*. International Telecommunication Union, Geneva, Switzerland, 1998.
- [29] THIEDE, T., TREURNIET, W. C., BITTO, R., et al. “PEAQ - The ITU Standard for Objective Measurement of Perceived Audio Quality”, *Journal of Audio Engineering Society*, v. 48, n. 1/2, pp. 3–29, Fev. 2000.
- [30] KABAL, P. *An Examination and Interpretation of ITU-R BS.1387: Perceptual Evaluation of Audio Quality*. Technical report, McGill University, 2002.
- [31] CAMPBELL, D., JONES, E., GLAVIN, M. “Audio Quality Assessment Techniques—A Review, and Recent Developments”, *Elsevier Signal Processing*, v. 89, n. 8, pp. 1489–1500, Ago. 2009.
- [32] VERA-CANDEAS, P., RUIZ-REYES, N., ROSA-ZURERA, M., et al. “Transient Modeling by Matching Pursuits with a Wavelet Dictionary for Parametric Audio Coding”, *IEEE Signal Processing Letters*, v. 11, n. 3, pp. 349–352, Mar. 2004.

- [33] GRIBONVAL, R., BACRY, E. “Harmonic Decomposition of Audio Signals With Matching Pursuit”, *IEEE Transactions on Signal Processing*, v. 51, n. 1, pp. 101–111, Jan. 2003.
- [34] HEUSDENS, R., VAFIN, R., KLEIJN, W. B. “Sinusoidal Modeling Using Psychoacoustic-Adaptive Matching Pursuits”, *IEEE Signal Processing Letters*, v. 9, n. 8, pp. 262–265, Ago. 2002.
- [35] JAGGI, S., CARL, W. C., MALLAT, S., et al. “High Resolution Pursuit for Feature Extraction”, *Applied and Computational Harmonic Analysis*, v. 5, n. 4, pp. 428–449, Out. 1998.
- [36] DONOHO, D. L., VETTERLI, M., DEVORE, R. A., et al. “Data Compression and Harmonic Analysis”, *IEEE Transactions on Information Theory*, v. 44, n. 6, pp. 2435–2476, Out. 1998.
- [37] ENGAN, K., AASE, S. O., HUSOY, J. H. “Multi-Frame Compression: Theory and Design”, *Elsevier Signal Processing*, v. 80, pp. 2121–2140, Out. 2000.
- [38] AL-SHAYKH, O. K., MILOSLAVSKY, E., NOMURA, T., et al. “Video Compression Using Matching Pursuits”, *IEEE Transactions on Circuits and Systems for Video Technology*, v. 9, n. 1, pp. 123–143, Fev. 1997.
- [39] CAETANO, R., DA SILVA, E. A. B., CIANCIO, A. G. “Matching pursuits video coding using generalized bit-planes”. In: *Proceedings of the ICIP’02 – International Conference on Image Processing*, v. 3, pp. 677–680, Rochester, USA, Set. 2002. IEEE.
- [40] GRIBONVAL, R. “From Projection Pursuit and CART to Adaptive Discriminant Analysis?” *IEEE Transactions on Neural Networks*, v. 16, n. 3, pp. 522–532, Mai. 2005.
- [41] TCHEOU, M. P., LOVISOLO, L., DA SILVA, E. A. B., et al. “Optimum Rate-Distortion Dictionary Selection for Compression of atomic decompositions of Electric Disturbance Signals”, *IEEE Signal Processing Letters*, v. 14, n. 2, pp. 81–84, Fev. 2007.
- [42] LOVISOLO, L., TCHEOU, M. P., DA SILVA, E. A. B., et al. “Modeling of Electric Disturbance Signals Using Damped Sinusoids via Atomic Decompositions and its Applications”, *EURASIP Journal on Advances in Signal Processing*, v. 2007, pp. Article ID 29507, 15 pp., 2007.

- [43] MALLAT, S., ZHANG, Z. *Matching Pursuits With Time-Frequency Dictionaries*. Technical report, Courant Institute of Mathematical Sciences, New York University, New York, USA, 1992.
- [44] SAYOOD, K. *Introduction to Data Compression*. 2 ed. San Francisco, USA, Morgan Kaufman, 2000.
- [45] ORTEGA, A., RAMCHANDRAN, K. “Rate-Distortion Methods for Image and Video Compression”, *IEEE Signal Processing Magazine*, v. 15, n. 6, pp. 23–50, Nov. 1998.
- [46] BONNANS, J. F., GILBERT, J. C., LEMARÉCHAL, et al. *Numerical Optimization: Theoretical and Practical Aspect*. 2 ed. Berlin, Germany, Springer-Verlag, 2006.
- [47] IZMAILOV, A., SOLODOV, M. *Otimização - Volume 1. Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*. 1 ed. Rio de Janeiro, Brasil, IMPA, 2005.
- [48] SCHWEITZER III, E. O., HOU, D. “Filtering for protective relays”. In: *Proceedings of the WESCANEX’93 – Communications, Computers and Power in the Modern Environment*, pp. 15–23, Saskatoon, Canada, Mai. 1993. IEEE.
- [49] WIOT, D. “A New Adaptive Transient Monitoring Scheme for Detection of Power System Events”, *IEEE Transactions on Power Delivery*, v. 19, n. 1, pp. 42–48, Jan. 2004.
- [50] LOBOS, T., RESMER, J., KOGLIN, H. J. “Analysis of power system transients using wavelets and prony method”. In: *Proceedings of the Porto Power Tech Conference*, v. 4, p. 4, Porto, Portugal, Set. 2001. IEEE.
- [51] TAWFIK, M. M., MORCOS, M. M. “ANN-Based Techniques for Estimating Fault Location on Transmission Lines using Prony Method”, *IEEE Transactions on Power Delivery*, v. 16, n. 2, pp. 219–224, Abr. 2001.
- [52] BUJANOWSKI, B., PIERRE, J., HIETPAS, S., et al. “A comparison of several system identification methods with application to power systems”. In: *Proceedings of the 36th Midwest Symposium on Circuits and Systems*, v. 1, pp. 64–67, Detroit, USA, Ago. 1993. IEEE.
- [53] GALLI, A. W., HEYDT, G. T., RIBEIRO, P. F. “Exploring the Power of Wavelet Analysis”, *IEEE Computer Applications in Power*, v. 9, n. 4, pp. 37–41, Out. 1996.

- [54] CHUNG, J., POWERS, E. J., GRADY, W. M., et al. “Electric power transient disturbance classification using wavelet-based hidden Markov models”. In: *Proceedings of the ICASSP’00 – International Conference on Acoustics, Speech, and Signal Processing*, v. 6, pp. 3662–3665, Istanbul, Turkey, Jun. 2000. IEEE.
- [55] PILLAY, P., BHATTACHARJEE, A. “Application of Wavelets to Model Short-Term Power System Disturbances”, *IEEE Transactions on Power Systems*, v. 11, n. 4, pp. 2031–2037, Nov. 1996.
- [56] POISSON, O., RIOUAL, P., MEUNIER, M. “Detection and Measurement of Power Quality Disturbances Using Wavelet Transforms”, *IEEE Transactions on Power Delivery*, v. 15, n. 3, pp. 1039–1044, Jul. 2000.
- [57] YANG, H., LIAO, C. “A De-Noising Scheme for Enhancing Wavelet-Based Power Quality Monitoring Systems”, *IEEE Transactions on Power Delivery*, v. 16, n. 3, pp. 353–360, Jul. 2001.
- [58] SANTOSO, S., GRADY, W. M., POWERS, E. J., et al. “Characterization of Distribution Power Quality Events with Fourier and Wavelets Transforms”, *IEEE Transactions on Power Delivery*, v. 15, n. 1, pp. 247–254, Jan. 2000.
- [59] KARIMI, M., MOKHTARI, H., IRAVANI, M. R. “Wavelet Based On-Line Disturbance Detection for Power Quality Applications”, *IEEE Transactions on Power Delivery*, v. 15, n. 4, pp. 1212–1220, Out. 2000.
- [60] IBRAHIM, W. R. A., MORCOS, M. M. “Artificial Intelligence and Advanced Mathematical Tools for Power Quality Applications: A Survey”, *IEEE Transactions on Power Delivery*, v. 17, n. 2, pp. 668–673, Abr. 2002.
- [61] COLLINS, J. J., HURLEY, W. G. “Application of expert systems and neural networks to the diagnosis of power quality problems”. In: *Proceedings of the PQA’94 – 3rd International Conference on Power Quality*, Amsterdam, A-2.03, Out. 1994.
- [62] GOSH, A. K., LUBKEMAN, D. L. “The Classification of Power System Disturbance Waveforms Using a Neural Network Approach”, *IEEE Transactions on Power Delivery*, v. 10, n. 1, pp. 109–115, Jan. 1995.
- [63] LOVISOLO, L. *Representações Coerentes de Sinais Elétricos*. Dissertação de Mestrado, PEE/COPPE, UFRJ, Rio de Janeiro, Brasil, 2001.

- [64] XU, W. “Component Modelling Issues for Power Quality Assessment”, *IEEE Power Engineering Review*, v. 21, n. 11, pp. 12–15, Nov. 2001.
- [65] FRIEDLANDER, B., PORAT, B. “Detection of Transient Signals By the Gabor Representation”, *IEEE Transactions on Acoustics, Speech, Signal Processing*, v. 37, n. 2, pp. 169–180, Fev. 1989.
- [66] FRIEDLANDER, B., ZEIRA, A. “Oversampled Gabor Representation for Transient Signals”, *IEEE Transactions on Signal Processing*, v. 43, n. 9, pp. 2088–2094, Set. 1995.
- [67] ZIBULSKI, M., ZEEVI, Y. Y. “Discrete Multiwindow Gabor-Type Transforms”, *IEEE Transactions on Signal Processing*, v. 45, n. 6, pp. 1428–1442, Jun. 1997.
- [68] KAUPPINEN, I., KAUPPINEN, J., SAARINEN, P. “A Method for Long Extrapolation of Audio Signals”, *Journal of Audio Engeneering Society*, v. 49, n. 12, pp. 1167–1179, Dez. 2002.
- [69] LU, Y., JOSHI, S., MORRIS, J. M. “Noise Reduction for NMR FID via Gabor Expansion”, *IEEE Transactions on Biomedical Engeneering*, v. 44, n. 6, pp. 512–528, Jun. 1997.
- [70] KAY, S. M. *Modern Spectral Estimation*. 1 ed. Englewood Cliffs, USA, Prentice-Hall, 1988.
- [71] TUFTS, D. W., KUMARESAN, R. “Improved Spectral Resolution”, *Proceedings of the IEEE*, v. 68, n. 3, pp. 419–420, Mar. 1980.
- [72] ABATZOGLOU, T. J. “A Fast Maximum Likelihood Algorithm for Frequency Estimation of a Sinusoid based on Newton’s Method”, *IEEE Transactions on Acoustics, Speech, Signal Processing*, v. 33, n. 1, pp. 77–89, Fev. 1985.
- [73] KUMARESAN, R., TUFTS, D. W. “Estimating the Parameters of Exponentially Damped Sinusoids and Pole-Zero Modelling in Noise”, *IEEE Transactions on Acoustics, Speech, Signal Processing*, v. 30, n. 6, pp. 833–840, Dez. 1982.
- [74] HUA, Y., SARKAR, T. K. “Matrix Pencil Method for Estimating the Parameters of Exponentially Damped/Undamped Sinusoids in Noise”, *IEEE Transactions on Acoustics, Speech, Signal Processing*, v. 38, n. 5, pp. 814–824, Mai. 1990.

- [75] HUA, Y. “Parameter Estimation of Exponentially Damped Sinusoids Using Higher Order Statistics and Matrix Pencil”, *IEEE Transactions on Acoustics, Speech, Signal Processing*, v. 39, n. 7, pp. 1691–1692, Jul. 1991.
- [76] PAPADOPOULOS, C. K., NIKIAS, C. L. “Parameter Estimation of Exponentially Damped Sinusoids Using Higher Order Statistics”, *IEEE Transactions on Acoustics, Speech, Signal Processing*, v. 38, n. 8, pp. 1424–1436, Ago. 1990.
- [77] RUIZ, D. P., CARRIÓN, M. C., GALLEGO, A., et al. “Parameter Estimation of Exponentially Damped Sinusoids Using a Higher Order Correlation-Based Approach”, *IEEE Transactions on Signal Processing*, v. 43, n. 11, pp. 2665–2677, Nov. 1995.
- [78] RIBEIRO, M. V., ROMANO, J. M. T., DUQUE, C. “An enhanced data compression method for applications in power quality analysis”. In: *Proceedings of the IECON’01 – 27th Annual Conference of the IEEE Industrial Electronics Society*, v. 1, pp. 676–681, Denver, USA, Nov. 2001. IEEE.
- [79] RIBEIRO, M. V., PARK, S. H., , et al. “A Novel MDL-Based Compression Method for Power Quality Applications”, *IEEE Transactions on Power Delivery*, v. 22, n. 1, pp. 27–36, Jan. 2007.
- [80] FERRANDO, S. E., KOLASA, L. A., KOVACEVIC, N. “Algorithm 820: A Flexible Implementation of Matching Pursuit for Gabor Functions on the Interval”, *ACM Transactions on Mathematical Software (TOMS)*, v. 28, n. 3, pp. 337–353, Set. 2002.
- [81] DEVORE, R. A., TEMLYAKOV, V. N. “Some Remarks in Greedy Algorithms”, *Advances in Computational Mathematics*, v. 5, n. 1, pp. 173–187, 1996.
- [82] DERRIEN, O. “Multi-scale frame-based analysis of audio signals for musical transcription using a dictionary of chromatic waveforms”. In: *Proceedings of the ICASSP’06 – International Conference on Acoustics, Speech and Signal Processing*, v. 5, pp. 57–60, Toulouse, France, Mai. 2006. IEEE.
- [83] LOVISOLO, L., RODRIGUES, M. A. M., DA SILVA, E. A. B., et al. “Coherent decompositions of power systems signals using damped sinusoids with applications to denoising”. In: *Proceedings of the ISCAS’03 – International Symposium on Circuits and Systems*, v. 5, pp. 685–688, Scottsdale, USA, Mai. 2002. IEEE.

- [84] KARIMI-GHARTEMANI, M., IRAVANI, M. R. “A Signal Processing Module for Power System Applications”, *IEEE Transactions on Power Delivery*, v. 18, n. 4, pp. 1118–1126, Out. 2003.
- [85] DUQUE, C. A., RIBEIRO, M. V., RAMOS, F. R., et al. “Power Quality Event Detection Based on the Divide and Conquer Principle and Innovation Concept”, *IEEE Transactions on Power Delivery*, v. 20, n. 4, pp. 2361–2369, Out. 2005.
- [86] CERQUEIRA, A. S., DUQUE, C. A., TRINDADE, R. M., et al. “Digital system for detection and classification of electrical events”. In: *Proceedings of the ISCAS’05 – International Symposium on Circuits and Systems*, v. 6, pp. 5417–5420, Kobe, Japan, Mai. 2005.
- [87] CANADIAN-AMERICAN EMTP USER GROUP. *EMTP Rule Book, Alternative Transients Rule Book*. West Linn, USA, Canadian-American EMTP User Group, 1987-1992.
- [88] MAEZONO, P. K. *Análise de Perturbações: Conceitos e Conhecimentos Básicos*. Schweitzer Engineering Laboratories, Jul. 2001.
- [89] RODRIGUES, M. A. M., MANTUANO FILHO, S., VILELA, J. M. F. “Ferramentas computacionais tradicionais e inteligentes para a análise de perturbações em sistemas de potência”. In: *XIV Seminário Nacional de Produção e Transmissão de Energia Elétrica*, Belém, Brasil, 1997.
- [90] SANTOSO, S., POWERS, E., GRADY, W. “Power Quality Disturbance Data Compression Using Wavelet Transform Methods”, *IEEE Transactions on Power Delivery*, v. 12, n. 3, pp. 1250–1257, Jul. 1997.
- [91] LITTLER, T. B., MORROW, D. J. “Wavelets for the Analysis and Compression of Power System Disturbances”, *IEEE Transactions on Power Delivery*, v. 14, n. 2, pp. 358–364, Abr. 1999.
- [92] DASH, P., PANIGRAHI, B., SAHOO, D., et al. “Power Quality Disturbance Data Compression, Detection, and Classification using Integrated Spline Wavelet and S-transform”, *IEEE Transactions on Power Delivery*, v. 18, n. 2, pp. 595–600, Abr. 2003.
- [93] HUANG, S.-J., JOU, M.-J. “Application of Arithmetic Coding for Electric Power Disturbance Data Compression with Wavelet Packet Enhancement”, *IEEE Transactions on Power Systems*, v. 19, n. 3, pp. 1334–1341, Ago. 2004.

- [94] WANG, Z., BOVIK, A. C., SHEIKH, H. R., et al. “Image Quality Assessment: From Error Visibility To Structural Similarity”, *IEEE Transactions on Image Processing*, v. 13, n. 4, pp. 600–612, Abr. 2004.
- [95] RODRIGUES, M. A. M. “Ferramenta de análise automática de faltas em linhas de transmissão”. . Comunicação interna.
- [96] KEZUNOVIC, M., POPOVIC, T., SEVCIK, D. R., et al. “Requirements for automated fault and disturbance data analysis”. In: *Proceedings of the CIGRE Colloquium*, Sydney, Australia, Set. 2003.
- [97] PHADKE, A. G., THORP, J. S. *Computer Relaying of Power Systems*. 2 ed. Chichester, England, Wiley, 1988.
- [98] TAKAGI, T., YAMAKOSHI, Y., BABA, J., et al. “A New Algorithm of an Accurate Fault Location for EHV/UHV Transmission Lines: Part I - Fourier Transformation Method”, *IEEE Transactions on Power Apparatus and Systems*, v. PAS-100, n. 3, pp. 1316–1323, Mar. 1981.
- [99] TAKAGI, T., YAMAKOSHI, Y., BABA, J., et al. “A New Algorithm of an Accurate Fault Location for EHV/UHV Transmission Lines: Part II - Laplace Transformation Method”, *IEEE Transactions on Power Apparatus and Systems*, v. PAS-101, n. 3, pp. 564–573, Mar. 1982.
- [100] GRAINGER, J. J., STEVENSON, W. D. *Power System Analysis*. 1 ed. USA, McGraw-Hill, 1994.
- [101] IEEE. *C37.111-1999: Standard Common Format for Transient Data Exchange (COMTRADE) for Power Systems*. Institute of Electrical and Electronics Engineering, Mar. 1999.
- [102] VERMA, T., MENG, T. “A 6Kbps to 85Kbps scalable audio coder”. In: *Proceedings of the ICASSP’00 – International Conference on Acoustics, Speech, and Signal Processing*, v. 2, pp. 877–880, Istanbul, Turkey, Jun. 2000. IEEE.
- [103] VAFIN, R., HEUSDEN, R., KLEIJN, W. “Modifying transients for efficient coding of audio”. In: *Proceedings of the ICASSP’01 – International Conference on Acoustics, Speech and Signal Processing*, v. 5, pp. 3285–3288, Salt Lake City, USA, Mai. 2001. IEEE.
- [104] RAVELLI, E., DAUDET, L. “Representations of audio signals in overcomplete dictionaries: what is the link between redundancy factor and coding pro-

- perties?” In: *Proceedings of the DAFx'06 – 9th International Conference on Digital Audio Effects*, pp. 267–270, Montreal, Canada, Set. 2006.
- [105] CHRISTENSEN, M. G., VAN DE PAR, S. “Efficient Parametric Coding of Transients”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 14, n. 4, pp. 1340–1351, Jul. 2006.
- [106] CHRISTENSEN, M. G., JAKOBSSON, A., ANDERSEN, S. V., et al. “Amplitude Modulated Sinusoidal Signal Decomposition for Audio”, *IEEE Signal Processing Letters*, v. 13, n. 7, pp. 389–392, Jul. 2006.
- [107] DAUDET, L. “Sparse and Structured Decompositions of Signals with Molecular Matching Pursuit”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 14, n. 5, pp. 1890–1902, Set. 2006.
- [108] CHRISTENSEN, M. G., JENSEN, S. H. “The cyclic matching pursuit and its application to audio modeling and coding”. In: *Proceedings of the ACSSC'07 – Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, USA, Nov. 2007.
- [109] GOODWIN, M. “Multichannel matching pursuit and applications to spatial audio coding”. In: *Proceedings of the ACSSC'06 – 40th Asilomar Conference on Signals, Systems and Computers*, pp. 1114–1118, Pacific Grove, USA, Out. 2006.
- [110] GRIBONVAL, R. “Sparse decomposition of stereo signals with matching pursuit and application to blind separation of more than two sources from a stereo mixture”. In: *Proceedings of the ICASSP'02 – International Conference on Acoustics, Speech, and Signal Processing*, v. 3, pp. 3057–3060, Orlando, USA, Mai. 2002. IEEE.
- [111] UMAPATHY, K., KRISHNAN, S., JIMAA, S. “Multigroup Classification of Audio Signals Using Time-Frequency Parameters”, *IEEE Transactions on Multimedia*, v. 7, n. 2, pp. 308–315, Abr. 2005.
- [112] HEUSDENS, R., VAN DE PAR, S. “Rate-distortion optimal sinusoidal modeling of audio and speech using psychoacoustical matching pursuits”. In: *Proceedings of the ICASSP'02 – International Conference on Acoustics, Speech, and Signal Processing*, v. 2, pp. 1809–1812, Orlando, USA, Mai. 2002. IEEE.
- [113] BJORNBERG, D., AGILI, S., MORALES, A. “Decomposition and recognition of a multi-channel audio source using matching pursuit algorithm”. In:

Proceedings of the ISCAS'04 – International Symposium on Circuits and Systems, v. 5, pp. 624–627, Vancouver, Canada, Mai. 2004. IEEE.

- [114] GOODWIN, M. “Matching pursuit with damped sinusoids”. In: *Proceedings of the ICASSP'97 – International Conference on Acoustics, Speech, and Signal Processing*, v. 3, pp. 2037–1997, Munich, Germany, Abr. 1997. IEEE.
- [115] KRSTULOVIC, S., GRIBONVAL, R., LEVEAU, P., et al. “A comparison of two extensions of the matching pursuit algorithm for the harmonic decomposition of sounds”. In: *Proceedings of the WASPAA'05 – Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 259 – 262, New Paltz, USA, Out. 2005. IEEE.
- [116] STURM, B., SHYNK, J., GAUGLITZ, S. “Agglomerative clustering in sparse atomic decompositions of audio signals”. In: *Proceedings of the ICASSP'08 – International Conference on Acoustics, Speech and Signal Processing*, pp. 97–100, Las Vegas, USA, Mar. 2008. IEEE.
- [117] KRSTULOVIC, S., GRIBONVAL, R. “Mptk: matching pursuit made tractable”. In: *Proceedings of the ICASSP'06 – International Conference on Acoustics, Speech and Signal Processing*, v. 3, pp. 496–499, Toulouse, France, Mai. 2006. IEEE.
- [118] KAUPPINEN, I., ROTH, K. “Audio signal extrapolation - theory and applications”. In: *Proceedings of the DAFX'02 – 5th International Conference on Digital Audio Effects*, pp. 105–110, Hamburg, Germany, Set. 2002.
- [119] PAINTER, T., SPANIAS, A. “Perceptual Coding of Digital Audio”, *Proceedings of the IEEE*, v. 88, n. 4, pp. 451–515, Abr. 2000.
- [120] OPOLKO, F., WAPNICK, J. “McGill University Master Samples Collection”. . Montreal, Q.C., Canada, McGill University, 1987.
- [121] GOTO, M., NISHIMURA, T. “RWC music database: music genre database and musical instrument sound database”. In: *Proceedings of the ISMIR'03 – International Conference on Music Information Retrieval*, pp. 229–230, Washington, USA, Out. 2003.
- [122] HANS, M., SCHAFER, R. “Lossless Compression of Digital Audio”, *IEEE Signal Processing Magazine*, v. 18, n. 4, pp. 21–32, Jul. 2001.

- [123] SCHIJNDEL, N. H. V., BENZA, J., CHRISTENSEN, M. G., et al. “Adaptive RD Optimized Hybrid Sound Coding”, *Journal of Audio Engineering Society*, v. 56, n. 10, pp. 787–809, Oct. 2008.
- [124] DEN BRINKER, A. C., BREEBAART, J., EKSTRAND, P., et al. “An Overview of the Coding Standard MPEG-4 Audio Amendments 1 and 2: HE-AAC, SSC, and HE-AAC v2”, *EURASIP Journal on Audio Speech Music Processing*, v. 2009, pp. 3:1–3:21, Jan. 2009.
- [125] CHRISTENSEN, M. G. *Estimation and Modeling Problems in Parametric Audio Coding*. Ph.D. Thesis, Aalborg University, Aalborg, Denmark, Jul. 2005.
- [126] RAVELLI, E., RICHARD, G., DAUDET, L. “Union of MDCT Bases for Audio Coding”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 16, n. 8, pp. 1361–1372, Nov. 2008.
- [127] RAUBER, T., RUNGER, G. *Parallel programming for multicore and cluster systems*. 1 ed. Berlin, Germany, Springer-Verlag, 2010.
- [128] BELL, T. C., WITTEN, I. H., CLEARY, J. G. *Text compression / Timothy C. Bell, John G. Cleary, Ian H. Witten*. Englewood Cliffs, USA, Prentice Hall, 1990.
- [129] SAID, A. *Introduction to Arithmetic Coding - Theory and Practice*. Technical report, HP Laboratories Palo Alto, 2004.
- [130] PRESS, W., FLANNERY, B., TEUKOLSKY, S., et al. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge, England, Cambridge, Oct. 1992.
- [131] MALLAT, S. “A Theory for Multiresolution Signal Decomposition: the Wavelet Representation”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 11, n. 7, pp. 674–693, Jul. 1989.
- [132] SHARIFI, K., LEON-GARCIA, A. “Estimation of Shape Parameter for Generalized Gaussian Distributions in Subband Decompositions of Video”, *IEEE Transactions on Circuits and Systems for Video Technology*, v. 5, n. 1, pp. 52–56, Feb. 1995.
- [133] BEERENDS, JOHN G.; STEMERDINK, J. A. “A Perceptual Audio Quality Measure Based on a Psychoacoustic Sound Representation”, *Journal of Audio Engineering Society*, v. 40, n. 12, pp. 963–978, 1992.

- [134] HUBER, R., KOLLMEIER, B. “PEMO-Q—A New Method for Objective Audio Quality Assessment Using a Model of Auditory Perception”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 14, n. 6, pp. 1902–1911, Nov. 2006.