REPRESENTAÇÕES DE SINAIS USANDO DICIONÁRIOS REDUNDANTES

Lisandro Lovisolo

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Aprovada por:

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

Prof. Paulo Sergio Ramirez Diniz, Ph.D.

Prof. Luiz Wagner Pereira Biscainho, D.Sc.

Prof. Marcos Craizer, D.Sc.

Prof. Ricardo Lopes de Queiroz, Ph.D.

Dr. Marco Antônio Macciola Rodrigues, D.Sc.

RIO DE JANEIRO, RJ - BRASIL
AGOSTO DE 2006

LOVISOLO, LISANDRO

Representações de Sinais Usando Dicionários Redundantes, *Signal Representations Using Overcomplete Dictionaries* [Rio de Janeiro] 2006

XV, 237 pp 29,7 cm (COPPE/UFRJ, D. Sc., Engenharia Elétrica, 2006)

Tese - Universidade Federal do Rio de Janeiro, COPPE

1. Representações de sinais 2. Representações adaptativas 3. Dicionários redundantes 4. Decomposições Coerentes 5. Quantização de decomposições vorazes 6. *Matching Pursuit* 7. Estatística dos ângulos no *Matching Pursuit* 8. *Frames* 9. Conteúdo tempo-freqüência de *frames* 10. *Frames* de Weyl-Heisenberg 11. Dicionários a partir de *frames* 12. Senóides amortecidas

I.COPPE/UFRJ II.Título (Série)

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

REPRESENTAÇÕES DE SINAIS USANDO DICIONÁRIOS REDUNDANTES

Lisandro Lovisolo

Agosto/2006

Orientadores: Eduardo Antônio Barros da Silva
Paulo Sérgio Ramirez Diniz

Programa: Engenharia Elétrica

Nesta tese, representações de sinais usando dicionários redundantes são empregadas a sinais de distúrbios elétricos. O processo de decomposição baseia-se num algoritmo iterativo e voraz conhecido como *Matching Pursuit* e utiliza um dicionário de senóides amortecidas. A cada iteração ele seleciona um átomo do dicionário para ser adicionado à representação. As representações obtidas são empregadas em tarefas de processamento de sinais como filtragem e compressão com perdas. Quando algoritmos de decomposição como o *Matching Pursuit* são empregados em esquemas de compressão com perdas, faz-se necessária a quantização eficiente dos coeficientes na representação. Apresentamos o projeto de quantizadores de Lloyd-Max para esses coeficientes. Tal projeto utiliza um modelo estatístico do ângulo entre o sinal a ser decomposto e o átomo escolhido pelo critério voraz. Os quantizadores de Lloyd-Max projetados apresentam bom desempenho taxa×distorção. Apresentamos um teorema que mostra que se o dicionário empregado no *Matching Pursuit* contém uma base ortonormal, então o *Matching Pursuit* possui uma probabilidade não-nula de produzir uma representação exata do sinal com um número finito de iterações. Definimos ainda o conteúdo tempo-freqüência de um *frame* como o somatório das distribuições de Wigner-Ville de seus elementos. Tal definição permite determinar se um dado conjunto de elementos em um espaço de Hilbert é um *frame* deste espaço. A análise de *frames* de Weyl-Heisenberg gerados a partir de uma função par permite povoar efetivamente dicionários formados por elementos de *frames* de Weyl-Heisenberg, fornecendo dicionários que obtêm bons desempenhos em algoritmos de decomposições vorazes.

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

SIGNAL REPRESENTATIONS USING OVERCOMPLETE DICTIONARIES

Lisandro Lovisolo

August/2006

Advisors: Eduardo Antônio Barros da Silva
Paulo Sérgio Ramirez Diniz

Department: Electrical Engineering

In this thesis, signal representations using overcomplete dictionaries are applied to the representation of electric disturbance signals. The decomposition process is based on an iterative greedy decomposition algorithm known as Matching Pursuit and employs a dictionary of damped sinusoids. At each iteration, it selects a new atom from the dictionary to be summed up in the signal representation. The signal representations obtained are then employed for signal processing tasks including filtering and lossy compression. When Matching-Pursuit-like decompositions algorithms are used for lossy signal compression, efficient quantization of the coefficients employed in the signal representation is required. We present the design of Lloyd-Max quantizers for these coefficients. This design uses a statistical model for the angle between the signal to be decomposed and the atom selected in the greedy loop. The Lloyd-Max quantizers designed are shown to perform very well in a rate×distortion sense. We provide a theorem showing that if the dictionary includes an orthonormal basis then the Matching Pursuit has a non-zero chance of obtaining exact signal expansions using a finite number of dictionary atoms. We also define the time-frequency content of a frame as the summation of the individual Wigner-Ville distributions of the elements in the frame. This definition allows to determine if a given set of elements in a Hilbert space is a frame in this space. The analysis of Weyl-Heisenberg frames generated from even prototype functions allows to effectively populate dictionaries formed by elements of Weyl-Heisenberg frames, yielding dictionaries with good rate×distortion performance in greedy decomposition algorithms.

Sumário

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

1.1 Representação Atômica de Sinais

Os gregos nos legaram o conceito de átomo – unidades indivisíveis de matéria. Uma metodologia largamente utilizada pela humanidade é: primeiro, dividir o mundo natural em pedaços ou partes - átomos; a seguir, escolher um subconjunto dessas partes (desprezando outras) e utilizá-lo para construir um modelo aproximado do mundo natural. Os impactos dessa metodologia na evolução da ciência são enormes; por exemplo, a idéia originalmente apresentada por Erwin Schrödinger em 1943 de que características hereditárias estariam armazenadas por combinações limitadas de cristais foi confirmada ao desvendar-se a estrutura do DNA. Representações atômicas de sinais assemelham-se a tal metodologia, pois objetivam extrair partes do sinal que permitam construir uma boa aproximação do mesmo.

A representação atômica de sinais consiste da utilização de formas de onda pré-definidas (átomos) para expressar sinais. Desta forma, um sinal $\mathbf{x}$ deverá ser representado por uma combinação linear de poucas formas de onda selecionadas a partir de um conjunto pré-definido, o dicionário $\mathcal{D}$. Os sinais $\mathbf{g}_i$, que estão incluídos em $\mathcal{D}$, são as formas de onda que podem ser utilizadas na combinação linear, e são chamados de átomos ou elementos. Algoritmos que obtêm representações atômicas de sinais escolhem um subconjunto de M elementos $\mathbf{g}_{i(m)}$ do dicionário $\mathcal{D}$ que aproxima um sinal $\mathbf{x}$ via

$$\mathbf{x} \approx \hat{\mathbf{x}} = \sum_{m=1}^{M} \gamma_m \mathbf{g}_{i(m)}, \quad \mathbf{g}_{i(m)} \in \mathcal{D}. \tag{1.1}$$

Cada átomo $\mathbf{g}_{i(m)}$, indexado por $i(m)$, utilizado na aproximação de $\mathbf{x}$ pode ser interpretado como uma característica intrínseca do sinal $\mathbf{x}$ de "peso" γ_m, possibilitando assim tanto a codificação compacta como a estimação de características do sinal.

Essa situação encontra paralelo na linguagem oral e escrita. Um dicionário contém as palavras de uma língua, e línguas são redundantes, pois permitem expressar uma idéia de diferentes formas. Um dicionário redundante contém uma quantidade de elementos maior que a necessária para gerar o espaço de sinais. A redundância permite representar um sinal de várias formas e escolher uma delas para expressar um sinal.

Os átomos (elementos da "tabela periódica") compõem todas as moléculas conhecidas. Apesar de os elementos da molécula de benzeno serem conhecidos, somente após Kekulé sonhar com uma cobra comendo o próprio rabo o ciclo hexacarbono foi compreendido. As propriedades de compostos químicos dependem em grande parte de como seus elementos estão ligados. Similarmente, em decomposições atômicas de sinais, não só os elementos selecionados são relevantes, mas também a forma como eles estão combinados; isto é, procura-se uma combinação linear, como a da equação (A.1), que forneça um boa representação do sinal.

Freqüentemente, classificamos as coisas hierarquicamente. Ao olhar o céu noturno, é raro não pensar sobre as constelações ou aglomerados de estrelas, a seguir sobre as estrelas em si, e depois, talvez, sobre os possíveis planetas que orbitam as estrelas. Utilizamos a Classificação Taxonômica das Espécies para classificar as formas de vida a partir do reino até a espécie. Sistemas de classificação hierárquica agrupam coisas conforme características comuns ou proximidade física, e o agrupamento de coisas a partir de características comuns pode ser interpretado como agrupamento por proximidade física num modelo/espaço conceitual. Decomposições atômicas permitem as duas formas de agrupamento. A proximidade física entre características do sinal pode ser obtida utilizando-se um dicionário que contenha a mesma forma de onda em diversas escalas geométricas. A classificação de sinais conforme suas características comuns pode ser obtida comparando as representações atômicas (os átomos e os pesos) de diferentes sinais.

Atualmente, pesquisas em neurociência e psicologia estão mapeando áreas especializadas do cérebro utilizadas para fins distintos e específicos. Quiçá a atomi-

zação, o particionamento e a classificação do mundo natural sejam inatos à nossa organização cerebral. Nestas tarefas duas questões recorrentes são: **i**) Quais são os blocos constituintes das coisas naturais? **ii**) Como esses blocos estão combinados? Representações atômicas utilizam combinações lineares de elementos pré-definidos para representar sinais. Os possíveis blocos constituintes são determinados pelos elementos do dicionário, enquanto algoritmos de decomposição atômica visam a obter tanto os blocos constituintes de um sinal quanto a forma como eles estão combinados.

Bases são freqüentemente construídas para salientar características específicas de sinais, e possuem uma miríade de aplicações. Entretanto, não há flexibilidade e tampouco liberdade na escolha da representação de um sinal em uma base; como a representação é única, se desejarmos expressar um sinal de maneira diferente, então outra base deverá ser empregada; em outras palavras, quando bases são utilizadas para representar sinais, os elementos utilizados na representação estão previamente determinados. Além disso, algumas aplicações podem requerer a expressão de sinais utilizando elementos que sejam linearmente dependentes (um dos elementos é uma combinação linear dos outros), o que as bases não permitem. Quando uma estratégia de decomposição usando um dicionário redundante é empregada, os átomos na representação do sinal podem ser escolhidos a posteriori, isto é, dependendo do sinal; e os átomos utilizados na representação podem ser linearmente dependentes.

Representações atômicas têm sido empregadas na filtragem e remoção de ruído de sinais [76, 84] e na análise de fenômenos físicos (reconhecimento de padrões e modelagem de sinais [58, 59, 64, 69, 72, 84, 107]), bem como em análises tempo-freqüência [84,85] e harmônica [37,58] de sinais. Representações atômicas são empregadas também em esquemas de compressão de sinais [2, 9, 37, 44, 88]. Recentemente, elas foram utilizadas para discriminar processos Gaussianos [63].

1.2 Expansões em *Frames*

O cérebro é uma grande rede de neurônios, e há grande evidência de que essa rede opera em paralelo. Vários neurônios recebem o mesmo estímulo simultaneamente, produzindo diferentes respostas que são entregues a outros neurônios em diferentes pontos da rede. No fim, a excitação ativa alguns neurônios que determi-

nam a resposta ao estímulo recebido. Ao invés de selecionar alguns átomos, todos os elementos de um dicionário poderiam ser utilizados para representar sinais. Isto traria uma flexibilidade similar à da rede neuronal cerebral, pois o padrão de entrada seria, assim, comparado a (projetado sobre) todos os elementos do dicionário. Tal abordagem leva ao conceito de *frame*.

Duffin e Schaeffer introduziram o conceito de *frame* em seu trabalho sobre séries de Fourier não harmônicas [38]. Nas duas últimas décadas, *frames* tornaram-se uma área de pesquisa muito ativa, que tem produzido diversos resultados teóricos e aplicações [17, 26, 27, 84]. *Frames* podem ser descritos como "bases" redundantes ou sobre-completas; um *frame* de um espaço $\mathbb{H}$ é um conjunto de elementos que gera $\mathbb{H}$. Logo, um *frame* $\mathcal{G} = \{\mathbf{g}_k\}_{k\in\mathcal{K}}$ de $\mathbb{H}$ pode ser utilizado para expressar qualquer $\mathbf{x} \in \mathbb{H}$ via

$$\mathbf{x} = \sum_{k\in\mathcal{K}} c_k \mathbf{g}_k, \tag{1.2}$$

onde c_k são chamados de coeficientes do *frame*. Como $\mathcal{G}$ é sobre-completo o conjunto de coeficientes do *frame* não é único. Uma forma de obter os coeficientes c_k é através da utilização do *frame* inverso ou dual $\tilde{\mathcal{G}} = \{\tilde{\mathbf{g}}_k\}_{k\in\mathcal{K}}$, que fornece as fórmulas de reconstrução

$$\mathbf{x} = \sum_{k\in\mathcal{K}} \langle \mathbf{x}, \mathbf{g}_k \rangle \tilde{\mathbf{g}}_k = \sum_{k\in\mathcal{K}} \langle \mathbf{x}, \tilde{\mathbf{g}}_k \rangle \mathbf{g}_k. \tag{1.3}$$

Como $\mathcal{G}$ é sobre-completo, $\tilde{\mathcal{G}}$, em geral, pode assumir diversas formas [17].

A projeção de um sinal em um *frame* fornece quanto de cada elemento do *frame* há no sinal, permitindo inferir características do sinal. Como os elementos de um *frame* não são obrigatoriamente linearmente independentes e, portanto, tampouco ortogonais, características parecidas entre si podem ser observadas a partir da expansão de um sinal em um *frame*. Quando os elementos de um *frame* são linearmente dependentes, o conjunto de coeficientes do *frame* pode ser alterado para destacar ou enfatizar determinadas características do sinal. Além disso, podemos projetar *frames* para aplicações específicas, dependendo das características que desejamos extrair dos sinais ou analisar neles.

Quando representamos um sinal $\mathbf{x}$ utilizando um *frame*, os elementos da representação estão selecionados a priori; entretanto, a expansão em *frames* permite utilizar elementos com propriedades especiais. O teorema Balian-Low [26] mostra

que não é possível construir uma base com elementos que sejam bem localizados no domínio do tempo e no da freqüência simultaneamente. Como a representação de sinais em um *frame* não necessita ser única, as imposições à definição dos elementos de um *frame* são menos restritivas que para os elementos de bases. Isso permite que *frames* possuam elementos melhor localizados no domínio do tempo e no da freqüência simultaneamente, quando comparados aos elementos de uma base.

As aplicações de *frames*, para citar algumas, estendem-se desde codificação de sinais [84], incluindo a amostragem de sinais [17,38,84], análise de sinais [27,84] e detecção de transitórios [51, 52], até o projeto de sistemas de comunicações [17, 99].

1.3 Estrutura da Tese

Esta tese trata de decomposições atômicas e de expansões em *frames* – representações de sinais usando dicionários redundantes. O Capítulo 2 revisa conceitos básicos sobre representações atômicas e *frames*.

O Capítulo 3 apresenta representações coerentes de sinais adquiridos em sistemas de distribuição de energia. Representações coerentes visam à obtenção de uma aproximação do sinal relacionada aos fenômenos físicos representados no sinal observado.

O Capítulo 4 realiza um estudo estatístico do algoritmo de decomposição de sinais conhecido como *Matching Pursuit*, que constrói a representação do sinal selecionando um elemento a cada iteração. O modelo estatístico é empregado no projeto de quantizadores para representações obtidas com o *Matching Pursuit*.

O Capítulo 5 introduz o conceito de conteúdo tempo-freqüência de *frames*. A partir deste, uma nova condição suficiente para que uma família de elementos seja um *frame* é apresentada. O conceito de conteúdo tempo-freqüência de *frames* é então aplicado na análise de *frames* de Weyl-Heisenberg.

O Capítulo 6 apresenta as conclusões.

Os capítulos 2 a 6 apresentam resumidamente o conteúdo dos apêndices A a F. Os apêndices (escritos em inglês) apresentam mais detalhes, deduções, explicações e resultados. Assim, a leitura desta tese pode restringir-se aos apêndices.

Capítulo 2

Conceitos Básicos

Neste capítulo são discutidos alguns aspectos relativos à representação de sinais utilizando dicionários redundantes. Começa-se pela apresentação de funcionalidades e características tanto importantes como desejáveis de representações atômicas. Depois, foca-se no algoritmo de *Matching Pursuit*, um algoritmo que constrói representações de sinais selecionando um átomo a cada iteração. A seguir, são discutidos dicionários para decomposições atômicas. Então, revisam-se aspectos básicos da teoria de *frames*. Discute-se ainda a utilização de *frames* como dicionários.

2.1 Representações Atômicas

A decomposição atômica de sinais pretende obter um subconjunto de M elementos $\mathbf{g}_{i(m)}$ de um dicionário $\mathcal{D}$ que aproxima um sinal $\mathbf{x}$ através da representação de M termos:

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

2.1.1 Distorção da Aproximação

Podemos definir a distorção em que incorremos ao utilizar uma representação de M termos para aproximar um sinal $\mathbf{x}$ como

$$d(\mathbf{x}, M, \mathcal{D}) = \|\mathbf{x} - \hat{\mathbf{x}}\| = \left\| \mathbf{x} - \sum_{m=1}^{M} \gamma_m \mathbf{g}_{i(m)} \right\|. \tag{2.2}$$

Tal distorção depende de três fatores: i) a quantidade de elementos M utilizados na aproximação; ii) os átomos $\mathbf{g}_{i(m)}$ usados para expressar o sinal; iii) e os pesos γ_i dos

átomos na representação de M termos. Como os átomos utilizáveis na representação de M termos são especificados por $\mathcal{D}$ a distorção depende de $\mathcal{D}$.

2.1.2 Completude do Dicionário

Dicionários devem ser completos. Para que combinações lineares de M termos formadas com átomos de um dicionário $\mathcal{D}$ sejam capazes de representar qualquer sinal $\mathbf{x} \in \mathbb{X}$ com distorção $d(\mathbf{x}, M, \mathcal{D})$ arbitrariamente pequena, $\mathcal{D}$ tem que ser completo em $\mathbb{X}$. Assim, deverá existir pelo menos uma combinação linear de elementos de $\mathcal{D}$ que produza $\hat{\mathbf{x}} = \mathbf{x}$, $\forall\ \mathbf{x} \in \mathbb{X}$, ou seja, $\mathcal{D}$ tem que gerar o espaço $\mathbb{X}$. Dicionários são normalmente ditos sobre-completos ou redundantes, já que, em geral, eles possuem mais elementos que os necessários para gerar o espaço. A redundância do dicionário permite usar diferentes combinações lineares de elementos de $\mathcal{D}$ para expressar um dado $\mathbf{x} \in \mathbb{X}$.

2.1.3 Adaptabilidade

É desejável que os átomos utilizados na representação de um sinal sejam escolhidos de acordo com o sinal, ou seja, adaptativamente. Por isso, algoritmos que obtêm representações por M termos são chamados de algoritmos de decomposição adaptativa de sinais [1, 29, 59, 62, 72, 84, 85, 105]. O uso de um dicionário redundante é um pré-requisito para obter representações adaptativas, pois assim pode-se escolher uma dentre diferentes combinações lineares dos elementos do dicionário para representar um dado sinal. O algoritmo empregado na decomposição influencia a representação por M termos obtida, pois diferentes critérios podem ser usados para selecionar os seus átomos.

2.1.4 Aproximações e Estimações

Aproximações adaptativas discriminam (ou separam) a informação relevante do ruído. A informação relevante é definida pelos elementos que podem ser utilizados na representação, ou seja, pelo dicionário. Assim, decomposições atômicas podem ser entendidas como estimadores de sinais que retêm somente a informação presente no sinal relacionada aos átomos selecionados para expressá-lo.

2.1.5 Compacidade ou Esparsidade

A representação mais compacta ou esparsa de $\mathbf{x}$ é aquela que utiliza o menor número de átomos [36, 59, 66, 105] com distorção nula. Na prática, não se busca necessariamente o menor número de átomos; um pequeno conjunto de átomos que aproxima o sinal com uma distorção aceitável pode ser suficiente. Assim, a esparsidade de uma representação está relacionada ao número de termos M, e à medida que M cresce, menos compacta ou esparsa é a representação.

2.1.6 Representações Coerentes

Quando lidamos com processos físicos, o sinal observado é uma mistura de componentes $\mathbf{p}_m$, que representam fenômenos físicos, dada por

$$\mathbf{x} = \sum_m \beta_m \mathbf{p}_m \ + \mathbf{n}, \tag{2.3}$$

onde $\mathbf{n}$ é o ruído inerente à observação. Quanto mais parecidos forem os átomos $\mathbf{g}_{i(m)}$ e seus pesos γ_m utilizados na representação de um sinal $\mathbf{x}$ com as componentes $\mathbf{p}_m$ e seus pesos β_m presentes em $\mathbf{x}$, melhor será a representação obtida para modelagem do sinal e reconhecimento de padrões; tais decomposições são ditas coerentes.

2.1.7 O Algoritmo de *Matching Pursuit*

O algoritmo de *Matching Pursuit* (MP) [84,85] é um algoritmo voraz [32,104] que aproxima sinais iterativamente. Seja $\mathcal{D} = \{\mathbf{g}_k\}$ com $k \in \{1, \ldots, \#_\mathcal{D}\}$, tal que $\|\mathbf{g}_k\| = 1 \ \forall k$ e $\#_\mathcal{D}$ é a cardinalidade de $\mathcal{D}$, isto é, a quantidade de elementos em $\mathcal{D}$. A cada iteração $n \geq 1$, o MP procura pelo átomo $\mathbf{g}_{i(n)} \in \mathcal{D}$, que possui maior produto interno com o sinal residual $\mathbf{r}_\mathbf{x}^{n-1}$ [84, 85], por isso este algoritmo é dito voraz. O átomo selecionado é subtraído do resíduo, obtendo-se

$$\mathbf{r}_\mathbf{x}^n = \mathbf{r}_\mathbf{x}^{n-1} - \gamma_n \mathbf{g}_{i(n)}; \text{ com } \gamma_n = \langle \mathbf{r}_\mathbf{x}^{n-1}, \mathbf{g}_{i(n)} \rangle, \tag{2.4}$$

o produto interno entre o resíduo a ser decomposto e o átomo selecionado $\mathbf{g}_{i(n)}$. O resíduo inicial é $\mathbf{r}_\mathbf{x}^0 = \mathbf{x}$. O MP obtém então uma representação/aproximação de $\mathbf{x}$ dada pela representação de M termos

$$\mathbf{x} \approx \hat{\mathbf{x}} \approx \sum_{n=1}^{M} \gamma_n \mathbf{g}_{i(n)}. \tag{2.5}$$

O erro ou distorção após M iterações do MP é o M-ésimo resíduo $\mathbf{r}_{\mathbf{x}}^{M} = \mathbf{x} - \hat{\mathbf{x}}$. Na prática as iterações do MP (o cálculo de γ_n, $i(n)$ e $\mathbf{r}_{\mathbf{x}}^{n}$) prosseguem até que um critério de distorção pré-estabelecido como $\|\mathbf{r}_{\mathbf{x}}^{n}\|$, um número máximo de passos n, ou um dado valor para uma medida de aproximação sejam atingidos.

O critério de seleção voraz garante a convergência do MP [29, 84, 112], i.e., $\lim_{n\to\infty} \|\mathbf{r}_{\mathbf{x}}^{n}\| = 0$. Defina o ângulo máximo entre um sinal qualquer $\mathbf{x}$ pertencente ao espaço de sinais $\mathbb{X}$ e o elemento mais próximo de $\mathcal{D}$

$$\Theta(\mathcal{D}) = \arccos\left(\min_{\mathbf{x}\in\mathbb{X}}\left[\max_{i\in\{1,\ldots,\#_{\mathcal{D}}\}}\left(\frac{|\langle \mathbf{x}, \mathbf{g}_i\rangle|}{\|\mathbf{x}\|}\right)\right]\right). \tag{2.6}$$

Mostra-se, ver seção B.2.2, que $\Theta(\mathcal{D})$ fornece um limitante superior para a distorção, num dado passo n do MP, dado por

$$e_n = \|\mathbf{r}_{\mathbf{x}}^{n}\| \leq \|\mathbf{x}\| \sin^n(\Theta(\mathcal{D})). \tag{2.7}$$

2.2 Dicionários

Conceitualmente, o dicionário mais simples é um conjunto finito de elementos $\mathcal{D} = \{\mathbf{g}_i\}_{i\in\mathcal{I}}$. Neste caso, $\mathbf{g}_i \in \mathcal{D}$ pode ser indexado através de $i \in \mathcal{I} = \{1, \ldots, \#_{\mathcal{D}}\}$, onde $\#_{\mathcal{D}}$ é o número de elementos em $\mathcal{D}$, a sua cardinalidade.

Um dicionário completo permite representar qualquer sinal com um erro de aproximação arbitrariamente pequeno. Entretanto, o emprego de um dicionário completo não garante a obtenção de uma representação compacta, tampouco coerente.

A probabilidade de encontrar um átomo no dicionário possuidor de grande semelhança com o sinal a ser decomposto aumenta com $\#_{\mathcal{D}}$ [54, 62, 84]. Logo, a utilização de dicionários com $\#_{\mathcal{D}}$ grande é desejável em algumas aplicações, possibilitando que $\mathcal{D}$ contenha átomos semelhantes a todas as componentes potenciais dos sinais a serem decompostos.

2.2.1 Dicionários Parametrizados

Os próprios átomos, além de $\#_{\mathcal{D}}$, também influenciam a probabilidade de encontrar um átomo em $\mathcal{D}$ possuidor de grande semelhança com as possíveis componentes de um sinal. Assim, um aspecto importante no projeto de $\mathcal{D}$ é a forma de

onda de seus elementos. Se a classe de componentes que podem estar representadas no sinal é conhecida então uma boa estratégia é utilizar átomos que se assemelhem a essas componentes [54,62,84]. Uma estratégia largamente empregada é a definição dos elementos de um dicionário a partir de funções protótipos. Tais dicionários são ditos parametrizados pois cada elemento $\mathbf{g}_\sigma$ é definido pelo valor de um conjunto de parâmetros

$$\sigma = (\sigma_0, \sigma_1, \ldots, \sigma_{K-1}) \in S, \tag{2.8}$$

onde K é a quantidade de parâmetros que definem $\mathbf{g}_\sigma$ e S é o conjunto de todos os valores possíveis de σ. Se S possui um número finito de pontos então o dicionário é indexável. Nesse caso, a codificação de uma representação em M termos (os átomos $\mathbf{g}_{i(m)}$ e seus pesos γ_m, $1 \leq m \leq M$) consiste na codificação dos pesos e dos parâmetros σ dos átomos.

Parâmetros Contínuos Em algumas casos é interessante utilizar parâmetros contínuos. Isso permite adaptar, a cada passo de decomposição n, o conjunto de parâmetros definidor do átomo $\sigma(n)$ ao resíduo a ser decomposto $\mathbf{r}_{\mathbf{x}}^{n-1}$. Dessa forma, $\sigma(n)$ pode assumir qualquer valor dentro de uma região do espaço S. O algoritmo de decomposição apresentado no Capítulo 3 obtém parâmetros contínuos para os átomos.

2.2.2 Compromissos de Dicionários

A cardinalidade do dicionário impacta:

i – A complexidade computacional: em geral, o custo computacional dos algoritmos que obtêm representações atômicas está relacionado a $\#_{\mathcal{D}}$, de forma tal que quanto maior for $\#_{\mathcal{D}}$, maior será, em geral, o custo computacional dos algoritmos.

ii – A taxa de codificação: a taxa necessária para codificar cada um dos átomos utilizados nas representações cresce com $\#_{\mathcal{D}}$.

Em decomposições vorazes, como o MP, alguns dos fatores que influenciam o erro máximo e a taxa de convergência do algoritmo são:

- $\Theta(D)$, ver equação (2.6), que limita a norma do erro em cada passo.

- O ângulo médio $\overline{\Theta}(\mathcal{D})$, entre os vetores provenientes de uma fonte $\mathcal{X}$, que possui densidade uniforme na superfície da esfera (a função de densidade de probabilidade de $\mathbf{x} \in \mathcal{X}$ depende somente de $\|\mathbf{x}\|$, ou seja, $\mathbf{x}$ possui a mesma função densidade de probabilidade em qualquer direção do espaço), e os átomos mais próximos em $\mathcal{D}$, isto é

$$\overline{\Theta}(\mathcal{D}) = E\left\{\arccos\left[\max_{i\in\{1,\ldots,\#_{\mathcal{D}}\}}\left(|\langle \mathcal{X}, \mathbf{g}_i\rangle|\right)\right]\right\}. \tag{2.9}$$

  Utilizaremos tanto $\Theta(D)$ como $\overline{\Theta}(\mathcal{D})$ para avaliar o desempenho de dicionários em algoritmos vorazes de decomposição de sinais.

2.3 *Frames*

O conceito de *frame* foi introduzido por Duffin e Schaeffer num estudo sobre séries de Fourier não harmônicas [38] em 1952. *Frames* só chamaram a atenção da comunidade científica em 1986, quando Daubechies, Grossman e Meyer [28] perceberam que *frames* podem ser entendidos de maneira semelhante à de expansões em bases e a relação entre *frames* e *wavelets* foi estabelecida [26, 27, 68].

Uma seqüência de elementos $\{g_k\}_{k\in\mathcal{K}}$ em um espaço $\mathbb{H}$ é um *frame* de $\mathbb{H}$ se existirem constantes $A, B > 0$ tais que [7, 17, 27, 84]

$$A\|x\|^2 \leq \sum_{k\in\mathcal{K}} |\langle x, g_k\rangle|^2 \leq B\|x\|^2, \quad \forall x \in \mathbb{H}. \tag{2.10}$$

Os números A e B são chamados de limitantes inferior e superior do *frame*, respectivamente. Diz-se que um *frame* está normalizado se $||g_k|| = 1,\ \forall k \in \mathcal{K}$ [27]. Um problema recorrente é como determinar se um conjunto de elementos g_k em $\mathbb{H}$ é um *frame* de $\mathbb{H}$, isto é, como caracterizar *frames* [17].

O *frame* inverso $\{\tilde{g}_k\}_{k\in\mathcal{K}}$ ao *frame* $\{g_k\}_{k\in\mathcal{K}}$ fornece as fórmulas de reconstrução

$$x = \sum_{k\in\mathcal{K}} \langle x, \tilde{g}_k\rangle g_k = \sum_{k\in\mathcal{K}} \langle x, g_k\rangle \tilde{g}_k. \tag{2.11}$$

É comum a definição de justeza ou aperto de um *frame* a partir de seus limitantes: quão mais próximos A e B forem entre si mais justo ou apertado é o *frame*. A justeza do *frame* indica qual a variação do escalamento das energias de sinais representados a partir de suas projeções sobre os elementos do *frame*.

2.3.1 *Frames* em Espaços de Dimensão Finita

Num espaço vetorial N-dimensional, em geral, utilizamos *frames* com um número finito de elementos K e a equação (2.10) torna-se

$$A\|\mathbf{x}\|^2 \leq \sum_{k=1}^{K} |\langle \mathbf{x}, \mathbf{g}_k \rangle|^2 \leq B\|\mathbf{x}\|^2, \qquad \mathbf{x} \in \mathbb{H}^N. \tag{2.12}$$

É comum definir o operador do *frame* [17]

$$\mathbf{S} : \mathbb{H}^N \to \mathbb{H}^N, \quad \mathbf{S}\{\mathbf{x}\} = \sum_{k=1}^{K} \langle \mathbf{x}, \mathbf{g}_k \rangle \mathbf{g}_k. \tag{2.13}$$

Em espaços vetoriais, $\mathbf{S}$ é uma matriz quadrada [17]. Se ρ_i são os auto valores de $\mathbf{S}$, então os limitantes do *frame* são dados por $A = \min_i \rho_i$, e $B = \max_i \rho_i$ [17]. Logo, se $\mathbf{S} = A\mathbf{I}_N$ ($\mathbf{I}_N$ é a matriz identidade de ordem N) teremos $A = B$ e $\mathbf{S}\mathbf{x} = A\mathbf{x}$, neste caso o *frame* é dito justo ou apertado (do inglês *tight*).

2.3.2 *Frames* de Weyl-Heisenberg

Um *frame* de Weyl-Heisenberg ou de Gabor é um *frame* de $L^2(\mathbb{R})$ obtido através de operações $\{E_{mb}T_{na}g(t)\}_{m,n\in\mathbb{Z}}$, sobre uma função fixa $g(t)$, com $a, b > 0$. As operações $E_{mb}g(t)$ e $T_{na}g(t)$ são definidas como

$$\text{Translação por } a \in \mathbb{R}, \quad T_a : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad (T_a g)(t) = g(t-a); \tag{2.14}$$

$$\text{Modulação por } b \in \mathbb{R}, \quad E_b : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad (E_b g)(t) = g(t)e^{2\pi \jmath bt}. \tag{2.15}$$

Os requisitos de a, b e $g(t)$ para que $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ seja um *frame* de $L^2(\mathbb{R})$ e a obtenção de *frames* em espaços discretos e vetoriais são discutidos na seção B.4.4.

2.3.3 Dicionários a Partir de *Frames*

Podemos usar um *frame* como dicionário em algoritmos de decomposição adaptativa de sinais. Considerando que $\{\mathbf{g}_k\}_{k\in\mathcal{K}}$ é um *frame* normalizado de $\mathbb{H}^N$, para todo $\mathbf{x} \in \mathbb{H}^N$ existe ao menos um $\mathbf{g}_k$ tal que $\langle \mathbf{x}, \mathbf{g}_k \rangle \neq 0$. Desta forma, sendo $\theta(\mathbf{x}, \mathbf{g}_k)$ o ângulo entre $\mathbf{x}$ e $\mathbf{g}_k$, para todo $\mathbf{x}$ tem-se que (ver seção B.5)

$$A\|\mathbf{x}\|^2 \leq \|\mathbf{x}\|^2 \sum_{k\in\mathcal{K}} |\cos(\theta(\mathbf{x}, \mathbf{g}_k))|^2 \leq B\|\mathbf{x}\|^2, \tag{2.16}$$

$$A \leq \sum_{k\in\mathcal{K}} |\cos(\theta(\mathbf{x}, \mathbf{g}_k))|^2 \leq B. \tag{2.17}$$

Quanto maior for o valor de $\sum_k |\cos(\theta(\mathbf{x}, \mathbf{g}_k))|^2$, maior será a "concentração" de elementos do dicionário em direções similares à de $\mathbf{x}$. Quanto mais próximos forem A e B, mais parecidas serão as "concentrações" de elementos do *frame* em todas as direções de $\mathbb{H}^N$.

2.4 Contribuições da Tese

Decomposições atômicas de sinais podem permitir representações eficientes e coerentes de sinais, reconhecimento de padrões e compressão de sinais, entre outras tarefas. A consecução dessas tarefas depende tanto do projeto do dicionário utilizado como do algoritmo que obtém a decomposição. No Capítulo 3 apresentamos representações coerentes de distúrbios em sistemas de distribuição de energia elétrica. Para obter uma representação significativa, do ponto de vista do sistema físico, um algoritmo que gera decomposições coerentes usando senóides amortecidas, relacionadas aos fenômenos físicos observáveis, é desenvolvido. A efetividade do método de decomposição apresentado é validada via testes em sinais artificiais e reais. Posteriormente mostra-se como utilizar a representação obtida para compressão de sinais com altas taxas de compressão conjugadas a boas razões sinal-ruído.

No Capítulo 4 projetam-se quantizadores para coeficientes de decomposições obtidas via *Matching Pursuit*. Para isso, os coeficientes são modelados a partir dos ângulos entre os resíduos e os átomos selecionados a cada iteração do algoritmo. Um estudo estatístico desses ângulos permite modelá-los, aproximadamente, como independentes e identicamente distribuídos a cada iteração. Esta modelagem permite estimar as estatísticas dos ângulos em iterações do *Matching Pursuit* a partir das estatísticas do ângulo na primeira iteração de decomposições de sinais provenientes de uma fonte Gaussiana sem memória. O modelo estatístico é aplicado no projeto de quantizadores de Lloyd-Max para coeficientes de decomposições obtidas via *Matching Pursuit*. As curvas operacionais de taxa×distorção obtidas para os quantizadores projetados são similares às do estado da arte em quantização de coeficientes do *Matching Pursuit*.

O Capítulo 5 define o conteúdo tempo-freqüência de *frames* a partir da distribuição de Wigner-Ville. Tal definição fornece uma nova condição suficiente para

que uma família de elementos seja um *frame*. A definição apresentada é utilizada para analisar *frames* de Weyl-Heisenberg gerados a partir de uma função protótipo simétrica. Tal análise mostra como "entrelaçar" *frames* de Weyl-Heisenberg de forma a gerar *frames* mais "justos". Este resultado é empregado para gerar dicionários a partir de *frames* de Weyl-Heisenberg "entrelaçados" que obtêm bom desempenho no *Matching Pursuit*.

Capítulo 3

Representações Coerentes e Eficientes de Sinais Elétricos

O entendimento dos distúrbios ocorridos em sistemas de distribuição de energia é uma necessidade crescente. A legislação exige a monitoração do sistema em diversos pontos. Na monitoração, utilizam-se equipamentos que digitalizam e armazenam medições, em particular das grandezas de corrente e tensão, o que se denomina oscilografia, para posterior análise por especialistas. Com a crescente quantidade de pontos monitorados, a quantidade de oscilogramas também cresce e provoca uma sobrecarga de informação. Neste contexto, surgem dois problemas:

1. Como lidar com a sobrecarga de informação?
2. Como projetar sistemas capazes de fornecer rapidamente indícios e análises sobre os distúrbios e falhas?

Este capítulo apresenta um método para a decomposição e representação de sinais elétricos que objetiva enfrentar as questões acima. O algoritmo de decomposição apresentado obtém uma representação coerente com as componentes do sinal, o que facilita a identificação do fenômeno, além de permitir a compactação dos arquivos de oscilografia com expressiva qualidade.

3.1 Modelo de Sinais Elétricos

Um bom modelo para representar sinais elétricos é uma soma de senóides (harmônicas de uma freqüência fundamental F) amortecidas [94]

$$x(t) = \sum_{m=1}^{M} \gamma_m \cos(2\pi k_m F t + \phi_m) e^{-\rho_m (t - t_{0_m})} \times [u(t - t_{0_m}) - u(t - t_{f_m})]. \quad (3.1)$$

Cada componente m, utilizada para representar $x(t)$, é determinada por uma sêxtupla (γ_m, $k_m F$, ρ_m, ϕ_m, t_{0_m}, t_{f_m}), onde γ_m é a amplitude da componente, $k_m F$ a sua freqüência (em múltiplos da freqüência fundamental F), ρ_m o amortecimento da exponencial, ϕ_m a fase da senóide, e t_{0_m} e t_{f_m} são, respectivamente, os instantes de início e fim da componente ($u(t)$ é a função degrau unitário). Um conjunto formado por M dessas sêxtuplas pode permitir uma representação compacta e acurada do sinal.

Apresentamos um algoritmo, baseado no *Matching Pursuit* (MP), que identifica as diferentes componentes senoidais amortecidas, dadas por sêxtuplas (γ_m, $k_m F$, ρ_m, ϕ_m, t_{0_m}, t_{f_m}) presentes em um sinal.

3.2 Algoritmo de Decomposição

A Figura 3.1 apresenta o diagrama em blocos do algoritmo de decomposição (nele estão indicadas as seções do Apêndice C onde cada um dos blocos é descrito). A cada iteração do algoritmo de decomposição encontra-se um átomo em conformidade com o modelo de sinal, equação (3.1). O processo de decomposição divide o problema de obtenção dos 5 parâmetros dos átomos ($k_m F$, ρ_m, ϕ_m, t_{s_m}, t_{e_m}) e seus pesos A_q de forma a obtê-los com um esforço computacional praticável. A seguir, descrevemos os objetivos e o funcionamento de cada um dos blocos.

***Matching Pursuit* com Átomos Gaussianos de Parâmetros Contínuos** A primeira etapa de cada iteração é a decomposição do resíduo corrente usando um dicionário de átomos de Gabor

$$g_\sigma(t) = \frac{K_\sigma}{\sqrt{s}} g\left(\frac{t-\tau}{s}\right) \cos(\xi t + \phi), \text{ com } g(t) = 2^{\frac{1}{4}} e^{-\pi t^2} \quad (3.2)$$

de parâmetros $\sigma = (s, \tau, \xi, \phi)$ contínuos e fase ϕ ótima. K_σ normaliza o átomo, ou seja, K_σ impõe $\|g_\sigma(t)\| = 1$. Para obter átomos de parâmetros contínuos, a cada

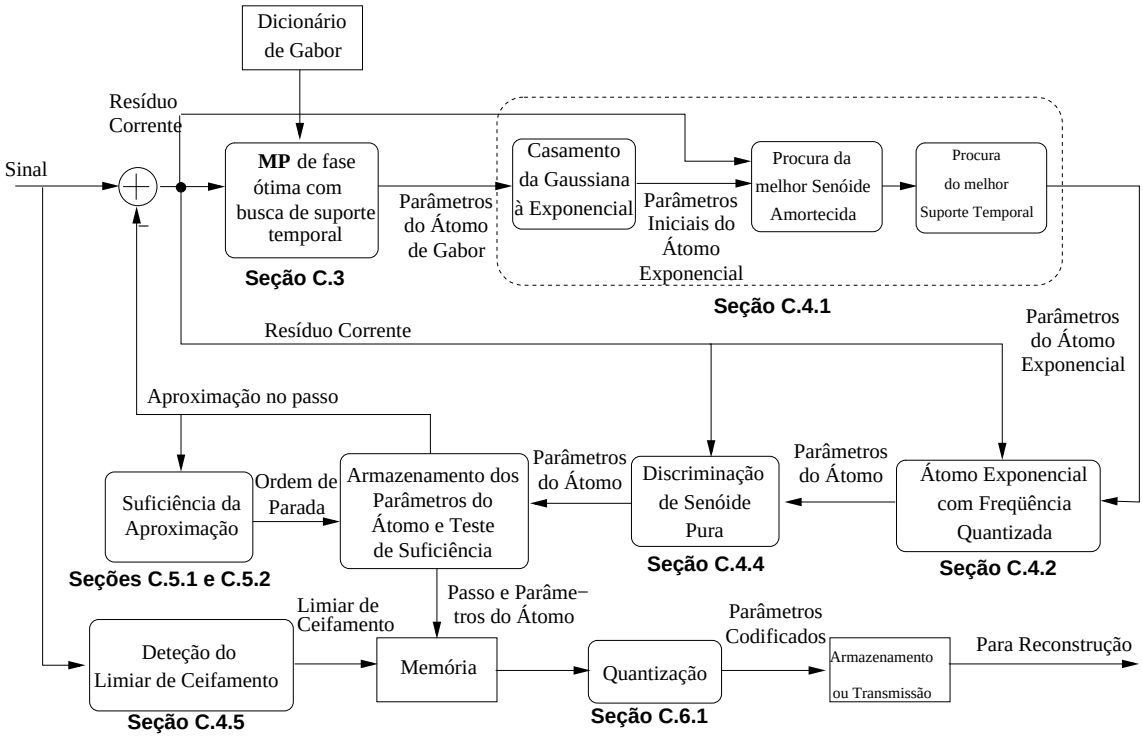


Figura 3.1: Diagrama em blocos do algoritmo de decomposição de sinais elétricos.

passo da decomposição, primeiro faz-se uma procura num dicionário de parâmetros quantizados [85]; os parâmetros quantizados são, então, utilizados como palpites iniciais na procura pelo melhor conjunto de parâmetros contínuos. Os átomos encontrados possuem fase ϕ ótima, esta é a que maximiza o produto interno do átomo com o resíduo para um dado valor do conjunto de parâmetros σ. Para encontrar a fase ótima a implementação emprega átomos complexos, e não reais como os da equação acima; a fase ótima do átomo real é obtida a partir das projeções do resíduo sobre as partes real e imaginária do átomo complexo. Os processos de obtenção de átomos com parâmetros contínuos e da respectiva fase ótima são descritos na seção C.2.

Procura do Melhor Suporte Temporal Devido à voracidade do MP, os átomos selecionados, em cada passo, podem introduzir artefatos de pré-eco e pós-eco (ruídos que aparecem antes e após a região de interesse do sinal, respectivamente). Por exemplo, a Figura 3.2 apresenta o sinal sintético *S1* (ver Apêndice C), e a Figura 3.3.(b) apresenta a aproximação obtida após 4 iterações do MP fazendo a procura do melhor átomo senoidal amortecido a cada iteração (a obtenção do átomo senoidal

será descrita a seguir). Nas Figuras 3.3.(a) e 3.3.(c) podemos observar o pré-eco (principalmente) e o pós-eco que surgem no sinal reconstruído. Para eliminar esses efeitos indesejados janelam-se os átomos via

$$g_{\sigma_l} = K_{\sigma_l} g_\sigma(t) \left[u(t - t_0) - u(t - t_f)\right], \tag{3.3}$$

onde t_0 e t_f são, respectivamente, os instantes de início e término do átomo, $\sigma_l = (\sigma, t_0, t_f)$ (com $\sigma = (s, \tau, \xi, \phi)$) e K_{σ_l} normaliza o átomo. Desenvolvemos um procedimento rápido para encontrar os valores dos parâmetros t_0 e t_f. O procedimento baseia-se numa métrica de erro na região de suporte do átomo $[t_0, t_f]$. O algoritmo de procura do melhor suporte temporal é descrito na seção C.2.2. Conforme mostram as Figuras 3.3.(c) e 3.3.(d), esse procedimento é eficaz na redução dos artefatos de pré-eco e pós-eco.

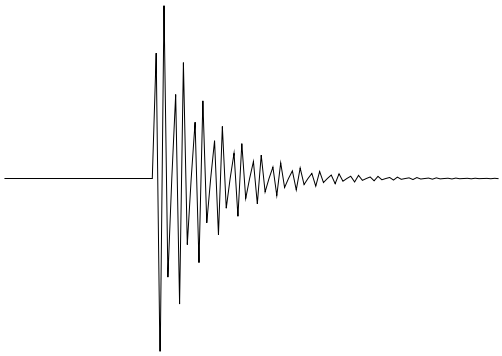
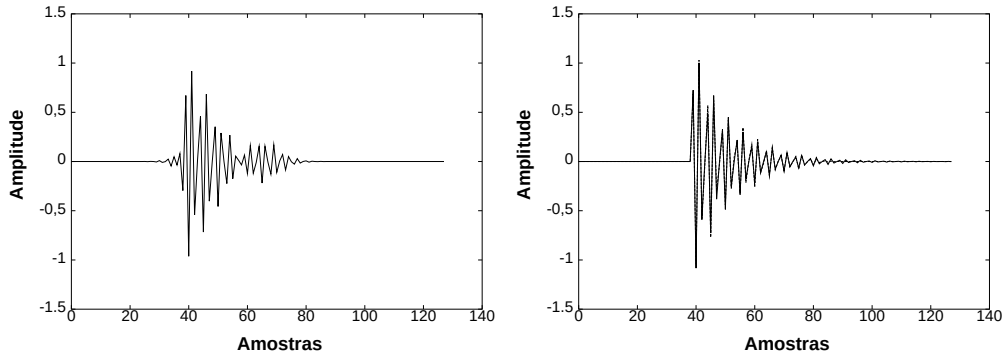


Figura 3.2: Sinal sintético *S1*.

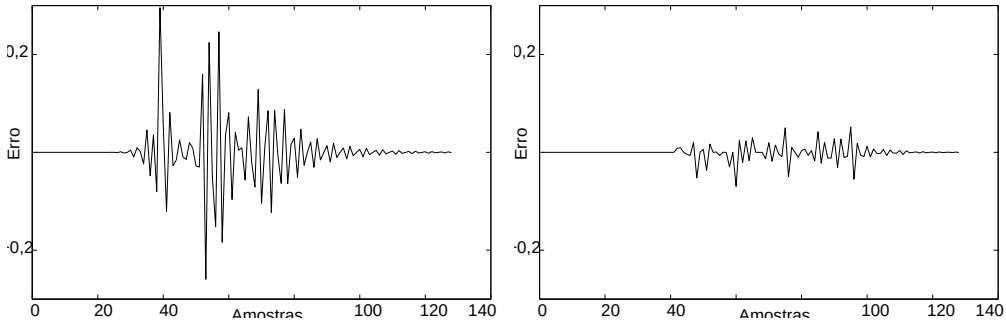
Procura por Átomos Senoidais Amortecidos O átomo de Gabor encontrado é utilizado para procurar a senóide amortecida que mais se aproxima do resíduo corrente. Os parâmetros do átomo Gaussiano são utilizados para obter uma estimativa inicial dos parâmetros do átomo senoidal amortecido

$$g(t) = K_g e^{-\rho(t-t_0)} \cos\left(\xi t + \phi\right)[u(t - t_0) - u(t - t_f)]. \tag{3.4}$$

Os parâmetros do átomo senoidal amortecido ρ, ξ, ϕ, t_0 e t_f são otimizados utilizando-se um processo similar ao utilizado para encontrar os parâmetros contínuos para o átomo de Gabor. Os detalhes deste procedimento são descritos na seção C.3.



(a) Aproximação sem procura (b) Aproximação com procura



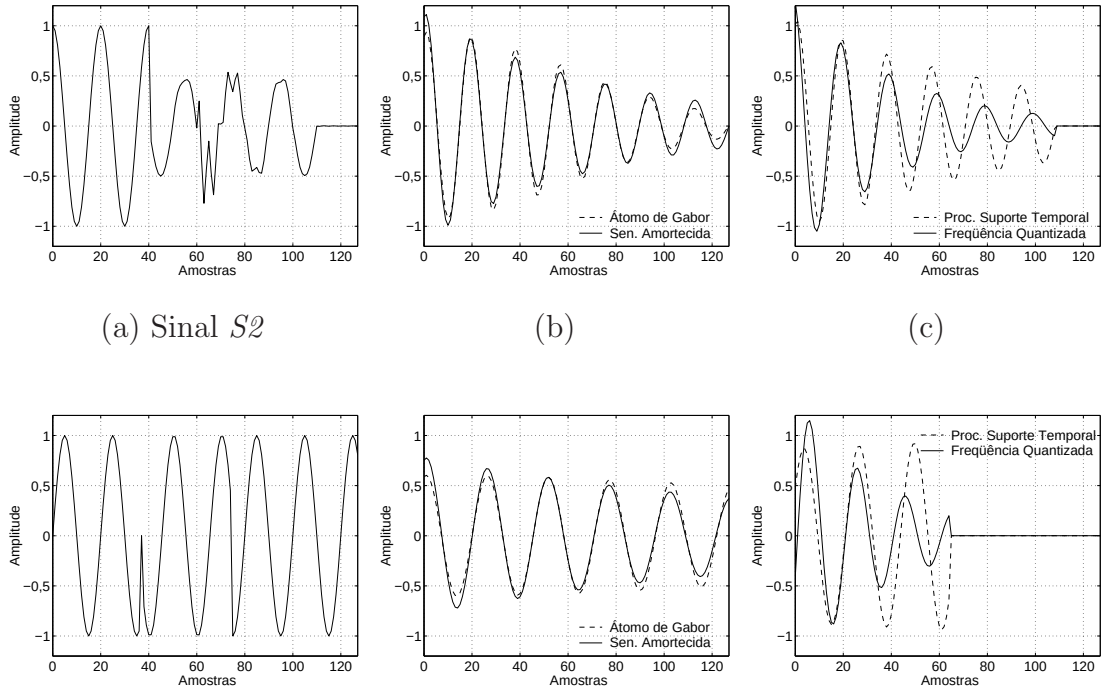
(c) Erro sem procura (d) Erro com procura

Figura 3.3: Síntese do sinal *S1* (Figura C.2) e erro de reconstrução após 4 passos de decomposição com o MP com dicionário de Gabor de parâmetros contínuos: (a) e (c) sem procura do melhor suporte temporal; (b) e (d) com procura do melhor suporte temporal.

Quantização da Freqüência O átomo senoidal amortecido não possui, em geral, uma freqüência múltipla da fundamental. De forma a obter-se uma freqüência múltipla da fundamental e átomos em concordância com o modelo da equação (3.1), a freqüência do átomo encontrado é quantizada, isto é, a freqüência encontrada para o átomo é aproximada pelo o múltiplo da freqüência fundamental mais próximo a ela. Após isso, os parâmetros dos átomos são otimizados mantendo-se a freqüência quantizada. Esse procedimento é descrito na seção C.3.2.

A Figura 3.4 ilustra o funcionamento dos procedimentos até aqui descritos para os sinais sintéticos *S2* e *S3*, veja Tabela 3.1. Nessa figura, temos: nas sub-figuras (a) e (d) os sinais originais; nas sub-figuras (b) e (e) os átomos de Gabor encontrados e as senóides amortecidas encontradas a partir dos mesmos; e nas sub-figuras (c) e (f) os átomos resultantes da procura pelo melhor suporte temporal para as senóides amortecidas, bem como os átomos resultantes após a quantização de suas

freqüências e a subseqüente otimização dos parâmetros.



(a) Sinal *S2* (b) (c)

(d) Sinal *S3* (e) (f)

Figura 3.4: Primeiro passo da decomposição dos sinais sintéticos *S2* e *S3* usando átomos senoidais amortecidos, freqüência quantizada e procura do melhor suporte temporal para o átomos. As sub-figuras (a) e (d) apresentam os sinais originais, as sub-figuras (b) e (e) os átomos de Gabor e átomos senoidais amortecidos encontrados no primeiro passo de decomposição e as sub-figuras (c) e (f) os átomos resultantes da busca da quantização da freqüência e da procura do suporte temporal.

Heurística para Obter Decomposições Coerentes Os sinais cujas decomposições foram apresentadas na Figura 3.4 foram gerados sinteticamente utilizando-se o modelo da equação (3.1). Os parâmetros utilizados na geração desses sinais são apresentados na Tabela 3.1. Comparando os átomos encontrados no primeiro passo do processo de decomposição, Figura 3.4, com os átomos utilizados na geração dos sinais, Tabela 3.1, podemos ver que, apesar de os átomos encontrados possuírem um bom casamento com os sinais (propriedade intrínseca ao MP devida à maximização do produto interno entre o resíduo e o átomo escolhido), os átomos encontrados não são exatamente coerentes com as componentes presentes nos sinais. Tais confusões do algoritmo de decomposição, ou seja, a escolha de átomos diferentes dos

Tabela 3.1: Geração dos sinais sintéticos *S2* e *S3*.

Sinal	F_s (Hz)	F (Hz)	m	γ_m	k_m	ϕ_m (^{o})	ρ_m	t_{s_m} (Seg.)	t_{e_m} (Seg.)	n_{s_m}	n_{e_m}
S2	1200	60	1	1,000	1	0	0	0	0,0333	0	40
			2	0,500	1	90	0	0,0333	0,0917	40	110
			3	0,200	6	-90	0,100	0,0500	0,1059	60	127
			4	0,050	3	-67	0	0,0417	0,0833	50	100
S3	1200	60	1	1,000	1	-90	0	0,0625	0,1059	75	127
			2	1,000	1	-90	0	0	0,0308	0	37
			3	1,000	1	135	0	0,0308	0,0625	37	75

empregados para gerar os sinais, são basicamente de dois tipos (ver seção C.3.2):

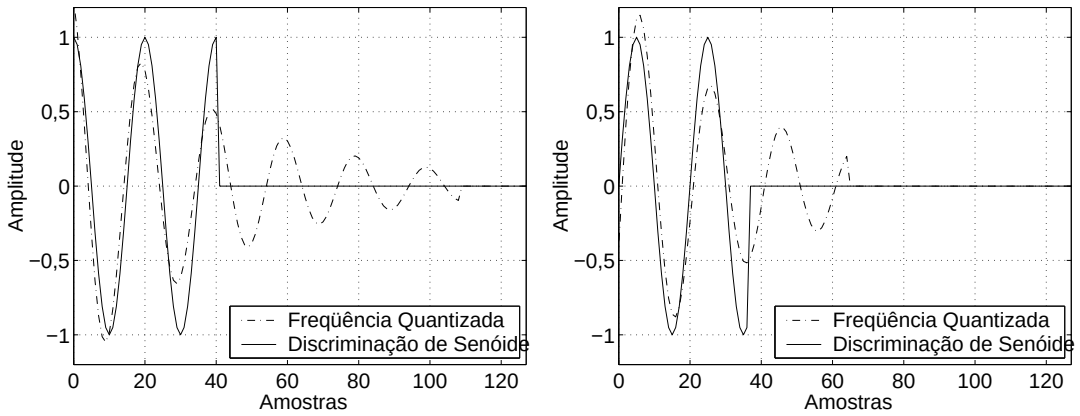
1. Sinais formados por segmentos de senóides de mesma fase e amplitudes distintas concatenados são identificados pelo algoritmo de decomposição como uma única senóide amortecida, como no caso do sinal *S2*, Figuras 3.4.(a)-(c).

2. Sinais formados por segmentos de senóides de mesma freqüência e fases distintas são aproximados pelo algoritmo de decomposição por uma única senóide, conforme se observa nas Figuras 3.4.(d)-(f) para o sinal *S3*.

De forma a corrigir as decomposições e obter decomposições coerentes, foi introduzida uma heurística dentro do laço de decomposição. Para desenvolver a heurística foram decompostos diversos sinais reais que por sua vez apresentaram problemas na identificação de componentes similares aos recém descritos para os sinais sintéticos. A heurística está representada no diagrama da Figura 3.1 pelo bloco "discriminação de senóide pura"; sua função é decidir se é conveniente utilizar uma senóide pura ao invés da senóide amortecida encontrada pelo algoritmo de decomposição. A heurística é explicada na seção C.3.4 e baseia-se no seguinte fato: mesmo se o produto interno não for o maior a cada passo de decomposição, ainda assim podemos aproximar o sinal. Entretanto, para decidir se é vantajoso utilizar uma senóide pura, a heurística utiliza uma medida de erro por amostra do resíduo resultante. Desta forma, só será utilizada uma senóide pura ao invés da amortecida quando determinadas condições do produto interno entre a senóide pura e o resíduo

e determinadas condições do erro por amostra do resíduo resultante da utilização da senóide pura forem satisfeitas quando comparadas às obtidas com o átomo senoidal amortecido. A Figura 3.5 apresenta o resultado da aplicação da heurística na primeira iteração das decomposições dos sinais *S2* e *S3*, onde podemos notar que a heurística obteve átomos coerentes com as componentes dos dois sinais.

Os átomos encontrados pelo algoritmo de decomposição com a heurística atuando dentro do laço de decomposição são coerentes com as componentes dos sinais. A Figura 3.6.(a) apresenta o sinal sintético *S4*; as componentes encontradas nos três primeiros passos de sua decomposição são apresentadas na Figura 3.6.(b); e o sinal reconstruído utilizando-se essas três componentes pode ser visto na Figura 3.6.(c). A Tabela 3.2 apresenta os parâmetros das componentes identificadas pelo algoritmo de decomposição no sinal *S4*, Figura 3.6.(a), bem como as estruturas utilizadas na geração do mesmo. A Tabela 3.2 apresenta também os parâmetros das componentes identificadas pelo algoritmo de decomposição e também os utilizados para gerar o sinal *S3*. Podemos ver, a partir desses resultados, que o algoritmo de decomposição identifica de forma satisfatória as componentes presentes nos sinais *S3* e *S4*.



Discriminação do seno no 1^o passo de decomposição do sinal *S2*

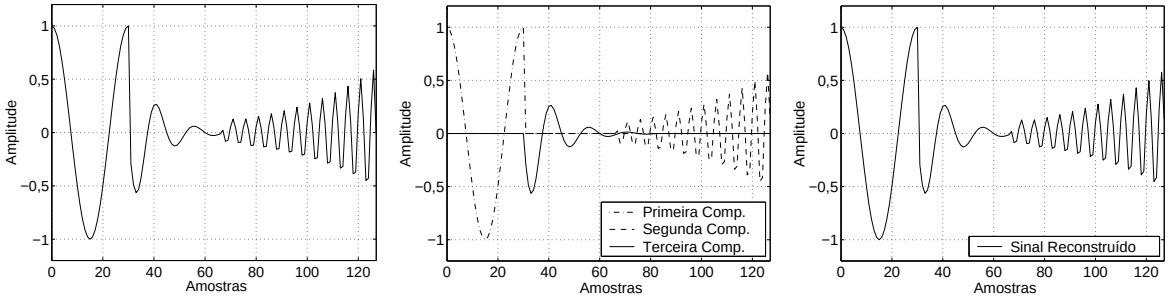
Discriminação do seno no 1^o passo de decomposição do sinal *S3*

Figura 3.5: Desempenho da heurística de discriminação de senóides no primeiro passo das decomposições dos sinais *S2* e *S3*.

Critério de Parada O processo de decomposição cessa quando o sinal não possuir informação coerente com o modelo de sinal adotado. Isso acontece quando os átomos contidos pelo dicionário não puderem fornecer uma boa aproximação do resíduo.

Tabela 3.2: Parâmetros das componentes encontradas pelo algoritmo de decomposição nos sinais *S3* e *S4*.

Sinal	Parâmetros dos átomos em *S3*					Parâmetros dos átomos em *S4*				
	$\frac{\xi}{2\pi}$ (Hz)	ϕ	ρ	n_s	n_e	$\frac{\xi}{2\pi}$ (Hz)	ϕ	ρ	n_s	n_e
Decomposição	60,00	270,00	0,00	0	36	50,00	0,00	0,00	0	30
	60,00	171,00	0,00	38	74	300,00	292,67	-0,029423	67	127
	60,00	270,00	0,00	75	127	100,00	90,00	0,100022	31	80
Geração	60,00	-90,00	0,00	0	37	50,00	0,00	0,00	0	30
	60,00	135,00	0,00	37	75	300,00	68,00	-0,03	67	127
	60,00	-90,00	0,00	75	127	100,00	90,00	0,1	30	80



(a) Sinal *S4* (b) Átomos identificados (c) Reconstrução de *S4*

Figura 3.6: Primeiras três componentes identificadas no sinal *S4*.

Para implementar um critério de parada considerando o exposto, foi utilizada a taxa de aproximação [29, 85]

$$\lambda(m) = \frac{|\langle \mathbf{r}_{\mathbf{x}}^{m-1}, \mathbf{g}_{\sigma(m)} \rangle|}{\|\mathbf{r}_{\mathbf{x}}^{m-1}\|}. \tag{3.5}$$

A taxa de aproximação mede quanto do resíduo $\mathbf{r}_{\mathbf{x}}^{m-1}$ é aproximado no passo m. A magnitude do coeficiente $\langle \mathbf{r}_{\mathbf{x}}^{m-1}, \mathbf{g}_{\sigma(m)} \rangle$ depende da energia do sinal pois os átomos possuem norma unitária. Assim, se o produto interno $\langle \mathbf{r}_{\mathbf{x}}^{m-1}, \mathbf{g}_{\sigma(m)} \rangle$ fosse utilizado como critério de parada, este dependeria da energia dos resíduos. A utilização da taxa de aproximação elimina tal dependência.

A taxa de aproximação depende do dicionário empregado. Para dicionários parametrizados ela depende fortemente da dimensão [30]. Por esse motivo, para

estabelecer um critério de parada estimamos a taxa de aproximação para uma fonte Gaussiana sem memória com coordenadas independentes e identicamente distribuídas (ruído branco Gaussiano), obtendo a estimativa λ_N – onde N é a dimensão do espaço de sinais. Devido a oscilações que ocorrem na taxa de aproximação, utilizou-se uma média móvel desta para implementar o critério de parada. O processo de decomposição pára quando a taxa de aproximação média for suficientemente próxima a λ_N, ou seja, quando a decomposição dos resíduos assemelhar-se à decomposição de ruído branco Gaussiano. A seção C.4 descreve os pormenores do critério de parada e a sua implementação. Os estudos estatísticos sobre o MP apresentados na seção 4.1 indicam que a escolha da fonte Gaussiana sem memória para esse propósito é uma boa opção.

Limiar de Ceifamento Como o sinal amostrado pode ultrapassar a faixa dos medidores, em alguns casos pode ser desejável ceifar o sinal recomposto. Antes do processo de decomposição, o maior valor absoluto entre os valores das amostras do sinal a ser decomposto é identificado. Este valor poderá ser ou não aplicado para ceifar o sinal reconstruído. O procedimento de identificação do limiar de ceifamento, bem como a sua aplicação, são descritos na seção C.3.5.

3.3 Aplicações

A decomposição apresentada foi empregada em diversas aplicações. Na seção C.4.3 mostra-se como o processo de decomposição pode ser utilizado para a redução de ruído. A redução de ruído é inerente ao critério de parada utilizado no processo de decomposição. Processos de filtragem tradicionais, seletivos em freqüência, não podem ser aplicados nos sinais de perturbações de sistemas de distribuição de energia elétrica, pois pode haver informação relevante sobre o distúrbio em toda a banda (de freqüência) operacional dos medidores.

Na seção C.3.7, apresentamos como eliminar a chamada componente "DC" que pode aparecer em algumas perturbações ocorridas na rede de distribuição de energia elétrica. A existência da componente "DC" prejudica a localização da falta na linha de transmissão por métodos fasoriais. Para eliminar a componente "DC" filtramos diretamente os átomos obtidos pelo processo de decomposição, isto é, átomos

são descartados na reconstrução do sinal em função de suas características. Faz-se, assim, uma filtragem não linear em função dos parâmetros dos átomos.

Outra aplicação das decomposições coerentes obtidas é na compressão de sinais. Para este fim, os parâmetros dos átomos são quantizados e codificados. A quantização dos parâmetros é descrita na seção C.5.1, e consiste basicamente da aplicação de um quantizador uniforme a cada um dos parâmetros, independentemente dos outros parâmetros. Este esquema de compressão difere fortemente dos tradicionais, pois a compressão é realizada quantizando-se não somente os coeficientes mas também os parâmetros do modelo de sinal. Desta forma, o sinal reconstruído é gerado via

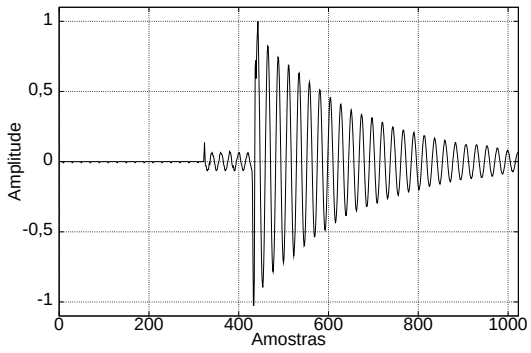
$$\hat{\mathbf{x}} = \sum_{m=1}^{M} Q[\gamma_m]\mathbf{g}_{Q[\sigma(m)]}, \tag{3.6}$$

onde $Q[\gamma_m]$ é o valor quantizado do coeficiente γ_m, $Q[\sigma(m)]$ é o valor quantizado do conjunto de parâmetros $\sigma(m)$ e M é o número de termos da representação do sinal. Isso faz com que as estruturas utilizadas na recomposição difiram das obtidas pelo algoritmo de decomposição. Embora seja simples, o processo de quantização empregado mostrou bom desempenho, permitindo verificar a capacidade do esquema de compressão proposto.

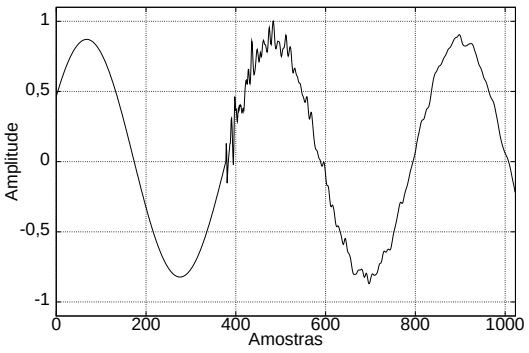
A avaliação do esquema de compressão é realizada em função da razão sinal-ruído

$$\mathrm{SNR} = 10\log_{10}\frac{\|\mathbf{x}\|^2}{\|\mathbf{x}-\hat{\mathbf{x}}\|^2}\ \mathrm{dB}, \tag{3.7}$$

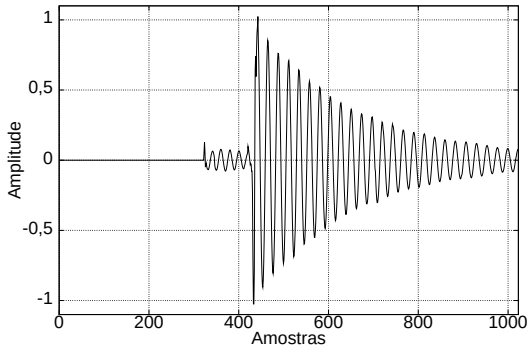
onde $\mathbf{x}$ é o sinal original, e o acrônimo SNR vem do termo inglês *Signal-to-Noise Ratio*. Apresentamos resultados para dois sinais reais *R1* e *R2* adquiridos por um sistema de monitoração, originalmente representados com 16 bits por amostra. O sinal *R1*, à esquerda na Figura 3.7, foi comprimido utilizando 1,035 bit/amostra com SNR de 28,11 dB; já o sinal *R2*, à direita na Figura 3.7, foi comprimido utilizando 0,584 bit/amostra com SNR de 31,13 dB. Podemos ver o bom desempenho do esquema de compressão a partir do pequeno erro dos sinais reconstruídos apresentados na Figura 3.7.



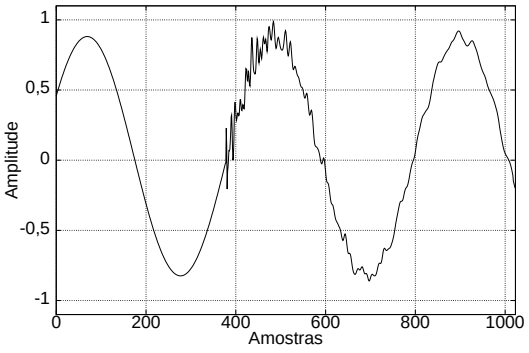
(a) *R1*



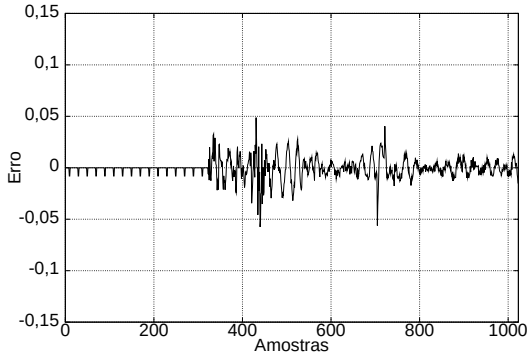
(d) *R2*



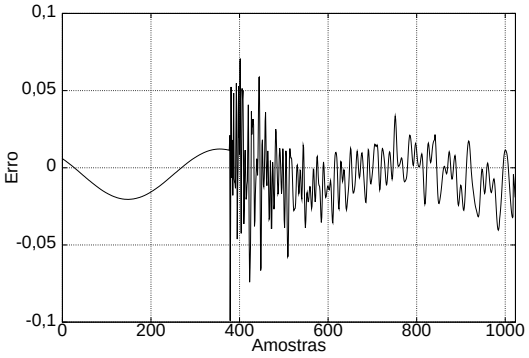
(b) Reconstrução de *R1*



(e) Reconstrução de *R2*



(c) Erro da Reconstrução *R1*



(c) Erro da Reconstrução de *R2*

Figura 3.7: Compressão do sinal *R1* através da quantização do parâmetros dos átomos a 1,035 bit/amostra com SNR de 28,11 dB, e do sinal *R2* a 0,584 bit/amostra com SNR de 31,13 dB

Capítulo 4

Quantização Lloyd-Max de Coeficientes *Matching Pursuit*

O desempenho do algoritmo de *Matching Pursuit* (MP), seção 2.1.7, em aplicações de compressão de sinais depende mormente de dois aspectos:

i) Dicionário – ele deverá incluir átomos que sejam similares às possíveis componentes dos sinais a serem comprimidos. Além disso, a cardinalidade do dicionário influencia a taxa necessária para codificar a representação do sinal.

ii) Quantização – para compressão, os coeficientes γ_n têm que ser quantizados.

Considerando uma regra de quantização $Q[\cdot]$ dos coeficientes γ_n o sinal é reconstruído através de

$$\hat{\mathbf{x}}_q = \sum_{n=1}^{M} Q[\gamma_n]\mathbf{g}_{i(n)}. \tag{4.1}$$

Para o projeto de quantizadores $Q[\cdot]$ eficientes faz-se necessário um modelo estatístico dos coeficientes MP. Davis, Mallat e Avellaneda observaram que resíduos MP possuem um comportamento caótico [30]. Ao invés de modelar diretamente os coeficientes MP, modelaremos estatisticamente os ângulos em iterações do MP. Isto é realizado a partir da observação empírica que os ângulos entre os resíduos e os átomos selecionados possuem estatísticas muito similares, a cada iteração do MP.

O ângulo entre o resíduo $\mathbf{r}_{\mathbf{x}}^{n-1}$ e o átomo selecionado $\mathbf{g}_{i(n)}$ na n-ésima iteração do MP é

$$\theta_n = \arccos\left(\frac{\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{i(n)}\rangle}{\|\mathbf{r}_{\mathbf{x}}^{n-1}\|}\right). \tag{4.2}$$

Verificou-se experimentalmente que as estatísticas de θ_n são aproximadamente independentes de n. Assim, os ângulos entre os resíduos e os átomos selecionados em iterações do MP podem ser modelados como estatisticamente independentes e identicamente distribuídos (iid). Utilizamos um modelo iid de ângulos em iterações MP para modelar estatisticamente coeficientes MP. Tal modelo permite projetar quantizadores de Lloyd-Max para coeficientes MP. Os quantizadores obtidos apresentam desempenho similar ao do estado da arte.

4.1 Ângulos em Iterações do *Matching Pursuit*

Utilizando a definição de θ_n, equação (4.2), os coeficientes produzidos pelo MP são

$$\gamma_1 = \|\mathbf{x}\| \cos(\theta_1), \tag{4.3}$$

$$\gamma_2 = \|\mathbf{x}\| \sin(\theta_1) \cos(\theta_2), \tag{4.4}$$

$$\vdots$$

$$\gamma_n = \|\mathbf{x}\| \left[\prod_{i=1}^{n-1} \sin(\theta_i)\right] \cos(\theta_n). \tag{4.5}$$

Se $\mathcal{D}$ contém os elementos $\mathbf{g}_k$ e $-\mathbf{g}_k$ para todo $k \in \{1, \ldots, \#_{\mathcal{D}}\}$, obtêm-se sempre coeficientes positivos. Consideraremos que $\mathbf{g}_k$ e $-\mathbf{g}_k$ pertencem a $\mathcal{D}$; se $\mathcal{D}$ não contém $-\mathbf{g}_k$ para algum k, então $-\mathbf{g}_k$ é incluído em $\mathcal{D}$ e $\#_{\mathcal{D}}$ (a cardinalidade do dicionário) é atualizada de acordo.

4.1.1 Estatística dos Ângulos no *Matching Pursuit*

Em [30], foi observado que resíduos MP possuem um comportamento caótico; assim, é razoável assumir que, após algumas iterações MP, os resíduos podem ter qualquer orientação. Mais precisamente, podemos assumir que a densidade de probabilidade dos resíduos depende somente de suas magnitudes. Assim, poderíamos assumir que os resíduos normalizados possuem uma função de densidade de probabilidade uniforme sobre a superfície da esfera unitária, ou seja, assumir que a orientação dos resíduos está uniformemente distribuída.

Considere uma fonte Gaussiana sem memória, independente e identicamente distribuída (iid), ou simplesmente uma fonte Gaussiana. Sendo

$\mathbf{x} = [x[1], x[2], \ldots, x[N]]^T$ uma realização dessa fonte, todos os $x[j]$ possuem a mesma distribuição normal $\mathcal{N}(0, \sigma^2)$. A densidade de probabilidade dessa fonte depende somente de $\|\mathbf{x}\|$ e sua orientação possui uma densidade de probabilidade uniforme sobre a esfera unitária. Assim, apesar da fonte Gaussiana poder não corresponder a fontes reais ou aos resíduos em iterações MP, ela possui a mesma função de densidade de probabilidade para qualquer orientação do sinal, isto é, ela não privilegia nenhuma orientação.

Podemos considerar que a orientação dos resíduos do MP possui uma densidade de probabilidade uniforme sobre a superfície da esfera tal qual a fonte Gaussiana iid sem memória. Assim, espera-se que as estatísticas dos ângulos em iterações MP ao decompor quaisquer resíduos sejam similares às obtidas para a decomposição de sinais provenientes de uma fonte Gaussiana iid sem memória.

A Figura 4.1 mostra os histogramas das variáveis aleatórias (VAs) Θ_n, correspondentes aos ângulos θ_n, resultantes da decomposição de uma fonte Gaussiana, utilizando um dicionário composto por 16 átomos aleatórios normalizados. Os átomos deste dicionário são obtidos a partir de realizações de uma fonte Gaussiana em $\mathbb{R}^4$ e posteriormente normalizados. Dicionários deste tipo, formados por $\#_{\mathcal{D}}$ sinais sorteados de uma fonte Gaussiana N-dimensional e normalizados, são aqui referidos por $DFGN(\#_{\mathcal{D}}, N)$; e logo, o dicionário empregado é denotado por $DFGN(16, 4)$. Esse dicionário tem a sua cardinalidade dobrada para que o MP obtenha sempre coeficientes positivos (incluem-se vetores de modo que $\mathbf{g}$ e $-\mathbf{g}$ pertençam ao dicionário). Os histogramas na Figura 4.1 foram obtidos a partir da decomposição MP de 50.000 sinais sorteados de uma fonte Gaussiana. Na Figura 4.1 nota-se que os histogramas de Θ_n são muito semelhantes. Assim modelaremos a fdp (função de densidade de probabilidade) dos ângulos $f_{\Theta_n}(\theta_n)$ como independentes e identicamente distribuídas a cada iteração. Os resultados apresentados em [30], acerca do comportamento caótico dos resíduos MP em determinadas condições, corroboram tal assunção.

A Figura 4.2 apresenta a média e a variância de Θ_n para vários valores de n, além das covariâncias entre ângulos MP. Nela observa-se que $\mathrm{Cov}[\Theta_i\Theta_k] = 0, \forall i \neq k$, isto é, os ângulos em diferentes iterações do MP são descorrelatados. Assim, assumir que as VAs Θ_n são independentes é razoável já que tal assunção não contradiz o comportamento observado na Figura 4.2.

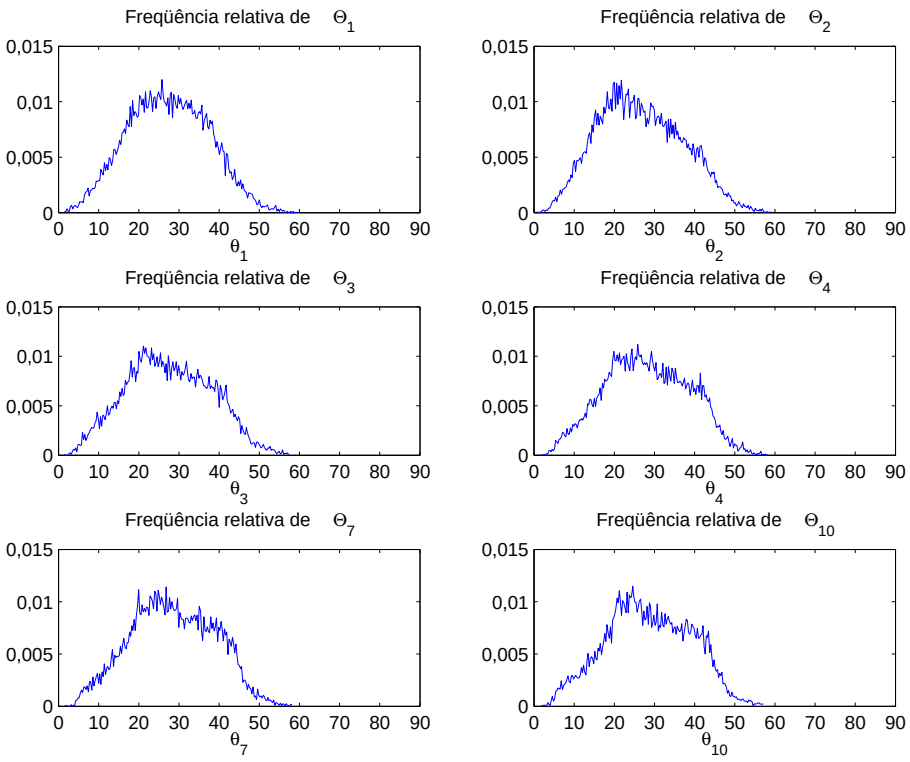


Figura 4.1: Histogramas de Θ_n para uma fonte Gaussiana em $\mathbb{R}^4$ usando um dicionário $DFGN(16, 4)$.

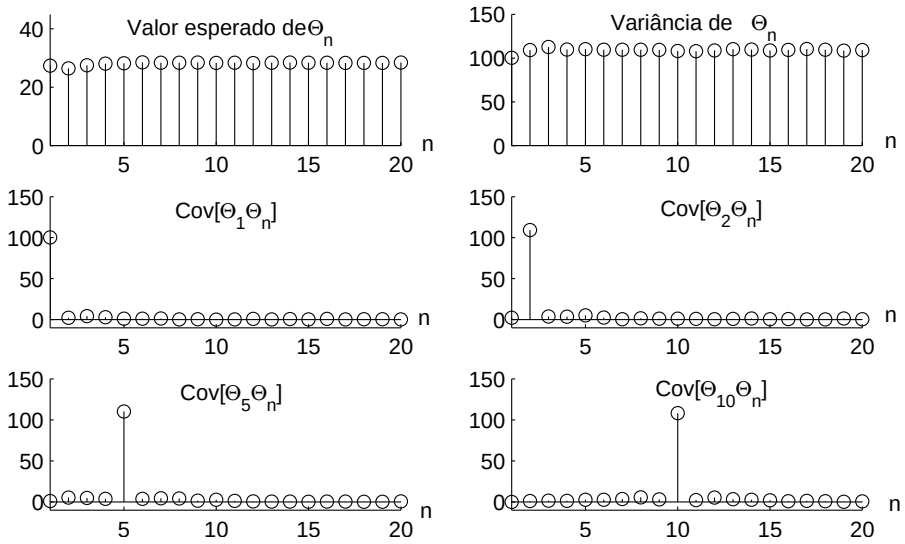


Figura 4.2: Média, variância e covariância de Θ_n para uma fonte Gaussiana em $\mathbb{R}^4$ usando o $DFGN(16, 4)$.

Os resultados apresentados até aqui foram obtidos utilizando um dicionário de cardinalidade e dimensão relativamente pequenas. Na prática, o MP é geralmente empregado em espaços de dimensão elevada com dicionários como o de Gabor [84,85]. Analisaremos, então, o comportamento dos ângulos MP para um dicionário de Gabor

de elementos reais com fases em múltiplos de $\frac{\pi}{V}$, V inteiro. Os valores das amostras $g_{j,p,v}[m]$ de cada átomo $\mathbf{g}_{j,p,v}$ do dicionário são dadas por [84]

$$g_{j,p,v}[m] = \begin{cases} \delta[m-p], & j = 0 \\ K_{(j,p,v)} f\left[\dfrac{m-p2^j}{2^j}\right] \cos\left[mk\pi 2^{1-j} + \dfrac{\pi v}{V}\right], & j \in (0, L) \\ \frac{1}{\sqrt{N}}, & j = L \end{cases}, \tag{4.6}$$

onde $f[m] = 2^{\frac{1}{4}} e^{-\pi m^2}$, $m \in [0, N-1]$ é a coordenada, $K_{(j,p,v)}$ normaliza os átomos, e $v \in [0, \ldots, V-1]$. Quanto aos parâmetros dos átomos tem-se que: j define a escala do átomo, p define a sua translação, e k a modulação do átomo. Para $L = \log_2(N)$ escalas, suas faixas de valores serão [85]: $j \in [0, L]$, $p \in [0, N2^{-j})$, $k \in [0, 2^j)$, e $v \in [0, V)$.

A Figura 4.3 apresenta $f_{\Theta_n}(\theta_n)$ (os histogramas dos ângulos), em alguns passos do MP, para um conjunto de 128.000 decomposições de sinais Gaussianos no $\mathbb{R}^{64}$ utilizando um dicionário de Gabor de 4 fases. A Figura 4.4 apresenta os mesmos gráficos mas para uma fonte cujas coordenadas possuem fdps Gama. Nota-se que para as fontes de sinais apresentadas as estatísticas dos ângulos diferem razoavelmente somente na primeira iteração do MP, sendo visualmente muito similares aos de outras iterações. Podemos notar também que: à medida que n aumenta os ângulos do MP possuem estatísticas similares mesmo para fontes de sinais distintas. Embora as estatísticas de Θ_n para $n > 1$ não sejam exatamente iguais às obtidas no primeiro passo de decomposição de realizações de uma fonte Gaussiana, elas são razoavelmente parecidas; logo $f_{\Theta_n}(\theta_n)$, para $n > 1$, pode ser razoavelmente aproximada por $f_{\Theta_1}(\theta_1)$ obtida para uma fonte Gaussiana iid sem memória. Esta assunção é razoável pois a fonte Gaussiana iid sem memória não privilegia nenhuma orientação, o que parece ser o caso também dos resíduos $\mathbf{r}_{\mathbf{x}}^{n-1}$ para $n > 1$.

4.1.2 Discussão dos Resultados

Verificamos que as estatísticas dos ângulos, após a primeira iteração do MP, podem ser consideradas invariantes à iteração. Para alguns dicionários, como o de Gabor, as fdps dos ângulos diferem ligeiramente da fdp do ângulo para uma fonte Gaussiana iid sem memória, o que sugere que os resíduos não possuem exatamente uma orientação uniformemente distribuída com a da fonte Gaussiana iid sem memó-

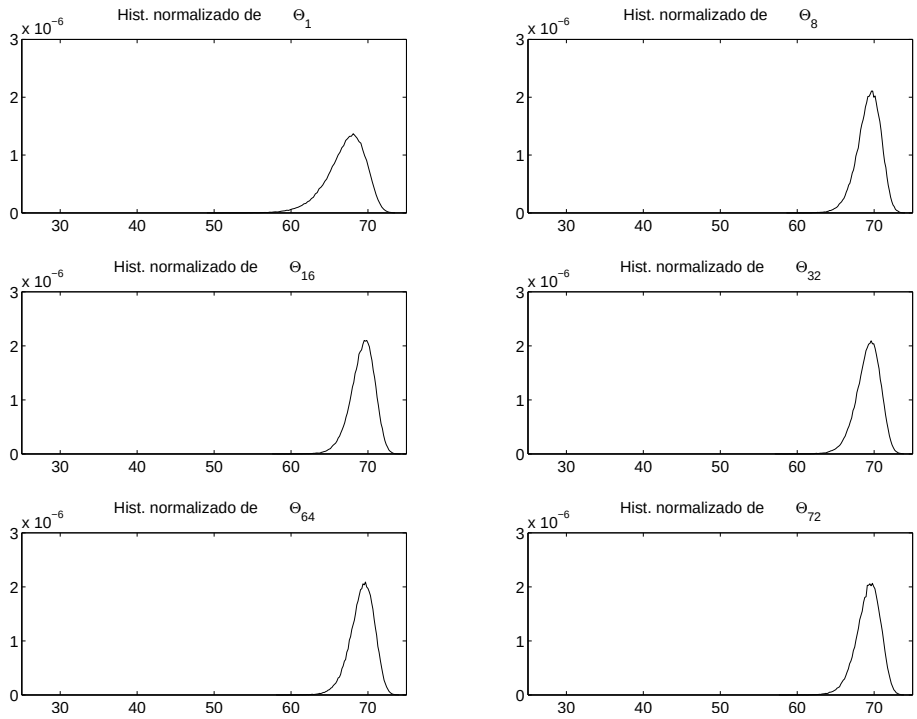


Figura 4.3: Histogramas normalizados dos ângulos MP para uma fonte Gaussiana no $\mathbb{R}^{64}$, utilizando 100 intervalos, para $n = \{1, 8, 16, 32, 64, 72\}$, para o dicionário de Gabor de 4 fases em $\mathbb{R}^{64}$.

ria. Entretanto, as estatísticas do ângulo no primeiro passo do MP para uma fonte Gaussiana iid sem memória são estimativas razoáveis das estatísticas dos ângulos em iterações do MP. Na seção D.1.4 é demonstrado que quando um dicionário contém ao menos uma base ortogonal as estatísticas dos ângulos mudam após um número de passos igual à dimensão do espaço de sinais. Isso ocorre porque para esse dicionário haverá uma probabilidade diferente de zero de se obter resíduos nulos $\mathbf{r}_{\mathbf{x}}^{n-1} = 0$, quando $n = N$ (N é a dimensão do espaço).

Utilizamos o modelo iid de ângulos MP para projetar quantizadores de Lloyd-Max para coeficientes MP. As estatísticas do ângulo na primeira iteração do MP, e portanto do primeiro coeficiente, dependem mais da fonte de sinais do que as obtidas em iterações subseqüentes. Logo, para utilizar o modelo iid de ângulos MP apropriadamente, o primeiro coeficiente (γ_1) será transmitido como informação lateral com erro desprezível.

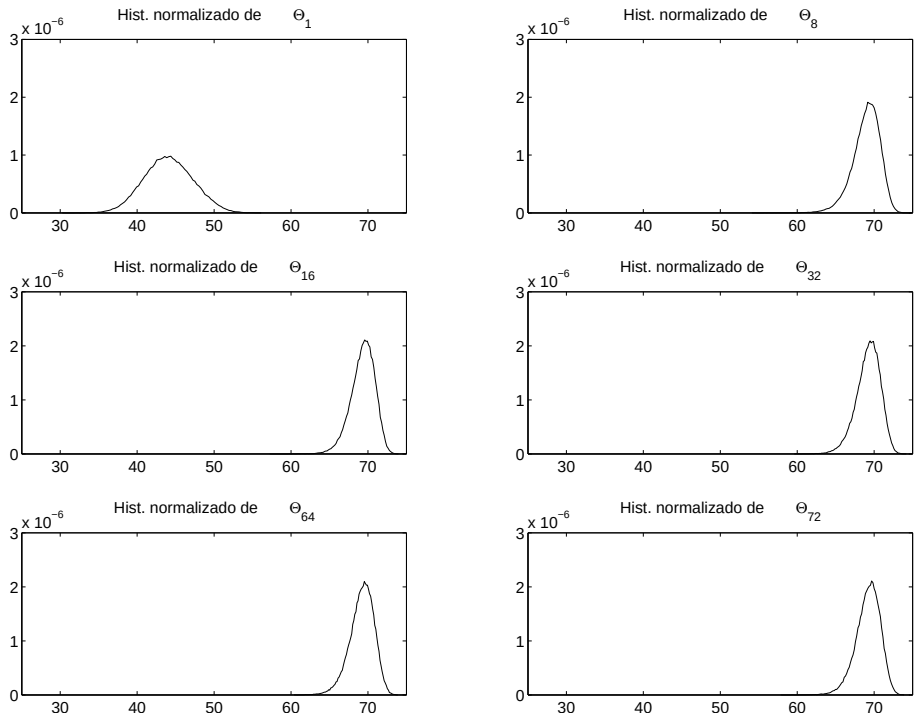


Figura 4.4: Histogramas normalizados dos ângulos MP para uma fonte de coordenadas com distribuição Gama no $\mathbb{R}^{64}$, utilizando 100 intervalos, para $n = \{1, 8, 16, 32, 64, 72\}$, para o dicionário de Gabor de 4 fases em $\mathbb{R}^{64}$.

4.2 Quantização da Decomposição em M Termos

Em aplicações de compressão, codificam-se versões quantizadas dos coeficientes. Uma estratégia comum é a quantização dos coeficientes fora do laço de decomposição, isto é, primeiro toda a decomposição é obtida e depois os coeficientes são quantizados. Tal estratégia permite realizar uma otimização taxa×distorção (RD, do inglês *rate*×*distortion*) muito simples, que consiste no teste de diferentes quantizadores e escolha daquele que satisfaz um critério RD pré-estabelecido.

Para projetar quantizadores de Lloyd-Max [73] (ótimos no sentido do erro quadrático médio) para os coeficientes utilizamos o modelo iid de ângulos MP. Os quantizadores que projetamos possuem o primeiro nível de reconstrução igual a zero; assim, coeficientes relativamente pequenos são quantizados como zero e não necessitam ser transmitidos, o que permite economizar taxa [54].

Se algum coeficiente ou índice da representação em M termos for perdido na transmissão, o decodificador deverá ser capaz de ignorar o termo perdido. Para que isto seja possível, o quantizador utilizado para um dado γ_n deverá ser independente

dos valores quantizados de outros γ_m ($m \neq n$).

4.2.1 Distorção devida à Quantização

Quando representações por M termos são quantizadas fora do laço de decomposição, cada coeficiente é substituído pela sua versão quantizada, e o sinal é reconstruído através da equação (4.1). Neste caso, podemos definir dois sinais de erro:

i) o erro relativo ao sinal original

$$\mathbf{d} = \mathbf{x} - \hat{\mathbf{x}}_q; \tag{4.7}$$

ii) o erro relativo ao M-termo

$$\mathbf{d}_M = \hat{\mathbf{x}} - \hat{\mathbf{x}}_q = \sum_{n=1}^{M} (\gamma_n - Q[\gamma_n])\mathbf{g}_{i(n)}. \tag{4.8}$$

Como $\mathbf{r}_{\mathbf{x}}^M = \mathbf{x} - \hat{\mathbf{x}}$, tem-se que $\mathbf{d} = \mathbf{d}_M + \mathbf{r}_{\mathbf{x}}^M$; logo, se $\mathbf{d}_M$ é nulo então teremos $\mathbf{d} = \mathbf{r}_{\mathbf{x}}^M$. Assim, utilizaremos $\mathbf{d}_M$ como critério de distorção no projeto de quantizadores, pois não se espera reduzir o resíduo a partir da quantização dos coeficientes.

A energia por amostra da distorção da representação por M termos quantizada é

$$d_M^2 = \frac{1}{N}\|\mathbf{d}_M\|^2 = \frac{1}{N}\sum_{n=1}^{M}\sum_{m=1}^{M}(\gamma_n - Q[\gamma_n])(\gamma_m - Q[\gamma_m])\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle, \tag{4.9}$$

onde N é o comprimento do sinal. Como os elementos de $\mathcal{D}$ possuem norma unitária, definindo-se

$$e_q(\gamma_n) = \gamma_n - Q[\gamma_n] \tag{4.10}$$

tem-se que

$$d_M^2 = \frac{1}{N}\left[\sum_{n=1}^{M} e_q^2(\gamma_n) + \sum_{n=1}^{M}\sum_{m=1,\neq n}^{M} e_q(\gamma_n)e_q(\gamma_m)\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle\right]. \tag{4.11}$$

A partir da equação (4.7) podemos obter

$$d^2 = \frac{1}{N}\|\mathbf{d}\|^2 = \frac{1}{N}\|\mathbf{x} - \hat{\mathbf{x}}_q\|^2. \tag{4.12}$$

Distorção para uma Fonte de Sinal Para uma fonte de sinal $\mathcal{X}$, podemos definir o valor esperado de d_M^2

$$E[d_M^2] = \frac{1}{N}\left\{\sum_{n=1}^{M} E\left[e_q^2(\Gamma_n)\right] + \sum_{n=1}^{M}\sum_{m=1,m\neq n}^{M} E\left[e_q(\Gamma_n)e_q(\Gamma_m)\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle\right]\right\}. \quad (4.13)$$

Acima, Γ_n é a variável aleatória (VA) correspondente ao coeficiente γ_n, com $1 \leq n \leq M$, quando se decompõem sinais originários de $\mathcal{X}$. Na equação (4.13), $E\left[e_q^2(\Gamma_n)\right]$ corresponde ao valor esperado do quadrado do erro de quantização da VA Γ_n; este erro está definido na equação (4.10).

4.2.2 Quantização de Coeficientes a partir dos Ângulos

Na equação (4.5), $\|\mathbf{x}\|$ é necessário para calcular os coeficientes MP. Se utilizamos o primeiro coeficiente γ_1, ao invés de $\|\mathbf{x}\|$, temos

$$\gamma_n = \gamma_1\delta_n,\ \delta_n = \mathrm{tg}(\theta_1)\left[\prod_{i=2}^{n-1}\mathrm{sen}(\theta_i)\right]\cos(\theta_n),\ n \geq 2. \quad (4.14)$$

Se γ_1 é conhecido, a fdp de Γ_n, para $n \geq 2$, é dada por $f_{\Gamma_n}(\gamma_n|\gamma_1) = f_{\Gamma_n}(\gamma_1\delta_n|\gamma_1) = f_{\Delta_n}(\delta_n|\gamma_1)$, onde Δ_n é a VA correspondente a δ_n, equação (4.14). A VA Δ_n depende de Θ_i, $1 < i \leq n$, logo, pode-se calcular a fdp de γ_n a partir das fdps de Θ_i, $1 < i \leq n$.

Dado γ_1, o valor esperado da distorção por amostra devida à quantização é

$$E[d_M^2|\gamma_1] = \frac{{\gamma_1}^2}{N}\left\{\sum_{n=2}^{M} E\left[e_q^2(\Delta_n)\right] + \sum_{n=2}^{M}\sum_{m=2,m\neq n}^{M} E\left[e_q(\Delta_n)e_q(\Delta_m)\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle\right]\right\}, \quad (4.15)$$

com $e_q(\Delta_n) = \Delta_n - Q[\Delta_n]$. Se as fdps das VAs Δ_n são conhecidas podemos calcular $E[d_M^2|\gamma_1]$ para qualquer regra de quantização aplicada às VAs Δ_n. Salientamos que como a quantização é aplicada a δ_n em vez de a γ_n, o valor de γ_1 precisa ser informado ao decodificador.

4.3 Quantização Lloyd-Max de Coeficientes MP

Quando o primeiro coeficiente γ_1 é conhecido, o projeto dos quantizadores dos coeficientes γ_n para $n \geq 2$ deverá objetivar a minimização da distorção da equação (4.15), projetando, assim, quantizadores para Δ_n, $n \geq 2$. Se somente um

quantizador for empregado para todos os coeficientes, deveremos projetá-lo para a VA gerada pela união das VAs Δ_n, isto é, para

$$\Delta = \cup_{n=2}^{M} \Delta_n. \tag{4.16}$$

Como iterações MP são disjuntas, a fdp de Δ é dada por

$$f_\Delta(\delta) = \frac{1}{M-1} \sum_{n=2}^{M} f_{\Delta_n}(\delta_n). \tag{4.17}$$

A distorção definida na equação (4.15) possui dois termos. O primeiro é a soma dos quadrados dos erros de quantização de Δ_n, e o segundo é a soma dos produtos internos entre os átomos do dicionário multiplicados pelos produtos dos respectivos erros de quantização. Na seção D.3.2 mostramos que a equação (4.15) pode ser simplificada para

$$E[d_M^2|\gamma_1] = \frac{M-1}{N} \int (\delta - Q[\delta])^2 f_\Delta(\delta) d\delta, \quad f_\Delta(\delta) = \frac{1}{M-1} \sum_{n=2}^{M} f_{\Delta_n}(\delta_n). \tag{4.18}$$

Esta expressão indica que o projeto de um quantizador ótimo aplicável a todos os coeficientes de representações por M termos, supondo γ_1 conhecido, equivale ao projeto de um quantizador ótimo para $\Delta = \cup_{n=2}^{M} \Delta_n$, o que é obtido por quantizadores de Lloyd-Max [73].

4.3.1 Projeto dos Quantizadores de Lloyd-Max

O projeto de quantizadores de Lloy-Max (QLM) requer $f_{\Delta_n}(\delta_n)$, para $2 \leq n \leq M$; como vimos, o modelo iid permite obter estas fdps a partir de $f_{\Theta_1}(\theta_1)$. Por sua vez $f_{\Theta_1}(\theta_1)$ é estimada aplicando-se uma iteração MP a um conjunto de sinais sorteados de uma fonte Gaussiana sem memória (sorteamos $\#_{\mathcal{D}} N^2$ sinais para este fim). As fdps $f_{\Delta_n}(\delta_n)$ estimadas são usadas para obter $f_\Delta(\delta)$ que é utilizada para projetar QLMs de b_{coef} bits ($L = 2^{b_{\text{coef}}}$ níveis de reconstrução). Os níveis de reconstrução e de decisão dos quantizadores são obtidos por um algoritmo iterativo [73]. Como a mesma lei de quantização é aplicada a todos os coeficientes e $f_\Delta(\delta)$ depende de M, cada valor de M implicará um quantizador distinto.

O projeto dos quantizadores independe de γ_1, sendo suficiente projetá-los para γ_1=1. Cópias dos quantizadores são armazenadas tanto no codificador como no decodificador. O codificador envia γ_1, o número de bits do quantizador, e o número

de termos M para o decodificador, em um cabeçalho. O valor de γ_1 é utilizado para escalar os níveis de decisão e de reconstrução do quantizador no codificador e no decodificador.

4.3.2 Desempenho da Quantização Lloyd-Max

O estado da arte em termos de quantização de coeficientes MP fora do laço de decomposição utiliza quantizadores uniformes nos quais o número de níveis de quantização e a faixa de quantização são adaptados após a quantização de cada coeficiente da representação por M termos [54]. Nos referiremos a este esquema por quantizador uniforme limitado adaptativamente (QULA). O QULA aloca um número diferente de bits para cada coeficiente. O QULA utiliza um quantizador uniforme de faixa dinâmica e número de níveis distinto para cada coeficiente; a faixa do quantizador para o l-ésimo coeficiente depende do valor quantizado do coeficiente anterior ($Q[\gamma_{l-1}]$), e o número de níveis do quantizador de cada coeficiente é obtido a partir de um procedimento de alocação de bits baseado em um multiplicador de Lagrange. Para aplicar o QULA, antes da quantização, os coeficientes são ordenados em ordem decrescente de magnitude. Como informação lateral, o QULA envia o valor do primeiro coeficiente e o número de níveis do quantizador utilizado para o segundo coeficiente.

Podem-se codificar facilmente os índices dos átomos de representações por M termos de forma que a taxa total do M-termo seja

$$R = S\left[\log_2\left(\#_{\mathcal{D}}\right)\right] + r_{\text{coef}}, \tag{4.19}$$

onde S é o número de termos que restam a ser transmitidos após a quantização (alguns coeficientes são quantizados para zero) e r_{coef} é a taxa gasta para codificar os coeficientes. Desta forma, a taxa em bits/amostra é dada por R/N, sendo N a dimensão do espaço de sinais. No QULA a diferença entre os índices de coeficientes sucessivos é utilizada para estimar a taxa do fluxo de bits das representações por M termos quantizadas [54]. Nas comparações entre o QULA e os quantizadores de Lloyd-Max (QLM), empregamos essa estratégia para estimar a taxa dos coeficientes quantizados do QULA e do QLM.

A Figura 4.5 mostra curvas RD de expansões MP quantizadas para três fontes

aleatórias distintas no $\mathbb{R}^{10}$ (uma fonte Gaussiana, uma fonte com fdp uniforme e uma com fdp Gama em cada coordenada – todas sem memória) para os QLMs e o QULA utilizando um dicionário $DFGN(128, 10)$ (gerado a partir de realizações de uma fonte Gaussiana normalizadas). Para cada taxa, o resultado apresentado corresponde à média sobre um conjunto de 100 realizações da fonte decompostas via MP e quantizadas. Para este experimento foram projetados QLMs com profundidade de bits variando de 1 a 8. A distorção utilizada é a da equação (4.12), ou seja, o erro por amostra da representação quantizada em relação ao sinal original. Podemos ver na Figura 4.5 que os dois esquemas de quantização possuem um desempenho similar para as três fontes; entretanto, os QLMs têm um desempenho ligeiramente superior a baixas taxas (abaixo de 8 bits/amostra).

A Figura 4.6 mostra as curvas RD para as três fontes de sinais do parágrafo acima em $\mathbb{R}^{64}$, para os QLMs e o QULA, para o dicionário de Gabor de 4 fases, equação (4.6). Para obter essas curvas, foram projetados QLMs com profundidade de bits entre 1 e 6 e permitiram-se no máximo 256 termos na representação. Os resultados apresentados são médias sobre 200 decomposições MP quantizadas de cada fonte.

As Figuras 4.5 e 4.6 mostram que o desempenho dos quantizadores de Lloyd-Max para coeficientes MP projetados a partir do modelo estatístico iid para ângulos MP possuem desempenho similar ao do estado da arte, independentemente do dicionário e da fonte de sinal. Nessas figuras, vê-se, ainda, que o QLM permite codificar sinais usando uma taxa menor (em bits por amostra) que a obtida utilizando o QULA.

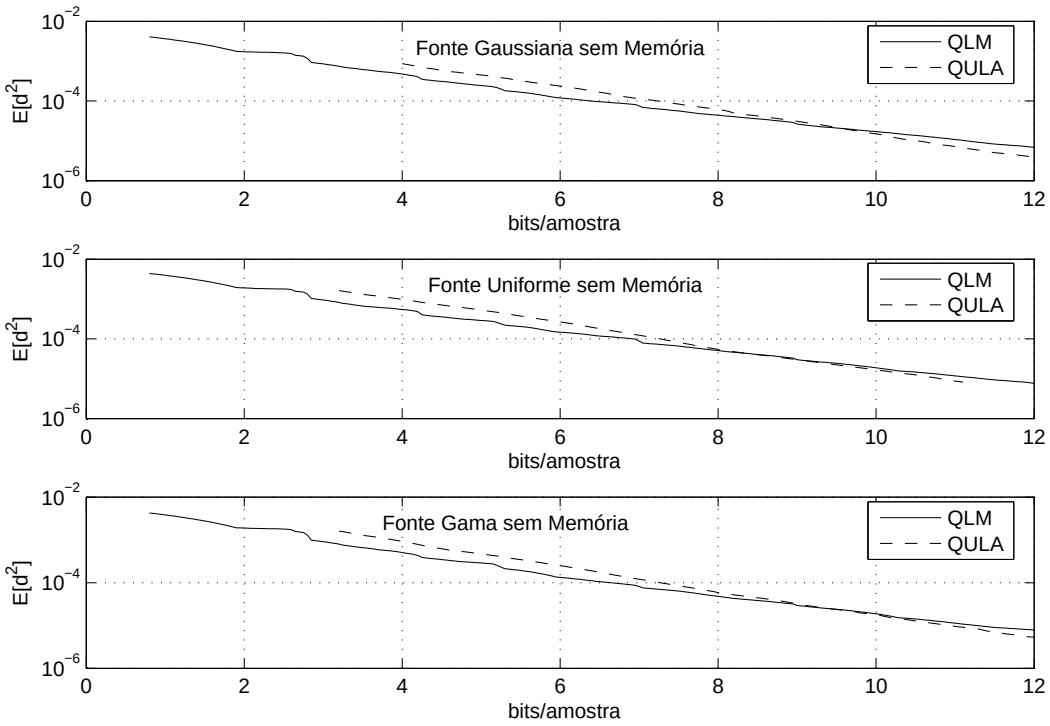


Figura 4.5: Curvas taxa×distorção para o QLM e o QULA para três fontes aleatórias utilizando um $DFGN(128, 10)$.

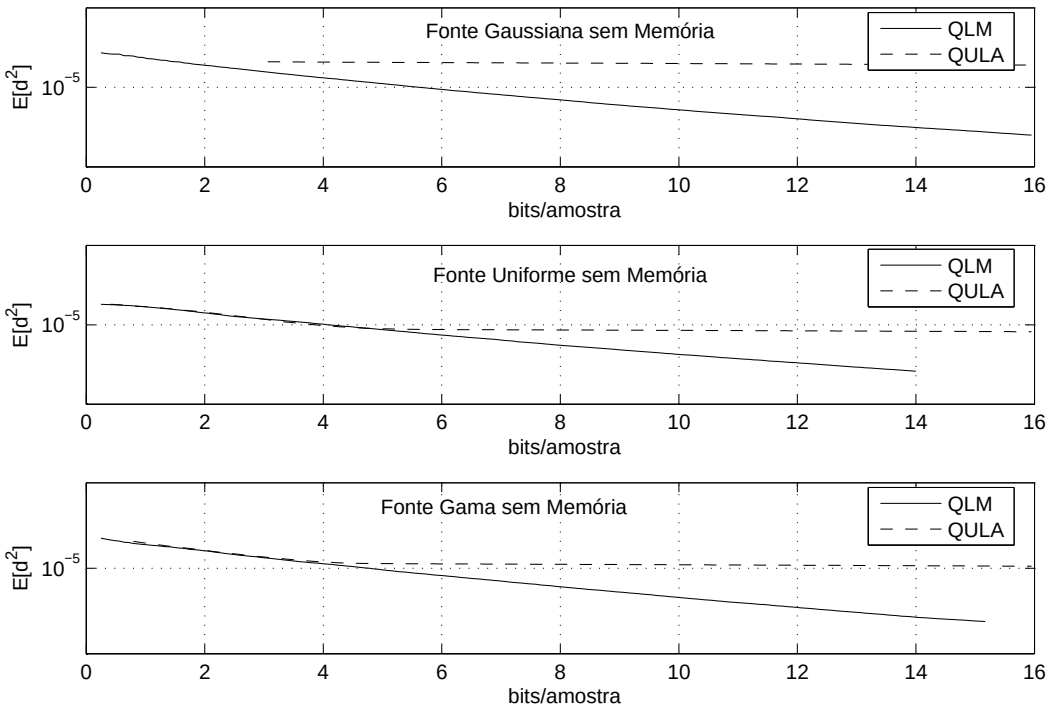


Figura 4.6: Curvas taxa×distorção para o QLM e o QULA para três fontes aleatórias utilizando um dicionário de Gabor com 4 fases em $\mathbb{R}^{64}$.

Capítulo 5

Conteúdo Tempo-Freqüência de *Frames*

Neste capítulo definimos o conteúdo tempo-freqüência de *frames* a partir da distribuição de Wigner-Ville. Mostramos que o conteúdo tempo-freqüência de um conjunto de sinais ser maior que zero é suficiente para que esse conjunto seja um *frame*. Analisamos, então, o conteúdo tempo-freqüência de *frames* de Weyl-Heisenberg, mostrando como intercalar ou entrelaçar no plano tempo-freqüência *frames* de Weyl-Heisenberg gerados a partir de uma função par, obtendo assim *frames* mais "justos". Os *frames* gerados por entrelaçamento são, então, avaliados como dicionários para algoritmos de decomposição vorazes.

5.1 Conteúdo Tempo-Freqüência de *Frames*

A distribuição de Wigner-Ville (WD) de um sinal $x(t)$ é definida como [84]

$$WD_x(t,f) = \int_{-\infty}^{+\infty} x\left(t+\frac{\tau}{2}\right)\overline{x\left(t-\frac{\tau}{2}\right)}e^{-2\pi jf\tau}d\tau, \tag{5.1}$$

onde $\overline{x(t)}$ denota a conjugação complexa de $x(t)$. A WD é provavelmente uma das ferramentas mais utilizadas e conhecidas para a análise tempo-freqüência de sinais, já que $WD_x(t,f)$ pode ser interpretada como uma medida da "densidade" da energia de $x(t)$ no tempo e na freqüência [84].

Definamos o conteúdo tempo-freqüência da expansão de um sinal $x(t)$ em um

frame $\mathcal{G} = \{g_k\}_{k\in\mathcal{K}}$ como

$$WD_{x(t)\text{ em }\mathcal{G}}(t,f) = \sum_{k\in\mathcal{K}} \langle x(t), g_k(t)\rangle\, WD_{g_k}(t,f) = \|x(t)\| \sum_{k\in\mathcal{K}} \alpha_k WD_{g_k}(t,f). \quad (5.2)$$

Se os elementos do *frame* possuem norma unitária, teremos que

$$\alpha_k = \left\langle \frac{x(t)}{\|x(t)\|}, g_k(t) \right\rangle \leq 1, \quad (5.3)$$

e portanto, neste caso,

$$WD_{x(t)\text{ em }\mathcal{G}}(t,f) \leq \|x(t)\| \sum_{k\in\mathcal{K}} WD_{g_k}(t,f). \quad (5.4)$$

Definição 5.1 *Definimos o conteúdo tempo-freqüência de um* frame $\mathcal{G} = \{g_k\}_{k\in\mathcal{K}}$ *via*

$$WD_{\mathcal{G}}(t,f) = \sum_{k\in\mathcal{K}} WD_{g_k}(t,f). \quad (5.5)$$

Teorema 5.1 *Seja* $\mathcal{G} = \{g_k\}_{k\in\mathcal{K}}$ *uma família de elementos e*

$$WD_{\mathcal{G}}(t,f) = \sum_{k\in\mathcal{K}} WD_{g_k}(t,f) \quad (5.6)$$

seu conteúdo tempo-freqüência. Uma condição suficiente para que $\mathcal{G}$ *seja um* frame *de* $L^2(\mathbb{R})$ *é*

$$0 < WD_{\mathcal{G}}(t,f) < \infty, \qquad \forall\ (t,f), \quad (5.7)$$

A prova pode ser vista na seção E.1. O Teorema 5.1 mostra que se o somatório das WDs de todos os elementos de um conjunto $\{g_k\}_{k\in\mathcal{K}}$ for limitado e positivo, então este conjunto é um *frame* de $L^2(\mathbb{R})$. Como este resultado surge diretamente das propriedades da WD e da definição de *frames*, ele é também válido em espaços de dimensão finita, ver seção E.1.

Exemplo 5.1 *Considere um* frame *do* $\mathbb{R}^N$ *formado pelos vértices "do cubo" N-dimensional. Isto é, seus elementos são* $\mathbf{g}_k = [g_k[0], \ldots, g_k[N-1]] = [\pm 1, \ldots, \pm 1]$. *Pode-se mostrar que para estes* frames $WD_{\mathcal{G}}(n,k) = N^2$, *ou seja, seu conteúdo tempo-freqüência é constante; este* frame *é justo. Podemos ver ainda que, para estes* frames, $\mathbf{S} = N^2\mathbf{I}_N$, *o que também mostra que este é um* frame *justo com limitantes* $A = B = N^2$.

No Apêndice E alguns resultados conhecidos sobre *frames* são desenvolvidos utilizando a definição de conteúdo tempo-freqüência de *frames* e o Teorema 5.1.

5.2 Conteúdo Tempo-Freqüência de *Frames* Weyl-Heisenberg

Para um *frame* de Weyl-Heisenberg (WH) $\mathcal{G} = \{E_{qb}T_{pa}g(t)\}_{p,q\in\mathbb{Z}}$ (ver seção 2.3.2), teremos

$$WD_{\mathcal{G}}(t,f) = \sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} WD_g(t-pa, f-qb). \tag{5.8}$$

Logo, para encontrar

$$\max_{(t,f)} WD_{\mathcal{G}}(t,f) \text{ e } \min_{(t,f)} WD_{\mathcal{G}}(t,f) \tag{5.9}$$

de um *frame* WH e portanto caracterizá-lo, basta analisar $WD_{\mathcal{G}}(t,f)$ na região $(t,f) \in [0,a)\times[0,b)$. Na seção E.2.3 mostra-se que para $g(t)$ par tem-se

$$\max_{t,f} WD_{\mathcal{G}}(t,f) = WD_{\mathcal{G}}(pa,qb) \text{ e} \tag{5.10}$$

$$\min_{t,f} WD_{\mathcal{G}}(t,f) = WD_{\mathcal{G}}\left(\left(p+\frac{1}{2}\right)a, \left(q+\frac{1}{2}\right)b\right). \tag{5.11}$$

Se desejarmos aumentar a capacidade de localização de sinais no plano tempo-freqüência a partir da sua projeção em um *frame* WH, podemos aumentar a densidade do *frame* no plano tempo-freqüência. Para um *frame* WH gerado a partir de uma função simétrica $g(t)$, tem-se duas opções para aumentar a sua densidade no plano tempo-freqüência: ou **i**) alterar os valores de a e/ou de b que geram o *frame*, o emprego de $a' < a$ gera um *frame* $\{E_{qb}T_{pa'}g(t)\}_{p,q\in\mathbb{Z}}$ mais denso no plano tempo-freqüência que $\{E_{qb}T_{pa}g(t)\}_{p,q\in\mathbb{Z}}$ (o mesmo ocorre se utilizamos $b' < b$); ou **ii**) empregar uma união de *frames* – uma possibilidade neste caso é unir os *frames* gerados por $\{E_{qb}T_{pa}g(t)\}_{p,q\in\mathbb{Z}}$ e por $\{E_{qb+\frac{b}{2}}T_{pa+\frac{a}{2}}g(t)\}_{p,q\in\mathbb{Z}}$. A segunda abordagem garante a obtenção de um *frame* mais apertado/justo que o original $\{E_{qb}T_{pa}g(t)\}_{p,q\in\mathbb{Z}}$. Tal estratégia é aplicável para gerar *frames* a partir de átomos exponenciais bilaterais ou átomos Gaussianos.

Ao situarmos os máximos dos novos elementos de um *frame* nos pontos de mínimo do *frame* original, minimizamos os produtos internos entre os elementos do *frame* original e do *frame* entrelaçado, quando a mesma função protótipo é utilizada para gerar ambos os *frames*. A idéia de intercalação de *frames* é apresentada na Figura 5.1. A Figura 5.1.a) mostra a localização dos máximos de $WD_{\mathcal{G}}(t,f)$ no plano tempo-freqüência, a Figura 5.1.b) o deslocamento do *frame* através da translação de

$a/2$ e da modulação por $b/2$ de todos os seus elementos, e a Figura 5.1.c) a união do *frame* original e de sua versão "entrelaçada".

Empregando a união de um *frame* WH e de sua versão "entrelaçada" sugerida acima obteremos

$$WD_{\mathcal{G}}(pa, qb) = WD_{\mathcal{G}}\left(pa + \frac{a}{2}, qb + \frac{b}{2}\right). \tag{5.12}$$

Se pudermos gerar *frames* WH $\{E_{qb}T_{pa}g(t)\}_{p,q\in\mathbb{Z}}$ e $\{E_{qb+\frac{b}{2}}T_{pa+\frac{a}{2}}g(t)\}_{p,q\in\mathbb{Z}}$ tais que sua união $\mathcal{G} = \{E_{qb}T_{pa}g(t)\}_{p,q\in\mathbb{Z}} \cup \{E_{qb+\frac{b}{2}}T_{pa+\frac{a}{2}}g(t)\}_{p,q\in\mathbb{Z}}$ possui conteúdo tempo-freqüência $WD_{\mathcal{G}}(t,f)$ com "oscilação" menor que a obtida para o *frame* original $\{E_{qb}T_{pa}g(t)\}_{p,q\in\mathbb{Z}}$, ou seja, se

$$\max_{t,f} WD_{\mathcal{G}}(t,f) - \min_{t,f} WD_{\mathcal{G}}(t,f) \tag{5.13}$$

é menor para o *frame* formado pelo entrelaçamento que para o *frame* original, então um *frame* mais apertado/justo é obtido. Como vimos nas equações (5.10) e (5.11) quando $g(t)$ é par a união descrita sempre produzirá um *frame* mais apertado/justo.

Exemplo 5.2 *A Figura 5.2 apresenta $WD_{\mathcal{G}}(t,f)$, na região $(t,f) \in [0,2a)\times[0,2b)$, para um* frame $\mathcal{G}$ *formado pela união de dois* frames *de WH entrelaçados gerados a partir de uma função-protótipo $g(t) = e^{-\beta|t|}$, com β =5, e parâmetros do* frame *WH original a =0,2 e b =2,5 (os resultados apresentados foram obtidos considerando-se 100 amostras por unidade de tempo). O* frame *WH original e sua versão entrelaçada são tais que seus conteúdos tempo-freqüência possuem máximos e mínimos iguais a 2,4564 e a 0,2274, respectivamente. Já o* frame *formado pela união dos dois possui conteúdo tempo-freqüência limitado por 1,5659 e 2,6838, sendo mais justo.*

5.3 Dicionários Parametrizados Formados por *Frames* Entrelaçados

Na seção 2.3.3 discutiu-se que *frames* podem ser usados como dicionários no *Matching Pursuit* (MP), e que quanto mais apertado ou justo for um *frame* mais parecida é a concentração de elementos do *frame* em todas as direções do espaço. O entrelaçamento de *frames* é eficiente no sentido de: (**i**) minimizar o produto interno entre os elementos acrescidos ao *frame* e os previamente existentes no *frame*; e

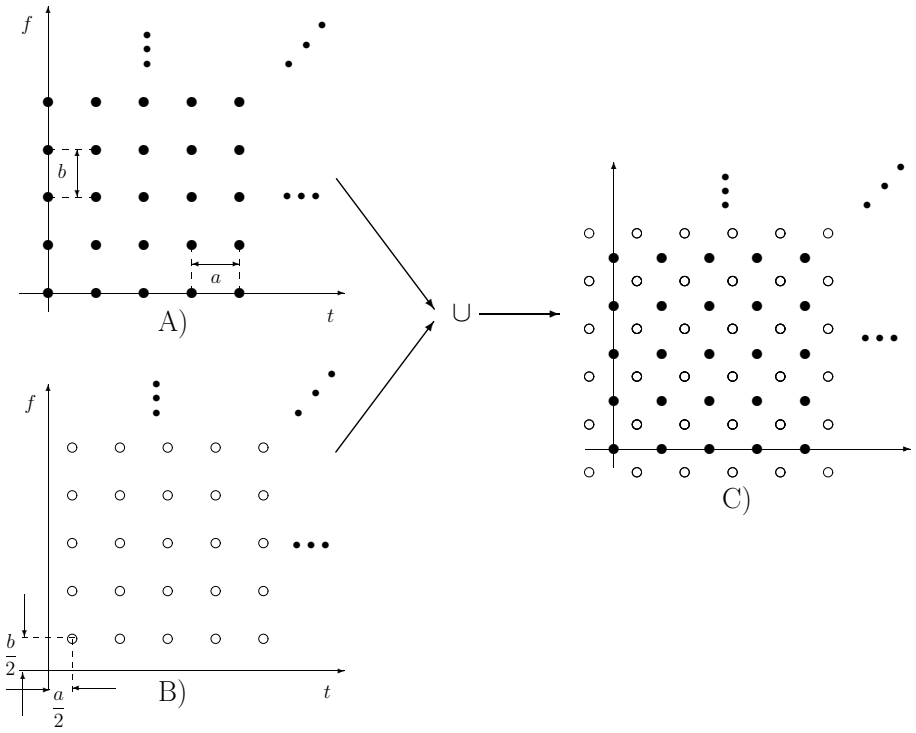


Figura 5.1: Intercalando *frames* WH.

(**ii**) gerar *frames* mais justos/apertados. O ítem (**i**) é relevante para a construção de dicionários para algoritmos de decomposição vorazes de sinais, pois nestes, não basta simplesmente aumentar $\#_{\mathcal{D}}$ (a cardinalidade do dicionário); devemos aumentá-la inserindo átomos correspondentes aos pontos com maior distância relativamente aos elementos já contidos no dicionário. Desta forma, sinais que eram mal-decompostos, que possuem um ângulo grande com o elemento mais próximo a ele no dicionário, serão melhor decompostos. Ao usar esta abordagem, o entrelaçamento de *frames* WH, os dicionários obtidos contêm elementos gerados a partir de um protótipo pré-definido, isto é, são parametrizados.

Como os dicionários obtidos pelo entrelaçamento de *frames* de WH se comportam quando utilizados no *Matching Pursuit* (MP)? Verificamos esse comportamento empiricamente através da análise dos ângulos em iterações MP obtidos por esses dicionários, seções 2.2.2 e 4.1.1. Para este experimento utilizaremos a decomposição de sinais provenientes de uma fonte Gaussiana sem memória em $\mathbb{R}^{64}$ utilizando 5 dicionários distintos que contêm elementos gerados a partir de *frames* WH de um sinal protótipo Gaussiano, ver seção 4.1.1 em escala um ($s = 1$) e com variância unitária ($\sigma^2 = 1$). Os elementos do dicionário são as partes reais (átomo Gaussiano de fase

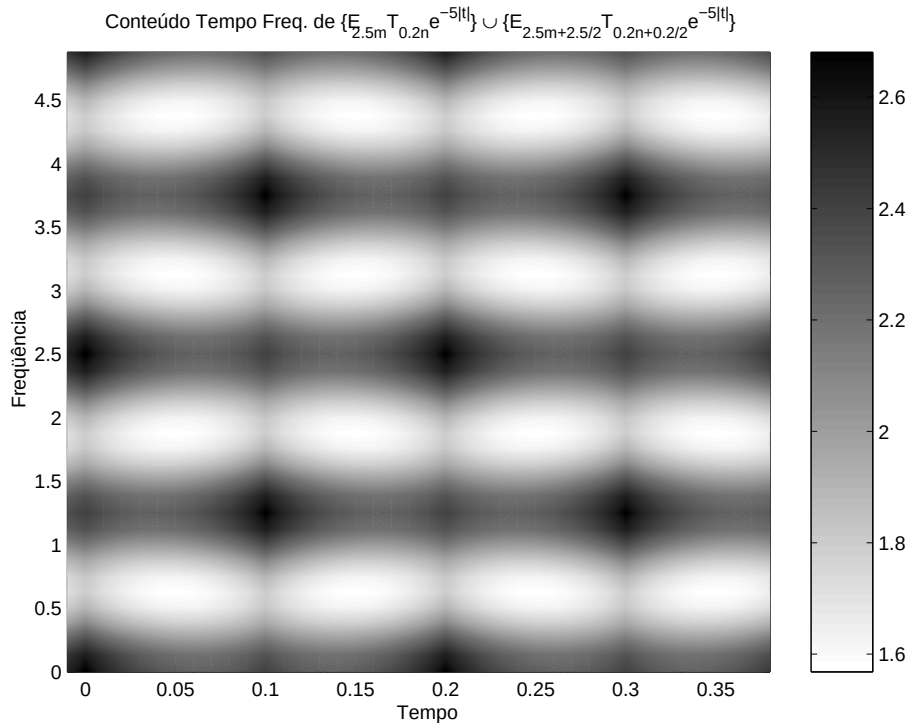


Figura 5.2: $WD_{\mathcal{G}}(t,f)$ com $g(t) = e^{-5|t|}$, a =.2 e b =2.5, ab=1/2, ver exemplo 5.2, $(t,f) \in [0,2a) \times [0,2b)$.

zero) e imaginárias (átomo Gaussiano de fase $\pi/2$) normalizadas de $E_{q\frac{1}{Q}}T_{p\frac{N}{P}}\mathbf{g}[n]$.

Para gerar o dicionário $\mathcal{A}$, P e Q são escolhidos de forma a fornecerem 16 pontos tanto no eixo do tempo como no da freqüência (os átomos são definidos no $\mathbb{R}^{64}$). O dicionário $\mathcal{C}$ é composto da união de $\mathcal{A}$ com sua versão deslocada no tempo por $N/2P$ e na freqüência por $1/2Q$, ou seja a união de $\mathcal{A}$ com o *frame* "entrelaçado" a $\mathcal{A}$. O dicionário $\mathcal{F}$ é gerado de forma semelhante à de $\mathcal{A}$, mas com parâmetros de deslocamento no tempo e na freqüência divididos por dois; assim, $\mathcal{F}$ tem cardinalidade quatro vezes maior que a de $\mathcal{A}$. O dicionário $\mathcal{G}$ é gerado a partir da união de $\mathcal{F}$ e sua versão "entrelaçada", como $\mathcal{C}$ em relação a $\mathcal{A}$. O dicionário $\mathcal{M}$ corresponde ao caso de amostragem máxima dos parâmetros de deslocamento no tempo e na freqüência, provendo, assim, 64 pontos em cada eixo. As cardinalidades dos dicionários são, então, $\#_{\mathcal{A}} = 512$, $\#_{\mathcal{C}} = 1024$, $\#_{\mathcal{F}} = 2048$, $\#_{\mathcal{G}} = 4096$, $\#_{\mathcal{M}} = 8192$.

A Figura 5.3 apresenta os histogramas (utilizando 250 intervalos) do ângulo

$$\Theta = \arccos\left\{\max_{\mathbf{g}_i \in \text{ Dicionário}}\left(\frac{\langle \mathbf{x}, \mathbf{g}_i\rangle}{\|\mathbf{x}\|}\right)\right\} \tag{5.14}$$

para sinais $\mathbf{x}$ provenientes de uma fonte Gaussiana iid sem memória (ver capítulo

4). Esses resultados foram obtidos a partir de 10.000 decomposições de 40 passos de realizações da fonte Gaussiana – vimos no Capítulo 4 que para uma fonte Gaussiana podemos assumir que os ângulos em iterações MP são aproximadamente estacionários. Conforme vimos na seção 2.2.2, $\overline{\Theta}(\mathcal{D})$ (o valor médio de Θ) e $\Theta(\mathcal{D})$ (o valor máximo de Θ) são relevantes no desempenho de dicionários no MP. Os histogramas de Θ para os dicionários $\mathcal{C}$ e $\mathcal{G}$ mostram que para ambos o entrelaçamento é capaz de melhorar razoavelmente a distribuição do ângulo e obter $\overline{\Theta}(\mathcal{D})$ e $\Theta(\mathcal{D})$ menores do que os obtidos para, respectivamente, $\mathcal{A}$ e $\mathcal{F}$. Vê-se ainda que $\mathcal{M}$ não obtém grande melhora em relação a $\mathcal{G}$. Observe que $\mathcal{G}$ é um subconjunto de $\mathcal{M}$ e $\mathcal{G}$ contém metade dos elementos presentes em $\mathcal{M}$ (assim como ocorre com $\mathcal{C}$ em relação a $\mathcal{F}$).

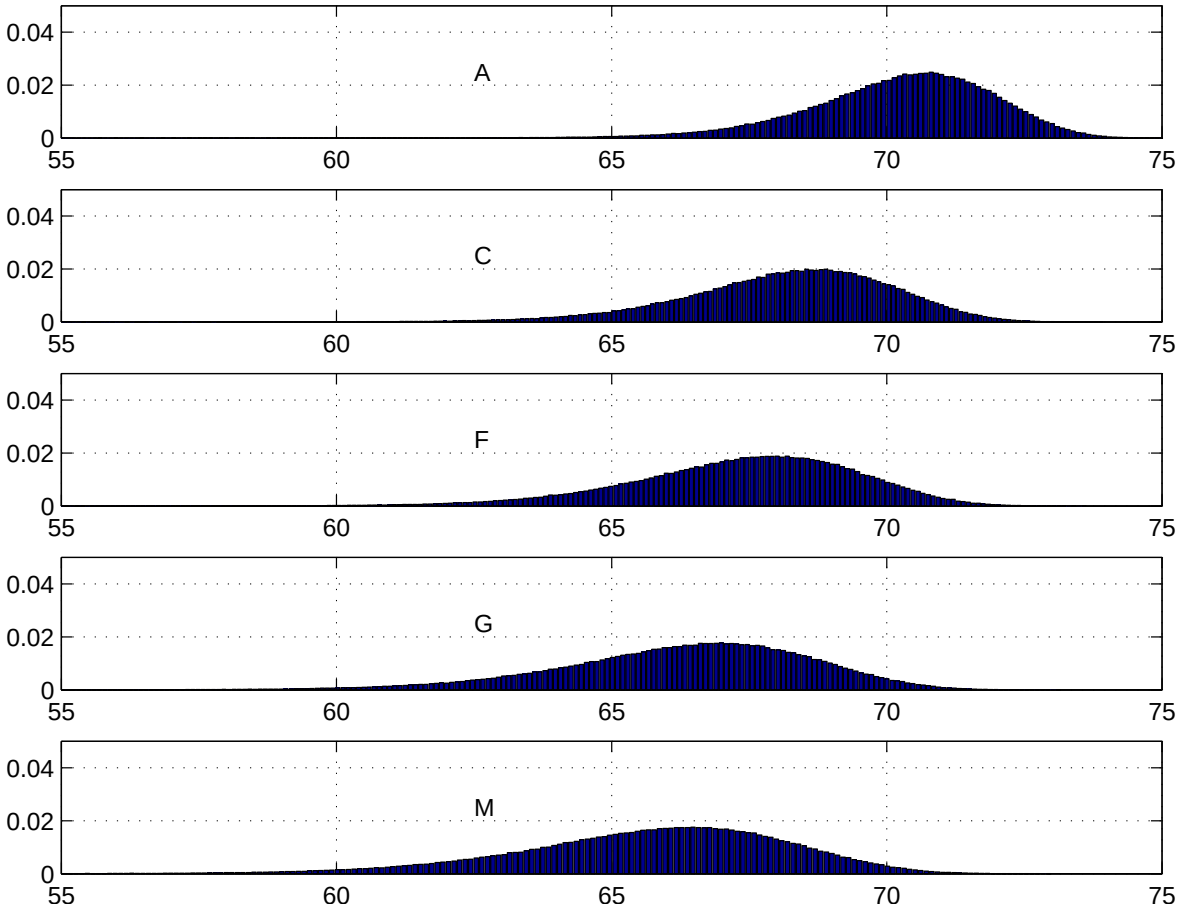


Figura 5.3: Histogramas dos ângulos entre os resíduos e os átomos selecionados no MP para aproximá-los para os dicionários $\mathcal{A}$, $\mathcal{C}$, $\mathcal{F}$, $\mathcal{G}$ e $\mathcal{M}$ em $\mathbb{R}^{64}$.

Para melhor avaliar os desempenhos dos dicionários devemos considerar também suas cardinalidades. Para esse fim realizamos um estudo taxa×distorção. Na análise taxa×distorção ignoraremos a quantização dos coeficientes e consideraremos que não são gastos bits para enviá-los. Tal suposição permite um bom estudo das

tendências e do comportamento das decomposições obtidas para dicionários distintos. Assim, a taxa gasta na codificação das decomposições obtidas via MP cresce linearmente com o número de termos da decomposição M, e podemos encontrar facilmente um código para codificar as decomposições com taxa

$$R = M\lceil \log_2(\#_{\mathcal{D}}) \rceil. \tag{5.15}$$

A Figura 5.4 apresenta, à esquerda, a média da norma do erro em função do passo M e, à direita, a média da norma do erro em função de $R = M\lceil \log_2(\#_{\mathcal{D}}) \rceil$. Os dicionários gerados por entrelaçamento no plano tempo-freqüência sempre possuem erros menores em função do passo, relativamente, como se vê nas curvas erro×passo para $\mathcal{A}$ e $\mathcal{C}$ e para $\mathcal{F}$ e $\mathcal{G}$. Um comportamento similar pode ser observado paras curvas erro×taxa. Entretanto, no último caso o ganho em utilizar $\mathcal{G}$ ao invés de $\mathcal{F}$ não é tão significativo como o de utilizar $\mathcal{C}$ no lugar de $\mathcal{A}$. Apesar de $\mathcal{G}$ não obter um ganho significativo no compromisso taxa×distorção de $\mathcal{F}$, como $\mathcal{C}$ consegue em relação a $\mathcal{A}$, $\mathcal{G}$ não tem desempenho taxa×distorção pior que o de $\mathcal{F}$. Assim, vemos que o entrelaçamento no plano tempo-freqüência é efetivo para gerar dicionários a partir de *frames* WH.

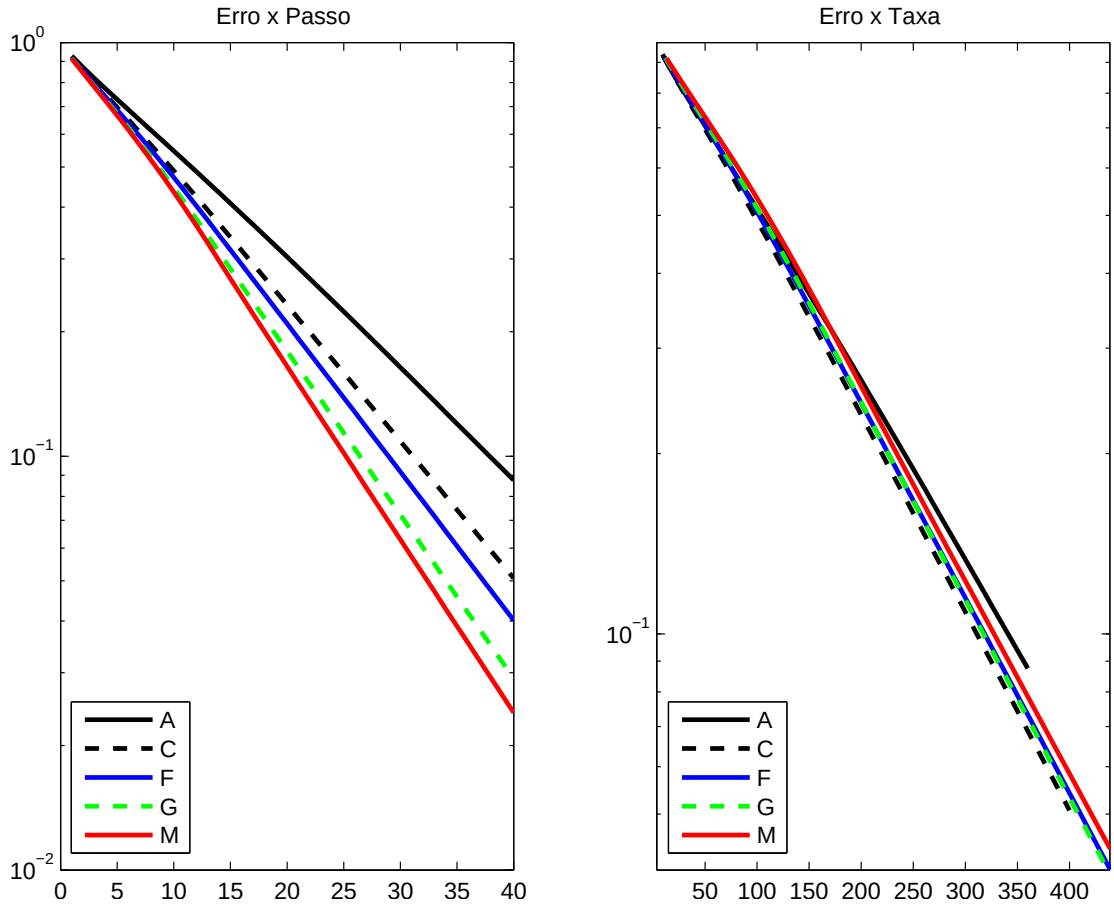


Figura 5.4: Erro no passo (esquerda) e erro×taxa (direita) para os dicionários $\mathcal{A}$, $\mathcal{C}$, $\mathcal{F}$, $\mathcal{G}$ e $\mathcal{M}$ em $\mathbb{R}^{64}$.

Capítulo 6

Conclusão

Nesta tese tratou-se de alguns aspectos relativos a representações de sinais usando dicionários redundantes. No Capítulo 2 apresentou-se uma rápida revisão teórica do tema. No Capítulo 3 vimos uma aplicação prática dessa estratégia na decomposição de sinais de distúrbios elétricos. A seguir projetamos quantizadores para coeficientes de representações adaptativas de sinais no Capítulo 4. Já no Capítulo 5 tratou-se de *frames* e seu conteúdo tempo-freqüência e vimos como gerar dicionários efetivos a partir de *frames*.

O esquema de decomposição de sinais apresentado no Capítulo 3 obtém representações coerentes utilizando átomos parametrizados e robustas a ruído, e além disso permite obter boas taxas de compressão de sinais aliadas a boa qualidade dos sinais comprimidos. O algoritmo, baseado no *Matching Pursuit*, decompõe o sinal adaptativamente utilizando senóides amortecidas. O sinal é representado por uma seqüência de coeficientes e de conjuntos de parâmetros que definem as estruturas identificadas no sinal. Procedimentos computacionalmente realizáveis e efetivos foram apresentados para encontrar os valores dos parâmetros.

Algoritmos de decomposição vorazes, como o *Matching Pursuit*, normalmente desviam-se da obtenção de representações fisicamente representativas. Para evitar tal desvio, foi dada "inteligência" ao algoritmo, de forma a permitir-lhe obter decomposições coerentes de sinais de distúrbios elétricos. Isso foi conseguido com a introdução de uma heurística dentro do laço de decomposição, a qual instrui o algoritmo a escolher um átomo apropriado. O critério de parada empregado no algoritmo de decomposição obtém automaticamente a quantidade de termos necessários às re-

presentações dos sinais.

As decomposições obtidas foram comprimidas quantizando não somente os coeficientes, mas também os parâmetros do modelo de sinal, isto é, o sinal comprimido reconstruído utiliza átomos diferentes dos obtidos pelo algoritmo de decomposição. Apesar de sua simplicidade, o processo de quantização aqui empregado mostrou um desempenho satisfatório.

Como na prática sinais de distúrbios elétricos são analisados por especialistas, como continuação do trabalho desenvolvido devemos submeter os sinais reconstruídos a processos de análises subjetivos. Assim poderemos avaliar melhor o desempenho do sistema de compressão proposto.

O projeto de quantizadores de Lloyd-Max para coeficientes MP foi apresentado no Capítulo 4. Esses quantizadores são projetados a partir das estatísticas dos ângulos entre resíduos e átomos selecionados em iterações do *Matching Pursuit*.

A análise empírica do ângulo entre o resíduo e o átomo selecionado em iterações do *Matching Pursuit* indicou que estes ângulos podem ser modelados como estatisticamente independentes e identicamente distribuídos. Assim, aproximadamente, tais ângulos podem ser considerados como possuidores de estatísticas invariantes à iteração do algoritmo. Desta forma, as estatísticas de ângulos em iterações do *Matching Pursuit* podem ser obtidas a partir das estatísticas do ângulo no primeiro passo do *Matching Pursuit* para fontes de sinais Gaussianas.

O modelo estatístico independente do passo de decomposição e identicamente distribuído para os ângulos no *Matching Pursuit* foi empregado para projetar quantizadores de Lloyd-Max para coeficientes de decomposições obtidas via *Matching Pursuit*. O projeto desses quantizadores só necessita de uma estimativa da função de densidade de probabilidade do ângulo no primeiro passo do *Matching Pursuit*. Os quantizadores de Lloyd-Max obtidos apresentaram desempenho similar ao do estado da arte em termos de quantização fora do laço de coeficientes *Matching Pursuit*. Além disso, os quantizadores de Lloyd-Max possuem intrinsecamente resiliência a erros, que resulta da aplicação da mesma regra de quantização sobre todos os coeficientes da representação. Vimos que quantizadores de Lloyd-Max obtidos para fontes Gaussianas podem ser utilizados também para fontes que não sejam Gaussianas.

Os quantizadores projetados para coeficientes de decomposições vorazes obti-

veram bom desempenho. Resta-nos testá-los embutidos em sistemas de codificação de sinais, como, por exemplo, áudio e vídeo, baseados em decomposições vorazes.

O conteúdo tempo-freqüência de um *frame* foi apresentado no Capítulo 5. Ele é definido como a soma das distribuições de Wigner-Ville dos elementos que compõem o *frame*. A análise do conteúdo tempo-freqüência de um *frame* permite caracterizá-lo. Apresentamos uma nova condição suficiente que diz que: se a soma das distribuições de Wigner-Ville de um conjunto de elementos for maior que zero, então o conjunto é um *frame*.

Assim como encontramos uma forma de caracterizar *frames* no domínio tempo-freqüência, conjecturamos ser possível fazer o mesmo no domínio escala-temporal, o que seria útil na caracterização de *frames* de *wavelets*.

A análise do conteúdo tempo-freqüência de *frames* de Weyl-Heisenberg gerados a partir de funções pares mostrou como definir os elementos de um segundo *frame* de Weyl-Heisenberg $\mathcal{H}$ em relação aos elementos de um *frame* de Weyl-Heisenberg $\mathcal{G}$ de forma a construir um *frame* "tipo" Weyl-Heisenberg $\mathcal{H} \cup \mathcal{G}$ mais apertado. O *frame* mais apertado é obtido através do entrelaçamento dos dois *frames* $\mathcal{H}$ e $\mathcal{G}$ no plano tempo-freqüência. Isso equivale à utilização de um reticulado losangular para gerar o *frame*, ao invés do retangular, no plano tempo-freqüência.

O entrelaçamento de *frames* de Weyl-Heisenberg foi utilizado para construir dicionários. Os dicionários foram avaliados usando as estatísticas dos ângulos produzidos pelos dicionários quando utilizados no *Matching Pursuit* para fontes Gaussianas. Os resultados experimentais indicam que para dicionários de mesma cardinalidade, aqueles formados por elementos de *frames* de Weyl-Heisenberg entrelaçados possuem um desempenho melhor que os dicionários constituídos pelos elementos de um único *frame* de Weyl-Heisenberg.

Apêndice A

Introduction

A.1 Atomic Signal Representations

A legacy of Greek thinking is the idea of atoms – indivisible units of matter. In science, in many different manners, we use the approach of dividing the natural world into pieces or parts. The methodology of partitioning the real-world, into small (sometimes not so small) pieces and parts or atoms, then selecting a subset of these parts (sometimes disregarding others), and concentrating on this subset to construct an approximation of the real-world has being largely employed. The impact of this methodology in science evolution is enormous. An example is the prime and original idea presented by Erwin Schrödinger in 1943, that goes as hereditary characteristics should be stored by combinations of a limited set of crystals. That was later realized by the discovery of the DNA structure. Atomic signal representations resemble the aforementioned methodology, since they focus on extracting relevant signal parts to construct a good signal approximation.

Atomic representation of signals is the idea encompassing the expression of signals using pre-defined waveforms. That is, signal $\mathbf{x}$ shall be represented by a linear combination of few waveforms selected from a pre-defined set, the dictionary $\mathcal{D}$. The signals $\mathbf{g}_i$, that are included in $\mathcal{D}$, are the waveforms allowed to be used in the linear combination that expresses $\mathbf{x}$, and are also called atoms or elements. Atomic signal decomposition algorithms select a subset of M dictionary elements

$\mathbf{g}_{i(m)}$ that approximate $\mathbf{x}$ as

$$\mathbf{x} \approx \hat{\mathbf{x}} = \sum_{m=1}^{M} \gamma_m \mathbf{g}_{i(m)}, \quad \mathbf{g}_{i(m)} \in \mathcal{D}. \tag{A.1}$$

Note that $i(m)$ indexes $\mathbf{g}_{i(m)}$ in the dictionary and γ_m is the weight of $\mathbf{g}_{i(m)}$. Each atom used to express $\mathbf{x}$ can be understood as an intrinsic signal feature, allowing for compact signal coding and estimation. A dictionary includes the words of a language. Languages are redundant as they allow us to express an idea in many different ways. An overcomplete dictionary is a redundant collection of elements, that is, it contains more elements than necessary to span the signal space. Redundancy allows us to select one among various signal representations.

The atoms in the periodic table, compose all the known molecules. However, the properties of chemical composites heavily depend on how their elements are interconnected. Although the elements in the benzine molecule were known, it was not until Kekule dreamed about a snake biting its own tail that the hex-carbon cycle was realized. Similarly, in atomic decompositions, not just the elements are relevant, but also how they are combined and interconnected. That is, one should seek a linear combination of atoms as in equation (A.1) that provides a good signal representation. Different atomic signal decomposition algorithms exist, all aim at finding a good linear combination to represent signals.

Humans, very frequently, classify things using hierarchical approaches. When looking at the night sky it is not rare to think about the constellations, then about the stars, and after, maybe, one wanders about planets orbiting the stars. The Taxonomic Classification of Species classifies live organisms, in a top-down approach, starting by kingdom down to the species. Hierarchical systems group things according to common characteristics or to physical proximity, but grouping things according to characteristics can be viewed as the grouping of things by proximity in a conceptual model/space. Atomic decompositions allow to group signals. The proximity between signal features can be accomplished by letting the dictionary include features in different geometric scales. The classification of signals according to common characteristics can be accomplished by comparing the atoms and their weights in the representation of different signals.

Currently, there is research in neuroscience and psychology to map specialized

areas in the brain that are used for different tasks. Atomizing, partitioning and classifying the natural world and its things seem to be innates to human kind. In these tasks, two common questions are: what or who are the building blocks of natural things; and how do these blocks combine? Atomic signal representations use linear combinations of pre-defined elements to represent signals. The possible building blocks are defined by the dictionary elements and atomic decomposition algorithms aim to find the building blocks composing a signal.

Bases have being largely employed to express vectors in a unique and convenient way. Bases are frequently constructed to highlight specific signal features and have an astonishing number of applications. However, there is neither freedom nor flexibility in the selection of the representation using a basis. Since the representation is unique, if one wants to express the signal in a different way then another basis must be used; that is, when bases are used to represent signals the elements are selected *a priori*. In addition, some applications may require signal representations using linearly dependent elements, which is a feature that is not allowed with bases. When an atomic decomposition strategy using a redundant dictionary is employed, the atoms in the signal representation can be chosen *a posteriori*, i.e., according to the signal. Further, the atoms used to represent the signal may be linearly dependent. These properties allow for a more appropriate and application oriented signal representation.

Atomic representations have been used for: signal filtering and denoising [76, 84]; analysis of the physical phenomena behind the observed signal together with pattern recognition and signal modelling [58,59,64,69,72,84,107]; and time-frequency [84, 85], and harmonic analyses [37,58]. Atomic representations can also provide for good signal compression tools [2, 9, 37, 44, 88]. Recently, atomic representations were proposed to discriminate outcomes from different Gaussian processes [63].

A.2 Frame Expansions

The brain is a large network of interconnected neurons, and great evidence exists that this network works in parallel. Several neurons receive the same stimulus simultaneously and deliver different responses to it, which are fed to other neurons

in different points in the network. At the end, the excitation activates some neurons that determine the response given to the stimulus provided. Instead of using a few dictionary elements to express signals, as in atomic representations, all the elements could be used to express signals. This approach gives raise to a flexibility that is similar to the one in the brain's neural network, since the input pattern can be compared to all the elements in the dictionary. The use of all the elements in the signal expression leads to the frame concept.

The frame concept remounts to Duffin and Schaeffer work on non-harmonic Fourier series [38]. In the last two decades, frames have been an active area of research that is producing many new theoretical results and applications [17, 26, 27, 84]. Frames can be regarded as redundant or overcomplete bases. A frame of a space $\mathbb{H}$ is a set of elements that spans $\mathbb{H}$. Therefore, a frame $\mathcal{G} = \{\mathbf{g}_k\}_{k\in\mathcal{K}}$ (a set of elements $\mathbf{g}_k$ indexed by $k \in \mathcal{K}$), can be used to express any $\mathbf{x} \in \mathbb{H}$ by means of

$$\mathbf{x} = \sum_{k\in\mathcal{K}} c_k \mathbf{g}_k, \tag{A.2}$$

where c_k are called the frame coefficients. Since $\mathcal{G}$ is overcomplete, the set of frame coefficients is not unique. There are several strategies to compute the frame coefficients. One strategy is the use of the inverse or dual frame, that provides the reconstruction formulas

$$\mathbf{x} = \sum_{k\in\mathcal{K}} \langle \mathbf{x}, \mathbf{g}_k \rangle \tilde{\mathbf{g}}_k = \sum_{k\in\mathcal{K}} \langle \mathbf{x}, \tilde{\mathbf{g}}_k \rangle \mathbf{g}_k, \tag{A.3}$$

where $\tilde{\mathcal{G}} = \{\tilde{\mathbf{g}}_k\}_{k\in\mathcal{K}}$ is the inverse or dual frame. Note that, since $\mathcal{G}$ is overcomplete, $\tilde{\mathcal{G}}$ is, in general, not unique [17].

A frame expansion of a signal provides "how much" of the signal is present in each frame element. This may allow to infer signal characteristics. Since the frame elements are not obligated to be linearly independent and therefore also not orthogonal, characteristics that are similar one to another can be observed by projecting the signal into different frame elements. These characteristics can be confused if the signal projection into a basis is used. Therefore, the set of frame coefficients can be altered in order to highlight desired signal features. In addition, frames can be designed for specific applications depending on the features that one desires to extract or analyze in the signal.

Although the elements used to express $\mathbf{x}$ are selected *a priori*, frame expansions allow for the use of elements with special properties. The Balian-Low theorem [26] shows that it is not possible to construct a basis with elements that are well localized in both time and frequency domains simultaneously. Frames do not impose uniqueness to the signal representation, therefore the definition of the frame elements is less restrictive than it is for basis elements. Hence, better localization of the frame elements simultaneously in time and frequency domains is achievable.

Frames applications range from signal coding [84], signal sampling [17,38,84], signal analysis [27, 84] and transient detection [51, 52], to communication systems design [17, 99].

A.3 Outline of the Thesis

This thesis is concerned with atomic decompositions and frame expansions – signal representations using redundant dictionaries. Appendix B treats signals representations using both frameworks. In general, atomic representations are said to be adaptive since the elements employed can be chosen according to the signal being decomposed. The main purpose of Appendix B is to provide a theoretical background on both atomic decompositions and frames.

Appendix C presents coherent representations of signals from electric power plants. These representations are obtained by employing an adaptive signal decomposition algorithm that obtains atomic signal representations. The main objective of a coherent decomposition is to provide a signal representation related to the physical phenomena represented in the measured signal.

Appendix D provides a statistical study of the Matching Pursuit algorithm, consisting of an iterative algorithm that builds atomic representations using a per-element selection criterion. The statistical model derived from the study is then employed to design quantizers for the coefficients of atomic decomposition obtained using Matching Pursuits for signal compression applications

Appendix E studies the time frequency content of frames. It provides a novel sufficient condition for a family of elements to be a frame.

Appendix F presents the conclusions.

Apêndice B

Background

This appendix discusses several aspects of signal representations using redundant dictionaries. It begin by presenting relevant features and characteristics of atomic decompositions. Then, it goes towards the algorithms that obtain atomic decompositions, in special the so-called Matching Pursuit algorithm, which is an iterative algorithm that selects the atoms of the atomic representation in a per-element fashion. Dictionaries to be used for atomic decompositions are then discussed. Then, this appendix also reviews basic aspects of frame theory, and discusses the problem of frame characterization. Further, methods commonly employed to construct frames from a prototype signal are addressed. It is also discussed, from an intuitive viewpoint, the connections between a frame and the dictionary composed using the elements of the frame. At the end, the review material is related to the remainder of this thesis.

B.1 Atomic Representations of Signals

The aim of atomic signal decomposition algorithms is to select a subset of M dictionary elements $\mathbf{g}_{i(m)}$ that approximates $\mathbf{x}$ using a linear combination such as

$$\mathbf{x} \approx \hat{\mathbf{x}} = \sum_{m=1}^{M} \gamma_m \mathbf{g}_{i(m)}, \quad \mathbf{g}_{i(m)} \in \mathcal{D}. \tag{B.1}$$

Equation (B.1) is also referred to as an M-term representation or expansion.

Atomic representations differ from classical transform based approaches used for signal representation, since the atoms used to represent a signal may be linearly

dependent and may not span the signal space. In addition, since, in general, $\mathcal{D}$ has more elements than necessary to span the signal space, the selection of the atoms may be signal dependent, leading to an adaptive signal decomposition. That is, while representations using basis utilize atoms that one defined *a priori*, adaptive signal representations select the atoms *a posteriori*. As previously discussed, this scheme can provide meaningful and compact approximations.

B.1.1 Distortion of the Approximation

The distortion of the M-term expansion of a signal $\mathbf{x}$ can be defined by

$$d(\mathbf{x}, M, \mathcal{D}) = \|\mathbf{x} - \hat{\mathbf{x}}\| = \left\| \mathbf{x} - \sum_{m=1}^{M} \gamma_m \mathbf{g}_{i(m)} \right\|. \tag{B.2}$$

The distortion of the M-term representation of $\mathbf{x}$ depends on three factors: i) the number of elements used in the signal representation M; ii) the atoms $\mathbf{g}_{i(m)}$ used to express the signal; iii) and the coefficients γ_i that weight the atoms in the M-term representation. Since the atoms that can be used in the M-term representation are limited by the elements that belong to the dictionary $\mathcal{D}$, the distortion also depends on $\mathcal{D}$.

B.1.2 Dictionary Requirement: Completeness

The most basic dictionary requirement is completeness. For M-terms using atoms from a dictionary $\mathcal{D}$ being capable of representing any signal $\mathbf{x} \in \mathbb{X}$ with an arbitrary distortion $d(\mathbf{x}, M, \mathcal{D})$, $\mathcal{D}$ must be complete in $\mathbb{X}$. That is, there shall be at least one linear combination of elements from $\mathcal{D}$ that gives $\hat{\mathbf{x}} = \mathbf{x}$, $\forall\ \mathbf{x} \in \mathbb{X}$, i.e. $\mathcal{D}$ must span $\mathbb{X}$. Dictionaries are often said to be overcomplete or redundant since, in general, they have more elements than necessary to span the signal space. Dictionary redundancy allows the use of distinct sets of dictionary elements to span $\mathbb{X}$, and the same applies, *a priori*, to any $\mathbf{x} \in \mathbb{X}$.

B.1.3 Adaptivity

It is desirable to select the atoms used in the representation according to the signal, allowing thus an adaptive representation. When the atoms used in the M-term expansion depend on the signal the decompositions are said to be adaptive. The

algorithms that produce the M-term representation are also referred to as adaptive signal decomposition algorithms [1, 29, 59, 62, 72, 84, 85, 105].

Since an overcomplete dictionary allows expressing the same signal using different M-term representations (meaning that the representation is not unique), an overcomplete or redundant dictionary is a requirement if adaptive signal decompositions are desired. Different decomposition algorithms may obtain signal expansions subject to distinct objectives or adapting criteria. Several adapting criteria have been used: **i**) the distortion defined in equation (B.2), **ii**) the smallest value of the peak error between $\mathbf{x}$ and its approximation $\hat{\mathbf{x}}$, **iii**) the smallest number of terms M for a given distortion, **iv**) the set of smallest energy coefficients using a given number of terms, **v**) the selection of a basis to represent the signal, among others. Mixtures of these criteria have also been applied. In general, the number of terms M, the atoms used in the representation, and the atoms weights depend on the employed criteria. Therefore, the algorithm used to obtain the M-term representation influences the linear combination obtained and the distortion of the representation. Several algorithms capable of obtaining M-term expansions are briefly described in subsection B.2.

B.1.4 Approximations and Estimations

Adaptive approximations can also be understood as discrimination strategies. Discrimination separates relevant information from noise, the relevant information is defined by the elements that are allowed to be used in the representation, the dictionary atoms. Therefore, the atomic representation can also be understood as signal estimators that retain just the relevant information, the information that is similar to the atoms used in the signal representation.

B.1.5 Compactness or Sparsity of Representations

The most compact or sparse representation of $\mathbf{x}$ is the one using the smallest number of atoms [36, 59, 66, 105] such that the distortion is null. In practice, not the smallest subset of dictionary elements, but rather a small number of terms providing an acceptable distortion may be enough. Hence, the sparsity of a representation is related to the number of terms M. As M grows, the less compact or sparse the

representation is.

The atomic signal representation is a linear combination of M elements, and, if the signal dimension is N, one may have: **i**) $M > N$ – a redundant signal representation; **ii**) $M = N$ – non-redundant and non-compact; **iii**) $M < N$ – fewer atoms than the necessary to span the signal space are used – a compact representation. For coding applications, compact representations ($M < N$) are often desirable since they allow for compacting the data.

B.1.6 Coherent Representations

Most signal processing applications deal with outcomes from physical processes. In these cases, the observed signal $\mathbf{x}$ is a mixture of components $\mathbf{p}_m$, representing physical phenomena, given by

$$\mathbf{x} = \sum_m \beta_m \mathbf{p}_m \ + \mathbf{n}, \tag{B.3}$$

where $\mathbf{n}$ is the noise, inherent to any physical measurement. From the perspective of signal modelling, it is interesting for the atoms $\mathbf{g}_{i(m)}$ used to approximate the signal to be similar to the phenomena $\mathbf{p}_i$ that are represented in the signal $\mathbf{x}$. The more the selected dictionary elements $\mathbf{g}_{i(m)}$ and weights γ_m are similar to the physical phenomena $\mathbf{p}_m$ and weights β_m, the better is the signal expansion for modelling and pattern recognition purposes. In this work, expansions such that the $\mathbf{g}_{i(m)}$ are similar to the components that are represented in the decomposed signal are said to be coherent. That is, if the signal expansion is a meaningful model, the representation is said to be coherent to the signal. Appendix C exemplifies and solves problems encountered when trying to obtain coherent decompositions by presenting efficient coherent representations of disturbance signals acquired in power systems.

B.1.7 Super Resolution

In general, the computational burden of obtaining atomic signal representations surpasses the computational demands of traditional expansions into basis, as, for example, the fast Fourier transform (FFT). Therefore, it is desirable to obtain a separation or resolution of signal features and of atom characteristics that is much higher for atomic decompositions than the one possible with traditional approaches.

B.2 Adaptive Signal Decomposition Algorithms

Practical signal processing applications and algorithms are often implemented in digital machines and the signal is considered to be a vector $\mathbf{x}$ that belongs to an N-dimensional vector space. Note that, in an N-dimensional vector space, the M-term can also be expressed, in matrix form, as

$$\hat{\mathbf{x}} \approx \mathbf{G}\vec{\gamma} = \begin{bmatrix} \mathbf{g}_1 & \cdots & \mathbf{g}_j & \cdots & \mathbf{g}_M \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_j \\ \vdots \\ \gamma_M \end{bmatrix} \tag{B.4}$$

$$\begin{bmatrix} \hat{x}[0] \\ \vdots \\ \hat{x}[j] \\ \vdots \\ \hat{x}[N] \end{bmatrix} \approx \begin{bmatrix} g_{i(1)}[0] & g_{i(2)}[0] & \cdots & g_{i(M)}[0] \\ \vdots & \vdots & \cdots & \vdots \\ g_{i(1)}[j] & g_{i(2)}[j] & \cdots & g_{i(M)}[j] \\ \vdots & \vdots & \cdots & \vdots \\ g_{i(1)}[N-1] & g_{i(2)}[N-1] & \cdots & g_{i(M)}[N-1] \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_j \\ \vdots \\ \gamma_M \end{bmatrix} \tag{B.5}$$

From equation (B.5) the best M-term representation can be obtained by a massive optimization. The massive optimization would try all possible combinations with M or less elements from $\mathcal{D}$ to find the linear combination that minimizes a prescribed error criterion. The generation of an atomic decomposition consists, thus, in finding the matrix $\mathbf{G}$ and the coefficient vector $\vec{\gamma}$ that provide a reasonable signal approximation according to a distortion criterion as the l^2 or the l^1 norm. However, massive optimization is not feasible. Even for a constrained $\mathcal{D}$, with finite cardinality, the solution of equation (B.5) is an NP-hard problem [29], that is, its complexity is not bounded by a polynomial of degree N.

Another possibility is to consider the dictionary to be a frame [17, 27, 84] (frames are discussed in section B.4). Using the frame concept, $\mathbf{G}$ in equation (B.5) has size $N \times \#_{\mathcal{D}}$. If $\mathbf{G}$ is a frame, the pseudo-inverse of $\mathbf{G}$ [17,84], $\mathbf{G}^{INV}$ can be used to find the coefficients. Several algorithms exist to calculate $\vec{\gamma} = \mathbf{G}^{INV}\mathbf{x}$ [17, 84]. However, the frame expansion $\vec{\gamma} = \mathbf{G}^{INV}\mathbf{x}$ does not allow to select the atoms used in the representation. For that purpose, the coefficients of the frame expansion could be ordered, for example in decreasing magnitude, and the expansion could use just

the M larger coefficients. However, limiting the number of terms changes $\mathbf{G}$ and implies a different set of coefficients. In addition, in some cases, the "pruned" $\mathbf{G}$ may not admit inversion.

The impracticability of massive optimization was exposed, and also the difficulty to select atoms (to adapt) using frame expansions. In order to obtain atomic representations several algorithms have been proposed. The decomposition algorithms presented to date can be classified into three different classes, irrespective of the interplay among the atoms selected by the algorithms. The first class is composed by algorithms that deliver a basis to represent the signal. The algorithms of the second class deliver a sequence of dictionary elements, that is, they employ a per-element selection criterion, such that the distortion of the approximation is reduced as the number of elements used increases. The algorithms in the third class impose that if a new atom is added to the M-term it can not be a linear combination of previously selected atoms, thus delivering subspaces that expand as the number of terms increases. Although the principles and objectives of the three classes of algorithms are different, the algorithms are said to be adaptive since their outcomes depend on the signal being decomposed. In what follows, some algorithms capable of obtaining atomic decompositions, from the three classes, are conceptually described. In broad sense, such algorithms have been classified as non-linear [29, 32, 84, 104] since the decomposition obtained for the sum of two given signals is not, in general, the sum of the decompositions of the two signals. It should also be pointed out that specific implementations are often proposed and employed considering the dictionary as an algorithmic constraint.

B.2.1 Basis Selection Algorithms

For selecting bases to represent signals two different strategies can be distinguished. The first strategy employs a dictionary composed of a union of bases and the algorithm selects one of the bases in the dictionary to express the signal [22, 84]. The second strategy builds a basis using dictionary elements [1, 15, 16, 84]. In general dictionaries are specially designed for each strategy; examples are wavelet packets or local cosine bases dictionaries [15, 84].

The Best Ortho Basis (BOB) algorithm [22] employs a dictionary that is

a union of K orthogonal basis $\mathcal{D} = \{\mathcal{B}_1, \ldots, \mathcal{B}_i, \ldots, \mathcal{B}_K\}$. The algorithm works roughly as follows. Let $\vec{\gamma}_i$ be the projection of $\mathbf{x}$ into the basis $\mathcal{B}_i$ and define an entropy $\varepsilon(\vec{\gamma}_i)$. The BOB selects the basis $\mathcal{B}_b$ that has the minimum entropy, that is, $b = \arg\left(\min_{\mathcal{B}_i \subset [1,\ldots,K]} \varepsilon(\vec{\gamma}_{\mathcal{B}_i})\right)$, to represent the signal. BOB was proposed using a dictionary of wavelet packets [22,27,84] that have a special tree-structure that allows for a fast algorithm. Since BOB searches for the minimum entropy coefficient vector it might obtain compact representations, that is, with few non-zero coefficients. However, BOB does not obtain compact representations when $\mathbf{x}$ is composed of non-orthogonal signals [84].

The Basis Pursuit (BP) is presented in [15,16], and, according to the authors, BP is a principle rather than an algorithm. The main idea of BP is to construct a basis, not necessarily orthogonal, to represent $\mathbf{x}$. In essence, the idea of BP is to search for the coefficient vector $\vec{\gamma}$ with minimum l_1 norm. One of the algorithms proposed to accomplish that is the BP-simplex [16]. The BP-simplex starts using a full rank representation of the signal, i.e., the projection of the signal into a basis, and iteratively minimizes the l_1 norm by swapping useless elements (those that do not have significant information respective to the signal being represented) by new ones.

In the basis selection algorithms above, specific and fast implementations are often proposed considering specific dictionary characteristics. Comparisons among several existent basis selection methods can be found in [1, 16, 84]. Although BOB and BP obtain an exact signal representation they are restricted to the selection of a basis to represent signals. In some applications this is undesirable, since the representation of a signal using a basis does not permit accurate signal modelling as the different phenomena represented in the signal are, in general, linearly dependent.

B.2.2 Atom Selection Algorithm – Matching Pursuit

The Matching Pursuit algorithm (MP) [84, 85] approximates signals iteratively. The basic idea of the MP is to find the best possible approximation at each iteration. This strategy is also called the Pure Greedy Algorithm (PGA) [32, 104]. Due to advances in computer resources, greedy algorithms have emerged in several scientific fields, for example, for signal processing [112], in statistics [53, 77] (where

they are also known as Projection Pursuits) and in control theory applications [14].

Let $\mathcal{D} = \{\mathbf{g}_k\}$ and $k \in \{1, \ldots, \#_\mathcal{D}\}$ such that $\|\mathbf{g}_k\| = 1 \ \forall k$, where $\#_\mathcal{D}$ is the dictionary cardinality, i.e. the number of elements in $\mathcal{D}$. In each decomposition step or iteration $n \geq 1$, the MP searches for the atom $\mathbf{g}_{i(n)} \in \mathcal{D}$, i.e. $i(n) \in \{1 \ldots \#_\mathcal{D}\}$, with largest inner product with the residual signal $\mathbf{r}_\mathbf{x}^{n-1}$ [84, 85], and the initial residue is set to be $\mathbf{r}_\mathbf{x}^0 = \mathbf{x}$. The selected atom $\mathbf{g}_{i(n)}$ is then subtracted from the residue to obtain a new residue $\mathbf{r}_\mathbf{x}^n = \mathbf{r}_\mathbf{x}^{n-1} - \gamma_n \mathbf{g}_{i(n)}$, where $\gamma_n = \langle \mathbf{r}_\mathbf{x}^{n-1}, \mathbf{g}_{i(n)} \rangle$ is the inner product between the residue and the selected atom $\mathbf{g}_{i(n)}$. Thus, at each step n, the resulting residue $\mathbf{r}_\mathbf{x}^n$ is orthogonal to $\mathbf{g}_{i(n)}$. The MP obtains the M-term signal representation/approximation

$$\hat{\mathbf{x}} \approx \sum_{n=1}^{M} \gamma_n \mathbf{g}_{i(n)}. \tag{B.6}$$

The approximation error or distortion is given by the M^{th} residue $\mathbf{r}_\mathbf{x}^M = \mathbf{x} - \hat{\mathbf{x}}$. In practice, the decomposition step (the calculation of γ_n, $i(n)$, and the residue $\mathbf{r}_\mathbf{x}^n$) is iterated until a prescribed distortion ($\|\mathbf{r}_\mathbf{x}^n\|$), a maximum number of steps, or a minimum for an approximation metric is reached [29, 84] (see subsection C.4.1).

The MP algorithm converges [29, 84, 112], i.e. $\lim_{n \to \infty} \|\mathbf{r}_\mathbf{x}^n\| = 0$. The maximum angle between any signal $\mathbf{x}$, that belongs to the signal space $\mathbb{X}$, and its closest atom in $\mathcal{D}$

$$\Theta(\mathcal{D}) = \arccos \left(\min_{\mathbf{x} \in \mathbb{X}} \left[\max_{i \in \{1, \ldots, \#_\mathcal{D}\}} \left(\frac{|\langle \mathbf{x}, \mathbf{g}_i \rangle|}{\|\mathbf{x}\|} \right) \right] \right) \tag{B.7}$$

determines the maximum error at a given MP step. If the angle between $\mathbf{r}_\mathbf{x}^{n-1}$ and $\mathbf{g}_{i(n)}$ is $\Theta(\mathcal{D})$ then the smallest possible reduction of the residue norm occurs at step n. If, at every step, the worst possible reduction of the residue norm after n MP steps occurs, therefore $|\mathbf{r}_\mathbf{x}^n|$ is upper bounded by

$$e_n = \|\mathbf{r}_\mathbf{x}^n\| \leq \|\mathbf{x}\| \sin^n (\Theta(\mathcal{D})). \tag{B.8}$$

Note, however, that this bound is weak since there is no guarantee that exists a signal such that the angle between its residuals and the atoms selected at each iteration are always $\Theta(\mathcal{D})$, when decomposed using the MP. In finite dimensional spaces the residue norm $\|\mathbf{r}_\mathbf{x}^n\|$ has an exponential decay rate as a function of n. However, in infinite dimensional spaces the convergence can be slow.

The choice of the atom with the largest inner product with the residue at each iteration may be suboptimal when the overall representation is considered.

The greedy selection criterion guarantees the reduction of the error at each step; however, this criterion can reintroduce components into the signal, examples are given in subsection C.2.2. In brief, there is no guarantee that the MP finds the best possible set of coefficients. In some cases, for the same atoms, a distinct set of coefficients can be found that leads to smaller distortion.

B.2.2.1 Variants of the MP Framework

A variant of the MP is the Weak Greedy Algorithm (WGA). The WGA [103] uses, instead of $i(n) = \arg\left\{\max_{k\in\{1,\ldots,\#_{\mathcal{D}}\}}\langle \mathbf{r}_{\mathbf{x}}^n, \mathbf{g}_k\rangle\right\}$, the selection criterion given by $i(n) = \arg_{k\in\{1,\ldots,\#_{\mathcal{D}}\}}\left\{\langle \mathbf{r}_{\mathbf{x}}^n, \mathbf{g}_k\rangle > t_n \sup_{\mathbf{g}_k\in\mathcal{D}}\langle \mathbf{r}_{\mathbf{x}}^n, \mathbf{g}_k\rangle\right\}$. The thresholds t_n may change or not with the iteration n and are optimality criteria. The WGA may reduce the computational demands since the search of the atom does not need to be exhaustive in $\mathcal{D}$. However, the thresholds t_n must be set-up and $\sup_{\mathbf{g}_k\in\mathcal{D}}\langle \mathbf{r}_{\mathbf{x}}^n, \mathbf{g}_k\rangle$ needs to be calculated. The Probabilistic MP (PMP) [46] is another variant of the MP. In the PMP, the atom choice depends not just on the inner product but also on the suitability (measured by a probability weight) of the atom to be used to approximate the signal at each step.

Quantized MP-like Decompositions MP resembles successive approximation vector quantization (SAVQ) [25, 81]. The difference is that, in MP, the coefficient can assume any value while, in SAVQ, the coefficients are given by $\gamma_n = {\alpha_0}^{n-1}$ being α_0 a constant and n the iteration. Therefore, SAVQ uses exponentially quantized coefficients. The so called Matching Pursuits using Generalized Bit Planes (MP-GBP) [8, 9] employs coefficients $\gamma_n = {\alpha_0}^{k_n}$, where k_n can be any integer value at each iteration n. The α-expansions [24,49] permit only one coefficient value, a preset value α. The three algorithms cited (SAVQ, MP-GBP, α-expansions) quantize the coefficient prior to computing the residue, and is referred as in-loop quantization. As for computing the residue the decomposition loop uses quantized coefficients.

Local Fitting The High Resolution Pursuits algorithm (HRP) [65, 72] was introduced in order to allow a better fitting between the atom and the signal for time-frequency atom dictionaries. Time-frequency atoms are designed to have their energies concentrated in specific time intervals and bandwidths. The MP searches

for the atom that best matches the overall signal what may provide a bad local fitting. The HRP multiplies the time-frequency atom found by the MP by B-spline windows in order to find the best local fitting with the residue. The local "fitting" strategy of the HRP tries to reconcile the overall match of the MP and localized signal features. In subsection C.2.2 a fast algorithm to accomplish local fitting in order to eliminate pre-echo and post-echo artifacts that often appear in MP-like algorithms is presented.

B.2.3 Subspaces Selection Algorithms

Two different approaches can be found to obtain subspace-based signal representations. The first strategy is to restrict $\mathbf{G}$ in equation (B.5) to have a fixed dimension $N \times M$ [59], with $M < N$, and to optimize $\mathbf{G}$ by trying different subsets of M elements from $\mathcal{D}$ in order to represent the signal. This optimization is feasible when $\mathbf{G}$ can be constructed from orthogonal vectors [59]. The second approach uses an MP-like algorithm but orthogonalizes the atom selected at each step with respect to the atoms previously used. That is, as each new atom is delivered by the algorithm, the dimension of the subspace generated by the atoms selected by the algorithm to represent the signal increases until the whole signal space dimension is reached [29, 90, 105]. Note that in these approaches if the number of terms used equals the signal space dimension, then a basis is used to represent the signal. However, these are placed apart form the basis selection algorithms since they may not use a basis.

In the Orthogonal Matching Pursuits algorithm (OMP) [29, 90] the atom chosen at each step is made orthogonal to the previously selected atoms using a Gram-Schmidt orthogonalization procedure. The orthogonalization procedure does not produce any new information to be transmitted. The subspaces generated by the atoms selected up to the n^{th} iteration V_n (the space generated by all linear combinations of the atoms selected up to step n) is such that $V_n \supset V_{n-1}$; that is the dimension of the subspace generated by the atoms selected by the OMP increases at each iteration. Thus, if the number of steps or terms equals the signal space dimension, then a basis is used to represent the signal. Note that this basis is such that its elements are given by linear combinations of dictionary atoms. The

orthogonalization procedure is, in general, not useful for phenomena identification and pattern recognition. In addition, the OMP may not lead to the smallest possible residue; to obtain the smallest residue using orthogonalized atoms, the Optimized OMP algorithm (OOMP) has been proposed [3, 93].

B.3 Dictionaries

Dictionary completeness allows to represent any signal with arbitrarily small reconstruction error, however, it does not guarantee the obtainment of neither a compact nor coherent representation. Dictionaries are constantly proposed for specific applications. A very commonly employed dictionary is the Gabor dictionary [27, 84, 85]. The Gabor dictionary is composed by Gaussian shaped atoms in different scales, with varying centers in time and multiplied by different complex sinusoids, the so-called time-frequency atoms [27, 84, 85]. Other dictionary examples are wavelet packets [8, 50, 84], local cosine bases [1, 16, 84] and curvelets dictionaries [8, 50]. Ridgelets [11] dictionaries have also been used to accomplish atomic representations of natural scenes [35]. Dictionaries composed by outcomes of random processes have also been used [54, 60, 62] mainly to investigate properties of decomposition algorithms. Training process have being employed to generate dictionaries [42, 43, 78]. For example, in [2, 87] training was used to find good dictionaries for representing video frame differences, after motion compensation for MPEG-4 coders. Mixtures of the dictionary construction approaches previously cited have also been used.

The impact of the dictionary definition, i.e., the atoms that belong to the dictionary, for atomic decompositions algorithms is tremendous. It was discussed that the dictionary has a direct influence on the distortion. In addition, the dictionary must contain atoms that are similar to the actual phenomena that may appear in the observed signal to allow for coherent expansions. Meanwhile, in order to provide compact representations, the dictionary may have to contain a very large number of elements allowing the selection of just a few elements to represent the signal with small distortion.

Conceptually, the most simple dictionary is a finite set of elements $\mathcal{D} =$

$\{\mathbf{g}_i\}_{i\in\mathcal{I}}$. In this case, each $\mathbf{g}_i \in \mathcal{D}$ can be indexed by $i \in \mathcal{I} = \{1, \ldots, \#_\mathcal{D}\}$, where $\#_\mathcal{D}$ is the number of elements in $\mathcal{D}$, the dictionary cardinality.

Intuitively, the probability of having an atom in the dictionary that is a good match to the signal being decomposed, *a priori*, increases with $\#_\mathcal{D}$, as mentioned in [54,62,84]. Therefore, in some applications a large cardinality dictionary is convenient such that the dictionary may contain atoms that are similar to all the potential components in the signals to be decomposed.

B.3.1 Parameterized Dictionaries

The probability of finding an atom in the dictionary that is a good match to the signal does not rely only on $\#_\mathcal{D}$, it depends also on the shape of the atoms. Thus, another relevant aspect in the definition of $\mathcal{D}$ is the actual waveform or shape of the atoms composing $\mathcal{D}$. If the class of components that may be represented in the observed signals to be decomposed is known *a priori* then it would be wise to use atoms that look like these components [54, 62, 84]. The use of atoms that resemble the possible signal components would augment the probability of finding a dictionary atom that matches well the signal being decomposed.

A strategy commonly used is to define the dictionary elements from a set of prototype functions/signals. In such dictionaries the actual waveforms of the dictionary atoms depend on a set of parameters that modifies the prototype signal. These dictionaries are said to be parameterized since each dictionary element $\mathbf{g}_\sigma$ is defined by a given value of the parameter set

$$\sigma = (\sigma_0, \sigma_1, \ldots, \sigma_{K-1}) \in S, \tag{B.9}$$

where K is the number of parameters that define $\mathbf{g}_\sigma$ and S is the set of all possible parameters. The Gabor dictionary previously mentioned is a parameterized dictionary.

If the set of possible atom parameters S has finite cardinality then the dictionary is therefore indexable. In this case the coding of the M-term representation, the storage or transmission of the atoms $\mathbf{g}_{i(m)}$ and weights γ_m $1 \leq m \leq M$, is accomplished by coding the weights of the atoms together with the σ or indices to the atoms. The atoms of parameterized dictionaries are often referred to as structu-

res [84, 85] and are unequivocally defined by the value of the parameter set σ. The stored parameters of the structures of the decomposition of a given signal compose the so-called structure book [84, 85, 94]. Thus, the structure book expresses the M-term signal representation in the parameter space.

Parameterized dictionaries are a good and natural choice when one intends to represent signals from a given class. This is so because one can define prototype signals according to the phenomena that are expected to be represented in signals from a given class. Parameterized dictionaries allow to look for the structures that best match to the signal being decomposed. These structures are derived from prototype signals which actual characteristics as band, center frequency, scale, phase, start time, attack, decay rate, etc depending on the parameter set values defining the dictionary atoms. Therefore, the use of a parameterized dictionary allows for estimating the signal and obtaining coherent decompositions. For example, parameterized dictionaries were employed for pattern recognition [72] and signal modelling [45,107] using atomic decompositions.

In addition, the values of the parameters that define the atoms in the dictionary can also be adapted to the signal source. In [29] a training process based on the Lind-Buzo-Gray algorithm [57] was used to reduce the number of elements in Gabor dictionaries for the specific task of representing chirp signals. In [39] a similar training algorithm is employed to both reduce the algorithm complexity and unbias the structure book for electroencephalogram data representation.

Summarizing, application oriented dictionaries are often used. However, in some applications finite cardinality parameter set dictionaries are not enough and the capability do adapt the atoms to the signal may be required.

Continuous Parameters Dictionaries In some cases it is interesting to change, adapt or fit the structures used in the signal representation, to the actual signal being decomposed. For that purpose, the value of the parameter set of an atom is allowed to be any point inside a region of the parameter set space. For example, if all the parameters are real then $S \subset \mathbb{R}^K$, S is nondenumerable and $\#_{\mathcal{D}} = \aleph_1$ ($\aleph_1$ is the cardinality of the real numbers). In this case we say that the parameters of the atoms are continuous. In general, to obtain continuous parameter atoms one uses optimization algorithms that find the values of the parameter sets defining

each atom in the M-term representation. When this occurs the structure book is densely quantized. In some cases this may require specialized quantization of the parameter set. Appendix C shows a strategy to obtain decompositions using continuous parameter atoms in the MP algorithm and also discusses the quantization of the parameter set.

B.3.2 Dictionary Trade-offs

The dictionary cardinality has a direct impact on:

i – Computational complexity: in general, the computational demands of the algorithms that provide atomic representations is related to $\#_{\mathcal{D}}$. The larger that $\#_{\mathcal{D}}$ is the higher is, in general, the computational demands of the decomposition algorithm;

ii – Coding cost: the rate needed to code each atom used in the atomic decomposition increases with $\#_{\mathcal{D}}$.

Therefore, in addition to the design of the atoms $\mathbf{g}_i \in \mathcal{D}$, a key problem is to control the value of $\#_{\mathcal{D}}$ while guaranteeing acceptable distortion at low data rate and reduced computational complexity.

Suppose that M-term representations of signals from a source $\mathcal{X}$ are obtained using a dictionary $\mathcal{D}_i$, then it is easy to find a code such that the total length of each coded representation is given by

$$R_i = R_{coef} + M * \log_2 \left(\#_{\mathcal{D}_i}\right) \text{ in bits,} \tag{B.10}$$

where R_{coef} is the number of bits used to code the M coefficients in the M-term representation. For coding purposes, it is worth using another dictionary $\mathcal{D}_k$ to represent signals from the source $\mathcal{X}$ instead of $\mathcal{D}_i$ if and only if the average distortion for the source of the representations using $\mathcal{D}_k$ is smaller than the average distortion using $\mathcal{D}_i$ at the same data length. Note that for fair comparison the same data length N should be considered rather than the same number of terms M. The computational burden of obtaining the decompositions using dictionaries $\mathcal{D}_i$ and $\mathcal{D}_k$ should also be compared.

For a continuous parameter dictionary $\mathcal{D}$ one has that $\#_{\mathcal{D}} = \aleph_1$. In this case the parameters of the structures should be quantized for coding the M-terms and the term $M * \log_2(\#_{\mathcal{D}_i})$ in equation (B.10) is replaced by the data length used to transmit/store the quantized structures. However, even when atom parameters adaptation takes place, it is still needed to design initial dictionaries that allow for this adaptation with reduced computational complexity.

B.3.3 Dictionary Evaluation Metrics

In greedy decomposition algorithms, some of the factors that determine both the minimum error obtainable at a given number of steps and the convergence rate of the approximation process are:

- The maximum angle $\Theta(D)$ between any signal in the space and its closest atom in $\mathcal{D}$ (see equation (B.7)) that bounds the error norm.

- The average of the angle among vectors drawn from a source $\mathcal{X}$, whose realizations have an uniformly distributed orientation, and their closest atoms in $\mathcal{D}$

$$\overline{\Theta}(\mathcal{D}) = E\left\{\arccos\left[\max_{i\in\{1,\ldots,\#_{\mathcal{D}}\}}\left(|\langle\mathcal{X},\mathbf{g}_i\rangle|\right)\right]\right\}. \tag{B.11}$$

The optimal dictionary must have the minimum attainable $\Theta(\mathcal{D})$ for a given cardinality $\#_{\mathcal{D}}$; subsection D.1.2 provides more discussions and examples on this topic.

Another dictionary metric is the coherence of a dictionary [105]

$$\mu(\mathcal{D}) = \max_{\gamma_j\in\Gamma}\left(\max_{\gamma_k\in\Gamma-\{\gamma_j\}}|\langle\mathbf{g}_{\gamma_j},\mathbf{g}_{\gamma_k}\rangle|\right). \tag{B.12}$$

which measures the largest similarity between two dictionary atoms. As dictionary atoms are normalized $\mu(\mathcal{D}) \leq 1$. If $\mu(\mathcal{D})$ is "close" to 1 it is said that the dictionary is coherent and for $\mu(\mathcal{D})$ small it is said that the dictionary is incoherent [105]. In [105] it is shown that $\mu(\mathcal{D})$ is related to the exact sparse problem. That is, if p elements of $\mathcal{D}$ are linearly combined to generate a signal $\mathbf{x}$, then the OMP [29] and the BP [16] can recover any such $\mathbf{x}$ if $p \leq \frac{1}{2}(\mu(\mathcal{D})^{-1} + 1)$ [105].

B.4 Frames

Frames were introduced by Duffin and Schaeffer for the study of non-harmonic Fourier series [38] in 1952. The basic idea is the use of more "points" or coefficients than the necessary to represent the signals from a given space. This idea is used in modern analog-to-digital converters (ADC) that sample signals at rates higher than the Nyquist rate with low resolution, the so-called sigma-delta ADC. Since the signal is sampled at high rate, the requirement on the sample precision can be relaxed.

It took a long time until frames came into play. The first book concerning the topic was written by Young in 1980 [111]. However, once Daubechies, Grossman and Meyer [28] in 1986 perceived the fact that frames can be understood in a way similar to expansions using orthonormal bases, and the link among frames and wavelets was established [26, 27, 68], new breath was given to the research on frames. Frames have been an active area of research producing a plethora of theoretical results and applications. Good surveys on frames can be found in [7, 26, 27, 84], and a recent book dedicated to this topic is [17].

B.4.1 Definition of Frames

Definition B.1 *A sequence of elements* $\{g_k\}_{k\in\mathcal{K}}$ *in a space* $\mathbb{H}$ *is a* frame *for* $\mathbb{H}$ *if there exist constants* $A, B > 0$ *such that* [7, 17, 27, 84]

$$A\|x\|^2 \leq \sum_{k\in\mathcal{K}} |\langle x, g_k\rangle|^2 \leq B\|x\|^2, \quad \forall x \in \mathbb{H}. \tag{B.13}$$

The numbers A and B are called the lower and upper *frame bounds*, respectively, and are not unique. The optimal lower frame bound is the supremo on all A and the optimal upper frame bound is the infimo on all B [17]. It is said that the frame is normalized if $||g_k|| = 1,\ \forall k \in \mathcal{K}$ [27].

A commonly addressed problem, related to frames, is how to determine if a given sequence of elements g_k in $\mathbb{H}$ is a frame of $\mathbb{H}$, the frame characterization problem [17]. This characterization is commonly accomplished by the frame bounds. Note that if for any signal $A = 0$ then the signal is orthogonal to all the elements $\{g_k\}_{k\in\mathcal{I}}$ and the signal can not be represented using these elements. That is the reason for the requirement that $A > 0$. If for a given signal there is no upper bound in equation (B.13) then this means that the elements are too "dense" for that signal.

An important concept is the inverse frame, that provides the reconstruction formulas

$$x = \sum_{k \in \mathcal{K}} \langle x, \tilde{g}_k \rangle g_k = \sum_{k \in \mathcal{K}} \langle x, g_k \rangle \tilde{g}_k. \tag{B.14}$$

The set $\{\tilde{g}_k\}_{k \in \mathcal{K}}$ is composed by the inverse frame elements.

Definition B.2 *Define the frame operator* $S\{\cdot\}$ *as*

$$S : \mathbb{H} \to \mathbb{H}, \; S\{x\} = \sum_{k \in \mathcal{K}} \langle x, g_k \rangle g_k. \tag{B.15}$$

A direct consequence of this definition is that any x can be synthesized by

$$x = \sum_{k \in \mathcal{K}} \langle x, g_k \rangle S^{-1}\{g_k\} \tag{B.16}$$

and S^{-1} is called the inverse frame operator. Note that $\tilde{g}_k = S^{-1}\{g_k\}$ are the elements of the inverse frame or the so-called canonical dual.

When $A = B$ the frame is said to be tight [17] and

$$S^{-1}\{\cdot\} = S\{\cdot\}/A. \tag{B.17}$$

Frames for which $A \approx B$ are said to be snug and for these $S^{-1}\{\cdot\} \approx S\{\cdot\}/A$.

B.4.2 Frame Bounds and Frame Tightness

Frames are often characterized in terms of their frame bounds. It is very common to define frame bounds ratio A/B [26]. It is also possible to define the "tightness" of a frame from its frame bounds ratio: the closer A and B are, the tighter the frame is. The frame bounds provide limits for the energy scaling of signals when represented using the projections into the frame elements.

Daubechies [27] shows that the frame operator gives

$$x = \frac{2}{A+B} \sum_{k \in \mathcal{K}} \langle x, g_k \rangle g_k + R\{x\}, \tag{B.18}$$

where $R\{x\}$ is the error incurred in reconstructing x from the projections of x into the frame elements instead of into the inverse frame elements (valid for discrete and vector spaces as well). In addition, she also shows that

$$R\{\cdot\} = I\{\cdot\} - \frac{2}{A+B} S\{\cdot\}, \tag{B.19}$$

where $I\{\cdot\}$ is the identity operator. Therefore one has that

$$-\frac{B-A}{B+A}I\{x\} \leq R\{x\} \leq \frac{B-A}{B+A}I\{x\} \quad \longrightarrow \quad \|R\{x\}\| \leq \frac{B-A}{B+A}\|x\|. \tag{B.20}$$

Hence, one sees that if B/A is close to one, then the error incurred in the reconstruction by equating the inverse frame to the direct frame is small. This error is smaller as A and B become closer. Therefore, from an analysis-synthesis perspective if the frame is tight there is no need to find the inverse. Tight frames are self-dual [62], that is, the inverse frame to a tight frame is a scaled version of the frame itself.

For normalized frames, Daubechies calls $\frac{A+B}{2}$ the frame redundancy. Several studies have been developed on normalized tight frames. For example [86] shows that these frames have noise reduction property. That is, if a given reconstruction distortion is required one can use less bits to represent the frame coefficients in comparison to the precision required to represent basis coefficients. Tight frames became a very popular and important class of frames [12, 40, 41, 48, 61, 62, 106]. Among other applications, tight frames are the most resistant to coefficients erasures [61]. Tight frames can also provide sparse signal decompositions [66, 106]. Due to these and other desirable properties a lot of attention has been given to the construction of tight frames [4, 12, 40, 41, 106].

If two families of elements $\{g_k\}_{k\in\mathcal{K}}$ and $\{h_l\}_{l\in\mathcal{L}}$ are frames of a space then the union of these families is also a frame. From the assumptions

$$A_1\|x\|^2 \leq \sum_{k\in\mathcal{K}} |\langle x, g_k\rangle|^2 \leq B_1\|x\|^2, \text{ and } A_2\|x\|^2 \leq \sum_{l\in\mathcal{L}} |\langle x, h_l\rangle|^2 \leq B_2\|x\|^2, \ \forall x \in \mathbb{H}, \tag{B.21}$$

then

$$(A_1 + A_2)\|x\|^2 \leq \sum_{k\in\mathcal{K}} |\langle x, g_k\rangle|^2 + \sum_{l\in\mathcal{L}} |\langle x, h_l\rangle|^2 \leq (B_1 + B_2)\,\|x\|^2, \quad \forall x \in \mathbb{H}. \tag{B.22}$$

If the families of $\{g_k\}_{k\in\mathcal{K}}$ and $\{h_l\}_{l\in\mathcal{L}}$ are carefully chosen the union of two frames can give rise to a tighter frame. However, defining the elements of the sequence $\{h_l\}_{l\in\mathcal{L}}$ with respect to sequence $\{g_k\}_{k\in\mathcal{K}}$ to achieve this goal is not trivial. Subsection B.4.4.1 introduces Weyl-Heisenberg frames that are constructed from translations and modulations of a prototype signal, and section E.3 shows an approach to define $\{g_k\}_{k\in\mathcal{K}}$ with respect to $\{h_l\}_{l\in\mathcal{L}}$ for the specific case of Weyl-Heisenberg frames in order to obtain tighter frames.

B.4.3 Frames in Finite Dimensional Spaces

In an N-dimensional vector space $\mathbb{H}^N$ one in general restricts the frame to have K elements; and thus equation (B.13) becomes

$$A\|\mathbf{x}\|^2 \leq \sum_{k=1}^{K} |\langle \mathbf{x}, \mathbf{g}_k \rangle|^2 \leq B\|\mathbf{x}\|^2, \qquad \mathbf{x} \in \mathbb{H}^N. \tag{B.23}$$

Let the synthesis operator be

$$\mathbf{T}\{\cdot\} : \mathbb{C}^K \to \mathbb{H}^N, \quad \mathbf{T}\{c_k\}_{k=1}^{k=K} = \sum_{k=1}^{K} c_k \mathbf{g}_k, \tag{B.24}$$

where $\mathbb{C}^K$ is a K-dimensional complex vector space. Let the analysis operator $\mathbf{T}^*\{\cdot\}$ (adjunct operator of $\mathbf{T}\{\cdot\}$) be given by

$$\mathbf{T}^*\{\cdot\} : \mathbb{H}^N \to \mathbb{C}^K, \quad \mathbf{T}^*\{\mathbf{x}\} = \{\langle \mathbf{x}, \mathbf{g}_k \rangle\}_{k=1}^{k=K}. \tag{B.25}$$

The operator $\mathbf{T}\{\cdot\}$ synthesizes $\mathbf{x}$ from the frame coefficients c_k that are obtained by the analysis operator of a dual frame given by

$$\tilde{\mathbf{T}}^*\{\cdot\} : \mathbb{H}^N \to \mathbb{C}^K, \quad \tilde{\mathbf{T}}^*\{\mathbf{x}\} = \{\langle \mathbf{x}, \tilde{\mathbf{g}}_k \rangle\}_{k=1}^{k=K}. \tag{B.26}$$

Using the analysis and synthesis operators the frame operator is then given by

$$\mathbf{S} : \mathbb{H}^N \to \mathbb{H}^N, \quad \mathbf{S}\{\mathbf{x}\} = \mathbf{T}\{\mathbf{T}^*\{\mathbf{x}\}\} = \sum_{k=1}^{K} \langle \mathbf{x}, \mathbf{g}_k \rangle \mathbf{g}_k. \tag{B.27}$$

In vector spaces the operators $\mathbf{T}\{\cdot\}$ and $\mathbf{T}^*\{\cdot\}$ can be interpreted as matrices [17], being $\mathbf{x}$ a column vector, one has $\mathbf{S} = \mathbf{T}\mathbf{T}^*$. Let ρ_i be the eigenvalues of $\mathbf{S}$, then the frame bounds are given by $A = \min_i \rho_i$, and $B = \max_i \rho_i$ [17]. Thus if $\mathbf{T}\mathbf{T}^* = A\mathbf{I}_N$ ($\mathbf{I}_N$ is the identity matrix of size N) the frame is tight ($A = B$). Hence, for a tight frame $\mathbf{S}^{-1}\mathbf{S} = \mathbf{I}_N$.

Note that in a vector space if the lower frame bound A were zero for a signal $\mathbf{x}$ then the frame elements would all be orthogonal to $\mathbf{x}$, and the frame would not be capable of representing $\mathbf{x}$. If the frame has a finite set of elements then necessarily, due to the Cauchy-Schwartz inequality, the upper frame bound condition will never be violated.

B.4.4 Frames from a Prototype Signal

If a given signal has to be analyzed or synthesized using a set pre-defined signals derived by operations over a prototype signal one should guarantee this set to be a frame. Since this frame can be used as dictionary, this may be used to generate parameterized dictionaries. There are several ways to construct frames from operations on a prototype signal, some of them are based on translations, modulations and dilations of a function $g(t)$ [17]:

Definition B.3

$$\textit{Translation by } a \in \mathbb{R}, \quad T_a : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad (T_a g)(t) = g(t-a); \tag{B.28}$$

$$\textit{Modulation by } b \in \mathbb{R}, \quad E_b : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad (E_b g)(t) = g(t) e^{2\pi j b t}; \tag{B.29}$$

$$\textit{Dilation by } c \in \mathbb{R} - \{0\}, \quad D_c : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad (D_c g)(t) = \frac{1}{\sqrt{c}} g\left(\frac{t}{c}\right). \tag{B.30}$$

The most common approaches to construct frames from a prototype are:

- A frame constructed by translations of a given function $g(t)$ through operations $\{T_{na} g(t)\}_{n \in \mathbb{Z}}$ is called a frame of translates [17];
- A Weyl-Heisenberg (WH) or Gabor frame is a frame in $L^2(\mathbb{R})$ obtained through operations $\{E_{mb} T_{na} g(t)\}_{m,n \in \mathbb{Z}}$, where $a, b > 0$ and $g(t)$ is a fixed function. These frames are also called Gabor Systems or Windowed Fourier Frames [27, 84, 85];
- A frame constructed by dilations and translations of a prototype (mother) function $g(t)$ by $\{T_{nac^j} D_{c^j} g\}_{j,n \in \mathbb{Z}} = \{c^{\frac{-j}{2}} g(c^{-j} t - na)\}_{j,n \in \mathbb{Z}}$, with $c > 1$, $a > 0$ and $g \in L^2(\mathbb{R})$, is called a wavelet frame [27, 84, 85].

B.4.4.1 Gabor Frames

While the introduction of the frame concept remounts to 1952 [38]; according to Christensen [17], the first mention of what are now called Gabor of Weyl-Heisenberg frames can be traced back to the work of Gabor on communications [55] in 1946 and to the book of von Neumann on quantum mechanics originally published in 1932 [109].

Definition B.4 *A Gabor frame is a frame for $L^2(\mathbb{R})$ obtained through operations $\{E_{mb}T_{na}g(t)\}_{m,n\in\mathbb{Z}}$, on a fixed function $g(t)$ with $a, b > 0$.*

Paraphrasing Christensen [17]: which conditions should hold on $g(t)$ in order to $\{E_{mb}T_{na}g(t)\}_{m,n\in\mathbb{Z}}$ being a frame? The answer depends on a complicated interplay between $g(t)$, a and b. For example in [74] a set of non-intuitive conditions was shown to hold on a and b to generate a Gabor System based on the characteristic function. In what follows, several results collected from the frame literature are presented, which provide either sufficient or necessary conditions for $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ to constitute a frame.

For $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ to compose a frame it is necessary that $ab \leq 1$ [17,27,84]. That is, if $ab > 1$ a frame will not be obtained; however, the assumption $ab \leq 1$ does not guarantee the generation of a frame for any $g(t)$, see for example [74]. It should be observed that ab is a measure of the density of the frame in the time-frequency plane [21, 27, 71, 84, 85]; the smaller ab is the denser is the frame.

If $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ constitutes a frame the frame bounds necessarily satisfy [17]

$$\forall t \in \mathbb{R},\ A \leq \frac{1}{b}\sum_{n} |g(t - na)|^2 \leq B \tag{B.31}$$

$$\forall \omega \in \mathbb{R},\ A \leq \frac{1}{b}\sum_{k} \left|\hat{g}\left(\omega - k\frac{b}{2\pi}\right)\right|^2 \leq B. \tag{B.32}$$

A well known sufficient condition for $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ to be a frame is presented in [68]. Let $a, b > 0$, $g(t) \in L^2(\mathbb{R})$, denote $\overline{g}(t)$ the complex conjugate of $g(t)$ and suppose that $\exists\ A, B > 0$ such that

$$A \leq \sum_{n\in\mathbb{Z}} |g(t - na)|^2 \leq B\ \forall t \in \mathbb{R} \text{ and} \tag{B.33}$$

$$\sum_{k\neq 0} \left|\left| \sum_{n\in\ \mathbb{Z}} T_{na} g T_{na+\frac{k}{b}} \overline{g} \right|\right|_{\infty} < A, \tag{B.34}$$

then $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a Gabor frame for $L^2(\mathbb{R})$.

A more general sufficient condition for the generation of a frame $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ for $a, b > 0$ given and $g(t) \in L^2(\mathbb{R})$ is [13, 17]: if

$$B := \frac{1}{b}\sup_{t\in[0,a]} \sum_{k\in\mathbb{Z}} \left| \sum_{n\in\mathbb{Z}} g(t - na)\, \overline{g\left(t - na - \frac{k}{b}\right)} \right| < \infty, \text{ and} \tag{B.35}$$

$$A := \frac{1}{b}\inf_{t\in[0,a]} \left[\sum_{n\in\mathbb{Z}} |g(t - na)|^2 - \sum_{k\neq 0} \left| \sum_{n\in\mathbb{Z}} g(t - na)\, \overline{g\left(t - na - \frac{k}{b}\right)} \right| \right] > 0 \tag{B.36}$$

then $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a frame for $L^2(\mathbb{R})$ with frame bounds A, B. Note that this result shows that if $g(t)$ has a limited support $[0, 1/x]$ then for any set $ab \leq 1$ and $b < x$ a Gabor frame is obtained.

Other conditions for the generation of Gabor frames exist (see for example [17, 27, 84, 113]). In [100] an extension of the results in equations (B.35) and (B.36) for irregularly sampled time-frequency parameters (when the set (an, bm) is replaced by any pair $(a_{n,m}, b_{n,m}) \in [na, (n+1)\,a] \times [mb, (m+1)\,b]$) is provided.

B.4.4.2 Wavelet Frames

Definition B.5 *For $c > 1$, $a > 0$ and $g \in L^2(\mathbb{R})$, a frame constructed by dilations and translations as $\{T_{nac^j}D_{c^j}g\}_{j,n\in\mathbb{Z}} = \{c^{\frac{j}{2}}g\,(c^j t - na)\}_{j,n\in\mathbb{Z}}$ is called a wavelet frame.*

In [27] both necessary and sufficient conditions are provided to construct wavelet frames. For example, if $\{T_{nac^j}D_{c^j}g\}_{j,n\in\mathbb{Z}} = \{c^{\frac{j}{2}}g\,(c^j t - na)\}_{j,n\in\mathbb{Z}}$ is a frame with frame bounds A and B then necessarily [18]

$$A \leq \frac{1}{a}\sum_{j\in\mathbb{Z}} \left|\hat{g}\left(c^j\omega\right)\right|^2 \leq B, \tag{B.37}$$

where $\hat{g}\,(\omega)$ is the Fourier transform of $g(t)$.

A sufficient condition to generate a wavelet frame [17] is: suppose that $c > 1$, $a > 0$ and $g(t) \in L^2(\mathbb{R})$ are given, if

$$B := \frac{1}{b}\sup_{|\omega|\in[1,c]} \sum_{j,n\in\mathbb{Z}} \left|\hat{g}\left(c^j\omega\right)\hat{g}\left(c^j\omega + \frac{n}{a}\right)\right| < \infty, \qquad \text{and} \tag{B.38}$$

$$A := \frac{1}{b}\inf_{|\omega|\in[1,c]} \left[\sum_{n\in\mathbb{Z}} \left|\hat{g}\left(c^j\omega\right)\right|^2 - \sum_{n\neq 0}\sum_{j\in\mathbb{Z}} \left|\hat{g}\left(c^j\omega\right)\hat{g}\left(c^j\omega + \frac{n}{a}\right)\right|\right] > 0 \tag{B.39}$$

then $\{T_{nac^j}D_{c^j}g\}_{j,n\in\mathbb{Z}}$ is a frame for $L^2(\mathbb{R})$ with bounds A, B given by the expressions above. In [100] an extension of this condition for irregularly sampled time-scale parameters, when the set (an, c^j) is replaced by any pair $(a_{n,j}, c_{n,j}) \in [c^j an, c^j a\,(n+1)] \times [c^j, c^{j+1}]$, is provided.

B.4.4.3 Discrete Spaces

Frames in discrete spaces ($l^2(\mathbb{Z})$) can be obtained by time-sampling the elements of frames in $L^2(\mathbb{R})$ [17]. If $g(t) \in L(\mathbb{R})$, if $g(t)$ is such that

$$\lim_{t\to 0} \sum_{k\in\mathbb{Z}} \frac{1}{\epsilon} \int_{-\frac{1}{2}\epsilon}^{\frac{1}{2}\epsilon} |g(k+t) - g(k)|^2 dt = 0, \tag{B.40}$$

and if $\{E_{m/Q}T_{n/P}g\}_{m,n\in\mathbb{Z}}$ with Q, $P \in \mathbb{N}$ is a frame for $L^2(\mathbb{R})$ with frame bounds A and B then $\{E_{m/Q}T_{n/P}\mathbf{g}_D\}_{n\in\mathbb{Z},\ m=0,1,\ldots,M-1}$, where $\mathbf{g}_D$ is the discretized version of $g(t)$ with one sample per time unit, i.e. $\mathbf{g}_D = g(j)_{j\in\mathbb{Z}}$, is a frame for $l^2(\mathbb{Z})$ with frame bounds A and B [17].

B.4.4.4 N-Dimensional Vector Spaces

In vector spaces $\mathbb{H}^N$ the simpler solution is to truncate or box-window the elements of a frame in the discrete space; however, this simple approach alters the frame bounds [98]. An alternative is to consider that the vector space is generated from an N-length section of an N-periodic $l^2(\mathbb{Z})$ space, where the translation operator is a circular shift. The circular shift of a signal does not change the signal norm; this way the frame in $\mathbb{H}^N$ has the same frame bounds as the frame in the N-periodic $l^2(\mathbb{Z})$.

B.5 Dictionaries From Frames

The initial requirement for atomic decompositions being capable of decomposing any signal $\mathbf{x} \in \mathbb{H}^N$ is that $\mathcal{D}$ must be complete in $\mathbb{H}^N$ [84]. Since frames are complete, they can be used to generate dictionaries for adaptive signal decompositions [17, 26, 40, 41, 84]. When generating dictionaries from frames, any signal is guaranteed to be represented by at least one linear combination of dictionary elements.

Greedy algorithms are executed in digital machines, meaning that the space in question is finite dimensional. In a vector space $\mathbb{H}^N$ any finite dictionary $\mathcal{D}$, i.e. $\#_{\mathcal{D}} < \infty$, is also a frame in $\mathbb{H}^N$. The argument for that is very simple. Let $\mathcal{D}$ be composed of elements $\mathbf{g}_k$, such that $||\mathbf{g}_k|| = 1\ \forall k$. Suppose that the set of elements

that compose $\mathcal{D}$ does not satisfy the frame definition of equation (B.23), then either

$$\sum_{k=1}^{\#_{\mathcal{D}}} |\langle \mathbf{x}, \mathbf{g}_k \rangle|^2 = 0 \tag{B.41}$$

or

$$\sum_{k=1}^{\#_{\mathcal{D}}} |\langle \mathbf{x}, \mathbf{g}_k \rangle|^2 = \infty \tag{B.42}$$

for at least one $\mathbf{x} \in \mathbb{H}^N$. Equation (B.41) is in contradiction with $\mathcal{D}$ being a dictionary in $\mathbb{H}^N$, since in this case for all $\mathbf{x}$ shall exist at least one element $\mathbf{g}_k \in \mathcal{D}$ such that $\langle \mathbf{x}, \mathbf{g}_k \rangle \neq 0$, otherwise $\mathbf{x}$ could not be decomposed using $\mathcal{D}$. Equation (B.42) also does not hold, if $\|\mathbf{x}\|^2 < \infty$, since all $\mathbf{g}_k$ have unit norm, then necessarily

$$\sum_{k=1}^{\#_{\mathcal{D}}} |\langle \mathbf{x}, \mathbf{g}_k \rangle|^2 \leq \#_{\mathcal{D}} \|\mathbf{x}\|^2 < \infty. \tag{B.43}$$

Therefore, any finite dictionary $\mathcal{D}$ in $\mathbb{H}^N$ is also a frame in $\mathbb{H}^N$.

Although a dictionary must have a lower frame bound A in equation (B.23), it does not need to have an upper frame bound B. In real vector spaces the dictionary that arise when continuous parameterized atoms are employed has an unbounded cardinality ($\#_{\mathcal{D}} = \aleph_1$), and the frame built by the atoms in such dictionaries will never have an upper frame bound. Since a dictionary has unit norm elements one has that

$$\sum_{k=1}^{\#_{\mathcal{D}}} \|\mathbf{g}_k\|^2 = \sum_{k=1}^{\#_{\mathcal{D}}} \sum_{i=1}^{N} g_k[i]^2 = \#_{\mathcal{D}}, \tag{B.44}$$

where $g_k[i]$ is the i^{th} sample of $\mathbf{g}_k$. Define the canonical basis of $\mathbb{R}^N$ by means of $\mathcal{B} = \{\mathbf{e}_0, \ldots, \mathbf{e}_{N-1}\}$, such that the values of the coordinates $n \in [0, \ldots, n-1]$ of the $\mathbf{e}_i$ are given by $e_i[n] = \delta[i-n]$, i.e. each $\mathbf{e}_i$ is a unit impulse at coordinate i. One can express each dictionary element using the canonical basis by means of

$$\mathbf{g}_k = \sum_{i=0}^{N-1} e_{i,k} \mathbf{e}_i, \quad e_{i,k} = \langle \mathbf{g}_k, \mathbf{e}_i \rangle = g_k[i]. \tag{B.45}$$

Then, the frame definition (see equation (B.23)) gives

$$A\|\mathbf{x}\|^2 \leq \sum_{k=1}^{\#_{\mathcal{D}}} |\langle \mathbf{x}, \sum_{i=0}^{N-1} e_{i,k} \mathbf{e}_i \rangle|^2 \leq B\|\mathbf{x}\|^2, \tag{B.46}$$

$$A\|\mathbf{x}\|^2 \leq \sum_{k=1}^{\#_{\mathcal{D}}} |\sum_{i=0}^{N-1} e_{i,k} \langle \mathbf{x}, \mathbf{e}_i \rangle|^2 \leq B\|\mathbf{x}\|^2, \tag{B.47}$$

$$A\|\mathbf{x}\|^2 \leq \sum_{k=1}^{\#_{\mathcal{D}}} |\sum_{i=0}^{N-1} e_{i,k} x[i]|^2 \leq B\|\mathbf{x}\|^2. \tag{B.48}$$

Supposing that $\mathbf{x}$ is non-zero just at coordinate i one has that

$$A\|\mathbf{x}\|^2 \leq \sum_{k=1}^{\#_{\mathcal{D}}} |x[i]e_{i,k}|^2 \leq B\|\mathbf{x}\|^2, \tag{B.49}$$

$$A\|\mathbf{x}\|^2 \leq \sum_{k=1}^{\#_{\mathcal{D}}} x[i]^2 e_{i,k}^2 \leq B\|\mathbf{x}\|^2, \tag{B.50}$$

$$A\|\mathbf{x}\|^2 \leq \|\mathbf{x}\|^2 \sum_{k=1}^{\#_{\mathcal{D}}} e_{i,k}^2 \leq B\|\mathbf{x}\|^2. \tag{B.51}$$

Therefore, for the upper frame bound B one must have that

$$\sum_{k=1}^{\#_{\mathcal{D}}} e_{i,k}^2 = \sum_{k=1}^{\#_{\mathcal{D}}} g_k[i]^2 \leq B. \tag{B.52}$$

Equation (B.44) is valid for any frame. Therefore, when $\#_{\mathcal{D}}$ is unbounded it is necessary for the summation $\sum_{k=1}^{\#_{\mathcal{D}}} g_k^2[i]$ to be unbounded for at least one coordinate i, leading to an unbounded equation (B.52). This result means that continuous parameter dictionaries in $\mathbb{R}^N$ are not frames of $\mathbb{R}^N$.

One way to generate finite cardinality dictionaries for adaptive signal decomposition algorithms is to consider a frame to be the dictionary. Consider that $\{\mathbf{g}\}_{k\in\mathcal{I}}$ is a frame of $\mathbb{H}^N$, such that $\|\mathbf{g}_k\| = 1,\ \forall k$. For any vector $\mathbf{x} \in \mathbb{H}^N$ there is at least one frame element such that $\langle \mathbf{x}, \mathbf{g}_k \rangle \neq 0$. This means that any signal will be approximated, by the greedy approach, using a frame as dictionary. Let $\theta(\mathbf{x}, \mathbf{g}_k)$ be the angle between $\mathbf{x}$ and $\mathbf{g}_k$, that is

$$\theta(\mathbf{x}, \mathbf{g}_k) = \arccos\left(\frac{\langle \mathbf{x}, \mathbf{g}_k \rangle}{\|\mathbf{x}\|}\right), \tag{B.53}$$

remind that $\|\mathbf{g}_k\| = 1$. Then, for any vector $\mathbf{x}$ one has that

$$A\|\mathbf{x}\|^2 \leq \sum_k \left| \|\mathbf{x}\|\|\mathbf{g}_k\| \cos\left(\theta(\mathbf{x}, \mathbf{g}_k)\right)\right|^2 \leq B\|\mathbf{x}\|^2 \tag{B.54}$$

$$A\|\mathbf{x}\|^2 \leq \|\mathbf{x}\|^2 \sum_k |\cos\left(\theta(\mathbf{x}, \mathbf{g}_k)\right)|^2 \leq B\|\mathbf{x}\|^2. \tag{B.55}$$

Therefore, one has that

$$A \leq \sum_k |\cos\left(\theta(\mathbf{x}, \mathbf{g}_k)\right)|^2 \leq B. \tag{B.56}$$

Larger $\sum_k |\cos\left(\theta(\mathbf{x}, \mathbf{g}_k)\right)|^2$ leads to concentrations of dictionary elements in the directions of the space that are similar to the direction of $\mathbf{x}$. The closer that A and

B are (the "tighter" the frame is) the more similar are the "densities" of the frame (the concentrations of frame elements) in all directions of $\mathbb{H}^N$. In general, tight frames do not provide the minimum $\Theta(\mathcal{D})$, however they indicate a reasonable approximation for all $\mathbf{x} \in \mathbb{H}^N$. For example, equiangular normalized tight frames [4, 12, 40] can achieve the lower bound on the dictionary coherence [99, 105] that is given by

$$\mu(\mathcal{D}) \leq \sqrt{\frac{\#_{\mathcal{D}} - N}{N(\#_{\mathcal{D}} - 1)}}. \tag{B.57}$$

B.6 Contributions of the Thesis

Adaptive decompositions have a large number of features that allow for efficient and coherent signal representations, pattern recognition, signal compression and other signal processing tasks. Because of that, adaptive decomposition algorithms are being used in a large range of applications. The key to accomplish the aforementioned tasks relies on both the design of the dictionary to be used and the algorithm used to obtain the representation. Appendix C presents coherent representations of signals from electric power plants. These representations are obtained by employing an adaptive signal decomposition algorithm. The main objective of a coherent decomposition is to provide a signal representation related to the physical phenomena involved in the production of the measured signal. The algorithm presented, based on the Matching Pursuits, represents a signal from its identified components. The dictionary employed is composed of damped sinusoids, in order to obtain signal components closely related to power systems phenomena. However, the use of a dictionary of damped sinusoids alone does not ensure that the decomposition will be meaningful in physical terms. To obtain a meaningful representation, a technique leading to efficient coherent damped-sinusoidal decompositions, that are closely related to the physical phenomena being observed, is developed. The effectiveness of the proposed method for compression of synthetic and natural signals is tested, obtaining high compression ratios along with high signal-to-noise ratio.

Appendix D studies the quantization of Matching Pursuit coefficients based on a statistical model for the angles in Matching Pursuit iterations. From the study of the angles between the residues and the selected atoms in Matching Pursuit iterations, it is conjectured that these angles can be modelled statistically as

approximately independent and identically distributed. This approximation permits the statistics of the angles in Matching Pursuits to be estimated from the statistics of the angles in the first decomposition step for a Gaussian source. The approximate statistical model of the angles presented is then employed to perform off-loop Lloyd-Max quantization of Matching Pursuit coefficients, where the quantization is performed outside the decomposition loop, for compression applications. The Lloyd-Max quantizer presented is compared to the state-of-the-art off-loop Matching Pursuit quantization scheme from [54]. Results show that the two quantizers have similar rate×distortion operational performance.

Appendix E studies the time frequency content of frames. It provides a novel sufficient condition for a family of elements to be a frame. The time-frequency content of frames is defined there using the Wigner-Ville distribution (WD), and it allows to characterize frames in the time-frequency domain. A theorem is provided, that shows that the summation of the WDs of the elements in a set of functions being greater than zero is sufficient for this set to be a frame. The condition is used to characterize the construction of frames of damped sinusoids. The main motivation for that arises from the close relation between damped sinusoids and the phenomena represented in actual signals, since damped sinusoids are solutions to ordinary differential equations. Then the analysis of frames constructed from translations and modulations of a symmetric signal is held. It is shown how to interlace such frames in both time and frequency in order to obtain new frames such that the new frame and its dual are more similar than the original frame is to its dual, thus obtaining tighter frames.

From the relations between frames and dictionaries one has the tendency to think that it may be effective to use normalized tight frames as dictionaries. The only known case of parameterizable normalized tight frames is the maximal oversampled WH frame in [113]. Nevertheless, this frame construction approach does not allow the control of the dictionary cardinality ($\#_{\mathcal{D}}$), as in this case $\#_{\mathcal{D}}$ will be unequivocally defined by the space dimension N and equal to N^2. The Weyl-Heisenberg frames interlacing approach presented is used to generate dictionaries in Appendix E. The resulting dictionaries are evaluated in terms of distortion, data rate incurred by their use and computational complexity when used in the

Matching Pursuit algorithm. The results show that the dictionaries constructed using the interlacing approach have good rate×distortion×complexity compromise when dictionaries composed by Weyl-Heisenberg frames are to be used.

Apêndice C

Efficient Coherent Representations of Power System Signals

Disturbance monitoring is increasingly being required in contemporaneous top-quality power plants. This monitoring permits the post-event analysis of disturbance signals. This analysis allows for the identification of patterns and characteristics of faults [95], in order to improve the knowledge of the system behavior and prevent future problems. The information is acquired by digitizing the signals corresponding to the voltage and/or current quantities with digital fault recorders at several points of the network. The acquired signals are stored for future transmission and analysis, in order to identify the different natural phenomena represented in the signal that took place in the course of the disturbance. In a broad sense, these techniques are called oscillography and can be divided into two categories [95]:

- Short-Time or Transient Oscillography – used for monitoring transient phenomena, protection systems and equipment malfunctions, as well as for locating faults in time and space and also for power quality assessment. Sampled versions of the voltage and/or the current waveforms at points of the network are used for these purposes;

- Long-Time Oscillography – monitoring of low frequency oscillations and slow transients, used to gather information about the dynamic behavior of the interconnected network. In this case, the power flow and the oscillations of the fundamental frequency of the system are of interest.

The number of points monitored by oscillographic systems is increasing rapidly because: **a**) The power system operation performance gets more critical as demand increases; **b**) At an interconnected system, with various players, it is necessary to establish precisely the causes of the disturbance as well as the parts responsible for the resulting effects. The storage and transmission of oscillographic signals may generate an information overload; even though storage cost is decreasing, the general tendency is to sample signals at higher rates and using longer windows. In addition, the number of signals acquired is also increasing. Thus, storage capacity and transmission bandwidth problems persist, demanding good compression schemes. In order to obtain high compression ratios, lossy compression must be used. However, the compression must be such that the loss of quality involved does not impair the signal analysis. Also, the information overload is a serious problem to disturbance analysis, as human experts (that perform the analysis) have in general difficulty to analyze a very large amount of data. This creates a demand for computational tools: **i**) That aid in the analysis of the phenomena; **ii**) That allow efficient transmission and storage of the information. Therefore, a technique that decomposes a signal in components that are coherent to power system phenomena modelling, would be welcome in automated analysis (for example using expert systems techniques [95]).

This appendix proposes a procedure aiming at efficient representation of power systems signals for short-time oscillography such that their analyses are not impaired. In order to achieve this goal it is used a damped sinusoid signal model that is related to the phenomena typically observed at power system plants. The different components of a signal, each one associated to a different phenomenon, are identified through an algorithm based on the Matching Pursuit (MP) algorithm, leading to an efficient representation of all relevant information. In this representation no important phenomenon is lost or distorted by the compression process, achieving high compression ratios allied to high signal to noise ratio (SNR).

The MP is capable of obtaining compact and efficient representations; the key to this lies in the choice of the dictionary $\mathcal{D}$, that should be coherent to the signal components. However, due to the greedy nature of MP, pre and post-echo artifacts may appear (see section C.2.2), along with inefficiencies with respect to phenomena identification (see sections C.3.2 and C.3.4). The dictionary is composed of damped

sinusoidal atoms which some parameters may assume values in a continuous range. In this work, a method is proposed to reduce the complexity of the search for $\sigma(n)$ in the case of damped sinusoid structures.

Two important aspects of signal representations obtained here need to be pointed out:

1. Are the structures used to represent the signal coherent? That is, are they related to the possible phenomena composing the signal? A coherent dictionary would be one in which the atoms $\mathbf{g}_k$ are good models for the phenomena inherent to a given signal, see subsection B.1.6. Section C.3 proposes the use of a dictionary of damped sinusoids for the representation of electric signals.

2. Are atoms being spent just to represent noise? Coherent representations retain only the significant atoms in the decomposition, those related to useful phenomena and not to noise. How to define which atoms correspond to coherent components and which ones correspond to noise? This idea is developed in subsection C.4.1.

Another important aspect concerns the signal compression scheme:

1. Is the representation consistent [62]? In [62] a consistent reconstruction is defined as: given a signal $\mathbf{x} \in \mathbb{X}$ and a function $f : \mathbb{X} \rightarrow \mathbb{Y},\ \mathbf{y} = f(\mathbf{x})$, if

$$f(\hat{\mathbf{x}}) = \mathbf{y} \tag{C.1}$$

   then $\hat{\mathbf{x}}$ is a consistent estimate of $\mathbf{x}$. Supposing the function f as an expert (human or system), a consistent reconstruction is obtained if the characteristics observed by an expert in the original signal and in its quantized representation are the same. Thus, consistency is related to how precisely the parameters representing the phenomena should be measured and/or quantized. Consistency of the electrical signal decomposition developed here, in the above sense, is addressed in sections C.3.7 and C.5, through real-world (natural) examples.

The power system signal model employed was proposed in [94]. A first version of the decomposition algorithm presented was implemented in [80]. However, the algorithm presented, in section C.4, is novel, as well as the qualitative and quantitative results of sections C.3.7 and C.5.

C.1 A Model for Signals From Power Systems

Regardless of the quantities measured, the net result of oscillography is the study of phenomena evolution in time. These phenomena are represented, in general, as sinusoidal oscillations of increasing or decreasing amplitudes, and are highly influenced by circuit switching, as well as non-linear power systems components and loads. In order to analyze and compress signals from power systems it is important to use a model capable of precisely representing the components that may compose those signals. Xu [110] discusses common phenomena in power systems:

- Harmonics – low frequency phenomena ranging from the system fundamental frequency (50/60 Hz) to 3000 Hz. Their main sources are semiconductor apparatuses (power electronic devices), arc furnaces, transformers (due to their non-linear flux-current characteristics), rotational machines, and aggregate loads (a group of loads treated as a single component).

- Transients – observed as impulses or high frequency oscillations superimposed to the voltages or currents of fundamental frequency (60/50 Hz) and also exponential DC and modulated components. The more common sources of transients are lightnings, transmission line and equipment faults, as well as switching operations, although, transients are not restricted to these sources. Their frequency range may span up to hundreds of thousands of Hz, although the measurement system (and the power line) usually filters components above few thousands of Hz, yielding frequency limited data.

- Sags and Swells – increments or decrements, respectively, in the RMS voltage of duration from half cycle to 1 minute (approximately).

In oscillography it is interesting to be capable of modelling and identifying these phenomena. A comparison between several methods of modelling and identification can be seen in [6]. Some of them are: Fourier filtering [97]; Prony analysis [79, 101]; auto regressive moving average models [6]; state space tracking methods [6]; Wavelets [19, 56, 79, 91]. These methods are normally employed along with expert systems, neural networks or any other artificial intelligence approach.

In a very simplistic way, power systems can be considered as built from transmission lines, sources and loads. Transmission lines can be modelled by a series

of distributed inductances, capacitances and resistances. In these circuits, currents or voltages follow closely the solutions of differential equations. To this model should be added the possibility to represent discontinuities due to circuit switching caused by operative maneuvers and by the protection system. The proposal here is to use a model for power system signals based on the solutions of linear differential equations that also considers these discontinuities. This model is given by an appropriate set of concatenated and superimposed damped sinusoids, each one having a well defined region of support. Therefore, the employed model is based on damped sinusoids [82, 94] and is given by

$$f(t) = \sum_{q=0}^{Q-1} \gamma_q \cos\left(2\pi f_q t + \phi_q\right) e^{-\rho_q (t - t_{s_q})} [u(t - t_{s_q}) - u(t - t_{e_q})], \tag{C.2}$$

where each component is a damped sinusoid represented by a 6-tuple (γ_q, f_q, ρ_q, ϕ_q, t_{s_q}, t_{e_q}). In the 6-tuple γ_q is the component amplitude, f_q is its frequency, ρ_q is the damping factor of the exponential, ϕ_q is the phase of the component, and t_{s_q} and t_{e_q} are respectively the component starting and ending times.

The well known Prony method [6,79,101] used in the analysis of power system signals (as well as in numerous other applications) employs a similar model. The Prony method obtains a representation of the signal as

$$f(t) = \sum_{q=0}^{Q-1} \gamma_q \cos\left(2\pi f_q t + \phi_q\right) e^{-\rho_q t}. \tag{C.3}$$

A drawback of the Prony method is that it does not take into account the discontinuities due to circuit switching. This is so because it does not consider that the damped sinusoids representing different components can start at different instants. Therefore, the proposed model can be seen as a generalization of the Prony model by adding a time localization feature. The algorithm employed here also avoids the numerical instabilities inherent to the Prony method, that might appear due to the matrix inversions required to compute the damping factors in it.

In this work, the analysis of power system signals will be restricted to frequencies which are integer multiples of a fundamental frequency F ($f_q = K_q F$). Such harmonic analysis is appropriate for pseudo-periodic signals such as power systems signals in question and, for example, audio signals. It also simplifies the model, and

therefore the algorithm used to obtain it, leading to

$$f(t) = \sum_{q=0}^{Q-1} \gamma_q \cos(2\pi k_q F t + \phi_q) e^{-\rho_q (t - t_{s_q})} [u(t - t_{s_q}) - u(t - t_{e_q})]. \tag{C.4}$$

Now, the 6-tuple is given by (γ_q, k_q, ρ_q, ϕ_q, t_{s_q}, t_{e_q}), where $k_q = f_q/F$ is an integer. This model can effectively describe the relevant phenomena in electric power systems signals that are worth analyzing. Therefore, the model of equation (C.4) is able to represent signals compactly and accurately in terms of the 6-tuples, and it would be a powerful tool for analysis.

Using a dictionary of unit norm vectors, the structures of equation (C.4) are uniquely determined by the 5-tuples ($k_q F$, ρ_q, ϕ_q, t_s, t_e). Observe that this dictionary contains a basis of impulses in time (5-tuples (0, 0, 0, t, t)), as well as a basis of sinusoids, that is impulses in frequency (5-tuples ($k_q F$, 0, 0, $-\infty$, ∞)), being thus overcomplete. However, in order to implement the MP algorithm, the parameters of the 5-tuple ($k_q F$, ρ_q, ϕ_q, t_s, t_e) must be quantized. One dictionary for which parameters quantization processes are well known and have been extensively used is the Gabor dictionary [27,84,85]. Then, at each step, it was decided to adopt the strategy of decomposing the signal using the MP with the Gabor dictionary, as an intermediate solution, and then, from the parameters of the chosen Gabor atom, searching for the damped sinusoid that best matches the signal at that step. This approach gives an MP based decomposition with good performance and feasible computational complexity. In section C.2, the MP using the Gabor dictionary is described. In section C.3, we describe how the damped sinusoid that best matches the residue is searched, at each step, starting from a Gabor atom.

A block diagram of the whole algorithm is shown in Figure C.1 (where the sections of this appendix in which each topic is treated are also indicated). The process breaks the problem of finding the parameters in a 5 dimensional space ($k_q F$, ρ_q, ϕ_q, t_s, t_e) into several smaller-dimensional problems, each demanding less computational effort, and thereby reduces the overall complexity. The first part of the algorithm is the decomposition of the current residue into a dictionary of Gabor atoms. The result of the decomposition at each step is used to match the current residue to a damped sinusoid instead of the Gabor atom. Then, for the damped sinusoid a frequency quantization is applied and so is an heuristic to select between

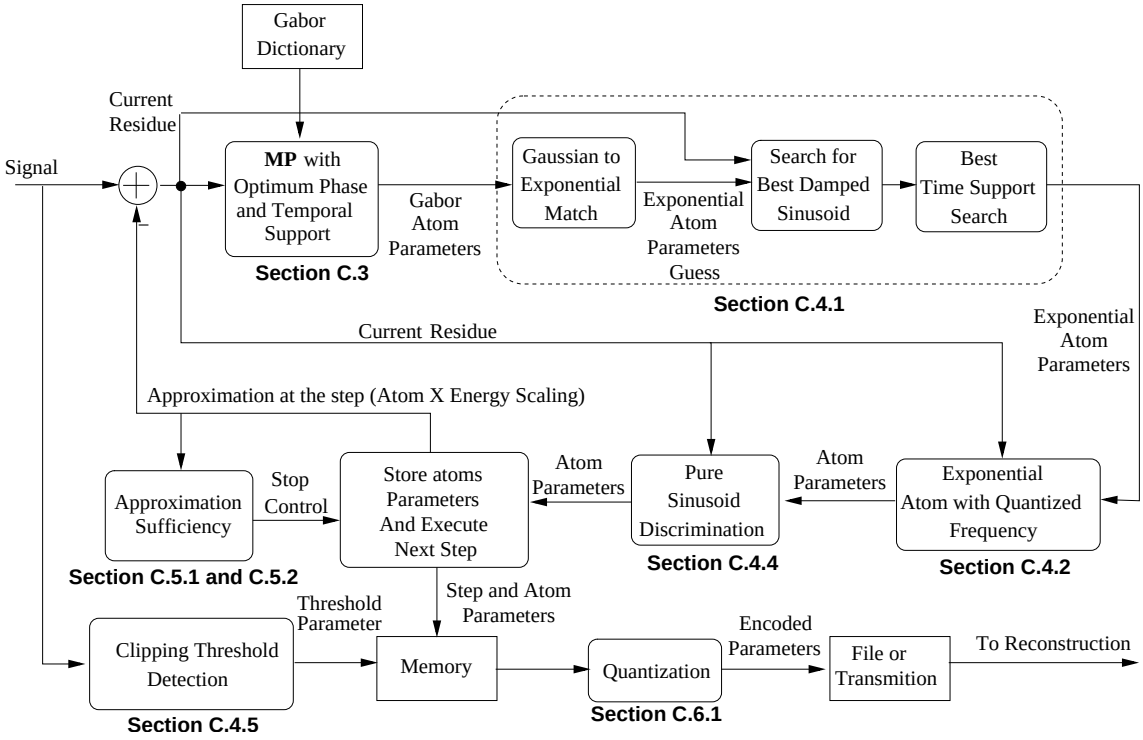


Figure C.1: Block diagram of the decomposition algorithm.

damped or pure sinusoids. Then the parameters of the resulting damped sinusoid structure are stored. The approximation sufficiency procedure is implemented to estimate whether the approximation already obtained is good enough or the decomposition must continue. The parameters are quantized and encoded at the end of the decomposition process using different quantizers for each parameter.

Note that the model in equation (C.4) by no means compactly represents "all" possible phenomena in power systems signals, as for example, inrush currents or inter-harmonic oscillations (for $f_q \neq k_q F$). However, since the proposed dictionary is complete, then any signal can be represented as a linear combination of the atoms, although this representation may not be compact.

C.2 Matching Pursuit with the Gabor Dictionary

To obtain signals decompositions according to the model proposed the Matching Pursuits algorithm (discussed in subsection B.2.2) was employed. The first step of the decomposition algorithm carries out the MP using a Gabor dictionary. The Gabor dictionary is generated by time shifting, scaling and modulation of a

Gaussian window $g(t) = 2^{\frac{1}{4}} e^{-\pi t^2}$ [84]. Each Gabor Dictionary atom is given by

$$g_\sigma(t) = \frac{K_\sigma}{\sqrt{s}} g\left(\frac{t-\tau}{s}\right) \cos\left(\xi t + \phi\right), \tag{C.5}$$

where $\sigma = (s, \tau, \xi, \phi)$ defines the atom, and each atom is a well localized time-frequency component [85]. The phase $\phi \in [0, \pi)$ and K_σ are such that $||g_\sigma(t)|| = 1$. For this $g(t)$, the parameters defining the atoms can be sampled obtaining a finite dictionary with discrete parameters which is complete, and indeed, capable of representing any signal [85]. The sampling of σ gives [85] $\sigma_s = [2^j, p2^j, k\pi 2^{1-j}, \phi]$, $j, p, k \in \mathbb{Z}$. Thus, σ_s can be represented uniquely by $\sigma_s = [j, p, k, \phi] \in \mathbb{Z}^3 \times \mathbb{R}$ (note that the phase remains continuous). The index j defines the atom scaling, p defines the time shift of the atom, and k the atom modulation.

Gabor Dictionary in $\mathbb{R}^N$ In order to implement the algorithm one needs to generate discrete atoms $\mathbf{g}_{\sigma_s}$. Being N the dimension of the signal space, define $L = \log_2(N)$, L is the number of scales in the dictionary, and the parameters ranges are given by $j \in [0, L],\ p \in [0, N2^{-j}),\ k \in [0, 2^j)$ [85]. Thus, each discrete atom $\mathbf{g}_{\sigma_s}$ has its samples $g_{\sigma_s}[m]$ given by

$$\begin{aligned} g_{\sigma_s}[m] &= g_j[m - p2^j] \cos\left[mk\pi 2^{1-j} + \phi\right], \\ g_j[m] &= \begin{cases} \delta[m], & \text{if } j = 0 \\ K_{\sigma_s} g\left[\frac{m}{2^j}\right], & \text{if } j \in (0, L), \\ \frac{1}{\sqrt{N}}, & \text{if } j = L \end{cases} \qquad \text{with } g[m] = 2^{\frac{1}{4}} e^{-\pi m^2}, \end{aligned} \tag{C.6}$$

where K_{σ_s} is equivalent to K_σ. Note that the total number of elements in the dictionary is $N \log_2(N)$; also, the optimum phase ϕ can be derived analytically from the triplet (s, τ, ξ).

Optimum Phase Computation Once the parameters $(s, \tau, \xi) = \sigma$ are found using the MP, the optimum phase for the atom can be computed [47,59]. The result presented in [47] is for continuous time atoms, but can be generalized to discrete ones. Any unit norm atom depending on a set of parameters $\sigma = (\eta, \xi, \phi)$, is given by

$$g_\sigma(t) = \frac{g_\eta(t) \cos\left(\xi t + \phi\right)}{||g_\eta(t) \cos\left(\xi t + \phi\right)||}, \tag{C.7}$$

where $g_\eta(t)$ can be any real function that depends on the set η (for Gabor atoms $\eta = (s, \tau)$, see equation (C.5) and for the damped sinusoids $\eta = (\rho, t_s, t_e)$, see equation (C.2). The function $g_\sigma(t)$ can be regarded as the real part of a complex function $G_\sigma(t) = g_\eta(t)e^{j\xi t+\phi}$ normalized to have unit energy. Defining $P_\sigma(t) = \mathrm{Re}\{G_\sigma(t)\}$ and $Q_\sigma(t) = \mathrm{Im}\{G_\sigma(t)\}$, at the n-th step of the MP the optimum phase $\phi_o \in [0, 2\pi)$ (obtaining always positive inner products), for the atom, is given by:

1. if $\xi \neq 0$ and $a \neq 0$, and $\dfrac{\langle r^{n-1}_{x(t)}, P_{\sigma(n)}(t)\rangle}{||P_{\sigma(n)}(t)||}$ $\begin{cases} > 0 \;\text{ then }\; \phi_o = \arctan\left(-\dfrac{b}{a}\right) \\ < 0 \;\text{ then }\; \phi_o = \arctan\left(-\dfrac{b}{a}\right) + \pi \end{cases}$;

2. if $\xi = 0$, and $-\dfrac{\langle r^{n-1}_{x(t)}, Q_{\sigma(n)}(t)\rangle}{||Q_{\sigma(n)}(t)||}$ $\begin{cases} > 0 \text{ then } \phi_o = 0 \\ < 0 \text{ then } \phi_o = \pi \end{cases}$;

3. if $a = 0$, and $\dfrac{\langle r^{n-1}_{x(t)}, P_{\sigma(n)}(t)\rangle a + \langle r^{n-1}_{x(t)}, Q_{\sigma(n)}(t)\rangle b}{||aP_{\sigma(n)}(t) + bQ_{\sigma(n)}(t)||}$ $\begin{cases} > 0 \text{ then } \phi_o = \frac{\pi}{2} \\ < 0 \text{ then } \phi_o = \frac{3\pi}{2} \end{cases}$.

where:

$$\begin{aligned} a &= \langle r^{n-1}_{x(t)}, P_{\sigma(n)}(t)\rangle \|Q_{\sigma(n)}(t)\|^2 - \langle r^{n-1}_{x(t)}, Q_{\sigma(n)}(t)\rangle \langle P_{\sigma(n)}(t), Q_{\sigma(n)}(t)\rangle \\ b &= \langle r^{n-1}_{x(t)}, Q_{\sigma(n)}(t)\rangle \|P_{\sigma(n)}(t)\|^2 - \langle r^{n-1}_{x(t)}, P_{\sigma(n)}(t)\rangle \langle P_{\sigma(n)}(t), Q_{\sigma(n)}(t)\rangle \end{aligned} \tag{C.8}$$

Allowing $\phi_o \in [0, 2\pi)$, the phase can be adjusted to provide always a positive inner product between the atom and the residue, unlike the approach in [47].

C.2.1 Continuous Parameters

From the discrete parameters of the atom found by the MP iteration, a local search using a pseudo-Newton method [92] can be applied to find the continuous parameters set σ that once used to generate $\mathbf{g}_\sigma$ produces the best match to the residue being decomposed. This achieves an MP using a Gabor dictionary with densely sampled parameters. For obtaining the continuous-parameters Gabor atom, there is a problem related to the application of the Newton method: Since there is no closed form for the inner product $\langle \mathbf{x}, \mathbf{g}_\sigma \rangle$, inner products ought to be calculated for each new update of the atom parameters. Directed by this observation, the strategy employed for the pseudo-Newton method was:

I. Independently increment each parameter of the atom by half of its value;

II. If the inner product increases then the parameter is updated, otherwise the increment is negated and halved (multiplied by -0.5) for the next iteration;

III. At each iteration the three parameters $[s, \tau, \xi]$ are tested, sequentially for updates;

IV. The iterations take place until the inner product increase is less then 1%, or the parameters increment is less then 10% of their actual values.

Note that when implementing an MP with a dictionary of continuous parameters, the fast MP algorithm [85] (based on correlation updates) is not applicable. We have implemented and tested both the discrete and "continuous-parameters" versions of the MP. The coefficients found were around 10% larger for the continuous case, providing thus a better match between the structure and the residue.

C.2.2 Pre-echo and Post-echo Reduction

When using the MP to decompose signals, undesirable artifacts of pre-echo and post-echo in general arise. As an example consider the decomposition of the synthetic signal *S1*, that is just the atom corresponding to the 6-tuple (1.000, 8, 0.080, -90, 0.0312, 0.1059), with fundamental frequency $F = 60$Hz and sampling frequency $F_s = 1200$Hz, see equation (C.4). The decomposition of *S1* using the discrete-parameters Gabor Dictionary with optimal phase, is shown at Figure C.2, with different number of steps (2, 4, and 8), where pre-echo and post-echo artifacts can be observed.

One way to suppress those artifacts is to include in the dictionary atoms having all possible temporal supports. This is equivalent to box-windowing the atoms by means of

$$g_{\sigma_l}(t) = K_l \frac{1}{\sqrt{s}} g\left(\frac{t-\tau}{s}\right) \cos(\xi t + \phi)[u(t-t_s) - u(t-t_e)], \tag{C.9}$$

where $g(t) = 2^{\frac{1}{4}} e^{-\pi t^2}$ and $u(t)$ is the step function. Defining $\sigma_l = (s, \tau, \xi, \phi, t_s, t_e)$, with $t_s < t_e$, where t_s is the starting time of the atom and t_e is its ending time, K_l is chosen such that $\|g_{\sigma_l}(t)\| = 1$. Finding t_s and t_e that minimize the error norm in the

support region of the atom would cope with the pre-echo and post-echo artifacts. Since this work restricts the attention to digitized signals, the error norm in the support region, at step n, is given by

$$e_{[m_s,m_e]} = ||(\mathbf{r}_{\mathbf{x}}^{n-1} - \langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma_l(n)} \rangle \mathbf{g}_{\sigma_l(n)})[\mathbf{u}[m - m_s] - \mathbf{u}[m - m_e]]||, \tag{C.10}$$

where m_s is the signal sample corresponding to t_s and m_e corresponds to t_e (note that, initially, for all the atoms $m_s = 0$ and $m_e = N-1$ where N is the signal length). Actually, minimizing the norm of the error is the same as finding the support interval that gives the largest inner product (the atoms with new time support must be scaled to unity norm). Then, a new support interval $[m'_s, m'_e]$ is considered for the atom only if

$$\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma_{l'}(n)} \rangle \geq \langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma_l(n)} \rangle \quad \text{and} \quad m'_e - m'_s < m_e - m_s, \tag{C.11}$$

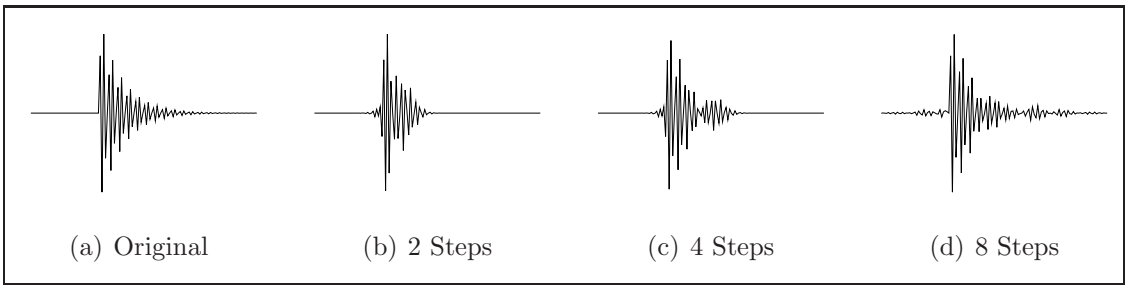
where $\sigma_{l'}(n) = (s, \tau, \xi, \phi, m'_s, m'_e)$. That is, the best time-support for the atom is the one yielding the largest inner product between the atom and the signal for the parameters subset (s, τ, ξ, ϕ).

Being $\sigma_l(n) = (s, \tau, \xi, \phi, m_s, m_e)$, it is easy to show that, if $\sigma_{l'}(n) = (s,\tau,$ $\xi,\phi,m_s,m_e - 1)$, then

$$\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma_{l'}(n)} \rangle = \frac{\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma_l(n)} \rangle - r_{\mathbf{x}}^{n-1}[m_e] g_{\sigma_l(n)}[m_e]}{\sqrt{1 - g^2_{\sigma_l(n)}[m_e]}}, \tag{C.12}$$

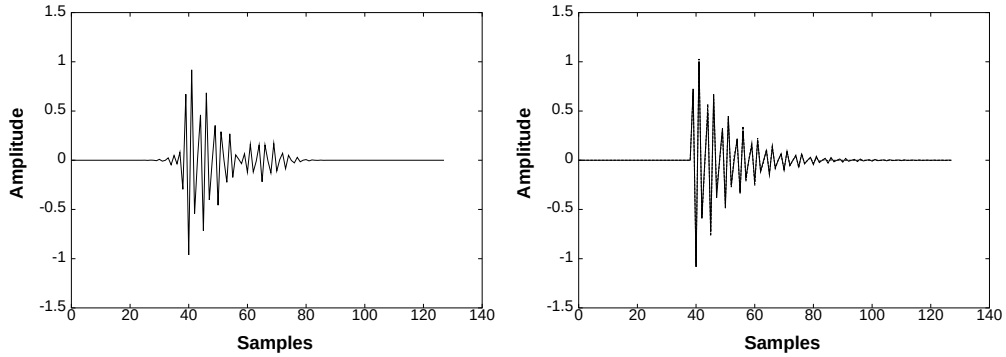
where $r_{\mathbf{x}}^{n-1}[m_e]$ is the m_e^{th} sample of $\mathbf{r}_{\mathbf{x}}^{n-1}$ and $g_{\sigma_l(n)}(m_e)$ the m_e^{th} sample of $\mathbf{g}_{\sigma_l(n)}$. Thus, if $\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma_l(n)} \rangle$ is known for a given value of m_e then the inner products for smaller values of m_e can be computed recursively, and, consequently, with low computational complexity. A similar result tells that an inner product for an atom starting at $m_s + 1$ can be computed from the one of the atom starting at m_s. Also, in order to further simplify this procedure, the search can be performed independently for m'_s and m'_e, leading to a fast algorithm for finding the best time support.

Figure C.3 shows the reconstruction of synthetic signal *S1* for a 4 step decomposition without and with temporal support search in the continuous-parameters Gabor dictionary. In this example, when temporal support search is applied, the reconstructed signal is visibly more similar to the original one, and the pre-echo and post-echo artifacts vanish; differences are noticed only after the 80^{th} sample,

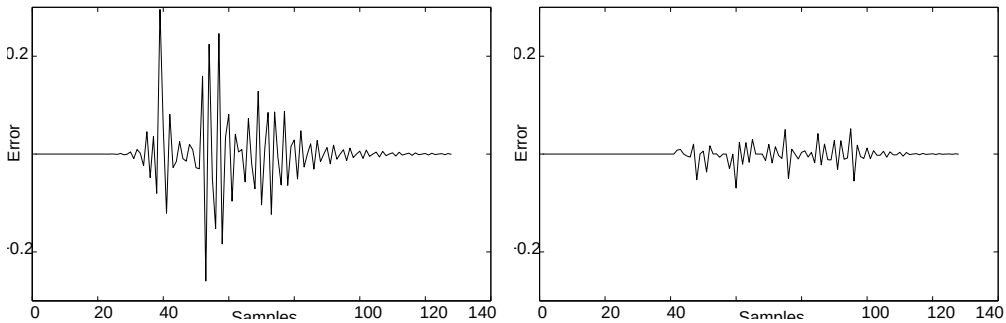


(a) Original (b) 2 Steps (c) 4 Steps (d) 8 Steps

Figure C.2: Reconstructed versions of the synthetic signal *S1* with the MP of continuous parameters and optimum phase with 2, 4 and 8 steps.



(a) Approximation without search (b) Approximation with search



(c) Error without search (d) Error with search

Figure C.3: Synthesis of synthetic signal *S1* (Figure C.2.(a)) after decomposition using 4 steps of the MP with densely sampled Gabor dictionary: (a) and (c) without, and (b) and (d) with the temporal support search of subsection C.2.2.

where the signal has little energy. When decomposing this signal with the Gabor dictionary of discrete parameters and optimum phase the inner product at the first step is 1.643, whereas using the continuous parameters it jumps to 2.0874 (a 27% increase). Further performing the search for the best time support the result becomes 2.145 (around 3% increase to the second and 30% increase to the former). Since the energy of the signal is equal to the sum of the energies of the inner products [84],

a larger inner product means a smaller reconstruction error. This shows that the closer visual similarity corresponds to a closer match in a mean-squared sense.

Unlike standard MP algorithms where the set of structures is static (discrete parameters and fixed dictionary), here, a dictionary of continuous parameters is used in the MP, allowing to fit the atoms to the signal, see subsection B.3.1. Also, the structures found have the best time support computed, in a fast fashion, eliminating pre and post-echo artifacts.

C.3 Damped Sinusoid Atomic Decompositions

In the proposed approach, an indirect search for the damped sinusoids was implemented. First the continuous parameters Gabor atom that maximizes the inner product is found. Then, the parameters of the Gabor atom are used to generate a guess for the best damped sinusoid. This guess is used as seed in a pseudo-Newton algorithm to find the parameters of the best damped sinusoid. The process is detailed in the sequel. Note that when such indirect damped sinusoid search is performed one can not use correlation updates (as in [85]), and therefore can not implement a fast MP algorithm because the dictionary being effectively used has continuous parameters, and thus infinite cardinality. Therefore, the computational complexity of the proposed method is increased when compared to the one of the classical MP; however, the good results obtained justify the increase in complexity.

C.3.1 Search for the Damped Sinusoids

The procedure to obtain the damped sinusoid that best represents the signal in a given step of the MP is as follows

1. Find the Gabor atom index $\sigma(n) = [s, \tau, \xi, \phi_o]$ such that $\sigma(n) = \arg\{\max_{\sigma \in S} |\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_\sigma \rangle|\}$ using the MP with optimum phase ϕ_o (see section C.2).

2. Obtain the initial guess for the damped sinusoid:

   (a) Find the half of the Gabor atom selected by the greedy search that has larger inner product with the signal:

- The right half (defined as the points to the right of the Gaussian window center τ) implies a decreasing exponential to be searched for;
- The left half (defined as the points to the left of the Gaussian window center τ) implies an increasing exponential to be searched for.

That is, the best half of the Gabor atom is the one that has larger inner product with the current residue.

(b) Having found the scale s, compute the initial guess for the damping factor ρ that best matches the half of the Gabor atom identified above. This initial value is given by $\rho = \sqrt{\pi/2s^3}$ for a decreasing exponential, and $-\sqrt{\pi/2s^3}$ for an increasing exponential. This value of ρ leads to the damped sinusoid having the same sample value of the Gaussian at its inflection point. As the signals have limited time support, this value needs to be corrected. This is done iteratively by a closed-form fast Newton algorithm.

(c) With the obtained ρ, τ and ξ as initial guesses, a pseudo-Newton method is applied to search for the atom

$$\mathbf{g}[m] = K_g e^{-\rho(m-m_s)} \cos\left[\xi m + \phi\right][\mathbf{u}[m-m_s] - \mathbf{u}[m-m_e]],\ m \in [0, \ldots, N-1] \tag{C.13}$$

where K_g is a normalization factor. As the modulated exponential can be increasing or decreasing ($\rho > 0$ or $\rho < 0$), strategies to search for the start and end samples of the damped sinusoid are needed. In the case of a decreasing exponential, it is assumed that $m_s = \tau$ and m_e equal to the signal size N. In the case of an increasing exponential, $m_s = 0$ and $m_e = \tau$. The pseudo-Newton algorithm is applied to obtain one of the 5-tuples $\sigma(n) = [\rho, \xi, \phi, \tau, N-1]$ for a damped sinusoids or $\sigma(n) = [\rho, \xi, \phi, 0, \tau]$ for an increasing sinusoid (τ can represent either the start or end time). The phase is obtained through the procedure of section C.2.

The pseudo-Newton method applied is similar to that presented for the Gabor atom (see section C.2) but for four parameters $[\rho, \xi, \phi, \tau]$.

3. Search for the best time support, as in subsection C.2.2, obtaining the 5-tuple $\sigma(n) = [\rho, \xi, \phi, m_s, m_e]$ that characterizes the atom. Note that:

   a) If $\rho > 0$ then a decreasing atom occurs and only m_e is searched;

   b) If $\rho < 0$ then the atom is an increasing one and the time support search must be done for m_s.

4. For the best time support found above, apply a new local search for ρ and ξ using the pseudo-Newton method, and compute ϕ.

5. With the 5-tuple $\sigma(n) = [\rho, \xi, \phi, m_s, m_e]$, the inner product of the corresponding atom with the current residue $\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma} \rangle$ and then the resulting residue $\mathbf{r}_{\mathbf{x}}^{n} = \mathbf{r}_{\mathbf{x}}^{n-1} - \langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma(n)} \rangle \mathbf{g}_{\sigma(n)}$ are computed , and the process is finished.

In the search process one ends up with damped sinusoids, as in equation (C.4), that can have any phase ϕ. These atoms are called dereferenced [59] as the atoms start time and phase are not related. The identification of damped sinusoids using this approach performed effectively. Illustrations of this effectiveness are given in section C.5, where results with natural signals are shown.

C.3.2 Frequency Quantization

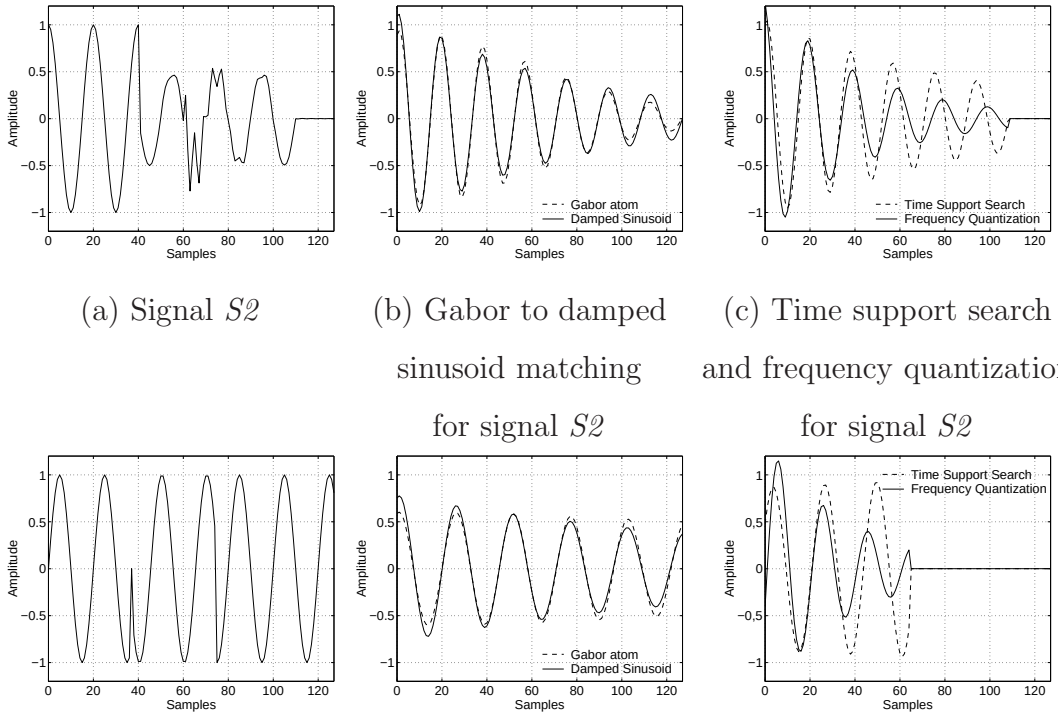
The algorithm presented so far represents a signal as a sum of damped sinusoids; however, unlike the model of equation (C.4) frequencies that are not multiple integers of a fundamental frequency are allowed. To obtain the signal model of equation (C.4) the frequencies of the structures must be quantized as integer multiples of the fundamental frequency. It is also important to take into account in the quantization process the sampling frequency F_s (the maximum possible frequency in the signal is $F_s/2$). The frequency quantization can then be performed as follows:

1. Compute the ratio between the sampling frequency and the fundamental, $r_f = \dfrac{F_s}{F}$. With this ratio design a quantizer for the frequency of the atom ξ using a linear quantizer of $b_{q(\xi)} = \lceil \log_2 r_f \rceil$ bits and step $\Delta_{q(\xi)} = \dfrac{\pi}{2^{b_{q(\xi)}}}$.

2. The quantized frequency for an atom is $\xi_q = \left\lfloor \dfrac{\xi + \frac{\Delta_{q(\xi)}}{2}}{\Delta_{q(\xi)}} \right\rfloor \times \Delta_{q(\xi)}$, where ξ is the original frequency of the atom.

3. After quantizing the frequency, a local optimization using the pseudo-Newton method is performed for the damping factor ρ.

Thus the frequency quantizer is designed a priori (points 1) and 2) above) and at each new atom the quantization is applied to its frequency.



(a) Signal *S2* (b) Gabor to damped sinusoid matching for signal *S2* (c) Time support search and frequency quantization for signal *S2*

(d) Signal *S3* (e) Gabor to damped sinusoid matching for signal *S3* (f) Time support search and frequency quantization for signal *S3*

Figure C.4: First step decomposition of synthetic signals *S2* and *S3* with time support search and frequency quantization.

Some examples are shown in order to point out problems regarding the identification of structures. Figure C.4 shows the behavior of the described process behavior at the first step of the decomposition of synthetic signals *S2* (Figure C.4.(a)) and *S3* (Figure C.4.(d)), whose parameters are given in Table C.1 (where m_{s_q} and m_{e_q} are the samples indexes of t_{s_q} and t_{e_q}, respectively, and q is the component number). These synthetic signals were chosen as they are representative of common phenomena in power systems [19, 56, 79, 94, 95, 101, 110]. Indeed, they indicate both the limitations of the MP algorithm and the effectiveness of the procedures presented for the desired application. In Figure C.4.(b) it can be observed the dam-

Table C.1: Generation of the synthetic signals *S2* and *S3* according to the model of equation (C.2).

Signal	F_s (Hz)	F (Hz)	q	γ_q	k_q	ϕ_q (Deg.)	ρ_q	t_{s_q} (Sec.)	t_{e_q} (Sec.)	m_{s_q}	m_{e_q}
S2	1200	60	1	1.000	1	0	0	0	0.0333	0	40
			2	0.500	1	90	0	0.0333	0.0917	40	110
			3	0.200	6	-90	0.100	0.0500	0.1059	60	127
			4	0.050	3	-67	0	0.0417	0.0833	50	100
S3	1200	60	1	1.000	1	-90	0	0.0625	0.1059	75	127
			2	1.000	1	-90	0	0	0.0308	0	37
			3	1.000	1	135	0	0.0308	0.0625	37	75

ped sinusoid obtained from the Gabor atom guess, that best matches the signal *S2* (Figure C.4.(a)). In Figure C.4.(c) the result of the time support search for the damped sinusoid atom and frequency quantization can be observed. The same is presented for signal *S3* in Figures C.4.(e) and C.4.(f) respectively. Note that prior to frequency quantization the frequency of the decomposing structure is different from the original one used to generate the signal component. The frequency quantization effectively corrects this error, that is caused by the greediness of the MP iteration together with the particularities of the dictionary used. Although this behavior may seem odd, the atom found prior to quantization is the one with higher inner product with the signal under decomposition. That is, this structure is the one that best represents the concatenation of the three structures in the signal. In simulations, it was observed that even if the frequency of the decomposing structure is forced to be equal to the one used in generating the atom, the inner product obtained is still smaller than the one found by the decomposition algorithm presented so far. In the present case, ideally, it is wanted to find the three structures composing the signal and not the one obtained by the decomposition algorithm presented so far.

From these examples, we see that the search for damped sinusoids and frequency quantization perform well. However, some false phenomena identification occur:

1. Signals formed by two or more sinusoids of same phases but different amplitu-

des are confused as one damped sinusoid as is the case of the decomposition of signal *S2* in Figure C.4.(b). Signal *S2* is formed by two sinusoids with same phases, one of them being corrupted by higher frequency components. However, the MP chooses damped sinusoid with full time support that is almost in phase with the higher energy sinusoid to represent the signal. Note that the phenomena switch (amplitude decrease) at the 40^{th} sample is not identified;

2. Sinusoids of same frequency and amplitude but different phase are identified as one sinusoid with phase that maximizes the inner product. The case of signal *S3*, with two phase discontinuities which are usually caused by switching, for which the 90^o phase change (around the 70^{th} sample) is located by the time support search but the smaller phase change (around the 40^{th} sample) is not;

The false phenomena identification that takes place when decomposing signals with MP-like algorithms have previously been described in [33]. There, the authors give an example of a signal composed by a sum of two dictionary structures, in which these structures are not found by the greedy decomposition. In subsection C.3.4 an algorithm is proposed which obtains better structure identification for the signal model presented here, addressing the limitations presented above.

C.3.3 Complexity Issues

Since a dictionary with densely sampled parameters is employed, the accurate establishment of the computational complexity is a hard task. This subsection tries to address the computational complexity of the employed decomposition method. Due to the nature of the problem at hand, frequency quantization, see equation (C.4), can be used. This contributes to decreasing the computational complexity, since it represents a significant decrease in size of the search space, and the sinusoids generated by these quantized frequencies can either be previously precomputed or generated on the fly using efficient algorithms as the FFT.

One important point affecting the computational complexity is that the recurrent formula for fast computation of the MP [84, 85]

$$\langle \mathbf{r}_{\mathbf{x}}^{n+1}, \mathbf{g}_{\sigma} \rangle = \langle \mathbf{r}_{\mathbf{x}}^{n}, \mathbf{g}_{\sigma} \rangle - \langle \mathbf{r}_{\mathbf{x}}^{n}, \mathbf{g}_{\sigma(n)} \rangle \langle \mathbf{g}_{\sigma(n)}, \mathbf{g}_{\sigma} \rangle \tag{C.14}$$

where $\sigma(n)$ is the parameter set of the atom selected by the MP at the n^{th} step and all the products $\langle \mathbf{g}_{\sigma_i}, \mathbf{g}_{\sigma_j} \rangle$ $(\sigma_i \neq \sigma_j)$, are not applicable to the present case. This is so because at the end of each MP step the signal is further decomposed using a continuous parameter atom. A dictionary with continuous parameters has infinite cardinality, and it is impossible to pre-store all the inner products between the dictionary atoms ($\langle \mathbf{g}_{\sigma(n-1)}, \mathbf{g}_{\sigma} \rangle$). The continuous parameters dictionary requires the use of a Newton method at each iteration, leading to more computational demands. Actually, the Newton method in a multivariate space is used twice, one for finding the best Gaussian atom, and the other for finding the best damped sinusoid. Note that, for signals of length N, if a dictionary with discrete-parameter damped sinusoids was used, its cardinality would be, approximately, $2^{b_{q(\xi)}} \times F \times N^2/2$, where $2^{b_{q(\xi)}}$ is the number of quantized frequency levels, F is the number of different possible damping factors ρ and $N^2/2$ refers to all the possible time-support intervals. On the other hand, the number of elements in the Gabor dictionary is given by $N \log_2(N)$. However, note that the use of a dictionary with a cardinality of $2^{b_{q(\xi)}} \times F \times N^2/2$ if not being prohibitive, is at least much more costly than using a dictionary of cardinality equal to $N \log_2(N)$; therefore, taking into consideration the fact that the continuous parameter dictionary has superior performance (see end of subsection C.2.2), the use of the Newton method is clearly a viable alternative for obtaining good decompositions with controlled computational burden.

C.3.4 Pure Sine Substitution

The structures found in the examples in subsection C.3.2, despite being the ones with largest inner product with the signals, do not exactly match the phenomena represented in the signal. This problem is inherent to any greedy decomposition algorithm [33]. Previous works have shown that a successful representation relies on a detailed analysis or on an appropriate distortion model in order to select the atoms [69, 70]. The proposal is similar to the criterion of subsection C.2.2, with extra conditions on the inner product. In order to obtain the selection of more appropriate component, either a damped or a pure (or plain) sinusoid will be tried as atoms and the one that gives smaller error in the region of support of the atom (this is closer to a shape similarity measure than the inner product) is selected. There are

two cases in which a pure sinusoid atom will be tested to model the signal instead of a damped one of same frequency. The pure sinusoids tried may have an inner product with the residue that is a fraction, smaller than one, of the inner product of the original damped sinusoid with the residue. However, in our case, we will accept the pure sinusoid to approximate the signal only if a shape similarity measure (the error per sample in a given region) is satisfied. This strategy is supported by the convergence of weak greedy algorithms [103].

Considering two sets of parameters $\sigma(n) = (k, \rho, \phi, m_s, m_e)$ (damped sinusoid) and $\sigma'(n) = (k, 0, \phi, m'_s, m'_e)$ (pure sinusoid), being $\mathbf{r}_{\mathbf{x}}^{n-1}$ the current residue, the two cases are:

i – If the inner product of the pure sinusoid with the signal is at least a fraction p_1 $(0 < p_1 < 1)$ of the inner product obtained with the damped sinusoid, then use the pure sinusoid. That is, if $|\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma'(n)} \rangle| \geq p_1 |\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma(n)} \rangle|$ then use $g_{\sigma'(n)}$ instead of $g_{\sigma(n)}$ to represent the current residue.

ii – If the inner product of the pure sinusoid with the signal is greater than a fraction p_2 $(0 < p_2 < p_1 < 1)$ of the inner product obtained with the original damped sinusoid, then it is tested if a pure sinusoid should be used to represent the signal instead of the damped sinusoid (see 1 and 2 below). More precisely: If case **i** above is not satisfied and $|\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma'(n)} \rangle| \geq p_2 |\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma(n)} \rangle|$ then it is verified if it is worth to use $g_{\sigma'(n)}$ instead of $g_{\sigma(n)}$ to represent the residue.

Obviously, p_1 is larger than p_2. Simulations, on natural and synthetic signals, have suggested for p_1 the value 0.99 and for p_2 the value 0.75. In brief, if case **i** is met, a pure sinusoid will be used; otherwise, a pure sinusoid will be tried to represent the current residue if case **ii** holds. Case **i** is straightforward since the use of a pure sinusoid will be only slightly worse, in terms of residue energy reduction, than the use of the damped one. In case **ii**, the use of the pure sinusoid is significantly worse than the damped one, but it is still valuable to verify the possibility of using a pure sinusoid instead of the damped one. This is done as follows:

1) From the frequency and the temporal limits (in samples) of the structure, a region where to search for a possible pure sinusoid to represent the signal is

computed by defining $\Delta_T = \frac{F_s}{F}$ (F is the fundamental frequency, see equation (C.4). The limits of the region where the pure sinusoid may be located are then given by $m'_s = m_s - \frac{\Delta_T}{2}$ and $m'_e = m_e + \frac{\Delta_T}{2}$. This range yields a possible region of support one cycle larger than the one of the original atom (half cycle before the start and half cycle after the end of the structure).

2) A pure sinusoid $g_{\sigma'(n)}$ ($\rho = 0$), of same frequency as the damped one $g_{\sigma(n)}$, will be a candidate to represent the signal at the current step, in the region of support $[m_s - \frac{\Delta_T}{2}, m_e + \frac{\Delta_T}{2}]$, if the two conditions below hold simultaneously:

   a) The error per sample in the region of support for the pure sinusoid is smaller than the error per sample for the damped sinusoid in the region of support of the pure sinusoid. That is

$$\frac{||(\mathbf{r}_{\mathbf{x}}^{n-1} - \langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma'(n)} \rangle \mathbf{g}_{\sigma'(n)})[\mathbf{u}[m - m'_s] - \mathbf{u}[m - m'_e]]||}{m'_e - m'_s - 1} < \frac{||(\mathbf{r}_{\mathbf{x}}^{n-1} - \langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma(n)} \rangle \mathbf{g}_{\sigma(n)})[\mathbf{u}[m - m'_s] - \mathbf{u}[m - m'_e]]||}{m'_e - m'_s - 1}, \tag{C.15}$$

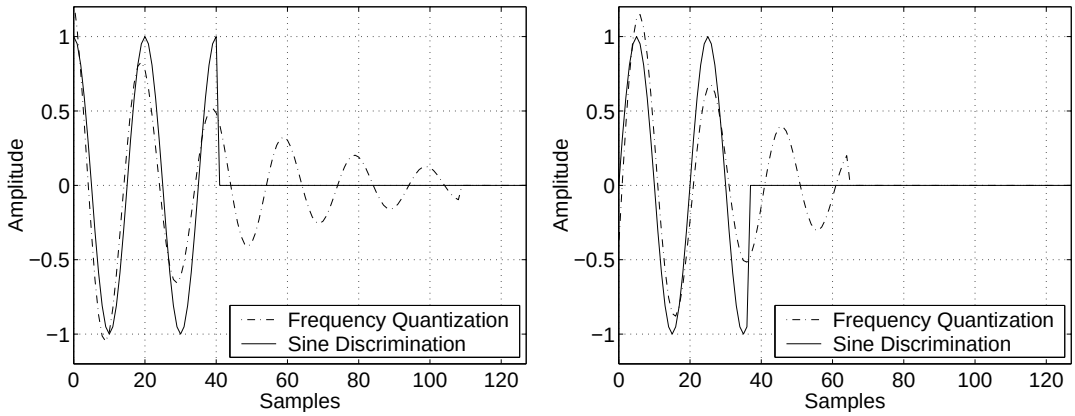
   where the primed variables correspond to the pure sinusoid and the non-primed ones to the damped sinusoid. Equation (C.15) compares the pure to the damped sinusoid in the region where the first is defined, providing a shape similarity measure in this region. As in the case of the pre-echo and post-echo suppression, described in section C.2.2, the expression in equation (C.15) can be computed recursively, leading to a fast algorithm.

   b) The error per sample for the pure sinusoid is smaller than or equal to half of the error per sample obtained with the damped sinusoid in its original region of support, that is

$$\frac{||(\mathbf{r}_{\mathbf{x}}^{n-1} - \langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma'(n)} \rangle \mathbf{g}_{\sigma'(n)})[\mathbf{u}[m - m'_s] - \mathbf{u}[m - m'_e]]||}{m'_e - m'_s - 1} \leq \frac{1}{2} \frac{||(\mathbf{r}_{\mathbf{x}}^{n-1} - \langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma(n)} \rangle \mathbf{g}_{\sigma(n)})[\mathbf{u}[m - m_s] - \mathbf{u}[m - m_e]]||}{m_e - m_s - 1}. \tag{C.16}$$

   This procedure compares the overall shape similarity of the two possible atoms in their respective regions of support. It is also imposed for the inner product of the current residue with the pure sinusoid to be at least 50% of what is obtained with the damped sinusoid, this constraint forces a reasonable and acceptable approximation ratio in the step.

Note that the search for the best time support can be performed independently for m_s and m_e. If $\rho > 0$, then it is expected that the start time of the atom is well defined since this is the region of the structure with largest energy. Thus, we first search for m_e applying the two conditions above and then we search for m_s using the same conditions. In the case $\rho < 0$, since m_e is expected to be well defined and then we first search for m_s and then for m_e.



Sine substitution for *S2* Sine substitution for *S3*

Figure C.5: Sine substitutions at the first step decomposition of synthetic signals *S2* and *S3*.

In Figure C.5 the pure sine substitution is verified when applied to the first step of the decompositions of signals *S2* and *S3*. Figure C.5 presents the sine discrimination together with the quantized frequency damped sinusoid identified so far (Figure C.4). Note that the pure sinusoidal structures found are coherent to the structures used to generate the atoms. Thus, the proposed sine discrimination procedure succeeded in correcting the erroneous decisions made by the MP algorithm. It was verified that this procedure does not impact the cases where the damping is representative of the phenomenon, as will be seen in examples to be shown.

C.3.5 Clipping Threshold

Power system oscillographic signals can be clipped mainly due to currents that surpass the range of the measurement device. However, this situation should not happen when the real-world monitoring takes place, i.e., ratings of measuring equipment should not be surpassed.

To model the effects of clipping, a clipping threshold, that is given by the maximum absolute value of the samples of $\mathbf{x}$, $T_c = \max_m |x[m]|$ for $m \in [0 \dots N-1]$, is identified. After the reconstruction of the signal from the structures, the signal samples are clipped if their absolute values are larger than the threshold, that is,

$$\hat{x}[m] = \begin{cases} \hat{x}[m], & \text{if } |\hat{x}[m]| \leq T_c, \\ \text{sign}(\hat{x}[m])T_c, & \text{if } |\hat{x}[m]| > T_c \end{cases} \tag{C.17}$$

where $\hat{x}[m]$ is the mth sample of the reconstructed signal $\hat{\mathbf{x}}$ and sign($\hat{x}[m]$) is the polarity of the m-th sample of $\hat{\mathbf{x}}$ (+ or -). Despite its simplicity, this procedure drastically improves the results in terms of signal to noise ratio for clipped signals and does not alter the results obtained for signals that are not clipped.

An example of the effectiveness of this procedure is depicted at Figure C.6 that shows the clipping of signal *S4* at 70% of the maximum sample value of the original signal. In this case all the 3 structures that compose *S4* are well identified, see Table C.2 for the generating structures, giving $(k, \rho, \phi, m_s, m_e) = (1, 0, 0, 0, 30)$ for the first structure (the other two structures are not relevant in this case as they are not clipped). The reconstructed signals can be evaluated also in using the SNR defined as

$$\text{SNR} = 10 \log_{10} \frac{\|\mathbf{x}\|^2}{\|\mathbf{x} - \hat{\mathbf{x}}_q\|^2} \text{ dB}, \tag{C.18}$$

where $\mathbf{x}$ is the original signal while $\hat{\mathbf{x}}_q$ is the reconstructed signal using q structures.

The signal *S4* clipped at 70% of the maximum sample value of the original signal is reconstructed with an SNR of 19.43 dB. When clipping is more restrictive, the identification of the parameters becomes more difficult and the suggested method does not perform as well as in the previous case. For example, clipping signal *S1*

Table C.2: Generation of the synthetic signal *S4* according to the model of equation (C.2).

Signal	F_s (Hz)	F (Hz)	q	γ_q	k_q	ϕ_q (Deg.)	ρ_q	t_{s_q} (Sec.)	t_{e_q} (Sec.)	m_{s_q}	m_{e_q}
S4	1500	50	1	1.000	1	0	0	0	0.0200	0	30
			2	0.250	2	90	0.1	0.0200	0.0533	30	80
			3	0.100	6	68	-0.03	0.0447	0.0847	67	127

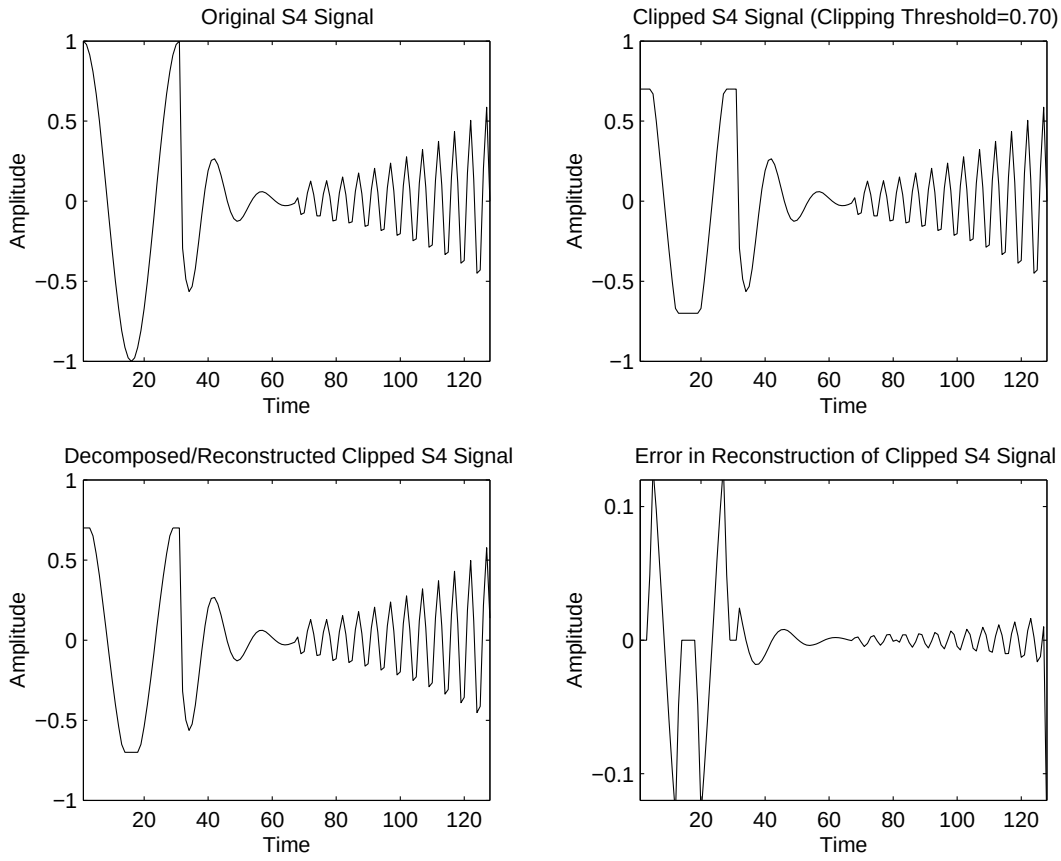


Figure C.6: Clipping of Signal *S4* and its decomposition.

with a threshold of 0.5 and then decomposing the clipped signal using, for example, 1, 2, 3, 8 and 20 structures, give reconstructed signals with SNR of 17.95, 20.54, 22.84, 36.26 and 45.03 dB, respectively. This example, with an 8 step decomposition, is depicted in Figure C.7. Just for comparison, the decomposition of the original signal *S1* is accomplished with just one structure with an SNR of 86.45. Note that despite the longer and worse decomposition, since the dictionary used is complete, the convergence of the decomposition is guaranteed, and it still gives an effective way to decompose clipped signals, as can be seen from the increase of the SNR as a function of the number of steps.

C.3.6 Results

From the details explained above the full decomposition algorithm (see the block diagram in Figure C.1) can be summarized as:

1. Identify the clipping threshold for the signal, section C.3.5.

2. Apply one step of the decomposition:

   (a) Apply the greedy iteration with the Gabor Dictionary;

   (b) From the parameters of the Gabor atom obtained in (a), search for a damped sinusoid as explained in section C.3.1;

   (c) Quantize the structure frequency, as in section C.3.2;

   (d) Apply the sine discrimination heuristic, as in section C.3.4.;

3. Repeat (2) iteratively until some stop criterion such as a fixed number of steps or a prescribed error is met (a better criterion to halt the decomposition process will be presented in section C.4).

4. At the end of these procedures we have a decomposition of the signal in terms of damped sinusoids. A reconstructed signal can be obtained from these structures. Then, we apply the clipping threshold procedure to it, as in subsection C.3.5.

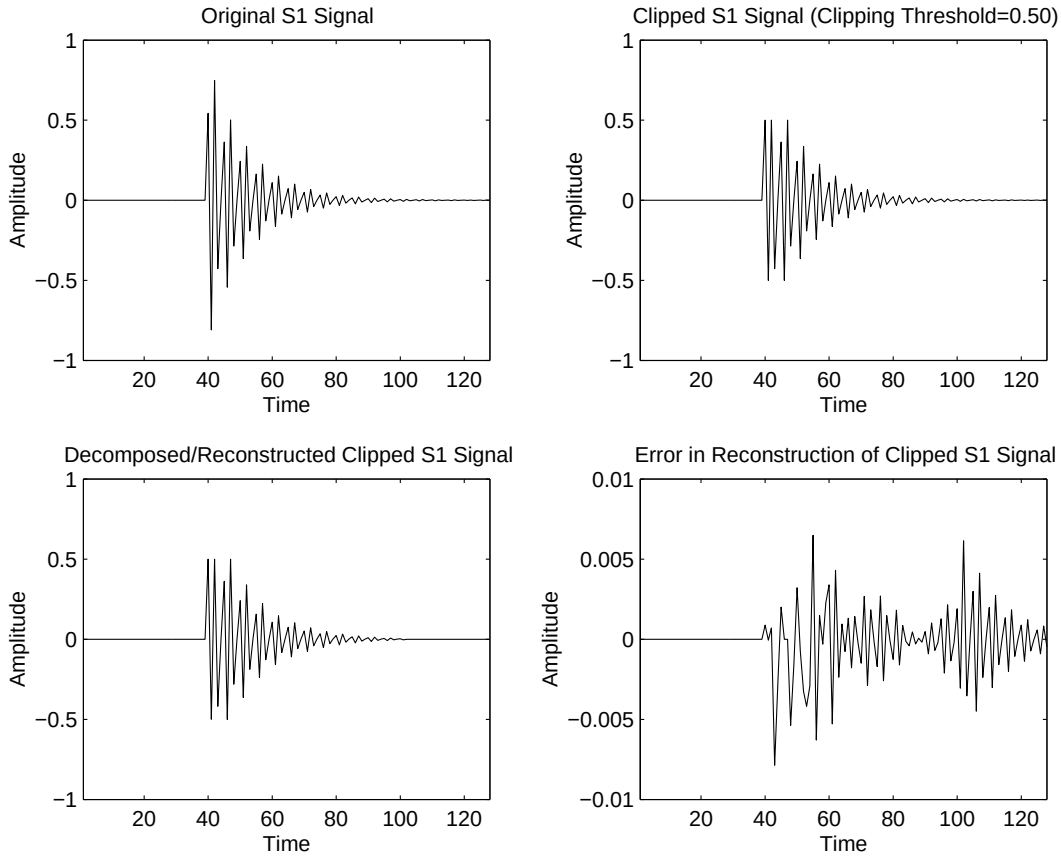


Figure C.7: Clipping of Signal *S1* and its decomposition.

Using the procedures described, the atoms found are physically meaningful (refer to Figures 4-6). This can be confirmed by the decomposition of signal *S4* (Figure C.8.(a)). The structures identified for that signal, using a three step decomposition, are presented in Figure C.8.(b), and the reconstructed signal in Figure C.8.(c). In Table C.3 the parameters of the structures extracted by the decomposition algorithm for signals *S3* and *S4* are shown together with the parameters used to generate them (repeated here for reader convenience). Signal *S4* is reconstructed with an SNR of 29.81 dB and signal *S3* with an SNR 88.58 dB (the SNR is defined in section C.5, equation (C.18)), implying that the estimation of the structure parameters is very good for these signals. Despite the good results obtained, a convincing validation of the proposed method can only be performed by decomposing natural signals [67] held in section C.5. The results shown just aim to verify the behavior of the proposed method with completely known signals.

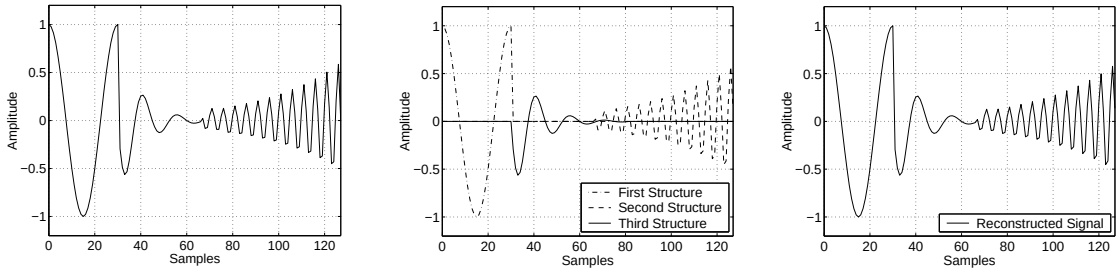
Table C.3: Decomposed structures parameters of signals *S3* and *S4*.

Signal	Atoms Parameters for *S3*					Atoms Parameters for *S4*				
	ξ (Hz)	ϕ	ρ	n_s	n_e	ξ (Hz)	ϕ	ρ	n_s	n_e
Decomposition	60.00	270.00	0.00	0	36	50.00	0.00	0.00	0	30
	60.00	171.00	0.00	38	74	300.00	292.67	-0.029423	67	127
	60.00	270.00	0.00	75	127	100.00	90.00	0.100022	31	80
Generation	60.00	-90.00	0.00	0	37	50.00	0.00	0.00	0	30
	60.00	135.00	0.00	37	75	300.00	68.00	-0.03	67	127
	60.00	-90.00	0.00	75	127	100.00	90.00	0.1	30	80

C.3.7 Application to the Filtering of DC Components

This subsection shows a study on the capability of the MP for filtering the exponential decay that sometimes appears in current quantities after disturbances, commonly known as the "DC Component" [97]. The exponential decay ("DC Component") appears added to the sinusoid and can be modelled as

$$Ae^{-\rho t}[u(t-t_s)-u(t-t_e)]+B\sin(2\pi ft+\phi) \tag{C.19}$$



(a) Signal *S4* (b) Structures identified for signal *S4* (c) Reconstructed signal *S4*

Figure C.8: Three steps decomposition of synthetic signal *S4*.

where t_s and t_e are the start and end times of the phenomenon (for simplicity the start and end times of the sinusoidal component were not presented) and ρ expresses the exponential decay constant. Observe that as equation (C.19) is a particular case of equation (C.4), it is expected for the decomposition developed to be capable of extracting/identifying the "DC Component". Therefore, once the signal is decomposed, the "DC Component" can be filtered out at the reconstruction process.

To accomplish the filtering of the "DC Component", all the low pass structures resulting from the signal decomposition and that are not of impulsive nature (time support not smaller than 10% of the fundamental frequency period), are eliminated in the reconstruction process. This filtering has shown to be effective when applied to synthetic and natural signals as well as signals obtained through the ATP-EMTP (Alternative Transient Program – Electromagnetic Transient Program) [10].

Most analysis of oscillographic signals are based on comparisons of the values of current and voltage quantities, often in phasor form. For that the signal is filtered to obtain just the fundamental frequency contribution [97]. Therefore, this measure was used to evaluate the proposed filtering. In all cases the proposed filtering worked very well and the modulus and angle of the phasor have been correctly estimated. This process was applied initially to synthetic signals (Figure C.9), generated using components as in equation (C.4). The components of the original signal were two sinusoids of 60Hz with amplitudes 1.0 and 2.0 and phases 0^o and 90^o, respectively, starting and ending at samples 0 and 50 for the first, and 50 and 100 for the last component. To this signal a "DC Component" was added from sample 50 to 100

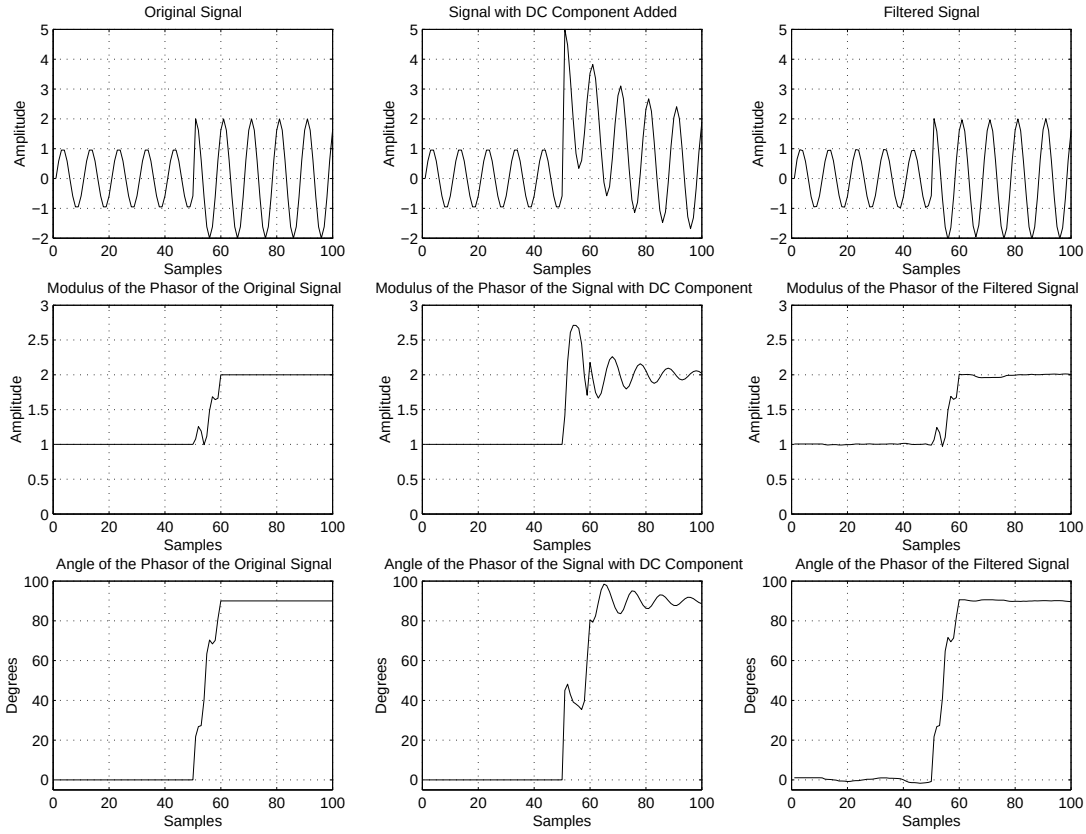


Figure C.9: Fourier filter applied after "DC Component" filtering of a synthetic signal.

with decay equal to 0.05 and amplitude 3.0 (see Figure C.9). In Figure C.9 it is seen that in the filtered signal (using the proposed decomposition and filtering) this component was eliminated. The results for an ATP-EMTP generated signal are shown in Figure C.10. Note that the only significant difference appears in the angle of the phasor after the elimination of the fault; however in this region the phasor is null, and this difference is not meaningful. Thus, it can be seen that the modulus variation due to the "DC Component" has been filtered without compromising the angle analysis.

C.4 Stopping Criterion for the Decomposition

In previous sections was presented a method to obtain coherent representations of electric signals. However, an important question remains to be answered: Is a given structure really necessary to represent the signal? Does the structure represent signal or noise? This question is now addressed.

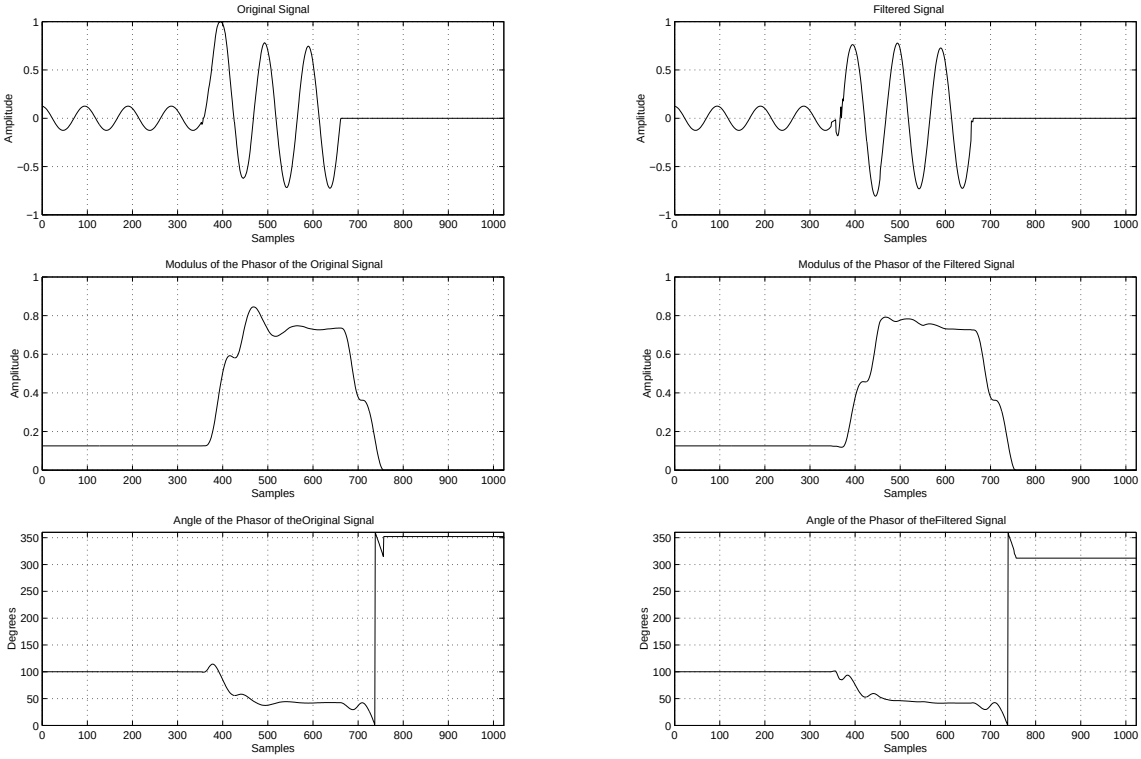


Figure C.10: Fourier filter applied after "DC Component" filtering of a signal simulated with the ATP-EMTP.

C.4.1 Approximation Ratio

The approximation ratio [29, 85] is defined as

$$\lambda(n) = \frac{|\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma(n)} \rangle|}{\|\mathbf{r}_{\mathbf{x}}^{n-1}\|}. \tag{C.20}$$

It is a measure of how much of the signal residue is approximated at step n. Figure C.11.(a) shows the behavior of the approximation ratio $\lambda(n)$ in the MP algorithm, with a Gabor dictionary of discrete parameters and continuous phase (section C.2), for signals of 128 samples. The values shown are for synthetic signals and for a noise signal (denoted as *g-128*) generated with independent identically distributed (i.i.d.) random Gaussian variables. The approximation ratio tends to be inside a neighborhood of a fixed value λ_0 for a sufficiently large number of steps (although the behavior around this neighborhood is quite irregular) [84]. It should be noted that λ_0 is independent of the signal being decomposed, it depends only on the dictionary [29,31,84]. The dictionary, in turn, depends only on the signal space dimension (length) in our algorithm, as the parameter space of the structures that compose the dictionary is parameterized based only on the signal length. It is also noticeable that

for the noise signal (i.i.d. Gaussian noise) the approximation ratio wanders around λ_0 even at the initial steps. Davis [29, 31] defines the structures with large $\lambda(n)$ as the coherent structures, since they are highly correlated with the signal. Therefore, one can assume that the residues of the signal being decomposed behave as noise after their approximation ratio is in the vicinity of λ_0. Thus, if the approximation ratio of the residues behaves like the one of noise, then there are no more coherent components to the dictionary in the residue, and the decomposition/approximation process can be stopped.

Note in Figure C.11.(a) that $\lambda(n)$ oscillates around the same value for all the signals for sufficiently large values of the step n. To filter out the oscillations around λ_0, we can consider the L steps forward mean approximation ratio

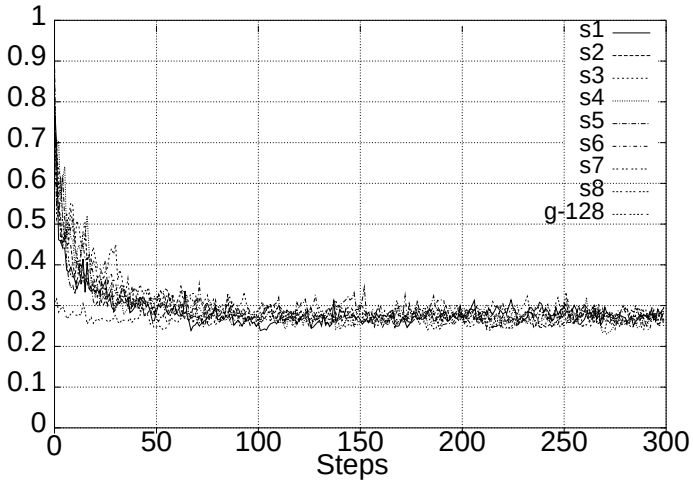
$$\lambda_{\text{mean}_f}(M) = \frac{1}{L} \sum_{n=M}^{M+L-1} \frac{|\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma_n} \rangle|}{\|\mathbf{r}_{\mathbf{x}}^{n-1}\|}. \tag{C.21}$$

At a given step M, if the moving mean approximation ratio in the next L steps is similar to the value of the noise signal mean approximation ratio, we can assume that we are approximating noise. The number of forward steps L, that are used to compute the moving mean approximation ratio, was set to be $\log_2(N)$. In other words, to avoid modelling such noise, the decomposition should be carried out only while

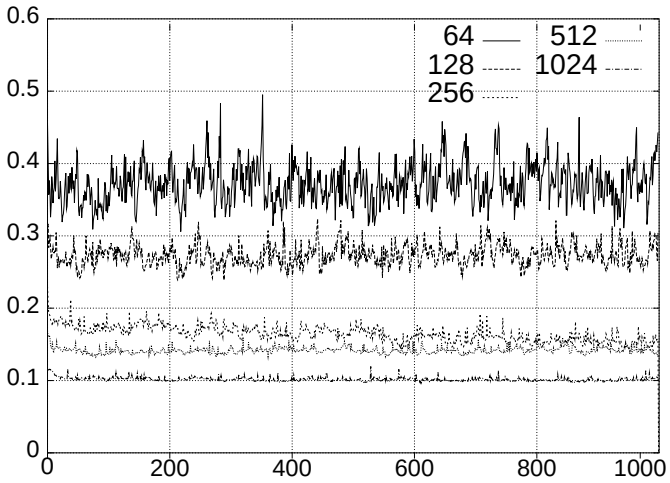
$$\lambda_{\text{mean}_f}(M) \geq \lambda_0 + \varepsilon, \tag{C.22}$$

where λ_0 is obtained for i.i.d. Gaussian noise signals and ε is a confidence constant. Thus, we need to compute, for each dictionary, the mean approximation ratio of the i.i.d Gaussian noise. In Figure C.11.(b) we see the behavior of the mean approximation ratio for different length i.i.d Gaussian noise signals of zero mean and unit variance. One can see that the value of λ_0 depends on the dictionary (that in this case is completely specified by the signal dimension), which is in agreement with [29, 31, 85]. When the time support search is applied (see section C.2.2), the approximation ratio changes because it is an intrinsic characteristic of the dictionary for a given signal dimension N. Time support search is equivalent to using a much larger dictionary, composed by the original Gabor dictionary and all its atoms with all possible time supports. Therefore, the time-support search increases the approximation ratio in the step (the search for best time support implies a greater than

or equal to inner product). However, it should be observed that with time support search the values obtained for λ_0 without this search can still be used as a lower limit for the approximation ratio as a halting criterion. In the results presented in this work ε was set to be 10% of λ_0.



(a) Approximation ratio for signals of size 128.



(b) Moving mean approximation ratio for different signal lengths N ($L = \log_2 N$).

Figure C.11: Approximation ratio behavior, of the MP using the discrete Gabor parameters dictionary with continuous phase, as a function of the step.

C.4.2 Stopping the Decomposition into Damped Sinusoids

The MP is carried out while the approximation ratio satisfies equation (C.22). This guarantees that the MP yields a model consisting of the structures that are coherent to signal phenomena and not to noise. Table C.4 shows the values of λ_0 for the dictionary of damped sinusoids that are used as the stopping criterion.

Note that the damped sinusoids dictionary employed here is fully specified by its dimension, and therefore the approximation ratio is a function only of the dimension of the signal space. Note that this a priori is independent of the fact that we use the Gaussian to damped sinusoid fitting and the procedure to decide to use either a pure or a damped sinusoid, since every outcome of the 5-tuple ($k_q F$, ρ, ϕ, m_s, m_e)) is possible.

Table C.4: Moving mean approximation ratio for different size noise signals in the MP with the damped sinusoids dictionary of continuous parameters obtained indirectly from the MP with Gabor atoms.

Signal Dimension	64	128	256	512	1024
λ_0	0.41	0.30	0.22	0.17	0.07

The values in Table C.4 are for the decomposition using the dictionary of damped sinusoids without frequency quantization and sine discrimination. As these operations decrease the inner products, they tend to reduce the approximation ratio. Therefore, a small modification in the computation of the approximation ratio in the iteration was performed: the approximation ratio is computed before the frequency quantization and pure sinusoid discrimination. Experimental results did show that this strategy performs well in identifying the coherent structures of the signal. For example, for signals *S2* and *S3* this procedure identified 4 and 17 structures respectively, with an SNR for the reconstructed signals of 83.42 dB and 109.33 dB respectively, using $L = \log_2 N$ steps to compute the mean approximation ratio. Nevertheless, it is important to point out that most of the structures found have amplitude near zero ($\gamma_q \approx 0$), that tend to be eliminated during quantization, as will be discussed in section C.5. Note that finding a large number of structures is inherent to the decomposition of most signals, either synthetic or natural, with the MP algorithm. This is so because, since the dictionary is overcomplete, each additional step of the MP, can reintroduce components removed in a previous iteration. This tends to generate decompositions with many small energy structures.

C.4.3 Noise Immunity

In order to assess the effectiveness of the proposed stop criterion we studied its performance to represent signals with different levels of added noise. This allows assessing the capability of the method for discriminating signal from noise. Table C.5 shows the energy of the added noise – Gaussian, i.i.d, noise signals of zero mean and energy values shown in Table C.5 were generated. It is also included in Table C.5 the SNR of the reconstructed signal and the number of structures identified by the algorithm. The algorithm identifies successfully and precisely the structures whose energy is larger than the noise energy. The corresponding signals can be observed in Figures C.12.(a) and C.12.(b).

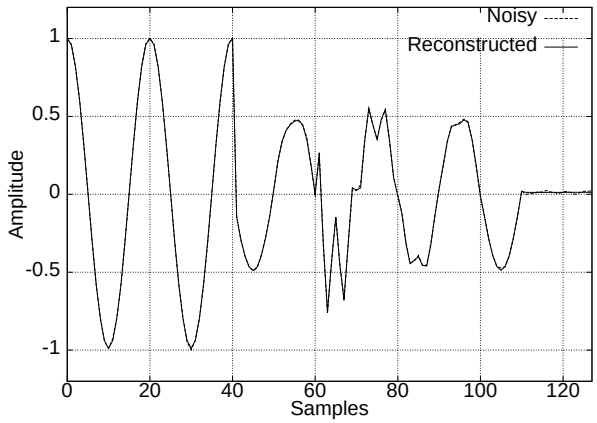
Table C.5: Decomposition of synthetic signal *S2*, using a damped sinusoids dictionary, with addition of noise.

Noise					
Energy	Norm	$\frac{\text{Signal Energy}}{\text{Noise Energy}}$	Number of Structures	SNR dB	Figure
0.0238	0.1543	1,277.4	6	36.60	C.12.(a)
0.3237	0.5689	93.92	3	19.73	C.12.(b)

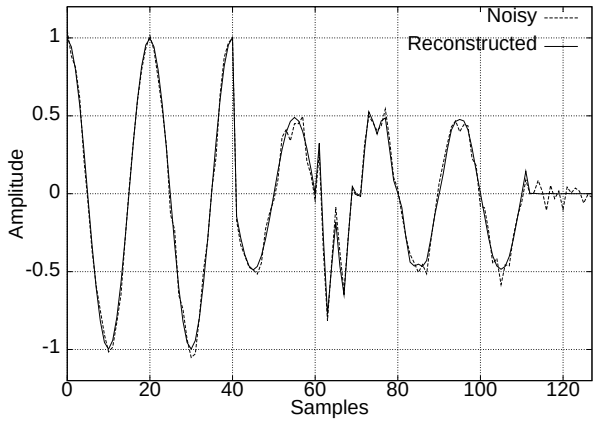
From these results we see that:

1. Small energy noise does not influence the choice of the atoms used to represent the signal. However, they influence the number of structures used, based on the mean approximation ratio, i.e., more added noise leads to earlier occurrence of noise like representation;

2. The algorithm is capable of sensing the presence of noise in the signal. This can be used to perform signal denoising by synthesis [84] – For example, if noise is added in the sampling process of power systems signals, we can use the damped sinusoids dictionary in the MP framework to represent the noisy signal and reconstruct it, from the adaptive decomposition, with much less noise, as can be seen in Figure C.12.



(a) corresponds to the case in first row of Table C.5



(b) corresponds to the second row of Table C.5

Figure C.12: Decomposition of synthetic signal *S2* with noise addition, in the damped sinusoids dictionary. For characteristics of the signal refer to Table C.5. The noisy signal is in dashed line while the reconstructed one is in continuous line.

C.5 Compression of Natural Signals

As mentioned, this decomposition is meant to represent the signals by obtaining their coherent structures, yielding good compression ratios with low distortion. Oscillographic signals are often post-analyzed by experts (human or system) in order to gain knowledge about the fault event. Thus, the compression must allow an accurate signal analysis, that is, the analysis made on the reconstructed signal must give the same results as the ones held on the original measured signal. This refers to the consistency concept, see equation (C.1). We present below a simple method to quantize and encode the parameters of the structures. This procedure is such that high SNR, high compression and accurate analysis are possible. Results confirming these conjectures for signals acquired from power systems are presented.

C.5.1 Structures Quantization

The damped sinusoids structure obtained by the decomposition proposed, at each step n, is characterized by the 6-tuple $\sigma(n) = [\gamma(n), \rho(n), \xi(n), \phi(n), t_s(n), t_e(n)]$. We present a simple scheme to quantize this 6-tuple:

- The starting and ending times (or samples) of the structures are quantized by a scalar quantizer of $\log_2 N$ bits, where N is the signal length.

- The frequency $\xi(n)$ is quantized with respect to the sampling frequency F_s and the fundamental frequency, as shown in subsection C.3.2.

- The inner product at a given step n of the decomposition, γ_n, the damping factor $\rho(n)$, and the phase $\phi(n)$ can be quantized by simple and independent fixed rate scalar quantizers.

  Those quantizers are designed according to the dynamic range (minimum and maximum values) of each parameter for all the structures found in the signal. This is done by setting up a number of bits for each of these parameters. The number of bits used to quantize the inner product (coefficients) of the structures is referred as b_{coef}; the number of bits for the damping factor is referred as b_ρ; and the number of bits for the phase is referred as b_ϕ. Given these number of bits, scalar quantizers can be designed for each parameter.

- As we store/transmit the signal we need to inform the models of the quantizers in a header, together with the signal size, the sampling frequency F_s and the fundamental F. In the experiments presented next, the header size is 183 bits long, when carrying also the clipping threshold to be applied to the signal. As stated in section C.3.5 the use of the clipping threshold may not be of practical use. Therefore, we present results for the compression ratio without considering it, what leads to a 150 bits long header.

C.5.2 Results

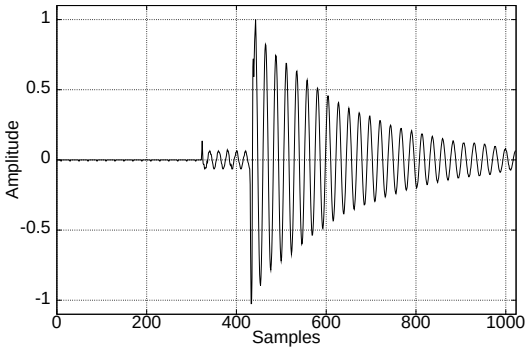
Measured signals collected from fault events in the Brazilian power system were decomposed and compressed using the scheme presented. Those signals are shown in Figures C.13.(a) (neutral current during a real fault – *R1*) and C.14.(a)

(voltage phase during a real fault – *R2*). Prior to quantization *R1* was decomposed using 19 structures and *R2* using 20 structures, see Table C.6. It is relevant to say that the first structure extracted from signal *R1* is $\sigma(1)$ = [7.451844, 0.004455, 956.947205, 11.489659, 432, 1023] where t_s and t_e are in samples, the fundamental frequency is 50 Hz and the sampling frequency 1000 Hz. After quantization it becomes $\sigma(1)$ = [7.440485, 0.004839, 950 , 10.683165, 432, 1023]. Table C.5.2 presents the ten structures that represent signal *R2* after quantization. The compression ratio is defined as the overall bit rate of the original signal divided by the total number of bits used to represent the structures of the signal and the header. The original signals were stored in single precision floating point (using 16 bits per sample - bps) and with no header, and the compression ratios were computed with respect to this format. The quality of the compressed signal *R1* can be verified visually in Figure C.13.(b) and from its reconstruction error in Figure C.13.(c). The same can be observed for signal *R2* in Figure C.14. Note that in these examples the reconstruction error is computed after the quantization of the structures and it can be diminished as the quantization is improved or the compression ratio is reduced.

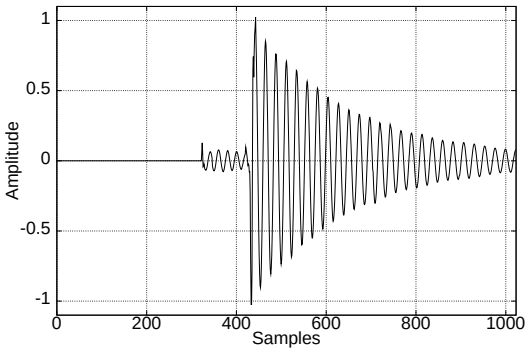
Table C.6: Quantization of natural signals decompositions.

Signal	b_{coef}	b_ρ	b_ϕ	Number of Structures	SNR (dB)	Compression Ratio
R1	6	7	7	19	28.10	15.00
R1	6	6	4	19	15.43	16.13
R2	5	4	4	7	28.32	36.01
R2	6	6	6	10	31.13	25.97

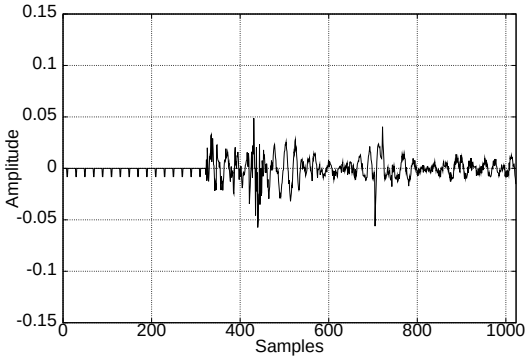
From these figures we see that the results obtained are promising in that high compression ratios can be achieved along with high SNR. In addition, the important phenomena represented in the signal (according to the damped sinusoids model of section C.1), are preserved. Since we use quantizers with a dead zone [96], the number of bits used to quantize the amplitude has a direct influence on the number of structures used to represent the signal after the quantization process. In addition, since b_ρ and b_ϕ determine the quantization error of the parameters ρ and ϕ, the representations improve as they are increased. Therefore, they yield



(a) *R1*



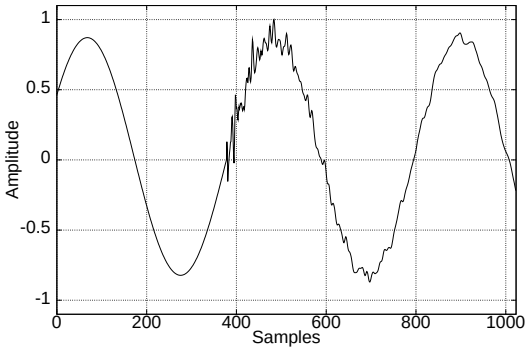
(b) *R1* Reconstruction



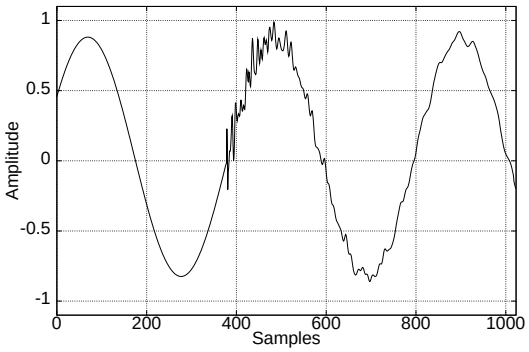
(c) *R1* Reconstruction Error

Figure C.13: Compression of signal *R1* by quantizing the atom parameters with $b_{coef} = 6$, $b_\rho = 7$ and $b_\phi = 7$, the compression ratio is 15.46 and the SNR is 28.11 dB at 1.035 bps.

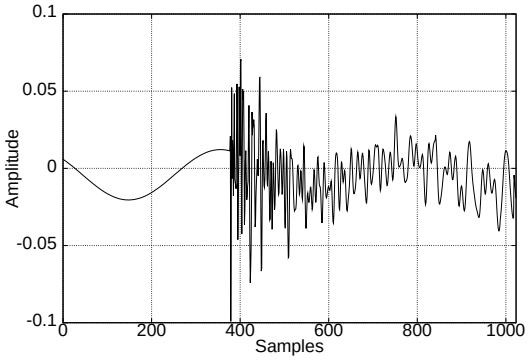
a compromise between compression ratio and reconstruction quality. The optimal selection of the number of bits used to quantize each of the three parameters is not straightforward and deserves further study. In addition, vector quantization [57] and context coding schemes [96] and content-based bit-allocation [108] could be investigated. For example, it is expected that frequency and damping factor have



(a) *R2*



(b) *R2* Reconstruction



(c) *R2* Reconstruction Error

Figure C.14: Compression of signal *R2* by quantizing the atom parameters with $b_{coef} = 6$, $b_\rho = 6$ and $b_\phi = 6$, the compression ratio is 27.40 and the SNR is 31.13dB at 0.584 bps.

some correlation between them, and exploitation of such correlation may lead to further rate-distortion improvements.

Note that for the proposed compression method if the sampling rates of the signal and of the dictionary change by the same amount, then the signal decomposition will undergo little change. In other words the bit stream obtained would be

Table C.7: Ten structures obtained after decomposing signal *R2* and quantizing the structures with $b_{coef} = 6$, $b_\rho = 6$ and $b_\phi = 6$ (see Fig C.14).

Structure	Amplitude	Freq. (Hz)	ρ	Phase (Deg.)	t_s (sample)	t_e (sample)
1	19.489243	60	0.000000	303.383514	0	1023
2	0.609039	0	0.000000	0.000000	0	1023
3	0.609039	2040	0.007807	113.768814	380	841
4	0.609039	600	0.000000	200.449829	397	1018
5	0.304519	3240	0.023422	92.098564	386	408
6	0.304519	2640	0.007807	281.713257	428	959
7	0.304519	900	0.000000	195.032257	395	973
8	0.304519	1680	0.003904	232.955185	378	919
9	0.304519	6960	0.238121	59.593189	379	382
10	0.304519	3900	0.000000	335.888855	389	472

the same (the only difference would be in the number of bits used to quantize the frequency – implying the addition of a small number of bits per structure). Thus the compression ratio increases almost linearly as a function of the sampling rate of the signal.

In Figure C.15 we present another natural signal, a voltage V from a transmission line where a fault has occurred. In Figure C.15.(a) is presented the original signal, in Figure C.15.(b) the compressed one (after quantization) and in Figure C.15.(c) the reconstruction error. In order to objectively assess the proposed compression algorithm, when used in a signal analysis framework, we have analyzed the fundamental frequency contribution in the original and the compressed signal of Figure C.15. For that purpose, the signals corresponding to the phasor before and after quantization are presented in Figure C.16. The maximum percentile errors in these computations where 1.25% and 4.75% for modulus and angle, respectively, which are indeed good results. One should note that the largest angle errors are irrelevant, as they occur near the modulus transition, which is not useful when using phasors for analysis. Figure C.17 shows the reconstruction obtained for all the three phases that compose the transmission line for the fault waveform of Figure C.15.

This shows that the decomposition behaves well, obtaining good reconstructions for all the three phases.

C.6 Chapter Summary

In this chapter, it was presented a signal decomposition scheme that yields a representation that is related to the phenomena represented in the signal, that is robust to noise, and that can be quantized effectively while maintaining signal characteristics normally extracted in automatized analysis. That is, the work presented finds a decomposition that is coherent with the signal phenomena. The algorithm decomposes a signal into a set of damped sinusoids through an adaptive algorithm based on the Matching Pursuit. That is, it represents a signal $\mathbf{x}$ as a sequence of inner products $\gamma_n = \langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{\sigma(n)} \rangle$ and indexes $\sigma(n)$, where n is the component number. Each damped sinusoid $\mathbf{g}_{\sigma(n)}$ is represented by a set of parameters $\sigma(n) = (f(n), \rho(n), \phi(n), t_s(n), t_e(n))$. Procedures to accomplish the search in this 5-parameter space in a computationally feasible and effective fashion were developed. In addition, it was presented a fast procedure to eliminate post-echo and pre-echo errors, that often appear when Matching Pursuit is employed, was presented. The decomposition algorithm that was presented obtains a signal representation based on a parameterized waveform model.

A set of novel heuristics that prevent MP from deviating from a physically meaningful decomposition was also presented. Since greedy algorithms deviate from obtaining meaningful decompositions, "intelligence" was given to the algorithm to

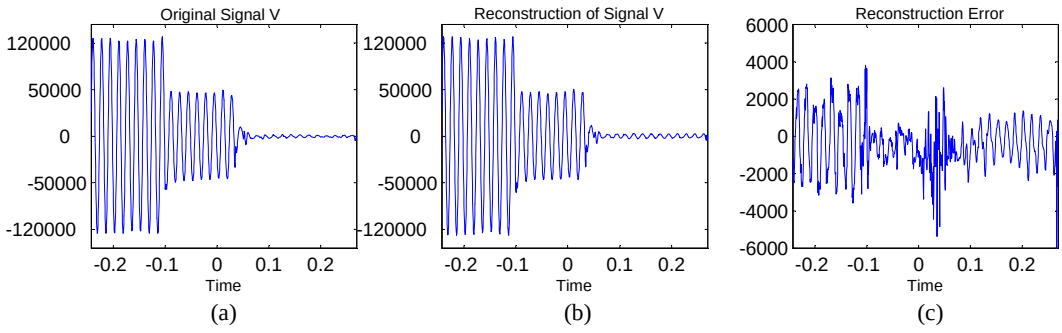


Figure C.15: Compression and reconstruction of the waveform of a fault. Compression Ratio = 68.27 and SNR = 26.13 dB.

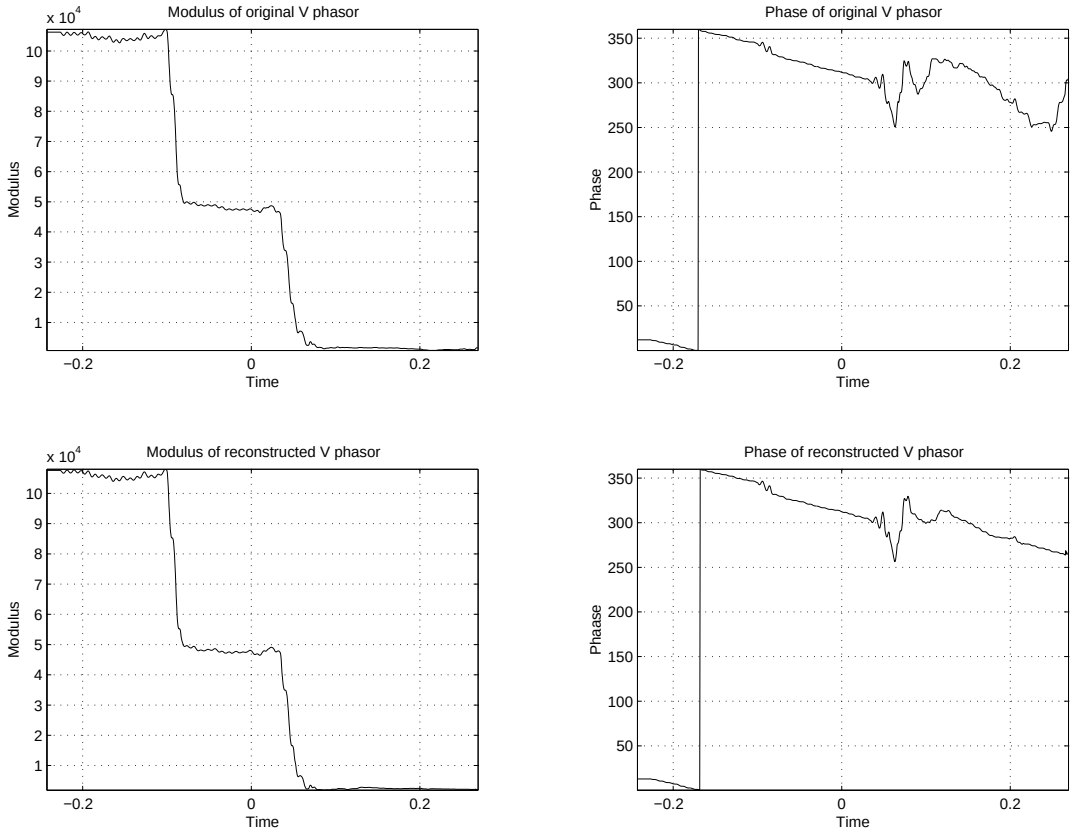


Figure C.16: Fundamental frequency contribution in the voltage signal of the compressed fault in Figure C.15.

keep it on track. This was accomplished by introducing a set of heuristics in the MP loop that instruct the MP for correct atom selection. These heuristics correct false phenomena identification which occurs when the standard MP algorithm is used.

The stopping criterion used for the decomposition process can also be understood as an estimator. This estimator tells when the decomposition is to be halted using the criterion that the residue behaves as Gaussian white noise, that is, there is no more representative information, in accordance to the signal model, to be extracted. In other words, it is decided how many atoms to use in order to decompose a signal by employing the statistics of the approximation ratio at the step. This allows to obtain an automatic coherent representation.

A filtering by synthesis application was presented. This is accomplished by using a pruned subset of the atoms found by the analysis process in the reconstruction of the signal. Many works have used the idea of filtering by thresholding, retaining just components that have energy above a pre-defined threshold [76, 84], for signal denoising. However, the filtering employed here differs from these in the

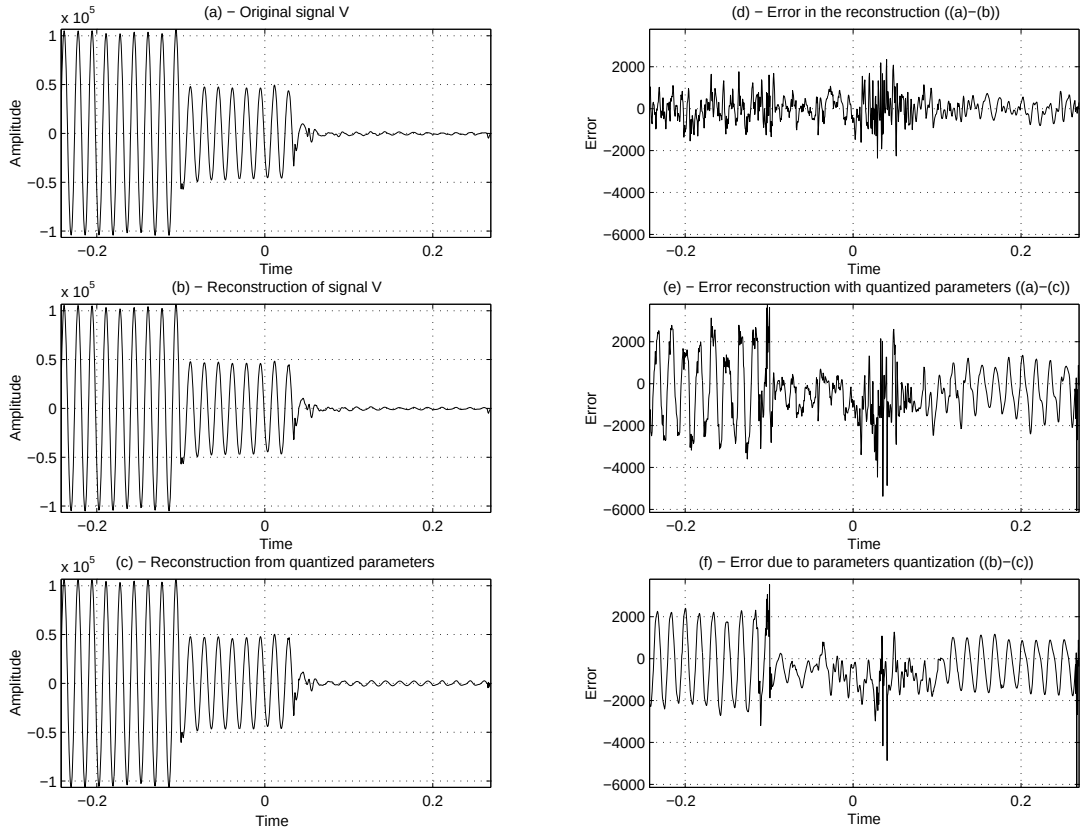


Figure C.17: Compression and reconstruction of voltage channels of a fault. (a) present the original signal, (b) the reconstructed signal decomposition, (c) the reconstructed signal after quantization of the atoms parameters, (c) the approximation error, (d) the approximation error after the parameter quantization and (e) the error due to parameter quantization.

sense that it selects the structures used in the synthesis not by their energy but by their characteristics, allowing an adaptive non-linear filtering. This approach was employed to eliminate "DC components" that often appear in oscillograms.

Almost all compression systems based on Transforms or on Matching Pursuit achieve signal compression by quantizing just the coefficients. The compression scheme presented strongly differs from other compression systems since the compression is achieved by quantizing the parameters of the signal model. That is, the reconstructed signal, after compression, is given by

$$\hat{\mathbf{x}} = \sum_{m=1}^{M} Q[\gamma_m]\mathbf{g}_{Q[\sigma(m)]}, \tag{C.23}$$

and the structures used to rebuild the signal are different from the ones obtained by the decomposition algorithm. Despite its simplicity, the quantization proce-

dure employed presented a satisfactory performance enabling the assessment of the compression capability of the proposed method. This compression obtains good compression ratios while preserving important features for signal analysis. In [102] a rate-distortion optimization procedure that can be applied to the compression scheme proposed here has been investigated.

The signal decomposition algorithm presented and the compression scheme that is applied to the resulting decompositions were evaluated with both synthetic and natural signals. The results with natural signals were very good. We often have more control of synthetic data than we have of natural data, and several algorithms have the performance diminished with natural data; however what was not the case for the proposed algorithm and compression scheme.

To accomplish an even more appropriate modelling of oscillographic signals it might be worthy to include into the model parameterized functions that represent inrush currents, as well as an approach using non quantized frequency to model inter-harmonic frequency components. This would enable the elimination of sub-synchronous components from the oscillograms, that are components of frequency near to the fundamental that are not multiples of it, which impair the analysis of oscillographic signals. In addition, a truly relevant continuation would be to submit the compressed signals to the analysis by system experts. This would enable a subjective assessment of the proposed compression method.

Apêndice D

Lloyd-Max Quantizer for Matching Pursuit Decompositions

The Matching Pursuit algorithm approximates a signal using a linear combination of pre-defined atoms from a dictionary. The linear combination is obtained using an iterative algorithm. The performance of the MP for signal compression applications depends heavily on two aspects:

i) Dictionary – it should include atoms that are good matches to the possible components of the signals to compress. In addition, the dictionary cardinality affects the data rate.

ii) Quantization – for compression the coefficients γ_n must be quantized.

Using a coefficient quantization rule $Q[\cdot]$ the compressed signal is retrieved by the quantized M-term representation

$$\hat{\mathbf{x}}_q = \sum_{n=1}^{M} Q[\gamma_n]\mathbf{g}_{i(n)}. \tag{D.1}$$

In order to design efficient quantizers $Q[\cdot]$, one needs a statistical model for MP coefficients. However, MP coefficients are difficult to model. For example, in [30] it has been observed that MP residues have a chaotic behavior. Here, instead of searching for a good model for MP coefficients, our modeling approach is based on the angles in Matching Pursuit iterations.

At each iteration of the MP angle between the residue $\mathbf{r}_{\mathbf{x}}^{n-1}$ and the selected atom $\mathbf{g}_{i(n)}$ is

$$\theta_n = \arccos\left(\frac{\left\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{i(n)} \right\rangle}{\|\mathbf{r}_{\mathbf{x}}^{n-1}\|}\right). \tag{D.2}$$

Experimentally, for a signal source $\mathbf{x} \in \mathcal{X}$ whose probability density function depends just on the signal norm $\|\mathbf{x}\|$, the statistics of θ_n seem to be independent of n. Here it is conjectured that the angles between the residues and the atoms in Matching Pursuit iterations can be statistically modeled as independent and identically distributed. Although this statistical model is an approximation, it is works well for a large class of dictionaries.

We show also that if the dictionary contains at least one orthonormal basis, then the use of a statistical model for the angles that is independent and identically distributed at each iteration is valid only if the number of the iterations is smaller than or equal to the signal space dimension. More specifically, dictionaries that include an orthonormal basis have a non-zero probability to produce residues such that $\mathbf{r}_{\mathbf{x}}^{n} = 0$, whenever the number of decomposition steps n is greater than or equal to the signal dimension. The proposed statistical model can be adapted for dictionaries including orthonormal basis by employing two sets of statistics, one set models the statistics whenever the step number is smaller than the signal dimension and another for the remaining cases.

The statistical model of the angles derived is employed to perform Lloyd-Max quantization of Matching Pursuit decompositions. The Lloyd-Max quantizer presented is compared to the state-of-the-art off-loop Matching Pursuit quantization scheme. Results show that both quantizers have similar rate-distortion performance. The good rate-distortion performance of the Lloyd-Max quantizer designed corroborates that the statistical model for MP angles proposed here is appropriate to be used in practice.

D.1 Angles in Matching Pursuit Iterations

Using the definition of θ_n, equation (D.2), the coefficients delivered by MP algorithm are given by

$$\begin{aligned}
\gamma_1 &= \|\mathbf{x}\| \cos(\theta_1), \\
\gamma_2 &= \|\mathbf{x}\| \sin(\theta_1) \cos(\theta_2), \\
&\vdots \\
\gamma_n &= \|\mathbf{x}\| \left[\prod_{i=1}^{n-1} \sin(\theta_i)\right] \cos(\theta_n).
\end{aligned} \tag{D.3}$$

In this work we assume that the atom and its negative belong to the dictionary $\mathcal{D}$ without any loss of generality. In this case, the coefficients γ_n can be made always positive without impacting the signal approximation. For that purpose, if $\mathcal{D}$ does not include $-\mathbf{g}_k$ then $-\mathbf{g}_k$ is included in $\mathcal{D}$ and $\#_{\mathcal{D}}$ is updated accordingly, obtaining then always positive coefficients. Note that, the data rate remains the same. Although it may be required one more bit to index the atoms in the M-term representation since at most we double $\#_{\mathcal{D}}$ to include every $-\mathbf{g}_k$ in $\mathcal{D}$, since there is no need to code any coefficient sign information saving thus one bit per coefficient in the M-term representation the rate remains the same.

D.1.1 Magnitude Decay of Matching Pursuits Coefficients

The maximum angle between any signal in the space and its closest atom in $\mathcal{D}$, defined in equation (B.7), bounds the error of the M-term representation, see equation (B.8). It should be noted that $\Theta(\mathcal{D})$ is the arc-cosine of the minimum of the approximation ratio, which has been discussed in subsection C.4.1, with respect to all signals in the space. However, subsequent MP coefficients may not be decreasing in magnitude, see equation (D.3). Theorem D.1 uses $\Theta(\mathcal{D})$ to find an upper bound for the number of MP iterations required for the decrease of the coefficient magnitude.

Theorem D.1 *After*

$$m = \left\lceil \frac{\log\left[\cos\left(\Theta(\mathcal{D})\right)\right]}{\log\left[\sin(\Theta(\mathcal{D}))\right]} \right\rceil \quad \textit{steps,} \tag{D.4}$$

the MP always produces coefficients $\gamma_{n+m} < \gamma_n$.

Proof: *Since*

$$\gamma_n = \|\mathbf{x}\| \left[\prod_{i=1}^{n-1} \sin(\theta_i)\right] \cos(\theta_n) \tag{D.5}$$

and

$$\gamma_{n+m} = \|\mathbf{x}\| \left[\prod_{i=1}^{n+m-1} \sin(\theta_i)\right] \cos(\theta_{n+m}) \tag{D.6}$$

it follows that

$$\gamma_{n+m} = \gamma_n \tan(\theta_n) \sin(\theta_{n+1}) \ldots \sin(\theta_{n+m-1}) \cos(\theta_{n+m}). \tag{D.7}$$

Suppose $\|\mathbf{r}_{\mathbf{x}}^{n-1}\|$ *is known, the smallest value of* γ_n *occurs if* $\theta_n = \Theta(\mathcal{D})$ *and is given by* $\gamma_n = \|\mathbf{r}_{\mathbf{x}}^{n-1}\| \cos(\Theta(\mathcal{D}))$. *Meanwhile, the largest possible value of* γ_{n+m} *is obtained if* $\theta_i = \Theta(\mathcal{D})\ \forall\ i \in [n, \ldots, n+m-1]$ *and* $\theta_{n+m} = 0$. *Thus, the largest number of steps* m *after which* $\gamma_{n+m} < \gamma_n$, *that is* $\frac{\gamma_{n+m}}{\gamma_n} < 1$, *is such that*

$$\tan(\Theta(\mathcal{D})) \left[\sin(\Theta(\mathcal{D}))\right]^{m-1} > 1, \tag{D.8}$$

what gives

$$\frac{\sin(\Theta(\mathcal{D}))^m}{\cos(\Theta(\mathcal{D}))} > 1 \tag{D.9}$$

and thus

$$m > \frac{\log\left[\cos(\Theta(\mathcal{D}))\right]}{\log\left[\sin(\Theta(\mathcal{D}))\right]}. \tag{D.10}$$

■

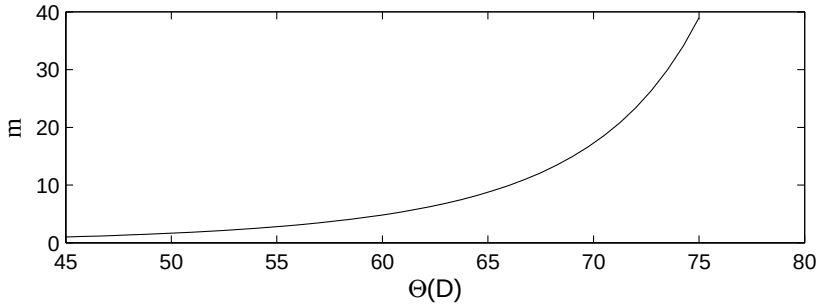


Figure D.1: Number of steps m that guarantees that $|\gamma_{n+m}| < |\gamma_n|$ in the MP as a function of $\Theta(\mathcal{D})$.

From Theorem D.1 one can predict, once $\Theta(\mathcal{D})$ is known, how many iterations m are needed to generate a coefficient that is guaranteed to be smaller than the coefficient of the current step. Note, however, that the bound provided by Theorem D.1 is weak since there is no guarantee that even exists a signal such that, when

decomposed using the MP, gives $\theta_n = \Theta(\mathcal{D})\ \forall\ n$. Therefore, in general, the number of steps m after which the magnitude decrease is smaller than the one provided by Theorem D.1.

Figure D.1 shows m as a function of $\Theta(\mathcal{D})$. If the dictionary is such that $\Theta(\mathcal{D}) \leq \pi/4$ then the magnitude of the coefficients decreases at each step. Such small $\Theta(\mathcal{D})$ can only be reached by dictionaries with large cardinality $\#_{\mathcal{D}}$ [23, 81] with respect to the space dimension N. For example, the first shell of the D_4 lattice [23] has 24 vectors (taking into account the two opposite vectors $\mathbf{g}_k$ and $-\mathbf{g}_k$) with $\Theta(\mathcal{D}) = \pi/4$. Raising the dimension to 8 the first dictionary that gives such a small $\Theta(\mathcal{D})$ is the first shell of the ϵ_8 lattice [23] that has 240 vectors (this dictionary also includes $\mathbf{g}_k$ and $-\mathbf{g}_k$). From these dictionaries one sees that $\#_{\mathcal{D}}$ does not grow linearly with N [23] if it is desired that $\Theta(\mathcal{D}) = \pi/4$. On the other hand, in order for the norm of MP residues to decay it is just necessary that $\Theta(\mathcal{D}) < \pi/2$, a condition met by any set of vectors spanning $\mathbb{R}^N$. There is no known solution to find $\Theta(\mathcal{D})$ for any given $\mathcal{D}$, actually this is a very difficult task. However, for practical applications estimates of $\Theta(\mathcal{D})$ may suffice.

D.1.2 Statistics of the Angles in MP Iterations

In [30], it has been observed that MP residues have a chaotic behavior, thus it would be reasonable to assume that, after some MP iterations, the residues can have any orientation. In more precise terms, it means that one can assume their probability distribution to be dependent just on their magnitudes. This implies that one could assume that the orientation of the residues may have an uniform probability density function on the unit-ball.

Consider now a memoryless independent and identically distributed (iid) Gaussian source (or simply, a Gaussian source). Being $\mathbf{x} = [x[1], x[2], \ldots, x[N)]^T$ an outcome from this source, it is such that all the $x[j]$ have the same Gaussian distribution $\mathcal{N}(0, \sigma^2)$. The probability distribution of this source depends just on the outcome magnitude and the orientation of this source has a uniform probability density function on the unit-ball. Although the Gaussian source may not match some actual signal sources or residues in MP iterations, it provides the same probability density function for any signal orientation and therefore does not exclude any

possible signal, i.e. it does not privilege any signal orientation.

Since the residues can be assumed to have a probability distribution that depends just on their norms and the Gaussian source also has this property, one may expect the angles in MP iterations arriving from the decomposition of MP residues to have a distribution that is similar to the distribution of the angle in the first iteration of the MP for a Gaussian source.

Figure D.2 shows the normalized histograms of the RVs (Random Variables) Θ_n of the angles at the n^{th} MP step or iteration for several values of n. They result from decompositions of realizations of a Gaussian source using a dictionary composed of 16 normalized Gaussian random atoms drawn from a Gaussian source in $\mathbb{R}^4$. Dictionaries of this kind, composed of $\#_{\mathcal{D}}$ normalized signals drawn from an N-dimensional Gaussian source, are referred here as $GSND(\#_{\mathcal{D}}, N)$ and thus the former dictionary is denoted by $GSND(16, 4)$. Figure D.3 shows the mean and the variance of Θ_n for several n and the covariance among some Θ_n for the same dictionary and source. The results in these figures were obtained using an ensemble of 50,000 MP decompositions of random signals drawn from a Gaussian source. In Figure D.2 one notes that the pdfs of the RVs Θ_n have a similar shape for all n. This leads to the conjecture that the pdfs $f_{\Theta_n}(\theta_n)$ are independent of the iteration number n and are identically distributed, i.e., their pdfs are the same, at each iteration. The results presented in [29, 30] corroborate this supposition. In these works it is shown that, under specific conditions, the MP angles θ_n have a chaotic behavior, i.e. the residue in step n maps, chaotically, to the residue in step $n + 1$.

Residues may have any orientation. Therefore, for a source with uniformly distributed angles, that is, with outcomes whose pdf depend only on their magnitudes, the angle RV, in the first MP step, Θ_1 may have a pdf that is similar to the ones of subsequent decomposition steps. That is, Θ_1 may have a pdf that is similar to the Θ_n for $n > 1$. Thus, it would be reasonable to assume that the angles at each step are statistically independent. Figure D.3 depicts the covariances between MP angles in different steps. From it we see that $\mathrm{Cov}[\Theta_i\Theta_k] = 0$, $\forall\ i \neq k$, that is, the angles are uncorrelated. In this chapter we use the hypothesis that the angles are independent RVs. This is not an unreasonable assumption, since it does not contradict the behavior observed in Figure D.3.

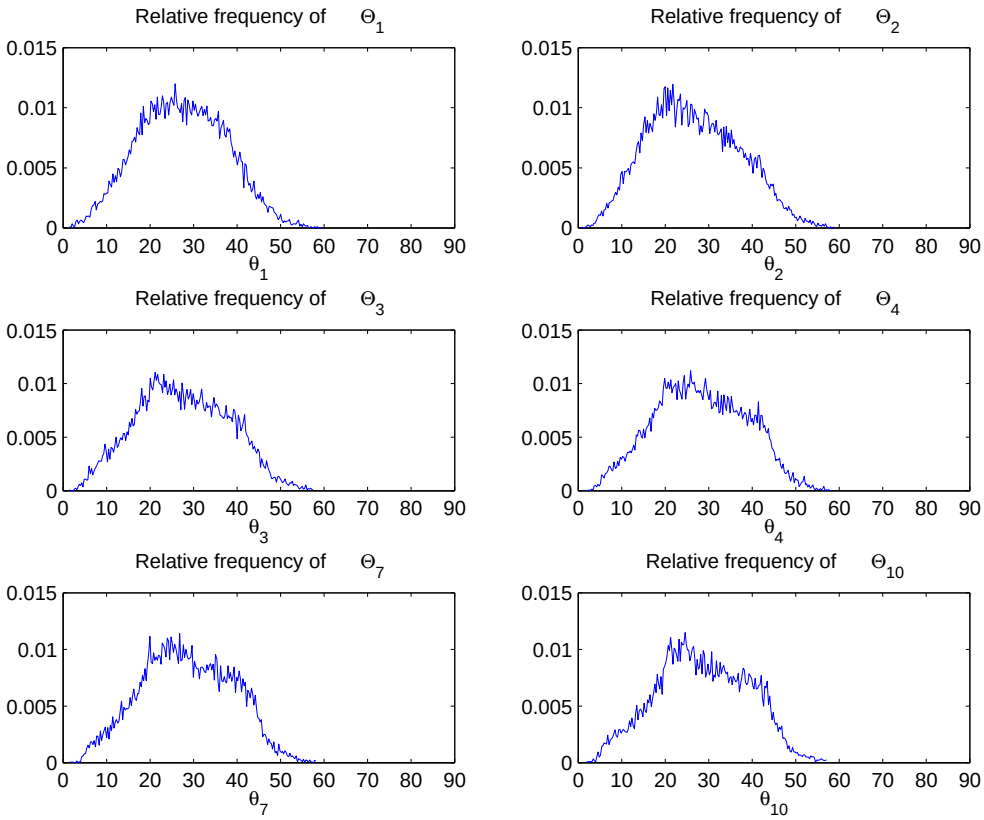


Figure D.2: Relative frequency histograms of Θ_n for a Gaussian source of $\mathbb{R}^4$ using the $GSND(16, 4)$.

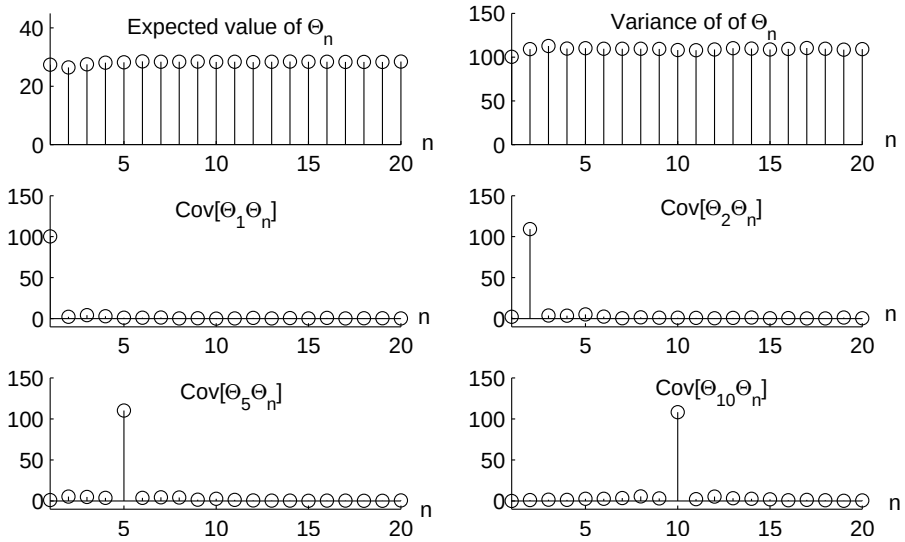


Figure D.3: Mean, variance and covariance of Θ_n for a Gaussian source in $\mathbb{R}^4$ using the $GSND(16, 4)$.

D.1.3 Angle Statistics for the Gabor Dictionary in $\mathbb{R}^{64}$

The results presented so far use dictionaries of relatively low dimension and cardinality. In practice the MP is commonly used in large dimensional spaces using

parameterized dictionaries as the Gabor dictionary, discussed in subsection C.2. The elements of this dictionary are defined by translations, modulations and dilations of a prototype signal. The most common choice for the prototype signal $\mathbf{f}[n]$ is the Gaussian window. The atoms of the Gabor dictionaries are complex, and the optimal phase for the atom can be computed, see subsection C.2. However, for compression applications, a quantized phase is required. Therefore, we analyze here a Gabor dictionary composed of atoms with phases being multiples of $\frac{\pi}{V}$. Each atom is then given by

$$g[n] = \begin{cases} \delta[n], & \text{if } j = 0 \\ K_{(j,p,v)} f\left[\dfrac{n - p2^j}{2^j}\right] \cos\left[nk\pi 2^{1-j} + \dfrac{\pi v}{V}\right], & \text{if } j \in (0, L) \\ \frac{1}{\sqrt{N}}, & \text{if } j = L \end{cases}, \tag{D.11}$$

where $f[n] = 2^{\frac{1}{4}} e^{-\pi n^2}$, n is the sample, $K_{(j,p,v)}$ provides a unit-norm atom, and $v \in [0, \ldots, V-1]$. In the definition above j defines the atom scaling, p defines the time shift, and k defines the atom modulation. For $L = \log_2(N)$ scales the atom parameters ranges are $j \in [0, L]$, $p \in [0, N2^{-j})$, $k \in [0, 2^j)$, and $v \in [0, V-1]$.

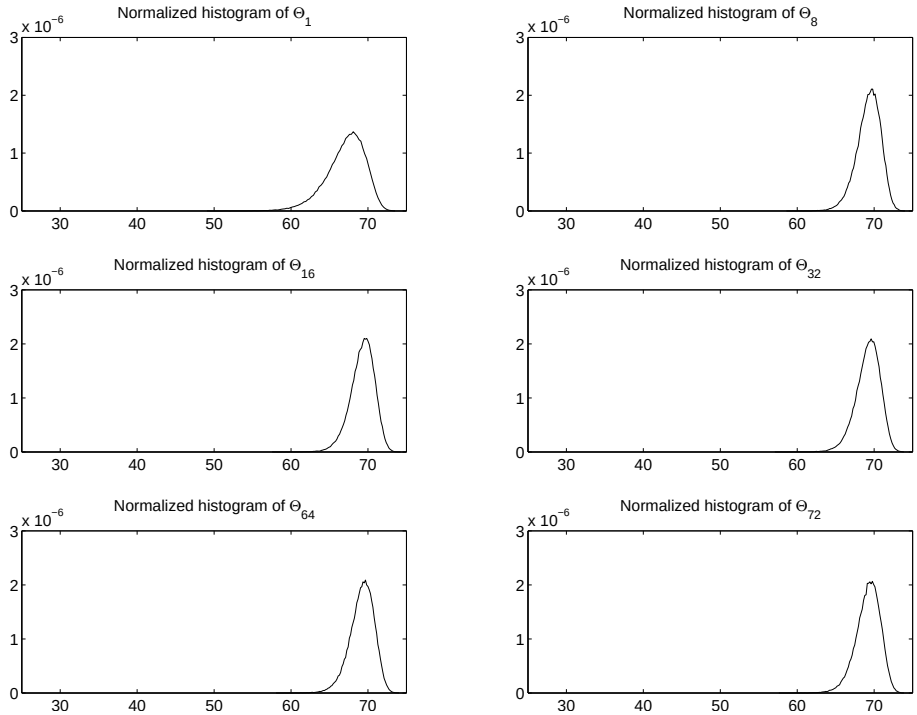


Figure D.4: Normalized histograms of MP angles for a Gaussian source in $\mathbb{R}^{64}$, using 100 bin, at $n = \{1, 8, 16, 32, 64, 72\}$, for the 4-phase Gabor dictionary in $\mathbb{R}^{64}$.

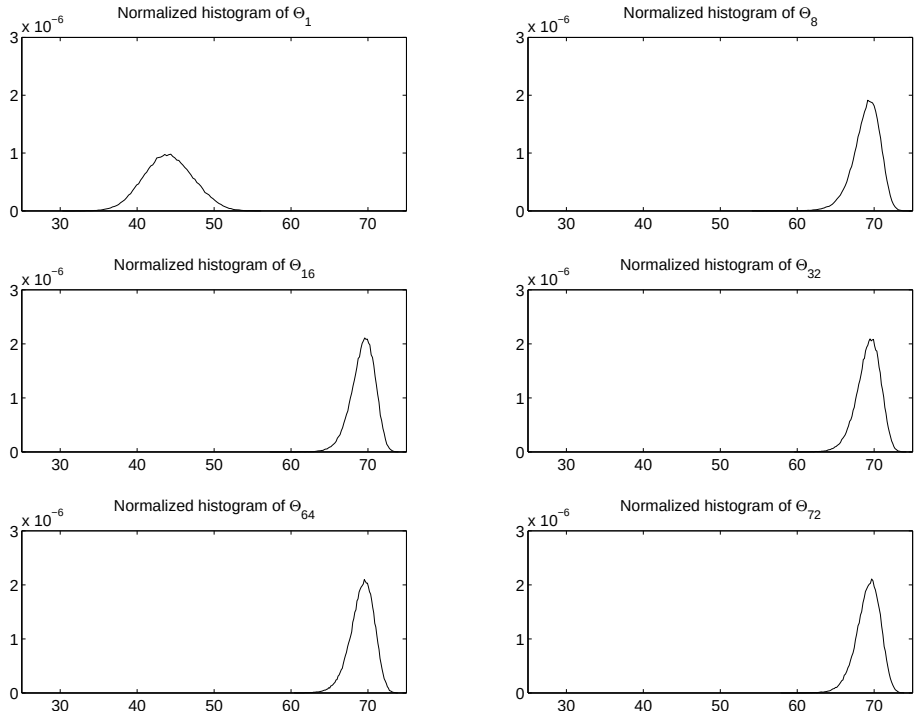


Figure D.5: Normalized histograms of MP angles for a source with coordinates driven from a gamma distribution, using 100 bin, at $n = \{1, 8, 16, 32, 64, 72\}$, for the 4-phase Gabor dictionary in $\mathbb{R}^{64}$.

Figure D.4 shows $f_{\Theta_n}(\theta_n)$, the probability density function of the RVs Θ_n – the normalized histograms of Θ_n, for some n, obtained for an ensemble of 128,000 decompositions of Gaussian signals in $\mathbb{R}^{64}$ using the Gabor dictionary with four phases. Figure D.5 shows $f_{\Theta_n}(\theta_n)$, for some n, obtained for an ensemble of 128,000 decompositions of signals driven from a memoryless source that has gamma distributed coordinates in $\mathbb{R}^{64}$ for the same dictionary. Note that the angles' statistics shown for each signal source differ only at the first MP iteration, being visually very similar for the other iterations. It can be noted that for this dictionary the angles in different MP steps $n \geq 2$ have similar statistics even for very different sources. Note that, although these statistics are not exactly equal to the statistics that are obtained for the first decomposition stage of Gaussian source, they are reasonably similar and therefore $f_{\Theta_n}(\theta_n)$ for $n > 1$ can be approximated by the $f_{\Theta_1}(\theta_1)$ obtained for a memoryless white Gaussian source. This is a reasonable assumption, since the memoryless white Gaussian source does not privilege any orientation, what seems to be the case for the residues $\mathbf{r}_{\mathbf{x}}^{n-1}$ for $n > 1$.

Figure D.6 presents the average value of the angle in MP iterations, the standard deviation of these angles and the largest value of the angle in MP iterations $\Theta_i(\mathcal{D})$ for the memoryless white Gaussian source. Figure D.7 presents the same results but for a memoryless gamma distributed source. It can be noted that for both sources that as n increases the angle statistics converge, $E[\Theta_i]$ and $\text{var}[\Theta_i]$, to fixed values, it can also be noted that these values are source independent although the speed of the convergence is faster for the Gaussian source than for the Gamma distributed one. In addition, note that $\Theta_i(\mathcal{D})$ seems to be constant. Figure D.8 presents the correlation and covariance between the angles in MP iterations for the decomposition of an ensemble from a Gaussian source. As one can observe the same behavior for the covariance than the one observed for the dictionary composed of 16 normalized outcomes of a Gaussian source in $\mathbb{R}^4$ – the $GSND(16, 4)$.

It should be noted, from the results presented, that for the $\mathbb{R}^{64}$ Gabor dictionary the value of $\Theta(\mathcal{D})$ is larger and $f_{\Theta_1}(\theta_1)$ is narrower than they are, for example, for the $GSND(16, 4)$. These observations indicate that the decay rate of the magnitude of MP coefficients for the $\mathbb{R}^{64}$ Gabor dictionary is not as fast as it is for the $GSND(16, 4)$.

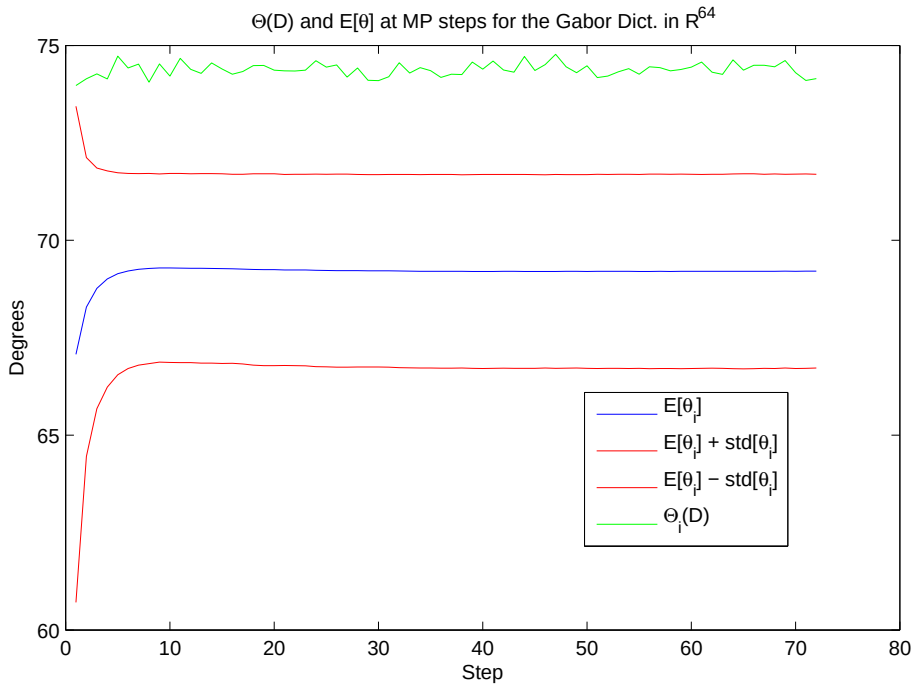


Figure D.6: Statistics of the angles in MP iterations for a Gaussian source in $\mathbb{R}^{64}$ as a function of the iteration for the Gabor dictionary in $\mathbb{R}^{64}$ of four phases.

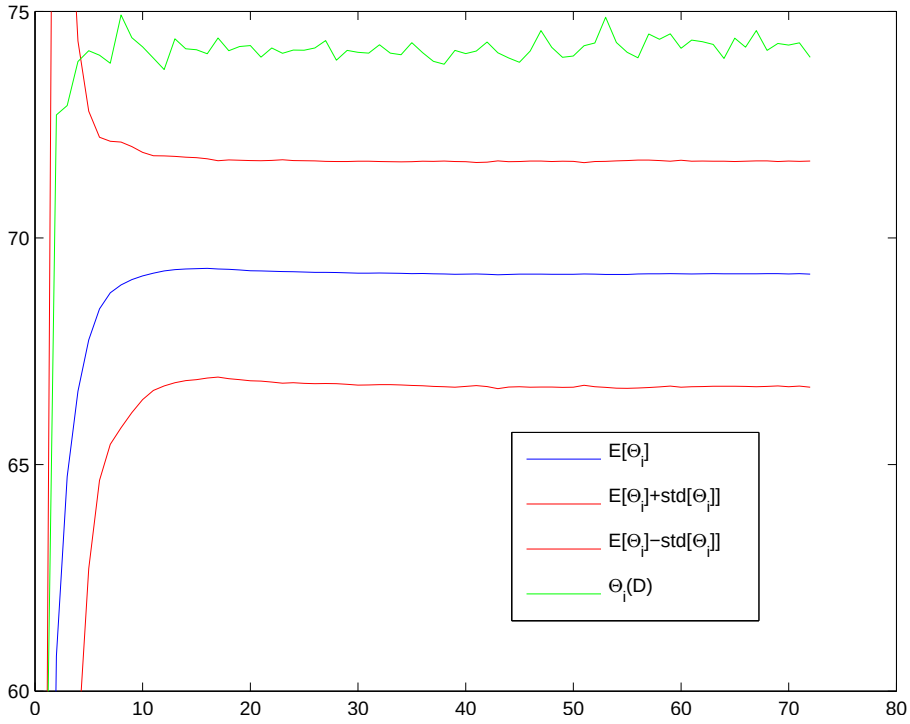


Figure D.7: Statistics of the angles in MP iterations for a Gamma distributed source in $\mathbb{R}^{64}$ as a function of the iteration for the Gabor dictionary in $\mathbb{R}^{64}$ of four phases.

D.1.4 Dictionaries Including Orthonormal Bases

The iid model for MP angles is well suited for a large number of dictionaries. Nevertheless, when the dictionary includes an orthonormal basis, the iid model requires a slight modification. We address this aspect now. An important contribution of the present work is the proof of the following proposition:

Null Residue Proposition *If the dictionary includes an orthonormal basis, then, under very general conditions, the MP algorithm has a non-zero probability of generating null residues in a finite number of steps.*

The proof of the "null residue proposition" is presented in Appendix G. In Appendix G it is shown that if $\mathcal{D}$ includes an orthonormal basis then the MP has a non zero chance of trapping the residuals of the MP decomposition of a given vector into successive subspaces that are orthogonal to the previous selected atoms. Therefore, since for any n it is known that $\mathbf{r}_{\mathbf{x}}^{n} \perp \mathbf{g}_{i(n)}$, if the source is iid Gaussian then there is a non-zero probability to produce a null residue, that is, to obtain $\mathbf{r}_{\mathbf{x}}^{n} = \vec{0}$, in a finite number of steps $n \geq N$.

Experimentally we observed that when a $GSND(20, 4)$ is used to decompose

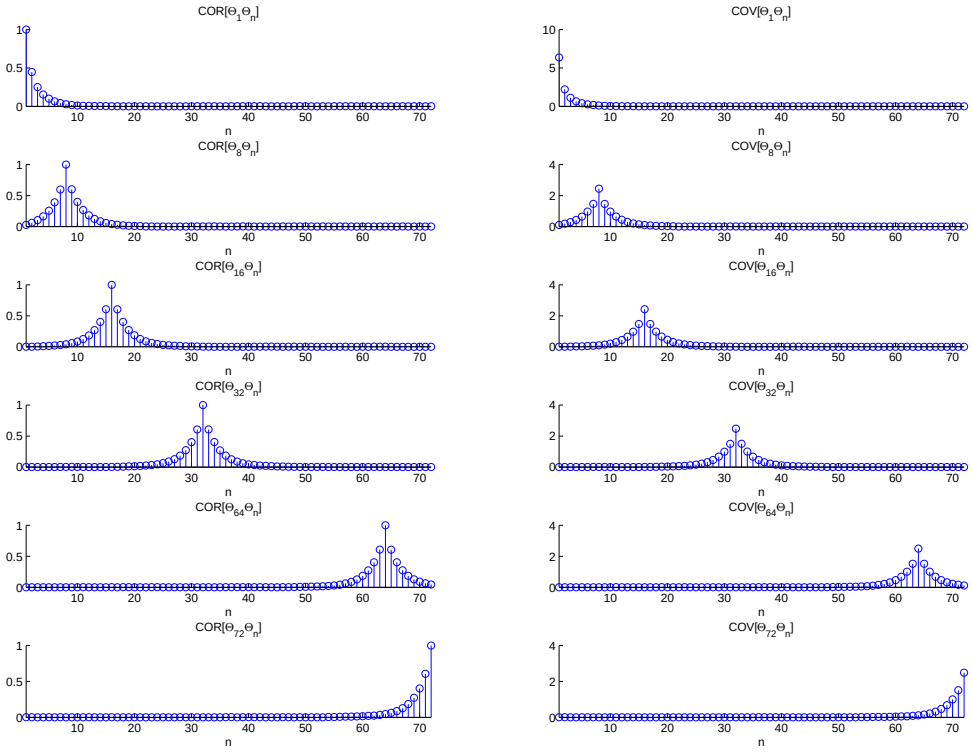


Figure D.8: Correlation and covariance between angles in MP iterations for a Gaussian source in $\mathbb{R}^{64}$ as a function of the iteration for the Gabor dictionary in $\mathbb{R}^{64}$ of four phases.

an ensemble of 25,600 signals drawn from a four dimensional memoryless Gaussian source, allowing at most 100 decomposition steps, none of the signal decomposition produces a null residue in a finite number of steps. However, when a set of 4 elements of the same $GSND(20, 4)$ is replaced by the canonical basis of $\mathbb{R}^4$ exact signal expansions with finite number of terms are obtained. Figure D.9 shows the histogram of the number of null residues produced by the MP in function of the MP iteration for an ensemble of 25,600 signals drawn from a four dimensional memoryless Gaussian source when the $GSND(20, 4)$ is modified to include the 4 vectors of the canonical basis. The last bin of the histogram accounts for the realizations that did not produce null residues in any of the 100 steps of the MP.

Corollary G.1 tells that for the iid Gaussian source the MP produces null residues after N (the signal space dimension) iterations. Figure D.9 shows that the MP has also a non zero probability to produce null residues when the source

is composed by outcomes of MP residues of the Gaussian source at any iteration. Therefore, it is reasonable to assume that MP residues may have any orientation what gives support to the analyses previously held on the distributions of MP angles.

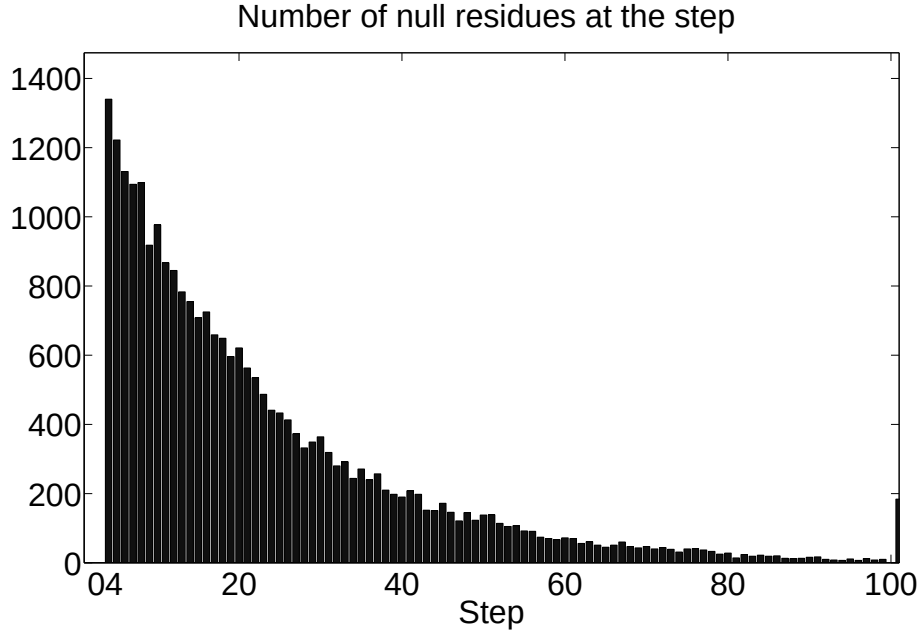


Figure D.9: Incidence of null residues using a $GSND(20, 4)$ with 4 elements replaced by the canonical basis of $\mathbb{R}^4$.

As stated by the Null Residue Proposition (see its proof in Theorem G.1) the modified $GSND(20, 4)$ produces null residues after 4 decomposition steps. In order to confirm that the null residues appear due to the modification of the $GSND(20, 4)$, by including a basis in it, we consider now what happens for the original $GSND(20, 4)$. Figure D.10 shows on its left-hand side the angle histograms at steps 1, 4, 7 and 10 for the decomposition process aforementioned using the original $GSND(20, 4)$. On its right-hand side, Figure D.10 presents the same graphs for the modified $GSND(20, 4)$. To compute the percentage histogram of the Θ_i in Figure D.10 (using 100 bins) at each step i only the angles generated by decompositions that did not produce null residues, i.e., that did not yield $\mathbf{r}_{\mathbf{x}}^{n-1} = \vec{0}$, for $n < i$, are considered. One observes in the left-hand side of Figure D.10 that for the original $GSND(20, 4)$ null angles do not occur, i.e. the residue and the selected atom never have the same direction. However, it is possible to verify in the right-hand side of Figure D.10 that for the modified $GSND(20, 4)$ when $n \geq 4$ null angles often occur. It is also interesting that the percentage of such angles as well as the whole histograms are

very similar across the iterations that are greater than or equal to 4.

The exact calculation of the probability of producing a null residue at a given step is a difficult task since its value depends on both $\mathcal{D}$ and the signal source. However, it is possible to estimate this probability from simulations since the percentage of null angles at a given step is a measure of the probability of producing null residues at that step. When $n \geq N$ steps, in the case $N = 4$, this probability seems to remain constant.

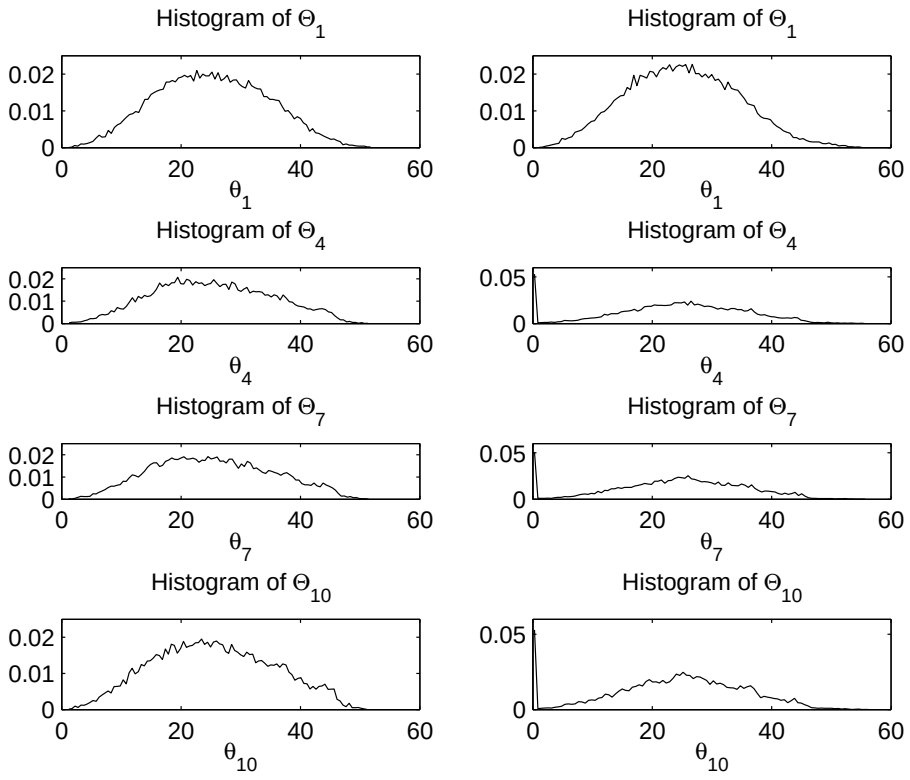


Figure D.10: Histograms of the angles at the MP steps 1, 4, 7 and 10 using the original $GSND(20, 4)$ (left) and the modified $GSND(20, 4)$ (right) – the $GSND(20, 4)$ with 4 of its elements replaced by the canonical basis of $\mathbb{R}^4$.

One may have dictionaries including several distinct orthonormal bases. For these dictionaries the probability of MP convergence in finite number of steps is expected to be larger than it is for a $\mathcal{D}$ that includes only one orthonormal basis. Another strategy is to generate dictionaries that are unions of orthonormal bases [66, 76, 84, 105]. An example of a dictionary composed by a union of bases is the D_{4,sh_1}, which is obtained by normalizing the elements of the first shell of the D_4 lattice [23, 25]. The $D_{4,sh1}$ has 24 elements in 12 directions with two opposite phases.

The dictionary formed by the normalized elements of the first shell of the

ϵ_8 lattice (the ϵ_{8,sh_1} dictionary) is composed by 240 elements (in 120 directions) in dimension 8, and for each of its elements there are either 126 or 110 orthogonal dictionary elements. Actually, the ϵ_{8,sh_1} can be shown to be the union of 30 bases for $\mathbb{R}^8$. However, the elements of the ϵ_{8,sh_1} can be grouped in order to provide a larger number of orthonormal bases than the original 30 bases, i.e., for a given dictionary element it is possible to select more than one subset of 7 elements from ϵ_{8,sh_1} that forms an orthonormal basis for $\mathbb{R}^8$, and these subsets are not just changes of sign of the elements (as was the case for the D_{4,sh_1}). Figure D.11 shows the histograms of the MP angles Θ_i, for $i \in [1, \ldots, 10]$, for the ϵ_{8,sh_1} obtained using an ensemble of 25,000 MP decompositions of Gaussian source signals. In this figure, it is possible to see that histograms of Θ_i, $i \in [1, \ldots, 7]$ are quite similar. Note also that, after a number of steps i greater than or equal to 8 (the space dimension) the histograms of the different Θ_i are also quite similar and have a large incidence of zero angles. In fact, the incidence of angles that equal zero remains constant, that is, the zero-angle incidence does not vary with i, implying a similar probability of convergence for $i \geq 8$ if the MP approximation did not converge previously. If we call the percentage of null angles p_0 then the total probability of the MP producing null residues at steps $n \geq N$ is given by

$$P(\mathbf{r}_{\mathbf{x}}^n = \vec{0}) = (1 - p_0)^{n-N-1} p_0. \tag{D.12}$$

It is important to point out that the results presented show that the first MP angle statistics for a memoryless Gaussian source are appropriate to model the angles in the first seven decomposition steps for the $\epsilon_{8,sh1}$. However, after the eighth step, they are not appropriate anymore. Nevertheless, from the results shown only two distinct pdfs are needed to model the MP angle for the $\epsilon_{8,sh1}$: one pdf being valid up to the seventh step and another being valid for step 8 and beyond.

D.1.5 Discussion

In [29, 30] the average value of the approximation ratio was studied. This value corresponds to the cosine of the MP angle at step n as defined in equations (C.20) and (D.2). In those works it is argued that the first order statistics of the approximation ratio (its mean) converge to a fixed value that depends on the signal space dimension N. Nevertheless, we have seen that the whole statistics of the angles

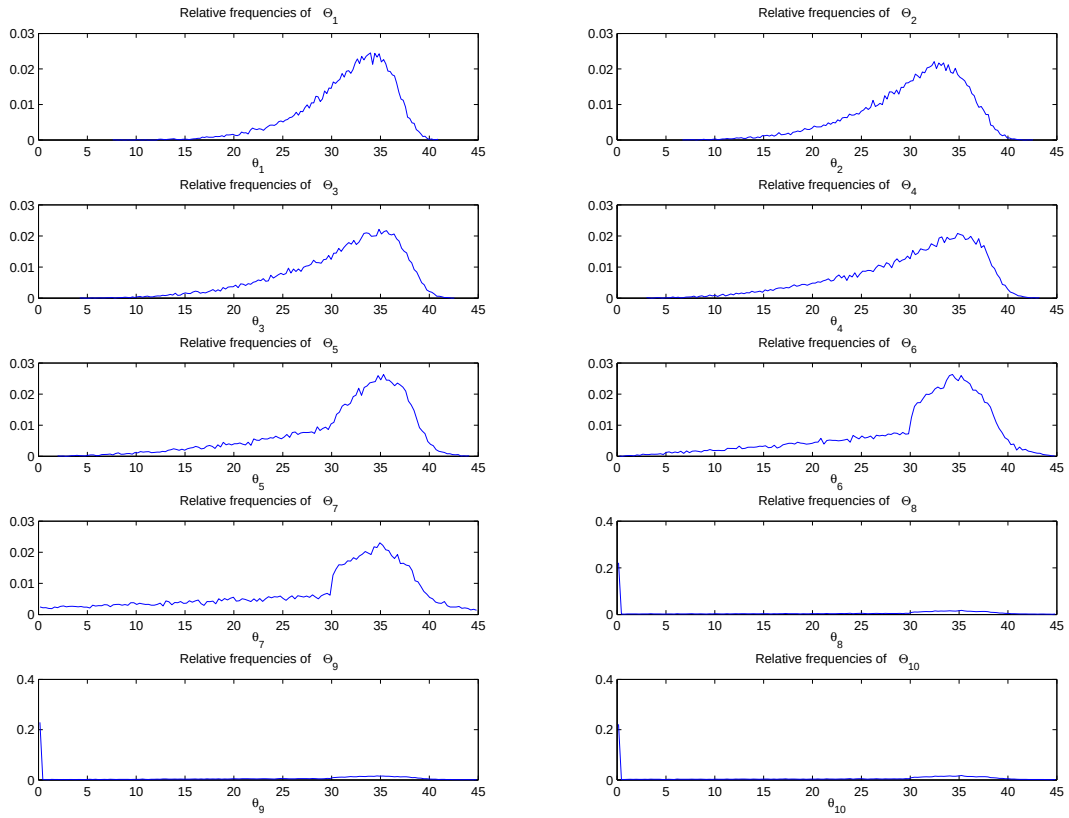


Figure D.11: Relative frequency histograms of Θ_n for a Gaussian source in $\mathbb{R}^8$ using the ϵ_{8,sh_1} dictionary.

in MP iterations can be considered to be invariant with respect to the step number. An exception occurs when the dictionary includes an orthonormal basis, when two different pdfs must be used to model the angles in MP iterations one valid at steps whose number are smaller than the signal dimension and another for the remaining steps.

The analysis of Figure D.4 shows that for the Gabor dictionary the angle pdf, after some steps, is slightly different from the angle pdf obtained for a memoryless Gaussian source. Visual inspection still indicates that one could use a Gaussian source to obtain good estimates of the pdf of the angle in MP iterations, that is the iid statistical model is not a bad assumption. In the sequel, this model is used to design Lloyd-Max quantizers of MP coefficients. This permits to verify the validity of the statistical model for the angles in MP iterations. It should be noted that the statistics of the first MP angle, and therefore of the first coefficient, are much more source dependent than they are at further MP steps. Therefore, in order, to make an appropriate use of the presented angle model, the first coefficient, γ_1, will

be quantized with negligible error and encoded as side information. The value of γ_1 can then be used instead of the value of the signal norm $\|\mathbf{x}\|$ which is usually transmitted as side information in practical MP based compression schemes [54, 88]. In the sequel, before using the statistical model for MP angles proposed for the design of Lloyd-Max quantizers for MP coefficients, we discuss some problems involved in their quantization.

D.2 Quantization of MP Coefficients

D.2.1 In-loop and Off-loop Quantization

In chapter C, the MP coefficients were quantized after the decomposition was carried out. We refer to this as off-loop. Another option would be to use in-loop quantization, that is, an atom coefficient γ_n is quantized before the corresponding residue is computed. This scheme is shown in Figure D.12 with switch B at the lower position (switch A is placed at the upper position only at the first MP iteration to make $\mathbf{r}_{\mathbf{x}}^0 = \mathbf{x}$, being moved to the lower position from then on). In the in-loop quantization scheme the coefficient γ_n and corresponding atom $\mathbf{g}_{i(n)}$ are found using the greedy iteration, then the coefficient is quantized obtaining $Q[\gamma_n]$, and $Q[\gamma_n]$ is used to compute the residue by means of

$$\mathbf{r}_{\mathbf{x}}^n = \mathbf{r}_{\mathbf{x}}^{n-1} - Q[\gamma_n]\mathbf{g}_{i(n)}. \tag{D.13}$$

As a result, the quantization error is fed back into the decomposition algorithm allowing the quantization error to be compensated in subsequent MP steps [88]. Nevertheless, there is no guarantee that the quantization error will be corrected since the coefficients of subsequent steps will also be quantized. This is further complicated by the fact that the MP decomposition is non-linear, that is, the MP decomposition of a sum of signals $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ is not, in general, the sum of the MP decompositions of $\mathbf{x}_1$ and $\mathbf{x}_2$ obtained separately. Another important fact about in-loop quantization is that it destroys a fundamental property of the MP: the orthogonality between $\mathbf{g}_{i(n)}$ and $\mathbf{r}_{\mathbf{x}}^n$. Therefore, since the greedy iteration is non-linear, quantizing the coefficient in-loop may result in M-term representations that are influenced by the quantizer used. This is particularly undesirable when signal mo-

deling is required since each distinct in-loop quantizer may generate a different signal model, that is, different atoms can be selected to express the same $\mathbf{x}$ depending upon the in-loop quantizer used.

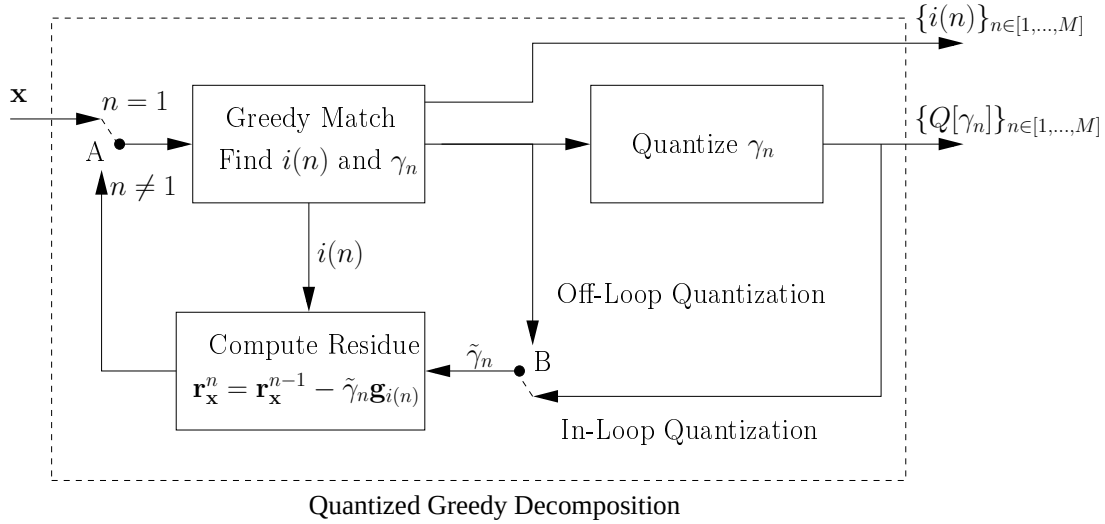


Figure D.12: In-loop and off-loop quantization of greedy decomposition coefficients.

Suppose a fixed quantizer is used for in-loop quantization of MP coefficients. Then, if the coefficient of step n is quantized to zero, all the subsequent coefficients are also quantized to zero even if without in-loop quantization the coefficient of a future step $n+m$ would had been larger than the coefficient of step n. This occurs because the residue is not updated. To guarantee the residue updating, a quantizer without a dead-zone could be used not allowing any coefficient to be quantized to zero. However, quantizers without dead-zone are not effective in a rate–distortion sense [54] as they can lead to a reconstruction error $\mathbf{x}-\hat{\mathbf{x}}_q$ that increases as M increases. Another possibility is the use of multiple-pass quantized MP decompositions [88], in which the MP algorithm is applied twice: the first pass trains the quantizer, and the second obtains the quantized M-term representation. Yet another way to secure the residue update could be based on the known result that MP coefficient magnitudes are upper bounded by a decaying exponential function of the iteration [32, 54, 104], although the coefficients may not be monotonically decreasing (see section D.1.1). Using this magnitude decay property, one could change the dynamic-range of the quantizer at each decomposition step; this dynamic-range-changing quantizer was proposed in [54] for off-loop quantization and is explained in section D.2.4. In addition to the schemes above, MP-like quantized decompositions

have also been proposed [8, 9], see subsection B.2.2.

In off-loop quantization the residues are computed without knowledge of the quantizer used (as in Figure D.12 with switch B placed at the upper position). For a given $\mathcal{D}$, this scheme always provides the same signal model for a given signal, independently of the quantizer used. This happens because the off-loop quantization scheme preserves the orthogonality between $\mathbf{r}_{\mathbf{x}}^{n}$ and $\mathbf{g}_{i(n)}$, not affecting the MP iteration. As opposed to in-loop quantization, the main drawback of off-loop quantization is that the quantization errors can not be compensated in subsequent steps. In this work, we minimize the quantization errors due to off-loop quantization of MP coefficients by designing Lloyd-Max quantizers using the iid model for MP angles.

D.2.2 On-line and Off-line Coding of Quantized M-Term Representations

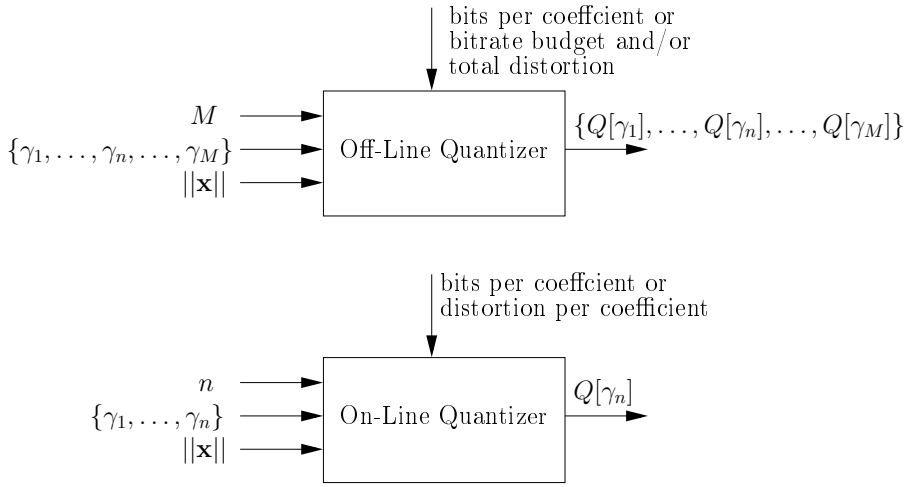


Figure D.13: Off-line (upper) and on-line (lower) quantization of M-term representations.

Off-loop quantization may be used in two different coding paradigms, see Figure D.13. In on-line coding each γ_n is quantized and transmitted as soon as it is available, that is, the coefficients are quantized without considering the coefficients of the following steps. In contrast, off-line coding is performed just after the whole M-term representation is available and can take into consideration the coefficients of all the M terms in the representation. Thus, on-line coding delivers the quantized coefficients earlier than off-line does. In addition, the on-line scheme outputs the bit-rate

of the quantized decomposition as the quantization of each coefficient occurs. Thus, in the on-line coding paradigm, if a coefficient is quantized to zero or the available rate is achieved then a halting order can be issued to the decomposition algorithm. Meanwhile, off-line coding allows for the conceptually simple rate–distortion (RD) optimization procedure which consists of trying different quantizers in order to find one quantizer meeting a prescribed RD criterion. For transmission applications, it is common practice to pre-store the signal decomposition and then quantize it in order to meet a desired RD criterion. A priori off-line coding of MP decompositions seems to have a larger computational cost than on-line coding since it requires all the M terms, i.e., M iterations of the MP need to be performed. However, MP iterations have a much larger computational complexity than commonly used coefficient quantization processes. In effect, if several versions of a signal need to be transmitted at different rates there is no need to apply the MP repeatedly when off-line coding is used thus reducing the computational complexity as compared to on-line coding.

D.2.3 Uniform Quantizer Examples

Some gain in RD performance can be achieved by quantizing the small coefficients to zero [54, 88] (that is, the quantizer has a dead-zone), because these coefficients do not need to be transmitted thus reducing the data size. Therefore the quantizer examples presented here have this property and the quantizers are designed just for positive coefficients as we consider that both the atom $\mathbf{g}_k$ and $-\mathbf{g}_k$ belong to the dictionary (see the beginning of section D.1).

A simple quantization approach is to utilize a uniform dead-zone quantizer (UDZQ), for which the quantized coefficients are given by

$$Q[\gamma_n] = \text{round}\left(\frac{\gamma_n}{\text{step}_Q}\right)\text{step}_Q. \tag{D.14}$$

For rate–distortion optimization, different values of step_Q can be tried in order to choose the one that meets an RD criterion, including the selected step_Q in a header.

The $\|\mathbf{x}\|$-UDZQ ($\|\mathbf{x}\|$ dependent uniform dead-zone quantizer) considers that the largest possible value of γ_n is the signal norm $\|\mathbf{x}\|$, this leads to

$$\text{step}_Q = \frac{\|\mathbf{x}\|}{2^{b_{\text{coef}}} - 1} \tag{D.15}$$

(here, b_{coef} is restricted to be a natural number). For $\|\mathbf{x}\|$-UDZQ quantized MP decompositions retrieval, two kinds of information need to be coded/stored: the quantized coefficients $Q[\gamma_n]$ and the atom indices $i(n)$. Thus, the total data rate depends on the number of bits spent to send the coefficients added to number of bits used to send the atom indices. Due to the dead-zone, some coefficients are quantized to zero, and thus instead of M terms only $S \leq M$ terms are sent, where S is the number of coefficients that are not quantized to zero. Thus, it is easy to find a code such the total bit rate of the coded M-term representation is

$$R = S\left[\log_2(\#_{\mathcal{D}}) + b_{\text{coef}}\right] + b_{\text{header}}, \tag{D.16}$$

where b_{coef} is the number of bits of the quantizer and $\#_{\mathcal{D}}$ is the dictionary cardinality (since all the coefficients are assumed to be positive $\mathbf{g}_k$ and $-\mathbf{g}_k$ should belong to $\mathcal{D}$). The header includes any relevant decoding information, and b_{header} is the header length (in bits). Assuming that 32-bit precision is used for $\|\mathbf{x}\|$, that b_{coef} can vary from 1 to B, and that the number of coded terms S varies from 1 to M, the header length is $32 + \lceil \log_2(B) \rceil + \lceil \log_2(M) \rceil$ bits.

We now introduce the max(γ)-UDZQ, a uniform dead-zone quantizer, that employs in its design the largest coefficient value of the M-term representation of the signal $\mathbf{x}$ instead of the signal norm. Thus, the max(γ)-UDZQ employs

$$\text{step}_Q = \frac{\max_{n\in[1,\ldots,M]} \gamma_n}{2^{b_{\text{coef}}} - 1}. \tag{D.17}$$

The largest coefficient does not need to be quantized but is sent in the header with negligible error instead. Therefore, when the max(γ)-UDZQ is used to quantize an M-term representation, the total bit rate is given by

$$R = S\left[\log_2(\#_{\mathcal{D}})\right] + (S-1)b_{\text{coef}} + b_{\text{header}}, \tag{D.18}$$

which is smaller than the rate of the $\|\mathbf{x}\|$-UDZQ. This is so because the header of the max(γ)-UDZQ coded decomposition has the same length as the $\|\mathbf{x}\|$-UDZQ, with $\|\mathbf{x}\|$ replaced by max(γ) as side information. The max(γ)-UDZQ has a better RD performance than the $\|\mathbf{x}\|$-UDZQ, because the largest energy atom is reconstructed exactly, and no data rate is spent to send its quantized coefficient. In addition, the dynamic range of the max(γ)-UDZQ is smaller than the dynamic range of the $\|\mathbf{x}\|$-UDZQ, reducing the quantization errors for every coefficient γ_n, in average, and

thus the overall distortion. The simple trick of using $\max(\gamma)$ to define the quantizer dynamic range instead of $\|\mathbf{x}\|$ improves the RD performance, as can be seen in subsection D.4.1.

Delayed On-Line Coding Using Off-loop Quantization The $\max(\gamma)$-UDZQ is applicable only after all the M terms of the representation are obtained since the maximum absolute value among the coefficients of the M-term representation is required. Thus, while the $\|\mathbf{x}\|$-UDZQ can be used in both on-line and off-line coding schemes, the $\max(\gamma)$-UDZQ would be only applicable for off-line coding. For a given dictionary, Theorem D.1 gives the number of MP iterations m guaranteeing that $|\gamma_{n+m}| < |\gamma_n|$. With this result an m-delayed on-line coding scheme using the $\max(|\gamma|)$-UDZQ is possible: since after m steps the largest coefficient is already available one can design the $\max(\gamma)$-UDZQ based just on the first m coefficients and start the quantization process.

D.2.4 State-of-the-art for MP Coefficients Quantization

The state-of-the-art for off-loop quantization of MP coefficients is presented in [54]. There, uniform quantizers whose number of quantization levels and range adapt according to the coefficients of the M-term representation are used. Here, this quantization scheme is referred to as adaptive bounded uniform quantizer (ABUQ). The ABUQ implements a bit-allocation per coefficient that relies on the known result that MP coefficient magnitudes decrease on average at each MP iteration. In the ABUQ, previous to quantization all the coefficients need to be sorted in decreasing magnitude order and the ABUQ is fed with the coefficients in this order. For each coefficient, the ABUQ employs a uniform quantizer of different range and number of levels; the quantizer range for the l^{th} coefficient depends on the actual quantized value of the $(l-1)^{th}$ coefficient, and the number of levels of each coefficient quantizer is decided using a criterion based on a Lagrangian multiplier – a bit-allocation procedure. For decoding the quantized decomposition the number of bits used to quantize the second coefficient in the decreasing magnitude order as well as the larger coefficient are sent as side information.

Delayed On-Line Coding Note that the result in Theorem D.1 can also be applied in the ABUQ, as was employed before for the UDZQ, in order to permit delayed on-line coding using the ABUQ.

Improvements The bit-allocation scheme of the ABUQ fails for dictionaries that have a large $\Theta(D)$ as is often the case for dictionaries in spaces of large dimension, i.e., $N \geq 64$. Section D.4.3 discusses this behavior. We propose a simple modification to improve the rate–distortion performance of the ABUQ. The modification is to allow the bit-stream generated by the ABUQ to be cut at any point, corresponding to any number of coded terms, what can provide precise rate control.

D.2.5 Error Resilience

For transmission, it is important that if coefficients and/or atom indices are lost, the decoder can just successfully ignore the lost terms when reconstructing the signal. Some MP coefficient quantization schemes depend on the value of the quantized coefficient $Q[\gamma_n]$ to derive the quantizer applied to the coefficient γ_{n+1}, as for example the ABUQ (see section D.2.4). In such cases, a lost coefficient would impair the recovery of all the subsequent coefficients. Therefore, in order to improve error resilience, it would be highly desirable that the quantizer for a given γ_m be independent of the quantized values of another γ_n $(m \neq n)$.

D.2.6 Proposed Quantizer

In this work, off-loop quantization is used as it does not influence the signal model, that is, the atoms employed to represent the signal are independent of the quantization employed. In addition, in off-loop quantization, the quantization rule can be easily optimized for RD constraints. Also, if a signal needs to be sent several times with different distortions and/or rates then distinct quantizers can be used to obtain different compressed signal versions. This avoids the necessity to compute several signal decompositions, one for each quantized signal version, as would be the case for in-loop quantization. Also, the same quantization rule is applied to all coefficients of the M-term representation, that is the quantization of γ_{m+1} does not depend on the quantization of γ_m. This provides some error resilience capability

indeed.

D.2.7 Distortion Due to Off-Loop Quantization

When the M-term representation is quantized off-loop each coefficient γ_n is replaced by its quantized version $Q[\gamma_n]$ and the signal approximation is retrieved using

$$\hat{\mathbf{x}}_q = \sum_{n=1}^{M} Q[\gamma_n]\mathbf{g}_{i(n)}. \tag{D.19}$$

There are alternatives for the distortion criterion to use in MP quantizer design:

i) the quantization error relative to the actual signal

$$\mathbf{d} = \mathbf{x} - \hat{\mathbf{x}}_q; \tag{D.20}$$

That leads to the energy distortion per sample (where N is the signal dimension)

$$d^2 = \frac{1}{N}\|\mathbf{d}\|^2 = \frac{1}{N}\|\mathbf{x} - \hat{\mathbf{x}}_q\|^2. \tag{D.21}$$

ii) the quantization error relative to the M-term representation

$$\mathbf{d}_M = \hat{\mathbf{x}} - \hat{\mathbf{x}}_q = \sum_{n=1}^{M}(\gamma_n - Q[\gamma_n])\mathbf{g}_{i(n)}. \tag{D.22}$$

In this work the second error measure is considered since it depends only on the quantization of the coefficients. The energy distortion per sample of the quantized M-term representation is given by

$$d_M^2 = \frac{1}{N}\|\mathbf{d}_M\|^2 = \frac{1}{N}\sum_{n=1}^{M}\sum_{m=1}^{M}(\gamma_n - Q[\gamma_n])(\gamma_m - Q[\gamma_m])\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle, \tag{D.23}$$

where N is the signal length. Noting that $\mathcal{D}$ is composed of unit norm vectors ($\|\mathbf{g}_{i(n)}\| = 1$), and defining the quantization error as

$$e_q(\gamma_m) = \gamma_n - Q[\gamma_n], \tag{D.24}$$

it follows that

$$d_M^2 = \frac{1}{N}\left[\sum_{n=1}^{M} e_q^2(\gamma_n) + \sum_{n=1}^{M}\sum_{m=1,m\neq n}^{M} e_q(\gamma_n)e_q(\gamma_m)\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle\right]. \tag{D.25}$$

If equation (D.25) is equal to zero, i.e., if $\|\mathbf{d}_M\| = 0$ then $\mathbf{d} = \mathbf{x} - \hat{\mathbf{x}}$ becomes the M^{th} residue $\mathbf{r}_{\mathbf{x}}^M$, and d^2 (equation (D.21)) is given by $d^2 = \|\mathbf{r}_{\mathbf{x}}^M\|^2/N$.

Distortion for a Given Source For a given signal source $\mathcal{X}$, one may consider the expected value of d_M^2

$$E[d_M^2] = \frac{1}{N}\left\{\sum_{n=1}^{M} E\left[e_q^2(\Gamma_n)\right] + \sum_{n=1}^{M}\sum_{m=1,m\neq n}^{M} E\left[e_q(\Gamma_n)e_q(\Gamma_m)\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle\right]\right\}. \tag{D.26}$$

Each Γ_n is a random variable (RV) that corresponds to γ_n, for $1 \leq n \leq M$, for signals drawn from $\mathcal{X}$, and thus $E\left[e_q^2(\Gamma_n)\right]$ stands for the expected value of the squared quantization errors of the RV Γ_n.

D.3 Lloyd-Max Quantizers for MP Coefficients

D.3.1 Quantizer Design Using the Angles in MP Iterations

In equation (D.3) the value of $\|\mathbf{x}\|$ is required to compute the coefficients. Alternatively, one can express the coefficients as a function of the first coefficient (γ_1) instead of as a function of $\|\mathbf{x}\|$. This way, equation (D.3) can be rewritten as

$$\gamma_n = \gamma_1\delta_n, \quad \delta_n = \tan(\theta_1)\left[\prod_{i=2}^{n-1}\sin(\theta_i)\right]\cos(\theta_n), \quad n \geq 2. \tag{D.27}$$

Thus, the pdfs of the coefficients γ_n can be computed from the pdfs of the angles θ_n. For a known γ_1, the pdf of the RV Γ_n, for $n \geq 2$, is given by

$$f_{\Gamma_n}(\gamma_n|\gamma_1) = f_{\Gamma_n}(\gamma_1\delta_n|\gamma_1) = f_{\Delta_n}(\delta_n|\gamma_1), \tag{D.28}$$

where Δ_n is the RV whose outcome is δ_n (defined in equation (D.27)). If an optimal quantizer Q is designed for the RV Y, then the optimal quantizer for $Z = cY$ (c is a constant) is simply a scaled version of Q. Therefore, considering that δ_n is quantized instead of γ_n and that γ_1 is known, from equation (D.27) it follows that equation (D.26) becomes

$$E[d_M^2|\gamma_1] = \frac{{\gamma_1}^2}{N}\left\{\sum_{n=2}^{M} E\left[e_q^2(\Delta_n)\right] + \sum_{n=2}^{M}\sum_{m=2,m\neq n}^{M} E\left[e_q(\Delta_n)e_q(\Delta_m)\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle\right]\right\}. \tag{D.29}$$

Thus, assuming γ_1 to be known, if the pdfs of the RVs Δ_n are known then $E[d_M^2|\gamma_1]$ can be computed for any quantization rule applied to Δ_n. Since the quantization

is applied to δ_n instead of γ_n, the value of γ_1 is required at the decoder. Indeed, in this work γ_1 is used to define γ_n, for $n \geq 2$. Note that the use of γ_1 to compute γ_n ($n \neq 1$) guarantees its correct value at the decoder and also reduces the data rate (assuming that the same number of bits would be spent to transmit both $\|\mathbf{x}\|$ or γ_1).

For designing a quantizer using equation (D.29) the pdfs of Δ_n are needed, these can be computed from the RVs Θ_i, $i \in [1, \ldots, n]$, corresponding to the angles between the residues and the atoms selected in MP steps.

For a given γ_1 the quantization of MP coefficients should aim to minimize the distortion per sample of the quantized M-term representation defined in equation (D.29). As stated above, in this case, one should design quantizers for Δ_n, $n \geq 2$. An interesting property is that if the same quantizer is applied to all the coefficients then we can have error resilience to coefficient losses as compared to the ABUQ (see section D.2.4), where the quantization of one coefficient depends on the quantization of the previous one. Taking this into consideration, in this work, we design a single Lloyd-Max quantizer for $\Delta = \cup_{n=2}^{M} \Delta_n$, the RV given by the union of the Δ_n ($n \geq 2$). This way, the designed quantizer will minimize the quantization error of the ensemble of all M terms. Since MP iterations are disjoint events, the pdf of $\Delta = \cup_{n=2}^{M} \Delta_n$ is given by

$$f_{\Delta}(\delta) = \frac{1}{M-1} \sum_{n=2}^{M} f_{\Delta_n}(\delta_n). \tag{D.30}$$

D.3.2 Distortion for an Optimal Quantizer

The distortion per sample in equation (D.29) has two terms. The first term is the sum of the squared quantization errors of Δ_n, whereas the second contains a sum of inner products between dictionary atoms weighted by the products of the quantization errors of the atoms involved in the inner products. As verified, in section D.1, the RVs of the angles Θ_n can be assumed to be uncorrelated. Although Δ_n and Δ_m may be correlated, when designing a quantizer for $\Delta = \cup_{n=2}^{M} \Delta_n$ the assumption that the quantization errors

$$e_q(\Delta_n) = \Delta_n - Q[\Delta_n] \tag{D.31}$$

are uncorrelated is reasonable. It is also reasonable to assume that the quantization errors products $e_q(\Delta_n)e_q(\Delta_m)$ are not correlated to the inner products $\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)} \rangle$. Using these assumptions the second term in equation (D.29) becomes

$$\sum_{n=2}^{M} \sum_{m=2, m\neq n}^{M} E\left[e_q(\Delta_n)\right] E\left[e_q(\Delta_m)\right] E\left[\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)} \rangle\right] \tag{D.32}$$

The atoms selected at different MP steps may be correlated. But due to the invariant nature of the angles statistics in different MP steps, it is plausible consider the expected value of the inner product between the atoms selected in any two different MP steps n and m to be also invariant. More specifically, we consider that $E\left[\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)} \rangle\right] = c$. Therefore, from equation (D.30) and (D.31) one has that

$$E\left[e_q(\Delta)\right] = \frac{1}{M-1} \sum_{n=2}^{M} E\left[e_q(\Delta_n)\right], \tag{D.33}$$

and equation (D.32) yields

$$\sum_{n=2}^{M} E\left[e_q(\Delta_n)\right] \sum_{m=2, m\neq n}^{M} E\left[e_q(\Delta_m)\right] c = \quad (M-1) E\left[e_q(\Delta)\right] \sum_{m=2, m\neq n}^{M} E\left[e_q(\Delta_m)\right] c. \tag{D.34}$$

Since Lloyd-Max quantizers are unbiased estimates of the input then the Lloyd-Max quantizer for Δ leads to $E[e_q(\Delta)] = 0$, making the expression above to vanish. As a result equation (D.29) becomes

$$E[d_M^2|\gamma_1] = \frac{{\gamma_1}^2}{N} \sum_{n=2}^{M} E[(\Delta_n - Q[\Delta_n])^2], \tag{D.35}$$

The above result is a sum of terms $E[(\Delta_n - Q[\Delta_n])^2]$, thus

$$E[d_M^2|\gamma_1] = \frac{\gamma_1^2}{N} \sum_{n=2}^{M} \int (\delta_n - Q[\delta_n])^2 f_{\Delta_n}(\delta_n) d\delta_n. \tag{D.36}$$

Since the same rule $Q[\cdot]$ is applied to all Δ_n, $2 \leq n \leq M$, and $\Delta = \cup_{n=2}^{M} \Delta_n$

$$E[d_M^2|\gamma_1] = \frac{M-1}{N} \int (\delta - Q[\delta])^2 f_{\Delta}(\delta) d\delta, \tag{D.37}$$

where $f_\Delta(\delta) = \frac{1}{M-1} \sum_{n=2}^{M} f_{\Delta_n}(\delta_n)$ and the $M-1$ multiplying the integral takes into account that $M-1$ coefficients are quantized. With the simplification above, applying the same quantization rule to MP coefficients entails the design of the optimal quantizer for $\Delta = \cup_{n=1}^{M} \Delta_n$, what is accomplished by Lloyd-Max quantizers [73].

D.3.3 Lloyd-Max Quantizer for MP Coefficients

From the simplification above, the aim of the quantizer design is to minimize

$$E[d_M^2|\gamma_1] = \frac{{\gamma_1}^2(M-1)}{N}\text{MSE}(\Delta), \tag{D.38}$$

$$\text{where MSE}(\Delta) = E\left[(\Delta - Q[\Delta])^2\right], \tag{D.39}$$

$$\text{subject to } f_\Delta(\delta) = \frac{1}{M-1}\sum_{n=2}^{M} f_{\Delta_n}(\delta_n), \quad \text{i.e. } \Delta = \cup_{n=2}^{M}\Delta_n; \tag{D.40}$$

$$\text{where } \Delta_n = \tan(\Theta_1)\left[\prod_{i=2}^{n-1}\sin(\Theta_i)\right]\cos(\Theta_n). \tag{D.41}$$

The minimization of equation (D.38) is obtained by a Lloyd-Max quantizer [73]. The quantizer design requires $f_{\Delta_n}(\delta_n)$, for $2 \leq n \leq M$, these are obtained from $f_{\Theta_1}(\theta_1)$. In turn, from section D.1, we see that $f_{\Theta_1}(\theta_1)$ can be estimated by applying one MP step to a large set of signals drawn from a Gaussian source. Then $f_{\Delta_n}(\delta_n)$ are estimated from $f_{\Theta_1}(\theta_1)$, which is an acceptable procedure since all Θ_n have similar statistics. The estimated $f_{\Delta_n}(\delta_n)$ are then used to calculate $f_\Delta(\delta)$, and $f_\Delta(\delta)$ is then used to obtain the Lloyd-Max quantizer of b_{coef} bits ($L = 2^{b_{\text{coef}}}$ levels).

The reconstruction levels r_k ($1 \leq k \leq L$) and thresholds t_k ($1 \leq k \leq L+1$) of the L-level Lloyd-Max quantizer are given by [73]

$$t_k = \frac{(r_k + r_{k-1})}{2}, \quad \text{and} \quad r_k = \frac{\displaystyle\int_{t_k}^{t_{k+1}} \delta f_\Delta(\delta)d\delta}{\displaystyle\int_{t_k}^{t_{k+1}} f_\Delta(\delta)d\delta}. \tag{D.42}$$

The thresholds and reconstruction levels of the LMQ are calculated using an iterative algorithm [73]. In the calculation, it is imposed that $r_1 = 0$, $t_1 = 0$ (just positive coefficients), and $t_{L+1} = \infty$. The pdf $f_\Delta(\delta)$ required in equation (D.42) is given by equation (D.40) and it is obtained from an estimate of $f_{\Theta_1}(\theta_1)$, which in turn is obtained using the angle in the first MP step for an ensemble of $\#_{\mathcal{D}}N^2$ realizations from Gaussian source.

The same quantizer law is used for all coefficients γ_n, therefore $f_\Delta(\delta)$ varies with M, as consequence each M leads to a different quantizer. The quantizer design is independent of γ_1 and it suffices to design quantizers for γ_1=1, storing copies of the quantizers in both encoder and decoder. The encoder sends γ_1, the number of bits of the quantizer, and the number of terms of the decomposition (M), in a

header, to the decoder. The parameter γ_1 is used to scale the quantizer in both coder and decoder, a simple strategy that makes good use of resources.

D.3.4 Validation of the Distortion Simplification

In the derivation of the Lloyd-Max quantizer for MP coefficients above, the quantization errors were assumed to have no correlation, and $E[\mathbf{g}_i, \mathbf{g}_k]$ to be constant. Those allowed for an approximate expression for the distortion per sample, that in turn is used to design Lloyd-Max quantizers. For the proposed quantizer $e_q^2(\gamma_1) = \gamma_1 - Q[\gamma_1] \approx 0$. Also, we have the simplification from equation (D.29) to equation (D.38) to be valid. Therefore, for a given source, the expected distortion per sample for M coded terms is given by the sum of the mean squared quantization errors (MSE) of the coefficients divided by the signal length

$$\frac{\text{MSE}}{N} = \frac{1}{N}\sum_{i=2}^{M} E\left[\gamma_1^2 e_q^2(\Delta_i)\right] = \frac{1}{N}\sum_{i=2}^{M} E\left[e_q^2(\Gamma_i)\right] = \frac{1}{N}\sum_{i=1}^{M} E\left[e_q^2(\Gamma_i)\right]. \tag{D.43}$$

In equations (D.20) and (D.22), we defined $\mathbf{d} = \mathbf{x} - \hat{\mathbf{x}}_q$ and $\mathbf{d}_M = \hat{\mathbf{x}} - \hat{\mathbf{x}}_q$. Since the MP is such that $\mathbf{x} = \hat{\mathbf{x}} + \mathbf{r}_{\mathbf{x}}^M$, therefore

$$\mathbf{d} = \hat{\mathbf{x}} + \mathbf{r}_{\mathbf{x}}^M - \hat{\mathbf{x}}_q, \text{ and}$$
$$\lim_{M\to\infty} d^2 = \lim_{M\to\infty} \|\mathbf{d}\|^2 = \lim_{M\to\infty} \|\hat{\mathbf{x}} + \mathbf{r}_{\mathbf{x}}^M - \hat{\mathbf{x}}_q\|^2. \tag{D.44}$$

Noting that

$$\|\hat{\mathbf{x}} + \mathbf{r}_{\mathbf{x}}^M - \hat{\mathbf{x}}_q\|^2 = \langle\hat{\mathbf{x}}, \hat{\mathbf{x}}\rangle + 2\langle\mathbf{r}_{\mathbf{x}}^M, \hat{\mathbf{x}}\rangle - 2\langle\hat{\mathbf{x}}, \hat{\mathbf{x}}_q\rangle - 2\langle\mathbf{r}_{\mathbf{x}}^M, \hat{\mathbf{x}}_q\rangle + \langle\hat{\mathbf{x}}_q, \hat{\mathbf{x}}_q\rangle + \langle\mathbf{r}_{\mathbf{x}}^M, \mathbf{r}_{\mathbf{x}}^M\rangle$$

and that $\lim_{M\to\infty} \mathbf{r}_{\mathbf{x}}^M = \vec{0}$ one obtains that

$$\lim_{M\to\infty} d^2 = \lim_{M\to\infty} \|\mathbf{d}\|^2 = \lim_{M\to\infty} \|\hat{\mathbf{x}} - \hat{\mathbf{x}}_q\|^2. \tag{D.45}$$

Therefore, if the simplification from equation (D.29) to equation (D.38) is valid, then one should obtain

$$\frac{1}{N}E\left[\lim_{M\to\infty} d^2\right] = \frac{1}{N}E\left[\lim_{M\to\infty} \|\mathbf{d}\|^2\right] = \frac{1}{N}E\left[\lim_{M\to\infty} \|\hat{\mathbf{x}} - \hat{\mathbf{x}}_q\|^2\right] = \frac{1}{N}E\left[\lim_{M\to\infty} e_q^2[\Gamma]\right]. \tag{D.46}$$

That is,

$$\lim_{M\to\infty} E\left[d^2\right] = \lim_{M\to\infty} \frac{MSE}{N}. \tag{D.47}$$

If equation above is true then the second term in equation (D.29), equation (D.32), can be considered to be null. We have assumed this condition for computing the distortion due to quantization of MP coefficients and the Lloyd-Max quantizer design.

Figure D.14 presents the distortion per sample d^2 of the reconstructed signal with respect to the original one (equation (D.21)) and MSE/N (equation (D.43)) against the number of coded terms M. The results shown are averages over a set of 20,000 MP decompositions of signals drawn from a Gaussian source, using the $GSND(16, 4)$ dictionary. It is possible to verify that as M increases the total distortion approximates the distortion due to the quantization of coefficients. When M is sufficiently large, this distortion is solely due to quantization as the residuals $\mathbf{r}_{\mathbf{x}}^{M}$ become very small. The graph indicates that, in this case, equation (D.47) is valid as both curves are the same, $E[d^2]$ and MSE/N. The same behavior was observed for other dictionaries and for different numbers of quantizer levels. These results corroborate the assumptions made in the simplification of the distortion due to off-loop quantization of MP coefficients.

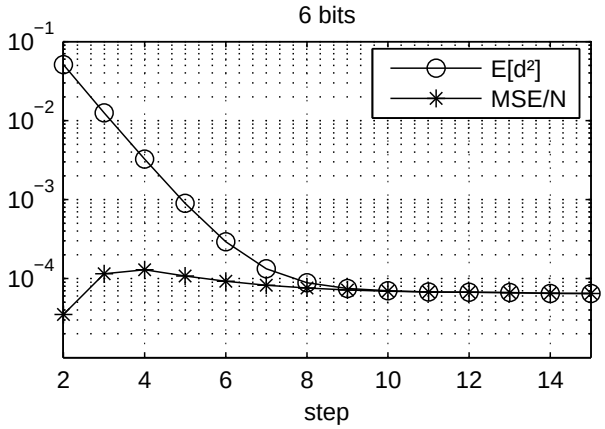


Figure D.14: Expected total distortion per sample (d^2) and coefficient quantization error ($\frac{\text{MSE}}{N}$) for a Gaussian source in $\mathbb{R}^4$ using a 6 bit Lloyd-Max quantizer, and M ranging from 2 to 15, for decompositions obtained using the $GSND(16, 4)$.

D.3.5 Practical Design of Lloyd-Max Quantizers

Since $f_{\Delta}(\delta)$ changes with M, therefore the b_{coef}-bits Lloyd-Max quantizer is expected to change with M. However, after a sufficiently large M the quantizer

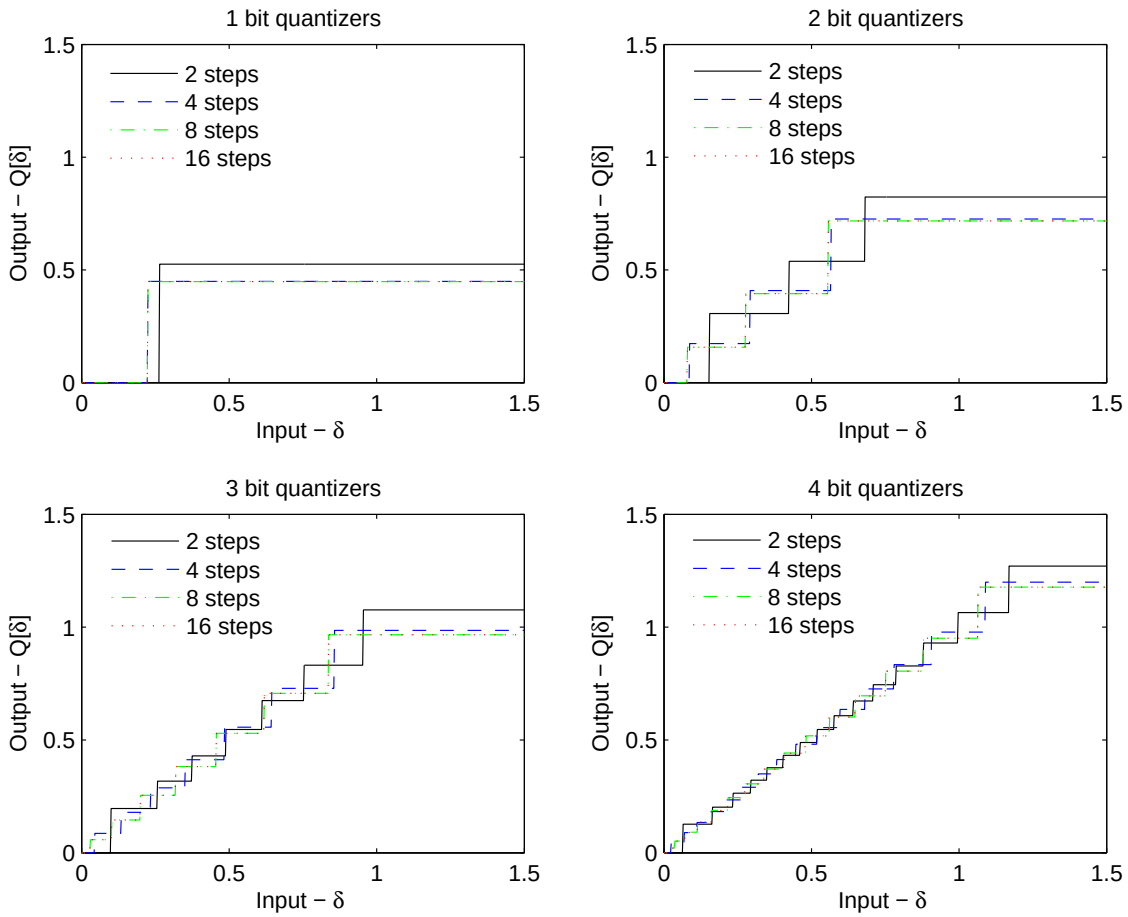


Figure D.15: Coefficients quantizer for a dictionary of 16 vectors uniformly distributed on the 4-dimensional hyper-sphere, for $\gamma_1 = 1$, for different number of terms for the M-term decomposition (2, 4, 8 and 16 terms).

remains almost the same. Figure D.15 shows the Lloyd-Max quantizers obtained for b_{coef} ranging from 1 to 4 bits and decomposition steps $M = \{2, 4, 8, 16\}$, for the $GSND(16, 4)$. Observe that the quantizer thresholds and reconstruction levels seem to be almost invariant after a given number of decomposition steps which depends on b_{coef}. That is, each L-level Lloyd-Max quantizer can be considered to be fixed after a sufficiently large M. Therefore, for practical applications the b_{coef}-bits quantizer does not need to be obtained for each value of M, but just on a range from 2 to J. For $M > J$ the quantizer obtained for $M = J$ can be used without major impacts on the quantizer performance. We justify this claim below.

The Lloyd-Max quantizer design aims to minimize

$$\frac{M-1}{N}\int(\delta - Q[\delta])^2 f_\Delta(\delta)d\delta = \frac{1}{N}\sum_{n=2}^{M}\int(\delta_n - Q[\delta_n])^2 f_{\Delta_n}(\delta_n)d\delta_n = \frac{1}{N}\sum_{n=2}^{M}E[e_q^2(\Delta_n)]. \tag{D.48}$$

From the definition of δ_n, see equation (D.27), and considering that the Θ_i are iid, it can be seen that as n increases the spread of the pdfs of the successive RVs Δ_n becomes smaller and $E[\Delta_n]$ tends to zero. In other words, the pdfs of successive RVs Δ_n get narrower and their centers approximate zero as n increases. This occurs because Δ_n is a product of $n-2$ terms of the form $\sin(\Theta_i)$ multiplied by $\cos(\Theta_n)\tan(\Theta_1)$, thus the expected value (the Θ_i are considered to be iid)

$$E[\Delta_n] = E[\tan(\Theta_1)]\left\{\prod_{i=2}^{n-1}E[\sin(\Theta_i)]\right\}E[\cos(\Theta_n)] = E[\tan(\Theta)]E[\sin(\Theta)]^{(n-2)}E[\cos(\Theta)] \tag{D.49}$$

decreases with n since $E[\sin(\Theta)] < 1$ (actually, note that $\delta_n \leq \sin(\Theta(\mathcal{D}))^{n-1}/\cos(\Theta(\mathcal{D}))$, see equation (D.5) and Theorem D.1). Therefore, the overall probability of each Δ_n being small increases as n increases. Since $\Delta = \cup_{n=2}^{M}\Delta_n$, the probability of Δ belonging to a given interval is the sum of the probabilities of each Δ_n belonging to that interval divided by $M-1$. Thus, the probability of Δ being close to zero increases with M. Sufficiently small δ are quantized to zero due to the quantizer dead-zone, thus for sufficiently large n, the quantization error of δ_n equals the value of δ_n itself. Therefore, as n increases, the contribution of $E[e_q^2(\Delta_n)]$ in equation (D.48) diminishes, that is $\lim_{n\to\infty}E[e_q^2(\Delta_n)] = E[\Delta_n^2] = 0$. As a result the distortion due to coefficient quantization $\sum_{n=2}^{M}E[e_q^2(\Delta_n)]$ changes very little after a given $M = J$, as Figures D.14 and D.17 illustrate. Therefore, after a given $n \geq J$ the RVs Δ_n do not largely contribute to the distortion in equation (D.48), and almost do not influence the quantizer design.

The number of terms $M = J$ for which the quantizer remains almost the same, depends on the pdf of $f_{\Theta_1}(\theta_1)$ (as the Θ_i are considered to be iid and uncorrelated), which in turn depends on the dictionary $\mathcal{D}$. The number of quantization levels L also influences this value of M, because a larger L results in a smaller dead-zone reducing the coefficient range mapped to the first reconstruction level (zero). In addition, as is further discussed in subsection D.4.2, there is little gain by indefinitely increasing the number of decomposition terms in the quantizer design. Therefore, to improve

the quantizer performance the number of quantization levels should be augmented.

Figure D.16 shows 3-bit quantizers obtained for different number of terms M for the Gabor dictionary in $\mathbb{R}^{64}$. In this figure the quantizer "convergence", discussed above, can also be observed.

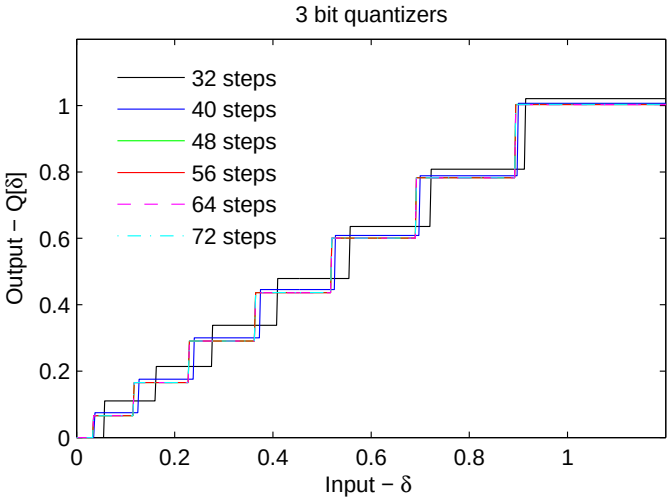


Figure D.16: 3-bit coefficient quantizers found for the Gabor dictionary with 4 phases in $\mathbb{R}^{64}$ for different number of terms for the M-term decomposition.

D.4 Performance of the Lloyd-Max Quantizers

D.4.1 Comparison to Other Fixed Quantizers

The Lloyd-Max quantizer is designed to be applied to all the coefficients of an M-term decomposition, i.e., it is the same (fixed) for all the coefficients of an M-term representation. Therefore, the first evaluation that will be held is to compare the Lloyd-Max quantizers to other fixed quantizers. For that purpose the Lloyd-Max quantizer performance is compared to the performances of both the $\|\mathbf{x}\|$-UDZQ and the $\max(\gamma)$-UDZQ which were described in subsection D.2.3. Figure D.17 shows the expected total distortion per sample $E[d^2]$, defined in equation (D.21), and also the mean squared error per sample ($\frac{\text{MSE}}{N} = \frac{1}{N}\left[\sum_{n=1}^{M} e_q^2(\gamma_n)\right]$) for 4-bit quantizers as a function of the number of terms of the quantized representation. The results presented are averages over an ensemble of 20,000 signals from a Gaussian source, using the $GSND(16, 4)$. As can be seen, the $\max(\gamma)$-UDZQ has superior performance than the $\|\mathbf{x}\|$-UDZQ, and the Lloyd-Max quantizer yields the best results.

Figure D.18 presents $E[d^2]$ against the number of coded terms M for quantizers using 2, 4, 6 and 8 bits, where it is also shown the expected distortion of the unquantized M-term representation. Since for small number of decomposition terms the residues are large there is not much difference among the three quantizers for small values of M. However, as the number of terms increases the residues decrease and the Lloyd-Max quantizer has better performance. Figure D.19 shows the same graphs for quantized M-term representations obtained using the $\epsilon_{8_{sh1}}$ utilizing quantizers with bit depth varying from 1 to 6; the results are averages over 50,000 signals from a memoryless Gaussian source. It is possible to note a similar behavior for the distortions there presented to the ones observed in Figure D.18. Quantizers designed with various number of levels for different dictionaries have all produced similar results. This leads to the conclusion that the max(γ)-UDZQ always outperforms the $\|\mathbf{x}\|$-UDZQ, an expected result – see subsection D.2.3, and that the Lloyd-Max outperforms both. As the number of coded terms increases the Lloyd-Max quantizer has an average distortion around 50% lower than the one obtained for the max(γ)-UDZQ, providing an average gain in the signal to noise ratio of 3dB for the Lloyd-Max quantizer as compared to the max(γ)-UDZQ.

Figure D.20 shows $E[d^2]$ and $\frac{\text{MSE}}{N}$ for the 4-bit Lloyd-Max quantizer, the 4-bit max($|\gamma|$)-UDZQ, and the 4-bit $\|\mathbf{x}\|$-UDZQ. Again, it is noticeable that the Lloyd-Max quantizer has a distortion that is below 50% of the distortion of the max(γ)-UDZQ. Note that due to the shape of $f_{\Theta_1}(\theta_1)$ the number of steps for which $E[d^2]$ equals MSE/N increases as compared to the results obtained for the $GSND(16, 4)$.

D.4.2 Rate–Distortion Optimization

The number of terms M and the number of bits b_{coef} define the Lloyd-Max quantizer. It is easy to find a code such that the data rate of the Lloyd-Max quantized MP decomposition is

$$R = M\left[\log_2\left(\#_{\mathcal{D}}\right)\right] + (M-1)\, b_{\text{coef}} + b_{\text{header}}. \tag{D.50}$$

For optimizing the overall RD performance, once the signal decomposition is obtained, several number of quantization levels $2^{b_{\text{coef}}}$ and decomposition terms M can be tried in order to meet a prescribed RD criterion. In other words, for each rate R one

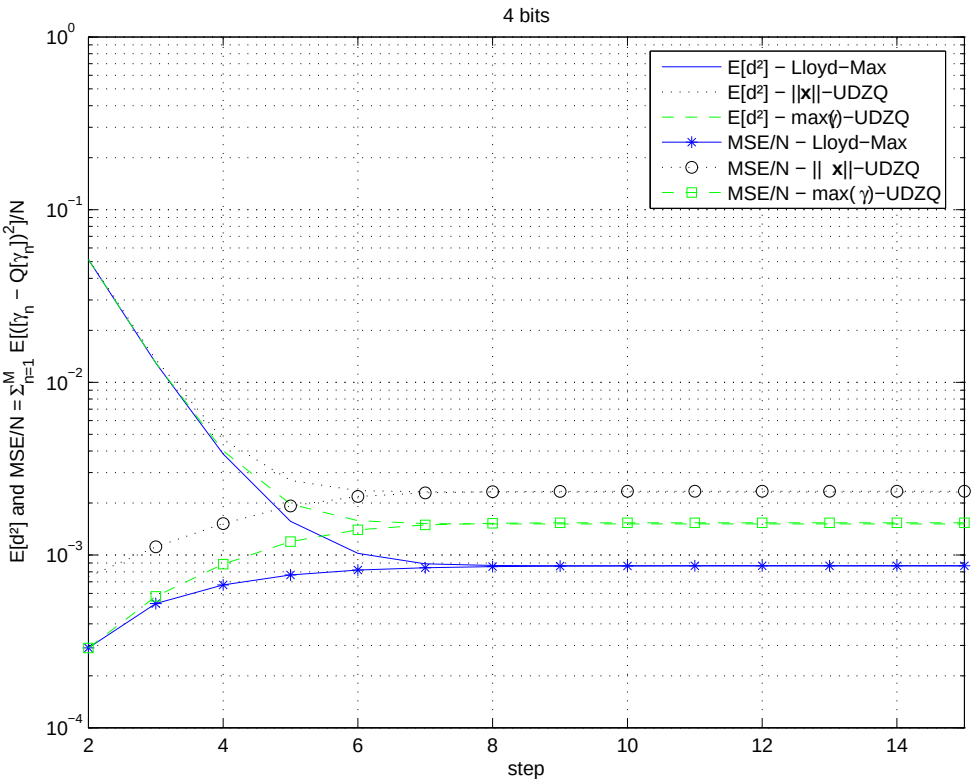


Figure D.17: Expected values of the distortion and of the sum of coefficient quantization errors, for 4-bit quantizers for the Lloyd-Max quantizer, the max(γ)-UDZQ and the $\|\mathbf{x}\|$-UDZQ, for M ranging from 2 to 15, for the $GSND(16, 4)$.

finds the (M, b_{coef}) pair that leads to the lowest distortion. This procedure can be applied off-line, searching for the pairs (M, b_{coef}) that lead to the best average RD performance for an ensemble of signals from a given source. This procedure is presented in Algorithm D.1. A graphical view of the expected result of this procedure is depicted in Figure D.21.

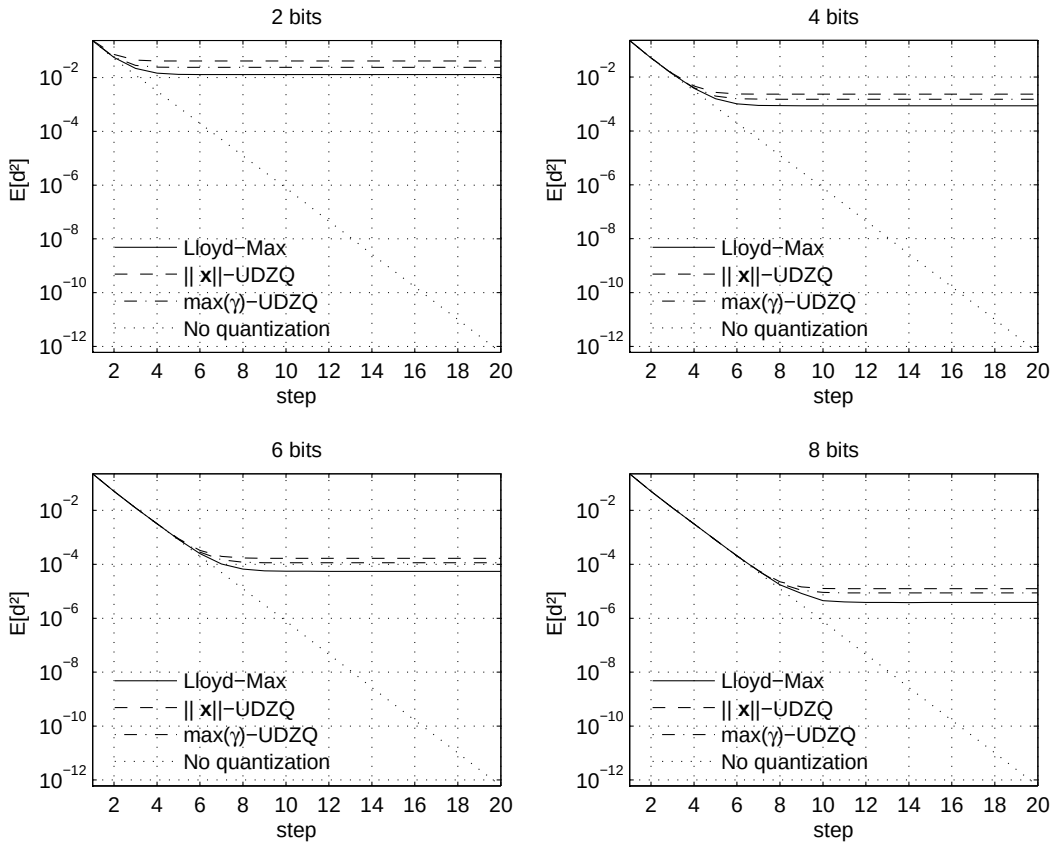


Figure D.18: Comparison of the expected distortion of the $\|\mathbf{x}\|$-UDZQ, $\max(\gamma)$-UDZQ, and Lloyd-Max quantizers for the $GSND(16, 4)$.

Algorithm D.1 RD optimization algorithm.

1. Design the quantizers for a set of pairs (M, b_{coef});
2. Create an empty table T indexed by (M, b_{coef});
3. For each pair (M, b_{coef}) compute the rate $R(M, b_{\text{coef}})$ and associated expected distortion $D(M, b_{\text{coef}})$ and include them in T;
4. Sort T in increasing rate order keeping the quantizer specifications, that is the pair (M, b_{coef}), and its correspondent distortion $D(M, b_{\text{coef}})$; That is, index T as function of the pairs (R, D);
5. For each rate R in T, retain in T only the quantizer (M, b_{coef}) leading to the smallest distortion D.
6. For each $(R, D) \in T$ if exists $(R', D') \in T$ such that $R' \leq R$ and $D' \leq D$ eliminate (R, D) from T;
7. T contains the quantizer specifications (M, b_{coef}) leading to the best RD performance.

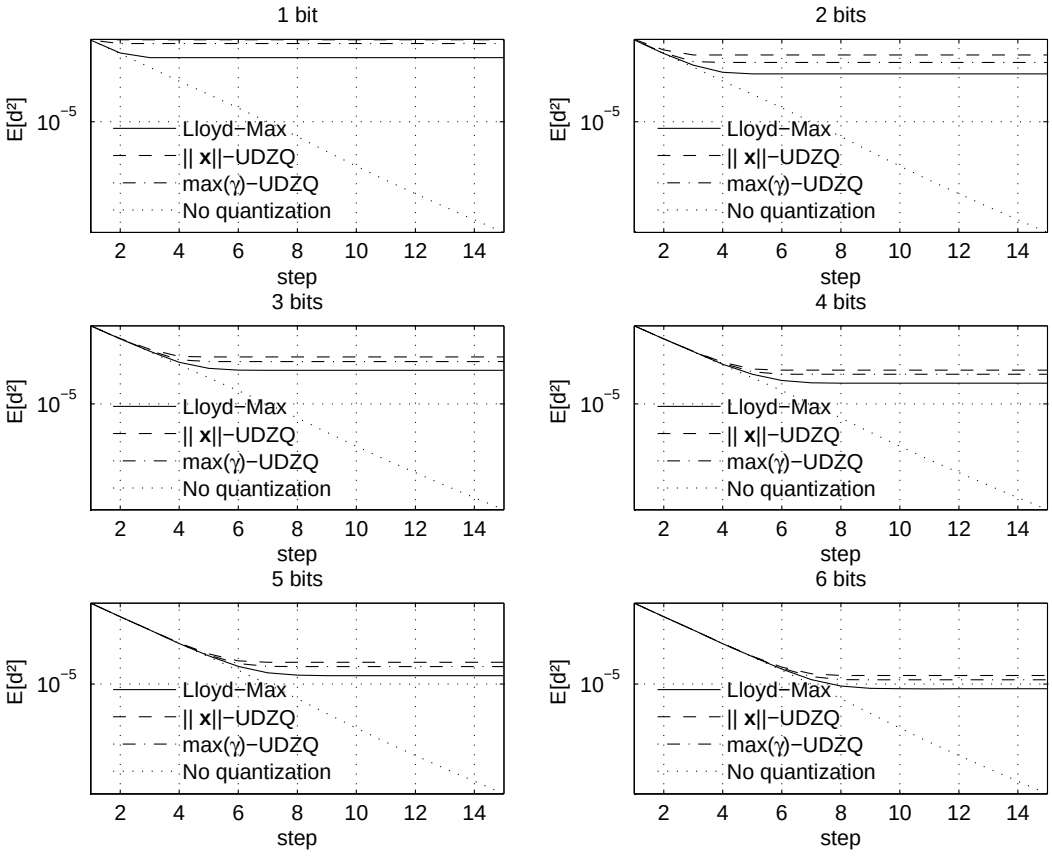


Figure D.19: Comparison of the expected distortion of the $\|\mathbf{x}\|$-UDZQ max(γ)-UDZQ and Lloyd-Max quantizers for the $\epsilon_{8_{sh1}}$ dictionary.

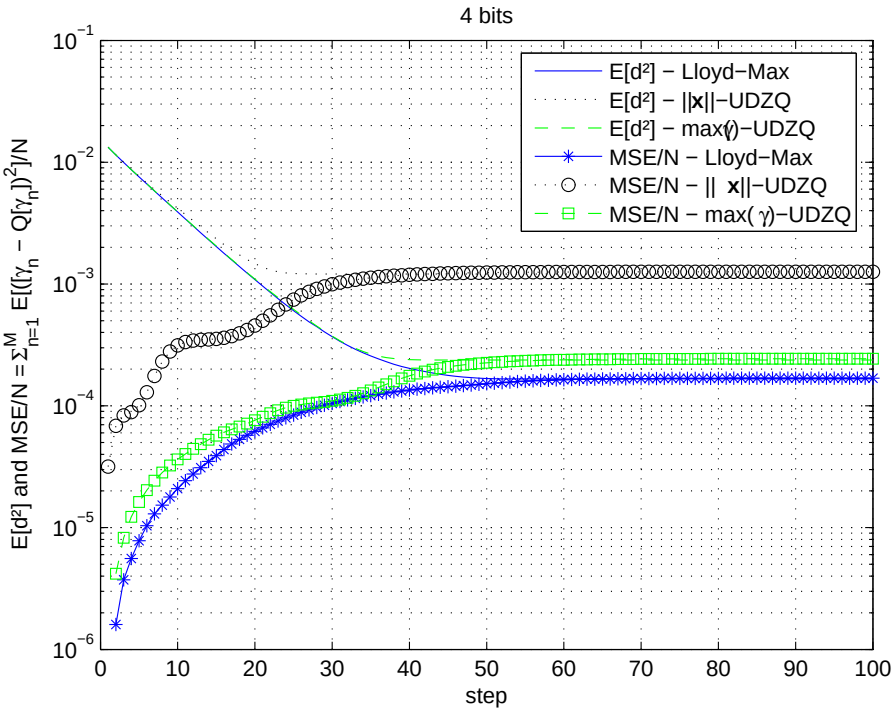


Figure D.20: Total distortion versus quantization error, for the Lloyd-Max quantizer, the max(γ)-UDZQ and the $\|\mathbf{x}\|$-UDZQ, all using 4 bits, for the 4-phase Gabor dictionary in $\mathbb{R}^{64}$ – $d^2 = \frac{1}{N}\|\mathbf{x} - \hat{\mathbf{x}}_q\|^2$ and $\frac{\text{MSE}}{N} = \frac{1}{N} E\left[\sum_{n=1}^{M} e_q^2(\gamma_n)\right]$.

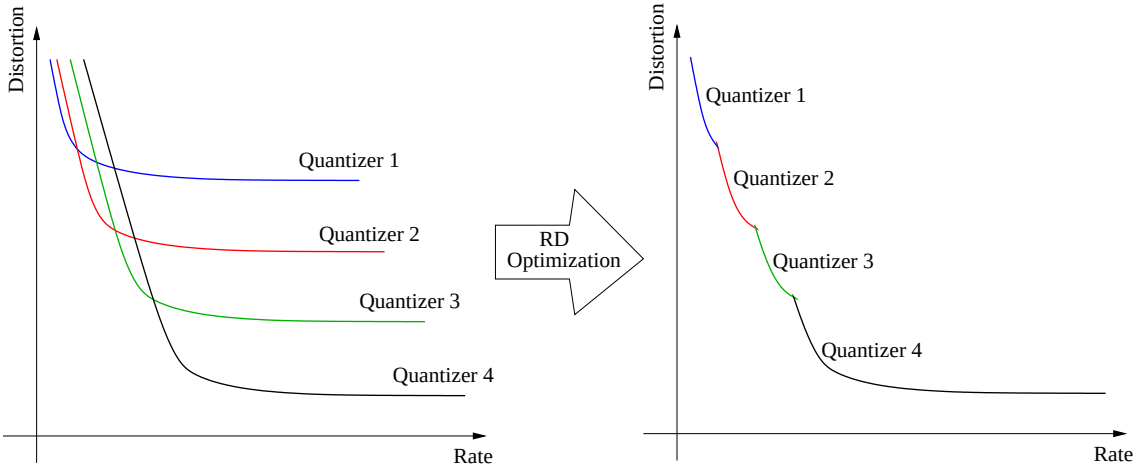


Figure D.21: Expected result of the RD optimization algorithm presented in Algorithm D.1.

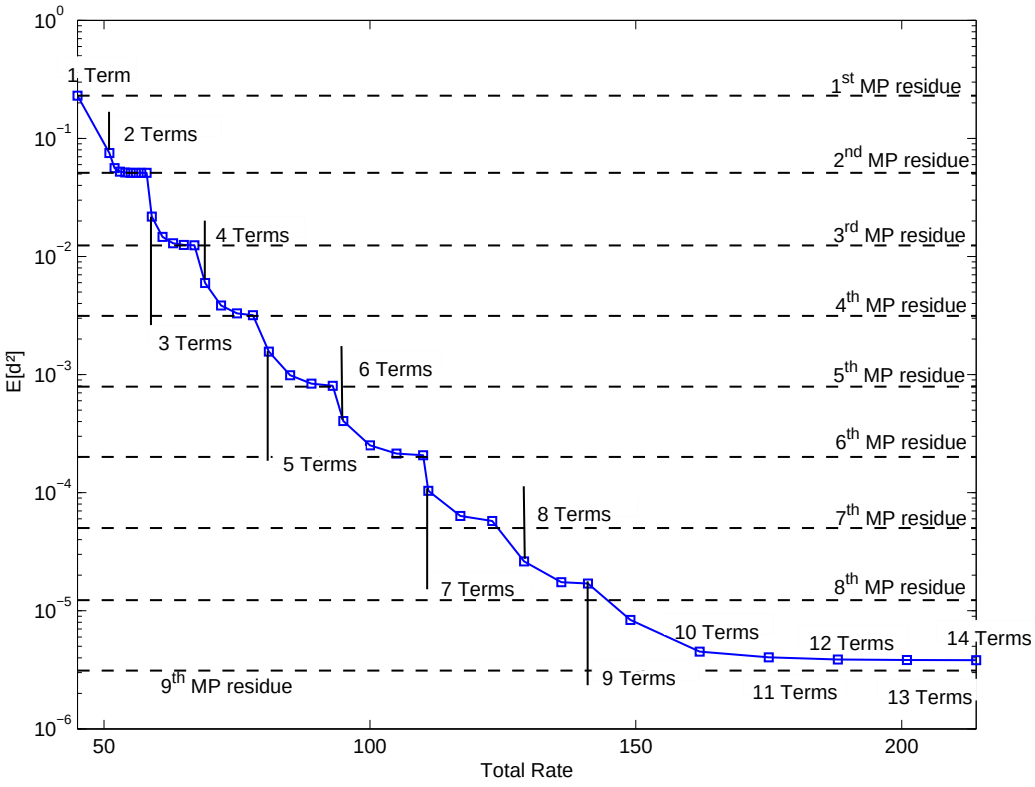


Figure D.22: RD performance for the 16 Gaussian elements dictionary in $\mathbb{R}^4$ explicitly showing the numbers of terms selected after the RD procedure (M) and the expected values of the residue norms (in dashed lines).

Analysis of the RD Optimization Figure D.22 shows the resulting RD operational curve for the $GSND(16, 4)$ for a Gaussian Source resulting from the RD optimization in Algorithm D.1. Figure D.22 shows also the average values of the MP residue norms at the steps (plotted in dashed lines), and the number of coded terms M in each point of the RD curve. Note that for $M = 7$ the quantization proposed achieves an $E[d^2]$ that is similar to the expected residue norm of the unquantized 7-step decompositions. However, after $M = 8$ none of the quantizers employed can achieve the expected value of the residue norm. That is, for $M \geq 8$ it would have been wiser to augment the quantizer bit depth. That is, as M increases b_{coef} should also increase, what is indeed quite obvious. For example, the results indicate that for $M = 4$ a 6-bit quantizer is enough to achieve a distortion that is similar to the expected norm of the residue after 4 decomposition steps, however the same results indicate that for $M = 6$ an 8-bit quantizer is required for obtaining a distortion close to the residue norm. This is quite obvious, since, in general, the coefficient values

decrease as the step increases, and to be capable of representing smaller coefficient values a higher quantizer resolution is required.

Table D.1: Number of bits of the quantizer (b_{coef}) and coded terms (M) selected by the RD optimization using the Lloyd-Max quantizer, for the $GSND(16, 4)$.

Total Rate - b_{header}	M	b_{coef}	Total Rate - b_{header}	M	b_{coef}	Total Rate - b_{header}	Steps	Bits
5	1	0	27	3	6	71	7	6
11	2	1	29	4	3	77	7	7
12	2	2	32	4	4	83	7	8
13	2	3	35	4	5	89	8	7
14	2	4	38	4	6	96	8	8
15	2	5	41	5	4	101	9	7
16	2	6	45	5	5	109	9	8
17	2	7	49	5	6	122	10	8
18	2	8	53	5	7	135	11	8
19	3	2	55	6	5	148	12	8
21	3	3	60	6	6	161	13	8
23	3	4	65	6	7	174	14	8
25	3	5	70	6	8	–	–	–

Table D.1 shows the values of b_{coef} and M that arise from the RD optimization of the $GSND(16, 4)$ as a function of the total bit-rate (not counting the header length) for b_{coef} varying from 2 to 8 and M varying from 1 to 20. Note that, in general, the pairs (S, b_{coef}) selected for the source RD are such that increments on the bit depth are selected only if the resulting increase in the data rate is smaller than the one incurred in augmenting the number of terms. For example, the pair $(3, 6)$ corresponds to the rate of 27 bits. If the quantizer resolution is increased to 7 bits then the rate is $3\left[\log_2(16) + 1\right] + (2)\,7 = 29$. However, note that the RD procedure chooses the $(4, 3)$ pair instead of the $(3, 7)$ pair, preferring to increase the number of coded terms rather than the quantization resolution. However, this is not always the case. The pair $(4, 2)$ leads to a data rate in between the data rates corresponding to the pairs $(3, 5)$ and $(3, 6)$. However the pair $(4, 2)$ is not selected by the RD optimization; why this choice does not occur? The average energy of the quantization errors of the second, third and fourth coefficients using a 2-bit quantizer is obviously larger than the average energy of the second and third

coefficients when the 5-bit or the 6-bit quantizer are used. If the total distortion of coding the second, the third and the fourth coefficients using the 2-bit quantizer had been larger than the total distortion of coding the second and third coefficients using the 6-bit quantizer then the pair $(4, 2)$ would have been selected. However, the sum of the energies of the second and third coefficients' quantization errors for the 2-bit quantizer is larger than the average energy of the non-quantized fourth coefficient. This explains why the pair $(4, 2)$ is not selected by the RD optimization. Observe that for the range allowed for b_{coef} (2 to 8) the RD procedure does not select any pair (M,b_{coef}) with $M > 14$ because the coefficients after the 14^{th} step are quantized to zero for all the b_{coef} allowed and these coefficients would just waste rate.

Table D.2 shows the pairs (S, b_{coef}) that result from the RD optimization procedure for $b_{\text{coef}} \in [1, \ldots, 6]$ and steps ranging from 1 to 100 for the Gabor dictionary of 4 phases in $\mathbb{R}^{64}$. This dictionary has a large cardinality and thus the impact in the data rate due to the increase of M is larger than for the $GSND(16, 4)$. In addition, the coefficients decay slower for this dictionary than for the $GSND(16, 4)$, as discussed in the end of section D.1.3. In some cases, these aspects make the RD optimization procedure, for this dictionary, to decrease the number of terms M used from one point of the RD curve to the next while b_{coef} is increased, as can be seen in Table D.2. Note that, due to the slow decay rate of the coefficients for the $\mathbb{R}^{64}$ Gabor dictionary the resulting RD pairs occupy all the allowed range for M differently of what occurs for the $GSND(16, 4)$.

On-line Quantization The (M, b_{coef}) pairs that give the best points in the RD curve for the Gaussian source can be stored in both coder and decoder. Then, for transmission of a given signal, the coder can simply choose the pair that meets the desired RD criterion. This allows on-line quantization of MP coefficients, i.e., the quantizer can be selected prior to the signal decomposition to meet an RD criterion and each coefficient can be quantized as soon as it is available.

Table D.2: Number of bits of the quantizer (b_{coef}) and coded terms (M) selected by the RD optimization using the Lloyd-Max quantizer, for the Gabor dictionary in $\mathbb{R}^{64}$ of 4 phases.

Total Rate - b_{header}	M	b_{coef}	Total Rate - b_{header}	M	b_{coef}	Total Rate - b_{header}	Steps	Bits
13	1	0	279	20	2	877	49	6
26	2	1	283	19	3	895	50	6
27	2	2	293	21	2	913	51	6
28	2	3	298	20	3	931	52	6
29	2	4	307	22	2	949	53	6
30	2	5	313	21	3	967	54	6
31	2	6	328	22	3	985	55	6
39	3	1	343	23	3	1003	56	6
41	3	2	358	24	3	1021	57	6
43	3	3	373	25	3	1039	58	6
45	3	4	388	26	3	1057	59	6
47	3	5	397	25	4	1075	60	6
49	3	6	403	27	3	1093	61	6
52	4	1	413	26	4	1111	62	6
55	4	2	418	28	3	1129	63	6
58	4	3	429	27	4	1147	64	6
61	4	4	433	29	3	1165	65	6
64	4	5	445	28	4	1183	66	6
65	5	1	448	30	3	1201	67	6
69	5	2	461	29	4	1219	68	6
73	5	3	477	30	4	1237	69	6
77	5	4	493	31	4	1255	70	6
78	6	1	509	32	4	1273	71	6
83	6	2	525	33	4	1291	72	6
88	6	3	541	34	4	1309	73	6
91	7	1	557	35	4	1327	74	6
97	7	2	573	36	4	1345	75	6
103	7	3	589	37	4	1363	76	6
104	8	1	605	38	4	1381	77	6
111	8	2	621	39	4	1399	78	6
117	9	1	637	40	4	1417	79	6
125	9	2	642	38	5	1435	80	6
130	10	1	653	41	4	1453	81	6
139	10	2	659	39	5	1471	82	6
148	10	3	669	42	4	1489	83	6
153	11	2	676	40	5	1507	84	6
163	11	3	693	41	5	1525	85	6
167	12	2	710	42	5	1543	86	6
178	12	3	727	43	5	1561	87	6
181	13	2	744	44	5	1579	88	6
193	13	3	761	45	5	1597	89	6
195	14	2	778	46	5	1615	90	6
208	14	3	795	47	5	1633	91	6
209	15	2	812	48	5	1651	92	6
223	16	2	829	49	5	1669	93	6
237	17	2	841	47	6	1687	94	6
251	18	2	846	50	5	1705	95	6
265	19	2	859	48	6	1723	96	6
268	18	3	863	51	5	1795	100	6

D.4.3 Comparison to the State of the Art

It is easy to find a code such that the total rate of MP quantized decomposition is given by

$$R = S\left[\log_2\left(\#_{\mathcal{D}}\right)\right] + r_{\text{coef}} \tag{D.51}$$

where S is the number of terms that remain after quantization (i.e., that have quantized coefficients that are not zero), and r_{coef} is the rate incurred in coding the quantized coefficients. Therefore, the rate in bits per sample is given by R/N, where N is the signal space dimension. The strategy employed to generate the ABUQ (see section D.2.4) coded bit-stream is to entropy code the differences between the quantization indices of successive coefficients. In the comparisons between Lloyd-Max quantization and ABUQ presented in the sequel, this strategy is also employed to code Lloyd-Max quantized M-term representations. Figure D.23 shows the RD curves of quantized MP expansions originated from three different random sources in $\mathbb{R}^{10}$ (a memoryless Gaussian, a memoryless uniform and a memoryless Gamma distributed sources) using both the Lloyd-Max quantization and the ABUQ for a $GSND(128, 10)$. For each distinct source the results are averages over an ensemble of 100 quantized MP decompositions of signals from each source. For this experiment the LMQs were designed with bit-depth ranging from 1 to 8. It can be seen in Figure D.23 that both quantizers have similar performance for all the three signal sources; however the LMQ tends to be better at low rates (below 8 bits/sample). In addition it can be noted that the LMQ is able to work at lower data rates as show their RD plots.

Figure D.24 shows the RD curves of three different realizations of a Gaussian source when their MP expansions are quantized using the Lloyd-Max quantization and the ABUQ for the $GSND(128, 10)$. For this experiment the Lloyd-Max quantizers were designed with bit-depth ranging from 1 to 8 bits and number of terms ranging from 1 to 16. The ABUQ was set so that its header info has the same length as the Lloyd-Max quantizer, that is, the second coefficient can use from 1 to 128 quantization levels. Note that, in the experiments presented, the ABUQ decompositions are not restricted to use 16 terms as in the Lloyd-Max quantized decompositions are. It can be seen in Figure D.24 that at rates below 12 bits/sample both methods achieve similar distortion; however, the Lloyd-Max quantizer allows more accurate and finer rate control at these rates, as it is demonstrated by the larger number of points at low R for the LMQ than for the ABUQ in the RD plots of the signals. For rates above 12 bits/sample, for some signals, the ABUQ outperforms the Lloyd-Max quantizer. However, for other signals (see the third one) the

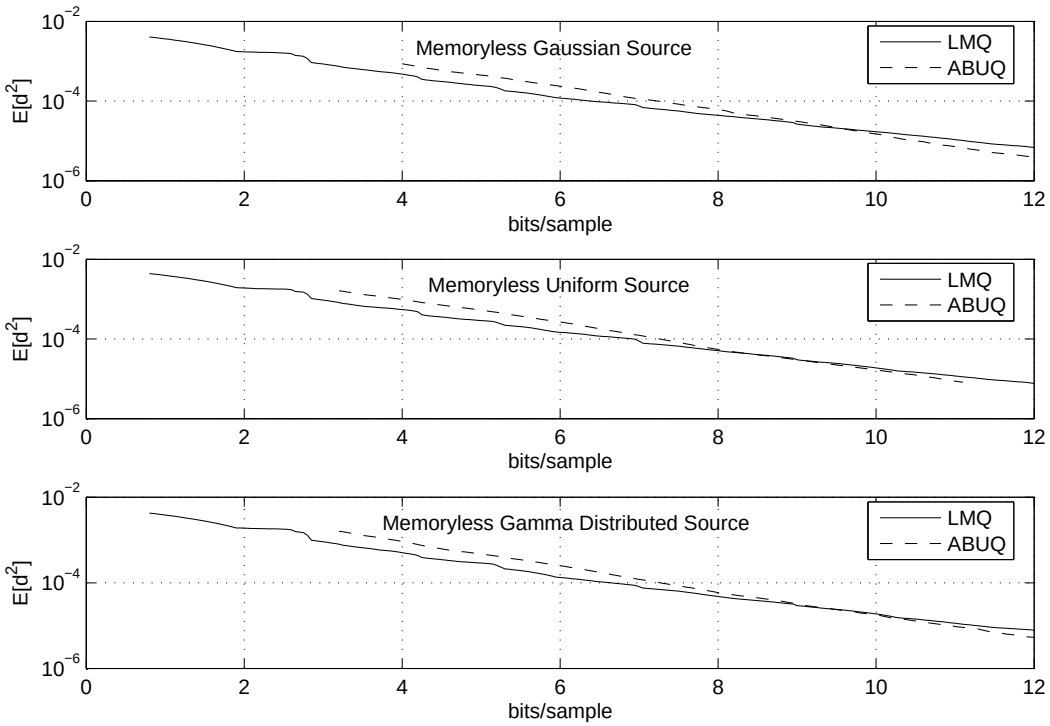


Figure D.23: LMQ and ABUQ RDs for three different random sources using the $GSND(128, 10)$.

ABUQ is not capable of obtaining smaller distortion by increasing the rate.

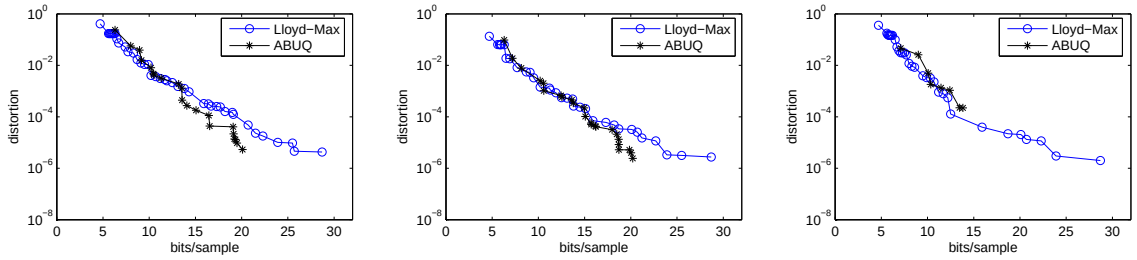


Figure D.24: Lloyd-Max quantizer compared to the ABUQ for the $GSND(128, 10)$ – RD curves for three different signals drawn from a Gaussian source.

Figure D.25 shows the RD curves of quantized MP expansions originated from the three different random sources above in $\mathbb{R}^{64}$ for the LMQ and the ABUQ for the Gabor dictionary of 4 phases in $\mathbb{R}^{64}$, see equation (D.11). For that purpose the LMQs were designed with bit-depth ranging from 1 to 6. The decompositions to be coded allowed a maximum of 256 terms. For each distinct source the results are averages over an ensemble of 200 quantized MP decompositions of signals from each source. The decompositions quantized with the ABUQ require a larger bit-rate because its bit-allocation scheme is not capable of controlling the rate. The reason

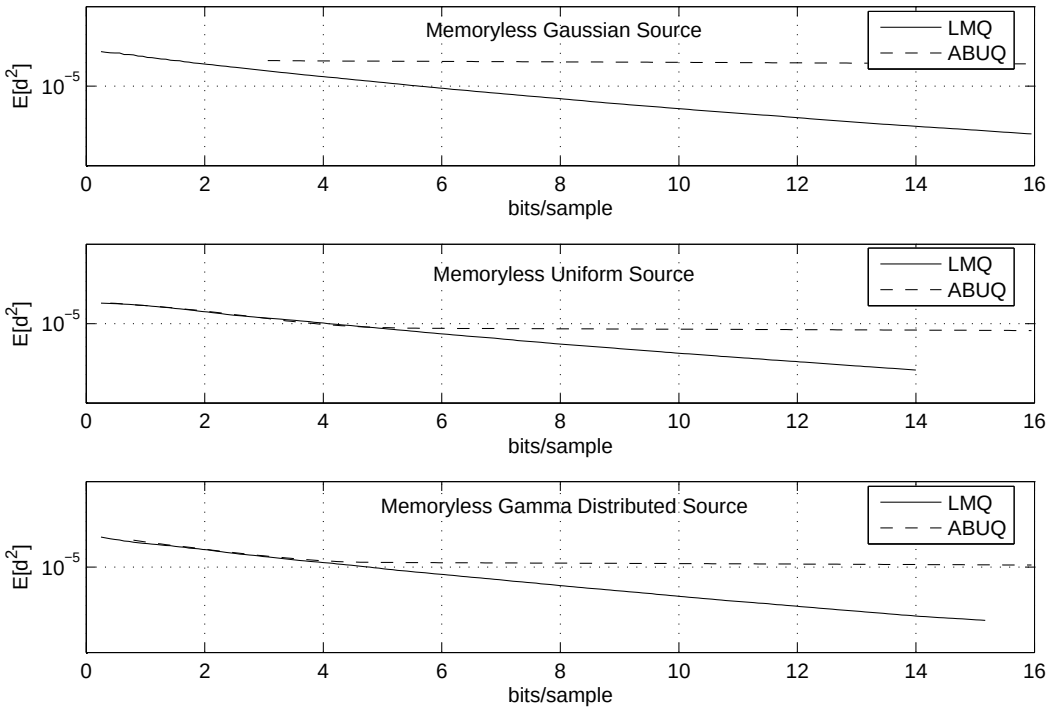


Figure D.25: LMQ and ABUQ RDs for three different random sources using the 4-phase Gabor dictionary in $\mathbb{R}^{64}$.

for that is further discussed next.

Figure D.26 presents the RD plots for three random Gaussian signals for the Gabor dictionary of 4 phases in $\mathbb{R}^{64}$. For that purpose the Lloyd-Max quantizers were designed with bit-depth ranging from 1 to 8 and number of terms ranging from 1 to 64. The ABUQ was set to allow a number of bits ranging from 1 to 256 for the second coefficient. The decompositions were set to allow a total of 1000 terms. Note that for this dictionary the ABUQ is not capable of implementing the rate control. This happens because the shape of $f_{\Theta_1}(\theta_1)$, see Figure D.4, for this dictionary implies a lower decay rate of the coefficients magnitudes when this dictionary is employed than when the $GSND(128, 10)$ is employed in the MP framework. If the bit-stream generated by the ABUQ is allowed to be truncated at any given number of terms, that is, the coding process can be stopped at any coefficient even if the number of quantization levels of the current coefficient is not null, then the RD curve obtained is shown in Figure D.27, where it is possible to observe that both the modified ABUQ and the Lloyd-Max quantizer have similar RD performance at low bit-rate.

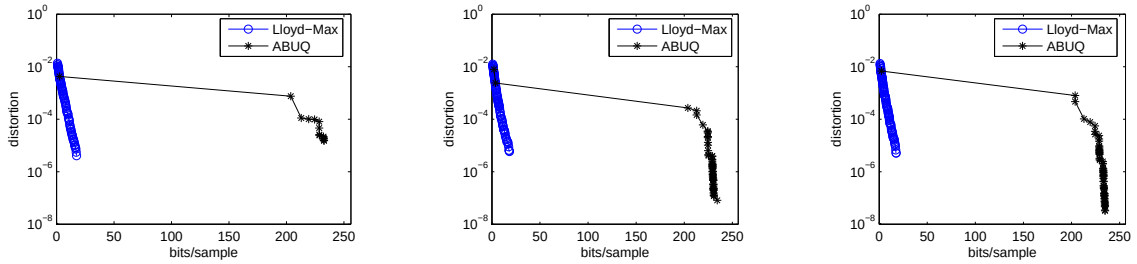


Figure D.26: Lloyd-Max quantizer compared to the ABUQ for the 4 phases Gabor dictionary in $\mathbb{R}^{64}$ – RD curves for three different signals drawn from a Gaussian source.

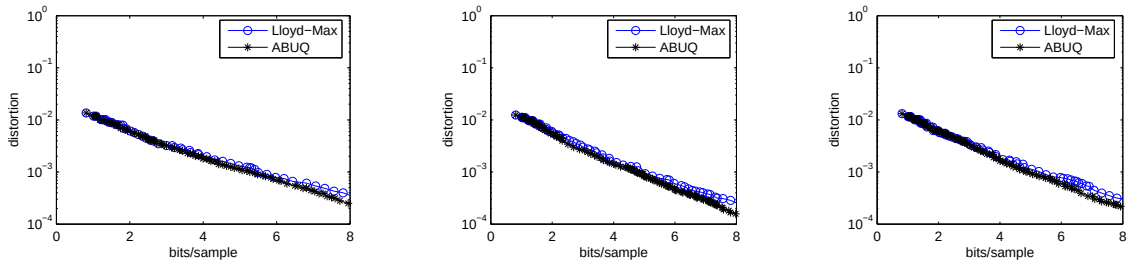


Figure D.27: Lloyd-Max quantizer compared to the ABUQ for the 4 phases Gabor dictionary in $\mathbb{R}^{64}$ – RD curves for three different signals drawn from a Gaussian source.

D.5 Chapter Summary

We have started this appendix by discussing the application of Matching Pursuits decompositions for signal compression. We have observed that an important point to compress signals using Matching expansions, besides the dictionary, is the quantization of the Matching Pursuit expansion coefficients.

The angle between the residue and the selected atom in Matching Pursuit iterations was defined. The empirical analysis of this angle led to a conjecture that these angles may be statistically modeled as independent and identically distributed (iid). Therefore, the angles in Matching Pursuit steps can approximately be considered statistically invariant with respect to the decomposition step. This way, the statistics of Matching Pursuit angles can be obtained from the statistics of the first Matching Pursuit angle for signals drawn from a Gaussian source. This model has shown to be adequate for a large class of dictionaries.

However, if the dictionary includes orthonormal bases the identical independent distributed statistical model is not appropriate after a number of steps greater

than or equal to the signal space dimension. In Appendix D, it was shown that if the dictionary includes at least one orthonormal basis then the Matching Pursuit has a finite non-null probability of convergence when the number of decomposition steps is greater than or equal to the signal space dimension. This result explains why the statistical model does not apply to these step numbers when the dictionary includes at least one orthonormal basis. However, it was verified that for dictionaries including orthonormal bases the independent identically distributed statistical model can be adapted by using two different statistical sets; one set applies whenever the step number is smaller than the signal dimension and the other applies when the number of the step is greater than or equal to the signal dimension.

The histograms of Matching Pursuits angles for the four-phases Gabor dictionary in $\mathbb{R}^{64}$ have shown to be very similar. However, this dictionary presents an intriguing aspect, that deserves to be mentioned: the histograms of Matching Pursuit angles, at step numbers that are larger than one, slightly differ from the histogram of the first angle for a Gaussian source. Therefore, one can see that the residues of Matching Pursuit iterations, for the particular dictionary, are not exactly uniform distributed on the hyper-sphere surface. Note that the residues for this dictionary seem to have approximately invariant statistics with respect to the decomposition step, however these statistics are a bit different from the ones obtained at the first decomposition step a Gaussian source. However, the Gaussian source still provides a very good approximate model for obtaining the aforementioned statistics.

The iid statistical model of MP angles was used to design Lloyd-Max quantizers for Matching Pursuit coefficients. This design is based only on the estimate of the probability density function of the first Matching Pursuit angle when an iid Gaussian signal source is decomposed. Each L-level quantizer designed changes with the number of terms in the signal representation M. However, it was observed that for practical purposes there is no need to design L-level quantizers for all values of M, since for each value of L, after a given number of terms in the representation, there is no gain in designing new quantizers.

The Lloyd-Max quantizers obtained were compared to two simple uniform-dead-zone quantizers (UDZQ), namely the $\|\mathbf{x}\|$-UDZQ and the $\max(|\gamma|)$-UDZQ. The dynamic range of the $\|\mathbf{x}\|$-UDZQ is defined by the signal modulus whereas for

the max $(|\gamma|)$-UDZQ it is defined by the magnitude of the largest coefficient of the signal decomposition. As was observed the max $(|\gamma|)$-UDZQ has better performance than the $\|\mathbf{x}\|$-UDZQ, but, the Lloyd-Max quantization presented performs better than both providing a reconstruction distortion that is below 50% of the one obtained by the max $(|\gamma|)$-UDZQ at high data rates. It was also discussed how to apply the max $(|\gamma|)$-UDZQ to on-line quantization, enabling for m-delayed on-line quantization.

The Lloyd-Max quantization presented was also compared to the state-of-the-art off-loop quantization scheme. Both quantization schemes presented similar rate–distortion performance. However, for some dictionaries the Lloyd-Max quantizer presented here provides more accurate rate control. In addition, the Lloyd-Max quantizers can be easily configured for on-line quantization of Matching Pursuit decompositions, and intrinsically provide error resilience that results from the use of the same quantization rule for all the coefficients of the Matching Pursuit decomposition.

If the source is not Gaussian the assumption that the first MP angle has similar statistics to the MP angles of other steps is not appropriate. One can not assume the first MP angle to be similar to the others MP angles because a source that is different from the Gaussian source does not have a uniform distributed projection on the hyper-sphere surface. However, for coding sources that are not Gaussian, since a priori the residues may have any direction, the Lloyd-Max quantizers obtained for the Gaussian source, as well as the pairs (S, b_{coef}) arriving from the Gaussian source RD optimization presented, can still be employed. In addition, the Lloyd-Max quantizers performance can be improved for any source by designing the appropriate quantization rule for the random variable Γ_1 corresponding to the first MP coefficient for the source being considered.

Apêndice E

Time-Frequency Content of Frames

Frames have been largely studied in the literature. In general, these studies rest on two approaches. In one of them, frames are considered as a branch of functional analysis, and general results are provided for frames in Hilbert spaces. In the other, frames are studied in view of special applications. In this chapter, results concerning both approaches are presented.

This chapter investigates the time-frequency content of frames. The time-frequency content of frames is defined using the Wigner-Ville distribution, and it allows the characterization of frames in the time-frequency domain. A theorem shows a sufficient condition for a set of elements being a frame using the sum of the time-frequency contents of the elements. The result is derived for $L^2(\mathbb{R})$, and is extensible for the $l^2(\mathbb{Z})$ as well as for finite vector spaces.

Special attention is given to the analysis of the time-frequency content of Weyl-Heisenberg frames. Using these results, Weyl-Heisenberg frames of damped sinusoids are characterized by means of their time-frequency content. Then, it is presented a strategy to generate "tighter" frames from a symmetric prototype signal through the "interlacing" of Weyl-Heisenberg frames in the time-frequency domain.

The frames that result from the "interlacing" of Weyl-Heisenberg frames are evaluated when generating dictionaries to be used in greedy decomposition algorithms. The results show that the "interlacing" of Weyl-Heisenberg frames as a dictionary generation method allows for good rate×distortion tradeoff in signal com-

pression applications.

E.1 Time-Frequency Content of Frames

The Wigner-Ville distribution (WD) is probably one of the most well-known and widely used tools for time-frequency analysis of signals. The WD of a signal $x(t)$ is defined as [20, 71, 84]

$$WD_x(t,f) = \int_{-\infty}^{+\infty} x\left(t+\frac{\tau}{2}\right)\overline{x\left(t-\frac{\tau}{2}\right)}e^{-2\pi jf\tau}d\tau, \tag{E.1}$$

where the $\overline{x(t)}$ denotes the complex conjugate of $x(t)$. Some applications have used the WD of signal decompositions in order to analyze the time-frequency content of signals [71, 84]. The Wigner-Ville distribution of a signal $x(t)$, $WD_x(t,f)$, is a "measure" of the energy density in the signal in both time (t) and frequency (f) simultaneously. However, it is just meaningful when regions of the time-frequency plane are considered, that is as local averages, and it can not be considered at a given time-frequency point (t', f') due to uncertainty principle [71, 84]. In addition, the WD of a signal has another drawback since it is not restricted to be positive, a mandatory characteristic for an energy density.

In Section B.4, we have seen that a signal x can be decomposed into a frame $\{g_k\}_{k\in I}$ or into its dual $\{\tilde{g}_k\}_{k\in I}$, and reconstructed by means of

$$x = \sum_{k\in\mathcal{I}}\langle x, \tilde{g}_k\rangle g_k = \sum_{k\in\mathcal{I}}\langle x, g_k\rangle \tilde{g}_k. \tag{E.2}$$

Once x is decomposed into the frame coefficients $\langle x, \tilde{g}_k\rangle$, its time-frequency content can be analyzed using

$$WD_{\tilde{x}}(t,f) = \sum_{k\in\mathcal{I}}\langle x(t), \tilde{g}_k(t)\rangle\, WD_{g_k}(t,f). \tag{E.3}$$

Using approaches like this, the inference of signal characteristics from signal decompositions is a common and powerful tool for signal analysis, detection and estimation. For instance, in [85] the time-frequency content of signals is estimated using the Matching Pursuits (MP) signal decomposition algorithm (see section B.2.2). For that purpose, the so-called Gabor dictionary is employed in the MP. This dictionary is formed by Gaussian functions in different scales with different

time and frequency shifts. From the MP expansion of signal $x(t)$ into the Gabor dictionary

$$x(t) \approx \sum_{n=1}^{M} \gamma_n g_{i(n)}(t), \tag{E.4}$$

the time-frequency content of $x(t)$ is analyzed by means of

$$\sum_{n=1}^{M} \gamma_n WD_{g_{i(n)}}(t, f). \tag{E.5}$$

In [89] the same approach is used, but using a modified version of the MP to obtain the signal expansion. Note the resemblance between the approach in equation (E.3) and the ones in equations (E.4) and (E.5).

In Section B.4, we have defined the frame operator as

$$S\{x\} = \sum_{k \in \mathcal{I}} \langle x, g_k \rangle g_k. \tag{E.6}$$

If instead of the coefficients obtained using the inverse frame, the signal projection into the frame elements g_k is considered, as is done in the frame operator, one can define

$$WD_{x(t) \text{ in } \mathcal{G}}(t, f) = \sum_{k} \langle x(t), g_k(t) \rangle \, WD_{g_k}(t, f). \tag{E.7}$$

Equation (E.7) is sum of the time-frequency contents of the projections of a signal $x(t)$ into the elements of the frame $\mathcal{G} = \{g_k\}_{k \in I}$. This is totally different than $WD_{\tilde{x}}(t, f)$ defined in equation (E.3). In addition, note that equation (E.7) does not provide an estimate of the signal time-frequency content. For that purpose the frame coefficients $\langle x(t), \tilde{g}_k(t) \rangle$ should be used instead of the projections $\langle x(t), g_k(t) \rangle$, as in equation (E.3).

E.1.1 Definition and Relevance

Definition E.1 *The time-frequency content of a frame $\mathcal{G} = \{g_k\}_{k \in I}$ is defined as*

$$WD_{\mathcal{G}}(t, f) = \sum_{k \in \mathcal{I}} WD_{g_k}(t, f). \tag{E.8}$$

The theorem below shows a condition on $WD_{\mathcal{G}}(t, f)$ that guarantees that a set of functions $\{g_k(t)\}_{k \in I}$ is a frame in $L^2(\mathbb{R})$.

Theorem E.1 *A sufficient condition for a set of elements $\mathcal{G} = \{g_k(t)\}_{k\in\mathcal{I}}$ to be a frame for $L^2(\mathbb{R})$ is that*

$$0 < WD_{\mathcal{G}}(t,f) < \infty, \quad \forall\ (t,f), \tag{E.9}$$

where $WD_{\mathcal{G}}(t,f)$ is defined in equation (E.8).

Proof: *The WD satisfies the Moyal relation [84]*

$$\langle WD_x(t,f), WD_g(t,f)\rangle = \int_t\int_f WD_x(t,f)WD_g(t,f)dfdt = |\langle x(t), g(t)\rangle|^2. \tag{E.10}$$

Also, the WD preserves the energy, since [84]

$$\int_t\int_f WD_x(t,f)dfdt = \int_t |x(t)|^2 dt = \|x(t)\|^2. \tag{E.11}$$

Applying the properties above in the frame definition

$$A\|x(t)\|^2 \le \sum_{k\in\mathcal{I}} |\langle x, g_k\rangle|^2 \le B\|x(t)\|^2, \quad \forall x \in \mathbb{H}, \tag{E.12}$$

see section B.4, it can verified that

$$A\int_t\int_f WD_x(t,f)dfdt \le \sum_{k\in\mathcal{I}} \langle WD_x, WD_{g_k}\rangle \le B\int_t\int_f WD_x(t,f)dfdt. \tag{E.13}$$

Therefore,

$$A\int_t\int_f WD_x(t,f)dfdt \le \int_t\int_f WD_x(t,f)WD_{\mathcal{G}}(t,f)dfdt \le B\int_t\int_f WD_x(t,f)dfdt. \tag{E.14}$$

A sufficient condition for equation (E.14) to hold for any signal $x(t)$ is that

$$A \le WD_{\mathcal{G}}(t,f) \le B \quad \forall\ (t,f). \tag{E.15}$$

Hence, if

$$0 < WD_{\mathcal{G}}(t,f) < \infty,\ \forall (t,f), \tag{E.16}$$

then the set $\mathcal{G} = \{g_k(t)\}_{k\in\mathcal{I}}$ is a frame in $L^2(\mathbb{R})$.

Note that if the time-frequency content of a frame $\mathcal{G}$ satisfies equation (E.15) then

$$\int_t\int_f WD_x(t,f)WD_{\mathcal{G}}(t,f)dfdt \ge \min_{t,f} WD_{\mathcal{G}}(t,f)\int_t\int_f WD_x(t,f)dfdt = \min_{t,f} WD_{\mathcal{G}}(t,f)\|x(t)\|^2, \tag{E.17}$$

$$\int_t\int_f WD_x(t,f)WD_{\mathcal{G}}(t,f)dfdt \le \max_{t,f} WD_{\mathcal{G}}(t,f)\int_t\int_f WD_x(t,f)dfdt = \max_{t,f} WD_{\mathcal{G}}(t,f)\|x(t)\|^2. \tag{E.18}$$

Therefore, although $\min_{t,f} WD_{\mathcal{G}}(t,f)$ *and* $\max_{t,f} WD_{\mathcal{G}}(t,f)$ *do not provide the exact frame bounds of frame* $\mathcal{G}$*, they bound the lower* A *and upper* B *frame bounds of* $\mathcal{G}$*, respectively.* ■

Corollary E.1 *The time-frequency content of a frame being constant is a sufficient condition for the frame to be tight* ($A = B$)*.*

Proof: *If*

$$WD_{\mathcal{G}}(t,f) = C \tag{E.19}$$

then from equation (E.14) we have that

$$A \int_t \int_f WD_x(t,f) df dt \leq C \int_t \int_f WD_x(t,f) df dt \leq B \int_t \int_f WD_x(t,f) df dt. \tag{E.20}$$

The above equation implies that

$$A = C = B. \tag{E.21}$$

■

E.1.2 Extension to Discrete Spaces

In discrete spaces, the WD [20] of a signal is defined as

$$WD_{\mathbf{x}}(l,\omega) = 2 \sum_{i=-\infty}^{\infty} e^{-\jmath\omega 2i} x[l+i]\,\overline{x[l-i]}. \tag{E.22}$$

Actually [20] presents several Wigner-Ville distribution definitions that satisfy both the Moyal and the energy preservation properties of the WD in $L^2(\mathbb{R})$. The last equation provides the time-frequency content of a signal in $l^2(\mathbb{Z})$. The aliasing due to the sampling operation can be overcome by oversampling the signal with a factor greater than 2 [20]. Since the WD definition in equation (E.22) satisfies the Moyal and the energy preservation properties, the result in Theorem E.1 is also valid in $l^2(\mathbb{Z})$.

E.1.3 Extension to Finite Dimensional Vector Spaces

In an N-dimensional vector space $\mathbb{H}^N$, the computation of the time-frequency content of signals can be done by considering that $\mathbb{H}^N$ is an N-length section of a N-periodic space $l^2(\mathbb{Z})$; as pointed out in section B.4.4.4 and in this case the frequency

is restricted to multiples of $\frac{2\pi}{N}$, i.e. $\Omega = \frac{2\pi k}{N}$, $k \in [0, \ldots, N-1]$. However, the assumption of an N-periodic space produces infinite summations in equation (E.22). To avoid these infinite summations, the signal samples can be considered to be zero outside the N samples that define the finite dimensional vector space $\mathbb{H}^N$. In doing so, for $\mathbf{x} \in \mathbb{H}^N$, the WD definition becomes

$$WD_{\mathbf{x}}(l, \Omega) = 2 \sum_{i=-N}^{N-1} e^{-\frac{j2\pi 2\Omega i}{2N}} x[l+i] \, \overline{x[l-i]}. \tag{E.23}$$

This WD definition also suffers from aliasing, that can be overcome as shown in [20] or [84]. For instance, this WD definition, in its fast version [84], is employed in this work to compute the numerical results presented on the time-frequency content of frames. The WD definition above, equation (E.23), also satisfies the Moyal and the energy conservation properties and thus Theorem E.1 is also valid to represent the time-frequency content of frames in $\mathbb{R}^N$.

Example E.1 *Let the vertices of the N-dimensional cube form a frame $\mathcal{G}$ in the N-dimensional space. That is, the frame $\mathcal{G}$ is composed by the elements that are given by*

$$\mathbf{g}_k = [g_k[0], \ldots, g_k[N-1]] = [\pm 1, \ldots, \pm 1]. \tag{E.24}$$

Using the definition in equation (E.23) it can be shown that for these frames $WD_{\mathcal{G}}(l, \Omega) = 2^N$. They have constant time-frequency content implying that these frames are tight. To confirm that, one can see that for these frames $\mathbf{T}\mathbf{T}^* = 2^N \mathbf{I}_N$. Therefore, since the frame operator $\mathbf{S}$ (see equation (B.27)) is an identity matrix of size N multiplied by $A = 2^N$ these frames are tight.

E.2 Time-Frequency Content of Weyl-Heisenberg Frames

As seen in Section B.4.4.1 a Weyl-Heisenberg (WH) frame is a frame for $L^2(\mathbb{R})$ obtained from a fixed function $g(t)$ with $a, b > 0$ through

$$\{E_{mb} T_{na} g(t)\}_{m,n \in \mathbb{Z}}; \tag{E.25}$$

where the translation by a a (T_a) and modulation by b (E_b) operations are given by

$$T_a g(t) = g\left(t - a\right); \tag{E.26}$$

$$E_b g(t) = g(t) e^{2\pi \jmath b t}. \tag{E.27}$$

Since it is known that [21, 71, 84]

$$WD_{T_a g}(t, f) = WD_g(t - a, f) \text{ and } WD_{E_b g}(t, f) = WD_g(t, f - b), \tag{E.28}$$

we can say that the elements of a Weyl-Heisenberg frame have the same time-frequency "shape" and "spread" and each one of the elements time-frequency content is located at a point (na, mb) of the time-frequency plane. Therefore, Weyl-Heisenberg (WH) frames are specially suited for time-frequency analysis of signals [84, 85].

Weyl-Heisenberg (WH) frames were also used for transient detection using the Gaborgram [51, 52, 114]. The Gaborgram is the plot of the coefficients

$$c_{m,n} = \langle x(t), E_{mb} T_{na} \tilde{g}(t) \rangle, \tag{E.29}$$

that obtain the signal synthesis using a WH frame $\{E_{mb} T_{na} g(t)\}$, as a function of the time-shift and frequency-shift indices n and m that generate the WH frame. The time-frequency content of Weyl-Heisenberg frames is addressed in what follows.

From equations (E.26) and (E.27) and the WD definition in equation (E.1) one has that

$$WD_{E_{mb} T_{na} g(t)}(t, f) = WD_g(t - na, f - mb) \tag{E.30}$$

Therefore, for a WH frame $\mathcal{G} = \{E_{mb} T_{na} g(t)\}_{m,n \in \mathbb{Z}}$ one has that

$$WD_{\mathcal{G}}\left(t, f\right) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} WD_g\left(t - na, f - mb\right). \tag{E.31}$$

Therefore

$$\begin{aligned} WD_{\mathcal{G}}\left(t, f\right) &= \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} WD_{E_{mb} T_{na} g(t)}(t, f) \\ &= \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \int_{-\infty}^{+\infty} g\left(t + \frac{\tau}{2} - na\right) \overline{g\left(t - \frac{\tau}{2} - na\right)} e^{-2\pi \jmath m b \tau} e^{-2\pi \jmath f \tau} d\tau. \\ &= \sum_{n \in \mathbb{Z}} \int_{-\infty}^{+\infty} g\left(t + \frac{\tau}{2} - na\right) \overline{g\left(t - \frac{\tau}{2} - na\right)} e^{-2\pi \jmath f \tau} \sum_{m \in \mathbb{Z}} e^{-2\pi \jmath m b \tau} d\tau \end{aligned} \tag{E.32}$$

Using the Poisson summation, one obtains that

$$WD_{\mathcal{G}}(t,f) = \frac{1}{b}\sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}}\int_{-\infty}^{+\infty} g\left(t+\frac{\tau}{2}-na\right)\overline{g\left(t-\frac{\tau}{2}-na\right)}e^{-2\pi jf\tau}\delta\left(\tau-\frac{m}{b}\right)d\tau,$$

$$WD_{\mathcal{G}}(t,f) = \frac{1}{b}\sum_{m\in\mathbb{Z}} e^{-2\pi jf\frac{m}{b}}\sum_{n\in\mathbb{Z}} g\left(t-na+\frac{m}{2b}\right)\overline{g\left(t-na-\frac{m}{2b}\right)}. \tag{E.33}$$

Equation (E.33) is simpler to compute than equation (E.31) since there is no need for integrations in the calculation of each WD_{g_k}. It can also be noted from equation (E.31) or (E.33) that $WD_{\mathcal{G}}(t,f)$ is 2D-periodic, with period a over t and b over f.

Since the term

$$\sum_{n\in\mathbb{Z}} g\left(t-na+\frac{m}{2b}\right)\overline{g\left(t-na-\frac{m}{2b}\right)}, \tag{E.34}$$

in equation (E.33), is a-periodic over t, then we have that

$$\sum_{n\in\mathbb{Z}} g\left(t-na+\frac{m}{2b}\right)\overline{g\left(t-na-\frac{m}{2b}\right)} = \sum_{k\in\mathbb{Z}} e_k e^{j\frac{2\pi}{a}kt}, \tag{E.35}$$

$$\tag{E.36}$$

where

$$\begin{aligned}
e_k &= \frac{1}{a}\int_0^a \sum_{n\in\mathbb{Z}} g\left(t-na+\frac{m}{2b}\right)\overline{g\left(t-na-\frac{m}{2b}\right)}e^{-j\frac{2\pi}{a}kt}dt\\
e_k &= \frac{1}{a}\sum_n \int_{na+\frac{m}{2b}}^{(n+1)a-\frac{m}{2b}} g(\tau)\,\overline{g\left(\tau-\frac{m}{b}\right)}e^{-j\frac{2\pi}{a}k\left(\tau+na-\frac{m}{2b}\right)}d\tau\\
e_k &= \frac{1}{a}\int_{-\infty}^{\infty} g(\tau)\,\overline{g\left(\tau-\frac{m}{b}\right)}e^{-j\frac{2\pi}{a}k\left(\tau-\frac{m}{2b}\right)}d\tau\\
e_k &= \frac{1}{a}\int_{-\infty}^{\infty} g\left(\tau+\frac{m}{2b}\right)\overline{g\left(\tau-\frac{m}{2b}\right)}e^{-j2\pi\frac{k}{a}\tau}d\tau.
\end{aligned} \tag{E.37}$$

Note that each e_k is the ambiguity function [71, 84]

$$AF_g(\tau,\nu) = \int_{-\infty}^{\infty} g\left(u+\frac{\tau}{2}\right)\overline{g\left(u-\frac{\tau}{2}\right)}e^{-j2\pi\nu u}du \tag{E.38}$$

at $\tau = m/b$ and $\nu = k/a$. Therefore,

$$WD_{\mathcal{G}}(t,f) = \frac{1}{ab}\sum_m\sum_k e^{-j2\pi\left(f\frac{m}{b}+t\frac{k}{a}\right)}AF_g\left(\frac{m}{b},\frac{k}{a}\right). \tag{E.39}$$

In this way, $WD_{\mathcal{G}}(t,f)$ can be obtained from the summation of samples of the ambiguity function of the prototype signal $AF_g(\tau,\nu)$ multiplied by complex exponentials.

E.2.1 Change of Scale

A well known result regarding WH frames is that if $\{E_{mb}T_{na}g(t)\}_{m,n\in\mathbb{Z}}$ is a WH frame with frame bounds A and B then the WH frame generated using a dilated version of the window $g(t)$

$$\{E_{mb}T_{na}D_cg(t)\}_{m,n\in\mathbb{Z}}, \quad D_cg(t) = \frac{1}{\sqrt{c}}g\left(\frac{t}{c}\right) \tag{E.40}$$

is also a WH frame with the same frame bounds [17]. This result is readily indicated from the time-frequency content of frames defined in equation (E.8). Since

$$WD_{\frac{1}{\sqrt{c}}g\left(\frac{t}{c}\right)} = WD_g\left(\frac{t}{c}, cf\right), \tag{E.41}$$

although the time-frequency content of $\{E_{mb}T_{na}D_cg(t)\}_{m,n\in\mathbb{Z}}$ is a time-expanded frequency-contracted ($c > 1$) version of $\{E_{mb}T_{na}g(t)\}_{m,n\in\mathbb{Z}}$ the maximum and minimum of the time-frequency contents of the original and the scaled frames are the same. Therefore, the time-frequency content of a WH frame and the one of its scaled version provide the same bounds on the upper and lower frame bounds.

E.2.2 Weyl-Heisenberg Frames from a Causal Prototype

As is shown by equation (E.31), a Weyl-Heisenberg frame has a 2D-periodic time-frequency content with period a in the time axis and period b in the frequency axis. The WD definition is such that if

$$g(t) = 0, \text{ for } t \in [t_0, t_1] \tag{E.42}$$

then

$$WD_g(t, f) = 0, \text{ for } t \in [t_0, t_1]. \tag{E.43}$$

If $g(t)$ is right sided, i.e. $g(t) = 0\ \forall t < 0$, then since the WD has bounded time support equation (E.31) becomes

$$WD_{\mathcal{G}}(t, f) = \sum_{m\in\mathbb{Z}}\sum_{n\leq t/a} WD_g(t - na, f - mb). \tag{E.44}$$

As equation (E.33) was derived from equation (E.31) one can derive, for a causal function, that for $t \in [0, a)$

$$WD_{\mathcal{G}}(t, f) = \frac{1}{b}\sum_{m\in\mathbb{Z}} e^{-2\pi\jmath f\frac{m}{b}} \sum_{n\leq 0} g\left(t - na + \frac{m}{2b}\right)\overline{g\left(t - na - \frac{m}{2b}\right)}. \tag{E.45}$$

This result is used in the next section to evaluate the frame bounds of WH frames generated from a decaying exponential.

E.2.2.1 Example: Frames of Damped Sinusoids

How can one generate sets of functions

$$f_k(t) = \sqrt{2\lambda_k} e^{-\lambda_k \left(t - t_{0_k}\right)} e^{+\jmath \xi_k t} u\left(t - t_{0_k}\right), \tag{E.46}$$

while guaranteeing that any signal can be synthesized using

$$x(t) = \sum_k \alpha_k f_k(t) = \sum_k \alpha_k \sqrt{2\lambda_k} e^{-\lambda_k \left(t - t_{0_k}\right)} e^{+\jmath \xi_k t} u\left(t - t_{0_k}\right)? \tag{E.47}$$

In Appendix C, this model, a damped sinusoids signal expansion, was used to represent electric oscillographic signals. Damped sinusoids were also employed for the detection of transient signals using Gaborgrams [51,52]. The large amount of potential applications of such signal model is further motivated by the fact that damped sinusoids are solutions for ordinary differential equations that often appear in physical system models, for examples of similar signal models see [59,64,75,83]. Hence, it is interesting and of practical relevance to investigate how the set of parameters $\sigma = [\lambda, t_0, \xi]$ can be sampled, i.e. how to define the mapping

$$\sigma(k) = [\lambda(k), t_0(k), \xi(k)], \ k : \mathbb{N} \mapsto \mathbb{R}^3, \tag{E.48}$$

while still providing a set of functions, generated from the application of the parameter mapping above into equation (E.46), that is capable of representing any signal. That is, the selection of the possible parameters that can be used for each signal component in equation (E.47) should permit both signal analysis and synthesis, what is accomplished by frames. If the parameters t_{0_k} and ξ_k using a regular lattice, that is using a linear "sampling" mapping

$$\sigma(k) = [\lambda, k\Delta t, k\Delta\xi], \tag{E.49}$$

where $\lambda, \Delta t, k\Delta\xi$ are constants, and generating damped sinusoids accordingly, one may obtain a WH frame

$$\left\{E_{mb} T_{na} \left(e^{-\lambda_k t} u(t)\right)\right\}_{m,n\in\mathbb{Z}}, \ a = \Delta t, \ b = \frac{\Delta\xi}{2\pi}. \tag{E.50}$$

Therefore, the question in equation (E.47) can be restated as: for the function $g(t) = \sqrt{2\lambda} e^{-\lambda t} u(t)$, which values of a and b allow for the construction of a WH frame from equation (E.50)? Next, we analyze these frames using the time-frequency content of frames.

Time-Frequency Content of Decaying Exponentials Gabor Frames The decaying exponential function

$$g(t) = \sqrt{2\lambda} e^{-\lambda t} u(t) \tag{E.51}$$

is zero for $t < 0$. Since the time-frequency content of a Weyl-Heisenberg frame $WD_{\mathcal{G}}(t,f)$ is a-periodic in t one just needs to compute it at $t \in [0,a)$. Therefore, for a WH frame $\mathcal{G}$ generated using a damped sinusoid as the one in equation (E.51) one obtains that its time-frequency content is given by

$$WD_{\mathcal{G}}(t,f) = \frac{2\lambda}{b} \sum_{m\in\mathbb{Z}} \sum_{n\leq 0} e^{-2\pi \jmath f \frac{m}{b}} e^{-\lambda\left(t-na+\frac{m}{2b}\right)} e^{-\lambda\left(t-na-\frac{m}{2b}\right)} u\left(t-na+\frac{m}{2b}\right) u\left(t-na-\frac{m}{2b}\right). \tag{E.52}$$

Separating the summation in $m \in \mathbb{Z}$, in equation (E.52) above, into $m \leq 0$, $m \geq 0$, eliminating the duplicity at $m = 0$, and changing the sign of the summation for $m \geq 0$, one has that equation (E.52) can be expressed as

$$\begin{aligned} WD_{\mathcal{G}}(t,f) =& \frac{2\lambda}{b} \sum_{m\leq 0} \sum_{n\leq 0} e^{-2\pi \jmath f \frac{m}{b}} e^{-\lambda\left(t-na+\frac{m}{2b}\right)} e^{-\lambda\left(t-na-\frac{m}{2b}\right)} u\left(t-na+\frac{m}{2b}\right) u\left(t-na-\frac{m}{2b}\right) + \\ &+ \frac{2\lambda}{b} \sum_{m\leq 0} \sum_{n\leq 0} e^{2\pi \jmath f \frac{m}{b}} e^{-\lambda\left(t-na-\frac{m}{2b}\right)} e^{-\lambda\left(t-na+\frac{m}{2b}\right)} u\left(t-na-\frac{m}{2b}\right) u\left(t-na+\frac{m}{2b}\right) + \\ &- \frac{2\lambda}{b} \sum_{n\leq 0} e^{-\lambda(t-na)} e^{-\lambda(t-na)} u(t-na)\, u(t-na). \end{aligned} \tag{E.53}$$

Therefore, one has that equation (E.52) turns into

$$\begin{aligned} WD_{\mathcal{G}}(t,f) = \frac{2\lambda}{b} \Bigg[& \sum_{m\leq 0} \sum_{n\leq 0} \left(e^{-2\pi \jmath f \frac{m}{b}} + e^{2\pi \jmath f \frac{m}{b}} \right) e^{-2\lambda(t-na)} u\left(t-na+\frac{m}{2b}\right) u\left(t-na-\frac{m}{2b}\right) \\ & - \sum_{n\leq 0} e^{2\lambda(t-na)} \Bigg]. \end{aligned} \tag{E.54}$$

The unit steps in equation (E.54) imposes that n must be such that

$$\begin{aligned} t - na + \frac{m}{2b} \geq 0 \rightarrow n \leq \frac{t}{a} + \frac{m}{2ab} \\ t - na - \frac{m}{2b} \geq 0 \rightarrow n \leq \frac{t}{a} - \frac{m}{2ab}. \end{aligned}$$

Since $t \in [0, a)$, $n \leq 0$ and $m \leq 0$ then

$$WD_{\mathcal{G}}(t,f) = \frac{2\lambda}{b}\left[\sum_{m\leq 0} 2\cos\left(2\pi f\frac{m}{b}\right)\sum_{n\leq \frac{t}{a}+\frac{m}{2ab}} e^{-2\lambda(t-na)} - \frac{e^{-2\lambda t}}{1-e^{-2\lambda a}}\right]$$

$$WD_{\mathcal{G}}(t,f) = \frac{2\lambda}{b}\left[e^{-2\lambda t}\sum_{m\leq 0} 2\cos\left(2\pi f\frac{m}{b}\right)\sum_{n=-\infty}^{\left\lfloor \frac{t}{a}+\frac{m}{2ab}\right\rfloor} e^{2\lambda na} - \frac{e^{-2\lambda t}}{1-e^{-2\lambda a}}\right]$$

$$WD_{\mathcal{G}}(t,f) = \frac{2\lambda e^{-2\lambda t}}{b}\left[\sum_{m\leq 0} 2\cos\left(2\pi f\frac{m}{b}\right)\frac{e^{2\lambda a\left\lfloor \frac{t}{a}+\frac{m}{2ab}\right\rfloor}}{1-e^{-2\lambda a}} - \frac{1}{1-e^{-2\lambda a}}\right]$$

$$WD_{\mathcal{G}}(t,f) = \frac{2\lambda e^{-2\lambda t}}{b\left(1-e^{-2\lambda a}\right)}\left[\sum_{m\leq 0} 2\cos\left(2\pi f\frac{m}{b}\right)e^{2\lambda a\left\lfloor \frac{t}{a}+\frac{m}{2ab}\right\rfloor} - 1\right] \tag{E.55}$$

For $\frac{1}{ab} = q, q \in \mathbb{N}$ one has that

$$\left\lfloor \frac{1}{2}\left(\frac{2t}{a}+\frac{m}{ab}\right)\right\rfloor = \left\lfloor \frac{1}{2}\left(2tbq+mq\right)\right\rfloor. \tag{E.56}$$

Let

$$\left\lfloor \frac{1}{2}\left(2tbq+mq\right)\right\rfloor = Q, \tag{E.57}$$

then

$$(2tbq+mq) = 2Q + r, \text{ with } 0 \leq r < 2. \tag{E.58}$$

Therefore,

$$\begin{aligned}\left\lfloor \frac{1}{2}\left(2tbq+mq\right)\right\rfloor = Q &= \frac{1}{2}\left(2tbq+mq\right) - \frac{r}{2} \\ &= \frac{1}{2}\left(2tbq+mq\right) - \frac{1}{2}\left(2tbq+mq\right)\text{mod}2.\end{aligned} \tag{E.59}$$

Since $t \in [0, a)$ then $2tbq \in [0, 2)$. In addition, mq is an integer smaller than or equal to zero. Therefore, for q even we have that

$$\left\lfloor \frac{1}{2}\left(2tbq+mq\right)\right\rfloor = \frac{1}{2}\left(2tbq+mq\right) - \frac{1}{2}2tbq = \frac{1}{2}mq. \tag{E.60}$$

Using the above result, valid for q even, the time-frequency content of a Weyl-Heisenberg frame of damped sinusoids is given by

$$\begin{aligned}WD_{\mathcal{G}}(t,f) &= \frac{2\lambda e^{-2\lambda t}}{b\left(1-e^{-2\lambda a}\right)}\left[\sum_{m\geq 0} 2\cos\left(2\pi f\frac{m}{b}\right)e^{-\frac{\lambda}{b}m} - 1\right] \\ &= \frac{2\lambda e^{-2\lambda t}}{b\left(1-e^{-2\lambda a}\right)}\left[2\frac{1-e^{-\frac{\lambda}{b}}\cos\left(2\pi\frac{f}{b}\right)}{1-2e^{-\frac{\lambda}{b}}\cos\left(2\pi\frac{f}{b}\right)+e^{-\frac{2\lambda}{b}}} - 1\right].\end{aligned} \tag{E.61}$$

From equation (E.61) one can see that the maximum of $WD_{\mathcal{G}}(t,f)$ occurs at $(t,f)=(0,0)$ and the minimum at $f=b/2$ and $t\to a$. At those points one has that

$$\frac{2\lambda}{b\left(1-e^{-2\lambda a}\right)}\frac{1-e^{-\lambda/b}}{1+e^{-\lambda/b}} \leq WD_{\mathcal{G}}(t,f) < \frac{2\lambda e^{-2\lambda a}}{b\left(1-e^{-2\lambda a}\right)}\frac{1+e^{-\lambda/b}}{1-e^{-\lambda/b}}. \tag{E.62}$$

Friedlander and Zeira showed in [52], using a condition for Gabor frames derived by Zibulski and Zeevi [113], based on the Zak transform [5,17,52,113], that for an integer oversampling factor $q\geq 1$, the decaying exponential generates a frame with frame bounds [52]:

$$B=\frac{2\lambda}{b\left(1-e^{-2\lambda a}\right)}\frac{1+e^{-\frac{\lambda}{b}}}{1-e^{-\frac{\lambda}{b}}}, \tag{E.63}$$

$$A=\frac{2\lambda e^{-2\lambda a}}{b\left(1-e^{-2\lambda a}\right)}\frac{1-e^{-\frac{\lambda}{b}}}{1+e^{-\frac{\lambda}{b}}}. \tag{E.64}$$

These are exactly the same bounds presented in equation (E.62). Numerical experiments using equation (E.55) suggest that equation (E.62) also holds for an odd value of q, this result is still to be derived.

Decaying Exponential Gabor Frames in Discrete Spaces Note that both the decreasing and the increasing exponential satisfy the condition in equation (B.40). In this case the oversampling factor, see section B.4.4.3, that is given by

$$\frac{1}{ab}=\frac{1}{MN},\ M,N\in\mathbb{N}, \tag{E.65}$$

will define the frame bounds.

E.2.3 Weyl-Heisenberg Frames from Even Prototypes

From the $WD_{\mathcal{G}}$ of a WH frame, equation (E.33), it is possible to obtain

$$\frac{\partial WD_{\mathcal{G}}(t,f)}{\partial t}=\frac{1}{b}\sum_{m\in\mathbb{Z}}e^{-2\pi\jmath f\frac{m}{b}}\sum_{n\in\mathbb{Z}}\quad\frac{\partial g\left(t-na+\frac{m}{2b}\right)}{\partial t}\overline{g\left(t-na-\frac{m}{2b}\right)}+$$
$$+\,g\left(t-na+\frac{m}{2b}\right)\frac{\partial\overline{g\left(t-na-\frac{m}{2b}\right)}}{\partial t} \tag{E.66}$$

$$\frac{\partial WD_{\mathcal{G}}(t,f)}{\partial f}=\frac{1}{b}\sum_{m\in\mathbb{Z}}-2\pi\jmath\frac{m}{b}e^{-2\pi\jmath f\frac{m}{b}}\sum_{n\in\mathbb{Z}}g\left(t-na+\frac{m}{2b}\right)\overline{g\left(t-na-\frac{m}{2b}\right)} \tag{E.67}$$

In order to obtain $\max_{(t,f)}WD_{\mathcal{G}}(t,f)$ and $\min_{(t,f)}WD_{\mathcal{G}}(t,f)$ it suffices to analyze $WD_{\mathcal{G}}(t,f)$ in the region $(t,f)\in[0,a)\times[0,b)$. As previously discussed,

for a WH frame $\mathcal{G}$ one has that $WD_{\mathcal{G}}$ is 2D-periodic with period $a \times b$. If $g(t)$ is symmetric then it can be shown from equations (E.66) and (E.67) that

$$\frac{\partial WD_{\mathcal{G}}(0,0)}{\partial t} = 0, \quad \frac{\partial WD_{\mathcal{G}}(0,0)}{\partial f} = 0, \quad \frac{\partial WD_{\mathcal{G}}\left(\frac{a}{2},\frac{b}{2}\right)}{\partial t} = 0 \quad \text{and} \quad \frac{\partial WD_{\mathcal{G}}\left(\frac{a}{2},\frac{b}{2}\right)}{\partial f} = 0. \tag{E.68}$$

Note that

$$\begin{aligned} WD_{\mathcal{G}}\left(\frac{a}{2},\frac{b}{2}\right) &= \frac{1}{b}\sum_{m\in\mathbb{Z}} e^{-2\pi\imath\frac{b}{2}\frac{m}{b}} \sum_{n\in\mathbb{Z}} g\left(\frac{a}{2}-na+\frac{m}{2b}\right)\overline{g\left(\frac{a}{2}-na-\frac{m}{2b}\right)} \\ &= \frac{1}{b}\sum_{m\in\mathbb{Z}}(-1)^m \sum_{n\in\mathbb{Z}} g\left(a\left(\frac{1}{2}+n\right)+\frac{m}{2b}\right) g\left(a\left(\frac{1}{2}+n\right)-\frac{m}{2b}\right), \\ WD_{\mathcal{G}}(0,0) &= \frac{1}{b}\sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} g\left(na+\frac{m}{2b}\right) g\left(na-\frac{m}{2b}\right), \end{aligned} \tag{E.69}$$

and therefore

$$\begin{aligned} WD_{\mathcal{G}}(0,0) - WD_{\mathcal{G}}\left(\frac{a}{2},\frac{b}{2}\right) =& \frac{1}{b}\sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}}\Big[g\left(na+\frac{m}{2b}\right) g\left(na-\frac{m}{2b}\right) \\ & -(-1)^m g\left(\left(n+\frac{1}{2}\right)a+\frac{m}{2b}\right) g\left(\left(n+\frac{1}{2}\right)a-\frac{m}{2b}\right)\Big]. \end{aligned} \tag{E.70}$$

If $g(t)$ is decreases with $|t|$, as is the case for the Gaussian and two-sided exponential atoms, it is easy to see that $WD_{\mathcal{G}}(0,0) - WD_{\mathcal{G}}\left(\frac{a}{2},\frac{b}{2}\right) \geq 0$. This suggests that:

- the maximum of $WD_{\mathcal{G}}$ occurs at $(t = na, f = mb)$, $\forall\ m,n \in \mathbb{Z}^2$;
- and the minimum of $WD_{\mathcal{G}}$ occurs at $(t = na+a/2, f = mb+b/2)$, $\forall\, m,n \in \mathbb{Z}^2$.

E.3 Interlacing Weyl-Heisenberg Frames

If a frame $\mathcal{G}$ is used to analyze a signal that has time-frequency content localized in regions around the points of minimum of $WD_{\mathcal{G}}(t,f)$, since

$$|\langle g(t), f(t)\rangle|^2 = \int_t\int_f WD_x(t,f)WD_g(t,f)dfdt, \tag{E.71}$$

then the inner products among the frame elements and the signal are small. This happens because the frame has small time-frequency content around its points of minimum.

To improve the analysis of the time-frequency content of signals using a WH frame one could increase the time-frequency density, using smaller time-shift and

time-frequency parameters a and b. One could also increase the frame "tightness", what is equivalent to populate the frame (put elements in the frame) with new elements placed near/at the points of minimum of the time-frequency content of the frame. One way to do this population is to "interlace" two WH frames generated from the same prototype. The resulting "interlaced" frame $\mathcal{F}$ is the union of two WH frames

$$\mathcal{F} = \mathcal{G} \cup \mathcal{H}, \tag{E.72}$$

$$\mathcal{G} = \{E_{mb}T_{na}g(t)\}_{m,n\in\mathbb{Z}} \text{ and } \mathcal{H} = \{E_{mb+\frac{b}{2}}T_{na+\frac{a}{2}}g(t)\}_{m,n\in\mathbb{Z}}. \tag{E.73}$$

Frames $\mathcal{G}$ and $\mathcal{H}$ are then defined, from the same prototype, in a way that the former is placed at the points of the time-frequency plane where the latter has poor localization capabilities and vice-versa. The "interlaced" frame provides more "similar" coefficients for any signal for both the analysis and the synthesis operators of a frame.

The idea of frames "interlacing", see Figure E.1, is applicable for constructing frames based on any symmetric prototypes. In [99] using the affine function it is shown that this approach leads to Grassmannian frames, which are unit norm frames whose elements are chosen such that their inner products are minimized.

For a symmetric prototype function $g(t)$ the minima of $WD_{\mathcal{G}}(t,f)$ occur at $WD_{\mathcal{G}}\left(na+\left(\frac{a}{2}\right), mb+\left(\frac{b}{2}\right)\right)$ and the maxima at $WD_{\mathcal{G}}(na, mb)$. By employing the union of frames presented in Figure E.1 one obtains

$$WD_{\mathcal{F}}(na, mb) = WD_{\mathcal{F}}\left(na+\frac{a}{2}, mb+\frac{b}{2}\right). \tag{E.74}$$

The obvious byproduct of frame interlacing is that if two Gabor systems $\mathcal{G} = \{E_{mb}T_{na}g(t)\}_{n,m\in\mathbb{Z}}$ and $\mathcal{H} = \{E_{mb+\frac{b}{2}}T_{na+\frac{a}{2}}g(t)\}_{n,m\in\mathbb{Z}}$ are generated such that their union gives a frame $\mathcal{F} = \mathcal{G} \cup \mathcal{H}$ for which $WD_{\mathcal{F}}(t,f)$ has smaller oscillations than $WD_{\mathcal{G}}(t,f)$ and $WD_{\mathcal{H}}(t,f)$ have themselves, then a "tighter" frame $\mathcal{F}$ is obtained. Experiments indicate that similar results can be derived for $g(t)$ anti-symmetric, although this is not explored in this work.

Example E.2 *Figure E.2 presents on the left the time-frequency content $WD_{\mathcal{G}}(t,f)$ of a WH frame $\{E_{mb}T_{na}g(t)\}_{n,m\in\mathbb{Z}}$, in the region $[0,2a)\times[0,2b)$, generated using a Gaussian prototype function $g(t) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-t^2/2\sigma^2}$, while Figure E.2.(b) presents*

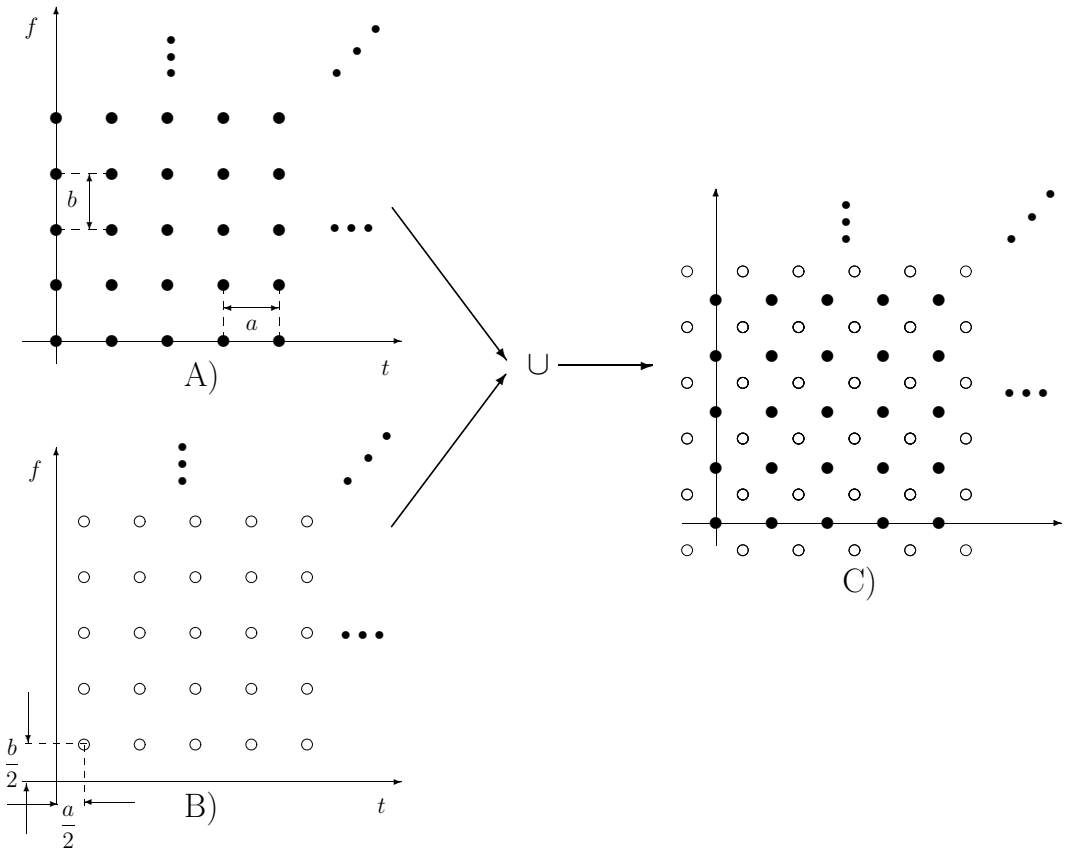


Figure E.1: Population of WH frames by "interlacing".

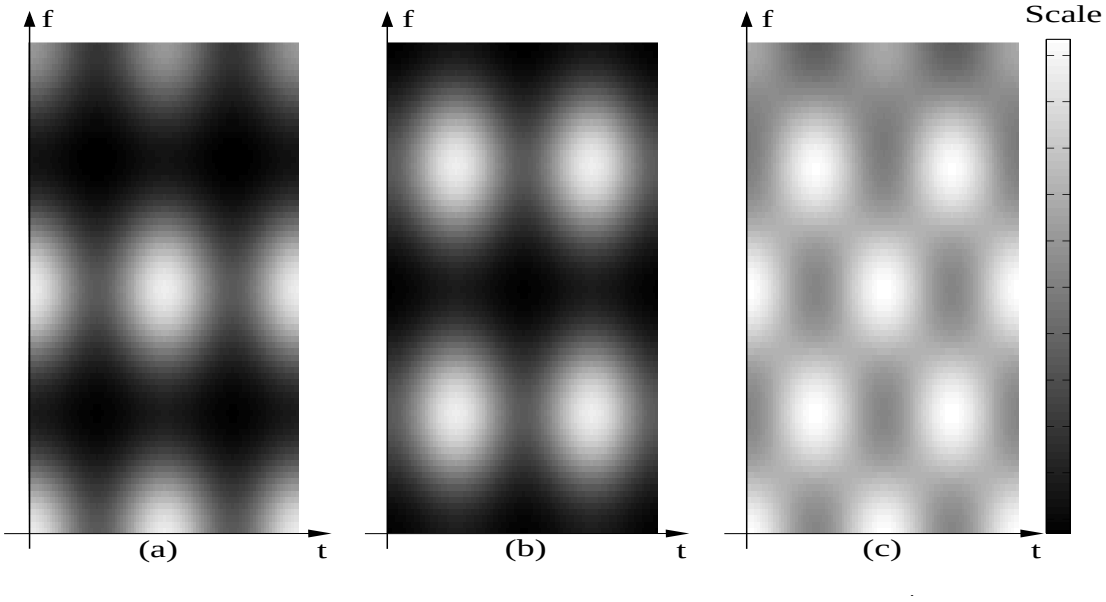


Figure E.2: Example of the "interlacing" of Weyl-Heisenberg frames in the interval $(t, f) \in [0, 2a) \times [0, 2b)$ – the horizontal axis represents time while the vertical one represents frequency.

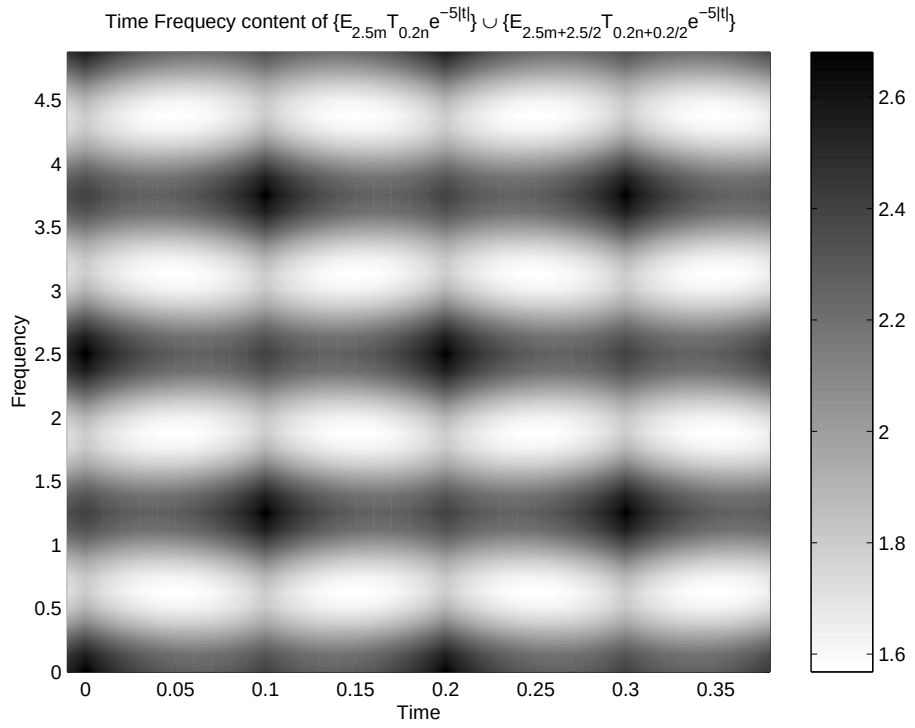


Figure E.3: $WD_{\mathcal{G}}(t,f)$ for $g(t) = e^{-5|t|}$, a =.2 and b =2.5, ab=1/2, $(t,f) \in [0,2a) \times [0,2b)$.

the time-frequency content of the frame $\{E_{mb+b/2}T_{na+a/2}g(t)\}_{n,m\in\mathbb{Z}}$. *Figure E.2.(c) shows the time-frequency content of the frame resulting from the union of the two frames. All three graphs shown in Figure E.2 were plotted using the same range. Note that for the frame* $\mathcal{G}$ *formed by the union of two interlaced frames the maximum and minimum of* $WD_{\mathcal{G}}(t,f)$ *become closer and a tighter frame is obtained.*

Example E.3 *Figure E.3 shows* $WD_{\mathcal{G}}(t,f)$ *in the region* $[0,2a)\times[0,2b)$ *for a frame* $\mathcal{G}$ *composed by the union of two interlaced WH frames of two-sided damped sinusoids* $g(t) = e^{-\beta|t|}$, *with* β =5, *time-shift parameter* a =.2 *and frequency-shift parameter* b =2.5 *(the data presented were obtained at a sampling rate of 100 samples per time unit). For each original frame the minima and maxima of* $WD_{\mathcal{G}}$ *are 0.2274 and 2.4564, however the frame formed by the union of the two frames the time-frequency content is bounded by 1.5659 and 2.6838. Thus, as predicted, the frame generated by the union of "interlaced" frames is tighter.*

E.3.1 Interlaced Weyl-Heisenberg Frames in N-Dimensional Spaces

Table E.1: Design of Weyl-Heisenberg frames in N-dimensional spaces.

Frame	Generation	n range	m range	number of elements
$\mathcal{A}$	$E_{m\frac{1}{Q}}T_{n\frac{N}{P}}\mathbf{g}$	$[0,1,\ldots,P-1]$	$[0,1,\ldots,Q-1]$	PQ
$\mathcal{B}$	$E_{m\frac{1}{Q}+\frac{1}{2Q}}T_{n\frac{N}{P}+\frac{N}{2P}}\mathbf{g}$	$[0,1,\ldots,P-1]$	$[0,1,\ldots,Q-1]$	PQ
$\mathcal{C}$	$A\cup B$			$2PQ$
$\mathcal{D}$	$E_{m\frac{1}{Q}}T_{n\frac{N}{2P}}\mathbf{g}$	$[0,1,\ldots,2P-1]$	$[0,1,\ldots,Q-1]$	$2PQ$
$\mathcal{E}$	$E_{m\frac{1}{2Q}}T_{n\frac{N}{P}}\mathbf{g}$	$[0,1,\ldots,2P-1]$	$[0,1,\ldots,Q-1]$	$2PQ$
$\mathcal{F}$	$E_{m\frac{1}{2Q}}T_{n\frac{N}{2P}}\mathbf{g}$	$[0,1,\ldots,2P-1]$	$[0,1,\ldots,2Q-1]$	$4PQ$
$\mathcal{G}$	$E_{m\frac{1}{Q}+\frac{1}{2Q}}T_{n\frac{N}{2P}+\frac{N}{2P}}\mathbf{g}$	$[0,1,\ldots,2P-1]$	$[0,1,\ldots,Q-1]$	$2PQ$
$\mathcal{H}$	$E_{m\frac{1}{2Q}+\frac{1}{2Q}}T_{n\frac{N}{P}+\frac{N}{2P}}\mathbf{g}$	$[0,1,\ldots,P-1]$	$[0,1,\ldots,2Q-1]$	$2PQ$

In order to achieve a better evaluation of the "interlacing" method to generate "tighter" frames, 8 different WH frames in N-dimensional spaces are now addressed. These are referred to by the calligraphic letters $\mathcal{A}$–$\mathcal{H}$. The designs of the 8 frames are presented in Table E.1. Frames $\mathcal{A}$ and $\mathcal{B}$ are given by $E_{mb}T_{na}\mathbf{g}$ and $E_{mb+b/2}T_{na+a/2}\mathbf{g}$ respectively, that is the structures of $\mathcal{B}$ are centered in the points of minimum of $WD_{\mathcal{A}}$ and vice-versa. The values of a and b were chosen to give rise to P points in the time axis and Q points in the frequency axis for both $\mathcal{A}$ and $\mathcal{B}$, this leads to $a = N/P$ and $b = 2\pi/Q$, allowing to construct $\mathcal{A}$ and $\mathcal{B}$ with the same number of elements (PQ). Frame $\mathcal{C}$ is the union of $\mathcal{A}$ and $\mathcal{B}$, that is $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$. Frames $\mathcal{D}$ and $\mathcal{E}$ are generated from $\mathcal{A}$ by doubling its cardinality in order to be the same of $\mathcal{C}$ ($2PQ$), to construct $\mathcal{D}$ the time-shift parameter a is halved, while to construct $\mathcal{E}$ the frequency-shift parameter b is halved. Frame $\mathcal{F}$ is generated simultaneously halving the time-shift parameter a and the frequency-shift parameter b, hence it has a time-frequency density and number of elements four times larger than the ones of $\mathcal{A}$ and $\mathcal{B}$, and two times larger than the ones of $\mathcal{C}$, $\mathcal{D}$ and $\mathcal{E}$. Frames $\mathcal{G}$ and $\mathcal{H}$ are generated from $\mathcal{B}$ doubling its cardinality as is done for $\mathcal{D}$ and $\mathcal{E}$ with respect to $\mathcal{A}$. To generate $\mathcal{G}$ the time-shift parameter a is halved and to generate $\mathcal{H}$ the frequency-shift parameter b is halved. Figure E.4 shows were the elements of frames

$\mathcal{A}$, $\mathcal{B}$, $\mathcal{D}$, $\mathcal{E}$, $\mathcal{G}$ and $\mathcal{H}$ are placed in the time-frequency plane. It is easy to see that all the elements of dictionaries $\mathcal{A}$–$\mathcal{E}$, $\mathcal{G}$ and $\mathcal{H}$ are contained in $\mathcal{F}$. Note from Figure E.4 that $\mathcal{D}$ is equivalent to $\mathcal{G}$ as $\mathcal{E}$ is to $\mathcal{H}$. All the frames are generated considering a periodic space (circular time shifts) and had their elements normalized. Since the elements are defined in $\mathbb{C}^N$ the normalization procedure does not change the frame bounds ratio A/B although it affects the frame bounds. The normalization of the elements will alter both frame bounds by the same amount: 1) let the elements norm be g, all the elements of the frame have the same norm as the space is periodic; 2) let the frame bounds to be A and B; 3) after normalizing the frame elements, the frame bounds become A/g^2 and B/g^2 and the frame bounds ratio remains the same.

Table E.2: Characteristics of Weyl-Heisenberg and interlaced Weyl-Heisenberg frames with N=32, P=8, and Q=8.

	Frame	$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$	$\mathcal{D}$	$\mathcal{E}$	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{H}$
σ^2	number of elements	64	64	128	128	128	256	128	128
1/2	A	0.0042	0.0042	1.7032	1.7032	0.0084	3.4064	1.7032	0.0084
	B	6.2926	6.2926	6.2968	6.2968	12.5851	12.5936	6.2968	12.5851
	A/B	0.0007	0.0007	**0.2705**	**0.2705**	0.0007	**0.2705**	**0.2705**	0.0007
1	A	0.1653	0.1653	3.3216	3.3216	0.3306	6.6433	3.3216	0.3306
	B	4.5131	4.5131	4.6784	4.6784	9.0261	9.3567	4.6784	9.0261
	A/B	0.0366	0.0366	**0.7100**	**0.7100**	0.0366	**0.7100**	**0.7100**	0.0366
2	A	0.8633	0.8633	3.9425	3.9398	1.7277	7.8849	3.9398	1.7277
	B	3.1958	3.1958	4.0591	4.0603	6.3874	8.1151	4.0603	6.3874
	A/B	0.2701	0.2701	**0.9713**	**0.9703**	0.2705	**0.9716**	**0.9703**	0.2705
4	A	1.6001	1.6001	3.9756	3.8531	3.3220	7.9992	3.8531	3.3220
	B	2.4251	2.4251	4.0253	4.1470	4.6789	8.0008	4.1470	4.6789
	A/B	0.6598	0.6598	**0.9877**	**0.9291**	0.7100	**0.9998**	**0.9291**	0.7100
8	A	1.4390	1.4390	3.9871	2.9200	3.9398	7.9946	2.9200	3.9398
	B	2.5793	2.5793	4.0183	5.0854	4.0603	8.0054	5.0854	4.0603
	A/B	0.5579	0.5579	**0.9922**	0.5742	**0.9703**	**0.9987**	0.5742	**0.9703**

In Table E.1 along with the design of the frames are also shown the ranges of

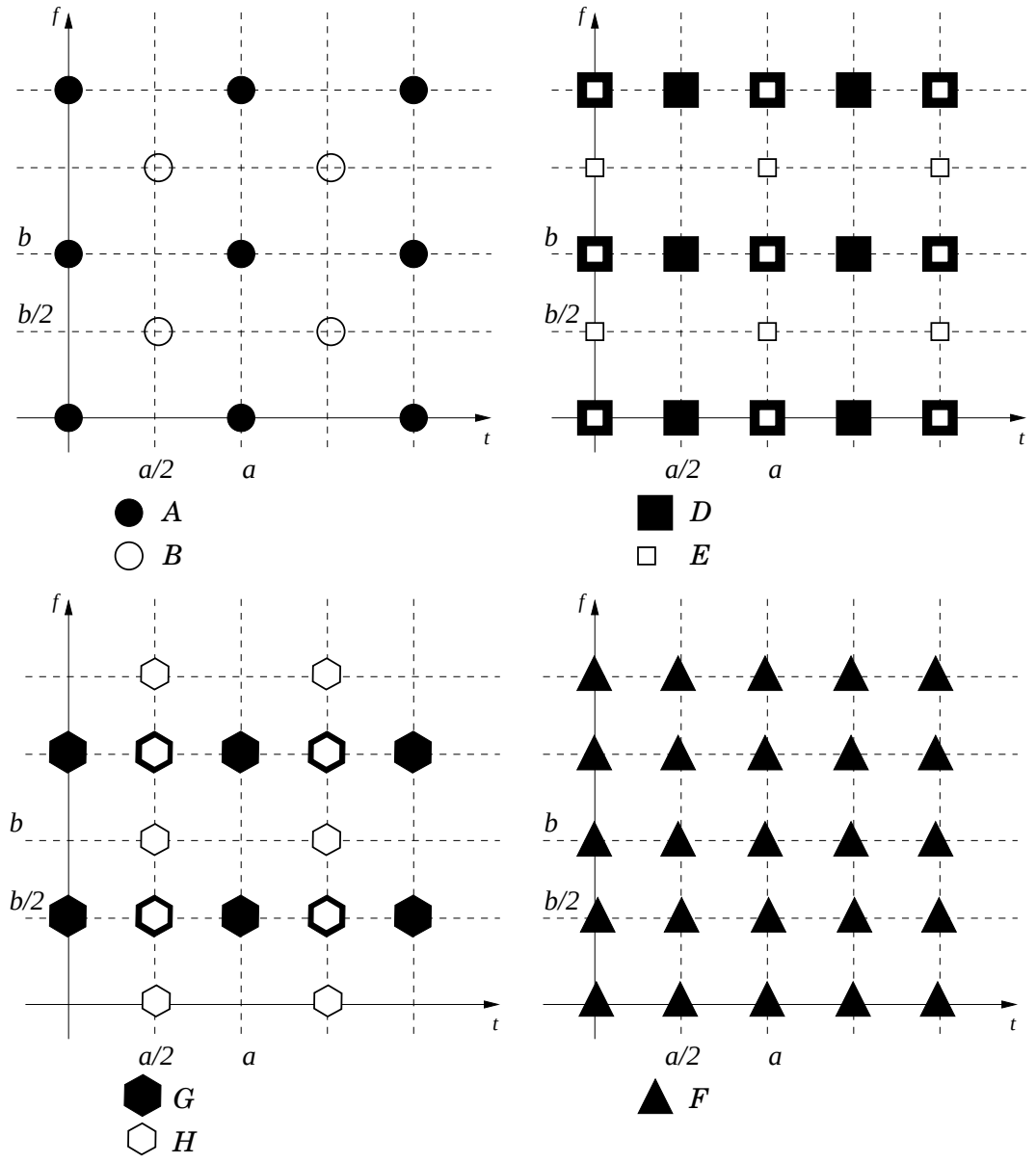


Figure E.4: Time-frequency localization of the elements of frames $\mathcal{A}$, $\mathcal{B}$, $\mathcal{D}$, $\mathcal{E}$, $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{H}$.

the indices n and m that generate each of them and the resulting number of elements in the frames. The construction structures in Table E.1 were employed to generate frames for $\mathbb{C}^{32}$ from a prototype

$$\mathbf{g} = [g[0], g[1], \ldots g[N-1]], \text{ with } g[k] = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-k^2/2\sigma^2} \tag{E.75}$$

for different values of σ^2. Table E.2 shows the frame bounds A and B and their ratio A/B for the 8 frames in Table E.1 for various σ^2 with P and Q set equal to 8. Table E.3 shows the same measures for various σ^2 with P and Q set equal to 16. When $P = Q = 16$ frame $\mathcal{F}$ has N (32) points in both time and frequency axes, that is $a = 1/N$ and $b = 2\pi/N$, it is a maximally oversampled WH frame and it is shown to be a tight [113]. The values in Tables E.2 and E.3 were obtained numerically. When they are obtained by computing the eigenvalues of the frame operator or by the time-frequency content of the frame they are very close (differing in the 12^{th} significant digit). From these figures one can see that the construction $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$ provides frames with frame bounds ratio very close to the one of $\mathcal{F}$ for all σ^2, with half the number of frame elements.

Figure E.4 answers why for low σ^2 frames $\mathcal{D}$ and $\mathcal{G}$ have frame bounds that are similar to the ones of $\mathcal{C}$ (as is seen in Tables E.2 and E.3) and for larger σ^2 $\mathcal{E}$ and $\mathcal{H}$ are other? The smaller σ^2 leads to larger time-support of the elements resulting in shorter frequency-support. As the densities of $\mathcal{D}$ and $\mathcal{G}$ are smaller in the frequency-axis than in the time-axis they generate tighter frames for small σ^2. A similar reasoning explains why the frames $\mathcal{E}$ and $\mathcal{H}$ give rise to tighter frames when σ^2 is large.

We have seen that the interlacing of WH frames is good for generating frames that are tighter than the non-interlaced ones. In some cases, the interlaced frames are tight, as can be seen for frame $\mathcal{C}$ in Table E.3.

Some effort has been spent toward finding "structured" tight frames. Actually, these frames have a broad definition. For example, in [48] these are defined to have some desired symmetry properties and in [106] to have previously defined norms. Here structured frames are considered to have the same prototype "shape". In [113] the generation of tight structured WH frames from one prototype is shown to be possible by employing the largest possible parameter oversampling, that is, by making the time-shift parameter of the WH frame equal to the dimension of the

signal space ($P = N$); the same should be done with the frequency-shift parameter, i.e., making $Q = N$. Although we have no formal proof, the results obtained numerically show that this can be achieved with parameter values halved with respect to the maximally oversampled ones, that is, with $P = Q = N/2$.

Table E.3: Characteristics of Weyl-Heisenberg and interlaced Weyl-Heisenberg frame for N=32, P=16, and Q=16.

	Frame	$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$	$\mathcal{D}$	$\mathcal{E}$	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{H}$
σ^2	number of elements	256	256	512	512	512	1024	512	512
1/8	A	0.0107	0.0107	16.0000	16.0000	0.0215	32.0000	16.0000	0.0215
	B	15.9893	15.9893	16.0000	16.0000	31.9785	32.0000	16.0000	31.9785
	A/B	0.0007	0.0007	**1.0000**	**1.0000**	0.0007	**1.0000**	**1.0000**	0.0007
1/2	A	3.4064	3.4064	16.0000	16.0000	6.8129	32.0000	16.0000	6.8129
	B	12.5936	12.5936	16.0000	16.0000	25.1871	32.0000	16.0000	25.1871
	A/B	0.2705	0.2705	**1.0000**	**1.0000**	0.2705	**1.0000**	**1.0000**	0.2705
1	A	6.6433	6.6433	16.0000	16.0000	13.2865	32.0000	16.0000	13.2865
	B	9.3567	9.3567	16.0000	16.0000	18.7135	32.0000	16.0000	18.7135
	A/B	0.7100	0.7100	**1.0000**	**1.0000**	0.7100	**1.0000**	**1.0000**	0.7100
2	A	7.8849	7.8849	16.0000	16.0000	15.7699	32.0000	16.0000	15.7699
	B	8.1151	8.1151	16.0000	16.0000	16.2301	32.0000	16.0000	16.2301
	A/B	0.9716	0.9716	**1.0000**	**1.0000**	0.9716	**1.0000**	**1.0000**	0.9716
4	A	7.9992	7.9992	16.0000	16.0000	15.9983	32.0000	16.0000	15.9983
	B	8.0008	8.0008	16.0000	16.0000	16.0017	32.0000	16.0000	16.0017
	A/B	0.9998	0.9998	**1.0000**	**1.0000**	0.9998	**1.0000**	**1.0000**	0.9998
8	A	7.9946	7.9946	16.0000	15.9893	16.0000	32.0000	15.9893	16.0000
	B	8.0054	8.0054	16.0000	16.0107	16.0000	32.0000	16.0107	16.0000
	A/B	0.9987	0.9987	**1.0000**	0.9987	**1.0000**	**1.0000**	0.9987	**1.0000**

E.3.2 Fast Analysis and Synthesis Operators

WH frames in N-dimensional spaces are interesting because one can find fast algorithms to compute their analysis and synthesis operators when the number of

"points" of the Weyl-Heisenberg frame in the frequency axis is such that $Q = N/r$, $r \in \mathbb{N}$. Note that

$$c_{n,m} = \langle \mathbf{x}, E_{m\frac{1}{Q}} T_{n\frac{N}{P}} \mathbf{g} \rangle = \sum_{l=0}^{N-1} x[l] e^{\frac{-\jmath 2\pi lm}{Q}} T_{n\frac{N}{P}} g[l]. \tag{E.76}$$

Defining

$$f_{n\frac{N}{P}}[k] = x[k] T_{n\frac{N}{P}} g[k] = x[k] g\left[\left(k - n\frac{N}{P}\right) \mathrm{mod} N\right] \tag{E.77}$$

$$\mathbf{f}_{n\frac{N}{P}} = \left[f_{n\frac{N}{P}}[0], \ldots, f_{n\frac{N}{P}}[N-1]\right] \tag{E.78}$$

we obtain

$$c_{n,m} = \langle \mathbf{x}, E_{m\frac{1}{Q}} T_{n\frac{N}{P}} \mathbf{g} \rangle = \sum_{l=0}^{N-1} f_{n\frac{N}{P}}[l] e^{\frac{-\jmath 2\pi lm}{Q}}. \tag{E.79}$$

For $Q = N/r, r \in \mathbb{N}$ we obtain

$$c_{n,m} = \langle \mathbf{x}, E_{m\frac{1}{Q}} T_{n\frac{N}{P}} \mathbf{g} \rangle = \sum_{l=0}^{N-1} f_{n\frac{N}{P}}[l] e^{\frac{-\jmath 2\pi l}{N} mr} = DFT\{\mathbf{f}_{n\frac{N}{P}}\}[mr], \tag{E.80}$$

where $DFT\{\mathbf{x}\}[k]$ is the k^{th} sample of the Discrete Fourier Transform of $\mathbf{x}$, which can be computed using fast algorithms (FFT) [34].

Using the result in equation (E.80) the analysis operator (which was discussed in section B.4.3) of such WH frames is given by

$$\mathbf{T}^*\{\cdot\} : \mathbb{H}^N \rightarrow \mathbb{C}^{PQ}, \tag{E.81}$$

$$\mathbf{T}^*\{\mathbf{x}\} : c_{n,m} = \{\langle \mathbf{x}, E_{m\frac{1}{Q}} T_{n\frac{N}{P}} \mathbf{g} \rangle\},\ n \in [0, P-1],\ m \in [0, Q-1] \tag{E.82}$$

$$c_{n,m} = FFT\{\mathbf{f}_{n\frac{N}{P}}\}[mr],\ f_{n\frac{N}{P}}[k] = x[k] g\left[\left(k - n\frac{N}{P}\right) \mathrm{mod} N\right] \tag{E.83}$$

Similarly, for the synthesis operator of such frames we have that

$$\mathbf{T}\{\cdot\} : \mathbb{C}^{PQ} \rightarrow \mathbb{H}^N,$$

$$\mathbf{T}\{c_{n,m}\} = [t[0], \ldots, t[N-1]] = \sum_{n=0}^{P-1} \sum_{m=0}^{Q-1} c_{n,m} E_{m\frac{1}{Q}} T_{n\frac{N}{P}} \mathbf{g}. \tag{E.84}$$

The last equation is the sum of PQ N-length vectors and can be computed using the Inverse Discrete Fourier Transform as below

$$t[l] = \sum_{n=0}^{P-1} \sum_{m=0}^{Q-1} c_{n,m} e^{\jmath \frac{2\pi ml}{Q}} f_{n\frac{N}{P}}[l] = \sum_{n=0}^{P-1} f_{n\frac{N}{P}}[l] \sum_{m=0}^{Q-1} c_{n,m} e^{\jmath \frac{2\pi}{N} lmr}. \tag{E.85}$$

Denoting $c_n(m) = c_{n,m}$ we can define

$$\mathbf{c}_n = [c_n[0], c_n[1], \ldots, c_n[Q-1]], \tag{E.86}$$

and its up-sampled version

$$U(\mathbf{c}_n)[k] = \begin{cases} c_n[k/r], \ k/r \in \{0, 1, \ldots, Q-1\} \\ 0, \ \text{otherwise} \end{cases}. \tag{E.87}$$

Thus, the synthesis operator is such that

$$\begin{aligned} t[l] &= \sum_{n=0}^{P-1} f_{n\frac{N}{P}}[l] \sum_{m=0}^{Q-1} U(\mathbf{c}_n)[mr] e^{\jmath \frac{2\pi}{N} lmr} \\ &= \sum_{n=0}^{P-1} f_{n\frac{N}{P}}[l] IFFT\{U(\mathbf{c}_n)\}[l]. \end{aligned} \tag{E.88}$$

Figure E.5 presents a way to compute the expansion coefficients or the reconstructed signal when a frame is used either as an analysis or as a synthesis operator. Note that in the outputs of each FFT block, in Figure E.5, there is a serial to parallel converter, while in the IFFT inputs a parallel to serial converter exists. Figure E.5 shows that all the frame coefficients can be obtained using $PN(1 + \log_2(N))$ operations, which can be reduced if one takes into account that just Q FFT/IFFT coefficients need to be computed at each FFT/IFFT branch.

E.3.3 Weyl-Heisenberg Frames with Real Elements

The preceding frames have complex elements. In some practical applications, instead of complex elements one may need to construct a frame with real elements with the same frame bounds.

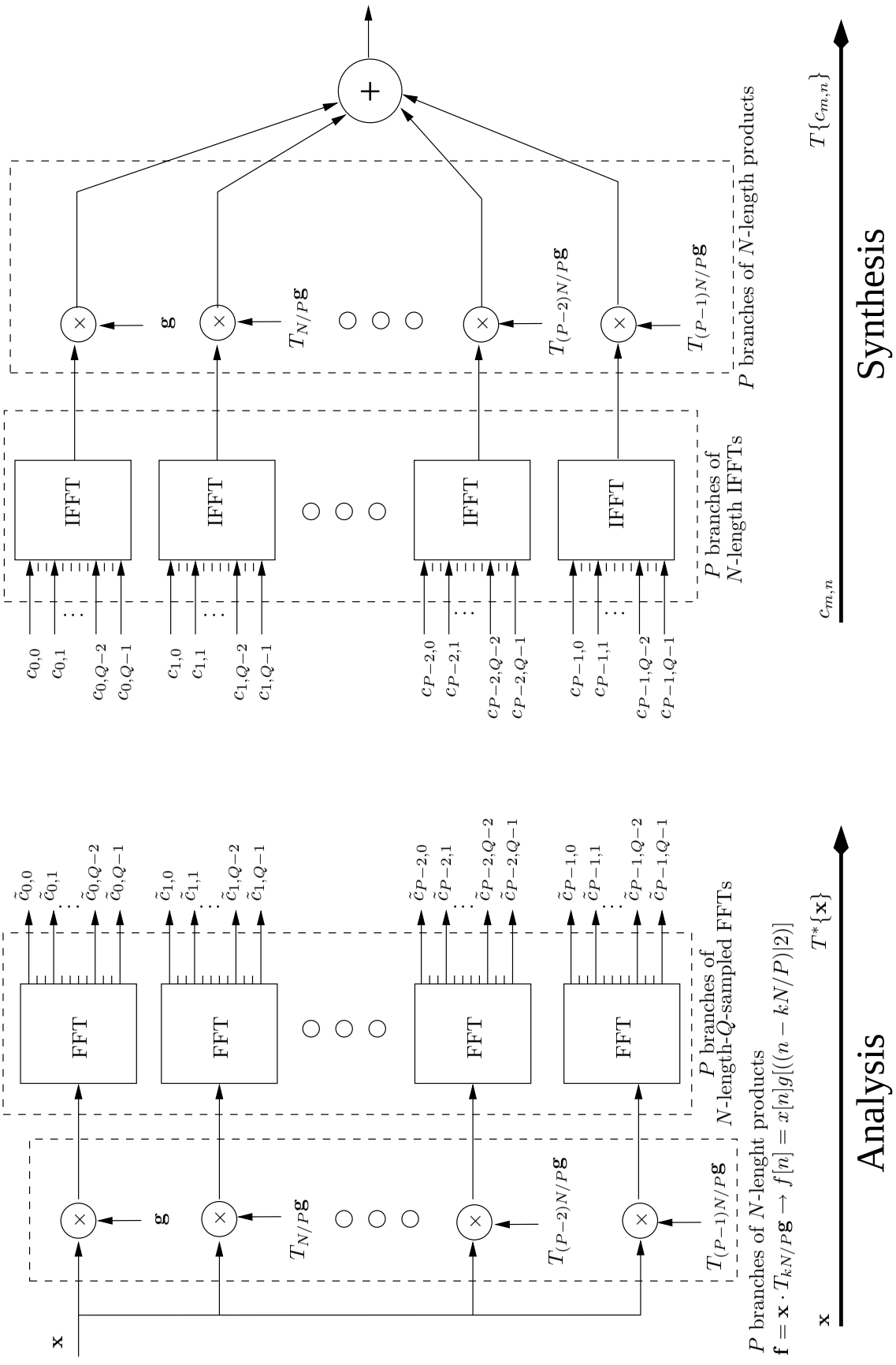
In order to do so, we first define the "real" modulation operator

$$E_{mb,\phi} g(t) = g(t) \cos(2\pi mb + \phi) \tag{E.89}$$

and generate the sequences

$$\mathcal{H}_1 = \{E_{mb,\phi_1} T_{na} g(t)\}_{m,n \in \mathbb{Z}} \tag{E.90}$$

$$\mathcal{H}_2 = \{E_{mb,\phi_2} T_{na} g(t)\}_{m,n \in \mathbb{Z}}. \tag{E.91}$$



Obs: The FFT block has a Serial to Parallel converter in its output. The IFFT has a Parallel to Serial converter in its input.

Figure E.5: Weyl-Heisenberg fast analysis and synthesis using FFTs and IFFTs.

In the sequel we show that if the phase difference between ϕ_1 and ϕ_2 is $\pi/2$ then $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$ is a frame with frame bounds that are equal to the frame bounds of the complex elements WH frame $\{E_{mb}T_{na}g(t)\}_{m,n\in\mathbb{Z}}$.

The Wigner-Ville distribution is a non-linear operator and therefore when applied to a sum of signals $x = x_1 + x_2$ it produces interference terms [84]. This way we have that

$$WD_{x_1+x_2}(t,f) = WD_{x_1}(t,f) + WD_{x_2}(t,f) + WD_{x_1,x_2}(t,f) + WD_{x_2,x1}(t,x) \tag{E.92}$$

where

$$WD_{x_1,x_2}(t,f) = \int_{-\infty}^{+\infty} x_1\left(t+\frac{\tau}{2}\right)\overline{x_2\left(t-\frac{\tau}{2}\right)}e^{-2\pi\jmath f\tau}d\tau \text{ and} \tag{E.93}$$

$$WD_{x_2,x_1}(t,f) = \int_{-\infty}^{+\infty} x_2\left(t+\frac{\tau}{2}\right)\overline{x_1\left(t-\frac{\tau}{2}\right)}e^{-2\pi\jmath f\tau}d\tau \tag{E.94}$$

Therefore, the elements of a phase-shifted real frame are given by

$$\begin{aligned} h_{na,mb,\phi}(t) &= g(t-na)\cos(mbt+\phi) = Re\{E_{mb}T_{na}g(t)e^{j\phi}\} = \\ &= \frac{1}{2}T_{na}g(t)\left(e^{j2\pi mbt}e^{\phi} + e^{-j2\pi mbt}e^{-\phi}\right). \end{aligned} \tag{E.95}$$

From the interference terms that appear in the WD of the sum of two signals it can be derived that

$$\begin{aligned} WD_{h_{na,mb,\phi}}(t,f) = \frac{1}{4}&\left[WD_g(t-na,f-mb) + WD_g(t-na,f+mb)\right] + \\ &+ \frac{1}{2}WD_g(t-na,f)\cos(2\pi mbt+2\phi). \end{aligned} \tag{E.96}$$

Suppose a frame $\mathcal{H}_1$ is generated with elements $h_{na,mb,\phi_1}(t)$, then the time frequency content of $\mathcal{H}$ is given by

$$\begin{aligned} WD_{\mathcal{H}_1}(t,f) &= \frac{1}{4}\sum_n\sum_m\left[WD_g(t-na,f-mb) + WD_g(t-na,f+mb)\right] + \\ &\quad + \frac{1}{2}\sum_n\sum_m WD_g(t-na,f)\cos(2\pi mbt+2\phi_1) \\ WD_{\mathcal{H}_1}(t,f) &= \sum_n\sum_m\frac{1}{2}WD_g(t-na,f-mb) + \\ &\quad + \frac{1}{2}\sum_n\sum_m WD_g(t-na,f)\cos(2\pi mbt+2\phi_1). \end{aligned} \tag{E.97}$$

Suppose another frame $\mathcal{H}_2$ is also generated whose elements are $h_{na,mb,\phi_2}(t)$. Now, the frame generated by the union of elements $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$ will have the time-frequency content $WD_{\mathcal{H}}(t,f) = WD_{\mathcal{H}_1}(t,f) + WD_{\mathcal{H}_2}(t,f)$ that is given by

$$WD_{\mathcal{H}}(t,f) = \sum_n \sum_m WD_g(t-na, f-mb) + \\ + \frac{1}{2} \sum_n \sum_m WD_g(t-na, f) \left[\cos\left(2\pi mbt + 2\phi_1\right) + \cos\left(2\pi mbt + 2\phi_2\right)\right] \tag{E.98}$$

To obtain $WD_{\mathcal{H}} = WD_{\mathcal{G}}$ (where $\mathcal{G}$ stands for the original WH frame, which has complex elements due to the modulation operator $E_{mb}g(t) = e^{-\jmath 2\pi mb}g(t)$) it suffices that

$$\frac{1}{2} \sum_n WD_g(t-na, f) \sum_m \left[\cos\left(2\pi mbt + 2\phi_1\right) + \cos\left(2\pi mbt + 2\phi_2\right)\right] = 0, \tag{E.99}$$

If $\sum_m \left[\cos\left(2\pi mbt + 2\phi_1\right) + \cos\left(2\pi mbt + 2\phi_2\right)\right] = 0$ then equation (E.99) holds. If for each individual m one has

$$\cos\left(2\pi mbt + 2\phi_1\right) + \cos\left(2\pi mbt + 2\phi_2\right) = 0, \tag{E.100}$$

then equation (E.99) is true. Therefore, either $\phi_2 = \phi_1 + \pi/2$ or $\phi_1 = \phi_2 + \pi/2$ are sufficient to ensure that $WD_{\mathcal{H}} = WD_{\mathcal{G}}$. It should be noted that this result is valid for any ranges of n and m in $\mathbb{Z}$.

Normalized Real Weyl-Heisenberg Frames As discussed in subsection E.3.1 when normalizing the elements of a frame composed of complex elements for $\mathbb{R}^N$, if the time-shifts are circular, A/B remains unchanged. When the frame has real elements defined solely from the sine or the cosine functions multiplied by the translated prototype the normalization of the elements changes the frame bounds; however, they do not change by the same amount anymore. This is so as, although the real WH frame will have the same frame bounds that the complex elements frame does, the elements of the real frame will have different norms, and then the resulting normalized frame will have frame bounds that are different than the ones of the original real frame.

Fast Analysis and Synthesis Operators If the phases used to generate frames with real elements are $\phi_1 = 0$ and $\phi_2 = \pi/2$, the analysis operators for real

WH frames can be computed using the same architecture used for complex elements real WH frames, see Figure E.5. Note that for these phases, to compute the real WH frame coefficients it suffices to take the real and complex parts of the outputs of the block diagram in Figure E.5. If the phases are not as above or the elements are normalized the simple modification of multiplying by the correct scaling constants the outputs (analysis operator) or the inputs (synthesis operator), also allows for fast computations. Hence, the analysis and synthesis operators of real WH frames also have a fast implementation algorithm.

E.4 Parameterized Dictionaries from Interlaced Weyl-Heisenberg Frames

It was discussed that, when it is desired to effectively represent a given class of signals using adaptive decompositions, parameterizable dictionaries are a good and natural choice. Parameterizable dictionaries have their elements $\mathbf{g}_i$ defined by a set of parameters, as in the Weyl-Heisenberg or wavelet frames construction approaches.

Sections B.3.2 and B.3.3 discussed dictionaries trade-offs and evaluation metrics, respectively. These discussions raise the question: How can one introduce new parameterizable atoms into a parameterizable dictionary $\mathcal{D}$ ("populate" $\mathcal{D}$) in order to improve the performance of the greedy decomposition algorithm, i.e., to obtain smaller reconstruction errors as the number of steps increases when compared to the original $\mathcal{D}$? The optimal solution would be to find all the vectors leading to the largest angles with their closest dictionary structures and incorporate these vectors in the dictionary. However, finding these vectors is a non-trivial task. Furthermore, supposing that one finds these vectors, it is most likely that they will not have the same parameterizable-shape as the structures that compose a parameterized dictionary.

Preferably, when new atoms introduced in a dictionary, not just the distortion at the step should be reduced but rather the rate×distortion tradeoff of signal representation is to be improved. If for example, $\#_{\mathcal{D}}$ is an exact power of two ($\#_{\mathcal{D}} = 2^W$), then the maximum increase in the dictionary cardinality, that has the lower impact on the coding rate, is to double $\#_{\mathcal{D}}$ (ignoring the coding rate of the coefficients).

Also, one can assume that, for each dictionary element $\mathbf{g}_\gamma$ there exists **at least one** vector $\mathbf{x}$ (and most likely more than one) among all the vectors that are coded using $\mathbf{g}_\gamma$, that produces the lower inner product $\langle \mathbf{x}, \mathbf{g}_\gamma \rangle$ (this signal is actually on the frontier of the Voronoi cell [23,57], see Definition G.1 in Section D.1.4). Then, in order to achieve an overall improvement of a dictionary in a rate×distortion sense, it would be interesting to introduce **at least** $\#_{\mathcal{D}}$ atoms in the dictionary. The introduction of $\#_{\mathcal{D}}$ elements can lead to an overall improvement of the rate×distortion if the new elements are correctly placed in the space.

It can be noted that the "frames interlacing" presented is efficient not just in the sense of generating "tighter" frames, but also in minimizing the inner product between each new element included in the dictionary and the elements already in the dictionary. The minimization of the inner product is relevant for the construction of dictionaries to be used in greedy signal decomposition algorithms. For these algorithms it is not enough to just enlarge $\#_{\mathcal{D}}$, it is interesting to do so while placing the new structures in the points having the largest distances from the elements already in the dictionary, what is actually related to the coherence metric of a dictionary

$$\mu(\mathcal{D}) = \max_{k \in [1,\ldots,\#_{\mathcal{D}}]} \left(\max_{j \in]1,\ldots,\#_{\mathcal{D}}] - \{k\}} |\langle \mathbf{g}_k, \mathbf{g}_j \rangle| \right). \tag{E.101}$$

New elements to be included in the dictionary must be designed should not impact the dictionary coherence substantially. Doing that, signals that were ill-decomposed using the original dictionary can be better decomposed (producing smaller residues at the MP steps) using the populated dictionary. Therefore, the "frame interlacing" approach presented in section E.3 is employed here to construct dictionaries, producing dictionaries composed of atoms that are parameterized and derived from a given predefined prototype signal.

In section E.3.1 we have seen that other two important metrics of dictionaries to be used in the MP algorithm are

$$\Theta(D) = \max_{\mathbf{x} \in \mathcal{X}} \left[\max_{i \in [1,\ldots,\#_{\mathcal{D}}]} (|\langle \mathbf{x}, \mathbf{g}_i \rangle|) \right] \text{ and} \tag{E.102}$$

$$\overline{\Theta}(\mathcal{D}) = E \left\{ \arccos \left[\max_{i \in [1,\ldots,\#_{\mathcal{D}}]} (|\langle \mathcal{X}, \mathbf{g}_i \rangle|) \right] \right\}, \tag{E.103}$$

where $\mathcal{D}$ is the dictionary and $\mathcal{X}$ is an N-dimensional iid Gaussian source (see section D.1.2). Table E.4 shows estimates of $\Theta(\mathcal{D})$ and $\overline{\Theta}(\mathcal{D})$ for the 9 frames presented

in Table E.1. The estimates in Table E.4 were obtained using the MP algorithm with 200 step decompositions of 1000 Gaussian noise signals (their orientations are uniformly distributed on the surface of the N-dimensional hyper-sphere, see section D.1.2). Table E.4 also shows the coherence of each of the 9 dictionaries. From the results in Table E.4 it can be seen that the dictionary $\mathcal{C}$ has a performance similar to the one of $\mathcal{F}$ with half the cardinality. Also, $\mathcal{C}$ has coherence measure very similar to the ones of $\mathcal{A}$ and $\mathcal{B}$. Note that $\mathcal{C}$ always has dictionary performance metrics $\Theta(\mathcal{D})$ and $\overline{\Theta}(\mathcal{D})$ close to the ones of $\mathcal{F}$. In some cases $\mathcal{D}$, $\mathcal{E}$, $\mathcal{G}$ and $\mathcal{H}$ may also lead to dictionary performance metrics close to the ones of $\mathcal{F}$ (depending on the Gaussian "spread" σ^2). Note that for "average" values of σ^2 (4 and 8) all dictionaries have similar performance metrics. However for these values of σ^2, $\mathcal{C}$ has lower $\mu(\mathcal{D})$ than $\mathcal{D}$–$\mathcal{H}$. The dictionary $\mathcal{C}$ has half of the elements of $\mathcal{F}$ and obtains performance metrics very close to the ones of $\mathcal{F}$ with lower coherence among its atoms.

Table E.4: Dictionary Performance Metrics.

	Metric/$\mathcal{D}$	$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$	$\mathcal{D}$	$\mathcal{E}$	$\mathcal{F}$	$\mathcal{G}$	$\mathcal{H}$	$\mathcal{M}$
σ^2	$\#_{\mathcal{D}}$	64	64	128	128	128	256	128	128	1024
2	$\Theta(\mathcal{D})$	79.0158	78.7032	**73.3615**	**73.3568**	77.8107	**72.7166**	**73.6985**	78.1283	**71.6059**
	$\overline{\Theta}(\mathcal{D})$	74.7854	74.8859	**67.2579**	**67.2414**	73.1671	**66.0883**	**67.2687**	73.2057	**63.8499**
	$\mu(\mathcal{D})$	**0.7346**	**0.7346**	**0.7346**	**0.7346**	0.9258	0.9258	**0.7346**	0.9258	0.9809
4	$\Theta(\mathcal{D})$	75.0269	75.1191	**73.5461**	**73.3922**	**74.4701**	**72.5942**	**73.7046**	**73.7399**	**72.5339**
	$\overline{\Theta}(\mathcal{D})$	69.3317	69.3767	**67.2123**	**67.3755**	**67.7916**	**65.5194**	**67.4083**	**67.7818**	**63.5583**
	$\mu(\mathcal{D})$	**0.5396**	**0.5396**	**0.6675**	0.7788	0.8571	0.8571	0.7788	0.8571	0.9622
8	$\Theta(\mathcal{D})$	75.3206	75.0785	**73.5876**	**73.7056**	**73.6926**	**73.5054**	**74.6438**	**74.2956**	**72.7682**
	$\overline{\Theta}(\mathcal{D})$	69.3851	69.4531	**67.2077**	**67.7792**	**67.2629**	**65.6424**	**67.8601**	**67.2451**	**63.6146**
	$\mu(\mathcal{D})$	**0.6065**	**0.6065**	**0.6483**	0.8825	0.7346	0.8825	0.8825	0.7346	0.9692
16	$\Theta(\mathcal{D})$	80.6769	80.9047	**74.1242**	79.6347	**73.9346**	**73.5657**	80.7496	**74.1546**	**72.7616**
	$\overline{\Theta}(\mathcal{D})$	77.1270	77.8409	**67.3896**	75.7875	**67.3841**	**66.4020**	77.1300	**67.3803**	**64.0668**
	$\mu(\mathcal{D})$	**0.7788**	**0.7788**	**0.7788**	0.9394	**0.7788**	0.9394	0.9394	**0.7788**	0.9845

One can gain further insight on the performances of the dictionary construction approaches discussed by changing P and Q (the time shift and frequency shift parameters of the WH frame) while maintaining the dictionary cardinality. Table

E.5 addresses this aspect where the dictionaries were generated using a Gaussian prototype function with $\sigma^2 = 4$ (this value was used as it led to dictionaries with similar performance, see Table E.4). The results in Table E.5 were estimated with the same procedure used to obtain the figures in Table E.4. The figures in Table E.5 show that $\mathcal{C}$ is more effective than the other dictionary generation structures as it leads to lower $\Theta(\mathcal{D})$ and $\overline{\Theta}(\mathcal{D})$ in average. It is worth noting that, in average, $\Theta(\mathcal{D})$ and $\overline{\Theta}(\mathcal{D})$ are lower for $\mathcal{C}$ than for $\mathcal{D}$, $\mathcal{E}$, $\mathcal{G}$ and $\mathcal{H}$ when P, Q and the dictionary cardinality are altered.

Table E.5: Dictionary evaluation changing P and Q while maintaining $\#_{\mathcal{D}}$.

	$P = 2$, $Q = 32$		$P = 4$, $Q = 16$		$P = 8$, $Q = 8$		$P = 16$, $Q = 4$	
	$\Theta(\mathcal{D})$	$\overline{\Theta}(\mathcal{D})$	$\Theta(\mathcal{D})$	$\overline{\Theta}(\mathcal{D})$	$\Theta(\mathcal{D})$	$\overline{\Theta}(\mathcal{D})$	$\Theta(\mathcal{D})$	$\overline{\Theta}(\mathcal{D})$
$\mathcal{C}$	**85.7066**	**83.1632**	**73.7664**	**67.7781**	**73.5461**	**67.2123**	**73.4896**	**67.3727**
$\mathcal{D}$	**85.7207**	**83.1723**	**73.9342**	**67.7929**	**73.3922**	**67.3755**	80.5967	76.1845
$\mathcal{E}$	89.7691	88.6939	85.6877	83.2296	**74.4701**	**67.7916**	**73.2987**	**67.3596**
$\mathcal{G}$	**85.6583**	**83.1490**	**74.0510**	**67.8011**	**73.7046**	**67.4083**	80.7627	77.2277
$\mathcal{H}$	89.8024	88.6631	85.7173	83.1736	**73.7399**	**67.7818**	**73.5004**	**67.3918**

It is easy to see from Figure E.4 that $\mathcal{A} \subset \mathcal{C}$, $\mathcal{B} \subset \mathcal{C}$ as well as $\mathcal{C} \subset \mathcal{F}$. Moreover, all the elements of dictionaries $\mathcal{A}$–$\mathcal{E}$, $\mathcal{G}$ and $\mathcal{H}$ are contained in $\mathcal{F}$. Each of the dictionaries $\mathcal{A}$, $\mathcal{B}$, $\mathcal{D}$, $\mathcal{E}$, $\mathcal{G}$ and $\mathcal{H}$ contains half of the elements that are in $\mathcal{C}$. On the other hand the other half of the elements of $\mathcal{C}$ are located at points that provide the smaller inner products, in the time-frequency plane, with the elements of dictionaries $\mathcal{A}$ and $\mathcal{B}$. This explains why $\mathcal{C}$ in general obtains better performance metrics than the others.

Table E.4 shows also the dictionary performance metrics for the dictionary $\mathcal{M}$. This dictionary is the maximally oversampled one, that is, it is generated when setting $P = Q = N$ what implies $\#(\mathcal{M}) = N^2$. It is noticeable that the dictionary evaluation metrics $\Theta(\mathcal{D})$ and $\overline{\Theta}(\mathcal{D})$ for dictionaries $\mathcal{C}$, $\mathcal{F}$ are reasonably close to ones of dictionary $\mathcal{M}$ for all values of σ^2. A natural question that arises is if dictionaries generated by $\mathcal{C}$ can provide better results in rate×distortion sense than $\mathcal{F}$ and $\mathcal{M}$ do.

Rate×Distortion Performance Now the performance of the "interlaced" dictionary is evaluated in a rate×distortion sense. For that we assume that the quantization errors of the coefficients are zero and that no bits are spent for sending them. Although this assumption is not practical, it allows for an investigation of the tendencies and behavior of the obtained decompositions. As a result, one can assume that the coding rate increases linearly with the number of elements M employed. Then, a trivial coding procedure gives a rate of

$$R = M\lceil \log_2(\#_{\mathcal{D}}) \rceil, \tag{E.104}$$

where s is the number of atoms used in the decomposition. For this rate×distortion experiment Gaussian signals in $\mathbb{R}^{64}$ are coded using 5 distinct dictionaries. Dictionary $\mathcal{A}$ is composed of the real and imaginary parts of the elements given by $E_{mb}T_{na}\mathbf{g}[k]$, as discussed in section E.3.3. The prototype function $\mathbf{g}$ is a Gaussian with $\sigma^2 = 4$, and with a and b being set to give 16 points in both time and frequency axes. In addition, the elements of $\mathcal{A}$ were also normalized. Dictionary $\mathcal{C}$ is formed by the union of $\mathcal{A}$ and its time- and frequency-shifted version by $a/2$ and $b/2$ respectively. Dictionary $\mathcal{F}$ is generated like $\mathcal{A}$ but with halved time- and frequency-shift parameters. $\mathcal{G}$ is generated from the union of $\mathcal{F}$ and its shifted version, that is, similarly to $\mathcal{C}$ with respect to $\mathcal{A}$. Dictionary $\mathcal{M}$ is the maximally oversampled with a and b set to provide N points in both time and frequency axes. The cardinalities of the designed dictionaries are then $\#(\mathcal{A}) = 512$, $\#(\mathcal{C}) = 1024$, $\#(\mathcal{F}) = 2048$, $\#(\mathcal{G}) = 4096$, $\#(\mathcal{M}) = 8192$.

Figure E.6 shows the histograms, with 250 bins, of the angle between an ensemble of signals drawn from an iid Gaussian source and their closest structures in the dictionary. These results were obtained using an ensemble of 10000 signals drawn from an iid Gaussian source as well as their residues in the first 40 iterations of their MP decompositions. The histograms of $\mathcal{C}$ and $\mathcal{G}$ show that, in both cases, the proposed scheme succeeds by improving the angle distribution, that is, they manage to reduce $\overline{\Theta}(\mathcal{D})$ as well as $\Theta(D)$ with respect to $\mathcal{A}$ and $\mathcal{F}$.

Figure E.7 shows on the left the mean error norm as a function of the step, while on the right it shows the mean error norm as a function of the rate – assumed here as the number of bits needed to code the dictionary element index multiplied by the number of steps. One can see that the dictionaries generated using the "interla-

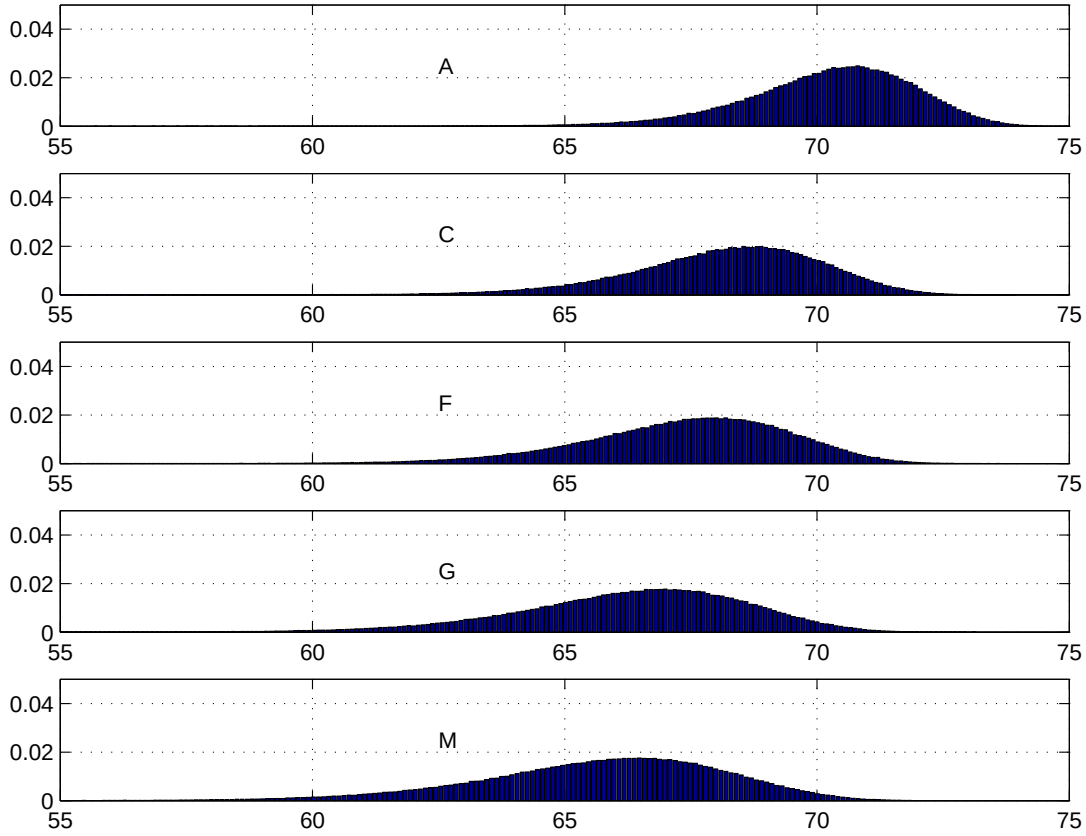


Figure E.6: Histograms of the angles between the residues and the structure used to approximate them for dictionaries $\mathcal{A}$, $\mathcal{C}$, $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{M}$ in $\mathbb{R}^{64}$.

cing" approach produce smaller residues as a function of the step when compared to the "non-interlaced" cases. Note that $\mathcal{G}$ does not produce a rate×distortion gain with respect to $\mathcal{F}$ as large as $\mathcal{C}$ obtains with respect to $\mathcal{A}$. This can be readily observed from the histograms in Figure E.6. From these histograms we can see that $\mathcal{G}$ does not obtain considerably lower $\Theta(\mathcal{D})$ and $\overline{\Theta}(\mathcal{D})$ with respect to the ones obtained by $\mathcal{F}$. In addition, in Figure E.7 we can see that $\mathcal{M}$ has a worse rate×distortion performance than $\mathcal{C}$, $\mathcal{F}$ and $\mathcal{G}$ have. Note that although $\mathcal{G}$ does not improves largely the rate×distortion characteristics of $\mathcal{F}$ as $\mathcal{C}$ does with respect to $\mathcal{A}$, $\mathcal{G}$ does not deteriorates the rate×performance.

E.5 Chapter Summary

In this chapter we presented the time-frequency content of frames, which is defined by the sum of the time-frequency content of the frame elements, where the time-frequency content of signals are measured using the Wigner-Ville distribution.

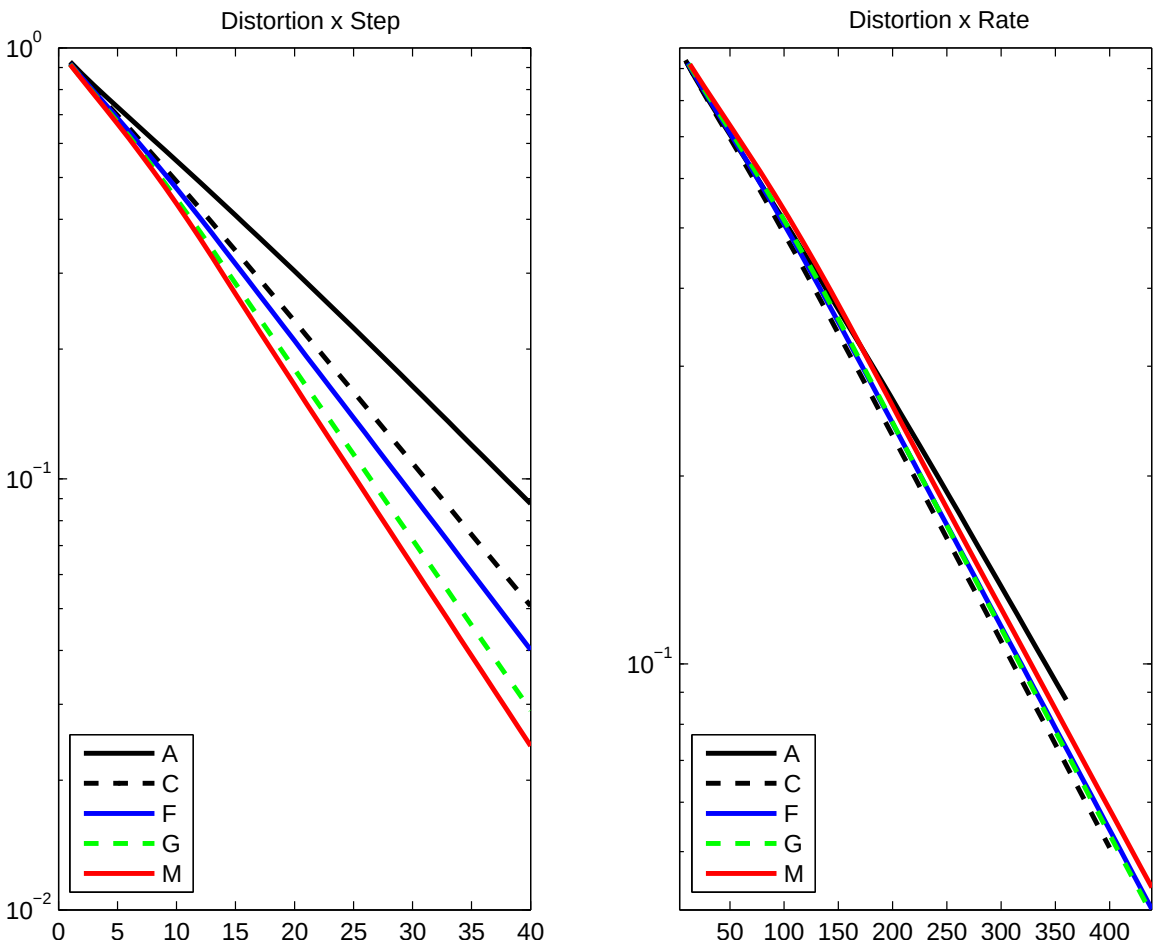


Figure E.7: Distortion at the step (left) and error×rate (right) for dictionaries $\mathcal{A}$, $\mathcal{C}$, $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{M}$ in $\mathbb{R}^{64}$.

A novel condition for frame characterization was presented that ensures that if the sum of the time-frequency content of the functions in a given set is positive definite then this set is a frame. The analysis of the time-frequency content of frames permits to characterize frames. The time-frequency content of frames was applied to study the generation of Weyl-Heisenberg frames from damped sinusoids.

The analysis of the time-frequency content of Weyl-Heisenberg frames generated from symmetric prototypes suggested ways to define the elements of a Weyl-Heisenberg frame $\mathcal{H}$ with respect to the elements of another Weyl-Heisenberg frame $\mathcal{G}$ in order to construct tighter Weyl-Heisenberg-like frames. This construction is accomplished by interlacing frames $\mathcal{H}$ and $\mathcal{G}$ in the time-frequency plane. This interlacing corresponds to a rhombus like lattice instead of a rectangular one in the time-frequency plane. Using the interlacing of Weyl-Heisenberg frames several frames were generated and evaluated in terms of their frame bounds. The results of

these evaluations show that the interlacing approach is effective for generating tighter frames. Experimental results have indicated that, when this interlacing is held in finite dimensional vector spaces, if the time- and frequency-shift parameters produce a number of points that equals half of the space dimension in each axis, then the interlaced frame is tight. This is still an unproven conjecture.

Although dictionaries do not need to satisfy the upper frame bound condition, any frame can be employed as a dictionary. It was conjectured that if two frames in an N-dimensional vector space have the same number of elements, then the tighter of the two frames is expected to provide the better results when used as a dictionary. The interlacing of Weyl-Heisenberg frames was used to generate dictionaries, and the experimental results, that evaluated dictionaries in terms of the statistics of the angle in Matching Pursuit decompositions for Gaussian sources, support this conjecture.

Indeed, the conjecture above suggests that the time-frequency interlacing of Weyl-Heisenberg frames permits to "populate" dictionaries constructed from Weyl-Heisenberg frames. The dictionaries then obtained using the rhombus like time-frequency parameters lattice ("interlaced") have a better performance, measured by the histograms of angle in Matching Pursuit steps, than the dictionaries of same cardinality constructed using a rectangular lattice.

Apêndice F

Conclusion

In this thesis some aspects related to signal decompositions using overcomplete dictionaries were addressed. In Chapter B a brief review of the theory behind these representations was presented. A practical application of signal representations using overcomplete dictionaries was developed and presented in Chapter C for the specific case of electric disturbance signals. When these signal representations are used for compression efficient quantization of the coefficients is required. In Chapter D we presented the design of Lloyd-Max quantizers for those coefficients. The time-frequency content of frames was defined and studied in Chapter E. The analysis of Weyl-Heisenberg frames using the the time-frequency content concept allowed to populate dictionaries formed by elements of Weyl-Heisenberg frames efficiently in terms of the rate×distortion performance.

The signal decomposition scheme presented in Chapter C provides coherent decompositions using a dictionary of parameterized atoms. These representations are robust to noise and achieve good signal compression ratios allied to an acceptable quality. The signal decomposition algorithm is based on the Matching Pursuits and it adaptively decomposes the signal using damped sinusoids. The signal representation provided by the algorithm is a sequence of coefficients and associated parameter sets which together define the structures identified in the signal. Feasible and effective computing procedures were presented to identify a signal structure at each algorithm iteration. In addition, a fast algorithm to eliminate pre-echo and post-echo artifacts that often appear in decomposition algorithms that are based on the Matching Pursuit was presented.

Greedy decomposition algorithms as the Matching Pursuit, often deviate from a physically meaningful (coherent) decomposition. In order to avoid such detour, that is, to obtain coherent decompositions, "intelligence" was included inside the decomposition loop. We included a heuristic inside the decomposition loop specifically designed for electric disturbance signals which instructs the greedy loop to select an appropriate atom according to the physical model of electric disturbance signals employed. The criterion employed to halt the decomposition process automatically detects the number of terms that should be employed to represent the signal using the damped sinusoids signal model.

The signal decompositions, obtained using the decomposition algorithm presented in Chapter C, are then used for lossy signal compression by quantizing both the coefficients of the signal representation and the parameter sets defining the structures in the signal representation. We should note that the compressed signal is reconstructed using parameter sets that differ from the ones that were obtained by the decomposition algorithm. Despite being very simple, this quantization procedure performed satisfactorily.

Since, in practice, electric disturbance signals are in general analyzed by system experts, as continuation we should submit the compressed signals to a subjective analysis. This way, we will be able to better evaluate the performance of the compression scheme presented.

Due to the nature of the signal model obtained by the decomposition algorithm in Chapter C, a weighted sum of damped sinusoids, it seems that these decompositions may be employed for other signal sources. This is another aspect to be investigated in the future, however, in this case, specific heuristics for the signal source must be designed and placed inside the decomposition loop, in order to obtain coherent decompositions.

In Chapter D, we addressed the design of Lloyd-Max quantizers for Matching Pursuit coefficients. This quantizers are designed using the statistics of the angles between the residue to be decomposed and the atom selected by the Matching Pursuit in each algorithm iteration. The empirical analysis of these angles using a Gaussian memoryless signal source indicated that they can be modeled as being independent and identically distributed across the Matching Pursuit iterations. In this

sense, we can consider that the statistics of these angles do not change in function of the Matching Pursuit iteration. Therefore, such angle statistics can be obtained from its statistics in the first Matching Pursuit iteration for the decomposition of signals drawn from a Gaussian memoryless signal source.

By observing the angles between the residue to be decomposed and the atoms selected in Matching Pursuit iterations one notes that when the dictionary includes a basis the Matching Pursuit may produce zero angles. This is equivalent to stating that the Matching Pursuit may obtain an exact signal representation using a finite number of terms. We provided a theorem proving the last claim.

The independent and identical distributed statistical model for Matching Pursuit angles was employed to design Lloyd-Max quantizers for Matching Pursuit coefficients. The design of these quantizers requires only an estimate of the probability density function of the Matching Pursuit angle for signals drawn from a Gaussian memoryless signal source. One should note that this requirement implies a dictionary dependent quantizer. The Lloyd-Max quantizers designed were shown to have a very good rate×distortion performance. In addition, the Lloyd-Max quantizer designed has error resilience intrinsically as a result of the use of the same quantization law for all the coefficients of the signal expansion.

The results obtained suggest that the Lloyd-Max quantizers designed using the statistics of Matching Pursuit angles for a Gaussian source can be also applied to other signal sources with very good rate×distortion performance. As suggestion for future work one could embed these quantizers in more specialized compression schemes that use greedy decompositions, such as the one presented in Chapter C or audio and video ones which abound in the literature.

In Chapter E we have defined the time-frequency content of a frame as the sum of the Wigner-Ville distributions of the frame elements. The time-frequency content of a frame permits to bound both the upper and the lower frame bounds. Applying this time-frequency content definition to a set of elements we can determine if this set is a frame. This is possible because we present a novel sufficient condition for knowing if a set of functions is a frame of a Hilbert space, which is: if the sum of the Wigner-Ville distributions of the elements in the set is greater than zero at any point of the time-frequency plane then the set is a frame of the space.

From the analysis of the time-frequency content of a Weyl-Heisenberg frame $\mathcal{G}$ generated from an even prototype function we envisioned a way to define another Weyl-Heisenberg-like frame $\mathcal{H}$ with respect to $\mathcal{G}$ such that the union of the elements of frames $\mathcal{G}$ and $\mathcal{H}$ is a tighter frame. This is accomplished by interlacing $\mathcal{H}$ and $\mathcal{G}$ in the time-frequency plane. The frame formed by the union of the two interlaced frames is equivalent to the use of a rhombus-like lattice for the time-frequency shift parameters that generate the frame instead of the rectangular one used to generate Weyl-Heisenberg frames.

The Weyl-Heisenberg frames interlacing procedure was employed to generate dictionaries to be employed in the Matching Pursuit. The dictionaries generated were evaluated using the statistics of the angle between signals drawn from a Gaussian source and the closest atom in the dictionary. In essence, we considered in the evaluation the largest and the average of that angle. The experimental results indicated that for dictionaries of same cardinality, the ones formed by the elements of interlaced Weyl-Heisenberg frames have better performance than the ones that are built up using the elements of a single Weyl-Heisenberg frame.

We conjecture that it is possible to find a similar definition and condition of frames content on a time-scale domain. This would provide a novel procedure to characterize wavelet frames. In addition, we are also in the quest for a procedure to "interlace" wavelet frames in a time-scale domain. This procedure could be used together with the interlacing of Weyl-Heisenberg frames to construct more efficient Gabor like dictionaries.

Another aspect that is to be investigated is the connection of the synthesis operator of Weyl-Heisenberg frames to multi-carrier modulation systems. This connection is suggested naturally in the derivation of the fast and efficient computation method for Weyl-Heisenberg frames analysis and synthesis operators in Chapter E, via the resemblance of the frame syntheses operator to Frequency Division Multiplexing systems.

Apêndice G

Proof of the Null Residue Proposition

Before proving the "null residue proposition" (see section D.1.4) we need some relevant definitions:

1. MP algorithm and dictionary set-up:

   (a) The MP algorithm is employed to decompose $\mathbf{x} \in \mathbb{R}^N$.

   (b) The dictionary cardinality $\#_{\mathcal{D}}$ is finite.

   (c) The dictionary is given by $\mathcal{D} \equiv \{\mathbf{g}_k\}_{k\in\{1,\ldots,\#_{\mathcal{D}}\}}$ and its elements have unit norms, i.e. $\|\mathbf{g}_k\| = 1 \quad \forall \quad k \in \{1, \ldots, \#_{\mathcal{D}}\}$.

   (d) All dictionary atoms are different, i.e. $|\langle \mathbf{g}_k, \mathbf{g}_j \rangle| < 1, \; \forall \, j \neq k \in \{1, \ldots, \#_{\mathcal{D}}\}$.

   (e) Since the atom selected at the n^{th} step $\mathbf{g}_{i(n)}$ and the resulting residue $\mathbf{r}_{\mathbf{x}}^n$ are orthogonal, i.e., $\mathbf{r}_{\mathbf{x}}^n \perp \mathbf{g}_{i(n)}$, then necessarily $i\,(n) \neq i\,(n+1)$, i.e., $\mathbf{g}_{i(n)} \neq \mathbf{g}_{i(n+1)}$.

2. Voronoi cells:

   Definition G.1 *The Voronoi region [57] associated to each dictionary element* $\mathbf{g}_k$ *is defined as*

   $$\mathcal{V}_k = \left\{ \mathbf{x} \in \mathbb{R}^N \mid \langle \mathbf{x}, \mathbf{g}_k \rangle > \langle \mathbf{x}, \mathbf{g}_j \rangle, \; j \in \{1, \ldots, \#_{\mathcal{D}}\} - \{k\} \right\}. \tag{G.1}$$

   (a) Since $\mathbf{g}_k \neq \mathbf{g}_j$ for $j \neq k$, $\|\mathbf{g}_k\| = 1$ and $\#_{\mathcal{D}}$ is finite the Voronoi region of each dictionary atom is not empty, that is, $\int_{\mathcal{V}_k} d\mathbf{x} > 0, \; \forall \, k \in \{1, \ldots, \#_{\mathcal{D}}\}$.

(b) Let

$$d = \min_{\mathbf{g}_j,\ \mathbf{g}_k \in \mathcal{D}} \|\mathbf{g}_j - \mathbf{g}_k\|, \tag{G.2}$$

then

$$\text{if } \|\mathbf{x}\| < \frac{d}{2} \text{ then } \mathbf{g}_j + \mathbf{x} \in \mathcal{V}_j,\ \forall \mathbf{g}_j \in \mathcal{D}. \tag{G.3}$$

(c) If $\mathbf{x} \in \mathcal{V}_k$ then $a\mathbf{x} \in \mathcal{V}_k$, $a \in \mathbb{R}_+^*$. This means that, the Voronoi regions generated by a dictionary of unit norm atoms are "conic" shaped with their vertexes at the origin $\vec{0}$.

(d) At the m^{th} MP step:

i. if $\mathbf{r}_{\mathbf{x}}^{m-1} \in \mathcal{V}_k$ then the MP selects $\mathbf{g}_{i(m)} = \mathbf{g}_k$, i.e. $i(m) = k$, for approximating $\mathbf{r}_{\mathbf{x}}^{m-1}$;

ii. if $\mathbf{r}_{\mathbf{x}}^{m-1}$ belongs to the boundary between two Voronoi regions any of the $\mathbf{g}_k$ defining these Voronoi regions can be used to approximate $\mathbf{r}_{\mathbf{x}}^{m-1}$.

3. Probability Distributions:

(a) The pdf of an outcome $\mathbf{x}$ from the source $\mathcal{X}$ is given by $f_{\mathcal{X}}(\mathbf{x})$.

(b) The pdf of the n^{th} residue is denoted using $f_{\mathcal{R}_{\mathbf{x}}^n}(\mathbf{r}_{\mathbf{x}}^n)$.

Definition G.2 *Let $P(n,k) = P\left(\mathbf{r}_{\mathbf{x}}^{n-1} \in \mathcal{V}_k\right)$ denote the probability of the residual to be decomposed in the n^{th} MP iteration being inside the Voronoi cell $\mathcal{V}_k$.*

$P(n,k)$ can be computed by means of

$$P(1,k) = \int_{\mathcal{V}_k} f_{\mathcal{X}}(\mathbf{x})\, d\mathbf{x}, \tag{G.4}$$

$$P(2,k) = \int_{\mathcal{V}_k} f_{\mathcal{R}_{\mathbf{x}}^1}(\mathbf{y})\, d\mathbf{y}, \tag{G.5}$$

$$\vdots$$

$$P(n,k) = \int_{\mathcal{V}_k} f_{\mathcal{R}_{\mathbf{x}}^{n-1}}(\mathbf{y})\, d\mathbf{y}. \tag{G.6}$$

4. Hyper-spheres:

Definition G.3 *Let $\alpha > 0$, $\epsilon > 0$, $\mathcal{B}_Q$ to be a set of Q vectors spanning a Q-dimensional hyper-plane $P_Q \subset \mathbb{R}^N$ and $\mathbf{g} \in P_Q$, we define*

$$S_Q(\alpha\mathbf{g}, \epsilon, N, \mathcal{B}_Q) \tag{G.7}$$

as a Q-dimensional hyper-sphere of radius ϵ centered at $\alpha\mathbf{g} \in \mathbb{R}^N$ ($\alpha \in \mathbb{R}$) contained in P_Q.

That is, any $\mathbf{x} \in P_Q$ is given by

$$\mathbf{x} = \sum_{i=1}^{Q} \beta_i \mathbf{g}_i,\ \mathbf{g}_i \in \mathcal{B}_Q \text{ and } \mathbf{g}_i \in \mathbb{R}^N \text{ such that } \langle \mathbf{g}_i, \mathbf{g}_j \rangle = \delta_{i,j},\ \forall\ i,\ j \in [1, \ldots, Q]. \tag{G.8}$$

For $\alpha > 0$ and $\mathbf{g} \in \mathbb{R}^N \cap P_Q$,

$$\text{if } \|\mathbf{x} - \alpha\mathbf{g}\| < \epsilon \text{ then } \mathbf{x} \in S_Q(\alpha\mathbf{g}, \epsilon, N, \mathcal{B}_Q). \tag{G.9}$$

(a) Note that if $\mathcal{B}_Q \subset \mathcal{B}_{Q+1}$ and $\mathbf{g} \in P_Q$ then

$$S_Q(\alpha\mathbf{g}, \epsilon, N, \mathcal{B}_Q) \subset S_{Q+1}(\alpha\mathbf{g}, \epsilon, N, \mathcal{B}_{Q+1}) \tag{G.10}$$

(b) We use $S_Q(\vec{0}, \epsilon, N, \mathcal{B}_Q)$ to denote a Q-dimensional hyper-sphere of radius ϵ centered at the origin.

After these definitions, we can state the "null residue proposition" as in the theorem below.

Theorem G.1 *Let a dictionary $\mathcal{D}$ include an orthonormal basis of $\mathbb{R}^N$ $\mathcal{B}_N = \{\mathbf{g}_{l_1}, \ldots, \mathbf{g}_{l_N}\} \in \mathcal{D}$, that is, there exists $\mathcal{I}_N = \{l_1, \ldots, l_N\}$ such that $\langle \mathbf{g}_{l_i}, \mathbf{g}_{l_j} \rangle = \delta(i - j)$, for i and $j \in \{1, \ldots, N\}$.*

If the signal source $\mathcal{X}$ has the following property: $\exists\ \mathbf{g}_{l_k}, l_k \in \mathcal{I}$, such that $\exists\ \alpha_{l_k},\ \epsilon_{l_k} > 0$ with

$$S_N(\alpha_{l_k}\mathbf{g}_{l_k}, \epsilon_{l_k}, N, \mathcal{B}_N) \subset \mathcal{V}_{l_k} \text{ and } f_{\mathcal{X}}(\mathbf{x}) > 0,\ \forall\ \mathbf{x} \in S_N(\alpha_{l_k}\mathbf{g}_{l_k}, \epsilon_{l_k}, N, \mathcal{B}_N). \tag{G.11}$$

Then

A. the MP has a non-zero probability of selecting only atoms that belong to $\mathcal{B}_N$ in the first $m \leq N$ steps, that is

$$P(i(1) \in \mathcal{I}_N, i(2) \in \mathcal{I}_N, \ldots, i(m) \in \mathcal{I}_N) > 0. \tag{G.12}$$

B. There is a non-zero probability of the residue being null after the N^{th} MP iteration, that is,

$$P(\mathbf{r}_{\mathbf{x}}^{N} = \vec{0}) > 0. \tag{G.13}$$

In order to prove this theorem we need some auxiliary results.

Lemma G.1 *Let a dictionary $\mathcal{D}$ include an orthonormal basis of $\mathbb{R}^N$, $\mathcal{B}_N = \{\mathbf{g}_{l_1}, \ldots, \mathbf{g}_{l_N}\} \in \mathcal{D}$, that is, there exists $\mathcal{I}_N = \{l_1, \ldots, l_N\}$ such that $\langle \mathbf{g}_{l_i}, \mathbf{g}_{l_j} \rangle = \delta(i-j)$, for i and $j \in \{1, \ldots, N\}$.*

Let the outcomes of the source $\mathcal{X}$ belong to an $(N-m)$ dimensional hyper-plane $P_{N-m} \subset \mathbb{R}^N$, with $0 < m < N$, which is generated by $\mathcal{B}_{N-m} \subset \mathcal{B}_N$ and represent the indices of the elements in $\mathcal{B}_{N-m}$ using $\mathcal{I}_{N-m}$. This means that

$$\mathbf{x} = \sum_{l_i \in \mathcal{I}_{N-m}} \beta_{l_i} \mathbf{g}_{l_i}. \tag{G.14}$$

If $\exists\ \epsilon > 0$ such that

$$f_{\mathcal{X}}(\mathbf{x}) > 0,\ \forall\ \mathbf{x} \in S_{N-m}(\vec{0}, \epsilon, N, \mathcal{B}_{N-m}), \tag{G.15}$$

then if the atom selected at the 1^{st} MP step is $\mathbf{g}_{l_k}$, i.e., $i(1) = l_k \in \mathcal{I}_{N-m}$, then

A. $\forall\ l_i \in \mathcal{I}_{N-m} - \{l_k\}$ there exists a hyper-sphere $S_{N-m-1}(\vec{0}, \epsilon_{l_k}, N, \mathcal{B}_{N-m-1})$, where $\mathcal{B}_{N-m-1} = \mathcal{B}_{N-m} - \{\mathbf{g}_{l_k}\}$, such that

$$f_{\mathcal{R}_{\mathbf{x}}^1}(\mathbf{y}) > 0,\ \forall\ \mathbf{y} \in S_{N-m-1}(\vec{0}, \epsilon_{l_k}, N, \mathcal{B}_{N-m-1}). \tag{G.16}$$

B. Therefore,

$$P(2, l_i \mid i(1) = l_k) > 0,\ \forall\ l_i \in \mathcal{I}_{N-m} - \{l_k\}. \tag{G.17}$$

Proof:

1. First we show that

$$\mathcal{V}_{l_i} \cap S_{N-m}(\vec{0}, \epsilon, N, \mathcal{B}_{N-m}) \neq \emptyset. \tag{G.18}$$

(a) Note that

$$\mathcal{V}_{l_i} \cap P_{N-m} \neq \emptyset \text{ and } \vec{0} \in \mathcal{V}_{l_i} \cap P_{N-m}. \tag{G.19}$$

(b) Note also that

$$\text{if } \mathbf{x} \in \mathcal{V}_{l_i} \cap P_{N-m} \text{ then } a\mathbf{x} \in \mathcal{V}_{l_i} \cap P_{N-m}. \tag{G.20}$$

(c) Let $\mathbf{x} \in \mathcal{V}_{l_i} \cap P_{N-m}$,

$$\text{if } a \leq \frac{\epsilon}{\|\mathbf{x}\|} \text{ then } a\mathbf{x} \in S_{N-m}(\vec{0}, \epsilon, N, \mathcal{B}_{N-m}). \tag{G.21}$$

□

2. *We now show that for each* $l_i \in \mathcal{I}_{N-m}$ *there is an* $(N-m)$ *dimensional hypersphere* $S_{N-m}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m})$ *such that*

$$\begin{aligned} S_{N-m}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m}) \cap \mathcal{V}_{l_i} = S_{N-m}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m}) \text{ and} \\ S_{N-m}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m}) \subset S_{N-m}(\vec{0}, \epsilon, N, \mathcal{B}_{N-m}). \end{aligned} \tag{G.22}$$

   (a) From Definition D.1 property (b) we have that for each $\mathbf{g}_{l_i} \in \mathcal{B}_{N-m}$ *exists* $d > 0$ *such that*

$$S_{N-m}(\mathbf{g}_{l_i}, d, N, \mathcal{B}_{N-m}) \cap \mathcal{V}_{l_i} = S_{N-m}(\mathbf{g}_{l_i}, d, N, \mathcal{B}_{N-m}). \tag{G.23}$$

   (b) From Definition D.1 property (c), we know that if (G.23) holds then

$$\forall\ a > 0,\ S_{N-m}(a\mathbf{g}_{l_i}, ad, N, \mathcal{B}_{N-m}) \cap \mathcal{V}_{l_i} = S_{N-m}(a\mathbf{g}_{l_i}, ad, N, \mathcal{B}_{N-m}). \tag{G.24}$$

   (c) It is known that $\forall\ \mathbf{y} \in S_{N-m}(a\mathbf{g}_{l_i}, ad, N, \mathcal{B}_{N-m})$, $\|\mathbf{y}\|$ *is such that*

$$|a(1-d)| \leq \|\mathbf{y}\| \leq |a(1+d)|\,. \tag{G.25}$$

   (d) Therefore, if we select

$$a < \frac{\epsilon}{|1+d|} \tag{G.26}$$

   then

$$S_{N-m}(a\mathbf{g}_{l_i}, ad, N, \mathcal{B}_{N-m}) \subset S_{N-m}(\vec{0}, \epsilon, N, \mathcal{B}_{N-m}). \tag{G.27}$$

   (e) Therefore, if we select $0 < \alpha_{l_i} < a$ *and* $0 < \epsilon_{l_i} < ad$ *then equation (G.22) is valid.*

□

3. *From item 1 we have that*

$$\mathcal{V}_{l_i} \cap S_{N-m}(\vec{0}, \epsilon, N, \mathcal{B}_{N-m}) \neq \emptyset \tag{G.28}$$

and from item 2 we know that $\exists\ S_{N-m}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m})$ *such that*

$$S_{N-m}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m}) \subset \left[\mathcal{V}_{l_i} \cap S_{N-m}(\vec{0}, \epsilon, N, \mathcal{B}_{N-m})\right]. \tag{G.29}$$

Therefore, from equation (G.15) we have that

$$f_{\mathcal{X}}(\mathbf{y}) > 0,\ \forall\ \mathbf{y} \in S_{N-m}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m}). \tag{G.30}$$

4. *Equation (G.30) implies that* $\forall\ l_i \in \mathcal{I}_{N-m}$

$$P(1, l_i) = \int_{\mathcal{V}_{l_i} \cap P_{N-m}} f_{\mathcal{X}}(\mathbf{y})d\mathbf{y} \geq \int_{S_{N-m}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m})} f_{\mathcal{X}}(\mathbf{y})d\mathbf{y} > 0. \tag{G.31}$$

5. *Since* $\mathcal{B}_{N-m} = \{\mathbf{g}_{l_1}, \ldots, \mathbf{g}_{l_{N-m}}\}$ *is an orthonormal basis of* $P_{N-m} \subset \mathbb{R}^N$*, any* $\mathbf{x} \in P_{N-m}$ *can be expressed as*

$$\mathbf{x} = \sum_{i=1}^{N-m} \gamma_{l_i}\mathbf{g}_{l_i} \tag{G.32}$$

This way, any $\mathbf{x} \in S_{N-m}(\alpha_{l_k}\mathbf{g}_{l_k}, \epsilon_{l_k}, N, \mathcal{B}_{N-m})$ *can be expressed as*

$$\mathbf{x} = \alpha_{l_k}\mathbf{g}_{l_k} + \sum_{i=1}^{N-m} \beta_{l_i}\mathbf{g}_{l_i} \text{ and then } {\epsilon_{l_k}}^2 \geq \|\mathbf{x} - \alpha_{l_k}\mathbf{g}_{l_k}\|^2 = \sum_{i=1}^{N-m} {\beta_{l_i}}^2. \tag{G.33}$$

6. *If the MP iteration selects* $\mathbf{g}_{l_k}$*, i.e.,* $i(1) = l_k$*, to approximate* $\mathbf{x}$ *then the resulting residue is* $\mathbf{r}^1_{\mathbf{x}} = \mathbf{x} - \gamma_{l_k}\mathbf{g}_k$*. From equation (G.33) it follows that* $\gamma_{l_k} = \langle \mathbf{x}, \mathbf{g}_{l_k} \rangle = \alpha_{l_k} + \beta_{l_k}$ *then*

$$\mathbf{r}^1_{\mathbf{x}} = \sum_{l_i \in \mathcal{I}_{N-m} - \{l_k\}} \beta_{l_i}\mathbf{g}_{l_i} \ \rightarrow\ \|\mathbf{r}^1_{\mathbf{x}}\| = \sqrt{\sum_{l_i \in \mathcal{I}_{N-m} - \{l_k\}} {\beta_{l_i}}^2} \leq \epsilon_{l_k}. \tag{G.34}$$

Therefore, $\mathbf{r}^1_{\mathbf{x}}$ *is contained in an* $(N-m-1)$*-dimensional hyper-plane generated by the basis* $\mathcal{B}_{N-m-1} = \mathcal{B}_{N-m} - \{\mathbf{g}_{l_k}\}$.

7. *Equation (G.34) shows that* $\mathbf{r}^1_{\mathbf{x}}$ *is inside an* $(N-m-1)$*-dimensional hyper-sphere of radius* ϵ_{l_k}*, i.e.,* $\mathbf{r}^1_{\mathbf{x}} \in S_{N-m-1}(\vec{0}, \epsilon_{l_k}, N, \mathcal{B}_{N-m-1})$*. Therefore, from equation (G.15) we have that*

$$f_{\mathcal{R}^1_{\mathbf{x}}}(\mathbf{y} \mid i(1) = l_k) > 0,\ \forall\ \mathbf{y} \in S_{N-m-1}(\vec{0}, \epsilon_{l_k}, N, \mathcal{B}_{N-m-1}). \tag{G.35}$$

This proves claim A.

8. *As explained in items 1-3 above, for each* $l_i \in \mathcal{I}_{N-m} - \{l_k\}$, $\exists\ S_{N-m-1}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m-1})$ *such that*

$$\begin{aligned} &S_{N-m-1}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m-1}) \cap \mathcal{V}_{l_i} = S_{N-m-1}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m-1}) \\ &\quad \text{and } S_{N-m-1}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m-1}) \subset S_{N-m-1}(\vec{0}, \epsilon_{l_k}, N, \mathcal{B}_{N-m-1}). \end{aligned} \tag{G.36}$$

9. *Therefore,*

$$\begin{aligned} P(2, l_i \mid i(1) = l_k) &= \int_{\mathcal{V}_{l_i}} f_{\mathcal{R}^1_{\mathbf{x}}}(\mathbf{y} \mid i(1) = l_k) \geq \\ &\geq \int_{S_{N-m-1}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i})} f_{\mathcal{R}^1_{\mathbf{x}}}(\mathbf{y} \mid i(1) = l_k) > 0. \end{aligned} \tag{G.37}$$

This proves claim B. ■

Lemma G.2 *Let a dictionary* $\mathcal{D}$ *include an orthonormal basis of* $\mathbb{R}^N$, $\mathcal{B}_N = \{\mathbf{g}_{l_1}, \ldots, \mathbf{g}_{l_N}\} \in \mathcal{D}$, *that is, there exists* $\mathcal{I}_N = \{l_1, \ldots, l_N\}$ *such that* $\langle \mathbf{g}_{l_i}, \mathbf{g}_{l_j} \rangle = \delta(i-j)$, *for* i *and* $j \in \{1, \ldots, N\}$.

If the signal source $\mathcal{X}$ *has the following property:* $\exists\ \mathbf{g}_{l_k}$, $l_k \in \mathcal{I}_N$, *such that* $\exists\ \alpha_{l_k},\ \epsilon_{l_k} > 0$ *with*

$$S_N(\alpha_{l_k}\mathbf{g}_{l_k}, \epsilon_{l_k}, N, \mathcal{B}_N) \subset \mathcal{V}_{l_k} \text{ and } f_{\mathcal{X}}(\mathbf{x}) > 0\ \forall\ \mathbf{x} \in S_N(\alpha_{l_k}\mathbf{g}_{l_k}, \epsilon_{l_k}, N, \mathcal{B}_N), \tag{G.38}$$

then:

A. *If the atom selected at the* 1^{st} *MP step is* $\mathbf{g}_{l_k}$, *i.e.,* $i(1) = l_k$, *then* $\forall\ l_i \in \mathcal{I}_N - \{l_k\}$, *i.e.,* $\forall\ \mathbf{g}_{l_i} \in \mathcal{B}_{N-1} = \mathcal{B}_N - \{\mathbf{g}_{l_k}\}$, *there exists a hyper-sphere* $S_{N-1}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-1})$ *such that*

$$f_{\mathcal{R}^1_{\mathbf{x}}}(\mathbf{y}) > 0,\ \forall\ \mathbf{y} \in S_{N-1}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-1}). \tag{G.39}$$

B. *Therefore,*

$$P(2, l_i) > 0,\ \forall\ l_i \in \mathcal{I}_N - \{l_k\}. \tag{G.40}$$

Proof: *The proof is a particular case of the proof of Lemma G.1.*

1. *The assumption in equation (G.38) is equivalent to equation (G.30). We can then apply the proof of Lemma G.1, from step 4, showing that claim A is valid and also that the claim B of Lemma G.1 also holds for the source considered in the current Lemma.*

2. *Since l_i, $l_i \in \mathcal{I}_N - \{l_k\}$ can be chosen in the second pass of the MP even if l_k has not been chosen in the first one, we have that $\forall\ l_i \in \mathcal{I}_N - \{l_k\}$*

$$P(2, l_i) \geq P(1, l_k) P(2, l_i \mid i(1) = l_k) > 0. \tag{G.41}$$

This proves claim B. ■

Proof of Theorem G.1:

- **Preliminaries:** The MP algorithm is such that $\mathbf{g}_{i(n)} \perp \mathbf{r}_{\mathbf{x}}^n$; therefore if the atoms selected in m successive MP steps, i.e., from steps 1 to m belong to the basis $\mathcal{B}_N$ then the residue $\mathbf{r}_{\mathbf{x}}^m$ is orthogonal to all the atoms selected in steps 1 to m. Therefore, if $i(1), i(2), \ldots, i(m) \in \mathcal{I}_N$ then necessarily $i(m+1) \notin \{i(1), \ldots, i(m)\}$.

  Defining

$$\mathcal{I}_m^C = \{i(1), \ldots, i(m)\} \text{ and } \mathcal{I}_{N-m} = \mathcal{I}_N - \mathcal{I}_m^C \tag{G.42}$$

  the left side in equation (G.12) can be expressed as

$$\begin{aligned} &P(i(1) \in \mathcal{I}_N, i(2) \in \mathcal{I}_N, \ldots, i(m) \in \mathcal{I}_N) = \\ &\quad = P(i(1) \in \mathcal{I}_N,\ i(2) \in \mathcal{I}_{N-1}, \ldots, i(m) \in \mathcal{I}_{N-m-1}). \end{aligned} \tag{G.43}$$

  To prove the result in equation (G.12), that is, to prove

$$P(i(1) \in \mathcal{I}_N, i(2) \in \mathcal{I}_N, \ldots, i(m) \in \mathcal{I}_N) > 0, \tag{G.44}$$

  it suffices to prove that

$$P(i(1) \in \mathcal{I}_N) > 0 \tag{G.45}$$

$$P(i(1) \in \mathcal{I}_N, i(2) \in \mathcal{I}_{N-1}) = \tag{G.46}$$
$$P(i(2) \in \mathcal{I}_{N-1} \mid i(1) \in \mathcal{I}_N) P(i(1) \in \mathcal{I}_N) > 0$$
$$\vdots$$

$$P(i(1) \in \mathcal{I}_N, i(2) \in \mathcal{I}_{N-1}, \ldots, i(m+1) \in \mathcal{I}_{N-m}) = \tag{G.47}$$
$$= P(i(m) \in \mathcal{I}_{N-m} \mid i(1), \ldots, i(m) \in \mathcal{I}_N) P(i(1), i(2), \ldots, i(m) \in \mathcal{I}_N) > 0$$

- **Proof:** The proof will be by induction.

1. From Lemma G.2 we have that

$$P(i(1) \in \mathcal{I}_N, i(2) \in \mathcal{I}_{N-1}) \geq P(i(1) = l_k, i(2) \in I_{N-1}) = \\ = P(1, l_k) \sum_{l_i \in \mathcal{I}_N - \{l_k\}} P(2, l_i \mid i(1) = l_k) > 0. \tag{G.48}$$

That is, equation (G.12) is valid for $m = 2$.

2. Now we prove that if

$$P(i(1), i(2), \ldots, i(m) \in \mathcal{I}_N) > 0 \tag{G.49}$$

then

$$P(i(m+1) \in \mathcal{I}_{N-m} \mid i(1), \ldots, i(m) \in \mathcal{I}_N) > 0. \tag{G.50}$$

(a) Applying Lemma G.2 once and Lemma G.1 $(m-1)$ times we have that

$$\forall\ l_i \in \mathcal{I}_{N-m},\ \exists\ S_{N-m}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m}) \text{ such that} \\ S_{N-m}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m}) \cap \mathcal{V}_{l_i} = S_{N-m}(\alpha_{l_i}\mathbf{g}_{l_i}, \epsilon_{l_i}, N, \mathcal{B}_{N-m}). \tag{G.51}$$

(b) Therefore, if $i(1), i(2), \ldots, i(m) \in \mathcal{I}_N$ then

$$P(i(m+1) \in \mathcal{I}_{N-m} \mid i(1), \ldots, i(m) \in \mathcal{I}_N) > 0. \tag{G.52}$$

(c) Since

$$P(i(1), i(2), \ldots, i(m+1) \in \mathcal{I}_N) = \tag{G.53}$$
$$P(i(m+1) \in \mathcal{I}_{N-m} \mid i(1), i(2), \ldots, i(m) \in \mathcal{I}_N) P(i(1), i(2), \ldots, i(m) \in \mathcal{I}_N),$$

then, from equation (G.52), we have that if

$$P(i(1), i(2), \ldots, i(m) \in \mathcal{I}_N) > 0 \tag{G.54}$$

then

$$P(i(1), i(2), \ldots, i(m+1) \in \mathcal{I}_N) > 0. \tag{G.55}$$

3. Equation (G.53) shows that if the result in equation (G.12) is valid for m then it is also valid for $m+1$. Since equation (G.48) shows that equation (G.12) is true for $m = 1$ then equation (G.12) is true for any m.

This proves claim A.

4. If $i(1), i(2), \ldots, i(N-1) \in \mathcal{I}_N$ then $\mathcal{I}_1 = \mathcal{I}_N - \mathcal{I}_{N-1}^C = \{\mathbf{g}_{\hat{l}}\}$ and the resulting residue is

$$\mathbf{r}_{\mathbf{x}}^{N-1} = \alpha \mathbf{g}_{\hat{l}}. \tag{G.56}$$

That is, $\mathbf{r}_{\mathbf{x}}^{N-1}$ lies on the space generated by $\mathbf{g}_{\hat{l}}$, that is, $\mathbf{r}_{\mathbf{x}}^{N-1}$ has the same direction as $\mathbf{g}_{\hat{l}}$. Therefore, the MP makes $i(N) = \hat{l}$ yielding a null residue. Therefore,

$$\begin{aligned} P(\mathbf{r}_{\mathbf{x}}^{N} = \vec{0}) &= P(i(1), i(2), \ldots, i(N) \in \mathcal{I}_N) = \\ &= P(i(1), i(2), \ldots, i(N-1) \in \mathcal{I}_2) > 0. \end{aligned} \tag{G.57}$$

This proves claim B. ■

Corollary G.1 *Let the dictionary include an orthonormal basis of* $\mathbb{R}^N$ $\mathcal{B}_N = \{\mathbf{g}_{l_1}, \ldots, \mathbf{g}_{l_N}\} \in \mathcal{D}$, *that is, there exists* $\mathcal{I}_N = \{l_1, \ldots, l_N\}$ *such that* $\langle \mathbf{g}_{l_i}, \mathbf{g}_{l_j} \rangle = \delta(i-j)$, *for* i *and* $j \in \{1, \ldots, N\}$. *If the signals to be decomposed using the MP come from an iid Gaussian source then*

$$P(\mathbf{r}_{\mathbf{x}}^{n} = \vec{0}) > 0, \ n = N. \tag{G.58}$$

Proof:

1. *The* N*-dimensional iid Gaussian source is such that any* $\mathbf{x} \in \mathbb{R}^N$ *has a non null pdf. That is,*

$$f_{\mathcal{X}}(\mathbf{x}) > 0, \ \forall \ \mathbf{x} \in \mathbb{R}^N. \tag{G.59}$$

Inside $\mathbb{R}^N$ *we can place a hyper-sphere of radius* ϵ_k *centered at any* $\alpha_k \mathbf{g}_k$, $k \in \{1, \ldots, \#_{\mathcal{D}}\}$ *such that (this is a straightforward result from equation (G.3))*

$$\mathcal{V}_k \cap S_N(\alpha_k \mathbf{g}_k, \epsilon_k, N, \mathcal{B}_N) = S_N(\alpha_k \mathbf{g}_k, \epsilon_k, N, \mathcal{B}_N), \tag{G.60}$$

therefore

$$\begin{aligned} &\forall \ k \in \{1, \ldots, \#_{\mathcal{D}}\}, \ \exists \ S_N(\alpha_k \mathbf{g}_k, \epsilon_k, N, \mathcal{B}_N) \in \mathcal{V}_k, \\ &\text{such that } f_{\mathcal{X}}(\mathbf{x}) > 0, \ \forall \ \mathbf{x} \in S_N(\alpha_k \mathbf{g}_k, \epsilon_k, N, \mathcal{B}_N). \end{aligned} \tag{G.61}$$

2. *Since* $\mathcal{I}_N \subset \{1, \ldots, \#_{\mathcal{D}}\}$ *for any* $l_k \in \mathcal{I}_N$ *one can find a hyper-sphere satisfying equation (G.61). This implies that the iid Gaussian source satisfies equation (G.11).*

3. Therefore,

$$P\left(\mathbf{r}_{\mathbf{x}}^{N-1} = \vec{0}\right) > 0. \tag{G.62}$$

■

Referências Bibliográficas

[1] Adler, J., Rao, B. D., and Kreutz-Delgado, K. Comparison of basis selection methods. In *Proceedings of the 30th Asilomar Conf. on Signals, Systems and Computers* (November 1996), vol. 1, pp. 252–257.

[2] Al-Shaykh, O. K., Miloslavsky, E., Nomura, T., Neff, R., and Zakhor, A. Video compression using matching pursuits. *IEEE Trans. on Circuits and Systems for Video Technology 9* (February 1997), 123–143.

[3] Andrle, M., Rebollo-Neira, L., and Sagianos, E. Backward-optimized orthogonal matching pursuit approach. *IEEE Signal Processing Letters 11* (2004), 705 – 708.

[4] Benedetto, J. J., and Fickus, M. Frame potentials. *Advances in Computational Math 18* (2003), 357–385.

[5] Bölcksei, H., Hlawatsh, F., and Feichtinger, H. G. Frame-theoretic analysis of oversampled filter banks. *IEEE Trans. on Signal Processing 46*, 12 (August 1998), 3256–3268.

[6] Bujanowski, B., Pierre, J., Hietpas, S., Sharpe, T., and Pierre, D. A comparison of several system identification methods with application to power systems. In *Proceedings of the 36th Midwest Symposium on Circuits and Systems* (1993).

[7] Burrus, C. S., Gopinath, R. A., and Guo, H. *Introduction To Wavelets and Wavelets Transforms A Primer*, 1 ed. Prentice Hall, Upper Saddle River, New Jersey 07458, USA, 1998.

[8] Caetano, R. *Codificação de vídeo usando planos de bits generalizados*. D. Sc. Thesis, COPPE/UFRJ, Rio de Janeiro, RJ, 2004.

[9] Caetano, R., da Silva, E. A. B., and Ciancio, A. G. Matching pursuits video coding using generalized bit-planes. In *IEEE Inter. Conf. on Image Processing* (Rochester, NY, USA, September 2002).

[10] Canadian-American EMTP User Group. *EMTP Rule Book, Alternative Transients Rule Book*. Canadian-American EMTP User Group, 1987-1992.

[11] Candes, E. J., and Donoho, D. L. Ridgelets: a key to higher-dimensional intermittency? *Philosophical Transactions: Mathematical, Physical and Engineering Sciences 357* (1999), 2495–2509.

[12] Casazza, P., and Kovacevic, J. Equal-norm tight frames with erasures. *Advances in Computational Mathematics–Especial Issue on Frames* (2002), 387–430.

[13] Casazza, P. G., and Christensen, O. Weyl-heisenberg frames for subspaces of $L^2(\mathbb{R})$. *Proc. Amer. Math. Soc. 129* (2001), 145–154.

[14] Chen, S., Billings, S. A., and Luo, W. Orthogonal least squares methods and their application to non-linear systems identification. *International Journal of Control 50*, 5 (1989), 1873–1896.

[15] Chen, S., and Donoho, D. Basis pursuit. In *Proceedings of the 28th Asilomar Conf. on Signals, Systems and Computers* (September 1994), vol. 1, pp. 41–44.

[16] Chen, S. S., Donoho, D. L., and Saunders, M. A. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput 20* (1998), 33–61.

[17] Christensen, O. *An Introduction To Frames And Riesz Bases*, 1 ed. Applied and Numerical Harmonic Analysis. Birkhäuser, Boston, Basel, Berlin, 2002.

[18] Chui, C., and Shi, X. Inequalities of Littlewod-Paley type for frames and wavelets. *Siam J. Math. Anal. 24* (1993), 263–277.

[19] Chung, J., Powers, E. J., Grady, W. M., and Bhatt, S. C. Electric power transient disturbance classification using wavelet-based hidden Markov models. In *IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing* (2000).

[20] Claasen, T., and Mecklenbrauker, W. The aliasing problem in discrete-time wigner distributions. *IEEE Trans. on Acoustics, Speech, and Signal Processing 31* (1983), 1067–1072.

[21] Cohen, L. Time-frequency distributions - a review. *Proceeding of the IEEE 77*, 7 (July 1989), 941–981.

[22] Coifman, R. R., and Wickerhauser, M. V. Entropy-based algorithms for best-basis selection. *IEEE Trans. on Information Theory 38* (1992), 713–718.

[23] Conway, J. H., and Sloane, N. J. A. *Sphere Packings, Lattices and Groups*, 2 ed. A Series of Comprehensive Studies in Mathematics. Springer Verlag, New York, New York 10010, USA, 1993.

[24] Craizer, M., Fonini Jr., D. A., and da Silva, E. A. B. Alpha-expansions: A class of frame decompositions. *Applied and Computational Harmonic Analysis 13* (2003), 103–115.

[25] da Silva, E. A. B. *Wavelets Transforms for Image Coding.* Phd. thesis, University of Essex, Essex, England, June 1995.

[26] Daubechies, I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. on Information Theory 36*, 5 (1990), 961–1005.

[27] Daubechies, I. *Ten Lectures on Wavelets.* Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA, 1991.

[28] Daubechies, I., Grossman, A., and Meyer, Y. Painless nonhorthogonal expansions. *J. Math. Phys. 27* (1986), 1271–1283.

[29] Davis, G. *Adaptive Nonlinear Approximations.* Ph. D. Thesis, New York University, 1994.

[30] Davis, G., Mallat, S., and Avellaneda, M. Adaptive greedy approximations. *Journal of Constructive Approximation 13* (1997), 57–98.

[31] Davis, G., Mallat, S., and Zhang, Z. Adaptive time-frequency approximations with matching pursuits. In *Technical Report* (Courant Institute of Mathematical Sciences, Computer Science Department, New York University, 251 Mercer Street, New York, NY 10012, USA, 1999).

[32] DeVore, R. A. Nonlinear approximation. *Acta Numerica* (1998), 51–150.

[33] DeVore, R. A., and Temlyakov, V. N. Some remarks in greedy algorithms. *Advances in Computational Mathematics 5* (1996), 173–187.

[34] Diniz, P. S. R., da Silva, E. A. B., and Netto, S. L. *Digital Signal Processing: System Analysis and Design.* Cambridge University Press, 2001.

[35] Donoho, D., and Flesia, A. Can recent innovations in harmonic analysis 'explain' key findings in natural image statistics? *Network: Computation in Neural Systems, Taylor & Francis 12* (2001), 371–393.

[36] Donoho, D. L., Elad, M., and Temlyakov, V. Stable recovery of sparse overcomplete representations in the presence of noise. In *Technical Report* (Statistics Department, Stanford University, Stanford, CA 94305, USA – donoho@stat.stanford.edu http://www-stat.stanford.edu/ donoho/ Reports/2004/StableSparse-Donoho-etal.pdf, 2004).

[37] Donoho, D. L., Vetterli, M., DeVore, R. A., and Daubechies, I. Data compression and harmonic analysis. *IEEE Trans. on Information Theory 44*, 6 (October 1998), 2435–2476.

[38] Duffin, R. J., and Schaeffer, A. C. A class of nonharmonic fourier series. *Trans. Amer. Math. Soc. 72* (1952), 341–366.

[39] Durka, P. J., Ircha, D., and Blinowska, K. J. Stochastic time-frequency dictionaries for matching pursuit. *IEEE Trans. on Signal Processing 49* (2001), 507 – 510.

[40] Eldar, Y. C., and Bölcskei, H. Geometrically uniform frames. *IEEE Trans. on Information Theory 49*, 4 (April 2003), 993–1006.

[41] Eldar, Y. C., and Forney Jr., G. D. Optimal tight frames and quantum measurement. *IEEE Trans. on Information Theory 48*, 3 (2002), 599–610.

[42] Engan, K., Aase, S. O., and Husoy, J. H. Designing frames for matching pursuit algorithms. In *Proceedings of ICASSP'98* (12 - 15 Maio 1998), vol. 3, pp. 1817 – 1820.

[43] Engan, K., Aase, S. O., and Husoy, J. H. Designing frames for matching pursuits algorithms. In *IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing* (May 1998), pp. 1817–1820.

[44] Engan, K., Aase, S. O., and Husoy, J. H. Multi-frame compression: Theory and design. *Elsevier Signal Processing 80* (2000), 2121–2140.

[45] Etemoglu, C. O., and Cuperman, V. Matching pursuits sinusoidal speech coding. *IEEE Trans. on Speech and Audio Proc. 11* (2003), 413 – 424.

[46] Ferrando, S. E., Doolitle, E. J., Bernal, A. J., and Bernal, L. J. Probabilistic matching pursuit with gabor dictionaries. *Elsevier Signal Processing 80* (2000), 2099–2120.

[47] Ferrando, S. E., Kolasa, L. A., and Kovacevic, N. Algorithm 820: A flexible implementation of matching pursuit for gabor functions on the interval. *ACM Trans. on Mathematical Software (TOMS) 28* (September 2002), 337–353.

[48] Fickus, M., Johnson, B. D., Kornelson, K., and Okoudjou, K. A. Convolutional frames and the frame potential. *Applied and Computational Harmonic Analysis 19* (2005), 77–91.

[49] Fonini Jr., D. A. *Quantização por Aproximações Sucessivas.* Tese de D. Sc., COPPE/UFRJ, 2003.

[50] Fonteneles, L. H. *Codificadores de Vídeo usando Ridgelets e Planos de Bits Generalizados.* M. Sc. Thesis, COPPE/UFRJ, Rio de Janeiro, RJ, June 2004.

[51] Friedlander, B., and Porat, B. Detection of transient signals by the gabor representation. *IEEE Trans. on Acoustics, Speech, and Signal Processing 37*, 2 (February 1989), 169–180.

[52] Friedlander, B., and Zeira, A. Oversampled gabor representation for transient signals. *IEEE Trans. on Signal Processing 43*, 9 (September 1995), 2088–2094.

[53] Friedman, J. H., and Suetzle, W. Projection pursuit regression. *Journal of the American Statiscal Association 76* (1981), 817–823.

[54] Frossard, P., Vandergheynst, P., Figueras I Ventura, R. M., and Kunt, M. A posteriori quantization of progressive matching pursuit streams. *IEEE Trans. on Signal Processing 52*, 2 (February 2004), 525–535.

[55] Gabor, D. Theory of communications. *IEE Journal 93* (1946), 429–457.

[56] Galli, A. W., Heydt, G. T., and Ribeiro, P. F. Exploring the power of wavelet analysis. *IEEE Computer Applications in Power 9*, 4 (October 1996), 37–41.

[57] Gersho, A., and Gray, R. M. *Vector quantization and signal compression*. Kluwer Academic Publishers, Boston, Massachussets, USA, 1992.

[58] Goodwin, M. M. *Adaptive Signal Models: Theory, Algorithms, and Audio Applications*, 1 ed. Kluwer Inter. Series in Engineering and Computer Science. Kluwer, New York, NY, USA, 1998.

[59] Goodwin, M. M., and Vetterli, M. Matching pursuits and atomic signal models based on recursive filters banks. *IEEE Trans. on Signal Processing 47*, 7 (July 1999), 1890–1902.

[60] Goyal, V. K. Quantized overcomplete expansions in $\mathbb{R}^N$: Analysis, synthesis, and algorithms. No. UCB/ERL M95/97. Electronics Research Laboratory, College of Engineering, University of California, Berkeley.

[61] Goyal, V. K., Kovacevic, J., and Kelner, J. Quantized frame expansions with erasures. *Appl. and Comput. Harmonic Analysis 10* (2001), 203–233.

[62] Goyal, V. K., Vetterli, M., and Thao, N. T. Quantized overcomplete expansions in $\mathbb{R}^N$: Analysis, synthesis, and algorithms. *IEEE Trans. on Information Theory 44*, 1 (January 1998), 16–31.

[63] Gribonval, R. From projection pursuit and cart to adaptive discriminant analysis? *IEEE Transactions on Neural Networks 16* (2005), 522 – 532.

[64] Gribonval, R., and Bacry, E. Harmonic decomposition of audio signals with matching pursuit. *IEEE Trans. on Signal Processing 51* (2003), 101–111.

[65] Gribonval, R., Bacry, E., Mallat, S., Depalle, P., and Rodet, X. Analysis of sound signals with high resolution matching pursuits. In *IEEE Conf. Time-Freq. and Time-Scale Anal (TFTS'96)* (Paris, France, June 1996).

[66] Gribonval, R., and Nielsen, M. Sparse representations in unions of bases. In *Publication interne n^o 1499* (Institut de Recherche en Informatique et Systèmes Aléatoires, Campus de Beaulieu, 35042 Rennes Cedex, France – http://www.irisa.fr, 2002).

[67] Haykin, S. Signal processing: where physics and mathematics meet. *IEEE Signal Processing Magazine 18*, 4 (2001), 6–7.

[68] Heil, C., and Walnut, D. Continuous and discrete wavelet transforms. *SIAM Review 31* (1989), 628–666.

[69] Heusdens, R., Vafin, R., and Kleijn, W. B. Sinusoidal modeling using psychoacoustic-adaptive matching pursuits. *IEEE Signal Processing Letters 9* (2002), 262–265.

[70] Heusdens, R., and Van De Par, S. Rate-distortion optimal sinusoidal modeling of audio and speech using psychoacoustical matching pursuits. In *IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing* (Orlando, Florida, USA, May 2002), vol. 2, pp. 1809–1812.

[71] Hlawatsch, F., and Boudreaux-Bartels, G. F. Linear and quadratic time-frequency signal representations. *IEEE Signal Processing Magazine 9* (April 1992), 21–67.

[72] Jaggi, S., Carl, W., Mallat, S., and Willsky, A. S. High resolution pursuit for feature extraction. *Appl. Comput. Harmon. Anal. 5*, 4 (October 1998), 428–449.

[73] Jain, A. K. *Fundamentals of Digital Image Processing*, 13 ed. Prentice Hall Information And System Sciences Series. Prentice Hall, Englewood Cliffs, NJ 07632, 1989.

[74] Janssen, A. J. E. M. *Zak Transforms With Few Zeros and The Tie*. in "Advances in Gabor Analysis"(H.G. Feichtinger, T.Strohmer, eds.). Birkhäuser, Boston, 2002.

[75] Kauppinen, I., Kauppinen, J., and Saarinen, P. A method for long extrapolation of audio signals. *J. Audio Engeneering Society 49*, 12 (December 2002), 1167–1179.

[76] Krim, H., Tucker, D., Mallat, S., and Donoho, D. On denoising and best signal representations. *IEEE Trans. on Information Theory 45*, 7 (November 1999), 2225–2238.

[77] Kruskal, J. B. Toward a practical method to hel uncover the structure of a set of multivariate observations by finding the linear transformation which optimizes a new "index of condensation". In *Statistical Computation*, R. C. Milton and J. A. Nelder (eds). Academic Press, 1969.

[78] Liu, Q., Wang, Q., and Wu, L. Size of the dictionary in matching pursuit algorithm. *IEEE Trans. on Sig. Proc. 52* (2004), 3403 – 3408.

[79] Lobos, T., Rezmer, J., and Koglin, H.-J. Analysis of power systems transients using wavelets and Prony method. In *IEEE Porto Power Tech Conf.* (Porto, Portugal, 2001).

[80] Lovisolo, L. *Representações Coerentes de Sinais Elétricos*. M. Sc. Thesis, COPPE/UFRJ, Rio de Janeiro, RJ, April 2001.

[81] Lovisolo, L., and da Silva, E. A. B. Uniform distributions of points on an hyper-sphere with applications to vector bit-plane enconding. *IEE Proceeding on Vision, Image and Signal Processing 148* (June 2001), 187 – 193.

[82] Lovisolo, L., da Silva, E. A. B., Rodrigues, M. A. M., and Diniz, P. S. R. Coherent decompositions of power systems signals using damped sinusoids with applications to denoising. In *IEEE Inter. Symposium on Circuit and Systems* (Scottsdale, Arizona, USA, May 2002), vol. V, pp. 685–688.

[83] Lu, Y., Joshi, S., and Morris, J. M. Noise reduction for nmr fid via gabor transform. *IEEE Trans. on Biomedical Engeneering 44*, 6 (1997), 512–528.

[84] Mallat, S. *A Wavelet Tour of Signal Processing*, 1^{st} ed. Academic Press, San Diego, California, USA, 1998.

[85] Mallat, S., and Zhang, Z. Matching pursuits with time-frequency dictionaries. *IEEE Trans. on Signal Processing 41*, 12 (December 1993), 3397–3415.

[86] Munch, N. J. Noise reduction in tight weyl-heisenberg frames. *IEEE Trans. on Information Theory 38*, 2 (March 1992), 608–616.

[87] Neff, R., and Zakhor, A. Very low bit-rate video coding based on matching pursuits. *IEEE Trans. on Circuits and Systems for Video Technology 7* (February 1997), 158–171.

[88] Neff, R., and Zakhor, A. Modulus quantization for matching-pursuit video coding. *IEEE Trans. on Circuits and Systems for Video Technology 10* (2000), 895–912.

[89] Papandreou-Suppappola, A., and S.B.Suppappola. Analysis and classification of time-varying signals with multiple time-frequency structures. *IEEE Signal Processing Letters 9* (2002), 92–95.

[90] Pati, Y. C., Rezaiifar, R., and Krishnaprasad, P. S. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decompositions. In *27^{th} Annual Asilomar Conf. on Signals, Systems and Computers* (November 1993).

[91] Pillay, P., and Bhattachrjee, A. Application of wavelets to model short-term power system disturbances. *IEEE Trans. on Power Systems 11*, 4 (November 1996), 2031–2037.

[92] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. *Numerical Recipes in C – Theory of Scientific Computing*, 2 ed. Cambridge University Press, New York, NY, USA, 1997.

[93] Rebollo-Neira, L., and Lowe, D. Optimized orthogonal matching pursuit approach. *IEEE Signal Processing Letters 9* (2002).

[94] Rodrigues, M. A. M. *Efficient Decompositions for Signal Coding.* Ph. D. Thesis, COPPE/UFRJ, Rio de Janeiro, RJ, March 1999.

[95] Rodrigues, M. A. M., de Figueiredo, M. V. F., Miranda, A. L. L., and Diniz, S. S. Oscillography for power system operational planning. In *VII Symposium of Specialists In Electric Operational And Expansion Planning - SEPOPE* (Curitiba, Brazil, 23 - 28 May 2000).

[96] Sayood, K. *Introduction to Data Compression*, 2 ed. Morgan Kauffman Publishers, San Francisco, USA, 2000.

[97] Schweitzer III, E. O., and Hou, D. Filtering for protective relays. In *47th Anual Georgia Tech Protective Relaying Conf.* (Atlanta, Georgia, EUA, April 1993).

[98] Strohmer, T. Numerical analysis of the non-uniform sampling problem. *Computational Applied Mathematics 122* (2000), 297–316.

[99] Strohmer, T., and Heath Jr., R. W. Grassmannian frames with applications to coding and comunications. *Applied and Computational Harmonic Analysis 14*, 3 (2003), 257–275.

[100] Sun, W., and Zhou, X. Irregular wavelet/gabor frames. *Appl. Comput. Harmon. Anal. 13* (2002), 63–76.

[101] Tawfik, M. M., and Morcos, M. M. ANN-based techniques for estimating fault location on transmission lines using Prony method. *IEEE Transaction on Power Delivery 16*, 2 (April 2001).

[102] Tcheou, M. P. *Representação e Análise de Sinais de Oscilogarfia Usando Decomposições Adaptativas Redundantes.* Dissertação de M. Sc., COPPE/UFRJ, 2005.

[103] Temlyakov, V. N. Weak greedy algorithms. *Advances in Computational Mathematics 12*, 2-3 (2000), 213 – 227.

[104] Temlyakov, V. N. Nonlinear methods of approximation. *Foundations of Computational Mathematics 3*, 1 (2003).

[105] Tropp, J. A. Greed is good: algorithmic results for sparse approximation. *IEEE Trans. on Information Theory 50* (2004), 2231–2241.

[106] Tropp, J. A., Dhillon, I. S., Heath Jr., R. W., and Strohmer, T. Designing structured tight frames via an alternating projection method. *IEEE Trans. on Information Theory 51* (2005), 188–209.

[107] Vera-Candeas, P., Ruiz-Reyes, N., Rosa-Zurera, M., Martinez-Munoz, D., and Lopez-Ferreras, F. Transient modeling by matching pursuits with a wavelet dictionary for parametric audio coding. *IEEE Signal Processing Letters 11* (2004), 349 – 352.

[108] Vleeschouwer, C. D., and Macq, B. Content based and perceptual bit-allocation using matching pursuits. *Elsevier Signal Processing: Image Communication 16* (2001), 611–626.

[109] von Neumann, J. *Mathematische Grundlagen der Quantenmechanik*. Springer Berlin, 1932. English version: Mathematical Foundations of Quantum Mechanichs, Princeton Univ. Press, 1955.

[110] Xu, W. Component modeling issues for power quality assessment. *IEEE Power Engineering Review 21*, 11 (November 2001), 12–15,17.

[111] Young, R. *An introduction to nonharmonic Fourier series*. Academic Press, New York, 1980.

[112] Zhang, Z. *Matching Pursuits*. Ph. D. dissertation, New York University, 1993.

[113] Zibulski, M., and Zeevi, Y. Y. Oversampling in the gabor scheme. *IEEE Trans. on Signal Processing 43*, 9 (August 1993), 2679–2687.

[114] Zibulski, M., and Zeevi, Y. Y. Discrete multiwindow gabor-type transforms. *IEEE Trans. on Signal Processing 45* (1997), 1428–1442.