



ESTIMADORES DE CANAL PARA SISTEMAS PLC-DMT BASEADOS EM
MÁQUINAS DE VETORES DE SUPORTE

João Terêncio Dias

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia Elétrica.

Orientadores: Marcello Luiz Rodrigues de
Campos
Ernesto Leite Pinto

Rio de Janeiro
Março de 2013

ESTIMADORES DE CANAL PARA SISTEMAS PLC-DMT BASEADOS EM
MÁQUINAS DE VETORES DE SUPORTE

João Terêncio Dias

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Examinada por:

Prof. Marcello Luiz Rodrigues de Campos, Ph.D.

Prof. Ernesto Leite Pinto, D.Sc.

Prof. Eduardo Antonio Barros da Silva, Ph.D.

Prof. João Baptista de Oliveira e Souza Filho, D.Sc.

Prof. Marco Antonio Grivet Mattoso Maia, Ph.D.

Prof. Moisés Vidal Ribeiro, D.Sc

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2013

Dias, João Terêncio

Estimadores de canal para Sistemas PLC-DMT baseados em Máquinas de Vetores de Suporte/João Terêncio Dias. – Rio de Janeiro: UFRJ/COPPE, 2013.

XXI, 94 p.: il.; 29, 7cm.

Orientadores: Marcello Luiz Rodrigues de Campos
Ernesto Leite Pinto

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2013.

Referências Bibliográficas: p. 88 – 94.

1. Comunicação por Linhas de Potência. 2. Máquinas de Vetores de Suporte. 3. Estimação de canal. 4. OFDM. I. Campos, Marcello Luiz Rodrigues de *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*A Deus, a Santa Clara, a minha
mãe Maria Úrsula, a minha
esposa Soraia Cristina e a minha
filha Emily.*

Agradecimentos

Meus agradecimentos a todos aqueles que direta ou indiretamente contribuíram para confecção deste trabalho, em especial: Aos professores Marcello Campos e Ernesto por suas orientações, dedicação e todos os conhecimentos transmitidos. Aos professores Pedreira e Moisés por suas colaborações e dicas. Aos demais professores do Departamento de Engenharia Elétrica da COPPE/UFRJ pelo incentivo e conhecimentos ministrados. Aos companheiros e amigos do LPS pelos momentos de descontração e valiosas contribuições. A COPPE/UFRJ e ao CEFET/RJ, pelo fornecimento dos meios necessários para que este trabalho pudesse ser realizado com sucesso.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

ESTIMADORES DE CANAL PARA SISTEMAS PLC-DMT BASEADOS EM MÁQUINAS DE VETORES DE SUPORTE

João Terêncio Dias

Março/2013

Orientadores: Marcello Luiz Rodrigues de Campos
Ernesto Leite Pinto

Programa: Engenharia Elétrica

Neste trabalho é proposta uma técnica para a estimação e rastreamento de canal linear com variação temporal periódica (LPTV) em sistemas de comunicação por meio da rede elétrica (PLC). A parte de estimação de canal é baseada em máquinas de vetores de suporte rápidas (*fast-SVM*) com atualização adaptativa dos coeficientes, permitindo o treinamento em tempo real e sua aplicação em canais variantes no tempo. A parte de rastreamento do canal faz uso da característica de convergência de um dos algoritmos *fast-SVM*, que apresenta uma convergência mais rápida que os demais algoritmos e permite reinicializar o estimador quando o canal varia rapidamente, sem degradar seu desempenho. Os resultados de simulações mostram que a técnica proposta é mais robusta ao ruído impulsivo, tem um significativo ganho no erro quadrático médio (MSE) e resulta em menor taxa de erro de bits (BER) quando comparado ao estimador de mínimos quadrados recursivo (RLS), a uma versão aproximada ao de erro quadrático de média mínima linear (aLMMSE), ao de mínima média quadrática (LMS), ao mapeamento com auto-organização (SOM) e ao de mínimos quadrados (LS). Tais comparações foram realizadas usando o critério de equalização de canal baseado em erro quadrático de média mínima (MMSE) e considerando a presença de ruído impulsivo, ruído colorido de fundo e que o sistema é usado para a transmissão de dados em banda larga, em ambientes internos. A técnica proposta supera os demais algoritmos, com complexidade computacional aceitável, em todos os cenários PLC.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

CHANNEL ESTIMATORS FOR PLC-DMT SYSTEMS BASED IN SUPPORT VECTOR MACHINES

João Terêncio Dias

March/2013

Advisors: Marcello Luiz Rodrigues de Campos
Ernesto Leite Pinto

Department: Electrical Engineering

In this work we propose a technique for estimating and tracking the Linear and Periodically Time-Varying (LPTV) channels for Power Line Communication (PLC) systems. The channel estimation part is based on fast Support Vector Machine (SVM) with adaptive update of coefficients, allowing its training in real time and its application in time-varying channels. The channel tracking part makes use of the feature at the convergence of one *fast-SVM* algorithm, therefore, has a higher convergence speed than the other adaptive algorithms and allows reboots of the estimator when the channel suddenly changes, without degrading its performance. The simulation results show that the proposed technique is more robust against impulsive noise and performs better with respect to Mean Square Error (MSE) and Bit Error Rate (BER) when compared to Recursive Least Squares (RLS), an approximate Linear Minimum Mean Square Error (aLMMSE), Least Mean Square (LMS), Self-Organizing Mapping (SOM) and Least Squares (LS) estimators. Such comparisons were performed using channel equalizers based on the Minimum Mean Squared Error (MMSE) criterion and considering the presence of additive impulsive noise, background colored noise and that the system is used for broadband data transmission in indoor environment. The proposed technique outperforms other algorithms with acceptable computational complexity in all PLC scenarios.

Sumário

Lista de Figuras	x
Lista de Tabelas	xii
Lista de Símbolos	xiii
Lista de Abreviaturas	xix
1 <i>Introdução</i>	1
1.1 Motivação	1
1.2 Objetivo da tese	3
1.3 Metodologia	3
1.4 Organização da tese	4
2 <i>Sistemas PLC</i>	5
2.1 Principais projetos, recomendações e padrões de sistemas PLC	5
2.2 Análise do sistema PLC-DMT	8
2.2.1 Modelagem do sistema PLC-DMT	8
2.2.2 Modelagem do canal em sistema PLC	9
2.2.3 Modelagem do ruído no canal em sistema PLC	12
3 <i>Regressor SVM</i>	20
3.1 Fundamentação Teórica	20
3.2 SVM para Regressão	25
3.2.1 Regressão linear	25
3.2.2 Regressão não-linear	33
3.3 SVM do Tipo Proximal	35
3.4 SVM do Tipo Proximal Não-Polarizado	37
3.5 Regressor SVM no Plano Complexo	38
4 <i>Regressor PSVM de valores complexos</i>	40
4.1 Função-custo	41
4.2 Algoritmo PSVM adaptativo	44

5	<i>Regressor UPSVM de valores complexos</i>	50
5.1	Algoritmo UPSVM adaptativo	50
6	<i>Estimadores utilizados na comparação de desempenho</i>	54
6.1	LS	54
6.2	RLS	54
6.3	LMMSE	55
6.4	LMS	55
6.5	SOM	56
7	<i>Complexidade Computacional</i>	58
8	Simulações e análise dos resultados	61
8.1	Considerações e parâmetros das simulações	61
8.2	Comparação de desempenho dos algoritmos	63
8.3	Técnica de reinicialização do estimador UPSVM	74
8.4	Análise de desempenho com a técnica de reinicialização do UPSVM .	81
9	Conclusões	86
	Referências Bibliográficas	88

Lista de Figuras

2.1	Diagrama de blocos simplificado do sistema PLC-DMT.	8
2.2	Uma realização de 34 ms do canal LTI.	12
2.3	Uma realização de 34 ms do canal LPTV com variação harmônica. . .	13
2.4	Uma realização de 34 ms do canal LPTV com variação comutada. . .	13
2.5	Um instantâneo da resposta em frequência do canal em um sistema PLC.	14
2.6	Realização de 34 ms da resposta de uma frequência do canal LPTV com variação harmônica.	14
2.7	Realização de 34 ms da resposta de uma frequência do canal LPTV com variação comutada.	15
2.8	Uma realização do ruído aditivo em um intervalo de aproximadamente 100 ms.	18
2.9	Amostra de ruído impulsivo de acordo com o modelo explicitado na equação 2.10.	18
2.10	Densidade espectral de potência do ruído aditivo.	19
3.1	relação entre h e o erro de estimação para $L_{emp}(f(\cdot)) = 0$, $L(f(\cdot)) = 0$ e a complexidade intrínseca de $f(\cdot)$	23
3.2	Hiperplano e margens gerados no treinamento do regressor SVM. . .	26
4.1	Amostras $(X_p(m), Y_p(m))$ e a função-custo de ε -Huber.	41
4.2	Regiões ε -insensível, de custo quadrático e linear.	42
4.3	Amostra de valor complexo $Y_p(m)$	43
8.1	Geometria de posicionamento dos tons pilotos nos símbolos OFDM. .	62
8.2	Comparação entre as convergências dos algoritmos com o conjunto de parâmetros I e $SNR = 40$ dB em canal LTI e ruído colorido de fundo. 64	64
8.3	Comparação entre as convergências dos algoritmos com o conjunto de parâmetros II e $SNR = 40$ dB em canal LTI e ruído colorido de fundo.	65
8.4	Curva de aprendizado dos algoritmos em canal LTI com ruído AIN e $SNR = 30$ dB.	66

8.5	MSE \times SNR em canal LTI com ruídos ACN e AIN.	67
8.6	BER \times SNR em canais LTI com ruído colorido de fundo e modulação 16-QAM para valores de separação entre pilotos iguais a 2, 4 e 6.	68
8.7	BER \times SNR em canais LTI com ruído colorido de fundo e modulação 64-QAM para valores de separação entre pilotos iguais a 2, 4 e 6.	69
8.8	MSE das estimações do canal LPTV com variação comutada e ruído ACN.	70
8.9	MSE das estimações do canal LPTV com variação harmônica e ruído ACN.	71
8.10	BER em função da SNR com modulação 16-QAM em canal LPTV com variação comutada e ruído ACN.	72
8.11	BER em função da SNR com modulação 16-QAM em canal LPTV com variação harmônica e ruído ACN.	73
8.12	BER para diferentes escolhas de limiar, SNR = 40 dB, modulação 16-QAM e ruído AIN.	76
8.13	MSE da estimação do canal LPTV com variação comutada e ruído AIN, para os algoritmos adaptativos com a técnica de reinicialização proposta.	78
8.14	MSE em função da SNR em canal LPTV com variação comutada e ruído AIN, para os algoritmos adaptativos com uso da técnica de reinicialização proposta.	79
8.15	BER em função da SNR em canal LPTV com variação comutada e ruído AIN, para os algoritmos adaptativos com uso da técnica de reinicialização proposta.	80
8.16	MSE da estimação do canal LPTV com variação harmônica e ruído ACN.	81
8.17	MSE da estimação do canal LPTV com variação comutada e ruído ACN.	82
8.18	BER em função da SNR com modulação 16-QAM.	83
8.19	BER em função da SNR com modulação 64-QAM.	84

Lista de Tabelas

2.1	Características dos cabos de uma rede elétrica em ambientes internos	11
2.2	Parâmetros da PSD do ruído colorido de fundo	15
7.1	Número de <i>flops</i> por iteração no algoritmo PSVM linear com valores complexos.	58
7.2	Número de <i>flops</i> por iteração no algoritmo UPSVM linear com valores complexos.	59
7.3	Número de <i>flops</i> por iteração no algoritmo LS com valores complexos.	59
7.4	Número de <i>flops</i> por iteração no algoritmo RLS com valores complexos.	59
7.5	Número de <i>flops</i> por iteração no algoritmo aLMMSE com valores complexos.	60
7.6	Número de <i>flops</i> por iteração no algoritmo LMS com valores complexos.	60
7.7	Número de <i>flops</i> por iteração no algoritmo SOM com valores complexos.	60
8.1	Parâmetros usados nas simulações	62

Lista de Símbolos

(\mathbf{x}_i, y_i)	<i>Conjunto de amostras rotuladas em que \mathbf{x}_i representa uma amostra e y_i denota o seu rótulo, p. 21</i>
\cdot^H	<i>Conjugado transposto de um vetor ou matriz, p. 38</i>
\cdot^*	<i>Conjugado de um valor complexo, p. 38</i>
A_i	<i>Amplitude da i-ésima fonte geradora de ruído, p. 16</i>
C	<i>Parâmetro de generalização, p. 27</i>
C_c	<i>Capacitância por unidade de comprimento, p. 10</i>
D	<i>Tempo em que a impedância é igual a Z_A para o canal LPTV com variação comutada, p. 12</i>
E_G	<i>Tensão no circuito equivalente de Thévenin, p. 10</i>
E_L	<i>Tensão na carga instalada na saída do circuito elétrico, p. 10</i>
E_h	<i>Espaço de hipóteses, p. 22</i>
G	<i>Condutância na frequência de ressonância, p. 10</i>
G_c	<i>Condutância por unidade de comprimento, p. 10</i>
G_i	<i>Fator que representa o quanto a variância do ruído impulsivo é maior que a do ruído de fundo, p. 16</i>
I	<i>Primeiro conjunto de parâmetros testados, p. 63</i>
II	<i>Segundo conjunto de parâmetros testados, p. 65</i>
I_L	<i>Corrente elétrica na carga, p. 10</i>
I_s	<i>Número de senóide amortecidas presente na simulação do ruído impulsivo, p. 16</i>
$K(\cdot)$	<i>Função kernel, p. 33</i>

K_0, K_1, f_0	<i>Parâmetros para geração do modelo de ruído de fundo, dependentes da localização das medidas, p. 15</i>
$L(\cdot)$	<i>Funcional, p. 21</i>
L_c	<i>Indutância por unidade de comprimento, p. 10</i>
$L_{emp}(\cdot)$	<i>Funcional de risco empírico, p. 22</i>
M	<i>Número de amostras do conjunto de treinamento, p. 21</i>
M_Q	<i>Número de centróides na constelação da modulação, p. 56</i>
$O(\cdot)$	<i>Complexidade computacional assintótica, p. 60</i>
P	<i>Número de sinais piloto em um símbolo OFDM, p. 56</i>
$P(\cdot)$	<i>Um ponto em um plano, p. 26</i>
Q	<i>Fator de qualidade que determina a seletividade de Z, p. 10</i>
R	<i>Resistência na frequência de ressonância, p. 10</i>
R_c	<i>Resistência por unidade de comprimento, p. 10</i>
T	<i>Tempo em que a impedância é igual a Z_B para o canal LPTV com variação comutada, p. 12</i>
T_0	<i>Período da senóide de tensão elétrica da rede, p. 11</i>
V_{bg}	<i>Ruído colorido de fundo, p. 15</i>
$Z(\omega)$	<i>Impedância seletiva em frequência, p. 10</i>
$Z(\omega, t)$	<i>Impedância seletiva em frequência e variante no tempo, p. 11</i>
$Z_A(\omega), Z_B(\omega)$	<i>Componentes da impedância no canal LPTV, p. 11</i>
Z_G	<i>Impedância no circuito equivalente de Thévenin, p. 10</i>
Z_L	<i>Impedância na carga instalada na saída do circuito elétrico, p. 10</i>
Z_c	<i>Impedância característica, p. 10</i>
Γ	<i>Matriz diagonal cujo os elementos são iguais a ϖ, p. 55</i>
\Re, \Im	<i>Partes real e imaginária de um número complexo, respectivamente, p. 43</i>

α_i^+, α_i^-	<i>Multiplicadores de Lagrange para a parte real dos erros positivos e negativos, respectivamente, p. 28</i>
$\bar{\mathbf{Y}}_p(m)$	<i>Ponto simétrico de $\mathbf{Y}_p(m)$ na modulação dos tons piloto, p. 56</i>
β_i^+, β_i^-	<i>Multiplicadores de Lagrange para a parte imaginária dos erros positivos e negativos, respectivamente, p. 38</i>
\mathbf{K}	<i>Matriz de kernel, p. 33</i>
\mathbf{w}	<i>Vetor contendo os coeficientes da SVM, p. 25</i>
$\mathbf{\Pi}_p, \mathbf{K}_p, \mathbf{S}_p$	<i>Matrizes geradas nos algoritmos adaptativos, p. 48</i>
\mathcal{H}	<i>Matriz formada pela inclusão de uma coluna de 1's ao vetor de sinais piloto transmitido, p. 36</i>
δ	<i>Parâmetro de inicialização da matriz \mathbf{S}_p, p. 63</i>
ϵ_{eq}	<i>Permissividade equivalente do meio dielétrico, p. 11</i>
γ	<i>Parâmetro de regularização da função-custo de Huber, p. 31</i>
γ_c	<i>Constante de propagação, p. 10</i>
$\hat{\mathbf{H}}_p(m)$	<i>Estimação da resposta em frequência do canal na posição dos sinais piloto na m-ésima amostra, p. 40</i>
$\hat{f}(\cdot)$	<i>Modelo de função, p. 22</i>
$\hat{h}(m)$	<i>Variável aleatória relacionada a norma do erro de estimação da saída do canal, p. 74</i>
ι	<i>Índice do neurônio perdedor, p. 56</i>
κ	<i>Constante associada ao tipo de modulação dos sinais piloto no algoritmo aLMMSE, p. 55</i>
λ, η	<i>Multiplicadores de Lagrange para a parte real e imaginária da função-custo, respectivamente, p. 45</i>
\mathbb{C}	<i>Conjunto dos números complexos, p. 38</i>
\mathbb{R}	<i>Conjunto dos números reais, p. 21</i>
μ	<i>Fator de janelamento exponencial, p. 44</i>
$\nu_{imp}(t)$	<i>Ruído impulsivo aperiódico, p. 17</i>

ν_{nb}	<i>Ruído em banda estreita, no domínio do tempo, p. 16</i>
ν_{ps}	<i>Ruído impulsivo periódico, p. 16</i>
ω_0	<i>Frequência angular de ressonância, p. 10</i>
ω_i	<i>Frequência angular da i-ésima fonte geradora de ruído, p. 16</i>
ϕ	<i>Referência de fase em relação ao sinal de tensão da rede, p. 11</i>
ϕ_{max}	<i>Limite máximo para o erro de fase do algoritmo SOM, p. 56</i>
ϕ_m	<i>Desvio de fase médio a cada símbolo no algoritmo SOM, p. 56</i>
$\Pi(\cdot)$	<i>Envelope do ruído impulsivo, p. 16</i>
ψ_i	<i>Multiplicadores de Lagrange de valores complexos, p. 38</i>
ρ	<i>Margem entre o hiperplano separador e os hiperplanos de suporte, p. 25</i>
σ	<i>Desvio padrão de um vetor, p. 33</i>
σ^2	<i>Variância de um vetor, p. 33</i>
τ_i	<i>Fator de amortecimento i-ésima senóide que compõe o ruído impulsivo, p. 16</i>
θ_i	<i>Fase da i-ésima fonte geradora de ruído, p. 16</i>
$\tilde{\mathbf{H}}_p$	<i>Vetor contendo um conjunto de respostas em frequência do canal para uma posição de sinal piloto, p. 42</i>
$\tilde{\mathbf{R}}_{\mathcal{H}_p, \mathcal{H}_p}(m)$	<i>Matriz de kernel linear do algoritmo PSVM, p. 48</i>
$\tilde{\mathbf{X}}_p(m)$	<i>Vetor contendo um conjunto de X_p referente a um subcanal e cada elemento multiplicado por um peso associado ao índice temporal, p. 52</i>
$\tilde{\mathbf{Y}}_p(m)$	<i>Vetor contendo um conjunto de Y_p referente a um subcanal e cada elemento multiplicado por um peso associado ao índice temporal, p. 47</i>
$\tilde{\mathcal{H}}_p(m)$	<i>Matriz contendo um conjunto \mathcal{H}_p referente a um subcanal e cada elemento multiplicado por um peso associado ao índice temporal, p. 47</i>
$\tilde{\mathbf{h}}_p(m)$	<i>Vetor com a amostra no instante m de $\tilde{\mathcal{H}}_p(m)$, p. 48</i>

$\tilde{R}_{\mathbf{x}_p \mathbf{x}_p}(m)$	<i>Termo do kernel linear do algoritmo UPSVM, p. 53</i>
Υ	<i>Índice do neurônio vencedor, p. 56</i>
ε	<i>Margem de erro, p. 22</i>
φ	<i>Taxa de aprendizado do neurônio vencedor do algoritmo SOM, p. 56</i>
φ_{max}	<i>Fator de máximo aprendizado do algoritmo SOM, p. 56</i>
ϖ	<i>Constante relacionada à variância do ruído no algoritmo aLMMSE, p. 55</i>
ς	<i>Taxa de aprendizado do neurônio perdedor do algoritmo SOM, p. 56</i>
ϑ	<i>Passo de aprendizado do algoritmo LMS, p. 55</i>
ξ	<i>Função-custo associada à estimativa de ruído presente na parte real das amostras de treinamento, p. 27</i>
ξ^+, ξ^-	<i>Funções-custo da parte real dos erros nos dois lados do hiperplano, p. 27</i>
ξ_i	<i>Função-custo associada à parte real do erro de cada amostra de treinamento, p. 27</i>
ζ^+, ζ^-	<i>Funções-custo da parte imaginária dos erros nos dois lados do hiperplano, p. 38</i>
b	<i>Termo de polarização da SVM, p. 25</i>
$d(\cdot)$	<i>Distância entre um ponto e uma reta, p. 26</i>
f	<i>Frequência, p. 10</i>
$f(\cdot)$	<i>Função, p. 21</i>
f_i	<i>Frequência da i-ésima senóide amortecida na geração do ruído impulsivo, p. 16</i>
m	<i>Índice temporal dos símbolos OFDM, p. 40</i>
$p(\mathbf{x}, y)$	<i>Distribuição de probabilidade que descreve a relação entre os dados e os seus rótulos, p. 21</i>
t	<i>Índice temporal da IFFT, p. 17</i>

$t_{arr,s}$	<i>Intervalo entre ruídos impulsivos consecutivos, p. 16</i>
$t_{w,s}$	<i>Duração de uma realização do ruído impulsivo, p. 16</i>
y_r, y_s	<i>Rótulo associado a cada SV em cada lado do hiperplano, p. 29</i>
H	<i>Resposta em frequência do canal, p. 10</i>
I_1, I_2, I_3, I_4	<i>Conjunto de índices que separam os dados no regressor SVM, p. 43</i>
$\mathbf{C}_{p,\iota}(m)$	<i>Vetor com os centróides dos cluster dos neurônios perdedores, p. 56</i>
$\mathbf{C}_{p,\gamma}(m)$	<i>Vetor com os centróides dos cluster dos neurônios vencedores, p. 56</i>
\mathbf{D}_{Ω}	<i>Matriz cujos elementos não-nulos se encontram na diagonal principal, p. 40</i>
$\mathbf{H}_p(m)$	<i>Resposta em frequência do canal na posição dos sinais piloto na m-ésima amostra, p. 40</i>
\mathbf{I}	<i>Matriz identidade, p. 55</i>
\mathbf{R}	<i>Matriz de kernel linear, p. 38</i>
\mathbf{R}_{pp}	<i>Valor esperado da matriz de correlação do canal nas posições dos tons piloto, p. 55</i>
$\mathbf{V}_p(m)$	<i>Vetor de ruído na posição dos sinais piloto na m-ésima amostra, p. 40</i>
$\mathbf{X}_p(m)$	<i>Vetor de sinais piloto transmitido na m-ésima amostra, p. 40</i>
$\mathbf{Y}_e(m)$	<i>Erro de estimação da saída do canal, p. 74</i>
$\mathbf{Y}_p(m)$	<i>Vetor de sinais piloto recebido na m-ésima amostra, p. 40</i>
\mathbf{a}	<i>Vetor de pesos do regressor fast-SVM, p. 44</i>
$\mathbf{e}_p(m)$	<i>Erro em cada iteração dos algoritmos adaptativos, p. 48</i>
$\mathbf{x}_r, \mathbf{x}_s$	<i>Vetores de suporte de cada lado do hiperplano, p. 29</i>
$\mathcal{L}(\cdot)$	<i>Função-custo, p. 21</i>
h	<i>Dimensão VC, p. 23</i>

Lista de Abreviaturas

ACN	<i>Additive Colored Noise</i> , p. 66
AC	<i>Alternate Current</i> , p. 2
AIN	<i>Additive Impulsive Noise</i> , p. 66
AM	<i>Amplitude Modulation</i> , p. 16
AWGN	<i>Additive White Gaussian Noise</i> , p. 2
BER	<i>Bit Error Rate</i> , p. 63
CEPCA	<i>Consumer Electronics Powerline Communications Alliance</i> , p. 6
CP-OFDM	<i>Cyclic Prefix - Orthogonal Frequency Division Multiplexing</i> , p. 48
CP	<i>Cyclic Prefix</i> , p. 8
DMT	<i>Discrete Multitone</i> , p. 8
ERBF	<i>Exponential Radial Base Function</i> , p. 34
ETSI-PLT	<i>European Telecommunications Standards Institute - Power Line Telecommunications</i> , p. 6
FFT	<i>Fast Fourier Transform</i> , p. 9
HDTV	<i>High Definition Television</i> , p. 6
IEEE	<i>Institute of Electrical and Electronics Engineers</i> , p. 6
IFFT	<i>Inverse Fast Fourier Transform</i> , p. 8
ISI	<i>Intersymbol Interference</i> , p. 8
ITU	<i>International Telecommunication Union</i> , p. 7

KKT	<i>Karush-Kuhn-Tucker</i> , p. 36
LCTV	<i>Linear and Cyclically Time Variant</i> , p. 3
LMMSE	<i>Linear Minimum Mean Square Error</i> , p. 55
LMS	<i>Least Mean Square</i> , p. 2
LPTV	<i>Linear and Periodically Time-Varying</i> , p. 2
LS	<i>Least Square</i> , p. 2
LTI	<i>Linear Time Invariant</i> , p. 2
LTV	<i>Linear Time Variant</i> , p. 2
MMSE	<i>Minimum Mean Square Error</i> , p. 61
MSE	<i>Minimum Square Error</i> , p. 63
OFDM	<i>Orthogonal Frequency Division Multiplexing</i> , p. 8
OPERA	<i>Open PLC European Research Alliance</i> , p. 5
PLC	<i>Power Line Communication</i> , p. 1
PSD	<i>Power Spectral Density</i> , p. 15
PSVM	<i>Proximal Support Vector Machine</i> , p. 3
PS	<i>Pilot Separation</i> , p. 61
QAM	<i>Quadrature Amplitude Modulation</i> , p. 61
RBF	<i>Radial Base Function</i> , p. 33
RLS	<i>Recursive Least Square</i> , p. 2
RNAs	<i>Redes Neurais Artificiais</i> , p. 20
SNR	Signal-to-Noise Ratio, p. 63
SOM	<i>Self-Organizing Mapping</i> , p. 3
SRM	<i>Structural Risk Minimization</i> , p. 24
SVM	<i>Support Vector Machine</i> , p. 3
SV	<i>Support Vector</i> , p. 29

TD	<i>Transform-Domain</i> , p. 2
UPA	<i>Universal Powerline Association</i> , p. 6
UPSVM	<i>Unbiased Proximal Support Vector Machine</i> , p. 3
VC	<i>Vapnik-Chervonenkis</i> , p. 23
VoIP	<i>Voice over Internet Protocol</i> , p. 6
aLMMSE	<i>approximate Linear Minimum Mean Square Error</i> , p. 55

Capítulo 1

Introdução

1.1 Motivação

Impulsionado por um aumento significativo de demanda, tem-se observado um crescente interesse pelo desenvolvimento de sistemas de comunicações capazes de suportar altas taxas de transmissão e, ao mesmo tempo, que sejam atrativos técnica e economicamente.

Novas técnicas são desenvolvidas visando baratear os custos dos meios de transmissão e buscando utilizar os canais de comunicação já existentes. Dentre estas tecnologias, a comunicação através da rede elétrica (PLC - *Power Line Communication*) tem se apresentado bastante atrativa, sendo bastante promissora para o Brasil, onde um programa do governo federal procura levar eletricidade a 100% das residências [1] e um outro programa procura promover a inclusão digital [2].

Além das vantagens mencionadas, o sistema PLC tem sido apontado como uma importante tecnologia de suporte a aplicações de redes elétricas inteligentes [3] (*SmartGrid*).

A tecnologia PLC transforma uma rede de distribuição elétrica em uma rede de comunicação pela superposição de um sinal de informação de baixa energia ao sinal de corrente elétrica de alta potência. Esta tecnologia pode oferecer uma taxa de transmissão de centenas de megabits por segundo. No entanto, os condutores da rede elétrica foram projetados para transmissão de energia em 50 ou 60 Hz, não para fins de comunicação, e a impedância característica da linha de transmissão varia de acordo com o tipo de condutor e em função da frequência. O descasamento de impedância na rede resulta em um efeito de multipercurso que frequentemente causa nulos em determinadas frequências. Aparelhos elétricos conectados à rede elétrica também causam significativo impacto nas características do canal de sistemas PLC [4].

Uma importante característica do sistema PLC é o fato que seu canal apresenta

uma variação periódica sincronizada com a frequência da rede elétrica. Em [5] foi proposto um modelo linear e periodicamente variante no tempo (LPTV - *Linear and Periodically Time-Varying*) para o canal do sistema PLC. Os resultados apresentados em [6], em relação à taxa de variação temporal do canal do sistema PLC, mostram que o “tempo de coerência” do canal pode atingir valores próximos a $600 \mu\text{s}$ (este “tempo de coerência” foi estimado através da medida do espalhamento espectral observado na saída do canal quando ele é excitado por uma senóide). Entretanto, é comum que esses canais exibam uma variação “chaveada” (chaveamento entre dois estados do canal) cicloestacionária sincronizada com a frequência da componente fundamental da rede elétrica. Esta variação é causada pela mudança de estado das cargas conectadas à rede quando a tensão, na frequência fundamental da rede elétrica, está próximo de zero volt, ou seja, próximo ao cruzamento com zero do sinal de tensão da rede elétrica. Então, a resposta ao impulso do canal do sistema PLC sofre uma variação cicloestacionária sincronizada com a metade do período da componente fundamental do sistema elétrico ($T_0/2$) [7].

Além destas características, o ruído presente no canal do sistema PLC não pode ser modelado como ruído aditivo branco (AWGN - *Additive White Gaussian Noise*) [8]. Os ruídos na rede elétrica podem ser classificados principalmente em duas categorias, o ruído colorido de fundo e o ruído impulsivo. O ruído colorido de fundo tem uma densidade espectral de potência que diminui com o aumento da frequência, e é geralmente estacionário por longos períodos [8]. O ruído impulsivo tem amplitudes mais elevadas e é considerado a principal causa de ocorrência de erros em rajadas na comunicação. Uma característica interessante de uma parcela significativa do ruído impulsivo é sua dependência da fase da corrente alternada (AC - *Alternate Current*) do sistema elétrico. Este tipo de ruído é muitas vezes conhecido como ruído impulsivo periódico, síncrono com a frequência da rede elétrica [9]. Também, é comum a presença de ruídos em banda estreita caracterizado por sinais senoidais modulados em amplitudes e causados, principalmente, pela indução dos sinais de rádio nos fios da rede elétrica na banda de frequência de interesse.

Em função das características mostradas e das vantagens que os sistemas PLC podem oferecer, têm-se observado um significativo número de trabalhos sobre estimação de canal para estes sistemas [10–22]. Em um recente trabalho [22] os autores investigaram o uso de técnicas não-adaptativas, tais como: mínimos quadrados (LS - *Least Square*) e domínio da transformada (TD - *Transform-Domain*), e adaptativas, como: mínima média quadrática (LMS - *Least Mean Square*) e mínimos quadrados recursivo (RLS - *Recursive Least Square*), no domínio da frequência, com uso de sinais piloto. Tal investigação considerou os modelos de canal linear invariante no tempo (LTI - *Linear Time Invariant*), linear variante no tempo (LTV - *Linear Time Variant*) com distribuição Gaussiana para a variação temporal e “tempo de

coerência” igual ao proposto em [6] e linear ciclicamente variante no tempo (LCTV - *Linear Cyclically Time Variant*), que inclui a variação “chaveada” sincronizada com a rede elétrica, corrompidos pela presença de ruídos Gaussiano branco e impulsivo (os modelos de canal usados em [22] são praticamente determinísticos). Foi mostrado que, em canais LCTV, o melhor desempenho é obtido quando um algoritmo não adaptativo é escolhido, neste caso o LS ou o TD, associado a um interpolador linear. Em um outro trabalho [17] os autores propuseram o uso do algoritmo de mapeamento com auto-organização (SOM - *Self-Organizing Mapping*) para estimar e rastrear o canal e adotaram um ajuste adaptativo independente para os passos de aprendizado dos “neurônios” vencedores e perdedores. Neste caso, eles mostram que a técnica proposta pode ter bom desempenho se a fase do canal não sofrer variações significativas. Assim como em [22], em [17] é usado um modelo de canal praticamente determinístico. Após extensa pesquisa bibliográfica foi possível concluir que algumas questões permanecem em aberto, dentre as quais destacamos o uso de algoritmos com convergência rápida, robusto a vários tipos de ruído e capazes de rastrear as variações abruptas nos canais do sistema PLC.

1.2 Objetivo da tese

O uso de Máquinas de Vetores de Suporte (SVM - *Support Vector Machine*) tem sido proposto para resolver uma variedade de problemas de comunicações digitais [23–26], e tem se mostrado um estimador robusto de canal invariante no tempo [27, 28]. Contudo, até onde se sabe, nenhuma proposta de SVM ou de suas soluções analíticas, ou seja, proximal (PSVM - *Proximal Support Vector Machine*) ou proximal não-polarizada (UPSVM - *Unbiased Proximal Support Vector Machine*), foi estudada na estimação da resposta ao impulso de canais do sistema PLC.

O objetivo desta tese é apresentar o desenvolvimento de uma técnica adaptativa de estimação de canal baseada em SVM com valores complexos e de uma técnica de rastreamento adequada às variações do canal LPTV, mantendo a robustez aos efeitos do ruído impulsivo.

1.3 Metodologia

Neste trabalho é analisado o comportamento do sistema PLC em ambientes internos, com os seguintes modelos de canal: LTI, onde apenas a seletividade em frequência é observada; LPTV com variação harmônica, que considera a presença de cargas cuja impedância resultante varia harmonicamente com a frequência da rede, e LPTV comutado, que inclui a variação “chaveada” cicloestacionária do canal em

sistemas PLC; além dos modelos de ruído colorido, impulsivo periódico síncrono e assíncrono. São realizados: uma revisão da teoria sobre SVM e suas variações do tipo proximal (PSVM e UPSVM), o desenvolvimento matemático das técnicas propostas e, através de simulações computacionais, é comparado o desempenho dessas técnicas com técnicas de estimação que atualmente são apontadas como de bom desempenho para canais do sistema PLC.

1.4 Organização da tese

Este trabalho está organizado da seguinte maneira: no capítulo II é descrito o sistema PLC e suas principais características; no capítulo III, os fundamentos de SVM e suas variações *fast-SVM* (PSVM e UPSVM) para aplicações como regressores, e uma abordagem de valores complexos são estudados; no capítulo IV, a primeira técnica de estimação proposta é descrita; no capítulo V é desenvolvida a segunda técnica de estimação proposta; no capítulo VI, as técnicas que serão utilizadas na comparação de desempenho são apresentadas; no capítulo VII é analisada a complexidade computacional de cada um dos algoritmos; no capítulo VIII, as simulações e análises dos resultados são apresentadas e, no capítulo IX, as conclusões do trabalho são realizadas.

Capítulo 2

Sistemas PLC

O uso da rede elétrica como meio de propagação para sinais de telecomunicações pode ser dividido nas seguintes partes: a utilização de cabos entre a subestação do transformador de média para baixa tensão e o medidor de consumo de energia elétrica (ambiente externo), a exploração de instalações elétricas intradomiciliares (ambiente interno) e nas aplicações em sistemas embarcados.

2.1 Principais projetos, recomendações e padrões de sistemas PLC

Muitos projetos e recomendações para a implantação de sistemas PLC estão surgindo e concorrendo entre si. Não está claro qual irá predominar, por isso vamos apenas fazer uma rápida menção a essas propostas e destacar as principais especificações para a camada física e, na próxima seção, descrever um cenário para o teste das técnicas propostas que pode ser facilmente adaptado as recomendações de qualquer padrão.

- **OPERA** (*Open PLC European Research Alliance*) - projeto que desenvolve procedimentos de aplicação da tecnologia PLC como acesso a internet banda larga [29].
- **PowerNet** - projeto que desenvolve e valida equipamentos de comunicação em banda-larga cognitivos, do tipo *plug and play*, para redes elétricas que atendam aos requisitos da regulamentação relativos a radiações eletromagnéticas, que sejam capazes de oferecer altas taxas de transmissão durante o período em que a rede não esteja sendo utilizada por outro usuário (apresenta baixa densidade espectral de potência para o sinal de comunicação) e trabalhem com baixa relação sinal-ruído [30].

- **UPA (*Universal Powerline Association*)** - une indústrias do mercado PLC, tanto na área de acesso (última milha) quanto em ambiente interno, para assegurar a implantação de interoperabilidade e coexistência de produtos para sistemas PLC. A UPA promove produtos baseados no padrão Digital *Home-UPA* para aplicações em ambientes internos e sob as especificações do projeto OPERA para aplicações na área de acesso [31].
- **CEPCA (*Consumer Electronics Powerline Communications Alliance*)** - desenvolve e promove tecnologias PLC de alta taxa de transmissão [32].
- **ETSI-PLT (*European Telecommunications Standards Institute - Power Line Telecommunications*)** - desenvolve padrões para permitir a interoperabilidade entre equipamentos de diferentes fabricantes e coexistência de múltiplos sistemas no mesmo ambiente. As normas e especificações devem cobrir a prestação de serviços de voz e dados sobre sistemas PLC em ambientes internos e externos [33].
- **HomePlug Powerline Alliance** - viabiliza e promove padrões-base para sistemas PLC em ambiente interno, interoperabilidade, produtos para adoção e implementação de estudos de custo-benefício. A *HomePlug Powerline Alliance* desenvolve os seguintes padrões:
 - **HomePlug 1.0** - Especificações para conexão de dispositivos através da rede PLC doméstica;
 - **HomePlug AV** - Especificações de projeto para transmissão de televisão de alta definição (HDTV - *High Definition Television*) e voz sobre protocolo de internet (VoIP - *Voice over Internet Protocol*) em sistema PLC;
 - **HomePlug BPL** - Especificações para a conexão a área de acesso (ambiente externo);
 - **Homeplug CC** - Especificações para dispositivos de Comando e Controle em baixa velocidade.
- **IEEE (*Institute of Electrical and Electronics Engineers*)** - estabelece as seguintes normas para os sistemas PLC:
 - **IEEE P1675** - “Norma para a instalação de Hardware e questões de segurança em sistemas PLC de banda-larga”.
 - **IEEE P1775** - “Requisitos de compatibilidade eletromagnética dos equipamentos PLC (Métodos de teste e medição)”.

- **IEEE P1901** - “Especificações para as camadas de controle de acesso ao meio (MAC) e física (PHY)”.
- **ITU (*International Telecommunication Union*)** - desenvolve o padrão **ITU-T G.hn** com especificações para as camadas MAC e PHY em ambientes internos.

As principais recomendações de parâmetros para a camada física são:

- **IEEE 1901**

- Técnica de comunicação: FFT - OFDM (*Fast Fourier Transform - Orthogonal Frequency Division Multiplexing*);
 - * Número de pontos da FFT: 3072, 6144;
 - * Frequência de amostragem [MHz]: 75, 150;
 - * Duração do símbolo OFDM [μ s]: 40,96;
 - * Intervalo de Guarda [μ s]: 5,56 (12%), 7,56 (16%), 47,12 (53%) - Variável de acordo com as condições da linha;
 - * Modulação por subportadora: 4, 16, 64, 256, 1024 ou 4096-QAM;
 - * Faixa de frequência [MHz]: 2-30 (2-48 ou 2-60 - opcionais).
- Técnica de comunicação: Wavelet OFDM;
 - * Número de pontos da transformada wavelet discreta: 512, 1024 e 1536;
 - * Frequência de amostragem [MHz]: 62,5, 125, 187,5;
 - * Duração do símbolo [μ s]: 8,192;
 - * Modulação por subportadora: BPSK, 4, 8, 16 ou 32-PSK;
 - * Faixa de frequência [MHz]: 2-28 (2-60 - opcional).

- **ITU-T G.hn**

- Técnica de comunicação: FFT - OFDM;
- Número de pontos da FFT: escalável com tamanho em potência de 2 (dependente da faixa de frequência);
- Faixa de frequência [MHz]:
 - * 2-50 (IFFT de até 2048 pontos);
 - * 2-100 (IFFT de até 4096 pontos);
 - * 100-200 (IFFT de até 1024 pontos).

2.2 Análise do sistema PLC-DMT

Nesta seção é apresentado o modelo PLC-DMT adotado nesta tese, destacando as características dos sistemas multiportadoras (DMT - *Discrete Multitone*), assim como os modelos de canal e ruído.

O diagrama de blocos simplificado do sistema DMT é mostrado na Figura 2.1. A descrição do sistema e a modelagem matemática são feitas na subseção seguinte.

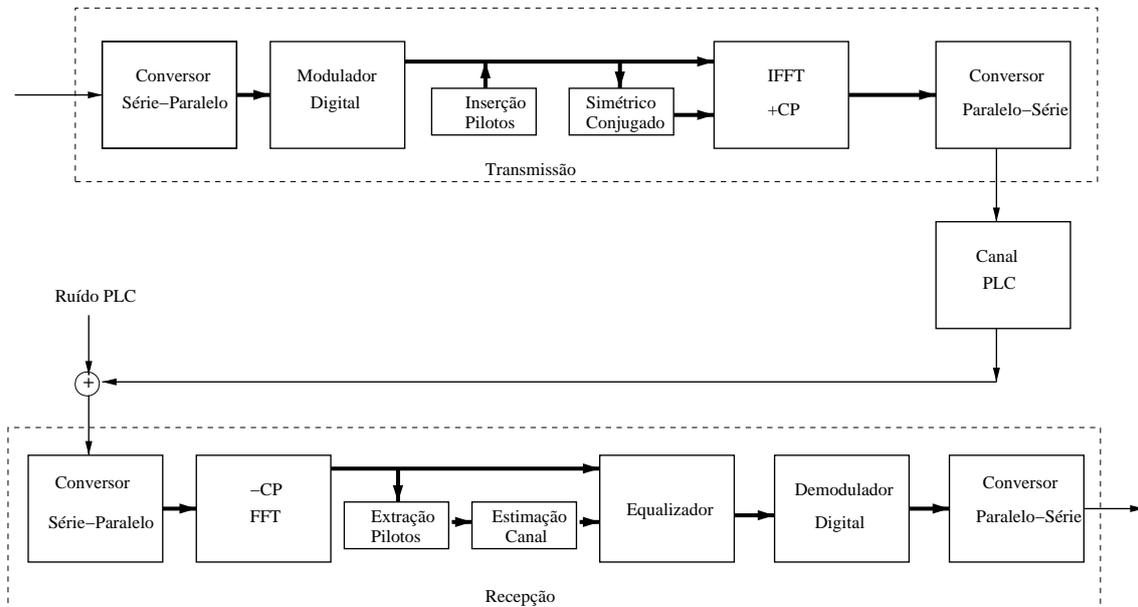


Figura 2.1: Diagrama de blocos simplificado do sistema PLC-DMT.

2.2.1 Modelagem do sistema PLC-DMT

DMT é um sistema de comunicação multiportadora igual ao OFDM (*Orthogonal Frequency Division Multiplexing*), onde o fluxo de dados com alta taxa de transmissão é dividido em fluxos de taxa mais baixa, transmitidos simultaneamente em um conjunto de subportadoras [34]. No DMT, um vetor de dados de tamanho N é acrescido de seu simétrico conjugado e enviado ao bloco da transformada inversa rápida de Fourier (IFFT - *Inverse Fast Fourier Transform*) de dimensão $2N$. Este procedimento converte os N símbolos de dados complexos no domínio da frequência em $2N$ amostras reais no domínio do tempo.

Um prefixo cíclico (CP - *Cyclic Prefix*) é acrescentado às $2N$ amostras no domínio do tempo para formar uma extensão cíclica no símbolo DMT. A duração do CP é escolhida de forma a ser maior que o atraso máximo de propagação do canal e dessa forma impedir a ocorrência de interferência intersimbólica (ISI - *Intersymbol Interference*), fazendo com que o símbolo DMT e o canal sejam convoluídos circularmente [35].

No receptor, antes da transformada rápida de Fourier (FFT - *Fast Fourier Transform*), o CP é removido fazendo com que a duração do símbolo DMT retorne às $2N$ amostras no domínio do tempo e, após a FFT, o número de amostras no domínio da frequência retorna a N pelo descarte das últimas N amostras.

Para o propósito de estimação de canal, podemos considerar que a seguinte relação entrada-saída na posição dos tons pilotos é obtida:

$$\begin{bmatrix} Y_1 \\ \vdots \\ \vdots \\ Y_P \end{bmatrix} = \begin{bmatrix} H_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & H_P \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ \vdots \\ X_P \end{bmatrix} + \begin{bmatrix} V_1 \\ \vdots \\ \vdots \\ V_P \end{bmatrix}, \quad (2.1)$$

em que $\{1, \dots, P\}$ são as posições dos tons pilotos e equivalem as posições $\{1, 1PS, \dots, (P-1)PS\}$ no símbolo DMT, sendo PS a separação entre pilotos. $[Y_1 \cdots Y_P]^T$, $[X_1 \cdots X_P]^T$ e $[V_1 \cdots V_P]^T$ são os vetores com as amostras do sinal recebido, transmitido e de ruído nestas posições, respectivamente, e $\mathbf{D}_{\{[H_1 \dots H_P]^T\}}$ é a matriz diagonal formada a partir do vetor de resposta em frequência do canal, também, nas posições dos tons pilotos.

A partir da forma matricial (2.1), pode-se ver que o símbolo recebido, no domínio da frequência, em cada subcanal é simplesmente uma versão escalada do símbolo transmitido mais ruído. Além disso, cada subcanal pode ser processado independentemente dos outros subcanais. Desta forma, a estimação do canal permite encontrar os coeficientes do equalizador que recupera o sinal transmitido.

2.2.2 Modelagem do canal em sistema PLC

Existem na literatura diversos trabalhos que se propõem a modelar o canal em sistema PLC, alguns baseados no conhecimento da topologia da rede, conhecidos como modelagem *bottom up* [36, 37], e outros baseados em campanhas de medidas, conhecidos como modelagem *top down* [5, 6, 38]. Os modelos de canais considerados neste trabalho são baseados em uma modelagem híbrida (parte determinística e parte estatística) proposta em [39], que classifica os modelos de canais em LTI e LPTV com variação temporal harmônica e comutada. Esses modelos são considerados mais realísticos que os demais encontrados na literatura, e a análise estatística apresentada em [39] é consistente com as medidas realizadas em canais reais apresentadas em [6].

Em [39], os autores consideram que a relação entrada-saída da rede pode ser

modelada como um quadripolo, caracterizado por sua matriz de parâmetros ABCD,

$$\begin{bmatrix} E_G \\ I_G \end{bmatrix} = \begin{bmatrix} 1 & Z_G \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E_L \\ I_L \end{bmatrix} \quad (2.2)$$

sendo,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \cosh(\gamma_c d) & Z_c \sinh(\gamma_c d) \\ \sinh(\gamma_c d)/Z_c & \cosh(\gamma_c d) \end{bmatrix}, \quad (2.3)$$

em que E_G e Z_G são a tensão e a impedância do circuito equivalente de Thévenin na entrada do canal, E_L e Z_L são a tensão e a impedância na carga instalada na saída do canal, Z_c é a impedância característica do circuito elétrico e γ_c é a constante de propagação. Z_c e γ_c estão relacionados a R_c , L_c , G_c , C_c (Resistência, Indutância, Condutância e Capacitância por unidade de comprimento, respectivamente) da linha de transmissão, de acordo com

$$\begin{cases} Z_c = \sqrt{\frac{R_c + j\omega L_c}{G_c + j\omega C_c}}, \\ \gamma_c = \sqrt{(R_c + j\omega L_c)(G_c + j\omega C_c)}, \end{cases} \quad (2.4)$$

e, sabendo que $I_L = E_L/Z_L$, a resposta em frequência do canal pode ser escrita como:

$$H(f) = \frac{E_L}{E_G} = \frac{Z_L(f)}{AZ_L(f) + B + CZ_L(f)Z_G + DZ_G}. \quad (2.5)$$

No caso de resposta seletiva em frequência do canal, um modelo adequado é o circuito RLC ressonante, cuja impedância pode ser descrita como:

$$Z(\omega) = \frac{R}{1 + jQ \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right)}, \quad (2.6)$$

sendo R a resistência na frequência de ressonância; ω_0 a frequência angular de ressonância e Q um fator de qualidade que determina a seletividade. Foi sugerido no capítulo 2 em [39], o seguinte conjunto de parâmetros para a simulação de canais LTI para ambientes internos na faixa de frequência até 30 MHz:

- Transmissor: impedâncias constantes de 50 Ω ;
- Comprimento das seções da rede: distribuição uniforme entre 0,5 m e 50 m;
- Tipos de cabos: escolhidos ao acaso, com distribuição uniforme, entre os tipos comercialmente usados em instalações elétricas em ambientes internos (considerando um conjunto de aproximadamente cinco tipos). Os valores para R_c , L_c , G_c , C_c e a constante dielétrica de isolamento devem ser obtidos para cada

tipo de cabo considerado. Neste trabalho usamos os valores apresentados na Tabela 2.1, propostos em [39], onde é considerado que a linha de transmissão não distorce o sinal e a atenuação é constante, logo a impedância característica é dada por um valor real, ou seja, $Z_c = \sqrt{L_c/C_c}$.

Tabela 2.1: Características dos cabos de uma rede elétrica em ambientes internos

Seção (mm^2)	1,5	2,5	4	6	10
ϵ_{eq}	1,45	1,52	1,56	1,73	2
$C_c(pF/m)$	15	17,5	20	25	33
$L_c(\mu H/m)$	1,08	0,96	0,87	0,78	0,68
R_1	12	9,34	7,55	6,25	4,98
G_1	30,9	34,7	38,4	42,5	49,3
$Z_c(\Omega)$	270	234	209	178	143

Sendo $R_c = R_1 \times 10^{-5} \sqrt{f}$ (Ω/m) e $G_c = G_1 \times 10^{-14} 2\pi f$ (S/m);

- Cargas conectadas à rede: $R \in [200, 1800]$ Ω ; $\omega_0/2\pi \in [2, 28]$ MHz e $Q \in [5, 25]$, todas com distribuição uniforme.

Para o modelo de canal LPTV com variação temporal harmônica, os autores consideram o mesmo procedimento adotado para o modelo LTI, mas a variação da impedância das cargas ao longo do ciclo da tensão na rede elétrica é modelada como [39],

$$Z(\omega, t) = Z_A(\omega) + Z_B(\omega) \left| \sin \left(\frac{2\pi}{T_0} t + \phi \right) \right|; \quad 0 \leq t \leq T_0 \quad (2.7)$$

em que $Z_A(\omega)$ é a componente da impedância invariante no tempo; $Z_B(\omega)$ é a amplitude máxima da componente da impedância com variação temporal e ϕ é o termo de fase que serve como referência para a variação em relação ao sinal de tensão da rede elétrica. Os valores dos parâmetros sugeridos são:

- $Z_A = 50 \Omega$;
- Z_B , uma função da impedância seletiva em frequência com parâmetros R , ω_0 e Q escolhidos aleatoriamente no conjunto apresentado para o caso LTI;
- ϕ , com distribuição uniforme entre 0 e π rad.

Para o modelo de canal LPTV com variação comutada, os autores consideram o mesmo procedimento adotado para o modelo LTI, mas o comportamento da impedância é representado pela comutação de dois valores Z_A e Z_B , síncrona com o período do sinal de tensão da rede elétrica (T_0). Neste caso os seguintes valores de parâmetros são indicados em [39]:

- Z_B , análogo ao caso LPTV com variação temporal harmônica;
- $Z_A = 0,5Z_B$;
- Período de tempo (T) que a impedância é igual a Z_B , em tempo discreto, com distribuição uniforme entre 1 período de amostragem e $T_0/4$;
- Período de tempo (D) que a impedância é igual a Z_A , em tempo discreto, com distribuição uniforme entre 0 e $T_0/2 - T$;

As Figuras 2.2, 2.3 e 2.4 ilustram uma realização de 34 ms para cada modelo de canal.

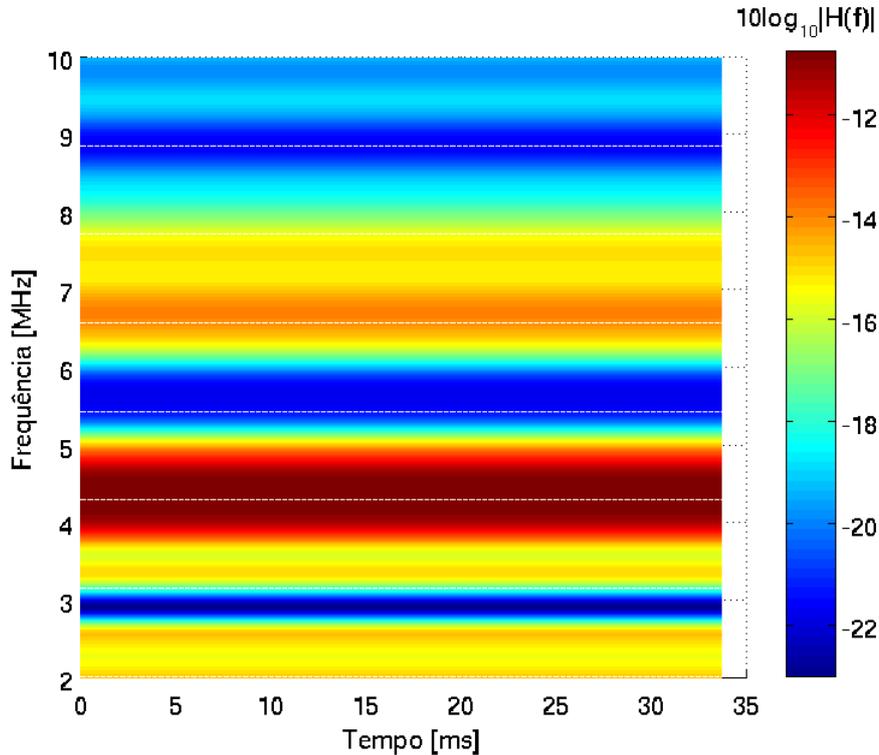


Figura 2.2: Uma realização de 34 ms do canal LTI.

A Figura 2.5 mostra um instantâneo da resposta em frequência do canal em um sistema PLC e as Figuras 2.6 e 2.7 ilustram uma realização de 34 ms da resposta de uma frequência do canal LPTV com variação harmônica e comutada, respectivamente.

2.2.3 Modelagem do ruído no canal em sistema PLC

As fontes de ruído são extremamente diversificadas nas linhas de baixa tensão. Em [8] foi feito um estudo, sendo os principais tipos de ruído observados em um canal em sistema PLC na faixa de 100 kHz a 30 MHz:

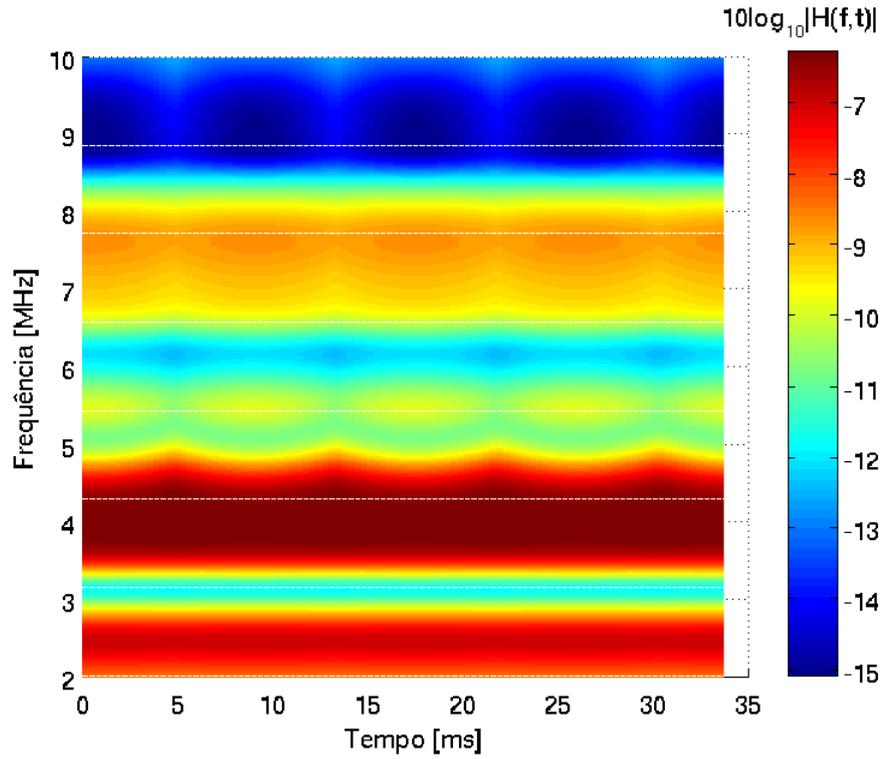


Figura 2.3: Uma realização de 34 ms do canal LPTV com variação harmônica.

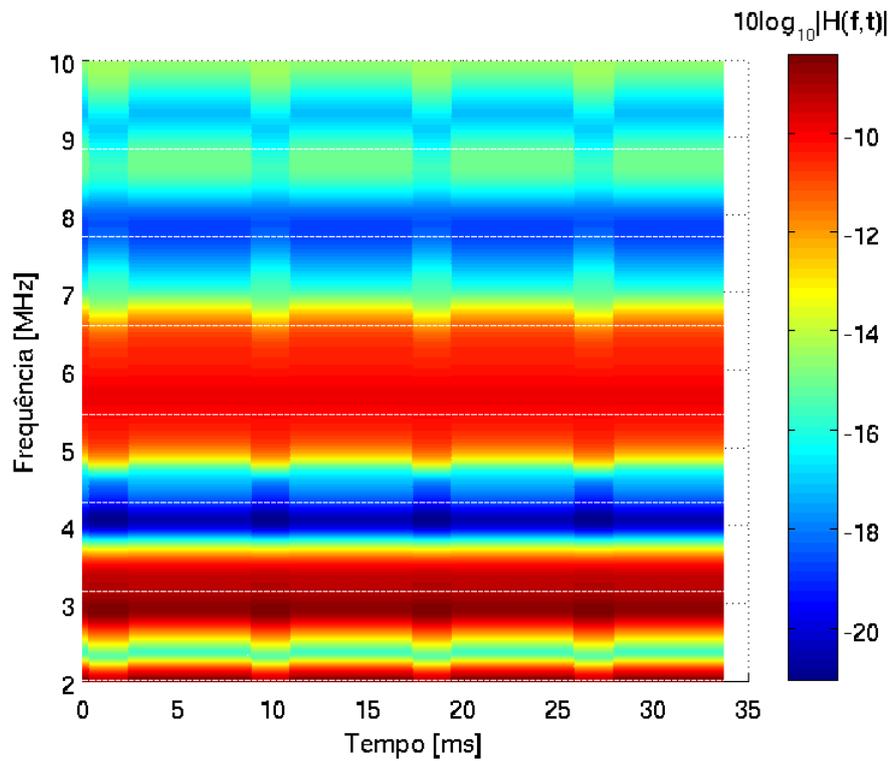


Figura 2.4: Uma realização de 34 ms do canal LPTV com variação comutada.

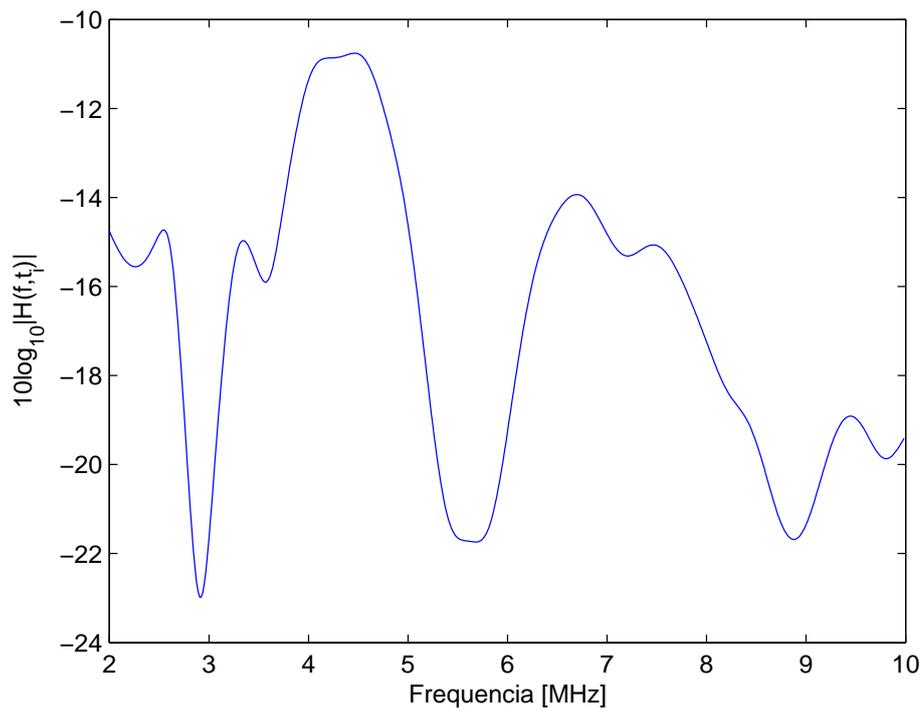


Figura 2.5: Um instantâneo da resposta em frequência do canal em um sistema PLC.

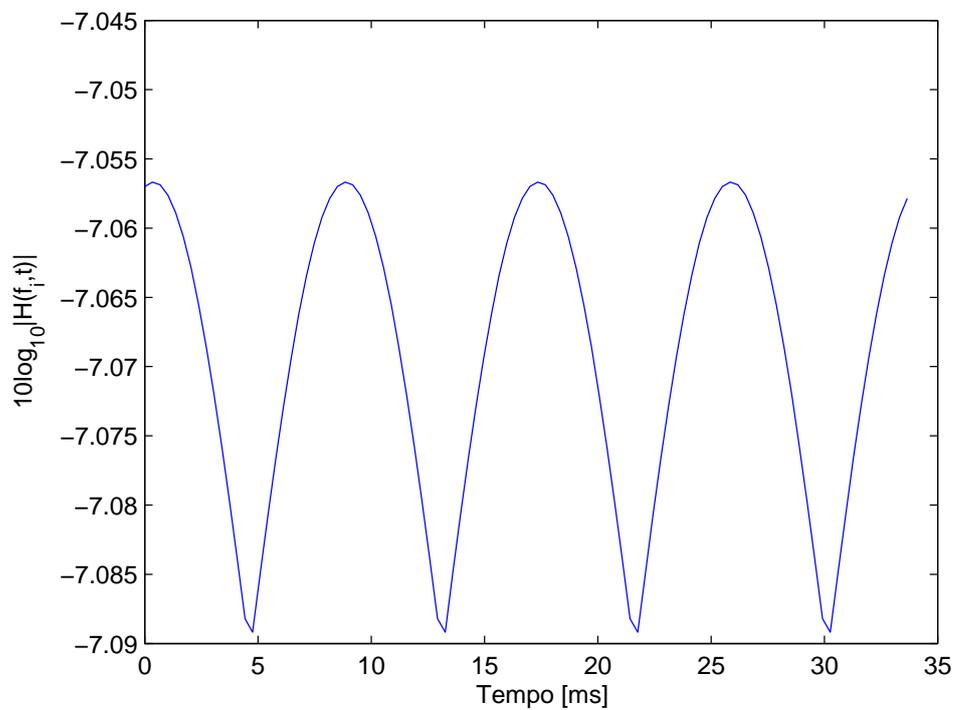


Figura 2.6: Realização de 34 ms da resposta de uma frequência do canal LPTV com variação harmônica.

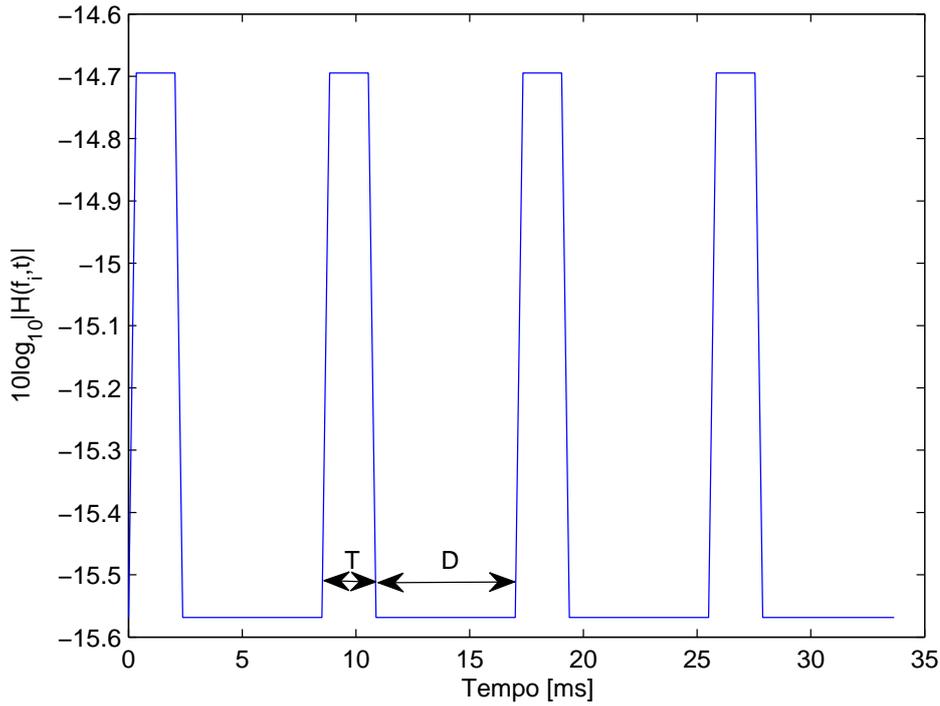


Figura 2.7: Realização de 34 ms da resposta de uma frequência do canal LPTV com variação comutada.

- **Ruído de fundo:** É o resultado da soma de vários ruídos gerados por diferentes fontes. Sua densidade espectral de potência (PSD - *Power Spectral Density*) varia muito pouco com o tempo, podendo ser considerada invariante por minutos ou até mesmo horas. Um modelo baseado em três parâmetros foi proposto por Esmailian [40] e é dado por

$$PSD_{bg}(f) = K_0 + K_1 |f|^{f_0} \text{ dBm/Hz}, \quad (2.8)$$

em que K_0, K_1, f_0 são parâmetros dependentes da localização das medidas e f é a frequência em MHz. Os valores de K_0, K_1, f_0 para o melhor e o pior cenário são mostrados na Tabela 2.2.

Tabela 2.2: Parâmetros da PSD do ruído colorido de fundo

Parâmetros	K_0	K_1	f_0
Melhor cenário	-140	38,75	-0,720
Pior cenário	-145	53,23	-0,337

- **Ruído em banda estreita:** É caracterizado por sinais senoidais modulados em amplitudes e causados, principalmente, pela indução dos sinais de rádio nos fios da rede elétrica na banda de frequência de interesse. O modelo proposto

por Dostert e Zimmermann [8] para esse tipo de ruído é

$$\nu_{nb}(t) = \sum_{i=1}^F A_i(t) \sin(\omega_i t + \theta_i). \quad (2.9)$$

Em que F é o total de fontes geradoras de ruído banda estreita no instante (t) e a amplitude da i -ésima fonte, $A_i(t)$, pode ser considerada constante ou modulada de forma a se obter a melhor representação dos sinais de rádio modulados em amplitude (AM - *Amplitude Modulation*).

- **Ruído Impulsivo Periódico:** É gerado pelos circuitos de varredura horizontal de televisores e monitores de computadores e pela utilização de aparelhos elétricos dotados de motores de escovas como aspiradores, liquidificadores, enceradeiras e furadeiras de mão. Estes eletrodomésticos injetam na linha (desconsiderando a parte irradiada) níveis de ruído de espectro amplo e variável. Estes tipos de ruído são classificados como ruído impulsivo periódico síncrono (de 60 Hz a 120 Hz) e assíncrono (de 5 kHz a 100 kHz) a frequência fundamental. O modelo descrito na Equação 2.10 foi proposto em [41], e expresso por

$$\nu_{ps}(t) = \sum_{i=1}^{I_s} A_i \sin(2\pi f_i(t - t_{arr,s}) + \theta_i) e^{-\frac{t-t_{arr,s}}{\tau_i}} \prod \left(\frac{t - t_{arr,s}}{t_{w,s}} \right), \quad (2.10)$$

em que I_s é número de senóides amortecidas presentes na simulação do ruído impulsivo e f_i é a frequência da i -ésima senóide amortecida. O termo $\prod(\cdot)$ é chamado de envelope do ruído impulsivo, sendo definido como um pulso retangular de duração $t_{w,s}$. $t_{arr,s}$ é o tempo que decorre entre os instantes de início de ruídos impulsivos consecutivos e A_i denota a amplitude da i -ésima sinusoidalidade. Assume-se que A_i é uma variável aleatória com distribuição $N(0, G_i \sigma_\nu^2)$ e $i = 1, \dots, I_s$, em que G_i representa o acréscimo sobre a variância do ruído de fundo e está na faixa de 20 a 30 dB acima da potência do ruído de fundo e τ_i denota o fator de amortecimento.

- **Ruído Impulsivo Aperiódico:** É causado, principalmente, por fontes chaveadas, reatores para iluminação e dimmers, dentre outras. Este ruído, muitas vezes, é similar ao ruído impulsivo periódico síncrono, têm duração de alguns micro-segundos à alguns mili-segundos e apresenta instantes de chegada aleatórios. A PSD para este tipo de ruído pode alcançar 50 dB acima da PSD dos ruídos de fundo. São altamente variantes no tempo e, juntamente com o ruído periódico, podem causar erros em alguns bits ou rajadas de bits transmitidos.

Nesta tese utilizamos um modelo simplificado para o ruído aditivo, similar ao adotado em [41, 42], que não considera o ruído em banda estreita, pois parte do princípio que este tipo de ruído pode ser facilmente mitigado através do uso de um filtro apropriado. Logo, o modelo de ruído adotado neste trabalho pode ser descrito no domínio do tempo por:

$$\nu(t) = \nu_{bg}(t) + \nu_{ps}(t) + \nu_{imp}(t), \quad (2.11)$$

sendo $\nu_{bg}(t)$ o ruído de fundo, $\nu_{ps}(t)$ o ruído impulsivo periódico síncrono e $\nu_{imp}(t)$ o ruído impulsivo aperiódico.

Na implementação do ruído aditivo, a duração de cada rajada do ruído impulsivo periódico síncrono é considerada como uma variável aleatória com distribuição log-normal de média igual a $50 \mu s$; e para o ruído aperiódico, é assumida uma distribuição similar de média $100 \mu s$. Essas considerações são consistentes com vários trabalhos encontrados na literatura sobre modelagem de ruído impulsivo para ambientes *indoor* [8, 9, 39, 43]. O tempo entre a chegada de dois impulsos aperiódicos consecutivos ($t_{arr,s}$) é considerado como uma variável aleatória exponencial, cujo valor médio é igual a 100 ms . No ruído impulsivo sincronizado com o sistema elétrico, o intervalo de chegada dos impulsos será $t_{arr,s} = T_0/2$, em que ($T_0 = 1/(60 \text{ Hz})$) [22].

A Figura 2.8 mostra uma realização do ruído aditivo em um intervalo de aproximadamente 100 ms (6 ciclos da frequência da rede), sendo possível observar a ocorrência de ruído impulsivo a cada intervalo de aproximadamente 8 ms , o que representa o ruído impulsivo periódico síncrono. A Figura 2.9 mostra, de forma ampliada, uma ocorrência de ruído impulsivo, permitindo observar o comportamento resultante do somatório de senóides amortecidas em sua composição, e a Figura 2.10 mostra a densidade espectral de potência do ruído aditivo, evidenciando que a frequência das senóides amortecidas na composição do ruído impulsivo encontra-se em 2 MHz e $3,5 \text{ MHz}$, e a PSD do ruído de fundo decresce com o aumento da frequência, caracterizando este modelo de ruído como ruído colorido.

Neste capítulo foram descritos: um modelo de sistema PLC-DMT, três modelos de canais (um invariante no tempo - LTI, um com variação harmônica e um com comutação entre dois estados da resposta impulsiva do canal, os dois últimos com variação sincronizada à frequência fundamental da rede elétrica - LPTV) adequados aos sistemas PLC *indoor* e um modelo de ruído aditivo composto de ruído colorido de fundo, ruído impulsivo síncrono e ruído impulsivo aperiódico.

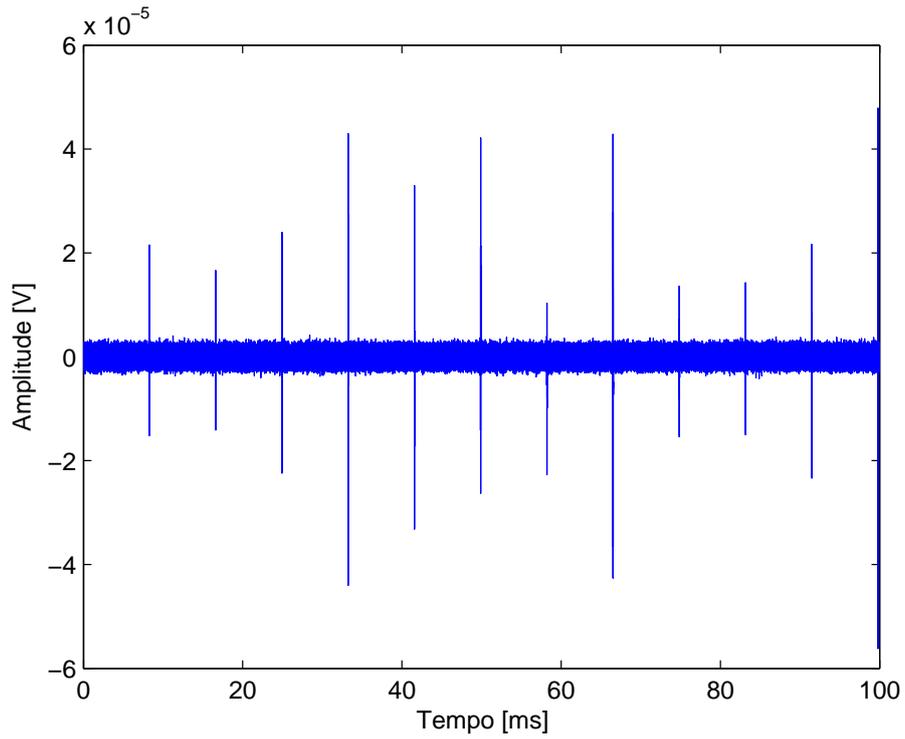


Figura 2.8: Uma realização do ruído aditivo em um intervalo de aproximadamente 100 ms.

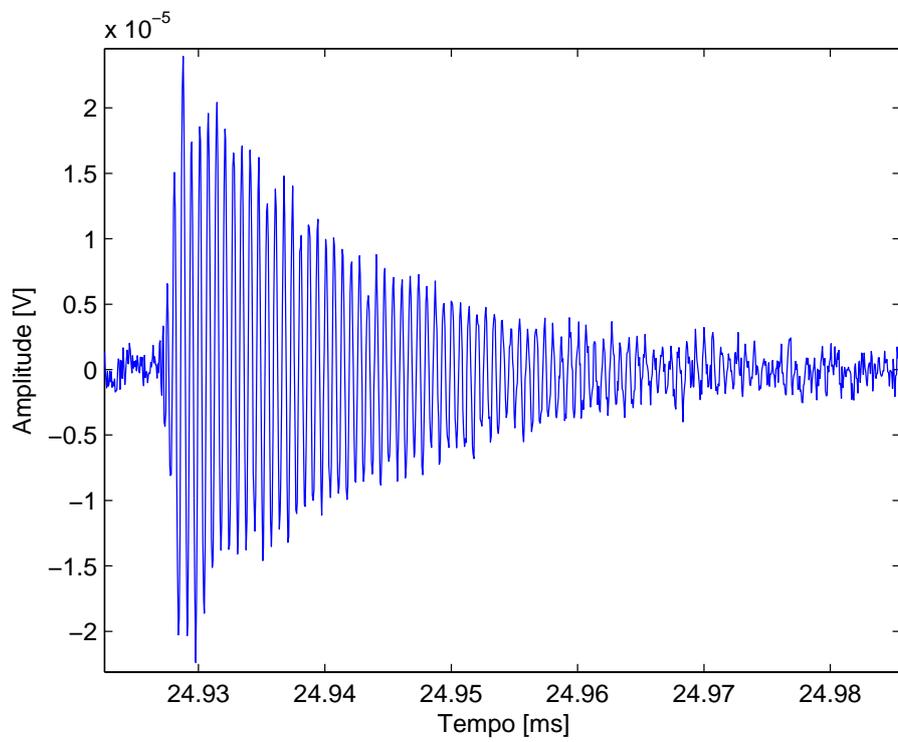


Figura 2.9: Amostra de ruído impulsivo de acordo com o modelo explicitado na equação 2.10.

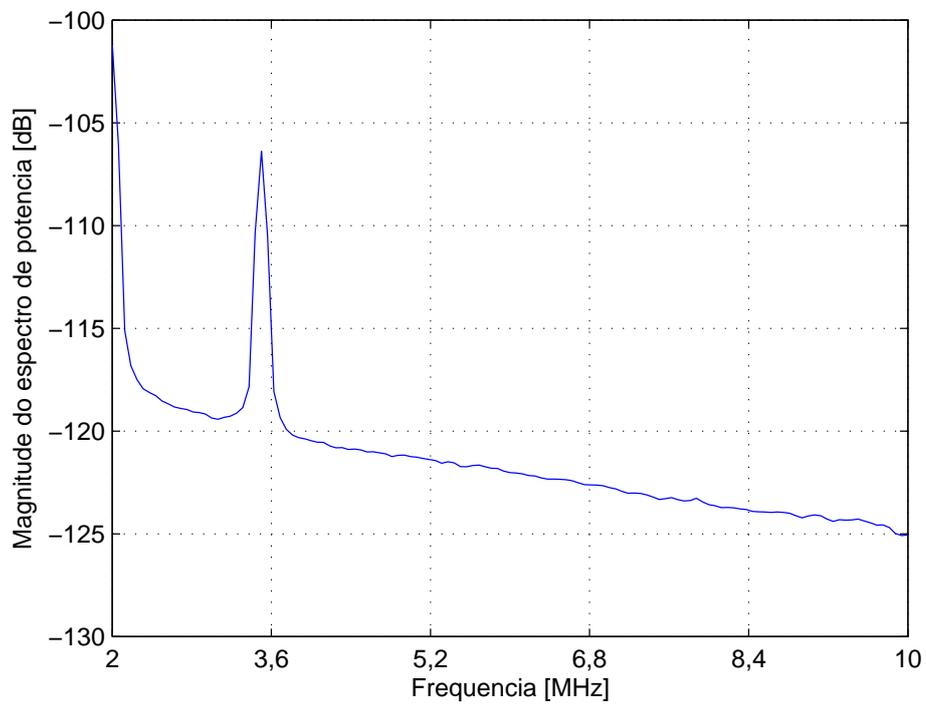


Figura 2.10: Densidade espectral de potência do ruído aditivo.

Capítulo 3

Regressor SVM

As máquinas de vetores de suporte (SVM - *Support Vector Machine*) formam uma técnica de aprendizado estatístico que vem recebendo crescente atenção da comunidade de aprendizado de máquina (*Machine Learning*) [44]. Resultados obtidos com esta técnica são comparáveis e frequentemente superiores aos obtidos por outros algoritmos de aprendizado estatístico, tais como as Redes Neurais Artificiais (RNAs) [45]. Tem-se observado aplicações de SVM em diversas áreas de pesquisa, como bioinformática [46], processamento de imagens [47, 48] e em sistemas de comunicações [49, 50].

Algoritmos para regressão, baseados em SVM, foram propostos em [27, 28, 51], em que a metodologia de solução da SVM leva a solução de um problema de otimização quadrática. Em [52] foi proposto um algoritmo para regressão baseado SVM, cuja metodologia de solução leva a um sistema de equações lineares (*fast-SVM*), contudo o algoritmo proposto em [52] só é adequado para situações cujo o ruído presente no sistema, pode ser modelado como Gaussiano. Neste capítulo, vamos apresentar dois algoritmos desenvolvidos para classificação (PSVM e UPSVM), no qual a solução também leva a um sistema de equações lineares e, a partir deles, iremos desenvolver duas novas propostas de algoritmos para regressão.

Neste trabalho, a análise da SVM será direcionada para a solução de problemas de regressão, apesar de sua proposta inicial ter sido feita por Vapnik [53] para resolver problemas de classificação. A fundamentação teórica de SVM será descrita de forma geral, uma vez que os conceitos apresentados são válidos tanto para classificação quanto para regressão.

3.1 Fundamentação Teórica

A teoria de aprendizado estatístico estabelece uma série de fundamentos para a obtenção de funções (classificadores e regressores) com boa generalização, ou seja, com capacidade de estimar corretamente sobre novos dados do mesmo domínio em

que houve o aprendizado.

As técnicas de aprendizado de máquina empregam um princípio de inferência denominado de indução, onde se obtém conclusões genéricas a partir de um conjunto particular de amostras. O aprendizado indutivo pode ser dividido, principalmente, em não-supervisionado e supervisionado.

No aprendizado não-supervisionado, o algoritmo apreende a agrupar as entradas submetidas segundo uma medida de qualidade. Essas técnicas são utilizadas, principalmente, quando o objetivo for encontrar padrões que auxiliem no entendimento dos dados.

No aprendizado supervisionado, tem-se o conhecimento do sistema através de um conjunto de amostras na forma entrada e saída desejada [45]. O algoritmo de aprendizado de máquina extrai a representação do conhecimento a partir dessas amostras. O objetivo é que a representação gerada seja capaz de produzir saídas corretas para novas entradas não apresentadas previamente.

Para o problema de aprendizado supervisionado, considera-se que os dados de treinamento, representados pelo conjunto de amostras rotuladas na forma (\mathbf{x}_i, y_i) , em que \mathbf{x}_i representa uma amostra, e y_i denota o seu rótulo, sejam independentes e identicamente distribuídos (*i.i.d*) e possuam distribuição de probabilidade $p(\mathbf{x}, y)$, que descreve a relação entre os dados e os seus rótulos [54].

Para estimar a taxa de acerto ou a taxa de erro obtidas pela função (classificador ou regressor) sobre novos dados, o conjunto de amostras é, em geral, dividido em dois subconjuntos disjuntos: de treinamento e de teste. O conjunto de treinamento é utilizado no aprendizado do conceito e o subconjunto de teste é utilizado para medir o grau de efetividade do conceito aprendido para novos dados.

Vapnik em [53] definiu a funcional $L(\cdot)$ como a probabilidade de classificação incorreta, tendo como argumento uma função $f(\cdot)$. Esta definição pode ser aplicada para argumentos na forma de função de regressão, e isto produz o que se chama de funcional de risco.

Seja $\mathcal{L}(\mathbf{x}, y, f(\mathbf{x})) : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$ uma função-custo, então a funcional de risco para funções de regressão $f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$ pode ser definida na forma:

$$L(f(\cdot)) = \int_{\mathbb{R}^n \times \mathbb{R}} \mathcal{L}(\mathbf{x}, y, f(\mathbf{x})) p(\mathbf{x}, y) d\mathbf{x} dy. \quad (3.1)$$

Considerando M amostras sujeitas a $p(\mathbf{x}, y)$, uma funcional de risco empírico pode ser dada por:

$$L_{emp}(f(\cdot)) = \frac{1}{M} \sum_{i=1}^M \mathcal{L}(\mathbf{x}_i, y_i, f(\mathbf{x}_i)), \quad (3.2)$$

que corresponde ao risco médio de treinamento.

No cálculo da funcional de risco costuma-se fazer uso de um valor chamado *margem de erro*, que é dado pela proporção de amostras de treinamento cuja distância da função $f(\cdot)$ é menor que um determinado valor $\varepsilon \in \mathbb{R}$. Esta estimativa empírica de risco conta as amostras como um erro se a aproximação da função é maior ou igual a ε .

O objetivo é minimizar a funcional de risco $L(f(\cdot))$ apresentada na Equação (3.1). Entretanto, $p(\mathbf{x}, y)$ é, geralmente, desconhecida. Assim, recorre-se à minimização da funcional de risco empírico $L_{emp}(f(\cdot))$ apresentada na Equação (3.2), a qual é baseada nos dados de treinamento.

Como o conjunto de dados de treinamento é finito, obter $L_{emp}(f(\cdot)) = 0$ não implica $L(f(\cdot)) = 0$.

Logo, deseja-se obter um modelo de função $\hat{f}(\cdot)$, a partir de um espaço de hipóteses ¹ E_h , que seja próximo, em relação a alguma medida de erro, à função fundamental $f(\cdot)$ no espaço objetivo ². Neste processo, dois tipos de erros são observados:

- Erro de Aproximação: é uma consequência do espaço de hipóteses não conter todo o espaço objetivo, de modo que a função fundamental $f(\cdot)$ pode residir fora deste espaço. Uma escolha ruim do E_h irá resultar em um erro de aproximação grande, o qual é referido como discordância do modelo.
- Erro de estimação: é o erro devido ao processo de aprendizado, que pode levar à seleção de um modelo que não seja o melhor possível dentro do espaço de hipóteses.

Em conjunto, estes erros formam o erro de generalização. Então, o que se busca na teoria de aprendizado estatístico é a minimização da funcional de risco empírico

$$\hat{f}(\cdot) = \arg \min_{f(\cdot) \in E_h} L_{emp}(f(\cdot)), \quad (3.3)$$

e esta minimização deve ser consistente no sentido de que

$$\lim_{M \rightarrow \infty} L_{emp}(f(\cdot)) = L(f(\cdot)). \quad (3.4)$$

Vapnik em [55] afirma que a Equação (3.4) é verdadeira a partir da lei dos grandes números. Entretanto, em virtude da existência do espaço de hipóteses E_h ,

¹Espaço formado pelo conjunto de todos os modelos de funções que um determinado algoritmo de aprendizado de máquina pode gerar a partir de um conjunto de amostras de treinamento.

²Espaço formado pelo conjunto de todas as funções que podem ser geradas a partir de um conjunto de amostras de treinamento.

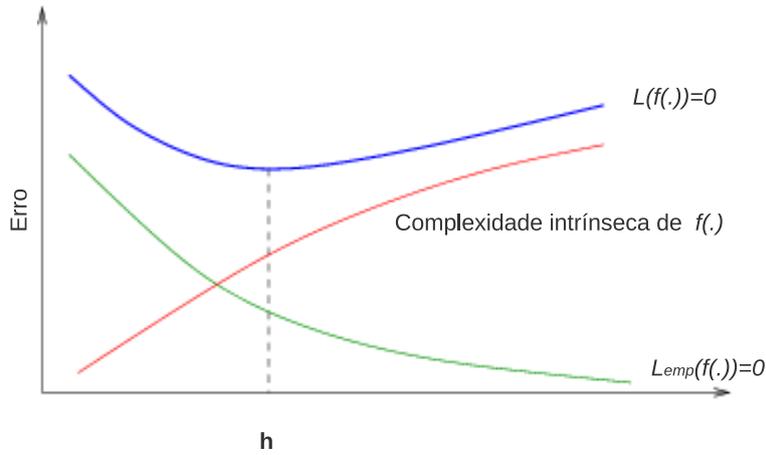


Figura 3.1: relação entre h e o erro de estimação para $L_{emp}(f(.)) = 0$, $L(f(.)) = 0$ e a complexidade intrínseca de $f(.)$.

a condição de consistência deve ser reescrita como

$$\lim_{M \rightarrow \infty} \min_{f(.) \in E_h} L_{emp}(f(.)) = \min_{f(.) \in E_h} L(f(.)), \quad (3.5)$$

que não é válida para qualquer E_h .

Considerando um número de amostras fixo na definição dos dados de treinamento, a complexidade intrínseca da função³ $f(.)$ vai determinar a relação entre $L_{emp}(f(.)) = 0$ e $L(f(.)) = 0$. Vapnik e Chervonenkis [53] definiram um índice escalar h que mede a complexidade intrínseca de uma classe de funções como *dimensão VC* (Vapnik-Chervonenkis). Na Figura 3.1, é ilustrado a relação entre h e o erro de estimação para $L_{emp}(f(.)) = 0$, $L(f(.)) = 0$ e a complexidade intrínseca de $f(.)$ [56].

Definição 1. *A dimensão VC de uma classe de funções é h se, e somente se, existe um conjunto de amostras $\{\mathbf{x}_i\}_{i=1}^h$ tal que, para qualquer uma das 2^h configurações possíveis de rotulação binária⁴, as amostras podem ser corretamente classificadas e não existe nenhum conjunto $\{\mathbf{x}_i\}_{i=1}^q$ com $q > h$ satisfazendo esta propriedade.*

³Medida de restrição do espaço de hipóteses que provê um limite a $L(f(.)) = 0$ e com isso evita o super-ajuste da função às amostras de treinamento.

⁴Identificação das amostras em duas classes

A minimização da funcional de risco, a qual vai estar associada ao desempenho de generalização do modelo resultante, requer que tanto a funcional de risco empírico como a razão entre a dimensão VC e o número de amostras de treinamento sejam pequenas. Como a funcional de risco empírico é usualmente uma função decrescente de h [53], para um dado número de amostras de treinamento, há um valor ótimo da dimensão VC. A escolha de um valor apropriado para h , que em muitos casos é controlado através do número de parâmetros livres do modelo, é crucial para se obter bom desempenho, especialmente quando o número de amostras de treinamento é pequeno.

A técnica de minimização do risco estrutural (SRM - *Structural Risk Minimization*), desenvolvida por Vapnik [57], é uma tentativa de tratamento do problema de escolha de uma dimensão VC apropriada. O princípio SRM, é baseado na observação de que um valor pequeno para a funcional de risco empírico $L_{emp}(f(\cdot))$ deve resultar em um valor pequeno para a funcional do risco esperado $L(f(\cdot))$. Para isso, a dimensão VC e a funcional de risco empírico devem ser minimizadas simultaneamente. Para implementar o princípio SRM é necessário propor uma estrutura aninhada para o espaço de hipóteses, na forma:

$$E_{h,1} \subset E_{h,2} \subset \dots \subset E_{h,k} \subset \dots \quad (3.6)$$

com a propriedade de que $h(k) \leq h(k+1)$ no qual $h(k)$ é a dimensão VC de $E_{h,k}$.

Então, o problema de otimização pode ser escrito na forma:

$$\min_{E_{h,k}} \left(L_{emp}(f(\cdot)) + \sqrt{\frac{h(k)}{M}} \right). \quad (3.7)$$

O princípio SRM é claramente bem fundamentado matematicamente, mas pode ser difícil de ser implementado pelas seguintes razões:

1. pode ser difícil calcular a dimensão VC de $E_{h,k}$, além do fato de que há somente um número pequeno de classes de funções para as quais é sabido como calcular a dimensão VC;
2. mesmo admitindo a viabilidade de obtenção da dimensão VC de $E_{h,k}$, o problema de minimização da expressão (3.7) pode ser de difícil solução.

A próxima seção vai mostrar que, embora não seja trivial controlar a dimensão VC da técnica de aprendizado durante a fase de treinamento, isto é adequadamente realizado pela abordagem SVM, a qual busca minimizar simultaneamente um limítante para a dimensão VC e uma funcional de risco empírico.

3.2 SVM para Regressão

As máquinas de vetores de suporte para o tratamento de problemas de regressão têm como propósito encontrar um hiperplano, dado por uma função $f(\cdot)$, de modo que as amostras \mathbf{x}_i dos conjuntos de treinamento e de teste se posicionem o mais próximas possível desta função, não importa de que lado do hiperplano essas amostras se encontrem.

3.2.1 Regressão linear

Considere o problema de representar um conjunto de dados de treinamento, composto pelas seguintes amostras de entrada-saída:

$$(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_M, y_M), \quad \mathbf{x} \in \mathbb{R}^n, \quad y \in \mathbb{R}, \quad (3.8)$$

por uma função linear

$$f(\mathbf{x}_i) = \mathbf{w}^T \mathbf{x}_i + b, \quad \text{com } \mathbf{w} \in \mathbb{R}^n, \quad b \in \mathbb{R}. \quad (3.9)$$

A função $f(\cdot)$ que atende o propósito de regressão e permite melhor generalização será obtida se, durante a fase de treinamento (aproximação da função), as distâncias de $f(\cdot)$ às amostras são mínimas e a distância de $f(\cdot)$ aos hiperplanos que confinam as amostras for máxima [58], ou seja, maximizar a margem $\rho(\mathbf{w}, b)$ que deve confinar as amostras na fase de teste. A Figura (3.2) ilustra esta situação.

Os vetores \mathbf{x}_i das amostras de treinamento devem obedecer a condição

$$|\mathbf{w}^T \mathbf{x}_i + b| \geq \frac{\rho(\mathbf{w}, b)}{2}, \quad (3.10)$$

uma vez que a função ótima deve ficar a igual distância das amostras mais distantes de ambos os lados.

Na Equação (3.10), $\rho(\mathbf{w}, b)$ pode ser tão grande quanto se queira, basta atuar na norma de \mathbf{w} e no valor de b . É necessário, portanto, impor alguma restrição, ou sobre a norma de \mathbf{w} , ou sobre a própria margem $\rho(\mathbf{w}, b)$. Vapnik mostrou em [57] que fixando a margem em $\rho(\mathbf{w}, b) = 2$, o problema de maximizar a margem se transforma em um problema equivalente de minimizar a norma de \mathbf{w} e resulta em um hiperplano canônico, no qual \mathbf{w} e b devem ser tais que

$$|\mathbf{w}^T \mathbf{x}_i + b| \geq 1, \quad \text{com } i = 1, \dots, M \quad (3.11)$$

e, sabendo que a distância⁵ $d(\mathbf{w}, b; \mathbf{x})$ de um vetor $\mathbf{x}_i \in \mathbb{R}^n$ à função $f(\mathbf{x}_i)$ pode ser

⁵Seja um ponto $P(x_1, y_1)$ e uma reta r definida por $ax + by + c = 0$, a distância do ponto

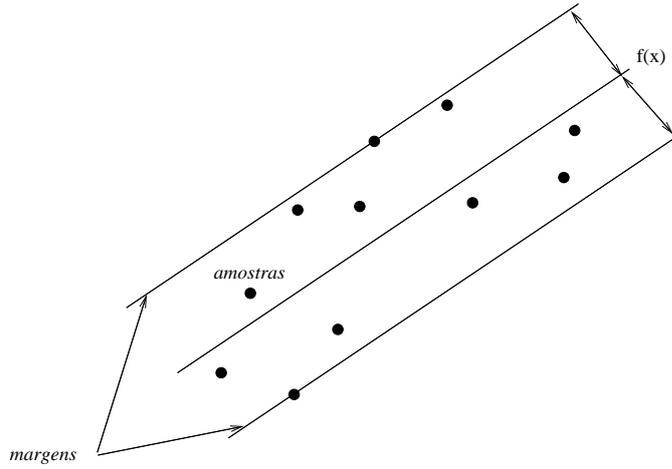


Figura 3.2: Hiperplano e margens gerados no treinamento do regressor SVM.

expressa na forma

$$d(\mathbf{w}, b; \mathbf{x}_i) = \frac{|(\mathbf{w}^T \mathbf{x}_i) + b|}{\|\mathbf{w}\|}. \quad (3.12)$$

A margem, sujeita às restrições da Expressão (3.11), é dada por:

$$\begin{aligned} \rho(\mathbf{w}, b) &= \min_{\{\mathbf{x}_i, y_i > f(\mathbf{x}_i)\}} d(\mathbf{w}, b; \mathbf{x}_i) + \min_{\{\mathbf{x}_j, y_j < f(\mathbf{x}_j)\}} d(\mathbf{w}, b; \mathbf{x}_j); \\ &= \min_{\{\mathbf{x}_i, y_i > f(\mathbf{x}_i)\}} \frac{|\mathbf{w}^T \mathbf{x}_i + b|}{\|\mathbf{w}\|} + \min_{\{\mathbf{x}_j, y_j < f(\mathbf{x}_j)\}} \frac{|\mathbf{w}^T \mathbf{x}_j + b|}{\|\mathbf{w}\|}; \\ &= \frac{1}{\|\mathbf{w}\|} (\min_{\{\mathbf{x}_i, y_i > f(\mathbf{x}_i)\}} |\mathbf{w}^T \mathbf{x}_i + b| + \min_{\{\mathbf{x}_j, y_j < f(\mathbf{x}_j)\}} |\mathbf{w}^T \mathbf{x}_j + b|); \\ &= \frac{2}{\|\mathbf{w}\|}. \end{aligned}$$

Logo, a função que maximiza a margem é aquela que minimiza:

$$L(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2. \quad (3.13)$$

Minimizar a expressão na Equação (3.13) é equivalente a implementar o princípio SRM [59].

Um problema observado nesta abordagem é o fato de não considerar a presença de ruído nas amostras de treinamento. No intuito de considerá-lo, Cortes & Vapnik [60]

$P(x_1, y_1)$ à reta r é definida por: $d(a, b; P(x_1, y_1)) = \frac{|ax_1 + by_1 + c|}{\sqrt{a^2 + b^2}}$

introduziram uma função de penalidade (funcional de risco empírico), na forma:

$$L(\xi) = \sum_{i=1}^M \xi_i, \quad \xi_i \geq 0, \quad (3.14)$$

em que cada ξ_i , $i = 1, \dots, M$, indica o erro de aproximação (função-custo) associado a cada amostra de treinamento. A função-objetivo do problema de otimização é agora composta de um termo associado à dimensão VC do regressor e um outro termo vinculado ao erro empírico. Sendo assim, o hiperplano ótimo generalizado é determinado pelo vetor \mathbf{w} , que minimiza a funcional:

$$L(\mathbf{w}, \xi^+, \xi^-) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \left(\sum_{i=1}^M \xi_i^+ + \sum_{i=1}^M \xi_i^- \right), \quad (3.15)$$

sujeito às restrições

$$\begin{cases} y_i - f(\mathbf{x}_i) \leq \xi_i^+ & \text{se } y_i - f(\mathbf{x}_i) > 0; \\ f(\mathbf{x}_i) - y_i \leq \xi_i^- & \text{se } y_i - f(\mathbf{x}_i) < 0, \end{cases} \quad (3.16)$$

em que C é um valor especificado pelo usuário, denominado parâmetro de generalização, ξ^+ e ξ^- (≥ 0) estão relacionadas ao erro de aproximação nos dois lados do hiperplano.

Neste caso, existe a necessidade de especificar uma função-custo associada à estimativa do ruído presente nas amostras. A seguir, são analisadas as principais funções-custo usadas para regressão em SVM.

• Função-custo de Laplace

Esta função-custo é pouco sensível a *outliers* devido a inclinação das duas semiretas gerada pela função e supondo que os *outliers* não vão se distanciar muito do hiperplano.

$$\xi_i = |y_i - f(\mathbf{x}_i)|. \quad (3.17)$$

Neste caso, o problema de otimização da Equação (3.15), juntamente com as restrições da Expressão (3.16), resultam num problema quadrático com restrições, cuja solução é o ponto de sela da funcional lagrangeano:

$$\begin{aligned} L(\mathbf{w}, b, \xi^+, \xi^-, \alpha^+, \alpha^-) &= \frac{1}{2} \|\mathbf{w}\|^2 + C \left(\sum_{i=1}^M \xi_i^+ + \sum_{i=1}^M \xi_i^- \right) \\ &+ \sum_{i=1}^M \alpha_i^+ [y_i - (\mathbf{w}^T \mathbf{x}_i + b) - \xi_i^+] + \sum_{i=1}^M \alpha_i^- [(\mathbf{w}^T \mathbf{x}_i + b) - y_i - \xi_i^-], \end{aligned} \quad (3.18)$$

sendo $\alpha_i^+, \alpha_i^- (i = 1, \dots, M)$ os multiplicadores de Lagrange para os erros positivos (\cdot^+) e negativos (\cdot^-), respectivamente. O lagrangeano da Equação (3.18) tem que ser minimizado com respeito a \mathbf{w} , b , ξ^+ e ξ^- e maximizado com respeito a α^+, α^- . O problema dual equivalente tem a forma

$$\max_{\alpha^+, \alpha^-} L_{pd}(\alpha^+, \alpha^-) = \max_{\alpha^+, \alpha^-} \left\{ \min_{\mathbf{w}, b, \xi^+, \xi^-} L(\mathbf{w}, b, \xi^+, \xi^-, \alpha^+, \alpha^-) \right\}. \quad (3.19)$$

O mínimo do lagrangeano com respeito a \mathbf{w} , b , ξ^+ e ξ^- é dado por:

$$\begin{aligned} \frac{\partial L}{\partial b} = 0 &\Rightarrow \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) = 0; \\ \frac{\partial L}{\partial \mathbf{w}} = 0 &\Rightarrow \mathbf{w} = \sum_{i=1}^M (\alpha_i^+ - \alpha_i^-) \mathbf{x}_i; \\ \frac{\partial L}{\partial \xi^+} = 0 &\Rightarrow C - \sum_{i=1}^M \alpha_i^+ = 0; \\ \frac{\partial L}{\partial \xi^-} = 0 &\Rightarrow C - \sum_{i=1}^M \alpha_i^- = 0. \end{aligned} \quad (3.20)$$

Das Equações (3.18), (3.19) e (3.20), resulta o problema dual:

$$\max_{\alpha^+, \alpha^-} \left[-\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M (\alpha_i^+ - \alpha_i^-) (\alpha_j^+ - \alpha_j^-) (\mathbf{x}_i^T \mathbf{x}_j) + \sum_{i=1}^M y_i (\alpha_i^+ - \alpha_i^-) \right], \quad (3.21)$$

com as restrições:

$$\begin{cases} 0 \leq \alpha_i^+ \leq C, & i = 1, \dots, M; \\ 0 \leq \alpha_i^- \leq C, & i = 1, \dots, M; \\ \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) = 0. \end{cases} \quad (3.22)$$

O problema de otimização pode ser simplificado, adotando $\alpha_i = \alpha_i^+ - \alpha_i^-$, assumindo a seguinte formulação:

$$\min_{\alpha} \left[\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M \alpha_i \alpha_j (\mathbf{x}_i^T \mathbf{x}_j) - \sum_{i=1}^M \alpha_i y_i \right], \quad (3.23)$$

sujeito a:

$$\begin{cases} -C \leq \alpha_i \leq C, & i = 1, \dots, M; \\ \sum_{i=1}^M \alpha_i = 0. \end{cases} \quad (3.24)$$

A função de regressão terá:

$$\begin{cases} \mathbf{w} = \sum_{i=1}^M \alpha_i \mathbf{x}_i; \\ b = -\frac{1}{2}((\mathbf{w}^T \mathbf{x}_r) + (\mathbf{w}^T \mathbf{x}_s)), \end{cases} \quad (3.25)$$

sendo \mathbf{x}_r e \mathbf{x}_s os vetores de suporte de cada lado do hiperplano, satisfazendo,

$$\alpha^+, \alpha^- > 0, \quad y_r > 0, \quad y_s < 0. \quad (3.26)$$

Logo, os vetores de suporte (SVs - *Support Vectors*) serão aqueles pontos nos quais $\alpha^+, \alpha^- > 0$.

- **Função-custo ε -insensível**

Esta função não penaliza as amostras que se encontram até uma distância ε do hiperplano e para distâncias maiores mantém as características da função-custo de Laplace.

$$\xi_i = \begin{cases} |y_i - f(\mathbf{x}_i)| - \varepsilon, & \text{se } |y_i - f(\mathbf{x}_i)| \geq \varepsilon; \\ 0, & \text{caso contrário.} \end{cases} \quad (3.27)$$

De forma análoga ao procedimento descrito para a função custo de Laplace, obtém-se o seguinte problema dual:

$$\begin{aligned} \max_{\alpha^+, \alpha^-} & \left[-\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M (\alpha_i^- - \alpha_i^+) (\alpha_j^- - \alpha_j^+) (\mathbf{x}_i^T \mathbf{x}_j) \right. \\ & \left. + \sum_{i=1}^M (\alpha_i^+ - \alpha_i^-) y_i - \sum_{i=1}^M (\alpha_i^+ + \alpha_i^-) \varepsilon \right], \end{aligned} \quad (3.28)$$

$$\begin{cases} 0 \leq \alpha_i^+ \leq C, \quad i = 1, \dots, M; \\ 0 \leq \alpha_i^- \leq C, \quad i = 1, \dots, M; \\ \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) = 0. \end{cases} \quad (3.29)$$

Solucionando o problema (3.28), com as restrições (3.29), podemos determinar os multiplicadores de Lagrange, α_i^+, α_i^- , e a função de regressão, onde:

$$\begin{cases} \mathbf{w} = \sum_{i=1}^M (\alpha_i^+ - \alpha_i^-) \mathbf{x}_i; \\ b = -\frac{1}{2}((\mathbf{w}^T \mathbf{x}_r) + (\mathbf{w}^T \mathbf{x}_s)). \end{cases} \quad (3.30)$$

As condições de otimalidade de Karush-Khun-Tucker (KKT) que são satisfeitas nesta solução são:

$$\alpha_i^+ \alpha_i^- = 0, \quad i = 1, \dots, M. \quad (3.31)$$

Portanto, os vetores de suporte vão corresponder exatamente às amostras para as quais um dos multiplicadores de Lagrange é maior que zero.

- **Função-custo Quadrática**

Esta função-custo corresponde ao critério convencional de minimização do erro quadrático.

$$\xi_i = (y_i - f(\mathbf{x}_i))^2. \quad (3.32)$$

Usando a função-custo quadrática o problema de otimização pode ser escrito como

$$\begin{aligned} \max_{\alpha^+, \alpha^-} & \left[-\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M (\alpha_i^- - \alpha_i^+) (\alpha_j^- - \alpha_j^+) (\mathbf{x}_i^T \mathbf{x}_j) \right. \\ & \left. + \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) y_i - \frac{1}{2C} \sum_{i=1}^M ((\alpha_i^+)^2 + (\alpha_i^-)^2) \right], \end{aligned} \quad (3.33)$$

sujeito a

$$\begin{cases} 0 \leq \alpha_i^+ \leq C, i = 1, \dots, M; \\ 0 \leq \alpha_i^- \leq C, i = 1, \dots, M; \\ \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) = 0. \end{cases} \quad (3.34)$$

Fazendo $\alpha_i = \alpha_i^+ - \alpha_i^-$, na equação (3.33), pode-se realizar uma simplificação via condição de KKT, o que implica $\alpha_i^2 = (\alpha_i^+)^2 + (\alpha_i^-)^2$. O problema de otimização resultante será dado por:

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M \alpha_i \alpha_j (\mathbf{x}_i^T \mathbf{x}_j) - \sum_{i=1}^M \alpha_i y_i + \frac{1}{2C} \sum_{i=1}^M \alpha_i^2, \quad (3.35)$$

sujeito a:

$$\begin{cases} -C \leq \alpha_i \leq C, i = 1, \dots, M; \\ \sum_{i=1}^M \alpha_i = 0. \end{cases} \quad (3.36)$$

A função de regressão terá:

$$\begin{cases} \mathbf{w} = \sum_{i=1}^M \alpha_i \mathbf{x}_i; \\ b = -\frac{1}{2} ((\mathbf{w}^T \mathbf{x}_r) + (\mathbf{w}^T \mathbf{x}_s)). \end{cases} \quad (3.37)$$

- **Função-custo ε -Quadrática**

Esta função não penaliza as amostras que se encontram até uma distância ε do hiperplano e para distâncias maiores mantém as características da função-

custo Quadrática.

$$\xi_i = \begin{cases} (|y_i - f(\mathbf{x}_i)| - \varepsilon)^2, & \text{se } |y_i - f(\mathbf{x}_i)| \geq \varepsilon; \\ 0, & \text{caso contrário.} \end{cases} \quad (3.38)$$

Usando a função-custo ε -Quadrática a formulação assume a configuração

$$\begin{aligned} \max_{\alpha^+, \alpha^-} & \left[-\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M (\alpha_i^- - \alpha_i^+) (\alpha_j^- - \alpha_j^+) (\mathbf{x}_i^T \mathbf{x}_j) \right. \\ & \left. + \sum_{i=1}^M \alpha_i^- (y_i - \varepsilon) - \alpha_i^+ (y_i + \varepsilon) \right] - \frac{1}{2C} \sum_{i=1}^M ((\alpha_i^+)^2 - (\alpha_i^-)^2), \end{aligned} \quad (3.39)$$

sujeito a:

$$\begin{cases} 0 \leq \alpha_i^+ \leq C, & i = 1, \dots, M; \\ 0 \leq \alpha_i^- \leq C, & i = 1, \dots, M; \\ \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) = 0, \end{cases} \quad (3.40)$$

e a solução é análoga a desenvolvida para a função-custo quadrática.

• Função-custo de Huber

A função-custo de Huber tem propriedades desejáveis quando a distribuição dos dados é desconhecida, pois possui uma região de minimização do erro quadrático e uma região com pouca sensibilidade a *outliers*.

$$\xi_i = \begin{cases} \frac{1}{2\gamma} (y_i - f(\mathbf{x}_i))^2, & \text{se } |y_i - f(\mathbf{x}_i)| < \gamma C; \\ C|y_i - f(\mathbf{x}_i)| - \frac{\gamma C^2}{2}, & \text{caso contrário.} \end{cases} \quad (3.41)$$

em que γ é um parâmetro livre, definido pelo usuário, que controla a forma da função-custo.

O problema de otimização será:

$$\begin{aligned} \max_{\alpha^+, \alpha^-} & \left[-\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M (\alpha_i^- - \alpha_i^+) (\alpha_j^- - \alpha_j^+) (\mathbf{x}_i^T \mathbf{x}_j) \right. \\ & \left. + \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) y_i - \frac{\gamma}{2} \sum_{i=1}^M ((\alpha_i^+)^2 - (\alpha_i^-)^2) \right], \end{aligned} \quad (3.42)$$

sujeito a:

$$\begin{cases} 0 \leq \alpha_i^+ \leq C, & i = 1, \dots, M; \\ 0 \leq \alpha_i^- \leq C, & i = 1, \dots, M; \\ \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) = 0. \end{cases} \quad (3.43)$$

A solução é obtida de maneira análoga ao procedimento descrito para a função-custo quadrática.

• **Função-custo ε -Huber**

Esta função não penaliza as amostras que se encontram até uma distância ε do hiperplano e para distâncias maiores mantém as características da função-custo de Huber.

$$\xi_i = \begin{cases} 0, & \text{se } |y_i - f(\mathbf{x}_i)| \leq \varepsilon; \\ \frac{1}{2\gamma}(|y_i - f(\mathbf{x}_i)| - \varepsilon)^2, & \text{se } \varepsilon < |y_i - f(\mathbf{x}_i)| < \gamma C; \\ C(|y_i - f(\mathbf{x}_i)| - \varepsilon) - \frac{\gamma C^2}{2}, & \text{caso contrário.} \end{cases} \quad (3.44)$$

A formulação assume a seguinte configuração:

$$\begin{aligned} & \max_{\alpha^+, \alpha^-} \left[-\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M (\alpha_i^- - \alpha_i^+) (\alpha_j^- - \alpha_j^+) (\mathbf{x}_i^T \mathbf{x}_j) \right. \\ & \left. + \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) y_i - \sum_{i=1}^M (\alpha_i^+ + \alpha_i^-) \varepsilon - \frac{\gamma}{2} \sum_{i=1}^M ((\alpha_i^+)^2 - (\alpha_i^-)^2) \right], \end{aligned} \quad (3.45)$$

com as restrições:

$$\begin{cases} 0 \leq \alpha_i^+ \leq C, \quad i = 1, \dots, M; \\ 0 \leq \alpha_i^- \leq C, \quad i = 1, \dots, M; \\ \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) = 0. \end{cases} \quad (3.46)$$

O solução é obtida de maneira análoga ao procedimento descrito para a função-custo quadrática.

Uma observação importante é que para problemas de regressão, é necessário selecionar adequadamente tanto a função-custo quanto o parâmetro que controla a capacidade de modelagem. Estas considerações devem ser baseadas em conhecimentos *a priori* acerca do problema e da distribuição do ruído. Na ausência desta informação, a função-custo de Huber apresenta-se como boa alternativa [53]. Vapnik [53] propôs a função-custo ε -insensível como um *trade-off* entre as funções-custo de Laplace e Huber e funções-custo que favorecem a redução do número de vetores de suporte. Entretanto, sua implementação é mais custosa computacionalmente e a região de ε -insensibilidade pode apresentar problemas.

Outra observação importante é que o número de vetores de suporte que são utilizados para construir o regressor depende da precisão desejada ε . Quanto menor a precisão de aproximação, isto é, quanto maior o valor de ε , menos vetores de suporte serão necessários. Com relação à capacidade de generalização esta também pode ser afetada, podendo diminuir ou não com o aumento do valor de ε .

Uma última observação é que a aplicação de SVM a dados ruidosos deve levar em consideração alguma técnica para a estimação da variância do ruído, uma vez que a variância do ruído maior que a precisão ε faz com que o número de vetores de suporte aumente drasticamente.

3.2.2 Regressão não-linear

Um modelo não-linear se torna necessário quando o modelo linear deixa de ser adequado para representar o mapeamento ($\mathbf{x} \in \mathbb{R}^n \rightarrow y \in \mathbb{R}$) no espaço original. Neste caso, segundo o teorema de Cover [45], o conjunto de treinamento, no espaço original, pode ser mapeado para um espaço de maior dimensionalidade, chamado espaço de características, onde sempre uma regressão linear poderá ser realizada.

Como o espaço de características pode ter dimensão muito elevada, a computação de uma função que mapeie os dados no espaço de características pode ser inviável. Porém, a abordagem baseada em funções *kernel* pode permitir tratar o problema da dimensionalidade [55].

A idéia da função *kernel* é permitir que as operações possam ser executadas no espaço original. Contudo, o esforço computacional ainda será dependente do número de dados de treinamento e, para proporcionar uma boa distribuição de dados em problemas de alta dimensionalidade, será necessário um grande conjunto de treinamento.

Para garantir a convexidade do problema de otimização e que o *kernel* represente um mapeamento no qual seja possível o cálculo de produtos escalares, utilizam-se funções *kernel* que seguem as condições estabelecidas pelo teorema de Mercer [54].

Um *kernel* que atende as condições de Mercer deve gerar matrizes positivas semi-definidas \mathbf{K} , onde cada elemento é definido por $K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$, para $i, j = 1, \dots, M$ [55].

A seguir são apresentadas algumas funções que satisfazem as condições de Mercer, sendo válidas para todo \mathbf{x}_i e \mathbf{x}_j .

1. Polinômios

Para construir uma função polinomial de grau g , é possível usar o seguinte modelo de geração do *kernel*:

$$\begin{cases} K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j)^g & \text{ou} \\ K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j + 1)^g. \end{cases} \quad (3.47)$$

O segundo *kernel* é preferível, pois este evita o problema de nulidade da Hessiana (matriz utilizada em técnicas de otimização envolvendo a função *kernel*).

2. Funções de base radial gaussiana

A Função *kernel* de base radial (RBF - Radial Base Function) é uma função definida positiva para qualquer σ fixo, e tende a zero quando $\|\mathbf{x}_i - \mathbf{x}_j\|$ vai para infinito e σ é finito. A função *kernel* RBF pode ser escrita como:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right). \quad (3.48)$$

Esta função local é atrativa no sentido de que os vetores de suporte contribuem com uma função gaussiana local.

3. Funções de base radial exponencial

Uma função de base radial exponencial tem a forma:

$$K(\mathbf{x}, \mathbf{x}_i) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_i\|}{2\sigma^2}\right). \quad (3.49)$$

O kernel ERBF (Exponential Radial Base Function) produz uma função linear por partes que pode ser atrativa quando se deseja aproximar funções com pouca suavidade.

Outras funções *kernels* podem ser encontradas na literatura [55] e *kernels* mais sofisticados podem ser obtidos pela formação de somatórios de *kernels*, uma vez que a soma de duas funções definidas positivas é definida positiva, produzindo:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \sum_k K_k(\mathbf{x}_i, \mathbf{x}_j). \quad (3.50)$$

Uma observação importante é que há muitos mapeamentos diferentes para um espaço de características, cada um associado a uma função *kernel* que atende ao teorema de Mercer. Para escolher o melhor mapeamento para um problema em particular existem, na literatura, vários métodos de seleção de *kernel*, dentre eles os mais populares são o *bootstrapping* e a *validação cruzada* [45].

Para construir a SVM para o espaço vetorial n -dimensional, $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$, costuma-se recorrer a um produto de *kernels* unidimensionais, denominado produto tensorial:

$$K(\mathbf{x}, \mathbf{x}_i) = \prod_{k=1}^n K(\mathbf{x}_k, \mathbf{x}_{ik}). \quad (3.51)$$

A formulação SVM, agora com *kernel* não-linear e usando a função-custo de Laplace, é dada por:

$$\max_{\alpha^+, \alpha^-} \left[-\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M (\alpha_i^- - \alpha_i^+) (\alpha_j^- - \alpha_j^+) K(\mathbf{x}_i, \mathbf{x}_j) + \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) y_i \right], \quad (3.52)$$

sujeito a:

$$\begin{cases} 0 \leq \alpha_i^+ \leq C, i = 1, \dots, M; \\ 0 \leq \alpha_i^- \leq C, i = 1, \dots, M; \\ \sum_{i=1}^M (\alpha_i^- - \alpha_i^+) = 0. \end{cases} \quad (3.53)$$

Resolvendo a equação (3.52) com suas restrições, obtêm-se os multiplicadores de Lagrange α_i^+ e α_i^- , e a função de regressão será dada por:

$$f(\mathbf{x}) = \sum_{i=1}^M (\alpha_i^+ - \alpha_i^-) K(\mathbf{x}, \mathbf{x}_i) + b, \quad (3.54)$$

em que:

$$\begin{cases} (\mathbf{w}^T \mathbf{x}) = \sum_{i=1}^M (\alpha_i^+ - \alpha_i^-) K(\mathbf{x}_i, \mathbf{x}_j); \\ b = -\frac{1}{2} \sum_{i=1}^M (\alpha_i^+ - \alpha_i^-) K(\mathbf{x}_i, \mathbf{x}_r) + K(\mathbf{x}_i, \mathbf{x}_s). \end{cases} \quad (3.55)$$

A restrição de igualdade pode ser abandonada se o *kernel* contiver um termo de polarização b . Sendo assim, a função de regressão passa a ser dada por:

$$f(\mathbf{x}) = \sum_{i=1}^M (\alpha_i^+ - \alpha_i^-) K(\mathbf{x}, \mathbf{x}_i). \quad (3.56)$$

A formulação SVM para outras funções-custo pode ser obtida de modo análogo, substituindo o produto interno por uma função *kernel*.

3.3 SVM do Tipo Proximal

Fung e Mangasarian [61] propuseram uma modificação do SVM tradicional, chamada “*Proximal Support Vector Machine*” (PSVM). A inequação de restrição, presente no SVM tradicional, é substituída por uma igualdade, fato que muda inteiramente a natureza dos hiperplanos de suporte, que agora passam a se chamar hiperplanos proximais.

Outra mudança observada no PSVM em relação ao SVM tradicional é a aproximação do hiperplano separador à origem, por meio da minimização da polarização b ao quadrado.

O parâmetro de generalização C , usado no SVM tradicional, continua a ser usado para ponderar a parcela relativa ao erro de aproximação dentro da função objetivo.

- **Formulação Matemática para o caso linear**

O algoritmo PSVM procura solucionar o seguinte problema:

$$\min_{\mathbf{w}, b} \left\{ \frac{1}{2} (\mathbf{w}^T \mathbf{w} + b^2) \right\}. \quad (3.57)$$

De forma análoga ao SVM tradicional, deve-se introduzir uma função-custo associada à estimativa do ruído presente nas amostras de treinamento. Con-

siderando a função de penalidade na forma proposta por Cortes & Vapnik [60],

$$L(\xi) = \sum_{i=1}^M \xi_i, \quad (3.58)$$

a funcional a ser minimizada será

$$\min_{\mathbf{w}, b, \xi} \left\{ \frac{1}{2} (\mathbf{w}^T \mathbf{w} + b^2) + \frac{1}{2} C \boldsymbol{\xi}^T \boldsymbol{\xi} \right\}. \quad (3.59)$$

Observe que o vetor de erro da função-custo ($\boldsymbol{\xi}$) pode ter elementos positivos ou negativos, ao contrário do que ocorre nos regressores SVM tradicionais. Isto simplifica a determinação do problema dual, uma vez que não é necessário atribuir um novo multiplicador de Lagrange para cada elemento do vetor de erros. A solução encontrada para este problema, após a aplicação dos multiplicadores de Lagrange ($\boldsymbol{\alpha}$) e das condições de Karush-Kuhn-Tucker (KKT), considerando uma função-custo quadrática, é dada pelas equações:

$$\begin{cases} \boldsymbol{\alpha} = \left(\frac{\mathbf{I}}{C} + \mathcal{H} \mathcal{H}^T \right)^{-1} \mathbf{y}; \\ \mathbf{w} = \mathbf{X} \boldsymbol{\alpha}; \\ b = \mathbf{y}^T \boldsymbol{\alpha}; \\ \boldsymbol{\xi} = \frac{\boldsymbol{\alpha}}{C}, \end{cases} \quad (3.60)$$

em que $\mathcal{H} = [\mathbf{X}^T \quad -\mathbf{1}]$, $\mathbf{1}$ é um vetor de 1's e $\mathbf{X} = [\mathbf{x}_1 \quad \cdots \quad \mathbf{x}_M]$ é a matriz de amostras de treinamento. Note que o vetor de erro é uma versão proporcional do vetor que contém os multiplicadores de Lagrange. Logo, serão considerados como vetores-suporte toda as observações presentes em \mathbf{X} associadas a um erro não nulo ($\xi_i \neq 0 \leftrightarrow \alpha_i \neq 0, 1 \leq i \leq M$).

A vantagem desta técnica em comparação com o SVM tradicional é a rapidez para obtenção da solução de forma analítica, por meio da solução de um sistema de equações lineares.

- **Formulação Matemática para o caso não-linear**

O algoritmo PSVM para o caso não-linear foi desenvolvido em [61] substituindo $\mathbf{X}^T \mathbf{w}$ por $\mathbf{K}(\mathbf{x}, \mathbf{x}_i) \alpha$ e $\mathbf{w}^T \mathbf{w}$ por $\alpha^T \alpha$. Na prática, isto implicou na substituição da matriz \mathbf{X} pela matriz de *kernel* $\mathbf{K}(\mathbf{x}, \mathbf{x}_i)$ e na substituição da matriz $\mathbf{X}^T \mathbf{X}$ pela matriz $\mathbf{K}^T(\mathbf{x}, \mathbf{x}_i) \mathbf{K}(\mathbf{x}, \mathbf{x}_i)$, na obtenção do vetor de multiplicadores de Lagrange. Logo, o vetor de multiplicadores de Lagrange para o PSVM não-linear continua sendo calculado pela Equação (3.60), porém fazendo $\mathcal{H} = [\mathbf{K}^T(\mathbf{x}, \mathbf{x}_i) \quad -\mathbf{1}]$. O valor de b continua sendo calculado como no caso linear e $\mathbf{K}^T(\mathbf{x}, \mathbf{x}_i)$ obtida durante o treinamento.

3.4 SVM do Tipo Proximal Não-Polarizado

Um classificador PSVM não-polarizado (UPSVM) foi proposto por Campos & Bastos [49] tanto para o caso linear quanto para o caso não-linear. Contudo, sua extensão como regressor ainda não havia sido analisada. A motivação para sua derivação está baseada na obtenção de um hiperplano separador não-polarizado, independente da escolha do parâmetro de regularização C .

- **Formulação Matemática do UPSVM Linear**

A principal diferença entre o PSVM e o UPSVM está em retirar o parâmetro de polarização b da função-objetivo dada pela Equação (3.59).

Desta forma o algoritmo UPSVM procura solucionar o seguinte problema:

$$\min_{\mathbf{w}, \xi} \left\{ \frac{1}{2} (\mathbf{w}^T \mathbf{w}) + \frac{1}{2} C \xi^T \xi \right\}. \quad (3.61)$$

O problema em questão continua a ser do tipo proximal, pois, devido à restrição imposta pela função-custo (ξ) à Equação (3.61), os hiperplanos obtidos e definidos em [61] como hiperplanos proximais são aqueles em torno dos quais as amostras se concentram em cada lado da função $f(\mathbf{x})$. Portanto, diferentemente do que ocorre no SVM tradicional com os chamados hiperplanos de suporte, os hiperplanos proximais não confinam necessariamente a maior parte das amostras.

A solução deste problema, após a aplicação dos multiplicadores de Lagrange (α) e das condições de KKT, considerando uma função-custo quadrática, será:

$$\begin{cases} b = \frac{\mathbf{1}^T (\frac{1}{C} + \mathbf{X}\mathbf{X}^T)^{-1} \mathbf{y}}{\mathbf{1}^T (\frac{1}{C} + \mathbf{X}\mathbf{X}^T)^{-1} \mathbf{1}}; \\ \alpha = (\frac{1}{C} + \mathbf{X}\mathbf{X}^T)^{-1} (\mathbf{y} - \mathbf{1}b); \\ \mathbf{w} = \mathbf{X}\alpha, \end{cases} \quad (3.62)$$

que é similar ao regressor *LS-SVM* proposto por Suykens em [52] para a função-custo quadrática.

- **Formulação Matemática do UPSVM Não-Linear**

A generalização do UPSVM linear para a obtenção de uma versão não-linear é realizada de maneira análoga ao que foi feito para a generalização do problema SVM original, ou seja, substituindo em todas as equações o *kernel* linear $\mathbf{X}^T \mathbf{X}$ por uma matriz $\mathbf{K}(\mathbf{x}, \mathbf{x}_i)$ (*kernel* não-linear).

Observe que não há necessidade de se conhecer o conjunto de transformações não-lineares, mas apenas a matriz *kernel* $\mathbf{K}(\mathbf{x}, \mathbf{x}_i)$.

3.5 Regressor SVM no Plano Complexo

Nas aplicações em comunicações, os sinais usados são, em sua maioria, representados por valores complexos (\mathbb{C}). Surge, então, a necessidade do uso de uma versão de valores complexos (\mathbb{C}) para as Máquinas de Vetores de Suporte (SVM).

Um Regressor SVM Linear de valor complexo foi proposto por Cristodoulou em [58] para o SVM tradicional. Como no regressor real, a funcional de risco consiste de dois termos: um referente ao risco estrutural (que compreende o limite na complexidade do modelo resultante), e um associado ao risco empírico (medido diretamente nos erros por uma função custo apropriada). Como os erros são valores complexos, torna-se conveniente minimizar simultaneamente as suas partes real e imaginária. Desta forma, torna-se equivalente a minimizar:

$$L(\mathbf{w}, b, \xi^+, \xi^-, \zeta^+, \zeta^-) = \frac{1}{2}(\mathbf{w}^T \mathbf{w}) + C \left(\sum_{i=1}^M \xi_i^+ + \sum_{i=1}^M \xi_i^- \right) + C \left(\sum_{i=1}^M \zeta_i^+ + \sum_{i=1}^M \zeta_i^- \right), \quad (3.63)$$

em que ξ^+ e ξ^- são as funções-custo da parte real dos erros positivos e negativos, respectivamente. De forma análoga, ζ^+ e ζ^- são as funções-custo da parte imaginária dos erros.

A solução deste problema, após a aplicação dos multiplicadores de Lagrange e das condições de KKT, será:

$$\mathbf{w} = \sum_{i=1}^M \psi_i^* \mathbf{x}_i^*, \quad (3.64)$$

sendo $\psi_i = \alpha_i^+ - \alpha_i^- + j(\beta_i^+ - \beta_i^-)$ e $(\cdot)^*$ o conjugado de um valor complexo. Este resultado é análogo ao problema do regressor SVM de valor real, exceto que neste caso são considerados os multiplicadores de Lagrange α_i e β_i para as componentes real e imaginária.

A norma dos coeficientes de valor complexo pode ser escrita como

$$\|\mathbf{w}\|^2 = \sum_{i=1}^M \sum_{j=1}^M \psi_i^* \psi_j \mathbf{x}_i^H \mathbf{x}_j = \Psi^H \mathbf{R} \Psi, \quad (3.65)$$

sendo $\mathbf{R}_{ij} = \mathbf{x}_i^H \mathbf{x}_j$ e $(\cdot)^H$ o conjugado transposto de um vetor ou matriz.

Substituindo a Equação (3.64) na funcional dual obtido a partir de (3.63), e procedendo de forma análoga ao caso de valor real, obtém-se a funcional dual a ser minimizada como:

$$L_d = -\frac{1}{2} \Psi^H \mathbf{R} \Psi + \Re[\Psi^H \mathbf{y}] - (\alpha^+ + \alpha^- + \beta^+ + \beta^-) \mathbf{1} \varepsilon. \quad (3.66)$$

A Expressão (3.66) é formalmente igual à expressão da funcional do Regressor SVM para valores reais. Esta forma quadrática pode ser resolvida usando as mesmas estratégias utilizadas para o SVM de valor real.

Neste capítulo foi apresentado os fundamentos de SVM e suas variações *fast-SVM* (PSVM e UPSVM) para aplicações como regressores, além da descrição de uma abordagem de SVM como regressor de valores complexos.

Capítulo 4

Regressor PSVM de valores complexos

Neste capítulo iremos descrever nossa primeira proposta para a estimação de canal adequada aos sistemas PLC.

O vetor de sinais piloto recebido na amostra (m) pode ser escrito a partir da Equação (2.1) como:

$$\mathbf{Y}_p(m) = \mathbf{D}_{\{\mathbf{H}_p(m)\}} \mathbf{X}_p(m) + \mathbf{V}_p(m), \quad (4.1)$$

sendo:

$$\begin{aligned} \mathbf{Y}_p(m) &= [Y_1(m), \dots, Y_P(m)]^T; \\ \mathbf{H}_p(m) &= [H_1(m), \dots, H_P(m)]^T; \\ \mathbf{X}_p(m) &= [X_1(m), \dots, X_P(m)]^T; \\ \mathbf{V}_p(m) &= [V_1(m), \dots, V_P(m)]^T, \end{aligned} \quad (4.2)$$

$\mathbf{D}_{\{\}}$ representa a matriz cujos elementos não-nulos se encontram na diagonal principal, $\{1, \dots, P\}$ são as posições dos tons pilotos e equivalem as posições $\{1, 1PS, \dots, (P-1)PS\}$ no símbolo DMT, sendo PS a separação entre pilotos e m é um índice temporal dos símbolos OFDM.

Nesta técnica, o desenvolvimento proposto em [58] para SVM com valores complexos é aplicado ao algoritmo *Proximal Support Vector Machine* (PSVM), proposto em [61], visando sua utilização como regressor. O objetivo desta análise é obter um estimador baseado em algoritmo *fast-SVM* que possa ser aplicado diretamente em números complexos e seja capaz de minimizar, simultaneamente, as partes real e imaginária dos erros de estimação.

A ideia básica de um regressor SVM, como visto na Seção 3.2, é encontrar uma função que represente os dados, neste caso $(\mathbf{X}_p(m), \mathbf{Y}_p(m))$, com um desvio menor ou igual a um determinado valor (parâmetro de tolerância ao erro, dado a partir

da limitação da margem) para cada par $(X_p(m), Y_p(m))$ e, ao mesmo tempo, tenha norma mínima $\|\mathbf{H}_p(m)\|$.

Em [61] os autores adicionaram um termo b^2 na função objetivo a ser minimizada, de acordo com suas análises, este fato aumenta a convexidade da função e assegura que a solução possua um único mínimo global. Em adição, é minimizado $\|[\mathbf{H}_p; b]\|$, ao invés de $\|\mathbf{H}_p\|$, como ocorre no SVM tradicional.

4.1 Função-custo

No entanto, em situações práticas, pode não ser possível forçar todos os erros serem menores ou iguais a este parâmetro (ε , no caso da função-custo ε -insensível). Para resolver este problema, um termo relacionado ao erro empírico deve ser incluído a partir de uma função-custo do modelo de resíduos apropriada. Uma função-custo que adicionalmente considera um intervalo de custo quadrático é a função-custo de ε -Huber, que neste cenário pode ser reescrita como:

$$\mathcal{L}(V_p(m)) = \begin{cases} 0, & \text{se } |V_p(m)| \leq \varepsilon; \\ \frac{1}{2\gamma}(|V_p(m)| - \varepsilon)^2, & \text{se } \varepsilon < |V_p(m)| \leq \gamma C + \varepsilon; \\ C(|V_p(m)| - \varepsilon) - \frac{1}{2}\gamma C^2, & \text{se } |V_p(m)| > \gamma C + \varepsilon, \end{cases} \quad (4.3)$$

em que γ e C são parâmetros que controlam o formato da função-custo. A Figura 4.1 ilustra algumas amostras $(X_p(m), Y_p(m))$ e a função-custo de ε -Huber, no qual $\xi_p(m)$ é função do erro na parte real do valor de saída. A Figura 4.2 mostra as regiões da função-custo de ε -Huber.

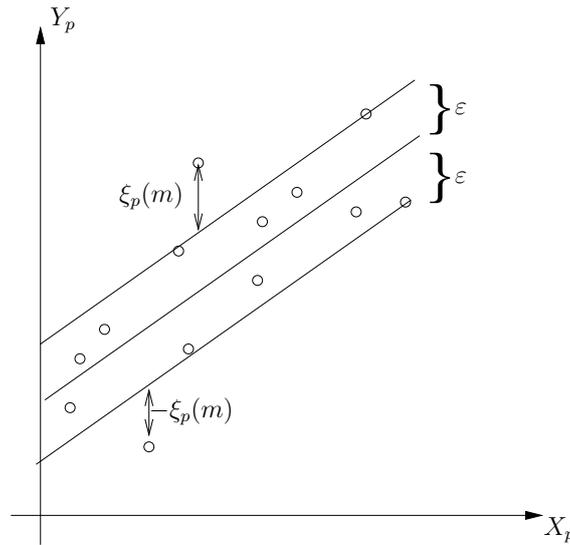


Figura 4.1: Amostras $(X_p(m), Y_p(m))$ e a função-custo de ε -Huber.

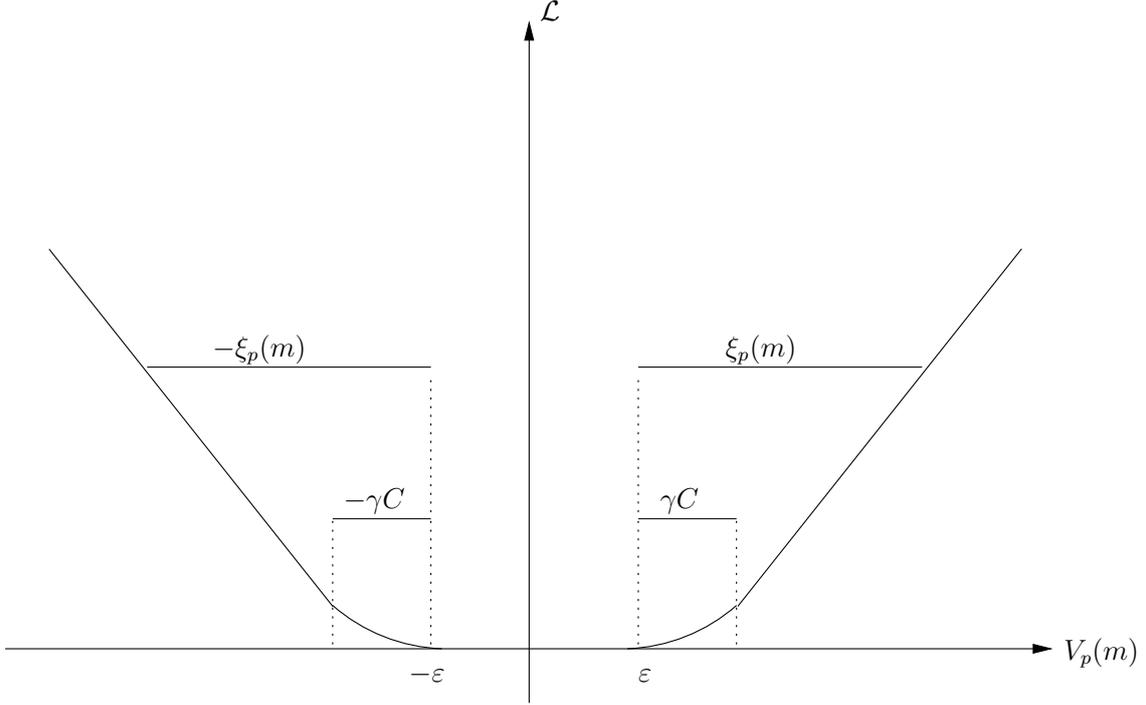


Figura 4.2: Regiões ε -insensível, de custo quadrático e linear.

Em redes de energia elétrica onde o ruído aditivo não pode ser modelado como AWGN essa função-custo pode ser vista como uma boa alternativa, uma vez que o valor de $\gamma C + \varepsilon$ pode ser ajustado de modo que o intervalo quadrático da função-custo possa englobar grande parte do erro causado pelo ruído de fundo e o intervalo linear considere o erro causado pelo ruído impulsivo como *outliers*, mostrando-se pouco sensível a este tipo de ruído.

Logo, é necessário um conjunto de M pares de observação $(X_p(m), Y_p(m))$ com $m = 1, \dots, M$ e a funcional a ser minimizada, para cada posição piloto no símbolo OFDM, torna-se:

$$\frac{1}{2} \|\tilde{\mathbf{H}}_p\|^2 + \sum_{m=1}^M (\mathcal{L}(\Re(V_p(m))) + \mathcal{L}(\Im(V_p(m)))) \quad (4.4)$$

sendo $\tilde{\mathbf{H}}_p = [H_p(1), \dots, H_p(M)]^T$ um vetor contendo o conjunto de M respostas em frequência do canal para uma posição de sinal piloto.

Separando os dados de acordo com os seguintes conjuntos de índices:

$$\begin{aligned} I_1 &= \{(X_p(m), Y_p(m)); \text{ se } |\xi_p(m)| \leq \gamma C\}; \\ I_2 &= \{(X_p(m), Y_p(m)); \text{ se } |\xi_p(m)| > \gamma C\}; \\ I_3 &= \{(X_p(m), Y_p(m)); \text{ se } |\zeta_p(m)| \leq \gamma C\}; \\ I_4 &= \{(X_p(m), Y_p(m)); \text{ se } |\zeta_p(m)| > \gamma C\}, \end{aligned} \quad (4.5)$$

em que $\zeta_p(m)$ é função do erro na parte imaginária do valor de saída (veja Fig. 4.3).

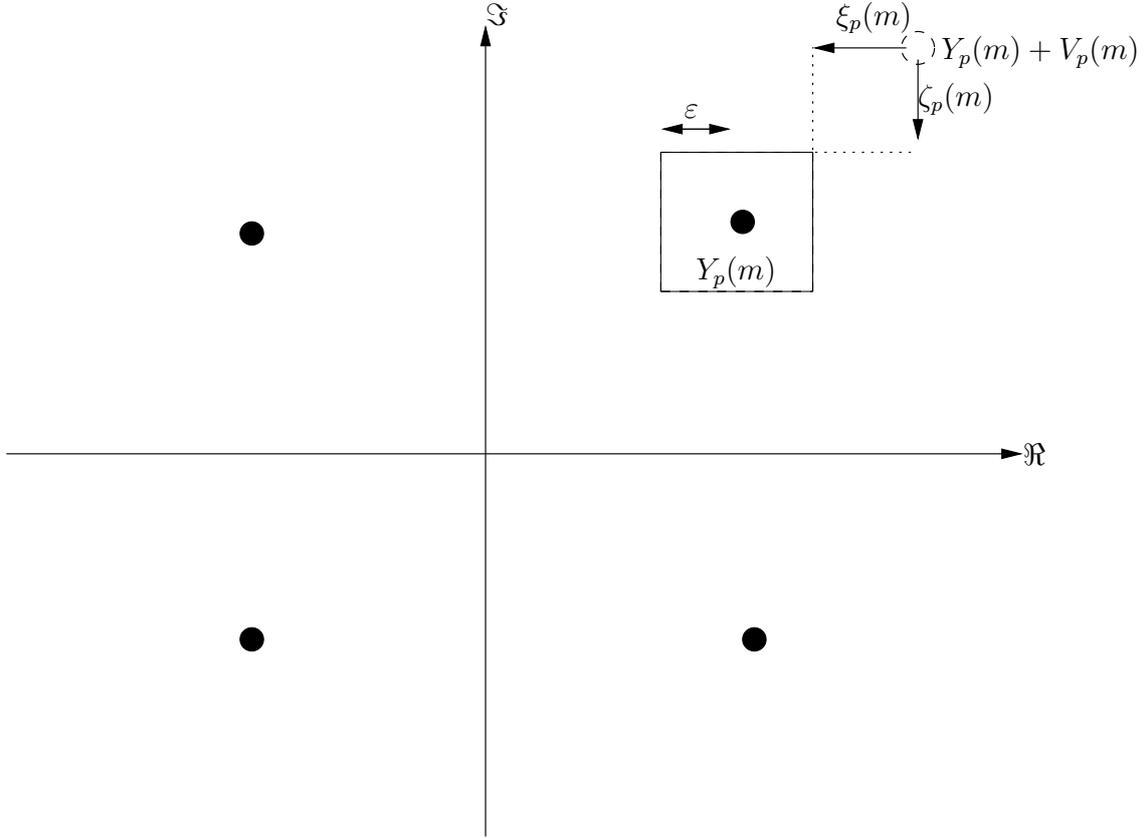


Figura 4.3: Amostra de valor complexo $Y_p(m)$.

Podemos reescrever a Equação 4.4 e, de forma equivalente, minimizar:

$$\begin{aligned} & \frac{1}{2} \|\tilde{\mathbf{H}}_p\|^2 + \frac{1}{2\gamma} \sum_{m \in \mathbf{I}_1} (\xi_p^2(m)) + C \sum_{m \in \mathbf{I}_2} (\xi_p(m)) \\ & + \frac{1}{2\gamma} \sum_{m \in \mathbf{I}_3} (\zeta_p^2(m)) + C \sum_{m \in \mathbf{I}_4} (\zeta_p(m)) - \sum_{m \in \mathbf{I}_2, \mathbf{I}_4} (\gamma C^2), \end{aligned} \quad (4.6)$$

sujeito a:

$$\begin{cases} \Re(Y_p(m) - (H_p^*(m)X_p(m))) = \varepsilon + \xi_p(m); \\ \Im(Y_p(m) - (H_p^*(m)X_p(m))) = \varepsilon + \zeta_p(m), \end{cases} \quad (4.7)$$

sendo que \Re e \Im representam as partes real e imaginária de um número complexo, respectivamente.

4.2 Algoritmo PSVM adaptativo

Uma importante mudança em relação ao SVM tradicional é o fato de que a desigualdade nas restrições foi substituída por uma igualdade na Equação (4.7), sem perda de generalidade. Esta alteração permite que o problema de otimização quadrática presente no SVM tradicional passe a ser visto como um sistema de equações lineares nos algoritmos *fast-SVM*. Uma outra consequência desta mudança nos algoritmos *fast-SVM* é que o vetor de erros pode ter elementos positivos e negativos, diferente do que ocorre nos algoritmos SVM clássicos. Isto simplifica a determinação do problema dual, uma vez que não é necessário especificar um vetor de multiplicadores de Lagrange para os elementos positivos do vetor de erros e outro para os elementos negativos.

De maneira usual, na metodologia SVM é possível transformar a minimização da funcional primal¹ da expressão (4.6), sujeita às restrições da Equação (4.7), na otimização da funcional dual². Primeiro, introduzem-se as restrições na funcional primal por meio dos Multiplicadores de Lagrange, então se verificam as condições de KKT (Karush-Kuhn-Tucker) [51]. A nova funcional deve ser minimizada com relação às variáveis primais e maximizada considerando as variáveis duais (problema Minimax).

Com o intuito de se obter uma implementação adaptativa, tal que a atualização em tempo real seja possível, foi adotado nesta tese um procedimento similar ao realizado em [63], que consiste na inclusão de um fator de janelamento exponencial μ^{m-i} que multiplica os erros ($\xi_{p,i}(m)$, $\zeta_{p,i}(m)$ e $Y_{p,i}(m) - (H_p^*(m)X_{p,i}(m))$) de modo que as amostras passadas, indicadas pelo aumento do índice (i), gradualmente percam sua importância a medida que elas se afastem temporalmente da amostra atual (m). Para se obter uma estimativa de $\mathbf{H}_p(m)$ robusta ao ruído foi usado um vetor de pesos $\mathbf{a}_{p,i}(m-1)$.

No desenvolvimento com valores complexos, a funcional primal-dual para o regressor PSVM, em cada instante de tempo (m) e para cada símbolo piloto (p), pode ser escrita como:

¹Funcional cujas variáveis a serem minimizadas fazem parte do problema original com restrições.

²Funcional cujas variáveis a serem minimizadas são os multiplicadores de Lagrange do problema primal [62].

$$\begin{aligned}
L_{pd}(m) &= \frac{1}{2} \|\tilde{\mathbf{H}}_p(m)\|^2 + \frac{1}{2\gamma} \sum_{i \in I_1} \mu^{m-i} (\xi_{p,i}^2(m)) + C \sum_{i \in I_2} \mu^{m-i} (\xi_{p,i}(m)) \\
&+ \frac{1}{2\gamma} \sum_{i \in I_3} \mu^{m-i} (\zeta_{p,i}^2(m)) + C \sum_{i \in I_4} \mu^{m-i} (\zeta_{p,i}(m)) - \sum_{i \in I_2, I_4} (\gamma C^2) + \frac{1}{2} b_p^2(m) \\
&\quad - \sum_{i=1}^m \mu^{m-i} (\lambda_{p,i}(m) \xi_{p,i}(m)) - \sum_{i=1}^m \mu^{m-i} (\eta_{p,i}(m) \zeta_{p,i}(m)) \\
&+ \sum_{i=1}^m \mu^{m-i} \alpha_{p,i}(m) [\mathfrak{R}(a_{p,i}(m-1)(Y_{p,i}(m) - (H_p^*(m)X_{p,i}(m)) - b_p(m))) \\
&- \varepsilon - \xi_{p,i}(m)] + \sum_{i=1}^m \mu^{m-i} \beta_{p,i}(m) [\mathfrak{S}(a_{p,i}(m-1)(Y_{p,i}(m) - (H_p^*(m)X_{p,i}(m)) \\
&\quad - b_p(m))) - \varepsilon - \zeta_{p,i}(m)],
\end{aligned} \tag{4.8}$$

em que os multiplicadores de Lagrange são introduzidos para cada restrição do problema primal. As condições de KKT [51] necessárias e suficientes para otimalidade no problema de restrição de igualdade são obtidas igualando a zero as derivadas parciais de $L_{pd}(m)$ em relação a $(H_p(m), b_p(m), \xi_{p,i}(m), \zeta_{p,i}(m), \alpha_{p,i}(m), \beta_{p,i}(m))$.

$$\left\{ \begin{aligned}
\frac{\partial L_{pd}(m)}{\partial H_p(m)} &= 0 \rightarrow H_p(m) = \sum_{i=1}^m \mu^{m-i} a_{p,i}(m-1) [\alpha_{p,i}(m) - j\beta_{p,i}(m)] X_{p,i}(m); \\
\frac{\partial L_{pd}(m)}{\partial b_p(m)} &= 0 \rightarrow b_p(m) = \sum_{i=1}^m \mu^{m-i} a_{p,i}(m-1) [\alpha_{p,i}(m) + j\beta_{p,i}(m)]; \\
\frac{\partial L_{pd}(m)}{\partial \xi_{p,i}(m)} &= 0 \rightarrow \frac{1}{\gamma} \sum_{i \in I_1} \mu^{m-i} \xi_{p,i}(m) + C \sum_{i \in I_2} \mu^{m-i} - \sum_{i=1}^m \mu^{m-i} \lambda_{p,i}(m) \\
&\quad - \sum_{i=1}^m \mu^{m-i} \alpha_{p,i}(m) = 0; \\
\frac{\partial L_{pd}(m)}{\partial \zeta_{p,i}(m)} &= 0 \rightarrow \frac{1}{\gamma} \sum_{i \in I_3} \mu^{m-i} \zeta_{p,i}(m) + C \sum_{i \in I_4} \mu^{m-i} - \sum_{i=1}^m \mu^{m-i} \eta_{p,i}(m) \\
&\quad - \sum_{i=1}^m \mu^{m-i} \beta_{p,i}(m) = 0; \\
\frac{\partial L_{pd}(m)}{\partial \alpha_{p,i}(m)} &= 0 \rightarrow \sum_{i=1}^m \mu^{m-i} [\mathfrak{R}(a_{p,i}(m-1)(Y_{p,i}(m) - (H_p^*(m)X_{p,i}(m)) \\
&\quad - b_p(m))) - \varepsilon - \xi_{p,i}(m)] = 0; \\
\frac{\partial L_{pd}(m)}{\partial \beta_{p,i}(m)} &= 0 \rightarrow \sum_{i=1}^m \mu^{m-i} [\mathfrak{S}(a_{p,i}(m-1)(Y_{p,i}(m) - (H_p^*(m)X_{p,i}(m)) \\
&\quad - b_p(m))) - \varepsilon - \zeta_{p,i}(m)] = 0;
\end{aligned} \right. \tag{4.9}$$

O gradiente de $L_{pd}(m)$ com relação a $\xi_{p,i}(m)$ e $\zeta_{p,i}(m)$ igualado a zero gera as seguintes restrições $(\lambda_{p,i}(m) + \alpha_{p,i}(m)) = C$ e $(\eta_{p,i}(m) + \beta_{p,i}(m)) = C$ para amostras pertencentes a I_2 e I_4 , e $(\lambda_{p,i}(m) + \alpha_{p,i}(m)) = \frac{1}{\gamma} \xi_{p,i}(m)$ e $(\eta_{p,i}(m) + \beta_{p,i}(m)) = \frac{1}{\gamma} \zeta_{p,i}(m)$

para amostras pertencentes a I_1 e I_3 . Isto mostra uma relação analítica entre os resíduos e os multiplicadores de Lagrange, dada por:

$$\alpha_{p,i}(m) = \begin{cases} C, & \text{se } \Re(V_{p,i}(m)) > e_C; \\ \frac{1}{\gamma}(\Re(V_{p,i}(m)) - \varepsilon), & \text{se } \varepsilon < \Re(V_{p,i}(m)) \leq e_C; \\ 0, & \text{se } -\varepsilon \leq \Re(V_{p,i}(m)) \leq \varepsilon; \\ \frac{1}{\gamma}(\Re(V_{p,i}(m)) + \varepsilon), & \text{se } -e_C \leq \Re(V_{p,i}(m)) < -\varepsilon; \\ -C, & \text{se } \Re(V_{p,i}(m)) < -e_C \end{cases} \quad (4.10)$$

e

$$\beta_{p,i}(m) = \begin{cases} C, & \text{se } \Im(V_{p,i}(m)) > e_C; \\ \frac{1}{\gamma}(\Im(V_{p,i}(m)) - \varepsilon), & \text{se } \varepsilon < \Im(V_{p,i}(m)) \leq e_C; \\ 0, & \text{se } -\varepsilon \leq \Im(V_{p,i}(m)) \leq \varepsilon; \\ \frac{1}{\gamma}(\Im(V_{p,i}(m)) + \varepsilon), & \text{se } -e_C \leq \Im(V_{p,i}(m)) < -\varepsilon; \\ -C, & \text{se } \Im(V_{p,i}(m)) < -e_C, \end{cases} \quad (4.11)$$

em que $e_C = \varepsilon + \gamma C$. Se $\xi_{p,i}(m), \zeta_{p,i}(m) \neq 0$ em seus conjuntos $(I_1, I_2$ e $I_3, I_4)$, então as condições de KKT geram $\lambda_{p,i}(m), \eta_{p,i}(m) = 0$. Considerando $\alpha_{p,i}(m) = \beta_{p,i}(m) = C$ para I_2 e I_4 , $\alpha_{p,i}(m) = \frac{1}{\gamma}\xi_{p,i}(m)$ e $\beta_{p,i}(m) = \frac{1}{\gamma}\zeta_{p,i}(m)$ para I_1 e I_3 , podemos escrever:

$$\begin{cases} \xi_{p,i}(m) = \gamma\alpha_{p,i}(m); \\ \zeta_{p,i}(m) = \gamma\beta_{p,i}(m). \end{cases} \quad (4.12)$$

A partir das duas últimas equações de (4.9), temos:

$$\begin{cases} \sum_{i=1}^m \mu^{m-i} (\Re(a_{p,i}(m-1)Y_{p,i}(m)) - \varepsilon) = \\ \sum_{i=1}^m \mu^{m-i} [\Re(a_{p,i}(m-1)(H_p^*(m)X_{p,i}(m) + b_p(m))) + \xi_{p,i}(m)]; \\ \sum_{i=1}^m \mu^{m-i} (\Im(a_{p,i}(m-1)Y_{p,i}(m)) - \varepsilon) = \\ \sum_{i=1}^m \mu^{m-i} [\Im(a_{p,i}(m-1)(H_p^*(m)X_{p,i}(m) + b_p(m))) + \zeta_{p,i}(m)]. \end{cases} \quad (4.13)$$

Substituindo as duas primeiras equações de (4.9) e (4.12) em (4.13), podemos reescrevê-la como:

$$\left\{ \begin{array}{l} \sum_{i=1}^m \mu^{m-i} [\Re(a_{p,i}(m-1)Y_{p,i}(m)) - \varepsilon] = \sum_{i=1}^m \mu^{m-i} [\Re(a_{p,i}(m-1) \\ (\mu^{m-i} a_{p,i}(m-1)(\alpha_{p,i}(m) + j\beta_{p,i}(m))X_{p,i}^*(m)X_{p,i}(m) \\ + \mu^{m-i} a_{p,i}(m-1)(\alpha_{p,i}(m) + j\beta_{p,i}(m)))) + \gamma\alpha_{p,i}(m)]; \\ \sum_{i=1}^m \mu^{m-i} [\Im(a_{p,i}(m-1)Y_{p,i}(m)) - \varepsilon] = \sum_{i=1}^m \mu^{m-i} [\Im(a_{p,i}(m-1) \\ (\mu^{m-i} a_{p,i}(m-1)(\alpha_{p,i}(m) + j\beta_{p,i}(m))X_{p,i}^*(m)X_{p,i}(m) \\ + \mu^{m-i} a_{p,i}(m-1)(\alpha_{p,i}(m) + j\beta_{p,i}(m)))) + \gamma\beta_{p,i}(m)]. \end{array} \right. \quad (4.14)$$

Considerando o vetor de multiplicadores de Lagrange em sua forma complexa:

$$\begin{aligned} \boldsymbol{\psi}_p(m) &= [(\alpha_{p,1}(m) + j\beta_{p,1}(m)), \dots, (\alpha_{p,m}(m) + j\beta_{p,m}(m))]^T \text{ e} \\ \tilde{\mathbf{Y}}_p(m) &= [(a_{p,1}(m-1)Y_{p,1}(m)), \dots, (a_{p,m}(m-1)Y_{p,m}(m))]^T; \\ \tilde{\mathcal{H}}_p(m) &= [(a_{p,1}(m-1)[X_{p,1}(m); 1]), \dots, (a_{p,m}(m-1)[X_{p,m}(m); 1])]^T, \end{aligned} \quad (4.15)$$

podemos reescrevê-lo a partir de (4.14) como:

$$\boldsymbol{\psi}_{p,PSVM}(m) = [\mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p(m) \tilde{\mathcal{H}}_p^H(m) + \gamma \mathbf{I}]^{-1} [\tilde{\mathbf{Y}}_p(m) - \mathbf{1}(\varepsilon + j\varepsilon)], \quad (4.16)$$

e a expressão para $\hat{H}_p(m)$ e \hat{b}_p , a partir de (4.9), como:

$$[\hat{H}_p(m); \hat{b}_p] = \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p^H(m) \boldsymbol{\psi}_{p,PSVM}(m). \quad (4.17)$$

Fazendo,

$$\left\{ \begin{array}{l} \mathbf{A} = \gamma \mathbf{I}; \\ \mathbf{B} = \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p(m); \\ \mathbf{C}^H = \tilde{\mathcal{H}}_p^H(m). \end{array} \right. \quad (4.18)$$

Podemos aplicar o Lema de Inversão de Matrizes, reescrever a matriz inversa da Equação (4.16)

$$\begin{aligned} [\mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p(m) \tilde{\mathcal{H}}_p^H(m) + \gamma \mathbf{I}]^{-1} &= [\mathbf{BC}^H + \mathbf{A}]^{-1} = \\ \mathbf{A}^{-1} - [\mathbf{I} + \mathbf{C}^H \mathbf{A}^{-1} \mathbf{B}]^{-1} \mathbf{A}^{-1} \mathbf{BC}^H \mathbf{A}^{-1} &= [\mathbf{I} - [\mathbf{C}^H \mathbf{B}^{-1} + \mathbf{A}]^{-1} \mathbf{BC}^H] \mathbf{A}^{-1} = , \\ \left[\mathbf{I} - [\tilde{\mathcal{H}}_p^H(m) \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p(m) + \gamma \mathbf{I}]^{-1} \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p(m) \tilde{\mathcal{H}}_p^H(m) \right] &(\gamma \mathbf{I})^{-1} \end{aligned} \quad (4.19)$$

e a Equação (4.17),

$$\begin{aligned} [\hat{H}_p(m); \hat{b}_p] &= \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p^H(m) [\mathbf{I} - [\tilde{\mathcal{H}}_p^H(m) \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p(m) + \\ \gamma \mathbf{I}]^{-1} \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p(m) \tilde{\mathcal{H}}_p^H(m)] &(\gamma \mathbf{I})^{-1} [\tilde{\mathbf{Y}}_p(m) - \mathbf{1}(\varepsilon + j\varepsilon)]. \end{aligned} \quad (4.20)$$

Manipulando a Equação (4.20), podemos reescrevê-la como:

$$\begin{aligned} [\hat{H}_p(m); \hat{b}_p] &= [\tilde{\mathcal{H}}_p^H(m) \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p(m) \\ &+ \gamma \mathbf{I}]^{-1} \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p^H(m) [\tilde{\mathbf{Y}}_p(m) - \mathbf{1}(\varepsilon + j\varepsilon)]. \end{aligned} \quad (4.21)$$

Além disso, o termo $\tilde{\mathcal{H}}_p^H(m) \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p(m)$ pode ser atualizado recursivamente. Logo, fazendo $\tilde{\mathbf{R}}_{\mathcal{H}_p \mathcal{H}_p}(m) = \tilde{\mathcal{H}}_p^H(m) \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathcal{H}}_p(m)$, temos

$$\tilde{\mathbf{R}}_{\mathcal{H}_p \mathcal{H}_p}(m) = \mu \tilde{\mathbf{R}}_{\mathcal{H}_p \mathcal{H}_p}(m-1) + \tilde{\mathbf{h}}_p(m) \tilde{\mathbf{h}}_p^H(m), \quad (4.22)$$

sendo $\tilde{\mathbf{h}}_p(m) = a_{p,m}(m-1)[X_{p,m}(m); 1]$, o que leva a uma solução RLS (recursive least squares) [64]. Uma observação importante é que será necessária uma formulação regularizada para o RLS, uma vez que (4.22) não provê diretamente o termo $\gamma \mathbf{I}$ em (4.20).

O algoritmo PSVM adaptativo adequado ao sistema CP-OFDM (OFDM com prefixo cíclico) pode ser descrito como o algoritmo 1, em que $\boldsymbol{\psi}_{p,PSVM}(m)$ são os multiplicadores de Lagrange que minimizam os erros, obtidos a partir das relações (4.10) e (4.11) com $\mathbf{V}_{p,PSVM}(m)$.

Algoritmo 1: PSVM adaptativo

Inicialize:

$$\begin{aligned} \hat{\mathbf{H}}_{p,PSVM}(m-1) &= \mathbf{0}; \\ \mathbf{S}_{p,PSVM}(m-1) &= \delta \mathbf{I}; \end{aligned}$$

Calcule o vetor de pesos:

$$\begin{aligned} \mathbf{V}_{p,PSVM}(m) &= \mathbf{Y}_p(m) - \mathbf{D}_{\{\hat{\mathbf{H}}_{p,PSVM}^*(m-1)\}} \mathbf{X}_p(m); \\ \boldsymbol{\psi}_{p,PSVM}(m) &= \boldsymbol{\alpha}_{p,PSVM}(m) + j \boldsymbol{\beta}_{p,PSVM}(m); \\ \mathbf{a}_{p,PSVM}(m) &= \mathbf{D}_{\{\mathbf{V}_{p,PSVM}(m)\}}^{-1} \boldsymbol{\psi}_{p,PSVM}(m) / \gamma; \end{aligned}$$

Atualize $\hat{\mathbf{H}}_{p,PSVM}(m)$ e $\mathbf{S}_{p,PSVM}(m)$:

$$\begin{aligned} \tilde{\mathbf{Y}}_{p,PSVM}(m) &= \mathbf{D}_{\{\mathbf{a}_{p,PSVM}(m)\}} \mathbf{Y}_p(m) - \mathbf{1}(\varepsilon + j\varepsilon); \\ \tilde{\mathcal{H}}_{p,PSVM}(m) &= \mathbf{D}_{\{\mathbf{a}_{p,PSVM}(m)\}} [\mathbf{D}_{\{\mathbf{x}_p(m)\}} \mathbf{1}]; \\ \boldsymbol{\Pi}_{p,PSVM}(m) &= \tilde{\mathcal{H}}_{p,PSVM}^H(m) \mathbf{S}_{p,PSVM}(m-1); \\ \mathbf{K}_{p,PSVM}(m) &= \boldsymbol{\Pi}_{p,PSVM}^H(m) / (\mu \mathbf{I} + \boldsymbol{\Pi}_{p,PSVM}(m) \tilde{\mathcal{H}}_{p,PSVM}(m)); \\ \mathbf{e}_{p,PSVM}(m) &= \tilde{\mathbf{Y}}_{p,PSVM}(m) - \tilde{\mathcal{H}}_{p,PSVM}^H(m) [\hat{\mathbf{H}}_{p,PSVM}(m-1); b(m-1)]; \\ [\hat{\mathbf{H}}_{p,PSVM}(m); b(m)] &= \\ &[\hat{\mathbf{H}}_{p,PSVM}(m-1); b(m-1)] + \mathbf{K}_{p,PSVM}(m) \mathbf{e}_{p,PSVM}(m); \\ \mathbf{S}_{p,PSVM}(m) &= (\mathbf{S}_{p,PSVM}(m-1) - \mathbf{K}_{p,PSVM}(m) \boldsymbol{\Pi}_{p,PSVM}(m)) / \mu; \end{aligned}$$

Pôde-se observar que este desenvolvimento *fast-SVM* para números complexos não apresentou nenhuma característica particular em relação aos desenvolvimentos para valores reais, e pode ser aplicado em ambos os casos.

Neste capítulo foi descrito o desenvolvimento do algoritmo PSVM adaptativo

para aplicações como regressor de valor complexo.

Capítulo 5

Regressor UPSVM de valores complexos

Neste capítulo descrevemos a segunda proposta de técnica para estimação de canal adequada aos sistemas PLC.

Nesta técnica, o desenvolvimento proposto em [58] para SVM com valores complexos é aplicado ao algoritmo *Umbiased Proximal Support Vector Machine* (UPSVM) proposto em [49] para ser usado como regressor. O objetivo desta análise é obter um estimador baseado em algoritmo *fast-SVM* que possa ser aplicado diretamente em números complexos e seja capaz de minimizar simultaneamente as partes real e imaginária dos erros de estimação.

Em [49] os autores afirmaram que a minimização do termo b^2 na função objetivo, como realizado no algoritmo PSVM, não leva a melhora na generalização ou a redução dos erros de treinamento. Ao invés disso, diminui a probabilidade de sucesso com os dados de teste e aumenta o número de erros de treinamento quando o parâmetro γ excede um determinado valor, fato este devido à polarização do hiperplano separador [49]. Em virtude disso, eles propuseram um novo algoritmo *fast-SVM* que não leva em conta o parâmetro b na função objetivo a ser minimizada.

5.1 Algoritmo UPSVM adaptativo

Em nosso desenvolvimento com valores complexos, a funcional primal-dual para o UPSVM a cada instante de tempo (m) e em cada símbolo piloto (p), pode ser formulada como:

$$\begin{aligned}
L_{pd}(m) &= \frac{1}{2} \|\tilde{\mathbf{H}}_p(m)\|^2 + \frac{1}{2\gamma} \sum_{i \in I_1} \mu^{m-i} (\xi_{p,i}^2(m)) + C \sum_{i \in I_2} \mu^{m-i} (\xi_{p,i}(m)) \\
&\quad + \frac{1}{2\gamma} \sum_{i \in I_3} \mu^{m-i} (\zeta_{p,i}^2(m)) + C \sum_{i \in I_4} \mu^{m-i} (\zeta_{p,i}(m)) - \sum_{i \in I_2, I_4} (\gamma C^2) \\
&\quad - \sum_{i=1}^m \mu^{m-i} (\lambda_{p,i}(m) \xi_{p,i}(m)) - \sum_{i=1}^m \mu^{m-i} (\eta_{p,i}(m) \zeta_{p,i}(m)) \tag{5.1} \\
&+ \sum_{i=1}^m \mu^{m-i} \alpha_{p,i}(m) [\Re(a_{p,i}(m-1)(Y_{p,i}(m) - (H_p^*(m)X_{p,i}(m)))) - \varepsilon - \xi_{p,i}(m)] \\
&+ \sum_{i=1}^m \mu^{m-i} \beta_{p,i}(m) [\Im(a_{p,i}(m-1)(Y_{p,i}(m) - (H_p^*(m)X_{p,i}(m)))) - \varepsilon - \zeta_{p,i}(m)].
\end{aligned}$$

Similar ao PSVM, a condição de KKT necessária e suficiente para otimalidade no problema com restrição de igualdade é obtida igualando a zero as derivadas parciais de $L_{pd}(m)$, com relação a $(H_p(m), \xi_{p,i}(m), \zeta_{p,i}(m), \alpha_{p,i}(m), \beta_{p,i}(m))$

$$\left\{ \begin{aligned}
\frac{\partial L_{pd}(m)}{\partial H_p(m)} &= 0 \rightarrow H_p(m) = \sum_{i=1}^m \mu^{m-i} a_{p,i}(m-1) [\alpha_{p,i}(m) - j\beta_{p,i}(m)] X_{p,i}(m); \\
\frac{\partial L_{pd}(m)}{\partial \xi_{p,i}(m)} &= 0 \rightarrow \frac{1}{\gamma} \sum_{i \in I_1} \mu^{m-i} \xi_{p,i}(m) + C \sum_{i \in I_2} \mu^{m-i} - \sum_{i=1}^m \mu^{m-i} \lambda_{p,i}(m) \\
&\quad - \sum_{i=1}^m \mu^{m-i} \alpha_{p,i}(m) = 0; \\
\frac{\partial L_{pd}(m)}{\partial \zeta_{p,i}(m)} &= 0 \rightarrow \frac{1}{\gamma} \sum_{i \in I_3} \mu^{m-i} \zeta_{p,i}(m) + C \sum_{i \in I_4} \mu^{m-i} - \sum_{i=1}^m \mu^{m-i} \eta_{p,i}(m) \\
&\quad - \sum_{i=1}^m \mu^{m-i} \beta_{p,i}(m) = 0; \\
\frac{\partial L_{pd}(m)}{\partial \alpha_{p,i}(m)} &= 0 \rightarrow \sum_{i=1}^m \mu^{m-i} [\Re(a_{p,i}(m-1)(Y_{p,i}(m) - (H_p^*(m)X_{p,i}(m)))) \\
&\quad - \varepsilon - \xi_{p,i}(m)] = 0; \\
\frac{\partial L_{pd}(m)}{\partial \beta_{p,i}(m)} &= 0 \rightarrow \sum_{i=1}^m \mu^{m-i} [\Im(a_{p,i}(m-1)(Y_{p,i}(m) - (H_p^*(m)X_{p,i}(m)))) \\
&\quad - \varepsilon - \zeta_{p,i}(m)] = 0.
\end{aligned} \right. \tag{5.2}$$

Inserindo a primeira equação de (5.2) e repetindo as considerações feitas em (4.12) nas duas últimas equações de (5.2), podemos reescrevê-las como:

$$\left\{ \begin{array}{l} \sum_{i=1}^m \mu^{m-i} [\Re(a_{p,i}(m-1)Y_{p,i}(m)) - \varepsilon] = \\ \sum_{i=1}^m \mu^{m-i} [\Re(a_{p,i}(m-1)\mu^{m-i}a_{p,i}(m-1)(\alpha_{p,i}(m) \\ + j\beta_{p,i}(m))X_{p,i}^*(m)X_{p,i}(m)) + \gamma\alpha_{p,i}(m)]; \\ \sum_{i=1}^m \mu^{m-i} [\Im(a_{p,i}(m-1)Y_{p,i}(m)) - \varepsilon] = \\ \sum_{i=1}^m \mu^{m-i} [\Im(a_{p,i}(m-1)\mu^{m-i}a_{p,i}(m-1)(\alpha_{p,i}(m) \\ + j\beta_{p,i}(m))X_{p,i}^*(m)X_{p,i}(m)) + \gamma\beta_{p,i}(m)]. \end{array} \right. \quad (5.3)$$

Usando as considerações (4.15) e sabendo que $0 < \mu < 1$, pode-se escrever os multiplicadores de Lagrange na forma vetorial complexa a partir de (5.3) como:

$$\boldsymbol{\psi}_{p,UPSV M}(m) = [\mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p(m) \tilde{\mathbf{X}}_p^H(m) + \gamma \mathbf{I}]^{-1} [\tilde{\mathbf{Y}}_p(m) - \mathbf{1}(\varepsilon + j\varepsilon)], \quad (5.4)$$

em que $\tilde{\mathbf{X}}_p(m) = [(a_{p,1}(m-1)X_{p,1}(m)), \dots, (a_{p,m}(m-1)X_{p,m}(m))]^T$.

A expressão para o canal estimado é:

$$\hat{H}_p(m) = \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p^H(m) \boldsymbol{\psi}_{p,UPSV M}(m). \quad (5.5)$$

Fazendo,

$$\left\{ \begin{array}{l} \mathbf{A} = \gamma \mathbf{I}; \\ \mathbf{B} = \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p(m); \\ \mathbf{C}^H = \tilde{\mathbf{X}}_p^H(m). \end{array} \right. \quad (5.6)$$

Podemos aplicar o Lema de Inversão de Matrizes, reescrever a matriz inversa da Equação (5.4)

$$\begin{aligned} [\mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p(m) \tilde{\mathbf{X}}_p^H(m) + \gamma \mathbf{I}]^{-1} &= [\mathbf{B} \mathbf{C}^H + \mathbf{A}]^{-1} = \\ \mathbf{A}^{-1} - [\mathbf{I} + \mathbf{C}^H \mathbf{A}^{-1} \mathbf{B}]^{-1} \mathbf{A}^{-1} \mathbf{B} \mathbf{C}^H \mathbf{A}^{-1} &= [\mathbf{I} - [\mathbf{C}^H \mathbf{B}^{-1} + \mathbf{A}]^{-1} \mathbf{B} \mathbf{C}^H] \mathbf{A}^{-1} = , \quad (5.7) \\ [\mathbf{I} - [\tilde{\mathbf{X}}_p^H(m) \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p(m) + \gamma \mathbf{I}]^{-1} \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p(m) \tilde{\mathbf{X}}_p^H(m)] &(\gamma \mathbf{I})^{-1} \end{aligned}$$

e a Equação (5.5),

$$\begin{aligned} \hat{H}_p(m) &= \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p^H(m) [\mathbf{I} - [\tilde{\mathbf{X}}_p^H(m) \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p(m) + \\ \gamma \mathbf{I}]^{-1} \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p(m) \tilde{\mathbf{X}}_p^H(m)] (\gamma \mathbf{I})^{-1} [\tilde{\mathbf{Y}}_p(m) - \mathbf{1}(\varepsilon + j\varepsilon)]. \end{aligned} \quad (5.8)$$

Manipulando a Equação (5.8), podemos reescrevê-la como:

$$\begin{aligned} \hat{H}_p(m) &= [\tilde{\mathbf{X}}_p^H(m) \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p(m) \\ &+ \gamma]^{-1} \tilde{\mathbf{X}}_p^H(m) \mathbf{D}_{\{\mu_p(m)\}} [\tilde{\mathbf{Y}}_p(m) - \mathbf{1}(\varepsilon + j\varepsilon)]. \end{aligned} \quad (5.9)$$

Como o termo $\tilde{\mathbf{X}}_p^H(m) \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p(m)$ pode ser atualizado recursivamente, fazendo $\tilde{R}_{\mathbf{X}_p \mathbf{X}_p}(m) = \tilde{\mathbf{X}}_p^H(m) \mathbf{D}_{\{\mu_p(m)\}} \tilde{\mathbf{X}}_p(m)$, temos

$$\tilde{R}_{\mathbf{X}_p \mathbf{X}_p}(m) = \mu \tilde{R}_{\mathbf{X}_p \mathbf{X}_p}(m-1) + \tilde{X}_p(m) \tilde{X}_p^H(m), \quad (5.10)$$

em que, $\tilde{X}_p(m) = a_{p,m}(m-1)X_{p,m}(m)$. Este procedimento também leva a uma solução RLS. Novamente uma solução RLS regularizada torna-se necessária devido ao fato de que (5.10) não provê diretamente o termo γ em (5.4).

O algoritmo UPSVM adaptativo adequado ao sistema CP-OFDM pode ser escrito como o algoritmo 2, sendo $\boldsymbol{\psi}_{p,UPSVM}(m)$ obtidos a partir das relações (4.10) e (4.11) com $\mathbf{V}_{p,UPSVM}(m)$.

Algoritmo 2: UPSVM adaptativo

Inicialize:

$$\hat{\mathbf{H}}_{p,UPSVM}(m-1) = \mathbf{0};$$

$$\mathbf{S}_{p,UPSVM}(m-1) = \delta \mathbf{I};$$

Calcule o vetor de pesos:

$$\mathbf{V}_{p,UPSVM}(m) = \mathbf{Y}_p(m) - \mathbf{D}_{\{\hat{\mathbf{H}}_{p,UPSVM}^*(m-1)\}} \mathbf{X}_p(m);$$

$$\boldsymbol{\psi}_{p,UPSVM}(m) = \boldsymbol{\alpha}_{p,UPSVM}(m) + j\boldsymbol{\beta}_{p,UPSVM}(m);$$

$$\mathbf{a}_{p,UPSVM}(m) = (\mathbf{D}_{\{\mathbf{V}_{p,UPSVM}(m)\}}^{-1} \boldsymbol{\psi}_{p,UPSVM}(m)) / \gamma;$$

Atualize $\hat{\mathbf{H}}_{p,UPSVM}(m)$ e $\mathbf{S}_{p,UPSVM}(m)$:

$$\tilde{\mathbf{Y}}_{p,UPSVM}(m) = \mathbf{D}_{\{\mathbf{a}_{p,UPSVM}(m)\}} \mathbf{Y}_p - \mathbf{1}(\varepsilon + j\varepsilon);$$

$$\tilde{\mathbf{X}}_{p,UPSVM}(m) = \mathbf{D}_{\{\mathbf{a}_{p,UPSVM}(m)\}} \mathbf{X}_p;$$

$$\boldsymbol{\Pi}_{p,UPSVM}(m) = \mathbf{D}_{\{\tilde{\mathbf{X}}_{p,UPSVM}^*(m)\}} \mathbf{S}_{p,UPSVM}(m-1);$$

$$\mathbf{K}_{p,UPSVM}(m) = \boldsymbol{\Pi}_{p,UPSVM}^*(m) / (\mu \mathbf{I} + \boldsymbol{\Pi}_{p,UPSVM}(m) \mathbf{D}_{\{\tilde{\mathbf{X}}_p(m)\}});$$

$$\mathbf{e}_{p,UPSVM}(m) = \tilde{\mathbf{Y}}_{p,UPSVM}(m) - \mathbf{D}_{\{\tilde{\mathbf{X}}_{p,UPSVM}^*(m)\}} \hat{\mathbf{H}}_{p,UPSVM}(m-1);$$

$$\hat{\mathbf{H}}_{p,UPSVM}(m) = \hat{\mathbf{H}}_{p,UPSVM}(m-1) + \mathbf{K}_{p,UPSVM}(m) \mathbf{e}_{p,UPSVM}(m);$$

$$\mathbf{S}_{p,UPSVM}(m) = (\mathbf{S}_{p,UPSVM}(m-1) - \mathbf{K}_{p,UPSVM}(m) \boldsymbol{\Pi}_{p,UPSVM}(m)) / \mu;$$

Pôde-se observar que este desenvolvimento *fast-SVM* para números complexos, também não apresentou nenhuma característica particular em relação aos desenvolvimentos para valores reais, e pode ser aplicado em ambos os casos.

Neste capítulo foi descrito o desenvolvimento do algoritmo UPSVM adaptativo para aplicações como regressor de valor complexo.

Capítulo 6

Estimadores utilizados na comparação de desempenho

Neste capítulo serão descritos alguns algoritmos que já foram aplicados para a estimação de canal em sistemas PLC. Estas técnicas serão utilizadas na comparação de desempenho com os algoritmos *fast-SVM* propostos.

6.1 LS

O LS (*Least Square*), que é tradicionalmente aplicado para a estimação de canal em sistemas PLC, provê a primeira estimação da função de transferência do canal dada a partir do algoritmo 3 [65].

Algoritmo 3: LS

A cada iteração calcule:

$$\hat{\mathbf{H}}_{p,LS}(m) = \mathbf{D}_{\{\mathbf{x}_p(m)\}}^{-1} \mathbf{Y}_p(m);$$

Similar a [22], usamos esta formulação para estimar a resposta em frequência do canal na posição dos tons pilotos p a cada instante de tempo (m) sem considerar as estimações nos instantes anteriores. Este procedimento permite monitorar à rápida variação do canal, mas as estimações têm alta variância devido a presença de ruído.

6.2 RLS

O RLS (*Recursive Least Square*), em sistemas OFDM com prefixo cíclico (CP-OFDM), pode ser escrito como no algoritmo 4, em que μ é um fator de esquecimento e $\mathbf{S}_p(m)$ é a matriz inversa da correlação dos dados observados que, neste caso, é uma matriz diagonal. Este algoritmo é considerado de bom desempenho quando aplicado a ambientes variantes no tempo [64].

Algoritmo 4: RLS

Inicialize:

$$\hat{\mathbf{H}}_{p,RLS}(m-1) = \mathbf{0};$$

$$\mathbf{S}_{p,RLS}(m-1) = \delta \mathbf{I};$$

Atualize $\hat{\mathbf{H}}_{p,RLS}(m)$ e $\mathbf{S}_{p,RLS}(m)$:

$$\mathbf{\Pi}_{p,RLS}(m) = \mathbf{D}_{\{\mathbf{x}_p(m)\}} \mathbf{S}_{p,RLS}(m-1);$$

$$\mathbf{K}_{p,RLS}(m) = \frac{\mathbf{\Pi}_{p,RLS}(m)}{(\mu \mathbf{I} + \mathbf{\Pi}_{p,RLS}(m) \mathbf{D}_{\{\mathbf{x}_p(m)\}})}$$

$$\mathbf{e}_{p,RLS}(m) = \mathbf{Y}_p(m) - \mathbf{D}_{\{\hat{\mathbf{H}}_{p,RLS}(m-1)\}} \mathbf{X}_p(m);$$

$$\hat{\mathbf{H}}_{p,RLS}(m) = \hat{\mathbf{H}}_{p,RLS}(m-1) + \mathbf{K}_{p,RLS}(m) \mathbf{e}_{p,RLS}(m);$$

$$\mathbf{S}_{p,RLS}(m) = (\mathbf{S}_{p,RLS}(m-1) - \mathbf{K}_{p,RLS}(m) \mathbf{\Pi}_{p,RLS}(m)) / \mu;$$

6.3 LMMSE

O algoritmo LMMSE (*Linear Minimum Mean Square Error*) é descrito em [66] como no algoritmo 5, sendo $\mathbf{R}_{pp} = E[\mathbf{H}_p \mathbf{H}_p^H]$ o valor esperado da matriz de correlação do canal nas posições dos tons piloto e $\sigma_{v,p}^2$ a variância do ruído na portadora p . Neste trabalho, $\mathbf{H}_p(m)$, para a obtenção de \mathbf{R}_{pp} , será estimado como

$$\mathbf{H}_p(m) = \frac{1}{m-1} \sum_{i=1}^{m-1} \hat{\mathbf{H}}_p(i).$$

Algoritmo 5: LMMSE

$$\hat{\mathbf{H}}_{p,LMMSE}(m) = \mathbf{R}_{pp} (\mathbf{R}_{pp} + \sigma_{v,p}^2 (\mathbf{X}_p \mathbf{X}_p^H)^{-1})^{-1} \hat{\mathbf{H}}_{p,LS}(m);$$

Para reduzir a complexidade do estimador foi proposta em [67] uma versão aproximada do algoritmo LMMSE (*approximate Linear Minimum Mean Square Error* - aLMMSE), onde o termo $(\mathbf{X}_p \mathbf{X}_p^H)^{-1}$ é substituído por seu valor esperado $E[(\mathbf{X}_p \mathbf{X}_p^H)^{-1}]$.

Usando a mesma modulação em todos os pilotos, definindo $\kappa = E[|X_p|^2]E[|1/X_p|^2]$ e $\mathbf{\Gamma} = \mathbf{D}_{\{E[|X_1|^2]/\sigma_{v,1}^2, \dots, E[|X_P|^2]/\sigma_{v,P}^2\}}$, o estimador aLMMSE pode ser descrito a partir do algoritmo 6.

Algoritmo 6: aLMMSE

$$\hat{\mathbf{H}}_{p,aLMMSE}(m) = \mathbf{R}_{pp} (\mathbf{R}_{pp} + \kappa \mathbf{\Gamma}^{-1})^{-1} \hat{\mathbf{H}}_{p,LS}(m);$$

Em [68] foi considerado $\kappa = 1$ para a modulação BPSK e $\mathbf{\Gamma} = \mathbf{I}\varpi$, sendo que \mathbf{I} é uma matriz identidade e ϖ fixado em 10000, independente da variância do ruído.

6.4 LMS

O algoritmo LMS (*Least Mean Square*) pode ser escrito como no algoritmo 7 [45], em que ϑ é o passo de aprendizado que determina a habilidade de rastreamento

e o mínimo de erro do algoritmo.

Algoritmo 7: LMS

$$\begin{aligned} \mathbf{e}_{p,LMS}(m) &= \mathbf{Y}_p(m) - \mathbf{D}_{\{\hat{\mathbf{H}}_{p,LMS}(m-1)\}} \mathbf{X}_p(m); \\ \hat{\mathbf{H}}_{p,LMS}(m) &= \hat{\mathbf{H}}_{p,LMS}(m-1) + 2\vartheta \mathbf{D}_{\{\mathbf{e}_{p,LMS}(m)\}} \mathbf{X}_p(m); \end{aligned}$$

6.5 SOM

Completando o conjunto de algoritmos que serão utilizados na comparação de desempenho, consideramos a técnica proposta em [17] e [18], onde os autores usaram uma rede neural competitiva para estimar e rastrear o canal PLC.

O algoritmo proposto é conhecido por mapeamento com auto-organização (SOM) e pode ser descrito como no algoritmo 8, em que $\mathbf{C}_{p,\mathcal{Y}}(m)$ é um vetor contendo os centróides dos *clusters* com índice temporal m e cada *cluster* está relacionado a uma subportadora p ; \mathcal{Y} é o índice do neurônio vencedor e $\varphi \in (0, 1)$ é a taxa de aprendizado dos neurônio vencedor.

Os neurônios perdedores são atualizados a partir de $\mathbf{C}_{p,\nu}(m)$, sendo $\bar{\mathbf{Y}}_p(m)$ o ponto simétrico de $\mathbf{Y}_p(m)$, localizado próximo à posição do neurônio perdedor ν ; ς é a taxa de aprendizado do neurônio perdedor e $\varsigma \in [0, \varphi]$.

Os centróides são inicializados com os tons pilotos nas posições p .

A estimação da resposta em frequência do canal é $\hat{\mathbf{H}}_{p,SOM}(m)$, no qual M_Q é o número de centróides na constelação da modulação.

Para rastrear a variação do canal, é calculado o termo ϕ_m a cada símbolo sucessivo m , e de forma adaptativa se escolhe a taxa de aprendizado φ_m e ς_m , sendo ϕ_{max} um limite máximo para o erro de fase e φ_{max} um fator de máximo aprendizado.

Os centróides devem ser reinicializados sempre que $\phi_m \geq \phi_{max}$, caso contrário o algoritmo pode divergir [18].

Neste capítulo foi apresentado os algoritmos que serão utilizados na comparação de desempenho com os algoritmos *fast-SVM* propostos.

Algoritmo 8: SOM

Inicialize:

$$\mathbf{C}_{p,r}(1) = \hat{\mathbf{H}}_{p,LS}(1);$$

$$\mathbf{C}_{p,t}(1) = -\hat{\mathbf{H}}_{p,LS}(1);$$

Para $m \geq 2$ **atualize** $\mathbf{C}_{p,r}(m)$ e $\mathbf{C}_{p,t}(m)$:

$$\mathbf{C}_{p,r}(m) = \mathbf{C}_{p,r}(m-1) + \varphi(\mathbf{Y}_p(m) - \mathbf{C}_{p,r}(m-1));$$

$$\mathbf{C}_{p,t}(m) = \mathbf{C}_{p,t}(m-1) + \varsigma(\bar{\mathbf{Y}}_p(m) - \mathbf{C}_{p,t}(m-1));$$

Calcule $\hat{\mathbf{H}}_{p,SOM}(m)$ e ϕ_m :

$$\hat{\mathbf{H}}_{p,SOM}(m) = (1/M_Q) \sum_{i=1}^{M_Q} \mathbf{C}_{p,i}(m);$$

$$\phi_m = \frac{1}{P} \left| \sum_{p=1}^P \angle \left(\frac{Y_p(m)}{C_{p,r}(m-1)} \right) \right|;$$

Faça:

$$\varphi_m = \frac{\varphi_{max}}{\phi_{max}} \phi_m, \quad \text{se } \phi_m < \phi_{max};$$

$$\varsigma_m = \frac{\varphi_m}{\phi_{max}} \phi_m = \frac{\varphi_{max}}{\phi_{max}^2} \phi_m^2, \quad \text{se } \phi_m < \phi_{max};$$

$$\varphi_m = \varphi_{max}, \quad \text{se } \phi_m \geq \phi_{max};$$

$$\varsigma_m = \varphi_m, \quad \text{se } \phi_m \geq \phi_{max};$$

Capítulo 7

Complexidade Computacional

O número de operações (multiplicações e adições) com ponto-flutuante (*flops*) calculado para os algoritmos PSVM, UPSVM, LS, RLS, aLMMSE, LMS e SOM (algoritmos (1), (2), (3), (4), (6), (7) e (8), respectivamente) operando com valores complexos, é mostrado nas Tabelas 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7 onde foi considerado que as multiplicações que envolvem o vetor de 1's podem ser expressas por somas.

Tabela 7.1: Número de *flops* por iteração no algoritmo PSVM linear com valores complexos.

Operações por iteração	Mult.	Ad.
$\mathbf{V}_{p,PSVM}(m) = \mathbf{Y}_p(m) - \mathbf{D}_{\{\mathbf{H}_{p,PSVM}^H(m-1)\}} \mathbf{X}_p(m)$	$4P$	$4P$
$\boldsymbol{\psi}_{p,PSVM}(m) = \boldsymbol{\alpha}_{p,PSVM}(m) + j\boldsymbol{\beta}_{p,PSVM}(m),$	–	–
$\mathbf{a}_{p,PSVM}(m) = (\mathbf{D}_{\{\mathbf{V}_{p,PSVM}(m)\}}^{-1} \boldsymbol{\psi}_{p,PSVM}(m)) / \gamma$	$6P$	$2P$
$\mathbf{Y}_{p,PSVM}(m) = \mathbf{D}_{\{\mathbf{a}_{p,PSVM}(m)\}} \mathbf{Y}_p(m) - \mathbf{1}(\varepsilon + j\varepsilon)$	$4P$	$4P$
$\mathcal{H}_{p,PSVM}(m) = \mathbf{D}_{\{\mathbf{a}_{p,PSVM}(m)\}} [\mathbf{D}_{\{\mathbf{X}_p(m)\}} - \mathbf{1}]$	$4P$	$3P$
$\mathbf{\Pi}_{p,PSVM}(m) = \tilde{\mathcal{H}}_{p,PSVM}^H(m) \mathbf{S}_{p,PSVM}(m-1)$	$8P$	$4P$
$\mathbf{K}_{p,PSVM}(m) = \frac{\mathbf{\Pi}_{p,PSVM}^H(m)}{(\mu \mathbf{I} + \mathbf{\Pi}_{p,PSVM}(m) \tilde{\mathcal{H}}_{p,PSVM}(m))}$	$16P^2 + 24P$	$4P^2 + 13P + 1$
$\mathbf{e}_{p,PSVM}(m) = \mathbf{Y}_{p,PSVM}(m) - \tilde{\mathcal{H}}_{p,PSVM}(m) [\hat{\mathbf{H}}_{p,PSVM}(m-1); b(m-1)]$	$8P$	$6P$
$[\hat{\mathbf{H}}_{p,PSVM}(m); b(m)] = [\hat{\mathbf{H}}_{p,PSVM}(m-1); b(m-1)] + \mathbf{K}_{p,PSVM}^H(m) \mathbf{e}_{p,PSVM}(m)$	$8P$	$5P + 1$
$\mathbf{S}_{p,PSVM}(m) = (\mathbf{S}_{p,PSVM}(m-1) - \mathbf{K}_{p,PSVM}(m) \mathbf{\Pi}_p(m)) / \mu$	$10P$	$6P$
total	$16P^2 + 76P$	$4P^2 + 47P + 2$

Tabela 7.2: Número de *flops* por iteração no algoritmo UPSVM linear com valores complexos.

Operações por iteração	Mult.	Ad.
$\mathbf{V}_{p,UPSVM}(m) = \mathbf{Y}_p(m) - \mathbf{D}_{\{\mathbf{H}_{p,UPSVM}^H(m-1)\}} \mathbf{X}_p(m)$	$4P$	$4P$
$\boldsymbol{\psi}_{p,UPSVM}(m) = \boldsymbol{\alpha}_{p,UPSVM}(m) + j\boldsymbol{\beta}_{p,UPSVM}(m),$	–	–
$\mathbf{a}_{p,UPSVM}(m) = (\mathbf{D}_{\{\mathbf{v}_{p,UPSVM}(m)\}}^{-1} \boldsymbol{\psi}_{p,UPSVM}(m)) / \gamma$	$6P$	$2P$
$\tilde{\mathbf{Y}}_{p,UPSVM}(m) = \mathbf{D}_{\{\mathbf{a}_{p,UPSVM}(m)\}} \mathbf{Y}_p - \mathbf{1}(\varepsilon + j\varepsilon)$	$4P$	$4P$
$\tilde{\mathbf{X}}_{p,UPSVM}(m) = \mathbf{D}_{\{\mathbf{a}_{p,UPSVM}(m)\}} \mathbf{X}_p$	$4P$	$2P$
$\mathbf{\Pi}_{p,UPSVM}(m) = \mathbf{D}_{\{\tilde{\mathbf{x}}_{p,UPSVM}^H(m)\}} \mathbf{S}_p(m-1)$	$4P$	$2P$
$\mathbf{K}_{p,UPSVM}(m) =$ $\mathbf{\Pi}_{p,UPSVM}^H(m) / (\mu \mathbf{I} + \mathbf{\Pi}_{p,UPSVM}(m) \mathbf{D}_{\{\tilde{\mathbf{x}}_{p,UPSVM}(m)\}})$	$8P$	$5P$
$\tilde{\mathbf{Y}}_{p,UPSVM}(m) - \mathbf{D}_{\{\tilde{\mathbf{x}}_{p,UPSVM}^H(m)\}} \mathbf{H}_{p,UPSVM}(m-1)$	$4P$	$4P$
$\mathbf{H}_{p,UPSVM}(m) =$ $\mathbf{H}_{p,UPSVM}(m-1) + \mathbf{K}_{p,UPSVM}(m) \mathbf{e}_{p,UPSVM}(m)$	$4P$	$4P$
$\mathbf{S}_{p,UPSVM}(m) =$ $(\mathbf{S}_{p,UPSVM}(m-1) - \mathbf{K}_{p,UPSVM}(m) \mathbf{\Pi}_{p,UPSVM}(m)) / \mu$	$6P$	$4P$
total	$44P$	$6P$

Tabela 7.3: Número de *flops* por iteração no algoritmo LS com valores complexos.

Operações por iteração	Mult.	Ad.
$\hat{\mathbf{H}}_{p,LS}(m) = \mathbf{D}_{\{\mathbf{x}_p(m)\}}^{-1} \mathbf{Y}_p(m)$	$4P$	$2P$
total	$4P$	$2P$

Tabela 7.4: Número de *flops* por iteração no algoritmo RLS com valores complexos.

Operações por iteração	Mult.	Ad.
$\mathbf{\Pi}_{p,RLS}(m) = \mathbf{D}_{\{\mathbf{x}_p(m)\}} \mathbf{S}_{p,RLS}(m-1)$	$4P$	$2P$
$\mathbf{K}_{p,RLS}(m) = \frac{\mathbf{\Pi}_{p,RLS}(m)}{(\mu \mathbf{I} + \mathbf{\Pi}_{p,RLS}(m) \mathbf{D}_{\{\mathbf{x}_p(m)\}})}$	$8P$	$5P$
$\mathbf{e}_{p,RLS}(m) = \mathbf{Y}_p(m) - \mathbf{D}_{\{\hat{\mathbf{H}}_{p,RLS}(m-1)\}} \mathbf{X}_p(m)$	$4P$	$4P$
$\hat{\mathbf{H}}_{p,RLS}(m) = \hat{\mathbf{H}}_{p,RLS}(m-1) + \mathbf{K}_{p,RLS}(m) \mathbf{e}_{p,RLS}(m)$	$4P$	$4P$
$\mathbf{S}_{p,RLS}(m) = (\mathbf{S}_{p,RLS}(m-1) - \mathbf{K}_{p,RLS}(m) \mathbf{\Pi}_{p,RLS}(m)) / \mu$	$6P$	$4P$
total	$26P$	$19P$

No algoritmo aLMMSE, como a matriz a ser invertida é positiva definida, pode-se aplicar o algoritmo de Choleski [69] para a obtenção de um sistema triangular equivalente ao original.

Tabela 7.5: Número de *flops* por iteração no algoritmo aLMMSE com valores complexos.

Operações por iteração	Mult.	Ad.
$\mathbf{R}_{pp} + \kappa\mathbf{\Gamma}^{-1}$	$(4P)^2 + 4P$	$(2P)^2 + 4P$
Choleski [69]	$(4P)^3/6$	$(2P)^3/6$
Subst. “Backward/foward” [69]	$(4P)^2 + 4P + 2$	$(2P)^2 + 2P$
$\hat{\mathbf{H}}_{p,aLMMSE}(m) = \mathbf{R}_{pp}(\mathbf{R}_{pp} + \kappa\mathbf{\Gamma}^{-1})^{-1}\hat{\mathbf{H}}_{p,LS}(m)$	$(4P)^3 + (4P)^2$	$(2P)^3 + (2P)^2$
total	$\frac{7(4P)^3}{6} + 3(4P)^2 + 8P + 2$	$\frac{7(2P)^3}{6} + 3(2P)^2 + 6P$

Tabela 7.6: Número de *flops* por iteração no algoritmo LMS com valores complexos.

Operações por iteração	Mult.	Ad.
$\mathbf{e}_{p,LMS}(m) = \mathbf{Y}_p(m) - \mathbf{D}_{\{\hat{\mathbf{H}}_{p,LMS}(m-1)\}}\mathbf{X}_p(m)$	$4P$	$4P$
$\hat{\mathbf{H}}_{p,LMS}(m) = \hat{\mathbf{H}}_{p,LMS}(m-1) + 2\vartheta \mathbf{D}_{\{\mathbf{e}_{p,LMS}(m)\}}\mathbf{X}_p(m)$	$7P$	$4P$
total	$11P$	$8P$

Tabela 7.7: Número de *flops* por iteração no algoritmo SOM com valores complexos.

Operações por iteração	Mult.	Ad.
$\mathbf{C}_{p,r}(m) = \mathbf{C}_{p,r}(m-1) + \varphi_m(\mathbf{Y}_p(m) - \mathbf{C}_{p,r}(m-1))$	$2P$	$4P$
$\mathbf{C}_{p,i}(m) = \mathbf{C}_{p,i}(m-1) + \varsigma_m(\mathbf{Y}_p(m) - \mathbf{C}_{p,i}(m-1))$	$2P$	$4P$
$\hat{\mathbf{H}}_{p,SOM}(m) = (1/M_Q) \sum_{i=1}^{M_Q} \mathbf{C}_{p,i}(m)$	$4P$	$4P$
$\phi_m = \frac{1}{P} \left \sum_{p=1}^P \angle \left(\frac{Y_p(m)}{C_{p,r}(m-1)} \right) \right $	$6P$	$4P$
$\varphi_m = \frac{\varphi_{max}}{\phi_{max}} \phi_m$	2	$-$
$\varsigma_m = \frac{\varphi_m}{\phi_{max}} \phi_m = \frac{\varphi_{max}}{\phi_{max}^2} \phi_m^2$	2	$-$
total	$18P$	$16P$

Pode-se observar, nas Tabelas apresentadas, que o número de operações com ponto-flutuante por iteração temporal, com exceção dos algoritmos aLMMSE e PSVM que possuem complexidade computacional $O(P^3)$ e $O(P^2)$, respectivamente, todos os outros algoritmos apresentaram complexidade assintótica $O(P)$. Desta forma podemos concluir que a complexidade computacional do algoritmo UPSVM, que também é $O(P)$, não inviabiliza sua aplicação *online*.

Neste capítulo foi mostrada a complexidade computacional de cada algoritmo, permitindo a comparação entre eles.

Capítulo 8

Simulações e análise dos resultados

8.1 Considerações e parâmetros das simulações

Com o objetivo de testar o desempenho dos algoritmos propostos e compará-los com os demais algoritmos descritos no capítulo anterior, foi simulado no ambiente MATLAB[®] um sistema PLC baseado na técnica OFDM. Para simular este sistema, foram usados os modelos de canal apresentados na Seção 2.2.2, bem como os modelos de ruído descritos na Seção 2.2.3.

Considerou-se, também, que o sistema possui sincronização perfeita, tanto de frequência e fase das portadoras quanto da temporização de símbolos e do janelamento dos *frames*, e que a duração do prefixo cíclico supera o atraso máximo do canal, o que elimina a interferência entre as portadoras e entre os símbolos.

Apesar dos principais padrões para a camada física recomendarem a faixa de frequência de 2 a 30 MHz, neste trabalho considerou-se a faixa de frequência de 2 a 10 MHz devido à complexidade envolvida nas simulações e ao fato de que os resultados observados nesta faixa podem ser facilmente estendidos para outras faixas de frequência.

A geometria de posicionamento dos tons pilotos adotada é a retangular, com espaçamento entre os pilotos no domínio da frequência representado por PS, conforme mostrada na Figura 8.1. Os tons pilotos são enviados em todos os símbolos OFDM.

Os parâmetros usados nas simulações estão descritos na Tabela 8.1.

O vetor $\hat{\mathbf{H}}(m)$ é obtido interpolando-se os pontos entre os coeficientes de $\hat{\mathbf{H}}_p(m)$. Neste trabalho adotou-se a interpolação linear, uma vez que em [22] foi feita uma comparação de desempenho entre diversas técnicas de interpolação (linear, *spline*, cúbica e *nearest*) em um cenário similar ao apresentado nesta tese e foi observado um melhor desempenho do interpolador linear.

Antes de iniciar as simulações, foi estudado o desempenho dos algoritmos basea-

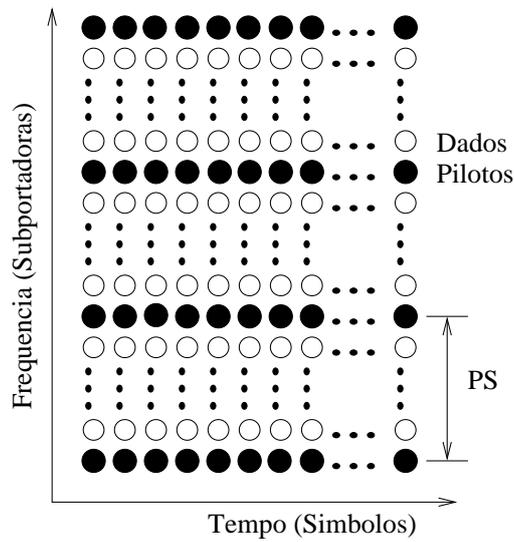


Figura 8.1: Geometria de posicionamento dos tons pilotos nos símbolos OFDM.

Tabela 8.1: Parâmetros usados nas simulações

Descrição	Valores
Número de subportadoras	512
Dimensão da IFFT	1024
Prefixo Cíclico	$8 \mu s$
Frequência Mínima	2 MHz
Frequência Máxima	10 MHz
Modulação	16-QAM e 64-QAM
Técnicas de Estimação	LS; RLS; aLMMSE; LMS; SOM; PSVM e UPSVM
Interpolação	Linear
Técnica de Equalização	MMSE
Geometria de posicionamento dos pilotos	Retangular com PS igual a 2, 4 e 6.

dos em SVM em um cenário com canal invariante no tempo e a presença de ruído impulsivo, para diferentes valores dos parâmetros C , γ e ε , avaliando-se o erro quadrático médio (MSE) resultante das estimações dos coeficientes do canal. Para isso, foram explorados os seguintes intervalos de valores para os parâmetros $C \in [1, 1000]$, $\gamma \in [10^{-4}, 0,5]$ e $\varepsilon \in [0, 1]$. Os melhores resultados foram obtidos para $C = 100$, $\gamma = 0,1$ e $\varepsilon = 0$. Após determinar os valores dos parâmetros SVM, eles foram mantidos fixos durante todos os testes.

As curvas de desempenho dos algoritmos são baseadas nas relações $\text{MSE} \times \text{Número de iterações}$, $\text{MSE} \times \text{SNR}$ e $\text{BER} \times \text{SNR}$, sendo SNR a razão entre a potência do sinal de dados e a potência do ruído colorido de fundo. Para a obtenção das curvas de desempenho foram executadas 100 realizações de cada tipo de canal, e cada realização foi rastreada durante aproximadamente 67 ms ($4T_0$), o que equivale a transmissão de, aproximadamente, 1000 símbolos OFDM para cada realização do canal.

8.2 Comparação de desempenho dos algoritmos

O erro quadrático médio (MSE) das estimações, a cada iteração m , mostrado nas Figuras (8.2, 8.3, 8.4, 8.5) foi estimado pela Equação:

$$\text{MSE}(m) = \frac{1}{P} \sum_{p=1}^P \|H_p(m) - \hat{H}_p(m)\|^2, \quad (8.1)$$

apenas para a análise de convergência e a comparação entre os algoritmos.

Os parâmetros dos algoritmos adaptativos são determinados, inicialmente, para observar a sua velocidade de convergência. Para isso foi usado $\mu = 0,99$ para RLS, PSVM e UPSVM, $\vartheta = 0,005$ para LMS, $\varphi_{max} = 0,2$ e $\phi_{max} = 0,12[\text{rad}]$ para o SOM. Este conjunto de parâmetros será chamado de I . Em todos os testes foram inicializados o vetor $\mathbf{H}_p(m-1)$ como $\mathbf{0}$, e a matriz $\mathbf{S}_p(m-1)$ como $\delta \mathbf{I}$, com $\delta = 0,1$, para todos os algoritmos adaptativos. No aLMMSE foi usado $\varpi = 10000$ em todos os experimentos, de acordo com [68]. A Figura (8.2) exhibe a convergência dos algoritmos com $\text{SNR} = 40$ dB em canal LTI e ruído colorido de fundo.

Pode-se observar na figura 8.2 que o algoritmo UPSVM apresenta convergência a partir do valor de MSE igual a -40 dB, similar aos valores de MSE observados para os algoritmos LS, aLMMSE e SOM. Este fato se deve à realização do cálculo do MSE no final da primeira iteração e o UPSVM calcular o vetor de pesos associado à função-custo em sua formulação no início desta iteração, além da ausência do termo b em sua funcional, como observado na seção 5.1. O PSVM não apresenta essa característica devido ao termo b , que deve ser minimizado junto com a funcional,

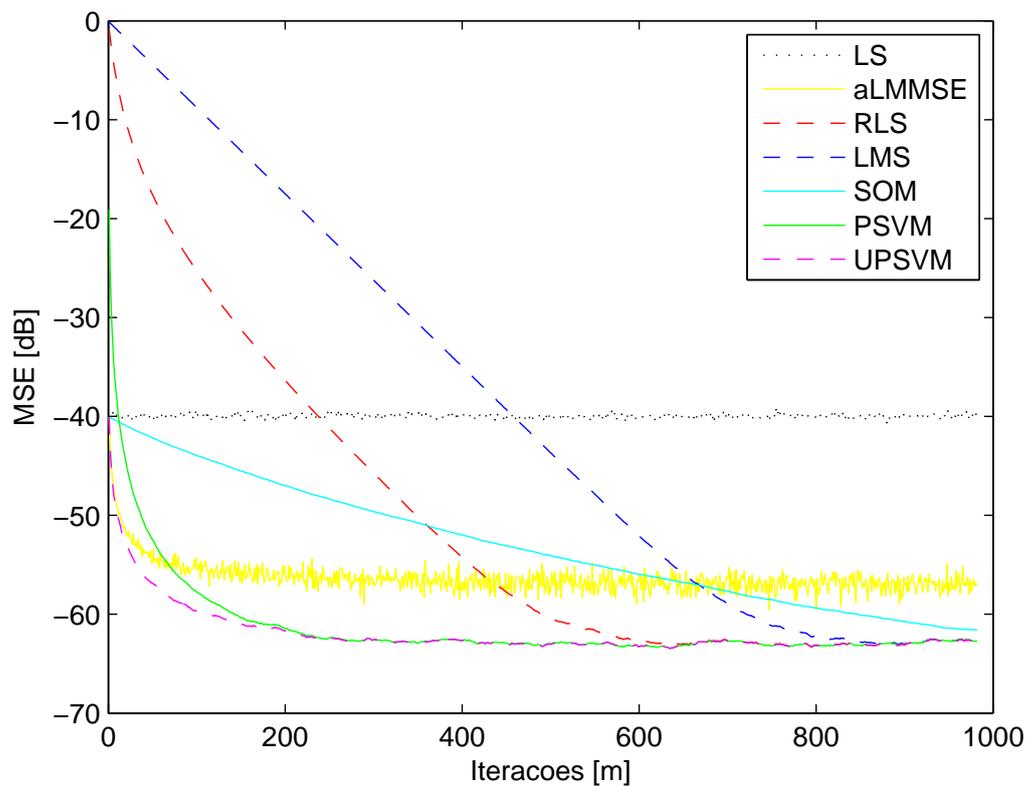


Figura 8.2: Comparação entre as convergências dos algoritmos com o conjunto de parâmetros I e $SNR = 40$ dB em canal LTI e ruído colorido de fundo.

mas o valor inicial de MSE é menor que o valor observado nos algoritmos LMS e RLS, também devido ao cálculo do vetor de pesos, como descrito na seção 4.2. O algoritmo aLMMSE não atinge o valor mínimo de MSE observado nos algoritmos adaptativos, possivelmente devido às considerações de aproximação em sua equação (seção 6.3, algoritmo 6) e à presença de ruído colorido nos dados recebidos. O algoritmo SOM é inicializado com o algoritmo LS, mas apresenta convergência muito lenta.

Para que o cálculo das curvas de taxa de erro de bits (BER - Bit Error Rate) pudesse ter início após a convergência dos algoritmos adaptativos e que a convergência ocorresse até o centésimo símbolo OFDM, os parâmetros dos algoritmos foram ajustados para $\mu = 0,94$ para o RLS, $\mu = 0,98$ para o PSVM, $\mu = 0,99$ para o UPSVM, $\vartheta = 0,033$ para o LMS, $\varphi_{max} = 0,2$ e $\phi_{max} = 0,01[rad]$ para o SOM, esse conjunto de parâmetros será chamado de *II*. A Figura 8.3 mostra o MSE das estimações dos algoritmos com estes novos parâmetros para $SNR = 40dB$ em canal LTI e ruído colorido de fundo.

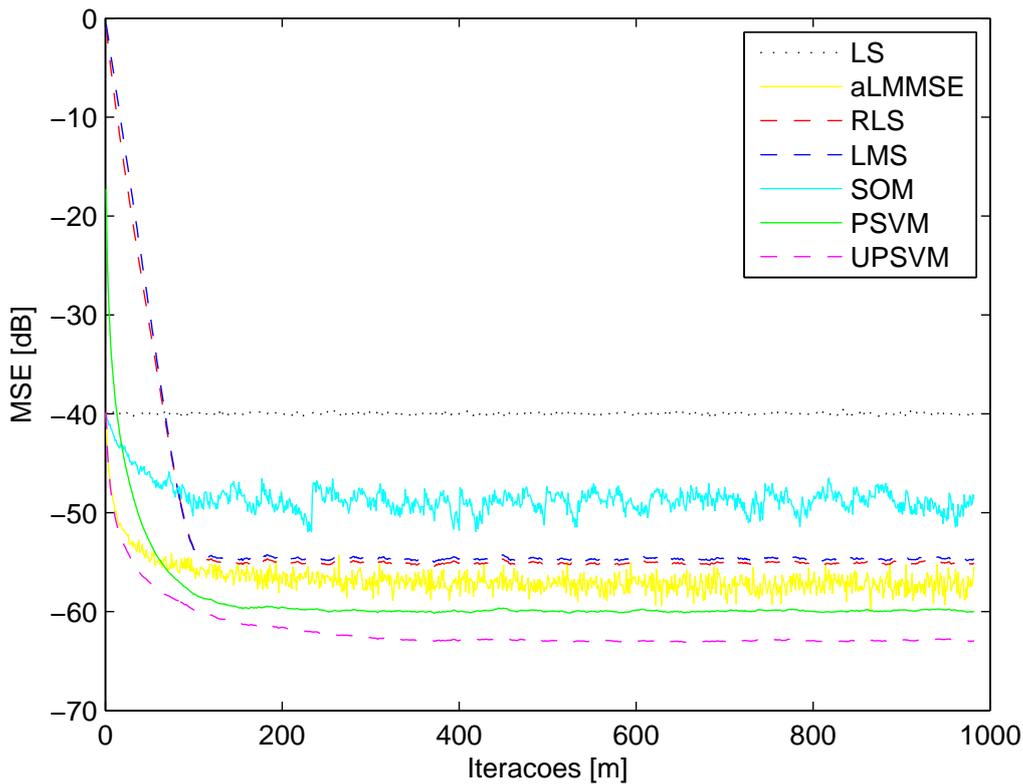


Figura 8.3: Comparação entre as convergências dos algoritmos com o conjunto de parâmetros *II* e $SNR = 40$ dB em canal LTI e ruído colorido de fundo.

Comparando as Figuras 8.2 e 8.3 podemos observar que existe uma relação entre a robustez ao ruído e a velocidade de convergência na escolha dos parâmetros dos algoritmos adaptativos, ou seja, quando mais rápida a convergência, maior o piso

de ruído observado.

A Figura 8.4 mostra a degradação do MSE na curva de aprendizado quando são considerados ruído impulsivo e $SNR = 30$ dB. Nos testes que consideram a presença de ruído impulsivo, este é gerado com variância 20 dB acima da variância do ruído colorido de fundo (esta consideração está de acordo com medições apresentadas em [8]).

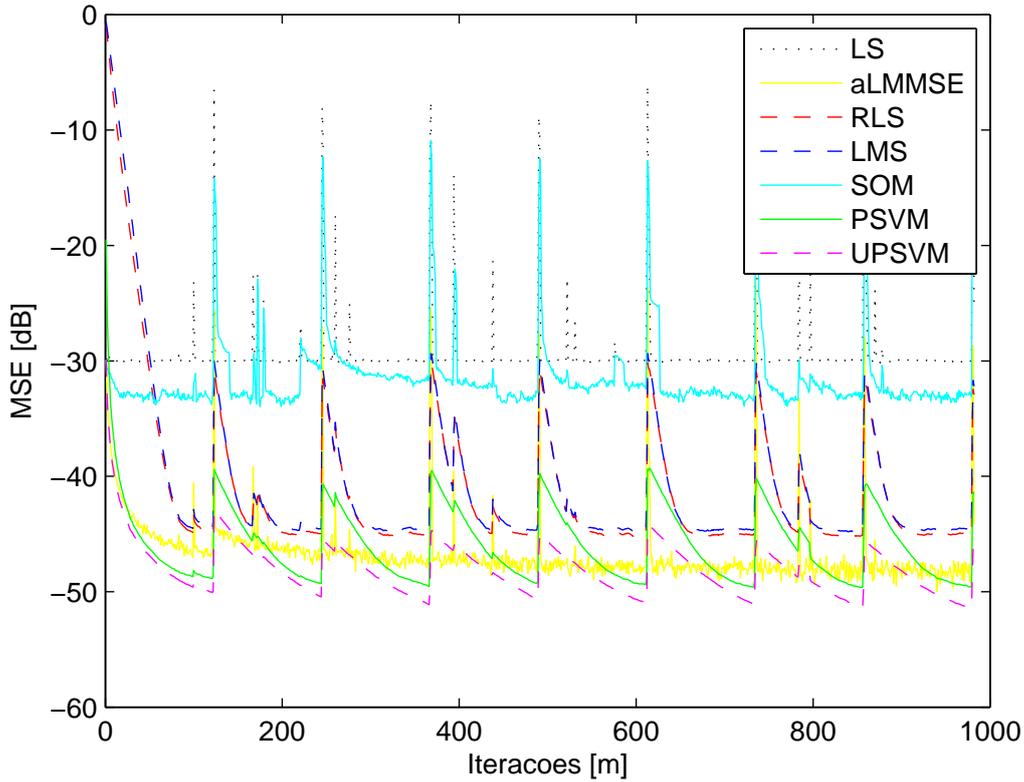


Figura 8.4: Curva de aprendizado dos algoritmos em canal LTI com ruído AIN e $SNR = 30$ dB.

Pode ser observado na Figura 8.4 que os algoritmos LS e SOM são sensíveis a presença de ruído, tanto ao ruído colorido de fundo quanto ao ruído impulsivo, os algoritmos aLMMSE, RLS e LMS apresentam baixa robustez ao ruído impulsivo e os algoritmos baseados em SVM apresentam boa robustez aos dois tipos de ruído, sendo o estimador UPSVM o que apresenta maior robustez a estes ruídos.

Com o intuito de se comparar o MSE das estimações dos algoritmos, após a convergência dos mesmos, para uma faixa de valores de SNR , foi obtido o MSE para valores de SNR de 0 a 30 dB. A Figura 8.5 mostra a comparação deste MSE para ruído colorido de fundo (ACN - *Additive Colored Noise*) e ruído colorido de fundo somado a ruído impulsivo (AIN - *Additive Impulsive Noise*).

A Figura 8.5 corrobora com as conclusões anteriores sobre a sensibilidade dos

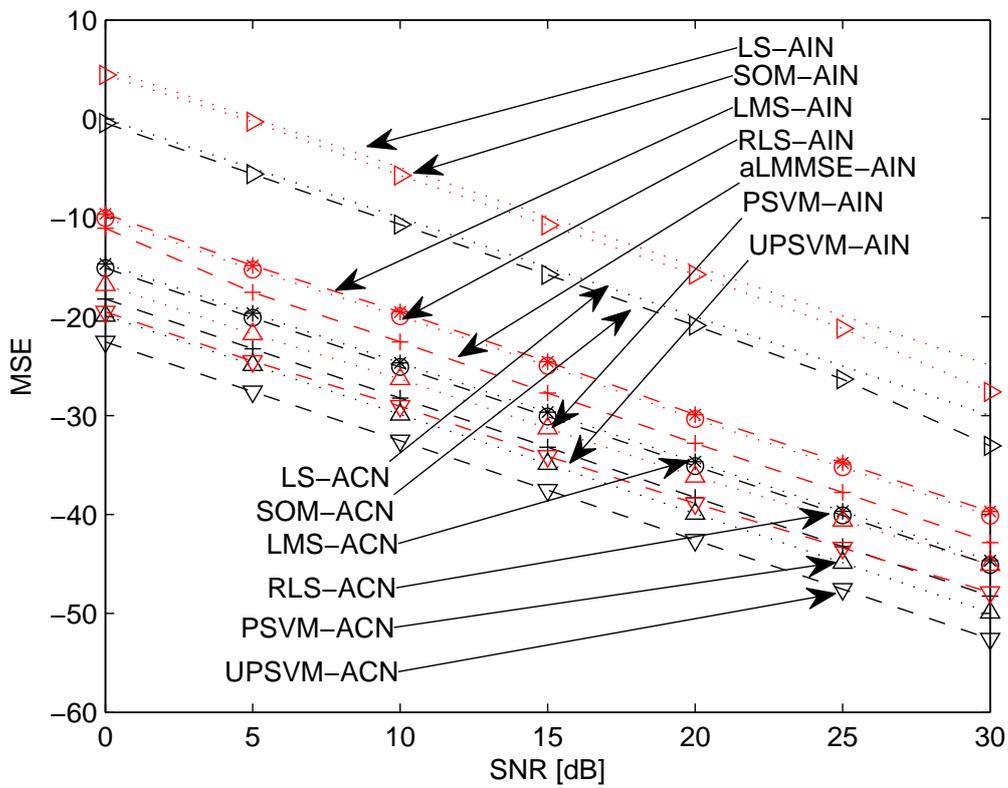


Figura 8.5: $MSE \times SNR$ em canal LTI com ruídos ACN e AIN.

algoritmos ao ruído e confirma o bom desempenho do estimador UPSVM para todos os valores de SNR avaliados.

Em um teste preliminar, usamos o modelo LTI de canal com ruído colorido de fundo para observar o desempenho da BER em relação à separação entre pilotos no domínio da frequência. Foram testadas as separações $PS = 2, 4$ e 6 . Tais valores foram escolhidos, iniciando-se com o de menor separação ($PS = 2$), observando o desempenho, e aumentando a separação até que uma degradação significativa da BER pudesse ser observada. As Figuras 8.6 e 8.7 mostram o resultado deste teste para as modulações 16-QAM e 64-QAM, onde se considerou como ideal o resultado obtido a partir do conhecimento pleno do canal ($\mathbf{H}(m)$), apenas para comparação, e aLMMSE, RLS, LMS, PSVM, UPSVM, LS e SOM os resultados obtidos com os coeficientes resultantes da interpolação linear de $\hat{\mathbf{H}}_p(m)$, obtidos por cada um dos respectivos estimadores.

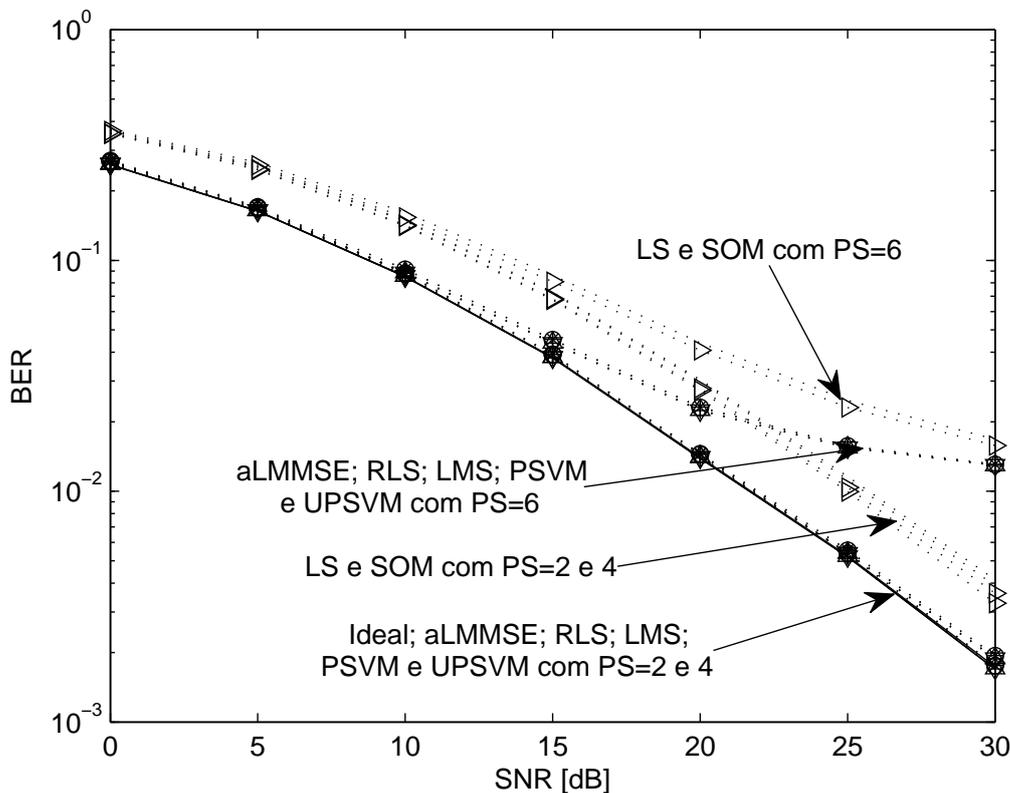


Figura 8.6: BER \times SNR em canais LTI com ruído colorido de fundo e modulação 16-QAM para valores de separação entre pilotos iguais a 2, 4 e 6.

As Figuras 8.6 e 8.7 mostram que $PS = 4$ é a maior separação de pilotos que não gera degradação adicional significativa na BER, para ambas as modulações. Por este fato, os testes seguintes irão considerar $PS = 4$, uma vez que apresenta a menor taxa de pilotos e com isso a melhor “vazão”.

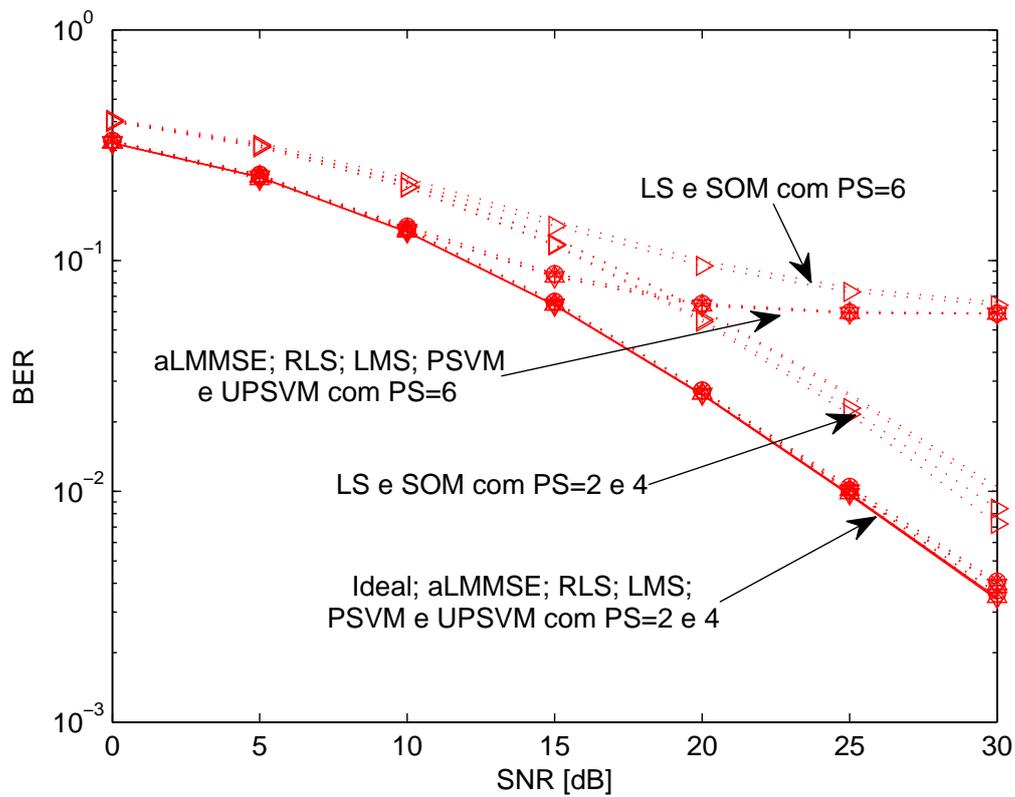


Figura 8.7: $BER \times SNR$ em canais LTI com ruído colorido de fundo e modulação 64-QAM para valores de separação entre pilotos iguais a 2, 4 e 6.

Para avaliar o desempenho dos algoritmos em canais variantes no tempo, foram obtidas as curvas de MSE considerando o canal LPTV com variação temporal alternando entre dois estados do canal (variação comutada) e variação temporal harmônica, e a presença de ruído ACN. As Figuras 8.8 e 8.9 mostram as curvas de MSE nos canais LPTV com variação comutada e harmônica, respectivamente, para valores de SNR de 0 a 30 dB.

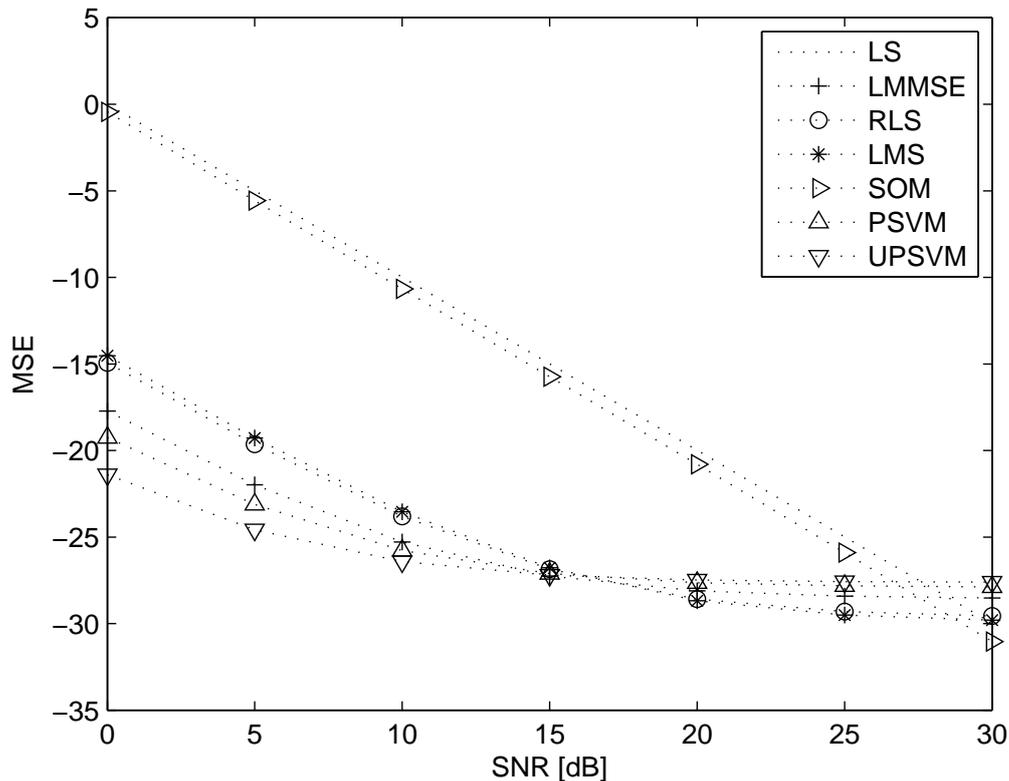


Figura 8.8: MSE das estimações do canal LPTV com variação comutada e ruído ACN.

Pode-se observar nas Figuras 8.8 e 8.9 que a rápida variação do canal LPTV impede que os algoritmos adaptativos consigam rastreá-lo. Nestes casos, os algoritmos RLS e LMS apresentam desempenho um pouco melhor que os algoritmos PSVM, UPSVM e aLMMSE, para valores de SNR superiores a 15 dB, quando se considera a variação comutada (Fig. 8.8); e superiores a 10 dB, quando se considera a variação harmônica (Fig. 8.9). Este fato se deve aos valores dos parâmetros destes algoritmos que, para estes valores de SNR , fazem com que a capacidade de rastreamento do canal resulte em melhor desempenho quando comparado ao obtido com maior robustez ao ruído de fundo. Os algoritmos não-adaptativos LS e SOM apresentam os melhores desempenhos para valores de SNR próximos de 30 dB, quando o canal LPTV apresenta variação comutada (Fig. 8.8), e acima de 25 dB para variação harmônica

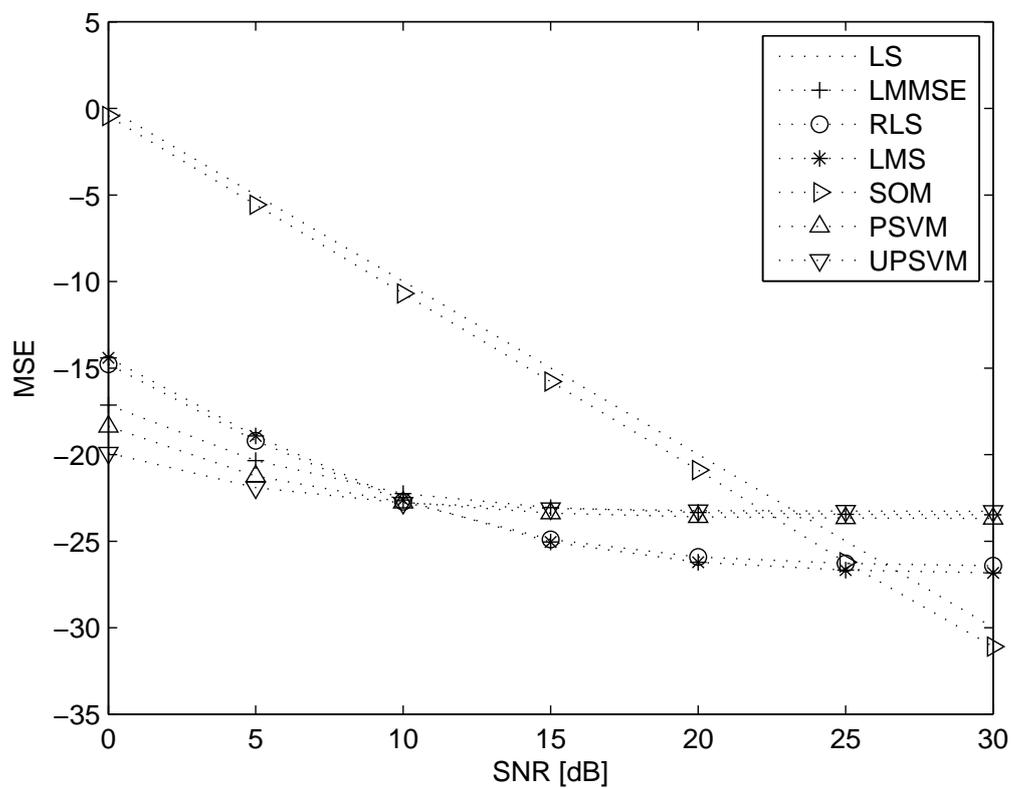


Figura 8.9: MSE das estimações do canal LPTV com variação harmônica e ruído ACN.

(Fig. 8.9). Contudo, os algoritmos RLS, LMS, LS e SOM apresentam significativa sensibilidade ao ruído impulsivo, visto na Figura 8.4.

Para avaliar o quanto a rápida variação do canal LPTV pode interferir na BER, foram calculadas as taxas de erro de bit para a modulação 16-QAM, considerando estes tipos de canais e ruído. As Figuras 8.10 e 8.11 mostram as curvas de BER nos canais LPTV com variação comutada e harmônica, respectivamente, para valores de SNR de 0 a 30 dB.

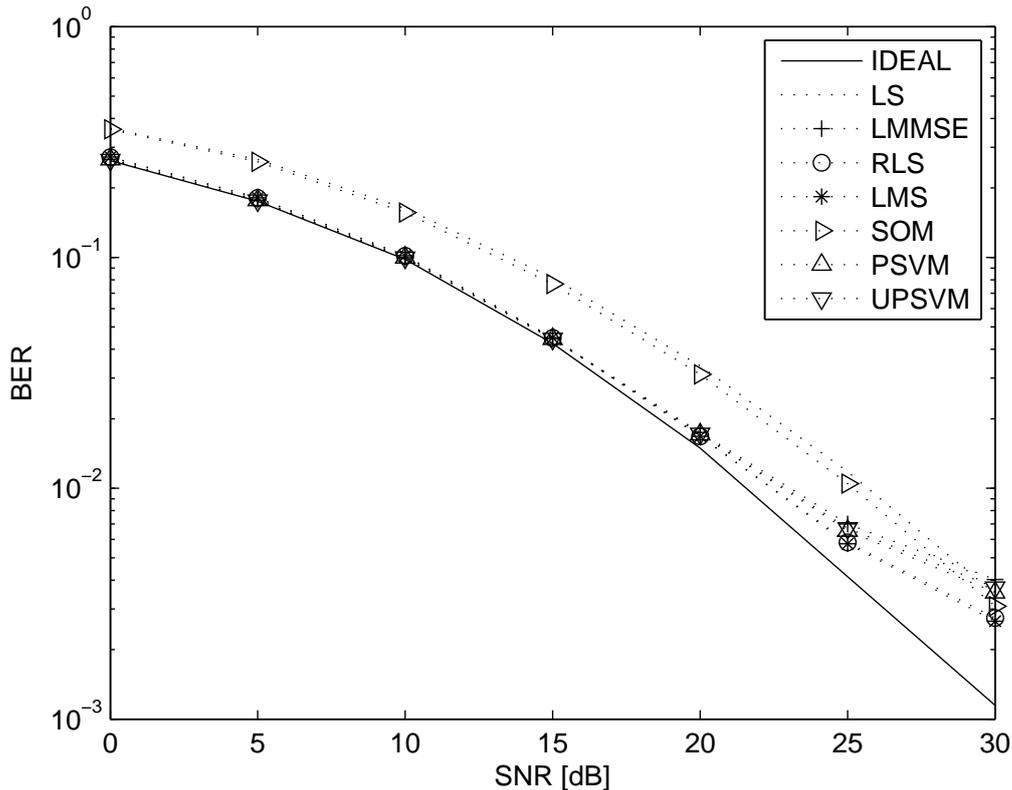


Figura 8.10: BER em função da SNR com modulação 16-QAM em canal LPTV com variação comutada e ruído ACN.

A Figura 8.10 mostra que quando o canal LPTV apresenta variação comutada, os algoritmos RLS e LMS apresentam um desempenho de BER um pouco melhor que os demais algoritmos, para SNR entre 20 e 30 dB, mas estes algoritmos não apresentam boa robustez ao ruído impulsivo, como visto anteriormente.

Na Figura 8.11 observamos que, quando se considera o canal LPTV com variação harmônica, os algoritmos RLS e LMS apresentam melhor desempenho de BER para valores de SNR entre 15 e 25 dB, sendo superado pelos algoritmos LS e SOM para valores de SNR superiores a 25 dB, e o desempenho de BER do algoritmo SOM é ligeiramente melhor que o LS. Novamente, vale lembrar que estes algoritmos são sensíveis ao ruído impulsivo.

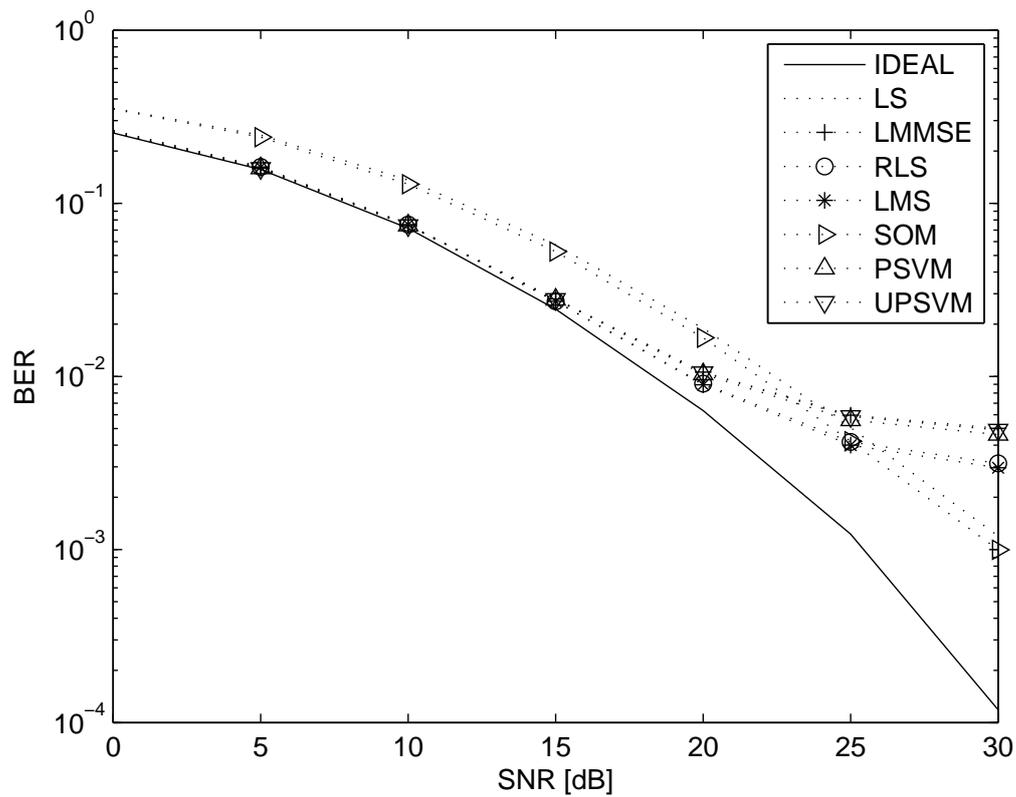


Figura 8.11: BER em função da SNR com modulação 16-QAM em canal LPTV com variação harmônica e ruído ACN.

8.3 Técnica de reinicialização do estimador UPSVM

Como foi mostrado, a rápida variação do canal LPTV degrada significativamente a BER quando se usam algoritmos adaptativos para a estimação do canal, e os algoritmos não-adaptativos são extremamente sensíveis ao ruído. Com o objetivo de contornar esse problema e fazendo uso da rápida convergência do algoritmo UPSVM, propomos o uso de uma técnica para obter um limiar para reinicialização do algoritmo.

Sabe-se que

$$\begin{aligned} \mathbf{Y}_p(m) &= \mathbf{D}_{\{\mathbf{H}_p(m)\}} \mathbf{X}_p(m) + \mathbf{V}_p(m) = \\ &\mathbf{D}_{\{\mathbf{X}_p(m)\}} [\mathbf{H}_p(m) + \mathbf{D}_{\{\mathbf{X}_p(m)\}}^{-1} \mathbf{V}_p(m)], \end{aligned} \quad (8.2)$$

se $\exists \mathbf{D}_{\{\mathbf{X}_p(m)\}}^{-1}$.

Seja $\hat{\mathbf{H}}_p(m)$ uma estimativa do canal obtida após um intervalo de tempo suficiente para que o estimador tenha convergido. Usando $\hat{\mathbf{H}}_p(m)$, a estimativa da saída do canal pode ser escrita por:

$$\hat{\mathbf{Y}}_p(m) = \mathbf{D}_{\{\mathbf{X}_p(m)\}} \hat{\mathbf{H}}_p(m), \quad (8.3)$$

e o erro de estimação da saída do canal dado por:

$$\begin{aligned} \mathbf{Y}e_p(m) &= \mathbf{Y}_p(m) - \hat{\mathbf{Y}}_p(m) = \mathbf{Y}_p(m) - \mathbf{D}_{\{\mathbf{X}_p(m)\}} \hat{\mathbf{H}}_p(m) = \\ &\mathbf{D}_{\{\mathbf{X}_p(m)\}} [\mathbf{H}_p(m) - \hat{\mathbf{H}}_p(m)] + \mathbf{V}_p(m). \end{aligned} \quad (8.4)$$

Definindo $\hbar(m)$, como:

$$\hbar(m) = \|\mathbf{D}_{\{\mathbf{X}_p(m)\}}^{-1} \mathbf{Y}e_p(m)\| = \|\mathbf{H}_p(m) - \hat{\mathbf{H}}_p(m) + \mathbf{D}_{\{\mathbf{X}_p(m)\}}^{-1} \mathbf{V}_p(m)\|, \quad (8.5)$$

e calculando $\hbar(m)$ em dois instantes de tempo, digamos m e m_s , temos:

$$\begin{aligned} \Delta \hbar(m) &= \hbar(m) - \hbar(m_s) = \|\mathbf{H}_p(m) - \hat{\mathbf{H}}_p(m) + \mathbf{D}_{\{\mathbf{X}_p(m)\}}^{-1} \mathbf{V}_p(m)\| \\ &- \|\mathbf{H}_p(m_s) - \hat{\mathbf{H}}_p(m_s) + \mathbf{D}_{\{\mathbf{X}_p(m_s)\}}^{-1} \mathbf{V}_p(m_s)\|. \end{aligned} \quad (8.6)$$

A desigualdade triangular no \Re^n [69] mostra que:

$$\|\mathbf{a}(m)\| - \|\mathbf{a}(m_s)\| \leq \|\mathbf{a}(m) - \mathbf{a}(m_s)\|. \quad (8.7)$$

Logo,

$$\begin{aligned} \Delta \tilde{h}(m) \leq & \| \mathbf{H}_p(m) - \mathbf{H}_p(m_s) - \hat{\mathbf{H}}_p(m) + \hat{\mathbf{H}}_p(m_s) + \\ & \mathbf{D}_{\{\mathbf{x}_p(m)\}}^{-1} \mathbf{V}_p(m) - \mathbf{D}_{\{\mathbf{x}_p(m_s)\}}^{-1} \mathbf{V}_p(m_s) \| \end{aligned} \quad (8.8)$$

Definindo:

$$\begin{cases} \Delta \mathbf{H}_p(m) = \mathbf{H}_p(m) - \mathbf{H}_p(m_s); \\ \Delta \hat{\mathbf{H}}_p(m) = \hat{\mathbf{H}}_p(m) - \hat{\mathbf{H}}_p(m_s); \\ \Delta \mathbf{V}_p(m) = \mathbf{V}_p(m) - \mathbf{V}_p(m_s), \end{cases} \quad (8.9)$$

e, considerando $\mathbf{D}_{\{\mathbf{x}_p(m)\}}^{-1} = \mathbf{D}_{\{\mathbf{x}_p(m_s)\}}^{-1}$, ou seja, que o mesmo vetor de pilotos seja enviado em todos os símbolos OFDM, a Equação (8.8) torna-se:

$$\Delta \tilde{h}(m) \leq \| \Delta \mathbf{H}_p(m) - \Delta \hat{\mathbf{H}}_p(m) + \mathbf{D}_{\{\mathbf{x}_p(m)\}}^{-1} \Delta \mathbf{V}_p(m) \|. \quad (8.10)$$

A partir da Equação (8.10), pode-se esperar que em estado estacionário, para canais invariante no tempo, após a convergência do algoritmo, o valor de $\Delta \tilde{h}(m)$ ultrapasse um determinado limiar caso ocorra variação abrupta do canal ou presença de ruído impulsivo.

A partir das observações anteriores, formularam-se duas hipóteses:

$$\begin{cases} Hip_0 = \text{O canal não sofre variação brusca e} \\ \text{o ruído não contém parcela de ruído impulsivo;} \\ Hip_1 = \text{O canal sofre variação brusca ou} \\ \text{o ruído contém parcela de ruído impulsivo,} \end{cases} \quad (8.11)$$

e, com o intuito de poder considerar o envio de diferentes vetores de tons pilotos ($\mathbf{D}_{\{\mathbf{x}_p(m)\}}^{-1} \neq \mathbf{D}_{\{\mathbf{x}_p(m_s)\}}^{-1}$), propôs-se realizar um teste de hipóteses usando como observação o quadrado da norma de $\mathbf{Y}\mathbf{e}_p(m)$, que pode ser expressa, supondo o uso de modulação M-PSK para os tons pilotos, por:

$$\begin{aligned} \|\mathbf{Y}\mathbf{e}_p(m)\|^2 = & \mathbf{A}^2 \|\mathbf{H}_p(m) - \hat{\mathbf{H}}_p(m)\|^2 + \|\mathbf{V}_p(m)\|^2 + \\ & 2\Re[\mathbf{V}_p^H(m) \mathbf{D}_{\{\mathbf{x}_p(m)\}} (\mathbf{H}_p(m) - \hat{\mathbf{H}}_p(m))], \end{aligned} \quad (8.12)$$

onde \mathbf{A} é a amplitude da constelação M-PSK.

Pode-se esperar que $E\{\|\mathbf{Y}\mathbf{e}_p\|^2 | Hip_1\} > E\{\|\mathbf{Y}\mathbf{e}_p\|^2 | Hip_0\}$, o que motiva o teste na forma:

$$\|\mathbf{Y}\mathbf{e}_p(m)\|^2 \underset{Hip_0}{\overset{Hip_1}{\geq}} \text{limiar}, \quad (8.13)$$

onde $\mathbf{Y}\mathbf{e}_p(m)$ é uma amostra do vetor aleatório $\mathbf{Y}\mathbf{e}_p$ e *limiar* é um limiar a ser

escolhido.

A proposta de escolha do limiar é dada por: $limiar = \varphi \|\mathbf{Ye}_p(m_s)\|^2$, onde $\|\mathbf{Ye}_p(m_s)\|^2$ é uma estimativa de $E\{\|\mathbf{Ye}_p\|^2 | Hip_0\}$ obtida previamente em um instante m_s , e $\varphi > 1$ é um fator a ser ajustado.

Para ajustar o valor φ , selecionou-se o fator de proporcionalidade entre $\|\mathbf{Ye}_p(m)\|^2$ e $\|\mathbf{Ye}_p(m_s)\|^2$ que resultasse em melhor desempenho da BER. Para isso, foram feitas simulações considerando o conjunto de fatores $\{1, 5, 1,75, 2, 3, 4\}$, como mostra a Figura 8.12. Para obtenção desta curva de desempenho foi mantido fixo o valor da SNR em 40 dB, executadas 100 realizações do canal LPTV com adição de ruído impulsivo (AIN) e cada realização foi rastreada durante aproximadamente 67 ms ($4T_0$).

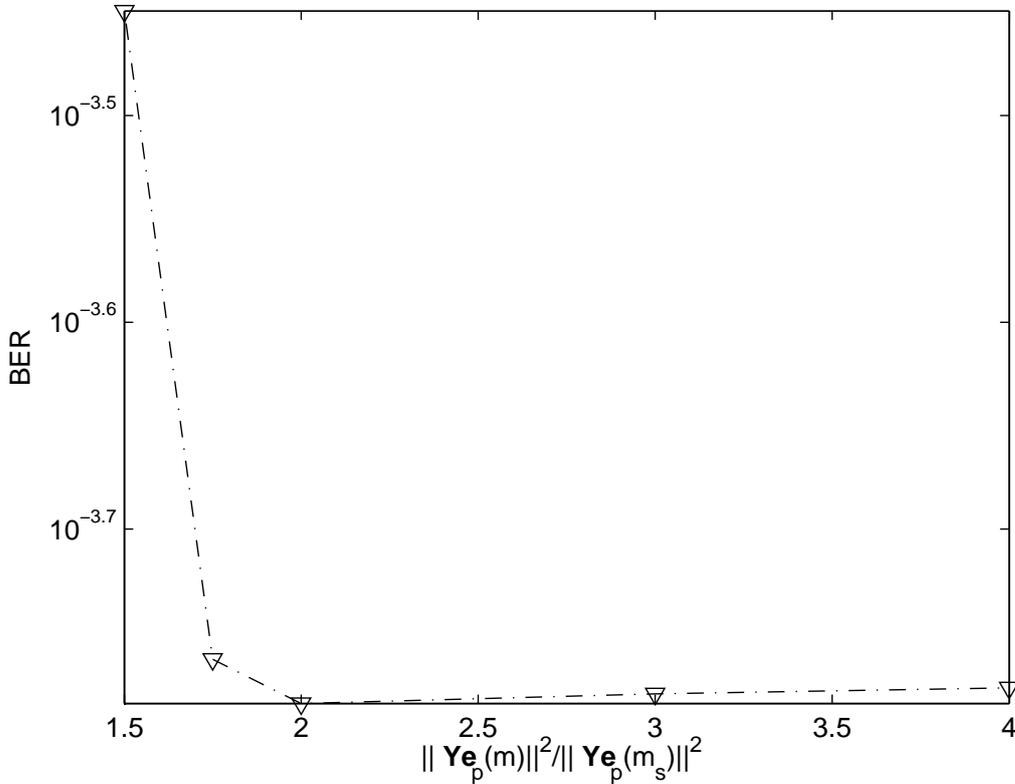


Figura 8.12: BER para diferentes escolhas de limiar, $SNR = 40$ dB, modulação 16-QAM e ruído AIN.

Como observado na curva da Figura 8.12, o fator que resulta em menor BER é $\varphi = 2$:

$$\|\mathbf{Ye}_p(m)\|^2 \underset{Hip_0}{\overset{Hip_1}{\geq}} 2\|\mathbf{Ye}_p(m_s)\|^2 \quad (8.14)$$

Vale lembrar que a escolha do limiar pode ser realizada através de um classifi-

cador baseado em SVM. Neste caso, seria necessário um conjunto de treinamento com “rótulos” para as amostras cuja norma quadrática representa a ocorrência de variação brusca do canal ou de ruído impulsivo, e “rótulos” para as amostras que representam a ausência dessas ocorrências. Como este procedimento deve ser feito *off-line* e apresenta um custo computacional elevado, optou-se por usar uma heurística mais simples.

A técnica proposta está descrita no algoritmo 9.

Algoritmo 9: Reinicialização do UPSVM

A cada iteração, calcule:

$$\Delta \|\mathbf{Y}\mathbf{e}_p(m)\|^2 = \|\mathbf{Y}\mathbf{e}_p(m)\|^2 - \|\mathbf{Y}\mathbf{e}_p(m_s)\|^2;$$

Se $\Delta \|\mathbf{Y}\mathbf{e}_p(m)\|^2 > \|\mathbf{Y}\mathbf{e}_p(m_s)\|^2$, faça:

$$\mathbf{H}_{p,UPSVM}(m-1) = \mathbf{0};$$

$$\mathbf{S}_{p,UPSVM}(m-1) = \delta \mathbf{I};$$

E refaça a estimação de $\mathbf{H}_{p,UPSVM}(m)$.

Para validação da técnica proposta foi verificado o desempenho do algoritmo UPSVM em ambiente com canal invariante no tempo (LTI) apenas com ruído colorido (ACN), e com canal variante no tempo (LPTV) com ruído impulsivo (AIN), como mostra a Figura 8.13.

Pode-se observar na Figura 8.13 que, para ambientes com canal LTI sem a presença de ruído impulsivo, a técnica proposta não altera o MSE, e para ambientes com canal LPTV e ruído impulsivo, tem-se uma significativa redução do MSE, principalmente para valores de *SNR* acima de 15 dB.

Com o intuito de verificar se a técnica proposta é adequada a outros algoritmos adaptativos, foram feitos testes aplicando a técnica de reinicialização aos algoritmos RLS, LMS, PSVM e UPSVM. As Figuras 8.14 e 8.15 mostram o desempenho do MSE e da BER para estes algoritmos, respectivamente.

Na Figura 8.14 é possível observar que a técnica proposta não é adequada aos algoritmos RLS e LMS, pois a reinicialização dos mesmos degrada sobremaneira o MSE. O que fica ainda mais crítico para *SNR* acima de 5 dB, onde, além do ruído impulsivo, as variações do canal produzem um expressivo número de reinicializações dos algoritmos. O algoritmo PSVM têm bom desempenho com a técnica proposta, possivelmente devido ao vetor de pesos associado à função-custo em sua formulação, obtido no início de cada iteração. Contudo, a necessidade de minimizar o parâmetro *b* em sua funcional, ao reinicializar o algoritmo, faz com que o MSE resultante seja maior que o observado no algoritmo UPSVM.

A Figura 8.15 corrobora com as observações anteriores e evidencia que a técnica proposta é mais adequada ao algoritmo UPSVM.

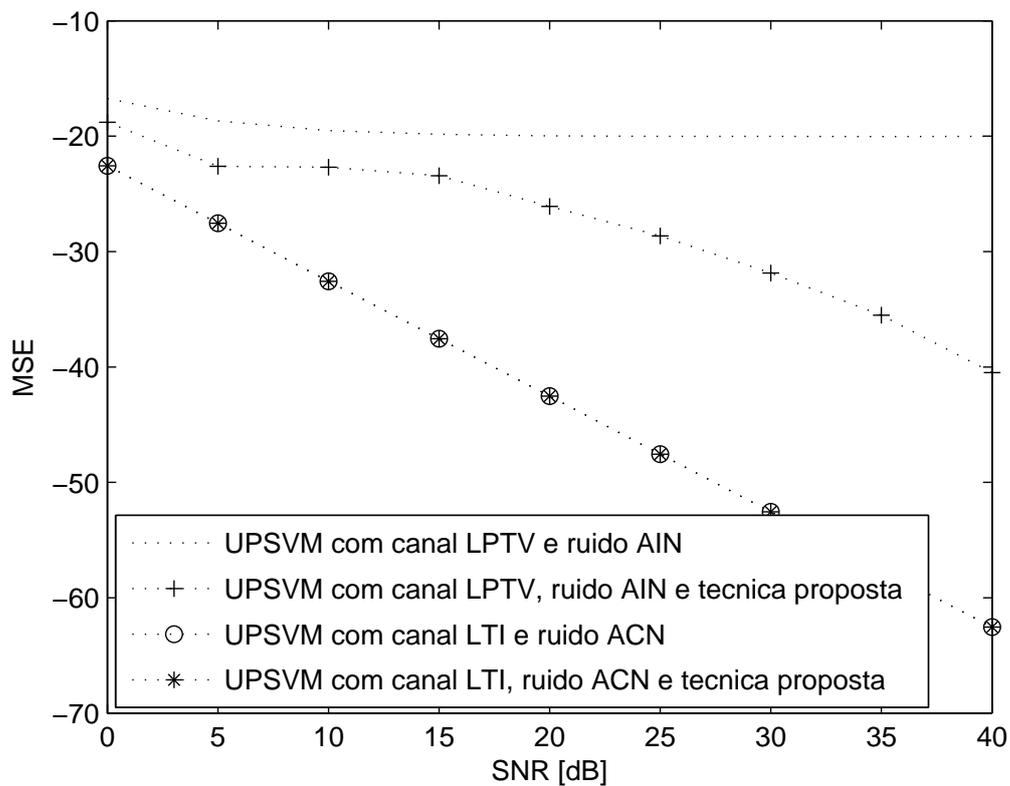


Figura 8.13: MSE da estimação do canal LPTV com variação comutada e ruído AIN, para os algoritmos adaptativos com a técnica de reinicialização proposta.

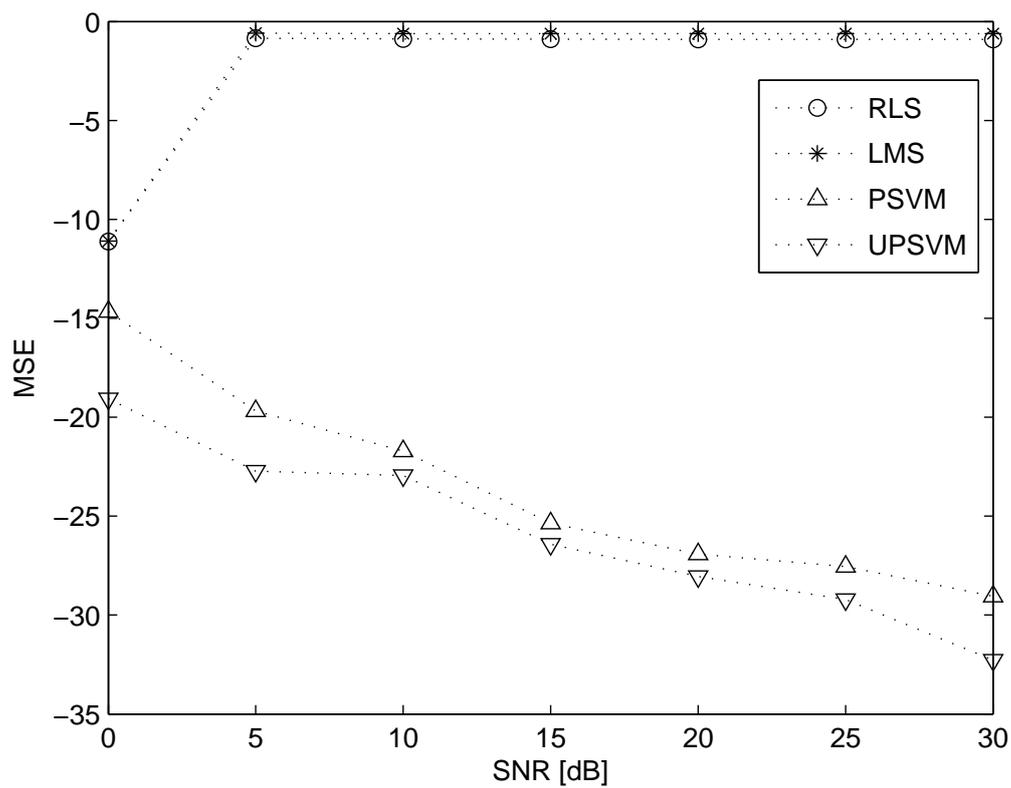


Figura 8.14: MSE em função da SNR em canal LPTV com variação comutada e ruído AIN, para os algoritmos adaptativos com uso da técnica de reinicialização proposta.

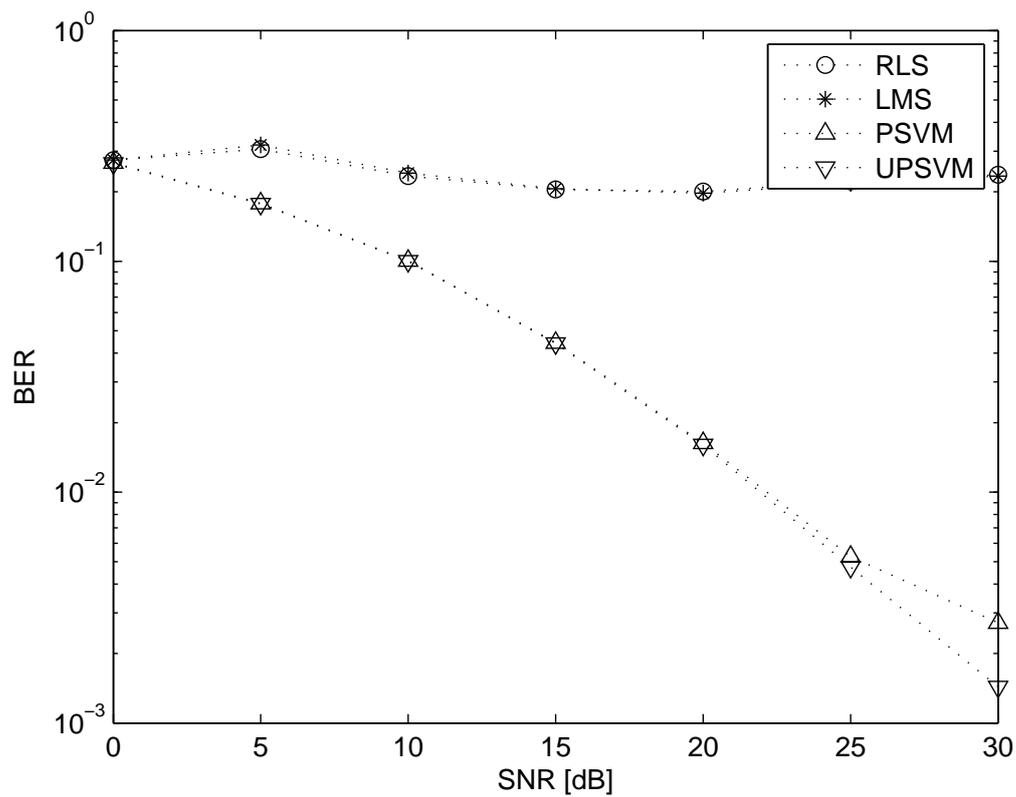


Figura 8.15: BER em função da SNR em canal LPTV com variação comutada e ruído AIN, para os algoritmos adaptativos com uso da técnica de reinicialização proposta.

8.4 Análise de desempenho com a técnica de reinicialização do UPSVM

As Figuras 8.16 e 8.17 mostram o MSE dos estimadores em canais LPTV com variação harmônica e comutada, respectivamente, ambos com ruído ACN. O algoritmo UPSVM usa a técnica de verificação da variação do erro de estimação e reinicialização proposta.

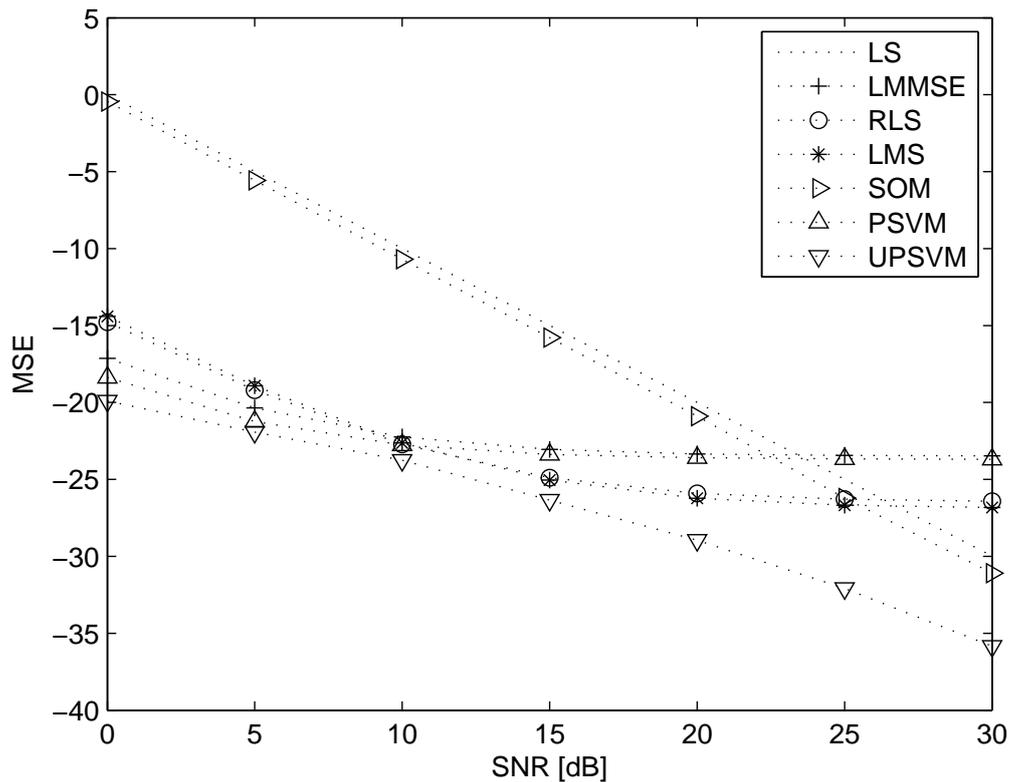


Figura 8.16: MSE da estimação do canal LPTV com variação harmônica e ruído ACN.

Comparando as Figuras 8.8 e 8.9 com as Figuras 8.16 e 8.17 podemos observar que o uso da técnica proposta reduz significativamente o MSE das estimações do algoritmo UPSVM para valores de SNR acima de 10 dB, mostrando que o algoritmo consegue rastrear com mais eficiência as variações do canal LPTV.

Para completar a análise, foi obtida a média da BER nos canais LPTV em ambas as variações (harmônica e comutada) na presença de ruído AIN, com as modulações 16QAM e 64-QAM. Os resultados estão nas Figuras 8.18 e 8.19.

As Figuras 8.18 e 8.19 mostram que a técnica proposta para o estimador UPSVM resulta em uma BER menor quando comparado aos outros estimadores para SNR acima de 20 dB. Estes resultados de BER para este tipo de canal são consistentes

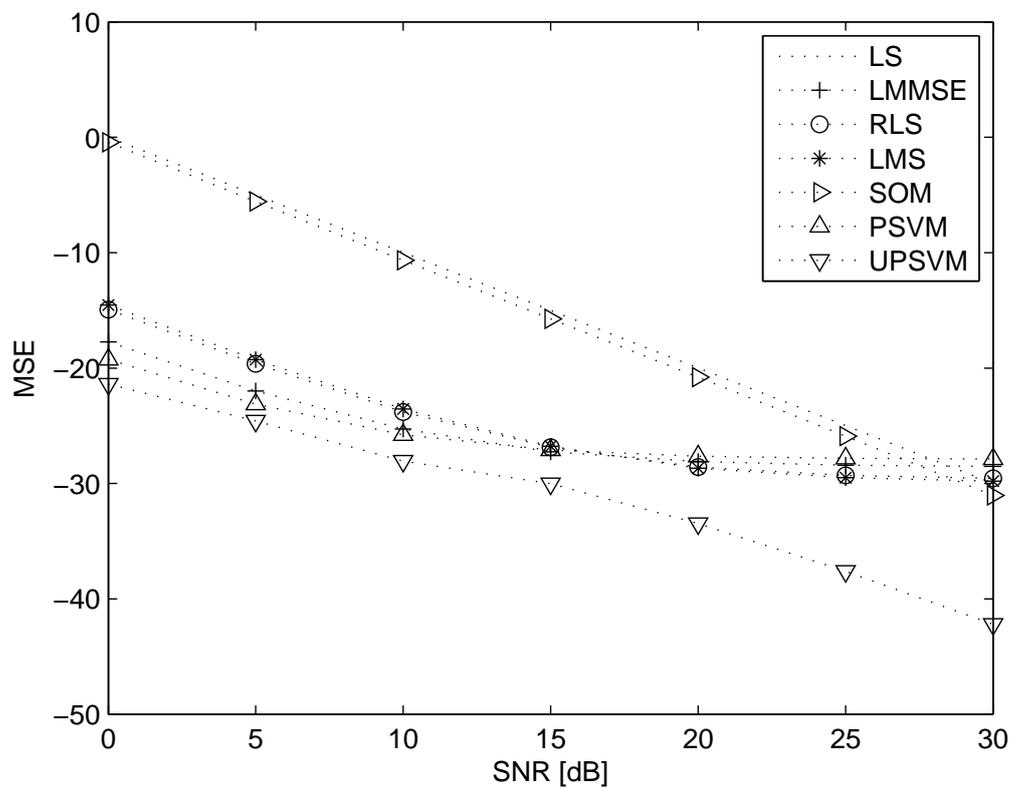


Figura 8.17: MSE da estimação do canal LPTV com variação comutada e ruído ACN.

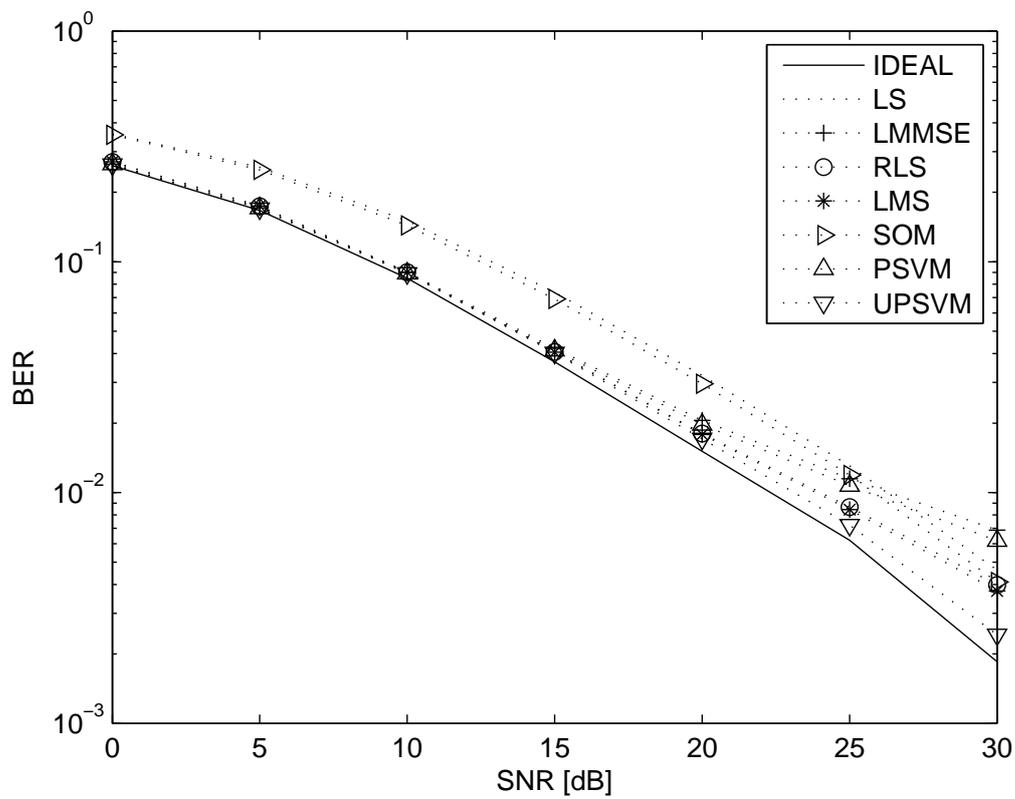


Figura 8.18: BER em função da SNR com modulação 16-QAM.

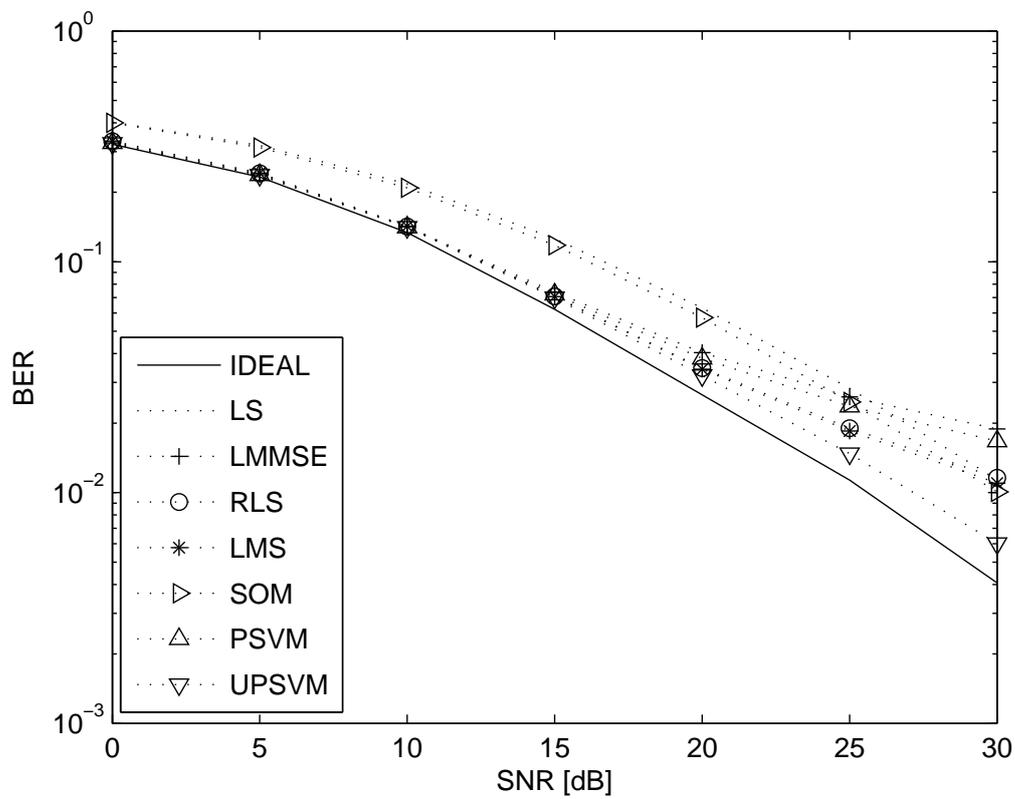


Figura 8.19: BER em função da SNR com modulação 64-QAM.

com diversos resultados apresentados na literatura, dentre eles [22, 70, 71].

Capítulo 9

Conclusões

Nesta tese foram descritos: um modelo de sistema PLC-DMT, três modelos de canais (um invariante no tempo - LTI, um com variação harmônica e um com comutação entre dois estados da resposta impulsiva do canal, os dois últimos com variação sincronizada à frequência fundamental da rede elétrica - LPTV) adequados aos sistemas PLC *indoor* e um modelo de ruído aditivo composto de ruído colorido de fundo, ruído impulsivo síncrono e ruído impulsivo aperiódico.

As principais contribuições foram a formulação de dois algoritmos *fast-SVM* (PSVM e UPSVM) como regressores de valores complexos para a estimação adaptativa de canal e a proposta de uma técnica de reinicialização do estimador UPSVM baseado em um limiar para a variação do erro de estimação do canal, medido em cada iteração.

Foi observado nas simulações que as técnicas *fast-SVM* superam os estimadores aLMMSE, RLS, LMS, SOM e LS em cenários com canais LTI e ruídos aditivos colorido de fundo (ACN) e impulsivo (AIN). Nos canais LPTV, foi observado que a rápida variação do canal impede que os algoritmos adaptativos consigam rastreá-lo. Neste caso, o algoritmo UPSVM com a técnica de reinicialização proposta apresentou um bom desempenho de MSE e de BER, mostrando ser adequado para a estimação de canal LPTV em ambientes com ruído impulsivo.

O algoritmo UPSVM apresenta um desempenho melhor que o algoritmo PSVM devido ao fato de sua inicialização não necessitar da minimização do parâmetro b , resultando em uma convergência mais rápida.

A técnica proposta permite uma implementação relativamente simples e a função-custo utilizada é robusta a diversos tipos de ruídos. As características da técnica proposta se mostraram interessantes para cenários PLC, onde se observa um comportamento ciclo-estacionário do canal e a presença de ruído não-Gaussiano.

Os tópicos que podem resultar em pesquisas futuras a partir desse trabalho são:

- Análise dos algoritmos *fast-SVM* como preditores de canal. Esta análise pode

permitir uma redução significativa de pilotos e o conseqüente aumento da vazão, uma vez que o envio de novos pilotos para a estimação do canal só se fará necessário em intervalos onde o alcance de predição atinge níveis inaceitáveis.

- Estudo da estimação e predição conjunta de canal utilizando os algoritmos *fast-SVM*. Este procedimento pode resultar em menor complexidade computacional quando comparado aos procedimentos de estimação e predição isolados e manter a mesmo desempenho.
- Estender a pesquisa dos algoritmos *fast-SVM* para a abordagem MIMO-OFDM.

Referências Bibliográficas

- [1] DAS MINAS E ENERGIA, M. “Programa Luz para Todos”. Website, 2003. <http://www.luzparatodos.mme.gov.br/luzparatodos> - visitado em Maio de 2012.
- [2] DO PROGRAMA DE INCLUSÃO DIGITAL, C. G. “Programa de inclusão digital”. Website, 2009. <http://www4.planalto.gov.br/brasilconectado/CGPID> - visitado em novembro de 2012.
- [3] LEITE, L. H. M. “Utilização de tecnologia PLC para suporte à aplicações Smart-Grid.” In: *ANEEL - Workshop Internacional Power Line Communication*, Brasília, Brasil, Maio 2007.
- [4] PAVLIDOU, N.; HAN VINCK, A. J. J. H. B. “Power line communications: State of art and future trends”, *IEEE Commun. Mag.*, v. 41, n. 4, pp. 34–40, Apr. 2003. ISSN: 0163-6804. doi: 10.1109/MCOM.2003.1193972.
- [5] CANETE, F.J.; CORTES, J. “Modeling and evaluation of the indoor power line transmission medium”, *IEEE Communications Magazine*, v. 41, n. 4, pp. 41–47, Apr. 2003. ISSN: 0163-6804. doi: 10.1109/MCOM.2003.1193973.
- [6] CORRIPIO, F., ARRABAL, J., DEL RIO, L., et al. “Analysis of the cyclic short-term variation of indoor power line channels”, *IEEE Journal on Selected Areas in Communications*, v. 24, n. 7, pp. 1327–1338, Jul. 2006. ISSN: 0733-8716. doi: 10.1109/JSAC.2006.874402.
- [7] UMEHARA, D.; HAYASAKI, T. D. S. M. M. “The influence of time-varying channels synchronized with commercial power supply on PLC equipments”. In: *IEEE International Symposium on Power Line Communications and its Application*, pp. 30–35, Jeju, Korea, Mar. 2008.
- [8] ZIMMERMANN, M.; DOSTERT, K. “Analysis and modeling of impulsive noise in broad-band powerline communications”, *IEEE Transactions on electromagnetic compatibility*, v. 44, n. 1, pp. 249–258, Feb. 2002. ISSN: 0018-9375. doi: 10.1109/15.990732.

- [9] CHAOUICHE, H.; GAUTHIER, F. Z. A. T. M. M. M. “Time domain modeling of powerline impulsive noise at its source”, *Journal of Electromagnetic Analysis and Applications*, v. 3, n. 9, pp. 359–367, Sep. 2011. ISSN: 1942-0749. doi: 10.4236/jemaa.2011.39057.
- [10] DING, X.; MENG, J. P. B. “Blind channel estimation for HomePlug power-line communications: A feasibility?” *IEEE Trans. on Power Delivery*, v. 26, n. 2, pp. 754–763, Apr. 2011.
- [11] SÁNCHEZ, G. L.; FERNÁNDEZ, B. “Blind channel estimation for power-line communications by a PSO-inspired algorithm”. In: *IEEE Latin-American Conference on Communications*, pp. 1–6, Medellin, Colombia, Sep. 2009.
- [12] BELLANTI, L.; RAUGI, M. T. M. “Estimate of cyclostationary PLC channels by a modified self-organizing map”. In: *IEEE International Symposium on Power Line Communications and its Application*, pp. 62–67, Dresden-Germany, Mar. 2009.
- [13] AFIFI, W.; ELKAMCHOUCI, H. “A new adaptive channel estimation for frequency selective time varying fading OFDM channels”. In: *IEEE International Conference on Computer Engineering & Systems*, pp. 573–578, Cairo- Egypt, Dec. 2009.
- [14] MO, R.; OH, S. Z. Y. L. S. “Channel estimation for power line OFDM systems with unknown colored noise exploiting real signal transmission”. In: *IEEE International Symposium on Power Line Communications and its Application*, pp. 149–153, Dresden- Germany, Mar. 2009.
- [15] DELL’AMICO, G.; GUERRINI, E. R. R. “Inter-packet channel estimation in OFDM systems”. In: *IEEE Global Telecommunications Conference*, pp. 1–6, Hawaii- USA, Nov. 2009.
- [16] RIVA, R.; ODoni, M. G. E. B. P. “Improved HomePlug AV channel estimation exploiting sounding procedure”. In: *IEEE International Symposium on Power Line Communications and its Application*, pp. 296–300, Dresden-Germany, Mar. 2009.
- [17] MUSOLINO, A.; RAUGI, M. T. M. “Cyclic short-time varying channel estimation in OFDM power-line communication”, *IEEE Trans. on Power Delivery*, v. 23, n. 1, pp. 157–163, Jan. 2008. ISSN: 0885-8977. doi: 10.1109/TPWRD.2007.910995.

- [18] RAUGI, M.; TUCCI, M. “Power-line communications channel estimation and tracking by a competitive neural network”, *IEEE Trans. on Consumer Electronics*, v. 52, n. 4, pp. 1213–1219, Nov. 2006. ISSN: 0098-3063. doi: 10.1109/TCE.2006.273136.
- [19] DÉGARDIN, V.; LIÉNARD, M. D. P. “Semi-blind estimation method in DMT-based transmission on indoor power line”. In: *IEEE International Symposium on Power-Line Comm. and its Applications*, pp. 195–199, Zaragoza, Espagne, Mar. 2004.
- [20] SCHAFHUBER, D.; MATZ, G. “MMSE and adaptive prediction of time-varying channels for OFDM systems”, *IEEE Trans. Wireless Commun.*, v. 4, n. 2, pp. 593–602, Mar. 2005. ISSN: 1536-1276. doi: 10.1109/TWC.2004.843055.
- [21] RAN, J.; GRUNHEID, R. R. H. B. E. K. R. “Decision-directed channel estimation method for OFDM systems with high velocities”. In: *IEEE Vehicular Technology Conf.*, pp. 2358–2361, Florida, USA, Oct. 2003.
- [22] PICORONE, A. A. M.; AMADO, L. R. R. M. V. “Linear and periodically time-varying PLC channel estimation in the presence of impulsive noise”. In: *IEEE International Symposium on Power Line Communications and its Application*, pp. 255–260, Rio de Janeiro, Brazil, Mar. 2010.
- [23] CHEN, S. AND GUNN, S., HARRIS, C. “Decision feedback equalizer design using support vector machines”, *IEE - Vis. Image Signal Process.*, v. 147, n. 3, pp. 213–219, Jun. 2000. ISSN: 1350-245X. doi: 10.1049/ip-vis:20000360.
- [24] SEBALD, D., BUCLEW, A. “Support vector machine techniques for non-linear equalization”, *IEEE - Trans. Signal Process.*, v. 48, n. 11, pp. 3217–3226, Nov. 2000. ISSN: 1053-587X. doi: 10.1109/78.875477.
- [25] CHEN, S., HANZO, L. “Support vector machine multiuser receiver for DS-CDMA signals in multipath channels”, *IEEE - Trans. Neural Netw.*, v. 12, n. 3, pp. 604–611, May. 2001. ISSN: 1045-9227. doi: 10.1109/72.925563.
- [26] RAHMAN, S., YAMAMOTO, H. “An MC-CDMA signal equalization and detection scheme based on support vector machines”. In: *Proc. 1st Int. Symp. Wireless Communication Systems*, pp. 11–15, Port Louis, Mauritius, Sep. 2004.

- [27] SÁNCHEZ-FERNÁNDEZ, M., PÉREZ-CRUZ, F. “SVM multiregression for nonlinear channel estimation in multiple-input multiple-output systems”, *IEEE - Trans. Signal Process.*, v. 52, n. 8, pp. 2298–2307, Aug. 2004. ISSN: 1053-587X. doi: 10.1109/TSP.2004.831028.
- [28] GARCÍA, M. J. F. G., MARTÍNEZ-RAMÓN, M. “Support vector machine for robust channel estimation in OFDM”, *IEEE - Signal Processing Letters.*, v. 13, n. 7, pp. 397–400, Jul. 2006. ISSN: 1070-9908. doi: 10.1109/LSP.2006.871862.
- [29] IPLC, F. “OPERA”. Website, 2007. <http://www.ipcf.org/company/opera/open-plc-european-research-alliance.html> - visitado em Maio de 2012.
- [30] RAO, S. *IST-POWERNET*. Relatório técnico, European Commission, 2006.
- [31] IPLC, F. “UPA”. Website, 2007. <http://www.ipcf.org/company/upa/universal-powerline-association.html> - visitado em Maio de 2012.
- [32] IPLC, F. “CEPCA”. Website, 2007. <http://www.ipcf.org/company/cepca> - visitado em Maio de 2012.
- [33] SAMY, R. “PLT Summary”. Website, 2012. <http://portal.etsi.org/plt-summary.asp> - visitado em Maio de 2012.
- [34] WEINSTEIN, S.; EBERT, P. “Data transmission by frequency-division multiplexing using the discrete Fourier transform”, *IEEE Trans. on Communications Technology*, v. 19, n. 5, pp. 628–634, Oct. 1971. ISSN: 0018-9332. doi: 10.1109/TCOM.1971.1090705.
- [35] POLLET, T.; PEETERS, M. M. V. L. “Equalization for DMT-based broadband modems”, *IEEE Trans. on Communications Mag.*, v. 38, n. 5, pp. 106–113, May. 2000. ISSN: 0163-6804. doi: 10.1109/35.841833.
- [36] MELLO, R.S.; GRIVET, M. “Modeling, simulation and estimation of PLC channels”. In: *Proc. International Telecommunications Symposium*, pp. 201–206, Ceara, Brazil, Sep. 2006.
- [37] TONELLO, A.M. ; VERSOLATTO, F. “Bottom-up statistical PLC channel modeling—part I: Random topology model and efficient transfer function computation”, *IEEE Transactions on Power Delivery*, Apr 2011.
- [38] ZIMMERMANN, M.; DOSTERT, K. “A multipath model for the powerline channel”, *IEEE Trans. on Communications*, Apr 2002.

- [39] FERREIRA, H. E. A. *Power line communications: Theory and applications for narrowband and broadband communications over power lines*. Wiley, 2010.
- [40] ESMAILIAN, T. *Multi mega bit per second data transmission over in-building power lines*. Tese de d.sc., University of Toronto, Toronto, Canada, 2003.
- [41] HORMIS, R.; BERENQUER, I. W. X. “A simple baseband transmission scheme for power line channels”, *IEEE Journal on Selected Areas in Communications*, v. 24, n. 7, pp. 1351–1363, Jul. 2006. ISSN: 0733-8716. doi: 10.1109/JSAC.2006.874429.
- [42] DE CAMPOS, F.P.V.; RIBEIRO, M. “Performance analysis of clustered-OFDM system with bitloading algorithm for broadband PLC”. In: *IEEE International Symposium on Power Line Communications and Its Applications*, pp. 345–350, Jeju city, Jeju Island, Apr. 2008.
- [43] CORTES, J. A.; DIEZ, L. C. F. J. “Analysis of the indoor broadband power-line noise scenario”, *IEEE Trans. on Electromagnetic Compatibility*, v. 52, n. 4, pp. 849–858, Nov. 2010. ISSN: 0018-9375. doi: 10.1109/TEMPC.2010.2052463.
- [44] MITCHELL, T. *Machine Learning*. 1 ed. EUA, McGraw-Hill, 1997.
- [45] HAYKIN, S. *Neural Networks - A comprehensive Foundation*. 2 ed. EUA, Prentice-Hall, 1999.
- [46] NOBLE, W. S. “Support vector machine applications in computational biology”, *Kernel Methods in computational biology*, 2004.
- [47] PONTIL, M. V. A. “Support vector machines for 3-D object recognition”, *IEEE trans. on Pattern Analysis and Machine Intelligence*, 1998.
- [48] ARAUJO, G. M. ; SILVA JÚNIOR, W. S. . D. S. E. A. B. . G. S. K. “Detecção de landmarks faciais usando SVM”. In: *Proc. XXIX Simpósio Brasileiro de Telecomunicações*, pp. 1–5, Curitiba, Brazil, Sep 2011.
- [49] BASTOS, F. A. C. *Classificador de modulações baseado em máquinas de vetores de suporte*. Tese de d.sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, Abril 2007.
- [50] DIAS, J. T., CAMPOS, F., RIBEIRO, M. V., et al. “Time-varying channel characterization based on SVM under the impulsive noise presence”. In: *Proc. International Symposium on Power Line Communications and Its Applications*, Rio de Janeiro, Brazil, Mar 2010.

- [51] VAPNIK, V., GOLOWICH, S., SMOLA, A. “Support vector method for function approximation, regression estimation, and signal processing”, *Neural Information Processing Systems*, v. 9, 1997.
- [52] SUYKENS, J.; VANDEWALE, J. “Least squares support vector machine”, *Neural Processing Letters*, pp. 293–300, 1999.
- [53] VAPNIK, V. *The nature of statistical learning theory*. 1 ed. Berlim, Springer-Verlag, 1995.
- [54] BURGESS, C. J. C. “A tutorial on support vector machines for pattern recognition”, *Knowledge Discovery and Data Mining*, 1998.
- [55] VAPNIK, V. *Statistical learning theory*. 1 ed. EUA, John Willey & Sons, 1998.
- [56] ABU-MOSTAFA, Y. S. *Learning From Data*. AMLBook, 2012.
- [57] VAPNIK, V. *Estimation of dependences based on empirical data*. 1 ed. Berlim, Springer-Verlag, 1982.
- [58] MARTÍNEZ-RAMÓN, M., CHRISTODOULOU, C. *Support vector machines for antenna array processing and electromagnetics*. 1 ed. EUA, Morgan & Claypool, 2006.
- [59] GUNN, S. “Support vector method for classification and regression”, *Image Speech & Intelligent Systems Group*, 1998.
- [60] VAPNIK, V. N., CORTES, C. “Support vector networks”. In: *Machine learning*, v. 20, pp. 273–297, 1995.
- [61] MANGASARIAN, O., FUNG, G. “Proximal support vector machine classifiers”. In: *Proceedings of the international Conference on Knowledge Discovery and Data Mining*, pp. 64–70, San Francisco, USA, Aug. 2001.
- [62] DAVID, G. L., Y. Y. *Linear and Nonlinear Programming*. Springer, 2008.
- [63] NAVIA-VÁZQUEZ, A.; PÉREZ-CRUZ, F. A.-R. A., FIGUEIRAS-VIDAS, A. R. “Weighted least squares training of support vector classifiers leading to compact and adaptive schemes”, *IEEE Transactions on Neural Networks*, v. 12, n. 5, pp. 1047–1059, Sep 2001.
- [64] DINIZ, P. S. R. *Adaptive Filtering: Algorithms and Practical Implementations*. Kluwer Academic Publisher, 2002.
- [65] OZDEMIR, H., ARSLAN, M. K. “Channel estimation for wireless OFDM systems”, *IEEE Commun. Surv. & Tutor.*, v. 9, n. 2, pp. 18–48, 2007.

- [66] LI, Y.; CIMINI, L. J., SOLLENBERGER, N. R. “Robust channel estimation for OFDM systems with rapid dispersive fading channels”, *IEEE Transactions on Communications*, v. 46, n. 7, pp. 902–915, 1998.
- [67] EDFOR, O.; SANDELL, M., BORJESSON, P. O. “OFDM channel estimation by singular value decomposition”, *IEEE Transactions on Communications*, v. 46, n. 7, pp. 931–939, 1998.
- [68] CORTÉS, J. A; TONELLO, A. M., DÍEZ, L. “Comparative analysis of pilot-based channel estimator for DMT systems over indoor power-line channels”. In: *Proc. of the IEEE International Symposium on Power Line Communications and its Application*, pp. 372–377, Mar 2007.
- [69] GOLUB, G. H., V. L. C. F. *Matrix Computations*. John Hopkins University Press, 1996.
- [70] TAHAT, A. A., N. P. GALATSANOS, N. P. “Channel estimation in a DMT based power-line communication system using sparse bayesian regression”. In: *Proceedings of the 11th WSEAS international conference on Multimedia systems & signal processing*, pp. 104–109, 2011.
- [71] KIM, Y.-H. “Mitigation of effect of impulsive noise for OFDM systems over power line channels”. In: *Proc. of the IEEE International Symposium on Power Line Communications and its Application*, pp. 30–35, Apr. 2008.