

Hugo Vinícius Leão e Silva

**Utilização de Filtragem Espacial e Otimização  
Numérica em um Método de Estimação de  
Direção de Chegada em Arranjos de Sensores**

Uberlândia

2017

Hugo Vinícius Leão e Silva

# **Utilização de Filtragem Espacial e Otimização Numérica em um Método de Estimação de Direção de Chegada em Arranjos de Sensores**

Tese apresentada ao Programa de Pós-Graduação *Strictu Sensu* em Engenharia Elétrica da Faculdade de Engenharia Elétrica da Universidade Federal de Uberlândia como recurso parcial para obtenção do título de Doutor em Ciências.

**Área de concentração:** Processamento da Informação

**Linha de Pesquisa:** Processamento Digital de Sinais

## **BANCA DE AVALIAÇÃO FINAL:**

**Edna Lúcia Flôres, Dr.<sup>a</sup> (UFU), Orientadora**

**Milena Bueno Pereira Carneiro, Dr.<sup>a</sup> (UFU)**

**Karine Barbosa Carbonaro, Dr.<sup>a</sup> (UFU)**

**Rodrigo Pinto Lemos, Dr. (UFG)**

**Leonardo da Cunha Brito, Dr. (UFG)**

**Cláudio Afonso Fleury, Dr. (IFG)**

Uberlândia

2017

Dados Internacionais de Catalogação na Publicação (CIP)  
Sistema de Bibliotecas da UFU, MG, Brasil.

---

L437u      Leão e Silva, Hugo Vinícius, 1984-  
2017      Utilização de filtragem espacial e otimização numérica em um  
método de estimação de direção de chegada em arranjos de sensores /  
Hugo Vinícius Leão e Silva. - 2017.  
189 f. : il.

Orientadora: Edna Lúcia Flôres.  
Tese (doutorado) - Universidade Federal de Uberlândia, Programa  
de Pós-Graduação em Engenharia Elétrica.  
Inclui bibliografia.

1. Engenharia elétrica - Teses. 2. Processamento de sinais - Teses. I.  
Flôres, Edna Lúcia, 1958- II. Universidade Federal de Uberlândia.  
Programa de Pós-Graduação em Engenharia Elétrica. III. Título.

CDU: 621.3

---

Hugo Vinícius Leão e Silva

# **Utilização de Filtragem Espacial e Otimização Numérica em um Método de Estimação de Direção de Chegada em Arranjos de Sensores**

Tese apresentada ao Programa de Pós-Graduação *Strictu Sensu* em Engenharia Elétrica da Faculdade de Engenharia Elétrica da Universidade Federal de Uberlândia como recurso parcial para obtenção do título de Doutor em Ciências.

Trabalho aprovado. Uberlândia, 30 de junho de 2017:

---

**Edna Lúcia Flôres**  
Orientadora

---

**Milena Bueno Pereira Carneiro**

---

**Karine Barbosa Carbonaro**

---

**Rodrigo Pinto Lemos**  
Convidado 1

---

**Leonardo da Cunha Brito**  
Convidado 2

---

**Cláudio Afonso Fleury**  
Convidado 3

Uberlândia  
2017



# AGRADECIMENTOS

Antes de mais nada, os meus sinceros agradecimentos vão à minha família. Meus pais – *Marcos* e *Lina* –, meu irmão *Júnior* e a minha amada noiva *Ludimila*, que tiveram que lidar em muitos aspectos com a minha ausência incontáveis vezes nesses mais de três anos. Eu tive que me abdicar de estar juntos deles mais vezes que eu desejei. Nessas horas eu vejo como a família provê um suporte imenso para que seja possível dar grandes saltos e preciso agradecê-los por isso. O salto que representa este doutoramento definitivamente exigiu de mim um isolamento por muitos períodos para superar os desafios que me foram imputados.

Entretanto, esse processo não seria possível de ser vencido sem o importantíssimo auxílio de meus orientadores e, porque não, mestres. Assim como é para a minha família, faço saber à Professora *Edna* e ao Professor *Rodrigo*: não há primeiro e segundo. Os dois dedicaram muito mais do que o tempo necessário para me guiar na pesquisa e na escrita de artigos e desta tese. Aos dois, o meu muito obrigado. Foram necessárias muitas reuniões em diversos horários, até mesmo aos fins-de-semana, para chegar neste resultado.

Por outro lado, isso me amadureceu bastante. Eu tenho ainda muita coisa para aprender, é verdade. No entanto, hoje eu me vejo mais confiante para desafiar outros problemas e não posso dizer outra coisa senão com certeza valeu a pena.

# RESUMO

O estimador SEAD (do inglês *Search of Direction by Differential Spectrum*) Aprimorado possui um robusto desempenho de estimação da direção de chegada (DOA) de fontes de sinal próximas entre si. Porém, ele apresenta algumas dificuldades na estimação DOA de fontes afastadas e em altas SNRs. Além disso, ele apresenta esforço computacional de ordem exponencial. Isso é devido aos algoritmos utilizados nos procedimentos de geração e refinamento de estimativas. Assim este trabalho propõe alterações neste estimador. A primeira delas é remover o limiar de amplitude no estágio de pré-seleção que classifica picos de sinal e picos de ruído. Dessa forma, é possível diminuir a probabilidade de desconsiderar picos de sinal na geração de estimativas iniciais. A segunda proposta é substituir o algoritmo discreto baseado no *Branch-and-Bound* no estágio de refinamento de estimativas por um método numérico de otimização local. Esse algoritmo discreto possui elevado esforço computacional e produz um limite inferior no desempenho de estimação maior que o Limite de Cramér-Rao (CRB). Assim, o SEAD Aprimorado não é um estimador assintoticamente eficiente. Nesse sentido, foram desenvolvidos o Método de Newton e um método quase-Newton que permitem reduzir de maneira significativa esse esforço computacional, além de produzir estimativas em um intervalo contínuo, removendo efetivamente o limite inferior no desempenho de estimação em altas SNRs. Finalmente, a técnica de máxima autofiltragem espacial permitiu ganhos significativos no desempenho de estimação DOA de fontes próximas entre si. Entretanto, esse autofiltro pode atenuar de maneira considerável as fontes afastadas quando a SNR diminui. Então, a terceira proposta é aplicar a filtragem espacial ao SEAD Aprimorado. Adicionalmente, este trabalho propõe dois novos filtros que superam as limitações do máximo autofiltro espacial na estimação DOA de fontes afastadas, enquanto apresentam um bom desempenho na situação de fontes próximas. O resultado é o SEAD Modificado, um estimador que apresenta um desempenho competitivo na estimação de fontes próximas e afastadas e possui um esforço computacional consideravelmente menor do que o SEAD Aprimorado e outro estimador proposto na literatura.

**Palavras-chave:** Processamento de sinais em arranjo, estimação da direção de chegada, estimação de máxima-verossimilhança, cancelamento de ruído, filtros espaciais, otimizadores de funções convexas, método de Newton.

# ABSTRACT

The Improved SEAD estimator has a robust direction-of-arrival (DOA) estimation performance of closely-spaced signal sources. However, it presents some difficulties on the DOA estimation of widely-spaced sources and high SNRs. Furthermore, it shows a computational effort of exponential order. This is due to the algorithms used on estimate generation and refinement procedures. So, this work proposes changes to that estimator. The first one is to remove the amplitude threshold in the pre-selection stage that divides signal peaks from noise peaks. In that way, one can lower the probability of disregarding signal peaks on generation of initial estimates. The second proposal is to replace the discrete algorithm based on Branch-and-Bound in the refinement stage with a local optimization numerical method. That discrete algorithm has a high computational effort and produces a lower bound on estimation performance higher than the Cramér-Rao Bound (CRB). Thus, the Improved SEAD is not an asymptotically efficient estimator. In that sense, the Newton's Method and a quasi-Newton method were developed, allowing for a significant reduction of that computational effort, besides producing DOA estimates in a continuous interval, effectively removing the lower bound on estimation performance at high SNRs. Finally, the technique of maximum spatial eigenfiltering recently presented allows for significant gains on the DOA estimation performance of closely-spaced sources. On the other hand, this eigenfilter may considerably attenuate widely-spaced sources when the SNR decreases. Then, the third proposal is applying spatial filtering to the Improved SEAD. Additionally, this work proposes two new filters that overcome the maximum spatial eigenfilter's limitation on the DOA estimation of widely-spaced sources and present a good estimation performance considering closely-spaced sources. The result is the Modified SEAD, an estimator that has a competitive DOA estimation performance for both closely and widely-spaced sources and has a considerably lower computational effort to that of the Improved SEAD and another DOA estimator proposed in the literature.

**Keywords:** Array signal processing, direction-of-arrival estimation, maximum likelihood estimation, noise cancelling, spatial filters, convex function optimizers, Newton's method.

# LISTA DE ILUSTRAÇÕES

Figura 1 – Representação bidimensional de um arranjo de sensores amostrando sinais de uma fonte. . . . .	26
Figura 2 – Superfície do DMLE para fontes de sinal situadas em $10^\circ$ e $15^\circ$ e sinais descorrelacionados na ausência de ruído. . . . .	36
Figura 3 – Superfície do SMLE para fontes de sinal situadas em $10^\circ$ e $15^\circ$ e sinais descorrelacionados na ausência de ruído. . . . .	36
Figura 4 – Superfície do DMLE para fontes de sinal situadas em $10^\circ$ e $15^\circ$ e sinais descorrelacionados na SNR de $-10$ dB. . . . .	37
Figura 5 – Superfície do SMLE para fontes de sinal situadas em $10^\circ$ e $15^\circ$ e sinais descorrelacionados na SNR de $-10$ dB. . . . .	37
Figura 6 – Espectro de autovalores de $\hat{\mathbf{R}}$ na ausência de ruído. . . . .	47
Figura 7 – Espectros diferenciais sobrepostos considerando fontes situadas em $10^\circ$ e $15^\circ$ na ausência de ruído (linhas pretas) e na SNR de $-10$ dB (linhas cinzas). . . . .	49
Figura 8 – Espectros diferenciais sobrepostos considerando fontes situadas em $10^\circ$ e $45^\circ$ na ausência de ruído (linhas pretas) e na SNR de $-10$ dB (linhas cinzas). . . . .	49
Figura 9 – Espectros diferenciais sobrepostos considerando fontes situadas em $10^\circ$ , $15^\circ$ e $45^\circ$ na ausência de ruído (linhas pretas) e na SNR de $-10$ dB (linhas cinzas). . . . .	50
Figura 10 – Espectro diferencial considerando fontes situadas em $10^\circ$ , $15^\circ$ e $45^\circ$ na ausência de ruído e o respectivo limiar de amplitude. . . . .	54
Figura 11 – Estimativas geradas e selecionadas durante o processo de refinamento de estimativas no caso de fontes situadas em $10^\circ$ , $15^\circ$ e $45^\circ$ na ausência de ruído. . . . .	55
Figura 12 – Respostas em frequência sobrepostas do máximo autofiltro espacial considerando fontes situadas em $10^\circ$ e $15^\circ$ na ausência de ruído (linhas pretas) e na SNR de $-10$ dB (linhas cinzas). . . . .	63
Figura 13 – Respostas em frequência sobrepostas do máximo autofiltro espacial considerando fontes situadas em $10^\circ$ e $45^\circ$ na ausência de ruído (linhas pretas) e na SNR de $-10$ dB (linhas cinzas). . . . .	64
Figura 14 – Respostas em frequência sobrepostas do máximo autofiltro espacial considerando fontes situadas em $10^\circ$ , $15^\circ$ e $45^\circ$ na ausência de ruído (linhas pretas) e na SNR de $-10$ dB (linhas cinzas). . . . .	64

Figura 15 – Superfície de um experimento do DMLE considerando fontes situadas em 10° e 45° e sinais descorrelacionados na ausência de ruído. . . . .	65
Figura 16 – Superfície de um experimento do DMLE usando o máximo autofiltro espacial para fontes de sinal situadas em 10° e 45° e sinais descorrelacionados na ausência de ruído. . . . .	65
Figura 17 – Superfície de outro experimento do DMLE usando o máximo autofiltro espacial para fontes de sinal situadas em 10° e 45° e sinais descorrelacionados na ausência de ruído. . . . .	66
Figura 18 – Espectro de autovalores de $\hat{\mathbf{R}}_I$ para fontes descorrelacionadas situadas em 10° e 15° na ausência de ruído. . . . .	67
Figura 19 – Espectro de autovalores de $\hat{\mathbf{R}}_I$ para fontes descorrelacionadas situadas em 10° e 45° na ausência de ruído. . . . .	67
Figura 20 – Espectro de autovalores de $\hat{\mathbf{R}}_I$ para fontes correlacionadas situadas em 10° e 45° na ausência de ruído. . . . .	68
Figura 21 – Respostas em frequência sobrepostas dos filtros considerando fontes situadas em 10° e 15° na ausência de ruído (linhas pretas) e na SNR de –10 dB (linhas cinzas). . . . .	71
Figura 22 – Respostas em frequência sobrepostas dos filtros considerando fontes situadas em 10° e 45° na ausência de ruído (linhas pretas) e na SNR de –10 dB (linhas cinzas). . . . .	72
Figura 23 – Respostas em frequência sobrepostas dos filtros considerando fontes situadas em 10°, 15° e 45° na ausência de ruído (linhas pretas) e na SNR de –10 dB (linhas cinzas). . . . .	73
Figura 24 – Superfície de um experimento do DMLE usando o filtro calculado a partir do subespaço de sinal considerando fontes situadas em 10° e 45° e sinais descorrelacionados na ausência de ruído. . . . .	74
Figura 25 – Superfície de um experimento do DMLE usando o filtro calculado a partir do espectro diferencial considerando fontes situadas em 10° e 45° e sinais descorrelacionados na ausência de ruído. . . . .	74
Figura 26 – Probabilidade de troca de subespaço para cada autovetor de sinal de $\hat{\mathbf{R}}_I$ considerando fontes situadas em 10° e 15°. . . . .	75
Figura 27 – Probabilidade de troca de subespaço para cada autovetor de sinal de $\hat{\mathbf{R}}_I$ considerando fontes situadas em 10° e 45°. . . . .	76
Figura 28 – Probabilidade de troca de subespaço para cada autovetor de sinal de $\hat{\mathbf{R}}_I$ considerando fontes situadas em 10°, 15° e 45°. . . . .	77
Figura 29 – Superfície de desempenho do algoritmo original do MODEX Modificado para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	92

Figura 30 – Superfície de desempenho do MODEX Modificado utilizando o máximo autofiltro espacial para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	93
Figura 31 – Superfície de desempenho do MODEX Modificado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . . .	93
Figura 32 – Superfície de desempenho do MODEX Modificado utilizando o filtro calculado a partir do espectro diferencial para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . .	94
Figura 33 – Superfície de desempenho do algoritmo original do MODEX Modificado para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	94
Figura 34 – Superfície de desempenho do MODEX Modificado utilizando o máximo autofiltro espacial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	95
Figura 35 – Superfície de desempenho do MODEX Modificado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	95
Figura 36 – Superfície de desempenho do MODEX Modificado utilizando o filtro calculado a partir do espectro diferencial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	96
Figura 37 – Superfície de desempenho do algoritmo original do SEAD Aprimorado para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	97
Figura 38 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o máximo autofiltro espacial para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . . .	97
Figura 39 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	98
Figura 40 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o filtro calculado a partir do espectro diferencial para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	98
Figura 41 – Superfície de desempenho do algoritmo original do SEAD Aprimorado para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	99

Figura 42 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o máximo autofiltro espacial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	99
Figura 43 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	100
Figura 44 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o filtro calculado a partir do espectro diferencial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	100
Figura 45 – Superfície de desempenho do SEAD Modificado para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	101
Figura 46 – Superfície de desempenho do SEAD Modificado utilizando o máximo autofiltro espacial para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	102
Figura 47 – Superfície de desempenho do SEAD Modificado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	103
Figura 48 – Superfície de desempenho do SEAD Modificado utilizando o filtro calculado a partir do espectro diferencial para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	103
Figura 49 – Superfície de desempenho do SEAD Modificado para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	104
Figura 50 – Superfície de desempenho do SEAD Modificado utilizando o máximo autofiltro espacial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	104
Figura 51 – Superfície de desempenho do SEAD Modificado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	105
Figura 52 – Superfície de desempenho do SEAD Modificado utilizando o filtro calculado a partir do espectro diferencial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente). . . . .	105
Figura 53 – Superfície de um experimento do DMLE usando o filtro espacial calculado a partir do subespaço de sinal considerando fontes de sinal situadas em $10^\circ$ e $15^\circ$ e sinais descorrelacionados na ausência de ruído. . . . .	106
Figura 54 – Superfície de um experimento do DMLE usando o filtro espacial calculado a partir do espectro diferencial considerando fontes de sinal situadas em $10^\circ$ e $15^\circ$ e sinais descorrelacionados na ausência de ruído. . . . .	106

# LISTA DE TABELAS

Tabela 1 – Combinações geradas no estágio de pré-seleção para o caso de fontes situadas em 10°, 15° e 45° na ausência de ruído. . . . .	53
Tabela 2 – Combinações geradas no novo estágio de pré-seleção para o caso de fontes situadas em 10°, 15° e 45° na ausência de ruído e avaliadas de acordo com o DMLE sem a aplicação de filtragem espacial. . . . .	88
Tabela 3 – Quantidade de combinações avaliadas no novo estágio de pré-seleção variando em função do número de fontes. . . . .	88
Tabela 4 – Tempos totais para 1.000 simulações do MODEX Modificado em função do número de fontes. . . . .	108
Tabela 5 – Tempos totais para 1.000 simulações do SEAD Aprimorado em função do número de fontes. . . . .	108
Tabela 6 – Tempos totais para 1.000 simulações do SEAD Modificado em função do número de fontes. . . . .	109
Tabela 7 – Percentil 97 dos tempos de execução (em segundos) dos estimadores avaliados em função do número de fontes e em função do filtro espacial utilizado. . . . .	109
Tabela 8 – Aumento nos tempos de execução relativos ao MODEX Modificado sem filtragem espacial em função do número de fontes. . . . .	110
Tabela 9 – Aumento nos tempos de execução relativos ao SEAD Aprimorado sem filtragem espacial em função do número de fontes. . . . .	110
Tabela 10 – Aumento nos tempos de execução relativos ao SEAD Modificado sem filtragem espacial em função do número de fontes. . . . .	110
Tabela 11 – Raiz do Erro Quadrático Médio em graus dos estimadores avaliados e o Limite de Cramér-Rao correspondente considerando duas fontes descorrelacionadas situadas em 10° e 15°. . . . .	187
Tabela 12 – Raiz do Erro Quadrático Médio em graus dos estimadores avaliados e o Limite de Cramér-Rao correspondente considerando duas fontes descorrelacionadas situadas em 10° e 45°. . . . .	188
Tabela 13 – Raiz do Erro Quadrático Médio em graus dos estimadores avaliados e o Limite de Cramér-Rao correspondente considerando duas fontes correlacionadas situadas em 10° e 15°. . . . .	188
Tabela 14 – Raiz do Erro Quadrático Médio em graus dos estimadores avaliados e o Limite de Cramér-Rao correspondente considerando duas fontes correlacionadas situadas em 10° e 45°. . . . .	189



# LISTA DE ABREVIATURAS E SIGLAS

CML	<i>Conditional Maximum Likelihood</i> – Máxima Verossimilhança Condicional
CRB	<i>Cramér-Rao Bound</i> – Limite de Cramér-Rao
CRU	Circunferência de Raio-Unitário
DMLE	<i>Deterministic Maximum Likelihood Estimator</i> – Critério Determinístico do Estimador de Máxima-Verossimilhança
DOA	<i>Direction-of-Arrival</i> – Direção de Chegada
EQM	Erro Quadrático Médio
FIR	<i>Finite Impulse Response</i> – Resposta ao Impulso Finita
ML	<i>Maximum Likelihood</i> – Máxima-Verossimilhança
MLE	<i>Maximum Likelihood Estimator</i> – Estimador de Máxima Verossimilhança
MODE	<i>Method of Direction Estimation</i>
MODEX	<i>MODEX with Extra Roots</i>
REQM	Raiz do Erro Quadrático Médio – <i>Root Mean Square Error</i>
SMLE	<i>Stochastic Maximum Likelihood Estimator</i> – Critério Estocástico do Estimador de Máxima-Verossimilhança
SEAD	<i>Search of Direction by Differential Spectrum</i> – Busca de Direção pelo Espectro Diferencial
SNR	<i>Signal-to-Noise Ratio</i> – Relação Sinal-Ruído
ULA	<i>Uniform Linear Array</i> – Arranjo Linear Uniforme

# LISTA DE SÍMBOLOS

$(\cdot)^*$	Complexo-conjugado de escalar, vetor ou matriz
$\delta_{(k,i)}(\cdot)$	Delta de Kronecker
$a$ ou $A$	Escalar
$E(\cdot)$	Esperança estatística
<b>A</b>	Matriz (letra maiúscula em negrito)
<b>I</b>	Matriz identidade
$\ \cdot\ $	Norma euclidiana
$M$	Número de fontes de sinal
$K$	Número de sensores no arranjo
$N$	Número de <i>snapshots</i> ou amostras de sinal
$\text{abs}(\cdot)$	Operador absoluto ou módulo
$\text{conv}(\cdot, \cdot)$	Operador convolução
$\text{diag}(\cdot)$	Operador diagonal
$\Im(\cdot)$	Operador parte imaginária
$\Re(\cdot)$	Operador parte real
$\text{posto}(\cdot)$	Operador posto matricial
$\text{tr}(\cdot)$	Operador traço matricial
$(\cdot)^+$	Pseudo-inversa de Moore-Penrose
$(\cdot)^H$	Transposição complexo-conjugada de vetor ou matriz (“hermitiano”)
$(\cdot)^T$	Transposição de vetor ou matriz
<b>a</b>	Vetor linha ou coluna (letra minúscula em negrito)
<b>0</b>	Vetor ou matriz com todos elementos iguais a zero
<b>1</b>	Vetor ou matriz com todos elementos iguais a um

# SUMÁRIO

<b>1</b>	<b>Introdução</b>	<b>18</b>
1.1	Motivação	21
1.2	Justificativa	21
1.3	Objetivos	22
1.4	Problema e Hipóteses	22
1.5	Estrutura deste Trabalho	23
1.6	Considerações Finais	24
<b>2</b>	<b>Modelo de Sinal e Estimador de Máxima-Verossimilhança</b>	<b>25</b>
2.1	Modelo de Sinal Considerando o Arranjo Linear Uniforme	25
2.1.1	Modelos estocásticos de sinal e ruído	29
2.1.2	Matriz de correlação espacial	30
2.2	CrITÉrios de Avaliação dos Estimadores DOA	31
2.3	O Estimador de Máxima-Verossimilhança	32
2.4	O Estimador de Máxima-Verossimilhança Determinístico	34
2.5	O Estimador de Máxima-Verossimilhança Estocástico	35
2.6	Simulações Numéricas	35
2.7	Limite de Cramér-Rao para o DMLE	38
2.8	Considerações Finais	38
<b>3</b>	<b>Estimadores Iterativos e Filtragem Espacial</b>	<b>40</b>
3.1	O Estimador MODEX Modificado	42
3.2	Reparametrização da Função-Objetivo do DMLE	42
3.2.1	O estimador MODE	44
3.2.2	O estimador MODEX Modificado	45
3.3	O Estimador SEAD Aprimorado	46
3.3.1	Espectro diferencial	47
3.3.2	Processo de seleção de estimativas	50
3.3.2.1	Pré-seleção	51
3.3.3	Refinamento das estimativas	53
3.4	Definição do Processo de Filtragem e da Nova Função-Objetivo	55
3.5	Filtro para Maximização da SNR	57
3.6	Considerações Finais	59
<b>4</b>	<b>Proposta de Alteração ao Estimador SEAD</b>	<b>61</b>
4.1	Análise da Resposta em Frequência do Máximo Autofiltro Espacial	63
4.2	Propostas de Filtro Espacial	68
4.2.1	Proposta 1: filtragem dos <i>snapshots</i> usando <i>todo</i> o subespaço de sinal	68

4.2.2	Proposta 2: filtragem dos <i>snapshots</i> usando espectro diferencial . . .	69
4.3	Análise da Resposta em Frequência dos Filtros Espaciais Propostos . . . .	70
4.4	Novo Estágio de Refinamento de Estimativas . . . . .	76
4.4.1	Fundamentação teórica . . . . .	76
4.4.2	Método de Newton para o DMLE sem filtragem espacial utilizando a matriz de predição linear . . . . .	81
4.4.3	Proposta 1: Método de Newton para o DMLE com filtragem espacial utilizando a matriz de predição linear . . . . .	83
4.4.4	Proposta 2: Método de Newton para o DMLE sem filtragem espacial utilizando a matriz de direção . . . . .	84
4.4.5	Proposta 3: Método de Newton para o DMLE com filtragem espacial utilizando a matriz de direção . . . . .	85
4.4.6	Propostas 4 e 5: Método quase-Newton BFGS para o DMLE com e sem filtragem espacial utilizando a matriz de direção . . . . .	86
4.5	Modificação no Estágio de Pré-seleção . . . . .	86
4.6	Considerações Finais . . . . .	88
<b>5</b>	<b>Resultados . . . . .</b>	<b>91</b>
5.1	Simulações Numéricas . . . . .	91
5.1.1	Desempenho de estimação DOA . . . . .	91
5.1.1.1	MODEX Modificado . . . . .	92
5.1.1.2	SEAD Aprimorado . . . . .	96
5.1.1.3	SEAD Modificado . . . . .	100
5.1.1.4	Discussão sobre os resultados . . . . .	104
5.1.2	Tempos de execução . . . . .	108
5.1.2.1	Discussão sobre os resultados . . . . .	111
5.2	Conclusões . . . . .	112
<b>6</b>	<b>Conclusões, Contribuições e Trabalhos Futuros . . . . .</b>	<b>114</b>
6.1	Contribuições deste Trabalho . . . . .	117
6.2	Divulgação Científica Efetivada . . . . .	118
6.3	Trabalhos Futuros . . . . .	118
	<b>Referências . . . . .</b>	<b>121</b>
	<b>Apêndices . . . . .</b>	<b>125</b>
<b>A</b>	<b>Derivação do Método de Newton para o DMLE sem Filtragem Espacial Usando a Matriz de Predição Linear . . . . .</b>	<b>126</b>
A.1	Introdução . . . . .	126
A.2	Derivada de <b>B</b> . . . . .	127

A.3	Derivada de $\mathbf{B}^+$	129
A.4	Derivada de $\mathbf{P}_\mathbf{B}$	129
A.5	Gradiente de $\text{tr}(\mathbf{P}_\mathbf{B}\hat{\mathbf{R}})$	130
A.6	Hessiana de $\text{tr}(\mathbf{P}_\mathbf{B}\hat{\mathbf{R}})$	132
A.7	Aproximação assintótica na hessiana de $\text{tr}(\mathbf{P}_\mathbf{B}\hat{\mathbf{R}})$	133
A.8	Prova 1: $\mathbf{B}^+\mathbf{B} = \mathbf{I}$	135
A.9	Prova 2: $(\mathbf{B}^+)^H = \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}$	135
A.10	Prova 3: $\mathbf{B}^+(\mathbf{B}^+)^H = (\mathbf{B}^H\mathbf{B})^{-1}$	135
A.11	Prova 4: $\frac{\partial \mathbf{P}_\mathbf{B}^\perp}{\partial \phi_k} = -\frac{\partial \mathbf{P}_\mathbf{B}}{\partial \phi_k}$	136
A.12	Prova 5: $\mathbf{B}^+\hat{\mathbf{R}}\mathbf{P}_\mathbf{B}^\perp \approx \mathbf{B}^+\mathbf{R}\mathbf{P}_\mathbf{B}^\perp = \mathbf{0}$	136
A.13	Prova 6: $\mathbf{P}_\mathbf{B}^\perp\hat{\mathbf{R}}\mathbf{P}_\mathbf{B}^\perp \approx \mathbf{P}_\mathbf{B}^\perp\mathbf{R}\mathbf{P}_\mathbf{B}^\perp = \mathbf{R}\mathbf{P}_\mathbf{B}^\perp$	137
A.14	Prova 7: $\mathbf{B}^+\hat{\mathbf{R}}(\mathbf{B}^+)^H \approx \mathbf{B}^+\mathbf{R}(\mathbf{B}^+)^H = \sigma(\mathbf{B}^H\mathbf{B})^{-1}$	138
<b>B</b>	<b>Derivação do Método de Newton para o DMLE com Filtragem Espacial</b>	
	<b>Usando a Matriz de Predição Linear</b>	<b>140</b>
B.1	Introdução	140
B.2	Derivada primeira de $\mathbf{U}$	140
B.3	Derivada segunda de $\mathbf{U}$	141
B.4	Derivada de $\mathbf{U}^+$	143
B.5	Derivada de $\mathbf{P}_\mathbf{\hat{B}}$	144
B.6	Gradiente de $\text{tr}(\mathbf{P}_\mathbf{\hat{B}}\hat{\mathbf{R}}_Z)$	145
B.7	Hessiana de $\text{tr}(\mathbf{P}_\mathbf{\hat{B}}\hat{\mathbf{R}}_Z)$	147
B.8	Aproximação assintótica e numérica na hessiana de $\text{tr}(\mathbf{P}_\mathbf{\hat{B}}\hat{\mathbf{R}}_Z)$	159
B.9	Prova 8: $\mathbf{P}_\mathbf{\hat{A}}^\perp = \mathbf{I} - \mathbf{U}\mathbf{U}^+ \triangleq \mathbf{P}_\mathbf{\hat{B}}$	160
B.10	Prova 9: $\mathbf{U}\mathbf{U}^+(\mathbf{U}^+)^H = (\mathbf{U}^+)^H$	161
B.11	Prova 10: $\mathbf{U}\mathbf{U}^+\mathbf{U} = \mathbf{U}$	162
B.12	Prova 11: $\mathbf{P}_\mathbf{\hat{B}}^\perp\mathbf{U}^+ = \mathbf{U}^+$	162
<b>C</b>	<b>Derivação do Método de Newton para o DMLE sem Filtragem Espacial</b>	
	<b>Usando a Matriz de Direção</b>	<b>163</b>
C.1	Introdução	163
C.2	Derivada primeira de $\mathbf{A}$	164
C.3	Derivada segunda de $\mathbf{A}$	165
C.4	Derivada de $\mathbf{A}^+$	166
C.5	Derivada de $\mathbf{P}_\mathbf{\hat{A}}^\perp$	166
C.6	Gradiente de $\text{tr}(\mathbf{P}_\mathbf{\hat{A}}^\perp\hat{\mathbf{R}})$	167
C.7	Hessiana de $\text{tr}(\mathbf{P}_\mathbf{\hat{A}}^\perp\hat{\mathbf{R}})$	169
C.8	Aproximação assintótica na hessiana de $\text{tr}(\mathbf{P}_\mathbf{\hat{A}}^\perp\hat{\mathbf{R}})$	171
C.9	Prova 12: $\mathbf{A}^+\mathbf{A} = \mathbf{I}$	173
C.10	Prova 13: $(\mathbf{A}^+)^H = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}$	173

C.11 Prova 14: $\mathbf{A}^+ (\mathbf{A}^+)^H = (\mathbf{A}^H \mathbf{A})^{-1}$	173
C.12 Prova 15: $\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{P}_A^\perp \mathbf{R} \mathbf{P}_A^\perp = \sigma \mathbf{P}_A^\perp$	174
C.13 Prova 16: $\mathbf{A}^+ \hat{\mathbf{R}} (\mathbf{A}^+)^H \approx \mathbf{A}^+ \mathbf{R} (\mathbf{A}^+)^H = \mathbf{P} + \sigma (\mathbf{A}^H \mathbf{A})^{-1}$	175
C.14 Prova 17: $\mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{A}^+ \mathbf{R} \mathbf{P}_A^\perp = \mathbf{0}$	175
C.15 Prova 18: $\mathbf{P}_A \mathbf{R} = \mathbf{P}_A \mathbf{R} \mathbf{P}_A = \mathbf{R} \mathbf{P}_A$	176
<b>D Derivação do Método de Newton para o DMLE com Filtragem Espacial</b>	
<b>Usando a Matriz de Direção</b>	<b>178</b>
D.1 Introdução	178
D.2 Derivada primeira de $\tilde{\mathbf{A}}$	178
D.3 Derivada segunda de $\tilde{\mathbf{A}}$	178
D.4 Derivada de $\tilde{\mathbf{A}}^+$	179
D.5 Derivada de $\mathbf{P}_A^\perp$	179
D.6 Gradiente de $\text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}}_Z)$	181
D.7 Hessiana de $\text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}}_Z)$	182
D.8 Aproximação numérica na hessiana de $\text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}}_Z)$	184
D.9 Prova 19: $\tilde{\mathbf{A}}^+ \tilde{\mathbf{A}} = \mathbf{I}$	185
D.10 Prova 20: $(\tilde{\mathbf{A}}^+)^H = \tilde{\mathbf{A}} (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1}$	185
D.11 Prova 21: $\tilde{\mathbf{A}}^+ (\tilde{\mathbf{A}}^+)^H = (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1}$	185
D.12 Prova 22: $\mathbf{P}_{\tilde{\mathbf{A}}} \mathbf{P}_{\tilde{\mathbf{A}}} = \mathbf{P}_{\tilde{\mathbf{A}}}$	185
<b>E Resultados Auxiliares</b>	<b>187</b>

# 1 INTRODUÇÃO

A estimação da direção de chegada (DOA, do inglês *Direction-of-Arrival*) de sinais amostrados por um arranjo de sensores é um tema muito pesquisado na literatura.

Primeiramente, os arranjos podem ser formados por microfones, antenas dipolo, acelerômetros, sismógrafos etc. (SORELIUS, 1999). Uma das mais importantes aplicações é a estimação DOA de sinais modulados digitalmente encontrados em sistemas de telecomunicação sem-fio multiusuário (KRUMMENAUER, 2011).

Nesse contexto, sabe-se que a demanda por serviços de transmissão de dados sem-fio é sempre crescente (BOAVENTURA, 1998), com número cada vez maior de dispositivos portáteis conectados à Internet. Uma das alternativas para se conseguir o aumento da capacidade desses serviços é com o uso de arranjos de sensores.

Esse aumento pode ser alcançado, pois arranjos de sensores são capazes de alterar dinamicamente o seu padrão de irradiação de forma a ajustar o ruído, a interferência e o multicaminho, além de rastrear usuários móveis (LIBERTI; RAPPAPORT, 1999). Consequentemente, o uso desses arranjos permite melhorar os sinais recebidos por uma Estação Rádio-Base (ERB) e formar feixes estreitos (do inglês *beamforming*) para transmissão de dados com um usuário móvel específico.

Assim, a utilização de arranjos de sensores pode resultar em (LIBERTI; RAPPAPORT, 1999):

- custos reduzidos de implantação, com número menor de ERBs para cobrir uma área;
- diminuição do número de zonas cegas;
- aumento da robustez do sistema a perturbações como o ruído;
- aumento da taxa de transmissão de dados, e
- aumento do número de usuários atendidos, pois os mesmos recursos espectrais podem ser reutilizados mais frequentemente.

Adicionalmente, pode-se verificar um aumento de eficiência energética com a possibilidade da redução da potência de transmissão nos dispositivos móveis. Isso é particularmente importante, pois a maior parte desses dispositivos é alimentada por bateria.

No entanto, os arranjos de sensores precisam conhecer *a priori* os ângulos DOA das fontes de sinal. A partir desse parâmetro, esse arranjo ajusta o ganho ou o deslocamento de fase para que o seu padrão de irradiação seja direcionado para a posição indicada pelo

DOA. Assim, sinais provenientes dessa direção têm sua recepção favorecida, ao mesmo tempo em que os sinais de outras direções são bastante atenuados (GODARA, 1997).

Um dos principais aspectos de um estimador DOA é que ele deve possuir alta resolução. Dessa forma, ele deve ser capaz de distinguir ângulos DOA de fontes de sinal espacialmente próximas entre si. Por outro lado, a situação de fontes afastadas entre si também deve ser considerada.

Também, um estimador ideal deve ser tão robusto ao ruído quanto possível, mesmo em condições como baixas relações sinal-ruído (ou SNR, do inglês *Signal-to-Noise Ratio*) ou pequeno número de amostras de sinal (ALVES, 2004). Quando esses valores caem abaixo de um determinado valor-limite, ocorre o efeito de limiar e o erro de estimação aumenta consideravelmente (GERSHMAN; STOICA, 1999). A partir desse ponto, o estimador não gera estimativas estatisticamente confiáveis.

Adicionalmente, a estimação DOA deve ser feita rapidamente. Por exemplo, o arranjo de sensores utilizado em um equipamento compatível com o padrão IEEE 802.16e (ou WiMAX, do inglês *Worldwide Interoperability for Microwave Access*) é recalibrado a cada cinco milissegundos (FUJITSU, INC.; CISCO, INC., 2008).

Um estimador pesquisado na literatura é o Estimador de Máxima-Verossimilhança (MLE, do inglês *Maximum Likelihood Estimator*). Ele pode ser usado em diversos problemas de estimação considerados difíceis (KAY, 1993) e, em teoria, possui desempenho de estimação superior (STARER; NEHORAI, 1992). Em contrapartida, a otimização da função de verossimilhança é computacionalmente intensa e nem sempre possui convergência garantida (STOICA; SHARMAN, 1990; LI; LU, 2007).

Dessa forma, foram desenvolvidos estimadores iterativos. O MODE (do inglês *Method of Direction Estimation* ou Método de Estimação de Direção) (STOICA; SHARMAN, 1990) emprega uma reparametrização da função-objetivo do MLE de maneira a otimizar os coeficientes de um polinômio cujas raízes estimam os ângulos DOA. Por isso, o MODE é conhecido por ser um estimador estatisticamente eficiente. Entretanto, ele é pouco robusto ao efeito de limiar, pois ele faz uma aproximação assintótica do MLE (GERSHMAN; STOICA, 1999).

O MODEX Modificado (do inglês *MODE with Extra Roots* ou MODE com raízes extras) (LOPES et al., 2003; ALVES, 2004) aprimorou significativamente o desempenho em baixas SNRs ao utilizar três soluções MODE geradas a partir de diferentes restrições de não-trivialidade. Como cada combinação de raízes produz uma estimativa DOA, o MODEX Modificado aplica um procedimento de máxima-verossimilhança (ML, do inglês *Maximum Likelihood*) para selecionar a melhor combinação delas. Entretanto, ao fazer isso, esse estimador exige um esforço computacional maior do que o MODE, especialmente quando aumenta o número de fontes de sinal.

O SEAD (do inglês *Search of Direction by Differential Spectrum* ou Busca de Direção



pelo Espectro Diferencial) (FERREIRA, 2005; FERREIRA; LEMOS, 2005a; FERREIRA; LEMOS, 2005b) foi proposto após os pesquisadores terem observado o comportamento dos autovalores da matriz de covariância espacial. O principal resultado dessa pesquisa foi o espectro diferencial. Esse espectro gera picos nas regiões dos ângulos DOA e permite gerar rapidamente estimativas iniciais próximas dos valores reais. Por outro lado, o algoritmo utilizado para refinar essas estimativas é muito simples e apresenta um esforço computacional muito elevado.

Então, o SEAD Aprimorado (SILVA, 2009; SILVA et al., 2013) foi proposto com o objetivo de diminuir esse esforço ao aplicar um novo método de busca discreta local baseado no algoritmo *Branch-and-Bound*. A redução alcançada foi bastante significativa, a ponto de tornar o SEAD um estimador viável, mesmo que o SEAD Aprimorado ainda apresente esforço computacional de ordem exponencial. Por outro lado, a sua resolução é limitada pelo número de iterações utilizado no método de busca. Por isso, o SEAD Aprimorado não é um estimador assintoticamente eficiente quando a SNR aumenta.

Os estimadores MODEX Modificado e o SEAD Aprimorado são baseados no MLE. Assim, eles dividem o espaço de medida dos dados observados em subespaços de sinal e de ruído (ou subespaço ortogonal) (LI; VACCARO; TUFTS, 1990). Tipicamente, todos estimadores de parâmetros baseados em subespaço apresentam o efeito de limiar (HAWKES; NEHORAI; STOICA, 2001). Esse efeito é devido a uma descontinuidade nas estimativas e é associado a uma maior probabilidade de que os dados amostrados são melhor aproximados por alguns componentes do subespaço ortogonal do que por alguns componentes do subespaço de sinal. Quando isso acontece, considera-se que ocorreu uma *troca de subespaço* (THOMAS; SCHARF; TUFTS, 1995; ABRAMOVICH; SPENCER, 2002).

Dessa forma, com o objetivo de diminuir o efeito de limiar no MODEX Modificado, foi proposto em (KRUMMENAUER et al., 2010) um filtro espacial ótimo que maximiza a SNR de saída (FORSTER; VEZZOSI, 1987; HAYKIN, 2001). Além disso, foi proposta uma nova função-objetivo do MLE para permitir a aplicação desse filtro. O resultado foi uma melhoria significativa no desempenho de estimação do MODEX Modificado especialmente na região de limiar.

Sob outra perspectiva, otimizar a função de verossimilhança é um problema não-linear e apresenta elevado esforço computacional (STARER; NEHORAI, 1992). Assim, é impraticável em uma aplicação em tempo real. Uma alternativa é obter uma solução aproximada utilizando métodos numéricos de busca local, como o Método de Newton e o BFGS (Broyden-Fletcher-Goldfarb-Shanno). Em (STARER, 1990) foi desenvolvida uma versão do primeiro para o MLE sem a aplicação da filtragem espacial e não foi encontrada na literatura a aplicação do BFGS à estimação DOA. Entretanto, outros métodos também podem ser usados, como o algoritmo Levenberg-Marquardt, embora seja preferível utilizar o Método de Newton se estiverem disponíveis o gradiente e a hessiana

(STARER; NEHORAI, 1992).

A proposta deste trabalho é melhorar o desempenho de estimação do SEAD Aprimorado utilizando filtragem espacial e métodos numéricos.

Este capítulo apresenta a motivação, a justificativa, os objetivos e a estrutura deste trabalho. Finalmente, são realizadas as considerações finais deste capítulo.

## 1.1 Motivação

Diversas tecnologias de transmissão de dados sem-fio com altas taxas de transmissão implementam técnicas de *beamforming*. Para que essas tecnologias apresentem melhor desempenho, é necessário que os estimadores DOA sejam cada vez mais confiáveis, mesmo em situações adversas, tais como: condições severas de ruído, alto fator de correlação entre sinais recebidos e fontes próximas entre si. Adicionalmente, os estimadores DOA devem apresentar baixo esforço computacional para tornar possível a estimação da direção de um número cada vez maior de usuários.

Assim, o aprimoramento dos estimadores DOA torna os sistemas de telecomunicação sem-fio mais rápidos, robustos e confiáveis.

## 1.2 Justificativa

O estimador DOA denominado SEAD Aprimorado é recente (SILVA et al., 2013) e ainda se encontra em desenvolvimento. Além disso, em (KRUMMENAUER et al., 2010) foi proposto um processo de filtragem espacial que ainda não foi aplicado a esse estimador. Adicionalmente, existem no SEAD Aprimorado diversos aspectos que devem ser melhorados.

Um desses aspectos é o método de busca local implementado, pois é um algoritmo com alto esforço computacional e que discretiza o espaço de busca. Uma alternativa para solucionar isso é o Método de Newton, que, embora já tenha sido desenvolvido em (STARER; NEHORAI, 1992) para o MLE, ele não considerou a filtragem espacial. Outro aspecto que merece atenção é o processo de geração de estimativas iniciais. Esse processo depende fortemente de um limiar de amplitude, podendo aumentar muito o erro de estimação principalmente na presença de fontes afastadas.

Dessa forma, esperam-se melhorias importantes em termos do desempenho de estimação e no esforço computacional do SEAD Aprimorado ao aplicar o processo de filtragem espacial, utilizando métodos de busca local mais apropriados e propor alterações ao processo de geração de estimativas iniciais.

## 1.3 Objetivos

O objetivo geral deste trabalho é tornar o SEAD Aprimorado um estimador assintoticamente eficiente, com menor esforço computacional, maior robustez ao ruído e confiável tanto na situação de fontes próximas quanto afastadas. Para isso, os objetivos específicos deste trabalho são:

- Avaliar o processo de filtragem espacial proposto em (KRUMMENAUER et al., 2010) na estimação DOA de fontes próximas e afastadas e, se for o caso, propor alterações relativas ao processo em si ou ao filtro espacial utilizado;
- Desenvolver *analiticamente* o Método de Newton considerando a ausência e a presença do processo de filtragem espacial;
- Aplicar o BFGS ao MLE considerando a ausência e a presença do processo de filtragem espacial;
- Aplicar alterações no SEAD Aprimorado que contribuam positivamente na estimação DOA de fontes afastadas;
- Substituir no SEAD Aprimorado o método de busca local baseado no algoritmo *Branch-and-Bound* pelo Método de Newton e pelo BFGS, avaliando-os na ausência e na presença do processo de filtragem espacial; e
- Comparar os resultados da nova versão do SEAD Aprimorado com a sua versão anterior e com o MODEX Modificado.

## 1.4 Problema e Hipóteses

O SEAD Aprimorado possui algumas características que o torna um estimador DOA sub-ótimo. Duas delas são relativas ao algoritmo de busca local utilizado e uma é sobre o processo de geração de estimativas iniciais.

O algoritmo de busca local do SEAD Aprimorado possui esforço computacional de ordem exponencial. A consequência disso é mais evidenciada quando se aumenta o número de fontes de sinal, com tempos de execução elevados, podendo ser incompatíveis com aplicações onde a estimação DOA é aplicada.

Ainda, esse algoritmo é baseado no *Branch-and-Bound* e, assim, realiza a busca em um intervalo discretizado. O resultado é que o erro de estimação é maior do que o Limite de Cramér-Rao quando a SNR é suficientemente alta. Isso impõe ao SEAD Aprimorado uma resolução máxima no desempenho de estimação e, por mais que a SNR aumente consideravelmente, o erro de estimação não decresce. Por outro lado, quando essa resolução é atingida, uma alternativa é aumentar a quantidade de iterações utilizadas no estágio

de busca. Ao fazer isso, o erro de estimação se aproxima do Limite de Cramér-Rao, mas aumenta *ainda mais* o esforço computacional.

Diferentemente, o Método de Newton e o BFGS são algoritmos que realizam essa busca local em um intervalo contínuo. Ainda que eles sejam métodos numéricos e, portanto, apresentam uma resolução máxima, eles são menos dependentes do número de iterações utilizadas do que algoritmo *Branch-and-Bound*. Adicionalmente, sabe-se que o Método de Newton apresenta maior esforço computacional do que o BFGS, embora o primeiro convirja para a solução do problema mais rapidamente do que o segundo. Portanto, a primeira hipótese desta tese é uma diminuição no esforço computacional e, em altas SNRs, uma diminuição no erro de estimação, tão logo o algoritmo de busca local utilizado no SEAD Aprimorado seja substituído pelo Método de Newton ou pelo BFGS.

Entretanto, é importante ressaltar que esses métodos exigem que as estimativas iniciais estejam suficientemente próximas da solução do problema. Isso ocorre de maneira frequente no processo de geração de estimativas iniciais já utilizado pelo SEAD Aprimorado. Contudo, o funcionamento desse processo é direcionado por um limiar de amplitude que pode ser mal dimensionado. Sempre que isso ocorre, causa severos desvios nas estimativas iniciais, afetando também o algoritmo de busca local, produzindo um maior erro de estimação. É importante ressaltar que esses desvios ocorrem mais frequentemente na presença de fontes afastadas. Portanto, a segunda hipótese é um menor erro de estimação nessa situação ao se alterar o processo de geração de estimativas iniciais de forma que não utilize o limiar de amplitude.

Sob outra perspectiva, a aplicação do processo de filtragem espacial utilizando o máximo autofiltro espacial aprimorou o desempenho de estimação do MODEX Modificado na situação de baixas SNRs e fontes próximas, como mostrado em (KRUMMENAUER et al., 2010). Assim, a última hipótese é que a filtragem espacial pode diminuir o erro de estimação do SEAD Aprimorado nas mais diversas situações aplicáveis à estimação DOA, principalmente em baixas SNRs.

De maneira geral e se confirmar todas essas hipóteses, espera-se que o SEAD Aprimorado se torne um estimador com baixo esforço computacional e com reduzido erro de estimação em qualquer valor de SNR e na presença ou na ausência de fontes afastadas.

## 1.5 Estrutura deste Trabalho

O Capítulo 2 apresenta os modelos de sinal e de ruído, o problema de estimação DOA e os critérios de avaliação dos estimadores que serão utilizados nas simulações numéricas realizadas neste trabalho. Além disso, são apresentados os critérios determinístico e estocástico do Estimador de Máxima-Verossimilhança e o Limite de Cramér-Rao para o critério determinístico.

A fundamentação teórica dos estimadores MODEX Modificado e SEAD Aprimorado é descrita no Capítulo 3. Adicionalmente, é apresentado o processo de filtragem utilizando o máximo autofiltro espacial projetado considerando a situação de fontes próximas.

O Capítulo 4 propõe dois novos filtros espaciais. A primeira proposta é a extensão do máximo autofiltro espacial com o objetivo de torná-lo adequado tanto para fontes próximas quanto afastadas. A segunda proposta é baseada no espectro diferencial. Além disso, são apresentadas três versões do Método de Newton e duas do BFGS aplicados ao MLE com e sem filtragem espacial.

Os resultados obtidos neste trabalho são mostrados no Capítulo 5.

O Capítulo 6 apresenta as conclusões, as contribuições deste trabalho e os trabalhos futuros que poderão ser realizados a partir desta tese.

## 1.6 Considerações Finais

Este capítulo apresentou a motivação, a justificativa, a hipótese e a estrutura deste trabalho.

O próximo capítulo define o problema da estimação DOA, apresentando o modelo de sinal, os critérios de avaliação dos estimadores DOA e a fundamentação teórica do Estimador de Máxima-Verossimilhança.

## 2 MODELO DE SINAL E ESTIMADOR DE MÁXIMA-VEROSSIMILHANÇA

Os métodos de estimação DOA utilizam modelos estatísticos que descrevem os sinais amostrados pelo arranjo de sensores. Esses modelos são baseados nas características do arranjo, das fontes de sinal e do meio de propagação (SILVA, 2007).

Além disso, esses modelos são a base para o Estimador de Máxima-Verossimilhança. A partir dos modelos de sinal é possível gerar duas versões (ou critérios) do MLE: o MLE Determinístico (DMLE, do inglês *Deterministic MLE*) e o MLE Estocástico (SMLE, do inglês *Stochastic MLE*).

Este capítulo apresenta os modelos de sinal e de ruído, o problema de estimação DOA e os critérios de avaliação dos estimadores que serão utilizados nas simulações numéricas realizadas neste trabalho. Ainda, são apresentados os critérios determinístico e estocástico do Estimador de Máxima-Verossimilhança. Adicionalmente, este capítulo apresenta o Limite de Cramér-Rao para o DMLE. Finalmente, são realizadas as considerações finais.

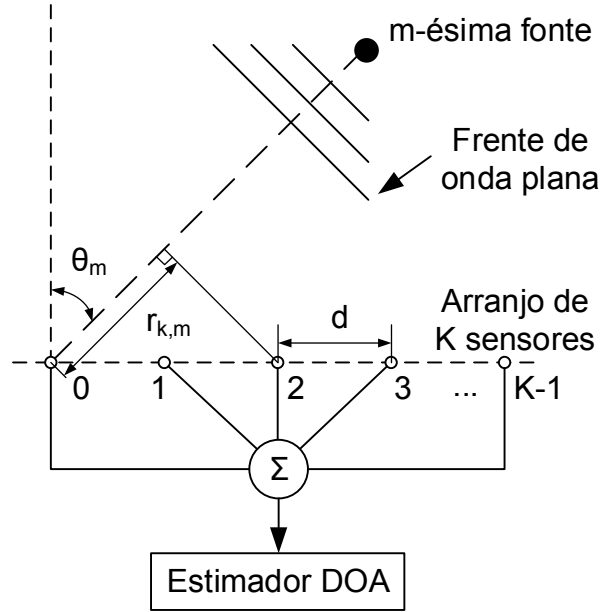
### 2.1 Modelo de Sinal Considerando o Arranjo Linear Uniforme

A estimação do parâmetro DOA é realizada a partir dos sinais amostrados por um arranjo formado por  $K$  sensores. Assume-se que existem  $M$  fontes de sinal.  $M$  é conhecido e  $M < K$ . Uma fonte de sinal pode emitir sinais acústicos, eletromagnéticos, tectônicos ou de outros tipos. Portanto, os sensores que formam um arranjo podem ser microfones, antenas dipolo, acelerômetros, sismógrafos etc. (SORELIUS, 1999) dispostos de maneira arbitrária (KRIM; VIBERG, 1996).

Para simplificar o modelo de sinal, foram feitas algumas considerações. A primeira delas é sobre o tipo de arranjo utilizado. Neste trabalho foi considerado o Arranjo Linear Uniforme (ou ULA, do inglês *Uniform Linear Array*) pela sua simplicidade no equacionamento e porque ele é bastante utilizado na literatura da área (SILVA, 2007). Todos os sensores são iguais, de baixo ganho e sem acoplamento mútuo entre eles (LIBERTI; RAPPAPORT, 1999). A segunda consideração é sobre a fonte de sinal, localizada em uma região muito distante do arranjo. Essa fonte é pontual. Assim, as suas frentes de onda podem ser consideradas planas (ALVES, 2004). A Figura 1 ilustra uma representação bidimensional de uma fonte emitindo sinais amostrados por uma ULA.

Pode-se observar na Figura 1 que a  $m$ -ésima fonte de sinal possui o ângulo DOA  $\theta_m$ . Os  $K$  sensores estão dispostos linearmente e espaçados em meio comprimento de onda,

Figura 1 – Representação bidimensional de um arranjo de sensores amostrando sinais de uma fonte.



$d = c/(2f)$ , em que:  $c = 3 \times 10^8$  m/s é a velocidade da luz no vácuo e  $f$  é a frequência do sinal em Hertz (Hz).  $r_{k,m} = kd \sin(\theta_m)$  é a distância percorrida entre as frentes de onda da  $m$ -ésima fonte e o  $k$ -ésimo sensor;  $m = 1, \dots, M$ ;  $k = 0, \dots, (K - 1)$ .

Suponha que cada uma das  $M$  fontes emitam sinais como mostrado na Equação (2.1) (ALVES, 2004).

$$f_m(t) = F_m(t) e^{j[\bar{\omega}_m t - \frac{2\pi}{\vartheta_m} r_m + \phi_m(t)]} \quad (2.1)$$

em que:  $F_m(t)$  é a amplitude do sinal variando no tempo;  $\bar{\omega}_m$  é a frequência da onda portadora;  $\theta_m$  é o comprimento de onda na direção de propagação;  $r_m$  é a distância da fonte até a origem de um sistema de coordenadas que representa a posição de um sensor do arranjo e  $\phi_m(t)$  é uma fase arbitrária.

A terceira consideração é que os sinais incidentes são de faixa estreita. Em outras palavras, os termos  $F_m(t)$  e  $\phi_m(t)$  variam lentamente em relação a  $e^{j\bar{\omega}_m t}$ . Assim,  $F_m(t)$  e  $\phi_m(t)$  apresentam os mesmos valores em todos os sensores, pois, tipicamente,  $d$  é da ordem do comprimento de onda. Dessa forma, o sinal amostrado da  $m$ -ésima fonte pelo  $k$ -ésimo sensor é apresentado na Equação (2.2) (ALVES, 2004).

$$x_{(k,m)}(t) = F_m(t) e^{j[\bar{\omega}_m t - \frac{2\pi}{\vartheta_m} r_{(k,m)} + \phi_m(t)]} \quad (2.2)$$

Então, considerando o princípio da superposição e a hipótese de ondas planas, o

sinal resultante amostrado em cada sensor do arranjo é mostrado na Equação (2.3).

$$\begin{aligned} x_k(t) &= \sum_{m=1}^M x_{(k,m)}(t) \\ &= \sum_{m=1}^M F_m(t) e^{j[\bar{\omega}_m t + \phi_m(t) - \frac{2\pi}{\vartheta_m} r_{(k,m)}]} \end{aligned} \quad (2.3)$$

É importante verificar que a diferença entre as fases do sinal amostrado por dois sensores distintos no arranjo é devido apenas à distância entre eles. Considerando que os sinais tenham a mesma portadora, tem-se que  $\vartheta_1 = \vartheta_2 = \dots = \vartheta_M = \vartheta$ . Dessa forma, o terceiro termo da exponencial da Equação (2.3) pode ser reduzido para  $\frac{2\pi}{\vartheta} k d \sin(\theta_m)$ . Então, definindo a frequência angular  $\omega_m \triangleq -\frac{2\pi}{\vartheta} d \sin(\theta_m)$  (KRUMMENAUER et al., 2013) em função do ângulo DOA da  $m$ -ésima fonte, a Equação (2.3) pode ser reescrita como apresentado na Equação (2.4).

$$x_k(t) = \sum_{m=1}^M F_m(t) e^{j[\bar{\omega}_m t + \phi_m(t) + k\omega_m]} \quad (2.4)$$

No entanto, o sinal é demodulado antes de realizar a estimação DOA. Assim, a Equação (2.5) apresenta o sinal amostrado em banda básica pelo  $k$ -ésimo sensor (ALVES, 2004).

$$x_k(t) = \sum_{m=1}^M F_m(t) e^{j[\phi_m(t) + k\omega_m]} \quad (2.5)$$

Esse sinal é composto por  $M$  exponenciais complexas com frequências angulares  $\omega_m$ , fases  $\phi_m$  e amplitudes  $F_m(t)$ . A Equação (2.5) mostra que os conceitos utilizados na estimação de frequência podem ser usados no problema de estimação DOA. Adicionalmente, pode-se simplificar essa equação como apresentado na Equação (2.6) (SILVA, 2007).

$$x_k(t) = \sum_{m=1}^M a_k(\omega_m) s_m(t) \quad (2.6)$$

em que:  $a_k(\omega_m) = e^{jk\omega_m}$  e  $s_m(t) = F_m(t) e^{j\phi_m(t)}$ .

Expandindo a Equação (2.6) para um arranjo linear formado por  $K$  sensores, obtém-se a Equação (2.7).

$$\mathbf{x}(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{(K-1)}(t) \end{bmatrix} = \sum_{m=1}^M \begin{bmatrix} a_0(\omega_m) \\ a_1(\omega_m) \\ \vdots \\ a_{(K-1)}(\omega_m) \end{bmatrix} s_m(t) = \sum_{m=1}^M \mathbf{a}(\omega_m) s_m(t) \quad (2.7)$$



em que:  $\mathbf{a}(\omega_m)$  é chamado de vetor de direção (do inglês *steering vector*), em analogia ao método clássico de conformação de feixes (do inglês *beamforming*). Nesse método, o arranjo é ajustado de tal forma a ter um ganho predeterminado na direção desejada (KRIM; VIBERG, 1996).

Adicionalmente, a Equação (2.7) pode ser expandida para a notação matricial, apresentada na Equação (2.8) (SILVA, 2007).

$$x(t) = \sum_{m=1}^M \mathbf{A}(\boldsymbol{\omega}) \mathbf{s}(t), \quad (2.8)$$

em que:  $\boldsymbol{\omega} = [\omega_1 \ \omega_2 \ \dots \ \omega_M]^T$  e  $\mathbf{s}(t) = [s_1(t) \ s_2(t) \ \dots \ s_M(t)]^T \in \mathbb{C}^{M \times 1}$  são respectivamente o vetor das frequências angulares e o vetor das amplitudes complexas dos sinais;  $(\cdot)^T$  é a operação de transposição vetorial ou matricial e  $\mathbf{A}(\boldsymbol{\omega})$  é a matriz contendo os vetores de direção de cada fonte, definida pela Equação (2.9).

$$\begin{aligned} \mathbf{A}(\boldsymbol{\omega}) &= [\mathbf{a}(\omega_1) \ \mathbf{a}(\omega_2) \ \dots \ \mathbf{a}(\omega_M)] \\ &= \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{j\omega_1} & e^{j\omega_2} & \dots & e^{j\omega_M} \\ e^{j2\omega_1} & e^{j2\omega_2} & \dots & e^{j2\omega_M} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(K-1)\omega_1} & e^{j(K-1)\omega_2} & \dots & e^{j(K-1)\omega_M} \end{bmatrix} \end{aligned} \quad (2.9)$$

A matriz de direção  $\mathbf{A}(\boldsymbol{\omega})$  tem a estrutura de Vandermonde (ALVES, 2004) e, embora esteja em função de  $\boldsymbol{\omega}$ , daqui em diante será referida apenas como  $\mathbf{A}$ .

Deve-se observar que o estimador DOA dispõe de  $N$  amostras dos sinais (ou *snapshots*) obtidas nos instantes  $t = n\Delta t, n = 1, \dots, N$ . Assim, este trabalho adota as notações simplificadas  $\mathbf{x}[n]$  e  $\mathbf{s}[n]$  para representar os sinais e suas amplitudes no  $n$ -ésimo *snapshot*. Além disso, o sinal é contaminado por ruído aditivo e a saída do arranjo pode ser expressa como mostrado na Equação (2.10).

$$\begin{aligned} \mathbf{y}[n] &= \mathbf{x}[n] + \mathbf{n}[n] \\ &= \mathbf{A}\mathbf{s}[n] + \mathbf{n}[n] \end{aligned} \quad (2.10)$$

em que:  $\mathbf{n}[n]$  é o vetor do ruído.

Finalmente, as Equações (2.11) a (2.13) definem as matrizes  $\mathbf{X}$ ,  $\mathbf{Y}$  e  $\mathbf{N}$ .

$$\mathbf{X} = [\mathbf{x}[1] \ \mathbf{x}[2] \ \dots \ \mathbf{x}[N]] \in \mathbb{C}^{K \times N} \quad (2.11)$$

$$\mathbf{Y} = [\mathbf{y}[1] \ \mathbf{y}[2] \ \dots \ \mathbf{y}[N]] \in \mathbb{C}^{K \times N} \quad (2.12)$$

$$\mathbf{N} = [\mathbf{n}[1] \ \mathbf{n}[2] \ \dots \ \mathbf{n}[N]] \in \mathbb{C}^{K \times N} \quad (2.13)$$

### 2.1.1 Modelos estocásticos de sinal e ruído

Este trabalho considera que  $\mathbf{s}[n]$  e  $\mathbf{n}[n]$  são processos aleatórios gaussianos. Esses processos são complexos, de média zero, estacionários, independentes e satisfazem às relações apresentadas nas Equações (2.14) a (2.17), para  $i \neq j$  (STOICA; NEHORAI, 1990; SILVA, 2007).

$$\mathbb{E} \left\{ \mathbf{s}[i] \mathbf{s}^H[j] \right\} = \mathbf{C} \delta_{(i,j)} \quad (2.14)$$

$$\mathbb{E} \left\{ \mathbf{s}[i] \mathbf{s}^T[j] \right\} = 0 \quad (2.15)$$

$$\mathbb{E} \left\{ \mathbf{n}[i] \mathbf{n}^H[j] \right\} = \sigma^2 \mathbf{I} \delta_{(i,j)} \quad (2.16)$$

$$\mathbb{E} \left\{ \mathbf{n}[i] \mathbf{n}^T[j] \right\} = 0 \quad (2.17)$$

em que:  $\mathbb{E}\{\cdot\}$  é o operador Esperança Estatística;  $\delta_{(i,j)} = 1, \forall i = j$  e  $\delta_{(i,j)} = 0, \forall i \neq j$  é a função Delta de Kronecker;  $(\cdot)^H$  é a operação de transposição e conjugação vetorial ou matricial (ou “hermitiano”);  $\sigma^2$  é a variância ou potência média do ruído;  $\mathbf{I}$  é a matriz identidade e  $\mathbf{C} \in \mathbb{R}^{M \times M}$  é a matriz de correlação dos sinais dentro de um mesmo *snapshot*, definida pela Equação (2.18).

$$\mathbf{C} = \begin{bmatrix} \rho_{(1,1)} & \rho_{(1,2)} & \cdots & \rho_{(1,M)} \\ \rho_{(2,1)} & \rho_{(2,2)} & \cdots & \rho_{(2,M)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{(M,1)} & \rho_{(M,2)} & \cdots & \rho_{(M,M)} \end{bmatrix} \quad (2.18)$$

em que:  $\rho_{(i,j)} = \rho_{(j,i)}, \forall i \neq j$ , é o fator de correlação entre as fontes  $i$  e  $j$ . Para  $i = j$ ,  $\rho_{(i,j)}$  indica a potência relativa do sinal da  $i$ -ésima fonte.

A Equação (2.14) indica a igualdade das variâncias entre as partes real e imaginária do sinal. Por outro lado, as Equações (2.15) e (2.17) mostram a ortogonalidade entre as partes real e imaginária do sinal e do ruído, respectivamente. Assim, ocorre uma completa decorrelação no tempo entre os sinais. Além disso, ocorre a decorrelação temporal e espacial entre as amostras dos vetores de ruído.

Quando o fator de correlação entre duas fontes  $i$  e  $j$  é  $\rho_{(i,j)} = 0, \forall i \neq j$ , os sinais dessas fontes são decorrelacionados e os vetores de amplitudes complexas são independentes entre si dentro do mesmo *snapshot*. Para  $0 < \rho_{(i,j)} < 1, \forall i \neq j$ , os sinais  $i$  e  $j$  são parcialmente correlacionados. Finalmente, considerando  $\rho_{(i,j)} = 1, \forall i \neq j$ , os sinais são considerados coerentes ou totalmente correlacionados. Nesse caso, o  $j$ -ésimo sinal é uma versão escalonada ou defasada do sinal  $i$  (ALVES, 2004).

Assim, quando  $\mathbf{C}$  é proporcional à matriz identidade de dimensões  $M \times M$ , todos os sinais são equipotentes e decorrelacionados. Então,  $\text{posto}(\mathbf{C}) = M$  (posto cheio) e todas as colunas de  $\mathbf{A}$  são linearmente independentes (ALVES, 2004).

Por outro lado, quando  $\mathbf{C}$  é uma matriz de dimensões  $M \times M$  formada apenas por 1's, todos os sinais são coerentes e as colunas de  $\mathbf{A}$  são linearmente dependentes. Nesse caso,  $\mathbf{C}$  assume posto unitário. Dessa forma, define-se o número de sinais não-coerentes  $\bar{M} = \min [M, \text{posto}(\mathbf{C})]$  (KRUMMENAUER et al., 2010).

Finalmente, existem os seguintes modelos estocásticos de sinal (STOICA; NEHORAI, 1990):

1. **Modelo condicional:** os sinais contidos em  $\mathbf{X}$  são variáveis aleatórias no tempo (ao longo dos *snapshots*) e no espaço (dentro do *snapshot*). Porém, nas demais execuções,  $\mathbf{X}$  é mantido e varia-se apenas o ruído aditivo  $\mathbf{N}$ . Um caso específico é o modelo determinístico, que utiliza a mesma matriz de sinais em todos os *snapshots*.
2. **Modelo incondicional:** os sinais são aleatórios no espaço e no tempo e a cada execução são geradas novas sequências de  $\mathbf{X}$  e  $\mathbf{N}$ . Esse é o modelo adotado nas simulações realizadas neste trabalho, pois o seu uso é difundido na literatura, além de ser o mais realístico (ALVES, 2004).

### 2.1.2 Matriz de correlação espacial

A matriz de correlação espacial é usada em todos os métodos de estimação pesquisados neste trabalho. Teoricamente, ela pode ser calculada a partir de infinitos *snapshots*, como é apresentado na Equação (2.19) (ALVES, 2004).

$$\begin{aligned} \mathbf{R}(i, j) &= \mathbb{E} \left\{ \mathbf{y}[i] \mathbf{y}^H[j] \right\} \\ &= \mathbf{A} \mathbb{E} \left\{ \mathbf{s}[i] \mathbf{s}^H[j] \right\} \mathbf{A}^H + \mathbb{E} \left\{ \mathbf{n}[i] \mathbf{n}^H[j] \right\} \end{aligned} \quad (2.19)$$

em que:  $(i, j) = 1, \dots, \infty$ .

Considerando as Equações (2.14) e (2.16) e a independência entre o sinal e o ruído,  $\mathbf{R}(i, j) \neq \mathbf{0}, \forall i = j$ . Assim, pode-se obter a matriz de correlação espacial mostrada na Equação (2.20).

$$\mathbf{R} = \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma^2 \mathbf{I} \in \mathbb{C}^{K \times K} \quad (2.20)$$

Por outro lado, considerando um número finito de *snapshots*, as esperanças estatísticas na Equação (2.19) são substituídas por médias aritméticas. Isso resulta na matriz de correlação espacial amostrada, apresentada na Equação (2.21).

$$\begin{aligned} \hat{\mathbf{R}} &= \mathbf{A} \left\{ \sum_{n=1}^N \mathbf{s}[n] \mathbf{s}^H[n] \right\} \mathbf{A}^H + \sum_{n=1}^N \mathbf{n}[n] \mathbf{n}^H[n] \\ &= \frac{1}{N} \sum_{n=1}^N \mathbf{y}[n] \mathbf{y}^H[n] \end{aligned} \quad (2.21)$$

Também, a matriz  $\hat{\mathbf{R}}$  pode ser decomposta em autovalores como mostrado na Equação (2.22).

$$\hat{\mathbf{R}} = \sum_{k=1}^K \hat{\lambda}_k \hat{\mathbf{e}}_k \hat{\mathbf{e}}_k^H = \hat{\mathbf{E}}_S \hat{\mathbf{\Lambda}}_S \hat{\mathbf{E}}_S^H + \hat{\mathbf{E}}_N \hat{\mathbf{\Lambda}}_N \hat{\mathbf{E}}_N^H \quad (2.22)$$

em que:  $\hat{\lambda}_k$  é o  $k$ -ésimo autovalor de  $\hat{\mathbf{R}}$  em ordem decrescente e  $\hat{\mathbf{e}}_k$  é o autovetor correspondente;  $\hat{\mathbf{\Lambda}}_S = \text{diag}(\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_{\bar{M}})$  é a matriz diagonal que contém os primeiros  $\bar{M}$  autovalores;  $\hat{\mathbf{E}}_S = [\hat{\mathbf{e}}_1 \ \hat{\mathbf{e}}_2 \ \dots \ \hat{\mathbf{e}}_{\bar{M}}]$  contém os autovetores correspondentes aos autovalores de sinal e abrange o subespaço de sinal de  $\hat{\mathbf{R}}$  juntamente com  $\hat{\mathbf{\Lambda}}_S$ ;  $\hat{\mathbf{\Lambda}}_N = \text{diag}(\hat{\lambda}_{(\bar{M}+1)}, \hat{\lambda}_{(\bar{M}+2)}, \dots, \hat{\lambda}_K)$  e  $\hat{\mathbf{E}}_N = [\hat{\mathbf{e}}_{(\bar{M}+1)} \ \hat{\mathbf{e}}_{(\bar{M}+2)} \ \dots \ \hat{\mathbf{e}}_K]$  contém os  $(K - \bar{M})$  autovalores restantes e os autovetores correspondentes. Da mesma forma,  $\hat{\mathbf{\Lambda}}_N$  e  $\hat{\mathbf{E}}_N$  formam o subespaço ortogonal ou de ruído (LI; VACCARO, 1991).

Adicionalmente, esse mesmo processo de autodecomposição pode ser realizado em  $\mathbf{R}$ , resultando nas matrizes equivalentes  $\mathbf{\Lambda}_S$ ,  $\mathbf{\Lambda}_N$ ,  $\mathbf{E}_S$  e  $\mathbf{E}_N$ . Nesse caso, os autovetores em  $\mathbf{E}_S$  geram o mesmo espaço que as colunas de  $\mathbf{A}$ , enquanto os autovetores em  $\mathbf{E}_N$  geram o complemento ortogonal, denominado espaço nulo de  $\mathbf{A}$ . É importante observar que  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{\bar{M}} > \lambda_{(\bar{M}+1)} = \lambda_{(\bar{M}+2)} = \dots = \lambda_K = \sigma^2$  (ALVES, 2004).

Além disso, o  $m$ -ésimo sinal possui a SNR dimensionada na forma apresentada pela Equação (2.23).

$$\text{SNR}_m = 10 \log_{10} \left\{ \frac{\mathbb{E} [|s_m[n]|^2]}{\sigma^2} \right\} \quad (2.23)$$

## 2.2 Critérios de Avaliação dos Estimadores DOA

Nesta seção são descritos os índices de mérito para a avaliação dos estimadores DOA.

O primeiro deles é o **Limite de Cramér-Rao** (CRB, do inglês *Cramér-Rao Bound*). O CRB limita inferiormente a variância do erro das estimativas geradas por um estimador não-polarizado. Esse limite é calculado em função de parâmetros como o número de sensores do arranjo, o número de sinais, a separação entre as fontes, a SNR, dentre outros. A expressão para o modelo de sinal condicional está na seção 2.7 e uma explicação mais detalhada sobre o tema pode ser encontrada em (STOICA; NEHORAI, 1990; VAN TREES, 2002).

A maneira mais comum de comparar o desempenho de estimação é pela avaliação da **curva de desempenho**. Essa curva é calculada a partir da Raiz do Erro Quadrático Médio (REQM) do conjunto das estimativas geradas a partir de  $E$  simulações de Monte-Carlo para cada SNR. A REQM para cada estimador é calculada como mostrado na

Equação (2.24).

$$\begin{aligned} \text{REQM} &= \sqrt{\text{EQM}} \\ &= \sqrt{\frac{1}{ME} \sum_{e=1}^E \sum_{m=1}^M (\theta_m - \hat{\theta}_m)^2} \end{aligned} \quad (2.24)$$

em que:  $\hat{\theta}_m$  é a estimativa do ângulo DOA  $\theta_m$  para a  $m$ -ésima fonte de sinal.

A curva de desempenho de um estimador *assintoticamente eficiente* decresce monotonicamente quando aumenta o número de *snapshots* ou a SNR (ALVES, 2004). Dessa forma, é desejável que essa curva decresça continuamente à medida que esses valores aumentam.

Por outro lado, quando a SNR e o número de *snapshots* diminuem até um determinado valor-crítico, o estimador falha consistentemente e a REQM aumenta de maneira considerável. Esse ponto é chamado de limiar de estimação. Como este trabalho avalia o comportamento assintótico dos estimadores em função da SNR, adota-se a **SNR de limiar**.

No entanto, é importante ressaltar que este trabalho avalia os estimadores em função da **separação das fontes de sinal**. Assim, cada estimador é avaliado a partir da **superfície de desempenho** formada pelo conjunto das curvas de desempenho associadas a cada valor de separação das fontes de sinal.

Os estimadores são avaliados em função do **fator de correlação** entre as fontes de sinal. Sabe-se que alguns estimadores não apresentam bom desempenho quando existem sinais coerentes, embora não seja o que ocorre com nenhum estimador estudado neste trabalho. Adicionalmente, avalia-se o **esforço computacional** de cada estimador, que deve ser o mais baixo possível, pois existem aplicações onde a estimação DOA é realizada em tempo real.

Em suma, um estimador ideal deve gerar estimativas não-polarizadas no menor esforço computacional possível. Ele também deve apresentar erro quadrático médio próximo do CRB quando a SNR aumenta, independentemente do número de sinais ou do fator de correlação entre eles. Além disso, ele deve apresentar a menor SNR de limiar possível em todos esses casos.

## 2.3 O Estimador de Máxima-Verossimilhança

O MLE provê uma estratégia sistemática a vários problemas de estimação de parâmetros (CRAMÉR, 1946) e pode ser usado em problemas considerados difíceis (KAY, 1993).

A distribuição de probabilidades do vetor de amostras  $\mathbf{y}[n]$  é expressa como uma função dos parâmetros no modelo de sinal e a função de verossimilhança  $L(\boldsymbol{\omega})$  é maximizada

em relação a esses parâmetros (OTTERSTEN, 1989).

O MLE consiste na estimação do vetor de frequências angulares  $\boldsymbol{\omega} \in \boldsymbol{\Omega}$  que maximiza a função de verossimilhança  $L(\boldsymbol{\omega})$  do vetor de amostras  $\mathbf{y}[n]$ .  $L(\boldsymbol{\omega})$  é definido como a função densidade de probabilidade  $p^N(\mathbf{y}[n]; \boldsymbol{\omega})$  dos  $N$  *snapshots*  $\mathbf{y}[n]$ , vista como uma função do vetor desconhecido de parâmetros  $\boldsymbol{\omega}$ . Uma vez obtida uma amostra específica,  $\mathbf{y}[n]$  torna-se fixo. Assim, a função densidade de probabilidade conjunta pode ser interpretada como uma função do vetor desconhecido de parâmetros, chamada função de verossimilhança  $L(\boldsymbol{\omega})$ , em que  $\boldsymbol{\omega}$  se torna variável (SILVA, 2007).

Assim, a expressão (ou função-objetivo) do MLE é calculada a partir de  $L(\boldsymbol{\omega})$  (ALVES, 2004; SILVA, 2007). O objetivo é determinar o vetor de estimativas  $\boldsymbol{\omega}$  em que  $L(\boldsymbol{\omega})$  possui o maior valor possível (BRESLER; MACOVSKI, 1986). Essa função-objetivo é calculada a partir da função erro quadrático médio (KUMARESAN; SCHARF; SHAW, 1986) e gera uma estimativa da amplitude complexa do sinal.

O MLE pode ser considerado um método quase-ótimo, pois ele possui um desempenho de estimação superior, além de prover estimativas assintoticamente não-polarizadas e eficientes. Por outro lado, a função de verossimilhança é complexa, multimodal, multivariada e altamente não-linear. Assim, é muito difícil de ser resolvida eficientemente (LI; LU, 2007). Então, o MLE demanda elevado esforço computacional, razão pela qual a sua implementação direta não é uma alternativa comumente utilizada (ALVES, 2004). Isso é devido à maneira pela qual é gerado o vetor de estimativas, exigindo busca  $M$ -dimensional pelo ótimo global. Na estimação DOA, esse ponto está associado ao mínimo global da função-objetivo.

As próximas seções deste capítulo mostram algumas expressões do MLE para estimação DOA considerando os dois modelos de sinal apresentados anteriormente. Esses modelos são importantes no cálculo dos critérios determinístico e estocástico, associados aos modelos de sinal condicional e incondicional, respectivamente (STARER, 1990; OTTERSTEN, 1989). Por outro lado, é importante ressaltar que qualquer um desses critérios pode ser usado independentemente do modelo de sinal considerado (STARER, 1990).

Adicionalmente, nenhum dos dois critérios é limitado à estimação de sinais não-coerentes, como ocorre com outros métodos de estimação (ALVES, 2004). Entretanto, nessa condição o SMLE é conhecido por gerar sempre estimativas mais precisas do que o DMLE, mas apresenta esforço computacional ainda maior (STARER, 1990). Além disso, o uso do DMLE é mais difundido na literatura. Adicionalmente, o DMLE é utilizado no método de estimação e nos filtros espaciais descritos nos próximos capítulos.

## 2.4 O Estimador de Máxima-Verossimilhança Determinístico

Para obter o DMLE, considera-se que os sinais estão de acordo com o modelo condicional, em que  $\mathbf{s}[n]$  é um processo aleatório gaussiano ao longo do tempo e do espaço. No entanto, ao longo dos *snapshots*,  $\mathbf{s}[n]$  é mantido constante e é alterado apenas o ruído aditivo  $\mathbf{n}[n]$ . O vetor  $\mathbf{n}[n]$  também é um processo aleatório gaussiano. Assim,  $\mathbf{y}[n]$  é um processo aleatório gaussiano, com média  $\mathbf{A}\mathbf{s}[n]$  e matriz de correlação espacial  $\sigma^2\mathbf{I}$ . Considerando isso, a função de verossimilhança é apresentada na Equação (2.25) (KRIM; VIBERG, 1996).

$$L_{\text{DMLE}}(\boldsymbol{\omega}, \mathbf{s}[n], \sigma^2) = \prod_{n=1}^N (\pi\sigma^2)^{-M} e^{-\|\mathbf{y}[n] - \mathbf{A}\mathbf{s}[n]\|^2/\sigma^2} \quad (2.25)$$

em que: o operador  $\|\cdot\|$  é a norma euclidiana.

Por conveniência, as estimativas ML são definidas como os argumentos que minimizam a Equação (2.25), como mostrado na Equação (2.26) (ALVES, 2004).

$$\bar{L}_{\text{DMLE}}(\boldsymbol{\omega}, \mathbf{s}[n], \sigma^2) = -\log [L_{\text{DMLE}}(\boldsymbol{\omega}, \mathbf{s}[n], \sigma^2)] \quad (2.26)$$

Após algumas simplificações e desconsiderando o termo constante, obtém-se a Equação (2.27).

$$\bar{L}_{\text{DMLE}}(\boldsymbol{\omega}, \mathbf{s}[n], \sigma^2) = M \log(\sigma^2) + \frac{1}{\sigma^2 N} \sum_{n=1}^N \|\mathbf{y}[n] - \mathbf{A}\mathbf{s}[n]\|^2 \quad (2.27)$$

em que as estimativas ML de  $\boldsymbol{\omega}$ ,  $\mathbf{s}[n]$  e  $\sigma^2$  são os valores que minimizam a Equação (2.27).

Para proceder com essa minimização, normalmente fixa-se  $\boldsymbol{\omega}$ , pois todos os parâmetros são independentes entre si. Em seguida, os valores mínimos são substituídos na Equação (2.27) e a minimização é realizada em  $\mathbf{s}[n]$ , obtendo-se as Equações (2.28) e (2.29) (KRIM; VIBERG, 1996).

$$\hat{\sigma}^2 = \frac{1}{K} \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}) \quad (2.28)$$

$$\hat{\mathbf{s}}[n] = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{y}[n] \quad (2.29)$$

em que:  $\text{tr}(\cdot)$  é o operador traço matricial;  $\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{P}_{\mathbf{A}} = \mathbf{I} - \mathbf{A}\mathbf{A}^+ \in \mathbb{C}^{K \times K}$  é um operador que projeta um vetor no espaço ortogonal ao espaço das colunas de  $\mathbf{A}$ . Finalmente,  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$  é a pseudo-inversa (ou a inversa generalizada) de Moore-Penrose.

Fazendo algumas substituições e observando que  $\frac{1}{N} \sum_{n=1}^N \|\mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{y}[n]\|^2 = \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}})$ , o vetor de estimativas das frequências angulares  $\hat{\boldsymbol{\omega}}$  é obtido a partir do problema de minimização apresentado na Equação (2.30) (KRIM; VIBERG, 1996).

$$\hat{\boldsymbol{\omega}} = \arg \min_{\boldsymbol{\omega}} [F_{\text{DMLE}}(\boldsymbol{\omega})] \quad (2.30)$$

em que:  $F_{\text{DMLE}}(\boldsymbol{\omega})$  é a função-objetivo do DMLE, como mostrado na Equação (2.31).

$$F_{\text{DMLE}}(\boldsymbol{\omega}) = \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}) \quad (2.31)$$

## 2.5 O Estimador de Máxima-Verossimilhança Estocástico

Para obter o SMLE, considera-se que os sinais são representados pelo modelo incondicional. Assim, os sinais e o ruído são processos aleatórios gaussianos ao longo do tempo e do espaço. Partindo desse pressuposto, a função de verossimilhança é apresentada na Equação (2.32).

$$L_{\text{SMLE}}(\boldsymbol{\omega}, \mathbf{C}, \sigma^2) = \prod_{n=1}^N \pi^{-K} |\mathbf{R}|^{-K} e^{-\mathbf{y}^T[n] \mathbf{R}^{-1} \mathbf{y}[n]} \quad (2.32)$$

Realizando o mesmo procedimento feito na seção 2.4 e sabendo que todos os parâmetros de  $L_{\text{SMLE}}$  são independentes, fixa-se  $\boldsymbol{\omega}$  de forma que as estimativas da matriz de correlação dos sinais  $\mathbf{C}$  e  $\sigma^2$  sejam calculadas como mostrado nas Equações (2.33) e (2.34) (KRIM; VIBERG, 1996).

$$\hat{\sigma}_{\text{SMLE}}^2 = \frac{1}{K - M} \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}) \quad (2.33)$$

$$\hat{\mathbf{C}}_{\text{SMLE}}(\boldsymbol{\omega}) = \left[ (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \right] (\hat{\mathbf{R}} - \hat{\sigma}_{\text{SMLE}}^2 \mathbf{I}) \left[ (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \right]^H \quad (2.34)$$

Substituindo as Equações (2.33) e (2.34) na função de verossimilhança e após algumas simplificações, as estimativas das frequências angulares podem ser calculadas como mostrado na Equação (2.35).

$$\hat{\boldsymbol{\omega}} = \arg \min_{\boldsymbol{\omega}} [F_{\text{SMLE}}(\boldsymbol{\omega})] \quad (2.35)$$

em que:  $F_{\text{SMLE}}(\boldsymbol{\omega})$  é a função-objetivo do SMLE, mostrada na Equação (2.36).

$$F_{\text{SMLE}}(\boldsymbol{\omega}) = \log_{10} \left[ \det \left( \mathbf{A} \hat{\mathbf{C}}_{\text{SMLE}}(\boldsymbol{\omega}) \mathbf{A}^H + \hat{\sigma}_{\text{SMLE}}^2 \mathbf{I} \right) \right] \quad (2.36)$$

## 2.6 Simulações Numéricas

Nesta seção são apresentadas algumas superfícies usando as funções-objetivo do DMLE e do SMLE. Essas superfícies foram calculadas utilizando os seguintes parâmetros:

- Arranjo Linear Uniforme formado por  $K = 10$  sensores espaçados em meio comprimento de onda, obtendo  $N = 100$  *snapshots* em cada experimento;
- Sinais descorrelacionados e de mesma potência, de acordo com o modelo de sinal incondicional;
- Fontes de sinal próximas entre si, com  $\theta_1 = 10^\circ$  e  $\theta_2 = 15^\circ$ . É importante observar que essa situação é bastante citada na literatura para avaliar o desempenho de estimadores de alta resolução.



Primeiramente, as Figuras 2 e 3 ilustram respectivamente as superfícies do DMLE e do SMLE para o mesmo experimento na ausência de ruído.

Figura 2 – Superfície do DMLE para fontes de sinal situadas em  $10^\circ$  e  $15^\circ$  e sinais descorrelacionados na ausência de ruído.

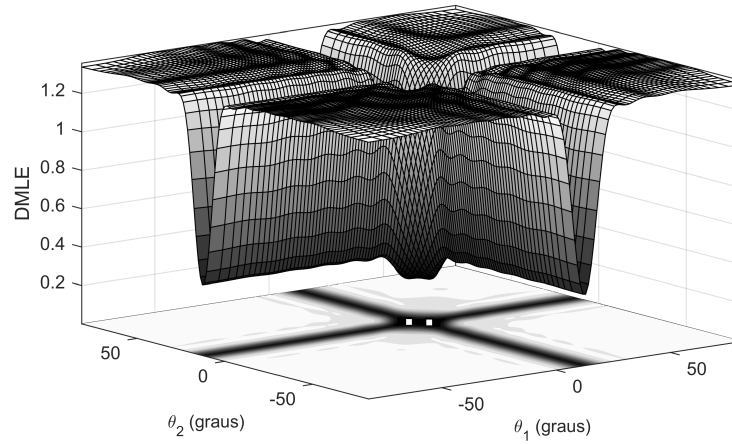
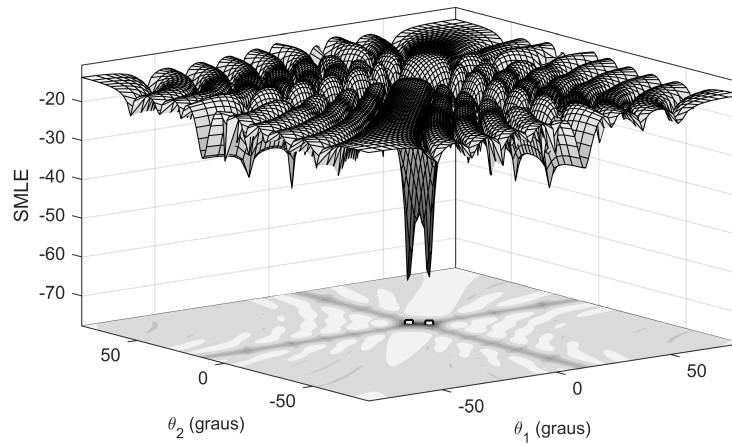


Figura 3 – Superfície do SMLE (Stochastic Maximum Likelihood Estimation) para fontes de sinal situadas em  $10^\circ$  e  $15^\circ$  e sinais descorrelacionados na ausência de ruído.

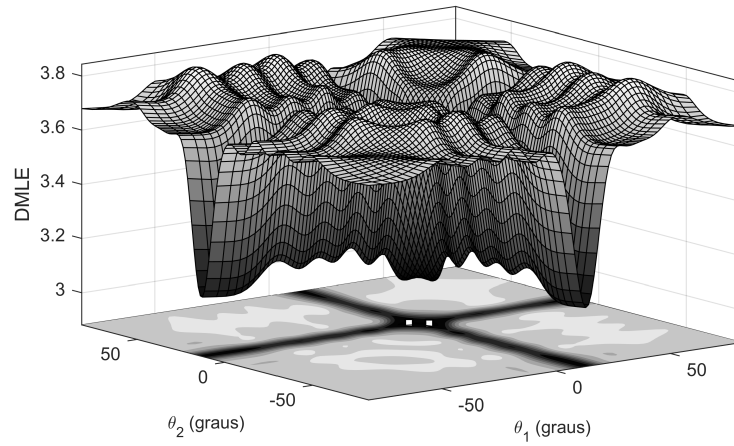


Pode-se observar na Figura 2 que as estimativas DOA segundo o critério determinístico são as coordenadas retangulares associadas aos mínimos globais. Como existe simetria na superfície, qualquer um dos dois mínimos globais representa a solução completa para o problema. Na ausência de ruído, estes pontos coincidem com as posições reais das fontes de sinal, que estão representadas por um quadrado na cor branca na parte inferior do gráfico.

Além disso, pode-se verificar na Figura 3 que o critério estocástico gera estimativas muito mais precisas do que o critério determinístico. Porém, o SMLE demanda maior esforço computacional (ALVES, 2004).

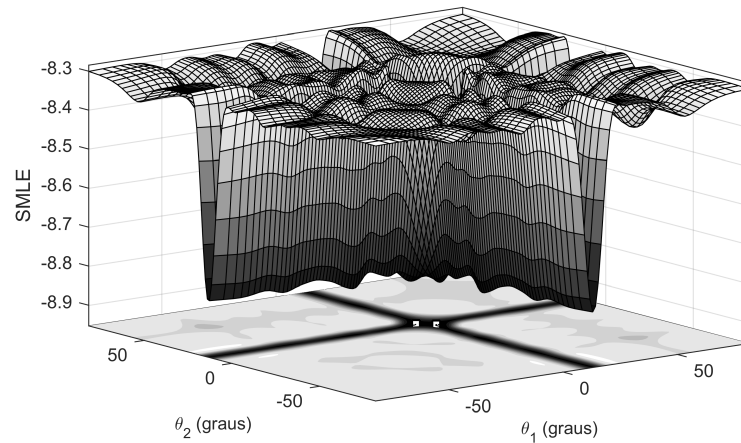
A Figura 4 mostra a superfície do DMLE na SNR de  $-10$  dB (decibéis). Pode-se observar nessa figura o aparecimento de um mínimo global distante das posições reais das fontes de sinal. Adicionalmente, os mínimos associados a essas posições foram ligeiramente deslocados. Assim, pode-se inferir que, dependendo de como ocorre a otimização numérica, o estimador pode fornecer estimativas DOA completamente imprecisas na pior situação. Na melhor situação, mesmo que a otimização convirja para os mínimos associados às posições reais, o estimador pode fornecer estimativas DOA com um pequeno erro de estimação.

Figura 4 – Superfície do DMLE para fontes de sinal situadas em  $10^\circ$  e  $15^\circ$  e sinais descorrelacionados na SNR de  $-10$  dB.



A Figura 5 ilustra a superfície do SMLE na SNR de  $-10$  dB. Pode-se verificar nessa figura que o SMLE não oferece mais ganho significativo em relação ao DMLE, pois a sua precisão diminuiu significativamente. Além disso, o SMLE foi perturbado pelo ruído de maneira considerável.

Figura 5 – Superfície do SMLE para fontes de sinal situadas em  $10^\circ$  e  $15^\circ$  e sinais descorrelacionados na SNR de  $-10$  dB.



Adicionalmente, é importante recordar que ambos os critérios são baseados nos

modelos de sinal apresentados neste capítulo. Assim, à medida que o nível de ruído aumenta, o sinal amostrado se assemelha cada vez menos a esses modelos. Consequentemente, quanto maior é o nível de ruído, mais difícil se torna a estimação DOA, com o deslocamento da posição dos mínimos globais ou ainda com o aparecimento de *outros* mínimos globais em posições distintas.

## 2.7 Limite de Cramér-Rao para o DMLE

O CRB foi apresentado na seção 2.2 como um limitante inferior para a variância do erro de estimação de um estimador não-polarizado.

As expressões para o Limite de Cramér-Rao são obtidas a partir do vetor de dados  $\mathbf{y}[n]$  para qualquer um dos seguintes parâmetros: amplitude complexa, frequência ou potência do ruído. Neste trabalho, o CRB utilizado é específico para a estimação de frequência (ALVES, 2004) devido à equivalência entre a estimação de frequência e a estimação DOA observada na seção 2.1. Os CRBs para o DMLE e para o SMLE foram apresentados em (STOICA; NEHORAI, 1990). Para o DMLE, o CRB é calculado pela Equação (2.37).

$$\mathbf{O}_{\text{DMLE}} = \frac{\sigma^2}{2N} [\Re(\mathbf{Q} \otimes \mathbf{C}^T)]^{-1} \{ \Re[\mathbf{Q} \otimes (\mathbf{C}\mathbf{M}\mathbf{C})^T] \} [\Re(\mathbf{Q} \otimes \mathbf{C}^T)]^{-1} \quad (2.37)$$

em que:  $\mathbf{O}_{\text{DMLE}} \in \mathbb{R}^{M \times M}$  é medido em  $\text{rad}^2$ ;  $\Re(\cdot)$  é o operador parte real;  $\otimes$  é o operador de multiplicação matricial ponto-a-ponto;  $\mathbf{Q} = \mathbf{D}^H \mathbf{P}_A^\perp \mathbf{D}$ ;  $\mathbf{M} = \mathbf{C}^{-1} + \sigma^2 \mathbf{C}^{-1} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{C}^{-1}$  e a matriz  $\mathbf{D}$  é a derivada de  $\mathbf{A}$  e é calculada como mostrado na Equação (2.38).

$$\mathbf{D} = \begin{bmatrix} \frac{\partial \mathbf{a}(\omega_1)}{\partial \omega_1} & \frac{\partial \mathbf{a}(\omega_2)}{\partial \omega_2} & \dots & \frac{\partial \mathbf{a}(\omega_{\bar{M}})}{\partial \omega_{\bar{M}}} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & \dots & 0 \\ j e^{j\omega_1} & j e^{j\omega_2} & \dots & j e^{j\omega_{\bar{M}}} \\ j 2 e^{j2\omega_1} & j 2 e^{j2\omega_2} & \dots & j 2 e^{j2\omega_{\bar{M}}} \\ \vdots & \vdots & \ddots & \vdots \\ j(K-1)e^{j(K-1)\omega_1} & j(K-1)e^{j(K-1)\omega_2} & \dots & j(K-1)e^{j(K-1)\omega_{\bar{M}}} \end{bmatrix} \quad (2.38)$$

O CRB calculado na Equação (2.37) considera  $N \rightarrow \infty$ . Assim, é uma aproximação assintótica. Por outro lado, o cálculo é mais simples e direto, podendo ser usado quando  $\mathbf{C}$  se aproxima da estimativa  $\hat{\mathbf{C}} = \frac{1}{N} \sum_{n=1}^N \mathbf{s}[n] \mathbf{s}[n]^H$ . É importante ressaltar que o CRB é o valor de todos os elementos na diagonal principal de  $\mathbf{O}_{\text{DMLE}}$  e é igual para os  $M$  sinais (ALVES, 2004).

## 2.8 Considerações Finais

Este capítulo apresentou os modelos de sinal, os critérios de avaliação de estimadores DOA e o Estimador de Máxima-Verossimilhança.

Os modelos de sinal são simplificados ao considerar um arranjo linear uniforme e ao considerar que as fontes de sinal faixa estreita estão em um campo distante. A partir desse modelo são calculados dois critérios do Estimador de Máxima-Verossimilhança: o DMLE e o SMLE, baseados nos modelos de sinal condicional e incondicional, respectivamente.

De maneira independente do critério utilizado, o MLE é um estimador assintoticamente eficiente em relação à SNR e ao número de *snapshots*, desde que o número de sensores do arranjo seja suficientemente maior do que o número de fontes (STOICA; NEHORAI, 1990). Adicionalmente, o SMLE produz estimativas mais precisas do que o DMLE, especialmente na estimação DOA de sinais coerentes. Ainda, o MLE possui aplicação limitada, pois o esforço computacional é elevado e ainda é maior para o critério estocástico.

Quando a SNR ou o número de *snapshots* é suficientemente grande, o MLE possui um mínimo global nos pontos que representam as combinações dos ângulos DOA reais das fontes de sinal. Também, em condições mais severas de SNR, ocorre o aparecimento de outros mínimos globais e o deslocamento dos mínimos associados às posições verdadeiras. Assim, o MLE não é um estimador confiável quando atinge a SNR de limiar, pois ele não consegue estimar corretamente os ângulos DOA. Quando isso ocorre, o Estimador de Máxima-Verossimilhança apresenta rápido aumento da variância do erro de estimação.

No entanto, especialmente quando a SNR é suficientemente alta, não é necessário buscar em todo o espaço de soluções do MLE para obter estimativas precisas. Quando isso é possível, as estimativas podem ser geradas de maneira mais eficiente com o uso de métodos iterativos.

Alguns desses métodos, baseados em uma reparametrização do DMLE, diminuem drasticamente o esforço computacional, enquanto ainda mantêm as qualidades de estimação do MLE. Por outro lado, eles podem ser ainda mais sensíveis ao ruído e apresentam SNRs de limiar razoavelmente maiores. Entretanto, para diminuí-las, pode-se usar um processo de filtragem dos dados de entrada.

O próximo capítulo apresenta os estimadores iterativos MODEX Modificado e SEAD Aprimorado. Além disso, é apresentado um processo de filtragem usando o máximo autofiltro espacial, com o objetivo de maximizar a SNR de saída, permitindo diminuir a influência do ruído na geração de estimativas DOA.

### 3 ESTIMADORES ITERATIVOS E FILTRAGEM ESPACIAL

O Estimador de Máxima-Verossimilhança apresentado na seção 2.3 possui grande capacidade de resolução, mas sua implementação direta exige elevado esforço computacional. Devido a isso, foram desenvolvidos estimadores iterativos derivados do DMLE. Esses estimadores apresentam desempenho de estimação muito próximo ao do DMLE, porém possuem um esforço computacional consideravelmente menor.

Dentre os estimadores clássicos derivados do DMLE, pode-se citar o *Method of Direction Estimation* (MODE) (STOICA; SHARMAN, 1990). O MODE aplica uma reparametrização da função-objetivo do DMLE de forma que a otimização é realizada sobre os coeficientes de um polinômio em que suas raízes estimam os ângulos DOA. Porém, sabe-se que o MODE apresenta o *efeito de limiar* (GERSHMAN; STOICA, 1999). Esse efeito é caracterizado por uma rápida degradação no desempenho de estimação quando a SNR ou o número de *snapshots* caem abaixo de um valor crítico. A razão disso é uma descontinuidade nas estimativas dos parâmetros que são funções dos dados processados. Esse fato ocorre tipicamente em todos os estimadores baseados em subespaço e é induzido pela divisão do espaço de medida em subespaços de sinal e de ruído (HAWKES; NEHORAI; STOICA, 2001). Adicionalmente, o efeito de limiar pode ser causado pelo fato dos autovetores do subespaço de sinal amostrado terem componentes significativos no subespaço de ruído (GERSHMAN; STOICA, 1999).

Diante disso, foi proposto o *MODE with Extra Roots* (MODEX) (GERSHMAN; STOICA, 1999). Esse estimador é capaz de melhorar significativamente o desempenho de estimação em condições como baixas SNRs ou poucas amostras de sinal. Isso foi possível com o aumento da ordem do polinômio usado no MODE. Por outro lado, passa a ser necessário um procedimento ML adicional para selecionar a melhor combinação de raízes. Dessa forma, o MODEX exige esforço computacional consideravelmente maior do que o do MODE, especialmente quando é grande o número de fontes de sinal.

Porém, o MODEX Modificado (LOPES et al., 2003; ALVES, 2004) foi proposto após ter sido observado que as raízes extras geradas pelo MODEX não são em geral boas estimativas DOA. Assim, o MODEX Modificado gera três estimativas MODE, cada uma baseada em uma restrição de não-trivialidade distinta. Então, as estimativas são combinadas e a melhor combinação é selecionada aplicando o mesmo procedimento ML usado no MODEX.

Por outro lado, a proposta do *Search of Direction by Differential Spectrum* (SEAD) (FERREIRA, 2005; FERREIRA; LEMOS, 2005a; FERREIRA; LEMOS, 2005b) foi moti-

vada pela determinação de um espectro mais robusto aos efeitos do ruído sobre a estimação DOA. A principal contribuição desse estimador foi a proposta de um estágio de busca angular em toda a abertura do arranjo. Essa busca é denominada espectro diferencial e é baseada na diferença entre os dois maiores autovalores de uma matriz de correlação espacial aumentada.

Embora tenha mostrado baixa capacidade de resolução, o espectro diferencial apresenta uma robustez razoável aos efeitos do ruído. Nesse espectro, os picos espúrios associados ao ruído são significativamente menores do que os picos correspondentes às fontes de sinal, mesmo em SNRs de até  $-10$  dB. Devido à sua baixa resolução (SILVA et al., 2013), é necessário um estágio de geração do vetor de estimativas iniciais. Logo após, o SEAD realiza um estágio de refinamento desse vetor ao redor dos picos mais significativos do espectro diferencial para fornecer estimativas DOA usando um procedimento ML. Entretanto, devido ao grande número de estimativas geradas, o SEAD apresentou tempo de execução incompatível com aplicações em tempo real. Como apresentado no Capítulo 1, o arranjo de sensores em um equipamento compatível com o WiMAX é recalibrado a cada cinco milissegundos.

Então, o SEAD Aprimorado (SILVA, 2009; SILVA et al., 2013) foi proposto com o objetivo de diminuir esse tempo de estimação. Para isso, foi utilizado um processo de seleção dividido em dois estágios. O primeiro estima a quantidade de fontes agrupadas em cada pico de sinal e o segundo estágio usa um processo de busca baseado no algoritmo *Branch-and-Bound*. O resultado foi uma redução considerável no número de estimativas geradas, mantendo competitivo o desempenho de estimação em relação ao SEAD.

Entretanto, foi observado em (KRUMMENAUER, 2007; KRUMMENAUER et al., 2010) que o procedimento ML utilizado no processo de seleção de raízes do MODEX Modificado nem sempre seleciona as melhores estimativas geradas. Isso é devido ao ruído presente em  $\hat{\mathbf{R}}$  provocar mudança significativa nesse processo de busca e é mais evidente à medida que a SNR diminui. Com isso, o erro de estimação aumenta de maneira considerável e, conseqüentemente, a SNR de limiar desse estimador é maior.

Assim, o desempenho de estimação do MODEX Modificado pode ser significativamente aprimorado ao reduzir o componente de ruído da matriz de covariância utilizada no procedimento ML. Uma maneira para fazer isso é submeter os dados da saída do arranjo  $\mathbf{y}[n]$  a um filtro espacial (FORSTER; VEZZOSI, 1987) adequado antes de realizar esse procedimento, lembrando que conceitos de estimação de frequência podem ser usados na estimação DOA.

Porém, para projetar esse filtro linear com resposta ao impulso finita (FIR, do inglês *Finite Impulse Response*) é necessário ter estimativas preliminares das frequências angulares. Em (KRUMMENAUER et al., 2010), essas estimativas são provenientes do máximo autovetor de uma matriz calculada a partir de  $\hat{\mathbf{R}}$ , ou seja, são provenientes do



autovetor associado ao máximo autovalor dessa matriz. Adicionalmente, é necessária uma nova função-objetivo para o DMLE que aplica esse filtro na geração de estimativas.

Este capítulo descreve a fundamentação teórica dos estimadores MODEX Modificado e SEAD Aprimorado. Além disso, é apresentado o processo de filtragem utilizando o máximo autofiltro espacial projetado considerando a situação de fontes próximas. Finalmente, são realizadas as considerações finais deste capítulo.

### 3.1 O Estimador MODEX Modificado

Esta seção explica o processo de geração de estimativas do MODEX Modificado. Para isso, é necessário apresentar o estimador MODE e a reparametrização da função-objetivo do DMLE que o fundamenta.

### 3.2 Reparametrização da Função-Objetivo do DMLE

A estimação de máxima-verossimilhança de  $\omega$  apresentada na Equação (2.30) é um problema não-linear considerado difícil. Porém, uma versão um pouco mais simples desse problema é gerada ao reparametrizá-lo em função do polinômio de predição linear  $b(z)$  mostrado na Equação (3.1) (STARER, 1990).

$$b(z) \triangleq \sum_{m=0}^M b_m z^{(M-m)} = b_0 \prod_{m=1}^M (z - e^{j\omega_m}) = b_0 z^M + b_1 z^{(M-1)} + \dots + b_M \quad (3.1)$$

em que:  $z = e^{j\omega_m}$ ,  $m = 1, 2, \dots, M$ , são as raízes do polinômio da Equação (3.1) e todas estão situadas sobre a circunferência de raio unitário (CRU), resultando na Equação (3.2) (SILVA, 2007).

$$\sum_{m=0}^M b_m e^{j(k-m)\omega_i} = 0 \quad (3.2)$$

em que:  $k = 0, 1, \dots, (K-1)$  e  $i = 1, 2, \dots, M$ .

Considere o vetor  $\mathbf{b} = [b_0 \ b_1 \ \dots \ b_M]^T$ , que contém os coeficientes do polinômio  $b(z)$  mostrado na Equação (3.1), e a matriz de predição linear  $\mathbf{B}$  apresentada na Equação (3.3).

$$\mathbf{B} = \begin{bmatrix} b_M & \dots & b_1 & b_0 & \mathbf{0} \\ & \ddots & & \ddots & \ddots \\ \mathbf{0} & & b_M & \dots & b_1 & b_0 \end{bmatrix}^H \quad (3.3)$$

Assim, a Equação (3.2) pode ser reescrita como mostrado na Equação (3.4) (KUMAR-RESAN; SCHARF; SHAW, 1986).

$$\mathbf{B}^H \mathbf{A} = \mathbf{0} \quad (3.4)$$

Considerando  $\text{posto}(\mathbf{A}) = \bar{M}$  e  $\text{posto}(\mathbf{B}) = (K - \bar{M})$ , tem-se que as colunas de  $\mathbf{B}$  pertencem ao espaço nulo de  $\mathbf{A}^H$  (BRESLER; MACOVSKI, 1986). Em outras palavras,  $\mathbf{B}$  abrange o espaço nulo de  $\mathbf{A}$ , permitindo que sejam escritas as Equações (3.5) e (3.6) (ALVES, 2004).

$$\mathbf{P}_B = \mathbf{B}\mathbf{B}^+ = \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H = \mathbf{I} - \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H = \mathbf{I} - \mathbf{A}\mathbf{A}^+ = \mathbf{P}_A^\perp \quad (3.5)$$

$$\mathbf{P}_A = \mathbf{A}\mathbf{A}^+ = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H = \mathbf{I} - \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H = \mathbf{I} - \mathbf{B}\mathbf{B}^+ = \mathbf{P}_B^\perp \quad (3.6)$$

em que:  $\mathbf{P}_B$  e  $\mathbf{P}_A^\perp$  são projetores ortogonais ao espaço das colunas de  $\mathbf{A}$ , resultando nas relações apresentadas nas Equações (3.7) e (3.8) (SILVA, 2007).

$$\mathbf{P}_B = \mathbf{P}_A^\perp \quad (3.7)$$

$$\mathbf{P}_A = \mathbf{P}_B^\perp \quad (3.8)$$

Dessa forma, o vetor de estimativas ML é obtido resolvendo o problema de otimização linear restrita apresentado na Equação (3.9) (STARER, 1990).

$$\hat{\mathbf{b}} = \arg \min_{\mathbf{b} \in D} [F_{\text{DMLE2}}(\mathbf{b})], \quad (3.9)$$

em que:  $F_{\text{DMLE2}}(\mathbf{b})$  é a função-objetivo reparametrizada do DMLE em termos de  $\mathbf{b}$ . Essa função é mostrada na Equação (3.10).

$$\begin{aligned} F_{\text{DMLE2}}(\mathbf{b}) &= \text{tr} \left[ \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H\hat{\mathbf{R}} \right] \\ &= \text{tr} (\mathbf{P}_B\hat{\mathbf{R}}) \end{aligned} \quad (3.10)$$

Adicionalmente,  $D$  representa o conjunto das restrições impostas ao parâmetro  $\mathbf{b}$ . De maneira simplificada, essas restrições são (ALVES, 2004):

- **$\mathbf{b} \neq \mathbf{0}$** : essa restrição de não-trivialidade garante que existe correspondência única entre  $\mathbf{b}$  e o vetor das frequências angulares  $\boldsymbol{\omega}$  e a invertibilidade de  $\mathbf{B}^H\mathbf{B}$  (BRESLER; MACOVSKI, 1986). Essa restrição pode ser aplicada ao impor que um dos elementos de  $\mathbf{b}$  seja igual a um ou ao impor que  $\mathbf{b}$  tenha norma unitária;
- **raízes de  $b(z)$  sobre a CRU**: essa restrição garante que o sinal é composto de exponenciais complexas não-amortecidas. Em (KUMARESAN; SCHARF; SHAW, 1986) foi avaliado que isso equivale a impor a restrição de simetria conjugada  $\{\mathbf{b}|b_M = b_{(M-m)}^*, m = 0, 1, \dots, M\}$ , em que:  $(\cdot)^*$  é o operador de conjugação complexa. Entretanto, essa restrição embora seja *necessária*, ela não é *suficiente* para garantir que as raízes de  $b(z)$  estejam sobre a CRU (STOICA; SHARMAN, 1990). Por outro lado, todo o polinômio  $b(z)$  com raízes sobre a CRU pode ser convertido em outro que possui simetria complexa conjugada mediante um fator multiplicativo (SILVA, 2007).



Assim como a Equação (2.30), a Equação (3.9) também é um problema não-linear, que impossibilita que seja obtida uma solução analítica para  $\hat{\mathbf{b}}$ . Entretanto, essa equação é mais fácil de ser manipulada algebricamente e é a base dos estimadores MODE e MODEX Modificado. Ao término da minimização, as raízes do polinômio descrito em  $\hat{\mathbf{b}}$  são extraídas para obter o vetor de estimativas das frequências angulares  $\hat{\boldsymbol{\omega}}$  associadas às estimativas DOA  $\hat{\boldsymbol{\theta}}$  (SILVA, 2007).

### 3.2.1 O estimador MODE

O MODE realiza a estimação DOA ao aproximar a função-objetivo do DMLE localmente em torno do vetor ótimo  $\mathbf{b}_o$  por uma função quadrática (ALVES, 2004). Isso é feito aplicando a decomposição em autovalores realizada na Equação (2.22) e considerando a assintoticidade em  $N$ , ou seja, um grande número de amostras de sinal.

Suponha que todos os autovetores do subespaço de ruído em  $\hat{\mathbf{E}}_N$  correspondam ao espaço nulo de  $\mathbf{A}$ . Quando isso acontece, todos os autovetores do subespaço de sinal em  $\hat{\mathbf{E}}_S$  descrevem o espaço de  $\mathbf{A}$ . Isso permite escrever as Equações (3.11) e (3.12) (SILVA, 2007).

$$\mathbf{P}_A = \hat{\mathbf{E}}_S \hat{\mathbf{E}}_S^H \quad (3.11)$$

$$\mathbf{P}_A^\perp = \hat{\mathbf{E}}_N \hat{\mathbf{E}}_N^H \quad (3.12)$$

O interesse pelo MODE está no comportamento local da Equação (3.10). Quando  $\mathbf{b}$  está próximo do vetor ótimo  $\mathbf{b}_o$ , naturalmente  $\|\mathbf{b} - \mathbf{b}_o\| \approx 0$ . Considerando que  $\mathbf{B}_o$  é construído a partir de  $\mathbf{b}_o$  da mesma maneira descrita na Equação (3.3),  $\|\mathbf{B} - \mathbf{B}_o\| \approx 0$ . Consequentemente,  $\mathbf{B}^H \hat{\mathbf{E}}_S = \hat{\mathbf{E}}_S^H \mathbf{B} = \mathbf{0}$ . Então, o MODE faz a estimação DOA minimizando a expressão mostrada na Equação (3.13) (SILVA, 2007).

$$\hat{\mathbf{b}} = \arg \min_{\mathbf{b} \in D} [F_{\text{MODE}}(\mathbf{b})], \quad (3.13)$$

em que:  $F_{\text{MODE}}(\mathbf{b})$  é a função-objetivo do MODE apresentada na Equação (3.14).

$$\begin{aligned} F_{\text{MODE}}(\mathbf{b}) &= \text{tr} \left[ \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \bar{\mathbf{R}} \right] \\ &= \text{tr} (\mathbf{P}_B \bar{\mathbf{R}}) \end{aligned} \quad (3.14)$$

em que:  $\bar{\mathbf{R}} = \hat{\mathbf{E}}_S \mathbf{W} \hat{\mathbf{E}}_S^H$  decorre da aplicação da restrição do subespaço de sinal em  $\hat{\mathbf{R}}$ ;  $\mathbf{W} = (\hat{\boldsymbol{\Lambda}}_S - \hat{\sigma}^2 \mathbf{I})^2 \hat{\boldsymbol{\Lambda}}_S^{-1} \in M \times M$  é a matriz ótima de ponderação com o objetivo de minimizar a variância do erro de estimação.

A minimização mostrada na Equação (3.13) é realizada iterativamente e são necessárias poucas iterações para garantir a convergência (GERSHMAN; STOICA, 1999). A primeira iteração utiliza  $\mathbf{B}^H \mathbf{B}$  inicializada como uma matriz definida-positiva, normalmente

considera-se  $\mathbf{B}^H \mathbf{B} = \mathbf{I}$ . Nas iterações seguintes,  $\hat{\mathbf{b}}$  passa a ser utilizado na atualização de  $\mathbf{B}^H \mathbf{B}$  (SILVA, 2007).

No fim desse processo, são extraídas as raízes do polinômio  $b(z)$  definido em  $\hat{\mathbf{b}}$  para obter as estimativas em termos das frequências angulares  $\hat{\omega}$ . Em sequência, o vetor  $\hat{\boldsymbol{\theta}}$  de estimativas DOA é calculado considerando a relação entre  $\omega_m$  e  $\theta_m$  mostrada na seção 2.1.

### 3.2.2 O estimador MODEX Modificado

Entretanto, quando a SNR é baixa ou o número de *snapshots* é pequeno, o MODE apresenta o efeito de limiar (GERSHMAN; STOICA, 1999). Isso ocorre quando um ou mais autovetores do subespaço de sinal possuem componentes significativos no subespaço de ruído. Nesse caso, o MODE tende a utilizar uma ou mais raízes de  $b(z)$  em regiões distantes das verdadeiras em resposta à forte intensidade do ruído no sinal amostrado (SILVA, 2007). Claramente, isso leva a uma considerável degradação no desempenho de estimação do MODE, pois não existe quantidade suficiente de raízes para serem associadas às fontes de sinal (GERSHMAN; STOICA, 1999).

Assim, foi proposto o MODEX (GERSHMAN; STOICA, 1999), que estende o MODE e usa  $X$  raízes extras com a finalidade de evitar esse fenômeno indesejado. Essas raízes permitem modelar os componentes de sinal e parte dos componentes de ruído. Essa extensão permitiu o MODEX gerar e avaliar novos pontos no espaço de soluções do DMLE, que possui diversos mínimos locais, como mostrado na seção 2.6. Então, o MODEX aprimorou de maneira significativa a estimação DOA em relação ao MODE, especialmente quando esse último apresenta o efeito de limiar.

Porém, nem todas as raízes extras do MODEX são adequadas para a estimação DOA. Isso ocorre, pois, *a priori*, o simples aumento da ordem do polinômio  $b(z)$  leva à produção de  $M$  raízes associadas ao subespaço de sinal próximas do mínimo global e  $X$  raízes associadas ao subespaço de ruído distantes do mínimo global. Em geral, esse último conjunto de raízes não contém boas estimativas DOA (LOPES et al., 2003).

Então, foi proposto o MODEX Modificado<sup>1</sup> (LOPES et al., 2003). Esse estimador é capaz de gerar raízes situadas mais próximas dos valores verdadeiros, gerando três vetores de estimativas equivalendo a  $3M$  raízes. Cada um desses vetores são calculados a partir de um vetor de estimativas MODE aplicando restrições diferentes para evitar a solução trivial  $\mathbf{b} = \mathbf{0}$  (LOPES et al., 2003).

Isso se deve à observação em (LOPES et al., 2003) de que, embora seja claro que o ruído perturbe as estimativas quando a SNR diminui, isso não ocorre da mesma maneira nos três vetores de estimativas. Assim, pressupõe-se que, dentre as  $3M$  estimativas calculadas, no mínimo  $M$  são de boa qualidade. Em outras palavras, ao aplicar restrições diferentes,

<sup>1</sup> Embora o nome sugira que esse estimador seja derivado do MODEX, na verdade, ele é derivado do MODE.

nem sempre *todas* as raízes são afetadas ao mesmo tempo. Além disso, o ruído não perturba a *mesma* estimativa DOA relativa a uma mesma fonte de sinal em todos os vetores de estimativas.

O funcionamento do MODEX Modificado é explicado de maneira breve. Primeiramente, ele executa o algoritmo completo do MODE impondo uma restrição de simetria complexa conjugada no vetor  $\hat{\mathbf{b}}$  (LOPES et al., 2003; KRUMMENAUER et al., 2010). Para  $M$  par,  $\hat{\mathbf{b}}$  possui a forma apresentada na Equação (3.15) e para  $M$  ímpar a forma mostrada na Equação (3.16) (SILVA, 2007).

$$\hat{\mathbf{b}} = \begin{bmatrix} \hat{b}_0 & \hat{b}_1 & \cdots & \hat{b}_{(\frac{M}{2}-1)} & \hat{b}_{\frac{M}{2}} & \hat{b}_{(\frac{M}{2}-1)}^* & \hat{b}_1^* & \hat{b}_0^* \end{bmatrix}^T \quad (3.15)$$

$$\hat{\mathbf{b}} = \begin{bmatrix} \hat{b}_0 & \hat{b}_1 & \cdots & \hat{b}_{(\frac{M-1}{2})} & \hat{b}_{(\frac{M-1}{2})}^* & \cdots & \hat{b}_1^* & \hat{b}_0^* \end{bmatrix}^T \quad (3.16)$$

A partir do vetor  $\hat{\mathbf{b}}$ , o segundo passo é gerar outras estimativas impondo duas restrições de não-trivialidade durante a minimização da Equação (3.13). A restrição linear real equivale a impor  $\Re(\hat{b}_0) = 1$  (ALVES, 2004). Entretanto, podem ocorrer instabilidades numéricas ao impor essa restrição, embora isso não seja frequente. Contudo, quando isso ocorre, basta repetir a minimização da Equação (3.13) sob a restrição  $\Im(\hat{b}_M) = 1$ , em que  $\Im(\cdot)$  é o operador parte imaginária. Essa restrição é chamada de restrição linear imaginária (ALVES, 2004). O uso dessas duas restrições lineares é o suficiente para abranger todas as situações práticas (SILVA, 2007) e uma dessas soluções sempre está correta para todos os valores de  $\mathbf{b}$  (LOPES et al., 2003).

Depois disso, são extraídas as  $3M$  raízes dos três polinômios gerados a partir do processo acima. Então, as raízes são combinadas em  $M$ -uplas (ou vetores de estimativas) e armazenadas na matriz  $\Omega_c$ . Assim, para selecionar a melhor combinação de raízes, é realizado um dos procedimentos ML apresentados nas Equações (3.17) e (3.18) (GERSHMAN; STOICA, 1999; SILVA, 2007).

$$\hat{\omega} = \arg \min_{\omega \in \Omega_c} [F_{\text{DMLE}}(\omega)] \quad (3.17)$$

$$\hat{\omega} = \arg \min_{\omega \in \Omega_c} [F_{\text{SMLE}}(\omega)] \quad (3.18)$$

em que:  $F_{\text{DMLE}}(\omega)$  e  $F_{\text{SMLE}}(\omega)$  são os critérios determinístico e estocástico do MLE, Equações (2.31) e (2.36) respectivamente. Em sequência, o vetor de estimativas  $\hat{\theta}$  em função dos ângulos DOA é calculado a partir de  $\hat{\omega}$  como apresentado na seção 2.1.

### 3.3 O Estimador SEAD Aprimorado

De maneira diferente do MODEX Modificado, o SEAD Aprimorado gera estimativas DOA em dois estágios como em um processo de otimização numérica. O primeiro estágio,

denominado pré-seleção, gera estimativas preliminares a partir do espectro diferencial, com baixo esforço computacional. O segundo estágio é responsável por realizar o refinamento dessas estimativas iniciais para obter as estimativas DOA de maneira mais precisa.

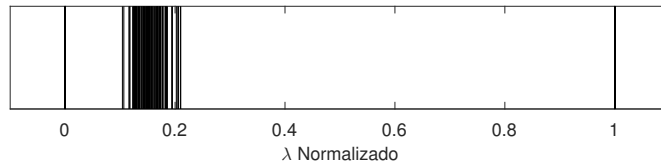
### 3.3.1 Espectro diferencial

Esta seção explica separadamente o espectro diferencial e como ele pode gerar estimativas iniciais.

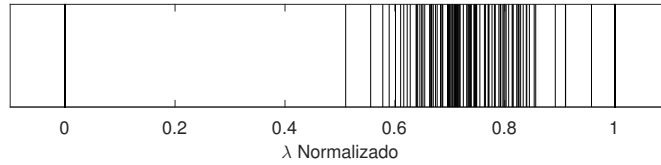
Para estimar as dimensões dos subespaços de sinal e de ruído de  $\hat{\mathbf{R}}$ , foi utilizada em (FERREIRA, 2005) a mesma técnica de avaliação do espalhamento dos autovalores usada em (ALVES; ASADA; MONTICELLI, 1999). Nessa pesquisa, foi observado que a diferença entre os dois maiores autovalores aumenta à medida que essas fontes de sinal se aproximam.

Essa diferença, considerando a ausência de ruído, é ilustrada nas Figuras 6a e 6b para as situações de fontes próximas e afastadas, respectivamente. Essas figuras unidimensionais foram geradas a partir da sobreposição de  $E = 100$  simulações independentes. Além disso, foram considerados um arranjo linear uniforme formado por  $K = 10$  sensores e  $N = 100$  *snapshots*.

Figura 6 – Espectro de autovalores de  $\hat{\mathbf{R}}$  na ausência de ruído.



(a) Considerando fontes situadas em  $10^\circ$  e  $15^\circ$



(b) Considerando fontes situadas em  $10^\circ$  e  $45^\circ$

Nas Figuras 6a e 6b, os espectros dos autovalores ilustrados foram normalizados dividindo todos os autovalores pelo máximo autovalor em cada um dos  $E = 100$  experimentos. Pode-se observar nessas figuras que o espectro possui três regiões principais. A primeira são todos os  $(K - \bar{M})$  autovalores de ruído para cada experimento. Como os experimentos foram realizados na ausência de ruído,  $\hat{\lambda}_{(\bar{M}+1)} = \hat{\lambda}_{(\bar{M}+2)} = \dots = \hat{\lambda}_K = 0$ . A segunda região é onde é encontrado o maior autovalor normalizado para cada experimento e, naturalmente,  $\hat{\lambda}_1 = 1$ .

Na terceira região, a intermediária, pode-se observar que  $\hat{\lambda}_2$  varia nos experimentos

devido à própria aleatoriedade do sinal. No entanto,  $\hat{\lambda}_2$  assume valores menores na situação de fontes próximas, como ilustrado na Figura 6a, se comparada à situação de fontes afastadas, como mostrado Figura 6b. Assim, pode-se verificar que a diferença ( $\hat{\lambda}_1 - \hat{\lambda}_2$ ) é, em geral, maior quando duas fontes de sinal estão afastadas entre si.

Então, esse comportamento dos dois maiores autovalores inspirou um estágio de busca angular. Nesse estágio é adicionada sinteticamente uma fonte de sinal de teste com o ângulo DOA  $\theta_T$  associado à frequência angular  $\omega_T$ . Para cada valor de  $\theta_T$ , é construída uma matriz de covariância espacial aumentada  $\mathbf{R}_T(\omega_T)$  como apresentado na Equação (3.19) (SILVA et al., 2013).

$$\mathbf{R}_T(\omega_T) = \hat{\mathbf{R}} + \mathbf{a}(\omega_T)\mathbf{a}(\omega_T)^H \quad (3.19)$$

em que:  $\mathbf{a}(\omega_T)$  é o vetor de direção associado à fonte de teste e  $\theta_T$  varia no intervalo  $[-90^\circ, 90^\circ]$ . Correspondentemente,  $-\pi \leq \omega_T \leq \pi$ .

Logo após, é calculada a diferença em função de  $\omega_T$  como apresentado na Equação (3.20).

$$D(\omega_T) = \bar{\lambda}_1 - \bar{\lambda}_2 \quad (3.20)$$

em que:  $\bar{\lambda}_1$  e  $\bar{\lambda}_2$  são os dois maiores autovalores de  $\mathbf{R}_T(\omega_T)$ . Após calcular todos os pontos do espectro diferencial, ele é normalizado de forma a se encaixar exatamente no intervalo  $[0, 1]$ .

Em vez de maximizar a potência da fonte de teste com a finalidade de se ajustar ao termo  $\mathbf{a}(\omega_T)\mathbf{a}(\omega_T)^H$ , o espectro diferencial maximiza  $D(\omega_T)$ . Assim, devido ao efeito de fontes sobrepostas, uma característica do espectro é que  $(\bar{\lambda}_1 - \bar{\lambda}_2)$  atinge um valor máximo quando  $\omega_T$  coincide com a frequência angular  $\omega_m$  de uma fonte de sinal (SILVA et al., 2013).

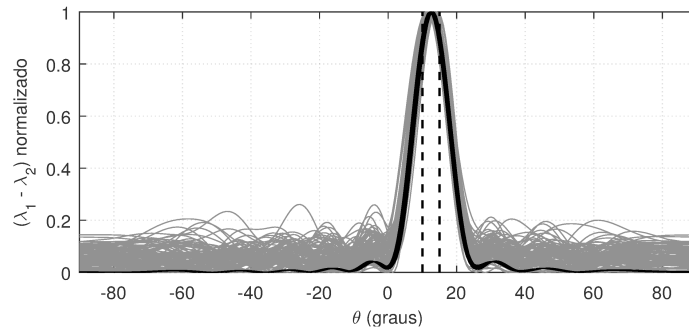
É importante ressaltar que teoricamente  $D(\omega_T)$  é imune ao ruído. Porém, na prática é usada a matriz de correlação espacial amostrada  $\hat{\mathbf{R}}$  em vez de  $\mathbf{R}$  e  $N$  é um valor finito. Consequentemente,  $D(\omega_T)$  não apresenta mais essa imunidade. Entretanto, avaliar esse termo em todo o intervalo  $[-\pi, \pi]$  resulta no espectro diferencial, que permite estimar os ângulos DOA a partir das frequências angulares associadas aos maiores picos espectrais.

As Figuras 7 a 9 ilustram o espectro diferencial em diversas situações de posicionamento de fontes de sinal na ausência e na presença de ruído. Em todas as situações foi utilizado um arranjo formado por  $K = 10$  sensores amostrando  $N = 100$  *snapshots* de sinais descorrelacionados. É importante ressaltar que foram calculados 100 espectros diferenciais na ausência de ruído e outros 100 considerando a SNR de  $-10$  dB. Todos esses espectros são independentes entre si e foram sobrepostos em cada figura para mostrar o comportamento médio em cada situação.

A Figura 7 mostra o espectro diferencial utilizando fontes próximas, com  $\theta_1 = 10^\circ$  e  $\theta_2 = 15^\circ$ . As linhas tracejadas verticais indicam as posições das fontes de sinal. Pode-se

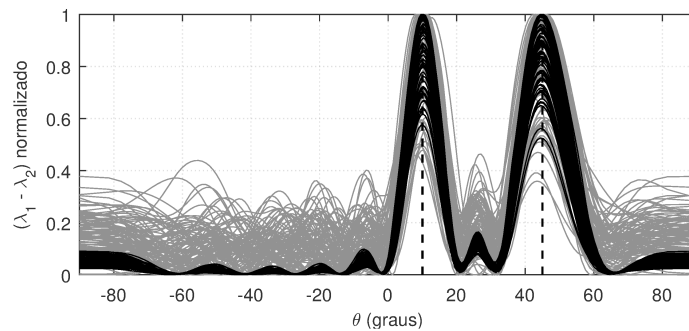
observar nessa figura que ocorre um único pico proeminente, com ápice em  $12^\circ$ , tanto na ausência de ruído (linhas cinzas), quanto na presença de ruído (linhas pretas). O aparecimento desse único pico é porque o arranjo não possui resolução suficiente para detectar cada fonte de sinal separadamente. Dessa forma, é necessário usar algum método que consiga gerar uma estimativa DOA para cada fonte de sinal a partir desse espectro. Além disso, é possível observar que, mesmo na SNR de  $-10$  dB, o pico associado às fontes de sinal apresenta amplitude consideravelmente maior do que os picos de ruído. Adicionalmente, nessa condição o pico pouco se deslocou de sua posição real.

Figura 7 – Espectros diferenciais sobrepostos considerando fontes situadas em  $10^\circ$  e  $15^\circ$  na ausência de ruído (linhas pretas) e na SNR de  $-10$  dB (linhas cinzas).



A Figura 8 ilustra o espectro diferencial usando fontes afastadas, com  $\theta_1 = 10^\circ$  e  $\theta_2 = 45^\circ$ . Na ausência de ruído, pode-se observar nessa figura dois picos proeminentes, um para cada fonte de sinal. Nesse caso, os ápices desses dois picos coincidem com as posições reais das fontes e representam as estimativas DOA para elas. Na SNR de  $-10$  dB, a diferença de amplitude entre os picos de sinal e de ruído diminui, embora na maior parte dos experimentos ainda seja possível distinguir picos de sinal de picos de ruído por suas amplitudes.

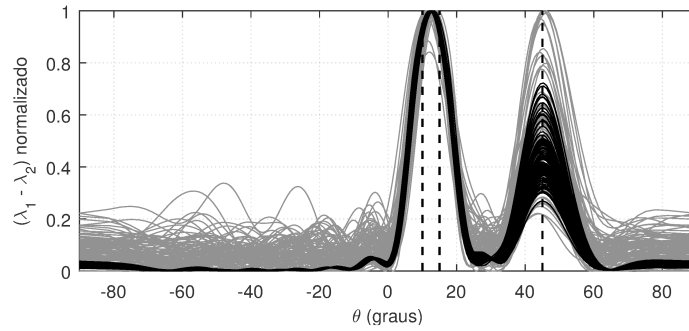
Figura 8 – Espectros diferenciais sobrepostos considerando fontes situadas em  $10^\circ$  e  $45^\circ$  na ausência de ruído (linhas pretas) e na SNR de  $-10$  dB (linhas cinzas).



A Figura 9 ilustra o espectro diferencial utilizando fontes situadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$ . Pode-se verificar nessa figura que, considerando a ausência de ruído, o pico associado às fontes em  $10^\circ$  e  $15^\circ$  possui amplitude maior do que o pico com ápice em  $45^\circ$ . Na SNR

de  $-10$  dB, pode-se observar que o pico associado às fontes situadas em  $10^\circ$  e  $15^\circ$  é significativamente menos afetado pelo ruído do que o pico em  $45^\circ$ . Inclusive, esse último pico pode se misturar aos picos de ruído e é difícil a sua detecção por amplitude.

Figura 9 – Espectros diferenciais sobrepostos considerando fontes situadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$  na ausência de ruído (linhas pretas) e na SNR de  $-10$  dB (linhas cinzas).



### 3.3.2 Processo de seleção de estimativas

O estimador SEAD é a base do SEAD Aprimorado. No entanto, o SEAD não será descrito neste trabalho, pois, mesmo que ele apresente desempenho de estimação bastante competitivo, ele possui esforço computacional significativamente maior. Consequentemente, a sua viabilidade em situações práticas é restrita, pois a estimação DOA pode ser usada em aplicações em tempo real (ALVES, 2004).

A razão disso é que o SEAD utiliza um processo de busca multidimensional considerado método de força-bruta (SILVA et al., 2013). Nesses métodos, a otimização ocorre por uma busca no espaço de soluções que exige um tempo de execução de ordem exponencial. Sempre que possível, os métodos de força-bruta devem ser evitados, entendendo mais profundamente o problema, o que possibilita gerar um algoritmo que de maneira ideal apresente esforço computacional de ordem polinomial (SIPSER, 1997; CARVALHO et al., 2001).

O SEAD Aprimorado foi proposto com o objetivo de diminuir esse esforço computacional e ao mesmo tempo apresentar desempenho de estimação competitivo. Essa redução foi alcançada substituindo o método de força-bruta existente por outro processo de busca baseado no algoritmo *Branch-and-Bound*. Esse algoritmo é comumente utilizado na otimização discreta e combinatorial e subdivide de maneira sistemática o espaço de busca em subconjuntos, permitindo desconsiderar as estimativas inadequadas.

Adicionalmente, o estágio de pré-seleção foi adaptado de forma que possa ser usado em outras situações de posicionamento e número de fontes de sinal. Isso é importante, pois a inicialização correta do processo de otimização por si só pode levar a uma considerável melhoria na estimação DOA (SILVA et al., 2013).



### 3.3.2.1 Pré-seleção

O objetivo do estágio de pré-seleção é gerar um vetor de estimativas iniciais a partir das frequências angulares dos  $Q$  picos com ápices acima de um limiar de amplitude. Esses picos são chamados picos detectados. Contudo, na seção 3.3.1 é possível observar que o espectro diferencial apresenta baixa resolução e não consegue distinguir fontes próximas. Assim, o estágio de pré-seleção também é responsável por estimar o número de fontes em cada pico detectado, pois o número de picos pode ser diferente do número de fontes de sinal (SILVA, 2009).

Assim, existem três situações ao aplicar a restrição pelo limiar de amplitude durante a fase de pré-seleção:  $Q > M$ ,  $Q = M$ ,  $Q < M$ . As duas primeiras situações são resolvidas da seguinte maneira:

1. Para  $Q > M$ , o SEAD Aprimorado realiza o procedimento ML utilizando as frequências angulares dos ápices dos picos detectados. O vetor de estimativas iniciais é a melhor combinação de frequências encontradas. Esse processo avalia  $C = \frac{Q!}{M!(M-Q)!}$  combinações;
2. Quando  $Q = M$ , o SEAD Aprimorado considera as frequências angulares dos ápices dos picos detectados como estimativas iniciais e não é necessário outro passo adicional. Dessa forma, o número de combinações avaliadas é  $C = 0$ .

Porém, na terceira situação,  $Q < M$ , é necessário realizar passos adicionais partindo do pressuposto de que todos os picos detectados são picos de sinal. O primeiro passo é armazenar as frequências angulares dos vales à esquerda e à direita ( $\omega_{(v,q,e)}$  e  $\omega_{(v,q,d)}$ ) e dos ápices ( $\omega_{(a,q)}$ ) do  $q$ -ésimo pico detectado no espectro diferencial, onde  $q = 1, 2, \dots, Q$  (SILVA et al., 2013).

Então, supõe-se que cada pico detectado corresponde a uma quantidade de no mínimo  $c_q = 1$  e no máximo  $c_q = (M - Q + 1)$  fontes de sinal. São consideradas apenas as combinações  $\{c_1, c_2, \dots, c_Q\}$  em que  $\sum_{q=1}^Q c_q = M$ . Assim, ao distribuir  $M$  fontes a  $Q$  picos detectados são produzidos  $C$  arranjos de distribuição de fontes de sinal por pico detectado, de acordo com a Equação (3.21).

$$C = \frac{(M-1)!}{(M-Q)!(Q-1)!} \quad (3.21)$$

Se  $c_q = 1$ , então é considerada apenas a frequência angular associada ao ápice do  $q$ -ésimo pico. Quando  $c_q$  é um número ímpar maior do que um, o ápice é considerado junto com outras as  $N_q = (c_q - 1)/2$  estimativas candidatas posicionadas de cada lado daquele pico. Se  $c_q$  é par, o algoritmo meramente aloca  $N_q = c_q/2$  estimativas candidatas em cada lado do  $q$ -ésimo pico, desconsiderando a frequência correspondente ao seu ápice.



As estimativas posicionadas à esquerda e à direita do pico são calculadas como mostrado nas Equações (3.22) e (3.23), respectivamente.

$$\hat{\omega}_{(e,q,n)} = \omega_{(a,q)} - n \times \frac{\Delta\omega_{(e,q)}}{2(N_q + 1)} \quad (3.22)$$

$$\hat{\omega}_{(d,q,n)} = \omega_{(a,q)} + n \times \frac{\Delta\omega_{(d,q)}}{2(N_q + 1)} \quad (3.23)$$

em que:  $n = 1, 2, \dots, N_q$ ;  $\Delta\omega_{(e,q)} = \omega_{(a,q)} - \omega_{(v,q,e)}$  é a largura à esquerda do pico e  $\Delta\omega_{(d,q)} = \omega_{(v,q,d)} - \omega_{(a,q)}$  é a largura à direita do pico.

O tamanho dos intervalos de busca é limitado às regiões dos picos detectados. Considere que um pico foi gerado por múltiplas fontes próximas. As frequências angulares dessas fontes tipicamente estão próximas da posição do ápice do pico. Dessa forma, as estimativas candidatas são geradas na metade interior de cada lado do pico.

Assim, são construídas e armazenadas em  $\Omega_c$  as  $M$ -uplas de acordo com os  $C$  arranjos de distribuição de fontes para cada pico e suas larguras esquerda e direita. Então, seleciona-se a melhor combinação usando um dos procedimentos ML mostrados nas Equações (3.24) e (3.25).

$$\hat{\omega} = \arg \min_{\omega \in \Omega_c} [F_{\text{DMLE}}(\omega)] \quad (3.24)$$

$$\hat{\omega} = \arg \min_{\omega \in \Omega_c} [F_{\text{SMLE}}(\omega)] \quad (3.25)$$

em que:  $F_{\text{DMLE}}(\omega)$  e  $F_{\text{SMLE}}(\omega)$  são os critérios determinístico e estocástico do MLE. Essa combinação é considerada o vetor de estimativas iniciais para o próximo estágio.

O exemplo a seguir é apresentado em função dos ângulos DOA. Considere fontes equipotentes, decorrelacionadas e posicionadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$  na ausência de ruído. Suponha também que os  $Q = 2$  picos de sinal tenham sido detectados corretamente. As posições dos ápices desses picos são  $\{\theta_{(a,1)}, \theta_{(a,2)}\} = [13^\circ \ 45^\circ]$  e as posições dos vales à esquerda e à direita de cada pico detectado são  $\{\theta_{(v,1,e)}, \theta_{(v,1,d)}, \theta_{(v,2,e)}, \theta_{(v,2,d)}\} = [0^\circ \ 27^\circ \ 31^\circ \ 64^\circ]$ .

Logo após são gerados dois arranjos de distribuição de fontes por pico:  $\{1, 2\}$  e  $\{2, 1\}$ . Em outras palavras, o primeiro arranjo corresponde à situação em que o primeiro pico está associado a uma fonte de sinal e o segundo pico a duas. No segundo arranjo ocorre o inverso: o primeiro pico está associado a duas fontes de sinal e o segundo a uma fonte.

Assim, duas estimativas são construídas como mostrado na Tabela 1. As estimativas selecionadas são  $[9,5^\circ \ 16,2^\circ \ 45,3^\circ]$ , pois elas minimizam a função-objetivo. Então, essa estimativa é considerada a estimativa inicial que será refinada no próximo estágio.

Tabela 1 – Combinações geradas no estágio de pré-seleção para o caso de fontes situadas em 10°, 15° e 45° na ausência de ruído.

Arranjo	Estimativas candidatas	DMLE
{1, 2}	[13° 41,5° 49,3°]	0,01244
{2, 1}	[9,5° 16,2° 45,3°]	0,00034

É importante ressaltar que o estágio de pré-seleção é muito dependente do cálculo correto do limiar de amplitude. Adicionalmente, um limiar fixo não é adequado em diversas situações. Se ele é muito alto, picos de sinal podem ser perdidos. Se ele é muito baixo, picos de ruído podem gerar estimativas.

Considerando que todo o estágio de pré-seleção é baseado na quantidade de picos detectados e no número de fontes de sinal, a detecção dos picos pode afetar negativamente de maneira significativa a estimação DOA. Além disso, o fato de fontes próximas gerarem picos maiores do que uma fonte afastada torna o cálculo do limiar de amplitude ainda mais importante (SILVA et al., 2013).

Então, foi proposto em (SILVA et al., 2013) o uso de um limiar de amplitude adaptativo que é inversamente proporcional ao número de fontes. Na prática, esse número deve ser estimado e técnicas convencionais para isso podem falhar (HAYKIN, 1990). Tipicamente, os picos de ruído são consideravelmente menores do que os picos de sinal e se mantêm abaixo desse limiar, mesmo quando  $M$  seja estimado para cima de maneira incorreta.

Ainda, a subestimação incorreta de  $M$  piora a situação de picos gerados por múltiplas fontes, pois pequenos picos de sinal podem ser desconsiderados. Assim, para diminuir essas ocorrências, o limiar de amplitude é diminuído de acordo com o complemento da amplitude normalizada dos picos. Dessa forma, o limiar de amplitude é calculado como apresentado na Equação (3.26).

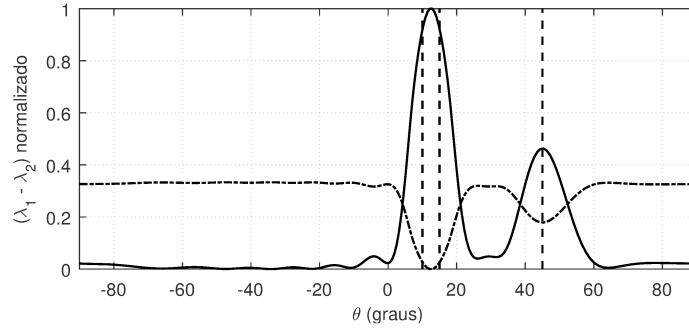
$$L_R(\omega_T) = \frac{1 - D(\omega_T)}{M} \quad (3.26)$$

A Figura 10 mostra um exemplo da aplicação do limiar de amplitude adaptativo considerando fontes de sinais descorrelacionados situadas em 10°, 15° e 45° na ausência de ruído. A linha traço-ponto nessa figura representa esse limiar.

### 3.3.3 Refinamento das estimativas

Apesar do estágio de pré-seleção ser capaz de estimar razoavelmente bem os ângulos DOA das fontes de sinal, sua resolução é limitada em cerca de 1°. Para aumentar a precisão das estimativas, foi proposto em (SILVA et al., 2013) realizar um refinamento utilizando um procedimento heurístico baseado no algoritmo *Branch-and-Bound*.

Figura 10 – Espectro diferencial considerando fontes situadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$  na ausência de ruído e o respectivo limiar de amplitude.



Esse procedimento ramifica (“*branch*”) cada estimativa  $\hat{\omega}_m$  contida no vetor de estimativas iniciais, gerando duas estimativas candidatas, uma à esquerda ( $\hat{\omega}_{(e,m)}$ ) e outra à direita ( $\hat{\omega}_{(d,m)}$ ) de  $\hat{\omega}_m$ . Assim, para cada fonte de sinal é construído um grupo com as frequências  $\{\hat{\omega}_{(e,m)}, \hat{\omega}_m, \hat{\omega}_{(d,m)}\}$ . Na primeira iteração,  $\hat{\omega}_{(e,m)}$  e  $\hat{\omega}_{(d,m)}$  são calculados respectivamente utilizando as Equações (3.27) e (3.28).

$$\hat{\omega}_{(e,m)} = \hat{\omega}_m - \frac{\Delta\omega_{(e,q)}}{4(N_q + 1)} \quad (3.27)$$

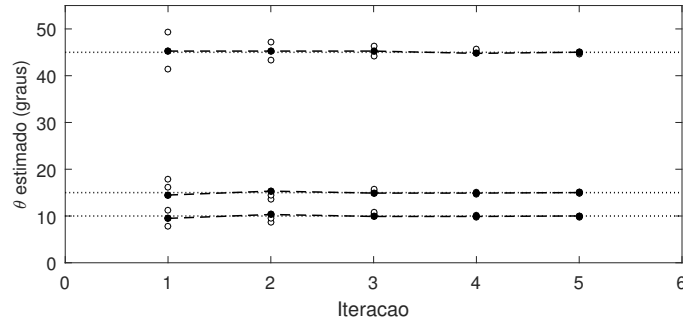
$$\hat{\omega}_{(d,m)} = \hat{\omega}_m + \frac{\Delta\omega_{(d,q)}}{4(N_q + 1)} \quad (3.28)$$

em que:  $q = 1, 2, \dots, Q$  e  $\Delta\omega_{(e,q)}$  e  $\Delta\omega_{(d,q)}$  estão associados ao  $q$ -ésimo pico. Os passos  $\Delta\omega_{(e,i)}/4(n+1)$  e  $\Delta\omega_{(d,i)}/4(n+1)$  correspondem à metade da largura entre as frequências do mesmo lado do pico, com a finalidade de evitar a sobreposição de grupos adjacentes.

Assim, esses grupos são combinados e geram novas estimativas, que são posteriormente armazenadas em  $\Omega_c$ . No entanto, é importante ressaltar que as frequências pertencentes ao mesmo grupo não são combinadas entre si de forma a evitar combinações ineficazes. Então, a melhor combinação em  $\Omega_c$  é selecionada usando o procedimento ML da Equação (3.24). Esse processo é repetido até um limite arbitrário e os passos são divididos pela metade a cada iteração para melhorar a resolução. Finalmente, o vetor de estimativas DOA  $\hat{\theta}$  é calculado a partir da combinação das frequências selecionadas.

A Figura 11 ilustra esse processo de refinamento de estimativas considerando fontes situadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$  na ausência de ruído. Nessa figura, as linhas pontilhadas indicam as posições reais das fontes de sinal, enquanto os círculos indicam as estimativas DOA em cada iteração. As estimativas selecionadas são aquelas preenchidas de preto. Pode-se observar nessa figura a convergência das estimativas para os valores reais.

Figura 11 – Estimativas geradas e selecionadas durante o processo de refinamento de estimativas no caso de fontes situadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$  na ausência de ruído.



### 3.4 Definição do Processo de Filtragem e da Nova Função-Objetivo

As próximas seções deste capítulo descrevem o processo de filtragem apresentado em (KRUMMENAUER, 2007; KRUMMENAUER et al., 2010).

Considerando o modelo de sinal mostrado na Equação (2.10), o espectro de sinal é definido por  $M$  componentes distintas nas frequências angulares  $\omega_m$  e o espectro de potência é constante em todo o intervalo  $[-\pi, \pi]$ .

Baseando-se nisso, é possível aplicar um filtro em cada *snapshot* de forma a aumentar a SNR de saída. Isso permite atenuar o ruído e preservar o sinal convoluindo os *snapshots* com a função do sistema que define esse filtro. Essa função é mostrada na Equação (3.29) (KRUMMENAUER, 2007).

$$h(z) = \hat{h}_0 + \hat{h}_1 z^{-1} + \dots + \hat{h}_{\hat{L}} z^{-\hat{L}} \quad (3.29)$$

em que:  $0 < \hat{L} < K$  é a ordem do filtro e  $\hat{\mathbf{h}}^H = [\hat{h}_{\hat{L}} \dots \hat{h}_0] \in \mathbb{C}^{(\hat{L}+1) \times 1}$  corresponde ao vetor de coeficientes do filtro (KRUMMENAUER et al., 2010).

Esse processo resulta no *snapshot* filtrado  $\mathbf{z}[n] = [z[n, 1] \dots z[n, (K + \hat{L})]]^T$ , em que a sua  $i$ -ésima amostra é calculada realizando a multiplicação apresentada na Equação (3.30).

$$z[n, i] = \hat{\mathbf{h}}^H \mathbf{y}[n, i] \quad (3.30)$$

em que:  $\mathbf{y}[n, i] = [y[i - \hat{L}] \dots y[i]]^T$ . Usando a Equação (2.10), obtém-se a Equação (3.31).

$$z[n, i] = \hat{\mathbf{h}}^H \mathbf{A}[i] \mathbf{x}[n] + \hat{\mathbf{h}}^H \mathbf{n}[n, i] \quad (3.31)$$

em que:  $\mathbf{A}[i]$  é composta de  $(\hat{L} + 1)$  linhas de  $\mathbf{A}$  a partir de  $[e^{j(i-\hat{L}-1)\omega_1} \dots e^{j(i-\hat{L}-1)\omega_M}]$  e  $\mathbf{n}[n, i] = [n[i - \hat{L}] \dots n[i]]^T$ .

No entanto, todo esse processo de convolução pode ser descrito de maneira mais simples escrevendo a Equação (3.30) em função de termos vetoriais como é mostrado na

Equação (3.32).

$$\begin{aligned}\mathbf{z}[n] &= \hat{\mathbf{H}}\mathbf{y}[n] \\ &= \hat{\mathbf{H}}\mathbf{A}\mathbf{x}[n] + \hat{\mathbf{H}}\mathbf{n}[n]\end{aligned}\tag{3.32}$$

em que:  $\hat{\mathbf{H}} \in \mathbb{C}^{(K+\hat{L}) \times K}$  é a matriz de convolução do filtro definida pela Equação (3.33).

$$\hat{\mathbf{H}} = \begin{bmatrix} \hat{h}_{\hat{L}} & \cdots & \hat{h}_1 & \hat{h}_0 & \mathbf{0} \\ & \ddots & & \ddots & \ddots \\ \mathbf{0} & & \hat{h}_{\hat{L}} & \cdots & \hat{h}_1 & \hat{h}_0 \end{bmatrix}^H \tag{3.33}$$

Assim, a nova matriz de covariância estimada  $\hat{\mathbf{R}}_z \in \mathbb{C}^{(K+\hat{L}) \times (K+\hat{L})}$  é calculada como apresentado na Equação (3.34).

$$\begin{aligned}\hat{\mathbf{R}}_z &= \frac{1}{N} \sum_{n=1}^N \mathbf{z}[n]\mathbf{z}^H[n] \\ &= \frac{1}{N} \sum_{n=1}^N \hat{\mathbf{H}}\mathbf{y}[n]\mathbf{y}^H[n]\hat{\mathbf{H}}^H \\ &= \hat{\mathbf{H}} \left[ \frac{1}{N} \sum_{n=1}^N \mathbf{y}[n]\mathbf{y}^H[n] \right] \hat{\mathbf{H}}^H \\ &= \hat{\mathbf{H}}\hat{\mathbf{R}}\hat{\mathbf{H}}^H\end{aligned}\tag{3.34}$$

É importante ressaltar que  $\hat{\mathbf{R}}$  é uma matriz definida-positiva e  $\text{posto}(\hat{\mathbf{H}}) = K$ . Dessa forma,  $\text{posto}(\hat{\mathbf{R}}_z) = K$  e  $\hat{\mathbf{R}}_z$  é uma matriz semidefinida-positiva. Adicionalmente, a multiplicação  $\hat{\mathbf{H}}\hat{\mathbf{R}}\hat{\mathbf{H}}^H$  resulta em dois fatos que devem ser considerados (KRUMMENAUER et al., 2010):

1. A matriz  $\hat{\mathbf{R}}_Z$  possui dimensões maiores do que  $\hat{\mathbf{R}}$  e, conseqüentemente, não é possível realizar a multiplicação  $\mathbf{P}_A^\perp \hat{\mathbf{R}}_Z$  na função de máxima-verossimilhança mostrada na Equação (2.31);
2. Considere a Equação (3.32). O modelo de sinal é modificado, pois a matriz  $\mathbf{A}$  perde a estrutura de Vandermonde. Adicionalmente, o ruído se torna correlacionado e os *snapshots* filtrados não são mais gaussianos, pois o termo  $\hat{\mathbf{H}}\hat{\mathbf{H}}^H$  não é definido-positivo.

A consequência disso é que a função de máxima-verossimilhança não é aplicável para dados filtrados e uma nova função-objetivo deve ser calculada a partir da distribuição estatística dos *snapshots* filtrados. Porém, a estratégia utilizada pelos pesquisadores em

(KRUMMENAUER et al., 2010) foi a de simplesmente substituir  $\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}$  na Equação (2.31) por seus correspondentes filtrados como apresentado na Equação (3.35).

$$\hat{\boldsymbol{\omega}} = \arg \min_{\boldsymbol{\omega} \in \Omega_c} [F_{\text{DMLEF}}(\boldsymbol{\omega})] \quad (3.35)$$

em que:  $F_{\text{DMLEF}}(\boldsymbol{\omega})$  é a função-objetivo do DMLE filtrado, como mostrado na Equação (3.36).

$$F_{\text{DMLEF}}(\boldsymbol{\omega}) = \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z) \quad (3.36)$$

em que:  $\mathbf{P}_{\hat{\mathbf{A}}}^\perp = \mathbf{I} - \tilde{\mathbf{A}}\tilde{\mathbf{A}}^+$ ,  $\tilde{\mathbf{A}}^+ = \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^H$  é a pseudo-inversa (ou a inversa generalizada) de Moore-Penrose e  $\tilde{\mathbf{A}} = \hat{\mathbf{H}}\mathbf{A}$  é a matriz de direção filtrada.

Finalmente, as estimativas geradas pelo MODEX Modificado e pelo SEAD Aprimorado são selecionadas usando a Equação (3.35) em vez das Equações (3.17) e (3.18).

### 3.5 Filtro para Maximização da SNR

Foi proposto em (KRUMMENAUER, 2007; KRUMMENAUER et al., 2010) filtrar os dados de forma a maximizar a SNR de saída. Para isso, durante o projeto do filtro é necessário avaliar a energia média do sinal quando todos os elementos da resposta impulsiva são diferentes de zero. Essa avaliação pode ser realizada tendo como referência a amostra  $z[n, \hat{L} + 1]$  do *snapshot* filtrado correspondente à linha  $(\hat{L} + 1)$  da matriz  $\hat{\mathbf{H}}$ .

Como o filtro é linear, o princípio da superposição se aplica, permitindo analisar os efeitos no sinal e no ruído separadamente (HAYKIN, 2001). Dessa forma, considere a Equação (3.37).

$$\begin{aligned} z[n, (\hat{L} + 1)] &= \bar{\mathbf{h}}^H \hat{\mathbf{y}}[n, (\hat{L} + 1)] \\ &= \bar{\mathbf{h}}^H \mathbf{A}[\hat{L} + 1] \mathbf{x}[n] + \bar{\mathbf{h}}^H \mathbf{n}[n, (\hat{L} + 1)] \\ &= z_s[n, (\hat{L} + 1)] + z_r[n, (\hat{L} + 1)] \end{aligned} \quad (3.37)$$

em que:  $\bar{\mathbf{h}}$  é o transposto-conjugado da  $(\hat{L} + 1)$ -ésima linha de  $\hat{\mathbf{H}}$ . Dessa forma, as energias médias do sinal e do ruído são calculadas como apresentado nas Equações (3.38) e (3.39).

$$\begin{aligned} E \{ z_s[n, (\hat{L} + 1)] z_s^*[n, (\hat{L} + 1)] \} &= \bar{\mathbf{h}}^H \mathbf{A}[\hat{L} + 1] E \{ \mathbf{x}[n] \mathbf{x}^H[n] \} \mathbf{A}^H[\hat{L} + 1] \bar{\mathbf{h}} \\ &= \bar{\mathbf{h}}^H \mathbf{A}[\hat{L} + 1] \mathbf{C} \mathbf{A}^H[\hat{L} + 1] \bar{\mathbf{h}} \end{aligned} \quad (3.38)$$

$$E \{ z_r[n, (\hat{L} + 1)] z_r^*[n, (\hat{L} + 1)] \} = \sigma^2 \bar{\mathbf{h}}^H \bar{\mathbf{h}} \quad (3.39)$$

A partir dessas equações é possível calcular a SNR média correspondente a  $z[n, (\hat{L}+1)]$ , como mostrado na Equação (3.40).

$$\text{SNR} = \frac{\bar{\mathbf{h}}^H \mathbf{A}[\hat{L}+1] \mathbf{C} \mathbf{A}^H[\hat{L}+1] \bar{\mathbf{h}}}{\sigma^2 \bar{\mathbf{h}}^H \bar{\mathbf{h}}} \quad (3.40)$$

Um assunto fundamental na teoria de telecomunicação que está associado a um problema de autovalores, é definir um filtro FIR ótimo que maximiza a SNR de saída (HAYKIN, 2001). Partindo do pressuposto de que  $\sigma_2$  não depende do filtro projetado, o filtro ótimo  $\mathbf{h}_{ol}$  é calculado resolvendo o problema de minimização apresentado na Equação (3.41) (KRUMMENAUER, 2007).

$$\mathbf{h}_{ol} = \arg \min_{\mathbf{h}} \frac{\bar{\mathbf{h}}^H \mathbf{A}[\hat{L}+1] \mathbf{C} \mathbf{A}^H[\hat{L}+1] \bar{\mathbf{h}}}{\bar{\mathbf{h}}^H \bar{\mathbf{h}}} \quad (3.41)$$

em que:  $\mathbf{h}_{ol}$  está sujeito a restrição da norma unitária  $\bar{\mathbf{h}}^H \bar{\mathbf{h}} = 1$  para ajustar o ganho do filtro (KRUMMENAUER et al., 2010).

Para qualquer valor de  $\hat{L}$  no intervalo  $0 < \hat{L} < K$ , a solução da Equação (3.41) é o *máximo* autofiltro (KRUMMENAUER et al., 2010). Por si só, um autofiltro é um filtro FIR cuja resposta impulsiva tem coeficientes iguais aos elementos de um autovetor. Por sua vez, o máximo autofiltro (ou filtro ótimo) é o autofiltro associado ao maior autovalor da matriz de correlação do componente de sinal na entrada do filtro (HAYKIN, 2001). Assumindo o caso teórico, o filtro ótimo é o máximo autovetor de  $\mathbf{A}[i] \mathbf{C} \mathbf{A}^H[i]$  (KRUMMENAUER et al., 2010).

Porém, não é possível calcular  $\mathbf{h}_{ol}$ , pois ele depende dos ângulos DOA verdadeiros, que são justamente os parâmetros a serem estimados. Entretanto, o modelo de sinal adotado neste trabalho considera ruído branco. Por isso, pode-se assumir a matriz mostrada na Equação (3.42).

$$\mathbf{R}[i] = \mathbf{A}[i] \mathbf{C} \mathbf{A}^H[i] + \sigma \mathbf{I} \quad (3.42)$$

em que:  $\mathbf{R}[i] \in \mathbb{C}^{(\bar{L}+1) \times (\bar{L}+1)}$ . É importante observar que  $\mathbf{R}[i]$  é uma das  $(K - \bar{L})$  submatrizes ao longo da diagonal principal de  $\mathbf{R}$  e qualquer uma delas podem ser escolhidas. Outra maneira é calcular a média de todas elas, como foi feito em (KRUMMENAUER et al., 2010) e apresentado na Equação (3.43).

$$\mathbf{R}_I = \frac{1}{K - \bar{L}} \sum_{i=\bar{L}+1}^K \mathbf{R}[i] \quad (3.43)$$

Então, o vetor de coeficientes é a solução ótima  $\mathbf{h}_{ol} = [h_{(ol, \hat{L})} \ \cdots \ h_{(ol, 0)}] = \mathbf{e}_1$ , em que:  $\mathbf{e}_1$  é o máximo autovetor da matriz escolhida, seja ela  $\mathbf{R}[i]$  ou  $\mathbf{R}_I$ . No entanto, as matrizes teóricas  $\mathbf{R}$ ,  $\mathbf{R}[i]$  e  $\mathbf{R}_I$  são desconhecidas durante o processo de estimação. Assim, o máximo autofiltro espacial deve ser estimado considerando a solução  $\mathbf{h}_{ol} = \hat{\mathbf{e}}_1$  calculada a partir das matrizes estimadas  $\hat{\mathbf{R}}$ ,  $\hat{\mathbf{R}}[i]$  e  $\hat{\mathbf{R}}_I$  (KRUMMENAUER et al., 2010).

### 3.6 Considerações Finais

Este capítulo apresentou a fundamentação teórica dos estimadores MODEX Modificado e SEAD Aprimorado, além de descrever o processo de filtragem espacial utilizando o máximo autofiltro espacial.

O MODE apresenta esforço computacional significativamente menor, porém apresenta valores maiores de SNR de limiar. Isso limita a sua aplicabilidade prática apenas em situações de alta SNR, quando ele apresenta desempenho de estimação semelhante ao Estimador de Máxima-Verossimilhança.

O MODEX Modificado parte do pressuposto de que o ruído influencia de maneira diferente determinadas restrições impostas na otimização da função-objetivo do MODE. Então, são geradas três estimativas MODE, totalizando  $3M$  raízes candidatas. Dessa forma, é necessário realizar um procedimento ML para escolher a melhor combinação de raízes. Devido a isso, o MODEX Modificado apresenta esforço computacional consideravelmente maior do que o MODE. Por outro lado, o MODEX Modificado possui um desempenho de estimação superior ao do MODE, com SNRs de limiar significativamente mais baixas.

O SEAD Aprimorado realiza a estimação DOA a partir de um vetor de estimativas iniciais geradas a partir do espectro diferencial. Esse vetor é posteriormente refinado por um estágio de otimização baseado no algoritmo *Branch-and-Bound*. Além disso, o estágio de pré-seleção é responsável pela geração de estimativas iniciais e é capaz de abranger diversas situações de posicionamento das fontes de sinal. Porém, ele é bastante dependente de um cálculo adequado do limiar de amplitude para distinguir os picos de sinal dos picos de ruído. Adicionalmente, o estágio de refinamento é responsável por aumentar a precisão das estimativas. No entanto, esse estágio é baseado no algoritmo *Branch-and-Bound* usado para otimização discreta e combinatorial. Assim, esse algoritmo possui um limite de resolução em função da quantidade de iterações utilizada e faz do SEAD Aprimorado um estimador assintoticamente ineficiente.

Por outro lado, foi observado em (KRUMMENAUER et al., 2010) que o procedimento ML utilizado no processo de seleção de estimativas pode ser aprimorado à medida que a SNR diminui. Isso pode ser explicado pelo fato do processo de seleção ser consideravelmente afetado pelo ruído presente na matriz de correlação. Dessa forma, foi projetado um filtro espacial de ordem  $\hat{L}$  com o objetivo de maximizar a SNR de saída. Esse filtro é baseado no máximo autofiltro definido em (HAYKIN, 2001).

Como a estimação DOA se aplica a fontes próximas e afastadas, o ideal é que o filtro seja multibanda, de forma que apresente ganho unitário nas frequências angulares associadas às fontes de sinal e ganho inferior nas outras faixas frequências. Isso deve ocorrer de maneira a preservar os sinais ao mesmo tempo em que atenua o ruído. Adicionalmente, é importante ressaltar que  $\hat{L}$  deve ser pequeno o suficiente para que não incorra em um



excessivo esforço computacional devido às matrizes com dimensões aumentadas. Além disso, deve-se permitir que a resposta em frequência do filtro detecte corretamente as bandas de passagem das fontes de sinal.

O máximo autofiltro espacial foi proposto considerando apenas a situação de fontes próximas. Entretanto, os pesquisadores afirmaram em (KRUMMENAUER et al., 2010) que foi observado um aprimoramento no desempenho de estimação em outras situações não esclarecidas considerando fontes de sinais correlacionados e descorrelacionados. As simulações realizadas neste trabalho mostraram um ganho considerável de desempenho para fontes próximas ao usar o filtro espacial. No entanto, ele apresentou perda significativa de desempenho na presença de fontes afastadas.

O próximo capítulo apresenta os motivos que permitem explicar essa perda de desempenho. Depois, são propostos dois novos filtros espaciais que são aplicáveis à estimação DOA tanto de fontes próximas quanto afastadas. Ainda, o Capítulo 4 apresenta as equações desenvolvidas neste trabalho para o Método de Newton e para o método quase-Newton BFGS aplicado ao Estimador de Máxima-Verossimilhança na ausência e na presença de filtragem espacial. Finalmente, é apresentada uma alteração no estágio de pré-seleção do SEAD Aprimorado que não exige o uso do limiar de amplitude para distinguir picos de sinal dos picos de ruído.

## 4 PROPOSTA DE ALTERAÇÃO AO ESTIMADOR SEAD

O Capítulo 3 apresentou estimadores iterativos que apresentam desempenho de estimação semelhante ao Estimador de Máxima-Verossimilhança com um esforço computacional significativamente menor. No entanto, foi observado que o SEAD Aprimorado possui em altas SNRs um desempenho de estimação limitado pelo estágio de refinamento de estimativas. Além disso, tanto o MODEX Modificado quanto o SEAD Aprimorado apresentam o efeito de limiar.

A partir do estudo realizado em (KRUMMENAUER et al., 2010), uma das maneiras de diminuir a SNR de limiar é reduzindo a influência do ruído no procedimento ML utilizado para selecionar a melhor combinação de estimativas. Isso foi realizado aplicando o máximo autofiltro espacial em uma versão modificada da função-objetivo do DMLE. Esse autofiltro é o filtro ótimo e ele maximiza a SNR de saída (HAYKIN, 2001). Durante este trabalho, as simulações mostraram um ganho significativo no desempenho de estimação do MODEX Modificado para fontes próximas. Por outro lado, as simulações também mostraram que ocorreu uma degradação significativa considerando fontes afastadas.

Adicionalmente, o SEAD Aprimorado apresenta limitação do desempenho de estimação em altas SNRs e se deve principalmente ao processo de busca utilizado no estágio de refinamento de estimativas. Essa busca é baseada no algoritmo *Branch-and-Bound* e produz estimativas em um intervalo discretizado com esforço computacional de ordem combinatorial. Ainda, o SEAD Aprimorado pode apresentar baixo desempenho de estimação considerando fontes afastadas em virtude de um limiar de amplitude sub ou superdimensionado. Na primeira situação, picos de ruído podem ser considerados como picos de sinal e, na segunda situação, picos de sinal podem ser descartados. Tanto uma quanto outra podem afetar severamente a geração da estimativa inicial.

De maneira geral, este capítulo propõe novos filtros espaciais, além de propor alterações ao SEAD Aprimorado. O objetivo é tornar esse estimador mais robusto ao ruído e apresentar menor esforço computacional de maneira independente do afastamento entre as fontes de sinal.

Para isso, este capítulo realiza um estudo sobre os motivos que levaram o máximo autofiltro espacial a apresentar um baixo desempenho de estimação considerando fontes afastadas. Com base nesse estudo são propostos dois novos filtros espaciais. A primeira proposta é a extensão do máximo autofiltro espacial com o objetivo de torná-lo adequado tanto para fontes próximas quanto afastadas. A segunda proposta é baseada no espectro diferencial, devido ao comportamento observado na seção 3.3.1 e à robustez ao ruído que

ele apresentou. Então, os filtros propostos são avaliados em relação às suas respostas em frequência na ausência e na presença de ruído.

Logo após, este capítulo apresenta o novo processo de busca utilizado no estágio de refinamento de estimativas. Como escrito na seção 2.3, a estimação ML constitui um problema altamente complexo, multimodal, multivariado e não-linear. Então, o processo de busca deve ser realizado numericamente. Na otimização de problemas não-lineares, o processo de busca deve ser inicializado de maneira adequada para que ele possa convergir para o ótimo global. Ainda, é importante lembrar que a estimação DOA no SEAD Aprimorado já é realizada numericamente, com a geração de estimativas iniciais que são posteriormente refinadas. Assim, basta substituir no estágio de refinamento de estimativas o processo de busca baseado no algoritmo *Branch-and-Bound* por outro que apresente menor esforço computacional e produza estimativas em um intervalo contínuo.

Uma alternativa que possui tais características é o Método de Newton. Ele se destaca pela rápida convergência local e apresenta esforço computacional de ordem polinomial (STARER, 1990), podendo ser considerado um algoritmo eficiente (CARVALHO et al., 2001). Em (STARER, 1990; STARER; NEHORAI, 1992), o Método de Newton foi derivado considerando a reparametrização do DMLE apresentada na Equação (3.10), que está em função da matriz de predição linear  $\mathbf{B}$ .

No entanto, ele não considera a aplicação de filtragem espacial, cuja função-objetivo foi apresentada na Equação (3.36) e está em termos da matriz de direção filtrada  $\tilde{\mathbf{A}}$ . Além disso, não foi encontrada na literatura uma implementação do Método de Newton para o Estimador de Máxima-Verossimilhança apresentado na Equação (2.31) que, por sua vez, utiliza a matriz de direção  $\mathbf{A}$ .

Neste trabalho foi feito todo o desenvolvimento da proposta original do Método de Newton em (STARER, 1990; STARER; NEHORAI, 1992), denominado Newton/B/sem filtro, e os mesmos resultados foram alcançados. Isso forneceu suporte a três novos desenvolvimentos: Newton/B/com filtro, Newton/A/sem filtro e Newton/A/com filtro.

Porém, para problemas com um número muito grande de variáveis, o Método de Newton possui alto esforço computacional, pois ele precisa calcular e *inverter* a matriz hessiana. Por outro lado, é possível diminuir esse esforço usando métodos quase-Newton. Um deles é denominado BFGS, sigla que contém as iniciais dos autores Charles George Broyden, Roger Fletcher, Donald Goldfarb e David Shanno (BROYDEN, 1970; FLETCHER, 1970; GOLDFARB, 1970; SHANNO, 1970). O BFGS, em vez de calcular a matriz hessiana, aproxima diretamente a sua inversa utilizando somente o vetor gradiente, apresentando um esforço computacional menor do que o Método de Newton. O custo disso é uma convergência mais lenta para o ótimo global. Nesse contexto, este capítulo apresenta dois novos desenvolvimentos: BFGS/A/sem filtro e BFGS/A/com filtro.

Contudo, a aplicação do BFGS no SEAD Aprimorado mostrou resultados melhores

que os do Método de Newton. Assim, embora sejam abordadas aqui três propostas do Método de Newton e duas do BFGS, este capítulo tem como proposta final substituir o processo de busca do SEAD Aprimorado pelo BFGS/A/sem filtro ou pelo BFGS/A/com filtro.

Por último, neste capítulo é proposta uma pequena alteração no estágio de pré-seleção. Essa alteração remove a dependência do SEAD Aprimorado em relação ao cálculo adequado do limiar de amplitude, uma vez que esse limiar não é mais utilizado. Isso permite produzir melhores estimativas iniciais na presença de fontes afastadas. Entretanto, na presença de fontes próximas, observou-se um pequeno aumento no erro de estimação.

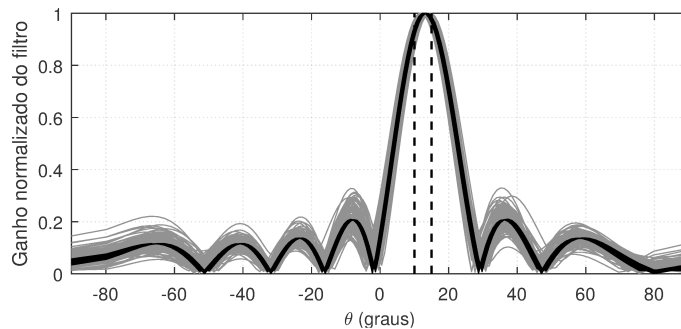
Finalmente, são realizadas as considerações finais deste capítulo.

## 4.1 Análise da Resposta em Frequência do Máximo Autofiltro Espacial

Esta seção avalia e discute a efetividade do máximo autofiltro espacial apresentado em (KRUMMENAUER et al., 2010) na seleção dos ângulos DOA verdadeiros. Foram realizados 200 experimentos da estimação desse autofiltro em três situações distintas considerando fontes decorrelacionadas situadas em: 1)  $10^\circ$  e  $15^\circ$  (fontes próximas); 2)  $10^\circ$  e  $45^\circ$  (fontes afastadas) e 3)  $10^\circ$ ,  $15^\circ$  e  $45^\circ$ . Todas as situações consideram a ausência e a presença de ruído na SNR de  $-10$  dB, um arranjo linear uniforme formado por  $K = 10$  sensores e  $N = 100$  *snapshots*.

A Figura 12 considera fontes próximas na ausência (linhas pretas) e na presença de ruído (linhas cinzas). Pode-se observar nessa figura que em todos os experimentos a banda de passagem do filtro foi corretamente detectada na região dos ângulos DOA verdadeiros, indicados pelas linhas tracejadas verticais. Adicionalmente, o filtro atenuou de maneira significativa os componentes de sinal com frequências angulares fora dessa região.

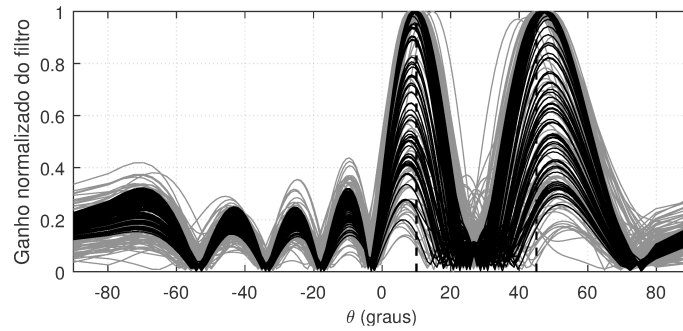
Figura 12 – Respostas em frequência sobrepostas do máximo autofiltro espacial considerando fontes situadas em  $10^\circ$  e  $15^\circ$  na ausência de ruído (linhas pretas) e na SNR de  $-10$  dB (linhas cinzas).



A situação de fontes afastadas é ilustrada na Figura 13. Pode-se verificar nessa figura

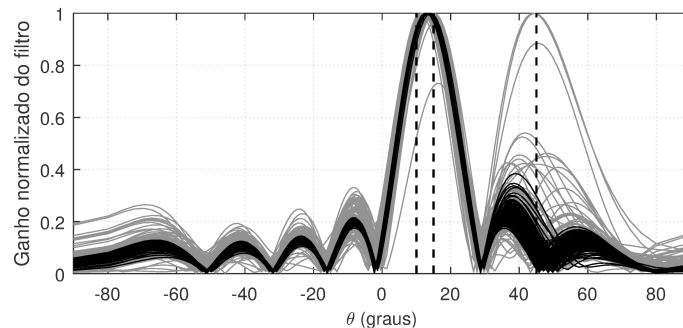
que o filtro apresentou duas bandas de passagem próximas dos ângulos DOA verdadeiros. É importante ressaltar que o máximo autofiltro espacial é um filtro multibanda. Entretanto, existem experimentos em que o filtro não detectou corretamente a banda de passagem para uma das fontes. Existem experimentos em que as bandas de passagem incluem as duas fontes de sinal, mas em outros não. Nesses experimentos, uma ou outra fonte é atenuada alternadamente de maneira considerável *mesmo na ausência de ruído*.

Figura 13 – Respostas em frequência sobrepostas do máximo autofiltro espacial considerando fontes situadas em  $10^\circ$  e  $45^\circ$  na ausência de ruído (linhas pretas) e na SNR de  $-10$  dB (linhas cinzas).



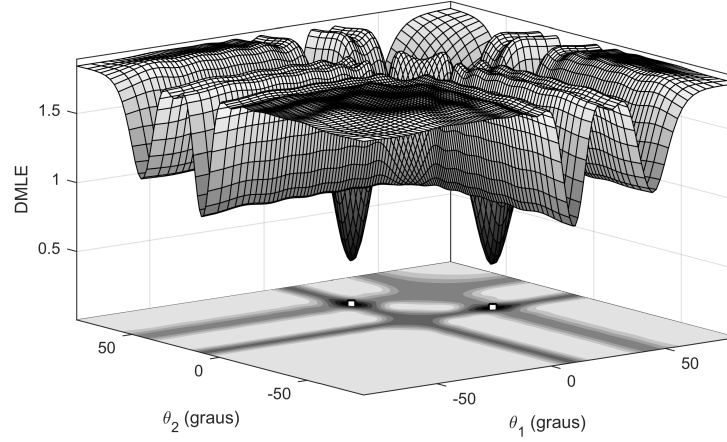
A situação de fontes situadas em  $\theta_1 = 10^\circ$ ,  $\theta_2 = 15^\circ$  e  $\theta_3 = 45^\circ$  ilustrada na Figura 14 é ainda pior do que na Figura 13. Considerando a ausência de ruído, mesmo que o filtro tenha detectado corretamente uma das bandas de passagem (na região de  $\theta_1$  e  $\theta_2$ ), ele atenuou consideravelmente a terceira fonte de sinal em todos os experimentos. No entanto, na SNR de  $-10$  dB, o filtro apresentou uma banda de passagem para a terceira fonte de sinal em alguns experimentos.

Figura 14 – Respostas em frequência sobrepostas do máximo autofiltro espacial considerando fontes situadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$  na ausência de ruído (linhas pretas) e na SNR de  $-10$  dB (linhas cinzas).



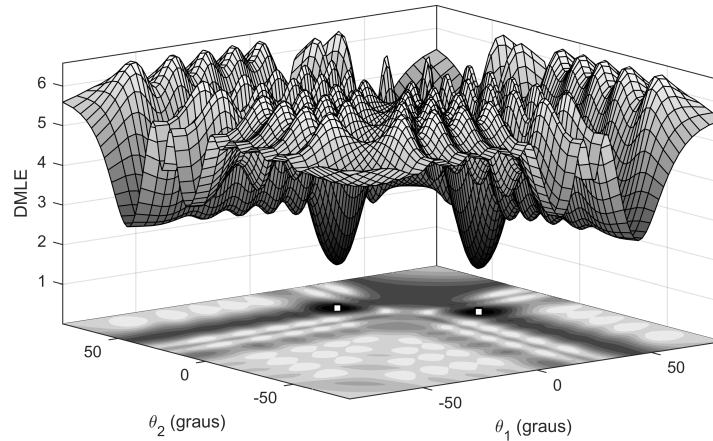
A consequência dessa forte atenuação de uma ou mais fontes de sinal é uma alteração significativa no espaço de soluções da função-objetivo do DMLE. Essa deformação pode deslocar o mínimo global. Para efeito de comparação, a Figura 15 ilustra a superfície DMLE sem a aplicação do máximo autofiltro espacial considerando fontes situadas em  $\theta_1 = 10^\circ$  e  $\theta_2 = 45^\circ$  na ausência de ruído.

Figura 15 – Superfície de um experimento do DMLE considerando fontes situadas em  $10^\circ$  e  $45^\circ$  e sinais descorrelacionados na ausência de ruído.



Pode-se observar na Figura 15 que a superfície apresentou mínimos globais bem definidos ao redor das posições reais das fontes de sinal, indicadas por um quadrado branco na parte inferior do gráfico. As Figuras 16 e 17 mostram as superfícies do DMLE usando o máximo autofiltro espacial em dois experimentos distintos, também na ausência de ruído.

Figura 16 – Superfície de um experimento do DMLE usando o máximo autofiltro espacial para fontes de sinal situadas em  $10^\circ$  e  $45^\circ$  e sinais descorrelacionados na ausência de ruído.

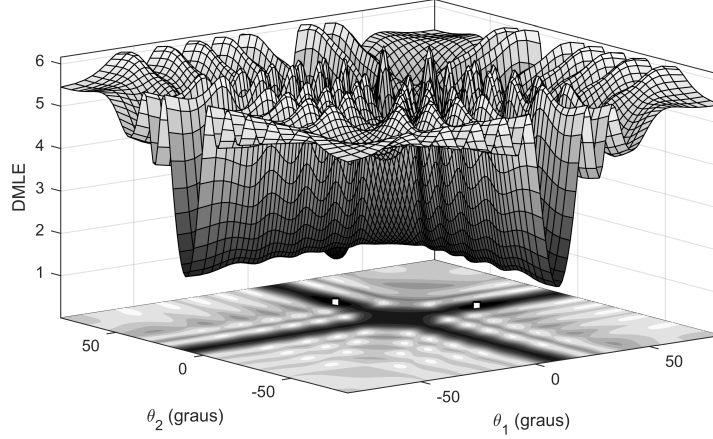


Pode-se verificar na Figura 16 que, devido à detecção correta das duas bandas de passagem, os mínimos globais estão bem definidos ao redor das posições reais das fontes de sinal. Porém, isso nem sempre ocorre. Na Figura 17, pode-se observar que o espaço de soluções foi consideravelmente alterado, mesmo que os mínimos globais estejam nas posições corretas. Isso é devido a uma forte atenuação da fonte de sinal posicionada em  $\theta_2 = 45^\circ$ .

Para entender a relação do afastamento entre as fontes de sinal e a seletividade do



Figura 17 – Superfície de outro experimento do DMLE usando o máximo autofiltro espacial para fontes de sinal situadas em  $10^\circ$  e  $45^\circ$  e sinais descorrelacionados na ausência de ruído.



autofiltro, considere a autodecomposição de  $\hat{\mathbf{R}}_I$  nos autovalores  $\hat{\lambda}_{(I,j)}$ ,  $j = 1, \dots, (K - \hat{L})$ , e seus respectivos autovetores  $\hat{\mathbf{e}}_{(I,j)}$ . Suponha também que os autovalores estejam em ordem decrescente e sejam pesos tal que  $\hat{\mathbf{R}}_I$  possa ser escrito como a soma de parcelas de acordo com o Teorema de Mercer ([MERCER, 1909](#)), como mostrado na Equação (4.1).

$$\hat{\mathbf{R}}_I = \hat{\lambda}_{(I,1)} \hat{\mathbf{e}}_{(I,1)} \hat{\mathbf{e}}_{(I,1)}^H + \hat{\lambda}_{(I,2)} \hat{\mathbf{e}}_{(I,2)} \hat{\mathbf{e}}_{(I,2)}^H + \dots + \hat{\lambda}_{[I,(K-\hat{L})]} \hat{\mathbf{e}}_{[I,(K-\hat{L})]} \hat{\mathbf{e}}_{[I,(K-\hat{L})]}^H \quad (4.1)$$

Os coeficientes do máximo autofiltro espacial são estimados pelo autovetor  $\hat{\mathbf{e}}_{(I,1)}$  correspondente ao maior autovalor de  $\hat{\mathbf{R}}_I$ . Na ausência de ruído, apenas as primeiras  $\bar{M}_I = \min [M, \text{posto}(\hat{\mathbf{R}}_I)]$  parcelas da Equação (4.1) são não-nulas. É importante ressaltar que essas parcelas correspondem ao subespaço de sinal e representam completamente a matriz  $\hat{\mathbf{R}}_I$ . Assim, pode-se concluir que embora a primeira parcela nessa equação seja a mais significativa, ela sozinha não representa todo o subespaço de sinal quando  $\bar{M}_I > 1$ .

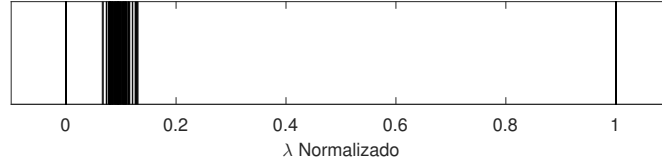
O motivo pelo qual  $\hat{\lambda}_{(I,1)} \hat{\mathbf{e}}_{(I,1)} \hat{\mathbf{e}}_{(I,1)}^H$  foi capaz de representar a correlação espacial na situação de fontes próximas pode ser encontrado em ([SILVA et al., 2013](#)). Assuma que  $\mathbf{R}_I$  seja a matriz construída de acordo com a Equação (3.43), porém usando a matriz teórica  $\mathbf{R}$ . Seguindo o desenvolvimento analítico de ([SILVA et al., 2013](#)), a diferença entre os dois maiores autovalores de  $\mathbf{R}_I$  considerando duas fontes descorrelacionadas com frequências angulares  $\omega_1$  e  $\omega_2$  pode ser obtida como apresentado na Equação (4.2).

$$\lambda_{(I,2)} - \lambda_{(I,1)} = 2 \sqrt{(\hat{L} + 1) + 2 \sum_{l=1}^{\hat{L}} k \cos [(\hat{L} - l + 1) (\omega_1 - \omega_2)]} \quad (4.2)$$

De maneira geral, a Equação (4.2) mostra que quanto mais próximas estão as fontes de sinal, maior é a diferença entre os dois maiores autovalores. Isso pode ser observado

nas Figuras 18 e 19, que mostram na ausência de ruído os espectros de autovalores de  $\hat{\mathbf{R}}_I$  considerando fontes próximas e afastadas, respectivamente.

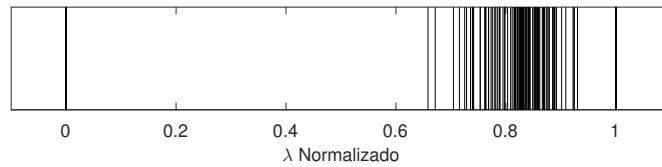
Figura 18 – Espectro de autovalores de  $\hat{\mathbf{R}}_I$  para fontes decorrelacionadas situadas em  $10^\circ$  e  $15^\circ$  na ausência de ruído.



Na Figura 18 pode-se verificar que  $\hat{\lambda}_{(I,1)} \gg \hat{\lambda}_{(I,2)}$  e a primeira parcela da Equação (4.1) aproxima suficientemente  $\hat{\mathbf{R}}_I$  de forma que as parcelas remanescentes do subespaço de sinal podem ser desconsideradas. Assim, como afirmado em (KRUMMENAUER et al., 2010), a resposta em frequência do autofiltro definido por  $\hat{\mathbf{e}}_{(I,1)}$  é capaz de preservar a banda de passagem dos ângulos DOA verdadeiros e atenuar as outras faixas de frequência.

Por outro lado, para a situação de duas fontes afastadas ilustrada na Figura 19,  $\hat{\lambda}_{(I,1)}$  e  $\hat{\lambda}_{(I,2)}$  se aproximam de forma que a parcela  $\hat{\lambda}_{(I,2)}\hat{\mathbf{e}}_{(I,2)}\hat{\mathbf{e}}_{(I,2)}^H$  não pode ser desconsiderada. Portanto, as duas primeiras parcelas da Equação (4.1) representam em conjunto o subespaço de sinal de  $\hat{\mathbf{R}}_I$  e a resposta em frequência do máximo autofiltro preserva as bandas de passagem de todos os ângulos DOA verdadeiros.

Figura 19 – Espectro de autovalores de  $\hat{\mathbf{R}}_I$  para fontes decorrelacionadas situadas em  $10^\circ$  e  $45^\circ$  na ausência de ruído.



Na presença de ruído branco, os  $\bar{M}_I$  maiores autovalores e seus autovetores correspondentes ainda correspondem ao subespaço de sinal de  $\hat{\mathbf{R}}_I$ , enquanto os *autoelementos*<sup>1</sup> remanescentes formam o subespaço de ruído. Para fontes próximas, é possível verificar que  $\hat{\lambda}_{(I,1)} \gg \hat{\lambda}_{(I,2)}$ . Dessa forma, o máximo autofiltro espacial preserva a banda de passagem dos ângulos DOA verdadeiros mesmo em baixas SNRs.

Entretanto, considerando fontes afastadas, não somente  $\hat{\lambda}_{(I,1)}$  se aproxima de  $\hat{\lambda}_{(I,2)}$  mas também os autovalores de ruído crescem à medida que a SNR diminui. Isso faz a primeira parcela da Equação (4.1) ser ainda menos significativa. Nessa situação, o autovetor  $\hat{\mathbf{e}}_{(I,1)}$  é inadequado na preservação das bandas de passagem dos ângulos DOA verdadeiros em baixa SNR.

Além disso, o Estimador de Máxima-Verossimilhança sem a aplicação de filtragem espacial é afetado pelo ruído à medida que a SNR diminui. Assim, o mínimo global pode

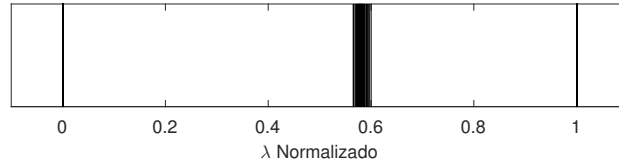
<sup>1</sup> Pares de autovalores/autovetores, do inglês *eigenelements* (HAWKES; NEHORAI; STOICA, 2001).



ser deslocado de sua posição verdadeira. Isso, em conjunto com a má representatividade do subespaço de sinal por parte do máximo autofiltro espacial, pode resultar em uma considerável degradação de desempenho.

Adicionalmente, como pode ser verificado na Figura 20, a correlação entre as fontes de sinal também afeta o espectro de autovalores. Quanto maior a correlação, menos significativos são os pesos  $\hat{\lambda}_{(I,2)}, \dots, \hat{\lambda}_{(I,\bar{M}_I)}$ , favorecendo o máximo autofiltro espacial.

Figura 20 – Espectro de autovalores de  $\hat{\mathbf{R}}_I$  para fontes correlacionadas situadas em  $10^\circ$  e  $45^\circ$  na ausência de ruído.



Em resumo, mesmo na presença de ruído, o máximo autofiltro espacial representa suficientemente bem o subespaço de sinal quando existe: 1) apenas uma fonte de sinal ou 2) duas ou mais fontes próximas. Para fontes afastadas e sinais descorrelacionados, o máximo autofiltro espacial não é capaz de representar de maneira suficiente o subespaço de sinal. Assim, uma ou mais fontes podem ser atenuadas. Por outro lado, para sinais correlacionados, a diferença entre  $\hat{\lambda}_{(I,1)}$  e  $\hat{\lambda}_{(I,2)}$  é maior, mesmo que as fontes estejam afastadas. Nessa situação, o máximo autofiltro espacial apresenta uma representatividade maior de todo o subespaço de sinal.

## 4.2 Propostas de Filtro Espacial

Esta seção apresenta duas propostas de filtros espaciais multibanda com o objetivo de melhorar a estimação DOA considerando fontes próximas e afastadas.

### 4.2.1 Proposta 1: filtragem dos *snapshots* usando *todo* o subespaço de sinal

Baseado na análise apresentada na seção 4.1, a solução natural é usar *todo* o subespaço de sinal para estimar um filtro que seja sensível a toda fonte. Para isso, é necessário associar em paralelo os autofiltros associados a cada autovetor pertencente ao subespaço de sinal.

Entretanto, esses autofiltros apresentam diferentes respostas em fase entre si e o filtro resultante pode não apresentar a resposta em frequência desejada. Então, para assegurar que todos os autofiltros tenham a mesma resposta em fase, cada autofiltro deve ser submetido ao processo de convolução mostrado na Equação (4.3).

$$\check{\mathbf{e}}_{(\hat{L},j)} = \text{conv} \left( \hat{\mathbf{e}}_{(\hat{L},j)}, \check{\mathbf{e}}_{(\hat{L},j)}^* \right) \quad (4.3)$$

em que:  $\text{conv}(\cdot)$  é o operador de convolução;  $\check{\mathbf{e}}_{(\hat{L},j)} = \check{\mathbf{I}}\hat{\mathbf{e}}_{(\hat{L},j)}$  é a versão espelhada de  $\hat{\mathbf{e}}_{(\hat{L},j)}$  e  $\check{\mathbf{I}}$  é a matriz de permutação apresentada na Equação (4.4).

$$\check{\mathbf{I}} = \begin{bmatrix} \mathbf{0} & & & 1 \\ & & 1 & \\ & \ddots & & \\ 1 & & & \mathbf{0} \end{bmatrix} \in (\hat{L} + 1) \times (\hat{L} + 1) \quad (4.4)$$

O filtro resultante é a soma dos  $\bar{M}_I$  autovetores  $\check{\mathbf{e}}_{(\hat{L},j)}$  referentes ao subespaço de sinal de  $\hat{\mathbf{R}}_I$ . Devido à convolução mostrada na Equação (4.3), esse filtro possui ordem  $(2\hat{L} - 1)$  e o vetor de coeficientes  $\hat{\mathbf{f}}^H = [\hat{f}_{2\hat{L}} \ \cdots \ f_0]$  é calculado como mostrado na Equação (4.5).

$$\hat{\mathbf{f}} = \hat{\lambda}_{(L,1)} \text{conv}(\hat{\mathbf{e}}_{(L,1)}, \check{\mathbf{e}}_{(\hat{L},1)}^*) + \hat{\lambda}_{(L,2)} \text{conv}(\hat{\mathbf{e}}_{(L,2)}, \check{\mathbf{e}}_{(\hat{L},2)}^*) + \cdots + \hat{\lambda}_{\bar{M}} \text{conv}(\hat{\mathbf{e}}_{\bar{M}}, \check{\mathbf{e}}_{(\hat{L},\bar{M})}^*) \quad (4.5)$$

Finalmente, a matriz  $\hat{\mathbf{H}}$  é substituída na Equação (3.36) pela matriz  $\hat{\mathbf{F}} \in \mathbb{C}^{(2\hat{L}+K) \times K}$  apresentada na Equação (4.6).

$$\hat{\mathbf{F}} = \begin{bmatrix} \hat{f}_{2\hat{L}} & \cdots & \hat{f}_1 & \hat{f}_0 & \mathbf{0} \\ & \ddots & & \ddots & \ddots \\ \mathbf{0} & \hat{f}_{2\hat{L}} & \cdots & \hat{f}_1 & \hat{f}_0 \end{bmatrix}^H \quad (4.6)$$

#### 4.2.2 Proposta 2: filtragem dos *snapshots* usando espectro diferencial

A seção 3.3.1 apresentou o espectro diferencial e como ele gera as estimativas iniciais para serem refinadas posteriormente. Entretanto, pode-se observar nas Figuras 7 a 9 que esse espectro gera picos proeminentes nas regiões dos ângulos DOA das fontes de sinal mesmo para valores de SNR de até  $-10$  dB. Além disso, pode-se verificar que o espectro diferencial se assemelha à resposta em magnitude de um filtro FIR.

Assim, a primeira tarefa no projeto do filtro é obter  $T$  amostras do espectro diferencial  $D(\omega)$  nas frequências  $\omega_t = [(2\pi t/T) - \pi], t = 0, \dots, (T - 1)$ , espaçadas igualmente em toda a abertura do arranjo. A partir disso,  $D(\omega)$  é considerado a densidade espectral de potência de um processo estocástico. Dessa forma, é possível calcular um filtro inverso cuja resposta em frequência é  $G(e^{j\omega_t})$ . Essa resposta, mostrada na Equação (4.7), permite reproduzir o comportamento do espectro diferencial.

$$S_d(\omega_t) = D(\omega_t) = |G(e^{j\omega_t})|^2 \sigma^2 = G(e^{j\omega_t})G^*(e^{j\omega_t})\sigma^2 \quad (4.7)$$

Portanto, para estimar a resposta em frequência do filtro inverso, considera-se  $G[t] = \sqrt{D(\omega_t)}$ . Assim, a resposta impulsiva  $g[n]$  correspondente é calculada a partir da Transformada Inversa de Fourier de  $G[t]$ .

Porém, como o máximo autofiltro espacial, esse filtro somente estima as bandas de passagem dos sinais e por isso não é necessária alta resolução. Assim, pode-se adotar uma pequena quantidade de amostras espectrais para representar completamente o espectro diferencial. Ao mesmo tempo, isso permite reduzir o esforço computacional, pois é necessário calcular uma decomposição em autovalores para cada  $D(\omega_t)$ . Entretanto,  $T$  deve ser escolhido cautelosamente com o objetivo de estabelecer um equilíbrio entre reduzir o esforço computacional e evitar o *aliasing* na resposta impulsiva do filtro.

Após múltiplas observações em diversas simulações, foi observado que  $g[n]$  possui até  $3K$  coeficientes complexos-conjugados diferentes de zero. Portanto, para  $K = 10$  sensores, utilizou-se  $T = 32$ . Adicionalmente, devido à simetria em  $g[n]$ , apenas a primeira metade é significativa, permitindo reduzir ainda mais a ordem do filtro ao truncar  $g[n]$  em  $n = \hat{L} = 7$ . É importante ressaltar que essa é a mesma ordem utilizada no máximo autofiltro espacial em (KRUMMENAUER et al., 2010).

Finalmente, o filtro possui o vetor de coeficientes  $\hat{\mathbf{g}}^H = [\hat{g}_{\hat{L}} \ \cdots \ g_0]$ . Esse vetor é usado na matriz  $\hat{\mathbf{G}} \in \mathbb{C}^{(\hat{L}+K) \times K}$  apresentada na Equação (4.8). Por sua vez,  $\hat{\mathbf{G}}$  substitui  $\hat{\mathbf{H}}$  na Equação (3.36).

$$\hat{\mathbf{G}} = \begin{bmatrix} \hat{g}_{\hat{L}} & \cdots & \hat{g}_1 & \hat{g}_0 & \mathbf{0} \\ & \ddots & & \ddots & \ddots \\ \mathbf{0} & & \hat{g}_{\hat{L}} & \cdots & \hat{g}_1 & \hat{g}_0 \end{bmatrix}^H \quad (4.8)$$

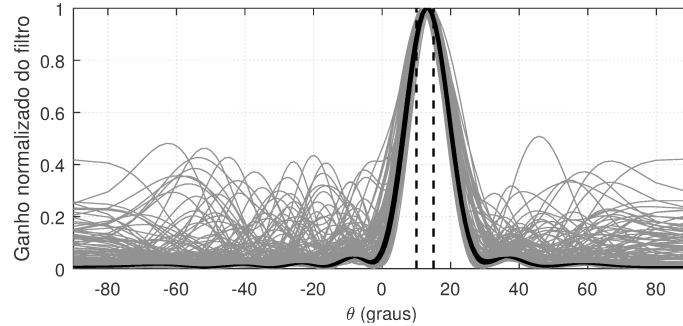
### 4.3 Análise da Resposta em Frequência dos Filtros Espaciais Propostos

As Figuras 21 a 23 sobrepõem 100 respostas em frequência dos filtros calculados a partir do subespaço de sinal e do espectro diferencial na ausência de ruído e outras 100 respostas com a SNR de  $-10$  dB. Assim como na seção 4.1, são consideradas três situações de posicionamento de fontes descorrelacionadas: 1) fontes próximas; 2) fontes afastadas e 3)  $10^\circ$ ,  $15^\circ$  e  $45^\circ$ . Além disso, é usado um arranjo linear uniforme formado por  $K = 10$  sensores e  $N = 100$  *snapshots*.

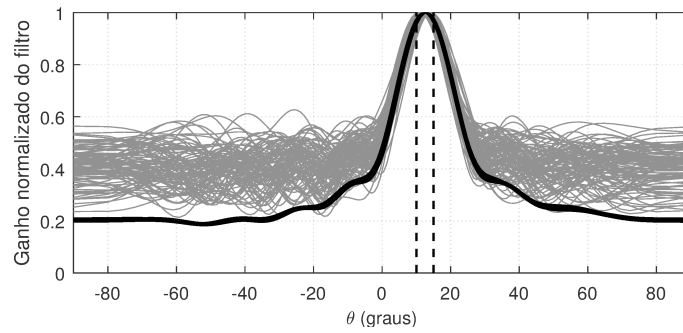
Pode-se observar na Figura 21 que na ausência de ruído ambos os filtros espaciais propostos detectaram a banda de passagem na região das fontes de sinal situadas em  $10^\circ$  e  $15^\circ$ . Entretanto, pode-se verificar que o ganho é maior na faixa de rejeição do filtro calculado a partir do espectro diferencial (Proposta 2). Isso ocorre devido à raiz quadrada em  $G[t] = \sqrt{D(\omega_t)}$  e à truncagem da ordem do filtro em  $n = \hat{L} = 7$ . Na presença de ruído, surgem picos espúrios no filtro calculado a partir do subespaço de sinal (Proposta 1) e diminui a atenuação na faixa de rejeição do filtro da Proposta 2.

A Figura 22 mostra a situação de fontes afastadas. Pode-se observar nessa figura

Figura 21 – Respostas em frequência sobrepostas dos filtros considerando fontes situadas em  $10^\circ$  e  $15^\circ$  na ausência de ruído (linhas pretas) e na SNR de  $-10$  dB (linhas cinzas).



(a) Proposta 1



(b) Proposta 2

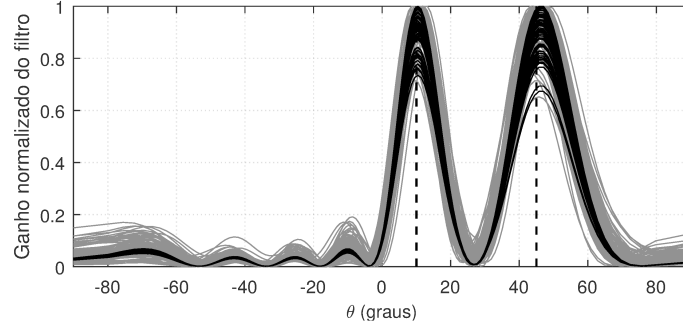
que na ausência de ruído os dois filtros propostos conseguiram detectar corretamente a banda de passagem para cada uma das fontes de sinal. Por outro lado, o filtro calculado a partir do espectro diferencial apresenta menor atenuação na faixa de rejeição, mas possui ganho maior nas bandas de passagem do que o filtro calculado a partir do subespaço de sinal. Exceto em alguns experimentos, o filtro da Proposta 2 apresentou comportamento semelhante considerando a SNR de  $-10$  dB.

A Figura 23 mostra as respostas em frequência dos filtros para fontes situadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$ . Pode-se verificar nessa figura que em todos os experimentos na ausência de ruído os dois filtros propostos foram capazes de detectar as bandas de passagem para as três fontes de sinal. No entanto, pode-se observar que em alguns experimentos na SNR de  $-10$  dB o filtro calculado a partir do espectro diferencial apresentou baixo ganho para a fonte situada em  $45^\circ$ .

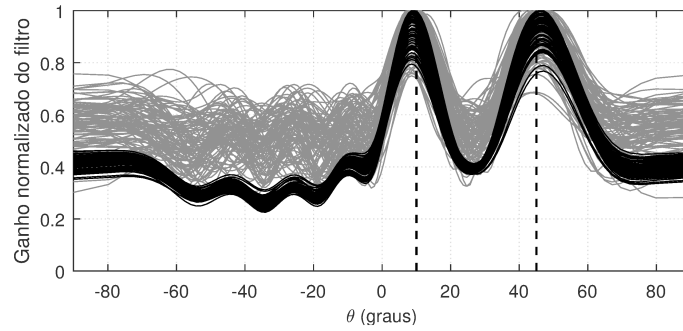
Finalmente, as Figuras 24 e 25 ilustram a superfície DMLE usando os dois filtros propostos neste trabalho. É importante ressaltar que todas as superfícies foram calculadas considerando o mesmo experimento da Figura 17.

Pode-se verificar na Figura 24 que o filtro calculado a partir do subespaço de sinal possui o mínimo global bem definido nas posições reais das fontes de sinal. Entretanto,

Figura 22 – Respostas em frequência sobrepostas dos filtros considerando fontes situadas em  $10^\circ$  e  $45^\circ$  na ausência de ruído (linhas pretas) e na SNR de  $-10$  dB (linhas cinzas).



(a) Proposta 1



(b) Proposta 2

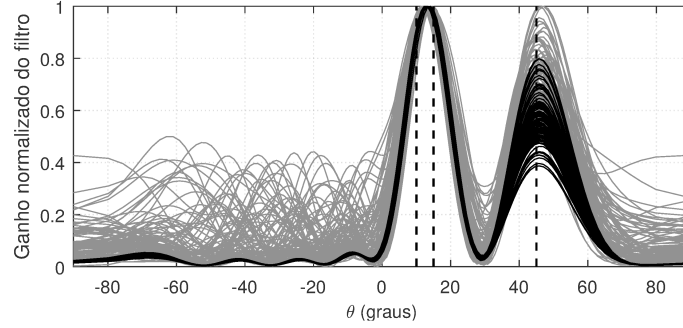
se comparado ao máximo autofiltro espacial, pode-se observar que ocorreu um aumento na quantidade de mínimos locais. Isso é devido ao aumento do tamanho do vetor de coeficientes de  $\hat{L}$  para  $2\hat{L}$ .

A superfície DMLE usando o filtro calculado a partir do espectro diferencial é ilustrada na Figura 25. Pode-se observar nessa figura que ela possui mínimos globais bem definidos e se assemelha muito à superfície DMLE sem a aplicação da filtragem espacial mostrada na Figura 15.

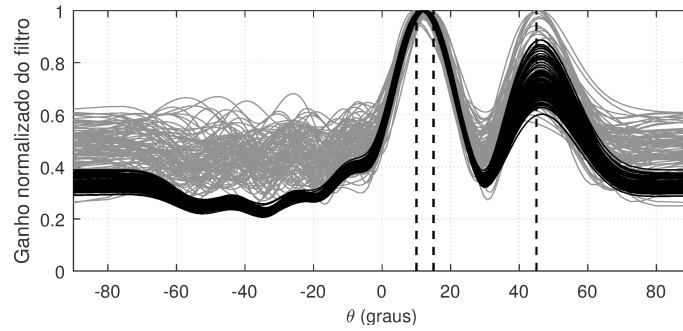
Adicionalmente, o filtro calculado a partir do subespaço de sinal preserva mais as bandas de passagens dos ângulos DOA verdadeiros do que o máximo autofiltro espacial. Entretanto, ele é também mais sensível a um fenômeno conhecido como troca de subespaço (do inglês *subspace swap*).

Segundo (HAWKES; NEHORAI; STOICA, 2001; SHAGHAGHI; VOROBYOV, 2015; THOMAS; SCHARF; TUFTS, 1995), a troca de subespaço ocorre quando a separação entre os autovalores de sinal e de ruído diminui a ponto de um ou mais autoelementos do subespaço de ruído serem considerados como pertencentes ao subespaço de sinal. Obviamente, quando ocorre essa troca é observada uma queda substancial no desempenho de estimação. De forma genérica, uma troca de subespaço resulta na Equação (4.9)

Figura 23 – Respostas em frequência sobrepostas dos filtros considerando fontes situadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$  na ausência de ruído (linhas pretas) e na SNR de  $-10$  dB (linhas cinzas).



(a) Proposta 1



(b) Proposta 2

(JOHNSON; ABRAMOVICH; MESTRE, 2008a; JOHNSON; ABRAMOVICH; MESTRE, 2008b).

$$\mathbf{e}_{(I,m)}^H \hat{\mathbf{E}}_{IN} \hat{\mathbf{E}}_{IN}^H \mathbf{e}_{(I,m)} > \mathbf{e}_{(I,m)}^H \hat{\mathbf{E}}_{IS} \hat{\mathbf{E}}_{IS}^H \mathbf{e}_{(I,m)}, \quad m = 1 \dots, \bar{M}_I \quad (4.9)$$

em que:  $\hat{\mathbf{E}}_{IN} = [\hat{\mathbf{e}}_{[I,(\bar{M}+1)]} \dots \hat{\mathbf{e}}_{[I,(\hat{L}+1)]}]$  e  $\hat{\mathbf{E}}_{IS} = [\hat{\mathbf{e}}_{(I,1)} \dots \hat{\mathbf{e}}_{(I,\bar{M})}]$  são respectivamente os autovalores de ruído e de sinal de  $\hat{\mathbf{R}}_I$  e o autovetor  $\mathbf{e}_{(I,m)}$  é o  $m$ -ésimo autovetor de sinal de  $\mathbf{R}_I$ .

As Figuras 26 a 28 mostram a probabilidade da troca de subespaço para cada autovetor de sinal. Para gerar essas figuras foram consideradas as três situações de posicionamento de fontes utilizadas neste trabalho e sinais descorrelacionados e correlacionados com fator  $\rho = 0,9$ . Cada ponto das curvas foi calculado a partir da média aritmética de  $E = 1000$  experimentos para cada valor de SNR variando de  $-15$  dB a  $15$  dB em passos de  $1,25$  dB.

A Figura 26 ilustra a situação de fontes próximas. É possível observar que a probabilidade de troca de subespaço aumenta significativamente para  $\hat{\mathbf{e}}_{(I,2)}$  em SNRs abaixo de  $-5$  dB para fontes descorrelacionadas. Considerando fontes correlacionadas, essa probabilidade sobe para SNRs abaixo de  $-2,5$  dB. Adicionalmente, o máximo autovetor de sinal  $\hat{\mathbf{e}}_{(I,1)}$  praticamente não apresentou troca de subespaço em  $-15$  dB tanto para sinais descorrela-

Figura 24 – Superfície de um experimento do DMLE usando o filtro calculado a partir do subespaço de sinal considerando fontes situadas em  $10^\circ$  e  $45^\circ$  e sinais descorrelacionados na ausência de ruído.

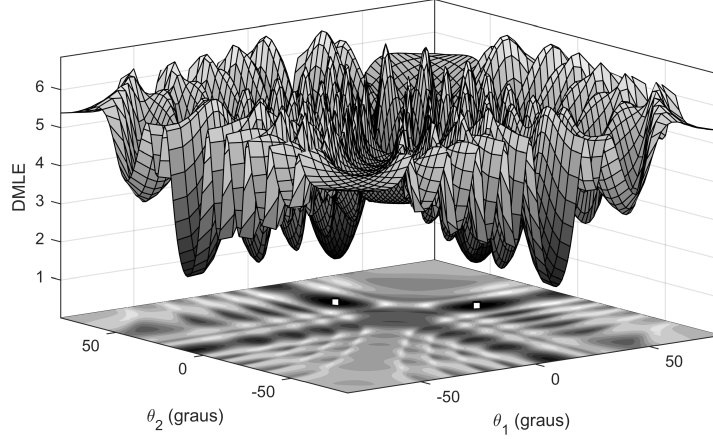
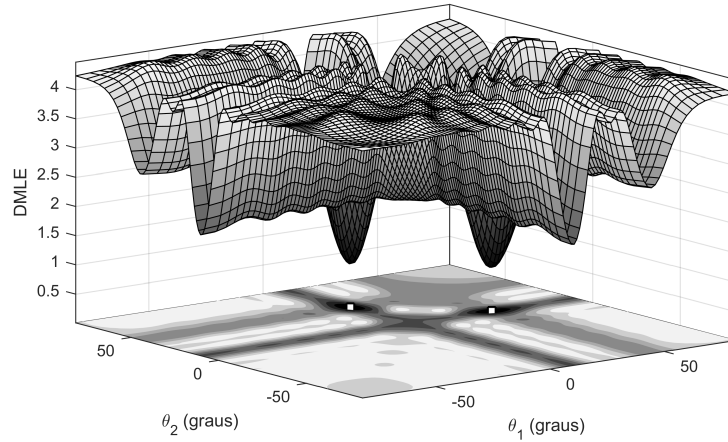


Figura 25 – Superfície de um experimento do DMLE usando o filtro calculado a partir do espectro diferencial considerando fontes situadas em  $10^\circ$  e  $45^\circ$  e sinais descorrelacionados na ausência de ruído.



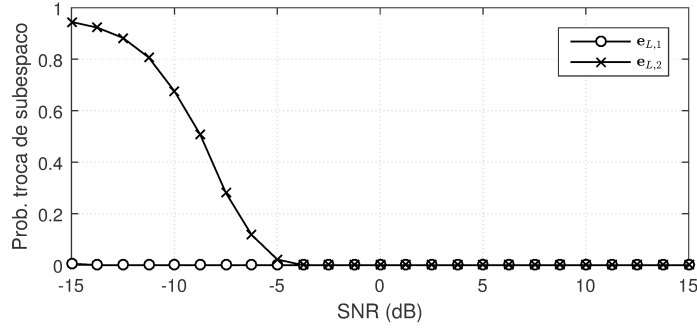
cionados quanto para sinais correlacionados. Nessa última condição, a probabilidade de troca de subespaço para esse autovetor foi nula em toda a faixa de SNRs avaliadas.

O resultado utilizando fontes afastadas é ilustrado na Figura 27. Pode-se observar nessa figura que, de maneira independente do fator de correlação entre os sinais, a probabilidade de troca de subespaço para ambos os autovetores de sinal aumenta quando a SNR assume valores abaixo de  $-11,25$  dB.

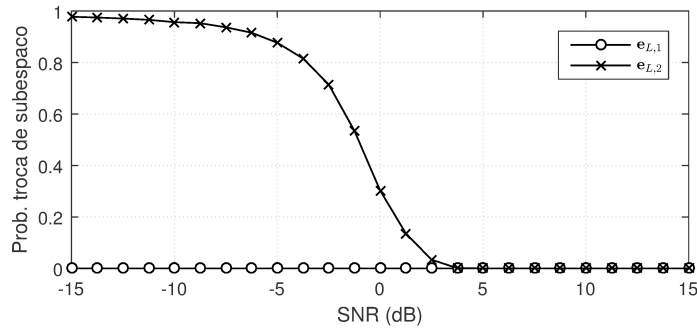
Finalmente, o resultado utilizando fontes situadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$  é mostrado na Figura 28. Pode-se verificar nessa figura que a troca de subespaço é mais provável de acontecer para  $\hat{e}_{(I,3)}$ . Considerando fontes descorrelacionadas, a probabilidade aumenta para valores de SNR menores do que  $-3,75$  dB. Para fontes correlacionadas, essa probabilidade aumenta quando a SNR é menor do que  $3,75$  dB. O segundo autovetor de sinal é afetado para



Figura 26 – Probabilidade de troca de subespaço para cada autovetor de sinal de  $\hat{\mathbf{R}}_I$  considerando fontes situadas em  $10^\circ$  e  $15^\circ$ .



(a) Fontes descorrelacionadas



(b) Fontes correlacionadas

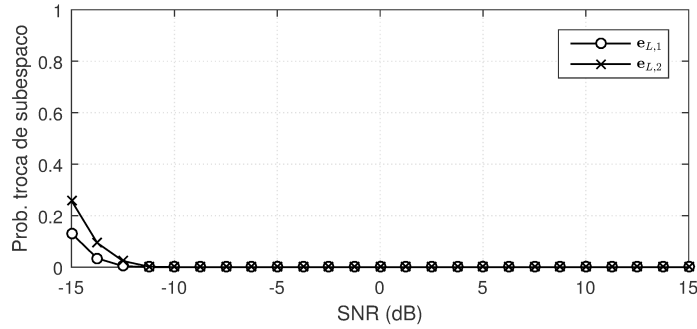
SNRs menores do que  $-11,25$  dB. O máximo autovetor de sinal apresentou probabilidade nula de troca de subespaço em toda a faixa de SNRs avaliada.

Em resumo, pode-se observar que o autovetor associado ao menor autovalor do subespaço de sinal é o mais susceptível à troca de subespaço, enquanto o máximo autovetor é o menos propenso a esse fenômeno. É importante ressaltar que esse é o motivo pelo qual o máximo autofiltro espacial apresenta maior robustez ao ruído na situação de fontes próximas. Além disso, é possível deduzir que o filtro calculado a partir do subespaço de sinal é mais susceptível ao ruído, pois ele utiliza todos os autovetores de sinal de  $\hat{\mathbf{R}}_I$ .

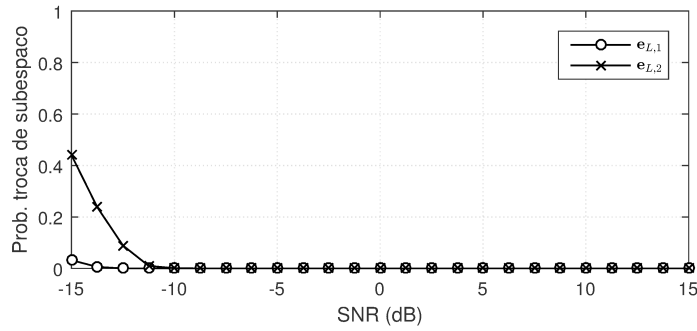
Por outro lado, o filtro calculado a partir do espectro diferencial apresenta robustez ao ruído relacionada ao aparecimento de picos espúrios nesse espectro. Pode-se verificar nas Figuras 7 a 9 que esse filtro apresenta grande robustez ao ruído em situação de fontes próximas, uma vez que o espectro gera picos de sinal com amplitude consideravelmente maior do que os picos de ruído. Adicionalmente, considerando fontes afastadas, na maioria dos casos ele ainda consegue distinguir os picos de sinal dos picos de ruído. Isso não ocorre com o máximo autofiltro espacial.



Figura 27 – Probabilidade de troca de subespaço para cada autovetor de sinal de  $\hat{\mathbf{R}}_I$  considerando fontes situadas em  $10^\circ$  e  $45^\circ$ .



(a) Fontes descorrelacionadas



(b) Fontes correlacionadas

## 4.4 Novo Estágio de Refinamento de Estimativas

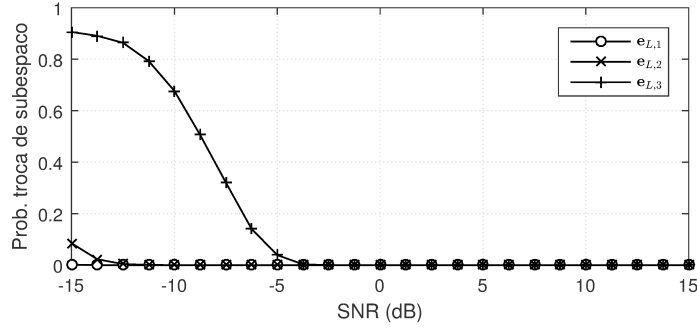
Esta seção apresenta cinco propostas para o estágio de refinamento de estimativas para diminuir o esforço computacional e melhorar o desempenho de estimação em altas SNRs no SEAD Aprimorado.

### 4.4.1 Fundamentação teórica

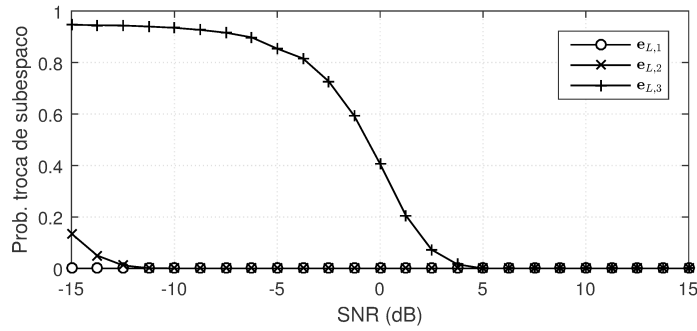
O Estimador de Máxima-Verossimilhança exige uma otimização. Otimização é um processo de busca por um ponto ótimo, que pode ser um mínimo ou um máximo global (CHAPRA; CANALE, 2010). Na seção 2.6, pode-se observar que a função-objetivo do DMLE apresentada na Equação (2.30) é não-linear,  $M$ -dimensional e apresenta diversos pontos mínimos (ALVES, 2004). Portanto, sua otimização dificilmente é um processo eficiente (LI; LU, 2007) e o vetor de estimativas  $\hat{\omega}$  deve ser calculado numericamente (STARER, 1990).

Esse processo de busca pelo mínimo global exige uma busca em  $M$  dimensões. A maneira mais óbvia de realizar isso é pela exaustiva busca em grade, onde a função-objetivo é avaliada em cada ponto. Esse método pode produzir garantidamente o ponto ótimo, desde que a grade possua resolução suficientemente alta, embora isso requiera intenso

Figura 28 – Probabilidade de troca de subespaço para cada autovetor de sinal de  $\hat{\mathbf{R}}_I$  considerando fontes situadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$ .



(a) Fontes descorrelacionadas



(b) Fontes correlacionadas

esforço computacional e deve ser evitada. Por outro lado, o ponto ótimo pode ser perdido se a resolução for suficientemente baixa. Dessa forma, é necessário estabelecer um ponto de equilíbrio entre o esforço computacional e a resolução da grade formada em torno do espaço de busca (STARER, 1990; ALVES, 2004).

A outra maneira é utilizar métodos de otimização global como o algoritmo genético (OTTERSTEN, 1989). Ao usar parâmetros escolhidos apropriadamente, esses métodos eventualmente convergem para o ponto ótimo de funções multimodais, como são chamadas as funções com diversos mínimos ou máximos locais. Contudo, a maioria desses métodos convergem de maneira razoavelmente rápida para a vizinhança do ponto ótimo, mas apresentam taxa de convergência local muito baixa (STARER; NEHORAI, 1992) e provavelmente fornecem respostas sub-ótimas (STARER, 1990). No contexto deste trabalho, esses métodos de otimização global se mostram ineficientes, pois, na maioria das situações, o estágio de pré-seleção do SEAD Aprimorado produz estimativas iniciais suficientemente precisas com um baixo esforço computacional.

Além disso, é possível utilizar algoritmos de otimização numérica local para refinar as estimativas iniciais fornecidas por um otimizador global (STARER; NEHORAI, 1992). No entanto, esses métodos são aplicáveis apenas em problemas unimodais. Para problemas multimodais assim como o MLE, esses algoritmos não podem ser usados sem as condições

*iniciais adequadas*. Assim, eles *devem* ser inicializados próximos do ponto ótimo (STARER, 1990). Alguns exemplos de algoritmos de otimização local são o Método de Newton e os métodos denominados quase-Newton.

Esses métodos podem utilizar a estratégia de busca em linha (ou *line search*), dentre outras. Nela, o algoritmo calcula uma *direção*  $\mathbf{p}_k$  e busca ao longo dessa direção a partir da estimativa atual  $\mathbf{x}_k$  para gerar uma nova estimativa  $\mathbf{x}_{(k+1)}$  mais próxima do ponto ótimo da função-objetivo  $f(\mathbf{x})$  (NOCEDAL; WRIGHT, 2006). A otimização é realizada com o objetivo de atingir um ponto onde  $f(\mathbf{x})$  é plana (YANG et al., 2005). Para isso, os métodos de busca em linha podem utilizar a informação que as derivadas parciais de primeira e segunda ordem fornecem, pois elas indicam se a função está sendo otimizada ou não. O vetor das derivadas parciais de primeira ordem (chamado *gradiente*)  $\mathbf{g}_k = \nabla f(\mathbf{x}_k)$  é igual a zero quando o processo de otimização atinge um ponto ótimo, que pode ser local ou global. A matriz das derivadas parciais de segunda ordem (chamada *hessiana*) mostra se esse ponto é um mínimo ou um máximo se ela for positiva ou negativa, respectivamente (CHAPRA; CANALE, 2010).

Em termos gerais, os métodos de busca em linha realizam uma expansão em Série de Taylor da função-objetivo em torno do ponto  $\mathbf{x}_k$ . Além disso, o Método de Newton considera uma expansão de segunda ordem (CHAPRA; CANALE, 2010), produzindo a Equação (4.10).

$$f(\mathbf{x}_{(k+1)}) \approx f(\mathbf{x}_k) + \mathbf{g}_k^T (\mathbf{x}_{(k+1)} - \mathbf{x}_k) + \frac{1}{2} (\mathbf{x}_{(k+1)} - \mathbf{x}_k)^T \mathbf{E}_k (\mathbf{x}_{(k+1)} - \mathbf{x}_k) \quad (4.10)$$

em que:  $\mathbf{E}_k$  é a *hessiana exata*,  $\nabla^2 f(\mathbf{x}_k)$ , da função-objetivo.

No ponto ótimo,  $\mathbf{g}_k = \mathbf{0}$ . Considerando  $\mathbf{E}_k$  uma matriz simétrica e não-singular, gera-se a Equação (4.11).

$$\mathbf{x}_{(k+1)} = \mathbf{x}_k + \mathbf{p}_k \quad (4.11)$$

em que a direção  $\mathbf{p}_k = -\mathbf{E}_k^{-1} \mathbf{g}_k$  também é chamada de *derivada diretiva*.

É importante ressaltar que a Equação (4.11) é calculada iterativamente para gerar novas estimativas e localizar o vetor de parâmetros  $\mathbf{x}$  que minimiza  $f(\mathbf{x})$  (STARER, 1990). Assim, ao final da  $k$ -ésima iteração,  $\mathbf{x}_k$  é corrigido e  $\mathbf{x}_{(k+1)}$  passa a ser  $\mathbf{x}_k$  na iteração seguinte. Esse processo é realizado até que se atinja uma quantidade predeterminada de iterações ou um ou mais critérios de convergência.

O Método de Newton converge quadraticamente em direção ao ponto ótimo, porém, existem três grandes inconvenientes ao calcular a Equação (4.11). Os dois primeiros são relativos ao esforço computacional: é necessário calcular todas as derivadas parciais de segunda ordem para construir a matriz  $\mathbf{E}_k$  e, depois, realizar sua inversão. Na prática, o

Método de Newton pode não ser muito útil para problemas com um grande número de variáveis (CHAPRA; CANALE, 2010). O terceiro inconveniente é que  $\mathbf{E}_k$  nem sempre é uma matriz positiva-definida. Portanto,  $\mathbf{p}_k$  nem sempre está em uma *direção descendente* (NOCEDAL; WRIGHT, 2006). Assim, a ideia discutida aqui de *minimizar*<sup>2</sup> localmente  $f(\mathbf{x})$  não é mais válida.

Existem estratégias para superar esses problemas e, dentre elas, pode-se citar os métodos quase-Newton (ou métodos de métrica variável). Eles são muito similares ao Método de Newton (CHAPRA; CANALE, 2010), mas não exigem o cálculo de derivadas parciais de segunda ordem (NOCEDAL; WRIGHT, 2006). Assim,  $\mathbf{E}_k^{-1}$  é aproximada iterativamente utilizando apenas o gradiente. Ao fazer isso, os métodos quase-Newton constroem um modelo com precisão suficiente para produzir convergência superlinear. Assim, esses métodos convergem mais lentamente para o ponto ótimo. Entretanto, como eles apresentam menor esforço computacional em relação ao Método de Newton, eles podem ser mais eficientes, realizando a otimização em um tempo menor. Adicionalmente, é importante ressaltar que nos métodos quase-Newton  $\mathbf{E}_k$  é, garantidamente, uma matriz definida-positiva.

O método quase-Newton mais popular é o algoritmo Broyden-Fletcher-Goldfarb-Shanno (BFGS) (NOCEDAL; WRIGHT, 2006). O BFGS faz um modelo quadrático da função objetivo no ponto  $\mathbf{x}_k$  como apresentado na Equação (4.12).

$$m_k(\mathbf{p}) = f_k + \mathbf{g}_k^T \mathbf{p} + \frac{1}{2} \mathbf{p}_k^T \mathbf{E}_k \mathbf{p} \quad (4.12)$$

Em vez de calcular  $\mathbf{E}_k$  completamente a cada iteração na Equação (4.12), ela é apenas atualizada levando em consideração a curvatura da função medida durante a última iteração, conforme o mostrado na Equação (4.13) (NOCEDAL; WRIGHT, 2006).

$$m_{(k+1)}(\mathbf{p}) = f_{(k+1)} + \mathbf{g}_{(k+1)}^T \mathbf{p} + \frac{1}{2} \mathbf{p}^T \mathbf{E}_{(k+1)} \mathbf{p} \quad (4.13)$$

Após impor algumas considerações, obtém-se a Equação (4.14), que calcula diretamente a *inversa* da hessiana aproximada (NOCEDAL; WRIGHT, 2006).

$$\mathbf{E}_k^{-1} = (\mathbf{I} - \rho_k \mathbf{s}_k \mathbf{y}_k^T) \mathbf{E}_{(k-1)}^{-1} (\mathbf{I} - \rho_k \mathbf{y}_k \mathbf{s}_k^T) + \rho_k \mathbf{s}_k \mathbf{s}_k^T \quad (4.14)$$

em que:  $\rho_k = \frac{1}{\mathbf{y}_k^T \mathbf{s}_k}$ ,  $\mathbf{s}_k = \mathbf{x}_{(k-1)} - \mathbf{x}_k$ ,  $\mathbf{y}_k = \mathbf{g}_{(k-1)} - \mathbf{g}_k$  e os termos com o índice  $(k-1)$  são aqueles calculados durante a iteração anterior. Além disso, na primeira iteração assume-se  $\mathbf{E}_{(k-1)}^{-1} = \mathbf{I}$ .

Entretanto, é uma prática comum ajustar o tamanho da correção dos parâmetros calculada por  $\mathbf{p}_k$ . Assim, tanto para o Método de Newton, quanto para os métodos quase-

<sup>2</sup> Observe que não é necessário distinguir minimização e maximização, pois, sem perda de generalidade, minimizar  $f(\mathbf{x})$  é equivalente a maximizar  $-f(\mathbf{x})$  (YANG et al., 2005).

Newton, a fórmula utilizada é a apresentada na Equação (4.15) (NOCEDAL; WRIGHT, 2006).

$$\mathbf{x}_{(k+1)} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \quad (4.15)$$

em que:  $\alpha_k > 0$  é chamado de *comprimento do passo* e é um valor real e escalar.

A convergência para um ponto ótimo depende fortemente de como  $\alpha_k$  e  $\mathbf{p}_k$  são calculados. O processo de cálculo de  $\alpha_k$  deve ser realizado rapidamente e o comprimento do passo deve minimizar substancialmente a função-objetivo. A escolha ideal seria o minimizador global da função univariada mostrada na Equação (4.16) (NOCEDAL; WRIGHT, 2006).

$$\phi(\alpha) = f(\mathbf{x}_k + \alpha \mathbf{p}_k) \quad (4.16)$$

Por outro lado, o cálculo desse valor apresenta um elevado esforço computacional. Mesmo o cálculo de um minimizador local de  $\phi(\alpha)$  com precisão moderada geralmente exige muitas avaliações da função-objetivo e do seu gradiente. Na prática, a estratégia utilizada é realizar uma busca em linha *inexata* para calcular  $\alpha_k$  que minimize  $f(\mathbf{x})$  a um custo mínimo.

Tipicamente, uma sequência de valores candidatos para  $\alpha_k$  é avaliada e o cálculo é interrompido quando certas condições são satisfeitas. Esse cálculo é realizado em dois estágios: uma fase de *bracketing*<sup>3</sup> encontra um intervalo contendo valores desejáveis para  $\alpha_k$  e uma fase de biseção ou interpolação que calcula um bom comprimento de passo dentro desse intervalo (NOCEDAL; WRIGHT, 2006).

Existem várias condições que mostram que  $\alpha_k$  não *precisa* se situar próximo dos minimizadores de  $\phi(\alpha_k)$ . Uma condição simples é impor ao comprimento de passo uma minimização da função-objetivo, ou seja,  $f(\mathbf{x}_k + \alpha_k \mathbf{p}_k) < f(\mathbf{x})$ . Porém, essa condição não é suficiente para produzir convergência para o minimizador de  $\phi(\alpha_k)$ .

Assim, é necessário aplicar uma condição de *minimização suficiente*. Essa condição estipula que  $\alpha_k$  minimize suficientemente a função-objetivo de acordo com o mostrado na Equação (4.17), às vezes chamada de *condição de Armijo* (NOCEDAL; WRIGHT, 2006).

$$f(\mathbf{x}_k + \alpha_k \mathbf{p}_k) \leq f(\mathbf{x}) + c_1 \alpha_k \mathbf{g}_k^T \mathbf{p}_k \quad (4.17)$$

em que:  $c_1 \in (0, 1)$  e, na prática, escolhe-se um valor bem pequeno, como  $c_1 = 10^{-4}$ . Em outras palavras, a minimização na função-objetivo deve ser proporcional ao comprimento do passo e à derivada direcional  $\mathbf{g}_k^T \mathbf{p}_k$ .

Contudo, essa condição não é suficiente para garantir que o algoritmo faça um progresso razoável, pois  $\alpha_k$  pode assumir valores *muito* pequenos. Para evitá-los, aplica-se

<sup>3</sup> As traduções que mais condizem com o significado de *bracketing* neste contexto são *cercar*, *envolver* ou *incluir* no sentido de determinar um intervalo que contenha o valor ótimo de  $\alpha_k$ .

a condição de curvatura, apresentada na Equação (4.18) (NOCEDAL; WRIGHT, 2006).

$$\nabla f(\mathbf{x}_k + \alpha_k \mathbf{p}_k)^T \mathbf{p}_k \geq c_2 \mathbf{g}_k^T \mathbf{p}_k \quad (4.18)$$

em que:  $c_2 \in (c_1, 1)$ . Essa condição garante que a inclinação de  $\phi(\alpha_k)$  seja  $c_2$  vezes maior do que a inclinação inicial  $\phi'(0)$ . Tipicamente, quando utilizado o Método de Newton ou algum método quase-Newton,  $c_2 = 0,9$ .

As Equações (4.17) e (4.18) são conhecidas em conjunto como as *condições fracas de Wolfe*. Alternativamente, nas *condições fortes de Wolfe*, a Equação (4.18) é substituída pela Equação (4.19).

$$\left| \nabla f(\mathbf{x}_k + \alpha_k \mathbf{p}_k)^T \mathbf{p}_k \right| \geq c_2 \left| \mathbf{g}_k^T \mathbf{p}_k \right| \quad (4.19)$$

Assim, a única diferença entre essas condições de Wolfe é que são excluídos os pontos distantes dos pontos estacionários de  $\phi(\alpha_k)$ .

#### 4.4.2 Método de Newton para o DMLE sem filtragem espacial utilizando a matriz de predição linear

Em (STARER, 1990), foi derivado o Método de Newton para o DMLE sem filtragem espacial, Equação (3.10), utilizando a matriz  $\mathbf{B}$  contendo os coeficientes do polinômio de predição linear  $b(z)$ . O desenvolvimento completo é demonstrado no Apêndice A e resumido nesta seção. A fórmula para a correção do vetor de parâmetros está expressa na Equação (4.20).

$$\boldsymbol{\phi}_{(k+1)} = \boldsymbol{\phi}_k - \mathbf{E}_k^{-1} \mathbf{g}_k \quad (4.20)$$

em que:  $\boldsymbol{\phi}_{(k+1)}$  e  $\boldsymbol{\phi}_k$  são vetores reais com dimensões  $2(M+1) \times 1$ . Isso ocorre, pois a Equação (3.10) está em função do vetor complexo  $\hat{\mathbf{b}}_k$ . Assim, o gradiente  $\mathbf{g}_k$  e a hessiana  $\mathbf{E}_k$  devem ser calculados em função de  $\boldsymbol{\phi}_k$ , parâmetro que separa as partes real e imaginária de  $\hat{\mathbf{b}}_k$ , como apresentado na Equação (4.21).

$$\boldsymbol{\phi}_k = \left[ \Re(\hat{\mathbf{b}}_k) \quad \Im(\hat{\mathbf{b}}_k) \right]^T = \left[ \phi_0 \quad \cdots \quad \phi_{(2M+1)} \right]^T \quad (4.21)$$

O gradiente e a hessiana são definidos nas Equações (4.22) e (4.23), respectivamente.

$$\mathbf{g}_k = \begin{bmatrix} \frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_0} \\ \vdots \\ \frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_{(2M+1)}} \end{bmatrix} \quad (4.22)$$

$$\mathbf{E}_k = \begin{bmatrix} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_0 \partial \phi_0} & \cdots & \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_0 \partial \phi_{(2M+1)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_{(2M+1)} \partial \phi_0} & \cdots & \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_{(2M+1)} \partial \phi_{(2M+1)}} \end{bmatrix} \quad (4.23)$$

Os elementos de  $\mathbf{g}_k$  e  $\mathbf{E}_k$  são mostrados nas Equações (4.24) e (4.25).

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_a} = 2 \Re \text{tr} \left( \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_a} \right) \quad (4.24)$$

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_b \partial \phi_a} = 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_b} \tilde{\mathbf{R}} \frac{\partial \mathbf{B}}{\partial \phi_a} \right] \quad (4.25)$$

em que:  $0 \leq (a, b) \leq (2M + 1)$ .

Finalmente,  $\mathbf{B}^+ = \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H$  é a pseudo-inversa (ou a inversa generalizada) de Moore-Penrose e  $\tilde{\mathbf{R}} \triangleq \mathbf{R} - \sigma \mathbf{I}$ .  $\frac{\partial \mathbf{B}}{\partial \phi_a}$  é a matriz apresentada na Equação (4.26).

$$\frac{\partial \mathbf{B}}{\partial \phi_a} = \begin{cases} \mathbf{Z}_c & 0 \leq a \leq M, c = a \\ -i\mathbf{Z}_c & (M + 1) \leq a \leq (2M + 1), c = (a - M - 1) \end{cases} \quad (4.26)$$

em que:  $\mathbf{Z}_c$  é a matriz de deslocamento mostrada na Equação (4.27) e  $0 \leq c \leq (2M + 1)$ .

$$\mathbf{Z}_c = \begin{bmatrix} \mathbf{0}_{(M+1-c) \times (K-M)} \\ \mathbf{I}_{(K-M) \times (K-M)} \\ \mathbf{0}_{(c-1) \times (K-M)} \end{bmatrix} \quad (4.27)$$

Pode-se observar na Equação (4.20) que não existe o termo  $\alpha_k$  multiplicando a derivada diretiva. Isso ocorre, pois essa equação está indiretamente em termos do vetor dos parâmetros estimados  $\hat{\mathbf{b}}_k$ . Esse vetor, por sua vez, possui uma relação não-linear com o vetor de direção  $\hat{\mathbf{a}}_k(\boldsymbol{\omega})$ . Assim, não é possível realizar de maneira simples a multiplicação  $\alpha_k \mathbf{E}_k^{-1} \mathbf{g}_k$  e obter o vetor dos parâmetros equivalentes em  $\mathbf{A}$ .

Dessa forma, ao final da otimização deve-se calcular o vetor final de parâmetros  $\hat{\mathbf{b}}_{(k+1)}$  sabendo que  $\mathbf{b} = \bar{\phi}_1 + i\bar{\phi}_2$ , em que  $\bar{\phi}_1$  e  $\bar{\phi}_2$  são respectivamente a primeira e a segunda metades de  $\boldsymbol{\phi}$ . Adicionalmente, são extraídas as raízes do polinômio  $b(z)$  definido em  $\hat{\mathbf{b}}_{(k+1)}$  para obter as estimativas em termos das frequências angulares  $\hat{\boldsymbol{\omega}}_{(k+1)}$ . Em sequência, o vetor  $\hat{\boldsymbol{\theta}}_{(k+1)}$  de estimativas DOA é calculado considerando a relação entre  $\omega$  e  $\theta$  mostrada na seção 2.1.

#### 4.4.3 Proposta 1: Método de Newton para o DMLE com filtragem espacial utilizando a matriz de predição linear

O desenvolvimento desta proposta é demonstrado no Apêndice B e utiliza a função-objetivo do DMLE com filtragem espacial definida em (KRUMMENAUER et al., 2010) e apresentada na Equação (3.36). Entretanto, essa equação está em função da matriz de direção filtrada  $\tilde{\mathbf{A}}$ . Porém, é realizada na seção B.9 a reparametrização dessa equação e a nova função-objetivo é apresentada na Equação (4.28).

$$F_{\text{DMLEF2}} = \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z) \quad (4.28)$$

em que:  $\mathbf{P}_{\tilde{\mathbf{B}}} = \mathbf{I} - \mathbf{U}\mathbf{U}^+$ ,  $\mathbf{U}^+ = (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H$  é a pseudo-inversa (ou a inversa generalizada) de Moore-Penrose,  $\mathbf{U} \triangleq \mathbf{H} \mathbf{P}_{\tilde{\mathbf{B}}}^\perp$  e  $\mathbf{H}$  é a matriz de convolução do filtro utilizado.

A fórmula de correção do vetor de parâmetros é a Equação (4.20) e, por isso, esta proposta do Método de Newton não faz um ajuste do comprimento do passo. O vetor gradiente é apresentado na Equação (4.29).

$$\mathbf{g}_k = \begin{bmatrix} \frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_0} \\ \vdots \\ \frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_{(2M+1)}} \end{bmatrix} \quad (4.29)$$

em que os seus elementos são definidos na Equação (4.30).

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_a} = -2 \Re \text{tr} \left( \mathbf{P}_{\tilde{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_a} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right) \quad (4.30)$$

em que:  $0 \leq a \leq (2M + 1)$ . A matriz hessiana é mostrada na Equação (4.31).

$$\mathbf{E}_k = \begin{bmatrix} \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_0 \partial \phi_0} & \cdots & \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_0 \partial \phi_{(2M+1)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_{(2M+1)} \partial \phi_0} & \cdots & \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_{(2M+1)} \partial \phi_{(2M+1)}} \end{bmatrix} \quad (4.31)$$

e os seus elementos são calculados a partir da Equação (4.32).

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_b \partial \phi_a} = 2 \Re \text{tr} \left[ \left( \mathbf{U}^+ \right)^H \frac{\partial \mathbf{U}^H}{\partial \phi_b} \mathbf{P}_{\tilde{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_a} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right] \quad (4.32)$$

em que:  $0 \leq (a, b) \leq (2M + 1)$ . Finalmente,  $\frac{\partial \mathbf{U}}{\partial \phi_a}$  é a matriz apresentada na Equação (4.33).

$$\frac{\partial \mathbf{U}}{\partial \phi_a} = -\mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_a} \mathbf{P}_{\tilde{\mathbf{B}}}^\perp - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_a} \mathbf{B}^+ \quad (4.33)$$

É importante ressaltar que é necessário converter o vetor  $\hat{\boldsymbol{\phi}}_k$  para  $\hat{\mathbf{b}}_k$ , extrair as raízes do polinômio  $b(z)$  resultante e obter as estimativas em termos de  $\boldsymbol{\omega}$  e  $\boldsymbol{\theta}$  como apresentado na seção 4.4.2.



#### 4.4.4 Proposta 2: Método de Newton para o DMLE sem filtragem espacial utilizando a matriz de direção

Esta proposta utiliza o parâmetro a matriz de direção  $\mathbf{A}$  e o seu desenvolvimento está no Apêndice C. Dessa forma, é utilizada diretamente a função-objetivo do critério determinístico do Estimador de Máxima-Verossimilhança apresentada na Equação (2.31). A fórmula de correção do vetor de parâmetros está expressa na Equação (4.34).

$$\boldsymbol{\omega}_{(k+1)} = \boldsymbol{\omega}_k - \alpha_k \mathbf{E}_k^{-1} \mathbf{g}_k \quad (4.34)$$

em que:  $\boldsymbol{\omega}_{(k+1)}$  e  $\boldsymbol{\omega}_k$  são vetores reais com dimensões  $M \times 1$  e estão na forma  $\boldsymbol{\omega} = [\omega_1 \ \cdots \ \omega_M]^T$ . O gradiente e a hessiana são definidos nas Equações (4.35) e (4.36), respectivamente.

$$\mathbf{g}_k = \begin{bmatrix} \frac{\partial \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_1} \\ \vdots \\ \frac{\partial \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_M} \end{bmatrix} \quad (4.35)$$

$$\mathbf{E}_k = \begin{bmatrix} \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_1 \partial \omega_1} & \cdots & \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_1 \partial \omega_M} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_M \partial \omega_1} & \cdots & \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_M \partial \omega_M} \end{bmatrix} \quad (4.36)$$

Os elementos de  $\mathbf{g}_k$  e  $\mathbf{E}_k$  são mostrados nas Equações (4.37) e (4.38).

$$\frac{\partial \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = -2 \Re \text{tr} \left( \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \right) \quad (4.37)$$

$$\frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = 2 \Re \text{tr} \left( \mathbf{C} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \right) \quad (4.38)$$

em que:  $1 \leq (a, b) \leq M$  e  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$  é a pseudo-inversa (ou a inversa generalizada) de Moore-Penrose. Finalmente,  $\frac{\partial \mathbf{A}}{\partial \omega_a}$  é a matriz apresentada na Equação (4.39).

$$\frac{\partial \mathbf{A}}{\partial \omega_a} = \begin{bmatrix} 0 & & & \\ & j e^{j\omega_a} & & \\ & \vdots & & \\ \mathbf{0}_{K \times (a-1)} & & & \mathbf{0}_{K \times (M-a)} \\ & j(K-2) e^{j(K-2)\omega_a} & & \\ & j(K-1) e^{j(K-1)\omega_a} & & \end{bmatrix} \quad (4.39)$$

O resultado colateral da derivação do Método de Newton para a Equação (2.31) é a diminuição no número de derivadas parciais a serem calculadas. Isso ocorre, pois a

função-objetivo está em função do vetor de parâmetros reais  $\omega$ . Assim, não é necessário calcular as derivadas parciais em relação às partes real e imaginária do vetor de parâmetros.

Além disso, a otimização é realizada diretamente em função de  $\omega$ , permitindo o uso simples das Condições de Wolfe para o ajuste do comprimento de passo para aprimorar a convergência do Método de Newton. Ao final desse processo, é necessário calcular as estimativas em termos de  $\hat{\theta}_{(k+1)}$  observando a relação entre  $\omega$  e  $\theta$  mostrada na seção 2.1.

#### 4.4.5 Proposta 3: Método de Newton para o DMLE com filtragem espacial utilizando a matriz de direção

Esta proposta é demonstrada no Apêndice D e utiliza a função-objetivo do DMLE com filtragem espacial definida em (KRUMMENAUER et al., 2010), apresentada na Equação (3.36).

A fórmula de correção do vetor de parâmetros é a Equação (4.34). O gradiente é apresentado na Equação (4.35).

$$\mathbf{g}_k = \begin{bmatrix} \frac{\partial \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_1} \\ \vdots \\ \frac{\partial \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_M} \end{bmatrix} \quad (4.40)$$

e os seus elementos são calculados à partir da Equação (4.41).

$$\frac{\partial \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_a} = -2 \Re \text{tr} \left( \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \mathbf{P}_{\mathbf{A}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \right) \quad (4.41)$$

em que:  $1 \leq a \leq M$ . A matriz hessiana é mostrada na Equação (4.42).

$$\mathbf{E}_k = \begin{bmatrix} \frac{\partial^2 \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_1 \partial \omega_1} & \dots & \frac{\partial^2 \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_1 \partial \omega_M} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_M \partial \omega_1} & \dots & \frac{\partial^2 \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_M \partial \omega_M} \end{bmatrix} \quad (4.42)$$

e os elementos de  $\mathbf{E}_k$  e são definidos na Equação (4.38).

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} = 2 \Re \text{tr} \left[ \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z (\tilde{\mathbf{A}}^+)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_b} \mathbf{P}_{\mathbf{A}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \right] \quad (4.43)$$

em que:  $1 \leq (a, b) \leq M$  e  $\tilde{\mathbf{A}}^+ = (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^H$  é a pseudo-inversa (ou a inversa generalizada) de Moore-Penrose. Finalmente,  $\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}$  é apresentada na Equação (4.44).

$$\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} = \mathbf{H} \frac{\partial \mathbf{A}}{\partial \omega_a} \quad (4.44)$$

Ao final da otimização, é necessário calcular as estimativas em termos de  $\hat{\theta}_{(k+1)}$  observando a relação entre  $\omega$  e  $\theta$  mostrada na seção 2.1.

#### 4.4.6 Propostas 4 e 5: Método quase-Newton BFGS para o DMLE com e sem filtragem espacial utilizando a matriz de direção

Como as propostas 2 e 3 realizam a otimização em termos de  $\omega$ , é possível modificá-las para utilizar o BFGS, gerando as propostas 4 e 5. Nessas duas propostas, a fórmula de correção do vetor de parâmetros é a Equação (4.34).

Entretanto, os cálculos da hessiana exata mostrados nas Equações (4.36) e (4.42) não são utilizados. Em vez disso, calcula-se diretamente a inversa da hessiana aproximada apresentada na Equação (4.14) usando, respectivamente, os gradientes nas Equações (4.35) e (4.40).

Considerando o comprimento do passo  $\alpha_k = 1$ , as propostas 4 e 5 podem produzir respostas sub-ótimas, especialmente porque na primeira iteração,  $\mathbf{E}_k = \mathbf{I}$ . Consequentemente, o termo  $\alpha_k \mathbf{E}_k^{-1} \mathbf{g}_k$  na Equação (4.34) é igual a  $\mathbf{g}_k$ . Assim, o BFGS apresenta na primeira iteração o mesmo comportamento do Método *Steepest Descent*. Dessa forma, para que o BFGS não escape da vizinhança do ponto ótimo, é de *extrema importância* o uso das Condições de Wolfe para o ajuste do comprimento de passo.

### 4.5 Modificação no Estágio de Pré-seleção

Na seção 3.3.2.1 foi observado que o estágio de pré-seleção do SEAD Aprimorado é muito dependente do cálculo correto do limiar de amplitude. Se esse limiar estiver muito baixo, picos de ruído podem produzir uma ou mais estimativas iniciais. Por outro lado, se este limiar estiver muito alto, picos de sinal podem ser ignorados e, assim, não geram estimativas iniciais.

Tanto uma condição quanto outra produzem grandes desvios na estimação DOA que nem sempre podem ser corrigidos no estágio de refinamento. Embora os picos de sinal apresentem amplitude significativamente maior do que os picos de ruído, a determinação de um limiar capaz de separá-los se torna cada vez mais difícil à medida que a SNR diminui. Adicionalmente, isso pode ser ainda mais difícil quando existem fontes próximas e afastadas, pois alguns picos de sinal podem possuir amplitudes significativamente menores.

Além disso, observou-se que o processo de determinação das frequências angulares dos vales à esquerda e à direita dos ápices dos picos pode produzir resultados insatisfatórios. Nele, o vale é o ponto mais baixo do espectro diferencial entre dois picos ou os limites do intervalo de varredura quando não existe um ponto mínimo à esquerda ou à direita de um pico. Nesse processo, os vales podem ficar muito distantes do ápice de um pico, levando o estágio de pré-seleção a gerar estimativas iniciais sub-ótimas.

Dessa forma, este trabalho propõe três alterações no estágio de pré-seleção do SEAD. A primeira permite remover completamente o limiar de amplitude ao custo de um maior

esforço computacional. Até então, o estágio de pré-seleção utiliza *apenas* os picos maiores do que o limiar de amplitude e considera que eles possuem *no mínimo* uma fonte de sinal. A alteração proposta aqui *sempre* utiliza os  $M$  maiores picos, mas considera que podem existir até  $(M - 1)$  picos que não estejam associados a *nenhuma* fonte de sinal. Com isso, não é mais necessário distinguir entre picos de sinal e picos de ruído utilizando um limiar.

Supõe-se que cada um dos  $M$  picos corresponde a uma quantidade de  $c_m = 0$  e no máximo  $c_m = M$  fontes de sinal, onde  $m = 1, \dots, M$ . Então, são construídas as combinações  $\{c_1, c_2, \dots, c_M\}$  e, dessas, são consideradas apenas as combinações cujo somatório  $\sum_{m=1}^M c_m = M$ , produzindo  $C$  arranjos de distribuição de fontes de sinal por pico, de acordo com a Equação (4.45).

$$C = 1 + \sum_{Q=1}^{(M+1)} \frac{M!(M-1)!}{[(M-Q)!]^2 Q!(Q-1)!} \quad (4.45)$$

A segunda alteração proposta é limitar a distância máxima entre os vales e o ápice dos picos. Para cada um dos  $M$  picos é realizado um rápido processo em que se busca um ponto mínimo à esquerda e à direita do pico *ou* o ponto que seja imediatamente maior do que 5% da amplitude do ápice desse pico, o que ocorrer primeiro. A terceira e última alteração é aplicar a filtragem espacial. Para isso, é utilizado o procedimento ML apresentado na Equação (3.35) para selecionar a melhor combinação de estimativas iniciais.

Excluindo essas três alterações, o processo de geração de estimativas iniciais é exatamente o mesmo. Para ilustrar, é utilizado o mesmo exemplo da seção 3.3.2.1 usando o novo procedimento de geração de estimativas iniciais descrito nesta seção.

Considere  $M = 3$  fontes equipotentes, descorrelacionadas e posicionadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$  na ausência de ruído. Nessa situação, apenas dois dos três maiores picos estão associados às fontes de sinal, pois existem duas fontes próximas e uma afastada. As posições dos ápices desses picos são  $\{\theta_{(a,1)}, \theta_{(a,2)}, \theta_{(a,3)}\} = [-4,5^\circ \ 12,8^\circ \ 45,3^\circ]$  e as posições dos vales à esquerda e à direita de cada pico detectado são  $\{\theta_{(v,1,e)}, \theta_{(v,1,d)}, \theta_{(v,2,e)}, \theta_{(v,2,d)}, \theta_{(v,3,e)}, \theta_{(v,3,d)}\} \approx [-11^\circ \ 0^\circ \ 1^\circ \ 25^\circ \ 31^\circ \ 61^\circ]$ .

A partir das posições dos ápices e vales são gerados  $C = 10$  arranjos de distribuição de fontes por pico:  $\{0, 0, 3\}$ ,  $\{0, 1, 2\}$ ,  $\{0, 2, 1\}$ ,  $\{0, 3, 0\}$ ,  $\{1, 0, 2\}$ ,  $\{1, 1, 1\}$ ,  $\{1, 2, 0\}$ ,  $\{2, 0, 1\}$ ,  $\{2, 1, 0\}$  e  $\{3, 0, 0\}$ . Logo após, são construídos os vetores das estimativas iniciais como apresentado na Tabela 2. Para comparação com os resultados apresentados na seção 3.3.2.1, as estimativas foram avaliadas utilizando a função-objetivo do DMLE sem a aplicação da filtragem espacial apresentada na Equação (3.24). Nesse caso, as estimativas selecionadas são  $[9,9^\circ \ 15,8^\circ \ 45,3^\circ]$ , que são posteriormente refinadas pelo Método de Newton ou BFGS.

É importante observar que esse novo estágio de pré-seleção avalia um número maior

Tabela 2 – Combinações geradas no novo estágio de pré-seleção para o caso de fontes situadas em  $10^\circ$ ,  $15^\circ$  e  $45^\circ$  na ausência de ruído e avaliadas de acordo com o DMLE sem a aplicação de filtragem espacial.

Arranjo	Estimativas candidatas	DMLE
$\{0, 0, 3\}$	$[41,6^\circ \ 45,3^\circ \ 48,8^\circ]$	0,9993
$\{0, 1, 2\}$	$[12,8^\circ \ 41,6^\circ \ 48,8^\circ]$	0,1232
$\{0, 2, 1\}$	$[9,9^\circ \ 15,8^\circ \ 45,3^\circ]$	0,0016
$\{0, 3, 0\}$	$[9,9^\circ \ 12,8^\circ \ 15,8^\circ]$	0,5260
$\{1, 0, 2\}$	$[-4,5^\circ \ 41,6^\circ \ 48,8^\circ]$	1,0519
$\{1, 1, 1\}$	$[-4,5^\circ \ 12,8^\circ \ 45,3^\circ]$	0,1428
$\{1, 2, 0\}$	$[-4,5^\circ \ 9,9^\circ \ 15,8^\circ]$	0,5972
$\{2, 0, 1\}$	$[-6,1^\circ \ -3,3^\circ \ 45,3^\circ]$	1,0202
$\{2, 1, 0\}$	$[-6,1^\circ \ -3,3^\circ \ 12,8^\circ]$	0,6619
$\{3, 0, 0\}$	$[-6,1^\circ \ -4,5^\circ \ -3,3^\circ]$	1,3646

de estimativas candidatas e, por isso, possui um esforço computacional maior. A Tabela 3 apresenta o número de combinações avaliadas nesse estágio para  $1 \leq M \leq 8$  de acordo com a Equação (4.45).

Tabela 3 – Quantidade de combinações avaliadas no novo estágio de pré-seleção variando em função do número de fontes.

M	1	2	3	4	5	6	7	8
N.º aval.	1	3	10	35	126	462	1716	6435

## 4.6 Considerações Finais

Este capítulo apresentou três propostas de alteração ao estimador SEAD Aprimorado.

A primeira delas se refere à aplicação de filtragem espacial ao Estimador de Máxima-Verossimilhança. Para isso, foi realizado um estudo sobre a seletividade do máximo autofiltro espacial em relação à separação das fontes de sinal. Foi observado que esse filtro detecta bandas de passagem para fontes próximas. Entretanto, ele pode filtrar uma ou mais fontes de sinal afastadas entre si.

Isso pode levar a uma degradação significativa no desempenho de estimação, pois ele altera de maneira considerável o espaço de soluções do Estimador de Máxima-Verossimilhança. O motivo para essa alteração é que esse filtro utiliza apenas o autovetor associado ao maior autovalor da matriz de covariância. Assim, o filtro é calculado usando apenas parte da informação contida no subespaço de sinal de  $\hat{\mathbf{R}}_I$ , levando-o a atenuar uma ou mais fontes de sinal.

A partir desse estudo foi possível propor dois novos filtros espaciais. O primeiro é o resultado da extensão do método utilizado na proposta do máximo autofiltro espacial. Em vez de usar somente o máximo autovetor, o filtro proposto é calculado a partir de todos os autovetores do subespaço de sinal da matriz  $\hat{\mathbf{R}}_I$ . Esse filtro é capaz de detectar bandas de passagem para as fontes de sinal, mesmo quando existe grande separação entre elas. Por outro lado, esse filtro também é mais sensível ao ruído, pois autovetores de sinal ficam cada vez mais susceptíveis à troca de subespaço à medida em que os seus autovalores se aproximam dos autovalores de ruído.

O outro filtro espacial proposto utiliza o espectro diferencial apresentado na seção 3.3.1 como a densidade espectral de potência de um processo aleatório. Devido à simetria dos coeficientes do filtro, foi possível usar uma pequena quantidade de amostras espectrais de forma a diminuir o esforço computacional. Essa abordagem também permitiu desenvolver um filtro multibanda que detecta bandas de passagem para todas as fontes de sinal de maneira independente do afastamento entre elas.

A segunda proposta de alteração ao método SEAD é a maior contribuição deste trabalho. O estágio de refinamento de estimativas do SEAD Aprimorado é baseado no algoritmo *Branch-and-Bound* utilizado em problemas de otimização discreta. Por isso, esse estágio possui em sua maior parte resolução máxima em função do número de iterações e, consequentemente, o erro de estimação não diminui sempre que a SNR aumenta. Assim, o SEAD Aprimorado não é um estimador assintoticamente eficiente.

Uma solução viável para o SEAD Aprimorado é o uso em conjunto do estágio de pré-seleção e de algum algoritmo de otimização numérica local. Esses métodos são capazes de maximizar ou minimizar uma função-objetivo com diversos pontos ótimos desde que sejam inicializados do ponto ótimo global. Assim, enquanto o estágio de pré-seleção gera boas estimativas iniciais, o algoritmo de otimização local refina essas estimativas.

O Método de Newton é um desses algoritmos. Ele faz uma aproximação de segundo grau da função-objetivo utilizando o gradiente e a hessiana. Entretanto, considerando problemas com muitas variáveis, ele possui um elevado esforço computacional porque ele precisa calcular diversas derivadas parciais e necessita inverter a hessiana. Além disso, a hessiana nem sempre é uma matriz definida-positiva e, quando isso acontece, o Método de Newton não é capaz de otimizar a função-objetivo. Assim, surgiram métodos quase-Newton. Esses métodos, em vez de calcular a hessiana exata, calculam diretamente a inversa de uma aproximação quadrática da hessiana, evitando realizar os dois passos separadamente. Ao fazer isso, os métodos quase-Newton possuem taxa de convergência menor do que a do Método de Newton, mas possuem esforço computacional consideravelmente menor. Adicionalmente, a aproximação da hessiana também sempre é, garantidamente, uma matriz definida-positiva. Dentre os métodos quase-Newton existentes, o mais utilizado é o BFGS.

Nesse contexto, o Método de Newton já foi aplicado ao Estimador de Máxima-

Verossimilhança utilizando a matriz de predição linear, mas sem a aplicação de filtragem espacial. É importante ressaltar que não foi encontrada na literatura a aplicação do Método de Newton ao DMLE utilizando a matriz de direção. Então, a primeira atividade foi reproduzir esse desenvolvimento já feito. A partir disso, foi possível desenvolver analiticamente o Método de Newton para o DMLE com e sem aplicação de filtragem espacial e utilizando a matriz de direção e a matriz de predição linear.

Isso permitiu que também fosse desenvolvido o BFGS aplicado ao DMLE com e sem filtragem espacial usando a matriz de direção. Porém, o BFGS deve ser utilizado em conjunto com as Condições de Wolfe para ajuste do comprimento de passo para produzir as estimativas adequadas. De todas essas propostas, a mais importante é a do BFGS, pois é a que apresenta o menor esforço computacional. Além disso, não foi encontrada na literatura a aplicação do BFGS à estimação DOA.

Finalmente, a terceira alteração proposta ao SEAD Aprimorado é retirar o limiar de amplitude do estágio de pré-seleção. Embora esse limiar na maioria das vezes permite separar os picos de sinal dos picos de ruído, ele não abrange todas as situações possíveis. Dessa forma, podem existir situações em que os picos de sinal são desconsiderados durante o processo de geração de estimativas iniciais. Ainda, existem situações em que os picos de ruído podem ser considerados durante esse processo. Tanto uma situação quanto outra afeta a estimativa inicial de tal maneira que dificulta bastante o estágio de refinamento de estimativas. Além disso, foi proposto um procedimento de detecção dos vales para melhorar a geração de estimativas iniciais impondo um limite de distância entre os vales do ápice de um pico.

O Capítulo 5 apresenta os resultados obtidos neste trabalho em termos do erro de estimação e do esforço computacional. Nesse capítulo, o SEAD Aprimorado é avaliado antes e depois das alterações propostas e esses resultados são comparados com o MODEX Modificado.

## 5 RESULTADOS

Este capítulo apresenta os resultados obtidos neste trabalho. Esses resultados foram obtidos a partir de simulações numéricas realizadas utilizando o Mathworks MATLAB R2016b.

### 5.1 Simulações Numéricas

Nesta seção são comparados os desempenhos de estimação DOA usando o MODEX Modificado, o SEAD Aprimorado e a versão do SEAD com as alterações propostas, denominado SEAD Modificado. Enquanto o MODEX Modificado é avaliado considerando o processo apresentado na seção 3.2.2, o SEAD Aprimorado é avaliado antes e depois das alterações propostas. Todos os estimadores utilizam o critério determinístico do Estimador de Máxima-Verossimilhança (DMLE) com e sem a aplicação da filtragem espacial e os filtros espaciais abordados neste trabalho.

#### 5.1.1 Desempenho de estimação DOA

Como citado na seção 2.1.1, assume-se o modelo de sinal incondicional. Como utilizado na literatura, são realizadas  $E = 1000$  simulações de Monte-Carlo para cada valor de SNR, que varia de  $-15$  dB a  $15$  dB em passos de  $1,25$  dB. Considera-se um arranjo linear uniforme formado por  $K = 10$  sensores, amostrando  $N = 100$  *snapshots* de fontes de sinais equipotentes e de faixa estreita. Também são consideradas fontes de sinais descorrelacionados e correlacionados, cujos fatores de correlação são  $\rho = 0$  e  $\rho = 0,9$ , respectivamente. Para avaliar o desempenho de estimação de maneira mais abrangente, as simulações foram realizadas considerando uma fonte de sinal posicionada em  $\theta_1 = 10^\circ$  e a segunda fonte variando no intervalo  $-40^\circ \leq \theta_2 \leq 60^\circ$  em passos de  $2,5^\circ$ . O MODEX Modificado utiliza duas iterações na minimização do polinômio  $b(z)$ . O SEAD Aprimorado e o SEAD Modificado utilizam dez iterações no estágio de refinamento de estimativas.

Todos os estimadores são avaliados de acordo com a superfície de desempenho (ou a superfície REQM), calculada usando a Equação (2.24). Cada superfície REQM é comparada com o Limite de Cramér-Rao (CRB) apresentado na Equação (2.37) em termos da SNR de limiar. Essa SNR é definida como o último valor de SNR onde a REQM é menor do que duas vezes o CRB.

Além disso, são destacadas nas superfícies usando linhas tracejadas em preto-e-branco as curvas de desempenho considerando fontes próximas e afastadas situadas em  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$  e  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , respectivamente. Finalmente, todos dados



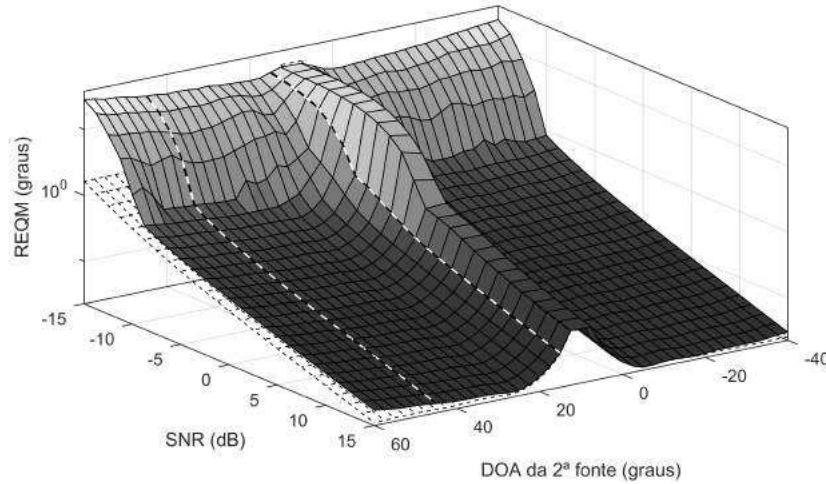
referentes a essas curvas estão disponíveis no Apêndice E.

#### 5.1.1.1 MODEX Modificado

As Figuras 29 a 32 ilustram as superfícies de desempenho do MODEX Modificado para duas fontes de sinais descorrelacionados.

A Figura 29 apresenta a superfície de desempenho para o algoritmo original do MODEX Modificado. Pode-se observar nessa figura que, à medida que as duas fontes de sinal se aproximaram, a variância do erro de estimação e o CRB aumentaram significativamente. Além disso, a variância do erro de estimação aumentou quando a SNR diminuiu. Como descrito na seção 2.2, os estimadores são avaliados em função da SNR de limiar. Para fontes afastadas, a SNR de limiar do MODEX Modificado variou entre  $-10$  dB e  $-8,75$  dB. Para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$  e  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , as SNRs de limiar foram  $-6,25$  dB e  $-10$  dB, respectivamente.

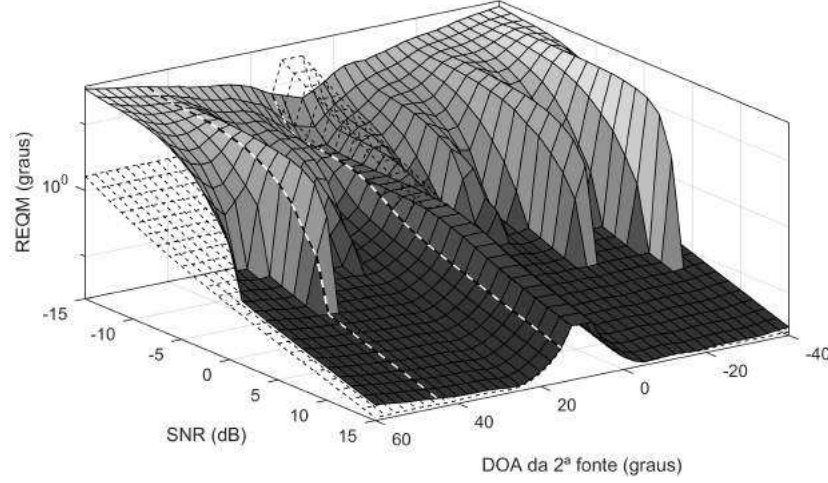
Figura 29 – Superfície de desempenho do algoritmo original do MODEX Modificado para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).



A Figura 30 considera o MODEX Modificado utilizando o máximo autofiltro espacial. Pode-se verificar nessa figura uma significativa redução da SNR de limiar na situação de fontes próximas e a variância do erro de estimação se manteve abaixo do CRB. Para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$ , não houve SNR de limiar na faixa de SNRs testadas. Por outro lado, para fontes afastadas ocorreu uma degradação significativa no desempenho, com um aumento considerável da SNR de limiar e, para  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , esse valor foi de  $3,75$  dB.

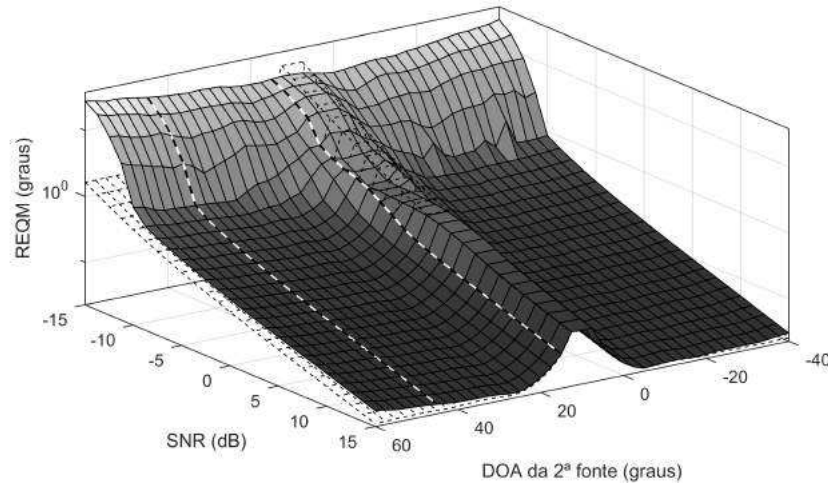
O desempenho do MODEX Modificado utilizando o filtro calculado a partir do subespaço de sinal é apresentado na Figura 31. Pode-se observar nessa figura que para fontes afastadas a SNR de limiar foi significativamente reduzida quando comparada à Figura 30. Na situação de fontes próximas, o erro de estimação é maior do que quando se aplica o

Figura 30 – Superfície de desempenho do MODEX Modificado utilizando o máximo autofiltro espacial para duas fontes decorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).



máximo autofiltro espacial, mas é consideravelmente menor do que quando não se aplica a filtragem espacial. Tanto para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$ , quanto para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$ , a SNR de limiar foi de  $-10$  dB.

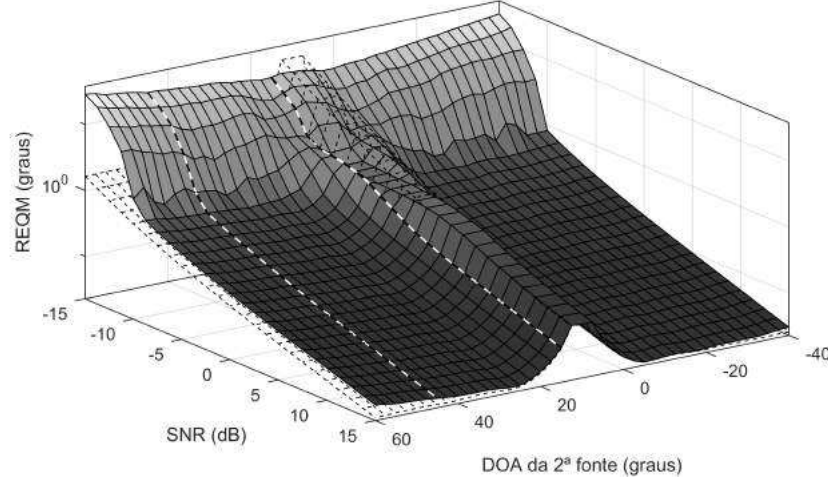
Figura 31 – Superfície de desempenho do MODEX Modificado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes decorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).



Finalmente, a Figura 32 mostra o desempenho do MODEX Modificado usando o filtro calculado a partir do espectro diferencial. Pode-se verificar nessa figura que a situação de fontes próximas apresentou menor erro de estimação do que o ilustrado na Figura 31, com SNR de limiar de  $-13,75$  dB para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$ . Por outro lado, a SNR de limiar para fontes afastadas foi de  $-10$  dB.

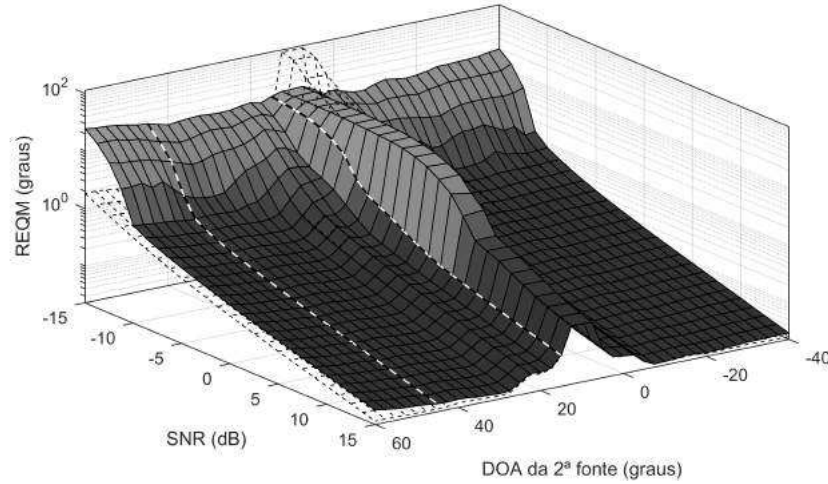
As Figuras 33 a 36 mostram as superfícies de desempenho do MODEX Modificado considerando duas fontes correlacionadas.

Figura 32 – Superfície de desempenho do MODEX Modificado utilizando o filtro calculado a partir do espectro diferencial para duas fontes decorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).



A Figura 33 considera o algoritmo original do MODEX Modificado. Pode-se observar nessa figura que a SNR de limiar para  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$  foi de  $-10$  dB e para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$  esse valor foi  $-5$  dB.

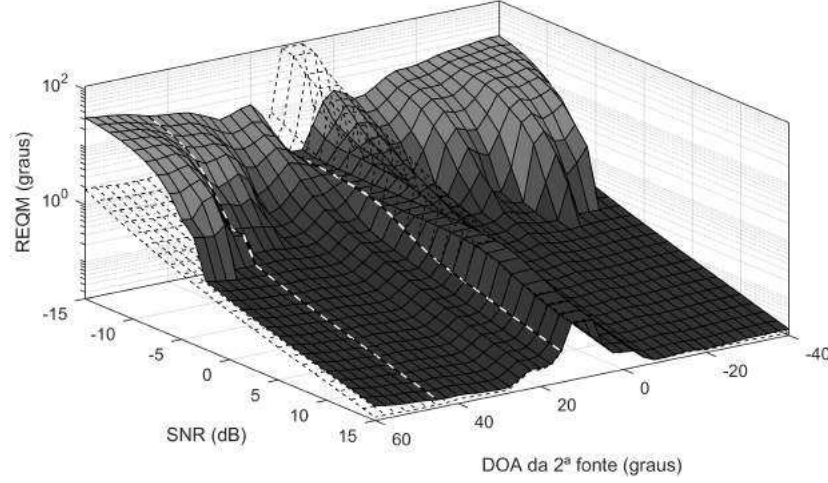
Figura 33 – Superfície de desempenho do algoritmo original do MODEX Modificado para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).



A Figura 34 ilustra a superfície de desempenho do MODEX Modificado utilizando o máximo autofiltro espacial. Como na Figura 30, pode-se verificar que a SNR de limiar para fontes afastadas aumentou consideravelmente e para,  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , esse valor foi de  $-3,75$  dB. No entanto, para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$ , o MODEX Modificado não apresentou SNR de limiar no intervalo das SNRs testadas.

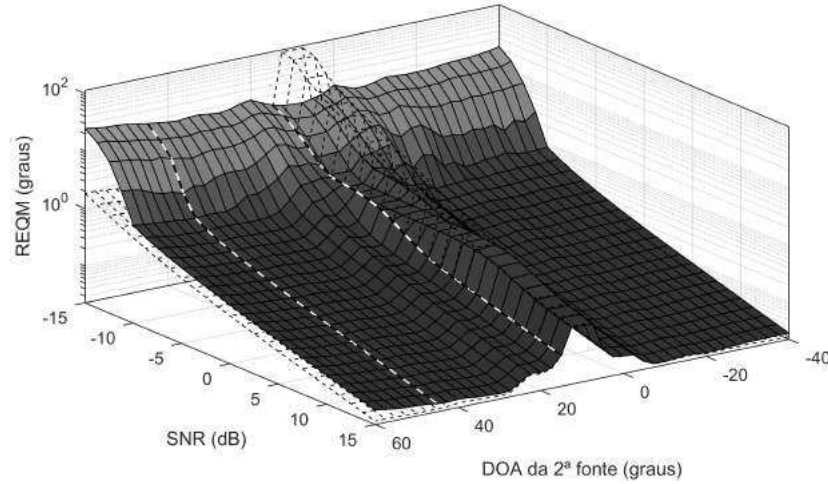
A Figura 35 considera o MODEX Modificado utilizando o filtro calculado a partir do subespaço de sinal. Pode-se observar nessa figura uma redução significativa na SNR

Figura 34 – Superfície de desempenho do MODEX Modificado utilizando o máximo autofiltro espacial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).



de limiar para fontes afastadas em relação ao máximo autofiltro espacial. A SNR de limiar para  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$  foi de  $-10$  dB, enquanto não houve SNR de limiar para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$  no intervalo de SNRs testadas.

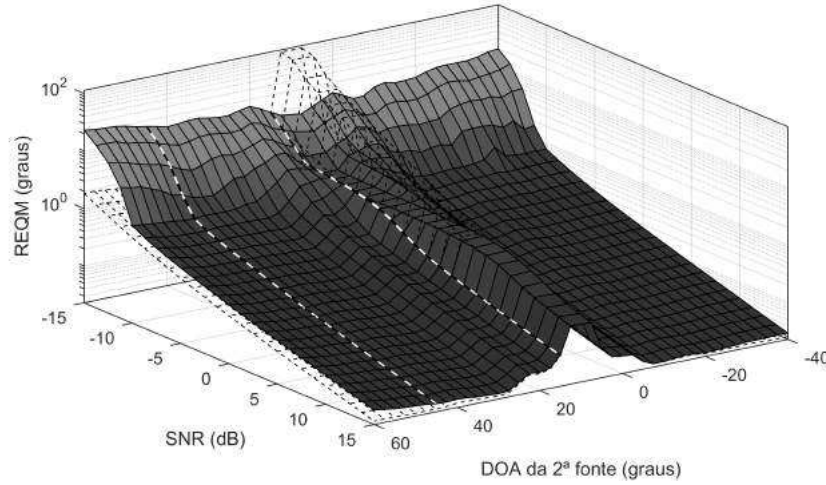
Figura 35 – Superfície de desempenho do MODEX Modificado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).



O desempenho de estimação do MODEX Modificado usando o filtro calculado a partir do espectro diferencial é ilustrado na Figura 36. Pode-se verificar nessa figura que, considerando  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$ , o erro de estimação se manteve abaixo do CRB até a SNR de  $-13,75$  dB e não houve SNR de limiar no intervalo das SNRs testadas. Para  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , a SNR de limiar foi de  $-10$  dB.



Figura 36 – Superfície de desempenho do MODEX Modificado utilizando o filtro calculado a partir do espectro diferencial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).



#### 5.1.1.2 SEAD Aprimorado

As Figuras 37 a 40 mostram o desempenho de estimação do SEAD Aprimorado considerando fontes de sinais descorrelacionados.

A superfície de desempenho considerando o algoritmo original do SEAD Aprimorado é ilustrada na Figura 37. Pode-se observar nessa figura que o desempenho de estimação para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$  foi de  $-11,25$  dB. No entanto, a SNR de limiar para fontes afastadas variou entre  $-8,75$  dB e  $-1,25$  dB e, para  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , esse valor foi de  $-7,50$  dB. Isso ocorreu devido ao estágio de pré-seleção, que considerou um ou mais picos de ruído como picos de sinal. Assim, em alguns experimentos, o estágio de refinamento utilizou estimativas ruins. Além disso, em altas SNRs, o SEAD Aprimorado apresentou erros de estimação ligeiramente maiores. Isso aconteceu, pois, como observado na seção 3.6, o estágio de refinamento é baseado no algoritmo *Branch-and-Bound*. Dessa forma, à medida que a SNR aumenta, são necessárias mais iterações para atingir o CRB.

A Figura 38 mostra a superfície de desempenho para o SEAD Aprimorado utilizando o máximo autofiltro espacial. Pode-se observar nessa figura que o desempenho de estimação para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$  foi ligeiramente melhorado e a SNR de limiar é  $-11,25$  dB. Entretanto, essa SNR poderia ter sido ainda menor. Isso não ocorreu porque o estágio de pré-seleção gerou estimativas iniciais ruins devido ao limiar de restrição de amplitude. Adicionalmente, para fontes afastadas pode-se observar uma degradação no desempenho de estimação e a SNR de limiar para  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$  foi de  $3,75$  dB.

As Figuras 39 e 40 ilustram as superfícies de desempenho para o SEAD Aprimorado usando os filtros calculados a partir do subespaço de sinal e do espectro diferencial, respectivamente. Em comparação à Figura 37, pode-se observar nessas figuras que a filtragem espacial praticamente não afetou o desempenho de estimação. Considerando

Figura 37 – Superfície de desempenho do algoritmo original do SEAD Aprimorado para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).

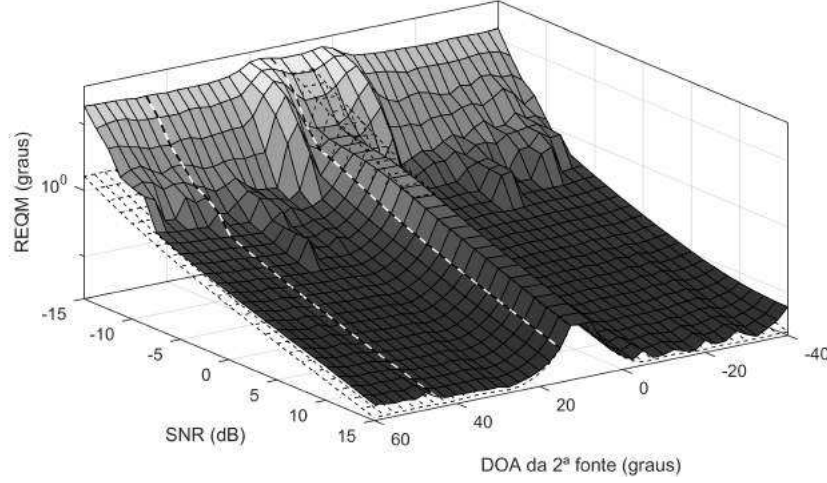
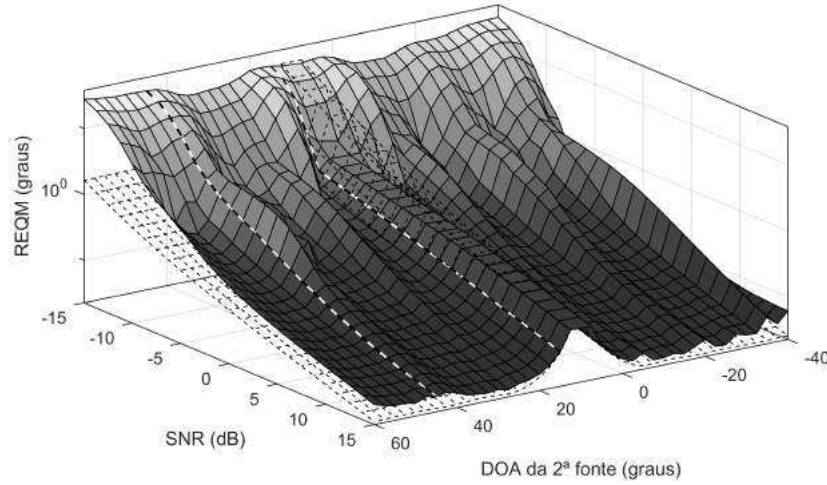


Figura 38 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o máximo autofiltro espacial para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).



fontes próximas, o erro de estimação foi sensivelmente menor, embora o valor da SNR de limiar tenha se mantido em  $-11,25$  dB. Por outro lado, para fontes afastadas, a SNR de limiar foi de  $-6,25$  dB. Essa SNR aumentou em  $1,25$  dB em relação ao algoritmo original do SEAD Aprimorado.

As Figuras 41 a 44 ilustram o desempenho de estimação do SEAD Aprimorado considerando fontes correlacionadas.

A superfície de desempenho do algoritmo original do SEAD Aprimorado é ilustrada na Figura 41. Pode-se observar nessa figura que o desempenho de estimação foi superior se comparado à situação de sinais descorrelacionados. Ainda, é possível verificar que o SEAD Aprimorado não apresentou o problema de seleção de picos de ruído considerando fontes afastadas, com SNR de limiar variando entre  $-11,25$  dB e  $-10$  dB. Para  $\{\theta_1, \theta_2\} =$

Figura 39 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes decorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).

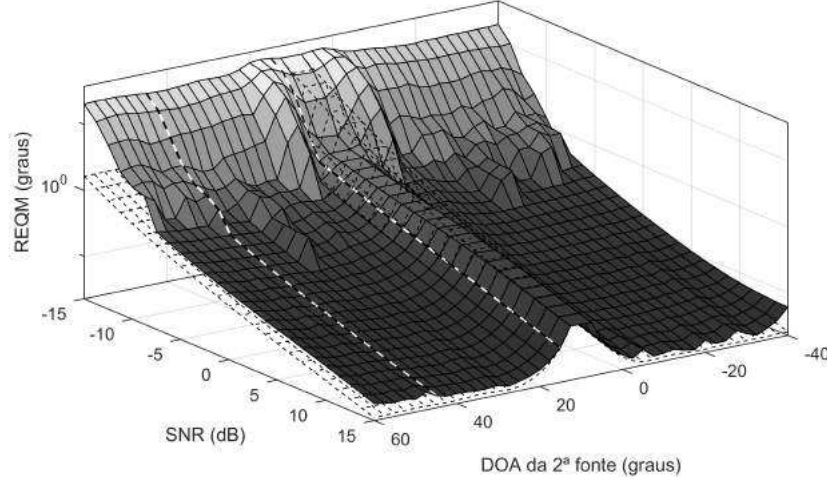
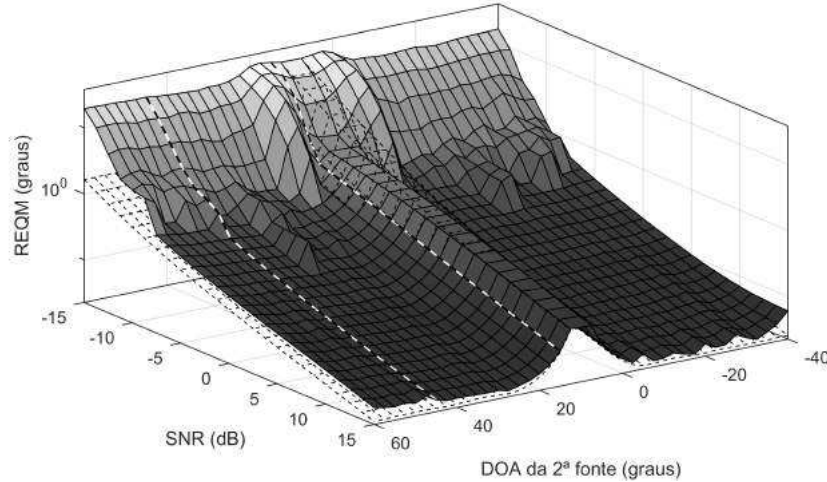


Figura 40 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o filtro calculado a partir do espectro diferencial para duas fontes decorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).



$[10^\circ \ 15^\circ]$ , esse valor foi de  $-13,75$  dB e, para  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , esse valor foi de  $-11,25$  dB. Além disso, pode-se verificar que, em algumas situações de afastamento de fontes devido ao estágio de refinamento de estimativas, o SEAD Aprimorado apresentou erros de estimação maiores em altas SNRs.

A Figura 42 apresenta a superfície REQM do SEAD Aprimorado utilizando o máximo autofiltro espacial. Como esperado, o desempenho de estimação foi degradado para fontes afastadas em relação ao algoritmo original do SEAD Aprimorado, com erros de estimação maiores do que quando não se aplica a filtragem espacial. Para  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , a SNR de limiar foi de  $-10$  dB. Para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$ , o SEAD Aprimorado não apresentou SNR de limiar no intervalo de SNRs testadas.

Figura 41 – Superfície de desempenho do algoritmo original do SEAD Aprimorado para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).

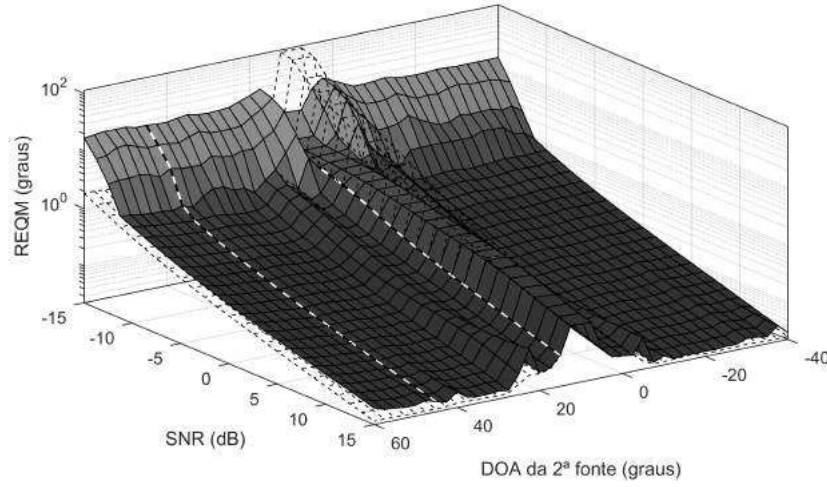
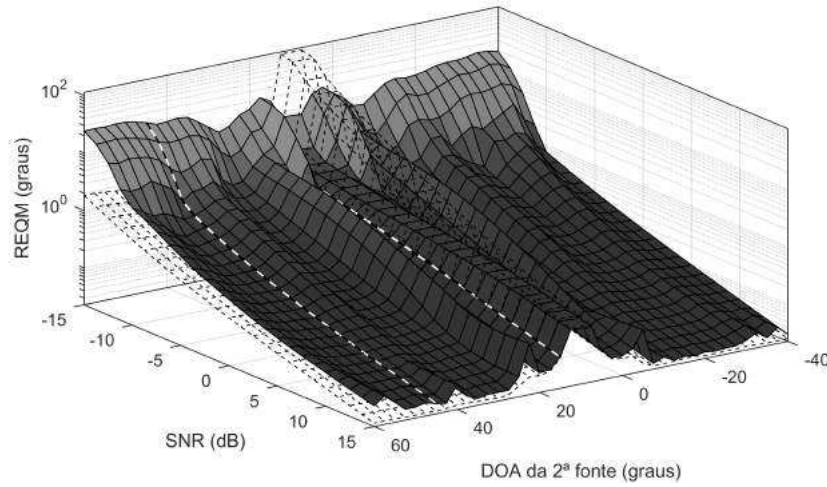


Figura 42 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o máximo autofiltro espacial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).

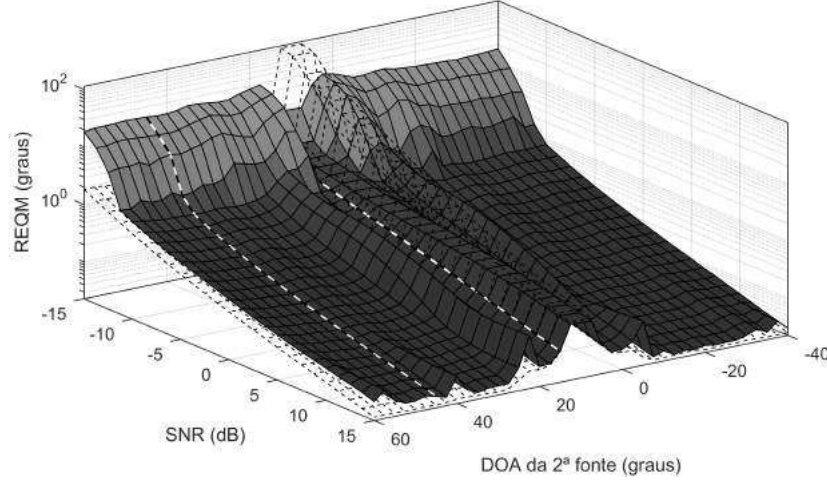


A Figura 43 mostra a superfície de desempenho do SEAD Aprimorado utilizando o filtro calculado a partir do subespaço de sinal. Pode-se observar nessa figura um ganho no desempenho de estimação para fontes próximas mesmo com a SNR de limiar se mantendo em  $-13,75$  dB para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$  em comparação ao algoritmo original do SEAD Aprimorado sem filtragem espacial. Para fontes afastadas, esse valor variou entre  $-11,25$  dB e  $-8,75$  dB e, considerando  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , a SNR de limiar foi de  $-11,25$  dB.

Finalmente, a superfície REQM do SEAD Aprimorado usando o filtro calculado a partir do espectro diferencial é apresentada na Figura 44. Pode-se verificar nessa figura que a SNR de limiar para fontes afastadas variou entre de  $-12,50$  dB e  $-10$  dB. Para fontes próximas, o desempenho de estimação foi sensivelmente melhorado em relação ao algoritmo original do SEAD Aprimorado para valores de SNRs acima de  $-10$  dB. Considerando

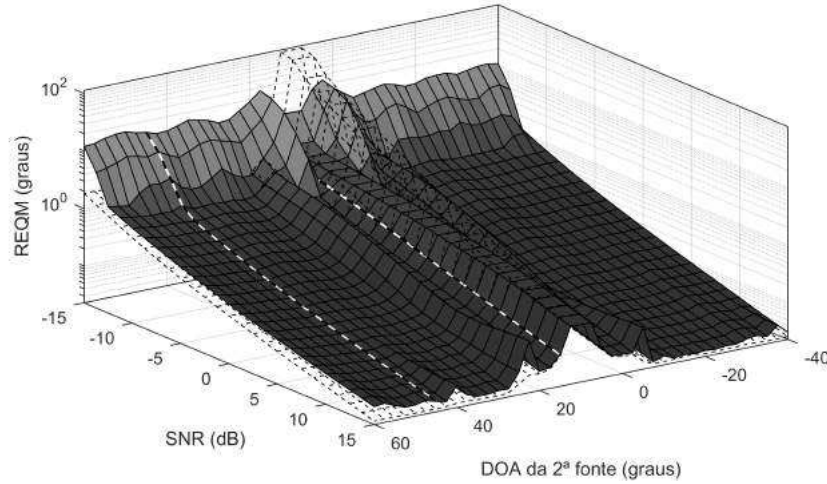


Figura 43 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).



$\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$  e  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , os valores de SNR de limiar foram de  $-13,75$  dB e  $-11,25$  dB, respectivamente.

Figura 44 – Superfície de desempenho do algoritmo original do SEAD Aprimorado utilizando o filtro calculado a partir do espectro diferencial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).



### 5.1.1.3 SEAD Modificado

Esta seção apresenta os resultados do SEAD Aprimorado após as alterações propostas, denominado SEAD Modificado. É importante ressaltar que foram avaliadas todas as cinco propostas para o estágio de refinamento de estimativas. Delas, a proposta 1, que implementa o Método de Newton aplicado ao DMLE com filtragem espacial utilizando a matriz de predição linear, apresentou um desempenho errático e abaixo de todas as outras propostas. As propostas 2 e 3, que implementam o Método de Newton aplicado ao DMLE com e sem filtragem espacial usando a matriz de direção, apresentaram bom desempenho. Entretanto,

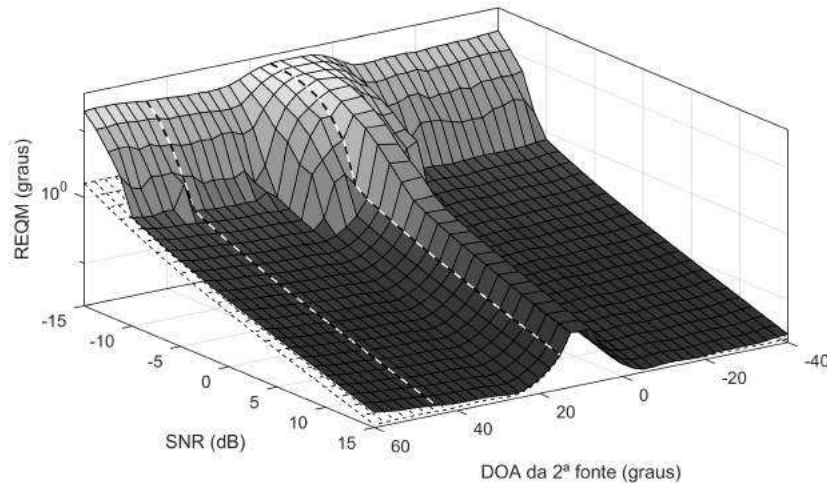
as propostas 4 e 5, que implementam o BFGS para o DMLE com e sem a filtragem espacial, apresentaram os melhores resultados. Dessa forma, para simplificar este capítulo, serão apresentados *apenas* os resultados referentes a essas propostas.

Além disso, as Figuras 45 e 49 apresentam o desempenho de estimação do SEAD Modificado *desconsiderando* a filtragem espacial. Isso foi realizado para evidenciar as diferenças no erro de estimação entre o algoritmo original do SEAD Aprimorado e o SEAD Modificado.

As Figuras 45 a 48 mostram o desempenho de estimação do SEAD Modificado considerando fontes de sinais descorrelacionados.

A superfície REQM do SEAD Modificado é ilustrada na Figura 45. Pode-se observar nessa figura que a SNR de limiar para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$  foi de  $-6,25$  dB. Esse valor é maior do que o apresentado pelo algoritmo original do SEAD Aprimorado. Isso se deve principalmente ao novo estágio de pré-seleção sem limiar de amplitude. No estágio de pré-seleção anterior, esse limiar impõe um limite rígido sobre as estimativas iniciais. Na situação de fontes próximas, esse limiar permite diminuir consideravelmente a SNR de limiar ao custo de um aumento no erro de estimação para fontes afastadas. Nessa situação, pode-se verificar que, em relação ao SEAD Aprimorado, a SNR de limiar para  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$  diminuiu de  $-7,50$  dB para  $-10$  dB. Ainda, é possível verificar que o erro de estimação diminuiu em altas SNRs com o uso de um algoritmo contínuo de otimização numérica.

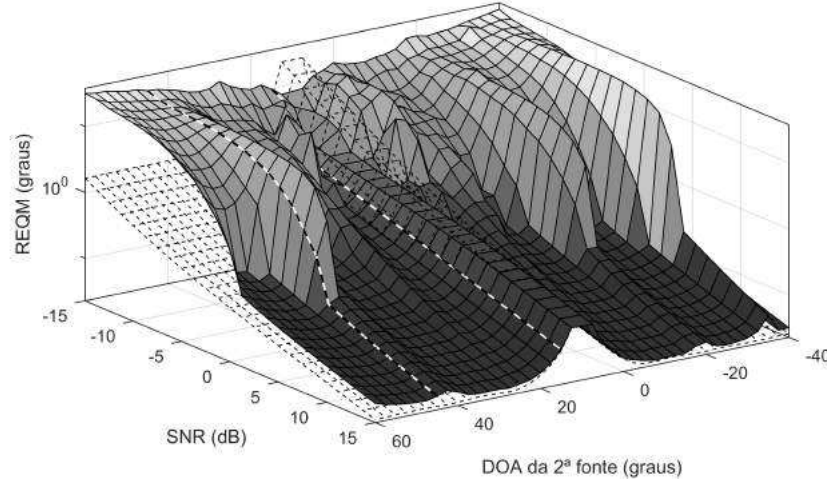
Figura 45 – Superfície de desempenho do SEAD Modificado para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).



A Figura 46 mostra a superfície de desempenho para o SEAD Modificado utilizando o máximo autofiltro espacial. Mais uma vez pode-se observar a grande degradação de desempenho para fontes afastadas. Considerando  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , a SNR de limiar foi de  $3,75$  dB. Esse valor é maior do que o apresentado pelo SEAD Modificado, devido ao novo estágio de pré-seleção. De maneira esperada, o erro de estimação considerando fontes

próximas é consideravelmente menor e, para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$ , o SEAD Aprimorado não apresentou uma SNR de limiar no intervalo das SNRs testadas.

Figura 46 – Superfície de desempenho do SEAD Modificado utilizando o máximo auto-filtro espacial para duas fontes descorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).



As Figuras 47 e 48 ilustram as superfícies REQM para o SEAD Modificado usando os filtros calculados a partir do subespaço de sinal e do espectro diferencial, respectivamente. Em relação ao SEAD Modificado sem a aplicação de filtragem espacial, pode-se observar que o uso desses filtros diminuiu sensivelmente o erro de estimação, embora a SNR de limiar tenha aumentado para  $-10$  dB e  $-8,75$  dB considerando  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$ . Adicionalmente, é possível verificar que o filtro calculado a partir do espectro diferencial apresentou melhores resultados do que filtro calculado a partir do subespaço de sinal, pois esse último introduz um erro de estimação para fontes próximas. Isso se deve ao número de iterações utilizado no estágio de refinamento de estimativas, uma vez que o BFGS apresenta maior sensibilidade ao ruído se comparado ao algoritmo anterior baseado no *Branch-and-Bound*. Por outro lado, pode-se verificar nessas figuras que as SNRs de limiar considerando  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$  foram respectivamente de  $-10$  dB e  $-8,75$  dB para os filtros calculados a partir do subespaço de sinal e a partir do espectro diferencial.

As Figuras 49 a 52 ilustram o desempenho de estimação do SEAD Modificado considerando fontes de sinais correlacionados.

A superfície de desempenho do SEAD Modificado é ilustrada na Figura 49. Pode-se observar nessa figura que, se comparado ao SEAD Aprimorado sem a aplicação de filtragem espacial, houve um aumento nas SNRs de limiar considerando fontes próximas e, para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$ , esse valor foi de  $-6,25$  dB. Entretanto, para fontes afastadas, pode-se verificar um ganho de até  $2,5$  dB na SNR de limiar e, para  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , o limiar de desempenho ocorreu em  $-10$  dB.

A Figura 50 apresenta a superfície REQM do SEAD Modificado utilizando o máximo

Figura 47 – Superfície de desempenho do SEAD Modificado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes decorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).

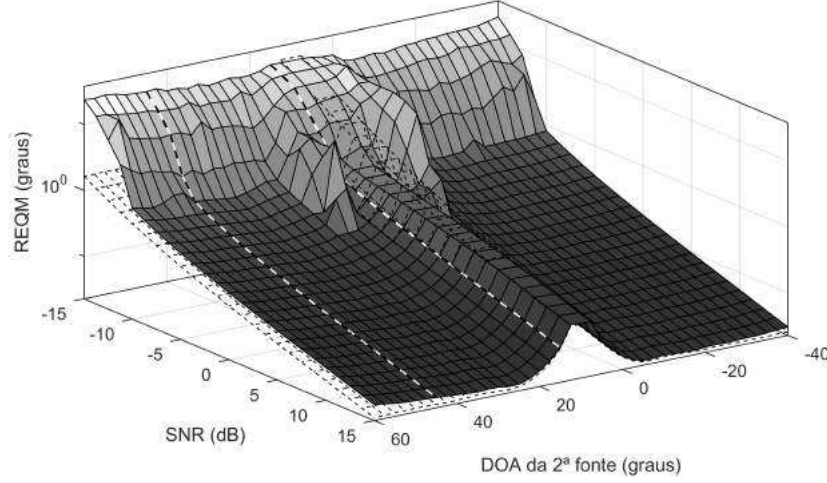
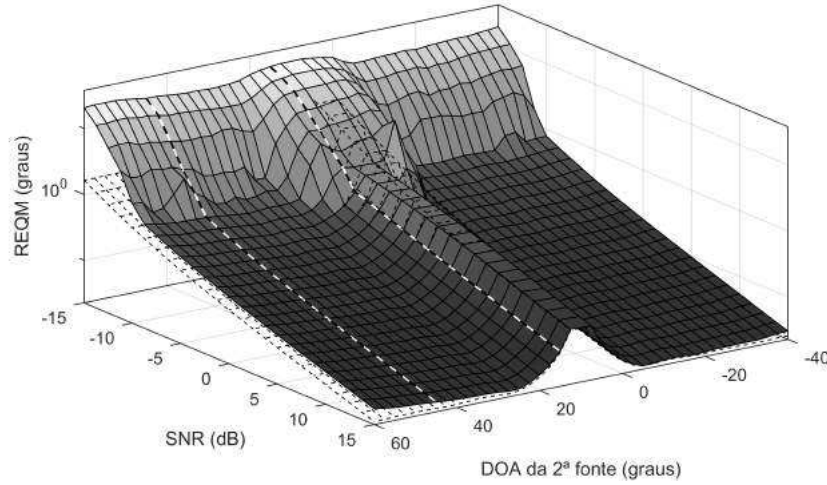


Figura 48 – Superfície de desempenho do SEAD Modificado utilizando o filtro calculado a partir do espectro diferencial para duas fontes decorrelacionadas (tons de cinza) e CRB correspondente (grade transparente).



autofiltro espacial. Como em todos os resultados com esse filtro, ocorreu um ganho para fontes próximas e uma degradação na situação de fontes afastadas quando comparado ao SEAD Modificado sem a aplicação de filtragem espacial. Para  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , a SNR de limiar foi  $-3,75$  dB. Para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$ , o SEAD Modificado não apresentou SNR de limiar no intervalo das SNRs testadas.

Finalmente, as Figuras 51 e 52 apresentam a superfície de desempenho do SEAD Modificado utilizando respectivamente os filtros calculados a partir do subespaço de sinal e do espectro diferencial. Pode-se verificar nessas figuras que o SEAD Modificado em conjunto com o filtro calculado a partir do subespaço de sinal apresenta desempenho melhor na situação de fontes próximas e a SNR de limiar para  $\{\theta_1, \theta_2\} = [10^\circ \ 15^\circ]$  foi de  $-12,50$  dB para ambos os filtros. Por outro lado, na situação de fontes afastadas, o



Figura 49 – Superfície de desempenho do SEAD Modificado para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).

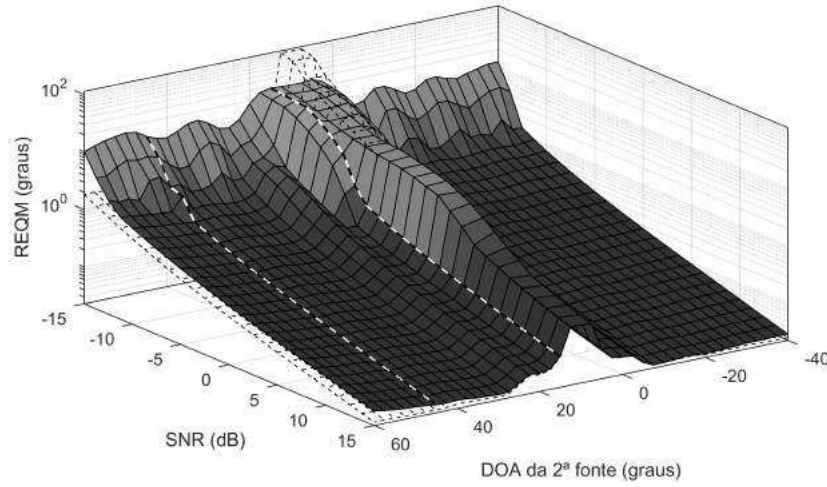
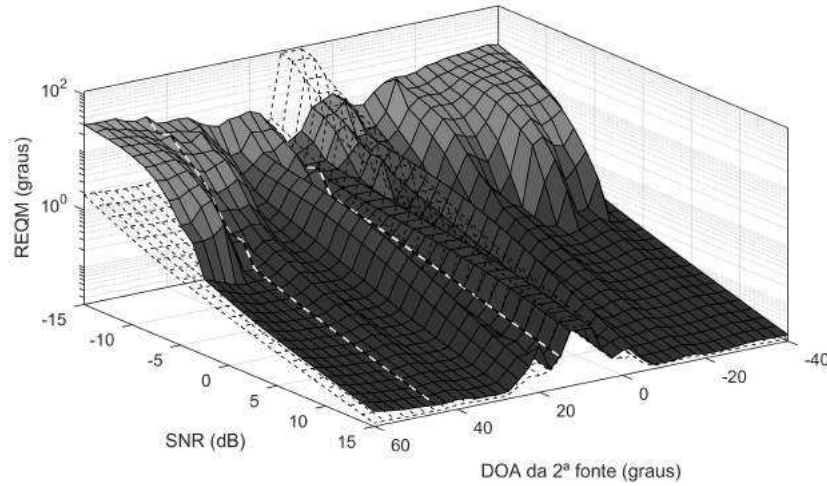


Figura 50 – Superfície de desempenho do SEAD Modificado utilizando o máximo autofiltro espacial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).



SEAD Modificado apresentou menores valores de SNR de limiar ao usar o filtro calculado a partir do espectro diferencial e, considerando  $\{\theta_1, \theta_2\} = [10^\circ \ 45^\circ]$ , a SNR de limiar foi de  $-10$  dB para ambos os filtros propostos.

#### 5.1.1.4 Discussão sobre os resultados

Primeiramente serão discutidos os resultados dos filtros propostos neste trabalho.

É possível observar nas seções 5.1.1.1 a 5.1.1.3 que o MODEX Modificado, o SEAD Aprimorado e o SEAD Modificado apresentam um erro de estimação consideravelmente menor para fontes afastadas do que para fontes próximas. Em todos os casos, a variância do erro de estimação aumenta de maneira razoavelmente constante até a SNR de limiar. Quando esse valor é atingido, o erro aumenta substancialmente e o estimador DOA não é mais confiável.

Figura 51 – Superfície de desempenho do SEAD Modificado utilizando o filtro calculado a partir do subespaço de sinal para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).

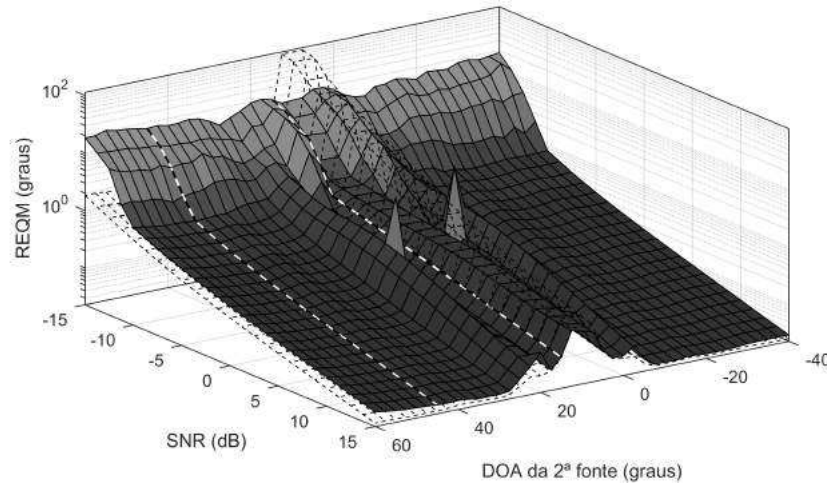
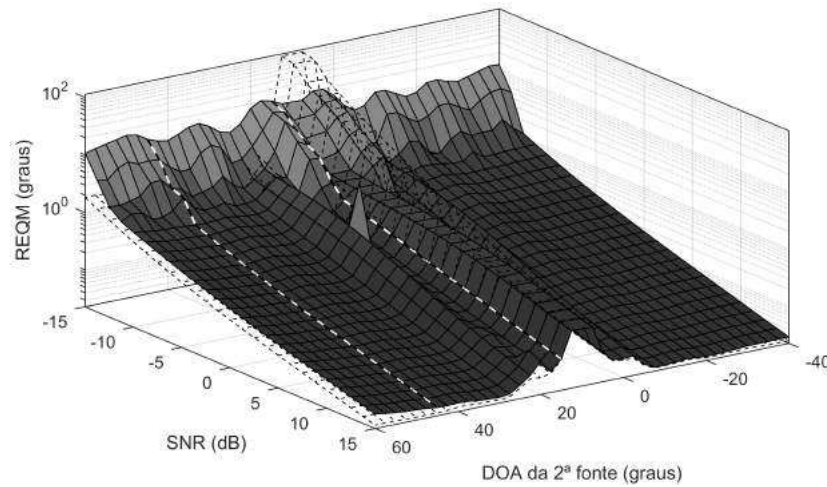


Figura 52 – Superfície de desempenho do SEAD Modificado utilizando o filtro calculado a partir do espectro diferencial para duas fontes correlacionadas (tons de cinza) e CRB correspondente (grade transparente).



Porém, ao aplicar os filtros espaciais apresentados neste trabalho, pode-se observar uma forte diminuição da variância do erro de estimação dos estimadores avaliados considerando fontes próximas. Para fontes afastadas em até  $10^\circ$ , o máximo autofiltro espacial apresentou os melhores resultados em termos do desempenho de estimação.

Por outro lado, quando consideradas fontes afastadas, esse filtro espacial *piorou* o desempenho de estimação do MODEX Modificado, do SEAD Aprimorado e do SEAD Modificado mesmo em SNRs razoavelmente altas. Entretanto, para SNRs suficientemente altas, o máximo autofiltro espacial não impediu o MODEX Modificado de produzir boas estimativas. Porém, esse autofiltro perturbou o estágio de refinamento do SEAD Aprimorado e do SEAD Modificado, gerando um aumento no erro de estimação mesmo na SNR de 10 dB.

Considerando os dois filtros espaciais propostos neste trabalho, pode-se verificar um desempenho muito semelhante entre eles, mas com algumas diferenças pontuais ao variar o estimador e/ou o fator de correlação. De maneira geral, os filtros propostos impuseram um aumento na SNR de limiar de *até* 2,50 dB na situação de fontes afastadas quando comparados aos estimadores sem a aplicação da filtragem espacial.

No entanto, para fontes próximas, o filtro calculado a partir do espectro diferencial apresentou os melhores resultados. Isso acontece, pois a aplicação do filtro calculado a partir do subespaço de sinal ao DMLE produz um espaço de soluções mais difícil de ser otimizado em comparação ao outro autofiltro. Isso pode ser verificado nas Figuras 53 e 54, que apresentam a superfície do DMLE para o mesmo experimento na ausência de ruído, considerando fontes situadas em  $10^\circ$  e  $15^\circ$  e sinais descorrelacionados.

Figura 53 – Superfície de um experimento do DMLE usando o filtro espacial calculado a partir do subespaço de sinal considerando fontes de sinal situadas em  $10^\circ$  e  $15^\circ$  e sinais descorrelacionados na ausência de ruído.

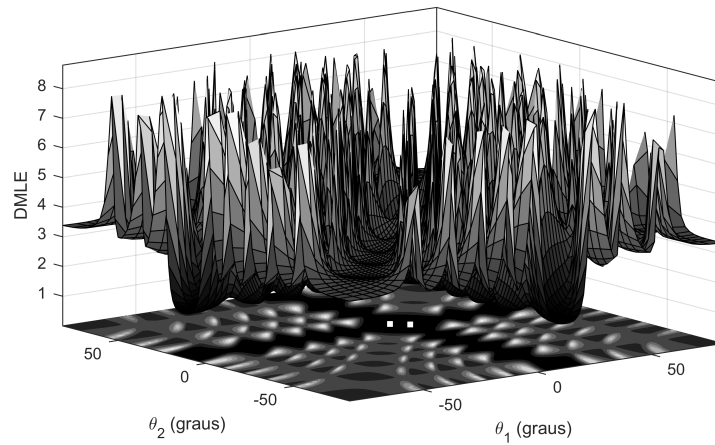
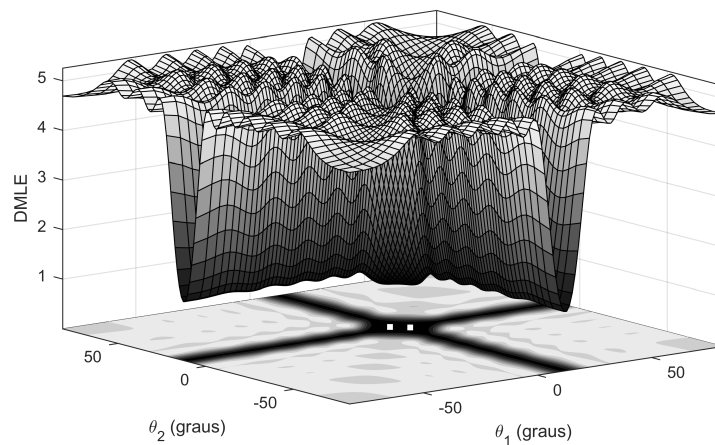


Figura 54 – Superfície de um experimento do DMLE usando o filtro espacial calculado a partir do espectro diferencial considerando fontes de sinal situadas em  $10^\circ$  e  $15^\circ$  e sinais descorrelacionados na ausência de ruído.



Nas duas figuras, o mínimo global está situado na região das fontes de sinal. Entretanto, na Figura 53 pode-se observar que os mínimos locais são tão baixos quanto o mínimo global e que há achatamento do mínimo global. Por outro lado, o mínimo global do espaço de soluções produzido pelo filtro calculado a partir do espectro diferencial ilustrado na Figura 54 é mais evidente e curvo. Consequentemente, é mais fácil a busca pelo mínimo global nessa superfície.

Além disso, é importante ressaltar que na situação de fontes próximas, as duas propostas de filtros espaciais produziram resultados melhores no desempenho de estimação ao comparar o MODEX Modificado, o SEAD Aprimorado e o SEAD Modificado sem a aplicação de filtragem espacial.

Em relação ao SEAD Modificado, pode-se verificar um erro de estimação na situação de fontes próximas e sem a aplicação de filtragem espacial se comparado ao SEAD Aprimorado. Esse aumento no erro se deve aos novos estágios de pré-seleção e refinamento de estimativas.

O procedimento realizado no estágio de pré-seleção não utiliza o limiar de amplitude. Esse limiar separa os picos de sinal dos picos de ruído e sua aplicação produz um erro de estimação consideravelmente menor na situação de ângulos próximos. Nessa situação, os picos de sinal tendem a ser significativamente maiores do que os picos de ruído. Entretanto, o limiar pode afetar de maneira muito negativa no desempenho de estimação de fontes afastadas, pois os picos referentes a essas fontes podem apresentar amplitudes semelhantes aos dos picos de ruído. Dessa forma, essas fontes podem ser completamente desconsideradas na geração de estimativas iniciais e, por isso, gera um erro de estimação significativo.

Ao não utilizar o limiar de amplitude, a probabilidade de gerar estimativas iniciais desconsiderando alguma fonte afastada diminuiu de maneira significativa. Entretanto, é necessário gerar e avaliar um número maior de estimativas para produzir o vetor de estimativas iniciais. Essa avaliação é realizada utilizando o Estimador de Máxima-Verossimilhança e isso aumenta a sensibilidade ao ruído. Por outro lado, com esse novo estágio de pré-seleção, o SEAD Modificado possui desempenho de estimação robusto em qualquer situação de fontes afastadas e consistentemente apresentou SNRs de limiar próximas de  $-10$  dB.

De maneira muito semelhante, o antigo estágio de refinamento de estimativas do SEAD Aprimorado impõe restrições rígidas que impedem que as estimativas divirjam muito das estimativas iniciais. Assim, considerando um vetor de estimativas iniciais próximas do mínimo global do DMLE, é provável que o SEAD Aprimorado produza boas estimativas após o refinamento, mesmo que o algoritmo realize uma otimização discreta. Entretanto, ao usar um número menor de iterações no estágio de refinamento, o SEAD Aprimorado apresenta um desempenho de estimação inferior em altas SNRs.

Por sua vez, o novo estágio de refinamento proposto realiza uma otimização sem



restrições no espaço de soluções do DMLE. De maneira geral, ele apresenta maior sensibilidade ao ruído, aumentando o erro de estimação em baixas SNRs e fontes próximas. Entretanto, uma vez que a otimização é realizada em um intervalo contínuo, o erro de estimação se aproximou do CRB em altas SNRs.

Finalmente, pode-se observar que, tanto em baixas quanto em altas SNRs, o SEAD Modificado apresenta desempenho de estimação semelhante ao do MODEX Modificado.

### 5.1.2 Tempos de execução

Os três estimadores abordados neste trabalho também foram avaliados em termos do tempo de execução com e sem os filtros espaciais. Foram amostrados os tempos de execução em um computador com o processador Intel Xeon E3-1220 e 32 GB de memória RAM (do inglês *Read Access Memory*) do tipo DDR3-1333E. As Tabelas 4 a 6 apresentam os tempos totais de  $E = 1000$  execuções do MODEX Modificado, do SEAD Aprimorado e do SEAD Modificado sem filtragem espacial e com filtragem espacial considerando os três filtros pesquisados neste trabalho.

Tabela 4 – Tempos totais para 1.000 simulações do MODEX Modificado em função do número de fontes.

N.º de fontes	Sem filtragem espacial	Máximo autofiltro	Proposta 1	Proposta 2
2	4,35 s	4,31 s	5,37 s	6,59 s
3	13,39 s	14,56 s	15,19 s	16,10 s
4	73,67 s	79,51 s	80,70 s	84,68 s
5	488,55 s	531,74 s	534,35 s	556,14 s
6	3.227,72 s	3.615,10 s	3.623,78 s	3.757,88 s
7	22.084,15 s	24.836,08 s	24.897,52 s	26.035,17 s
8	149.708,79 s	167.027,47 s	167.707,35 s	174.856,92 s

Tabela 5 – Tempos totais para 1.000 simulações do SEAD Aprimorado em função do número de fontes.

N.º de fontes	Sem filtragem espacial	Máximo autofiltro	Proposta 1	Proposta 2
2	14,95 s	15,96 s	17,40 s	17,93 s
3	26,33 s	28,50 s	29,24 s	30,36 s
4	67,96 s	72,91 s	73,98 s	77,19 s
5	208,03 s	225,49 s	227,19 s	236,37 s
6	643,19 s	717,43 s	721,71 s	747,27 s
7	2.103,73 s	2.353,62 s	2.366,52 s	2.465,85 s
8	6.692,41 s	7.460,42 s	7.510,18 s	7.824,93 s

Em relação aos estimadores, pode-se verificar que o tempo de execução cresce à medida que o número de fontes aumenta. Entretanto, é possível observar que, entre os três estimadores, o MODEX Modificado apresenta o maior tempo total de execução. Por sua vez,

Tabela 6 – Tempos totais para 1.000 simulações do SEAD Modificado em função do número de fontes.

N.º de fontes	Sem filtragem espacial	Máximo autofiltro	Proposta 1	Proposta 2
2	12,37 s	13,93 s	13,81 s	15,41 s
3	12,46 s	13,20 s	14,04 s	5,09 s
4	16,72 s	17,63 s	18,72 s	19,80 s
5	34,32 s	37,73 s	39,51 s	39,55 s
6	119,25 s	130,57 s	134,59 s	132,09 s
7	755,65 s	797,00 s	814,43 s	799,28 s
8	1.675,23 s	1.750,40 s	1.782,98 s	1.753,48 s

o SEAD Aprimorado também apresenta um esforço computacional significativo. Entretanto, o tempo total de execução dele é cerca de 22 vezes menor se comparado ao MODEX Modificado estimando oito fontes. O SEAD Modificado diminui ainda mais o tempo total de execução, sendo entre 89 e 100 vezes mais rápido do que o MODEX Modificado e cerca de quatro vezes mais rápido do que o SEAD Aprimorado, também considerando  $M = 8$ . No entanto, o esforço computacional apresentado pelo SEAD Modificado ainda é de ordem combinatorial.

Adicionalmente, os tempos de execução dos estimadores foram avaliados em relação ao percentil 97 como apresentado na Tabela 7 em função do número de fontes e do filtro espacial utilizado: máximo autofiltro (H1), filtro calculado a partir do subespaço de sinal (H2) e o filtro calculado a partir do espectro diferencial (H3).

Tabela 7 – Percentil 97 dos tempos de execução (em segundos) dos estimadores avaliados em função do número de fontes e em função do filtro espacial utilizado.

N.º de fontes	MODEX Mod	MODEX Mod H1	MODEX Mod H2	MODEX Mod H3	SEAD Aprim	SEAD Aprim H1	SEAD Aprim H2	SEAD Aprim H3	SEAD Mod	SEAD Mod H1	SEAD Mod H2	SEAD Mod H3
2	0,0160	0,0047	0,0117	0,0283	0,0360	0,0376	0,0463	0,0465	0,0336	0,0378	0,0338	0,0438
3	0,0161	0,0152	0,0167	0,0219	0,0302	0,0386	0,0383	0,0380	0,0141	0,0141	0,0152	0,0188
4	0,0757	0,0816	0,0862	0,0902	0,0710	0,0756	0,0793	0,0861	0,0175	0,0184	0,0197	0,0208
5	0,5047	0,5575	0,5614	0,5797	0,2205	0,2324	0,2418	0,2503	0,0357	0,0391	0,0408	0,0439
6	3,2867	3,6701	3,6770	3,8092	0,6682	0,7371	0,7421	0,7702	0,1213	0,1332	0,1372	0,1347
7	22,3552	25,1142	25,1906	26,4026	2,1439	2,3974	2,4092	2,5095	0,7811	0,8222	0,8416	0,8233
8	151,0126	168,7119	169,2398	177,1867	6,7761	7,5569	7,6098	7,9333	1,5624	1,6380	1,6689	1,6354

Pode-se verificar na Tabela 7 que em 97% das amostras o tempo de cada execução do MODEX Modificado se manteve abaixo de um segundo para  $2 \leq M \leq 5$ . Além disso, considerando  $M > 5$ , o tempo de execução subiu de maneira bastante significativa. O SEAD Aprimorado apresentou tempos de execução abaixo de um segundo considerando  $2 \leq M \leq 6$  e, para  $M = 7$  e  $M = 8$ , houve um aumento nos tempos de execução, mas não da ordem do MODEX Modificado. Adicionalmente, pode-se observar que o SEAD Modificado apresentou tempos de execução inferiores a um segundo considerando o número de fontes menor ou igual a sete e, para  $M = 8$ , o tempo de execução é significativamente menor do que dos outros estimadores.

Finalmente, os filtros espaciais são avaliados isoladamente tendo como referência as Tabelas 4 a 6, onde pode-se verificar que a aplicação da filtragem espacial resulta em um aumento nos tempos de execução. As Tabelas 8 e 9 mostram esses aumentos nos tempos de estimação relativos aos MODEX Modificado, SEAD Aprimorado e SEAD Modificado sem a aplicação de filtragem espacial.

Tabela 8 – Aumento nos tempos de execução relativos ao MODEX Modificado sem filtragem espacial em função do número de fontes.

N.º de fontes	Máximo autofiltro	Proposta 1	Proposta 2
2	-1,01%	23,38%	51,43%
3	8,76%	13,49%	20,28%
4	7,92%	9,54%	14,94%
5	8,84%	9,37%	13,83%
6	12,00%	12,27%	16,43%
7	12,46%	12,74%	17,89%
8	11,57%	12,02%	16,80%
<b>Média</b>	<b>8,65%</b>	<b>13,26%</b>	<b>21,66%</b>

Tabela 9 – Aumento nos tempos de execução relativos ao SEAD Aprimorado sem filtragem espacial em função do número de fontes.

N.º de fontes	Máximo autofiltro	Proposta 1	Proposta 2
2	6,72%	16,34%	19,90%
3	8,25%	11,05%	15,32%
4	7,28%	8,86%	13,58%
5	8,39%	9,21%	13,62%
6	11,54%	12,21%	16,18%
7	11,88%	12,49%	17,21%
8	11,48%	12,22%	16,92%
<b>Média</b>	<b>9,36%</b>	<b>11,77%</b>	<b>16,11%</b>

Tabela 10 – Aumento nos tempos de execução relativos ao SEAD Modificado sem filtragem espacial em função do número de fontes.

N.º de fontes	Máximo autofiltro	Proposta 1	Proposta 2
2	12,65%	11,67%	24,55%
3	5,95%	12,66%	21,16%
4	5,44%	11,95%	18,43%
5	9,91%	15,10%	15,24%
6	9,49%	12,86%	10,77%
7	5,47%	7,78%	5,77%
8	4,49%	6,43%	4,67%
<b>Média</b>	<b>7,63%</b>	<b>11,21%</b>	<b>14,37%</b>

Pode-se verificar nas Tabelas 8 a 10 que o máximo autofiltro espacial foi responsável por um aumento médio no tempo de execução que variou entre 8% e 9%. O filtro calculado

a partir do subespaço de sinal (Proposta 1) apresentou um aumento médio entre 11% e 13% e o filtro calculado a partir do espectro diferencial (Proposta 2) aumentou o tempo de estimação entre 14% e 22%.

Finalmente, pode-se observar que o SEAD Modificado apresentou os menores aumentos médios nos tempos de execução com a aplicação da filtragem espacial se comparado ao MODEX Modificado e ao SEAD Aprimorado.

#### 5.1.2.1 Discussão sobre os resultados

Na seção 5.1.2, pode-se verificar que a filtragem espacial incorre em um aumento no esforço computacional à estimação DOA. Nos estimadores pesquisados neste trabalho, observou-se que o aumento médio nos tempos de execução é de até 22%. Isso é devido à estimação de cada um dos filtros espaciais antes de iniciar a estimação DOA e, principalmente, às dimensões maiores das matrizes utilizadas na função-objetivo do DMLE filtrado.

O máximo autofiltro espacial é o mais simples de ser calculado e, conseqüentemente, ele apresentou os menores aumentos médios no tempo de execução. Por sua vez, o filtro calculado a partir do subespaço de sinal é uma extensão do máximo autofiltro espacial para todo o subespaço de sinal. Naturalmente, são necessárias mais operações para estimar o filtro. Por último, o filtro calculado a partir do espectro diferencial exige ainda mais operações ao calcular o espectro diferencial e depois os coeficientes do filtro. Contudo, ele é capaz de produzir um filtro com uma ordem menor do que o filtro da Proposta 1.

Em relação aos estimadores DOA, o MODEX Modificado apresentou o maior esforço computacional dentre os três estimadores abordados neste trabalho. À medida que o número de fontes cresce, o esforço computacional do MODEX Modificado aumenta bastante. Assim, ele não é um estimador viável para aplicações em tempo real quando o número de fontes é grande. Nessas aplicações, ele está limitado à estimação DOA de um pequeno número de fontes de sinal, situação em que ele apresenta tempos de execução reduzidos. O SEAD Aprimorado possui ordem exponencial, como apresentado na seção 3.6, e isso é devido ao estágio de refinamento de estimativas. Entretanto, o esforço computacional é menor do que o apresentado pelo MODEX Modificado. Ainda assim, ele possui aplicabilidade limitada ao número de fontes.

Ao propor neste trabalho substituir o estágio de refinamento de estimativas do SEAD Aprimorado pelo BFGS, o SEAD Modificado apresentou uma considerável diminuição nos tempos de execução e no esforço computacional. Isso o torna mais apto às aplicações em tempo real que exigem a estimação DOA das fontes de sinal. Entretanto, ao propor alterar o estágio de pré-seleção, removendo o limiar de amplitude, passa a ser necessário testar um número maior de estimativas para gerar o vetor de estimativas iniciais que serão refinadas pelo BFGS.

O número de combinações que esse novo estágio de pré-seleção realiza está descrito na Equação (4.45) e apresentado na Tabela 3. Assim, pode-se verificar que a função que descreve o número de estimativas testadas possui ordem combinatorial. Entretanto, é importante ressaltar que, ainda assim, o SEAD Modificado possui o menor esforço computacional.

## 5.2 Conclusões

Este capítulo apresentou os resultados obtidos a partir das alterações propostas neste trabalho ao estimador SEAD Aprimorado. Em primeiro lugar, foram definidos os parâmetros das simulações numéricas realizadas. Posteriormente, foram apresentados os resultados em termos do desempenho de estimação DOA e dos tempos de execução. Finalmente, esses resultados foram analisados e discutidos.

Como existem duas propostas de filtros espaciais e propostas de alteração ao SEAD Aprimorado. A análise e a discussão foram feitas em duas partes. A primeira delas refere-se aos filtros propostos.

O desempenho de estimação foi avaliado a partir da variância do erro de estimação em função do afastamento das fontes e da SNR, gerando superfícies de desempenho. Essas superfícies ressaltaram uma melhora significativa proporcionada pelo máximo autofiltro espacial no desempenho de estimação para duas fontes próximas em SNRs baixas. Por outro lado, à medida que a SNR diminui, pode-se observar uma nítida degradação do desempenho de estimação para duas fontes afastadas.

Adicionalmente, foi possível verificar que os dois filtros espaciais propostos neste trabalho tiveram também um melhoramento significativo no desempenho de estimação para fontes próximas em relação ao MODEX Modificado, SEAD Aprimorado e SEAD Modificado sem a aplicação de filtragem espacial. Entretanto, também pode-se verificar que os filtros espaciais propostos não afetaram negativamente de maneira significativa o desempenho de estimação para fontes afastadas. Em relação ao máximo autofiltro espacial, é importante ressaltar que nessa situação as duas propostas de filtros espaciais resultaram em uma melhora significativa no desempenho de estimação.

Em relação às alterações propostas ao SEAD Aprimorado, pode-se observar que ele foi menos afetado pela aplicação dos filtros espaciais, devido ao estágio de pré-seleção utilizar um limiar de restrição de amplitude. Além disso, esse estimador apresenta uma resolução máxima associada à quantidade de iterações utilizadas no estágio de refinamento de estimativas. Isso acontece porque esse estágio utiliza um algoritmo com elevado esforço computacional e é utilizado na otimização de problemas discretos. Embora tenha apresentado resultados robustos, ele não é apropriado para a estimação DOA. Adicionalmente, o estágio de pré-seleção pode produzir erros de estimação, pois ele pode não considerar

todos os picos de sinal para produzir uma estimativa inicial devido à aplicação do limiar de amplitude.

Nesse sentido, é possível observar que o SEAD Modificado, denominação do SEAD Aprimorado após as alterações propostas, superou esses problemas. O novo estágio de pré-seleção não utiliza o limiar de amplitude. Por isso, ele apresenta um esforço computacional maior. Principalmente, o estágio de refinamento possui esforço computacional significativamente menor e é apropriado na otimização de problemas contínuos, como é a estimação DOA. Com ou sem filtragem espacial, o SEAD Modificado possui um desempenho de estimação muito próximo do MODEX Modificado, tido como referência neste trabalho. Contudo, é importante ressaltar que a aplicação ao SEAD Modificado dos filtros espaciais propostos possibilita ganhos expressivos na estimação DOA de fontes próximas e um pequeno aumento na SNR de limiar na situação de fontes afastadas.

Além disso, pode-se verificar que a filtragem espacial incorre em aumentos nos tempos de execução. Isso é devido ao tempo gasto para estimar os filtros e às matrizes de dimensões maiores utilizadas na função-objetivo do DMLE filtrado. Em relação ao máximo autofiltro espacial, usado como referência neste trabalho, os filtros espaciais propostos apresentaram um pequeno aumento nos tempos de execução. Por outro lado, esse aumento é compensado pelo desempenho de estimação observado quando aplicados aos estimadores abordados neste trabalho. Não foi encontrado nenhum filtro espacial que, ao ser aplicado ao DMLE, resultasse em um desempenho de estimação robusto tanto para fontes próximas quanto afastadas. Assim, os filtros propostos permitem o uso da filtragem espacial na estimação DOA de máxima-verossimilhança em situações adversas.

Finalmente, pode-se concluir que o SEAD Modificado possui desempenho de estimação compatível com o do MODEX Modificado, mas, com um esforço computacional significativamente menor. Ainda, o SEAD Aprimorado em conjunto com os filtros espaciais propostos possui desempenho de estimação superior na situação de fontes próximas e fontes afastadas.

O próximo capítulo apresenta as conclusões, as contribuições deste trabalho e os trabalhos futuros que poderão ser realizados a partir desta tese.

## 6 CONCLUSÕES, CONTRIBUIÇÕES E TRABALHOS FUTUROS

Este trabalho propõe alterações ao estimador SEAD Aprimorado para melhorar o seu desempenho de estimação em diversas situações.

Para isso, foram apresentados a modelagem utilizada para os sinais e para o arranjo de sensores e o Estimador de Máxima-Verossimilhança. Contudo, esse estimador apresenta um grande esforço computacional. Dessa forma, o seu uso é inviável em aplicações em tempo real onde a estimação DOA pode ser utilizada. Assim, foram apresentados dois estimadores iterativos: o MODEX Modificado e o SEAD Aprimorado.

O MODEX Modificado apresenta desempenho de estimação muito próximo ao do MLE com um esforço computacional significativamente menor. Isso foi alcançado gerando três estimativas MODE utilizando restrições distintas. Logo após, é selecionada a melhor combinação usando um procedimento ML. Por sua vez, o SEAD Aprimorado utiliza um processo de geração de estimativas DOA dividido em dois estágios: pré-seleção e refinamento. O primeiro é responsável por gerar estimativas iniciais que são posteriormente refinadas pelo segundo estágio. O resultado é um estimador muito robusto na estimação DOA de fontes próximas, com esforço computacional menor do que o MODEX Modificado.

Tanto o MODEX Modificado estimador quanto o SEAD Aprimorado apresentam o efeito de limiar por dividirem o espaço de medida dos dados observados em subespaço de sinal e subespaço ortogonal. Assim, quando a SNR diminui suficientemente, o procedimento ML pode não selecionar a melhor combinação de estimativas *mesmo* que ela faça parte do grupo de combinações. Isso é devido ao ruído presente na matriz de covariância afetar de maneira significativa esse procedimento.

Contudo, uma maneira de diminuir o efeito de limiar é realizar a filtragem espacial dos dados recebidos pelo arranjo antes de proceder com a estimação DOA. Porém, a função-objetivo do Estimador de Máxima-Verossimilhança deve ser alterada, pois as matrizes possuem dimensões maiores e o modelo de sinal é modificado com a aplicação do filtro.

Assim, este trabalho propõe alterações ao SEAD Aprimorado, resultando no SEAD Modificado. A primeira alteração aplica a filtragem espacial para diminuir o efeito de limiar no desempenho de estimação. A nova função-objetivo foi apresentada em (KRUMMENAUER et al., 2010) juntamente com o máximo autofiltro espacial. O resultado é uma melhoria significativa do desempenho de estimação do MODEX Modificado na situação de fontes próximas. Então, o máximo autofiltro espacial foi utilizado como referência e é calculado a partir do autovetor associado ao máximo autovalor da matriz de covariância.



Porém, pode-se observar que em diversas situações esse autofiltro usa apenas parte da informação contida no subespaço de sinal. Isso é devido ao máximo autofiltro espacial utilizar apenas o máximo autovetor. Dessa forma, esse autofiltro funciona adequadamente somente quando existe uma fonte de sinal ou duas ou mais fontes próximas entre si. Por outro lado, na situação de fontes afastadas, o máximo autofiltro espacial atenua consideravelmente uma delas.

Essa atenuação causa uma alteração significativa no espaço de soluções e pode introduzir erros de estimação, especialmente em SNRs mais baixas. Assim, para evidenciar essa característica, o máximo autofiltro espacial foi avaliado no MODEX Modificado, no SEAD Aprimorado e no SEAD Modificado em função do afastamento entre duas fontes de sinal. Pode-se observar que o SEAD Aprimorado foi afetado *negativamente* pela aplicação do máximo autofiltro espacial na situação de fontes afastadas. Isso é devido ao estágio de pré-seleção desse estimador gerar uma quantidade muito reduzida de estimativas. Adicionalmente, o estágio de refinamento aplica rígidas restrições em uma eventual divergência na estimação. Essas duas características limitaram o impacto que o filtro gerou na seleção de estimativas. Por outro lado, o MODEX Modificado e o SEAD Modificado, ambos em conjunto com esse autofiltro, apresentaram melhor desempenho de estimação na situação de fontes próximas. No entanto, para fontes afastadas, eles apresentaram resultados piores do que sem a filtragem espacial.

Assim, foram propostos dois filtros multibanda adequados à estimação DOA de fontes próximas e afastadas. Um filtro é a extensão do máximo autofiltro espacial para todo o subespaço de sinal. O outro é calculado a partir do espectro diferencial utilizado no SEAD Aprimorado.

Na situação de fontes afastadas e em relação ao máximo autofiltro espacial, os dois filtros espaciais propostos apresentaram uma melhoria considerável no desempenho de estimação quando aplicados ao MODEX Modificado e ao SEAD Modificado. Adicionalmente, foi verificado que ocorreu apenas uma pequena perda de desempenho quando comparado a esses estimadores sem a aplicação de filtragem espacial. Na situação de fontes próximas, os dois filtros propostos apresentaram resultados próximos aos do autofiltro. Ainda, foi observado que, ao aplicar os filtros propostos, a estimação DOA é um pouco mais sensível ao ruído. Assim, confirmou-se a hipótese de que o erro de estimação do SEAD Modificado seria reduzido com a aplicação de filtragem espacial. Contudo, essa redução somente ocorreu com a proposição de novos filtros espaciais.

Adicionalmente, foram identificados dois problemas no SEAD Aprimorado. O primeiro é devido ao estágio de pré-seleção utilizar um limiar de amplitude que pode falhar em algumas situações. O segundo problema é devido ao estágio de refinamento ser baseado no algoritmo discreto denominado *Branch-and-Bound*. Dessa forma, a variância do erro de estimação não diminui continuamente à medida que a SNR aumenta, tornando o

SEAD Aprimorado em um estimador assintoticamente ineficiente. Além disso, o estágio de refinamento possui esforço computacional de ordem exponencial.

Assim, a segunda alteração proposta ao SEAD Aprimorado é substituir o estágio de refinamento de estimativas. Como são produzidas estimativas iniciais antes desse estágio, uma alternativa viável é utilizar um método numérico de otimização local como o Método de Newton ou algum método quase-Newton. O Método de Newton para o DMLE já foi proposto em (STARER; NEHORAI, 1992) usando a matriz de predição linear, porém, ele não aplica a filtragem espacial. Além disso, a filtragem espacial da forma proposta utiliza a matriz de direção. Então, a primeira atividade foi reproduzir no Apêndice A o equacionamento do Método de Newton já proposto, em que foram obtidos os mesmos resultados.

A partir disso, foi desenvolvida uma reparametrização da função-objetivo do DMLE usando a matriz de predição linear para adicionar a filtragem espacial. Logo após, o Método de Newton foi derivado para essa nova função-objetivo (Proposta 1, Apêndice B). No entanto, ele apresentou desempenho errático e nem sempre converge para o mínimo global da função mesmo em altas SNRs.

Então, como não foi encontrado na literatura da área, foi desenvolvido o Método de Newton para a função-objetivo original do DMLE, que utiliza a matriz de direção e não aplica a filtragem espacial (Proposta 2, Apêndice C). Ele apresentou desempenho compatível com o proposto em (STARER; NEHORAI, 1992), mas exige uma quantidade menor de derivadas parciais, pois a variável independente é real em vez de ser complexa. Adicionalmente, foi possível aplicar o método quase-Newton denominado BFGS para reduzir ainda mais o esforço computacional despendido (Proposta 4). Isso é possível, pois o BFGS não exige a hessiana exata e é capaz de calcular diretamente a inversa da hessiana aproximada. Por último, foram derivados o Método de Newton (Proposta 3, Apêndice D) e o BFGS (Proposta 5) para a função-objetivo do DMLE com aplicação de filtragem espacial em função da matriz de direção.

Normalmente, o BFGS possui uma taxa de convergência local menor do que o Método de Newton devido às aproximações que ele considera. Entretanto, os resultados obtidos no decorrer deste trabalho indicaram que o BFGS apresenta um desempenho sensivelmente superior ao do Método de Newton. Mas, para isso, é necessário fazer um ajuste do comprimento de passo. Nesse sentido, o BFGS e o Método de Newton utilizaram as condições de Wolfe. Por esse motivo, somente foram apresentados os resultados da substituição do estágio de refinamento do SEAD Aprimorado pelo BFGS. Porém, convém ressaltar que as Propostas 4 e 5 utilizaram parte do equacionamento resultante do desenvolvimento das Propostas 2 e 3.

Então, ao aplicar as Propostas 4 e 5 ao SEAD Modificado, ele passou a ser um estimador assintoticamente eficiente, com um esforço computacional muito inferior, confir-

mando a primeira hipótese afirmada neste trabalho. Contudo, ele ainda utilizava o limiar de amplitude na geração de estimativas iniciais. Assim, foram propostas duas alterações ao estágio de pré-seleção para remover esse limiar e para auxiliar na geração de melhores estimativas iniciais. Com isso, o novo estágio de pré-seleção apresentou maior esforço computacional, uma vez que ele requer um número maior de avaliações da função-objetivo para gerar o vetor de estimativas iniciais. Além disso, ele é um pouco mais sensível ao ruído na situação de fontes próximas e baixas SNRs. No entanto, é importante ressaltar que o desempenho de estimação para fontes afastadas é robusto e é comparável ao do MODEX Modificado. Assim, confirma-se a hipótese de menor erro de estimação na situação de fontes afastadas ao não utilizar o limiar de amplitude no estágio de pré-seleção. Adicionalmente, mesmo com o novo estágio de pré-seleção apresentando maior esforço computacional, o SEAD Modificado apresentou tempos de execução significativamente menores do que o SEAD Aprimorado e o MODEX Modificado.

Em resumo, as três alterações propostas foram: 1) aplicação de filtragem espacial, 2) substituição do estágio de refinamento anterior pelo BFGS e 3) a retirada do limiar de amplitude no estágio de pré-seleção tornaram o SEAD Modificado um estimador muito competitivo em termos do desempenho de estimação DOA de fontes próximas e afastadas e em termos do esforço computacional. Pode-se afirmar que essas alterações atingiram o objetivo geral proposto neste trabalho.

Finalmente, é importante ressaltar que, para as quatro funções-objetivo avaliadas, este trabalho documentou *todo* o desenvolvimento analítico do Método de Newton.

## 6.1 Contribuições deste Trabalho

Pode-se listar como contribuições deste trabalho, em ordem decrescente de relevância:

- O desenvolvimento analítico do Método de Newton para a função-objetivo do DMLE com filtragem espacial em função da matriz de direção filtrada  $\tilde{\mathbf{A}}$  apresentada na Equação (3.35). Isso, além de realizar a otimização local dessa equação, permitiu aplicar o método quase-Newton BFGS, que possui menor esforço computacional;
- O desenvolvimento analítico do Método de Newton para a função objetivo do DMLE sem filtragem espacial em função da matriz de direção  $\mathbf{A}$  apresentada na Equação (2.30). Da mesma forma que na contribuição acima, esse desenvolvimento permitiu aplicar o BFGS;
- A proposição de dois novos filtros espaciais multibanda adequados à estimação DOA de fontes próximas e afastadas, com esforço computacional competitivo em relação ao máximo autofiltro espacial;

- A remoção do limiar de amplitude do estágio de pré-seleção do SEAD Modificado, tornando mais confiável a estimação DOA independentemente do espaçamento entre as fontes de sinal.

## 6.2 Divulgação Científica Efetivada

Este trabalho produziu a seguinte divulgação científica:

1. LEMOS, R.P.; SILVA, H.V.L.; FLORES, E.L.; KUNZLER, J.A.; BELTRÁN, D.F.B. Spatial Filtering Based On Differential Spectrum For Improving ML DOA Estimation Performance. *IEEE Signal Processing Letters*. v. 23, n. 12, pp. 1811-1815. Agosto 2016;
2. BELTRÁN, D.F.B.; LEMOS, R.P.; SILVA, H.V.L.; KUNZLER, J.A.; FLORES, E.L. Adaptive Beamforming for Moving Targets Using Genetic Algorithms. *Ingeniería*, v. 21, n. 2, pp. 214, Agosto 2016;
3. BELTRÁN, D.F.B.; KUNZLER, J.A.; LEMOS, R.P.; SILVA, H.V.L. Adaptive Beamforming for Moving Targets Using Genetic Algorithms. In: *2015 Workshop on Engineering Applications International Congress on Engineering (WEA)*, Outubro 2015, Bogotá, Colômbia, pp. 1-5;
4. BELTRÁN, D.F.B.; LEMOS, R.P.; KUNZLER, J.A.; SILVA, H.V.L. Adaptive Beamforming for Moving Targets Using Genetic Algorithms and a CDMA reference signal. In: *2015 IEEE Colombian Conference on Communications and Computing (COLCOM)*, Maio 2015, Popayan, Colômbia, pp. 1-5;
5. KUNZLER, J.A.; LEMOS, R.P.; BELTRÁN, D.F.B.; MACHADO, P.C.; DEUS, G.A.; SILVA, H.V.L.; FERREIRA, Y.R. Further Investigation on Frobenius Spectrum for DOA estimation. In: *2015 IEEE 6th Latin American Symposium on Circuits and Systems (LASCAS 2015)*, Fevereiro 2015, Montevideu, Uruguai, v. 1. pp. 1-4; e
6. CRUZ, G.; KUNZLER, J.A.; LEMOS, R.P.; BELTRÁN, D.F.B.; SILVA, H.V.L.; FERREIRA, Y.R. A New Approach for the Differential Spectrum Using the Frobenius Norm. In: *2014 2nd International Conference on Artificial Intelligence, Modelling and Simulation (AIMS)*, Novembro 2014, Madrid, Espanha, v. 1. pp. 350-354.

## 6.3 Trabalhos Futuros

Existem quatro linhas de pesquisa possíveis para o aprimoramento continuado do estimador SEAD Modificado:

1. **Derivação de uma função-objetivo para o DMLE filtrado e recálculo do Método de Newton e BFGS** – ao longo deste trabalho, observou-se que a função-objetivo para o DMLE filtrado proposta em (KRUMMENAUER et al., 2010) não foi calculada considerando o modelo de sinal filtrado. À medida que a SNR aumenta, mesmo que essa função possua os mínimos globais próximos dos valores reais, ela não representa o Estimador de Máxima-Verossimilhança no sentido estrito. Assim, ainda que os resultados obtidos neste trabalho tenham sido bastante satisfatórios, pode-se observar a presença de algumas instabilidades numéricas que podem ser devido a isso. Espera-se que a determinação de um MLE considerando o novo modelo de sinal apresente resultados melhores. Naturalmente, isso exigirá em outro momento o desenvolvimento do Método de Newton e do BFGS;
2. **Desenvolvimento de otimizadores locais baseados em região de confiança** – este trabalho abordou o uso de otimizadores locais utilizando a busca em linha. Além desses, existem outros otimizadores que são baseados em região de confiança, como o Levenberg-Marquardt. Entretanto, eles não foram avaliados neste trabalho. A implementação e a avaliação desses métodos pode trazer resultados importantes em termos do desempenho de estimação e do esforço computacional;
3. **Proposição de novos filtros espaciais** – observou-se que o filtro calculado a partir do subespaço de sinal produz um espaço de soluções que não é apropriado para os métodos de otimização, pois os mínimos locais assumem valores próximos do mínimo global. Além disso, na situação de fontes próximas, o mínimo global é plano e isso dificulta a convergência de otimizadores baseados no gradiente. Adicionalmente, esse filtro apresenta ordem significativamente maior para possuir fase nula de forma que permite a combinação em paralelo dos filtros referentes a cada uma das bandas de passagem. Entretanto, isso aumenta o esforço computacional despendido. Por outro lado, o filtro calculado a partir do espectro diferencial apresenta maior esforço computacional, pois exige o cálculo de diversas amostras do espectro diferencial. Mesmo que esse número de amostras seja reduzido, o tempo de estimação do filtro é maior do que despendido pelos outros filtros abordados neste trabalho. Assim, há espaço para aprimoramentos na proposição de novos filtros espaciais multibanda que sejam mais adequados à estimação DOA de fontes próximas e afastadas ao mesmo tempo em que não elevado possuam esforço computacional;
4. **Alterações futuras ao estágio de pré-seleção** – embora a aplicação do limiar de amplitude no estágio de pré-seleção favoreça muito a estimação DOA de fontes próximas, ele limita bastante a sua aplicação em outras situações. A classificação de picos em picos de sinal e picos de ruído unicamente por suas amplitudes é um processo difícil na presença de diversas fontes próximas e uma afastada mesmo em altas SNRs. Além disso, essa classificação é ainda mais difícil ao observar que um limiar de

amplitude alto pode excluir os picos de sinal e um limiar baixo pode incluir os picos de ruído. O novo estágio de pré-seleção proposto neste trabalho finalmente retirou o limiar de amplitude e aprimorou consideravelmente a estimação de fontes afastadas. Entretanto, houve uma perda de desempenho na situação de fontes próximas e agora é necessário um número maior de operações para gerar o vetor de estimativas iniciais. Dessa forma, um caminho para solucionar este problema *pode* passar por técnicas inspiradas pelo reconhecimento de padrões, com a determinação de um ou mais parâmetros de avaliação dos picos além de suas amplitudes.

## REFERÊNCIAS

- ABRAMOVICH, Y. I.; SPENCER, N. K. DOA estimation performance breakdown: A new approach to prediction and cure. In: *Proceedings of 2002 European Signal Processing Conference - EUSIPCO 2002*. Toulouse, França: [s.n.], 2002. p. 89–92. Citado na página 20.
- ALVES, A. C. B.; ASADA, E.; MONTICELLI, A. Critical evaluation of direct and iterative methods for solving  $ax = b$  systems in power flow calculations and contingency analysis. *IEEE Transactions on Power Systems*, v. 14, n. 2, p. 702–708, 1999. Citado na página 47.
- ALVES, C. A. *Análise Teórica e Experimental de Métodos de Estimação de DOA e de Estimação de Frequência com Alta Resolução*. Tese (Doutorado) — Universidade Estadual de Campinas, Campinas, Brasil, Julho 2004. Citado 20 vezes nas páginas 19, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 36, 38, 40, 43, 44, 46, 50, 76 e 77.
- BOAVENTURA, A. M. S. *Avaliação de Métodos de Estimação da Direção de Chegada de Sinais em Sistemas de Comunicações Celulares*. Dissertação (Mestrado) — Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brasil, 1998. Citado na página 18.
- BRESLER, Y.; MACOVSKI, A. Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 34, n. 5, p. 1081–1089, 1986. Citado 2 vezes nas páginas 33 e 43.
- BROYDEN, C. G. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA Journal of Applied Mathematics*, v. 6, n. 1, p. 76–90, Março 1970. Citado na página 62.
- CARVALHO, M. H. de et al. *Uma Introdução Sucinta a Algoritmos de Aproximação*. [S.l.], 2001. Disponível em: <<http://www.ime.usp.br/dcc/livros/aprox>>. Citado 2 vezes nas páginas 50 e 62.
- CHAPRA, S. C.; CANALE, R. P. *Numerical Methods for Engineers*. 6<sup>a</sup>. ed. [S.l.]: McGraw-Hill, 2010. Citado 3 vezes nas páginas 76, 78 e 79.
- CRAMÉR, H. *Mathematical Methods of Statistics*. Princeton, EUA: Princeton University Press, 1946. Citado na página 32.
- FERREIRA, Y. R. *Método de Estimação de Ângulos DOA*. Dissertação (Mestrado) — Universidade Federal de Goiás, Goiânia, Brasil, Abril 2005. Citado 3 vezes nas páginas 20, 40 e 47.
- FERREIRA, Y. R.; LEMOS, R. P. A new DOA estimation algorithm based on angle search through the difference between the principal singular values. In: *Proceedings of the International Microwave and Optoelectronics Conference*. Brasília, Brasil: [s.n.], 2005. p. 283–286. Citado 2 vezes nas páginas 20 e 40.



- FERREIRA, Y. R.; LEMOS, R. P. A new DOA estimation algorithm based on differential spectrum. In: *Proceedings of the Eighth International Symposium on Signal Processing and Its Applications*. Sydney, Austrália: [s.n.], 2005. p. 303–307. Citado 2 vezes nas páginas [20](#) e [40](#).
- FLETCHER, R. A new approach to variable metric algorithms. *The Computer Journal*, v. 13, n. 3, p. 317–322, Agosto 1970. Citado na página [62](#).
- FORSTER, P.; VEZZOSI, G. Application of spheroidal sequences to array processing. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '87)*. Dallas, EUA: [s.n.], 1987. p. 2268–2271. Citado 2 vezes nas páginas [20](#) e [41](#).
- FUJITSU, INC.; CISCO, INC. *Beamforming Boots the Range and Capacity of WiMAX Networks*. [S.l.], 2008. Citado na página [19](#).
- GERSHMAN, A. B.; STOICA, P. New MODE-based techniques for direction finding with an improved threshold performance. *Signal Processing*, v. 76, p. 221–235, 1999. Citado 5 vezes nas páginas [19](#), [40](#), [44](#), [45](#) e [46](#).
- GODARA, L. C. Applications of antenna arrays to mobile communications, part ii: Performance improvement, feasibility and systems considerations. v. 85, n. 7, p. 1031–1060, Julho 1997. Citado na página [19](#).
- GOLDFARB, D. A family of variable-metric methods derived by variational means. *Mathematics of Computation*, v. 24, n. 109, p. 23–26, Janeiro 1970. Citado na página [62](#).
- GOLUB, G. H.; PEREYRA, V. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM Journal on Numerical Analysis*, v. 10, n. 2, p. 413–432, Abril 1973. Citado 4 vezes nas páginas [129](#), [143](#), [166](#) e [179](#).
- HAWKES, M.; NEHORAI, A.; STOICA, P. Performance breakdown of subspace-based methods: Prediction and cure. In: *Proceedings of 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '01)*. Salt Lake City, EUA: [s.n.], 2001. p. 4005–4008. Citado 4 vezes nas páginas [20](#), [40](#), [67](#) e [72](#).
- HAYKIN, S. *Modern Filters*. New York, EUA: Macmillan, 1990. Citado na página [53](#).
- HAYKIN, S. *Adaptive Filter Theory*. 4ª. ed. Englewood Cliffs, EUA: Prentice-Hall, 2001. Citado 5 vezes nas páginas [20](#), [57](#), [58](#), [59](#) e [61](#).
- JOHNSON, B. A.; ABRAMOVICH, Y. I.; MESTRE, X. MUSIC, G-MUSIC, and maximum-likelihood performance breakdown. *IEEE Transactions on Signal Processing*, v. 56, n. 8, p. 3944–3958, Agosto 2008. Citado na página [73](#).
- JOHNSON, B. A.; ABRAMOVICH, Y. I.; MESTRE, X. The role of subspace swap in maximum likelihood estimation performance breakdown. In: *Proceedings of 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '08)*. Las Vegas, EUA: [s.n.], 2008. p. 2469–2472. Citado na página [73](#).
- KAY, S. M. *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Englewood Cliffs, EUA: Prentice Hall, 1993. Citado 2 vezes nas páginas [19](#) e [32](#).
- KRIM, H.; VIBERG, M. Two decades of array signal processing. *IEEE Signal Processing Magazine*, p. 67–97, Julho 1996. Citado 5 vezes nas páginas [25](#), [28](#), [34](#), [35](#) e [163](#).

KRUMMENAUER, R. *Filtragem Ótima na Estimação de Direção de Chegada de Ondas Planas Usando Arranjo de Sensores*. Dissertação (Mestrado) — Universidade Estadual de Campinas, Campinas, Brasil, Julho 2007. Citado 4 vezes nas páginas 41, 55, 57 e 58.

KRUMMENAUER, R. *Sobre Modelos e Métodos de Estimação de Direção de Chegada de Sinais em Sistemas Digitais de Comunicação Sem Fio*. Tese (Doutorado) — Universidade Estadual de Campinas, Campinas, Brasil, Agosto 2011. Citado na página 18.

KRUMMENAUER, R. et al. Improving the threshold performance of maximum likelihood estimation of direction of arrival. *Signal Processing*, v. 90, n. 11, p. 1582–1590, Novembro 2010. Citado 23 vezes nas páginas 20, 21, 22, 23, 30, 41, 46, 55, 56, 57, 58, 59, 60, 61, 63, 67, 70, 83, 85, 114, 119, 140 e 178.

KRUMMENAUER, R. et al. Maximum likelihood-based direction-of-arrival estimator for discrete sources. *Circuits Systems and Signal Processing*, v. 32, p. 2423–2443, Março 2013. Citado na página 27.

KUMARESAN, R.; SCHARF, L. L.; SHAW, A. An algorithm for pole-zero modeling and spectral analysis. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. 34, n. 3, p. 637–640, 1986. Citado 3 vezes nas páginas 33, 42 e 43.

LI, F.; VACCARO, R. J. Unified analysis for DOA estimation algorithms in array signal processing. *Signal Processing*, v. 25, p. 147–169, 1991. Citado na página 31.

LI, F.; VACCARO, R. J.; TUFTS, D. W. Unified performance analysis of subspace-based estimation algorithms. In: *Proceedings of 1990 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '90)*. Albuquerque, EUA: [s.n.], 1990. p. 2575–2578. Citado na página 20.

LI, M.; LU, Y. A refined genetic algorithm for accurate and reliable DOA estimation with a sensor array. *Wireless Personal Communications*, v. 43, n. 2, p. 533–547, 2007. ISSN 1572-834X. Citado 3 vezes nas páginas 19, 33 e 76.

LIBERTI, J. C.; RAPPAPORT, T. S. *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*. Upper Saddle River, EUA: Prentice-Hall, 1999. Citado 2 vezes nas páginas 18 e 25.

LOPES, A. et al. Improving the MODEX algorithm for direction estimation. *Signal Processing*, v. 83, n. 9, p. 2047–2051, Setembro 2003. Citado 4 vezes nas páginas 19, 40, 45 e 46.

MERCER, J. Functions of positive and negative type and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society of London Series A*, v. 209, n. 441–458, p. 415–446, Janeiro 1909. Citado na página 66.

NOCEDAL, J.; WRIGHT, S. J. *Numerical Optimization*. 2ª. ed. Nova York, EUA: Springer, 2006. Citado 4 vezes nas páginas 78, 79, 80 e 81.

OTTERSTEN, B. *Parametric Subspace Fitting Methods for Array Signal Processing*. Tese (Doutorado) — Stanford University, Stanford, EUA, Dezembro 1989. Citado 2 vezes nas páginas 33 e 77.

PETERSEN, K. B.; PEDERSEN, M. S. *The Matrix Cookbook*. 2012. Citado 4 vezes nas páginas 126, 151, 155 e 160.

- SHAGHAGHI, M.; VOROBYOV, S. A. Subspace leakage analysis and improved DOA estimation with small sample size. *IEEE Transactions on Signal Processing*, v. 63, n. 12, p. 3251–3265, Junho 2015. Citado na página 72.
- SHANNO, D. Conditioning of quasi-newton methods for function minimization. *Mathematics of Computation*, v. 24, n. 111, p. 647–656, Julho 1970. Citado na página 62.
- SILVA, F. J. da. *Uso de Filtragem em Métodos de Estimação DOA Através de Arranjo de Sensores*. Dissertação (Mestrado) — Universidade Estadual de Campinas, Campinas, Brasil, Julho 2007. Citado 10 vezes nas páginas 25, 27, 28, 29, 33, 42, 43, 44, 45 e 46.
- SILVA, H. V. L. et al. A branch-and-bound inspired technique to improve the computational efficiency of DOA estimation. *Signal Processing*, v. 93, n. 4, p. 947–956, Abril 2013. ISSN 0165-1684. Citado 8 vezes nas páginas 20, 21, 41, 48, 50, 51, 53 e 66.
- SILVA, H. V. L. e. *Redução da Complexidade Computacional do Método de Estimação de Ângulos de Incidência Através da Diferença Entre os Valores Singulares da Matriz de Covariância Espacial*. Dissertação (Mestrado) — Universidade Federal de Goiás, Goiânia, Brasil, Março 2009. Citado 3 vezes nas páginas 20, 41 e 51.
- SIPSER, M. F. *Introduction to the Theory of Computation*. Boston, EUA: PWS Publishing Company, 1997. Citado na página 50.
- SORELIUS, J. *Subspace-Based Parameter Estimation Problems in Signal Processing*. Tese (Doutorado) — Uppsala University, Uppsala, Suécia, 1999. Citado 2 vezes nas páginas 18 e 25.
- STARER, D. *Algorithms for Polynomial-Based Signal Processing*. Tese (Doutorado) — Yale University, New Haven, EUA, Maio 1990. Citado 12 vezes nas páginas 20, 33, 42, 43, 62, 76, 77, 78, 81, 126, 127 e 171.
- STARER, D.; NEHORAI, A. Newton algorithms for conditional and unconditional maximum likelihood estimation of the parameters of exponential signals in noise. *IEEE Transactions on Signal Processing*, v. 40, n. 6, p. 1528–1534, Junho 1992. Citado 9 vezes nas páginas 19, 20, 21, 62, 77, 116, 126, 127 e 171.
- STOICA, P.; NEHORAI, A. Performance study of conditional and unconditional direction-of-arrival estimation. *IEEE Transactions on Acoustic, Speech and Signal Processing*, v. 38, n. 10, p. 1783–1795, Outubro 1990. Citado 5 vezes nas páginas 29, 30, 31, 38 e 39.
- STOICA, P.; SHARMAN, K. C. Novel eigenanalysis method for direction estimation. *IEEE Proceedings F – Radar and Signal Processing*, v. 137, n. 1, p. 19–26, Fevereiro 1990. ISSN 0956-375X. Citado 3 vezes nas páginas 19, 40 e 43.
- THOMAS, J. K.; SCHARF, L. L.; TUFTS, D. W. The probability of a subspace swap in the SVD. *IEEE Transactions on Signal Processing*, v. 43, n. 3, p. 730–736, Março 1995. Citado 2 vezes nas páginas 20 e 72.
- VAN TREES, H. L. *Optimum Array Processing. Part IV of Detection, Estimation and Modulation Theory*. New York, EUA: John Wiley and Sons, 2002. Citado na página 31.
- YANG, W. Y. et al. *Adaptive Filter Theory*. Hoboken, EUA: Wiley-Interscience, 2005. Citado 2 vezes nas páginas 78 e 79.

## Apêndices

# A DERIVAÇÃO DO MÉTODO DE NEWTON PARA O DMLE SEM FILTRAGEM ESPACIAL USANDO A MATRIZ DE PREDIÇÃO LINEAR

## A.1 Introdução

A função-objetivo do Critério Determinístico do Estimador de Máxima Verossimilhança (DMLE) sem filtragem espacial e usando a matriz de predição linear  $\mathbf{B}$  é (STARER, 1990; STARER; NEHORAI, 1992):

$$F_{\text{DMLE2}} = \text{tr}(\mathbf{P}_\mathbf{B} \hat{\mathbf{R}})$$

em que:  $\mathbf{P}_\mathbf{B} = \mathbf{B}\mathbf{B}^+$ . Considerando o Arranjo Linear Uniforme,  $\mathbf{B}$  é a seguinte matriz Toeplitz construída a partir do polinômio  $\mathbf{b}$ :

$$\mathbf{B} = \begin{bmatrix} b_M^* & \mathbf{0} \\ \vdots & b_M^* \\ b_0^* & \vdots & b_M^* \\ & b_0^* & \vdots \\ \mathbf{0} & & b_0^* \end{bmatrix} \in \mathbb{C}^{K \times (K-M)}$$

em que:  $\mathbf{b} = [b_0 \ \dots \ b_M]^T \in \mathbb{C}^{(M+1) \times 1}$  é o vetor contendo os coeficientes do polinômio cujas raízes são as frequências angulares dos DOA estimados usando a relação apresentada na seção 2.1.

No entanto,  $\mathbf{b}$  é uma variável complexa. Assim, toda a derivação do Método de Newton é realizada em função do parâmetro  $\boldsymbol{\phi} \triangleq [\Re(\mathbf{b}) \ \Im(\mathbf{b})]^T \in \mathbb{R}^{2(M+1) \times 1}$ . Dessa forma, os contadores  $j$  e  $k$  variam entre 0 e  $2M + 1$ . Então, para  $0 \leq (j, k) \leq M$ , as derivações estão em função da parte real de  $\mathbf{b}$  e, para  $(M + 1) \leq j, k \leq (2M + 1)$ , as derivações correspondem à parte imaginária de  $\mathbf{b}$ .

Assim, os vetores de parâmetros  $\boldsymbol{\phi}_0$  e  $\boldsymbol{\phi}_1$  apresentam dimensões  $2(M + 1) \times 1$ . Consequentemente, a matriz hessiana  $\mathbf{E}_k$  e o vetor gradiente  $\mathbf{g}_k$  possuem respectivamente as dimensões  $2(M + 1) \times 2(M + 1)$  e  $2(M + 1) \times 1$  e apresentam uma estrutura que permite realizar o produto complexo  $r = pq$  da seguinte forma (PETERSEN; PEDERSEN, 2012):

$$r = \begin{bmatrix} \Re r \\ \Im r \end{bmatrix} = \begin{bmatrix} \Re p & -\Im p \\ \Im p & \Re p \end{bmatrix} = \begin{bmatrix} \Re q \\ \Im q \end{bmatrix}$$

Nesse contexto, a fórmula de correção do Método de Newton é:

$$\hat{\phi}_{(k+1)} = \hat{\phi}_k - \mathbf{E}_k^{-1} \mathbf{g}_k$$

Ou, na forma matricial:

$$\begin{bmatrix} \Re \hat{\mathbf{b}}_1 \\ \Im \hat{\mathbf{b}}_1 \end{bmatrix} = \begin{bmatrix} \Re \hat{\mathbf{b}}_0 \\ \Im \hat{\mathbf{b}}_0 \end{bmatrix} - \begin{bmatrix} \Re \bar{\mathbf{E}}_k & -\Im \bar{\mathbf{E}}_k \\ \Im \bar{\mathbf{E}}_k & \Re \bar{\mathbf{E}}_k \end{bmatrix}^{-1} \begin{bmatrix} \Re \bar{\mathbf{g}}_k \\ \Im \bar{\mathbf{g}}_k \end{bmatrix}$$

em que:  $\Re \bar{\mathbf{E}}_k$ ,  $\Im \bar{\mathbf{E}}_k$ ,  $\Re \bar{\mathbf{g}}_k$  e  $\Im \bar{\mathbf{g}}_k$  são partições da hessiana e do gradiente relativos às partes real e imaginária da correção do vetor de parâmetros complexos  $\mathbf{b}$ . Assim, ao final da otimização, a estimativa em termos de  $\mathbf{b}$  é calculada somando a primeira metade do vetor  $\hat{\phi}_1$  à sua segunda metade, porém multiplicada por  $\sqrt{-1}$ . Finalmente, a estimativa em termos das frequências angulares  $\hat{\omega}$  é obtida utilizando o procedimento descrito na seção 3.2. Ainda, utiliza-se a relação apresentada na seção 2.1 para calcular as estimativas em função dos ângulos DOA.

É importante ressaltar que o Método de Newton apresentado neste Apêndice foi desenvolvido em (STARER, 1990; STARER; NEHORAI, 1992). Contudo, neste trabalho ele foi derivado novamente, sendo obtidos os mesmos resultados alcançados pelos autores originais. O propósito aqui é documentar todo esse desenvolvimento para consulta posterior.

## A.2 Derivada de B

Considerando que:

$$\mathbf{B} = \begin{bmatrix} b_M^* & & \mathbf{0} \\ \vdots & b_M^* & \\ b_0^* & \vdots & b_M^* \\ & b_0^* & \vdots \\ \mathbf{0} & & b_0^* \end{bmatrix}$$

Então:

$$\frac{\partial \mathbf{B}}{\partial \phi_j} = \begin{bmatrix} \frac{\partial b_M^*}{\partial \phi_j} & & \mathbf{0} \\ \vdots & \frac{\partial b_M^*}{\partial \phi_j} & \\ \frac{\partial b_0^*}{\partial \phi_j} & \vdots & \frac{\partial b_M^*}{\partial \phi_j} \\ & \frac{\partial b_0^*}{\partial \phi_j} & \vdots \\ \mathbf{0} & & \frac{\partial b_0^*}{\partial \phi_j} \end{bmatrix}$$

Assim, cada derivada parcial é calculada considerando a independência entre as partes real e imaginária de  $b_c^*$ ,  $0 \leq c \leq M$  e que  $\Re z^* = \Re z$  e  $\Im z^* = -\Im z$ :

$$\frac{\partial b_c^*}{\partial \phi_j} = \frac{\partial \Re b_c^*}{\partial \phi_j} + i \frac{\partial \Im b_c^*}{\partial \phi_j} = \frac{\partial \Re b_c}{\partial \phi_j} - i \frac{\partial \Im b_c}{\partial \phi_j}$$

O desenvolvimento abaixo está separado para a parte real de  $\mathbf{b}$ ,  $0 \leq j \leq M$ , apresentado à esquerda, e para a parte imaginária de  $\mathbf{b}$ ,  $(M+1) \leq j \leq 2(M+1)$ , apresentado à direita. É possível observar que, respectivamente, as partes imaginária e real são independentes de  $\phi_j$  no desenvolvimento à esquerda e à direita, assim:

$$\left. \begin{array}{l} \frac{\partial b_c^*}{\partial \phi_j} = \frac{\partial \Re b_c}{\partial \phi_j} \end{array} \right| \begin{array}{l} \frac{\partial b_c^*}{\partial \phi_j} = -i \frac{\partial \Im b_c}{\partial \phi_j} \end{array}$$

Considerando  $j = c$  e  $j = c + M + 1$ , tem-se respectivamente que  $\phi_j = \Re b_c$  e  $\phi_j = \Im b_c$  e, assim, obtemos:

$$\left. \begin{array}{l} \frac{\partial b_c^*}{\partial \phi_j} = \frac{\partial \Re b_c}{\partial \phi_j} = 1 \end{array} \right| \begin{array}{l} \frac{\partial b_c^*}{\partial \phi_j} = -i \frac{\partial \Im b_c}{\partial \phi_j} = -i \end{array}$$

Então, para  $\phi_0$  (à esquerda) e  $\phi_{(M+1)}$  (à direita):

$$\left. \begin{array}{l} \frac{\partial \mathbf{B}}{\partial \phi_0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \right| \begin{array}{l} \frac{\partial \mathbf{B}}{\partial \phi_{(M+1)}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & -i \end{bmatrix} \end{array}$$

E, para  $\phi_M$  (à esquerda) e  $\phi_{2(M+1)}$  (à direita):

$$\left. \begin{array}{l} \frac{\partial \mathbf{B}}{\partial \phi_M} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array} \right| \begin{array}{l} \frac{\partial \mathbf{B}}{\partial \phi_{2(M+1)}} = \begin{bmatrix} -i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & -i \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

Finalmente, para  $0 \leq j \leq M$  e  $j = c$  no caso real à esquerda, e para  $(M+1) \leq j \leq 2(M+1)$  e  $c = j + M + 1$  no caso imaginário à direita, tem-se:

$$\left. \boxed{\frac{\partial \mathbf{B}}{\partial \phi_j} = \mathbf{Z}_c} \right| \boxed{\frac{\partial \mathbf{B}}{\partial \phi_j} = -i \mathbf{Z}_c}$$



em que:  $\mathbf{Z}_c$  é uma matriz de deslocamento que contém 1's ao longo da  $(M - c)$ -ésima diagonal e 0 no restante, como definido abaixo:

$$\mathbf{Z}_c = \begin{bmatrix} \mathbf{0}_{(M+1-c) \times (K-M)} \\ \mathbf{I}_{(K-M) \times (K-M)} \\ \mathbf{0}_{(c-1) \times (K-M)} \end{bmatrix}$$

### A.3 Derivada de $\mathbf{B}^+$

A derivada da pseudo-inversa de Moore-Penrose é (GOLUB; PEREYRA, 1973):

$$\frac{\partial \mathbf{B}^+}{\partial \phi_j} = -\mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{B}^+ (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{I} - \mathbf{B} \mathbf{B}^+) + (\mathbf{I} - \mathbf{B}^+ \mathbf{B}) \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \mathbf{B}^+$$

Mas  $\mathbf{B}^+ \mathbf{B} = \mathbf{I}$ , então:

$$\frac{\partial \mathbf{B}^+}{\partial \phi_j} = -\mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{B}^+ (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{I} - \mathbf{B} \mathbf{B}^+)$$

Entretanto,  $\mathbf{P}_\mathbf{B}^\perp = (\mathbf{I} - \mathbf{P}_\mathbf{B}) = (\mathbf{I} - \mathbf{B} \mathbf{B}^+)$ , assim:

$$\frac{\partial \mathbf{B}^+}{\partial \phi_j} = -\mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{B}^+ (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_\mathbf{B}^\perp$$

No entanto,  $\mathbf{B}^+ (\mathbf{B}^+)^H = (\mathbf{B}^H \mathbf{B})^{-1}$ , então:

$$\frac{\partial \mathbf{B}^+}{\partial \phi_j} = -\mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_\mathbf{B}^\perp$$

Reordenando:

$$\boxed{\frac{\partial \mathbf{B}^+}{\partial \phi_j} = (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_\mathbf{B}^\perp - \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+}$$

### A.4 Derivada de $\mathbf{P}_\mathbf{B}$

Considerando que  $\mathbf{P}_\mathbf{B} = \mathbf{B} \mathbf{B}^+$ , então:

$$\frac{\partial \mathbf{P}_\mathbf{B}}{\partial \phi_j} = \frac{\partial \mathbf{B} \mathbf{B}^+}{\partial \phi_j}$$

Aplicando a regra da derivada do produto:

$$\frac{\partial \mathbf{P}_\mathbf{B}}{\partial \phi_j} = \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{B} \frac{\partial \mathbf{B}^+}{\partial \phi_j}$$

Mas  $\frac{\partial \mathbf{B}^+}{\partial \phi_j} = (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_\mathbf{B}^\perp - \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+$ , portanto:

$$\frac{\partial \mathbf{P}_\mathbf{B}}{\partial \phi_j} = \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{B} \left[ (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_\mathbf{B}^\perp - \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \right]$$

Efetuando a multiplicação:

$$\frac{\partial \mathbf{P}_B}{\partial \phi_j} = \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp - \mathbf{B} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+$$

Mas  $(\mathbf{B}^+)^H = \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1}$ , assim:

$$\frac{\partial \mathbf{P}_B}{\partial \phi_j} = \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp - \mathbf{B} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+$$

Reordenando:

$$\frac{\partial \mathbf{P}_B}{\partial \phi_j} = \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ - \mathbf{B} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp$$

Colocando um termo em evidência e considerando que  $\mathbf{A}\mathbf{I} = \mathbf{I}\mathbf{A} = \mathbf{A}$ :

$$\frac{\partial \mathbf{P}_B}{\partial \phi_j} = (\mathbf{I} - \mathbf{B} \mathbf{B}^+) \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp$$

Mas  $\mathbf{P}_B^\perp = (\mathbf{I} - \mathbf{P}_B) = (\mathbf{I} - \mathbf{B} \mathbf{B}^+)$ , então:

$$\frac{\partial \mathbf{P}_B}{\partial \phi_j} = \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp$$

Reordenando:

$$\boxed{\frac{\partial \mathbf{P}_B}{\partial \phi_j} = (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp + \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+}$$

## A.5 Gradiente de $\text{tr}(\mathbf{P}_B \hat{\mathbf{R}})$

Derivando a função-objetivo produz:

$$\nabla_{\phi_j} \text{tr}(\mathbf{P}_B \hat{\mathbf{R}}) = \frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j}$$

Entretanto,  $\frac{\partial \text{tr} \mathbf{X}}{\partial x} = \text{tr} \left( \frac{\partial \mathbf{X}}{\partial x} \right)$ , então:

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = \text{tr} \left( \frac{\partial \mathbf{P}_B \hat{\mathbf{R}}}{\partial \phi_j} \right)$$

Aplicando a regra da derivada do produto:

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = \text{tr} \left( \frac{\partial \mathbf{P}_B}{\partial \phi_j} \hat{\mathbf{R}} + \mathbf{P}_B \frac{\partial \hat{\mathbf{R}}}{\partial \phi_j} \right)$$

Mas  $\hat{\mathbf{R}}$  é constante, assim:

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = \text{tr} \left( \frac{\partial \mathbf{P}_B}{\partial \phi_j} \hat{\mathbf{R}} \right)$$

Porém,  $\frac{\partial \mathbf{P}_B}{\partial \phi_j} = (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp + \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+$ , então:

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = \text{tr} \left\{ \left[ (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp + \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \right] \hat{\mathbf{R}} \right\}$$

Efetutando a multiplicação:

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = \text{tr} \left[ (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp \hat{\mathbf{R}} + \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \hat{\mathbf{R}} \right]$$

Mas  $\mathbf{P}_B^\perp = \mathbf{P}_B^{\perp H}$  e  $\hat{\mathbf{R}} = \hat{\mathbf{R}}^H$ , então:

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = \text{tr} \left[ (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^{\perp H} \hat{\mathbf{R}}^H + \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \hat{\mathbf{R}} \right]$$

Aplicando  $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$ :

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = \text{tr} \left[ (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^{\perp H} \hat{\mathbf{R}}^H \right] + \text{tr} \left( \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \hat{\mathbf{R}} \right)$$

Aplicando  $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BCA}) = \text{tr}(\mathbf{CAB})$ :

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = \text{tr} \left[ (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^{\perp H} \hat{\mathbf{R}}^H \right] + \text{tr} \left( \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \right)$$

Aplicando  $(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H$ :

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = \text{tr} \left[ \left( \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \right)^H \right] + \text{tr} \left( \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \right)$$

Porém,  $\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}^T)$  e  $\mathbf{A}^H = \mathbf{A}^{T*}$ , então  $\text{tr}(\mathbf{A}^H) = \text{tr}[(\mathbf{A}^T)^*] = \text{tr}(\mathbf{A}^*) = \text{tr}(\mathbf{A})^*$ .  
Consequentemente:

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = \text{tr} \left( \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \right)^* + \text{tr} \left( \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \right)$$

Mas  $z + z^* = 2\Re z$ , dessa forma:

$$\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = 2\Re \text{tr} \left( \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \right)$$

Aplicando  $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BCA}) = \text{tr}(\mathbf{CAB})$ :

$$\boxed{\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = 2\Re \text{tr} \left( \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right)}$$

## A.6 Hessiana de $\text{tr}(\mathbf{P}_B \hat{\mathbf{R}})$

Derivando o gradiente da função-objetivo produz:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} &= \frac{\partial \nabla_{\phi_j} \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k} \\ &= \frac{\partial}{\partial \phi_k} \left[ \nabla_{\phi_j} \text{tr}(\mathbf{P}_B \hat{\mathbf{R}}) \right] \\ &= \frac{\partial}{\partial \phi_k} \left[ \frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} \right] \end{aligned}$$

Porém,  $\frac{\partial \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_j} = 2 \Re \text{tr}(\mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j})$ , dessa forma:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = \frac{\partial}{\partial \phi_k} \left[ 2 \Re \text{tr} \left( \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right) \right]$$

Aplicando  $\frac{\partial c\mathbf{X}}{\partial x} = c \frac{\partial \mathbf{X}}{\partial x}$  para  $c$  constante:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \left\{ \frac{\partial}{\partial \phi_k} \left[ \text{tr} \Re \left( \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right) \right] \right\}$$

Aplicando  $\frac{\partial \Re(\mathbf{X})}{\partial x} = \Re \left( \frac{\partial \mathbf{X}}{\partial x} \right)$ :

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \left\{ \frac{\partial}{\partial \phi_k} \left[ \text{tr} \left( \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right) \right] \right\}$$

Aplicando  $\frac{\partial \text{tr}(\mathbf{X})}{\partial x} = \text{tr} \left( \frac{\partial \mathbf{X}}{\partial x} \right)$ :

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ \frac{\partial}{\partial \phi_k} \left( \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right) \right]$$

Aplicando a derivada do produto:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} &= 2 \Re \text{tr} \left[ \frac{\partial \mathbf{B}^+}{\partial \phi_k} \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} + \mathbf{B}^+ \frac{\partial \hat{\mathbf{R}}}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} + \mathbf{B}^+ \hat{\mathbf{R}} \frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} \frac{\partial \mathbf{B}}{\partial \phi_j} + \right. \\ &\quad \left. + \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial}{\partial \phi_k} \left( \frac{\partial \mathbf{B}}{\partial \phi_j} \right) \right] \end{aligned}$$

Mas  $\hat{\mathbf{R}}$  e  $\frac{\partial \mathbf{B}}{\partial \phi_j}$  são constantes em relação a  $\phi_k$ , então:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left( \frac{\partial \mathbf{B}^+}{\partial \phi_k} \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} + \mathbf{B}^+ \hat{\mathbf{R}} \frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} \frac{\partial \mathbf{B}}{\partial \phi_j} \right)$$

Mas  $\frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} = -\frac{\partial \mathbf{P}_B}{\partial \phi_k}$ , portanto:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left( \frac{\partial \mathbf{B}^+}{\partial \phi_k} \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \hat{\mathbf{R}} \frac{\partial \mathbf{P}_B}{\partial \phi_k} \frac{\partial \mathbf{B}}{\partial \phi_j} \right)$$

Mas  $\frac{\partial \mathbf{P}_B}{\partial \phi_k} = (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp + \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+$ , portanto:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left\{ \frac{\partial \mathbf{B}^+}{\partial \phi_k} \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \hat{\mathbf{R}} \left[ (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp + \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \right] \frac{\partial \mathbf{B}}{\partial \phi_j} \right\}$$

Efetuada a multiplicação:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ \frac{\partial \mathbf{B}^+}{\partial \phi_k} \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \hat{\mathbf{R}} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \right]$$

Mas  $\frac{\partial \mathbf{B}^+}{\partial \phi_k} = (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp - \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+$ , assim:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left\{ \left[ (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp - \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \right] \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \right. \\ \left. - \mathbf{B}^+ \hat{\mathbf{R}} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \right\} \end{aligned}$$

Efetuada a multiplicação:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \right. \\ \left. - \mathbf{B}^+ \hat{\mathbf{R}} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \right] \end{aligned}$$

Reordenando:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \right. \\ \left. - \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \hat{\mathbf{R}} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right] \end{aligned}$$

## A.7 Aproximação assintótica na hessiana de $\text{tr}(\mathbf{P}_B \hat{\mathbf{R}})$

A hessiana é:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \right. \\ \left. - \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \hat{\mathbf{R}} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right] \end{aligned}$$

Considerando a seção A.12,  $\mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{B}^+ \mathbf{R} \mathbf{P}_B^\perp = \mathbf{0}$ , assim:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \hat{\mathbf{R}} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right]$$

Entretanto, considerando a seção A.13,  $\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{P}_B^\perp \mathbf{R} \mathbf{P}_B^\perp = \mathbf{R} \mathbf{P}_B^\perp$ , dessa forma:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{R} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \mathbf{B}^+ \hat{\mathbf{R}} \left( \mathbf{B}^+ \right)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right]$$

Ainda, considerando a seção A.14,  $\mathbf{B}^+ \hat{\mathbf{R}} \left( \mathbf{B}^+ \right)^H \approx \mathbf{B}^+ \mathbf{R} \left( \mathbf{B}^+ \right)^H = \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1}$ , assim:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{R} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right]$$

Mas  $\mathbf{P}_B^\perp = (\mathbf{I} - \mathbf{B} \mathbf{B}^+)$  e  $\mathbf{R} = \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}$ , então:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} &= 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \left( \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I} \right) \left( \mathbf{I} - \mathbf{B} \mathbf{B}^+ \right) \frac{\partial \mathbf{B}}{\partial \phi_j} - \right. \\ &\quad \left. - \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right] \end{aligned}$$

Realizando a multiplicação, considerando que  $\mathbf{I} \mathbf{I} = \mathbf{I}$ :

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} &= 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \left( \mathbf{A} \mathbf{C} \mathbf{A}^H - \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{B} \mathbf{B}^+ + \sigma - \sigma \mathbf{B} \mathbf{B}^+ \right) \frac{\partial \mathbf{B}}{\partial \phi_j} - \right. \\ &\quad \left. - \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right] \end{aligned}$$

Realizando as multiplicações à esquerda e à direita:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} &= 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{A} \mathbf{C} \mathbf{A}^H \frac{\partial \mathbf{B}}{\partial \phi_j} - \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{B} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} + \right. \\ &\quad + \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \frac{\partial \mathbf{B}}{\partial \phi_j} - \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{B} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} - \\ &\quad \left. - \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right] \end{aligned}$$

Mas  $\mathbf{A}^H \mathbf{B} = \mathbf{0}$ , pois  $\mathbf{B}$  abrange o espaço nulo de  $\mathbf{A}^H$ , assim:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} &= 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{A} \mathbf{C} \mathbf{A}^H \frac{\partial \mathbf{B}}{\partial \phi_j} + \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \frac{\partial \mathbf{B}}{\partial \phi_j} - \right. \\ &\quad \left. - \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{B} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} - \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right] \end{aligned}$$

Colocando dois termos em evidência:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} &= 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{A} \mathbf{C} \mathbf{A}^H \frac{\partial \mathbf{B}}{\partial \phi_j} + \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \left( \mathbf{I} - \mathbf{B} \mathbf{B}^+ \right) \frac{\partial \mathbf{B}}{\partial \phi_j} - \right. \\ &\quad \left. - \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right] \end{aligned}$$

Mas  $\mathbf{P}_B^\perp = (\mathbf{I} - \mathbf{B}\mathbf{B}^+)$ , então:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} &= 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{A} \mathbf{C} \mathbf{A}^H \frac{\partial \mathbf{B}}{\partial \phi_j} + \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} - \right. \\ &\quad \left. - \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \right] \end{aligned}$$

Realizando a subtração:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{A} \mathbf{C} \mathbf{A}^H \frac{\partial \mathbf{B}}{\partial \phi_j} \right]$$

E pode ser reescrita como:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \left( \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I} - \sigma \mathbf{I} \right) \frac{\partial \mathbf{B}}{\partial \phi_j} \right]$$

Mas  $\mathbf{R} = \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}$ , portanto:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{R} - \sigma \mathbf{I}) \frac{\partial \mathbf{B}}{\partial \phi_j} \right]$$

Definindo  $\tilde{\mathbf{R}} \triangleq \mathbf{R} - \sigma \mathbf{I}$  produz:

$$\boxed{\frac{\partial^2 \text{tr}(\mathbf{P}_B \hat{\mathbf{R}})}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \tilde{\mathbf{R}} \frac{\partial \mathbf{B}}{\partial \phi_j} \right]}$$

## A.8 Prova 1: $\mathbf{B}^+ \mathbf{B} = \mathbf{I}$

Como  $\mathbf{B}$  possui posto cheio, realiza-se o mesmo procedimento mostrado na seção C.9, assim:

$$\boxed{\mathbf{B}^+ \mathbf{B} = \mathbf{I}}$$

## A.9 Prova 2: $(\mathbf{B}^+)^H = \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1}$

Realiza-se o mesmo procedimento mostrado na seção C.10, assim:

$$\boxed{(\mathbf{B}^+)^H = \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1}}$$

## A.10 Prova 3: $\mathbf{B}^+ (\mathbf{B}^+)^H = (\mathbf{B}^H \mathbf{B})^{-1}$

Realiza-se o mesmo procedimento mostrado na seção C.11, assim:

$$\boxed{\mathbf{B}^+ (\mathbf{B}^+)^H = (\mathbf{B}^H \mathbf{B})^{-1}}$$



### A.11 Prova 4: $\frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} = -\frac{\partial \mathbf{P}_B}{\partial \phi_k}$

Considerando que  $\mathbf{P}_B^\perp = \mathbf{I} - \mathbf{P}_B$ , então:

$$\frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} = \frac{\partial (\mathbf{I} - \mathbf{P}_B)}{\partial \phi_k}$$

Mas  $\mathbf{I}$  é constante em relação a  $\phi_k$ , assim:

$$\frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} = \frac{\partial (-\mathbf{P}_B)}{\partial \phi_k}$$

Aplicando  $\frac{\partial(c\mathbf{A})}{\partial a} = c\frac{\partial \mathbf{A}}{\partial a}$ :

$$\boxed{\frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} = -\frac{\partial \mathbf{P}_B}{\partial \phi_k}}$$

### A.12 Prova 5: $\mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{B}^+ \mathbf{R} \mathbf{P}_B^\perp = \mathbf{0}$

Sabendo que  $\mathbf{B}$  possui posto cheio, ao aproximar  $\hat{\mathbf{R}}$  por  $\mathbf{R}$  e ao considerar que  $\mathbf{B}$  está próximo do valor ótimo obtém-se:

$$\mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{B}^+ \mathbf{R} \mathbf{P}_B^\perp$$

Mas  $\mathbf{B}^+ = (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H$ ,  $\mathbf{P}_B^\perp = (\mathbf{I} - \mathbf{P}_B) = (\mathbf{I} - \mathbf{B} \mathbf{B}^+) = [\mathbf{I} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H]$  e  $\mathbf{R} = \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}$ , então:

$$\mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx [(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H] (\mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}) [\mathbf{I} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H]$$

Realizando a multiplicação:

$$\mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx [(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{A} \mathbf{C} \mathbf{A}^H + (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \sigma \mathbf{I}] [\mathbf{I} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H]$$

Realizando a multiplicação:

$$\begin{aligned} \mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx & (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{A} \mathbf{C} \mathbf{A}^H + (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \sigma \mathbf{I} - (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{B} - \\ & - (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \sigma \mathbf{I} \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \end{aligned}$$

Mas  $\mathbf{B}^H \mathbf{A} = \mathbf{0}$  e  $\mathbf{A}^H \mathbf{B} = \mathbf{0}$  pois  $\mathbf{B}$  abrange o espaço nulo de  $\mathbf{A}^H$ , então:

$$\mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \sigma \mathbf{I} - (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \sigma \mathbf{I} \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H = \mathbf{0}$$

Mas  $c\mathbf{I}\mathbf{A} = c\mathbf{A}$  e  $c\mathbf{A} = \mathbf{A}c$ , portanto:

$$\mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \sigma (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H - \sigma (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H = \mathbf{0}$$

Entretanto,  $\mathbf{A} \mathbf{A}^{-1} = \mathbf{A}^{-1} \mathbf{A} = \mathbf{I}$ , assim:

$$\mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \sigma (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H - \sigma (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H = \mathbf{0}$$

Realizando a subtração:

$$\boxed{\mathbf{B}^+ \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{0}}$$

### A.13 Prova 6: $\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{P}_B^\perp \mathbf{R} \mathbf{P}_B^\perp = \mathbf{R} \mathbf{P}_B^\perp$

Sabendo que  $\mathbf{B}$  possui posto cheio, ao aproximar  $\hat{\mathbf{R}}$  por  $\mathbf{R}$  e ao considerar que  $\mathbf{B}$  está próximo do valor ótimo obtém-se:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{P}_B^\perp \mathbf{R} \mathbf{P}_B^\perp$$

Mas  $\mathbf{B}^+ = (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H$  e  $\mathbf{P}_B^\perp = (\mathbf{I} - \mathbf{P}_B) = (\mathbf{I} - \mathbf{B} \mathbf{B}^+) = [\mathbf{I} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H]$ , dessa forma:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx [\mathbf{I} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H] \mathbf{R} [\mathbf{I} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H]$$

Realizando a multiplicação:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx [\mathbf{R} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{R}] [\mathbf{I} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H]$$

Mas  $\mathbf{R} = \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}$ , portanto:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx [\mathbf{R} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H (\mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I})] [\mathbf{I} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H]$$

Realizando a multiplicação e sabendo que  $c \mathbf{I} \mathbf{A} = c \mathbf{A}$  e  $c \mathbf{A} = \mathbf{A} c$  para  $c$  escalar:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx [\mathbf{R} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{A} \mathbf{C} \mathbf{A}^H - \sigma \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H] [\mathbf{I} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H]$$

Mas  $\mathbf{B}^H \mathbf{A} = \mathbf{0}$  e  $\mathbf{A}^H \mathbf{B} = \mathbf{0}$  pois  $\mathbf{B}$  abrange o espaço nulo de  $\mathbf{A}^H$ , então:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx [\mathbf{R} - \sigma \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H] [\mathbf{I} - \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H]$$

Realizando a multiplicação e sabendo que  $c \mathbf{I} \mathbf{A} = c \mathbf{A}$  e  $c \mathbf{A} = \mathbf{A} c$  para  $c$  escalar:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} - \mathbf{R} \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H - \sigma \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H + \sigma \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H$$

Entretanto  $\mathbf{A} \mathbf{A}^{-1} = \mathbf{A}^{-1} \mathbf{A} = \mathbf{I}$ , assim:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} - \mathbf{R} \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H - \sigma \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H + \sigma \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H$$

Realizando a subtração:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} - \mathbf{R} \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H$$

Mas  $\mathbf{R} = \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}$ , portanto:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} - (\mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}) \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H$$

Realizando a multiplicação e sabendo que  $c \mathbf{I} \mathbf{A} = c \mathbf{A}$  e  $c \mathbf{A} = \mathbf{A} c$  para  $c$  escalar:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} - \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H - \sigma \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H$$

Mas  $\mathbf{B}^H \mathbf{A} = \mathbf{0}$  e  $\mathbf{A}^H \mathbf{B} = \mathbf{0}$  pois  $\mathbf{B}$  abrange o espaço nulo de  $\mathbf{A}^H$ , então:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} - \sigma \mathbf{B} \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H$$

Sabendo que  $\mathbf{B}$  possui posto cheio, ao aproximar  $\hat{\mathbf{R}}$  por  $\mathbf{R}$  e ao considerar que  $\mathbf{B}$  está próximo do valor ótimo obtém-se:

$$\hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} \mathbf{P}_B^\perp$$

Mas  $\mathbf{B}^+ = \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H$  e  $\mathbf{P}_B^\perp = (\mathbf{I} - \mathbf{P}_B) = (\mathbf{I} - \mathbf{B} \mathbf{B}^+) = \left[ \mathbf{I} - \mathbf{B} \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H \right]$ , dessa forma:

$$\hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} \left[ \mathbf{I} - \mathbf{B} \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H \right]$$

Realizando a multiplicação:

$$\hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} - \mathbf{R} \mathbf{B} \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H$$

Mas  $\mathbf{R} = \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}$ , portanto:

$$\hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} - \left( \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I} \right) \mathbf{B} \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H$$

Realizando a multiplicação e sabendo que  $c \mathbf{I} \mathbf{A} = c \mathbf{A}$  e  $c \mathbf{A} = \mathbf{A} c$  para  $c$  escalar:

$$\hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} - \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{B} \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H - \sigma \mathbf{B} \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H$$

Mas  $\mathbf{B}^H \mathbf{A} = \mathbf{0}$  e  $\mathbf{A}^H \mathbf{B} = \mathbf{0}$  pois  $\mathbf{B}$  abrange o espaço nulo de  $\mathbf{A}^H$ , então:

$$\hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{R} - \sigma \mathbf{B} \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H$$

Finalmente, pode-se considerar que:

$$\mathbf{P}_B^\perp \hat{\mathbf{R}} \mathbf{P}_B^\perp \approx \mathbf{P}_B^\perp \mathbf{R} \mathbf{P}_B^\perp = \mathbf{R} \mathbf{P}_B^\perp$$

### A.14 Prova 7: $\mathbf{B}^+ \hat{\mathbf{R}} \left( \mathbf{B}^+ \right)^H \approx \mathbf{B}^+ \mathbf{R} \left( \mathbf{B}^+ \right)^H = \sigma \left( \mathbf{B}^H \mathbf{B} \right)^{-1}$

Sabendo que  $\mathbf{B}$  possui posto cheio, ao aproximar  $\hat{\mathbf{R}}$  por  $\mathbf{R}$  e ao considerar que  $\mathbf{B}$  está próximo do valor ótimo obtém-se:

$$\mathbf{B}^+ \hat{\mathbf{R}} \left( \mathbf{B}^+ \right)^H \approx \mathbf{B}^+ \mathbf{R} \left( \mathbf{B}^+ \right)^H$$

Mas  $\mathbf{R} = \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}$ ,  $\mathbf{B}^+ = \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H$  e  $\left( \mathbf{B}^+ \right)^H = \mathbf{B} \left( \mathbf{B}^H \mathbf{B} \right)^{-1}$ , então:

$$\mathbf{B}^+ \hat{\mathbf{R}} \left( \mathbf{B}^+ \right)^H \approx \left( \mathbf{B}^H \mathbf{B} \right)^{-1} \mathbf{B}^H \left( \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I} \right) \mathbf{B} \left( \mathbf{B}^H \mathbf{B} \right)^{-1}$$

Realizando as multiplicações e sabendo que  $c\mathbf{IA} = c\mathbf{A}$  e  $c\mathbf{A} = \mathbf{A}c$  para  $c$  escalar:

$$\mathbf{B}^+ \hat{\mathbf{R}} (\mathbf{B}^+)^{\mathbf{H}} \approx (\mathbf{B}^{\mathbf{H}} \mathbf{B})^{-1} \mathbf{B}^{\mathbf{H}} \mathbf{A} \mathbf{C} \mathbf{A}^{\mathbf{H}} \mathbf{B} (\mathbf{B}^{\mathbf{H}} \mathbf{B})^{-1} + \sigma (\mathbf{B}^{\mathbf{H}} \mathbf{B})^{-1} \mathbf{B}^{\mathbf{H}} \mathbf{B} (\mathbf{B}^{\mathbf{H}} \mathbf{B})^{-1}$$

Mas  $\mathbf{B}^{\mathbf{H}} \mathbf{A} = \mathbf{0}$  e  $\mathbf{A}^{\mathbf{H}} \mathbf{B} = \mathbf{0}$  pois  $\mathbf{B}$  abrange o espaço nulo de  $\mathbf{A}^{\mathbf{H}}$ :

$$\mathbf{B}^+ \hat{\mathbf{R}} (\mathbf{B}^+)^{\mathbf{H}} \approx \sigma (\mathbf{B}^{\mathbf{H}} \mathbf{B})^{-1} \mathbf{B}^{\mathbf{H}} \mathbf{B} (\mathbf{B}^{\mathbf{H}} \mathbf{B})^{-1}$$

Mas  $\mathbf{A} \mathbf{A}^{-1} = \mathbf{A}^{-1} \mathbf{A} = \mathbf{I}$  e  $\mathbf{A} \mathbf{I} = \mathbf{I} \mathbf{A} = \mathbf{A}$ , portanto:

$$\boxed{\mathbf{B}^+ \hat{\mathbf{R}} (\mathbf{B}^+)^{\mathbf{H}} \approx \sigma (\mathbf{B}^{\mathbf{H}} \mathbf{B})^{-1}}$$

# B DERIVAÇÃO DO MÉTODO DE NEWTON PARA O DMLE COM FILTRAGEM ESPACIAL USANDO A MATRIZ DE PREDIÇÃO LINEAR

## B.1 Introdução

A função-objetivo do Critério Determinístico do Estimador de Máxima Verossimilhança (DMLE) com filtragem espacial definida em (KRUMMENAUER et al., 2010) usa a matriz de direção filtrada  $\tilde{\mathbf{A}}$ . Entretanto, é realizada na seção B.9 a reparametrização dessa função em termos da matriz de predição linear filtrada  $\tilde{\mathbf{B}}$ . Assim, após esse processo, a função-objetivo é:

$$F_{\text{DMLEF2}} = \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)$$

em que:  $\mathbf{P}_{\tilde{\mathbf{B}}} = \mathbf{I} - \mathbf{U}\mathbf{U}^+$  e  $\mathbf{U} \triangleq \mathbf{H}\mathbf{P}_{\tilde{\mathbf{B}}}^\perp$ . Assim como apresentado no Apêndice A, são utilizados os vetores de parâmetros  $\hat{\phi}_0$  e  $\hat{\phi}_1$  com dimensões  $[2(M+1) \times 1]$ . Assim, a hessiana e o gradiente possuem dimensões  $[2(M+1) \times 2(M+1)]$  e  $[2(M+1) \times 1]$ , respectivamente.

## B.2 Derivada primeira de $\mathbf{U}$

Definindo  $\mathbf{U} \triangleq \mathbf{H}\mathbf{P}_{\tilde{\mathbf{B}}}^\perp$ , assim:

$$\frac{\partial \mathbf{U}}{\partial \phi_j} = \frac{\partial \mathbf{H}\mathbf{P}_{\tilde{\mathbf{B}}}^\perp}{\partial \phi_j}$$

Aplicando a regra da derivada do produto:

$$\frac{\partial \mathbf{U}}{\partial \phi_j} = \frac{\partial \mathbf{H}}{\partial \phi_j} \mathbf{P}_{\tilde{\mathbf{B}}}^\perp + \mathbf{H} \frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}^\perp}{\partial \phi_j}$$

Mas  $\mathbf{H}$  é constante, assim:

$$\frac{\partial \mathbf{U}}{\partial \phi_j} = \mathbf{H} \frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}^\perp}{\partial \phi_j}$$

No entanto,  $\frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}^\perp}{\partial \phi_j} = -\frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j}$ , assim:

$$\frac{\partial \mathbf{U}}{\partial \phi_j} = -\mathbf{H} \frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j}$$

Mas  $\frac{\partial \mathbf{P}_B}{\partial \phi_j} = (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp + \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+$ , então:

$$\frac{\partial \mathbf{U}}{\partial \phi_j} = -\mathbf{H} \left[ (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp + \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \right]$$

Realizando a multiplicação:

$$\frac{\partial \mathbf{U}}{\partial \phi_j} = -\mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp - \mathbf{H} \mathbf{P}_B^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+$$

Mas  $\mathbf{U} = \mathbf{H} \mathbf{P}_B^\perp$ , assim:

$$\boxed{\frac{\partial \mathbf{U}}{\partial \phi_j} = -\mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+}$$

### B.3 Derivada segunda de U

Derivando a primeira derivada de U produz:

$$\frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} = \frac{\partial}{\partial \phi_k} \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \right)$$

Mas  $\frac{\partial \mathbf{U}}{\partial \phi_j} = -\mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+$ , então:

$$\frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} = \frac{\partial}{\partial \phi_k} \left[ -\mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \right]$$

Aplicando a regra da derivada do produto, sabendo que  $\frac{\partial \mathbf{X} + \mathbf{Y}}{\partial \mathbf{x}} = \frac{\partial \mathbf{X}}{\partial \mathbf{x}} + \frac{\partial \mathbf{Y}}{\partial \mathbf{x}}$ :

$$\begin{aligned} \frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} = & -\frac{\partial \mathbf{H}}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp - \mathbf{H} \frac{\partial (\mathbf{B}^+)^H}{\partial \phi_k} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp - \mathbf{H} (\mathbf{B}^+)^H \frac{\partial}{\partial \phi_k} \left( \frac{\partial \mathbf{B}^H}{\partial \phi_j} \right) \mathbf{P}_B^\perp - \\ & - \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} - \frac{\partial \mathbf{U}}{\partial \phi_k} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ - \mathbf{U} \frac{\partial}{\partial \phi_k} \left( \frac{\partial \mathbf{B}}{\partial \phi_j} \right) \mathbf{B}^+ - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \frac{\partial \mathbf{B}^+}{\partial \phi_k} \end{aligned}$$

Mas  $\mathbf{H}$  é constante, assim:

$$\begin{aligned} \frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} = & -\mathbf{H} \frac{\partial (\mathbf{B}^+)^H}{\partial \phi_k} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp - \mathbf{H} (\mathbf{B}^+)^H \frac{\partial}{\partial \phi_k} \left( \frac{\partial \mathbf{B}^H}{\partial \phi_j} \right) \mathbf{P}_B^\perp - \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_k} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ - \mathbf{U} \frac{\partial}{\partial \phi_k} \left( \frac{\partial \mathbf{B}}{\partial \phi_j} \right) \mathbf{B}^+ - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \frac{\partial \mathbf{B}^+}{\partial \phi_k} \end{aligned}$$

Ainda,  $\frac{\partial \mathbf{B}}{\partial \phi_j}$  e  $\frac{\partial \mathbf{B}^H}{\partial \phi_j}$  são constantes, dessa forma:

$$\frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} = -\mathbf{H} \frac{\partial (\mathbf{B}^+)^H}{\partial \phi_k} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp - \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} - \frac{\partial \mathbf{U}}{\partial \phi_k} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \frac{\partial \mathbf{B}^+}{\partial \phi_k}$$

Mas  $\frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} = (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp - \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+$ , então:

$$\begin{aligned} \frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} = & -\mathbf{H} \left[ (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp - \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \right]^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_B^\perp - \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \frac{\partial \mathbf{P}_B^\perp}{\partial \phi_k} - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_k} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \left[ (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_B^\perp - \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \right] \end{aligned}$$





Realizando a multiplicação:

$$\begin{aligned} \frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} = & -\mathbf{H} \mathbf{P}_{\mathbf{B}}^{\perp} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^{\perp} + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^{\perp} + \\ & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^{\perp} + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^{\perp} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ + \\ & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^{\perp} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ - \\ & - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^{\perp} + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \end{aligned}$$

Entretanto,  $\mathbf{U} = \mathbf{H} \mathbf{P}_{\mathbf{B}}^{\perp}$ , assim:

$$\begin{aligned} \frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} = & -\mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^{\perp} + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^{\perp} + \\ & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^{\perp} + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^{\perp} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ + \\ & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^{\perp} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ - \\ & - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^{\perp} + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \end{aligned}$$

Reordenando os termos:

$$\begin{aligned} \frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} = & \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^{\perp} - \\ & - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^{\perp} + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^{\perp} + \\ & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^{\perp} + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^{\perp} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \\ & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^{\perp} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \end{aligned}$$

## B.4 Derivada de $\mathbf{U}^+$

A derivada da pseudo-inversa é (GOLUB; PEREYRA, 1973):

$$\frac{\partial \mathbf{U}^+}{\partial \phi_j} = -\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} (\mathbf{I} - \mathbf{U} \mathbf{U}^+) + (\mathbf{I} - \mathbf{U}^+ \mathbf{U}) \frac{\partial \mathbf{U}^H}{\partial \phi_j} (\mathbf{U}^+)^H \mathbf{U}^+$$

Mas  $\mathbf{P}_{\tilde{\mathbf{B}}} = \mathbf{I} - \mathbf{U} \mathbf{U}^+$ , dessa forma:

$$\frac{\partial \mathbf{U}^+}{\partial \phi_j} = -\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} \mathbf{P}_{\tilde{\mathbf{B}}} + (\mathbf{I} - \mathbf{U}^+ \mathbf{U}) \frac{\partial \mathbf{U}^H}{\partial \phi_j} (\mathbf{U}^+)^H \mathbf{U}^+$$

Realizando a multiplicação e considerando que  $\mathbf{U}^+ \mathbf{U} \neq \mathbf{I}$ , pois  $\mathbf{U}$  possui número de condição infinito. Assim, o último termo não pode ser anulado:

$$\frac{\partial \mathbf{U}^+}{\partial \phi_j} = -\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} \mathbf{P}_{\tilde{\mathbf{B}}} + \frac{\partial \mathbf{U}^H}{\partial \phi_j} (\mathbf{U}^+)^H \mathbf{U}^+ - \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_j} (\mathbf{U}^+)^H \mathbf{U}^+$$

## B.5 Derivada de $\mathbf{P}_{\tilde{\mathbf{B}}}$

Calculando  $\frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial b_i}$  e considerando que  $\mathbf{P}_{\tilde{\mathbf{B}}} = \mathbf{I} - \mathbf{H}\mathbf{P}_{\mathbf{B}}^{\perp} (\mathbf{H}\mathbf{P}_{\mathbf{B}}^{\perp})^+ = \mathbf{I} - \mathbf{U}\mathbf{U}^+$  e sabendo que  $\frac{\partial \mathbf{X} + \mathbf{Y}}{\partial \mathbf{x}} = \frac{\partial \mathbf{X}}{\partial \mathbf{x}} + \frac{\partial \mathbf{Y}}{\partial \mathbf{x}}$ :

$$\begin{aligned} \frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} &= \frac{\partial}{\partial \phi_j} \{ \mathbf{I} - \mathbf{U}\mathbf{U}^+ \} \\ &= \frac{\partial \mathbf{I}}{\partial \phi_j} - \frac{\partial \mathbf{U}\mathbf{U}^+}{\partial \phi_j} \end{aligned}$$

Mas  $\mathbf{I}$  é constante, assim:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} = - \frac{\partial \mathbf{U}\mathbf{U}^+}{\partial \phi_j}$$

Aplicando a regra da derivada do produto:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} = - \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \mathbf{U} \frac{\partial \mathbf{U}^+}{\partial \phi_j} \right)$$

Mas  $\frac{\partial \mathbf{U}^+}{\partial \phi_j} = -\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \mathbf{U}^+ (\mathbf{U}^+)^{\mathbf{H}} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} - \mathbf{U}^+ (\mathbf{U}^+)^{\mathbf{H}} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} \mathbf{U}\mathbf{U}^+ + \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} (\mathbf{U}^+)^{\mathbf{H}} \mathbf{U}^+ - \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} (\mathbf{U}^+)^{\mathbf{H}} \mathbf{U}^+$ , então:

$$\begin{aligned} \frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} &= - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - \mathbf{U} \left[ -\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \mathbf{U}^+ (\mathbf{U}^+)^{\mathbf{H}} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} - \mathbf{U}^+ (\mathbf{U}^+)^{\mathbf{H}} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} \mathbf{U}\mathbf{U}^+ + \right. \\ &\quad \left. + \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} (\mathbf{U}^+)^{\mathbf{H}} \mathbf{U}^+ - \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} (\mathbf{U}^+)^{\mathbf{H}} \mathbf{U}^+ \right] \end{aligned}$$

Realizando a multiplicação:

$$\begin{aligned} \frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} &= - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \mathbf{U}\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - \mathbf{U}\mathbf{U}^+ (\mathbf{U}^+)^{\mathbf{H}} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} + \mathbf{U}\mathbf{U}^+ (\mathbf{U}^+)^{\mathbf{H}} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} \mathbf{U}\mathbf{U}^+ - \\ &\quad - \mathbf{U} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} (\mathbf{U}^+)^{\mathbf{H}} \mathbf{U}^+ + \mathbf{U}\mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} (\mathbf{U}^+)^{\mathbf{H}} \mathbf{U}^+ \end{aligned}$$

Mas, considerando a seção B.10,  $\mathbf{U}\mathbf{U}^+ (\mathbf{U}^+)^{\mathbf{H}} = (\mathbf{U}^+)^{\mathbf{H}}$ , assim:

$$\begin{aligned} \frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} &= - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \mathbf{U}\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - (\mathbf{U}^+)^{\mathbf{H}} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} + (\mathbf{U}^+)^{\mathbf{H}} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} \mathbf{U}\mathbf{U}^+ - \\ &\quad - \mathbf{U} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} (\mathbf{U}^+)^{\mathbf{H}} \mathbf{U}^+ + \mathbf{U}\mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} (\mathbf{U}^+)^{\mathbf{H}} \mathbf{U}^+ \end{aligned}$$

Ainda, considerando a seção B.11,  $\mathbf{U}\mathbf{U}^+ \mathbf{U} = \mathbf{U}$ , então:

$$\begin{aligned} \frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} &= - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \mathbf{U}\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - (\mathbf{U}^+)^{\mathbf{H}} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} + (\mathbf{U}^+)^{\mathbf{H}} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} \mathbf{U}\mathbf{U}^+ - \\ &\quad - \mathbf{U} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} (\mathbf{U}^+)^{\mathbf{H}} \mathbf{U}^+ + \mathbf{U} \frac{\partial \mathbf{U}^{\mathbf{H}}}{\partial \phi_j} (\mathbf{U}^+)^{\mathbf{H}} \mathbf{U}^+ \end{aligned}$$

Realizando a subtração:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} = -\frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} + (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} \mathbf{U} \mathbf{U}^+$$

Reordenando os termos:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} = -\frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} \mathbf{U} \mathbf{U}^+$$

Mas  $\mathbf{U} \mathbf{U}^+ = (\mathbf{U} \mathbf{U}^+)^H$ , portanto:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} = -\frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} (\mathbf{U} \mathbf{U}^+)^H$$

Aplicando  $(\mathbf{A}\mathbf{B})^H = \mathbf{B}^H \mathbf{A}^H$

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} = -\frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \left( \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H$$

Colocando um termo em evidência:

$$\boxed{\frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} = - \left[ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - \left( \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H \right]}$$

## B.6 Gradiente de $\text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)$

Derivando a função-objetivo produz:

$$\nabla_{\phi_j} \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z) = \frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j}$$

Entretanto,  $\frac{\partial \text{tr} \mathbf{X}}{\partial x} = \text{tr} \left( \frac{\partial \mathbf{X}}{\partial x} \right)$ , então:

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = \text{tr} \left( \frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z}{\partial \phi_j} \right)$$

Aplicando a regra da derivada do produto:

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = \text{tr} \left( \frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} \hat{\mathbf{R}}_Z + \mathbf{P}_{\tilde{\mathbf{B}}} \frac{\partial \hat{\mathbf{R}}_Z}{\partial \phi_j} \right)$$

Mas  $\hat{\mathbf{R}}_Z$  é constante, assim:

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = \text{tr} \left( \frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} \hat{\mathbf{R}}_Z \right)$$

No entanto,  $\frac{\partial \mathbf{P}_{\tilde{\mathbf{B}}}}{\partial \phi_j} = - \left[ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - \left( \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H \right]$ , dessa forma:

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = \text{tr} \left\{ - \left[ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - \left( \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H \right] \hat{\mathbf{R}}_Z \right\}$$

Entretanto,  $\text{tr}(-\mathbf{A}) = -\text{tr}(\mathbf{A})$ , então:

$$\frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = -\text{tr} \left\{ \left[ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ + \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - \left( \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H \right] \hat{\mathbf{R}}_Z \right\}$$

Realizando as multiplicações:

$$\begin{aligned} \frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = & -\text{tr} \left[ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \right. \\ & \left. - \left( \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H \hat{\mathbf{R}}_Z \right] \end{aligned}$$

No entanto,  $\hat{\mathbf{R}}_Z = \hat{\mathbf{R}}_Z^H$ , assim:

$$\begin{aligned} \frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = & -\text{tr} \left[ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H \hat{\mathbf{R}}_Z^H - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \right. \\ & \left. - \left( \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \right)^H \hat{\mathbf{R}}_Z^H \right] \end{aligned}$$

Aplicando  $(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H$ :

$$\begin{aligned} \frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = & -\text{tr} \left[ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} \hat{\mathbf{R}}_Z^H - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \right. \\ & \left. - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} (\mathbf{U} \mathbf{U}^+)^H \hat{\mathbf{R}}_Z^H \right] \end{aligned}$$

Aplicando  $\text{tr}(\mathbf{ABU}) = \text{tr}(\mathbf{BUA}) = \text{tr}(\mathbf{UAB})$ :

$$\begin{aligned} \frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = & -\text{tr} \left[ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \hat{\mathbf{R}}_Z^H (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \right. \\ & \left. - \hat{\mathbf{R}}_Z^H (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_j} (\mathbf{U} \mathbf{U}^+)^H \right] \end{aligned}$$

Aplicando  $(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H$ :

$$\begin{aligned} \frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = & -\text{tr} \left[ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right)^H - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \right. \\ & \left. - \left( \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right)^H \right] \end{aligned}$$

Aplicando  $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$ :

$$\begin{aligned} \frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = & -\text{tr} \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right) - \text{tr} \left[ \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right)^H \right] + \\ & + \text{tr} \left( \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right) + \text{tr} \left[ \left( \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right)^H \right] \end{aligned}$$

Mas  $\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}^T)$  e  $\mathbf{A}^H = \mathbf{A}^{T*}$ , então  $\text{tr}(\mathbf{A}^H) = \text{tr}[(\mathbf{A}^T)^*] = \text{tr}(\mathbf{A}^*) = \text{tr}(\mathbf{A})^*$ . Assim:

$$\begin{aligned} \frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} &= -\text{tr}\left(\frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right) - \text{tr}\left(\frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right)^* + \text{tr}\left(\mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right) + \\ &\quad + \text{tr}\left(\mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right)^* \end{aligned}$$

Mas  $z + z^* = 2\Re z$ , assim:

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = -2\Re \text{tr}\left(\frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right) + 2\Re \text{tr}\left(\mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right)$$

Colocando um termo em evidência:

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = -2\Re \left[ \text{tr}\left(\frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right) - \text{tr}\left(\mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right) \right]$$

Aplicando  $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$ :

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = -2\Re \text{tr}\left(\frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right)$$

Colocando um termo em evidência:

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = -2\Re \text{tr}\left[\left(\mathbf{I} - \mathbf{U} \mathbf{U}^+\right) \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right]$$

Mas  $\mathbf{P}_{\tilde{\mathbf{B}}} = \mathbf{I} - \mathbf{U} \mathbf{U}^+$ , então:

$$\boxed{\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = -2\Re \text{tr}\left(\mathbf{P}_{\tilde{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right)}$$

## B.7 Hessiana de $\text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)$

Derivando o gradiente da função-objetivo produz:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} &= \frac{\partial \nabla_{\phi_j} \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k} \\ &= \frac{\partial}{\partial \phi_k} \left[ \nabla_{\phi_j} \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z) \right] \\ &= \frac{\partial}{\partial \phi_k} \left[ \frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} \right] \end{aligned}$$

Mas  $\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_j} = -2\Re \text{tr}\left(\mathbf{P}_{\tilde{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right) = -2\Re \text{tr}\left[\left(\mathbf{I} - \mathbf{U} \mathbf{U}^+\right) \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right]$ , conseqüentemente:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = \frac{\partial}{\partial \phi_k} \left\{ -2\Re \text{tr}\left[\left(\mathbf{I} - \mathbf{U} \mathbf{U}^+\right) \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z\right] \right\}$$

Aplicando  $\frac{\partial c\mathbf{X}}{\partial x} = c\frac{\partial \mathbf{X}}{\partial x}$  para  $c$  constante:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}}\hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = -2 \frac{\partial}{\partial \phi_k} \left\{ \Re \text{tr} \left[ (\mathbf{I} - \mathbf{U}\mathbf{U}^+) \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right] \right\}$$

Aplicando  $\frac{\partial \Re \mathbf{X}}{\partial x} = \Re \frac{\partial \mathbf{X}}{\partial x}$ :

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}}\hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = -2 \Re \frac{\partial}{\partial \phi_k} \left\{ \text{tr} \left[ (\mathbf{I} - \mathbf{U}\mathbf{U}^+) \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right] \right\}$$

Aplicando  $\frac{\partial \text{tr}(\mathbf{X})}{\partial x} = \text{tr} \left( \frac{\partial \mathbf{X}}{\partial x} \right)$ :

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}}\hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = -2 \Re \text{tr} \left\{ \frac{\partial}{\partial \phi_k} \left[ (\mathbf{I} - \mathbf{U}\mathbf{U}^+) \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right] \right\}$$

Realizando a multiplicação:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}}\hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = -2 \Re \text{tr} \left[ \frac{\partial}{\partial \phi_k} \left( \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U}\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right) \right]$$

Aplicando a regra da derivada do produto, sabendo que  $\frac{\partial \mathbf{X} + \mathbf{Y}}{\partial \mathbf{x}} = \frac{\partial \mathbf{X}}{\partial \mathbf{x}} + \frac{\partial \mathbf{Y}}{\partial \mathbf{x}}$ :

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}}\hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left( \frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \hat{\mathbf{R}}_Z}{\partial \phi_k} - \right. \\ & - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U}\mathbf{U}^+ \frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & \left. - \mathbf{U}\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z - \mathbf{U}\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \hat{\mathbf{R}}_Z}{\partial \phi_k} \right) \end{aligned}$$

Mas  $\hat{\mathbf{R}}_Z$  é constante, assim:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}}\hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left( \frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \right. \\ & \left. - \mathbf{U} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U}\mathbf{U}^+ \frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U}\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z \right) \end{aligned}$$

Mas:

$$\begin{aligned} \frac{\partial^2 \mathbf{U}}{\partial \phi_k \partial \phi_j} = & \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp - \\ & - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp + \\ & - \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \\ & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \end{aligned}$$

Assim:

$$\begin{aligned}
 \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \left[ \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ - \right. \right. \\
 & - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp + \\
 & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp + \\
 & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \left. \right] \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\
 & - \mathbf{U} \mathbf{U}^+ \left[ \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ - \right. \\
 & - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp + \\
 & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp + \\
 & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \left. \right] \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\
 & \left. - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z \right\}
 \end{aligned}$$



Realizando as multiplicações:

$$\begin{aligned}
 \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left[ \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z - \right. \\
 & - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \frac{\partial \mathbf{U}}{\partial \phi_j} \hat{\mathbf{R}}_Z - \\
 & - \mathbf{U} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \mathbf{U} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \mathbf{U} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\
 & - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\
 & - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\
 & - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\
 & \left. - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z \right]
 \end{aligned}$$

Aplicando a propriedade  $\mathbf{A} \mathbf{A}^+ \mathbf{A} = \mathbf{A}$  da pseudo-inversa de Moore-Penrose ([PETERSEN](#);

PEDERSEN, 2012):

$$\begin{aligned}
 \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left[ \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z - \right. \\
 & - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z - \\
 & - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\
 & - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\
 & + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\
 & - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\
 & - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\
 & - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_j} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\
 & \left. - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \frac{\partial \mathbf{B}}{\partial \phi_k} \mathbf{B}^+ \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z \right]
 \end{aligned}$$

Mas, após realizado um teste numérico,  $\mathbf{B}^+ \mathbf{U}^+ = \mathbf{0}$ , assim:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left[ -\mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \right. \\ & -\mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & \left. - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z \right] \end{aligned}$$

Reordenando os termos:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left[ \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \right. \\ & -\mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_k} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & -\mathbf{U} \frac{\partial \mathbf{B}}{\partial \phi_j} (\mathbf{B}^H \mathbf{B})^{-1} \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \mathbf{U} \mathbf{U}^+ \mathbf{H} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z - \\ & \left. - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z \right] \end{aligned}$$

Realizando as subtrações:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \mathbf{U} \mathbf{U}^+ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \mathbf{U} \mathbf{U}^+ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & \left. - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z \right] \end{aligned}$$

Colocando um termo em evidência:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ (\mathbf{I} - \mathbf{U} \mathbf{U}^+) \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right] + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z - \\ & \left. - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z \right\} \end{aligned}$$

No entanto,  $\mathbf{P}_{\hat{\mathbf{B}}} = \mathbf{I} - \mathbf{U} \mathbf{U}^+$ , assim:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \mathbf{P}_{\hat{\mathbf{B}}} \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right] + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z - \\ & \left. - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^+}{\partial \phi_k} \hat{\mathbf{R}}_Z \right\} \end{aligned}$$

Porém,  $\frac{\partial \mathbf{U}^+}{\partial \phi_k} = -\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ + \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} + \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ - \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+$ ,

então:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \mathbf{P}_{\hat{\mathbf{B}}} \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right] + \frac{\partial \mathbf{U}}{\partial \phi_j} \left[ -\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ + \right. \\ & + \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} + \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ - \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \left. \right] \hat{\mathbf{R}}_Z - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \left[ -\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ + \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} + \right. \\ & + \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ - \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \left. \right] \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \left[ -\mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ + \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} + \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ - \right. \\ & \left. \left. - \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \right] \hat{\mathbf{R}}_Z \right\} \end{aligned}$$

Realizando as multiplicações:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \mathbf{P}_{\hat{\mathbf{B}}} \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right] - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \mathbf{U} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z - \\ & \left. - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z \right\} \end{aligned}$$

Aplicando a propriedade  $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$  da pseudo-inversa de Moore-Penrose ([PETERSEN](#);



Realizando a subtração:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \mathbf{P}_{\hat{\mathbf{B}}} \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right] - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z - \\ & - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right\} \end{aligned}$$

Reordenando os termos:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \mathbf{P}_{\hat{\mathbf{B}}} \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right] - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z - \\ & - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \mathbf{U} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right\} \end{aligned}$$





Colocando um termo em evidência:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \mathbf{P}_{\hat{\mathbf{B}}} \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right] - \\ & - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right\} \end{aligned}$$

Porém, considerando a seção B.12,  $\mathbf{P}_{\hat{\mathbf{B}}}^\perp \mathbf{U}^+ = \mathbf{U}^+$ , então:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \mathbf{P}_{\hat{\mathbf{B}}} \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right] - \\ & - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right\} \end{aligned}$$

Mas  $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$ , assim:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \mathbf{P}_{\hat{\mathbf{B}}} \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z \left. \right] \left. \right\} - \\ & - 2 \Re \text{tr} \left[ - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\hat{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right] \end{aligned}$$

Contudo,  $\text{tr}(\mathbf{ABU}) = \text{tr}(\mathbf{BUA}) = \text{tr}(\mathbf{UAB})$ , então:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}}\hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \hat{\mathbf{R}}_Z - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z + \\ & + \left. \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \hat{\mathbf{R}}_Z \right] \mathbf{P}_{\tilde{\mathbf{B}}} \Big\} - \\ & - 2 \Re \text{tr} \left[ - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\tilde{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right] \end{aligned}$$

Colocando um termo em evidência:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}}\hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{U}^+ + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{U}^+ - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ + \\ & + \left. \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\tilde{\mathbf{B}}} + \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \right] \hat{\mathbf{R}}_Z \mathbf{P}_{\tilde{\mathbf{B}}} \Big\} - \\ & - 2 \Re \text{tr} \left[ - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\tilde{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right] \end{aligned}$$

## B.8 Aproximação assintótica e numérica na hessiana de $\text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}}\hat{\mathbf{R}}_Z)$

A hessiana é:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}}\hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = & -2 \Re \text{tr} \left\{ \left[ \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} \mathbf{U}^+ + \right. \right. \\ & + \mathbf{H}(\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_j} (\mathbf{B}^+)^H \frac{\partial \mathbf{B}^H}{\partial \phi_k} \mathbf{U}^+ - \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ - \\ & - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \frac{\partial \mathbf{U}}{\partial \phi_k} \mathbf{U}^+ - \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \mathbf{U} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ + \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\tilde{\mathbf{B}}} + \\ & + \left. \frac{\partial \mathbf{U}}{\partial \phi_j} \frac{\partial \mathbf{U}^H}{\partial \phi_k} (\mathbf{U}^+)^H \mathbf{U}^+ \right] \hat{\mathbf{R}}_Z \mathbf{P}_{\tilde{\mathbf{B}}} \Big\} - 2 \Re \text{tr} \left[ - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\tilde{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right] \end{aligned}$$

Mas, assintoticamente,  $\hat{\mathbf{R}}_Z \mathbf{P}_{\tilde{\mathbf{B}}} \approx \mathbf{R}_Z \mathbf{P}_{\tilde{\mathbf{B}}}$ . Ainda, considerando que  $\mathbf{P}_{\tilde{\mathbf{B}}}$  está próximo do valor ótimo,  $\mathbf{R}_Z \mathbf{P}_{\tilde{\mathbf{B}}} = \mathbf{0}$ , verificado em um teste numérico, assim:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}}\hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = -2 \Re \text{tr} \left[ - (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\tilde{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right]$$

Colocando um termo em evidência:

$$\boxed{\frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{B}}} \hat{\mathbf{R}}_Z)}{\partial \phi_k \partial \phi_j} = 2 \Re \text{tr} \left[ (\mathbf{U}^+)^H \frac{\partial \mathbf{U}^H}{\partial \phi_k} \mathbf{P}_{\tilde{\mathbf{B}}} \frac{\partial \mathbf{U}}{\partial \phi_j} \mathbf{U}^+ \hat{\mathbf{R}}_Z \right]}$$

## B.9 Prova 8: $\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \mathbf{U}\mathbf{U}^+ \triangleq \mathbf{P}_{\tilde{\mathbf{B}}}$

Iniciando o desenvolvimento:

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \tilde{\mathbf{A}} \left( \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} \right)^{-1} \tilde{\mathbf{A}}^H$$

Entretanto, considerando a seção D.11,  $\tilde{\mathbf{A}}^+ \left( \tilde{\mathbf{A}}^+ \right)^H = \left( \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} \right)^{-1}$ , então:

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \tilde{\mathbf{A}} \tilde{\mathbf{A}}^+ \left( \tilde{\mathbf{A}}^+ \right)^H \tilde{\mathbf{A}}^H$$

Aplicando  $(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H$ :

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \tilde{\mathbf{A}} \tilde{\mathbf{A}}^+ \left( \tilde{\mathbf{A}} \tilde{\mathbf{A}}^+ \right)^H$$

Mas  $\tilde{\mathbf{A}} = \mathbf{H}\mathbf{A}$ , assim:

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \tilde{\mathbf{A}} (\mathbf{H}\mathbf{A})^+ \left[ \tilde{\mathbf{A}} (\mathbf{H}\mathbf{A})^+ \right]^H$$

Aplicando a propriedade  $(\mathbf{AB})^+ = (\mathbf{A}^+ \mathbf{AB})^+ (\mathbf{ABB}^+)^+$  da pseudo-inversa de Moore-Penrose (PETERSEN; PEDERSEN, 2012):

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \tilde{\mathbf{A}} \left( \mathbf{H}^+ \mathbf{H} \mathbf{A} \right)^+ \left( \mathbf{H} \mathbf{A} \mathbf{A}^+ \right)^+ \left[ \tilde{\mathbf{A}} \left( \mathbf{H}^+ \mathbf{H} \mathbf{A} \right)^+ \left( \mathbf{H} \mathbf{A} \mathbf{A}^+ \right)^+ \right]^H$$

No entanto, da mesma forma que o mostrado para  $\tilde{\mathbf{A}}$  na seção D.9,  $\mathbf{H}^+ \mathbf{H} = \mathbf{I}$ , dessa forma:

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \tilde{\mathbf{A}} \mathbf{A}^+ \left( \mathbf{H} \mathbf{A} \mathbf{A}^+ \right)^+ \left[ \tilde{\mathbf{A}} \mathbf{A}^+ \left( \mathbf{H} \mathbf{A} \mathbf{A}^+ \right)^+ \right]^H$$

Mas  $\tilde{\mathbf{A}} = \mathbf{H}\mathbf{A}$ , assim:

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \mathbf{H} \mathbf{A} \mathbf{A}^+ \left( \mathbf{H} \mathbf{A} \mathbf{A}^+ \right)^+ \left[ \mathbf{H} \mathbf{A} \mathbf{A}^+ \left( \mathbf{H} \mathbf{A} \mathbf{A}^+ \right)^+ \right]^H$$

Ainda,  $\mathbf{P}_{\mathbf{A}} = \mathbf{A} \mathbf{A}^+$ , então:

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \mathbf{H} \mathbf{P}_{\mathbf{A}} (\mathbf{H} \mathbf{P}_{\mathbf{A}})^+ \left[ \mathbf{H} \mathbf{P}_{\mathbf{A}} (\mathbf{H} \mathbf{P}_{\mathbf{A}})^+ \right]^H$$

Definindo  $\mathbf{E} \triangleq \mathbf{H} \mathbf{P}_{\mathbf{A}}$ :

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \mathbf{E} \mathbf{E}^+ \left( \mathbf{E} \mathbf{E}^+ \right)^H$$

Da mesma forma que o apresentado na seção D.12,  $\mathbf{P}_{\tilde{\mathbf{A}}} = \mathbf{P}_{\tilde{\mathbf{A}}}^H = \mathbf{P}_{\tilde{\mathbf{A}}} \mathbf{P}_{\tilde{\mathbf{A}}} = \mathbf{P}_{\tilde{\mathbf{A}}} \mathbf{P}_{\tilde{\mathbf{A}}}^H = \tilde{\mathbf{A}} \tilde{\mathbf{A}}^+$ ,  $\mathbf{E} \mathbf{E}^+ = (\mathbf{E} \mathbf{E}^+)^H = \mathbf{E} \mathbf{E}^+ \mathbf{E} \mathbf{E}^+ = \mathbf{E} \mathbf{E}^+ (\mathbf{E} \mathbf{E}^+)^H$ , então:

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \mathbf{E} \mathbf{E}^+$$

Mas  $\mathbf{E} = \mathbf{H} \mathbf{P}_{\mathbf{A}}$ , portanto:

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \mathbf{H} \mathbf{P}_{\mathbf{A}} (\mathbf{H} \mathbf{P}_{\mathbf{A}})^+$$

Entretanto,  $\mathbf{P}_{\mathbf{A}} = \mathbf{P}_{\mathbf{B}}^\perp$ , assim:

$$\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \mathbf{H} \mathbf{P}_{\mathbf{B}}^\perp (\mathbf{H} \mathbf{P}_{\mathbf{B}}^\perp)^+$$

Definindo  $\mathbf{U} \triangleq \mathbf{H} \mathbf{P}_{\mathbf{B}}^\perp$ :

$$\boxed{\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = \mathbf{I} - \mathbf{U} \mathbf{U}^+ \triangleq \mathbf{P}_{\tilde{\mathbf{B}}}}$$

## B.10 Prova 9: $\mathbf{U} \mathbf{U}^+ (\mathbf{U}^+)^H = (\mathbf{U}^+)^H$

Iniciando o desenvolvimento, sabendo que  $\mathbf{U}$  não possui posto cheio:

$$\mathbf{U} \mathbf{U}^+ (\mathbf{U}^+)^H = \mathbf{U} \mathbf{U}^+ (\mathbf{U}^+)^H$$

Porém,  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ , então:

$$\mathbf{U} \mathbf{U}^+ (\mathbf{U}^+)^H = \mathbf{U} (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H \left[ (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H \right]^H$$

Aplicando  $\mathbf{B}^H \mathbf{A}^H = (\mathbf{A} \mathbf{B})^H$ , sabendo que  $(\mathbf{U}^H \mathbf{U})^{-1} = \left[ (\mathbf{U}^H \mathbf{U})^{-1} \right]^H$ :

$$\mathbf{U} \mathbf{U}^+ (\mathbf{U}^+)^H = \mathbf{U} (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H \mathbf{U} (\mathbf{U}^H \mathbf{U})^{-1}$$

Realizando a multiplicação numericamente instável:

$$\mathbf{U} \mathbf{U}^+ (\mathbf{U}^+)^H = \mathbf{U} (\mathbf{U}^H \mathbf{U})^{-1}$$

Aplicando  $\mathbf{B}^H \mathbf{A}^H = (\mathbf{A} \mathbf{B})^H$ , sabendo que  $(\mathbf{U}^H \mathbf{U})^{-1} = \left[ (\mathbf{U}^H \mathbf{U})^{-1} \right]^H$ :

$$\mathbf{U} \mathbf{U}^+ (\mathbf{U}^+)^H = \left[ (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H \right]^H$$

Mas  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ , assim:

$$\boxed{\mathbf{U} \mathbf{U}^+ (\mathbf{U}^+)^H = (\mathbf{U}^+)^H}$$

## B.11 Prova 10: $\mathbf{U}\mathbf{U}^+\mathbf{U} = \mathbf{U}$

Iniciando o desenvolvimento, sabendo que  $\mathbf{U}$  não possui posto cheio:

$$\mathbf{U}\mathbf{U}^+\mathbf{U} = \mathbf{U}\mathbf{U}^+\mathbf{U}$$

Mas  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ , assim:

$$\mathbf{U}\mathbf{U}^+\mathbf{U} = \mathbf{U} (\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H \mathbf{U}$$

Realizando a multiplicação numericamente instável:

$$\boxed{\mathbf{U}\mathbf{U}^+\mathbf{U} = \mathbf{U}}$$

## B.12 Prova 11: $\mathbf{P}_B^\perp \mathbf{U}^+ = \mathbf{U}^+$

Iniciando o desenvolvimento:

$$\mathbf{P}_B^\perp \mathbf{U}^+ = \mathbf{P}_B^\perp \mathbf{U}^+$$

Mas  $\mathbf{P}_B^\perp = \mathbf{I} - \mathbf{B}\mathbf{B}^+$ , assim:

$$\mathbf{P}_B^\perp \mathbf{U}^+ = (\mathbf{I} - \mathbf{B}\mathbf{B}^+) \mathbf{U}^+$$

Realizando a multiplicação:

$$\mathbf{P}_B^\perp \mathbf{U}^+ = \mathbf{U}^+ - \mathbf{B}\mathbf{B}^+ \mathbf{U}^+$$

Entretanto, após realizado um teste numérico,  $\mathbf{B}^+ \mathbf{U}^+ = \mathbf{0}$ , então:

$$\boxed{\mathbf{P}_B^\perp \mathbf{U}^+ = \mathbf{U}^+}$$

# C DERIVAÇÃO DO MÉTODO DE NEWTON PARA O DMLE SEM FILTRAGEM ESPACIAL USANDO A MATRIZ DE DIREÇÃO

## C.1 Introdução

A função-objetivo do Critério Determinístico do Estimador de Máxima Verossimilhança (DMLE) sem filtragem espacial e usando a matriz de direção  $\mathbf{A}$  é (KRIM; VIBERG, 1996):

$$F_{\text{DMLE}} = \text{tr} \left( \mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}} \right)$$

em que:  $\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{P}_{\mathbf{A}} = \mathbf{I} - \mathbf{A}\mathbf{A}^+$  e  $\mathbf{A}^+ = \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \mathbf{A}^H$  é a pseudo-inversa de Moore-Penrose.

Considerando o Arranjo Linear Uniforme,  $\mathbf{A} = \mathbf{A}(\boldsymbol{\omega})$  é a matriz de direção calculada em função dos parâmetros estimados:

$$\mathbf{A} = \begin{bmatrix} e^{j0\omega_1} & \dots & e^{j0\omega_m} \\ e^{j1\omega_1} & \dots & e^{j1\omega_m} \\ \vdots & \ddots & \vdots \\ e^{j(K-2)\omega_1} & \dots & e^{j(K-2)\omega_m} \\ e^{j(K-1)\omega_1} & \dots & e^{j(K-1)\omega_m} \end{bmatrix} = \begin{bmatrix} 1 & \dots & 1 \\ e^{j\omega_1} & \dots & e^{-j\omega_m} \\ \vdots & \ddots & \vdots \\ e^{j(K-2)\omega_1} & \dots & e^{j(K-2)\omega_m} \\ e^{j(K-1)\omega_1} & \dots & e^{j(K-1)\omega_m} \end{bmatrix} \in \mathbb{C}^{K \times M}$$

Finalmente,  $\boldsymbol{\omega} = [\omega_1 \dots \omega_M]^T \in \mathbb{R}^{M \times 1}$  é o vetor das estimativas das frequências angulares. Para se obter as estimativas em termos dos ângulos DOA, deve-se utilizar a relação  $\omega_m \triangleq -\frac{2\pi}{d} \sin(\theta_m)$  apresentada na seção 2.1. A partir de um vetor de parâmetros iniciais  $\hat{\boldsymbol{\omega}}_0 \in \mathbb{R}^{M \times 1}$ ,  $\hat{\boldsymbol{\omega}}_1 \in \mathbb{R}^{M \times 1}$  é calculado iterativamente utilizando a seguinte fórmula de correção do Método de Newton:

$$\hat{\boldsymbol{\omega}}_{(k+1)} = \hat{\boldsymbol{\omega}}_k - \alpha_k \mathbf{E}_k^{-1} \mathbf{g}_k$$

em que:  $\hat{\boldsymbol{\omega}}_k$  é o vetor das estimativas atuais,  $\alpha_k$  é um fator multiplicativo do passo de correção,  $\mathbf{E}_k \in \mathbb{R}^{M \times M}$  é a matriz hessiana contendo todas as derivadas parciais de segunda ordem e  $\mathbf{g}_k \in \mathbb{R}^{M \times 1}$  é o vetor gradiente contendo todas as derivadas parciais de primeira ordem.

## C.2 Derivada primeira de $\mathbf{A}$

Considerando que:

$$\mathbf{A} = \begin{bmatrix} e^{j0\omega_1} & \dots & e^{j0\omega_m} \\ e^{j1\omega_1} & \dots & e^{j1\omega_m} \\ \vdots & \ddots & \vdots \\ e^{j(K-2)\omega_1} & \dots & e^{j(K-2)\omega_m} \\ e^{j(K-1)\omega_1} & \dots & e^{j(K-1)\omega_m} \end{bmatrix} \in \mathbb{C}^{K \times M}$$

Assim:

$$\frac{\partial \mathbf{A}}{\partial \omega_a} = \begin{bmatrix} \frac{\partial e^{j0\omega_1}}{\partial \omega_a} & \dots & \frac{\partial e^{j0\omega_m}}{\partial \omega_a} \\ \frac{\partial e^{j1\omega_1}}{\partial \omega_a} & \dots & \frac{\partial e^{j1\omega_m}}{\partial \omega_a} \\ \vdots & \ddots & \vdots \\ \frac{\partial e^{j(K-2)\omega_1}}{\partial \omega_a} & \dots & \frac{\partial e^{j(K-2)\omega_m}}{\partial \omega_a} \\ \frac{\partial e^{j(K-1)\omega_1}}{\partial \omega_a} & \dots & \frac{\partial e^{j(K-1)\omega_m}}{\partial \omega_a} \end{bmatrix}$$

Mas, para  $\omega_a \neq \omega_b$ :

$$\frac{\partial e^{j\omega_b}}{\partial \omega_a} = 0$$

Considerando  $\omega_a = \omega_b$  e  $\frac{\partial e^x}{\partial x} = x' e^x$ :

$$\begin{aligned} \frac{\partial e^{j\omega_a}}{\partial \omega_a} &= \{j\omega_a\}' e^{j\omega_a} \\ &= jce^{j\omega_a}, c = 0, \dots, (K-1) \end{aligned}$$

Assim, para  $a = 1$ :

$$\frac{\partial \mathbf{A}}{\partial \omega_1} = \begin{bmatrix} 0 \\ je^{j\omega_1} \\ \vdots \\ j(K-2)e^{j(K-2)\omega_1} \\ j(K-1)e^{j(K-1)\omega_1} \end{bmatrix} \mathbf{0} \in \mathbb{C}^{K \times M}$$

E, para  $a = M$ :

$$\frac{\partial \mathbf{A}}{\partial \omega_m} = \mathbf{0} \begin{bmatrix} 0 \\ je^{j\omega_m} \\ \vdots \\ j(K-2)e^{j(K-2)\omega_m} \\ j(K-1)e^{j(K-1)\omega_m} \end{bmatrix} \in \mathbb{C}^{K \times M}$$



### C.3 Derivada segunda de $\mathbf{A}$

Derivando a primeira derivada de  $\mathbf{A}$  produz:

$$\begin{aligned}\frac{\partial^2 e^{j\omega_a}}{\partial \omega_b \partial \omega_a} &= \frac{\partial}{\partial \omega_b} \left( \frac{\partial \mathbf{A}}{\partial \omega_a} \right) \\ &= \frac{\partial j c e^{j\omega_a}}{\partial \omega_b}\end{aligned}$$

Para  $\omega_a \neq \omega_b$ :

$$\frac{\partial^2 e^{j\omega_a}}{\partial \omega_b \partial \omega_a} = \frac{\partial j c e^{j\omega_a}}{\partial \omega_b} = 0$$

Para  $\omega_a = \omega_b$  e considerando que  $\frac{\partial e^x}{\partial x} = x' e^x$ :

$$\begin{aligned}\frac{\partial^2 e^{j\omega_a}}{\partial \omega_a \partial \omega_a} &= \frac{\partial j c e^{j\omega_a}}{\partial \omega_a} \\ &= j c (j c \omega_a)' e^{j\omega_a} \\ &= j c (j c) e^{j\omega_a} \\ &= -c^2 e^{j\omega_a}, c = 0, \dots, (K-1)\end{aligned}$$

Assim, para  $a = b = 1$ , temos:

$$\frac{\partial^2 \mathbf{A}}{\partial \omega_1 \partial \omega_1} = \begin{bmatrix} 0 \\ -e^{j1\omega_1} \\ \vdots \\ -(K-2)^2 e^{j(K-2)\omega_1} \\ -(K-1)^2 e^{j(K-1)\omega_1} \end{bmatrix} \mathbf{0} \in \mathbb{C}^{K \times M}$$

E, para  $a = b = M$ :

$$\frac{\partial^2 \mathbf{A}}{\partial \omega_m \partial \omega_m} = \begin{bmatrix} 0 \\ -e^{j\omega_m} \\ \vdots \\ -(K-2)^2 e^{j(K-2)\omega_m} \\ -(K-1)^2 e^{j(K-1)\omega_m} \end{bmatrix} \mathbf{0} \in \mathbb{C}^{K \times M}$$

Finalmente, considerando  $a \neq b$ , obtém-se:

$$\frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a} = \mathbf{0}_{K \times M}$$

## C.4 Derivada de $\mathbf{A}^+$

A derivada da pseudo-inversa de Moore-Penrose é (GOLUB; PEREYRA, 1973):

$$\frac{\partial \mathbf{A}^+}{\partial \omega_a} = -\mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ + \mathbf{A}^+ (\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} (\mathbf{I} - \mathbf{A} \mathbf{A}^+) + (\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \frac{\partial \mathbf{A}^H}{\partial \omega_a} (\mathbf{A}^+)^H \mathbf{A}^+$$

Mas  $\mathbf{A}^+ \mathbf{A} = \mathbf{I}$ , assim:

$$\frac{\partial \mathbf{A}^+}{\partial \omega_a} = -\mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ + \mathbf{A}^+ (\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} (\mathbf{I} - \mathbf{A} \mathbf{A}^+)$$

Mas  $\mathbf{P}_\mathbf{A}^\perp = (\mathbf{I} - \mathbf{P}_\mathbf{A}) = (\mathbf{I} - \mathbf{A} \mathbf{A}^+)$ , então:

$$\frac{\partial \mathbf{A}^+}{\partial \omega_a} = -\mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ + \mathbf{A}^+ (\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp$$

Mas  $\mathbf{A}^+ (\mathbf{A}^+)^H = (\mathbf{A}^H \mathbf{A})^{-1}$ , portanto:

$$\frac{\partial \mathbf{A}^+}{\partial \omega_a} = -\mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ + (\mathbf{A}^H \mathbf{A})^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp$$

Reordenando os termos:

$$\boxed{\frac{\partial \mathbf{A}^+}{\partial \omega_a} = (\mathbf{A}^H \mathbf{A})^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp - \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+}$$

## C.5 Derivada de $\mathbf{P}_\mathbf{A}^\perp$

Considerando que  $\mathbf{P}_\mathbf{A}^\perp = (\mathbf{I} - \mathbf{P}_\mathbf{A})$ , assim:

$$\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = \left( \frac{\partial \mathbf{I}}{\partial \omega_a} - \frac{\partial \mathbf{P}_\mathbf{A}}{\partial \omega_a} \right)$$

Como  $\mathbf{I}$  é constante em relação a  $\omega_a$ , temos:

$$\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = -\frac{\partial \mathbf{P}_\mathbf{A}}{\partial \omega_a}$$

Mas  $\mathbf{P}_\mathbf{A} = \mathbf{A} \mathbf{A}^+$ , então:

$$\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = -\frac{\partial \mathbf{A} \mathbf{A}^+}{\partial \omega_a}$$

Aplicando a regra da derivada do produto:

$$\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = -\left( \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ + \mathbf{A} \frac{\partial \mathbf{A}^+}{\partial \omega_a} \right)$$

Mas  $\frac{\partial \mathbf{A}^+}{\partial \omega_a} = (\mathbf{A}^H \mathbf{A})^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp - \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+$ , portanto:

$$\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = -\frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ - \mathbf{A} \left[ (\mathbf{A}^H \mathbf{A})^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp - \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \right]$$

Realizando a multiplicação:

$$\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = -\frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ - \mathbf{A} \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp + \mathbf{A} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+$$

Mas  $(\mathbf{A}^+)^H = \mathbf{A} \left( \mathbf{A}^H \mathbf{A} \right)^{-1}$ , então:

$$\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = -\frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ - (\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp + \mathbf{A} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+$$

Reordenando:

$$\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = -\frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ + \mathbf{A} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ - (\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp$$

Colocando um termo em evidência:

$$\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = -\left( \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ - \mathbf{A} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \right) - (\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp$$

Considerando  $\mathbf{A} \mathbf{I} = \mathbf{I} \mathbf{A} = \mathbf{A}$  e colocando dois termos em evidência:

$$\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = -(\mathbf{I} - \mathbf{A} \mathbf{A}^+) \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ - (\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp$$

Mas  $\mathbf{P}_\mathbf{A}^\perp = (\mathbf{I} - \mathbf{P}_\mathbf{A}) = (\mathbf{I} - \mathbf{A} \mathbf{A}^+)$ , dessa forma:

$$\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = -\mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ - (\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp$$

Reordenando:

$$\boxed{\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} = -(\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_\mathbf{A}^\perp - \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+}$$

## C.6 Gradiente de $\text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})$

Derivando a função-objetivo produz:

$$\nabla_{\phi_j} \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}}) = \frac{\partial \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_a}$$

Entretanto,  $\frac{\partial \text{tr} \mathbf{X}}{\partial x} = \text{tr} \left( \frac{\partial \mathbf{X}}{\partial x} \right)$ , então:

$$\frac{\partial \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_a} = \text{tr} \left( \frac{\partial \mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}}}{\partial \omega_a} \right)$$

Aplicando a regra da derivada do produto:

$$\frac{\partial \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_a} = \text{tr} \left( \frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_a} \hat{\mathbf{R}} + \mathbf{P}_\mathbf{A}^\perp \frac{\partial \hat{\mathbf{R}}}{\partial \omega_a} \right)$$

Entretanto,  $\hat{\mathbf{R}}$  é constante em relação a  $\omega_a$ , assim:

$$\frac{\partial \operatorname{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = \operatorname{tr} \left( \frac{\partial \mathbf{P}_A^\perp}{\partial \omega_a} \hat{\mathbf{R}} \right)$$

Mas  $\frac{\partial \mathbf{P}_A^\perp}{\partial \omega_a} = -(\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_A^\perp - \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+$ , assim:

$$\frac{\partial \operatorname{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = \operatorname{tr} \left\{ \left[ -(\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_A^\perp - \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \right] \hat{\mathbf{R}} \right\}$$

Realizando a multiplicação:

$$\frac{\partial \operatorname{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = \operatorname{tr} \left[ -(\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_A^\perp \hat{\mathbf{R}} - \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right]$$

Mas  $\mathbf{P}_A^\perp = \mathbf{P}_A^{\perp H}$  e  $\hat{\mathbf{R}} = \hat{\mathbf{R}}^H$ , então:

$$\frac{\partial \operatorname{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = \operatorname{tr} \left[ -(\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_A^{\perp H} \hat{\mathbf{R}}^H - \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right]$$

Colocando um termo em evidência e assumindo  $\operatorname{tr}(c\mathbf{A}) = c \operatorname{tr} \mathbf{A}$  para  $c$  escalar:

$$\frac{\partial \operatorname{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = -\operatorname{tr} \left[ (\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_A^{\perp H} \hat{\mathbf{R}}^H + \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right]$$

Aplicando  $\operatorname{tr}(\mathbf{A} + \mathbf{B}) = \operatorname{tr}(\mathbf{A}) + \operatorname{tr}(\mathbf{B})$ :

$$\frac{\partial \operatorname{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = - \left\{ \operatorname{tr} \left[ (\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_A^{\perp H} \hat{\mathbf{R}}^H \right] + \operatorname{tr} \left( \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right) \right\}$$

Aplicando  $\operatorname{tr}(\mathbf{ABC}) = \operatorname{tr}(\mathbf{BCA}) = \operatorname{tr}(\mathbf{CAB})$ :

$$\frac{\partial \operatorname{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = - \left\{ \operatorname{tr} \left[ \hat{\mathbf{R}}^H (\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_a} \mathbf{P}_A^{\perp H} \right] + \operatorname{tr} \left( \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right) \right\}$$

Aplicando  $(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H$ :

$$\frac{\partial \operatorname{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = - \left\{ \operatorname{tr} \left[ \left( \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right)^H \right] + \operatorname{tr} \left( \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right) \right\}$$

Mas  $\operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{A}^T)$  e  $\mathbf{A}^H = \mathbf{A}^{T*}$ , então  $\operatorname{tr}(\mathbf{A}^H) = \operatorname{tr}[(\mathbf{A}^T)^*] = \operatorname{tr}(\mathbf{A}^*) = \operatorname{tr}(\mathbf{A})^*$ .

Consequentemente:

$$\frac{\partial \operatorname{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = - \left[ \operatorname{tr} \left( \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right)^* + \operatorname{tr} \left( \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right) \right]$$

Mas  $z + z^* = 2\Re z$ , portanto:

$$\frac{\partial \operatorname{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = -2\Re \operatorname{tr} \left( \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right)$$

Aplicando  $\operatorname{tr}(\mathbf{ABC}) = \operatorname{tr}(\mathbf{BCA}) = \operatorname{tr}(\mathbf{CAB})$ :

$$\boxed{\frac{\partial \operatorname{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_a} = -2\Re \operatorname{tr} \left( \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \right)}$$

## C.7 Hessiana de $\text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})$

Derivando o gradiente da função-objetivo produz:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} &= \frac{\partial \nabla_{\omega_a} \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_b} \\ &= \frac{\partial}{\partial \omega_b} \left[ \nabla_{\omega_a} \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}}) \right] \\ &= \frac{\partial}{\partial \omega_b} \left[ \frac{\partial \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_a} \right] \end{aligned}$$

Mas  $\frac{\partial \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_a} = -2 \Re \text{tr} \left( \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right)$  então:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = \frac{\partial}{\partial \omega_b} \left[ -2 \Re \text{tr} \left( \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right) \right]$$

Aplicando  $\frac{\partial c\mathbf{X}}{\partial x} = c \frac{\partial \mathbf{X}}{\partial x}$  para  $c$  constante:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = -2 \left\{ \frac{\partial}{\partial \omega_b} \left[ \text{tr} \Re \left( \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right) \right] \right\}$$

Aplicando  $\frac{\partial \Re(\mathbf{X})}{\partial x} = \Re \left( \frac{\partial \mathbf{X}}{\partial x} \right)$ :

$$\frac{\partial^2 \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = -2 \Re \left\{ \frac{\partial}{\partial \omega_b} \left[ \text{tr} \left( \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right) \right] \right\}$$

Aplicando  $\frac{\partial \text{tr}(\mathbf{X})}{\partial x} = \text{tr} \left( \frac{\partial \mathbf{X}}{\partial x} \right)$ :

$$\frac{\partial^2 \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = -2 \Re \text{tr} \left[ \frac{\partial}{\partial \omega_b} \left( \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} \right) \right]$$

Aplicando a derivada do produto:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} &= -2 \Re \text{tr} \left[ \frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_b} \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} + \mathbf{P}_\mathbf{A}^\perp \frac{\partial}{\partial \omega_b} \left( \frac{\partial \mathbf{A}}{\partial \omega_a} \right) \mathbf{A}^+ \hat{\mathbf{R}} + \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \frac{\partial \mathbf{A}^+}{\partial \omega_b} \hat{\mathbf{R}} + \right. \\ &\quad \left. + \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \frac{\partial \hat{\mathbf{R}}}{\partial \omega_b} \right] \end{aligned}$$

Mas  $\hat{\mathbf{R}}$  é constante em relação a  $\omega_b$ , então:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = -2 \Re \text{tr} \left( \frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_b} \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} + \mathbf{P}_\mathbf{A}^\perp \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} + \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \frac{\partial \mathbf{A}^+}{\partial \omega_b} \hat{\mathbf{R}} \right)$$

Mas  $\frac{\partial \mathbf{P}_\mathbf{A}^\perp}{\partial \omega_b} = -(\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_\mathbf{A}^\perp - \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+$ , assim:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_\mathbf{A}^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} &= -2 \Re \text{tr} \left\{ \left[ -(\mathbf{A}^+)^H \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_\mathbf{A}^\perp - \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \right] \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} + \right. \\ &\quad \left. + \mathbf{P}_\mathbf{A}^\perp \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a} \mathbf{A}^+ \hat{\mathbf{R}} + \mathbf{P}_\mathbf{A}^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \frac{\partial \mathbf{A}^+}{\partial \omega_b} \hat{\mathbf{R}} \right\} \end{aligned}$$



## C.8 Aproximação assintótica na hessiana de $\text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})$

Adotando a mesma estratégia utilizada em (STARER, 1990; STARER; NEHORAI, 1992) é possível realizar uma aproximação assintótica na hessiana ao considerar  $N \rightarrow \infty$  e, assim, substituir a matriz de correlação estimada  $\hat{\mathbf{R}}$  por sua versão teórica  $\mathbf{R} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \sigma\mathbf{I}$ . Adicionalmente, é importante ressaltar que a hessiana pode ser tratada como uma matriz de ponderação com o propósito de aumentar a velocidade de convergência. Dessa forma, essa aproximação não influencia na acurácia do método, mas reduz o esforço computacional de maneira considerável.

Também é assumido que o vetor de parâmetros estimados esteja próximo de seu valor ótimo. Essa é uma afirmação válida, pois o objetivo do Método de Newton é apenas minimizar localmente a função-objetivo. Com essas considerações em vista, obtém-se o seguinte desenvolvimento:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left[ \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \right. \\ & \left. - \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{A}^+ \hat{\mathbf{R}} \left( \mathbf{A}^+ \right)^H \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} + \right. \\ & \left. + \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a} \right] \end{aligned}$$

Considerando a seção C.12,  $\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{P}_A^\perp \mathbf{R} \mathbf{P}_A^\perp = \sigma \mathbf{P}_A^\perp$  e  $a\mathbf{B}\mathbf{C} = \mathbf{B}a\mathbf{C} = \mathbf{B}\mathbf{C}a$  para  $a$  escalar, então:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left[ \sigma \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \right. \\ & \left. - \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{A}^+ \hat{\mathbf{R}} \left( \mathbf{A}^+ \right)^H \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} + \right. \\ & \left. + \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a} \right] \end{aligned}$$

Considerando a seção C.13,  $\mathbf{A}^+ \hat{\mathbf{R}} \left( \mathbf{A}^+ \right)^H \approx \mathbf{A}^+ \mathbf{R} \left( \mathbf{A}^+ \right)^H = \mathbf{P} + \sigma \left( \mathbf{A}^H \mathbf{A} \right)^{-1}$ :

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left\{ \sigma \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \right. \\ & \left. - \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} - \left[ \mathbf{P} + \sigma \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \right] \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} + \right. \\ & \left. + \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a} \right\} \end{aligned}$$

Realizando a multiplicação:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left[ \sigma \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \right. \\ & - \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{C} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \\ & \left. - \sigma \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} + \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a} \right] \end{aligned}$$

Reordenando os termos:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left[ \sigma \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \sigma \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \right. \\ & - \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{C} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} + \\ & \left. + \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a} \right] \end{aligned}$$

Realizando a subtração:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left( - \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} - \right. \\ & \left. - \mathbf{C} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} + \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a} \right) \end{aligned}$$

Reordenando os termos:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left( - \mathbf{C} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \right. \\ & \left. - \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} + \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a} \right) \end{aligned}$$

Aplicando  $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BCA}) = \text{tr}(\mathbf{CAB})$ :

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left( - \mathbf{C} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} - \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_b} - \right. \\ & \left. - \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_b} \mathbf{A}^+ \frac{\partial \mathbf{A}}{\partial \omega_a} + \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a} \right) \end{aligned}$$

Considerando a seção C.14,  $\mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{A}^+ \mathbf{R} \mathbf{P}_A^\perp = \mathbf{0}$ :

$$\frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = -2 \Re \text{tr} \left( - \mathbf{C} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \right)$$

Assumindo  $\text{tr}(c\mathbf{A}) = c \text{tr} \mathbf{A}$  para  $c$  escalar:

$$\boxed{\frac{\partial^2 \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}})}{\partial \omega_b \partial \omega_a} = 2 \Re \text{tr} \left( \mathbf{C} \frac{\partial \mathbf{A}^H}{\partial \omega_b} \mathbf{P}_A^\perp \frac{\partial \mathbf{A}}{\partial \omega_a} \right)}$$



## C.9 Prova 12: $\mathbf{A}^+ \mathbf{A} = \mathbf{I}$

Iniciando o desenvolvimento, sabendo que  $\mathbf{A}$  possui posto cheio:

$$\mathbf{A}^+ \mathbf{A} = \mathbf{A}^+ \mathbf{A}$$

Mas  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ , então:

$$\mathbf{A}^+ \mathbf{A} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A}$$

Entretanto,  $\mathbf{X}^{-1} \mathbf{X} = \mathbf{X} \mathbf{X}^{-1} = \mathbf{I}$ , portanto:

$$\boxed{\mathbf{A}^+ \mathbf{A} = \mathbf{I}}$$

## C.10 Prova 13: $(\mathbf{A}^+)^H = \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1}$

Iniciando o desenvolvimento:

$$(\mathbf{A}^+)^H = (\mathbf{A}^+)^H$$

Mas  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ , assim:

$$(\mathbf{A}^+)^H = \left[ (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \right]^H$$

Aplicando  $(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H$ :

$$(\mathbf{A}^+)^H = \mathbf{A} \left[ (\mathbf{A}^H \mathbf{A})^{-1} \right]^H$$

Entretanto,  $(\mathbf{A}^H \mathbf{A})^{-1} = \left[ (\mathbf{A}^H \mathbf{A})^{-1} \right]^H$ , então:

$$\boxed{(\mathbf{A}^+)^H = \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1}}$$

## C.11 Prova 14: $\mathbf{A}^+ (\mathbf{A}^+)^H = (\mathbf{A}^H \mathbf{A})^{-1}$

Iniciando o desenvolvimento, sabendo que  $\mathbf{A}$  possui posto cheio:

$$\mathbf{A}^+ (\mathbf{A}^+)^H = \mathbf{A}^+ (\mathbf{A}^+)^H$$

Considerando  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$  e  $(\mathbf{A}^+)^H = \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1}$ , então:

$$\mathbf{A}^+ (\mathbf{A}^+)^H = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1}$$

Entretanto,  $\mathbf{X}^{-1} \mathbf{X} = \mathbf{X} \mathbf{X}^{-1} = \mathbf{I}$  e  $\mathbf{IX} = \mathbf{XI} = \mathbf{X}$ , portanto:

$$\boxed{\mathbf{A}^+ (\mathbf{A}^+)^H = (\mathbf{A}^H \mathbf{A})^{-1}}$$

## C.12 Prova 15: $\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{P}_A^\perp \mathbf{R} \mathbf{P}_A^\perp = \sigma \mathbf{P}_A^\perp$

Ao aproximar  $\hat{\mathbf{R}}$  por  $\mathbf{R}$  e ao considerar que  $\mathbf{A}$  está próximo do valor ótimo obtém-se:

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{P}_A^\perp \mathbf{R} \mathbf{P}_A^\perp$$

Mas  $\mathbf{R} = \mathbf{A} \mathbf{P}_A^H + \sigma \mathbf{I}$ , então:

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{P}_A^\perp (\mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}) \mathbf{P}_A^\perp$$

Realizando a multiplicação:

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{P}_A^\perp \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_A^\perp + \sigma \mathbf{P}_A^\perp \mathbf{P}_A^\perp$$

Mas  $\mathbf{P}_A^\perp = \mathbf{I} - \mathbf{A} \mathbf{A}^+$ , assim:

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{P}_A^\perp \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_A^\perp + \sigma (\mathbf{I} - \mathbf{A} \mathbf{A}^+) (\mathbf{I} - \mathbf{A} \mathbf{A}^+)$$

Realizando a multiplicação:

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{P}_A^\perp \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_A^\perp + \sigma (\mathbf{I} - \mathbf{A} \mathbf{A}^+ - \mathbf{A} \mathbf{A}^+ + \mathbf{A} \mathbf{A}^+ \mathbf{A} \mathbf{A}^+)$$

Mas  $\mathbf{A}^+ \mathbf{A} = \mathbf{I}$ :

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{P}_A^\perp \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_A^\perp + \sigma (\mathbf{I} - \mathbf{A} \mathbf{A}^+ - \mathbf{A} \mathbf{A}^+ + \mathbf{A} \mathbf{A}^+)$$

Realizando a subtração:

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{P}_A^\perp \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_A^\perp + \sigma (\mathbf{I} - \mathbf{A} \mathbf{A}^+)$$

Mas  $\mathbf{P}_A^\perp = \mathbf{I} - \mathbf{A} \mathbf{A}^+$ , dessa forma:

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx (\mathbf{I} - \mathbf{A} \mathbf{A}^+) \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_A^\perp + \sigma (\mathbf{I} - \mathbf{A} \mathbf{A}^+)$$

Realizando a multiplicação:

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_A^\perp - \mathbf{A} \mathbf{A}^+ \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_A^\perp + \sigma (\mathbf{I} - \mathbf{A} \mathbf{A}^+)$$

Mas  $\mathbf{A}^+ \mathbf{A} = \mathbf{I}$ :

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_A^\perp - \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_A^\perp + \sigma (\mathbf{I} - \mathbf{A} \mathbf{A}^+)$$

Realizando a subtração:

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \sigma (\mathbf{I} - \mathbf{A} \mathbf{A}^+)$$

Mas  $\mathbf{P}_A^\perp = \mathbf{I} - \mathbf{A} \mathbf{A}^+$ , então:

$$\mathbf{P}_A^\perp \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \sigma \mathbf{P}_A^\perp$$

### C.13 Prova 16: $\mathbf{A}^+ \hat{\mathbf{R}} (\mathbf{A}^+)^H \approx \mathbf{A}^+ \mathbf{R} (\mathbf{A}^+)^H = \mathbf{P} + \sigma (\mathbf{A}^H \mathbf{A})^{-1}$

Ao aproximar  $\hat{\mathbf{R}}$  por  $\mathbf{R}$  e ao considerar que  $\mathbf{A}$  está próximo do valor ótimo obtém-se:

$$\mathbf{A}^+ \hat{\mathbf{R}} (\mathbf{A}^+)^H \approx \mathbf{A}^+ \mathbf{R} (\mathbf{A}^+)^H$$

Mas  $\mathbf{R} = \mathbf{A} \mathbf{P} \mathbf{A}^H + \sigma \mathbf{I}$ , então:

$$\mathbf{A}^+ \hat{\mathbf{R}} (\mathbf{A}^+)^H \approx \mathbf{A}^+ (\mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}) (\mathbf{A}^+)^H$$

Realizando a multiplicação à esquerda:

$$\mathbf{A}^+ \hat{\mathbf{R}} (\mathbf{A}^+)^H \approx (\mathbf{A}^+ \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{A}^+) (\mathbf{A}^+)^H$$

Realizando a multiplicação à direita:

$$\mathbf{A}^+ \hat{\mathbf{R}} (\mathbf{A}^+)^H \approx \mathbf{A}^+ \mathbf{A} \mathbf{C} \mathbf{A}^H (\mathbf{A}^+)^H + \sigma \mathbf{A}^+ (\mathbf{A}^+)^H$$

Mas  $\mathbf{A}^+ \mathbf{A} = \mathbf{I}$ :

$$\mathbf{A}^+ \hat{\mathbf{R}} (\mathbf{A}^+)^H \approx \mathbf{C} \mathbf{A}^H (\mathbf{A}^+)^H + \sigma \mathbf{A}^+ (\mathbf{A}^+)^H$$

Mas  $\mathbf{A}^+ (\mathbf{A}^+)^H = (\mathbf{A}^H \mathbf{A})^{-1}$ :

$$\mathbf{A}^+ \hat{\mathbf{R}} (\mathbf{A}^+)^H \approx \mathbf{C} \mathbf{A}^H (\mathbf{A}^+)^H + \sigma (\mathbf{A}^H \mathbf{A})^{-1}$$

Aplicando  $\mathbf{B}^H \mathbf{A}^H = (\mathbf{A} \mathbf{B})^H$ :

$$\mathbf{A}^+ \hat{\mathbf{R}} (\mathbf{A}^+)^H \approx \mathbf{C} (\mathbf{A}^+ \mathbf{A})^H + \sigma (\mathbf{A}^H \mathbf{A})^{-1}$$

Mas  $\mathbf{A}^+ \mathbf{A} = \mathbf{I}$ ,  $\mathbf{I}^H = \mathbf{I}$  e  $\mathbf{I} \mathbf{X} = \mathbf{X} \mathbf{I} = \mathbf{X}$ :

$$\mathbf{A}^+ \hat{\mathbf{R}} (\mathbf{A}^+)^H \approx \mathbf{P} + \sigma (\mathbf{A}^H \mathbf{A})^{-1}$$

### C.14 Prova 17: $\mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_\mathbf{A}^\perp \approx \mathbf{A}^+ \mathbf{R} \mathbf{P}_\mathbf{A}^\perp = \mathbf{0}$

Ao aproximar  $\hat{\mathbf{R}}$  por  $\mathbf{R}$  e ao considerar que  $\mathbf{A}$  está próximo do valor ótimo obtém-se:

$$\mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_\mathbf{A}^\perp \approx \mathbf{A}^+ \mathbf{R} \mathbf{P}_\mathbf{A}^\perp$$

Mas  $\mathbf{R} = \mathbf{A} \mathbf{P} \mathbf{A}^H + \sigma \mathbf{I}$ , então:

$$\mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_\mathbf{A}^\perp \approx \mathbf{A}^+ (\mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I}) \mathbf{P}_\mathbf{A}^\perp$$

Realizando as multiplicações:

$$\mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_\mathbf{A}^\perp \approx \mathbf{A}^+ \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_\mathbf{A}^\perp + \sigma \mathbf{A}^+ \mathbf{P}_\mathbf{A}^\perp$$

Mas  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ , assim:

$$\mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{P}_A^\perp + \sigma (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{P}_A^\perp$$

Entretanto,  $\mathbf{P}_A^\perp = \mathbf{I} - \mathbf{A} \mathbf{A}^+$ , então:

$$\mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A} \mathbf{C} \mathbf{A}^H (\mathbf{I} - \mathbf{A} \mathbf{A}^+) + \sigma (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H (\mathbf{I} - \mathbf{A} \mathbf{A}^+)$$

Realizando as multiplicações:

$$\begin{aligned} \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp &\approx (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A} \mathbf{C} \mathbf{A}^H - (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{A} \mathbf{A}^+ + \\ &+ \sigma \left[ (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H - (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A} \mathbf{A}^+ \right] \end{aligned}$$

Mas  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ , assim:

$$\begin{aligned} \mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp &\approx (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A} \mathbf{C} \mathbf{A}^H - (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A} \mathbf{C} \mathbf{A}^H \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H + \\ &+ \sigma \left[ (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H - (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \right] \end{aligned}$$

Entretanto,  $\mathbf{X}^{-1} \mathbf{X} = \mathbf{X} \mathbf{X}^{-1} = \mathbf{I}$  e  $\mathbf{I} \mathbf{X} = \mathbf{X} \mathbf{I} = \mathbf{X}$ , portanto:

$$\mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{C} \mathbf{A}^H - \mathbf{C} \mathbf{A}^H + \sigma \left[ (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H - (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \right]$$

Realizando as subtrações:

$$\boxed{\mathbf{A}^+ \hat{\mathbf{R}} \mathbf{P}_A^\perp \approx \mathbf{0}}$$

## C.15 Prova 18: $\mathbf{P}_A \mathbf{R} = \mathbf{P}_A \mathbf{R} \mathbf{P}_A = \mathbf{R} \mathbf{P}_A$

Iniciando o desenvolvimento:

$$\mathbf{P}_A \mathbf{R} = \mathbf{P}_A \mathbf{R}$$

Mas  $\mathbf{R} = \mathbf{A} \mathbf{P}_A^H + \sigma \mathbf{I}$ , então:

$$\mathbf{P}_A \mathbf{R} = \mathbf{A} \mathbf{A}^+ (\mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I})$$

Mas  $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ , assim:

$$\mathbf{P}_A \mathbf{R} = \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H (\mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{I})$$

Realizando as multiplicações:

$$\mathbf{P}_A \mathbf{R} = \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{A} \mathbf{C} \mathbf{A}^H + \sigma \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$$

Entretanto,  $\mathbf{X}^{-1}\mathbf{X} = \mathbf{X}\mathbf{X}^{-1} = \mathbf{I}$  e  $\mathbf{IX} = \mathbf{XI} = \mathbf{X}$ , portanto:

$$\mathbf{P}_\mathbf{A}\mathbf{R} = \mathbf{A}\mathbf{C}\mathbf{A}^\mathbf{H} + \sigma\mathbf{A}\left(\mathbf{A}^\mathbf{H}\mathbf{A}\right)^{-1}\mathbf{A}^\mathbf{H}$$

Iniciando o desenvolvimento:

$$\mathbf{P}_\mathbf{A}\mathbf{R}\mathbf{P}_\mathbf{A} = \mathbf{P}_\mathbf{A}\mathbf{R}\mathbf{P}_\mathbf{A}$$

Mas  $\mathbf{R} = \mathbf{A}\mathbf{P}\mathbf{A}^\mathbf{H} + \sigma\mathbf{I}$ , então:

$$\mathbf{P}_\mathbf{A}\mathbf{R}\mathbf{P}_\mathbf{A} = \mathbf{A}\mathbf{A}^+ \left( \mathbf{A}\mathbf{C}\mathbf{A}^\mathbf{H} + \sigma\mathbf{I} \right) \mathbf{A}\mathbf{A}^+$$

Mas  $\mathbf{A}^+ = \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H}$ , assim:

$$\mathbf{P}_\mathbf{A}\mathbf{R}\mathbf{P}_\mathbf{A} = \mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H} \left( \mathbf{A}\mathbf{C}\mathbf{A}^\mathbf{H} + \sigma\mathbf{I} \right) \mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H}$$

Realizando a multiplicação:

$$\mathbf{P}_\mathbf{A}\mathbf{R}\mathbf{P}_\mathbf{A} = \left[ \mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H} \mathbf{A} \mathbf{C} \mathbf{A}^\mathbf{H} + \sigma \mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H} \right] \mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H}$$

Entretanto,  $\mathbf{X}^{-1}\mathbf{X} = \mathbf{X}\mathbf{X}^{-1} = \mathbf{I}$  e  $\mathbf{IX} = \mathbf{XI} = \mathbf{X}$ , portanto:

$$\mathbf{P}_\mathbf{A}\mathbf{R}\mathbf{P}_\mathbf{A} = \left[ \mathbf{A}\mathbf{C}\mathbf{A}^\mathbf{H} + \sigma\mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H} \right] \mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H}$$

Realizando a multiplicação:

$$\mathbf{P}_\mathbf{A}\mathbf{R}\mathbf{P}_\mathbf{A} = \mathbf{A}\mathbf{C}\mathbf{A}^\mathbf{H}\mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H} + \sigma\mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H}\mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H}$$

Entretanto,  $\mathbf{X}^{-1}\mathbf{X} = \mathbf{X}\mathbf{X}^{-1} = \mathbf{I}$  e  $\mathbf{IX} = \mathbf{XI} = \mathbf{X}$ , portanto:

$$\mathbf{P}_\mathbf{A}\mathbf{R}\mathbf{P}_\mathbf{A} = \mathbf{A}\mathbf{C}\mathbf{A}^\mathbf{H} + \sigma\mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H}$$

Iniciando o desenvolvimento:

$$\mathbf{R}\mathbf{P}_\mathbf{A} = \mathbf{R}\mathbf{P}_\mathbf{A}$$

Mas  $\mathbf{R} = \mathbf{A}\mathbf{P}\mathbf{A}^\mathbf{H} + \sigma\mathbf{I}$ , então:

$$\mathbf{R}\mathbf{P}_\mathbf{A} = \left( \mathbf{A}\mathbf{C}\mathbf{A}^\mathbf{H} + \sigma\mathbf{I} \right) \mathbf{A}\mathbf{A}^+$$

Realizando a multiplicação:

$$\mathbf{R}\mathbf{P}_\mathbf{A} = \mathbf{A}\mathbf{C}\mathbf{A}^\mathbf{H}\mathbf{A}\mathbf{A}^+ + \sigma\mathbf{A}\mathbf{A}^+$$

Mas  $\mathbf{A}^+ = \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H}$ , assim:

$$\mathbf{R}\mathbf{P}_\mathbf{A} = \mathbf{A}\mathbf{C}\mathbf{A}^\mathbf{H}\mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H} + \sigma\mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H}$$

Entretanto,  $\mathbf{X}^{-1}\mathbf{X} = \mathbf{X}\mathbf{X}^{-1} = \mathbf{I}$  e  $\mathbf{IX} = \mathbf{XI} = \mathbf{X}$ , portanto:

$$\mathbf{R}\mathbf{P}_\mathbf{A} = \mathbf{A}\mathbf{C}\mathbf{A}^\mathbf{H} + \sigma\mathbf{A} \left( \mathbf{A}^\mathbf{H}\mathbf{A} \right)^{-1} \mathbf{A}^\mathbf{H}$$

Finalmente, pode-se considerar que:

$$\mathbf{P}_\mathbf{A}\mathbf{R} = \mathbf{P}_\mathbf{A}\mathbf{R}\mathbf{P}_\mathbf{A} = \mathbf{R}\mathbf{P}_\mathbf{A}$$

# D DERIVAÇÃO DO MÉTODO DE NEWTON PARA O DMLE COM FILTRAGEM ESPACIAL USANDO A MATRIZ DE DIREÇÃO

## D.1 Introdução

A função-objetivo do Critério Determinístico do Estimador de Máxima Verossimilhança (DMLE) com filtragem espacial e usando a matriz de direção filtrada  $\tilde{\mathbf{A}}$  é (KRUMME-NAUER et al., 2010):

$$F_{\text{DMLEF}} = \text{tr} \left( \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z \right)$$

em que:  $\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} = (\mathbf{I} - \mathbf{P}_{\tilde{\mathbf{A}}}) = (\mathbf{I} - \tilde{\mathbf{A}}\tilde{\mathbf{A}}^+)$  e  $\hat{\mathbf{R}}_Z = \mathbf{H}\hat{\mathbf{R}}\mathbf{H}^H$ . Adicionalmente,  $\tilde{\mathbf{A}} = \mathbf{H}\mathbf{A}$  e  $\mathbf{H}$  é a matriz de convolução do filtro estimado. Assim como mostrado no Apêndice C, são utilizados os vetores de parâmetros  $\hat{\omega}_0$  e  $\hat{\omega}_1$  com dimensões  $(M \times 1)$ . Assim, a hessiana e o gradiente possuem dimensões  $(M \times M)$  e  $(M \times 1)$ , respectivamente.

## D.2 Derivada primeira de $\tilde{\mathbf{A}}$

Considerando que  $\tilde{\mathbf{A}} = \mathbf{H}\mathbf{A}$ , então:

$$\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} = \frac{\partial \mathbf{H}\mathbf{A}}{\partial \omega_a}$$

Aplicando a derivada do produto:

$$\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} = \frac{\partial \mathbf{H}}{\partial \omega_a} \mathbf{A} + \mathbf{H} \frac{\partial \mathbf{A}}{\partial \omega_a}$$

Mas  $\mathbf{H}$  é constante em relação a  $\omega_a$ , assim:

$$\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} = \mathbf{H} \frac{\partial \mathbf{A}}{\partial \omega_a}$$

## D.3 Derivada segunda de $\tilde{\mathbf{A}}$

Derivando a primeira derivada de  $\tilde{\mathbf{A}}$  produz:

$$\frac{\partial^2 \tilde{\mathbf{A}}}{\partial \omega_b \partial \omega_a} = \frac{\partial}{\partial \omega_b} \left( \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \right)$$

Mas  $\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} = \mathbf{H} \frac{\partial \mathbf{A}}{\partial \omega_a}$ , assim:

$$\frac{\partial^2 \tilde{\mathbf{A}}}{\partial \omega_b \partial \omega_a} = \frac{\partial \left( \mathbf{H} \frac{\partial \mathbf{A}}{\partial \omega_a} \right)}{\partial \omega_b}$$

Aplicando a derivada do produto:

$$\frac{\partial^2 \tilde{\mathbf{A}}}{\partial \omega_b \partial \omega_a} = \frac{\partial \mathbf{H}}{\partial \omega_b} \frac{\partial \mathbf{A}}{\partial \omega_a} + \mathbf{H} \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a}$$

Mas  $\mathbf{H}$  é constante em relação a  $\omega_b$ , assim:

$$\boxed{\frac{\partial^2 \tilde{\mathbf{A}}}{\partial \omega_b \partial \omega_a} = \mathbf{H} \frac{\partial^2 \mathbf{A}}{\partial \omega_b \partial \omega_a}}$$

## D.4 Derivada de $\tilde{\mathbf{A}}^+$

A derivada da pseudo-inversa de Moore-Penrose é (GOLUB; PEREYRA, 1973):

$$\frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_a} = -\tilde{\mathbf{A}}^+ \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ + \tilde{\mathbf{A}}^+ \left( \tilde{\mathbf{A}}^+ \right)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \left( \mathbf{I} - \tilde{\mathbf{A}} \tilde{\mathbf{A}}^+ \right) + \left( \mathbf{I} - \tilde{\mathbf{A}}^+ \tilde{\mathbf{A}} \right) \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \left( \tilde{\mathbf{A}}^+ \right)^H \tilde{\mathbf{A}}^+$$

Mas  $\tilde{\mathbf{A}}^+ \tilde{\mathbf{A}} = \mathbf{I}$ , então:

$$\frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_a} = -\tilde{\mathbf{A}}^+ \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ + \tilde{\mathbf{A}}^+ \left( \tilde{\mathbf{A}}^+ \right)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \left( \mathbf{I} - \tilde{\mathbf{A}} \tilde{\mathbf{A}}^+ \right)$$

Mas  $\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = (\mathbf{I} - \mathbf{P}_{\tilde{\mathbf{A}}}) = (\mathbf{I} - \tilde{\mathbf{A}} \tilde{\mathbf{A}}^+)$ , assim:

$$\frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_a} = -\tilde{\mathbf{A}}^+ \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ + \tilde{\mathbf{A}}^+ \left( \tilde{\mathbf{A}}^+ \right)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \mathbf{P}_{\tilde{\mathbf{A}}}^\perp$$

Mas  $\tilde{\mathbf{A}}^+ \tilde{\mathbf{A}}^{+H} = \left( \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} \right)^{-1}$ , portanto:

$$\frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_a} = -\tilde{\mathbf{A}}^+ \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ + \left( \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} \right)^{-1} \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \mathbf{P}_{\tilde{\mathbf{A}}}^\perp$$

Reordenando os termos:

$$\boxed{\frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_a} = \left( \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} \right)^{-1} \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \mathbf{P}_{\tilde{\mathbf{A}}}^\perp - \tilde{\mathbf{A}}^+ \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+}$$

## D.5 Derivada de $\mathbf{P}_{\tilde{\mathbf{A}}}^\perp$

Considerando que  $\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = (\mathbf{I} - \mathbf{P}_{\tilde{\mathbf{A}}})$ , então:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = \left( \frac{\partial \mathbf{I}}{\partial \omega_a} - \frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}}{\partial \omega_a} \right)$$

Como  $\mathbf{I}$  é constante em relação a  $\omega_a$ , temos:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = -\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}}{\partial \omega_a}$$

Mas  $\mathbf{P}_{\tilde{\mathbf{A}}} = \tilde{\mathbf{A}}\tilde{\mathbf{A}}^+$ , então:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = -\frac{\partial \tilde{\mathbf{A}}\tilde{\mathbf{A}}^+}{\partial \omega_a}$$

Aplicando a regra da derivada do produto:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = -\left(\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+ + \tilde{\mathbf{A}}\frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_a}\right)$$

Mas  $\frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_a} = \left(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}\right)^{-1}\frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a}\mathbf{P}_{\tilde{\mathbf{A}}}^\perp - \tilde{\mathbf{A}}^+\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+$ , dessa forma:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = -\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+ - \tilde{\mathbf{A}}\left[\left(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}\right)^{-1}\frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a}\mathbf{P}_{\tilde{\mathbf{A}}}^\perp - \tilde{\mathbf{A}}^+\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+\right]$$

Realizando a multiplicação:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = -\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+ - \tilde{\mathbf{A}}\left(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}\right)^{-1}\frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a}\mathbf{P}_{\tilde{\mathbf{A}}}^\perp + \tilde{\mathbf{A}}\tilde{\mathbf{A}}^+\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+$$

Mas  $\left(\tilde{\mathbf{A}}^+\right)^H = \tilde{\mathbf{A}}\left(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}\right)^{-1}$ , assim:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = -\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+ - \left(\tilde{\mathbf{A}}^+\right)^H\frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a}\mathbf{P}_{\tilde{\mathbf{A}}}^\perp + \tilde{\mathbf{A}}\tilde{\mathbf{A}}^+\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+$$

Reordenando:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = -\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+ + \tilde{\mathbf{A}}\tilde{\mathbf{A}}^+\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+ - \left(\tilde{\mathbf{A}}^+\right)^H\frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a}\mathbf{P}_{\tilde{\mathbf{A}}}^\perp$$

Colocando um termo em evidência:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = -\left(\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+ - \tilde{\mathbf{A}}\tilde{\mathbf{A}}^+\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+\right) - \left(\tilde{\mathbf{A}}^+\right)^H\frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a}\mathbf{P}_{\tilde{\mathbf{A}}}^\perp$$

Colocando um termo em evidência e considerando  $\mathbf{A}\mathbf{I} = \mathbf{I}\mathbf{A} = \mathbf{A}$ :

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = -\left(\mathbf{I} - \tilde{\mathbf{A}}\tilde{\mathbf{A}}^+\right)\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+ - \left(\tilde{\mathbf{A}}^+\right)^H\frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a}\mathbf{P}_{\tilde{\mathbf{A}}}^\perp$$

Mas  $\mathbf{P}_{\tilde{\mathbf{A}}}^\perp = (\mathbf{I} - \mathbf{P}_{\tilde{\mathbf{A}}}) = (\mathbf{I} - \tilde{\mathbf{A}}\tilde{\mathbf{A}}^+)$ , assim:

$$\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = -\mathbf{P}_{\tilde{\mathbf{A}}}^\perp\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+ - \left(\tilde{\mathbf{A}}^+\right)^H\frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a}\mathbf{P}_{\tilde{\mathbf{A}}}^\perp$$

Reordenando:

$$\boxed{\frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_a} = -\left(\tilde{\mathbf{A}}^+\right)^H\frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a}\mathbf{P}_{\tilde{\mathbf{A}}}^\perp - \mathbf{P}_{\tilde{\mathbf{A}}}^\perp\frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a}\tilde{\mathbf{A}}^+}$$



## D.6 Gradiente de $\text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)$

Derivando a função-objetivo produz:

$$\nabla_{\phi_j} \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z) = \frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_a}$$

Entretanto,  $\frac{\partial \text{tr} \mathbf{X}}{\partial x} = \text{tr} \left( \frac{\partial \mathbf{X}}{\partial x} \right)$ , então:

$$\frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_a} = \text{tr} \left( \frac{\partial \mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z}{\partial \omega_a} \right)$$

Aplicando a regra da derivada do produto:

$$\frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_a} = \text{tr} \left( \frac{\partial \mathbf{P}_{\hat{\mathbf{A}}}^\perp}{\partial \omega_a} \hat{\mathbf{R}}_Z + \mathbf{P}_{\hat{\mathbf{A}}}^\perp \frac{\partial \hat{\mathbf{R}}_Z}{\partial \omega_a} \right)$$

Mas  $\hat{\mathbf{R}}_Z$  é constante em relação a  $\omega_a$ , assim:

$$\frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_a} = \text{tr} \left( \frac{\partial \mathbf{P}_{\hat{\mathbf{A}}}^\perp}{\partial \omega_a} \hat{\mathbf{R}}_Z \right)$$

Porém,  $\frac{\partial \mathbf{P}_{\hat{\mathbf{A}}}^\perp}{\partial \omega_a} = -(\tilde{\mathbf{A}}^+)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \mathbf{P}_{\hat{\mathbf{A}}}^\perp - \mathbf{P}_{\hat{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+$ , então:

$$\frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_a} = \text{tr} \left\{ \left[ -(\tilde{\mathbf{A}}^+)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \mathbf{P}_{\hat{\mathbf{A}}}^\perp - \mathbf{P}_{\hat{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \right] \hat{\mathbf{R}}_Z \right\}$$

Realizando a multiplicação:

$$\frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_a} = \text{tr} \left[ -(\tilde{\mathbf{A}}^+)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z - \mathbf{P}_{\hat{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right]$$

Considerando  $\mathbf{P}_{\hat{\mathbf{A}}}^\perp = \mathbf{P}_{\hat{\mathbf{A}}}^{\perp H}$  e  $\hat{\mathbf{R}}_Z = \hat{\mathbf{R}}_Z^H$  obtém-se:

$$\frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_a} = \text{tr} \left[ -(\tilde{\mathbf{A}}^+)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \mathbf{P}_{\hat{\mathbf{A}}}^{\perp H} \hat{\mathbf{R}}_Z^H - \mathbf{P}_{\hat{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right]$$

Colocando um termo em evidência e assumindo  $\text{tr}(c\mathbf{A}) = c \text{tr} \mathbf{A}$  para  $c$  escalar:

$$\frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_a} = -\text{tr} \left[ (\tilde{\mathbf{A}}^+)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \mathbf{P}_{\hat{\mathbf{A}}}^{\perp H} \hat{\mathbf{R}}_Z^H + \mathbf{P}_{\hat{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right]$$

Aplicando  $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$ :

$$\frac{\partial \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_a} = -\left\{ \text{tr} \left[ (\tilde{\mathbf{A}}^+)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \mathbf{P}_{\hat{\mathbf{A}}}^{\perp H} \hat{\mathbf{R}}_Z^H \right] + \text{tr} \left( \mathbf{P}_{\hat{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right) \right\}$$

Aplicando  $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BCA}) = \text{tr}(\mathbf{CAB})$ :

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_a} = - \left\{ \text{tr} \left[ \hat{\mathbf{R}}_Z^H (\tilde{\mathbf{A}}^+)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_a} \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp H} \right] + \text{tr} \left( \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right) \right\}$$

Aplicando  $(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H$ :

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_a} = - \left\{ \text{tr} \left[ \left( \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right)^H \right] + \text{tr} \left( \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right) \right\}$$

Entretanto,  $\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}^T)$  e  $\mathbf{A}^H = \mathbf{A}^{T*}$ , assim,  $\text{tr}(\mathbf{A}^H) = \text{tr}[(\mathbf{A}^T)^*] = \text{tr}(\mathbf{A}^*) = \text{tr}(\mathbf{A})^*$ . Finalmente:

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_a} = - \left[ \text{tr} \left( \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right)^* + \text{tr} \left( \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right) \right]$$

Mas  $z + z^* = 2 \Re z$ , dessa forma:

$$\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_a} = -2 \Re \text{tr} \left( \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right)$$

Aplicando  $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BCA}) = \text{tr}(\mathbf{CAB})$ :

$$\boxed{\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_a} = -2 \Re \text{tr} \left( \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \right)}$$

## D.7 Hessiana de $\text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)$

Derivando o gradiente da função-objetivo produz:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} &= \frac{\partial \nabla_{\omega_a} \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_b} \\ &= \frac{\partial}{\partial \omega_b} \left[ \nabla_{\omega_a} \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z) \right] \\ &= \frac{\partial}{\partial \omega_b} \left[ \frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_a} \right] \end{aligned}$$

Mas  $\frac{\partial \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_a} = -2 \Re \text{tr} \left( \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right)$ , portanto:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} = \frac{\partial}{\partial \omega_b} \left[ -2 \Re \text{tr} \left( \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right) \right]$$

Aplicando  $\frac{\partial c\mathbf{X}}{\partial x} = c \frac{\partial \mathbf{X}}{\partial x}$  para  $c$  constante:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} = -2 \left\{ \frac{\partial}{\partial \omega_b} \left[ \text{tr} \Re \left( \mathbf{P}_{\tilde{\mathbf{A}}}^{\perp} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right) \right] \right\}$$

Aplicando  $\frac{\partial \Re(\mathbf{X})}{\partial x} = \Re\left(\frac{\partial \mathbf{X}}{\partial x}\right)$ :

$$\frac{\partial^2 \text{tr} \left( \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z \right)}{\partial \omega_b \partial \omega_a} = -2 \Re \left\{ \frac{\partial}{\partial \omega_b} \left[ \text{tr} \left( \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right) \right] \right\}$$

Aplicando  $\frac{\partial \text{tr}(\mathbf{X})}{\partial x} = \text{tr} \left( \frac{\partial \mathbf{X}}{\partial x} \right)$ :

$$\frac{\partial^2 \text{tr} \left( \mathbf{P}_{\mathbf{\hat{A}}}^\perp \mathbf{\hat{R}}_Z \right)}{\partial \omega_b \partial \omega_a} = -2 \Re \text{tr} \left[ \frac{\partial}{\partial \omega_b} \left( \mathbf{P}_{\mathbf{\hat{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}} + \mathbf{\hat{R}}_Z \right) \right]$$

Aplicando a derivada do produto:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left[ \frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_b} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z + \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial}{\partial \omega_b} \left( \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \right) \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z + \right. \\ & \left. + \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_b} \hat{\mathbf{R}}_Z + \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \frac{\partial \hat{\mathbf{R}}_Z}{\partial \omega_b} \right] \end{aligned}$$

Mas  $\hat{\mathbf{R}}_Z$  é constante em relação a  $\omega_b$ , então:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} = -2 \Re \text{tr} \left[ \frac{\partial \mathbf{P}_{\hat{\mathbf{A}}}^\perp}{\partial \omega_b} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z + \mathbf{P}_{\hat{\mathbf{A}}}^\perp \frac{\partial}{\partial \omega_b} \left( \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \right) \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z + \mathbf{P}_{\hat{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_b} \hat{\mathbf{R}}_Z \right]$$

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\hat{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} = -2 \Re \text{tr} \left[ \frac{\partial \mathbf{P}_{\hat{\mathbf{A}}}^\perp}{\partial \omega_b} \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z + \mathbf{P}_{\hat{\mathbf{A}}}^\perp \frac{\partial^2 \tilde{\mathbf{A}}}{\partial \omega_b \partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z + \mathbf{P}_{\hat{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_b} \hat{\mathbf{R}}_Z \right]$$

$$\text{Mas } \frac{\partial \mathbf{P}_{\tilde{\mathbf{A}}}^\perp}{\partial \omega_b} = - \left( \tilde{\mathbf{A}}^+ \right)^{\text{H}} \frac{\partial \tilde{\mathbf{A}}^{\text{H}}}{\partial \omega_b} \mathbf{P}_{\tilde{\mathbf{A}}}^\perp - \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_b} \tilde{\mathbf{A}}^+, \text{ consequentemente:}$$

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left\{ \left[ -(\tilde{\mathbf{A}}^+)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_b} \mathbf{P}_{\tilde{\mathbf{A}}}^\perp - \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_b} \tilde{\mathbf{A}}^+ \right] \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z + \right. \\ & \left. + \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial^2 \tilde{\mathbf{A}}}{\partial \omega_b \partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z + \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_b} \hat{\mathbf{R}}_Z \right\} \end{aligned}$$

Realizando a multiplicação:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left[ -(\tilde{\mathbf{A}}^+)^{\text{H}} \frac{\partial \tilde{\mathbf{A}}^{\text{H}}}{\partial \omega_b} \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z - \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_b} \tilde{\mathbf{A}}^+ \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z + \right. \\ & \left. + \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial^2 \tilde{\mathbf{A}}}{\partial \omega_b \partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z + \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_b} \hat{\mathbf{R}}_Z \right] \end{aligned}$$

Reordenando:

$$\begin{aligned} \frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} = & -2 \Re \text{tr} \left[ \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \frac{\partial \tilde{\mathbf{A}}^+}{\partial \omega_b} \hat{\mathbf{R}}_Z - (\tilde{\mathbf{A}}^+)^{\text{H}} \frac{\partial \tilde{\mathbf{A}}^{\text{H}}}{\partial \omega_b} \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z - \right. \\ & \left. - \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_b} \tilde{\mathbf{A}}^+ \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z + \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial^2 \tilde{\mathbf{A}}}{\partial \omega_b \partial \omega_a} \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z \right] \end{aligned}$$



Não foi possível realizar nenhuma aproximação assintótica como foi apresentado na Seção C.8. No entanto, foi feita uma análise numérica da hessiana considerando todas as combinações de um, dois, três, quatro e cinco termos. Nessa análise foi observado que o quarto termo é responsável pela convergência do Método de Newton. Como o propósito dessa aproximação é diminuir o esforço computacional, a hessiana se torna:

$$\frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} = -2 \Re \text{tr} \left[ -\tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z (\tilde{\mathbf{A}}^+)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_b} \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \right]$$

Colocando um termo em evidência e assumindo  $\text{tr}(c\mathbf{A}) = c \text{tr} \mathbf{A}$  para  $c$  escalar:

$$\boxed{\frac{\partial^2 \text{tr}(\mathbf{P}_{\tilde{\mathbf{A}}}^\perp \hat{\mathbf{R}}_Z)}{\partial \omega_b \partial \omega_a} = 2 \Re \text{tr} \left[ \tilde{\mathbf{A}}^+ \hat{\mathbf{R}}_Z (\tilde{\mathbf{A}}^+)^H \frac{\partial \tilde{\mathbf{A}}^H}{\partial \omega_b} \mathbf{P}_{\tilde{\mathbf{A}}}^\perp \frac{\partial \tilde{\mathbf{A}}}{\partial \omega_a} \right]}$$

## D.9 Prova 19: $\tilde{\mathbf{A}}^+ \tilde{\mathbf{A}} = \mathbf{I}$

Como  $\tilde{\mathbf{A}}$  possui posto cheio, realiza-se o mesmo procedimento mostrado na seção C.9, assim:

$$\boxed{\tilde{\mathbf{A}}^+ \tilde{\mathbf{A}} = \mathbf{I}}$$

## D.10 Prova 20: $(\tilde{\mathbf{A}}^+)^H = \tilde{\mathbf{A}} (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1}$

Realiza-se o mesmo procedimento mostrado na seção C.10, assim:

$$\boxed{(\tilde{\mathbf{A}}^+)^H = \tilde{\mathbf{A}} (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1}}$$

## D.11 Prova 21: $\tilde{\mathbf{A}}^+ (\tilde{\mathbf{A}}^+)^H = (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1}$

Realiza-se o mesmo procedimento mostrado na seção C.11, assim:

$$\boxed{\tilde{\mathbf{A}}^+ (\tilde{\mathbf{A}}^+)^H = (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1}}$$

## D.12 Prova 22: $\mathbf{P}_{\tilde{\mathbf{A}}} \mathbf{P}_{\tilde{\mathbf{A}}} = \mathbf{P}_{\tilde{\mathbf{A}}}$

Iniciando o desenvolvimento, sabendo que  $\tilde{\mathbf{A}}$  possui posto cheio:

$$\mathbf{P}_{\tilde{\mathbf{A}}} \mathbf{P}_{\tilde{\mathbf{A}}} = \mathbf{P}_{\tilde{\mathbf{A}}} \mathbf{P}_{\tilde{\mathbf{A}}}$$

Mas  $\mathbf{P}_{\tilde{\mathbf{A}}} = \tilde{\mathbf{A}} \tilde{\mathbf{A}}^+ = \tilde{\mathbf{A}} (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^H$ , então:

$$\mathbf{P}_{\tilde{\mathbf{A}}} \mathbf{P}_{\tilde{\mathbf{A}}} = \tilde{\mathbf{A}} (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^H$$

Realizando a multiplicação:

$$\mathbf{P}_{\tilde{\mathbf{A}}} \mathbf{P}_{\tilde{\mathbf{A}}} = \tilde{\mathbf{A}} \left( \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} \right)^{-1} \tilde{\mathbf{A}}^H$$

No entanto,  $\mathbf{P}_{\tilde{\mathbf{A}}} = \tilde{\mathbf{A}} \tilde{\mathbf{A}}^+ = \tilde{\mathbf{A}} \left( \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} \right)^{-1} \tilde{\mathbf{A}}^H$ , assim:

$$\boxed{\mathbf{P}_{\tilde{\mathbf{A}}} \mathbf{P}_{\tilde{\mathbf{A}}} = \mathbf{P}_{\tilde{\mathbf{A}}}}$$

## E RESULTADOS AUXILIARES

Este apêndice apresenta os erros de estimação dos estimadores e filtros abordados nesta tese nas situações de fontes próximas ( $10^\circ$  e  $15^\circ$ ) e afastadas ( $10^\circ$  e  $45^\circ$ ). É importante ressaltar que esses dados são referentes às curvas de desempenho tracejadas em preto-e-branco nas Figuras 29 a 48. Além disso, foi incluído nas tabelas o Limite de Cramér-Rao (CRB) para comparação.

Tabela 11 – Raiz do Erro Quadrático Médio em graus dos estimadores avaliados e o Limite de Cramér-Rao correspondente considerando duas fontes descorrelacionadas situadas em  $10^\circ$  e  $15^\circ$ .

[illegible]





Tabela 14 – Raiz do Erro Quadrático Médio em graus dos estimadores avaliados e o Limite de Cramér-Rao correspondente considerando duas fontes correlacionadas situadas em  $10^\circ$  e  $45^\circ$ .

SNR (dB)	CRB	MODEX Mod	SEAD Aprim	SEAD Mod	MODEX Mod H1	SEAD Aprim H1	SEAD Mod H1	MODEX Mod H2	SEAD Aprim H2	SEAD Mod H2	MODEX Mod H3	SEAD Aprim H3	SEAD Mod H3
-15,00	1,673	17,948	16,740	11,214	20,692	20,148	24,752	17,924	19,894	15,928	16,010	12,649	11,300
-13,75	1,302	10,677	9,278	5,985	17,549	11,418	15,623	11,424	11,949	9,900	9,858	7,111	6,042
-12,50	1,023	5,843	4,534	2,224	14,572	4,739	17,545	6,585	6,611	4,945	4,842	3,042	2,389
-11,25	0,811	1,892	1,242	1,887	11,472	1,634	10,526	2,038	1,561	1,957	1,919	1,256	2,017
-10,00	0,649	0,878	0,821	0,821	7,902	1,101	7,114	0,920	0,896	0,894	0,900	0,812	0,813
-8,75	0,525	0,651	0,657	0,657	4,381	0,846	4,196	0,681	0,699	0,700	0,678	0,661	0,662
-7,50	0,429	0,522	0,533	0,533	2,009	0,666	1,855	0,538	0,561	0,562	0,553	0,545	0,545
-6,25	0,354	0,429	0,438	0,439	0,921	0,531	0,931	0,443	0,459	0,459	0,460	0,451	0,452
-5,00	0,295	0,357	0,365	0,364	0,882	0,429	0,890	0,367	0,379	0,379	0,387	0,377	0,377
-3,75	0,247	0,300	0,306	0,305	0,323	0,351	0,350	0,309	0,317	0,317	0,323	0,316	0,316
-2,50	0,209	0,255	0,257	0,257	0,269	0,289	0,289	0,261	0,267	0,267	0,271	0,267	0,267
-1,25	0,177	0,217	0,217	0,218	0,225	0,241	0,240	0,222	0,226	0,226	0,228	0,225	0,226
0,00	0,151	0,185	0,185	0,186	0,190	0,202	0,201	0,190	0,192	0,193	0,192	0,191	0,192
1,25	0,129	0,159	0,158	0,159	0,162	0,170	0,170	0,162	0,164	0,165	0,163	0,163	0,164
2,50	0,111	0,136	0,137	0,137	0,138	0,145	0,144	0,139	0,143	0,141	0,139	0,141	0,140
3,75	0,095	0,117	0,118	0,117	0,119	0,124	0,123	0,119	0,122	0,122	0,119	0,121	0,120
5,00	0,082	0,101	0,102	0,101	0,102	0,107	0,105	0,102	0,106	0,105	0,103	0,104	0,103
6,25	0,071	0,087	0,088	0,087	0,088	0,092	0,090	0,088	0,092	0,090	0,088	0,090	0,089
7,50	0,061	0,075	0,076	0,075	0,076	0,078	0,078	0,076	0,080	0,078	0,076	0,078	0,077
8,75	0,053	0,065	0,066	0,065	0,065	0,068	0,067	0,066	0,069	0,067	0,066	0,067	0,066
10,00	0,046	0,056	0,058	0,056	0,057	0,060	0,058	0,057	0,061	0,058	0,057	0,059	0,057
11,25	0,040	0,049	0,051	0,049	0,049	0,053	0,050	0,049	0,054	0,050	0,049	0,053	0,049
12,50	0,034	0,042	0,044	0,042	0,042	0,046	0,043	0,042	0,048	0,044	0,042	0,045	0,043
13,75	0,030	0,036	0,039	0,036	0,037	0,040	0,037	0,037	0,041	0,038	0,037	0,040	0,037
15,00	0,026	0,032	0,034	0,032	0,032	0,034	0,032	0,032	0,036	0,033	0,032	0,035	0,032