

UNIVERSIDADE FEDERAL FLUMINENSE
ESCOLA DE ENGENHARIA
MESTRADO EM ENGENHARIA DE TELECOMUNICAÇÕES

JOÃO ALFREDO CAL BRAZ

MÁQUINA DE VETOR DE RELEVÂNCIA
APLICADA À ESTIMAÇÃO DO CANAL
MULTIPERCURSO FAIXA-LARGA

NITERÓI
2010

JOÃO ALFREDO CAL BRAZ

MÁQUINA DE VETOR DE RELEVÂNCIA
APLICADA À ESTIMAÇÃO DO CANAL
MULTIPERCURSO FAIXA-LARGA

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Telecomunicações da Universidade Federal Fluminense, como requisito parcial para obtenção do título de Mestre. Área de Concentração: Sistemas de Telecomunicações.

Professores Orientadores:

Leni Joaquim de Matos, D. Sc. (TET/UFF)

Edson Luiz Cataldo Ferreira, D. Sc. (GMA/UFF)

Niterói
2010

Ficha Catalográfica elaborada pela Biblioteca da Escola de Engenharia e Instituto de Computação da UFF

B827 Braz, João Alfredo Cal.

Máquina de vetor de relevância aplicada à estimação do canal multipercurso faixa-larga / João Alfredo Cal Braz - Niterói, RJ : [s.n.], 2010.

101 f.

Orientadores: Leni Joaquim de Matos, Edson Cataldo Ferreira.
Dissertação (Mestrado em Engenharia de Telecomunicações) -
Universidade Federal Fluminense, 2010.

1. Sistema de comunicação móvel. 2. Sondagem. 3. Teoria Bayesiana de decisão estatística. 4. Máquina de vetor de relevância.
I. Título.

CDD 621.38456

JOÃO ALFREDO CAL BRAZ

**MÁQUINA DE VETOR DE RELEVÂNCIA
APLICADA À ESTIMAÇÃO DO CANAL
MULTIPERCURSO FAIXA-LARGA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Telecomunicações da Universidade Federal Fluminense, como requisito parcial para obtenção do título de Mestre. Área de Concentração: Sistemas de Telecomunicações.

Aprovada em outubro de 2010.

BANCA EXAMINADORA

Prof.^a Dr.^a LENI JOAQUIM DE MATOS - Orientadora
Universidade Federal Fluminense

Prof. Dr. EDSON LUIZ CATALDO FERREIRA - Co-orientador
Universidade Federal Fluminense

Prof. Dr. GLÁUCIO LIMA SIQUEIRA
Pontifícia Universidade Católica do Rio de Janeiro

Prof. Dr. RODOLFO SABOIA LIMA DE SOUZA
Inmetro

Prof.^a Dr.^a JACQUELINE SILVA PEREIRA
Universidade Federal Fluminense

Niterói
2010

A Deus, por sempre iluminar meu caminho.
A minha família, presente em todas as minhas conquistas.

Agradecimentos

Aos meus orientadores, professores Leni Joaquim de Matos e Edson Cataldo, por toda a disponibilidade e presteza. Seus conselhos foram fundamentais para a elaboração deste trabalho.

Ao Inmetro e ao CETUC/PUC-Rio, por disponibilizarem os equipamentos necessários para a realização das medições.

Aos meus colegas: Pedro Castellanos, cuja experiência contribuiu com valiosas sugestões; Tiago Francisco e Pedro Vieira, com quem pude ter boas conversas e receber boas idéias.

A minha mãe, Dolores, pela pessoa admirável que é, e por ter oferecido todo o apoio e carinho ao longo desta jornada. À Rejane, minha noiva, em cujo sucesso me espelho para alcançar meus objetivos.

Resumo

O conhecimento das características do canal rádio é fundamental para o projeto de sistemas de comunicação sem fio de alto desempenho. Neste trabalho, é implementada uma sonda para a obtenção da resposta faixa-larga do canal, que visa fornecer dados experimentais sobre as condições de propagação do sinal no ambiente em estudo. A correta interpretação dos dados obtidos na fase de sondagem possui papel importante na construção de modelos representativos do canal, de maneira que torna-se importante a utilização de métodos que realizem esta tarefa de forma precisa e confiável. A solução implementada utiliza o algoritmo denominado Máquina de Vetor de Relevância (*Relevance Vector Machine, RVM*), que consiste de um método de regressão baseado na Estatística Bayesiana, para identificar os parâmetros constitutivos da resposta ao impulso do canal. Os resultados fornecidos pelo RVM possibilitam a construção de modelos estatísticos representativos do ambiente.

Palavras-chave: sondagem, modelagem do canal, Estatística Bayesiana, Máquina de Vetor de Relevância.

Abstract

The knowledge of the characteristics of the channel is fundamental to the design of high performance wireless communication systems. Herein is described the design of a wideband channel sounder, intended to supply experimental data regarding the propagation conditions of the signal over the scenario under study. The proper interpretation of the data acquired in the sounding stage plays an important role in the construction of channel models, and the use of methods that perform this task in accurate and reliable means becomes important. The implemented solution makes use of the Relevance Vector Machine (RVM) algorithm, that consists of a regression method based on Bayesian Statistics, to identify the constitutive parameters of the channel impulse response. The outputs of the estimation algorithm enables the construction of statistical models that properly represent the studied environment.

Keywords: sounding, channel modeling, Bayesian Statistics, Relevance Vector Machine.

Sumário

1	Introdução	1
1.1	Objetivos da dissertação	3
2	Caracterização do canal rádio faixa larga	5
2.1	Canais determinísticos	6
2.1.1	Funções de espalhamento de retardo e espalhamento Doppler	6
2.1.2	Funções de transferência variante no tempo e de retardo / espalhamento-Doppler	8
2.2	Canais aleatórios variáveis no tempo	10
2.2.1	Considerações sobre canais reais	11
2.3	Caracterização de canais em pequena e larga escala	15
2.3.1	Caracterização em pequena escala	15
2.3.2	Parâmetros de dispersão do canal rádio móvel	16
2.3.3	Caracterização em larga escala	18
3	Análise Bayesiana	19
3.1	Estatística Clássica	19
3.2	Estatística Bayesiana	22
3.3	Modelagem Bayesiana	26
3.4	Regressão Bayesiana e Aprendizagem Supervisionada	28
4	Técnicas de sondagem	37
4.1	Técnica de sondagem no domínio da frequência	38
4.2	Técnica de sondagem no domínio do tempo	38

4.2.1	Sondagem por pulsos periódicos	39
4.2.2	Sondagem por compressão de pulsos	39
5	Sistema de sondagem e cenário das medições	44
5.1	Projeto da sonda faixa larga	44
5.2	Locais de medição	51
6	Ferramentas de Processamento	54
6.1	Implementação do filtro casado	54
6.2	A Máquina de Vetor de Relevância Aplicada à Estimação do Canal Multi- percurso	57
7	Resultados e Discussão	69
7.1	Avaliação do desempenho do algoritmo RVM	69
7.2	Modelagem estatística do canal	76
8	Conclusões e Trabalhos Futuros	83
8.1	Conclusões	83
8.2	Sugestões para Trabalhos Futuros	85

Lista de Figuras

2.1	Linha de retardos com derivações que representa a função de espalhamento de retardo.	7
2.2	Banco de filtros e conversores de frequência.	8
2.3	Relações entre as formas da função que caracteriza o canal rádio variável no tempo.	10
2.4	Relações entre as funções correlação do canal.	12
2.5	Relações entre as funções correlação em canais WSSUS.	14
3.1	Esquema interpretativo das inferências clássica e bayesiana.	25
3.2	Esquema interpretativo do conceito da navalha de Occam na comparação de modelos.	28
3.3	Geração do par entrada-alvo do problema de ajuste polinomial.	30
3.4	Ajuste polinomial. As curvas verdes representam a função geradora g ; os pontos azuis representam os valores alvo t ; as curvas vermelhas representam funções de ajuste, para polinômios de diferentes ordens. (a) $M = 0$, (b) $M = 1$, (c) $M = 3$, (d) $M = 9$	31
3.5	Esquemático ilustrativo da distribuição condicional gaussiana para t dado x	33
4.1	Sondagem no domínio da frequência através do analisador de rede vetorial.	38
4.2	Técnica de sondagem por varredura de retardo de tempo e correlação cruzada.	41
4.3	Técnica de sondagem por convolução por filtro casado.	42
5.1	Sistema empregado na sondagem do canal sem fio multipercorso.	45
5.2	Trecho de uma sequência pseudo-aleatória típica.	46

5.3	Gráficos da (a) densidade espectral de potência e (b) autocorrelação da sequência PN transmitida.	48
5.4	Levantamento da curva de linearidade do amplificador de potência para a frequência de 1,95 GHz.	48
5.5	Medição da perda de retorno da antena utilizada.	49
5.6	Fotografia do sistema de recepção da sonda.	51
5.7	Planta do local das medições. O triângulo vermelho representa a posição da antena transmissora. Os círculos azuis representam os pontos de medição.	52
5.8	Fotografias do edifício utilizado na sondagem do canal, relacionadas aos pontos de medição da planta da Fig. 5.7. (a) ponto 3, (b) ponto 4, (c) ponto 6, (d) ponto 8.	53
6.1	Sinal na saída do filtro casado. (a) Ponto de medição em LOS. (b) Ponto de medição em NLOS.	55
6.2	Modelo banda base equivalente da sondagem faixa larga do canal rádio.	57
6.3	Processo iterativo de aprendizagem dos parâmetros do modelo do canal multipercurso.	62
6.4	Diagrama de atividades do algoritmo implementado.	67
6.5	Exemplos do resultado da aplicação do RVM a amostras dos dados colhidos na sondagem do canal. (a) situação LOS, (b) situação NLOS.	68
7.1	Função distribuição cumulativa do TDM para dois valores de ρ . (a) sem adição de ruído ($RSR > 30$), (b) $RSR = 30$ dB, (c) $RSR = 25$ dB, (d) $RSR = 20$ dB.	74
7.2	Função distribuição cumulativa do TDM para dois valores de f_s . (a) sem adição de ruído ($RSR > 30$), (b) $RSR = 30$ dB, (c) $RSR = 25$ dB, (d) $RSR = 20$ dB.	74
7.3	Função distribuição cumulativa do TDM para três valores de P . (a) sem adição de ruído ($RSR > 30$ dB), (b) $RSR = 30$ dB, (c) $RSR = 25$ dB, (d) $RSR = 20$ dB.	75
7.4	Gráfico da função de convergência em função do número de realizações da resposta do canal. Configuração utilizada: $RSR = 20$ dB, $\rho = 0,90$, $f_s = 12,5$ amostras/bit.	76
7.5	Grade de medições com 36 posições.	77

7.6	Gráfico de dispersão do espalhamento de retardo calculado na posição 15 da grade, para os 12 locais de medição. A numeração dos pontos corresponde ao local de medição especificado na planta da Fig. 5.7. (a)LOS e (b)NLOS.	77
7.7	Gráfico de dispersão da média do espalhamento de retardo com 95% de nível de confiança para os 12 locais medidos. (a)LOS e (b)NLOS.	78
7.8	Histograma do espalhamento de retardo e curvas de ajuste de distribuições de probabilidade. (a)LOS e (b)NLOS.	79
7.9	Histograma do tempo de interchegada de multipercursos e curvas de ajuste de distribuições de probabilidade. (a)LOS e (b)NLOS.	81
7.10	Histograma da fração da energia total concentrada na componente de multipercorso mais forte e curvas de ajuste de distribuições de probabilidade. (a)LOS e (b)NLOS.	82

Lista de Tabelas

7.1	Resultados da avaliação de desempenho do algoritmo. A primeira linha de dados corresponde à configuração de referência.	73
7.2	Especificações das distribuições na modelagem de τ_{rms} em LOS.	80
7.3	Especificações das distribuições na modelagem de τ_{rms} em NLOS.	80
7.4	Especificações da distribuição não paramétrica na modelagem do tempo de interchegada nos casos LOS e NLOS.	81
7.5	Especificações da distribuição não paramétrica na modelagem do tempo de interchegada nos casos LOS e NLOS.	82

Capítulo 1

Introdução

A comunicação sem fio é, hoje, uma das áreas em maior desenvolvimento no campo das telecomunicações. O aumento drástico da demanda por conectividade livre de fios, aliado a progressos na ciência e tecnologia, como a integração de circuitos em grande escala (*very large scale integration*, VLSI) e a implementação de sofisticados algoritmos de processamento e técnicas de codificação, além do grande sucesso das tecnologias de comunicações móveis de segunda e terceira geração, são demonstrações de que idéias advindas de teoria das comunicações podem ter um impacto significativo na prática [1]. Tais avanços fornecem um conjunto de perspectivas ainda mais ricas sobre os futuros desdobramentos das comunicações sem fio e sua influência no cotidiano das pessoas.

Os sistemas celulares, por exemplo, têm experimentado um crescimento exponencial nas últimas décadas, com a expectativa de alcançar 5 bilhões de usuários em todo mundo, ao fim de 2010 [2], tornando-se uma ferramenta indispensável para uma parcela significativa da população mundial. De forma semelhante, as redes locais sem fio tornaram-se populares e têm complementado ou substituído as redes cabeadas em muitas residências, escritórios e *campi*. Aplicações como redes de sensores, fábricas automatizadas, casas inteligentes e tele-medicina são exemplos emergentes que fazem uso de tecnologias de comunicação sem fio [3].

A eficiente implementação de um sistema de comunicação sem fio exige que, para o local em questão, seja realizada a caracterização do canal de radiopropagação, a fim de que possam ser projetadas soluções como sistemas de equalização adequados nos equi-

pamentos transmissores e receptores. Em última instância, a correta caracterização do canal possibilita um planejamento mais preciso da rede, de forma que sejam especificados, de maneira acurada, a região de cobertura, a capacidade do enlace e que possam ser implementadas técnicas de combate aos efeitos nocivos da interferência inter-simbólica, decorrentes da propagação em multipercurso, típicos do canal rádio, como recepção em diversidade (receptores Rake), modulação multiportadora (modulação OFDM) ou espalhamento espectral (salto em frequência e sequência direta).

A caracterização do canal é comumente realizada através da análise estatística de um grande número de medições realizadas no ambiente de interesse. Desta forma, é possível compreender o efeito de perda no percurso e da dispersão temporal que influenciam na propagação do sinal. Esta última é um fenômeno de particular relevância em locais fechados (locais *indoor*), visto que a existência de múltiplos percursos entre transmissor e receptor afetam o desempenho do sistema de forma acentuada.

A comunicação rádio *indoor* cobre uma ampla variedade de situações, desde comunicação entre indivíduos em residências ou escritórios comerciais, supermercados ou *shopping centers*, até o envio de mensagens a robôs em movimento em linhas de montagem das unidades fabris, num futuro próximo. Um típico ambiente de radiopropagação *indoor* consiste de uma estação base fixa, composta por uma antena em posição elevada, comunicando-se com dispositivos portáteis, localizados a alturas iguais ou inferiores a 2 metros, em um escritório, por exemplo. Devido à existência de grande quantidade de objetos que funcionam como obstáculos à propagação do sinal, ocasionando fenômenos de reflexão, refração e espalhamento das ondas pelas estruturas no escritório, o sinal transmitido percorre vários caminhos e chega ao receptor em tempos diferentes, causando o chamado desvanecimento por multipercurso. Este comportamento do sinal em propagação levou Turin [4] à modelagem da resposta impulsiva do canal (RIC), cujos parâmetros são aleatoriamente variáveis no tempo.

Vários autores [5, 6, 7] já se dedicaram ao estudo do canal de radiopropagação *indoor*. Hashemi [8] foi um dos pioneiros a explorar o assunto, descrevendo o comportamento de parâmetros como o retardo médio em excesso, $\bar{\tau}$, e o espalhamento de retardo eficaz,

τ_{rms} , em ambientes fechados de escritórios. Rappaport [6] realizou medições importantes de caracterização de canal faixa-larga em ambientes fabris. Saleh e Valenzuela [5] também realizaram medições *indoor* e propuseram um modelo de perfil de potência de retardo baseado em *clusters*, que correspondem a agrupamentos de componentes de multipercursos provenientes de um mesmo espalhador.

A relevância dos parâmetros que podem ser extraídos do perfil de potência de retardo levou à busca por métodos de processamento da informação que consigam fazer a identificação dos ecos válidos, em meio à contribuição do processo estocástico do ruído inerente ao canal de radiopropagação. Entre as técnicas comumente empregadas destaca-se o algoritmo CLEAN [9], introduzido para aplicações em rádio-astronomia [10, 11], e que tem sido aplicado em medidas de canal faixa-larga e UWB [12, 13, 10]. Ele é baseado em um processo de iterações que identifica os instantes do retardos e suas amplitudes, baseados nos pontos de máximo da função de correlação cruzada entre o sinal medido e a resposta de uma medição de referência, tomada em visada em relação ao transmissor. O algoritmo SAGE [14] constitui outro método de estimação conjunta de parâmetros do canal como retardo e frequência Doppler. Foi formulado por Dempster et al. [15] como um método iterativo para solução de problemas de estimação de máxima verossimilhança, em situações onde parte dos dados está ausente ou censurada. Outro método com considerável destaque é o MUSIC [16], que se utiliza do tratamento de autovalores e autovetores, gerados a partir de resposta em frequência do canal, para estimar o número de componentes de multipercurso, utilizando limiares ótimos de patamar de ruído, em vez de limiares fixos, como usados em métodos convencionais.

1.1 Objetivos da dissertação

O objetivo primordial deste trabalho consiste na utilização do método de maximização da evidência para realizar a modelagem do canal de radiopropagação. Este método utiliza-se das bases da teoria Bayesiana para a construção de um algoritmo de aprendizagem de máquina, denominado Máquina de Vetor de Relevância (*Relevance Vector Machine*, RVM). Diante dos desafios expostos sobre o estudo do canal de radiopropaga-

ção *indoor* faixa-larga, as seguintes tarefas foram realizadas:

- Medições no canal, a partir da implementação de um sistema de sondagem no domínio do tempo, com alta resolução de multipercursos;
- Demonstração da eficácia da aplicação da análise Bayesiana, usando a maximização da evidência, no problema de estimação do canal de radiopropagação;
- Determinação dos valores dos parâmetros de dispersão temporal, com o objetivo de modelar o canal de radiopropagação em questão.

A dissertação está estruturada em oito capítulos, incluindo o capítulo introdutório:

- Capítulo 2: apresenta os conceitos básicos de canais de radiopropagação faixa-larga.
- Capítulo 3: introduz os conceitos da Estatística Bayesiana e sua utilização na regressão de dados.
- Capítulo 4: relaciona os métodos de sondagem do canal rádio, com ênfase na técnica utilizada.
- Capítulo 5: descreve o sistema de sondagem que foi montado e a caracterização do ambiente de medição.
- Capítulo 6: apresenta as ferramentas de processamento aplicadas aos dados obtidos nas medições.
- Capítulo 7: apresenta os resultados alcançados e sua discussão.
- Capítulo 8: apresenta as conclusões e propostas para futuros trabalhos.

Capítulo 2

Caracterização do canal rádio faixa larga

O cenário típico de um sistema de comunicação sem fio, em ambientes fechados, é composto por um elemento base realizando troca de dados com uma estação ou um conjunto de estações espalhadas pelo ambiente, posicionadas tanto em visada quanto fora de visada. Os diversos fenômenos de propagação como sombreamento, reflexões, difrações e espalhamentos impõem limitações claras ao desempenho de sistemas de comunicações sem fio. Neste panorama, são distinguidos dois tipos de manifestações atuantes no canal de propagação, denominados desvanecimentos [17]: 1) o desvanecimento de grande escala, que ocorre devido à existência de obstáculos de dimensões grandes, que fazem com que o receptor esteja sombreado pelo contorno destas proeminências. As estatísticas proveem uma forma de estimar a perda no percurso em função da distância do enlace, sendo dada em termos da perda média no percurso e da variação em torno da média; 2) o desvanecimento de pequena escala, que corresponde às drásticas alterações em amplitude e fase do sinal. Ele se manifesta através de dois mecanismos, chamados espalhamento temporal do sinal e comportamento variante no tempo do canal.

Relativo aos desvanecimentos de pequena escala, o espalhamento temporal do sinal é decorrente da chamada propagação em multipercursos. Isto ocorre devido aos trajetos de comprimentos distintos que os raios podem fazer entre o transmissor e receptor, que

causarão a chegada, em instantes diferentes, de cada componente do sinal na estação de destino. Já a variância do canal, vista no domínio do tempo, advém da diferença de posicionamento entre o transmissor e o receptor, que resulta em mudanças nos caminhos de propagação, ou pelo simples movimento dos espalhadores no ambiente. Identifica-se que para espalhadores localizados sobre uma elipse que contém transmissor e receptor como focos, os raios espalhados sobre cada um destes objetos chegarão no mesmo instante de tempo ao receptor. Estes raios podem ser diferenciados apenas pelo seu ângulo de chegada [18].

A fim de melhor compreender os parâmetros que descrevem as características do canal, Bello [19], aprimorou os trabalhos desenvolvidos em [20, 21] e propôs a modelagem do canal de radiopropagação como um filtro linear variante no tempo. Foi, então, introduzida a função do sistema, que descreve o comportamento do canal e pode ser descrita em função da dispersividade temporal ou de frequência, bastando para isto que seja empregada uma transformada dupla de Fourier.

Primeiramente, será apresentada a formulação para canais que variam no tempo de forma determinística. Este desenvolvimento será usado como base para a apresentação dos canais aleatórios variáveis no tempo.

2.1 Canais determinísticos

2.1.1 Funções de espalhamento de retardo e espalhamento Doppler

Por questão de conveniência, será utilizada a representação de envoltória complexa passa-baixa para reproduzir sinais reais passa-faixa. A relação entre as duas formas pode, sumariamente, ser expressa pela Eq. (2.1):

$$x(t) = \text{Re}[z(t)e^{j2\pi f_0 t}], \quad (2.1)$$

onde $x(t)$ representa o sinal passa-faixa e $z(t)$, sua conversão para envoltória complexa. A frequência da portadora é dada por f_0 .

Nosso interesse inicial é representar o sinal na saída do sistema como função de sua resposta ao impulso e do sinal de entrada. Visto que o canal é variante no tempo, a resposta impulsiva do filtro que representa o canal também deve variar no tempo. A envoltória complexa da resposta impulsiva do filtro com estas especificações é dada por $h(t, \tau)$, onde τ representa a variável retardo.

A envoltória complexa do sinal à saída do filtro relaciona-se com o sinal à sua entrada através da Eq. (2.2):

$$\begin{aligned} w(t) &= z(t) * h(t, \tau) \\ &= \int_{-\infty}^{\infty} z(t - \tau) h(t, \tau) d\tau. \end{aligned} \quad (2.2)$$

Fisicamente, $h(t, \tau)$ pode ser interpretado como a resposta ao impulso no instante t , devido a uma entrada impulsiva aplicada há τ segundos atrás. Para que o canal seja fisicamente realizável, não pode haver saída antes de a entrada ter sido aplicada. Assim, restringe-se $h(t, \tau)$ para $\tau \geq 0$. Logo, para um período de observação igual a T , os limites de integração da equação passam a ser $(0, T)$.

A Eq. (2.2) representa o canal como uma soma contínua de espalhadores estacionários [19], onde $h(t, \tau) d\tau$ representa a modulação complexa produzida pelos hipotéticos espalhadores elementares, que proveem retardos na faixa de $(\tau, \tau + \Delta\tau)$. A Fig. 2.1 ilustra a interpretação física como uma linha de retardos com múltiplas derivações.

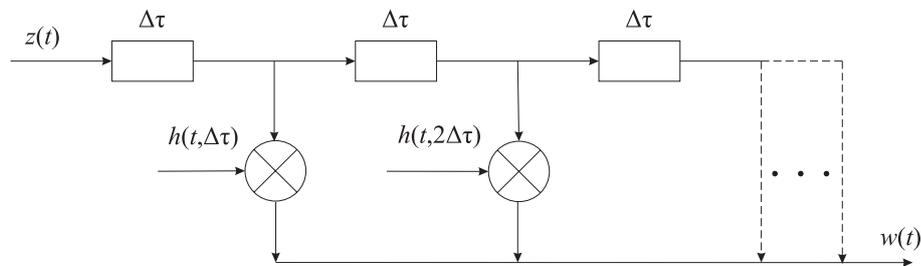


Figura 2.1: Linha de retardos com derivações que representa a função de espalhamento de retardo.

No domínio da frequência, a caracterização do canal é possível através da função dual à resposta impulsiva variante no tempo, $h(t, \tau)$. A função $H(f, \nu)$, chamada função de espalhamento Doppler, relaciona o espectro de saída do canal ao espectro de entrada, assim como $h(t, \tau)$ relaciona as funções de entrada e saída no domínio do tempo. A representação do espectro de saída é dada como a superposição de réplicas do espectro de entrada $Z(f)$ com diferentes deslocamentos Doppler e filtradas, resultando em $W(f)$, dada pela Eq. (2.3):

$$W(f) = \int_{-\infty}^{\infty} Z(f - \nu)H(f - \nu, \nu) d\nu. \quad (2.3)$$

De maneira análoga à linha de retardos da função de espalhamento de retardo, a Eq. (2.3) possibilita que o canal seja interpretado como uma cadeia densa de conversão de frequência, como exhibe a Fig. 2.2. A coleção de funções $H(f, \nu) d\nu$ representam o banco de filtros (operação dual à modulação), seguidos por conversores de frequência de deslocamento Doppler, que produzem deslocamentos na faixa de $(\nu, \nu + \Delta\nu)$ hertz.

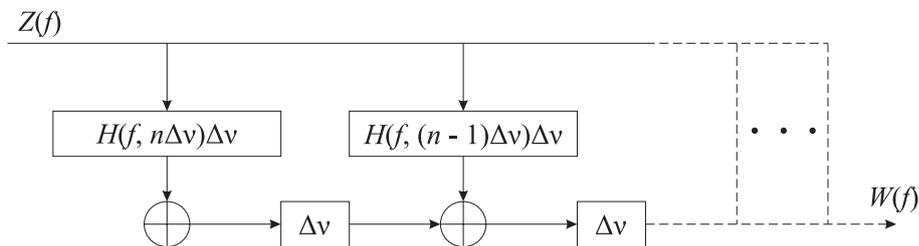


Figura 2.2: Banco de filtros e conversores de frequência.

2.1.2 Funções de transferência variante no tempo e de retardo / espalhamento-Doppler

O comportamento do canal também pode ser expresso através da chamada função de transferência variante no tempo. Ela descreve a variação do canal na frequência, ao longo do tempo. É, portanto, a transformada de Fourier da função de espalhamento de retardo em relação à variável retardo. Considerando $h(t, \tau)$ a resposta ao impulso do canal variante no tempo e $T(f, t)$, a função de transferência variante no tempo, temos a

Eq. (2.4):

$$\begin{aligned} T(f, t) &= \mathcal{F}_\tau[h(t, \tau)] \\ &= \int_{-\infty}^{\infty} h(t, \tau) e^{-j2\pi f\tau} d\tau. \end{aligned} \quad (2.4)$$

Esta função, introduzida por [20], estabelece a resposta espectral de um sistema variante no tempo, como exhibe a Eq. (2.5):

$$w(t) = \int_{-\infty}^{\infty} Z(f)T(f, t)e^{j2\pi ft} df. \quad (2.5)$$

A quarta forma de se representar o canal faixa larga é através da função no domínio do retardo e do deslocamento Doppler. A partir da função do sistema no domínio retardo-Doppler, obtém-se a função de espalhamento no tempo através da transformada inversa de Fourier em respeito à variável desvio Doppler.

$$h(t, \tau) = \int_{-\infty}^{\infty} S(\tau, \nu) e^{j2\pi\nu t} d\nu. \quad (2.6)$$

Realizando a substituição na Eq. (2.2), chega-se à Eq. (2.7):

$$w(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t - \tau) S(\tau, \nu) e^{j2\pi\nu t} d\nu d\tau. \quad (2.7)$$

Esta expressão apresenta a saída do filtro como uma soma de sinais atrasados e com deslocamento Doppler. Cada componente apresenta uma amplitude de espalhamento diferencial $S(\tau, \nu) d\nu d\tau$.

Como se pode verificar na Fig. 2.3, as quatro formas de se expressar a função de sistema são interrelacionadas através da transformada de Fourier em relação a diferentes domínios.

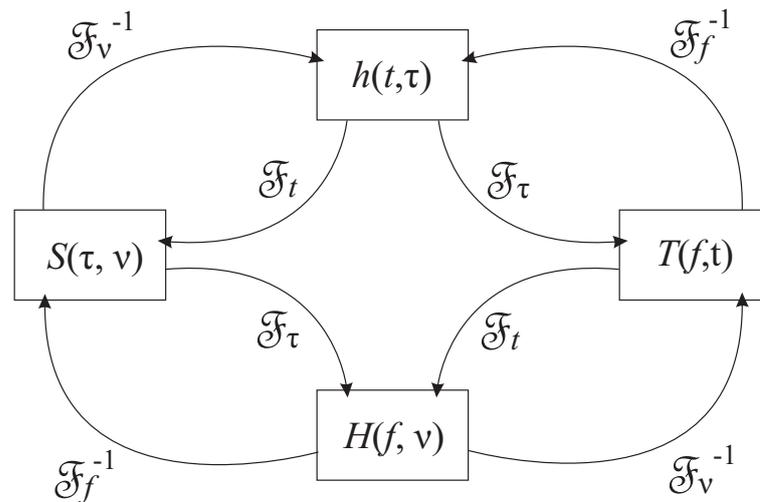


Figura 2.3: Relações entre as formas da função que caracteriza o canal rádio variável no tempo.

2.2 Canais aleatórios variáveis no tempo

Em verdade, o canal rádio móvel não é um canal determinístico variante no tempo. Em vez disto, ele apresenta variação aleatória, de forma que as funções do sistema apresentadas para o caso determinístico transformam-se em processos estocásticos. A caracterização estatística exata do canal com estas características é dada em termos de funções densidade de probabilidade multidimensionais, as quais requerem mais conhecimento do que os dados disponíveis em situações práticas. O modo alternativo, eficaz e mais prático, envolve a caracterização estatística em termos das correlações das funções do sistema, já que o conhecimento destas possibilita a determinação da função autocorrelação da saída do canal.

Se $R_w(t, s)$ é a autocorrelação do processo estocástico da saída do sistema, considerando o sinal de entrada $z(t)$ determinístico, partindo da Eq. (2.1), chega-se a:

$$\begin{aligned}
R_w(t, s) &= \mathbb{E}[w(t)w^*(s)] \\
&= \mathbb{E} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t - \tau)z^*(s - \eta)h(t; \tau)h^*(s; \eta) d\tau d\eta \right] \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t - \tau)z^*(s - \eta)\mathbb{E}[h(t; \tau)h^*(s; \eta)] d\tau d\eta \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t - \tau)z^*(s - \eta)R_h(t, s; \tau, \eta) d\tau d\eta. \tag{2.8}
\end{aligned}$$

De maneira semelhante, podem ser obtidas as seguintes funções:

$$R_w(\tau, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(\tau - \nu)z^*(\eta - \mu)R_S(\tau, \eta; \nu, \mu) d\nu d\mu, \tag{2.9}$$

$$R_w(\nu, \mu) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(\nu - f)z^*(\mu - m)R_H(f, m; \nu, \mu) df dm, \tag{2.10}$$

$$R_w(f, m) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(f - t)z^*(m - s)R_T(f, m; t, s) dt ds, \tag{2.11}$$

onde o valor esperado é representado por $\mathbb{E}[\cdot]$ e o conjugado é representado por $*$. As variáveis t e s representam instantes de tempo, τ e η denotam variáveis de retardo no tempo, f e m representam variáveis de frequência, ν e μ representam deslocamentos em frequência. Enquanto as funções de sistema se relacionavam através de transformadas de Fourier simples, as funções de correlação das funções de sistema se relacionam através de transformadas de Fourier duplas.

2.2.1 Considerações sobre canais reais

Os canais reais, modelados como aleatórios, abordados de forma genérica na seção 2.2, podem ser tratados de forma mais específica se submetidos a certas restrições. Tais restrições se devem ao fato de os canais reais possuírem estatísticas de desvane-

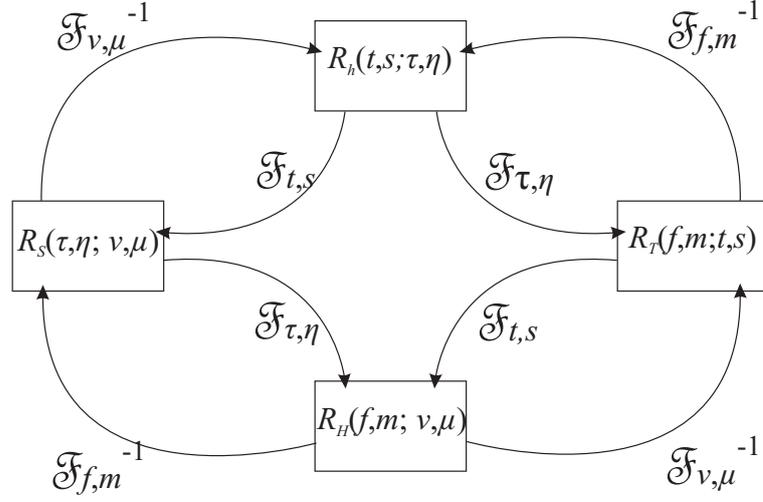


Figura 2.4: Relações entre as funções correlação do canal.

cimento que podem ser consideradas estacionárias sobre curtos intervalos de tempo ou sobre o deslocamento de pequenas distâncias (da ordem de grandeza do comprimento de onda). Sob a consideração de que os canais práticos podem ser considerados estacionários no sentido amplo (*wide sense stationary*, WSS), diz-se que a média estatística do processo aleatório associado ao canal é constante e que a sua função autocorrelação é invariável a uma translação no eixo do tempo. Aplicando estas propriedades, introduzimos a variável diferença de tempo $\xi = s - t$ e as funções de correlação R_h e R_T podem ser reescritas de forma direta, pelas Eqs. (2.12) e (2.13):

$$R_h(t, t + \xi; \tau, \eta) = R_h(\xi, \tau, \eta), \quad (2.12)$$

$$R_T(f, m; t, t + \xi) = R_T(f, m; \xi). \quad (2.13)$$

Sabendo-se que R_H e R_S estão ligadas às Eqs. (2.12) e (2.13) através da transformada de Fourier, após desenvolvimento chega-se às Eqs. (2.14) e (2.15):

$$R_S(\tau, \eta; \nu, \mu) = \delta(\nu - \mu) \int_{-\infty}^{\infty} R_h(\xi; \tau, \eta) e^{-j2\pi\nu\xi} d\xi, \quad (2.14)$$

$$R_H(f, m; \nu, \mu) = \delta(\nu - \mu) \int_{-\infty}^{\infty} R_T(f; m, \xi) e^{-j2\pi\nu\xi} d\xi. \quad (2.15)$$

Verifica-se que as Eqs. (2.14) e (2.15) terão valor nulo sempre que $\nu \neq \mu$. Este comportamento das funções de correlação do canal em relação às variáveis de deslocamento Doppler sugere que, para um modelo de canal composto por um número de espalhadores elementares que produzem retardo e deslocamento Doppler, as contribuições destes espalhadores serão descorrelacionadas. Identificam-se as integrais destas equações como as transformadas de Fourier das funções de correlação do sistema denominadas densidade espectral de potência cruzada retardo-Doppler (P_S) e densidade espectral de frequência-Doppler (P_H). Reescrevemos as Eqs. (2.14) e (2.15) como as Eqs. (2.16) e (2.17):

$$R_S(\tau, \eta; \nu, \mu) = \delta(\nu - \mu)P_S(\tau, \eta; \nu), \quad (2.16)$$

$$R_H(f, m; \nu, \mu) = \delta(\nu - \mu)P_H(f, m; \nu). \quad (2.17)$$

Seguindo o conceito de espalhadores descorrelacionados (*uncorrelated scatterers*, US), de forma análoga ao realizado para a obtenção das Eqs. (2.14) e (2.15), as funções autocorrelação das funções do canal também serão nulas sempre que $\eta \neq \tau$, como exibem as Eqs. (2.18) e (2.19):

$$R_h(t, s; \tau, \eta) = \delta(\eta - \tau)P_h(t, s; \tau) \quad (2.18)$$

$$R_S(\tau, \eta; \nu, \mu) = \delta(\eta - \tau)P_S(\tau, \nu; \mu) \quad (2.19)$$

Estas funções autocorrelação relacionam-se com as densidades espectrais de potência de retardo e de retardo-Doppler através das Eqs. (2.20) e (2.21):

$$P_h(t, s; \tau) = \int_{-\infty}^{\infty} R_T(\Omega; t, s) e^{j2\pi\tau\Omega} d\Omega, \quad (2.20)$$

$$P_S(\tau; \nu, \mu) = \int_{-\infty}^{\infty} R_H(\Omega; \nu, \mu) e^{j2\pi\tau\Omega} d\Omega. \quad (2.21)$$

Outra importante contribuição fornecida por Bello [19] foi a demonstração de que os conceitos de espalhadores descorrelacionados e estacionariedade no sentido amplo são duais no tempo-frequência. Logo, um canal US em relação aos retardos pode ser conside-

rado como portador de estatísticas WSS na variável frequência, de forma que as funções autocorrelação dependem puramente da diferença de frequência $\Omega = m - f$.

Chega-se, então, a uma classe de canais denominados WSSUS, que descreve grande parte dos canais rádio. Com a aplicação das restrições pertinentes a este modelo chega-se à forma final das funções de correlação das funções de sistema do canal, dadas pelas Eqs. (2.22) a (2.25).

$$R_h(t, t + \xi; \tau, \eta) = \delta(\eta - \tau)P_h(\xi; \tau), \quad (2.22)$$

$$R_H(f, f + \Omega; \nu, \mu) = \delta(\nu - \mu)P_H(\Omega; \nu), \quad (2.23)$$

$$R_T(f, f + \Omega; t, t + \xi) = R_T(\Omega; \xi), \quad (2.24)$$

$$R_S(\tau, \eta; \nu, \mu) = \delta(\eta - \tau)\delta(\nu - \mu)P_S(\tau; \nu). \quad (2.25)$$

Estas funções se interrelacionam através de transformadas de Fourier simples, como esquematizado na Fig. 2.5.

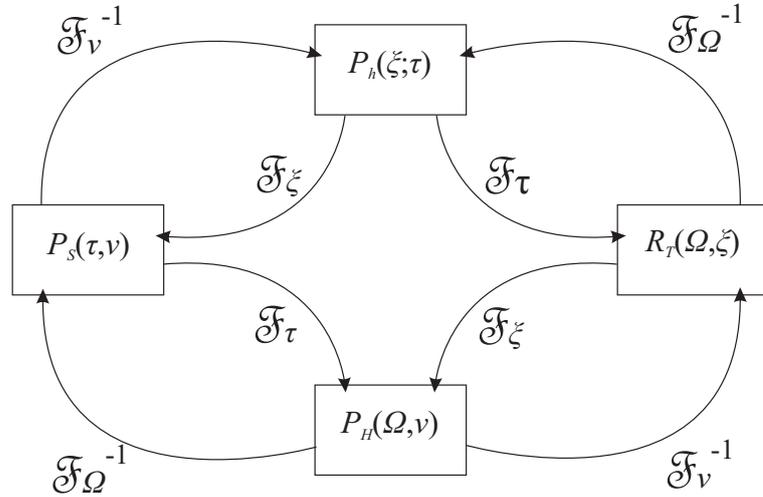


Figura 2.5: Relações entre as funções correlação em canais WSSUS.

2.3 Caracterização de canais em pequena e larga escala

Para a efetiva caracterização de canais reais, mostrou-se necessário que fosse assumida a estacionariedade em pequenos intervalos de tempo para que simplificações pudessem ser feitas às expressões genéricas. Em [19], Bello realiza a completa caracterização do canal em duas etapas. Primeiramente, é feita a análise dentro de um curto período de tempo, intervalo para o qual os centros de espalhamento não mudaram de posição e as funções do canal podem ser consideradas estacionárias. Posteriormente, investiga-se o comportamento em larga escala, de acordo com as estatísticas dos resultados para pequena escala. Os canais em que estas considerações são válidas são chamados *quasi-wide-sense stationary* (QWSS).

2.3.1 Caracterização em pequena escala

As funções de maior interesse para a determinação dos parâmetros de dispersividade temporal do canal são aquelas que derivam da função espalhamento de retardo, $h(t, \tau)$, e seletividade em frequência, $T(f, t)$. Foi demonstrado, nas seções anteriores, que a envoltória complexa do sinal recebido pode ser obtida da autocorrelação do espalhamento de retardo, $R_h(t, s; \tau, \eta)$. Considerando um canal WSSUS, substituímos a Eq. (2.22) na Eq. (2.8) para relacionar R_w com P_h e chegar à Eq. (2.26):

$$R_w(t, t + \xi) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t - \tau) z^*(t + \xi - \eta) \delta(\eta - \tau) P_h(\xi; \tau) d\tau d\eta. \quad (2.26)$$

Quando o tempo de separação entre as observações é nulo, ou seja, num instante de observação, $\xi = 0$ e a densidade espectral cruzada de potência $P_h(\xi; \tau)$ torna-se uma densidade espectral de retardos simples $P_h(\tau)$. A Eq. (2.26) reduz-se à Eq. (2.27):

$$R_w(t, t) = \int_{-\infty}^{\infty} |z(t - \tau)|^2 P_h(\tau) d\tau. \quad (2.27)$$

Para uma entrada impulsional em $t = \tau$, obtém-se a Eq. (2.28):

$$R_w(\tau, \tau) = P_h(\tau). \quad (2.28)$$

Esta equação, válida para canais WSSUS, porta a importante informação de que a autocorrelação da saída do canal é descrita pela distribuição temporal da potência recebida, chamada perfil de potência de retardo.

O comportamento seletivo em frequência do canal móvel pode ser observado através do estudo da autocorrelação da função de transferência variante no tempo, $R_T(\Omega, \xi)$. A partir da informação contida nesta função, é possível verificar a máxima separação em frequência entre sinais aplicados ao sistema cujos sinais de saída exibirão alta correlação, isto é, a resposta do canal para estas frequências apresentará ganho aproximadamente constante e fase linear. Como indica o diagrama da Fig. 2.5, as funções $R_T(\Omega, \xi)$ e $P_h(\xi, \tau)$ se relacionam através da transformada de Fourier.

Fazendo $\xi = 0$, o período de observação torna-se um instante de tempo e a função $R_T(\Omega, \xi)$, reduzida para $R_T(\Omega)$, pode ser chamada de função de correlação de frequência e é escrita em função da densidade espectral de retardos como:

$$R_T(\Omega) = \int_{-\infty}^{\infty} P_h(\tau) e^{-j2\pi\Omega\tau} d\tau. \quad (2.29)$$

O desenvolvimento de técnicas de sondagem do canal visa buscar métodos capazes de alcançar algumas das funções que descrevem o canal, comumente $P_h(t)$ ou $R_T(\Omega)$. A partir da obtenção de uma destas funções, pode-se chegar às demais e a completa caracterização do canal pode ser realizada.

2.3.2 Parâmetros de dispersão do canal rádio móvel

A dispersão do sinal no canal pode ocorrer no domínio do tempo e no domínio da frequência. A dispersão temporal do sinal se deve à existência de espalhadores que fazem com que cheguem ao receptor réplicas atenuadas do sinal transmitido e com diferentes

retardos. Os parâmetros que caracterizam este tipo de dispersão são o retardo médio, o espalhamento de retardo e a banda de coerência. Já a dispersão em frequência ocorre devido à mobilidade do meio ou da estação receptora, que ocasionará a variação dos ângulos de chegada dos multipercursos e o chamado efeito Doppler. Os parâmetros que descrevem a dispersão em frequência são o deslocamento Doppler, o espalhamento Doppler e o tempo de coerência.

Neste trabalho, pelo fato de não ocorrer deslocamento entre a estação transmissora e a receptora, a dispersão relevante para a caracterização do canal é a dispersão temporal. Como a variação temporal da posição dos espalhadores é mínima, visto que as medições foram realizadas em horários em que não havia trânsito de pessoas, pode-se considerar os efeitos do deslocamento Doppler insignificantes. Diante disto, os parâmetros relacionados à dispersão em frequência não serão aqui abordados.

O conhecimento do perfil de potência de retardo de um canal possibilita a extração direta de dois parâmetros, de grande interesse prático. São os dois primeiros momentos de $P_h(\tau)$, identificados como retardo médio e espalhamento de retardo. Eles são expressos pelas Eqs. (2.30) e (2.31):

$$\bar{\tau} = \frac{\int_0^{\infty} \tau P_h(\tau) d\tau}{\int_0^{\infty} P_h(\tau) d\tau}, \quad (2.30)$$

$$\tau_{\text{rms}} = \sqrt{\frac{\int_0^{\infty} (\tau - \bar{\tau})^2 P_h(\tau) d\tau}{\int_0^{\infty} P_h(\tau) d\tau}}. \quad (2.31)$$

A partir da função de correlação em frequência, obtém-se, diretamente, o terceiro parâmetro descritivo da característica de dispersão temporal do canal, denominado banda de coerência. Este parâmetro é definido como o menor valor de Ω para o qual $R_T(\Omega)$ é igual a um coeficiente de correlação desejado, tipicamente 0,5 ou 0,9, e funciona como uma medida da faixa de frequências sobre a qual o canal pode ser considerado plano, ou

seja, a correlação entre as amplitudes das componentes espectrais é elevada. A banda de coerência é útil na avaliação de desempenho e limitação de esquemas de modulação e técnicas de recepção em diversidade.

A comparação da banda de coerência do canal com a faixa que o sinal ocupa permite a classificação do sinal quanto à largura de sua banda. Em um sinal banda larga, a largura de banda do sinal é maior que a largura do canal, no sinal banda estreita, o sinal possui largura de banda menor que a do canal.

A determinação dos parâmetros de dispersão temporal do canal também possibilita o conhecimento do tipo de desvanecimento que afetará o sinal [17]. Sob a análise no domínio do tempo, quando o retardo máximo em excesso característico do canal supera o tempo de símbolo do sinal que trafega neste meio, ocorrerá desvanecimento seletivo em frequência. No caso contrário, o sinal será afetado por desvanecimento plano. No domínio da frequência, quando a banda de coerência do canal for menor que a banda do sinal, ocorrerá desvanecimento seletivo. No outro caso, ocorrerá desvanecimento plano, significando que as componentes espectrais serão afetadas pelo canal de forma semelhante. Ambos os tipos de desvanecimento ocasionam perda de relação sinal-ruído, mas o desvanecimento seletivo é responsável pela introdução de interferência inter-simbólica no sinal.

2.3.3 Caracterização em larga escala

Ao se considerar grandes deslocamentos, as variações do terreno e do ambiente ocasionam a não estacionariedade do canal. Desta forma, não é possível aplicar diretamente as funções de autocorrelação para a análise estatística dos parâmetros do canal. Neste caso, é feito um estudo estatístico dos parâmetros em áreas WSS contíguas e espacialmente homogêneas. Este trabalho não tem como objetivo realizar a caracterização do canal em larga escala, assim, este tema não será abordado.

Capítulo 3

Análise Bayesiana

Neste capítulo, serão apresentados os fundamentos da Estatística Bayesiana e como esta difere da chamada Estatística Clássica. Em seguida, será apresentado o problema da regressão bayesiana e a estrutura genérica dos algoritmos de aprendizagem de máquina. O conhecimento destes temas possibilitará a compreensão da técnica utilizada na estimação do canal sem fio multipercurso.

O problema fundamental da Estatística reside na geração de inferências ou afirmativas sobre características desconhecidas de um sistema físico, a partir de dados gerados por este mesmo sistema e que foram devidamente observados. Entenda-se aqui o conceito de inferência estatística como a análise em termos probabilísticos dos fenômenos de interesse produzidos pelo sistema físico e que serve de suporte para a tomada de ação prática no contexto de uma dada situação.

Antes de prosseguir com a apresentação do paradigma bayesiano, é importante conhecer as bases da corrente clássica, de forma a clarificar a diferenciação que existe entre as duas.

3.1 Estatística Clássica

No quadro clássico, o principal objetivo da inferência estatística é determinar generalizações que podem ser feitas sobre a população a partir da amostra que da mesma foi recolhida. O significado de amostra é sinônimo a observações ou dados estatísticos

resultantes de experiências ou inquéritos repetidos sob as mesmas condições, enquanto população corresponde ao conjunto de todas as observações possíveis de serem feitas nestas condições.

Representa-se a amostra por z , quando representa um escalar, ou $\mathbf{z} = (z_1, z_2, \dots, z_n)$, quando representa uma grandeza vetorial de dimensão n . O conjunto \mathcal{Z} de amostras possíveis, tal que $z \in \mathcal{Z}$, designa-se por espaço-amostra e, usualmente, $\mathcal{Z} \subseteq \mathbb{R}$ para o caso da amostra tratar-se de um escalar ou $\mathcal{Z} \subseteq \mathbb{R}^n$, no caso de tratar-se de um vetor.

Para que sejam estabelecidas inferências na Estatística Clássica, é importante ter em mente que os dados observados formam apenas um dos muitos conjuntos que poderiam ter sido obtidos nas mesmas circunstâncias. Nesta perspectiva, a interpretação dos dados não depende apenas do particular conjunto observado, mas também das hipóteses adotadas acerca dos possíveis conjuntos alternativos de dados. Esta consideração leva a aceitar os dados como observação de uma variável aleatória Z ou $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ com uma função de densidade de probabilidade (f.d.p.) F_0 . No presente trabalho, o termo “distribuição” de uma variável aleatória ou vetor aleatório será utilizado com significado equivalente a função densidade de probabilidade desta variável ou vetor.

Naturalmente, F_0 não é perfeitamente conhecida, mas normalmente existe algum conhecimento inicial sobre a natureza do fenômeno aleatório gerador de dados, o que leva a propor uma família de distribuições \mathcal{F} , a qual pertence F_0 , e que compõe o modelo estatístico para Z . As distribuições que compõem \mathcal{F} são rotuladas por um parâmetro w , com domínio em um conjunto \mathcal{W} , de forma que pode-se escrever a Eq. (3.1):

$$\mathcal{F} = \{f(z|w) : w \in \mathcal{W}\}, \quad z \in \mathcal{Z}. \quad (3.1)$$

Quando as n variáveis aleatórias do vetor $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ são supostas independentes condicionalmente em w , o modelo estatístico pode representar-se em termos

das densidades marginais das variáveis. Podemos reescrever a Eq. (3.1) como a Eq. (3.2):

$$\mathcal{F} = \left\{ f(z|w) = \prod_{i=1}^n f_i(z_i|w) : w \in \mathcal{W} \right\}, \quad z \in \mathcal{Z}. \quad (3.2)$$

Para o caso de as variáveis serem independentes e identicamente distribuídas (i.i.d.), a seguinte simplificação é válida: $f_i(\cdot|w) = f(\cdot|w)$ para $i = 1, 2, \dots, n$. Temos, então, a Eq. (3.3):

$$\mathcal{F} = \left\{ f(z|w) = \prod_{i=1}^n f(z_i|w) : w \in \mathcal{W} \right\}, \quad z \in \mathcal{Z}. \quad (3.3)$$

A escolha da família \mathcal{F} resulta de fatores como conhecimento experimental obtido anteriormente e considerações teóricas sobre os objetivos do estudo e sobre a natureza do fenômeno envolvido.

Após a importante etapa de modelagem e parametrização, a inferência clássica contém uma vasta gama de procedimentos destinados a explorar as conclusões que podem ser extraídas das amostras a respeito do modelo representativo da população. De forma concreta, a inferência clássica procura responder às seguintes questões:

- Os dados z estão de acordo ou são compatíveis com a família \mathcal{F} ?
- Que conclusões podem ser extraídas do valor particular de w_0 que indexa a função de distribuição F_0 que descreve apropriadamente as condições que se observam as variáveis aleatórias?

Os procedimentos de julgamento e interpretação dos resultados na Estatística Clássica é feita à luz da *amostragem repetida*. Segundo este princípio, os métodos estatísticos devem ser verificados através do seu comportamento num número indefinido de repetições realizadas sob as mesmas condições. Ela baseia-se na regularidade estatística das frequências relativas e sustenta que a probabilidade de um dado evento (ou acontecimento) pode ser medida observando a frequência relativa do mesmo acontecimento numa sucessão numerosa de provas ou experiências idênticas e independentes.

A conjectura do verdadeiro valor do parâmetro w é feita através do emprego dos

chamados *estimadores* que correspondem a operações realizadas sobre as observações coletadas $T(Z_1, Z_2, \dots, Z_n)$.

3.2 Estatística Bayesiana

Até agora, a Estatística foi vista sob o ponto de vista clássico, em que probabilidades são vistas em termos de frequências de eventos aleatórios e repetíveis. Concentraremos agora na perspectiva bayesiana, na qual as probabilidades fornecem uma quantificação de incerteza ou grau de credibilidade.

O teorema de Bayes, apresentado adiante, é uma proposição sobre probabilidades condicionadas que são indiscutíveis bastando que sejam aceitas as leis tradicionais do cálculo e a axiomática de Kolmogorov. Partindo-se do espaço de probabilidade (Δ, \mathcal{A}, P) , identifica-se:

- Δ é um espaço que contém os chamados eventos elementares δ , $\delta \in \Delta$ e subconjuntos A , $A \subseteq \Delta$, denominados eventos;
- \mathcal{A} é a família dos eventos dotados de probabilidade não nula;
- P é a medida de probabilidade definida para os eventos $A \subseteq \Delta$, $A \in \mathcal{A}$, em que $P(A)$ é a probabilidade do evento A .

Considere um conjunto de partições de Δ munidas de probabilidade não nula e disjuntos dois a dois:

$$A_1, A_2, \dots, A_m, \quad P(A_i) > 0, \quad A_i \cap A_j = \emptyset, \quad i \neq j, \quad \cup_i A_i = \Delta. \quad (3.4)$$

Dado um outro evento qualquer B , com $P(B) > 0$, que pode ser decomposto na união dos conjuntos disjuntos $B = \cup_i (A_i \cap B)$, pelas regra de aditividade da probabilidade e da definição de probabilidade condicional, tem-se:

$$P(B) = \sum_i P(A_i \cap B) = \sum_i P(B|A_i)P(A_i), \quad (3.5)$$

$$P(A_i \cap B) = P(B|A_i)P(A_i) = P(A_i|B)P(B). \quad (3.6)$$

Resolvendo a Eq. (3.6) em função de $P(A_i|B)$, chega-se ao Teorema de Bayes, dado pela Eq. (3.7):

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_i P(B|A_i)P(A_i)}. \quad (3.7)$$

Os eventos A_i são hipóteses às quais o investigador atribui graus de credibilidade, que correspondem às chamadas distribuições *a priori* $P(A_i)$, $i = 1, 2, \dots, m$, de natureza subjetiva. Depois da informação adicional, que consiste em saber que o acontecimento B se realizou, o investigador revê suas probabilidades *a priori* através da fórmula de Bayes e passa a atribuir aos A_i , $i = 1, 2, \dots, m$ as probabilidades *a posteriori* $P(A_i|B)$, $i = 1, 2, \dots, m$.

O método bayesiano apresenta uma divergência fundamental em relação ao modelo clássico. No modelo clássico, o parâmetro w , $w \in \mathcal{W}$, é um escalar ou vetor desconhecido, mas fixo, ou seja, igual ao valor particular que indexa a distribuição da família \mathcal{F} que descreve apropriadamente o processo aleatório gerador das observações. No modelo bayesiano, o parâmetro w , $w \in \mathcal{W}$, é tomado como uma variável aleatória, escalar ou vetorial, não observável. De forma genérica, a filosofia bayesiana consiste em localizar os parâmetros desconhecidos do sistema, no caso, o parâmetro w , e quantificar a sua incerteza em termos de probabilidade [22].

Assim, para o caso de w ser um parâmetro contínuo, que será o caso de interesse prático, $h(w)$ representa a função densidade de probabilidade *a priori* que transporta consigo toda a informação inicial de caráter subjetivo a respeito do parâmetro não observável.

A família de distribuições \mathcal{F} também é um item integrante do modelo bayesiano. A distinção no emprego deste caso para o caso clássico é que aqui os elementos $f(z|w)$ de \mathcal{F} são supostos ter interpretação subjetiva, assim como as distribuições *a priori*. A diferença básica é que, no caso da $f(z|w)$, sua modelagem dispõe de um conjunto de observações (z_1, z_2, \dots, z_n) .

Supondo-se que na saída do processo aleatório seja observado $Z = z$ e considerando um elemento $f(z|w)$ de \mathcal{F} e a distribuição *a priori* $h(w)$, o Teorema de Bayes modifica-se

da Eq. (3.7) para a forma da Eq. (3.8):

$$h(w|z) = \frac{f(z|w)h(w)}{\int_{\mathcal{W}} f(z|w)h(w) dw}, \quad w \in \mathcal{W}, \quad (3.8)$$

onde $h(w|z)$ é a distribuição *a posteriori* de w depois de conhecer a saída $Z = z$. Assim, após o reconhecimento dos dados z , a informação inicial $h(w)$ sobre o parâmetro não observável modifica-se para $h(w|z)$.

O denominador da Eq. (3.8) é uma distribuição marginal de Z , para qualquer w . Para o caso de uma amostra que seja um vetor aleatório ($Z_1 = z_1, Z_2 = z_2, \dots, Z_n = z_n$) a Eq. (3.8) converte-se na Eq. (3.9):

$$h(w|z_1, z_2, \dots, z_n) = \frac{\prod_i f(z_i|w)h(w)}{\int_{\mathcal{W}} \prod_i f(z_i|w)h(w) dw}, \quad w \in \mathcal{W}. \quad (3.9)$$

Um conceito existente na Estatística Clássica que é de fundamental importância no quadro bayesiano é o princípio da verossimilhança, que diz que apenas o valor observado de z deve ser relevante para a extração de conclusões sobre w [23]. O conceito chave no princípio de verossimilhança é a função verossimilhança. Considere f a função densidade de probabilidade da variável aleatória Z . Supondo que tenha sido observado o evento $A : Z = z$, a função verossimilhança é obtida variando w , como exhibe a Eq. (3.10):

$$L(w) = L(w|A) = f(z|w). \quad (3.10)$$

Esta é uma função com domínio em \mathcal{W} e, para cada $w \in \mathcal{W}$, exprime a verossimilhança ou plausibilidade que lhe é atribuída dado o evento A . A verossimilhança não é uma probabilidade, de forma que não faz sentido realizar uma adição de verossimilhanças. Já a razão de verossimilhanças tem significado na avaliação do peso da plausibilidade de w diante de outro parâmetro, ambos decorrentes da observação $Z = z$.

Retomando a Eq. (3.8), identificamos $f(z|w)$ como a verossimilhança de w e o denominador, a distribuição marginal (ou preditiva) de z . Este é um fator que não depende

de w , de forma que esta equação pode ser reescrita na forma da Eq. (3.11).

$$h(w|z) \propto f(z|w)h(w), \quad w \in \mathcal{W} \quad (3.11)$$

distribuição *a posteriori* \propto verossimilhança \times distribuição *a priori*

Em última análise, a verossimilhança tem o papel, na fórmula do Teorema de Bayes, de transformar o conhecimento inicial sobre w através dos dados z . E, desta forma, a distribuição *a posteriori* transporta toda a informação disponível sobre o parâmetro e todos os procedimentos de inferência são baseados exclusivamente em $h(w|z)$.

Como já mencionado, o objetivo da inferência é realizar afirmativas sobre o parâmetro não observável w . O que difere é como isto é implementado nas correntes clássica e bayesiana. As inferências clássicas são baseadas em probabilidades associadas com as diferentes amostras z que poderiam ocorrer para algum valor fixo, mas desconhecido, de w , e no uso de diferentes estimadores para gerar conclusões a respeito da variável não observável. Já as inferências bayesianas fornecem solução única, baseadas nas probabilidades subjetivas e na credibilidade *a posteriori* associada com diferentes valores do parâmetro w e condicionada pelo particular valor de z observado, satisfazendo o princípio da verossimilhança [24].

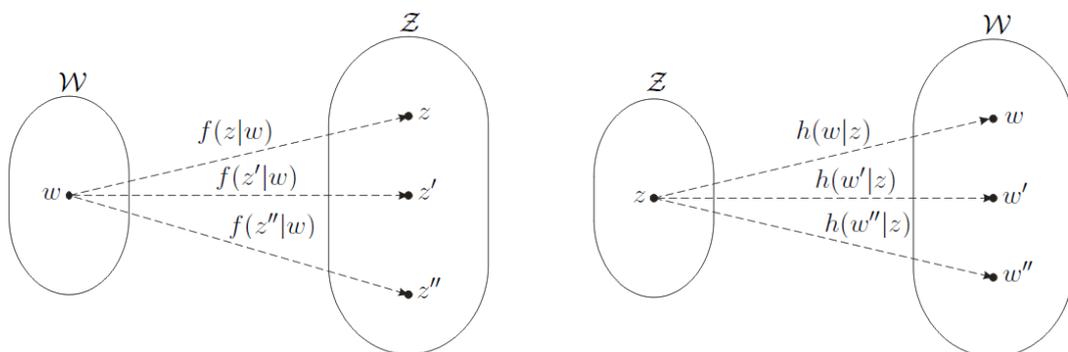


Figura 3.1: Esquema interpretativo das inferências clássica e bayesiana.

A Fig. 3.1 ilustra como ocorre o processo de inferência clássico e bayesiano. Na corrente clássica, a distribuição f , indexada pelo valor do parâmetro desconhecido w , representa a distribuição que descreve de forma apropriada o processo gerador da variável

aleatória z , considerando-se todos os valores possíveis do espaço-amostra \mathcal{Z} . Na corrente bayesiana, dada a observação da saída do sistema físico, representada pela realização da variável aleatória $Z = z$, a distribuição a posteriori $h(w|z)$ porta a informação da incerteza sobre a variável aleatória não observável ω , necessária para a realização de inferências.

3.3 Modelagem Bayesiana

Considere um modelo paramétrico \mathcal{H} constituído de um vetor de parâmetros não observáveis \mathbf{w} e seja \mathbf{z} o vetor representativo da saída observável do experimento. Este modelo é representado por uma distribuição de probabilidade *a priori* $h(\mathbf{w}|\mathcal{H})$ do vetor de parâmetros não observáveis \mathbf{w} ; e pela verossimilhança $f(\mathbf{z}|\mathbf{w}, \mathcal{H})$. O objetivo é inferir a hipótese \mathcal{H} e os parâmetros correspondentes \mathbf{w} que maximizem a distribuição *a posteriori* da Eq. (3.12):

$$\{\hat{\mathbf{w}}, \hat{\mathcal{H}}\} = \max_{\mathbf{w}, \mathcal{H}} \{h(\mathbf{w}, \mathcal{H}|\mathbf{z})\}. \quad (3.12)$$

A solução da Eq. (3.12) pode ser obtida reescrevendo-se a distribuição *a posteriori* e maximizando-se sequencialmente cada um dos termos, como exhibe a Eq. (3.13):

$$h(\mathbf{w}, \mathcal{H}|\mathbf{z}) = h(\mathbf{w}|\mathbf{z}, \mathcal{H})h(\mathcal{H}|\mathbf{z}). \quad (3.13)$$

Esta técnica é chamada método da estimação marginal. Dois níveis de inferência podem ser distinguidos no processo de modelagem de dados. No primeiro nível, assume-se que uma hipótese de modelo particular é verdadeira, e é realizado o ajuste aos dados. Isto é feito através do procedimento de inferência aplicado aos parâmetros não observáveis do modelo, resumidos nos valores de parâmetros mais prováveis e os erros associados a eles. Este processo é repetido para cada hipótese de modelo. O segundo nível de inferência é o teste de comparação de modelos. O objetivo é comparar os modelos e designar algum tipo de preferência ou classificação das alternativas.

1. Ajuste de modelo: nesta fase, assumimos que o modelo \mathcal{H} é verdadeiro, e inferimos quais devem ser os parâmetros \mathbf{w} , à luz dos dados \mathbf{z} . Utilizando o Teorema de Bayes

para obter a distribuição *a posteriori* dos parâmetros \mathbf{w} , escrevemos a Eq. (3.14):

$$h(\mathbf{w}|\mathbf{z}, \mathcal{H}) = \frac{f(\mathbf{z}|\mathbf{w}, \mathcal{H})h(\mathbf{w}|\mathcal{H})}{f(\mathbf{z}|\mathcal{H})}. \quad (3.14)$$

A Eq. (3.14) possui a forma da Eq. (3.8). O termo $f(\mathbf{z}|\mathcal{H})$ funciona como uma constante normalizadora e possui pouca importância nesta fase do processo de inferência. Métodos de otimização baseados no método de gradiente são utilizados para encontrar o estimador máximo *a posteriori* (MAP) da distribuição acima. Esta distribuição fica, portanto, caracterizada pelo valor de \mathbf{w} que maximiza a distribuição *a posteriori* da Eq. (3.14), \mathbf{w}_{MAP} , e de intervalos de confiança associados.

2. Comparação de modelos: na segunda fase da inferência, deseja-se saber qual modelo é o mais plausível à luz dos dados. A distribuição *a posteriori* de cada modelo é dada por:

$$h(\mathcal{H}|\mathbf{z}) \propto f(\mathbf{z}|\mathcal{H})h(\mathcal{H}). \quad (3.15)$$

Identifica-se o termo $f(\mathbf{z}|\mathcal{H})$ como a verossimilhança marginal de \mathcal{H} que, alternativamente, é denominada evidência de \mathcal{H} , termo que será utilizado daqui a diante. O segundo termo, $h(\mathcal{H})$, é a distribuição *a priori* subjetiva sobre a hipótese e expressa quão plausível o modelo é, antes de os dados serem observados.

Consideradas duas hipóteses de modelos \mathcal{H}_1 e \mathcal{H}_2 , relaciona-se a plausibilidade do modelo \mathcal{H}_1 , frente a do modelo \mathcal{H}_2 , da forma expressa na Eq. (3.16):

$$\frac{h(\mathcal{H}_1|\mathbf{z})}{h(\mathcal{H}_2|\mathbf{z})} = \frac{h(\mathcal{H}_1)}{h(\mathcal{H}_2)} \frac{f(\mathbf{z}|\mathcal{H}_1)}{f(\mathbf{z}|\mathcal{H}_2)}. \quad (3.16)$$

A primeira fração do lado direito da equação mede o quanto o conhecimento inicial a respeito do experimento favorece \mathcal{H}_1 sobre \mathcal{H}_2 . A segunda fração, chamada de fator de Bayes, expressa o quão bem os dados observados foram preditos por \mathcal{H}_1 , comparado a \mathcal{H}_2 .

A fase de comparação de modelos incorpora o conceito da navalha de Occam [25]. Considere que o modelo definido pela hipótese \mathcal{H}_1 é mais simples (possui menor quantidade de parâmetros) do que \mathcal{H}_2 . Baseado na motivação de estética do modelo, seria possível

induzir uma tendência favorável ao modelo mais simples manipulando a primeira razão da Eq. (3.16), puramente subjetiva. No entanto, na decisão por não inserir preferência a algum dos modelos, a fração de distribuições *a posteriori* da Eq. (3.16) explica o conceito da navalha de Occam. Modelos mais simples tendem a fazer melhores predições, ainda que para um espaço de dados (E.D.) mais restrito. Verificamos através da Fig. 3.2 que, se \mathcal{H}_2 é um modelo mais complexo, ele irá espalhar sua densidade de probabilidade de predição de forma mais plana e sobre um espaço de dados de z maior, comparado a \mathcal{H}_1 . Assim, para o caso de os dados serem compatíveis com ambas as teorias, o modelo mais simples torna-se-á o mais provável, sem a necessidade de expressar nossa predileção por modelos menos complexos.

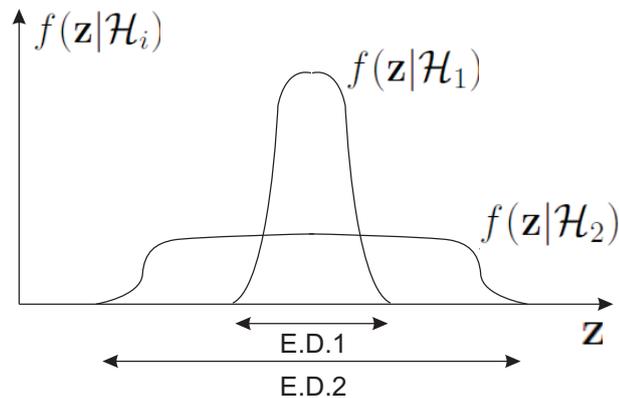


Figura 3.2: Esquema interpretativo do conceito da navalha de Occam na comparação de modelos.

3.4 Regressão Bayesiana e Aprendizagem Supervisionada

Uma das aplicações da modelagem é a interpolação de um conjunto de dados no qual existe ruído inserido [26]. Seguindo a plataforma bayesiana, este constitui um trabalho de inferência em dois níveis: no primeiro nível é encontrado o interpolador de melhor ajuste de cada uma das hipóteses, enquanto no segundo nível de inferência é feita a classificação dos modelos e a decisão pelo que melhor representa o conjunto de dados.

É importante notar que a comparação de modelos não é tarefa simples, visto que

não é possível escolher o modelo que simplesmente encaixa-se melhor ao conjunto de dados [26]. Isto porque modelos complexos sempre podem ter melhor ajuste, assim a escolha pelo modelo de máxima verossimilhança conduziria a uma solução super parametrizada. No entanto, decorrente da Eq. (3.16), os métodos bayesianos incorporam naturalmente a navalha de Occam que penaliza as hipóteses mais complexas.

O problema de interpolação é exposto através de N vetores de dados de entrada, que compõem um conjunto de treinamento e são representados por $\{\mathbf{x}_n\}_{n=1}^N$. Para cada vetor de entrada existe um valor alvo, tal que o conjunto de valores alvo são representados por $\{t_n\}_{n=1}^N$. Através de um processo de treinamento (ou aprendizagem) deseja-se obter um modelo que represente a dependência do valor alvo ao vetor de entrada, com o objetivo de fazer previsões precisas dos valores de t para valores ainda não vistos de \mathbf{x} . O problema de ajuste polinomial de uma curva é um exemplo típico de funções chamadas modelos de regressão [27], em que os dados de entrada são escalares $\{x_n\}_{n=1}^N$, e será utilizado para ilustrar o tema.

Suponha que nos foi dado um conjunto de treinamento composto de N observações de x , juntamente com as correspondentes observações dos valores de t . A Fig. 3.3 exhibe os pares (x_n, t_n) , que foram produzidos por simulação computacional, gerando valores para x espaçados uniformemente no intervalo $[0, 1]$ e os valores alvo obtidos computando os valores correspondentes, através da função $x \mapsto g(x) = \sin(2\pi x)$ e então adicionando um pequeno nível de ruído aleatório ϵ com distribuição gaussiana, como exhibe a Eq. (3.17):

$$t = g(x) + \epsilon. \quad (3.17)$$

O objetivo é explorar o conjunto de treinamento, a fim de fazer previsões sobre o valor de t da variável alvo para novos valores de x , o que, em resumo, consiste em descobrir a função geradora $x \mapsto g(x)$ através de uma função de ajuste $x \mapsto y(x, \mathbf{w})$ que obedeça à Eq. (3.18):

$$t = y(x, \mathbf{w}) + \epsilon, \quad (3.18)$$

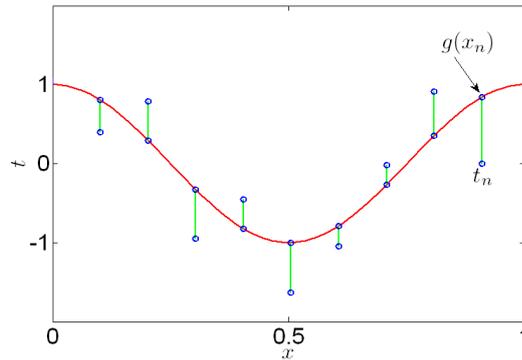


Figura 3.3: Geração do par entrada-alvo do problema de ajuste polinomial.

onde \mathbf{w} representa o vetor com os parâmetros de ajuste de y . De maneira simplificada, pode-se fazer o ajuste dos dados usando a função polinomial da forma:

$$y(x, \mathbf{w}) = w_0 + w_1x + w_2x^2 + \dots + w_Mx^M = \sum_{j=0}^M w_jx^j, \quad (3.19)$$

onde M é a ordem do polinômio e $\mathbf{w} = w_0, w_1, \dots, w_M$. Observando a Fig. 3.4, que exibe os resultados do ajuste polinomial para as ordens $M=0, 1, 3$ e 9 , nota-se que os polinômios de mais baixa ordem fornecem ajuste pobre aos dados, por consequência, representações ruins da função g . O polinômio de terceira ordem parece fornecer o melhor ajuste à função. Já o polinômio de mais alta ordem fornece um ajuste excelente sobre os dados de treinamento, visto que o polinômio passa exatamente sobre os pontos de dados, porém sua curva oscila bruscamente e fornece uma representação pobre de g , o que indica uma pobre *generalização* para novos pontos de dados [27].

O interessante comportamento observado pode ser amenizado utilizando-se um conjunto de dados de entrada com mais elementos. Este é um caso típico de *sobre-ajuste* e pode ser entendido como uma propriedade da verossimilhança máxima. Com a adoção da metodologia bayesiana, o problema do sobre-ajuste pode ser evitado, pois o número de parâmetros adapta-se automaticamente ao número de elementos do conjunto de dados de entrada.

A Eq. (3.19) representa um caso simples de regressão, função linear dos coeficientes \mathbf{w} . É possível estender esta classe de modelos lineares, considerando combinações lineares fixas das variáveis de entrada, como na Eq. (3.20):

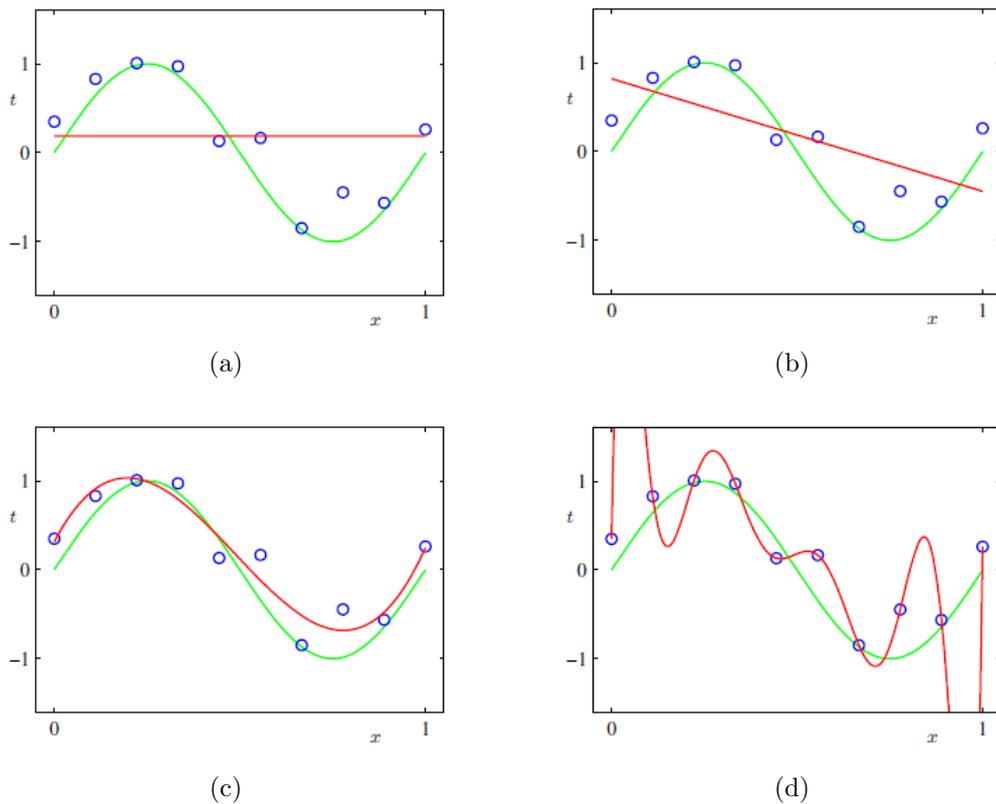


Figura 3.4: Ajuste polinomial. As curvas verdes representam a função geradora g ; os pontos azuis representam os valores alvo t ; as curvas vermelhas representam funções de ajuste, para polinômios de diferentes ordens. (a) $M = 0$, (b) $M = 1$, (c) $M = 3$, (d) $M = 9$.

$$y(x, \mathbf{w}) = w_0 + \sum_{j=1}^{M-1} w_j \phi_j(\mathbf{x}), \quad (3.20)$$

onde $\phi_j(x)$ é conhecida como *função base* [27, 28, 29].

A técnica clássica (não bayesiana) utiliza um estimador para determinar o valor específico do vetor de parâmetros \mathbf{w} . Um dos exemplos mais simples, exibido na Eq. (3.21) é a função *erro da soma dos quadrados*, calculado para cada par (x_n, t_n) dos dados de treinamento.

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N |y(x_n, \mathbf{w}) - t_n|^2. \quad (3.21)$$

A minimização desta função com respeito a \mathbf{w} leva a uma estimativa \mathbf{w}^* que pode ser utilizada para fazer previsões para novos valores de x , avaliando $y(x, \mathbf{w}^*)$.

O método clássico para solucionar o problema de sobre-ajuste exibido na Fig. 3.4 é a regularização, que adiciona um termo de penalidade $\Omega(\mathbf{w})$ à função erro, como expressa

a Eq. (3.22).

$$\tilde{E}(\mathbf{w}) = E(\mathbf{w}) + \lambda\Omega(\mathbf{w}), \quad (3.22)$$

onde o parâmetro Ω , chamado regularizador, desencoraja o sobre-ajuste, penalizando valores altos dos parâmetros peso w_i . Já o parâmetro λ faz o controle da troca entre o ajuste aos dados, reduzindo a função erro, e evitando as variações bruscas de y .

A regularização dos mínimos quadrados pode ser abordada do ponto de vista probabilístico. Os valores alvos observados t_n são assumidos ter sido gerados pela função y adicionados a ruído gaussiano, de forma que na Eq. (3.18) os valores de ruído ϵ_n seguem a uma distribuição normal com média zero e variância σ^2 , como expressa a Eq. (3.23):

$$\begin{aligned} h(\epsilon|\sigma^2) &= \mathcal{N}(\epsilon|0, \beta^{-1}) \\ &= \left(\frac{\beta}{2\pi}\right)^{1/2} \exp\left\{-\frac{\beta}{2}\epsilon^2\right\}, \end{aligned} \quad (3.23)$$

onde a notação $\mathcal{N}(\epsilon|\mu, \beta^{-1})$ especifica uma distribuição gaussiana sobre ϵ com média μ e precisão (inverso da variância) β . Variáveis como μ e β são chamadas *hiperparâmetros*, já que elas controlam a distribuição dos parâmetros.

Também podemos expressar nossa incerteza sobre o valor da variável alvo utilizando uma distribuição de probabilidade. Com este propósito, assumimos que, dado um valor para x , a f.d.p. que representa a variável alvo t correspondente será uma distribuição gaussiana com valor esperado igual à função geradora y , calculada no ponto x , como na Eq. (3.24):

$$f(t|x, \mathbf{w}, \beta) = \mathcal{N}(t|y(x, \mathbf{w}), \beta^{-1}). \quad (3.24)$$

Como não é nossa intenção modelar os dados de entrada, x , podemos, a título de simplificação da notação, omitir esta variável, reescrevendo $f(t|x, \mathbf{w}, \beta)$ como $f(t|\mathbf{w}, \beta)$. Seu esquema interpretativo está representado na Fig. 3.5.

Agora, os dados de treinamento $\{x, t\}_{n=1}^N$ são utilizados para determinar os valores dos parâmetros desconhecidos \mathbf{w} e β por máxima verossimilhança. Se os dados são

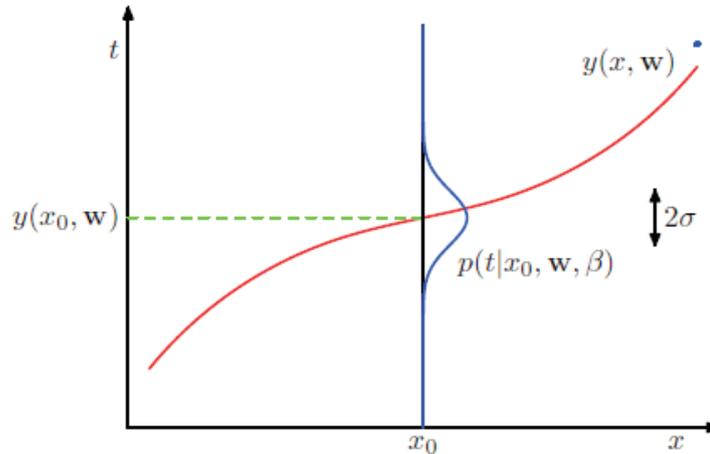


Figura 3.5: Esquemático ilustrativo da distribuição condicional gaussiana para t dado x .

assumidos i.i.d., a função verossimilhança é dada pela Eq. (3.25).

$$f(\mathbf{t}|\mathbf{w}, \beta) = \prod_{n=1}^N \mathcal{N}(t_n | y(x_n, \mathbf{w}), \beta^{-1}). \quad (3.25)$$

A maximização da expressão acima para os coeficientes do polinômio fornecerá a solução de máxima verossimilhança, denotada por \mathbf{w}_{ML} . Verifica-se que a maximização da verossimilhança é equivalente, na determinação de \mathbf{w} , à minimização da função erro da soma dos quadrados definida na Eq. (3.21) [29]. Este seria o estimador pontual utilizado pela Estatística Clássica para apontar o valor mais adequado de \mathbf{w} . Também pode-se utilizar a máxima verossimilhança para determinar o parâmetro de precisão β_{ML} .

De acordo com o tratamento bayesiano, a caracterização da incerteza sobre \mathbf{w} deve ser expressa através de uma distribuição de probabilidade $h(\mathbf{w})$. A definição de uma distribuição *a priori* que expressa nossa incerteza sobre \mathbf{w} levando consigo toda a informação, excluindo os dados em si, sem perda de generalidade, pode ser expressa na forma dada pela Eq. (3.26):

$$h(\mathbf{w}|\alpha) \propto \exp \{ -\alpha \Omega(\mathbf{w}) \}, \quad (3.26)$$

onde α pode ser tratado como um hiperparâmetro e Ω , o regularizador, cuja escolha

habitual, expressa pela Eq. (3.27), será adotada [27]:

$$\Omega(\mathbf{w}) = \frac{\|\mathbf{w}\|^2}{2}. \quad (3.27)$$

Uma escolha possível para a distribuição gaussiana para $h(\mathbf{w}|\alpha)$ tem a forma da Eq. (3.28):

$$h(\mathbf{w}|\alpha) = \left(\frac{\alpha}{2\pi}\right)^{M/2} \exp\left\{-\frac{\alpha}{2}\|\mathbf{w}\|^2\right\}. \quad (3.28)$$

Ela expressa a preferência por modelos mais suaves, através da declaração de que pesos menores são valores mais prováveis nesta distribuição *a priori*.

Agora o Teorema de Bayes pode ser utilizado para expressar a distribuição *a posteriori* para \mathbf{w} , como o produto da distribuição *a priori* (Eq. (3.28)) e a função verossimilhança (Eq. (3.25)):

$$h(\mathbf{w}|\mathbf{t}, \alpha, \beta) \propto h(\mathbf{w}|\alpha)f(\mathbf{t}|\mathbf{w}, \beta). \quad (3.29)$$

Esta expressão é análoga à apresentada na Eq. (3.14), referente à fase de ajuste de modelo. Aqui, em vez de aprender um valor único para \mathbf{w} , pode-se inferir uma distribuição sobre todos os valores possíveis. Com efeito, o conhecimento *a priori* sobre os valores dos parâmetros foi atualizado à luz da informação provida pelos dados \mathbf{t} .

O tratamento bayesiano na realização de predições pode ser feito por duas maneiras distintas. A primeira utilizando uma estimativa para \mathbf{w} como o valor mais provável para a distribuição *a posteriori* $h(\mathbf{w}|\mathbf{t}, \alpha, \beta)$. Com base nesta estimativa, é possível fazer predições utilizando $f(t_*|\mathbf{w}_{MAP}, \beta)$, onde t_* representa o valor alvo t correspondente para valores ainda não vistos de x . Pelo segundo método, a predição é feita integrando a distribuição *a posteriori* sobre o espaço de parâmetro \mathbf{w} , em vez de utilizar um valor estimado específico de \mathbf{w} , como exige a Eq. (3.30):

$$h(t_*|\mathbf{t}, \alpha, \beta) = \int_{\mathcal{W}} h(\mathbf{w}|\mathbf{t}, \alpha, \beta)f(t_*|\mathbf{w}, \beta) d\mathbf{w}. \quad (3.30)$$

Esta corresponde à distribuição preditiva de t , obtida pela marginalização sobre \mathbf{w} . Tal integração pode ser de difícil solução analítica e requerer métodos sofisticados para

sua resolução. Por outro lado, o tratamento bayesiano elimina a questão do sobre-ajuste e normalmente resulta em capacidade preditiva aprimorada [28].

Na metodologia bayesiana, em modelos de regressão linear como o apresentado acima, o conjunto de dados de treinamento é utilizado para se determinar a distribuição *a posteriori* do vetor de parâmetros e, após esta fase, os dados de treinamento são descartados e as predições das novas entradas são baseadas puramente no vetor de parâmetros aprendido \mathbf{w} . No entanto, existe uma classe de técnicas de reconhecimento de padrões utilizado em problemas de regressão linear que mantém os pontos dos dados de treinamento e os utiliza na fase de predição. Ela é chamada de método baseado em memória [27] e é considerada mais rápida na fase de treinamento, ainda que seja mais lenta na realização de predições. Nesta abordagem, os modelos paramétricos lineares fazem uso das chamadas *funções núcleo*, que têm a forma da Eq. (3.31):

$$k(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x})^T \phi(\mathbf{x}'). \quad (3.31)$$

Muitas das funções núcleo possuem a propriedade de ser função apenas da diferença entre os argumentos, como é o caso no assunto abordado na presente dissertação, tal que $k(\mathbf{x}, \mathbf{x}') = k(\mathbf{x} - \mathbf{x}')$, que são chamados de núcleos estacionários, visto que são invariantes a translações ao espaço dos dados de entrada.

Dentre as técnicas para o tratamento de problemas de regressão que se utilizam das funções núcleo $k(\mathbf{x}, \mathbf{x}')$, concentramos nosso estudo na chamada na Máquina de Vetor de Relevância (*Relevance Vector Machine*, RVM), introduzida por [30]. Esta se constitui em uma das técnicas das chamadas máquinas de núcleo esparsas, pelo fato de que suas inferências resultam em poucos parâmetros w_i não nulos. A maioria dos parâmetros são automaticamente fixados em zero durante o processo de aprendizagem, gerando um processo que é eficaz em discernir as funções núcleo que são relevantes para fazer boas predições, além de bom desempenho de generalização [30, 31]. O RVM faz predições

baseadas na função dada pela Eq. (3.32):

$$y(\mathbf{x}, \mathbf{w}) = w_0 + \sum_{n=1}^N w_n K(\mathbf{x}, \mathbf{x}_n), \quad (3.32)$$

e gera saídas probabilísticas baseadas nas distribuições *a posteriori*. Uma de suas modificações principais é a distribuição *a priori* sobre parâmetros. Diferentemente da Eq. (3.28), agora é introduzido um hiperparâmetro α_i separado para cada parâmetro w_i , em vez de um hiperparâmetro compartilhado [27]. Assim, a distribuição *a priori* dos parâmetros peso tem a forma da Eq. (3.33):

$$h(\mathbf{w}|\boldsymbol{\alpha}) = \prod_{i=1}^M \mathcal{N}(w_i|0, \alpha_i^{-1}), \quad (3.33)$$

onde α_i representa a precisão do parâmetro correspondente w_i , e $\boldsymbol{\alpha}$ denota $(\alpha_1, \dots, \alpha_M)$. Veremos que, com a maximização da evidência com respeito aos hiperparâmetros, uma parte significativa deles irá para infinito, e os parâmetros peso correspondentes terão distribuições *a posteriori* concentrados em zero. As funções base associadas a estes parâmetros não terão papel na predição e serão excluídos, dando origem à característica esparsa do modelo.

Capítulo 4

Técnicas de sondagem

A disseminação dos sistemas de comunicação sem-fio de alta capacidade torna a caracterização adequada do canal de propagação faixa larga de estratégica importância. O conhecimento preciso da resposta do canal ao impulso proporciona o conhecimento de parâmetros que podem ser levados em consideração na implementação de equalizadores adaptativos ou seletores de frequência otimizados, que minimizem os efeitos negativos do canal multipercurso [32]. A caracterização dos modelos de canal faixa larga pode ser feita através de diferentes abordagens. A modelagem determinística é baseada na simulação eletromagnética, através da resposta ao impulso derivada de um ambiente simplificado ou de traçado de raios. Este método é limitado pelos recursos computacionais e pelas simplificações necessárias ao ambiente de propagação. Estes fatores tornam atrativo o uso da abordagem empírica, baseada na análise estatística de um grande número de medidas obtidas através da sondagem do canal. Assim, os métodos de sondagem faixa-larga, incluindo as técnicas de sondagem no domínio do tempo e no domínio da frequência, constituem os meios empíricos capazes de fornecer subsídios para a obtenção dos parâmetros descritivos do canal rádio. Sondas de canal são usadas para se obter dados experimentais a partir dos quais estes parâmetros podem ser extraídos. O resultado de cada uma destas técnicas gera uma função que representa a característica dispersiva do canal. As funções obtidas por um grupo ou outro estão relacionadas pela transformada de Fourier.

4.1 Técnica de sondagem no domínio da frequência

O princípio da técnica de caracterização do canal faixa-larga na frequência é baseado na emissão de uma série de ondas contínuas, idealmente compostas por apenas uma componente espectral, sequencialmente, com pequena separação de frequência entre elas, de forma a varrer toda a banda desejada. O sinal recebido é registrado em cada uma destas frequências. Este método, ilustrado na Fig. 4.1, é bastante utilizado em ambientes fechados e é realizado através do uso do Analisador Vetorial de Rede. Com ele, são medidas as informações de fase e amplitude em cada ponto de frequência discreta, dentro da banda utilizada na sondagem. A matriz dos valores obtidos, para cada frequência, dá origem à função de transferência variável no tempo, $T(f, t)$, representativa do canal [33]. Diversos autores já utilizaram esta técnica de sondagem [34, 35, 33, 36].

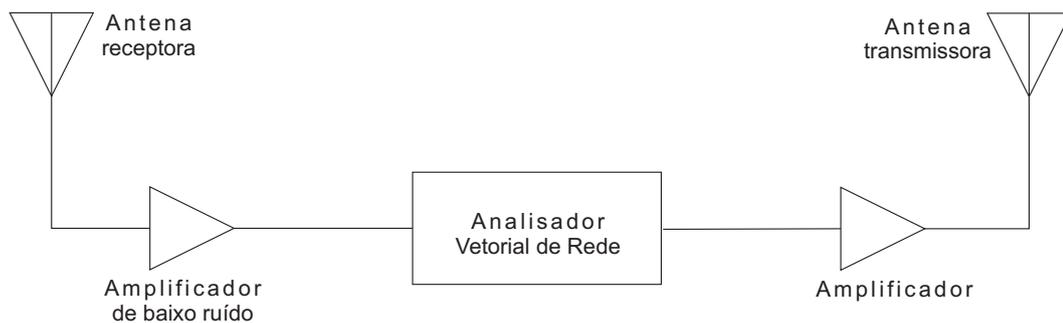


Figura 4.1: Sondagem no domínio da frequência através do analisador de rede vetorial.

As principais limitações desta técnica são: 1) a dificuldade de realização de medição em grandes distâncias, visto que as antenas transmissora e receptora devem estar conectadas ao analisador de rede; 2) o tempo de varredura da banda considerada influencia na medição, sendo necessário considerar o canal estacionário ao longo da varredura.

4.2 Técnica de sondagem no domínio do tempo

As limitações apresentadas na sondagem no domínio da frequência, acima descritas, motivam o desenvolvimento de outras técnicas, como as sondagens no domínio do tempo, que utilizam a transmissão de um sinal que, efetivamente, ocupa uma larga banda de frequências. Este grupo apresenta como benefício principal a obtenção de forma direta do

perfil de retardos, entidade que carrega importantes informações do desvanecimento em pequena escala do canal.

4.2.1 Sondagem por pulsos periódicos

A forma mais simples de sondagem no domínio do tempo consiste na transmissão de pulsos regularmente espaçados. Estes pulsos devem ter duração muito curta, de forma que se assemelhem a impulsos. O sinal, coletado por um osciloscópio de armazenamento digital, equivale à convolução do pulso de sondagem pela resposta impulsiva do canal [37, 38]. A duração do impulso indica a menor diferença de retardos que duas componentes de multipercursos, que chegam à antena receptora, podem ter, de forma a serem identificadas como ecos distintos. A periodicidade dos pulsos deve possuir um valor mínimo que garanta que as contribuições de multipercurso tenham chegado ao receptor antes da transmissão do próximo pulso. Este método, empregado por [40], apresenta desvantagens como a dificuldade de se gerar, no transmissor, pulsos de duração curtíssima e com significativa relação potência de pico-potência média que possibilitem a detecção de ecos mais fracos e também a grande largura de faixa do filtro passa-banda que torna a técnica sujeita à interferência e ao ruído [33, 41].

4.2.2 Sondagem por compressão de pulsos

A técnica de sondagem por compressão de pulsos visa contornar as deficiências existentes no método anterior e é baseada na teoria de sistemas lineares. Considere o canal um sistema linear $h(t)$ ao qual é aplicado, em sua entrada, ruído branco $n(t)$. A saída $y(t)$ do canal, na entrada do receptor, é dada pela convolução apresentada na Eq. (4.1):

$$y(t) = \int h(\xi)n(t - \xi) d\xi. \quad (4.1)$$

Se realizarmos a correlação cruzada da saída com uma versão atrasada do sinal de

entrada, obteremos:

$$\begin{aligned}
 E[y(t)n^*(t-\tau)] &= E \left[\int h(\xi)n(t-\xi)n^*(t-\tau) d\xi \right] \\
 &= \int h(\xi)R_n(\tau-\xi) d\xi \\
 &= N_0h(\tau).
 \end{aligned} \tag{4.2}$$

O seguinte resultado indica que, ao aplicarmos ruído branco à entrada do sistema, obteremos, após o correlator cruzado, uma saída proporcional à resposta impulsiva do sistema, avaliada no instante do retardo. Com a impossibilidade de se gerar um sinal genuinamente branco, utiliza-se uma forma de onda determinística com características de ruído e cujas propriedades são bem conhecidas. Trata-se das sequências binárias pseudo-aleatórias (sequências PN) [42]. Sua simples implementação, através da utilização de registradores de deslocamento, suas conhecidas características de autocorrelação [43] e a possibilidade de transmissão do sinal a potências mais elevadas tornam esta abordagem bastante atraente. Duas técnicas vêm sendo largamente realizadas, nas medições em faixa larga, com a utilização de sequências PN. São estas a Varredura de Retardo de Tempo e Correlação Cruzada e a Convolução com Filtro Casado.

Varredura de retardo de tempo e correlação cruzada

O projeto de uma sonda de varredura de retardo de tempo e correlação cruzada, cuja estrutura está esquematizada na Fig. 4.2, dá-se através da implementação de um sistema transmissor que contém um gerador de sequência pseudo-aleatória cuja duração de bit é T_c . Este sinal, multiplicado pelo sinal de uma portadora centrada na frequência f_c , gerará um sinal modulado a ser transmitido pela antena.

O bloco receptor conta com um filtro banda larga, que recebe o sinal que chega à antena. O sinal faixa larga recebido é multiplicado por uma sequência pseudo-aleatória semelhante à da transmissão, porém com frequência de bit ligeiramente diferente. Na saída do multiplicador, posiciona-se um integrador, de forma que o conjunto multiplicador-integrador componha um correlator. Neste ponto, o sinal de saída contém o necessário

para a determinação da forma da resposta ao impulso do canal. A diferença de frequências entre os osciladores caracteriza o chamado correlator deslizante, descrito por [44]. Este único correlator substitui o processo de correlação em tempo real, que requereria um banco de correladores com diferenças de retardo infinitesimais entre eles.

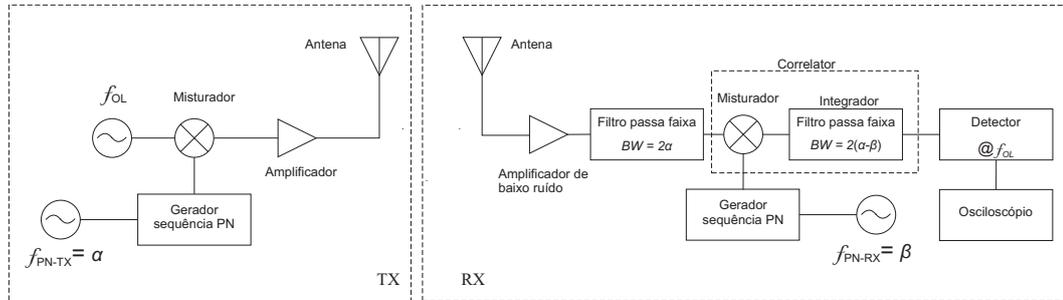


Figura 4.2: Técnica de sondagem por varredura de retardo de tempo e correlação cruzada.

Esta técnica, implementada por diversos autores [45, 46, 47, 48, 49], apresenta como vantagem sobre as demais técnicas de sondagem no domínio do tempo a eliminação da necessidade por rápida amostragem e processamento digital, no caso de medições com o sistema receptor em movimento [50].

Deficiências intrínsecas desta estratégia de sondagem que derivam da diferença de tempo entre as sequências PN são notáveis, como: 1) o fato de as medidas obtidas serem dadas não em tempo real, mas escaladas pelo fator de deslizamento das sequências; 2) a demora na aquisição dos perfis devido ao tempo decorrido até a coincidência das sequências e a necessidade de se considerar o canal estacionário durante este intervalo, o que nem sempre é verdadeiro [50].

Convolução por filtro casado

Outra técnica popular que utiliza o conceito de compressão de pulsos é a técnica de sondagem com um filtro casado à sequência transmitida. Diferentemente da abordagem do item anterior, agora o sistema receptor produz uma sequência idêntica à do bloco transmissor para fazer a operação de convolução. A estrutura básica deste sistema de sondagem é exibida na Fig. 4.3.

Considere $s(t)$ o sinal transmitido e $y(t)$ a representação da envoltória complexa do sinal banda passante que chega à antena receptora, considerando a resposta impulsiva

do canal igual a modelada por Turin [4]. A operação desempenhada pelo filtro casado irá produzir o sinal $z(t)$, seguindo a operação descrita abaixo:

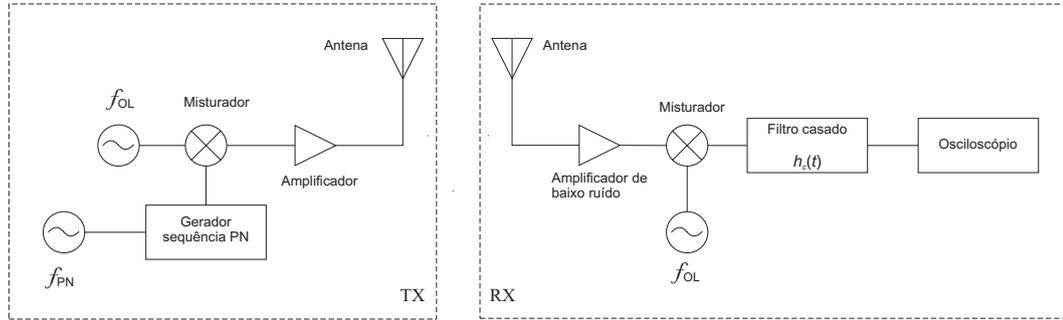


Figura 4.3: Técnica de sondagem por convolução por filtro casado.

$$y(t) = \sum_{l=1}^L w_l s(t - \tau_l). \quad (4.3)$$

Se o filtro casado tem resposta ao impulso $h_c(t) = s^*(-t)$, teremos a saída $z(t)$ com a forma dada pela Eq. (4.4):

$$\begin{aligned} z(t) &= h_c(t) * y(t) \\ &= h_c(t) * \sum_{l=1}^L w_l s(t - \tau_l) \\ &= s^*(-t) * \sum_{l=1}^L w_l s(t - \tau_l) \\ &= \sum_{l=1}^L w_l \int_{-\infty}^{\infty} s(-\tau) s(t - \tau - \tau_l) d\tau \\ &= \sum_{l=1}^L w_l R_s(t - \tau_l). \end{aligned} \quad (4.4)$$

Se o sinal $s(t)$ for apropriadamente escolhido, de tal forma que sua função autocorrelação seja semelhante à função impulso, como é o caso das sequências pseudo aleatórias, a saída do filtro casado gerará uma curva semelhante à forma da resposta impulsiva do canal de radiopropagação, descrita na modelagem de um canal multipercurso estabelecida por Turin.

Uma das formas usualmente empregadas para implementar o filtro casado é através

de linhas de retardo de onda acústica de superfície (em inglês, *surface acoustic wave* - SAW) [18, 51]. Esta espécie de dispositivo, entretanto, impõe algumas dificuldades sobre o processo de sondagem, pois é difícil obter longas sequências e ocorre a geração de sinais acústicos espúrios que dão origem a fenômenos de reflexão e espalhamento, que culminarão na criação de lóbulos laterais na saída do filtro casado e resultarão em redução da faixa dinâmica do sistema [18, 50].

A abordagem alternativa, que substitui o uso da linha de retardo SAW, é a realização do filtro casado em *software*, isento das não linearidades típicas dos dispositivos analógicos. As vantagens desta técnica a tornam mais atrativa, e, por isso, já foi implementada em vários trabalhos [46, 52, 53, 54].

O projeto da sequência pseudo-aleatória transmitida irá determinar importantes questões de desempenho da sonda. A taxa de bits está diretamente relacionada à diferença mínima de tempo de chegada entre duas versões do sinal vindas de diferentes espalhadores. Também, o número de bits da sequência, que estabelece a duração do sinal transmitido até uma nova repetição, define o eco com máximo retardo que pode ser detectado, sem ambigüidade, pelo sistema. Estas restrições de tempo mínimo e máximo de chegada entre as componentes de multipercurso, além da sensibilidade reduzida dos osciloscópios, são as principais restrições desta técnica. Por outro lado, graças aos avanços dos mecanismos de aquisição e armazenamento de dados, a necessidade de altas taxas de amostragem não figura como grande empecilho às sondas atuais.

Dentre todos os métodos exibidos, a técnica de filtro casado com correlação realizada em *software* foi a escolhida para ser empregada no presente trabalho, visto as vantagens que esta possui sobre as demais. Além disto, com este método é possível obter os dados necessários para aplicação da técnica de estimação de parâmetros a ser utilizada.

Capítulo 5

Sistema de sondagem e cenário das medições

Este capítulo trata dos assuntos que dizem respeito às medições realizadas para obtenção da resposta ao impulso (RIC) do canal de radiopropagação *indoor*. O sistema de sondagem faixa larga desenvolvido será apresentado e seus parâmetros de projeto, relatados. Serão apresentados, também, os ambientes onde foram realizadas as medições e a metodologia adotada.

5.1 Projeto da sonda faixa larga

Tendo em vista as vantagens descritas no capítulo 4 do sistema de sondagem no domínio do tempo, foi projetada uma sonda por compressão de pulsos, na faixa de frequências entre 1870 a 2030 MHz, que utiliza convolução por filtro casado para que fosse obtida a resposta impulsiva do canal. Esta faixa foi escolhida devido à limitação dos equipamentos e dispositivos utilizados.

O sistema pode ser dividido em: bloco transmissor, que realiza a transmissão do sinal de sondagem, e bloco receptor, que demodula e coleta amostras do sinal propagado no ambiente. Pode-se visualizar o esquema ilustrativo da sonda na Fig. 5.1. Ambos os blocos serão apresentados a seguir.

O bloco transmissor é composto por um equipamento gerador/modulador, que

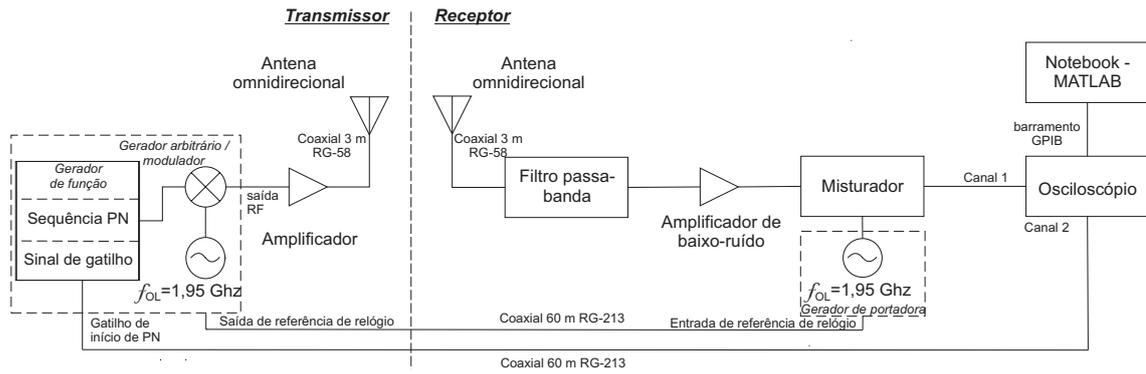


Figura 5.1: Sistema empregado na sondagem do canal sem fio multipercurso.

produzirá o sinal em RF para ser amplificado e transmitido através da antena para o ambiente, como ilustra o lado esquerdo da Fig. 5.1.

O primeiro elemento componente do bloco transmissor é o gerador vetorial de forma de onda *Anritsu*, modelo MG3700A, que funciona como um gerador de sinais arbitrários modulando uma portadora em RF. Ele foi utilizado na geração da sequência pseudo-aleatória (PN) transmitida na sondagem do canal. A sequência PN utilizada foi elaborada em MATLAB[®] e, então, transferida para a memória interna do equipamento.

As características da sequência utilizada são determinantes nas especificações de detecção de multipercursos da sonda. A Fig. 5.2 representa uma sequência pseudo-aleatória típica, gerada a partir de um polinômio primitivo de ordem $m = 8$ e que pode ser representado por um circuito de registradores de deslocamento com realimentação [42]. A sequência ilustrada apresenta 2 níveis (1, -1), composta por $M = 255$ bits de duração $T_b = 12,5$ ns e que se repete com período $T = 3,1875$ μ s.

O primeiro parâmetro relevante no projeto da PN é a duração de bit. O valor escolhido irá determinar, primeiramente, a largura de faixa a ser sondada, visto que a densidade espectral de potência de um sinal com modulação BPSK tem a seguinte forma [55]:

$$S_{BPSK}(f) = 2E_b \text{sinc}^2(T_b f), \quad (5.1)$$

onde E_b representa a energia de bit do sinal, e T_b , a duração do bit. A expressão acima mostra que a maior parte da energia do sinal está concentrada sob o lóbulo principal da função sinc^2 e que a largura deste lóbulo é inversamente proporcional ao valor de T_b .

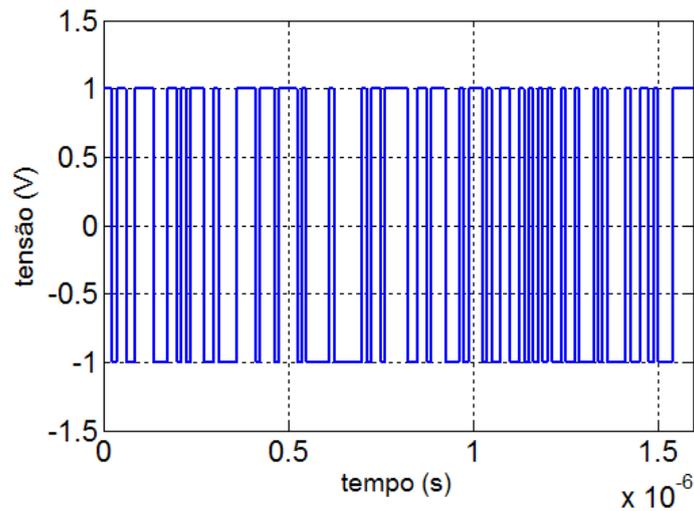


Figura 5.2: Trecho de uma sequência pseudo-aleatória típica.

A duração de bit também irá determinar a menor diferença de tempo de chegada entre dois ecos que podem ser distinguidos pela sonda (resolução temporal). Isto se dá pelo fato do pico da função autocorrelação de uma sequência pseudo-aleatória, livre de ruídos, apresentar formato triangular de largura $2T_b$. Assim, se um eco B chega ao receptor T_b segundos após o eco A, o seu pico de correlação ocorrerá sobre o valor de mínimo da correlação devido ao eco A e a interferência será mínima.

O segundo parâmetro é o comprimento da sequência pseudo aleatória, M . A quantidade de bits que compõem a PN está associada a sua duração, $T_{PN} = MT_b$. Para um sistema que transmite ciclicamente o mesmo sinal, se uma componente de multipercurso chegar com atraso superior a T_{PN} , não será possível identificar se este eco possui retardo τ ou $T_{PN} + \tau$. O comprimento da PN também irá determinar a faixa dinâmica do sistema, pois observa-se que o valor do pico de autocorrelação da PN é proporcional ao número de bits que a compõe. A faixa dinâmica do sistema determina a maior diferença de potências entre multipercursos que pode ser detectada pela sonda, de acordo com a Eq. (5.2):

$$FD = 20 \log(m) \quad (\text{dB}). \quad (5.2)$$

No projeto da sequência pseudo-aleatória, outro fator que deve ser levado em consideração é o número de amostras presentes em cada bit. Quanto maior este valor, mais os pulsos transmitidos se aproximam da forma retangular e, no domínio da frequência, mais

a densidade espectral aproxima-se da equação (5.1) e a relação entre os lóbulos principal e secundários torna-se maior.

A sequência elaborada apresenta os seguintes valores: $T_b = 12,5$ ns e $M = 255$ bits. Estes valores levam às seguintes especificações da sonda construída:

- largura da faixa de frequência de sondagem: 160 MHz de faixa (1870 a 2030 MHz);
- retardo mínimo: 12,5 ns (equivalente à resolução espacial de 3,75 m);
- retardo máximo: 3,1875 μ s (equivalente a 956,25 m);
- faixa dinâmica: 48 dB.

Entende-se por resolução espacial da sonda a diferença de percurso mínima realizada por dois ecos que, por consequência, chegam ao receptor em instantes distintos e que podem ser eficazmente distinguidos pelo sistema de sondagem. Relaciona-se diretamente com a resolução temporal através da velocidade de propagação da luz no ar.

A escolha da taxa da sequência levou em consideração o compromisso entre o número de amostras por bit e a resolução de multipercursos da sonda. Tendo em vista que o gerador arbitrário possui frequência de amostragem máxima de 160×10^6 amostras por segundo na produção do sinal, foi necessário realizar um estudo da sequência mais vantajosa a ser empregada na sonda. O teste consistiu em, tendo os sistemas transmissor e receptor ligados por cabo coaxial, para cada configuração de sinal projetado e carregado no equipamento transmissor, fazer a coleta do sinal por um *notebook*. Em seguida, era realizada em *software*, a autocorrelação das sequências. O sinal com taxa igual a 80 Mbits/s e 2 amostras/bit foi considerado o mais vantajoso, visto que fornecia boa resolução de multipercursos e obedecia às propriedades de autocorrelação típicas das sequências pseudo-aleatórias.

Os gráficos da Fig. 5.3 exibem a autocorrelação da sequência PN escolhida e a densidade espectral de potência da mesma, calculados a partir do sinal coletado no experimento. A densidade espectral de potência apresenta os lóbulos secundários atenuados, devido à filtragem que ocorre no equipamento modulador/transmissor. Trata-se de um filtro passa baixa com frequência de corte de 100 MHz.

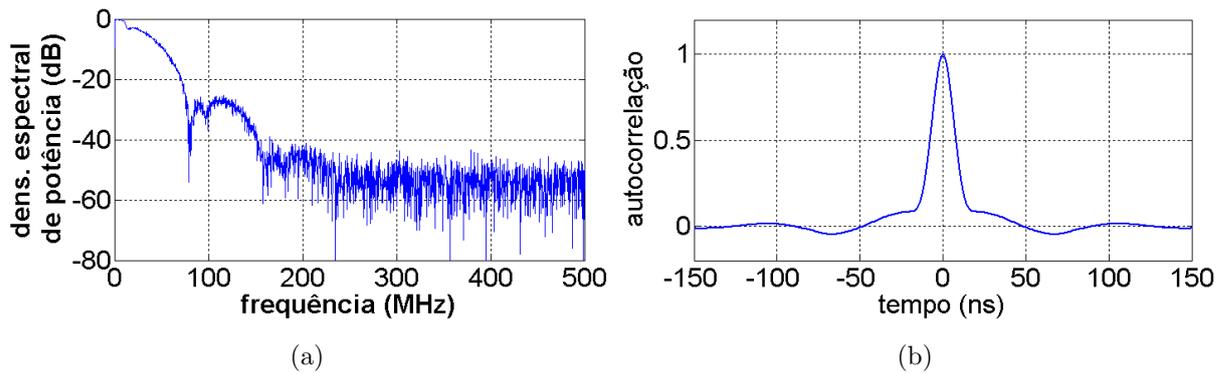


Figura 5.3: Gráficos da (a) densidade espectral de potência e (b) autocorrelação da sequência PN transmitida.

O sinal, já modulado, que deixa o gerador arbitrário, precisa ser amplificado para que possa se propagar pelo ambiente e chegar à antena receptora com potência suficiente para compensar a baixa sensibilidade do osciloscópio. Assim, foi ligado à saída do gerador um amplificador de potência *Comtech*, modelo AR8829-20, que fornece ao sinal de entrada uma amplificação de, aproximadamente, 44 dB na faixa de frequências em torno de 1,95 GHz. Foi realizado o levantamento da curva do ganho do amplificador, para que fosse garantido que a potência de -10 dBm na entrada do equipamento pertencesse à região linear, como é verificado na Fig. 5.4.

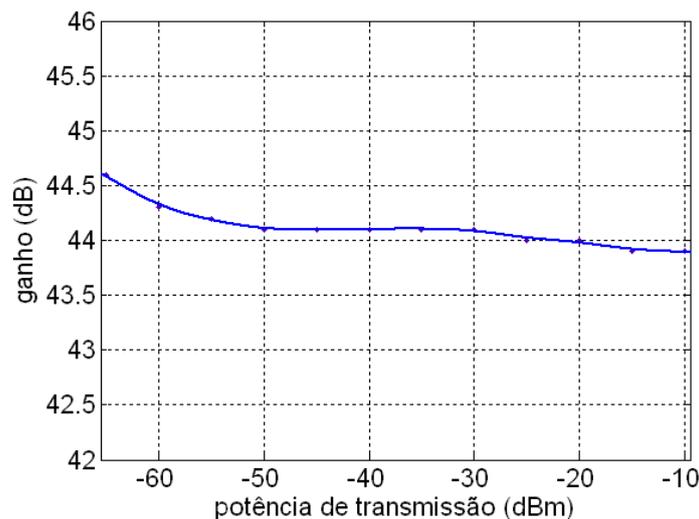


Figura 5.4: Levantamento da curva de linearidade do amplificador de potência para a frequência de 1,95 GHz.

O bloco transmissor encerra-se com uma antena monopolo *Antenex*, modelo EXCN 1920SM. Ela apresenta diagrama omnidirecional no plano horizontal, o que garante que

o sinal seja distribuído igualmente em todas as direções, fazendo com que espalhadores localizados em diferentes posições no ambiente influenciem nas medições. A polarização utilizada é a vertical e o ganho é de 2 dBi.

O bloco receptor é responsável por capturar o sinal de RF propagado no ambiente, realizando a conversão de frequência para a banda base, antes de armazenar a forma de onda para pós-processamento. O esquemático do receptor é exibido no lado direito da Fig. 5.1.

Este bloco tem início na antena receptora. Ela é idêntica à utilizada na transmissão, com diagrama horizontal omnidirecional. Foram realizadas medições nestas antenas, com o objetivo de verificar se estas apresentavam desempenho satisfatório para a faixa de frequência de operação da sonda. Para tal, mediu-se a perda de retorno através de um analisador de rede vetorial *Anritsu* MS2034A. Como indicado no gráfico da Fig. 5.5, foram encontrados valores médios de perda de retorno inferiores a 15 dB, o que garante mais de 97% da energia disponível entregue pela antena.

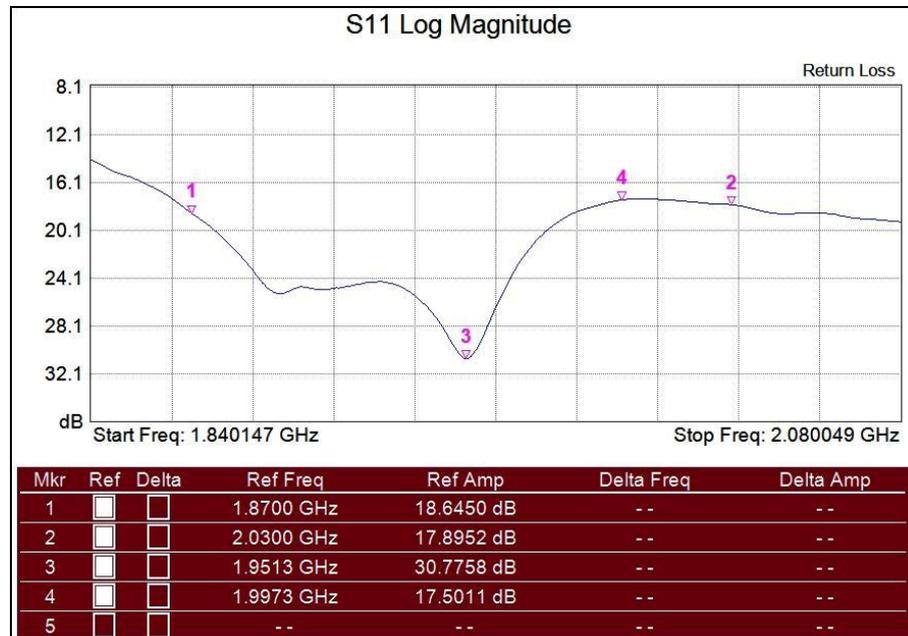


Figura 5.5: Medição da perda de retorno da antena utilizada.

Conectado à antena, foi utilizado um amplificador de baixo ruído *Mini-Circuits*, modelo ZEL-1724LN, capaz de fornecer ganho aproximado de 22,5 dB na faixa de 1700 a 2400 MHz. O funcionamento deste amplificador exigiu o uso de uma fonte de tensão

regulada *Minipa*, modelo MPL-3303, ajustada em 15 V.

Com o objetivo de eliminar as interferências de outras fontes eletromagnéticas, a saída do amplificador foi diretamente conectada a um filtro passa-faixa *Mini-Circuits*, modelo VBFZ-2000, que insere perdas inferiores a 2,3 dB somente às componentes espectrais em sua faixa passante de frequências (1730 a 2270 MHz).

Optou-se pela técnica de demodulação coerente do sinal de RF, através do uso de um misturador *Mini-Circuits* ZFM-2000+, que insere perda de conversão média de 7,5 dB, considerando-se que na sua porta para entrada do oscilador local foi posicionado um gerador *Rohde & Schwarz* sintonizado na portadora 1,95 GHz com potência de +7 dBm.

O sincronismo entre as portadoras de 1,95 GHz dos blocos transmissor e receptor foi estabelecido através da passagem de um sinal de sincronismo de 10 MHz, por meio de um cabo coaxial RG-213 que unia os sistemas (do gerador *Anritsu*, na transmissão, ao oscilador *R&S*, na recepção).

A coleta da sequência PN recebida foi feita por meio de um osciloscópio digital *Agilent*, modelo DSO6034A. Para realizar o gatilho da sequência no osciloscópio foi necessário desenvolver um sinal de marcação no equipamento gerador. Este acionava um pulso de curta duração a cada início de sequência. Ligou-se, então, a interface de saída do sinal de marcação, na traseira do gerador, ao canal número 2 do osciloscópio, através de um segundo cabo coaxial, modelo RG-213. Desta forma, o sinal do canal 2 realizava o gatilho da PN, recebida no canal 1.

O osciloscópio empregado apresenta largura de banda de 300 MHz. O conhecimento da resposta em frequência assegura que a taxa de Nyquist [56] é respeitada, logo o sinal será amostrado satisfatoriamente, e que a componente da demodulação coerente em $f_s + f_{OL}$ será profundamente atenuada, dispensando a utilização de um filtro passa baixas na saída do misturador [57, 58].

Um *notebook*, conectado através de um barramento de dados GPIB ao osciloscópio digital, realizou o controle deste equipamento. Esta tarefa foi desempenhada com o desenvolvimento de rotina em MATLAB[®], cujo código é apresentado na função *captura*, no Apêndice deste trabalho. Foi utilizada a caixa de ferramentas de controle de instrumen-



Figura 5.6: Fotografia do sistema de recepção da sonda.

tos do *software* para acessar a interface de comunicação fornecida pelo driver IVI [59] do osciloscópio e, então, realizar a automação da aquisição e transferência de dados.

A Fig. 5.6 apresenta uma fotografia do sistema receptor da sonda construída.

5.2 Locais de medição

O local escolhido para as medições foi o prédio da Divisão de Eletricidade do Instituto Nacional de Normalização Metrologia e Qualidade Industrial (Inmetro). O interior deste prédio é constituído por apenas um andar, com um corredor no formato da letra U, com extensão aproximada de 115 m, altura de 2,94 m e largura de 2,77 m. Ao longo deste corredor há portas de entrada para 46 salas, onde existem laboratórios, escritórios e salas de máquinas. Na região central do prédio, cuja planta baixa está exibida na Fig. 5.7, está o *hall* de entrada e uma área de espera, separados pelo balcão da recepção. Uma porta de vidro, nos fundos da área de espera, dá acesso a um jardim descoberto de, aproximadamente, 1100 m².

O prédio é construído com paredes e laje de concreto de grossa espessura, que bloqueiam grande parte das interferências eletromagnéticas externas. Já as paredes que estão de frente para o jardim, assim como a porta de entrada do prédio são feitas de vidro.

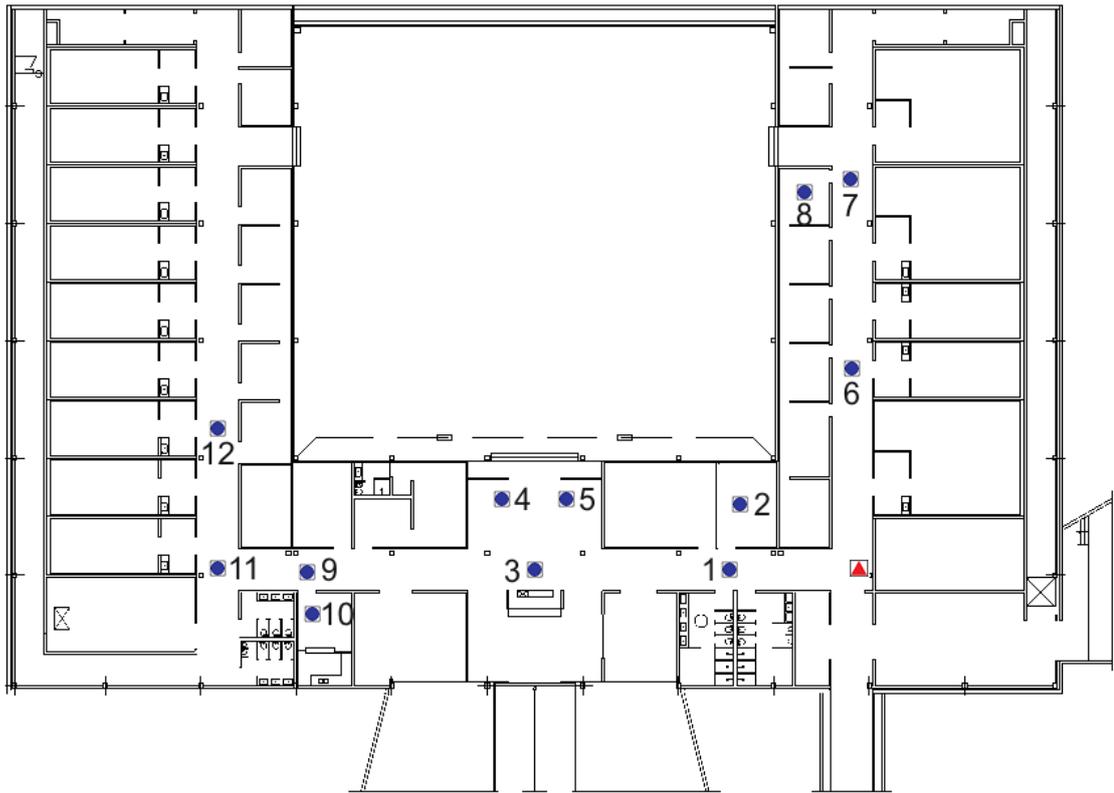


Figura 5.7: Planta do local das medições. O triângulo vermelho representa a posição da antena transmissora. Os círculos azuis representam os pontos de medição.

A razão da escolha por este local para as medições foi por tratar-se de um ambiente fechado com corredor conectado a várias salas, cenário que representa um ambiente típico de comunicação sem fio *indoor* em uma sala de escritórios, campus universitário, etc. A configuração do ambiente possibilitou a realização de medições em visada, ao longo do corredor, e fora de visada, quando o receptor é posicionado no interior de um laboratório ou escritório, ou está além da esquina do corredor.

A Fig. 5.8 apresenta fotografias do edifício onde foram realizadas as medições de sondagem.



(a)



(b)



(c)



(d)

Figura 5.8: Fotografias do edifício utilizado na sondagem do canal, relacionadas aos pontos de medição da planta da Fig. 5.7. (a) ponto 3, (b) ponto 4, (c) ponto 6, (d) ponto 8.

Capítulo 6

Ferramentas de Processamento

A campanha de medições, descrita no capítulo 5, resultou na aquisição das formas de onda recebidas em cada ponto medido pelo sistema receptor. Para que se possa obter informações sobre as características de propagação do canal, é necessário tratar adequadamente esse banco de dados obtidos pelo processo de sondagem.

6.1 Implementação do filtro casado

A implementação de um filtro casado possibilita a geração de um curva semelhante à resposta impulsiva do canal, em cada ponto de medição, conforme foi descrito na subseção 4.2.2. Então, será possível aplicar técnicas de processamento a estes dados, baseadas na teoria do capítulo 3, e os parâmetros descritivos do canal se tornarão acessíveis.

O filtro casado foi realizado via *software*, sendo necessária, então, a definição da sua resposta ao impulso de maneira que o processo de convolução resultasse na expressão da Eq. (4.4). Esta foi obtida em laboratório, através da modificação do *setup* original da Fig. 5.1, onde as antenas foram retiradas e os blocos transmissor e receptor foram diretamente conectados, através de um cabo coaxial calibrado, de maneira semelhante àquela descrita na seção 5.1 para os testes de projeto da sequência PN. O sinal coletado, livre das distorções do canal rádio, pode ser utilizado na convolução com os sinais obtidos das campanhas de medição [52].

Os gráficos da Fig. 6.1 exibem dois sinais resultantes da convolução por filtro ca-

sado, originários de pontos de medição distintos. A Fig. 6.1a foi gerada a partir de um ponto onde existe visada direta entre os sistemas transmissor e receptor. Ela claramente exhibe uma componente de multipercurso de maior amplitude, a que percorre uma linha reta entre as duas antenas, seguida de outras componentes de menor amplitude. A Fig. 6.1b revela um estado em que não há visada, onde muitas componentes alcançam o receptor e não há diferença de amplitude relevante entre elas.

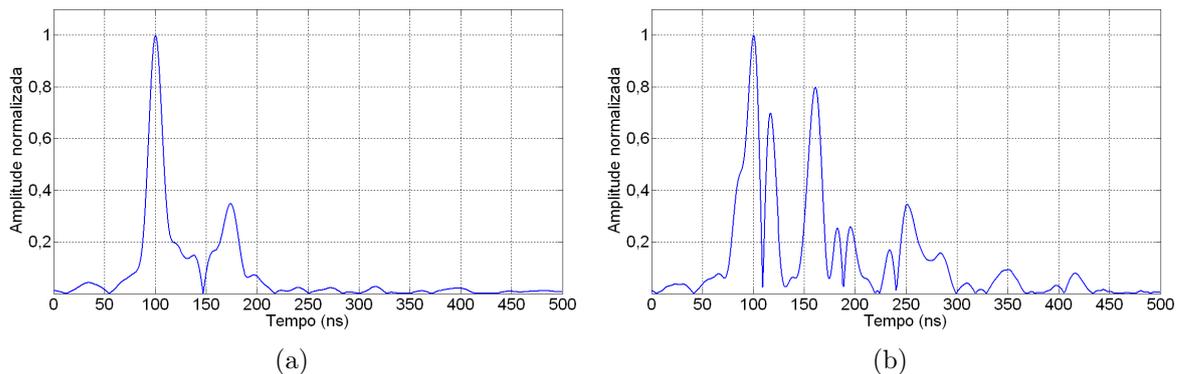


Figura 6.1: Sinal na saída do filtro casado. (a) Ponto de medição em LOS. (b) Ponto de medição em NLOS.

Na atual circunstância, a resposta ao impulso do canal, $h(t)$, pode ser extraída dos dados provenientes da sondagem. Um modelo amplamente disseminado, que descreve canais multipercursos, foi sugerido por Turin [4]. Este modelo tem sido utilizado de forma recorrente na caracterização do comportamento de canais *indoor* [8]. Seguindo este modelo, a resposta ao impulso de um canal invariante no tempo é dada pela Eq. (6.1):

$$\begin{aligned}
 h(t) &= \sum_{l=1}^L a_l \delta(t - \tau_l) e^{j\theta_l} \\
 &= \sum_{l=1}^L w_l \delta(t - \tau_l),
 \end{aligned} \tag{6.1}$$

onde L representa o número de multipercursos; w_l , a amplitude complexa do l -ésimo multipercurso; τ_l , o retardo de propagação deste multipercurso, e δ é a função delta de Dirac. O conhecimento da resposta impulsiva de um sistema estacionário possibilita a determinação de sua saída para a transmissão de um sinal qualquer, conforme mostra a

Eq. (6.2):

$$\begin{aligned}
 y(t) &= \int_{-\infty}^{\infty} u(t')h(t-t') dt' + \eta(t) \\
 &= \sum_{l=1}^L w_l u(t - \tau_l) + \eta(t).
 \end{aligned} \tag{6.2}$$

O termo aditivo $\eta(t)$ representa um ruído aleatório, modelado como gaussiano, com densidade espectral N_0 . Considerando $u(t)$ uma sequência pseudo-aleatória, o sinal na saída do filtro é obitdo por processo semelhante ao descrito na Eq. (4.4). Desta forma, chega-se à Eq. (6.3):

$$z(t) = \sum_{l=1}^L w_l R_{uu}(t - \tau_l) + \xi(t), \tag{6.3}$$

onde $R_{uu}(t)$ e $\xi(t)$ são a função de autocorrelação da sequência PN transmitida e a correlação cruzada da sequência com o ruído aditivo, respectivamente. Pode-se, também, extrair a relação expressa pela Eq. (6.4):

$$R_{\xi\xi}(t) = E[\xi(t')\xi^*(t+t')] = N_0 R_{uu}(t). \tag{6.4}$$

É notável a semelhança entre o modelo do canal, na Eq. (6.1), e a saída do filtro casado, na Eq. (6.3), levando em consideração que $R_{uu}(t)$ possui um formato triangular estreito, que tende à função $\delta(t)$, e que, no resultado da sondagem, haverá um termo originário do ruído do canal, $\xi(t)$. Estratégias de estimação, como o algoritmo abaixo descrito, têm o objetivo de determinar, de maneira precisa, os parâmetros do modelo que descreve a resposta do canal.

6.2 A Máquina de Vetor de Relevância Aplicada à Estimação do Canal Multipercurso

Em condições ideais, a resposta do canal é composta por uma estrutura esparsa, ou seja, são verificados picos correspondentes aos instantes de chegada dos multipercursos. No entanto, esta característica é ofuscada pelo ruído aditivo e pela dispersão temporal, devido à banda finita dos *hardwares* do transmissor e receptor. Isto motiva a aplicação de algoritmos capazes de recuperar a estrutura da função $h(t, \tau)$ do canal.

As características, já citadas, pertencentes à estrutura da Máquina de Vetor de Relevância, como a característica esparsa pelo uso reduzido de funções base, além das propriedades herdadas por utilizar-se da plataforma bayesiana, como a preferência por hipóteses simplificadas que modelam os dados observados, fazem do RVM uma técnica adequada para a aplicação na identificação dos parâmetros da resposta do canal multipercorso sem fio.

O procedimento de sondagem pode ser, resumidamente, ilustrado pelo diagrama da Fig. 6.2, onde $s(t)$ corresponde à transmissão de uma sequência PN que se repete continuamente.

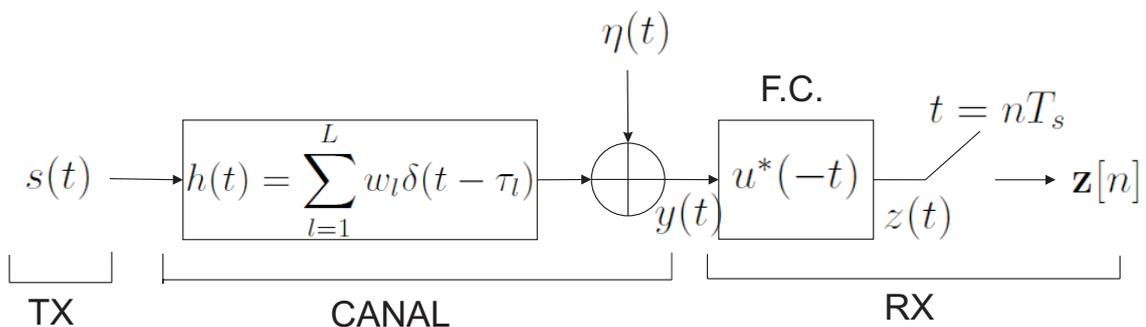


Figura 6.2: Modelo banda base equivalente da sondagem faixa larga do canal rádio.

A Eq. (6.3) estabelece que a saída do filtro casado, representada pelo vetor \mathbf{z} , é uma combinação linear de L funções núcleo retardadas e ponderadas, adicionadas ao ruído $\xi(t)$. Para uma taxa de amostragem $\frac{1}{T_s}$, onde T_s é o intervalo de amostragem na saída do filtro casado e, T_p sendo a duração de bit da PN, $T_s \leq T_p$, são geradas N amostras na saída do filtro casado. Assim, é possível reescrever a Eq. (6.3) na forma vetorial, como na

Eq. (6.5):

$$\mathbf{z} = \mathbf{K}\mathbf{w} + \boldsymbol{\xi}. \quad (6.5)$$

Define-se:

$$\mathbf{z} = [z[0], z[1], \dots, z[N-1]]^T, \quad (6.6)$$

$$\mathbf{w} = [w_1, w_2, \dots, w_L]^T, \quad (6.7)$$

$$\boldsymbol{\xi} = [\xi[0], \xi[1], \dots, \xi[N-1]]^T. \quad (6.8)$$

O vetor ruído aditivo possui as características estatísticas dadas pelas Eqs. (6.9) e (6.10), a serem exploradas.

$$\mathbb{E}[\boldsymbol{\xi}] = 0, \quad (6.9)$$

$$\mathbb{E}[\boldsymbol{\xi}\boldsymbol{\xi}^H] = \boldsymbol{\Sigma} = N_0\boldsymbol{\Lambda}, \quad \Lambda_{i,j} = R_{uu}\left((i-j)T_S\right). \quad (6.10)$$

A matriz $\boldsymbol{\Sigma}$ representa a matriz covariância de $\boldsymbol{\xi}$ e é decorrente da Eq. (6.4). A matriz \mathbf{K} é denominada matriz de projeto. Ela acumula versões amostradas e deslocadas da função núcleo $R_{uu}(t)$. Apresenta-se estruturada em vetores \mathbf{r}_l , que representam funções núcleo, cada uma associada a um deslocamento temporal τ_l específico.

$$\mathbf{K} = [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_L] \quad (6.11)$$

$$\mathbf{r}_l = \left[R_{uu}\left(-\tau_l\right), R_{uu}\left(T_S - \tau_l\right), \dots, R_{uu}\left((N-1)T_S - \tau_l\right) \right]^T. \quad (6.12)$$

É notável a correspondência direta entre a Eq. (6.5) e o problema de regressão da Eq. (3.18), caracterizando o problema como uma questão de regressão linear que pode ser efetivamente solucionada através da aplicação da técnica da Máquina de Vetor de Relevância.

O problema de estimação do canal a ser resolvido surge do seguinte cenário: dado o vetor \mathbf{z} , determinar a ordem L do modelo, que representa o número de multipercursos, e estimar de forma otimizada o vetor \mathbf{w} , chamado vetor de parâmetros ou vetor de pesos, que armazena as amplitudes dos multipercursos presentes no perfil de potência de retardo,

eliminada a contribuição do ruído. Além disto, deverá ser determinado, com uma precisão definida inicialmente, o espaço de busca dos retardos de multipercursos τ_l . O espaço de busca dos multipercursos $\mathcal{T} = \{T_1, \dots, T_{L_0}\}$ compõe a hipótese de modelo inicial \mathcal{H}_0 , manifestada através das L_0 colunas da matriz de projeto \mathbf{K} .

O objetivo é, para um conjunto de hipóteses \mathcal{H} , realizar um processo iterativo que alterne o ajuste de modelos, estimando os parâmetros \mathbf{w} da hipótese em questão, e a seleção da hipótese, a fim de encontrar o modelo de ordem $L \leq L_0$ que represente a resposta ao impulso do canal.

Após a caracterização do sistema que compõe o processo de sondagem, devem ser caracterizados os parâmetros que definem as hipóteses de modelo \mathcal{H} , denominados hiperparâmetros. A partir da Eq. (6.10), verificamos que a matriz covariância do ruído é proporcional a N_0 , que deve ser estimado a partir dos dados. Desta forma, o modelo \mathcal{H} deve incluir o termo N_0 . Cria-se então $\beta = N_0^{-1}$ como uma variável aleatória que existe no intervalo $[0; \infty)$ e sua f.d.p. é definida na Eq. (6.13) como uma distribuição gama.

$$h(\beta|\kappa, \nu) = \frac{\kappa^\nu}{\Gamma(\nu)} \beta^{\nu-1} e^{-\kappa\beta}. \quad (6.13)$$

Esta distribuição possui os parâmetros κ, ν que definem a informação *a priori* a respeito de N_0 . Na ausência de conhecimento prévio a respeito, pode-se utilizar uma distribuição não informativa, que é representada por uma distribuição que tende à distribuição uniforme. Isto é alcançado atribuindo-se valores pequenos para os parâmetros da distribuição da Eq. (6.13), tipicamente $\kappa = \nu = 10^{-4}$.

O segundo hiperparâmetro associado à hipótese, que guia o mecanismo de seleção de modelo é o hiperparâmetro de relevância. É constituído de um vetor de parâmetros $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_{L_0}\}$, com cada elemento associado a cada coluna da matriz \mathbf{K} . Este parâmetro controla a contribuição (ou relevância) do correspondente peso w_i na explicação dos dados z . Se os pesos \mathbf{w} podem ser considerados i.i.d. (ou seja, a contribuição dos multipercursos é considerada estatisticamente independente), $h(\mathbf{w}|\boldsymbol{\alpha})$ tem a forma exibida

na Eq. (3.33), e é reescrita como a Eq. (6.14):

$$h(\mathbf{w}|\boldsymbol{\alpha}) = \prod_{l=1}^{L_0} \frac{\alpha_l}{\pi} e^{-|w_l|^2 \alpha_l}. \quad (6.14)$$

Altos valores de α_l forçam o peso correspondente w_l a valores próximos de zero. Nesta situação, a l -ésima coluna da matriz de projeto deve ser excluída e o modelo, rejeitado. Adicionalmente, o retardo correspondente T_l é removido do espaço de busca de multipercursos \mathcal{T} .

Para finalizar a especificação do modelo, define-se a distribuição *a priori* do hiperparâmetro $\boldsymbol{\alpha}$, também considerada uma distribuição gama.

$$h(\boldsymbol{\alpha}|\zeta, \epsilon) = \prod_{l=1}^L \frac{\zeta^\epsilon}{\Gamma(\epsilon)} \alpha_l^{\epsilon-1} e^{-\zeta \alpha_l}. \quad (6.15)$$

Como no caso do parâmetro β , será feito $\zeta = \epsilon = 10^{-4}$ para tornar a distribuição de $\boldsymbol{\alpha}$ não informativa.

O processo de aprendizagem consiste em inferir os parâmetros de \mathbf{w}_i e da hipótese \mathcal{H}_i (seus hiperparâmetros $\boldsymbol{\alpha}_i$ e β), maximizando a distribuição *a posteriori*, escrita na Eq. (3.12). Como na Eq. (3.13), pelo método da estimação marginal, escreve-se a Eq. (6.16):

$$h(\mathbf{w}, \boldsymbol{\alpha}, \beta|\mathbf{z}) = h(\mathbf{w}|\mathbf{z}, \boldsymbol{\alpha}, \beta)h(\boldsymbol{\alpha}, \beta|\mathbf{z}). \quad (6.16)$$

O primeiro termo do ramo direito da equação é identificado como uma distribuição *a posteriori* de parâmetro, enquanto o segundo termo é uma distribuição *a posteriori* de hipótese.

O primeiro passo do processo de inferência corresponde ao ajuste de modelo e consiste na determinação dos valores de \mathbf{w} de uma hipótese particular \mathcal{H}_i que maximizem a distribuição *a posteriori* de parâmetro. Utilizando-se do procedimento descrito nas Eqs. (3.25)-(3.29) e observando a equivalência entre a variável alvo \mathbf{t} e os dados observados na saída do filtro casado \mathbf{z} , tem-se que a verossimilhança de \mathbf{w} e β tem a forma da

distribuição gaussiana, dada pela Eq. (6.17):

$$f(\mathbf{z}|\mathbf{w}, \beta) = \frac{\exp\{-(\mathbf{z} - \mathbf{K}\mathbf{w})^H \beta \mathbf{\Lambda}^{-1} (\mathbf{z} - \mathbf{K}\mathbf{w})\}}{\pi^N |\beta^{-1} \mathbf{\Lambda}|}. \quad (6.17)$$

Assim, o produto $h(\mathbf{w}|\boldsymbol{\alpha})f(\mathbf{z}|\mathbf{w}, \beta)$ também será uma distribuição gaussiana. Pela Eq. (3.29), verifica-se que este produto é proporcional à distribuição *a posteriori* de parâmetro $h(\mathbf{w}|\mathbf{z}, \boldsymbol{\alpha}, \beta)$. A aplicação do estimador máximo *a posteriori* fornece os resultados dados pelas Eqs. (6.18) e (6.19):

$$\boldsymbol{\Phi} = (\mathbf{A} + \beta \mathbf{K}^H \mathbf{\Lambda}^{-1} \mathbf{K})^{-1}, \quad (6.18)$$

$$\boldsymbol{\mu} = \beta \boldsymbol{\Phi} \mathbf{K}^H \mathbf{\Lambda}^{-1} \mathbf{z}. \quad (6.19)$$

Identifica-se $\boldsymbol{\mu}$ como o estimador MAP dos parâmetros \mathbf{w} da hipótese \mathcal{H}_i e $\boldsymbol{\Phi}$ a matriz covariância associada. Reconhece-se, também, $\mathbf{A} = \text{diag}(\boldsymbol{\alpha})$.

Na fase de comparação de modelos, deve-se determinar a hipótese \mathcal{H}_i , ou mais precisamente, seus parâmetros $\boldsymbol{\alpha}$ e β que maximizem o segundo termo da Eq. (6.16), $h(\boldsymbol{\alpha}, \beta|\mathbf{z})$. Aplicando o Teorema de Bayes, reescreve-se esse termo como:

$$h(\boldsymbol{\alpha}, \beta|\mathbf{z}) \propto f(\mathbf{z}|\boldsymbol{\alpha}, \beta)h(\boldsymbol{\alpha}, \beta) = f(\mathbf{z}|\boldsymbol{\alpha}, \beta)h(\boldsymbol{\alpha})h(\beta). \quad (6.20)$$

Esta equação tem a forma especificada na Eq. (3.15), e segue a estrutura:

a posteriori do modelo \propto evidência \times *a priori* do modelo.

Arbitrando-se $\zeta = \epsilon = \kappa = \nu = 0$, faz-se a escolha por distribuições *a priori* não informativas. Desta maneira, maximizar a distribuição *a posteriori* do modelo equivale a maximizar o termo evidência. Este é o chamado procedimento de evidência [26] e é obtido marginalizando sobre o espaço de parâmetros, como feito na Eq. (3.30).

$$\begin{aligned} f(\mathbf{z}|\boldsymbol{\alpha}, \beta) &= \int f(\mathbf{z}|\mathbf{w}, \beta)h(\mathbf{w}|\boldsymbol{\alpha}) d\mathbf{w} \\ &= \frac{\exp(-\mathbf{z}^H (\beta^{-1} \mathbf{\Lambda} + \mathbf{K} \mathbf{A}^{-1} \mathbf{K}^H) \mathbf{z})}{\pi^N |\beta^{-1} \mathbf{\Lambda} + \mathbf{K} \mathbf{A}^{-1} \mathbf{K}^H|}. \end{aligned} \quad (6.21)$$

Na Eq. (6.21), foram utilizadas as características de independência entre as variáveis para simplificar o integrando. O termo acima também é chamado verossimilhança marginal e sua maximização em relação aos hiperparâmetros é reconhecida como verossimilhança máxima tipo 2 [23]. Colocando os termos da Eq. (6.21) em função de $\boldsymbol{\mu}$ e $\boldsymbol{\Phi}$, derivando a expressão e fazendo-a igual a zero, obtemos, nas Eqs. (6.22) e (6.23), os valores maximizantes dos hiperparâmetros.

$$\alpha_l = \frac{1}{\Phi_{ll} + |\boldsymbol{\mu}_l|^2}, \quad (6.22)$$

$$\beta^{-1} = \frac{\text{tr}[\boldsymbol{\Phi}\mathbf{K}^H\boldsymbol{\Lambda}^{-1}\mathbf{K}] + (\mathbf{z} - \mathbf{K}\boldsymbol{\mu})^H\boldsymbol{\Lambda}^{-1}(\mathbf{z} - \mathbf{K}\boldsymbol{\mu})}{N}. \quad (6.23)$$

Desta forma, o algoritmo de aprendizagem consiste de, para uma particular hipótese \mathcal{H}_i e valores iniciais para $\boldsymbol{\alpha}_i$ e β , realizar a aplicação repetida das Eqs. (6.18) e (6.19), alternada com a atualização dos parâmetros de evidência pelas Eqs. (6.22) e (6.23), até que o sistema chegue a um ponto de convergência, após j iterações. Este processo é ilustrado na Fig. 6.3.

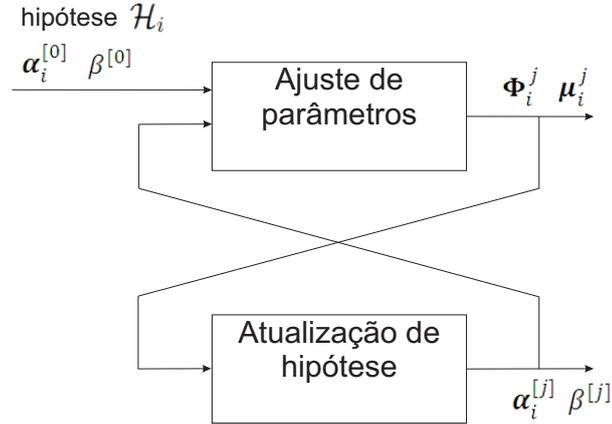


Figura 6.3: Processo iterativo de aprendizagem dos parâmetros do modelo do canal multipercursos.

A detecção de componentes de multipercursos pode ser aprimorada caso sejam realizadas medições adicionais do canal, seja pela captura de mais de uma sequência de sondagem transmitida ou pelo uso de antenas de múltiplos sensores. Para estes casos, a análise acima descrita pode ser estendida para incorporar as observações adicionais. Neste novo cenário, existem P vetores de saída do filtro, \mathbf{z}_p , originários de observações individuais

$p = 1 \cdots P$. De maneira semelhante, haverá P matrizes de projeto, que serão idênticas, caso seja considerado o uso do mesmo filtro casado para todas as observações, e P vetores de parâmetros \mathbf{w} . No entanto, um único vetor de hiperparâmetros $\boldsymbol{\alpha}$ será utilizado no controle dos vetores de parâmetros. O uso de um hiperparâmetro α_i , para controlar um conjunto de P pesos, garante a dependência entre as observações consecutivas do canal. Neste caso, as Eqs. (6.18), (6.19), (6.22) e (6.23) podem ser reescritas como as Eqs. (6.24) (6.25), (6.26), (6.27):

$$\boldsymbol{\Phi}_p = (\mathbf{A} + \beta \mathbf{K}^H \boldsymbol{\Lambda}_p^{-1} \mathbf{K})^{-1}, \quad (6.24)$$

$$\boldsymbol{\mu}_p = \beta \boldsymbol{\Phi}_p \mathbf{K}^H \boldsymbol{\Lambda}_p^{-1} \mathbf{z}_p, \quad p = 1 \cdots P \quad (6.25)$$

$$\alpha_l = \frac{P}{\sum_{p=1}^P \boldsymbol{\Phi}_{p,ll} + |\boldsymbol{\mu}_{p,l}|^2}, \quad (6.26)$$

$$\beta^{-1} = \frac{1}{NP} \left(\sum_{p=1}^P \text{tr}[\boldsymbol{\Phi}_p \mathbf{K}^H \boldsymbol{\Lambda}_p^{-1} \mathbf{K}] (\mathbf{z}_p - \mathbf{K} \boldsymbol{\mu}_p)^H \boldsymbol{\Lambda}_p^{-1} (\mathbf{z}_p - \mathbf{K} \boldsymbol{\mu}_p) \right). \quad (6.27)$$

Como descrito no capítulo 3, o resultado do processo iterativo fará com que uma quantidade significativa dos elementos do vetor de hiperparâmetros $\boldsymbol{\alpha}$ assumam valores muito altos e, conseqüentemente, os parâmetros peso correspondentes se concentrarão em torno de zero. A experiência empírica sugere que o estabelecimento de um valor limite para α , a partir do qual pode-se excluir a função núcleo correspondente, mostra-se eficaz somente em elevados regimes de relação sinal-ruído. Em vez disto, o trabalho descrito em [60] desenvolve uma análise estatística que estabelece critérios para a exclusão de funções núcleo e, por conseguinte, realiza a seleção de modelo.

Para iniciar esta análise, é preciso realizar algumas simplificações. É necessário considerar a existência de apenas uma componente de multipercorso, ou seja, $L = 1$, com retardo conhecido τ . Isto implica numa matriz de projeto $\mathbf{K} = [\mathbf{r}(\tau)]$, onde $\mathbf{r}(\tau) = [R_{uu}(-\tau), R_{uu}(T_s - \tau), \dots, R_{uu}((N-1)T_s - \tau)]^T$ e o hiperparâmetro associado a esta componente é denotado por α . Em seguida, é necessário determinar a solução α_∞ , chamada de solução de regime estacionário, para o hiperparâmetro deste cenário simplificado. As Eqs. (6.26) e (6.27) se simplificam nas Eqs. (6.28) e (6.29):

$$\phi = \left(\alpha + \mathbf{r}(\tau)^H \mathbf{B} \mathbf{r}(\tau) \right)^{-1}, \quad (6.28)$$

$$\mu_p = \phi \mathbf{K}^H \mathbf{B} \mathbf{z}_p = \frac{\mathbf{r}(\tau)^H \mathbf{B} \mathbf{z}_p}{\alpha + \mathbf{r}(\tau)^H \mathbf{B} \mathbf{r}(\tau)}, \quad p = 1 \cdots P. \quad (6.29)$$

Inserindo as Eqs. (6.28) e (6.29) na Eq. (6.26), chega-se ao valor de α_∞ , exibido na Eq. (6.30):

$$\alpha_\infty = \frac{\left(\mathbf{r}(\tau)^H \mathbf{B} \mathbf{r}(\tau) \right)^2}{\frac{1}{P} \sum_p |\mathbf{r}(\tau)^H \mathbf{B} \mathbf{z}_p|^2 - \mathbf{r}(\tau)^H \mathbf{B} \mathbf{r}(\tau)} \quad (6.30)$$

Relembrando a Eq. (6.14), verifica-se que α corresponde à variância da distribuição *a priori* de \mathbf{w} . Isto significa que a Eq. (6.26) só irá convergir para a Eq. (6.30) caso o resultado desta equação seja um valor positivo, ou mais especificamente, seu denominador seja positivo. A não realização desta condição implica que o processo iterativo irá divergir, isto é, $\alpha_\infty = \infty$. Assim, para uma dada medição \mathbf{z}_p e uma matriz de ruído conhecida \mathbf{B} , pode-se identificar, imediatamente, se a função base avaliada $\mathbf{r}(\tau)$ deve ser excluída do processo de iteração.

Este procedimento possui o papel de realizar a eliminação inicial das funções base consideradas irrelevantes. Entretanto, uma análise mais detalhada deve ser realizada para que não ocorra uma sobre-estimação do número de componentes de multipercorso. Nesta análise aprofundada, devem ser alteradas as considerações sobre o sistema. Agora, assume-se que o retardo do único multipercorso presente τ é desconhecido. A matriz de projeto terá a forma da Eq. (6.31):

$$\mathbf{K} = \begin{bmatrix} \mathbf{r}_l \end{bmatrix}$$

$$\mathbf{r}_l = \left[R_{uu}(-T_l), \dots, R_{uu}((N-1)T_s - T_l) \right]^T, \quad (6.31)$$

onde esta função base é associada ao retardo $T_l \in \mathcal{T}$. Reescrevemos a Eq. (6.5) para este caso, conforme a Eq. (6.32):

$$\mathbf{z}_p = w_p \mathbf{r}(\tau) + \boldsymbol{\xi}_p. \quad (6.32)$$

Nesta expressão assumimos o caso genérico, em que ocorre a observação de um conjunto de P sequências PN. Realizando a combinação da Eq. (6.32) com a Eq. (6.30),

chega-se a um caso particular da expressão de α_∞^{-1} dada pela Eq. (6.33):

$$\begin{aligned} \alpha_\infty^{-1} = & \frac{|\mathbf{r}_l^H \mathbf{B} \mathbf{r}(\tau)|^2 \sum_p |w_p|^2}{P |\mathbf{r}_l^H \mathbf{B} \mathbf{r}_l|^2} + \frac{2 \sum_p \operatorname{Re}[w_p \mathbf{r}_l^H \mathbf{B} \mathbf{r}(\tau) \boldsymbol{\xi}_p^H \mathbf{B} \mathbf{r}_l]}{P |\mathbf{r}_l^H \mathbf{B} \mathbf{r}_l|^2} \\ & + \frac{\mathbf{r}_l^H \mathbf{B} (\sum_p \boldsymbol{\xi}_p \boldsymbol{\xi}_p^H) \mathbf{B} \mathbf{r}_l}{P |\mathbf{r}_l^H \mathbf{B} \mathbf{r}_l|^2} - \frac{1}{\mathbf{r}_l^H \mathbf{B} \mathbf{r}_l}. \end{aligned} \quad (6.33)$$

Conforme Shutin [60, 61], a partir da Eq. (6.33), a análise é dividida em dois cenários. O primeiro cenário corresponde ao caso da componente de multipercurso avaliada coincidir com um elemento do espaço de busca, ou seja, $T_l \in \mathcal{T}$. O segundo cenário refere-se ao caso de o espaço de busca não contemplar o instante de chegada do multipercurso avaliado.

Para o primeiro caso, no qual considera-se $\tau = T_l$ e $\mathbf{r}(\tau) = \mathbf{r}_l$, a separação dos quatro termos que compõem a Eq. (6.33) possibilita visualizar que esta expressão pode ser dividida em dois grupos: 1- termos que consideram a contribuição pura do ruído, representados por α_n^{-1} ; 2- termos que levam em consideração somente a contribuição dos multipercursos e excluem a contribuição do ruído, representados por α_s^{-1} . Para este caso, a solução é regida pela Eq. (6.34):

$$f_{\alpha_n^{-1}}(x) = P_n \delta(x) + \ell^+(x) f_{\tilde{\alpha}_n^{-1}}(x | P, \sigma_c^2), \quad (6.34)$$

onde $\delta(\cdot)$ é a função delta de Dirac, $\ell^+(x)$ é a função indicatriz definida na Eq. (6.35), P_n representa a integral da Eq. (6.36) e $\tilde{\alpha}_n^{-1}$ é uma variável aleatória de média zero com f.d.p. definida pela Eq. (6.37):

$$\ell^+(x) = \begin{cases} 0 & x \leq 0, \\ 1 & x > 0. \end{cases} \quad (6.35)$$

$$P_n = \int_{-1/\mathbf{r}_l^H \mathbf{B} \mathbf{r}_l}^0 f_{\tilde{\alpha}_n^{-1}}(x|P, \sigma_c^2) d\tau. \quad (6.36)$$

$$f_{\tilde{\alpha}_n^{-1}}(x|P, \sigma_c^2) = \frac{\left(x - \frac{1}{\mathbf{r}_l^H \mathbf{B} \mathbf{r}_l}\right)^{P-1}}{\Gamma(P)(\sigma_c^2)^P} \exp\left(-\left(x - \frac{1}{\mathbf{r}_l^H \mathbf{B} \mathbf{r}_l}\right)/\sigma_c^2\right) \quad (6.37)$$

O problema de exclusão de uma função base reside na decisão se o valor do hiperparâmetro α^{-1} correspondente foi gerado pela f.d.p. de α_n^{-1} , dada pela Eq. (6.34), chamada de “hipótese nula”, ou pela f.d.p. de $\alpha_s^{-1} + \alpha_n^{-1}$, chamada “hipótese alternativa”. Isto significa avaliar se esta função base possui apenas influência do ruído ou se também existe contribuição de alguma componente de multipercurso. Entretanto, a dificuldade na determinação de uma f.d.p. para a hipótese alternativa faz com que se utilize uma abordagem diferente para a solução deste problema.

Assumindo que, ao final do processo iterativo, conhece-se o valor de $N_0 = \beta^{-1}$, então, para qualquer função base \mathbf{r}_l da matriz de projeto \mathbf{K} e o correspondente hiperparâmetro α_l , pode-se decidir, com uma probabilidade ρ especificada *a priori*, que α_l é gerada pela hipótese nula dada pela Eq. (6.34). Assim, define-se α_{th}^{-1} como o valor da variável aleatória α^{-1} dada pela f.d.p. definida na Eq. (6.34), tal que $P(\alpha^{-1} \leq \alpha_{th}^{-1}) = \rho$. Assim, o processo de decisão consistirá de comparar os valores dos elementos α_l^{-1} do vetor de hiperparâmetros ao valor limite α_{th}^{-1} . A ocorrência de $\alpha_l^{-1} > \alpha_{th}^{-1}$ significa dizer, com probabilidade ρ , que esta função base foi gerada pela hipótese alternativa e, portanto, possui contribuição de multipercurso; já a ocorrência de $\alpha_l^{-1} \leq \alpha_{th}^{-1}$ não garante que a função base avaliada foi gerada pela hipótese alternativa e, assim, deve ser excluída.

O segundo caso corresponde à situação em que leva-se em consideração a existência de um desalinhamento entre o instante de ocorrência de um multipercurso, τ , e os instantes de tempo contemplados pelo espaço de busca, T_l . Este estudo, realizado em [60], consiste em avaliar a influência deste descasamento, partindo da expressão de α_∞^{-1} , na Eq. (6.33). É observado que, quanto maior o descasamento entre T_l e τ , menor será o peso da contribuição de α_s^{-1} para possibilitar que esta componente de multipercurso seja detectada. Mais precisamente, isto significa que, para um descasamento especificado por $|\tau - T_l| < T_p$, a contribuição de α_s^{-1} deve ser considerada; para o caso de $|\tau - T_l| \geq T_p$, o

descasamento é maior que a duração de bit e o valor de α_{th}^{-1} é calculado usando somente a Eq. (6.34). É possível mostrar que a contribuição aproximada, causada pelo possível desalinhamento entre T_l e τ , pode ser calculada pela Eq. (6.38) e o novo limite para a exclusão de funções base é obtido através da Eq. (6.39):

$$\hat{\alpha}_s^{-1} = \frac{3}{8} \frac{\sum_{p=0}^{P-1} |\mu_p|^2}{P}, \quad (6.38)$$

$$\hat{\alpha}_{th}^{-1} = \hat{\alpha}_s^{-1} + \alpha_{th}^{-1}. \quad (6.39)$$

O desenvolvimento de uma rotina computacional, que obedece aos procedimentos descritos nesta seção, possibilitou a aplicação do RVM aos dados medidos. A estrutura de funcionamento do algoritmo é descrita no diagrama de atividades da Fig. 6.4.

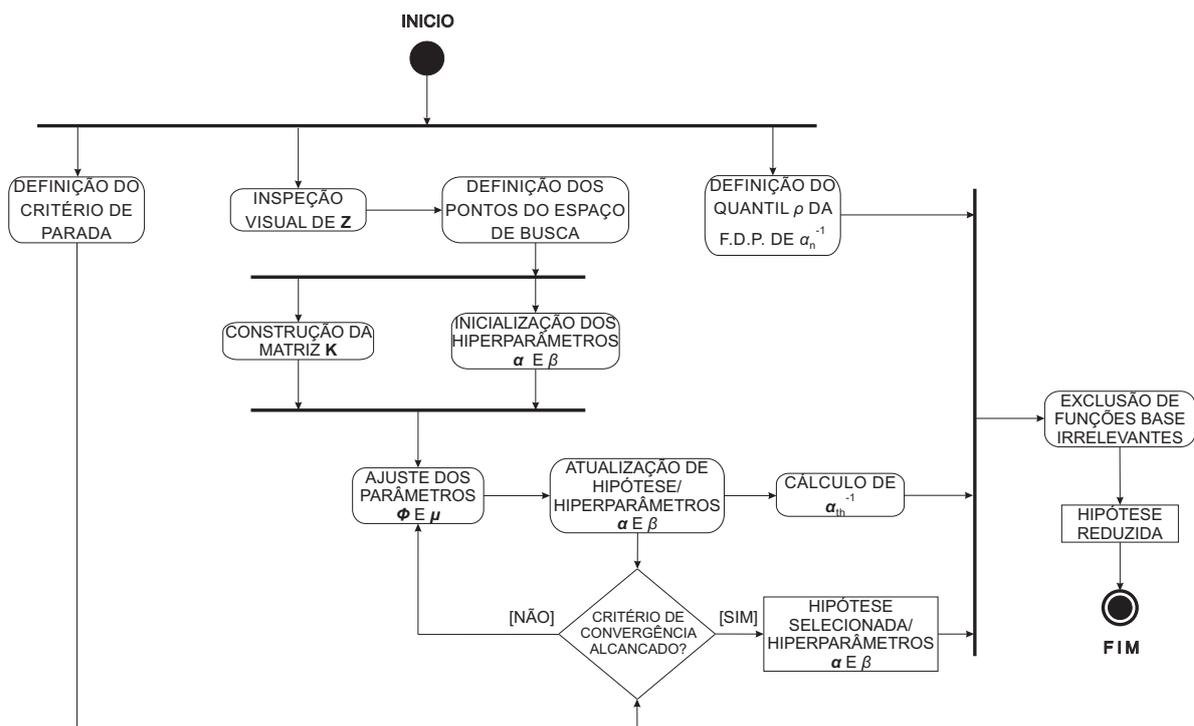


Figura 6.4: Diagrama de atividades do algoritmo implementado.

Com o fornecimento do vetor \mathbf{z} como entrada e ajustando corretamente os parâmetros do algoritmo, o processo iterativo de seleção de hipótese fornecerá os parâmetros constitutivos da função $h(t)$. Os gráficos superiores da Fig. 6.5 exibem exemplos do sinal de saída do filtro casado, que são usados como dados de entrada ao RVM. Os gráficos inferiores mostram o resultado da aplicação do algoritmo, onde podem ser identificados

a quantidade de multipercursos detectados, seus instantes de chegada e amplitudes relativas. A abordagem estatística da estratégia de eliminação das funções base irrelevantes, pelo RVM, permite a não necessidade de adoção de um valor mínimo de amplitude das componentes de multipercursos que pode ser detectado pelo algoritmo, o chamado limiar mínimo de detecção, existente em outros métodos [9, 62].

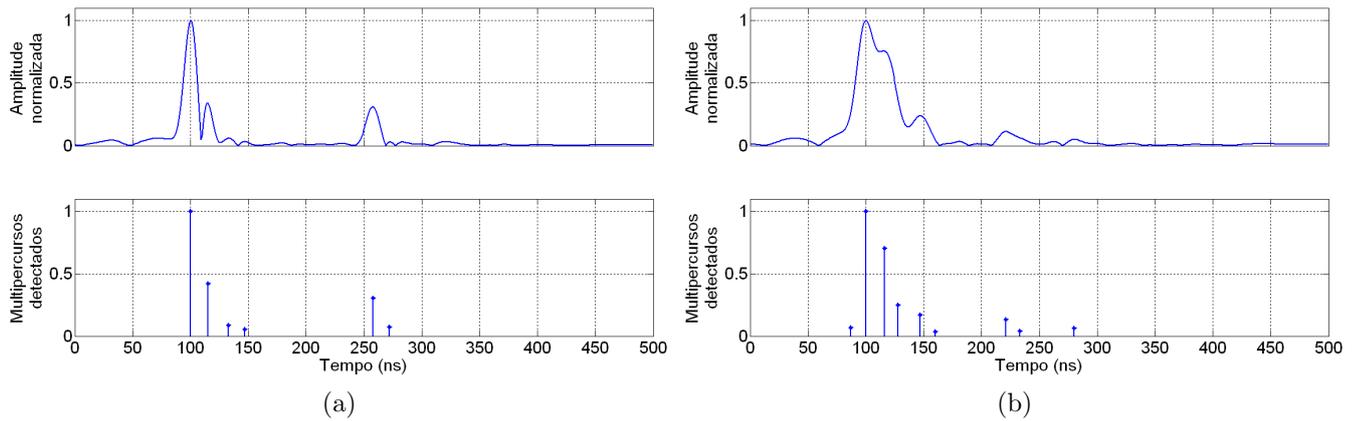


Figura 6.5: Exemplos do resultado da aplicação do RVM a amostras dos dados colhidos na sondagem do canal. (a) situação LOS, (b) situação NLOS.

Capítulo 7

Resultados e Discussão

Este capítulo apresenta os resultados da aplicação das ferramentas de processamento aos dados obtidos na sondagem da resposta do canal. A identificação de parâmetros relevantes e a análise estatística dos dados coletados neste trabalho visam fornecer informações relevantes sobre as características do canal em estudo.

7.1 Avaliação do desempenho do algoritmo RVM

Uma vez verificado que o algoritmo RVM, apresentado no capítulo 6, cumpre o objetivo de identificação dos parâmetros do modelo da resposta impulsiva do canal, torna-se de interesse uma investigação mais detalhada do desempenho deste algoritmo.

A inspeção dos dados coletados, a partir da sondagem do canal, possibilitou a verificação de que os dados obtidos deram origem a respostas impulsivas bastante características de ambientes *indoor*, como a presença de multipercursos oriundos de espalhadores próximos, especialmente em pontos de medição fora de visada. Além disto, foi possível verificar uma alta relação sinal-ruído; isto é uma característica típica de ambientes confinados e, também, deve-se ao projeto da sonda e à técnica de processamento empregados, que possibilitaram a obtenção da resposta do canal de maneira direta e com baixa inserção de ruído. Esta condição de baixo patamar de ruído representa uma condição que facilita a correta detecção dos parâmetros do modelo do canal pelo algoritmo. Torna-se, portanto, uma questão de interesse a observação do desempenho do RVM em canais que apresentam

potência de ruído superior às observadas no ambiente medido.

Como exibido no diagrama de atividades da Fig. 6.4, o algoritmo RVM possibilita o ajuste de certos parâmetros que influenciam no processo iterativo e nos critérios de decisão de escolha de modelo. Enumeramos os seguintes parâmetros que podem ser ajustados:

- f_s : a variação da taxa de amostragem do sinal na saída do filtro casado influencia no resultado dos estimadores envolvidos no processo de regressão. Além disto, controla o esforço computacional requerido, devido ao aumento de tamanho das matrizes envolvidas no algoritmo.
- ρ : representa o quantil da f.d.p. definida na Eq. (6.34). O aumento do valor deste parâmetro torna o sistema mais rigoroso na decisão por admitir a existência de um novo multipercurso.
- P : equivale ao número de PNs capturadas em sequência pelo osciloscópio, ou, em outras palavras, ao número de realizações da resposta do canal, armazenadas no vetor \mathbf{z} , que são utilizadas como entrada no algoritmo de estimação.

A primeira avaliação do desempenho do algoritmo de estimação é dada através da realização de uma matriz de testes. Para o preenchimento desta matriz, o RVM é aplicado aos dados colhidos nos locais de medição, em diferentes configurações dos parâmetros acima descritos. O objetivo destes testes é mensurar a eficiência de estimação do RVM para respostas do canal com diferentes valores de relação sinal-ruído.

Para a realização destes testes, foi feito uso de medições específicas. Estas consistiam da captura, no osciloscópio, de quinze sequências PN em cada um dos doze locais avaliados. Com este procedimento, torna-se possível avaliar como o número de realizações de \mathbf{z} , fornecidas como dados de entrada ao algoritmo, influencia no processo de detecção.

Conforme mencionado, o ambiente que serviu como cenário das medições ocasionou a geração de perfis com baixos níveis de ruído. O interesse em avaliar o desempenho do algoritmo em condições em que existe reduzida relação sinal-ruído (RSR) causa a necessidade de se gerar respostas do canal com diferentes níveis de RSR. A geração destas respostas, com tais características, foi obtida da seguinte maneira: aos sinais coletados

na campanha de medições, foi adicionado ruído aleatório com distribuição gaussiana e com determinado nível de potência, de forma que a resposta ao impulso tivesse a RSR desejada.

Por fim, torna-se necessário adotar um critério de qualidade de detecção, a fim de que os resultados obtidos possam ser mensurados e comparados. Para isto, devem-se identificar quais tipos de erro de detecção podem ocorrer. São eles: (a) erro de inserção (e_i): é a detecção errônea de uma componente não existente; (b) erro de deleção (e_d): é a não detecção de uma componente existente. Quando o algoritmo detecta corretamente uma componente existente, este é chamado um acerto. Conhecendo os tipos de erro, é conveniente a adoção de um critério que combine a ocorrência destes eventos e forneça uma medida da qualidade de detecção do algoritmo. A chamada “taxa de detecção de multipercursos” [60] é dada pela Eq. (7.1):

$$TDM = \frac{\text{no. acertos}}{L + e_i}, \quad (7.1)$$

onde L representa o número de multipercursos verdadeiros existentes em uma medição da resposta do canal. Note que, quando não ocorrem erros de inserção e o número de acertos é L , $TDM = 1$. Os erros de deleção incorrem na redução do número de acertos e, conseqüentemente, $TDM < 1$.

A verificação do número de erros de detecção, para perfis com diferentes níveis de ruído, depende da adoção de uma referência. Esta referência é extraída do resultado da aplicação do RVM no mesmo ponto de medição e sendo executado sob determinada configuração de seus parâmetros, na qual se considera que todos os multipercursos existentes no perfil são corretamente detectados. Isto torna os resultados apresentados relativos àqueles sob a configuração de referência.

Naturalmente, as RICs que não tiveram adição de ruído gaussiano devem ser referências, visto que estas oferecem condições facilitadas de detecção. Agora, deve-se determinar a configuração do RVM que servirá de configuração padrão para gerar as saídas de referência. O trabalho desenvolvido em [60] procurou avaliar a influência dos parâmetros do algoritmo na qualidade de detecção, porém trabalhou-se apenas com canais simulados,

onde os parâmetros do modelo, como número e posição dos multipercursos eram conhecidos *a priori*. Levando em consideração os resultados apresentados em [60], o aumento do valor de f_s possibilita aumento de desempenho. De forma semelhante, o aumento de P também proporciona melhoras. O parâmetro ρ age da seguinte forma: quanto mais próximo de 1, mais dificilmente o algoritmo irá gerar um erro de inserção, porém a probabilidade de erros de deleção é aumentada. Valores moderados de ρ variam entre 0,9 e 0,999. Decidiu-se assim, pela seguinte configuração do RVM para a geração de resultados de referência: $f_s = 12,5$ amostras/bit, $P = 10$ e $\rho = 0,9$, sendo o algoritmo aplicado a RICs sem adição de ruído.

Em seguida, determinam-se os testes por realizar. Serão avaliados RICs sem adição de ruído, e com valores de SNR iguais a 30, 25 e 20 dB. Quanto aos parâmetros do algoritmo, a amostragem f_s será feita nos valores de 12,5 e 5 amostras por bit; a P serão atribuídos os valores, 1, 3 e 10; por fim, ρ poderá assumir os valores 0,9 e 0,99.

A matriz gerada para avaliar a influência da variação dos parâmetros acima citados totaliza 48 testes para cada um dos doze pontos de medição. A extensa fase de testes culminou no conteúdo da tabela 7.1, que possui todos os resultados obtidos. A tabela encontra-se estruturada da seguinte forma: as quatro colunas da esquerda identificam a configuração utilizada dos parâmetros do algoritmo. As doze colunas que representam os pontos de medição analisados são preenchidas com os dados: número total de multipercursos detectados, número de erros de deleção e número de erros de inserção.

A partir dos dados da matriz de testes, podem ser desenvolvidas análises que facilitem a geração de conclusões sobre o desempenho do RVM. Os gráficos da Fig. 7.1, 7.2 e 7.3 contêm funções de distribuição cumulativas (f.d.c.) do TDM obtidas para os 12 locais de medição e visam facilitar a avaliação da influência da variação de cada um dos parâmetros do algoritmo.

A partir do conjunto de gráficos apresentado na Fig. 7.1, é possível verificar uma sensível melhora de detecção para $\rho = 0,90$. Este comportamento apresenta-se constante para os diferentes valores de RSR avaliados. Já os gráficos apresentados na Fig. 7.2 mostram uma melhora de detecção mais notável diante do aumento de f_s . De maneira

Tabela 7.1: Resultados da avaliação de desempenho do algoritmo. A primeira linha de dados corresponde à configuração de referência.

	f_s	P	ρ	Ponto 1		Ponto 2		Ponto 3		Ponto 4		Ponto 5		Ponto 6		Ponto 7		Ponto 8		Ponto 9		Ponto 10		Ponto 11		Ponto 12																						
				L	erros	L	erros	L	erros	L	erros	L	erros	L	erros																																	
				e_d	e_i	e_d	e_i	e_d	e_i	e_d	e_i	e_d	e_i	e_d	e_i																																	
Inf	12,5	30	0,9	8	X	X	9	X	X	7	X	X	14	X	X	15	X	X	11	X	X	11	X	X	15	X	X	8	X	X	10	X	X	10	X	X	10	X	X									
			0,99	8	0	0	7	2	0	7	0	0	14	0	0	14	0	0	14	0	0	11	0	0	11	0	0	14	0	0	8	0	0	10	0	0	7	3	0	9	1	0						
			0,9	8	0	0	9	0	0	8	0	0	14	0	0	14	0	0	14	0	0	10	0	0	10	0	0	14	0	0	8	0	0	11	0	0	10	0	0	9	1	0	0					
			0,99	8	0	0	7	2	0	7	0	0	12	2	0	14	1	0	10	1	0	10	1	0	10	1	0	14	1	0	7	1	0	11	0	1	7	3	0	9	1	0	0					
			0,9	9	0	1	7	2	0	7	0	0	11	3	0	12	3	0	9	2	0	11	0	0	11	0	0	15	1	1	7	1	0	11	0	1	7	3	0	8	2	0	0					
			0,99	7	1	0	5	4	0	6	1	0	8	6	0	12	3	0	7	4	0	11	0	0	11	0	0	11	4	0	7	1	0	9	1	0	6	4	0	7	3	0	0	0				
			0,9	8	0	0	7	2	0	7	0	0	12	2	0	14	2	0	10	1	0	10	1	0	10	1	0	12	2	0	10	0	2	10	1	1	8	3	1	9	1	0	0	0				
			0,99	8	0	0	6	4	1	6	1	0	12	3	1	15	1	1	10	1	0	9	3	1	14	1	0	8	2	2	11	1	2	9	3	2	9	3	2	9	1	0	0	0				
			0,9	7	2	1	7	3	1	6	3	2	10	5	1	13	3	1	10	1	0	10	4	3	13	3	1	7	3	2	11	1	2	6	4	0	9	1	0	0	0	0	0	0				
			0,99	6	2	0	4	5	0	3	4	0	8	6	0	11	5	1	6	5	0	8	4	1	11	4	0	8	4	1	11	4	0	5	4	1	8	2	0	5	5	0	6	4	0	0		
25	20	3	0,9	11	0	3	9	2	2	7	1	1	12	2	0	14	2	1	11	0	0	9	4	2	15	1	1	7	3	2	10	1	1	9	3	2	11	1	2	0	0	0	0	0				
			0,99	7	1	0	7	2	0	6	1	0	10	4	0	12	3	0	9	2	0	7	4	0	14	1	0	5	3	0	8	3	1	6	4	0	7	3	0	0	0	0	0	0				
			0,9	9	0	1	9	2	2	7	1	1	9	5	0	13	3	1	7	4	0	8	4	1	13	3	1	8	2	2	7	4	1	9	3	2	9	2	1	0	0	0	0	0	0	0	0	
			0,99	8	0	0	7	2	0	5	2	0	8	6	0	12	4	1	6	5	0	6	5	0	11	4	0	7	3	2	7	4	1	6	5	1	6	5	1	6	4	0	0	0	0	0		
			0,9	10	1	3	8	3	2	7	1	1	8	6	0	10	5	0	6	6	1	9	5	3	13	4	2	5	4	1	8	3	1	6	5	1	6	5	1	6	5	1	6	5	1	0	0	
			0,99	5	3	0	5	5	1	2	5	0	5	9	0	8	7	0	5	6	0	6	5	0	10	4	0	4	4	0	6	4	0	5	5	0	5	5	0	5	5	0	5	5	0	0	0	
			0,9	8	1	1	7	3	1	3	4	0	9	5	0	9	6	0	7	5	1	7	4	0	11	4	0	5	4	1	8	2	0	5	5	0	5	5	0	6	5	0	5	5	0	0	0	
			0,99	9	2	3	7	3	1	4	4	1	10	7	3	10	5	0	7	5	1	8	4	1	15	3	3	4	4	0	10	3	3	8	5	3	8	5	3	7	5	2	0	0	0	0		
			0,9	5	3	0	4	5	0	2	5	0	8	8	2	9	6	0	5	6	0	6	5	0	11	4	0	4	4	0	6	4	0	6	6	2	6	5	1	8	4	2	0	0	0	0		
			0,99	4	4	0	5	4	0	3	4	0	9	8	3	9	7	1	6	6	1	9	4	2	13	3	1	4	5	1	9	3	2	5	6	1	8	4	2	0	0	0	0	0	0	0		
Inf	5	10	0,9	8	0	0	6	3	0	7	0	0	10	4	0	11	4	0	9	2	0	9	2	0	12	3	0	7	2	1	8	2	0	5	5	0	8	2	0	0	0	0	0	0	0			
			0,99	7	1	0	6	3	0	7	0	0	9	5	0	11	4	0	9	2	0	9	2	0	12	3	0	6	3	1	8	2	0	5	5	0	8	2	0	0	0	0	0	0	0	0	0	
			0,9	7	1	0	6	3	0	7	0	0	9	5	0	11	4	0	9	2	0	9	2	0	13	2	0	6	3	1	8	2	0	5	5	0	8	2	0	0	0	0	0	0	0	0	0	0
			0,99	7	1	0	6	3	0	6	1	0	8	6	0	11	4	0	9	2	0	7	4	0	11	4	0	5	4	1	8	2	0	5	5	0	6	4	0	0	0	0	0	0	0	0	0	0
			0,9	7	1	0	5	4	0	6	1	0	8	6	0	10	5	0	7	4	0	8	3	0	12	3	0	4	4	0	8	2	0	5	5	0	6	4	0	0	0	0	0	0	0	0	0	0
			0,99	5	3	0	4	5	0	4	3	0	7	7	0	9	6	0	6	5	0	7	4	0	11	4	0	4	4	0	7	3	0	3	7	0	3	7	0	5	5	0	5	5	0	0	0	
			0,9	8	0	0	7	2	0	5	2	0	8	6	0	10	5	0	8	3	0	8	3	0	12	3	0	4	4	0	8	2	0	4	6	0	4	6	0	6	4	0	0	0	0	0	0	
			0,99	7	1	0	6	3	0	5	2	0	8	6	0	10	5	0	8	3	0	7	4	0	11	4	0	4	4	0	6	4	0	3	7	0	3	7	0	6	4	0	0	0	0	0	0	
			0,9	6	2	0	5	4	0	6	2	1	8	6	0	10	5	0	8	3	0	7	4	0	11	4	0	6	4	2	8	3	1	6	5	0	6	4	0	2	8	0	5	5	0	0	0	
			0,99	6	2	0	5	4	0	3	4	0	6	8	0	9	6	0	6	5	0	7	4	0	11	4	0	8	3	0	5	5	2	6	4	0	2	8	0	4	6	0	4	6	0	0	0	
25	20	1	0,9	5	3	0	5	4	0	5	2	0	8	6	0	11	5	1	6	5	0	9	3	1	11	4	0	5	3	0	7	3	0	3	7	0	5	6	1	0	0	0	0	0	0	0		
			0,99	4	4	0	2	7	0	3	4	0	6	8	0	8	7	0	5	6	0	6	5	0	11	4	0	3	5	0	5	5	0	2	8	0	4	6	0	0	0	0	0	0	0	0		
			0,9	8	0	0	4	5	0	3	4	0	9	5	0	9	6	0	8	3	0	7	4	0	11	4	0	4	4	0	4	4	0	3	0	3	0	3	0	2	8	0	4	6	0	0	0	0
			0,99	5	3	0	3	6	0	3	4	0	8	6	0	9	6	0	9	6	0	7	4	0	11	4	0	3	5	0	5	5	0	2	8	0	2	8	0	4	6	0	0	0	0	0	0	0
			0,9	8	0	0	2	7	0	2	5	0	4	0	6	9	0	8	3	0	7	4	0	11	4	0	4	4	0	4	4	0	3	0	3	0	3	0	2	8	0	7	4	1	0	0	0	
			0,99	5	3	0	3	6	0	3	4	0	8	6	0	9	6	0	6	5	0	7	4	0	11	4	0	3	5	0	6	4	0	2	8	0	6	4	0									

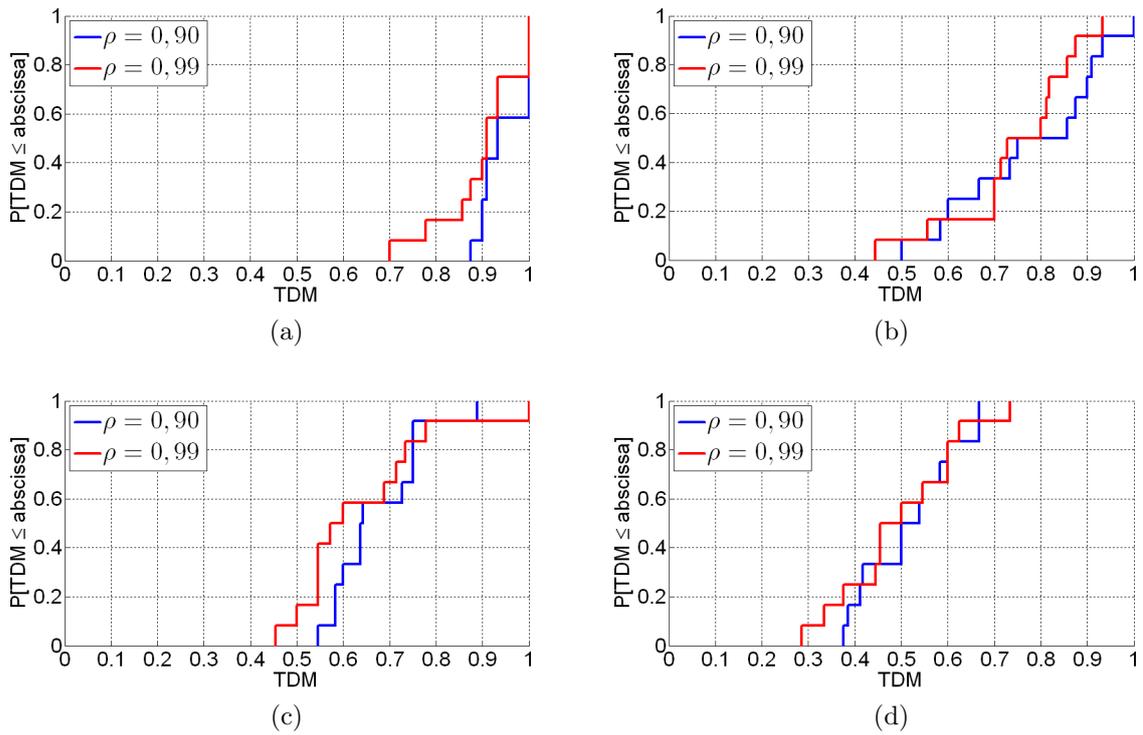


Figura 7.1: Função distribuição cumulativa do TDM para dois valores de ρ . (a) sem adição de ruído (RSR > 30), (b) RSR = 30 dB, (c) RSR = 25 dB, (d) RSR = 20 dB.

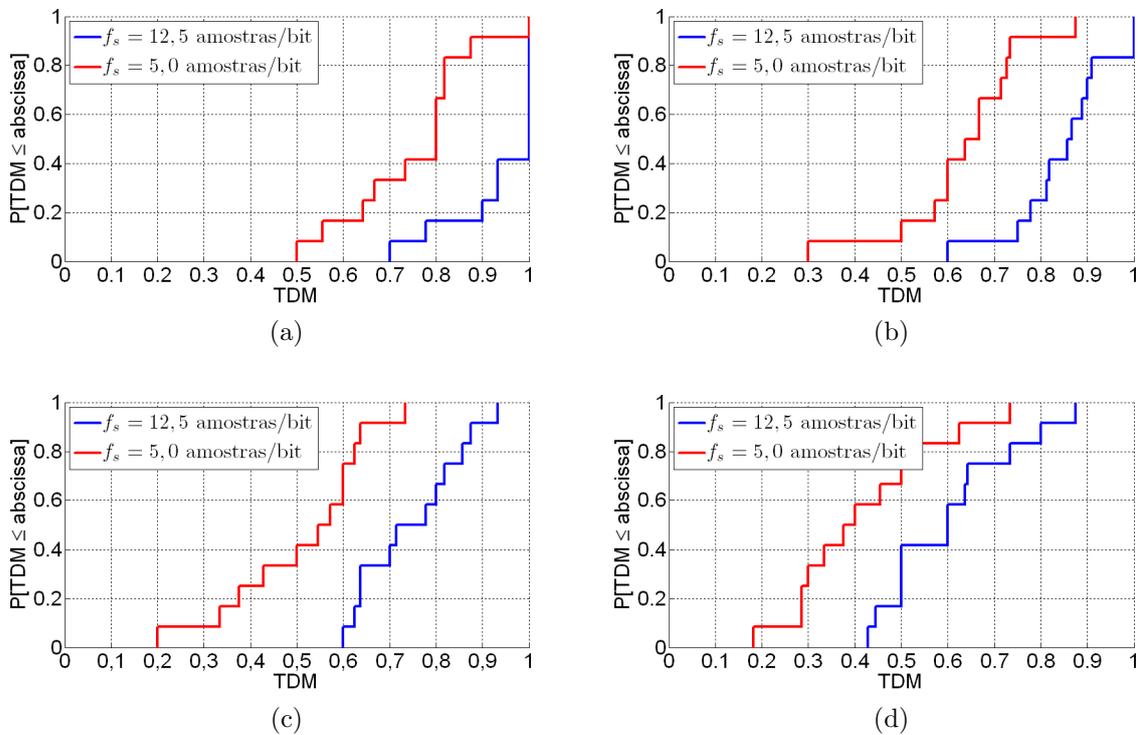


Figura 7.2: Função distribuição cumulativa do TDM para dois valores de f_s . (a) sem adição de ruído (RSR > 30), (b) RSR = 30 dB, (c) RSR = 25 dB, (d) RSR = 20 dB.

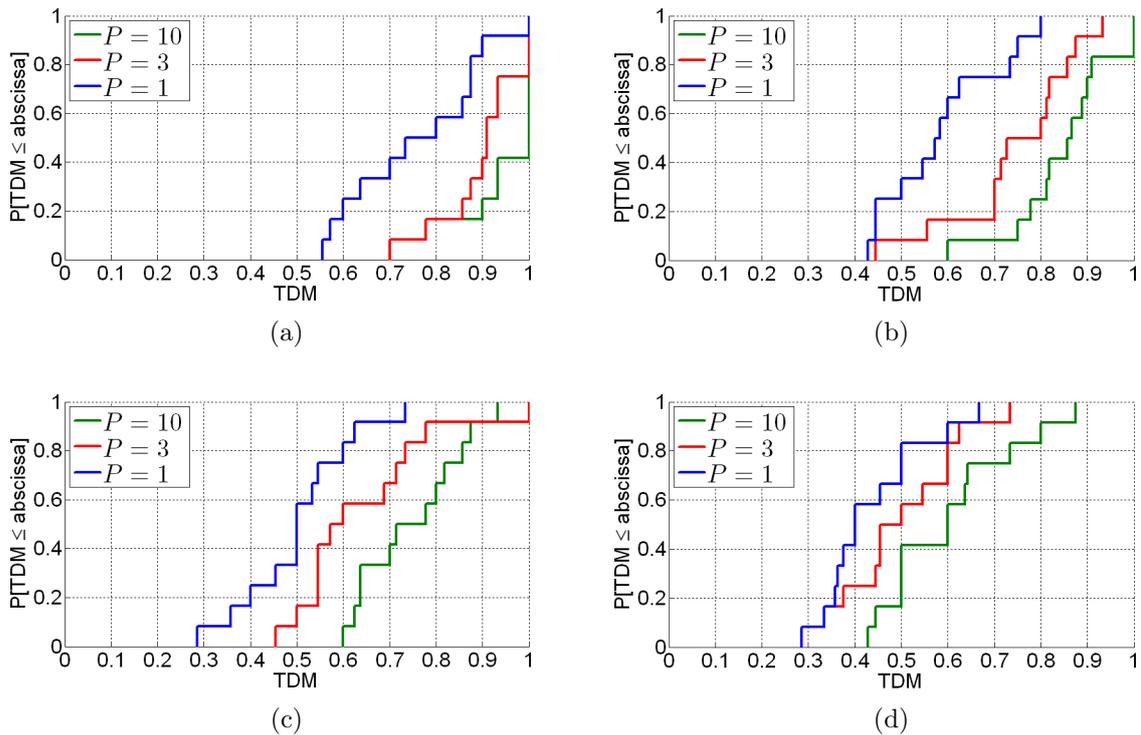


Figura 7.3: Função distribuição cumulativa do TDM para três valores de P . (a) sem adição de ruído ($\text{RSR} > 30$ dB), (b) $\text{RSR} = 30$ dB, (c) $\text{RSR} = 25$ dB, (d) $\text{RSR} = 20$ dB.

semelhante, o aumento do número de P proporciona melhora na taxa de detecção, como mostram os gráficos da Fig. 7.3.

Dos parâmetros avaliados, o número de realizações da resposta do canal apresenta particular importância, visto que é o único que precisa ser dimensionado para a realização das medições de sondagem. Apresentada a importância deste parâmetro, mostra-se relevante a realização de um estudo mais aprofundado sobre seu comportamento. O método aplicado consiste em determinar o número mínimo de realizações de \mathbf{z} necessário para que se obtenha um nível de convergência satisfatório das RICs, geradas a partir das saídas do algoritmo. A função de convergência aplicada é definida pela Eq. (7.2):

$$\text{Conv}(p) = \left(h_{\text{ref}}(\tau) - h_p(\tau) \right)^2, \quad (7.2)$$

onde $p \mapsto \text{Conv}(p)$ é a função de convergência, $\tau \mapsto h_{\text{ref}}(\tau)$ é a função resposta ao impulso do canal, gerada com a aplicação do RVM em sua configuração de referência, e $\tau \mapsto h_p(\tau)$ é a função resposta ao impulso do canal, resultado da aplicação do RVM, executado com

p realizações da resposta do canal.

A Fig. 7.4 mostra o gráfico da função de convergência em função do número de realizações fornecidas ao algoritmo. Verifica-se que, a partir de $p = 9$, o nível de convergência decai de forma mais lenta. Este é, portanto, um valor para o número de realizações da resposta do canal que apresenta uma vantajosa relação entre o tempo de captura na fase de sondagem e a eficácia do algoritmo.

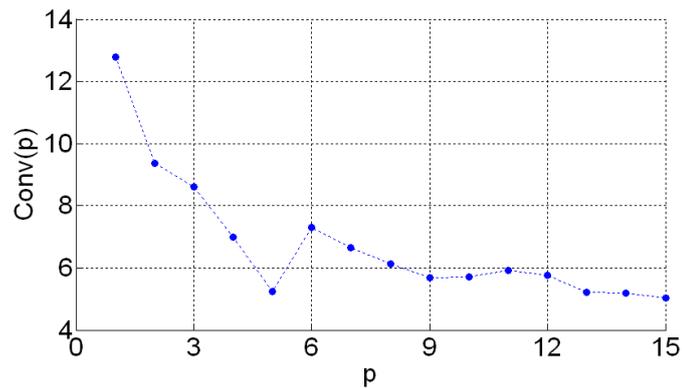


Figura 7.4: Gráfico da função de convergência em função do número de realizações da resposta do canal. Configuração utilizada: RSR = 20 dB, $\rho = 0,90$, $f_s = 12,5$ amostras/bit.

7.2 Modelagem estatística do canal

A aplicação do algoritmo RVM à massa de dados colhida, durante a fase de sondagem, possibilita a realização de estudos sobre a caracterização faixa larga do canal rádio.

As medições que serviram de base para este estudo obedeceram a uma metodologia específica. Os locais de medição escolhidos são idênticos àqueles apresentados no capítulo 5, sendo que, fez-se uso de uma grade com 36 posições, como ilustra a Fig. 7.5. Assim, em cada local, a antena receptora deveria ser posicionada sobre cada uma das 36 posições para realizar a captura do sinal recebido, sendo que cada captura consistia de apenas 1 realização da resposta do canal, ou seja, $P = 1$.

O estudo da modelagem do canal rádio terá como ponto de partida o levantamento do espalhamento de retardo, calculado para apenas uma posição da grade de medições, em cada um dos doze locais avaliados. A Fig. 7.6 apresenta os gráficos do espalhamento

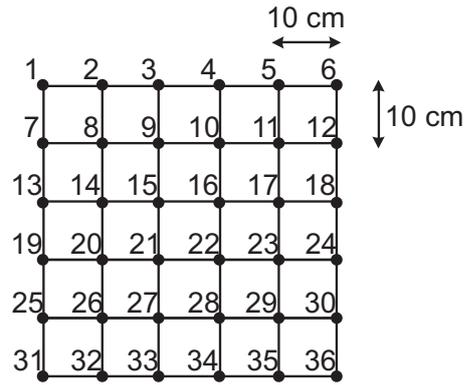


Figura 7.5: Grade de medições com 36 posições.

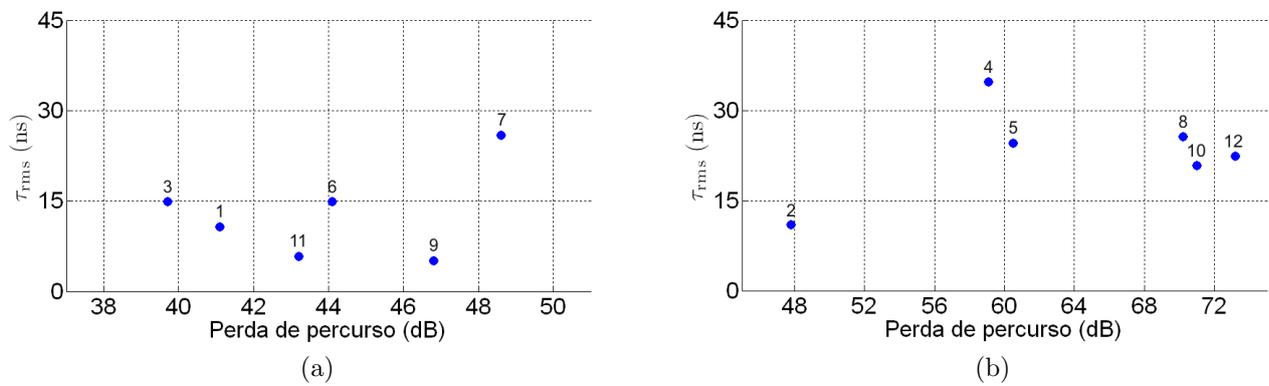


Figura 7.6: Gráfico de dispersão do espalhamento de retardo calculado na posição 15 da grade, para os 12 locais de medição. A numeração dos pontos corresponde ao local de medição especificado na planta da Fig. 5.7. (a)LOS e (b)NLOS.

de retardo considerando-se o elemento receptor fixo na posição 15 da grade. Os resultados apresentados estão divididos em dois casos, (a) com os sistemas transmissor e receptor em visada e (b) com os sistemas fora de visada. Verificam-se, em média, valores menores de espalhamento de retardo na situação em LOS, quando comparado aos valores obtidos em NLOS.

O segundo passo, no sentido da caracterização do ambiente, foi realizado através da especificação do espalhamento de retardo do canal para uma região com área aproximada $A = 0,36 \text{ m}^2$, que corresponde à área da grade de medições utilizada. O interesse deste cálculo é avaliar como varia o τ_{rms} para um elemento receptor, cuja posição pode variar em uma região de área A . A Fig. 7.7 exhibe os gráficos de dispersão da média do espalhamento de retardo e os intervalos de confiança associados, calculados a partir das 36 posições da grade, para os 12 locais de medição. Estes resultados exprimem que o valor médio

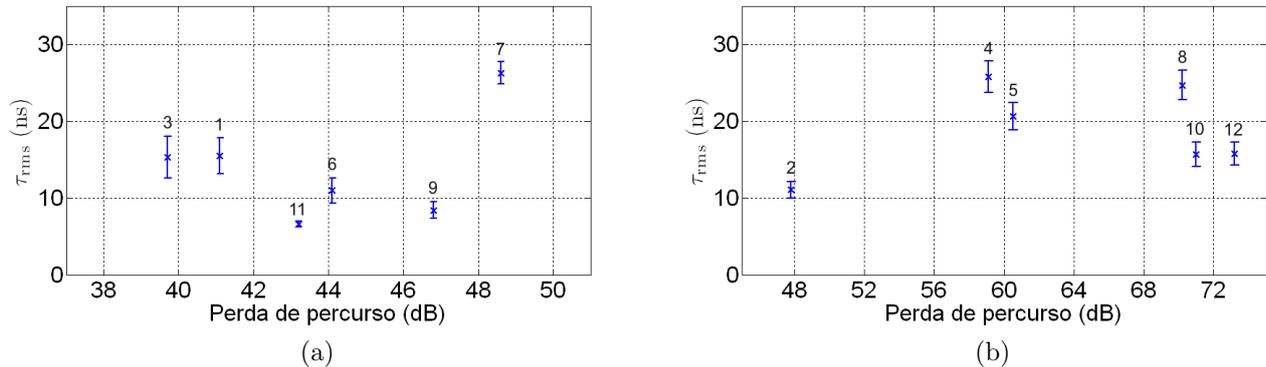


Figura 7.7: Gráfico de dispersão da média do espalhamento de retardo com 95% de nível de confiança para os 12 locais medidos. (a)LOS e (b)NLOS.

do espalhamento de retardo para amostras aleatórias, tomadas no mesmo local e numa região de área da ordem de A , possui probabilidade igual a 95% de estar contida nos limites especificados pelo intervalo de confiança. Como no caso anterior, os resultados encontram-se divididos em (a) LOS e (b) NLOS.

De forma geral, são verificados valores de espalhamento de retardo que variam entre 5 e 40 ns. Estes resultados estão em concordância com os obtidos em outros trabalhos onde foi realizada sondagem de ambientes *indoor* [39, 40, 63]. As Figs. 7.6 e 7.7 exibem pouca correlação entre os valores de τ_{rms} e a perda de percurso. Outros trabalhos obtiveram resultados semelhantes [52, 63], porém contrapõem-se aos obtidos em [39, 7].

É verificado, ainda, que os valores da média do espalhamento de retardo são menores para o caso de visada direta, quando comparado aos casos sem visada, de uma forma geral. Tal constatação também foi observada em [40, 64]. Quanto à amplitude do intervalo de confiança, não é notada diferença relevante entre os dois casos.

A modelagem do canal faixa larga é complementada com a análise da dispersividade temporal sobre todos os locais de medição conjuntamente, com o objetivo de apresentar resultados que representem de forma genérica o ambiente onde foram realizadas as medições.

A Fig. 7.8 apresenta histogramas dos valores de espalhamento de retardo obtidos em todas as medições na grade. Na situação em LOS, verifica-se que a maior concentração dos valores de espalhamento de retardo ocorre para valores inferiores a 15 ns, enquanto no

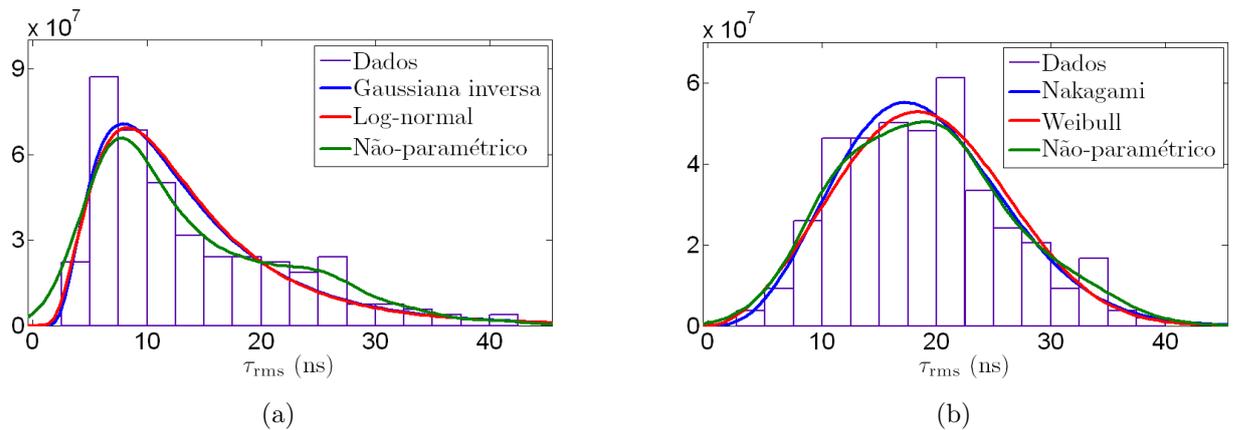


Figura 7.8: Histograma do espalhamento de retardo e curvas de ajuste de distribuições de probabilidade. (a)LOS e (b)NLOS.

caso NLOS, verifica-se maior uniformidade na distribuição dos valores de espalhamento de retardo, na faixa entre 10 e 25 ns. Com a finalidade de modelar o comportamento de τ_{rms} , foi realizado o ajuste de distribuições de probabilidade sobre os histogramas. As Figs. 7.8a e 7.8b exibem f.d.p.s selecionadas no teste de hipótese de Kolmogorov-Smirnov [65] e que representam o comportamento de τ_{rms} para os casos LOS e NLOS. As tabelas 7.2 e 7.3 fornecem os valores que especificam as f.d.p.s de melhor ajuste.

Identificam-se μ como a média e λ o parâmetro de forma da distribuição gaussiana inversa. Os parâmetros da distribuição lognormal, μ e σ , são, respectivamente, a média e o desvio padrão da variável aleatória $X = \ln(\tau_{\text{rms}})$, que, por definição, tem distribuição normal. Na distribuição Weibull, λ é o parâmetro de escala e k , o parâmetro de forma. A f.d.p Nakagami é apresentada na Eq. (7.3) e os seus parâmetros, nas Eqs. (7.4).

$$f(x|\mu, \omega) = 2 \left(\frac{\mu}{\omega} \right)^{\mu} \frac{1}{\Gamma(\mu)} x^{(2\mu-1)} e^{-\frac{\mu}{\omega} x^2} \quad (7.3)$$

$$\mu = \frac{(E[X^2])^2}{\text{Var}[X^2]} \quad (7.4a)$$

$$\omega = E[X^2] \quad (7.4b)$$

O melhor ajuste é alcançado pela distribuição não-paramétrica, implementada no

MATLAB[®] pela função *ksdensity*. Esta distribuição é alcançada através do uso de estimadores de densidade baseados em núcleo [66]. O estimador \hat{f} de uma variável unidimensional contínua baseada no vetor amostra $\{x_1, \dots, x_n\}$, avaliado no ponto x , tem a forma apresentada na Eq. (7.5).

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n K(x - x_i; h), \quad (7.5)$$

onde K é a função núcleo, representada por uma f.d.p. simétrica, cuja variância é controlada pelo parâmetro h , chamado de largura de banda.

Existem trabalhos que apontam a modelagem do espalhamento de retardo por f.d.p.s como distribuições normais [39] ou log-normais [67], generalizados para ambientes com ou sem visada.

Tabela 7.2: Especificações das distribuições na modelagem de τ_{rms} em LOS.

Gaussiana inversa	μ $1,38 \times 10^{-8}$	λ $3,50 \times 10^{-8}$
Lognormal	μ -18,27	σ 0,59
Não paramétrico	Núcleo normal	h $2,57 \times 10^{-9}$

Tabela 7.3: Especificações das distribuições na modelagem de τ_{rms} em NLOS.

Weibull	λ $2,13 \times 10^{-8}$	k 2,85
Nakagami	μ 1,84	ω $4,10 \times 10^{-16}$
Não paramétrico	Núcleo normal	h $2,68 \times 10^{-9}$

Um parâmetro importante, e cujo comportamento foi estudado neste trabalho, é o tempo de interchegada de multipercursos. Ele se refere ao período de tempo decorrido entre a chegada de dois multipercursos consecutivos. A Fig. 7.9 exhibe gráficos com histogramas do tempo de interchegada de multipercursos, separados nos casos (a) LOS e (b) NLOS. Verificam-se comportamentos semelhantes do tempo de interchegada para os casos LOS e NLOS. Ambos os histogramas apresentam um pico inicial e, em seguida, queda monotônica. Resultados semelhantes foram obtidos em [68]. Distribuições não paramétricas foram utilizadas na modelagem dos histogramas, e suas especificações estão

descritas na tabela 7.4.

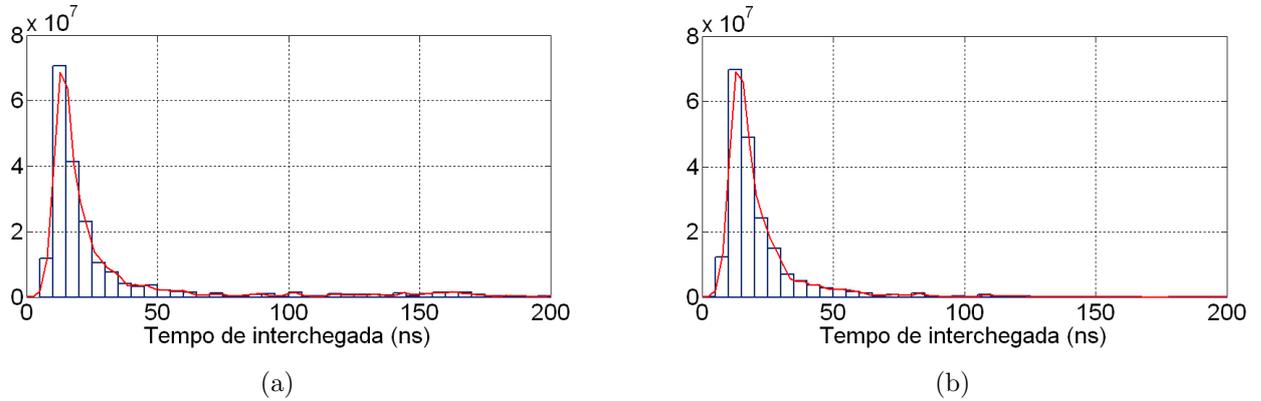


Figura 7.9: Histograma do tempo de interchegada de multipercursos e curvas de ajuste de distribuições de probabilidade. (a)LOS e (b)NLOS.

Tabela 7.4: Especificações da distribuição não paramétrica na modelagem do tempo de interchegada nos casos LOS e NLOS.

Não paramétrico	Núcleo	h
LOS	normal	$1,82 \times 10^{-9}$
NLOS	normal	$1,72 \times 10^{-9}$

Outro parâmetro relevante para análise estatística é a fração da energia total concentrada na componente de multipercurso mais forte. Ele fornece informações sobre a predominância de uma componente na resposta do canal, para um dado ambiente. Nos gráficos da Fig. 7.10 são exibidos histogramas dos valores da fração de energia nos casos LOS e NLOS. Verifica-se uma diferença de comportamento marcante entre os dois gráficos. No caso LOS, tem-se uma concentração expressiva entre 0,2 e 0,6 da energia total na componente mais forte. Enquanto isso, no caso NLOS, a maior concentração está em valores inferiores a 0,3. As distribuições não paramétricas utilizadas são especificadas na tabela 7.5. Testes de natureza semelhante foram realizados em [69], e resultados apontaram a concentração de energia em menos componentes de multipercurso no caso LOS.

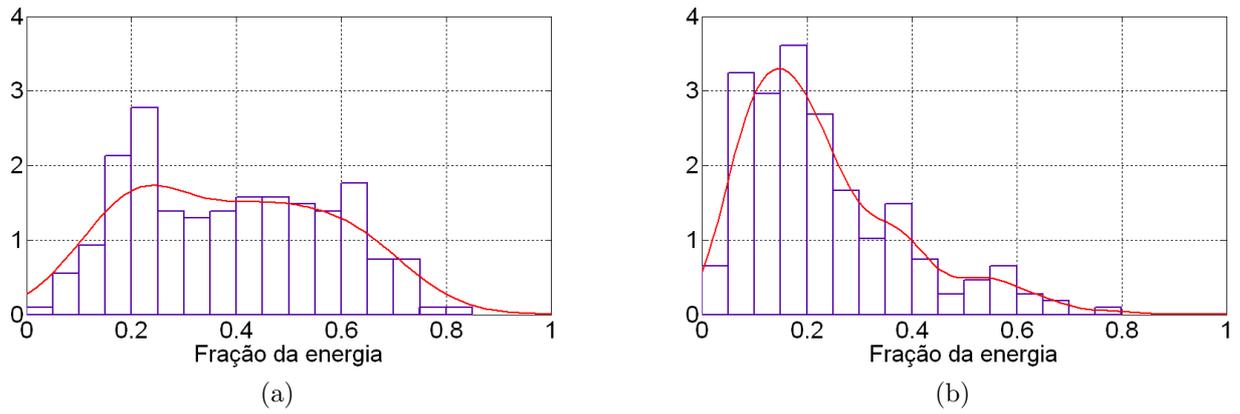


Figura 7.10: Histograma da fração da energia total concentrada na componente de multipercurso mais forte e curvas de ajuste de distribuições de probabilidade. (a)LOS e (b)NLOS.

Tabela 7.5: Especificações da distribuição não paramétrica na modelagem do tempo de interchegada nos casos LOS e NLOS.

Não paramétrico	Núcleo	h
LOS	normal	$8,66 \times 10^{-2}$
NLOS	normal	$4,44 \times 10^{-2}$

Capítulo 8

Conclusões e Trabalhos Futuros

8.1 Conclusões

Esta dissertação abordou o estudo do canal rádio multipercurso faixa-larga. Tendo em vista que, nos tempos modernos, o conhecimento preciso do canal rádio tornou-se imprescindível para o desenvolvimento de sistemas de comunicação sem fio de alta eficiência, os esforços realizados neste trabalho foram direcionados no sentido de oferecer contribuições para a modelagem destes ambientes.

O estudo desenvolvido visou aliar os conhecimentos científicos consolidados na caracterização de canais aleatórios e nos métodos de sondagem da resposta faixa-larga do canal aos desenvolvimentos recentes da Estatística Bayesiana na solução dos problemas de regressão. Através desta combinação, foi apresentada a implementação de uma técnica empregada na detecção de multipercursos em um canal de propagação real. Para o cumprimento deste objetivo, a elaboração da dissertação pôde ser dividida em três fases. A fase de planejamento, construção do aparato de medição e coleta dos dados experimentais do ambiente; a fase de estudo, implementação e validação do algoritmo de estimação; e por fim, a fase de aplicação do método, compilação e análise dos resultados.

Os capítulos iniciais deste trabalho trataram de fundamentos teóricos. Neles, foram apresentadas: 1) a caracterização dos canais faixa-larga, através da apresentação das suas funções de sistema, e 2) as características básicas e as premissas da Estatística Bayesiana, seguida pela sua aplicação na modelagem e na regressão. A seguir, foi descrita a

implementação em *hardware* de um sistema de sondagem no domínio do tempo, que foi projetado para realizar medições no ambiente *indoor* do prédio de Metrologia Elétrica do Inmetro, que abriga diversos laboratórios e salas de escritório.

O algoritmo de estimação RVM, cujo princípio de funcionamento foi detalhadamente descrito, demonstrou compor um mecanismo eficaz na identificação de parâmetros como o número de multipercursos, suas amplitudes complexas e instantes de chegada, que compõem o modelo de melhor ajuste para representar a resposta impulsiva do canal.

A definição dos valores dos parâmetros ajustáveis do RVM mostrou desempenhar papel importante na qualidade de detecção. Os estudos apresentados apontaram o grau de influência causado pela variação destes parâmetros. Além disto, foi possível verificar o nível de redução da qualidade de detecção para a queda da RSR até 20 dB. Este fato sugere que o algoritmo pode apresentar comportamento comprometido em ambientes onde a RSR seja inferior a 25 dB.

A utilização dos dados produzidos pelo RVM para gerar a modelagem estatística do canal constituiu o objetivo final deste trabalho. O comportamento de elementos representativos da dispersão temporal dos multipercursos foi analisada, como o espalhamento de retardo (para o caso pontual e para pequenas áreas), o tempo de interchegada de multipercursos e a distribuição de energia entre as componentes de multipercurso. Este estudo evidenciou comportamentos característicos de ambientes confinados. Os locais com visada direta entre as antenas transmissora e receptora, quando comparados àqueles sem visada, apresentaram comportamento diferenciado, como menores valores de espalhamento de retardo e parcela considerável da potência do sinal concentrada na componente de multipercurso mais forte. Distribuições de probabilidade foram especificadas para representar o comportamento estatístico destes parâmetros.

O conjunto dos dados obtidos e, a partir destes, as análises que puderam ser traçadas, são indicadores de que o algoritmo RVM de estimação da resposta do canal é uma técnica que realiza, de forma satisfatória, a detecção dos multipercursos válidos em meio ao sinal ruidoso, que é capturado do sistema de sondagem. O mecanismo de inferência bayesiano e as estratégias para exclusão das funções base irrelevantes implicam na neces-

cidade de desenvolvimento de um código computacional elaborado. Por outro lado, a não necessidade de adoção de limiar mínimo de detecção, comum em outros métodos, figura como uma de suas vantagens. Além disto, sua robustez torna o mecanismo confiável e com desempenho bastante satisfatório.

8.2 Sugestões para Trabalhos Futuros

Partindo das realizações alcançadas neste trabalho, é possível identificar alguns tópicos que podem ser explorados.

- A utilização da sonda construída para a realização de medições em ambientes *indoor* que sejam estruturalmente diferentes do ambiente que serviu de cenário de medição neste trabalho. Tal estudo fornecerá dados concretos sobre a variação da dispersividade do canal em função das características estruturais do edifício, a distribuição e a proximidade dos possíveis espalhadores do sinal.
- A realização de medições, utilizando o mesmo *setup*, que explorem a transição *outdoor-indoor* do sinal propagado.
- Realização de medições em deslocamento, a fim de estudar o comportamento do algoritmo em um canal rádio-móvel.
- A utilização do banco de dados, gerado a partir das medições, para realizar a comparação do RVM com outras técnicas de estimação dos parâmetros da resposta do canal, como SAGE, MUSIC, etc.

Referências Bibliográficas

- [1] TSE, D.; VISWANATH, P. *Fundamentals of wireless communication*. Cambridge: Cambridge University Press, 2005. 586 p.
- [2] BAILEY, D. More mobile phone users in 2010. *United Nations*. fev. 2010. Disponível em: <http://www.unmultimedia.org/radio/english/detail/90889.html>. Acesso em: 06 ago 2010.
- [3] GOLDSMITH, A. *Wireless communications*. Cambridge: Cambridge University Press, 2005. 672 p.
- [4] TURIN, G. L. Communication through noisy, random-multipath channels. *IRE Convention Record*. v. 4, p. 154-166, mar. 1956.
- [5] SALEH, M.; VALENZUELA, R. A. A statistical model for indoor multipath propagation. *IEEE Journal on Selected Areas in Communications*. v. SAC-5, n. 2, p. 128-137, fev. 1987.
- [6] RAPPAPORT, T. S.; SEIDEL, S. Y. Statistical channel impulse response models for factory and open plan building radio communication system design. *IEEE Transactions on Communications*. v. 39, n. 5, p. 794-807, maio 1991.
- [7] DEVASIRVATHAM, D. M. J. A comparison of time delay spread and signal level measurements within two dissimilar office buildings. *IEEE Transactions on Antennas and Propagation*. v. AP-35, n. 3, p. 319-324, mar. 1987.
- [8] HASHEMI, H. The indoor radio propagation channel. *Proceedings of the IEEE*. v. 81, n. 7, p. 943-968, jul. 1993.

- [9] CRAMER, J. M.; SCHOLTZ; WIN, M. Z. Evaluation of an ultra-wideband propagation channel. *IEEE Transactions on Antennas and Propagation*. v. 5, p. 561-570, n. 5, p. 561-570, maio 2002.
- [10] YANO, S. M. Investigating the ultra-wideband indoor wireless channel. *Conference IEEE VTC Spring*, v. 3, n. 55, p. 1200-1204, maio 2002.
- [11] HÖGBOM, J. A. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astron. and Astrophys. Suppl ser.* v. 15, p. 417-426, jun. 1974.
- [12] CRAMER, R. J. -M; WIN M. Z.; SCHOLTZ, R. A. Evaluation of the multipath characteristics of the impulse radio channel. *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*. v. 2, p. 864-868, Boston, sep. 1998.
- [13] CRAMER, R. J. -M; SCHOLTZ, R. A; WIN M. Z. Spatio-temporal diversity in ultra-wideband radio. *Proc. IEEE Wireless Communications and Networking Conference*. v. 2, p. 888-892, New Orleans, 1999.
- [14] FLEURY, B. H. et al. Channel parameter estimation in mobile radio environments using the SAGE algorithm. *IEEE Journal on Selected Areas in Communications*. v. 17, n. 3, p. 434-450, mar. 1999.
- [15] DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Statist. Soc., Ser. B*. vol. 39, n. 1, p. 1-38, 1977.
- [16] LO, T.; LITVA, J.; LEUNG, H. A new approach for estimating indoor radio propagation characteristics. *IEEE Transactions on Antennas and Propagation*. v. 42, n. 10, p. 1369-1376, out. 1994.
- [17] SKLAR, B. Rayleigh fading channels in mobile digital communication systems: part I: characterization. *IEEE Communications Magazine*. v. 35, n. 7, p. 90-100, jul 1997.

- [18] PARSONS, J. D. *The mobile radio propagation channel*. 2 ed. Chichester: John Wiley & Sons, 2000. 436 p.
- [19] BELLO, P. A. Characterization of randomly time-variant linear channels. *IEEE Transactions on Communication Systems*, v. CS-11, p. 360-393, dez. 1963.
- [20] ZADEH, L. A. Frequency analysis of variable networks. *Proceedings I.R.E.*, v. 38, p. 291-299, 1950.
- [21] KAILATH, J. Time-variant communication channels. *IEEE Transactions on Information Theory*, v. 9, n. 4, p. 233-237, out. 1963.
- [22] PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. *Estatística Bayesiana*. Lisboa: Fundação Calouste Gulbenkian, 2003, 447 f.
- [23] BERGER, J. O. *Statistical Decision Theory and Bayesian Analysis*. New York: Springer, 1980, 648 f.
- [24] MACKAY, D. J. C. Probable networks and plausible predictions - a review of practical bayesian methods for supervised neural networks. *Network: Computation in Neural Systems*, v. 6, n.6, p. 469-505, ago. 1995.
- [25] MACKAY, D. J. C. *Information Theory, Inference & Learning Algorithms*. Cambridge: Cambridge University Press, 2002. 550 p.
- [26] MACKAY, D. J. C. Bayesian interpolation. *Neural Computation*, v. 4, n. 3, p. 415-447, 1992.
- [27] BISHOP, C. M. *Pattern Recognition and Machine Learning*. New York: Springer, 2006, 738 f.
- [28] BISHOP, C. M.; TIPPING, M. E. Bayesian regression and classification. *Advances in Learning Theory: Methods, Models and Applications*, v. 190, p. 267-285, IOS Press, NATO Science Series III: Computer and Systems Sciences, 2003.

- [29] TIPPING, M. E. Bayesian inference: an introduction to principles and practice in machine learning. *Advanced Lectures on Machine Learning*, v. 3176, p. 41-62, 2004.
- [30] TIPPING, M. E. The relevance vector machine. *Advances in Neural Information Processing Systems 12*, p. 652-658, 2000.
- [31] TIPPING, M. E. Sparse bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, v. 1, p. 211-244, 2001.
- [32] JANSSEN, G. J. M.; VRIENS, J. A. M. High resolution coherent radio channel measurements using direct sequence spread spectrum modulation. *IEEE Proceedings MELECON'91*, v. 1, p. 720-727, maio 1991.
- [33] SOUZA, C. F. *Análise da dispersão temporal de canais de banda ultralarga (UWB) através de medidas realizadas em ambientes internos e externos*. Niterói, 2006. 251 f. Dissertação (Mestrado em Engenharia de Telecomunicações) - Escola de Engenharia, Universidade Federal Fluminense, Niterói, 2006.
- [34] GENTILE, C.; KIK, A. A comprehensive evaluation of indoor ranging using ultra-wideband technology. *EURASIP Journal on Wireless Communications and Networking*, v. 2007, 10 p., 2007.
- [35] RENZO, M et al. The ultra-wide bandwidth outdoor channel: from measurement campaign to statistical modelling. *Mobile Networks and Applications*, v. 11, n. 4, p. 451-467, ago. 2006.
- [36] BARROS, F. J. B. *Medidas e Análise da Dispersão Temporal do Canal de Propagação UWB Indoor em Vários Tipo de Ambientes*. Rio de Janeiro, 2005. 122 f. Dissertação (Mestrado em Engenharia de Telecomunicações) - Centro de Estudos em Telecomunicações, PIC-Rio, Rio de Janeiro, 2005.
- [37] RAPPAPORT, T. S. Characterization of UHF multipath radio channels in factory buildings. *IEEE Transactions on Antennas and Propagation*. v. 37, n. 8, p. 1058-1069, ago. 1989.

- [38] RAPPAPORT, T. S.; SEIDEL, S. Y.; SINGH, R. 900 MHz multipath propagation measurements for US digital cellular radiotelephone. *IEEE Transactions on Vehicular Technology*. v. 39, p. 132-139, maio 1990.
- [39] HASHEMI, H.; THOLL, D. Analysis of the rms delay spread of indoor radio propagation channels. *Proceedings of the IEEE International Conference on Communications*, Chicago, p. 875-881, jun. 1992.
- [40] PAHLAVAN, K.; GANESH, R.; HOTALING, T. Multipath propagation measurements on manufacturing floors at 910 MHz. *Electronig Letters*, v. 25, n. 3, p. 225-227, fev. 1989.
- [41] MOLISCH, A. F. Ultrawideband propagation channels - theory, measurements, and modeling. *IEEE Transactions on Vehicular Technology*. v. 54, n. 5, set. 2005.
- [42] GOLOMB, S. W. *Shift Register Sequences*. Laguna Hills: Aegean Park Press, 1981. 257 p.
- [43] MORGAN, D. Autocorrelation function of sequential M-bit words taken from an N-bit shift register (PN) sequence. *IEEE Transactions on Computers*. v. C-29, n. 5, p. 408-410, maio 1980.
- [44] DIXON, R. C. *Spread Spectrum Systems*. 2 ed. New York: John Wiley & Sons, 1984, 440 f.
- [45] MATOS, L. J. *Influência da vegetação na dispersão dos sinais rádio-móveis*. Rio de Janeiro, 2005. 215 f. Tese (Doutorado em Engenharia Elétrica) - Centro de Estudos em Telecomunicações, PUC-Rio, Rio de Janeiro, 2005.
- [46] VASQUEZ, E. J. A. *Caracterização do Canal Móvel em Faixa Larga*. Rio de Janeiro, 2000. 209 f. Tese (Doutorado em Engenharia Elétrica) - Centro de Estudos em Telecomunicações, PUC-Rio, Rio de Janeiro, 2000.

- [47] COX, D. C. 910 MHz urban mobile radio propagation: multipath characteristics in New York city. *IEEE Transactions on Communications*, v. COM-21, n. 11, p. 1188-1194, nov. 1973.
- [48] PIRKL, R. J. *A sliding correlator channel sounder for ultra-wideband measurements*. Atlanta, 2007. 161 f. Dissertação (Mestrado em Engenharia Elétrica) - School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, 2007.
- [49] BULTITUDE, R. J. C; MAHMOUD, S. A; SULLIVAN, W. A. A comparison of indoor radio propagation characteristics at 910 MHz and 1.75 GHz. *IEEE Journal on Selected Areas in Communications*, v. 7, n. 1, p. 20-30, jan. 1989.
- [50] KIVINEN, J. et al. Wideband radio channel measurement system at 2 GHz. *IEEE Transactions on Instrumentation and Measurement*, v. 48, n. 1, p. 39-44, fev. 1999.
- [51] BROCATO, R. et al. UWB communication using SAW correlators. *Proceedings of IEEE Radio Wireless Conference*, Atlanta, p. 267-270, set. 2004.
- [52] CICCOGNANI, W.; DURANTINI, A.; CASSIOLI, D. Time domain propagation measurements of the UWB indoor channel using PN-sequence in the FCC-compliant band 3.6-6GHz. *IEEE Transactions on Antennas and Propagation*, v. 53, n. 4, p. 1542-1549, abr. 2005.
- [53] STEWART, K. A.; LABEDZ, G. P.; SOHRABI, K. Wideband channel measurements at 900 MHz. *Proceedings of IEEE Vehicular Technology Conference*, Chicago, p. 236-240, jul. 1995.
- [54] WEPMAN, J.; HOFFMAN, J. R.; LOEW, L. H. Analysis of impulse response measurements for PCS channel modelling applications. *IEEE Transactions on Vehicular Technology*, v. 44, n. 3, p. 613-620, ago. 1995.
- [55] HAYKIN, S. *Communication Systems*. 5 ed. John Wiley & Sons, 2009. 440 f.
- [56] LATHI, B. P. *Modern Digital and Analog Communication Systems*. 3 ed. Oxford: Oxford University Press, 1998. 800 f.

- [57] AGILENT TECHNOLOGIES. *Data Sheet - Agilent Technologies InfiniiVision 6000 Series Oscilloscopes*. mar. 2009. Disponível em: <http://cp.literature.agilent.com/litweb/pdf/5989-2000EN.pdf>. Acesso em: 20 ago 2009.
- [58] AGILENT TECHNOLOGIES. *Application Note 1587 - Evaluating oscilloscope sample rate vs. sampling fidelity: how to make the most accurate digital measurements*. fev. 2008. Disponível em: <http://cp.literature.agilent.com/litweb/pdf/5989-5732EN.pdf>. Acesso em: 20 ago 2009.
- [59] FRANKLIN, P.; RYLAND, J. IVI instrument driver guided tour. *Proceedings of Autotestcon*. n. 40, p. 167-173, set. 2004.
- [60] SHUTIN, D.; KUBIN, G; FLEURY B. H. Application of the evidence procedure to the estimation of wireless channels. *EURASIP Journal on Applied Signal Processing*, v. 2007, 23 p., 2007.
- [61] SHUTIN, D. *Multipath Tracking and Prediction for Multiple-Input Multiple-Output Wireless Channels*. Graz, 2006. 199 f. Tese (Ph.D. em Engenharia Elétrica) - Faculdade de Engenharia Elétrica e Tecnologia da Informação, Graz University of Technology, Graz, Austria, 2006.
- [62] SOUSA, E. S.; JOVANOVIĆ, V. M.; DAIGNEAULT, C. Delay spread measurements for the digital cellular channel in Toronto. *IEEE Transactions on Vehicular Technology*, v. 43, n. 4, p. 837-847, nov. 1994.
- [63] YANG, H.; SMULDERS, F. M.; HERBEN, M. H. A. J. Indoor channel measurements and analysis in the frequency bands 2 GHz and 60 GHz. *IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications*, p. 579-583, set. 2005.
- [64] NOORI, N.; KARIMZADEH-BAEE, R.; ABOLGHASEMI, A. An empirical ultra wideband channel model for indoor laboratory environments. *Radioengineering*, v. 18, n. 1, p. 68-74, abr. 2009.

- [65] CHAKRAVARTI, I. M.; LAHA, R. G.; ROY, J. *Handbook of Methods of Applied Statistics*. New York: Wiley, 1967, vol I, p. 392-394.
- [66] BOWMAN, A. W.; AZZALINI, A. *Applied Smoothing Techniques for Data Analysis*. Oxford: Oxford University Press, 1997, 208 f.
- [67] OESTGES, C.; VANHOENACKER-JANVIER, D.; CLERCKX, B. Channel characterization of indoor wireless personal area networks. *IEEE Transactions on Antennas and Propagation*, v. 54, n. 11, nov. 2006.
- [68] MCKINSTRY, D. R. *Ultra-wideband Small Scale Channel Modeling and its Application to Receiver Design*. Blacksburg, 2003. 238 f. Dissertação (Mestrado em Engenharia Elétrica) - Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, 2003.
- [69] GHASSEMZADEH, S. S.; JANA, R.; RICE, C. W.; TURIN, W.; TAROKH, V. Measurement and modeling of an ultra-wide bandwidth indoor channel. *IEEE Transactions on Communications*, v. 52, n. 10, p. 1786-1796, out. 2004.

Apêndice

Programas em MATLAB desenvolvidos para a dissertação.

```
function captura(ponto)
% OBJETIVO: realiza a captura da forma de onda exibida no
osciloscópio,
% através da interface GPIB. Armazena a forma de onda no diretório
% repositório.

ponto = num2str(ponto);
fd = 80e6;
td = 1/fd;      %tempo de bit
tamPN = 255;   %numero de bits da PN
tPN = td*tamPN; %duração da PN

tstamp = datestr(clock, 'yyyymmddTHHMMSS');
banco_caminho_com =
strcat('C:\banco_medidas\', 'P', ponto, '_', tstamp);
mkdir(banco_caminho_com);
P = input('Número de observações do sinal de sondagem: ');

obj = icdevice('Agilent546XX.mdd', 'GPIB::7::INSTR');

% realiza a captura de P realizações do canal e armazena em wave:
for p = 1:P
    connect(obj)
    groupobj = get(obj, 'Measurement');
    chan1 = groupobj(1);
    set(obj.Trigger(1), 'Level', 2e-2);
    xrecord = get(obj.Acquisition(1), 'TimePerRecord');
    set(obj.Acquisition(1), 'StartTime', 0);
    [wave_discard, initx, incrx] = invoke(chan1, 'ReadWaveform', 1);
    wave = [];
    tsweep = 0;
    isweep = -1.4*tPN;
    numvarred = 0;
    while (tsweep+xrecord) < 3.0*tPN
        tsweep = tsweep + xrecord;
        set(obj.Acquisition(1), 'StartTime', isweep);
    end
end
```

```

        [wave2, initx2, incrxw] = invoke(chan1, 'FetchWaveform');
        wave = [wave wave2];
        isweep = isweep + xrecord;
        numvarred = numvarred+1;
    end

    t = zeros(1,length(wave));
    for i = 2:length(t)
        t(i) = t(i-1)+incrx;
        wave(2,i)=t(i);
    end

    %salva a variável wave no diretório repositório:
    arquivo_caminho_com = strcat(banco_caminho_com, '\wave_',...
        num2str(p), '.mat');
    save(arquivo_caminho_com, 'wave');
    disconnect(obj)
    pause (.01);
end

figure; plot(t,wave(1,:));grid on;
% gera .txt com informações adicionais da captura:
captura_info_name = strcat('\captura_info',tstamp, '.txt');
captura_info_fullpath = strcat(banco_caminho_com,...
    captura_info_name);
fid = fopen(captura_info_fullpath,'wt');
fprintf(fid, 'P\t%d\n', P);
fprintf(fid, 'fd\t%.2e\n', fd);
fprintf(fid, 'tamPN\t%d\n', tamPN);
fprintf(fid, 'xrecord\t%.2e\n', xrecord);
fclose(fid);
end

```

```

function RVMjointSIMOtesteNoise
% OBJETIVO:
% Algoritmo RVM com inicializacao conjunta. Aplicado aos dados provenientes das
medicoes de sondagem.

%% PERFIS: PARAMETROS DO ALGORITMO RVM - RETIRE O COMENTARIO DA OPÇÃO UTILIZADA:

% perfil 1
%critparada = 0.0025;
%rho = 1 - 1e-1;

% perfil 2
% critparada = 0.0025;
% rho = 1 - 1e-2;

%% SELEÇÃO DO VETOR Z (CONTÉM A CORRELAÇÃO DA PN OBSERVADA NO OSCILOSCÓPIO COM A PN
DE REFERÊNCIA):
format long g
RVMcodeFolder = pwd;
FolderName = uigetdir('.', 'Selecionar diretório de vetores z_data');
if FolderName
    prompt = ('Número de observações da sequência a considerar:');
    name = 'Mensagem';
    numlines = 1;
    defAns = {'1'};
    P_cell = inputdlg(prompt, name, numlines, defAns);
    P = str2double(P_cell{1});
    for p = 1:P
        FileName = strcat('z_data_', num2str(p), '.mat');
        z_data_fullpath = strcat(FolderName, '\', FileName);
        load(z_data_fullpath);
        z(:,p) = z_data(1,:);
        t_z(:,p) = z_data(2,:);
    end
else
    error('A seleção do diretório de vetores z_data foi cancelada.')
end

info_file_num = randi(P); %selecao aleatoria de um dos arquivos info .txt
info_filename = strcat('gerarPDP_info_', num2str(info_file_num), '.txt');
add_info_fullpath = strcat(FolderName, '\', info_filename);
fid = fopen(add_info_fullpath, 'r');
add_info_val = fscanf(fid, ['%s' '%f']);
FDS = add_info_val(1);
SNR = add_info_val(4);

switch FDS %FDS = fator de subsampling ao vetor z.
    case 5
        r_eb = 8;
    case 2
        r_eb=18;
end
Ruusim_data_name = strcat('Ruusim_data_FDS_', num2str(FDS), '.mat');
load(Ruusim_data_name);
Ruusim = Ruusim_data(:,1);
tRuusim = Ruusim_data(:,2);
[Ruusim_max, i_Ruusimmax] = max(Ruusim);
tRuu = tRuusim(1:i_Ruusimmax);

N = length(tRuu); %N = samp_z;
N_eb = floor(N*3/4); %no. de amostras do vetor z que será feita a
busca.

```

```

samp_eb = 1:r_eb:N_eb;           %no. das amostras do vetor z que pertencem ao
espaco de busca.
L0 = length(samp_eb);

z_eb = zeros(length(t_z(:,1)),1);
z_eb(samp_eb) = z(samp_eb,1);   %amostras do vetor z que pertencem ao espaco
de busca.

figure; plot(t_z,z); hold on; grid on;
stem(t_z(:,1),z_eb,'Color','r'); hold off
screen_size = get(0, 'ScreenSize');
set(gcf, 'Position', [0 0 screen_size(3) screen_size(4) ] );

% AJUSTE DO ESPACO DE BUSCA:
prompt = ('Carregar dados do espaco de busca? [S(sim)/N(não)/C(converter)]');
name = 'Mensagem';
numlines = 1;
defAns = {'S'};
carregar_cell = inputdlg(prompt,name,numlines,defAns);
carregar_cmp = carregar_cell{1};
carregar_bool = strcmpi(carregar_cmp,'S');
ajustar_bool = strcmpi(carregar_cmp,'N');
converter_bool = strcmpi(carregar_cmp,'C');
if carregar_bool || converter_bool
    cd (FolderName);
    [samp_eb_FileName,samp_eb_FolderName,samp_eb_FilterIndex] = ...
        uigetfile('..\samp_eb*.mat','Selecionar vetor samp_eb');
    if samp_eb_FilterIndex
        samp_eb_fullpath = strcat(samp_eb_FolderName,samp_eb_FileName);
        load(samp_eb_fullpath);
        RVM_folder = strcat(FolderName,'\RVM\P_',num2str(P),'');
        mkdir(RVM_folder);
        tstamp = datestr(clock,'yyyymmddTHHMMSS');
        if converter_bool
            samp_bit_FDS_2 = 25/2;
            samp_bit_FDS_5 = 25/5;
            if FDS == 5
                conv_eb = samp_bit_FDS_2/samp_bit_FDS_5;
            elseif FDS == 2
                conv_eb = samp_bit_FDS_5/samp_bit_FDS_2;
            end
            samp_eb = floor(samp_eb./conv_eb);
        end
        L0 = length(samp_eb);
        z_eb = zeros(1,length(t_z(:,1)));
        z_eb(samp_eb) = z(samp_eb,1);

        figure; plot(t_z,z); hold on; grid on;
        stem(t_z(:,1),z_eb,'Color','r'); hold off
        screen_size = get(0, 'ScreenSize');
        set(gcf, 'Position', [0 0 screen_size(3) screen_size(4) ] );

% AJUSTE DO PERFIL CARREGADO:
prompt = ('Ajustar dados do espaco de busca carregado? [S(sim)/N(não)]');
name = 'Mensagem';
numlines = 1;
defAns = {'N'};
reajustar_cell = inputdlg(prompt,name,numlines,defAns);
reajustar_cmp = reajustar_cell{1};
reajustar_bool = strcmpi(reajustar_cmp,'S');
if reajustar_bool
    ajustar_bool = 1;

```

```

        z_eb = zeros(length(t_z(:,1)),1);
        z_eb(samp_eb) = z(samp_eb,1);
    else

        eb_name = strcat('eb',tstamp);
        eb_fullpath = strcat(RVM_folder,eb_name);
        saveas(gcf,eb_fullpath,'fig')

        samp_eb_name = strcat('samp_eb',tstamp);
        samp_eb_fullpath = strcat(RVM_folder,samp_eb_name);
        save(samp_eb_fullpath,'samp_eb');
    end
else
    error('O carregamento dos dados do espaço de busca foi cancelado.')
end
end
if ajustar_bool
    continuar_bool = 1;
    i = 4;
    while continuar_bool
        prompt = {'1- No. da amostra do espaço de busca a modificar:', '1- Numero de
amostras a deslocar (pos/neg/x):', ... % x = remover elemento
                '2- No. da amostra do espaço de busca a modificar:', '2- Numero de amostras
a deslocar (pos/neg/x):', ...
                '3- No. da amostra do espaço de busca a modificar:', '3- Numero de amostras
a deslocar (pos/neg/x):', ...
                '4- No. da amostra do espaço de busca a modificar:', '4- Numero de amostras
a deslocar (pos/neg/x):'};
        name = 'Ajuste do espaço de busca.';
        numlines = 1;
        defAns = {'0','0','0','0','0','0','0','0'};
        options.Resize='on';
        options.WindowStyle='normal';
        answer=inputdlg(prompt,name,numlines,defAns,options);
        pos_eb = zeros(i,1);
        pos_eb_mod = zeros(i,1);
        dele = 0;
        for m = 1:i
            pos_eb(m) = str2double(answer{2*m-1});
            pos_eb_mod(m) = str2double(answer{2*m});
            if ~isnan(pos_eb_mod(m)) && pos_eb(m) % caso não seja dado comando de
exclusão e seja selecionado elemento a deslocar
                samp_eb(pos_eb(m)) = samp_eb(pos_eb(m))+pos_eb_mod(m);
                z_eb(samp_eb(pos_eb(m)) - pos_eb_mod(m)) = 0;
                z_eb(samp_eb(pos_eb(m))) = z(samp_eb(pos_eb(m)),1);
            end
            if cell2mat(answer{2*m}) == 'x' % caso seja dado comando de exclusao
                dele(m) = pos_eb(m); % registro da posição do elemento a ser
excluido
            end
        end
        if ~isempty(dele)
            pos_dele = find(dele);
            z_eb(samp_eb(dele(pos_dele))) = 0;
            samp_eb(dele(pos_dele)) = [];
        end
        samp_eb = sort(samp_eb);
        L0 = length(samp_eb);

        figure; plot(t_z,z); hold on; grid on;
        stem(t_z(:,1),z_eb,'Color','r'); hold off
        screen_size = get(0, 'ScreenSize');
        set(gcf, 'Position', [0 0 screen_size(3) screen_size(4) ] );
    end
end
end

```

```

    prompt = {'Ajustar mais elementos do espaço de busca? [S/N]'};
    name = 'Mensagem';
    numlines = 1;
    defAns = {'N'};
    continuar_cell = inputdlg(prompt,name,numlines,defAns);
    continuar_cmp = continuar_cell{1};
    continuar_bool = strcmpi(continuar_cmp,'S');
end
RVM_folder = strcat(FolderName,'\RVM\P_',num2str(P),'\');
mkdir(RVM_folder);
tstamp = datestr(clock,'yyyymmddTHHMMSS');
eb_name = strcat('eb',tstamp);
eb_fullpath = strcat(RVM_folder,eb_name);
saveas(gcf,eb_fullpath,'fig')

samp_eb_name = strcat('samp_eb',tstamp);
samp_eb_fullpath = strcat(RVM_folder,samp_eb_name);
save(samp_eb_fullpath,'samp_eb');
end

% CONSTRUCAO DOS PARÂMETROS DO RVM:
cd (RVMcodeFolder);
% construcao da matriz K
K0 = zeros(N,L0);
for i = 1:L0
    K0(:,i) = Ruusim((i_Ruusimmax - (samp_eb(i)-1)):(end - (samp_eb(i)-1)));
end

% construcao da matriz lambda
lamb = zeros(N);
for i = 0:(i_Ruusimmax -1)
    lamb(:,i+1) = Ruusim(i_Ruusimmax - i:end - i);
end

N_z_tail = N-floor(.80.*N)+1; %numero de elementos da cauda do PDP
z_tail = zeros(P*N_z_tail,1);
for p = 1:P
    z_sel = z(:,p);
    z_tail_add = z_sel(floor(.80.*N):N);
    z_tail(N_z_tail*(p-1)+1:N_z_tail*(p-1)+N_z_tail) = z_tail_add;
end

N0 = estima_ruido(z_tail); % Estimação do ruído a partir da cauda da CIR
sigma0 = N0*lamb;
B0 = sigma0\eye(size(sigma0)); % == B0 = inv(sigma0)

%cálculo do ponto estacionario do hiperparâmetro alfa (estimativa inicial)
A0 = zeros(L0);

numalfa = zeros(L0,1);
denalfa = zeros(L0,1);
for l = 1:L0
    numalfa(l) = (K0(:,l)'*B0*K0(:,l)).^2;
    denalfa(l) = 1/P*sum(abs(K0(:,l)'*B0*z).^2,2) - K0(:,l)'*B0*K0(:,l);
    A0(l,1) = numalfa(l)./denalfa(l);
end

[Arow,Acol] = find(A0>0);
% Usando matrizes completas:
A = A0;
K = K0;

```

```

ap = Acol;      %somente as colunas dos alfa positivos;
L = length(Acol);
A(A<0) = Inf;
beta = 1/N0;

%%% ESTIMACAO CONJUNTA %%%
alfadifmax = critparada + 1;
med = zeros(L0,P);
cov = zeros(L0);
Anovo = zeros(L0);
Anovo(A==Inf) = Inf;
Adif = zeros(L0,1);
j = 1;
invop3 = zeros(N,P);
invop5 = zeros(N,P);
denomalfa = zeros(1,P);
denombeta = zeros(1,P);
while (alfadifmax > critparada) && (j < 1000)
    invop1 = lamb*K(:,ap); invop2 = A(ap,ap) + beta*K(:,ap)'*invop1;...
    cov(ap,ap) = invop2\eye(size(invop2));
    for p = 1:P
        invop3(:,p) = lamb*z(:,p); med(ap,p) =
beta*cov(ap,ap)*K(:,ap)'*invop3(:,p);
    end
    for i = 1:L
        for p = 1:P
            denomalfa(p) = cov(ap(i),ap(i)) + abs(med(ap(i),p)).^2;
        end
        Anovo(ap(i),ap(i)) = P./sum(denomalfa,2);
        Adif(ap(i),j) = abs(Anovo(ap(i),ap(i)) - A(ap(i),ap(i)))./A(ap(i),ap(i));
    end
    A = Anovo;
    alfadifmax = max(Adif(:,j));
    invop4 = lamb*K(:,ap);
    for p = 1:P
        invop5(:,p) = lamb*(z(:,p)-K(:,ap)*med(ap,p));
        denombeta(p) = (z(:,p)-K(:,ap)*med(ap,p))'*invop5(:,p);
    end
    beta = N*P/(P*trace(cov(ap,ap)*K(:,ap)'*invop4) + sum(denombeta,2));
    j = j+1;
end

N0 = 1/beta; %valor de N0 atualizado após o processo iterativo.
sigma = N0*lamb;
B = sigma\eye(size(sigma));
Ainv = zeros(L0);
for i = 1:L0
    Ainv(i,i) = 1./A(i,i);
end

% Determinação do valor de alfa limite para deleção das funções base irrelevantes:
sigc2 = zeros(L0);
medinvalfan=zeros(L0);
Pn = zeros(L0,1);
invalfath = zeros(L0,1);
invalfas = zeros(L0,1);
invalfathchap = zeros(L0,1);
for i = 1:L
    sigc2(ap(i)) = 1./(P*K(:,ap(i))'*B*K(:,ap(i)));
    medinvalfan(ap(i)) = P*sigc2(ap(i));
    Pn(ap(i)) = gamcdf(medinvalfan(ap(i)),P,sigc2(ap(i)));
    invalfath(ap(i)) = gaminv(rho,P,sigc2(ap(i)))-medinvalfan(ap(i));
    invalfas(ap(i)) = 3/8*((med(ap(i))).^2)./P;
    invalfathchap(ap(i)) = invalfath(ap(i))+invalfas(ap(i));
end

```

```

end

for i = 1:L
    if (Ainv(ap(i),ap(i)) < invalfathchap(ap(i)))
        Ainv(ap(i),ap(i)) = 0;
    end
end

[mpath,trash] = find(Ainv ~= 0);
mpathup = samp_eb(mpath);
medmpath = med(mpath,1);
mpathdet = zeros(N,1);
mpathdet(mpathup) = medmpath;

z_max = max(abs(z(:,1)));          % normalizando os pesos med a amplitude
mpathdet_max = max(abs(mpathdet)); % do maior multipercurso
norm_z = z_max./mpathdet_max;
mpathdet(mpathup) = norm_z.*medmpath;

mpathdet_data = zeros(N,2);
mpathdet_data(:,1) = mpathdet;
mpathdet_data(:,2) = t_z(:,1);

figure; plot(t_z,z); hold on; grid on;
title(['FDS = ',num2str(FDS),'; SNR = ',num2str(SNR),...
      '; r eb = ',num2str(r_eb),'; critparada = ',num2str(critparada),...
      '; \rho = ',num2str(rho,'%0.6f\n'),'; P = ',num2str(P)])
stem(t_z(:,1),mpathdet,'Color','r'); hold off
RVM_name = strcat('RVM',tstamp);
RVM_fullpath = strcat(RVM_folder,RVM_name);
saveas(gcf,RVM_fullpath,'fig')

mpathdet_data_name = strcat('mpath_data',tstamp);
mpathdet_data_fullpath = strcat(RVM_folder,mpathdet_data_name);
save(mpathdet_data_fullpath,'mpathdet_data');

add_info_name = strcat('add_info_RVM',tstamp,'.txt');
add_info_fullpath = strcat(RVM_folder,add_info_name);
fid = fopen(add_info_fullpath,'wt');
fprintf(fid,'ENTRADA:\n');
fprintf(fid,'P\t%d\n', P);
fprintf(fid,'FDS\t%d\n', FDS);
fprintf(fid,'rho\t%.6f\n', rho);
fprintf(fid,'critparada\t%.4f\n', critparada);
fprintf(fid,'r_eb\t%d\n', r_eb);
fprintf(fid,'SAIDA:\n');
fprintf(fid,'j\t%d\n', j);
fprintf(fid,'alfadifmax\t%.5f\n', alfadifmax);
fclose(fid);

if j == 500
    fid = fopen(add_info_fullpath,'a');
    fprintf(fid,'%s', 'Parametros nao convergiram.');
```