

FLÁVIO RENÊ MIRANDA PAVAN

Sobre a desconvolução multiusuário e a separação de fontes

São Paulo
2016

FLÁVIO RENÊ MIRANDA PAVAN

Sobre a desconvolução multiusuário e a separação de fontes

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Mestre em Ciências

São Paulo
2016

FLÁVIO RENÊ MIRANDA PAVAN

Sobre a desconvolução multiusuário e a separação de fontes

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Mestre em Ciências

Área de Concentração:
Sistemas Eletrônicos

Orientadora:
Prof.^a Dr.^a Maria D. Miranda

São Paulo
2016

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, _____ de _____ de _____

Assinatura do autor: _____

Assinatura do orientador: _____

Catálogo-na-publicação

Pavan, Flávio Renê Miranda

Sobre a desconvolução multiusuário e a separação de fontes / F. R. M.
Pavan -- versão corr. -- São Paulo, 2016.

117 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Telecomunicações e Controle.

1.Processamento digital de sinais 2.Engenharia elétrica I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Telecomunicações e Controle II.t.

AGRADECIMENTOS

Algumas pessoas tiveram um papel fundamental para que este trabalho se concretizasse.

Gostaria de agradecer à Prof.^a Maria Miranda, minha orientadora, pela dedicação e esforço incansáveis na realização deste trabalho, pela constante preocupação em me manter motivado mesmo nos dias mais difíceis, pelo profissionalismo inspirador e pelo grande exemplo de excelência acadêmica que tem me cativado desde as aulas de Processamento Digital de Sinais durante a graduação, até os dias de hoje.

Agradeço ao Prof. Magno Silva pelas suas contribuições na parte de análise de algoritmos, pela amizade e pela grande disponibilidade e atenção com que sempre me atendeu.

Agradeço aos Professores Luiz Monteiro, Romis Attux e Aline Panazio, pela revisão cuidadosa deste trabalho e pelas críticas e sugestões valiosas. Sem as suas ideias, não teria sido possível concluir o trabalho desta forma.

Um agradecimento especial aos meus pais, Eloiza e Flávio, pela grande ajuda ao longo dos anos e pela preocupação quanto a minha formação acadêmica e a minha realização profissional.

Agradeço também à querida amiga Estela Basso, pela amizade de mais de uma década, pelas boas risadas e por sua atenção e preocupação, mesmo que à distância, durante os momentos difíceis.

RESUMO

Os problemas de separação cega de fontes e desconvolução cega multiusuário vêm sendo intensamente estudados nas últimas décadas, principalmente devido às inúmeras possibilidades de aplicações práticas. A desconvolução multiusuário pode ser compreendida como um problema particular de separação de fontes em que o sistema misturador é convolutivo, e as estatísticas das fontes, que possuem alfabeto finito, são bem conhecidas. Dentre os desafios atuais nessa área, cabe destacar que a obtenção de soluções adaptativas para o problema de separação cega de fontes com misturas convolutivas não é trivial, pois envolve ferramentas matemáticas avançadas e uma compreensão aprofundada das técnicas estatísticas a serem utilizadas. No caso em que não se conhece o tipo de mistura ou as estatísticas das fontes, o problema é ainda mais desafiador. Na área de Processamento Estatístico de Sinais, soluções vêm sendo propostas para resolver casos específicos. A obtenção de algoritmos adaptativos eficientes e numericamente robustos para realizar separação cega de fontes, tanto envolvendo misturas instantâneas quanto convolutivas, ainda é um desafio. Por sua vez, a desconvolução cega de canais de comunicação vem sendo estudada desde os anos 1960 e 1970. A partir de então, várias soluções adaptativas eficientes foram propostas nessa área. O bom entendimento dessas soluções pode sugerir um caminho para a compreensão aprofundada das soluções existentes para o problema mais amplo de separação cega de fontes e para a obtenção de algoritmos eficientes nesse contexto.

Sendo assim, neste trabalho (i) revisitam-se a formulação dos problemas de separação cega de fontes e desconvolução cega multiusuário, bem como as relações existentes entre esses problemas, (ii) abordam-se as soluções existentes para a desconvolução cega multiusuário, verificando-se suas limitações e propondo-se modificações, resultando na obtenção de algoritmos com boa capacidade de separação e robustez numérica, e (iii) relacionam-se os critérios de desconvolução cega multiusuário baseados em curtose com os critérios de separação cega de fontes.

Palavras-chave: Desconvolução cega multiusuário. Separação cega de fontes. Algoritmos adaptativos. Processamento Estatístico de Sinais.

ABSTRACT

Blind source separation and blind deconvolution of multiuser systems have been intensively studied over the last decades, mainly due to the countless possibilities of practical applications. Blind deconvolution in the multiuser case can be understood as a particular case of blind source separation in which the mixing system is convolutive, and the sources, which exhibit a finite alphabet, have well known statistics. Among the current challenges in this area, it is worth noting that obtaining adaptive solutions for the blind source separation problem with convolutive mixtures is not trivial, as it requires advanced mathematical tools and a thorough comprehension of the statistical techniques to be used. When the kind of mixture or source statistics are unknown, the problem is even more challenging. In the field of statistical signal processing, solutions aimed at specific cases have been proposed. The development of efficient and numerically robust adaptive algorithms in blind source separation, for either instantaneous or convolutive mixtures, remains an open challenge. On the other hand, blind deconvolution of communication channels has been studied since the 1960s and 1970s. Since then, various types of efficient adaptive solutions have been proposed in this field. The proper understanding of these solutions can suggest a path to further understand the existing solutions for the broader problem of blind source separation and to obtain efficient algorithms in this context.

Consequently, in this work we (i) revisit the problem formulation of blind source separation and blind deconvolution of multiuser systems, and the existing relations between these problems, (ii) address the existing solutions for blind deconvolution in the multiuser case, verifying their limitations and proposing modifications, resulting in the development of algorithms with proper separation performance and numeric robustness, and (iii) relate the kurtosis based criteria of blind multiuser deconvolution and blind source separation.

Keywords: Blind deconvolution of multiuser systems. Blind source separation. Adaptive algorithms. Statistical signal processing.

LISTA DE FIGURAS

Figura 1 – Esquema simplificado do problema de BSS, envolvendo a separação de N_u fontes a partir da observação de L misturas.	24
Figura 2 – Modelo equivalente de tempo discreto de um sistema de comunicação MIMO.	32
Figura 3 – (a) Parte real do erro $e_i(n)/\gamma = \bar{d}_i(n) - y_i(n)$ em função de $y_{i,R}(n)$. (b) Parte real de $d_i(n)$ em função de $y_{i,R}(n)$ para constelação transmitida 64-QAM. As coordenadas da parte real dos símbolos da constelação correspondentes aos valores de $e_{i,R}(n)/\gamma$ e $d_{i,R}(n)$ são indicadas por pontos pretos.	44
Figura 4 – Diagramas de constelação para as duas saídas do equalizador MU-DM-QRD-SWA, no Cenário 3.5–A, com SNR = 35 dB nas saídas do canal e constelação 4-QAM. São consideradas três janelas de 1×10^4 amostras, cada uma terminando na n -ésima iteração.	55
Figura 5 – Curvas de MSE para sinais 4-QAM no Cenário 3.5–A, obtidas a partir de 50 realizações independentes.	56
Figura 6 – Diagramas de constelação para as duas saídas do equalizador MU-DM-QRD-SWA, no Cenário 3.5–B, com SNR = 40 dB nas saídas do canal e constelação 64-QAM. São consideradas três janelas de 1×10^4 amostras, cada uma terminando na n -ésima iteração.	57
Figura 7 – Curvas de MSE para sinais 64-QAM no Cenário 3.5–B, obtidas a partir de 50 realizações independentes.	58
Figura 8 – Validação do modelo para a esperança de $ \epsilon_i(n) ^2$ no Cenário 4.2–A, para sinal 16-QAM em um ambiente estacionário, com média de 100 realizações independentes para a curva experimental de cada equalizador. As curvas estão em função de $1 - \lambda$, sendo λ o fator de esquecimento adotado. No gráfico são mostradas curvas do modelo da Equação (B.56) (deduzido nesse trabalho) e o modelo da Equação (4.92) (segundo a abordagem de (LUO; CHAMBERS, 2002)).	78
Figura 9 – Validação do modelo do EMSE em regime ζ_i (Equação (4.85)) no Cenário 4.2–A, para sinal 16-QAM, com média de 100 realizações independentes para a curva experimental de cada equalizador. As curvas estão em função de $1 - \lambda$, sendo λ o fator de esquecimento adotado. (a) Ambiente estacionário. (b) Ambiente com <i>tracking</i> e $\sigma_q = 1 \times 10^{-4}$. (c) Ambiente com <i>tracking</i> e $\alpha_q = 1 \times 10^{-5}$	79

Figura 10 – Validação do modelo para a esperança de $|\epsilon_i(n)|^2$ no Cenário 4.2–B, com sinal 16-QAM em um ambiente estacionário, com média de 100 realizações independentes para a curva experimental de cada equalizador. As curvas estão em função de $1 - \lambda$, sendo λ o fator de esquecimento adotado. No gráfico são mostradas curvas do modelo da Equação (B.56) (deduzido nesse trabalho) e o modelo da Equação (4.92) (segundo a abordagem de (LUO; CHAMBERS, 2002)). 80

Figura 11 – Validação do modelo do EMSE em regime ζ_i (Equação (4.85)) no Cenário 4.2–B, para sinal 16-QAM, com média de 100 realizações independentes para a curva experimental de cada equalizador. As curvas estão em função de $1 - \lambda$, sendo λ o fator de esquecimento adotado. (a) Ambiente estacionário. (b) Ambiente com *tracking* e $\sigma_q = 1 \times 10^{-4}$. (c) Ambiente com *tracking* e $\alpha_q = 1 \times 10^{-5}$ 82

LISTA DE TABELAS

Tabela 1 – Parâmetros principais e inicialização do MU-DM-SWA.	48
Tabela 2 – Equações que implementam o MU-DM-SWA.	49
Tabela 3 – Parâmetros principais e inicialização do MU-DM-QRD-SWA.	51
Tabela 4 – Equações que implementam o MU-DM-QRD-SWA.	52
Tabela 5 – Polinômios $H_{i,\ell}(z)$ do canal MIMO no Cenário 3.5–A.	53
Tabela 6 – Polinômios $H_{i,\ell}(z)$ do canal MIMO no Cenário 3.5–B.	53
Tabela 7 – Parâmetros dos três algoritmos nos Cenários 3.5–A e B.	54
Tabela 8 – Recuperação das fontes no Cenário 3.5–A.	55
Tabela 9 – Recuperação das fontes no Cenário 3.5–B.	56
Tabela 10 – Polinômios $H_{i,\ell}(z)$ do canal MIMO no Cenário 4.2–A.	75
Tabela 11 – Polinômios $H_{i,\ell}(z)$ do canal MIMO no Cenário 4.2–B.	76
Tabela 12 – Parâmetros dos MU-DM-SWA nos Cenários 4.2–A e B.	76
Tabela 13 – Recuperação das fontes no Cenário 4.2–A.	77
Tabela 14 – Recuperação das fontes no Cenário 4.2–B.	80

LISTA DE ABREVIATURAS E SIGLAS

BSS	<i>Blind source separation</i> (Separação cega de fontes)
CMA	<i>Constant modulus algorithm</i> (Algoritmo de módulo constante)
DM	<i>Dual-mode</i> (Modo duplo)
EMSE	<i>Excess mean square error</i> (Erro quadrático médio em excesso)
FFT	<i>Fast Fourier transform</i>
FIR	<i>Finite impulse response</i> (Resposta ao pulso unitário finita)
ICA	<i>Independent component analysis</i> (Análise de Componentes Independentes)
i.i.d.	independente e identicamente distribuída
JADE	<i>Joint approximate diagonalization of eigenmatrices</i>
MIMO	<i>Multiple-input and multiple-output</i> (Múltiplas entradas e múltiplas saídas)
MMA	<i>Multimodulus algorithm</i> (Algoritmo multimódulo)
MSE	<i>Mean square error</i> (Erro quadrático médio)
MU	<i>Multiuser</i> (Multiusuário)
QAM	<i>Quadrature amplitude modulation</i> (Modulação de amplitude em quadratura)
QRD	<i>QR decomposition</i> (Decomposição QR)
RLS	<i>Recursive least squares</i>
RoI	<i>Region of interest</i> (Região de interesse)
SISO	<i>Single-input and single-output</i> (Uma entrada e uma saída)
SNR	<i>Signal-to-noise ratio</i> (Relação sinal-ruído)
SWA	<i>Shalvi-Weinstein algorithm</i> (Algoritmo de Shalvi-Weinstein)

LISTA DE SÍMBOLOS

$a_i(n)$	sinal da i -ésima fonte no instante n
i	índice de fonte ou de fonte estimada
\mathbb{C}	conjunto dos números complexos
N_u	número de fontes
$u_\ell(n)$	sinal da ℓ -ésima mistura no instante n
ℓ	índice de mistura ou de sensor
L	número de misturas ou de sensores
$\mathbf{a}(n)$	vetor de fontes no instante n
$(\cdot)^\top$	transposição de um vetor ou de uma matriz
\mathbb{C}^A	conjunto de vetores de A entradas complexas
$\mathbf{u}(n)$	vetor de misturas ou regressor no instante n
$\mathcal{H}\{\cdot\}$	mapeamento de mistura
$y_i(n)$	sinal da i -ésima fonte estimada no instante n
$\mathbf{y}(n)$	vetor de fontes estimadas no instante n
$\mathcal{W}\{\cdot\}$	mapeamento de separação
$(\cdot)^{-1}$	inverso de um mapeamento, de um escalar ou de uma matriz
u_ℓ	ℓ -ésima mistura instantânea
a_i	i -ésima fonte instantânea
$h_{\ell,i}$	coeficiente associando a i -ésima fonte à ℓ -ésima mistura
\mathbf{u}	vetor de misturas para o caso instantâneo
\mathbf{H}	matriz de coeficientes do sistema misturador instantâneo
\mathbf{a}	vetor de fontes para o caso instantâneo
$\mathbb{C}^{A \times B}$	conjunto das matrizes $A \times B$ com entradas complexas
\mathbf{y}	vetor de fontes estimadas para o caso instantâneo

\mathbf{W}	matriz de coeficientes do sistema separador instantâneo
f	função densidade de probabilidade
Λ	matriz diagonal com entradas escalares
\mathbf{P}	matriz de permutação
$h_{\ell,i}(k)$	coeficiente associando a i -ésima fonte à ℓ -ésima mistura, para atraso k
$*$	convolução de tempo discreto
$\mathbf{H}(k)$	matriz de coeficientes do sistema misturador convolutivo, para atraso k
$\mathbf{H}[z]$	matriz de sistema de mistura
$\mathbb{C}[z]$	conjunto de polinômios com potências inteiras em z e coeficientes complexos
$\mathbb{C}[z]^{A \times B}$	conjunto de matrizes $A \times B$ de entradas polinomiais, com potências inteiras em z e coeficientes complexos
$\mathbf{u}[z]$	transformada z do vetor de misturas
$\mathbb{C}[z]^A$	conjunto vetores de A entradas polinomiais, com potências inteiras em z e coeficientes complexos
$\mathbf{a}[z]$	transformada z do vetor de fontes
$\mathbf{W}[z]$	matriz de sistema de separação
\mathbf{I}_A	matriz identidade $A \times A$
$\mathbb{R}^{A \times B}$	conjunto das matrizes $A \times B$ com entradas reais
$\Lambda[z]$	matriz diagonal de polinômios em z
K_c	número de coeficientes de cada filtro FIR do canal de comunicações
$H_{i,\ell}(z)$	polinômio em z que representa o canal que associa a i -ésima fonte ao ℓ -ésimo sensor
$\eta_\ell(n)$	ruído gaussiano branco aditivo recebido na ℓ -ésima mistura, no instante n
K_t	inteiro associado à diversidade temporal de cada filtro FIR do equalizador
M	número de coeficientes de cada filtro FIR do equalizador
\mathbf{w}_i	vetor de coeficientes do i -ésimo equalizador

$\bar{\mathbf{H}}$	matriz de convolução do canal de comunicações MIMO
$\boldsymbol{\eta}(n)$	vetor de ruído gaussiano branco aditivo no instante n
$\mathbf{H}_{i,\ell}$	matriz de convolução correspondente ao canal que associa a i -ésima fonte ao ℓ -ésimo sensor
θ_i	ângulo de rotação no plano complexo do sinal recuperado pelo i -ésimo equalizador
τ_i	atraso temporal do sinal recuperado pelo i -ésimo equalizador
j	unidade imaginária
$\mathbf{s}_{zf,i}$	vetor de resposta combinada da cascata do canal de comunicação com o i -ésimo equalizador
$H_i(z)$	i -ésimo polinômio em z para teste de equalização perfeita
$\hat{J}_i(n)$	função custo multimódulo para o i -ésimo equalizador, no instante n
λ	fator de esquecimento
\mathcal{R}	constante de dispersão
$y_i(\ell, n)$	saída do i -ésimo equalizador, com regressor no instante ℓ e vetor de coeficientes no instante n
$\mathbf{w}_i(n)$	vetor de coeficientes do i -ésimo equalizador, no instante n
$\text{Re}\{\cdot\}, (\cdot)_{\text{R}}$	parte real
$\text{Im}\{\cdot\}, (\cdot)_{\text{i}}$	parte imaginária
ϱ	constante positiva de ponderação do termo que penaliza a correlação cruzada
$ \cdot $	módulo (valor absoluto)
k_1	inteiro associado ao atraso de espalhamento
$c_{i,m}(n, k)$	correlação cruzada determinista entre as saídas do i -ésimo e m -ésimo equalizadores, com atraso k , no instante n
$\text{E}[\cdot]$	esperança matemática
$(\cdot)^*$	complexo conjugado
$(\cdot)^{\text{H}}$	hermitiano (transposto conjugado) de um vetor ou de uma matriz

$\frac{\partial(\cdot)}{\partial \mathbf{x}}$	derivada parcial de função escalar em relação ao vetor \mathbf{x} (operador cogradiante)
$\nabla_{\mathbf{x}}(\cdot)$	gradiente complexo de função escalar em relação ao vetor \mathbf{x}
$\mathbf{0}$	vetor nulo
$\widehat{\mathbf{R}}(n)$	estimativa da matriz de covariância do regressor no instante n
$\mathbf{p}_i^y(n)$	vetor de correlação cruzada correspondente ao termo de dispersão da função custo, no instante n
$\mathbf{p}_i^c(n)$	vetor de correlação cruzada correspondente ao termo de correlação cruzada da função custo, no instante n
$\hat{c}_{i,m}(n, k)$	estimativa da correlação cruzada entre as saídas do i -ésimo e m -ésimo equalizadores, com atraso k , no instante n
γ	fator de normalização do erro do algoritmo de Shalvi-Weinstein
$\bar{e}_i(n)$	erro do algoritmo de Shalvi-Weinstein, para o i -ésimo equalizador, no instante n
$e_i(n)$	erro devido à minimização da dispersão, para o i -ésimo equalizador, no instante n
$\epsilon_i(n)$	erro devido à minimização da correlação cruzada, para o i -ésimo equalizador, no instante n
$\bar{d}_i(n)$	termo que compreende a não linearidade do erro do algoritmo de Shalvi-Weinstein, para o i -ésimo equalizador, no instante n
$d_i(n)$	termo que compreende a não linearidade do erro devido à minimização de dispersão do algoritmo de Shalvi-Weinstein, para o i -ésimo equalizador, no instante n
$\sqrt{(\cdot)}$	raiz quadrada
β	constante igual a 3 para o caso real e igual a 2 para o caso não real
$\mathbf{y}_i(n)$	vetor de linhas de atraso de saídas de todos os equalizadores, exceto o i -ésimo equalizador, no instante n
$\hat{\mathbf{c}}_i(n)$	vetor de linhas de atraso de correlações cruzadas entre o i -ésimo e os demais equalizadores, no instante n

δ	escalar que multiplica a matriz identidade correspondente à inicialização de $\widehat{\mathbf{R}}^{-1}(n)$
$\xi_i(n)$	erro $\bar{e}_i(n)$ do algoritmo de Shalvi-Weinstein para o i -ésimo equalizador, dividido pelo fator de normalização γ
$\mathcal{C}(n)$	fator de Cholesky de $\widehat{\mathbf{R}}(n)$
ς	raiz quadrada do fator de conversão
$\bar{\mathbf{p}}_i(n)$	vetor transformado de coeficientes do algoritmo QR, para o i -ésimo equalizador, no instante n
$\Theta(n)$	matriz de rotação do algoritmo QR, no instante n
$\mathbf{b}(n)$	vetor regressor transformado do algoritmo QR, no instante n
$\text{MSE}_i(n)$	erro quadrático médio do i -ésimo equalizador, no instante n
$\ (\cdot)\ $	norma Euclidiana de um vetor ou norma espectral de uma matriz
$\mathbf{u}(n)$	sequência de tempo discreto correspondente ao vetor regressor $\mathbf{u}(n)$
α_1	escalar que multiplica a matriz identidade correspondente ao limite inferior de $\widehat{\mathbf{R}}(n)$
α_2	escalar que multiplica a matriz identidade correspondente ao limite superior de $\widehat{\mathbf{R}}(n)$
$B_{\mathbf{u}}$	limite superior da norma do vetor regressor
$B_{\mathbf{d}}$	limite superior correspondente ao sinal $d_i(n)$
B_{ϵ}	limite superior correspondente ao sinal $\epsilon_i(n)$
ζ_i	erro quadrático médio em excesso do i -ésimo equalizador, em regime
$e_{a,i}(n)$	erro <i>a priori</i> do i -ésimo equalizador, no instante n
$\tilde{\mathbf{w}}_i(n)$	vetor de ponderação de erro do i -ésimo equalizador, no instante n
$\mathbf{w}_{o,i}(n)$	solução ótima do problema de mínimos quadrados para o i -ésimo equalizador, no instante n
\mathbb{N}	conjunto dos números naturais
$v_i(n)$	ruído para o i -ésimo equalizador, considerado na análise em regime para levar em conta as imperfeições do modelo, no instante n

$\mathbf{q}_i(n)$	vetor aleatório i.i.d. de média nula do modelo de <i>random-walk</i> , para o i -ésimo equalizador, no instante n
\mathbf{Q}_i	matriz de covariância do vetor $\mathbf{q}_i(n)$
$\ \mathbf{x}\ _{\mathbf{A}}$	norma Euclidiana no vetor \mathbf{x} , ponderada pela matriz positiva definida \mathbf{A}
$\mathbf{R}_{\mathbf{x}}$	matriz de covariância do vetor aleatório \mathbf{x}
\mathbf{p}_i	vetor de correlação cruzada entre a i -ésima fonte recuperada e o vetor regressor
$\varepsilon_i(n)$	perturbação considerada no modelo de análise em regime, para o i -ésimo equalizador, no instante n
$\text{Tr}(\cdot)$	traço de uma matriz
mom	momento conjunto
cum	cumulante conjunto
$(\cdot)!$	fatorial
ω_i	i -ésima frequência angular real da transformada de Fourier multidimensional
$\boldsymbol{\omega}$	vetor de frequências angulares reais da transformada de Fourier multidimensional
$\Phi_{\mathbf{x}}$	primeira função característica do vetor aleatório \mathbf{x}
$\Psi_{\mathbf{x}}$	segunda função característica do vetor aleatório \mathbf{x}
$\ln(\cdot)$	logaritmo natural
$\text{COV}(\cdot, \cdot)$	covariância
C_p^x	cumulante de ordem p da variável aleatória real x
$C_{p,q}^x$	cumulante de ordem $p + q$ da variável aleatória complexa x
$\text{VAR}[\cdot]$	variância
κ_x	curtose da variável aleatória x
J_{SW}	função custo de Shalvi-Weinstein
J_{G}	função custo de Godard

\mathfrak{R}_p^x	constante de dispersão de ordem p da variável aleatória x
$\hat{J}_G(n)$	função custo determinista de Godard, no instante n
J_{GMM}	função custo de Godard multimódulo
\mathfrak{R}_p^x	constante de dispersão multimódulo de ordem p da variável aleatória x
J_{SWMM}	função custo de Shalvi-Weinstein multimódulo
$r_{i,m}(k)$	correlação cruzada entre as saídas do i -ésimo e m -ésimo equalizadores, com atraso k
$J_{\text{SW},i}$	função custo de Shalvi-Weinstein multiusuário, para o i -ésimo equalizador
$J_{G,i}$	função custo de Godard multiusuário, para o i -ésimo equalizador
$J_{\text{SWMM},i}$	função custo de Shalvi-Weinstein multimódulo e multiusuário, para o i -ésimo equalizador
$J_{\text{GMM},i}$	função custo de Godard multimódulo e multiusuário, para o i -ésimo equalizador
$\hat{J}_{\text{GMM},i}(n)$	função custo determinista de Godard, multimódulo e multiusuário, para o i -ésimo equalizador, no instante n
ϕ°	contraste ortogonal
$\mathbf{u}[e^{j\omega}]$	transformada de Fourier de tempo discreto do vetor de misturas
$\mathbf{H}[e^{j\omega}]$	resposta em frequência do sistema misturador
$\mathbf{a}[e^{j\omega}]$	transformada de Fourier de tempo discreto do vetor de fontes
$\frac{\partial(\cdot)}{\partial x}$	derivada parcial de função escalar em relação ao escalar x
r_α^f	esperança do módulo de f elevado a α
$\delta_{\ell,m}$	delta de Kronecker

SUMÁRIO

1	INTRODUÇÃO	19
1.1	Desafios atuais e Justificativa	20
1.2	Objetivos do trabalho	21
1.3	Organização do trabalho	21
2	OS PROBLEMAS DE SEPARAÇÃO DE FONTES E DES- CONVOLUÇÃO MULTIUSUÁRIO	23
2.1	Formulação do problema de separação de fontes	23
2.2	Modelos usuais de misturas lineares e fontes	25
2.2.1	Misturas instantâneas e fontes independentes entre si	25
2.2.2	Misturas convolutivas e fontes com independência espacial e temporal	28
2.3	Formulação do problema de desconvolução multiusuário	31
2.4	Conclusão	34
3	O ALGORITMO DE SHALVI-WEINSTEIN MULTIUSUÁRIO	35
3.1	Soluções existentes	35
3.2	Dedução do SWA multiusuário	36
3.3	Operação em <i>dual-mode</i>	43
3.4	Considerações sobre a implementação	46
3.4.1	Recursões preliminares	46
3.4.2	MU-DM-SWA	47
3.4.3	MU-DM-QRD-SWA	50
3.5	Resultados de simulação	53
3.6	Conclusão	57
4	ANÁLISES DO ALGORITMO DE SHALVI-WEINSTEIN	59
4.1	Análise determinista	59
4.1.1	Inicialização dentro e fora da região de interesse	61
4.1.2	Transição entre modos de operação	65
4.2	Análise em regime	66
4.2.1	Indicadores de desempenho e hipóteses iniciais	67
4.2.2	Considerações sobre o modelo completo	69
4.2.3	Expressão para o erro quadrático médio em excesso	71
4.2.4	Validação do modelo	75
4.3	Conclusão	83

5	CRITÉRIOS DE SEPARAÇÃO DE FONTES BASEADOS EM CURTOSE	85
5.1	Cumulantes e curtose	85
5.2	Desconvolução cega baseada em cumulantes	89
5.2.1	Caso SISO	89
5.2.2	Caso SISO multimódulo	90
5.2.3	Caso MIMO	91
5.3	Separação de fontes baseada em cumulantes	93
5.4	Abordagens para misturas convolutivas	95
5.5	Conclusão	96
6	CONCLUSÕES	98
	REFERÊNCIAS BIBLIOGRÁFICAS	100
	APÊNDICE A – CÁLCULO DE WIRTINGER VETORIAL	106
A.1	Operador cogradiente	106
A.2	Gradiente complexo	108
	APÊNDICE B – DEDUÇÃO DE MOMENTOS DA ANÁLISE EM REGIME	109
B.1	Definições	109
B.2	Resultados preliminares	110
B.3	Momentos da análise em regime	114

1 INTRODUÇÃO

Basicamente, o problema de separação cega de fontes (BSS, do inglês, *blind source separation*) consiste em separar individualmente um conjunto de sinais ou fontes, somente com base na observação de diferentes misturas desses sinais (CARDOSO, 1998; HYVÄRINEN; OJA, 2000). Esse problema é classificado como cego visto que geralmente não se tem conhecimento *a priori* das fontes, ou, na melhor situação, sabe-se pouco sobre as suas características estatísticas. Além disso, o conhecimento sobre o sistema físico responsável por realizar as misturas também é, em geral, escasso.

Em uma primeira análise, poder-se-ia conjecturar que, devido a essa falta de informações tanto sobre as fontes quanto sobre o sistema misturador, seria impossível resolver esse problema. No entanto, a depender da situação, pode-se lançar mão de ferramentas estatísticas que permitam separar e recuperar as fontes, mesmo sem conhecê-las *a priori* ou sem saber, a fundo, como as misturas são feitas (CARDOSO, 1998; HYVÄRINEN; OJA, 2000). O ser humano, por exemplo, quando consegue identificar e compreender a voz de uma pessoa em um ambiente barulhento, sem saber de antemão o que ela vai dizer nem como as ondas sonoras são misturadas, está resolvendo um problema de separação cega de fontes.

O problema de BSS foi formulado pela primeira vez no início dos anos 1980, por Bernard Ans, Jeanny Héroult e Christian Jutten, no contexto de modelamento neural para decodificação de movimentos em vertebrados (HÉRAULT; JUTTEN; ANS, 1985). No entanto, os princípios teóricos que asseguram a existência de soluções foram somente compreendidos posteriormente (COMON, 1994). As primeiras publicações relacionadas à separação cega de fontes apareceram em torno de 1985 e chamaram a atenção de pesquisadores principalmente na França e no resto da Europa (JUTTEN; COMON, 2010).

Inicialmente, o problema foi tratado para casos em que o sistema misturador pudesse ser modelado como linear e instantâneo. Posteriormente, começou a surgir interesse no estudo de sistemas misturadores mais complicados, *e.g.*, sistemas convolutivos e não lineares, embora de grande relevância prática. Um dos conjuntos de técnicas mais importantes de BSS, conhecido como análise de componentes independentes (ICA, do inglês, *independent component analysis*), foi introduzido em 1987 e formalizado matematicamente para misturas lineares e instantâneas por Comon em 1991 (COMON, 1992; COMON, 1994). As técnicas de ICA são, em geral, baseadas na independência estatística das fontes desconhecidas e na sua não gaussianidade (CARDOSO, 1998; HYVÄRINEN; OJA, 2000). Desde os anos 1990, o interesse na área de BSS e nas suas aplicações tem sido muito grande na comunidade científica internacional (JUTTEN; COMON, 2010). Em particular, métodos de BSS têm

sido usados amplamente em aplicações biomédicas (LI et al., 2014; ADALI; ANDERSON; FU, 2014), de áudio (VINCENT et al., 2014; EWERT et al., 2014; RIVET et al., 2014) e de comunicações (CHEVALIER; CHEVREUIL, 2010). A necessidade de se separar fontes desconhecidas em misturas aparece, por exemplo, em sistemas de múltiplos sensores, como redes de microfones ou antenas, redes de sensores químicos (DUARTE; MOUSSAOUI; JUTTEN, 2014), análise de sinais neuronais (BROWN; YAMADA; SEJNOWSKI, 2001; OWEISS, 2010), dentre outros.

Em particular, uma aplicação de interesse prático em BSS é a equalização cega de canais de comunicação de múltiplas entradas e múltiplas saídas (MIMO, do inglês, *multiple-input and multiple-output*), também conhecida como desconvolução cega multiusuário. Esse problema consiste em separar sinais de telecomunicação desconhecidos, misturados por diferentes canais de comunicação e captados posteriormente em um conjunto de antenas ou sensores. O sistema misturador, nesse caso, costuma ser modelado como linear e convolutivo, pois canais de comunicação apresentam interferência intersimbólica (DING; LI, 2001). Trata-se de um modelo de mistura mais sofisticado do que uma simples mistura linear e instantânea. Entretanto, o conhecimento sobre as estatísticas das fontes que se deseja separar é maior do que em problemas gerais de separação cega de fontes: na desconvolução cega multiusuário, existe conhecimento completo sobre as estatísticas dos símbolos transmitidos, visto que eles pertencem, em geral, a um alfabeto finito. Devido a esse conhecimento mais apurado sobre as fontes, a desconvolução cega multiusuário é também conhecida como um problema semicego (JUTTEN; COMON, 2010; ZARZOSO; COMON; SLOCK, 2010).

As primeiras soluções para desconvolução cega monousuário, que não consiste em um problema de separação de fontes, foram propostas nos anos 1960 e 1970 (LUCKY, 1966; SATO, 1975). Para o problema multiusuário, soluções adaptativas *online* baseadas em equalizadores cegos MIMO foram propostas na literatura nos anos 1990 e 2000, com o objetivo de se melhorar o uso eficiente da banda do canal (PAPADIAS; PAULRAJ, 1997a; SILVA; MIRANDA; LICCIARDI, 2004). Devido a problemas de convergência e instabilidade numérica, essas soluções multiusuário continuam sendo investigadas na área de Processamento Estatístico de Sinais (SHAH; ABED-MERAIM; AL-NAFFOURI, 2015).

1.1 DESAFIOS ATUAIS E JUSTIFICATIVA

Dentre os desafios atuais na área de separação cega de fontes e desconvolução cega multiusuário, cabe destacar que:

- A obtenção de soluções em bloco para o problema de separação cega de fontes com misturas convolutivas não é trivial, pois envolve ferramentas matemáticas avançadas e uma compreensão aprofundada das técnicas estatísticas a serem utilizadas (HAYKIN,

2000b; ZARZOSO; COMON; SLOCK, 2010; CASTELLA; CHEVREUIL; PESQUET, 2010). No caso em que não se conhece o tipo de mistura ou as estatísticas das fontes, o problema é ainda mais desafiador (HAYKIN, 2000a; CASTELLA; CHEVREUIL; PESQUET, 2010).

- A obtenção de algoritmos adaptativos eficientes e numericamente robustos para realizar separação cega de fontes, tanto envolvendo misturas instantâneas quanto convolutivas, ainda é um desafio. Soluções vêm sendo propostas para resolver casos específicos na área de Processamento Estatístico de Sinais (HAYKIN, 2000a; JUTTEN; COMON, 2010).

Além disso, observa-se que:

- A relação matemática existente entre os problemas de desconvolução cega multiusuário e separação cega de fontes ainda precisa ser melhor compreendida, para que seja possível vislumbrar novas técnicas e soluções eficientes.
- A compreensão do problema de desconvolução cega multiusuário, das soluções adaptativas existentes para resolvê-lo e das suas limitações, pode auxiliar na compreensão e obtenção de soluções adaptativas eficientes para o problema de separação cega de fontes.

1.2 OBJETIVOS DO TRABALHO

Os objetivos deste trabalho são elencados a seguir.

- Compreender a formulação dos problemas de separação cega de fontes e desconvolução cega multiusuário, bem como as principais relações existentes entre eles.
- Investigar as soluções existentes para a desconvolução cega multiusuário, verificar suas limitações e propor modificações, visando obter algoritmos com boa capacidade de separação e robustez numérica.
- Compreender os critérios de desconvolução cega multiusuário baseados em curtose e suas relações com critérios de separação cega de fontes.

1.3 ORGANIZAÇÃO DO TRABALHO

Esta dissertação está estruturada em 6 capítulos, cujos conteúdos são resumidos a seguir.

No Capítulo 2, os problemas de separação cega de fontes e desconvolução cega multiusuário são apresentados. Inicialmente, é feita a formulação da separação cega de

fontes de forma genérica, sem se assumir conhecimento prévio sobre a estrutura do problema posto. Posteriormente, são apresentados modelos e hipóteses usuais adotados para os sistemas misturadores e para as fontes. Em seguida, o problema de desconvolução cega multiusuário é formulado e interpretado como um caso particular de separação cega de fontes para misturas convolutivas.

No Capítulo 3, partindo-se de uma formulação MIMO do problema de mínimos quadrados, é obtida uma solução adaptativa para a desconvolução cega multiusuário. Essa solução é posteriormente aprimorada, a fim de se evitar problemas de instabilidade numérica. Em seguida, são feitas considerações sobre a implementação das soluções obtidas e, a fim de se verificar seu funcionamento, são apresentados resultados de simulações numéricas.

No Capítulo 4, é feita a análise da solução para o problema de desconvolução cega multiusuário deduzida no Capítulo 3, quanto à estabilidade e ao desempenho em regime.

No Capítulo 5, são apresentados os critérios usuais de desconvolução cega baseados em curtose. Esses critérios são relacionados com critérios de separação cega de fontes mais abrangentes.

Por fim, no Capítulo 6, são apresentadas a conclusão do trabalho realizado e as sugestões de trabalhos futuros.

2 OS PROBLEMAS DE SEPARAÇÃO DE FONTES E DESCONVOLUÇÃO MULTIUSUÁRIO

Neste capítulo, os problemas de separação cega de fontes e desconvolução cega multiusuário são apresentados. O capítulo é organizado conforme descreve-se a seguir. Na Seção 2.1, o problema de BSS é apresentado de forma genérica, sem se assumir conhecimento prévio sobre as fontes ou sobre o sistema misturador. Na Seção 2.2, esse problema é particularizado para dois tipos de sistemas misturadores lineares comuns em aplicações práticas. Em seguida, na Seção 2.3, é apresentada a formulação do problema de desconvolução cega multiusuário, o qual pode ser interpretado como um caso particular de separação cega de fontes para misturas convolutivas. Por fim, a conclusão do capítulo é feita na Seção 2.4.

2.1 FORMULAÇÃO DO PROBLEMA DE SEPARAÇÃO DE FONTES

Sejam os sinais das fontes não observados denotados por $a_i(n) \in \mathbb{C}$, com $i = 1, 2, \dots, N_u$. Por sua vez, os sinais das misturas observadas são denotados por $u_\ell(n) \in \mathbb{C}$, com $\ell = 1, 2, \dots, L$. A partir dos sinais de cada fonte e de cada mistura, definem-se o vetor de fontes como

$$\mathbf{a}(n) = [a_1(n) \quad a_2(n) \quad \cdots \quad a_{N_u}(n)]^T \in \mathbb{C}^{N_u}, \quad (2.1)$$

em que $(\cdot)^T$ denota a transposta de um vetor ou matriz, e o vetor de misturas como

$$\mathbf{u}(n) = [u_1(n) \quad u_2(n) \quad \cdots \quad u_L(n)]^T \in \mathbb{C}^L. \quad (2.2)$$

Em geral, considera-se que os sinais estão representados em tempo discreto. Sendo assim, a variável independente n representa o instante de tempo. No entanto, a depender da forma como se pretende resolver o problema de BSS, pode ser necessário considerar que a variável independente corresponde à frequência de alguma representação espectral dos sinais, por exemplo. A relação existente entre o vetor de fontes $\mathbf{a}(n)$ e o vetor de misturas $\mathbf{u}(n)$ pode ser dada, genericamente, por (JUTTEN; COMON, 2010)

$$\mathbf{u}(n) = \mathcal{H}\{\mathbf{a}(\cdot)\}, \quad (2.3)$$

em que \mathcal{H} é um mapeamento desconhecido de \mathbb{C}^{N_u} em \mathbb{C}^L , responsável por misturar as fontes que se deseja, posteriormente, separar. A notação de (2.3) indica que o vetor $\mathbf{u}(n)$ depende de uma transformação \mathcal{H} aplicada ao conjunto de vetores $\{\mathbf{a}(\cdot)\}$ à entrada do sistema, para instantes de tempo menores ou iguais a n . Em outras palavras, essa notação leva em conta a eventual diversidade temporal ou memória que o mapeamento de mistura

possa ter. Caso esse mapeamento seja instantâneo, *i.e.*, sem diversidade temporal, (2.3) se reduz a (JUTTEN; COMON, 2010)

$$\mathbf{u}(n) = \mathcal{H}\{\mathbf{a}(n)\}. \quad (2.4)$$

Em um problema de BSS, busca-se reconstruir os sinais das fontes $a_i(n)$ a partir das misturas observadas $u_\ell(n)$. Portanto, é usual levar em conta o sistema responsável pela separação, bem como os sinais reconstruídos, na formulação do problema. Os sinais correspondentes às fontes estimadas ou reconstruídas são denotados por $y_i(n) \in \mathbb{C}$, com $i = 1, 2, \dots, N_u$. Por conveniência, constitui-se um vetor de fontes estimadas dado por

$$\mathbf{y}(n) = [y_1(n) \ y_2(n) \ \dots \ y_{N_u}(n)]^T \in \mathbb{C}^{N_u}, \quad (2.5)$$

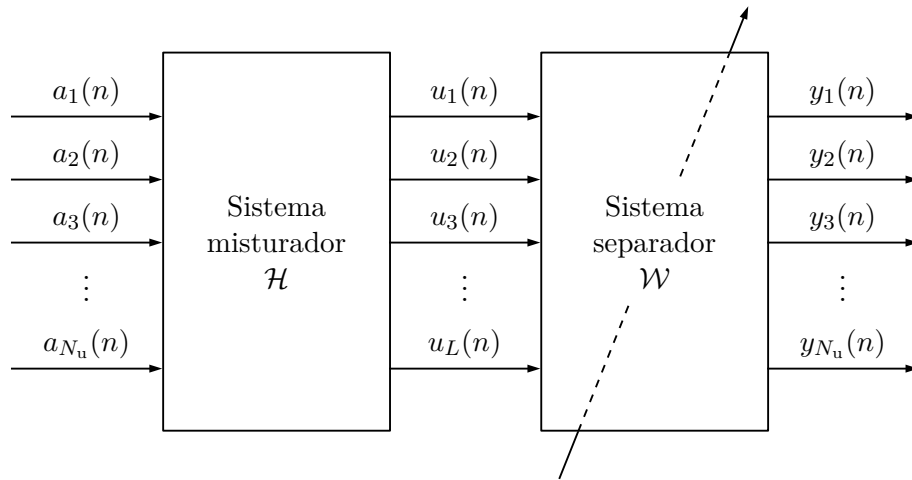
que deve satisfazer a equação

$$\mathbf{y}(n) = \mathcal{W}\{\mathbf{u}(\cdot)\}, \quad (2.6)$$

em que \mathcal{W} é um mapeamento de \mathbb{C}^L em \mathbb{C}^{N_u} , responsável por separar as fontes, ou seja, recuperá-las a partir das misturas observadas.

Na Figura 1, é apresentado um diagrama de blocos para descrever o problema de BSS. Observa-se que o modelo considerado para esse problema envolve, no caso geral, sistemas MIMO, tanto na etapa de mistura quanto na etapa de separação. Além disso, se o mapeamento \mathcal{H} for conhecido e existir mapeamento inverso \mathcal{H}^{-1} , então a escolha de $\mathcal{W} = \mathcal{H}^{-1}$ recuperaria as fontes a partir das misturas observadas. No entanto, nos problemas de separação cega de fontes, usualmente não se conhece o mapeamento \mathcal{H} responsável por gerar as misturas, e seu inverso pode não existir.

Figura 1 – Esquema simplificado do problema de BSS, envolvendo a separação de N_u fontes a partir da observação de L misturas.



Fonte: autoria própria.

A formulação feita até aqui considerou, de forma genérica, que existem N_u fontes e L misturas. Quando $N_u = L$, o problema de BSS é dito igualmente determinado. No

caso em que há mais fontes do que misturas, *i.e.*, $N_u > L$, o problema é dito subdeterminado. Quando $L > N_u$ e há mais misturas do que fontes, diz-se que o problema é sobredeterminado (JUTTEN; COMON, 2010).

Se não há algum conhecimento prévio nem hipótese razoável sobre o sistema misturador ou sobre as estatísticas das fontes, o problema de BSS tem mais incógnitas do que equações. Em outras palavras, tem-se uma quantidade de incógnitas a serem determinadas, *e.g.*, os parâmetros do sistema misturador que se deseja inverter, mas não há equações que permitam encontrá-las. Assim, não é possível encontrar as fontes a partir, por exemplo, da solução direta de um sistema de equações. A fim de se contornar esse problema, na literatura foram propostas soluções que usam conhecimento prévio ou hipóteses razoáveis sobre o sistema misturador e sobre as estatísticas das fontes (CARDOSO, 1998; HYVÄRINEN; OJA, 2000).

2.2 MODELOS USUAIS DE MISTURAS LINEARES E FONTES

Para que exista alguma solução para problema de BSS formulado na Seção 2.1, é preciso fazer suposições preliminares razoáveis sobre a forma como as misturas são feitas e sobre as estatísticas das fontes (CARDOSO, 1998; HYVÄRINEN; OJA, 2000). Se houver algum conhecimento prévio sobre a estrutura do sistema misturador ou sobre as distribuições das fontes, pode-se utilizá-lo na obtenção de soluções (CARDOSO, 1998). Essas informações adicionais, que costumam acompanhar o problema de BSS, podem ser interpretadas como restrições que permitem resolvê-lo. A seguir, são apresentados dois modelos usuais de misturas lineares e as respectivas hipóteses sobre as fontes que costumam acompanhá-los.

2.2.1 Misturas instantâneas e fontes independentes entre si

O modelo mais simples para o sistema misturador consiste em considerá-lo linear e instantâneo (*i.e.*, sem memória) (CARDOSO, 1998; JUTTEN; COMON, 2010). Trata-se do primeiro modelo de mistura a ser estudado na área de BSS (JUTTEN; COMON, 2010). Sua simplicidade vem do fato de não haver diversidade temporal na composição das misturas. Quando as misturas são instantâneas, ou seja, respeitam (2.4), a saída do sistema misturador, em cada instante, depende somente dos valores das fontes nesse dado instante.

Em problemas de BSS com misturas lineares e instantâneas, a diversidade temporal das fontes é usualmente ignorada (CARDOSO, 1998; HYVÄRINEN; OJA, 2000). Em outras palavras, deixa-se de considerar que as fontes são processos estocásticos e passa-se a interpretá-las como sorteios independentes, ao longo do tempo, de uma mesma variável aleatória. Portanto, nesse contexto, o índice de tempo n dos sinais das fontes, das misturas

e das fontes recuperadas costuma ser omitido (HYVÄRINEN; OJA, 2000; STONE, 2004; JUTTEN; COMON, 2010).

As misturas lineares e instantâneas u_ℓ são obtidas a partir das fontes a_i de acordo com a relação (CARDOSO, 1998)

$$u_\ell = \sum_{i=1}^{N_u} h_{\ell,i} a_i, \quad (2.7)$$

para $\ell = 1, 2, \dots, L$. As constantes $h_{\ell,i} \in \mathbb{C}$ correspondem aos coeficientes das combinações lineares das fontes, que, por sua vez, compõem as misturas. Por conveniência, a Equação (2.7) pode ser reescrita matricialmente como

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_L \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,N_u} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,N_u} \\ \vdots & \vdots & \ddots & \vdots \\ h_{L,1} & h_{L,2} & \cdots & h_{L,N_u} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{N_u} \end{bmatrix}, \quad (2.8)$$

ou na correspondente forma compacta

$$\mathbf{u} = \mathbf{H}\mathbf{a}, \quad (2.9)$$

em que o vetor de misturas \mathbf{u} é dado pela Equação (2.2), a matriz de coeficientes do sistema misturador $\mathbf{H} \in \mathbb{C}^{L \times N_u}$ é igual a

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,N_u} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,N_u} \\ \vdots & \vdots & \ddots & \vdots \\ h_{L,1} & h_{L,2} & \cdots & h_{L,N_u} \end{bmatrix}, \quad (2.10)$$

e o vetor de fontes \mathbf{a} é dado pela Equação (2.1).

Quando o sistema misturador é linear e instantâneo, é usual adotar um sistema separador que também o seja (HYVÄRINEN; OJA, 2000). Com isso, a relação genérica de separação (2.6) pode ser particularizada como

$$\mathbf{y} = \mathbf{W}\mathbf{u}, \quad (2.11)$$

em que o vetor de fontes estimadas \mathbf{y} é dado pela Equação (2.5) e o mapeamento \mathcal{W} foi substituído pela matriz de coeficientes do sistema separador $\mathbf{W} \in \mathbb{C}^{N_u \times L}$. No caso igualmente determinado, se \mathbf{H} for não singular, o problema de BSS pode ser resolvido com

$$\mathbf{W} = \mathbf{H}^{-1}. \quad (2.12)$$

Observa-se que a matriz \mathbf{H} é determinada se todas as suas N_u^2 entradas forem conhecidas. Se as entradas de \mathbf{H} são desconhecidas, então são necessárias N_u^2 equações linearmente

independentes para determiná-las. Infelizmente, não convém utilizar as equações do modelo do sistema misturador (2.8), pois as fontes $a_i(n)$ não são conhecidas, *i.e.*, observadas. Por outro lado, se houver algum conhecimento *a priori* ou hipótese razoável sobre as estatísticas das fontes, essa informação pode ser usada para compor as N_u^2 equações linearmente independentes que permitem determinar, matematicamente, as incógnitas procuradas.

No contexto de misturas lineares e instantâneas, a principal informação *a priori* considerada para as fontes é a independência estatística entre elas (HYVÄRINEN; OJA, 2000). A hipótese de que as fontes a_i são independentes equivale a considerar que a função densidade de probabilidade conjunta do vetor de fontes \mathbf{a} é igual ao produto das funções densidade de probabilidade marginais de cada fonte a_i , ou seja (PAPOULIS; PILLAI, 2002)

$$f_{a_1, a_2, \dots, a_{N_u}}(a_1, a_2, \dots, a_{N_u}) = \prod_{i=1}^{N_u} f_{a_i}(a_i). \quad (2.13)$$

Cabe observar que a suposição de independência é mais forte e restritiva do que, por exemplo, a não correlação. Para sistemas físicos, muitas vezes a hipótese de independência é razoável para as fontes que se deseja separar. No entanto, ela não precisa ser exatamente verdadeira na prática para ser considerada e permitir a separação adequada das fontes (HYVÄRINEN; OJA, 2000). Além disso, para que seja possível encontrar uma solução para o problema de BSS, considerar somente a independência estatística entre as fontes não é suficiente. É possível provar, matematicamente, a necessidade de se adotar uma hipótese adicional de não gaussianidade das fontes (COMON, 1994; CARDOSO, 1998).

Um resultado matemático importante para misturas lineares e instantâneas diz que a recuperação da independência estatística entre as fontes permite separá-las (HYVÄRINEN; OJA, 2000). Esse resultado sobre a separabilidade das fontes foi apresentado no trabalho de Comon (COMON, 1992), com base em teoremas provados no contexto de Análise Fatorial por Darmois (DARMOIS, 1953). Segundo Comon, dispondo-se inicialmente de um vetor de fontes independentes \mathbf{a} , com no máximo uma delas sendo gaussiana, se as saídas y_i do sistema separador forem independentes, então essas saídas correspondem às fontes separadas, a menos de permutações e mudanças de escala.

Comon apresenta uma formulação analítica desse resultado. Seja $\mathbf{u} = \mathbf{H}\mathbf{a}$, em que \mathbf{H} é uma matriz não singular e \mathbf{a} é um vetor de fontes com elementos (ou componentes) estatisticamente independentes, e no máximo um de seus elementos é gaussiano. Então, $\mathbf{y} = \mathbf{W}\mathbf{u}$ é um vetor aleatório com componentes mutuamente independentes se, e somente se,

$$\mathbf{W}\mathbf{H} = \mathbf{\Lambda}\mathbf{P}, \quad (2.14)$$

em que $\mathbf{\Lambda}$ é uma matriz diagonal e \mathbf{P} é uma matriz de permutação (COMON, 1992). Esse

resultado permite compreender a existência de dois tipos de ambiguidade que ocorrem quando se deseja fazer a separação cega de fontes para misturas lineares e instantâneas: a ambiguidade de permutação, representada pela matriz \mathbf{P} , e a ambiguidade de escala ou energia, representada pela matriz $\mathbf{\Lambda}$.

2.2.2 Misturas convolutivas e fontes com independência espacial e temporal

Em algumas aplicações práticas, os sistemas misturadores são lineares mas possuem diversidade temporal. Em outras palavras, as misturas são obtidas, em cada instante de tempo, a partir de combinações lineares de inúmeras fontes consideradas nesse instante, mas também em instantes anteriores. Nesse caso, diz-se que o sistema misturador é linear e convolutivo, *i.e.*, com memória. Em BSS para áudio, por exemplo, esse modelo de mistura pode ser adequado para considerar efeitos de eco e reverberação, que dependem do passado das fontes (CASTELLA; CHEVREUIL; PESQUET, 2010). As misturas lineares e convolutivas começaram a ser estudadas no início dos anos 1990, após o estudo do caso mais simples de misturas lineares e instantâneas (JUTTEN; COMON, 2010).

Como o sistema misturador possui memória, a diversidade temporal das fontes deve ser levada em conta, de alguma forma, na solução do problema de BSS para misturas convolutivas. Portanto, diferentemente do caso de misturas instantâneas, nesse caso as fontes são usualmente consideradas como processos estocásticos e não simples variáveis aleatórias.

Para sistemas misturadores convolutivos, cada mistura em um dado instante pode sofrer o efeito de quaisquer fontes avaliadas nesse instante e em instantes passados. Para simplificar, inicialmente considera-se o caso em que são tomadas $N - 1$ amostras anteriores ao instante sendo considerado. Nesse caso, as misturas $u_\ell(n)$ são obtidas a partir das amostras das fontes $\{a_i(k)\}_{k \leq n}$ de acordo com a relação

$$u_\ell(n) = \sum_{i=1}^{N_u} h_{\ell,i}(0) a_i(n) + \sum_{i=1}^{N_u} h_{\ell,i}(1) a_i(n-1) + \dots + \sum_{i=1}^{N_u} h_{\ell,i}(N-1) a_i(n-N+1),$$

para $\ell = 1, 2, \dots, L$, em que $h_{\ell,i}(k) \in \mathbb{C}$ são os coeficientes do sistema misturador para compor a ℓ -ésima mistura a partir da i -ésima fonte com atraso k . Reduzindo-se com um somatório a expressão obtida, chega-se a

$$u_\ell(n) = \sum_{k=0}^{N-1} \sum_{i=1}^{N_u} h_{\ell,i}(k) a_i(n-k) \triangleq \sum_{i=1}^{N_u} h_{\ell,i}(k) * a_i(k), \quad (2.15)$$

em que $*$ denota a operação de convolução de tempo discreto. Uma representação ainda mais geral, em que não se especificam os instantes de tempo nos quais as fontes são tomadas, é dada por

$$u_\ell(n) = \sum_{k \in \mathbb{Z}} \sum_{i=1}^{N_u} h_{\ell,i}(k) a_i(n-k), \quad (2.16)$$

para $\ell = 1, 2, \dots, L$. Por conveniência, define-se a matriz de coeficientes $\mathbf{H}(k) \in \mathbb{C}^{L \times N_u}$ do sistema misturador, para o k -ésimo atraso, como

$$\mathbf{H}(k) = \begin{bmatrix} h_{1,1}(k) & h_{1,2}(k) & \cdots & h_{1,N_u}(k) \\ h_{2,1}(k) & h_{2,2}(k) & \cdots & h_{2,N_u}(k) \\ \vdots & \vdots & \ddots & \vdots \\ h_{L,1}(k) & h_{L,2}(k) & \cdots & h_{L,N_u}(k) \end{bmatrix}, \quad (2.17)$$

com $k \in \mathbb{Z}$. O conjunto de matrizes $\{\mathbf{H}(k)\}_{k \in \mathbb{Z}}$ corresponde à resposta ao pulso unitário do sistema misturador convolutivo (CASTELLA; CHEVREUIL; PESQUET, 2010). Pode-se, então, reescrever (2.16) vetorialmente como

$$\mathbf{u}(n) = \sum_{k \in \mathbb{Z}} \mathbf{H}(k) \mathbf{a}(n-k) \triangleq \mathbf{H}(n) * \mathbf{a}(n), \quad (2.18)$$

sendo os vetores $\mathbf{u}(n)$ e $\mathbf{a}(n)$ dados respectivamente pelas Equações (2.2) e (2.1). Nota-se que, no desenvolvimento feito até aqui, o sistema misturador considerado é linear e invariante no tempo. Para facilitar a análise desse sistema, pode-se considerar o seu comportamento no domínio transformado z . A transformada z da resposta ao pulso unitário do sistema misturador é definida como (CASTELLA; CHEVREUIL; PESQUET, 2010)

$$\mathbf{H}[z] \triangleq \sum_{k \in \mathbb{Z}} \mathbf{H}(k) z^{-k}.$$

A matriz $\mathbf{H}[z] \in \mathbb{C}[z]^{L \times N_u}$ é denominada matriz de sistema e possui, como entradas, polinômios em z . O símbolo $\mathbb{C}[z]$ denota o conjunto de polinômios com potências inteiras em z e com coeficientes complexos. A matriz $\mathbf{H}[z]$ representa, de forma compacta, o conjunto de matrizes $\{\mathbf{H}(k)\}_{k \in \mathbb{Z}}$. Adicionalmente, as transformadas z dos vetores $\mathbf{u}(n)$ e $\mathbf{a}(n)$ são dadas, respectivamente, por

$$\mathbf{u}[z] \triangleq \sum_{k \in \mathbb{Z}} \mathbf{u}(k) z^{-k} \in \mathbb{C}[z]^L \quad \text{e} \quad \mathbf{a}[z] \triangleq \sum_{k \in \mathbb{Z}} \mathbf{a}(k) z^{-k} \in \mathbb{C}[z]^{N_u}.$$

Os vetores $\mathbf{u}[z]$ e $\mathbf{a}[z]$ são vetores de polinômios em z e representam, de forma compacta, os respectivos conjuntos de vetores $\{\mathbf{u}(k)\}_{k \in \mathbb{Z}}$ e $\{\mathbf{a}(k)\}_{k \in \mathbb{Z}}$. Utilizando-se o fato de que a convolução no domínio do tempo representa o produto das transformadas no domínio z , pode-se, por fim, aplicar a transformada z em (2.18), o que resulta

$$\mathbf{u}[z] = \mathbf{H}[z] \mathbf{a}[z]. \quad (2.19)$$

Na representação de sistemas MIMO, costuma-se utilizar uma notação alternativa àquela da Equação (2.19), dada por (REGALIA, 1995)

$$\mathbf{u}(n) = \mathbf{H}[z] \mathbf{a}(n). \quad (2.20)$$

Essa notação indica que o vetor $\mathbf{u}(n)$ é a saída, no instante n , de um sistema linear e invariante no tempo com matriz de sistema $\mathbf{H}[z]$, quando excitado pelo conjunto de vetores $\{\mathbf{a}(k)\}_{k \leq n}$.

Quando o sistema misturador é linear e convolutivo, costuma-se adotar um sistema separador que também o seja (CASTELLA; CHEVREUIL; PESQUET, 2010). Desse modo, a relação de separação para o caso instantâneo (2.11) pode ser generalizada como

$$\mathbf{y}(n) = \mathbf{W}[z] \mathbf{u}(n), \quad (2.21)$$

em que $\mathbf{y}(n)$ é o vetor de fontes estimadas dado pela Equação (2.5) e $\mathbf{W}[z] \in \mathbb{C}[z]^{N_u \times L}$ é a matriz de sistema do sistema separador.

Em geral, a solução do problema de BSS para misturas convolutivas é mais complicada do que para o caso de misturas instantâneas. Dado um sistema misturador MIMO linear e invariante no tempo, nem sempre o respectivo sistema inverso existe (CASTELLA; CHEVREUIL; PESQUET, 2010). Mesmo no caso igualmente determinado, e em que o sistema misturador é estável, causal e possui resposta ao pulso unitário finita (FIR, do inglês, *finite impulse response*), a existência de um sistema inverso está condicionada à irreducibilidade ou invertibilidade de $\mathbf{H}[z]$ (CASTELLA; CHEVREUIL; PESQUET, 2010).

Para o caso de misturas convolutivas igualmente determinado, se $\mathbf{H}[z]$ possui uma inversa à esquerda $\mathbf{G}[z]$, ou seja,

$$\mathbf{G}[z] \mathbf{H}[z] = \mathbf{I}_{N_u}, \quad (2.22)$$

em que $\mathbf{I}_{N_u} \in \mathbb{R}^{N_u \times N_u}$ denota a matriz identidade, então a escolha de

$$\mathbf{W}[z] = \mathbf{G}[z] \quad (2.23)$$

resolveria o problema de BSS. Nota-se, de forma intuitiva, que o número de incógnitas a serem determinadas em relação ao caso instantâneo é maior, visto que o sistema misturador é descrito por um conjunto de matrizes (equivalentemente, por uma matriz de polinômios) e não só por uma única matriz, como no caso instantâneo. Portanto, para se resolver o problema de BSS para misturas convolutivas, além de se considerar a hipótese de independência espacial (*i.e.*, entre fontes), é usual considerar também que cada fonte é independente e identicamente distribuída (i.i.d.). Quando é feita a suposição de que as fontes são i.i.d., ou seja, considera-se que elas apresentam independência temporal, adicionam-se restrições ou equações que permitem resolver o problema de BSS. No entanto, para que o problema de BSS tenha solução no caso de misturas convolutivas, não é suficiente considerar independência espacial e temporal das fontes. Assim como no caso de misturas instantâneas, ainda é preciso considerar, por exemplo, a não gaussianidade das fontes (CASTELLA; CHEVREUIL; PESQUET, 2010).

Por fim, cabe notar que o resultado matemático sobre separabilidade para misturas instantâneas pode ser estendido para misturas convolutivas (YELLIN; WEINSTEIN, 1994). Essa extensão diz que as fontes também podem ser recuperadas para misturas convolutivas ao se impor a independência estatística na saída do sistema separador. No entanto, as

ambiguidades de recuperação devido ao problema cego podem ser mais graves para as misturas convolutivas do que para as misturas instantâneas.

Em outras palavras, se $\mathbf{u}(n) = \mathbf{H}[z]\mathbf{a}(n)$, em que $\mathbf{H}[z]$ é uma matriz invertível cujas entradas são polinômios em z , e $\mathbf{a}(n)$ é um vetor de fontes estatisticamente independentes, sendo no máximo uma delas gaussiana, então $\mathbf{y}(n) = \mathbf{W}[z]\mathbf{u}(n)$ é um vetor aleatório com componentes mutuamente independentes se e somente se

$$\mathbf{W}[z] \mathbf{H}[z] = \mathbf{\Lambda}[z] \mathbf{P}, \quad (2.24)$$

em que $\mathbf{\Lambda}[z]$ é uma matriz diagonal de polinômios em z e \mathbf{P} é uma matriz de permutação (YELLIN; WEINSTEIN, 1994). Assim como para misturas instantâneas, a recuperação também se dá com ambiguidade de permutação para misturas convolutivas. Essa ambiguidade é representada pela matriz \mathbf{P} . No entanto, a matriz $\mathbf{\Lambda}[z]$ indica que a recuperação de independência permite separar as fontes a menos de uma filtragem linear no tempo. Essa limitação existente no caso convolutivo é mais grave do que a correspondente ambiguidade de escala para o caso instantâneo, e pode ser indesejável em aplicações práticas.

2.3 FORMULAÇÃO DO PROBLEMA DE DESCONVOLUÇÃO MULTIUSUÁRIO

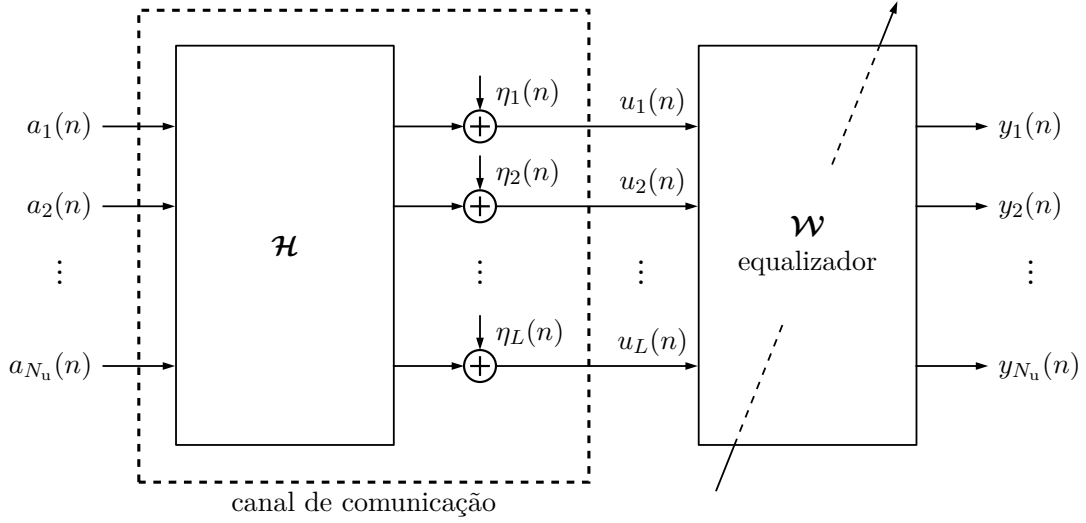
A desconvolução cega de canais de comunicação MIMO pode ser compreendida como um problema particular de separação cega de fontes. No caso do problema de desconvolução, além de o modelo adotado para o sistema misturador ser linear e convolutivo, tem-se conhecimento considerável das estatísticas das fontes porque, na grande maioria das aplicações, o alfabeto transmitido é finito.

O modelo equivalente de tempo discreto de um sistema de comunicação MIMO com N_u fontes e $L \geq N_u$ sensores (*e.g.*, antenas) é mostrado na Figura 2. As sequências correspondentes às fontes são denotadas por $a_i(n)$, com $i = 1, 2, \dots, N_u$. Em geral, considera-se que as fontes são independentes e identicamente distribuídas (i.i.d.), não gaussianas e possuem média nula. Costuma-se assumir, também, que as fontes apresentam as mesmas estatísticas e são independentes umas das outras (HAYKIN, 2000b; DING; LI, 2001).

As fontes são transmitidas por canais de comunicação que causam distorção nos sinais, a partir da adição de ruído e da atuação de interferências cocanal e intersimbólica. O canal que associa a i -ésima fonte ao ℓ -ésimo sensor é modelado por um filtro FIR com K_c coeficientes. Os coeficientes de cada canal são representados por polinômios $H_{i,\ell}(z)$. O ruído gaussiano branco aditivo, nas saídas do canal, é denotado por $\eta_\ell(n)$, com $\ell = 1, 2, \dots, L$. As saídas dos L sensores são processadas por N_u equalizadores FIR espaço-temporais, cada um com diversidade temporal K_t e $M = LK_t$ coeficientes. A saída do i -ésimo equalizador pode ser escrita como

$$y_i(n) = \mathbf{u}^\top(n) \mathbf{w}_i, \quad (2.25)$$

Figura 2 – Modelo equivalente de tempo discreto de um sistema de comunicação MIMO.



Fonte: autoria própria.

com

$$\mathbf{u}(n) = [\check{\mathbf{u}}_1^T(n) \quad \check{\mathbf{u}}_2^T(n) \quad \cdots \quad \check{\mathbf{u}}_L^T(n)]^T,$$

$$\check{\mathbf{u}}_\ell(n) = [u_\ell(n) \quad u_\ell(n-1) \quad \cdots \quad u_\ell(n-K_t+1)]^T,$$

para $\ell = 1, 2, \dots, L$, sendo \mathbf{w}_i o vetor de coeficientes do i -ésimo equalizador. Segundo o modelo equivalente apresentado na Figura 2, o vetor regressor $\mathbf{u}(n)$ pode ser escrito como

$$\mathbf{u}(n) = \bar{\mathbf{H}}^T \mathbf{a}(n) + \boldsymbol{\eta}(n), \quad (2.26)$$

em que $\bar{\mathbf{H}}$ é a matriz de convolução do canal MIMO, dada por

$$\bar{\mathbf{H}}^T = \begin{bmatrix} \mathbf{H}_{1,1}^T & \mathbf{H}_{2,1}^T & \cdots & \mathbf{H}_{N_u,1}^T \\ \mathbf{H}_{1,2}^T & \mathbf{H}_{2,2}^T & \cdots & \mathbf{H}_{N_u,2}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{1,L}^T & \mathbf{H}_{2,L}^T & \cdots & \mathbf{H}_{N_u,L}^T \end{bmatrix}, \quad (2.27)$$

sendo que $\mathbf{H}_{i,\ell}$ é a matriz de convolução correspondente ao canal que associa a i -ésima fonte ao ℓ -ésimo sensor, $\mathbf{a}(n)$ é o vetor de linhas de atraso de fontes concatenadas, escrito como

$$\mathbf{a}(n) = [\check{\mathbf{a}}_1^T(n) \quad \check{\mathbf{a}}_2^T(n) \quad \cdots \quad \check{\mathbf{a}}_{N_u}^T(n)]^T,$$

$$\check{\mathbf{a}}_i(n) = [a_i(n) \quad a_i(n-1) \quad \cdots \quad a_i(n-K_t-K_c+2)]^T,$$

com $i = 1, 2, \dots, N_u$, e $\boldsymbol{\eta}(n)$ corresponde ao vetor de linhas de atraso de ruído nas saídas do canal, *i.e.*,

$$\boldsymbol{\eta}(n) = [\check{\boldsymbol{\eta}}_1^T(n) \quad \check{\boldsymbol{\eta}}_2^T(n) \quad \cdots \quad \check{\boldsymbol{\eta}}_L^T(n)]^T,$$

$$\check{\boldsymbol{\eta}}_\ell(n) = [\eta_\ell(n) \quad \eta_\ell(n-1) \quad \cdots \quad \eta_\ell(n-K_t+1)]^T,$$

com $\ell = 1, 2, \dots, L$.

A partir dos sinais de mistura recebidos e das propriedades estatísticas previamente conhecidas dos sinais transmitidos, o papel dos equalizadores consiste em reduzir ao máximo as interferências intersimbólica e cocanal introduzidas pelos canais. Em seu funcionamento adequado, o equalizador deve recuperar todas as N_u fontes distintas. Devido às limitações inerentes do problema cego sendo considerado, o equalizador é capaz de recuperar as fontes a menos de um atraso temporal e de uma rotação complexa (DING; LI, 2001). Em outras palavras, sob correto funcionamento, o equalizador deve obter estimativas de

$$e^{j\theta_i} a_i(n - \tau_i), \quad (2.28)$$

para $i = 1, 2, \dots, N_u$, em que $j^2 = -1$, θ_i é um ângulo de rotação no plano complexo e τ_i é um inteiro positivo correspondente ao atraso da sequência recuperada em relação à sequência transmitida. Além disso, outra limitação decorrente do problema cego é a ambiguidade de permutação, *i.e.*, pode haver uma eventual mudança na ordenação dos sinais recuperados em relação aos sinais transmitidos.

Nesse contexto, define-se que ocorre equalização perfeita, *i.e.*, condição de *zero-forcing*, quando a resposta combinada $\mathbf{s}_{zf,i}$ da cascata do canal de comunicação com o i -ésimo equalizador puder ser escrita como (DING; LI, 2001)

$$\mathbf{s}_{zf,i} = \left[\underbrace{0 \ \dots \ 0}_{\tau_i \text{ zeros}} \ e^{j\theta_i} \ 0 \ \dots \ 0 \right], \quad (2.29)$$

para $i = 1, 2, \dots, N_u$. Observa-se que, no caso multiusuário, o seguinte par de condições é suficiente para que haja equalização perfeita (PAPADIAS; PAULRAJ, 1997a):

Condição 2.3–A. O número de coeficientes de cada filtro do equalizador multiusuário deve atender à relação

$$M \geq \frac{N_u(K_c - 1)}{1 - N_u/L}. \quad (2.30)$$

Condição 2.3–B. Os canais com coeficientes dados por polinômios $H_i(z)$, com $i = 1, 2, \dots, N_u$, definidos como

$$H_i(z) = \sum_{\ell=1}^{N_u} H_{\ell,i}(z^{N_u}) z^{\ell-1}, \quad (2.31)$$

não devem ter zeros em comum.

Quanto à Condição 2.3–A, observa-se que o denominador de (2.30) vale zero quando se tem o mesmo número de fontes e de sensores, *i.e.*, $N_u = L$. Nesse caso igualmente determinado, a equalização perfeita com filtros FIR é impossível, já que seriam necessários infinitos coeficientes para atingi-la.

2.4 CONCLUSÃO

Nesse capítulo, os problemas de separação cega de fontes e desconvolução cega multiusuário foram apresentados. A formulação do problema de BSS foi feita de forma genérica na Seção 2.1, sem se assumir conhecimento prévio sobre as fontes ou sobre o sistema misturador. Na Seção 2.2, o problema de BSS foi particularizado para dois tipos de sistemas misturadores lineares, *i.e.*, instantâneo e convolutivo, relevantes em diversas aplicações práticas. Essa particularização permitiu compreender as principais hipóteses usualmente adotadas na solução do problema para cada tipo de mistura considerado. Em seguida, na Seção 2.3, foi feita a formulação do problema de desconvolução cega multiusuário. Em desconvolução multiusuário, dispõe-se de conhecimento completo das estatísticas das fontes, visto que o alfabeto transmitido é finito e conhecido de antemão pelo receptor. Em BSS, por outro lado, muitas vezes não há um alfabeto limitado e conhecido a ser transmitido, como acontece em diversas aplicações de áudio, imagens médicas, sinais neuronais, e assim por diante (VINCENT et al., 2014; ADALI; ANDERSON; FU, 2014; BROWN; YAMADA; SEJNOWSKI, 2001).

Nota-se que o conhecimento *a priori* que se tem sobre as estatísticas das fontes e sobre a estrutura do sistema misturador possui grande influência na forma como o problema de BSS é abordado. Dada a abrangência desse problema e as inúmeras frentes de estudo existentes nessa área, escolheu-se, nesse trabalho, voltar o enfoque à desconvolução cega multiusuário para possibilitar, posteriormente, a compreensão mais aprofundada das soluções existentes em BSS. Além disso, o estudo de desconvolução cega multiusuário se justifica pois as soluções adaptativas existentes apresentam problemas de convergência e instabilidade numérica, e ainda são estudadas na área de Processamento Estatístico de Sinais (SILVA; MIRANDA; LICCIARDI, 2004; SHAH; ABED-MERAIM; AL-NAFFOURI, 2015).

3 O ALGORITMO DE SHALVI-WEINSTEIN MULTIUSUÁRIO

Com o objetivo de incentivar o entendimento dos princípios teóricos de separação cega de fontes, bem como a compreensão de suas soluções, neste capítulo são apresentadas soluções adaptativas para o problema de desconvolução cega multiusuário. Essas soluções são obtidas inspirando-se nas soluções existentes para o caso monousuário (SHALVI; WEINSTEIN, 1993; YANG; WERNER; DUMONT, 2002; MIRANDA; SILVA; NASCIMENTO, 2008).

A seguir, descreve-se a organização do capítulo. Na Seção 3.1, são feitas algumas considerações sobre soluções adaptativas já existentes para o problema de desconvolução cega multiusuário. Na Seção 3.2, partindo-se de uma formulação do problema de mínimos quadrados, é feita a dedução de uma solução adaptativa para o problema de equalização cega multiusuário: o algoritmo de Shalvi-Weinstein (SWA, do inglês, *Shalvi-Weinstein algorithm*). Esse algoritmo, por apresentar problemas de instabilidade numérica, é aprimorada na Seção 3.3. Nela, lança-se mão de um princípio de funcionamento denominado *dual-mode*, que evita uma das causas de instabilidade numérica do SWA. Na Seção 3.4, são obtidas versões numericamente robustas de algoritmos baseados no SWA, que evitam a divergência a partir da operação em *dual-mode* e de uma implementação com decomposição QR-MIMO. Posteriormente, na Seção 3.5, são apresentados resultados de simulação que evidenciam as propriedades de estabilidade numérica das versões dos algoritmos propostos, mesmo sob cenários adversos. Finalmente, na Seção 3.6, são feitas as considerações finais sobre o capítulo.

3.1 SOLUÇÕES EXISTENTES

A fim de se melhorar o uso eficiente da banda do canal e permitir a recuperação das fontes em tempo real, soluções adaptativas baseadas em equalizadores espaço-temporais têm sido propostas na literatura para o problema de desconvolução cega multiusuário. Muitas dessas soluções são baseadas no trabalho de Papadias e Paulraj (PAPADIAS; PAULRAJ, 1997a), em que o algoritmo de módulo constante (CMA, do inglês, *constant modulus algorithm*) (GODARD, 1980; TREICHLER; AGEE, 1983) foi estendido para o caso multiusuário, levando em conta a penalização da correlação cruzada entre as sequências de saída. Como mostrado em (PAPADIAS; PAULRAJ, 1997a), essa penalização permite a recuperação conjunta e simultânea de todos os sinais transmitidos. Além do CMA, outro algoritmo relevante para equalização cega de sistemas de comunicação SISO (do inglês, *single-input and single-output*) foi proposto originalmente por Shalvi e Weinstein em (SHALVI; WEINSTEIN, 1993). Trata-se do algoritmo de Shalvi-Weinstein (SWA).

Sob certas condições, como mostrado em (REGALIA, 1999), o SWA e o CMA buscam alcançar os mesmos pontos estacionários de funções custo distintas. Diante disso, o SWA pode ser interpretado como um algoritmo de módulo constante que inclui, na sua equação de adaptação, uma aproximação para a matriz Hessiana. Dessa forma, ele pode ser considerado como um algoritmo do tipo quase-Newton (SILVA; MIRANDA, 2004). Devido à multimodalidade das funções custo a partir das quais o CMA e o SWA são obtidos, certas situações podem levar ambos os algoritmos a divergir (*i.e.*, a norma do vetor de coeficientes tender ao infinito) ou convergir para mínimos locais indesejados. Isso pode ocorrer no caso de, por exemplo, uma escolha inadequada de fator de esquecimento (respectivamente, passo de adaptação) para o SWA (respectivamente, CMA), uma inicialização distante da solução de *zero-forcing*, ou uma relação sinal-ruído (SNR, do inglês, *signal-to-noise ratio*) baixa. Nesse contexto e para canais SISO, mostrou-se em (MIRANDA; SILVA; NASCIMENTO, 2008) que a divergência do SWA pode ser causada por inconsistências na estimativa não linear dos sinais transmitidos ou pela perda de robustez numérica na atualização da inversa da matriz de autocorrelação. Para lidar com o primeiro problema de instabilidade, o trabalho de (MIRANDA; SILVA; NASCIMENTO, 2008) introduziu um funcionamento duplo para o SWA, conhecido como *dual-mode*. Para lidar com o segundo problema de instabilidade e reduzir o custo computacional do algoritmo, ele também foi implementado com uma estrutura em treliça.

No que diz respeito a algoritmos de equalização para canais MIMO, uma versão multiusuário do SWA foi apresentada em (SILVA; MIRANDA, 2005). Nessa versão, levou-se em conta a penalização da correlação cruzada das sequências de saída, inicialmente proposta para o CMA multiusuário em (PAPADIAS; PAULRAJ, 1997a). No entanto, o algoritmo proposto em (SILVA; MIRANDA, 2005) sofre com os mesmos problemas de divergência do SWA para o caso SISO, o que motiva a busca por soluções numericamente estáveis, feita no decorrer deste capítulo. Em (SILVA; MIRANDA, 2005), o SWA multiusuário foi obtido a partir de uma abordagem envolvendo o gradiente estocástico. A seguir, é apresentada uma dedução alternativa desse algoritmo, agora sob uma perspectiva de problema de mínimos quadrados aplicada ao caso MIMO.

3.2 DEDUÇÃO DO SWA MULTIUSUÁRIO

O algoritmo de Shalvi-Weinstein foi proposto em (SHALVI; WEINSTEIN, 1993) no contexto de equalização cega de canais de comunicação SISO, a partir da minimização da razão entre os cumulantes de quarta e segunda ordem da saída do equalizador. Esse algoritmo foi deduzido em (MIRANDA; SILVA; NASCIMENTO, 2008) com base na função custo de módulo constante (GODARD, 1980), mas sob uma perspectiva de mínimos quadrados. Essa dedução permitiu evidenciar as aproximações nas estimativas das estatísticas de ordem superior, necessárias para a obtenção do SWA.

Além disso, uma versão multiusuário do SWA foi introduzida em (SILVA; MIRANDA, 2004), levando em conta a penalização das correlações cruzadas das sequências de saída (PAPADIAS; PAULRAJ, 1997a), mas sem uma dedução formal. Também para o caso SISO, (YANG; WERNER; DUMONT, 2002) apresentou um algoritmo que evita rotações complexas que não sejam múltiplas de 90° na saída de equalizadores cegos, conhecido como algoritmo multimódulo (MMA, do inglês, *multimodulus algorithm*). Esse tipo de algoritmo é interessante pois elimina a necessidade de se utilizar um estágio adicional de rotação na saída do equalizador, sem o qual a taxa de erros de símbolo no receptor pode aumentar significativamente.

Inspirando-se no fato de que o SWA pode ser interpretado como um algoritmo de módulo constante (REGALIA, 1999), a seguir, sua dedução determinista para o caso SISO, apresentada em (MIRANDA; SILVA; NASCIMENTO, 2008), é estendida para o caso MIMO, de modo a se obter uma versão multimódulo do algoritmo proposto em (SILVA; MIRANDA, 2004).

A fim de se equalizar o canal MIMO com um algoritmo multimódulo, busca-se minimizar

- A dispersão das partes real e imaginária da saída do equalizador, separadamente, de forma análoga ao algoritmo multimódulo de (YANG; WERNER; DUMONT, 2002) para o caso SISO.
- A correlação cruzada entre as saídas dos equalizadores, inspirando-se na versão multiusuário introduzida em (PAPADIAS; PAULRAJ, 1997a), mas sem fazer uma separação das partes real e imaginária. No caso SISO, essa separação é considerada no MMA para resolver o problema de ambiguidade de fase. No entanto, uma rotação complexa de múltiplos de 90° pode ainda ocorrer em algoritmos do tipo MMA, como mostrado em (YUAN; TSAI, 2005).

No caso MIMO, uma rotação de 90° também pode ocorrer, e esse problema se torna crítico quando se penaliza a correlação cruzada entre partes reais e a correlação cruzada entre partes imaginárias separadamente. Por exemplo, se dois equalizadores recuperarem a mesma fonte, mas uma rotacionada de 90° em relação à outra, o resultado de uma penalização “separada” seria muito próximo de zero. Isso ocorre pois uma rotação de 90° troca as partes real e imaginária da fonte transmitida, e para uma mesma fonte com simetria circular, as partes real e imaginária são estatisticamente independentes. Portanto, nessa situação, o equalizador MIMO não seria capaz de recuperar duas fontes distintas.

Desse modo, define-se a função custo multimódulo para o i -ésimo equalizador, com

$i = 1, 2, \dots, N_u$, como

$$\widehat{J}_i(n) \triangleq (1 - \lambda) \sum_{\ell=0}^n \lambda^{n-\ell} \left\{ [\mathcal{R} - y_{i,R}^2(\ell, n)]^2 + [\mathcal{R} - y_{i,I}^2(\ell, n)]^2 \right\} + \varrho \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} |c_{i,m}(n, k)|^2, \quad (3.1)$$

em que $y_{i,R}(\ell, n)$ e $y_{i,I}(\ell, n)$ são, respectivamente, as partes real e imaginária de $y_i(\ell, n) \triangleq \mathbf{u}^\top(\ell) \mathbf{w}_i(n)$, ou seja,

$$y_i(\ell, n) = y_{i,R}(\ell, n) + jy_{i,I}(\ell, n) \triangleq \mathbf{u}^\top(\ell) \mathbf{w}_i(n), \quad (3.2)$$

$\mathcal{R} = E[a_{i,R}^4(n)] / E[a_{i,R}^2(n)] = E[a_{i,I}^4(n)] / E[a_{i,I}^2(n)]$ é a constante de dispersão, $E[\cdot]$ denota a esperança matemática, $a_{i,R}(n)$ e $a_{i,I}(n)$ são, respectivamente, as partes real e imaginária de $a_i(n)$, ou seja

$$a_i(n) = a_{i,R}(n) + ja_{i,I}(n), \quad (3.3)$$

$0 \ll \lambda < 1$ é um fator de esquecimento,

$$c_{i,m}(n, k) = (1 - \lambda) \sum_{p=0}^n \lambda^{n-p} y_i(p, n) y_m^*(p - k, n) \quad (3.4)$$

é a correlação cruzada determinista entre as saídas do i -ésimo e do m -ésimo equalizadores, $(\cdot)^*$ denota o complexo conjugado, ϱ é uma constante positiva e k_1 é um inteiro associado ao atraso de espalhamento. Cabe notar que, no método de mínimos quadrados, consideram-se as observações do instante 0 até o instante n para se estimar o vetor de coeficientes $\mathbf{w}_i(n)$.

Segundo a convenção para o Cálculo de Wirtinger multivariável apresentada no Apêndice A, o cogradiente de $\widehat{J}_i(n)$ em relação a $\mathbf{w}_i(n)$, tratando $\mathbf{w}_i^H(n)$ como uma constante, em que $(\cdot)^H$ denota o transposto conjugado, resulta

$$\begin{aligned} \frac{\partial \widehat{J}_i(n)}{\partial \mathbf{w}_i(n)} &= -2(1 - \lambda) \sum_{\ell=0}^n \lambda^{n-\ell} \left\{ [\mathcal{R} - y_{i,R}^2(\ell, n)] y_{i,R}(\ell, n) - j [\mathcal{R} - y_{i,I}^2(\ell, n)] y_{i,I}(\ell, n) \right\} \mathbf{u}^\top(\ell) \\ &+ \varrho(1 - \lambda) \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} c_{i,m}^*(n, k) \sum_{p=0}^n \lambda^{n-p} y_m^*(p - k, n) \mathbf{u}^\top(p). \end{aligned} \quad (3.5)$$

O gradiente complexo de $\widehat{J}_i(n)$ em relação a $\mathbf{w}_i(n)$, de acordo com a definição (A.17), é dado por

$$\nabla_{\mathbf{w}_i(n)} \widehat{J}_i(n) \triangleq \left[\frac{\partial \widehat{J}_i(n)}{\partial \mathbf{w}_i(n)} \right]^H. \quad (3.6)$$

Da definição estabelecida, resulta

$$\begin{aligned} \nabla_{\mathbf{w}_i(n)} \widehat{J}_i(n) &= -2(1-\lambda) \sum_{\ell=0}^n \lambda^{n-\ell} \left\{ [\mathcal{R} - y_{i,R}^2(\ell, n)] y_{i,R}(\ell, n) + j [\mathcal{R} - y_{i,I}^2(\ell, n)] y_{i,I}(\ell, n) \right\} \mathbf{u}^*(\ell) \\ &\quad + \varrho(1-\lambda) \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} c_{i,m}(n, k) \sum_{p=0}^n \lambda^{n-p} y_m(p-k, n) \mathbf{u}^*(p). \end{aligned} \quad (3.7)$$

Fazendo-se $\nabla_{\mathbf{w}_i(n)} \widehat{J}_i(n) = \mathbf{0}$, chega-se à igualdade

$$\mathcal{R} \widehat{\mathbf{R}}(n) \mathbf{w}_i(n) = \mathbf{p}_i^y(n) + \mathbf{p}_i^c(n), \quad (3.8)$$

em que

$$\widehat{\mathbf{R}}(n) \triangleq \sum_{\ell=0}^n \lambda^{n-\ell} \mathbf{u}^*(\ell) \mathbf{u}^\top(\ell), \quad (3.9)$$

$$\mathbf{p}_i^y(n) \triangleq \sum_{\ell=0}^n \lambda^{n-\ell} \left[y_{i,R}^3(\ell, n) + j y_{i,I}^3(\ell, n) \right] \mathbf{u}^*(\ell), \quad (3.10)$$

$$\mathbf{p}_i^c(n) \triangleq \frac{\varrho}{2} \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} c_{i,m}(n, k) \sum_{p=0}^n \lambda^{n-p} y_m(p-k, n) \mathbf{u}^*(p). \quad (3.11)$$

A fim de se obter uma equação de atualização para $\mathbf{w}_i(n)$ a partir de (3.8), a seguir são encontradas relações de recorrência para $\widehat{\mathbf{R}}(n) \mathbf{w}_i(n)$, $\mathbf{p}_i^y(n)$ e $\mathbf{p}_i^c(n)$.

- Inicialmente, para se encontrar uma recursão para $\widehat{\mathbf{R}}(n) \mathbf{w}_i(n)$, define-se a diferença

$$\Delta \mathbf{w}_i(n) \triangleq \mathbf{w}_i(n) - \mathbf{w}_i(n-1). \quad (3.12)$$

Partindo-se de (3.9), pode-se escrever

$$\widehat{\mathbf{R}}(n) = \sum_{\ell=0}^{n-1} \lambda^{n-\ell} \mathbf{u}^*(\ell) \mathbf{u}^\top(\ell) + \mathbf{u}^*(n) \mathbf{u}^\top(n). \quad (3.13)$$

Novamente a partir de (3.9), vale

$$\sum_{\ell=0}^{n-1} \lambda^{n-\ell} \mathbf{u}^*(\ell) \mathbf{u}^\top(\ell) = \lambda \widehat{\mathbf{R}}(n-1). \quad (3.14)$$

Substituindo-se o primeiro termo de (3.13) de acordo com a igualdade (3.14), chega-se a uma recursão para a atualização da matriz $\widehat{\mathbf{R}}(n)$:

$$\widehat{\mathbf{R}}(n) = \lambda \widehat{\mathbf{R}}(n-1) + \mathbf{u}^*(n) \mathbf{u}^\top(n). \quad (3.15)$$

Utilizando (3.12) e (3.15), o vetor $\widehat{\mathbf{R}}(n) \mathbf{w}_i(n)$ pode ser reescrito como

$$\widehat{\mathbf{R}}(n) \mathbf{w}_i(n) = \lambda \widehat{\mathbf{R}}(n-1) \mathbf{w}_i(n-1) + \widehat{\mathbf{R}}(n) \Delta \mathbf{w}_i(n) + y_i(n) \mathbf{u}^*(n). \quad (3.16)$$

- Devido às não linearidades em termos de $y(n)$ presentes em $\mathbf{p}_i^y(n)$, obter uma recursão para o seu cálculo de forma exata não é evidente. A fim de se obter uma recursão de forma aproximada, definem-se

$$\Delta \{y_{i,R}^3(\ell, n)\} \triangleq y_{i,R}^3(\ell, n) - y_{i,R}^3(\ell, n-1), \quad (3.17)$$

$$\Delta \{y_{i,I}^3(\ell, n)\} \triangleq y_{i,I}^3(\ell, n) - y_{i,I}^3(\ell, n-1). \quad (3.18)$$

Com as definições (3.17) e (3.18), o vetor $\mathbf{p}_i^y(n)$, definido em (3.10), pode ser reescrito como

$$\begin{aligned} \mathbf{p}_i^y(n) &= \sum_{\ell=0}^n \lambda^{n-\ell} [y_{i,R}^3(\ell, n-1) + jy_{i,I}^3(\ell, n-1)] \mathbf{u}^*(\ell) \\ &+ \sum_{\ell=0}^n \lambda^{n-\ell} [\Delta \{y_{i,R}^3(\ell, n)\} + j\Delta \{y_{i,I}^3(\ell, n)\}] \mathbf{u}^*(\ell). \end{aligned} \quad (3.19)$$

Adicionalmente, considera-se a seguinte hipótese simplificadora.

Hipótese 3.2–A. Sejam os cumulantes de ordem 4 definidos como

$$\begin{aligned} \boldsymbol{\kappa}_{i,R} &\triangleq \mathbb{E} [y_{i,R}^3(n) \mathbf{u}^*(n)] - 3 \mathbb{E} [y_{i,R}^2(n)] \mathbb{E} [y_{i,R}(n) \mathbf{u}^*(n)], \\ \boldsymbol{\kappa}_{i,I} &\triangleq \mathbb{E} [y_{i,I}^3(n) \mathbf{u}^*(n)] - 3 \mathbb{E} [y_{i,I}^2(n)] \mathbb{E} [y_{i,I}(n) \mathbf{u}^*(n)]. \end{aligned}$$

A soma

$$\boldsymbol{\kappa}_i \triangleq \boldsymbol{\kappa}_{i,R} + j\boldsymbol{\kappa}_{i,I} \quad (3.20)$$

pode ser aproximada, no instante n , por

$$\widehat{\boldsymbol{\kappa}}_i(n, n_0) = \sum_{\ell=0}^n \lambda^{n-\ell} [y_{i,R}^3(\ell, n_0) + jy_{i,I}^3(\ell, n_0)] \mathbf{u}^*(\ell) - 3 \mathbb{E} [a_{i,R}^2(n)] \widehat{\mathbf{R}}(n) \mathbf{w}_i(n_0), \quad (3.21)$$

em que as esperanças matemáticas foram substituídas por médias no tempo, exceto pelas esperanças $\mathbb{E} [y_{i,R}^2(n)]$ e $\mathbb{E} [y_{i,I}^2(n)]$, que foram ambas substituídas por $\mathbb{E} [a_{i,R}^2(n)] = \mathbb{E} [a_{i,I}^2(n)]$ (SHALVI; WEINSTEIN, 1993; MIRANDA; SILVA; NASCIMENTO, 2008). Para λ suficientemente próximo de um e n suficientemente grande, $\widehat{\boldsymbol{\kappa}}_i(n, n)$, calculado com $\mathbf{w}_i(n)$, pode ser adequadamente aproximado por $\widehat{\boldsymbol{\kappa}}_i(n, n-1)$, calculado com $\mathbf{w}_i(n-1)$. Nesse caso, tem-se

$$\boldsymbol{\kappa}_i \approx \widehat{\boldsymbol{\kappa}}_i(n, n) \approx \widehat{\boldsymbol{\kappa}}_i(n, n-1). \quad (3.22)$$

A partir de (3.22) (Hipótese 3.2–A), e das definições (3.17) e (3.18), obtém-se a aproximação

$$\sum_{\ell=0}^n \lambda^{n-\ell} [\Delta \{y_{i,R}^3(\ell, n)\} + j\Delta \{y_{i,I}^3(\ell, n)\}] \mathbf{u}^*(\ell) \approx 3 \mathbb{E} [a_{i,R}^2(n)] \widehat{\mathbf{R}}(n) \Delta \mathbf{w}_i(n). \quad (3.23)$$

Expressando o primeiro termo do lado direito de (3.19) por uma soma com ℓ indo de 0 até $n - 1$, lembrando que denota-se $y_{i,R}^3(n, n - 1) = y_{i,R}^3(n)$ (analogamente para a parte imaginária correspondente), utilizando (3.10) para o instante $n - 1$ e (3.23), conclui-se que vale a aproximação

$$\mathbf{p}_i^y(n) \approx \widehat{\mathbf{p}}_i^y(n) = \lambda \widehat{\mathbf{p}}_i^y(n - 1) + [y_{i,R}^3(n) + jy_{i,I}^3(n)] \mathbf{u}^*(n) + 3 \text{E} [a_{i,R}^2(n)] \widehat{\mathbf{R}}(n) \Delta \mathbf{w}_i(n). \quad (3.24)$$

- Falta, ainda, encontrar uma recursão para $\mathbf{p}_i^c(n)$. Devido às não linearidades na saída dos equalizadores presentes nesse vetor, utiliza-se a seguinte hipótese simplificadora.

Hipótese 3.2–B. Considerando-se o termo $\mathbf{p}_i^c(n)$, cuja expressão é dada pela Equação (3.11), vale a seguinte aproximação

$$\mathbf{p}_i^c(n) \approx \widehat{\mathbf{p}}_i^c(n) = \lambda \widehat{\mathbf{p}}_i^c(n - 1) + \frac{\rho}{2} \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} \widehat{c}_{i,m}(n, k) y_m(n - k) \mathbf{u}^*(n), \quad (3.25)$$

sendo

$$\widehat{c}_{i,m}(n, k) = (1 - \lambda) \sum_{p=0}^n \lambda^{n-p} y_i(p) y_m^*(p - k) \quad (3.26)$$

uma estimativa da correlação cruzada $c_{i,m}(n, k)$ definida na Equação (3.4). Cabe notar que $y_i(p)$ e $y_m^*(p - k)$ são definidos com o vetor de coeficientes nos instantes $p - 1$ e $p - k - 1$, respectivamente.

Além disso, como as não linearidades presentes em $\mathbf{p}_i^c(n)$ são diferentes daquelas presentes em $\mathbf{p}_i^y(n)$, não é possível obter (3.25) ao se aplicar, diretamente à Equação (3.11), uma suposição similar à Hipótese 3.2–A. Tendo isso em vista e baseando-se no método de inovações (ver, *e.g.*, (KAILATH; SAYED; HASSIBI, 2000)), pode-se obter uma recorrência usual para $\widehat{\mathbf{p}}_i^c(n)$ ao se fazer a soma do estimador no instante anterior $\widehat{\mathbf{p}}_i^c(n - 1)$ com a observação adicional (uma função da inovação), que por sua vez é dada pela soma do lado direito de (3.25). Nota-se que essa soma aparece no MU-CMA (PAPADIAS; PAULRAJ, 1997a) e nos algoritmos de (LUO; CHAMBERS, 2002; SILVA; MIRANDA; LICCIARDI, 2004; SILVA; MIRANDA, 2005).

Finalmente, substituindo-se (3.16), (3.24), (3.25) e (3.26) em (3.8), e notando-se que

$$\mathcal{R} \widehat{\mathbf{R}}(n - 1) \mathbf{w}_i(n - 1) \approx \widehat{\mathbf{p}}_i^y(n - 1) + \widehat{\mathbf{p}}_i^c(n - 1), \quad (3.27)$$

obtém-se

$$\gamma \widehat{\mathbf{R}}(n) \Delta \mathbf{w}_i(n) \approx \bar{e}_i(n) \mathbf{u}^*(n), \quad (3.28)$$

sendo

$$\bar{e}_i(n) \triangleq e_i(n) - \epsilon_i(n), \quad (3.29)$$

$$e_i(n) \triangleq e_{i,R}(n) + j e_{i,I}(n) = [\mathcal{R} - y_{i,R}^2(n)] y_{i,R}(n) + j [\mathcal{R} - y_{i,I}^2(n)] y_{i,I}(n), \quad (3.30)$$

$$\epsilon_i(n) \triangleq \frac{\rho}{2} \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} \hat{c}_{i,m}(n, k) y_m(n - k), \quad (3.31)$$

$$\gamma \triangleq 3 \text{E} [a_{i,R}^2(n)] - \mathcal{R}. \quad (3.32)$$

Isolando-se $\Delta \mathbf{w}_i(n)$ em (3.28) e substituindo-se a definição de (3.12), chega-se à equação de atualização dos coeficientes do algoritmo de Shalvi-Weinstein multiusuário e multimódulo (aqui abreviado como MU-SWA, do inglês, *multiuser SWA*), dada por

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + \frac{\bar{e}_i(n)}{\gamma} \widehat{\mathbf{R}}^{-1}(n) \mathbf{u}^*(n). \quad (3.33)$$

Esse algoritmo apresenta algumas vantagens em relação aos algoritmos de equalização cega multiusuário já existentes. Em primeiro lugar, pelo fato de se tratar de um algoritmo multimódulo, consegue-se evitar rotações complexas não múltiplas de 90° nos sinais recuperados à saída do equalizador. Com isso, a taxa de erro de símbolos para um decisor fixo pode ser reduzida. Além disso, a filosofia multimódulo faz com que não seja necessário utilizar um segundo estágio de equalização para desfazer a rotação da constelação (YANG; WERNER; DUMONT, 2002), economizando recursos. Em segundo lugar, pelo fato de esse algoritmo incluir na equação de atualização uma aproximação para a matriz Hessiana, ele apresenta, em geral, uma convergência mais rápida do que o CMA (SILVA; MIRANDA, 2005).

Infelizmente, assim como ocorre no caso SISO (MIRANDA; SILVA; NASCIMENTO, 2008), a implementação da versão MIMO do SWA como em (3.33) apresenta problemas de instabilidade numérica. Como mostrado em (MIRANDA; SILVA; NASCIMENTO, 2008) para o caso SISO, a divergência do SWA pode ser causada por

- (i) inconsistências nas estimativas não lineares dos sinais transmitidos ou
- (ii) perda de robustez numérica na atualização da inversa da matriz de autocorrelação (HAYKIN, 2014; APOLINÁRIO JR., 2009).

É importante observar que essas duas causas de instabilidade numérica também afetam o funcionamento do algoritmo SWA para o caso MIMO. A seguir, a operação em *dual-mode* proposta por (MIRANDA; SILVA; NASCIMENTO, 2008) é estendida para o caso MIMO, com o objetivo de se evitar a divergência devido à causa (i).

3.3 OPERAÇÃO EM DUAL-MODE

Para contornar os problemas de instabilidade numérica decorrentes de inconsistências na estimativa não linear dos sinais transmitidos do MU-SWA descrito em (3.33), e inspirando-se em (MIRANDA; SILVA; NASCIMENTO, 2008), reescreve-se o erro de (3.29) isolando-se a não linearidade, ou seja

$$\frac{\bar{e}_i(n)}{\gamma} = \bar{d}_i(n) - y_i(n), \quad (3.34)$$

em que

$$\bar{d}_i(n) = d_i(n) - \frac{\epsilon_i(n)}{\gamma}, \quad (3.35)$$

$$d_i(n) = \frac{e_i(n)}{\gamma} + y_i(n) = x_{i,R}(n) y_{i,R}(n) + j x_{i,I}(n) y_{i,I}(n) = d_{i,R}(n) + j d_{i,I}(n), \quad (3.36)$$

sendo

$$x_{i,R}(n) = \frac{3 \text{E} [a_{i,R}^2(n)] - y_{i,R}^2(n)}{3 \text{E} [a_{i,R}^2(n)] - \mathcal{R}} = \frac{3 \text{E} [a_{i,R}^2(n)] - y_{i,R}^2(n)}{\gamma}, \quad (3.37)$$

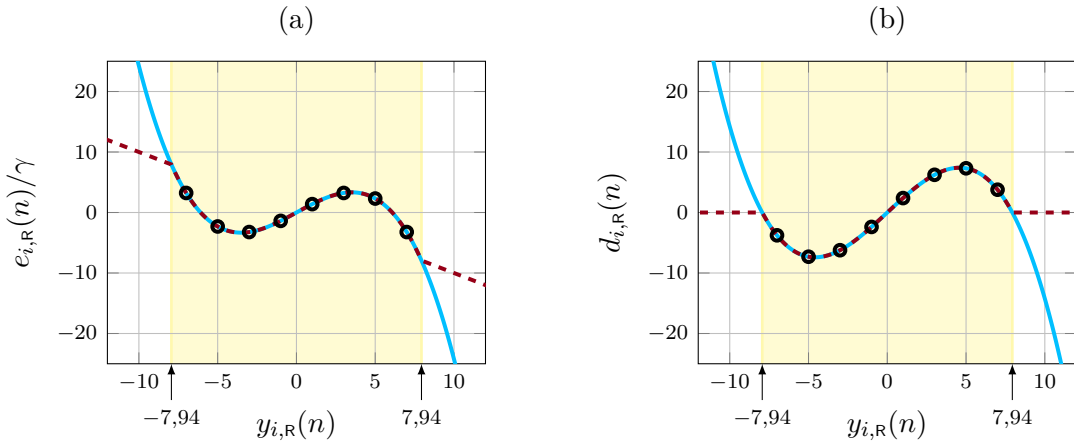
$$x_{i,I}(n) = \frac{3 \text{E} [a_{i,I}^2(n)] - y_{i,I}^2(n)}{3 \text{E} [a_{i,I}^2(n)] - \mathcal{R}} = \frac{3 \text{E} [a_{i,I}^2(n)] - y_{i,I}^2(n)}{\gamma}. \quad (3.38)$$

Nota-se que o sinal $\bar{d}_i(n)$, definido em (3.35), contém não linearidades em relação a $y_i(n)$. Essas não linearidades resultam do gradiente de $\hat{J}_i(n)$, e são representadas em dois termos: $d_i(n)$ e $\epsilon_i(n)/\gamma$. Além disso, a parte real do termo $d_i(n)$, definido em (3.36), pode ser escrita como o produto de $y_{i,R}(n)$ pelo fator adimensional $x_{i,R}(n)$ (analogamente para a parte imaginária). Como $3 \text{E} [a_{i,R}^2] = 3 \text{E} [a_{i,I}^2] > \mathcal{R}$, pode-se observar a partir de (3.37) e (3.38) que, se as condições $y_{i,R}^2(n) < 3 \text{E} [a_{i,R}(n)^2]$ e $y_{i,I}^2(n) < 3 \text{E} [a_{i,I}(n)^2]$ são satisfeitas, então $x_{i,R}(n)$ e $x_{i,I}(n)$ são positivos. Portanto, nesse caso, $d_{i,R}(n)$ e $d_{i,I}(n)$, como definidos em (3.36), têm os mesmos sinais de $y_{i,R}(n)$ e $y_{i,I}(n)$, respectivamente.

Para interpretar as Equações (3.34) a (3.38), considera-se inicialmente um sinal com modulação de amplitude em quadratura (QAM, do inglês, *quadrature amplitude modulation*) de 64 símbolos, *i.e.*, 64-QAM. Na Figura 3(a), a curva contínua representa $e_{i,R}(n)/\gamma = d_{i,R}(n) - y_{i,R}(n)$ em função de $y_{i,R}(n)$. Os erros nas posições correspondentes às coordenadas da parte real dos símbolos da constelação são indicados por pontos pretos. Na Figura 3(b), a curva contínua corresponde à parte real de $d_i(n)$. Quando $|y_{i,R}(n)| < \sqrt{3 \text{E} [a_{i,R}^2(n)]} = \sqrt{3 \times 21} \approx 7,94$, todos os símbolos da constelação estão contemplados em $y_{i,R}(n)$. Nessa situação, ambos $y_{i,R}(n)$ e $d_{i,R}(n)$ têm o mesmo sinal. Quando $|y_{i,R}(n)| > 7,94$, o erro representado na Figura 3(a) cresce cubicamente. Considerando-se $\bar{d}_i(n) = 0$, então $e_i(n)/\gamma = -y_i(n)$. Esse caso é representado nas Figuras 3(a) e (b) com

a curva tracejada. Nota-se que, na Figura 3(a), o erro correspondente à curva tracejada cresce linearmente quando $|y_{i,R}(n)| > 7,94$, sendo, portanto, assintoticamente mais lento do que o crescimento do erro representado pela curva contínua, ou seja, quando se considera a não linearidade representada em $\bar{d}_i(n)$.

Figura 3 – (a) Parte real do erro $e_i(n)/\gamma = \bar{d}_i(n) - y_i(n)$ em função de $y_{i,R}(n)$. (b) Parte real de $d_i(n)$ em função de $y_{i,R}(n)$ para constelação transmitida 64-QAM. As coordenadas da parte real dos símbolos da constelação correspondentes aos valores de $e_{i,R}(n)/\gamma$ e $d_{i,R}(n)$ são indicadas por pontos pretos.



Fonte: autoria própria.

Levando em conta as observações anteriores, definem-se dois modos de operação para o MU-SWA. No primeiro modo, a atualização dos coeficientes é feita com (3.33) reescrita como

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + [\bar{d}_i(n) - y_i(n)] \widehat{\mathbf{R}}^{-1}(n) \mathbf{u}^*(n). \quad (3.39)$$

No segundo modo, $\bar{d}_i(n)$ é considerado nulo e a atualização dos coeficientes é feita como

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) - y_i(n) \widehat{\mathbf{R}}^{-1}(n) \mathbf{u}^*(n). \quad (3.40)$$

A motivação para se definir esses dois modos de operação consiste em remeter a função não linear $\bar{d}_i(n)$ à entrada do algoritmo, como sugerido em (3.34).

A escolha do modo de operação, como em (MIRANDA; SILVA; NASCIMENTO, 2008), deve ser feita em função dos valores das saídas dos equalizadores. Define-se a região de interesse (RoI, do inglês, *region of interest*) como a faixa de valores de $y_i(n)$ que contempla todos os símbolos da constelação. Assim, considerando-se inicialmente apenas a i -ésima saída do equalizador, pode-se definir a RoI de $y_i(n)$ como $|y_i(n)| \leq \sqrt{3E[|a_i(n)|^2]}$.

Porém, no MU-SWA, os vetores $\mathbf{w}_i(n)$ são atualizados para $i = 1, 2, \dots, N_u$, sujeitos ao mesmo vetor regressor $\mathbf{u}(n)$. Além disso, devido aos termos $\epsilon_i(n)$ que penalizam a correlação cruzada, as saídas dos equalizadores estão relacionadas entre si. Sendo assim,

uma questão natural que surge é: como considerar quando o conjunto das estimativas $y_i(n)$ pertence à RoI? Por exemplo, se apenas a saída de um equalizador qualquer se afastar pouco da região de interesse, a atualização dos coeficientes apenas desse equalizador, feita fora da região de interesse, pode causar um efeito prejudicial para os demais equalizadores.

Diante disso, aplica-se a condição de pertinência à RoI não em cada saída $y_i(n)$ individualmente, mas no conjunto dos $y_i(n)$. Portanto, considera-se que o algoritmo está dentro da RoI se

$$\frac{1}{N_u} \sum_{i=1}^{N_u} |y_i(n)|^2 \leq 3 \text{E} [|a_i(n)|^2] = 3(4 - \beta) \text{E} [a_{i,R}^2(n)], \quad (3.41)$$

em que $\beta = 3$ para o caso real e $\beta = 2$ para o caso não real. Nessa situação, os vetores de coeficientes são atualizados segundo (3.39). Se a condição (3.41) não for satisfeita para o conjunto das saídas $y_i(n)$, os vetores de coeficientes devem ser atualizados segundo (3.40).

No exemplo dado nessa seção para o 64-QAM, o símbolo de maior energia está contido no intervalo $|y_i(n)| < \sqrt{3 \text{E} [|a_i(n)|^2]}$. Uma generalização do resultado que relaciona o símbolo de maior energia com a variância da constelação é enunciada a seguir.

Teorema 3.3–A. *Seja um sinal de constelação simétrica e equiprovável, representado por $a(n) = a_R(n) + ja_I(n)$. Se a constelação é retangular do tipo S -QAM, com $S = 2^{2M}$ e $M = 1, 2, 3, \dots$, e a parte real do símbolo de maior energia é denotada como α , vale a seguinte desigualdade:*

$$\alpha^2 < 3 \text{E} [a_R^2]. \quad (3.42)$$

Demonstração. Para a constelação retangular S -QAM, as partes reais dos símbolos pertencerão ao conjunto $\{-\sqrt{S} + 1, \dots, -3, -1, 1, 3, \dots, \sqrt{S} - 1\}$. Logo, o símbolo de maior energia possuirá parte real $\alpha = \sqrt{S} - 1 = 2^M - 1 \Rightarrow \alpha^2 = 2^{2M} - 2^{M+1} + 1$. Por sua vez, a energia média da parte real dos símbolos da constelação é dada por

$$\begin{aligned} \text{E} [a_R^2] &= \frac{2}{\sqrt{S}} \sum_{k=1}^{\sqrt{S}/2} (2k - 1)^2 \\ &= \frac{2}{2^M} \sum_{k=1}^{2^{M-1}} (2k - 1)^2 \\ &= \frac{2}{2^M} \left[\frac{1}{3} 2^{M-1} (2^{2M} - 1) \right], \end{aligned}$$

em que a fórmula da soma dos quadrados dos números ímpares (JEFFREY; DAI, 2008, ver Eq. (2) da Sec. 1.2.3.2) foi usada na última igualdade. Logo, a energia média da parte real pode ser reescrita como

$$\text{E} [a_R^2] = \frac{2^{2M} - 1}{3} \Rightarrow 3 \text{E} [a_R^2] = 2^{2M} - 1.$$

Ambos α^2 e $3 \text{E} [a_R^2]$ possuem o termo 2^{2M} em comum. Por outro lado, como $2^M > 1$ para $M = 1, 2, 3, \dots$, resulta que $-2^{M+1} + 1 < -1$ e, portanto, $\alpha^2 < 3 \text{E} [a_R^2]$. ■

Com a generalização provada no Teorema 3.3–A, conclui-se que, para qualquer constelação retangular e simétrica QAM, os símbolos da parte real (ou imaginária) da constelação sempre serão contemplados em $d_{i,R}(n)$ (ou $d_{i,I}(n)$).

3.4 CONSIDERAÇÕES SOBRE A IMPLEMENTAÇÃO

Nesta seção, são apresentadas duas implementações que tratam causas de divergência do algoritmo de Shalvi-Weinstein multiusuário (MU-SWA) deduzido na Seção 3.2. A primeira implementação utiliza a filosofia *dual-mode* para evitar a divergência devido a inconsistências na estimativa não linear dos sinais transmitidos. A segunda implementação, além de operar em *dual-mode*, também evita outra causa de divergência: a perda de robustez numérica na estimativa da inversa da matriz de covariância.

3.4.1 Recursões preliminares

Adotando-se a filosofia *dual-mode*, a atualização dos coeficientes $\mathbf{w}_i(n)$, para $i = 1, 2, \dots, N_u$, é dada pelas Equações (3.39) e (3.40), dentro e fora da RoI, respectivamente. No entanto, a fim de se evitar tanto a inversão da matriz de covariância $\widehat{\mathbf{R}}(n)$, definida em (3.9), quanto o cálculo direto do somatório que resulta em $\epsilon_i(n)$, definido na Equação (3.31), são obtidas as recursões apresentadas a seguir.

Recursão 3.4–A. Parte-se da recursão de $\widehat{\mathbf{R}}(n)$ dada pela Equação (3.15) e reproduzida a seguir por conveniência:

$$\widehat{\mathbf{R}}(n) = \lambda \widehat{\mathbf{R}}(n-1) + \mathbf{u}^*(n) \mathbf{u}^\top(n). \quad (3.43)$$

Uma recursão para atualização de $\widehat{\mathbf{R}}^{-1}(n)$ vem da aplicação direta do lema de inversão matricial na Equação (3.43) (HAYKIN, 2000b; SAYED, 2008), resultando em

$$\widehat{\mathbf{R}}^{-1}(n) = \frac{1}{\lambda} \left[\widehat{\mathbf{R}}^{-1}(n-1) + \frac{\widehat{\mathbf{R}}^{-1}(n-1) \mathbf{u}^*(n) \mathbf{u}^\top(n) \widehat{\mathbf{R}}^{-1}(n-1)}{\lambda + \mathbf{u}^\top(n) \widehat{\mathbf{R}}^{-1}(n-1) \mathbf{u}^*(n)} \right], \quad (3.44)$$

em que $\widehat{\mathbf{R}}^{-1}(-1)$ corresponde à inicialização da inversa da matriz de covariância.

Recursão 3.4–B. Considera-se a Equação (3.31), repetida a seguir por conveniência:

$$\epsilon_i(n) \triangleq \frac{\rho}{2} \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} \hat{c}_{i,m}(n,k) y_m(n-k). \quad (3.45)$$

Definindo-se

$$\check{\mathbf{y}}_m(n) \triangleq \begin{bmatrix} y_m(n) & y_m(n-1) & \cdots & y_m(n-k_1) \end{bmatrix}^\top, \quad (3.46)$$

$$\mathbf{y}_i(n) \triangleq \begin{bmatrix} \check{\mathbf{y}}_1^\top(n) & \cdots & \check{\mathbf{y}}_{i-1}^\top(n) & \check{\mathbf{y}}_{i+1}^\top(n) & \cdots & \check{\mathbf{y}}_{N_u}^\top(n) \end{bmatrix}^\top, \quad (3.47)$$

$$\check{\mathbf{c}}_{i,m}(n) \triangleq \begin{bmatrix} \hat{c}_{i,m}(n,0) & \hat{c}_{i,m}(n,1) & \cdots & \hat{c}_{i,m}(n,k_1) \end{bmatrix}^\top, \quad (3.48)$$

$$\check{\mathbf{c}}_i(n) \triangleq \begin{bmatrix} \check{\mathbf{c}}_{i,1}^\top(n) & \cdots & \check{\mathbf{c}}_{i,i-1}^\top(n) & \check{\mathbf{c}}_{i,i+1}^\top(n) & \cdots & \check{\mathbf{c}}_{i,N_u}^\top(n) \end{bmatrix}^\top, \quad (3.49)$$

a Equação (3.45) pode ser reescrita como um produto interno de vetores, ou seja,

$$\epsilon_i(n) = \frac{\rho}{2} \hat{\mathbf{c}}_i^\top(n) \mathbf{y}_i(n). \quad (3.50)$$

Os elementos $\hat{c}_{i,m}(n, k)$ do vetor $\check{\mathbf{c}}_{i,m}(n)$, definidos previamente em (3.26), são reescritos aqui por conveniência:

$$\hat{c}_{i,m}(n, k) = (1 - \lambda) \sum_{p=0}^n \lambda^{n-p} y_i(p) y_m^*(p - k). \quad (3.51)$$

Nota-se que, a partir das definições (3.46) a (3.49) e (3.26), vale a recursão

$$\hat{\mathbf{c}}_i(n) = \lambda \hat{\mathbf{c}}_i(n - 1) + (1 - \lambda) y_i(n) \mathbf{y}_i^*(n). \quad (3.52)$$

Por fim, cabe notar que na operação em *dual-mode*, ao entrar novamente na região de interesse, o algoritmo não deve considerar as informações das saídas passadas dos equalizadores, necessárias para calcular os termos $\epsilon_i(n)$ que penalizam a correlação cruzada entre as saídas.

A partir das recursões encontradas para $\widehat{\mathbf{R}}^{-1}(n)$ e $\hat{\mathbf{c}}_i(n)$ e da consideração feita sobre a operação em *dual-mode*, a seguir são apresentadas as tabelas de dois algoritmos propostos para evitar a divergência do algoritmo de Shalvi-Weinstein multiusuário e multimódulo deduzido na Seção 3.2. Tratam-se do MU-DM-SWA (do inglês, *multiuser dual-mode SWA*) e do MU-DM-QRD-SWA (do inglês, *multiuser dual-mode QR decomposition SWA*).

3.4.2 MU-DM-SWA

O MU-DM-SWA consiste em uma implementação do algoritmo de Shalvi-Weinstein deduzido na Seção 3.2, seguindo a filosofia *dual-mode* apresentada na Seção 3.3. A operação em *dual-mode* evita a divergência em virtude da inconsistência numérica nas estimativas de grandezas não lineares do algoritmo.

Na Tabela 1, são apresentados, de forma resumida, os parâmetros principais e a inicialização para o algoritmo. Nota-se que a inicialização de $\widehat{\mathbf{R}}^{-1}(n)$, que corresponde a uma estimativa da inversa da matriz de covariância do vetor regressor, foi estipulada como sendo uma matriz diagonal com entradas proporcionais ao inverso da variância dos símbolos transmitidos, *i.e.*,

$$\widehat{\mathbf{R}}^{-1}(-1) = \delta \mathbf{I}_M, \text{ com } \delta = \frac{1 - \lambda}{(4 - \beta) \text{E} [a_{i,R}^2(n)]}, \quad (3.53)$$

em que \mathbf{I}_M denota a matriz identidade $M \times M$. Além disso, adota-se a inicialização usual de $\mathbf{w}_i(n)$ em desconvolução cega, *i.e.*, um *center-spike* de amplitude unitária. Finalmente, os vetores para estimação de correlação cruzada entre as fontes, ou seja, $\hat{\mathbf{c}}_i(n)$ e $\mathbf{y}_i(n)$, são

inicializados com zeros. As equações que implementam o MU-DM-SWA, para N iterações, são agrupadas na Tabela 2. Nota-se que foi utilizada a notação

$$\xi_i(n) = \frac{\bar{e}_i(n)}{\gamma} = \bar{d}_i(n) - y_i(n). \quad (3.54)$$

Tabela 1 – Parâmetros principais e inicialização do MU-DM-SWA.

Parâmetros principais:

$M = LK_t$: número de coeficientes de cada um dos N_u filtros adaptativos

$\mathcal{R} = \text{E}[a_{\text{R}}^4(n)] / \text{E}[a_{\text{R}}^2(n)] = \text{E}[a_1^4(n)] / \text{E}[a_1^2(n)]$

$\beta = 3$: caso real, $\beta = 2$: caso não real

$\gamma = \beta \text{E}[a_{i,\text{R}}^2(n)] - \mathcal{R}$

λ : fator de esquecimento do algoritmo

g : ponderação da parcela de penalização da correlação cruzada no erro

Inicialização:

$\widehat{\mathbf{R}}^{-1}(-1) = \delta \mathbf{I}_M$

$\delta = \frac{1 - \lambda}{(4 - \beta) \text{E}[a_{i,\text{R}}^2(n)]}$

Limiar = $3(4 - \beta) \text{E}[a_{i,\text{R}}^2(n)]$

Para $i = 1, 2, \dots, N_u$, faça: $\left\{ \begin{array}{l} \mathbf{w}_i(-1) = [0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0]^T \\ \widehat{\mathbf{c}}_i(-1) = \mathbf{0} \\ \mathbf{y}_i(-1) = \mathbf{0} \end{array} \right\}$

Fonte: autoria própria.

Tabela 2 – Equações que implementam o MU-DM-SWA.

Para $n = 0, 1, \dots, N - 1$, calcule: $\left\{ \right.$

Para $i = 1, 2, \dots, N_u$, calcule: $\left\{ \right.$

$$y_i(n) = \mathbf{u}^\top(n) \mathbf{w}_i(n-1) \left. \right\}$$

$$\widehat{\mathbf{R}}^{-1}(n) = \frac{1}{\lambda} \left[\widehat{\mathbf{R}}^{-1}(n-1) - \frac{\widehat{\mathbf{R}}^{-1}(n-1) \mathbf{u}^*(n) \mathbf{u}^\top(n) \widehat{\mathbf{R}}^{-1}(n-1)}{\lambda + \mathbf{u}^\top(n) \widehat{\mathbf{R}}^{-1}(n) \mathbf{u}^*(n)} \right]$$

Para $i = 1, 2, \dots, N_u$, calcule: $\left\{ \right.$

$$\widehat{\mathbf{c}}_i(n) = \lambda \widehat{\mathbf{c}}_i(n-1) + (1 - \lambda) y_i(n) \mathbf{y}_i^*(n)$$

$$\epsilon_i(n) = \frac{\rho}{2} \widehat{\mathbf{c}}_i^\top(n) \mathbf{y}_i(n)$$

$$d_{i,R}(n) = [\gamma + \mathcal{R} - y_{i,R}^2(n)] y_{i,R}(n)$$

$$d_{i,I}(n) = [\gamma + \mathcal{R} - y_{i,I}^2(n)] y_{i,I}(n)$$

$$\bar{d}_i(n) = \gamma^{-1} [d_{i,R}(n) + j d_{i,I}(n) - \epsilon_i(n)] \left. \right\}$$

Se $\frac{1}{N_u} \sum_{i=0}^{N_u} |y_i(n)|^2 > \text{Limiar}$, faça

Para $i = 1, 2, \dots, N_u$, faça: $\left\{ \right.$

$$\bar{d}_i(n) = 0$$

$$\widehat{\mathbf{c}}_i(n) = \mathbf{0}$$

$$\mathbf{y}_i(n) = \mathbf{0} \left. \right\}$$

Para $i = 1, 2, \dots, N_u$, calcule: $\left\{ \right.$

$$\xi_i(n) = \bar{d}_i(n) - y_i(n)$$

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + \widehat{\mathbf{R}}^{-1}(n) \xi_i(n) \mathbf{u}^*(n) \left. \right\}$$

Fonte: autoria própria.

3.4.3 MU-DM-QRD-SWA

O MU-DM-QRD-SWA consiste em uma implementação do algoritmo de Shalvi-Weinstein, deduzido na Seção 3.2, com decomposição QR-MIMO, seguindo a filosofia *dual-mode* apresentada na Seção 3.3. Utilizando-se a decomposição QR como mostrado a seguir, é possível evitar outra causa de divergência em virtude de erros numéricos no cálculo da inversa da matriz de covariância, ou seja, $\widehat{\mathbf{R}}^{-1}(n)$ (HAYKIN, 2014; APOLINÁRIO JR., 2009).

Inicialmente, nota-se que a atualização dos coeficientes do MU-DM-SWA, dada pela Equação (3.39), possui a mesma estrutura da atualização de coeficientes do algoritmo RLS (do inglês, *recursive least squares*) (HAYKIN, 2014). Portanto, (3.39) pode ser reescrita utilizando-se sua representação em espaço de estados (REGALIA, 1993; MIRANDA; SILVA; NASCIMENTO, 2008) e (3.54), na forma

$$\begin{bmatrix} \mathbf{w}_i(n) \\ \xi_i(n) \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{R}}^{-1}(n) \widehat{\mathbf{R}}(n-1) & \widehat{\mathbf{R}}^{-1}(n) \mathbf{u}^*(n) \\ -\lambda^{-1} \mathbf{u}^\top(n) & 1 \end{bmatrix} \begin{bmatrix} \lambda \mathbf{w}_i(n-1) \\ \bar{d}_i(n) \end{bmatrix}. \quad (3.55)$$

Na literatura, há diferentes técnicas para garantir a robustez numérica na atualização de $\widehat{\mathbf{R}}^{-1}(n)$. Por exemplo, no caso de algumas versões de algoritmos RLS, utiliza-se decomposição QR para se garantir, implicitamente, a existência de $\widehat{\mathbf{R}}^{-1}(n)$ mesmo para sequências de entrada próximas da ausência de excitação persistente (REGALIA, 1993; HAYKIN, 2014; APOLINÁRIO JR., 2009). Por exemplo, seja uma transformação de similaridade dada por

$$\begin{bmatrix} \mathbf{C}(n) & \mathbf{0} \\ \mathbf{0}^\top & \varsigma \end{bmatrix}, \quad (3.56)$$

em que $\mathbf{C}(n)$ é o fator de Cholesky de $\widehat{\mathbf{R}}(n)$ e ς é a raiz quadrada do fator de conversão (HAYKIN, 2014; APOLINÁRIO JR., 2009). Aplicando-se essa transformação aos dois lados de (3.55), após manipulações algébricas, resulta

$$\begin{bmatrix} \bar{\mathbf{p}}_i(n) \\ \varsigma \xi_i(n) \end{bmatrix} = \Theta(n) \begin{bmatrix} \sqrt{\lambda} \bar{\mathbf{p}}_i(n-1) \\ \bar{d}_i(n) \end{bmatrix}, \quad (3.57)$$

em que

$$\bar{\mathbf{p}}_i(n) = \mathbf{C}(n) \mathbf{w}_i(n-1), \quad (3.58)$$

$$\Theta(n) = \begin{bmatrix} \sqrt{\lambda} \mathbf{C}^{-\top}(n) \mathbf{C}^\top(n-1) & \mathbf{C}^{-\top}(n) \mathbf{u}^*(n) \\ -\sqrt{\lambda} \varsigma \mathbf{b}^\top(n) & \varsigma \end{bmatrix}, \quad (3.59)$$

$$\mathbf{b}^\top(n) = \mathbf{u}^\top(n) \mathbf{C}^{-1}(n-1). \quad (3.60)$$

Sabe-se que a matriz unitária de rotação $\Theta(n)$ possui um papel importante na atualização da matriz \mathbf{C} (REGALIA, 1993; HAYKIN, 2014; APOLINÁRIO JR., 2009), ou seja,

$$\begin{bmatrix} \mathbf{C}(n) \\ \mathbf{0}^\top \end{bmatrix} = \Theta(n) \begin{bmatrix} \sqrt{\lambda} \mathbf{C}(n-1) \\ \mathbf{u}^\top(n) \end{bmatrix}. \quad (3.61)$$

Além disso, a saída do i -ésimo equalizador pode ser calculada como

$$y_i(n) = \mathbf{u}^T(n) \mathbf{w}_i(n-1) = \mathbf{b}^T(n) \bar{\mathbf{p}}_i(n-1). \quad (3.62)$$

Finalmente, cabe notar que a matriz $\mathbf{C}(n)$ resulta da decomposição QR da matriz de dados (HAYKIN, 2014; APOLINÁRIO JR., 2009). Estendendo-se essa transformação para todos os filtros que compõem o equalizador MIMO, obtém-se o MU-DM-QRD-SWA.

Na Tabela 3, são apresentados, de forma resumida, os parâmetros principais e a inicialização para o algoritmo. As inicializações são equivalentes àsquelas do MU-DM-SWA apresentadas na Tabela 1, tendo em vista que $\mathbf{C}(n)$ corresponde ao fator de Cholesky de $\hat{\mathbf{R}}(n)$. Com isso, resulta

$$\mathbf{C}(-1) = \frac{1}{\sqrt{\delta}} \mathbf{I}_M, \quad (3.63)$$

sendo δ dado pela Equação (3.53). As equações que implementam o MU-DM-QRD-SWA, para N iterações, são agrupadas na Tabela 4.

Tabela 3 – Parâmetros principais e inicialização do MU-DM-QRD-SWA.

Parâmetros principais:

$M = LK_t$: número de coeficientes de cada um dos N_u filtros adaptativos

$$\mathcal{R} = \mathbb{E}[a_R^4(n)] / \mathbb{E}[a_R^2(n)] = \mathbb{E}[a_i^4(n)] / \mathbb{E}[a_i^2(n)]$$

$\beta = 3$: caso real, $\beta = 2$: caso não real

$$\gamma = \beta \mathbb{E}[a_{i,R}^2(n)] - \mathcal{R}$$

$\sqrt{\lambda}$: fator de esquecimento do algoritmo

ϱ : ponderação da parcela de penalização da correlação cruzada no erro

Inicialização:

$$\mathbf{C}(-1) = \mathbf{I}_M / \sqrt{\delta}$$

$$\delta = \frac{1 - \lambda}{(4 - \beta) \mathbb{E}[a_{i,R}^2(n)]}$$

$$\text{Para } i = 1, 2, \dots, N_u, \text{ faça: } \left\{ \begin{array}{l} \bar{\mathbf{p}}_i(-1) = [0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0]^T \\ \hat{\mathbf{c}}_i(-1) = \mathbf{0} \\ \mathbf{y}_i(-1) = \mathbf{0} \end{array} \right\}$$

Fonte: autoria própria.

Tabela 4 – Equações que implementam o MU-DM-QRD-SWA.

Para $n = 0, 1, \dots, N - 1$, calcule: $\left\{ \right.$

$$\Theta(n) = \left[\theta_{i,j}(n) \right]_{\substack{1 \leq i \leq M+1 \\ 1 \leq j \leq M+1}} \quad \text{tal que} \quad \begin{bmatrix} \mathbf{C}(n) \\ \mathbf{0}^\top \end{bmatrix} = \Theta(n) \begin{bmatrix} \sqrt{\lambda} \mathbf{C}(n-1) \\ \mathbf{u}^\top(n) \end{bmatrix}$$

$$\mathbf{x}^\top(n) = \left[\theta_{M+1,j}(n) \right]_{1 \leq j \leq M}$$

$$\mathbf{b}^\top(n) = -\frac{\sqrt{\lambda}}{\theta_{M+1,M+1}(n)} \mathbf{x}^\top(n)$$

$$\begin{bmatrix} y_1(n) & y_2(n) & \dots & y_{N_u}(n) \end{bmatrix} = \mathbf{b}^\top(n) \begin{bmatrix} \bar{\mathbf{p}}_1(n-1) & \bar{\mathbf{p}}_2(n-1) & \dots & \bar{\mathbf{p}}_{N_u}(n-1) \end{bmatrix}$$

Para $i = 1, 2, \dots, N_u$, calcule: $\left\{ \right.$

$$\hat{\mathbf{c}}_i(n) = \lambda \hat{\mathbf{c}}_i(n-1) + (1 - \lambda) y_i(n) \mathbf{y}_i^*(n)$$

$$\epsilon_i(n) = \frac{\rho}{2} \hat{\mathbf{c}}_i^\top(n) \mathbf{y}_i(n)$$

$$d_{i,R}(n) = [\gamma + \mathcal{R} - y_{i,R}^2(n)] y_{i,R}(n)$$

$$d_{i,I}(n) = [\gamma + \mathcal{R} - y_{i,I}^2(n)] y_{i,I}(n)$$

$$\bar{d}_i(n) = \gamma^{-1} [d_{i,R}(n) + j d_{i,I}(n) - \epsilon_i(n)] \left. \right\}$$

Se $\frac{1}{N_u} \sum_{i=0}^{N_u} |y_i(n)|^2 > \text{Limiar}$, faça

Para $i = 1, 2, \dots, N_u$, faça: $\left\{ \right.$

$$\bar{d}_i(n) = 0$$

$$\hat{\mathbf{c}}_i(n) = \mathbf{0}$$

$$\mathbf{y}_i(n) = \mathbf{0} \left. \right\}$$

$$\begin{bmatrix} \bar{\mathbf{p}}_1(n) & \bar{\mathbf{p}}_2(n) & \dots & \bar{\mathbf{p}}_{N_u}(n) \\ \varsigma \xi_1(n) & \varsigma \xi_2(n) & \dots & \varsigma \xi_{N_u}(n) \end{bmatrix} = \Theta(n) \begin{bmatrix} \sqrt{\lambda} \bar{\mathbf{p}}_1(n-1) & \sqrt{\lambda} \bar{\mathbf{p}}_2(n-1) & \dots & \sqrt{\lambda} \bar{\mathbf{p}}_{N_u}(n-1) \\ \bar{d}_1(n) & \bar{d}_2(n) & \dots & \bar{d}_{N_u}(n) \end{bmatrix}$$

$$\left. \right\}$$

Fonte: autoria própria.

3.5 RESULTADOS DE SIMULAÇÃO

As simulações foram feitas em dois cenários distintos, baseados no modelo da Figura 2:

- **Cenário 3.5–A:** Nesse cenário, consideram-se $N_u = 2$ fontes e $L = 3$ sensores. As fontes são sinais 4-QAM de valores $\{a_R + ja_I\}$, sendo $a_R, a_I = \pm 1$ com $\beta = 2$, $\mathcal{R} = 1$ e $\gamma = 2$. O canal de comunicação MIMO é real. Cada canal associando a i -ésima fonte ao ℓ -ésimo sensor é um filtro FIR com $K_c = 2$ coeficientes. Os polinômios $H_{i,\ell}(z)$ do canal MIMO são apresentados na Tabela 5.

Tabela 5 – Polinômios $H_{i,\ell}(z)$ do canal MIMO no Cenário 3.5–A.

$i \backslash \ell$	1	2	3
1	$-0,6 + 1,2z^{-1}$	$0,5 - 1,0z^{-1}$	$0,4 - 0,2z^{-1}$
2	$0,1 - 0,2z^{-1}$	$-0,6 + 0,9z^{-1}$	$-0,1 + 0,4z^{-1}$

Fonte: autoria própria.

- **Cenário 3.5–B:** Nesse cenário, consideram-se $N_u = 3$ fontes e $L = 4$ sensores. As fontes são sinais 64-QAM de valores $\{a_R + ja_I\}$, sendo $a_R, a_I = \pm 1, \pm 3, \pm 5, \pm 7$, com $\beta = 2$, $\mathcal{R} = 37$ e $\gamma = 26$. O canal de comunicação MIMO é complexo, com $K_c = 2$ e coeficientes $H_{i,\ell}(z)$ fornecidos na Tabela 6.

Tabela 6 – Polinômios $H_{i,\ell}(z)$ do canal MIMO no Cenário 3.5–B.

$i \backslash \ell$	1	2	3	4
1	$-0,6 + j0,4$ $+(1,2 - j0,2)z^{-1}$	$0,1 + j0,7$ $-(0,2 + j0,5)z^{-1}$	$0,5 + j0,4$ $+(-1,0 + j0,3)z^{-1}$	$-0,7 - j0,1$ $+(0,5 - j0,3)z^{-1}$
2	$-0,6 + j0,8$ $+(0,9 - j0,1)z^{-1}$	$0,4 - j0,3$ $+(-0,2 + j0,2)z^{-1}$	$-0,1 + j0,8$ $+(0,4 + j0,1)z^{-1}$	$-0,7 + j0,2$ $+(0,8 - j0,2)z^{-1}$
3	$-0,8 + j0,5$ $+(0,4 - j0,8)z^{-1}$	$-0,2 - j0,3$ $+(-0,3 + j0,4)z^{-1}$	$0,5 - j0,6$ $+(-0,2 + j0,5)z^{-1}$	$0,1 + j0,6$ $+(0,9 - j0,3)z^{-1}$

Fonte: autoria própria.

Em cada cenário, são confrontadas as versões do algoritmo de Shalvi-Weinstein que evitam a divergência (MU-DM-SWA e MU-DM-QRD-SWA), apresentadas na Seção 3.4, com a versão original do algoritmo (MU-SWA) deduzida na Seção 3.2. Os parâmetros utilizados para os três algoritmos são indicados na Tabela 7.

A inicialização dos algoritmos, de acordo com as Tabelas 1 e 3, é detalhada a seguir:

Tabela 7 – Parâmetros dos três algoritmos nos Cenários 3.5–A e B.

Parâmetro	Cenário 3.5–A	Cenário 3.5–B
Diversidade temporal do equalizador (K_t)	4	4
Número de coeficientes de cada filtro (M)	12	16
Fator de esquecimento (λ)	0,9995	0,9997
Ponderação do termo de correlação cruzada (ρ)	1	1
Atraso de espalhamento (k_1)	3	3

Fonte: autoria própria.

- No Cenário 3.5–A, os equalizadores MU-SWA e MU-DM-SWA são inicializados com $\widehat{\mathbf{R}}^{-1}(-1) = 2,5 \times 10^{-4} \mathbf{I}_{12}$, de acordo com a Equação (3.53), e vetores $\mathbf{w}_1(-1)$ e $\mathbf{w}_2(-1)$ com pino único de valor 1 nas respectivas posições empiricamente estipuladas 3 e 6. O equalizador MU-DM-QRD-SWA é inicializado, equivalentemente, com $\mathcal{C}(-1) = 63,2456 \mathbf{I}_{12}$, de acordo com a Equação (3.63), e vetores $\bar{\mathbf{p}}_1(-1)$ e $\bar{\mathbf{p}}_2(-1)$ iguais, respectivamente, aos vetores $\mathbf{w}_1(-1)$ e $\mathbf{w}_2(-1)$.
- No Cenário 3.5–B, o MU-SWA e o MU-DM-SWA são inicializados com $\widehat{\mathbf{R}}^{-1}(-1) = 7,1429 \times 10^{-6} \mathbf{I}_{16}$, de acordo com a Equação (3.53), e vetores $\mathbf{w}_1(-1)$, $\mathbf{w}_2(-1)$ e $\mathbf{w}_3(-1)$ com pino único de valor 1 nas respectivas posições empiricamente escolhidas 14, 6 e 9. Por sua vez, o equalizador MU-DM-QRD-SWA é inicializado com $\mathcal{C}(-1) = 374,1657 \mathbf{I}_{16}$, de acordo com a Equação (3.63), e vetores $\bar{\mathbf{p}}_1(-1)$, $\bar{\mathbf{p}}_2(-1)$ e $\bar{\mathbf{p}}_3(-1)$ iguais, respectivamente, a $\mathbf{w}_1(-1)$, $\mathbf{w}_2(-1)$ e $\mathbf{w}_3(-1)$.

A fim de se verificar a robustez à divergência, os algoritmos são simulados sob condições adversas, *i.e.*, presença de perturbação no vetor regressor e relação sinal-ruído elevada nas saídas do canal. O indicador de desempenho utilizado nas simulações é o erro quadrático médio (MSE, do inglês, *mean square error*), definido como

$$\text{MSE}_i(n) \triangleq \text{E} \left[\left| y_i(n) - e^{j\theta_m} a_m(n - \tau_m) \right|^2 \right], \text{ com } i = 1, 2, \dots, N_u, \quad (3.64)$$

em que θ_m e τ_m são, respectivamente, a rotação complexa e o atraso temporal da m -ésima fonte recuperada pelo i -ésimo equalizador. Os valores de θ_m e τ_m são determinados verificando-se a rotação complexa e o atraso com que cada fonte é recuperada em uma simulação prévia do algoritmo com os parâmetros escolhidos. Nas simulações, o MSE ao longo das iterações é estimado a partir de médias de conjunto.

Resultados: Cenário 3.5–A

Nesse cenário, foram consideradas relações sinal-ruído nas saídas do canal de 35 dB e $+\infty$ (*i.e.*, sem ruído). Sob ambas SNRs, no instante $n = 5 \times 10^4$ foi aplicada uma

perturbação de valor 100 em cada regressor $u_\ell(n)$, com $\ell = 1, 2, 3$. A recuperação das fontes, tanto antes quanto depois da perturbação aplicada, para os três tipos de algoritmos considerados e ambas SNRs, se deu de acordo com a Tabela 8.

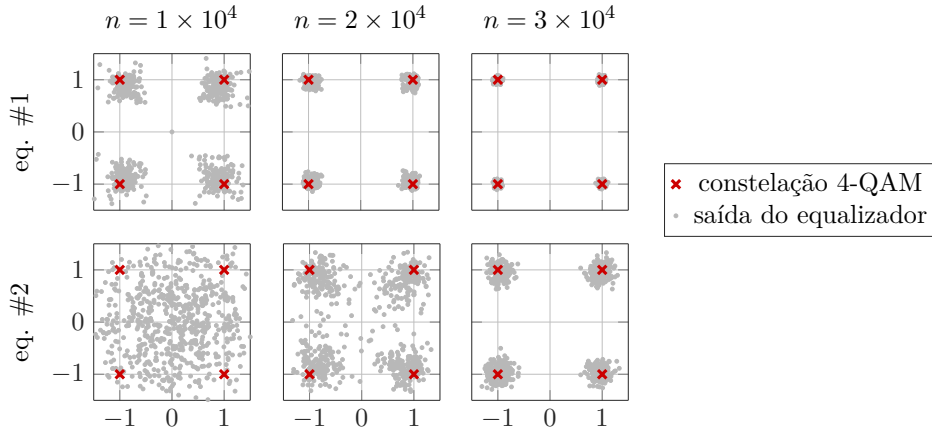
Tabela 8 – Recuperação das fontes no Cenário 3.5–A.

Equalizador (i)	1	2
Fonte recuperada (m)	1	2
Rotação complexa (θ_m)	0°	0°
Atraso temporal (τ_m)	3	2

Fonte: autoria própria.

Na Figura 4, são mostrados os diagramas de constelação para as duas saídas do equalizador MU-DM-QRD-SWA, com SNR = 35 dB nas saídas do canal. Ao longo das colunas, são mostrados os resultados durante três janelas temporais distintas. Cada janela considerada possui duração de 1×10^4 amostras e termina nos instantes indicados ao topo de cada coluna.

Figura 4 – Diagramas de constelação para as duas saídas do equalizador MU-DM-QRD-SWA, no Cenário 3.5–A, com SNR = 35 dB nas saídas do canal e constelação 4-QAM. São consideradas três janelas de 1×10^4 amostras, cada uma terminando na n -ésima iteração.

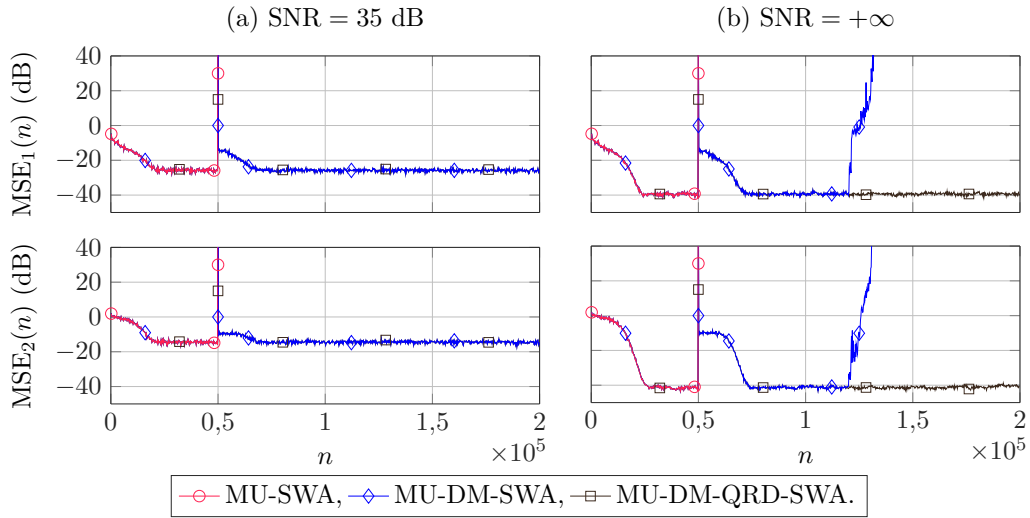


Fonte: autoria própria.

As curvas de MSE em função do tempo, para cada equalizador, são mostradas na Figura 5. O MSE foi estimado a partir de 50 realizações independentes. Cabe observar que perturbação aplicada em $n = 5 \times 10^4$ provoca uma das causas de divergência nos algoritmos sem operação *dual-mode*, devido às inconsistências nas estimativas não lineares dos sinais transmitidos (MIRANDA; SILVA; NASCIMENTO, 2008). Desse modo, após a perturbação, o MU-SWA diverge. Já o MU-DM-SWA e o MU-DM-QRD-SWA conseguem suportar a perturbação aplicada, logo retornando ao seu funcionamento adequado, em aproximadamente 2×10^4 iterações. No entanto, sob SNR infinita nas saídas do canal, os

erros numéricos no cálculo da inversa da matriz de covariância são relativamente mais pronunciados, sendo essa outra causa de divergência (MIRANDA; SILVA; NASCIMENTO, 2008). Como o MU-DM-SWA não possui robustez numérica nesse sentido, ele diverge após um número suficiente de iterações em que os erros numéricos acumulados se tornam muito elevados. Por outro lado, a implementação QR-MIMO do MU-DM-QRD-SWA é numericamente robusta e consegue também evitar essa outra causa de divergência.

Figura 5 – Curvas de MSE para sinais 4-QAM no Cenário 3.5–A, obtidas a partir de 50 realizações independentes.



Fonte: autoria própria.

Resultados: Cenário 3.5–B

Nesse cenário, foram consideradas relações sinal-ruído de 40 dB e $+\infty$ (*i.e.*, sem ruído) nas saídas do canal. Sob ambas SNRs, no instante $n = 5 \times 10^4$ foi aplicada uma perturbação de valor $100 - j100$ em cada regressor $u_\ell(n)$, com $\ell = 1, 2, 3, 4$. A recuperação das fontes, tanto antes quanto depois da perturbação aplicada, para os três tipos de algoritmos considerados e ambas SNRs, se deu de acordo com a Tabela 9.

Tabela 9 – Recuperação das fontes no Cenário 3.5–B.

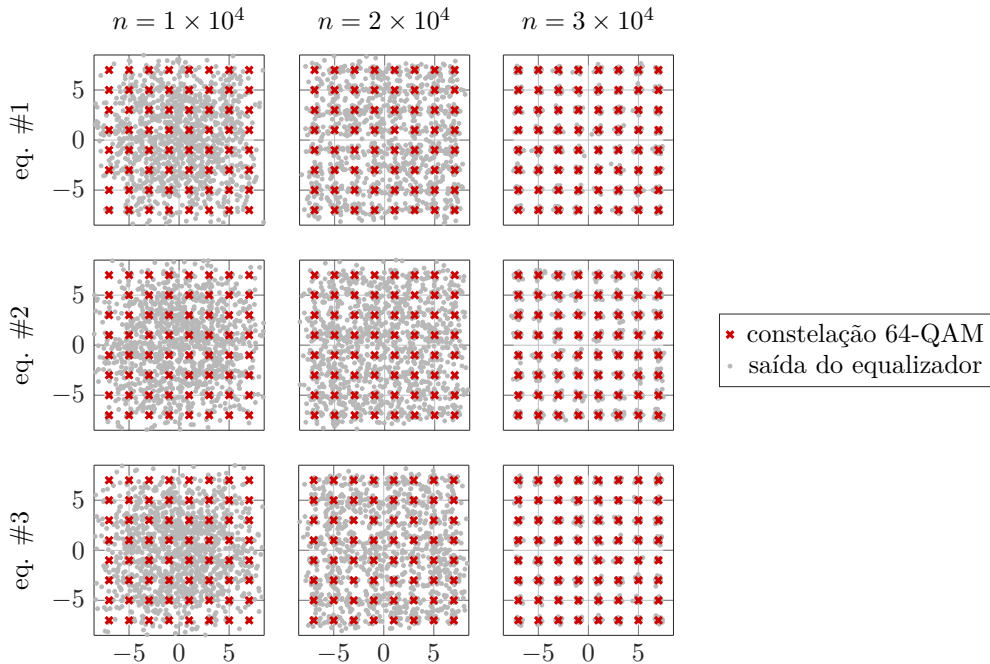
Equalizador (i)	1	2	3
Fonte recuperada (m)	3	1	2
Rotação complexa (θ_m)	0°	90°	90°
Atraso temporal (τ_m)	2	1	0

Fonte: autoria própria.

Na Figura 6, são mostrados os diagrama de constelação para as três saídas do equalizador MU-DM-QRD-SWA, com $\text{SNR} = 40$ dB nas saídas do canal. Ao longo das

colunas, são mostrados os resultados durante três janelas temporais distintas. Cada janela considerada possui duração de 1×10^4 amostras e termina nos instantes indicados ao topo de cada coluna.

Figura 6 – Diagramas de constelação para as duas saídas do equalizador MU-DM-QRD-SWA, no Cenário 3.5-B, com SNR = 40 dB nas saídas do canal e constelação 64-QAM. São consideradas três janelas de 1×10^4 amostras, cada uma terminando na n -ésima iteração.



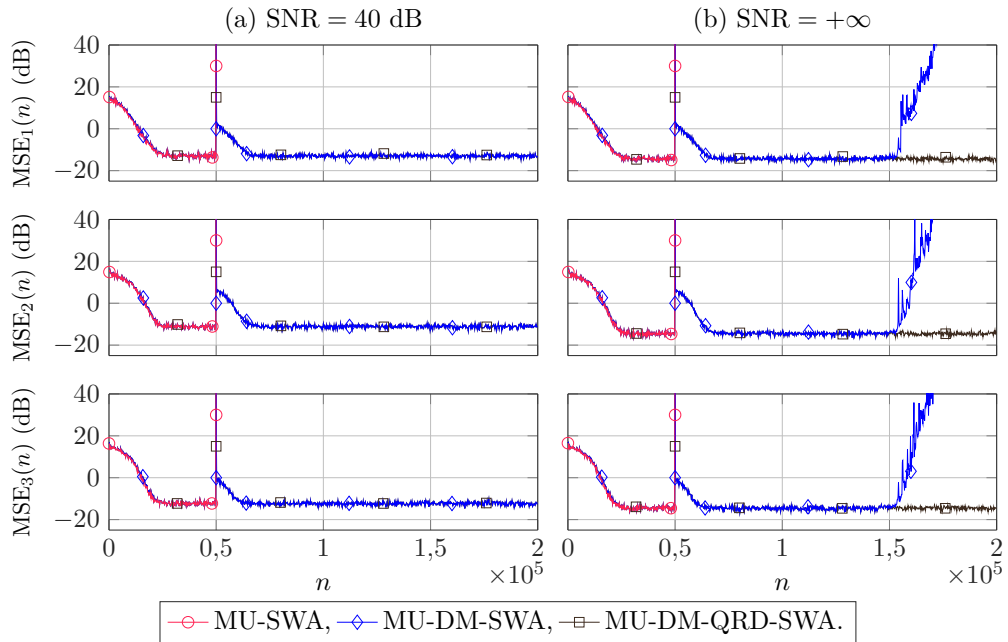
Fonte: autoria própria.

As curvas de MSE em função do tempo, para cada equalizador, são mostradas na Figura 7. O MSE foi estimado a partir de 50 realizações independentes. Novamente, devido à perturbação aplicada, o MU-SWA diverge por não conseguir evitar a instabilidade numérica em virtude das inconsistências nas estimativas não lineares dos sinais transmitidos. Embora o MU-DM-SWA seja imune a essa causa de divergência, sob SNR infinita nas saídas do canal, os erros numéricos no cálculo da inversa da matriz de covariância são relativamente mais pronunciados, fazendo com que ele divirja. Somente o MU-DM-QRD-SWA conseguiu suportar tanto a primeira causa de divergência, com operação em *dual-mode*, quanto a segunda causa, com uma implementação QR-MIMO para contornar os erros numéricos na estimativa da inversa da matriz de covariância.

3.6 CONCLUSÃO

Nesse capítulo, foram propostas soluções adaptativas para a desconvolução cega multiusuário com base no algoritmo de Shalvi-Weinstein (SWA) multiusuário, com o objetivo de incentivar o entendimento dos princípios teóricos de separação cega de fontes, bem como a compreensão de suas soluções.

Figura 7 – Curvas de MSE para sinais 64-QAM no Cenário 3.5-B, obtidas a partir de 50 realizações independentes.



Fonte: autoria própria.

Essas soluções foram obtidas inspirando-se nas soluções existentes para o caso monousuário (MIRANDA; SILVA; NASCIMENTO, 2008; YANG; WERNER; DUMONT, 2002). Com o objetivo de se evitar a divergência do MU-SWA, foram propostos dois algoritmos, *i.e.*, o MU-DM-SWA e o MU-DM-QRD-SWA. O primeiro algoritmo evita uma das causas de divergência relacionada a inconsistências nas estimativas não lineares dos sinais transmitidos, presentes no MU-SWA. Para tanto, propõe-se uma operação em dois modos, denominada *dual-mode*. O segundo algoritmo consegue evitar tanto essa causa de divergência quanto outra causa relacionada à perda de robustez numérica (HAYKIN, 2014) na estimativa da inversa da matriz de covariância. Isso é feito com uma implementação do algoritmo de Shalvi-Weinstein utilizando decomposição QR-MIMO.

As soluções foram testadas com simulações numéricas, e suas propriedades de resistência a perturbações e robustez numérica foram comprovadas. O fato de se conseguir evitar a divergência do SWA devido a perturbações abruptas, com as soluções propostas nesse capítulo, pode ser um grande feito, visto que algoritmos de separação de fontes baseados em maximização de curtose, tais como o SWA, são muito sensíveis à influência de *outliers* (HYVÄRINEN; OJA, 2000).

4 ANÁLISES DO ALGORITMO DE SHALVI-WEINSTEIN

No Capítulo 3, foi obtida uma solução adaptativa eficiente para o problema de desconvolução cega multiusuário: o algoritmo de Shalvi-Weinstein multiusuário (MU-SWA). Além disso, essa solução foi aprimorada a fim de se evitar problemas de instabilidade numérica, com base em uma operação denominada *dual-mode*. Neste capítulo, são feitas análises determinista e em regime do algoritmo de Shalvi-Weinstein com operação em *dual-mode* (MU-DM-SWA).

O capítulo é organizado conforme descreve-se a seguir. Na Seção 4.1, é feita a análise determinista de estabilidade do MU-DM-SWA. Posteriormente, na Seção 4.2, é feita uma análise do erro em regime para esse algoritmo, com operação dentro da região de interesse (RoI), permitindo realizar a escolha de parâmetros que garanta o seu funcionamento adequado em média. Nessa seção, o modelo obtido é validado por meio de simulações numéricas. Por fim, na Seção 4.3, são apresentadas as conclusões deste capítulo.

4.1 ANÁLISE DETERMINISTA

Na análise determinista de estabilidade a seguir, busca-se assegurar que o MU-DM-SWA, quando adequadamente implementado (*i.e.*, com operação em *dual-mode* conforme apresentada no Capítulo 3) evita a divergência em precisão infinita, e, sendo assim, é um algoritmo numericamente estável. Rigorosamente, nesta análise, considera-se que o algoritmo é estável quando a função de Lyapunov, no caso definida como $\|\mathbf{w}_i(n)\|^2$ para o i -ésimo equalizador, for limitada em um intervalo de iterações $[n_0, n]$ arbitrariamente grande (SOLO; KONG, 1995).

Supõe-se, inicialmente, que o algoritmo opera dentro da região de interesse a partir de um instante $n_0 \geq 0$. Por conveniência, reescreve-se a Equação (3.39) em uma representação em espaço de estados, ou seja,

$$\mathbf{w}_i(n) = [\mathbf{I}_M - \widehat{\mathbf{R}}^{-1}(n) \mathbf{u}^*(n) \mathbf{u}^\top(n)] \mathbf{w}_i(n-1) + \bar{d}_i(n) \widehat{\mathbf{R}}^{-1}(n) \mathbf{u}^*(n), \quad (4.1)$$

em que \mathbf{I}_M denota a matriz identidade $M \times M$, com $M = LK_t$, sendo L o número de sensores e K_t a diversidade temporal adotada para cada filtro FIR do equalizador. Observa-se que as não linearidades são colocadas como entradas do algoritmo. Multiplicando-se ambos os lados dessa equação à esquerda por $\widehat{\mathbf{R}}(n)$ e reconhecendo que $\lambda \widehat{\mathbf{R}}(n-1) = \widehat{\mathbf{R}}(n) - \mathbf{u}^*(n) \mathbf{u}^\top(n)$, tem-se

$$\mathbf{v}_i(n) = \lambda \mathbf{v}_i(n-1) + \bar{d}_i(n) \mathbf{u}^*(n), \quad (4.2)$$

sendo $\mathbf{v}_i(n) = \widehat{\mathbf{R}}(n) \mathbf{w}_i(n)$. Trata-se de uma equação de diferenças de primeira ordem, cuja

solução é dada por

$$\mathbf{v}_i(n) = \lambda^{n-n_0+1} \mathbf{v}_i(n_0 - 1) + \sum_{\ell=n_0}^n \lambda^{n-\ell} \bar{d}_i(\ell) \mathbf{u}^*(\ell), \quad (4.3)$$

em que $\mathbf{v}_i(n_0 - 1)$ é uma condição inicial e $n_0 \geq 0$. Multiplicando-se ambos os lados dessa igualdade por $\widehat{\mathbf{R}}^{-1}(n)$, obtém-se a solução da Equação (4.1), dada por

$$\mathbf{w}_i(n) = \widehat{\mathbf{R}}^{-1}(n) \left[\lambda^{n-n_0+1} \widehat{\mathbf{R}}(n_0 - 1) \mathbf{w}_i(n_0 - 1) + \sum_{\ell=n_0}^n \lambda^{n-\ell} \bar{d}_i(\ell) \mathbf{u}^*(\ell) \right]. \quad (4.4)$$

Tomando-se a norma Euclidiana do vetor $\mathbf{w}_i(n)$, aplicando-se a desigualdade triangular e lançando-se mão das propriedades de normas de vetores e matrizes, resulta

$$\|\mathbf{w}_i(n)\| \leq \|\widehat{\mathbf{R}}^{-1}(n)\| \left[\lambda^{n-n_0+1} \|\widehat{\mathbf{R}}(n_0 - 1)\| \|\mathbf{w}_i(n_0 - 1)\| + \sum_{\ell=n_0}^n \lambda^{n-\ell} |\bar{d}_i(\ell)| \|\mathbf{u}^*(\ell)\| \right] \quad (4.5)$$

(dentro da RoI),

em que $\|\widehat{\mathbf{R}}^{-1}(n)\|$ e $\|\widehat{\mathbf{R}}(n_0 - 1)\|$ denotam as normas espectrais das matrizes $\widehat{\mathbf{R}}^{-1}(n)$ e $\widehat{\mathbf{R}}(n_0 - 1)$, respectivamente.

Por outro lado, caso o algoritmo esteja operando fora da região de interesse a partir de um instante $n_0 \geq 0$, a atualização dos coeficientes é feita como em (3.40), sendo possível reescrevê-la na forma

$$\mathbf{w}_i(n) = [\mathbf{I}_M - \widehat{\mathbf{R}}^{-1}(n) \mathbf{u}^*(n) \mathbf{u}^\top(n)] \mathbf{w}_i(n - 1). \quad (4.6)$$

A norma Euclidiana do vetor $\mathbf{w}_i(n)$, nesse caso, possui o limite superior

$$\|\mathbf{w}_i(n)\| \leq \lambda^{n-n_0+1} \|\widehat{\mathbf{R}}^{-1}(n)\| \|\widehat{\mathbf{R}}(n_0 - 1)\| \|\mathbf{w}_i(n_0 - 1)\| \quad (4.7)$$

(fora da RoI).

Deseja-se demonstrar que a norma Euclidiana do vetor de coeficientes $\|\mathbf{w}_i(n)\|$ é limitada independentemente do modo de operação, desde um instante n_0 qualquer até um instante n suficientemente grande, e para quaisquer condições iniciais limitadas e qualquer $i = 1, 2, \dots, N_u$. O fato de a norma $\|\mathbf{w}_i(n)\|$ ser limitada no intervalo $[n_0, n]$ significa que a função de Lyapunov $\|\mathbf{w}_i(n)\|^2$ possui valor finito. Para provar que isso ocorre, é preciso inicialmente assumir a hipótese de excitação persistente (SOLO; KONG, 1995) para a sequência $u(n)$ correspondente ao vetor regressor $\mathbf{u}(n)$, comum aos equalizadores. Essa hipótese é enunciada a seguir.

Hipótese 4.1–A. Seja a sequência de tempo discreto $u(n)$ constituída pela concatenação das K_t amostras mais recentes de cada um dos K_c regressores correspondentes, por sua vez,

a cada um dos sensores. Essa sequência é persistentemente excitante (ou suficientemente rica) se existem α_1 e α_2 tais que¹

$$\alpha_1 \mathbf{I}_M \leq \widehat{\mathbf{R}}(n) \leq \alpha_2 \mathbf{I}_M, \quad \forall n \geq -1.$$

com $0 < \alpha_1 \leq \alpha_2 < +\infty$.

Como consequências da Hipótese 4.1–A, válidas para todo $n \geq -1$, resultam:

- as normas espectrais de $\widehat{\mathbf{R}}^{-1}(n)$ e de $\widehat{\mathbf{R}}(n)$ satisfazem

$$\|\widehat{\mathbf{R}}^{-1}(n)\| \leq \alpha_1^{-1} \quad \text{e} \quad \|\widehat{\mathbf{R}}(n)\| \leq \alpha_2; \quad (4.8)$$

- a norma Euclidiana do vetor regressor $\mathbf{u}(n)$ é limitada superiormente, ou seja,

$$0 \leq \|\mathbf{u}(n)\| \leq B_u < +\infty, \quad B_u > 0. \quad (4.9)$$

A prova de estabilidade é feita considerando-se inicialmente $n_0 = 0$, ou seja, partindo-se da primeira iteração do MU-DM-SWA. Depois, considera-se que, em um instante $n_0 > 0$, o algoritmo transita entre os dois modos de operação. Com isso, pode-se provar que o algoritmo não diverge para qualquer sequência de modos de operação que possa ocorrer na prática.

4.1.1 Inicialização dentro e fora da região de interesse

Considerando-se $n_0 = 0$, tem-se:

- Se a matriz $\widehat{\mathbf{R}}^{-1}(n)$ é inicializada como $\delta \mathbf{I}_M$, com $0 < \delta < +\infty$, então

$$\|\widehat{\mathbf{R}}^{-1}(-1)\| = \delta^{-1} < +\infty. \quad (4.10)$$

- Considerando-se uma inicialização finita do vetor de coeficientes, vale

$$\|\mathbf{w}_i(-1)\| < +\infty, \quad (4.11)$$

sendo que $\|\mathbf{w}_i(-1)\| = 1$ quando a técnica de inicialização usada é o *center spike* de amplitude 1.

- Assumindo-se a Hipótese 4.1–A, nota-se, a partir da Equação (4.8), que $\|\widehat{\mathbf{R}}^{-1}(n)\| \leq \alpha_1^{-1}$.

¹ Para duas matrizes \mathbf{A} e \mathbf{B} , ambas $M \times M$, a desigualdade $\mathbf{A} \geq \mathbf{B}$ denota que a subtração $\mathbf{A} - \mathbf{B}$ é positiva semidefinida.

Portanto, as Equações (4.5) e (4.7) podem ser reescritas, respectivamente, como

$$\|\mathbf{w}_i(n)\| \leq \alpha_1^{-1} \left[\lambda^{n+1} \delta^{-1} \|\mathbf{w}_i(-1)\| + \sum_{\ell=0}^n \lambda^{n-\ell} |\bar{d}_i(\ell)| \|\mathbf{u}^*(\ell)\| \right] \quad (4.12)$$

(dentro da RoI),

$$\|\mathbf{w}_i(n)\| \leq \lambda^{n+1} \alpha_1^{-1} \delta^{-1} \|\mathbf{w}_i(-1)\| \quad (4.13)$$

(fora da RoI).

Para provar que $\|\mathbf{w}_i(n)\| < +\infty$ dentro da RoI, nota-se, a partir da desigualdade (4.9) da Hipótese 4.1–A, que

$$\sum_{\ell=0}^n \lambda^{n-\ell} |\bar{d}_i(\ell)| \|\mathbf{u}^*(\ell)\| \leq B_u \sum_{\ell=0}^n \lambda^{n-\ell} |\bar{d}_i(\ell)|. \quad (4.14)$$

Agora, utiliza-se a desigualdade enunciada a seguir.

Lema 4.1–A. *Supondo-se operação do MU-DM-SWA dentro da RoI no intervalo $[0, n]$, com $n \geq 0$, vale a desigualdade*

$$\sum_{\ell=0}^n \lambda^{n-\ell} |\bar{d}_i(\ell)| < \frac{1}{1-\lambda} \left(B_d + \frac{B_\epsilon}{\gamma} \right), \quad (4.15)$$

sendo

$$B_d = \frac{\sqrt{N_u(4-\beta)}}{\gamma} [1 + (4-\beta)N_u] \left(3 \mathbb{E} [a_{i,R}^2(n)] \right)^{3/2} \quad (4.16)$$

e

$$B_\epsilon = \frac{\varrho}{2} (N_u - 1)(k_1 + 1) \left(3N_u(4-\beta) \mathbb{E} [a_{i,R}^2(n)] \right)^{3/2}. \quad (4.17)$$

Demonstração. Aplicando-se a desigualdade triangular em (3.35), tem-se um limite superior para $|\bar{d}_i(\ell)|$, assim,

$$\sum_{\ell=0}^n \lambda^{n-\ell} |\bar{d}_i(\ell)| \leq \underbrace{\sum_{\ell=0}^n \lambda^{n-\ell} |d_i(\ell)|}_{(I)} + \frac{1}{\gamma} \underbrace{\sum_{\ell=0}^n \lambda^{n-\ell} |\epsilon_i(\ell)|}_{(II)}. \quad (4.18)$$

A seguir, são avaliados os limites superiores para os termos (I) e (II) de (4.18).

- **Termo (I).** Usando a Equação (3.36), resulta

$$|d_i(\ell)|^2 = x_{i,R}^2(\ell) y_{i,R}^2(\ell) + x_{i,I}^2(\ell) y_{i,I}^2(\ell). \quad (4.19)$$

Devido à condição (3.41) dentro da região de interesse, resultam as desigualdades

$$y_{i,R}^2(\ell) \leq |y_i(\ell)|^2 \leq \sum_{i=1}^{N_u} |y_i(\ell)|^2 \leq 3N_u \mathbb{E} [|a_i(n)|^2] = 3N_u(4-\beta) \mathbb{E} [a_{i,R}^2(n)]. \quad (4.20)$$

Aplicando-se a desigualdade triangular em (3.37) e considerando-se o limite superior de (4.20) para $y_{i,R}^2(\ell)$, tem-se

$$|x_{i,R}(n)| \leq \frac{3 \mathbb{E} [a_{i,R}^2(n)] + y_{i,R}^2(n)}{\gamma} \leq \frac{3 \mathbb{E} [a_{i,R}^2(n)] [1 + (4 - \beta)N_u]}{\gamma}, \quad (4.21)$$

em que $\beta = 3$ para o caso real e $\beta = 2$ para o caso não real. No caso em que $\beta = 2$, usando os mesmos argumentos, é possível concluir que o limite superior de $|x_{i,R}(n)|$ vale também para $|x_{i,I}(n)|$. Com os limites superiores de $|x_{i,R}(n)|$ e $|x_{i,I}(n)|$ aplicados em (4.19), obtém-se a desigualdade

$$|d_i(\ell)|^2 \leq \left\{ \frac{3 \mathbb{E} [a_{i,R}^2(n)] [1 + (4 - \beta)N_u]}{\gamma} \right\}^2 |y_i(\ell)|^2. \quad (4.22)$$

Utilizando o limite superior para $|y_i(\ell)|^2$ fornecido por (4.20), conclui-se que

$$0 \leq |d_i(\ell)| \leq B_d < +\infty, \quad (4.23)$$

sendo B_d definido conforme (4.16). Logo, pode-se obter o limite superior para o termo (I) de (4.18), dado por

$$\sum_{\ell=0}^n \lambda^{n-\ell} |d_i(\ell)| \leq B_d \sum_{\ell=0}^n \lambda^{n-\ell} = B_d \frac{1 - \lambda^{n+1}}{1 - \lambda} < \frac{B_d}{1 - \lambda}, \quad (4.24)$$

em que a última desigualdade se justifica pois, para $0 \ll \lambda < 1$ e $n \geq 0$, vale

$$\frac{\lambda^{n+1}}{1 - \lambda} > 0. \quad (4.25)$$

- **Termo (II).** Inicialmente, considerando-se as Equações (3.50) e (3.52), escreve-se $\epsilon_i(\ell)$ de forma fechada, e para $n_0 = 0$, como

$$\epsilon_i(\ell) = \frac{\rho}{2} \lambda^{\ell+1} \widehat{\mathbf{c}}_i^T(-1) \mathbf{y}_i(\ell) + \frac{\rho}{2} (1 - \lambda) \sum_{k=0}^{\ell} \lambda^{\ell-k} y_i(k) \mathbf{y}_i^H(k) \mathbf{y}_i(\ell), \quad (4.26)$$

sendo $\widehat{\mathbf{c}}_i(-1)$ uma condição inicial. Utilizando-se a desigualdade triangular e a desigualdade de Cauchy-Schwarz, obtém-se um limite superior para o módulo de $\epsilon_i(\ell)$, dado por

$$|\epsilon_i(\ell)| \leq \frac{\rho}{2} \lambda^{\ell+1} \|\widehat{\mathbf{c}}_i(-1)\| \|\mathbf{y}_i(\ell)\| + \frac{\rho}{2} (1 - \lambda) \sum_{k=0}^{\ell} \lambda^{\ell-k} |y_i(k)| |\mathbf{y}_i^H(k) \mathbf{y}_i(\ell)|. \quad (4.27)$$

Além disso, para $n_0 = 0$, valem:

- ◊ Segundo a forma de inicialização do algoritmo apresentada na Seção 3.4.2, a norma da condição inicial do vetor que estima a correlação cruzada vale

$$\|\widehat{\mathbf{c}}_i(-1)\| = 0 < +\infty. \quad (4.28)$$

- ◇ Quanto à parcela $|y_i(k)|$, a partir do limite superior para $|y_i(\ell)|^2$ dado em (4.20), pode-se escrever

$$|y_i(k)| \leq \sqrt{3N_u(4 - \beta) E [a_{i,R}^2(n)]} < +\infty. \quad (4.29)$$

- ◇ O módulo do produto interno $|\mathbf{y}_i^H(k)\mathbf{y}_i(\ell)|$, também a partir do limite superior para $|y_i(\ell)|^2$ dado em (4.20), possui o seguinte limite superior

$$|\mathbf{y}_i^H(k)\mathbf{y}_i(\ell)| \leq 3N_u(4 - \beta)(N_u - 1)(k_1 + 1) E [a_{i,R}^2(n)] < +\infty. \quad (4.30)$$

Dessa forma, reescreve-se o limite superior de (4.27) como

$$|\epsilon_i(\ell)| \leq B_\epsilon(1 - \lambda) \sum_{k=0}^{\ell} \lambda^{\ell-k}, \quad (4.31)$$

sendo B_ϵ definido em (4.17).

Finalmente, com a desigualdade (4.31), resulta o seguinte limite superior para o termo (II) da Equação (4.18):

$$\frac{1}{\gamma} \sum_{\ell=0}^n \lambda^{n-\ell} |\epsilon_i(\ell)| \leq \frac{1}{\gamma} B_\epsilon(1 - \lambda) \sum_{\ell=0}^n \sum_{k=0}^{\ell} \lambda^{n-k}. \quad (4.32)$$

Observa-se que

$$0 < \sum_{\ell=0}^n \sum_{k=0}^{\ell} \lambda^{n-k} = \frac{1}{(1 - \lambda)^2} - \frac{\lambda^{n+1}}{(1 - \lambda)^2} - \frac{(n + 1)\lambda^{n+1}}{1 - \lambda} < \frac{1}{(1 - \lambda)^2}, \quad (4.33)$$

em que a última desigualdade se justifica pois, para $0 \ll \lambda < 1$ e $n \geq 0$, valem

$$\frac{\lambda^{n+1}}{(1 - \lambda)^2} > 0 \quad (4.34)$$

e

$$\frac{(n + 1)\lambda^{n+1}}{1 - \lambda} > 0. \quad (4.35)$$

Portanto, para $n \geq 0$, obtém-se o seguinte limite superior para (4.32):

$$\frac{1}{\gamma} \sum_{\ell=0}^n \lambda^{n-\ell} |\epsilon_i(\ell)| < \frac{B_\epsilon}{\gamma(1 - \lambda)}. \quad (4.36)$$

Aplicando-se os limites superiores de (4.24) e (4.36) em (4.18), chega-se à desigualdade (4.15). ■

Utilizando-se o Lema 4.1–A para dentro da RoI e a Equação (4.13) para fora, pode-se provar o seguinte teorema que garante a estabilidade do MU-DM-SWA para $n_0 = 0$, ou seja, com inicialização tanto dentro quanto fora da RoI.

Teorema 4.1–A. *As normas $\|\mathbf{w}_i(n)\|$ de (4.12) e (4.13), no intervalo $[0, n]$ e com $n \geq 0$, satisfazem, respectivamente*

$$\|\mathbf{w}_i(n)\| < \alpha_1^{-1} \left[\lambda^{n+1} \delta^{-1} \|\mathbf{w}_i(-1)\| + \frac{B_u}{(1-\lambda)} \left(B_d + \frac{B_\epsilon}{\gamma} \right) \right] < +\infty \quad (4.37)$$

(dentro da RoI),

$$\|\mathbf{w}_i(n)\| \leq \lambda^{n+1} \alpha_1^{-1} \delta^{-1} \|\mathbf{w}_i(-1)\| < +\infty \quad (4.38)$$

(fora da RoI),

sendo B_u , B_d e B_ϵ definidos respectivamente em (4.9), (4.16) e (4.17). Além disso, para $n \rightarrow +\infty$ e considerando-se $0 \ll \lambda < 1$, tem-se

$$\|\mathbf{w}_i(n)\| < \frac{B_u \alpha_1^{-1}}{(1-\lambda)} \left(B_d + \frac{B_\epsilon}{\gamma} \right) < +\infty \quad (4.39)$$

(dentro da RoI),

$$\|\mathbf{w}_i(n)\| \rightarrow 0 \quad \text{exponencialmente rápido} \quad (4.40)$$

(fora da RoI).

Em outras palavras, a norma $\|\mathbf{w}_i(n)\|$ é limitada dentro da RoI e, fora da RoI, o algoritmo é contrativo, visto que possui um ponto de equilíbrio exponencialmente estável em $\mathbf{0}$.

Para os dois modos de operação, a demonstração é direta. Para o caso em que o algoritmo opera dentro da RoI, a partir de (4.12) e utilizando-se (4.14) e (4.15), chega-se a (4.37). Para fora da RoI, basta fazer $n \rightarrow +\infty$ na Equação (4.38).

4.1.2 Transição entre modos de operação

Agora, considera-se que o algoritmo transita entre modos pela primeira vez em um instante $n_0 > 0$. Nessa situação, valem:

- Assumindo-se a Hipótese 4.1–A, a partir da Equação (4.8), resulta $\|\widehat{\mathbf{R}}(n_0 - 1)\| = \alpha_2$. Além disso, continua valendo $\|\widehat{\mathbf{R}}^{-1}(n)\| \leq \alpha_1^{-1}$.
- Considerando-se o Teorema 4.1–A, pode-se afirmar que $\|\mathbf{w}_i(n_0 - 1)\| < +\infty$.

Portanto, para um intervalo $[n_0, n]$, com $n \geq n_0$, em que o algoritmo mantém o modo de operação para o qual transitou, as Equações (4.5) e (4.7) podem ser reescritas, respectiva-

mente, como

$$\|\mathbf{w}_i(n)\| \leq \alpha_1^{-1} \left[\lambda^{n-n_0+1} \alpha_2 \|\mathbf{w}_i(n_0 - 1)\| + \sum_{\ell=n_0}^n \lambda^{n-\ell} |\bar{d}_i(\ell)| \|\mathbf{u}^*(\ell)\| \right] \quad (4.41)$$

(fora para dentro da RoI),

$$\|\mathbf{w}_i(n)\| \leq \lambda^{n-n_0+1} \alpha_1^{-1} \alpha_2 \|\mathbf{w}_i(n_0 - 1)\| \quad (4.42)$$

(dentro para fora da RoI).

Quando o algoritmo passa de fora para dentro da RoI, devido ao seu funcionamento, vale $\|\hat{\mathbf{c}}_i(n_0 - 1)\| = 0$ e a prova de que $\|\mathbf{w}_i(n)\| < +\infty$ para um intervalo $[n_0, n]$, com $n \geq n_0$, é análoga à do caso em que $n_0 = 0$. Quando o algoritmo transita de dentro para fora da RoI, as provas de que $\|\mathbf{w}_i(n)\| < +\infty$ para um intervalo $[n_0, n]$ e de que $\|\mathbf{w}_i(n)\| \rightarrow 0$ para $n \rightarrow \infty$ são também análogas às feitas quando $n_0 = 0$.

As provas podem ser facilmente estendidas para qualquer sequência possível de transições entre modos, desde o instante inicial $n = 0$. Com isso, comprova-se que, sob a hipótese de excitação persistente, a norma do vetor de coeficientes $\mathbf{w}_i(n)$ do MU-DM-SWA, sob precisão infinita e com $i = 1, 2, \dots, N_u$, mantém-se limitada. Portanto, garante-se que, nesse sentido, o MU-DM-SWA não diverge.

Em contrapartida, a análise determinista feita nessa seção não permite avaliar o efeito da variação dos parâmetros do algoritmo em seu desempenho. Por isso, a seguir é feita a análise em regime do MU-DM-SWA, dentro da RoI, que permite verificar a influência dos parâmetros do algoritmo sobre o seu desempenho.

4.2 ANÁLISE EM REGIME

A análise em regime (*i.e.*, análise estocástica) permite avaliar a escolha de parâmetros para garantir o funcionamento do algoritmo com um desempenho desejado. Utilizando os argumentos de conservação de energia de (RUPP; SAYED, 1996) e (SAYED, 2008, Cap. 21), além dos resultados de análises em regime do CMA, SWA e suas versões multiusuário (FIJALKOW; MANLOVE; JOHNSON, 1998; MAI; SAYED, 2000; SILVA; MIRANDA, 2004; SAYED, 2008; MENDES-FILHO; MIRANDA; SILVA, 2012; LUO; CHAMBERS, 2002; SILVA; MIRANDA, 2005), nesta seção é obtido um modelo teórico, em regime, para o erro quadrático médio em excesso do MU-DM-SWA, operando dentro da RoI e em um ambiente não estacionário. A análise é iniciada com a definição do indicador de desempenho e a descrição das hipóteses simplificadoras que são utilizadas. Posteriormente, é obtida uma expressão analítica para o erro quadrático médio em excesso em regime do algoritmo. Finalmente, a expressão obtida é validada por meio de simulações com médias de conjunto sob dois cenários distintos.

4.2.1 Indicadores de desempenho e hipóteses iniciais

Para se medir o desempenho de cada equalizador em regime, o erro quadrático médio em excesso (EMSE, do inglês, *excess mean square error*) do i -ésimo equalizador é definido como

$$\zeta_i \triangleq \lim_{n \rightarrow +\infty} \mathbb{E} \left[|e_{a,i}(n)|^2 \right], \text{ com } i = 1, 2, \dots, N_u, \quad (4.43)$$

em que

$$e_{a,i}(n) = \mathbf{u}^\top(n) \tilde{\mathbf{w}}_i(n-1) \quad (4.44)$$

é o erro *a priori*

$$\tilde{\mathbf{w}}_i(n) = \mathbf{w}_{o,i}(n) - \mathbf{w}_i(n) \quad (4.45)$$

é o vetor de ponderação de erro e $\mathbf{w}_{o,i}(n)$ é a solução ótima do problema de mínimos quadrados.

Nesta análise, lança-se mão das seguintes hipóteses simplificadoras:

Hipótese 4.2–A. Para todo n , os símbolos transmitidos por cada fonte i , denotados por $a_i(n)$, têm média nula e satisfazem à condição de simetria da constelação, ou seja,

$$\mathbb{E} \left[a_i(n) |a_i(n)|^{2m} \right] = 0 \quad \forall m \in \mathbb{N}. \quad (4.46)$$

Se os dados forem complexos, então $\mathbb{E} [a_i^2(n)] = 0$ e as partes real e imaginária de $a_i(n)$ são não correlacionadas.

Além disso, considera-se que o processo $a_i(n)$ é subgaussiano, independente e identicamente distribuído (i.i.d.) e suas estatísticas não variam com o tempo, como é o caso da maioria das constelações usadas em comunicações digitais (SAYED, 2008; SHALVI; WEINSTEIN, 1993; PROAKIS; SALEHI, 2007). Por fim, considera-se que $a_i(n)$ e $a_m(n)$, com $i \neq m$, possuem as mesmas estatísticas e são independentes entre si.

Hipótese 4.2–B. Considera-se que os sinais transmitidos $a_i(n - \tau_i)$, com $i = 1, 2, \dots, N_u$ são relacionados ao vetor regressor $\mathbf{u}(n)$ pela expressão

$$e^{j\theta_i} a_i(n - \tau_i) = \mathbf{u}^\top(n) \mathbf{w}_{o,i}(n-1) + v_i(n), \quad (4.47)$$

em que θ_i são rotações complexas inerentes à equalização cega, múltiplas de 90° , e os sinais $v_i(n) = v_{i,R}(n) + jv_{i,I}(n)$ fazem o papel de ruído incluído para levar em conta as imperfeições do modelo considerado com a solução $\mathbf{w}_{o,i}(n)$. Para todo n , o ruído $v_i(n)$ têm média nula e simetria circular, *i.e.*,

$$\mathbb{E} \left[v_i(n) |v_i(n)|^{2m} \right] = 0 \quad \forall m \in \mathbb{N}. \quad (4.48)$$

Se os dados forem complexos, então $\mathbb{E} [v_i^2(n)] = 0$ e as partes real e imaginária de $v_i(n)$ são não correlacionadas.

Além disso, considera-se que o processo $v_i(n)$ é i.i.d. e gaussiano, de modo que

$$\mathbb{E} \left[v_{i,R}^6(n) \right] = 15 \mathbb{E}^3 \left[v_{i,R}^2(n) \right] \quad (4.49)$$

e

$$\mathbb{E} \left[v_{i,R}^4(n) \right] = 3 \mathbb{E}^2 \left[v_{i,R}^2(n) \right]. \quad (4.50)$$

Considera-se ainda que $v_i(n)$ e $v_m(n)$, com $i \neq m$, são não correlacionados entre si.

Para tornar a análise de desempenho mais tratável, assume-se que a sequências $\{\mathbf{u}(n)\}$ e $\{v_i(n)\}$ são estacionárias e que os processos $v_i(n)$ são independentes de $\{\mathbf{u}(\ell)\}$, com $\ell \leq n$ (e não somente não correlacionados). Também é assumida a independência entre $v_i(n)$, $e^{j\theta_i} a_i(n - \tau_i)$ e $e_{a,i}(n)$, bem como a não correlação entre as partes real e imaginária de $e_{a,i}(n)$, em regime.

Hipótese 4.2–C. Em um ambiente não estacionário, as soluções ótimas seguem um modelo de *random-walk* (SAYED, 2008), ou seja,

$$\mathbf{w}_{o,i}(n+1) = \mathbf{w}_{o,i}(n) + \mathbf{q}_i(n), \quad (4.51)$$

em que $\mathbf{q}_i(n)$ são vetores aleatórios i.i.d. de média nula, com matriz de covariância positiva definida dada por $\mathbf{Q}_i = \mathbb{E} [\mathbf{q}_i^*(n) \mathbf{q}_i^T(n)]$, considerados independentes das condições iniciais $\{\mathbf{w}_{o,i}(-1), \mathbf{w}_i(-1)\}$ e também dos sinais $\{\mathbf{u}(\ell), e^{j\theta_i} a_i(\ell - \tau_i)\}$ para todo $\ell \leq n$, com $i = 1, 2, \dots, N_u$.

Embora seja possível concluir, a partir desse modelo, que a matriz de covariância de $\mathbf{w}_{o,i}(n)$ cresce indefinidamente, ele tem sido utilizado na literatura para simplificar a análise de sistemas adaptativos em ambientes não estacionários (SAYED, 2008). Para o caso de um equalizador de taxa de símbolos, as soluções ótimas $\mathbf{w}_{o,i}(n)$ são as próprias soluções de Wiener. Por outro lado, para um equalizador com diversidade temporal e espacial, sem a presença de ruído, os vetores $\mathbf{w}_{o,i}(n)$ representam as soluções de *zero-forcing* que levam à equalização perfeita (TREICHLER; FIJALKOW; JOHNSON, 1996; MAI; SAYED, 2000).

Hipótese 4.2–D. Em regime, considera-se que a norma Euclidiana ponderada ao quadrado do vetor regressor, dada por $\|\mathbf{u}(n)\|_{\widehat{\mathbf{R}}^{-1}(n)}^2 \triangleq \mathbf{u}^H(n) \widehat{\mathbf{R}}^{-1}(n) \mathbf{u}(n)$, é independente de $e_{a,i}(n)$. Na análise em regime do algoritmo RLS, essa hipótese é conhecida como o princípio de separação (SAYED, 2008). Uma hipótese similar foi adotada nas análises em regime do CMA (MAI; SAYED, 2000) e do MMA (MENDES-FILHO; MIRANDA; SILVA, 2012).

Hipótese 4.2–E. O fator de esquecimento λ é suficientemente próximo de um, tal que a estimativa da inversa da matriz de covariância $\widehat{\mathbf{R}}^{-1}(n)$, em regime, pode ser aproximada por sua média, *i.e.*, $\mathbb{E} [\widehat{\mathbf{R}}^{-1}(n)] \approx (1 - \lambda) \mathbf{R}_u^{-1}$, em que

$$\mathbf{R}_u = \mathbb{E} [\mathbf{u}^*(n) \mathbf{u}^T(n)] \quad (4.52)$$

é a matriz de covariância do vetor regressor (SAYED, 2008).

Hipótese 4.2–F. O número de coeficientes de cada equalizador é muito menor que o comprimento da janela exponencial, *i.e.*, $M \ll 1/(1 - \lambda)$.

Hipótese 4.2–G. Para $n \rightarrow +\infty$, $a_i(n - \tau_i)$ e $v_i(n)$ são independentes de $e_{a,m}(n)$, com $i \neq m$. Isso ocorre pois observa-se que as flutuações em regime do erro *a priori* são insensíveis aos sinais transmitidos (MAI; SAYED, 2000; YOUSEF; SAYED, 1999; LUO; CHAMBERS, 2002). Além disso, também para $n \rightarrow +\infty$, vale $E[e_{a,i}(n) e_{a,m}^*(n - k)] = 0$, com $i \neq m$. De acordo com (LUO; CHAMBERS, 2002), essa suposição é realista quando o equalizador é longo o suficiente de modo a garantir que a sua saída converge assintoticamente para a correspondente estimativa da fonte, em regime. Como consequência, em regime, considera-se que é válida a expressão $E[y_i(n) y_m^*(n - k)] = 0$.

Hipótese 4.2–H. Para $n \rightarrow +\infty$, momentos de ordem superior ou igual a 3 de $e_{a,i}(n)$ são muito próximos de zero e podem ser desprezados.

Hipótese 4.2–I. Para $n \rightarrow +\infty$, a esperança $E[|e_{a,i}(n)|^2]$ pode ser desprezada face à esperança $E[|a_i(n)|^2]$, pois $E[|a_i(n)|^2] \gg E[|e_{a,i}(n)|^2]$. Consequentemente, em regime, também vale a relação para as partes reais, ou seja, $E[a_{i,R}^2(n)] \gg E[e_{a,i,R}^2(n)]$.

As Hipóteses de 4.2–D até 4.2–I tornam a análise mais tratável. Em compensação, não se deve esperar uma concordância razoável entre a análise e a simulação quando o fator de esquecimento λ não for suficientemente próximo de um.

4.2.2 Considerações sobre o modelo completo

O modelo (4.47), apresentado na Hipótese 4.2–B, é comumente utilizado no contexto de identificação de sistemas, sendo denominado modelo de regressão linear (SAYED, 2008). No entanto, ele pode ser utilizado na análise de algoritmos adaptativos em uma configuração de equalização, como mostrado em (RUPP, 2011). Esse modelo permite escrever a saída do i -ésimo equalizador $y_i(n)$ em termos de seus erros *a priori* $e_{a,i}(n)$. Portanto, utilizando (4.45), (4.44) e a Hipótese 4.2–B, obtém-se

$$\begin{aligned} y_i(n) &= \mathbf{u}^T(n) \mathbf{w}_i(n - 1) = \mathbf{u}^T(n) [\mathbf{w}_{o,i}(n - 1) - \tilde{\mathbf{w}}_i(n - 1)] \\ &= e^{j\theta_i} a_i(n - \tau_i) - e_{a,i}(n) - v_i(n). \end{aligned} \quad (4.53)$$

Esse modelo mais sofisticado foi uma contribuição da análise de algoritmos adaptativos cegos de (MENDES-FILHO; MIRANDA; SILVA, 2012). Trata-se de uma generalização do modelo utilizado na análise em regime do CMA de (MAI; SAYED, 2000), em que a hipótese de equalização perfeita foi considerada. Nota-se que, no caso de equalização perfeita, $v_i(n) = 0 \forall i, n$.

O modelo completo prevê que diferentes equalizadores podem ter EMSEs distintos, a depender da “proximidade” com que cada um se encontra da solução ótima. De fato, é

possível verificar com base em simulações que, em muitas situações nas quais a relação sinal-ruído é baixa ou a equalização perfeita não é atingida, os equalizadores podem apresentar EMSEs diferentes em regime. Embora seja mais complicado, o modelo completo torna a análise em regime mais completa.

Para aplicar esse modelo, é preciso ter conhecimento dos momentos $E[|v_i(n)|^2]$ de cada equalizador. A seguir, o procedimento para a obtenção desses momentos é detalhado. Inicialmente, tomando-se a esperança do quadrado da Equação (4.47) em regime, novamente sob a Hipótese 4.2–B, resulta

$$E[|a_i(n)|^2] = E[|v_i(n)|^2] + E[\mathbf{w}_{o,i}^H(n-1) \mathbf{u}^*(n) \mathbf{u}^T(n) \mathbf{w}_{o,i}(n-1)], \quad (4.54)$$

Considerando-se o modelo de *random-walk* apresentado na Hipótese 4.2–C, nota-se que vale $E[\mathbf{w}_{o,i}(n-1)] = \mathbf{w}_{o,i}(-1)$. Logo, em regime, utiliza-se a aproximação

$$E[|v_i(n)|^2] \approx E[|a_i(n)|^2] - \mathbf{w}_{o,i}^H(-1) \mathbf{R}_u \mathbf{w}_{o,i}(-1), \quad (4.55)$$

em que \mathbf{R}_u é dada pela Equação (4.52).

A partir da Equação (4.55), os momentos $E[|v_i(n)|^2]$ utilizados no modelo completo podem ser calculados, desde que sejam conhecidos \mathbf{R}_u e as condições iniciais do vetor de coeficientes ótimos, *i.e.*, $\mathbf{w}_{o,i}(-1)$. A solução ótima $\mathbf{w}_{o,i}(-1)$ é encontrada resolvendo-se o sistema linear de equações normais, dado por

$$\mathbf{R}_u \mathbf{w}_{o,i}(-1) = \mathbf{p}_i, \quad (4.56)$$

em que

$$\mathbf{p}_i = E[e^{j\theta_i} a_i(n - \tau_i) \mathbf{u}^*(n)] \quad (4.57)$$

é o vetor de correlação cruzada entre a i -ésima fonte recuperada e o vetor regressor. Tendo-se conhecimento dos coeficientes do canal de comunicação MIMO cujo modelo é mostrado na Figura 2, é possível calcular diretamente \mathbf{R}_u e \mathbf{p}_i . A partir do modelo para $\mathbf{u}(n)$ dado pela Equação (2.26), as Equações (4.52) e (4.57), correspondentes a \mathbf{R}_u e \mathbf{p}_i , respectivamente, resultam em

$$\mathbf{R}_u = \bar{\mathbf{H}}^H \mathbf{R}_a \bar{\mathbf{H}} + \mathbf{R}_\eta, \quad (4.58)$$

$$\mathbf{p}_i = \bar{\mathbf{H}}^H E[e^{j\theta_i} a_i(n - \tau_i) \mathbf{a}^*(n)], \quad (4.59)$$

sendo $\mathbf{R}_a = E[\mathbf{a}^*(n) \mathbf{a}^T(n)]$ e $\mathbf{R}_\eta = E[\boldsymbol{\eta}^*(n) \boldsymbol{\eta}^T(n)]$. A partir das Equações (4.58) e (4.59), calcula-se $\mathbf{w}_{o,i}(-1)$ resolvendo-se a Equação (4.56). Com $\mathbf{w}_{o,i}(-1)$ e \mathbf{R}_u , encontra-se $E[|v_i(n)|^2]$ lançando-se mão da Equação (4.55).

A partir das Equações (4.56) a (4.59), observa-se que o vetor de correlação cruzada \mathbf{p}_i depende do atraso τ_i e da rotação complexa θ_i da fonte recuperada pelo i -ésimo equalizador, com $i = 1, 2, \dots, N_u$. Logo, a solução ótima $\mathbf{w}_{o,i}(-1)$ encontrada na resolução

da Equação (4.56), bem como a variância $E[|v_i(n)|^2]$ dada pela Equação (4.55), também dependem desses parâmetros da fonte recuperada. Portanto, é preciso saber qual equalizador recupera cada fonte para então determinar $E[|v_i(n)|^2]$ de cada equalizador. É importante ressaltar que, para se fazer uso do modelo cuja dedução é apresentada na Seção 4.2.3 a seguir, deve-se saber de antemão:

- qual é a fonte recuperada por cada equalizador;
- a rotação complexa de cada fonte recuperada;
- o atraso de cada fonte recuperada.

Para ter esse conhecimento, é preciso simular o algoritmo algumas vezes com os parâmetros a serem utilizados na validação do modelo. Isso deve ser feito de modo a colher as informações necessárias para a determinação do $E[|v_i(n)|^2]$ de cada equalizador. Essas três informações a respeito da fonte recuperada dependem da inicialização do algoritmo e da geometria da função custo que está sendo minimizada. Por conveniência, e sem perda de generalidade, assume-se na análise que a ambiguidade de permutação não ocorre, ou seja, que o i -ésimo equalizador busca recuperar a fonte $a_i(n)$.

4.2.3 Expressão para o erro quadrático médio em excesso

Uma relação de recorrência para o vetor de ponderação de erro pode ser obtida subtraindo-se ambos os lados da Equação (3.33) de $\mathbf{w}_{o,i}(n)$ e utilizando-se (4.51), *i.e.*,

$$\tilde{\mathbf{w}}_i(n) = \tilde{\mathbf{w}}_i(n-1) - \widehat{\mathbf{R}}^{-1}(n) \mathbf{u}^*(n) [\bar{d}_i(n) - y_i(n)] + \mathbf{q}_i(n). \quad (4.60)$$

Agora, as normas ponderadas de ambos os lados de (4.60) são calculadas usando $\widehat{\mathbf{R}}(n)$ como matriz de ponderação e tomando-se a esperança com $n \rightarrow +\infty$. Sob a Hipótese 4.2-C, chega-se a

$$\begin{aligned} E \left[\|\tilde{\mathbf{w}}_i(n)\|_{\widehat{\mathbf{R}}(n)}^2 \right] &\approx E \left[\|\tilde{\mathbf{w}}_i(n-1)\|_{\widehat{\mathbf{R}}(n)}^2 \right] + E \left[\|\mathbf{u}^*(n)\|_{\widehat{\mathbf{R}}^{-1}(n)}^2 |\bar{d}_i(n) - y_i(n)|^2 \right] \\ &\quad - 2 \operatorname{Re} \left\{ E \left[e_{a,i}^*(n) [\bar{d}_i(n) - y_i(n)] \right] \right\} + E \left[\|\mathbf{q}_i(n)\|_{\widehat{\mathbf{R}}(n)}^2 \right]. \end{aligned} \quad (4.61)$$

Assumindo-se que os equalizadores operam em regime, tem-se

$$E \left[\|\tilde{\mathbf{w}}_i(n)\|_{\widehat{\mathbf{R}}(n)}^2 \right] \approx E \left[\|\tilde{\mathbf{w}}_i(n-1)\|_{\widehat{\mathbf{R}}(n)}^2 \right], \text{ para } n \rightarrow +\infty.$$

Sob essas condições, (4.61) se reduz a

$$E \left[\|\mathbf{u}^*(n)\|_{\widehat{\mathbf{R}}^{-1}(n)}^2 |\bar{d}_i(n) - y_i(n)|^2 \right] + E \left[\|\mathbf{q}_i(n)\|_{\widehat{\mathbf{R}}(n)}^2 \right] = 2 \operatorname{Re} \left\{ E \left[e_{a,i}^*(n) [\bar{d}_i(n) - y_i(n)] \right] \right\}. \quad (4.62)$$

Como (4.60) possui a mesma estrutura do algoritmo RLS, a relação de variâncias (4.62) obtida é similar àquela de (SAYED, 2008, Eq.(21.42)). Lançando-se mão de (4.53), (4.44), e definindo-se

$$\varepsilon_i(n) \triangleq \bar{d}_i(n) - [e^{j\theta_i} a_i(n - \tau_i) - v_i(n)], \quad (4.63)$$

o erro $\bar{d}_i(n) - y_i(n)$ pode ser reescrito como

$$\bar{d}_i(n) - y_i(n) = e_{a,i}(n) + \varepsilon_i(n). \quad (4.64)$$

Da Equação (4.47), é possível notar que $e^{j\theta_i} a_i(n - \tau_i) - v_i(n) = \mathbf{u}^T(n) \mathbf{w}_{o,i}(n-1)$. Interpreta-se $\bar{d}_i(n)$ como o sinal desejado, sendo que $\varepsilon_i(n)$ faz o papel de uma perturbação similar àquela considerada na filtragem adaptativa supervisionada (ver, *e.g.*, (SAYED, 2008, p. 231)). No entanto, nesse caso, $\bar{d}_i(n)$ é uma estimativa não linear de $e^{j\theta_i} a_i(n - \tau_i)$ e depende de $y_i(n)$. Portanto, em contraste com o caso supervisionado, $\varepsilon_i(n)$ e $\mathbf{u}(n)$ são correlacionados e $\varepsilon_i(n)$ e $e_{a,i}(n)$ não são independentes.

Substituindo-se (4.64) em (4.62), obtém-se

$$\mathbb{E} \left[\|\mathbf{u}^*(n)\|_{\hat{\mathbf{R}}^{-1}(n)}^2 |\varepsilon_i(n) + e_{a,i}(n)|^2 \right] + \mathbb{E} \left[\|\mathbf{q}_i(n)\|_{\hat{\mathbf{R}}(n)}^2 \right] \approx 2 \operatorname{Re} \left\{ \mathbb{E} \left[e_{a,i}^*(n) [\varepsilon_i(n) + e_{a,i}(n)] \right] \right\}. \quad (4.65)$$

Sob as Hipóteses 4.2-D e 4.2-E, (4.65) se reduz a

$$(1 - \lambda)M\zeta_i + 2(1 - \lambda)M \operatorname{Re} \left\{ \mathbb{E} \left[e_{a,i}^*(n) \varepsilon_i(n) \right] \right\} + (1 - \lambda)M \mathbb{E} \left[|\varepsilon_i(n)|^2 \right] + \frac{\operatorname{Tr}(\mathbf{Q}_i \mathbf{R}_u)}{1 - \lambda} \approx 2\zeta_i + 2 \operatorname{Re} \left\{ \mathbb{E} \left[e_{a,i}^*(n) \varepsilon_i(n) \right] \right\}. \quad (4.66)$$

em que $\operatorname{Tr}(\cdot)$ indica o traço de uma matriz. Sob a Hipótese 4.2-F, os dois primeiros termos do lado esquerdo de (4.66) podem ser desprezados em relação aos dois primeiros termos do lado direito, resultando em

$$(1 - \lambda)M \mathbb{E} \left[|\varepsilon_i(n)|^2 \right] + \frac{\operatorname{Tr}(\mathbf{Q}_i \mathbf{R}_u)}{1 - \lambda} \approx 2\zeta_i + 2 \operatorname{Re} \left\{ \mathbb{E} \left[e_{a,i}^*(n) \varepsilon_i(n) \right] \right\}. \quad (4.67)$$

Para completar a análise, falta obter expressões analíticas para os momentos $\mathbb{E} \left[e_{a,i}^*(n) \varepsilon_i(n) \right]$ e $\mathbb{E} \left[|\varepsilon_i(n)|^2 \right]$. A dedução de expressões aproximadas para esses dois momentos, lançando-se mão das Hipóteses 4.2-A, 4.2-B, 4.2-G, 4.2-H e 4.2-I, é feita no Apêndice B. Os resultados obtidos são reproduzidos a seguir.

Segundo o Resultado 2.3-I do Apêndice B, o momento $\mathbb{E} \left[e_{a,i}^*(n) \varepsilon_i(n) \right]$ pode ser aproximado por

$$\mathbb{E} \left[e_{a,i}^*(n) \varepsilon_i(n) \right] \approx \frac{\alpha_i \zeta_i}{\gamma}, \quad (4.68)$$

sendo

$$\alpha_i = \alpha_{v_i} + \frac{\alpha_{m_i}}{2}, \quad (4.69)$$

$$\alpha_{v_i} = 3 \mathbb{E}[v_{i,R}^2(n)], \quad (4.70)$$

$$\alpha_{m_i} = \varrho(4 - \beta)(1 - \lambda)(k_1 + 1) \left[(N_u - 1) \mathbb{E}[a_{i,R}^2(n)] + S_{2,i} \right], \quad (4.71)$$

$$S_{2,i} = \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \mathbb{E}[v_{m,R}^2(n)], \quad (4.72)$$

com $\beta = 3$ para o caso real e $\beta = 2$ para o caso não real. O termo α_{v_i} é nulo quando ocorre equalização perfeita. Por sua vez, o termo α_{m_i} é não nulo somente no caso de desconvolução cega para canais MIMO.

Segundo o Resultado 2.3–J do Apêndice B, o momento $\mathbb{E}[|\varepsilon_i(n)|^2]$ pode ser aproximado por

$$\mathbb{E}[|\varepsilon_i(n)|^2] \approx \frac{(4 - \beta)\psi_i + \chi_i \zeta_i}{\gamma^2}, \quad (4.73)$$

em que

$$\psi_i = \psi_{a_i} + \psi_{v_i} + \psi_{m_i}, \quad (4.74)$$

$$\psi_{a_i} = \mathbb{E}[a_{i,R}^6(n)] - \mathcal{R} \mathbb{E}[a_{i,R}^4(n)], \quad (4.75)$$

$$\psi_{v_i} = \mathbb{E}[v_{i,R}^2(n)] \left\{ \mathcal{R}^2 + 3 \mathbb{E}[a_{i,R}^4(n)] + 45 \mathbb{E}[a_{i,R}^2(n)] \mathbb{E}[v_{i,R}^2(n)] - 6\mathcal{R} \mathbb{E}[v_{i,R}^2(n)] + 15 \mathbb{E}^2[v_{i,R}^2(n)] \right\}, \quad (4.76)$$

$$\psi_{m_i} = \left\{ 6 \mathbb{E}[a_{i,R}^2(n)] - \mathcal{R} + 3 \mathbb{E}[v_{i,R}^2(n)] \right\} \mathbb{E}[v_{i,R}^2(n)] \alpha_{m_i} + \frac{\varrho^2}{4} (1 - \lambda)^2 \left\{ \mathbb{E}[a_{i,R}^2(n)] + \mathbb{E}[v_{i,R}^2(n)] \right\} \phi_{m_i}, \quad (4.77)$$

$$\chi_i = \chi_{a_i} + \chi_{v_i} + \chi_{m_i}, \quad (4.78)$$

$$\chi_{a_i} = 9 \left\{ \mathbb{E}[a_{i,R}^4(n)] - 2 \mathbb{E}^2[a_{i,R}^2(n)] \right\}, \quad (4.79)$$

$$\chi_{v_i} = 6 \mathbb{E}[v_{i,R}^2(n)] \left\{ 18 \mathbb{E}[a_{i,R}^2(n)] - \mathcal{R} \right\}, \quad (4.80)$$

$$\chi_{m_i} = 6 \mathbb{E}[a_{i,R}^2(n)] \alpha_{m_i}. \quad (4.81)$$

Adicionalmente, o fator ϕ_{m_i} , que aparece na Equação (4.77), vale

$$\phi_{m_i} = (4 - \beta)^2 (k_1 + 1) \frac{\lambda^2}{1 - \lambda^2} \left[(N_u - 1) \mathbb{E}^2[a_{i,R}^2(n)] + 2 \mathbb{E}[a_{i,R}^2(n)] S_{2,i} + \Lambda_{4,i} \right], \quad (4.82)$$

sendo

$$\Lambda_{4,i} = \sum_{\substack{m=1 \\ k \neq i}}^{N_u} \mathbb{E}^2[v_{m,R}^2(n)]. \quad (4.83)$$

Nota-se que, a partir das Equações (4.74) e (4.78), respectivamente, os termos ψ_{v_i} e χ_{v_i} são nulos quando ocorre equalização perfeita. Além disso, os termos ψ_{m_i} e χ_{m_i} são não nulos somente no caso de desconvolução cega MIMO.

Finalmente, voltando à Equação (4.67) e substituindo nela os momentos aproximados $E[e_{a,i}^*(n)\varepsilon_i(n)]$ e $E[|\varepsilon_i(n)|^2]$ pelas respectivas expressões (4.68) e (4.73), resulta

$$(1 - \lambda)M \frac{(4 - \beta)\psi_i + \chi_i\zeta_i}{\gamma^2} + \frac{\text{Tr}(\mathbf{Q}_i\mathbf{R}_u)}{1 - \lambda} \approx 2\zeta_i + 2\frac{\alpha_i\zeta_i}{\gamma}. \quad (4.84)$$

Manipulando-se (4.84), chega-se à expressão desejada para o erro quadrático médio em excesso ζ_i , dada por

$$\zeta_i \approx \frac{(1 - \lambda)M(4 - \beta)\psi_i + \gamma^2(1 - \lambda)^{-1} \text{Tr}(\mathbf{Q}_i\mathbf{R}_u)}{2\gamma^2 + 2\gamma\alpha_i - (1 - \lambda)M\chi_i}. \quad (4.85)$$

Cabe ressaltar que essa expressão é para o erro quadrático médio em excesso em regime ζ_i do i -ésimo equalizador. Por outro lado, não necessariamente esse equalizador irá recuperar a i -ésima fonte. Além disso, a expressão vale para o caso de desconvolução cega de canais MIMO e leva em conta casos em que não ocorre equalização perfeita, representados pelo termo $v_i(n)$. No entanto, para alguns casos particulares já apresentados na literatura, essa expressão pode ser simplificada. Alguns desses cenários em que (4.85) pode ser reduzida são elencados a seguir.

- **Canal SISO com equalização quase perfeita.** Nesse caso, os termos de (4.85) correspondentes ao caso MIMO são nulos, ou seja, $\alpha_{m_i} = \psi_{m_i} = \chi_{m_i} = 0 \forall i$. Como a equalização é quase perfeita, pode-se considerar que $v_i(n) \approx 0 \forall i, n$. Disso, decorre que os termos dependentes de momentos de $v_i(n)$, *i.e.*, α_{v_i} , ψ_{v_i} e χ_{v_i} são nulos. Consequentemente, a Equação (4.85) reduz-se a

$$\zeta_i \approx \frac{(1 - \lambda)M(4 - \beta)\psi_{a_i} + \gamma^2(1 - \lambda)^{-1} \text{Tr}(\mathbf{Q}_i\mathbf{R}_u)}{2\gamma^2 - (1 - \lambda)M\chi_{a_i}}. \quad (4.86)$$

- **Canal SISO sem equalização perfeita.** Nessa situação, somente os termos de (4.85) correspondentes ao caso MIMO são nulos, ou seja, $\alpha_{m_i} = \psi_{m_i} = \chi_{m_i} = 0 \forall i$. Portanto, a Equação (4.85) reduz-se a

$$\zeta_i \approx \frac{(1 - \lambda)M(4 - \beta)(\psi_{a_i} + \psi_{v_i}) + \gamma^2(1 - \lambda)^{-1} \text{Tr}(\mathbf{Q}_i\mathbf{R}_u)}{2\gamma^2 + 2\gamma\alpha_{v_i} - (1 - \lambda)M(\chi_{a_i} + \chi_{v_i})}. \quad (4.87)$$

- **Canal MIMO com equalização quase perfeita para $N_u = 2$ fontes.** Como a equalização é quase perfeita, pode-se considerar que $v_i(n) \approx 0 \forall i, n$ e os termos α_{v_i} , ψ_{v_i} e χ_{v_i} são nulos. Além disso, levando-se em conta $N_u = 2$ fontes, decorre imediatamente que $N_u - 1 = 1$ e os termos $S_{2,i}$ e $\Lambda_{4,i}$ deixam de ser somatórios. Com isso, a Equação (4.85) reduz-se a

$$\zeta_i \approx \frac{(1 - \lambda)M(4 - \beta)(\psi_{a_i} + \psi_{m_i}) + \gamma^2(1 - \lambda)^{-1} \text{Tr}(\mathbf{Q}_i\mathbf{R}_u)}{2\gamma^2 + 2\gamma\alpha_{m_i} - (1 - \lambda)M(\chi_{a_i} + \chi_{m_i})}, \quad (4.88)$$

com

$$S_{2,i} = \mathbb{E}[|v_{m,\mathcal{R}}(n)|^2], \quad m \neq i, \quad (4.89)$$

$$\Lambda_{4,i} = \mathbb{E}^2[v_{m,\mathcal{R}}^2(n)], \quad m \neq i, \quad (4.90)$$

e

$$\phi_{m_i} = (4 - \beta)^2(k_1 + 1) \frac{\lambda^2}{1 - \lambda^2} \left[\mathbb{E}^2[a_{i,\mathcal{R}}^2(n)] + 2 \mathbb{E}[a_{i,\mathcal{R}}^2(n)] S_{2,i} + \Lambda_{4,i} \right]. \quad (4.91)$$

4.2.4 Validação do modelo

Nesta seção, são apresentados resultados de simulação numérica com o objetivo de validar o modelo deduzido na Seção 4.2.3 para o erro quadrático médio em excesso do MU-DM-SWA, operando dentro da RoI e em regime.

A comparação é feita sob dois cenários distintos baseados no modelo da Figura 2. Em ambas as situações, as fontes são sinais 16-QAM de valores $\{a_{\mathcal{R}} + ja_i\}$, sendo $a_{\mathcal{R}}, a_i = \pm 1, \pm 3$ com $\beta = 2$, $\mathcal{R} = 8,2$ e $\gamma = 6,8$.

- **Cenário 4.2–A:** Nesse cenário, consideram-se $N_u = 2$ fontes e $L = 3$ sensores. O canal de comunicação MIMO é complexo e cada canal associando a i -ésima fonte ao ℓ -ésimo sensor é um filtro FIR com $K_c = 2$ coeficientes. Os coeficientes de cada canal são representados por polinômios $H_{i,\ell}(z)$, apresentados na Tabela 10. A SNR considerada nas saídas do canal é de 50 dB.

Tabela 10 – Polinômios $H_{i,\ell}(z)$ do canal MIMO no Cenário 4.2–A.

$i \backslash \ell$	1	2	3
1	$-0,6 + j0,4$ $+(1,2 - j0,2)z^{-1}$	$0,1 + j0,7$ $-(0,2 + j0,5)z^{-1}$	$0,5 + j0,4$ $+(-1,0 + j0,3)z^{-1}$
2	$-0,6 + j0,8$ $+(0,9 - j0,1)z^{-1}$	$0,4 - j0,3$ $+(-0,2 + j0,2)z^{-1}$	$-0,1 + j0,8$ $+(0,4 + j0,1)z^{-1}$

Fonte: autoria própria.

- **Cenário 4.2–B:** Consideram-se $N_u = 2$ fontes e $L = 2$ sensores. O canal de comunicação MIMO é complexo, com $K_c = 2$ e coeficientes $H_{i,\ell}(z)$ fornecidos na Tabela 11. Nesse cenário, a SNR nas saídas do canal vale 15 dB.

Em cada cenário, deseja-se validar o modelo variando-se o fator de esquecimento λ do algoritmo. Os parâmetros adotados para o MU-DM-SWA nas simulações são fornecidos na Tabela 12.

Tabela 11 – Polinômios $H_{i,\ell}(z)$ do canal MIMO no Cenário 4.2-B.

$i \backslash \ell$	1	2
1	$-0,6 + j0,4 + (1,2 - j0,2)z^{-1}$	$0,1 + j0,7 - (0,2 + j0,5)z^{-1}$
2	$-0,6 + j0,8 + (0,9 - j0,1)z^{-1}$	$0,4 - j0,3 + (-0,2 + j0,2)z^{-1}$

Fonte: autoria própria.

Tabela 12 – Parâmetros dos MU-DM-SWA nos Cenários 4.2-A e B.

Parâmetro	Cenário 4.2-A	Cenário 4.2-B
Diversidade temporal do equalizador (K_t)	3	15
Número de coeficientes de cada filtro (M)	9	30
Intervalo do fator de esquecimento (λ)	[0,997, 0,99995]	[0,997, 0,99995]
Ponderação do termo de correlação cruzada (ρ)	1	1
Atraso de espalhamento (k_1)	3	10

Fonte: autoria própria.

Cabe notar que o MU-DM-SWA é inicializado de acordo com a Tabela 1. No Cenário 4.2-A, adotam-se vetores $\mathbf{w}_1(-1)$ e $\mathbf{w}_2(-1)$ com pino único de valor 1 nas respectivas posições empiricamente estipuladas 3 e 4. No Cenário 4.2-B, os vetores $\mathbf{w}_1(-1)$ e $\mathbf{w}_2(-1)$ têm pino único de valor 1 nas respectivas posições 7 e 25, adotadas empiricamente. Além disso, nos Cenários 4.2-A e B, para cada λ considerado, adotam-se inicializações, respectivamente, de $\widehat{\mathbf{R}}^{-1}(-1) = (1 - \lambda)\mathbf{I}_9/10$ e $\widehat{\mathbf{R}}^{-1}(-1) = (1 - \lambda)\mathbf{I}_{30}/10$, segundo a Equação (3.53).

Observa-se que, no Cenário 4.2-A, o modelo para ζ_i é validado sob condições favoráveis à obtenção de equalização perfeita: considera-se uma constelação transmitida 16-QAM e um canal de comunicação complexo com $N_u < L$ e SNR alta em suas saídas. Por outro lado, no Cenário 4.2-B, busca-se validar o modelo sob condições mais adversas: considera-se novamente uma constelação transmitida 16-QAM e um canal de comunicação complexo similar ao do cenário anterior, porém em configuração igualmente determinada, *i.e.*, $N_u = L$ e com SNR relativamente baixa nas saídas do canal. Nota-se que, para se atingir a equalização perfeita no caso igualmente determinado, cada filtro do equalizador precisa ter infinitos coeficientes segundo a Condição 2.3-A (Equação (2.30)), apresentada na Seção 3.1. Somando-se isso ao fato de a SNR ser baixa nas saídas do canal, espera-se que a taxa de erro de símbolo seja aceitável no Cenário 4.2-B somente com uma ordem do equalizador significativamente alta, daí a necessidade de se utilizar $M = 30$ conforme mostrado na Tabela 12. Além disso, espera-se que o modelo completo (4.47), que prevê a ocorrência de equalização imperfeita, tenha maior relevância no Cenário 4.2-B.

Como passo intermediário a fim de se verificar a correção do desenvolvimento algébrico feito no Apêndice B, em ambos os cenários é inicialmente feita a validação do modelo dado pela Equação (B.56) para a esperança do termo $|\epsilon_i(n)|^2$. Essa validação inicial é feita em ambiente estacionário, *i.e.*, sem *tracking*. O modelo para $E[|\epsilon_i(n)|^2]$, deduzido nos Resultados 2.2–G e 2.2–H do Apêndice B, utiliza hipóteses menos restritivas que o modelo análogo previamente obtido por (LUO; CHAMBERS, 2002). Como resultado, a aproximação para a esperança de $|\epsilon_i(n)|^2$, segundo a abordagem de (LUO; CHAMBERS, 2002, ver a Eq. (42), *mutatis mutandis*), resulta mais simples e é dada por

$$E[|\epsilon_i(n)|^2] \approx \frac{\varrho^2}{4} \left(\frac{1-\lambda}{1+\lambda} \right) (N_u - 1)(k_1 + 1) E^3[|a_i(n)|^2]. \quad (4.92)$$

Posteriormente, o desempenho do modelo para o EMSE é observado em cada cenário considerando-se um ambiente estacionário e dois ambientes com tipos de *tracking* distintos. Em cada situação, a comparação é feita com os valores de médias de conjunto do erro quadrático em excesso do MU-DM-SWA.

Resultados: Cenário 4.2–A

A recuperação das fontes para os três tipos de ambientes considerados (sem e com *tracking*) se deu de acordo com a Tabela 13.

Tabela 13 – Recuperação das fontes no Cenário 4.2–A.

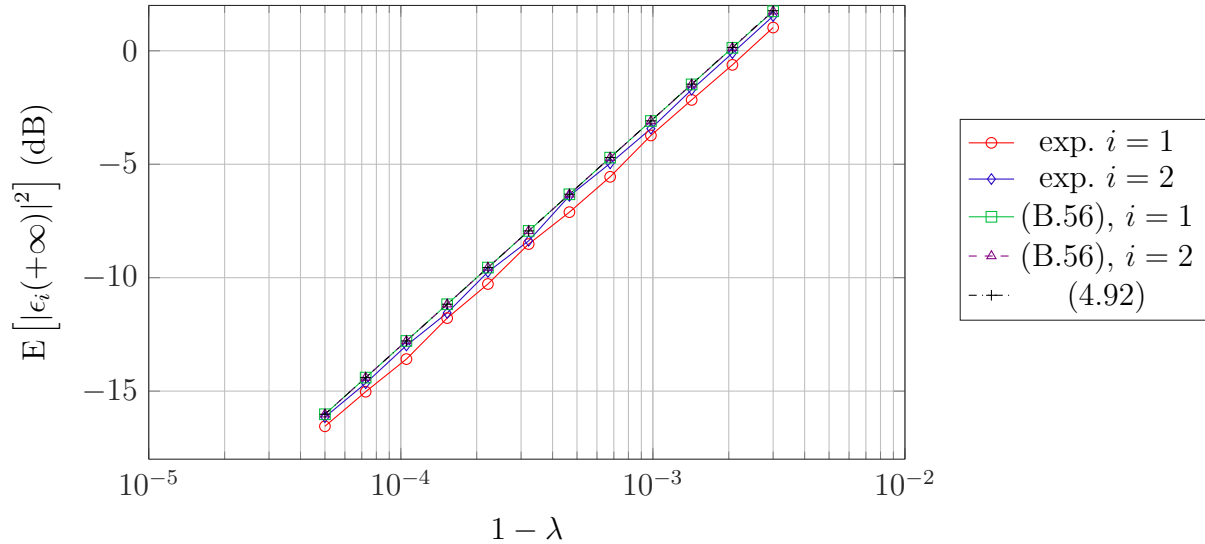
Equalizador (i)	1	2
Fonte recuperada	2	1
Rotação complexa (θ_i)	0°	90°
Atraso temporal (τ_i)	3	0
$E[v_i(n) ^2]$ segundo (4.55)	$1,2850 \times 10^{-3}$	$5,0629 \times 10^{-4}$

Fonte: autoria própria.

Na Figura 8, valores de médias de conjunto de 100 realizações independentes de $|\epsilon_i(n)|^2$, em regime, são comparados com o respectivo modelo deduzido nesse trabalho, dado pela Equação (B.56), e também com o modelo segundo a abordagem de (LUO; CHAMBERS, 2002), dado pela Equação (4.92). Nessa simulação intermediária, considerou-se um ambiente estacionário. Nota-se a concordância entre os modelos (B.56) e (4.92) com os valores experimentais.

Na Figura 9, valores de médias de conjunto de 100 realizações independentes de $|e_{a,i}(n)|^2$, em regime, são comparados com o modelo para o EMSE ζ_i , deduzido nesse trabalho e dado pela Equação (4.85). As simulações foram feitas em um ambiente estacionário e em dois ambientes com *tracking*. As curvas experimentais são comparadas com (4.85)

Figura 8 – Validação do modelo para a esperança de $|\epsilon_i(n)|^2$ no Cenário 4.2-A, para sinal 16-QAM em um ambiente estacionário, com média de 100 realizações independentes para a curva experimental de cada equalizador. As curvas estão em função de $1 - \lambda$, sendo λ o fator de esquecimento adotado. No gráfico são mostradas curvas do modelo da Equação (B.56) (deduzido nesse trabalho) e o modelo da Equação (4.92) (segundo a abordagem de (LUO; CHAMBERS, 2002)).

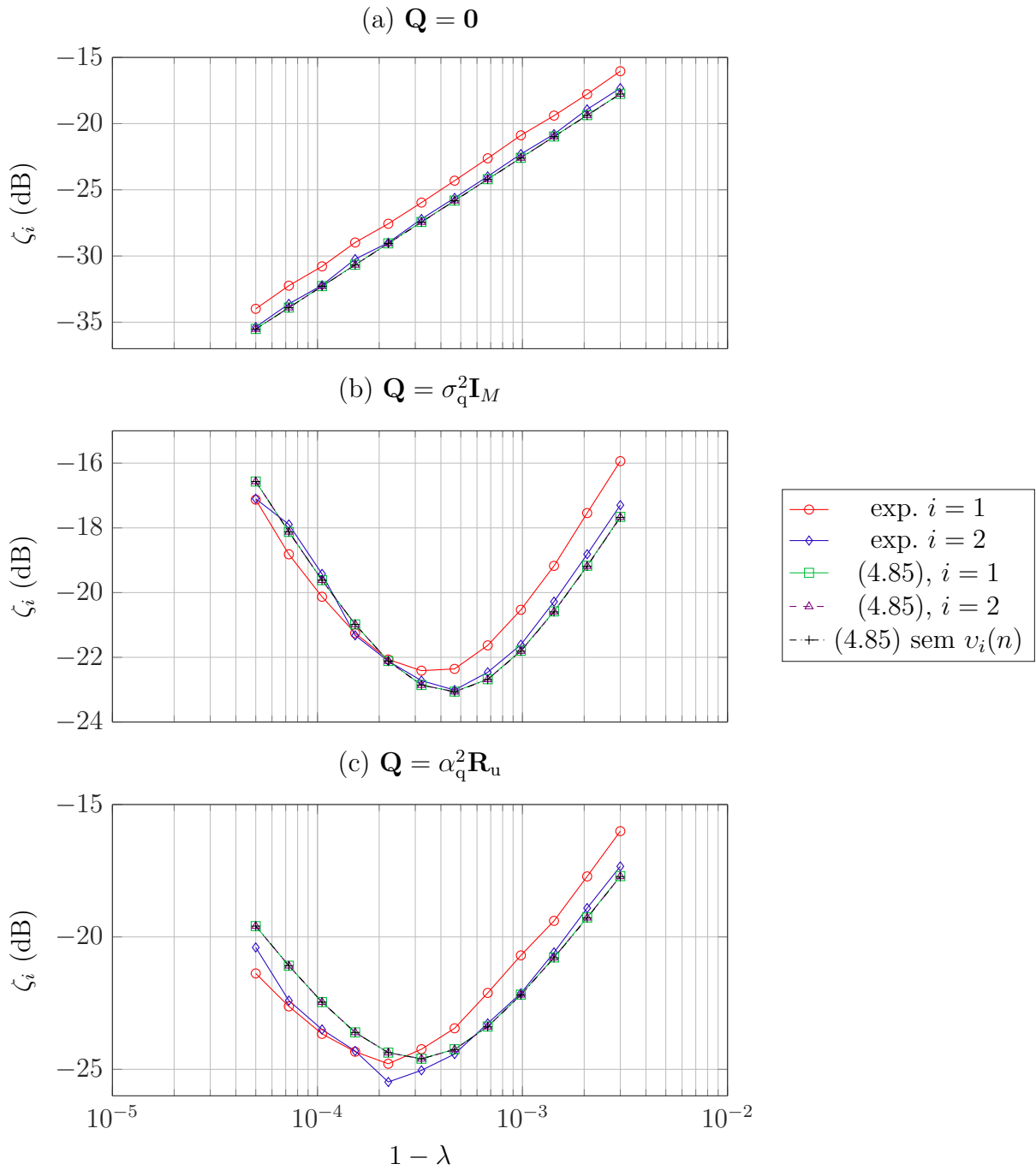


Fonte: autoria própria.

considerando-se os modelos com e sem $v_i(n)$, a fim de se verificar a relevância da inclusão desse termo no modelo.

Nota-se que ambos os modelos (4.85) com e sem $v_i(n)$ são adequados em face do EMSE experimental obtido. Isso ocorre pois, nesse cenário, o equalizador consegue se aproximar da equalização perfeita e $v_i(n) \approx 0$. Observa-se também, por meio de simulações, que essa concordância entre os modelos e as curvas experimentais é válida para $\lambda > 0,997$, visto que, para fatores de esquecimento menores no cenário considerado, as hipóteses assumidas na análise em regime deixam de ser válidas, além de os equalizadores não serem capazes de separar as fontes adequadamente (*i.e.*, com baixa taxa de erro de símbolos).

Figura 9 – Validação do modelo do EMSE em regime ζ_i (Equação (4.85)) no Cenário 4.2–A, para sinal 16-QAM, com média de 100 realizações independentes para a curva experimental de cada equalizador. As curvas estão em função de $1 - \lambda$, sendo λ o fator de esquecimento adotado. (a) Ambiente estacionário. (b) Ambiente com *tracking* e $\sigma_q = 1 \times 10^{-4}$. (c) Ambiente com *tracking* e $\alpha_q = 1 \times 10^{-5}$.



Fonte: autoria própria.

Resultados: Cenário 4.2-B

Nesse cenário, a recuperação das fontes para os três tipos de ambientes considerados (sem e com *tracking*) se deu conforme a Tabela 14.

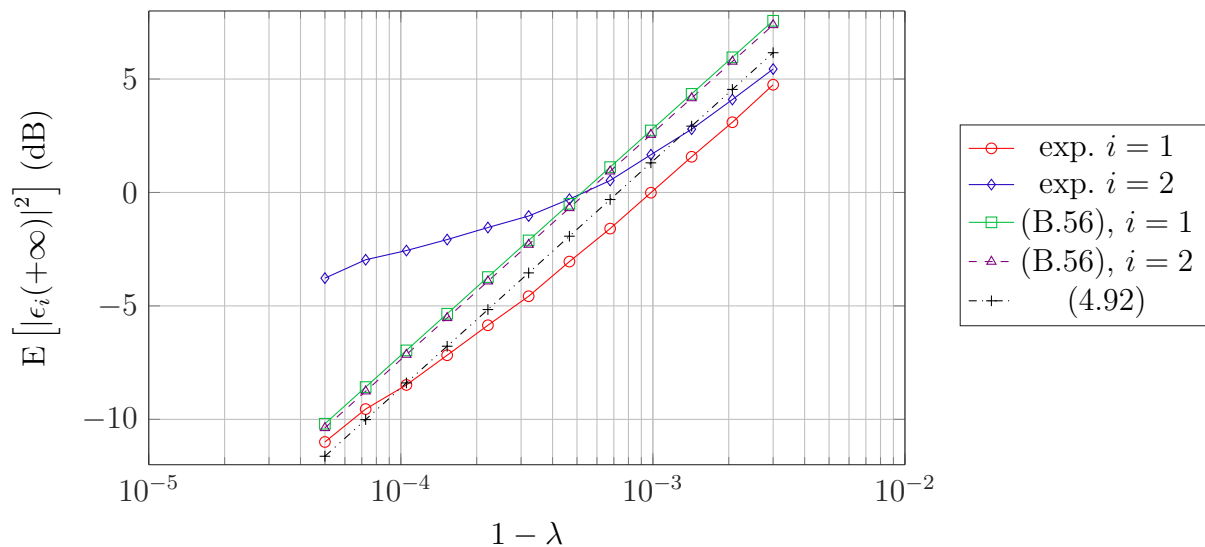
Tabela 14 – Recuperação das fontes no Cenário 4.2-B.

Equalizador (i)	1	2
Fonte recuperada	1	2
Rotação complexa (θ_i)	0°	-90°
Atraso temporal (τ_i)	7	9
$E[v_i(n) ^2]$ segundo (4.55)	0,8514	1,3223

Fonte: autoria própria.

Novamente, como passo intermediário a fim de se verificar a correção do desenvolvimento algébrico feito no Apêndice B, foi feita a validação do modelo dado pela Equação (B.56) para a esperança do termo $|\epsilon_i(n)|^2$ em ambiente estacionário. Esse modelo também foi comparado com o modelo segundo a abordagem de (LUO; CHAMBERS, 2002), dado pela Equação (4.92). Na Figura 10, valores de médias de conjunto de 100 realizações independentes de $|\epsilon_i(n)|^2$, em regime, são comparados com os dois modelos em questão.

Figura 10 – Validação do modelo para a esperança de $|\epsilon_i(n)|^2$ no Cenário 4.2-B, com sinal 16-QAM em um ambiente estacionário, com média de 100 realizações independentes para a curva experimental de cada equalizador. As curvas estão em função de $1 - \lambda$, sendo λ o fator de esquecimento adotado. No gráfico são mostradas curvas do modelo da Equação (B.56) (deduzido nesse trabalho) e o modelo da Equação (4.92) (segundo a abordagem de (LUO; CHAMBERS, 2002)).



Fonte: autoria própria.

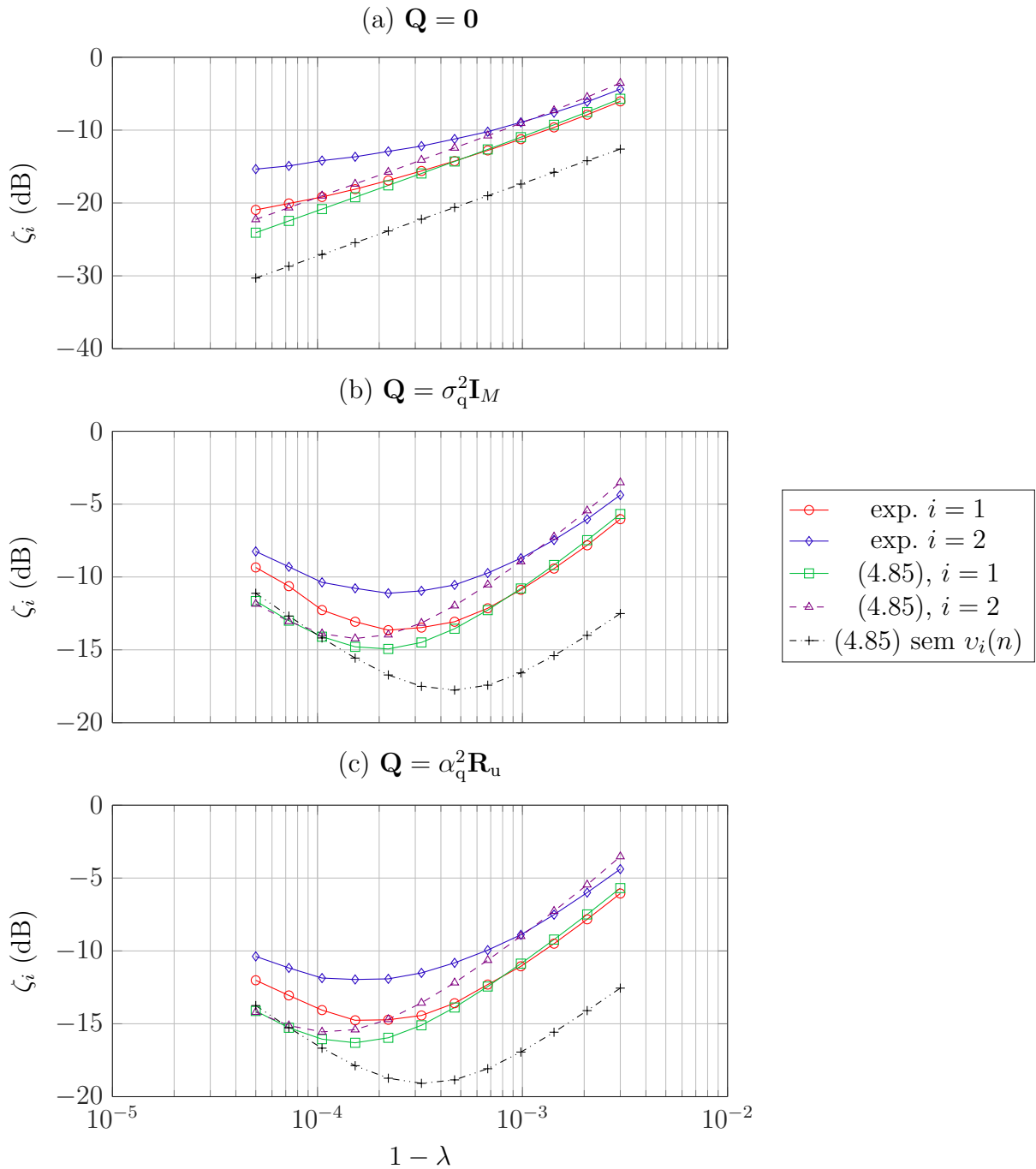
Nota-se a concordância entre os modelos (B.56) e (4.92) com os valores experimentais, exceto na curva azul (com $i = 2$) para $\lambda > 0,9995$, pois devido à baixa relação sinal-ruído nas saídas do canal e ao alto fator de esquecimento, esse equalizador não é capaz de acompanhar as rápidas variações da entrada.

Na Figura 11, valores de médias de conjunto de 100 realizações independentes de $|e_{a,i}(n)|^2$, em regime, são comparados com o modelo para o EMSE ζ_i , deduzido nesse trabalho e dado pela Equação (4.85). As simulações foram feitas em um ambiente estacionário e em dois ambientes com *tracking*. As curvas experimentais são comparadas com o modelo de (4.85) considerando $v_i(n)$ tanto com a variância mostrada na Tabela 14 quanto com variância nula, a fim de verificar a sua relevância no modelo.

Para cada equalizador, o modelo (4.85) com $v_i(n)$ não nulo apresenta uma maior concordância com o EMSE experimental em relação a esse mesmo modelo considerando-se $v_i(n) = 0$. Isso ocorre pois, nesse cenário menos favorável à equalização perfeita, o valor de $v_i(n)$ não é desprezível. Além disso, cabe notar que, para $\lambda > 0,9993$, passa a haver uma gradual discordância entre o modelo (4.85) com $v_i(n)$ não nulo e os valores de EMSE experimental, para ambos os equalizadores. Como a relação sinal-ruído nas saídas do canal é baixa nesse cenário, o equalizador não consegue acompanhar as rápidas variações da entrada com fator de esquecimento alto, o que não é previsto pelo modelo. Adicionalmente, nesse cenário pouco favorável à equalização perfeita, as hipóteses consideradas na análise em regime podem deixar de ser válidas, o que influencia na qualidade do modelo.

Por fim, nota-se por meio de simulações que a concordância observada entre os modelos e as curvas experimentais é válida para $\lambda > 0,997$, pois para fatores de esquecimento menores no cenário considerado, além de as hipóteses assumidas na análise em regime deixarem de ser válidas, os equalizadores não são capazes de separar as fontes adequadamente.

Figura 11 – Validação do modelo do EMSE em regime ζ_i (Equação (4.85)) no Cenário 4.2-B, para sinal 16-QAM, com média de 100 realizações independentes para a curva experimental de cada equalizador. As curvas estão em função de $1 - \lambda$, sendo λ o fator de esquecimento adotado. (a) Ambiente estacionário. (b) Ambiente com *tracking* e $\sigma_q = 1 \times 10^{-4}$. (c) Ambiente com *tracking* e $\alpha_q = 1 \times 10^{-5}$.



Fonte: autoria própria.

4.3 CONCLUSÃO

Nesse capítulo, foram feitas análises determinista e em regime do algoritmo de Shalvi-Weinstein com operação em *dual-mode* (MU-DM-SWA). Além de as soluções para desconvolução cega multiusuário serem mais sofisticadas do que aquelas existentes para o caso SISO, nota-se que a análise dessas soluções quanto à estabilidade e ao desempenho em regime também é mais sofisticada e laboriosa.

Com a análise determinista em precisão infinita feita na Seção 4.1, foi possível provar que a norma do vetor de coeficientes do MU-DM-SWA sempre se mantém limitada, desde que satisfeita a condição de excitação persistente sobre o regressor. A estabilidade foi provada para funcionamento dentro e fora da RoI, bem como quando ocorre transição entre modos de operação. Isso corrobora os resultados encontrados para simulações numéricas do MU-DM-SWA na Seção 3.5 do Capítulo 3. É importante notar que o funcionamento em *dual-mode* do MU-DM-SWA facilita os cálculos da análise determinista, visto que as não linearidades do algoritmo são remetidas à entrada.

Com a análise em regime da Seção 4.2, obteve-se um modelo para o erro quadrático médio em excesso (EMSE) do MU-DM-SWA operando dentro da RoI, dado pela Equação (4.85). Esse modelo permite realizar a escolha de parâmetros que garantem o funcionamento médio adequado do MU-DM-SWA. Com o auxílio de simulações numéricas, verificou-se a validade do modelo obtido para dois cenários distintos, desde que o fator de esquecimento do algoritmo seja próximo de um e as hipóteses estocásticas consideradas sejam razoáveis. Embora possua essas limitações inerentes a modelos de algoritmos cegos, nota-se que em ambientes com *tracking*, o modelo permite encontrar o fator de esquecimento que resulta no EMSE em regime mínimo. Em um ambiente favorável à equalização perfeita, o modelo obtido concorda com resultados experimentais, conforme observado na Figura 9. Já em ambientes mais adversos, comuns em aplicações práticas, a equalização imperfeita deve ser levada em conta no modelo. Isso é feito considerando-se um ruído para modelar a equalização imperfeita, segundo as Equações (4.47) e (4.53), o que traz uma dificuldade adicional na obtenção do modelo. No entanto, fica evidenciada a vantagem do modelo “completo” sobre o modelo mais simples quando o algoritmo opera em cenário adverso, como é possível concluir a partir da Figura 11.

Por fim, cabe ressaltar que a análise determinista é suficiente para garantir que o algoritmo não diverge, mas não leva em conta a qualidade da convergência, e não é capaz, por exemplo, de prever uma possível convergência indesejável do algoritmo para mínimos locais. Em contrapartida, a análise estocástica permite avaliar a qualidade do desempenho médio do algoritmo, sem no entanto garantir que o algoritmo não diverge, pois trata-se de uma análise em média (RUPP, 2015).

Após o estudo do problema de desconvolução multiusuário e a obtenção de soluções

que evitam a divergência, no próximo capítulo retorna-se ao problema de separação cega de fontes, com uma apresentação de critérios de separação baseados em curtose.

5 CRITÉRIOS DE SEPARAÇÃO DE FONTES BASEADOS EM CURTOSE

Neste capítulo, apresentam-se critérios para a desconvolução cega multiusuário, baseados em curtose, que levam em conta os casos MIMO e multimódulo. Além disso, também são apresentados critérios de BSS baseados em cumulantes. O objetivo é compreender melhor sob que condições é possível estender os algoritmos eficientes de desconvolução cega, propostos no Capítulo 3, para algoritmos mais abrangentes de BSS. Isso pode ser interessante, já que deseja-se que esses algoritmos tenham as boas características das soluções propostas no Capítulo 3.

Inicialmente, na Seção 5.1, são definidos os conceitos estatísticos de cumulante e curtose de uma variável aleatória. Na Seção 5.2, são apresentados critérios usuais de desconvolução cega baseados em maximização de curtose. Esses critérios são relacionados ao critério da função custo determinista apresentada no Capítulo 3. Posteriormente, na Seção 5.3, são apresentados critérios de separação cega de fontes baseados em maximização de curtose, para o caso de misturas instantâneas. Já para misturas convolutivas, as abordagens usuais para a resolução do problema são apresentadas na Seção 5.4. Por fim, a conclusão deste capítulo é feita na Seção 5.5.

5.1 CUMULANTES E CURTOSE

Seja um conjunto de variáveis aleatórias contínuas x_1, x_2, \dots, x_m completamente descrito por uma função densidade de probabilidade conjunta

$$f_{\mathbf{x}}(\mathbf{x}) = f_{x_1, x_2, \dots, x_m}(x_1, x_2, \dots, x_m), \quad (5.1)$$

em que

$$\mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_m]^T \quad (5.2)$$

é o vetor aleatório constituído pelas variáveis x_i , com $i = 1, 2, \dots, m$. Os momentos conjuntos de ordem n de \mathbf{x} , em torno da origem, são definidos como (PICINBONO, 1993; PAPOULIS; PILLAI, 2002)

$$\text{mom}(x_1^{q_1}, x_2^{q_2}, \dots, x_m^{q_m}) = \text{E} \left[\prod_{k=1}^m x_k^{q_k} \right], \quad (5.3)$$

sendo q_k expoentes inteiros positivos, com $k = 1, 2, \dots, m$, que satisfazem a igualdade

$$\sum_{k=1}^m q_k = n. \quad (5.4)$$

Utilizando-se a definição de esperança matemática, (5.3) pode ser reescrita como (PICINBONO, 1993; PAPOULIS; PILLAI, 2002)

$$\text{mom}(x_1^{q_1}, x_2^{q_2}, \dots, x_m^{q_m}) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_{\mathbf{x}}(\mathbf{x}) \prod_{k=1}^m x_k^{q_k} dx_1 \dots dx_m. \quad (5.5)$$

Cabe notar que o vetor aleatório \mathbf{x} , de m variáveis aleatórias, possui no máximo¹

$$\frac{(n-1)!}{(n-m)!(m-1)!} \quad (5.7)$$

momentos conjuntos de ordem n distintos.

Por conveniência, define-se a primeira função característica de \mathbf{x} como (PICINBONO, 1993; PAPOULIS; PILLAI, 2002)

$$\Phi_{\mathbf{x}}(\boldsymbol{\omega}) = \text{E} [e^{j\boldsymbol{\omega}^T \mathbf{x}}] = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_{\mathbf{x}}(\mathbf{x}) e^{j\boldsymbol{\omega}^T \mathbf{x}} dx_1 \dots dx_m, \quad (5.8)$$

sendo

$$\boldsymbol{\omega} = [\omega_1 \quad \omega_2 \quad \cdots \quad \omega_m]^T \quad (5.9)$$

o vetor de frequências angulares. A primeira função característica $\Phi_{\mathbf{x}}(\boldsymbol{\omega})$ pode ser interpretada como a transformada de Fourier da função densidade de probabilidade conjunta de \mathbf{x} , com os sinais das frequências trocados. Se \mathbf{x} for um vetor de variáveis aleatórias discretas, as integrais de (5.8) podem ser substituídas por somatórios. Nesse caso, $\Phi_{\mathbf{x}}(\boldsymbol{\omega})$ pode ser interpretada como a transformada de Fourier de tempo discreto² de $f_{\mathbf{x}}(\mathbf{x})$, com os sinais das frequências trocados. Tanto para conjuntos de variáveis aleatórias contínuas ou discretas, a primeira função característica sempre existe, mesmo que funções densidade de probabilidade não possam ser definidas (PEEBLES, 2000; PAPOULIS; PILLAI, 2002). Além disso, em ambos os casos, a primeira função característica é definida para variáveis contínuas $\omega_1, \omega_2, \dots, \omega_m \in (-\infty, +\infty)$.

Pode-se utilizar a primeira função característica de \mathbf{x} para a obtenção dos momentos conjuntos desse vetor aleatório, caso eles existam, segundo a relação (PICINBONO, 1993; PAPOULIS; PILLAI, 2002)

$$\text{mom}(x_1^{q_1}, x_2^{q_2}, \dots, x_m^{q_m}) = (-j)^n \left. \frac{\partial^n \Phi_{\mathbf{x}}(\boldsymbol{\omega})}{\partial \omega_1^{q_1} \partial \omega_2^{q_2} \dots \partial \omega_m^{q_m}} \right|_{\boldsymbol{\omega}=0}. \quad (5.10)$$

Também é conveniente definir a segunda função característica de \mathbf{x} como o logaritmo natural de $\Phi_{\mathbf{x}}(\boldsymbol{\omega})$, *i.e.*,

$$\Psi_{\mathbf{x}}(\boldsymbol{\omega}) = \ln \Phi_{\mathbf{x}}(\boldsymbol{\omega}). \quad (5.11)$$

¹ O fatorial de um número natural ℓ é definido como

$$\ell! \triangleq \begin{cases} \ell(\ell-1)!, & \text{se } \ell > 0, \\ 1, & \text{se } \ell = 0. \end{cases} \quad (5.6)$$

² Embora a transformada de Fourier de tempo discreto seja usualmente definida para sinais no domínio do tempo, nesse caso o domínio de integração é o conjunto enumerável de valores assumidos pela variável aleatória discreta.

A partir de (5.11), os cumulantes conjuntos de ordem n de \mathbf{x} são definidos como (NIKIAS; PETROPULU, 1993; PICINBONO, 1993)

$$\text{cum}(x_1^{q_1}, x_2^{q_2}, \dots, x_m^{q_m}) = (-j)^n \frac{\partial^n \Psi_{\mathbf{x}}(\boldsymbol{\omega})}{\partial \omega_1^{q_1} \partial \omega_2^{q_2} \dots \partial \omega_m^{q_m}} \Big|_{\boldsymbol{\omega}=0}. \quad (5.12)$$

Os cumulantes possuem propriedades matemáticas interessantes: são operadores lineares e gozam da propriedade de aditividade para partições estatisticamente independentes de um conjunto de variáveis aleatórias. Além disso, o cumulante conjunto resulta nulo se aplicado em um conjunto de variáveis aleatórias que tenha duas ou mais partições, todas estatisticamente independentes entre si. Por fim, cabe notar que, para variáveis aleatórias conjuntamente gaussianas, os cumulantes de ordem superior a dois são sempre nulos (NIKIAS; PETROPULU, 1993).

É usual expressar os cumulantes em função de momentos conjuntos das variáveis envolvidas. É possível demonstrar que são válidas as seguintes relações para os cumulantes de ordem dois, três e quatro de variáveis aleatórias com distribuições distintas, respectivamente (KENDALL; STUART, 1977; MCCULLAGH, 1987):

$$\text{cum}(x_1, x_2) = E[x_1 x_2] - E[x_1] E[x_2] = \text{COV}(x_1, x_2), \quad (5.13)$$

$$\begin{aligned} \text{cum}(x_1, x_2, x_3) &= E[x_1 x_2 x_3] - E[x_1] E[x_2 x_3] - E[x_2] E[x_1 x_3] - E[x_3] E[x_1 x_2] \\ &\quad + 2 E[x_1] E[x_2] E[x_3], \end{aligned} \quad (5.14)$$

$$\begin{aligned} \text{cum}(x_1, x_2, x_3, x_4) &= E[x_1 x_2 x_3 x_4] - E[x_1] E[x_2 x_3 x_4] - E[x_2] E[x_1 x_3 x_4] \\ &\quad - E[x_3] E[x_1 x_2 x_4] - E[x_4] E[x_1 x_2 x_3] - E[x_1 x_2] E[x_3 x_4] \\ &\quad - E[x_1 x_3] E[x_2 x_4] - E[x_1 x_4] E[x_2 x_3] + 2 E[x_1] E[x_2] E[x_3 x_4] \\ &\quad + 2 E[x_1] E[x_3] E[x_2 x_4] + 2 E[x_1] E[x_4] E[x_2 x_3] + 2 E[x_2] E[x_3] E[x_1 x_4] \\ &\quad + 2 E[x_2] E[x_4] E[x_1 x_3] + 2 E[x_3] E[x_4] E[x_1 x_2] \\ &\quad - 6 E[x_1] E[x_2] E[x_3] E[x_4], \end{aligned} \quad (5.15)$$

em que $\text{COV}(\cdot, \cdot)$ denota a covariância entre as variáveis consideradas. Para o caso em que todas as variáveis aleatórias possuem média nula, os cumulantes (5.13), (5.14) e (5.15) podem ser simplificados, resultando respectivamente em

$$\text{cum}(x_1, x_2) = E[x_1 x_2], \quad (5.16)$$

$$\text{cum}(x_1, x_2, x_3) = E[x_1 x_2 x_3], \quad (5.17)$$

$$\text{cum}(x_1, x_2, x_3, x_4) = E[x_1 x_2 x_3 x_4] - E[x_1 x_2] E[x_3 x_4] - E[x_1 x_3] E[x_2 x_4] - E[x_1 x_4] E[x_2 x_3]. \quad (5.18)$$

Considerando-se o número de variáveis aleatórias m igual à ordem n do cumulante conjunto, nota-se, a partir da Equação (5.4), que $q_k = 1$ para $k = 1, 2, \dots, m$. Nessa situação e para variáveis aleatórias identicamente distribuídas, costuma-se utilizar as

seguintes notações:

$$\text{cum}(\underbrace{x, x, x, \dots, x}_{p \text{ vezes}}) = \text{cum}(x : p), \quad (5.19)$$

$$C_p^x = \text{cum}(x : p), \quad (5.20)$$

$$C_{p,q}^x = \text{cum}(x : p, x^* : q). \quad (5.21)$$

Particularmente para o caso em que x é uma variável aleatória real, vale $C_{p,q}^x = C_{p+q}^x$.

Para uma variável aleatória x de média nula, os cumulantes de segunda e quarta ordem podem ser representados como:

$$C_{1,1}^x = \text{E} [|x|^2] = \text{VAR}[x], \quad (5.22)$$

$$C_{2,2}^x = \text{E} [|x|^4] - 2 \text{E}^2 [|x|^2] - \text{E} [x^2] \text{E} [(x^*)^2], \quad (5.23)$$

em que $\text{VAR}[\cdot]$ denota a variância. O cumulante de ordem quatro normalizado, dado por

$$\kappa_x = \frac{C_{2,2}^x}{(C_{1,1}^x)^2}, \quad (5.24)$$

é denominado curtose (PICINBONO, 1993). A curtose de uma variável aleatória está associada, em termos gerais, ao decaimento das caudas da sua função densidade de probabilidade. Para variáveis aleatórias gaussianas, como decorrência da propriedade dos cumulantes, a curtose é nula. Por esse motivo, seu uso recebe destaque em critérios de separação cega de fontes. Em geral, deseja-se recuperar a não gaussianidade das fontes transmitidas para poder separá-las, visto que, pelo Teorema do Limite Central, as misturas devem ser mais gaussianas do que as fontes (HYVÄRINEN; OJA, 2000). Usualmente, as variáveis aleatórias com curtose positiva são denominadas supergaussianas, enquanto aquelas com curtose negativa são ditas subgaussianas.

Para o caso particular em que x possui média nula e é real, utilizando-se (5.22) e (5.23), a curtose pode ser reescrita como

$$\kappa_x = \frac{\text{E} [x^4]}{\text{E}^2 [x^2]} - 3. \quad (5.25)$$

Já se x possuir média nula, mas for complexa e satisfizer a condição de simetria circular usual em comunicações digitais (HAYKIN, 2000b), *i.e.*, $\text{E}[x^2] = 0$, então a curtose pode ser expressa como

$$\kappa_x = \frac{\text{E} [|x|^4]}{\text{E}^2 [|x|^2]} - 2. \quad (5.26)$$

Os cumulantes de ordem superior, e em especial a curtose, são ferramentas importantes para a resolução de problemas de separação cega de fontes. Grande parte

dos algoritmos existentes para resolver os problemas de BSS buscam recuperar a independência estatística das fontes a partir de critérios baseados em estatísticas de ordem superior (SHALVI; WEINSTEIN, 1990; CARDOSO, 1998). A seguir, são apresentados os principais critérios baseados em estatísticas de ordem superior para o problema de desconvolução cega multiusuário, em que as estatísticas das fontes são bem conhecidas.

5.2 DESCONVOLUÇÃO CEGA BASEADA EM CUMULANTES

Nesta seção, considera-se o problema de desconvolução cega de acordo com a formulação feita na Seção 2.3 do Capítulo 2. Primeiramente, na Seção 5.2.1, são apresentados os principais critérios de desconvolução cega SISO, considerando-se $N_u = 1$ fonte e $L = 1$ sensor. Em seguida, na Seção 5.2.2, são apresentados os critérios SISO com abordagem multimódulo. Posteriormente, na Seção 5.2.3, é apresentada a extensão desses critérios para o caso MIMO.

5.2.1 Caso SISO

O problema de desconvolução cega para o caso SISO vem sendo estudado desde os anos 1960 e 1970 (LUCKY, 1966; SATO, 1975). Inúmeras estratégias para resolver esse problema resultaram em critérios diferentes, utilizados posteriormente na proposição de algoritmos em bloco e adaptativos. Em 1990, Shalvi e Weinstein propuseram um critério para desconvolução SISO baseado em estatísticas de ordem superior (SHALVI; WEINSTEIN, 1990; SHALVI; WEINSTEIN, 1994). Esse critério foi obtido a partir da comparação entre os cumulantes das fontes e dos sinais recuperados, impondo-se a condição de equalização perfeita (SHALVI; WEINSTEIN, 1990), abordada na Seção 2.3 do Capítulo 2 e dada pela Equação (2.29). O critério resultante busca maximizar o cumulante da saída, mantendo sua variância igual àquela dos sinais transmitidos. Considerando-se inteiros não nulos p e q tais que $p + q > 2$, esse critério pode ser colocado na forma (SHALVI; WEINSTEIN, 1994)

$$\begin{aligned} & \text{máx } \left| C_{p,q}^{y(n)} \right| & (5.27) \\ & \text{sujeito a } C_{1,1}^{y(n)} = C_{1,1}^{a(n)}, \end{aligned}$$

em que $a(n)$ representa o sinal transmitido e $y(n)$ corresponde à saída do equalizador, dada pela Equação (2.25). A restrição de (5.27) pode ser englobada em uma única função custo normalizada J_{SW} a ser maximizada, resultando em

$$J_{\text{SW}} = \frac{\left| C_{p,q}^{y(n)} \right|}{\left(C_{1,1}^{y(n)} \right)^{\frac{p+q}{2}}}. \quad (5.28)$$

Usualmente, assume-se $p = q = 2$ e (5.28) pode ser escrita como (SHALVI; WEINSTEIN, 1990)

$$J_{\text{SW}} = \frac{|C_{2,2}^{y(n)}|}{(C_{1,1}^{y(n)})^2}. \quad (5.29)$$

Outro critério para o caso SISO foi proposto por Godard e é conhecido como critério de módulo constante (GODARD, 1980). Ele é baseado na minimização de uma função custo estocástica de dispersão de ordem p , sendo p um inteiro positivo (GODARD, 1980). Esse critério pode ser representado pela minimização da função custo J_G dada por

$$J_G = \text{E} \left[\left(|y(n)|^p - \mathfrak{R}_p^a \right)^2 \right], \quad (5.30)$$

em que

$$\mathfrak{R}_p^a = \frac{\text{E} [|a(n)|^{2p}]}{\text{E} [|a(n)|^p]} \quad (5.31)$$

é denominada constante de dispersão. No caso usual, utiliza-se $p = 2$, de modo que a função custo de Godard resulta

$$J_G = \text{E} \left[\left(|y(n)|^2 - \mathfrak{R}_2^a \right)^2 \right]. \quad (5.32)$$

Cabe notar que o critério (5.32) é baseado em momentos de ordem superior. Além disso, é possível demonstrar que, sob certas condições, as funções custo de Godard e Shalvi-Weinstein possuem os mesmos pontos de estacionários, sendo equivalentes nesse sentido (REGALIA, 1999), conforme abordado na Seção 3.1 do Capítulo 3. Portanto, o algoritmo de Shalvi-Weinstein pode ser interpretado como um algoritmo de módulo constante que inclui, na sua equação de adaptação, uma aproximação para a matriz Hessiana, sendo considerado como um algoritmo do tipo quase-Newton (SILVA; MIRANDA, 2004).

No trabalho de (MIRANDA; SILVA; NASCIMENTO, 2008), a função custo de Godard (5.32) foi apresentada em sua forma determinista para a dedução do algoritmo de Shalvi-Weinstein no caso SISO. No lugar da esperança matemática, tomam-se médias temporais com janela exponencial para deduzir o SWA a partir do método de mínimos quadrados. Assim, expressa-se (5.32) na forma

$$\hat{J}_G(n) = \sum_{\ell=0}^n \lambda^{n-\ell} \left(|y(\ell, n)|^2 - \mathfrak{R}_2^a \right)^2, \quad (5.33)$$

em que $0 \ll \lambda < 1$ é o fator de esquecimento e $y(\ell, n) \triangleq \mathbf{u}^T(\ell) \mathbf{w}(n)$.

5.2.2 Caso SISO multimódulo

A minimização da função custo (5.32) permite desfazer a interferência intersimbólica introduzida pelo canal de comunicação (GODARD, 1980). No entanto, notou-se que a

recuperação das fontes se dá com um atraso temporal e uma eventual rotação complexa da constelação, de acordo com a condição de equalização perfeita. A fim de se minimizar os erros de detecção devido a essa rotação complexa, foram propostos na literatura estágios de rotação posteriores ao equalizador. Por outro lado, (YANG; WERNER; DUMONT, 2002) apresentou um critério de desconvolução que evita automaticamente as rotações complexas não múltiplas de 90° na saída do equalizador. Esse critério, conhecido como multimódulo, consiste na minimização da função custo (5.32), decomposta nas partes real e imaginária da saída do equalizador, ou seja,

$$J_{\text{GMM}} = \text{E} \left[\left(y_{\text{R}}^2(n) - \mathcal{R}_2^{a_{\text{R}}} \right)^2 \right] + \text{E} \left[\left(y_{\text{I}}^2(n) - \mathcal{R}_2^{a_{\text{I}}} \right)^2 \right], \quad (5.34)$$

em que $y(n) = y_{\text{R}}(n) + jy_{\text{I}}(n)$ corresponde à saída do equalizador, e as constantes de dispersão são dadas, nesse caso, por

$$\mathcal{R}_2^{a_{\text{R}}} = \frac{\text{E} [a_{\text{R}}^4(n)]}{\text{E} [a_{\text{R}}^2(n)]} \quad \text{e} \quad \mathcal{R}_2^{a_{\text{I}}} = \frac{\text{E} [a_{\text{I}}^4(n)]}{\text{E} [a_{\text{I}}^2(n)]}. \quad (5.35)$$

Diante da equivalência dos pontos estacionários das funções custo (5.29) e (5.32), e do critério que evita rotações complexas dado por (5.34), sugere-se a seguinte função custo a ser maximizada, inspirada no critério de Shalvi-Weinstein, e que também evita rotações complexas:

$$J_{\text{SWMM}} = \frac{|C_{2,2}^{y_{\text{R}}(n)}|}{\left(C_{1,1}^{y_{\text{R}}(n)}\right)^2} + \frac{|C_{2,2}^{y_{\text{I}}(n)}|}{\left(C_{1,1}^{y_{\text{I}}(n)}\right)^2}. \quad (5.36)$$

5.2.3 Caso MIMO

No caso de desconvolução cega multiusuário, a recuperação das fontes segundo o critério de equalização perfeita, dado pela Equação (2.29), requer a remoção da interferência intersimbólica introduzida pelo canal dispersivo, bem como da interferência cocanal presente no problema MIMO.

Nesse cenário, Papadias e Paulraj estabeleceram condições necessárias e suficientes para a equalização perfeita, adicionando uma condição às já existentes para o caso SISO previamente estabelecidas por Shalvi e Weinstein (PAPADIAS; PAULRAJ, 1997b; SHALVI; WEINSTEIN, 1990). A condição adicional requer que as saídas do equalizador MIMO sejam não correlacionadas entre si, para garantir que cada um dos filtros que compõem o equalizador convirja para fontes distintas. No contexto estocástico, ela pode ser colocada na forma

$$r_{i,m}(k) = \text{E} [y_i(n) y_m^*(n-k)] = 0, \quad \text{para } i \neq m \text{ e } k = 0, 1, \dots, k_1, \quad (5.37)$$

em que $y_i(n)$ é a i -ésima saída do equalizador MIMO com $i = 1, 2, \dots, N_{\text{u}}$, sendo N_{u} o número de fontes transmitidas e k_1 o atraso de espalhamento necessário para que a

convergência para a mesma fonte possa ser identificada. A escolha adequada de k_1 é função dos atrasos temporais com que cada filtro do equalizador recupera cada fonte.

Para se considerar a nova condição de descorrelação para o caso MIMO, pode-se adicionar um termo de penalização da correlação cruzada às funções custo já obtidas para o caso SISO, como proposto em (PAPADIAS; PAULRAJ, 1997a; PAPADIAS; PAULRAJ, 1997b). O algoritmo obtido em um contexto determinista, no Capítulo 3, sugere que a função custo de Shalvi-Weinstein estocástica, a ser maximizada para o i -ésimo equalizador e dada pela Equação (5.29), pode ser reescrita no caso MIMO na forma de uma função custo a ser minimizada $J_{\text{SW},i}$, dada por

$$J_{\text{SW},i} = -\frac{|C_{2,2}^{y_i(n)}|}{(C_{1,1}^{y_i(n)})^2} + \varrho \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} |r_{i,m}(k)|^2, \quad (5.38)$$

em que ϱ é a ponderação adotada para a parcela de descorrelação.

Equivalentemente para a função custo de Godard, a extensão para o caso MIMO de (5.32) pode ser colocada como uma função custo a ser minimizada $J_{\text{G},i}$ dada por

$$J_{\text{G},i} = \text{E} \left[\left(|y_i(n)|^2 - \mathfrak{R}_2^a \right)^2 \right] + \varrho \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} |r_{i,m}(k)|^2. \quad (5.39)$$

Diante do exposto, e inspirando-se no algoritmo de Shalvi-Weinstein apresentado no Capítulo 3, considera-se aqui a seguinte função custo multimódulo a ser minimizada $J_{\text{SWMM},i}$, escrita como

$$J_{\text{SWMM},i} = -\frac{|C_{2,2}^{y_{i,\text{R}}(n)}|}{(C_{1,1}^{y_{i,\text{R}}(n)})^2} - \frac{|C_{2,2}^{y_{i,\text{I}}(n)}|}{(C_{1,1}^{y_{i,\text{I}}(n)})^2} + \varrho \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} |r_{i,m}(k)|^2. \quad (5.40)$$

Cabe notar que, na abordagem multimódulo, o termo adicional de descorrelação não deve ser decomposto em partes real e imaginária, pois isso impossibilitaria a verificação de correlação quando há rotação complexa de uma das saídas recuperadas, como observa-se na Seção 3.2 do Capítulo 3.

Por sua vez, a função custo multiusuário de Godard da Equação (5.39), segundo uma abordagem multimódulo, pode ser escrita como

$$J_{\text{GMM},i} = \text{E} \left[\left(y_{\text{R}}^2(n) - \mathcal{R}_2^{\text{aR}} \right)^2 \right] + \text{E} \left[\left(y_{\text{I}}^2(n) - \mathcal{R}_2^{\text{aI}} \right)^2 \right] + \varrho \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} |r_{i,m}(k)|^2. \quad (5.41)$$

Nota-se que essa é uma versão estocástica da função custo determinista apresentada na Equação (3.1) do Capítulo 3, apresentada novamente a seguir, por conveniência:

$$\hat{J}_{\text{GMM},i}(n) \triangleq (1 - \lambda) \sum_{\ell=0}^n \lambda^{n-\ell} \left\{ \left[\mathcal{R}_2^{\text{aR}} - y_{i,\text{R}}^2(\ell, n) \right]^2 + \left[\mathcal{R}_2^{\text{aI}} - y_{i,\text{I}}^2(\ell, n) \right]^2 \right\} + \varrho \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{k=0}^{k_1} |c_{i,m}(n, k)|^2, \quad (5.42)$$

em que $y_{i,r}(\ell, n)$ e $y_{i,i}(\ell, n)$ são, respectivamente, as partes real e imaginária de $y_i(\ell, n) \triangleq \mathbf{u}^\top(\ell) \mathbf{w}_i(n)$, $0 \ll \lambda < 1$ é o fator de esquecimento e $c_{i,m}(n, k)$ é a correlação cruzada determinista definida em (3.4).

Portanto, o mesmo algoritmo de Shalvi-Weinstein deduzido no Capítulo 3 pode também ser obtido a partir da minimização de (5.41) com o método de quase-Newton, utilizando-se uma estimativa com janela exponencial para a matriz de covariância do vetor regressor. A vantagem da dedução determinista apresentada no Capítulo 3, baseada no problema de mínimos quadrados, é que ela evidencia as aproximações feitas para se obter o algoritmo de Shalvi-Weinstein MIMO – o que não fica explícito minimizando-se (5.41) com o método de quase-Newton. Além disso, a interpretação da função custo determinista, a partir da ótica estocástica, pode auxiliar na compreensão e na obtenção de algoritmos de separação cega de fontes mais abrangentes.

5.3 SEPARAÇÃO DE FONTES BASEADA EM CUMULANTES

No âmbito da desconvolução cega, o conhecimento adicional das estatísticas das fontes, devido ao fato de o alfabeto transmitido ser finito, é usado em favor da obtenção de critérios que evitam problemas de ambiguidade e recuperação encontrados em soluções mais gerais de BSS. Sendo assim, as funções custo apresentadas na Seção 5.2 sempre dependem de conhecimento exato das estatísticas das fontes, do qual não se dispõe em problemas mais amplos de BSS.

Em geral, assume-se que as fontes são independentes entre si e não gaussianas em BSS. Nesse caso, os critérios de separação a serem estabelecidos devem ser mais amplos. Embora Benveniste, Goursat e Ruget tenham mostrado, em 1980, que o problema de separação de fontes não gaussianas poderia ser resolvido igualando-se as funções densidade de probabilidade das fontes e dos sinais recuperados (BENVENISTE; GOURSAT; RUGET, 1980), nos problemas usuais de BSS, não se tem conhecimento das distribuições estatísticas das fontes.

Por outro lado, critérios que garantem a recuperação da independência estatística das fontes conseguem separá-las, como demonstrado nos trabalhos seminais de Comon em 1992 e 1994 (COMON, 1992; COMON, 1994), com base em resultados encontrados por Darmois em 1953 (DARMOIS, 1953). No entanto, para se garantir a independência na saída do sistema separador, é preciso aplicar restrições em estatísticas de todas as ordens, ao contrário da imposição de decorrelação, que requer somente condições sobre os momentos de ordem dois. Felizmente, em 1982, Lii e Rosenblatt indicaram, no contexto de desconvolução cega, que é possível realizar a recuperação cega das fontes utilizando-se somente restrições sobre momentos de ordem dois e sobre algum outro momento de ordem superior a dois (LII; ROSENBLATT, 1982). Cabe notar que, posteriormente, isso permitiu

a obtenção de algoritmos eficientes para a resolução de problemas de BSS baseados em momentos de ordem superior.

Em separação cega de fontes, as funções custo que busca-se minimizar para separar as fontes são denominadas contrastes (MOREAU; COMON, 2010). Os principais contrastes existentes são baseados em máxima verossimilhança e teoria da informação. Quando há conhecimento *a priori* que pode ser incorporado no modelo para melhorar a qualidade da separação, costuma-se utilizar critérios de máxima verossimilhança (CARDOSO, 1998). Dentre os critérios baseados em teoria da informação, destacam-se o Infomax (BELL; SEJNOWSKI, 1995) e critérios baseados em informação mútua. É possível demonstrar que existe uma equivalência entre a minimização de critérios baseados em verossimilhança e critérios baseados em teoria da informação (CARDOSO, 1998).

Como alternativa ao uso desses contrastes, que muitas vezes envolvem alto custo computacional e requerem conhecimento das funções densidade de probabilidade das fontes, é possível aproximá-los utilizando estatísticas de ordem superior. Em outras palavras, estatísticas de ordem superior podem ser utilizadas para definir funções contraste que consistem em simples aproximações dos contrastes de máxima verossimilhança e baseados em teoria da informação. Usualmente, isso é feito com o auxílio de cumulantes. Do ponto de vista prático, essas aproximações são essenciais para viabilizar a obtenção de algoritmos. Assim, esses contrastes aproximados são extremamente importantes.

Na abordagem de separação cega de fontes, é usual considerar uma etapa preliminar de pré-branqueamento ou descorrelação das misturas, antes de realizar a etapa de separação (HYVÄRINEN; OJA, 2000). Se o pré-branqueamento for realizado antes de se utilizar o algoritmo de separação, podem ser elencados dois contrastes principais na área de BSS para misturas instantâneas e fontes reais. O primeiro, denominado como contraste de ICA (do inglês, *independent component analysis*), corresponde a uma aproximação para o critério de informação mútua e é dado por (CARDOSO, 1998)

$$\phi_{\text{ICA}}^{\circ}(y_1, y_2, \dots, y_{N_u}) = \sum_{(i,j,k,\ell) \neq (i,i,i,i)} \text{cum}^2(y_i, y_j, y_k, y_\ell) = - \sum_{i=1}^{N_u} (C_4^{y_i})^2 + \text{constante}, \quad (5.43)$$

em que y_i , com $i = 1, 2, \dots, N_u$, são as fontes estimadas a partir do sistema separador. O segundo contraste, conhecido como JADE (do inglês, *joint approximate diagonalization of eigenmatrices*), testa a independência sobre um conjunto menor de cumulantes cruzados, e é dado por (CARDOSO, 1998)

$$\phi_{\text{JADE}}^{\circ}(y_1, y_2, \dots, y_{N_u}) = \sum_{(i,j,k,\ell) \neq (i,j,k,k)} \text{cum}^2(y_i, y_j, y_k, y_\ell). \quad (5.44)$$

Note que ambos os contrastes (5.43) e (5.44) não requerem conhecimento prévio de cumulantes das fontes, mas necessitam de estimativas de curtose das saídas, que podem ser feitas utilizando-se médias empíricas (HYVÄRINEN; OJA, 2000). Caso esse conhecimento adicional esteja disponível, os contrastes podem ser simplificados (CARDOSO, 1998).

Do ponto de vista estatístico, os critérios (5.43) e (5.44) fazem sentido, pois cumulantes, e em particular a curtose, são medidas de não gaussianidade. Pelo Teorema do Limite Central, ao se misturar fontes não gaussianas, as misturas devem ser “mais” gaussianas do que as fontes. Portanto, ao se descobrir a direção de máxima curtose da distribuição das saídas do sistema separador, pode-se inferir qual é a matriz que compôs as misturas. Isso é feito em algoritmos do tipo *projection pursuit* com deflação (HYVÄRINEN; OJA, 2000; STONE, 2004).

Embora a estimação de curtose seja direta, ela é uma medida muito sensível a *outliers* (HYVÄRINEN; OJA, 2000). Trata-se de uma característica que pode refletir uma desvantagem prática ao se utilizar curtose como critério de separação de fontes. No entanto, cabe observar que a solução numericamente robusta encontrada no Capítulo 3 para equalização cega multiusuário, utilizando uma operação em *dual-mode* para evitar divergência devido a erros em estimativas não lineares, e com propriedades numéricas e de convergência boas, pode ser investigada para ser aplicada nesse contexto de BSS.

5.4 ABORDAGENS PARA MISTURAS CONVOLUTIVAS

Os critérios de separação para misturas convolutivas são mais complicados que os critérios para misturas instantâneas, sendo muitas vezes considerados na literatura para o caso particular de $N_u = 2$ fontes e $L = 2$ sensores (YELLIN; WEINSTEIN, 1994). Os principais critérios para misturas convolutivas envolvendo estatísticas de ordem superior se dividem em abordagens no domínio do tempo e no domínio da frequência (MALMIRI, 2002).

No domínio do tempo, algumas abordagens envolvendo critérios baseados em cumulantes consistem em (MALMIRI, 2002):

- Aplicar extensões para o caso convolutivo de algoritmos adaptativos para misturas instantâneas, buscando, por exemplo, minimizar a correlação cruzada entre não linearidades ímpares de saídas distintas atrasadas entre si (JUTTEN et al., 1991).
- Maximizar cumulantes de quarta ordem da saída do equalizador (TUGNAIT, 1998).
- Minimizar cumulantes de quarta ordem cruzados entre fontes atrasadas entre si (NGUYEN THI; JUTTEN, 1995).
- Cancelar poliespectros de ordem dois ou três das saídas (YELLIN; WEINSTEIN, 1994).

Por outro lado, é possível lançar mão de representações no domínio da frequência para resolver o problema de separação de misturas convolutivas. Partindo-se das representações no domínio transformado z para o problema de BSS, segundo as Equações (2.19),

(2.2.2) e (2.2.2), apresentadas no Capítulo 2, é possível fazer, sob certas condições, a substituição $z = e^{j\omega}$ nessas equações e analisá-las no domínio da frequência angular normalizada ω . As representações em ω , ou transformadas de Fourier de tempo discreto, são úteis em aplicações práticas pois podem ser estimadas utilizando-se algoritmos usuais, como a FFT (do inglês, *fast Fourier transform*). Dessa forma, o processo de mistura pode ser escrito como

$$\mathbf{u}[e^{j\omega}] = \mathbf{H}[e^{j\omega}]\mathbf{a}[e^{j\omega}], \quad (5.45)$$

em que

$$\mathbf{u}[e^{j\omega}] = \sum_{n=-\infty}^{+\infty} \mathbf{u}(n) e^{-j\omega n}, \quad (5.46)$$

$$\mathbf{H}[e^{j\omega}] = \sum_{n=-\infty}^{+\infty} \mathbf{H}(n) e^{-j\omega n}, \quad (5.47)$$

$$\mathbf{a}[e^{j\omega}] = \sum_{n=-\infty}^{+\infty} \mathbf{a}(n) e^{-j\omega n}, \quad (5.48)$$

correspondem, respectivamente, às transformadas de Fourier do vetor de misturas $\mathbf{u}(n)$, dado pela Equação (2.2), da resposta ao pulso unitário do sistema convolutivo MIMO $\mathbf{H}(n)$, dada pela Equação (2.17), e do vetor de fontes $\mathbf{a}(n)$, dado pela Equação (2.1). A partir da Equação (5.45), nota-se que, para cada frequência angular ω fixada, a mistura convolutiva pode ser interpretada como instantânea no domínio da frequência. Assim, busca-se resolver o problema de separação de fontes lançando-se mão de algoritmos de ICA para misturas instantâneas que, por sua vez, devem ser aplicados para resolver o problema em cada frequência do domínio transformado. Nesse caso, deve-se tomar cuidado especial com as ambiguidades de permutação do ICA instantâneo, que podem fazer com que as fontes sejam recuperadas em ordens diferentes para cada frequência, impossibilitando a antitransformação correta do resultado para o domínio do tempo (MALMIRI, 2002).

5.5 CONCLUSÃO

A compreensão das relações entre os critérios de separação de fontes e as soluções de desconvolução multiusuário não é uma tarefa trivial. Em (CAVALCANTE, 2004; ROMANO et al., 2011), as relações entre os critérios de BSS foram adequadamente tratadas do ponto de vista de funções contraste. Além disso, foram estabelecidas as relações com alguns critérios de desconvolução cega de canais de comunicação. Nesse trabalho de mestrado, inspirando-se em (CAVALCANTE, 2004; ROMANO et al., 2011) tentou-se particularizar essas relações para o caso de funções contraste baseadas em cumulantes de ordem superior. No Capítulo 3, os algoritmos eficientes de desconvolução cega foram obtidos para misturas convolutivas e supondo conhecidas as estatísticas das fontes, caracterizadas por um alfabeto finito. No caso de BSS, o sistema misturador é convolutivo, porém não necessariamente

são conhecidas as estatísticas das fontes. Isso dificulta o processo de separação, pois torna-se necessário estimar os cumulantes das fontes de alguma forma. Portanto, pode ser interessante estender critérios de desconvolução cega para problemas mais gerais de BSS.

A obtenção de algoritmos adaptativos para o problema de desconvolução cega de canais de comunicação multiusuário não é trivial. Os algoritmos cegos estendidos para o caso MIMO precisam de uma modificação que possibilite que equalizadores diferentes convirjam para fontes distintas em tempo real. Uma possível estratégia, usada nesse trabalho e abordada na Seção 5.2.3, consiste em considerar um termo que penaliza a correlação cruzada entre as saídas do equalizador no erro do algoritmo. Cabe notar que, com o correto ajuste dos parâmetros relacionados a esse termo, é possível forçar a convergência para fontes distintas. Conjectura-se a possibilidade de se utilizar um termo similar em funções custo de algoritmos para separação cega de fontes em trabalhos futuros. Adicionalmente, a passagem de funções custo de desconvolução cega multiusuário, como as das Equações (5.39) e (5.40), para o caso de separação cega de fontes, exige conhecimento estatístico das fontes que pode ser inferido a partir de técnicas cegas baseadas em aproximações de funções densidade de probabilidade, como sugerido em (CAVALCANTE, 2004).

6 CONCLUSÕES

Separação cega de fontes (BSS) é um ramo de pesquisa extremamente amplo e diverso dentro da área de Processamento Estatístico de Sinais. Além de serem inúmeras as formas de se colocar o problema de BSS, as soluções existentes e as suas aplicações são muito variadas. Cabe notar ainda que compreensão do problema e de suas soluções é desafiadora, pois envolve conceitos matemáticos e estatísticos avançados, dada a dificuldade do problema em si: recuperar fontes a partir somente da observação de suas misturas. Para adentrar o universo de separação cega de fontes e começar a compreender seus fundamentos teóricos e soluções, é fundamental delimitar um caminho de estudo que seja menos abrangente. Caso contrário, a tarefa seria impossível de ser feita em um tempo finito.

Tendo isso em vista, nesse trabalho o enfoque foi dado para uma aplicação de separação cega de fontes na área de comunicações digitais conhecida como desconvolução cega multiusuário. Esse problema é compreendido como um caso particular de separação cega de fontes de misturas convolutivas em que as estatísticas das fontes transmitidas são bem conhecidas. Durante o trabalho, buscou-se compreender o problema de desconvolução no contexto de separação cega de fontes, investigar as soluções existentes para esse problema e aprimorá-las. Isso tudo com o intuito de possibilitar a posterior abordagem e compreensão teórica de problemas mais gerais e abrangentes.

No Capítulo 2, os problemas de separação cega de fontes e desconvolução cega multiusuário foram apresentados. A formulação do problema de BSS foi inicialmente feita de forma genérica, e posteriormente particularizada para misturas lineares instantâneas e convolutivas. Essa particularização permitiu compreender as principais hipóteses usualmente adotadas na solução do problema para cada tipo de mistura considerado, e algumas limitações existentes na forma como resolvê-lo. Em seguida, o problema de desconvolução cega multiusuário foi formulado no contexto de BSS.

Devido à abrangência do problema de BSS e às inúmeras frentes de estudo existentes nessa área, voltou-se o enfoque à desconvolução cega multiusuário nos Capítulos 3 e 4.

No Capítulo 3, foram propostas soluções adaptativas para a desconvolução cega multiusuário com base no algoritmo de Shalvi-Weinstein (SWA) multiusuário. Nesse capítulo, buscou-se aprimorar os algoritmos encontrados em termos de boa capacidade de separação e robustez numérica, lançando-se mão de um tratamento para as não linearidades do algoritmo e de uma implementação robusta baseada em decomposição QR. Um algoritmo do tipo gradiente estocástico, inspirado nos algoritmos propostos nesse trabalho e mantendo as boas propriedades numéricas e de capacidade de separação, resultou na apresentação e

publicação de um artigo (PAVAN; SILVA; MIRANDA, 2016) no *workshop* internacional *IEEE Sensor Array and Multichannel Signal Processing (SAM)* de 2016.

No Capítulo 4, fez-se a análise da solução obtida no capítulo anterior em termos de estabilidade numérica (análise determinista) e desempenho em regime.

Após esse estudo mais aprofundado do problema específico de desconvolução multiusuário, no Capítulo 5, voltou-se a abordar o problema de separação cega de fontes em uma apresentação das principais funções custo baseadas em curtose de cada um desses problemas. Isso foi feito com o objetivo de permitir a extensão dos resultados de desconvolução cega obtidos para problemas mais amplos de BSS, que pretende-se fazer futuramente.

SUGESTÕES DE TRABALHOS FUTUROS

Após esse trabalho, pode-se concluir que ainda há muito mais a se estudar na área de separação cega de fontes. Com base no trabalho realizado, algumas sugestões para trabalhos futuros são:

- Compreender com maior profundidade as relações existentes entre as soluções de desconvolução cega multiusuário e separação cega de fontes baseadas em curtose, aventando a possibilidade de se estender os resultados de desconvolução multiusuário para casos de BSS mais gerais.
- Verificar a influência das considerações *a priori* para o problema de BSS e o seu efeito nas soluções do problema e no seu desempenho, tanto para o caso de misturas instantâneas quanto convolutivas.
- Obter algoritmos eficientes para separação cega de fontes do ponto de vista de boa capacidade de separação, baixo custo computacional e robustez numérica.

REFERÊNCIAS BIBLIOGRÁFICAS

- ADALI, T.; ANDERSON, M.; FU, G.-S. Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging. *IEEE Signal Processing Magazine*, v. 31, n. 3, p. 18–33, maio 2014.
- APOLINÁRIO JR., J. A. (Ed.). *QRD-RLS Adaptive Filtering*. New York City: Springer, 2009.
- BELL, A. J.; SEJNOWSKI, T. J. An information-maximisation approach to blind separation and blind deconvolution. *Neural Computation*, v. 7, n. 6, p. 1129–1159, 1995.
- BENVENISTE, A.; GOURSAT, M.; RUGET, G. Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communications. *IEEE Transactions on Automatic Control*, AC-25, n. 3, p. 385–399, 1980.
- BRANDWOOD, D. H. A complex gradient operator and its application in adaptive array theory. *Communications, Radar and Signal Processing, IEE Proceedings F*, v. 130, n. 1, p. 11–16, fev. 1983.
- BROWN, G. D.; YAMADA, S.; SEJNOWSKI, T. J. Independent component analysis at the neural cocktail party. *Trends in Neurosciences*, v. 24, n. 1, p. 54–63, jan. 2001.
- BROWN, J. W.; CHURCHILL, R. V. *Complex Variables and Applications*. 9. ed. New York City: McGraw-Hill Education, 2013.
- CARDOSO, J.-F. Blind signal separation: statistical principles. *Proceedings of the IEEE*, v. 86, n. 10, p. 2009–2025, out. 1998.
- CASTELLA, M.; CHEVREUIL, A.; PESQUET, J.-C. Convolutional mixtures. In: COMON, P.; JUTTEN, C. (Ed.). *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Oxford: Academic Press, 2010. cap. 8, p. 281–324.
- CAVALCANTE, C. C. *Sobre separação cega de fontes: proposições e análise de estratégias para processamento multi-usuário*. Tese (Doutorado) — Universidade Estadual de Campinas (UNICAMP), 2004.
- CHEVALIER, P.; CHEVREUIL, A. Application to telecommunications. In: COMON, P.; JUTTEN, C. (Ed.). *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Oxford: Academic Press, 2010. cap. 17, p. 683–725.
- COMON, P. Independent component analysis. *Higher Order Statistics*, p. 29–38, 1992.
- COMON, P. Independent component analysis, A new concept? *Signal Processing*, v. 36, n. 3, p. 287–314, abr. 1994.
- DARMOIS, G. Analyse générale des liaisons stochastiques: étude particulière de l'analyse factorielle linéaire. *Revue de l'Institut International de Statistique*, v. 21, n. 1/2, p. 2–8, 1953.

DING, Z.; LI, Y. *Blind Equalization and Identification*. New York City: Marcel Dekker, 2001. (Signal Processing and Communication Series).

DUARTE, L. T.; MOUSSAOUI, S.; JUTTEN, C. Source separation in chemical analysis: Recent achievements and perspectives. *IEEE Signal Processing Magazine*, v. 31, n. 3, p. 135–146, maio 2014.

EWERT, S. et al. Score-informed source separation for musical audio recordings: An overview. *IEEE Signal Processing Magazine*, v. 31, n. 3, p. 116–124, maio 2014.

FIJALKOW, I.; MANLOVE, C. E.; JOHNSON, C. R. Adaptive fractionally spaced blind CMA equalization: excess MSE. *IEEE Transactions on Signal Processing*, v. 46, n. 1, p. 227–231, jan. 1998.

GODARD, D. Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *IEEE Transactions on Communications*, v. 28, n. 11, p. 1867–1875, nov. 1980.

HAYKIN, S. (Ed.). *Unsupervised Adaptive Filtering. Volume I: Blind Source Separation*. New Jersey: John Wiley & Sons, 2000a. (Adaptive and Learning Systems for Signal Processing, Communications, and Control).

HAYKIN, S. (Ed.). *Unsupervised Adaptive Filtering. Volume II: Blind Deconvolution*. New Jersey: John Wiley & Sons, 2000b. (Adaptive and Learning Systems for Signal Processing, Communications, and Control).

HAYKIN, S. *Adaptive Filter Theory*. 5. ed. New York City: Pearson Education, 2014.

HÉRAULT, J.; JUTTEN, C.; ANS, B. Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In: *Actes du Xème colloque GRETSI*. Nice, France: Groupe d'Etudes du Traitement du Signal et des Images (GRETSI), 1985. p. 1017–1022.

HYVÄRINEN, A.; OJA, E. Independent component analysis: algorithms and applications. *Neural Networks*, v. 13, n. 4–5, p. 411–430, 2000.

JEFFREY, A.; DAI, H.-H. *Handbook of Mathematical Formulas and Integrals*. 4. ed. Oxford: Academic Press, 2008.

JUTTEN, C.; COMON, P. Introduction. In: COMON, P.; JUTTEN, C. (Ed.). *Handbook of Blind Source Separation: Independent Component analysis and Applications*. Oxford: Academic Press, 2010. cap. 1, p. 1–28.

JUTTEN, C. et al. Blind separation of sources: an algorithm for separation of convolutive mixtures. In: *Higher Order Statistics, 1991*. Chamrousse, France: Proceedings of the International Signal Processing Workshop on Higher Order Statistics, 1991. p. 273–276.

KAILATH, T.; SAYED, A.; HASSIBI, B. *Linear Estimation*. 1. ed. New York City: Pearson, 2000.

KENDALL, M.; STUART, A. *The Advanced Theory of Statistics, Distribution Theory*. London: Charles Griffin, 1977. v. 1–3.

- KREUTZ-DELGADO, K. *The Complex Gradient Operator and the CR-calculus*. 2009. Disponível em: <<http://arxiv.org/abs/0906.4835>>. Acesso em: 1 jul. 2016.
- LI, Y. et al. Sparse representation for brain signal processing: A tutorial on methods and applications. *IEEE Signal Processing Magazine*, v. 31, n. 3, p. 96–106, maio 2014.
- LII, K.-S.; ROSENBLATT, M. Deconvolution and estimation of transfer function phase and coefficients for non-gaussian linear processes. *The Annals of Statistics*, v. 10, n. 4, p. 1195–1208, 1982.
- LUCKY, R. W. Techniques for adaptive equalization of digital communication systems. *Bell Systems Technical Journal*, v. 45, p. 255–286, fev. 1966.
- LUO, Y.; CHAMBERS, J. A. Steady-state mean-square error analysis of the cross-correlation and constant modulus algorithm in a MIMO convolutive system. *IEE Proceedings – Vision, Image and Signal Processing*, v. 149, n. 4, p. 196–203, ago. 2002.
- MAI, J.; SAYED, A. H. A feedback approach to the steady-state performance of fractionally spaced blind adaptive equalizers. *IEEE Transactions on Signal Processing*, v. 48, n. 1, p. 80–91, jan. 2000.
- MALMIRI, M. B. *On blind source separation in convolutive and nonlinear mixtures*. Tese (Doutorado) — Institut National Polytechnique de Grenoble (Grenoble INP), 2002.
- MCCULLAGH, P. *Tensor methods in statistics*. London: Chapman and Hall, 1987.
- MENDES-FILHO, J.; MIRANDA, M. D.; SILVA, M. T. M. A regional multimodulus algorithm for blind equalization of QAM signals: introduction and steady-state analysis. *Signal Processing*, v. 92, n. 11, p. 2643–2656, nov. 2012.
- MIRANDA, M. D.; SILVA, M. T. M.; NASCIMENTO, V. H. Avoiding divergence in the Shalvi-Weinstein algorithm. *IEEE Transactions on Signal Processing*, v. 56, n. 11, p. 5403–5413, nov. 2008.
- MOREAU, E.; COMON, P. Contrasts. In: COMON, P.; JUTTEN, C. (Ed.). *Handbook of Blind Source Separation: Independent Component analysis and Applications*. Oxford: Academic Press, 2010. cap. 3, p. 65–105.
- NGUYEN THI, H. L.; JUTTEN, C. Blind sources separation for convolutive mixtures. *Signal Processing*, v. 45, p. 209–229, 1995.
- NIKIAS, C. L.; PETROPULU, A. P. *Higher-Order Spectra Analysis. A Nonlinear Signal Processing Framework*. Upper Saddle River, New Jersey: Prentice Hall, 1993.
- OWEISS, K. G. (Ed.). *Statistical Signal Processing for Neuroscience and Neurotechnology*. 1. ed. Oxford: Academic Press, 2010.
- PAPADIAS, C. B.; PAULRAJ, A. J. A constant modulus algorithm for multiuser signal separation in presence of delay spread using antenna arrays. *IEEE Signal Processing Letters*, v. 4, n. 6, p. 178–181, jun. 1997a.
- PAPADIAS, C. B.; PAULRAJ, A. J. Blind separation of independent co-channel signals. In: *13th International Conference on Digital Signal Processing, 1997*. Santorini, Greece: Proceedings of the 13th International Conference on Digital Signal Processing, 1997b. v. 1, p. 139–142.

- PAPOULIS, A.; PILLAI, S. U. *Probability, Random Variables and Stochastic Processes*. 4. ed. London: McGraw-Hill Europe, 2002.
- PAVAN, F. R. M.; SILVA, M. T. M.; MIRANDA, M. D. Avoiding divergence in the constant modulus algorithm for blind equalization of MIMO systems. In: *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), 2016*. Rio de Janeiro, Brazil: Proceedings of the 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop, 2016.
- PEEBLES, P. Z. *Probability, Random Variables, and Random Signal Principles*. 4. ed. New York City: McGraw-Hill, 2000.
- PICINBONO, B. *Random Signals and Systems*. Upper Saddle River, New Jersey: Prentice Hall, 1993.
- PROAKIS, J.; SALEHI, M. *Digital Communications*. 5. ed. New York City: McGraw-Hill, 2007.
- REGALIA, P. A. Numerical stability properties of a QR-based fast least squares algorithm. *IEEE Transactions on Signal Processing*, v. 41, p. 2096–2109, jun. 1993.
- REGALIA, P. A. Subspace methods in system identification and source localization. In: *Algorithms and Parallel VLSI Architectures III*. Leuven, Belgium: Elsevier, 1995. p. 13–23.
- REGALIA, P. A. On the equivalence between the Godard and Shalvi-Weinstein schemes of blind equalization. *Signal Processing*, v. 73, p. 185–190, jan. 1999.
- RIVET, B. et al. Audiovisual speech source separation: An overview of key methodologies. *IEEE Signal Processing Magazine*, v. 31, n. 3, p. 125–134, maio 2014.
- ROMANO, J. M. T. et al. *Unsupervised signal processing: channel equalization and source separation*. Florida: CRC Press, 2011.
- RUPP, M. Convergence properties of adaptive equalizer algorithms. *IEEE Transactions on Signal Processing*, v. 59, n. 6, p. 2562–2574, jun. 2011.
- RUPP, M. Adaptive filters: stable but divergent. *EURASIP Journal on Advances in Signal Processing*, v. 2015, n. 1, p. 1–15, 2015.
- RUPP, M.; SAYED, A. H. A time-domain feedback analysis of filtered-error adaptive gradient algorithms. *IEEE Transactions on Signal Processing*, v. 44, n. 6, p. 1428–1439, jun. 1996.
- SATO, Y. A method of self-recovering equalization for multi-level amplitude modulation. *IEEE Transactions on Communications*, v. 23, p. 679–682, jun. 1975.
- SAYED, A. H. *Adaptive Filters*. 1. ed. New Jersey: Wiley-IEEE Press, 2008.
- SHAH, S. A. W.; ABED-MERAIM, K.; AL-NAFFOURI, T. Y. Multi-modulus algorithms using hyperbolic and givens rotations for blind deconvolution of MIMO systems. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015*. Brisbane: Proceedings of the ICASSP, 2015. p. 2155–2159.

- SHALVI, O.; WEINSTEIN, E. New criteria for blind deconvolution of nonminimum phase systems (channels). *IEEE Transactions on Information Theory*, v. 36, n. 2, p. 312–321, mar. 1990.
- SHALVI, O.; WEINSTEIN, E. Super-exponential methods for blind deconvolution. *IEEE Transactions on Information Theory*, v. 39, n. 2, p. 504–519, mar. 1993.
- SHALVI, O.; WEINSTEIN, E. Universal methods for blind deconvolution. In: HAYKIN, S. (Ed.). *Blind Deconvolution*. Upper Saddle River, New Jersey: Prentice Hall, 1994. cap. 4.
- SILVA, M. T. M.; MIRANDA, M. D. Tracking issues of some blind equalization algorithms. *IEEE Signal Processing Letters*, v. 11, n. 9, p. 760–763, set. 2004.
- SILVA, M. T. M.; MIRANDA, M. D. Tracking analysis of some space-time blind equalization algorithms. In: *IEEE/SP 13th Workshop on Statistical Signal Processing, 2005*. Bordeaux, France: Proceedings of the SSP, 2005. p. 167–172.
- SILVA, M. T. M.; MIRANDA, M. D.; LICCIARDI, A. N. J. A robust algorithm for blind space-time equalization. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2004*. Montreal: Proceedings of the ICASSP, 2004. v. 4, p. 857–860.
- SOLO, V.; KONG, X. *Adaptive Signal Processing Algorithms: Stability and Performance*. Upper Saddle River, New Jersey: Prentice Hall, 1995.
- STONE, J. V. *Independent Component Analysis: A Tutorial Introduction*. Cambridge, Massachussets: MIT Press, 2004. (A Bradford Book).
- TREICHLER, J.; AGEE, B. A new approach to multipath correction of constant modulus signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 31, n. 2, p. 459–472, abr. 1983.
- TREICHLER, J. R.; FIJALKOW, I.; JOHNSON, C. R. Fractionally spaced equalizers. *IEEE Signal Processing Magazine*, v. 13, n. 3, p. 65–81, maio 1996.
- TUGNAIT, J. K. On blind separation of convolutive mixtures of independent linear signals in unknown additive noise. *IEEE Transactions on Signal Processing*, v. 46, n. 11, p. 3117–3123, nov. 1998.
- VINCENT, E. et al. From blind to guided audio source separation: How models and side information can improve the separation of sound. *IEEE Signal Processing Magazine*, v. 31, n. 3, p. 107–115, maio 2014.
- YANG, J.; WERNER, J. J.; DUMONT, G. A. The multimodulus blind equalization and its generalized algorithms. *IEEE Journal on Selected Areas in Communications*, v. 20, n. 5, p. 997–1015, jun. 2002.
- YELLIN, D.; WEINSTEIN, E. Criteria for multichannel signal separation. *IEEE Transactions on Signal Processing*, v. 42, n. 8, p. 2158–2168, ago. 1994.
- YOUSEF, N. R.; SAYED, A. H. A feedback analysis of the tracking performance of blind adaptive equalization algorithms. In: *38th IEEE Conference on Decision and Control, 1999*. Phoenix, Arizona: Proceedings of the 38th IEEE Conference on Decision and Control, 1999. v. 1, p. 174–179.

YUAN, J.-T.; TSAI, K.-D. Analysis of the multimodulus blind equalization algorithm in QAM communication systems. *IEEE Transactions on Communications*, v. 53, n. 9, p. 1427–1431, set. 2005.

ZARZOSO, V.; COMON, P.; SLOCK, D. Semi-blind methods for communications. In: COMON, P.; JUTTEN, C. (Ed.). *Handbook of Blind Source Separation: Independent Component analysis and Applications*. Oxford: Academic Press, 2010. cap. 15, p. 593–638.

APÊNDICE A – CÁLCULO DE WIRTINGER VETORIAL

A função custo $\widehat{J}_i(n)$ de (3.1) pode ser interpretada como uma função real e escalar, sempre maior ou igual a zero, dos vetores complexos $\mathbf{w}_i(n)$ e $\mathbf{w}_i^H(n)$ que compõem as saídas $y_i(\ell, n)$. No problema de otimização considerado, busca-se obter os pontos estacionários de $\widehat{J}_i(n)$. Uma alternativa seria levantar os pontos em que as derivadas de $\widehat{J}_i(n)$ em relação a cada elemento de $\mathbf{w}_i(n)$ e $\mathbf{w}_i^H(n)$ são nulas.

No entanto, cabe observar que $\widehat{J}_i(n)$ é uma função que assume valores reais, mas possui variáveis independentes que podem assumir valores complexos. Logo, é possível mostrar que $\widehat{J}_i(n)$ não é uma função analítica em domínio complexo, e, por isso, não é diferenciável no sentido padrão adotado em cursos de Análise Complexa (BROWN; CHURCHILL, 2013).

Felizmente, os pontos estacionários de $\widehat{J}_i(n)$ podem ser obtidos utilizando-se o Cálculo de Wirtinger. Esse cálculo estende o conceito de derivada complexa para funções que não satisfazem as condições de Cauchy-Riemann (BRANDWOOD, 1983; KREUTZ-DELGADO, 2009; HAYKIN, 2014), como é o caso de $\widehat{J}_i(n)$.

A abordagem do cálculo de Wirtinger, para o caso de funções de uma variável complexa, consiste em interpretar uma função real de variável complexa $z = x + jy$ como uma função real de variáveis reais x e y . Com esse “truque”, pode-se utilizar uma definição de derivada mais ampla que aquela adotada em Análise Complexa. Pode-se garantir que essa “nova” derivada exista até mesmo para funções não analíticas em domínios complexos. Com isso, pode-se calcular os pontos estacionários de funções não analíticas em um domínio complexo (KREUTZ-DELGADO, 2009).

A seguir, os principais resultados do Cálculo de Wirtinger multivariável são apresentados. É importante notar que existem diversas formas de se elaborar a teoria apresentada a seguir, mantendo-se sua consistência (KREUTZ-DELGADO, 2009; HAYKIN, 2014). A forma como a teoria é elaborada não influencia nos resultados finais, *i.e.*, os pontos estacionários que se deseja calcular.

A.1 OPERADOR COGRADIENTE

Seja o vetor coluna

$$\mathbf{z} = [z_1 \quad z_2 \quad \cdots \quad z_N]^T \in \mathbb{C}^N, \quad (\text{A.1})$$

em que $z_i = x_i + jy_i$, com $i = 1, 2, \dots, N$. Equivalentemente, pode-se escrever

$$\mathbf{z} = \mathbf{x} + j\mathbf{y}, \quad (\text{A.2})$$

em que

$$\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_N]^\top \in \mathbb{R}^N, \quad (\text{A.3})$$

$$\mathbf{y} = [y_1 \ y_2 \ \cdots \ y_N]^\top \in \mathbb{R}^N. \quad (\text{A.4})$$

A partir de (A.2), o complexo conjugado de $\mathbf{z} \in \mathbb{C}^N$ vale

$$\mathbf{z}^* = \mathbf{x} - j\mathbf{y}. \quad (\text{A.5})$$

O operador cogradiente é definido como a um vetor linha dado por

$$\frac{\partial}{\partial \mathbf{z}} \triangleq \left[\frac{\partial}{\partial z_1} \ \frac{\partial}{\partial z_2} \ \cdots \ \frac{\partial}{\partial z_N} \right], \quad (\text{A.6})$$

em que z_i , com $i = 1, 2, \dots, N$, correspondem às coordenadas do vetor \mathbf{z} definido em (A.1), e os operadores de derivada parcial em cada componente são definidos como

$$\frac{\partial}{\partial z_i} \triangleq \frac{1}{2} \left(\frac{\partial}{\partial x_i} - j \frac{\partial}{\partial y_i} \right), \quad (\text{A.7})$$

para $i = 1, 2, \dots, N$. Vetorialmente, pode-se reescrever (A.7) como

$$\frac{\partial}{\partial \mathbf{z}} \triangleq \frac{1}{2} \left(\frac{\partial}{\partial \mathbf{x}} - j \frac{\partial}{\partial \mathbf{y}} \right). \quad (\text{A.8})$$

É importante notar que, quando se aplica o operador cogradiente $\frac{\partial}{\partial \mathbf{z}}$ a uma função escalar $f(\mathbf{z}, \mathbf{z}^*)$, o vetor \mathbf{z}^* é tratado como uma constante (BRANDWOOD, 1983; KREUTZ-DELGADO, 2009).

Alguns resultados de cogradientes de funções escalares $f(\mathbf{z}, \mathbf{z}^*)$ de variáveis complexas são elencados a seguir (KREUTZ-DELGADO, 2009; HAYKIN, 2014):

$$\frac{\partial}{\partial \mathbf{z}} \mathbf{a}^\top \mathbf{z} = \mathbf{a}^\top, \quad (\text{A.9})$$

$$\frac{\partial}{\partial \mathbf{z}} \mathbf{a}^\top \mathbf{z}^* = \mathbf{0}^\top, \quad (\text{A.10})$$

$$\frac{\partial}{\partial \mathbf{z}} \mathbf{z}^\top \mathbf{a} = \mathbf{a}^\top, \quad (\text{A.11})$$

$$\frac{\partial}{\partial \mathbf{z}} \mathbf{z}^H \mathbf{a} = \mathbf{0}^\top, \quad (\text{A.12})$$

$$\frac{\partial}{\partial \mathbf{z}} \mathbf{z}^\top \mathbf{z} = 2\mathbf{z}^\top, \quad (\text{A.13})$$

$$\frac{\partial}{\partial \mathbf{z}} \mathbf{z}^H \mathbf{z} = \mathbf{z}^H, \quad (\text{A.14})$$

$$\frac{\partial}{\partial \mathbf{z}} \mathbf{z}^\top \mathbf{A} \mathbf{z} = \mathbf{z}^\top (\mathbf{A} + \mathbf{A}^\top), \quad (\text{A.15})$$

$$\frac{\partial}{\partial \mathbf{z}} \mathbf{z}^H \mathbf{A} \mathbf{z} = \mathbf{z}^H, \quad (\text{A.16})$$

em que $\mathbf{z}, \mathbf{a} \in \mathbb{C}^N$ são vetores coluna de mesma dimensão, \mathbf{a} independe de \mathbf{z} , o vetor nulo é denotado por $\mathbf{0} \in \mathbb{C}^N$ e $\mathbf{A} \in \mathbb{C}^{N \times N}$ é uma matriz quadrada que independe de \mathbf{z} .

A.2 GRADIENTE COMPLEXO

Seja $f(\mathbf{z}, \mathbf{z}^*)$ uma função real escalar tal que $f : \mathbb{C}^N \rightarrow \mathbb{R}$. Mantendo-se a consistência com as definições estabelecidas na Seção A.1, o operador gradiente complexo usual de f é definido como (KREUTZ-DELGADO, 2009)

$$\nabla_{\mathbf{z}} f \triangleq \left(\frac{\partial f}{\partial \mathbf{z}} \right)^H. \quad (\text{A.17})$$

Por fim, se $f : \mathbb{C}^N \rightarrow \mathbb{R}$, é possível provar que uma condição necessária e suficiente para que f possua um ponto estacionário em \mathbb{C}^N é que (HAYKIN, 2014)

$$\nabla_{\mathbf{z}} f = \mathbf{0}. \quad (\text{A.18})$$

Portanto, em problemas de otimização, pode-se utilizar o gradiente complexo para encontrar pontos estacionários de funções custo reais com variáveis independentes complexas.

APÊNDICE B – DEDUÇÃO DE MOMENTOS DA ANÁLISE EM REGIME

Neste apêndice, inicialmente são determinados momentos aproximados de funções de $y_i(n)$ em termos de momentos de $a_i(n)$, $v_i(n)$ e $e_{a,i}(n)$. Posteriormente, a partir desses momentos obtidos, os termos de interesse $E[e_{a,i}^*(n)\varepsilon_i(n)]$ e $E[|\varepsilon_i(n)|^2]$ são representados em função de momentos de $a_i(n)$, $v_i(n)$ e $e_{a,i}(n)$.

B.1 DEFINIÇÕES

A fim de facilitar o cálculo dos momentos e sintetizar a notação durante as deduções, a seguinte notação é estabelecida. Seja f um processo aleatório qualquer. Definem-se:

$$\begin{aligned} f_{\text{R}} &\triangleq \text{Re}\{f\}, \\ f_{\text{I}} &\triangleq \text{Im}\{f\}, \\ r_{\alpha}^f &= E[|f|^{\alpha}], \\ \delta_{\ell,m} &= \begin{cases} 1, & \text{se } \ell = m \\ 0, & \text{caso contrário} \end{cases}. \end{aligned}$$

Além disso, para reduzir a notação, os índices de fonte e as variáveis de tempo são omitidos quando possível. Por conveniência, reescreve-se a Equação (4.63) para o i -ésimo equalizador como

$$\varepsilon = \bar{d} - a + v, \quad (\text{B.1})$$

em que, sem perda de generalidade, assume-se que $\theta_i = \tau_i = 0$. Considerando-se o funcionamento do algoritmo dentro da RoI (ver Seção 3.3), definem-se

$$N_{\text{t}} \triangleq (k_1 + 1)(N_{\text{u}} - 1), \quad (\text{B.2})$$

$$y = y_{\text{R}} + jy_{\text{I}}, \quad (\text{B.3})$$

$$y_{\text{R}} = a_{\text{R}} - e_{a,\text{R}} - v_{\text{R}}, \quad (\text{B.4})$$

$$y_{\text{I}} = a_{\text{I}} - e_{a,\text{I}} - v_{\text{I}}, \quad (\text{B.5})$$

$$\bar{d} = d - \frac{\epsilon}{\gamma}, \quad (\text{B.6})$$

$$d = d_{\text{R}} + jd_{\text{I}}, \quad (\text{B.7})$$

$$d_{\text{R}} = \frac{(\gamma + r - y_{\text{R}}^2) y_{\text{R}}}{\gamma} = \frac{(3E[a_{\text{R}}^2] - y_{\text{R}}^2) y_{\text{R}}}{\gamma}, \quad (\text{B.8})$$

$$d_{\text{I}} = \frac{(\gamma + r - y_{\text{I}}^2) y_{\text{I}}}{\gamma} = \frac{(3E[a_{\text{I}}^2] - y_{\text{I}}^2) y_{\text{I}}}{\gamma}. \quad (\text{B.9})$$

Usando as Equações (3.50) e (3.52), reescreve-se ϵ de forma fechada, também por conveniência, como

$$\epsilon = \frac{\rho}{2}(1 - \lambda) \sum_{\ell=0}^n \lambda^{n-\ell} y(\ell) \mathbf{y}^H(\ell) \mathbf{y}(n), \quad (\text{B.10})$$

em que $\mathbf{y}(n)$ é o vetor definido em (3.47). Nota-se que os índices de tempo e de fonte foram suprimidos quando possível. A seguir, os Resultados 2.2–A a 2.2–H são deduzidos na Seção B.2. Posteriormente, eles são usados na Seção B.3 para a obtenção dos Resultados 2.3–I e 2.3–J, correspondentes respectivamente às expressões para $E[e_{a,i}^*(n) \varepsilon_i(n)]$ e $E[|\varepsilon_i(n)|^2]$.

B.2 RESULTADOS PRELIMINARES

Resultado 2.2–A. *A partir de (B.4), e considerando-se as Hipóteses 4.2–A e 4.2–B, tem-se*

$$E[y_R e_{a,R}] = E[(a_R - e_{a,R} - v_R) e_{a,R}] = -r_2^{e_{a,R}}, \quad (\text{B.11})$$

$$E[y_R a_R] = r_2^{a_R}, \quad (\text{B.12})$$

$$E[y_R v_R] = -r_2^{v_R}, \quad (\text{B.13})$$

$$E[a^*(n) y(\ell)] = r_2^a \delta_{n,\ell}, \quad (\text{B.14})$$

$$E[e_a^*(n) y(\ell)] = -\zeta \delta_{n,\ell}, \quad (\text{B.15})$$

$$E[v^*(n) y(\ell)] = -r_2^v \delta_{n,\ell}, \quad (\text{B.16})$$

$$E[y_R(n) y^*(\ell)] = r_2^{y_R} \delta_{n,\ell}, \quad (\text{B.17})$$

$$E[y^*(\ell) y(m)] = r_2^y \delta_{\ell,m}. \quad (\text{B.18})$$

O momento de ordem 2 de y_R pode ser obtido diretamente a partir dos termos (B.11), (B.12) e (B.13), pois

$$r_2^{y_R} = E[y_R a_R] - E[y_R e_{a,R}] - E[y_R v_R].$$

Fazendo-se as substituições e considerando-se a Hipótese 4.2–I, o momento $r_2^{y_R}$ pode ser aproximado por

$$r_2^{y_R} \approx r_2^{a_R} + r_2^{v_R}. \quad (\text{B.19})$$

Resultado 2.2–B. *A partir de (B.4), resulta¹*

$$y_R^3 = a_R^3 - 3a_R^2 e_{a,R} - 3a_R^2 v_R + 3a_R e_{a,R}^2 + 6a_R e_{a,R} v_R + 3a_R v_R^2 - e_{a,R}^3 - 3e_{a,R}^2 v_R - 3e_{a,R} v_R^2 - v_R^3. \quad (\text{B.20})$$

Considerando-se (B.20) e as Hipóteses 4.2–A, 4.2–B e 4.2–H, resultam

$$E[y_R^3 e_{a,R}] = -3(r_2^{a_R} + r_2^{v_R}) r_2^{e_{a,R}}, \quad (\text{B.21})$$

$$E[y_R^3 a_R] = r_4^{a_R} + 3r_2^{a_R} r_2^{e_{a,R}} + 3r_2^{a_R} r_2^{v_R}, \quad (\text{B.22})$$

$$E[y_R^3 v_R] = -3r_2^{a_R} r_2^{v_R} - 3r_2^{e_{a,R}} r_2^{v_R} - r_4^{v_R}, \quad (\text{B.23})$$

$$E[y_R^3(n) y^*(\ell)] = r_4^{y_R} \delta_{n,\ell}. \quad (\text{B.24})$$

¹ Contas feitas com o auxílio do software Mathematica.

O momento de ordem 4 de y_R pode ser obtido diretamente a partir de (B.21), (B.22) e (B.23), pois

$$r_4^{y_R} = E \left[y_R^3 a_R \right] - E \left[y_R^3 e_{a,R} \right] - E \left[y_R^3 v_R \right].$$

Fazendo-se as substituições e considerando-se a Hipótese 4.2-I, o momento $r_4^{y_R}$ pode ser aproximado por

$$r_4^{y_R} \approx r_4^{a_R} + r_4^{v_R} + 6r_2^{a_R} r_2^{v_R} + 6r_2^{a_R} r_2^{e_{a,R}}. \quad (\text{B.25})$$

Resultado 2.2-C. A partir de (B.4), resulta²

$$\begin{aligned} y_R^6 = & a_R^6 - 6a_R^5 e_{a,R} - 6a_R^5 v_R + 15a_R^4 e_{a,R}^2 + 30a_R^4 e_{a,R} v_R + 15a_R^4 v_R^2 - 20a_R^3 e_{a,R}^3 - 60a_R^3 e_{a,R}^2 v_R \\ & - 60a_R^3 e_{a,R} v_R^2 - 20a_R^3 v_R^3 + 15a_R^2 e_{a,R}^4 + 60a_R^2 e_{a,R}^3 v_R + 90a_R^2 e_{a,R}^2 v_R^2 + 60a_R^2 e_{a,R} v_R^3 + 15a_R^2 v_R^4 \\ & - 6a_R e_{a,R}^5 - 30a_R e_{a,R}^4 v_R - 60a_R e_{a,R}^3 v_R^2 - 60a_R e_{a,R}^2 v_R^3 - 30a_R e_{a,R} v_R^4 - 6a_R v_R^5 + e_{a,R}^6 \\ & + 6e_{a,R}^5 v_R + 15e_{a,R}^4 v_R^2 + 20e_{a,R}^3 v_R^3 + 15e_{a,R}^2 v_R^4 + 6e_{a,R} v_R^5 + v_R^6. \end{aligned} \quad (\text{B.26})$$

Tomando-se a esperança de (B.26) e considerando-se as Hipóteses 4.2-A, 4.2-B, 4.2-H e 4.2-I, resulta

$$r_6^{y_R} \approx r_6^{a_R} + r_6^{v_R} + 15 (r_4^{a_R} r_2^{v_R} + r_2^{a_R} r_4^{v_R}) + 15 (r_4^{a_R} + 6r_2^{a_R} r_2^{v_R}) r_2^{e_{a,R}}. \quad (\text{B.27})$$

Resultado 2.2-D. A partir da representação do vetor $\mathbf{y}(n)$ como em (3.47), tem-se

$$\alpha_0 \triangleq E \left[\mathbf{y}^H(n) \mathbf{y}(n) \right] = \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{\tau=0}^{k_1} E \left[|y_m(n - \tau)|^2 \right] = (k_1 + 1) \sum_{\substack{m=1 \\ m \neq i}}^{N_u} r_2^{y_m}, \quad (\text{B.28})$$

sendo $r_2^{y_m} = E \left[|y_m(n - \tau)|^2 \right]$. Considerando-se as Hipóteses 4.2-A, 4.2-B e 4.2-I, resulta $r_2^{y_m} \approx (4 - \beta) \left(r_2^{a_R} + r_2^{v_{m,R}} \right)$, portanto,

$$\alpha_0 \approx (4 - \beta)(k_1 + 1) [(N_u - 1)r_2^{a_R} + S_{2,i}], \quad (\text{B.29})$$

em que

$$S_{2,i} \triangleq \sum_{\substack{m=1 \\ m \neq i}}^{N_u} r_2^{v_{m,R}}. \quad (\text{B.30})$$

Os Resultados 2.2-E, 2.2-F e 2.2-G, apresentados a seguir, foram incluídos para facilitar a obtenção do Resultado 2.2-H.

² Contas feitas com o auxílio do software Mathematica.

Resultado 2.2–E. A partir da Equação (B.10), pode-se escrever

$$\mathbb{E}[\epsilon a^*] = \frac{\varrho}{2}(1 - \lambda) \sum_{\ell=0}^n \lambda^{n-\ell} \mathbb{E}[a^*(n) y(\ell) \mathbf{y}^H(\ell) \mathbf{y}(n)], \quad (\text{B.31})$$

$$\mathbb{E}[\epsilon e_a^*] = \frac{\varrho}{2}(1 - \lambda) \sum_{\ell=0}^n \lambda^{n-\ell} \mathbb{E}[e_a^*(n) y(\ell) \mathbf{y}^H(\ell) \mathbf{y}(n)], \quad (\text{B.32})$$

$$\mathbb{E}[\epsilon v^*] = \frac{\varrho}{2}(1 - \lambda) \sum_{\ell=0}^n \lambda^{n-\ell} \mathbb{E}[v^*(n) y(\ell) \mathbf{y}^H(\ell) \mathbf{y}(n)], \quad (\text{B.33})$$

$$\mathbb{E}[y_R \epsilon^*] = \frac{\varrho}{2}(1 - \lambda) \sum_{\ell=0}^n \lambda^{n-\ell} \mathbb{E}[y_R(n) y^*(\ell) \mathbf{y}^T(\ell) \mathbf{y}^*(n)], \quad (\text{B.34})$$

$$\mathbb{E}[y_R^3 \epsilon^*] = \frac{\varrho}{2}(1 - \lambda) \sum_{\ell=0}^n \lambda^{n-\ell} \mathbb{E}[y_R^3(n) y^*(\ell) \mathbf{y}^T(\ell) \mathbf{y}^*(n)], \quad (\text{B.35})$$

Nota-se que o vetor $\mathbf{y}(n)$ não inclui a saída do i -ésimo equalizador. Logo, considerando-se a Hipótese 4.2–G, valem

$$\mathbb{E}[a^*(n) y(\ell) \mathbf{y}^H(\ell) \mathbf{y}(n)] = \mathbb{E}[a^*(n) y(\ell)] \mathbb{E}[\mathbf{y}^H(\ell) \mathbf{y}(n)], \quad (\text{B.36})$$

$$\mathbb{E}[e_a^*(n) y(\ell) \mathbf{y}^H(\ell) \mathbf{y}(n)] = \mathbb{E}[e_a^*(n) y(\ell)] \mathbb{E}[\mathbf{y}^H(\ell) \mathbf{y}(n)], \quad (\text{B.37})$$

$$\mathbb{E}[v^*(n) y(\ell) \mathbf{y}^H(\ell) \mathbf{y}(n)] = \mathbb{E}[v^*(n) y(\ell)] \mathbb{E}[\mathbf{y}^H(\ell) \mathbf{y}(n)], \quad (\text{B.38})$$

$$\mathbb{E}[y_R(n) y^*(\ell) \mathbf{y}^T(\ell) \mathbf{y}^*(n)] = \mathbb{E}[y_R(n) y^*(\ell)] \mathbb{E}[\mathbf{y}^T(\ell) \mathbf{y}^*(n)], \quad (\text{B.39})$$

$$\mathbb{E}[y_R^3(n) y^*(\ell) \mathbf{y}^T(\ell) \mathbf{y}^*(n)] = \mathbb{E}[y_R^3(n) y^*(\ell)] \mathbb{E}[\mathbf{y}^T(\ell) \mathbf{y}^*(n)]. \quad (\text{B.40})$$

Inicialmente, substituem-se os resultados (B.14), (B.15), (B.16), (B.17) e (B.24) respectivamente em (B.36), (B.37), (B.38), (B.39) e (B.40). Considerando-se que $r_2^a = (4 - \beta)r_2^{aR}$ e $r_2^v = (4 - \beta)r_2^{vR}$, conclui-se que (B.31), (B.32), (B.33), (B.34) e (B.35) podem ser reescritas, respectivamente, como

$$\mathbb{E}[\epsilon a^*] = \frac{\varrho}{2}(1 - \lambda)(4 - \beta)r_2^{aR}\alpha_0, \quad (\text{B.41})$$

$$\mathbb{E}[\epsilon e_a^*] = -\frac{\varrho}{2}(1 - \lambda)\alpha_0\zeta, \quad (\text{B.42})$$

$$\mathbb{E}[\epsilon v^*] = -\frac{\varrho}{2}(1 - \lambda)(4 - \beta)r_2^{vR}\alpha_0, \quad (\text{B.43})$$

$$\mathbb{E}[y_R \epsilon^*] = \frac{\varrho}{2}(1 - \lambda)r_2^{yR}\alpha_0, \quad (\text{B.44})$$

$$\mathbb{E}[y_R^3 \epsilon^*] = \frac{\varrho}{2}(1 - \lambda)r_4^{yR}\alpha_0, \quad (\text{B.45})$$

sendo α_0 conforme definido em (B.28) e aproximado em (B.29).

Resultado 2.2–F. Levando-se em conta a representação do vetor $\mathbf{y}(n)$ como em (3.47), a expressão $\mathbb{E}[|\mathbf{y}^H(\ell) \mathbf{y}(n)|^2]$ pode ser expressa como

$$\mathbb{E}[|\mathbf{y}^H(\ell) \mathbf{y}(n)|^2] = \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{\substack{m'=1 \\ m' \neq i}}^{N_u} \sum_{k=1}^{k_1} \sum_{k'=1}^{k_1} y_m^*(\ell - k) y_m(n - k) y_{m'}(n - k') y_{m'}^*(\ell - k').$$

Assim, considerando-se as Hipóteses 4.2-A, 4.2-B, 4.2-H, tem-se

$$\mathbb{E} \left[|\mathbf{y}^H(\ell) \mathbf{y}(n)|^2 \right] \approx \begin{cases} (k_1 + 1) \sum_{\substack{m=1 \\ m \neq i}}^{N_u} (r_2^{y_m})^2, & \text{se } \ell \neq n, \quad m = m', \quad e \quad k = k' \\ (k_1 + 1) \sum_{\substack{m=1 \\ m \neq i}}^{N_u} r_4^{y_m}, & \text{se } \ell = n, \quad m = m', \quad e \quad k = k' \\ k_1(k_1 + 1) \sum_{\substack{m=1 \\ m \neq i}}^{N_u} (r_2^{y_m})^2, & \text{se } \ell = n, \quad m = m', \quad e \quad k \neq k' \\ (k_1 + 1)^2 \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \sum_{\substack{m'=1 \\ m' \neq i \\ m' \neq m}}^{N_u} r_2^{y_{m'}} r_2^{y_m}, & \text{se } \ell = n \text{ e } m \neq m' \\ 0, & \text{caso contrário.} \end{cases} \quad (\text{B.46})$$

Resultado 2.2-G. *Seja*

$$\phi_m \triangleq \sum_{\ell=0}^n \lambda^{2(n-\ell)} \mathbb{E} \left[|\mathbf{y}^H(\ell) \mathbf{y}(n)|^2 \right] = \sum_{\ell=0}^{n-1} \lambda^{2(n-\ell)} \mathbb{E} \left[|\mathbf{y}^H(\ell) \mathbf{y}(n)|^2 \right] + \mathbb{E} \left[|\mathbf{y}^H(n) \mathbf{y}(n)|^2 \right]. \quad (\text{B.47})$$

Usando-se o resultado de (B.46) e considerando-se $n \rightarrow +\infty$, resulta

$$\phi_m \approx (k_1 + 1) \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \left[\left(\frac{\lambda^2}{1 - \lambda^2} + k_1 \right) (r_2^{y_m})^2 + r_4^{y_m} + (k_1 + 1) \sum_{\substack{m'=1 \\ m' \neq i \\ m' \neq m}}^{N_u} r_2^{y_{m'}} r_2^{y_m} \right]. \quad (\text{B.48})$$

Considerando-se que, com $0 \ll \lambda < 1$, valem

$$\left(\frac{\lambda^2}{1 - \lambda^2} + k_1 \right) (r_2^{y_m})^2 \gg r_4^{y_m} + (k_1 + 1) \sum_{\substack{m'=1 \\ m' \neq i \\ m' \neq m}}^{N_u} r_2^{y_{m'}} r_2^{y_m}, \quad (\text{B.49})$$

e

$$\frac{\lambda^2}{1 - \lambda^2} + k_1 \approx \frac{\lambda^2}{1 - \lambda^2}, \quad (\text{B.50})$$

a Equação (B.48) pode ser aproximada por

$$\phi_m \approx (k_1 + 1) \frac{\lambda^2}{1 - \lambda^2} \sum_{\substack{m=1 \\ m \neq i}}^{N_u} (r_2^{y_m})^2. \quad (\text{B.51})$$

Nota-se que, com as Hipóteses 4.2-A e 4.2-B, o somando que aparece na Equação (B.51) pode ser escrito como

$$(r_2^{y_m})^2 = (4 - \beta)^2 (r_2^{y_{m,R}})^2 \approx (4 - \beta)^2 \left[(r_2^{a_R})^2 + 2r_2^{a_R} r_2^{v_{m,R}} + (r_2^{v_{m,R}})^2 \right], \quad (\text{B.52})$$

Substituindo-se (B.52) em (B.51) e utilizando-se novamente a Hipótese 4.2-B, a expressão de (B.51) pode ser novamente aproximada como

$$\phi_m \approx (4 - \beta)^2 (k_1 + 1) \frac{\lambda^2}{1 - \lambda^2} \left[(N_u - 1) \text{E}^2[a_{i,R}^2(n)] + 2 \text{E}[a_{i,R}^2(n)] S_{2,i} + \Lambda_{4,i} \right], \quad (\text{B.53})$$

sendo $S_{2,i}$ como definido em (B.30) e

$$\Lambda_{4,i} \triangleq \sum_{\substack{m=1 \\ m \neq i}}^{N_u} \left(r_2^{v_{m,R}} \right)^2.$$

Resultado 2.2-H. Considerando-se inicialmente (B.10), pode-se escrever

$$|\epsilon|^2 = \frac{\varrho^2}{4} (1 - \lambda)^2 \sum_{m=0}^n \sum_{\ell=0}^n \lambda^{2n-\ell-m} y^*(\ell) y(m) \mathbf{y}^H(n) \mathbf{y}(\ell) \mathbf{y}^H(m) \mathbf{y}(n). \quad (\text{B.54})$$

Tomando-se a esperança de (B.54), resulta

$$r_2^\epsilon = \text{E} \left[|\epsilon|^2 \right] = \frac{\varrho^2}{4} (1 - \lambda)^2 \sum_{m=0}^n \sum_{\ell=0}^n \lambda^{2n-\ell-m} \text{E} \left[y^*(\ell) y(m) \mathbf{y}^H(n) \mathbf{y}(\ell) \mathbf{y}^H(m) \mathbf{y}(n) \right]. \quad (\text{B.55})$$

Observa-se que o vetor $\mathbf{y}(n)$ não inclui a saída do i -ésimo equalizador. Logo, considerando-se a Hipótese 4.2-G, vale

$$\text{E} \left[y^*(\ell) y(m) \mathbf{y}^H(n) \mathbf{y}(\ell) \mathbf{y}^H(m) \mathbf{y}(n) \right] = \text{E} \left[y^*(\ell) y(m) \right] \text{E} \left[\mathbf{y}^H(n) \mathbf{y}(\ell) \mathbf{y}^H(m) \mathbf{y}(n) \right].$$

Portanto, considerando-se o Resultado 2.2-G e (B.18), pode-se escrever

$$r_2^\epsilon = \frac{\varrho^2}{4} (1 - \lambda)^2 (4 - \beta) r_2^{y_R} \phi_m, \quad (\text{B.56})$$

sendo ϕ_m conforme definido em (B.47) e aproximado em (B.53). Essa aproximação para ϕ_m é mais completa que a aproximação análoga apresentada em (LUO; CHAMBERS, 2002), devido à inclusão dos termos das correlações cruzadas quando $\ell = n$. Na Seção 4.2.4, valores de médias de conjunto de $|\epsilon|^2$ são comparados com os modelos de r_2^ϵ dados por (B.56) e (LUO; CHAMBERS, 2002), evidenciando a melhoria da estimativa ao se fazer uma aproximação mais completa para ϕ_m .

B.3 MOMENTOS DA ANÁLISE EM REGIME

Resultado 2.3-I. A fim de se reescrever $\text{E} [e_a^* \epsilon]$ conforme desejado, considerando-se a Equação (B.1) e as Hipóteses 4.2-A e 4.2-B, resulta

$$\text{E} [e_a^* \epsilon] = \text{E} \left[e_a^* (\bar{d} - a + v) \right] = \text{E} \left[e_a^* \bar{d} \right]. \quad (\text{B.57})$$

Substituindo-se \bar{d} como em (B.6), considerando-se a independência entre as partes real e imaginária de e_a , notando-se que, por simetria, vale a igualdade $\text{E} [d_{i,e_a}] = (3 - \beta) \text{E} [d_{R,e_a,R}]$ e usando d_R como em (B.8), resulta

$$\text{E} [e_a^* \epsilon] = \frac{1}{\gamma} \left[(4 - \beta) \left(3r_2^{a_R} \text{E} [y_R e_{a,R}] - \text{E} [y_R^3 e_{a,R}] \right) - \text{E} [\epsilon e_a^*] \right]. \quad (\text{B.58})$$

Substituindo-se na Equação (B.58) os momentos (B.11), (B.21) e (B.42), e notando-se que $(4 - \beta)r_2^{\epsilon_{a,R}} = \zeta$, após manipulações algébricas, resulta

$$\mathbb{E}[e_a^* \varepsilon] \approx \left(3r_2^{v_R} + \frac{\varrho}{2}(1 - \lambda)\alpha_0 \right) \frac{1}{\gamma},$$

sendo α_0 conforme definido em (B.28) e aproximado em (B.29). Denotando-se

$$\alpha_v = 3r_2^{v_R}, \quad (\text{B.59})$$

$$\alpha_m = \varrho(1 - \lambda)\alpha_0, \quad (\text{B.60})$$

$$\alpha = \alpha_v + \frac{\alpha_m}{2}, \quad (\text{B.61})$$

resulta

$$\mathbb{E}[e_a^* \varepsilon] \approx \frac{\alpha \zeta}{\gamma}. \quad (\text{B.62})$$

Resultado 2.3–J. A fim de se reescrever o momento $\mathbb{E}[|\varepsilon|^2]$ conforme desejado, inicialmente considera-se (B.1) e as Hipóteses 4.2–A e 4.2–B, assim

$$r_2^\varepsilon = \mathbb{E}[|\varepsilon|^2] = r_2^a - 2 \operatorname{Re} \left\{ \mathbb{E}[\bar{d}a^*] \right\} + r_2^{\bar{d}} + 2 \operatorname{Re} \left\{ \mathbb{E}[\bar{d}v^*] \right\} + r_2^v. \quad (\text{B.63})$$

Sabendo-se que $r_2^a = (4 - \beta)r_2^{a_R}$ e $r_2^v = (4 - \beta)r_2^{v_R}$, pode-se escrever

$$r_2^\varepsilon = (4 - \beta)r_2^{a_R} - 2 \operatorname{Re} \left\{ \mathbb{E}[\bar{d}a^*] \right\} + r_2^{\bar{d}} + 2 \operatorname{Re} \left\{ \mathbb{E}[\bar{d}v^*] \right\} + (4 - \beta)r_2^{v_R}. \quad (\text{B.64})$$

A seguir, são calculados os momentos $\mathbb{E}[\bar{d}a^*]$, $\mathbb{E}[\bar{d}v^*]$ e $r_2^{\bar{d}}$.

► A partir da Equação (B.6) e das Hipóteses 4.2–A e 4.2–B, chega-se à expressão

$$\mathbb{E}[\bar{d}a^*] = \mathbb{E} \left[\left(d_R + jd_1 - \frac{\epsilon}{\gamma} \right) a^* \right] = \mathbb{E}[d_R a_R] + \mathbb{E}[d_1 a_1] - \frac{1}{\gamma} \mathbb{E}[\epsilon a^*]. \quad (\text{B.65})$$

Notando-se que $\mathbb{E}[d_1 a_1] = (3 - \beta) \mathbb{E}[d_R a_R]$, usando-se d_R como em (B.8) e substituindo-se os momentos (B.12), (B.22) e (B.41) na Equação (B.65), após manipulações algébricas, resulta

$$\mathbb{E}[\bar{d}a^*] = \frac{4 - \beta}{\gamma} \left[3(r_2^{a_R})^2 - r_4^{a_R} - 3r_2^{a_R} r_2^{\epsilon_{a,R}} - 3r_2^{a_R} r_2^{v_R} - \frac{\varrho}{2}(1 - \lambda)r_2^{a_R} \alpha_0 \right]. \quad (\text{B.66})$$

sendo α_0 conforme definido em (B.29) e aproximado em (B.29).

► Multiplicando-se (B.6) por v^* , tomando-se a esperança da expressão resultante, notando-se que $\mathbb{E}[d_1 v_1] = (3 - \beta) \mathbb{E}[d_R v_R]$ e substituindo-se d_R como em (B.8), resulta

$$\mathbb{E}[\bar{d}v^*] = \frac{1}{\gamma} \left[(4 - \beta) \left(3r_2^{a_R} \mathbb{E}[y_R v_R] - \mathbb{E}[y_R^3 v_R] \right) - \mathbb{E}[\epsilon v^*] \right]. \quad (\text{B.67})$$

Substituindo-se os momentos (B.13), (B.23) e (B.43) na Equação (B.67) e fazendo a simplificação da expressão resultante, tem-se

$$\mathbb{E}[\bar{d}v^*] = \frac{4 - \beta}{\gamma} \left(3r_2^{\epsilon_{a,R}} r_2^{v_R} + r_4^{v_R} + \frac{\varrho}{2}(1 - \lambda)r_2^{v_R} \alpha_0 \right). \quad (\text{B.68})$$

► A partir da Equação (B.6), chega-se à expressão

$$|\bar{d}|^2 = d_R^2 + d_1^2 - \frac{2}{\gamma} \operatorname{Re} \{ (d_R + jd_1) \epsilon^* \} + \frac{|\epsilon|^2}{\gamma^2}. \quad (\text{B.69})$$

Aplicando-se a esperança matemática em (B.69) e notando-se que $r_2^{d_1} = (3 - \beta)r_2^{d_R}$ e $E[jd_1\epsilon^*] = (3 - \beta)E[d_R\epsilon^*]$, resulta

$$r_2^{\bar{d}} = (4 - \beta)r_2^{d_R} - \frac{2}{\gamma}(4 - \beta) \operatorname{Re} \{ E[d_R\epsilon^*] \} + \frac{r_2^\epsilon}{\gamma^2}. \quad (\text{B.70})$$

A seguir, são calculados os momentos $r_2^{d_R}$ e $E[d_R\epsilon^*]$.

◊ A partir da Equação (B.8), escreve-se

$$d_R^2 = \frac{1}{\gamma^2} [y_R^6 - 6r_2^{a_R}y_R^4 + 9(r_2^{a_R})^2y_R^2]. \quad (\text{B.71})$$

Tomando-se a esperança de (B.71), tem-se

$$r_2^{d_R} = \frac{1}{\gamma^2} [r_6^{y_R} - 6r_2^{a_R}r_4^{y_R} + 9(r_2^{a_R})^2r_2^{y_R}]. \quad (\text{B.72})$$

◊ Utilizando-se (B.8) e (B.10), pode-se escrever

$$E[d_R\epsilon^*] = \frac{1}{\gamma} (3r_2^{a_R} E[y_R\epsilon^*] - E[y_R^3\epsilon^*]). \quad (\text{B.73})$$

Agora, lançando-se mão de (B.44) e (B.45), resulta

$$E[d_R\epsilon^*] = -\frac{\varrho}{2\gamma}(1 - \lambda)(r_4^{y_R} - 3r_2^{a_R}r_2^{y_R})\alpha_0. \quad (\text{B.74})$$

Substituindo-se os momentos (B.72), (B.74) e (B.56) na Equação (B.70) e fazendo a simplificação da expressão resultante, tem-se

$$r_2^{\bar{d}} = \frac{4 - \beta}{\gamma^2} \left[r_6^{y_R} - 6r_2^{a_R}r_4^{y_R} + 9(r_2^{a_R})^2r_2^{y_R} + \varrho(1 - \lambda)(r_4^{y_R} - 3r_2^{a_R}r_2^{y_R})\alpha_0 + \frac{\varrho^2}{4}(1 - \lambda)r_2^{y_R}\phi_m \right]. \quad (\text{B.75})$$

Finalmente, substituem-se na Equação (B.64) os momentos $E[\bar{d}a^*]$, $E[\bar{d}v^*]$ e $r_2^{\bar{d}}$ pelas respectivas expressões (B.66), (B.75) e (B.68). Considerando-se os momentos da parte real de y (Resultados 2.2-D, 2.2-H e 2.3-I) e lembrando-se que $(4 - \beta)r_2^{\epsilon_{a,R}} = \zeta$, chega-se à seguinte expressão aproximada³,

$$r_2^\epsilon = E[|\epsilon(n)|^2] \approx \frac{(4 - \beta)\psi + \chi\zeta}{\gamma^2}, \quad (\text{B.76})$$

³ As manipulações algébricas foram feitas com o auxílio do software Mathematica versão 9.0.1.0.

em que

$$\psi = \psi_a + \psi_v + \psi_m, \quad (\text{B.77})$$

$$\psi_a = r_6^{aR} - \mathcal{R}r_4^{aR}, \quad (\text{B.78})$$

$$\psi_v = r_2^{vR} \left[\mathcal{R}^2 + 3r_4^{aR} + 45r_2^{aR}r_2^{vR} - 6\mathcal{R}r_2^{vR} + 15(r_2^{vR})^2 \right], \quad (\text{B.79})$$

$$\psi_m = (6r_2^{aR} - \mathcal{R} + 3r_2^{vR})r_2^{vR}\alpha_m + \frac{\varrho^2}{4}(1 - \lambda)^2 (r_2^{aR} + r_2^{vR})\phi_m, \quad (\text{B.80})$$

$$\chi = \chi_a + \chi_v + \chi_m, \quad (\text{B.81})$$

$$\chi_a = 9 \left[r_4^{aR} - 2(r_2^{aR})^2 \right], \quad (\text{B.82})$$

$$\chi_v = 6r_2^{vR} (18r_2^{aR} - \mathcal{R}), \quad (\text{B.83})$$

$$\chi_m = 6r_2^{aR}\alpha_m, \quad (\text{B.84})$$

sendo α_m dado por (B.60) e ϕ_m conforme definido em B.47 e aproximado em (B.53).