CRIPTOFONIA APLICADA A SISTEMAS MODERNOS DE COMUNICAÇÕES MÓVEIS

José Francisco de Andrade Junior

DISSERTAÇÃO 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 MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Aprovada por:

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

Prof. José Antonio Apolinário Junior, D.Sc.

Prof. Luiz Wagner Pereira Biscainho, D.Sc.

Prof. Sérgio Lima Netto, Ph.D.

Prof. Maurício Henrique Costa Dias, D.Sc.

RIO DE JANEIRO, RJ - BRASIL

OUTUBRO DE 2008

ANDRADE JR., JOSÉ FRANCISCO DE

Criptofonia Aplicada a Sistemas Modernos de Comunicações Móveis [Rio de Janeiro] 2008

XVI, 112 p. 29,7cm (COPPE/UFRJ, M.Sc., Engenharia Elétrica, 2008)

Dissertação - Universidade Federal do Rio de Janeiro, COPPE

1. Criptofonia
2. Scramblers
3. Comunicações Móveis
4. AMR
5. GSM
6. Voz

I. COPPE/UFRJ II. Título (série)

Agradecimentos

Em primeiro lugar, a Deus por tudo que conquistei na vida.

À minha esposa Ozaneide, pelo inestimável apoio, dedicação e compreensão, sem os quais eu não poderia ter concluído este trabalho; e ao meu filho Gabriel pelas infindáveis horas revitalizantes de alegria.

Aos meus pais, Andrade e Clotilde Fortunato (*in memoriam*), pelo zelo e dedicação a mim conferidos.

Aos meus avós Benedito (*in memoriam*) e Nazaré de Andrade, pelo amor, pela minha criação e educação, que, na falta da minha mãe, para mim foram tudo.

De forma especial, aos meus Orientadores, Professores Marcello Campos e José Apolinário, pelo incentivo, compreensão, confiança, amizade, ensinamentos e forma tranqüila com que conduziram a árdua tarefa de orientar.

Aos meus amigos, pelo constante incentivo e compreensão, sem os quais não poderia suportar os longos períodos de ausência do convívio fraternal.

À Marinha do Brasil, por ter permitido e apoiado a realização deste curso, em particular, agradeço ao Capitão-de-Fragata (EN) Rogério Correa Manso, pelo incentivo, amizade e orientação administrativa.

Aos Professores Luiz Wagner P. Biscainho, Antonio Petraglia e Mariane Petraglia, pelos ensinamentos e paciência que tiveram em relação às minhas perguntas e questionamentos realizados durante as aulas.

Aos membros da Banca examinadora, por terem aceitado o convite para fazer parte deste processo de avaliação.

Ao responsável pelo Laboratório de Voz do Instituto Militar de Engenharia (IME), por ter cedido arquivos de voz necessários à consecução deste trabalho.

Aos colegas Diego Haddad e Jorge Costa Pires Filho, pela amizade, companheirismo e proveitosas discussões sobre os tópicos ministrados nas disciplinas de Processamento de Sinais.

Finalmente, agradeço a todos, incluindo professores e funcionários do PEE, que, de alguma forma, colaboraram para o desenvolvimento deste trabalho.

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

CRIPTOFONIA APLICADA A SISTEMAS MODERNOS DE COMUNICAÇÕES MÓVEIS

José Francisco de Andrade Junior

Outubro/2008

Orientadores: Marcello Luiz Rodrigues de Campos

José Antonio Apolinário Junior

Programa: Engenharia Elétrica

Técnicas de criptofonia são utilizadas para transformar um sinal de voz em sinal ininteligível, cujo propósito é evitar escutas não autorizadas. Quando se deseja implementar sigilo em sistemas comerciais que empregam CODECs, tal como AMR (*Adaptive Multirate*) CODEC, a encriptação digital é uma opção adequada devido à necessidade de alterações internas de *hardware* e *software*. Se o sinal encriptado por técnicas digitais for aplicado diretamente ao CODEC, devido às suas características espectrais diferirem bastante daquelas apresentadas por um sinal de voz, a codificação pode resultar em um sinal de baixa qualidade. Por outro lado, cifradores analógicos podem ser empregados antes de codificadores de voz sem causar grandes alterações no desempenho do processo de codificação. Cifradores analógicos são adequados para sistemas de comunicações em que o grau de privacidade requerido não é crítico e as modificações de *hardware* seriam proibitivas em decorrência do elevado custo. Esta dissertação investiga o uso de diferentes técnicas de criptofonia aplicadas às comunicações móveis que empregam VOCODER. Especificamente para cifradores analógicos no domínio da freqüência, são apresentados resultados objetivos de qualidade para sinais aplicados ao CODEC AMR. Estes resultados são obtidos em termos de distâncias de Itakura e Cepstral e de valores PESQ. As distâncias espectrais permitem avaliar, tanto a inteligibilidade residual do sinal cifrado, quanto à qualidade do sinal decifrado. Os resultados de qualidade medidos pelo algoritmo PESQ são empregados para avaliar a qualidade do sinal decifrado. Este trabalho também propõe uma metodologia simples de seleção de chaves para criptofonia.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

SPEECH PRIVACY FOR MODERN MOBILE COMMUNICATION SYSTEMS

José Francisco de Andrade Junior

October/2008

Advisors: Marcello Luiz Rodrigues de Campos

José Antonio Apolinário Junior

Department: Electrical Engineering

Speech-privacy techniques are used to scramble clear speech into an unintelligible signal in order to avoid eavesdropping. Some analog speech-privacy equipments (scramblers) have been replaced by digital encryption devices (COMSEC), which have higher degree of security but require complex implementations and large bandwidth for transmission. However, if speech privacy is wanted in a mobile phone using a modern commercial CODEC, such as the AMR (Adaptive Multirate) CODEC, digital encryption may not be an option due to the fact that it requires internal hardware and software modifications. If encryption is applied before the CODEC, poor voice quality may result, for the VOCODER would handle digitally encrypted signal resembling noise. On the other hand, analog scramblers may be placed before the voice encoder without causing much penalty to its performance. Analog scramblers are intended in applications where the degree of security is not too critical and hardware modifications would be prohibitive due to its high cost. This thesis investigates the use of different techniques of voice scramblers applied to mobile communications VOCODERs. Specifically for frequency-domain speech scramblers, results for objective evaluation of quality are presented. Spectral distances perform residual intelligibility evaluation of scrambled signals and quality evaluation of recovered plain signals. The PESQ values could be employed to evaluate the quality of recovered plain signal. This work also proposes a new simple methodology to select keys for frequency-domain speech scramblers.

Sumário

Lista de Figuras

Lista de Tabelas

Lista de Nomenclaturas

3GPP - *The 3rd Generation Partnership Project.*

AMR - (*Adaptive Multi-Rate*) CODEC otimizado para codificação de voz. Este esquema de codificação/decodificação é o padrão utilizado para sinais de voz pelo 3GPP desde 1998. O CODEC AMR se baseia na qualidade do enlace para selecionar a taxa de compressão mais adequada.

DCT - (*Discrete Cosine Transform*) A Transforma Discreta de Cossenos é uma transformação matemática baseada em funções cossenos, sendo bastante empregada nos campos do processamento digital de sinais e compressão de dados. A DCT de comprimento N para um sinal para $x[n]$ pode ser definida como:
$C[k] = \alpha[k] \sum_{n=0}^{N-1} x[n] \cos\left[\frac{\pi(n+0{,}5)k}{N}\right]$, para $0 \leq k \leq N-1$.
onde
$$\alpha[k] = \begin{cases} \sqrt{\frac{1}{N}} & k = 0; \\ \sqrt{\frac{2}{N}} & 1 \leq k \leq N-1. \end{cases}$$

CODEC - Acrônimo para Codificador/Decodificador, dispositivo de hardware ou software que codifica/decodifica sinais.

COMSEC -(*Communications Security*) ou Comunicação Segura, que pode ser garantida por meio de métodos de criptofonia ou criptografia e demais elementos de segurança das comunicações

COTS - (*Commercial off-the-shelf*) denominação dada aos componetes de hardware e software e equipamentos comerciais de "prateleira".

DFT - (*Discrete Fourier Transform*) A Transformada Discreta de Fourier é uma das transformadas de Fourier, cuja aplicação transforma uma função resultante de um sinal amostrado no domínio do tempo (amostras) $x[n]$ em uma função discreta no domínio da freqüência $X[k]$. A DFT pressupõe um número finito de amostras não

nulas do sinal. A definição de DFT é a seguinte:
$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j(\frac{2\pi}{N}kn)}$, $k = 0, \ldots, N-1$.

DTFT - (*Discrete-Time Fourier Transform*)Transforma uma seqüência discreta no domínio do tempo $x[n]$ em uma função no domínio da freqüência $X(\omega)$. A DTFT é definida como: $X(\omega) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}$

FIR - (*Finite Impulse Response*) Sigla que representa filtros de resposta ao impulso finita. Este tipo de filtro digital é caracterizado por uma resposta ao impulso que se torna nula após um tempo finito, em contraste com os filtros IIR. Os filtros FIR apresentam algumas propriedades úteis tornando-os preferíveis frente aos filtros IIR: a) são intrinsecamente estáveis; b) não fazem uso de realimentação e, em consequência, os erros de arredondamento não se propagam; c) podem apresentar fase linear; e d) podem apresentar fase mínima.

FRS - (*Family Radio Service*) São transceptores portáteis que possuem 12 canais na faixa de UHF e utilizam modulação FM. A potência de transmissão é limitada a 500 mW, o que isenta o usuário da necessidade de licença de utilização emitida pela Agência Nacional de Telecomunicações.

GSM - (*Global System for Mobile Communication*) ou Sistema Global para Comunicações Móveis, que é uma tecnologia de comunicações móveis e o padrão mais difundido na telefonia celular. Os telefones GSM são utilizados por mais de 3 bilhão de pessoas em mais de 200 países.

IDCT - (*Inverse Discrete Cosine Transform*) Transformada Discreta de Cossenos Inversa. Para um sinal $C[k]$, a IDCT de comprimento N é definida como:
$x[n] = \sum_{k=0}^{N-1} \alpha[k]C[k] \cos\left[\frac{\pi(n+0,5)k}{N}\right]$, para $0 \leq n \leq N-1$.
onde $\alpha[k] = \begin{cases} \sqrt{\frac{1}{N}} & k = 0; \\ \sqrt{\frac{2}{N}} & 1 \leq k \leq N-1. \end{cases}$

IDFT - (*Inverse Discrete Fourier Transform*) A Transformada Discreta de Fourier Inversa transforma uma função discreta no domínio da freqüência $X[k]$ em uma

função discreta no domínio do tempo $x[n]$. A definição de IDFT é a seguinte: $x[n] = \frac{1}{N} \sum_{n=0}^{N-1} X[k] e^{j(\frac{2\pi}{N} kn)}$, $n = 0, \ldots, N-1$.

IDTFT - (*Inverse Discrete Time Fourier Transform*) Transformada inversa de Fourier para de sinais discretos é a função inversa da DTFT e definida como: $x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{j\omega n} d\omega$

IIR - (*Infinite Impulse Response*) Sigla que representa filtros de resposta ao impulso de duração infinita.

MODEM - Acrônimo para Modulador e Demodulador. Em geral, é um dispositivo de hardware ou software que modula uma forma de onda analógica com sinal digital adequando-a à transmissão analógica, e que demodula o sinal analógico e o reconverte para o formato digital original. Existem MODEMs que possibilitam esquemas de modulações e demodualaçoes digitais.

VOCODER - Abreviatura de Voice Coder ou codificador de voz. É um dispositivo destinado à codificação e decodificação de sinais de voz.

Capítulo 1

Introdução

Cada vez mais os sistemas de comunicações móveis de voz são utilizados para tratar de assuntos importantes, onde o sigilo se faz necessário e os torna alvos importantes de escutas não autorizadas.

Com o advento das comunicações móveis, onde cada usuário é um elemento integrante de redes de comunicações sem fio, a utilização de dispositivos de segurança para salvaguardar o sigilo das informações, outrora restritos aos sistemas de comunicações militares e governamentais, tornou-se mandatória. Este requisito ganha importância à medida que casos de escutas clandestinas tornam-se uma realidade recorrente.

Durante as Olimpíadas de Atenas, em 2004, mais de uma centena de telefones celulares foram "grampeados", incluindo os celulares do Primeiro Ministro e de sua esposa, do Ministro da Defesa, do Ministro da Justiça, do Ministro das Relações Exteriores, do Chefe do Estado-Maior da Marinha e de empregados da Embaixada Americana em Atenas, dentre outros. Este acontecimento ficou conhecido como *The Athens Affair* [1], que, em decorrência de problemas técnicos aliados à falta de procedimentos adequados de controle de acesso, continua sem solução.

A existência de concorrência acirrada nos setores da indústria motiva a espionagem industrial e torna o pessoal que ocupa "posições chave" nestes setores alvos de escutas e "grampos". Diante deste tipo de ameaça, cresce a premência da implementação de contramedidas no sentido de impedir e/ou dificultar a obtenção de informações privilegiadas por concorrentes.

Os modernos sistemas de comunicações governamentais, em geral, incorporam técnicas de criptofonia/criptografia, cuja aplicação garante o sigilo das informações de voz. O problema surge quando se tenta implementar mecanismos de criptofonia em

sistemas do tipo COTS (*Commercial Off-The-Self*) ou em sistemas legados, cujas alterações técnicas não são passíveis de execução ou são bastante onerosas. Como exemplo de sistemas COTS podem ser citados: telefones celulares, sistemas de comunicações por satélite, PoC (*Push-to-Talk over Cellular*) etc.

Quando se trata de sistemas móveis de comunicações de voz COTS, a existência do VOCODER [2] acrescenta uma variável adicional ao problema, pois, necessariamente, o sinal criptofonado deve possuir características de um sinal de voz [3]. Este requisito impede a utilização de técnicas de criptofonia digital, que geralmente expandem a banda do sinal original de voz, visando transformá-lo em um sinal com características espectrais de ruído dentro da banda de observação.

1.1 Propósito e Motivação

O propósito deste trabalho é estudar, por meio de simulações e testes de avaliação qualitativa, sistemas de criptofonia analógicos[1] aplicados a sistemas móveis de comunicações que empregam CODEC AMR/GSM (*Adaptive Multirate*) [4].

Como motivação para o presente trabalho, tem-se a necessidade de se prover um grau de privacidade, mesmo que seja este táctico ou casual, para sistemas de comunicações COTS que empregam VOCODER.

1.2 Estrutura da Dissertação

A presente Dissertação está estruturada em seis capítulos e quatro apêndices. As deduções e explanações que demandam cálculos ou detalhamento de assuntos teóricos acessórios são apresentados na forma de apêndices, enquanto que os tópicos mais diretamente correlacionados com o objeto do trabalho são apresentados na forma de capítulos, conforme detalhamento a seguir:

[1]Embora denominados sistemas de criptofonia analógicos, a implementação desses sistemas e a elaboração dos *softwares* de simulação e avaliação fazem uso de técnicas de processamento digital de sinais.

Capítulo 1: Este capítulo introduz e discute, de forma sucinta, os principais aspectos que motivaram o desenvolvimento e a formalização do problema objeto desta Dissertação. O propósito do trabalho, as motivações e a estrutura da Dissertação também fazem parte desta introdução.

Capítulo 2: Este capítulo apresenta as principais técnicas de criptofonia, dando enfoque à aplicação de cada uma das técnicas a sistemas de comunicações que empregam CODEC. Cada modalidade de criptofonia apresentada é classificada em relação ao nível de segurança (resistência à criptoanálise) e quanto à preservação da banda em relação ao sinal de voz original. Com o intuito de permitir uma comparação rápida e visual dos sinais cifrados, são apresentados espectrogramas comparativos entre os sinais originais e aqueles cifrados pelas respectivas técnicas. Para concluir o capítulo, é proposta uma metodologia de escolha de chaves para criptofonia.

Capítulo 3: Neste capítulo, é realizada uma descrição sucinta sobre sincronismos de amostra e quadro aplicados a sistemas de criptofonia. A abordagem adotada contempla o emprego de seqüências pseudo-aleatórias no papel de palavras de sincronismo. Noções básicas sobre a modulação AFSK e os efeitos do CODEC AMR sobre o sincronismo de quadro são detalhados. Para finalizar o capítulo, resultados de simulações para as diversas taxas de compressão do CODEC AMR e palavras de sincronismo de 60 e 110 ms são apresentados.

Capítulo 4: Neste capítulo, são discutidos os conceitos de medidas objetivas de qualidade para avaliação de sinais de voz. As medidas objetivas estão divididas em dois grupos: medidas não-perceptuais e medidas perceptuais. O grupo de medidas não-perceptuais é representado pelas distâncias de Itakura e Cepstral. Como medidas objetivas perceptuais são apresentados os algoritmos PSQM, PSQM+ e PESQ. As relações entre os resultados de cada algoritmo e o índice subjetivo MOS também são discutidas.

Capítulo 5: Neste capítulo, é apresentada uma descrição da metodologia e dos dados

utilizados para realização das simulações. Os resultados de cinco simulações são apresentados e discutidos. As três primeiras simulações permitem realizar comparações entre as técnicas denominadas CSI-F (ver Capítulo 2). As duas outras simulações apresentam resultados referentes ao período de troca de chaves e à inteligibilidade residual do sinal. Para finalizar, é realizada uma análise crítica dos resultados apresentados.

Capítulo 6: Neste capítulo, são apresentados os diversos resultados e conclusões obtidas ao longo do trabalho, bem como um resumo geral da Dissertação. Como complemento, uma breve discussão sobre possibilidades e sugestões de continuidade do trabalho são apresentadas.

Apêndice A: Este apêndice apresenta os fundamentos teóricos básicos para bancos de filtros. Um enfoque especial é dado aos bancos de filtros em termos de DFT uniforme, cuja implementação se dá por meio de componentes polifásicas. A complexidade computacional desta implementação é confrontada com a complexidade da implementação direta.

Apêndice B: Neste apêndice, é realizada a dedução da fórmula analítica para o cálculo do ângulo máximo $\Phi_{\mathrm{I}}^{\mathrm{max}}$.

Apêndice C: Neste apêndice, são apresentados os fundamentos básicos da análise, geração e detecção de sinais FSK. O tipo de detecção que abordada é a detecção ótima realizada por meio de correlatores. O resultado da detecção é, então, expresso em termos de razão de verossimilhança.

Apêndice D: Neste apêndice, são apresentadas características básicas do CODEC AMR.

Capítulo 2

Técnicas de Criptofonia

2.1 Introdução

A utilização de sistemas de criptofonia tem como marco inicial a Primeira Guerra Mundial. Em decorrência dos conflitos posteriores, estes sistemas começaram a ser utilizados pelos Governos, Forças Armadas, companhias telefônicas e Missões Diplomáticas. Com o advento dos semicondutores, foi possível construir sistemas de criptofonia mais seguros e que podiam operar de maneira mais amigável.

Os sistemas de criptofonia, de uma maneira macro, podem ser divididos em duas grandes classes (ver Figura 2.1):

a) Cifradores Analógicos ou Misturadores; e

b) Cifradores Digitais.

Os misturadores, também denominados *scramblers*, são sistemas de criptofonia analógicos no que se refere à informação produzida (sinal cifrado), embora atualmente todo o processamento do sinal seja digital. O *scrambler* foi inventado pouco antes da Segunda Grande Guerra pelos laboratórios Bell. O projeto, basicamente, realizava o produto de dois sinais (modulação) ou, em outras situações, realizava a subtração entre o sinal de informação e um sinal padrão predefinido (máscara), que geralmente era formado por um ruído. O processo de recuperação do sinal (*descrambler*) realizava-se por meio da aplicação das operações inversas correspondentes, e para tanto, o sinal padrão deveria ser conhecido.

Com a evolução, os *scramblers* começaram a realizar alterações em características

do sinal (amplitude, freqüência, espectro, seqüência temporal ect.). Como conseqüência direta, a forma de onda do sinal passou a ser alvo dos processos de criptoanálise [5].

Os cifradores analógicos apresentam níveis de segurança que variam de casual a táctico[1] e devem ser empregados somente em situações que não exijam níveis de segurança estratégicos.

Os cifradores digitais são conhecidos como sistemas de criptofonia digital ou sistemas COMSEC. Estes sistemas, ao invés de transmitirem partes do sinal de voz, enviam apenas os parâmetros produzidos na fase de análise do processo de codificação [2], o que permite a aplicação direta de técnicas de criptografia ao conjunto de parâmetros citado. Os cifradores digitais podem ser classificados em duas modalidades:

a) Categoria I - Informação codificada na forma digital e transmissão não-codificada na forma analógica. Este tipo de cifrador fornece os dados encriptados diretamente ao MODEM, que realiza o processo de modulação em banda base para codificar o sinal de voz e adequá-lo à largura de banda do canal e demais características analógicas do transmissor; e

b) Categoria II - Informação codificada (digital) e transmissão codificada (digital). Este tipo de cifrador se beneficia da capacidade do transmissor de receber dados no formato digital e, desta forma, fornece o sinal encriptado diretamente ao modulador.

Independentemente da classe do sistema de criptofonia utilizado, alguns requisitos importantes devem ser atendidos:

a) Largura de banda do sinal crifrado compatível com o canal de transmissão utilizado;

b) O sinal cifrado (voz) deve ser ininteligível ao ouvido humano, o que é equivalente a uma baixa inteligibilidade residual;

[1]Os níveis de segurança são classificados como **Casual**, **Táctico** ou **Estratégico**, de acordo com os recursos computacionais e o tempo necessários para realizar o processo de criptoanálise e conseqüente obtenção da respectiva informação.

c) A voz decifrada deve apresentar boa inteligibilidade e preservar as características (timbre e altura) de voz do locutor;

d) Baixo retardo (*delay*) nos processo de cifragem e decifragem do sinal, devendo, para os sistemas comunicações *full-duplex*, estar limitado ao máximo retardo permitido pelo sistema;

e) Resistência à criptoanálise adequada ao nível de segurança alcançado; e

f) Custo de implementação aceitável e compatível com o nível de segurança pretendido.

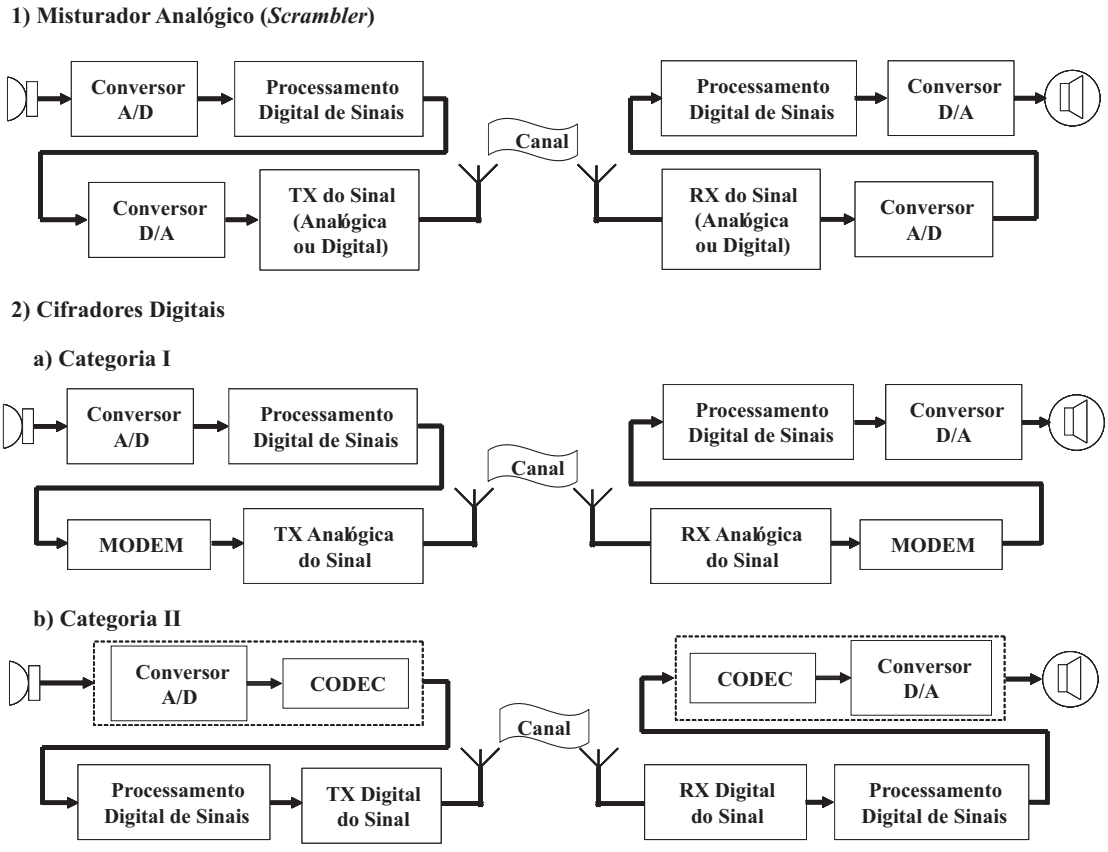

Figura 2.1: Classificação simples dos sistemas de criptofonia.

Tomando como base a classificação apresentada em [5] para os sistemas de criptofonia, este trabalho adotará as seguintes denominações para os sistemas de criptofonia:

I. Cifradores Analógicos (*Scramblers*)

Criptofonia por Segmentação de Informações (CSI):

i. CSI no Domínio do Tempo (CSI-T);

ii. CSI no Domínio da Freqüência (CSI-F):

- ❒ CSI-F em termos de Bancos de Filtros; e
- ❒ CSI-F em termos de Transformadas Ortogonais;

iii. CSI Bidimensionais:

- ❒ CSI Tempo-Freqüência (CSI-TF);
- ❒ CSI baseada em Matrizes de Hadamard (CSI-Hadamard)

II. Cifradores Digitais

a) Criptofonia Bit a Bit (CBB); e

b) Criptofonia por Parâmetros Analíticos (CPA).

2.2 Criptofonia por Segmentação da Informação (CSI)

Esta classe de criptofonia é formada por sistemas capazes de realizar a manipulação de elementos de informação[2] para níveis que não permitam ao ouvinte identificar a mensagem, produzindo uma baixa inteligibilidade residual[3].

Com o objetivo de aumentar a resistência à criptoanálise, a manipulação dos elementos de informação do sinal tenta tornar o espectro do sinal cifrado o mais plano possível.

2.2.1 CSI no Domínio do Tempo

Esta técnica realiza a criptofonia por meio de alterações na posição de segmentos de amostras temporais que compõem o sinal [7]. O nível de segurança resultante deste

[2]Amplitude, freqüência, espectro, seqüência de amostras temporal etc.

[3]Expressa a similaridade existente entre sinal cifrado é o sinal original. A inteligibilidade residual possui natureza subjetiva; contudo, em [6] foram propostos métodos objetivos para a sua medida indireta.

método pode variar de casual a tático.

A forma mais comum da implementação da CSI-T consiste em dividir o sinal de voz digitalizado $x(n)$ em blocos com duração típica[4] igual a $N \times 20$ ms, onde N é o número de segmentos de permutação utilizados. Cada bloco é dividido em N segmentos que, então, são permutados para formar os blocos cifrados. Antes de se realizar o processo de transmissão, deve-se converter o sinal de volta ao formato analógico.

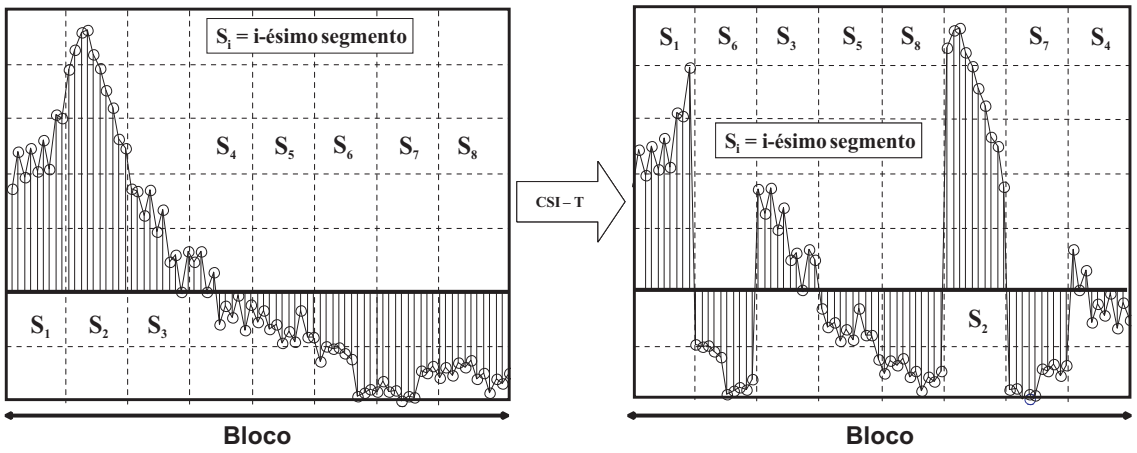

Figura 2.2: Bloco de um sinal de voz segmentado e cifrado com CSI-T.

Para um sinal de voz com M blocos, $\vec{\boldsymbol{x}}_m$, $m = 1, \ldots, M$, em que cada bloco possui N segmentos, cada segmento contendo R amostras do sinal. O i-ésimo bloco pode ser representado pelo vetor $\vec{\boldsymbol{x}}_i = [\vec{\boldsymbol{s}}_1{}^T \, \vec{\boldsymbol{s}}_2{}^T \cdots \vec{\boldsymbol{s}}_N{}^T]^T$, onde o i-ésimo segmento é definido como $\vec{\boldsymbol{s}}_j = \left[x^i_{\frac{(j-1)R}{N}+1} \; x^i_{\frac{(j-1)R}{N}+2} \; \cdots x^i_{\frac{(j-1)R}{N}+R} \right]^T$

Os elementos componentes do vetor $\vec{\boldsymbol{x}}_i$ podem ser rearranjados na forma matricial:

$$\boldsymbol{X}_i = [\vec{\boldsymbol{s}}_1 \; \vec{\boldsymbol{s}}_2 \; \cdots \; \vec{\boldsymbol{s}}_N]^T \tag{2.1}$$

[4]Para uma freqüência de amostragem $fs = 8$kHz, um segmento com 20 ms possui 160 amostras.

$$= \begin{bmatrix} x_1^i & x_{R+1}^i & \cdots & x_N^i \\ x_2^i & x_{R+2}^i & \cdots & \vdots \\ \vdots & \vdots & \ddots & x_{N\times R-1}^i \\ x_R^i & x_{2R}^i & \cdots & x_{N\times R}^i \end{bmatrix}_{R\times N} . \tag{2.2}$$

Pode-se, então, definir uma matriz de permutação $\boldsymbol{P}$ de ordem $N \times N$, cuja composição admite apenas um elemento não nulo em cada linha e em cada coluna. Para se garantir a preservação da energia do sinal, a norma da matriz de permutação P_i deve ser unitária, e para tanto, o elemento não nulo deve ter valor igual à unidade.

Realizando o produto das matrizes $\boldsymbol{P}$ e $\boldsymbol{X}_i$ e concatenando as linhas da matriz resultante, chega-se ao bloco do sinal de voz cifrada $\vec{\boldsymbol{y}}_i$:

$$\boldsymbol{Y}_i = \boldsymbol{P}\boldsymbol{X}_i = [\vec{\boldsymbol{y}}_1\, \vec{\boldsymbol{y}}_2\, \cdots\, \vec{\boldsymbol{y}}_N]^T \tag{2.3}$$

$$\boldsymbol{Y}_i = \begin{bmatrix} y_1^i & y_{R+1}^i & \cdots & y_N^i \\ y_2^i & y_{R+2}^i & \cdots & \vdots \\ \vdots & \vdots & \ddots & y_{N\times R-1}^i \\ y_R^i & y_{2R}^i & \cdots & y_{N\times R}^i \end{bmatrix}_{R\times N} \tag{2.4}$$

$$\vec{\boldsymbol{y}}_i = [y_1^i\, y_2^i\, \cdots\, y_{N\times R}^i]^T \tag{2.5}$$

O processo para decifrar o sinal é semelhante ao processo de cifragem, onde a matriz $\boldsymbol{P}$ é substituída por sua inversa, conforme detalhamento a seguir:

$$\boldsymbol{X}_i = \boldsymbol{P}^{-1}\boldsymbol{Y}_i = \boldsymbol{P}^{-1}\boldsymbol{P}\boldsymbol{X}_i \tag{2.6}$$

No receptor, $\boldsymbol{Y}_i$ pode ser obtida rearranjando-se o vetor $\vec{\boldsymbol{y}}_i$ na forma de uma matriz de ordem $R \times N$. Então $\vec{\boldsymbol{x}}_i$ é obtido pela cocatenação das linhas de $\boldsymbol{X}_i$ calculada pela Equação (2.6).

Por introduzir atrasos excessivos, a CSI-T não é adequada para o propósito deste trabalho, tendo sido apresentada somente com fins didáticos. O atraso é inevitável quando se faz uso de sistemas de CSI-T, pois o cifrador precisa dispor de um número N de segmentos para permutação antes da transmissão; isto causa um atraso de, no

mínimo, N vezes o comprimento do segmento. De maneira semelhante, a recuperação do sinal também introduz atrasos significativos.

A estimativa precisa do atraso provocado pelos esquemas de CSI-T depende do nível de segurança exigido, o que demonstra que o problema do atraso excessivo não pode ser tratado de maneira isolada. Um sistema típico, com N segmentos de T_s ms, apresenta um atraso total de $2NT_s$, que, para $N = 8$ e $T_s = 20$ ms, perfaz 320 ms.

Por outro lado, se segmentos menores que 20 ms forem utilizados não haverá preservação da banda do sinal de voz original [8].

Os seguintes fatores limitam a aplicação da CSI-T:

a) Introdução de atrasos demasiadamente grandes e que aumentam com o comprimento da chave de cifragem (número de permutações);

b) Processo de sincronismo crítico; e

c) Baixo número de chaves capazes de produzir inteligibilidade residual baixa [9].

2.2.2 CSI no Domínio da Freqüência

Os primeiros cifradores de CSI-F empregaram a técnica de inversão de freqüência, que consiste na inversão do espectro do sinal ou de parte deste com o intuito de tornar o sinal ininteligível aos ouvintes que não possuam receptores capazes de desfazer a inversão espectral do sinal. Estes inversores, devido à simplicidade de se desfazer o processo de criptofonia, não são mais empregados, exceto em rádios domésticos do tipo FRS, conhecidos comercialmente como *Talk-About*.

Com o surgimento de novos circuitos DSP, capazes de realizar tarefas complexas com alto nível de miniaturização, foi possível projetar sistemas de CSI-F implementados com bancos de filtros e transformadas ortogonais [6].

Se o número de sub-bandas (ou subfaixas) for suficientemente pequeno, o sinal apresentará inteligibilidade residual. Para superar este problema, deve-se escolher um número mínimo de sub-bandas e uma chave (permutação) dentre aquelas que geram baixa inteligibilidade residual. Os critérios para escolha de chaves serão abordados

na Seção 2.4. Outra forma de melhorar o desempenho dos sistemas CFI-F é realizar alterações nas chaves de maneira periódica e aleatória, de acordo com um polinômio gerador de seqüências pseudo-aleatórias.

Um sistema de CSI-F possui nível de segurança que varia de casual a tático e, para o caso em que se empregam seqüências pseudo-aleatórias de chaves, consegue-se melhorar a segurança pouco acima do nível tático.

2.2.2.1 CSI-F baseada em Bancos de Filtros

O diagrama da Figura 2.3 representa um banco de filtros (para maiores detalhes ver Apêndice A) com M subfaixas capazes de cobrir todo o espectro de sinal de voz a ser cifrado. Após a filtragem pelo conjunto de filtros de análise, $H_k(z)$, e decimação crítica por um fator M, as subfaixas são permutadas de acordo com a matriz de permutação $\boldsymbol{P}$.

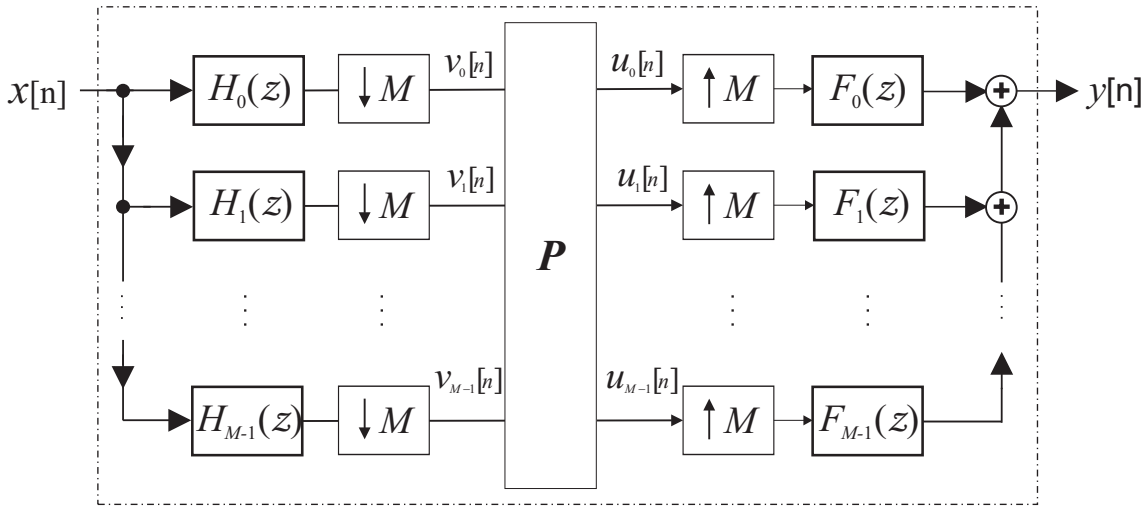

Figura 2.3: CSI-F baseado em banco filtros.

Considerando um banco de filtros do tipo DFT uniforme (ver Apêndice A), para um sinal de voz dividido em blocos, sendo o i-ésimo bloco representado pelo vetor $\boldsymbol{x}_i$, pode-se expressar matematicamente o processo de cifragem/decifragem.

As amostras $V_k^i[z]$ pertencentes ao i-ésimo bloco representam a k-ésima sub-banda

no domínio z, expressa como:

$$V_k^i[z] = H_k(z^{\frac{1}{M}})X_k^i[z^{\frac{1}{M}}] \quad ,\forall\ k = 0, 1, \ldots, M-1. \tag{2.7}$$

Os vetores $V_k^i[z]$ são, então, rearranjados na forma de uma matriz com $M \times N$ elementos, onde $N-1$ é a ordem do filtro $H_k(z)$:

$$\boldsymbol{V}_i = [V_0^i(z)\ V_2^i(z)\ \cdots\ V_{M-1}^i(z)]^T. \tag{2.8}$$

A multiplicação da matriz $\boldsymbol{V}_i$ pela matriz de permutação $\boldsymbol{P}$, expressa pela Equação (2.9), resulta na matriz $\boldsymbol{U}_i$, cujas linhas estão dispostas segundo a permutação aplicada.

Cada linha de $\boldsymbol{U}_i$ representa uma das sub-bandas do i-ésimo bloco cifrado, conforme:

$$\boldsymbol{U}_i = \boldsymbol{P}\boldsymbol{V_i} \tag{2.9}$$

$$\boldsymbol{U}_i = [U_0^i(z)\ U_2^i(z)\ \cdots\ U_{M-1}^i(z)]^T \tag{2.10}$$

$$Y_k^i[z] = F_k(z^M)U_k^i[z^M] \quad ,\ k = 0, 1, \ldots, M-1 \tag{2.11}$$

$$\boldsymbol{Y}_i = [Y_0^i(z)\ Y_2^i(z)\ \cdots\ Y_{M-1}^i(z)]^T. \tag{2.12}$$

As Equações (2.11) e (2.11) expressam o i-ésimo bloco do cifrado após ser processado pelos bancos de síntese. O sinal cifrado referente ao i-ésimo bloco é, portanto, obtido aplicando-se a inversa da transformada $\mathscr{Z}$ às linhas de $\boldsymbol{Y}_i$ após a interpolação, e efetuando-se o somatório elemento-a-elemento de cada linha de $\boldsymbol{Y}_i$:

$$\boldsymbol{y}_i(n) = \sum_{k=0}^{M-1} \mathscr{Z}^{-1}\left\{Y_k^i[z]\right\} = \sum_{k=0}^{M-1}\sum_{m=0}^{N-1} u_k^i[n]f_k(m - Mn). \tag{2.13}$$

Na Equação (2.13), $f_k(m - Mn)$ corresponde à resposta ao impulso do filtro $F_k(z^M)$.

Para decifrar o sinal criptofonado, pode-se utilizar o mesmo processo supracitado, tomando-se apenas o cuidado de substituir a matriz de permutação $\boldsymbol{P}$ por sua inversa $\boldsymbol{P}^{-1}$.

A Figura 2.4 apresenta espectrogramas do sinal de voz original e sua versão cifrada pela técnica CSI-F baseada em bancos de filtros do tipo DFT uniforme com 8 sub-bandas ($M = 8$).

A adoção de filtros com atenuação abrupta a partir da freqüência de corte confere a esta modalidade de CSI um importante diferencial, que é a imunidade à perda de sincronismo de quadro [10], tornando-a bastante atraente para projetos de baixo custo aplicados a equipamentos de arquitetura fechada.

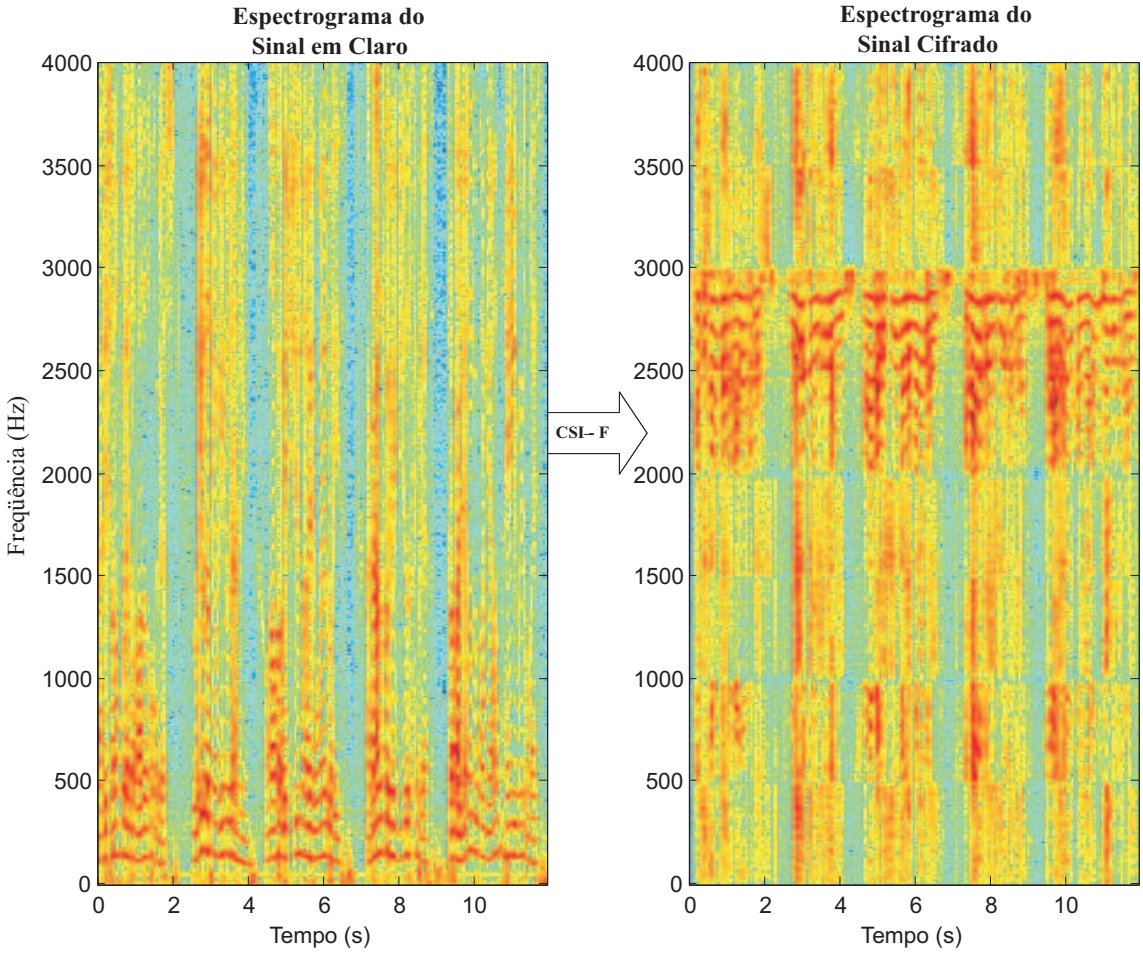

Figura 2.4: Espectrogramas de um sinal de voz e de sua versão cifrada obtida com CSI-F baseada em bancos de filtros.

2.2.2.2 CSI-F baseada em Transformadas Ortogonais

Crifradores baseados em transformações ortogonais [6]-[11] são também denominados *scramblers* no domínio da transformada. Como premissa, este trabalho de limitará a abordar as transformações ortogonais, diretas e inversas, que transformam sinais do domínio do tempo para o domínio da freqüência e vice-versa; contudo, a aplicação

do método não está limitada ao domínio da freqüência, dependendo apenas da transformação utilizada.

O processo de transformação do sinal, denominado transformada, é realizado por meio da multiplicação de cada bloco do sinal pela matriz de transformação. Cada bloco resultante, no domínio da transformada (freqüência), é dividido em N segmentos. Estes segmentos são permutados e reagrupados na forma de blocos com NM amostras do sinal, que, então, são transformados de volta ao domínio do tempo, conforme esquematizado pela Figura 2.5.

Matematicamente, tem-se:

$$\vec{\boldsymbol{v}}_i = \boldsymbol{T}\vec{\boldsymbol{x}}_i, \tag{2.14}$$

onde $\vec{\boldsymbol{x}}_i$ representa o i-ésimo bloco do sinal e contém NM amostras. A matriz $\boldsymbol{T}$ é uma matriz de transformação ortogonal com $NM \times NM$ elementos.

O vetor $\vec{\boldsymbol{v}}_i$ pode ser dividido em N segmentos e rearranjado na forma da matriz $\boldsymbol{V}_i$ com dimensão $N \times M$ cujas linhas representam os segmentos no domínio da transformada. A permutação é realizada, portanto, fazendo-se a multiplicação de $\boldsymbol{V}_i$ pela matriz de permutação $\boldsymbol{P}_{N\times N}$:

$$\boldsymbol{U}_i = \boldsymbol{P}\boldsymbol{V}_i. \tag{2.15}$$

A matriz de permutação é formada somente por uns e zeros e possui apenas um elemento não-nulo por linha e coluna, implicando que a matriz $\boldsymbol{U}_i$ seja resultante da permutação das colunas da matriz $\boldsymbol{V}_i$.

O sinal cifrado $\vec{\boldsymbol{y}}_i$ é obtido aplicando-se a transformação inversa $\boldsymbol{T}^{-1}$ ao vetor $\vec{\boldsymbol{u}}_i$, que é formado pela concatenação das linhas da matriz $\boldsymbol{U}_i$:

$$\vec{\boldsymbol{y}}_i = \boldsymbol{T}^{-1}\vec{\boldsymbol{u}}_i. \tag{2.16}$$

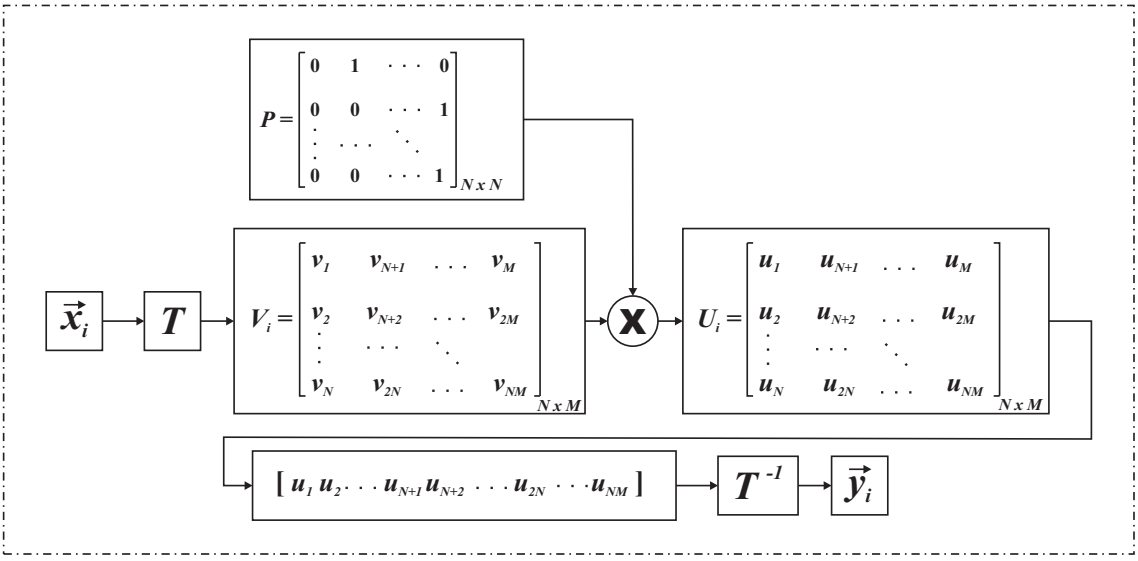

Figura 2.5: CSI-F baseada em transformadas ortogonais.

Para recuperar o sinal original, a cada bloco do sinal, aplica-se a mesma metodologia utilizada no processo de cifragem, com a substituição da matriz $\boldsymbol{P}$ por sua inversa $\boldsymbol{P^{-1}}$:

$$\tilde{\boldsymbol{x}}_i = \boldsymbol{T}^{-1}\boldsymbol{P}^{-1}\boldsymbol{T}\vec{\boldsymbol{y}}_i. \tag{2.17}$$

É importante mencionar que a escolha de transformações ortogonais unitárias assegura que o ruído adicionado pelo canal não tenha o seu valor amplificado durante o processo de recuperação do sinal, além de permitir o cálculo rápido[5].

Considerando η como sendo o ruído adicionado pelo canal e $\vec{\boldsymbol{y}}_i$ é o i-ésimo bloco do sinal transmitido, no receptor tem-se:

$$\tilde{\boldsymbol{y}}_i = \vec{\boldsymbol{y}}_i + \eta\,(t)\,. \tag{2.18}$$

O sinal decifrado é expressso como:

$$\tilde{\boldsymbol{x}}_i = \boldsymbol{\Phi}^{-1}\left\{\vec{\boldsymbol{y}}_i + \eta(t)\right\} \tag{2.19}$$

$$\tilde{\boldsymbol{x}}_i = \vec{\boldsymbol{x}}_i + \boldsymbol{\Phi}^{-1}\eta(t), \tag{2.20}$$

[5]Se uma matriz $\boldsymbol{M}$ é ortogonal e unitária, a sua inversa $\boldsymbol{M}^{-1}$ pode ser calculada como $\boldsymbol{M}^{T}$.

onde $\boldsymbol{\Phi}^{-1}$ e igual a:

$$\boldsymbol{\Phi}^{-1} = \boldsymbol{T}^{-1}\boldsymbol{P}^{-1}\boldsymbol{T} \tag{2.21}$$

Como a transformação $\boldsymbol{\Phi}^{-1}$ é ortogonal e, portanto, tem norma unitária $\left(\|\boldsymbol{\Phi}^{-1}\| = 1\right)$, o ruído no receptor não é afetado pelo processo de criptofonia. Pode-se então afirmar que se a matriz de transformação [6] é ortogonal, conforme demonstrado pela Equação (2.22), a energia do ruído não é afetada na recuperação do sinal.

$$\left\|\boldsymbol{\Phi}^{-1}\eta(t)\right\| = \|\eta(t)\|\,. \tag{2.22}$$

A Figura 2.6 apresenta espectrogramas do sinal de voz original e sua versão cifrada pela técnica CSI-F baseada na *Discrete Cosine Transform* (DCT) com 8 segmentos (sub-bandas, $M = 8$). Uma diferença perceptível entre os espectrogramas do sinal cifrado apresentados pelas Figuras 2.4 e 2.6 é a fronteira entre as sub-bandas, que na primeira é mais acentuada em decorrência da maior seletividade do banco de filtros.

Embora transformadas aplicadas a blocos do sinal realizem o mesmo trabalho de um banco de filtros com reconstrução perfeita [12], o sistema descrito nesta subseção é suscetível à perda de sincronismo. A principal diferença reside na seletividade dos filtros de cada subfaixa. Tomando-se como base a DCT, pode-se verificar que a sensibilidade ao sincronismo é decorrente da característica de filtragem pouco seletiva que esta transformada realiza em cada subfaixa.

[6] Se $\boldsymbol{T}$ é $\boldsymbol{P}$ são ortogonais e unitárias, $\boldsymbol{\Phi}^{-1} = \boldsymbol{T}^{-1}\boldsymbol{P}^{-1}\boldsymbol{T}$ também goza desta propriedade.

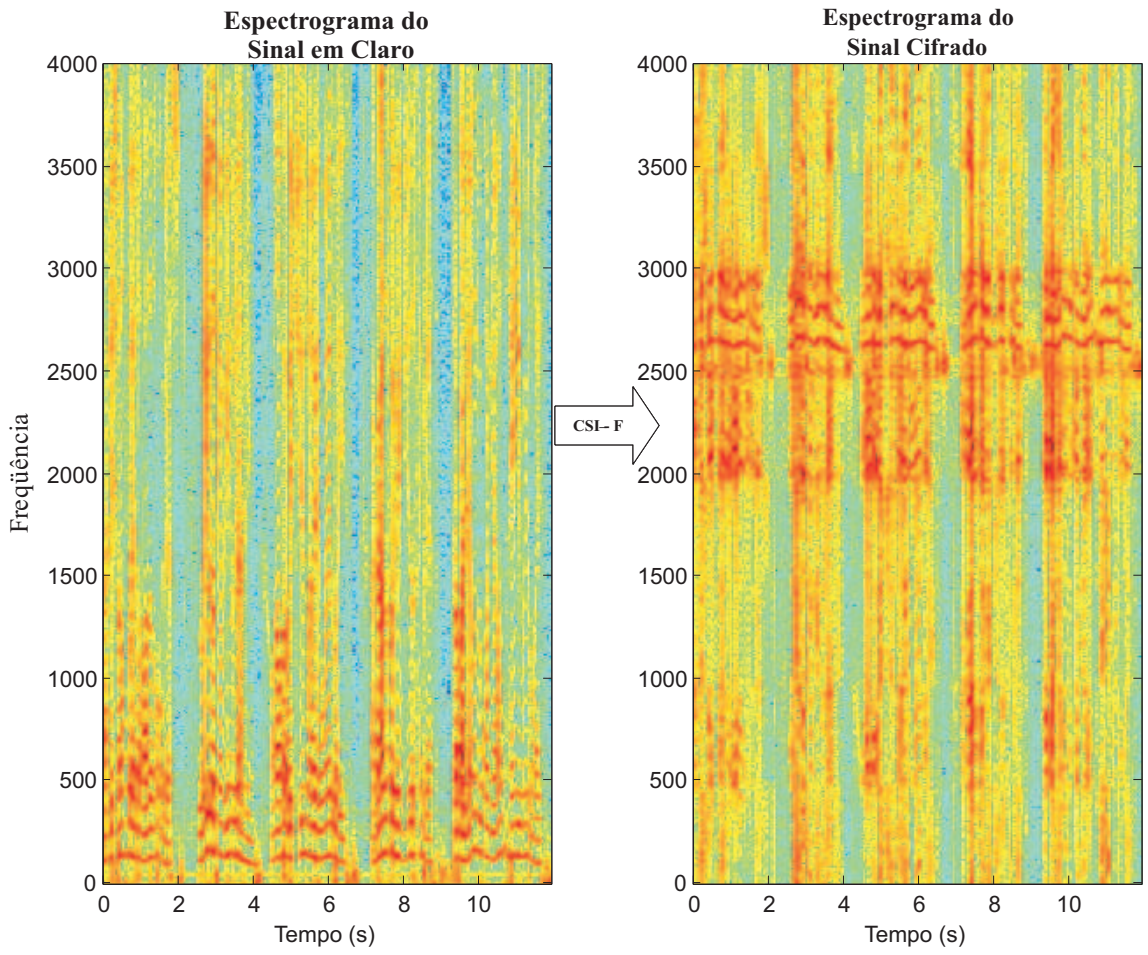

Figura 2.6: Espectrogramas de um sinal de voz e de sua versão cifrada obtida com CSI-F baseada em transformadas ortogonais.

2.2.3 CSI Bidimensionais

Os sistemas de CSI bidimensionais apresentam, dentre os métodos de CSI, os melhores resultados em termos de inteligibilidade residual e resistência à criptoanálise, permanecendo como uma alternativa aos cifradores digitais quando existem impedimentos de ordem técnica ou a relação custo-benefício não permite o seu emprego.

Os sistemas de CSI bidimensionais apresentam grande susceptibilidade à perda de sincronismo de quadro, sendo, portanto, necessária a implementação de mecanismos de sincronismo, que em muitos casos representam aumento de custo e complexidade do

projeto.

A seguir serão apresentadas duas categorias de CSI bidimensionais: CSI Tempo-Freqüência (CSI-TF) e CSI baseada em Matrizes de Hadamard (CSI-Hadamard) [13].

2.2.3.1 CSI Tempo-Freqüência (CSI-TF)

A CSI-TF possui características comuns às CSI-T e CSI-F, sendo a sua implementação realizada em duas etapas. Primeiramente, o sinal é dividido em blocos e cada bloco é dividido em N segmentos temporais. Estes segmentos são, então, submetidos à filtragem por um banco de filtros com M subfaixas. O resultado é representado na forma de uma matriz denominada $\boldsymbol{TF}$, cuja i-ésima coluna contém as M subfaixas do i-ésimo bloco. Cada linha da matriz corresponde a uma subfaixa dos N blocos.

Como pode ser visto na Figura 2.7, os elementos da matriz $\boldsymbol{TF}$ são ordenados pelo processo *First-In, First-Out* (FIFO) e permutados. Depois os elementos são reagrupados na forma matricial e cada segmento (coluna) é processado pelo banco de síntese, cujo resultado é o sinal cifrado pelo método de CSI-TF.

Analogamente aos sistemas de CSI-T, esta modalidade de criptofonia não é adequada ao propósito deste estudo, pois também introduz grandes atrasos, além de necessitar de um esquema de sincronismo elaborado.

Em decorrência da similaridade desta metodologia com aquelas explicitadas para os sistemas de CSI-T e CSI-F, não será apresentado o seu respectivo detalhamento matemático.

A Figura 2.8 apresenta os espectrogramas para um sistema CSI-TF. Comparado aos espectrogramas obtidos pelos métodos de CSI-F, o espectrograma obtido pelo método bidimensional CSI-TF denota uma distribuição de energia mais uniforme pelas sub-bandas. Este fato diminui a inteligibilidade residual, o que aumenta a resistência à criptoanálise.

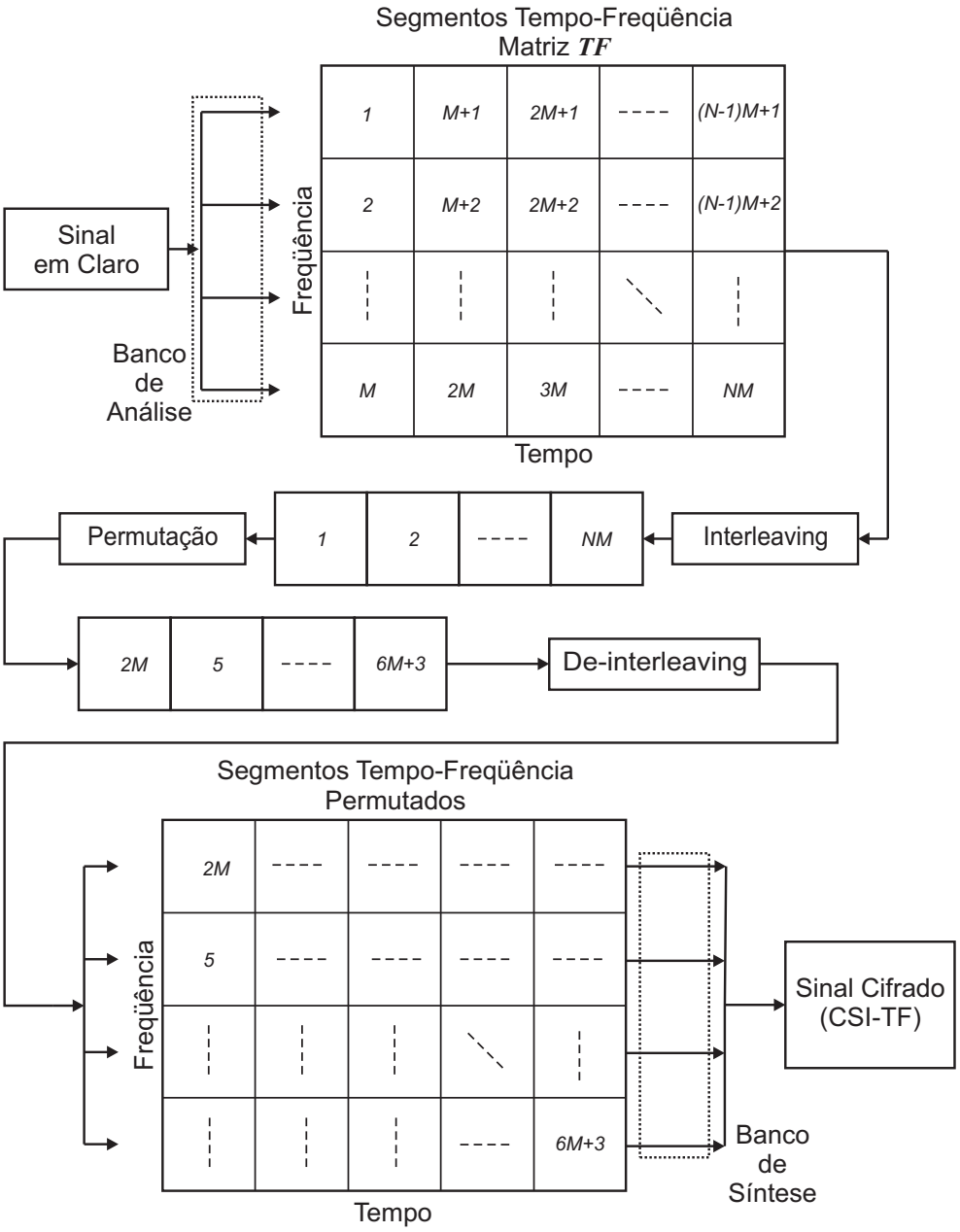

Figura 2.7: Diagrama de blocos exemplificando um sistema de CSI-TF. Neste diagrama, pode-se verificar que a filtragem é realizada por segmento, ao contrário dos sistemas de CSI-F, que realizam a filtragem por bloco.

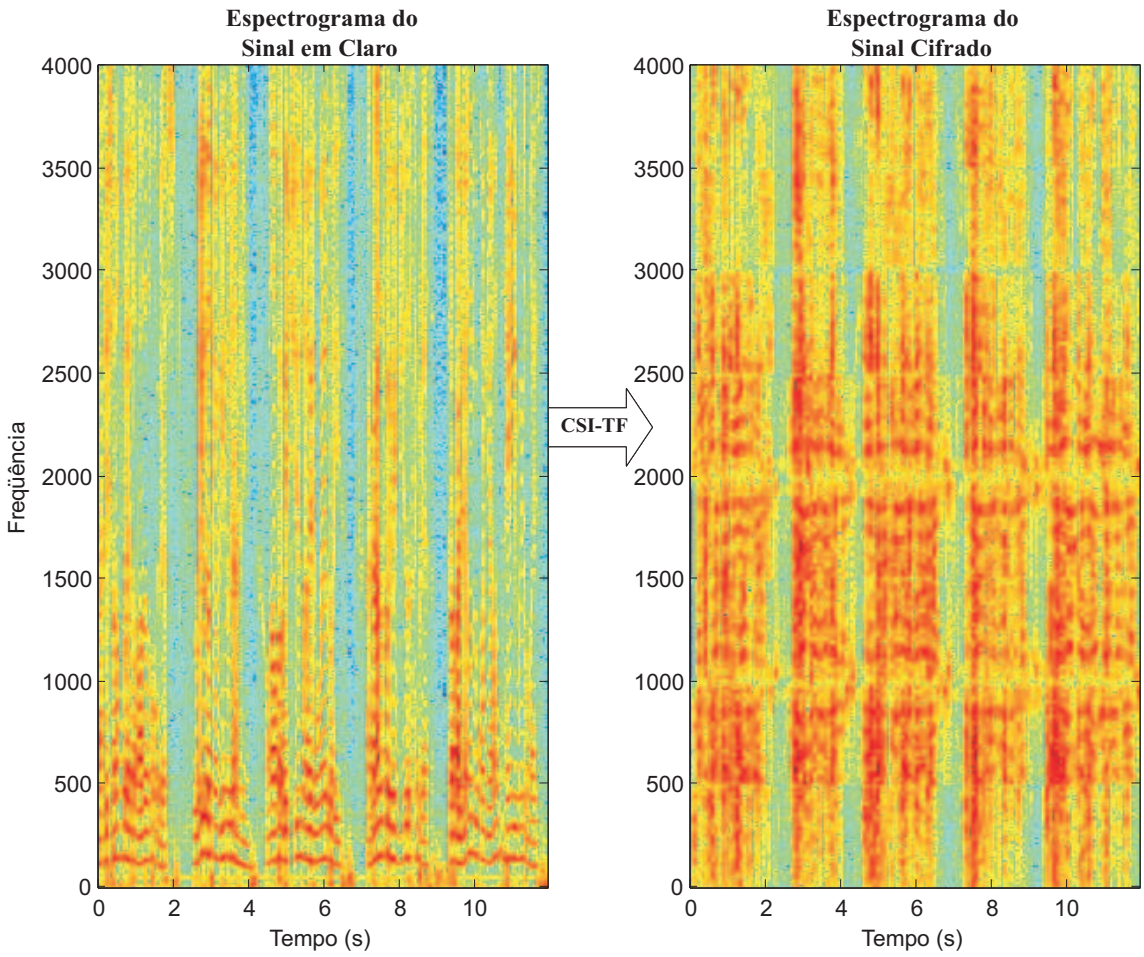

Figura 2.8: Espectrogramas de um sinal de voz e de sua versão cifrada obtida com CSI-TF. O número de segmentos tempo-freqüência utilizados foi $NM = 64$, sendo 8 segmentos (tempo) e 8 subfaixas (freqüência).

2.2.3.2 CSI baseada em Matrizes de Hadamard (CSI-Hadamard)

Existem duas possíveis aplicações para matrizes de Hadamard [12] no campo dos cifradores analógicos. A primeira aplicação é como sistema de CSI-F baseado em transformada ortogonal (ver Seção 2.2.2.2), que, em decorrência da pobre filtragem, resulta em *aliasing* inter-bandas após a permutação [6]. Esta não é, portanto, uma opção adequada e não será abordada neste trabalho. Por outro lado, a aplicação de matrizes de Hadamard na geração de matrizes de permutação introduz um conceito diferente dos demais já apresentados nesta seção. Esta metodologia se enquadra como bidimensional

pelo fato de alterar não somente a distribuição das amostras, como também as suas amplitudes, podendo ser empregada tanto no domínio do tempo como no domínio da freqüência.

Em contraste com os outros sistemas de CSI, que tem como característica comum a preservação das características essenciais do sinal original, na abordagem baseada em matrizes de Hadamard [9] cada amostra do segmento do sinal de voz cifrado é formada por meio de combinações lineares de todas as amostras pertencentes ao respectivo segmento.

Como conseqüência direta deste fato, observa-se:

a) Menor inteligibilidade residual;

b) Maior resistência à criptoanálise; e

c) Maior número de chaves (permutações).

Por definição, matrizes de Hadarmard possuem apenas elementos iguais a -1 e $+1$ e colunas e linhas ortogonais entre si. Desta forma, a inversa de uma matriz de Hadamard $\boldsymbol{H}$ pode ser calculada como:

$$\boldsymbol{H}^{-1} = \frac{1}{N}\boldsymbol{H}^{T}, \tag{2.23}$$

onde N é a ordem da matriz, e os valores de se seus elementos estão restritos ao conjunto formado pelos elementos 1, 2 ou $4n$, $\forall n \in \mathbb{Z}^{+}$. O fato de a matriz inversa de $\boldsymbol{H}$ ser obtida pela simples operação de transposição, contribui para a eficiência do processo e, desta maneira, não aumenta significativamente a complexidade computacional do processo de cifragem/decifragem.

Os procedimentos de cifragem/decifragem do sinal de voz são idênticos aos já apresentados para os outros sistemas de CSI, à exceção da matriz de permutação, que é definida como:

$$\boldsymbol{S} = \frac{1}{\sqrt{N}}\boldsymbol{P}\boldsymbol{H} \tag{2.24}$$

Matrizes de Hadamard podem ser formadas a partir de outras matrizes de Hadamard pela simples permutação de linhas ou colunas ou pela multiplicação de uma linha ou coluna por -1. As matrizes resultantes dessas operações não denominadas H-equivalentes. Estas matrizes contribuem para aumentar o número de permutações (chaves) possíveis. Segundo [9], o número de matrizes de Hadamard é dado pela desigualdade:

$$\mathcal{N}_H > 2^{2N-1}(N-1)! = 2^{2N-1}\{\frac{\mathcal{N}_P}{N}\}, \tag{2.25}$$

onde $\mathcal{N}_P = N!$, que corresponde ao número de permutações existentes para as demais CSI.

Resultante da aplicação da CSI-Hadamard, a Figura 2.12 evidencia com clareza que os segmentos cifrados não preservam a amplitude dos segmentos originais.

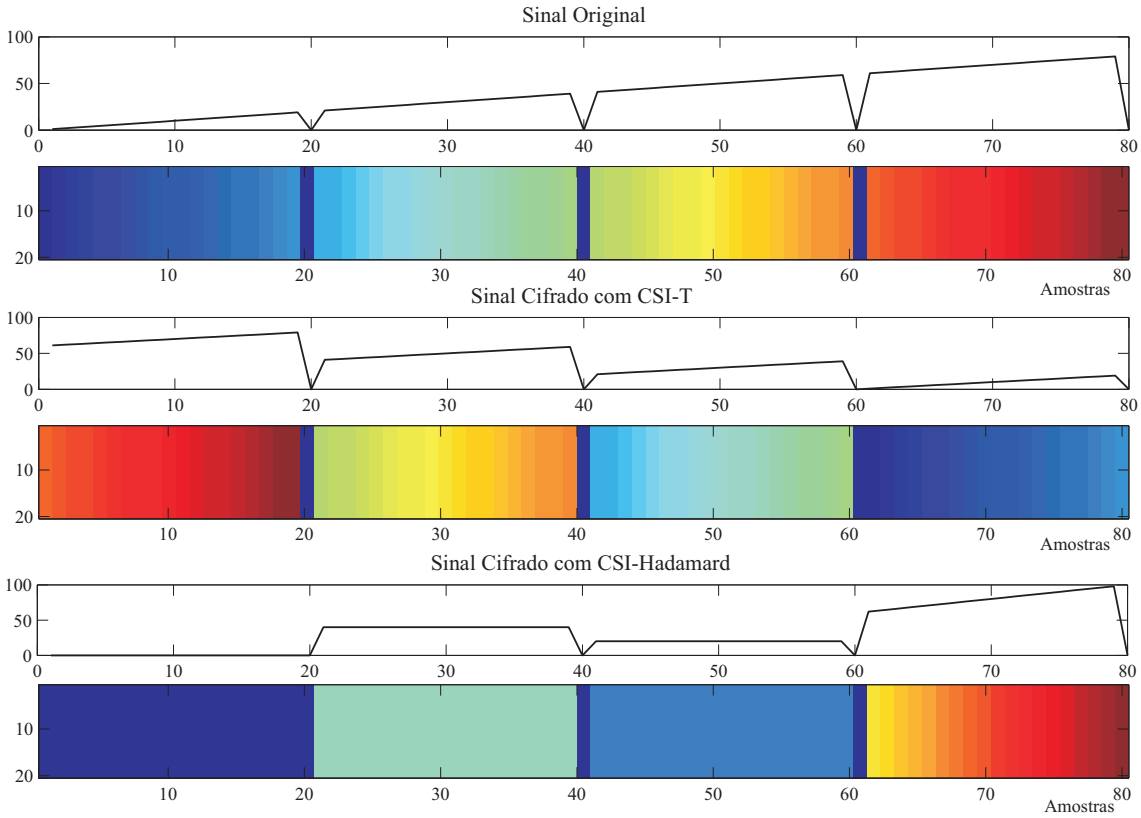

Figura 2.9: Exemplo simples (didático) de CSI-Hadamard de ordem $N = 4$.

Resumo das caracteríticas dos sistemas de CSI-Hadamard:

a) Conservação da banda do sinal;

b) Boa eficiência computacional;

c) Menor inteligibilidade residual e maior resistência à criptoanálise;

d) Maior número de chaves (permutações); e

e) Em razão da linearidade e da ortogonalidade do método, não há amplificação do ruído nem da distorção provocada pelo canal.

Podendo ser considerado um método bidimensional do tipo tempo-amplitude, a CSI-Hadamard não se demonstrou adequada à aplicação que é objeto deste trabalho. Dois fatores limitam a sua aplicação em sistemas com CODEC: a necessidade de um esquema preciso de sincronismo e as pequenas alterações de amplitude provocadas no sinal pelo CODEC. Na Figura 2.10, o espectrograma do sinal cifrado exibe um padrão bem diferente do espectrograma do sinal original, fato que decorre da bidimensionalidade do processo.

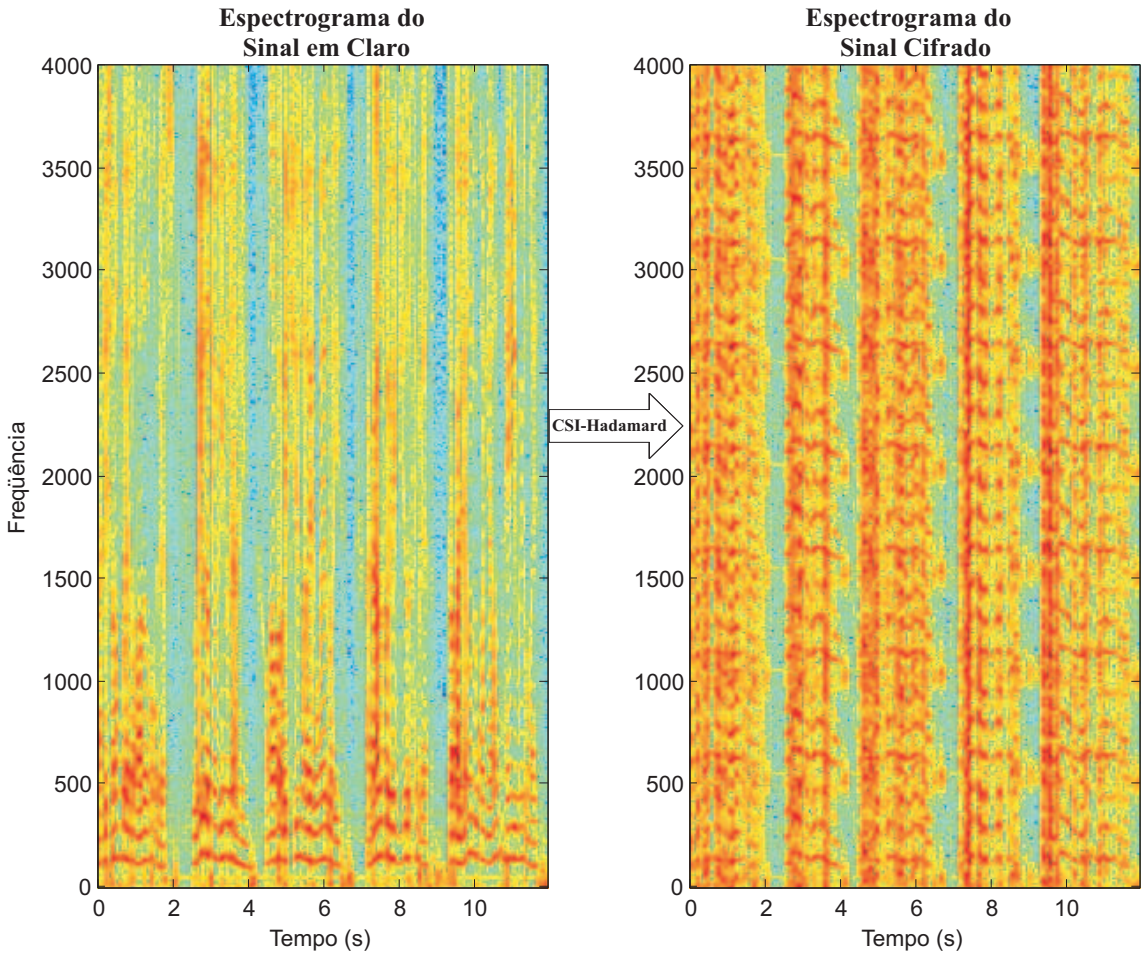

Figura 2.10: Espectrogramas de um sinal de voz e de sua versão cifrada no domínio da freqüência obtida com CSI-Hadamard ($N = 8$).

2.3 Criptofonia Digital

2.3.1 Criptofonia Bit a Bit (CBB)

Os sistemas CBB podem alcançar excelentes níveis de segurança, geralmente ao custo do aumento da banda de transmissão.

As implementações mais comuns de sistemas CBB fazem usos dos seguintes elementos:

a) Categoria I

- ❒ Codificador simples de voz;
- ❒ Módulos para encriptar e decriptar; e
- ❒ MODEM;

b) Categoria II

- ❒ Codificador simples de voz; e
- ❒ Módulos para encriptar e decriptar.

A CBB não apresenta inteligibilidade residual, pois para o ouvinte o sinal transmitido se assemelha a um ruído. A resistência à criptoanálise, portanto, depende apenas do algoritmo de criptografia empregado para encriptar os bits referentes à codificação do sinal.

Este tipo de sistema não cumpre o propósito do estudo aqui apresentado, pois não é passível de implementação sem modificações de monta no *hardware* do sistema de transcepção.

Na Figura 2.11 é apresentado o espectrograma de um sinal cifrado (encriptado) pelo método CBB, onde cada amostra do sinal foi submetida a uma criptografia simples pelo método "Ou-exclusivo" (XOR) com chave de 16 bits.

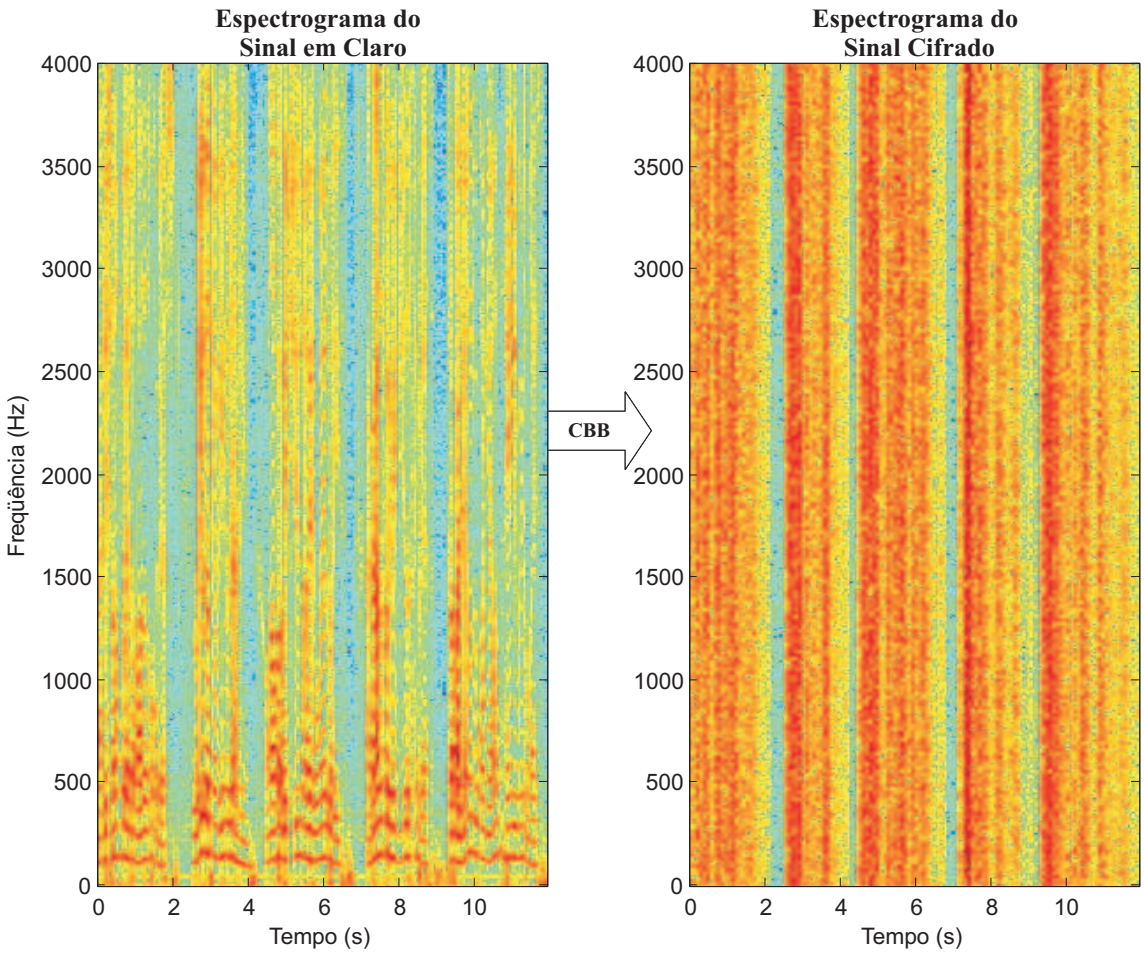

Figura 2.11: Espectrogramas de um sinal de voz e de sua versão cifrada obtida com CBB.

2.3.2 Criptofonia por Parâmetros Analíticos (CPA)

Os sistemas de CPA fazem uso de codificadores de voz [2] para tornar o sinal ininteligível. O processo de cifragem se dá pela encriptação e/ou manipulação dos parâmetros obtidos na codificação do sinal de voz, mais especificamente durante a análise. Na recepção, o processo de recuperação do sinal ocorre após se decriptar os parâmetros recebidos e realizar a síntese do sinal de voz.

A CPA se enquadra na classe de cifradores digitais com elevada segurança, podendo atingir até o nível estratégico dependendo dos requisitos de projeto.

Na prática, os sistemas CPA são implementados com VOCODER, pois possibili-

tam maior compressão do sinal em relação aos demais métodos. Valores típicos para as taxas de codificação variam de 2400 a 9600 bps. Esta caracterítica permite a utilização desses sistemas em canais de rádio HF, VHF, UHF etc.

Na transmissão, cada bloco tem seus parâmetros analíticos encriptados, codificados e transmitidos. No receptor, antes da reconstrução do sinal (síntese) os parâmetros são decriptados, e o sinal é sintetizado pelo processo inverso.

A utilização de CPA está limitada aos sistemas que permitem acessar os parâmetros de codificação do sinal antes das etapas de codificação de canal e modulação (transmissão), e depois da demodulação e decodificação na recepção.

A seguir, na Figura 2.12, é apresentado um protótipo de sistema de CPA implementado no SIMULINK. O codificador utilizado nesta aplicação é do tipo RELP (*Residual Excited Linear Prediction*) [2].

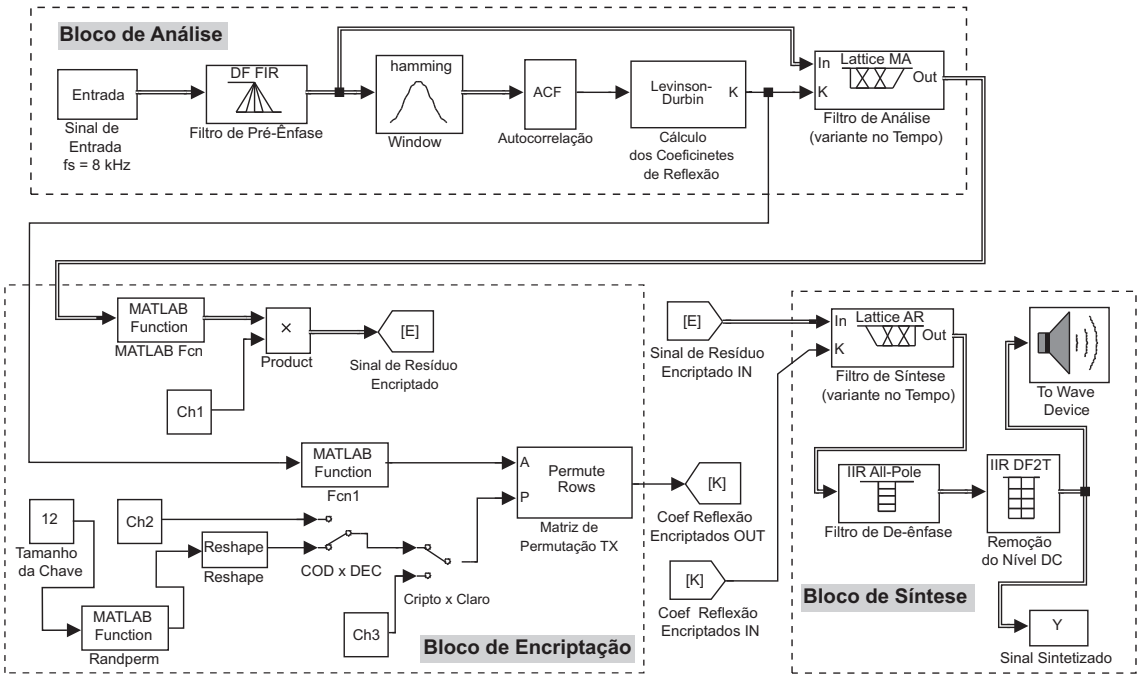

Figura 2.12: Protótipo simples de CPA utilizando codificador RELP.

A Figura Figura 2.13 apresenta o espectrograma do sinal cifrado por meio do método CPA. Conforme pode ser observado, a CPA produz um sinal cujo espectro não

apresenta "vestígios" dos formantes do sinal original, como já era esperado em razão da manipulação (encriptação) dos coeficientes de predição linear (*Linear Prediction Coefficients*-LPC) [2].

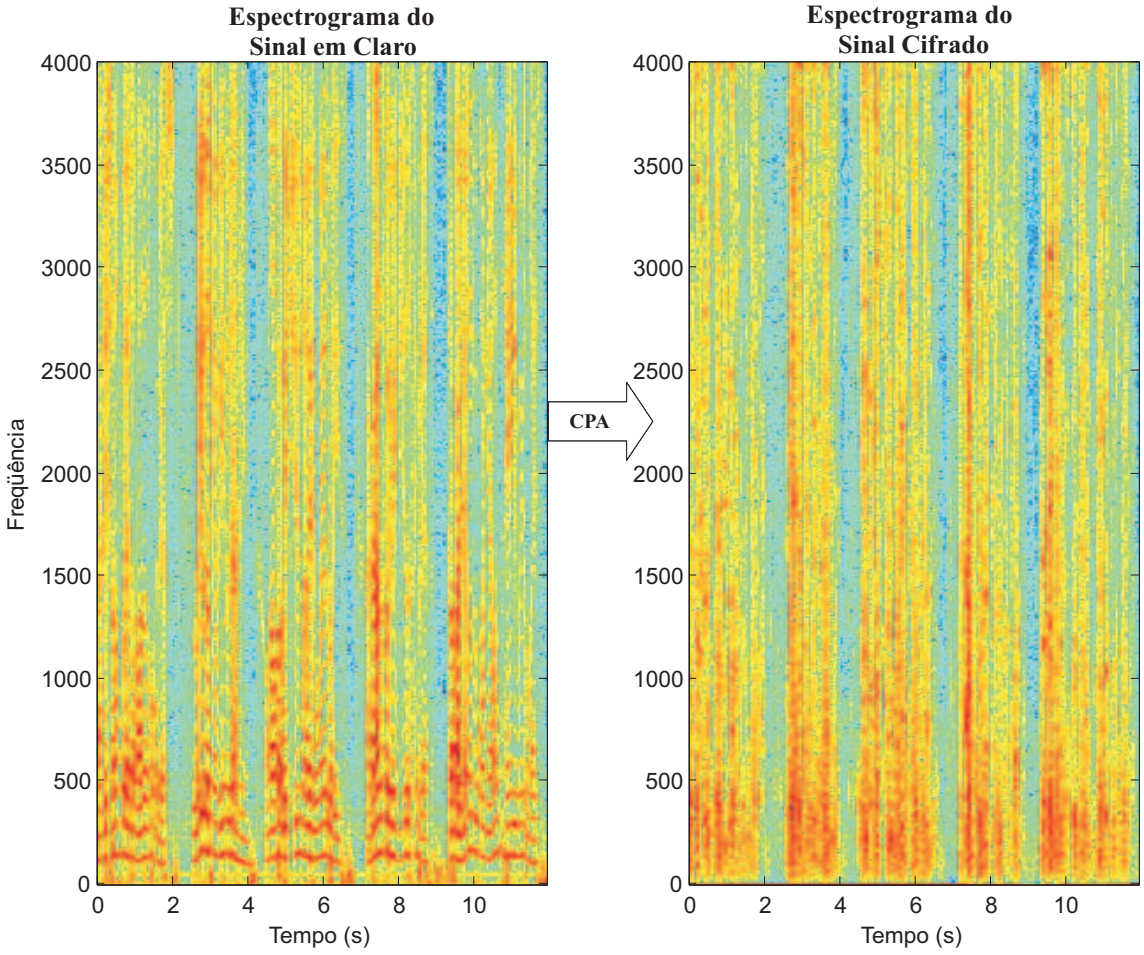

Figura 2.13: Espectrogramas de um sinal de voz e de sua versão cifrada com protótipo apresentado na Fig 2.8.

Este trabalho se limitará à apresentação de resultados objetivos de qualidade para sinais cifrados e decifrados pelo emprego de CSI-F em conjunto com CODEC AMR, conforme Capítulo 5.

2.4 Chaves para Criptofonia

A escolha das matrizes de permutação, também denominadas chaves para criptofonia, é uma importante etapa do processo de criptofonia. Dela, em grande parte, depende a dificuldade de se decifrar o sinal em um dado sistema de criptofonia.

Dentro do universo possível de $\mathcal{N}_P$ chaves, apenas um pequeno percentual produz sinais com inteligibilidade residual e resistência à criptoanálise [14] adequadas, o que torna a escolha de chaves uma tarefa crítica.

Para um conjunto $\boldsymbol{U}$ que contém todas as $\mathcal{N}_P = N!$ chaves, podem ser estabelecidos dois critérios para a escolha das chaves [15]:

- Critério I: Todas as chaves $\boldsymbol{P}_i$ pertencentes ao subconjunto $\boldsymbol{S}$ devem produzir baixa inteligibilidade residual; e
- Critério II: Para cada chave $\boldsymbol{P}_i \in \boldsymbol{S}$, deverá existir somente uma chave $\boldsymbol{P}_i^{-1} \in \boldsymbol{U}$ capaz de recuperar o sinal cifrado. Se outra chave for empregada no processo de decifragem, o sinal produzido deverá ser ininteligível.

O critério I está associado à inteligibilidade residual, enquanto a resistência à criptoanálise pode ser medida indiretamente pela aplicação do critério II.

Embora a inteligibilidade residual seja subjetiva e difícil de ser mensurada, pode-se definir a "distância" $D(\boldsymbol{P}_i, \boldsymbol{I})$ como medida indireta para inteligibilidade residual, onde $\boldsymbol{I}$ é a matriz identidade da mesma ordem de $\boldsymbol{P}_i$. O critério I pode ser reescrito como:

$$D(\boldsymbol{P}_i, \boldsymbol{I}) > \mathcal{L}_I \quad , \; \forall \; \boldsymbol{P}_i \in \boldsymbol{U} \tag{2.26}$$

O limiar $\mathcal{L}_I$ deve ser estimado de maneira a garantir um baixo valor para inteligibilidade residual.

De maneira semelhante, o segundo critério pode ser expresso:

$$D(\boldsymbol{P}_j^{-1}, \boldsymbol{P}_i) > \mathcal{L}_{II} \quad , \; \forall \; \{\boldsymbol{P}_i, \boldsymbol{P}_j\} \in \boldsymbol{S} \;\; \text{e} \;\; i \neq j \tag{2.27}$$

A utilização da *Distância de Hamming* (DH) para o cálculo de $D(\boldsymbol{P}_i, \boldsymbol{I})$ [14], [15] tem como resultado o número de elementos que são movidos das suas posições originais

após a permutação. Quanto maior o resultado obtido para DH, menor a inteligibilidade residual. O limiar proposto em [16] é de 90%.

Neste trabalho é proposta uma nova abordagem que considera a permutação como sendo uma rotação de eixos dos espaços vetoriais $\mathbb{R}^N \rightarrow \mathbb{R}^N$, onde N é o tamanho da chave. Pode-se, então, considerar a matriz de permutação $\boldsymbol{P_i}$ resultante da rotação da matriz identidade $\boldsymbol{I}$ segundo a chave de permutação $\boldsymbol{V}_N^{P_i}$, cuja definição será explicitada *a posteriori*. Desta forma, a "distância" $D(\boldsymbol{P}_i, \boldsymbol{I})$ pode ser calculada como sendo a rotação entre $\boldsymbol{P}_i$ e $\boldsymbol{I}$.

Os valores obtidos para o ângulo $D(\boldsymbol{P}_i, \boldsymbol{I})$, por definição, são diretamente proporcionais à rotação da matriz de permutação P_i; portanto, quanto maior o limiar $\mathcal{L}_I$, maior o percentual de chaves pertencentes a $\boldsymbol{S}$ capazes de transladar os segmentos dentro do bloco para a metade oposta em relação ao segmento central. Visto que existem chaves diferentes com valores de Φ_I idênticos, a metodologia apresentada deve se restringir à seleção de conjuntos de chaves, pois o seu resultado é válido somente como medida indireta da inteligibilidade residual média do conjunto de chaves.

Seja $\boldsymbol{V}_N = [1 \;\; 2 \cdots N]^T_{(N\times 1)}$, então o ângulo de rotação de $\boldsymbol{P}_i$ em relação a $\boldsymbol{I}$ pode ser definido como:

$$D(\boldsymbol{P}_i, \boldsymbol{I}) = \Phi_I = \arccos\left\{\frac{(\boldsymbol{P}_i\boldsymbol{V}_N)^T \cdot \boldsymbol{V}_N}{\|\boldsymbol{V}_N\|^2}\right\} \quad , \quad \forall \quad \boldsymbol{P}_i \in \boldsymbol{U} \tag{2.28}$$

$$D(\boldsymbol{P}_j^{-1}, \boldsymbol{P}_i) = \Phi_{II} = \arccos\left\{\frac{(\boldsymbol{P}_j^{-1}\boldsymbol{V}_N)^T \cdot (\boldsymbol{P}_i\boldsymbol{V}_N)}{\|\boldsymbol{V}_N\|^2}\right\} \quad , \quad \forall \;\; \{\boldsymbol{P}_i, \boldsymbol{P}_j\} \in \boldsymbol{S} \;\; \text{e} \;\; i \neq j. \tag{2.29}$$

O cálculo efetuado pela Equação (2.28) é um cálculo indireto para a rotação provocada sobre a matriz $\boldsymbol{P}_i$. O valor calculado expressa numericamente o ângulo entre os vetores $\boldsymbol{V}_N$ e $\boldsymbol{V}_N^{P_i} = \boldsymbol{P}_i\boldsymbol{V}_N$, onde $\boldsymbol{V}_N^{P_i}$ é o vetor permutado segundo a matriz $\boldsymbol{P}_i$ e corresponde à chave de permutação.

Para cada tamanho de chave N, existe um valor máximo Φ_I^{Max} decorrente da aplicação da matriz de permutação $\boldsymbol{P}^{\text{Max}}$, cujos elementos estão dispostos na diagonal secundária (ver Apêndice B).

$$\Phi_I^{\text{Max}}(N) = \arccos\left\{\frac{N+2}{2N+1}\right\} \tag{2.30}$$

Se o limiar $\mathcal{L}_{\mathcal{I}}$ for escolhido suficientemente grande, as chaves resultantes permutarão a maioria dos segmentos pertencentes a um bloco para a metade oposta à sua posição original. Com base nos resultados apresentados pela Figura 2.14, foi observado que $\mathcal{L}_{\mathcal{I}} = 0,85\Phi_I^{\mathrm{Max}}(N)$ representa um valor adequado.

Conforme dados constantes da Figura 2.15, pode-se fazer uma correlação entre os valores do limiar $\mathcal{L}_{\mathcal{I}}$ e da DH [15], isto é, valores grandes de $\mathcal{L}_{\mathcal{I}}$ implicam valores grandes de DH média (Figura 2.17) para o conjunto de chaves considerado e, portanto, uma baixa inteligibilidade residual para este conjunto.

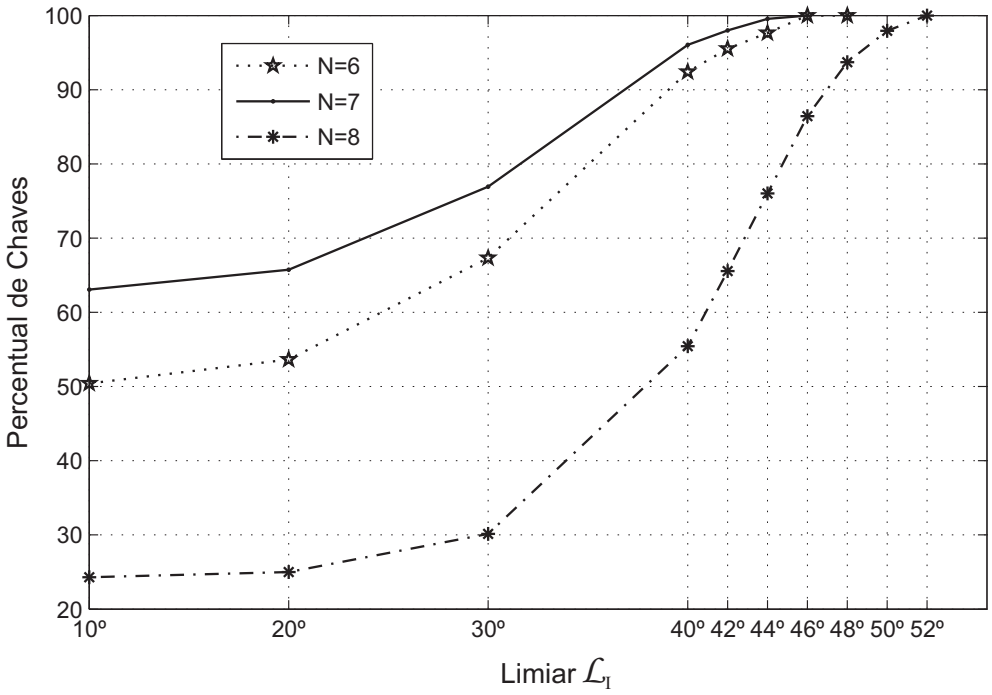

Figura 2.14: Percentual de chaves capazes de permutar pelo menos um segmento para metade oposta do bloco.

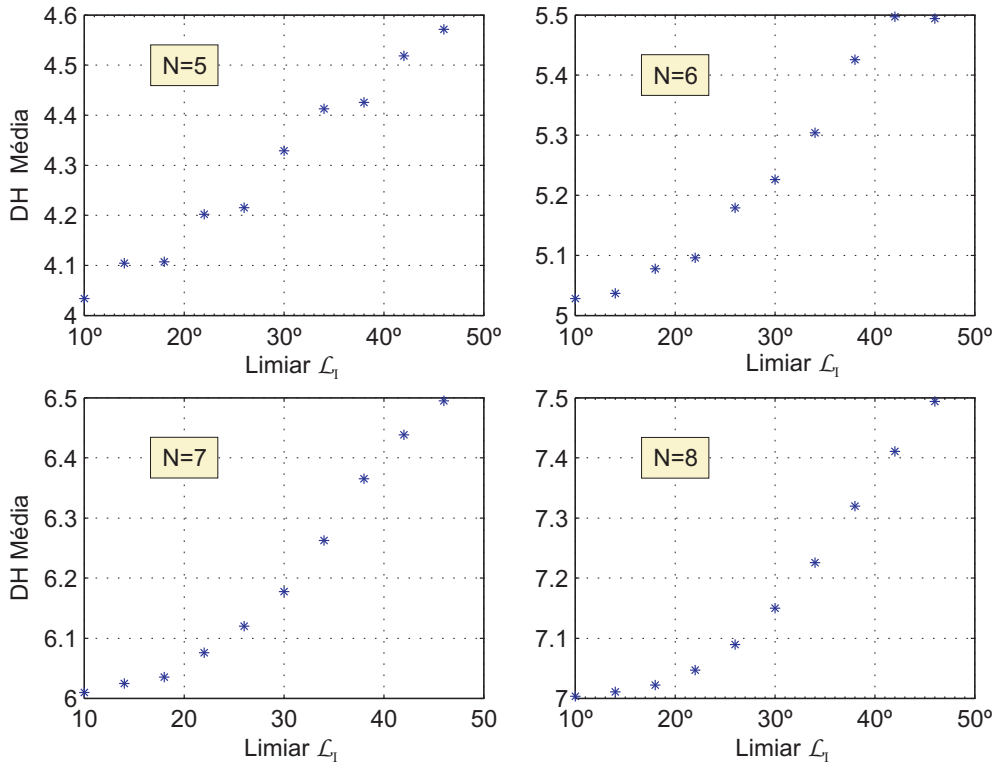

Figura 2.15: DH média versus limiar $\mathcal{L}_{\mathcal{I}}$.

Tabela 2.1: Número de chaves que atendem ao critério I ($4 \leq N \leq 10$)

N	**Nº Total de Chaves**	$\boldsymbol{\Phi_I^{\mathrm{Max}}(N)}$	$\boldsymbol{\mathcal{L}_{\mathcal{I}}}$	**Nº de Chaves para** $\boldsymbol{\Phi_I \geq \mathcal{L}_{\mathcal{I}}}$
4	24	48,19º	40,96º	5 (20,83%)
5	120	50,49º	42,91º	27 (22,50%)
6	720	52,02º	44,22º	128 (17,78%)
7	5.040	53,13º	45,16º	672 (13,33%)
8	40.320	53,97º	45,87º	4.900 (12,15%)
9	362.880	54,62º	46,43º	35.163 (9,69%)
10	3.628.800	55,15º	46,89º	301.704 (8,31%)

De acordo com o conteúdo da Tabela 2.1, é possível verificar que valores de N menores que 8 não devem ser empregados em sistemas de criptofonia, em decorrência do deduzido número de chaves com baixa inteligibilidade residual. Da mesma forma,

quando se utiliza CSI-F em conjunto com sistemas que fazem uso de CODEC, deve-se evitar valores muito elevados de N, pois a permutação de um grande número de subfaixas pode produzir um sinal cujas características espectrais divirjam das características espectrais de um sinal de voz, comprometendo, desta forma, os processos de codificação e decodificação realizados pelo CODEC AMR.

Uma maneira de aumentar a segurança dos sistemas de CSI sem ter que aumentar demasiadamente o valor de N é utilizar chaves cujos valores são modificados periodicamente. Este tipo de implementação pressupõe a existência de mecanismos de sincronismo precisos para auxiliar a troca de chaves simultaneamente no transmissor e receptor.

O critério II é bem mais restritivo que o I, pois seleciona as chaves dentro do subconjunto $\boldsymbol{S}$, o que implica automático atendimento ao critério I. A utilização deste critério deve ser avaliada com parcimônia, pois diminui ainda mais o número de chaves disponíveis. Para ilustrar a redução no número de chaves com baixa inteligibilidade residual que a adoção do critério II traria, na Tabela 2.2 são apresentados os números de chaves segundo o critério $\mathcal{L}_{\mathcal{II}} = 0,5\Phi_{II}^{\mathrm{Max}}(N)$.

Tabela 2.2: Número de chaves que atendem ao critério II ($4 \leq N \leq 8$)

N	**Nº Total de Chaves - Critério I**	$\boldsymbol{\Phi_{II}^{\mathrm{Max}}(N)}$	$\boldsymbol{\mathcal{L}_{\mathcal{II}}}$	**Nº de Chaves para $\Phi_{II} \geq \mathcal{L}_{\mathcal{II}}$**
4	5	33,56º	40,96º	4 (16,67%)
5	27	43,34º	42,91º	21 (17,50%)
6	128	44,42º	44,22º	90 (12,56%)
7	672	45,57º	45,16º	486 (9,64%)
8	4.900	46,66º	45,87º	3.788 (9.39%)

Na Figura 2.16 são apresentadas as doze matrizes de permutação com maior inteligibilidade residual para o atendimento do critério I ($\mathcal{L}_{\mathcal{I}} = 0,85\Phi_{I}^{\mathrm{Max}}(N)$, para $N = 8$) e que correspondem às chaves com maior susceptibilidade à criptoanálise. Como pode ser observado neste conjunto de chaves, as piores chaves são [8 7 3 4 1 6 2 5] e

[8 7 3 4 2 5 1 6], cujos valores de DH são 5 e 6, respectivamente. Ambas as chaves permutam 4 segmentos para a metade oposta do bloco.

Na Figura 2.17 são apresentadas as doze matrizes de permutação com menor inteligibilidade residual para o atendimento do critério I ($\mathcal{L}_{\mathcal{I}} = 0,85\Phi_I^{\text{Max}}(N)$, para $N = 8$). Este conjunto de chaves correspondem às chaves com maior resistância à criptoanálise. Como pode ser observado neste conjunto de chaves, as melhores chaves são [8 7 6 5 4 3 2 1] e [7 8 6 5 4 3 2 1], cujos valores de DH valem 8. Ambas chaves permutam 4 segmentos para a metade oposta do bloco. Quando da utilização de chaves fixas, as chaves correspondentes à matriz de permutação $\boldsymbol{P}^{\text{Max}}$ devem ser evitadas, pois, embora possuam baixa inteligibilidade residual, são testadas compulsoriamente pelos processos de criptoanálise.

A metodologia apresentada nesta seção não tem a pretensão de esgotar a problemática da escolha de chaves de criptofonia para cifradores analógicos; pelo contrário, constitui um método complementar ao apresentado pela referência [15] e visa, tão somente, a possibilitar a seleção preliminar de chaves dentro do conjunto N. Uma metodologia mais abrangente e completa para a solução deste problema é descrita em [17], onde é apresentado um método objetivo para quantificação da inteligibilidade residual.

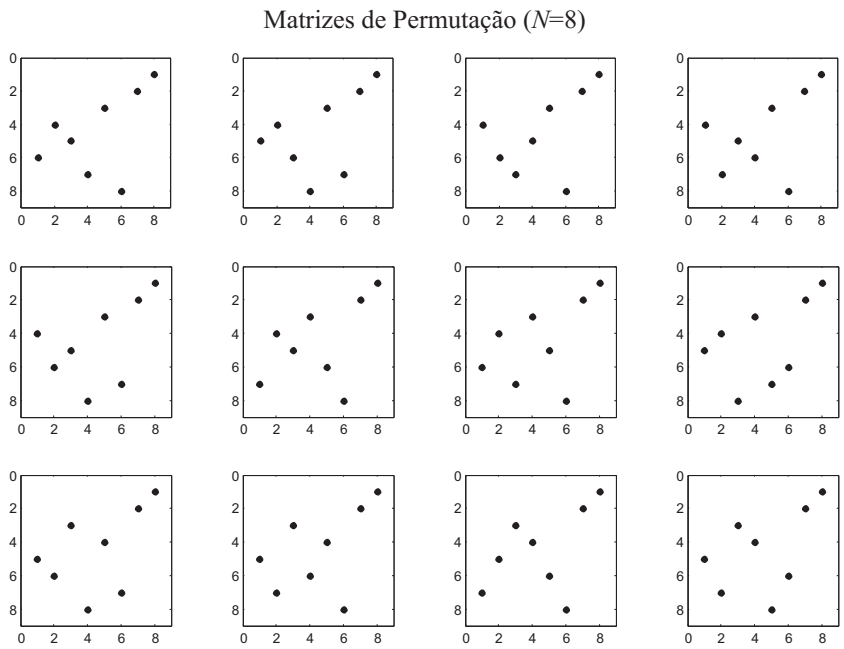

Figura 2.16: As doze matrizes de permutação ($N = 8$) com maiores valores de inteligibilidade residual dentre as chaves que atendem ao critério I.

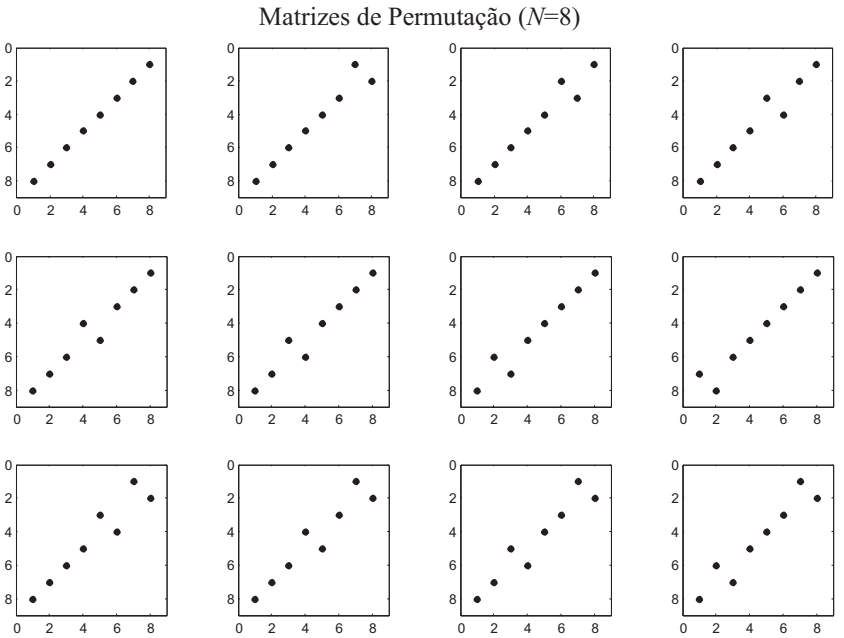

Figura 2.17: As doze matrizes de permutação ($N = 8$) com menores valores de inteligibilidade residual dentre as chaves que atendem ao critério I.

Capítulo 3

Sincronismo em Sistemas de Criptofonia

3.1 Introdução

Em decorrência dos efeitos introduzidos pelo canal de comunicações, para decifrar o sinal de maneira correta faz-se necessário o emprego de mecanismos de sincronismo de bit e de quadro. O sincronismo pressupõe que o sinal de *clock* no receptor possua a mesma fase e freqüência do *clock* utilizado na geração do sinal. Em sistemas reais, o sinal experimenta efeitos causados pelo canal, tais como variações na freqüência e na fase da seqüência de bits transmitida. A adoção de esquemas de sincronismo em sistemas de criptofonia permite ainda alterar periodicamente a chave utilizada para cifrar o sinal, diminuindo, desta forma, a inteligibilidade residual do sinal cifrado (ver Capítulo 2).

A ausência do sincronismo em sistemas de comunicações móveis que fazem uso de criptofonia pode ser solucionada com o emprego de técnicas de sincronismo de bit (amostras) e sincronismo de quadro, conforme detalhamento constante deste capítulo.

3.2 Sincronismo de Bit (Amostra)

Seja S_b um sinal digital formado por uma seqüência de bits tal que $S_b = \{a_k\}_{k=0}^{N-1}$, onde a_k pode assumir os valores discretos -1 e $+1$. Este sinal, após ser transmitido, é contaminado por ruído e sofre atrasos inerentes ao canal de transmissão, podendo ser expresso como:

$$r(t) = M(t) \sum_{k=0}^{N-1} s(t; a_k; \varepsilon) + \eta(t). \tag{3.1}$$

Na Equação (3.1), o fator $M(t)$ é responsável pela distorção de amplitude causada pelo canal e $s(t; a_k; \varepsilon)$ representa a informação após incorporar a forma do pulso adequado ao canal de transmissão. O ruído adicionado pelo canal é do tipo Aditivo Gaussiano Branco (AWGN), e representado aqui por $\eta(t)$.

Desprezando-se as distorções de amplitude provocadas pelo canal, pode-se expressar o sinal transmitido como:

$$x(t) = \sum_{k=0}^{N-1} a_k g(t; \varepsilon), \tag{3.2}$$

onde $g(t; \varepsilon)$ é um pulso cuja forma de onda é escolhida com base nas características do canal. Esta escolha deve ser realizada de maneira a garantir a minimização de erros e interferências inter-simbólicas. Na Equação (3.2), a variável ε representa os atrasos provocados pelo canal.

Considerando os sinais $x(t)$ e $y(t)$ como amostras dos processos estacionários de segunda ordem $X(t)$ e $Y(t)$, pode-se demonstrar [18] que a correlação cruzada entre os dois sinais é função apenas da diferença dos instantes de observação. Se a ergodicidade [18] for satisfeita conjuntamente para os processos $X(t)$ e $Y(t)$, as médias estatísticas se tornam iguais às médias temporais correspondentes e, portanto, a correlação cruzada dos processos $X(t)$ e $Y(t)$ se confunde com a correlação temporal cruzada das amostras $x(t)$ e $y(t)$:

$$R_{xy}(t_1, t_2) = E[X(t_1)Y(t_2)] = A[x(t-\tau)y(t)] \tag{3.3}$$

onde $E[\,\cdot\,]$ e $A[\,\cdot\,]$ são as médias estatística e temporal, respectivamente, e $\tau = t_2 - t_1$.

$$R_{xy}(\tau) = \int_{k=0}^{kT} x(t-\tau)y(t)dt \quad , \text{ com } T = \frac{1}{N} \tag{3.4}$$

Para se obter o sincronismo de amostras entre os sinais $x(t)$ e $y(t)$, deve-se calcular o valor máximo de $R_{xy}(\tau)$ e, então, determinar o atraso correspondente a $\tau = \tau_m$.

Para fim de exemplo, sejam $x(t)$ e $y(t)$ duas seqüências binárias apresentadas na Figura 3.1, onde o sinal $y(t)$ é uma cópia de $x(t)$ atrasada de $\tau = \tau_m$.

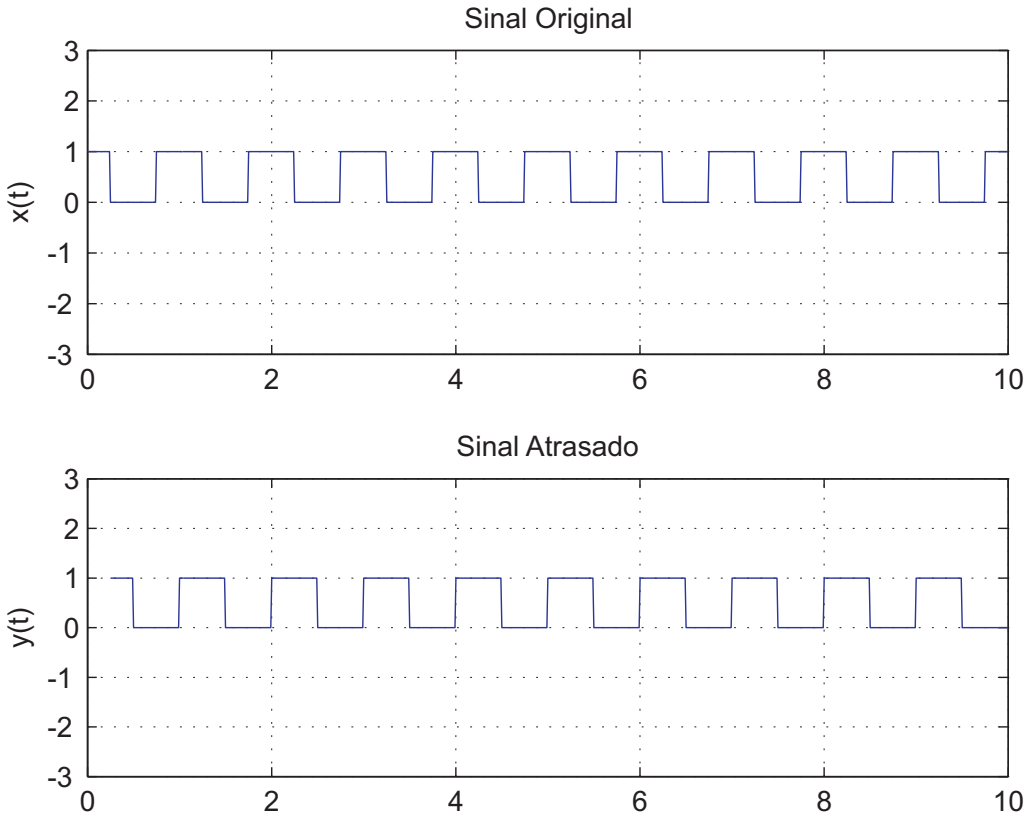

Figura 3.1: Defasagem entre os sinais transmitido e o recebido.

A correlação cruzada entre $x(t)$ e $y(t)$ pode ser calculada por meio da Equação (3.4).

$$R_{xy}(\tau) = \frac{[T - \tau]}{T} \tag{3.5}$$

$$\tau_m = T[1 - R_{xy}(\tau_m)] \tag{3.6}$$

Aplicando-se a Equação (3.6) às formas de onda constantes da Figura 3.1, obtém-se um valor máximo para $R_{xy}(\tau)$ igual a 0,75, o que corresponde a um atraso τ_m =0,25 (T =1,0).

3.3 Sincronismo de Quadro

A metodologia aqui apresentada para alcançar o sincronismo de quadro emprega seqüências predefinidas e com propriedades estatísticas específicas. Estas seqüências, também denominadas *Palavras de Sincronismo* (PS), são periodicamente inseridas no

sinal transmitido. As Palavras de Sincronismo podem ser inseridas no início e/ou final de cada quadro, a cada M quadros ou, ainda, somente no início de cada transmissão. O comprimento, a duração e o número de repetições da PS são estimados com base nos parâmetros abaixo listados:

- Precisão de sincronismo requerida, expressa em número mínimo de amostras que podem ser sincronizados pelo processo;
- Taxa máxima de transmissão disponível, de maneira a não exceder a taxa de Nyquist; e
- Atraso máximo estimado para o sistema.

Seja S uma seqüência contendo N amostras, cujo prévio conhecimento permite a sua exata identificação durante o processo de recepção do sinal, tal que:

$$S = [s_1\ s_2\ s_3 \cdots s_N]^T. \tag{3.7}$$

Esta seqüência, empregada como PS, pode ser interpretada como um conjunto finito de k símbolos (amostras), que, no limite ($k = N - 1$), possuem correlação $R_{SS}(\tau) = \frac{1}{N}$. A função de correlação cruzada desta seqüência S com ela mesma atrasada de k amostras pode ser expressa como:

$$R_{SS}(k) = \sum_{i=1}^{N-k} S_i S_{i+k}^* \quad , \quad k = 0, 1, \cdots, (N-1). \tag{3.8}$$

Se o sinal no receptor $r(t)$ contém o mesmo padrão de informação presente na PS, a menos do atraso e da distorção provocada pelo canal e outros subsistemas (VOCODER etc), sem perda de generalidade, a Equação (3.8) pode ser reescrita de forma a propiciar o cálculo da correlação cruzada entre $r(t)$ e S:

$$R_{rS}(k) = \sum_{i=1}^{N-k} r(i) S^*(i+k) \quad , \quad k = 0, \cdots, (N-1). \tag{3.9}$$

O cálculo da correlação, exclusivamente, não conduz a bons resultados para o processo de sincronismo de quadros [19], [20]; desta forma, há necessidade de se levar em consideração as transições aleatórias que ocorrem na fronteira entre a PS e o sinal $r(t)$.

Outro fator a ser considerado é a grande faixa dinâmica do sinal em relação à PS. Para minimizar estes efeitos, pode-se normalizar o valor da correlação cruzada calculada na Equação (3.9), conforme proposto em [19], [20]:

$$M_{\text{Norm}} = \frac{|R_{rS}(k)|}{\sqrt{\sum_{j=0}^{k} |r(k-j)|^2}} \quad , \quad k = 0, \cdots, (N-1). \tag{3.10}$$

Uma importante escolha para garantir o sincronismo de quadro se faz pela seleção adequada da PS, cujas características estatísticas devem ser cuidadosamente estudadas [21]. O requisito principal para que uma seqüência possa ser empregada como PS é possuir uma baixa autocorrelação aperiódica[1]. A seguir são citadas algumas propriedades desejáveis para seqüências candidatas a PS:

- Devem assumir apenas dois valores discretos {0 ou 1, −1 ou 1};
- Devem ter uma função de autocorrelação com um único pico estreito, para ajudar na sincronização do código;
- Devem ter funções de correlação cruzada com valores baixos; e
- Devem ser balanceadas (equilibradas) entre 0 e 1 {ou −1 e 1}, para que a densidade espectral de potência esteja bem distribuída pelas bandas de freqüência utilizadas.

Dentre os códigos ou seqüências que atendem, parcial ou totalmente, às propriedades supracitadas, podem ser mencionados os seguintes:

- Walsh-Hadamard;

[1]A função de correlação aperiódica entre duas seqüências pseudo-aleatórias demonstra o grau de correlação entre elas para um intervalo de tempo considerado menor que o período das seqüências, conforme a definição:

$$R_{S_a S_b}(k) = \begin{cases} \sum_{i=0}^{N-1-k} S_a(i) S_b(i+k) & , \quad 0 \le k \le N-1 \\ \sum_{i=0}^{N-1+k} S_a(i-k) S_b(i) & , \quad 1-N \le k < 0 \\ 0 & , \quad |k| \ge 0. \end{cases}$$

- Barker [21];
- Neuman-Hofman [22];
- Seqüências PN de comprimento máximo;
- Códigos de Gold; e
- Códigos de Kasami.

Dos códigos supracitados, apenas os códigos de Walsh-Hadamard são classificados como códigos ou seqüências ortogonais; os demais são códigos/seqüências não-ortogonais. Em razão da simplicidade e das propriedades referentes à autocorrelação aperiódica, este trabalho adotará o emprego das seqüências de Barker.

3.3.1 Seqüências de Barker

Uma seqüência de Barker é uma seqüência de N valores $S_i = \pm 1, \forall\, i = 1, 2, \cdots, N$, tal que $|\sum_{i=1}^{N-k} S_i S_{i+k}^*| \leq 1, \forall\, \{1 \leq k \leq N-1\}$. Tal definição é equivalente a dizer que a autocorrelação aperiódica de S satisfaz a desigualdade $R_{SS}(k) \leq 1$. São conhecidas somente as seguintes seqüências de Barker:

Tabela 3.1: Codificação de Barker para Sincronismo de Quadros

N	**Seqüência Codificada**	**Correlação Aperiódica**
1	$\{+1\}$	{1}
2	$\{+1, +1\}$ ou $\{+1, -1\}$	{2, 1} ou {2, -1}
3	$\{+1, +1, -1\}$	{3, 0, -1}
4	$\{+1, +1, +1, -1\}$ ou $\{+1, +1, -1, +1\}$	{4, 1, 0, 1} ou {4, -1, 0, 1}
5	$\{+1, +1, +1, -1, +1\}$	{5, 0, 1, 0, 1}
7	$\{+1, +1, +1, -1, -1, +1, -1\}$	{7, 0, -1, 0, -1, 0, -1}
11	$\{+1, +1, +1, -1, -1, -1, +1, -1, -1, +1, -1\}$	{11, 0, -1, 0, -1, 0, -1, 0, -1, 0, -1}
13	$\{+1, +1, +1, +1, +1, -1, -1, +1, +1, -1, +1, -1, +1\}$	{13, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1}

3.4 Modulação da Palavra de Sincronismo

A inserção da Palavra de Sincronismo (PS) como preâmbulo dos quadros do sinal de áudio não pode ser realizada sem que a sua forma de onda tenha sido alterada para um formato adequado às necessidades do sub-sistema de banda-base, que no caso aqui analisado é o CODEC Adaptive Multi-Rate (AMR) [4]. Uma maneira prática de se adequar a forma de onda da PS à transmissão via CODEC AMR é o emprego de uma modulação digital como *Frequency-shift keying* (FSK) [23]. A modulação escolhida para aplicação aqui estudada foi a *Audio Frequency-shift keying* (*A*FSK), que difere da FSK somente pelo fato de o processo de modulação ocorrer na banda-base do sinal (áudio-freqüência).

Na modulação *A*FSK a informação contida no sinal digital é representada por meio de mudanças de *pitch* de um sinal senoidal cujas freqüências pertencem à faixa de áudio. O sinal modulado resultante possui características espectrais adequadas à transmissão via rádio, telefones e outros sistemas, cujos pontos de acesso são canais de voz. Neste texto, as modulações *A*FSK e FSK serão tratadas indistintamente.

A modulação *A*FSK atribui freqüências diferentes para a portadora, dependendo do valor do símbolo que é transmitido . Conforme exemplificado pela Figura 3.2, quando um símbolo **0** é transmitido, a portadora assume a freqüência correspondente f_0. Quando um símbolo **1** é transmitido, a freqüência da portadora assume a freqüência correspondente f_1.

Pode-se utilizar um número maior de freqüências de transmissão, cada uma correspondendo a um símbolo. Este modo é chamado de M-FSK, onde M representa o número de símbolos empregados. A utilização da modulação M-FSK aumenta a taxa de símbolos transmitidos; em contrapartida, necessita de uma maior banda de transmissão.

O emprego da modulação *A*FSK se limita às aplicações de baixa velocidade, e sua eficiência, em termos de potência e banda, é pequena em relação a outras modulações digitais. Contudo, devido à sua simplicidade, muitas são as suas aplicações nos campos das comunicações via rádio, telefonia, transmissão de música e voz via rede de alimentação etc. Em complemento, o Apêndice C apresenta detalhes sobre os processos

de análise, geração e detecção de sinais FSK.

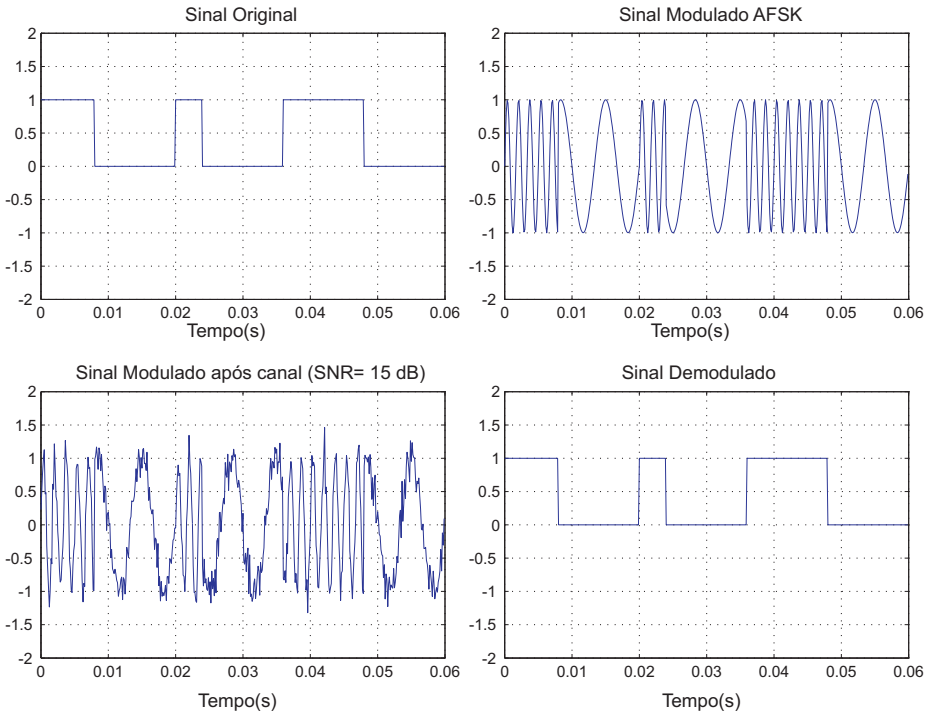

Figura 3.2: Formas de onda da modulação FSK.

3.5 Efeitos dos CODEC AMR/GSM Sobre o Sincronismo de Quadro

Após codificação e decodificação da seqüência de sincronismo pelos CODEC AMR e GSM [4], [24], foi observado que as amostras iniciais e finais da seqüência sofreram distorção em nível suficiente para impedir a demodulação pelo detector FSK (Figura 3.3). Este fato provoca uma perda sincronismo de quadros, principalmente em decorrência da natureza variável do número de amostras que são afetadas.

Nesta seção serão apresentados resultados das distorções para as diversas taxas do CODEC AMR, o que permite fazer uma estimativa conservadora para o número de amostras afetadas. A partir do valor estimado para o número de amostras distorcidas, é possível formular uma solução para que o sincronismo seja alcançado.

Em decorrência dos resultados semelhantes obtidos para os CODEC AMR (taxa de 12,2 kbps) e GSM (Full Rate), não serão apresentados resultados referentes ao CODEC GSM.

Para se estimar o número de amostras que são distorcidas no processo de codificação/decodificação, devem ser considerados os atrasos médios do CODEC AMR para as diversas taxas, que são de 55 amostras para 4,75 kbps e de 40 amostras para as demais taxas.

A realização da estimativa do número de amostras afetadas destrutivamente pelo CODEC foi realizada experimentalmente, conforme resultados apresentados pelas Figuras 3.4 a 3.9, utilizando-se um conjunto composto por vinte seqüências de Barker de comprimento $N = 11$, totalizando uma PS com 220 amostras. Esta estimativa foi realizada com o emprego do detector ótimo FSK descrito no Apêndice C.

Devido a o número de amostras afetadas pelo CODEC não ser constante, não é possível realizar o sincronismo de quadro sem que se realize uma pequena alteração na PS. A solução proposta é a inclusão de um preâmbulo antes da PS contendo uma seqüência de zeros, que, após a modulação FSK, se transforma em uma senóide, cuja freqüência é distinta em relação às freqüências utilizadas para representar os Símbolos **0** e **1**. Isto permite uma discriminação mais segura. Aplicando-se este método, o efeito transitório de distorção causado pelo CODEC AMR é, então, "sentido" somente pelas amostras do preâmbulo.

Por se tratar de um sinal determinístico, o preâmbulo possui a sua função de autocorrelação constante, o que permite facilmente a identificação da fronteira entre o preâmbulo e a PS. Com a aplicação das técnicas apresentadas na Seção 3.3 consegue-se alcançar o sincronismo de quadro, desde que se conheça o atraso introduzido pelo CODEC e o comprimento do preâmbulo. Na Seção 3.7, serão apresentados resultados simulados para as diversas taxas do CODEC AMR.

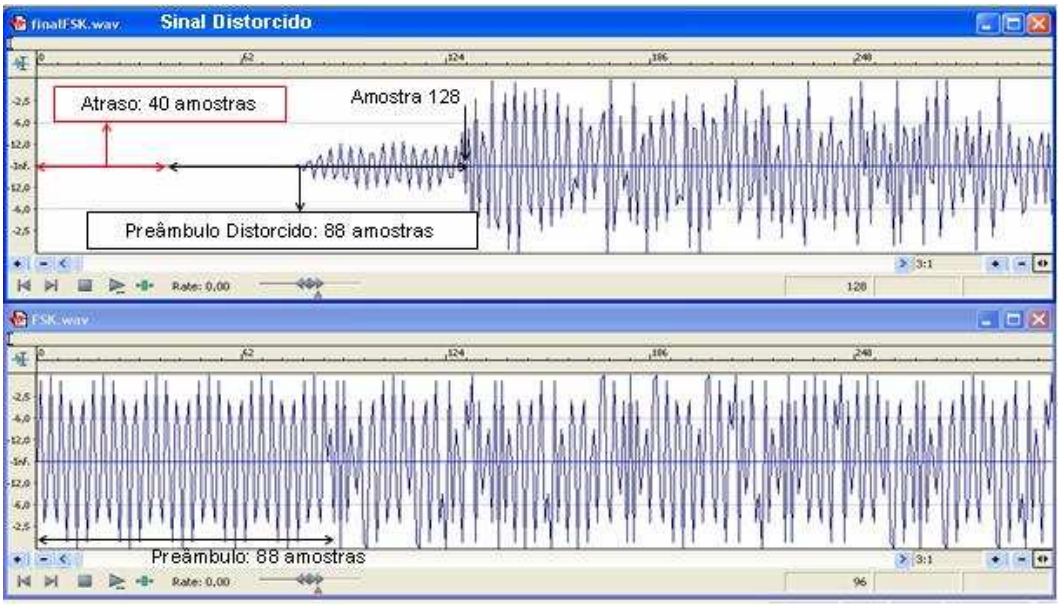

Figura 3.3: Efeitos do CODEC AMR sobre o Sincronismo de Quadro.

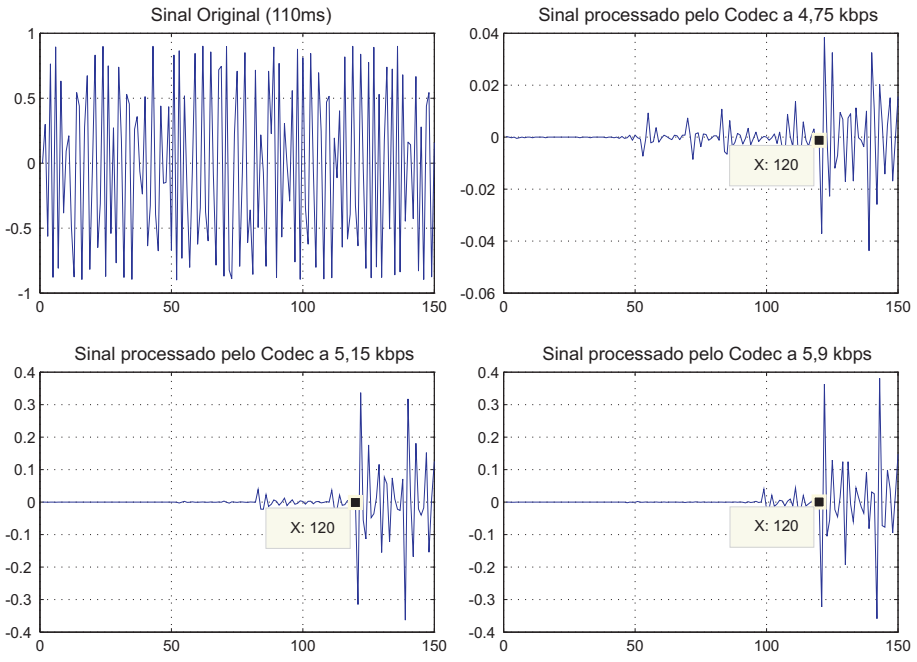

Figura 3.4: Amostras distorcidas para PS de 110 ms e taxas 4,75; 5,15; e 5,9 kbps.

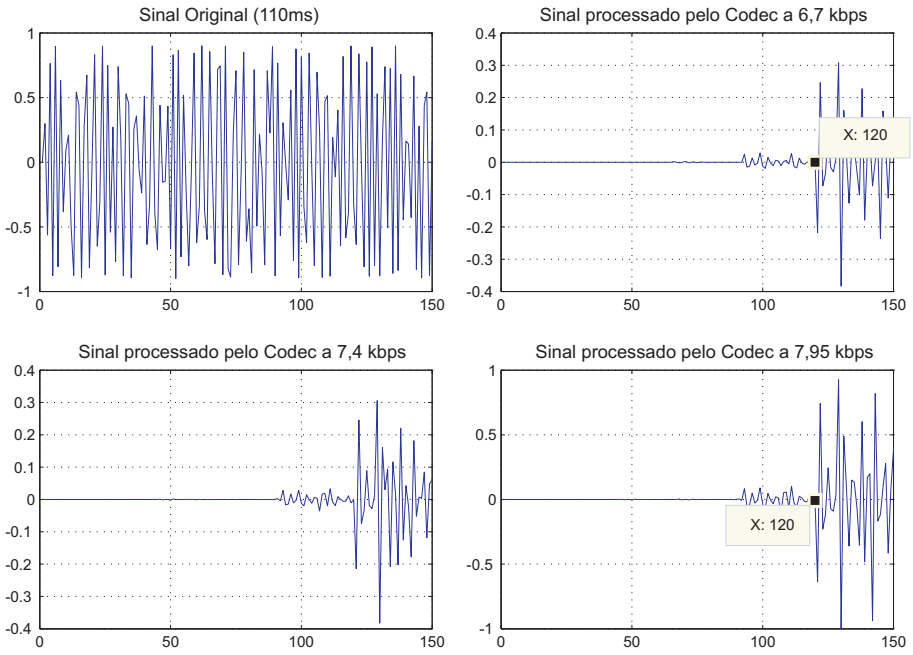

Figura 3.5: Amostras distorcidas para PS de 110 ms e taxas 6,7; 7,4; e 7,95 kbps.

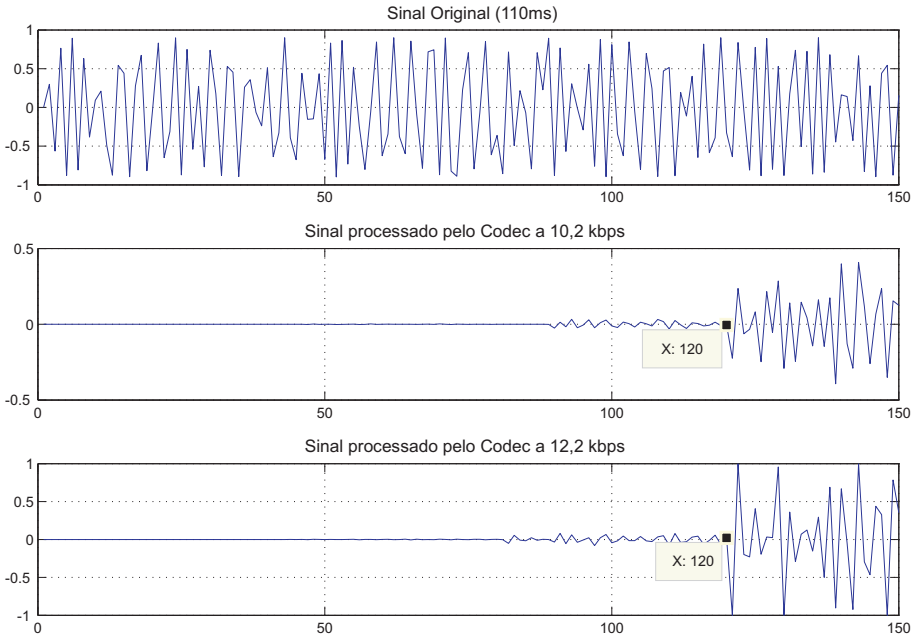

Figura 3.6: Amostras distorcidas para PS de 110 ms e taxas 10,2; e 12,2 kbps.

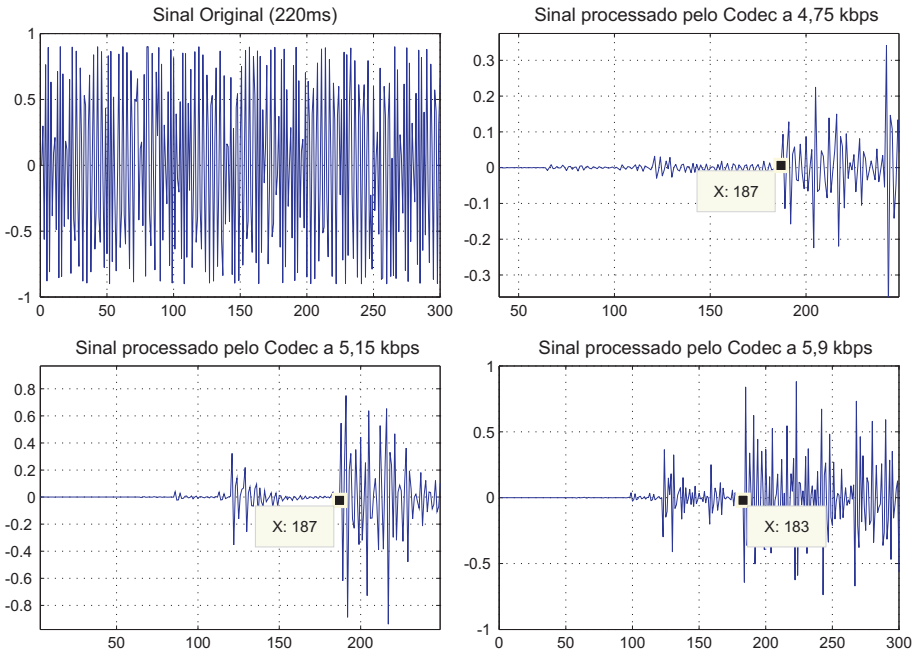

Figura 3.7: Amostras distorcidas para PS de 220 ms e taxas 4,75; 5,15; e 5,9 kbps.

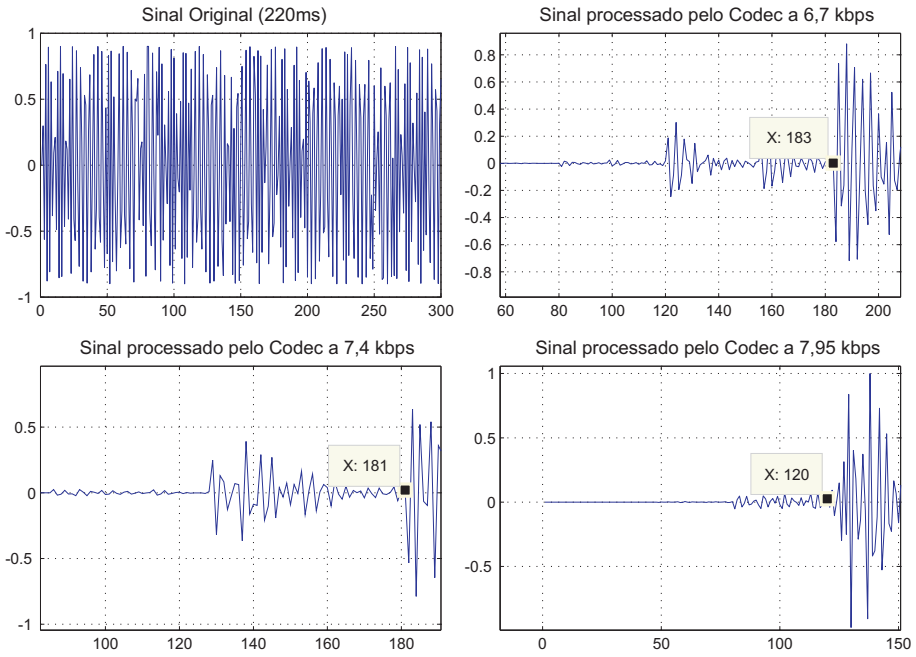

Figura 3.8: Amostras distorcidas para PS de 220 ms e taxas 6,7; 7,4; e 7,95 kbps.

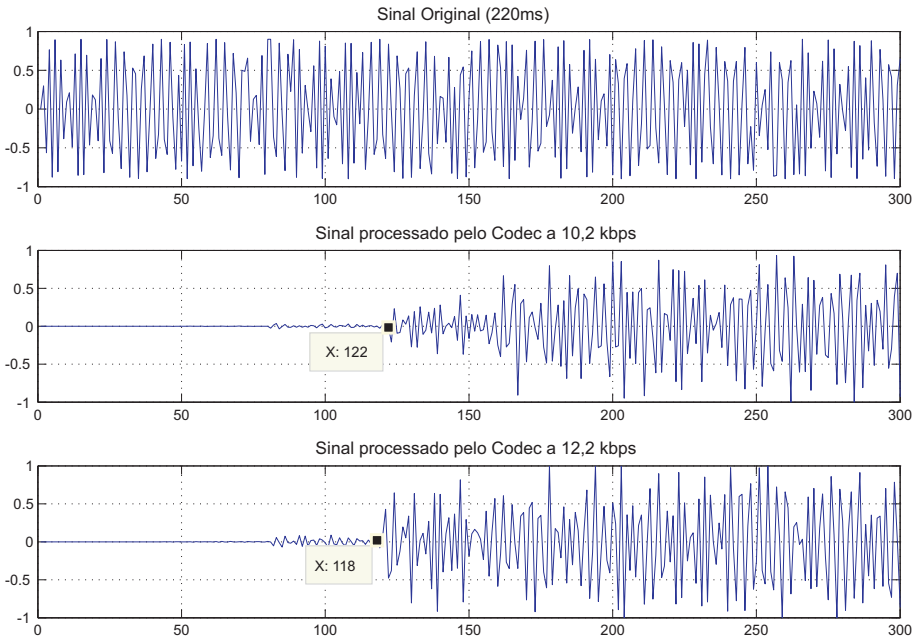

Figura 3.9: Amostras distorcidas para PS de 220 ms e taxas 10,2; e 12,2 kbps.

3.6 Requisitos para Implementação

A implementação eficiente de um esquema de sincronismo de quadro para sistemas de comunicações com criptofonia pressupõe requisitos mandatórios, que influenciam diretamente o desempenho do processo de sincronismo. Inicialmente, devem-se estabelecer os requisitos a serem alcançados com o sincronismo, que, para o caso de cifradores analógicos (Secção 2.2), se resumem à precisão de sincronismo e ao máximo retardo que pode ser introduzido antes do quadro inicial. A precisão de sincronismo, neste trabalho, é definida como o número máximo de amostras atrasadas para o qual ainda se pode alcançar o sincronismo de quadro.

No limite ideal, a precisão de sincronismo deveria ser igual a uma amostra, o que, expressa em termos de tempo, corresponde a $T_b = \frac{1}{f_s}$. Em decorrência da limitação imposta pela taxa de Nyquist, a precisão de uma amostra não pode ser alcançada pelo método aqui apresentado. Neste caso, o valor mínimo de $T_b = T_{\text{mín}}$ é função do máximo desvio de freqüência utilizado na modulação AFSK (ver Apêndice C).

A máxima precisão é dada pela equação:

$$T_{\text{mín}} = \frac{2}{f_{\texttt{FSKmáx}} - f_{\texttt{FSKmín}}}, \tag{3.11}$$

o que, expresso em número de amostras, é:

$$N_{\text{mín}} = \frac{2f_s}{f_{\texttt{FSKmáx}} - f_{\texttt{FSKmín}}} = \frac{T_{\text{mín}}}{T_s}. \tag{3.12}$$

De acordo com a Equação 3.12 e levando-se em consideração a taxa de Nyquist, que estabelece que $f_{\texttt{FSKmáx}} - f_{\texttt{FSKmín}} \leq \frac{f_s}{2}$, chega-se a uma precisão máxima de 4 amostras. Após ser estabelecido o sincronismo de quadro, o resultado pode ser melhorado com a aplicação de técnicas de sincronismo de amostra (ver Seção 3.2).

3.7 Resultados

Nesta seção, serão apresentados resultados de simulações realizadas cujo propósito é testar a validade da aplicação do método proposto na Secção 3.3. As simulações foram realizadas para as diversas taxas do CODEC AMR, sendo, portanto, a sua aplicação também adequada aos CODEC GSM (*Half-Rate e Enhanced Full-Rate*) [25]-[24].

Para as simulações, a separação de freqüências empregada para modulação *A*FSK foi de 2 kHz, com $f_{\texttt{FSKmín}} = 1150$ kHz e $f_{\texttt{FSKmáx}} = 3150$ kHz. O detector FSK utilizado foi do tipo *Detector Ótimo* implementado com correlatores.

Foram empregadas PS de comprimentos 60 e 110 ms, que em decorrência dos atrasos introduzidos, não são adequadas à solução do sincronismos inter-quadro, aplicando-se, apenas, ao sincronismo de quadro inicial. Os valores de 60 e 110 ms foram escolhidos experimentalmente com base no atraso introduzido e nos resultados obtidos para a autocorrelação.

Para o sincronismo inter-quadro devem ser empregadas PS que provoquem retardos desprezíveis e que, quando da demodulação da informação, sejam imperceptíveis. Para facilitar o processo de mascaramento do sinal de sincronismo pode-se, após a extração da informação necessária à sincronização, fazer uso de filtros *Notch* [26] com *notches* nas freqüências $f_{\texttt{FSKmín}}$ e $f_{\texttt{FSKmáx}}$.

A PS de 60 ms possui 480 amostras, sendo formada de 72 seqüências de Barker de ordem $N = 5$ com 120 amostras de preâmbulo. Esta configuração de PS alcança uma precisão de sincronismo de 4 amostras e permite detectar atrasos de até 360 $(480 - 120)$ amostras.

A PS de 110 ms possui 880 amostras, sendo formada de 64 seqüências de Barker de ordem $N = 11$ com 176 amostras de preâmbulo. Esta configuração de PS alcança uma precisão de sincronismo de 4 amostras e permite detectar atrasos de até 704 $(880 - 176)$ amostras.

As Tabelas 3.2 e 3.3 apresentam os valores dos atrasos reais provocados pelo CODEC AMR para as PS de 60 ms e 110 ms, cujos resultados foram mais satisfatórios com o emprego da segunda palavra de sincronismo (PS de 110 ms). Os resultados apresentados na Tabela 3.3 contêm erro apenas para a taxa de $4,75kbps$, que é decorrente da forte distorção provocada pelo CODEC nesta taxa de codificação. De acordo com estes resultados, pode-se concluir que a utilização da PS de 110 ms, seguida da aplicação do método descrito na Seção 3.2, é adequada à solução do problema de sincronismo apresentado neste capítulo.

Tabela 3.2: Atrasos obtidos para PS com 60 ms de duração, composta de 72 seqüências de Barker de ordem $N = 5$ e 120 amostras de preâmbulo, perfazendo 480 amostras

Taxa (kbps)	τ_m para máxima Correlação cruzada Normalizada (amostras)	Atraso Real (amostras)	Atraso Calculado (amostras)	Erro (amostras)
4, 75	244	55	124	+69
5, 15	176	40	56	+16
5, 90	160	40	40	0
6, 70	160	40	40	0
7, 40	160	40	40	0
7, 95	160	40	40	0
10, 20	160	40	40	0
12, 20	200	40	80	+40

Tabela 3.3: Atrasos obtidos para PS com 110 ms de duração, composta de 64 seqüências de Barker de ordem $N = 11$ e 176 amostras de preâmbulo, perfazendo 880 amostras

Taxa (kbps)	τ_m para máxima Correlação cruzada Normalizada (amostras)	Atraso Real (amostras)	Atraso Calculado (amostras)	Erro (amostras)
4, 75	240	55	64	+9
5, 15	216	40	40	0
5, 90	216	40	40	0
6, 70	216	40	40	0
7, 40	216	40	40	0
7, 95	216	40	40	0
10, 20	216	40	40	0
12, 20	216	40	40	0

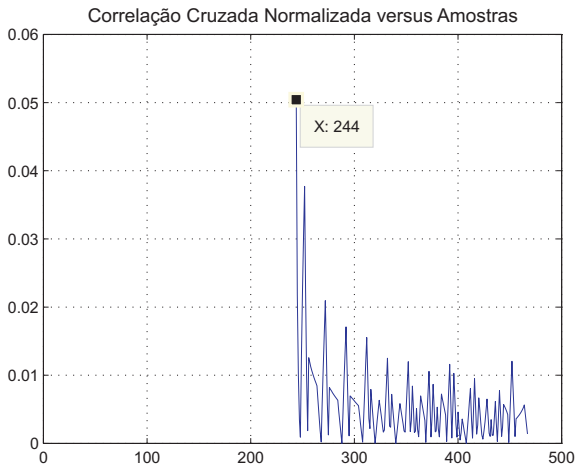

Figura 3.10: Correlação cruzada para PS de 60 ms e taxa de 4,75 kbps.

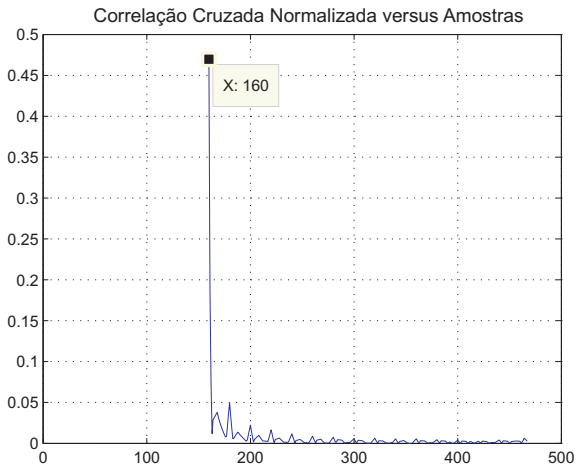

Figura 3.12: Correlação cruzada para PS de 60 ms e taxa de 5,9 kbps.

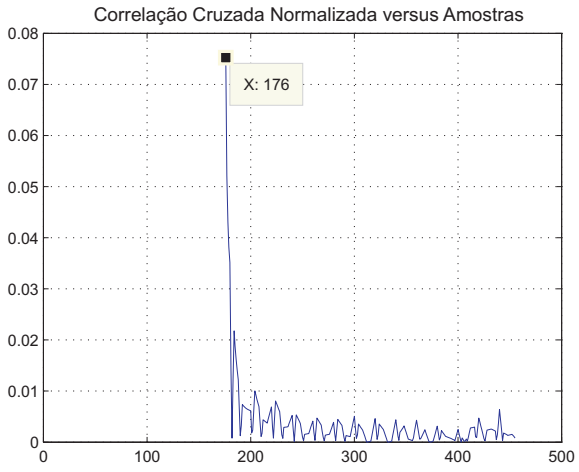

Figura 3.11: Correlação cruzada para PS de 60 ms e taxa de 5,15 kbps.

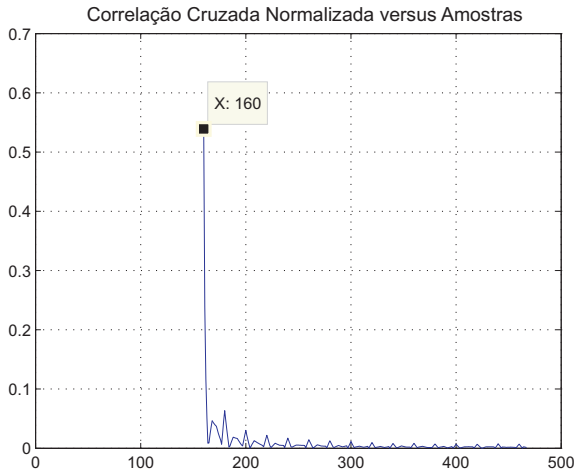

Figura 3.13: Correlação cruzada para PS de 60 ms e taxa de 6,7 kbps.

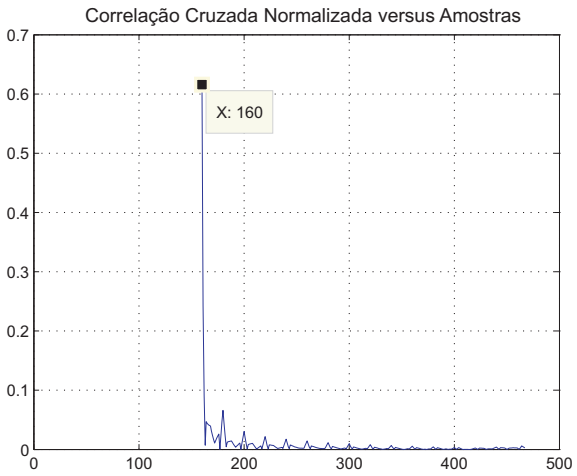

Figura 3.14: Correlação cruzada para PS de 60 ms e taxa de 7,4 kbps.

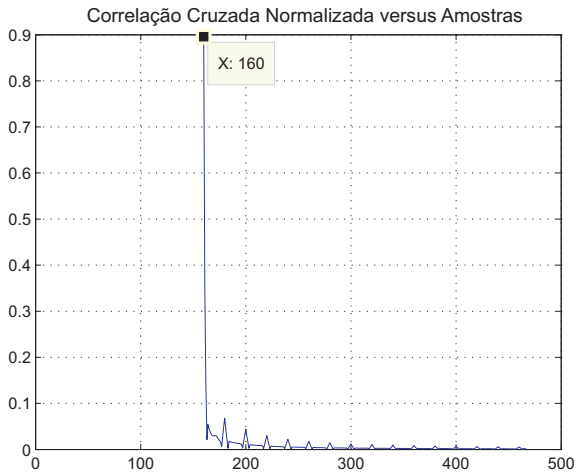

Figura 3.16: Correlação cruzada para PS de 60 ms e taxa de 10,2 kbps.

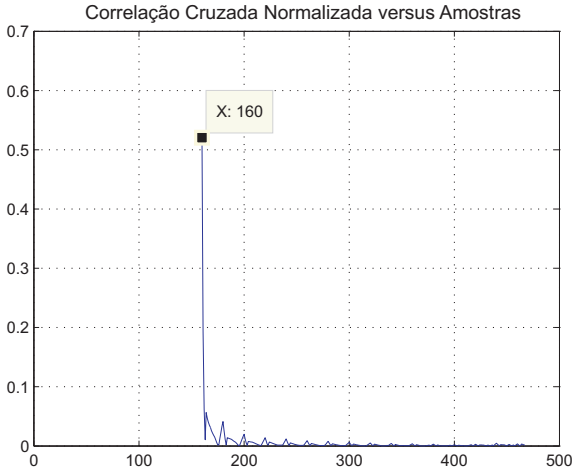

Figura 3.15: Correlação cruzada para PS de 60 ms e taxa de 7,95 kbps.

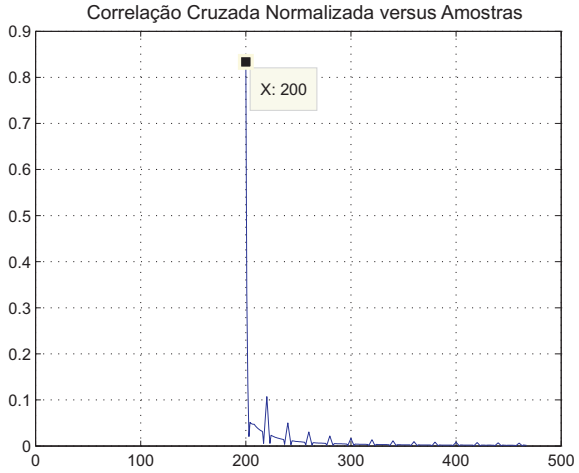

Figura 3.17: Correlação cruzada para PS de 60 ms e taxa de 12,2 kbps.

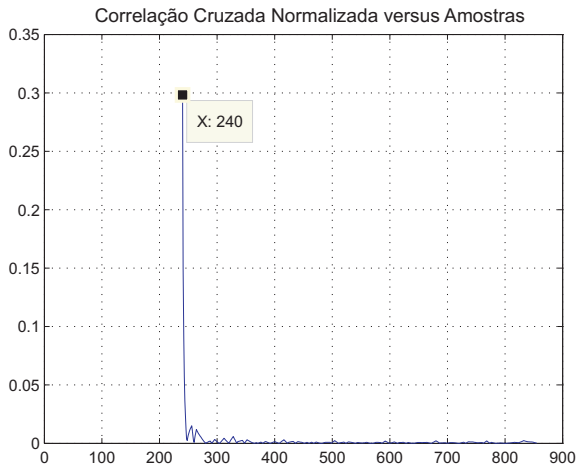

Figura 3.18: Correlação cruzada para PS de 110 ms e taxa de 4,75 kbps.

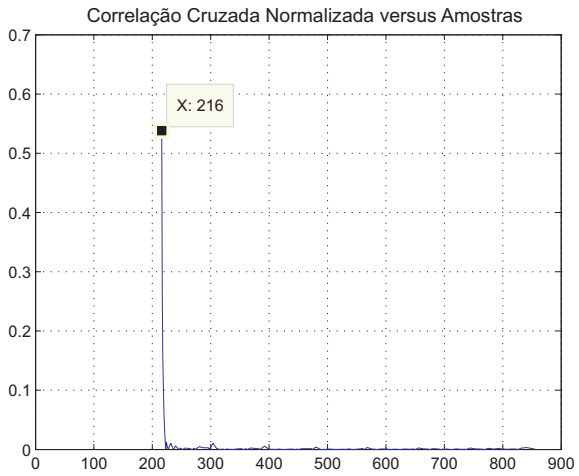

Figura 3.20: Correlação cruzada para PS de 110 ms e taxa de 5,9 kbps.

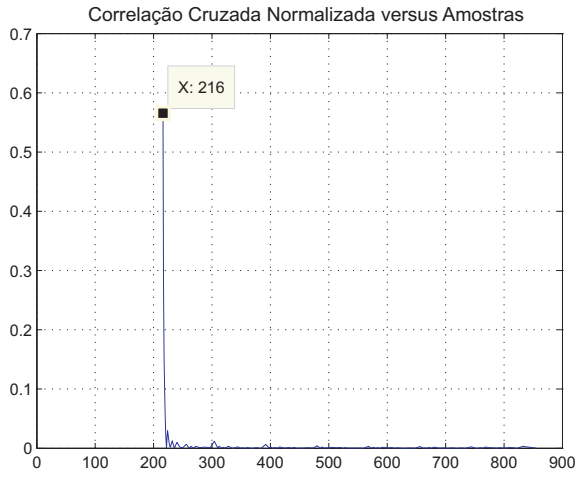

Figura 3.19: Correlação cruzada para PS de 110 ms e taxa de 5,15 kbps.

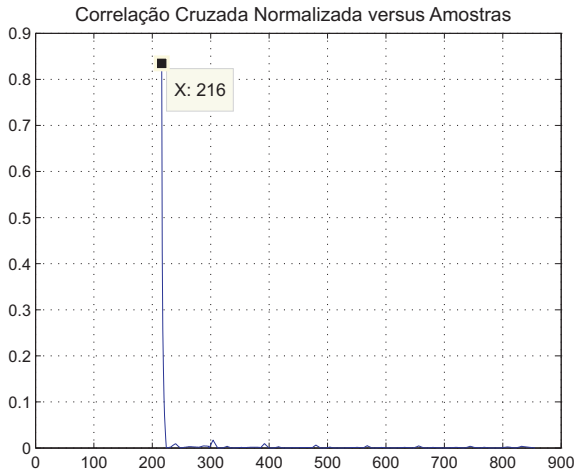

Figura 3.21: Correlação cruzada para PS de 110 ms e taxa de 6,7 kbps.

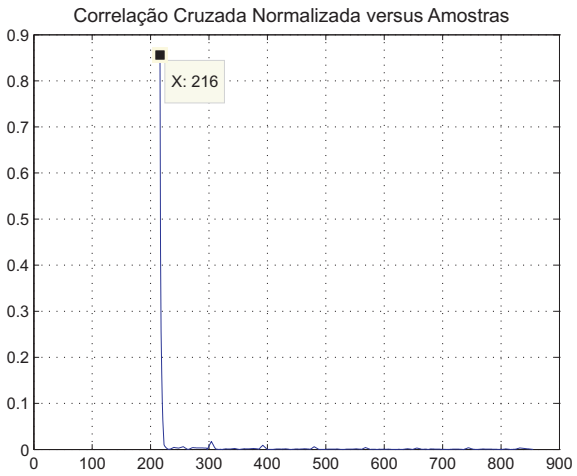

Figura 3.22: Correlação cruzada para PS de 110 ms e taxa de 7,4 kbps.

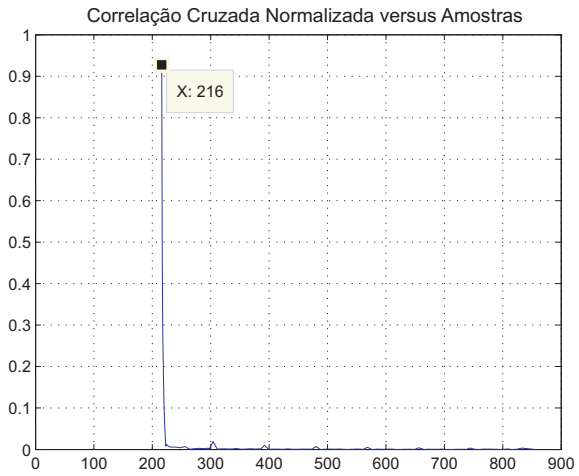

Figura 3.24: Correlação cruzada para PS de 110 ms e taxa de 10,2 kbps.

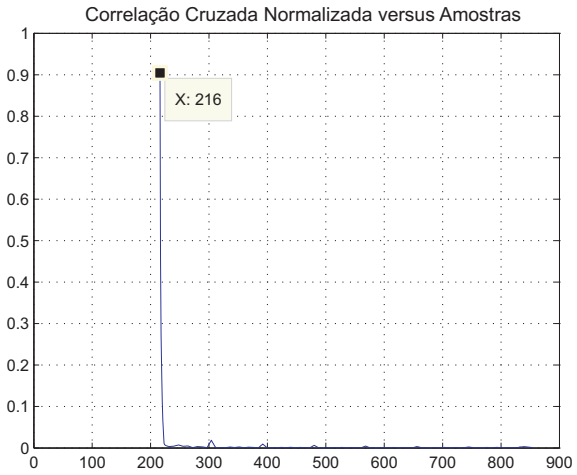

Figura 3.23: Correlação cruzada para PS de 110 ms e taxa de 7,95 kbps.

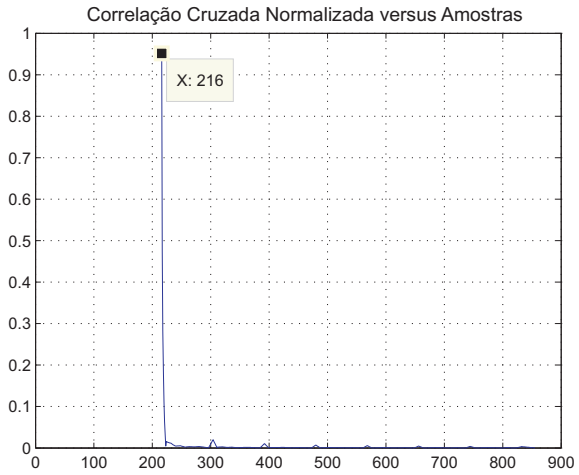

Figura 3.25: Correlação cruzada para PS de 110 ms e taxa de 12,2 kbps.

Capítulo 4

Medidas Objetivas de Qualidade

4.1 Introdução

No Capítulo 5 serão apresentados os resultados comparativos entre os arquivos em claro, cifrados e decifrados pelas técnicas CSI-F. A avaliação de qualidade dos arquivos decifrados e da inteligibilidade residual dos arquivos cifrados é realizada por meio da aplicação de técnicas de medidas objetivas de qualidade, pois os métodos de avaliação subjetiva são dispendiosos e demandam muito tempo [2].

As metodologias para avaliação objetiva de qualidade podem ser classificadas de diversas formas. Este trabalho se limitará a classificar as medidas objetivas como perceptuais e não-perceptuais.

4.2 Medidas Não-Perceptuais (Distâncias)

Para determinar a diferença entre dois vetores, representando aqui blocos de sinais de voz, faz-se necessário o uso de medidas objetivas que expressem o quão semelhantes estes vetores são um do outro. Estas medidas são, genericamente, denominadas *distâncias*.

Na abordagem empregada neste capítulo, o conceito de distância define o quão semelhantes são dois "segmentos temporais" de realizações pertecentes a um processo estocástico que assume valores vetoriais, ou, alternativamente, quão semelhantes são dois "segmentos temporais" pertencentes a vetores oriundos de processos estocásticos distintos. Esta é a situação correspondente, por exemplo, à comparação entre blocos do sinal original e blocos resultantes do processo de cifragem.

Sejam $\boldsymbol{x}$ e $\boldsymbol{y}$ vetores pertencentes ao espaço vetorial real N-dimensional, denotado

por $\mathbb{R}^N$, então uma métrica definida $d(\boldsymbol{x}, \boldsymbol{y})$ pertencente ao espaço $\mathbb{R}$ é uma função real que atende as seguintes propriedades [5]:

a) $d(\boldsymbol{x}, \boldsymbol{y}) \geq 0$;

b) $d(\boldsymbol{x}, \boldsymbol{y}) = 0$, se, e somente se, $\boldsymbol{x} = \boldsymbol{y}$; e

c) $d(\boldsymbol{x}, \boldsymbol{y}) \ \leq \ d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{x}, \boldsymbol{z}), \ \forall \ \boldsymbol{z} \ \in \ \mathbb{R}^N$

Para se determinar a distância entre os espectros de dois blocos de sinais de voz, faz-se necessário o levantamento de coeficientes capazes de descrever o espectro de um sinal por meio de uma análise a curto tempo. Algumas abordagens amplamente empregadas utilizam medidas de distância espectral baseadas em conjuntos de coeficientes de predição linear (LPC - *Linear Prediction Coefficients*), conforme detalhamento constante da Seção 4.2.1.

4.2.1 Cálculo dos Coeficientes de Predição Linear (LPC)

O cálculo dos coeficientes de predição linear consiste na obtenção dos coeficientes de um filtro cuja resposta em freqüência seja o inverso do espectro do sinal de voz.

O aparelho fonador humano pode ser modelado como sendo a saída de um filtros "só pólos" excitado por um trem de pulsos quase periódico ou por um ruído aleatório [2], conforme esquema apresentado na Figura 4.1.

O filtro $H(z)$ pode ser representado como:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{G}{A(z)} = \frac{G}{1 - \sum_{i=1}^{P} \hat{\boldsymbol{a}}(i) z^{-i}} \tag{4.1}$$

O propósito desta modelagem, conhecida como autorregressiva (AR), é determinar o conjunto de coeficientes $\hat{\boldsymbol{a}}$. Para tanto, pode-se aplicar técnicas para minimizar o erro médio quadrático de predição entre a amostra atual $y[n]$ e a amostra predita $\hat{y}[n]$. O erro de predição pode ser expresso como:

$$e[n] = Gx[n] = y[n] - \hat{y}[n] = y[n] - \sum_{i=1}^{P} \hat{\boldsymbol{a}}(i) y[n-i] \tag{4.2}$$

Algumas abordagens clássicas podem ser utilizadas para a minimização do erro quadrático de predição; entre elas, podem ser citadas o método da covariância e método da autocorrelação. Para os métodos citados algumas soluções são conhecidas, tais como a decomposição de Cholesky para o método da covariância e a solução recursiva de Levinson-Durbin para o método da autocorrelação. Maiores detalhes sobre os métodos supracitados podem ser encontrados em [2].

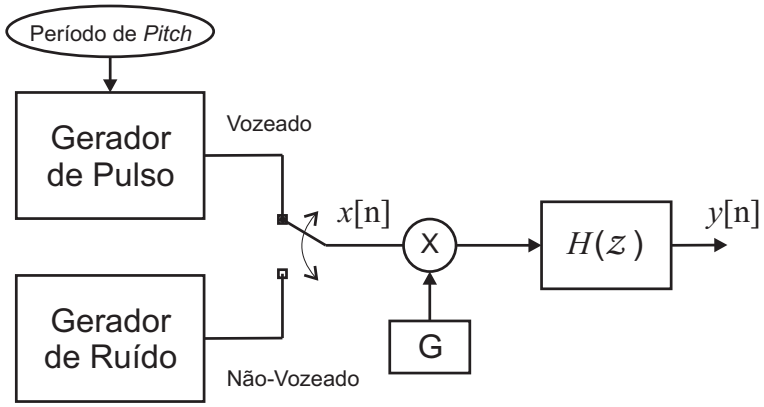

Figura 4.1: Modelagem simplificada para produção de voz.

4.2.2 Distância de Itakura

Dois blocos (ou quadros) pertencentes aos sinais distintos $\boldsymbol{x}$ e $\boldsymbol{y}$ produzem conjuntos de coeficientes LPC [2] diferentes. Pode-se, então, definir uma métrica adequada que expresse o quão diferentes são os referidos blocos. Uma métrica que produz resultados consistentes e que compara coeficientes de predição linear é a distância de Itakura [27]. Esta métrica se apóia na premissa de que o ruído, aliado às imprecisões do modelo de predição linear para sinais de voz, resulta na impossibilidade de se obter os "verdadeiros" coeficientes de predição linear associados a cada bloco do sinal de voz. O conjunto de coeficientes é, então, estimado. Desta forma, o cerne da proposta de Itakura é determinar a probabilidade de o conjunto de parâmetros LPC $\hat{\boldsymbol{a}}$ ser estimado a partir de um bloco do sinal de voz cujos coeficientes "verdadeiros" são os elementos de $\boldsymbol{a}$.

A distribuição de probabilidades da estimativa $\hat{\boldsymbol{a}}$ é uma distribuição de probabilidade gaussiana multidimensional com média $\boldsymbol{a}$ [2], portanto a probabilidade condicional

de ocorrência dos parâmetros LPC $\hat{\boldsymbol{a}}$, dado o conjunto de parâmetros $\boldsymbol{a}$, é expressa como:

$$P(\hat{\boldsymbol{a}}/\boldsymbol{a}) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{\Lambda}_{\hat{a}}|}} \exp\left\{ -\frac{1}{2}(\hat{\boldsymbol{a}} - \boldsymbol{a})\boldsymbol{\Lambda}_{\hat{a}}^{-1}(\hat{\boldsymbol{a}} - \boldsymbol{a})^T \right\}, \tag{4.3}$$

onde $\boldsymbol{\Lambda}_{\hat{a}}$ é a matriz de covariância do bloco considerado e N corresponde ao número de elementos do bloco. A matriz de covariância pode ser definida em função da correlação de $\hat{\boldsymbol{a}}$, $\boldsymbol{R}_{\hat{a}}$:

$$\boldsymbol{\Lambda}_{\hat{a}} = \frac{\boldsymbol{R}_{\hat{a}}^{-1}}{N}\left\{ \hat{\boldsymbol{a}}\boldsymbol{R}_{\hat{a}}\hat{\boldsymbol{a}}^T \right\} \tag{4.4}$$

Uma abordagem baseada no erro quadrático médio pode ser empregada para obter uma formulação simples para a distância de Itakura. Seja $e_y[n] = y[n] - \sum_{i=1}^{P} \hat{\boldsymbol{a}}(i)y[n-i]$; então, o erro médio quadrático de predição, $E[e_y^2[n]]$, pode ser expresso como:

$$E[e_y^2[n]] = \sum_{n=0}^{N-1} \left(y[n] - \sum_{i=1}^{P} \hat{\boldsymbol{a}}(i)y[n-i] \right)^2 \tag{4.5}$$

$$= \sum_{n=0}^{N-1} y^2[n] - 2\sum_{i=1}^{P} \hat{\boldsymbol{a}}(i) \sum_{n=0}^{N-1} y[n]y[n-i] + \sum_{i=1}^{P}\sum_{j=1}^{P} \hat{\boldsymbol{a}}(i)\hat{\boldsymbol{a}}(j) \sum_{n=0}^{N-1} y[n-i]y[n-j] \tag{4.6}$$

$$= \sum_{n=0}^{N-1} \phi_{00} - 2\sum_{i=1}^{P} \hat{\boldsymbol{a}}(i)\phi_{0i} + \sum_{i=1}^{P}\sum_{j=1}^{P} \hat{\boldsymbol{a}}(i)\hat{\boldsymbol{a}}(j)\phi_{ij} \tag{4.7}$$

$$= [1 \ -\hat{\boldsymbol{a}}(1) \ -\hat{\boldsymbol{a}}(2) \ldots -\hat{\boldsymbol{a}}(P)] \begin{bmatrix} \phi_{00} & \phi_{01} & \cdots & \phi_{0P} \\ \phi_{10} & \phi_{11} & \cdots & \phi_{1P} \\ \phi_{20} & \phi_{21} & \cdots & \phi_{2P} \\ \vdots & \ldots & \ddots & \vdots \\ \phi_{P0} & \phi_{P1} & \cdots & \phi_{PP} \end{bmatrix} \begin{bmatrix} -1 \\ -\hat{\boldsymbol{a}}(1) \\ -\hat{\boldsymbol{a}}(2) \\ \vdots \\ -\hat{\boldsymbol{a}}(P) \end{bmatrix} \tag{4.8}$$

$$E[e_y^2[n]] = \hat{\boldsymbol{a}}\boldsymbol{\Phi}_{\hat{a}}\hat{\boldsymbol{a}}^T. \tag{4.9}$$

Definindo a matrix $\boldsymbol{R}_{\hat{a}} = \boldsymbol{\Phi}_{\hat{a}}$ como sendo a matriz de autocorrelação, tem-se:

$$= \hat{\boldsymbol{a}}\boldsymbol{R}_{\hat{a}}\hat{\boldsymbol{a}}^T. \tag{4.10}$$

De uma maneira análoga, pode-se obter o erro quadrático médio de predição para os parâmetros conhecidos $\boldsymbol{a}$ como:

$$E[e_x^2[n]] = \boldsymbol{a}\boldsymbol{R}_a\boldsymbol{a}^T. \tag{4.11}$$

Desta forma, a distância de Itakura pode ser definida como:

$$\boxed{d(\hat{\boldsymbol{a}}, \boldsymbol{a}) = d(\boldsymbol{y}, \boldsymbol{x}) = \left\{ \frac{\hat{\boldsymbol{a}} \boldsymbol{R}_{\hat{a}} \hat{\boldsymbol{a}}^T}{\boldsymbol{a} \boldsymbol{R}_{\boldsymbol{a}} \boldsymbol{a}^T} \right\}}. \tag{4.12}$$

O numerador da Equação (4.12) representa a energia na saída do filtro inverso tendo como sinal de entrada o sinal $\boldsymbol{y}$. O denominador desta equação representa o erro mínimo de predição.

4.2.3 Distância Cepstral

O cepstro real de um sinal de voz $y[n]$ é definido como:

$$c_y[n] = \mathfrak{F}^{-1}\{ \ln |\mathfrak{F}(y[n])| \} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln |Y(e^{j\omega})| e^{j\omega} d\omega, \tag{4.13}$$

onde $\mathfrak{F}(\cdot)$ denota a DTFT do sinal. Considerando-se a modelagem do sinal de voz apresentada na Figura 4.1, pode-se representar o logarítmo da magnitude do espectro de $y[n]$ como:

$$\begin{aligned} \ln |Y(e^{j\omega})|^2 &= \ln \left| \frac{GX(e^{j\omega})}{A(e^{j\omega})} \right|^2 \\ &= 2 \ln G - 2 \ln \left| 1 - \sum_{i=1}^{P} \hat{\boldsymbol{a}}(i) e^{-j\omega} \right| \end{aligned} \tag{4.14}$$

Para um par de espectros, uma distância representativa é definida como o valor médio quadrático da diferença dos logaritmos de cada densidade espectral [5].

$$d(\hat{\boldsymbol{a}}, \boldsymbol{a}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} | \ \ln |S_{\hat{a}}(e^{j\omega})| - \ln |S_a(e^{j\omega})| \ |^2 d\omega \tag{4.15}$$

Aplicando-se o teorema de *Parseval* à Equação (4.9), encontra-se a distância $d(\hat{\boldsymbol{a}}, \boldsymbol{a})$ em termos dos coeficientes cepstrais:

$$d(\hat{\boldsymbol{a}}, \boldsymbol{a}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} | \ \ln |S_{\hat{a}}(e^{j\omega})| - \ln |S_a(e^{j\omega})| \ |^2 d\omega = \sum_{n=-\infty}^{\infty} (c_{\hat{a}}[n] - c_a[n])^2 \tag{4.16}$$

Em decorrência do cepstro ser real, i.e., $c_a[n] = c_a[-n]$, tem-se:

$$\boxed{d(\hat{\boldsymbol{a}}, \boldsymbol{a}) = (c_{\hat{a}}[0] - c_a[0])^2 + 2 \sum_{n=1}^{\infty} (c_{\hat{a}}[n] - c_a[n])^2} \tag{4.17}$$

Os coeficientes cepstrais podem ser obtidos a partir dos coeficientes de predição linear[1].

4.3 Medidas Perceptuais

Medidas perceptuais são medidas obtidas por meio de algoritmos que fazem uso de modelos psico-acústicos com o propósito de reproduzir parcialmente características do ouvido humano.

Embora este trabalho aplique somente o algoritmo PESQ (*Perceptual Evaluation of Speech Quality*) [28] como método de avaliação peceptual de qualidade de sinal, antes de apresentar detalhes sobre o algoritmo PESQ será realizada uma pequena introdução sobre os algoritmos PSQM (*Perceptual Speech Quality Measure*) e PSQM+. Estes algoritmos foram propostos com o objetivo de se avaliar a qualidade de voz em VOCODER e sistemas de telefonia de banda estreita e podem ser considerados como algoritmos predecessores do algoritmo PESQ.

4.3.1 PSQM

O algoritmo PSQM foi desenvolvido pela empresa holandesa de telecomunicações KPN em 1997 e tem a sua especificação constante da recomendação ITU-T P.861 [29]. As medidas da qualidade realizadas pelo algoritmo PSQM fazem uso de um modelo psico-acústico que reproduz parcialmente características perceptivas do ouvido humano. O sinal é convertido para o domínio psico-acústico através de três operações:

a) Mapeamento tempo-frequência implementado por meio de FFT em conjunto com janela de *Hanning*;

b) Alteração na escala de freqüências de Hertz para uma escala em *Bark* [30];

[1] $\begin{cases} c_y[n] & = -a_y[n] - \frac{1}{n}\sum_{k=1}^{n-1}(n-k)c_y[n-k]a_y[k] \\ c_y[0] & = 2\ln G \\ a_y[0] & = 1 \end{cases}$

c) Compressão da amplitude do sinal de acordo com a sensibilidade auditiva (*loudness*).

Como resultado da comparação entre o sinal original e o sinal perturbado (distorcido) tem-se o fator denominado perturbação de ruído. A distorção é calculada a cada 256 amostras do sinal com 50% de overlap. O valor obtido pela aplicação do algoritmo é denominado PSQM e indica o grau de degradação, numa escala que varia de 0 a 6,5. O valor 0 corresponde a um sinal idêntico ao original, sem degradação, e o valor 6,5 corresponde à degradação máxima.

O valor PSQM pode ser convertido para a escala *Mean Opinion Score* [2] (MOS) [31], de acordo com a equação:

$$MOS = \frac{4}{1 + e^{[0,66PSQM-2]}} + 1 \tag{4.18}$$

Tabela 4.1: Escala MOS

MOS	**Qualidade**
5	Excelente
4	Bom
3	Razoável
2	Pobre
1	Ruim

[2]O *Mean Opinion Score* (MOS), resultado de medidas subjetivas de avaliação, é o índice mais aplicado na avaliação de qualidade de voz.

Tabela 4.2: Valores MOS típicos considerando a locução na língua espanhola e diferentes CODECs [32]

CODEC	Taxa(kbps)	MOS
GSM FR	12,2	3,16
G729a	8,0	3,69
GSM EFR	12,0	3,99
G726 ADPCM	16,0	2,56
G729	8,00	3,80
AMR	4,75	3,06
AMR	12,2	3,92
G711(Lei-A)	64,0	4,34

4.3.2 PSQM+

O algoritmo PSQM+ foi proposto com o objetivo de aprimorar os resultados produzidos pelo algoritmo PSQM. Ele leva em consideração o valor PSQM e a energia do sinal. O maior problema do PSQM é o sincronismo do sinal original com o sinal medido, pois o retardo provocado pelo CODEC é desconhecido e pode sofrer variações. Para que o algoritmo realize a comparação dos sinais de maneira correta, é necessário que haja um perfeito sincronismo (alinhamento temporal) entre eles. Portanto, no intuito de produzir uma avaliação mais realista, o algoritmo PSQM+ não contabiliza a degradação decorrente dos atrasos e variações de atrasos existentes.

4.3.3 Perceptual Evaluation of Speech Quality - PESQ

O algoritmo PESQ (*Perceptual Evaluation of Speech Quality*) é o atual algoritmo padrão da ITU para medida de qualidade de voz em sistemas de telefonia, cuja descrição se encontra na recomendação da ITU-T P.862 [28]. Este padrão foi criado em conjunto pelas empresas *KPN Research* e *British Telecommunications PLC* a partir dos algoritmos PSQM+ e *Perceptual Analysis Measurement System* (PAMS), tendo como propósito a inclusão de recursos que permitissem a avaliação de novos sistemas de telefo-

nia como GSM, VoIP e ISDN, visto que o algoritmo descrito pela recomendação IUT-T P.861 [29] não é eficiente para tratar dos problemas específicos de redes. Um fator de relevância no algoritmo PESQ é que a medida de qualidade é apresentada diretamente na escala MOS.

Em síntese, o algoritmo PESQ segue os mesmos procedimentos do algoritmo PSQM acrescido de algumas alterações para melhoria de desempenho. As principais alterações realizadas foram:

a) Equalização da energia dos sinais;

b) Filtragem dos sinais, de forma que o sinal a ser avaliado tenha características semelhantes a sinais recebidos através de redes telefônicas; e

c) Sincronização no domínio do tempo entre o sinal original e o sinal a ser avaliado.

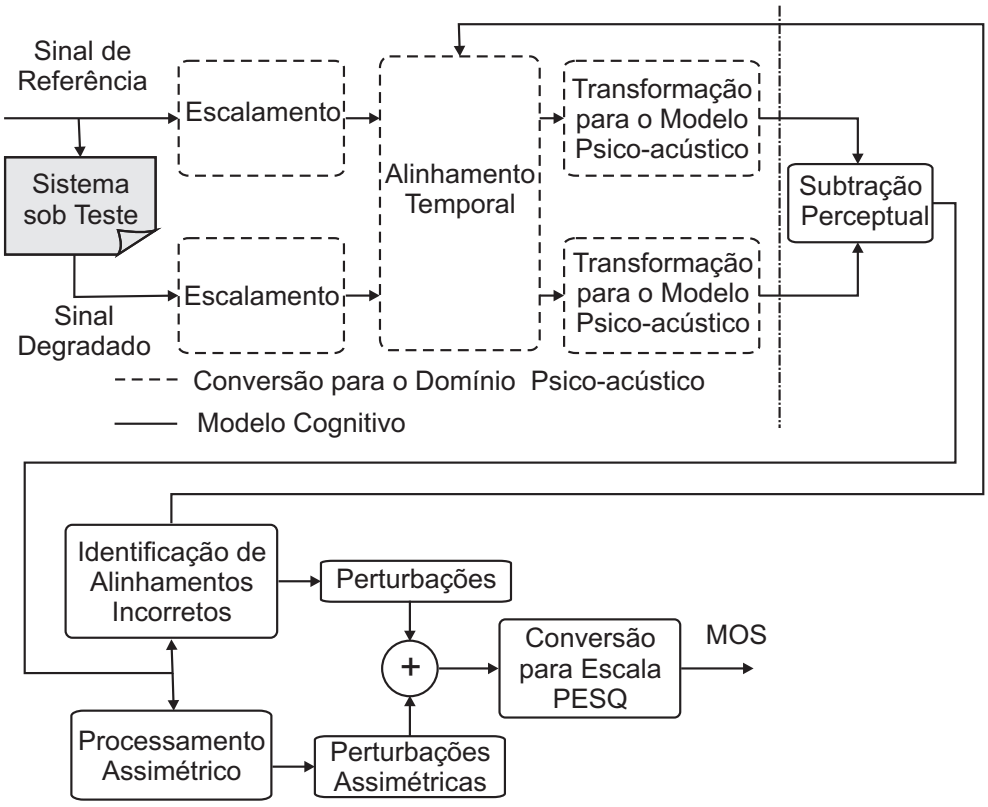

Figura 4.2: Diagram de blocos simplificado do algoritmo PESQ.

Na Figura 4.2 são apresentados os principais blocos funcionais componentes do algoritmo PESQ. Uma importância especial deve ser dada ao bloco de alinhamento temporal, que é responsável pelo sincronismo entre o sinal de referência e o sinal a ser avaliado. Este sincronismo é extrema importância para a avaliação de qualidade de sinais processados por sistemas, cujos atrasos introduzidos não são desprezíveis, como por exemplo CODECs, redes de telecomunicações etc.

Devido a seu bom desempenho em relação aos outros algoritmos apresentados para avaliação da qualidade dos sinais de voz, o algoritmo PESQ tornou-se o padrão (ITU-T P862) para avaliação de qualidade de voz em redes de telefonia de banda estreita e CODEC de sinais de voz.

Capítulo 5

Simulações e Resultados

5.1 Introdução

Neste capítulo serão apresentados detalhes sobre as simulações e seus respectivos resultados. Os resultados comparativos entre o sinal original ("em claro") e os sinais cifrado e decifrado serão apresentados em termos de medidas objetivas. As técnicas de criptofonia abordadas foram as modalidades de CSI-F implementadas por meio de bancos de filtros, CSI-F(BF), e transformada discreta de cossenos, CSI-F(DCT). A técnica de CSI-F(BF) foi implementada em termos de componentes polifásicas (ver Apêndice A).

As demais técnicas apresentadas no Capítulo 2 apresentaram um sinal decifrado ininteligível, quando submetidas aos CODEC AMR/GSM-FR, e, desta forma, não serão objeto das simulações apresentadas neste trabalho.

5.2 Descrição da Metodologia de Simulação

O emprego da criptofonia aplicada a sistemas de comunicações com VOCODER é esquematizado na Figura 5.1. À exceção dos efeitos provocados pelo canal, as simulações apresentadas neste capítulo reproduzem o sistema esquematizado. Os efeitos causados pelo canal são tratados pelo rádio, mais especificamente, pela codificação de canal e códigos corretores de erro do transceptor, transcendendo, portanto, o propósito desta dissertação. Resultados da influência do canal sobre sistemas de criptofonia podem ser encontrados em [33].

Para as simulações aqui apresentadas, foram empregadas 200 frases foneticamente equilibradas para o português do Brasil. Estas sentenças foram fonadas por 40 locutores

do sexo masculino, perfazendo 5 frases por locutor. A duração das 200 frases totalizou 8 minutos. O processo de aquisição do áudio foi realizado ambiente de baixo ruído e fez uso de microfones com cápsula de eletreto.

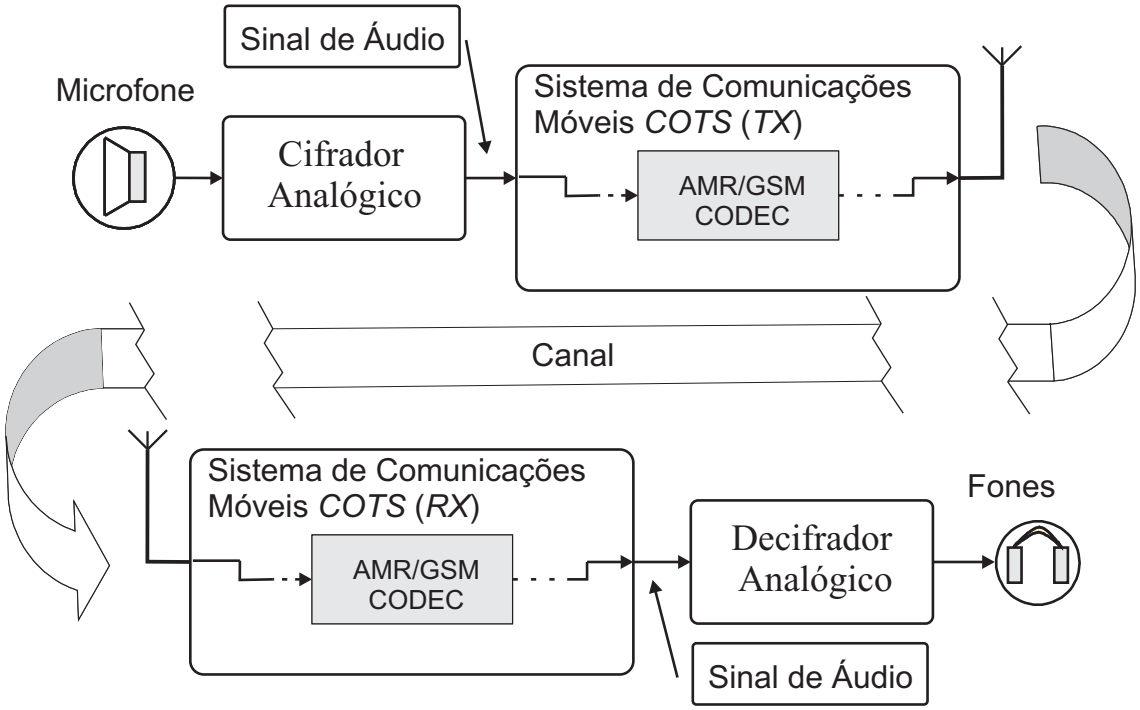

Figura 5.1: Criptofonia aplicada a sistemas de comunicações móveis com VOCODER.

Com a finalidade de simular uma situação real, quando as variações da qualidade do enlace promovem mudanças na taxa de compressão do sinal, quatro taxas de codificação do CODEC AMR foram experimentadas.

Para os testes comparativos entre as técnicas, optou-se por empregar chaves fixas. Contudo, resultados obtidos para CSI-F(DCT) com alteração periódica de chave serão apresentados, onde a diminuição da inteligibilidade residual é evidenciada.

Desde que se pretenda utilizar chaves fixas ou se realize o sincronismo para troca de chaves de maneira local, a adoção da técnica de CSI-F(BF) dispensa o emprego de esquemas de sincronismo por meio de palavras de sincronismo (ver Capítulo 3).

O sincronismo local pode, então, ser realizado com o uso de relógios locais de alta precisão. Para a técnica de CSI-F(DCT), foi necessária a sincronização inicial do sinal,

com base na metodologia descrita no Capítulo 3.

Para demonstrar a influência da alteração periódica de chaves de criptofonia sobre qualidade e inteligibilidade residual dos sinais decifrado e cifrado, respectivamente, serão apresentados resultados MOS, obtidos pelo algoritmo PESQ, distâncias espectrais e o espectrograma do sinal cifrado.

5.3 Resultados

Os dados listados a seguir são comuns às simulações que serão apresentadas nesta seção:

a) Freqüência de amostragem: 8 kHz;

b) Duração do bloco de voz: 20 ms;

c) Número mínimo de frases por locutor: 5 frases;

d) Tipo do banco de filtros: DFT Uniforme;

e) Filtro-protótipo (Figura 5.2):

   – Tipo: FIR (real);

   – Ordem: 158; e

   – Fase: Linear;

f) Número de pontos utilizado no cálculo da DCT para cada bloco: 160;

g) CODEC: AMR; e

h) Taxas de codificação utilizadas: $4,75$; $5,90$; $7,95$; e $12,90$ kbps.

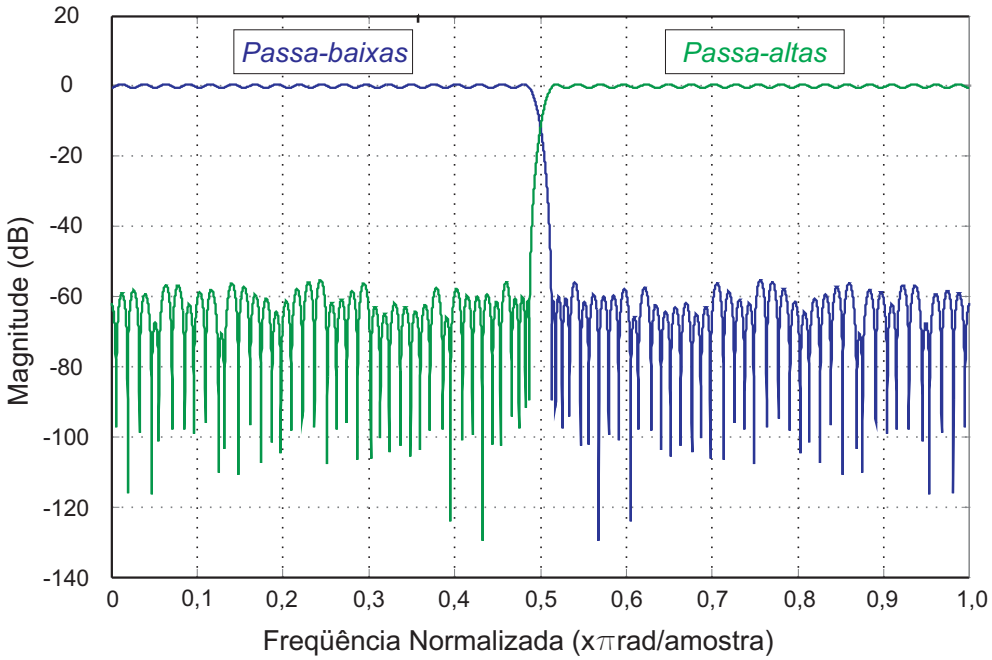

Figura 5.2: Resposta em freqüência dos filtros-protótipo utilizados para implementação da técnica de CSI-F(BF).

5.3.1 Simulação I

Esta simulação teve como propósito a realização de medidas indiretas da inteligibilidade residual do sinal cifrado e avaliação da qualidade do sinal recuperado em sistemas CSI-F com 8 sub-bandas. Nesta seção, serão apresentados os resultados das medidas objetivas de qualidade, obtidos para os sinais cifrado e decifrado.

Dados utilizados na simulação:

a) Técnicas utilizadas: CSI-F(BF) e CSI-F(DCT);

b) Número de sub-bandas para CSI-F(BF): 8;

c) Número de segmentos (subfaixas) por bloco para CSI-F(DCT): 8; e

d) Rotação provocada pela matriz de permutação usada: $\Phi_I = \Phi_I^{\text{Max}} = 53,97^{\text{o}}$.

Tabela 5.1: Medidas indiretas da inteligibilidade residual do sinal cifrado para 8 sub-bandas/segmentos

Método	Taxa do CODEC (AMR)	Distância de Itakura (dB)	Distância Cepstral (dB)
CSI-F(DCT)	4,75 kbps	5,81	5,83
CSI-F(BF)	4,75 kbps	5,22	5,51
CSI-F(DCT)	5,90 kbps	5,80	5,84
CSI-F(BF)	5,90 kbps	5,16	5,49
CSI-F(DCT)	7,95 kbps	5,83	5,90
CSI-F(BF)	7,95 kbps	5,18	5,54
CSI-F(DCT)	12,2 kbps	5,90	5,89
CSI-F(BF)	12,2 kbps	5,04	5,43

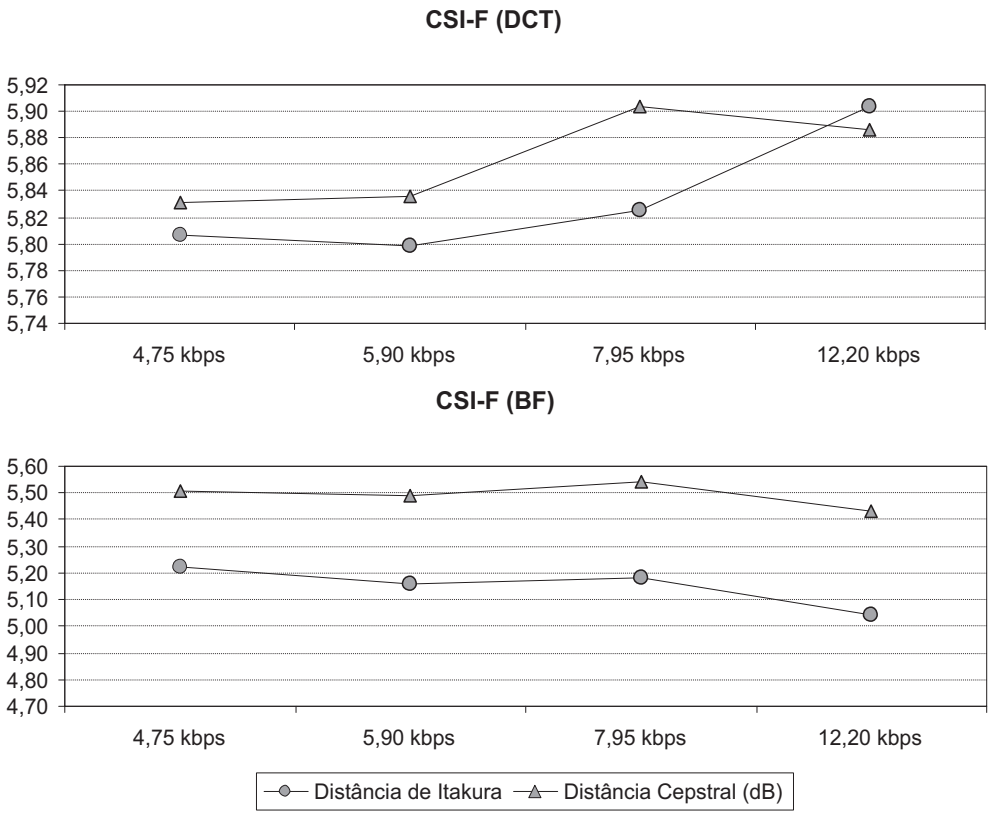

Figura 5.3: Medidas indiretas da inteligibilidade residual o sinal cifrado em função da taxa de compressão (8 sub-bandas/segmentos).

Tabela 5.2: Medidas objetivas de avaliação de qualidade do sinal decifrado 8 sub-bandas/segmentos

Método	Taxa do CODEC (AMR)	Distância de Itakura (dB)	Distância Cepstral (dB)	PESQ
CSI-F(DCT)	4,75 kbps	0,69	−0,17	1,98
CSI-F(BF)	4,75 kbps	1,06	0,10	2,19
CSI-F(DCT)	5,90 kbps	0,58	−0,47	2,03
CSI-F(BF)	5,90 kbps	0,78	−0,41	2,24
CSI-F(DCT)	7,95 kbps	0,45	−0,87	2,18
CSI-F(BF)	7,95 kbps	0,55	−0,83	2,43
CSI-F(DCT)	12,2 kbps	0,22	−2,48	2,82
CSI-F(BF)	12,2 kbps	0,39	−1,33	2,90

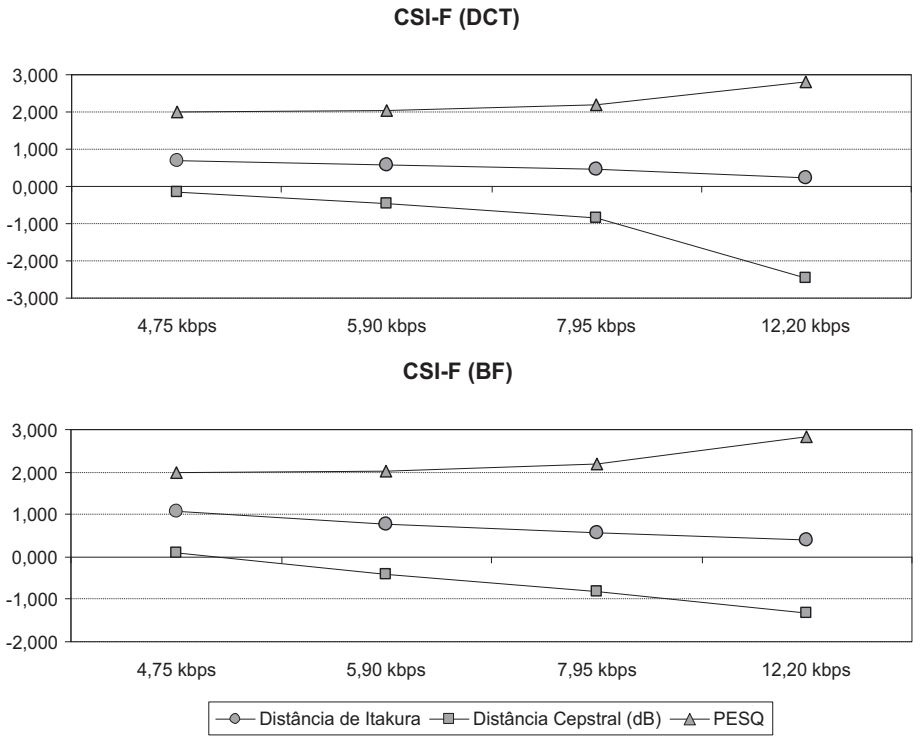

Figura 5.4: Medidas objetivas de qualidade do sinal decifrado em função da taxa de compressão (8 sub-bandas/segmentos).

5.3.2 Simulação II

Esta simulação teve como propósito a realização de medidas indiretas da inteligibilidade residual do sinal cifrado e avaliação da qualidade do sinal recuperado em sistemas CSI-F com 16 subfaixas. Nesta seção, serão apresentados os resultados das medidas objetivas de qualidade obtidos para os sinais cifrado e decifrado.

Dados utilizados na simulação:

a) Técnicas utilizadas: CSI-F(BF) e CSI-F(DCT);

b) Número de sub-bandas para CSI-F(BF): 16;

c) Número de segmentos (subfaixas) por bloco para CSI-F(DCT): 16; e

d) Rotação provocada pela matriz de permutação usada: $\Phi_I = \Phi_I^{\text{Max}} = 56,94^{\text{o}}$.

Tabela 5.3: Medidas indiretas da inteligibilidade residual do sinal cifrado para 16 sub-bandas/segmentos.

Método	Taxa do CODEC (AMR)	Distância de Itakura (dB)	Distância Cepstral (dB)
CSI-F(DCT)	4,75 kbps	5,66	5,78
CSI-F(BF)	4,75 kbps	5,39	5,61
CSI-F(DCT)	5,90 kbps	5,63	5,79
CSI-F(BF)	5,90 kbps	5,31	5,59
CSI-F(DCT)	7,95 kbps	5,65	5,83
CSI-F(BF)	7,95 kbps	5,31	5,64
CSI-F(DCT)	12,2 kbps	5,73	5,85
CSI-F(BF)	12,2 kbps	5,16	5,50

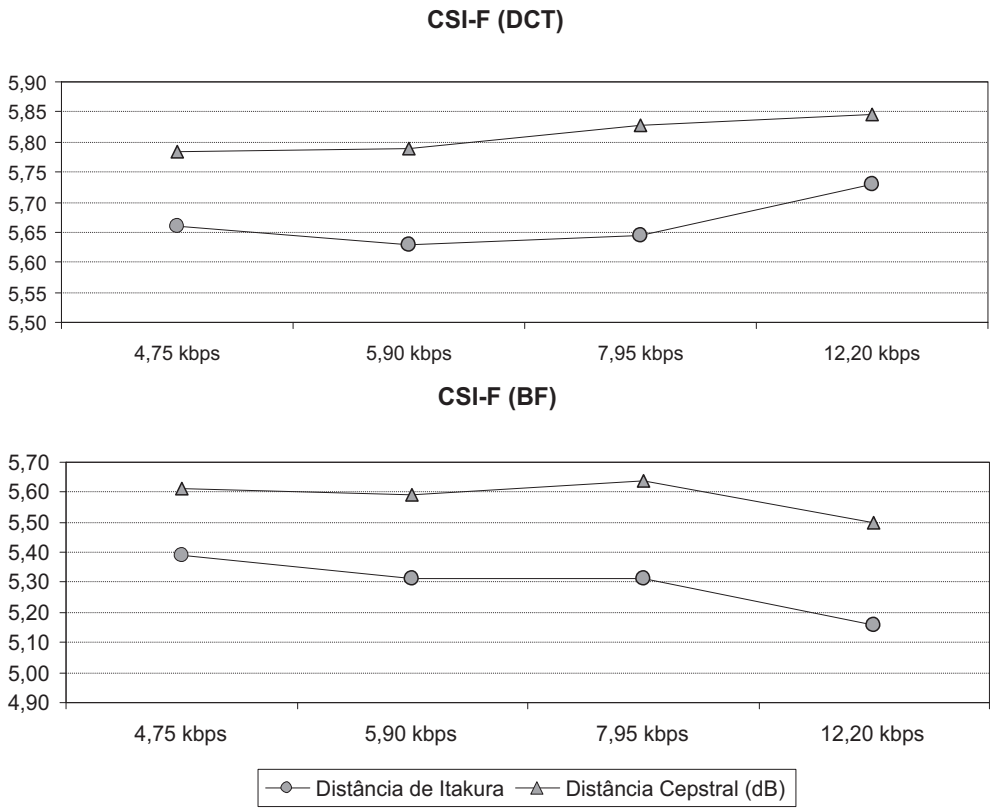

Figura 5.5: Medidas indiretas da inteligibilidade residual o sinal cifrado em função da taxa de compressão (16 sub-bandas/segmentos).

Tabela 5.4: Medidas objetivas de avaliação de qualidade do sinal decifrado para 16 sub-bandas/segmentos

Método	Taxa do CODEC (AMR)	Distância de Itakura (dB)	Distância Cepstral (dB)	PESQ
CSI-F(DCT)	4,75 kbps	0,87	−0,33	1,95
CSI-F(BF)	4,75 kbps	2,02	2,24	1,93
CSI-F(DCT)	5,90 kbps	0,73	−0,69	2,03
CSI-F(BF)	5,90 kbps	0,85	0,03	2,08
CSI-F(DCT)	7,95 kbps	0,45	−1,27	2,24
CSI-F(BF)	7,95 kbps	0,66	−0,31	2,30
CSI-F(DCT)	12,2 kbps	0,19	−2,96	3,02
CSI-F(BF)	12,2 kbps	0,48	−1,04	3,09

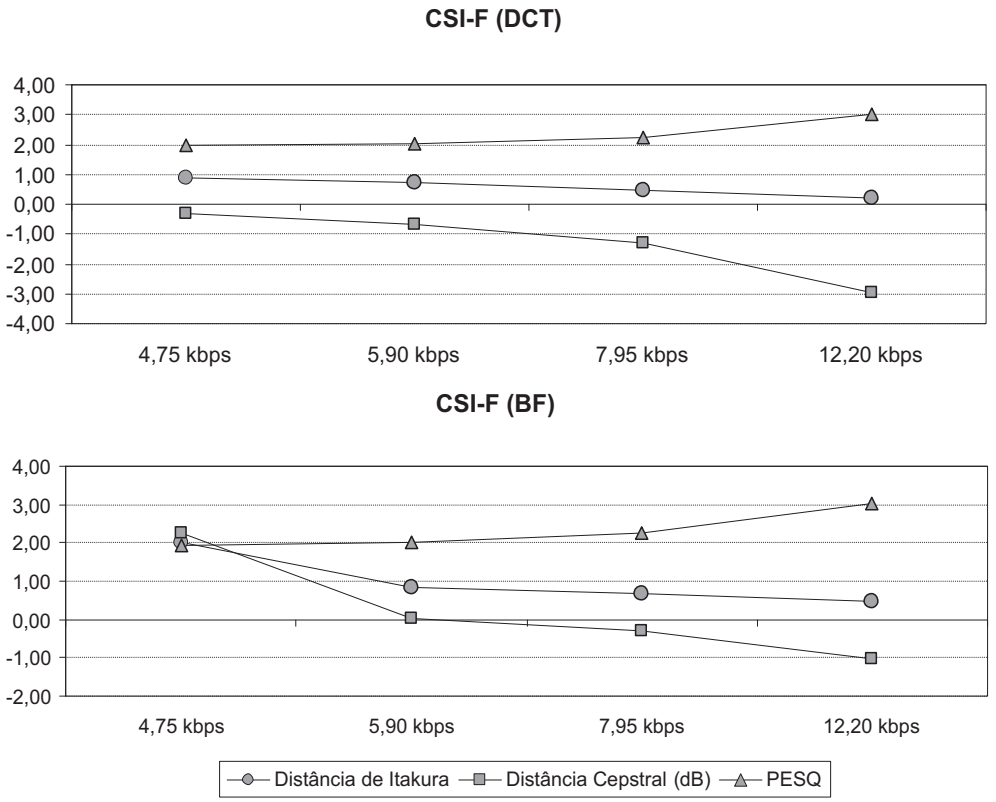

Figura 5.6: Medidas objetivas de qualidade do sinal decifrado em função da taxa de compressão (16 sub-bandas/segmentos).

5.3.3 Simulação III

Esta simulação teve como propósito a realização de medidas indiretas da inteligibilidade residual do sinal cifrado e avaliação da qualidade do sinal recuperado em sistemas CSI-F com 8 sub-bandas e troca periódica de chaves. Procurando-se contemplar o pior caso e para tanto empregou-se a CSI-F(DCT), cujo sincronismo já é crítico para chaves fixas, e o Período de Troca de Chaves (PTC) a cada bloco do sinal, i.e., PTC=1. Esta condição é de difícil implementação prática, tendo apenas como propósito estabelecer um limite comparativo para os demais valores indiretos de inteligibilidade residual obtidos para PTC > 1. As chaves utilizadas foram escolhidas aleatoriamente dentro do subconjunto $\boldsymbol{S}$, para $\mathcal{L}_{\mathcal{I}} = 0,85\Phi_I^{\mathrm{Max}}(N)$ (critério I).

Dados utilizados na simulação:

a) Técnica utilizada: CSI-F(DCT);

b) Número de segmentos (subfaixas) por bloco: 8;

c) Período de troca de chaves: a cada bloco do sinal (PTC=1);

d) Rotação provocada pela matriz de permutação: $0,85\Phi_I^{\text{Max}}(8) \leq \Phi_I \leq \Phi_I^{\text{Max}}(8) = 53,97^{\text{o}}$.

Tabela 5.5: Medidas indiretas da inteligibilidade residual para o sinal cifrado com alteração periódica do valor da chave (PTC=1)

Método	Taxa do CODEC (AMR)	Distância de Itakura (dB)	Distância Cepstral (dB)
CSI-F(DCT)	4,75 kbps	6,09	5,57
CSI-F(DCT)	5,90 kbps	6,08	5,54
CSI-F(DCT)	7,95 kbps	6,06	5,51
CSI-F(DCT)	12,2 kbps	6,05	5,50

Tabela 5.6: Medidas indiretas da inteligibilidade residual do sinal cifrado.

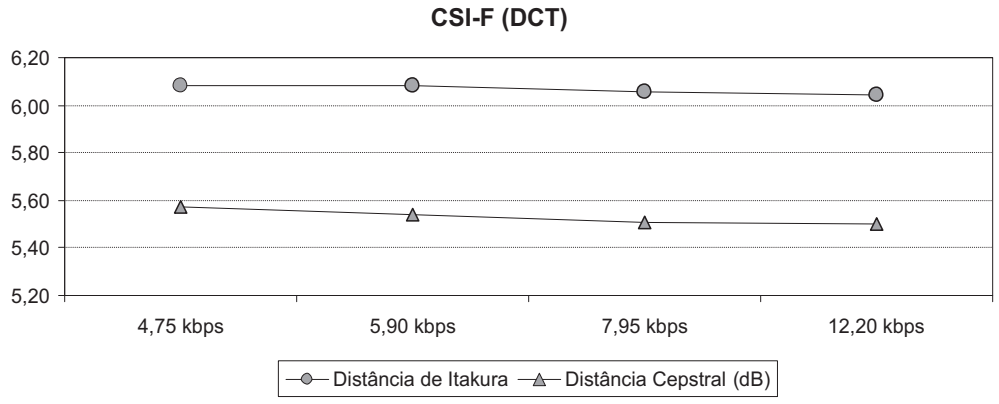

Tabela 5.7: Medidas objetivas de avaliação de qualidade do sinal decifrado

Método	Taxa do CODEC (AMR)	Distância de Itakura (dB)	Distância Cepstral (dB)	PESQ
CSI-F(DCT)	4,75 kbps	1,23	1,36	1,41
CSI-F(DCT)	5,90 kbps	1,08	1,16	1,52
CSI-F(DCT)	7,95 kbps	0,85	0,79	1,71
CSI-F(DCT)	12,2 kbps	0,47	−0,48	2,34

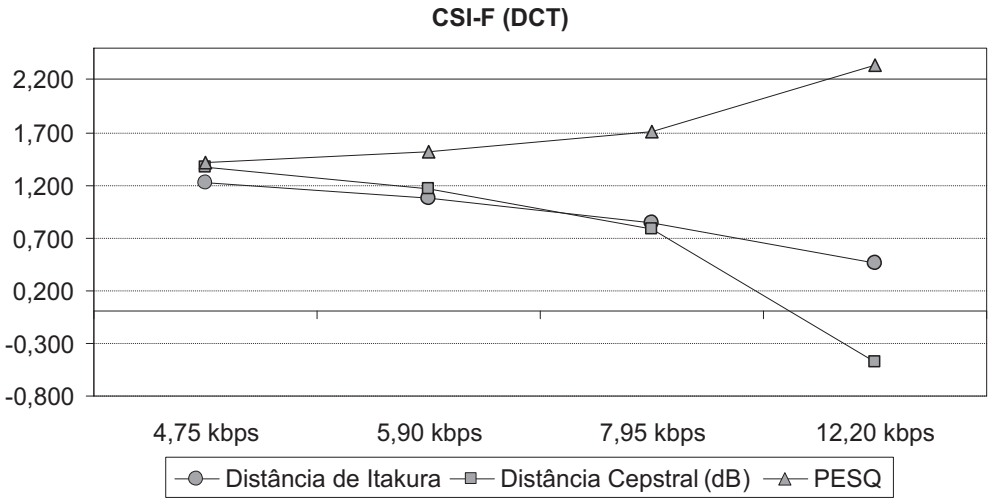

Figura 5.7: Medidas objetivas para o sinal decifrado com mudança periódica de chave a cada bloco do sinal de voz (PTC=1).

5.3.4 Simulação IV

Esta simulação teve como propósito a realização de medidas indiretas da inteligibilidade residual média do sinal em função do ângulo da rotação provocada pela matriz de permutação. Em decorrência da existência de chaves distintas (matrizes de permutação) que provocam rotações idênticas, optou-se por fazer uso de um conjunto de 60 chaves, para um dado ângulo de rotação, e tomar a média dos resultados parciais como resultado final para cada ângulo.

Dados utilizados na simulação:

a) Técnica utilizada: CSI-F(BF) ;

b) Número de sub-bandas: 8;

c) Rotações provocadas pelas matrizes de permutação: $\Phi_I = 15,05^o;\ 25,58^o;\ 30,05^o;$ e $45,10^o$; e

d) Taxa de compressão: 12, 20 kbps.

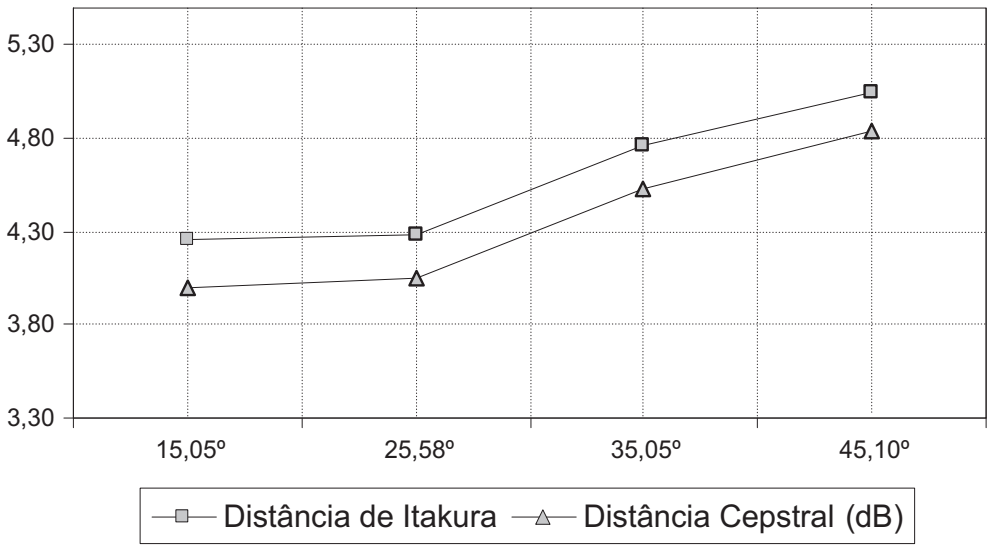

Figura 5.8: Medida indireta da inteligibilidade residual média em função do ângulo de rotação Φ_I.

5.3.5 Simulação V

Esta simulação teve como meta obter resultados objetivos de qualidade do sinal em função do período de troca de chaves.

Dados utilizados na simulação:

a) Técnica empregada: CSI-F(DCT);

b) Número de segmentos (subfaixas) por bloco: 8;

c) Períodos de troca de chave: 1, 2, 4 e 8 blocos.

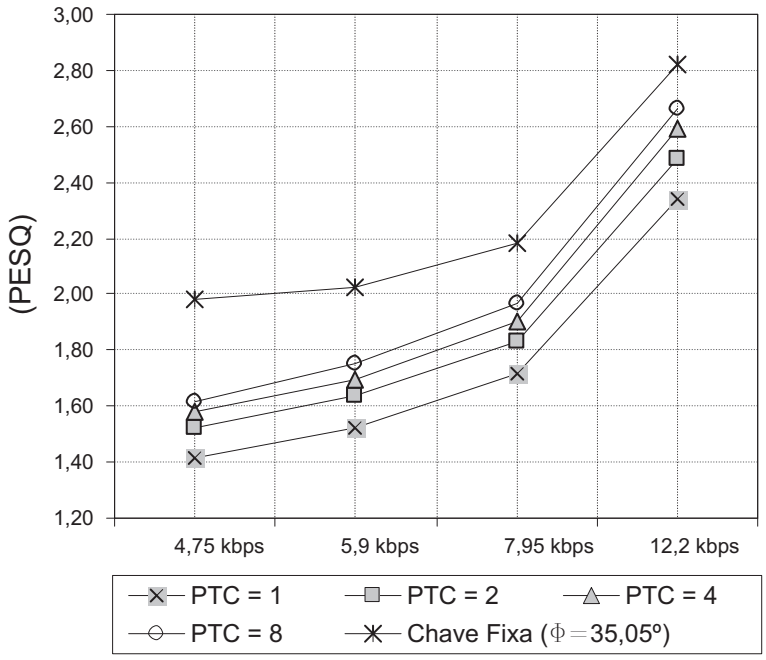

Figura 5.9: Medida PESQ em função da taxa de compressão e do período de troca de chaves (PTC).

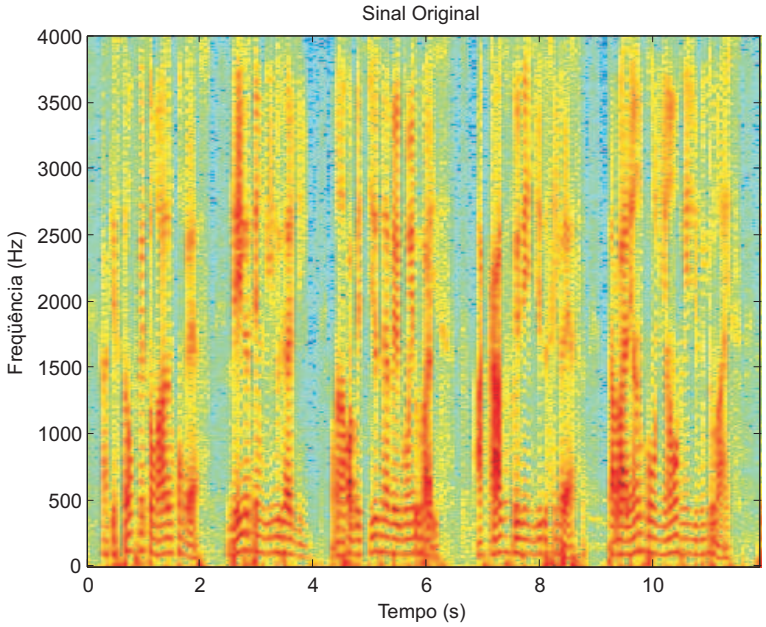

Figura 5.10: Espectrograma do sinal em claro.

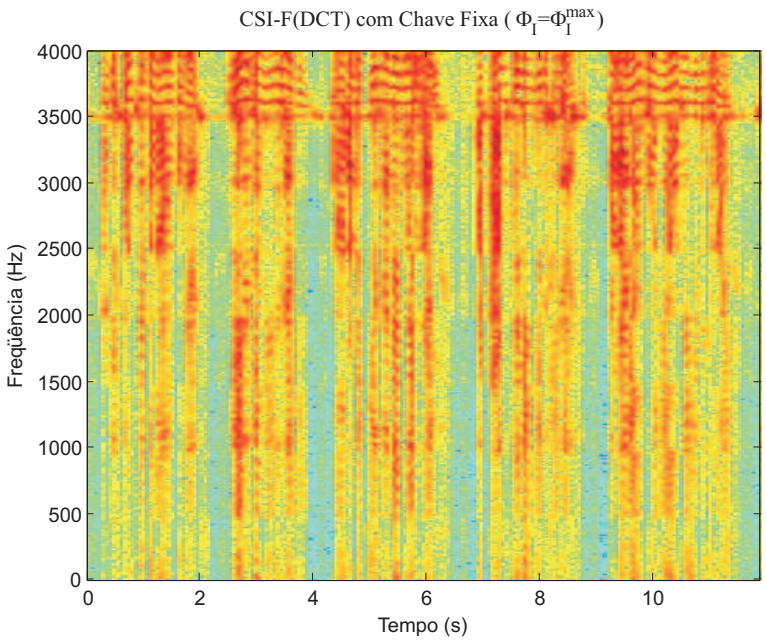

Figura 5.11: Espectrograma do sinal cifrado bruto com chave fixa.

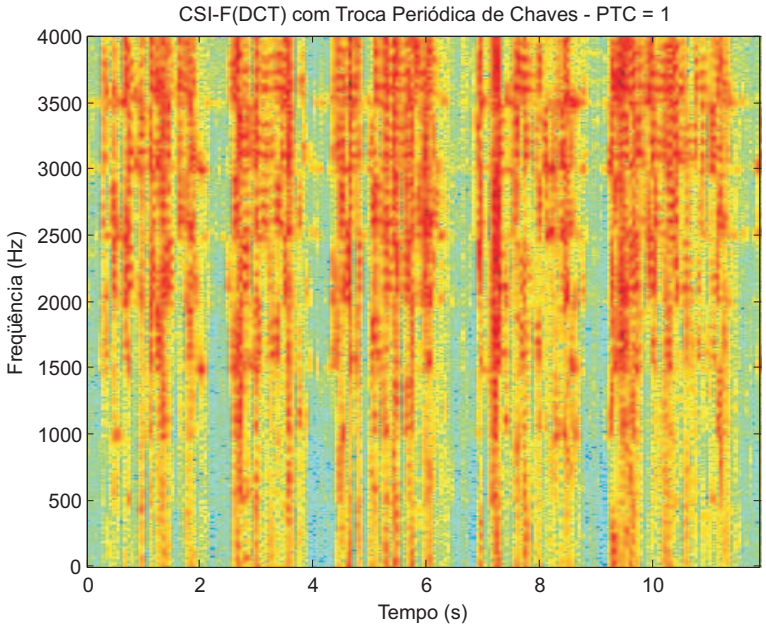

Figura 5.12: Espectrograma do sinal cifrado bruto com mudança periódica de chave (PTC=1).

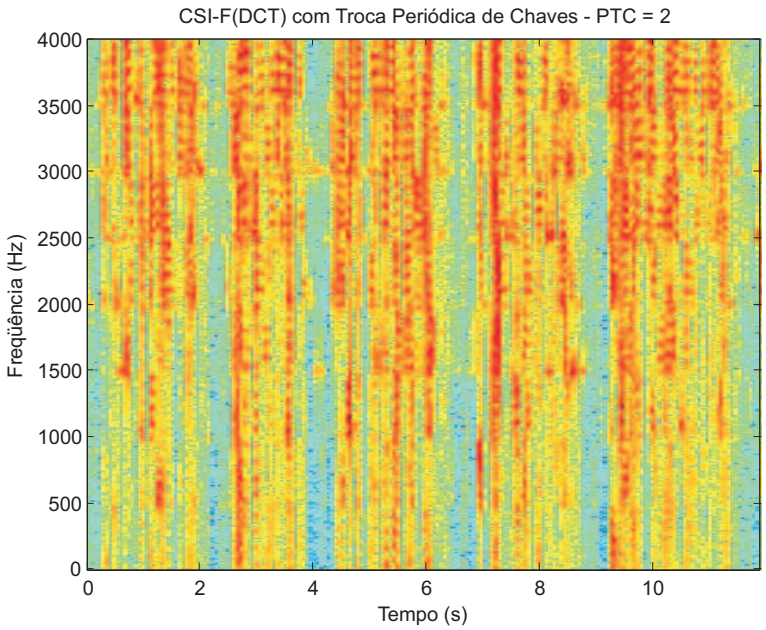

Figura 5.13: Espectrograma do sinal cifrado bruto com mudança periódica de chave (PTC=2).

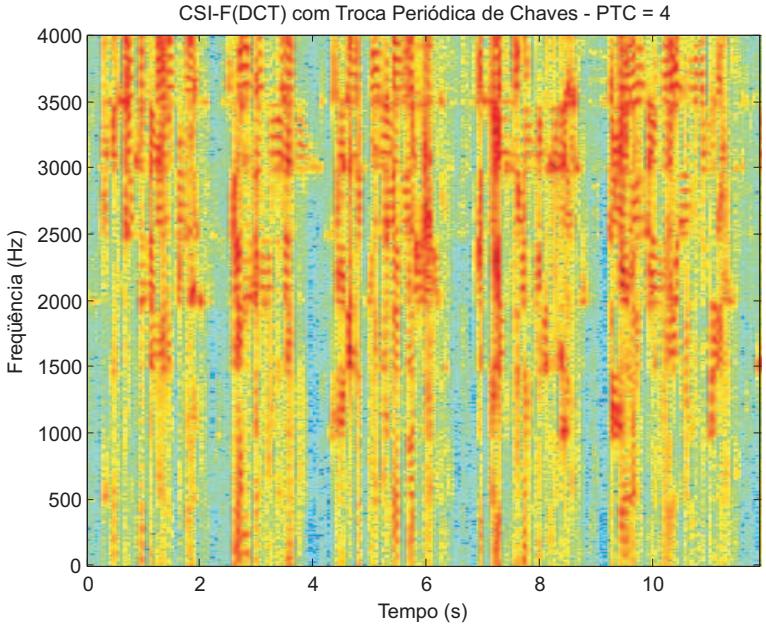

Figura 5.14: Espectrograma do sinal cifrado bruto com mudança periódica de chave (PTC=4).

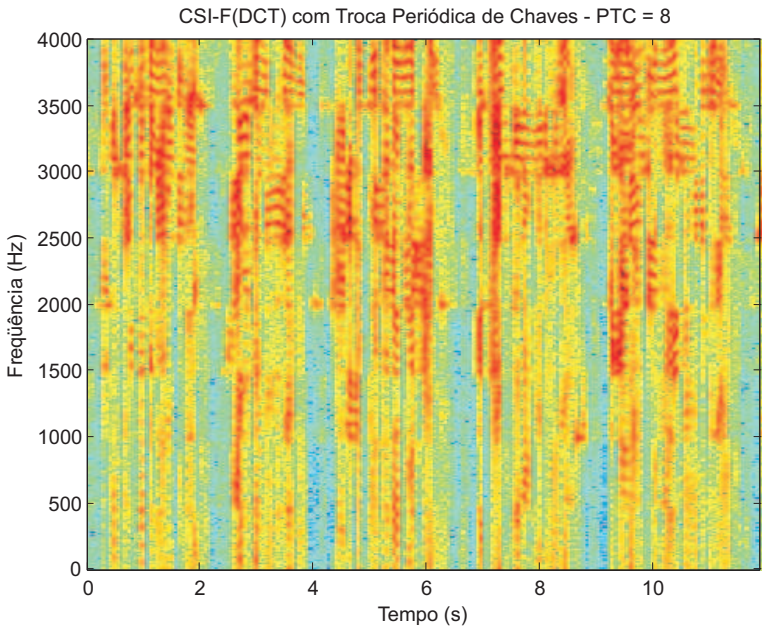

Figura 5.15: Espectrograma do sinal cifrado bruto com mudança periódica de chave (PTC=8).

5.4 Análise dos Resultados

Os resultados apresentados na seção anterior permitem evidenciar os pontos favoráveis e desfavoráveis das técnicas de criptofonia apresentadas. Do ponto de vista da qualidade e inteligibilidade residual dos sinais, pode-se afirmar que as técnicas CSI-F(DCT) e CSI-F(BF) produzem resultados semelhantes, divergindo em valores inferiores a 1 dB.

Nas Simulações I e II, os resultados obtidos para as distâncias de Itakura e Cepstral apresentaram boa concordância com aqueles apresentados pela Referência [6]. É importante citar que quaisquer resultados acima de 4,0 dB para as distâncias de Itakura e Cepstral já refletem uma inteligibilidade residual adequada para sinais cifrados.

Para a avaliação de qualidade do sinal decifrado, foram obtidos valores de medidas objetivas que evidenciaram maior adequação das técnicas testadas para CODECs com taxas de codificação superiores a 7 kbps. Para os sinais decifrados, os valores adequados para as distâncias de Itakura e Cepstral devem estar situados abaixo de 1 dB. Para os

resultados obtidos por meio do algoritmo PESQ, os valores MOS acima de 2,5, após a audição, foram considerados bons, sendo aceitáveis valores MOS acima de $2,0$. Embora na escala MOS, os valores 2 e 3 sejam classificados como "pobre" e razoável, respectivamente, tem-se como premissa para este trabalho a não alteração de *hardware*/*software* dos sistemas de comunicações COTS; portanto, diante deste contingenciamento e da não disponibilidade de dispositivos comerciais que solucionem o problema de sigilo destes sistemas, a adoção de sinais com valores de MOS entre 2 e 3 torna-se uma solução aceitável.

A variação no número de sub-bandas de permutação, aqui limitadas a 8 e 16, não produziu resultados substancialmente diferentes do ponto de vista de qualidade, apontando para o uso preferencial de 16 faixas de permutação, podendo-se, quando a qualidade do enlace for degradada, reduzir o número de subfaixas para 8. O número de 16 subfaixas deve, sempre que possível, ser adotado, em decorrência do maior número de chaves disponíveis, o que implica maior resistência à criptoanálise.

As medidas indiretas de inteligibilidade apresentadas pelos gráficos das Figuras 5.3 e 5.5 denotam uma pequena superioridade da CSI-F(DCT). Este fato se dá devido às características de filtragem da DCT em relação ao banco de filtros. A característica menos seletiva da DCT produz um espectro cujas subfaixas adjacentes se interpõem, produzindo, desta forma, um espectro mais "misturado", o que resulta em maiores valores de medias espectrais. Por outro lado, a maior seletividade do banco de filtros permite prescindir de esquemas de sincronismo. Esta maior seletividade produz possui um aspecto indesejável: uma "assinatura" espectral, o que permite identificar o número de subfaixas utilizadas para permutação (ver Figura 2.4). Para minimizar este problema, pode-se diminuir a seletividade dos filtros-protótipo, o que acarreta uma maior susceptibilidade à perda de sincronismo. Portanto, uma relação de compromisso entre a seletividade dos filtros-protótipo e o contingenciamento do sincronismo deve ser estabelecida.

A Simulação III produziu resultados compatíveis com o esperado, especialmente no que se refere à diminuição da inteligibilidade residual. Este efeito, no entanto, é alcançado à custa de implementações complexas de esquemas de sincronismo de quadro,

pois, como as características espectrais do sinal mudam a cada quadro, as distorções provocadas pelo CODEC também se alteram nesta freqüência, provocando flutuações no sincronismo de quadro. Este fato explica a baixa qualidade do sinal recuperado para PTC=1. Para o experimento em questão, foram empregadas técnicas de sincronismo de quadro somente para os primeiros quadros do sinal.

Para contornar este problema, pode-se adotar um esquema de sincronismo local para troca de chaves em conjunto com a técnica de CSI-F(BF). O sincronismo local necessário à troca periódica de chaves pode ser alcançado com auxílio de GPS (*Global Position System*), que disponibiliza um sinal de tempo que é altamente preciso, da ordem de $1,5.10^{-8}$ s [34], dependendo do tipo do receptor utilizado esta precisão pode cair para 10^{-6} s.

O emprego da técnica de CSI-F(BF) com mudança periódica de chaves deve ser limitado às situações que permitam longos PTC, pois, como os filtros são elementos que possuem "memória", não é factível a mudança de chaves em períodos curtos.

A Simulação IV demonstrou a aplicabilidade da metodologia proposta na Seção 2.4, sendo importante relembrar que o método proposto é valido somente para inteligibilidade residual média, em consonância com o que ocorre para a distância de Hamming [14], [15]. Uma metodologia mais geral de seleção de chaves para criptofonia, que leva em consideração as localizações dos formantes da voz, é apresentada em [17].

Os resultados produzidos pela Simulação V são importantes para a correta seleção do PTC. Deve-se avaliar criteriosamente a utilização de baixos valores de PTC, pois, conforme demonstrado pelo gráfico da Figura 5.9, para pequenos valores de PTC a qualidade do sinal decifrado é inversamente proporcional ao valor do PTC e da taxa do empregada pelo CODEC. Valores de PTC baixos implicam inteligibilidades residuais baixas, mas para se garantir a recuperação do sinal codificado a baixas taxas de codificação, deve-se estabelecer um valor mínimo para o PTC, isto pode ser realizado por meio de uma relação de proporcionalidade inversa entre o PTC e a taxa operada pelo CODEC. Os efeitos de degradação do sinal decorrem das flutuações no sincronismo de quadro supramencionado.

Subjetivamente, após a audição dos inúmeros resultados das simulações, pode-se afirmar que os sinais resultantes possuem qualidade subjetiva compatível com a aplicação pretendida, i.e., o sinal cifrado é ininteligível, enquanto que o sinal decifrado possui boa inteligibilidade, principalmente para as taxas de codificação $7,95$ e $12,20$ kbps. Dentre as técnicas simuladas, a que apresentou melhor resultado subjetivo de qualidade foi a CSI-F(BF).

Capítulo 6

Conclusões e Sugestão para Trabalhos Futuros

6.1 Resumo e Principais Conclusões

O propósito desta dissertação foi apresentar uma solução para o problema de ausência de sigilo comum em equipamentos de comunicações móveis comerciais e cujas características de codificação do sinal de voz impedem a utilização de criptofonia digital. No Capítulo 1, foram apresentadas situações reais que justificam o presente estudo e a importância do sigilo nas comunicações móveis pessoais.

No Capítulo 2, foi realizada uma revisão das principais técnicas de criptofonia e estabelecidos os requisitos necessários aos sistemas de criptofonia adequados ao cumprimentos do propósito deste trabalho, quando ficou evidenciada possibilidade de emprego das técnicas de CSI-F. A seguir, buscou-se uma técnica simples de seleção de chaves para criptofonia, cujo resultado permite, por meio de simples operações de produto escalar, selecionar subconjuntos de matrizes de permutação capazes de gerar sinais ininteligíveis.

O Capítulo 3 abordou o problema do sincronismo em sistemas de criptofonia e a sua dificuldade de implementação. Nesse capítulo, foram apresentadas técnicas básicas para tratar o problema, bem como resultados de simulações que caracterizam o problema em questão e apontam algumas soluções para minimizar a perda de sincronismo. Os efeitos de atrasos e distorção causados pelo CODEC AMR, cujo resultado está diretamente ligado à perda de sincronismo, também foram discutidos.

Para avaliação objetiva de qualidade e inteligibilidade residual dos sinais decifrado e cifrado, respectivamente, no Capítulo 4 foram apresentadas duas classes de medidas

objetivas de avaliação: distâncias espectrais, que são medidas objetivas não-perceptuais; e o algoritmo PESQ, que leva em consideração as características perceptuais do som.

Os resultados constantes do Capítulo 5 foram obtidos de 5 simulações distintas, tendo as duas primeiras o propósito comparativo entre CSI-F com 8 e 16 subfaixas de permutação, todavia nunca deixando de realizar uma análise comparativa entre as técnicas CSI-F(BF) e CSI-F(CDT). As demais simulações procuraram dar enfoque a outros aspectos julgados importantes e abordados durante o desenvolvimento deste trabalho, tais como: escolha de chaves e mudança periódica de chaves. Por fim, foram discutidos e analisados os resultados das simulações supracitadas.

A seguir, são sumarizadas conclusões e sugestões para implementações de sistemas de criptofonia que visem a solucionar o problema em foco:

- A técnica CSI-F(BF), quando implementada com filtros adequados, é imune a atrasos sofridos pelo sinal, sendo, desta forma, também imune à perda de sincronismo;

- A técnica CSI-F(DCT), em decorrência das características de filtragem da DCT, é vulnerável a atrasos sofridos pelo sinal e, conseqüentemente, não prescinde de esquemas de sincronismo de amostra/quadro;

- As técnicas de CSI-F com chave fixa devem se limitar a aplicações cujo grau de sigilo requerido seja tático;

- Quando for requerido um grau de sigilo superior a tático, deve-se empregar técnicas de CSI-F com troca periódica de chaves;

- No sentido de se preservar parte das características espectrais do sinal de voz, deve-se limitar o número máximo de subfaixas utilizadas na permutação a 16 subfaixas; e

- As técnicas de CSI-T e CSI-Hadamard não são adequadas ao propósito deste estudo, pois resultam em grandes atrasos do sinal, além de necessitarem de esquemas precisos de sincronismo de amostra.

Com base nos resultados apresentados no Capítulo 5 e na audição dos arquivos resultantes das simulações, pode-se concluir que as técnicas de CSI-F são adequadas ao cumprimento do propósito deste trabalho. No intuito de se chegar à técnica mais aceitável para solução do problema, ou seja; a técnica que cumpre o propósito empregando o menor montante de recursos, elege-se a necessidade de implementação de esquemas de sincronismo como fator de desempate para o critério aceitabilidade. Desta forma, pode-se concluir que a técnica de CSI-F(BF) implementada por meio de componentes polifásicas é a técnica mais aceitável para solução do problema objeto deste trabalho.

6.2 Sugestões para Futuros Trabalhos

Como sugestões de trabalhos futuros na mesma linha de pesquisa desta Dissertação podem ser citados os seguintes tópicos:

a) Novas metodologias de seleção e geração automática de chaves para criptofonia baseada em modelos perceptivos;

b) Efeitos do canal GSM sobre sistemas de criptofonia;

c) Modelamento em banda-base do efeito de canal GSM aplicado a sistemas de criptofonia;

d) Transmissão *full-duplex* de voz codificada através do canal de dados dos sistemas GSM; e

e) Esquemas eficientes de sincronismo aplicados a sistemas de criptofonia com troca periódica de chaves.

Referências Bibliográficas

[1] PREVELAKIS, V., SPINELLIS, D., "The Athens Affair", *IEEE Spectrum Magazine*, v. 44, n. 7, pp. 26–33, July 2007.

[2] DELLER, J. R., PROAKIS, J. G., HANSEN, J. H. L., *Discrete-Time Processing of Speech Signals*. New York, USA, Macmilan, 1993.

[3] ANDRADE Jr., J. F., CAMPOS, M. L. R., APOLINÁRIO Jr., J. A., "Speech privacy for modern mobile communication systems". In: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2008)*, pp. 1777–1780, Nevada, USA, April 2008.

[4] "Technical Specification 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Mandatory speech CODEC speech processing functions; AMR speech CODEC", 3GPP TS 26.071 V6.0.0 (2004-12), 2004.

[5] APOLINÁRIO JR., J., *Criptoanálise de Sinais de Voz Cifrados por Permutação de Segmentos Temporais*. Tese de M.Sc., Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília-DF, Brasil, Junho 1993.

[6] GOLDBURG, B., SRIDHARAN, S., "Design and cryptanalysis of transform-based scramblers", *IEEE Jornal on Selected Areas on Communications*, v. 11, n. 5, pp. 735–744, June 1993.

[7] JAYANT, N., B. MCDERMOTT, S. C., QUINN, A., "A comparison of four methods for analog speech privacy", *IEEE Transactions on Communications*, v. COM-29, n. 1, pp. 18–23, July 1981.

[8] BEKER, H. J., PIPER, F. C., *Secure Speech Communications*. London, UK, Academic Press, 1985.

[9] SENK, V., V. D. DELIC, V. S. M., "A new speech scrambling concept based on Hadamard matrices", *IEEE Signal Processing Letters*, v. 4, n. 6, pp. 161–163, June 1997.

[10] LEE, L. S., G .C. CHOU, C. S. C., "New frequency domain speech scrambling system which does not require frame synchronization", *IEEE Transaction Communication*, v. COM-32, n. 4, pp. 444–456, April 1984.

[11] EHSANI, M. S., BOROUJENY, S. E., "Fast Fourier transform speech scrambler", *IEEE First International Symposium Intelligent Systems*, pp. 248–251, September 2002.

[12] DINIZ, P. S. R., da SILVA, E. A. B., NETTO, S. L., *Digital Signal Processing: System Analysis and Design*. Cambridge, UK, Cambridge University Press, 2002.

[13] MILOSEVIC, V. S., V. D. DELIC, V. S., "Hadamard Transform application in speech scrambling". In: *13th International Conference on Digital Signal Processing Proceedings, DSP 97*, v. 1, pp. 361–363, Santorini, Greece, July 1997.

[14] WOO, R. W., LEUNG, C., "A new key generation method for frequency-domain speech scramblers", *IEEE Transactions on Communications*, v. 45, n. 7, pp. 749–752, July 1997.

[15] SAKURAI, K., KOGA, K., MURATANI, T., "A speech scrambler using the fast Fourier transform technique", *IEEE Journal on Selected Areas in Communications*, v. 2, n. 3, pp. 434 – 442, May 1984.

[16] MATSUNAGA, A., KOGA, K., OHKAWA, M., "An analog speech scrambling system using the FFT technique with high-hevel security", *IEEE Journal on Selected Areas in Communications*, v. 7, n. 4, pp. 540–547, May 1989.

[17] BORZINO, A. M. C. R., APOLINÁRIO Jr., J. A., da SILVA, D. G., "An efficient objective intelligibility measure for frequency domain scramblers", *EURASIP Journal on Information Security*, v. 2007, n. 32028, 2007.

[18] PEEBLES, P. Z., *Probability, Random Variables and Random Signals Principles*. Fouth ed. , New York, USA, McGraw-Hill, 2000.

[19] MASSEY, J. L., "Optimum frame synchronization", *IEEE Transactions on Communications*, v. 20, pp. 115–119, April 1972.

[20] BUMILLER, G., LAMPE, L., "Fast burst synchronization for power line communication systems", *EURASIP Journal on Advances in Signal Processing*, v. 2007, n. 1, pp. 166–166, 2007.

[21] BARKER, R. H., "Group synchronization of binary digital systems", *in Communication Theory*, pp. 273–287, 1953.

[22] NEUMAN, F., HOFMAN, L., "New pulse sequences with desirable correlation properties". In: *IEEE National Telemetry Conference (NTC '71)*, pp. 272–282, Washington, USA, April 1971.

[23] PROAKIS, J., *Digital Communications*. Fouth ed. , New York, USA, McGraw-Hill, 2001.

[24] "Technical Specification 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Digital cellular telecommunications system (Phase 2+); Enhanced Full Rate (EFR) speech transcoding", GSM 06.60 version 8.0.1 (2009-11), 2000.

[25] "Technical Specification 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Digital cellular telecommunications system (Phase 2+); Half Rate speech transcoding", 3GPP TS 06.20 V8.0.1 (2000-11), 2000.

[26] HAYKIN, S., *Adaptive Filter Theory*. Fouth ed. , New Jersey, USA, Prentice Hall, 2002.

[27] ITAKURA, F., "Minimum prediction residual principle applied to speech recognition", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. ASSP-23, n. 1, pp. 67–72, February 1975.

[28] "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs. ITU-T Recommendation P.862", International Telecommunication Union (ITU-T), 2001.

[29] "Objective quality measurement of telephone band (300 - 3400 Hz) speech codecs. ITU-T Recommendation P.861", International Telecommunication Union (ITU-T), 1996.

[30] BOSI, M., GOLDBERG, R. E., *Introduction to Digital Audio Coding and Standards.* , Norwell, USA, Kluwer, 2002.

[31] "Mean Opinion Score (MOS) terminology P.800.1", International Telecommunication Union (ITU-T), 2003.

[32] "Application guide for objective quality measurement based on Recommendations P.862, P.862.1 and P.862.2 ITU-T Recommendation P.862.3", International Telecommunication Union (ITU-T), 2005.

[33] ANDRADE Jr., J. F., CAMPOS, M. L. R., APOLINÁRIO Jr., J. A., "Sistemas de Criptofonia sob Influência de Canais de Comunicações Móveis", *XXVI Simpósio Brasileiro de Telecomunicações (SBrT'08)*, pp. 1–5, Setembro 2008.

[34] PETOVELLO, M. G., LACHAPELLE, G., "Estimation of Clock Stability Using GPS", *GPS Solutions*, v. 4, n. 1, pp. 21–33, July 2000.

[35] MITRA, S. K., *Digital Signal Processing: A computer Based Approach.* New York, USA, Mcgraw-Hill, 1998.

[36] VAIDYANATHAN, P. P., *Mulitrate Systems and Filter Banks.* New York, USA, Prentice-Hall, 1993.

Apêndice A

Bancos de Filtros de DFT Uniforme

A.1 Introdução

Banco de filtros de DFT uniforme é a denominação dada a uma implementação eficiente de bancos de filtros digitais, cujo detalhamento será desenvolvido no decorrer deste Apêndice. Genericamente, os bancos de filtros digitais são formados por conjuntos de filtros digitais do tipo passa-faixa, cujo propósito é permitir oprocessamento do sinal em M subfaixas (ou M sub-bandas) separadas. Os bancos de filtros podem ser classificados como bancos de filtros de análise e bancos de filtros de síntese.

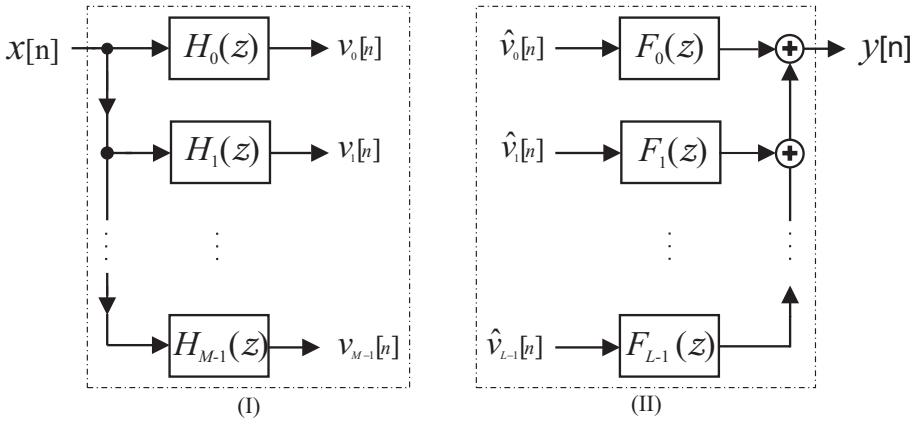

Figura A.1: (I) Banco de filtros de análise e (II) Banco de filtros de síntese.

Os bancos de filtros de análise são empregados para decompor o sinal $x[n]$ em um conjunto de M subfaixas, conforme mostrado pela Fig. A.1. Como resultado da filtragem realizada por cada filtro $H_k(z)$ tem-se a componente $\nu_k[n]$, que representa a k-ésima subfaixa do espectro do sinal original.

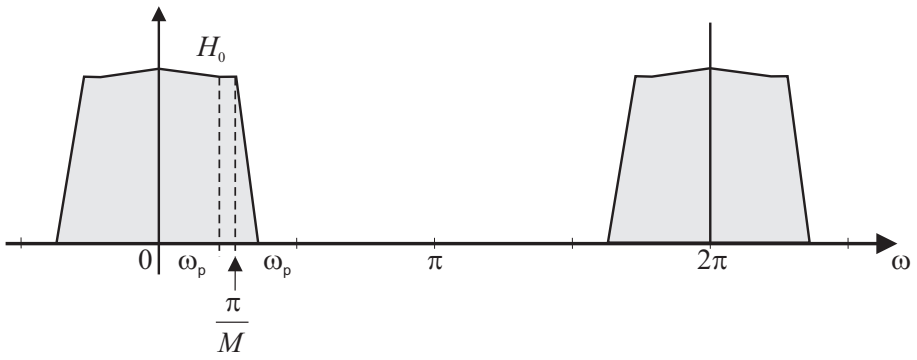

Figura A.2: Resposta em freqüência do filtro protótipo $H_0(z)$.

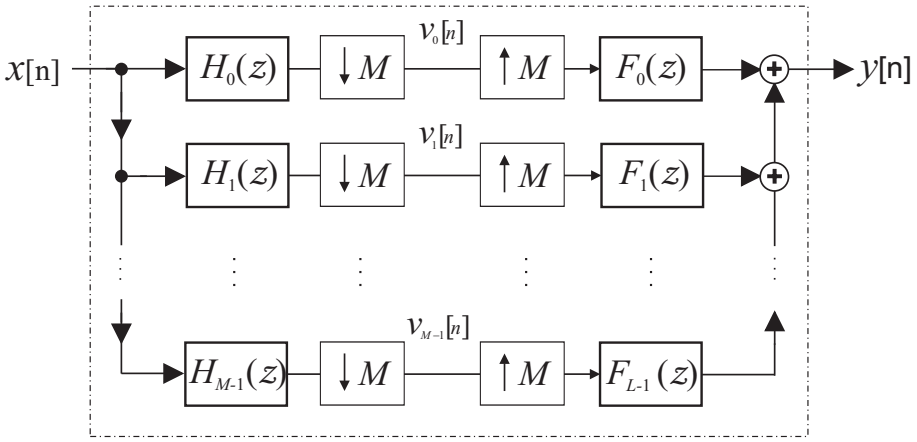

Figura A.3: Diagrama de banco de filtros com M subfaixas.

Conforme pode ser observado na Fig. A.1, o número de componentes do sinal resultante é expandido por um fator M, que, em muitos casos, provoca um indesejável aumento na banda do sinal. Para solucionar este aumento do número total de amostras do sinal, pode-se realizar a decimação[1] [12] de cada componente $\nu_k[n]$ ($k = 0, 1, \ldots, M-1$) por um fator igual a M.

O conjunto de filtros destinados à recuperação do sinal, a partir das M subfaixas $\hat{\nu}_k[n]$, denominado banco de síntese, efetua a operação de interpolação [12] por um

[1]Se a decimação ocorrer por um fator igual ao número de faixas do banco de filtros, diz-se que o banco é criticamente decimado.

fator[2] L, filtra cada subfaixa e adiciona as componentes resultantes, produzindo o sinal restaurado $y[n]$.

Se o sinal de entrada pode ser recuperado completamente a partir de suas M subfaixas, a estrutura é chamada de banco de filtros com reconstrução perfeita de M subfaixas.

A.2 Bancos de Filtros de DFT Uniforme

Seja um filtro-protótipo passa-baixas $H_0(z)$, com resposta ao impulso igual a $h_0[n]$. Sem perda de generalidade, pode-se representar $H_0(z)$ como sendo um filtro digital causal FIR, do tipo:

$$H_0\left(z\right) = \sum_{n=0}^{\infty} h_0\left[n\right] z^{-n}. \tag{A.1}$$

Conforme mostrado na Fig. A.2, a faixa de passagem e a freqüência de corte de $H_0(z)$ são representadas por ω_p e ω_s, respectivamente. Então, com o propósito de determinar a função de transferência dos filtros das M-1 subfaixas restantes, em função de $H_0(z)$, pode-se definir $h_k(z)$ como:

$$h_k\left(z\right) = h_0\left(z\right) \mathrm{e}^{j\left(\frac{2kn\pi}{M}\right)} \;\;,\; k = 0, 1, \ldots, M-1. \tag{A.2}$$

Para simplificar a notação e torná-la similar à definição de DFT [12], pode-se fazer $\mathrm{e}^{-j\left(\frac{2\pi}{M}\right)} = W_M$, desta forma:

$$h_k\left(z\right) = h_0\left(z\right) W_M^{-kn} \;\;,\; k = 0, 1, \ldots, M-1, \tag{A.3}$$

que no domínio da transformada $\mathscr{Z}$ é representada como:

$$\begin{aligned} H_k(z) = \sum_{k=0}^{\infty} h_k\left[n\right] z^{-n} = \sum_{n=0}^{\infty} h_0\left[n\right] \left(zW_M^k\right)^{-n} = \\ H_0(zW_M^k) \;\;,\; k = 0, 1, \ldots, M-1. \end{aligned} \tag{A.4}$$

A resposta em freqüência de $H_k(z)$ pode ser obtida substituindo-se o valor de z por $e^{j\omega}$ na Equação (A.4).

$$H_k(z = e^{j\omega}) = H_0\big(e^{j\left[\omega - \frac{2k\pi}{M}\right]}\big). \tag{A.5}$$

[2]Quando não ocorre mudança na taxa de amostragem, entre entrada e saída, tem-se $L = M$.

De maneira prática, a resposta em freqüência de $H_k(e^{j\omega})$ é obtida pelo deslocamento da resposta de $H_0(z = e^{j\omega})$ de um valor igual a $\frac{2k\pi}{M}$. Em decorrência de a resposta em módulo de $H_k(z)$ ser igual à resposta em módulo de $H_0(z)$ deslocada no espectro, este tipo de banco de filtros é denominado uniforme.

$$\left|H_k(e^{j\omega})\right| = \left|H_0\big(e^{j[\omega-\frac{2k\pi}{M}]}\big)\right|. \tag{A.6}$$

A denominação DFT advém do fato de que o deslocamento em freqüência, referente às subfaixas, é realizado pela multiplicação da resposta ao impulso do filtro protótipo $h_0[n]$ pelos elementos que formam a matriz DFT [35], aqui representados por W_M^{-kn}.

Embora as funções de transferência apresentadas utilizem somente a notação referente aos filtros de análise, este desenvolvimento pode, indistintamente, ser aplicado ao banco de síntese.

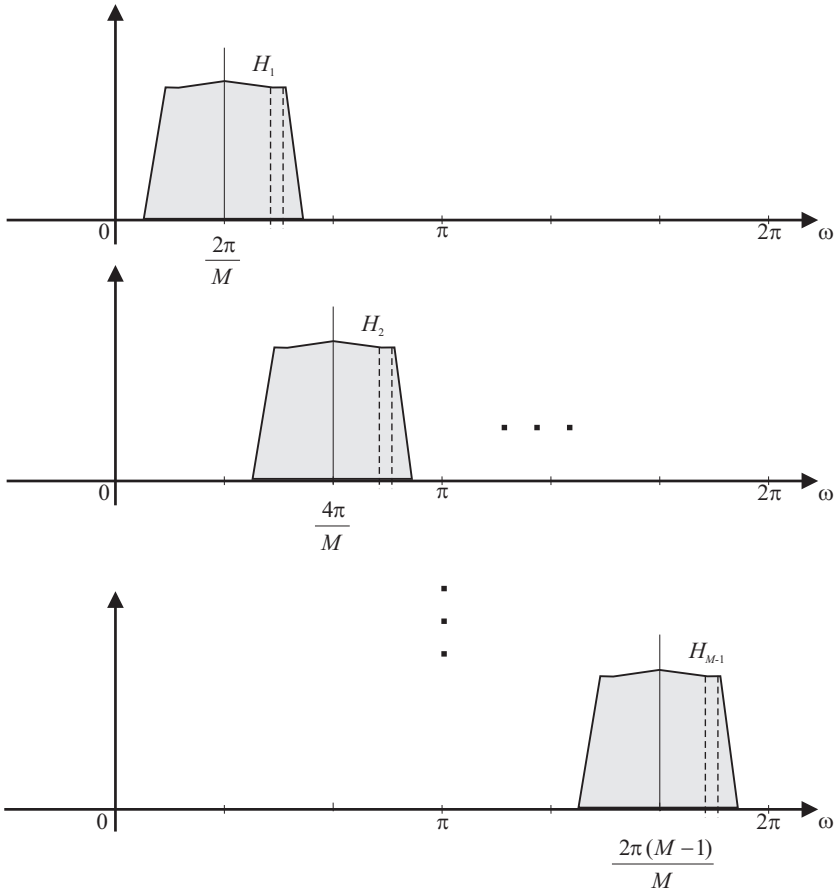

Figura A.4: Resposta em freqüência de banco de filtros com M faixas distribuídas uniformente ($H_k(z)$, $k = 0, \ldots, M-1$).

A.2.1 Implementação em termos de Componentes Polifásicas

A implementação de bancos de filtros de DFT uniforme, por meio de componentes polifásicas, se dá pela substituição dos filtros de análise e síntese por suas respectivas componentes polifásicas [36].

Usualmente, os filtros de análise e síntese são substituídos por componentes polifásicas Tipo I e Tipo II, respectivamente. Esta forma de implementação visa ao aumento da eficiência, pois apresenta menor complexidade computacional em relação à implementação convencional (forma direta).

As funções de transferência dos filtros-protótipo de análise e síntese ($H_0(z)$ e $F_0(z)$) pertencentes a um banco de filtros com M bandas podem ser escritas em função de suas componentes polifásicas Tipo I e Tipo II, respectivamente:

$$H_0(z) = \sum_{l=0}^{M-1} z^{-l} E_l\big(z^M\big) \quad , \ (l = 0, 1, \ldots, M-1) \tag{A.7}$$

$$F_0(z) = \sum_{l=0}^{M-1} z^{-(M-1-l)} R_l\big(z^M\big) \quad , \ (l = 0, 1, \ldots, M-1), \tag{A.8}$$

onde as componentes polifásicas são definidas como:

$$E_l(z) = \sum_{n=0}^{\infty} h_0[nM + l] z^{-n}, \quad (l = 0, 1, \ldots, M-1) \tag{A.9}$$

$$R_l(z) = E_{M-1-l}(z) = \sum_{n=0}^{\infty} h_0[nM + (M-1-l)] z^{-n}, \quad (l = 0, 1, \ldots, M-1) \tag{A.10}$$

A decomposição polifásica para os demais filtros pode ser obtida com base na Equação (A.4), substituindo-se o valor de z nas Equações (A.7) e (A.8) por zW_M^k.

$$H_k(z) = \sum_{l=0}^{M-1} z^{-l} W_M^{-kl} E_l\big(z^M W_M^{kM}\big) = \sum_{l=0}^{M-1} z^{-l} W_M^{-kl} E_l\big(z^M\big), \quad (k = 0, 1, \ldots, M-1) \tag{A.11}$$

$$F_k(z) = \sum_{l=0}^{M-1} \big(zW_M^k\big)^{-(M-1-l)} R_l\big(z^M\big) \quad , \ (k = 0, 1, \ldots, M-1) \tag{A.12}$$

As Equações (A.11) e (A.12) podem ser reescritas na forma matricial como:

$$H_k(z) = \begin{bmatrix} 1 & W_M^{-k} & W_M^{-2k} & \cdots & W_M^{-(M-1)k} \end{bmatrix} \begin{bmatrix} E_0(z^M) \\ z^{-1}E_1(z^M) \\ z^{-2}E_2(z^M) \\ \vdots \\ z^{-(M-1)}E_{M-1}(z^M) \end{bmatrix} \tag{A.13}$$

$$F_k(z) = \begin{bmatrix} 1 & W_M^{k} & W_M^{2k} & \cdots & W_M^{(M-1)k} \end{bmatrix} \begin{bmatrix} R_{M-1}(z^M) \\ z^{-1}R_{M-2}(z^M) \\ z^{-2}R_{M-3}(z^M) \\ \vdots \\ z^{-(M-1)}R_0(z^M) \end{bmatrix}. \tag{A.14}$$

Para $(k = 0, 1, \ldots, M-1)$, as Equações (A.13) e (A.14) transformam-se em:

$$\begin{bmatrix} H_0(z) \\ H_1(z) \\ H_2(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_M^{-1} & W_M^{-2} & \cdots & W_M^{-(M-1)} \\ 1 & W_M^{-2} & W_M^{-4} & \cdots & W_M^{-2(M-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_M^{-(M-1)} & W_M^{-2(M-1)} & \cdots & W_M^{-(M-1)^2} \end{bmatrix} \begin{bmatrix} E_0(z^M) \\ z^{-1}E_1(z^M) \\ z^{-2}E_2(z^M) \\ \vdots \\ z^{-(M-1)}E_{M-1}(z^M) \end{bmatrix} \tag{A.15}$$

$$\begin{bmatrix} F_0(z) \\ F_1(z) \\ F_2(z) \\ \vdots \\ F_{M-1}(z) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_M^{1} & W_M^{2} & \cdots & W_M^{(M-1)} \\ 1 & W_M^{2} & W_M^{4} & \cdots & W_M^{2(M-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_M^{(M-1)} & W_M^{2(M-1)} & \cdots & W_M^{(M-1)^2} \end{bmatrix} \begin{bmatrix} R_{M-1}(z^M) \\ z^{-1}R_{M-2}(z^M) \\ z^{-2}R_{M-3}(z^M) \\ \vdots \\ z^{-(M-1)}R_0(z^M) \end{bmatrix}. \tag{A.16}$$

Em termos de matriz da DFT, tem-se:

$$\begin{bmatrix} H_0(z) \\ H_1(z) \\ H_2(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} = M\boldsymbol{D}_M^{-1} \begin{bmatrix} E_0(z^M) \\ z^{-1}E_1(z^M) \\ z^{-2}E_2(z^M) \\ \vdots \\ z^{-(M-1)}E_{M-1}(z^M) \end{bmatrix} \tag{A.17}$$

$$\begin{bmatrix} F_0(z) \\ F_1(z) \\ F_2(z) \\ \vdots \\ F_{M-1}(z) \end{bmatrix} = \boldsymbol{D}_M \begin{bmatrix} R_{M-1}(z^M) \\ z^{-1}R_{M-2}(z^M) \\ z^{-2}R_{M-3}(z^M) \\ \vdots \\ z^{-(M-1)}R_0(z^M) \end{bmatrix}, \tag{A.18}$$

onde $\boldsymbol{D}_M$ é matriz da DFT de ordem M:

$$\boldsymbol{D}_M = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_M^1 & W_M^2 & \cdots & W_M^{(M-1)} \\ 1 & W_M^2 & W_M^4 & \cdots & W_M^{2(M-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_M^{(M-1)} & W_M^{2(M-1)} & \cdots & W_M^{(M-1)^2} \end{bmatrix}. \tag{A.19}$$

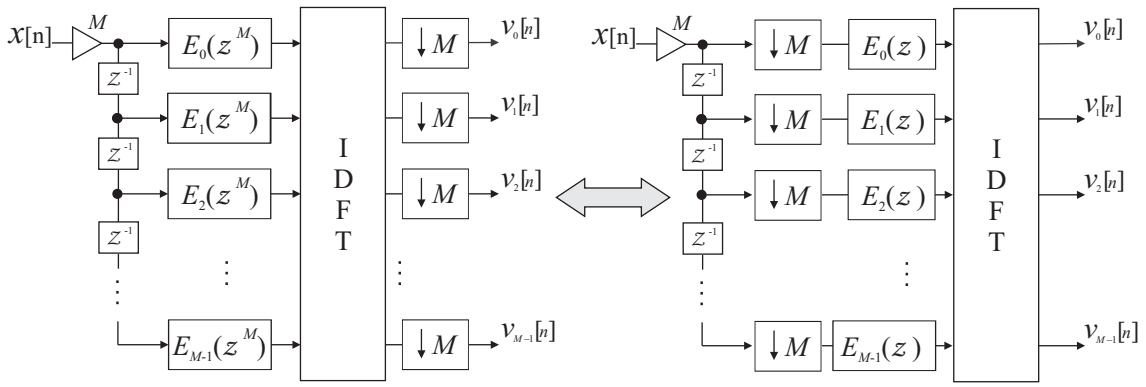

Figura A.5: Implementação de banco de análise utilizando decomposição polifásica, onde $H_k(z) = \frac{V_k(z)}{X_k(z)}$.

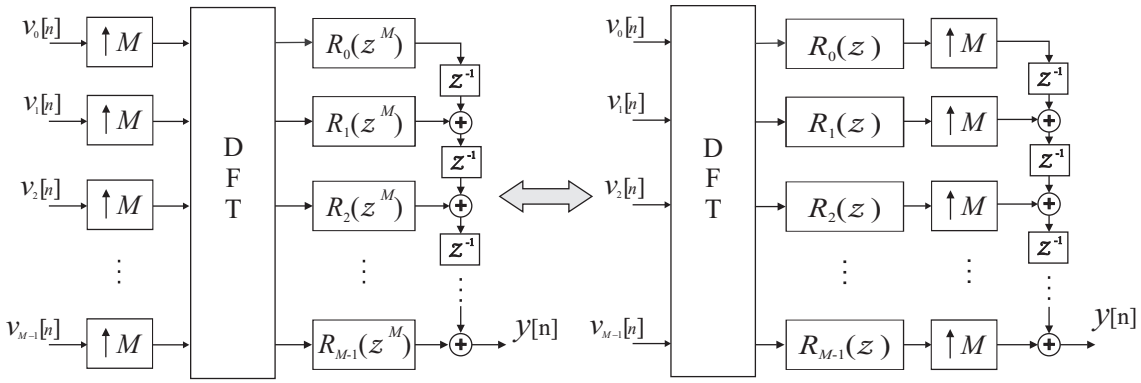

Figura A.6: Implementação de banco de síntese utilizando decomposição polifásica, onde $F_k(z) = \frac{Y_k(z)}{V_k(z)}$.

A complexidade computacional das implementações mostradas na Figuras A.5 e A.6 é bem inferior àquela necessária para implementação direta (ver Fig. A.2), que, para um banco de análise com M subfaixas e filtro passa-baixas de ordem N, possui um número de multiplicações da ordem $N \times M$. Quando se emprega a decomposição polifásica, são necessárias $N + \frac{M}{2}\log_2(M)$ multiplicações, sendo N multiplicações para os M filtros e $\frac{M}{2}\log_2(M)$ multiplicações para o cálculo da DFT com M pontos, o que denota a eficiência muito superior deste método (ver Figura. A.7).

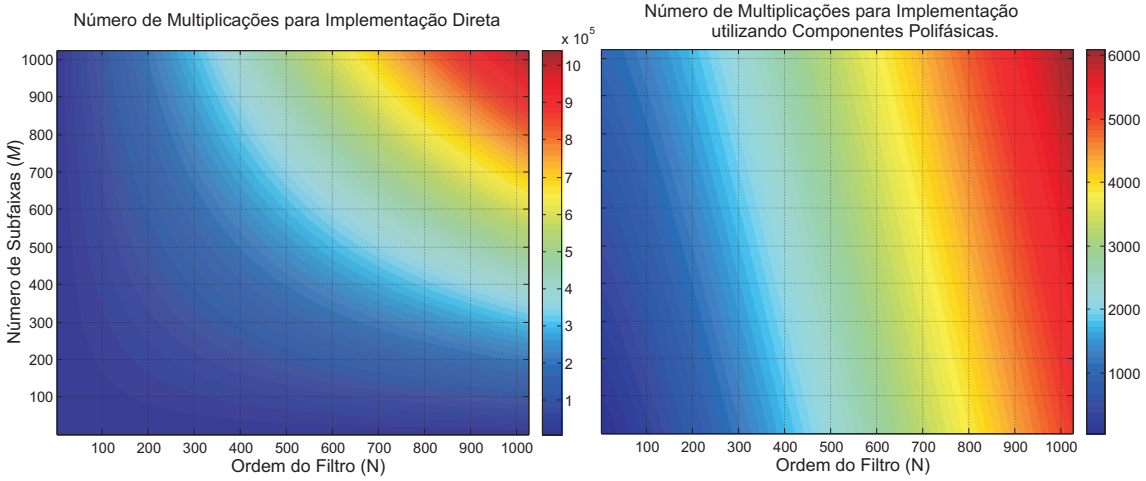

Figura A.7: Complexidade computacional das implementações apresentadas.

Apêndice B

Cálculo do Ângulo Máximo Φ_I^{Max}

Para cada tamanho de chave N, existe um valor máximo Φ_I^{Max} decorrente da aplicação da matriz de permutação $\boldsymbol{P}^{90^{\mathrm{o}}}$. Por definição, esta é uma matriz diagonal secundária de norma unitária:

$$\boldsymbol{P}^{Max} = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \iddots & 1 & 0 \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ 0 & 1 & \iddots & \iddots & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}_{N\times N} . \tag{B.1}$$

Aplicando-se a permutação ao vetor $\boldsymbol{V}_N = [1 \;\; 2 \cdots N]^T_{(N\times 1)}$, obtém-se:

$$\boldsymbol{V}_N^{Max} = \boldsymbol{P}^{Max}\boldsymbol{V}_N = \begin{bmatrix} N & (N-1) & \cdots & 2 & 1 \end{bmatrix}^T . \tag{B.2}$$

Para o cálculo do ângulo:

$$\Phi_I^{\mathrm{Max}} = \arccos\left\{ \frac{(\boldsymbol{V}_N^{Max})^T \boldsymbol{V}_N}{\|\boldsymbol{V}_N\|^2} \right\}, \tag{B.3}$$

faz-se necessário determinar o produto escalar $\boldsymbol{V}_N^{Max} \cdot \boldsymbol{V}_N$:

$$\begin{aligned} \boldsymbol{V}_N^{Max} \cdot \boldsymbol{V}_N &= \Big\{ N + 2(N-1) + 3(N-2) + \ldots + N[N-(N-1)] \Big\} \\ &= N\sum_{k=1}^{N} k - \left\{ \sum_{k=1}^{N} k^2 - \sum_{k=1}^{N} k \right\} = (N+1)\sum_{k=1}^{N} k - \sum_{k=1}^{N} k^2. \end{aligned} \tag{B.4}$$

Os somatórios constantes da Equação (B.4) são duas séries conhecidas: série aritmética[1]

[1] $S_N = \sum_{k=1}^{N} k = \frac{N(N+1)}{2}$

e série quadrática[2].

$$\boldsymbol{V}_N^{Max} \cdot \boldsymbol{V}_N = (N+1)\left\{\frac{N(N+1)}{2}\right\} - \frac{N(N+1)}{6}\left\{2(N+1)-1\right\} \tag{B.5}$$

$$= \left\{\frac{N(N+1)}{6}\right\}(N+2). \tag{B.6}$$

O quadrado da norma de $\boldsymbol{V}_N$ dado por:

$$\|\boldsymbol{V}_N\|^2 = 1 + 2^2 + 3^2 + \ldots + N^2 = \left\{\frac{N(N+1)}{6}\right\}(2N+1). \tag{B.7}$$

Substituindo os resultados das Equações (B.6) e (B.7) na Equação (B.3), obtém-se o valor para Φ_I^{Max} em função de N:

$$\Phi_I^{\mathrm{Max}}(N) = \arccos\left\{\frac{N+2}{2N+1}\right\}. \tag{B.8}$$

Em graus:

$$\boxed{\Phi_I^{\mathrm{Max}}(N) = \frac{180}{\pi}\arccos\left\{\frac{N+2}{2N+1}\right\}}. \tag{B.9}$$

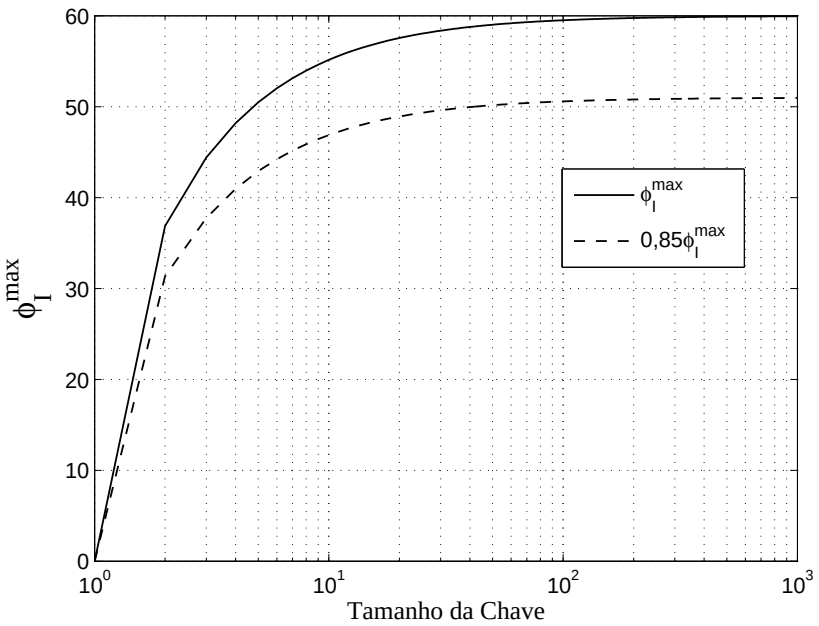

Figura B.1: Valores de Φ_I^{Max} em função de N.

[2] $S_N = \sum_{k=1}^{N} k^2 = \frac{N(N+1)}{6}\left\{2(N+1)-1\right\}$

Apêndice C

Análise, Geração e Detecção de Sinais FSK

C.1 Análise e Geração de Sinais FSK

Em princípio, a modulação FSK pode ser obtida pela aplicação direta do sinal digital, com a banda de freqüência limitada, na entrada de um *Voltage-Controlled Oscillator* (VCO), conforme diagrama da Figura C.1. As variações de amplitude do sinal digital forçam o VCO a variar a sua freqüência entre valores discretos, que podem ser determinados com auxílio da curva de conversão freqüência-tensão do VCO.

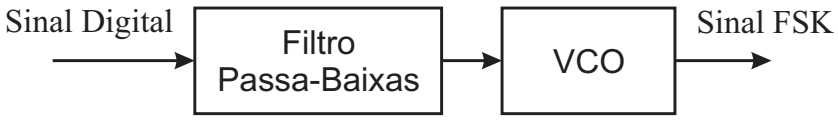

Figura C.1: Geração do Sinal FSK.

Com o emprego de técnicas de Processamento Digital de Sinais, a geração de sinais FSK pode ser realizada por meio de operações matemáticas; esta será a abordagem adotada no decorrer desta seção.

A seguir será apresentada a análise matemática para uma modulação FSK, cujo sinal modulante possui apenas dois níveis: **0** e **1**, ou seja, 2-FSK ou BFSK (*Binary* FSK). O sinal BFSK admite duas freqüências e, portanto, as formas de onda correspondentes aos estados **0** e **1** podem ser escritas como:

$$E^0_{\mathtt{FSK}}(t) = E_0 \cos(\omega_1 t) \tag{C.1}$$

$$E^1_{\texttt{FSK}}(t) = E_0 \cos(\omega_2 t) \tag{C.2}$$

Considerando $\omega_2 > \omega_1$, pode-se, então, definir a portadora virtual ω_0 e o desvio ω_d como:

$$\omega_0 = \frac{\omega_1 + \omega_2}{2} \tag{C.3}$$

$$\omega_d = \frac{\omega_2 - \omega_1}{2} \tag{C.4}$$

Para um sinal modulador representado por uma onda quadrada de com período T e ciclo de trabalho de 50%, o sinal modulado pode ser considerado como uma composição de dois sinais OOK (*On/Off Key*)[1]. Esta consideração será útil na obtenção do espectro de freqüências do sinal modulado.

O sinal resultante pode ser considerado como a superposição linear do sinal OOK_1 com o sinal OOK_2.

O sinal OOK_1 fornece:

Estado 1: 0

Estado 0: $E_m(t) = E_0 \cos(\omega_1 t)$

O espectro de $E^0_m(t)$ pode ser expresso como:

$$E^0_m(t) = E_0 \sum_{n=-\infty}^{+\infty} \left\{ \frac{\sin(\frac{n\pi}{2})}{\frac{n\pi}{2}} \right\} e^{j(\omega_1 + n\omega)t}. \tag{C.5}$$

O sinal OOK_2 fornece:

Estado 1: $E_m(t) = E_0 \cos(\omega_2 t)$

Estado 0: 0

O sinal OOK_2 possui um retardo de $\frac{T}{2}$ em relação ao sinal OOK_1, então o espectro resultante pode ser escrito como:

$$E^1_m(t) = E_0 \sum_{n=-\infty}^{+\infty} e^{-jn\omega(\frac{T}{2})} \left\{ \frac{\sin(\frac{n\pi}{2})}{\frac{n\pi}{2}} \right\} e^{j(\omega_2 + n\omega)t}. \tag{C.6}$$

[1]A modulação OOK (*On/Off Key*) é um caso especial da modulação ASK (*Amplitude Shift Key*), cujas transmissões dos símbolos **0** são representadas pela ausência de portadora.

$$E_m(t) = E_m^0(t) + E_m^1(t) = E_0 \sum_{n=-\infty}^{+\infty} \left\{ \frac{\sin(\frac{n\pi}{2})}{\frac{n\pi}{2}} \right\} e^{j(\omega_1 + n\omega)t} + E_0 \sum_{n=-\infty}^{+\infty} e^{-jn\omega(\frac{T}{2})} \left\{ \frac{\sin(\frac{n\pi}{2})}{\frac{n\pi}{2}} \right\} e^{j(\omega_2 + n\omega)t}. \tag{C.7}$$

Considerando a ortogonalidade do sinal BFSK, só se faz necessária a detecção de um sinal OOK de cada vez; desta forma, é possível determinar a largura de banda necessária à transmissão deixando-se passar pelo menos cada subportadora e as suas respectivas raias adjacentes.

$$BW = 2\omega_d + 2\omega = 2(\omega_d + \omega). \tag{C.8}$$

A largura de banda de um sinal FSK também pode ser calculada em função da taxa de transmissão e da separação existente entre as freqüências correspondentes aos estados **0** e **1**:

$$BW = V_m(1 + r) + \frac{(\omega_2 - \omega_1)}{2\pi}, \tag{C.9}$$

onde:

V_m é a velocidade de transmissão em bits por segundo (bps);

r é o fator de filtragem do filtro formatador de pulso, cuja função é suavizar a transição inter-pulsos;

ω_2 é a freqüência angular referente ao símbolo **1**; e

ω_1 é a freqüência angular referente ao símbolo **0**.

O desvio de freqüência utilizado, que é a diferença entre as freqüências correspondentes aos estados **0** e **1**, está relacionado com a velocidade de transmissão. Normalmente, usa-se um desvio de freqüência limitado pela metade e o dobro da velocidade de transmissão, em bps. Por exemplo: para uma velocidade de 1 kbps, pode-se utilizar um valor para o desvio entre 500 Hz e 2 kHz. Quanto maior o desvio, maior será a imunidade a ruídos; em contrapartida, haverá um aumento na largura de banda do sinal resultante.

C.2 Detecção Ótima de Sinais FSK.

Detectores que realizam detecção ótima [23] de sinais se baseiam nos valores das funções de distribuição de probabilidades condicionais [18]. Estas distribuições, também

conhecidas como probabilidades *a posteriori*, expressam matematicamente a probabilidade de se receber um determinado conjunto de símbolos $\boldsymbol{S}_m$, dado que na saída dos correlatores[2] (Figura C.2) tem-se o vetor $\boldsymbol{r}$. Para o caso especial da modulação BFSK, o vetor $\boldsymbol{r}$ é definido como $\boldsymbol{r} = [r_{1I}, r_{1Q}, r_{2I}, r_{2Q}]$, com $\boldsymbol{r}_1 = r_{1I} + \boldsymbol{j} r_{1Q}$ e $\boldsymbol{r}_2 = r_{2I} + \boldsymbol{j} r_{2Q}$, onde os sub-índices I e Q significam "em fase" e "em quadratura", respectivamente.

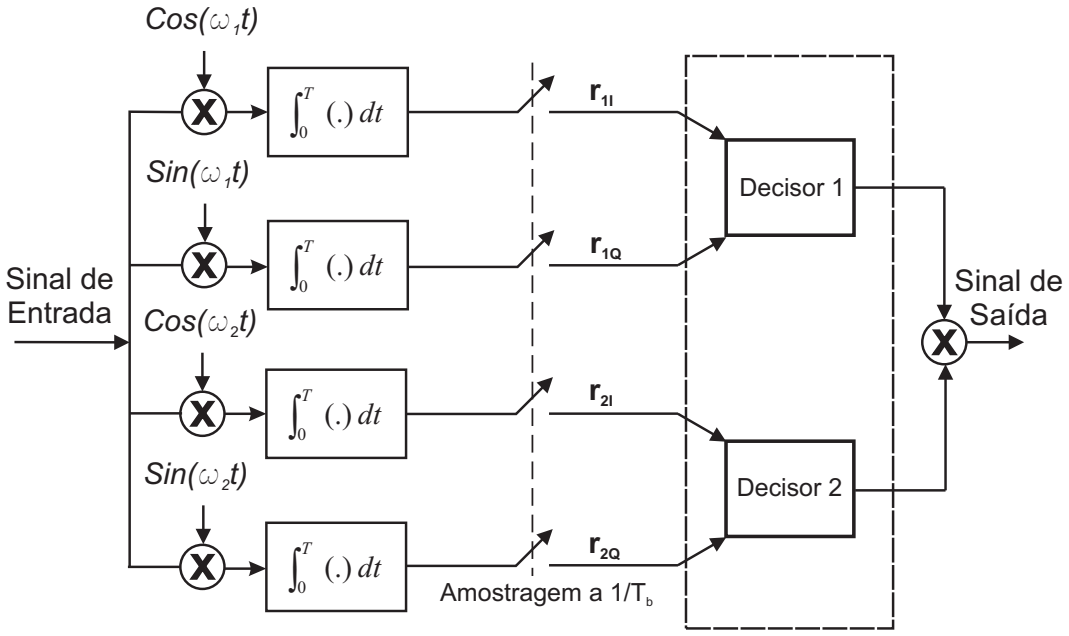

Figura C.2: Detector Ótimo para sinais BFSK.

A decisão de qual símbolo é recebido é baseada na expressão da probabilidade *a posteriori*:

$$p(\boldsymbol{S}_m/\boldsymbol{r}) = \frac{p(\boldsymbol{r}/\boldsymbol{S}_m)p(\boldsymbol{S}_m)}{p(\boldsymbol{r})} \ , \ m = 1, 2 \ , \tag{C.10}$$

$$\frac{p(\boldsymbol{S}_1/\boldsymbol{r})}{p(\boldsymbol{S}_2/\boldsymbol{r})} \underset{\boldsymbol{S}_2}{\overset{\boldsymbol{S}_1}{\gtrless}} \frac{p(\boldsymbol{S}_1)}{p(\boldsymbol{S}_2)} \ . \tag{C.11}$$

A Função Densidade de Probabilidade (PDF) $p(\boldsymbol{r}/\boldsymbol{S}_m)$ para uma portadora com fase aleatória ϕ pode ser escrita em termos da PDF marginal:

$$p_{\boldsymbol{r},\boldsymbol{S}_m}(\boldsymbol{r}/\boldsymbol{S}_m) = \int_0^{2\pi} p_{\boldsymbol{r},\boldsymbol{S}_m,\phi}(\boldsymbol{r}/\boldsymbol{S}_m, \phi) \ d\phi \ . \tag{C.12}$$

[2]Em geral, os detectores ótimos são implementados com filtros casados; contudo, de maneira alternativa, pode-se fazer uso de correlatores para se alcançar resultados semelhantes [23].

Para o caso especial do sinal BFSK, as saídas dos correlatores (Figura C.2) são:

$$\boldsymbol{r}_1 = r_{1I} + \boldsymbol{j} r_{1Q} = 2\varepsilon\cos(\phi) + n_{1I} + \boldsymbol{j}[2\varepsilon\sin(\phi) + n_{1Q}] \ , \tag{C.13}$$

$$\boldsymbol{r}_2 = r_{2I} + \boldsymbol{j} r_{2Q} = 2\varepsilon|\rho|\cos(\phi) + n_{2I} + \boldsymbol{j}[2\varepsilon|\rho|\sin(\phi - \alpha_0) + n_{2Q}] \ , \tag{C.14}$$

onde $\rho = |\rho|\exp(\boldsymbol{j}\alpha)$ é o coeficiente de correlação cruzada dos sinais $\boldsymbol{S}_1$ e $\boldsymbol{S}_2$. As variáveis n_{1I}, n_{1Q}, n_{2I} e n_{2Q} no modelo são variáveis aleatórias (VA) gaussianas mutuamente descorrelacionadas que representam o ruído introduzido pelo canal, cuja média é nula e a variância unitária. A grandeza ε representa a energia do sinal.

Considerando a característica ortogonal do sinal BFSK, o que implica um coeficiente de correlação nulo ($\rho = 0$), as equações para $\boldsymbol{r}_1$ e $\boldsymbol{r}_2$ transformam-se em:

$$\boldsymbol{r}_1 = 2\varepsilon\cos(\phi) + n_{1I} + \boldsymbol{j}[2\varepsilon\sin(\phi) + n_{1Q}] \ , \tag{C.15}$$

$$\boldsymbol{r}_2 = n_{2I} + \boldsymbol{j} n_{2Q} \ . \tag{C.16}$$

Em decorrência da independência estatística das VA n_{1I}, n_{1Q}, n_{2I} e n_{2Q} em relação à fase e a elas próprias, a Função de Densidade de Probabilidade (PDF) pode ser escrita como o produto das PDF marginais de cada VA.

$$p(r_{1I}, r_{1Q}/S_1, \phi) = \frac{1}{2\pi}\exp\left\{ - \frac{[r_{1I} - 2\varepsilon\cos(\phi)]^2 + [r_{1Q} - 2\varepsilon\sin(\phi)]^2}{2\sigma^2} \right\} , \tag{C.17}$$

$$p(r_{1I}, r_{1Q}) = \frac{1}{2\pi}\exp\left\{ - \frac{[r_{2I}^2 + r_{2Q}^2]}{2\sigma^2} \right\} , \tag{C.18}$$

onde $\sigma^2 = 2\varepsilon N_0$ é a variância do sinal recebido e representa a energia do sinal contaminado pelo ruído do canal.

Se a VA ϕ possuir distribuição uniforme no intervalo $[0, 2\pi]$, a sua PDF é dada por $\frac{1}{2\pi}$. Substituindo este resultado e o resultado da Equação (C.15) na Equação (C.12), chega-se a:

$$p(r_{1I}, r_{1Q}/S_1) = \tfrac{1}{2\pi}\int_0^{2\pi} p(r_{1I}, r_{1Q}/S_1, \phi)\ d\phi$$

$$= \tfrac{1}{2\pi\sigma^2}\exp\left\{ - \tfrac{r_{1I}^2 + r_{1Q}^2 + 4\varepsilon^2}{2\sigma^2} \right\} \int_0^{2\pi} \exp\left\{ \tfrac{2\varepsilon[r_{1I}\cos(\phi) + r_{1Q}\sin(phi)]}{2\pi\sigma^2} \right\} d\phi \quad . \tag{C.19}$$

A integral constante da Equação (C.19) corresponde à função de Bessel modificada de ordem zero I_0; desta forma:

$$p(r_{1I}, r_{1Q}/S_1) = \frac{1}{2\pi\sigma^2}\exp\left\{ - \frac{r_{1I}^2 + r_{1Q}^2 + 4\varepsilon^2}{2\sigma^2} \right\} I_0\left(\frac{2\varepsilon\sqrt{r_{1I}^2 + r_{1Q}^2}}{\sigma^2} \right). \tag{C.20}$$

De maneira análoga, para o caso onde o sinal $\boldsymbol{S}_2$ é transmitido, pode-se escrever:

$$p(r_{2I}, r_{2Q}/S_2) = \frac{1}{2\pi\sigma^2}\exp\left\{ - \frac{r_{2I}^2 + r_{2Q}^2 + 4\varepsilon^2}{2\sigma^2} \right\} I_0\left(\frac{2\varepsilon\sqrt{r_{2I}^2 + r_{2Q}^2}}{\sigma^2} \right). \tag{C.21}$$

No intuito de se chegar a uma expressão mais simples para os elementos decisores, o primeiro passo é reescrever as Equações (C.11) e (C.12) em termos de razão de verossimilhança:

$$\Lambda(\boldsymbol{r}) = \frac{p(\boldsymbol{S}_1/\boldsymbol{r})}{p(\boldsymbol{S}_2/\boldsymbol{r})}. \tag{C.22}$$

Aplicando-se o teorema de Bayes à Equação (C.22), tem-se:

$$\Lambda(\boldsymbol{r}) = \frac{p(\boldsymbol{r}/\boldsymbol{S}_1)p(\boldsymbol{S}_1)}{p(\boldsymbol{r})} \frac{p(\boldsymbol{r})}{p(\boldsymbol{r}/\boldsymbol{S}_2)p(\boldsymbol{S}_2)}. \tag{C.23}$$

Para sinais binários $\boldsymbol{S}_1$ e $\boldsymbol{S}_2$ equiprováveis, i.e., $p(\boldsymbol{S}_1) = p(\boldsymbol{S}_2)$, a Equação (C.23) se reduz a:

$$\Lambda(\boldsymbol{r}) = \frac{p(\boldsymbol{r}/\boldsymbol{S}_1)}{p(\boldsymbol{r}/\boldsymbol{S}_2)} \tag{C.24}$$

$$= \frac{\exp\left\{ - \frac{r_{1I}^2 + r_{1Q}^2 + 4\varepsilon^2}{2\sigma^2} \right\} I_0\left(\frac{2\varepsilon\sqrt{r_{1I}^2 + r_{1Q}^2}}{\sigma^2} \right)}{\exp\left\{ - \frac{r_{2I}^2 + r_{2Q}^2 + 4\varepsilon^2}{2\sigma^2} \right\} I_0\left(\frac{2\varepsilon\sqrt{r_{2I}^2 + r_{2Q}^2}}{\sigma^2} \right)} \tag{C.25}$$

De maneira simplificada, pode-se escrever:

$$\Lambda(\boldsymbol{r}) = \frac{I_0\left(\frac{2\varepsilon\sqrt{r_{1I}^2 + r_{1Q}^2}}{\sigma^2} \right)}{I_0\left(\frac{2\varepsilon\sqrt{r_{2I}^2 + r_{2Q}^2}}{\sigma^2} \right)} \underset{\boldsymbol{S}_2}{\overset{\boldsymbol{S}_1}{\gtrless}} \frac{p(\boldsymbol{S}_1)}{p(\boldsymbol{S}_2)} \tag{C.26}$$

O detector ótimo tem como saídas as duas envoltórias $\sqrt{r_{1I}^2 + r_{1Q}^2}$ e $\sqrt{r_{2I}^2 + r_{2Q}^2}$. Na Equação (C.26), a variância não é conhecida *a priori* e, portanto, não se pode calcular a razão de verossimilhança somente com o resultado proveniente do detector. Para

superar o problema, pode-se fazer uso do fato de que a função de Bessel modificada de ordem zero é monotônica [23]; desta maneira, pode-se simplificar o processo decisório realizando-se a comparação entre a razão das envoltórias produzidas pelo detector e a razão das probabilidades de ocorrência dos símbolos $\boldsymbol{S}_1$ e $\boldsymbol{S}_2$.

$$\Lambda(\boldsymbol{r}) = \frac{\sqrt{r_{1I}^2 + r_{1Q}^2}}{\sqrt{r_{2I}^2 + r_{2Q}^2}} \underset{\boldsymbol{S}_2}{\overset{\boldsymbol{S}_1}{\gtrless}} \frac{p(\boldsymbol{S}_1)}{p(\boldsymbol{S}_2)} \tag{C.27}$$

Como $\boldsymbol{S}_1$ e $\boldsymbol{S}_2$ são equiprováveis, tem-se:

$$\Lambda(\boldsymbol{r}) = \frac{\sqrt{r_{1I}^2 + r_{1Q}^2}}{\sqrt{r_{2I}^2 + r_{2Q}^2}} \underset{\boldsymbol{S}_2}{\overset{\boldsymbol{S}_1}{\gtrless}} 1 \tag{C.28}$$

Como pôde ser observado, o cálculo das envoltórias do sinal recebido não depende das fases das respectivas portadoras, o que implica a imunidade deste tipo de detector a ruídos de fase.

Apêndice D

CODEC AMR

O CODEC AMR (*Adaptive Multirate*) foi originalmente desenvolvido para uso em celulares 3G, mas acabou sendo aplicado ao sistema GSM. Este CODEC permite que diferentes taxas de bits transportem a conversação, dependendo da qualidade do enlace, que pode ser traduzida em largura de banda disponível e taxa de erros de bit (BER).

Os quadros processados pelo CODEC AMR possuem duração de 20 ms e 160 amostras. Diferentes técnicas são empregadas pelo CODEC AMR [4]:

- *Discontinuous Transmission* (DTX);
- *Voice Activity Detection* (VAD); e
- *Comfort Noise Generation* (CNG).

O uso das técnicas DTX, VAD e CNG permite reduzir a largura de banda utilizada durante os períodos de silêncio do sinal.

O CODEC AMR ajusta dinamicamente a taxa de codificação de acordo com a qualidade do enlace de rádio. À medida que as condições do enlace se tornam mais críticas, a taxa de codificação é comutada para a taxa imediatemente inferior. A diminuição da qualidade do sinal, causada pelo aumento de compressão, é recompensada pelo aumento em 4 a 6 dB da razão sinal-ruído do enlace.

O CODEC AMR é baseado no modelo de predição linear com excitação por código (*code-excited linear predictive*-CELP). O modelo CELP tem como premissas:

- Emprego do modelo de fonte-filtro para a produção da fala, por meio de predição linear (LP);

- Uso de códigos fixos e adaptáveis para a entrada do modelo de predição linear (LP); e
- Aplicação de quantização vetorial.

Conforme explicitado na Tabela D.1, o CODEC AMR emprega um total de 14 taxas de codificação, sendo 8 pertencentes ao modo *Full-Rate* (FR) [24] e 6 pertencentes ao modo *Half-Rate* (HR) [25].

Tabela D.1: Taxas de codificação AMR.

Modo de Operação	Taxa (*kbps*)	Canal	Nº de Bits do Bloco Codificado	CODEC Compatível
AMR-SID	1,80	FR/HR	-	-
AMR-4,75	4,75	FR/HR	95	-
AMR-5,15	5,15	FR/HR	103	-
AMR-5,90	5,90	FR/HR	118	-
AMR-6,70	6,70	FR/HR	134	ARIB 6.7 kbit/s *enhanced full rate*
AMR-7,40	7,40	FR/HR	148	TIA/EIA IS-641 TDMA *enhanced full rate*
AMR-7,95	7,95	FR/HR	159	-
AMR-10,20	10,20	FR	204	-
AMR-12,20	12,20	FR	244	ETSI GSM *enhanced full rate*

A seguir são resumidas as principais caracteríticas do CODEC AMR:

- Freqüência de amostragem: 8 kHz, i.e., 160 amostras para cada bloco de 20 ms;
- Pré-filtragem: $200 - 3400$ Hz;
- Tipo de codificação: Codificador híbrido de voz do tipo *Algebraic Code Excited Linear Prediction* (ACELP);
- Atraso total do algoritmo: 20 ms + 5 ms; e

- Complexidade do algoritmo (CA): 5 (para PCM (G.711), CA=1);
- Medida PSQM para AMR (12, 20 kbps) em condições ideais: 4, 45; e
- Medida PSQM para AMR (12, 20 kbps) em condições reais: 3, 75.