
SOBRE CRITÉRIOS PARA EQUALIZAÇÃO NÃO-SUPERVISIONADA

Aline de Oliveira Neves*

aline.panazio@ufabc.edu.br

Ricardo Suyama[†]

rsuyama@decom.fee.unicamp.br

Romis Ribeiro de Faissol Attux[†]

romisri@decom.fee.unicamp.br

Maria D. Miranda[‡]

maria@lcs.poli.usp.br

João Marcos Travassos Romano[†]

romano@decom.fee.unicamp.br

*Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas - Universidade Federal do ABC, Santo André, SP

[†]DSPCom - Laboratório de Processamento de Sinais para Comunicações - Faculdade de Engenharia Elétrica e de Computação (FEEC) - UNICAMP, Campinas, SP

[‡]Laboratório de Comunicações e Sinais - Universidade de São Paulo, São Paulo, SP

ABSTRACT

In this work, we study the criteria used to solve the blind equalization problem. Two approaches are considered in detail: the constant modulus and the Shalvi-Weinstein criteria. In the course of our exposition, a more recent and less studied technique, the generalized constant modulus criterion, is also discussed. Some of the most important results found in the literature are presented together with some recent contributions related to the comparison between blind criteria and between unsupervised techniques and the Wiener criterion.

KEYWORDS: Adaptive Filtering, Blind Equalization, Constant Modulus Criterion.

RESUMO

Neste artigo são abordados critérios usados para resolver o problema da equalização cega também conhecida como

autodidata. Consideram-se os critérios clássicos do módulo constante e o do Shalvi-Weinstein. Apresentaremos os principais resultados existentes na literatura e alguns resultados mais recentes, que dizem respeito ao estudo do algoritmo do módulo constante generalizado (GCMA) e à comparação entre os critérios citados e destes com o critério de Wiener.

PALAVRAS-CHAVE: Filtragem adaptativa, Equalização cega, Critério do Módulo Constante.

1 INTRODUÇÃO

Desde os primórdios das comunicações digitais, o problema da equalização tem despertado significativo interesse por parte da comunidade científica, o que pode ser atestado pelo grande número de esforços de pesquisa envolvendo seus dois principais aspectos: a escolha da estrutura do filtro (equalizador) e o método de ajuste de seus parâmetros.

No que diz respeito ao primeiro aspecto, vemos uma interessante evolução que, a partir de estruturas lineares relativamente simples, chega às mais modernas técnicas de inteligência computacional, tais como redes neurais e filtros nebulosos, passando por estruturas recursivas e pelos

Artigo submetido em 01/11/2005

1a. Revisão em 13/04/2006

2a. Revisão em 27/07/2006

Aceito sob recomendação do Editor Associado

Prof. José Roberto Castilho Piqueira

consagrados filtros polinomiais. O outro aspecto, por sua vez, surge como um problema típico de otimização: uma vez escolhida a estrutura, como obter dela o melhor desempenho possível do ponto de vista da tarefa que se deseja realizar? As primeiras abordagens, como é natural, basearam-se em informações sobre o canal de comunicação ou na informação transmitida, que, note o leitor, também é o que se busca restaurar. Esses esforços pioneiros deram origem à condição *zero-forcing* e à *teoria de Wiener* (Quresh, 1985).

Entretanto, em muitos casos era difícil o acesso à informação demandada pelas técnicas vigentes. Essa motivação foi o gérmen de uma nova linha de pesquisa: *a teoria de equalização cega ou não-supervisionada*. Como os próprios nomes indicam, a essência das novas abordagens era substituir a informação detalhada fornecida por uma referência ou pelo pleno conhecimento do canal por informações menos restritivas, como, por exemplo, as características gerais ou as estatísticas da mensagem enviada.

Dentre as técnicas e os resultados teóricos que foram se sucedendo com o passar dos anos, duas abordagens mereceram particular atenção: o critério do módulo constante (Godard, 1980), (Treichler and Agee, 1983), dito também *de Godard* e o *critério de Shalvi-Weinstein* (Shalvi and Weinstein, 1990). A importância dessas técnicas foi crescendo à medida que alguns resultados teóricos lhes eram favoráveis em comparação com outros critérios como, por exemplo, o de decisão-direta (DD). Posteriormente, mostrou-se que ambas possuíam uma estreita relação (Regalia, 1999).

Embora haja artigos versando sobre a análise dos critérios supracitados, acreditamos haver ainda uma carência na literatura de trabalhos que procurem expor, de maneira sistemática, a evolução do *corpus* analítico vigente. É exatamente essa tarefa que procuraremos empreender neste trabalho: prover uma visão geral do esforço de pesquisa direcionado ao entendimento dos critérios do módulo constante e de Shalvi-Weinstein, desde sua proposta até os dias atuais. Além disso, no contexto adequado, apresentaremos contribuições recentes ao estudo do algoritmo do módulo constante generalizado (GCMA) (Cavalcanti, 1999) (Cavalcanti et al., 1998b) e também à análise das relações entre os critérios do módulo constante (CM) e de Wiener.

Começaremos, então, apresentando os critérios e algoritmos a eles associados, bem como os primeiros esforços realizados no sentido de se analisá-los, baseados principalmente no estudo das derivadas da função custo, o que será chamado de análise clássica. Em seguida, apresentaremos o importante paradigma geométrico de análise e abordaremos o problema da relação entre os diversos algoritmos autodidatas e destes com o clássico critério de Wiener. O estudo de suas

equivalências e diferenças nos permite obter uma melhor compreensão dos critérios e de seus respectivos algoritmos, o que facilita a escolha do mais adaptado para a resolução de diferentes problemas. Nesse contexto, teremos a oportunidade de expor contribuições recentes: incluímos o critério do módulo constante generalizado (GCM) na análise da relação entre os critérios autodidatas e apresentamos uma forma bastante simples de se enxergar essas relações a partir dos vetores gradiente, em vez da função custo. No que se refere à relação com o critério de Wiener, discutimos uma forma bastante simples e direta de relacioná-lo ao critério do módulo constante (CM). Para tanto, é preciso estabelecer um limitante superior para a função custo do critério CM a partir de um critério supervisionado de quarta ordem e encontrar uma relação entre os pontos de mínimo dos dois critérios.

Este artigo está organizado da seguinte forma: na Seção 2, são apresentados os conceitos básicos da teoria de equalização, incluindo o critério de Wiener. Em seguida, expomos dois resultados da área de equalização não-supervisionada: os teoremas de Benveniste-Goursat-Ruget (BGR) e de Shalvi-Weinstein (SW), essenciais para fornecer um panorama daquilo que será apresentado posteriormente. Na Seção 3, são abordados o critério do módulo constante e os principais resultados provenientes da análise clássica, que baseia-se no estudo direto das derivadas da função custo. A família de critérios de Shalvi-Weinstein e os primeiros frutos de sua análise são apresentados na Seção 4. Na Seção 5, discute-se o critério do módulo constante generalizado, uma extensão ainda pouco explorada da proposta original de Godard. Os fundamentos e principais resultados decorrentes da análise geométrica são discutidos na Seção 6. Na Seção 7, tratamos da equivalência entre critérios autodidatas. Por sua vez, a equivalência destes com o critério de Wiener é estudada na Seção 8. A Seção 9 conclui este artigo.

2 EQUALIZAÇÃO ADAPTATIVA

O principal objetivo de um sistema de comunicação é garantir o adequado intercâmbio de mensagens entre um transmissor e um receptor cuja interconexão se faz através de um canal ou meio de comunicação. Tal meio, no entanto, é usualmente um sistema que distorce a informação de entrada, o que termina por ocasionar uma mudança nas características do sinal enviado que, por sua vez, pode levar a um inaceitável aumento da taxa de bits errados (BER - *Bit Error Rate*) (Proakis, 1995). Em seguida, apresentaremos o modelo do sistema que será utilizado e apresentaremos no que consiste o problema da equalização.

2.1 Características do Canal de Comunicação

Passemos, em primeiro lugar, a uma breve análise das duas principais conseqüências da ação nociva do canal: a *interferência intersimbólica (IIS)* e o *ruído*.

A IIS é um fenômeno que decorre do espalhamento temporal dos pulsos transmitidos devido à limitação da largura de banda do canal. Isto pode ser visto escrevendo-se a saída $x(n)$ de um canal linear e invariante no tempo, tendo $s(n)$ como sinal transmitido (Haykin, 1996) :

$$x(n) = \sum_{k=-\infty}^{\infty} h(k)s(n-k) \quad (1)$$

sendo $h(n)$ a resposta ao impulso do canal. Assumindo que nos interessa o sinal $s(n-d)$, sendo d um atraso inteiro arbitrário, podemos reescrever (1) como:

$$x(n) = h(d)s(n-d) + \sum_{k=-\infty, k \neq d}^{\infty} h(k)s(n-k) \quad (2)$$

Observe que, se os valores da resposta ao impulso do canal não forem todos nulos fora do instante de interesse, haverá um resíduo que se sobreporá ao sinal que se deseja recuperar (o segundo termo, que representa a interferência). Caso tal componente tenha uma potência significativa, pode haver um aumento indesejável da taxa de erro de bit.

Além da IIS é usual considerar no modelo do canal o ruído como um processo aditivo, branco e gaussiano (AWGN - *Additive White Gaussian Noise*). Como o próprio nome já indica, tal modelo considera a existência de um sinal aleatório que se soma àquele que resulta dos demais efeitos do sistema.

A partir dessas definições, chegamos ao modelo padrão que adotaremos ao longo do trabalho:

$$x(n) = h(d)s(n-d) + \sum_{k=-\infty, k \neq d}^{\infty} h(k)s(n-k) + r(n) \quad (3)$$

onde $r(n)$ representa o ruído.

Degradações dessa sorte precisam ser mitigadas para que a qualidade da informação reconstruída atinja um patamar mínimo requerido pela aplicação. Uma contramedida comum a tais efeitos é recorrer a um filtro de recepção denominado *equalizador*. A Figura 1 ilustra o sistema em questão.

Nesta seção inicial, lançaremos as bases teóricas para a análise das diversas soluções do problema de escolher adequadamente os parâmetros de tal filtro.

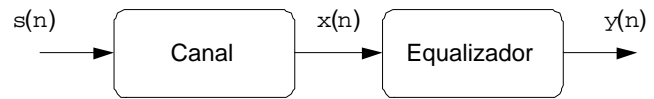


Figura 1: Esquema simplificado de um sistema de comunicação

2.2 O Problema da Equalização

Tendo definido o problema de anular (ou reduzir) a IIS e o ruído, é hora de nos determos na análise da ferramenta que escolhemos para solucioná-lo: o equalizador. Idealmente, gostaríamos que o equalizador, que atua sobre o sinal $x(n)$ representado na equação (2), produzisse uma saída

$$y(n) = \alpha s(n-d) \quad (4)$$

onde α é uma constante arbitrária que não influi no processo de decisão. Em outras palavras, desejamos que o nosso filtro elimine tanto o resíduo produzido pela interferência entre pulsos quanto o ruído. A escolha da estrutura que deverá cumprir essa tarefa é muito importante. É preciso estabelecer um compromisso entre potencial de aproximação e simplicidade prática e analítica do dispositivo. Um candidato que corresponde a uma boa opção na maioria das aplicações práticas, e que doravante adotaremos, é o filtro linear e transversal (FLT) com resposta ao impulso finita, cuja relação entre entrada e saída é dada por:

$$y(n) = \sum_{k=0}^{N-1} w^*(k)x(n-k) \quad (5)$$

onde $w(n)$ é a resposta ao impulso associada ao conjunto de N pesos do filtro e o sobrescrito * significa conjugado. Em notação vetorial, podemos reescrever (5) da seguinte forma:

$$y(n) = \mathbf{w}^H \mathbf{x}(n) \quad (6)$$

onde \mathbf{w} é o vetor de pesos do filtro e $\mathbf{x}(n) = [x(n) \ x(n-1) \ \dots \ x(n-N+1)]^T$ é o vetor de entrada, sendo N o tamanho do filtro. O sobrescrito H significa conjugado e transposto.

É exatamente o vetor \mathbf{w} que determina o comportamento do dispositivo, passível de análise tanto no domínio do tempo quanto no da frequência. Portanto, o projeto de um equalizador desse tipo se resume à escolha dos elementos que formam tal conjunto de parâmetros, como veremos em breve. Desconsiderando por um momento a existência do ruído, e tendo em vista que canal e equalizador são filtros

lineares, podemos escrever:

$$\begin{aligned} y(n) &= [h(n) * w(n)] * s(n) \\ &= c(n) * s(n) \\ &= \sum_{k=-\infty}^{\infty} c(k) s(n-k) \end{aligned} \quad (7)$$

sendo $c(n) = h(n) * w(n)$ a chamada resposta combinada canal + equalizador. Tal formulação é bastante atraente por prover uma ponte direta entre a saída do equalizador e o sinal de entrada. Em particular, ela permite que nos demos conta de um fato interessante: o equalizador só conseguirá atingir a condição que impusemos como sendo de equalização perfeita se $c(n)$ for proporcional a $\delta(n-d)$, sendo $\delta(\cdot)$ a função impulso discreta ou delta de Kronecker. Quando isso ocorre, o equalizador consegue forçar um valor nulo a $c(n)$ em todos os instantes diferentes daquele de interesse: tal é a condição que recebe o nome de *zero-forcing* (ZF).

2.3 Equalização Supervisionada

Apresentado o problema da equalização podemos retornar a questão: como deve ser feito o ajuste do vetor de parâmetros? Cabe observar que se trata, em linhas gerais, de um problema de otimização. Torna-se, portanto, imperativo que escolhamos adequadamente a peça-chave de todo o esquema: o critério, que nos dará a base para avaliar os possíveis candidatos.

Uma possibilidade imediata, em harmonia com a equação (4), seria buscar uma medida da proximidade entre a saída do equalizador e o sinal transmitido. Uma opção comum é a média das distâncias euclidianas entre as amostras de ambos sinais. Se considerarmos que a diferença entre $y(n)$ e $s(n-d)$ pode ser encarada como um *sinal de erro* $e(n)$, chegaremos a uma função custo baseada no *erro quadrático médio* (EQM)

$$J_{Wiener} = E [|s(n-d) - y(n)|^2] = E [|e(n)|^2] \quad (8)$$

onde $E[\cdot]$ representa o operador esperança matemática. O critério apresentado é conhecido como *critério de Wiener*, o mais clássico da teoria de filtragem ótima (Haykin, 1996). A simplicidade de sua formulação, aliada à conveniência, do ponto de vista analítico, de se ter uma função quadrática, faz com que J_{Wiener} seja um termo de comparação inevitável com as demais abordagens. Portanto, convém que nos detenhamos um pouco em seu estudo.

Para que cheguemos a uma solução, devemos minimizar a função J_{Wiener} com respeito ao vetor de parâmetros \mathbf{w} . Felizmente, tal tarefa não será assaz complexa nesse caso, pois a função possui um único mínimo. Calculando o

gradiente de (8) e igualando a zero, chegamos na denominada *solução de Wiener* (Haykin, 1996)

$$\mathbf{w} = \mathbf{R}_x^{-1} \mathbf{p}_d \quad (9)$$

onde $\mathbf{R}_x = E [\mathbf{x}(n) \mathbf{x}^H(n)]$ é a chamada matriz de autocorrelação e $\mathbf{p}_d = E [\mathbf{x}(n) s^*(n-d)]$ é o vetor de correlação cruzada.

Um ponto por nós considerado sobremaneira importante é que a noção de unimodalidade da função de Wiener *só vale caso o atraso de equalização seja fixo*. Isso quer dizer que, em tese, há uma solução de Wiener para cada possível atraso d . Mais ainda, tais soluções não são equivalentes a priori, mas podem ter a si associados erros quadráticos médios significativamente díspares (Johnson et al., 1998) (Suyama et al., 2003).

2.4 Equalização Não-Supervisionada

A proposta de Wiener constitui uma formulação emblemática do problema de *equalização supervisionada*, ou seja, realizada com a ajuda de um sinal de referência. No entanto, apesar de ser conveniente, a presença de uma seqüência de valores desejados pode não ser viável do ponto de vista prático.

A constatação de tal inviabilidade em várias aplicações práticas fez com que houvesse diversas propostas de critérios *não-supervisionados* ou *cegos*. A base de todas essas técnicas é o uso de informações estatísticas acerca do sinal transmitido, que devem ter um papel similar àquele desempenhado pelas amostras de referência no caso da abordagem de Wiener.

Embora algumas propostas pioneiras da área fossem baseadas principalmente em idéias isoladas bastante engenhosas (Lucky et al., 1968) (Sato, 1975) (Godard, 1980), houve dois resultados muito importantes que podem ser considerados a base da teoria de equalização não-supervisionada: os teoremas de Benveniste-Goursat-Ruget (BGR) (Benveniste et al., 1980) e de Shalvi-Weinstein (SW) (Shalvi and Weinstein, 1990). É preciso que, antes de passarmos ao estudo dos critérios que nos dispusemos a analisar, apresentemos essas importantes contribuições.

2.4.1 O Teorema de Benveniste-Goursat-Ruget

A tônica de qualquer técnica cega consiste em substituir o conhecimento das amostras do sinal desejado por informações sobre suas características gerais. É sob tal ponto de vista que o teorema de Benveniste-Goursat-Ruget (BGR) deve ser entendido, pois ele permite que se defina a condição de equalização ideal *sem que se faça menção aos valores do sinal transmitido*. Vejamos como isso é possível.

Consideremos, inicialmente, três condições:

1. O sinal transmitido é formado por amostras independentes e identicamente distribuídas (i.i.d.);
2. Canal e equalizador são filtros lineares e não há adição de ruído;
3. Há a possibilidade de inversão perfeita, o que define as soluções de interesse: configurações ZF, ou seja, do tipo $c(n) = \delta(n - d)$.

Definidas as premissas, resta o problema central: como obter, a partir delas, um critério adequado que dependa apenas de uma fonte de informação conhecida *a priori* no receptor? Estatísticas de segunda ordem, já sabiam os autores, não seriam suficientes para resolver o problema genérico por não portarem informação de fase, essencial para a inversão de um canal de fase não-mínima. Seria necessário, portanto, um ente que, de alguma maneira, contivesse informação acerca de estatísticas de ordem superior a dois. Nesse contexto, afigura-se-nos natural a escolha de Benveniste, Goursat e Ruget: trabalhar com a densidade de probabilidade do sinal transmitido, $p_s(s)$.

A idéia básica passa a ser comparar essa densidade àquela associada à saída do equalizador. Em primeiro lugar, é preciso assumir que $p_s(s)$ não é gaussiana, pois um processo gaussiano filtrado por um sistema linear não-trivial é ainda gaussiano, o que reduz a comparação a um mero ajuste de potência. Resta saber se, quando o sinal de entrada não é gaussiano, a comparação entre densidades de probabilidade pode ser um critério seguro para a obtenção das condições ZF. Exatamente sobre essa questão versa o teorema BGR, que enunciamos a seguir:

Teorema 1 (Teorema BGR) *Sob as condições 1, 2 e 3 previamente definidas, caso as densidades de probabilidade do sinal transmitido (não-gaussiano) e do sinal na saída do equalizador sejam iguais, garante-se que $c(n) = \alpha\delta(n - d)$, onde α representa uma constante de módulo unitário.*

O teorema BGR pode ser considerado um resultado fundamental da teoria de equalização não-supervisionada, pois estabelece solidamente um critério que não leva em conta amostras do sinal piloto. Obedecendo as condições 1, 2 e 3, basta analisar as densidades de probabilidade de entrada e de saída para que se saiba se há ou não perfeita inversão do canal. Como todas as estatísticas de uma variável aleatória podem ser calculadas a partir de sua densidade, implicitamente faz-se uso de todas as estatísticas do sinal.

2.4.2 O Teorema de Shalvi-Weinstein

O teorema de Shalvi-Weinstein (SW) pode ser entendido como uma espécie de refinamento do teorema BGR. As mesmas condições expostas na Seção 2.4.1 são o ponto de partida, mas os autores consideraram a possibilidade de lidar com uma fonte de informação menos restritiva que a presente no teorema BGR. A chave para isso foi empregar uma grandeza denominada *curtose*. Antes de apresentá-la, no entanto, precisamos apresentar brevemente o que são cumulantes.

Cumulantes são grandezas estatísticas derivadas da função característica (Nikias and Petropulu, 1993). Denotamos um cumulante de ordem (p, q) como $C_{p,q}^y$, o que é equivalente a escrever $cum(y : p, y^* : q)$. Os cumulantes de ordem um, dois e três são iguais aos momentos, e portanto, o cumulante de ordem dois $C_{1,1}^y$, por exemplo, é equivalente à variância de y . Os cumulantes de ordem 4, $C_{2,2}^y$, são chamados de *curtose*. Sua definição baseada em momentos é dada por

$$K(y) = E[|y|^4] - 2E^2[|y|^2] - |E[y^2]|^2. \quad (10)$$

Além disso, um sinal é dito sub-gaussiano se sua curtose é negativa. Essa característica é satisfeita na maioria das aplicações relacionadas à comunicação digital, o que é importante, pois alguns critérios que serão apresentados só funcionam para sinais que apresentam essa característica (Godard, 1980).

Tendo definido $K(y)$, podemos enunciar o teorema SW:

Teorema 2 (Teorema SW) *Sob as condições 1, 2 e 3, caso $E[|y(n)|^2] = E[|s(n)|^2]$ e $|K[y(n)]| = |K[s(n)]|$, então necessariamente $c(n) = \alpha\delta(n - d)$, onde α representa uma constante de módulo unitário.*

O teorema de SW estabelece uma condição que depende apenas de duas estatísticas: a variância e a curtose, ou seja, uma condição bem mais simples que aquela subjacente ao teorema de Benveniste-Goursat-Ruget, o qual, como apontamos, tacitamente depende de todos os momentos dos sinais envolvidos.

Cabe observar que os principais critérios para equalização cega dependem, de maneira explícita ou implícita, dos teoremas apresentados. Para nós, terá particular importância neste trabalho o teorema de Shalvi-Weinstein, base do critério homônimo e, como mostraremos na seção 7, do critério do módulo constante, cujo estudo iniciaremos a seguir.

3 O CRITÉRIO DO MÓDULO CONSTANTE

Ainda sem receber o nome que o consagrou, o critério do módulo constante (CMC) surgiu pela primeira vez na literatura como um caso particular de uma proposta de Dominique Godard (Godard, 1980). O objetivo do pesquisador era buscar estratégias viáveis para a adaptação de equalizadores no contexto de redes multiponto. Para atingir tal meta, propôs uma função custo cuja razão de ser era reduzir os efeitos da distorção do canal para que as decisões do receptor se tornassem seguras a ponto de permitirem o emprego de algoritmos convencionais de decisão-direta. Tal função possui uma característica muito especial: é capaz de caracterizar o nível de interferência intersimbólica na saída do equalizador independentemente da fase da portadora. Vejamos sua expressão matemática:

$$J_{Godard} = E \left[(R_p - |y(n)|^p)^2 \right] \quad (11)$$

sendo

$$R_p = \frac{E \left[|s(n)|^{2p} \right]}{E \left[|s(n)|^p \right]} \quad (12)$$

uma constante que contém toda informação necessária acerca do sinal transmitido. A função apresentada em (11) também recebe o nome de "dispersão de ordem p ". Para deduzir seu valor, Godard assumiu circularidade do processo estocástico transmitido e independência entre suas amostras.

Calculando o gradiente de (11), e aplicando a clássica aproximação estocástica (Haykin, 1996) - que substitui a esperança matemática por uma estimativa instantânea -, Godard chegou à seguinte expressão para o algoritmo que hoje leva seu nome:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e^*(n) \mathbf{x}(n) \quad (13)$$

sendo μ o passo de adaptação e

$$e(n) = \left(R_p - |y(n)|^2 \right) y(n) \quad (14)$$

Dois casos particulares do algoritmo de Godard merecem particular atenção: $p = 1$, que dá origem a uma versão modificada do algoritmo de Sato (Sato, 1975) (Haykin, 1996), e $p = 2$, que nos leva ao *algoritmo do módulo constante* (CMA - *Constant Modulus Algorithm*), objeto de nossa análise. O nome "*constant modulus algorithm*" não pode ser encontrado no trabalho original de Godard, havendo sido utilizado pela primeira vez no seminal artigo de Treichler e Agee (Treichler and Agee, 1983), outra fonte de informação sobre os esforços pioneiros na área de equalização autodidata.

Tendo apresentado o critério CM e o algoritmo dele decorrente, é hora de passarmos a um estudo das primeiras

análises de convergência da técnica, efetuadas sob uma hipótese de trabalho muito útil e pouco realista: *comprimento duplamente infinito do equalizador*.

3.1 Análise de Convergência: Equalizadores Duplamente Infinitos

Equalizadores duplamente infinitos são aqueles cuja resposta ao impulso é infinita tanto na parte causal quanto na anticausal, ou seja, se estende de $n = -\infty$ a $n = +\infty$. Tais filtros serão, para nós, apenas abstrações matemáticas cunhadas para permitir uma análise mais direta de técnicas cegas. Entretanto, por cerca de uma década, as principais análises da estrutura da função custo do critério CM se basearam nessa hipótese por um motivo muito simples: ela permitia que fossem obtidos, no domínio da resposta combinada canal+equalizador, os mínimos do critério CM (Foschini, 1985). A grande vantagem de se estudar um equalizador duplamente infinito é que ele permite a inversão perfeita de um sistema FIR (*Finite Impulse Response*), certamente o modelo mais usual de canal de comunicação, sendo tal condição semelhante à hipótese de "inversão perfeita" apresentada na discussão dos teoremas BGR e SW (Benveniste et al., 1980) (Shalvi and Weinstein, 1990).

Já em seu artigo pioneiro, Godard empreendeu tal sorte de análise e chegou à conclusão de que, quando o equalizador é duplamente infinito, todos os mínimos da função custo do critério CM são soluções ZF, ou seja, levam à recuperação ideal do sinal transmitido. Esse resultado significa que o critério não é frágil, pois numa situação ideal leva a resultados coerentes com os objetivos da equalização.

Em 1985, Foschini publicou um célebre trabalho (Foschini, 1985) que esmiuçou as conseqüências da hipótese que vimos analisando. Seu ponto de partida foi, como é a regra no contexto em questão, lidar não com os parâmetros do equalizador, mas sim com os da resposta combinada canal+equalizador. Em tal espaço, Foschini considera interessantes, como era de se esperar, as configurações ZF. Entretanto, ele não se ateve apenas a tal classe de soluções. Empreendendo uma importante análise da matriz hessiana, ele demonstrou a existência de três possíveis tipos de pontos de equilíbrio do CMA no domínio da resposta combinada $c(n)$:

1. Um máximo local, correspondente ao vetor nulo;
2. Mínimos globais, já detectados por Godard, correspondentes a uma solução com apenas um elemento não-nulo;
3. Pontos de sela, correspondentes a soluções com mais de um elemento não-nulo.

Em suma, todos os mínimos são globais e ideais, e os demais pontos de equilíbrio não são estáveis, ou seja, não podem ser considerados parte integrante do regime permanente do algoritmo.

Apesar de ter analisado com profundidade o cenário duplamente infinito, Foschini reconhecia que o estudo mais importante seria aquele realizado no domínio dos parâmetros do equalizador, no qual ocorre efetivamente a adaptação. Entretanto, considerava que a análise empreendida no domínio da resposta combinada deveria, de alguma forma, apresentar pontos de contato com o caso real. No contexto dessa discussão, emergem duas idéias interessantes:

- A noção de que buscar, periodicamente, uma configuração na qual o coeficiente de maior peso esteja no "centro de massa" do vetor de parâmetros pode ajudar a aproximar o comportamento do filtro do ideal. Essa é a essência da inicialização *center-spike*, que consiste em fazer com que o único parâmetro não-nulo do vetor de coeficientes \mathbf{w} esteja na posição central do filtro (Godard, 1980) (Foschini, 1985) (Li and Ding, 1995).
- A idéia de que o aumento da ordem do equalizador poderia permitir que se atingisse um comportamento arbitrariamente próximo do ideal, correspondente ao caso duplamente infinito.

Podemos afirmar que o trabalho de Foschini encerra o ciclo pioneiro de análise do CMA. O próximo passo seria mergulhar numa busca pelos mínimos da função no espaço dos parâmetros, busca esta marcada pela descoberta de uma complexidade maior que a esperada *a priori*.

3.2 Constatação da Convergência para Mínimos Sub-ótimos

O quadro pintado por Godard e Foschini é bastante próximo do ideal: mínimos globais e demais pontos de equilíbrio instáveis. Entretanto, em 1991, Ding e outros (Ding et al., 1991) publicaram um trabalho que revelou quão complexa poderia ser a função custo associada ao CMC. O principal resultado contido nesse artigo é, sem dúvida, a constatação da convergência do CMA para mínimos locais, ou seja, sub-ótimos. Tais soluções representam uma ruptura com os trabalhos pioneiros e também expõem a deficiência do método de análise baseado na resposta combinada.

Basicamente, Ding e outros revelaram que, num caso bastante particular, o CMA poderia convergir para mínimos incapazes de reduzir interferência intersimbólica *mesmo se a estrutura tivesse ordem suficiente para equalizar perfeitamente o canal*. Vejamos como isso se dá. Considere

o canal escolhido com uma função de transferência do tipo:

$$H(z) = \frac{1}{1 + \alpha z^{-N}} \quad (15)$$

onde $|\alpha| < 1$. Para um equalizador de ordem N , seria possível inverter perfeitamente tal sistema, desde que o vetor de parâmetros fosse:

$$\mathbf{w}_{eq} = [1 \ 0 \ \dots \ \alpha]^T \quad (16)$$

Isso quer dizer que, do ponto de vista estrutural, é possível levar a bom termo a tarefa de equalização. Entretanto, os autores também procuraram uma classe de maus mínimos, ou seja, mínimos incapazes de reduzir a IIS:

$$\mathbf{w}_{eq} = \pm g_{eq} [0 \ 0 \ \dots \ \alpha]^T \quad (17)$$

onde g_{eq} representa um ganho.

Impondo a condição de gradiente nulo, é demonstrado que, para que \mathbf{w}_{eq} seja um ponto de equilíbrio do CMA, deve valer a condição:

$$g_{eq} = \left(\frac{E[|x(n)|^p]}{E[|x(n)|^{2p}] R_p} \right)^{1/p} \quad (18)$$

Caso seja obedecida uma condição muito pouco restritiva acerca da constelação transmitida, demonstra-se, a partir da matriz hessiana, que os pontos em questão serão *mínimos da função CM*. Isso prova que o critério pode levar a soluções totalmente ineficientes sob o ponto de vista fundamental de redução da IIS. Por meio de simulações, prova-se também a existência de mínimos locais quando o modelo de canal é MA (*Moving Average*), embora tais soluções não sejam mínimos tão ruins quanto aqueles apresentados na equação (17).

A grande contribuição do trabalho de Ding e outros foi demonstrar a existência de uma nova classe de pontos de equilíbrio do algoritmo CMA: a dos mínimos locais. Entretanto, as conclusões do famoso artigo acabaram por cristalizar em alguns círculos a noção de que o critério CM era inferior a técnicas como a de Wiener por apresentar soluções sub-ótimas. Uma releitura, em nossa opinião, questiona tal idéia por dois motivos:

1. O critério de Wiener também pode ser considerado, como comentamos na seção 2.3, multimodal (Suyama et al., 2003). Há várias soluções de Wiener para um mesmo problema de equalização, algumas delas incapazes de reduzir a IIS a níveis aceitáveis.
2. Canais AR (Auto-Regressivo) e MA são essencialmente distintos, não sendo razoável estabelecer *a priori* nenhuma relação entre a qualidade dos mínimos nos dois casos.

3.3 Validade da Análise no Espaço Combinado

A análise no espaço combinado canal+equalizador foi um expediente muito comum nos primeiros trabalhos sobre técnicas para equalização cega. Entretanto, apenas em 1992 ficou estabelecido o domínio de validade da abordagem, através de um trabalho escrito por Ding e outros (Ding et al., 1992). A principal conclusão de tal artigo é que a análise no espaço combinado canal + equalizador só equivale à análise do domínio dos parâmetros do filtro caso a matriz de convolução tenha um espaço nulo trivial. A matriz de convolução pode ser encontrada, por exemplo, escrevendo a resposta combinada canal+equalizador, definida em (7), na forma vetorial:

$$\mathbf{c} = \mathbf{H}\mathbf{w} \quad (19)$$

sendo $\mathbf{c} = [c_0 \ c_1 \ \dots \ c_{M+N-1}]^T$, com $M - 1$ a ordem do canal e $N - 1$ a ordem do equalizador, e \mathbf{H} é a matriz de convolução do canal com $(M + N)$ linhas:

$$\mathbf{H}^T = \begin{bmatrix} h_0 & h_1 & \dots & h_{M-1} & 0 & \dots & 0 \\ 0 & h_0 & h_1 & \dots & h_{M-1} & 0 & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & h_0 & h_1 & \dots & h_{M-1} \end{bmatrix} \quad (20)$$

No que se refere à tarefa de equalização, \mathbf{H} só possuirá um espaço nulo trivial em dois casos:

- Quando o canal é um simples ganho;
- Quando o equalizador é duplamente infinito.

O primeiro caso é relativamente simples, e o segundo é justamente o caso pouco prático que norteou as conclusões de Godard e Foschini. Isso revela a enorme limitação do antigo paradigma de análise e sacramenta a ruptura iniciada com a constatação da existência de mínimos locais na função CM.

3.4 Classes de Mínimos do Critério CM

Em 1995, Johnson e Anderson (Johnson and Anderson, 1995) escreveram um artigo que consideramos uma espécie de síntese dos primeiros esforços direcionados à análise dos pontos de equilíbrio do CMA. Através de um elegante tratamento matemático, eles dividiram tais pontos em três classes, cada qual caracterizada pela forma como anula o vetor gradiente da função custo:

- A classe dos máximos é composta apenas pela origem do espaço de parâmetros do equalizador, sempre um

ponto de equilíbrio do algoritmo. Por se tratar de um vetor formado apenas por zeros, o elemento desta classe anula o gradiente da forma mais trivial possível.

- A classe de Mazo é composta pelos mínimos globais e pelos pontos de sela, os únicos pontos, além do máximo, que guardam semelhança com os pontos críticos encontrados no caso duplamente infinito.
- A classe de Ding é composta pelos mínimos locais, que surgem quando a matriz de convolução do canal não tem um espaço trivial nulo. Esses mínimos correspondem a soluções como aquelas discutidas na Seção 3.2.

3.5 A Localização dos Pontos de Mínimo

Em 1996, Li e Ray Liu (Li and Liu, 1996) obtiveram resultados interessantes sobre a localização dos mínimos do critério CM. Definindo uma função custo genérica J_{GEN} como:

$$J_{GEN} = E \{ \Phi [y(n)] \} \quad (21)$$

eles mostram que, caso a saída esteja muito próxima do sinal transmitido (o que é o ideal da equalização), é possível expressar o mínimo que produz tal estado através de uma expressão do tipo:

$$\mathbf{w}_{glob} = f(0)\mathbf{R}_f^{-1}\mathbf{h} \quad (22)$$

que tem uma forma notavelmente similar à da solução de Wiener. A matriz \mathbf{R}_f depende das características do canal e do sinal transmitido e \mathbf{h} é um vetor que contém coeficientes do canal. Esse é o ponto alto de toda a análise. No entanto, destacamos que (22) decorre de uma aproximação de $y(n)$ em torno de $s(n)$, cuja solidez depende da proximidade entre essas duas variáveis.

4 A FAMÍLIA DE CRITÉRIOS DE SHALVI E WEINSTEIN

Uma outra família de critérios muito importante em equalização cega é a que foi obtida diretamente do teorema de Shalvi e Weinstein. Partindo da equação (7), Shalvi e Weinstein, em (Shalvi and Weinstein, 1990) (Shalvi and Weinstein, 1994), calculam o cumulante de ordem (p, q) de $y(n)$ em função do cumulante do sinal transmitido $s(n)$ e da resposta combinada canal+equalizador:

$$C_{p,q}^y = C_{p,q}^s \sum_l c_l^p (c_l^*)^q \quad (23)$$

Retomando o teorema de SW (Seção 2.4.2), se $C_{1,1}^y = C_{1,1}^s$, usando (23), $\sum |c_l|^2 = 1$, e portanto, para $p + q > 2$, $|C_{p,q}^y| \leq |C_{p,q}^s|$, com igualdade se e somente se

a equalização for perfeita (solução ZF). Podemos então encontrar os pesos do equalizador de forma a maximizar $C_{p,q}^y$. Matematicamente, o critério é:

$$\max |C_{p,q}^y| \text{ sujeito a: } C_{1,1}^y = C_{1,1}^s \quad (24)$$

sendo chamado de *critério de Shalvi-Weinstein com restrição* (Shalvi and Weinstein, 1994). Como (24) só depende do conhecimento da variância de $s(n)$, ou seja, $C_{1,1}^s$, e da informação de que seu cumulante de ordem (p, q) é não-nulo para algum p e q tais que $p + q > 2$, o critério proposto é universal no sentido de não impor nenhuma outra restrição à distribuição de probabilidade do sinal de entrada, diferentemente do critério CM, onde o sinal deve ser sub-gaussiano. Por ser universal, o critério proposto pode ser usado em uma grande quantidade de problemas de desconvolução.

A restrição no critério dado por (24) pode ser substituída por uma normalização, o que leva à seguinte família de critérios normalizados (Shalvi and Weinstein, 1994):

$$\max \left\{ \frac{|C_{p,q}^y|}{(C_{1,1}^y)^{\frac{p+q}{2}}} \right\} \quad (25)$$

Supondo que $p = q$, o critério (25) pode ser escrito em termos dos símbolos transmitidos, $s(n)$, da seguinte forma:

$$\frac{|C_{2p}^y|}{(C_2^y)^p} = \frac{|C_{2p}^s|}{(C_2^s)^p} \underbrace{\left(\frac{\sum_l |c_l|^{2p}}{(\sum_l |c_l|^2)^p} \right)}_{f_{2p}(c)} \quad (26)$$

Observe que a função $f_{2p}(c)$ pertence ao intervalo $(0, 1]$ para qualquer c , sendo que o limite superior é alcançado se e somente se c possuir um único elemento não-nulo (solução ZF) (Donoho, 1981) (Shalvi and Weinstein, 1990). Como as estatísticas do sinal transmitido $s(n)$ podem ser consideradas conhecidas, a maximização de (25) é equivalente à maximização de $f_{2p}(c)$, que, por sua vez, é o próprio critério proposto por Donoho (Donoho, 1981) no contexto de sinais reais, do qual um caso particular com $p = 2$ é o critério de desconvolução proposto inicialmente por Wiggins (Wiggins, 1978).

É interessante observar que, apesar da normalização, não existe em (25) nenhuma condição que relacione a potência do sinal na saída do equalizador à do sinal de entrada do sistema, como acontece em (24). Observa-se que, apenas quando se ajusta $C_{1,1}^y$ para ser igual a $C_{1,1}^s$, o critério (25) coincide com (24) (Tugnait, 1992). Entretanto, se esta igualdade não for considerada, e como $f_{2p}(c) = f_{2p}(\alpha c)$, onde α é uma constante, o critério apresentará outros pontos de máximo, havendo portanto uma ambigüidade de ganho (Regalia and Mboup, 1999).

Em vez da restrição existente em (24), pode-se também adicionar ao critério um funcional da potência do sinal de saída do equalizador $y(n)$ como termo de penalidade. Tem-se então a seguinte família de critérios sem restrição:

$$\max \{ |C_{p,q}^y| + f(C_{1,1}^y) \} \quad (27)$$

onde $f : [0, \infty) \rightarrow \Re$ é uma função real e contínua por partes que deve respeitar algumas restrições. Considerando a seguinte escolha para $f(\cdot)$:

$$f(C_{1,1}^y) = \sum_{i=0}^M k_i (C_{1,1}^y)^i \quad (28)$$

onde k_0, k_1, \dots, k_M são constantes reais, e fazendo $k_0 = r^2, k_1 = 2r, k_2 = -2$ e $k_i = 0$ para $i = 3, \dots, M$ temos um caso particular de (27), que resulta no critério:

$$\max \{ |C_{2,2}^y| - 2(C_{1,1}^y)^2 + 2rC_{1,2}^y + r^2 \} \quad (29)$$

onde r é uma constante real. Se $s(n)$ for uma variável aleatória complexa, com media nula e sub-gaussiana ($C_{2,2}^y < 0$), então o critério acima se reduz a (Shalvi and Weinstein, 1990):

$$\min E \left[\left(|y(n)|^2 - r \right)^2 \right] \quad (30)$$

Note que, devido à condição de sub-gaussianidade, os pontos estacionários que nos interessam são os pontos de mínimo da função custo e não mais os pontos de máximo. Se escolhermos $r = R_p$ (12), chegaremos ao critério proposto por Godard (Godard, 1980), definido em (11).

4.1 Algoritmos do Tipo Super-Exponencial

O algoritmo de Shalvi-Weinstein, obtido a partir do critério com restrição (24), é conhecido também como algoritmo super-exponencial (ASE). Diferentemente do CMA, este algoritmo não é deduzido simplesmente se realizando uma aproximação estocástica do gradiente da função custo. A técnica foi proposta em (Shalvi and Weinstein, 1993), e, apesar de haver tido seu bom desempenho comprovado por simulações, foi deduzida num contexto bastante restrito. Uma extensão deste trabalho foi realizada por P. Regalia e M. Mboup em (Regalia and Mboup, 1999). Assim, basear-nos-emos neste artigo para, em seguida, explicar brevemente no que consiste o algoritmo.

Denotaremos por T o conjunto de todas as respostas combinadas canal+equalizador c (19) que podem ser obtidas a partir de um equalizador finito. Esse conjunto é chamado de *conjunto de estados alcançáveis*. Assim, se, por exemplo, tivermos um canal que não possa ser invertido por um

equalizador FIR, sabemos que a solução ZF nunca poderá ser atingida, e, portanto, não pertencerá a T .

Em seguida podemos definir o operador P_T , que nada mais é do que o operador de projeção do conjunto de todos os valores possíveis de \mathbf{c} em T . Matematicamente este pode ser escrito como: $P_T = \mathbf{H} (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H$, onde \mathbf{H} é a matriz de convolução definida em (20). Assim, se T contém todos os valores possíveis de \mathbf{c} , incluindo a solução ZF, P_T será igual a matriz identidade, e diremos que, neste caso, o equalizador possui ordem suficiente. No entanto, se T não contém a solução ZF (caso bem mais comum), P_T será diferente da matriz identidade, e o equalizador passará a ser chamado de *sub-parametrizado*.

Precisamos então encontrar quais são os pontos de máximo de $f_{2p}(\mathbf{c})$, definida em (26). Em (Regalia and Mboup, 1999), os autores mostram que $\mathbf{c} \in T$ é um ponto estacionário de $f_{2p}(\mathbf{c})$ se e somente se

$$P_T \left(\mathbf{c}^{\odot(2p-1)} \right) = \alpha \mathbf{c} \quad (31)$$

para algum escalar α . Na expressão (31), $\mathbf{c}^{\odot m} = [c_0^m \ c_1^m \ \dots \ c_{M+L-1}^m]$ o que é chamado de m -ésima potência de Hadamard de \mathbf{c} . Denotaremos por $\mathbf{c}^{\odot <2p-1>}$ o vetor $\mathbf{c}^{\odot p} \odot (\mathbf{c}^*)^{\odot <p-1>}$ cuja k -ésima componente pode ser escrita como $[\mathbf{c}^{\odot <2p-1>}]_k = |c_k|^{2p-2} c_k$.

Considerando o caso do equalizador com ordem suficiente, ou seja, em que $P_T = I$, o problema já está resolvido, pois (31) se reduz a $\mathbf{c}^{\odot <2p-1>} - \alpha \mathbf{c} = 0$, que, em termos de componentes, pode ser escrito como:

$$c_k \left(|c_k|^{2p-2} - \alpha \right) = 0, \text{ para todo } k \quad (32)$$

Isso significa que todos os termos não-nulos possuem mesma amplitude, dada por $\sqrt[2p-2]{\alpha}$. Godard já havia chegado nesta conclusão em (Godard, 1980), mas de uma forma que não incluía o caso do equalizador sub-parametrizado. Neste caso, $\mathbf{c}^{\odot <2p-1>} - \alpha \mathbf{c} \neq 0$, não sendo, portanto, possível encontrar uma solução com forma fechada para \mathbf{c} . Uma saída é encontrar os pontos de máximo de $f_{2p}(\mathbf{c})$ de maneira iterativa. O seguinte mapeamento não-linear $D_{2p}(\mathbf{c})$ possui, como pontos fixos, os mesmos pontos estacionários de $f_{2p}(\mathbf{c})$ (Regalia and Mboup, 1999):

- Considere $\mathbf{c} \in T$, com norma unitária ($\|\mathbf{c}\| = 1$). A escolha da norma é arbitrária;
- Projete sua potência de Hadamard em T : $\mathbf{v} = P_T \left(\mathbf{c}^{\odot(2p-1)} \right)$;
- Faça com que o resultado possua norma unitária: $\mathbf{q} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ (mesma norma utilizada no passo 1);

- Os pontos fixos deste mapeamento, ou seja, $\mathbf{q} = D_{2p}(\mathbf{c}) = \mathbf{c}$ são pontos que satisfazem (31), maximizando $f_{2p}(\mathbf{c})$.

Em (Mboup and Regalia, 1999) e (Mboup and Regalia, 2000), os autores mostram que o algoritmo super-exponencial exposto acima é equivalente a um algoritmo de busca pelo gradiente, cuja função custo pertence à família de funções de Donoho ($f_{2p}(\mathbf{c})$), podendo ser escrito como:

$$\begin{aligned} \mathbf{v} &= \mathbf{c}(n) + \mu P_T \nabla f_{2p}(\mathbf{c}(n)) \\ \mathbf{c}(n+1) &= \frac{\mathbf{v}}{\|\mathbf{v}\|} \end{aligned} \quad (33)$$

sendo $\mu = \frac{1}{2p f_{2p}(\mathbf{c}(n))}$ o passo de adaptação. Assumiu-se um canal real, estável, causal e que não exista adição de ruído.

O algoritmo (33) é equivalente ao mapeamento $D_{2p}(\mathbf{c})$. Utilizando essa nova abordagem, a convergência do algoritmo pode ser demonstrada de uma forma bem mais simples do que a realizada em (Regalia and Mboup, 1999).

Tendo então o algoritmo no espaço da resposta combinada canal+equalizador, e fazendo uso de (19), é possível encontrar o algoritmo no espaço de parâmetros do equalizador. As expressões abaixo mostram os resultados considerando-se cumulantes de quarta ordem (curtose):

$$\mathbf{w}_{n+1} = \mathbf{w}_{n+1} + \frac{\beta}{\delta} \mathbf{Q}_n \mathbf{x}_n^* y_n \left[|y_n|^2 - \frac{E(|s_n|^4)}{E(|s_n|^2)} \right] \quad (34)$$

onde β é um escalar constante, $\gamma = C_{1,2}^s / C_{1,1}^s$, $\delta = C_{2,2}^s / C_{1,1}^s$ e \mathbf{Q} é proporcional à inversa da matriz de autocorrelação do sinal de saída do canal, $x(n)$. Sua expressão recursiva é dada por:

$$\mathbf{Q}_{n+1} = \frac{1}{1-\beta} \left(\mathbf{Q}_n - \frac{\beta \mathbf{Q}_n \mathbf{x}_n^* \mathbf{x}_n^T \mathbf{Q}_n}{1-\beta + \beta \mathbf{x}_n^T \mathbf{Q}_n \mathbf{x}_n^*} \right) \quad (35)$$

Uma explicação detalhada do processo de obtenção destas expressões pode ser encontrada em (Shalvi and Weinstein, 1994). Cabe observar entretanto, que só é possível chegar a (34) com aproximações empíricas para os cumulantes envolvidos. Vamos nos referir a (34) como ASE (Algoritmo Super-Exponencial).

Observe que o algoritmo super-exponencial procura os pontos de máximo de $f_{2p}(\mathbf{c})$, mas considerando a restrição de potência, ou seja, $C_{1,1}^y = C_{1,1}^s$, o que evita a ambigüidade de ganho. Dessa forma, o algoritmo pode ser visto como uma forma de otimizar o critério dado por (24).

Além disso, a partir de (33) é possível se mostrar uma característica muito importante do ASE: seu passo de adaptação μ é ótimo no sentido da velocidade de

convergência. Este passo, definido em (33), é equivalente, no domínio de parâmetros do equalizador, a $\frac{\beta}{\delta} \mathbf{Q}$, usado em (34) (Mboup and Regalia, 2000).

4.2 Análise de Convergência

Como no caso do critério CM, os primeiros estudos de convergência realizados para os critérios de Shalvi-Weinstein consideraram o caso ideal de um equalizador duplamente infinito. Além disso, a análise é feita no espaço da resposta combinada canal+equalizador. Em (Shalvi and Weinstein, 1990), os autores mostram que nessas condições, o critério com restrição possui um único ponto de máximo (global) que corresponde à solução ZF. Todos os outros pontos estacionários existentes serão pontos de mínimo ou de sela.

Em (Shalvi and Weinstein, 1994), os autores vão um pouco mais além, e concluem também que, para os casos em que $p \neq q$, a função custo apresenta menos pontos de sela. Além disso, na vizinhança destes pontos estacionários, a função se mostra consideravelmente plana (sendo essencialmente quadrática), o que pode levar um algoritmo baseado na busca do gradiente a ter uma convergência lenta.

Já com relação ao critério normalizado sem restrição, o único ponto de máximo ocorre quando $c_n = g\delta(n-d)$, onde d é um atraso e g é um ganho (Shalvi and Weinstein, 1994). Isso porque a função custo é insensível à multiplicação da resposta combinada por uma constante. Neste caso então, ao invés de um ponto de máximo global, a função irá possuir *um vale de soluções possíveis*.

Sob condições práticas, onde o equalizador é finito, existe adição de ruído e o canal pode conter não linearidades, Shalvi e Weinstein (Shalvi and Weinstein, 1994) comentam brevemente que pontos de sela podem vir a se tornar máximos locais. Segundo (Ding et al., 1992), deve-se tomar cuidado ao se estender resultados de análises realizadas no espaço da resposta combinada canal+equalizador para o espaço de parâmetros do equalizador. Análises sob condições reais, com equalizadores finitos, foram realizadas ou através do paradigma geométrico, ou através de equivalências entre os critérios de Shalvi-Weinstein e o critério CM. Estas duas abordagens serão tratadas com detalhes na segunda parte deste artigo.

Quanto ao critério sem restrição, para a escolha de $f(\cdot)$ como mostrado em (27), valem todas as análises já realizadas para o critério CM.

5 O CRITÉRIO DO MÓDULO CONSTANTE GENERALIZADO

O critério e o algoritmo do módulo constante generalizado (GCM - *Generalized Constant Modulus*) foram primeiramente propostos por Cavalcanti e outros em (Cavalcanti et al., 1998b) (Cavalcanti, 1999). Os autores partem do princípio de que, para sinais que já possuem módulo constante, como no caso de modulações PSK (*Phase Shift Keying*), qualquer sinal de módulo constante que for obtido na saída do equalizador representa uma equalização bem-sucedida, independentemente do valor deste módulo. Portanto, o critério corresponde a uma generalização do critério CM, penalizando desvios do sinal de saída do equalizador com relação a um módulo constante não especificado. Note que, assim, o critério proposto possuirá um número virtualmente infinito de soluções. O critério minimiza a seguinte função custo:

$$J_{GCM} = \frac{E \left[\left(|y(n)|^2 - E \left[|y(n)|^2 \right] \right)^2 \right]}{\left(E \left[|y(n)|^2 \right] \right)^2} \quad (36)$$

O algoritmo que busca otimizar este critério, deduzido em (Cavalcanti et al., 1998b), é obtido de uma forma bastante particular. Em vez de calcular a aproximação estocástica do gradiente da função custo dada em (36), como é o caso na obtenção do CMA, o autor calcula o gradiente a partir da função custo instantânea. O algoritmo resultante é dado por:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu \frac{4}{\bar{z}^2(n)} \left((|y(n)|^2 - \bar{z}(n)) y(n) \mathbf{x}(n)^* - \mu \frac{4}{\bar{z}^2(n)} \left(\frac{|y(n)|^4}{\bar{z}(n)} - |y(n)|^2 \right) \mathbf{b}(n) \right) \quad (37)$$

onde $\bar{z}(n)$ e $\mathbf{b}(n)$ são obtidos através de médias temporais :

$$\begin{aligned} \bar{z}(n+1) &= \lambda \bar{z}(n) + (1-\lambda) |y(n)|^2 \\ \mathbf{b}(n+1) &= \lambda \mathbf{b}(n) + (1-\lambda) y(n) \mathbf{x}^*(n) \end{aligned} \quad (38)$$

Este algoritmo possui uma velocidade de convergência entre a do algoritmo CMA (mais lento) e o ASE (mais rápido), com um custo computacional mais baixo que o ASE já que não existe a necessidade de uma inversão de matriz (Neves et al., 2002).

Dado que o critério GCM é mais recente do que os critérios CM e SW, não foi realizada uma análise de sua convergência como a apresentada para os outros critérios. De fato, tal análise começou com sua comparação com os outros critérios, o que será apresentado na seção 7.

6 ANÁLISE GEOMÉTRICA SOBRE POSIÇÃO E CLASSIFICAÇÃO DE MÍNIMOS

A análise sob uma abordagem geométrica é mais recente do que as que foram vistas até agora. Tal abordagem teve início com um artigo de Li e Ding (Li and Ding, 1995), onde os autores propõem uma forma de análise da convergência de algoritmos autodidatas baseada nos chamados cones de mínimo global único. Os autores obtêm resultados muito interessantes no que diz respeito à convergência do algoritmo de Shalvi-Weinstein e à localização de mínimos para o caso do equalizador de comprimento finito, concluindo com a proposta de um algoritmo para inicialização do equalizador. Devido às equivalências existentes entre os critérios de Shalvi-Weinstein e Godard, que serão discutidas na próxima seção, os resultados mostrados a seguir também são válidos para o critério do módulo constante.

A análise de Li e Ding se baseia essencialmente em três conceitos:

1. *Conjunto de estados alcançáveis* T : é o conjunto de todas as configurações c no espaço conjunto canal-equalizador que podem ser obtidas a partir de um equalizador finito.

$$T : \{c; c_n = \sum_{k=N_1}^{N_2} w_k h_{n-k}, w_k \in \mathbb{R}\} \quad (39)$$

onde w_k são os coeficientes do equalizador e h é a resposta impulsiva do canal.

A idéia do conjunto de estados alcançáveis é englobar as possíveis respostas conjuntas produzidas pela operação de convolução, que faz o papel de uma espécie de mapeamento. Por exemplo, se tanto o canal quanto o equalizador forem filtros FIR (*Finite Impulse Response*), será impossível que obtenhamos uma situação de inversão perfeita, ou seja, nenhuma solução *zero-forcing* (ZF, solução que permite a recuperação exata do sinal transmitido) pertencerá ao conjunto T .

2. *Cones com mínimo global único* S :

$$\begin{aligned} S_n^+ & : \{c \in \ell^1(\mathbb{R}) / c_n > 0 \text{ e } c_n > |c_k|, k \neq n\} \\ S_n^- & : \{c \in \ell^1(\mathbb{R}) / c_n < 0 \text{ e } -c_n > |c_k|, k \neq n\} \end{aligned} \quad (40)$$

$$(41)$$

onde $\ell^1(\mathbb{R})$ é um espaço vetorial real.

Conforme a definição nos mostra, os cones são definidos essencialmente pela posição e pelo sinal do elemento de maior magnitude do vetor c . Decorre de

tal construção que cada cone possuirá um único mínimo global (ZF), a saber, $c = [0 \ 0 \ \dots \ \pm 1 \ \dots \ 0 \ 0]^T$, ou seja, o único elemento não nulo ocupa uma posição n .

3. *Super-Esfera* $\Phi(r)$, com $r > 0$:

$$\Phi(r) = c \in \ell^1(\mathbb{R}) : \|c\| = r \quad (42)$$

onde $\Phi(r)$ é uma esfera de raio r no espaço conjunto canal+equalizador. O termo “super” simplesmente torna a definição aplicável a espaços com dimensão maior que três.

Tendo definido esses três conceitos, discutiremos a seguir os principais resultados do artigo, começando por definir a *curtose normalizada* como:

$$K_{norm}(b) = \frac{K(b)}{\sigma_b^4} \quad (43)$$

onde b é uma variável aleatória, σ_b^2 é a sua variância e $K(\cdot)$ é a curtose, definida em (10).

Li e Ding mostram então que, se a seguinte condição for satisfeita

$$\frac{K_{norm}(y_n)}{K_{norm}(s_n)} > 0.5 \quad (44)$$

onde y_n é o sinal de saída do equalizador e s_n é o sinal transmitido, e $c_{inicial}$ estiver contido em um cone de mínimo global único S_n^\pm , sob um passo de adaptação suficientemente pequeno, o vetor c convergirá para um ponto de mínimo dentro de $T \cap S_n^\pm$. Assim sendo, a curtose normalizada no instante inicial de adaptação tem grande influência no estado final do equalizador.¹

A demonstração, em linhas gerais, mostra que, caso a condição (44) seja satisfeita e $c_{inicial}$ pertença a um cone S_n^\pm , então o custo inicial será menor do que o menor custo dentre todos os valores da fronteira do cone. Assim, se for usado um método *steepest descent*, c não cruzará a fronteira, pois isso demandaria uma migração rumo a soluções com custo maior.

Essa nova abordagem lança ainda algumas luzes sobre o problema da localização dos mínimos. Denotando por ZF_n^+ uma solução do tipo $c = [\dots 0 \ 1 \ (\text{posição } n) \ 0 \dots]$ e ZF_n^- uma solução do tipo $c = [\dots 0 \ -1 \ (\text{posição } n) \ 0 \dots]$, se $ZF_n^\pm \in T \cap S_n^\pm$, então esta solução será o único ponto de mínimo de $T \cap S_n^\pm$, e não haverá pontos de mínimo em $T \cap B_n^\pm$, onde B_n^\pm é a fronteira de S_n^\pm . Caso haja possibilidade de perfeita inversão do canal, soluções ZF farão parte do conjunto de estados alcançáveis, e, portanto,

¹Regalia e Mboup em (Regalia and Mboup, 1999) mostram que existem casos de cones $T \cap S_n^\pm$ que não possuem pontos de convergência e que, se isto ocorrer, o algoritmo buscará soluções em outros cones.

$ZF_n^\pm \in T \cap S_n^\pm$ será satisfeita para algum valor de n , não existindo nenhum mínimo em $T \cap B_n^\pm$, ou seja, nos pontos de fronteira do conjunto de estados alcançáveis. Este resultado apenas indica que mínimos locais existentes não estarão associados ao mesmo cone que contém o mínimo global ou à sua fronteira, não havendo contradição com trabalhos como o de (Ding et al., 1991), que mostram que o algoritmo do módulo constante (CMA- *Constant Modulus Algorithm*) pode convergir para soluções ruins do ponto de vista de redução de interferência intersimbólica (IIS) se não forem adequadamente inicializados.

E se a inversão perfeita do canal não for possível? Li e Ding mostram que, se a solução ZF_n^\pm estiver próxima a, mas não contida em, $T \cap S_n^\pm$, então deve haver somente um mínimo em $T \cap S_n^\pm$, próximo de ZF_n^\pm , enquanto outros possíveis mínimos deverão estar próximos à fronteira do cone S_n^\pm .

Com relação a inicialização do algoritmo, tendo por base os conceitos acima, os autores propõem um processo baseado em três etapas:

1. A resposta conjunta inicial do sistema, $c_{inicial}$, deve estar contida em algum cone de mínimo local único S_n
2. A condição inicial deve levar a saída a satisfazer (44)
3. A solução ZF_n deve estar contida em (caso possa haver equalização perfeita) ou muito próxima a $T \cap S_n^\pm$.

Em termos um pouco mais práticos, temos que a inicialização *center-spike*, ou seja, aquela em que o equalizador possui um único coeficiente diferente de zero no "centro de massa" do filtro, garantirá obediência a 1) caso o valor máximo da resposta ao impulso do canal ocorra para um único valor de n , o que é usual. Não se pode garantir obediência a 2) *a priori*, pois o canal é desconhecido, e 3) pode ser satisfeita com a escolha de um equalizador de comprimento suficientemente grande. Os autores sugerem então a inicialização *center-spike* aliada a procedimentos de correção periódica do "centro de massa" dos coeficientes do equalizador e a eventuais alterações no comprimento do filtro, o que pode levar a uma situação próxima das condições ideais 1, 2 e 3 para uma ampla gama de canais.

Em (Li et al., 1996), os autores mostram que os pontos críticos estáveis indesejáveis para os quais os algoritmos cegos podem convergir são originados por dois fatores, pertencendo, portanto, a dois conjuntos distintos. O primeiro surge a partir do fato de utilizarmos filtros FIR como equalizadores de canais FIR, sendo assim mínimos dependentes do comprimento do equalizador, enquanto que o segundo decorre da própria função custo utilizada como critério, sendo, portanto, mínimos dependentes do custo.

Partimos, então, de duas condições genéricas relativas a um critério de equalização qualquer:

- C1 As soluções do tipo ZF são os únicos mínimos globais do critério em questão no espaço conjunto canal+equalizador, o que não significa que estas soluções pertençam sempre ao conjunto de estados alcançáveis.
- C2 A função custo do critério é contínua no espaço conjunto real canal+equalizador

Os autores mostram então que todo equalizador finito adaptado por um critério autodidata que satisfaça C1 e C2 apresenta mínimos locais. Essa afirmação é, certamente, muito forte, pois só não atinge os critérios que possuem restrições (Kennedy and Ding, 1992), já que eles normalmente não satisfazem C1, e equalizadores finitos com símbolos de entrada não-gaussianos. Observe que a condição C1 não é satisfeita para distribuição gaussiana. Como esses mínimos locais estão associados ao tamanho da resposta ao impulso do equalizador, são chamados *mínimos locais dependentes do comprimento*.

A outra classe de mínimos definida pelos autores está relacionada à própria estrutura da função custo considerada. São mínimos associados a técnicas que empregam não-linearidades de natureza descontínua como o algoritmo de decisão-direta, que possui a não-linearidade do decisor, ou o algoritmo *stop-and-go* (SGA) (Picchi and Prati, 1987), sendo que ainda podemos mencionar, em termos estruturais, o próprio *decision feedback equalizer* (DFE) (Proakis, 1995). Nesses casos, os mínimos locais ocorrem mesmo para o equalizador duplamente infinito, ou seja, não dependem do comprimento do equalizador, sendo denominados *mínimos locais dependentes do custo*. Isso mostra que é possível distinguir os diversos métodos existentes entre *aqueles que possuem só mínimos dependentes do comprimento (como os critérios CM e de Shalvi-Weinstein)* e *aqueles que possuem ambos os tipos de mínimos locais*.

O trabalho trata ainda da influência da adição de ruído branco e gaussiano no comportamento dos algoritmos CMA e *standard cumulant algorithm* (SCA) (Donoho, 1981), do qual o algoritmo de Shalvi-Weinstein é um caso particular. É demonstrado que, com a adição de ruído, as soluções ZF deixam de ser os mínimos globais das funções custo no espaço conjunto canal+equalizador, o que é razoável, pois a condição ZF está relacionada à IIS, e não ao ruído. Mostra-se, no entanto, que para altas relações sinal-ruído (SNR- *Signal to Noise Ratio*), o desvio da solução ótima em relação à solução ZF é muito pequeno.

Como foi dito brevemente no início desta seção, os resultados aqui mostrados são válidos tanto para o critério

do módulo constante (CM) e o CMA como para o critério de Shalvi-Weinstein (SW) e seus algoritmos, devido à relação existente entre os dois, que será tratada com mais detalhes na próxima seção. Nesta análise, incluiremos também o critério do módulo constante generalizado (GCM) (Cavalcanti et al., 1998a), que possui uma relação muito próxima com o critério de SW.

7 EQUIVALÊNCIA ENTRE OS CRITÉRIOS DE GODARD, SHALVI-WEINSTEIN E GCM

Analisaremos aqui a relação de equivalência entre os critérios de Godard, Shalvi-Weinstein e GCM sob dois pontos de vista diferentes: a partir de suas funções custo e a partir de seus vetores gradiente.

7.1 Análise baseada na função custo dos critérios

O primeiro trabalho que abordou a questão da comparação entre os critérios de Godard e Shalvi-Weinstein foi o de Li e Ding, já discutido na seção anterior (Li and Ding, 1995). Utilizando a mesma abordagem baseada em cones de mínimo global único, os autores mostram que existe uma correspondência biunívoca entre os pontos de mínimo dos dois critérios. Entretanto, *isso não significa que ambos sejam iguais ou que os respectivos algoritmos correspondentes o sejam.*

Para o caso do equalizador duplamente infinito, essa equivalência é facilmente mostrada, uma vez que as funções custo só possuem mínimos globais (Foschini, 1985; Shalvi and Weinstein, 1990), cada qual correspondente a uma solução ZF. Faltava estabelecê-la para o caso do equalizador finito, e foi justamente isso que Li e Ding fizeram. Supondo sinais reais, a demonstração utiliza os conceitos definidos na Seção 6, e consiste em mostrar que, se c_0 é um ponto de mínimo do critério SW, então ele estará direta e unicamente associado a um ponto de mínimo do critério CM. Em seguida, o mesmo é realizado no sentido inverso, ou seja, é mostrado que um mínimo do critério CM também se associa de forma única a um mínimo do critério SW.

O problema da equivalência voltou a ser abordado, em 1999, por Phillip Regalia (Regalia, 1999), de uma forma, aliás, bastante diferente da abordagem utilizada por Li e Ding. De maneira muito engenhosa, o autor escreve as funções custo CM e SW em coordenadas polares, adotando uma coordenada radial ρ , ângulos de rotação θ e vetores de norma unitária. No caso de sinais complexos, o vetor θ deve incluir as fases dos coeficientes do filtro.

Escrevendo então o vetor de pesos como $\mathbf{w} = \rho \bar{\mathbf{w}}$, onde $\bar{\mathbf{w}}$ é o vetor de coeficientes do equalizador com norma unitária, ou seja, $\|\bar{\mathbf{w}}\| = 1$, podemos fazer:

$$y(n) = \rho \bar{y}(n) \quad (45)$$

Substituindo (45) na função custo de Shalvi-Weinstein normalizada dada por (25) observa-se que a mesma é insensível à componente radial do vetor de coeficientes, dependendo, portanto, apenas dos ângulos θ . Fazendo o mesmo para o critério CM, dado por (11) com $p = 2$, encontra-se uma função que depende tanto do raio ρ como de θ . Como então comparar as duas funções, se dependem de variáveis diferentes? Através da *superfície de erro reduzida* de Godard, $J_{CMred}(\theta)$, que nada mais é do que a otimização da função dependente de ρ e θ com respeito à componente radial ρ , ou seja, $J_{CM}(\rho_{opt}, \theta)$.

Obtemos então duas funções dependentes somente das componentes angulares. O principal resultado formal do artigo é expresso por:

$$J_{CMred}(\theta) = R_2^2 \left(1 - \frac{1}{\lambda + J_{SW}(\theta)} \right) \quad (46)$$

onde λ é igual a 3 para $y(n)$ real e igual a 2 para $y(n)$ complexo, sendo que neste caso, $y(n)$ também deve ser simétrico e circular, satisfazendo $E[\bar{y}_n^2] = E[(\bar{y}_n^*)^2] = 0$.

Nas palavras de Regalia, temos que *a superfície de erro reduzida associada ao critério CM é uma deformação da superfície de erro associada ao critério SW*. O operador responsável pela deformação é de natureza monotônica e preserva ordem. Portanto os pontos críticos da função custo CM são também pontos críticos da função SW e mantêm a sua classificação (mínimo, máximo ou sela).

Aproveitando uma discussão feita pelo próprio autor, podemos afirmar que os métodos CM e SW têm, basicamente, em comum: susceptibilidade à existência de mínimos locais biunivocamente associados, mesma capacidade de reduzir IIS e mesma degradação média do desempenho na presença de ruído ou de correlação da fonte. Pode, no entanto, haver diferenças no desajuste (*misadjustment*) ou na velocidade de convergência de seus algoritmos, em decorrência do tipo de deformação que relaciona as funções. É interessante notar que a validade da análise realizada é muito ampla, sendo preciso pouco mais que a condição de estacionariedade até quarta ordem do sinal de entrada do equalizador, não importando se são independentes e identicamente distribuídos (i.i.d.) ou correlacionados, se o canal é ou não linear e se o ruído é aditivo ou multiplicativo.

Já com relação ao critério GCM (Cavalcanti et al., 1998a), partindo de (36), sabemos que sua função custo pode ser

escrita da seguinte forma (Neves, 2001):

$$J_{GCM} = \frac{E[|y(n)|^4]}{(E[|y(n)|^2])^2} - 1 \quad (47)$$

Comparando essa função com a função custo do critério de Shalvi-Weinstein, podemos concluir que $J_{GCM} = J_{SW} + \lambda - 1$. Assim, diferentemente do que ocorria na equivalência entre os critérios de Godard e Shalvi-Weinstein mostrada por Regalia (Regalia, 1999), agora não só os pontos críticos das duas funções são os mesmos: as superfícies de erro são realmente iguais a menos de um *offset*. A grande diferença entre os algoritmos resultantes dos dois critérios está na forma como foram obtidos: o algoritmo super-exponencial (ASE), resultante do critério de Shalvi-Weinstein, possui um passo de adaptação ótimo no sentido da velocidade de convergência, mas seu custo computacional é bastante elevado devido a necessidade de uma inversão de matriz; o GCMA (algoritmo do módulo constante generalizado), por sua vez, é obtido através de uma aproximação estocástica pouco usual para o vetor gradiente, e apresenta em geral uma velocidade de convergência relativamente melhor que o CMA com um custo computacional menor do que o ASE (Neves, 2001).

7.2 Análise baseada nos vetores gradiente de cada critério

Uma outra forma de comparação é, em vez de se partir das funções custo dos critérios, comparar seus vetores gradiente. Estudaremos o caso real (a extensão para o caso complexo é direta).

Calculando-se o vetor gradiente para a função custo de Shalvi-Weinstein, dada por (25) com $p = q$, e para o critério de Godard dado por (11) com $p = 2$, em relação a $\mathbf{w}(n)$, temos:

$$\frac{\partial J_{SW}}{\partial \mathbf{w}} = \frac{4}{(E[y_n^2])^2} (E[y_n^3 \mathbf{x}_n] - r^y E[y_n \mathbf{x}_n]) \quad (48)$$

$$\frac{\partial J_{God}}{\partial \mathbf{w}} = 4 (E[y_n^3 \mathbf{x}_n] - r E[y_n \mathbf{x}_n]) \quad (49)$$

sendo $r^y = \frac{E[y_n^4]}{E[y_n^2]^2}$ e $r = \frac{E[s_n^4]}{E[s_n^2]^2}$.

Como as funções custo do critério de SW e do critério GCM só diferem por um *offset*, o vetor gradiente deste último também será dado por (48). Comparando (48) e (49) vemos que são expressões muito parecidas, que possuem a mesma estrutura. Entretanto, r^y só será igual a r no caso de

equalização perfeita pois

$$r^y = \frac{E[s_n^4] \sum_l |c_l|^4 + 3E^2[s_n^2] \left[\left(\sum_l |c_l|^2 \right)^2 - \sum_l |c_l|^4 \right]}{E[s_n^2] \sum_l |c_l|^2}$$

e, só quando $\mathbf{c} = [0 \dots 0 \ 1 \ 0 \ \dots 0]^T$ (solução ZF), vale a expressão

$$r^y = r = \frac{E[s_n^4]}{E[s_n^2]^2} \quad (50)$$

onde também se considerou que $s(n)$ é uma seqüência i.i.d.

Portanto, caso (50) seja satisfeita, como o fator $1/E^2[y^2(n)]$ não interfere na obtenção dos pontos críticos, (48) e (49) possuirão os mesmos pontos estacionários. Note que a igualdade $r^y = r$ nada mais é do que a restrição de potência existente no critério de SW com restrição.

Mas no que resulta se ter $r \neq r^y$? Observando as duas superfícies de erro, mostradas nas Figuras 2 e 3, vemos que, enquanto o critério de Godard possui pontos de mínimo (local ou global), o de Shalvi-Weinstein (e portanto o GCM também) possui um vale de soluções que levam tanto ao mínimo global como ao mínimo local da função custo. Esses vales de soluções contêm os pontos estacionários do critério CM, e aparecem devido a ausência da restrição de potência (Neves, 2001).

É interessante notar aqui que o algoritmo super-exponencial é derivado do critério SW com restrição, e, portanto, ele convergirá para os mesmos pontos de mínimo que o CMA. Entretanto, o GCMA não possui essa restrição, convergindo para os vales de soluções mostrados na Figura 3.

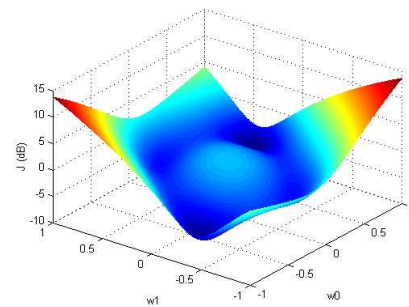


Figura 2: Superfície de erro do critério CM para o canal $h(z) = 1 - 0.6z^{-1}$

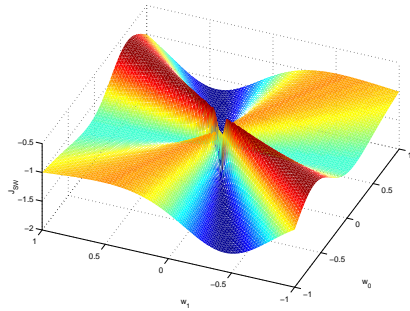


Figura 3: Superfície de erro do critério SW para o canal $h(z) = 1 - 0.6z^{-1}$

7.3 Considerações Sobre as Regiões de Atração dos Pontos Estacionários

O fato dos critérios de Godard e Shalvi-Weinstein possuírem os mesmos pontos críticos, mas resultarem em algoritmos diferentes, e que possivelmente realizam trajetórias diferentes antes da convergência, pode ficar mais claro com a análise do artigo (Gu and Tong, 2001), onde M. Gu e L. Tong fazem uma análise das regiões de atração dos pontos estacionários do critério de SW. A motivação maior do artigo é trazer novas informações sobre a questão da inicialização: como inicializar o algoritmo para que a convergência ocorra para o mínimo global? Isso está diretamente relacionado à região de atração do ponto crítico desejado. Em toda a análise, é suposto que a fonte é sub-gaussiana, os símbolos transmitidos e a resposta impulsiva do canal são reais, a matriz de convolução do canal possui posto completo e o ruído é branco e gaussiano.

Região de atração Ψ nada mais é do que a região que contém o ponto estacionário \mathbf{w}_{opt} , de forma que qualquer condição inicial \mathbf{w}_0 pertencente a Ψ levará o algoritmo de busca do gradiente, com passo de adaptação suficientemente pequeno, a gerar uma seqüência \mathbf{w}_k satisfazendo $J(\mathbf{w}_{k+1}) < J(\mathbf{w}_k) \forall k$, permanecer em Ψ e convergir para \mathbf{w}_{opt} . Caso a busca seja por pontos de máximo da função custo, basta inverter a relação $J(\mathbf{w}_{k+1}) > J(\mathbf{w}_k)$ e considerar o algoritmo no mesmo sentido do vetor gradiente.

Observe que (Gu and Tong, 2001) possui resultados interessantes. No contexto de receptores de Shalvi-Weinstein, é mostrado que, no caso de fontes i.i.d. com símbolos de mesma potência e canais ortonormais, Ψ não se altera com a redução ou aumento da relação sinal-ruído, ou seja, a dificuldade para se inicializar o algoritmo na região de atração da solução desejada será a mesma para uma situação pouco ou muito ruidosa. Já para canais ortogonais e fontes não identicamente distribuídas, o tamanho de Ψ depende da constelação, da potência do sinal transmitido, do nível de

ruído e do canal. Considerando a variância do ruído fixa, o tamanho de Ψ irá diminuir conforme se diminui a potência do sinal de entrada ou o valor absoluto do seu cumulante.

Entretanto, o resultado que mais nos interessa na análise deste trabalho é a comparação entre as regiões de atração dos critérios de Shalvi-Weinstein e de Godard. Os autores mostram, para o caso de fonte i.i.d. e sendo possível inverter o canal, que, não havendo ruído, as regiões de atração para os dois receptores são iguais. O mesmo continua valendo caso exista ruído e o canal seja ortonormal. Já se o canal for ortogonal, as regiões de atração se tornam diferentes, o que faz com que os algoritmos possuam transientes diferentes em seus processos de convergência. Note, entretanto, que este é um resultado bastante restrito devido às hipóteses consideradas.

8 RELAÇÃO DOS CRITÉRIOS AUTODIDATAS COM WIENER E ZF

No trabalho inicial de Godard (Godard, 1980), já havia sido comentado que as soluções encontradas com o critério CM estariam relativamente próximas às soluções de Wiener. Apesar de ser uma observação sem grande fundamentação teórica, ela serviu como forte argumento para a aceitação da técnica proposta, uma vez que seria possível obter soluções próximas à de Wiener mesmo sem a utilização de uma seqüência de treinamento. Assim, evidenciar de forma teórica uma possível relação entre os mínimos dos dois critérios seria extremamente útil para a validação da técnica não-supervisionada.

8.1 Resultados existentes na literatura

Um artigo que tratou mais detalhadamente a relação entre equalizadores obtidos com os dois critérios foi escrito por H. Zeng, L. Tong e C. Johnson, em 1998 (Zeng et al., 1998). Os autores consideram sinais binários, i.i.d. e canal AWGN (*Additive White Gaussian Noise*). O trabalho enfocou basicamente três questões:

1. O critério CM possui mínimos locais? Se sim, onde estão?
2. Qual o desempenho, em termos de erro quadrático médio, dos equalizadores obtidos através do critério CM?
3. Qual a relação entre os equalizadores obtidos por Wiener e os obtidos pelo critério CM?

A primeira pergunta se aplica num contexto de equalizadores fracionários, não contemplados neste trabalho. Detalhes

sobre equalizadores fracionários podem ser encontrados em (Treichler et al., 1996). O caso do equalizador cujos coeficientes estão espaçados de um período de símbolo (*baud rate*), como foi o caso até aqui, já havia sido respondida de maneira afirmativa em diversos trabalhos. As demais questões estão diretamente relacionadas a uma necessidade de validação do critério CM em si, conforme comentamos. Como os diferentes receptores são ótimos segundo os critérios que lhes deram origem, é necessário estabelecer uma base comum para compará-los, o que pode ser feito em termos do valor erro quadrático médio dos receptores.

Utilizando uma abordagem geométrica, baseada no chamado teorema de Weierstrass, os autores procuram encontrar uma região compacta B com uma fronteira ∂B e uma referência interna w_{ref} , de tal forma que o custo (critério CM) na fronteira seja maior do que no ponto de referência. Assim, tem-se pelo menos um ponto de mínimo em B .

Desta forma é mostrado que o critério CM apresenta mínimos locais para o caso em que existe ruído, sempre considerando regiões próximas das soluções de Wiener. Se o mínimo erro quadrático médio for menor do que um certo valor, garante-se a existência de uma solução do critério CM na vizinhança de uma solução de Wiener. É interessante notar que a idéia de erro quadrático médio está associada exclusivamente a Wiener, e, no entanto, aqui ela é usada para analisar um equalizador cego. Estes são chamados de mínimos locais "bons", pois resultam em um baixo erro quadrático médio (EQM). Um limitante superior para o EQM destes últimos também é encontrado, o qual depende das características da solução de Wiener associada.

Quanto à questão que realmente mais nos interessa, foi concluído que os equalizadores obtidos pelo critério CM e por Wiener são aproximadamente colineares, ou seja, idênticos a menos de uma constante multiplicativa. É calculado também o valor do EQM em excesso do equalizador CM em relação ao de Wiener.

Um segundo trabalho, escrito pelos mesmos autores (Zeng et al., 1999), pode ser visto como uma extensão do trabalho anterior ao caso de sinais reais e não mais binários. A análise realizada segue a mesma idéia do artigo anterior, e os resultados obtidos são bastante similares aos já comentados, a saber: estabelecimento de um limitante superior, em termos de erro quadrático médio (EQM), para o desempenho de equalizadores obtidos através do critério CM; considerações sobre o erro em excesso deste em relação ao de Wiener; constatação de uma tendência a colinearidade entre os receptores obtidos pelos dois critérios.

Pode-se concluir que equalizadores obtidos pelo critério CM podem ter, para altas SNRs, desempenho semelhante ao de equalizadores de Wiener (supervisionados) em termos de

taxa de erro de bits. Isso ajuda a dissipar a visão de que o CM é um critério "ruim" por apresentar mínimos locais.

Em (Gu and Tong, 1999), a análise é estendida para equalizadores obtidos pelo critério de SW. Pela equivalência já mostrada nas seções anteriores entre este e o critério CM, muitos dos resultados já discutidos continuam sendo válidos. No artigo, os autores consideram uma situação bastante restritiva, engendrada por canais cuja matriz de convolução H possui colunas ortogonais e respostas combinadas canal+equalizador bidimensionais. Além disso, são considerados sinais i.i.d. e ruído AWGN.

Os autores concluem, assim, que os equalizadores obtidos pelos critérios CM, SW, Wiener e ZF são colineares e unicamente correspondentes se e somente se não existir ruído ou o canal for ortogonal. Além disso, para respostas combinadas bidimensionais, é mostrado que para todo equalizador CM/SW existe uma solução de Wiener próxima, ou seja, não existem soluções espúrias (sem associação com alguma solução de Wiener). Dessa forma, os três critérios são bastante semelhantes e levam a desempenhos próximos. Outro resultado bastante interessante é que a diminuição da SNR aproxima a relação sinal-interferência (SIR) do receptor CM/SW da apresentada pelo receptor de Wiener, tornando mais próximos os desempenhos dos dois receptores.

Também analisando os receptores SW, Regalia e Mboup, em (Regalia and Mboup, 1999), mostram que os máximos da função custo do critério de Shalvi-Weinstein normalizado (25) estarão tanto mais próximos das correspondentes soluções de Wiener quanto maior for o valor de p empregado, sendo iguais para o caso limite em que p tende a infinito. Isto indica que o uso de estatísticas de ordem cada vez mais elevada leva a pontos críticos cada vez mais próximos dos receptores de Wiener. Entretanto, um dos motivos pelo qual não são usadas estatísticas de ordem muito elevada é que a otimização dos critérios passa a envolver estimadores com variância muito grande, o que pode ser muito problemático do ponto de vista prático.

8.2 Uma Outra Abordagem

Existe também uma forma simples de estabelecer uma ponte entre os receptores de Wiener e CM, exposta em (Suyama et al., 2003), e baseada no estabelecimento de limitantes para a superfície de erro do critério CM.

Considere o caso em que o sinal transmitido é i.i.d. e assume valores $+1$ e -1 com igual probabilidade (sinal BPSK-*Binary Phase Shift Keying*). Nessa condição a função custo CM (11) com $p = 2$ assume a seguinte forma:

$$J_{CM} = E [y(n)^2 - 1]^2 \quad (51)$$

pois a constante R_2 é igual a 1. A equação (51) pode ainda ser reescrita como:

$$\begin{aligned} J_{CM} &= E [y(n)^2 - 1]^2 \\ &= E \left\{ [y(n) - s(n - n_d)]^2 [y(n) + s(n - n_d)]^2 \right\} \end{aligned} \quad (52)$$

onde n_d é o atraso de equalização.

Utilizando a desigualdade de Cauchy-Schwarz (Papoulis, 1991):

$$E [ab]^2 \leq E [a^2] \cdot E [b^2] \quad (53)$$

onde a e b são variáveis aleatórias quaisquer, e sabendo que $J_{CM} \geq 0$, chegamos a:

$$J_{CM} \leq \sqrt{E \{ [y(n) + s(n - n_d)]^4 \} E \{ [y(n) - s(n - n_d)]^4 \}} \quad (54)$$

A equação (54) estabelece um limitante superior para a função custo CM a partir de uma espécie de “função de Wiener de quarta ordem” (Walach and Widrow, 1984). Trata-se, portanto, de um elo importante entre uma técnica autodidata e uma supervisionada.

Embora não se trate de uma relação direta entre o critério CM e o de Wiener, ela evidencia o papel do atraso de equalização, fator chave do critério de Wiener. Voltaremos a este ponto mais adiante.

Vejamus um exemplo. Para um canal simples, definido por:

$$h_1(z) = 1 + 0,4z^{-1} \quad (55)$$

traçamos as curvas limitantes correspondentes aos atrasos $n_d = 0, 1$ e 2 .

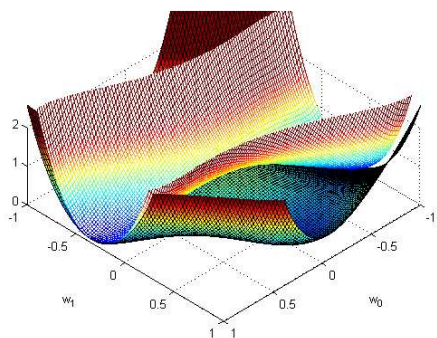


Figura 4: Superfície limitante para o canal h_1 , $n_d = 0$

Na Figura 4, temos a superfície da função custo CM limitada por uma outra, correspondente ao limitante apresentado em (54) para um atraso $d = 0$. Observa-se que o limitante

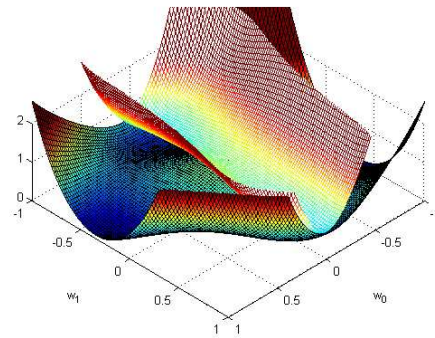


Figura 5: Superfície limitante para o canal h_1 , $n_d = 1$

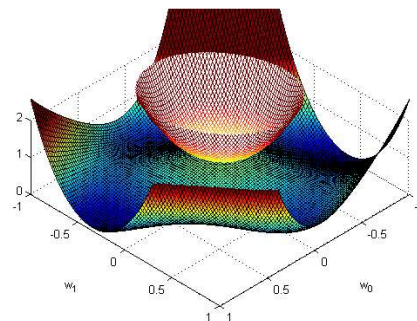


Figura 6: Superfície limitante para o canal h_1 , $n_d = 2$

superior se concentra sobre um par específico de mínimos do critério CM. O mesmo ocorre para $d = 1$ (Figura 5), para um outro par de mínimos. Estes mínimos estão próximos das soluções de Wiener para $d = 0$ e $d = 1$.

Um fato interessante ocorre para $d = 2$ (Figura 6). Pelo gráfico, percebe-se que o limitante correspondente se situa sobre o máximo da função custo CM. A despeito do que possa parecer, o limitante em questão não é um parabolóide, mas sim uma superfície que apresenta um par de mínimos muito próximos a zero.

Esta visão de limitantes pode nos ajudar a compreender por que na maioria dos casos, aparentemente, apenas algumas das possíveis soluções de Wiener (em função do atraso de equalização) estão representadas na função custo CM. Considerando nosso exemplo, podemos traçar um limitante geral composto pelas superfícies obtidas para $d = 0, 1$ e 2 , obtendo a superfície mostrada na Figura 7. Observa-se, no entanto, que não há indícios da curva correspondente ao atraso $d = 2$. Isso ocorre porque as superfícies correspondentes a $d = 0$ e $d = 1$ (bons atrasos de equalização) representam um limite mais restritivo para

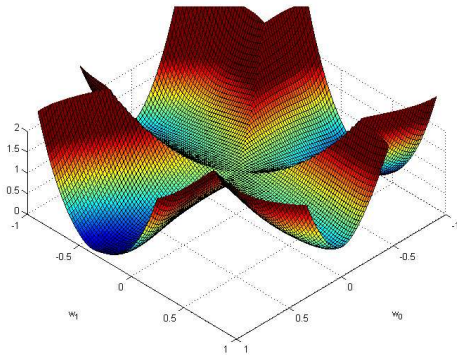


Figura 7: Limitante formado pelas superfícies correspondentes a $n_d = 0, 1$ e 2

J_{CM} . Não surpreende, portanto, que o mínimo relacionado ao atraso de equalização igual a 2 não esteja presente no critério CM.

A relação (54) ainda mostra claramente que, havendo a possibilidade de inversão perfeita do canal, ou seja, $y(n) - s(n - n_d)$ ou $y(n) + s(n - n_d)$ (lembrando que a inversão de sinal na saída do equalizador não representa um problema), a função custo CM se anulará, o que indica que a solução ZF será um mínimo do critério CM. Em outras palavras, numa situação ideal, a relação (54) torna-se uma igualdade para $\mathbf{w} = \mathbf{w}_{ZF}$ (equalizador zero forcing), o que mostra que as superfícies do limitante e do custo CM se tangenciam exatamente nos pontos de mínimo.

Para os demais casos, existirá sempre uma separação entre as superfícies, porém esta distância será menor em regiões formadas por soluções que levem a boas condições de equalização, e maior em regiões correspondentes a soluções ruins.

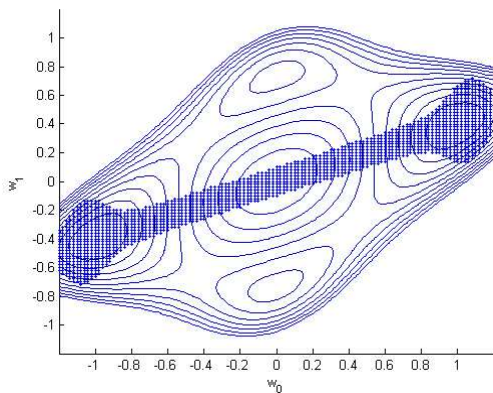


Figura 8: Canal AR com 1 pólo em $-0, 4$

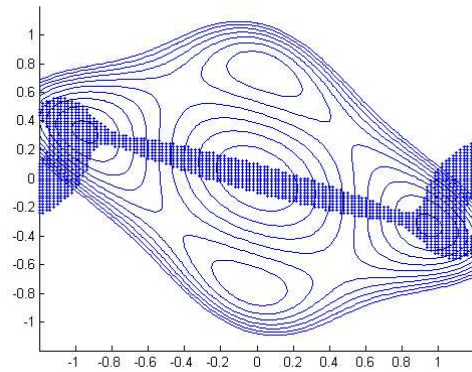


Figura 9: Canal MA com 1 zero em $-0, 4$

Na Figura 8, apresentamos as curvas de nível do custo CM para um canal AR (auto-regressivo), caso em que a equalização perfeita é possível, e destacamos as regiões nas quais a distância entre a superfície e seu limitante é inferior a 0,1. De maneira similar, na Figura 9, mostramos as curvas referentes ao canal h_1 , dado por (55). Nota-se que, nas regiões próximas aos mínimos globais, a distância entre as superfícies é pequena, mesmo não havendo possibilidade de inversão perfeita, como no primeiro caso.

Tendo em vista que nosso maior interesse é a localização dos mínimos dos critérios, é importante verificar a qualidade do limitante obtido nas regiões em torno dos mesmos. Considere então que estejamos trabalhando em uma condição próxima a ZF. Nessa situação, conforme vimos através dos exemplos, as superfícies estarão próximas entre si, e podemos descrever a região nas proximidades do mínimo CM como:

$$J_{CM} \cong \sqrt{E \{ [y(n) + s(n - n_d)]^4 \} E \{ [y(n) - s(n - n_d)]^4 \}} \quad (56)$$

Podemos observar que a função é dividida em duas partes: uma delas está associada a um certo mínimo \mathbf{w}_1 , enquanto que a outra está ligada ao seu simétrico, $-\mathbf{w}_1$. Dessa forma, o primeiro termo do lado direito da equação (56) será preponderante na determinação do valor de J_{CM} quando estivermos considerando regiões próximas a \mathbf{w}_1 , enquanto o segundo termo terá maior influência no valor da função custo quando estivermos em regiões em torno de $-\mathbf{w}_1$.

Levando em conta que nas regiões consideradas o valor da função J_{CM} é pequeno, podemos considerar uma aproximação do tipo :

$$E \{ [y(n) \pm s(n - n_d)]^4 \} \cong E \{ [y(n) \pm s(n - n_d)]^2 \}^2 \quad (57)$$

Note que a igualdade ocorre para ocasiões em que é possível atingir a condição ZF. Aplicando este resultado à equação (56) obtemos o seguinte :

$$J_{CM} \cong E \left\{ [y(n) + s(n - n_d)]^2 \right\} E \left\{ [y(n) - s(n - n_d)]^2 \right\} \quad (58)$$

ou seja

$$J_{CM}(\mathbf{w}) \cong J_{Wiener}(\mathbf{w}, d) \cdot J_{Wiener}(-\mathbf{w}, d) \quad (59)$$

Temos agora uma relação direta entre a função custo CM e a função custo de Wiener, a qual deve ser válida em condições de boa equalização, ou seja, próximas à condição ZF. A partir de (59) é possível, então, derivar uma expressão relacionando a localização dos mínimos dos dois critérios.

Uma condição para que um ponto seja um mínimo é que o gradiente desta função calculado no ponto em questão seja zero. Sabendo que J_{Wiener} pode ser escrito como (Haykin, 1996):

$$J_{Wiener}(\mathbf{w}, d) = \mathbf{w}^T \mathbf{R}_x \mathbf{w} - 2\mathbf{w}^T \mathbf{p}_d + 1 \quad (60)$$

onde \mathbf{R}_x é a matriz de autocorrelação do sinal de saída do canal, x_n , e \mathbf{p}_d é o vetor de correlação cruzada entre os sinais x_n e s_{n-d} . Substituindo (60) em (59) e aplicando o operador gradiente, obtemos, após alguma manipulação:

$$\frac{\partial J_{CM}(\mathbf{w})}{\partial \mathbf{w}^T} = \mathbf{R}_x \mathbf{w} (\mathbf{w}^T \mathbf{R}_x \mathbf{w}) + \mathbf{R}_x \mathbf{w} - 2\mathbf{p}_d \mathbf{w}^T \mathbf{p}_d \frac{1}{2} \quad (61)$$

Supondo que \mathbf{w}_{CM} seja um mínimo do critério CM, e multiplicando por \mathbf{R}_x^{-1} à esquerda, chegamos à seguinte condição:

$$\mathbf{w}_{CM} (\mathbf{w}_{CM}^T \mathbf{R}_x \mathbf{w}_{CM}) + \mathbf{w}_{CM} - 2\mathbf{w}_{Wiener,d} \mathbf{w}_{CM}^T \mathbf{p}_d = \mathbf{0} \quad (62)$$

Isolando $\mathbf{w}_{Wiener,d}$, obtemos uma expressão interessante:

$$\mathbf{w}_{Wiener,d} \cong \frac{\mathbf{w}_{CM} (\mathbf{w}_{CM}^T \mathbf{R}_x \mathbf{w}_{CM} + 1)}{2\mathbf{w}_{CM}^T \mathbf{p}_d} \quad (63)$$

Tal relação mapeia de maneira explícita uma solução CM em um mínimo de Wiener. Um fato interessante é que a aproximação exposta em (63) sugere a colinearidade entre soluções, em harmonia com idéias já discutidas na Seção 8.1 (Zeng et al., 1998; Schniter and Johson, 2000).

Utilizando o mesmo canal definido em (55) como exemplo, obtivemos os seus receptores CM de 2 coeficientes através de uma estimação do gradiente da função custo. Com estes resultados, aplicamos a relação obtida em (63) e formamos a Tabela 1, na qual podemos observar a qualidade da estimação

de \mathbf{w}_{Wiener} . O mesmo foi feito para um segundo canal dado por:

$$h_2(z) = 1 + 0,6z^{-1} + 0,4z^{-2} \quad (64)$$

operando com um equalizador de 3 coeficientes. Os resultados relativos a este canal são mostrados na Tabela 2.

Com base nos resultados obtidos, podemos dizer que as aproximações consideradas são bastante razoáveis, pois as estimativas dos receptores de Wiener estão bem próximas dos valores exatos. Isso confirma a validade da abordagem apresentada e garante a solidez da relação estabelecida entre os critérios.

coef.	\mathbf{w}_{CM}	\mathbf{w}_{Wiener}	\mathbf{w}_{Wiener} dado \mathbf{w}_{CM}	EQM
\mathbf{w}_0	0,9497	0,9784	0,9609	0,14694
\mathbf{w}_1	-0,3282	-0,3374	-0,3321	
\mathbf{w}_0	0,0399	0,0540	0,0437	0,36736
\mathbf{w}_1	0,7515	0,8435	0,8230	

Tabela 1: Aproximação do receptor de Wiener dado \mathbf{w}_{CM} para h_1 - Equalizador com 2 coeficientes

coef.	\mathbf{w}_{CM}	\mathbf{w}_{Wiener}	\mathbf{w}_{Wiener} dado \mathbf{w}_{CM}	EQM
\mathbf{w}_0	0,8945	0,9507	0,9208	0,22209
\mathbf{w}_1	-0,5256	-0,5573	-0,5410	
\mathbf{w}_2	0,0532	0,0578	0,0547	
\mathbf{w}_0	0,0111	0,0131	0,0115	0,22948
\mathbf{w}_1	0,8821	0,9395	0,9100	
\mathbf{w}_2	-0,4932	-0,5226	-0,5088	

Tabela 2: Aproximação do receptor de Wiener dado \mathbf{w}_{CM} para h_2 - Equalizador com 3 coeficientes

9 CONCLUSÃO

Primeiramente, apresentamos os critérios de Godard e de Shalvi-Weinstein, bem como os principais resultados decorrentes da abordagem analítica primordial, à qual associamos o epíteto clássica. Em síntese, a análise de ambos os critérios seguiu cursos semelhantes, partindo de uma análise fundada na hipótese pouco prática de comprimento infinito do equalizador até a constatação da existência de mínimos locais. É interessante notar que essa análise, em alguns momentos, precisou tocar na equivalência entre soluções obtidas através dos critérios. Por fim, apresentamos o critério do módulo constante generalizado, que pode ser entendido como uma extensão da proposta de Godard.

Em seguida, detivemo-nos na análise geométrica dos critérios e algoritmos autodidatas. Mostramos no que consiste esta análise e quais os resultados importantes que

foram obtidos na literatura a partir dela, principalmente no que se refere à definição e localização de pontos estacionários de critérios e de sua contribuição no estabelecimento de equivalências entre eles.

Para enriquecer a discussão, mostramos alguns resultados bastante recentes. Incluímos o critério GCM na análise da equivalência entre os critérios autodidatas, mostrando a sua estreita relação com o critério de Shalvi-Weinstein. Vimos também como é possível observar a relação entre os critérios CM, SW e GCM se partirmos de seus vetores gradiente, em vez de trabalharmos com as funções custo.

Em seguida, discutimos a relação existente entre os critérios autodidatas e o critério de Wiener. Primeiramente, vimos alguns resultados mais conhecidos da literatura, e, em seguida, expusemos uma de nossas contribuições recentes no sentido de relacionar os critérios de Wiener e CM. Foi determinado um limitante superior para a função custo do critério CM, o que permitiu a obtenção de uma aproximação para J_{CM} dependente do atraso de equalização.

É nossa opinião que a análise da relação entre critérios ainda é um tema que merece ser explorado e estabelecido em bases ainda mais sólidas, principalmente no que se refere às associações entre técnicas cegas e supervisionadas. Em particular, parece-nos de significativa importância que haja mais esforços no sentido de analisar o grau de conexão existente entre os mínimos do critério CM e as diversas soluções de Wiener. Estudos dessa natureza serviriam para aproximar os "mundos distintos" de critérios não-supervisionados como o de decisão-direta e de Godard, e seriam ainda úteis para que fosse possível estabelecer uma comparação mais sólida entre as múltiplas técnicas de que dispõe no presente o engenheiro de sistemas.

AGRADECIMENTOS

Os autores agradecem ao CNPq e à Fapesp o apoio financeiro.

REFERÊNCIAS

Benveniste, A., Goursat, M. and Ruget, G. (1980). Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communications, *IEEE Trans. on Automatic Control* **AC-25**(3): 385–399.

Cavalcanti, F., Brandão, A. and Romano, J. (1998a). A generalized constant modulus algorithm for blind equalization, *Globecom, Sidney, Austrália*.

Cavalcanti, F. R. (1999). *Antenas Inteligentes e Processamento Espaço-Temporal para Sistemas de*

Comunicação sem fio, Tese de doutorado, UNICAMP.

Cavalcanti, F. R., Brandão, A. and Romano, J. M. (1998b). A generalized constant modulus algorithm for blind equalization, *Proc. of Globecom, Sidney, Australia*.

Ding, Z., Jonhson, C. and Kennedy, R. (1992). On the (non)existence of undesirable equilibria of Godard blind equalizers, *IEEE Trans. on Signal Processing* **40**(10): 2425–2432.

Ding, Z., Kennedy, R., Anderson, B. and Jonhson, C. (1991). III- convergence of Godard blind equalizers in data communication systems, *IEEE Trans. on Communications* **39**(9): 1313–1327.

Donoho, D. (1981). On minimum entropy deconvolution, in *D. F. Findley (ed.), Applied Time Series Analysis II, Academic Press*.

Foschini, G. (1985). Equalization without altering or detecting data, *Bell Syst. Tech. J.* **64**(8): 1885–1911.

Godard, D. (1980). Self-recovering equalization and carrier tracking in two-dimensional data communication systems, *IEEE Trans. on Communications* **28**(11): 1867–1875.

Gu, M. and Tong, L. (1999). Geometrical characterizations of constant modulus receivers, *IEEE Trans. on Signal Processing* **47**(10): 2745–2756.

Gu, M. and Tong, L. (2001). Domains of attraction of Shalvi-Weinstein receivers, *IEEE Trans. on signal Processing* **49**(7): 1397–1408.

Haykin, S. (1996). *Adaptive Filter Theory*, 3 edn, Prentice Hall.

Johnson, C. and Anderson, B. (1995). Godard blind equalizer error surface characteristics: White, zero-mean, binary source case, *International Journal of Adaptive Control and Signal Processing* (9): 301–324.

Johnson, C., Schniter, P., Endres, T., Behm, J., Brown, D. and Casas, R. (1998). Blind equalization using the constant modulus criterion: A review, *Proc. of the IEEE* **86**(10): 1927–1950.

Kennedy, R. and Ding, Z. (1992). Blind adaptive equalizers for quadrature amplitude modulated communication systems based on convex cost functions, *Optical Engeneering* **31**(6): 1189–1199.

Li, Y. and Ding, Z. (1995). Convergence analysis of finite length blind adaptive equalizers, *IEEE Trans. on Signal Processing* **43**(9): 2121–2129.

- Li, Y. and Liu, K. R. (1996). Static and dynamic convergence behavior of adaptive blind equalizers, *IEEE Trans. on Signal Processing* **44**(11): 2726–2735.
- Li, Y., Liu, K. R. and Ding, Z. (1996). Length- and cost-dependent local minima of unconstrained blind channel equalizers, *IEEE Trans. on Signal Processing* **44**(11): 2726–2735.
- Lucky, R., Salz, J. and Weldon, E. (1968). *Principles of Data Communication*, MacGraw-Hill, Nova York.
- Mboup, M. and Regalia, P. (1999). On the equivalence between the super-exponential algorithm and a gradient search method, *Proc. of IEEE Int. Conference on Acoustic, Speech and Signal Processing*, Vol. 5, Phoenix, pp. 2643–2646.
- Mboup, M. and Regalia, P. (2000). A gradient search interpretation of the super-exponential algorithm, *IEEE Trans. on Information Theory* **46**(7): 2731–2734.
- Neves, A. (2001). *Uma abordagem unificada para algoritmos de equalização autodidata*, Tese de mestrado, Unicamp.
- Neves, A., Miranda, M. and Romano, J. (2002). New issues on the criterion properties and algorithm convergence of the generalized CM approach, *Proc. of the International Telecommunication Symposium, ITS*, Natal, Brasil.
- Nikias, C. and Petropulu, A. (1993). *Higher Order Spectra Analysis - A Nonlinear Signal Processing Framework*, 1 edn.
- Papoulis, A. (1991). *Probability, Random Variables, and Stochastic Process*, 3 edn.
- Picchi, G. and Prati, G. (1987). Blind equalization and carrier recovery using a "stop-and-go" decision-directed algorithm, *IEEE Trans. on Communications* **COM-35**: 877–887.
- Proakis, J. (1995). *Digital Communications*, 3 edn.
- Quresh, S. (1985). Adaptive equalization, *Proceedings of the IEEE* **73**(9): 1349–1387.
- Regalia, P. (1999). On the equivalence between the Godard and Shalvi Weinstein schemes of blind equalization, *Signal Processing* **73**(1-2): 185–190.
- Regalia, P. and Mboup, M. (1999). Undermodeled equalization: A characterization of stationary points for a family of blind criteria, *IEEE Trans. on Signal Processing* **47**(3): 760–770.
- Sato, Y. (1975). A method for self recovering equalization, *IEEE Trans. on Communications* **COM-23**(6): 679–682.
- Schniter, P. and Johson, C. (2000). Bounds for the MSE performance of constant modulus estimators, *IEEE Trans. on Information Theory* **46**(7): 2544–2560.
- Shalvi, O. and Weinstein, E. (1990). New criteria for blind deconvolution of nonminimum phase systems (channels), *IEEE Trans. on Information Theory* **36**(2): 312–321.
- Shalvi, O. and Weinstein, E. (1993). Super-exponential methods for blind deconvolution, *IEEE Trans. on Information Theory* **39**(2): 504–519.
- Shalvi, O. and Weinstein, E. (1994). *Blind Deconvolution*, Prentice Hall, chapter Universal Methods for Blind Deconvolution.
- Suyama, R., Attux, R. R. F., Romano, J. M. T. and Bellanger, M. (2003). Relations entre les critères du module constant et de wiener, *19e Colloque GRETSI - Paris - França*.
- Treichler, J. and Agee, B. (1983). New approach to multipath correction of constant modulus signals, *IEEE Trans. on Acoust., Speech and Signal Processing* **ASSP-31**(2): 459–472.
- Treichler, J., Fijalkow, I. and Johnson, C. (1996). Fractionally spaced equalizers, *IEEE Signal Processing Magazine* pp. 65–81.
- Tugnait, J. (1992). Comments on 'New approach for blind deconvolution of nonminimum phase systems (channels)', *IEEE Trans. on Information Theory* **28**: 210–213.
- Walach, E. and Widrow, B. (1984). The least mean fourth (LMF) adaptive algorithm and its family, *IEEE Transactions on Information Theory* **IT-30**(2): 275–283.
- Wiggins, R. (1978). Minimum entropy deconvolution, *Geoexplorations* **16**: 21–35.
- Zeng, H., Tong, L. and Johnson, C. (1998). Relationships between the constant modulus and wiener receivers, *IEEE Trans. on Information Theory* **44**(4): 1523–1538.
- Zeng, H., Tong, L. and Johnson, C. (1999). An analysis of constant modulus receivers, *IEEE Trans. on Signal Processing* **47**(11): 2990–2999.