

Neuro-Equalizador para Canais Variantes no Tempo com Algoritmo de Treinamento do Tipo Gauss-Newton

Claudio José Camera Santos

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal da Bahia, como parte dos requisitos para obtenção do grau de Mestre em Engenharia Elétrica.

Área de Concentração: Processamento de Sinais

Orientador: Antonio Cezar de Castro Lima

Banca examinadora

Antonio Cezar de Castro Lima - UFBA

Luis Alberto Luz de Almeida - UFBA

Evandro Ottoni Teatini Salles - UFES

Agradecimentos

Ao professor Antonio Cezar Castro Lima pela orientação, paciência, incentivo e pragmatismo, oferecidos e demonstrados não só no desenvolvimento deste trabalho, mas também, durante o tempo em que nos conhecemos.

Aos meus amados pais pelo amor incondicional e todo o apoio dispensado em todas as fases de minha vida, não há palavras suficientes para expressar o meu agradecimento.

A minha amada esposa Mônica, pelo amor, carinho, apoio e paciência ao longo do nosso convívio e desta jornada. Se a sua compreensão fosse estabelecida por um algoritmo, este teria ordem de convergência infinita.

A minha filha Maria Lúcia pelo seu sorriso e pelo futuro.

Aos professores e colegas do Mestrado de Engenharia Elétrica da UFBA pela troca de experiências técnicas e pessoais.

A todos os meus verdadeiros amigos, vocês sabem quem vocês são.

Homenagem

Dedico este trabalho à memória de José Carlos Sousa dos Santos, um dos engenheiros mais criativos que conheci e que sempre trazia consigo um humor contagiante, ele foi um excelente colega de trabalho e amigo. José Carlos me influenciou e me incentivou a ingressar e a perseverar no curso de mestrado, entretanto seu maior legado para mim são os ensinamentos proverbiais que ele aplicava as mais diversas situações cotidianas, dele e destes, lembrarei até o fim da minha vida.

Resumo

A utilização de técnicas de Redes Neurais Artificiais (RNA) tornou-se bastante comum na equalização de diversos tipos de canais de comunicações. RNAs se apresentam em diversas topologias. A escolha de uma topologia para equalização de um determinado canal pode ser realizada considerando diferentes critérios tais como: velocidade de convergência do método, Taxa de Erro de Bits (BER), complexidade computacional, dentre outros. Neste trabalho é investigado o comportamento em canais variáveis no tempo de um tipo de Equalizador com Realimentação de Decisão (DFE) formado apenas por um *perceptron* recorrente e treinado através do algoritmo de Gauss-Newton (GN) para estimar os pesos sinápticos tanto na fase de aprendizado e como na fase de testes. Apesar da simplicidade de implementação e baixo custo computacional, a topologia adotada apresentou bom desempenho quando comparada com estruturas mais complexas baseadas em Redes Neurais Recorrentes (RRN) ou Multicamadas de *Perceptrons* (MLP) usando filtros de Kalman.

Abstract

Artificial Neural Networks (ANN) techniques have become very common as equalization solutions in several types of communication channels. These Neural Networks are presented in many topologies. The suitable choice of a topology for equalization purpose depends on different criteria such as: convergence rate, Bit Error Rate (BER), computational complexity, among many others. In this work, it is investigated the behavior of a structure similar to a Decision Feedback Equalizer (DFE) employed to equalize a time varying channels. The structure is based on simplified Recurrent Neural Network (RNN) which ended up being only one recurrent perceptron. The Gauss-Newton algorithm has been used to estimate the synaptic weights of the perceptron during the training and testing phases. Despite the simplicity of implementation and low computational cost, it has been shown that the proposed topology presents good comparative performance related to more complex structures based on Recurrent Neural Networks (RNN) and Multilayer Perceptrons (MLP) using Kalman Filters.

Conteúdo

Introdução	10
1.1 Organização desta Dissertação.....	13
Equalizadores.....	15
2.1 Equalizadores Lineares.	16
2.2 Equalizadores Não Lineares.....	18
2.3 Equalizadores Adaptativos.....	19
2.4 Neuro-Equalizadores.....	22
2.4.1 Implementação do Neuro-Equalizador DFE.	23
Canais Variáveis no Tempo.....	26
3.1 Modelagem como Processos Aleatórios	30
3.2 Simulação de SLVT	32
3.2.1 Modelo associado a filtro transversal linear.....	33
3.2.2 Simplificação do Modelo	34
3.2.3 Geração dos coeficientes do filtro	35
3.3 Caracterização Estatística de Canais Variáveis no Tempo	35
Algoritmo de Gauss-Newton.....	41
Procedimentos de Teste.....	46
5.1 Modelos de Canais	46
5.1.1 Canal modelo 1	46
5.1.2 Canal modelo 2	50
5.1.3 Canal modelo 3	51
Simulações e Resultados.....	54
6.1.1 Canal modelo 1	56
6.1.2 Canal modelo 2	59
6.1.3 Canal modelo 3	62
Conclusão	65
Referências Bibliográficas.....	66

Lista de Figuras

Figura 2.1 - Estrutura de um DFE.....	19
Figura 2.2 - Diagrama em bloco de um equalizador adaptativo	20
Figura 2.3. - Modos de operação de um equalizador adaptativo	21
Figura 2.4. - Estrutura de um DFE para o sistema digital de comunicações	21
Figura 2.5- Comparação entre um perceptron e filtro adaptativo.....	22
Figura 2.6 - Esquemático do DFE.....	23
Figura 2.7 - DFE Implementado por um perceptron.....	25
Figura 3.1 - Resposta de um SLVT a um tom aplicado à entrada com deslocamento em frequência Δf espalhamento em frequência ΔB	30
Figura 3.2 - Espalhamento de frequências em SLVT	32
Figura 3.3 - Modelo de canal variável no tempo por filtro transversal linear	34
Figura 3.4 - Perfil de intensidade de multipercursos	37
Figura 3.5 - Função de correlação da diferença de frequência $R(\Delta f)$ e banda de coerência do canal B_c	38
Figura 3.6 - função de correlação de espaçamento de tempo $R(\Delta t)$ e tempo de coerência T_c	40
Figura 5.1 – Canal modelo 1 , variação dos coeficientes para espalhamento Doppler 0,5 Hz.	47
Figura 5.2 – Diagrama de espalhamento canal modelo 1 para espalhamento Doppler de 0,5 Hz....	48
Figura 5.3 – Canal modelo 1 , variação dos coeficientes para espalhamento Doppler 1 Hz.....	49
Figura 5.4 – Diagrama de espalhamento canal modelo 1 para espalhamento Doppler de 1 Hz.....	49
Figura 5.5 – Canal modelo 2 , variação dos coeficientes para espalhamento Doppler 0,5 Hz.....	50
Figura 5.6 – Diagrama de espalhamento canal modelo 2 para espalhamento Doppler de 0,5 Hz....	50
Figura 5.7 – Canal modelo 3 , variação dos coeficientes para desvio padrão 0,1.	51
Figura 5.8 – Diagrama de espalhamento canal modelo 3 para desvio padrão 0,1.....	52
Figura 5.9 – Canal modelo 3 , variação dos coeficientes para desvio padrão 0,3.	52
Figura 5.10 – Diagrama de espalhamento do canal modelo 3 , variação dos coeficientes para desvio padrão 0,3.....	53
Figura 6.1 – Canal modelo 1 , comparação da BER e ação do GN nos coeficientes no período de teste para espalhamento Doppler 0,5 Hz.....	56
Figura 6.2 – Canal modelo 1 , comparação da BER e ação do GN nos coeficientes no período de teste para espalhamento Doppler 1 Hz.....	57
Figura 6.3 – Canal modelo 1 , comparação de desempenho BER do GN com outras topologias	

para espalhamento Doppler 0,5 Hz.	58
Figura 6.4 – Canal modelo 1 , comparação de desempenho BER do GN com outras topologias para espalhamento Doppler 1 Hz.	58
Figura 6.5 – Canal modelo 2 comparação da BER e ação do GN nos coeficientes no período de teste para espalhamento Doppler 0,5 Hz.	59
Figura 6.6 – Canal modelo 2 comparação da BER e ação do GN nos coeficientes no período de teste para espalhamento Doppler 1 Hz.	60
Figura 6.7 – Canal modelo 2 , comparação de desempenho BER do DFE-GN com outras topologias para espalhamento Doppler 0,5 Hz.	61
Figura 6.8 – Canal modelo 3 , comparação de desempenho BER e ação do GN nos coeficientes no período de teste variando desvio padrão.	62
Figura 6.9 – Canal modelo 3 , comparação de desempenho BER do DFE-GN e outras topologias variando desvio padrão.	63
Figura 6.10 – Canal modelo 3 , comparação de desempenho BER do DFE-GN e outras topologias.	64

Abreviaturas

AWGN	Ruído Branco Aditivo Gaussiano (Additive White Gaussian Noise)
BER	Taxa de Erro de Bit (Bit Error Rate)
DFE	Equalizador com Decisão Realimentada (Decision Feedback Equalizer)
DFE-GN	Neuro-Equalizador DFE com Algoritmo de Treinamento Gauss-Newton
DFRNE	Neuro-Equalizador Recursivo com Decisão Realimentada (Decision Feedback Recursive Neuralequalizer)
EKF	Filtro de Kalman Estendido (Extended Kalman Filter)
FNN	Rede Neural Direta (Feedforward Neural Network)
FIR	Resposta ao Impulso com Duração Finita (Finite Impulse Response)
GN	Método de Gauss-Newton
IIR	Resposta ao Impulso com Duração Infinita (Infinite Impulse Response)
ISI	Interferência Intersimbólica
LMS	Mínimos Quadrados Médios (Least Mean Square)
MLSE	Estimativa de Seqüência pela Máxima Verossimilhança (Maximum Likelihood Sequence Estimation)
MSE	Erro Quadrático Mínimo (Mean Square Error)
MLP	Múltiplas Camadas de <i>Perceptrons</i> (Multilayer Perceptron)
NN	Rede Neural (Neural Network)
PN	Pseudo-Aleatória (Pseudo Noise)
PSD	Densidade Espectral de Potência (Power Spectral Density)
QAM	Modulação em Amplitude e Quadratura (Quadrature Amplitude Modulation)
RLS	Mínimos Quadrados Recursivos (Recursive Least Square)
RNA	Redes Neurais Artificiais
RNE	Equalizador Neural Recursivo (Recursive Neural Equalizer)
RNN	Rede Neural Recursiva (Recursive Neural Network)
RTRL	Aprendizado Recorrente em Tempo Real (Real Time Recurrent Learning)
SER	Taxa de Erro de Símbolo (Symbol Error Rate)
SLVT	Sistema Linear Variante no Tempo
SLIT	Sistema Linear Invariante no Tempo
SNR	Relação Sinal Ruído (Signal to Noise Ratio)
UKF	Filtro de Kalman Unscented (Unscented Kalman Filter)
ZFE	Equalizador Forçador a Zero (Zero Forcing Equalizer)

Capítulo 1

Introdução

A cada dia, os sistemas de comunicações estão mais presentes em nossas vidas. Podemos ser encontrados a qualquer momento pelo telefone celular; rádios e televisores estão em todos os cantos das residências; computadores nos conectam a Internet seja em casa, na escola ou no escritório. Através da Internet temos acesso constante a notícias, pagamos contas bancárias, podemos falar ou trocar mensagens instantâneas com nosso vizinho ou com alguém no outro lado do mundo. Tão comuns são as ferramentas de comunicações que a maioria das pessoas sequer imagina a complexidade que está por trás do seu funcionamento (Haykin, 2001).

Para que o processo de transmissão e recepção se realize é necessário que tenhamos presentes os elementos de um sistema de comunicações: A informação; a transformação da informação em sinal, que pode necessitar de codificação e/ou modulação; um canal de transmissão, que pode ser um meio físico ou a atmosfera; a recepção do sinal, que pode ser associada a um processo de demodulação e/ou decodificação; e por último a transformação do sinal em uma estimativa da informação original. (Proakis, 2001).

A transmissão e recepção de um sinal podem ser analógicas ou digitais. Em qualquer uma das duas, a mensagem recebida sempre será diferente da mensagem original devido a diversos fatores tais como: imperfeições na conversão do sinal original em sinal elétrico, ruído introduzido pelos equipamentos envolvidos e pelo canal de comunicações, imperfeições no canal, interferências, entre muitos outros. Porém, é interessante que a mensagem recebida seja a mais próxima possível da mensagem transmitida e que os erros sejam minimizados. A transmissão digital é preferida por diversos motivos: circuitos digitais são menos suscetíveis a distorções e interferências que os analógicos; circuitos digitais são mais flexíveis, confiáveis e, atualmente, com menor custo que circuitos analógicos; a transmissão digital possui maior imunidade ao ruído permitindo uma troca exponencial da largura de banda com a figura de mérito enquanto as técnicas analógicas permitem, no máximo, uma relação quadrática; a transmissão digital permite a detecção e correção de erros na mensagem através de técnicas e algoritmos de codificação adequados; transmissões digitais são mais seguras, permitindo criptografia e cifragem.

Não há transmissão e recepção de um sinal sem que haja um canal entre eles. Um canal ideal não possuiria ruído e deveria apresentar uma resposta em frequência com duas características essenciais: ganho constante em toda a faixa de frequência e fase linear. Infelizmente, tal canal não existe em situação prática e é exatamente no canal onde o projetista tem menor controle e onde se encontram as maiores fontes de erro nos sistemas de comunicações.

Os dois tipos mais comuns de erros provenientes do canal são o ruído e a interferência intersimbólica (ISI). Os dados digitais têm um amplo espectro, normalmente com conteúdo de baixa frequência mais significativo do que o conteúdo nas altas frequências. A transmissão de sinais digitais requer o uso de um canal com largura de banda suficiente para acomodar as componentes essenciais de frequência do sinal transmitido, entretanto, tipicamente o canal é dispersivo, sua resposta em frequência não apresenta as características de um canal ideal. O resultado é que os pulsos transmitidos sofrem modificações em sua forma e são afetados por pulsos adjacentes, dando origem a uma forma de interferência bastante comum denominada Interferência Intersimbólica (ISI) (Haykin, 2001).

Para complicar, no canal também existe o ruído que se mistura ao sinal, modificando suas características. Uma prática bastante comum é considerar o ruído no canal como AWGN, ruído aditivo branco gaussiano, um ruído presente em todo o espectro de frequência de interesse e com sua amplitude sendo uma variável aleatória com função de distribuição gaussiana.

Para combater os efeitos do ruído aditivo branco gaussiano, o receptor deve incorporar um filtro casado e para combater os efeitos da interferência intersimbólica, o receptor deve incorporar um equalizador.

Na prática, o ruído e a interferência intersimbólica estão presentes simultaneamente no canal, afetando os dados transmitidos de maneira combinada. Desta forma, o receptor otimizado deve ser projetado para tratar conjuntamente o ruído e os efeitos dispersivos do canal. (Haykin, 2001).

Este trabalho propõe-se a apresentar um método de combate à interferência intersimbólica ISI em canais contaminados com ruído branco aditivo gaussiano. Será visto que existem, basicamente, duas categorias de equalizadores: os equalizadores que utilizam a estimativa de sequência pelo método da máxima verossimilhança e os equalizadores que utilizam filtragem. O segundo grupo apresenta a vantagem de ter menor complexidade computacional e também, por isso, é objeto deste

trabalho.

Será visto também que os equalizadores lineares possuem uma forte limitação quando trabalham em canais que possuem componentes espectrais nulos em sua resposta em frequência. Esta dificuldade não é tão pronunciada quando se utilizam equalizadores não lineares, como é o caso do equalizador com decisão realimentada DFE. Quando redes neurais são incorporadas ao DFE, se obtém resultados ainda melhores em termos de velocidade de convergência e erro quadrático médio quando comparados com DFEs sem redes neurais ou neuro-equalizadores sem realimentação de decisão (Choi, Lima e Haykin , 2005) Uma das principais características de canais variáveis no tempo é apresentar nulos em sua resposta em frequência. (Sklar, 2001).

Devido a sua simplicidade de implementação, redes neurais diretas FFN, *Feed Forward Networks*, comumente são aplicadas como soluções para compor neuro-equalizadores (Zerguine *et al*, 2001), (Choi, Lima e Haykin, 2005). Em contrapartida, alguns estudos apresentam as Redes Neurais Recursivas RNN - *Recursive Neural Network* como superiores em desempenhos aos FNNs (Choi, Lima e Haykin , 2005). Estes estudos se baseiam em algumas limitações das FNN, principalmente às relacionadas com as baixas taxas de convergências dos métodos baseados no gradiente descendente e em consequência da necessidade de seqüências de treinamento extremamente longas. Por outro lado, RNN são normalmente implementadas por redes neurais multicamadas que utilizam algoritmos de treinamento baseados no Filtro de Kalman. Estes algoritmos são bastante complexos e demandam um esforço computacional muito grande. (Cortes *et al*, 2006).

Para contornar o problema da baixa convergência associada ao método do gradiente descendente baseado em primeira diferenciação, e portando um método de primeira ordem, o método de segunda diferenciação proposto por Newton seria bastante desejável, pois alcança taxas de convergência de segunda ordem. Porém, a grande desvantagem do método de Newton reside no esforço computacional de cálculo e inversão da matriz Hessiana. Para contornar este problema, métodos Quasi-Newton são propostos de maneira a atingir uma convergência próxima de segunda ordem do método de Newton, contornando o esforço computacional devido ao cálculo da Hessiana e sua inversa (Luenberg, 2003). Um método alternativo é o método de Gauss-Newton, que pode ser visto como uma modificação do método de Newton e será utilizado no neuro-equalizador proposto neste trabalho.

O objetivo deste trabalho é apresentar uma proposta de neuro-equalizador FNN composto por apenas um *perceptron* recorrente supervisionado com algoritmo de treinamento utilizando o método de Gauss-Newton. Este trabalho nasceu de uma proposta de estudo futuro de (Cortes, 2006) e (Ludwig and Lima, 2005). Nestes trabalhos foram avaliados desempenhos de um neuro-equalizador DFE com algoritmo de treinamento Gauss-Newton em diversos modelos de canais fixos, porém deixando abertos o comportamento e o desempenho da estrutura proposta para canais variáveis no tempo. Este trabalho dá continuidade àqueles estudos, adaptando o algoritmo de forma a proceder à equalização em canais variantes no tempo. Para efeito de comparação de desempenho, foram utilizados três modelos de canais retirados de (Ling e Proakis, 1985), (Choi, Lima e Haykin, 2005) e (Choi, Bouchard e Yeap, 2005). Estes três canais se mostraram adequados por englobarem situações de desvanecimento lento e desvanecimento rápido, naturalmente necessárias para o estudo em questão.

Ao final das simulações é possível perceber que, com o método proposto, obtemos um excelente compromisso entre desempenho do equalizador em termos de BER e esforço computacional exigido quando comparado com outras topologias.

1.1 Organização desta Dissertação

Esta dissertação está dividida em sete capítulos:

- O capítulo 1 inicia com uma introdução a sistemas de comunicações, o papel do canal e os problemas dele decorrentes, como ruído branco aditivo gaussiano (AWGN) e interferência intersimbólica (ISI), além de levantar a necessidade de técnicas para mitigar estes efeitos. Desta forma introduz a motivação para equalizações de canais e finalmente apresenta a proposta do trabalho.
- O capítulo 2 apresenta, com mais detalhes, os conceitos fundamentais de equalizadores; introduz algumas topologias básicas; apresenta, resumidamente, o conceito de equalização adaptativa e Redes Neurais nas técnicas de equalização e apresenta a estrutura do neuro-equalizador DFE com algoritmo de treinamento do tipo Gauss-Newton proposta neste trabalho.
- O capítulo 3 apresenta a teoria básica de canais variantes no tempo, explora conceitos

fundamentais como desvanecimento e espalhamento Doppler, modelagem, simulação e caracterização estatística de canais variantes no tempo.

- O capítulo 4 faz uma breve descrição do método de Gauss-Newton e discute a sua aplicação no neuro-equalizador proposto neste trabalho.
- O capítulo 5 apresenta os procedimentos de testes e a caracterização dos modelos de canais utilizados nas simulações.
- O capítulo 6 apresenta as simulações e os resultados obtidos, bem como realiza comparações com resultados de artigos pertinentes.
- O capítulo 7 apresenta a conclusão do trabalho e finaliza a dissertação, com sugestão de trabalho futuro.

Capítulo 2

Equalizadores

Os equalizadores podem ser divididos, basicamente, em duas grandes categorias: Estimativa de Sequência pela Máxima Verossimilhança e Equalização por filtragem.

O método da Estimativa de Sequência pela Máxima Verossimilhança (MLSE) procura estimar a resposta ao impulso do canal pela estatística das seqüências de dados enviadas. O objetivo é ajustar o receptor para a seqüência enviada, de maneira a possibilitar ao detector realizar boas estimativas da seqüência de pulsos distorcidos pelo canal proveniente do demodulador. O receptor MLSE não se baseia na tentativa de eliminar ou compensar as distorções na seqüência de pulsos, ao invés disto, tenta se ajustar ao canal, estimando a sua resposta ao impulso, de maneira a lidar melhor com as seqüências distorcidas, um exemplo é a equalização com algoritmo de Viterbi. (Sklar, 2001).

A complexidade computacional de um equalizador MLSE cresce exponencialmente com a dispersão do canal. Se o tamanho do alfabeto de símbolos é M e o número de símbolos contribuindo para a ISI é L , o algoritmo de Viterbi, por exemplo, terá uma métrica computacional de $M^{(L+1)}$ para cada símbolo recebido (Proakis, 2001). A complexidade computacional guarda uma relação com o custo de implementação da solução e, portanto, a depender das condições do canal, pode ser mais adequado utilizar uma solução que seja, ao mesmo tempo, satisfatória em termos de desempenho com relação à taxa de erro e que necessite de baixo esforço computacional, o que conseqüentemente oferecerá uma melhor relação custo benefício.

A segunda categoria é a equalização por filtragem. Esta técnica utiliza filtros para compensar as distorções dos símbolos. Nesta categoria, o detector recebe os símbolos depois de demodulados e modificados pelo equalizador cujo objetivo é minimizar os efeitos da ISI. (Sklar, 2001).

Na equalização por filtragem, o custo computacional é menor, pois não há a necessidade de estimar a resposta ao impulso do canal. Principalmente pela menor complexidade computacional e facilidade de implementação de hardware, a equalização por filtros é a técnica mais utilizada. Os filtros utilizados neste processo de equalização podem ser lineares do tipo Resposta ao Impulso de Duração Finita (FIR) e Resposta ao Impulso de Duração Infinita (IIR), ou não-lineares, onde se

introduz uma função não-linear na saída do equalizador. Os equalizadores que utilizam a topologia IIR com função não-linear do tipo *hard-limit*, na saída são denominados Equalizadores com Decisão Realimentada ou DFE - *Decision Feedback Equalizers* (Sklar, 2001).

Devido ao menor custo computacional e à simplicidade, este trabalho será baseado na técnica de utilização por filtragem. No item a seguir, será apresentado um breve resumo da teoria dos mesmos.

Outro fator importante na consideração dos equalizadores por filtragem é a taxa de amostragem do filtro. Sendo T_s a duração do símbolo, se o espaçamento para cada atraso dos coeficientes do filtro for T_s , o equalizador é chamado de espaçamento por símbolo. Se o espaçamento para cada atraso dos coeficientes do filtro for menor que T_s , o equalizador é chamado de espaçamento fracionado.

Para um equalizador com espaçamento fracionado cada atraso é dado por

$$T_s' = \frac{T_s}{1 + \alpha} \quad (2.1)$$

Onde α representa o excesso da largura de banda escolhida, e T_s' representa o tempo de amostragem.

O objetivo é que o tempo de amostragem deve ser escolhido de forma que a função de transferência do equalizador contemple o espectro do símbolo. Veja que a saída do equalizador continua com o tempo de amostragem igual a T_s , enquanto que a amostragem de entrada é T_s' . Desta forma, o equalizador atua nos símbolos recebidos antes que suas componentes de frequência sejam sobrepostas.

A implementação da técnica de espaçamento fracionado demanda maior banda, maior complexidade e maior custo computacional.

Como o objetivo da estrutura proposta neste trabalho é ser a mais simples possível e, uma vez que as estruturas propostas nos artigos que usaremos como comparação são baseadas na técnica de espaçamento por símbolo. Esta será a utilizada neste trabalho.

2.1 Equalizadores Lineares.

A primeira abordagem, e de conceito mais simples para um equalizador linear, é baseada na

minimização da distorção de pico, também conhecida como equalizador forçador a zero ou ZFE - *Zero Forcing Equalizer*. Considerando que o formato do pulso no transmissor é conhecido e que o canal é a única parte do sistema fora de controle, seu princípio é bastante simples: Sendo a resposta do canal $H(z)$, bastaria que o equalizador fosse construído de forma que sua resposta em frequência fosse $H^{-1}(z)$ avaliado em $z = e^{-j\omega}$, ou seja, o sistema formado pelo canal e equalizador teria uma resposta plana em frequência, desta maneira, a ISI seria eliminada. Entretanto, para esta abordagem, alguns problemas de ordem prática se colocam como complicadores. O equalizador ZFE só poderia ser aplicado em canais de fase mínima e, além disto, o ZFE ignora o efeito do ruído na composição da resposta do canal (Haykin, 2001).

Uma abordagem mais refinada que o ZFE é o Equalizador com Erro Quadrático Médio Minimizado MMSE - *Minimum Mean-Square Equalizer*, cujo princípio é baseado em ajustar os coeficientes do filtro, de forma que o erro quadrático médio entre o símbolo transmitido e o símbolo estimado seja minimizado, considerando a equação (2.2),

$$\varepsilon(k) = I(k) - \hat{I}(k) \quad (2.2)$$

Onde $\varepsilon(k)$ é o erro, $I(k)$ é o símbolo transmitido e $\hat{I}(k)$ é a estimativa do símbolo transmitido todos no instante k . O objetivo é minimizar a função custo $J = E |\varepsilon(k)|^2 = E |I(k) - \hat{I}(k)|^2$, o que pode ser feito por um método de otimização, entre os métodos mais comuns, figuram o algoritmo LMS e o algoritmo RLS. Claramente o equalizador MMSE apresenta a vantagem sobre o ZFE de considerar a presença do ruído na estimativa do símbolo.

Quando os canais são bem comportados e não apresentam nulos na resposta em frequência, o desempenho dos equalizadores lineares é satisfatório. Entretanto, em canais com ISI mais severas, com nulos na resposta em frequência, o desempenho dos equalizadores lineares é comprometido devido ao fenômeno conhecido por “*noise enhancement*” (Proakis, 2001). Este fenômeno pode ser melhor entendido, analisando o comportamento do ZFE. Em virtude de tentar inverter a resposta em frequência do canal $H(z)$, o ZFE promove um ganho elevado no ruído em frequências onde a resposta do canal $H(z)$ tenha uma magnitude pequena, ou seja, em regiões de nulos espectrais. (Haykin, 2001).

A limitação dos equalizadores lineares para lidar com ISI severas e com canais com nulos na resposta em frequência motivou um grande número de pesquisas com equalizadores não lineares de

baixa complexidade computacional. O equalizador DFE surgiu como uma solução efetiva para esta limitação (Proakis, 2001).

2.2 Equalizadores Não Lineares.

O DFE foi resultado de uma tese de doutorado de M.E. Austin submetida ao *Massachusetts Institute of Technology* em 1967. A proposta do DFE é inserir um estágio equalizador aos chamados equalizadores convencionais (*feedforward*) de maneira que as decisões dos símbolos anteriores são utilizadas para a decisão do símbolo presente.(Austin, 1967).

A estrutura de um DFE consiste em um filtro casado, seguido por um filtro transversal linear que forma o estágio equalizador direto “*feedforward*”, este estágio é seguido por um decisor, cuja saída (decisões prévias) entra em outro filtro transversal linear que forma o estágio de realimentação ou “*feedback*” como pode ser visto na figura 2.1.

A entrada do filtro direto é dada pelo sinal recebido da saída do filtro casado, assim, os coeficientes do filtro direto se ajustam de maneira a compensar os efeitos provenientes da adição de ruído nos símbolos (Austin, 1967). Na prática, os coeficientes do filtro direto se ajustam de forma a incorporar o filtro casado ao equalizador (Haykin, 2001). Após o filtro direto o sinal é submetido a um dispositivo decisor, este dispositivo na proposta de Austin é um “*hard decision*”, embora, no nosso trabalho utilizaremos uma abordagem diferente para o decisor como será discutida posteriormente. A saída do decisor é uma estimativa dos símbolos que será encaminhada para o segundo estágio que é o filtro de realimentação, os coeficientes deste filtro se ajustam de forma a eliminar a distorção da estimativa atual, subtraindo a interferência dos símbolos anteriores no símbolo atual (Austin, 1967).

A figura 2.1. apresenta a estrutura original do DFE, claramente o elemento decisor introduz uma não linearidade no sistema, fazendo do DFE s um equalizador não linear.

Os filtros direto e reverso são tipicamente estruturas FIR. Na implementação do DFE é importante determinar três parâmetros m , n e d que são respectivamente a ordem do filtro direto, a ordem do filtro de realimentação e o atraso da decisão com relação a entrada, estes números caracterizam o DFE, as entradas consistem portanto das entradas diretas $\mathbf{r}(\mathbf{k}) = [r(k), r(k-1), \dots, r(k-m+1)]^T$ e das entradas de realimentação $\mathbf{u}(\mathbf{k}) = [u(k), u(k-1), \dots, u(k-n)]^T$. A saída do DFE é $\mathbf{y}(\mathbf{k})$ que é

então passada através do dispositivo de decisão. Para determinação do símbolo estimado $\hat{s} = (k - d)$, é suficiente usar uma ordem de realimentação igual a $n = N + m - d - 2$, uma vez que os símbolos que contribuem para a decisão no instante de tempo k são $\mathbf{s}(\mathbf{k}) = [s(k), s(k-1), \dots, s(k-m-N+2)]^T$ para a ordem do filtro direto $m=d+1$ (Choi, Lima e Haykin, 2005) note que o valor de d está relacionado com a ordem do canal discreto equivalente N .

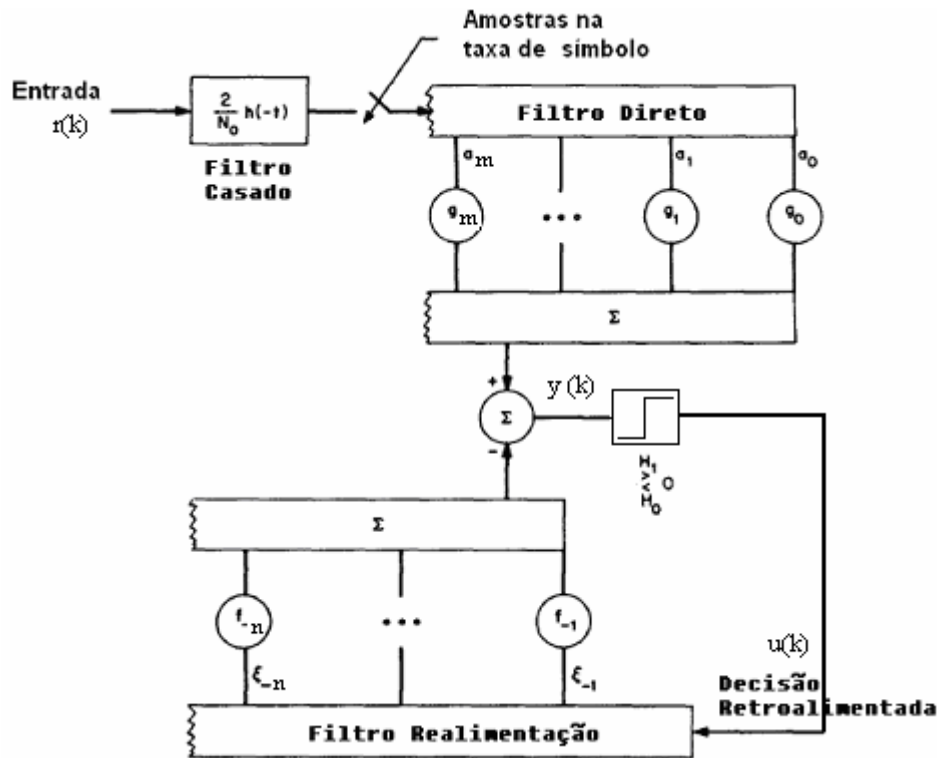


Figura 2.1 - Estrutura de um DFE – Adaptado de (Austin, 1967).

2.3 Equalizadores Adaptativos

Na discussão sobre equalizador DFE vimos que os coeficientes dos filtros diretos e de realimentação se ajustam para mitigar os efeitos do ruído e da interferência intersimbólica. Para a maioria das aplicações práticas é conveniente que os coeficientes dos equalizadores se modifiquem automaticamente com as condições dos canais. A utilização de filtros fixos seria inviável até para canais que possuem uma variação relativamente tão lenta que poderiam ser chamados de canais fixos. Na prática, não existem canais fixos, uma vez que variações de condições ambientais tais como temperatura e a degradação dos componentes com o tempo ocasionariam mudanças na resposta do canal. Além disso, na maioria das vezes não é possível determinar as condições dos canais a priori, por exemplo, um canal de telefonia fixa comutada terá características diferentes a cada ligação realizada, uma vez que ligações diferentes passam por rotas diferentes (Proakis, 2001).

Equalizadores ou filtros que tem seus coeficientes se adaptando de forma a otimizar uma resposta desejada são denominados adaptativos.

Basicamente um equalizador adaptativo é composto por duas partes, um filtro digital e um algoritmo que através de um critério de otimização atualiza os coeficientes do filtro.

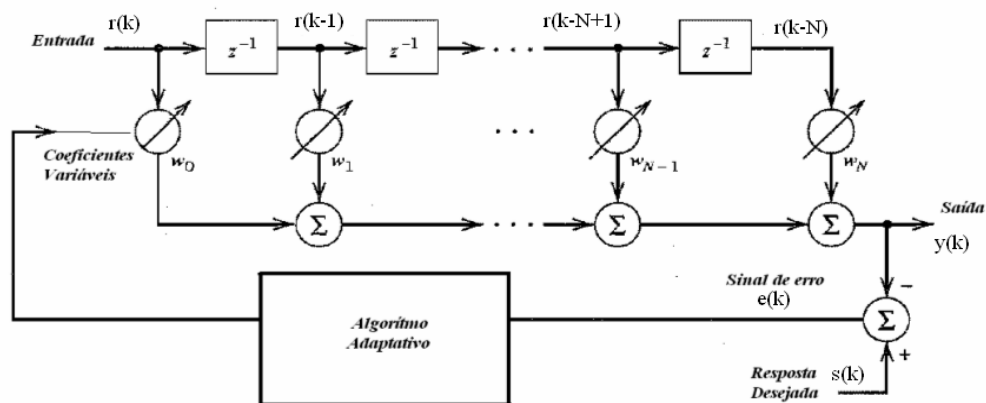


Figura 2.2 - Diagrama em bloco de um equalizador adaptativo - *Adaptado de (Haykin, 2001)*

A figura 2.2. mostra o digrama em bloco de um equalizador adaptativo linear. Basicamente o sinal na entrada passa pelo filtro e é comparado com um sinal desejado. O sinal erro proveniente é a variável de entrada, ou seja, a função custo que deve ser minimizada por um algoritmo através da modificação dos valores dos coeficientes do filtro. Ao minimizar o sinal erro, a saída do filtro se torna uma estimativa da resposta desejada. Porém, este esquema pressupõe que uma resposta desejada esteja disponível para comparação, esta última afirmação parece ser ambígua, pois o objetivo do equalizador é exatamente prover esta resposta que na verdade é o sinal da fonte de informação distorcido pelo canal e apresentado à entrada do equalizador.

O que na verdade acontece, é que a resposta desejada é fornecida durante um certo tempo conhecido como período de treinamento. Durante o período de treinamento, uma sequência conhecida é gerada na fonte e transmitida através do canal. Uma versão desta sequência é gerada no receptor com o devido sincronismo, considerando inclusive o atraso proveniente do filtro. Estas seqüências são conhecidas como pseudo-aleatórias ou seqüências PN *pseudo-noise*. Uma seqüência PN é uma seqüência determinística e periódica que apresenta características de ruído (Haykin, 2001).

Quando o processo de treinamento é concluído, o sistema passa a seu modo de operação ou modo de teste. No modo de teste, o transmissor passa a transmitir o sinal de interesse, enquanto no receptor, duas situações podem ocorrer: os coeficientes do filtro são congelados ou o sinal de comparação com a saída do filtro passa a ser a estimativa do equalizador. A última situação é conhecida como modo de decisão direta, neste modo, o algoritmo continua a modificar os coeficientes em uma tentativa de rastrear as mudanças do canal.

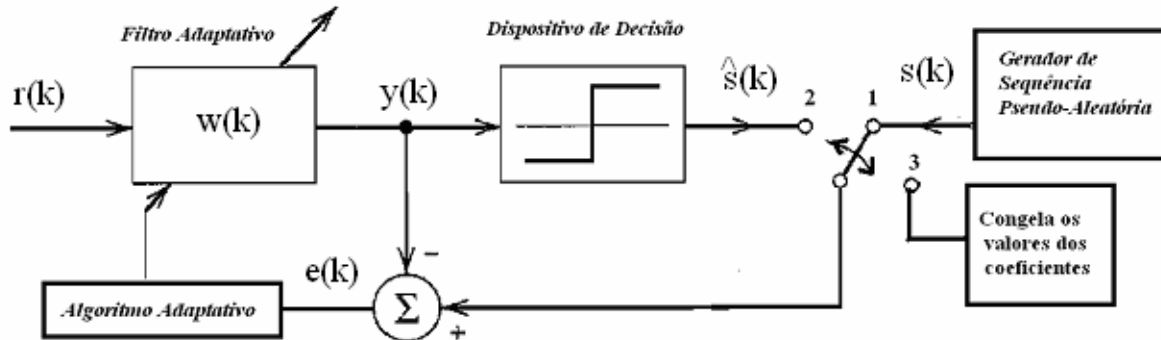


Figura 2.3. - Modos de operação de um equalizador adaptativo: No modo de treinamento a chave está na posição 1, no modo de rastreamento a chave é colocada na posição 2 e no modo de congelar os coeficientes a chave é colocada na posição 3. *Adaptada de (Haykin, 2001)*

O esquema apresentado na figura 2.3 mostra os modos de operação de um equalizador linear adaptativo. Conforme comentado, o equalizador linear não apresenta um desempenho satisfatório em canais que não sejam de fase mínima ou que possuam nulos espectrais. Neste caso, a solução é a utilização de um equalizador não linear como o DFE. A figura 2.4 mostra a estrutura de um DFE em um sistema de comunicações.

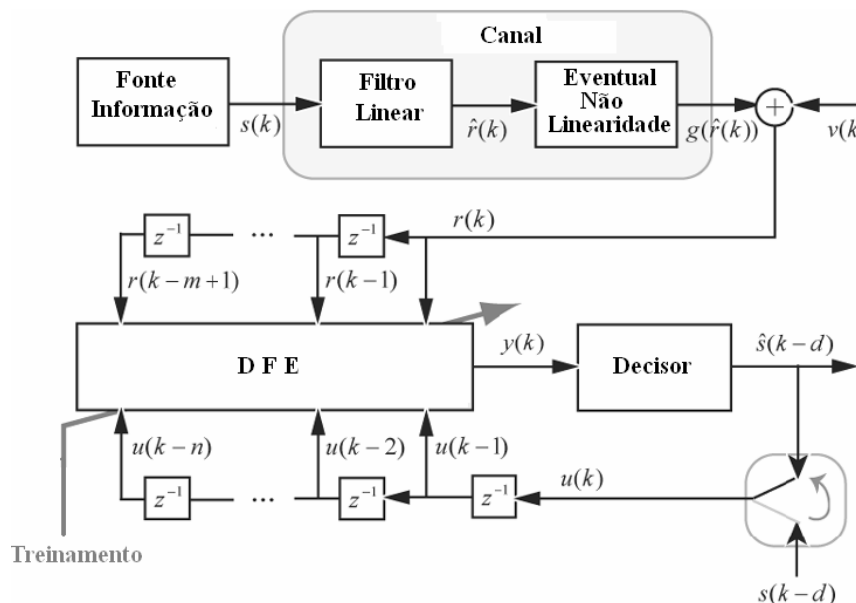


Figura 2.4. - Estrutura de um DFE para o sistema digital de comunicações – *Extraído e adaptado de (Choi, Lima e Haykin, 2005)*

A inclusão de uma não linearidade no dispositivo de decisão e a introdução do processo de treinamento que ajusta os valores dos coeficientes do filtro (período de aprendizagem), faz com que o conceito de redes neurais seja lembrado. De fato, a estrutura do filtro adaptativo apresentado é na verdade um *perceptron*. Note que as entradas são dados pelo sinal de recepção passando pelos respectivos atrasos do filtro, os pesos sinápticos são os valores dos coeficientes do filtro, e a função de ativação é dada pelo dispositivo de decisão. A figura 2.5 compara a estrutura de um *perceptron* e de um filtro adaptativo.

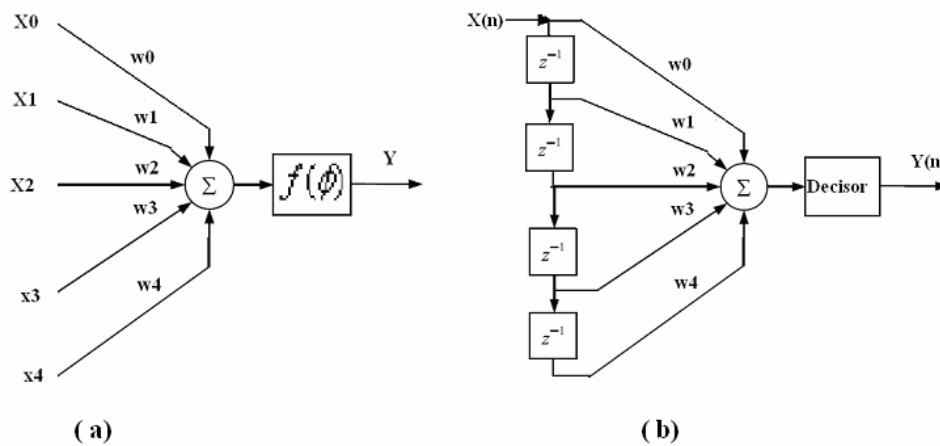


Figura 2.5- Comparação entre um *perceptron* e filtro adaptativo. (a) *Perceptron* (b) *Filtro adaptativo*

2.4 Neuro-Equalizadores

No item anterior vimos que a inclusão da não linearidade melhora a capacidade de mapeamento das características dos canais ou dos símbolos recebidos, tornando o equalizador mais adequado para lidar com canais variáveis no tempo onde nulos espectrais aparecem com frequência.

Considerando os dois pontos de vista de equalização, tanto o critério da máxima verossimilhança quanto o uso de filtragem, o uso de Redes Neurais para realizar equalização é plenamente justificado: As Redes Neurais Artificiais (RNA) são conhecidas pela sua capacidade de aproximar funções, no mapeamento de características não-lineares (Haykin, 1994). A capacidade das RNA na classificação de padrões através da formação de limites de decisões complexos e não-lineares é amplamente conhecida. Estas características motivaram a utilização de Redes Neurais na solução de problemas de equalização (Ibnkahla, 2000) e (Parisi *et al*, 1997). Outros trabalhos demonstraram que estruturas recorrentes, RNN, apresentam um desempenho superior àquelas de alimentação

direta, FNN (Kechriotis *et al*, 1994). Nas RNN, a informação contida nos símbolos previamente detectados é utilizada para melhorar a estimativa dos símbolos mais recentes. Arquiteturas mais simples, constituídas de apenas uma camada com um (Chen and Lima, 2004) ou mais neurônios (Ong *et al*, 1997) têm sido testadas e apresentaram bom desempenho frente a estruturas mais complexas de Multicamadas de *Perceptrons* (MLP) e Multicamadas Recorrentes de *Perceptrons* (RMLP).

2.4.1 Implementação do Neuro-Equalizador DFE.

Este trabalho apresenta um DFE implementado com um módulo de decisão não-linear do tipo tangente hiperbólica na fase de determinação dos coeficientes do mesmo. Esta evolução equivale a uma espécie de *perceptron* recorrente com função de ativação não-linear utilizadas por (Chen e Lima, 2004) e (Cortes *et al*, 2006) em problemas de equalização de canais fixos. Neste tipo de arquitetura, os coeficientes dos filtros direto e de realimentação formam os pesos sinápticos do *perceptron* recorrente. Para classificar os símbolos na saída do filtro direto é empregada uma função tangente hiperbólica durante a fase de treinamento (*soft-decision*) e uma função sinal na fase de teste (*hard-decision*). A figura 2.6 mostra o esquemático deste tipo DFE ou *perceptron* recorrente, indicando que estruturas simples de RNA podem realizar o trabalho de equalização de sinais com pouco esforço computacional. Para testar o desempenho destes equalizadores, foram utilizados modelos de canais discretos com três coeficientes e parâmetros de equalização do DFE: $m=3$, $n=2$ e $d=2$.

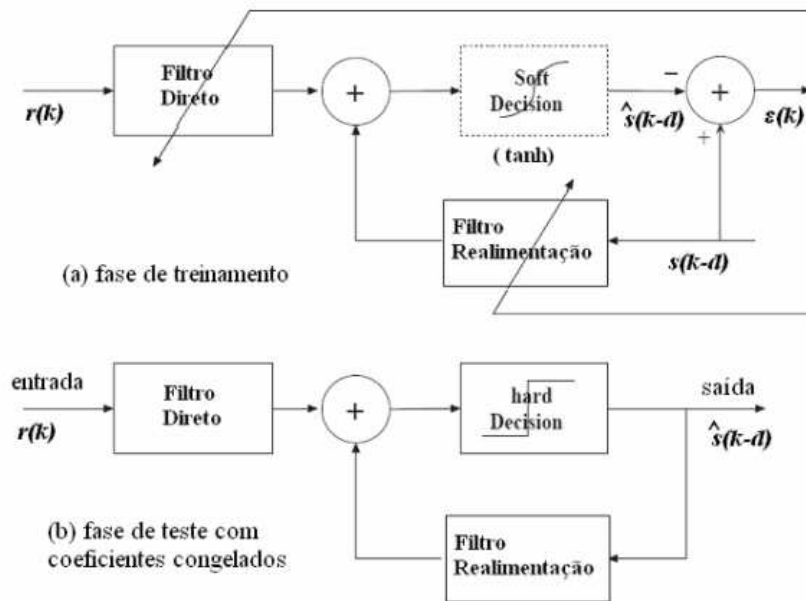


Figura 2.6 - Esquemático do DFE, (a) na fase de treinamento a função de ativação realiza “soft decision” através de uma função tangente hiperbólica (b) na fase de teste a função de ativação realiza um “hard decision” através de uma função sinal-adaptado de (Chen, Z. e de C. Lima, 2004).

O sinal derivado da fonte de informação $s(k)$ é um sinal 4-QAM banda base gerado nas simulações pela função `qammod` do MATLAB[®]. O sinal $s(k)$ passa pelo modelo de canal a ser estudado e é acrescido de ruído AWGN somente após passar pelo canal como é recomendado por (Tranter, *et al*, 2004), o sinal após os efeitos do canal chamado $\mathbf{r(k)}$ é então apresentado a entrada do neuro-equalizador. O sinal $\mathbf{r(k)}$ é convoluído com os coeficientes do filtro direto e apresentado na saída com um atraso $d=2$. Após passar pela função de decisão o sinal estimado $\mathbf{x(k)} = [r(k), r(k-1), r(k-2), u(k-1), u(k-2)]^T$ passa então pelos coeficientes do filtro de realimentação sendo chamado de $\mathbf{u(k)}$, portanto, fica claro que o sinal de entrada do *perceptron* recorrente pode ser expresso pelo vetor $\mathbf{x(k)}$ segundo a equação (2.3).

$$\mathbf{x(k)} = [r(k), r(k-1), r(k-2), u(k-1), u(k-2)]^T \quad (2.3)$$

Sendo W a matriz dos pesos dos coeficientes, a saída do DFE é dada pela equação (2.4).

$$\mathbf{y(k)} = W\mathbf{x(k)} \quad (2.4)$$

Na fase de treinamento os dados da resposta desejada são conhecidos a priori, portanto o sinal de realimentação $\mathbf{u(k)}$ deve ser o mesmo que o sinal da saída desejada, porém com o atraso devido ao filtro direto que no nosso caso é $d=2$. Assim $u(k) = s(k-2)$, portanto ao implementar a simulação os dois primeiros valores do vetor que representa $\mathbf{u(k)}$ devem ter o valor zero. Deste modo, o vetor $\mathbf{x(k)}$ é formado conforme equação (2.5).

$$\mathbf{x(k)} = [r(k), r(k-1), r(k-2), s(k-3), s(k-4)]^T \quad (2.5)$$

A saída $\mathbf{y(k)}$ passa pelo dispositivo de decisão que é uma função tangente hiperbólica, para resultar na estimativa do sinal desejado, lembrando que o número de atraso do filtro direto $d=2$ a estimativa da saída $\hat{\mathbf{s(k-2)}}$ é dada pela equação (2.6).

$$\hat{\mathbf{s(k-2)}} = \tanh[\mathbf{y(k)}] = \tanh[W\mathbf{x(k)}] \quad (2.6)$$

Os coeficientes W do DFE-Neural são atualizados pelo algoritmo de treinamento visando minimizar o sinal de erro $\varepsilon(k)$ dado pela equação (2.7).

$$\varepsilon(k) = \hat{\mathbf{s(k-2)}} - \mathbf{s(k-2)} \quad (2.7)$$

Na fase de operação ou teste os dados da resposta não são conhecidos a priori, portanto o sinal de realimentação $u(k)$ deve ser igual à estimativa do sinal de saída considerando o atraso $d=2$. Assim, $u(k) = \hat{s}(k-2)$ e neste caso, ao implementar a simulação, devemos ter o cuidado de observar que no primeiro e segundo símbolo de testes o valor de $u(k)$ ainda está relacionado com os dois últimos símbolos de treinamento. A partir do terceiro símbolo de teste podemos, portanto, definir o vetor $x(k)$ conforme a equação (2.8).

$$x(k) = [r(k), r(k-1), r(k-2), \hat{s}(k-3), \hat{s}(k-4)]^T \quad (2.8)$$

A saída $y(k)$ passa pelo dispositivo de decisão desta feita feita uma função sinal, para resultar na estimativa do sinal desejado, novamente lembrando que o numero de atraso do filtro direto é $d=2$ e a estimativa da saída $\hat{s}(k-2)$ é dada pela equação (2.9).

$$\hat{s}(k-2) = \text{sig}[y(k)] = \text{sig}[Wx(k)] \quad (2.9)$$

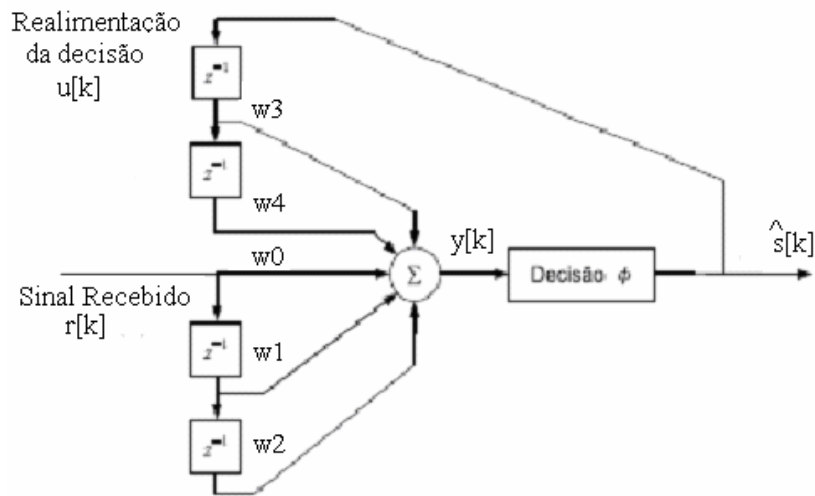


Figura 2.7 - DFE Implementado por um *perceptron* – adaptado de (Cortes, 2006)

Capítulo 3

Canais Variáveis no Tempo

Os Sistemas Lineares Variantes no Tempo (SLVT) podem ser modelados e simulados no domínio do tempo ou no domínio da frequência. A modelagem no domínio do tempo consiste em um filtro transversal linear cujos coeficientes são variáveis com o tempo. Este modelo é muito fácil de implementar para finalidades de simulação e é computacionalmente muito eficiente se a resposta ao impulso for relativamente curta. Muitos dos conceitos aplicados aos Sistemas Lineares Invariantes no Tempo (SLIT) se aplicam aos SLVT, porém com algumas diferenças importantes. Atenção particular deve ser dada à taxa de amostragem usada na simulação, uma vez que um aumento na taxa da amostragem é requerido devido à expansão da largura de faixa resultante das variações temporais. Um efeito da expansão da largura de faixa é o "espalhamento Doppler" . (Tranter *et al*, 2004).

Muitas das técnicas de análise nos domínios do tempo frequência usados nos SLIT podem ser usadas nos SLVT, com ligeiras modificações. Entretanto, é preciso cautela em simplificar os diagramas de bloco dos SLVT uma vez que estes não obedecem necessariamente todas as propriedades dos SLIT. Caso o Sistema Variante no Tempo seja linear, a aplicação da propriedade da superposição e a operação de convolução continuam válidas, entretanto, novamente é preciso cautela, pois as mesmas não são aplicáveis a Sistemas Não Lineares. Análise detalhada entre as propriedades dos SLIT e aplicações ou restrições destas nos SLVT, bem como, técnicas de simulação para sistemas não lineares podem ser encontradas em (M.C. Jeruchim *et al*, 2000) e em (Tranter *et al*, 2004).

Considerando no domínio do tempo a resposta ao impulso complexa $\tilde{h}(\tau)$ para um SLIT, utilizaremos a envoltória de $\tilde{h}(\tau)$, onde a variável τ representa a diferença entre o tempo em que o impulso é aplicado na entrada do sistema e o tempo em que a resposta ao impulso é medida na saída do sistema. A equação (3.1) indica a relação entre a saída do sistema $\tilde{y}(t)$ e a sua entrada $\tilde{x}(t)$:

$$\tilde{y}(t) = \int_{-\infty}^{+\infty} \tilde{h}(\tau) \tilde{x}(t - \tau) d\tau \quad (3.1)$$

Passando para o domínio da frequência encontramos a equação (3.2).

$$\tilde{Y}(f) = \tilde{H}(f)\tilde{X}(f) \quad (3.2)$$

Tomando a inversa de $\tilde{Y}(f)$ encontramos:

$$\tilde{y}(t) = \int_{-\infty}^{+\infty} \tilde{Y}(f) \exp(j2\pi ft) df = \int_{-\infty}^{+\infty} \tilde{X}(f) \tilde{H}(f) \exp(j2\pi ft) df \quad (3.3)$$

Os SLVT também são caracterizados no domínio de tempo por uma resposta ao impulso, porém com a forma $\tilde{h}(\tau, t)$, ou seja, o impulso é aplicado na entrada no tempo $t - \tau$ e a resposta é medida no tempo t . Uma vez que o sistema é variável no tempo, a resposta ao impulso mudará em função do tempo em que o impulso é aplicado, e do tempo em que a saída é medida. Para um sistema invariante no tempo, a resposta ao impulso será estritamente uma função do tempo decorrido e é representada simplesmente por $\tilde{h}(\tau)$. Por outro lado, a resposta ao impulso de um sistema variante no tempo é uma função de duas variáveis. Podemos resumidamente dizer que enquanto a resposta ao impulso de um SLIT depende apenas da diferença de tempo em que a saída é medida e o impulso é aplicado à entrada. A resposta ao impulso em um SLVT depende também de quando o impulso foi aplicado. (Tranter *et al*, 2004).

Matematicamente podemos expressar para um SLIT,

$$h(\tau, t1) = h(\tau, t2) = h(\tau) \quad (3.4)$$

Já para um SLVT

$$h(\tau, t1) \neq h(\tau, t2) \quad (3.5)$$

Tomando a transformada de Fourier de $\tilde{h}(\tau, t)$ achamos a função de transferência do SLVT, embora tenhamos o cuidado de observar que a função de transferência possui uma variação temporal.

$$\tilde{H}(f, t) = \int_{-\infty}^{+\infty} \tilde{h}(\tau, t) \exp(-j2\pi f\tau) d\tau \quad (3.6)$$

Na equação (3.6) a variável t descreve a natureza do sistema. Se as características do sistema são fortemente afetadas por t , temos um sistema com variação rápida. Geralmente, a natureza de variação temporal do sistema é modelada como um processo aleatório em t . Se o processo for estacionário, então as variações do tempo podem ser modeladas por uma função apropriada do

autocorrelação no domínio de tempo ou por uma densidade espectral de potência no domínio da frequência. A constante de tempo da função de autocorrelação ou a largura de faixa da densidade espectral de potência são os parâmetros principais que descrevem se o SLVT está variando lentamente ou rapidamente. (Tranter *et al*, 2004)

Para o domínio do tempo podemos expressar a relação entrada e saída de um SLVT através da integral de convolução:

$$\tilde{y}(t) = \int_{-\infty}^{+\infty} \tilde{h}(\tau, t) \tilde{x}(t - \tau) d\tau = \int_{-\infty}^{+\infty} \tilde{h}(t - \tau, t) \tilde{x}(\tau) d\tau \quad (3.7)$$

A análise em frequência para um SLVT é um pouco mais complicada que a análise no domínio do tempo.

Como ponto inicial vamos definir uma transformada de Fourier bidimensional onde o significado das frequências f_1 e f_2 será explicado adiante

$$\tilde{H}(f_1, f_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{h}(\tau, t) \exp(-j2\pi f_1 \tau - j2\pi f_2 t) d\tau dt \quad (3.8)$$

Ao definir esta transformada de Fourier bidimensional, assumimos que a energia da função da resposta ao impulso é finita de forma a assegurar a existência da transformada de Fourier. Porém, como será apresentado mais tarde o canal que modela a resposta ao impulso será tratado como um processo aleatório estacionário em t . Neste caso, a transformada de Fourier da resposta ao impulso pode não existir com respeito a t , entretanto a função de autocorrelação existirá e um procedimento apropriado deve definir a autocorrelação da resposta ao impulso com respeito à variável t . Daí então deve-se tomar a sua transformada de Fourier para obter a função de densidade espectral de potência do processo aleatório, que será a representação no domínio da frequência. (Tranter *et al*, 2004).

Na “função transferência” bidimensional definida em (3.8), a frequência f_1 é variável e é associada com a variável do tempo, podendo ser vistos como análogos à frequência f variável na função de transferência $H(f)$ de SLIT. Entretanto, a relação entrada/saída para SLVT no domínio da frequência, envolve uma convolução no domínio da frequência na segunda variável da função de transferência $H(f_1, f_2)$. Esta convolução explica o efeito da natureza do sistema de ser variável no tempo no domínio da frequência (Tranter *et al*, 2004). Se a entrada a um sistema de SLIT for um

tom em $f_c + f_0$ dado por

$$x(t) = \cos[2\pi(f_c + f_0)t] \quad (3.9)$$

Cuja envoltória complexa é

$$\tilde{x}(t) = A \exp(j2\pi f_0 t) \quad (3.10)$$

A relação entrada/saída no domínio da frequência é dada por

$$\tilde{X}(f) = A\delta(f - f_0) \quad (3.11)$$

e

$$\tilde{Y}(f) = \tilde{H}(f)\tilde{X}(f) = H(f_0)A\delta(f - f_0) \quad (3.12)$$

A envoltória complexa na saída do sistema é definida como

$$\tilde{y}(t) = A |\tilde{H}(f_0)| \exp[j2\pi f_0 t + \angle \tilde{H}(f_0)] \quad (3.13)$$

E o sinal passa-faixa no domínio do tempo é dado por

$$y(t) = A |\tilde{H}(f_0)| \cos[2\pi(f_c + f_0)t + \angle \tilde{H}(f_0)] \quad (3.14)$$

As relações apresentadas demonstram que para um SLIT quando a entrada do sistema é um tom complexo de frequência f_0 então o sistema produz na saída um tom na mesma frequência com a amplitude e fase na saída afetada pela resposta em amplitude e fase do SLIT para aquela frequência (Tranter *et al*, 2004).

Agora consideraremos o comportamento de um SLVT quando é aplicada a sua entrada o mesmo sinal dado em (3.9).

Após desenvolvimento matemático (Tranter *et al*, 2004) mostra-se que a densidade espectral de potência na saída do SLVT no domínio da frequência é dada então por

$$S_{\tilde{y}\tilde{y}}(f) = A^2 S_{\tilde{h}\tilde{h}}(f - f_0) \quad (3.15)$$

Onde $S_{\tilde{h}\tilde{h}}(f)$ é a média da densidade espectral de potência, obtida tomando-se a transformada de Fourier da autocorrelação média do sinal ao passar pelo SLVT.

A equação anterior mostra que a saída não consiste em um único tom, mas em de uma faixa de

freqüência. A figura 3.1. ilustra o ocorrido.

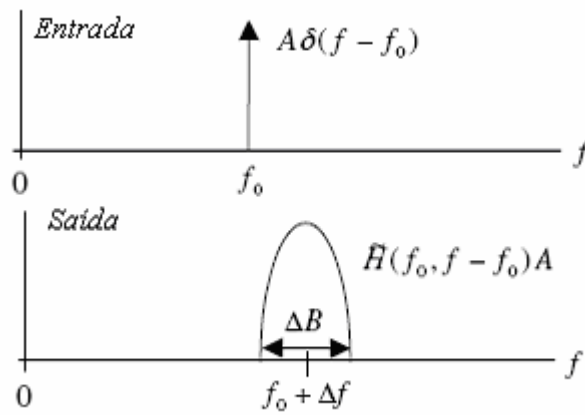


Figura 3.1 - Resposta de um SLVT a um tom aplicado à entrada com deslocamento em freqüência Δf espalhamento em freqüência ΔB – Fonte (M.C. Jeruchim, et al,2000).

Em geral, a saída de um SLVT responde a um tom de freqüência f_0 aplicado a sua entrada pode ter esta freqüência deslocada e também espalhada, este fenômeno nos canais de comunicação móveis é o que chamamos respectivamente de deslocamento Doppler e espalhamento Doppler, e é normalmente devido ao movimento relativo (velocidade relativa) entre as antenas transmissoras e receptoras ou por mudanças nas características do canal, devido, por exemplo, a multipercursos. (M.C. Jeruchim et al, 2000).

3.1 Modelagem como Processos Aleatórios

Normalmente as mudança nas características dos SLVT ocorrem aleatoriamente o que torna interessante a modelagem de SLVTs como processos aleatórios. Para modelos lineares as variações aleatórias das características do sistema em função do tempo podem ser obtidas pela resposta ao impulso, caracterizando a mesma como um processo aleatório em t . Tal aproximação é usada freqüentemente para modelar canais de comunicações variantes no tempo, e de fato será a abordagem utilizada neste trabalho nos três modelos de canais utilizados para simulação e que serão apresentados com maiores detalhes no capítulo 5.

Vamos considerar um sistema em que a saída é uma versão atenuada e atrasada da entrada, onde a atenuação varia aleatoriamente com o tempo.

O sistema é caracterizado por uma resposta do impulso:

$$h(\tau, t) = a(t)(\tau - t_0) \quad (3.16)$$

Onde $a(t)$ é a atenuação que varia aleatoriamente com o tempo, a saída do sistema é dada em função da entrada dada por

$$\tilde{y}(t) = \int_{-\infty}^{+\infty} \tilde{h}(\tau, t) \tilde{x}(t - \tau) d\tau = a(t) \tilde{x}(t - t_0) \quad (3.17)$$

A atenuação $a(t)$ pode ser modelada como um processo estacionário, o que irá permitir o uso da função de autocorrelação e conseqüentemente da função de densidade espectral de potência, além da utilização de um modelo de resposta ao impulso variante no tempo (Tranter *et al*, 2004).

Considerando, portanto o processo estacionário a função de autocorrelação é obtida por

$$R_{\tilde{y}\tilde{y}}(a) = E[\tilde{y}^*(t) \tilde{y}(t + a)] = E[\tilde{a}^*(t) \tilde{x}^*(t - t_0) \tilde{a}(t + a) \tilde{x}(t - t_0 + a)] \quad (3.18)$$

Como a entrada do sistema e a atenuação são variáveis não correlacionadas, podemos simplificar a expressão para a autocorrelação.

$$R_{\tilde{y}\tilde{y}}(a) = R_{\tilde{a}\tilde{a}}(a) R_{\tilde{x}\tilde{x}}(a) \quad (3.19)$$

Com a expressão simplificada da autocorrelação disponível, podemos encontrar a função de densidade espectral de potência na saída do sistema tomando a sua transformada de Fourier e utilizar o teorema da convolução.

$$S_{\tilde{y}\tilde{y}}(f) = S_{\tilde{a}\tilde{a}}(f) \otimes S_{\tilde{x}\tilde{x}}(f) \quad (3.20)$$

Se a entrada do sistema for uma exponencial complexa com fase aleatória, a convolução da equação anterior pode resultar em deslocamento e espalhamento de frequência, ou seja,

$$\tilde{x}(t) = A \exp(j2\pi f_0 t + j\theta(t)) \quad (3.21)$$

Sendo $\theta(t)$ uma variável uniformemente distribuída no intervalo $[-\pi, \pi]$, então as expressões da autocorrelação, densidade espectral de potência na entrada e densidade espectral de potência na saída são respectivamente

$$R_{\tilde{x}\tilde{x}}(a) = A^2 \exp(j2\pi f_0 a) \quad (3.22)$$

$$S_{\tilde{x}\tilde{x}}(f) = A^2 \delta(f - f_0) \quad (3.23)$$

$$S_{\tilde{y}\tilde{y}}(f) = A^2 S_{\tilde{h}\tilde{h}}(f - f_0) \quad (3.24)$$

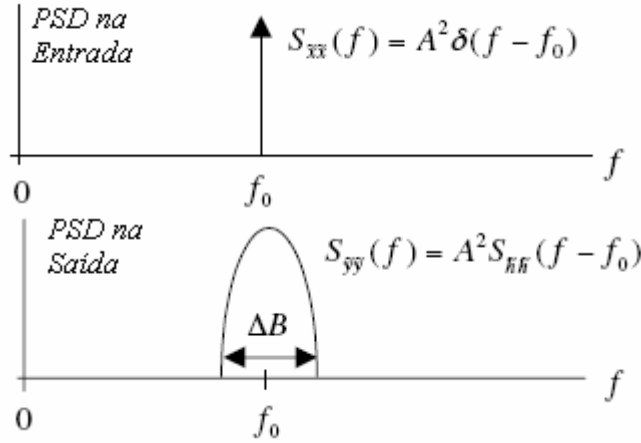


Figura 3.2 - Espalhamento de freqüências em SLVT – fonte (Tranter, et al, 2004).

3.2 Simulação de SLVT

Dado uma resposta ao impulso de um canal SLVT com uma determinada banda passante, pode-se derivar um modelo de simulação utilizando o teorema da amostragem partindo da relação entrada / saída que é dada pela integral de convolução:

$$\tilde{y}(t) = \int_{-\infty}^{+\infty} \tilde{h}(\tau, t) \tilde{x}(t - \tau) d\tau \quad (3.25)$$

Pelo teorema da amostragem é possível representar um sinal passa-baixa $w(\tau)$ limitado em banda por B Hz em termos de seus valores amostrados por:

$$\tilde{y}(\tau) = \sum_{n=-\infty}^{+\infty} \tilde{x}(t - nT) \tilde{g}_n(t) \quad (3.26)$$

$$\tilde{g}_n(t) = \int_{-\infty}^{+\infty} \tilde{h}(\tau, t) \frac{\text{sen}(2\pi B(\tau - nT))}{2\pi B(\tau - nT)} d\tau \quad (3.27)$$

Onde $1/T$ é a taxa de amostragem determinada pela taxa de Nyquist igual a $2B$, o valor mínimo de $2B$ é escolhido para que o custo computacional do modelo seja mínimo (*Tranter, et al, 2004*). Usando a representação acima e considerando que $\tilde{x}(t - \tau) = \tilde{w}(\tau)$, podemos substituir $\tilde{x}(t - \tau)$ na integral de convolução.

$$\tilde{x}(t - \tau) = \sum_{n=-\infty}^{+\infty} \tilde{x}(t - nT) \frac{\text{sen}(2\pi B(\tau - nT))}{2\pi B(\tau - nT)} \quad (3.28)$$

Assim,

$$\tilde{y}(\tau) = \int_{-\infty}^{+\infty} \tilde{h}(\tau, t) \left\{ \sum_{n=-\infty}^{+\infty} \tilde{x}(t - nT) \frac{\text{sen}(2\pi B(\tau - nT))}{2\pi B(\tau - nT)} \right\} d\tau \quad (3.29)$$

$$\tilde{y}(\tau) = \sum_{n=-\infty}^{+\infty} \tilde{x}(t - nT) \int_{-\infty}^{+\infty} \tilde{h}(\tau, t) \left\{ \frac{\text{sen}(2\pi B(\tau - nT))}{2\pi B(\tau - nT)} \right\} d\tau \quad (3.30)$$

De onde podemos escrever

$$\tilde{y}(\tau) = \sum_{n=-\infty}^{+\infty} \tilde{x}(t - nT) \tilde{g}_n(t) \quad (3.31)$$

onde

$$\tilde{g}_n(t) = \int_{-\infty}^{+\infty} \tilde{h}(\tau, t) \frac{\text{sen}(2\pi B(\tau - nT))}{2\pi B(\tau - nT)} d\tau \quad (3.32)$$

3.2.1 Modelo associado a filtro transversal linear

Reconhecemos a equação (3.31) como um filtro transversal linear cuja saída é $\tilde{y}(\tau)$, entrada é $\tilde{x}(\tau)$ e os pesos dos coeficientes dados por $\tilde{g}_n(\tau)$, assim representado na figura 3.3.

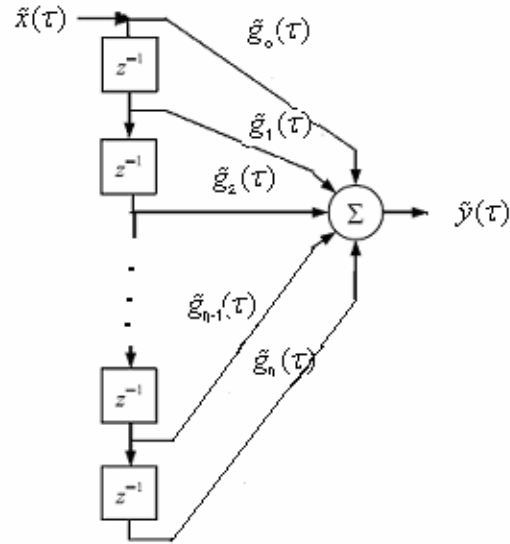


Figura 3.3 - Modelo de canal variável no tempo por filtro transversal linear

3.2.2 Simplificação do Modelo

O modelo proposto pode ser simplificado a luz de diversas considerações e aproximações, freqüentemente podemos considerar os ganhos dos coeficientes como não-correlacionados, assim conforme (Tranter, et al, 2004) podemos considerar a aproximação:

$$\tilde{g}_n(t) = T\tilde{h}(nT, t) \quad (3.33)$$

Nesta aproximação, os ganhos dos coeficientes representam amostras da resposta ao impulso do SLVT, onde as amostras são realizadas na resposta ao impulso que naturalmente varia com o tempo.

A segunda aproximação é obtida limitando o número de coeficientes da resposta ao impulso, desprezando os coeficientes com valores menores, em outras palavras, se $E\{|h(\tau, t)|^2\} \rightarrow 0$ e $|\tau| \geq mT$, então:

$$\tilde{y}(\tau) = \sum_{n=-\infty}^{+\infty} \tilde{x}(t - nT) \tilde{g}_n(t) = \sum_{n=-m}^{+m} \tilde{x}(t - nT) \tilde{g}_n(t) \quad (3.34)$$

Assim, o modelo do filtro transversal linear passa a ter um número finito de coeficientes, podendo ser representado por um filtro FIR com coeficientes variáveis no tempo (Tranter, et al, 2004).

3.2.3 Geração dos coeficientes do filtro

Os coeficientes dos filtros podem ser considerados como processos aleatórios estacionários com função de densidade de probabilidade Gaussiana de média zero, podendo assumir diversos valores de variância e com funções de densidades espectral de potências arbitrariamente definidas, porém idênticas, neste caso, os coeficientes podem ser gerados filtrando um ruído branco gaussiano, a função de transferência do filtro é escolhida de tal maneira que esta produza o espalhamento Doppler desejado, conforme se observa na equação abaixo.

$$S_{\tilde{g}\tilde{g}}(f) = S_{\tilde{d}}(f) = S_{\tilde{\omega}\tilde{\omega}}(f) |\tilde{H}(f)|^2 = |\tilde{H}(f)|^2 \quad (3.35)$$

Onde $S_{\tilde{\omega}\tilde{\omega}}(f)$ é a densidade de potência do processo de entrada que é o ruído branco gaussiano e portanto pode ser normalizado para 1, e $S_{\tilde{g}\tilde{g}}(f)$ é a densidade espectral de potência que fará com que os coeficientes dos filtros variem de forma a produzir o espalhamento Doppler desejado (Tranter, et al, 2004).

Portanto, considerando um canal de comunicações com uma determinada largura de faixa, podemos definir o espalhamento Doppler como sendo a frequência de corte do filtro que estará relacionado a largura de banda do canal.

Um filtro, por exemplo, Butterworth, pode ser usado para gerar a densidade espectral de potência dos coeficientes do filtro que resultará no espalhamento Doppler desejado. Um filtro de Butterworth de segunda ordem tem sido usado extensivamente em diversos trabalhos (Cowan e S. Semnani , 1998), (Ling e Proakis, 1985), , (Choi, Lima e Haykin , 2005), (Choi, Bouchard e Yeap, 2005).

3.3 Caracterização Estatística de Canais Variáveis no Tempo

Sabemos que a frequência Doppler é definida como:

$$f_d = \frac{v}{\lambda} \cos \alpha \quad (3.36)$$

Onde f_d é a frequência Doppler, v é a velocidade relativa entre transmissor e receptor, λ é o comprimento de onda da frequência de transmissão e α é o ângulo formado entre a horizontal e a onda incidente no receptor.

A situação mais crítica ocorre quando o ângulo α é nulo, neste caso temos a frequência Doppler caracterizada como a razão entre a velocidade relativa entre o transmissor e receptor.

Portanto, se, por exemplo, desejarmos simular um canal de telefonia móvel, devemos usar a frequência Doppler como frequência de corte do filtro e a taxa de amostragem como no mínimo duas vezes a largura de banda do canal.

Outro aspecto interessante que pode ser observado pela equação (3.36) é que quanto maior a frequência de operação, menor será o comprimento de onda e conseqüentemente maior a frequência Doppler. Consideremos por exemplo o sistema GSM de telefonia móvel que pode operar em frequências na faixa de 900 MHz e 1,8 GHz, um telefone móvel movendo-se a uma determinada velocidade experimentará uma frequência Doppler duas vezes maior em 1,8 GHz que em 900 MHz.

Entretanto, a noção de quanto a variação do canal dificultará uma dada comunicação digital não fica clara simplesmente definindo o espalhamento Doppler. Naturalmente que intuitivamente quanto maior for o espalhamento Doppler maior é a variação das características de um canal, o que se reflete claramente quando consideramos um SLIT que possui espalhamento Doppler nulo.

É evidente a necessidade de definição de outros parâmetros, mas antes é interessante chamar atenção para o fato de que a caracterização do canal depende do domínio de análise, se tempo ou frequência. Basicamente, no domínio da frequência temos o desvanecimento plano e o desvanecimento seletivo e no domínio do tempo temos o desvanecimento rápido e o desvanecimento lento (Haykin, 2001).

Chegaremos à definição de dois parâmetros fundamentais para a classificação dos canais, são eles: A banda de coerência para análise no domínio do tempo, e o tempo de coerência para análise no domínio da frequência. Para isto, vamos analisar o comportamento do canal no domínio do tempo e frequência.

Considere um impulso aplicado a um canal dispersivo. Como este canal não é ideal, ou seja, não tem largura de banda infinita, após a aplicação do impulso, ainda haverá sinal na saída do canal

durante um intervalo de tempo, mesmo considerando o atraso natural do canal. Isto significa que o receptor receberá uma densidade de potência do sinal durante este intervalo de tempo. Se, além disso, considerarmos que existem multipercursos e que o mesmo pulso chegará com atenuações e atrasos diferentes no receptor, teremos no receptor um perfil de sinal onde existem variações construtivas e destrutivas, como resultado teremos um aspecto “serrilhado” ao sinal recebido. A figura 3.4 ilustra esta situação, onde $S(\tau)$ é chamado perfil de intensidade do sinal na recepção ou espectro de potência de atraso.

Para um simples impulso, o tempo T_m , entre o primeiro e o último componente do sinal recebido representa o “atraso máximo”. Este intervalo de tempo é determinado considerando o instante em que o sinal decorrente de dispersão e multipercursos chegue ao receptor com valores normalmente determinados em um limiar de -10 dB ou -20 dB abaixo do componente de maior potência. (Sklar,2001)

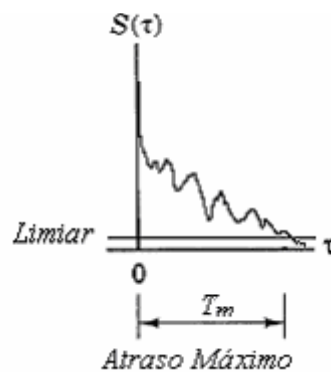


Figura 3.4 - Perfil de intensidade de multipercursos

A relação entre a duração do símbolo T_s e o atraso do canal T_m , resultam em duas categorias de degradação com relação às componentes de frequência: Se $T_m > T_s$ decorre que os componentes de multipercursos recebidos de um símbolo se estendem além da duração do tempo do símbolo transmitido resultando em ISI. Neste caso, ocorre o chamado desvanecimento seletivo em frequência.

Por outro lado, se $T_m < T_s$, todos os componentes de multipercursos recebidos de um símbolo chegam dentro da duração do símbolo, nesta situação não ocorre ISI, pois não há influência de símbolos anteriores no símbolo atual, mas por outro lado, o símbolo sofre degradação devido a interferências construtivas ou destrutivas o que pode levar o detector a uma decisão errada. Nesta situação, ocorre o chamado desvanecimento plano em frequência.

Entretanto é estranho estarmos nos referindo ao comportamento do desvanecimento em frequência

com uma análise temporal do sinal. Portanto, para que a situação fique mais clara, é interessante caracterizarmos o espalhamento temporal no domínio da frequência.

Tomando a transformada de Fourier de $S(\tau)$, encontraremos a função de correlação da diferença de frequência $R(\Delta f)$, onde a figura 3.5 mostra a função $R(\Delta f)$. Portanto, o espalhamento temporal pode ser visto como se fosse o resultado de um processo de filtragem em frequência do sinal. O conhecimento de $R(\Delta f)$ ajuda a identificar a correlação entre dois sinais recebidos que estão espaçados por dois valores de frequências que é dado por $\Delta f = f_1 - f_2$. A função $R(\Delta f)$ pode ser encontrada transmitindo um par de sinais senoidais separados em frequência por Δf , fazendo a correlação cruzada dos espectros dos dois sinais recebidos. Repetindo-se o processo diversas vezes para separações de frequência Δf diferentes, fazendo uma varredura de $\Delta f = 0$ e valores positivos e ou negativos, uma vez que pelo gráfico nota-se que $R(\Delta f)$ é uma função par. Quanto maior for a correlação entre valores de frequências mais afastados, mais próximas são as características do canal de um canal ideal.

A banda de coerência, B_c , é uma medida estatística da faixa de frequências a qual o canal se aproxima de um canal ideal, ou seja, a resposta do canal aos componentes espectrais tem aproximadamente o mesmo ganho e fase aproximadamente linear, significando que o canal possui uma forte correlação em amplitude. Isto é, na faixa de frequência correspondente a B_c , o sinal é afetado de maneira semelhante, de maneira que aproximadamente podemos dizer que $B_c < \frac{1}{T_s} \approx W$

(Sklar, 2001)

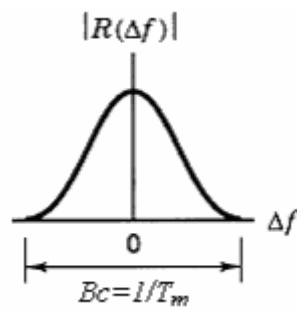


Figura 3.5 - Função de correlação da diferença de frequência $R(\Delta f)$ e banda de coerência do canal B_c .

Um canal é dito seletivo em frequência se $B_c < \frac{1}{T_s} \approx W$, onde a taxa de símbolo $\frac{1}{T_s}$, é normalmente igual a taxa de sinalização ou a banda do sinal W . Na prática, W pode ser diferente de $\frac{1}{T_s}$ devido

aos filtros no transmissor e receptor ou a técnica de modulação empregada. O desvanecimento seletivo ocorre quando as componentes espectrais de potência são afetadas diferentemente pelo canal, componentes fora da banda de coerência serão afetados diferentemente, quando comparados com os componentes que caem dentro da banda de coerência. (Sklar, 2001).

A dispersão do sinal e a banda de coerência caracterizam as propriedades de espalhamento temporal do canal. Entretanto, não fornecem nenhuma informação sobre a natureza das variações do canal causadas pelo movimento relativo do transmissor e receptor ou por variações dos obstáculos no canal. Em comunicações móveis, o canal é variante no tempo devido às mudanças nos caminhos de propagação entre transmissor e receptor. Tais mudanças ocasionam variações na amplitude e fase do sinal no receptor. Se todos os espalhamentos que ocorrem no canal forem considerados estacionários, enquanto não houver nenhum movimento relativo entre transmissor e receptor ou nenhuma mudança nos obstáculos, ou condições do canal como, por exemplo, variações de condições atmosféricas, a amplitude e fase no sinal recebido serão constantes e o canal poderá ser considerado como invariante no tempo, também chamado fixo. Caso exista algum movimento relativo entre transmissor e receptor, mudança de obstáculos ou de condições do canal, haverá variação dos valores de amplitude e fase do sinal e o canal é considerado como variante no tempo.

Para caracterizar melhor as variações temporais do canal é utilizada a função de correlação de espaçamento de tempo $R(\Delta t)$, que representa a autocorrelação da resposta do canal a um tom senoidal. Esta função especifica a faixa temporal na qual há correlação entre a resposta do canal a um tom senoidal enviado no tempo t_1 e a resposta à outro tom de mesma frequência e fase enviado no tempo t_2 , onde $\Delta t = t_2 - t_1$. O tempo de coerência T_c é a faixa de tempo na qual a resposta do canal se comporta como se fosse invariante. Note que para medir a banda de coerência, foram usadas várias frequências, em contrapartida, para medir o tempo de coerência é utilizado a mesma frequência enviada em intervalos de tempo distintos, A função $R(\Delta t)$ e o parâmetro T_c fornecem informações sobre a velocidade de desvanecimento do canal. (Sklar, 2001). Note que para um canal ideal invariante no tempo, $R(\Delta t)$ seria uma constante e T_c seria infinito. A figura 3.6. mostra a função de correlação de espaçamento de tempo $R(\Delta t)$.

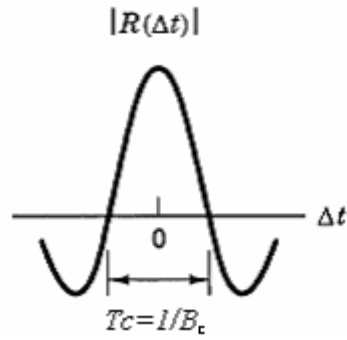


Figura 3.6 - função de correlação de espaçamento de tempo $R(\Delta t)$ e tempo de coerência T_c .

A natureza de variação no tempo para a velocidade do desvanecimento no canal pode ser visto em termos de duas categorias: desvanecimento rápido e desvanecimento lento.

O desvanecimento rápido ocorre em canais em que $T_c < T_s$, onde T_c é o tempo de coerência e T_s é a duração do símbolo transmitido. O desvanecimento rápido descreve uma condição onde o tempo em que um canal age de forma correlacionada é menor que a duração de um símbolo. Portanto, pode ser esperado que durante a transmissão de um símbolo, o canal mudará suas características de maneira a distorcer a forma do símbolo, analogamente à ISI, onde a distorção ocorre devido as componentes do sinal não estarem correlacionadas no tempo. Por outro lado, o desvanecimento lento ocorre em canais em que $T_c > T_s$. O desvanecimento lento descreve uma condição na qual o tempo em que o canal age de forma correlacionada é maior que duração de um símbolo. Assim pode-se esperar que o canal não apresente mudanças significativas de suas características no período de duração do símbolo e, portanto, não haverá degradação significativa na forma do símbolo (Sklar, 2001).

Do exposto na teoria acima, vemos que alguns parâmetros são importantes no estudo da variação de canais, a frequência Doppler f_d , a banda de coerência B_c , o tempo de coerência T_c e o espalhamento temporal T_m . Porém, em termos de condições de variações, estes parâmetros só significam alguma coisa quando comparados com o tempo de duração do símbolo T_s . Assim, o mesmo canal pode ser considerado como de desvanecimento rápido quando submetido a uma determinada taxa de transmissão, e ser considerado de desvanecimento lento para uma outra taxa de transmissão maior. Outro fator importante, inclusive na simulação, é a taxa de amostragem. Esta observação é extremamente importante para este trabalho uma vez que em todas as simulações é utilizada a taxa de uma amostra por símbolo, dado que o equalizador é de espaçamento simbólico, maiores detalhes quanto às simulações e modelo dos canais são dados no capítulo 5.

Capítulo 4

Algoritmo de Gauss-Newton

A utilização e justificativa do algoritmo de Gauss-Newton no treinamento do neuro-equalizador DFE para canais fixos pode ser encontrada em detalhes em (Cortes, 2006), (Cortes, *et al*, 2006) e em (Ludwig e Lima, 2005). Neste capítulo, veremos os conceitos mais importantes do método para possibilitar continuidade da leitura e compreensão do restante do trabalho.

Como foi visto no capítulo 2, os pesos dos coeficientes dos equalizadores adaptativos devem ser determinados por algoritmos de otimização. Métodos baseados no gradiente apresentam taxas de convergência de 1ª ordem e podem ser muito lentos para uma aplicação em equalizadores. Os métodos de Newton por sua vez garantem uma taxa de convergência de 2ª ordem, porém apresentam o inconveniente do cálculo da Hessiana e de sua inversa o que quando implementado em algoritmo exige extremo esforço computacional. Para contornar os inconvenientes do cálculo da Hessiana e de sua inversa encontramos na literatura uma grande quantidade de métodos alternativos chamados Quasi-Newton que se aproximam da 2ª ordem na taxa de convergência e contornam de alguma forma, o inconveniente do cálculo da Hessiana e sua inversa, na maioria das vezes através de estimativas das mesmas por métodos recursivos.

O método de Gauss-Newton leva este nome por ser devido a uma proposta de modificação do método de Newton pelo matemático alemão Carl Friedrich Gauss.

A iteração de Newton é dada pela equação (4.1).

$$x(k+1) = x(k) - [\nabla^2 f(x(k))]^{-1} \nabla f(x(k)) \quad (4.1)$$

Considerando que:

$$\nabla f(x) = J(x)^T f(x) \quad (4.2)$$

Onde $f(x)$ é a função custo a ser minimizada e $J(x)$ é a matriz Jacobiana definida por

$$J(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \dots & \frac{\partial f_1(x)}{\partial x_n} \\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \dots & \frac{\partial f_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(x)}{\partial x_1} & \frac{\partial f_n(x)}{\partial x_2} & \dots & \frac{\partial f_n(x)}{\partial x_n} \end{pmatrix} \quad (4.3)$$

A Hessiana pode ser escrita em termos da matriz Jacobiana por

$$\nabla^2 f(x) = J(x)^T J(x) + \sum_{i=1}^m f_i(x) \nabla^2 f_i(x) \quad (4.4)$$

O termo $\sum_{i=1}^m f_i(x) \nabla^2 f_i(x)$ normalmente é muito menor do que o termo $J(x)^T J(x)$ de forma que é possível desprezar o primeiro. Esta aproximação dá origem ao método de Gauss-Newton cuja iteração é dada pela equação (4.5).

$$x(k+1) = x(k) - [J(k)^T J(k)]^{-1} J(k)^T f(x(k)) \quad (4.5)$$

4.1 Aplicação do Algoritmo de Gauss-Newton no Neuro-equalizador.

O algoritmo de Gauss-Newton é uma aproximação do método de Newton, ou seja, o objetivo é a minimização pelo método dos mínimos quadráticos de uma determinada função custo. Como foi visto no início do capítulo, estamos interessados em encontrar os valores otimizados para os coeficientes do neuro-equalizador adaptativo, logo a função custo neste problema é o erro na determinação destes coeficientes dado pela equação (4.6).

$$e[\omega(k)] = \hat{s}(k) - s(k-d) \quad (4.6)$$

A matriz Jacobiana é calculada na fase de treinamento ou na fase de teste quando houver tentativa de rastreamento dos coeficientes. Nestas duas situações, $\hat{s}(k)$ é o sinal na saída do detector operando em “soft decision” cuja função de não linearidade na saída do detector é a tanh, logo a expressão do erro é dada por:

$$e[\omega(k)] = \tanh[Wx(k)] - s(k-d) \quad (4.7)$$

Cada linha da matriz Jacobiana corresponde a cada símbolo de treinamento m , para o vetor de pesos da iteração k . Portanto, a matriz Jacobiana tem como número de linhas o número de símbolos de treinamento e como número de colunas o número de coeficientes do equalizador. Cada linha da matriz Jacobiana pode ser obtida de:

$$\frac{\partial e(m)}{\partial \vec{\omega}} = \{1 - \tanh^2[Wx(k)]\}x(k) \quad (4.8)$$

Considerando-se que na aplicação deste trabalho o DFE tem cinco coeficientes, a matriz Jacobiana $J(k)$ pode ser expressa por:

$$J(k) = \begin{pmatrix} \frac{\partial e(1)}{\partial \omega_1} & \frac{\partial e(1)}{\partial \omega_2} & \dots & \frac{\partial e(1)}{\partial \omega_5} \\ \frac{\partial e(2)}{\partial \omega_1} & \frac{\partial e(2)}{\partial \omega_2} & \dots & \frac{\partial e(2)}{\partial \omega_5} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e(5)}{\partial \omega_1} & \frac{\partial e(5)}{\partial \omega_2} & \dots & \frac{\partial e(5)}{\partial \omega_5} \end{pmatrix}_{\omega=\omega(k)} \quad (4.9)$$

Os coeficientes são atualizados conforme a iteração do método de Gauss-Newton, fornecidos por:

$$\omega(k+1) - \omega(k) = \Delta\omega = -\{[J(k)^T J(k)]^{-1} J(k)^T\} \{\hat{s}(k) - s(k-d)\} \quad (4.10)$$

Portanto podemos resumir a parte do algoritmo de treinamento do equalizador como no quadro abaixo:

```

Defina r vetor com 5 elementos
Defina sr vetor de 2 elementos ( símbolos de realimentação)
Defina w vetor de 5 elementos inicialmente com valores nulos ( Pesos iniciais do
perceptron)
Defina nbittrain = Número de símbolos de treinamento
Defina iter (Número de iterações do método para o cálculo da Jacobiana)
Defina m como a seqüência de símbolos de treinamento conhecida

Faça k de 1 a iter
  Faça i de 1 a nbittrain
    u = [r(i:i+2) sr] ( é o sinal de entrada do perceptron)
    y =  $\tanh(\text{real}(\mathbf{w} * \mathbf{u}')) + j * \tanh(\text{imag}(\mathbf{w} * \mathbf{u}'))$  ( sinal de saída do perceptron)
    temp = [m(i) sr] ( vetor temporário para calculo do novo sinal de realimentação)
    sr = temp(1:2) ( utiliza-se o sinal de entrada como referência)
    err(i) = m(i)-y ( Calcula o erro)

    J(i,:) = [((1-y2)*conj(u)).'] ( Calcula a Jacobiana)

  Fim do Faça
  w=w+(Inverte-Matriz(J'*J)*J'*err);
Fim do Faça

```

Analisando o pseudocódigo acima, se verifica que o algoritmo proposto trabalha em lotes na atualização dos coeficientes, diferindo de outros algoritmos que normalmente atualizam os coeficientes símbolo a símbolo. Este fato pode representar uma desvantagem sobre certos aspectos, mas por outro lado representa uma grande economia em termos de custos computacionais, uma vez que reduz consideravelmente o número de operações com atualizações dos coeficientes.

A análise assintótica do algoritmo revela que a ordem de complexidade computacional é dada por $O(l)$ onde l é o número de símbolos tratados pelo algoritmo. (Cortes, 2006), (Cortes, *et al*, 2006) e em (Ludwig e Lima, 2005).

Para a utilização do algoritmo para rastreamento do canal na fase de teste, pequenas alterações são necessárias ao algoritmo mostrado acima. Na fase de teste a seqüência de treinamento conhecida não existe mais e deve ser substituída pela estimativa do sinal. Neste caso, o sinal de erro passa a ser calculado como a diferença entre os sinais de saída do “*soft decision*” pelo símbolo estimado equivalente, valor que é proveniente da saída de um “*hard decision*” (naturalmente esta mesma saída é a estimativa dos símbolos). Considerando que a decisão é feita corretamente, a

probabilidade de erro é baixa, o que se torna um excelente recurso para que o equalizador continue se ajustando às características do canal mesmo na fase de teste. Obviamente que tal ajuste só ocorre satisfatoriamente até determinadas condições de variação do canal como será visto no Capítulo 6.

Em (Cortes, 2006), (Cortes, *et al*, 2006) e em (Ludwig e Lima, 2005) é mostrado que somente duas iterações são suficientes para a atualização dos procedimentos. Este dado foi confirmado empiricamente nas simulações com canais variantes no tempo e será repetido neste trabalho.

Capítulo 5

Procedimentos de Teste

O procedimento de teste visa obter valores de parâmetros importantes do comportamento do equalizador como BER e diagramas de espalhamento (*Scatter Plots*).

Foi utilizada simulação em computador utilizando o software MATLAB[®], aplicando técnicas de simulação de Monte Carlo. Os dados de testes foram gerados por fontes aleatórias adequadas e os testes repetidos o suficiente para fosse possível obter dados consistentes estatisticamente. O número de iterações ou repetições de cada teste depende do modelo do canal testado e do número de símbolos utilizados para treinamento e testes.

Foram executados testes em diversos canais cuja modelagem será detalhada neste capítulo. Os números de símbolos de treinamento e de testes variam conforme as características de cada canal simulado. Normalmente o parâmetro de interesse é o valor de BER em função da SNR para cada canal de teste. Basicamente, foram realizados testes variando a relação sinal ruído dos canais e medindo a BER a intervalos regulares de SNR. As variações de SNR foram realizadas por faixas de valores de acordo com o procedimento adotado nos artigos que foram utilizados para comparação.

Os modelos de canais utilizados foram extraídos de artigos que utilizamos como referências. Conforme veremos no próximo item neste capítulo, os resultados obtidos foram comparados com os resultados destes artigos provenientes de outras topologias de neuro-equalizadores que foram apresentadas nestes artigos como topologias consagradas ou promissoras.

5.1 Modelos de Canais

5.1.1 Canal modelo 1

Este modelo de canal discreto consiste de três coeficientes com defasagens iguais ao período de um símbolo. Este modelo é descrito pela seguinte função de transferência:

$$H_1(z) = a_0(t) + a_1(t)z^{-1} + a_2(t)z^{-2} \quad (5.1)$$

Onde os coeficientes $a_i(t)$ são variáveis no tempo e gerados através da passagem de um sinal AWGN por um filtro de Butterworth de segunda ordem com frequência de corte ou largura de

banda de 3 dB normalizada de acordo com o espalhamento Doppler desejado.

Este mesmo modelo de canal foi utilizado nos trabalhos de (Ling e Proakis, 1985), e (Choi, Lima, e Haykin, 2005). Os modelos propostos nestes artigos foram canais de largura de banda entre 2 KHz e 3 KHz , considerando a taxa de transmissão de 2400 símbolos/seg e espalhamento Doppler de 0,5Hz e 1 Hz .

Um ponto relevante que deve ser considerado para simulação na geração dos coeficientes deste canal é que o produto da variância do sinal AWGN com o ganho do filtro de Butterworth deve ser controlado de forma que os coeficientes do canal sejam normalizados para que, ao longo do tempo, a média do ganho da função de transferência do canal fique em torno de 0 dB. Desta maneira, os coeficientes refletem apenas variações do canal, incluindo os nulos na função de transferências, sem contudo introduzir um fator de atenuação alto e sempre presente, o que levaria o símbolo a não ser reconhecido, mesmo em intervalos de tempo onde as condições de resposta em frequência se apresentasse plana, este procedimento seria equivalente a um controle automático de ganho em um sistema real. Este aspecto foi avaliado e aplicado em todas as simulações deste modelo de canal.

A figura 5.1. mostra a variação dos valores dos coeficientes em função do tempo de amostragem para o espalhamento Doppler de 0,5Hz e a figura 5.3 mostra a variação dos valores dos coeficientes em função do tempo de amostragem para o espalhamento Doppler 1 Hz .

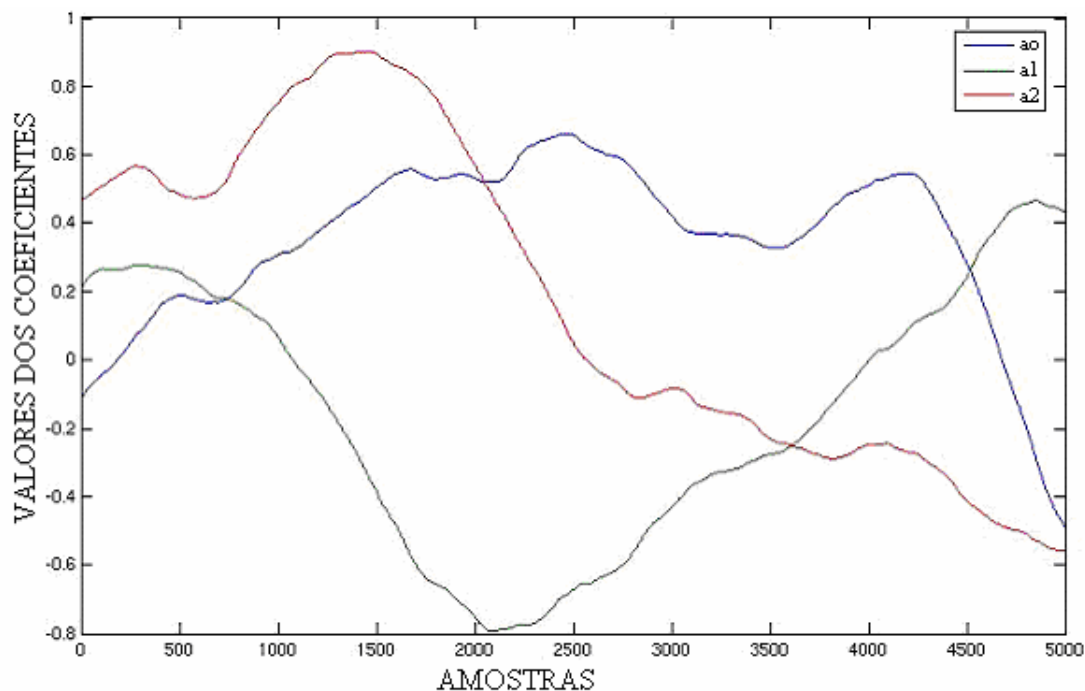


Figura 5.1 – Canal modelo 1 , variação dos coeficientes para espalhamento Doppler 0,5 Hz.

A visualização da variação dos coeficientes do canal não fornece intrinsecamente a idéia de quanto severo é este canal para a detecção correta dos símbolos, ou de como os símbolos se apresentam no receptor. Para que se possa melhorar esta percepção neste e nos outros modelos de canais, são apresentados os diagramas de espalhamento “*scatter plots*” dos símbolos recebidos. Obviamente estes diagramas se referem aos símbolos antes do equalizador. A figura 5.2 apresenta os diagramas de espalhamento para o canal modelo 1 com espalhamento Doppler de 0,5 Hz, enquanto a figura 5.4 apresenta os diagramas de espalhamento para o canal modelo 1 com espalhamento Doppler de 1 Hz, em ambos os casos a simulação é realizada considerando as SNR no canal de 12 e 15 dBs. Note que estes valores de SNR já se encontram no final da faixa testada e ainda assim a quantidade de símbolos invadindo a região do outro é muito grande, o que tornaria inviável a recuperação da mensagem sem utilização da técnica de equalização.

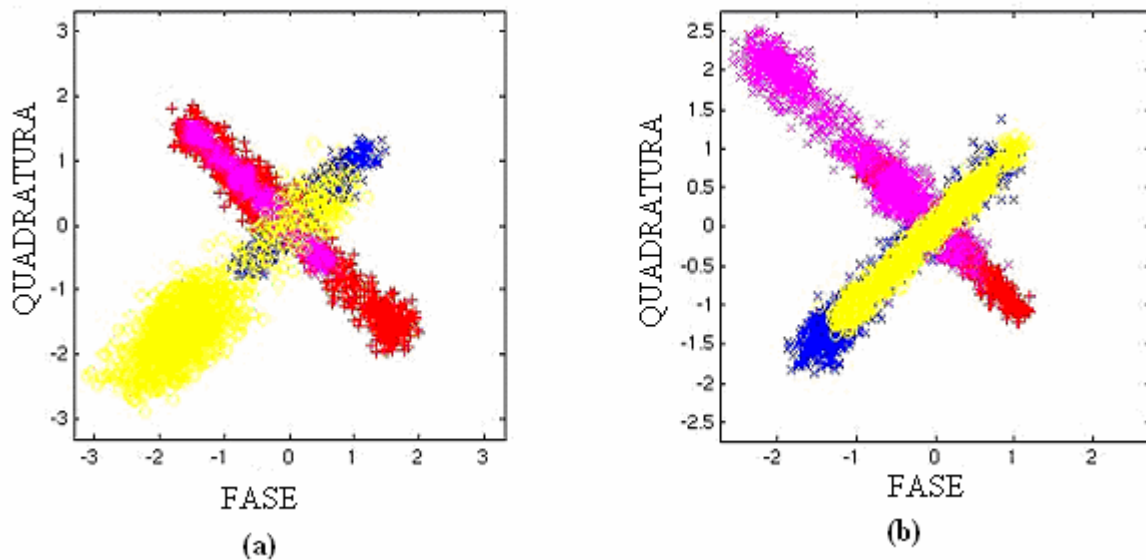


Figura 5.2 – Diagrama de espalhamento de símbolos canal modelo 1 para espalhamento Doppler de 0,5 Hz ; a) Para SNR de 12 dB; b) Para SNR de 15 dB.

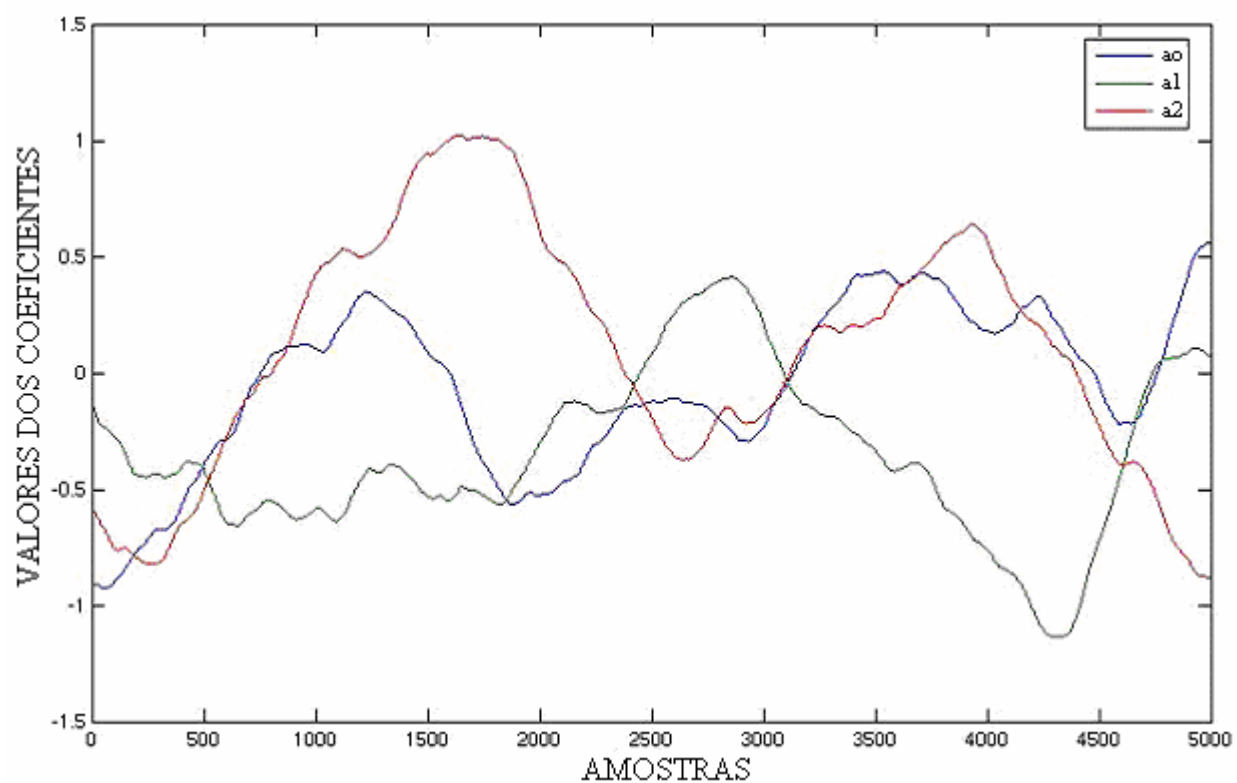


Figura 5.3 – Canal modelo 1 , variação dos coeficientes para espalhamento Doppler 1 Hz.

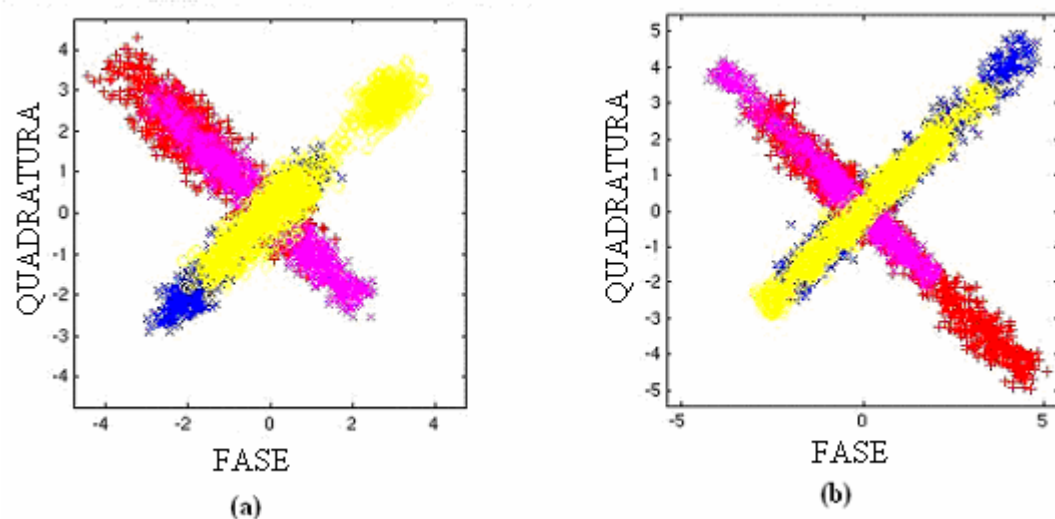


Figura 5.4 – Diagrama de espalhamento de símbolos canal modelo 1 para espalhamento Doppler de 1 Hz ; a) Para SNR de 12 dB; b) Para SNR de 15 dB.

5.1.2 Canal modelo 2

Este modelo de canal apresenta as mesmas características do canal modelo 1 e portanto valem as mesmas considerações, exceto pelo fato do primeiro coeficiente ser constante e unitário. Este canal foi proposto em (Choi, Bouchard e Yeap, 2005) para uma frequência Doppler de 0,5 Hz e é descrito pela seguinte função de transferência:

$$H_2(z) = 1 + a_1(t)z^{-1} + a_2(t)z^{-2} \quad (5.2)$$

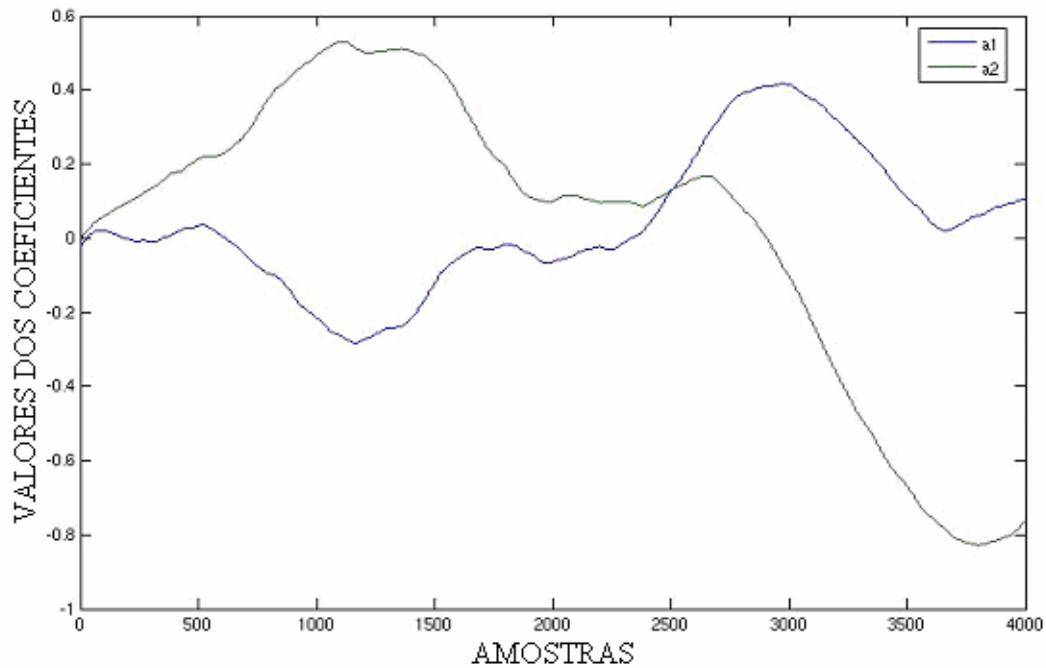


Figura 5.5 – Canal modelo 2 , variação dos coeficientes para espalhamento Doppler 0,5 Hz.

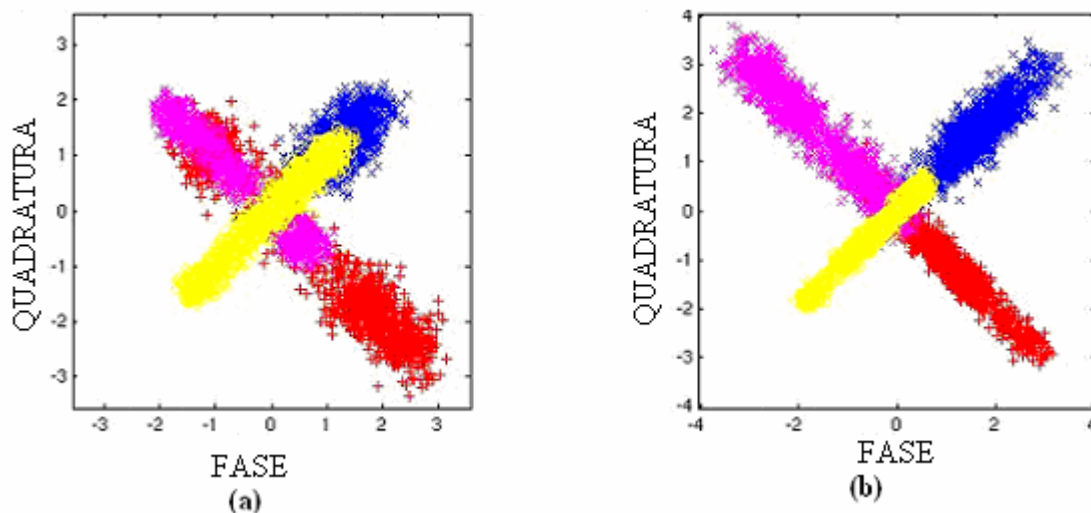


Figura 5.6 – Diagrama de espalhamento de símbolos canal modelo 2 para espalhamento Doppler de 0,5 Hz ; a) Para SNR de 12 dB; b) Para SNR de 15 dB.

5.1.3 Canal modelo 3

Este modelo de canal discreto consiste de três coeficientes com defasagens iguais ao período de um símbolo e é descrito pela seguinte função de transferência:

$$H_3(z) = [c_0 + a_0(t)] + [c_1 + a_1(t)] z^{-1} + [c_2 + a_2(t)] z^{-2} \quad (5.3)$$

Onde $c_0 = 0.3482$; $c_1 = 0.8704$ e $c_2 = 0.3482$.

Este modelo um canal é derivado de um modelo fixo amplamente usado para testes de equalizadores recomendado pela ITU e freqüentemente citado na literatura (Proakis, 2001), para avaliar características de variação no tempo onde são adicionados variáveis aleatórias aos valores dos seus coeficientes. Os valores de variação são gerados através de uma cadeia de Markov de segunda ordem, na qual um sinal AWGN passa através de um filtro de Butterworth de segunda ordem com frequência de corte normalizada 0,1. Este canal foi proposto em (Choi, Lima e Haykin, 2005). Para este tipo de canal será simulada a BER para vários valores de desvio padrão do sinal AWGN para que os resultados possam ser comparados com os do respectivo artigo supracitado.

As figuras 5.7 e 5.9 apresentam as características de variação dos coeficientes para desvios padrões do ruído AWGN de 0,1 e 0,3 respectivamente, enquanto as figuras 5.8 e 5.10 apresentam os respectivos diagramas de espalhamentos para SNR de 6 dB e 12 dB.

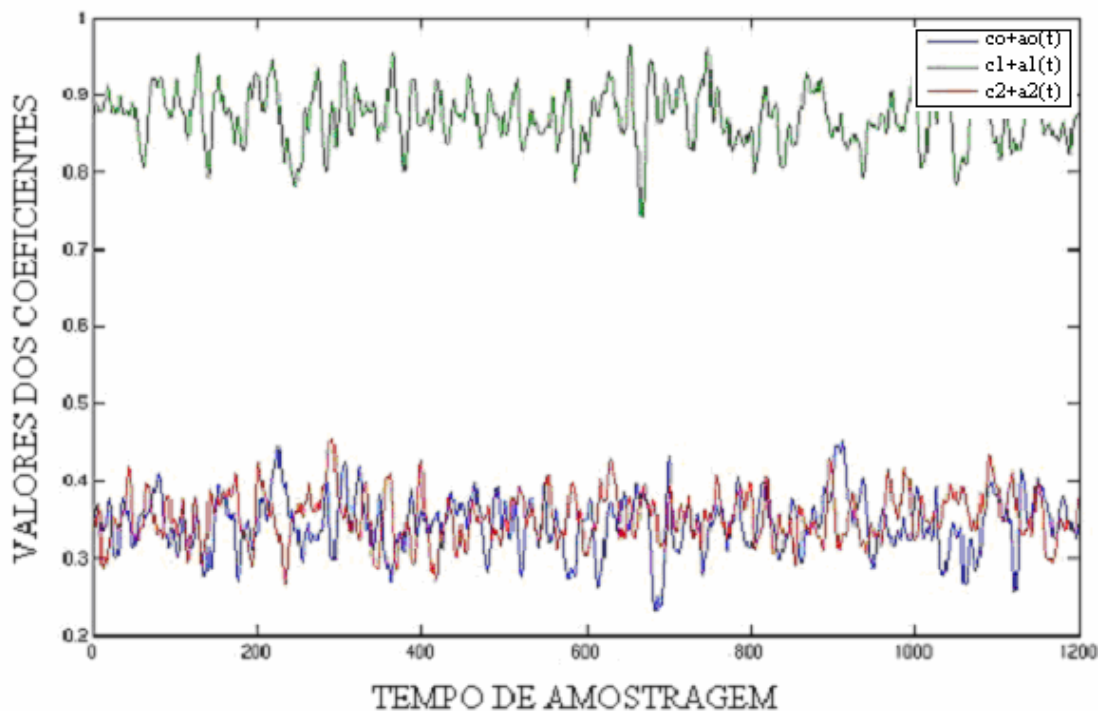


Figura 5.7 – Canal modelo 3 , variação dos coeficientes para desvio padrão 0,1.

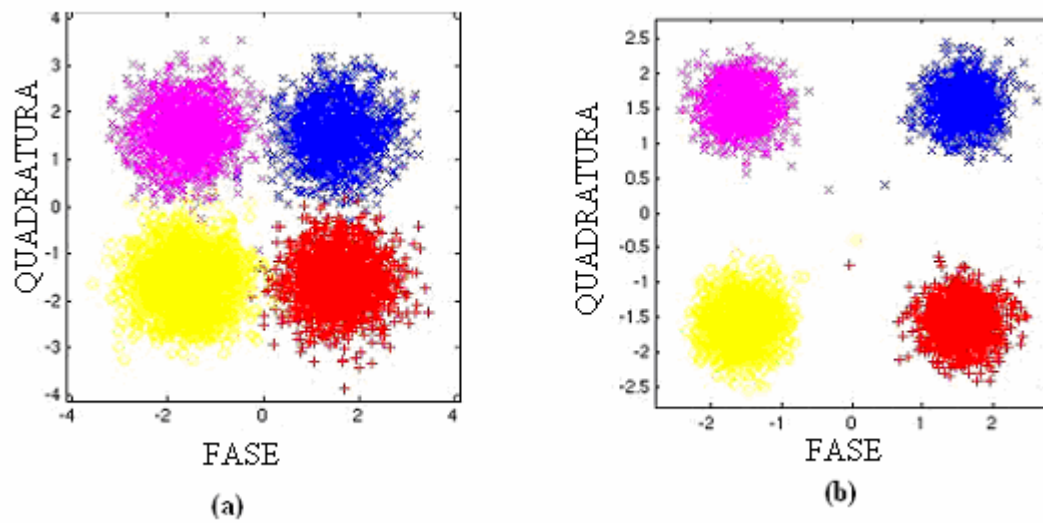


Figura 5.8 – Diagrama de espalhamento canal modelo 3 para desvio padrão 0,1. ; a) Para SNR de 6 dB; b) Para SNR de 12 dB.

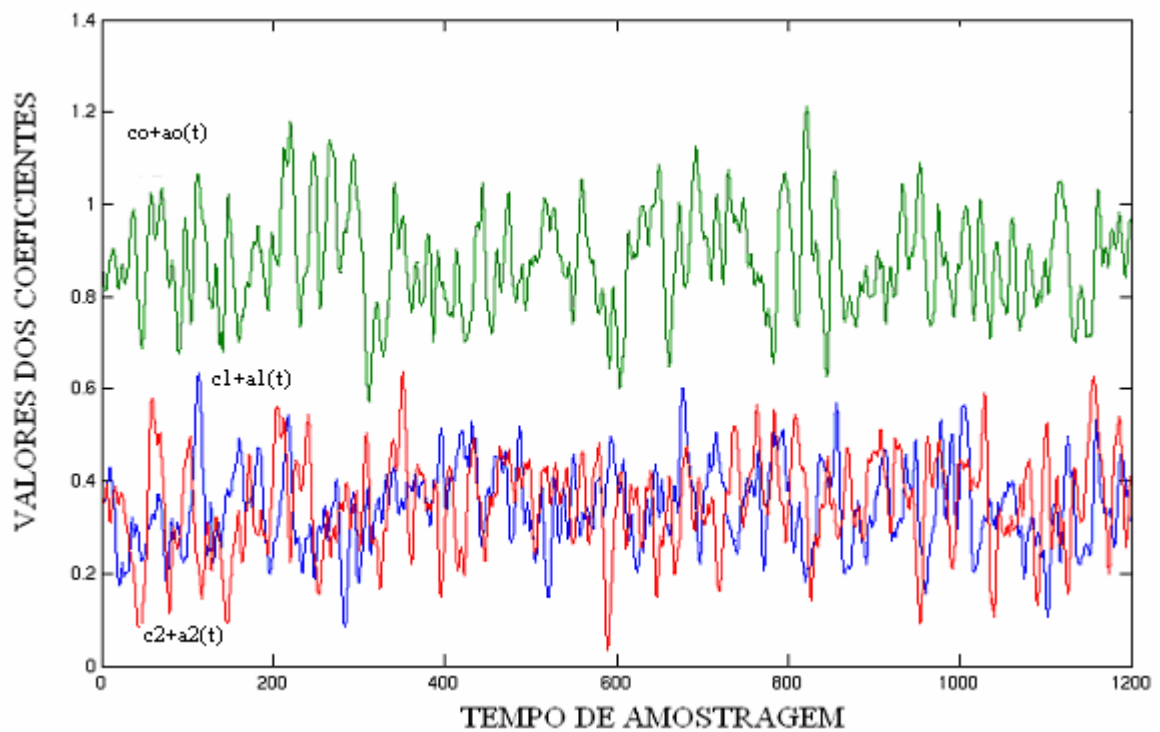


Figura 5.9 – Canal modelo 3 , variação dos coeficientes para desvio padrão 0,3.

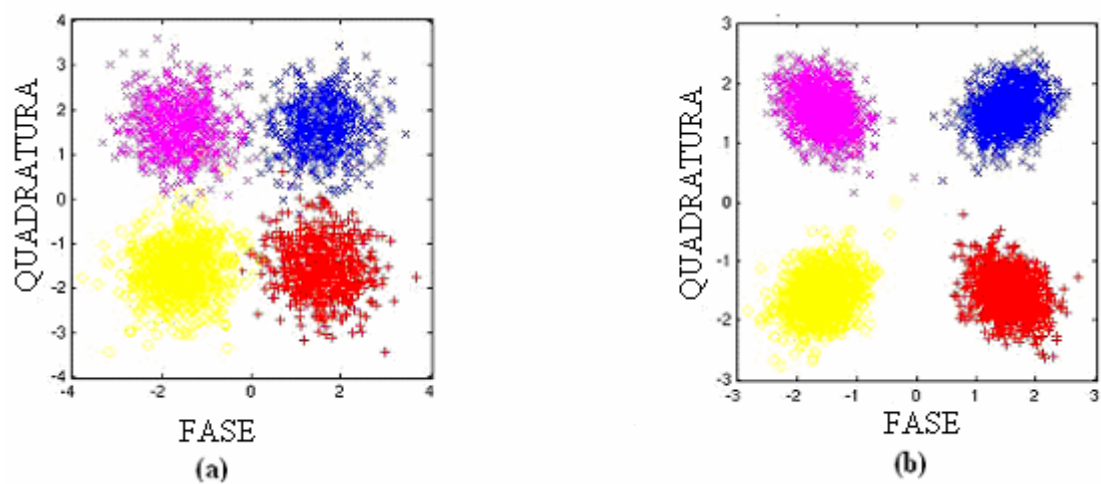


Figura 5.10 – Diagrama de espalhamento de símbolos do canal modelo 3 , variação dos coeficientes para desvio padrão 0,3. ; a) Para SNR de 6 dB; b) Para SNR de 12 dB.

Capítulo 6

Simulações e Resultados

Para avaliar o desempenho e aplicabilidade da estrutura proposta foram utilizadas para comparações outras topologias de neuro-equalizadores utilizadas nos trabalhos de (Choi, Bouchard e Yeap 2005), (Choi, Lima e Haykin, 2005), (Cortes *et al*, 2006) e (Ludwig e Lima, 2005). As topologias utilizadas para comparação foram: DFRNE (*Decision FeedBack Recursive Neural Equalizer*), RTRL (*Real Time Recurrent Learning Algorithm*), RNE-EKF (*Recursive Neural Equalizer Extended Kalman Filter*), RNE-UKF (*Recursive Neural Equalizer Unscented Kalman Filter*), GEKF (*Global Extended Kalman Filter*).

Como foi visto no capítulo 4, no algoritmo de treinamento Gauss-Newton a construção de cada linha da matriz Jacobiana se dá através do registro de variação de erro de cada coeficiente do filtro em relação a cada bit de treinamento, sendo que o número de linhas da matriz corresponde ao número de símbolos de treinamento. Desta forma, a atualização dos coeficientes não se dá de bit em bit, como em outros métodos como o LMS, mas por lotes de símbolos de treinamento. Esta característica do método faz com que ao aumentar a velocidade de variação do canal, seja necessário reduzir o número de símbolos de treinamento para que as informações da matriz Jacobiana responsável pela atualização dos coeficientes do filtro continuem válidas para as características do canal. Conforme foi visto no capítulo 3 e de maneira intuitiva, podemos perceber que há um compromisso entre o número de símbolos de treinamento e o erro dos coeficientes com a velocidade de variação do canal. Em outras palavras, para determinada característica de variação de canal, a partir de um determinado número de símbolos de treinamento, o aumento do mesmo acarreta em um maior erro dos coeficientes, fato devido à diminuição de correlação dos coeficientes do canal durante o período dos símbolos de treinamento. Claramente, percebemos aqui que a minimização dos erros dos coeficientes não é uma solução ótima, pois conforme (Cortes *et al*, 2006) e (Ludwig e Lima, 2005) considerando o comportamento da topologia proposta em canais fixos, seriam necessários 60 símbolos de treinamento para que os erros dos coeficientes atingissem sua variação mínima em uma solução sub-ótima.

Uma primeira abordagem para encontrar o número de símbolos de treinamento que atingiria o menor erro de coeficientes para cada modelo de canal seria achar a autocorrelação da resposta ao

impulso de cada canal. Entretanto, não foi encontrado na literatura o valor específico de autocorrelação que manteria um erro adequado para os coeficientes. Por este motivo, foi adotado um procedimento empírico para determinar qual o número de símbolos de treinamento seria mais adequado. O procedimento consiste em testar o desempenho da BER para diversos valores de SNR e diversos comprimentos de sequência de treinamento, adotando o resultado mais adequado para simulação do canal. Este procedimento foi realizado para todos os canais como uma fase prévia de simulação, visto que as simulações, de fato, dependiam da definição do comprimento da sequência de treinamento.

De forma semelhante, o comprimento da sequência de teste também poderia ser determinado pela característica de cada modelo de canal. Porém, há ainda um compromisso entre a razão símbolos de testes e símbolos de treinamento. Também empiricamente, foi proposta uma relação adequada para cada canal. Ao aumentar o valor desta relação, foi percebida uma degradação no desempenho do neuro-equalizador. Além de todas as considerações expostas acima, o comprimento da sequência, além de manter um desempenho de BER satisfatório, precisaria manter ou superar as relações entre comprimento das sequências de treinamento e testes das topologias propostas nos artigos “*benchmarks*” utilizados. As curvas apresentadas mostram a BER simulada para valores de SNR e Desvio Padrão, deve ser notado que quando não há erro na simulação, o gráfico aparece interrompido.

Foram utilizadas duas abordagens para a fase de símbolos de testes:

Na primeira abordagem, os coeficientes do neuro-equalizador são congelados após o período de treinamento e então os símbolos de testes são enviados.

Na segunda abordagem, os coeficientes do neuro-equalizador são atualizados pelo algoritmo de Gauss-Newton em uma tentativa de rastrear as variações do canal. Para tanto, a cada lote de símbolos de teste de quantidade idêntica ao número de símbolos de treinamento, uma nova matriz Jacobiana é computada e utilizada para atualização dos coeficientes. A atualização sempre é seguida de novos símbolos de teste.

A melhoria de desempenho obtida pela utilização da segunda abordagem, quando obtida, é discutida separadamente para cada modelo de canal proposto.

6.1.1 Canal modelo 1

Para este modelo de canal, o melhor resultado empírico obtido para o número de símbolos de treinamento foi de 20. Este resultado foi o mesmo para as duas frequências Doppler de 0,5 Hz e 1 Hz. Para que os resultados fossem comparáveis com os artigos, a relação entre símbolos de teste e de treinamento foi fixada em 10. Portanto, para cada valor de SNR, foram enviadas 500 seqüências, cada uma com 20 símbolos de treinamento mais 200 símbolos de teste, o que resulta em 110.000 símbolos testados para cada valor de SNR.

Foram realizadas simulações para as duas abordagens de tratamento para os símbolos de teste, com os coeficientes do neuro-equalizador congelados e com tentativa de rastreamento do canal. A figura 6.1. mostra a curva obtida para as duas simulações, evidenciando que a atualização dos coeficientes na fase de teste atinge um melhor resultado do que com os coeficientes congelados.

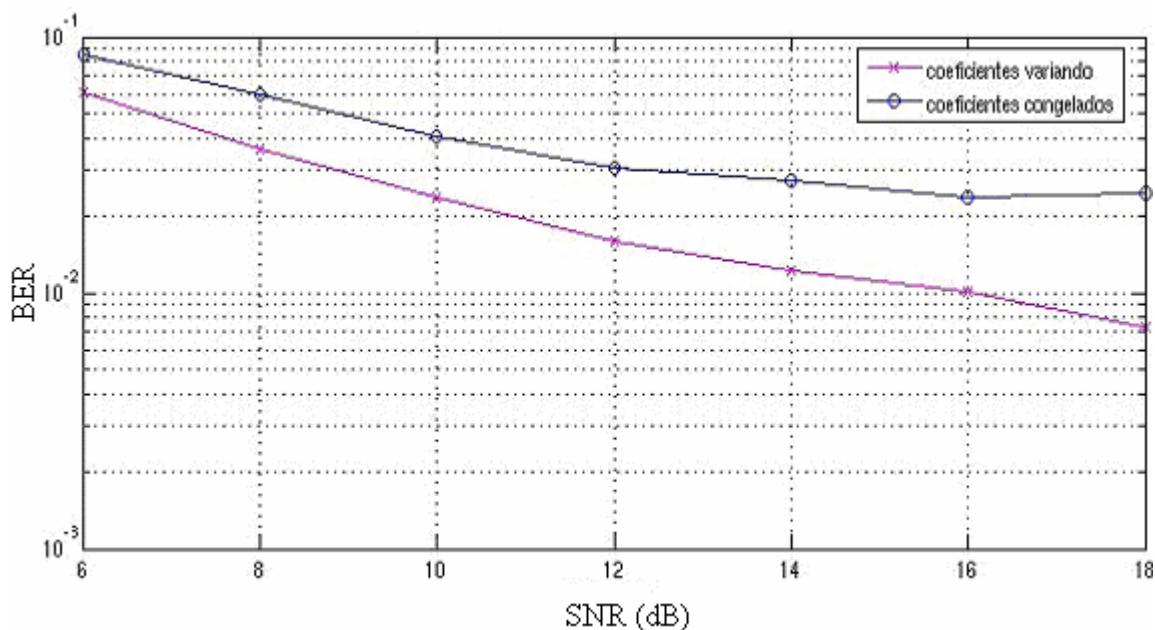


Figura 6.1 – Canal modelo 1 , comparação da BER e ação do Gauss-Newton nos coeficientes no período de teste a medida que se aumenta a SNR para espalhamento Doppler 0,5 Hz.

Na análise das figuras 6.1 e 6.2, verifica-se que neste caso, o ganho obtido com a atualização dos coeficientes é bastante considerável. Este ganho naturalmente ocorre em troca do custo computacional envolvido. Considerando que a cada 20 símbolos de teste temos um novo ajuste de coeficientes, em 200 símbolos de treinamento fazemos nove ajustes, para estes ajustes o maior custo computacional é o cálculo de $J(k)^T J(k)$, já que o cálculo de sua inversa exige apenas 104 operações (Côrtes *et al*, 2006). Para o caso em que $J(k)$ tem 20 linhas temos aproximadamente 995 operações.

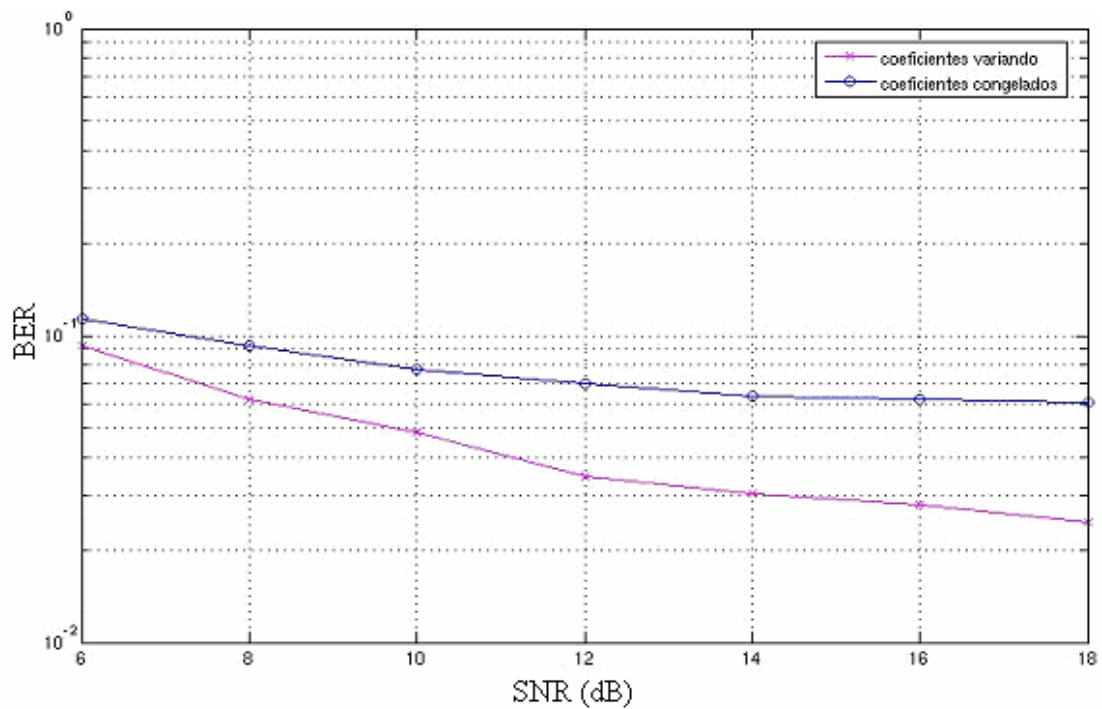


Figura 6.2 – Canal modelo 1 , comparação da BER e ação do Gauss-Newton nos coeficientes no período de teste a medida que se aumenta a SNR para espalhamento Doppler 1 Hz.

Para a comparação com outras topologias foi feita a opção de utilizar rastreamento dos coeficientes na fase de testes devido ao melhor resultado obtido. As figuras 6.3 e 6.4 mostram os valores quando comparados com as topologias propostas por (Choi, Lima e Haykin , 2005) para espalhamento Doppler de 0,5 Hz e 1 Hz respectivamente.

Pela análise dos resultados obtidos podemos verificar que:

Para o canal modelo 1 com espalhamento Doppler de 0,5 Hz, DFE-GN apresenta melhor desempenho que as topologias comparadas- RNE-EKF, RNE-UKF e DFRNE - para valores de SNR até aproximadamente 12 dB sendo superado pelos RNE-EKF, RNE-UKF para valores de SNR superiores.

Para o canal modelo 1 com espalhamento Doppler de 1 Hz, o DFE-GN apresenta melhor desempenho que as topologias comparadas- RNE-EKF, RNE-UKF e DFRNE - para valores de SNR até aproximadamente 11 dB sendo superado pelos RNE-EKF, RNE-UKF para valores de SNR superiores.

O resultado obtido indica que para canais modelo 1 o neuro-equalizador com algoritmo de treinamento Gauss-Newton só deve ser escolhido se a expectativa de valores de SNR do canal for

muito baixa, pelos resultados abaixo de 11 dB.

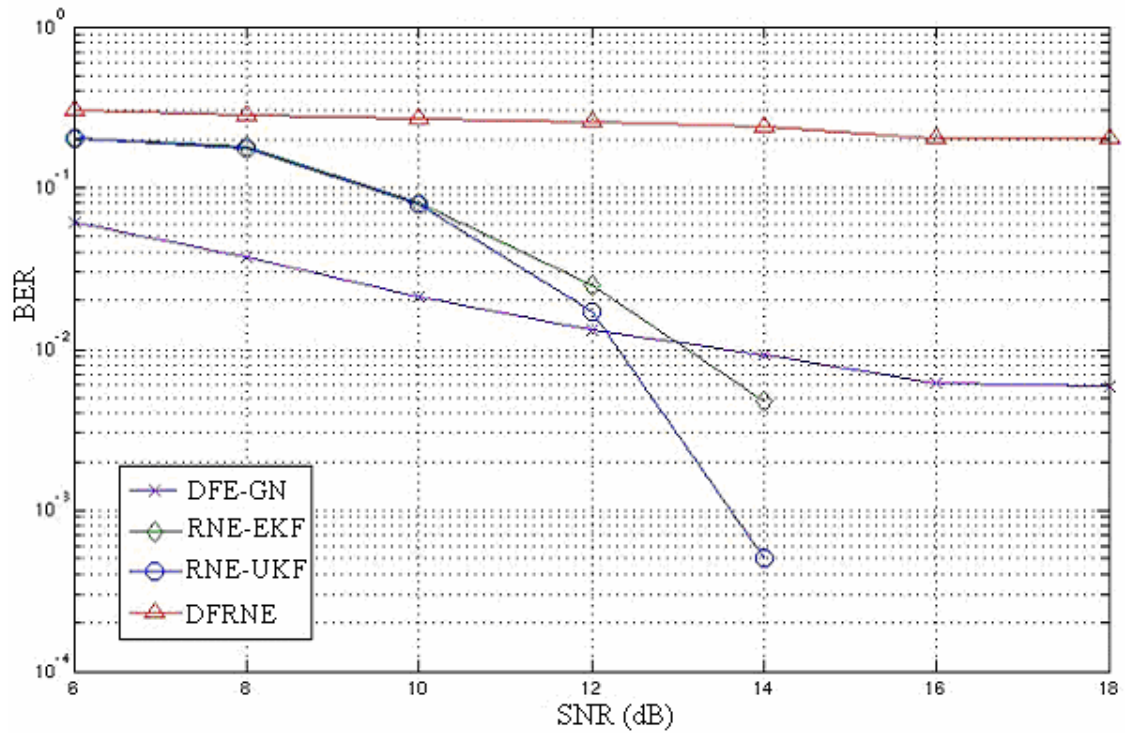


Figura 6.3 – Canal modelo 1 , comparação de desempenho BER do DFE-GN com outras topologias a medida que se aumenta a SNR para espalhamento Doppler 0,5 Hz.

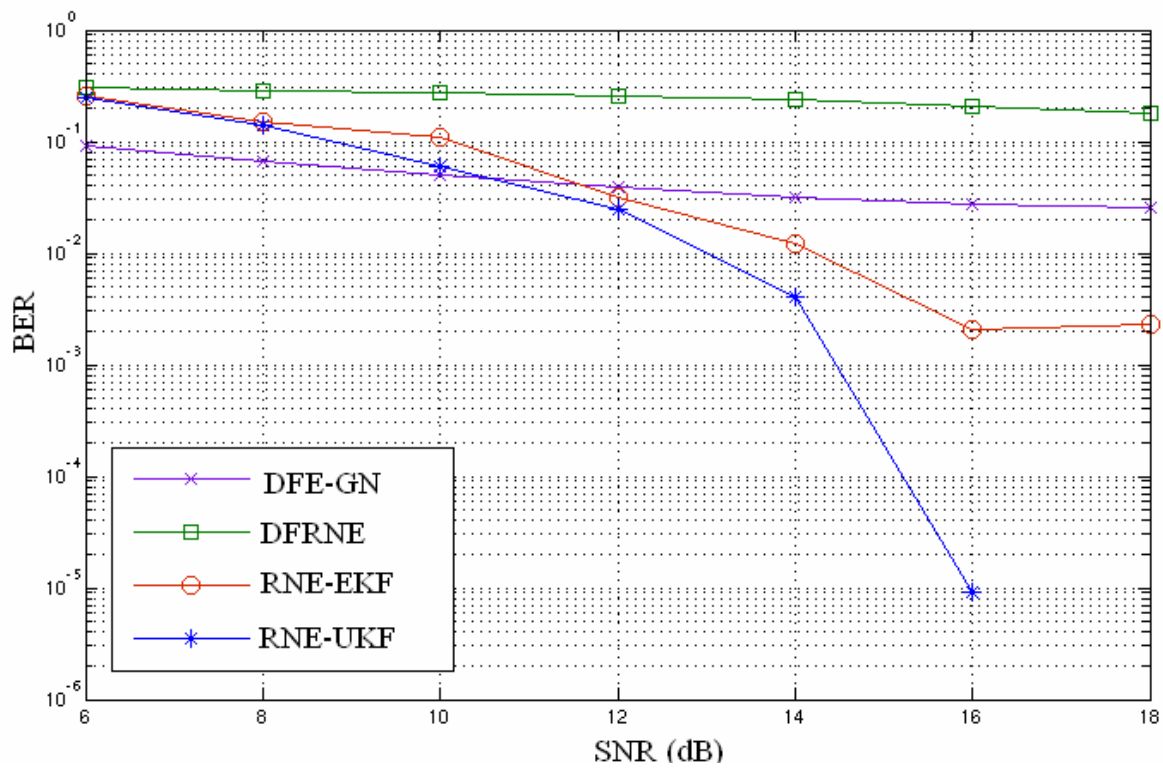


Figura 6.4 – Canal modelo 1 , comparação de desempenho BER do DFE-GN com outras topologias a medida que se aumenta a SNR para espalhamento Doppler 1 Hz.

6.1.2 Canal modelo 2

Para este modelo de canal, o comprimento da sequência de símbolos de treinamento também foi definido empiricamente. O melhor resultado empírico obtido para o número de símbolos de treinamento foi equivalente para valores entre 20 símbolos e 30 símbolos para a frequência Doppler de 0,5 Hz, porém este valor ficou em 20 símbolos para a frequência Doppler de 1 Hz. De acordo com os resultados obtidos, foi escolhido o valor de 20 símbolos de treinamento para as duas frequências. Novamente, para que os resultados fossem comparáveis com o artigo (Choi, Bouchard e Yeap 2005), a relação entre símbolos de teste e de treinamento foi fixada em 10. Portanto para cada SNR, foram enviadas sequências de 20 símbolos de treinamento mais 200 símbolos de testes, repetidas por 500 vezes, desta forma, para cada valor de SNR foram testados 110.000 símbolos.

Novamente foram realizadas simulações com os coeficientes congelados após o período de treinamento e com os coeficientes tentando rastrear as condições de canal. A figura 6.5 apresenta os resultados das simulações para os dois comportamentos dos coeficientes do neuro-equalizador no período de testes para espalhamento Doppler de 0,5 Hz. Vemos que ocorre uma inversão de desempenho em SNRs próxima a 10dB . Para valores de SNR inferiores a SNRs próximas a 10 dB o desempenho do neuro-equalizador com os coeficientes variando com o algoritmo de Gauss-Newton é superior do que o desempenho com os coeficientes sendo congelados no período de teste, enquanto que para valores de SNR superiores a SNR próximas a 10 dB o desempenho do neuro-equalizador com os coeficientes variando com o algoritmo de Gauss-Newton é inferior que o desempenho com os coeficientes sendo congelados no período de teste.

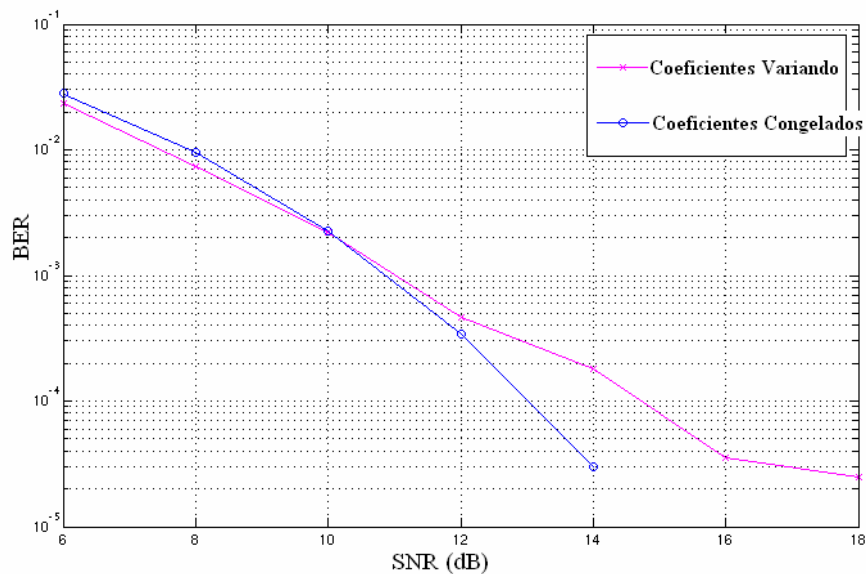


Figura 6.5 – Canal modelo 2 comparação da BER e ação do Gauss-Newton nos coeficientes no período de teste a medida que se aumenta a SNR para espalhamento Doppler 0,5 Hz.

A figura 6.6 apresenta os resultados das simulações para os dois comportamentos dos coeficientes do neuro-equalizador no período de testes para espalhamento Doppler de 1 Hz . Vemos que para este valor de espalhamento Doppler, o desempenho do equalizador com os coeficientes variando com o algoritmo de Gauss-Newton é superior do que o desempenho do equalizador com os coeficientes sendo congelados no período de teste para todos os valores de SNR testadas.

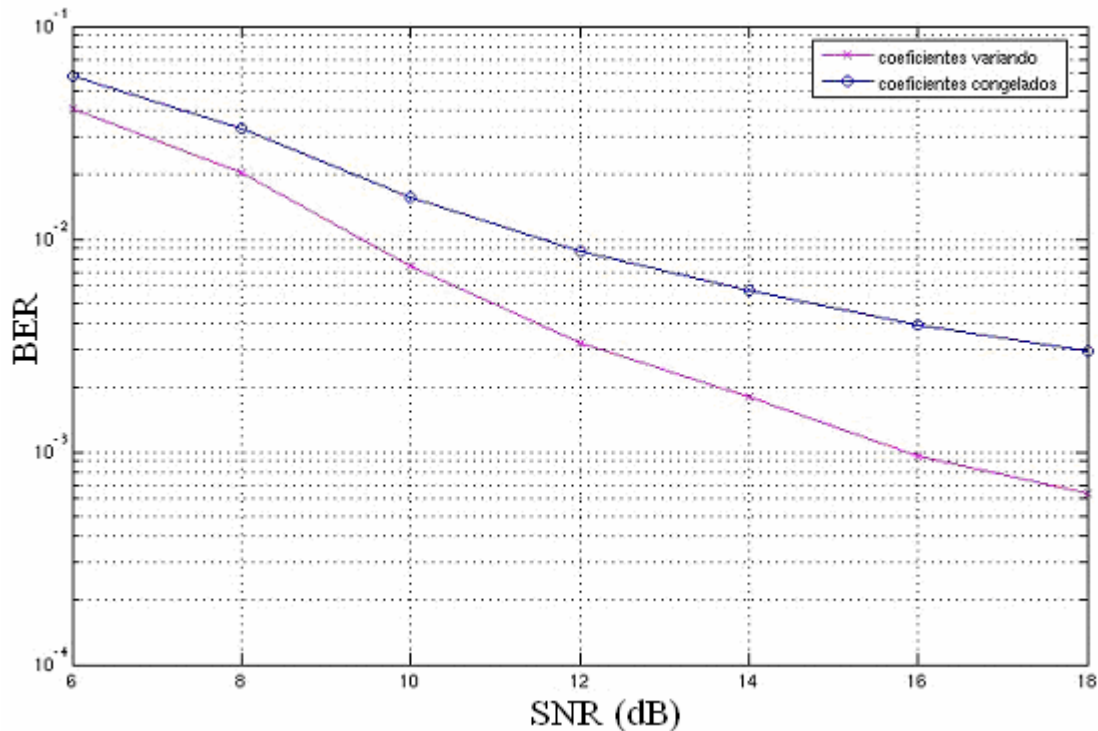


Figura 6.6 – Canal modelo 2 comparação da BER e ação do Gauss-Newton nos coeficientes no período de teste a medida que se aumenta a SNR para espalhamento Doppler 1 Hz.

Comparando os resultados para os dois espalhamentos Doppler, e ainda lembrando do canal modelo 1 que possui condições de variações mais severas que este modelo de canal, intuitivamente chegamos a conclusão que a medida em que as condições de variação do canal se tornam mais severas, ou seja, neste caso, a medida que o espalhamento Doppler aumenta, o desempenho do equalizador melhora utilizando o rastreamento dos coeficientes. Obviamente este resultado só é válido para determinado limite. Por outro lado, ao passo que o espalhamento Doppler diminui, ou as condições de variações são menos severas, o desempenho do equalizador com coeficientes congelados é superior ao desempenho quando há tentativa de rastrear variações do canal. A justificativa deste comportamento pode ser alcançada se considerarmos o resultado obtido por (Cortes *et al*, 2006) quando foi demonstrado que para canais fixos, o desempenho do equalizador diminui quando o número de iterações do método de Gauss-Newton ultrapassa o valor de 4. O

cálculo do Jacobiano a cada número de símbolos de testes correspondentes ao comprimento da sequência de treinamento pode ser interpretado como uma iteração a mais em canais onde as condições de variação são brandas. Assim, quando temos um canal cuja variação é lenta, seria como se estivéssemos aumentando o número de iterações do algoritmo para obtenção dos coeficientes. Neste caso, o erro obtido é maior do que se considerássemos somente as iterações do algoritmo para o período de treinamento. Por outro lado, em canais cuja variação temporal é maior, as novas iterações no período de teste tendem a de certa forma, corrigir os valores dos coeficientes, diminuindo o erro nos seus valores.

Utilizaremos para comparação do canal modelo 2 as topologias GEKF e RTRL apresentados no do artigo de (Choi, Bouchard e Yeap, 2005). Neste artigo foi utilizado o valor para espalhamento Doppler de 0,5 Hz. Para este valor de espalhamento Doppler, foi feita a opção de utilizar os coeficientes congelados na fase de teste, uma vez que a melhoria obtida para SNRs inferiores a 10 dB utilizando o Gauss-Newton na fase de teste não é muito significativa, diferentemente do que ocorre para SNR superiores a 10 dB quando utilizamos os coeficientes congelados.

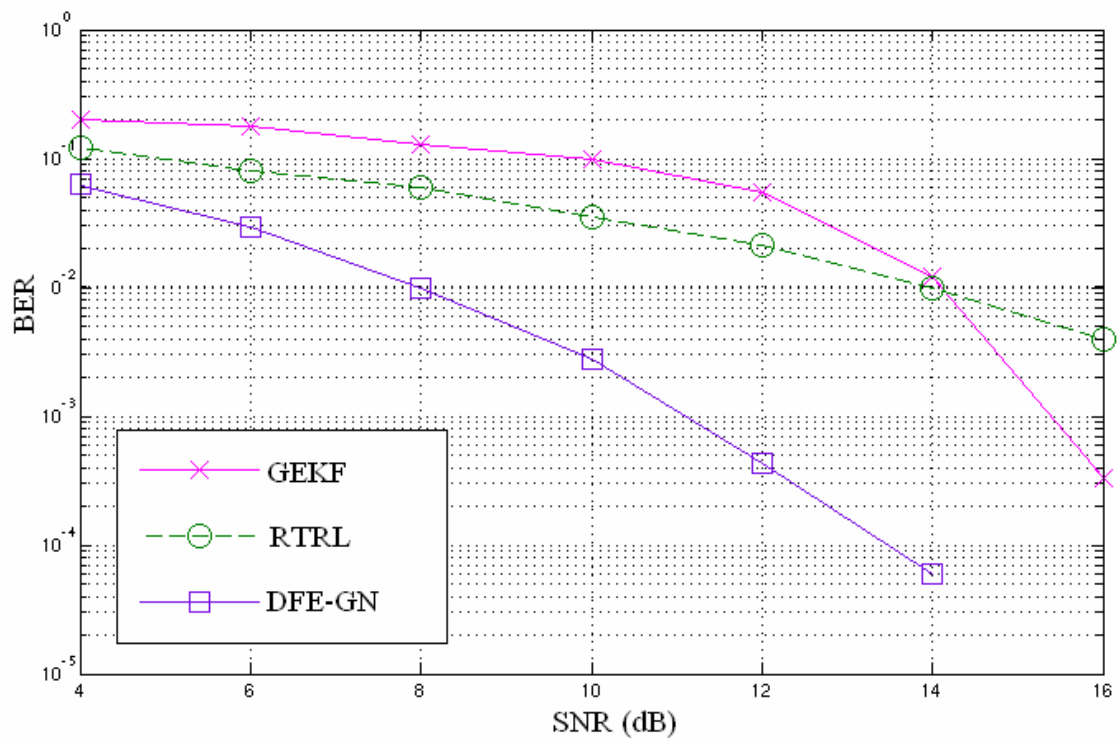


Figura 6.7 – Canal modelo 2 , comparação de desempenho BER do DFE-GN com outras topologias a medida que se aumenta a SNR para espalhamento Doppler 0,5 Hz.

Na análise da figura 6.7, percebemos que o desempenho do DFE-GN superou de maneira indiscutível as topologias GEKF e RTRL para toda a faixa de SNR em teste. Devido a sua extrema simplicidade de implementação e custo computacional significativamente mais baixos que estas topologias, o neuro-equalizador proposto mostra-se adequado ao canal modelo 2.

6.1.3 Canal modelo 3

Para este canal, inicialmente foi realizada a simulação fixando a relação sinal ruído em 16 dB e variando o desvio padrão da fonte AWGN da cadeia de Markov. Foram testados valores de desvio padrão de 0.05 ; 0.1 ; 0.15 ; 0.2 ; 0.25 e 0.3 utilizando-se 100 símbolos para treinamento, 1000 para testes e 200 iterações de forma semelhante como foi realizado em (Choi, Lima e Haykin , 2005).

A figura 6.8. apresenta os resultados das simulações para os coeficientes do neuro-equalizador congelados e tentando rastrear o canal na fase de teste. Novamente, existe uma melhoria no desempenho quando os coeficientes tentam acompanhar as variações do canal, porém visto que a melhoria não é tão significativa, para esta relação entre símbolos de treinamento e testes a utilização do algoritmo na tentativa de rastreamento do canal parece não ser compensadora quando consideramos o incremento de custo computacional.

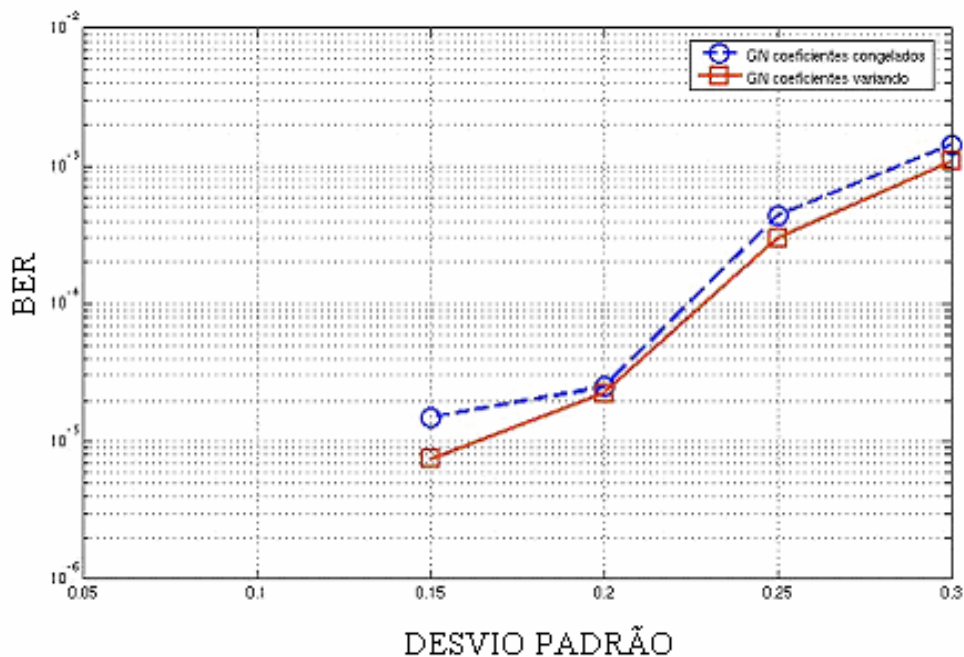


Figura 6.8 – Canal modelo 3 , comparação de desempenho BER e ação do Gauss-Newton nos coeficientes no período de teste variando desvio padrão.

A figura 6.9. mostra o desempenho no canal modelo 3 do DFE-GN quando comparado com os resultados apresentados em (Choi, Lima e Haykin , 2005) para as topologias DFRNE, RNE-EKF e RNE-UKF.

Percebemos que o desempenho do neuro-equalizador com algoritmo de treinamento Gauss-Newton

superou de maneira indiscutível as topologias DFRNE, RNE-EKF e RNE-UKF para toda a faixa de desvio padrão em teste.

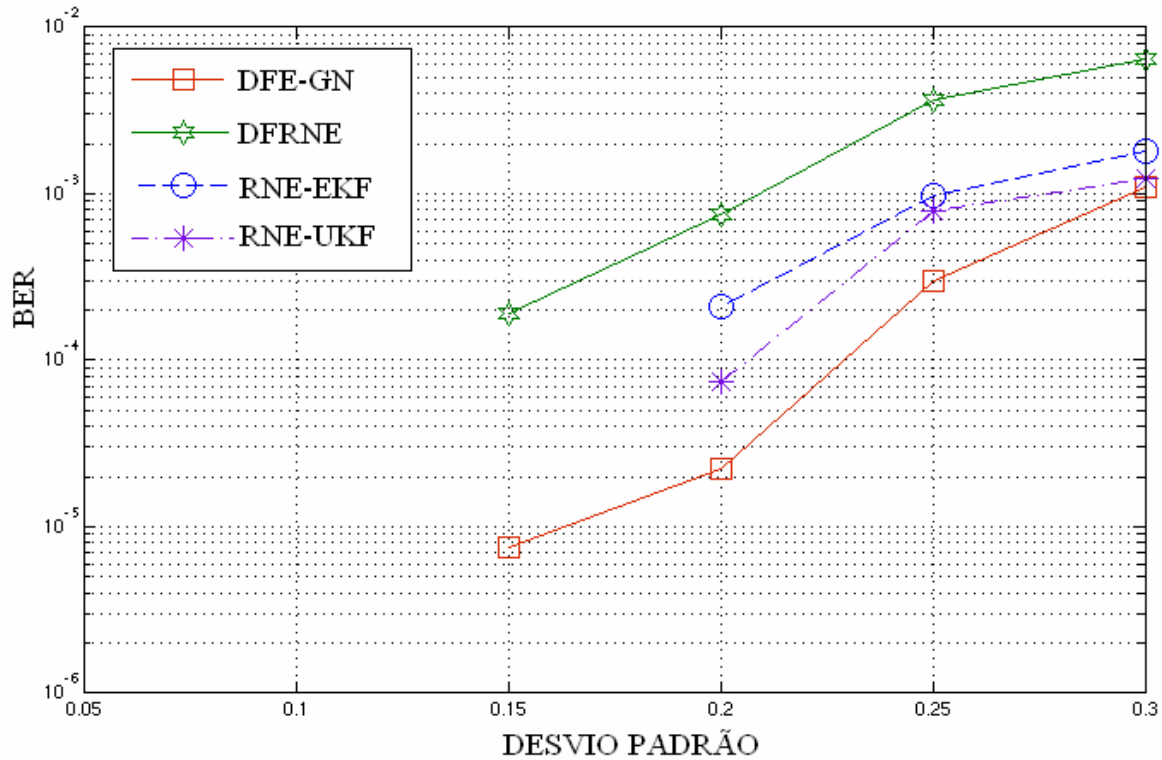


Figura 6.9 – Canal modelo 3 , comparação de desempenho BER do DFE-GN e outras topologias variando desvio padrão.

Em seguida foi realizado o teste de desempenho de BER do canal. A faixa de relação sinal ruído utilizada foi de 4 a 18 dB tomadas a intervalos de 2dB, o desvio padrão foi fixado em 0.1, o número de símbolos de treinamento fixados em 100, os coeficientes do equalizador são congelados após a fase de treinamento e então são enviados 100.000 símbolos de testes. Após 10 repetições é apurado o valor médio da BER. Este procedimento equivale ao realizado em (Choi, Lima e Haykin , 2005). Em seguida comparamos o resultado obtido com os valores de (Choi, Lima e Haykin , 2005) para as topologias DFRNE , RNE-EKF e RNE-UKF o que é mostrado na figura 6.10.

Aqui novamente o desempenho do DFE-GN superou as topologias DFRNE, RNE-EKF e RNE-UKF para toda a faixa de SNR em teste. Novamente considerando a simplicidade da estrutura e o custo computacional envolvido, temos que o neuro-equalizador com algoritmo de treinamento Gauss-Newton é uma solução perfeitamente adequada para o canal modelo 3.

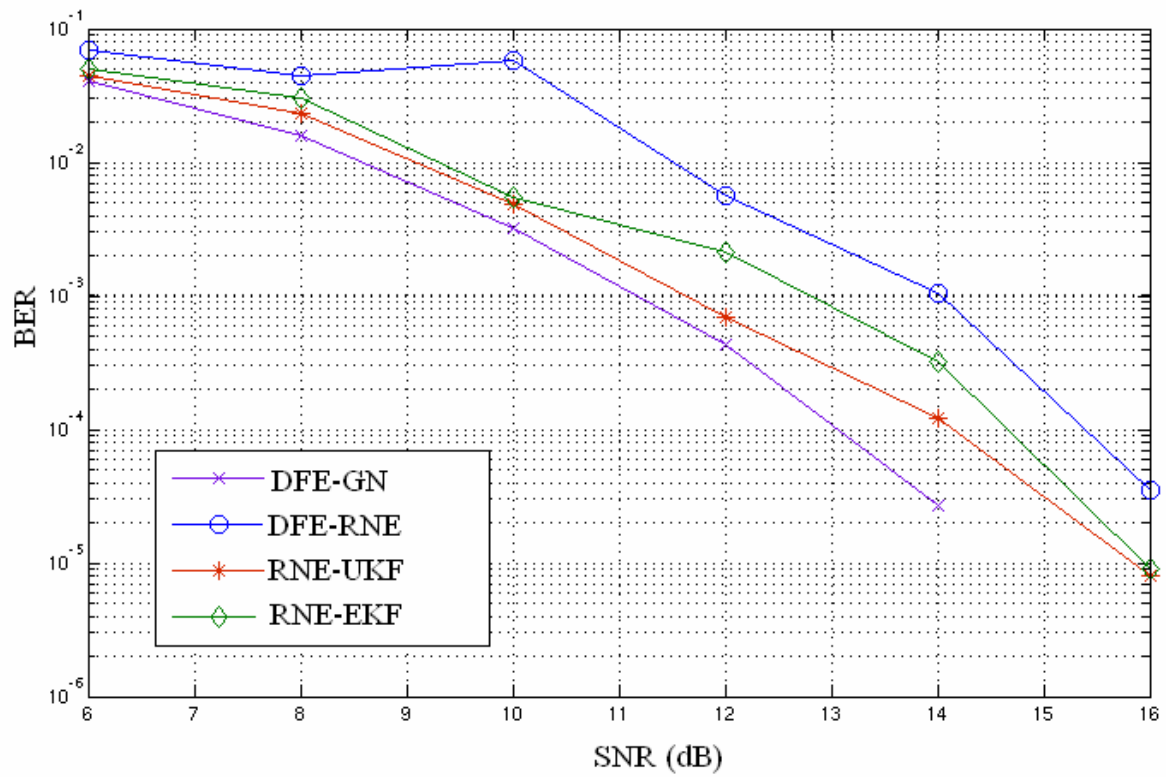


Figura 6.10 – Canal modelo 3 , comparação de desempenho BER do DFE-GN e outra topologias.

Capítulo 7

Conclusão

Este trabalho apresentou o comportamento em canais variantes no tempo de uma topologia de equalizador DFE composto por um *perceptron* recorrente utilizando o algoritmo de Gauss-Newton na fase de treinamento e também como alternativa de rastreamento do canal na fase de teste.

Da análise do desempenho da topologia proposta para os modelos de canais testados, podemos concluir que para canais com variações temporais moderadas “*slow-fading*” o neuro-equalizador estudado se apresenta como uma solução satisfatória, superando em desempenho e custo computacional outras topologias como GEKF e RTRL para o canal modelo 2 e DFRNE, RNE-EKF e RNE-UKF para o canal modelo 3. Por outro lado, para canais com variações mais severas “*fast-fading*” o neuro-equalizador estudado não se apresenta como solução viável, esta situação foi mostrada no canal modelo 1 onde o DFE-GN superou os DFRNE, RNE-EKF e RNE-UKF somente para SNR abaixo da faixa de 10 dB a 12 dB , sendo amplamente superado para SNR superiores a 12 dB.

Em estudo futuro, podem ser estudadas, estruturas de *perceptrons* multicamadas que utilizem o algoritmo de Gauss-Newton e aproveitem suas propriedades de mapeamento de funções e não linearidades na tentativa de obter melhores resultados em canais com condições de variações mais severas.

Referências Bibliográficas

- Austin, M. E. (1967). Decision-feedback equalization for digital communication over dispersive channels, Technical report, Massachusetts Institute of Technology - Research Laboratory of Electronics.
- Chen, Z. and Lima, A. C. de C. (2004). A new neural equalizer for decision-feedback equalization, IEEE Signal Processing Society Workshop, pp. 675– 684.
- Choi, J., Bouchard, M. and Yeap, T. (2005). Decision feedback recurrent neural equalization with fast convergence rate, IEEE Trans. on Neural Net. 16(3): 699–708.
- Choi, J., Lima, A. C. de C. and Haykin, S. (2005). Kalman filter-trained recurrent neural equalizer for time-varying channels, IEEE Trans. on Commun. 53(3): 472–480.
- Cortes, P.N., Junior, O.L. and Lima, A.C. de C. (2006). Neuro-equalizador com algoritmo de treinamento do tipo Gauss-Newton, Anais do XVI CBA, number 532, Salvador, Bahia, Brasil.
- Cortes P.N. (2006). Neuro-Equalizador com Algoritmo de Treinamento do Tipo Gauss Newton, Dissertação de Mestrado – Departamento de Engenharia Elétrica UFBA.
- C. Cowan and S. Semnani, (1998). Time-variant equalization using a novel nonlinear adaptive structure, Int. J. Adapt. Control Signal Process., vol.12, no. 2, pp. 195–206.
- F. Ling and J. G. Proakis, (1995). Adaptive lattice decision-feedback equalizer - Their performance and application to time-variant multipath channels, IEEE Trans. Commun. vol. COM-33, pp. 348–356, Apr. 1985.
- Ibnkahla, M. (2000). Applications for neural networks to digital communications - A survey, Signal Process., 80: 1185–1215.
- Kechriotis, G., Zervas, E. and Manolakos, E. S. (1994). Using recurrent neural networks for adaptive communication channel equalizations, IEEE Trans. Neural Netw. 5: 267–278.

Parisi, R., Claudio, E. D. D., Orlandi, G. and Rao, B. D. (1997). Fast adaptive digital equalization by recurrent neural networks, IEEE Trans. Signal Process. 45: 2731–2739.

Haykin, S. (2001). Communications Systems, 4th edn, Willey.

Haykin, S. (1994). Neural Network: A Comprehensive Foundation, 1st edn, Prentice Hall.

Ludwig O. and Lima A.C. de C. (2005). Transaction on Neural Network, IEEE, vol.1 n° 11.

Luenberguer, David G.(2003). Linear and Nonlinear Programming, 2nd edn, Springer.

M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, Simulation of Communications Systems, 2nd edn., New York: Kluwer Academic/Plenum Publishers, 2000.

Ong, S. Choi, S, You, C. and Hong, D. (1997). A complex version of a decision feedback recurrent neural equalizer as an infinite impulse response filter, in Proc. GLOBECOM'97, Nov. 3–8, pp. 57–61.

Proakis, J. G. (2001). Digital Communications, 4th edn, McGraw-Hill.

Sklar, B. (2001). Digital Communications: Fundamentals and Applications, 2nd edn, Prentice Hall .

Tranter, W. H. ; Shanmugan ,K.S.;Rappaport,T. S. and Kosbar, K.L . (2004). Principles of Communication Systems Simulation with Wireless Applications , 1st edn , Prentice Hall.

Zerguine, A., Shafi, A., and Bettayeb, M. (2001). Multilayer perceptron-based DFE with lattice structure, IEEE Trans. Neural Net., 12(4): 532–545.