

4

Receptores baseados na Mínima Taxa de Erro

Nos últimos anos, receptores multiusuário que empregam o critério de mínimo erro médio quadrático (MMSE) [3, 53, 54] se tornaram bem sucedidos devido ao bom desempenho e simples implementação adaptativa [3, 54, 55]. Contudo, é sabido que a função custo baseada no erro médio quadrático (MSE) não é a mais adequada em aplicações de comunicações digitais, onde a função custo mais apropriada para minimização é a taxa de erro dos símbolos transmitidos e no caso particular deste trabalho, que emprega símbolos binários, a mínima taxa de erro de bits (MBER).

Técnicas que usam a BER como critério de desempenho apareceram pela primeira vez no contexto de equalização linear [56] e foram posteriormente estendidas para o caso de equalização com decisão realimentada [57]. Entretanto, esses métodos requerem o conhecimento do canal e uma grande complexidade computacional. O primeiro algoritmo adaptativo que se aproxima do equalizador de mínima taxa de erro foi proposto em [58], onde as estimativas do canal, da potência do ruído e da saída do canal sem ruído são usadas para aproximar um algoritmo baseado no gradiente estocástico. O algoritmo em [58] é significativamente mais complexo do que o LMS e sofre de problemas de convergência, uma vez que a superfície de BER de um canal de comunicações pode ser irregular com mínimos locais, de modo que a convergência de um algoritmo do tipo gradiente para o mínimo global não é garantida. Recentemente, foram propostos novos algoritmos adaptativos baseados no gradiente estocástico para ajustar os parâmetros de um receptor de comunicações [59]-[64]. A técnica proposta em [59] para um receptor DS-CDMA usa uma aproximação diferencial para o gradiente da probabilidade de erro de uma amostra do sinal e move os coeficientes do receptor na direção negativa do gradiente estocástico estimado. Apesar de apresentar resultados superiores ao LMS, o método em [59] apresenta uma taxa de convergência consideravelmente lenta já que ele ajusta os parâmetros do receptor apenas quando um erro é cometido. O algoritmo proposto em [60] no contexto de equalização foi posteriormente

estendido para receptores multiusuário em sistemas DS-CDMA em [61] e é chamado de *Approximate Minimum Bit Error Rate* (AMBER). O AMBER é um algoritmo bastante eficaz devido à baixa complexidade computacional, é baseado no gradiente estocástico e seu funcionamento é similar ao algoritmo *signed error* LMS, exceto pelo fato de que nas proximidades da região de decisão os parâmetros do receptor continuam a ser atualizados. A técnica descrita em [62] para equalização e denominada *Least Bit Error Rate* (LBER) foi posteriormente utilizada no projeto de receptores para sistemas DS-CDMA [63, 64]. O LBER também é um algoritmo baseado no gradiente estocástico que emprega técnicas de estimação de densidade de *kernel* para aproximar a BER como função dos dados. A vantagem do LBER é que um erro não precisa ser observado a fim de garantir uma estimativa da BER e a função suavizada é uma abordagem conveniente para algoritmos baseados no gradiente estocástico.

Em geral, os algoritmos baseados na minimização da BER requerem seqüências de treinamento mais longas do que o usual para que possam convergir para valores de BER inferiores aos dos métodos que utilizam o MMSE como critério de desempenho. A literatura indica que os algoritmos AMBER e LBER são as técnicas mais bem sucedidas e adequadas para implementação adaptativa de algoritmos que minimizam a BER, desde que a aplicação permita o emprego de uma seqüência de treinamento longa. Uma outra possibilidade interessante para os algoritmos de minimização da BER é o emprego na fase de operação do receptor, posteriormente ao treinamento, onde são usadas estimativas dos símbolos para continuar atualizando os parâmetros do filtro de recepção e rastreando a interferência e o canal. Neste capítulo, as técnicas AMBER e LBER são estendidas para receptores multiusuário com decisão realimentada e novas abordagens para acelerar a convergência do AMBER e do LBER são apresentadas, reduzindo o comprimento necessário para a seqüência de treinamento.

É considerado um receptor com decisão realimentada do tipo *one-shot*, como na Figura 4.1, onde o sinal na saída do filtro casado ao pulso do chip é amostrado à taxa de chip e o ruído é do tipo gaussiano. Os receptores para CDMA com decisão realimentada utilizam as decisões dos outros usuários associados ao sistema para cancelar a IMA, melhorando o desempenho do sistema em comparação com um MUD linear [2, 12, 65, 66]. A saída de um MUD *one-shot* com decisão realimentada é dada por:

$$x_k(i) = \mathbf{w}_k^T(i)\mathbf{r}(i) - \mathbf{f}_k^T(i)\hat{\mathbf{b}}(i) \quad (4-1)$$

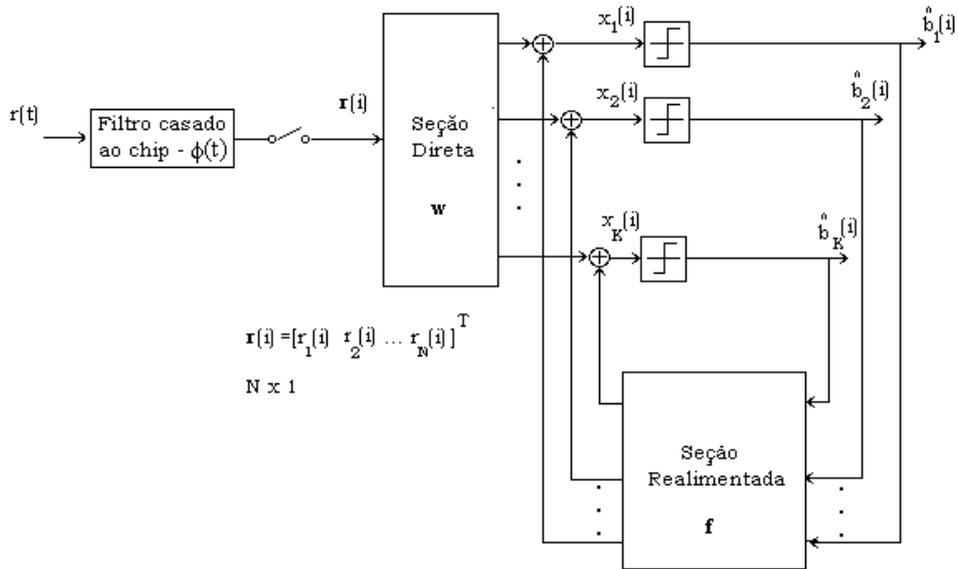


Figura 4.1: Diagrama em blocos do receptor multiusuário com decisão realimentada.

onde $\mathbf{r}(i)$ é o vetor recebido com elementos reais (somente neste capítulo supõe-se sinais reais, por simplicidade) $\hat{\mathbf{b}}(i)$ é o vetor de decisões com dimensão $K \times 1$. A matriz da seção direta do receptor $\mathbf{w}(i)$ tem dimensão $N \times K$, a matriz de realimentação $\mathbf{f}(i) = [\mathbf{f}_1(i) \ \dots \ \mathbf{f}_K(i)]^T$ possui dimensão $K \times K$ e sua estrutura tem como restrição o fato de que a diagonal principal de \mathbf{f} tem zeros para evitar o cancelamento dos símbolos desejados. Neste capítulo, é utilizada uma matriz inteira \mathbf{f} com exceção da diagonal, que corresponde à técnica de decisão realimentada paralela descrita em [66]. Uma das vantagens da decisão realimentada do tipo paralela é que ela produz desempenho em termos da BER aproximadamente uniforme sobre a população de usuários, ao contrário da técnica sucessiva relatada em [67], que fornece um desempenho não uniforme. O símbolo detectado para o receptor multiusuário com decisão realimentada é dado por :

$$\hat{b}_k(i) = \text{sgn}(x_k(i)) \tag{4-2}$$

onde $x_k(i)$ é i -ésimo símbolo estimado para o usuário k e $\text{sgn}(\cdot)$ é a função sinal.

4.1

Minimização da BER em Receptores com Decisão Realimentada

Dada a seqüência de treinamento \mathbf{d}_k transmitida pelo usuário k , a probabilidade de erro de bit $P(\epsilon|\mathbf{d}_k)$, para o receptor multiusuário com decisão realimentada é expressa por:

$$P(\epsilon|\mathbf{d}_k) = P_{\epsilon_k} = P(d_k(i) \operatorname{sgn}(x_k(i)) = -1)$$

$$P_{\epsilon_k} = P(\operatorname{sgn}(d_k(i)x_k(i)) = -1) = P(d_k(i)x_k(i) < 0) \quad (4-3)$$

onde $x_k(i)$ é dado por (4-1) e $d_k(i)$ é o símbolo desejado retirado de uma seqüência de treinamento para o usuário k e o símbolo i .

4.1.1

A Técnica AMBER

A solução para os receptores multiusuário lineares ($\mathbf{f}_k(i) = 0$) que utiliza o critério de desempenho BER através da técnica conhecida como AMBER [60, 61] emprega a função vetorial $g(\mathbf{w}_k(i))$ para aproximar a expressão do gradiente para um vetor de coeficientes $\mathbf{w}_k(i)$ que consegue um desempenho próximo à mínima taxa de erro (MBER) para receptores lineares é descrito por:

$$g(\mathbf{w}_k(i)) = E \left[Q \left(\frac{d_k(i) \mathbf{w}_k^T(i) \mathbf{s}(i)}{\|\mathbf{w}_k(i)\| \sigma} \right) d_k(i) \mathbf{s}(i) \right] \quad (4-4)$$

onde $d_k(i)$ é o símbolo transmitido desejado para o usuário k , tirado da seqüência de treinamento, $Q(\cdot)$ é a função de erro Gaussiana e $\mathbf{s}(i) = \sum_{k=1}^K A_k \mathbf{H}_k(i) \mathbf{S}_k \mathbf{b}_k(i)$ contém as amostras recebidas sem ruído retiradas das saídas dos filtros casados ao pulso de *chip*. Para receptores lineares a quantidade $Q\left(\frac{d_k(i) \mathbf{w}_k^T(i) \mathbf{s}(i)}{\|\mathbf{w}_k(i)\| \sigma}\right)$ dentro do operador valor esperado em (4-4) corresponde à probabilidade de erro condicional dado o produto $d_k(i) \mathbf{s}(i)$. Essa quantidade pode ser substituída em (4-4) por uma função indicadora de erro $i_{d_k(i)}$ dada por:

$$i_{d_k(i)} = \frac{1}{2} (1 - \operatorname{sgn}(d_k(i)x_k(i))) \quad (4-5)$$

onde $x_k(i)$ é o símbolo estimado. O gradiente da função custo se reduz à seguinte expressão:

$$g(\mathbf{w}_k(i)) = E \left[E[i_{d_k(i)} | d_k(i) \mathbf{s}(k)] d_k(i) \mathbf{s}(k) \right] = E[i_{d_k(i)} d_k(i) \mathbf{s}(k)] \quad (4-6)$$

Como $\mathbf{s}(i) = \mathbf{r}(i) - \mathbf{n}(i)$, e $i_{d_k}(i)$ e $d_k(i)$ são estatisticamente independentes, temos $E[i_{d_k}(i)d_k(i)\mathbf{n}(i)] = E[d_k(i)]E[i_{d_k}(i)\mathbf{n}(i)] = 0$ e então o valor instantâneo da função em (4-6), para receptores lineares, é dado por:

$$\hat{g}(\mathbf{w}_k(i)) = i_{d_k}(i)d_k(i)\mathbf{r}(i) \quad (4-7)$$

4.1.2 A Técnica LBER

A taxa de erro de bits (BER) observada na saída do receptor multiusuário é dependente da distribuição da variável de decisão $x_k(i)$, que é função dos parâmetros do receptor. Considerando a probabilidade de erro dada por (4-3), a variável de decisão ajustada pelo sinal para o receptor multiusuário com decisão realimentada $x_{s_k}(i) = d_k(i)x_k(i)$, retirada de uma mistura Gaussiana é dada por:

$$x_{s_k}(i) = \text{sgn}(d_k(i)) \left(\mathbf{w}_k^T(i)\mathbf{s}(i) - \mathbf{f}_k^T(i)\hat{\mathbf{b}}(i) + \mathbf{w}_k^T(i)\mathbf{n}(i) \right) \quad (4-8)$$

onde o primeiro termo de (4-8) $\mathbf{s}(i) = \sum_{k=1}^K A_k \mathbf{H}_k(i) \mathbf{S}_k \mathbf{b}_k(i)$ contém as amostras do sinal transmitido e as quantidades do somatório são dadas pelo modelo síncrono para o enlace reverso descrito na Seção 2.3. O vetor de amostras do ruído é dado por $\mathbf{n}(i) = [n_1(i) \dots n_N(i)]^T$.

$$x_{s_k}(i) = \text{sgn}(d_k(i))x'_k(i) + n'(i) \quad (4-9)$$

onde o primeiro termo de (4-9) é a saída sem ruído e com sinal ajustado do receptor multiusuário. O segundo termo $n'(i)$ é um ruído gaussiano com média zero ($E[n'(i)] = 0$) e variância dada por $\sigma^2 \mathbf{w}_k^T \mathbf{w}_k$.

Considere que J amostras dos símbolos transmitidos $b_k(i)$ e J amostras dos símbolos estimados $x_k(i)$ estão disponíveis através do conhecimento das amostras $d_i(k) = b_i(k)$ de uma seqüência de treinamento. Uma estimativa de densidade de *kernel* [62]-[64] é dada por:

$$\hat{p}_{x_{s_k}}(x_{s_k}) = \frac{1}{J\sqrt{2\pi}\rho\sqrt{\mathbf{w}_k^T \mathbf{w}_k}} \sum_{i=1}^J \exp\left(\frac{-(x_{s_k} - \text{sgn}(d_k(i))x_k(i))^2}{2\rho^2 \mathbf{w}_k^T \mathbf{w}_k}\right) \quad (4-10)$$

onde ρ é o parâmetro de raio da estimativa de densidade de *kernel* [62]-[64].

Substituindo a estimativa por múltiplas amostras dada por (4-10) por

uma estimativa de *kernel* de uma única amostra [62]-[64] tem-se:

$$\hat{p}_{x_{s_k}}(x_{s_k}) = \frac{1}{\sqrt{2\pi\rho}\sqrt{\mathbf{w}_k^T \mathbf{w}_k}} \exp\left(\frac{-(x_{s_k} - \text{sgn}(d_k(i))x_k(i))^2}{2\rho^2 \mathbf{w}_k^T \mathbf{w}_k}\right) \quad (4-11)$$

A probabilidade de erro para o usuário k é estimada por:

$$P_{\epsilon_k} = P(x_{s_k} < 0) = \int_{-\infty}^0 \hat{p}_{x_{s_k}}(x_{s_k}) dx_{s_k} = Q\left(\frac{\text{sgn}(d_k(i))x_k(i)}{\rho(\mathbf{w}_k^T \mathbf{w}_k)^{1/2}}\right) \quad (4-12)$$

Os termos do gradiente de P_{ϵ} em relação a $\mathbf{w}_k(i)$ e $\mathbf{f}_k(i)$ são:

$$\frac{\partial P_{\epsilon_k}}{\partial \mathbf{w}_k} = \frac{\exp\left(\frac{-x_k(i)^2}{2\rho^2 \mathbf{w}_k^T \mathbf{w}_k}\right) \text{sgn}(d_k(i))}{\sqrt{2\pi\rho}} \left(\frac{-\mathbf{r}(i)}{(\mathbf{w}_k^T \mathbf{w}_k)^{1/2}} + \frac{\mathbf{w}_k x_k(i)}{(\mathbf{w}_k^T \mathbf{w}_k)^{3/2}} \right) \quad (4-13)$$

$$\frac{\partial P_{\epsilon_k}}{\partial \mathbf{f}_k} = \frac{1}{\sqrt{2\pi\rho}\sqrt{\mathbf{w}_k^T \mathbf{w}_k}} \exp\left(\frac{-x_k(i)^2}{2\rho^2 \mathbf{w}_k^T \mathbf{w}_k}\right) \text{sgn}(d_k(i)) \hat{\mathbf{b}}(i) \quad (4-14)$$

4.2 Algoritmos Adaptativos

Nessa seção são descritos algoritmos baseados no gradiente estocástico (SG) e numa aproximação do método de Newton [45, 46, 67], denominado neste trabalho gradiente-Newton, que ajustam os parâmetros dos receptores com base na minimização do erro médio quadrático e da BER. Note que no caso de estimação MMSE com algoritmos adaptativos, a técnica gradiente-Newton equivale ao método LMS-Newton descrito em [46, 68]. É importante ressaltar que as duas técnicas mencionadas mostraram-se adequadas à minimização da BER, cuja superfície exibe mínimos locais e que necessita de uma escolha criteriosa do parâmetro de convergência de modo a alcançar taxa de convergência satisfatória e desempenho próximo da mínima BER.

4.2.1 Algoritmos baseados no Gradiente Estocástico

São apresentados algoritmos para ajuste dos parâmetros do receptor baseados no gradiente estocástico. Essas técnicas são simples, de fácil implementação e apresentam uma taxa de convergência relativamente lenta [45, 46]. Além disso, esses algoritmos são suscetíveis à dispersão dos autovalores da matriz autocovariância de observação, o que pode acarretar uma diminuição da taxa de convergência.

A solução de mínimo erro médio quadrático (MMSE) para o receptor multiusuário com decisão realimentada emprega o sinal de erro $e_k(i) = d_k(i) - x_k(i)$ e é descrita por:

$$\mathbf{w}_k(i+1) = \mathbf{w}_k(i) + \mu_w e_k(i) \mathbf{r}(i) \quad (4-15)$$

$$\mathbf{f}_k(i+1) = \mathbf{f}_k(i) - \mu_f e_k(i) \hat{\mathbf{b}}(i) \quad (4-16)$$

onde μ_w e μ_f são os parâmetros de convergência das recursões para a seção direta e a seção realimentada do MUD, respectivamente.

A solução motivada pela técnica AMBER para a mesma estrutura de recepção é obtida através uma otimização com gradiente estocástico e expressa por:

$$\mathbf{w}_k(i+1) = \mathbf{w}_k(i) + \mu_w g(\mathbf{w}_k(i), \mathbf{f}_k(i)) = \mathbf{w}_k(i) + \mu_w i_{d_k}(i) d_k(i) \mathbf{r}(i) \quad (4-17)$$

$$\mathbf{f}_k(i+1) = \mathbf{f}_k(i) - \mu_f g(\mathbf{w}_k(i), \mathbf{f}_k(i)) = \mathbf{f}_k(i) - \mu_f i_{d_k}(i) d_k(i) \hat{\mathbf{b}}(i) \quad (4-18)$$

onde μ_w e μ_f são os passos do algoritmo e $i_{d_k}(i) = \frac{1}{2}(1 - \text{sgn}(d_k(i)x_k(i) - \tau))$ é a função indicadora de erro modificada cujo limiar τ aumenta a taxa de convergência e $x_k(i)$ é dado por (4-1). Note que para a obtenção da expressão em (4-18) para a seção realimentada do receptor usamos uma analogia com a recursão LMS, que inverte o sinal e emprega o vetor de decisões $\hat{\mathbf{b}}(i)$, e cuja expressão pode ser vista em [62].

O algoritmo LBER para o receptor com decisão realimentada em questão é obtido substituindo-se os termos da expressão do gradiente de (4-13) e (4-14) em $\mathbf{w}_k(i+1) = \mathbf{w}_k(i) - \mu_w \left[\frac{\partial P_{\epsilon_i}}{\partial \mathbf{w}_k} \right]_i$ e $\mathbf{f}_k(i+1) = \mathbf{f}_k(i) - \mu_f \left[\frac{\partial P_{\epsilon_i}}{\partial \mathbf{f}_k} \right]_i$ e ajustando os coeficientes do receptor de modo que $\mathbf{w}_k^T(i) \mathbf{w}_k(i) = 1$:

$$\begin{aligned} \mathbf{w}_k(i+1) &= \mathbf{w}_k(i) + \mu_w \frac{1}{\sqrt{2\pi\rho}} \exp\left(\frac{-(x_k(i))^2}{2\rho^2}\right) \text{sgn}(d_k(i)) \\ &\quad \times (\mathbf{r}(i) - \mathbf{w}_k(i)x_k(i)) \end{aligned} \quad (4-19)$$

$$\mathbf{f}_k(i+1) = \mathbf{f}_k(i) - \mu_f \frac{1}{\sqrt{2\pi\rho}} \exp\left(\frac{-(x_k(i))^2}{2\rho^2}\right) \text{sgn}(d_k(i)) \hat{\mathbf{b}}(i) \quad (4-20)$$

onde ρ é relacionado ao desvio padrão do ruído σ .

4.2.2 Algoritmos baseados no Gradiente-Newton

Algoritmos que aproximam o método de Newton [46, 67], como por exemplo os algoritmos gradiente-Newton [46, 68] e RLS [45, 46], incorporam estatísticas de segunda ordem dos sinais de entrada, o que leva ao aumento da taxa de convergência. Esses métodos apresentam em geral uma taxa de convergência mais rápida do que algoritmos baseados no gradiente estocástico, ainda que eles requeiram uma complexidade computacional maior. Além disso, a convergência dos algoritmos gradiente-Newton não depende da dispersão dos autovalores da matriz de autocorrelação do vetor de observações. A equação de atualização do método de Newton é dada por:

$$\mathbf{w}_k(i+1) = \mathbf{w}_k(i) - \frac{1}{2} \mathbf{R}^{-1}(i) z_{\mathbf{w}_k}(i) \quad (4-21)$$

onde $\mathbf{R}(i)$ é a matriz de autocorrelação do vetor de observações $\mathbf{r}(i)$ e $z_{\mathbf{w}_k}(i)$ é o vetor de gradiente.

Na prática, apenas estimativas da matriz $\mathbf{R}(i)$ de autocorrelação do vetor de observações e do vetor de gradiente $g_{\mathbf{w}_i}(k)$ estão disponíveis. Essas estimativas podem ser aplicadas ao método de Newton para obter uma regra de atualização dada por:

$$\mathbf{w}_k(i+1) = \mathbf{w}_k(i) - \delta \hat{\mathbf{R}}^{-1}(i) \hat{z}_{\mathbf{w}_k}(i) \quad (4-22)$$

O fator de convergência δ é introduzido para proteger o algoritmo de divergência, que é originária do uso de estimativas ruidosas de $\mathbf{R}(i)$ e $z_{\mathbf{w}_k}(i)$.

Para obter uma estimativa despolarizada da matriz $\mathbf{R}(i)$, utiliza-se a seguinte soma ponderada:

$$\hat{\mathbf{R}}(i) = \alpha \mathbf{r}(i) \mathbf{r}^T(i) + (1 - \alpha) \hat{\mathbf{R}}(i-1) \quad (4-23)$$

onde α é uma pequena constante cujo valor é escolhido na faixa $0 < \alpha \leq 0,1$.

Para evitar a requerida inversão da matriz $\hat{\mathbf{R}}(i)$, emprega-se o lema de inversão de matrizes [45, 46] descrito por:

$$\hat{\mathbf{R}}^{-1}(i) = \frac{1}{1 - \alpha} \left[\hat{\mathbf{R}}^{-1}(i-1) - \frac{\hat{\mathbf{R}}^{-1}(i-1) \mathbf{r}(i) \mathbf{r}^T(i) \hat{\mathbf{R}}^{-1}(i-1)}{\frac{1-\alpha}{\alpha} + \mathbf{r}^T(i) \hat{\mathbf{R}}^{-1}(i-1) \mathbf{r}(i)} \right] \quad (4-24)$$

A solução via algoritmo LMS-Newton [68] para o receptor multiusuário com decisão realimentada utiliza as técnicas descritas em (4-22)-(4-24) para

obter as seguintes expressões:

$$\mathbf{w}_k(i+1) = \mathbf{w}_k(i) + \delta_w \hat{\mathbf{R}}_r^{-1}(i) e_k(i) \mathbf{r}(i) \quad (4-25)$$

$$\mathbf{f}_k(i+1) = \mathbf{f}_k(i) - \delta_f \hat{\mathbf{R}}_b^{-1}(i) e_k(i) \hat{\mathbf{b}}(i) \quad (4-26)$$

onde $e_k(i) = d_k(i) - x_k(i)$, as matrizes $\hat{\mathbf{R}}_r^{-1}(i)$ e $\hat{\mathbf{R}}_b^{-1}(i)$ são estimativas das matrizes inversas de autocorrelação do vetor de observação $\mathbf{r}(i)$ e do vetor de decisões $\hat{\mathbf{b}}(i)$, respectivamente, e δ_w e δ_f são os passos dos algoritmos para a seção direta $\mathbf{w}_k(i)$ e realimentada $\mathbf{f}_k(i)$, respectivamente.

A solução usando a técnica AMBER para a mesma estrutura de recepção, através de uma otimização com o método gradiente-Newton, será denominada gradiente-Newton-AMBER e é obtida de uma forma análoga à sua versão com gradiente estocástico:

$$\mathbf{w}_k(i+1) = \mathbf{w}_k(i) + \delta_w \hat{\mathbf{R}}_r^{-1}(i) i_{d_k}(i) d_k(i) \mathbf{r}(i) \quad (4-27)$$

$$\mathbf{f}_k(i+1) = \mathbf{f}_k(i) - \delta_f \hat{\mathbf{R}}_b^{-1}(i) i_{d_k}(i) d_k(i) \hat{\mathbf{b}}(i) \quad (4-28)$$

onde $i_{d_k}(i) = \frac{1}{2}(1 - \text{sgn}(d_k(i)x_k(i) - \tau))$ é a função indicadora de erro modificada presente em (4-17) e (4-18).

O algoritmo LBER usando o método gradiente-Newton, chamado gradiente-Newton-LBER, para o receptor com decisão realimentada em questão é obtido substituindo-se os termos da expressão do gradiente de (4-13) e (4-14) em $\mathbf{w}_k(i+1) = \mathbf{w}_k(i) - \delta_w \hat{\mathbf{R}}_r^{-1}(i) \left[\frac{\partial P_{\epsilon_i}}{\partial \mathbf{w}_k} \right]_i$ e $\mathbf{f}_k(i+1) = \mathbf{f}_k(i) - \delta_f \hat{\mathbf{R}}_b^{-1}(i) \left[\frac{\partial P_{\epsilon_i}}{\partial \mathbf{f}_k} \right]_i$ e ajustando os coeficientes do receptor de modo que $\mathbf{w}_k^T(i) \mathbf{w}_k(i) = 1$ de maneira similar à sua versão com gradiente estocástico:

$$\begin{aligned} \mathbf{w}_k(i+1) &= \mathbf{w}_k(i) + \delta_w \hat{\mathbf{R}}_r^{-1}(i) \frac{1}{\sqrt{2\pi\rho}} \exp\left(\frac{-(x_k(i))^2}{2\rho^2}\right) \text{sgn}(d_k(i)) \\ &\quad \times (\mathbf{r}(i) - \mathbf{w}_k(i)x_k(i)) \end{aligned} \quad (4-29)$$

$$\mathbf{f}_k(i+1) = \mathbf{f}_k(i) - \delta_f \hat{\mathbf{R}}_b^{-1}(i) \frac{1}{\sqrt{2\pi\rho}} \exp\left(\frac{-(x_k(i))^2}{2\rho^2}\right) \text{sgn}(d_k(i)) \hat{\mathbf{b}}(i) \quad (4-30)$$

onde ρ é relacionado ao desvio padrão do ruído com desvio padrão σ .

4.3

Método de *Averaging* para Aceleração da Convergência

Nessa seção é descrito um método conhecido como *averaging* que será usado para acelerar a convergência dos algoritmos propostos nas seções anteriores em estruturas de recepção lineares ($\mathbf{f} = \mathbf{0}$). Essa técnica foi proposta inicialmente por Polyak [69] e permite a aceleração da convergência através do uso de passos de convergência maiores. A contribuição desta seção é o uso do conceito de *averaging* em conjunto com algoritmos que minimizam a BER. Para a obtenção de versões dos algoritmos de MBER com a técnica de *averaging* introduz-se a seguinte recursão:

$$\bar{\mathbf{w}}_k(i+1) = (1 - \beta)\bar{\mathbf{w}}_k(i) + \beta\mathbf{w}_k(i+1) \quad (4-31)$$

onde β é o fator de *averaging*.

Se considerarmos o vetor médio de coeficientes do receptor linear ($\mathbf{w}(i)$ e $\mathbf{f} = \mathbf{0}$) e que, para uma adaptação suficientemente lenta dos coeficientes, a matriz $\mathbf{r}(i)\mathbf{r}^T(i)$ se comporta de forma similar ao seu valor esperado $\mathbf{R}(i) = E[\mathbf{r}(i)\mathbf{r}^T(i)]$, e que o vetor $\mathbf{r}(i)$ é suposto independente dos coeficientes anteriores do filtro receptor $\mathbf{w}_k(i-1)$. De acordo com os resultados descritos por Ljung em [70], o uso de (4-31) com um algoritmo do tipo gradiente estocástico e para i grande, o algoritmo se comporta conforme

$$\bar{\mathbf{w}}_k(i+1) = (1 - \beta)\bar{\mathbf{w}}_k(i) + \beta E[\mathbf{r}(i)\mathbf{r}^T(i)]^{-1}\mathbf{z}_w(i) \quad (4-32)$$

onde $\mathbf{z}_w(i)$ é o vetor de gradiente. Note que (4-32) é independente do fator de convergência do algoritmo e depende apenas de β .

Considere que \mathbf{w}_k^* denota o vetor de coeficientes ótimo, que corresponde ao filtro de Wiener se o critério de desempenho do estimador for o MMSE e ao filtro que alcança a mínima taxa de erro para as técnicas AMBER e LBER caso o critério de desempenho seja o de mínima BER. O vetor de erro dos coeficientes $\boldsymbol{\varepsilon}(i)$ denota a diferença entre o vetor de coeficientes do filtro no instante i , $\mathbf{w}_k(i)$, e o vetor ótimo \mathbf{w}_k^* , ou seja $\boldsymbol{\varepsilon}(i) = \mathbf{w}_k(i) - \mathbf{w}_k^*$. Dessa forma, temos

$$\boldsymbol{\varepsilon}(i) = (1 - \beta)\bar{\mathbf{w}}_k(i) - (1 - \beta)\bar{\mathbf{w}}_k^* - \beta\bar{\mathbf{w}}_k^* + \beta\mathbf{R}(i)^{-1}\mathbf{r}(i)\mathbf{r}^T(i) \quad (4-33)$$

$$\boldsymbol{\varepsilon}(i) = (1 - \beta)\boldsymbol{\varepsilon}(i) - \beta\bar{\mathbf{w}}_k^* + \beta\mathbf{R}(i)^{-1}\mathbf{r}(i)\mathbf{r}^T(i) \quad (4-34)$$

Aplicando o operador de valor esperado aos dois lados de (4-34), pode-se

constatar como em [70] que $\mathbf{w}_k(i)$ converge para \mathbf{w}_k^* , ou seja, $E[\mathbf{w}_k(i)] \rightarrow \mathbf{w}_k^*$, na medida que $i \rightarrow \infty$, $\varepsilon(i) \rightarrow 0$ e que essa convergência não depende dos autovalores da matriz autocorrelação do vetor de observação \mathbf{R} . Para os algoritmos gradiente-Newton, o conceito de *averaging* também é capaz de propiciar melhorias na taxa de convergência do algoritmo.

4.4 Simulações

Nessa seção são mostrados experimentos de simulação que foram conduzidos a fim de avaliar o desempenho de convergência e em termos de BER dos algoritmos e estruturas propostos neste capítulo. Devido à grande dificuldade de analisar a convergência de algoritmos que minimizam a BER, e em especial quando estes algoritmos são usados em estruturas não lineares, a investigação dos algoritmos propostos foi baseada em simulações.

Inicialmente, é considerado um experimento em que são comparadas as estruturas com decisão realimentada (DF) e linear operando com os algoritmos baseados no gradiente estocástico com as técnicas LMS, AMBER e LBER em um sistema DS-CDMA no enlace reverso. São considerados canais lineares e seqüências de Gold nas simulações. Além disso, usamos o limiar da técnica AMBER $\tau = 0.2$ para acelerar a convergência do algoritmo e $\rho = 2\sigma^2$ para o método LBER. Os detectores empregam N elementos adaptativos na seção direta e um filtro com K coeficientes na seção realimentada para cada usuário. Note que os receptores lineares são obtidos a partir das expressões dos mesmos com decisão realimentada fazendo-se $\mathbf{f} = \mathbf{0}$ e $\mu_f = 0$. Supõe-se também que o sistema tem controle de potência perfeito.

Para analisar a convergência dos receptores adaptativos com as técnicas LMS, AMBER e LBER, foram realizados experimentos onde se calcula a BER a cada símbolo transmitido. As curvas de convergência foram obtidas com 2000 símbolos de treinamento, e tiradas as médias de 100 experimentos independentes. Os MUDs operam com $K = 4$ usuários no sistema, $E_b/N_0 = 10dB$ e usam os passos $\mu_w = 0,005$ e $\mu_f = 0,005$. Os resultados mostrados na Figura 4.2 indicam que as estruturas com DF são superiores às lineares e que o método LBER é superior ao AMBER, que é melhor do que o LMS. Os esquemas com DF realizam cancelamento de interferência através da seção realimentada do receptor e, desta forma, conseguem um desempenho melhor do que os receptores lineares. Note também que os algoritmos propostos para a estrutura DF baseados nas técnicas LBER e

AMBER superam os mesmos para estruturas lineares, em função do cancelamento dos usuários associados.

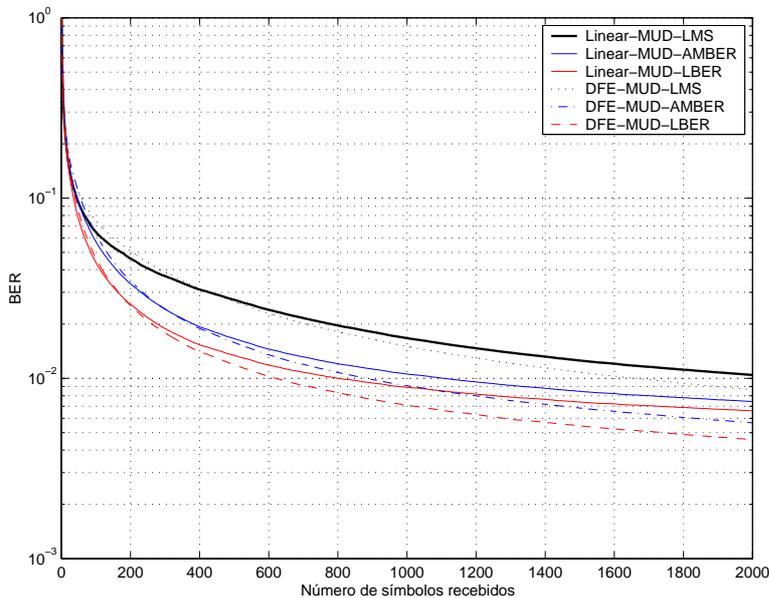


Figura 4.2: Convergência dos algoritmos com receptores lineares e com decisão realimentada para $E_b/N_0 = 10dB$ e canal com função de transferência $H(z) = 1, 2 + 1, 1z^{-1} - 0, 2z^{-2}$, onde o operador z^{-1} introduz um retardo de um *chip* no sinal transmitido.

Os resultados das simulações em termos de BER foram obtidos usando-se 1000 símbolos de treinamento e 10^4 símbolos de dados e tirando-se a média de 100 experimentos independentes. Todos os receptores operam com $N = 7$, o sistema possui $K = 4$ usuários e os algoritmos empregam os passos $\mu_w = 0,0025$ e $\mu_f = 0,0025$ durante a fase de treinamento e a adaptação não ocorre no modo de dados.

A Figura 4.3 ilustra o desempenho em BER dos receptores lineares e DF usando os algoritmos LMS, AMBER e LBER para ajuste dos parâmetros do receptor. As estruturas DF com os algoritmos que minimizam a BER superam em desempenho as estruturas lineares com os mesmos algoritmos e o receptor DF que usa o critério de desempenho MMSE através do LMS. O algoritmo LBER com o MUD DF apresentou desempenho superior ao LMS e ao AMBER com a mesma estrutura DF, e operando com alta E_b/N_0 . Com baixa E_b/N_0 o MUD DF com a técnica AMBER obteve melhor resultado do que os demais. Para os MUDs lineares e alta E_b/N_0 , o algoritmo LBER superou o LMS e o AMBER, enquanto que para baixa E_b/N_0 o método AMBER foi melhor do que os outros algoritmos. De fato, as estruturas DF conseguem economizar até 1 dB em comparação com as lineares, para o mesmo desempenho em BER.

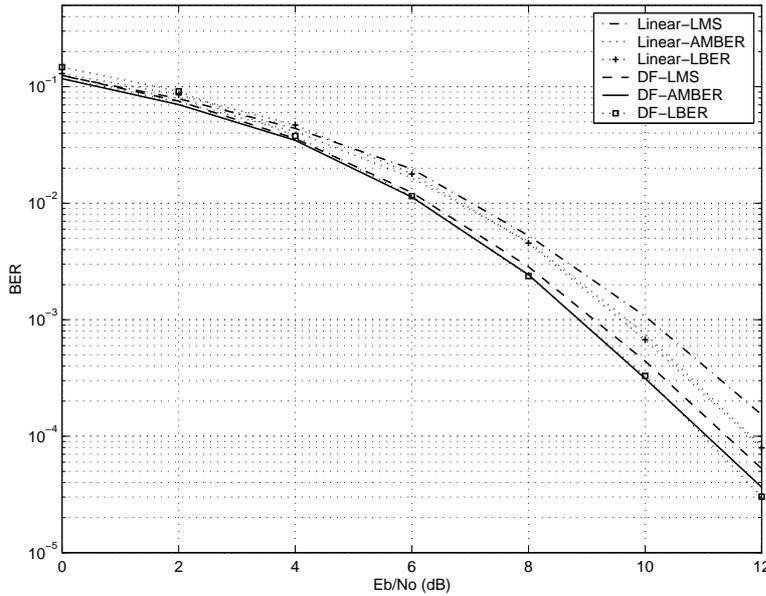


Figura 4.3: Desempenho em termos de BER dos MUDs para o canal com $H(z) = 1, 2 + 1, 1z^{-1} - 0, 2z^{-2}$.

Em uma outra situação, o desempenho em termos de BER versus o número de usuários dos MUDs é mostrado com $E_b/N_0 = 8dB$. Os resultados, mostrados na Figura 4.4, indicam que o MUD DF com o algoritmo LBER alcança o melhor desempenho em BER, seguido do MUD DF com o AMBER e o LMS, e as estruturas lineares com os algoritmos LBER, o AMBER e LMS. De fato, o uso de algoritmos MBER e estruturas DF podem aumentar a capacidade de um sistema DS-CDMA, para um dado valor de BER. Considerando o desempenho do detector convencional (SUD), o aumento de capacidade é bastante significativo para os MUDs DF com algoritmos MBER.

A seguir apresentamos simulações que ilustram o desempenho dos algoritmos gradiente-Newton, que são comparados às técnicas baseadas no gradiente estocástico em receptores lineares. Nesses experimentos o sistema DS-CDMA opera com $N = 8$ no enlace direto, seqüências de espalhamento aleatórias, filtros adaptativos com N elementos e supõe-se que a mesma potência é transmitida para todos os usuários.

Para avaliar o desempenho dos algoritmos em convergência, foram usados 2000 símbolos de dados e tiradas as médias de 100 experimentos independentes. Os parâmetros dos algoritmos foram sintonizados de forma a obter uma convergência rápida e o melhor desempenho em termos de BER. Os algoritmos baseados no gradiente estocástico usam $\mu_w = 0,005$, enquanto as técnicas gradiente-Newton adotam $\alpha = 0,001$ e $\delta_w = 0,0001$.

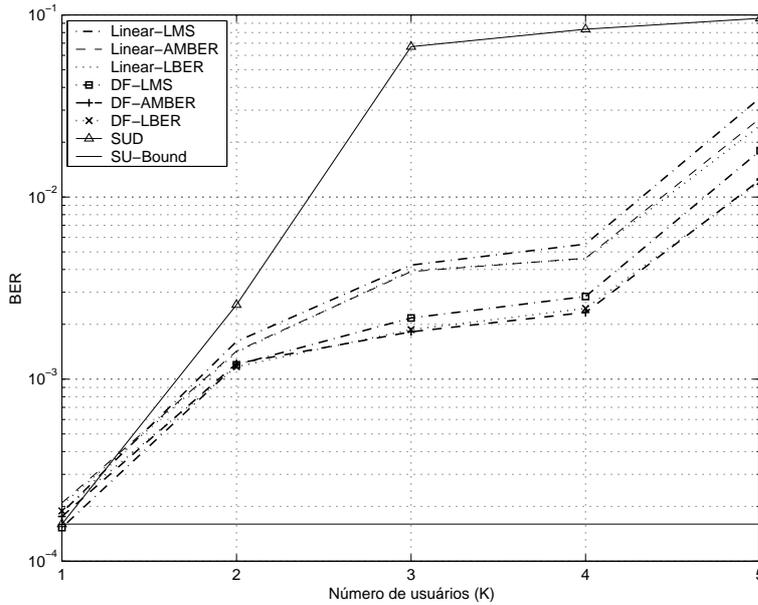


Figura 4.4: Desempenho em BER dos receptores versus número de usuários com $E_b/N_0 = 8dB$ e canal $H(z) = 1, 2 + 1, 1z^{-1} - 0, 2z^{-2}$.

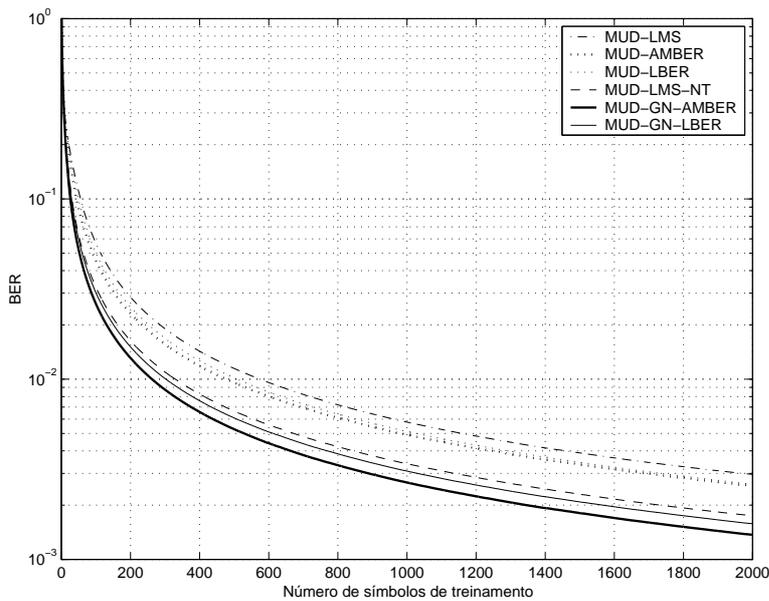


Figura 4.5: Desempenho em convergência dos algoritmos para um receptor linear em um canal com $H(z) = 1, 1 + 1, 2z^{-1} - 0, 2z^{-2}$ e $\frac{E_b}{N_0} = 10$ dB.

Para avaliar o desempenho dos algoritmos em um MUD linear no enlace direto de um sistema DS-CDMA, consideramos um canal com função de transferência $H(z) = 1, 1 + 1, 2z^{-1} - 0, 2z^{-2}$, onde o operador z^{-1} introduz um retardo de um *chip* no sinal transmitido, e a BER é medida a cada símbolo recebido. Os resultados indicam que o algoritmo gradiente-Newton-AMBER (GN-AMBER) obtém o melhor desempenho,

superando o gradiente-Newton-LBER (GN-LBER), o LMS-Newton (LMS-NT), o AMBER, o LBER e LMS com $E_b/N_0 = 10dB$, como mostrado na Figura 4.5. De fato, os algoritmos baseados no método gradiente-Newton conseguem uma convergência mais rápida e para valores inferiores de BER do que as técnicas baseadas no gradiente estocástico.

Foram também realizados experimentos para obter o desempenho em termos de BER versus E_b/N_0 , onde os receptores processaram 200 símbolos de treinamento e 10^4 símbolos de dados, cujas curvas foram obtidas tirando-se a média de 100 experimentos independentes. O número de usuários no sistema era $K = 4$, os MUDs ajustaram os seus parâmetros durante a seqüência de treinamento e não houve adaptação durante a fase de dados. Os parâmetros dos algoritmos foram sintonizados de forma a obter uma convergência rápida e o melhor desempenho em termos de BER. Os algoritmos baseados no gradiente estocástico usam $\mu_w = 0,005$, enquanto as técnicas gradient-Newton adotam $\alpha = 0,001$ e $\delta_w = 0,0001$.

As curvas de desempenho, mostradas na Figura 4.6, indicam que o algoritmo GN-AMBER obteve o melhor desempenho em BER, para baixa e média E_b/N_0 . Para valores mais altos de E_b/N_0 , o GN-LBER apresenta melhor resultado. Note que o algoritmo GN-LBER mostra um desempenho ruim para baixos valores de E_b/N_0 , e bom desempenho para valores altos de E_b/N_0 . Por outro lado, os algoritmos baseados na técnica AMBER têm bom desempenho para baixa e alta E_b/N_0 . O método GN-AMBER é superior ao LMS-Newton para toda a faixa de E_b/N_0 , e requerem uma complexidade computacional inferior já que devido à presença da função indicadora de erro, a atualização dos parâmetros do receptor ocorre com menos freqüência. Além disso, os algoritmos propostos baseados na técnica gradiente-Newton e que minimizam a BER requerem seqüências de treinamento menores do que os seus equivalentes baseados no gradiente estocástico.

O desempenho dos receptores DF em um sistema DS-CDMA no enlace reverso com os algoritmos adaptativos foi avaliado em um canal com 3 percursos seletivo em freqüência e com desvanecimento de Rayleigh cujos coeficientes são dados por $h_{k,l}(i) = a_l|\alpha_{k,l}(i)|$ ($l = 0, 1, 2$ e $k = 1, \dots, K$), onde $\alpha_{k,l}(i)$, $l = 0, 1, 2$, são seqüências independentes de variáveis aleatórias independentes de Rayleigh com potência unitária ($E[\alpha_{k,l}^2(i)] = 1$), que mudam a cada símbolo transmitido. Em todas as situações, os receptores operam com seqüências de espalhamento de Gold com comprimento $N = 15$, processam 200 símbolos em modo de treinamento (TR) e em seguida são colocados em modo *decision-directed* (DD). Além disso, os MUDs empregam N elementos adaptativos na seção direta e um filtro com K coeficientes

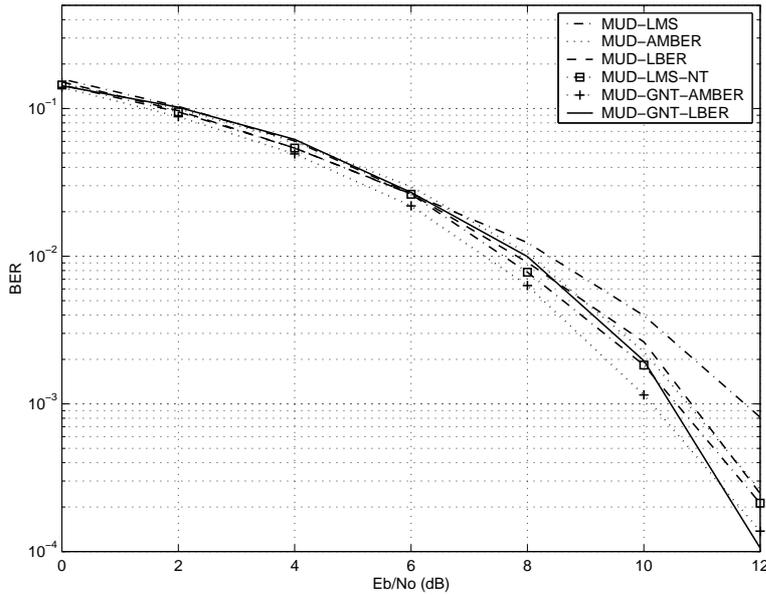


Figura 4.6: Desempenho em BER versus E_b/N_0 para MUDs lineares em um canal $H(z) = 1, 1 + 1, 2z^{-1} - 0, 2z^{-2}$ e AWGN .

na seção realimentada para cada usuário. Os parâmetros otimizados dos algoritmos são: $\mu_w = 0.0075$, $\mu_f = 0.0025$, $\delta_w = 0.00025$, $\delta_f = 0.0001$, $\alpha = 0.0025$, $\rho = 4\sigma^2$ e $\tau = 0.1$. Os parâmetros do canal são $a_0 = 1$, $a_1 = 0,6$ e $a_2 = 0,4$ e supõe-se que o sistema tem controle de potência perfeito.

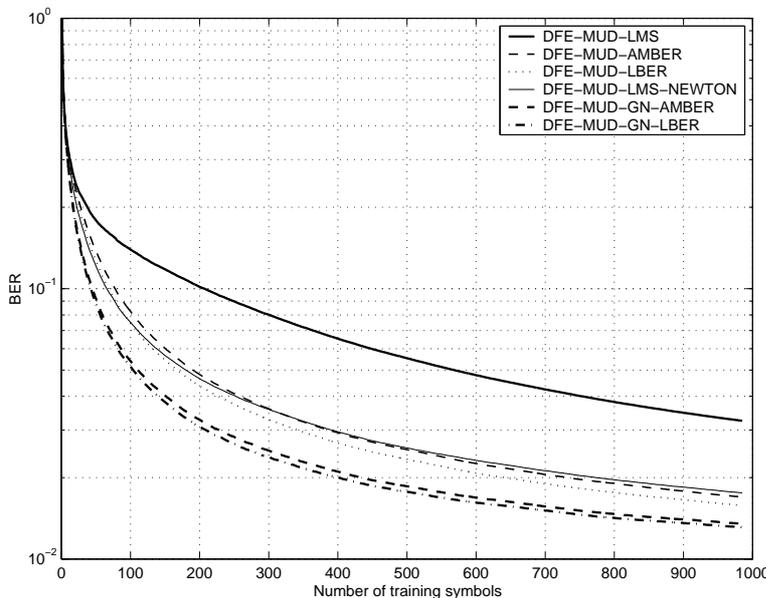


Figura 4.7: Desempenho em convergência dos algoritmos com $E_b/N_0 = 10dB$ em um canal seletivo em frequência com desvanecimento.

O desempenho em convergência dos algoritmos para um sistema com $K = 4$ usuários é mostrado na Figura 4.7, onde os receptores processam

200 símbolos em TR e 800 símbolos em DD, e são tiradas as médias de 100 experimentos independentes. O desempenho médio em BER versus E_b/N_0 é ilustrado na Figura 4.8 para um sistema com $K = 4$ usuários, onde cada MUD processa 10^3 símbolos e são tiradas as médias de 100 experimentos independentes. Na Figura 4.9 é mostrado o desempenho médio em BER versus o número de usuários do sistema. Os resultados indicam que os algoritmos MBER com a técnica gradiente-Newton são superiores aos demais métodos, economizando potência de transmissão e aumentando a capacidade do sistema.

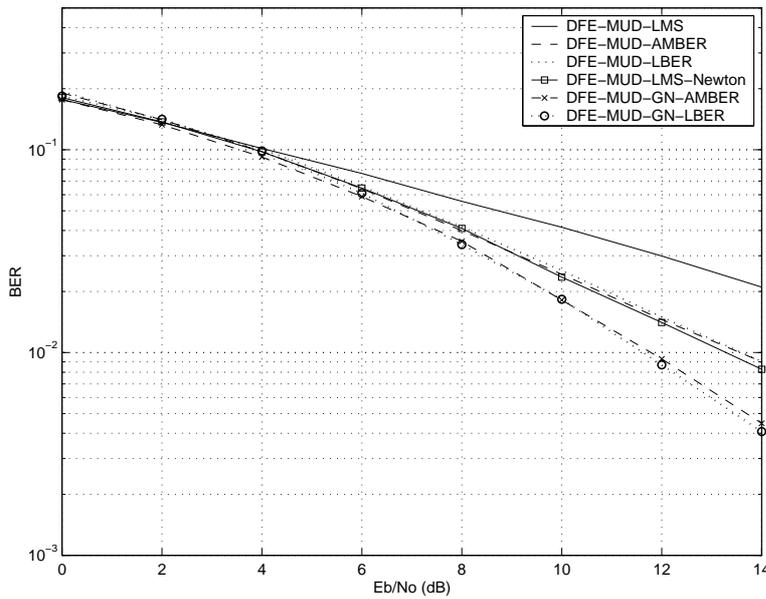


Figura 4.8: Desempenho em BER dos receptores DF com $K = 4$.

Para avaliar os algoritmos com aceleração da convergência dada pelo conceito de *averaging*, é considerado um sistema DS-CDMA no enlace direto com seqüências de espalhamento de comprimento $N = 15$ e frequência da portadora igual a 1900 MHz. A seqüência de coeficientes do canal $h_l(i) = p_l|\alpha_l(i)|$ ($l = 0, 1, 2$), onde $\alpha_l(i)$, $l = 0, 1, 2$, é uma seqüência de variáveis aleatórias complexas gaussianas correlacionadas obtidas passando-se ruído branco gaussiano complexo por um filtro com função de transferência aproximada $\gamma/\sqrt{1 - (f/f_d)^2}$ onde γ é uma constante de normalização, $f_d = v/\lambda_c$ é o máximo deslocamento Doppler, λ_c é o comprimento de onda da frequência da portadora e v é a velocidade do terminal móvel. Os parâmetros do canal são $p_0 = 1$, $p_1 = 0,5$ e $p_2 = 0,3$. Em todas as situações a seguir os MUDs operam com seqüências de Gold com comprimento $N = 15$, processam 200 símbolos em TR e depois são chaveados para o modo DD.

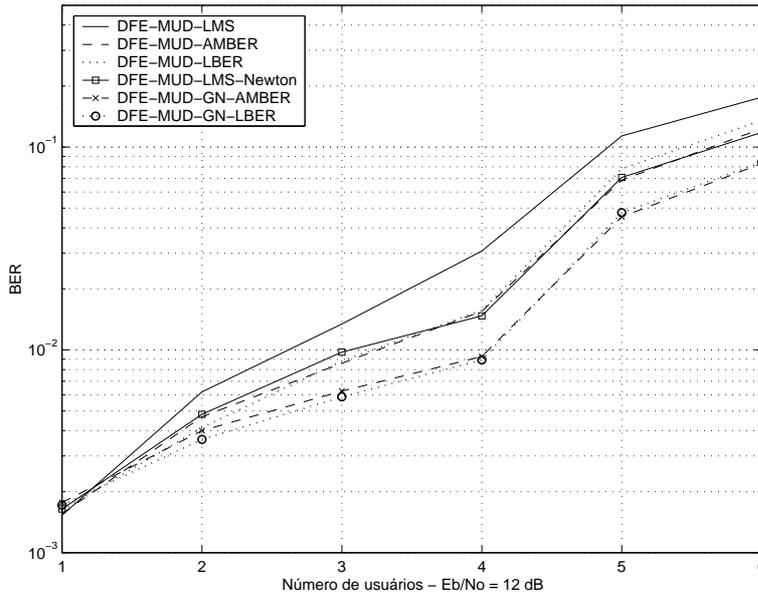


Figura 4.9: Desempenho em BER versus número de usuários dos MUDs DF com $E_b/N_0 = 12$ dB.

Além disso, supõe-se que a mesma potência é transmitida para todos os usuário.

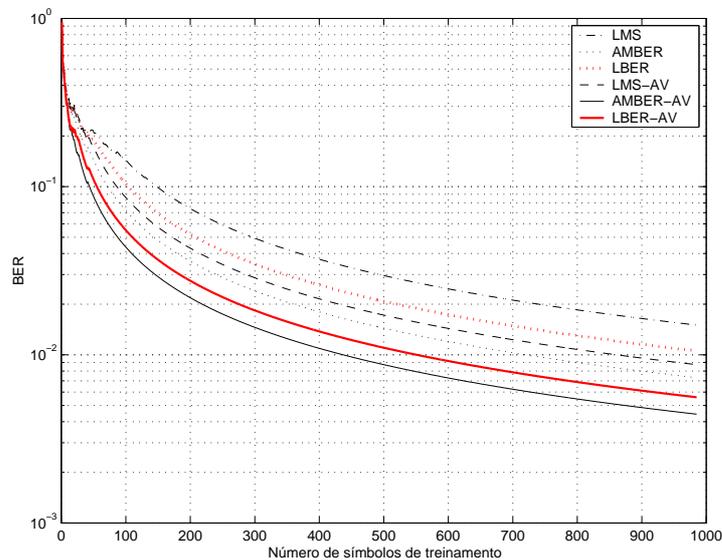


Figura 4.10: Convergência dos algoritmos do tipo gradiente estocástico com $K = 4$ usuários, o móvel move-se com 80 km/h com $E_b/N_0 = 12$ dB. Parâmetros: $\mu_w = 0,001$ para algoritmos convencionais e $\mu_w = 0,0025$ para algoritmos com *averaging*, $\rho = 8\sigma^2$, $\beta = 0,15$ e $\tau = 0,15$.

O desempenho em convergência dos algoritmos com *averaging* em um sistema com $K = 4$ usuários é mostrado nas Figuras 4.10 e 4.11, onde os MUDs processam 200 símbolos em TR e 800 símbolos em DD, e são tiradas

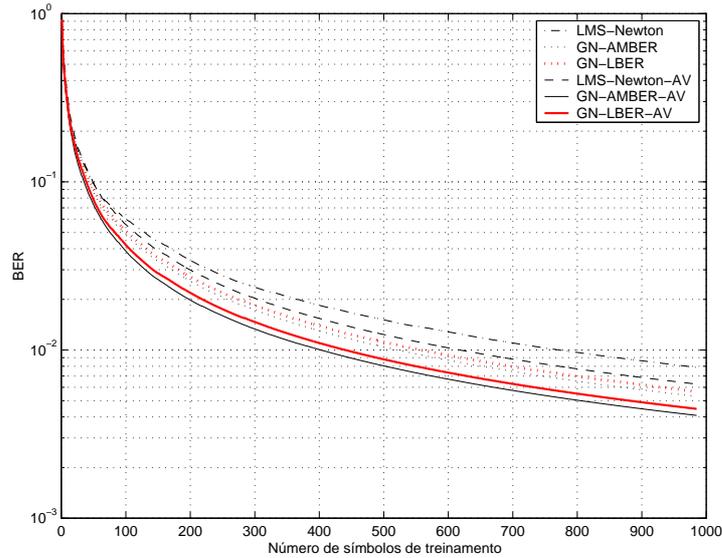


Figura 4.11: Convergência dos algoritmos do tipo gradiente estocástico com $K = 4$ usuários, o móvel move-se com 80 km/h com $E_b/N_0 = 12$ dB. Parâmetros: $\alpha = 0,01$, $\delta = 0,0005$ para algoritmos convencionais e $\delta = 0,001$ para algoritmos com *averaging*, $\rho = 8\sigma^2$, $\beta = 0,15$ e $\tau = 0,15$.

as médias de 100 experimentos independentes. A razão para a escolha de uma seqüência de treinamento com 200 símbolos é que experimentos indicaram que esse era o comprimento mínimo necessário para que os algoritmos MBER superassem os algoritmos com o critério MMSE. O critério usado para selecionar o fator de *averaging* β foi também obtido por simulação e se trata do valor capaz de aproximar a inversa da matriz de autocorrelação do vetor recebido sem que o algoritmo se torne instável. Dessa maneira, sugerimos um valor em torno de $\beta = 0.15$ já que valores maiores podem tornar a técnica instável, enquanto valores menores resultam em um desempenho inferior.

O desempenho em BER versus o número de usuários é mostrado nas Figuras 4.12 e 4.13 onde os receptores processam 10^3 símbolos e tiradas as médias de 100 experimentos independentes. Os resultados mostram que os algoritmos com o método de *averaging* são capazes de acelerar a convergência dos algoritmos, já que eles permitem que os algoritmos usem passos de convergência maiores sem o risco de perder o rastreamento do canal, economizando potência de transmissão e aumentando a capacidade do sistema.

Para os algoritmos baseados no gradiente estocástico, os métodos de *averaging* aumentam consideravelmente o desempenho em convergência, enquanto que para os algoritmos do tipo gradiente-Newton os ganhos em desempenho são menos significantes. Os resultados para os algoritmos do tipo

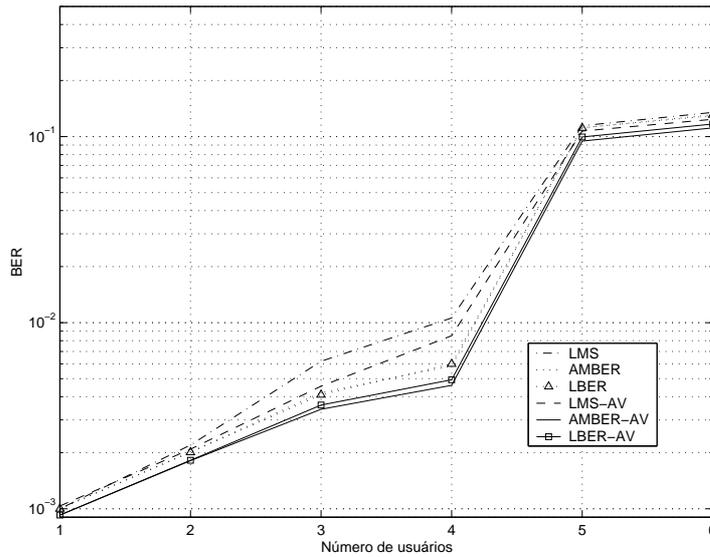


Figura 4.12: Desempenho em BER versus número de usuários para algoritmos do tipo gradiente estocástico, em um cenário onde o móvel move-se com 80 km/h com $E_b/N_0 = 10$ dB. Parâmetros: $\mu_w = 0,001$ para algoritmos convencionais e $\mu_w = 0,0025$ para algoritmos com *averaging*, $\rho = 4\sigma^2$, $\beta = 0,15$ e $\tau = 0,15$.

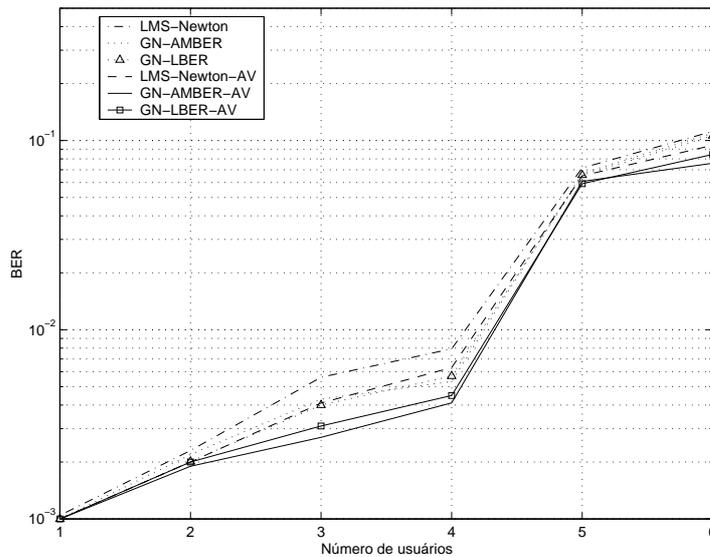


Figura 4.13: Desempenho em BER versus número de usuários para algoritmos do tipo gradiente-Newton, em um cenário onde o móvel move-se com 80 km/h com $E_b/N_0 = 10$ dB. Parâmetros: $\alpha = 0,01$, $\delta = 0,0005$ para algoritmos convencionais e $\delta = 0,001$ para algoritmos com *averaging*, $\rho = 8\sigma^2$, $\beta = 0,15$ e $\tau = 0,15$.

gradiente estocástico, mostrados nas Figuras 4.10 e 4.12, mostram que a técnica AMBER-AV possui o melhor desempenho, seguida da LBER-AV, a AMBER, o LMS-AV, o LBER e o LMS. Para os algoritmos gradiente-Newton, a técnica GN-AMBER é a melhor, seguida da GN-LBER-AV, a GN-AMBER, a GN-LBER, o LMS-Newton-AV e o LMS-Newton, conforme

ilustrado nas Figuras 4.11 e 4.13. Em termos de complexidade computacional, os métodos de *averaging* requerem $2N$ multiplicações e $N + 1$ adições além das operações necessárias pelos algoritmos baseados no gradiente estocástico e gradiente-Newton.