



DISSERTAÇÃO DE MESTRADO

**ESTIMADOR DE DIREÇÃO DE CHEGADA EM TEMPO REAL  
COM ARRANJO DE MICROFONES**

André Tomaz Gontijo

Brasília, Agosto de 2010

**UNIVERSIDADE DE BRASÍLIA**

FACULDADE DE TECNOLOGIA

**UNIVERSIDADE DE BRASÍLIA  
FACULDADE DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**ESTIMADOR DE DIREÇÃO DE CHEGADA EM TEMPO REAL  
COM ARRANJO DE MICROFONES**

**ANDRÉ TOMAZ GONTIJO**

**ORIENTADOR: RICARDO ZELENOVSKY**

**DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA**

**PUBLICAÇÃO: PPGENE.DM – 428/2010**

**BRASÍLIA/DF, AGOSTO – 2010**

**UNIVERSIDADE DE BRASÍLIA  
FACULDADE DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**DISSERTAÇÃO DE MESTRADO**

**ESTIMADOR DE DIREÇÃO DE CHEGADA EM TEMPO REAL  
COM ARRANJO DE MICROFONES**

ANDRÉ TOMAZ GONTIJO

DISSERTAÇÃO DE MESTRADO ACADÊMICO SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE.

**APROVADA POR:**

---

RICARDO ZELENOVSKY, DOUTOR - PUC-RJ, UnB/ENE

(ORIENTADOR)

---

JOÃO PAULO CARVALHO LUSTOSA DA COSTA, Dr.-Ing. - Technische Universitaet Ilmenau (TU Ilmenau), Alemanha, UnB/ENE

(EXAMINADOR INTERNO)

---

JOSÉ ANTONIO APOLINÁRIO JÚNIOR, D.Sc. - COPPE/UFRJ, IME

(EXAMINADOR EXTERNO)

BRASÍLIA, 27 DE AGOSTO DE 2010.

## FICHA CATALOGRÁFICA

GONTIJO, ANDRÉ TOMAZ

Estimador de Direção de Chegada em Tempo Real com Arranjo de Microfones [Distrito Federal], 2010.

v. 1, 138p., 210x297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2010).

Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia Departamento de Engenharia Elétrica.

- |                                |                        |
|--------------------------------|------------------------|
| 1. Estimção DOA                | 2. Arranjo de sensores |
| 3. Filtragem Espacial          | 4. Sinais de áudio     |
| 5. Processamento em tempo real | 6. SRP-PHAT            |

I. ENE/FT/UnB

## REFERÊNCIA BIBLIOGRÁFICA

GONTIJO, ANDRÉ TOMAZ (2010). Estimador de Direção de Chegada em Tempo Real com Arranjo de Microfones. Dissertação de Mestrado em Engenharia Elétrica, Publicação PGENE.DM-428/2010, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 138 p.

## CESSÃO DE DIREITOS

AUTOR: André Tomaz Gontijo.

TÍTULO: Estimador de Direção de Chegada em Tempo Real com Arranjo de Microfones.

GRAU/ANO: Mestre/2010.

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e

científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito do autor.

André Tomaz Gontijo  
Brasília – DF – Brasil.

*À minha família.*

## **AGRADECIMENTOS**

*Agradeço a Deus porque tem nos concedido capacidade e graça para realizar este projeto e todas as outras coisas.*

*À minha família pelo apoio cotidiano e incondicional durante todos os anos que até aqui se sucederam. Em especial, ao meu pai pelo companheirismo, paciência e participação nos trabalhos.*

*Ao meu orientador Prof. Ricardo Zelenovsky pelo apoio e dedicação. Aos professores João Paulo e Apolinário pelo trabalho de revisão.*

## RESUMO

O presente trabalho trata do estudo de métodos para a estimação da direção de chegada – DOA (*Direction of Arrival*) – de uma frente de ondas continuamente. São apresentadas diferentes técnicas para a estimação DOA como: DS, CAPON, MUSIC e SRP-PHAT.

Para avaliar o desempenho dos métodos de estimação, são realizadas simulações em diferentes cenários. A fim de validar os resultados através de medições, foi construído um hardware específico composto por um arranjo de microfones para captação de ondas sonoras. Desta forma, além de se avaliar as técnicas em termos de acuracidade, também obtém-se uma estimativa preliminar do tempo de processamento.

Como principais contribuições pode-se citar a extensão do SRP-PHAT para o acompanhamento em tempo real de fonte sonora em movimento (*DOA tracking*). Para validar o algoritmo proposto, utiliza-se o hardware construído neste trabalho. São realizados testes com o hardware e software projetados e apresentados os resultados.

## ABSTRACT

The present work deals with the study of DOA (Direction of Arrival) continuous estimation methods for a sound wavefront. Different techniques for DOA estimation like DS, CAPON, MUSIC, ESPRIT and SRP-PHAT are presented.

In order to evaluate the performance of DOA techniques in a real environment, simulations are performed to study the different methods. For validating results through real measures, a hardware prototype was developed to capture sound waves through a uniform linear array (ULA) of microphones. In a preliminary way, processing time is also taken in account for real-time implementation.

As contributions are the implementation of SRP-PHAT algorithm for real-time DOA tracking in a real environment and the development of hardware and software for it. Different scenarios are tested with the hardware/software developed and results are shown.

## SUMÁRIO

1	INTRODUÇÃO .....	1
1.1	MOTIVAÇÃO.....	1
1.2	OBJETIVOS .....	2
1.3	OVERVIEW E CONTRIBUIÇÕES.....	3
2	DIREÇÃO DE CHEGADA EM ARRANJO DE SENSORES .....	5
2.1	ADEQUAÇÃO DO SISTEMA PARA ESTIMAÇÃO.....	5
2.2	VIABILIDADE DE ESTIMAÇÃO DOA PARA SISTEMAS REAIS .....	6
2.3	ARRANJOS DE SENSORES.....	8
2.4	DISTÂNCIA ENTRE SENSORES .....	16
2.5	A BANDA DE FREQUÊNCIA DOS SINAIS.....	19
3	MÉTODOS DE ESTIMAÇÃO DA DIREÇÃO DE CHEGADA.....	24
3.1	DS- ATRASO E SOMA .....	25
3.2	CAPON .....	28
3.3	MUSIC.....	32
3.4	SRP-PHAT .....	36
3.5	OUTRAS ABORDAGENS PARA ESTIMAÇÃO DOA.....	67
4	ESTIMAÇÃO DOA .....	77
4.1	HARDWARE: AQUISIÇÃO, CONDICIONAMENTO E AMPLIFICAÇÃO .....	78
4.2	IMPLEMENTAÇÃO DA ESTIMAÇÃO DOA.....	90
5	RESULTADOS EXPERIMENTAIS .....	108
5.1	ENSAIOS ESTÁTICOS.....	110
5.2	ENSAIOS DINÂMICOS .....	120
6	CONCLUSÃO E TRABALHO FUTURO.....	126
	REFERÊNCIAS BIBLIOGRÁFICAS.....	129
	ANEXO I .....	133
	ANEXO II .....	135

## LISTA DE FIGURAS

Figura 2.1: Sinal $s(t)$ incidente num ULA de M elementos, segundo o azimute $\theta$ .....	9
Figura 2.2: Geometria para frentes de onda planas.....	9
Figura 2.3: Geometria para frentes planas.....	13
Figura 2.4: Frente de ondas (seta negra) incidindo sobre um arranjo planar uniforme.....	14
Figura 2.5 Largura do pico de potência em função da separação entre sensores.....	17
Figura 2.6 Largura do pico máximo em função da separação dos elementos do arranjo .....	18
Figura 2.7 Pico máximo de potência obtido para diferentes distâncias entre sensores.....	19
Figura 2.8 Resposta em frequência de um verbete de voz .....	20
Figura 2.9 Resposta em frequência de um seno de 200 Hz.....	20
Figura 2.10 Resposta em frequência teórica.....	21
Figura 2.11 Resposta em frequência real.....	22
Figura 3.1 – DOA para um sinal de incidência .....	26
Figura 3.2 – Aplicação da filtragem espacial na recuperação de um sinal incidente em $30^\circ$ .....	27
Figura 3.3 – Estimação de DOA usando o método CAPON para 2 sinais de incidência .....	29
Figura 3.4 – Filtragem espacial baseada na estimação DOA com o emprego do estimador CAPON.....	30
Figura 3.5 – Entrada com vários sinais e saída do método CAPON juntamente com a FFT da saída.....	31
Figura 3.6 – Estimação de DOA método MUSIC para 3 sinais - voz em $0^\circ$ , ruído em $30^\circ$ e música a $60^\circ$ . ( $SNR = 0,82dB$ ) .....	34
Figura 3.7 – Estimação de DOA com o método MUSIC para um sinal senoidal incidindo a $30^\circ$ .....	35
Figura 3.8: Pré-filtragem dos sinais em tempo discreto.....	37
Figura 3.9: Sinal sonoro utilizado na simulação .....	44
Figura 3.10: Função de correlação sem transformada de fase de um sinal com atrasos correspondentes a 10 e 15 amostras .....	45
Figura 3.11: Função de correlação com transformada de fase para um sinal com atrasos de 10 e 15 amostras.....	45
Figura 3.12: Trecho sonoro considerado nas simulações.....	47
Figura 3.13: Trecho sonoro filtrado com filtro digital considerado nas simulações .....	48
Figura 3.14: Tom senoidal de 1 kHz considerado nas simulações .....	48
Figura 3.15: Variação do EMA com relação ao SNR para um trecho sonoro: MUSIC, CAPON e DS.....	49
Figura 3.16: Dispersão do erro com relação ao SNR para um trecho sonoro: MUSIC, CAPON e DS.....	50
Figura 3.17: Variação do Erro com relação ao SNR para um trecho sonoro SRP e SRP-PHAT .....	51
Figura 3.18: Dispersão do erro com relação ao SNR para um trecho sonoro SRP e SRP-PHAT .....	51
Figura 3.19: Variação do EMA com relação ao SNR para um trecho sonoro filtrado em 1 kHz: MUSIC, CAPON e DS.....	52
Figura 3.20: Dispersão do erro com relação ao SNR para um trecho sonoro filtrado em 1 kHz: MUSIC, CAPON e DS.....	53
Figura 3.21: Variação do EMA com relação ao SNR para um trecho sonoro filtrado em 1 kHz SRP e SRP-PHAT .....	54
Figura 3.22: Dispersão do erro com relação ao SNR para um trecho sonoro filtrado em 1 kHz: SRP e SRP-PHAT .....	54
Figura 3.23: Variação do Erro com relação ao SNR para um tom senoidal de 1 kHz para MUSIC, CAPON e DS.....	55
Figura 3.24: Dispersão do erro com relação ao SNR para um tom senoidal de 1 kHz: MUSIC, CAPON e DS .....	56
Figura 3.25: Variação do Erro com relação ao SNR para um tom senoidal de 1 kHz SRP e SRP-PHAT.....	56
Figura 3.26: Dispersão do EMA com relação ao SNR para um tom senoidal de 1 kHz: SRP e SRP-PHAT .....	57
Figura 3.27: Variação do EMA com relação à potência do eco para um trecho sonoro: MUSIC, CAPON e DS.....	58
Figura 3.28: Variação do EMA com relação à potência do eco para um trecho sonoro: SRP e SRP-PHAT.....	59
Figura 3.29: Variação do EMA com relação à potência do eco para um trecho sonoro filtrado em 1 kHz: MUSIC, CAPON e DS.....	60
Figura 3.30: Variação do EMA com relação à potência do eco para um trecho sonoro filtrado em 1 kHz: SRP e SRP-PHAT .....	61
Figura 3.31: Variação do EMA com relação à potência do eco para um tom senoidal de 1 kHz: MUSIC, CAPON e DS .....	62
Figura 3.32: Variação do EMA com relação à potência do eco para um tom senoidal de 1 kHz: SRP e SRP-PHAT.....	62
Figura 3.33: Variação do EMA com relação à separação do interferente para um trecho sonoro: MUSIC, CAPON e DS.....	63
Figura 3.34: Variação do EMA com relação à separação do interferente para um trecho sonoro: SRP e SRP-PHAT .....	63
Figura 3.35: Variação do EMA com relação à separação do interferente para um trecho sonoro filtrado em 1 kHz: MUSIC, CAPON e DS.....	64
Figura 3.36: Variação do EMA com relação à separação do interferente para um trecho sonoro filtrado em 1 kHz: SRP e SRP-PHAT .....	65
Figura 3.37: Variação do EMA em relação à separação do interferente para um tom senoidal de 1 kHz: MUSIC, CAPON e DS .....	66
Figura 3.38: Variação do erro com relação à separação do interferente para um tom senoidal de 1 kHz: SRP e SRP-PHAT. ....	67
Figura 3.39: Arranjo de Doublets .....	75
Figura 4.1: Visão geral do sistema de estimação DOA .....	77
Figura 4.2: Curva de resposta de magnitude x frequência para o eletreto Panasonic WM-61A .....	79
Figura 4.3: Microfones WM-61A.....	79
Figura 4.4: Especificações LME49740 segundo seu fabricante.....	80
Figura 4.5: Curva de fase e atraso de grupo LME49740 .....	80
Figura 4.6: Diagrama esquemático do estágio de amplificação .....	82
Figura 4.7: Resposta de magnitude, fase e atraso de grupo para o filtro Bessel ordem 2 .....	84
Figura 4.8: Topologia Sallen-Key .....	84
Figura 4.9: Simulação para o bloco de filtragem analógica .....	86
Figura 4.10: Circuito para fonte de tensão +9V regulada.....	87

Figura 4.11: Circuito para fonte de tensão -9V regulada.....	87
Figura 4.12: Dispositivo de digitalização USB-1608FS .....	88
Figura 4.13: Placa de circuito impresso para amplificação e filtragem .....	90
Figura 4.14: Ruído ambiente com escala de 100 mV .....	91
Figura 4.15: Sinal sonoro com escala de 1 V .....	91
Figura 4.16: <i>Snapshot</i> típico instantâneo de um microfone.....	92
Figura 4.17: Dispersão das estimativas DOA em função do tamanho do snapshot .....	93
Figura 4.18: Correlação teórica simulada para um atraso de 3 amostras entre canais .....	94
Figura 4.19: Correlação real obtida para um atraso de 3 amostras entre canais .....	95
Figura 4.20: Correlação teórica simulada para um atraso de 6 amostras entre canais .....	95
Figura 4.21: Correlação real obtida para um atraso de 6 amostras entre canais .....	96
Figura 4.22: Máximo de atraso entre canais adjacentes: onda incidente em 90°.....	97
Figura 4.23: Correlação real obtida para um atraso de três amostras entre canais .....	98
Figura 4.24: Correlação real com interpolação obtida para um atraso de três amostras entre canais .....	99
Figura 4.25: Correlação real obtida para um atraso de seis amostras entre canais .....	99
Figura 4.26: Correlação real com interpolação obtida para um atraso de seis amostras entre canais .....	99
Figura 4.27: Ambiguidade na estimação DOA para o arranjo linear .....	100
Figura 4.28: Correlação real com interpolação obtida para um atraso de três amostras entre canais .....	102
Figura 4.29: Correlação real com interpolação obtida para um atraso de 6 amostras entre canais .....	102
Figura 4.30: Função de correlação final para um <i>delay</i> teórico de $t = 3$ amostras. ....	104
Figura 4.31: Janela deslizante com esquecimento exponencial .....	106
Figura 4.32: Fator de esquecimento – Janela Deslizante .....	107
Figura 5.1: Arranjo Uniforme Linear de Microfones.....	108
Figura 5.2: <i>Tracking</i> DOA .....	109
Figura 5.3: DOA estimada para alvo fixo 45° E.....	111
Figura 5.4: Histograma DOA sem limiar de potência e sliding-window – 45° E. Erro Médio: 4,46° .....	111
Figura 5.5: Histograma DOA com limiar de potência e sliding-window – 45° E .....	112
Figura 5.6: DOA estimada para alvo fixo 30° E. Erro médio: 3,63°.....	112
Figura 5.7: Histograma DOA sem limiar de potência e janela deslizante – 30° E .....	113
Figura 5.8: Histograma DOA com limiar de potência e janela deslizante – 30° E .....	113
Figura 5.9: DOA estimada para alvo fixo 15°E. Erro médio: 3,62°.....	114
Figura 5.10: Histograma DOA sem limiar de potência e janela deslizante – 15°E .....	114
Figura 5.11: Histograma DOA com limiar de potência e janela deslizante – 15°E.....	115
Figura 5.12: DOA estimada para alvo fixo 0°. Erro médio: 3,61°.....	115
Figura 5.13: Histograma DOA sem limiar de potência e janela deslizante – 0° .....	115
Figura 5.14: Histograma DOA com limiar de potência e janela deslizante – 0° .....	116
Figura 5.15: DOA estimada para alvo fixo 15°D. Erro médio: 1,90° .....	116
Figura 5.16: Histograma DOA sem limiar de potência e janela deslizante– 15°D.....	117
Figura 5.17: Histograma DOA com limiar de potência e janela deslizante – 15°D.....	117
Figura 5.18: DOA estimada para alvo fixo 30°D. Erro médio: 1,99° .....	117
Figura 5.19: Histograma DOA sem limiar de potência e janela deslizante – 30°D.....	118
Figura 5.20: Histograma DOA com limiar de potência e janela deslizante– 30°D .....	118
Figura 5.21: DOA estimada para alvo fixo 45°D. Erro médio: 1,99° .....	119
Figura 5.22: Histograma DOA sem limiar de potência e janela deslizante – 45°D.....	119
Figura 5.23: Histograma DOA com limiar de potência e janela deslizante – 45°D.....	119
Figura 5.24: Movimento angular uniforme de 0 a 45° para direita do arranjo .....	121
Figura 5.25: Movimento angular uniforme de 0 a 45° para esquerda do arranjo.....	122
Figura 5.26: Movimento angular uniforme de -45° a 45° .....	122
Figura 5.27: Movimento angular uniforme de -45°D a 30°D a 30°E .....	123

**LISTA DE TABELAS**

Tabela 1. Resumo dos ensaios estatísticos - <i>Tracking</i> .....	120
Tabela 2. Resumo dos ensaios dinâmicos .....	124
Tabela 3. Resumo dos ensaios dinâmicos – Porcentagem de erros .....	124

## LISTA DE SÍMBOLOS

### Símbolos Gregos

$\theta$	Azimute de incidência	[rad]
$\phi$	Ângulo de elevação	[rad]
$\varphi$	Defasagem angular	[rad]
$\lambda$	Comprimento de onda	[m]
$\Phi$	Matriz diagonal de defasamentos entre canais no ESPRIT	
$\Psi$	Matriz de transformação	

### Outros Símbolos

$d$	Distância entre elementos de um arranjo de sensores	[m]
$v$	Velocidade da onda	[m/s]
$f$	Frequência da onda	[Hz]
$t$	Tempo	[s]
$\omega$	Velocidade angular	[rad/s]
$X$	Eixo de plano cartesiano no modelo do URA	
$Y$	Eixo de plano cartesiano no modelo do URA	
$Z$	Eixo de plano cartesiano no modelo do URA	
$\mathbf{u}$	Vetor unitário na direção X	

$\mathbf{v}$	Vetor unitário na direção Y
$\mathbf{w}$	Vetor unitário na direção Z
$\mathbf{A}(\theta)$	Matriz Diretora
$\mathbf{a}(\theta)$	Vetor volante, ou vetor diretor, ou <i>steering vector</i>
$\mathbf{S}$	Vetor ou matriz de sinais incidentes no arranjo de sensores
$\mathbf{N}$	Vetor ou matriz de ruído branco aditivo (AWGN)
$DOA$	Vetor tridimensional representando a direção de chegada
$i$	Linha do plano de sensores no URA
$j$	Coluna do plano de sensores no URA
$M$	Quantidade de sensores do ULA ou quantidade de sensores em uma linha horizontal do URA
$DP$	Distância percorrida pelo sinal até o sensor (i,j)
$D$	Distância entre dois elementos no URA
$\theta_k$	Direção de chegada da fonte de sinais para o sensor k
$K$	Quantidade de fontes de sinais
$\mathbf{E}_n$	Vetor ou Matriz de autovetores de ruído
$\sigma_n^2$	Variância do ruído
$\mathbf{R}_{xx}$	Matriz de correlação da saída dos sensores

$R_{ss}$  Matriz de correlação da saída dos sinais

$I$  Matriz identidade

$E\{ \}$  Valor Esperado

$P_{DS}$  Potência no algoritmo de estimação DS

$P_{CA}$  Potência no algoritmo de estimação CAPON

$P_{MUS}$  Potência no algoritmo de estimação MUSIC

### Sobrescritos

· Variação temporal

\_ Valor médio

† Operador Hermitiano (conjugado do transposto)

\* Transposto conjugado

### Siglas

ADC *Analogic to digital converter*

A/D Analógico para digital

ARM *Advances Risc Machines Ltda.*

BW *Bandwidth*

CAPON Algoritmo de estimação de direção de chegada

CDMA	<i>Code Division Multiple Access</i>
D/A	Digital para analógico
DC	<i>Direct Current</i>
DS	<i>Delay and Sum</i>
DOA	<i>Direction of Arrival</i>
DSP	<i>Digital Signal Processor</i>
ESPRIT	<i>Estimation of Signal Parameters via Rotational Invariance Technique</i>
FFT	<i>Fast Fourier Transform</i>
MATLAB <sup>®</sup>	É marca registrada da MathWorks, Inc.
MUSIC	<i>Multiple Signal Classification</i>
MUX	Multiplexador
PAI	Prótese Auditiva inteligente
PCI	<i>Peripheral Component Interconnect</i>
PCM	<i>Pulse code modulation</i>
PHAT	<i>Phase Transform</i>
RNID	<i>Royal National Institute for Deaf People</i>
SDMA	<i>Spatial Division Multiple Access</i>
SMD	<i>Surface Mounting Device</i>

SNR	<i>Signal – Noise Range</i>
SPI	<i>Serial Peripheral Interface</i> - Interface serial de periférico
SRP	<i>Steered Response Power</i>
THD	<i>Total Harmonic Distortion</i>
ULA	<i>Uniform Linear Array</i>
URA	<i>Uniform rectangular array</i>
USB	<i>Universal Serial Bus</i> - Barramento serial universal

# 1 INTRODUÇÃO

## 1.1 MOTIVAÇÃO

A utilização de arranjo de sensores para a localização de fontes emissoras de ondas tem despertado grande interesse por sua vasta gama de aplicações. Entre elas estão a prótese auditiva inteligente [6, 12, 13, 22], sistemas de teleconferência com detecção automática do interlocutor [23, 26, 27, 37], radares com antenas fixas e sonares [36], exames médicos de precisão [6], detecção de vazamentos em dutos [19], sistemas de acesso múltiplo por divisão no espaço (SDMA) [20, 30, 44], sistemas de telecomunicação de terceira e quarta gerações, entre outros [1, 6, 20, 30, 27, 31, 37, 44].

O projeto de sistemas de comunicação sem-fio, para atender as necessidades de aumento de largura de banda e do número de usuários, precisa maximizar a utilização do meio eletromagnético (espectro de frequência) definido para a interface aérea. Dada a limitação de banda disponível para cada serviço, um sistema de filtragem espacial, aliado a um sistema de rastreamento (*tracking*), pode aumentar a capacidade do sistema por permitir outro grau de liberdade sobre os já existentes: arranjos de antenas podem estimar a direção da fonte do sinal, o que viabiliza um sistema que direcione o diagrama de irradiação do conjunto de antenas em tempo real, permitindo uma relação sinal-ruído melhorada, maior eficiência e capacidade de seletividade espacial [12, 13, 20, 22, 30, 36, 44].

Os arranjos de microfones, um caso específico de arranjos de sensores de banda base, os quais serão discutidos nesta dissertação nos capítulos 2, 4 e 5, possuem aplicações diretas na recuperação e tratamento de sinais sonoros em ambientes ruidosos [4, 23] ou com alto nível de reverberação [7, 16, 45, 37].

Desta forma, as técnicas de estimação da direção de chegada (DOA) têm grande importância no contexto dos sistemas de rastreamento e localização de fontes de ondas em geral.

A estimação da DOA com confiabilidade, robustez e precisão constitui premissa importante para os mecanismos de filtragem espacial. Uma vez determinada a direção de chegada, pode-se, facilmente, e com baixo esforço computacional, implementar algoritmos para filtros espaciais em uma, duas ou  $R$  dimensões [4, 11].

## 1.2 OBJETIVOS

O objetivo do presente trabalho é o estudo e avaliação das técnicas de estimação e do hardware e software necessários para o desenvolvimento de um sistema de rastreamento da direção de chegada (DOA *tracking*) de sinais sonoros por um arranjo linear e uniforme de sensores (microfones) em tempo real.

O projeto de hardware inclui a transdução e condicionamento do sinal, a amplificação em níveis e relação sinal-ruído adequadas para digitalização e filtragem analógica na faixa de maior potência do espectro de voz. Para viabilizar a estimação da direção de chegada, o projeto dos circuitos elétricos deve atender a importantes restrições como: alta impedância de entrada e baixos níveis de distorção, fase linear na banda passante, baixo acoplamento entre canais e melhor relação sinal-ruído possível. Neste projeto, são utilizados amplificadores ativos de alta linearidade e baixa distorção, além de microfones com resposta plana para frequências entre 20  $Hz$  e 20  $kHz$ . Este representa uma evolução do hardware anteriormente desenvolvido em [22].

No desenvolvimento do projeto de software é utilizado o ambiente MATLAB<sup>®</sup> para o processamento de sinais digitais e análise de técnicas e dos resultados. O dispositivo USB-1608FS (*Measurement Computing*) é utilizado para a aquisição dos sinais analógicos.

Para a análise das diferentes técnicas de estimação DOA são realizadas simulações com vistas a caracterizar os algoritmos. Em ambiente de simulação, o desempenho dos algoritmos é avaliado em diferentes perspectivas: em condições ideais de operação; em diferentes relações SNR (sinal-ruído); em níveis crescentes de reverberação; para sinais interferentes em ângulos próximos ao sinal de interesse.

Consideradas, em conjunto, as respostas dos algoritmos às condições de operação, é escolhido o algoritmo SRP-PHAT (*Steered Response Power with Phase Transform*) para a estimação da direção de chegada da fonte sonora em ambiente real. Diversos testes são realizados para fixação de parâmetros como a taxa de amostragem, o número de amostras para cada quadro (*snapshot*), a faixa de variação do sinal analógico, etc.

O desempenho do sistema é avaliado em diferentes situações para o caso estático (fonte fixa) e dinâmico. É desenvolvida, no MATLAB<sup>®</sup>, interface gráfica para visualização, em tempo real, da posição instantânea e das últimas posições estimadas da fonte, o que constitui a informação básica necessária e suficiente para posterior filtragem espacial dos sinais recebidos.

### **1.3 OVERVIEW E CONTRIBUIÇÕES**

No Capítulo 2 são apresentados os conceitos envolvidos nas técnicas de estimação da direção de chegada (DOA) por arranjo de sensores. São detalhadas as premissas adotadas e as restrições que devem ser consideradas para o uso dos estimadores.

No Capítulo 3 são apresentados os estimadores DS, CAPON, MUSIC e SRP-PHAT para estimação DOA, além de revisão de literatura sobre o tema. Neste capítulo, além do

detalhamento do funcionamento de cada técnica de estimação, são apresentadas simulações computacionais com vistas a exemplificar o uso de cada uma delas.

O Capítulo 4 está dividido em duas seções. Na primeira, é detalhado o desenvolvimento do hardware de aquisição e condicionamento dos sinais analógicos sonoros, incluindo a captação, amplificação e filtragem analógica. Ainda nesta seção, são justificadas as escolhas de projeto e apresentadas premissas para o hardware que devem ser atendidas para o bom funcionamento dos métodos de estimação.

Na segunda seção, é apresentada a descrição completa do sistema de rastreamento da direção de chegada. É apresentado um diagrama funcional do sistema e detalhadas, bloco a bloco, todas as etapas do processo de estimação após a amplificação e filtragem dos sinais analógicos.

No Capítulo 5 são apresentados os resultados experimentais obtidos com o sistema de rastreamento DOA. Testes são realizados com a fonte de ondas espacialmente fixa e móvel.

O Capítulo 6 apresenta nossas conclusões e sugestões de trabalho futuro. Como principais contribuições citamos o projeto de circuitos amplificadores e filtros analógicos de alta fidelidade e grande linearidade de fase; o comparativo, por simulações de diferentes cenários, de técnicas de estimação DOA; a implementação e avaliação, em ambiente real, de um sistema de rastreamento da direção de chegada de ondas sonoras em tempo real com o uso da técnica SRP-PHAT.

## 2 DIREÇÃO DE CHEGADA EM ARRANJO DE SENSORES

### 2.1 ADEQUAÇÃO DO SISTEMA PARA ESTIMAÇÃO

O desenvolvimento das técnicas de estimação de direção de chegada – DOA – baseia-se em três conceitos fundamentais e complementares: arranjo de sensores, processamento de sinais e métodos adaptativos.

Primeiramente, os arranjos de sensores, que podem ser compostos por diversos tipos de transdutores, são utilizados para recepção ou transmissão de sinais, exibindo um diagrama de radiação ou captação característico (*beamformer*). Tais arranjos podem assumir diversas geometrias para se adaptar a diferentes aplicações. Os transdutores utilizados no âmbito deste trabalho são modelados como omnidirecionais.

Técnicas de estimação de DOA, aliadas a um tratamento estatístico adequado, permitem formar um conjunto de estimativas confiável para a estimação da posição angular da fonte. Essas estimativas constituem a informação básica para a formação de diagramas de captação adaptativos (antenas inteligentes ou *smart antennas*) [20, 44].

O uso de diagramas adaptativos provê uma relação sinal/ruído (SNR) melhorada e pode diminuir a interferência co-canal e de multipercurso, pelo recurso da filtragem espacial [4, 6].

Para a modelagem dos arranjos de sensores e dos sinais que serão apresentados ao longo deste capítulo, supõe-se que:

- 1) O número de sinais incidentes é finito, o que fornecerá uma solução única na solução das equações lineares sob as quais o problema é definido. Sendo assim, o número de sensores utilizado no arranjo de sensores determina o número máximo de sinais cuja direção de chegada é passível de determinação.

- 2) Os emissores de onda estão suficientemente distantes de forma que as frentes de onda incidentes podem ser consideradas planas. Para sinais sonoros, em especial, considera-se que o afastamento de 15 vezes a distância entre dois elementos adjacentes é suficiente para a aproximação de frente de onda plana. Por exemplo, para um arranjo linear uniforme compacto, composto por um número reduzido de sensores, com distância de  $15\text{cm}$  entre sensores adjacentes, a fonte sonora deve estar, no mínimo, a  $2.25\text{m}$  [26].
- 3) Assim, considera-se a distância mínima da fonte em relação ao arranjo de sensores de  $15d$ , em que  $d$  é a distância entre elementos adjacentes do arranjo linear. Para este caso, o erro devido à aproximação por ondas planas é menor que dez graus.
- 4) Considera-se que os transdutores e canais elétricos de filtragem e amplificação sejam idênticos e que não exista acoplamento eletromagnético entre eles.

## 2.2 VIABILIDADE DE ESTIMAÇÃO DOA PARA SISTEMAS REAIS

Diferentes tipos de sistemas podem utilizar técnicas de direção de chegada combinadas com arranjos de sensores. Trata-se de uma área abrangente na qual as aplicações podem envolver tanto ondas eletromagnéticas de alta frequência (rádio AM/FM, *wi-fi*, *wi-max*, celular, *bluetooth*, etc.) quanto ondas mecânicas de baixa frequência, como música e voz.

A aplicação da teoria de arranjos, na prática, deve levar em consideração limitações inerentes aos sistemas reais. Para microfones, existe o problema do sinal de interesse possuir banda larga, o que torna a estimação mais difícil que até para sinais UWB (*ultra wideband*). Além disso, deve ser levada em consideração a reverberação, que existe em praticamente todos os ambientes reais. Para antenas, devem ser considerados fatores como:

perdas por acoplamento mútuo, desvanecimento e interferências, cuja relevância e intensidade variam com a frequência típica da onda e com os ambientes de operação.

### 2.2.1 Sistemas de rádio – transmissão eletromagnética

Em sistemas de rádio, a estimação da direção de chegada assume papel importante na busca por uma maior eficiência espectral e economia de potência. Além disso, um diagrama de radiação adaptativo melhora a relação sinal/ruído e minimiza o desvanecimento e a interferência co-canal [4, 44].

A eficiência espectral vem com o aumento do reuso da frequência, vez que é possível direcionar os lóbulos das antenas de transmissão/recepção somente para o usuário desejado, viabilizando técnicas de múltiplo acesso por divisão do espaço (SDMA ou *Spatial Division Multiple Access*). Da mesma forma, minimiza-se a potência necessária para transmissão, uma vez que as perdas por desvanecimento são diminuídas [4, 6, 20, 44] e os multipercursos minimizados pela filtragem espacial [20, 44].

### 2.2.2 Sistemas de áudio – transmissão mecânica

Sistemas que se baseiam na propagação de ondas mecânicas como as de áudio, de ultra-som, sonares e sismologia de reflexão poderiam ser otimizados por meio do uso de técnicas de estimação DOA e arranjo de sensores. Os principais benefícios da adoção desse tipo de técnica para as ondas mecânicas estão na maximização da relação sinal-ruído (*SNR*), na rejeição do ruído espacial e na estimação da direção de chegada da onda mecânica. A estimativa de posição das fontes geradoras de ondas mecânicas tem especial importância nos sismógrafos, sonares e equipamentos médicos como ultrasonógrafo.

Nesta dissertação, as ondas mecânicas de interesse são as perturbações que resultam em ondas sonoras. Neste caso, uma das principais dificuldades na modelagem do sistema são as interferências por reverberação, causadas por múltiplas reflexões. Além disso, a

banda larga dos sinais sonoros também é considerada como um entrave para o uso de algumas técnicas. Tais limitações devem ser levadas em consideração e sua influência nos resultados mensurada por meio de testes em protótipos.

## 2.3 ARRANJOS DE SENSORES

Arranjos de sensores têm despertado interesse, na comunidade científica nos usuários que usam aplicações com arranjos, já que permitem que sistemas façam uso adaptativo da diversidade espacial. Entre os potenciais ganhos com a utilização desse tipo de elemento estão: *beam steering*, *beamforming*, *steering vector*, *null beamforming*, *diversity combining*. Há ainda, na literatura, indícios de que é possível um superganho do arranjo [21].

Existem muitas geometrias passíveis de utilização para arranjo de sensores, entre as principais estão a linear uniforme, a retangular e a circular. A geometria circular se mostra mais adequada quando se deseja evitar a ambiguidade de sinais que são oriundos de ângulos maiores que  $90^\circ$  ou menores que  $-90^\circ$ , pois a resolução para o arranjo linear é de apenas  $180^\circ$ . Comparativamente, o arranjo circular precisa de menos elementos mantendo-se a resolução de  $360^\circ$ . Nessa seção, serão apresentadas duas geometrias: o arranjo linear e o retangular.

### 2.3.1 Arranjo linear uniforme

Para este tipo de arranjo, os elementos estão igualmente espaçados de uma distância  $d$  formando um segmento de reta. Dessa maneira, há apenas um ângulo de direção de chegada a considerar: o azimute. O ângulo de azimute é definido pela interseção entre a reta perpendicular ao eixo do arranjo e a direção do sinal incidente considerado. A Figura 2.1

traz a representação dessa geometria, apresentando o azimute  $\theta$ , em que  $s(t)$  é o sinal incidente e  $x_i(t)$  é a saída do  $i$ -ésimo sensor.

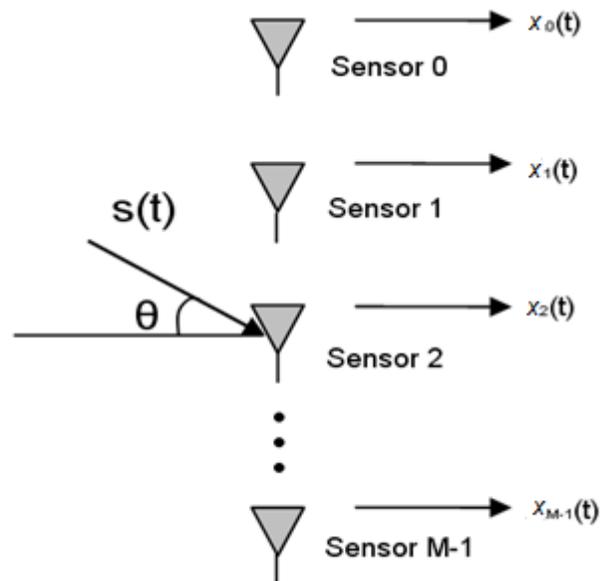


Figura 2.1: Sinal  $s(t)$  incidente num ULA de  $M$  elementos, segundo o azimute  $\theta$

Quando uma onda plana, de um sinal banda estreita, vinda de uma determinada direção incide sobre um arranjo, cada elemento transdutor é sensibilizado por versões defasadas da mesma. A defasagem pode ser compreendida a partir da Figura 2.2, na qual  $s$  é a diferença de comprimento do caminho percorrido pela frente de onda entre o sensor 0 e o sensor 1.

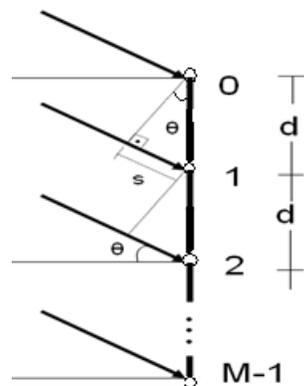


Figura 2.2: Geometria para frentes de onda planas

O sinal incidirá primeiramente no sensor 0 que será considerado como referência. O sinal percebido pelo sensor 1 terá uma defasagem temporal, ou retardo, em relação ao sensor 0, que depende da velocidade de propagação da onda, como pode ser visto na equação (2.1), na qual  $v$  é a velocidade de propagação da frente de onda,  $\Delta d$  é a distância percorrida pela onda entre o sensor 0 e o sensor 1 e  $t$  é o retardo.

$$\Delta t = \frac{\Delta d}{v} \quad (2.1)$$

A distância  $\Delta d$  pode ser calculada por geometria a partir da Figura 2.2. O atraso  $\Delta t$  é dado pela equação (2.2), na qual  $\theta$  é a direção do sinal e  $d$  é a distância entre os sensores.

$$\Delta t = \frac{\Delta d \cdot \text{sen}(\theta)}{v} \quad (2.2)$$

Essa defasagem no tempo pode ser convertida para defasagem angular de um sinal a partir da velocidade angular  $\omega$  na equação (2.3), onde  $f$  é a frequência da onda incidente.

$$\varphi = \omega \cdot \Delta t = 2\pi f \cdot \frac{\Delta d \cdot \text{sen}(\theta)}{v} \quad (2.3)$$

Finalmente, a defasagem angular pode então ser avaliada em termos do comprimento de onda  $\lambda = \frac{v}{f}$ , como na equação (2.4).

$$\varphi = 2\pi \cdot \frac{\Delta d \cdot \text{sen}(\theta)}{\lambda} \quad (2.4)$$

A defasagem apresentada na equação (2.4) é a do sensor 1 em relação ao sensor 0 para um único sinal incidente  $s(t)$  com um ângulo  $\theta$ . Como o sensor 2 também é espaçado de  $d$  do sensor 1, temos que a defasagem entre o sensor 0 e o sensor 2 é o dobro da primeira defasagem. Esse raciocínio pode ser estendido aos  $M$  sensores, resultando na equação

matricial em (2.5). Nessa equação, foi adicionada a contribuição de ruído gaussiano branco  $n(t)$  em cada elemento do arranjo.

As equações são válidas apenas para sinais  $s(t)$  de banda estreita (*narrowband*) e analíticos, i.e., do tipo  $s(t) = b(t)e^{j\omega_0 t}$  em que  $b(t)$  é limitado em banda com frequência máxima tal que  $\omega_0 \gg 2\pi f_n$  em que  $f_n$  é a frequência predominante (referência) do sinal.

$$\begin{bmatrix} x_0(t) \\ x_1(t) \\ x_2(t) \\ \vdots \\ x_{M-1}(t) \end{bmatrix} = \begin{bmatrix} 1 \\ e^{-j\varphi} \\ e^{-2j\varphi} \\ \vdots \\ e^{-(M-1)j\varphi} \end{bmatrix} s(t) + \begin{bmatrix} n_0(t) \\ n_1(t) \\ n_2(t) \\ \vdots \\ n_{M-1}(t) \end{bmatrix} \quad (2.5)$$

Quando existem vários sinais provenientes de diferentes direções, o vetor de defasagem é representado por uma matriz  $M \times K$ , na qual  $K$  é o número de sinais presentes. A equação 2.6 explicita o equacionamento resultante para  $K$  sinais.

$$\begin{bmatrix} x_0(t) \\ x_1(t) \\ x_2(t) \\ \vdots \\ x_{M-1}(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-j\varphi_0} & e^{-j\varphi_1} & & e^{-j\varphi_{K-1}} \\ e^{-2j\varphi_0} & e^{-2j\varphi_1} & & e^{-2j\varphi_{K-1}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-(M-1)j\varphi_0} & e^{-(M-1)j\varphi_1} & \dots & e^{-(M-1)j\varphi_{K-1}} \end{bmatrix} \cdot \begin{bmatrix} s_0(t) \\ s_1(t) \\ s_2(t) \\ \vdots \\ s_{K-1}(t) \end{bmatrix} + \begin{bmatrix} n_0(t) \\ n_1(t) \\ n_2(t) \\ \vdots \\ n_{M-1}(t) \end{bmatrix} \quad (2.6)$$

Alternativamente a relação acima pode ser representada na forma da equação (2.7).

$$\mathbf{X}(t) = \mathbf{A}(\varphi(\theta))\mathbf{S}(t) + \mathbf{N}(t) \quad (2.7)$$

A matriz  $\mathbf{A}(\theta)_{[M \times K]} = \mathbf{A}(\varphi(\theta))_{[M \times K]}$  é formada pelos vetores coluna  $\mathbf{a}(\theta_k)$ ,  $0 \leq k < K$ , os quais representam a resposta do arranjo em cada direção  $\theta_k$ , para cada sensor de 0 a  $M-1$ . Cada um desses vetores também é chamado de vetor volante (*steering vector*) ou vetor diretor e é apresentado na equação 2.8.

$$a(\theta_k) = \begin{bmatrix} 1 \\ e^{-j\varphi_k} \\ e^{-2j\varphi_k} \\ \vdots \\ e^{-(M-1)j\varphi_k} \end{bmatrix} \quad (2.8)$$

A equação 2.6 representa os sinais em tempo contínuo. Quando esses sinais são amostrados por meio de  $M$  sensores, em tempo discreto, os dados são representados por uma matriz  $M \times N$ , na qual  $N$  é o número de amostras do quadro considerado, em que  $t = n \cdot t_0$  e  $n$  varia de 1 até  $N$ , sendo  $t_0$  é o período de amostragem. Assim, ter-se-ia para o  $i$ -ésimo bloco de  $\mathbf{X}$  a expressão conforme equação (2.9):

$$\mathbf{X}_{[M \times N]}(i) = \mathbf{A}(\theta)_{[M \times K]} \mathbf{S}_{[K \times N]}(i) + \mathbf{N}_{(M \times N)}(i) \quad (2.9)$$

### 2.3.1.1 Ambigüidade na estimação de DOA

Para evitar que a equação 2.4 tenha mais de uma solução possível, deve-se garantir que o módulo do ângulo de defasagem seja menor ou igual a 180 graus ( $\pi$  radianos). Aplicando essa condição à equação 2.4 obtêm-se como resultado a equação 2.10. Como o módulo do *seno* é sempre menor que a unidade, chega-se à equação 2.11, versão espacial do teorema da amostragem de *Nyquist*.

$$d \leq \frac{\lambda}{2 \cdot |\text{sen}(\theta)|} \quad (2.10)$$

$$d \leq \frac{\lambda}{2} \quad (2.11)$$

Sendo assim, o ângulo da direção de chegada pode ser estimado a partir do defasamento entre os elementos do arranjo, desde que o espaçamento entre esses seja

menor que a metade do menor comprimento de onda de todos os sinais incidentes. Há ainda os casos em que se interessa conhecer tanto o ângulo de azimute como o de elevação. São exemplos os sonares e radares, para estes casos poder-se-ia utilizar a geometria de um arranjo planar.

### 2.3.2 Arranjo planar

A segunda geometria que apresentamos é o arranjo retangular uniforme (URA - *Uniform Rectangular Array*). Esse arranjo pode ser entendido como um conjunto de arranjos lineares uniformes, formando uma área retangular nos eixos  $x, y$  ou  $y, z$ . A Figura 2.3 apresenta esse arranjo no plano  $y, z$  e as duas distâncias relevantes para sua modelagem: a distância horizontal  $d_h$  e a distância vertical  $d_v$ .

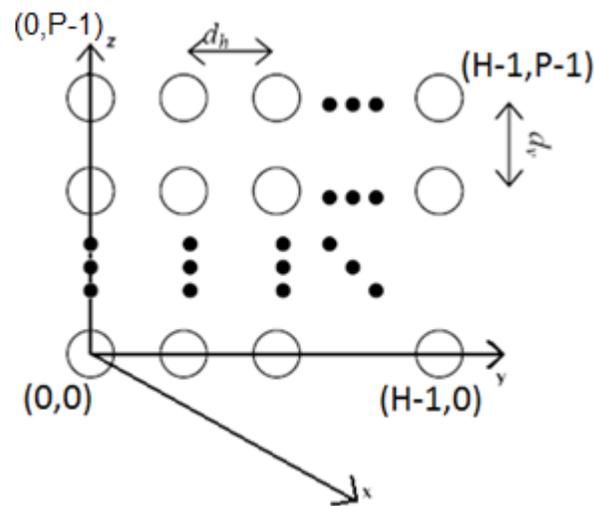


Figura 2.3: Geometria para frentes planas

No URA existem dois ângulos a serem considerados na direção de chegada: o ângulo de elevação  $\phi$  e o ângulo de azimute  $\theta$ . A elevação é medida a partir da projeção da direção do sinal sobre o plano  $x, y$  e o azimute é o ângulo formado entre essa projeção e o eixo  $x$ , como pode ser visto na Figura 2.4.

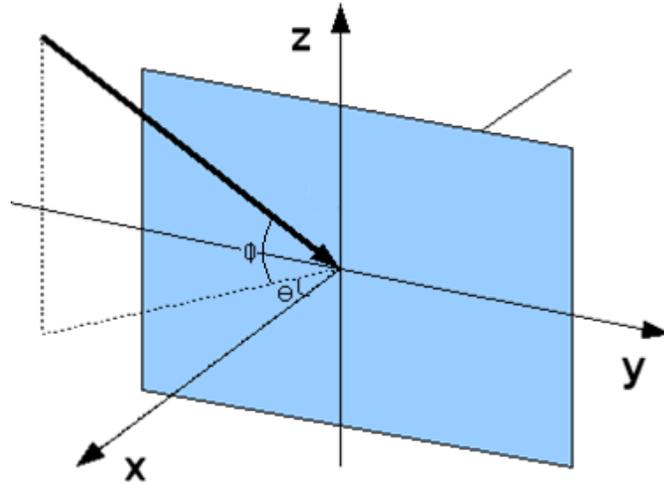


Figura 2.4: Frente de ondas (seta negra) incidindo sobre um arranjo planar uniforme

Para o cálculo de distâncias no arranjo retangular, utilizar-se-ão os vetores unitários na direção  $x$ , na direção  $y$  e na direção  $z$ . Esses são os versores,  $\mathbf{u}$ ,  $\mathbf{v}$ ,  $\mathbf{w}$ .

Assim, pode-se dizer que a direção de chegada  $DOA(\phi, \theta)$  e a distância dada entre duas coordenadas do vetor posição entre os sensores  $i$  e  $j$  –  $D(i, j)$  – são representadas, respectivamente, pelas equações (2.12) e (2.13).

$$DOA(\phi, \theta) = -\cos(\phi) \cdot \cos(\theta) \cdot \mathbf{u} - \cos(\phi) \cdot \sin(\theta) \cdot \mathbf{v} + \sin(\theta) \cdot \mathbf{w} \quad (2.12)$$

$$D(i, j) = i \cdot dh \cdot \mathbf{v} + j \cdot dv \cdot \mathbf{w} \quad (2.13)$$

Na equação 2.13,  $i$  e  $j$  denotam a posição do elemento no arranjo, na qual  $i$  varia entre 0 e o número de sensores  $H-I$  do arranjo na horizontal, e  $j$  varia entre 0 e o número de sensores  $P-I$  na vertical.

O produto escalar das equações 2.12 e 2.13 fornece a distância percorrida  $DP$  por cada sinal até o elemento na posição  $(i, j)$  no arranjo. Essa operação tem como resultado a equação 2.14.

$$DP(i, j) = \cos(\phi) \cdot \sin(\theta) \cdot i \cdot dh + \sin(\theta) \cdot j \cdot dv \quad (2.14)$$

De posse da distância percorrida por cada sinal até cada sensor e utilizando a equação 2.1, encontra-se de forma semelhante à equação 2.3, a defasagem angular do URA mostrado na equação 2.15.

$$\varphi(i, j) = \omega \cdot t = 2 \cdot \pi \cdot f \cdot \frac{DP(i, j)}{v} \quad (2.15)$$

Por fim, a defasagem em termos do comprimento de onda do sinal incidente no arranjo retangular é apresentada na equação 2.16.

$$\varphi(i, j) = \frac{2\pi}{\lambda} \cdot (\cos(\phi) \cdot \text{sen}(\theta) \cdot i \cdot d_h + \text{sen}(\theta) \cdot j \cdot d_v) \quad (2.16)$$

Dado o exposto, fica comprovado que a defasagem depende tanto do ângulo de azimute quanto do de elevação. A matriz diretora  $A(\theta, \phi)$  é apresentada na equação (2.17).

$$\mathbf{A}_k(\theta, \phi) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-j\varphi_{(0,0)}} & e^{-j\varphi_{(1,0)}} & & e^{-j\varphi_{(H-1,0)}} \\ e^{-2j\varphi_{(0,1)}} & e^{-j2\varphi_{(1,1)}} & & e^{-2j\varphi_{(H-1,1)}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-(H-1)j\varphi_{(0,P-1)}} & e^{-(H-1)j\varphi_{(1,P-1)}} & \dots & e^{-(H-1)j\varphi_{(H-1,P-1)}} \end{bmatrix} \quad (2.17)$$

Logo, a matriz pertence ao espaço vetorial complexo de dimensão  $H \times P$  -  $\mathbb{C}^{H \times P}$ . Analogamente à equação 2.7, a matriz de saída  $\mathbf{X}(t)$  é representada na equação 2.18, em que  $M = H \times P$ :

$$\mathbf{X}(t) = \mathbf{A}(\theta, \phi)_{M \times K} \mathbf{S}(t)_{K \times N} + \mathbf{N}(t)_{M \times N} \quad (2.18)$$

## 2.4 DISTÂNCIA ENTRE SENSORES

Uma vez que quanto maior o espaçamento entre os sensores, maior a defasagem entre os sinais percebidos por elementos distintos do arranjo, a distância entre os sensores constitui importante parâmetro na estimação da direção de chegada.

É nominada a grandeza frequência espacial:  $f_t = \frac{d_{nyquist}}{d} \frac{1}{\lambda} [\frac{ciclos}{m}]$ . Esta define uma medida de quanto uma determinada estrutura se repete para uma unidade de distância. Para um arranjo arbitrário de sensores igualmente espaçados, interpretamos a frequência espacial como a frequência em que o arranjo atingirá a máxima sensibilidade a uma onda incidente com comprimento de onda  $\lambda$ , considerado que o meio de propagação é um ambiente isotrópico.

Para o caso sonoro, a velocidade do som no meio aéreo é  $v = 340 \frac{m}{s}$ , tem-se pela versão espacial do Teorema de *Nyquist* que a distância entre os sensores deverá ser  $d \leq \frac{\lambda}{2}$ , ou seja,  $d \leq \frac{v}{2f}$ . Por hipótese, se for considerado que a frequência máxima do espectro sonoro é  $f = 1 \text{ kHz}$ , então tem-se  $d \leq 17 \text{ cm}$  (vide equação 2.11).

Deste modo, a frequência espacial de um arranjo separado por  $17 \text{ cm}$  é de  $1 \text{ kHz}$  e sua sensibilidade máxima ocorre para os sons recebidos em torno de  $1 \text{ kHz}$ . Se variarmos a distância de separação também estamos variando a frequência para a qual o arranjo tem sua melhor captação. Por exemplo, considere dois espaçamentos distintos entre os microfones, de  $10$  e  $5 \text{ cm}$ . A frequência espacial para o arranjo em  $10 \text{ cm}$  é  $f_t = \frac{17}{10} \cdot 1 \text{ kHz} = 1,7 \frac{ciclos}{m}$  e para o arranjo de  $5 \text{ cm}$  é  $f_t = \frac{17}{5} \cdot 1 \text{ kHz} = 3,4 \frac{ciclos}{m}$ .

Como forma de ilustrar o comportamento da seletividade do arranjo em função da separação dos sensores, é apresentada na Figura 2.5 uma simulação para um arranjo de oito sensores, com o espaçamento entre os sensores variando de 1 *cm* a 17 *cm*.

Nota-se como o pico de potência espectral se dispersa à medida que se afasta da frequência espacial ideal para o arranjo. O sinal incidente é um tom senoidal de 1 *kHz* e a relação sinal-ruído mantida em 20 *dB*.

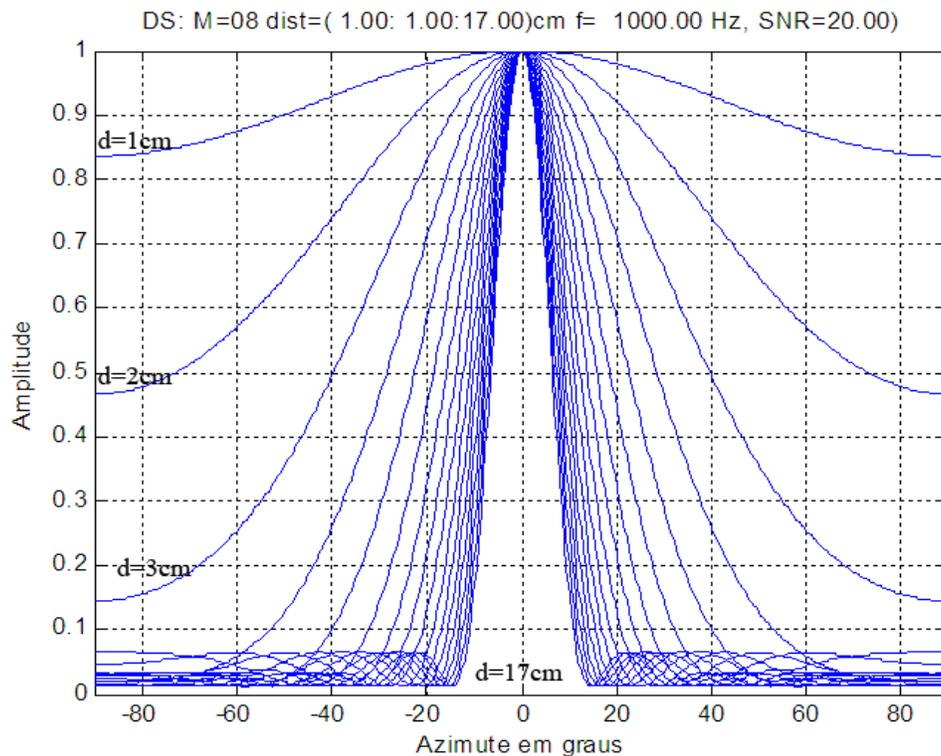


Figura 2.5 Largura do pico de potência em função da separação entre sensores

Do cálculo anterior, tem-se que a distância ótima entre sensores para um tom recebido na frequência  $f = 1 \text{ kHz}$  é de aproximadamente 17 *cm*. Note como os lóbulos de potência espectral apresentam maior dispersão, menor convergência e menor seletividade à medida que a distância utilizada entre os sensores diminui.

Em seguida, apresentamos na Figura 2.6, para a mesma simulação, a largura do pico de potência espacial espectral, tomados como referência os pontos em que os valores da

potência assumem 90% da potência máxima, como forma de medir a dispersão da estimação da direção de chegada à medida que a distância entre os sensores aumenta.

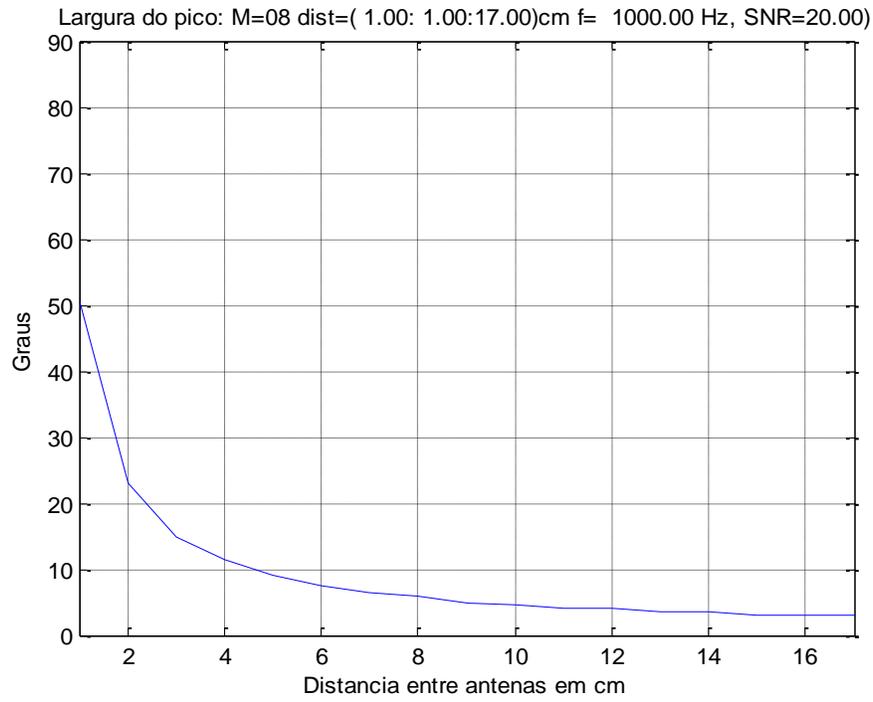


Figura 2.6 Largura do pico máximo em função da separação dos elementos do arranjo

Nota-se que a seletividade do estimador decai exponencialmente à medida que aproximamos os sensores, mantendo-se fixa a frequência do emissor ( $f = 1 \text{ kHz}$ ). Este fato reforça a necessidade de escolher o espaçamento adequado entre sensores sob o risco da estimação da direção de chegada não propiciar a seletividade desejada.

Para o mesmo conjunto de dados, apresentamos a direção de chegada relacionada a cada espaçamento entre sensores, com  $d$  variando desde 1 até 17  $\text{cm}$ . Deste modo, obtém-se o gráfico da Figura 2.7 da qual nota-se que, em condições ideais, a estimativa da direção de chegada permanece constante e praticamente independe do espaçamento entre sensores. Este é um comportamento esperado visto que os sensores utilizados são isotrópicos.

Apesar disso, os sensores utilizados na prática não são perfeitamente isotrópicos nem omnidirecionais. Isto pode ocasionar diferenças entre as potências recebidas pelos diversos ângulos de incidência no arranjo. Neste trabalho, consideramos que a potência recebida pelo arranjo é constante em todas as direções.

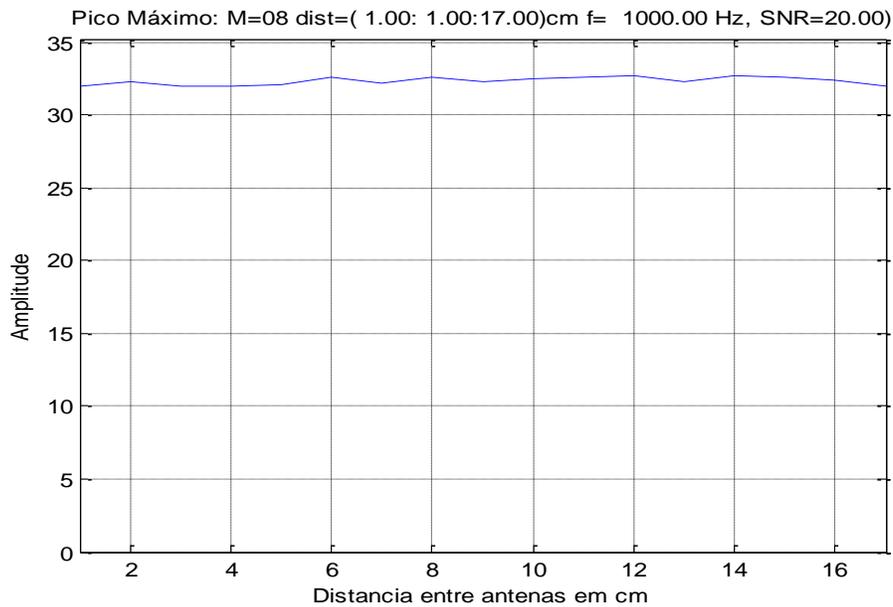


Figura 2.7 Pico máximo de potência obtido para diferentes distâncias entre sensores

## 2.5 A BANDA DE FREQUÊNCIA DOS SINAIS

Uma variável de grande importância na estimação da direção de chegada para aplicações em ambiente real é a banda de frequência do sinal. Para um sinal sonoro ou de voz, tem-se que os componentes espectrais de maior potência se encontram entre  $100\text{ Hz}$  e  $5\text{ kHz}$ , aproximadamente. Assim, calculando-se, para um trecho de sinal de voz de aproximadamente  $0,2\text{ s}$  para o verbete de voz “ou”, têm-se o espectro de frequências como indicado na Figura 2.8 aproximadamente dado por:

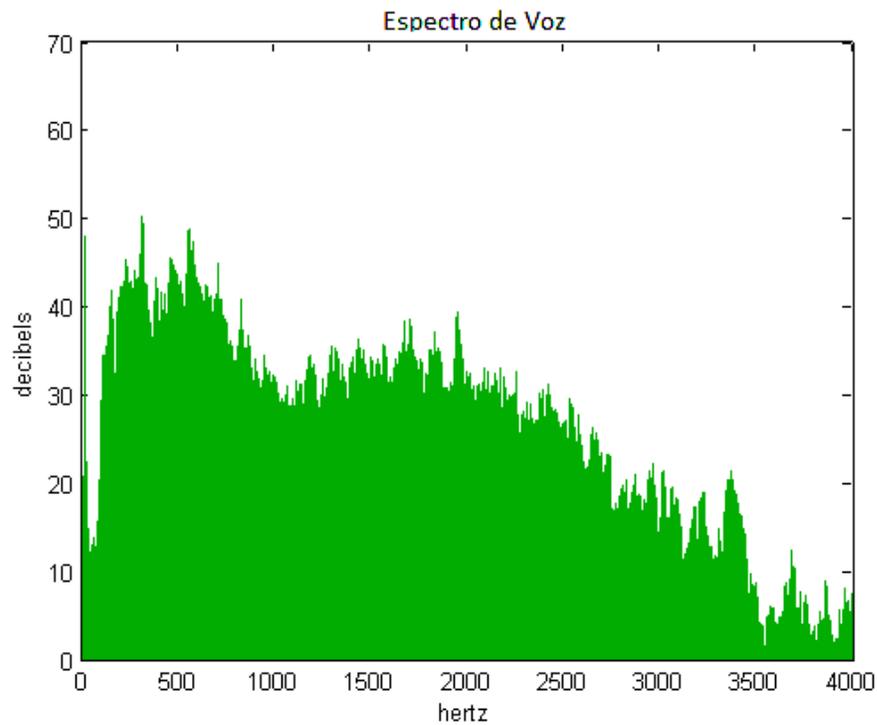


Figura 2.8 Resposta em frequência de um verbete de voz

Enquanto que, para um sinal senoidal de 200 Hz, tem-se o espectro de frequências aproximado pelo gráfico da Figura 2.9:

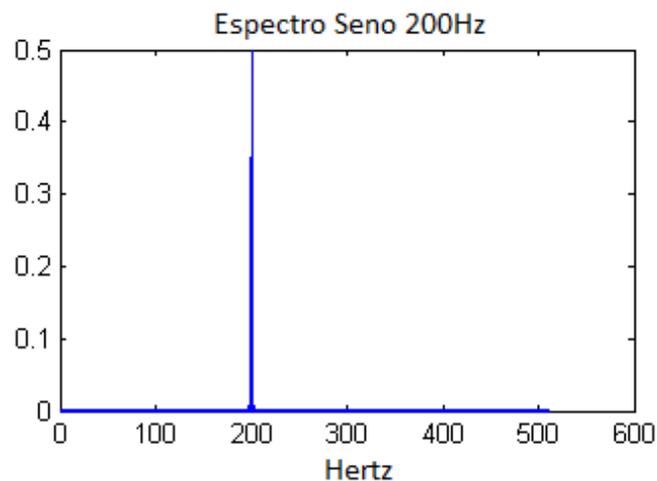


Figura 2.9 Resposta em frequência de um seno de 200 Hz

Como será descrito no Capítulo 3, a maioria dos métodos de estimação da direção de chegada têm a premissa de que o sinal recebido pelo arranjo de sensores é de banda

estreita. Ou seja, para se estimar a direção de chegada, deve-se considerar uma frequência de referência  $f_c$  sobre a qual será estimada a função de densidade de potência espacial.

Já a implementação do método SRP-PHAT proposta, por realizar o cálculo da correlação cruzada entre os sinais no domínio da frequência, leva em consideração todo o espectro de frequência do sinal recebido, e por isso, é considerado um método de estimação banda-larga (*wideband*).

Para ambos os métodos banda-estreita e banda-larga, ainda há uma importante restrição a ser considerada no desenho do *hardware* de aquisição: a resposta de fase do circuito de condicionamento e aquisição do sinal.

Para a situação ideal, teríamos um circuito com resposta de fase linear em toda a banda passante do sinal. Isto é, para todas as frequências do sinal recebido, o circuito introduz a mesma defasagem. Para algoritmos de estimação de banda-estreita, ainda é desejável, e necessário, que o sinal possua uma banda passante relativamente estreita se comparada com a frequência de referência  $f_c$  [9, 28, 32, 34]. Esta situação está ilustrada na figura 2.10.

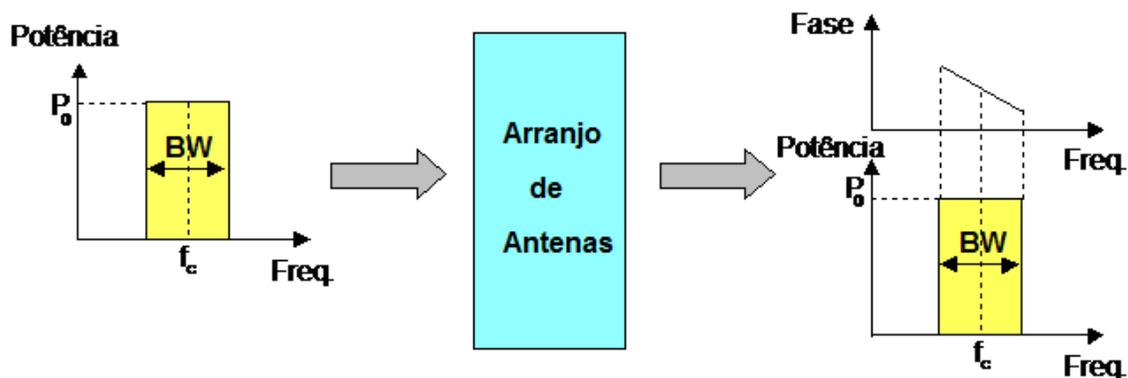


Figura 2.10 Resposta em frequência teórica.

Ou seja, são características dos estimadores banda-estreita:

- Banda estreita:  $BW \ll f_c$ ;

- A densidade espectral de potência não é alterada;
- O deslocamento de fase não é alterado.

Embora para ambos os métodos de banda-estreita e banda-larga seja desejável que se tenha um circuito de aquisição de fase completamente linear, devido as imperfeições dos modelos de projeto e dos próprios componentes, é impossível conceber um circuito com tal característica. Para uma situação real, ter-se-ia o que está mostrado na figura 2.11.

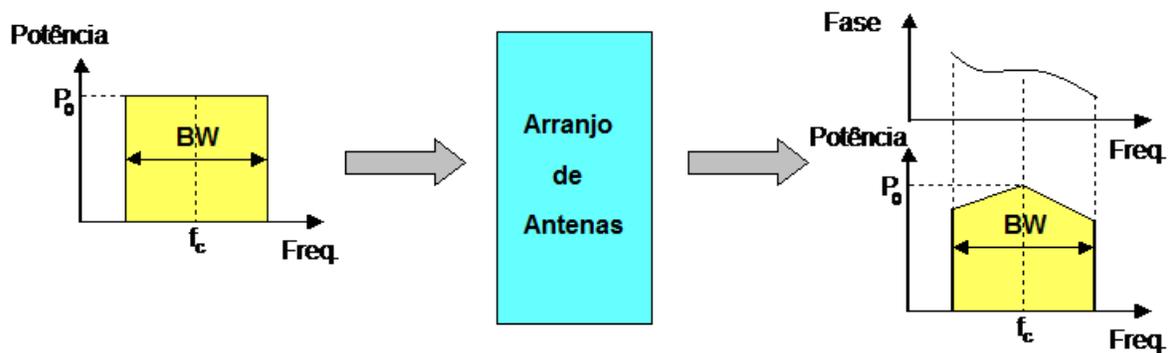


Figura 2.11 Resposta em frequência real

Para ilustrar de que forma este comportamento afeta o resultado dos algoritmos de estimação DOA, considere um sinal sonoro banda-larga centrado em  $1\text{ kHz}$ . Por hipótese, utiliza-se um arranjo de oito microfones separados por  $1\text{ cm}$  de distância e se estipula o erro máximo de  $1^\circ$  na estimação DOA.

Assim,

$$f_c = 1\text{ kHz}$$

$$\frac{\lambda}{2} = 17\text{ cm.} \quad (2.19)$$

Tempo para percorrer a distância entre dois elementos adjacentes do arranjo:

$$T_a = \frac{0,07[m]}{340[\frac{m}{s}]} = 0,206\text{ ms.}$$

Para uma variação de fase de  $1^\circ$ , ter-se-ia:

$$2\pi[(fc + BW) - (fc - BW)].Ta = \frac{\pi}{180} \quad (2.20)$$

Para uma precisão de  $1^\circ$ ,  $BW = 6,7 \text{ Hz}$ , em que  $BW$  é a banda do sinal. Para uma precisão de  $10^\circ$ ,  $BW = 67 \text{ Hz}$ . Ou seja, se a frequência central da banda do sinal é de  $1 \text{ kHz}$ , para um erro de  $1^\circ$  na fase, a banda de voz deveria estender-se apenas de  $993,3 \text{ Hz}$  a  $1006,7 \text{ Hz}$  o que tornaria o som de voz recebido absolutamente ininteligível.

Como demonstrado nos capítulos subseqüentes, este problema pode ser resolvido, em parte, por meio da utilização de um filtro digital de banda estreita e de fase linear após a digitalização dos sinais e antes da execução dos algoritmos de DOA de banda-estreita. Esta será a solução adotada quando queremos utilizar os métodos de banda-estreita para cálculo da direção de chegada de sinais banda-larga [15, 32, 46].

### 3 MÉTODOS DE ESTIMAÇÃO DA DIREÇÃO DE CHEGADA

Nesta seção, alguns métodos utilizados para a estimação da direção de chegada de um sinal serão apresentados. Os métodos serão aplicados considerando a geometria linear para a disposição dos sensores (ULA).

Os métodos apresentados serão o DS [3], CAPON [9], MUSIC [34], ESPRIT [28] e SRP-PHAT [6, 7, 16, 17, 26, 25, 29, 37, 45]. Os dois primeiros são baseados na estimação da densidade espacial de potência (PSD – *Power Spatial Density*). Eles se caracterizam por estimar a distribuição de potência do espectro espacial dos sinais e, a partir daí, estimar a direção de chegada como aquela em que há a maior potência espectral. Os dois seguintes são baseados na teoria de subespaços e utilizam-se da auto-estrutura das matrizes de correlação dos sensores e dos sinais para a estimação da direção de chegada. O último, SRP-PHAT, estima a DOA usando a correlação temporal.

Para apresentar esses métodos, definimos  $\mathbf{R}_{ss}$  como a matriz de autocorrelação dos sinais, apresentada na equação 3.1, na qual o operador  $\dagger$  representa o conjugado do transposto da matriz,  $E\{\cdot\}$  representa o operador valor esperado e  $\mathbf{S}$  representa o sinal incidente na variável de tempo discreta  $n$ .

$$\mathbf{R}_{ss} = E\{\mathbf{S}(n)\mathbf{S}(n)^\dagger\} \quad (3.1)$$

A matriz  $\mathbf{R}_{ss}$  surge no cálculo da matriz de autocorrelação  $\mathbf{R}_{xx}$  dos sinais recebidos pelos microfones apresentada na equação 2.7.

$$\mathbf{R}_{xx} = \mathbf{X}(n)\mathbf{X}(n)^\dagger \quad (3.2)$$

Desenvolvendo a equação (3.2), chegamos às equações 3.3 e 3.4, nas quais  $\sigma^2$  denota a potência do ruído. Para tanto, considera-se que  $\mathbf{S}$  e  $\mathbf{N}$  não são espacialmente correlacionados e que  $\mathbf{N}$  é ruído branco.

$$\begin{aligned} \mathbf{R}_{XX} &= E\{[A(\theta)S(n) + N(n)][A(\theta)S(n) + N(n)]^\dagger\} \\ &= \mathbf{A} E\{S(n)S(n)^\dagger\} \mathbf{A}^\dagger + \sigma^2 \mathbf{I} \end{aligned} \quad (3.3)$$

$$\mathbf{R}_{XX} = \mathbf{A} R_{SS} \mathbf{A}^\dagger + \sigma^2 \mathbf{I} \quad (3.4)$$

### 3.1 DS- ATRASO E SOMA

Nesse método denominado *beamformer* tradicional ou clássico [3], os sinais são defasados e subseqüentemente somados de forma a avaliar em que direção a soma das correlações das saídas é máxima. Esse é o motivo da denominação DS, do inglês *Delay-and-Sum*. A estimação é feita a partir da potência total da soma dos sinais dos sensores, representada por  $y(n)$ . Essa soma é ponderada pelo conjugado do transposto do vetor diretor, conforme equação (2.8). O resultado é o produto escalar apresentado na equação 3.5.

$$P_{DS} = E\{|y(n)|^2\} = E\{|\mathbf{a}(\theta)^\dagger \mathbf{X}(n)|^2\} \quad (3.5)$$

Expandindo essa equação e usando a equação 2.7, obtemos a equação 3.6.

$$P_{DS} = \mathbf{a}(\theta)^\dagger \mathbf{R}_{XX} \mathbf{a}(\theta) \quad (3.6)$$

Os picos de potência da equação 3.6 são as estimativas dos ângulos de direção de chegada. Dessa forma, o método consiste em variar  $\theta$  em passos discretos entre  $\frac{\pi}{2}$  e  $-\frac{\pi}{2}$ , buscando pelos máximos da equação 3.6.

O DS possui a vantagem de ser simples, exigindo um baixo custo computacional e permitindo inferir o número de sinais presentes pela quantidade de picos encontrados. No

entanto, oferece vulnerabilidade com relação à largura dos picos encontrados e aos lóbulos secundários, o que resulta numa baixa resolução.

Para ilustrar as características do método DS, são apresentadas algumas simulações:

### 3.1.1 Simulação da densidade de potência espectral para um sinal incidente

- $M = 8$  sensores
- $d = 0,5 \lambda$ ,  $d$  é o espaçamento entre os microfones em relação ao comprimento de onda
- Assinatura espacial do sinal gerado =  $23^\circ$
- Frequência do sinal gerado =  $1 \text{ kHz}$
- Amplitude do sinal gerado = 1
- Passo angular =  $0,5^\circ$
- Frequência de amostragem =  $3 \text{ kHz}$
- Número de amostras = 100.000
- Não existe ruído

O gráfico a seguir apresenta o resultado do estimador para as condições dadas:

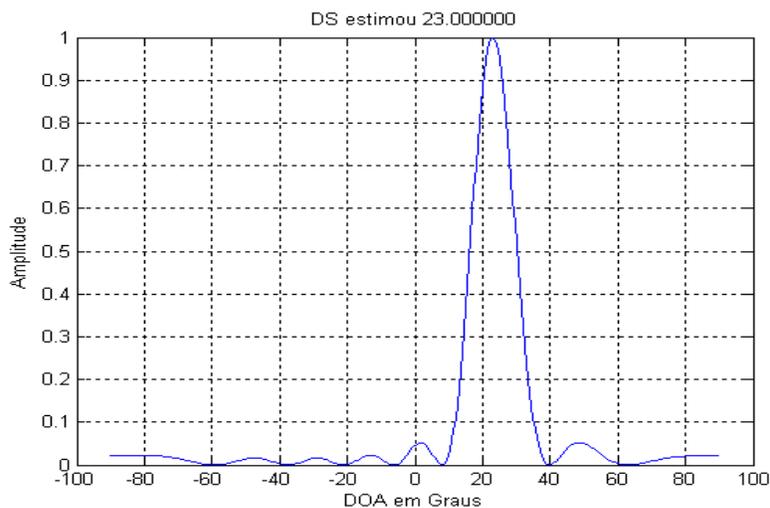


Figura 3.1 – DOA para um sinal de incidência

Como os valores candidatos para a estimação da direção de chegada são discretizados e avaliados em passos de  $0,5^\circ$ , o estimador não é capaz de diferenciar uma DOA de, por exemplo,  $23,05^\circ$ . Isto é, um aumento de resolução implica em aumento de complexidade.

### 3.1.2 Simulação de filtragem espacial pelo método DS

Nesta simulação foram utilizados três sinais senoidais superpostos com direções de incidência de  $30^\circ$ ,  $45^\circ$  e  $60^\circ$  e mesma potência, com vistas a simular um ambiente com interferências. Neste caso, utilizou-se a fonte de  $30^\circ$  como sinal desejado e aplicou-se a filtragem espacial sobre o sinal composto de forma a avaliar a capacidade de direcionar o lóbulo de recepção das antenas para uma direção desejada.

Para o ambiente de simulação e ondas senoidais, o resultado da Figura 3.2 demonstra que a filtragem espacial foi de fato realizada e, a menos do sinal, o sinal de interesse é recuperado. Apesar disso, a filtragem oferecida pelo método DS, por utilizar somente os atrasos estimados para cada sinal, é vulnerável à interferentes próximos e apresenta baixa seletividade angular.

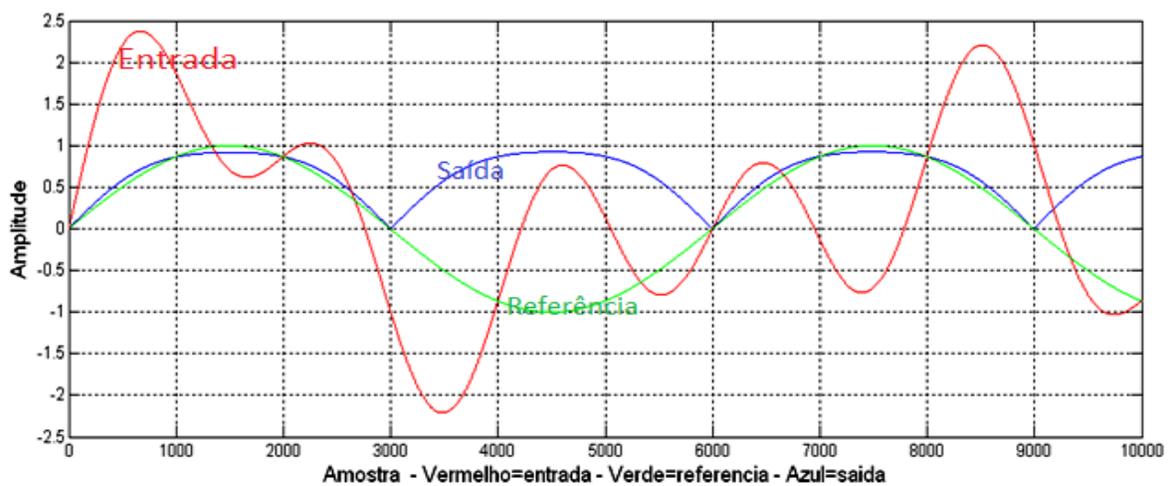


Figura 3.2 – Aplicação da filtragem espacial na recuperação de um sinal incidente em  $30^\circ$

Observa-se que a saída obtida é bastante próxima da resposta teórica esperada (referência). Nesta simulação, a saída foi tomada em módulo e por isso os sentidos negativos foram representados como sentidos positivos.

### 3.2 CAPON

J. CAPON [9] introduziu esse método de estimativa da direção de chegada utilizando a premissa de que deseja-se maximizar a potência na direção de chegada (*look direction*). Para tanto, é imposta uma restrição na função de densidade espectral espacial de potência, cujo resultado é visto na equação 3.7:

$$P_{CA} = \frac{1}{\mathbf{a}(\theta)^\dagger \cdot R_{XX}^{-1} \mathbf{a}(\theta)} \quad (3.7)$$

Analogamente ao DS, a busca da direção de chegada é efetuada iterativamente buscando-se o ângulo  $\theta$  que maximiza a expressão 3.7.

O CAPON oferece resolução melhor que o DS, mas possui um custo computacional adicional devido à necessidade de se inverter a matriz  $R_{xx}$ . Esse método também apresenta restrições quanto ao nível de correlação dos sinais interferentes com o sinal de interesse, devido à presença da matriz  $R_{XX}^{-1}$  no denominador da equação 3.7.

#### 3.2.1 Simulação do método CAPON com sinal interferente correlatado

- $M=8$  sensores;
- $d = 0,5 \lambda$ ;
- $ang = [37^\circ \ 60^\circ]$ , na qual  $ang$  é um vetor com o ângulo de incidência dos sinais gerados;
- $F = [1 \ 1]$ ,  $F$  o vetor de frequência dos sinais gerados em  $kHz$ ;
- $A = [1 \ 0,8]$ , vetor de amplitude dos sinais gerados;

- $f_s = 10$ , em que  $f_s$  é a frequência de amostragem  $kHz$ ;
- $N = 10.000$ ,  $N$  é o número de amostras utilizadas;
- Passo angular =  $0,5^\circ$ ;
- Ruído igual a zero

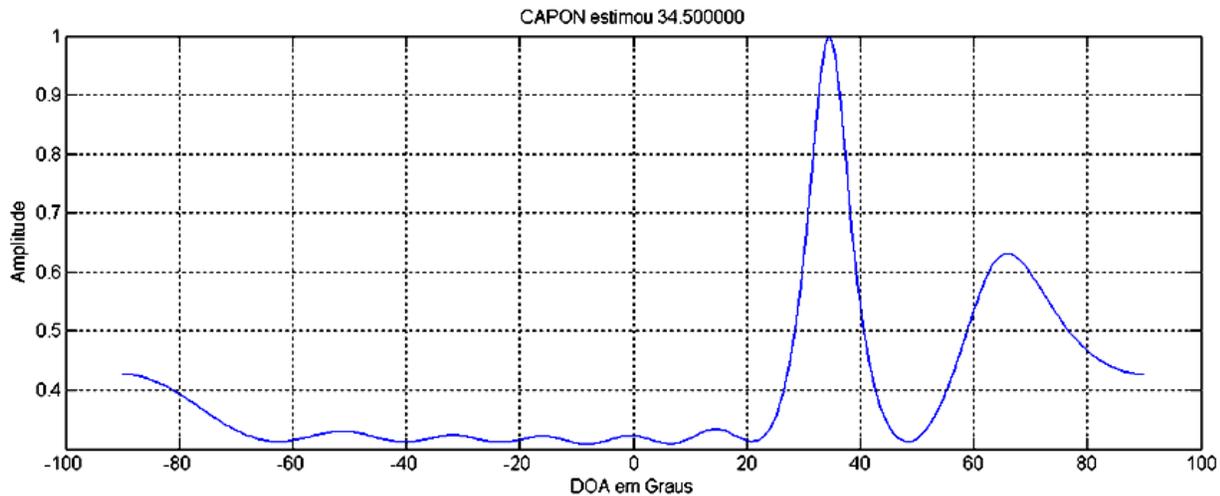


Figura 3.3 – Estimação de DOA usando o método CAPON para 2 sinais de incidência

Do resultado da simulação observa-se que devido à imposição de que a potência do sinal proveniente de outras direções que não a desejada deve ser nula, o método CAPON apresenta um pico de potência mais acentuado e menos disperso.

Contudo, pela presença de um sinal interferente a 60 graus, com 80% da amplitude do sinal senoidal desejado, o estimador apresenta um erro de  $2,5^\circ$  e detecta um lóbulo secundário em aproximadamente  $65^\circ$ .

### 3.2.2 Simulação do método CAPON com sinal interferente e ruído branco

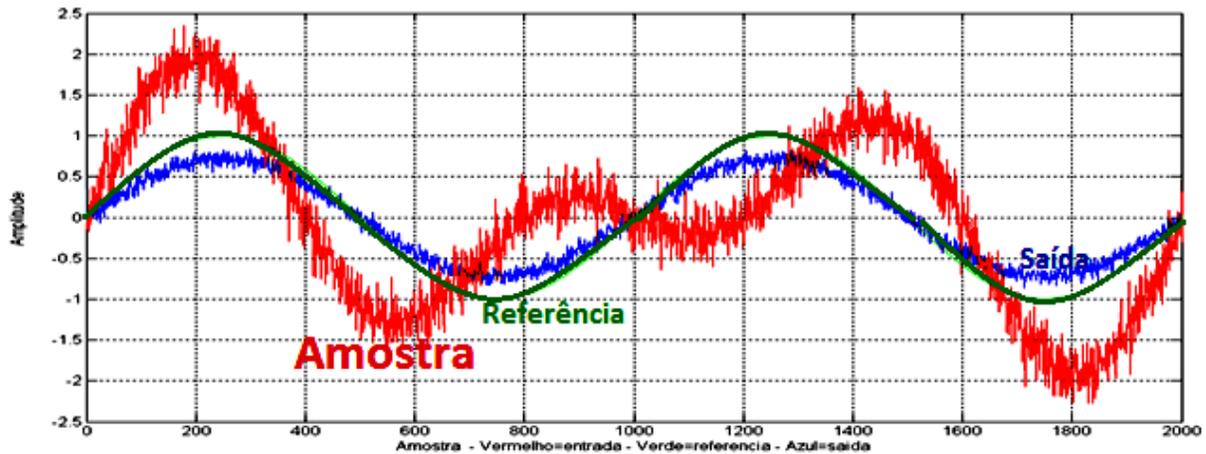


Figura 3.4 – Filtragem espacial baseada na estimação DOA com o emprego do estimador CAPON

Para esta simulação foram gerados sinais senoidais nos azimutes de  $30^\circ$  e  $45^\circ$ , adicionados de ruído branco com 20% da potência do sinal de interesse, com vistas a estudar o comportamento do método CAPON quando da presença de interferentes e ruído ambiente. A recuperação do sinal é realizada pela aplicação de um filtro espacial com pesos calculados para cada direção pelo método CAPON.

Note que, apesar do sinal interferente a  $45^\circ$ , a partir do sinal ruidoso pôde se recuperar o sinal desejado com fidelidade. O sinal interferente foi gerado com 80% da potência do sinal de referência, para simular uma reflexão em ambiente reverberante. Por meio da simulação, podemos comprovar que, para condições ideais de captação e tratamento dos sinais, o algoritmo de filtragem espacial de fato reduz a amplitude do ruído ambiente percebido melhorando a relação sinal-ruído, além de suprimir quase que totalmente o sinal interferente.

### 3.2.3 Simulação do método CAPON com sinal interferente não correlatado

Nesta simulação foi utilizado um sinal de voz com os dizeres “direção de chegada” como sinal de interesse – assinado espacialmente para  $30^\circ$  com  $f_s = 1 \text{ kHz}$  e

espaçamento de  $d = 0,5 \lambda$ . Os interferentes adicionados foram sinais sonoros de ruído ambiente, assinado espacialmente para  $0^\circ$ , e um sinal sonoro musical, assinado espacialmente para  $60^\circ$ , ambos com a mesma potência do sinal de interesse.

Esta simulação tem como objetivo avaliar, no caso de captação e tratamento ideal dos sinais pelo sistema de aquisição e condicionamento, e em ambiente real, qual seria a eficiência em se recuperar os verbetes de um interlocutor.

Para realizar a filtragem espacial, considera-se, como hipótese, que é conhecida a direção de chegada do sinal de interesse.

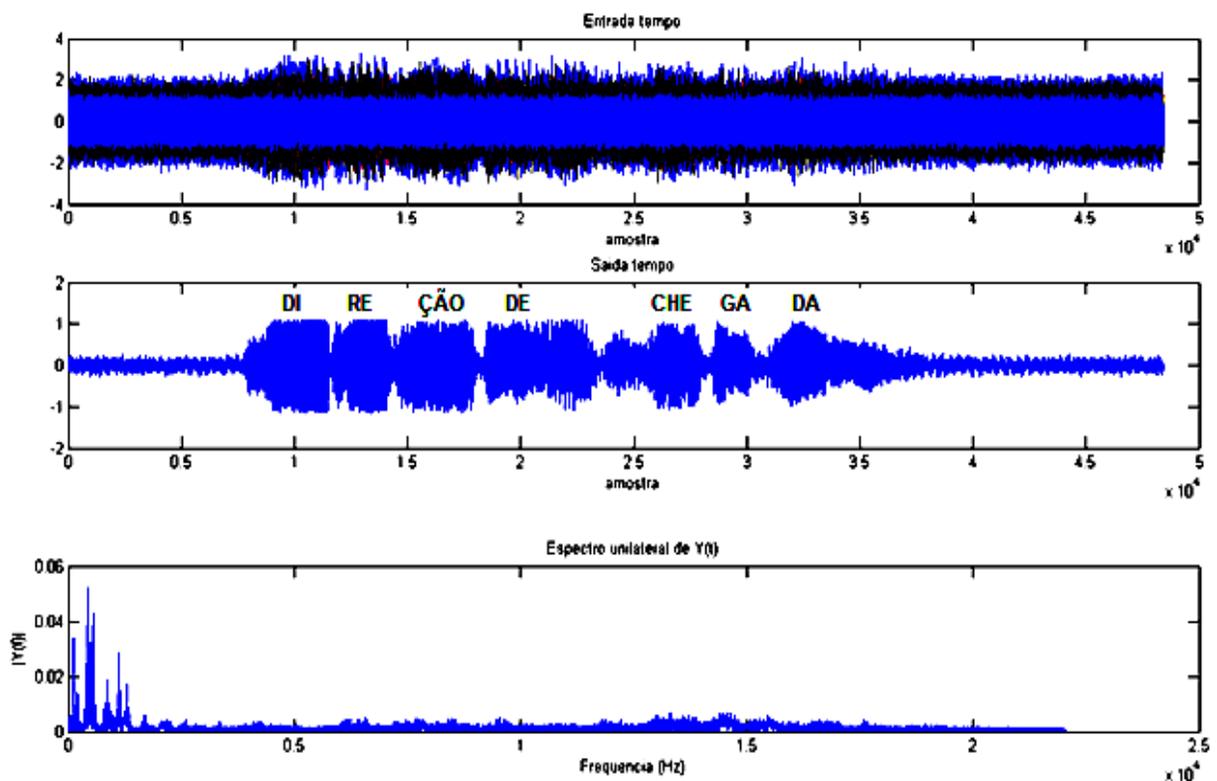


Figura 3.5 – Entrada com vários sinais e saída do método CAPON juntamente com a FFT da saída

Pode-se observar no gráfico da saída a seletividade do método, uma vez que este recupera praticamente o sinal íntegro original, de modo que quase todo o ruído introduzido foi rejeitado. Observando-se a *FFT* da saída pode-se ainda verificar que a potência do sinal

de saída concentra-se predominantemente entre  $0\text{ Hz}$  e  $1500\text{ Hz}$  (banda predominante de frequências de voz).

### 3.3 MUSIC

O método MUSIC [34] (*Multiple Signal Classification*) é baseado nas propriedades da auto-estrutura da resposta do arranjo, ou *Array Manifold*. Trata-se de uma decomposição em autovalores e autovetores da matriz de correlação cruzada do sinal nos subespaços de sinal e de ruído.

Na teoria de estimação de DOA, esse *Array Manifold* é caracterizado por todos os valores que o vetor diretor pode assumir para aquele arranjo. Estes valores determinam uma região no espaço quando variamos o ângulo  $\theta$ , conforme equação 2.8, de  $-\frac{\pi}{2}$  a  $\frac{\pi}{2}$ .

Supondo a incidência de ondas planas no arranjo e observando que todos os elementos do vetor dependem de uma única variável, então o vetor diretor é unidimensional. Assim, o *Array Manifold* é uma corda no espaço  $\mathcal{R}^M$  em que  $M$  é a quantidade de sensores do arranjo.

Uma característica do método MUSIC é que o ângulo  $\theta$  a ser estimado depende da frequência do sinal observado [25]. Isto o caracteriza como um método de banda estreita (*narrowband*).

Os sinais que são captados pelos sensores, de acordo com a equação 2.7, definem um subespaço vetorial  $S_x$ , cuja intersecção com a corda do *Array Manifold* permite a estimação da direção de chegada.

Para  $K$  sinais não correlacionados, os autovalores da matriz  $\mathbf{R}_{xx}$  da equação 3.2 são compostos por  $K$  autovalores referentes a cada sinal, e  $M - K$  autovalores referentes ao

ruído. Tal fato constitui uma limitação para o método, visto que se torna necessário saber com antecedência quantos sinais correlacionados e quantos não correlacionados existem, e esta informação, geralmente, não está disponível.

Os autovalores do ruído são geralmente menores que os autovalores do sinal para condições normais de operação. O espaço complexo de dimensão  $M$  ( $\mathbb{C}^M$ ) pode ser decomposto nos subespaços do ruído, representado por  $E_n$ , e do subespaço do sinal. O primeiro será composto pelos autovetores relacionados com os menores autovalores, enquanto o subespaço do sinal será composto pelos autovetores relacionados com os maiores autovalores.

Uma estimação para a matriz de correlação  $R_{xx}$  equação (3.2) é dada pela equação (3.8).

$$\hat{R}_{XX} = \frac{1}{N} \mathbf{X}(n)\mathbf{X}(n)^\dagger \quad (3.8)$$

Com essa estimativa e os autovalores associados, é possível decompor a matriz de correlação dos sinais para formar a estimativa do subespaço do ruído. De posse da matriz  $E_n$  de autovetores, a potência de densidade espectral pode ser calculada pela equação (3.9).

$$P_{MUS} = \frac{1}{\mathbf{a}^\dagger(\theta)\mathbf{E}_n\mathbf{E}_n^\dagger\mathbf{a}(\theta)} \quad (3.9)$$

De forma semelhante aos métodos já apresentados, a estimação da DOA é obtida encontrando-se o ângulo  $\theta$  que maximiza a equação anterior.

O algoritmo MUSIC permite alta resolução na estimação da direção de chegada, uma vez que o pico resultante da estimação é, para alguns casos, mais estreito que o dos métodos DS ou CAPON. Apresenta, também, robustez frente a ruído e possibilita a

estimação do número de sinais incidentes. No entanto, pelas operações de decomposição e inversão de matrizes, possui alto custo computacional.

Ainda, em função da estrutura matemática do algoritmo, este método apresenta grandes erros de estimação quando os sinais incidentes estão correlatados.

### 3.3.1 Simulação do método MUSIC com sinal interferente correlatado

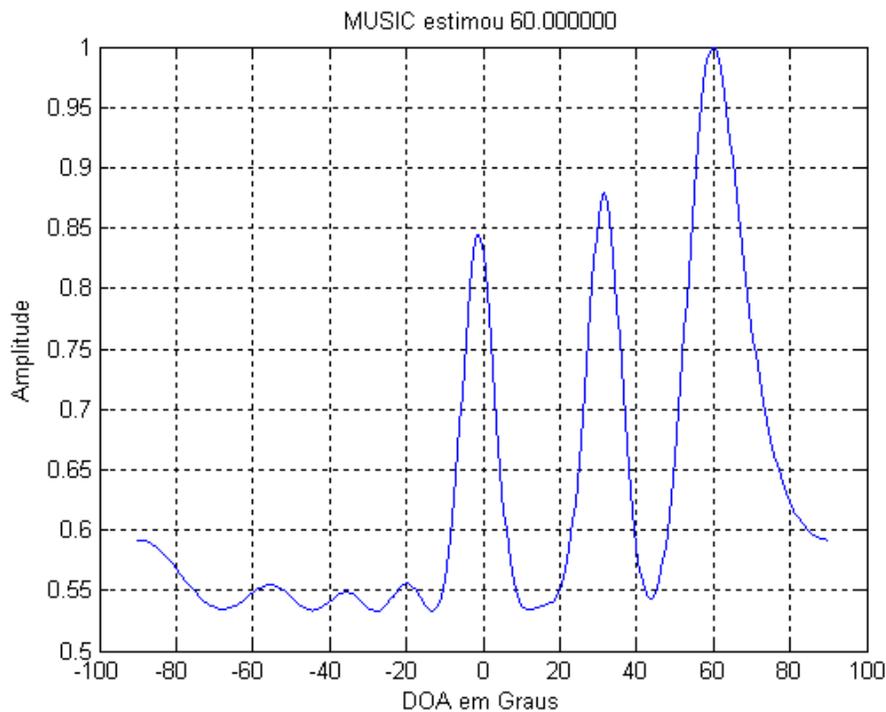


Figura 3.6 – Estimação de DOA método MUSIC para 3 sinais - voz em  $0^\circ$ , ruído em  $30^\circ$  e música a  $60^\circ$ . ( $SNR = 0,82dB$ )

Os resultados a seguir foram obtidos utilizando-se sinais incidentes de três trechos de áudio, para os quais a relação entre o sinal de interesse e os interferentes é de 10 dB. Têm-se um sinal sonoro de voz como de interesse e como sinais interferentes sinais sonoros assinados espacialmente para  $30^\circ$  e  $60^\circ$ , todos de banda-larga.

Nota-se que, apesar do sinal de interesse estar em  $0^\circ$  e possuir maior potência, o fato dos sinais interferentes possuírem alta correlação com o sinal de interesse faz com que

resultado da estimação da direção de chegada seja equivocado, uma vez que o pico de maior amplitude está em  $60^\circ$ .

Apesar da característica de banda-estreita do método MUSIC, em ambiente de simulação, sinais de banda-larga tem sua direção de chegada corretamente estimada por esse algoritmo pelo fato do atraso ser simulado na forma  $e^{-j\varphi}$ .

Desta forma, todos os componentes de frequência são igualmente defasados, o que, conforme será demonstrado, não acontece para dados adquiridos a partir de ambientes reais.

### 3.3.2 Simulação do método MUSIC para um único sinal

Nesta simulação foi dada assinatura espacial para um trecho sonoro e calculada a direção de chegada. Neste caso, o sinal incidia a  $30^\circ$ . Note que para um único sinal incidente tem-se uma ótima seletividade com o uso do método MUSIC.

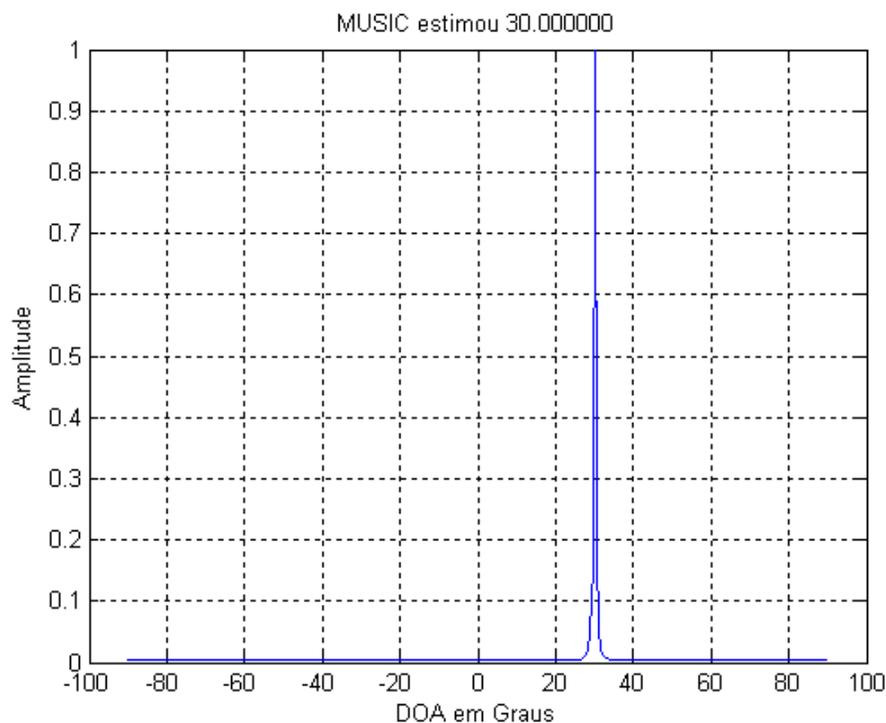


Figura 3.7 – Estimação de DOA com o método MUSIC para um sinal senoidal incidindo a  $30^\circ$

### 3.4 SRP-PHAT

Embora os métodos DS, CAPON, MUSIC e ESPRIT ofereçam resolução satisfatória para os casos apresentados até então, eles possuem uma limitação: são métodos de banda estreita (*narrowband*). Quando se considera que, na maioria dos casos práticos, os sinais abordados por este trabalho são banda-larga (*wideband*), é necessário investigar um método banda-larga para fins de implementação.

Nesta seção, descrevemos a estrutura do algoritmo SRP (*Steered Response Power*) e como seu funcionamento pode ser melhorado por um artifício matemático: a transformada de fase (*PHAT – Phase Transform*) [6, 7, 16, 17, 25, 26, 29, 37].

Considere o sinal recebido por dois sensores  $i$  e  $j$  pertencentes ao arranjo. Definimos a correlação cruzada para tempo contínuo entre esses dois sinais pela equação (3.10):

$$R_{x_i x_j}(\tau) = E[x_i^*(t)x_j(t + \tau)] \quad (3.10)$$

Em que  $x_i^*$  denota o complexo conjugado de  $x_i$ . Pela dificuldade prática em se utilizar o operador valor esperado  $E[\cdot]$ , usa-se um estimador para a função de correlação cruzada, que, para processos ergódicos, é dado por [29]:

$$\hat{R}_{x_i x_j}(\tau) = \int_{-\infty}^{\infty} x_i^*(t)x_j(t + \tau) dt \quad (3.11)$$

Para o caso em tempo discreto, definimos a correlação cruzada na forma da equação (3.12):

$$R_{x_i x_j}(\tau) = E[x_i^*(n)x_j(n + \tau)] \quad (3.12)$$

Novamente, utilizamos um estimador para esta correlação cruzada, que para este caso, pode ser dado por (3.13):

$$\hat{R}_{x_i x_j}(\tau) = \sum_{n=-\infty}^{n=\infty} x_i^*(n) x_j(n + \tau) \quad (3.13)$$

Para o caso discreto, temos perda de resolução na estimativa da correlação cruzada dado que  $\tau$  é um número inteiro. Esta imprecisão pode ser melhorada, em parte, com a interpolação da função de correlação cruzada, conforme será visto no Capítulo 4.

De maneira análoga aos métodos anteriores, queremos investigar a correlação cruzada entre as saídas dos elementos do arranjo, de forma a determinar para que atraso  $\tau$  a função de correlação cruzada atinge seu máximo. Por conhecer a velocidade da onda e a geometria do arranjo de sensores, e de acordo com o exposto em [29], o máximo da função de correlação cruzada nos dará uma estimativa para a direção de chegada.

Em algumas aplicações, podemos considerar uma filtragem dos sinais  $x_i(n)$  e  $x_j(n)$  antes do cômputo da correlação cruzada. Desta forma, teríamos [29]:

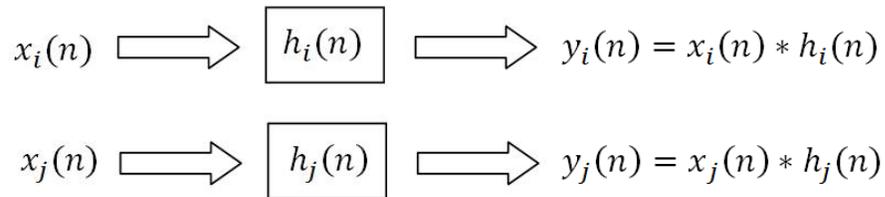


Figura 3.8: Pré-filtragem dos sinais em tempo discreto

Em que  $h_i(n)$  e  $h_j(n)$  representam a resposta impulsional de pré-filtros que podem ser aplicados antes do cálculo da correlação entre os sinais  $x_i(n)$  e  $x_j(n)$  e o símbolo  $*$  indica o operador convolução.

Definimos, então, a função de correlação cruzada generalizada (*GCC – Generalized Cross-Correlation*) [5, 29], em tempo discreto, como:

$$\bar{R}_{x_i x_j}(\tau) = \hat{R}_{y_i y_j}(\tau) = \sum_{n=-\infty}^{\infty} [y_i(n)^* y_j(n + \tau)]$$

$$= \sum_{n=-\infty}^{\infty} \left[ (x_i(n) * h_i(n))^* (x_j(n + \tau) * h_j(n + \tau)) \right] \quad (3.14)$$

Para o algoritmo SRP-PHAT, os pré-filtros realizam a transformada de fase (*PHAT* - *Phase transform*) que será detalhada mais adiante [29, 45].

Pela definição da Densidade Espectral de Potência (PSD - *Power Spectrum Density*) e utilizando o símbolo  $\mathcal{F}\{\}$  para indicar o operador transformada de Fourier, temos que:

$$PSD = \Phi_x(e^{j\omega}) = \mathcal{F}\{R_x(\tau)\} = \sum_{\tau=-\infty}^{\infty} R_x(\tau) e^{-j\omega\tau} \quad (3.15)$$

E então, conversamente:

$$R_x(\tau) = \mathcal{F}^{-1}\{\Phi_x(e^{j\omega})\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_x(e^{j\omega}) e^{j\omega\tau} d\omega \quad (3.16)$$

De modo similar, a densidade espectral de potência cruzada entre  $y_i$  e  $y_j$  é dada por:

$$\Phi_{y_i y_j}(e^{j\omega}) = \mathcal{F}\{R_{y_i y_j}(\tau)\} = \sum_{\tau=-\infty}^{\infty} R_{y_i y_j}(\tau) e^{-j\omega\tau} \quad (3.17)$$

E então temos:

$$R_{y_i y_j}(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{y_i y_j}(e^{j\omega}) e^{j\omega\tau} d\omega \quad (3.18)$$

De maneira análoga ao exposto anteriormente, o estimador para a função de correlação cruzada entre  $y_i$  e  $y_j$  é dado por (3.19):

$$\hat{R}_{y_i y_j}(\tau) = \sum_{n=-\infty}^{\infty} [y_i(n)^\dagger y_j(n + \tau)] \quad (3.19)$$

Denotando-se por  $H_i(e^{j\omega})$  e  $H_j(e^{j\omega})$ ,  $X_i(e^{j\omega})$  e  $X_j(e^{j\omega})$  as transformadas de Fourier dos termos de pré-filtragem e dos sinais, respectivamente, a função de correlação cruzada generalizada pode ser expressa por [5,29]:

$$\begin{aligned} \hat{R}_{y_i y_j}(\tau) &= \bar{R}_{x_i x_j}(\tau) = \mathcal{F}^{-1}\{\hat{\Phi}_{y_i y_j}(e^{j\omega})\} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} H_i(e^{j\omega}) X_i(e^{j\omega}) H_j(e^{j\omega})^* X_j(e^{j\omega})^* e^{j\omega\tau} d\omega \end{aligned} \quad (3.20)$$

No anexo I é demonstrado como se obter a equação (3.20) a partir da equação (3.15).

Usando-se a relação (3.21), passa-se a depender da variável  $k$ , que se torna o indexador de frequência:

$$\omega = \frac{2\pi}{N}k \quad (3.21)$$

E então (3.20) é reescrita da forma (3.22) utilizando o operador IDFT (*Inverse Discrete Fourier Transform*):

$$\hat{R}_{y_i y_j}(\tau) = IDFT\{\hat{\Phi}_{y_i y_j}(k)\} \quad (3.22)$$

Ambos pré-filtros  $H_i(k)$  e  $H_j(k)$  podem ser combinados em um único filtro,  $\gamma_{ij}(k)$ , deste modo:

$$\gamma_{ij}(k) = H_i(k)^* H_j(k) \quad (3.23)$$

Assim temos:

$$\hat{\Phi}_{y_i y_j}(k) = H_i(k) X_i(k) H_j(k)^* X_j(k)^* = \gamma_{ij}(k) X_i^*(k) X_j(k) \quad (3.24)$$

Finalmente, uma estimativa da correlação cruzada generalizada pode ser calculada pela Transformada Inversa de Fourier Discreta da equação 3.24 [29, 37], ou seja:

$$\bar{R}_{x_i x_j}(\tau) = \hat{R}_{y_i y_j}(\tau) = \frac{1}{N} \sum_{k=0}^{N-1} \gamma_{ij}(k) X_i(k) X_j^*(k) e^{j\frac{2\pi k\tau}{N}} \quad (3.25)$$

Para calcular a correlação no domínio da frequência, deve-se definir um número de pontos para a DFT de tal forma que a correlação circular seja igual a correlação linear.

A escolha do termo de pré-filtragem  $\gamma_{ij}(k)$  depende do tipo de ponderação que se deseja conferir ao espectro espacial. Para situações com baixa relação sinal-ruído, deseja-se uma função  $\gamma_{ij}(k)$  tal que esta atenuie principalmente as frequências que compõem o ruído. Para o caso de filtros na faixa de frequência de voz, por exemplo, pode-se rejeitar frequências nas bandas superiores a 6 kHz, o que reduz a interferência de ruídos de alta frequência como os emitidos por equipamentos eletrônicos.

Existem diferentes tipos de pré-filtros como ML (*Maximum Likelihood*), SCOT (*Smoothed Coherence Transform*), Roth e PHAT (*Phase Transform*). Para a aplicação considerada, o PHAT é o que apresenta as propriedades mais adequadas, conforme se verá adiante [10, 19, 45].

### 3.4.1 Phase Transform (PHAT)

Considere o sinal captado por dois sensores distintos,  $i$  e  $j$ , com um *delay* de  $D$  amostras entre eles. Neste caso,

$$x_i(n) = s(n) + n_i(n) \quad (3.26)$$

$$x_j(n) = s(n + D) + n_j(n) \quad (3.27)$$

Em que o termo  $n_i(n)$  modela o ruído ambiente captado pelos sensores, sendo, neste trabalho, considerado aditivo gaussiano branco (*AWGN*).

Aplicamos a transformada de Fourier nos sinais para obter:

$$X_i(e^{j\omega}) = S(e^{j\omega}) + N_i(e^{j\omega}) \quad (3.28)$$

$$X_j(e^{j\omega}) = S(e^{j\omega})e^{j\omega D} + N_j(e^{j\omega}) \quad (3.29)$$

Considerando  $\gamma_{ij}(\omega) = 1$  – vide (3.23) –, temos que a densidade espectral de potência é dada por (3.30):

$$\begin{aligned} \hat{\Phi}_{x_i x_j}(e^{j\omega}) &= \mathcal{F}\{\hat{R}_{x_i x_j}(\tau)\} = \sum_{\tau=-\infty}^{\infty} \hat{R}_{x_i x_j}(\tau) e^{-j\omega\tau} \\ &= \sum_{\tau=-\infty}^{\infty} \left( \sum_{n=-\infty}^{\infty} [x_i(n)^* x_j(n + \tau)] \right) e^{-j\omega\tau} \\ &= X_i^*(e^{j\omega}) X_j(e^{j\omega}) = [S(e^{j\omega}) + N_i(e^{j\omega})]^* [S(e^{j\omega})e^{j\omega D} + N_j(e^{j\omega})] \\ &= S^*(e^{j\omega})S(e^{j\omega})e^{j\omega D} + S^*(e^{j\omega})N_j(e^{j\omega}) + S(e^{j\omega})e^{j\omega D}N_i^*(e^{j\omega}) + \\ &N_i^*(e^{j\omega})N_j(e^{j\omega}) \end{aligned} \quad (3.30)$$

O termo  $S(\omega)S^*(\omega)$ , que corresponde a uma estimativa da densidade de potência espectral do sinal, será representado por  $\hat{\Phi}_{SS}$ . Deste modo,

$$\widehat{\Phi}_{x_i x_j}(e^{j\omega}) = \widehat{\Phi}_{SS} e^{j\omega D} + S^*(e^{j\omega}) N_j(e^{j\omega}) + S(e^{j\omega}) e^{j\omega D} N_i^*(e^{j\omega}) + N_i^*(e^{j\omega}) N_j(e^{j\omega}) \quad (3.31)$$

Considerando-se um quadro de  $N$  amostras do sinal  $e$ , supondo-se a premissa que os sinais de ruído  $N_i(e^{j\omega})$  e  $N_j(e^{j\omega})$  são descorrelatados entre si e também descorrelatados com o sinal de interesse, tem-se que para valores de  $N$  elevados:

$$\begin{aligned} S^*(e^{j\omega}) N_j(e^{j\omega}) &\approx 0 \\ S(e^{j\omega}) e^{j\omega D} N_i^*(e^{j\omega}) &\approx 0 \\ N_i^*(e^{j\omega}) N_j(e^{j\omega}) &\approx 0 \end{aligned} \quad (3.32)$$

Então, a função de correlação cruzada entre os sinais  $x_1$  e  $x_2$  é função somente do atraso entre eles e é dada por:

$$\widehat{R}_{x_i x_j}(\tau) = \mathcal{F}^{-1}[\widehat{\Phi}_{SS} e^{j\omega D}] = \widehat{R}_{SS}(\tau) * \delta(\tau + D) = \widehat{R}_{SS}(\tau + D) \quad (3.33)$$

Com base nesse resultado será estimado o atraso  $\tau$ , em número de amostras, que maximiza a função de correlação cruzada dos sinais. Uma vez obtido  $D$ , podemos facilmente obter o ângulo da direção de chegada (DOA), conforme será explicitado nas seções posteriores.

Em um ambiente reverberante, ao invés de obtermos um impulso apenas na equação 3.33, obtemos uma série de impulsos com diferentes defasagens (ecos do sinal), atenuados por um fator  $\alpha$ . Para este caso, a função de correlação cruzada entre os sinais é dada por:

$$\widehat{R}_{x_i x_j}(\tau) = \widehat{R}_{SS}(\tau) * \sum_i \alpha_i \delta(\tau + D_i) = \sum_i \widehat{R}_{SS}(\tau + D_i) \quad (3.34)$$

Sendo assim, em ambientes de grande reverberação, temos múltiplos picos para a função de correlação cruzada, que correspondem aos múltiplos sinais incidentes no arranjo, mas que, de fato, são reflexões de um único sinal. Esses picos serão somados na função de correlação cruzada final, o que dificulta, ou mesmo inviabiliza, a estimação do atraso  $D$  nessas circunstâncias [7, 16, 29].

Para minimizar o efeito de espalhamento espectral ocasionado pelas reflexões do sinal em ambiente reverberante, escolhemos a função de pré-filtragem  $\gamma(k)$  como (3.30):

$$\gamma(k) = \frac{1}{|\hat{\varphi}_{x_i x_j}(e^{j\omega})|} \quad (3.35)$$

Esta função de pré-filtragem, PHAT, tem a propriedade de reduzir a função de correlação cruzada a um impulso, como se pode verificar:

$$\begin{aligned} \hat{R}_{x_i x_j}^{PHAT}(\tau) &= \mathcal{F}^{-1}\{\hat{\varphi}_{x_i x_j}(e^{j\omega})\} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \frac{1}{|\hat{\varphi}_{x_i x_j}(e^{j\omega})|} \right\} \hat{\varphi}_{x_i x_j}(e^{j\omega}) e^{j\omega\tau} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \frac{1}{|\hat{\varphi}_{ss}(e^{j\omega})e^{j\omega D}|} \right\} \hat{\varphi}_{ss}(\omega) e^{j\omega D} e^{j\omega\tau} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j\omega D} e^{j\omega\tau} d\omega \\ \hat{R}_{x_i x_j}^{PHAT}(\tau) &= \delta(\tau + D) \end{aligned} \quad (3.36)$$

Ou seja, desde que o ruído seja decorrelatado, podemos aplicar a transformada de fase (PHAT) na função de correlação cruzada para obter como resposta um impulso

centrado no atraso estimado  $D$ . Esta técnica é utilizada para minimizar a dispersão espectral causada pelas reflexões do sinal em ambientes reverberantes [45].

Para ilustrar o efeito de concentração do espectro da função de densidade espectral espacial, é apresentada a função de correlação cruzada para um sinal sonoro de referência com e sem a utilização da transformada de fase (PHAT).

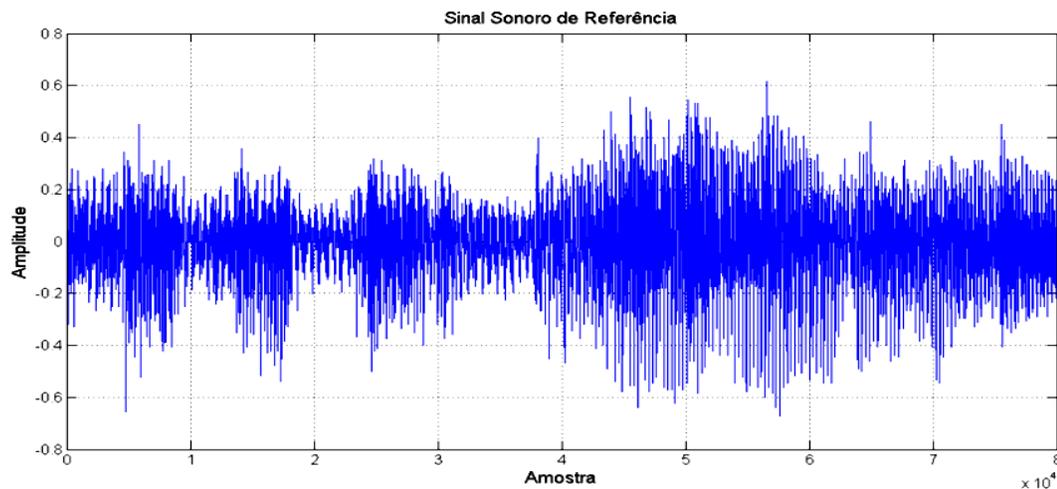


Figura 3.9: Sinal sonoro utilizado na simulação

Para simular um ambiente reverberante, o sinal sonoro acima (figura 3.9) foi preparado com um atraso  $D = 10$  ( $\text{DOA} = 31,54^\circ$ ) e  $D = 15$  ( $\text{DOA} = 51,70^\circ$ ) com as mesmas amplitudes, com  $f_s = 44.1 \text{ kHz}$  e  $d = 15 \text{ cm}$ . Quando o sinal não é pré-filtrado com a técnica PHAT, o método SRP aponta a direção de chegada  $40,44^\circ$ , conforme gráfico a seguir:

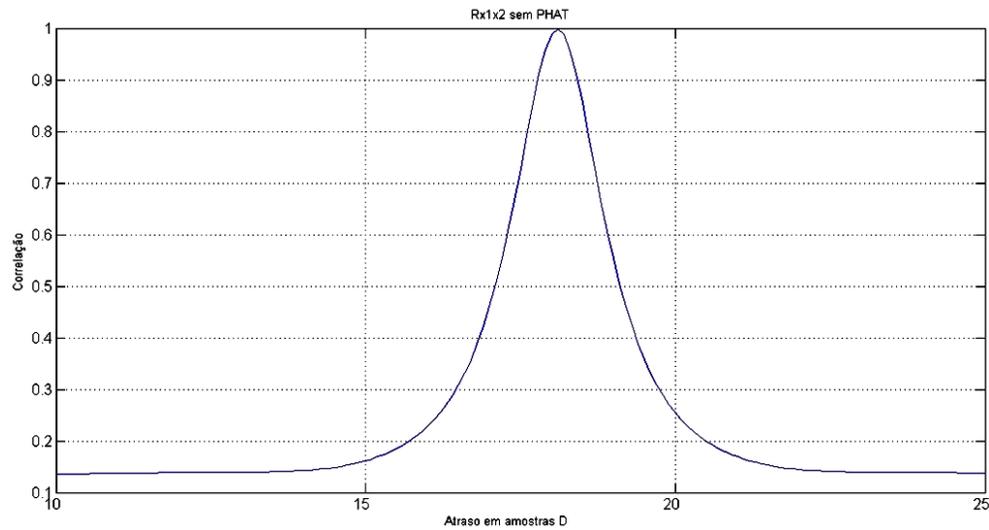


Figura 3.10: Função de correlação sem transformada de fase de um sinal com atrasos correspondentes a 10 e 15 amostras

Para os mesmos sinais, com a aplicação da técnica PHAT, tem-se a função de correlação apresentada no gráfico abaixo. Neste caso, o cálculo da correlação cruzada mostra, corretamente, picos em  $31,54^\circ$  (10 amostras) e  $51,70^\circ$  (15 amostras):

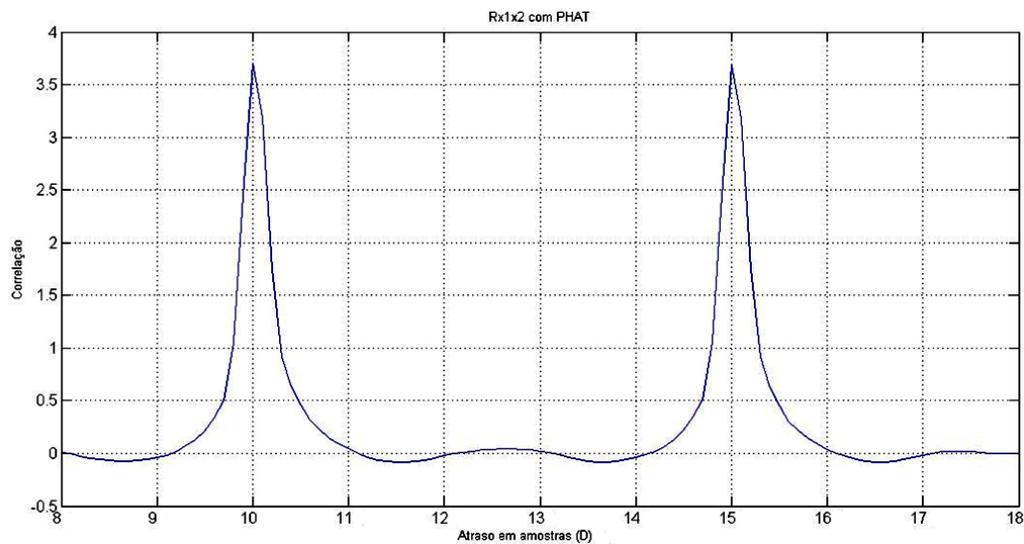


Figura 3.11: Função de correlação com transformada de fase para um sinal com atrasos de 10 e 15 amostras

Podemos notar que no primeiro caso, como existiam atrasos em 10 e 15 amostras, houve uma notável superposição dos picos de correlação e o estimador aponta, erroneamente, para um ângulo intermediário entre  $31,54^\circ$  e  $51,70^\circ$ , ou seja,  $40,44^\circ$ .

Já com a aplicação do método PHAT e pelos motivos anteriormente citados há uma separação clara, na função de correlação, dos atrasos provenientes do sinal com atraso de 10 amostras e do sinal com atraso de 15 amostras.

Apesar da aplicação da transformada de fase produzir picos mais estreitos na função de correlação, esta possui a característica de ponderar de maneira igual todas as frequências do espectro considerado. Sob condições de baixa relação sinal-ruído, elementos de frequência que são predominantemente originados pelo ruído terão os mesmos pesos relativos que elementos que compõem o sinal, o que pode ocasionar erros consideráveis. Deste modo, o uso da técnica PHAT deve ser evitado nos casos em que o ruído ambiente for elevado [45].

### **3.5 COMPARATIVO ENTRE OS MÉTODOS**

O problema da estimação da direção de chegada de fontes de onda por meio de arranjo de sensores tem sido tema de diversas pesquisas na área acadêmica. Do desenvolvimento do tema surgiram os diversos algoritmos, dentre os quais, para este trabalho citamos DS, CAPON, MUSIC, SRP e SRP-PHAT.

Nesta seção são realizadas diferentes simulações com vistas a caracterizar o comportamento dos estimadores concebidos pelos diferentes métodos frente a diversos aspectos como: relação sinal-ruído (SNR), reverberação, interferentes em outras direções e separação do interferente do sinal desejado.

A caracterização dos métodos é fator essencial para a escolha do método a ser implementado para estimação da direção de chegada em ambiente real.

Para as simulações realizadas, são considerados como sinais de referência:

## Sonoro

Trecho sonoro extraído de música, com  $f_s = 44.1 \text{ kHz}$ ,  $N = 82.660$  amostras e codificado em 16 bits.

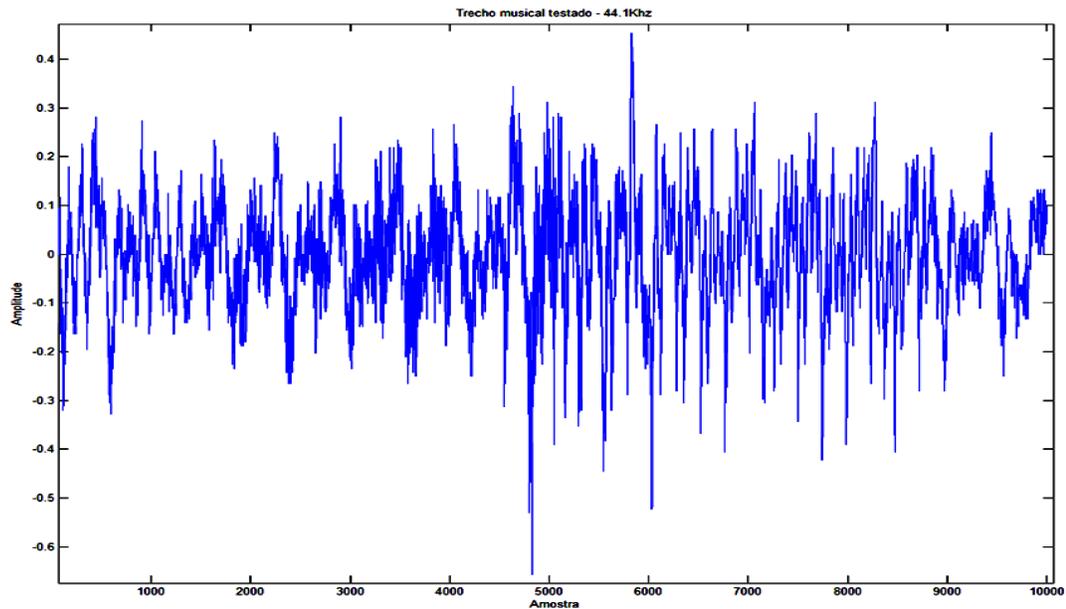


Figura 3.12: Trecho sonoro considerado nas simulações

### Trecho sonoro filtrado em 1 kHz com banda de 100 Hz

Trecho sonoro extraído de música, com  $f_s = 44,1 \text{ kHz}$ ,  $N = 82.660$  amostras e codificado em 16 bits. Filtrado com filtro digital (IIR) tipo *Chebyshev*, com 4 pólos, fator de *roll-off* 80dB, passa-faixas, entre 950 e 1050 Hz.

Escolhemos o filtro *Chebyshev* (IIR) por este possuir uma curva de resposta mais íngreme em amplitude que filtros *FIR* por ser possível sua implementação com fase linear por meio da função *filtfilt* do MATLAB<sup>®</sup>.

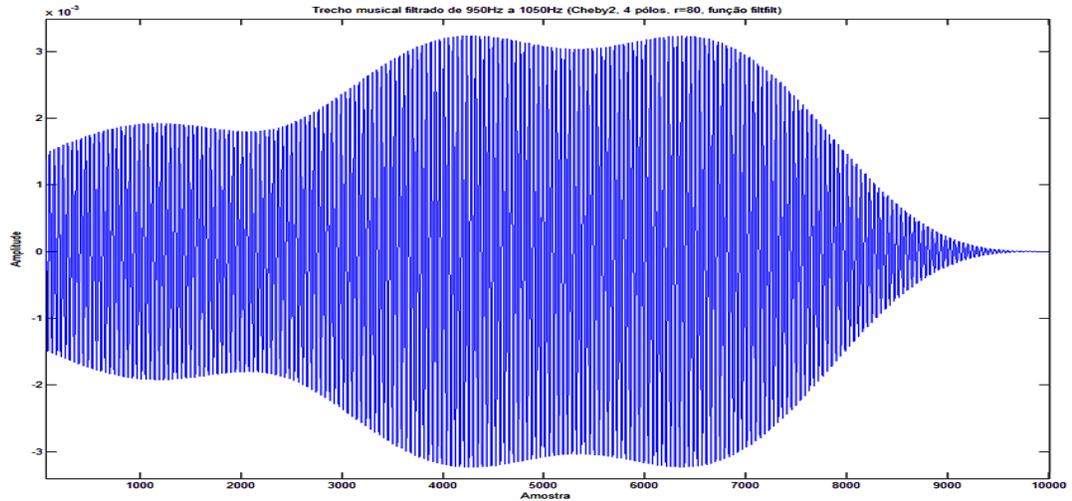


Figura 3.13: Trecho sonoro filtrado com filtro digital considerado nas simulações

### Tom senoidal de 1kHz

Tom senoidal gerado em 1 kHz, com  $f_s = 44,1 \text{ kHz}$ ,  $N = 88.200$  amostras e codificado com 16 bits.

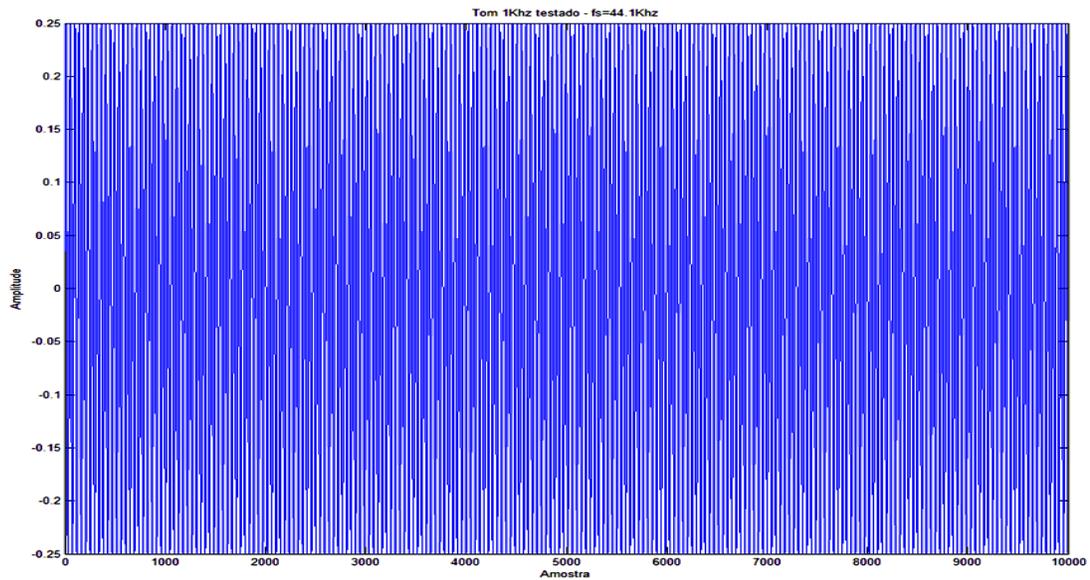


Figura 3.14: Tom senoidal de 1 kHz considerado nas simulações

### 3.5.1 Relação Sinal-Ruído (*SNR*)

Esta simulação é realizada de modo a caracterizar o comportamento dos algoritmos frente a uma relação sinal-ruído (*SNR*) decrescente. A variação da relação *SNR* é de  $-75\text{ dB}$  a  $75\text{ dB}$ . Como referência, um nível de conversação normal tem *SNR* em torno de  $40\text{ dB}$ .

Para cada valor de *SNR*, foram feitos 500 ensaios. O erro é obtido pela média dos erros absolutos (EMA) das 500 estimativas, enquanto que o desvio-padrão é calculado com base no vetor das 500 medidas da EMA. O ângulo para o qual o sinal simulado é defasado representa DOA de  $31,54^\circ$ .

#### 3.5.1.1 Trecho sonoro

#### MUSIC, CAPON e DS

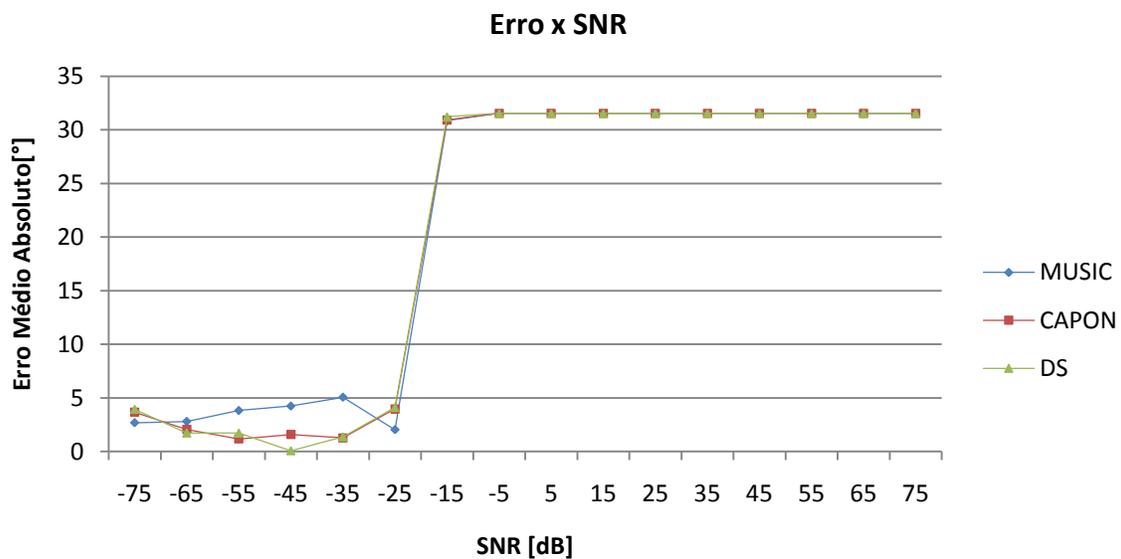


Figura 3.15: Variação do EMA com relação ao SNR para um trecho sonoro: MUSIC, CAPON e DS

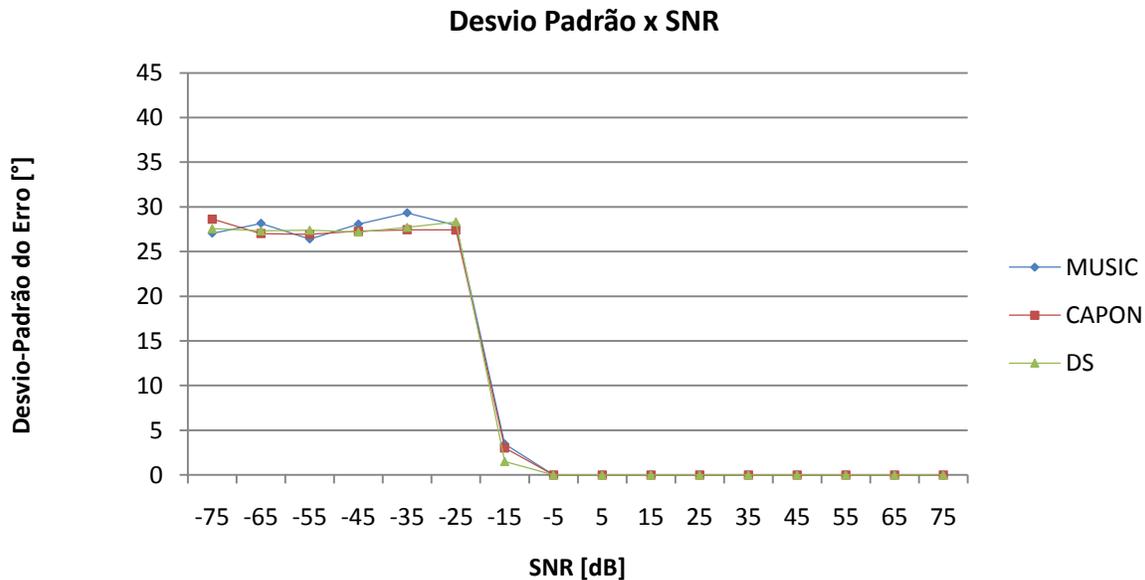


Figura 3.16: Dispersão do erro com relação ao SNR para um trecho sonoro: MUSIC, CAPON e DS

Das Figuras 3.15 e 3.16 nota-se que para baixas relações de  $SNR$ , de  $-75$  a  $-5$   $dB$ , os algoritmos MUSIC, CAPON e DS, que são algoritmos de banda estreita, parecem convergir para um erro médio absoluto em torno de  $4^\circ$ .

No entanto, pela análise conjunta dos gráficos do erro e desvio-padrão, observa-se que, de fato, os algoritmos não apresentam boa resolução quando em condições de SNR baixo. O que ocorre é que, pela elevada dispersão das estimativas (em torno de  $27^\circ$ ), o erro médio absoluto, que é calculado pela média dos valores absolutos dos erros para cada estimativa, assume estatisticamente um valor baixo, próximos de zero.

Da análise do gráfico do desvio-padrão dos erros, observa-se que os algoritmos só convergem na estimação da direção de chegada para relações de SNR superiores a  $-5$   $dB$ . Embora haja convergência, neste caso, o EMA é de aproximadamente  $30^\circ$ . Como a DOA do sinal simulado é de  $31,57^\circ$ , conclui-se que os algoritmos estão estimando, sistematicamente, DOA próximo a  $0^\circ$ . Isto se deve ao uso de algoritmos de banda-estreita para sinais de banda larga sem pré-filtragem.

## SRP E SRP-PHAT

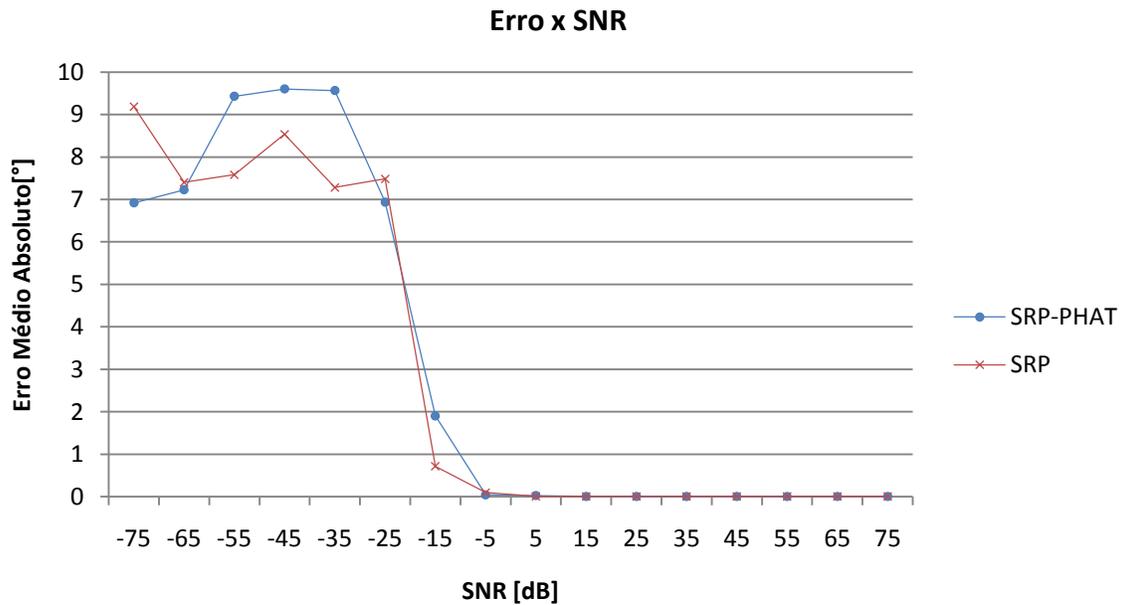


Figura 3.17: Variação do Erro com relação ao SNR para um trecho sonoro SRP e SRP-PHAT

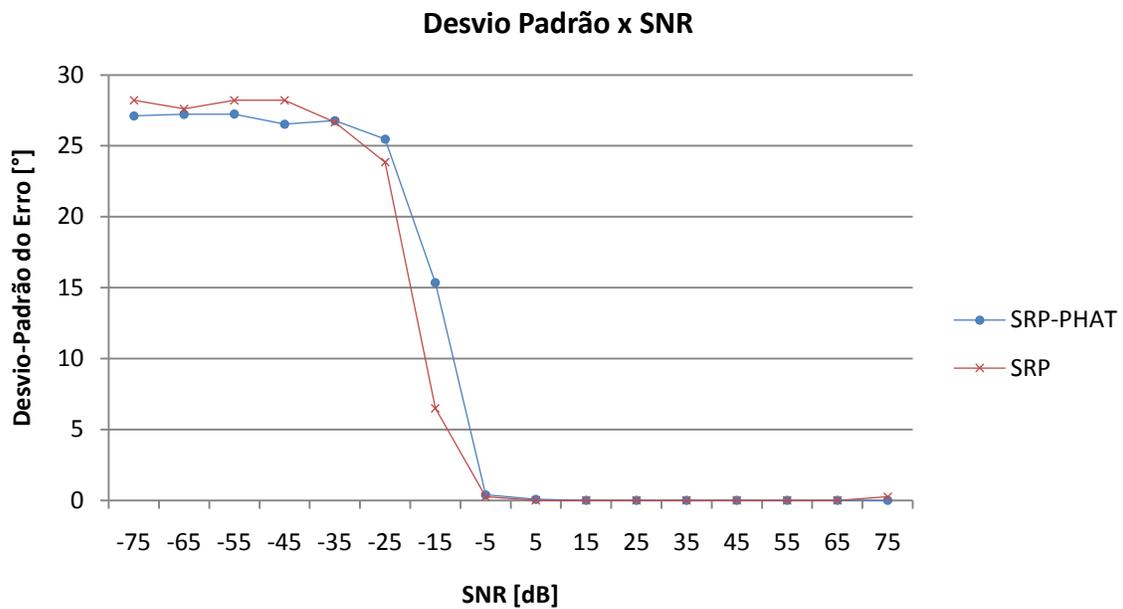


Figura 3.18: Dispersão do erro com relação ao SNR para um trecho sonoro SRP e SRP-PHAT

Já para os algoritmos SRP e SRP-PHAT, que são algoritmos banda-larga, nota-se que há convergência dos algoritmos a partir de valores SNR maiores que  $-5$  dB. Para

valores de SNR inferiores a este limiar, além do algoritmo não convergir – o desvio-padrão do erro sobe para próximo a  $27^\circ$ , e o erro médio das estimativas aumenta consideravelmente, o que desaconselha o uso desses métodos para valores de SNR abaixo do limiar estabelecido. Logo, concluímos que tanto técnicas de banda estreita como banda larga não são indicadas para regime de SNR baixo. No entanto, as técnicas de banda estreita tem uma tolerância maior frente ao ruído, uma vez que, como podemos observar pelos resultados, estas convergem para valores de SNR a partir de aproximadamente  $-20$  dB, enquanto as técnicas de banda larga só convergem para valores de aproximadamente  $5$  dB. Quanto à robustez frente ao ruído, note que a técnica da transformada de fase (PHAT) não introduz melhora significativa nos resultados obtidos.

### 3.5.1.2 Trecho sonoro filtrado em 1 kHz

#### MUSIC, CAPON e DS

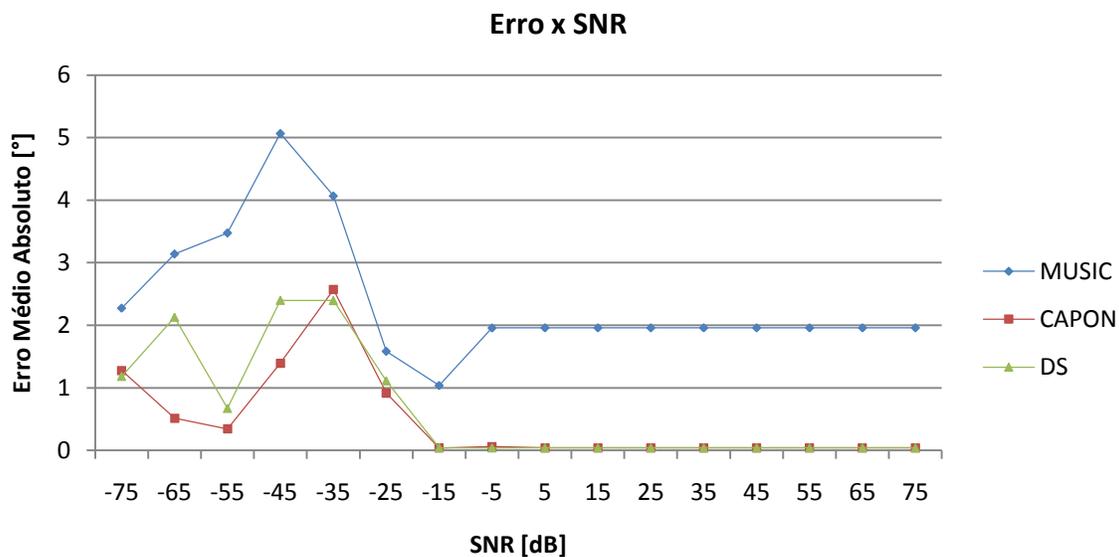


Figura 3.19: Variação do EMA com relação ao SNR para um trecho sonoro filtrado em 1 kHz: MUSIC, CAPON e DS

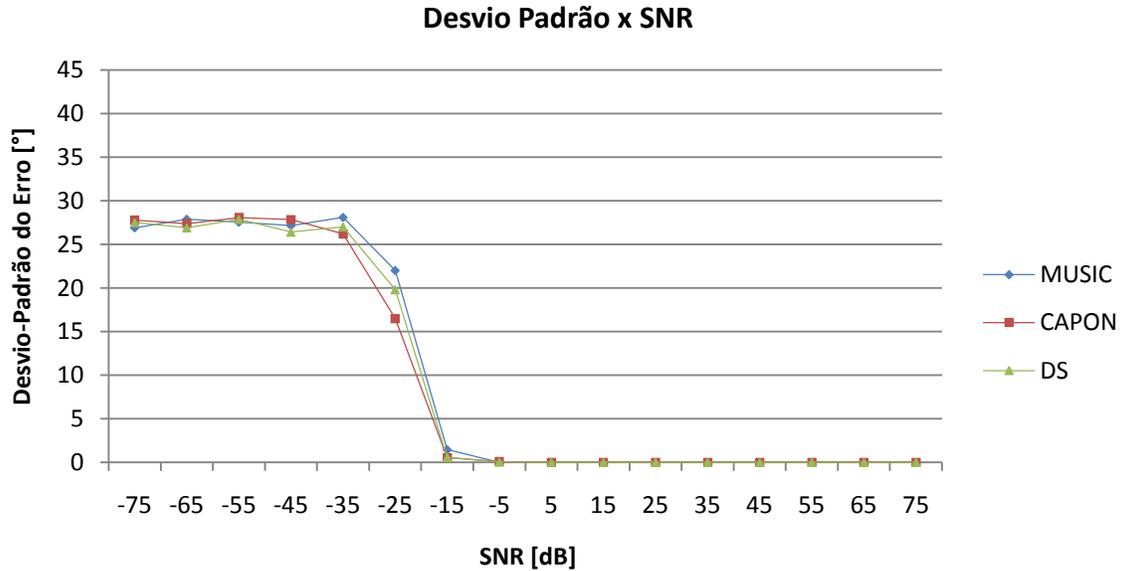


Figura 3.20: Dispersão do erro com relação ao SNR para um trecho sonoro filtrado em 1 kHz: MUSIC, CAPON e DS

Já para o trecho filtrado com filtro digital passa-faixas, entre 950 e 1050 Hz, observa-se a convergência dos algoritmos MUSIC, CAPON e DS, que são algoritmos de banda estreita, a partir de valores de SNR maiores que  $-15\text{ dB}$ .

Para esta situação, os algoritmos CAPON e DS convergem para valores de erro mínimos, enquanto o algoritmo MUSIC permanece com um viés (*bias*) na estimação de aproximadamente  $2^\circ$ , que pode ser parcialmente explicado por imperfeições na estimação dos autovetores e autovalores do ruído.

Esses resultados sugerem que os algoritmos de banda estreita podem ser utilizados para estimação da direção de chegada de sinais banda-larga, após filtragem em bandas estreitas por filtros digitais passa-faixas. No entanto, como explicitado em seções posteriores, para ambientes reais, essa solução não apresenta resultados satisfatórios.

## SRP E SRP-PHAT

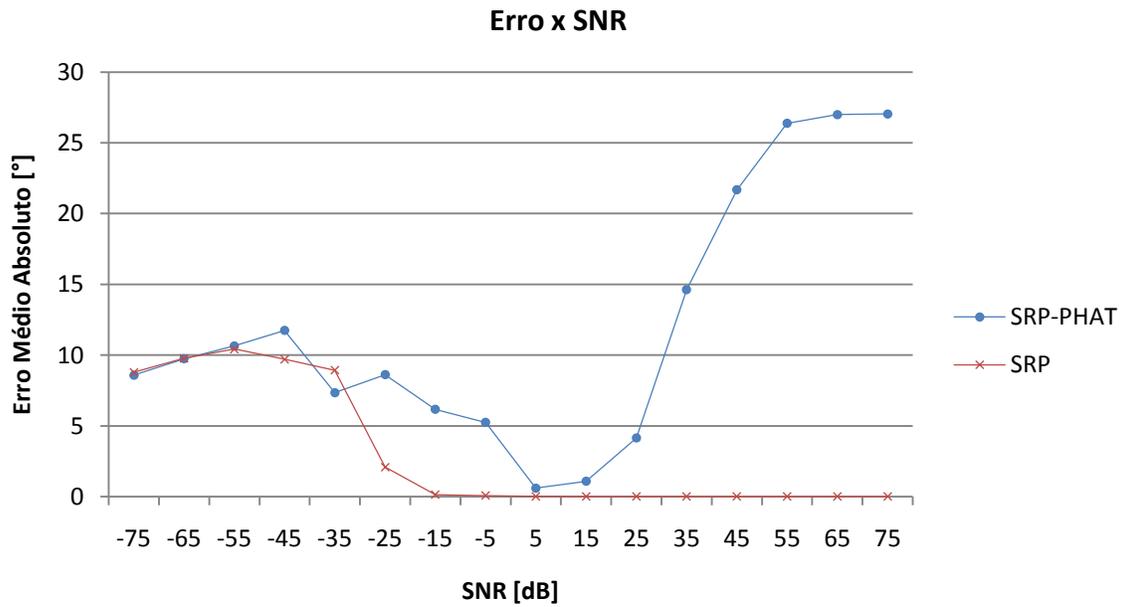


Figura 3.21: Variação do EMA com relação ao SNR para um trecho sonoro filtrado em 1 kHz SRP e SRP-PHAT

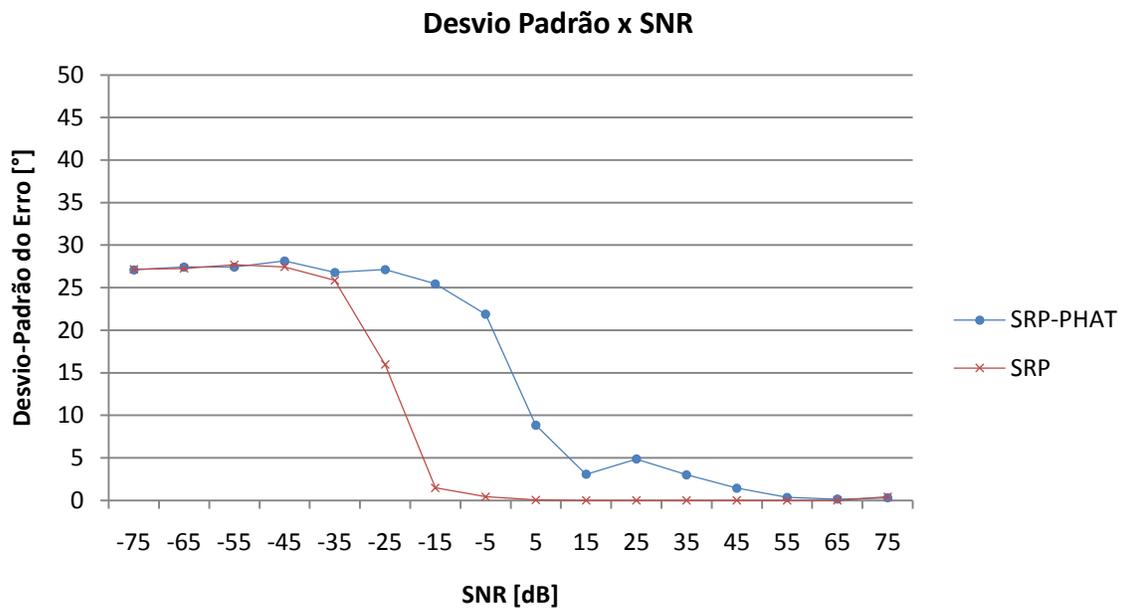


Figura 3.22: Dispersão do erro com relação ao SNR para um trecho sonoro filtrado em 1 kHz: SRP e SRP-PHAT

Para sinais de banda estreita, o algoritmo SRP tem melhor desempenho que o algoritmo SRP-PHAT. O SRP converge para valores de SNR maiores que  $-5\text{ dB}$  e apresenta erros mínimos a partir de valores maiores que  $-15\text{ dB}$ .

O algoritmo SRP-PHAT não tem boa resolução para sinais de banda-estreita. Conforme exposto no Capítulo 3, a transformada de fase tem por característica ponderar de maneira equânime todas as componentes de frequência do sinal. Neste caso, por ter-se adicionado ruído ao sinal, pesa-se de maneira equivalente o único componente de frequência do sinal ( $1\text{ kHz}$ ) com todos os demais componentes de frequência que compõem o ruído, o que causa a perda completa de resolução do método. Sendo assim, o algoritmo SRP-PHAT não deve ser utilizado para sinais de banda-estreita.

### 3.5.1.3 Tom senoidal em $1\text{ kHz}$

#### MUSIC, CAPON e DS

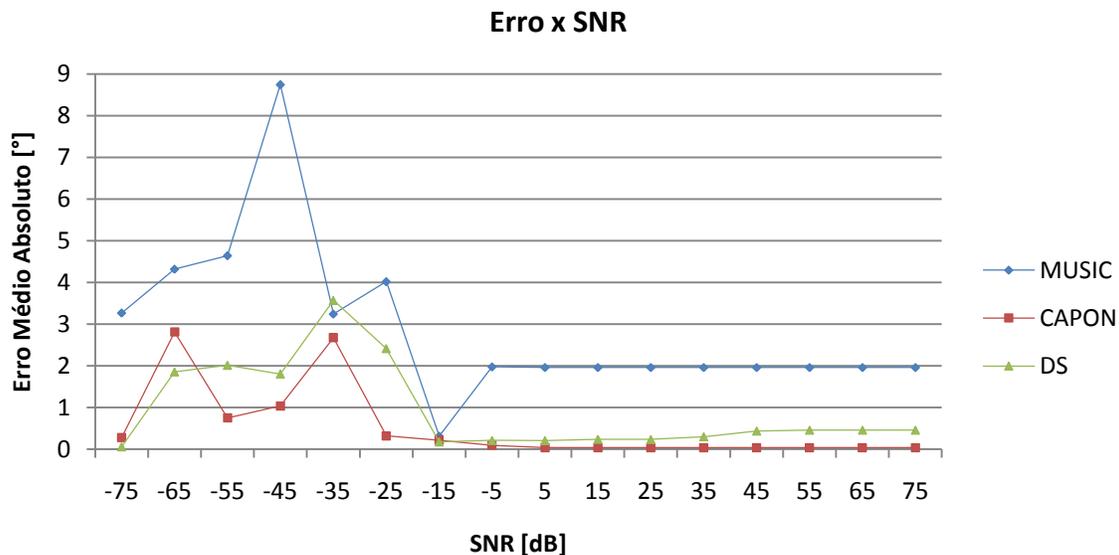


Figura 3.23: Variação do Erro com relação ao SNR para um tom senoidal de  $1\text{ kHz}$  para MUSIC, CAPON e DS

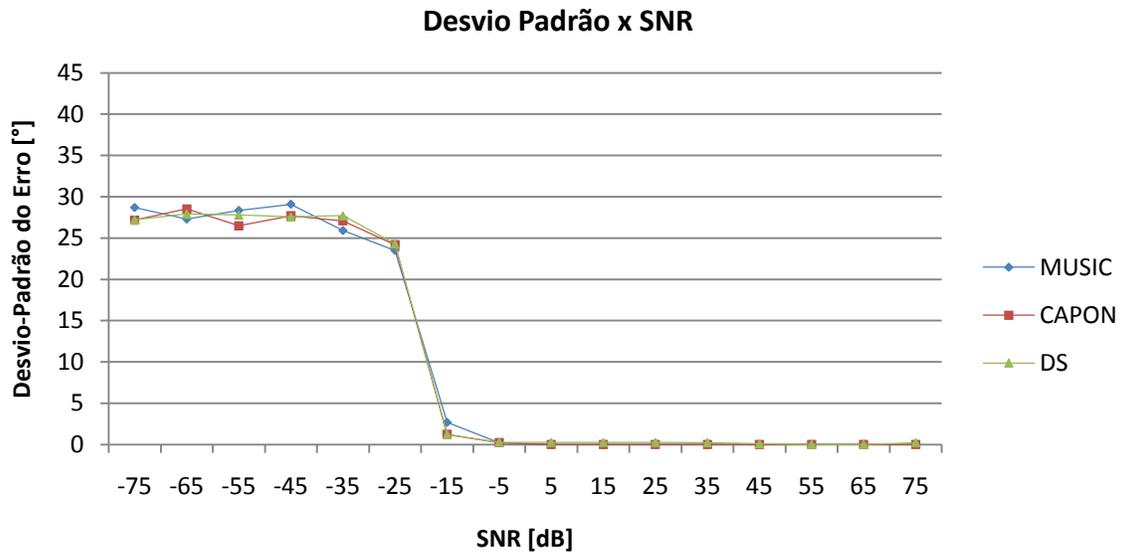


Figura 3.24: Dispersão do erro com relação ao SNR para um tom senoidal de 1 kHz: MUSIC, CAPON e DS

### SRP E SRP-PHAT

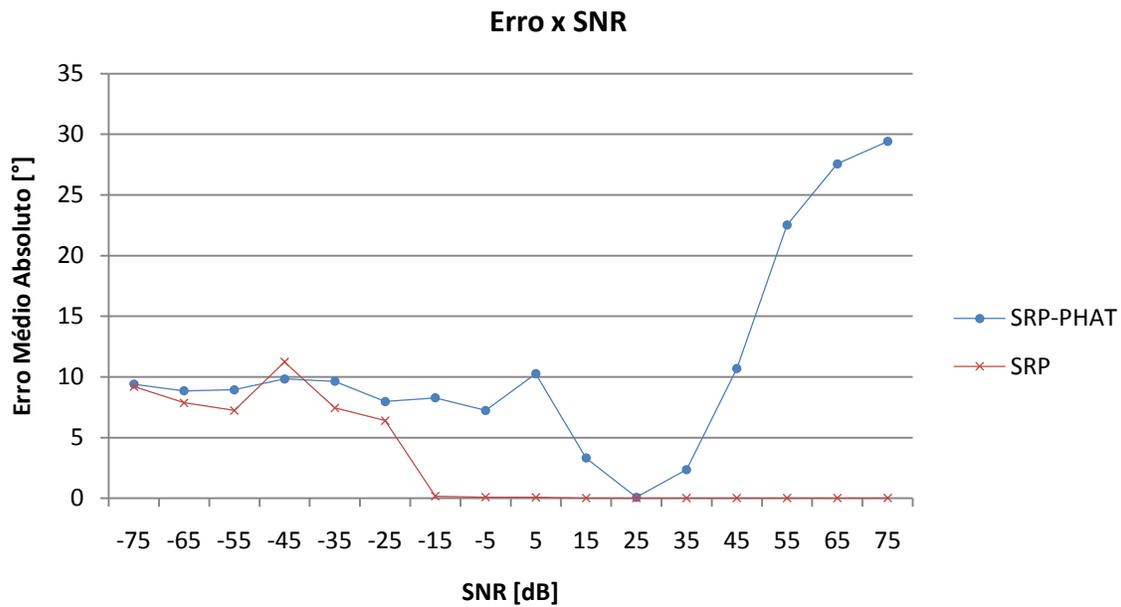


Figura 3.25: Variação do Erro com relação ao SNR para um tom senoidal de 1 kHz SRP e SRP-PHAT

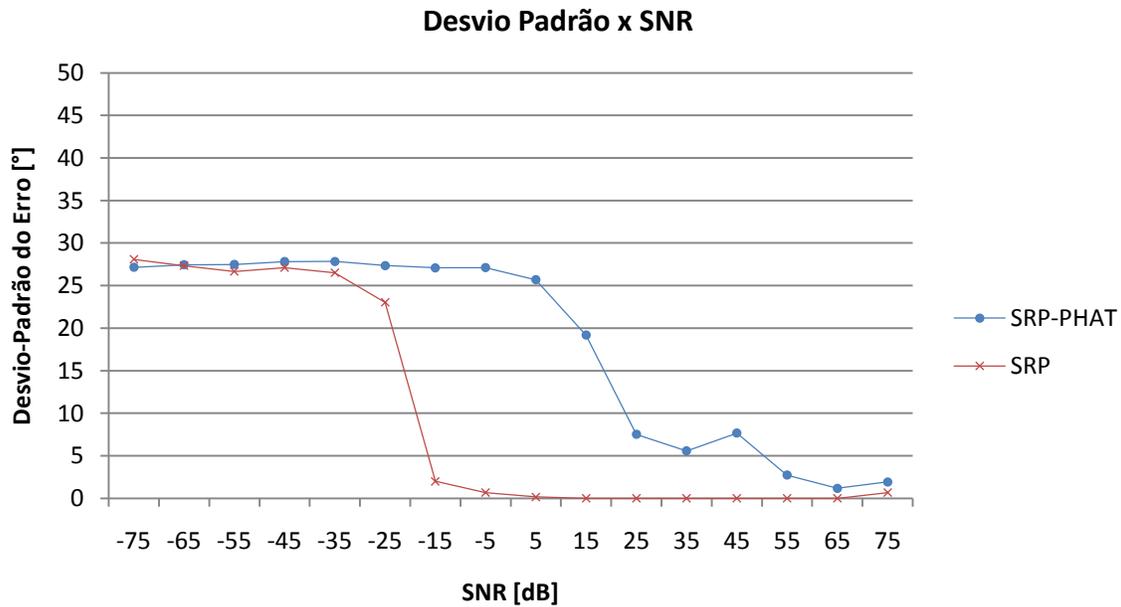


Figura 3.26: Dispersão do EMA com relação ao SNR para um tom senoidal de 1 kHz: SRP e SRP-PHAT

Para o caso do tom senoidal, valem as mesmas observações dos itens anteriores. A convergência para os métodos de banda estreita MUSIC, CAPON e DS se dá por volta de  $-5\text{ dB}$ . Para os métodos banda-larga, somente o SRP apresenta convergência, a partir de  $5\text{ dB}$  de SNR.

### 3.5.2 Reverberação

Para avaliar o comportamento dos métodos de estimação DOA em ambiente reverberante, são realizadas simulações com os sinais de referência, em que há um sinal de interesse na direção de chegada  $31,57^\circ$ , que representa um atraso de 10 amostras. Ao sinal de interesse, é somado um sinal interferente de DOA igual a  $51,70^\circ$ , simulado por um atraso de 15 amostras. Por simplificação, considera-se que há apenas uma reflexão do sinal (eco).

Define-se a relação eco-sinal pela equação (3.37):

$$RES = 10 \log_{10} \frac{VAR[sinal\ interferente]}{VAR[sinal\ interesse]} \quad (3.37)$$

A RES é variada de  $-72dB$  a  $72 dB$ , em passos de  $3 dB$ , para que seja analisado, em condições ideais, a partir de que valores de potência relativa os possíveis sinais interferentes gerados pelo efeito de reverberação ambiente afetam a estimação da direção de chegada.

### 3.5.2.1 Trecho sonoro

#### MUSIC, CAPON e DS

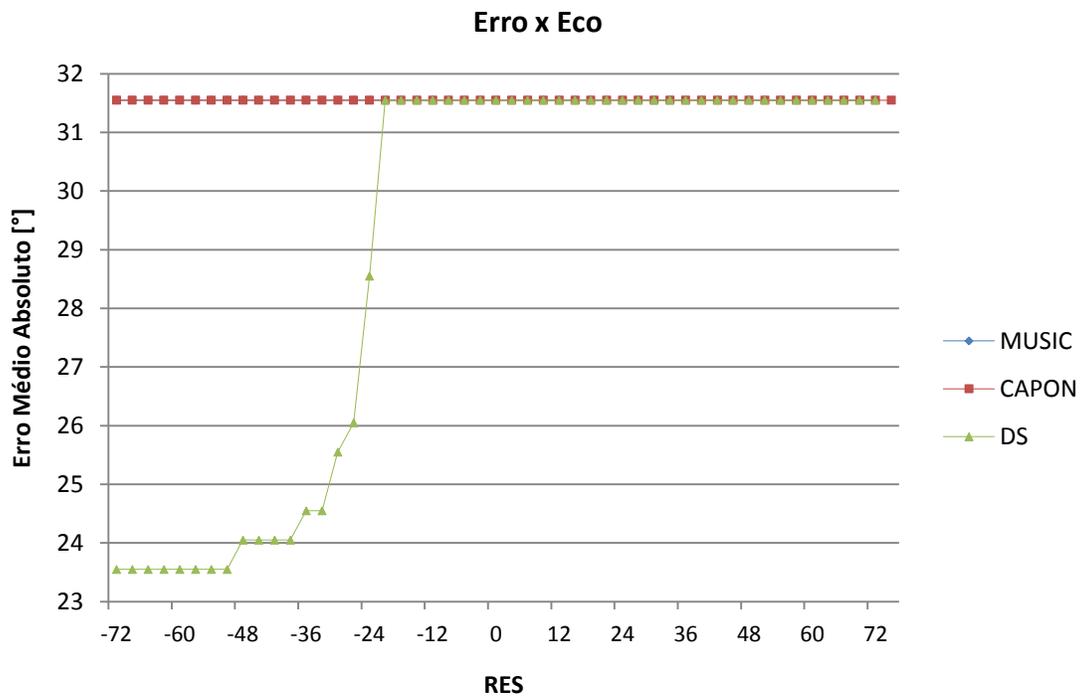


Figura 3.27: Variação do EMA com relação à potência do eco para um trecho sonoro: MUSIC, CAPON e DS

Como já explicitado, os algoritmos MUSIC, CAPON e DS não apresentam resultados satisfatórios para sinais banda-larga, assim o EMA apresenta-se sempre mais elevado.

## SRP E SRP-PHAT

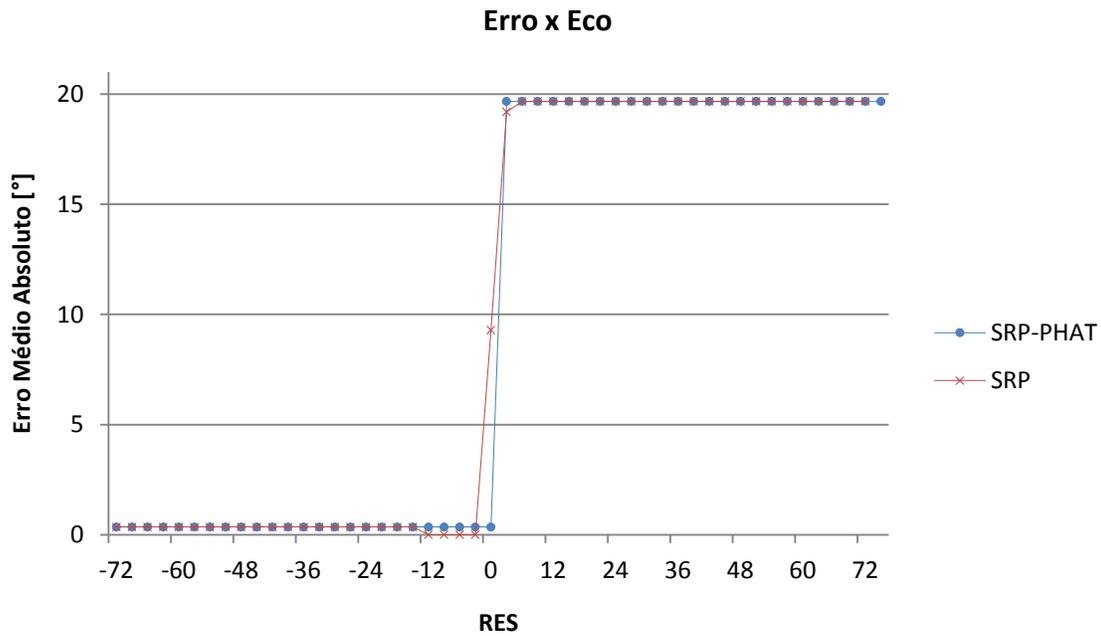


Figura 3.28: Variação do EMA com relação à potência do eco para um trecho sonoro: SRP e SRP-PHAT

Para os métodos SRP e SRP-PHAT, o estimador DOA apresenta erro EMA em torno de  $0,35^\circ$  até a relação sinal/eco de aproximadamente  $0\text{ dB}$ . Abaixo de  $0\text{ dB}$  o sinal interferente tem maior potência relativa e o estimador passa imediatamente a apontar para a direção do interferente ( $\text{DOA}=51,70^\circ$ ).

## 3.5.2.2 Trecho sonoro filtrado em 1 kHz com banda de 100 Hz

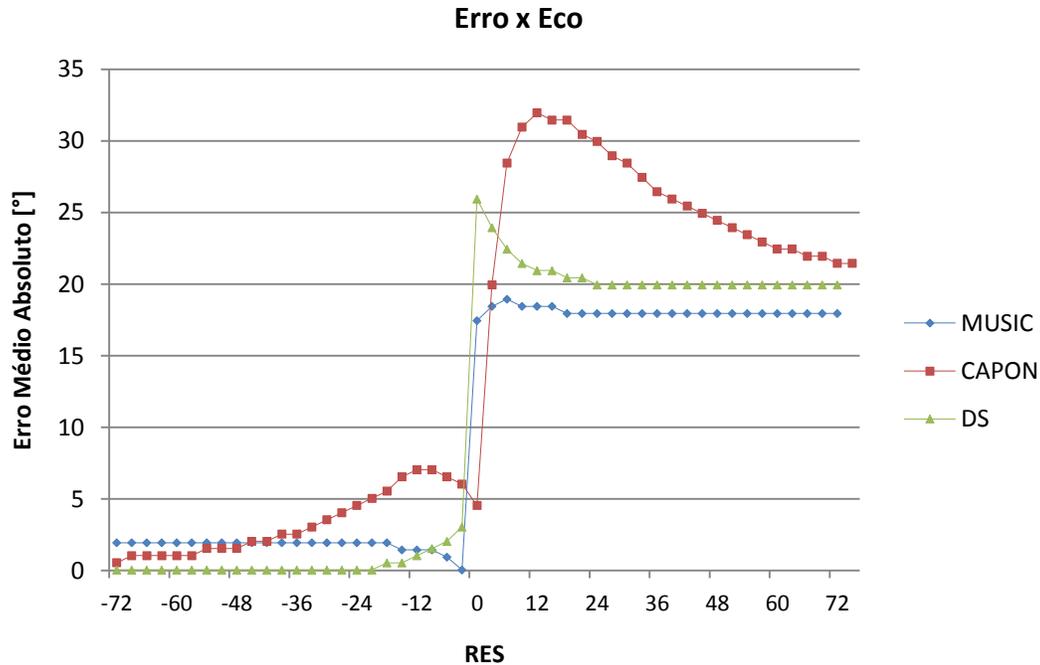
**MUSIC, CAPON e DS**

Figura 3.29: Variação do EMA com relação à potência do eco para um trecho sonoro filtrado em 1 kHz: MUSIC, CAPON e DS

Quando utiliza-se um filtro digital de banda estreita, os algoritmos MUSIC, CAPON e DS apresentam resultados satisfatórios. Para valores de sinal/eco menores que 0 dB o algoritmo MUSIC apresenta erro médio de 1,95°, CAPON de 3,28°, DS de 0,39°. Para valores acima de 0 dB todos passam a estimar o sinal interferente.

Há um sobressalto no algoritmo CAPON devido sua estrutura funcional que leva todos os valores de direção de chegada, exceto o momentaneamente calculado, para zero, o que acarreta uma convergência mais gradual - Figura 3.29.

## SRP E SRP-PHAT

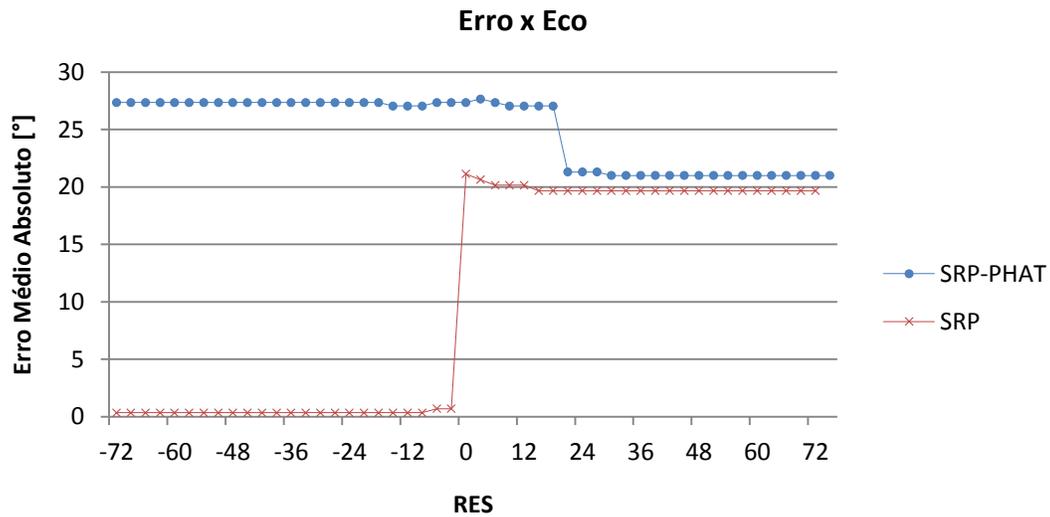


Figura 3.30: Variação do EMA com relação à potência do eco para um trecho sonoro filtrado em 1 kHz: SRP e SRP-PHAT

Para sinais de banda estreita, o algoritmo SRP-PHAT não converge. O método SRP tem erro médio de  $0.37^\circ$  e aponta para o sinal interferente a partir de  $0\text{ dB}$  para a relação eco/sinal.

### 3.5.2.3 Tom senoidal em 1 kHz

## MUSIC, CAPON e DS

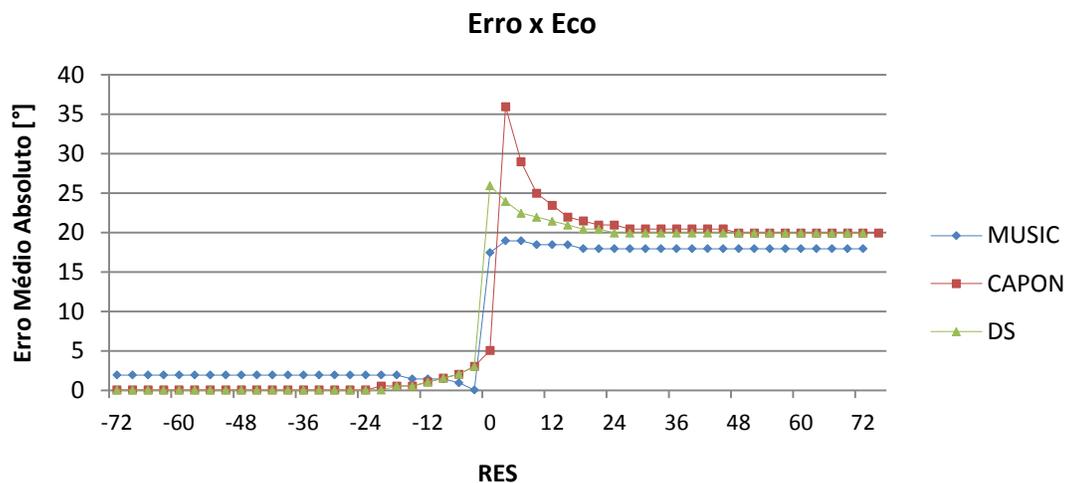


Figura 3.31: Variação do EMA com relação à potência do eco para um tom senoidal de 1 kHz: MUSIC, CAPON e DS

Quando se tem um tom senoidal como entrada, para valores de sinal/eco menores que  $0\text{ dB}$  o algoritmo MUSIC apresenta erro médio de  $1,84^\circ$ , CAPON de  $0,61^\circ$ , DS de  $0,29^\circ$ . Para valores maiores que  $0\text{ dB}$  todos passam a estimar o sinal interferente.

### SRP E SRP-PHAT

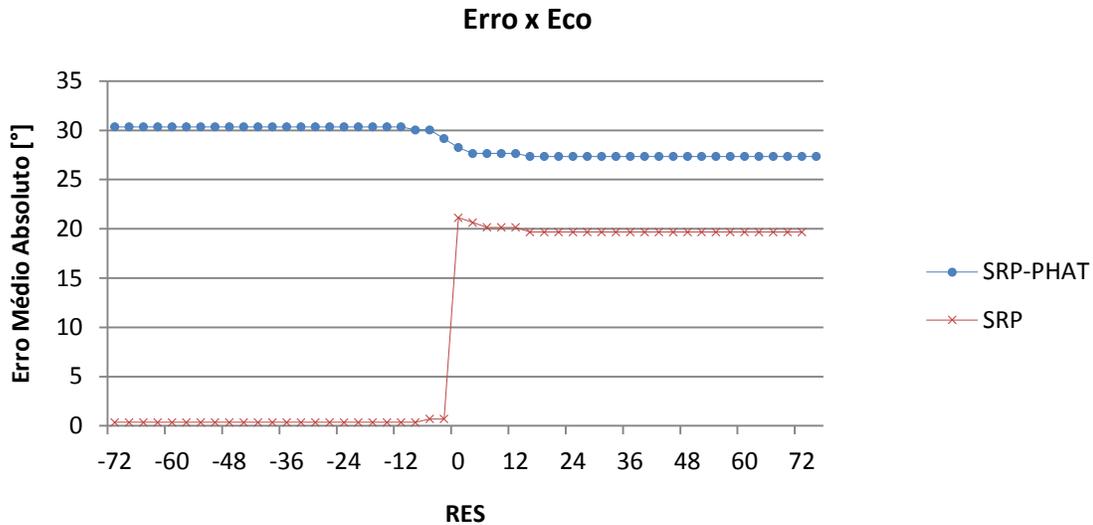


Figura 3.32: Variação do EMA com relação à potência do eco para um tom senoidal de 1 kHz: SRP e SRP-PHAT

Para um tom senoidal, o método SRP tem erro médio de  $0,38^\circ$  e aponta para o sinal interferente a partir de  $0\text{ dB}$  para a relação eco/sinal. Isto é, sempre que a potência relativa do eco for maior que a do sinal de interesse, o algoritmo irá apontar para DOA do sinal interferente.

#### 3.5.3 Separação do Interferente

Nas próximas simulações, será analisado o comportamento dos algoritmos estudados – MUSIC, CAPON, DS, SRP e SRP-PHAT – quando existe, somado ao sinal de interesse, ( $\text{DOA}=31,54^\circ$ ), uma versão defasada do sinal, com sua direção de chegada variando desde  $61,54^\circ$  até  $34,54^\circ$ .

O objetivo é caracterizar o comportamento dos métodos quando há um sinal interferente, com amplitude de 99% do sinal de interesse, se aproximando da fonte de referência. Ou seja, queremos investigar se os estimadores de direção de chegada perdem resolução à medida que fontes interferentes se aproximam da DOA do sinal de referência.

### 3.5.3.1 Trecho Sonoro

#### MUSIC, CAPON e DS

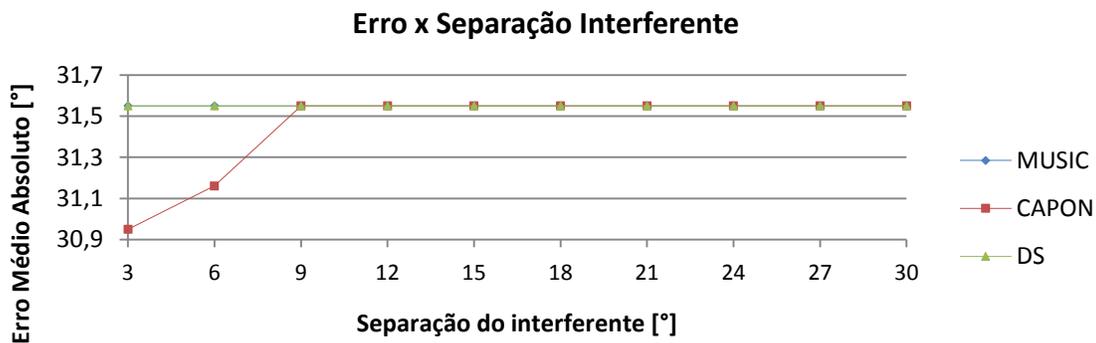


Figura 3.33: Variação do EMA com relação à separação do interferente para um trecho sonoro: MUSIC, CAPON e DS

Novamente se observa que os métodos de banda estreita não apresentam resolução satisfatória para sinais de banda larga. Além disso, nota-se que esses métodos não apresentam resultados consistentes para quase todas as realizações.

#### SRP E SRP-PHAT

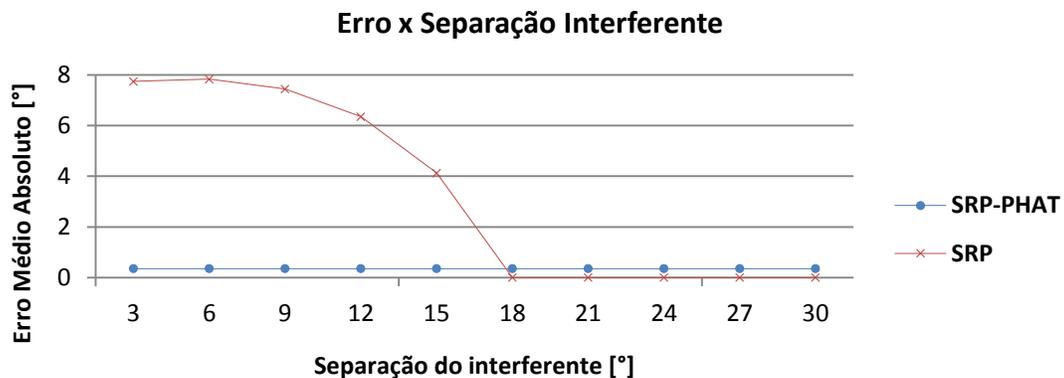


Figura 3.34: Variação do EMA com relação à separação do interferente para um trecho sonoro: SRP e SRP-PHAT

Já os algoritmos SRP e SRP-PHAT apresentam boa resolução para o caso banda-larga. Note que o algoritmo SRP tem sua resolução severamente afetada nos casos em que o sinal interferente se encontra espaçado de  $18^\circ$  ou menos da fonte de interesse. Isto se deve, em parte, à soma dos lóbulos de potência dos sinais no âmbito da função de densidade de potência espectral. Desse modo, não é aconselhável o uso dessa técnica quando houver interferentes espacialmente próximos ao sinal de interesse.

Este problema é superado pela adoção da transformada de fase (PHAT). A transformada PHAT torna os picos dos atrasos, na função de densidade espectral espacial, mais estreitos. Isto permite que o estimador DOA apresente boa resolução mesmo quando os possíveis sinais interferentes estão arbitrariamente próximos à fonte de interesse. Esta importante característica para aplicações em ambientes reais será levada em consideração na implementação do rastreamento DOA.

### 3.5.3.2 Trecho Sonoro filtrado em 1 kHz

#### MUSIC, CAPON e DS

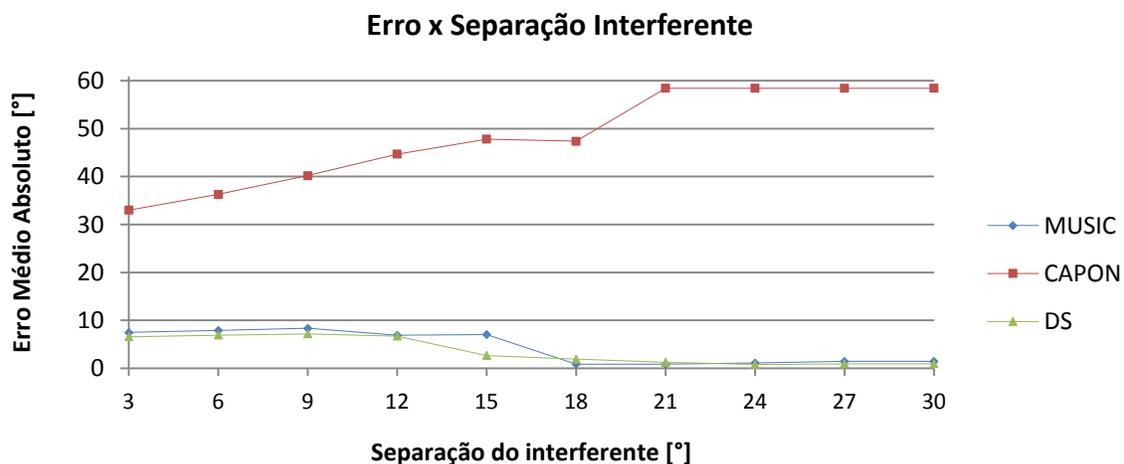


Figura 3.35: Variação do EMA com relação à separação do interferente para um trecho sonoro filtrado em 1 kHz: MUSIC, CAPON e DS

Para o caso do trecho sonoro filtrado digitalmente, o algoritmo MUSIC apresenta um erro médio de  $6,44^\circ$  nas estimativas acima de  $18^\circ$  de separação espacial, o método DS apresenta erro médio de  $6,02^\circ$  a partir de  $15^\circ$  de separação e o método CAPON não apresenta convergência satisfatória quando há um sinal interferente espacialmente próximo.

### SRP E SRP-PHAT

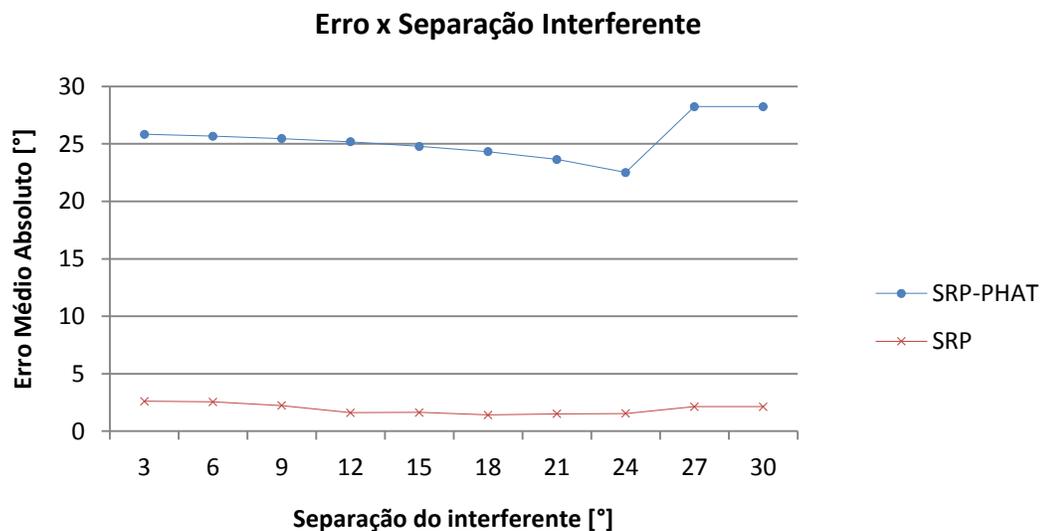


Figura 3.36: Variação do EMA com relação à separação do interferente para um trecho sonoro filtrado em 1 kHz: SRP e SRP-PHAT

Por ponderar as frequências de maneira equânime, o algoritmo SRP-PHAT não apresenta convergência para sinais de banda estreita. Já a técnica SRP apresenta erro médio de  $1,93^\circ$  em toda a faixa testada, o que recomenda seu uso sob essas circunstâncias.

## 3.5.3.3 Tom senoidal de 1 kHz

## MUSIC, CAPON e DS

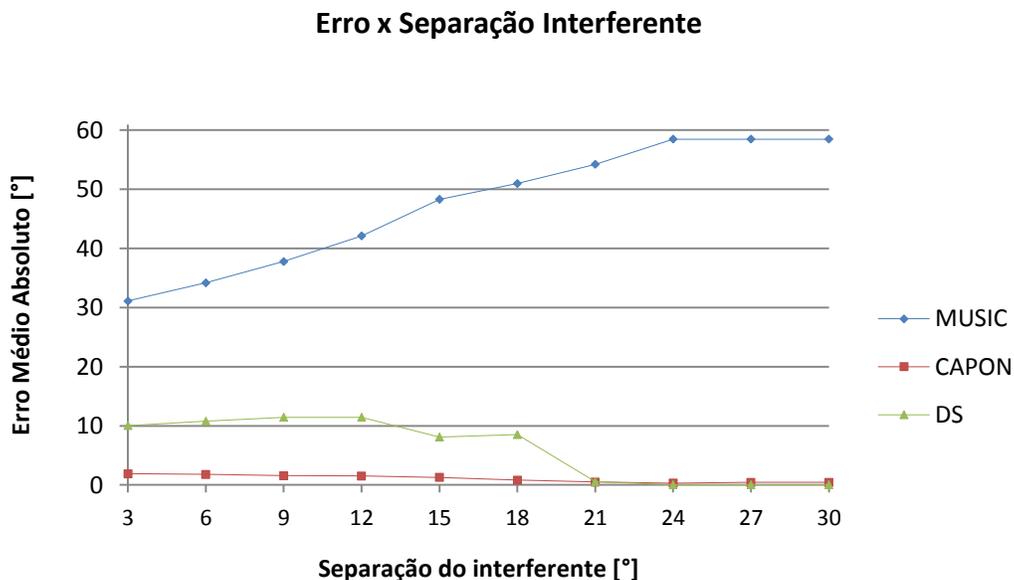


Figura 3.37: Variação do EMA em relação à separação do interferente para um tom senoidal de 1 kHz: MUSIC, CAPON e DS

Quando empregado com um tom senoidal como sinal de referência, o algoritmo DS apresenta erro médio de  $8,70^\circ$  a partir de  $21^\circ$  ou menos de separação espacial, o CAPON tem erro médio de  $1,05^\circ$  em toda a faixa testada e o algoritmo MUSIC não converge. Este é um indicativo de que, quando se trata de sinais de banda muito estreita a utilização do CAPON é adequada.

## SRP E SRP-PHAT

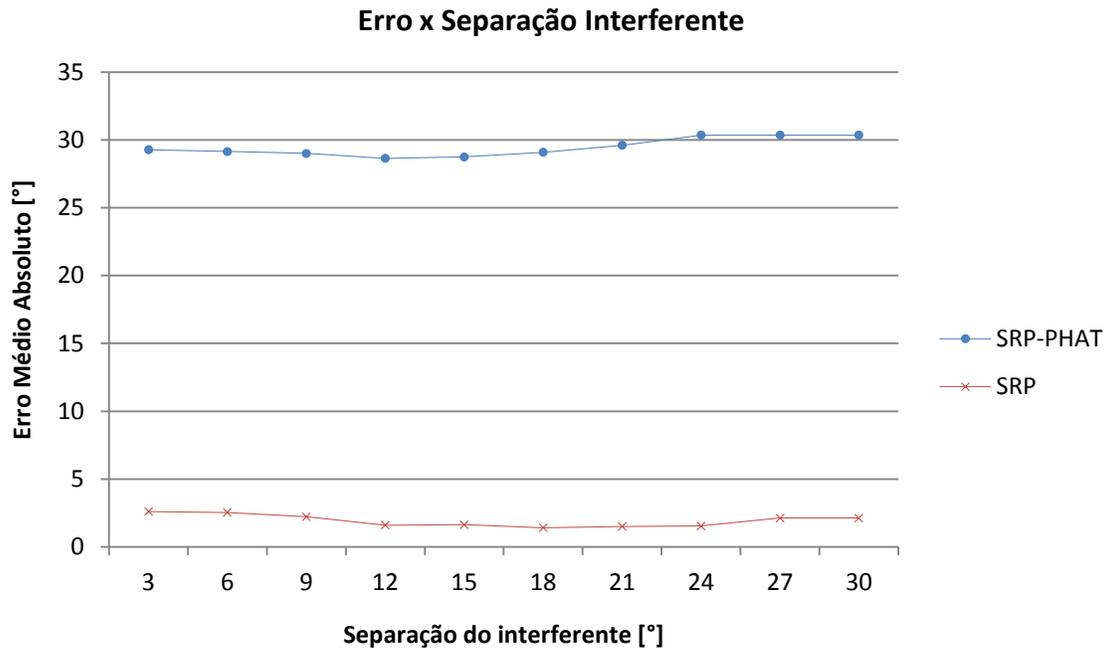


Figura 3.38: Variação do erro com relação à separação do interferente para um tom senoidal de 1 kHz: SRP e SRP-PHAT.

Para o caso banda estreita, o algoritmo SRP-PHAT não converge e o algoritmo SRP apresenta erro médio de  $1,93^\circ$  na faixa espacial considerada para o teste.

### 3.5 OUTRAS ABORDAGENS PARA ESTIMAÇÃO DOA

#### 3.5.1 Abordagens alternativas

Vários outros métodos foram desenvolvidos para o problema da estimação da direção de chegada em arranjo de sensores. Exemplos desses esforços são os algoritmos TOPS [42], WAVES [15], C-SPRIT [32], I-MUSIC [46].

YU *et al.* [43] apresentam um método para estimação DOA chamado TOFS (*Test of Orthogonality of Frequency Subspaces*) que consiste na procura pelos vetores volantes de cada possível direção de chegada para cada possível frequência, sem que seja necessário o

cálculo de uma matriz de foco (*focusing matrix*) para as frequências. Tal método se mostra não polarizado quando o SNR é alto e não apresenta picos falsos. É baseado no método I-MUSIC (*incoherent MUSIC*) que é uma adaptação da estimação DOA com o método MUSIC para o caso de sinais de banda larga.

YAN e FAN [41] apresentam uma abordagem para a cálculo em tempo real da direção de chegada (*DOA tracking*) para sinais banda-larga provenientes de múltiplas fontes móveis explorando a propriedade de ciclo-estacionariedade. Os autores supõem que os sinais captados possuem ciclo-estacionariedade de forma a ser possível perseguir de maneira seletiva o emissor. O princípio utilizado é o de que interferências e ruídos que não compartilham da mesma frequência de ciclo dos sinais desejados, ou que não exibem ciclo-estacionariedade, podem ser suprimidos, de forma a melhorar a estimação DOA. Foi utilizado o algoritmo de mínimos combinado com o filtro de *Kalman* para melhorar a performance da estimação DOA.

HASSANPOUR e KAHAEI [23] propõem um algoritmo para estimar a direção de chegada (DOA) de um locutor utilizando dois microfones. O algoritmo se baseia na minimização de uma função de custo que se altera segundo o *array* de dados – no domínio da frequência – adquiridos pelos microfones. A partir daí é derivada uma procura por funções de custo monotônicas, para a qual a convergência do algoritmo é garantida de maneira iterativa (*step-by-step*). O algoritmo de procura pelo mínimo da função de custo é do tipo gradiente descendente.

DO-HONG e RUSSER [18] apresentam uma análise para a estimação da direção de chegada (DOA) de fontes espacialmente próximas, com geometria de sensores arbitrária, levando em consideração o efeito do erro no modelo de arranjo de sensores (erros de posição, ganho e fase), o que constitui uma análise importante em implementações práticas.

Esses erros são introduzidos por sensores não idênticos, pela resposta dependente da frequência do circuito de aquisição/tratamento, por condições ambientais, entre outros. Os erros de posição, ganho e fase são modelados por distribuições Gaussianas em torno do valor correto. Pelo método MNT (*Modified Nominal Transformation*), combinado com a técnica MUSIC (MNT-MUSIC), é conseguido um erro de ordem 5 vezes menor com a presença de erros de modelo no arranjo de sensores do que com a técnica convencional.

WEI *et al.* [39] introduzem um método de baixa complexidade computacional para estimação DOA para sistemas DS/CDMA, visto que algoritmos de alta resolução como MUSIC, ESPRIT têm complexidade computacional proibitiva para uso em aplicações em tempo real. O novo método de estimação DOA (*tracking*) é utilizado para rastrear DOA de comunicações terra-ar e utiliza feixe duplo e passos variáveis para estimar a direção de chegada. O princípio do algoritmo é comparar a amplitude de modulação do sinal durante cada ciclo de *tracking* DOA, ajustando dois feixes gradualmente para a direção de chegada do usuário. Se a amplitude da modulação é igual ou quase igual para os dois feixes, então a DOA está na direção do eixo entre os dois feixes.

Segundo VIBERG e LUNDGREN [38], o modelo matemático de resposta do arranjo de sensores tem grande influência na estimação DOA, principalmente em algoritmos de alta resolução como MUSIC. Na prática, o modelo de arranjo está sujeito à perturbações por variações na posição dos sensores, canais elétricos não ideais e acoplamento mútuo. Para mitigar esse problema, é possível medir a resposta do arranjo para fontes em locais conhecidos. No modelo apresentado pelos autores, é considerado que se conhece um modelo nominal para o arranjo, que é utilizado para interpolar os dados de calibração do arranjo por meio do método de Aproximação Local por Polinômios (*Local Polynomial Approximation*). Além disso, é derivado uma variação do método MUSIC

denominado W-MUSIC. Nesta hipótese, os erros de modelo dos sensores são considerados individualmente, o que potencialmente reduz a sensibilidade da estimação DOA aos mesmos. A tese central dos autores é multiplicar a matriz de vetores volantes por uma matriz de correção, que pode depender ou não do ângulo. Simulações com perturbação gaussiana nos parâmetros mostram um ganho de aproximadamente 30% pelo uso dessa técnica.

ONUMA *et. al.* [32] apresentam um algoritmo de estimação DOA de alta resolução para sinais banda larga empregando o algoritmo C-SPRIT e a transformada de Fourier discreta (DFT). O método C-SPRIT permite estimar a direção de chegada de ondas correlacionadas sem que haja o indesejado efeito de espalhamento espacial do resultado. A correlação de sinais pode ocorrer por efeito *multipath* ou por reflexões, por exemplo. Os algoritmos MUSIC e ESPRIT são frequentemente utilizados para sinais de banda estreita (*narrowband*). No entanto, na abordagem proposta é possível utilizar o algoritmo ESPRIT para sinais do tipo banda larga (*wideband*).

MENDONZA, COVARRUBIAS e LÓPES [30] introduzem um método de localização espacial de fontes em sistemas de comunicações móveis para celular que permite, em tempo real, estimar continuamente a localização espacial das fontes. Para enfrentar o problema da complexidade computacional do método MUSIC, são utilizados algoritmos que se baseiam na atualização sucessiva do subespaço amostrado mediante novas amostras: PAST (*Projection Approximation Subspace Tracking*) e Bi-SVD (*Bi-Iteration Singular Value Decomposition*). Os resultados obtidos demonstram uma redução significativa do tempo de processamento, além de ganho na precisão da estimação DOA para o ambiente dinâmico, situação em que o algoritmo MUSIC convencional apresenta grande perda de precisão.

Uma modificação do algoritmo MUSIC para estimação DOA foi proposta por ZHOU, ZHAO e CUI [46]. O algoritmo MUSIC apresenta baixa resolução em diversos casos: na detecção de sinais quando as fontes estão espacialmente próximas e a relação SNR (*signal to noise ratio*) é baixa; quando há um número relativamente baixo de *snapshots*; quando da expansão do subespaço em condições não ideais. Para sinais banda larga, os vetores volantes e os padrões de direção para frequências diferentes não são constantes, o que introduz erros na estimação.

Conforme proposto pelos autores, para mitigar os referidos problemas e gerar uma versão do algoritmo MUSIC que possa ser utilizada na estimação de sinais banda larga, realiza-se a ponderação dos autovetores no subespaço de forma que a gravidade específica (peso relativo) dos autovetores no espectro espacial seja ajustada, conforme abaixo:

$$P_{MUSIC} = \frac{a^H(\theta)a(\theta)}{a^H(\theta)U_N\Psi\Psi U_N^H a(\theta)} \quad (3.38)$$

$$\text{Em que } \Psi\Psi = \frac{\varphi}{\sqrt{\left(\frac{1}{\lambda_{N+1}}\right)^2 + \left(\frac{1}{\lambda_{N+2}}\right)^2 + \dots + \left(\frac{1}{\lambda_M}\right)^2}} \quad (3.39)$$

$$\text{E } \varphi =$$

$$\text{diag}\left\{\frac{1}{\lambda_{N+1}}, \frac{1}{\lambda_{N+2}}, \dots, \frac{1}{\lambda_M}\right\} \quad (3.40)$$

Tal alteração faz com que os autovetores relacionados aos maiores autovalores ganhem mais importância que os autovetores relacionados aos menores autovalores. Esse ajuste da “gravidade” de cada autovetor no espectro espacial reduz a indesejada influência ocasionada pela expansão do subespaço de ruído devido em condições de baixa SNR.

YAMADA [40] argumenta que uma estratégia comum para a resolução de estimação DOA por algoritmos de alta resolução como MUSIC é a transformação de um problema de estimação DOA para um problema de otimização global não linear: achar todos os mínimos ou máximos, locais ou não, de certas funções de uma variável definidas como valores no círculo complexo unitário. Esses valores estão definidos em polinômios auto-recíprocos de Laurent de uma variável.

Desta maneira, propõe um algoritmo finito para computar a solução completa de distribuição de mínimos e máximos para as referidas funções no círculo unitário, isto é, o número exato de mínimos e máximos localizados num sub-arco do círculo unitário. A idéia chave do algoritmo é o uso da sequência de *Sturm* generalizada. Assim, o algoritmo pode vir a ser uma ferramenta matemática poderosa para a estimação DOA com algoritmos de alta resolução e vir a ser utilizada no desenvolvimento futuro de técnicas para estimação DOA [40].

CAO, LIU e XU [8] apresentam uma técnica de matrizes de foco para métodos banda-estreita como MUSIC pela utilização estimação CSM (*Coherent Signal-subspace Method*), que consiste em explorar as propriedades de auto-correlação e correlação cruzada para propor uma transformação das matrizes de sinais para cada frequência (*frequency bin*). Deste modo, é conseguido um método que não requer valores iniciais para o cálculo e apresenta bons resultados para sinais banda-larga.

SONG *et al.* [35] utilizam uma abordagem bayesiana para a filtragem espacial por meio de estimação DOA quando estão presentes incertezas nas medidas devido ao ruído, imperfeições nos circuitos e sensores e reflexões indesejadas. Por meio da inferência da relação sinal-ruído (SNR) instantânea e quando em boas condições de SNR, dá-se maior ênfase nas estimativas DOA. Para SNR baixas, as estimativas DOA têm menor importância

e a estimativa final considera com maior peso a informação *a priori* da fonte de sinal. Considera-se que é conhecida a função de densidade de probabilidade da estimativa DOA e que ambos sinal e ruído seguem a distribuição Guassiana estacionária descorrelacionada de média zero.

DO e SILVERMAN [17] apresentam um método de cálculo para o algoritmo SRP-PHAT para três dimensões de menor custo computacional: SPF (*stochastic particle filtering*). O método apresentado tem custo computacional da ordem de 0,03% da procura *full-grid*, o que torna esta solução uma candidata para implementação SRP-PHAT em sistemas de tempo real. Introduzem, também, uma variação da transformada tradicional PHAT, B-PHAT.

Segundo os autores, a procura pela direção de uma fonte de ondas pela estimação DOA em um espaço 3D passa pela determinação do máximo de uma função multivariável não-linear com vários máximos locais. A varredura em todos os pontos do espaço possíveis (*full-grid search*) torna a implementação do método SRP-PHAT inviável para sistemas em tempo real. Como forma de minimizar o custo computacional, o método SPF utiliza as premissas do método SRC (*Stochastic Region Contraction*) para selecionar pontos de avaliação da função a ser maximizada no espaço que possuem alta probabilidade de indicarem o máximo global com uma probabilidade desprezível de erro [17].

Outra modificação proposta por Do e Silverman é a introdução de um fator beta na transformada de fase (B-PHAT) de forma a representá-la como:

$$\beta - PHAT \ X\Psi_{ij}(k) \equiv \frac{1}{|X_i(k)X_j(k)|^\beta} \quad (3.41)$$

O fator B faz com que a função resultante SRP-PHAT seja suavizada e máximos locais de correlação correspondentes ao ruído sejam amenizados [17].

AZIMI-SADJADI *et al.* [2] apresentam métodos para rastreamento de veículos e grupos de veículos por um arranjo circular de microfones. Utilizam três variações diferentes do método CAPON para analisar o comportamento do estimador DOA para sinais de banda larga. Como forma de estimar a direção de chegada de sinais banda-larga com o algoritmo CAPON (banda estreita), são geradas sub-bandas de frequência pela transformada de Fourier discreta (DFT). A partir daí, diferentes métodos de ponderação das direções de chegada de sub-bandas de frequência são utilizados, como a média aritmética, a média geométrica e a média harmônica.

### 3.5.2 ESPRIT

O método ESPRIT (*Estimation of Signal Parameters via Rotational Invariance Technique*) oferece robustez e eficiência espectral [28, 32]. Este método apresenta resultados mais precisos para a estimação dos parâmetros do ruído quando comparado ao MUSIC, o que se reflete numa melhora significativa da estimação DOA comparativamente a esse.

Neste caso, é utilizado uma versão de arranjo linear de sensores modificado em relação aos demais métodos. No ESPRIT os sensores são compostos por *doublets*. Trata-se de um par de sensores  $x_i$  e  $y_i$  separados por uma distância constante  $\Delta$ , que possuem a mesma orientação. Esse vetor garante a invariância espacial entre os sinais dos diversos pares. A figura 3.39 apresenta uma descrição deste arranjo.

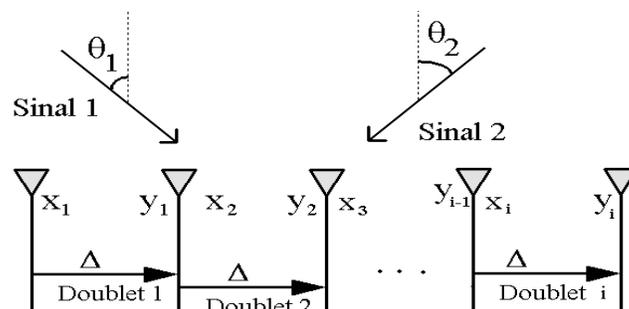


Figura 3.39: Arranjo de Doublets

O arranjo pode ser dividido em dois sub-arranjos  $Z_x$  e  $Z_y$ , idênticos e separados pelo vetor de deslocamento. Assim, as saídas  $X(t)$  e  $Y(t)$  dos dois arranjos são representadas pelas equações (3.41) e (3.42):

$$X(t) = \mathbf{A}S(t) + N_x(t) \quad (3.42)$$

$$Y(t) = \mathbf{A}\Phi S(t) + N_y(t) \quad (3.43)$$

Em que  $\Phi$  é a matriz diagonal que representa o defasamento entre os dois arranjos.

Dessa forma, a saída completa é a matriz  $Z(t)$  da equação 3.43 que pode ser representada pela equação 3.44.

$$Z(t) = \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{A}\Phi \end{bmatrix} S(t) + \begin{bmatrix} N_x \\ N_y \end{bmatrix} \quad (3.44)$$

$$Z(t) = \mathbf{A}S(t) + N_z(t) \quad (3.45)$$

Esse modelo tem a vantagem de não exigir conhecimento da resposta espacial do arranjo. De forma semelhante ao MUSIC, calculam-se os autovalores das matrizes de correlação dos dois arranjos, encontrando as matrizes diagonais  $\mathbf{E}_x$  e  $\mathbf{E}_y$  dos autovalores. Estas estão relacionadas por uma única matriz de transformação não singular, como apresentado na equação 3.43.

São os autovalores ( $\Phi_i$ ) dessa última matriz que permitem estimar a DOA como apresentado na equação 3.46.

$$\mathbf{E}_x \Psi = \mathbf{E}_y \quad (3.46)$$

$$\theta = \sin^{-1} \left\{ \frac{\lambda}{2\pi f \Delta} \arg(\Phi_i) \right\} \quad (3.47)$$

Por último, para estimar  $\Psi$ , um dos métodos mais comuns é o TLS-ESPRIT, apresentado na equação 3.48.

$$\Psi = E_y E_x^\dagger^{-1} \quad (3.48)$$

O ESPRIT é capaz de estimar a DOA com boa resolução, sendo também bastante robusto ao ruído e dispensando a maximização da função de densidade de potência espectral, uma vez que utiliza uma expressão analítica para estimar a direção de chegada, vide equação 3.46. Suas desvantagens são a complexidade e problemas quando os sinais são correlatados.

Algoritmos como o MUSIC tem uma resolução melhor por serem do tipo ML. Porém, uma das vantagens do ESPRIT é o fato de ter um custo computacional menor que técnicas como MUSIC e CAPON. Além disso, o ESPRIT tem uma complexidade de duas EVDs que é considerada adequada para implementações em tempo real com um número de sensores não muito alto.

## 4 ESTIMAÇÃO DOA

Neste capítulo serão apresentados o *hardware* de aquisição e uma implementação do algoritmo SRP-PHAT para estimação robusta DOA em tempo real.

Será explicitado como foi desenvolvido o sistema completo de rastreamento da direção de chegada de sinais sonoros por arranjo de microfones. O capítulo está dividido em duas partes essenciais e complementares: *hardware* e implementação de algoritmo de estimação DOA (*DOA tracking*).

Como forma de prover uma visão geral da estimação da direção de chegada, considere o fluxograma de etapas necessárias para a implementação de um sistema de rastreamento DOA, a ser detalhada nas próximas seções:

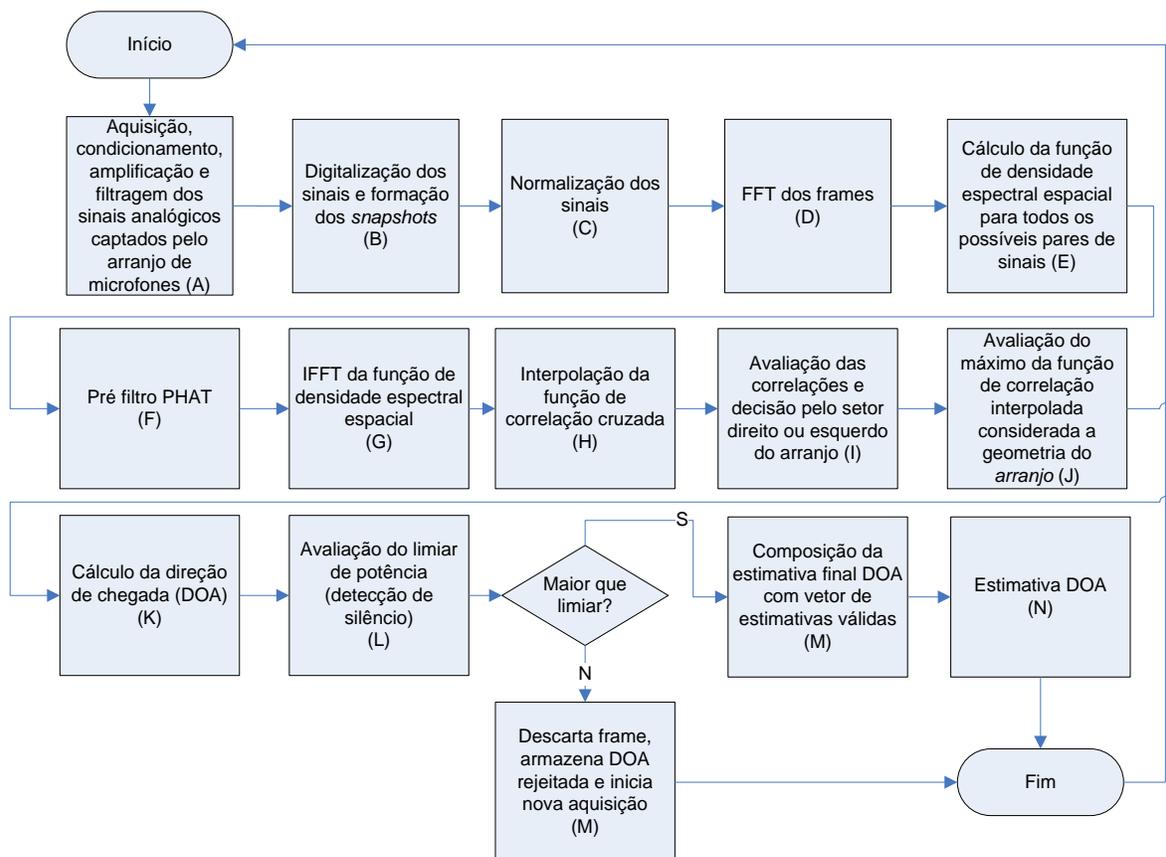


Figura 4.1: Visão geral do sistema de estimação DOA

No decorrer das próximas seções será referenciado cada bloco constituinte da estimação DOA pelos seus índices alfabéticos (A,B,...,M).

#### **4.1 HARDWARE: AQUISIÇÃO, CONDICIONAMENTO E AMPLIFICAÇÃO**

Nesta seção serão apresentadas as premissas e condicionantes para o desenvolvimento do *hardware* de aquisição, condicionamento, amplificação e filtragem dos sinais analógicos captados pelo arranjo linear de microfones. Gontijo e Costa (2007) iniciaram o desenvolvimento deste hardware [22].

As premissas e condições que devem ser atendidas são:

- Sensores e componentes idênticos;
- Amplificação linear de altíssima fidelidade, baixa distorção e *offset* DC desprezível;
- Amplificação com fase linear: atraso de grupo constante em toda a banda passante;
- Filtragem analógica do tipo passa-baixa com a máxima linearidade de fase;
- Alta relação sinal-ruído (SNR);
- Ruído de alimentação (correlatado) menor possível.

##### 4.1.1 Projeto de Hardware – (A)

###### **Microfones (sensores)**

Os microfones utilizados no arranjo são cápsulas de eletreto omnidirecionais, de 6 *mm*. Abaixo, apresenta-se a curva de resposta em frequência para este componente:

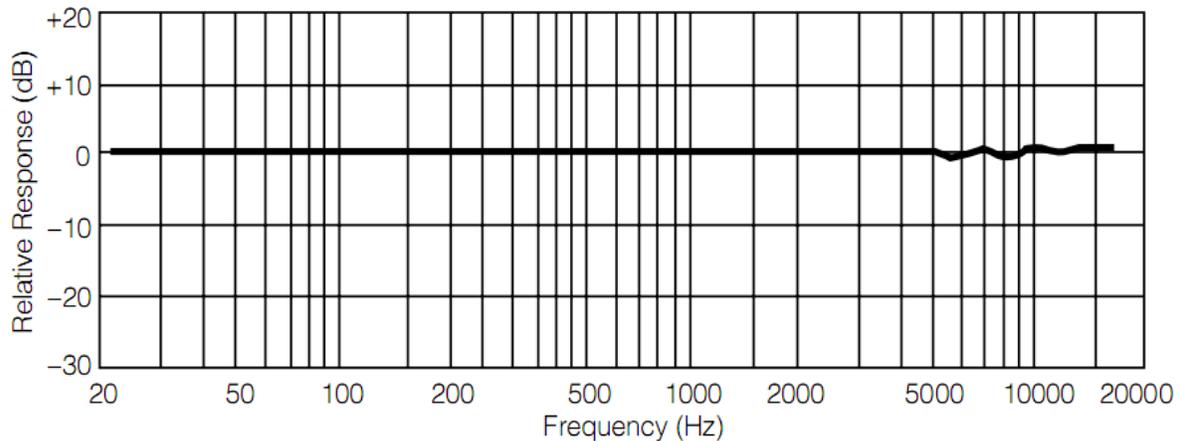


Figura 4.2: Curva de resposta de magnitude x frequência para o eletreto Panasonic WM-61A

A relação sinal-ruído (SNR) para as condições de operação  $V_s = 2,0\text{ V}$  e  $R_L = 2,2\text{ k}\Omega$  ohms é de  $62\text{ dB}$ . A sensibilidade é de  $-35 \pm 4\text{ dB}$  para  $1\text{ kHz}$ .

Esse microfone é escolhido pela alta linearidade da resposta em frequência, pela sensibilidade e relação sinal-ruído adequada à aplicação.



Figura 4.3: Microfones WM-61A

## Amplificadores

Os amplificadores utilizados para os circuitos de amplificação e filtragem analógica são os chips LME49740 (1), *National Semiconductor*. Cada chip possui quatro amplificadores operacionais. As principais especificações desse amplificador estão listadas a seguir na Figura 4.4:

(1) <http://www.national.com/pf/LM/LME49740.html>

## Key Specifications

■ Power Supply Voltage Range	$\pm 2.5V$ to $\pm 17V$
■ THD+N ( $A_V = 1$ , $V_{OUT} = 3V_{RMS}$ , $f_{IN} = 1kHz$ )	
$R_L = 2k\Omega$	0.00003% (typ)
$R_L = 600\Omega$	0.00003% (typ)
■ Input Noise Density	$2.7nV/\sqrt{Hz}$ (typ)
■ Slew Rate	$\pm 20V/\mu s$ (typ)
■ Gain Bandwidth Product	55MHz (typ)
■ Open Loop Gain ( $R_L = 600\Omega$ )	140dB (typ)
■ Input Bias Current	10nA (typ)
■ Input Offset Voltage	0.1mV (typ)
■ DC Gain Linearity Error	0.000009%

Figura 4.4: Especificações LME49740 segundo seu fabricante

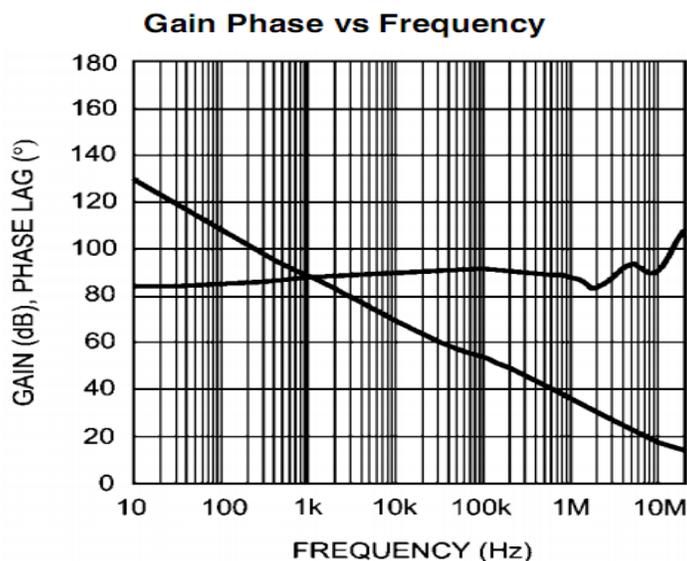


Figura 4.5: Curva de fase e atraso de grupo LME49740

Uma das principais características do LME49740 é sua linearidade de fase, conforme Figura. 4.5. Esta é uma especificação essencial para o amplificador utilizado na construção do circuito para estimação da direção de chegada, visto que a distorção de fase introduzida pelo canal elétrico altera a estimativa de fase para o sinal captado pelo arranjo.

Isto posto, todo o projeto do circuito elétrico será condicionado de forma que o canal elétrico resultante possua a maior linearidade de fase possível, implicando na utilização de componentes de uso profissional, com alta precisão.

### **Estágio de amplificação**

O sinal gerado pelos microfones de eletreto possui baixa amplitude, geralmente na faixa de  $\pm 5$  mV. Além disso, a impedância característica do microfone é da ordem de  $2k\Omega$ .

Deste modo, o estágio de amplificação deve ser projetado de forma a prover ganho elevado ao mesmo tempo em que mantém a impedância de entrada do circuito elevada, de forma a manter a integridade do sinal captado pelos eletretos do arranjo.

Com vistas a atender as restrições anteriormente citadas, foi escolhido o modelo de amplificador de instrumentação diferencial, por este possuir: níveis mínimos de *offset* DC; impedância de entrada elevada; distorção de fase mínima e ganho diferencial, compatibilidade com as características do microfone e capacidade de eliminar ruído comum entre as entradas (*CMRR* elevado).

Além disso, as características desse tipo de circuito permitem que este possua elevados ganhos com boa linearidade. Os dois amplificadores na configuração *buffer* propiciam alta impedância de entrada. O diagrama esquemático para o estágio de amplificação está representado na figura 4.6.

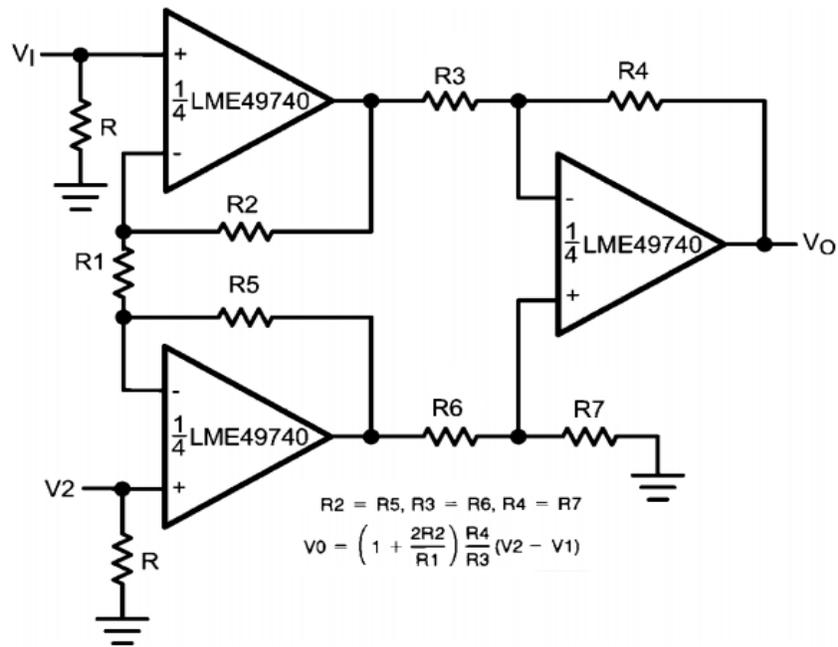


Figura 4.6: Diagrama esquemático do estágio de amplificação

### Filtragem analógica

A etapa de filtragem analógica servirá para suprimir eventuais ruídos de alta frequência indesejados, além de prover um mecanismo *anti-aliasing* antes da etapa de digitalização. É escolhido o tipo de filtro passa-baixos, com frequência de corte 5,5 kHz como forma de restringir a banda de frequência dos sinais captados na faixa de voz, que será a faixa de interesse para o presente estudo.

A resposta de fase para filtros ativos geralmente não é linear, o que traz a necessidade de se projetar o filtro que ofereça a máxima linearidade de fase possível. Para que tenhamos maior linearidade de fase e atraso de grupo constante em toda a banda passante, foi escolhida a classe de filtros de Bessel. Dentre as possíveis configurações de filtro, os filtros tipo Bessel são os que oferecem maior linearidade de fase.

Os filtros passa-baixas de Bessel têm a função de transferência característica

$$H(s) = \frac{\theta_n(s)}{\theta_n\left(\frac{s}{\omega_0}\right)} \quad (4.1)$$

Na qual  $\theta_n(s)$  é um polinômio de Bessel e  $\omega_0$  é a frequência de corte do filtro. Para essa classe de filtros, o atraso de grupo nas frequências próximas a zero é dada por  $\frac{1}{\omega_0}$ . A função de transferência de um filtro de Bessel é uma função racional em que o denominador é um polinômio de Bessel:

$$\theta_n(s) = \sum_{k=0}^{\infty} a_k s^k \quad (4.2)$$

$$a_k = \frac{(2n-k)!}{2^{n-k} k! (n-k)!}, k = 0, 1, \dots, n \quad (4.3)$$

Um filtro Bessel passa-baixos de ordem  $n$  tem a característica que todos os primeiros  $n - 1$  termos na expansão da série de Taylor para a expressão do atraso de grupo são zeros, o que provê a máxima constância para o atraso de grupo para valores próximos de  $\omega = 0$ . Esta é uma característica fundamental para o *hardware* de estimação DOA, uma vez que a faixa de frequências de interesse se estende de 0Hz a 5.5 kHz e deseja-se *group delay* constante em toda a banda passante.

Para a faixa de frequência considerada e  $fc = 5.5 \text{ kHz}$ , as curvas de magnitude e fase para um filtro de Bessel, implementado com a topologia *Sallen-Key*, ordem 2, são apresentadas abaixo:

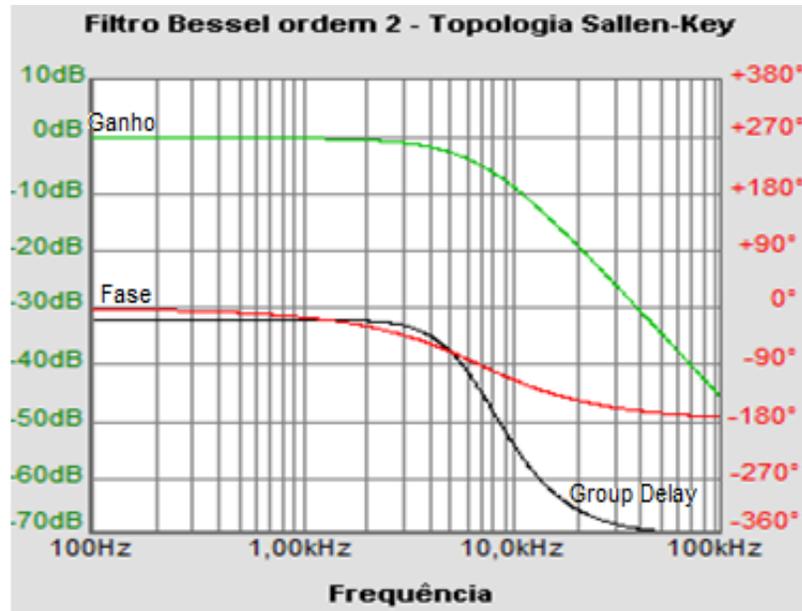


Figura 4.7: Resposta de magnitude, fase e atraso de grupo para o filtro Bessel ordem 2

Pela curva do atraso de grupo, percebemos que este é praticamente constante em toda a banda passante. Quanto à magnitude, é requerido que o filtro possua ganho unitário, uma vez que a combinação de uma configuração de ganho juntamente com a filtragem analógica introduziria o indesejado efeito de fase não linear.

A topologia para implementação do circuito escolhida é a *Sallen-Key* por sua simplicidade.

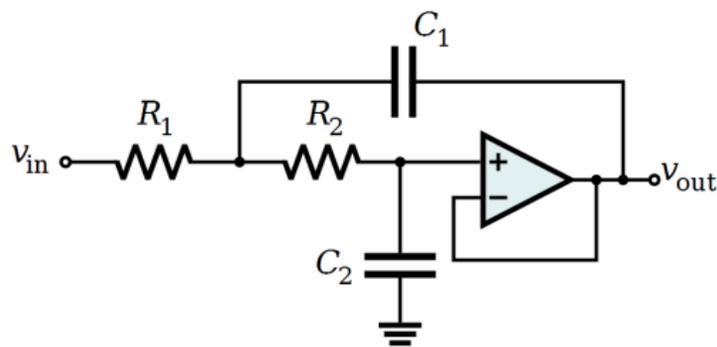


Figura 4.8: Topologia Sallen-Key

Para a filtragem de sinais oriundos dos microfones de eletreto, é necessário que haja alta impedância de entrada de forma a não distorcer o sinal captado pelo arranjo. Esta

topologia é uma versão adaptada da topologia VCVS (*voltage-controlled voltage-source*) e tem como principal característica o uso de um amplificador operacional na configuração *buffer*, o que provê alta impedância de entrada no circuito de filtragem.

Para esta configuração, têm-se a função de transferência:

$$H(s) = \frac{\overbrace{(2\pi f_c)^2}^{\omega_c^2}}{s^2 + \underbrace{2\pi \frac{f_c}{Q}}_{\frac{\omega_c}{Q} = 2\zeta\omega_c} s + \underbrace{(2\pi f_c)^2}_{\omega_c^2}} \quad (4.4)$$

Em que:

$$s = j\omega = (\sqrt{-1}) 2\pi f \quad (4.5)$$

$$f_c = \frac{1}{2\pi \sqrt{R_1 R_2 C_1 C_2}} \quad (4.6)$$

$$Q = \frac{\sqrt{R_1 R_2 C_1 C_2}}{C_2 (R_1 + R_2)} \quad (4.7)$$

Para a implementação proposta do *hardware*, temos  $f_c = 5.5 \text{ kHz}$  e  $Q = 577,35$ , em que  $Q$  representa a largura e altura do pico da resposta em frequência do filtro. Todos os componentes passivos utilizados na montagem do filtro ativo são componentes de tolerância 1%. Esta é uma medida necessária para que se minimize as diferenças de fase entre os canais adquiridos pelo circuito digitalizador (homogeneidade dos canais elétricos). Uma diferença de fase introduzida pelo canal elétrico pode resultar em uma alteração da estimativa da direção de chegada, como veremos na seção posterior.

Como forma de estimar o comportamento real que será obtido após a etapa de filtragem analógica, foi realizada simulação no software *CircuitMaker* considerando um tom senoidal de  $1\text{ kHz}$  como entrada cuja amplitude pico-a-pico  $V_{pp} = 1\text{ mV}$ .

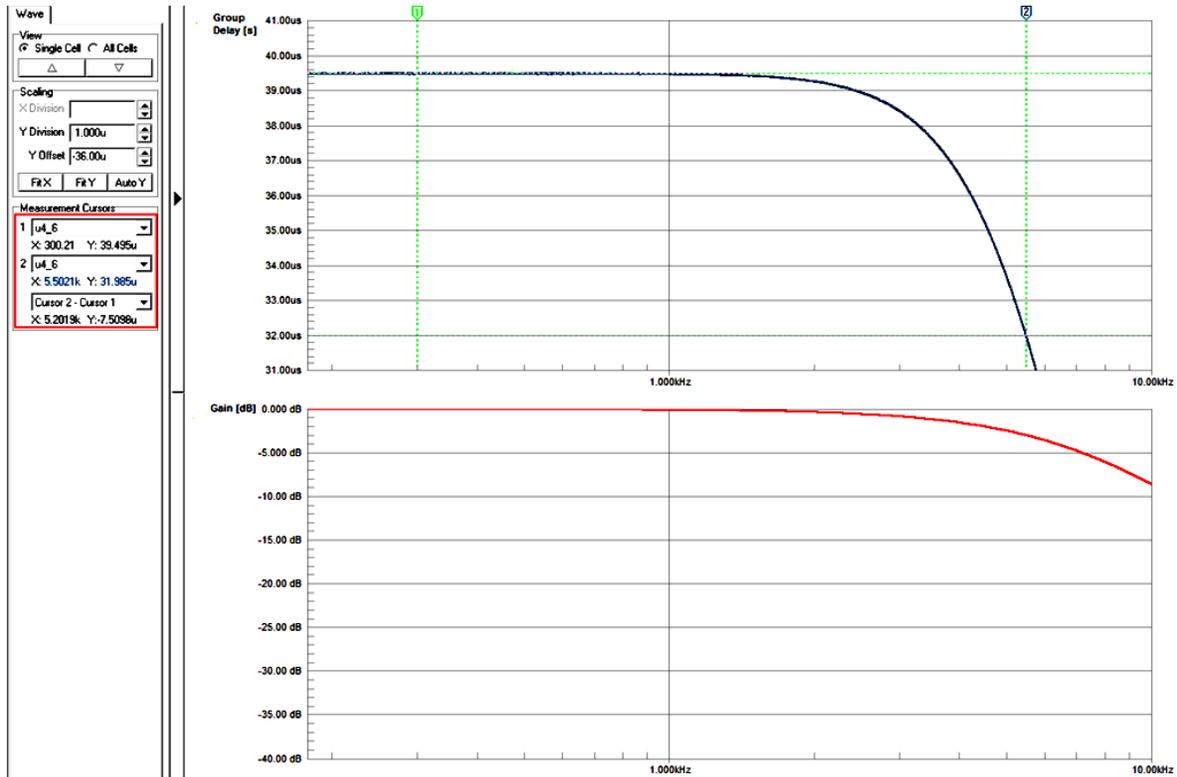


Figura 4.9: Simulação para o bloco de filtragem analógica

Note que o atraso de grupo introduzido entre a faixa de  $300\text{ Hz}$  e  $5\text{ kHz}$  é de apenas  $7,5\text{ }\mu\text{s}$ . Ainda, na banda  $5\text{ kHz}$  o atraso é de menos de  $1\text{ }\mu\text{s}$ . Considera-se que o atraso de grupo devido ao bloco de amplificação pode ser desprezado, uma vez que todos os elementos que compõem o circuito, exceto amplificadores operacionais, são resistivos.

### Alimentação DC

Os chips LME49740 são alimentados por fontes DC simétricas, que podem variar entre  $\pm 2,7\text{ V}$  e  $\pm 15\text{ V}$ . Para que não se introduza um ruído no sinal de interesse, foram projetadas fontes de  $\pm 9\text{ V}$  estabilizadas por reguladores de tensão da família TL7X09.

Ambos circuitos para construção das fontes de tensão positivas e negativas contam com uma ponte de diodos retificadora completa, capacitores de estabilização, LED's de indicação, proteção contra corrente reversa e filtro de ruído.

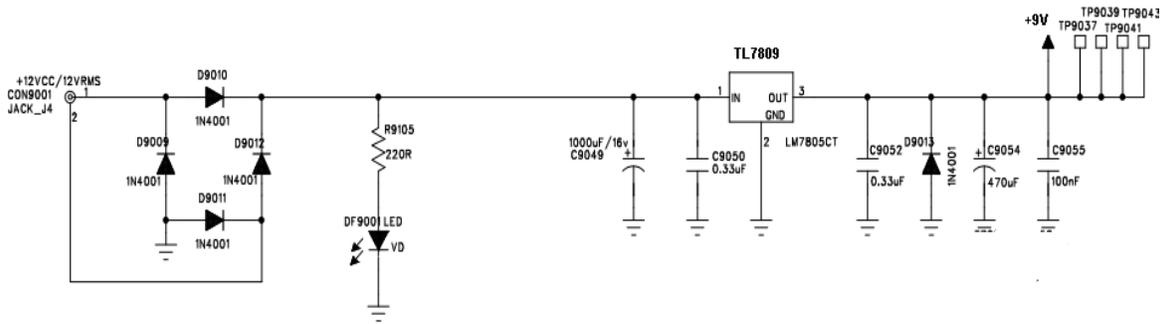


Figura 4.10: Circuito para fonte de tensão +9V regulada

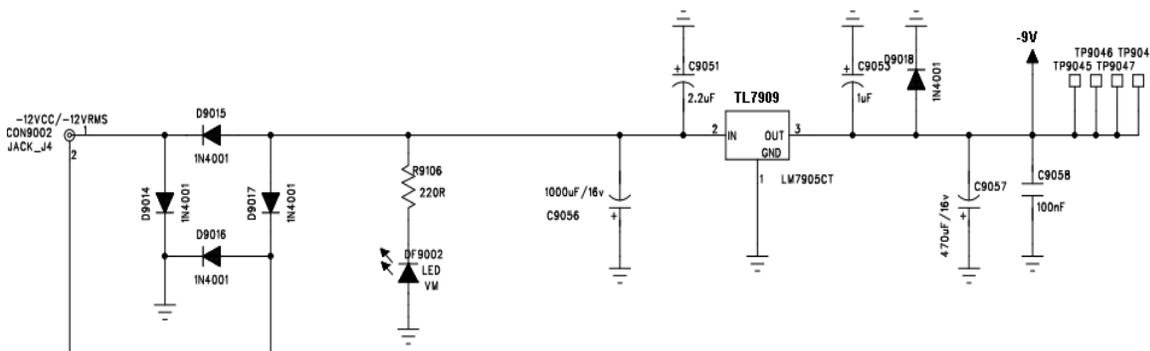


Figura 4.11: Circuito para fonte de tensão -9V regulada

Uma das premissas para o êxito da estimação DOA é a inexistência de ruído correlatado nos canais de aquisição. Ou seja, deve-se ter boa relação sinal-ruído (SNR) e o ruído, quando houver, deve ser não correlacionado.

Ocorre que o ruído que fontes de alimentação introduzem são, via de regra, correlatados e, para o caso da estimação DOA, ruídos correlatados são interpretados como sinais altamente correlacionados oriundos da direção angular  $0^\circ$ . Deste modo, é imperativo que o circuito de alimentação não introduza ruído nos canais de aquisição de dados, sob o risco de todo o sistema de estimação ter um viés forte e permanente para o ângulo  $0^\circ$ .

## Digitalização – (B)

Uma vez adquirido, amplificado e filtrado o sinal analógico, este é digitalizado por meio do dispositivo USB 1608-FS, *Measurement Computing*. A principal característica desse dispositivo é a digitalização com *sample-and-hold* (um conversor A/D por canal, para 8 canais), o que é fundamental para estimação da direção de chegada visto que os sinais são amostrados todos simultaneamente, o que evita que distorções de fase e atraso sejam introduzidas.



Figura 4.12: Dispositivo de digitalização USB-1608FS

Características do dispositivo:

- *Sampling* de 8 canais simultâneos;
- Input de  $\pm 1 V$  a  $\pm 10 V$ ;
- 16 bits de resolução;
- Taxa de amostragem máxima:  $(200 \text{ kS/s}) / (\# \text{ de canais})$ ;
- Crosstalk:  $-80 \text{ dB}$ ;
- Impedância de entrada:  $100 \text{ M}\Omega$ , mínimo;
- Acurácia:  $2,98 \text{ mV}$  na faixa de  $\pm 5 V$ ;
- USB 2.0;
- Interface com *MATLAB*<sup>®</sup> (2).

As especificações estão aderentes às necessidades do projeto para estimação DOA.

O amostragem simultânea garante a integridade de fase, a resolução de 16 bits e a acurácia

(2) funções MATLAB: `getdata()`

são adequados para o cálculo da correlação cruzada, os níveis de *crosstalk* são mínimos e pode-se, pelo MATLAB<sup>®</sup>, adquirir os 8 canais simultaneamente em tempo real.

Pela limitação de *throughput* do dispositivo em [(200 kS/s)/(# de canais)] e limitação do *buffer* físico, com a utilização de 7 microfones, conseguimos uma taxa máxima de amostragem  $f_s = 14,285 \text{ kHz}$ . Esta taxa permite a aquisição de sinais com banda passante de aproximadamente 7 kHz, valor suficiente para os sinais sonoros utilizados nesse trabalho.

#### 4.1.2 Placa para Amplificação e filtragem

Consideradas as restrições e condicionantes da seção anterior, foi projetado um protótipo de placa de circuito impresso com alimentação, amplificação e filtragem analógica dos sinais.

Na Figura 4.13, apresentamos o *layout* PCB da placa, destacando os componentes que formam o conjunto alimentação (vermelho) e o conjunto amplificação e filtragem analógica (verde). O projeto da placa foi realizado para amplificação simultânea de 8 canais. Os sinais de dados trafegam no sentido horizontal da face da placa, enquanto os sinais de alimentação e terra estão no sentido vertical pelo verso.

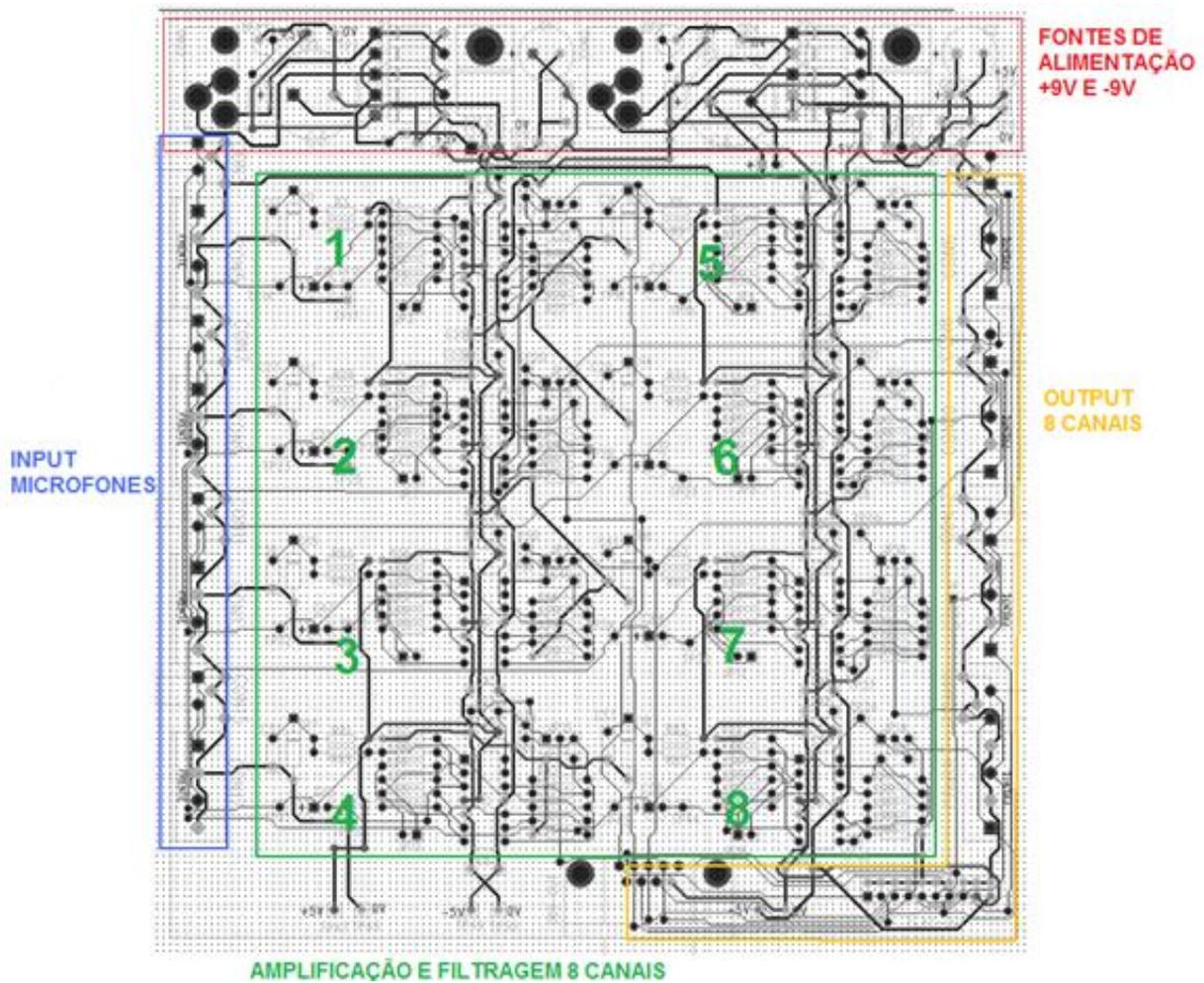


Figura 4.13: Placa de circuito impresso para amplificação e filtragem

## 4.2 IMPLEMENTAÇÃO DA ESTIMAÇÃO DOA

Nesta seção serão descritas todas as etapas do processo de estimação da direção de chegada (DOA) por arranjo de sensores após a amplificação e filtragem dos sinais analógicos.

As letras referem-se ao esquema geral para estimação DOA - Figura. 4.1.

### Digitalização, normalização e FFT dos *frames* - (B),(C),(D)

A digitalização dos sinais é comandada pelo MATLAB<sup>®</sup> e realizada pelo dispositivo USB-1608FS. O ruído ambiente captado se traduz em um sinal de aproximadamente 100 mV, enquanto um sinal sonoro captado tem amplitude pico-a-pico de aproximadamente 2V.

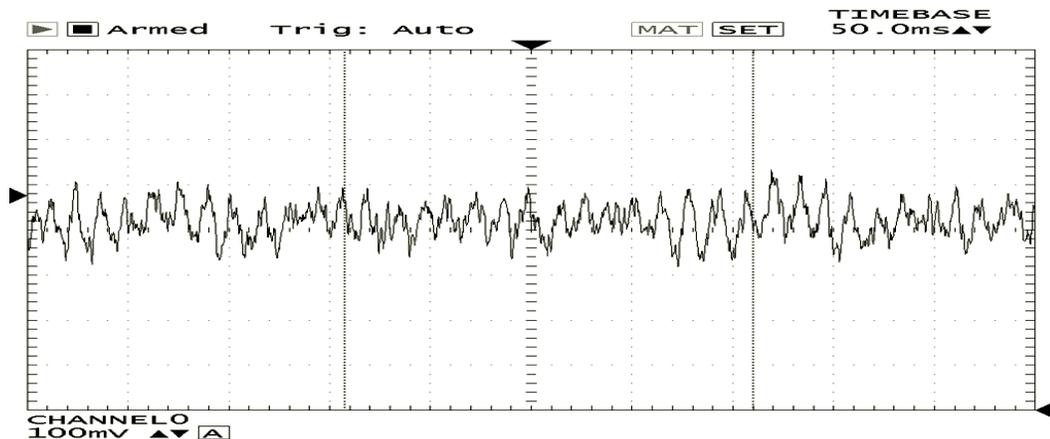


Figura 4.14: Ruído ambiente com escala de 100 mV

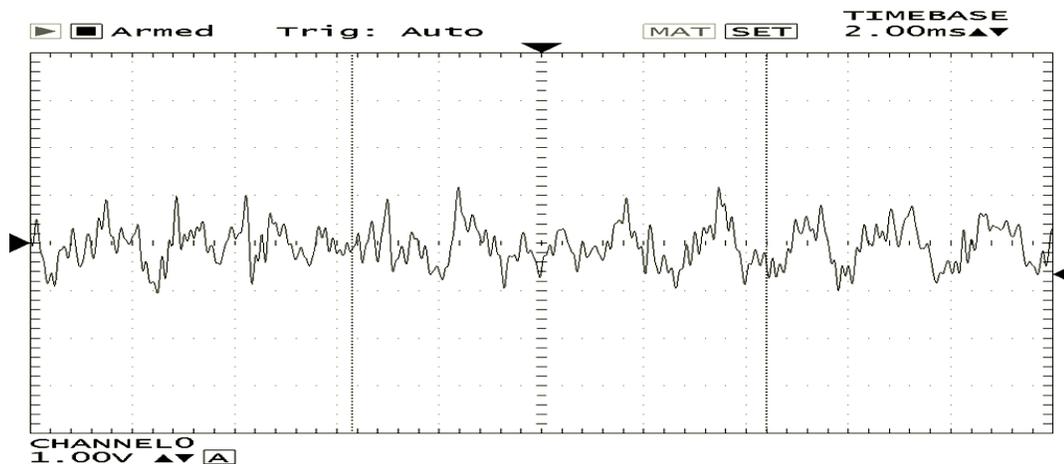


Figura 4.15: Sinal sonoro com escala de 1 V

Após a captura, os *snapshots* instantâneos são formados e cada estimativa da direção de chegada será realizada considerando um conjunto de  $N = 5.000$  amostras. Os *snapshots* são normalizados para que a amplitude não influencie no cálculo da correlação cruzada e são gerados vetores com a FFT de cada canal.

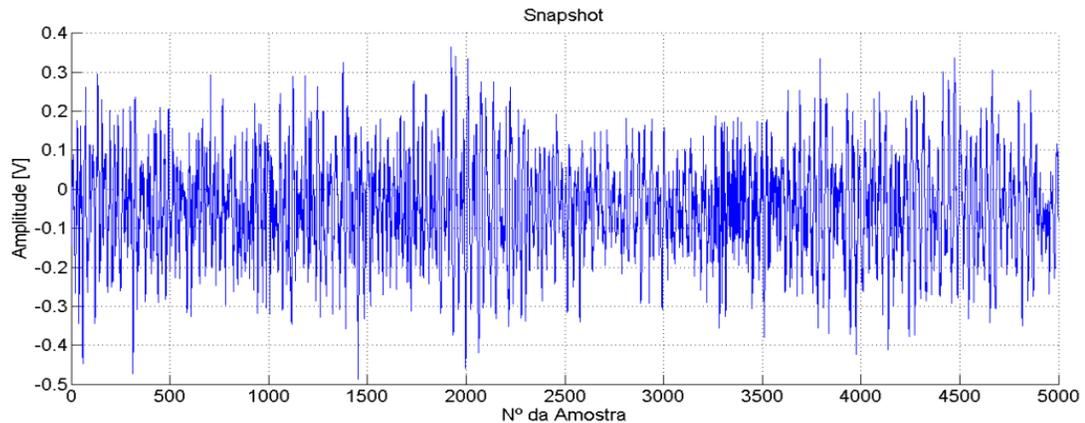


Figura 4.16: *Snapshot* típico instantâneo de um microfone

Pelo fato da estimação da direção de chegada estar baseada num cálculo de correlação entre sinais de diferentes canais, quanto mais amostras disponíveis para o cálculo da correlação, mais precisa esta será. Entretanto, para a aplicação de métodos como DS e CAPON a voz deve ser analisada em um intervalo curto, no qual pode-se considerar que o sinal é estacionário. Para sinais sonoros este intervalo se encontra entre 20 e 40 *ms*.

Além disso, o aumento do número de amostras do *snapshot* aumenta consideravelmente o esforço computacional necessário para o rastreamento em tempo real da fonte sonora. O número de 5000 amostras foi escolhido após estudo da dispersão das estimativas em função do tamanho do quadro, conforme Figura 4.17.

Embora haja uma convergência crescente das estimativas para quadros maiores, há que se considerar que quanto maior o quadro considerado, menor é a estacionariedade do sinal, o que pode implicar na degradação dos resultados obtidos com métodos de banda estreita. Para 5000 amostras, cada quadro tem a duração de 350 *ms*.

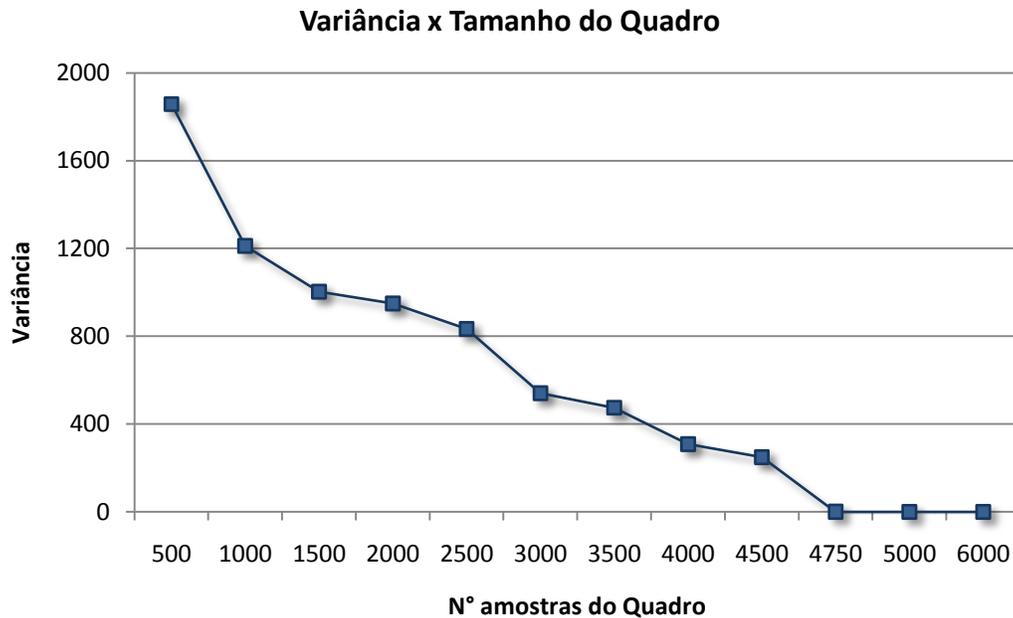


Figura 4.17: Dispersão das estimativas DOA em função do tamanho do snapshot

### **Cálculo da função densidade espectral espacial, pré-filtro PHAT, IFFT da função de densidade espectral espacial - (E,F,G)**

Nesta etapa é calculada a função de densidade espectral espacial. Esta função descreve o nível de correlação entre pares de sinais para um determinado atraso estipulado  $\tau$ . Cada valor da função de densidade espectral é obtido pelo cálculo da correlação cruzada entre os sinais no domínio da frequência [29].

No arranjo construído para este trabalho, existem sete microfones. Desta forma, existem 21 pares possíveis de microfones para o cálculo da correlação cruzada. A correlação cruzada é calculada para todos os pares possíveis, de forma a prover redundância e robustez à estimativa.

Apesar de já terem sido apresentadas em 3.5.1, apresentamos um resumo das principais equações para o cálculo da função de densidade espectral.

Para o termo de pré-filtragem PHAT, conforme equação 4.8:

$$\gamma_{ij}(k) = \frac{1}{|X_i(k)X_j^*(k)|} \quad (4.8)$$

E para o cálculo da correlação cruzada entre um par de sinais, conforme equação (4.9):

$$\bar{R}_{x_i x_j}(\tau) = \hat{R}_{y_i y_j}(\tau) = \frac{1}{N} \sum_{k=0}^{N-1} \gamma_{ij}(k) X_i(k) X_j^*(k) e^{j \frac{2\pi k \tau}{N}} \quad (4.9)$$

Abaixo, apresentamos a função de correlação cruzada, no domínio do tempo, para uma simulação com sinal sonoro e para um sinal sonoro captado no arranjo de sensores real, considerando-se que o sinal é proveniente de um ângulo que gera um atraso de 3 amostras entre microfones adjacentes 1 e 2:

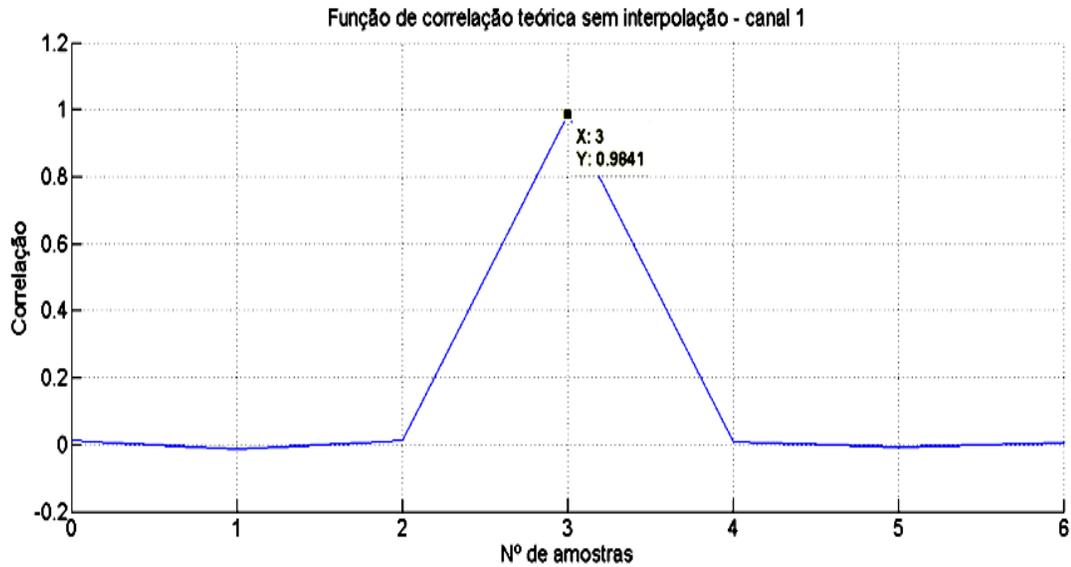


Figura 4.18: Correlação teórica simulada para um atraso de 3 amostras entre canais

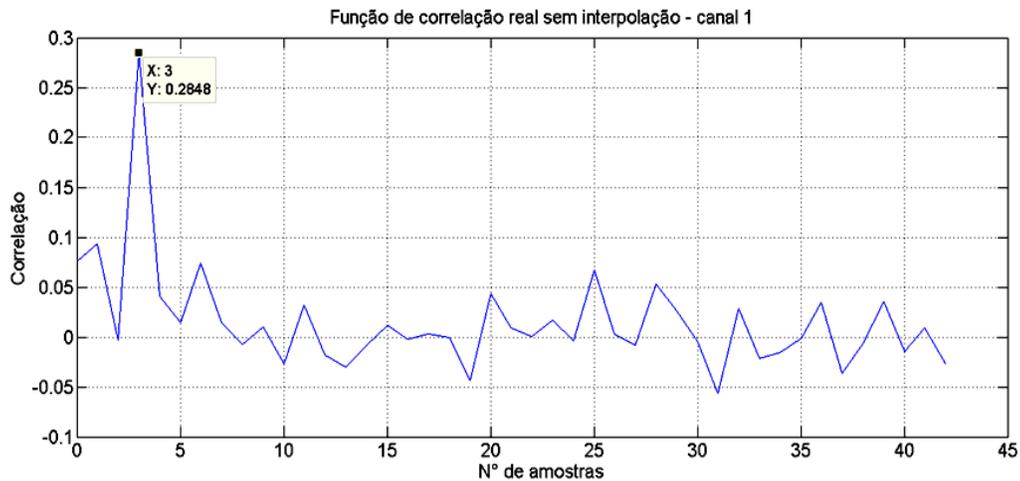


Figura 4.19: Correlação real obtida para um atraso de 3 amostras entre canais

Note que na situação teórica, quando não existe ruído ambiente, imperfeições dos canais elétricos ou reverberação, a correlação tem um pico claro e definido em  $\tau = 3$ . Já no ambiente real, surgem picos secundários que podem afetar a precisão da estimação da direção de chegada, mas mesmo assim foi possível estimar  $\tau = 3$ .

Para a mesma situação e para sinais de sensores que estão separados de dois espaçamentos, teríamos um atraso em  $\tau = 6$ . Neste caso, temos para a situação teórica e real respectivamente Figuras 4.20 e 4.21:

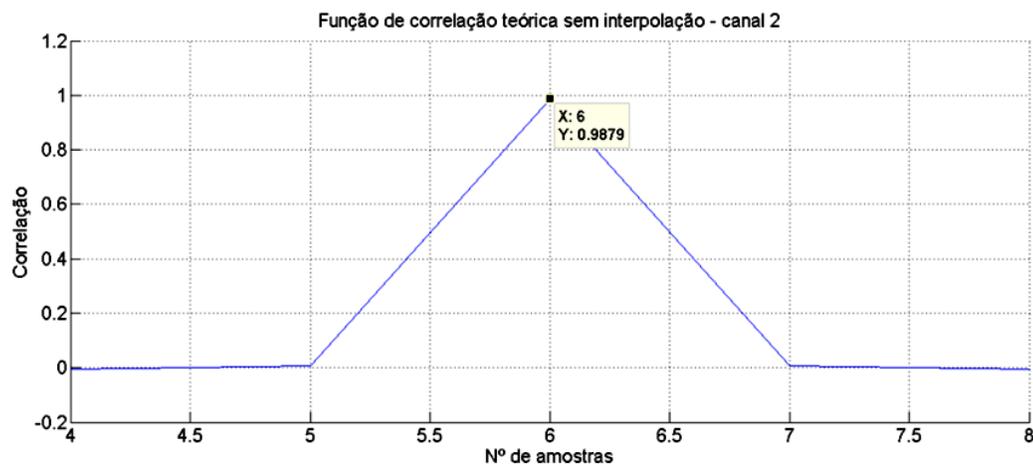


Figura 4.20: Correlação teórica simulada para um atraso de 6 amostras entre canais

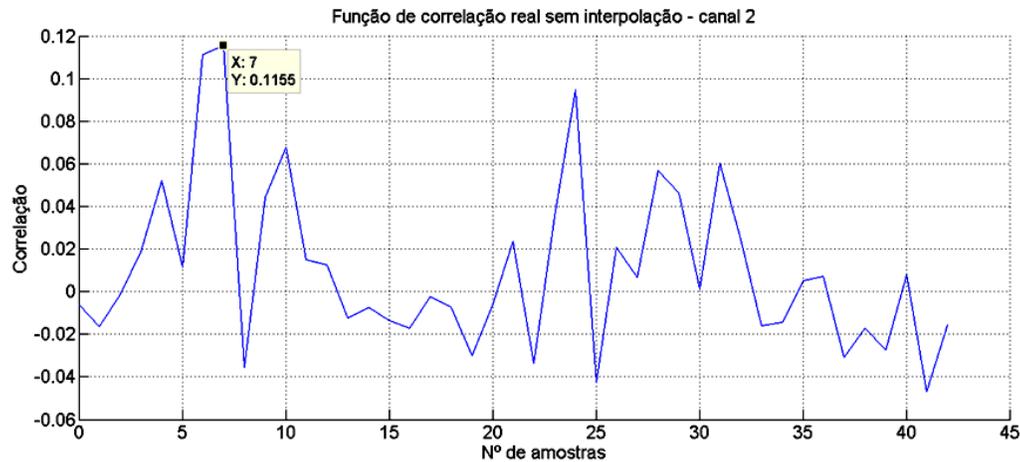


Figura 4.21: Correlação real obtida para um atraso de 6 amostras entre canais

Para microfones espaçados de  $2d$  no arranjo físico, além de haver um erro da estimativa do atraso  $\tau$  de uma amostra, encontra-se uma função de correlação cruzada real com picos falsos que têm cerca de 90% do valor do pico verdadeiro, o que reforça a necessidade de calcular-se a correlação para todos os pares possíveis, de forma a prover confiabilidade ao resultado.

### Interpolação da função de correlação cruzada (H)

No arranjo de microfones utilizado temos 7 elementos separados de uma distância  $d = 15 \text{ cm}$ . Considerando a velocidade da onda sonora como  $c = 346,1 \text{ m/s}$  e utilizando-se uma taxa de amostragem  $f_s = 14,285 \text{ kHz}$ , temos no mínimo 0 amostras ( $\text{DOA} = 0^\circ$ ) e no máximo 6 amostras ( $\text{DOA} = 76^\circ$ ) entre canais adjacentes.

$$n^\circ \text{ amostras delay} = \frac{d}{c} \cdot f_s \cdot \sin(\theta) \quad (4.10)$$

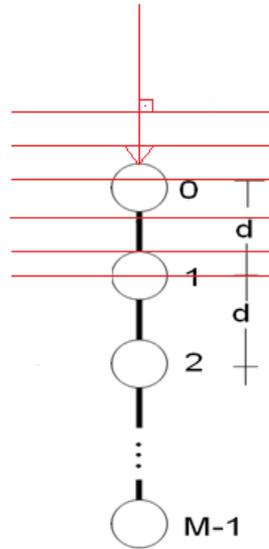


Figura 4.22: Máximo de atraso entre canais adjacentes: onda incidente em 90°

Desta forma, os picos da função de correlação cruzada poderiam ocorrer numa faixa de 6 valores discretos, o que levaria a uma precisão máxima de  $\frac{90}{6} = 15^\circ$  para a estimação DOA que não é uma precisão satisfatória para aplicação de rastreamento de sinais sonoros.

A solução adotada para incremento da precisão é a interpolação da função de correlação cruzada pelo método *Cubic Splines*, detalhado posteriormente. Escolhendo-se um fator de 10 vezes para a interpolação, têm-se que haverá 10 vezes mais pontos na função de correlação cruzada, o que significa dizer que a precisão da estimação DOA será elevada para  $\frac{90}{6 \times 10} = 1,5^\circ$ . Esta resolução do estimador DOA já pode ser considerada satisfatória visto que a própria fonte sonora, geralmente, possui dimensões maiores que um setor de  $1,5^\circ$  determinado.

### Cubic Splines

*Cubic Splines* [14] é um método de interpolação de valores por polinômios, contínuos por parte, de 3ª ordem. Considere uma tabela de pares de valores  $[x_i, y_i]$  para  $i = 0, 1, \dots, n$  para a função  $y = f(x)$ . Esses valores geram  $n + 1$  pontos e  $n$  intervalos entre

eles. A interpolação *cubic splines* é uma curva contínua por partes que passa por cada um dos valores da tabela. Cada intervalo possui um polinômio cúbico próprio, cada polinômio com seus próprios coeficientes:

$$S_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \text{ para } x \in [x_i, x_{i+1}]. \quad (4.11)$$

Juntos, esses segmentos de polinômios são denominados  $S(x)$  ou *spline*. A característica mais importante do método *Cubic Splines* é que os diversos polinômios que formam a curva interpolante final (*Spline*) são contínuos em seus limites, e como forma de garantir que a curva seja a mais suave possível, as derivadas de primeiro e segundo grau são contínuas nas fronteiras dos polinômios. Ou seja:

$$\begin{aligned} S_i(x_i) &= y_i, S_i(x_{i+1}) = y_{i+1} \\ S_{i-1}'(x_i) &= S_i'(x_i), S_{i-1}''(x_i) = S_i''(x_i) \end{aligned} \quad (4.12)$$

O resultado da imposição dessas condições é uma curva interpolante suave e contínua. Como veremos adiante, a interpolação *Splines* é adequada para a função de correlação cruzada.

A seguir, mostramos as mesmas funções de correlação para canais adjacentes e canais separados por dois espaçamentos com e sem interpolação de fator 10 do tipo *Cubic Splines*:

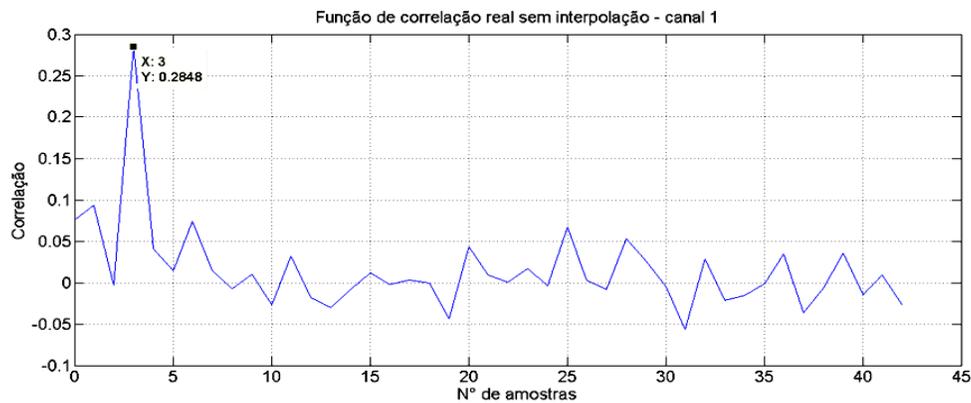


Figura 4.23: Correlação real obtida para um atraso de três amostras entre canais

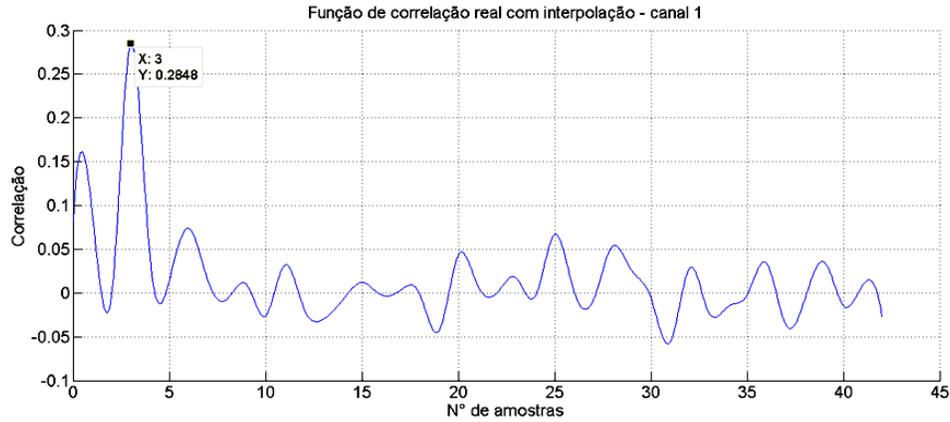


Figura 4.24: Correlação real com interpolação obtida para um atraso de três amostras entre canais

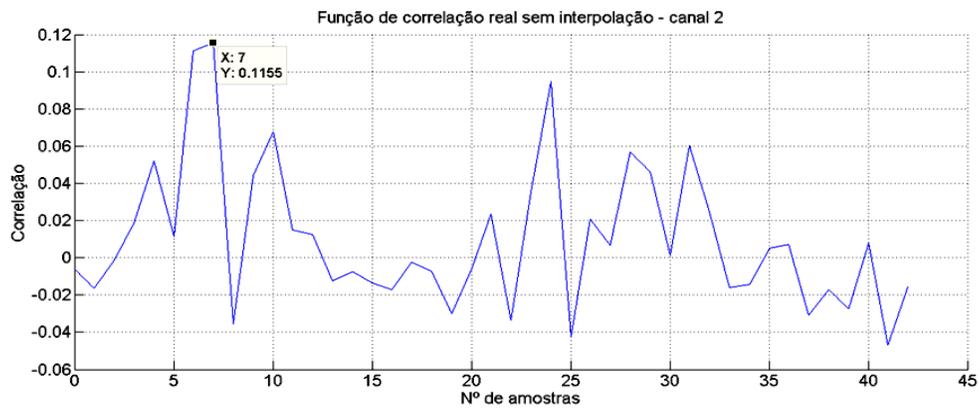


Figura 4.25: Correlação real obtida para um atraso de seis amostras entre canais

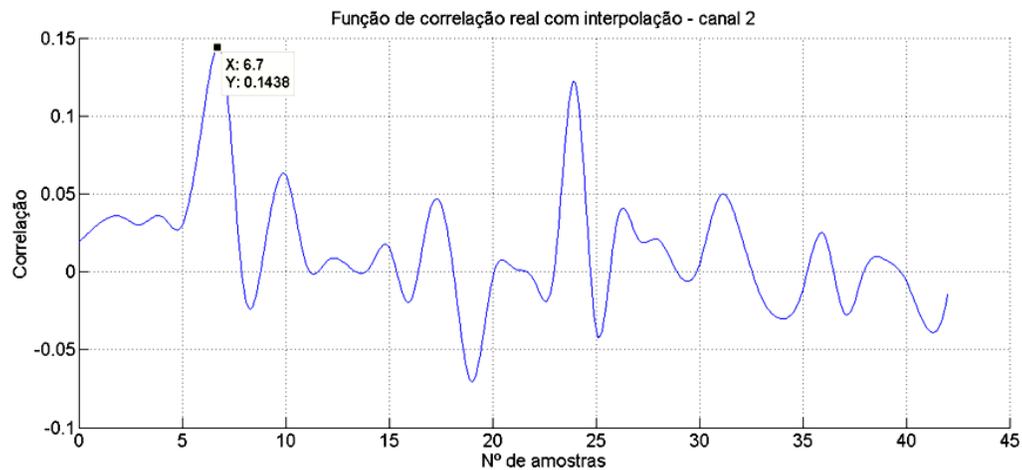


Figura 4.26: Correlação real com interpolação obtida para um atraso de seis amostras entre canais

Da análise dos gráficos com e sem interpolação, nota-se que é mantida a forma da curva da função de correlação cruzada com o interpolador *Cubic Splines*. Além disso, por

sua característica de continuidade, a interpolação por *Splines* não introduz descontinuidades na função de correlação, o que é fundamental para a estimação da direção de chegada. Tais descontinuidades poderiam levar a grandes erros de estimativa [14].

A interpolação provê um ganho substancial, de 10 vezes, no refinamento da estimativa do atraso  $\tau$  por meio da função de correlação, e conseqüentemente na precisão da estimação da direção de chegada. No entanto, o uso de interpoladores em aplicações de tempo real deve ser realizado com cautela uma vez que esses algoritmos introduzem uma carga computacional adicional considerável, o que pode degradar o desempenho do estimador como um todo.

#### Avaliação das correlações e decisão por setor direito ou esquerdo do arranjo (I)

A geometria linear do arranjo de microfones traz uma ambigüidade na estimativa da direção de chegada. Conforme vemos na figura abaixo, para ângulos iguais a  $\pm\theta$ , o atraso, em número de amostras, é o mesmo para  $+\theta$  e para  $-\theta$ .

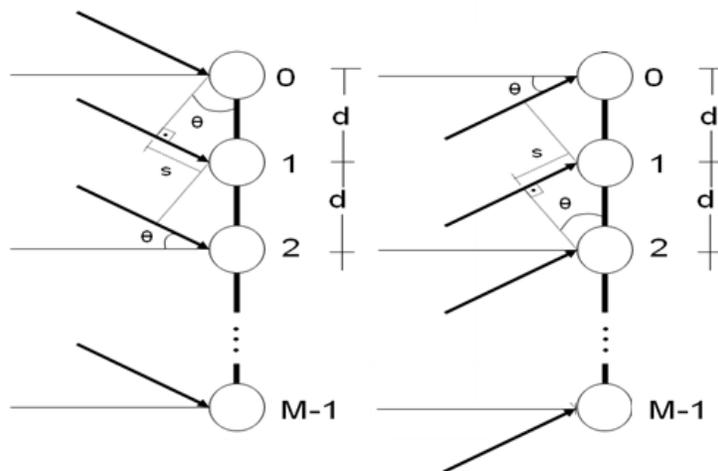


Figura 4.27: Ambigüidade na estimação DOA para o arranjo linear

Ou seja, tanto para  $+\theta$  quanto para  $-\theta$  a distância percorrida pela frente de onda  $s$  é igual. Isto significa que no cálculo da correlação cruzada, o atraso em número de amostras será igual para  $+\theta$  e para  $-\theta$ .

Esta ambiguidade é retirada do problema de estimação da direção de chegada fazendo-se o cálculo da correlação total entre os canais em duas etapas: primeiro, calcula-se a soma da correlação cruzada total entre os 21 pares possíveis de sinais considerando-se o microfone 1 como referência, armazena-se o valor; logo em seguida, calcula-se a soma da correlação cruzada total entre os 21 pares possíveis de sinais considerando-se o microfone 7 como referência e se compara os valores obtidos para o primeiro e segundo cálculo.

Deste modo, se a soma da correlação total é maior para o primeiro caso, a fonte está, com bastante confiabilidade, no setor à direita do eixo de referência perpendicular ao arranjo. Caso a soma para o segundo caso seja maior, a fonte de ondas está localizada no setor à esquerda do eixo referência.

#### **Avaliação do máximo da função de correlação interpolada considerada a geometria do arranjo e cálculo da direção de chegada DOA (J,K)**

Uma vez determinadas as 21 possíveis funções de correlação entre os pares de sinais, deseja-se formar, a partir das 21 estimativas para o atraso individual, uma estimativa final para o atraso  $\tau$  em amostras.

Considere as funções de correlação cruzada obtidas a partir da comparação do sinal obtido pelo sensor 1 com o sensor 2 (espaçamento igual a  $d$  e atraso igual a  $\tau$ ) e pelo cálculo a partir dos sinais obtidos pela comparação do sinal do sensor 1 com o sensor 3 (espaçamento igual a  $2d$  e atraso igual a  $2\tau$ ).

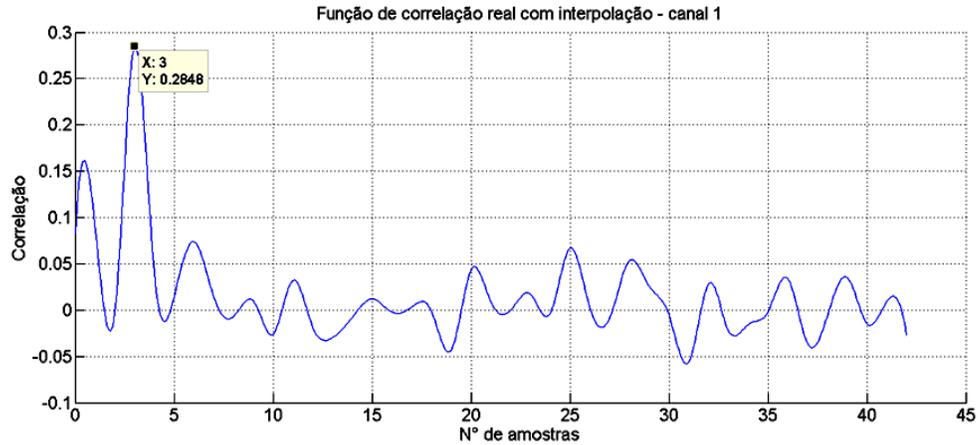


Figura 4.28: Correlação real com interpolação obtida para um atraso de três amostras entre canais

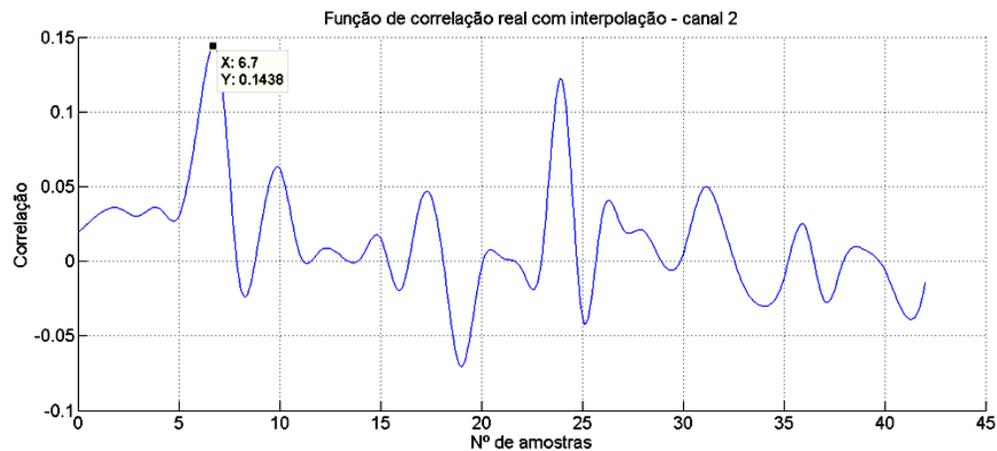


Figura 4.29: Correlação real com interpolação obtida para um atraso de 6 amostras entre canais

Observe que o atraso estimado é um múltiplo de  $\tau$ . Ou seja, para um espaçamento entre sensores de  $(n.d)$ ,  $n = 1, 2 \dots 6$ , temos para as funções de correlação cruzada um máximo em um múltiplo de  $\tau$ , conforme o espaçamento entre os pares de sinais em questão.

Para um arranjo de 7 microfones, haverá 1 estimativa com  $6\tau$  de atraso entre amostras, 2 estimativas com  $5\tau$ , 3 estimativas com  $4\tau$ , 4 estimativas com  $3\tau$ , 5 estimativas com  $2\tau$  e 6 estimativas com  $1\tau$  de atraso, e sabe-se exatamente qual função de correlação corresponde a qual par de sinais.

Sendo assim, devido à geometria linear do arranjo tem-se que o atraso em amostras das funções de correlação cruzada será sempre múltiplo do atraso unitário. Desta forma, para conjugar todas as funções de correlação cruzada em uma estimativa final apenas, basta que se avalie para cada conjunto de múltiplos possíveis, qual a soma total das correlações obtidas em cada função gerada por cada par de sinais.

**Algoritmo:**

Para cada par de sinais  $x_i$  e  $x_j$ , advindos dos microfones  $i$  e  $j$ , obtenha:

1. A função de correlação cruzada para cada um dos 21 pares possíveis.
2. Para cada função de correlação, correspondente aos diferentes espaçamentos entre sensores desde 1 até 6  $d$ , obtenha o valor da correlação para os valores de  $\tau$ , com  $\tau$  variando desde  $\tau = 0$  até  $\tau = \tau_{m\acute{a}x}$ .
3. Some ponto a ponto os valores de correlação para todas as funções. Busque pelo valor de  $\tau$  que maximiza a soma das correlações obtidas. Este valor de  $\tau$  é a estimativa final para o atraso em amostras.
4. Uma vez obtido o atraso em amostras  $\tau$ , converta-o para o ângulo correspondente.

A seguir, apresentamos a função de correlação final obtida pelo algoritmo supracitado para o caso de uma fonte sonora incidindo no arranjo a 30°:



Figura 4.30: Função de correlação final para um *delay* teórico de  $t = 3$  amostras.

Como pode se observar, o máximo de correlação obtida, considerando todas as 21 funções de correlação cruzada combinadas com a geometria linear, com espaçamentos iguais a  $d$  do arranjo, ocorre para o atraso de  $\tau = 3,2$  amostras, o que corresponde a DOA de  $31,20^\circ$ .

### **Avaliação do limiar de potência (detecção de silêncio) (L)**

Da teoria de estimação da direção de chegada pelo algoritmo SRP-PHAT é sabido que quando não há correlação significativa entre os canais, a DOA estimada será  $0^\circ$ . Daí decorre a necessidade de detectar um limiar de potência para cada quadro de forma a descartar aqueles que não atingirem um mínimo de potência necessário para a estimação DOA.

A vantagem da implantação de um limiar de potência para validação da estimação da direção de chegada está na estabilização das estimativas da direção de chegada. Caso não seja implementado o descarte de quadros com baixa potência, o estimador apontará freqüentemente para  $0^\circ$  e terá um viés permanente para este ângulo, o que não é desejável.

Para medir a potência de um frame, considera-se que:

$$Pot = \sum_{k=1}^7 Var[X(n)] \quad (4.13)$$

Em que  $X(n)$  representa o vetor de valores, no tempo, adquirido pelo canal  $k$  para um determinado quadro. Experimentalmente, determinou-se que:

$$Pot_{sinal} \sim 0.18$$

$$Pot_{referência} \sim 0.0020$$

O termo  $Pot_{referência}$  indica, na verdade, a potência do ruído “de fundo” quando não há sinal incidente sobre os microfones. No ruído “de fundo” está considerado o ruído elétrico de diversas origens presentes no circuito, junto ao ruído sonoro ambiente. Assim, é possível elaborar uma relação entre sinal e ruído de fundo dada por.

$$RSR_f = 10 \log \left( \frac{Pot_{sinal}}{Pot_{referência}} \right) \sim 19 \text{ dB} \quad (4.14)$$

O limiar de potência considerado para o descarte dos quadros é de  $RSR_f < -10 \text{ dB}$  do valor de referência. Ou seja, quadros que possuam  $Pot_{sinal} < Pot_{referência}$  serão descartados pelo estimador e não irão compor a estimativa final para a direção de chegada.

### **Composição da estimativa final DOA com vetor de estimativas válidas (M,N)**

Visto que o processo de estimação da direção de chegada é estocástico, estimativas de DOA incorretas podem surgir devido à interferências não desejadas como: ruído ambiente, fontes interferentes, reverberação, imperfeições nos componentes e canais elétricos não idênticos.

Como forma de minimizar a influência dos efeitos indesejados no desempenho do estimador de direção de chegada, é proposto um método de ponderação de estimativas como forma de conferir robustez ao sistema.

### Janela deslizante (*Sliding window*) com esquecimento exponencial

A janela deslizante, ou *sliding window*, é um método de ponderação de estimativas por algum critério matemático. Para o caso do estimador de direção de chegada, serão consideradas as últimas quatro estimativas DOA válidas.

A ponderação com decaimento exponencial é um mecanismo para prover ao sistema de estimação, como um todo, estabilidade e robustez frente aos possíveis erros que podem ocorrer. O custo computacional levou à decisão de se usar apenas as últimas quatro estimativas. O fator de esquecimento não é crítico e todas as ponderações ensaiadas trouxeram resultados semelhantes. Escolheu-se um fator de esquecimento de aproximadamente 50% entre duas amostras contíguas e válidas, após validação pelo limiar de potência de referência, conforme equação 4.20:

$$\widehat{DOA}_K = 0,5DOA_K + 0,25DOA_{K-1} + 0,125DOA_{K-2} + 0,0625DOA_{K-3} \quad (4.15)$$

Esses pesos foram aproximados pela função apresentada em 4.15.

De maneira esquemática, podemos representar a técnica pela figura a seguir:

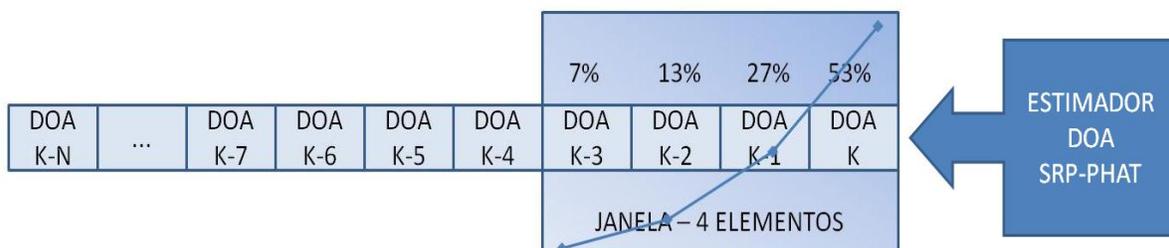


Figura 4.31: Janela deslizante com esquecimento exponencial

O estimador da direção de chegada (DOA) gera continuamente estimativas que estão sujeitas aos mais variados erros. A janela destacada pondera as últimas 4 estimativas que tiveram potência maior que o limiar de silêncio, ponderando a estimativa mais recente com 53%, a penúltima com 27%, a antepenúltima com 13% e a última quarta estimativa com 7%. A função de esquecimento da janela é exponencial, conforme equação abaixo:

$$f(x) = 1,0667e^{-0,693x} \quad (4.16)$$

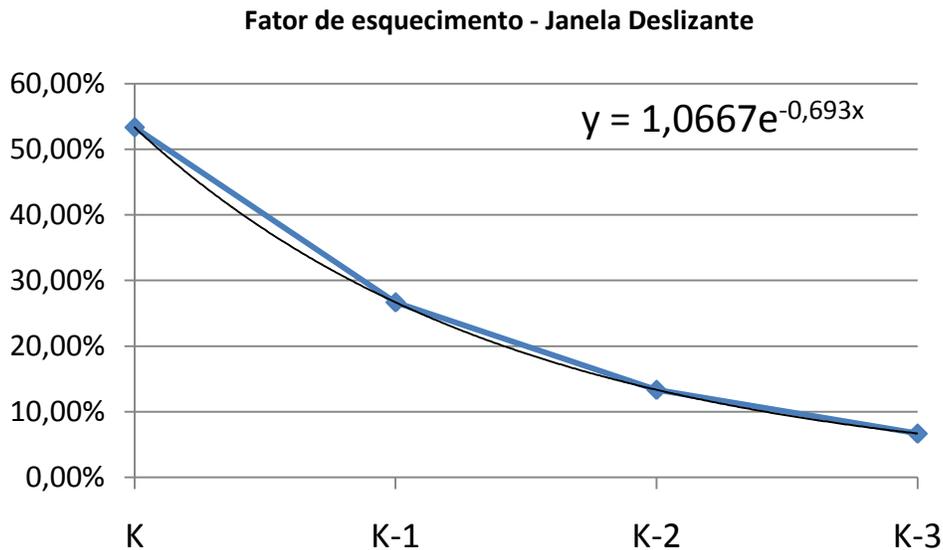


Figura 4.32: Fator de esquecimento – Janela Deslizante

O fator de esquecimento é escolhido de forma que a estimativa subsequente tenha metade do peso da estimativa anterior. Ou seja,  $Peso(K) = 2Peso(K - 1)$ . Isto permite que, mesmo que uma ou mais estimativas qualquer apresente um grande erro, o valor final apresentado pelo estimador mantenha um bom nível de confiança e estabilidade. Embora a janela possa assumir qualquer comprimento em estimativas, optou-se pelo número de quatro estimativas para que o resultado pudesse convergir rapidamente quando da mudança do ângulo da fonte a ser rastreada.

Além disso, uma janela de tamanho maior poderia trazer um “arrasto” das estimativas passadas. A escolha pela função exponencial na ponderação das estimativas se deve à necessidade de dar maior importância às estimativas mais recentes e para garantir que, no rastreamento, o estimador acompanhe o alvo em tempo real, convergindo rapidamente para novos valores de DOA quando o alvo está em movimento.

## 5 RESULTADOS EXPERIMENTAIS

Neste capítulo serão apresentados diversos resultados experimentais obtidos com o sistema de rastreamento da direção de chegada para sinais sonoros por meio de arranjo de microfones. Desejamos avaliar o comportamento do sistema estimador quando da estimação da direção de chegada de sinais sonoros de banda-larga, em ambientes com relação sinal-ruído moderada e reverberação média.

O arranjo de microfones é composto por sete cápsulas de eletreto circulares de 6 *mm* espaçadas de 15 *cm*, presas a uma haste de alumínio de 90 *cm* de comprimento.

A parte traseira da haste é preenchida com material isolante acústico, uma vez que os microfones de eletreto são omnidirecionais e a geometria do arranjo não permite diferenciar se os sons são provenientes da parte frontal ou traseira.

Os cabos que conduzem os sinais são eletricamente blindados e o chassi do arranjo é fixado de forma a minimizar a propagação de vibrações mecânicas indesejadas.



Figura 5.1: Arranjo Uniforme Linear de Microfones

Foi desenvolvida em ambiente MATLAB<sup>®</sup> uma interface gráfica de visualização para acompanhamento da trajetória do alvo. Esta interface contém as últimas 4 estimativas

válidas de acordo com o critério de limiar de potência, a estimativa instantânea final para a DOA e o histórico das últimas 20 estimativas, conforme Figura 5.2.

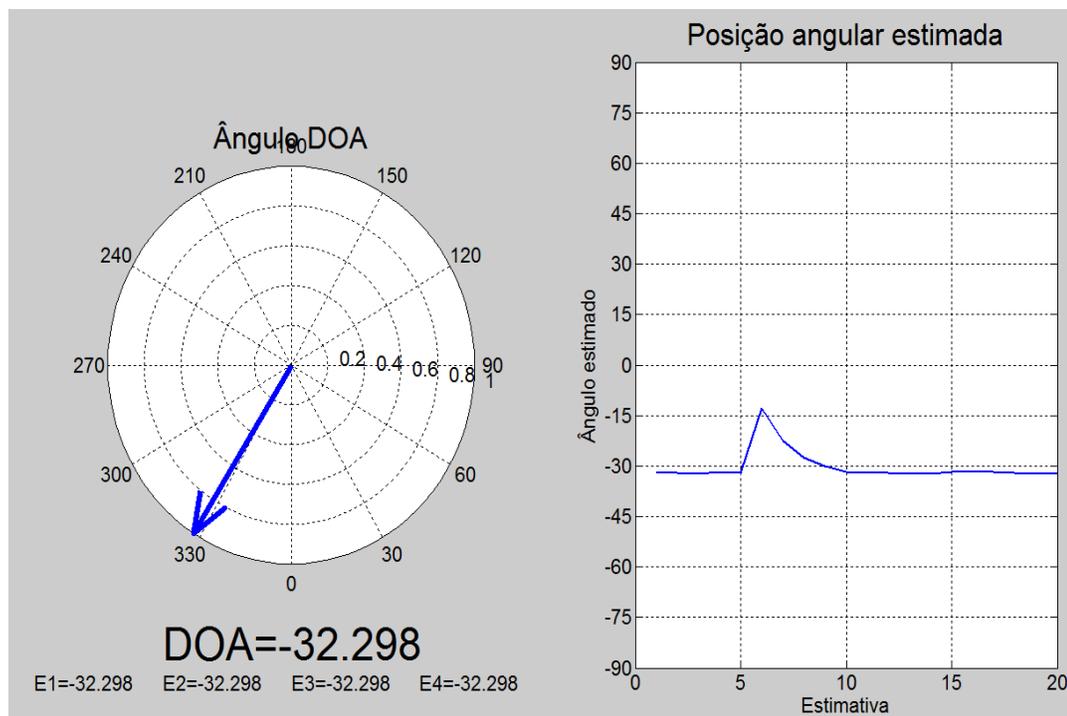


Figura 5.2: *Tracking DOA*

Os sinais são gravados em ambiente real e processados em tempo real pelo ambiente MATLAB<sup>®</sup>. O ambiente de gravação utilizado é uma sala de dimensões 9,0 m x 6,0 m x 2,80 m, sendo as dimensões respectivamente comprimento, largura e altura.

O sinal sonoro utilizado para as gravações são trechos de músicas variadas, para as quais o SNR apresentado pelo circuito de aquisição é de aproximadamente 19 dB. A fonte sonora, para todos os testes realizados, encontra-se nivelada no plano vertical com o arranjo de microfones, distando 3 m deste.

O método utilizado é o SRP-PHAT e são utilizadas as técnicas de limiar de potência e de janela deslizante com esquecimento exponencial (vide Capítulo 4). Conforme mostrado a seguir, para alguns testes, será avaliado o comportamento sem a técnica de

janela móvel com vistas a comparar o desempenho do estimador com e sem o uso desse artifício.

Os dados são coletados a uma taxa de amostragem por canal de  $f_s = 14,285 \text{ kHz}$  e com resolução de 16 bits. Os sete canais são amostrados simultaneamente pelo dispositivo USB 1608-FS (*sample-and-hold*) e os quadros para processamento são compostos por 5000 amostras, o que representa um trecho sonoro de aproximadamente 350 ms para cada *snapshot*. O atraso observado entre a movimentação da fonte sonora e a atualização do resultado para estimação DOA é de cerca de 0,5s.

Os resultados estão divididos em duas categorias: ensaios estáticos e ensaios dinâmicos. Ensaios estáticos estão relacionados à caracterização do comportamento do sistema como um todo quando sobre o arranjo incidem sinais cuja fonte está espacialmente fixa. Já os ensaios dinâmicos buscam caracterizar o comportamento do sistema quando a fonte está se movendo com determinada velocidade angular.

## 5.1 ENSAIOS ESTÁTICOS

Com vistas a caracterizar o comportamento do estimador de direção de chegada sob a hipótese de uma fonte espacialmente fixa, foram realizados ensaios experimentais com uma fonte sonora contínua para os ângulos 0°, 15°, 30° e 45° para os setores à direita – designados por “D”, e à esquerda – designados por “E” – do arranjo linear.

Nas análises a seguir, são apresentados ensaios para rastreamento (DOA *tracking*) e histogramas da estimação DOA para as estimativas obtidas com e sem a utilização do limiar de potência e janela deslizante.

### Sector de 0 a 90° à esquerda do arranjo

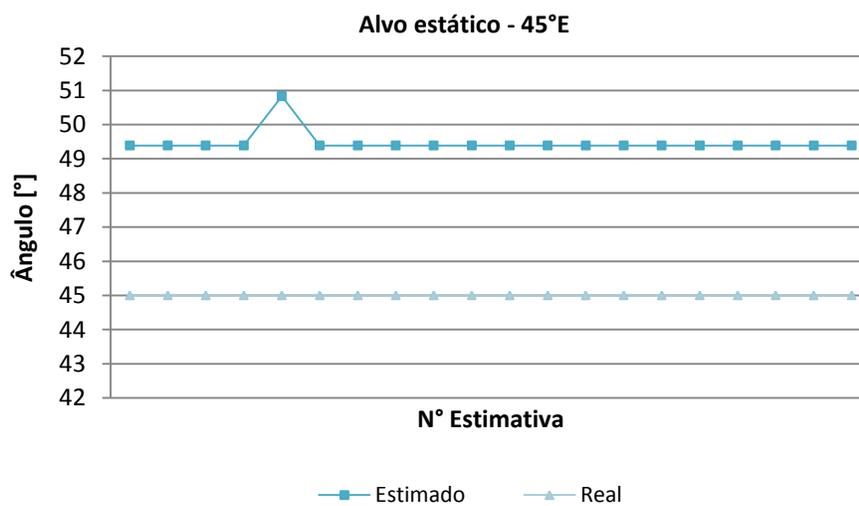


Figura 5.3: DOA estimada para alvo fixo 45° E

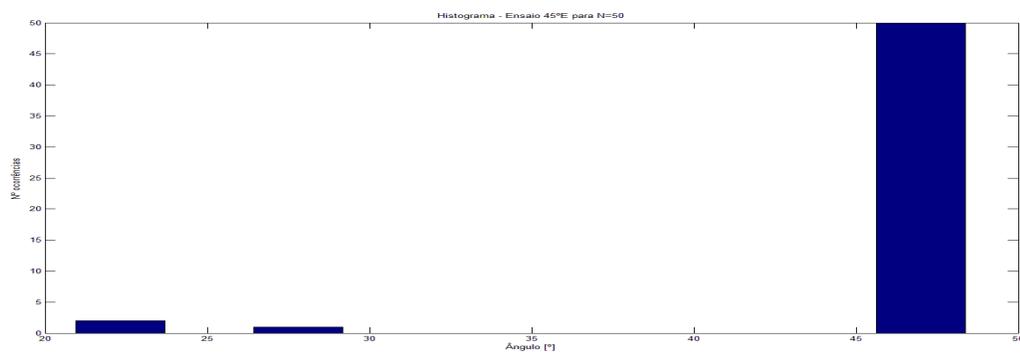


Figura 5.4: Histograma DOA sem limiar de potência e sliding-window – 45° E. Erro Médio: 4,46°

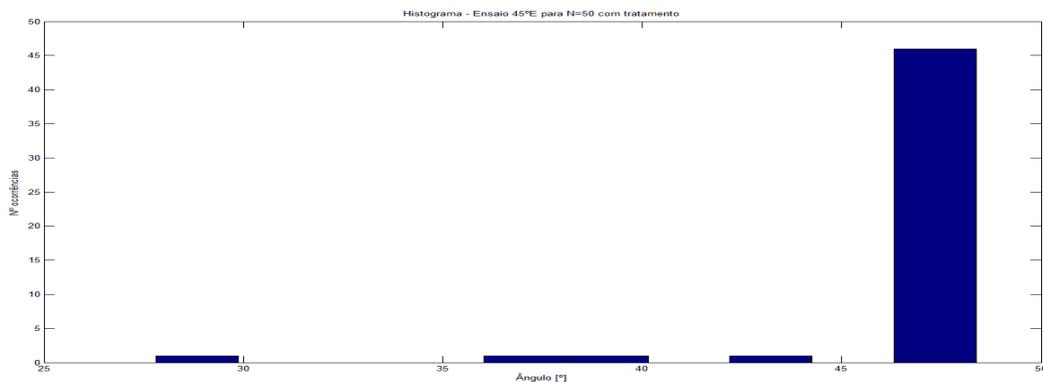


Figura 5.5: Histograma DOA com limiar de potência e sliding-window – 45° E

Os trechos sonoros utilizados são compostos por sinais sonoros musicais variados, e, portanto, não necessariamente os mesmos. Para os ensaios do tipo acompanhamento da fonte sonora são realizadas 20 estimativas, e os histogramas são compostos por dados de 60 realizações.

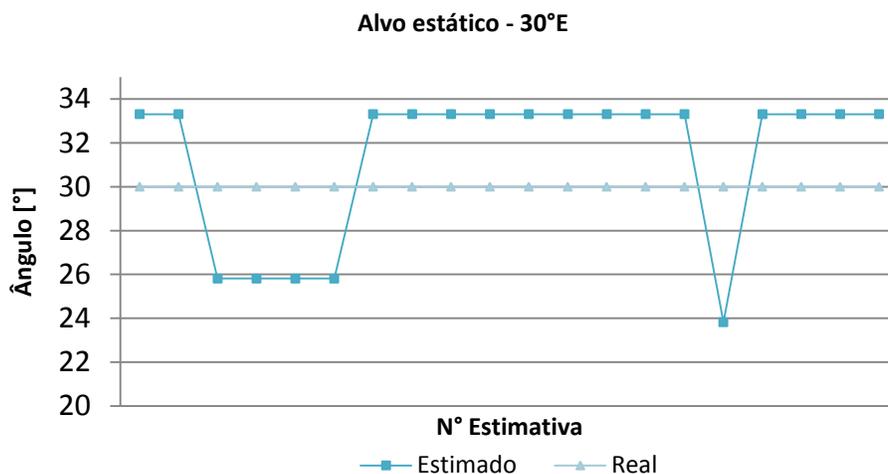


Figura 5.6: DOA estimada para alvo fixo 30° E. Erro médio: 3,63°.

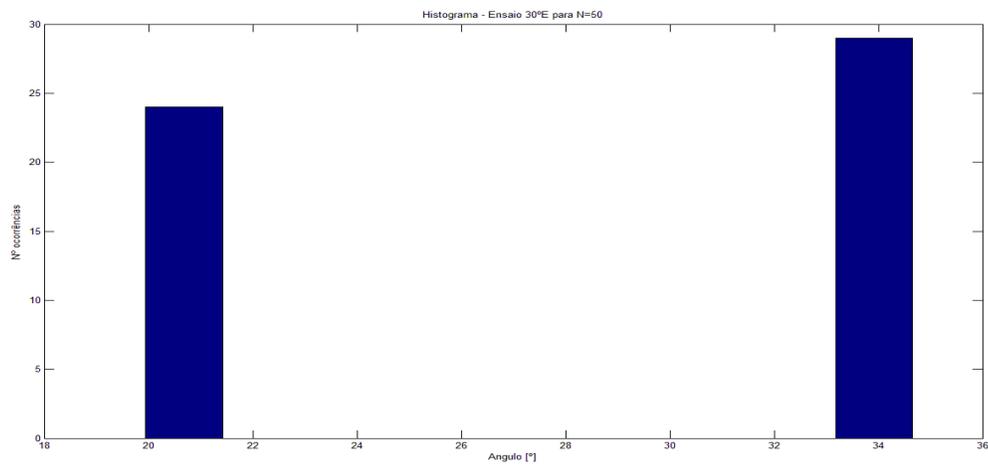


Figura 5.7: Histograma DOA sem limiar de potência e janela deslizante – 30° E

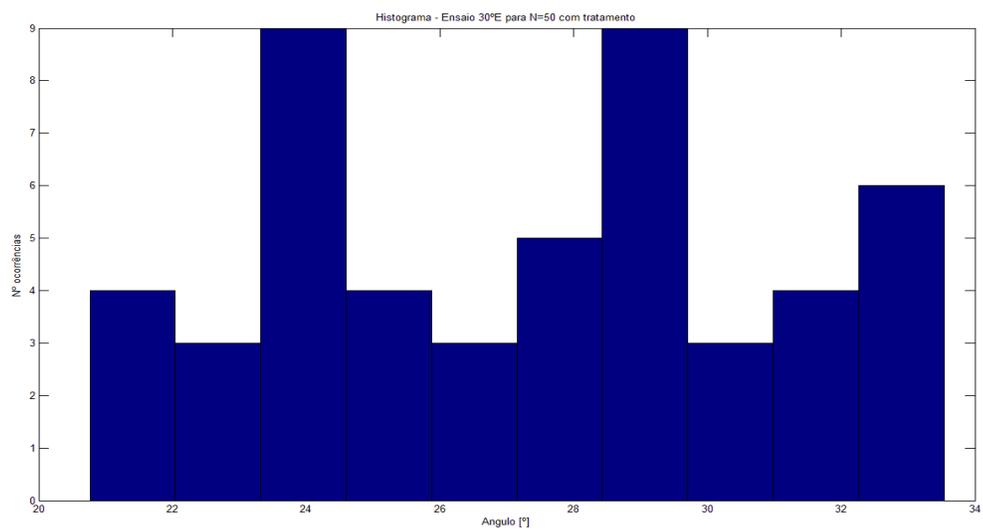


Figura 5.8: Histograma DOA com limiar de potência e janela deslizante – 30° E

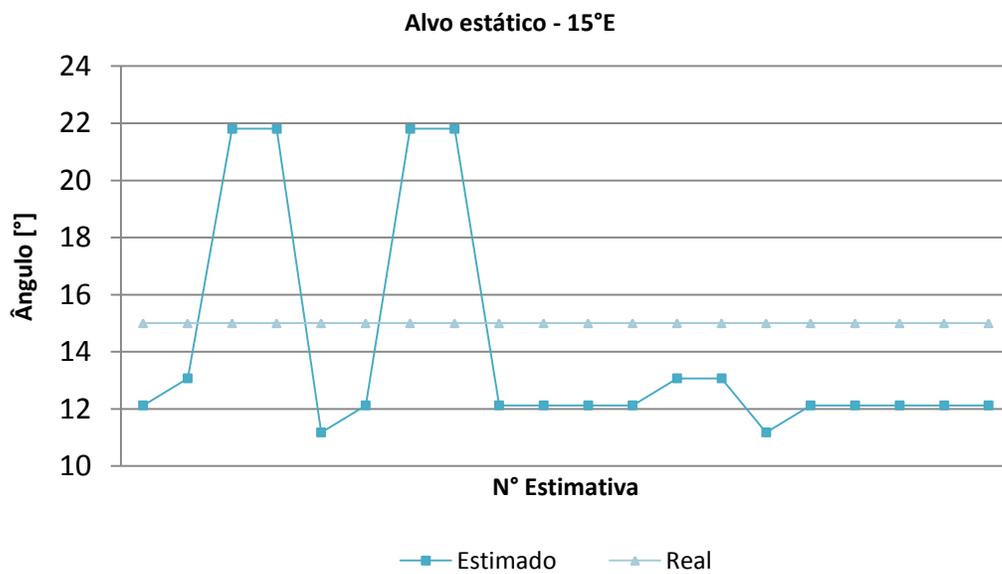


Figura 5.9: DOA estimada para alvo fixo 15°E. Erro médio: 3,62°.

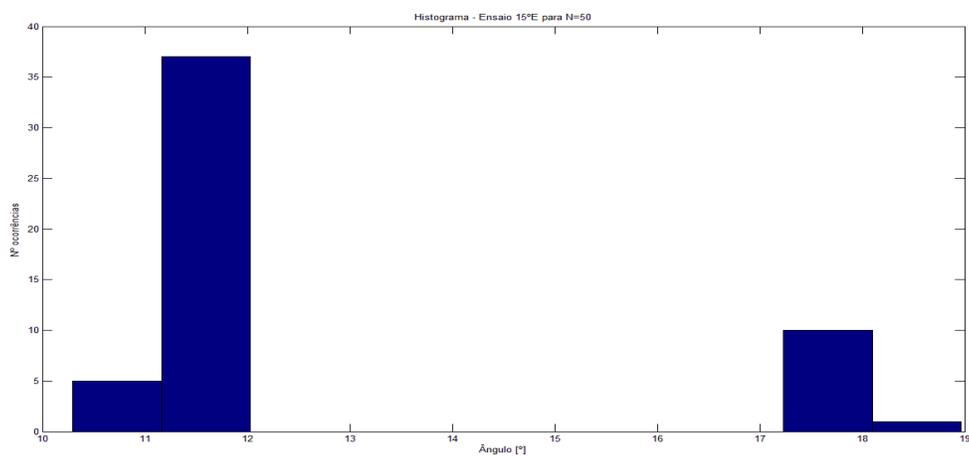


Figura 5.10: Histograma DOA sem limiar de potência e janela deslizante – 15°E

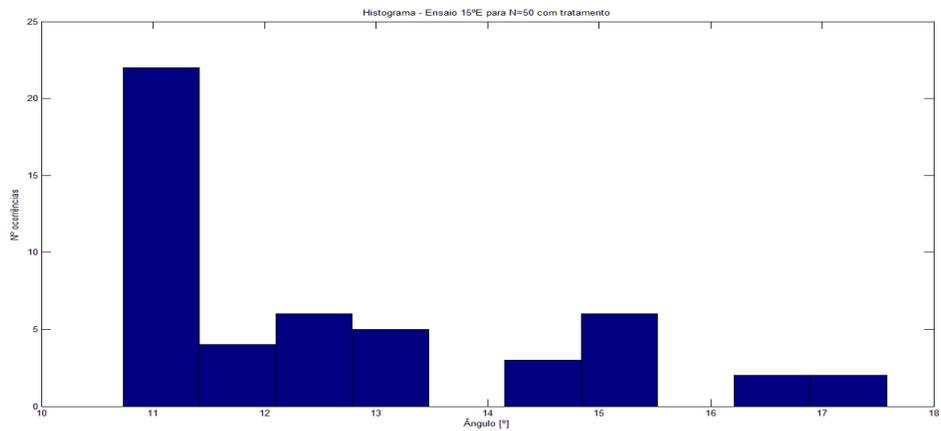


Figura 5.11: Histograma DOA com limiar de potência e janela deslizante – 15°E

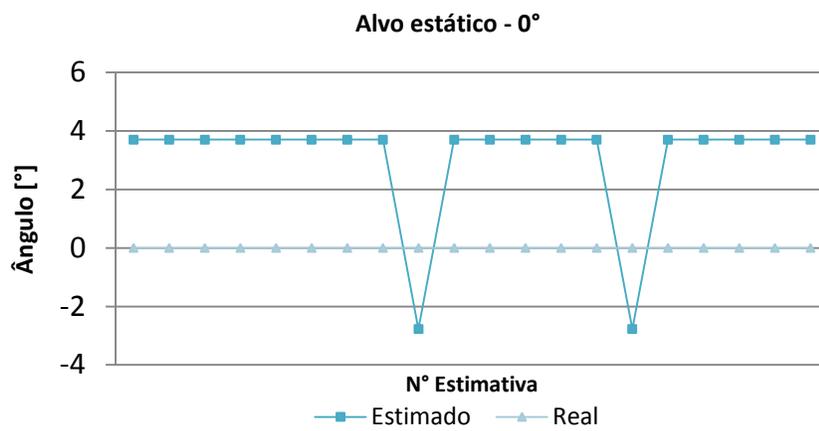


Figura 5.12: DOA estimada para alvo fixo 0°. Erro médio: 3,61°.

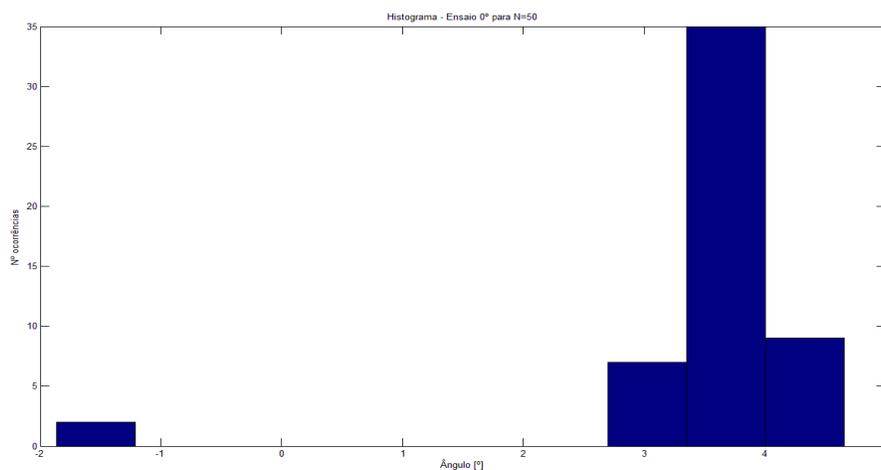


Figura 5.13: Histograma DOA sem limiar de potência e janela deslizante – 0°

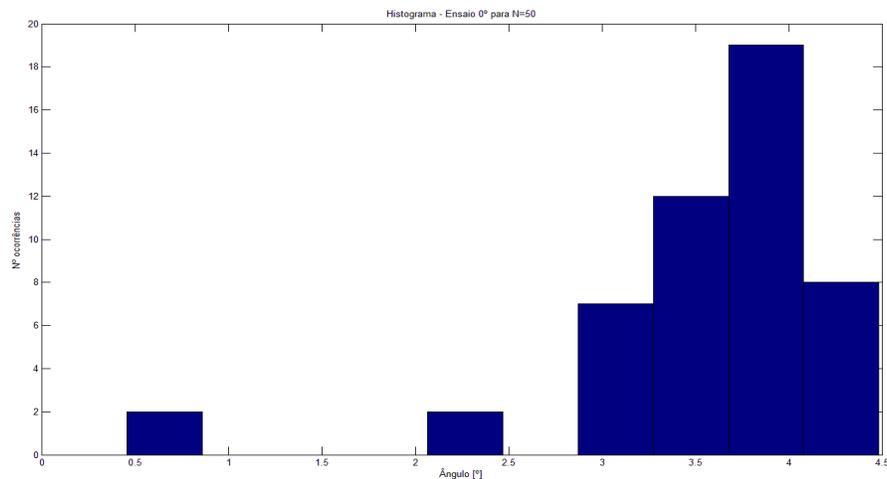


Figura 5.14: Histograma DOA com limiar de potência e janela deslizante – 0°

Este ensaio experimental demonstra o efeito da janela deslizante: pelo uso desta, as estimativas se concentram em torno do valor correto, o que minimiza distorções e dificulta a geração de estimativas com valores grandes de erro absoluto.

### Setor de 0 a 90° à direita do arranjo

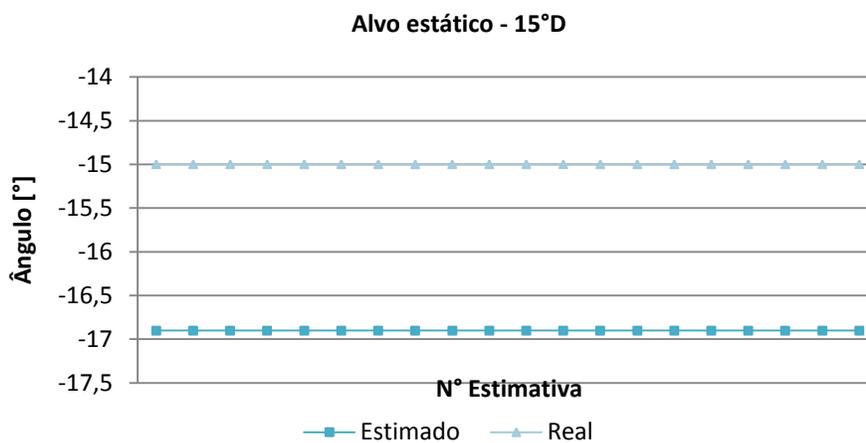


Figura 5.15: DOA estimada para alvo fixo 15°D. Erro médio: 1,90°.

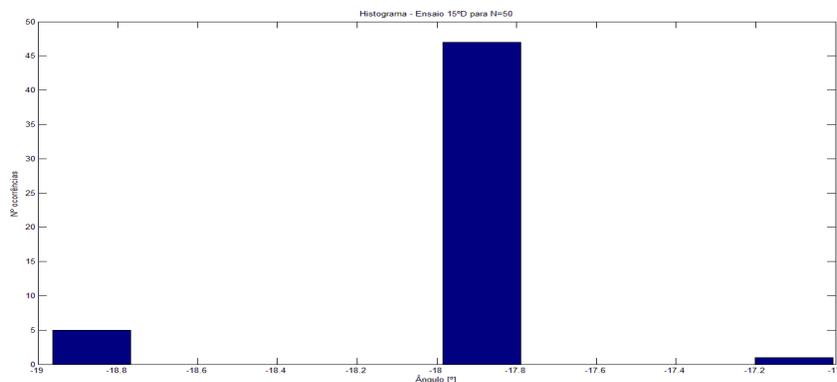


Figura 5.16: Histograma DOA sem limiar de potência e janela deslizante – 15°D

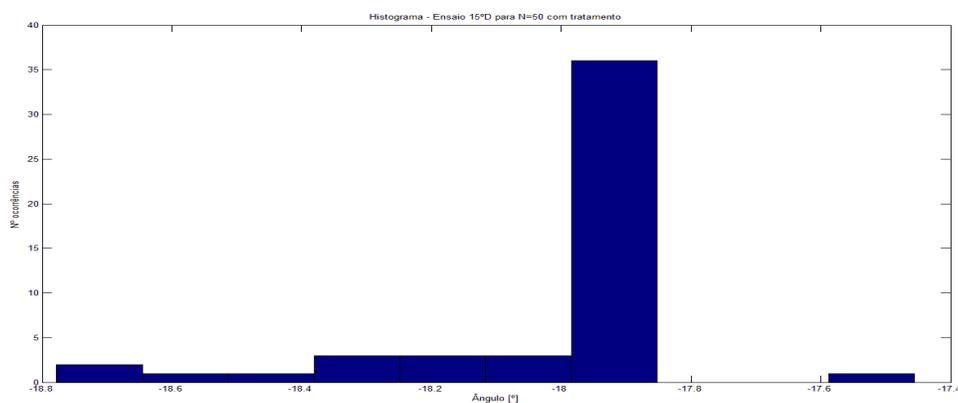


Figura 5.17: Histograma DOA com limiar de potência e janela deslizante – 15°D

### Alvo estático - 30°D

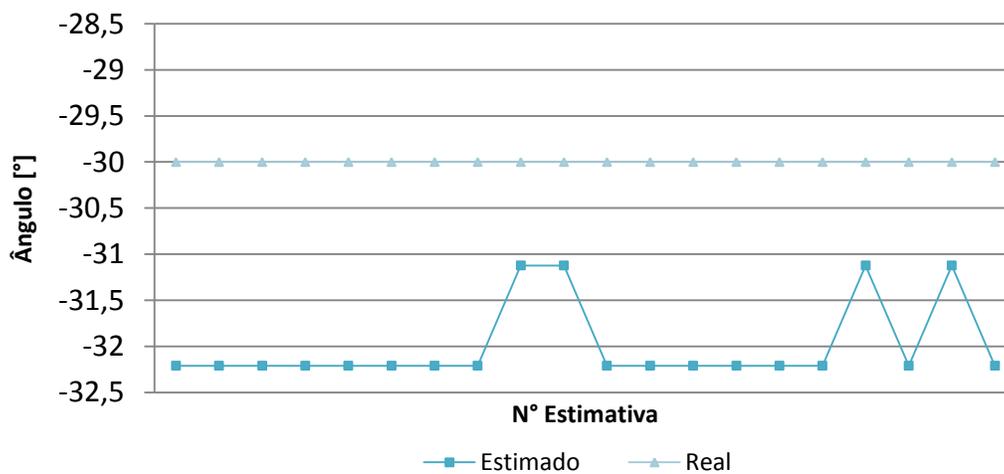


Figura 5.18: DOA estimada para alvo fixo 30°D. Erro médio: 1,99°.

Neste ensaio observa-se um viés de aproximadamente  $2^\circ$  na estimativa, que pode ser por parte ocasionado pela imprecisão do posicionamento da fonte sonora que foi realizado com medições com erro em torno de  $\pm 5 \text{ cm}$ .

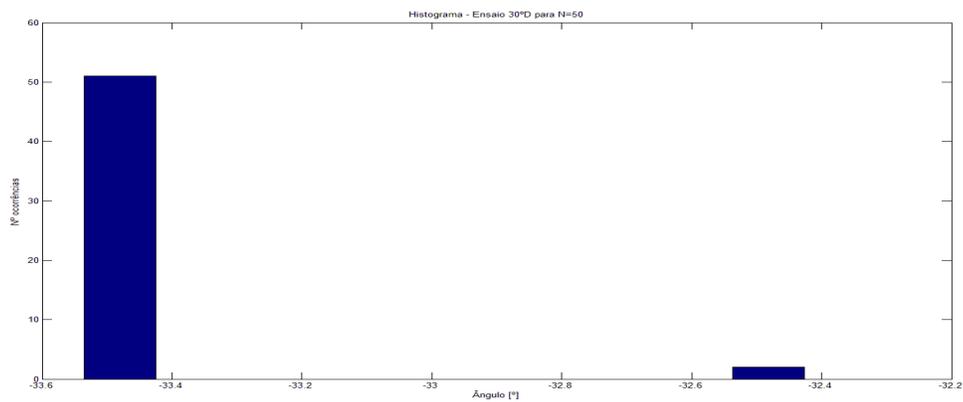


Figura 5.19: Histograma DOA sem limiar de potência e janela deslizante –  $30^\circ\text{D}$

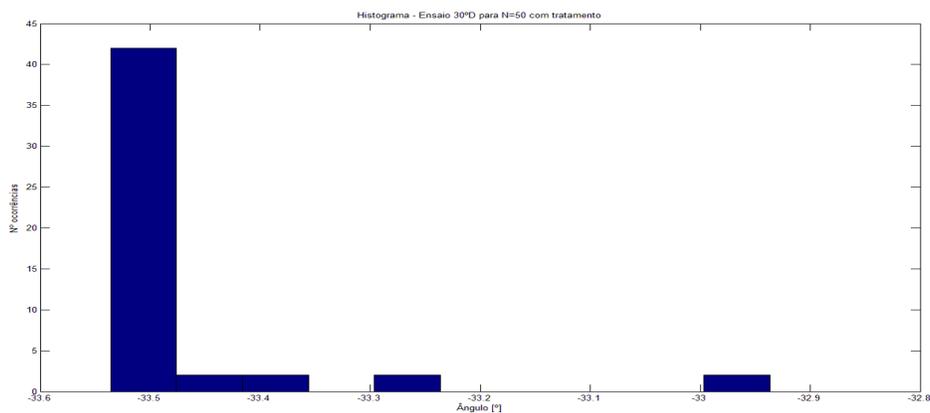


Figura 5.20: Histograma DOA com limiar de potência e janela deslizante –  $30^\circ\text{D}$

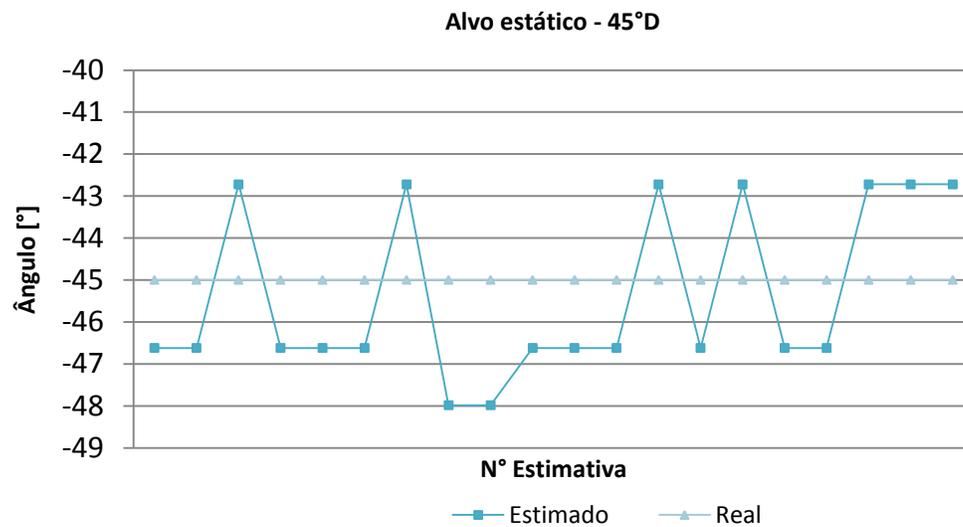


Figura 5.21: DOA estimada para alvo fixo 45°D. Erro médio: 1,99°.

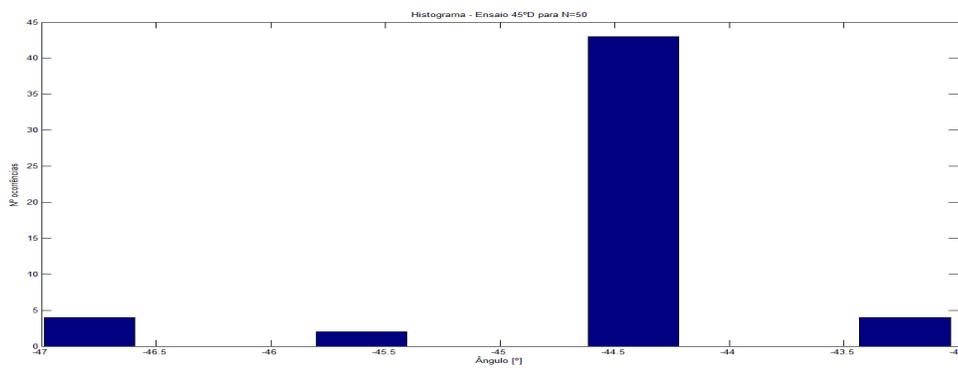


Figura 5.22: Histograma DOA sem limiar de potência e janela deslizante – 45°D

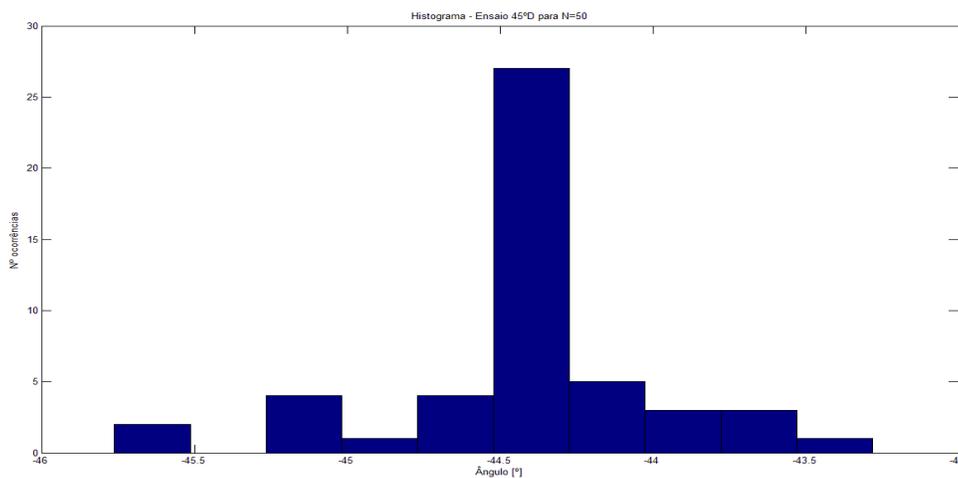


Figura 5.23: Histograma DOA com limiar de potência e janela deslizante – 45°D

### Resumo dos ensaios estáticos – *Tracking*

Ensaio	Erro médio	Desvio-Padrão
45° E	4,46°	0,32°
30° E	3,62°	0,70°
15° E	3,61°	1,71°
0°	3,61°	0,29°
15° D	1,90°	0,13°
30° D	1,99°	0,45°
45° D	1,98°	0,46°

Tabela 1: Resumos dos ensaios estatísticos- *Tracking*

Pela análise dos gráficos acima, observa-se que a precisão do estimador para alvo fixo é de cerca de 3°. As medidas apresentam boa convergência com desvio-padrão em torno de 0,5°, apenas.

Da análise dos histogramas pode-se notar o efeito positivo do uso do limiar de potência e da janela de estimativas: as medidas apresentam menor dispersão com o uso desses dispositivos, o que causa uma maior concentração das medidas em torno do valor correto. Isto confere maior robustez e confiabilidade às estimativas geradas pelo sistema de rastreamento DOA.

## 5.2 ENSAIOS DINÂMICOS

De forma a caracterizar o comportamento dinâmico do sistema de rastreamento da direção de chegada, foram realizados experimentos variando-se o tipo de movimentação que é realizado pela fonte sonora.

### Movimento angular uniforme de 0 a 45° para direita do arranjo

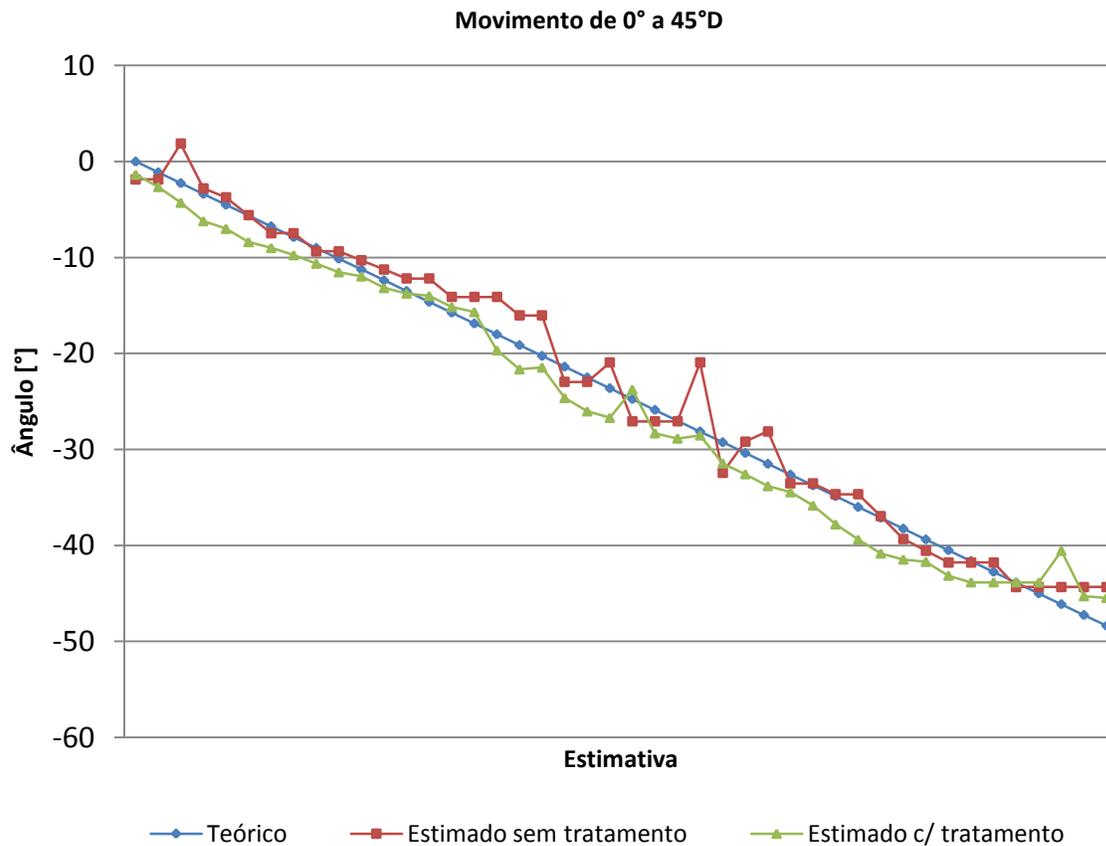


Figura 5.24: Movimento angular uniforme de 0 a 45° para direita do arranjo

Erro médio sem tratamento: 1,65°.

Erro médio com tratamento: 2,02°.

Desvio-Padrão do erro sem tratamento: 1,48°.

Desvio-Padrão do erro com tratamento: 1,08°.

### Movimento angular uniforme de 0 a 45° para esquerda do arranjo

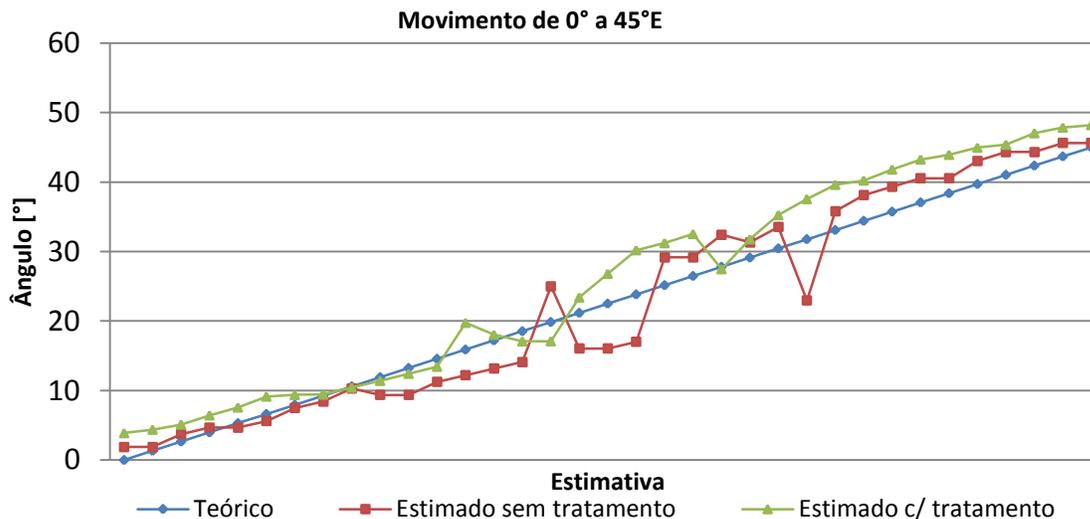


Figura 5.25: Movimento angular uniforme de 0 a 45° para esquerda do arranjo

Erro médio sem tratamento: 3,01°.

Erro médio com tratamento: 3,43°.

Desvio-Padrão do erro sem tratamento: 1,98°.

Desvio-Padrão do erro com tratamento: 2,07°.

### Movimento angular uniforme de -45 a 45°

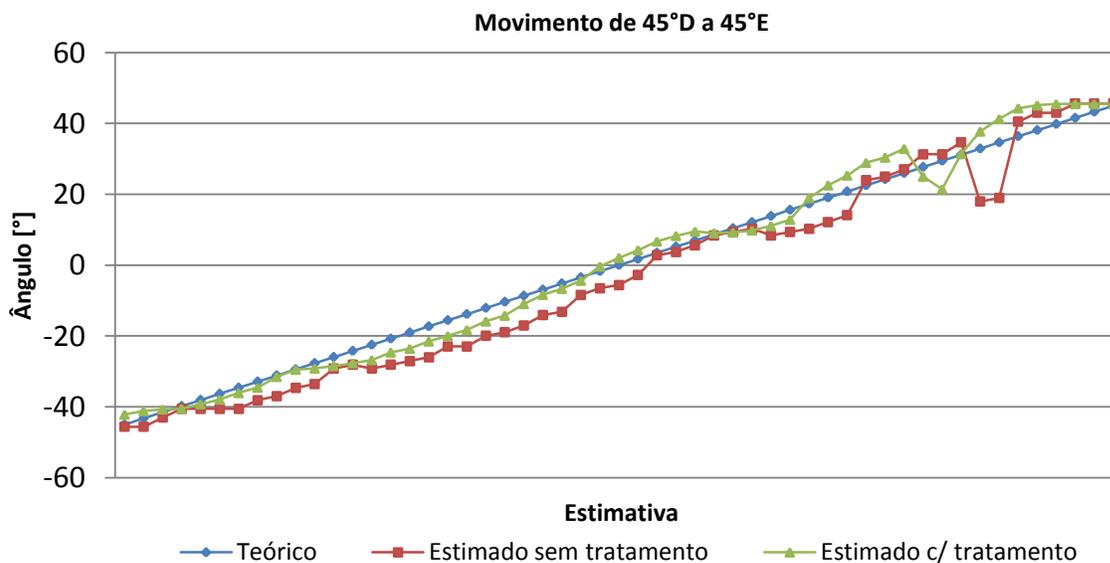


Figura 5.26: Movimento angular uniforme de -45° a 45°

Erro médio sem tratamento: 4,82°.

Erro médio com tratamento: 3,07°.

Desvio-Padrão do erro sem tratamento: 3,33°.

Desvio-Padrão do erro com tratamento: 2,06°.

### Movimento angular uniforme de -45°D a 30°D a 30°E

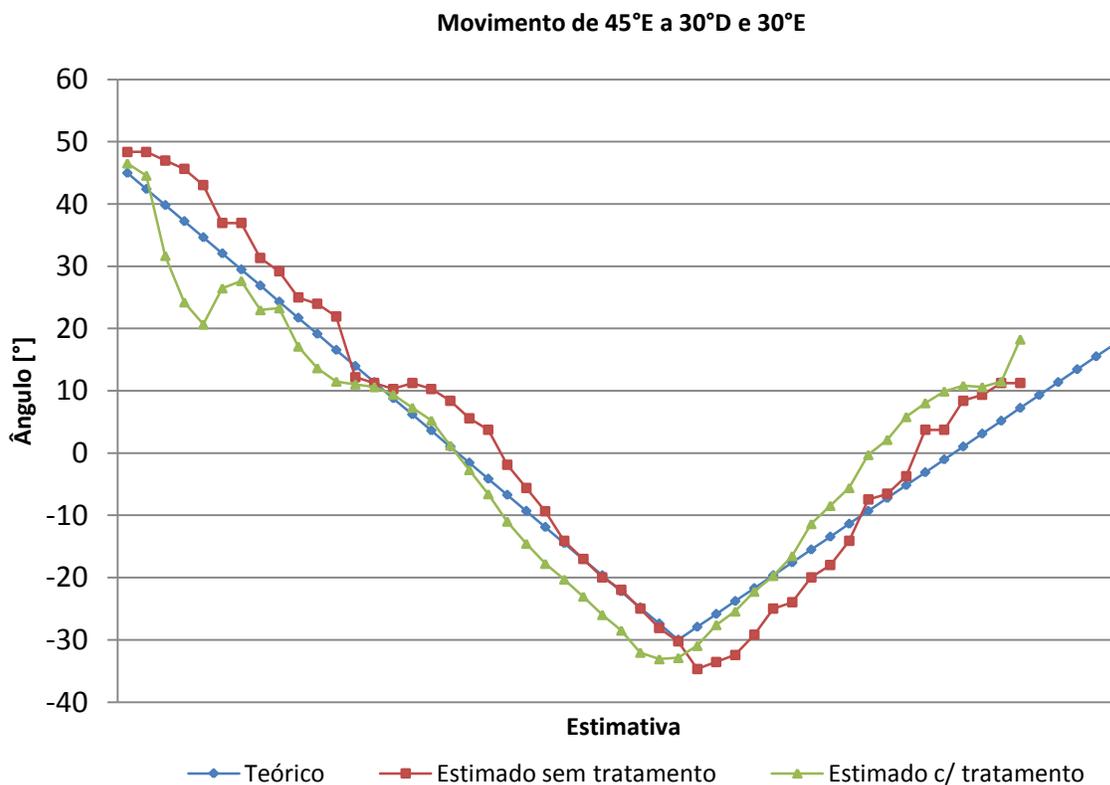


Figura 5.27: Movimento angular uniforme de -45°D a 30°D a 30°E

Erro médio sem tratamento: 4,98°.

Erro médio com tratamento: 6,27°.

Desvio-Padrão do erro sem tratamento: 3,10°.

Desvio-Padrão do erro com tratamento: 4,44°.

### Resumo dos ensaios dinâmicos

Ensaio	Erro médio com tratamento	Erro médio sem tratamento	DP com tratamento	DP sem tratamento
0° a 45° D	2,02°	1,65°	1,08°	1,48°
0° a 45° E	3,43°	3,01°	2,07°	1,98°
45° D a 45° E	3,07°	4,82°	2,06°	3,33°
45° E a 30° D e 30° E	6,27°	4,98°	4,44°	3,10°
<b>Média</b>	<b>3,70°</b>	<b>3,62°</b>	<b>2,41°</b>	<b>2,47°</b>

Tabela 2: Resumo dos ensaios dinâmicos

Ensaio	Porcentagem erros < 3° sem tratamento	Porcentagem erros < 5° sem tratamento	Porcentagem erros < 3° com tratamento	Porcentagem erros < 5° com tratamento
0° a 45° D	80%	96%	82%	96%
0° a 45° E	47%	83%	44%	69%
45° D a 45° E	31%	54%	56%	83%
45° E a 30° D e 30° E	30%	50%	30%	42%
<b>Média</b>	<b>47%</b>	<b>71%</b>	<b>53%</b>	<b>73%</b>

Tabela 3: Resumo dos ensaios dinâmicos- Porcentagem de erros

Os diversos experimentos realizados com o alvo em movimento visam a caracterizar o comportamento do sistema de estimação proposto para um emissor de ondas sonoras em movimento. A velocidade angular utilizada para a movimentação da fonte sonora é aproximadamente constante de 5 °/s e a velocidade linear de 0,25 m/s. Esta velocidade se mostra compatível com a movimentação de interlocutores em aplicações de rastreamento por voz em videoconferências ou próteses auditivas, por exemplo.

Considerando todos os ensaios em conjunto, pode-se notar que, para uma precisão de 5° na estimativa, obtemos uma taxa de confiabilidade de aproximadamente 70%. Este é um resultado satisfatório dado as limitações da taxa de amostragem e poder computacional do hardware em questão.

## 6 CONCLUSÃO E TRABALHO FUTURO

Várias são as aplicações possíveis para a estimação da direção de chegada por arranjo de sensores, dentre as quais se destacam as próteses auditivas inteligentes, sonares, radares, artefatos militares, equipamentos médicos e antenas diretivas. Em várias dessas aplicações, a estimação da direção de chegada é insumo essencial.

Neste estudo foram avaliados diferentes algoritmos para a estimação da direção de chegada. Para a estimação, por arranjo de microfones, da direção de chegada de sinais sonoros de banda-larga, em ambientes com relação sinal-ruído moderada e reverberação média, conclui-se que a técnica SRP-PHAT tem o melhor desempenho dentre os algoritmos analisados.

Além disso, é proposta uma forma de cálculo que leva em consideração todas as combinações possíveis entre pares de microfones em conjunto com a geometria do arranjo de sensores para formar uma estimativa final do ângulo de chegada. Isto é, calculamos a correlação para todas as combinações de microfones possíveis aquela que apresenta a maior correlação.

Uma nova técnica de ponderação das estimativas por janela deslizante móvel com fator de esquecimento ajustável é apresentada, de modo a prover robustez e confiabilidade ao resultado.

Propomos um limiar de variância em que cada quadro do sinal adquirido é avaliado em relação ao limiar. Esse limiar proporciona maior estabilidade nas estimativas uma vez que pequenos trechos de silêncio, inerentes ao sinal de voz, não causam viés (*bias*) do estimador para o ângulo zero.

Apresentamos um projeto de *hardware*, para até oito canais, em circuito impresso, para a aquisição, condicionamento, amplificação e filtragem de sinais sonoros obtidos a partir de sensores de eletreto. Os circuitos apresentados atendem às restrições de sinal-ruído, distorção e fase que os estimadores analisados requerem.

A implementação do sistema rastreador DOA é capaz de detectar uma fonte móvel no espaço com erro médio de  $3,7^\circ$  e desvio-padrão de  $2,4^\circ$ , sendo que 70% dos erros são menores que  $5^\circ$ . Para o caso estático, o sistema detecta o alvo com erro médio de  $3^\circ$  e desvio-padrão de  $0,6^\circ$ .

Por limitações no hardware de conversão analógico-digital, a taxa de amostragem utilizada foi  $14,2\text{ kHz}$ . Identificamos que este é um fator de melhoria para desenvolvimento futuro, vez que maiores taxas de amostragem possibilitam maior precisão no cálculo da função de correlação e menor fator de interpolação. Para uma precisão de  $1^\circ$  na estimativa e um espaçamento entre sensores de  $15\text{ cm}$ , deveríamos ter uma taxa de amostragem elevada para, no mínimo,  $133\text{ kHz}$ . Para um espaçamento de sensores de  $1\text{ cm}$ , a taxa de amostragem ideal seria de  $2\text{ MHz}$ .

Propomos para desenvolvimento futuro a investigação de outros métodos de ponderação das estimativas como os filtros de *Kalman*, de partículas, entre outros métodos de filtragem estocástica, e sua comparação com a janela deslizante apresentada. Outro aspecto a ser observado é o grande *overhead* computacional introduzido pelo ambiente MATLAB<sup>®</sup>. Em desenvolvimento futuro, sugerimos portar o sistema para processamento por DSP's.

Além disso, seria adequado incorporar ao cálculo da estimação DOA características espectrais de voz. Isto é, poder-se-ia, para aplicação de rastreamento de voz humana por estimação DOA, incorporar técnicas de detecção de voz nos quadros como VAD (*Voice*

*Activity Detection*) [33]. Esta técnica permitiria o descarte de quadros que não tivessem o conteúdo predominantemente formado por sinais de voz, o que provavelmente traria uma melhora sensivelmente no desempenho do estimador.

Embora o SRP-PHAT se mostre um algoritmo viável e adequado para sinais de banda larga, como são o caso dos sinais sonoros, outros algoritmos de banda estreita, como MUSIC, CAPON, TOPS e WAVES, poderiam ser utilizados em desenvolvimento futuro. Para tanto, seria necessário um banco de filtros digitais de banda estreita e um critério de junção das estimativas de cada faixa. Por operarem eficientemente em relações sinal-ruído menores do que o SRP-PHAT e melhor resolução sugerimos que, futuramente, esses algoritmos também sejam analisados.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ALMEIDA, K.; IÓRIO, M. *“Próteses Auditivas - Fundamentos teóricos e Aplicações clínicas”*, [S.l.]: Ed. Lovise, São Paulo, 1996.
- [2] AZIMI-SADJADI, M.R.; PEZESHKI, A.; ROSEVEARE, N. *“Wideband DOA estimation algorithms for multiple moving sources using unattended acoustic sensors”*, Aerospace and Electronic Systems, IEEE Transactions, v. 44, n.4, p.1585-1599. Out. 2008.
- [3] B.D., Van Veen; BUCKLEY, K.M., “Beamforming: A versatile approach to spatial filtering”, IEEE ASSP Mag., v. 5, n. 2, p.4-24, Abr. 1988.
- [4] BIN, C.Y.; JIE, W.;FA, L.Y. “Broadband Adaptive Beamforming Using UIA-MVDR Algorithm”, Wireless Communications, Networking and Mobile Computing, 2007.
- [5] BRACEWELL, R. *“Pentagram Notation for Cross Correlation”*, The Fourier Transform and Its Applications. New York: McGraw-Hill, p. 46 e 243, 1965.
- [6] BRANDSTEIN, M. S., WARD, D. B. *“Microphone Arrays: Signal Processing Techniques and Applications”*, Nova Iorque: Springer Verlag, 2001.
- [7] BRANDSTEIN, M., SILVERMAN. H, *“A robust method for speech signal time-delay estimation in reverberant romos”*, in Proc. IEEE Int. Conf. Acoust. Speech, Signal Processin. Munique, Alemanha, p. 375-378. Abr. 1997.
- [8] CAO, J.; LIU, Z.; XU,Y. *“New algorithm requiring no preprocessing for wideband DOA estimation”*, Signal Processing - 9th International Conference, p. 394-397. Out. 2008.
- [9] CAPON J. *“High-resolution frequency-wavenumber spectrum analysis”*, v.57, p.1408-1418, 1969.
- [10] CARTER, G.C.; NUTTALL, A. H.; CABLE, P. G. *“The smoothed coherence transform (SCOT)”*, Tech. Rep., Naval Underwater Systems Center, New London Lab. 1972.
- [11] DA COSTA, J. P. C. L., *“Parameter Estimation Techniques for Multi-dimensional Array Signal Processing”*, Ed. Shaker, Aachen, Alemanha, Mar. 2010.
- [12] DA COSTA, J. P. C. L., ZELENOVSKY, R. PUTTINI, *“Simulação da Prótese Auditiva Inteligente”*. In: Congresso Ibero-Americano sobre Tecnologias de Apoio a Portadores de Deficiência (IBERDISCAP), 2006, Vitória. Anais do Congresso Ibero-Americano sobre Tecnologias de Apoio a Portadores de Deficiência (IBERDISCAP), 2006.
- [13] DA COSTA, J. P. C. L.; ZELENOVSKY, R. PUTTINI, *“Estudo de Viabilidade da Prótese Auditiva Inteligente”*. In: Congresso Ibero-Americano sobre Tecnologias de Apoio

a Portadores de Deficiência (IBERDISCAP), 2006, Vitória. Anais do Congresso Ibero-Americano sobre Tecnologias de Apoio a Portadores de Deficiência (IBERDISCAP), 2006.

[14] DE BOOR, C. "*A Practical Guide to Splines*", New York: Springer Verlag, 318p, 1978.

[15] DI CLAUDIO, E.; PARISI, R. "*WAVES: Weighted average of signal subspaces for robust wideband direction finding*", IEEE Trans. On Signal processing, v. 49, n. 10, p. 2179-2190, Out. 2001.

[16] DIBIASE, J. H.; SILVERMAN, H. F.; BRANDSTEIN, M. "*Robust Localization in Reverberant Room*", Microphone Arrays, Springer-Verlag, 2001.

[17] DO, H; SILVERMAN, H.F. "*Stochastic Particle Filtering: A fast SRP-PHAT single source localization algorithm*", in Proceedings of the IEEE Workshop on Applications of Acoustics, Speech, and Signal Processing, p.213-216. Nova Iorque, 2009.

[18] DO, HONG, T.; RUSSEK, P. "*An analysis of wideband direction-of-arrival estimation for closely-spaced sources in the presence of array model errors*", Microwave and Wireless Components Letters, IEEE , v. 13, n.8, p. 314- 316, Ago. 2003.

[19] GAO, Y; BRENNAN, M.J.; JOSEPH, P.F. "*A comparison of time delay estimators for the detection of leak noise signals in plastic water distribution pipes*", Journal of Sound and Vibration, v. 292, p.552-570. Mai. 2006.

[20] GODARA, L. C. "*Application of antenna arrays to mobile communications, part II: Beamforming and direction-of-arrival considerations*", Proc. of IEEE, v. 85, n. 8, p. 1195-1245. Ago. 1997.

[21] GODARA, L. C. "*Application of antenna arrays to mobile communications, part I: Performance Improvement, Feasibility, and System Considerations*", Proc. of IEEE, v. 85, n. 7, p. 1031-1060, Jul. 1997.

[22] GONTIJO, A. T. e COSTA, M. V. S. "*Desenvolvimento do Hardware e Estudo dos Métodos Utilizados na Estimção de DOA por Meio de Arranjos de Sensores: Prótese Auditiva Inteligente*", Monografia de Graduação, Publicação ENE 01/2007, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 89p, Jan. 2007.

[23] HASSANPOUR, M.; KAHAEI, M.H. "*Adaptive DOA tracking using microphone arrays*", 9th International Symposium, Signal Processing and Its Applications, p.1-4, Fev. 2007.

[24] IVRLAC, M.T.; NOSSEK, J.A.; "*High-efficiency super-gain antenna arrays*", Smart Antennas (WSA), 2010 International ITG Workshop on, p.369-374, Fev. 2010.

- [25] JAHANSON, A. “*Implementations strategies for array localization algorithms based on the SRP-PHAT*”, BlekingeTekniska Högskolan, Ronneby, Suíça, 2001.
- [26] JAHANSON, A., GRBIC, N., NORDHOLM, S., “*Speaker Localisation using the far-field SRP-PHAT in conference telephony*”, IEEE Internacional Symposium On Intelligent Signal Processing and Communication Systems, 2002.
- [27] JIAN, M.; KOT, A. C.; ER, M. H. “*Doa estimation of speech source with microphone arrays*”, In: IEEE, p. v293–v296, 1998.
- [28] KAILATH, “*Direction of arrival estimation of subspace rotation methods-ESPRIT*”, In. Proc. ICASSP. Tóquio, 1986.
- [29] KNAPP, C.H.; CARTER, C. “*The generalized correlation method for estimation of time delay*”, IEEE Transactions on ASSP, 24, p. 320-327, Ago. 1976.
- [30] MENDONZA, F.M.; COVARRUBIAS, D.H.R.; LÓPEZ, C.A.M. “*Localización espacial de fuentes en comunicaciones móviles empleando actualización sucesiva de subespacio*”, IEEE Latin America Transactions, v.6, Jun. 2008.
- [31] MUNHOZ, R. A.; TORRES, A. M. V. L. “*Vez da Voz – Audição e surdez: O que é a perda auditiva e como tratá-la.*”, [S.l.]: Serviço de Indicação e Adaptação de Próteses Auditivas do Núcleo de Audiologia e Próteses Auditivas Campinas Ltda – Unitron, 2003.
- [32] ONUMA, Y.; ICHIGE, K.; ARAI, H. “*DOA estimation for wideband signals using modified C-SPRIT method*”, TENCON 2007 - IEEE Region 10 Conference, p.1-4, Nov. 2007.
- [33] RAMÍREZ, J.; J. M. GÓRRIZ, J. C. Segura. “*Voice Activity Detection. Fundamentals and Speech Recognition System Robustness*”, In M. Grimm and K. Kroschel. Robust Speech Recognition and Understanding, p. 1–22, 2007.
- [34] SCHMIDT, R.O. “*Multiple Emitter Location and Signal Parameter Estimation*”, IEEE Trans. Antennas Propagation, v. AP-34, p.276-280, Mar. 1986.
- [35] SONG, Xin; WANG, Jinkuan; HAN, Yinghua; MENG, Yan. “*Robust Adaptive Beamforming Using a Bayesian Approach*”, TENCON 2006. IEEE Region 10 Conference, p.1-4, Nov. 2006.
- [36] SUCHITA W.V.; KULAT, K. D. “*Robust Algorithms for DOA Estimation and Adaptive Beamforming for Smart Antenna Application*”, Second International Conference on Emerging Trends in Engineering and Technology, ICETET-09.

- [37] VARMA, K.; IKUMA, T.; BEEEX, A. A. "Robust TDE-based DOA estimation for compact audio arrays", IEEE Sensor Array and Multichannel Signal Proc. Workshop (SAM), Ago, 2002.
- [38] VIBERG, M.; LUNDGREN, A. "Array interpolation based on Local Polynomial Approximation with application to DOA estimation using weighted MUSIC", Acoustics, Speech and Signal Processing in IEEE International Conference, p.2145-2148, Abr. 2009.
- [39] WEI, C.;RUI-JUAN,Y.;MEI-RONG, H.;ZHAO-YONF, F. "An Improved Fast Approach for DOA Tracking Estimation", WRI World Congress on Computer Science and Information Engineering, v. 1, p.63-67, 2009.
- [40] YAMADA, I. "High-resolution DOA estimation by algebraic phase unwrapping algorithm" Circuits and Systems, IEEE International Symposium, v.3, p. 2413- 2416, Mai.2005.
- [41] YAN, H.; FAN, H.H. "A Wideband Kalman DOA Tracking Algorithm", Sensor Array and Multichannel Processing. Fourth IEEE Workshop, p.45-49, Jul. 2006.
- [42] YONN, Y.S.; KAPLAN, L.;MCCLELLAN, J. "TOPS: New DOA estimator for wideband signals", IEEE Transactions on Signal processing, v. 54, n.6, Jun.2006.
- [43] YU, H. *et al.*, "A New Method for Wideband DOA Estimation", Int. Conf. on Wireless Communications, Networking, and Mobile Comuputing, p. 598-601, Set. 2007.
- [44] ZELENOVSKY, R. "Emprego de arranjo de antenas na recuperação de dados digitais em ambiente CDMA", Tese (Doutorado em Engenharia Elétrica). Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2001.
- [45] ZHANG, C.; FLORENCIO, D.; ZHANG, Z. "Why does PHAT work well in lownoise, reverberative environments?", Acoustics, Speech and Signal Processing, p.2565 – 2568, 2008.
- [46] ZHOU, Lin; ZHAO, Yong-jun; CUI, Hao. "High resolution wideband DOA estimation based on modified MUSIC algorithm", Information and Automation - International Conference, p.20-22, Jun.2008.

## ANEXO I

### Lema

Dado que

$$\Phi_x(e^{j\omega}) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j\omega\tau} d\tau \quad (\text{I.1})$$

Em que

$$R_x(\tau) = \int_{-\infty}^{\infty} [x_i(n) * h_i(n)]^* [x_j(n + \tau) * h_j(n + \tau)] dn \quad (\text{I.2})$$

Então

$$\Phi_x(e^{j\omega}) = H_i^*(e^{j\omega}) X_i^*(e^{j\omega}) H_j(e^{j\omega}) X_j(e^{j\omega}) \quad (\text{I.3})$$

### Demonstração

Substituindo (I.2) em (I.1), e utilizando a definição de convolução temos (I.4):

$$\Phi_x(e^{j\omega}) = \int_{\tau=-\infty}^{\infty} \int_{n=-\infty}^{\infty} \int_{\alpha=-\infty}^{\infty} x_i^*(\alpha) h_i^*(n - \alpha) d\alpha \int_{\beta=-\infty}^{\infty} x_j(\beta) h_j(n + \tau - \beta) d\beta dn e^{-j\omega\tau} d\tau \quad (\text{I.4})$$

Note que  $x_i^*(\alpha)$  é constante para os integrais de  $n, \beta$  e  $\tau$ .

Note também que  $x_j(\beta)$  é constante para os integrais em  $n$  e  $\tau$ .

Logo, podemos trocar as posições dos integrais, obtendo:

$$\Phi_x(e^{j\omega}) = \int_{\alpha=-\infty}^{\infty} x_i^*(\alpha) \left\{ \int_{\beta=-\infty}^{\infty} x_j(\beta) \left[ \int_{n=-\infty}^{\infty} h_i^*(n - \alpha) \left( \int_{\tau=-\infty}^{\infty} h_j(n + \tau - \beta) e^{-j\omega\tau} d\tau \right) dn \right] d\beta \right\} d\alpha \quad (\text{I.5})$$

Resolvendo os integrais de dentro para fora:

$$\int_{\tau=-\infty}^{\infty} h_j(n + \tau - \beta) e^{-j\omega\tau} d\tau = H_j(e^{j\omega}) e^{j\omega n} e^{-j\omega\beta} \quad (\text{I.6})$$

Note que para  $f(x)$  real, temos que:

$$f^*(x) \rightarrow F^*(-e^{j\omega}) \quad (\text{I.7})$$

Então:

$$\int_{n=-\infty}^{\infty} h_i^*(n - \alpha) e^{j\omega n} dn = H_i^*(e^{j\omega}) e^{j\omega\alpha} \quad (\text{I.8})$$

Substituindo, tem-se que:

$$\Phi_x(e^{j\omega}) = \int_{\alpha=-\infty}^{\infty} x_i^*(\alpha) \int_{\beta=-\infty}^{\infty} x_j(\beta) H_i^*(e^{j\omega}) e^{j\omega\alpha} H_j(e^{j\omega}) e^{-j\omega\beta} d\beta d\alpha \quad (\text{I.9})$$

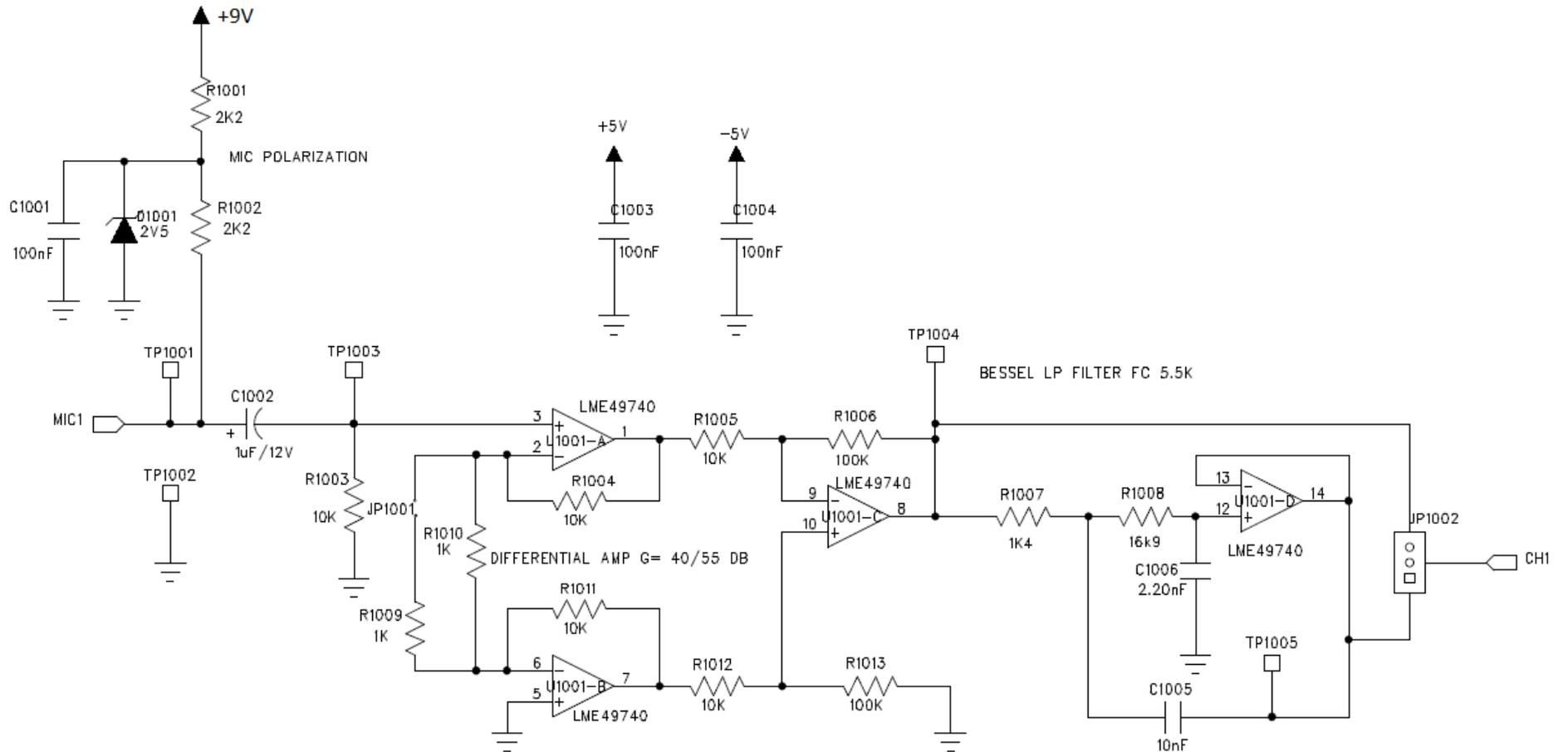
$$\Phi_x(e^{j\omega}) = H_i^*(e^{j\omega}) H_j(e^{j\omega}) \int_{\alpha=-\infty}^{\infty} x_i^*(\alpha) e^{j\omega\alpha} d\alpha \int_{\beta=-\infty}^{\infty} x_j(\beta) e^{-j\omega\beta} d\beta \quad (\text{I.10})$$

E, finalmente:

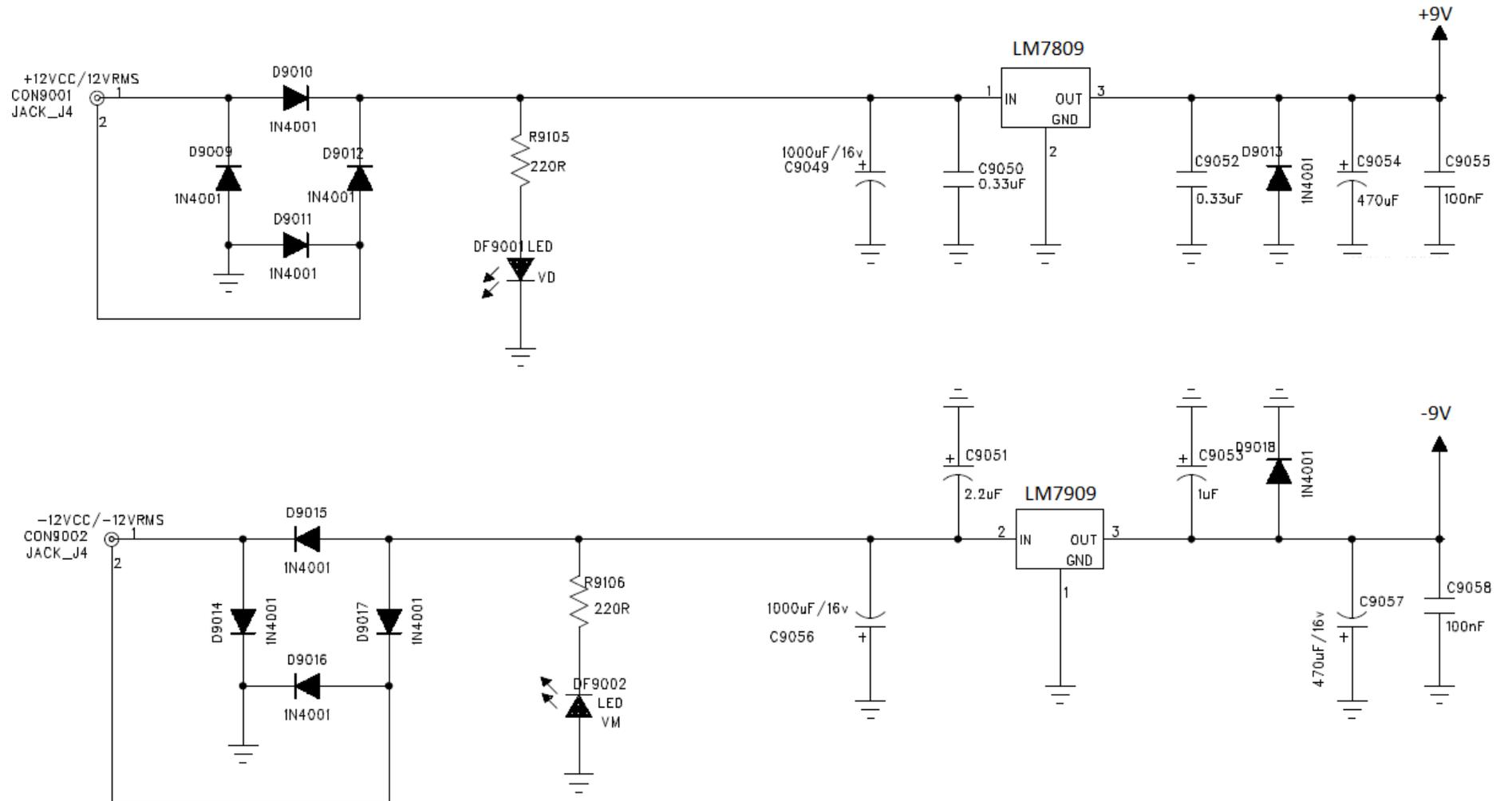
$$\Phi_x(e^{j\omega}) = H_i^*(e^{j\omega}) H_j(e^{j\omega}) X_i^*(e^{j\omega}) X_j(e^{j\omega}) \quad (\text{I.11})$$

## ANEXO II

### I.I Canais analógicos de 1 a 8 (CH1 a CH8)

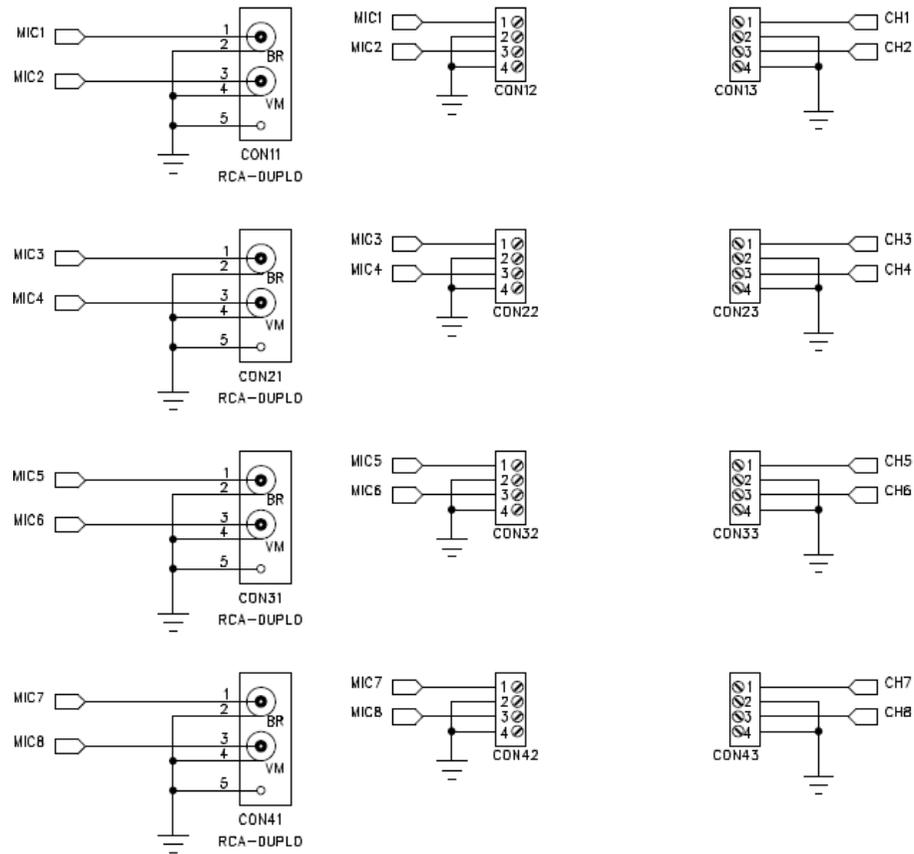


## I.II Fonte simétrica $\pm 9V$

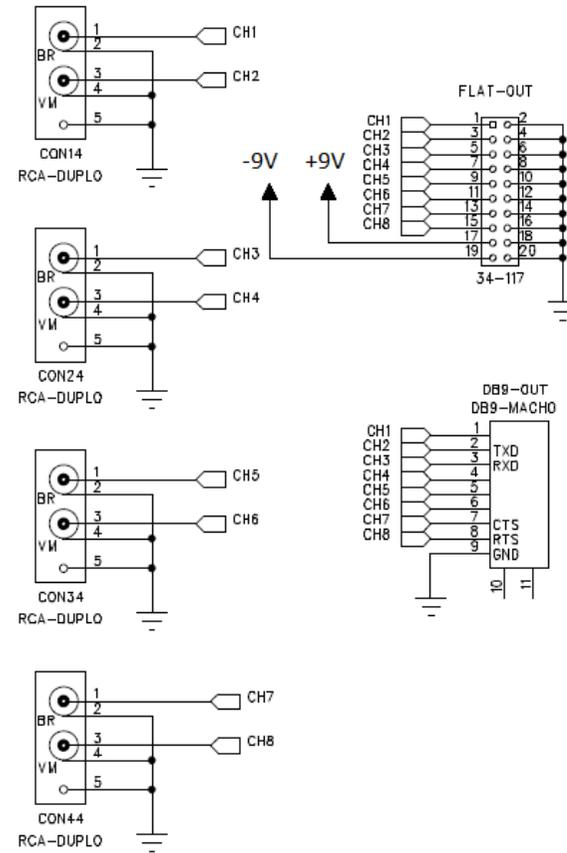


### I.III Conectores de entrada e saída

#### ENTRADAS



#### SAÍDAS



## I.IV Circuito impresso

