

Victor Guimarães Furtado

**Modelagem Não Linear e Controlador  
Ator-Crítico para Manobrabilidade de um USV  
baseado em DLQT-I e Programação Dinâmica  
Heurística Dependente de Ação**

São Luís - MA

26 de Março de 2025

Victor Guimarães Furtado

Universidade Federal do Maranhão – UFMA

Centro de Ciências Exatas e Tecnologias

Departamento de Engenharia Elétrica

Programa de pós-graduação em Engenharia Elétrica

**Modelagem Não Linear e Controlador Ator-Crítico para  
Manobrabilidade de um USV baseado em DLQT-I e  
Programação Dinâmica Heurística Dependente de Ação**

Dissertação de mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal do Maranhão como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica. Área de concentração: Automação e Controle.

São Luís - MA

26 de Março de 2025

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).  
Diretoria Integrada de Bibliotecas/UFMA

Furtado, Victor Guimarães.

Modelagem Não Linear e Controlador Ator-Crítico para Manobrabilidade de um USV baseado em DLQT-I e Programação Dinâmica Heurística Dependente de Ação / Victor Guimarães Furtado. - 2025.

136 f.

Coorientador(a) 1: Yan Ferreira da Silva.

Orientador(a): João Viana da Fonseca Neto.

Dissertação (Mestrado) - Programa de Pós-graduação em Engenharia Elétrica/ccet, Universidade Federal do Maranhão, São Luís, 2025.

1. Derramamentos de Derivados de Petróleo. 2. Manobrabilidade de Usv. 3. Control Ótimo Online. 4. Programação Dinâmica Adaptativa. 5. Aprendizado Por Reforço. I. Neto, João Viana da Fonseca. II. Silva, Yan

Victor Guimarães Furtado

# **Modelagem Não Linear e Controlador Ator-Crítico para Manobrabilidade de um USV baseado em DLQT-I e Programação Dinâmica Heurística Dependente de Ação**

Dissertação de mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal do Maranhão como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica. Área de concentração: Automação e Controle.

Trabalho apresentado em: São Luís - MA, 26 de Março de 2025.

**Prof. Dr. João Viana da Fonseca Neto**  
Orientador

**Prof. Dr. Yan Ferreira da Silva**  
Coorientador

**Prof. Dr. Francisco das Chagas de Souza**  
Universidade Federal do Maranhão  
Examinador Interno

**Prof. Dr. José Pinheiro de Moura**  
Universidade Estadual do Maranhão  
Examinador Externo

São Luís - MA  
26 de Março de 2025

*Este trabalho é dedicado à minha família.*

## AGRADECIMENTOS

Agradeço à Deus pela vida, saúde, força e sabedoria durante esta caminhada.

Agradeço aos meus pais, Irlando José Furtado e Nildemi Fonseca Guimarães Furtado, sei que nada conseguiria sem a educação, orientação e cuidados dos meus pais. Agradeço pelo incentivo, apoio e compreensão ao longo destes anos. Agradeço por sempre me incentivarem aos estudos, por terem me apoiado nos momentos difíceis, e por serem a minha base à qual sempre eu posso contar. Agradeço pelo que tanto fizeram e fazem por mim durante todas as etapas da minha vida.

Ao meu orientador João Viana e ao meu coorientador Yan Ferreira pela orientação e dedicação aplicados neste trabalho. Meu agradecimento pelo acompanhamento, compreensão, auxílio e tempo despendido durante todo curso.

À minha noiva Érica Paloma pelo apoio, incentivo, conselhos que foram essenciais nessa jornada. Agradeço por sempre estar ao meu lado em todos os momentos.

Ao meu irmão Peterson Guimarães, que sempre me incentivou e apoiou em minhas escolhas. Agradeço ao meu sobrinho Noah Guimarães por continuar a me dar força e alegria pra continuar.

Aos meus companheiros do LCP, Alam, Alan, Evandro, Monik, Renato e Tayná pela ajuda, por todos os ensinamentos que compartilhamos e momentos de descontração.

Aos meus amigos Alissoney Diniz, Denis Vinicius e Raul Barbosa pelo companheirismo e amizade durante todos esses anos.

A todos os professores e colaboradores do Programa de Pós-Graduação em Engenharia de Eletricidade e da Universidade Federal do Maranhão, que contribuíram com o desenvolvimento deste trabalho. À Agência Nacional do Petróleo, Gás Natural e Biocombustíveis em nome do Programa de Formação de Recursos Humanos 54.1 e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pelo suporte financeiro.

# Resumo

Os derramamentos de derivados de petróleo provocam degradação ambiental, problemas socioeconômicos e graves danos à saúde humana. Em decorrência das dificuldades do monitoramento dos eventos em grandes áreas, a utilização de veículos aquáticos de superfície não tripulados (*Unmanned Surface Vehicles* - USVs) tornou-se imprescindível para a tarefa. A aplicação desses dispositivos demandam a sua instrumentalização com sensores para análise da qualidade da água, além de um sistema de orientação, navegação e controle (*Guidance, Navigation and Control* - GNC). Especificamente para viabilizar a demanda de manobrabilidade, apresenta-se o desenvolvimento da proposta de um sistema de controle *online* baseado em Programação Dinâmica Adaptativa (PDA) e Aprendizado por Reforço (AR). O desenvolvimento de um sistema de controle para manobrabilidade de USVs, apresenta empecilhos em sua implementação em decorrência das dificuldades da identificação do sistema para a modelagem matemática do processo por se tratar de um modelo não linear complexo, acoplado e com incertezas. As perturbações externas e possíveis variações paramétricas aumentam os entraves para sua elaboração. Nesse contexto, esta dissertação apresenta o projeto de um controlador ótimo adaptativo baseado em dados para controle das velocidades de USV completamente atuado, como controle de manobrabilidade. A metodologia desenvolvida para o projeto do controlador ótimo *online* agrega PDA e AR na técnica Programação Dinâmica Heurística Dependente de Ação (*Action Dependent Heuristic Dynamic Programming* - ADHDP), visando o cálculo da solução *online* implícita da equação de *Hamilton-Jacobi-Bellman* (HJB) na forma da Equação Algébrica de *Riccati* Discreta (*Discrete Algebraic Riccati Equation* - DARE). O sistema utiliza o Aprendizado por Reforço por meio da estrutura ator-crítico, com uma abordagem de controle ótimo para cálculo da lei de controle ótima baseado nos dados da ação de controle e estados do processo. O controlador desenvolvido é um Rastreador Linear Quadrático Discreto com ação integral (DLQT-I) denominado ADHDP-DLQT-I e os seus resultados são avaliados em simulador do modelo não linear, resolvido pelo método *Runge-Kutta*. O desempenho do sistema ADHDP-DLQT-I é comparado com um controlador PID adaptativo sintonizado pelo método dos Mínimos Quadrados (*Least Mean Square* - LMS).

**Palavras-chave:** Derramamentos de derivados petróleo; Manobrabilidade de USV; Controle ótimo *online*; PDA; Aprendizado por Reforço; ADHDP-DLQT-I.

# Abstract

Oil spills cause environmental degradation, socio-economic problems, and severe harm to human health. Due to the challenges of monitoring such events over large areas, the use of Unmanned Surface Vehicles (USVs) has become essential for this task. The application of these devices requires equipping them with sensors for water quality analysis, as well as a Guidance, Navigation, and Control (GNC) system. Specifically, to enable maneuverability, this work presents the development of an online control system proposal based on Adaptive Dynamic Programming (ADP) and Reinforcement Learning (RL). The development of a control system for USV maneuverability presents implementation challenges due to the difficulties in system identification for mathematical modeling, as it involves a complex, nonlinear, coupled model with uncertainties. External disturbances and possible parametric variations further increase the obstacles to its design. In this context, this dissertation presents the design of an adaptive optimal controller based on data for controlling the speeds of a fully actuated USV, ensuring maneuverability. The methodology developed for designing the online optimal controller integrates ADP and RL within the Action-Dependent Heuristic Dynamic Programming (ADHDP) technique, aiming to compute the implicit online solution of the Hamilton-Jacobi-Bellman (HJB) equation in the form of the Discrete Algebraic Riccati Equation (DARE). The system employs Reinforcement Learning using an actor-critic structure with an optimal control approach to compute the optimal control law based on control action data and process states. The developed controller is a Discrete Linear Quadratic Tracker with integral action (DLQT-I), referred to as ADHDP-DLQT-I, and its results are evaluated using a nonlinear model simulator solved via the Runge-Kutta method. The performance of the ADHDP-DLQT-I system is compared with an adaptive PID controller tuned using the Least Mean Square (LMS) method.

**Keywords:** Oil Spills; USV maneuverability; Online optimal control; ADP; Reinforcement Learning; ADHDP-DLQT-I.

# Lista de ilustrações

Figura 2.1 – Diagrama de blocos do sistema de controle ADHDP-DLQT-I. . . . .	29
Figura 3.1 – Veículos não-autônomos (1), autônomos (2) e não tripulados de superfície essenciais para monitoramento de derramamentos de petróleo no mar, detectam e mapeiam manchas de óleo na superfície. Os tipos UUVs (3) e AUVs (4) analisam a dispersão subaquática e inspecionam infraestruturas críticas. Os ROVs (5) realizam inspeções detalhadas e auxiliam na contenção e limpeza. . . . .	31
Figura 3.2 – Configurações de cascos de USVs: casco único (1) e catamarã (2). . . . .	32
Figura 3.3 – Estrutura generalizada do sistema GNC de um USV (Fonte: Adaptado de (LIU et al., 2016)). . . . .	33
Figura 3.4 – Variáveis de movimento no sistema referencial de corpo fixo para USV. . . . .	34
Figura 3.5 – Sistemas de Coordenadas de Referência Centrados na Terra e Sistemas de Coordenadas de referência geográfica (Fonte: (FOSSEN, 2011)). . . . .	35
Figura 3.6 – Eixos e pontos do sistema de coordenadas de corpo fixo (Fonte: (FOSSEN, 2011)) . . . . .	36
Figura 3.7 – Representação do movimento horizontal do USV (3DOF). . . . .	41
Figura 3.8 – Estrutura ator-crítico do Aprendizado por reforço. Essa estrutura fornece métodos para aprender soluções de controle ótimas on-line com base em dados medidos ao longo das trajetórias do sistema. (Fonte: Adaptado de (LEWIS; VRABIE, 2009)) . . . . .	52
Figura 3.9 – Solução e inclinações do método Runge-Kutta de quarta ordem para a equação diferencial $y' = -2xy$ , com passo $h = 0,5$ . . . . .	57
Figura 3.10 – Diagrama de blocos do sistema de controle PID-LMS. . . . .	59
Figura 4.1 – Diagrama de blocos do sistema de controle ADHDP-DLQT-I baseado no modelo ator-crítico com dois atores. . . . .	64
Figura 4.2 – Diagrama de blocos do sistema MIMO em malha fechada com controlador PID-LMS. . . . .	71
Figura 4.3 – Diagrama de blocos da matriz de função de transferência do modelo matemático completo do <i>Cybership I</i> . . . . .	72
Figura 4.4 – Diagrama de bloco de um controlador com desacoplamento do processo. . . . .	73
Figura 4.5 – Diagrama de blocos de um sistema TITO com desacoplamento do processo. . . . .	73
Figura 4.6 – Diagrama de bloco da matriz de desacoplamento. . . . .	74
Figura 4.7 – Diagrama de bloco do modelo completo do <i>Cybership I</i> em malha fechada com controlador. . . . .	77
Figura 4.8 – Diagrama de bloco do sistema desacoplado em malha fechada com controlador PID-LMS. . . . .	77

Figura 5.1 – Resposta das velocidades $u$ -surge e $v$ -sway do USV para o controlador ADHDP-DLQT-I . . . . .	79
Figura 5.2 – Resposta da velocidade $r$ -yaw do USV para o controlador ADHDP-DLQT-I. . . . .	80
Figura 5.3 – Ações de controle $\tau_1$ e $\tau_2$ do controlador ADHDP-DLQT-I. . . . .	81
Figura 5.4 – Ação de controle $\tau_3$ do controlador ADHDP-DLQT-I. . . . .	82
Figura 5.5 – Convergência da matriz de ganhos de realimentação de estados $K_1$ . . . . .	83
Figura 5.6 – Convergência da matriz de ganhos de realimentação de saída $K_2$ . . . . .	84
Figura 5.7 – Resposta ao degrau unitário do controlador PID-LMS da velocidade $u$ do USV. . . . .	85
Figura 5.8 – Erro do sistema com controlador PID-LMS para velocidade $u$ . . . . .	86
Figura 5.9 – Sinais de entrada do combinador linear do treinamento LMS. . . . .	87
Figura 5.10–Convergência dos ganhos do controlador PID-LMS para a velocidade $u$ . . . . .	88
Figura 5.11–Sinal de controle $U_c$ para a velocidade $u$ . . . . .	89
Figura 5.12–Resposta ao degrau unitário do sistema com controlador PID-LMS para velocidade $v$ . . . . .	90
Figura 5.13–Erro do sistema com controlador PID-LMS para velocidade $v$ . . . . .	91
Figura 5.14–Sinais de entrada do combinador linear de treinamento LMS para velocidade $v$ . . . . .	92
Figura 5.15–Convergência dos ganhos do controlador PID-LMS para a velocidade $v$ . . . . .	93
Figura 5.16–Sinal de controle $U_c$ para a velocidade $v$ . . . . .	94
Figura 5.17–Resposta ao degrau unitário do sistema com controlador PID-LMS para velocidade $r$ . . . . .	95
Figura 5.18–Erro do sistema com controlador PID-LMS para velocidade $r$ . . . . .	96
Figura 5.19–Sinais de entrada do combinador linear de treinamento LMS para velocidade $r$ . . . . .	97
Figura 5.20–Convergência dos ganhos do controlador PID-LMS para a velocidade $r$ . . . . .	98
Figura 5.21–Sinal de controle $U_c$ para a velocidade $r$ . . . . .	99
Figura 5.22–Avaliação das velocidades $u$ e $v$ para o modelo não linear. . . . .	100
Figura 5.23–Avaliação das velocidades $u$ e $v$ para o modelo não linear. . . . .	100
Figura 5.24–Respostas do controlador ADHDP-DLQT-I para as velocidades $u$ , $v$ e $r$ utilizando o modelo completo do Cybership I. . . . .	101
Figura 5.25–Ação de controle do controlador ADHDP-DLQT-I para as velocidades $u$ , $v$ e $r$ utilizando o modelo completo do Cybership I. . . . .	102
Figura 5.26–Convergência dos ganhos da matriz $K_1$ para o modelo completo do <i>Cybership I</i> . . . . .	103
Figura 5.27–Convergência dos ganhos da matriz $K_2$ para o modelo completo do <i>Cybership I</i> . . . . .	103

Figura 5.28–Resposta ao degrau unitário do controlador PID-LMS da velocidade $u$ do USV. . . . .	105
Figura 5.29–Erro do sistema com controlador PID-LMS para velocidade $u$ . . . . .	106
Figura 5.30–Sinais de entrada do combinador linear do treinamento LMS. . . . .	107
Figura 5.31–Convergência dos ganhos do controlador PID-LMS para a velocidade $u$ . . . . .	108
Figura 5.32–Sinal de controle $U_c$ para a velocidade $u$ . . . . .	109
Figura 5.33–Resposta ao degrau unitário do sistema com controlador PID-LMS para velocidade $v$ . . . . .	110
Figura 5.34–Erro do sistema com controlador PID-LMS para velocidade $v$ . . . . .	111
Figura 5.35–Sinais de entrada do combinador linear de treinamento LMS para velocidade $v$ . . . . .	112
Figura 5.36–Convergência dos ganhos do controlador PID-LMS para a velocidade $v$ . . . . .	113
Figura 5.37–Sinal de controle $U_c$ para a velocidade $v$ . . . . .	114
Figura 5.38–Resposta ao degrau unitário do sistema com controlador PID-LMS para velocidade $v$ . . . . .	115
Figura 5.39–Erro do sistema com controlador PID-LMS para velocidade $v$ . . . . .	116
Figura 5.40–Sinais de entrada do combinador linear de treinamento LMS para velocidade $r$ . . . . .	117
Figura 5.41–Convergência dos ganhos do controlador PID-LMS para a velocidade $r$ . . . . .	118
Figura 5.42–Sinal de controle $U_c$ para a velocidade $r$ . . . . .	119
Figura 5.43–Resposta comparativa dos controladores para a velocidade $u$ . . . . .	120
Figura 5.44–Resposta comparativa dos controladores para a velocidade $v$ . . . . .	121
Figura 5.45–Resposta comparativa dos controladores para a velocidade $r$ . . . . .	122
Figura 5.46–Resposta comparativa das ações de controle para as velocidades $u$ , $v$ e $r$ . . . . .	123

# Lista de tabelas

Tabela 3.1 – Notação SNAME de Forças, momentos, velocidades e posições para embarcações. . . . .	34
Tabela 4.1 – Parâmetros do Cybersip I . . . . .	61

## LISTA DE ABREVIATURAS E SIGLAS

ADHDP	<i>Action Dependent Heuristic Dynamic Programming</i> (Programação Dinâmica Heurística Dependente de Ação)
AP	Algoritmo Principal
AR	Aprendizado por Reforço
ARE	<i>Algebraic Riccati Equation</i> (Equação Algébrica de Riccati)
ASV	<i>Autonomous Surface Vehicles</i> (Veículo Aquático de Superfície Autônomo)
CB	Centro de Flutuabilidade
CF	Centro de Flutuação
CG	Centro de Gravidade
CO	Centro de origem
DARE	<i>Discrete Algebraic Riccati Equation</i> (Equação Algébrica de Riccati Discreta)
DLQR	<i>Discrete Linear Quadratic Regulator</i> (Regulador Linear Quadrático Discreto)
DLQT-I	<i>Discrete Linear Quadratic Tracker with Integral action</i> (Rastreador Linear Quadrático Discreto com ação Integral)
DOF	<i>Degrees Of Freedom</i> (Graus de Liberdade)

ECEF	<i>Earth-centered Earth-Fixed</i> (Coordenadas de Referência Centrado na Terra e Fixo à Terra)
ECI	<i>Earth-centered Earth-Fixed</i> (Coordenada Inercial Centrada na Terra)
GNC	<i>Guidance, Navigation and Control</i> (Orientação, Navegação e Controle)
HDP	<i>Heuristic Dynamic Programming</i> (Programação Dinâmica Heurística)
HFV	<i>Hypersonic Flight Vehicle</i> (Veículo de Voo Hipersônico)
HJB	<i>Hamilton-Jacobi-Bellman</i>
LMS	<i>Least Mean Square</i> (Mínimos Quadrados Médios)
LQG	Linear Quadrático Gaussiano
LQR	<i>Linear Quadratic Regulator</i> (Regulador Linear Quadrático)
MDP	<i>Markov Decision Process</i> (Processo de Decisão de Markov)
MIMO	<i>Multiple Input Multiple Output</i> (Entrada Múltipla Saída Múltipla)
MPC	<i>Model Predictive Control</i> (Controlador Preditivo Baseado em Modelo)
MRAC	<i>Model Reference Adaptive Control</i> (Controle Adaptativo por Modelo de Referência)
NED	<i>North-East-Down</i> (Norte-Leste-Baixo)
NTNU	<i>Norwegian University of Science and Technology</i> (Universidade Norueguesa de Ciência e Tecnologia)
PDA	Posicionamento Dinâmico
PDA	Programação Dinâmica Adaptativa
PID	Proporcional Integral Derivativo
RK4	Runge-Kutta de quarta ordem
RLOTC	<i>Reinforcement Learning Optimal Tracking Control</i> (Controlador Ótimo de Rastreamento via Aprendizado por Reforço)
RLS	<i>Recursive Least Squares</i> (Mínimos Quadrados Recursivos)
ROV	<i>Remotely Operated Vehicles</i> (Veículos Operados Remotamente)

RSSF	Rede de Sensores Sem Fio
SNAME	<i>Society of Naval Architects and Marine Engineers</i> (Sociedade de Arquitetos Navais e Engenheiros Marítimos)
SISO	<i>Single Input Single Output</i> (Única Entrada Única Saída)
TITO	<i>Two Inputs Two Outputs</i> (Duas Entradas Duas Saídas)
UKF	<i>Unscented Kalman Filter</i>
USV	<i>Unmanned Surface Vehicles</i> (Veículo Aquático de Superfície Não-tripulado)
UUV	<i>Unmanned Underwater Vehicles</i> (Veículo Subaquático Não-tripulado)

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>17</b>
<b>1.1</b>	<b>Justificativa</b>	<b>20</b>
<b>1.2</b>	<b>Objetivos</b>	<b>20</b>
1.2.1	Objetivo Geral	21
1.2.2	Objetivos Específicos	21
<b>1.3</b>	<b>Organização do trabalho</b>	<b>21</b>
<b>2</b>	<b>ESTADO DA ARTE E PROPOSTA</b>	<b>23</b>
<b>2.1</b>	<b>Estado da Arte</b>	<b>23</b>
<b>2.2</b>	<b>Proposta para Controle de Manobrabilidade</b>	<b>26</b>
2.2.1	Descrição do Problema	27
2.2.2	Formulação do Problema	27
2.2.3	Solução Proposta	28
<b>3</b>	<b>MODELAGEM MATEMÁTICA, CONTROLE ÓTIMO <i>ONLINE</i> E MÉTODO DE RUNGE-KUTTA</b>	<b>30</b>
<b>3.1</b>	<b>Veículos Aquáticos de Superfície Não-tripulados</b>	<b>30</b>
<b>3.2</b>	<b>Modelo Matemático do USV</b>	<b>33</b>
3.2.1	Sistemas de Coordenadas de Referência	34
3.2.2	Sistemas de Coordenadas de Referência Centrados na Terra	35
3.2.3	Sistemas de Coordenadas de Referência Geográfica	35
3.2.4	Transformações entre BODY e NED	37
3.2.5	Dinâmica de embarcações	37
3.2.5.1	Matriz de Inércia	38
3.2.5.2	Matriz Coriolis-Centrípeta	38
3.2.5.3	Matriz de amortecimento	39
3.2.5.4	Termos adicionais da equação	40
3.2.6	Modelo para Três Graus de Liberdade (3-DOF)	40
<b>3.3</b>	<b>Controle Ótimo</b>	<b>42</b>
3.3.1	O Problema de Controle Ótimo	43
3.3.2	O Problema do Regulador Linear Quadrático Discreto	44
3.3.3	O problema do Rastreador Linear Quadrático Discreto	46
<b>3.4</b>	<b>Formulação do Rastreador Linear Quadrático Discreto com Ação Integral</b>	<b>48</b>
<b>3.5</b>	<b>Programação Dinâmica Adaptativa</b>	<b>51</b>
<b>3.6</b>	<b>Formulação da Sintonia ADHDP para o caso LQR Discreto</b>	<b>53</b>

3.6.1	<i>Q-learning</i> para DLQR . . . . .	55
<b>3.7</b>	<b>Método de Runge-Kutta . . . . .</b>	<b>56</b>
3.7.1	Método de Runge-Kutta de quarta ordem . . . . .	57
<b>3.8</b>	<b>Controlador Adaptativo PID-LMS . . . . .</b>	<b>58</b>
<b>4</b>	<b>PROJETO DO CONTROLADOR ATOR-CRÍTICO PARA MA- NOBRABILIDADE . . . . .</b>	<b>60</b>
<b>4.1</b>	<b>Modelo Matemático do <i>CyberShip I</i> . . . . .</b>	<b>60</b>
4.1.1	A Representação do Modelo Linear em Espaço de Estados . . . . .	61
<b>4.2</b>	<b>Controle ADHDP-DLQT com ação integral . . . . .</b>	<b>62</b>
4.2.1	Algoritmo ADHDP-DLQT-I . . . . .	64
<b>4.3</b>	<b>Simulador Não Linear . . . . .</b>	<b>69</b>
<b>4.4</b>	<b>Projeto do Controlador PID Adaptativo LMS Desacoplado . . . . .</b>	<b>70</b>
<b>4.5</b>	<b>Projeto do Controlador PID Adaptativo LMS com Desacopla- mento . . . . .</b>	<b>71</b>
<b>5</b>	<b>RESULTADOS E ANÁLISE DO CONTROLADOR DE MANO- BRABILIDADE . . . . .</b>	<b>78</b>
<b>5.1</b>	<b>Resultados do Controlador ADHDP-DLQT-I no Modelo <math>xg_0</math> . . . . .</b>	<b>79</b>
<b>5.2</b>	<b>Resultados do Controlador PID-LMS para o modelo <math>xg_0</math> . . . . .</b>	<b>85</b>
<b>5.3</b>	<b>Resultados do Modelo Não Linear . . . . .</b>	<b>99</b>
<b>5.4</b>	<b>Resultados para o Modelo Completo com Controlador ADHDP- DLQTI . . . . .</b>	<b>101</b>
<b>5.5</b>	<b>Resultados para o Modelo Completo com Controlador PID-LMS</b>	<b>104</b>
<b>5.6</b>	<b>Comparação entre os métodos ADHDP-DLQT-I e PID-LMS . . . . .</b>	<b>119</b>
<b>6</b>	<b>CONCLUSÃO . . . . .</b>	<b>124</b>
<b>6.1</b>	<b>Trabalhos Futuros . . . . .</b>	<b>124</b>
<b>6.2</b>	<b>Publicações . . . . .</b>	<b>125</b>
6.2.1	Principal Publicação . . . . .	125
6.2.2	Outras Publicações Relacionadas ao Tema . . . . .	125
	<b>REFERÊNCIAS . . . . .</b>	<b>126</b>
	<b>APÊNDICES . . . . .</b>	<b>132</b>
	<b>A – MATRIZES DO DESACOPLAMENTO . . . . .</b>	<b>133</b>

# CAPÍTULO 1

## INTRODUÇÃO

A exploração *offshore* de petróleo envolve atividades com alto índice de potencial poluidor do mar decorrente de eventuais derramamentos de óleo e seus derivados, causados por vazamentos em dutos e plataformas, além de acidentes durante a produção e transporte (EUZEBIO; RANGEL; MARQUES, 2019). Esses derramamentos provocam sérios danos à saúde humana, problemas socioeconômicos e degradação ambiental. O meio ambiente é gravemente afetado por danos à qualidade da água, do solo, do ar e à vida dos seres vivos dependentes desses fatores de forma direta ou indireta (SILVA et al., 2021).

O monitoramento das águas em regiões propensas a derramamentos é essencial para o dimensionamento da abrangência e possíveis impactos desses acidentes em primeira instância, a fim de auxiliar em medidas mitigadoras para redução das consequências. Áreas de derramamentos são locais insalubres e por conta dos riscos associados, é mais adequado que o monitoramento seja realizado de forma remota, como por exemplo, utilizando uma Rede de Sensores Sem Fio (RSSF) híbrida com nós estáticos e dinâmicos (SELMIC; PHOHA; SERWADDA, 2016) (SILVA, 2024). Entre os tipos de nós dinâmicos para monitoramento marítimo, destacam-se os veículos aquáticos de superfície não tripulados do tipo USV (LIU et al., 2016). Esses veículos têm aplicações em diversas áreas como pesquisas científicas (ZHOU et al., 2024a), missões ambientais (MENDOZA-CHOK et al., 2022), exploração de recursos oceânicos (AHMED et al., 2024), aplicações militares (ALIFERIS, 2021), transporte de cargas (SHAO et al., 2019), busca e salvamento (LI et al., 2023), entre outras.

Os USV's podem ser historicamente referenciados, inicialmente, pela patente obtida por Nikola Tesla em 1898 para controle remoto de barcos ou veículos por rádio (NIKOLA, 1898). Entretanto, os USV's como definidos atualmente foram desenvolvidos e operados durante a Segunda Guerra Mundial, como barcos drones para limpar minas aquáticas e obstáculos e para treinamentos de prática de tiro (BREIVIK, 2010). A marinha dos

Estados Unidos caracteriza esses veículos a partir de duas definições (NAVY, 2007):

- **Não-tripulado:** Capaz de operação não tripulada. Pode ser tripulado para uso duplo ou teste e avaliação. Possui vários graus de autonomia.
- **Veículo de Superfície:** Veículo que desloca água em repouso. Opera com contato quase contínuo com a superfície da água. A interface do veículo com a superfície é um fator essencial.

Em geral, são embarcações que realizam tarefas em ambientes aquáticos variados, sem tripulação e apresentam dinâmica altamente não linear (BREIVIK, 2010). Os sistemas de orientação, navegação e controle, chamado de sistema GNC, é um componente essencial para a operação dos USV's. Cada unidade desempenha uma função específica para o deslocamento da embarcação. Em resumo, a orientação está vinculada à trajetória, a navegação envolve as condições de deslocamento, como sensoriamento da velocidade e obstáculos, e o sistema de controle é responsável pela manobrabilidade, controlando as velocidades da embarcação, geradas pela força produzida pelos propulsores (LIU et al., 2016).

As técnicas para o desenvolvimento de um sistema de controle de USV variam conforme os objetivos do sistema e compreendem uma ampla diversidade de abordagens. Alguns tipos de controladores encontrados na literatura são: controlador adaptativo com filtro de Kalman UKF (*Unscented Kalman Filter*) (PENG; HAN; HUANG, 2009), adaptativo por modelo de referência (*Model Reference Adaptive Control* - MRAC) (KRAGELUND et al., 2013), *backstepping* (DO; JIANG; PAN, 2002)(GADRE et al., 2012)(SONNENBURG, 2012)(SONNENBURG; WOOLSEY, 2013), Regulador Linear Quadrático (*Linear Quadratic Regulator* - LQR) (LEFEBER; PETTERSEN; NIJMEIJER, 2003), Linear Quadrático Gaussiano (LQG) (ANNAMALAI; MOTWANI, 2013), controlador preditivo baseado em modelo (*Model Predictive Control* - MPC) (ANNAMALAI et al., 2014), Proporcional Integral Derivativo (PID) (WONDERGEM et al., 2010).

O desenvolvimento do sistema de controle de um USV requer considerações da complexidade da dinâmica do processo, como variações das condições operacionais e perturbações externas do ambiente (correntes oceânicas e ventos). Essas características do processo comprometem a operação e o desempenho de controladores clássicos baseados em ganhos fixos tornando-os inadequados. Nesse sentido, a teoria de controle adaptativo possui a capacidade de compensar variações paramétricas da planta, incertezas de modelagem e distúrbios do processo, preservando o desempenho do sistema com projetos de controladores *online*, não baseados em apenas um ponto de operação, mas em diferentes pontos (MOURA; REGO; NETO, 2019).

O controle adaptativo é capaz de modificar seu projeto em resposta às mudanças

na dinâmica do processo e nas perturbações, através de parâmetros reguláveis e um mecanismo para esse ajuste (ÅSTRÖM; WITTENMARK, 2008). Entretanto, controladores adaptativos, geralmente, não são projetados com a característica de serem ótimos. Algumas abordagens que solucionam esse dilema agregam conceitos de controle ótimo e controle adaptativo baseado nos paradigmas de AR e PDA.

A PDA baseia-se nas formulações de *Bellman*, que contempla de sistemas não-lineares a sistemas lineares baseados em funções de custo quadrática, para uma implementação *online* utilizando um mecanismo "para frente no tempo" em busca de uma política de controle ótima, adaptando suas estruturas paramétricas de adaptativo-crítico, para encontrar a solução da Equação HJB a partir de sucessivas aproximações. A rede neural de ação calcula as ações de controle e a rede neural crítica aproxima a função valor. Este método avalia o efeito do controle para o desempenho futuro, e fornece orientações para aprimorar a lei de controle. As abordagens que combinam PDA e AR solucionam problemas de controle ótimo de forma *online* com sistemas de controle que não dependem do modelo do processo de forma explícita, são baseados apenas em dados (FERREIRA et al., 2016).

Um dos métodos de PDA introduzido por Werbos (WERBOS, 1992), utiliza o aprendizado por reforço baseado no aprendizado da função de valor escalar denominado Programação Dinâmica Heurística (*Heuristic Dynamic Programming* - HDP). Além disso, apresentou uma variação denominada Programação Dinâmica Heurística Dependente de Ação (ADHDP), equivalente ao aprendizado Q (*Q-learning*) para o processo de decisão de Markov de estado discreto, esse método aprende a função Q e permite que se execute o aprendizado por reforço sem nenhum conhecimento da dinâmica do sistema (LEWIS; VRABIE, 2009).

A ADHDP é uma variação dos métodos críticos adaptativos mais comuns, constituindo uma estrutura ator-crítico. O ator define uma lei de controle (política de controle) parametrizada e o crítico fornece uma representação parametrizada da função valor. A função avalia o impacto do controle no desempenho futuro e fornece orientações para aprimorar a lei de controle (RÊGO, 2014). Esse método é capaz de aprender a solução da Equação HJB a partir de aproximações e parametrizações da Equação de Bellman, sem a identificação explícita do sistema, baseado somente em dados de entrada (esforço de controle) e saída (estados), que possibilita a aplicação de um controlador ótimo em um sistema de dinâmica desconhecida e não-linear. O método fornece a solução da Equação Algébrica de Riccati (ARE) de forma *online*, resultando em um controlador ótimo adaptativo *online* capaz de superar variações paramétricas, mudanças na dinâmica do sistema e erros de modelagem (FURTADO; NETO, 2024).

O método apresentado nesse trabalho, utiliza a ADHDP para a implementação de um rastreador ótimo *online* do tipo DLQT (*Discrete Linear Quadratic Tracker*) com ação integral (DLQT-I). O método apresenta a ADHDP associada ao problema do LQR para

sistemas discretos por meio das suas parametrizações utilizando o modelo linear do USV Cybership I, desenvolvido pela NTNU, e reformula a implementação para o DLQT-I, o que resulta no controlador ator-crítico denominado ADHDP-DLQT-I.

## 1.1 Justificativa

Derramamentos de petróleo e seus derivados no mar representam uma grande ameaça ao ecossistema marinho e à saúde humana (EUZEBIO; RANGEL; MARQUES, 2019). O problema se agrava quando esses óleos atingem a linha costeira, onde a produtividade biológica é maior e o óleo encalhado pode permanecer por longos períodos. Esses episódios estão associados ao transporte, ruptura de tanques, exploração *offshore* e vazamento de oleodutos submarinos (ASIF et al., 2022). Uma resposta rápida é crítica em casos de derramamentos, pois a janela de oportunidade para gerir a situação é bastante limitada. A rápida dispersão do óleo torna o tempo crucial para uma contenção eficiente e outros tipos de medidas (ELMAKIS et al., 2024). Dada a gravidade dos impactos e os riscos desses eventos, é essencial monitorar as regiões propensas a derramamentos de forma remota a fim de auxiliar a mitigação dos impactos reduzindo o tempo de exposição do ecossistema marinho (ZHANG; YANG; REN, 2024) (SILVA, 2024).

Os USV's oferecem uma solução prática para monitoramento dinâmico de derramamentos pois sua alta carga útil e capacidade substancial da bateria permitem operações contínuas (ELMAKIS et al., 2024). O desenvolvimento do sistema de controle desses veículos é essencial para a expansão e o aprimoramento da tecnologia. Sua aplicação no monitoramento amplia as possibilidades de prevenção e redução dos impactos, proporcionando uma resposta rápida e eficaz, principalmente devido à maior área de cobertura e à capacidade de alcançar áreas de difícil acesso, em comparação aos nós estáticos.

A aplicação de um controlador ADHDP-DLQT-I para o desenvolvimento de um sistema de controle de USV, é uma área pouco explorada que pode agregar significativamente devido às suas características ótimo *online* e baseado em dados visando superar as complexidades do modelo dinâmico da planta, suas incertezas e as perturbações externas associadas.

## 1.2 Objetivos

Nesta seção são descritos os objetivos gerais e específicos necessários para a conclusão da proposta do trabalho.

### 1.2.1 Objetivo Geral

O objetivo geral é desenvolver um controlador ótimo *online* do tipo ator-crítico baseado em dados ADHDP-DLQT-I, para manobrabilidade de um USV por meio do controle das suas velocidades e avaliar seu desempenho em um simulador não-linear.

### 1.2.2 Objetivos Específicos

Para atingir o objetivo principal desta pesquisa, é necessário atingir os seguintes objetivos específicos:

- Levantar o modelo matemático linear e não linear de um USV;
- Desenvolver o projeto de um controlador adaptativo do tipo PID-LMS para a manobrabilidade do USV;
- Desenvolver uma metodologia de controlador ator-crítico ótimo adaptativo do tipo ADHDP-DLQT-I para rastreamento de velocidade para o USV, baseado em Aprendizado por Reforço e Programação Dinâmica Adaptativa;
- Desenvolver um simulador do modelo não linear do USV, com solução baseada no método de *Runge-Kutta*;
- Avaliar o desempenho do controlador desenvolvido comparando com o controlador PID-LMS e com o simulador não-linear;

## 1.3 Organização do trabalho

Nos capítulos remanescentes desta dissertação são apresentados o estado da arte, a proposta e o arcabouço teórico que descreve os conceitos e formulações das principais abordagens envolvidas na solução da proposta. Ademais, são apresentados capítulos com a metodologia proposta, os principais resultados computacionais baseados e comentários finais que abordam a conclusão e trabalhos futuros. Nos próximos parágrafos são apresentadas as descrições dos capítulos.

No Capítulo 2, são apresentados o estado da arte e a proposta para o controle de manobrabilidade do USV. Este capítulo aborda a descrição, formulação e solução do problema.

No Capítulo 3, são apresentados os principais paradigmas para o desenvolvimento da pesquisa. Em termos, este capítulo apresenta o referencial teórico, abordando conceitos sobre USVs, dinâmica e modelagem não linear do processo, controle ótimo do tipo regulador e rastreador com ganhos em regime permanente, programação dinâmica adaptativa, sintonia de controladores, ADHDP e método *Runge-Kutta*.

A metodologia para Projeto do Controlador Ator-Crítico que é orientado para USV é apresentada no Capítulo 4. A formulação do controle ADHDP-DLQT-I e seu algoritmo são apresentados neste capítulo. Também, apresenta-se a dinâmica do sistema com seu modelo matemático com os parâmetros do *CyberShip* I e sua descrição em espaço de estados.

Os resultados e análises das estratégias de controle ator-critico de manobrabilidade do USV são apresentados no Capítulo 5, que também inclui os algoritmos e resultados para diversas situações do desempenho dos controladores que são provenientes de um modelo não-linear.

Os comentários conclusivos sobre a metodologia abordada para desenvolvimento do controlador ADHDP-LQT-I e sugestões de trabalhos futuros são apresentados no Capítulo 6. Também, apresenta-se um apêndice com as funções de transferência das matrizes de desacoplamento.

## CAPÍTULO 2

# ESTADO DA ARTE E PROPOSTA

O Estado arte abrange os tópicos de controle ótimo adaptativo baseado em PDA, especificamente a abordagem ator-critico no contexto da ADHDP, DLQT-I e aplicações do controle ótimo para para manobrabilidade de um USV. A proposta compreende os principais paradigmas da problemática por meio da sua descrição, formulação e solução.

### 2.1 Estado da Arte

A PDA tem sido amplamente explorada para otimização e controle em sistemas dinâmicos. Modelos tradicionais, como a HDP, utilizam uma rede neural crítica para estimar a função de valor com base em sinais de reforço discretos do ambiente. No entanto, esses métodos frequentemente dependem de conhecimento prévio para definir um sinal de reforço eficiente, limitando sua adaptabilidade. Por sua vez, a ADHDP utiliza apenas sinais de entrada e os estados da planta para calcular o ganho ótimo baseado em dados (LEWIS; VRABIE; SYRMOS, 2012).

Em Ni, He e Wen (2013), é apresentada uma abordagem de PDA baseada em uma ADHDP com rede crítica dupla, sendo uma rede de referência que fornece de forma adaptativa um sinal de reforço. uma representação de meta interna para ajudar no aprendizado e otimização dos sistemas. Diferente dos métodos tradicionais que dependem de um sinal de reforço fixo do ambiente, a rede de referência gera um sinal de meta interna ajustável, permitindo uma adaptação dinâmica ao estado e à ação do sistema. Os resultados demonstram que a rede crítica dupla supera o método HDP típico em termos de erro acumulado e custo computacional. O estudo também sugere possibilidades futuras para a extensão para múltiplas redes de referência e substituição do método de retropropagação para algoritmos mais avançados para otimização da atualização dos pesos da rede neural.

Em Mu et al. (2017), é proposto uma abordagem de controle suplementar orientada

por dados para rastreamento da velocidade e altitude de um veículo hipersônico de respiração atmosférica, baseada em ADHDP. A estratégia combina controle de modo deslizante e ADHDP para ajuste dos parâmetros *online* para compensar incertezas e perturbações sem depender de um modelo matemático preciso. A ADHDP analisa os erros entre estados reais e desejados fornecendo ações de controle suplementares para ajuste em torno da condição de operação normal do sinal de controle gerado pelo controle de modo deslizante. O controle suplementar aproveita a filosofia de aprendizagem baseada em ADHDP e desenvolve um controle adaptativo baseado em dados para o veículo.

Em [Shuai, Daqian e Bin \(2022\)](#), investiga-se o problema do controle de rastreamento para um sistema incerto de veículo de voo hipersônico (*Hypersonic Flight Vehicle - HFV*). O esquema proposto combina um controlador adaptativo de linha de base e um controlador de otimização auxiliar, visando o desenvolvimento de um controlador adaptativo de aprendizagem *online* capaz de lidar com as incertezas nos parâmetros e no ambiente da planta. O controlador de otimização auxiliar é projeto por meio de PDA, especificamente a ADHDP para estabelecer uma estratégia de controle direto baseada em aprendizagem orientada por dados e minimização da função de custo.

Em [Cerqueira, Lopes e Fonseca \(2024\)](#), é apresentado o projeto de controlador ótimo discreto do tipo DLQR aplicando o método ADHDP em um sistema MIMO de um modelo matemático de um F-16. O controlador projetado se caracteriza como um controlador adaptativo ótimo ao obter a solução da Equação de *Riccati* de forma *online*. Além disso, é desenvolvida uma metodologia heurística para sintonia das matrizes  $QR$  do regulador linear quadrático discreto, onde apresenta convergência para valores diversos de  $Q$ , mas indica não correlação entre a variação de  $Q$  e o número de iterações.

Em [Silva e Neto \(2021\)](#), é apresentado um método de reconstrução de estado baseado em Programação Dinâmica Adaptativa e Aprendizado por Reforço. A formulação utilizando ADP-RL realizada está relacionada a um controlador ótimo discreto com realimentação de saída. Para isso, é proposto um regulador linear quadrático discreto baseado em dados, onde o aprendizado por reforço é utilizado para estimar as ações de controle necessárias à reconstrução dos estados. O aprendizado por diferença temporal foi implementado por meio do algoritmo de iteração de valor na etapa de estimativa das ações de controle. Além disso, a aproximação matricial desenvolvida acelera o desempenho geral do algoritmo ao evitar a resolução de sistemas de equações não lineares, reduzindo a complexidade computacional. Com essa abordagem, a reconstrução de estado pode ser realizada sem a necessidade de resolver equações complexas em cada etapa do aprendizado, otimizando a eficiência do método. Uma das principais vantagens da metodologia apresentada é que ela não depende do modelo dinâmico explícito do sistema. Além disso, o método apresenta alta adaptabilidade a influências ambientais, como perturbações ou variações nos parâmetros do sistema, absorvendo essas mudanças por meio do aprendizado e das interações com o

ambiente. Assim, o controlador ADP-RL proporciona uma reconstrução eficiente e robusta do estado do sistema.

Em [Silva e Neto \(2020\)](#), é apresentada uma heurística de sintonia e análise da convergência do algoritmo de aprendizado por reforço, aplicado a um controlador baseado em dados de entrada e saída gerados por um modelo. A heurística proposta ajusta os parâmetros do gerador de dados e cria superfícies para ajudar na análise de convergência e robustez do controle ótimo *online*. Os testes são realizados em um Regulador Linear Quadrático Discreto com realimentação de saída, baseado em aprendizado por reforço e aprendizado de diferença temporal utilizando iteração de política. O algoritmo aplica o método de Mínimos Quadrados Recursivo (RLS) para estimar parâmetros *online* relacionados ao DLQR. Após a aplicação das heurísticas é possível observar a influência dos parâmetros e facilitar a análise de convergência. O trabalho mostra que o uso de heurísticas para ajustar parâmetros e construir superfícies para seleção de parâmetros oferece uma abordagem promissora para melhorar a geração de dados de maneira mais confiável, além de apoiar a análise de convergência e robustez do DLQR com realimentação de saída em sistemas baseados em aprendizado por reforço.

Em [Moura, Rego e Neto \(2019\)](#), é apresentada uma abordagem para o projeto de controladores ótimos *online* para o processo de retomada de carga a granel por uma retomadora de roda de caçamba. O objetivo principal é controlar a velocidade da retomadora para manter o fluxo de minério desejado durante o processo. O controlador desenvolvido é um DLQR com ação integral, baseado em ADHDP, para aprender de forma *online* a solução de controle ótimo. Esse método demonstra um controlador independente de modelo da planta, com ganhos autoajustáveis, que atua em tempo real utilizando apenas sinais de entrada e saída, além dos estados medidos ao longo da trajetória. O sistema apresentado garante a convergência dos parâmetros da planta, demonstrando que a solução de *Riccati* e o ganho calculado para o regime permanente são equivalentes aos obtidos pelo método de Schur. O controlador ADHDP-DLQR-IA desenvolvido mostrou ser mais eficiente do que controladores do tipo PID sintonizados por Ziegler-Nichols, *fuzzy*, e de um controlador de modo deslizante sintonizado por uma RBF.

Em [Chen, Dai e Dong \(2022\)](#), é apresentado um algoritmo de controle ótimo adaptativo de rastreamento, denominado *Reinforcement Learning Optimal Tracking Control* (RLOT), baseado em aprendizado por reforço. O controle é desenvolvido para um USV subatuado sujeito a incertezas de modelagem e perturbações externas variantes no tempo. O método integra técnica de *backstepping* a um projeto de controle ótimo. A robustez do sistema de controle da embarcação é aprimorada utilizando aproximadores baseados em redes neurais, que modelam as incertezas da dinâmica. Redes ator-crítico são contruídas para encontrar a solução aproximada da Equação de Hamilton-Jacobi-Bellman (HJB) a fim de alcançar o controle otimizado. Em comparação com os métodos existentes de

controle por aprendizado por reforço, o algoritmo RLOTC demonstrou a capacidade de compensar as dinâmicas incertas da embarcação e os distúrbios desconhecidos, alcançando o desempenho otimizado de controle ao considerar a otimização em cada projeto de *backstepping*.

O projeto apresentado em Ni, He e Wen (2013) apresenta um sinal de reforço adaptativo que melhora o desempenho do sistema, no entanto, apresenta grande esforço computacional devido à necessidade de otimização em múltiplas redes de referência. Em Mu et al. (2017), o controlador se destaca pela sua capacidade de lidar diretamente com o modelo não-linear, porém esse possui uma dependência do controle de modo deslizante, que representa uma limitação em relação à flexibilidade em comparação a outras abordagens baseadas exclusivamente em ADHDP. O controlador apresentado em Shuai, Daqian e Bin (2022) possui a vantagem de lidar com incertezas tanto nos parâmetros quanto no ambiente da planta porém uma alta complexidade computacional de implementação. O estudo apresentado em Cerqueira, Lopes e Fonseca (2024) destaca a relevância da sintonia das matrizes  $Q$  e  $R$ , no entanto, o controlador abordado é um regulador e não um seguidor. A combinação de ADP-RL apresentada em Silva e Neto (2021) possui como vantagem alta adaptabilidade a perturbações externas, com boa eficiência computacional, no entanto, a abordagem necessita de grandes volumes de dados para um aprendizado eficaz, o que impacta o desempenho em sistemas com pouca informação disponível. Em Silva e Neto (2020), o uso de heurísticas para sintonia de um DLQR com realimentação de saída baseado em aprendizado por reforço se destaca pela capacidade de ajustar parâmetros online de forma eficiente, no entanto, apresenta a aplicação voltada para um controlador do tipo regulador. Em Moura, Rego e Neto (2019) o controlador desenvolvido possui a vantagem da ação integral no seguidor para eliminar o erro de regime e como desvantagem a sua alta complexidade computacional. Em Chen, Dai e Dong (2022) a principal vantagem é a robustez do controlador em relação a incertezas de modelagem e distúrbios externos, mas a complexidade computacional de implementação do algoritmo de controle é uma desvantagem.

## 2.2 Proposta para Controle de Manobrabilidade

O projeto da estratégia de controle para manobrabilidade proposto nesse trabalho, é um controlador ótimo *online* do tipo ator-crítico denominado ADHDP-DLQT-I. O objetivo do controlador é realizar o rastreamento de um sinal de referência por meio de uma lei de controle ótimo, que minimize a função de custo quadrática, tendo como principal desafio a solução *online* implícita da ARE para o cálculo do ganho ótimo.

O controlador proposto utiliza PDA e AR para calcular a solução da ARE por meio de aproximações. Durante o processo, os ganhos são atualizados de forma adaptativa para

implementação do controle ótimo DLQT-I. O sistema utiliza a AR com modelo ator-crítico (*Q-learning*) como uma abordagem de controle ótimo adaptativo via ADHDP.

### 2.2.1 Descrição do Problema

A manobrabilidade de USVs está intrinsecamente ligada aos sistemas de controle e navegação. Os controladores para sistemas de manobrabilidade de USVs são amplamente investigados por diversos pesquisadores. As técnicas aplicadas são: controle adaptativo, *backstepping*, controle de lógica Fuzzy, controle preditivo não linear, controle robusto, controle de modo deslizante, entre outras (HALVORSEN, 2008) (AN; HE; WANG, 2022).

O desenvolvimento da teoria de controle possibilitou a aplicação de técnicas avançadas para aprimorar os sistemas de controle dos USVs e conseqüentemente a sua navegação. Esse desenvolvimento aumentou suas capacidades de manobrabilidade, autonomia e eficiência em operações. Entretanto, alguns problemas são comuns no desenvolvimento desses controladores. Aplicações reais de USVs possuem uma variedade de condições, como incertezas do sensor, dinâmica não modelada e variação de massa (ANNAMALAI et al., 2014). A presença dessas incertezas insere dinâmicas não modeladas que afetam o desempenho de controladores baseados em modelo. Apesar disso, a dinâmica e os parâmetros do sistema são geralmente assumidos como conhecidos pois é difícil e caro obter parâmetros precisos do sistema (LIU et al., 2016).

A autonomia de energia do sistema é uma característica essencial na análise do tipo de controlador de um USV. Nesse contexto, Controladores ótimos se destacam por satisfazer o desempenho especificado do sistema, minimizando o índice de custo que representa o equilíbrio entre o desempenho e custos energéticos disponíveis, garantindo uma resposta ótima.

### 2.2.2 Formulação do Problema

Seja o modelo linear do USV dado por

$$\dot{x}(t) = Ax(t) + Bu(t), \quad (2.1)$$

$$y(t) = Cx(t) + Du(t), \quad (2.2)$$

e seu índice de desempenho dado por

$$V(x(t_0), t_0) = \phi(x(T), T) + \int_{t_0}^T L(x, u, t)dt, \quad (2.3)$$

o objetivo do controlador é determinar uma lei de controle ótimo  $u^*(t)$  de forma *online* no intervalo  $[t_0, T]$  que minimize a função de custo  $V(x, t)$  e leve um dado estado inicial  $x(t_0)$  a um estado final desejado. Dessa forma,

$$V^* = \min_u V(x(t_0), t_0), \quad (2.4)$$

representa o valor ótimo, com política ótima de controle dada por

$$u^* = \arg \min_{u_k} V(x(t_0), t_0). \quad (2.5)$$

A complexidade hidrodinâmica, não linearidades, perturbações do ambiente e o acoplamento multi-variável são grandes desafios para a identificação do sistema, o que dificulta o desenvolvimento de controladores baseado em modelo. Entretanto, avaliações de desempenho de controladores em um modelo linearizado não representam completamente a resposta da dinâmica da planta.

A avaliação do desempenho de um controle linear de manobrabilidade é mais próxima de uma aplicação real, caso o esforço de controle calculado para o sistema linear  $u = \phi(x(t), u(t), Q, R)$ , seja aplicado à modelagem genérica não linear do USV dada por

$$\dot{x}(t) = f(x(t), u(t), t), \quad (2.6)$$

$$y(t) = g(x(t), u(t), t), \quad (2.7)$$

sendo  $x(t)$  o vetor de estados,  $u(t)$  o esforço de controle calculado para o sistema linearizado, e  $y(t)$  a saída do sistema não linear.

O controle de manobrabilidade de USV é realizado no sentido do seu deslocamento horizontal, dinâmica descrita somente por três velocidades do modelo 6DOF. Dentre as velocidades, duas são lineares e a outra é angular, *surge*, *sway* e *yaw*, respectivamente. Logo, a manobrabilidade é estabelecida a partir de um controlador dessas três velocidades.

### 2.2.3 Solução Proposta

O desenvolvimento de um controlador para manobrabilidade de USV que considera e compensa as incertezas do modelo de forma adaptativa, com uma resposta ótima para o desempenho e a autonomia através da diminuição do consumo energético.

Um controlador ótimo adaptativo baseado em dados (não depende explicitamente das matrizes do modelo, somente sinais de entrada e saída) que supera as limitações dos controladores baseados em modelo, com desempenho avaliado em dinâmica não-linear do *Cybership* I por um simulador com método de RK, e garante uma resposta ótima em regime permanente por meio de PDA e AR, mais especificamente ADHDP.

A sintonia do controlador é realizada por meio da Equação HJB, onde, para um índice de desempenho generalizado, a função valor  $V(x, t)^*$  representa o custo ótimo que satisfaz a Equação HJB dada por

$$\frac{\partial V^*}{\partial t} + \min_u \left[ L(x, u, t) + \left( \frac{\partial V^*}{\partial x} \right)^T \cdot f(x, u, t) \right] = 0. \quad (2.8)$$

O controlador proposto tem o modelo ator-crítico com a estrutura apresentada na Figura 2.1.

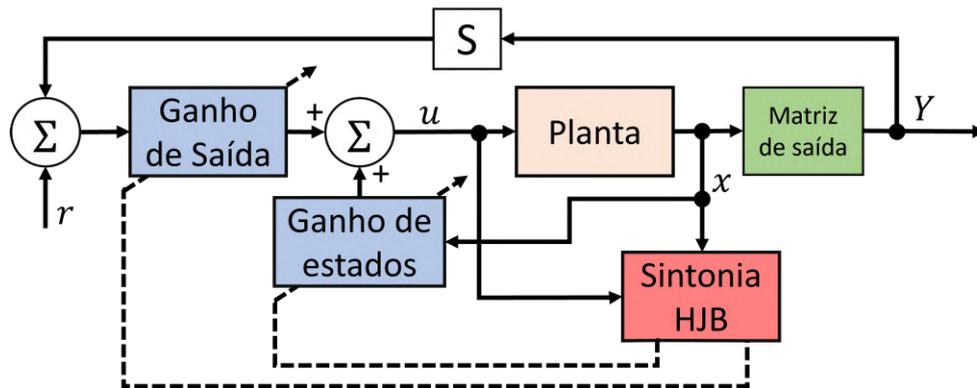


Figura 2.1 – Diagrama de blocos do sistema de controle ADHDP-DLQT-I.

A Figura 2.1 exibe a estrutura do controlador com dois blocos de ganhos, esses blocos são os atores e a sintonia HJB é o crítico do modelo ator-crítico que calcula a sintonia dos ganhos a partir dos estados  $x$  e das entradas  $u$ . Os ganhos representam uma realimentação de estados e uma realimentação de saída. O objetivo do controlador é fazer a saída  $Y$  seguir a referência  $r$ . O bloco S são os sensores do sistema.

## CAPÍTULO 3

# MODELAGEM MATEMÁTICA, CONTROLE ÓTIMO *ONLINE* E MÉTODO DE RUNGE-KUTTA

O entendimento do processo é uma etapa fundamental para o desenvolvimento do controlador. A modelagem matemática de um processo representa as suas características físicas e particularidades. Neste capítulo são apresentados conceitos da dinâmica de um USV, controle ótimo *online*, PDA, AR, ADHDP e PID-LMS.

### 3.1 Veículos Aquáticos de Superfície Não-tripulados

O crescente interesse mundial em questões comerciais, científicas, militares nos oceanos e águas rasas, motivou um crescimento no desenvolvimento de Veículos Aquáticos de Superfície Não-tripulados (*Unmanned Surface Vehicles - USV*), conforme exibido na Figura 3.1, com capacidades avançadas de orientação, navegação e controle (*Guidance, Navigation and Control - GNC*) (LIU et al., 2016).



Figura 3.1 – Veículos não-autônomos (1), autônomos (2) e não tripulados de superfície essenciais para monitoramento de derramamentos de petróleo no mar, detectam e mapeiam manchas de óleo na superfície. Os tipos UUVs (3) e AUVs (4) analisam a dispersão subaquática e inspecionam infraestruturas críticas. Os ROVs (5) realizam inspeções detalhadas e auxiliam na contenção e limpeza.

Os veículos aquáticos são divididos em superfície e subaquáticos, os veículos de superfície podem ser semi-autônomos (USVs-1) ou completamente autônomos (*Autonomous Surface Vehicles* - ASVs-2) (LIU et al., 2016), os subaquáticos (*Unmanned Underwater Vehicles* - UUVs - 3) são divididos em autônomos (*Autonomous Underwater Vehicles* - AUVs - 4) e remotos (*Remotely Operated Vehicles* - ROVs - 5) (ZANOLI; CONTE, 2003), são exibidos nas Figuras 3.1 e 3.2.

Os USVs são embarcações robóticas projetadas para navegar em superfícies aquáticas independente de uma tripulação, e podem ser autônomos ou operados remotamente (semi-autônomos) (GHAZALI; SATAR; RAHIMAN, 2024). Esses dispositivos possuem a capacidade de realizar tarefas longas e críticas com menos riscos, são compactos com grande manobrabilidade e menor custo comparado à veículos subaquáticos ou autônomos (BAE; HONG, 2023). Em geral, são capazes de aumentar a compreensão da situação, reduzir a carga de trabalho humana e melhorar o desempenho em missões, além de oferecer persistência e versatilidade no cumprimento do seu objetivo (ALIFERIS, 2021).



Figura 3.2 – Configurações de cascos de USVs: casco único (1) e catamarã (2).

Um elemento fundamental para a dinâmica do USV é o seu tipo de casco, sendo os mais comuns: caiaques (único casco) (ZHOU et al., 2024b) (1) exibido na Figura 3.2, catamarãs (cascos duplos) (DIN et al., 2024) (2) exibido na Figura 3.2 e trimarãs (cascos triplos) (MARTÍNEZ et al., 2024). O casco está diretamente ligado com seu sistema de propulsão (LIU et al., 2016). A sua configuração tradicional de atuadores inclui um motor de propulsão principal e um leme, caracterizando um sistema subatado. Um USV é considerado completamente atuado quando dispõe de dois ou mais propulsores transversais, com pelo menos um localizado na proa e outro na popa. Propulsores transversais são muito utilizados na atracação de navios e Posicionamento Dinâmico (PD) (HALVORSEN, 2008).

O sistema GNC é composto por orientação, navegação e controle (Figura 3.3). O sistema de orientação é responsável por gerenciar as tarefas que serão realizadas, planejar o caminho, gerar e atualizar comandos de trajetórias viáveis e ideais para o sistema de controle e navegação a partir das informações obtidas pelo sistema de navegação, em muitos casos, são utilizadas técnicas de otimização para definir a melhor rota, levando em consideração, por exemplo, otimização de energia, menor tempo de atividade e prevenção de colisões.

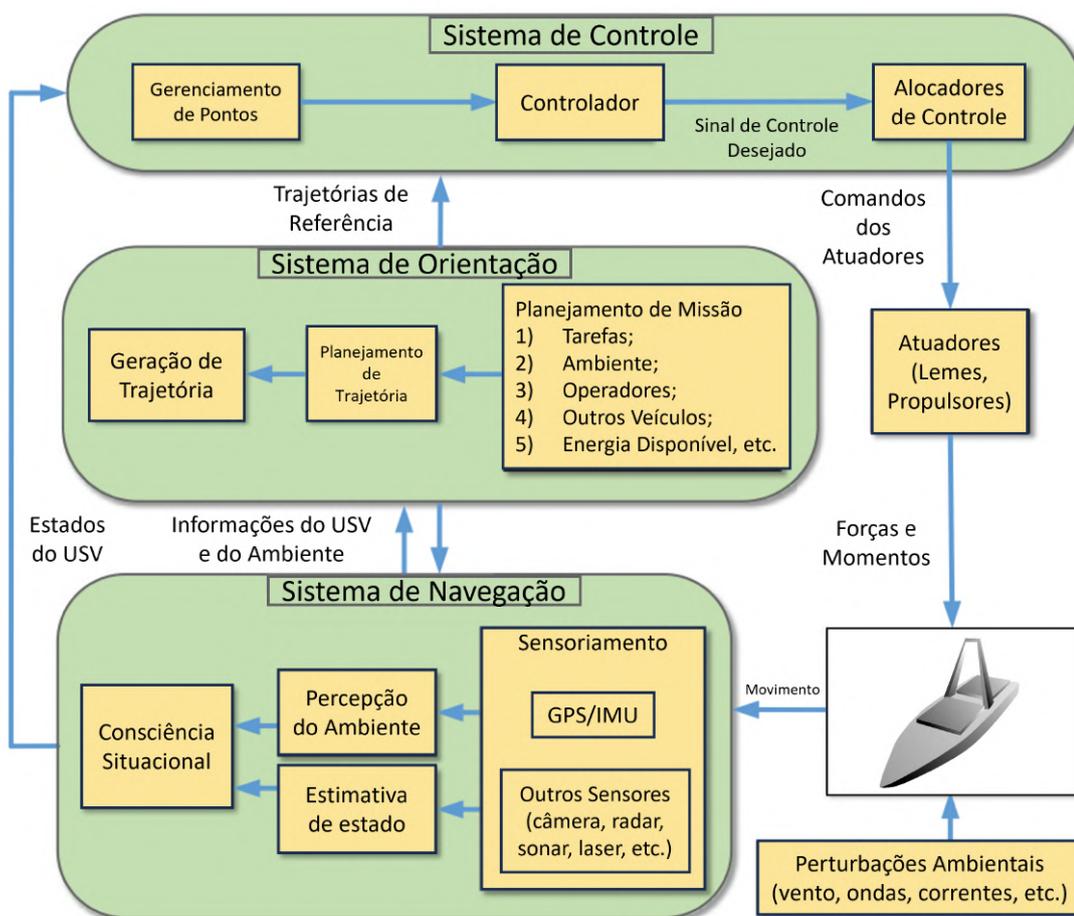


Figura 3.3 – Estrutura generalizada do sistema GNC de um USV (Fonte: Adaptado de (LIU et al., 2016)).

O sistema de navegação, apresentado na Figura 3.3, identifica os estados atuais e futuros do USV (como posição, orientação, velocidade e aceleração) e seu ambiente ao redor com base nos estados passados e atuais do USV e informações do ambiente, como correntes oceânicas e velocidade do vento obtidas pelos sensores de bordo. O sistema de controle determina as forças e momentos de controle a serem gerados de acordo com as instruções fornecidas pelos sistemas de orientação e navegação (LIU et al., 2016).

## 3.2 Modelo Matemático do USV

O movimento de uma embarcação marítima possui 6 DOFs, que representam os deslocamentos e rotações independentes responsáveis por definir sua posição e orientação. Para descrever completamente o movimento, são necessárias seis coordenadas independentes (FOSSÉN, 2011). As três primeiras coordenadas e suas derivadas temporais, correspondem à posição e ao deslocamento linear ao longo dos eixos  $X$ ,  $Y$  e  $Z$ , enquanto as três últimas coordenadas e suas derivadas temporais descrevem a orientação e o deslocamento rotacional em torno desses eixos (FOSSÉN, 2002). As seis componentes do movimento, segundo a

notação da *Society of Naval Architects and Marine Engineers* (SNAME) para embarcações marítimas (SNAME, 1950), são apresentadas na Figura 3.4.

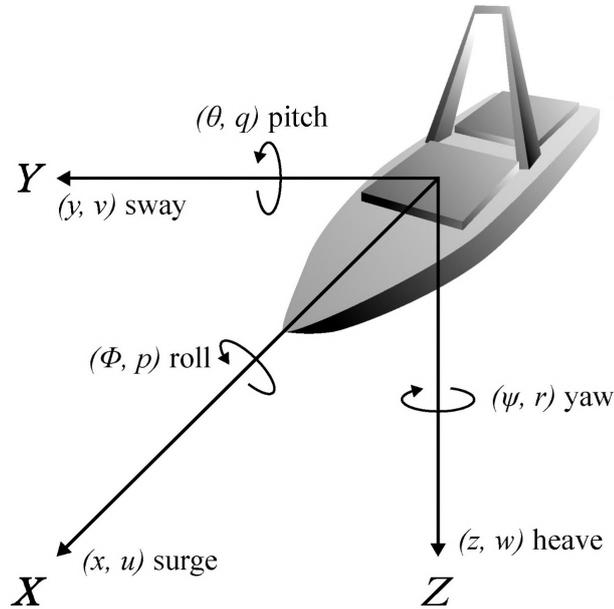


Figura 3.4 – Variáveis de movimento no sistema referencial de corpo fixo para USV.

As componentes de movimento apresentadas na Figura 3.4 são definidas como: *surge* (deslocamento longitudinal), *roll* (rotação em torno do eixo longitudinal), *sway* (deslocamento lateral/deslocamento no eixo transversal), *pitch* (rotação em torno do eixo transversal) *heave* (deslocamento vertical) e *yaw* (rotação em torno do eixo vertical) (FOSSEN, 2011). A Tabela 3.1 apresenta a notação da SNAME para embarcações marítimas.

Tabela 3.1 – Notação SNAME de Forças, momentos, velocidades e posições para embarcações.

DOF		Forças e momentos	Velocidades linear e angular	Posições e ângulos de Euler
1	<i>surge</i>	$X$	$u$	$x$
2	<i>sway</i>	$Y$	$v$	$y$
3	<i>heave</i>	$Z$	$w$	$z$
4	<i>roll</i>	$K$	$p$	$\phi$
5	<i>pitch</i>	$M$	$q$	$\theta$
6	<i>yaw</i>	$N$	$r$	$\psi$

### 3.2.1 Sistemas de Coordenadas de Referência

Ao analisar o movimento de veículos marítimos em 6DOFs, é necessário definir dois sistemas de coordenadas centrados na Terra e ao menos dois sistemas de coordenadas de

referência geográfica para descrever o movimento da embarcação (FOSSEN, 2002).

### 3.2.2 Sistemas de Coordenadas de Referência Centrados na Terra

O Sistema de Coordenada Inercial Centrado na Terra (*Earth-Centered Inertial* - ECI), é um sistema de coordenada para navegação terrestre representado, por  $\{i\} = (x_i, y_i, z_i)$ , é um sistema de coordenadas de referência não acelerado essencial para sistemas de navegação inercial (FOSSEN, 2011). A origem do sistema  $\{i\}$  é localizado no centro da Terra ( $o_i$ ) e seus eixos estão orientados conforme exibido na Figura 3.5.

O Sistema de Coordenadas de Referência Centrado na Terra e Fixo à Terra (*Earth-centered Earth-Fixed* - ECEF), apresentado na Figura na Figura 3.5, é representado por  $\{e\} = (x_e, y_e, z_e)$ , possui origem fixada no centro da Terra ( $o_e$ ), enquanto seus eixos giram em relação ao referencial inercial ECI, que é fixo no espaço. A taxa angular de rotação da Terra é  $w_e = 7,2921 \times 10^{-5}$  rad/s. Em casos onde as embarcações marítimas se deslocam a velocidades relativamente baixas, a rotação da Terra pode ser desprezada, possibilitando que  $\{e\}$  aproximado para um referencial inercial. No entanto, essa rotação não deve ser negligenciada. O sistema ECEF é amplamente utilizado para orientação global, navegação e controle, sendo essencial para descrever o movimento e a localização de embarcações em trânsito entre continentes (FOSSEN, 2011).

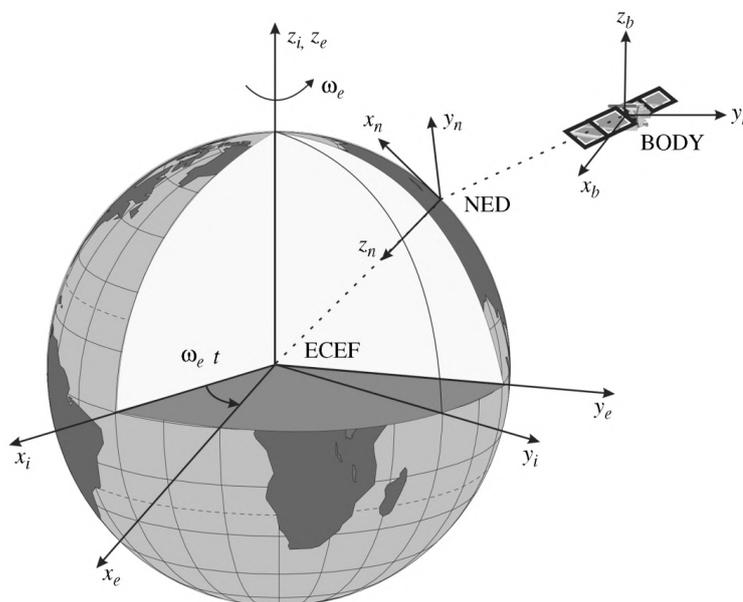


Figura 3.5 – Sistemas de Coordenadas de Referência Centrados na Terra e Sistemas de Coordenadas de referência geográfica (Fonte: (FOSSEN, 2011)).

### 3.2.3 Sistemas de Coordenadas de Referência Geográfica

O sistema de coordenadas *North-East-Down*(NED), apresentado na Figura na Figura 3.5, representado por  $\{n\} = (x_n, y_n, z_n)$  com origem  $o_n$  é o sistema utilizado

no dia-a-dia, definido como o plano tangente à superfície da Terra. Nesse sistema, o eixo  $x$  aponta para o Norte verdadeiro, o eixo  $y$  aponta para o Leste e o eixo  $z$  aponta para baixo, perpendicular à superfície da Terra. A posição de  $\{n\}$  em relação ao sistema  $\{e\}$  é determinada por dois ângulos,  $\lambda$  e  $\mu$ , que representam, a longitude e a latitude respectivamente. Em situações onde a longitude e a latitude são consideradas constantes, adota-se um plano tangente fixo à superfície da Terra para navegação. Essa técnica é comumente denominada navegação na Terra plana e pode se assumir que  $\{n\}$  é um referencial inercial (FOSSEN, 2011).

O sistema de coordenadas de corpo fixo representado por  $\{b\} = (x_b, y_b, z_b)$  com origem  $o_b$  é um sistema de coordenadas móvel associado à embarcação. A posição e orientação da embarcação são descritas em relação a um referencial inercial ( $\{e\}$  ou  $\{n\}$ ), enquanto as velocidades linear e angular são expressas no referencial fixo ao corpo. A origem do sistema ( $o_b$ ), denominada Centro de Origem (CO), é posicionada no centro da embarcação. Os eixos do corpo ( $x_b, y_b, z_b$ ) são alinhados aos eixos principais de inércia, conforme a Figura 3.6.

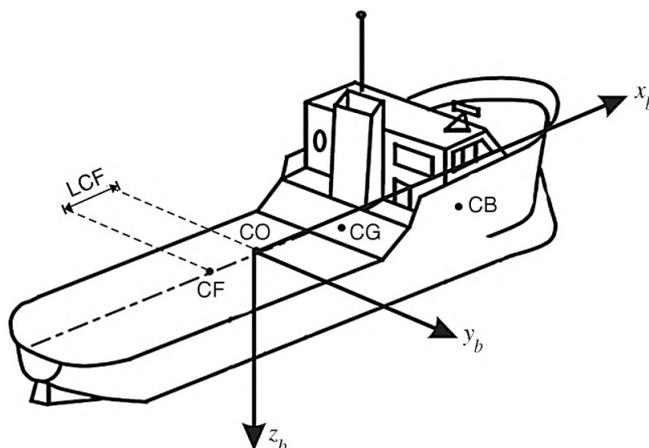


Figura 3.6 – Eixos e pontos do sistema de coordenadas de corpo fixo (Fonte: (FOSSEN, 2011))

Os eixos e pontos apresentados na Figura 3.6 são definidos como:

- $x_b$  - eixo longitudinal (direcionado da popa para frente);
- $y_b$  - eixo transversal (direcionado para estibordo);
- $z_b$  - eixo normal (direcionado para baixo);
- CO - Centro de Origem;
- CG - Centro de Gravidade;
- CB - Centro de Flutuabilidade;
- CF - Centro de Flutuação (localizado a uma distância LCF de CO na direção  $x_b$ ).

### 3.2.4 Transformações entre BODY e NED

A matriz de Rotação  $\mathbf{R}$  entre dois sistemas de coordenadas  $a$  e  $b$ , representado por  $\mathbf{R}_b^a$ , é um elemento  $SO(3)$  (*special orthogonal group of order 3*), ou seja, um elemento do grupo especial ortogonal de ordem 3:

$$SO(3) = \{ \mathbf{R} | \mathbf{R} \in \mathbb{R}^{3 \times 3}, \mathbf{R} \text{ é ortogonal e } \det \mathbf{R} = 1 \}. \quad (3.1)$$

O grupo  $SO(3)$  é um subconjunto de todas as matrizes ortogonais de ordem 3, ou seja,  $SO(3) \subset O(3)$ , onde  $O(3)$  é definido como:

$$O(3) := \{ \mathbf{R} | \mathbf{R} \in \mathbb{R}^{3 \times 3}, \mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I} \}. \quad (3.2)$$

As matrizes de rotação são úteis ao derivar as equações cinemáticas de movimento para uma embarcação. A partir das definições em (3.1) e (3.2), as seguintes propriedades podem ser declaradas:

**Propriedade da Matriz de Rotação** A matriz de rotação  $\mathbf{R} \in SO(3)$  satisfaz

$$\mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I}, \quad \det \mathbf{R} = 1,$$

implica que  $\mathbf{R}$  é ortogonal. Consequentemente, a matriz de rotação inversa é dada por  $\mathbf{R}^{-1} = \mathbf{R}^T$ . O uso da matriz de rotação entre  $\{n\}$  e  $\{b\}$  ( $\mathbf{R}_b^n$ ) necessita da definição da matriz anti-simétrica, onde uma matriz  $\mathbf{S}$  é dita anti-simétrica se:

$$\mathbf{S} = -\mathbf{S}^T, \quad (3.3)$$

nesse caso os elementos fora da diagonal principal de  $\mathbf{S}$  satisfazerm  $s_{ij} = -s_{ji}$  para  $i \neq j$  enquanto os elementos da diagonal são zero (FOSSSEN, 2011).

### 3.2.5 Dinâmica de embarcações

A seguinte convenção vetorial, introduzida por Fossen(FOSSSEN, 1991) em 1991 (HALVORSEN, 2008), expressa a equação dinâmica não-linear de movimento em 6 graus de liberdade, como apresentado em (FOSSSEN, 1994):

$$\mathbf{M}\dot{\nu} + \mathbf{C}(\nu)\nu + \mathbf{Damp}(\nu)\nu + \mathbf{g}(\eta) = \tau + \mathbf{g}_0 + \mathbf{w}, \quad (3.4)$$

sendo

$$\eta = [x, y, z, \phi, \theta, \psi]^T, \quad (3.5)$$

$$\nu = [u, v, w, p, q, r]^T, \quad (3.6)$$

os vetores de velocidades e posição/ângulos de Euler, respectivamente, como apresentado na Tabela 3.1. Essa representação permite a modelagem avançada de embarcações. As diferentes matrizes e vetores e suas respectivas propriedades são definidas nos próximos tópicos.

### 3.2.5.1 Matriz de Inércia

A matriz de inércia  $\mathbf{M}$  da Equação (3.4) inclui a inércia de corpo rígido e a inércia de massa adicionada, definidas como:

$$\mathbf{M} = \mathbf{M}_{RB} + \mathbf{M}_A. \quad (3.7)$$

A matriz de corpo rígido  $\mathbf{M}_{RB}$  é composta pela massa e momentos de inércia da embarcação, já a massa adicionada  $\mathbf{M}_A$  é devido a inércia do fluido circundante e só contribuirá quando o corpo da embarcação estiver acelerando em relação às massas de água. Como apresentado em (FOSSSEN, 2011), a matriz  $\mathbf{M}_{RB}$  generalizada para 6DOF é definida como:

$$\mathbf{M}_{RB} = \begin{bmatrix} m\mathbf{I}_{3 \times 3} & -m\mathbf{S}(\mathbf{r}_g^b) \\ m\mathbf{S}(\mathbf{r}_g^b) & \mathbf{I}_b \end{bmatrix} = \begin{bmatrix} m & 0 & 0 & 0 & mz_g & -my_g \\ 0 & m & 0 & -mz_g & 0 & mx_g \\ 0 & 0 & m & my_g & -mx_g & 0 \\ 0 & -mz_g & my_g & I_x & -I_{xy} & -I_{xz} \\ mz_g & 0 & -mx_g & -I_{yx} & I_y & -I_{yz} \\ -my_g & mx_g & 0 & I_{zx} & -I_{zy} & I_z \end{bmatrix}, \quad (3.8)$$

sendo  $m$  a massa do corpo,  $\mathbf{r}_g^b = [x_g, y_g, z_g]^T$  o vetor de  $o_b$  para o centro de gravidade no sistema de coordenadas móvel  $\{b\}$ ,  $\mathbf{S}$  é a matriz anti-simétrica,  $I$  o momento de inércia em torno dos respectivos eixos e  $I_{xy} = I_{yx}$ ,  $I_{xz} = I_{zx}$  e  $I_{yz} = I_{zy}$ , são os produtos de inércia entre os eixos subscritos, ou seja a matriz de inércia  $\mathbf{I}_b = \mathbf{I}_b^T$ . A matriz  $\mathbf{M}_{RB}$  é única e satisfaz as propriedades:  $\mathbf{M}_{RB} = \mathbf{M}_{RB}^T > 0$  e  $\dot{\mathbf{M}}_{RB} = \mathbf{0}_{6 \times 6}$ .

### 3.2.5.2 Matriz Coriolis-Centrípeta

O efeito Coriolis,  $C(\nu)$  apresentada na Equação da Equação (3.4), descreve a mudança na força necessária para alcançar a mesma aceleração em um sistema de referência rotacionado, em comparação com um sistema de referência inercial. Esse efeito ocorre quando um corpo está se movendo em um sistema de referência que está em rotação, o que desvia a trajetória do corpo devido à rotação do sistema, como por exemplo um corpo se movendo sobre a Terra. O termo coriolis  $\mathbf{C}(\nu)$  da Equação (3.4), depende da matriz de inércia  $\mathbf{M}$  e do vetor de velocidade  $\nu$ . A sua representação não é única.

Seja  $\mathbf{M}$  uma matriz de inércia 6x6 definida como

$$\mathbf{M} = \mathbf{M}^T = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{11} & \mathbf{M}_{12} \end{bmatrix} > 0, \quad (3.9)$$

então, a matriz Coriolis-Centrípeta pode ser parametrizada de modo que  $\mathbf{C}(\nu) = -\mathbf{C}^T(\nu)\nu$ , dada por

$$\mathbf{C}(\nu) = \begin{bmatrix} \mathbf{0}_{3 \times 3} & -\mathbf{S}(\mathbf{M}_{11}\nu_1 + \mathbf{M}_{12}\nu_2) \\ -\mathbf{S}(\mathbf{M}_{11}\nu_1 + \mathbf{M}_{12}\nu_2) & -\mathbf{S}(\mathbf{M}_{21}\nu_1 + \mathbf{M}_{22}\nu_2) \end{bmatrix}, \quad (3.10)$$

sendo  $\nu_1$  e  $\nu_2$ , subvetores do vetor  $\nu = [u, v, w, p, q, r]^T$ , tal que  $\nu_1 := [u, v, w]^T$  e  $\nu_2 := [p, q, r]^T$ . O vetor  $\nu = [u, v, w, p, q, r]^T$  representa as velocidades lineares ( $\nu_1$ ) e angulares ( $\nu_2$ ) conforme apresentado na Tabela 3.1. O efeito Coriolis também é dividido em termos de efeito de corpo rígido e massa adicionada, dado por

$$\mathbf{C}(\nu) = \mathbf{C}_{RB}(\nu) + \mathbf{C}_A(\nu), \quad (3.11)$$

a matriz apresentada em 3.10 engloba os dois efeitos.

### 3.2.5.3 Matriz de amortecimento

A matriz de amortecimento, **Damp** apresentada na Equação (3.4), representa as forças resultantes causadas pelo amortecimento potencial, atrito do casco, amortecimento de ondas e do desprendimento de vórtices. Essa matriz é não linear dependente do vetor de velocidade  $\nu$ . No entanto, para operações em baixa velocidade, uma matriz de amortecimento linear pode ser aplicada. Enquanto para manobras de alta velocidade pode ser aplicada uma matriz de amortecimento quadrática. No caso de amortecimento linear, ele é causado principalmente pelo termo de atrito do casco, esse termo é causado devido à teoria da camada limite laminar e é importante para movimento em baixa frequência da embarcações (FALTINSEN; SORTLAND, 1987). Além disso, haverá uma contribuição de alta frequência devido a uma camada limite turbulenta, geralmente chamado de atrito do casco quadrático ou não linear.

A matriz de amortecimento linear em **CO** é composta por coeficientes de amortecimento potencial  $D_p$  e amortecimento viscoso linear  $D_v$ , como pode ser visto a seguir:

$$\mathbf{Damp} = D_p + D_v = - \begin{bmatrix} X_u & 0 & 0 & 0 & 0 & 0 \\ 0 & Y_v & 0 & Y_p & 0 & Y_r \\ 0 & 0 & Z_w & 0 & Z_q & 0 \\ 0 & K_v & 0 & K_p & 0 & K_r \\ 0 & 0 & M_w & 0 & M_q & 0 \\ 0 & N_v & 0 & N_p & 0 & N_r \end{bmatrix}, \quad (3.12)$$

sendo  $X_u$  a força de amortecimento na direção  $X$  causada pela velocidade  $u$ , e os outros termos da matriz de forma equivalente, veja as definições da SNAME na Tabela 3.1 (SØRENSEN, 2018).

#### 3.2.5.4 Termos adicionais da equação

Essa seção apresenta a descrição dos termos restantes da Equação (3.4), como  $\mathbf{g}(\eta)$ ,  $\tau$ ,  $\mathbf{g}_0$  e  $w$ . Os termos adicionais são:

- $\tau$  - Força do atuador: Todas as forças resultantes devido ações dos atuadores, é o vetor de entrada da Equação (3.4), esse vetor é definido como  $\tau = [\tau_1, \tau_2, \tau_3]^T$ ;
- $w$  - forças ambientais: A soma de todas as forças de perturbação do sistema, como vento, ondas e correntezas atuando no corpo;
- $g_0$  - Forças gravitacionais: A gravidade sempre atua no corpo da embarcação em CG. Ela é definida como  $g_0 = mg$ , sendo  $g$  o vetor de aceleração da gravidade;
- $g(\eta)$  - Forças de empuxo: Forças que atuam na embarcação a partir da água através do empuxo. A força é maior se a embarcação desloca um volume maior de fluido. Insere um torque no corpo dependendo da configuração angular da embarcação. Em repouso, a força é igual à sua oposta força gravitacional  $g_0$  de modo a se anularem.

#### 3.2.6 Modelo para Três Graus de Liberdade (3-DOF)

O modelo para manobrabilidade, com as devidas considerações de flutuabilidade, simetria e manobras em baixa velocidade, é reduzido para 3 graus de liberdade (3-DOF), contemplando somente o movimento horizontal (FOSSSEN, 2011). Visto que, o movimento horizontal de um USV é descrito somente pelas componentes de movimento em *surge*, *sway* e *yaw*, os vetores de estado são dados por  $\nu = [u, v, r]^T$  e  $\eta = [N, E, \psi]^T$ . Dentre as considerações estão (FOSSSEN, 2011)(SETIAWAN et al., 2021):

- Os movimentos de rotação (*roll*), inclinação (*pitch*) e levantamento (*heave*) são desconsiderados, ou seja,  $w = p = q = 0$ ;
- A embarcação possui uma distribuição de massa homogênea e simetria no plano  $X - Z$ , tal que  $I_{xy} = I_{yz} = 0$ ;
- O centro da força gravitacional (CG) e o centro de flutuabilidade (CF) estão localizados em uma linha vertical, que é o eixo  $Z$ .

O deslocamento horizontal, apresentado na Figura 3.7, é descrito pelas velocidades nas componentes *u-surge*, *v-sway* e *r-yaw*, sendo  $u$  e  $v$  as velocidades lineares nas direções

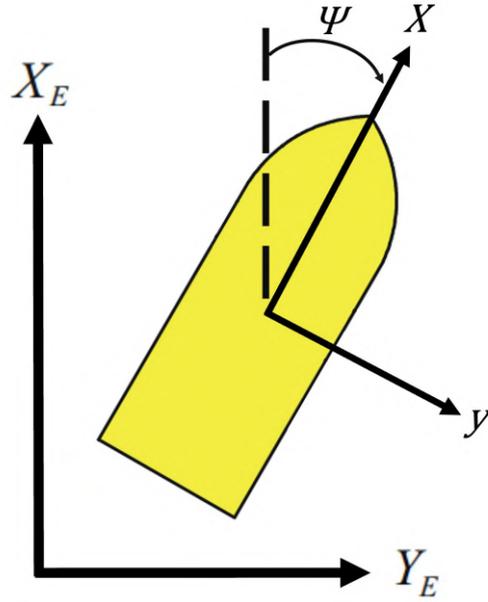


Figura 3.7 – Representação do movimento horizontal do USV (3DOF).

$X$  e  $Y$ , respectivamente, e  $r$  a velocidade angular em torno do eixo  $Z$  (Figura 3.6). Na Figura 3.7 é observado que essas velocidades são referentes às direções de corpo rígido da embarcação, sendo  $X_E$  e  $Y_E$  as coordenadas referenciais da Terra. O movimento horizontal reduz as equações cinemáticas da expressão geral da Equação (3.4) em 6-DOF, para uma rotação principal em torno do eixo  $Z$ . A equação da cinemática do USV em 3-DOF é dada por (FOSSSEN, 2011):

$$\dot{\eta} = R(\psi)\nu, \quad (3.13)$$

sendo

$$R(\psi) = \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3.14)$$

A Equação (3.13) transforma a dinâmica do sistema em referencia BODY{b} para referência NED{n}. O termo  $\eta = [N, E, \psi]^T$  possui  $N$  e  $E$  como posição ao norte e leste respectivamente,  $\psi$  é o ângulo de orientação da embarcação (rumo), definido como a diferença angular entre o eixo  $X_b$  e  $X_e$ . A matriz rotacional  $R(\psi)$  do sistema para 3-DOF é dada pela equação 3.14 e possui as características apresentadas na seção 3.2.4.

Após as devidas considerações para 3DOF os termos da Equação (3.4) são reescritos, além disso, os termos de flutuabilidade e gravidade são desconsiderados visto que o objetivo é o movimento horizontal da embarcação. Considerando um amortecimento linear, o modelo não-linear é reescrito na forma (FOSSSEN, 2011):

$$\mathbf{M}\dot{\nu} + \mathbf{C}(\nu)\nu + \mathbf{Damp}\nu = \tau + \mathbf{w}, \quad (3.15)$$

sendo

$$M = \begin{bmatrix} m - X_{\dot{u}} & 0 & 0 \\ 0 & m - Y_{\dot{v}} & mx_g - Y_{\dot{r}} \\ 0 & mx_g - Y_{\dot{r}} & I_z - N_{\dot{r}} \end{bmatrix}, \quad (3.16)$$

$$C(\nu) = \begin{bmatrix} 0 & 0 & -(m - Y_{\dot{v}})v - (mx_g - Y_{\dot{r}})r \\ 0 & 0 & (m - X_{\dot{u}})u \\ (m - Y_{\dot{v}})v + (mx_g - Y_{\dot{r}})r & -(m - X_{\dot{u}})u & 0 \end{bmatrix}, \quad (3.17)$$

$$D = \begin{bmatrix} -X_u & 0 & 0 \\ 0 & -Y_v & -Y_r \\ 0 & -N_v & -N_r \end{bmatrix}, \quad (3.18)$$

a matriz anti-simétrica do efeito Coriolis  $\mathbf{C}(\nu)$  é escolhida segundo [Fossen e Fjellstad \(1995\)](#).

### 3.3 Controle Ótimo

O controle ótimo é um método de controle moderno que tem como objetivo determinar uma ação de controle capaz de otimizar um índice de desempenho. Em sistemas de controle clássicos, as especificações de projeto, como as figuras de mérito, são definidas de forma explícita no início do projeto. Entretanto, no controle ótimo, a função de utilidade que compõe o índice de desempenho compreende os requisitos de projeto para o sistema controlado de forma implícita. A seleção de uma função de utilidade é uma etapa crítica no projeto de controladores ótimos, pois ela carrega todas as informações para o controlador atender os objetivos desejados no projeto. As características e o desempenho do controlador estão diretamente relacionados com essa função ([LEWIS; VRABIE; SYRMOS, 2012](#))([RÊGO, 2014](#)).

A formulação de uma função de utilidade inclui as especificações de desempenho como: tempo mínimo para alcançar um estado de operação particular, energia mínima necessária para controlar o sistema dinâmico, erro mínimo de operação sujeito a restrições sobre a energia de controle a ser gasta, máxima relação de desempenho em comparação a um sistema ideal, entre outras. O valor do índice é desconhecido até o fim do processo de otimização. Em relação ao significado do índice de desempenho, o problema de otimização pode ser de minimização ou maximização. No entanto, a maioria dos problemas de controle ótimo lida com a minimização de erros e do esforço de controle ([LEWIS; VRABIE; SYRMOS, 2012](#)). Um controlador que minimiza o índice de desempenho, é considerado ótimo não necessariamente em comparação a sistemas de controle clássico, mas no sentido de minimizar as características incluídas no índice o considerando o melhor sistema possível de um tipo particular.

De modo geral, as técnicas de controle ótimo focam em determinar o controlador que alcançará o melhor valor possível para o índice de desempenho através de sua minimização sujeita a restrições dinâmicas, sendo a solução do problema dependente de uma planta específica a ser controlada e o índice de desempenho a ser otimizado (MEINSMA; SCHAFT, 2023).

A teoria moderna de controle ótimo surgiu a partir de 1960 para englobar sistemas modernos, complexos e multivariáveis que não eram contemplados na teoria clássica de controle. Esses sistemas mesmo sendo não-lineares, podem ser linearizados em torno de um ponto ou trajetória de operação desejada. A abordagem clássica para solução de problemas de controle ótimo é baseada na teoria do cálculo variacional que encontra as condições necessárias para otimalidade na forma de equações de Euler-Lagrange. Há basicamente dois objetivos de projeto de controle, o problema do regulador para estabilizar o sistema tornando seus estados e saídas pequenos, e o problema do seguidor para controlar o sistema por meio do rastreamento de trajetórias prescritas mantendo os estados limitados (FERREIRA et al., 2016).

### 3.3.1 O Problema de Controle Ótimo

Seja um sistema dinâmico planta descrita pela equação discreta não-linear e variante no tempo:

$$x_{k+1} = f(x_k, u_k, k), \quad (3.19)$$

sendo  $x_k \in \mathbb{R}^n$  o vetor de estados,  $u_k \in \mathbb{R}^m$  o vetor de entrada de controle, com condição inicial  $x_0$ . A equação 3.19 representa a restrição, pois determina o estado no tempo  $k + 1$  dado estado e controle no tempo  $k$ . Seja o índice de desempenho escalar escolhido para o sistema dinâmico, dado por

$$V_i = \phi(N, x_N) + \sum_{k=i}^{N-1} L(x_k, u_k, k), \quad (3.20)$$

sendo o intervalo de tempo  $[i, N]$  discreto com passo de amostragem fixo. A função  $\phi$  de tempo e estado final, e  $L$  uma função do estado, da entrada de controle e do tempo em cada instante  $k \in [i, N]$  (LEWIS; VRABIE; SYRMOS, 2012).

O problema do controle ótimo é encontrar a lei de controle  $u_k^*$  no intervalo  $[i, N]$  ( $u_k^*, \forall k \in [i, N]$ ) que leva o sistema 3.19 ao longo de uma trajetória  $x_k^*$  tal que o valor do índice de desempenho seja otimizado. O objetivo é maximizar ou minimizar o índice de desempenho, dependendo do seu significado associado. A sua solução clássica é baseada na teoria de multiplicadores de Lagrange (Cálculo Variacional). O problema do controle ótimo será apresentado no contexto do regulador e do rastreador (LEWIS; VRABIE; SYRMOS, 2012).

### 3.3.2 O Problema do Regulador Linear Quadrático Discreto

Analisando o caso particular do Regulador Linear Quadrático, os índices de desempenho são considerados quadráticos e podem ser considerados como superfícies quadráticas de dimensão  $(n + m)$ . O problema do regulador discreto é um caso particular do problema de controle ótimo linear quadrático discreto. Seja a planta a ser controlada descrita pela equação de estado linear e variante no tempo:

$$x_{k+1} = A_k x_k + B_k u_k, \quad (3.21)$$

com  $x_k \in \mathbb{R}^n$  e  $u_k \in \mathbb{R}^m$ , com condição inicial conhecida  $x_i$ . O objetivo é encontrar uma lei de controle  $u_k$  que minimiza o índice de desempenho quadrático, definido no intervalo de interesse  $[i, N]$ , dado por

$$V_i = \frac{1}{2} x_N^T S_N x_N + \frac{1}{2} \sum_{k=i}^{N-1} (x_k^T Q_k x_k + u_k^T R_k u_k), \quad (3.22)$$

com restrição à Equação dinâmica (3.21). As matrizes  $Q_k = Q_k^T \geq 0$  e  $R_k = R_k^T > 0$  são matrizes variantes no tempo de ponderação de estado e controle, respectivamente. A matriz  $S = S^T \geq 0$  pondera o desvio de estado final no instante  $N$ . As matrizes  $Q$  e  $S$  possuem dimensão  $(n \times n)$  e a matriz  $R$   $(m \times m)$ . No caso do regulador, a lei de controle ótima deve conduzir a planta de um estado não nulo  $x_i$  no instante  $i$  para o estado nulo no instante  $N$ , levando em consideração a energia mínima de controle (LEWIS; VRABIE; SYRMOS, 2012). De forma generalizada, a minimização desse índice é dada por

$$V^*(x_k) = \min V(x_k, u_k, k), \quad (3.23)$$

sendo  $V^*(x_k)$  a função valor ótima no instante de tempo  $k$ . Analisando o índice de desempenho para um horizonte infinito ou  $N$  muito grande, e liberando o primeiro termo do somatório quando  $i = 1$ , o índice é reescrito,

$$V = \frac{1}{2} (x_k^T Q_k x_k + u_k^T R_k u_k) + \frac{1}{2} \sum_{k=i+1}^{\infty} (x_k^T Q_k x_k + u_k^T R_k u_k), \quad (3.24)$$

sendo o segundo termo da Equação (3.24) o próprio índice de desempenho com início em  $k + 1$  (LEWIS; VRABIE; SYRMOS, 2012). Logo, a Equação é reescrita,

$$V_i = \frac{1}{2} (x_k^T Q_k x_k + u_k^T R_k u_k) + V(x_{k+1}). \quad (3.25)$$

A Equação (3.25), é a Equação de Bellman para o DLQR. Assumindo que a função valor é quadrática para o estado, então a solução da função valor no instante de tempo  $k$  é

$$V(x_k) = \frac{1}{2} x_k^T P_k x_k, \quad (3.26)$$

sendo  $P_k \geq 0$  uma matriz de ponderação semidefinida positiva para todo  $k$ . Substituindo a Equação (3.26) em (3.25), a Equação de Bellman é reescrita como

$$2V(x_k) = x_k^T P_k x_k = x_k^T Q_k x_k + u_k^T R_k u_k + x_{k+1}^T P_{k+1} x_{k+1}, \quad (3.27)$$

substituindo a Equação (3.21) de estado na Equação (3.27)

$$2V(x_k) = x_k^T P_k x_k = x_k^T Q_k x_k + u_k^T R_k u_k + (Ax_k + Bu)^T P_{k+1} (Ax_k + Bu). \quad (3.28)$$

Assumindo uma lei de controle com política de realimentação de estado constante na forma

$$u_k = -K_k x_k, \quad (3.29)$$

sendo  $K_k$  os ganhos do controlador, a Equação (3.28) pode ser escrita como

$$x_k^T P_k x_k = x_k^T Q_k x_k + x_k^T K_k^T R_k K_k x_k + x_k^T (A - BK_k)^T P_{k+1} (A - BK_k) x_k. \quad (3.30)$$

essa equação vale para todas as trajetórias e com as devidas manipulações, tem-se

$$P_k = (A - BK_k)^T P_{k+1} (A - BK_k) + Q_k + K_k^T R_k K_k, \quad (3.31)$$

essa é uma Equação de Lyapunov que possui como solução os elementos  $P_k$  e  $P_{k+1}$ . Para encontrar a lei de controle ótima  $u_k^*$  que minimize a Equação (3.28) considerando a restrição do sistema a sua equação de estados, o mínimo de  $V(x_k)$  é encontrado fazendo

$$\frac{\partial V(x_k, u_k)}{\partial u_k} = R_k u_k + B^T P_{k+1} (Ax_k + Bu_k) = 0, \quad (3.32)$$

reorganizando a Equação (3.40) para encontrar a lei de controle ótima, tem-se

$$u_k^* = -(B^T P_{k+1} B + R_k)^{-1} B^T P_{k+1} A x_k, \quad (3.33)$$

comparando as Equações (3.33) e (3.29), o ganho ótimo é dado por

$$K_k = -(B^T P_{k+1} B + R_k)^{-1} B^T P_{k+1} A, \quad (3.34)$$

e a lei de controle representada por

$$u_k^* = -K_k x_k, \quad (3.35)$$

substituindo a Equação de ganho ótimo (3.34) na Equação de Lyapunov (3.31), obtém-se a recorrência de *Riccati* dada por (LEWIS; VRABIE; SYRMOS, 2012)

$$A^T P_{k+1} A - P_k + Q_k - A^T P_{k+1} B (B^T P_{k+1} B + R_k)^{-1} B^T P_{k+1} A = 0. \quad (3.36)$$

### 3.3.3 O problema do Rastreador Linear Quadrático Discreto

O objetivo do projeto de controle de um rastreador ótimo discreto (DLQT) é desenvolver um seguidor (ou rastreador) discreto de uma trajetória desejada ao longo de um intervalo de tempo usando uma lei de controle de malha fechada (LEWIS; VRABIE; SYRMOS, 2012). Seja um sistema linear com Equação de estado (3.21) e Equação de saída dada por

$$y_k = Cx_k, \quad (3.37)$$

sendo a saída  $y_k \in R^r$  o vetor de saída, tendo como objetivo seguir um sinal de referência desejado  $r_k$  ao longo de um intervalo de tempo  $[0, N]$ , então, o objetivo é minimizar a função de custo

$$J_i = \frac{1}{2}(Cx_N - r_N)^T P(Cx_N - r_N) + \frac{1}{2} \sum_{k=i}^{N-1} [(Cx_k - r_k)^T Q(Cx_k - r_k) + u_k^T R u_k], \quad (3.38)$$

sendo as matrizes  $P \geq 0, Q \geq 0, R > 0$ . O coestado do sistema é dado por

$$\lambda_k = A^T \lambda_{k+1} + C^T Q C x_k - C^T Q r_k, \quad (3.39)$$

onde pode ser observado que o sistema hamiltoniano resultante não é homogêneo e é conduzido por uma função  $-C^T Q r_k$  dependente da trajetória desejada (LEWIS; VRABIE; SYRMOS, 2012). A condição de estacionariedade é

$$0 = B^T \lambda_{k+1} + R u_k, \quad (3.40)$$

e as condições de contorno

$$\text{Condição inicial - } x_0, \quad (3.41)$$

$$\lambda_N = C^T P(Cx_N - r_N). \quad (3.42)$$

A partir da condição de estacionariedade do controle ótimo dada pela Equação (3.40), tem-se que

$$u_k = -R^{-1} B^T \lambda_{k+1}, \quad (3.43)$$

o sistema Hamiltoniano não homogêneo é dado por

$$\begin{bmatrix} x_k + 1 \\ \lambda_k \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^T \\ C^T Q C & A^T \end{bmatrix} \begin{bmatrix} x_k \\ \lambda_{k+1} \end{bmatrix} + \begin{bmatrix} 0 \\ -C^T Q \end{bmatrix} r_k. \quad (3.44)$$

Essa versão do sistema de controle ótimo não pode ser implementada, pois as condições de contorno estão divididas entre os instantes  $k = 0$  e  $k = N$  (LEWIS; VRABIE; SYRMOS, 2012). Esse modelo de controle não pode ser expresso como uma realimentação

do estado. No entanto,  $u_k$  pode ser expresso como uma combinação de realimentação linear do estado mais um termo adicional dependente de  $r_k$  (LEWIS; VRABIE; SYRMOS, 2012).

A partir da Equação (3.42), é assumido que para todo  $k \leq N$

$$\lambda_k = S_k x_k - v_k, \quad (3.45)$$

sendo  $S_k \in \mathbb{R}^{n \times n}$  e  $v_k \in \mathbb{R}^n$ . Substituindo a Equação (3.45) na primeira linha da Equação (3.44), tem-se que

$$x_{k+1} = Ax_k - BR^{-1}B^T S_{k+1} x_{k+1} + BR^{-1}B^T v_{k+1}, \quad (3.46)$$

isolando  $x_{k+1}$  na Equação (3.46), tem-se que

$$x_{k+1} = (I + BR^{-1}B^T S_{k+1})^{-1} (Ax_k + BR^{-1}B^T v_{k+1}), \quad (3.47)$$

substituindo a Equação (3.46) na segunda linha da Equação (3.44) e por sua vez substituindo na Equação (3.45), tem-se que

$$S_k x_k - v_k = C^T Q C x_k + A^T S_{k+1} (I + BR^{-1}B^T S_{k+1})^{-1} \times (Ax_k + BR^{-1}B^T v_{k+1}) - A^T v_{k+1} - C^T Q r_k, \quad (3.48)$$

reescrevendo a Equação (3.48), tem-se que

$$\begin{aligned} & [-S_k + A^T S_{k+1} (I + BR^{-1}B^T S_{k+1})^{-1} A + C^T Q C] x_k \\ & + [v_k + A^T S_{k+1} (I + BR^{-1}B^T S_{k+1})^{-1} BR^{-1}B^T v_{k+1} \\ & - A^T v_{k+1} - C^T Q r_k] = 0. \end{aligned} \quad (3.49)$$

Esta equação é válida para todas as sequências de estados  $x_k$  para qualquer condição inicial  $x_0$ . Aplicando o lema da inversa na Equação (3.49), é possível obter os valores de  $S_k$  e  $v_k$ , dados por

$$S_k = A^T [S_{k+1} - S_{k+1} B (B^T S_{k+1} B + R)^{-1} B^T S_{k+1}] A + C^T Q C, \quad (3.50)$$

$$v_k = [A^T - A^T S_{k+1} B (B^T S_{k+1} B + R)^{-1} B^T] v_{k+1} + C^T Q r_k. \quad (3.51)$$

Realizando uma análise por comparação entre as Equações (3.45) e (3.42), as condições de contorno são determinadas como

$$S_N = C^T P C, \quad (3.52)$$

$$v_N = C^T P r_N. \quad (3.53)$$

Visto que as variáveis auxiliares  $S_k$  e  $v_k$  podem ser calculadas, a suposição dada na Equação (3.45) é válida e o controle ótimo é dado por

$$u_k = -R^{-1}B^T\lambda_{k+1} = -R^{-1}B^T(S_{k+1}x_{k+1} - v_{k+1}), \quad (3.54)$$

substituindo a Equação de estado (3.21), pré-multiplicando por  $R$  e resolvendo para  $u_k$ , tem-se

$$u_k = (B^T S_{k+1} B + R)^{-1} B^T (-S_{k+1} A x_k + v_{k+1}), \quad (3.55)$$

o ganho de realimentação  $K_k$  é composto pelos elementos que multiplicam os estados  $x_k$  na Equação (3.55), e o ganho de avanço  $K_k^v$  os elementos que multiplicam a variável auxiliar  $v_{k+1}$  (LEWIS; VRABIE; SYRMOS, 2012), logo

$$K_k = (B^T S_{k+1} B + R)^{-1} B^T S_{k+1} A, \quad (3.56)$$

$$K_k^v = (B^T S_{k+1} B + R)^{-1} B^T, \quad (3.57)$$

a lei de controle é reescrita como

$$u_k = -K_k x_k + K_k^v v_{k+1}. \quad (3.58)$$

### 3.4 Formulação do Rastreador Linear Quadrático Discreto com Ação Integral

O controlador DLQT-I possui propriedades próximas a um controlador PID na forma PI (ação proporcional e integral), pois diferente do DLQT este possui uma ação integral, além da proporcional e esta característica elimina o erro de regime que pode acontecer principalmente em sistemas do tipo 0, como é o caso da planta utilizada. Uma das principais diferenças entre o controlador ótimo MIMO DLQT-I e um controlador PI convencional é que o controlador ótimo utiliza realimentação de saída e estado, enquanto o PI possui apenas realimentação de saída e é aplicado a sistemas de entrada e saída única (SISO) (RUSCIO, 2012).

A sua implementação formula o rastreador como um problema de regulador, por meio de uma reformulação da função de custo do DLQR e das matrizes de estados do sistema para introdução de uma variável ou vetor de erro como um novo estado (RUSCIO, 2012). O novo índice de desempenho é dado por

$$V_i = \frac{1}{2} x_N^T S x_N + \frac{1}{2} \sum_{k=i}^{N-1} ((y_k - r)^T Q (y_k - r) + \Delta u_k^T R \Delta u_k), \quad (3.59)$$

sendo  $S = S^T \geq 0$ ,  $Q = Q^T \geq 0$  e  $R = R^T > 0$  as matrizes de ponderação de desvio de estado final, de estado e controle respectivamente. Além disso,  $\Delta u_k = u_k - u_{k-1}$  é o incremento de controle e a referência  $r_k$  é considerada uma constante  $r$ .

Analisando o problema para horizontes grandes ou infinitos ( $N = \infty$ ) ou quando  $S$  é escolhido como a solução de *Riccati*, então a Equação (3.60) pode ser reescrita como

$$V_i = \frac{1}{2} \sum_{k=i}^{\infty} ((y_k - r)^T Q (y_k - r) + \Delta u_k^T R \Delta u_k). \quad (3.60)$$

A solução segundo [Ruscio \(2012\)](#) considera uma modelagem para expansão do sistema onde o sistema é dado por

$$\Delta x_{k+1} = A_d \Delta x_k + B_d \Delta u_k, \quad (3.61)$$

$$y_k = y_{k-1} + C_d \Delta x_k, \quad (3.62)$$

sendo  $\Delta x_k = x_k - x_{k-1}$ . Expandindo o sistema de equações, inserindo a Equação (3.62) em (3.61), tem-se o novo sistema em espaço de estados dado por

$$\begin{bmatrix} \Delta x_{k+1} \\ y_k \end{bmatrix} = \begin{bmatrix} A_d & 0_{n \times r} \\ C_d & I_{r \times r} \end{bmatrix} \begin{bmatrix} \Delta x_k \\ y_{k-1} \end{bmatrix} + \begin{bmatrix} B_d \\ 0_{r \times m} \end{bmatrix} \Delta u_k \quad (3.63)$$

$$y_k = \begin{bmatrix} C_d & I_{r \times r} \end{bmatrix} \begin{bmatrix} \Delta x_k \\ y_{k-1} \end{bmatrix}. \quad (3.64)$$

O índice de desempenho da Equação (3.60) com  $r = 0$  e o modelo em espaço de estados dado pelas Equações (3.63) e (3.64), representam o problema de controle linear quadrático padrão. Assumindo que  $r$  é uma referência constante diferente de zero, a equação de medição (3.64) é reescrita como

$$y_k - r = y_{k-1} - r + C_d \Delta x_k. \quad (3.65)$$

A partir da Equação (3.65), incluindo o erro do sistema, as Equações de estado (3.63) e saída (3.64) são reescritas como

$$\underbrace{\begin{bmatrix} \Delta x_{k+1} \\ y_k - r \end{bmatrix}}_{\tilde{x}_{k+1}} = \underbrace{\begin{bmatrix} A_d & 0_{n \times r} \\ C_d & I_{r \times r} \end{bmatrix}}_{\tilde{A}_d} \underbrace{\begin{bmatrix} \Delta x_k \\ y_{k-1} - r \end{bmatrix}}_{\tilde{x}_k} + \underbrace{\begin{bmatrix} B_d \\ 0_{r \times m} \end{bmatrix}}_{\tilde{B}_d} \Delta u_k, \quad (3.66)$$

$$\underbrace{y_k - r}_{\tilde{y}_k} = \underbrace{\begin{bmatrix} C_d & I_{r \times r} \end{bmatrix}}_{\tilde{C}_d} \underbrace{\begin{bmatrix} \Delta x_k \\ y_{k-1} - r \end{bmatrix}}_{\tilde{x}_k}. \quad (3.67)$$

O novo modelo em espaço de estados é dado na forma

$$\tilde{x}_{k+1} = \tilde{A} \tilde{x}_k + \tilde{B} \Delta u_k, \quad (3.68)$$

$$\tilde{y}_k = \tilde{C}_d \tilde{x}_k, \quad (3.69)$$

e o índice de desempenho

$$V_i = \frac{1}{2} \sum_{k=i}^{\infty} (\tilde{y}_k^T Q \tilde{y}_k + \Delta u_k^T P \Delta u_k) = \frac{1}{2} \sum_{k=i}^{\infty} (\tilde{x}_k^T \tilde{Q} \tilde{x}_k + \Delta u_k^T P \Delta u_k), \quad (3.70)$$

sendo a matriz de ponderação  $\tilde{Q} = \tilde{C}^T Q \tilde{C}$  (RUSCIO, 2012).

A solução para o problema de controle ótimo linear quadrático padrão formado pelas Equações (3.68), (3.69) e (3.70), consiste em minimizar o índice de desempenho (3.70) em relação ao desvio de controle  $\Delta u_k$  e sujeito à restrição dada pela Equação (3.68). A solução é dada pela realimentação de estado

$$\Delta u_k = K \tilde{x}_k, \quad (3.71)$$

sendo a matriz de ganho de realimentação  $K$ , dada de forma análoga à apresentada na Seção 3.3.2 na Equação (3.34), dada por

$$K = -(R + \tilde{B}^T P \tilde{B})^{-1} \tilde{B}^T P \tilde{A}, \quad (3.72)$$

sendo  $P$  a solução positiva para a Equação Algébrica de *Riccati* em tempo discreto (DARE), dada por

$$P = \tilde{Q} + K^T R K + (\tilde{A} + \tilde{B} K)^T P (\tilde{A} + \tilde{B} K). \quad (3.73)$$

Expandindo os termos da Equação (3.71), tem-se o controlador na forma incremental dado por

$$\Delta u_k = \underbrace{\begin{bmatrix} K_1 & K_2 \end{bmatrix}}_K \underbrace{\begin{bmatrix} \Delta x_k \\ y_{k-1} - r \end{bmatrix}}_{\tilde{x}_k}, \quad (3.74)$$

e pode ser reescrita como  $u_k = u_{k-1} + \Delta u_k$  ou como

$$u_k = u_{k-1} + K_1 \Delta x_k + K_2 (y_{k-1} - r_k), \quad (3.75)$$

sendo  $r = r_k$  na Equação (3.4) para obter a equação do controlador (3.75). Essa lei de controle apresenta as similaridades com um controlador PI convencional (RUSCIO, 2012).

O controlador DLQT-I fornece um erro em regime permanente nulo, ou seja,  $y = r$  em estado estacionário. Para isso, o sistema de malha fechada deve ser estável e as variáveis de controle devem estar nos limites permitidos, sem saturação (RUSCIO, 2012).

### 3.5 Programação Dinâmica Adaptativa

A Programação Dinâmica é um método para determinar soluções de controle ótimo utilizando o princípio de otimalidade de Bellman (BELLMAN, 1958) "olhando para trás" no tempo a partir de estados e objetivos desejados. Projetos baseados em PD possuem a desvantagem de apresentarem a solução de forma *offline*. A implementação de um sistema capaz de resolver o problema de controle ótimo de forma *online* requer a resolução da equação de Bellman. A solução dessa equação avalia o desempenho das políticas de controle atuais e fornece as orientações para sua melhoraria. Alguns métodos que englobam Aprendizado por Reforço (AR) e Programação Dinâmica Adaptativa (PDA) possibilitam a implementação de controladores ótimos *online* (LEWIS; VRABIE; SYRMOS, 2012).

O AR é inspirado nos mecanismos naturais de aprendizado, onde animais adaptam suas ações baseado em estímulos de recompensa ou punição recebidos a partir do ambiente. O AR trata de uma classe de métodos de aprendizado que possibilitam o projeto de controladores adaptativos que aprendem *online* e encontrar soluções para problemas como o controle ótimo. A otimização ocorre por meio de uma estrutura composta por um ator ou agente, que interage com o ambiente e modifica suas ações, ou políticas de controle, baseado em estímulos recebidos em resposta às suas ações, essas estruturas são conhecidas como ator-crítico (LEWIS; VRABIE; SYRMOS, 2012).

As estruturas ator-crítico são um tipo de extensão de controladores adaptativos diretos e indiretos, pois utilizam uma variação das técnicas direta e indireta em conjunto, identificando o valor de desempenho da política de controle atual e, a partir dessa informação, atualiza o controlador (RÊGO, 2014).

As estruturas ator-crítico, apresentadas na Figura 3.8, representam um tipo de algoritmo de aprendizado por reforço. Essas estruturas operam de forma "para frente" no tempo, visando calcular decisões ótimas para uma implementação *online*. Nesse processo, o ator aplica uma ação, ou política de controle ao ambiente, enquanto o elemento crítico avalia o valor dessa ação, fornecendo diretrizes para o aprimoramento da política.

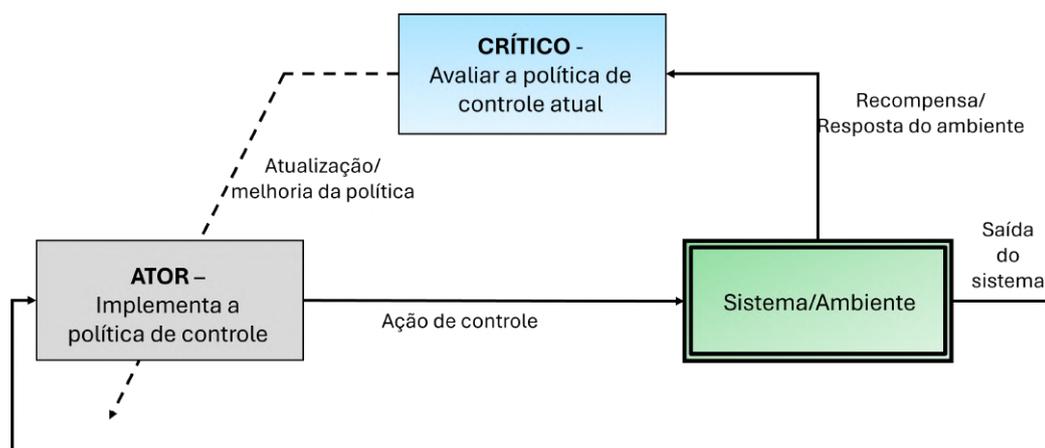


Figura 3.8 – Estrutura ator-crítico do Aprendizado por reforço. Essa estrutura fornece métodos para aprender soluções de controle ótimas on-line com base em dados medidos ao longo das trajetórias do sistema. (Fonte: Adaptado de (LEWIS; VRABIE, 2009))

O mecanismo de aprendizado apresentado na Figura 3.8 opera em duas etapas, a avaliação de política realizada pelo crítico onde o sistema analisa os resultados da aplicação das ações de controle atuais no ambiente. Esses resultados são avaliados por meio do índice de desempenho, que quantifica a distância da ação atual para a ação ótima. A partir da avaliação realizada pelo crítico, o ator realiza a melhoria da política. Diversas estratégias podem ser utilizadas para aprimorar a política de controle, de forma que a nova política apresente um desempenho superior à anterior.

A PDA compõe uma família de técnicas de aprendizado por reforço para o controle por realimentação (LEWIS; VRABIE, 2009). A PDA pode ser vista como parte do controle adaptativo pois devido ao "olhar para frente" (*lookahead*) implícito, alcança estabilidade sob condições mais fracas do que as formas bem conhecidas de controle adaptativo direto e indireto, como também pode ser formulado como uma parte do controle ótimo pois busca métodos gerais viáveis do ponto de vista computacional para o caso estocástico não linear (WERBOS, 2012).

Apesar da solução de controle ótimo poder ser calculada usando PD, é um procedimento de retrocesso no tempo e não é indicado para projeto *offline* devido ao alto custo computacional associado com a sua implementação *online*, principalmente para sistemas de ordem elevada (RÊGO, 2014). A PDA formula métodos de aprendizado por reforço *online* para resolver a Equação de Bellman, a partir da Iteração de Política e Iteração de Valor, em busca de solucionar o problema de controle ótimo sem o modelo explícito do processo, com base apenas em dados medidos ao longo das trajetórias do sistema. A solução é obtida aplicando críticos adaptativos para realizar aproximações de programação dinâmica. Os dois principais elementos desse método é o erro de diferença temporal e a aproximação da função valor (LEWIS; VRABIE, 2009).

A partir dos métodos de PDA, destaca-se a Programação Dinâmica Heurística Dependente de Ação (ADHDP) que é direcionada para a implementação *online* de controle ótimo do tipo linear quadrático discreto. O objetivo principal é aproximar soluções de programação dinâmica a partir de métodos de aproximação e esquemas de iteração de política para encontrar a solução da equação HJB-Riccati de forma implícita (RÊGO, 2014).

A ADHDP é baseada na abordagem de aprendizagem  $Q$ . A função  $Q$  representa os custos instantâneo e futuro parametrizado para o DLQR. Esse método consiste em estimar a função  $Q$  para qualquer política, seja ela ótima ou não. A formulação ADHDP necessita somente de amostras da função de recompensa instantânea  $r$  (RÊGO, 2014).

### 3.6 Formulação da Sintonia ADHDP para o caso LQR Discreto

O desenvolvimento do controle ótimo DLQT-I apresentado na Seção 3.4 é baseado no problema do controle ótimo DLQR, logo, a formulação da ADHDP é apresentada para o caso discreto do regulador quadrático linear (LEWIS; VRABIE, 2009). Seja o sistema discreto invariante no tempo descrito em espaço de estados por

$$x_{k+1} = A_d x_k + B_d u_k, \quad (3.76)$$

$$y_k = C_d x_k, \quad (3.77)$$

sendo  $x_k \in R^m$  os estados,  $u_k \in R^n$  a entrada de controle e  $y_k \in R^s$  a saída, e  $k$  o índice de tempo discreto. A ação de controle para realimentação de estados é definida por

$$u_k = h(x_k) = -K x_k, \quad (3.78)$$

sendo  $h(x_k)$  uma política de controle definida como uma função do espaço de estados para o espaço de controle  $h(\cdot) : R^m \rightarrow R^n$ , para cada estado  $x_k$  é definida uma ação de controle  $u_k$ . A noção de controle ótimo surge da definição de um índice de desempenho ou uma função de custo que deve ser minimizada (LEWIS; VRABIE; SYRMOS, 2012), nesse caso apresentada como:

$$V_h(x_k) = \sum_{i=0}^{\infty} \gamma^i r_{k+i} = \sum_{i=k}^{\infty} \gamma^{i-k} r(x_i, u_i) \quad (3.79)$$

sendo  $\gamma$  o fator de desconto ( $0 < \gamma \leq 1$ ) que representa a influência dos custos futuros sobre o custo atual, e  $r(x_k, u_k)$  é o custo instantâneo ou função utilidade que determina o custo de controle para um passo. A Equação (3.79) é o *cost-to-go*, o custo necessário para sair de um estado atual até um estado final por uma determinada política de controle, é definido como o somatório dos custos futuros descontados do tempo atual  $k$  para o horizonte infinito futuro.

O objetivo do controle ótimo é encontrar a política que minimize a função de custo para se obter

$$V^*(x_k) = \min_{u_k} \left( \sum_{i=k}^{\infty} \gamma^{i-k} r(x_i, u_i) \right), \quad (3.80)$$

conhecido como custo ótimo ou valor ótimo. Logo, a política de controle ótima é dada por

$$u_k^* = \arg \min_{u_k} \left( \sum_{i=k}^{\infty} \gamma^{i-k} r(x_i, u_i) \right), \quad (3.81)$$

expandindo a Equação (3.79) a partir da liberação do primeiro termo do somatório ( $i = k$ ), tem-se:

$$V_h(x_k) = r(x_k, u_k) + \gamma \sum_{i=k+1}^{\infty} \gamma^{i-(k+1)} r(x_i, u_i), \quad (3.82)$$

reescrevendo a Equação (3.82) como uma equação a diferença equivalente:

$$V_h(x_k) = r(x_k, u_k) + \gamma V_h(x_{k+1}), \quad V_h(0) = 0. \quad (3.83)$$

Para obter uma política atual  $u_k = h(x_k)$ , resolve-se uma equação a diferença ao invés de um somatório infinito. A Equação (3.83) é conhecida como Equação de Bellman.

O fator de desconto é considerado  $\gamma = 1$ , ou seja, o custo futuro é considerado de forma completa, e o custo instantâneo é parametrizado como uma função quadrática  $r(x_k, u_k) = x_k^T Q x_k + u_k^T R u_k$  com matrizes de ponderação  $Q = Q^T \geq 0$  e  $R = R^T > 0$ . O objetivo do projeto é selecionar o ganho de realimentação de estado  $K$  e a política de controle que minimize o *cost-to-go*  $V_h(x_k) = V_K(x_k)$ , para todos os estados atuais  $x_k$ . Assumindo que o valor ótimo é quadrático no estado, então ele é dado por

$$V^*(x_k) = \frac{1}{2} x_k^T P x_k, \quad (3.84)$$

para uma matriz  $P > 0$  a ser determinada. Dessa forma, substituindo a parametrização de  $r(x_k, u_k)$  e a Equação (3.84) em (3.83), a Equação de Bellman para o DLQR é dada de forma aproximada por

$$x_k^T P x_k = x_k^T Q x_k + u_k^T R u_k + x_{k+1}^T P x_{k+1}, \quad (3.85)$$

substituindo a Equação (3.76) de estado em (3.85), tem-se que

$$x_k^T P x_k = x_k^T Q x_k + u_k^T R u_k + (Ax_k + Bu_k)^T P (Ax_k + Bu_k), \quad (3.86)$$

logo, a função Hamiltoniana para o DLQR é dada por

$$H(x_k, u_k) = x_k^T Q x_k + u_k^T R u_k + (Ax_k + Bu_k)^T P (Ax_k + Bu_k) - x_k^T P x_k, \quad (3.87)$$

a condição de estacionariedade  $\partial H(x_k, u_k) / \partial u_k = 0$  é uma condição essencial para a otimalidade (LEWIS; VRABIE; SYRMOS, 2012). Aplicando essa condição na Equação (3.87) obtemos o controle ótimo:

$$u_k = -Kx_k = -(R + B^T PB)^{-1} B^T PAx_k, \quad (3.88)$$

substituindo a Equação (3.88) em (3.87):

$$A^T PA - P + Q - A^T PB(R + B^T PB)^{-1} B^T PA = 0, \quad (3.89)$$

a Equação (3.89) é não linear em P e conhecida como Equação Algébrica de *Riccati* (ARE) que possui como solução a matriz P. A Equação (3.89) também pode ser vista como a Equação de Bellman para o caso DLQR (LEWIS; VRABIE; SYRMOS, 2012).

### 3.6.1 Q-learning para DLQR

A Equação de Bellman pode ser dada de forma aproximada para o *Q-learning* por

$$Q(x_k, u_k) = r(x_k, u_k) + \gamma V_h(x_{k+1}), \quad (3.90)$$

logo

$$Q_K(x_k, u_k) = x_k^T Q x_k + u_k^T R u_k + x_{k+1}^T P x_{k+1}, \quad (3.91)$$

portanto

$$Q_K(x_k, u_k) = x_k^T Q x_k + u_k^T R u_k + (Ax_k + Bu_k)^T P (Ax_k + Bu_k), \quad (3.92)$$

com P sendo a solução de *Riccati*, a função Q para o DLQR em representação matricial pode ser dada por

$$Q_K(x_k, u_k) = \begin{bmatrix} x_k \\ u_k \end{bmatrix}^T \begin{bmatrix} Q + A^T PA & A^T PB \\ B^T PA & R + B^T PB \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix} \quad (3.93)$$

$$Q_K(x_k, u_k) = z_k^T H z_k. \quad (3.94)$$

A função  $Q_K$  para o DLQR é quadrática em  $(x_k, u_k)$  e escrita utilizando produto de Kronecker como

$$Q_K(x_k, u_k) = \bar{H}^T \bar{z}_k, \quad (3.95)$$

sendo

$$\bar{H} = \text{vec}(H) \quad \text{e} \quad \bar{z}_k = z_k \otimes z_k = \begin{bmatrix} x_k \\ u_k \end{bmatrix} \otimes \begin{bmatrix} x_k \\ u_k \end{bmatrix}, \quad (3.96)$$

sendo  $\bar{z}_k$  o conjunto de base quadrática composto pelos estados e entrada de controle. O termo  $\bar{H}$  é calculado de forma *online* utilizando aprendizado por reforço pela atualização da política de iteração dada por (LEWIS; VRABIE, 2009)

$$\frac{\partial}{\partial u} Q_K(x_k, u) = 0, \quad (3.97)$$

a partir da Equação (3.99), tem-se que

$$Q_K(x_k, u_k) = z_k^T H z_k = \begin{bmatrix} x_k \\ u_k \end{bmatrix}^T \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix}, \quad (3.98)$$

desenvolvendo a Equação (3.98), tem-se que

$$Q_K(x_k, u_k) = x_k^T H_{xx} x_k + u_k^T H_{ux} x_k + x_k^T H_{xu} u_k + u_k^T H_{uu} u_k, \quad (3.99)$$

aplicando na Equação (3.99), a atualização da política de iteração apresentada na Equação (3.97) é representada por:

$$H_{ux} x_k + H_{xu}^T x_k + H_{uu} u_k + H_{uu}^T u_k = 0. \quad (3.100)$$

Pela Equação (3.94), tem-se que  $H_{xu}^T = H_{ux}$  e  $H_{uu}^T = H_{uu}$ . Logo, substituindo as relações na Equação (3.100), obtem-se que

$$2H_{ux} x_k + 2H_{uu} u_k = 0, \quad (3.101)$$

resolvendo para  $u_k$ , a política ótima é dada por

$$u_k = -K x_k = -(H_{uu})^{-1} H_{ux} x_k, \quad (3.102)$$

substituindo os valores de  $H_{uu}$  e  $H_{ux}$ , tem-se

$$u_k = -(B^T P B + R)^{-1} B^T P A x_k. \quad (3.103)$$

## 3.7 Método de Runge-Kutta

Os métodos de *Runge-Kutta* formam uma família de técnicas numéricas explícitas de passo simples para a resolução de equações diferenciais Ordinárias de primeira ordem. Dado um subintervalo  $[x_i, x_{i+1}]$  de uma função  $f(x, y)$ , o valor de  $y_{i+1}$  é dado por

$$y_{i+1} = y_i + \text{Inclinação} \cdot h, \quad (3.104)$$

sendo  $i$  o número de iterações,  $h$  o passo  $h = (x_{i+1} - x_i)$  e a *Inclinação* uma constante obtida pelo cálculo da inclinação em vários pontos da função no interior do intervalo. Os

métodos de *Runge-Kutta* são classificados de acordo com a sua ordem. A ordem indica o número de pontos usados em um subintervalo para determinar o valor da *Inclinação* (GILAT; SUBRAMANIAM, 2009), como apresentado na Figura 3.9.

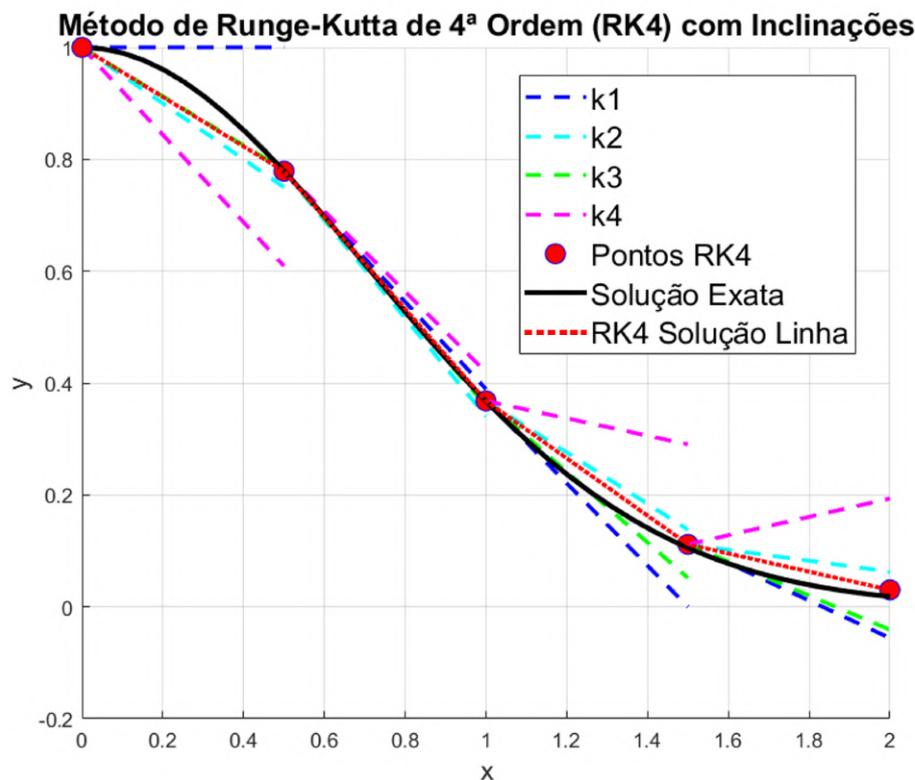


Figura 3.9 – Solução e inclinações do método Runge-Kutta de quarta ordem para a equação diferencial  $y' = -2xy$ , com passo  $h = 0,5$ .

A ordem também indica o erro de truncamento global, por exemplo, o método de segunda ordem possui um erro de truncamento local de  $O(h^3)$  e um erro de truncamento global de  $O(h^2)$ . O erro de truncamento local está relacionado ao erro cometido em cada passo individual do método, o erro global está associado ao erro acumulado ao longo de todos os passos (GILAT; SUBRAMANIAM, 2009).

Esse método é mais preciso do que o método explícito de Euler e sua precisão aumenta à medida que a ordem do método também aumenta. A ordem mais utilizada é quarta ordem devido seu equilíbrio entre precisão e complexidade.

### 3.7.1 Método de Runge-Kutta de quarta ordem

A forma geral do método de *Runge-Kutta* de quarta ordem (RK4) é:

$$y_{i+1} = y_i + (c_1k_1 + c_2k_2 + c_3k_3 + c_4k_4)h, \quad (3.105)$$

sendo

$$k_1 = f(x_i, y_i), \quad (3.106)$$

$$k_2 = f(x_i + a_2h, y_i + b_{21}k_1h), \quad (3.107)$$

$$k_3 = f(x_i + a_3h, y_i + b_{31}k_1h + b_{32}k_2h), \quad (3.108)$$

$$k_4 = f(x_i + a_4h, y_i + b_{41}k_1h + b_{42}k_2h + b_{43}k_3h), \quad (3.109)$$

sendo  $a$ ,  $b$  e  $c$  constantes. Uma das formas de se obter essas constantes é por uma expansão em série de Taylor, sendo a forma geral para o método de *Runge-Kutta*.

$$y_{i+1} = y_i + h \left. \frac{dy}{dx} \right|_{x_i} + \frac{h^2}{(2!)} \left. \frac{d^2y}{dx^2} \right|_{x_i} + \dots + \frac{h^{(k+1)}}{(k+1)!} \left. \frac{d^{(k+1)}y}{dx^{(k+1)}} \right|_{x_i}, \quad (3.110)$$

sendo  $k$  a ordem do método *Runge-Kutta*. As equações (3.106), (3.107), (3.108) e (3.109) também são expandidas em série de Taylor e substituídas na Equação (3.105) então é formado um sistema de Equação com (3.110) onde é obtido as constantes  $a$ ,  $b$  e  $c$ .

Os valores das constante para o método RK4 clássico são:

$$c_1 = c_4 = \frac{1}{6}, \quad c_2 = c_3 = \frac{2}{6}, \quad a_2 = a_3 = b_{21} = b_{32} = \frac{1}{2}, \quad (3.111)$$

$$a_4 = b_{43} = 1, \quad b_{31} = b_{41} = b_{42} = 0, \quad (3.112)$$

substituindo os valores das constantes, tem-se:

$$y_{i+1} = y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)h, \quad (3.113)$$

sendo

$$k_1 = f(x_i, y_i), \quad (3.114)$$

$$k_2 = f(x_i + \frac{1}{2}h, y_i + \frac{1}{2}k_1h), \quad (3.115)$$

$$k_3 = f(x_i + \frac{1}{2}h, y_i + \frac{1}{2}k_2h), \quad (3.116)$$

$$k_4 = f(x_i + h, y_i + k_3h). \quad (3.117)$$

### 3.8 Controlador Adaptativo PID-LMS

O controlador PID é a principal técnica de controle para aplicações industriais. Diferentes tipos de regras de sintonia são aplicadas para esse controlador (OGATA, 2011), entretanto grande parte dessas sintonias são baseadas em análises experimentais. A sua estrutura padrão é baseada em ganhos fixos, o que causa uma limitação quanto à sua aplicação.

O controlador adaptativo PID-LMS, possui ajuste de ganho do controlador PID baseado em uma estrutura de rede neural do tipo perceptron. O sinal de erro é pré-processado para a entrada do combinador linear que representa a estrutura de ganho PID. Os ganhos são ajustados por uma unidade neural treinada pelo método LMS (NUNES et al., 2007).

O pré-processamento do erro é a entrada do neurônio e consiste no próprio erro, sua integral e sua derivada, representado pelos sinais  $x_P$ ,  $x_I$  e  $x_D$  respectivamente. Esses sinais são ponderados pelos ganhos do controlador. A saída da rede neural é uma combinação linear entre os sinais pré processados e os pesos (ABREU; NETO, 2019), como apresentado na Figura 3.10.

A atualização dos ganhos do controlador baseado no algoritmo LMS é dado por (NUNES et al., 2007)

$$\hat{\omega}_{PID}(k+1) = \hat{\omega}_{PID}(k) + \eta_{PID}e(k)\hat{x}_{PID}(k), \quad (3.118)$$

sendo  $\hat{\omega}_{PID}(k)$  os ganhos estimados,  $\eta_{PID}$  a taxa de aprendizado,  $e$  o erro do sistema e  $\hat{x}_{PID}(k)$  os sinais de erro, com sua integral e derivada. A estrutura do controlador é apresentada na Figura 3.10.

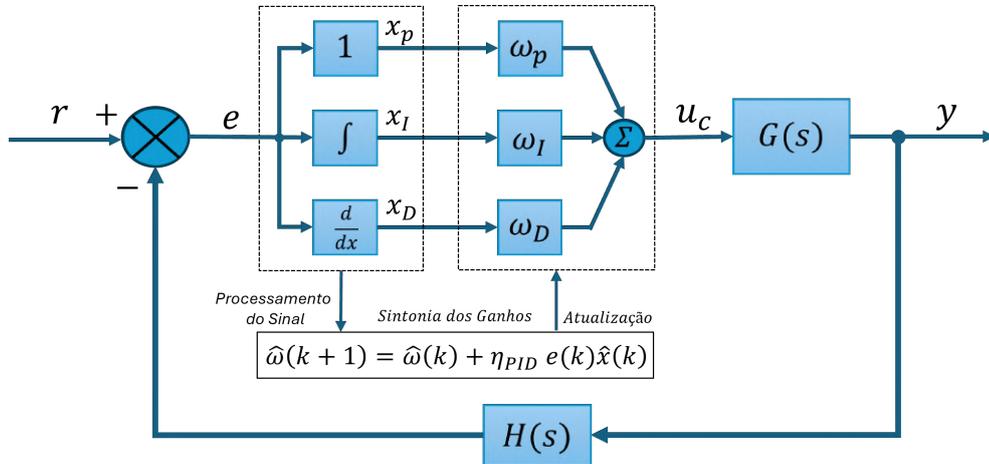


Figura 3.10 – Diagrama de blocos do sistema de controle PID-LMS.

No diagrama de blocos apresentado na Figura 3.10,  $r$  é a referência,  $e$  o erro,  $G(s)$  a planta controlada,  $y$  a saída,  $H(s)$  a realimentação e  $u_c$  é a saída do combinador linear, isto é, o sinal de controle dado por

$$u_c(k) = \hat{\omega}_P(k)x_P(k) + \hat{\omega}_I(k)x_I(k) + \hat{\omega}_D(k)x_D(k), \quad (3.119)$$

sendo  $\hat{\omega}_P(k)$ ,  $\hat{\omega}_I(k)$  e  $\hat{\omega}_D(k)$  os pesos da rede ou nesse caso os ganhos do controlador PID.

## CAPÍTULO 4

# PROJETO DO CONTROLADOR ATOR-CRÍTICO PARA MANOBRABILIDADE

Nesse Capítulo é apresentado o modelo matemático do CyberShip I para três graus de liberdade e deslocamento horizontal. A formulação do controlador ADHDP-DLQT-I e seu algoritmo, o simulador do modelo não linear baseado em *Runge-Kutta* de quarta ordem e o desacoplamento do modelo para implementação PID.

### 4.1 Modelo Matemático do *CyberShip I*

O modelo matemático utilizado para a implementação do controlador e simulações computacionais é o modelo do *CyberShip I*, desenvolvido pelo Departamento de Engenharia Cibernética da *Norwegian University of Science and technology* (NTNU) em Trondheim - Noruega. Esse modelo é a representação de um navio *supply offshore* controlado por propulsores, criado na escala de 1:70, com massa  $m = 17,6$  kg e comprimento  $L = 1,19$  m.

O modelo não linear do *CyberShip I* para deslocamento horizontal é descrito pela Equação (3.15). Utilizando métodos hidrodinâmicos e identificação de sistemas foram determinados os parâmetros do modelo para três graus de liberdade. Os parâmetros são apresentados em: McGookin (1997), Strand (1999) e Cid (2003). Esse modelo também foi adotado em: Pettersen e Fossen (1998), Strand (1999), Alfaro-Cid, McGookin e Murray-Smith (2001), Jeong et al. (2010), Tomera e Podgórski (2021) e Zhong et al. (2023).

Os parâmetros do CyberShip I são:

Tabela 4.1 – Parâmetros do Cybersip I

Parâmetro	Valor	Unidade(SI)
$m$	17,6	kg
$I_z$	1,80	kg·m <sup>2</sup>
$x_g$	-0,04	m
$X_{\dot{u}}$	-1,40	kg
$Y_{\dot{v}}$	-17,6	kg
$N_{\dot{r}}$	-0,18	kg·m <sup>2</sup>
$Y_{\dot{r}}$	0	kg·m
$N_{\dot{v}}$	0	kg·m
$X_u$	-4,0	kg/s
$Y_v$	-6,0	kg/s
$N_r$	-1,0	kg·m <sup>2</sup> /s

Substituindo os valores dos parâmetros apresentados na Tabela 4.1, nas Equações (3.16), (3.17) e (3.18), tem-se:

$$M = \begin{bmatrix} 19 & 0 & 0 \\ 0 & 35,2 & -0,704 \\ 0 & -0,704 & 1,98 \end{bmatrix}, C(\nu) = \begin{bmatrix} 0 & 0 & -35,2\nu \\ 0 & 0 & 19u \\ 35,2\nu & -19u & 0 \end{bmatrix}, \text{Damp} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.1)$$

Um segundo modelo, denominado  $x_{g0}$ , também é utilizado nas implementações do controlador. Esse modelo consiste em uma adaptação, apresentada em McGookin (1997). Como pode ser observado pela substituição das matrizes da Equação (4.1) na Equação (3.15), há um acoplamento entre as acelerações e velocidades. A adaptação altera levemente a dinâmica do sistema e causa o desacoplamento existente entre as velocidades  $v$  e  $r$  no modelo linear apresentado na Seção 4.1.1, pois, a matriz  $M$  se torna diagonal e dada por

$$M = \begin{bmatrix} 19 & 0 & 0 \\ 0 & 35,2 & 0 \\ 0 & 0 & 1,98 \end{bmatrix}. \quad (4.2)$$

Essa alteração consiste em aproximar CG e CO como pontos coincidentes no espaço, e conseqüentemente torna a distância  $x_g$  igual a zero. Essa abordagem é essencial para a implementação do método RK4 no modelo não linear, onde é necessário obter cada Equação de estado da dinâmica de modo independente.

#### 4.1.1 A Representação do Modelo Linear em Espaço de Estados

O modelo linear do USV surge da escolha das matrizes que representam sua dinâmica e da consideração dos efeitos para aplicação, como o caso do efeito Coriolis. Especificamente

para aplicações como Posicionamento Dinâmico (PD), pequenos deslocamentos e manobras em baixa velocidade relativa, o amortecimento é considerado linear e o efeito Coriolis é desprezado. Nesse caso, o modelo então é reduzido para (FOSSSEN, 2011):

$$\mathbf{M}\dot{\nu} + \mathbf{Damp}\nu = \tau, \quad (4.3)$$

sabendo-se que a forma canônica da representação em espaço de estados é dada por:

$$\dot{x}(t) = Ax(t) + Bu(t), \quad (4.4)$$

reorganizando a Equação (4.3), tem-se:

$$\dot{\nu} = -\mathbf{M}^{-1}\mathbf{Damp}\nu + \mathbf{M}^{-1}\tau, \quad (4.5)$$

por analogia entre as equações (4.3) e (4.4), tem-se que  $x(t) = \nu$ ,  $u(t) = \tau$ ,  $A = -\mathbf{M}^{-1}\mathbf{Damp}$  e  $B = \mathbf{M}^{-1}$ . A equação de saída é dada por

$$y(t) = Cx(t), \quad (4.6)$$

sendo  $y(t)$  os valores de saída do sistema, que são as próprias velocidades dadas pelo vetor de estados, logo, a matriz  $C$  é a matriz identidade  $I$  e a matriz de transmissão  $D = 0$ .

## 4.2 Controle ADHDP-DLQT com ação integral

A planta do USV é do tipo 0, isto é, não possui integrador implícito ao modelo, conseqüentemente, apresenta erro de regime em estado estacionário para implementações somente com ganho proporcional como o caso do DLQT, por esses motivos, o projeto é desenvolvido com uma parcela de ação integral. Seja o sistema dinâmico representado pelas Equações (4.5) e (4.6), um novo sistema é formulado a partir das matrizes do sistema, como apresentado na Seção (3.4). A implementação do DLQT-I é formulado como um problema de DLQR na ADHDP, como apresentado na Seção (3.6).

A partir da formulação apresentada na Seção (3.4), tem-se que a lei de controle do controlador ADHDP-DLQT-I é dada por

$$u_k = u_{k-1} + K_{ADHDP}\tilde{x}_k, \quad (4.7)$$

sendo  $K_{ADHDP} = [K_{ADHDP1} \ K_{ADHDP2}]$  o ganho calculado de forma *online*, composto pelo ganho de realimentação de estados  $K_1$  e o ganho de realimentação de saída  $K_2$ . Logo, a nova lei de controle é dada por

$$u_k = u_{k-1} + K_{ADHDP1}(x_k - x_{k-1}) + K_{ADHDP2}(y_{k-1} - r_k), \quad (4.8)$$

sendo saída  $y_k$  do sistema obtida por meio da seleção dos estados de interesse que é dada pela equação

$$y_k = C_d x_k, \quad (4.9)$$

sendo  $C_d$  a matriz  $C$  discretizada. O cálculo dos ganhos pela ADHDP é reformulado para a expansão de um novo vetor de estados e da matriz  $H$ , dado por

$$K_{ADHDP}(\tilde{H}) = -\tilde{H}_{uu}^{-1} \tilde{H}_{ux}, \quad (4.10)$$

sendo  $\tilde{H}$  a matriz expandida de dimensão  $(n+r) \times (n+r)$ ,  $\tilde{H}_{uu}$  e  $\tilde{H}_{ux}$  as matrizes blocadas da matriz de aprendizado  $\tilde{H}$ . A função valorada é dada por

$$\tilde{Q}_k(\tilde{x}_k, u_k) = \tilde{z}_k^T \tilde{H} \tilde{z}_k, \quad (4.11)$$

sendo  $\tilde{z}_k^T = [\tilde{x}_k^T \ u_k^T]$  e  $\tilde{x}_k^T = [\Delta x_k^T \ (y_{k-1} - r)^T]$ . Para a implementação da ADHDP e cálculo da matriz  $\tilde{H}$ , a Equação (3.85) é parametrizada como um vetor alvo (*target*) (NETO; LOPES, 2011), representado por

$$d_{tar} = \tilde{x}_k^T \tilde{Q} \tilde{x}_k + u_k^T R u_k + w^T \tilde{H} w, \quad (4.12)$$

sendo  $w = [\tilde{x}_{k+1} \ K_{ADHDP} \tilde{x}_{k+1}]$ . O vetor base como produto de Kronecker  $\tilde{z}$ , é remodelado para uma matriz de produto de Kronecker sem redundâncias, representado por  $\tilde{\tilde{z}}$ . Logo, a aproximação da solução da DARE é dada por

$$\tilde{\tilde{z}} \cdot vH = d_{tar}, \quad (4.13)$$

sendo  $vH$  a vetorização de  $\tilde{H}$ . A remoção das redundâncias do produto de Kronecker, reduz as dimensões de  $\tilde{\tilde{z}}$  e consequentemente de  $d_{tar}$  e  $vH$ , a nova dimensão é dada por

$$n_{kron} = \frac{(\tilde{n} + m)(\tilde{n} + m + 1)}{2} = \frac{(6 + 3)(6 + 3 + 1)}{2} = 45, \quad (4.14)$$

sendo  $\tilde{n}$  a quantidade de estados do sistema expandido e  $m$ , a quantidade de entradas. O sistema remodelado possui 6 estados pois o modelo pré-expansão possui 3 estados e 3 saídas. Logo,  $\tilde{\tilde{z}}$  é uma matriz  $45 \times 45$ , e  $d_{tar}$  e  $vH$  são vetores coluna de 45 posições.

Realizando as manipulações dos elementos utilizando a teoria para vetorização e produto de Kronecker, a solução de *Riccati* é aproximada por um sistema de equações lineares para os coeficientes da matriz  $\tilde{H}$ , que possui como solução:

$$vH = \tilde{\tilde{z}}^{-1} d_{tar}. \quad (4.15)$$



## ALGORITMO PRINCIPAL - ADHDP-DLQT-I

## ▷ Setup - Inicialização dos Parâmetros

- 
- 1 ▷ Inicialização dos parâmetros de simulação
  - 2  $(T_s, t_s, N, h, N_{it}) \leftarrow \text{PARAM\_SIM}$
  - 3 ▷ Definir modelo linear do USV
  - 4  $(A, B, C, D) \leftarrow \text{MODELO\_USV}$
  - 5 ▷ Discretização do sistema
  - 6  $(A_d, B_d, C_d, D_d) \leftarrow \text{C2D}(A, B, C, D, T_s, \text{'zoh'})$
  - 7 ▷ Inicialização dos parâmetros ADHDP, DLQR e condições iniciais
  - 8  $(\tilde{H}, K_{ADHDP}, Q, R, u_0, x_0, y_0) \leftarrow \text{INIC}$
- 

## ▷ Processo Iterativo

- 
- 9 **Para**  $i$  de 1 até  $N_{it}$  **Faça**
  - 10 ▷ Implementação do ruído de controle
  - 11  $e_i(:, i) \leftarrow [ \ ]$
  - 12 ▷ Ação de Controle
  - 13  $u(:, i) \leftarrow u(:, i-1) + K_1[x(1:3, i) - x(1:3, i-1)] + K_2[y(:, i-1) - r] + e_i(:, i)$
  - 14 ▷ Simulação da dinâmica do sistema
  - 15  $x(:, i+1) \leftarrow A_d x(:, i) + B_d u(:, i)$
  - 16  $y(:, i) \leftarrow C_d x(:, i)$
  - 17 ▷ Montagem dos vetores auxiliares
  - 18  $x_2(:, i) \leftarrow [(x(:, i) - x(:, i-1)) ; (y(:, i-1) - r)]$
  - 19  $x_2(:, i+1) \leftarrow [(x(:, i+1) - x(:, i)) ; (y(:, i) - r)]$
  - 20  $u_2(:, i) \leftarrow [u(:, i) - u(:, i-1)]$
  - 21 ▷ Montagem do vetor alvo
  - 22  $d_{tar}(1 : (n_{kron} - 1), :) \leftarrow d_{tar}(2 : n_{kron}, :)$
  - 23  $d_{tar}(n_{kron}, 1) \leftarrow [y(:, i-1) - r]^T Q [y(:, i-1) - r] + [u_2(:, i)]^T R [u_2(:, i)] + \dots$   
 $\dots [x_2(:, i+1); K_{ADHDP} x_2(:, i+1)]^T \tilde{H} [x_2(:, i+1); K_{ADHDP} x_2(:, i+1)]$
  - 24 ▷ Conjunto Base - produto de Kronecker
  - 25  $\tilde{z}(:, 1 : (n_{kron} - 1)) \leftarrow \tilde{z}(:, 2 : (n_{kron}))$
  - 26  $\tilde{z}(:, n_{kron}) \leftarrow [x_2(1, i)^2; x_2(1, i)x_2(2, i); \dots; x_2(n, i)u_2(1, i); \dots; u_2(m, i)^2]$
  - 27  $i_k \leftarrow i_k + 1$
  - 28 **Se**  $i_k = n_{kron}$  **Então**
  - 29  $vH \leftarrow (\tilde{z}^T)^{-1} d_{tar}$
- 

▷ Montagem da Matriz  $\tilde{H}$ 

---


$$30 \quad \tilde{H} \leftarrow \begin{bmatrix} vH_1 & vH_2/2 & vH_3/2 & \dots & vH_{np}/2 \\ vH_2/2 & vH_{np+1} & vH_{np+2}/2 & \dots & vH_{2np-1}/2 \\ vH_3/2 & vH_{np+2} & vH_{2np}/2 & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & vH_{nkron-1}/2 \\ vH_{np}/2 & vH_{2np-1}/2 & \dots & vH_{nkron-1}/2 & vH_{nkron} \end{bmatrix}$$


---

## ▷ Atualização do Ganho

- 
- 31 ▷ Ganho ótimo  $K_{ADHDP}$
  - 32  $\tilde{H}_{ux} \leftarrow \tilde{H}(m+1 : m+n, 1 : m)$
  - 33  $H_{uu} \leftarrow \tilde{H}(m+1 : m+n, m+1 : m+n)$
  - 34  $K_{ADHDP} \leftarrow -(\tilde{H}_{uu})^{-1}(\tilde{H}_{ux})$
  - 35  $i_k \leftarrow 0$
- 

## ▷ Fim do Processo Iterativo

---

O setup do algoritmo é composto por subrotinas de inicialização que são chamadas pelo Algoritmo Principal (AP). Na linha 2, a subrotina PARAM\_SIM define os parâmetros da simulação intervalo de amostragem  $T_s$ , tempo de simulação  $t_f$ , o vetor de tempo  $t$  e o número de iterações  $N_{it}$ . A subrotina é apresentada abaixo

---

PARAM\_SIM

---

- 1  $T_s \leftarrow [ \ ]$
  - 2  $t_f \leftarrow [ \ ]$
  - 3  $t \leftarrow 0 : T_s : t_f$
  - 4  $N_{it} \leftarrow \text{dimensão de } t$
- 

Na linha 4 do AP, a subrotina MODELO\_USV estabelece as matrizes de espaço de estados do modelo linear do USV. A definição é realizada substituindo os valores dos parâmetros e calculando as matrizes como apresentado na Seção.

---

MODELO\_USV

---

```

1  ▷ Inicialização dos Parâmetros do USV
2   $m \leftarrow 17.6$ 
3   $I_z \leftarrow 1.80$ 
4   $X_{uu} \leftarrow -1.40$ 
5   $Y_{vv} \leftarrow -17.60$ 
6   $N_{rr} \leftarrow -0.18$ 
7   $x_g \leftarrow -0.04$ 
8   $Y_{rr} \leftarrow 0$ 
9   $N_{vv} \leftarrow 0$ 
10  $X_u \leftarrow -4$ 
11  $Y_v \leftarrow -6$ 
12  $N_r \leftarrow -1$ 
13 ▷ Definição da matriz de inércia
14  $M \leftarrow \begin{bmatrix} m - X_{uu} & 0 & 0 \\ 0 & m - Y_{vv} & m \cdot x_g \\ 0 & m \cdot x_g & I_z - N_{rr} \end{bmatrix}$ 
15 ▷ Definição da matriz de amortecimento
16  $Damp \leftarrow \begin{bmatrix} -X_u & 0 & 0 \\ 0 & -Y_v & 0 \\ 0 & 0 & -N_r \end{bmatrix}$ 
17 ▷ Construção das matrizes do modelo
18  $A \leftarrow -M^{-1} \cdot Damp$ 
19  $B \leftarrow M^{-1}$ 
20  $C \leftarrow I$ 
21  $D \leftarrow 0$ 

```

---

Na linha 6 do AP as matrizes do sistema são discretizadas pelo método ZOH com tempo de amostragem  $T_s$ . A subrotina INIC na linha 8 do AP inicializa parâmetros da ADHDP, matrizes de ponderação do DLQT-I e condições iniciais.

---

 INIC
 

---

- 1 ▷ Condições iniciais
  - 2  $x_0 \leftarrow [ \quad ], u_0 \leftarrow [ \quad ], y_0 \leftarrow [ \quad ]$
  - 3 ▷ Matrizes de ponderação
  - 4  $Q \leftarrow [ \quad ], R \leftarrow [ \quad ]$
  - 5 ▷ Cálculo da dimensão do conjunto base
  - 6  $n \leftarrow$  quantidade de estados
  - 7  $m \leftarrow$  quantidade de entradas
  - 8  $n_p \leftarrow m + n$
  - 9  $n_{kron} \leftarrow \frac{n_p(n_p+1)}{2}$
  - 10 ▷ Inicialização da matriz  $\tilde{H}$
  - 11  $\tilde{H}_0 \leftarrow [ \quad ]$
  - 12 ▷ Inicialização do Ganho
  - 13  $K_{0ADHDP} \leftarrow -(\tilde{H}_{uu})^{-1}\tilde{H}_{ux}$
- 

**Processo Iterativo:**

O processo iterativo do AP, inicia com a implementação do ruído de controle  $e_i$  na linha 13, é utilizado um ruído branco gaussiano na ação de controle por questões de excitação persistente do método. Na linha 13 é definida a lei de controle, durante as primeiras iterações, os ganhos são inicializados com zero, tal que, o valor da ação de controle será determinada somente pelo ruído, até o cálculo do primeiro valor dos ganhos. Nas linhas 15 e 16, a dinâmica do sistema é simulada. As linhas 18 a 20 apresentam a montagem dos vetores auxiliares, um artifício para simplificar a implementação nos cálculos de  $d_{tar}$  e  $\tilde{z}$ . Na linha 22 os valores do vetor alvo calculados anteriormente são deslocados para uma linha acima, após o deslocamento, é calculado o valor para a última posição do vetor como pode ser visto na linha 23 do AP. Na linha 25, é realizado um procedimento com  $\tilde{z}$  análogo ao efetuado anteriormente com  $d_{tar}$ , entretanto, por se tratar de uma matriz, os valores das colunas, são deslocados uma coluna à esquerda. Na linha 24 é calculado o valor atual do produto de Kronecker e armazenado na última coluna da matriz  $\tilde{z}$ .

**Atualização dos ganhos:**

Na linha 27, há o incremento unitário do índice  $i_k$ , na linha 28 é realizada uma comparação de  $i_k$  com  $n_{kron}$ , essa condição indica que o vetor  $d_{tar}$  e a matriz  $\tilde{z}$  foram completamente preenchidas para atualização dos ganhos. Logo, é inicializado o processo para cálculo da matriz  $\tilde{H}$  e atualização dos Ganhos. Na linha 29 é realizado o cálculo de

$vH$ , e na linha 30 a construção da matriz  $\widetilde{H}$  a partir de  $vH$ . Nas linhas 32 e 33 as matrizes  $\widetilde{H}_{ux}$  e  $\widetilde{H}_{uu}$  são extraídas de  $\widetilde{H}$ . Na linha 34 é realizado o cálculo dos novos ganhos para a sua atualização. Na linha 35 o índice, que determina quando será a próxima atualização dos ganhos, é zerado e o processo iterativo é reinicializado até a sua condição de parada  $i = N_{it}$

### 4.3 Simulador Não Linear

A avaliação do desempenho do controlador é realizada com o intuito de obter a resposta mais fiel à dinâmica não-linear do USV por meio de um controlador baseado no modelo linearizado. A dinâmica aproximada é simulada pelo modelo não linear do sistema, considerando o efeito Coriolis da matriz dependente de estado  $\mathbf{C}(v)$  (Equação 3.17).

Em paralelo à sintonia e simulação do modelo linear, a ação de controle calculada a cada instante do algoritmo ADHDP-DLQT-I é enviado para um simulador da dinâmica não linear do USV. O simulador utiliza o esforço de controle do modelo linear para avaliar a resposta do sistema não linear resolvido pelo método RK4, implementado manualmente.

A implementação do método RK4 necessita que as equações de estado sejam desacopladas, nesse caso é utilizado o modelo do CyberShip I  $xg_0$ . Desenvolvendo as equações de estado da dinâmica não-linear, para o modelo  $xg_0$ , tem-se

$$m_{11}\dot{u} = \frac{1}{m_{11}}(m_{22}vr + m_{23}r^2 - d_{11}u + \tau_1), \quad (4.16)$$

$$m_{22}\dot{v} = \frac{1}{m_{22}}(-m_{11}ur - d_{22}v + \tau_2), \quad (4.17)$$

$$m_{33}\dot{r} = \frac{1}{m_{33}}((m_{11} - m_{22})uv - m_{23}ur - d_{33}r + \tau_3), \quad (4.18)$$

sendo a solução para a velocidade  $u$  dada por

$$\begin{aligned} k_{u1} &= h \left[ \frac{1}{m_{11}} \left( m_{22}vr + m_{23}r^2 - d_{11}u + \tau_1 \right) \right], \\ k_{u2} &= h \left[ \frac{1}{m_{11}} \left( m_{22} \left( v + \frac{h}{2} \right) \left( r + \frac{h}{2} \right) + m_{23} \left( r + \frac{h}{2} \right)^2 - d_{11} \left( u + \frac{k_{u1}}{2} \right) + \tau_1 \right) \right], \\ k_{u3} &= h \left[ \frac{1}{m_{11}} \left( m_{22} \left( v + \frac{h}{2} \right) \left( r + \frac{h}{2} \right) + m_{23} \left( r + \frac{h}{2} \right)^2 - d_{11} \left( u + \frac{k_{u2}}{2} \right) + \tau_1 \right) \right], \\ k_{u4} &= h \left[ \frac{1}{m_{11}} \left( m_{22} (v + h) (r + h) + m_{23} (r + h)^2 - d_{11} (u + k_{u3}) + \tau_1 \right) \right], \\ u_{k+1} &= u + \frac{1}{6}(k_{u1} + 2k_{u2} + 2k_{u3} + k_{u4}), \end{aligned} \quad (4.19)$$

a solução para a velocidade  $v$  dada por

$$\begin{aligned}
k_{v1} &= h \left[ \frac{1}{m_{22}} (-m_{11}ur - d_{22}v + u_2) \right] \\
k_{v2} &= h \left[ \frac{1}{m_{22}} \left( -m_{11} \left( u + \frac{h}{2} \right) \left( r + \frac{h}{2} \right) - d_{22} \left( v + \frac{k_{v1}}{2} \right) + u_2 \right) \right] \\
k_{v3} &= h \left[ \frac{1}{m_{22}} \left( -m_{11} \left( u + \frac{h}{2} \right) \left( r + \frac{h}{2} \right) - d_{22} \left( v + \frac{k_{v2}}{2} \right) + u_2 \right) \right] \\
k_{v4} &= h \left[ \frac{1}{m_{22}} (-m_{11} (u + h) (r + h) - d_{22} (v + k_{v3}) + u_2) \right] \\
v_{k+1} &= v + \frac{1}{6} (k_{v1} + 2k_{v2} + 2k_{v3} + k_{v4}), \tag{4.20}
\end{aligned}$$

a solução para a velocidade  $r$  dada por

$$\begin{aligned}
k_{r1} &= h \left[ \frac{1}{m_{33}} ((m_{11} - m_{22})uv - m_{23}ur - d_{33}r + u_3) \right] \\
k_{r2} &= h \left[ \frac{1}{m_{33}} \left( (m_{11} - m_{22}) \left( u + \frac{h}{2} \right) \left( v + \frac{h}{2} \right) - m_{23} \left( u + \frac{h}{2} \right) \left( r + \frac{k_{r1}}{2} \right) - d_{33} \left( r + \frac{k_{r1}}{2} \right) + u_3 \right) \right] \\
k_{r3} &= h \left[ \frac{1}{m_{33}} \left( (m_{11} - m_{22}) \left( u + \frac{h}{2} \right) \left( v + \frac{h}{2} \right) - m_{23} \left( u + \frac{h}{2} \right) \left( r + \frac{k_{r2}}{2} \right) - d_{33} \left( r + \frac{k_{r2}}{2} \right) + u_3 \right) \right] \\
k_{r4} &= h \left[ \frac{1}{m_{33}} ((m_{11} - m_{22}) (u + h) (v + h) - m_{23} (u + h) (r + k_{r3}) - d_{33} (r + k_{r3}) + u_3) \right] \\
r_{k+1} &= r + \frac{1}{6} (k_{r1} + 2k_{r2} + 2k_{r3} + k_{r4}), \tag{4.21}
\end{aligned}$$

sendo as Equações (4.19), (4.20) e (4.21) o valor dos estados  $u$ ,  $v$  e  $r$ , respectivamente, no próximo instante de tempo  $k$ ,  $u_1$ ,  $u_2$  e  $u_3$  as entradas de controle calculadas no Algoritmo ADHDP-DLQT-I, e  $h$  o tamanho do passo. A solução é desenvolvida conforme apresentado na Seção 3.7. Os valores dos parâmetros do modelo  $x_{g0}$  são apresentados na Seção 4.1. O algoritmo implementado é comparado com a função `ode45` do MATLAB®, que utiliza o método RK4 em comparação ao RK5 para uma diminuição da quantidade de iterações por meio de um tamanho de passo  $h$  variável.

## 4.4 Projeto do Controlador PID Adaptativo LMS Desacoplado

O controlador PID Adaptativo LMS é desenvolvido para comparação dos resultados obtidos pelo ADHDP-DLQT-I. Esse método é aplicado nas duas plantas utilizadas. Para o modelo  $x_{g0}$  não há dificuldades pois o modelo é desacoplado, entretanto, para aplicar no modelo acoplado é necessário realizar o desacoplamento da planta.

A matriz de Funções de Transferência do modelo  $x_{g0}$  é dada por

$$G_{x_{g0}}(s) = \begin{bmatrix} \frac{0.05263}{s+0.2105} & 0 & 0 \\ 0 & \frac{0.02841}{s+0.1705} & 0 \\ 0 & 0 & \frac{0.5051}{s+0.5051} \end{bmatrix}. \tag{4.22}$$

A implementação do controlador em diagrama de blocos é apresentada na Figura 4.7.

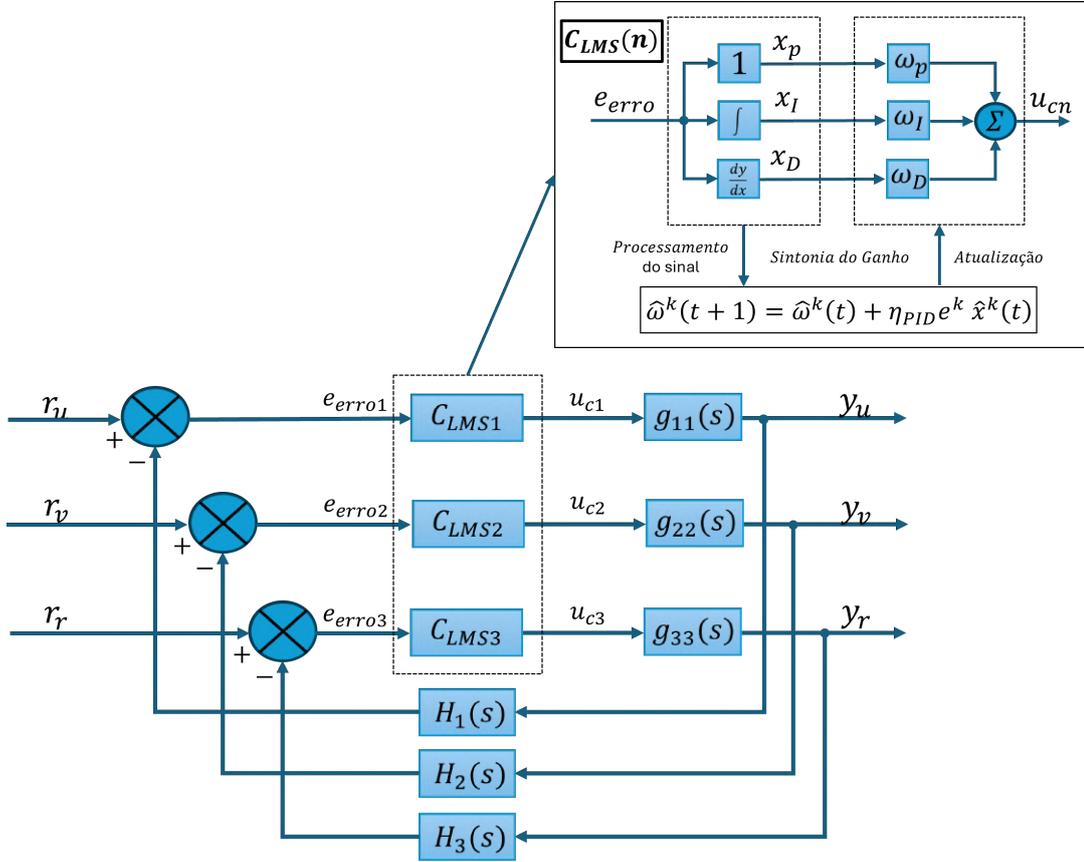


Figura 4.2 – Diagrama de blocos do sistema MIMO em malha fechada com controlador PID-LMS.

Os elementos  $g_{11}(s)$ ,  $g_{22}(s)$  e  $g_{33}$  apresentados na Figura 4.2 são as Funções de Transferência para cada velocidade,  $r_u$ ,  $r_v$  e  $r_r$  são as referências e  $y_u$ ,  $y_v$  e  $y_r$  são velocidades de saída. As realimentações  $H(s)$  são consideradas unitárias.

## 4.5 Projeto do Controlador PID Adaptativo LMS com Desacoplamento

A matriz de função de transferência do modelo completo apresentada na Seção 4.1.1 é dada por

$$G(S) = \begin{bmatrix} \frac{0.0526}{s + 0.2104} & 0 & 0 \\ 0 & \frac{0.0284s + 0.01424}{s^2 + 0.6755s + 0.0854} & \frac{0.0101s}{s^2 + 0.6755s + 0.0867} \\ 0 & \frac{0.0101s}{s^2 + 0.6755s + 0.0867} & \frac{0.5051s + 0.0854}{s^2 + 0.6755s + 0.0854} \end{bmatrix}. \quad (4.23)$$

O diagrama de blocos do sistema a partir da matriz de funções de transferência é apresentado na Figura 4.3.

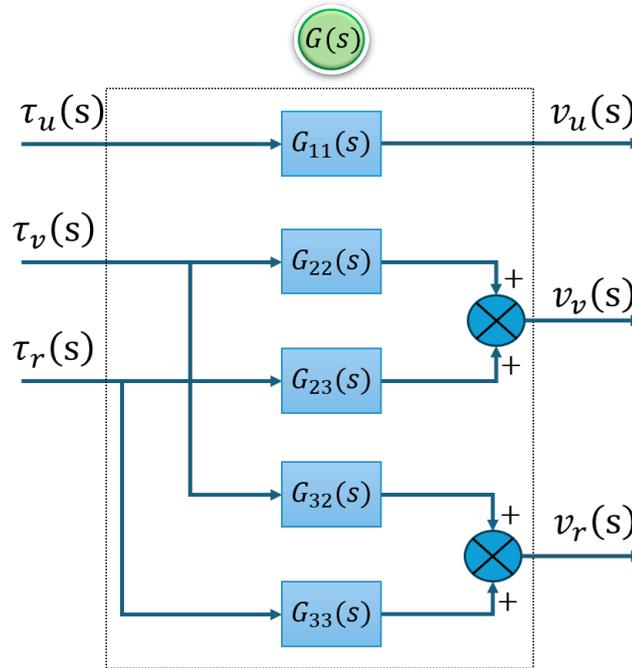


Figura 4.3 – Diagrama de blocos da matriz de função de transferência do modelo matemático completo do *Cybership I*.

As entradas dos blocos apresentados na Figura 4.3 são as forças geradas pelos propulsores, e as saídas são as três velocidades  $u$ ,  $v$  e  $r$ . Observa-se na Figura 4.3 que há um acoplamento entre a entrada  $\tau_v$  e a saída  $v_r$ , assim como entre a entrada  $\tau_r$  e a saída  $v_v$  causada pelos termos  $G_{23}$  e  $G_{32}$ . Esse acoplamento causa uma influência da entrada 2 na saída 3, e entre a entrada 3 e a saída 2, dessa forma, o PID não pode ser aplicado diretamente, visto que este foi projetado para sistemas de entrada e saída únicas (SISO). Logo, para aplicar um PID é necessário realizar o desacoplamento do sistema. Os métodos de desacoplamento visam reduzir os efeitos que uma entrada exerce sobre outra saída, possibilitando uma análise mais simplificada do problema ao considerar cada conjunto de entradas e saídas de forma independente como três sistemas SISO's (PAULO, 2015). Na Figura 4.4 é apresentado um diagrama de bloco para o esquema de controle com desacoplamento e realimentação unitária.

O desacoplamento do sistema consiste em inserir um novo bloco entre o controlador e o processo, de modo que a interação entre o desacoplador e o processo compõe um novo bloco com o processo desacoplado, como pode ser visto na Figura 4.4. A Figura 4.5 apresenta um exemplo de um sistema acoplado de duas entradas e duas saídas (*Two-Input, Two-Output* - TITO).

O controlador, desacoplador e a planta são representados pelas matrizes  $C(s)$ ,  $D(s)$  e  $G(s)$ , respectivamente. A matriz do processo desacoplado, formada pelos blocos do

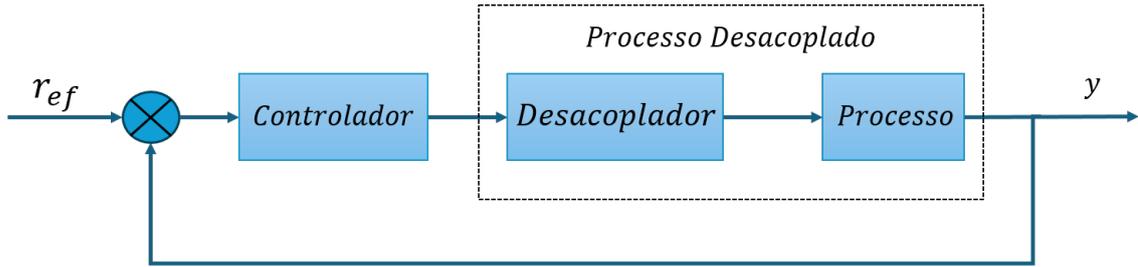


Figura 4.4 – Diagrama de bloco de um controlador com desacoplamento do processo.

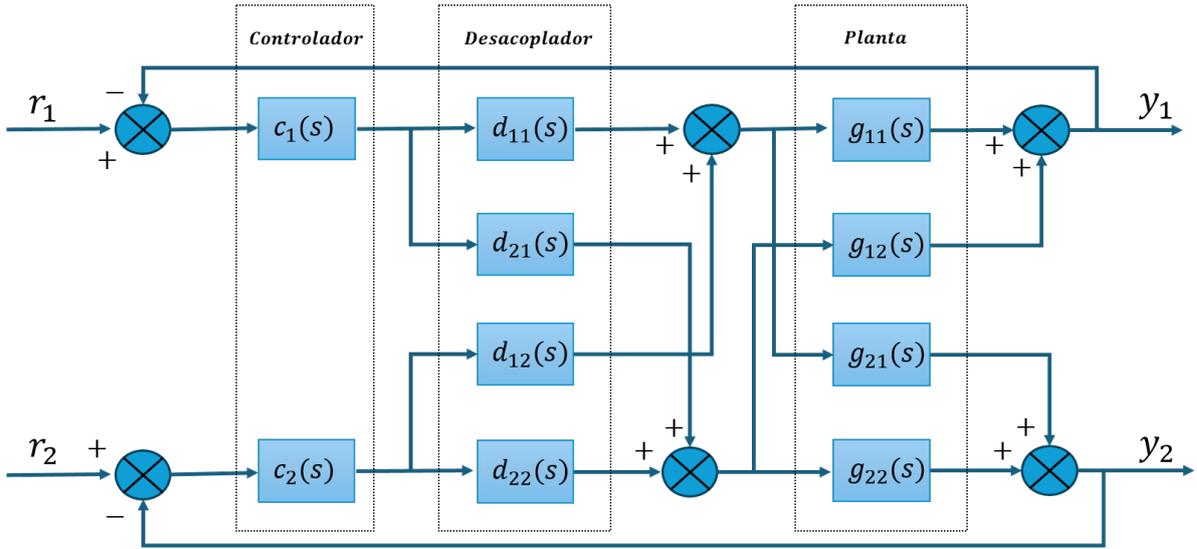


Figura 4.5 – Diagrama de blocos de um sistema TITO com desacoplamento do processo.

desacoplador e processo (Figura 4.4), é representada por  $Q(s)$ . Essas matrizes são dadas por

$$D(s) = \begin{bmatrix} d_{11}(s) & d_{12}(s) \\ d_{21}(s) & d_{22}(s) \end{bmatrix}, \quad C(s) = \begin{bmatrix} c_{11}(s) & 0 \\ 0 & c_{22}(s) \end{bmatrix},$$

$$G(s) = \begin{bmatrix} g_{11}(s) & g_{12}(s) \\ g_{21}(s) & g_{22}(s) \end{bmatrix}, \quad Q(s) = \begin{bmatrix} q_{11}(s) & 0 \\ 0 & q_{22}(s) \end{bmatrix},$$

sendo

$$Q(s) = G(s)D(s). \quad (4.24)$$

O modelo do *Cybership I*, no sentido de acoplamento, pode ser considerado um sistema TITO, pois somente duas saídas e duas entradas estão acopladas. Para realizar o desacoplamento do processo, foi utilizado o desacoplamento invertido (GARRIDO; VÁZQUEZ; MORILLA, 2011), esse método pode ser aplicado para quaisquer processos  $n \times n$ . A equação que define um desacoplamento ideal é dada por

$$D(s) = G^{-1}(s)Q(s), \quad (4.25)$$

a matriz de desacoplamento  $D(s)$  é dividida em duas matrizes  $D_d(s)$  de ramo direto entre a saída do controlador ( $c$ ) e entrada do processo ( $u$ ). A outra submatriz  $D_o(s)$  no ramo de realimentação entre a entrada do processo e a saída do controlador. A Figura 4.6 apresenta o diagrama de bloco da matriz de desacoplamento.

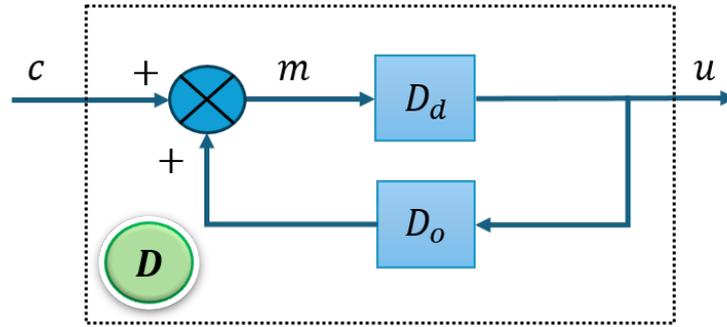


Figura 4.6 – Diagrama de bloco da matriz de desacoplamento.

A matriz  $D_o$  realimenta o desacoplador com a entrada do processo ( $u$ ), e a matriz  $D_d$  fornece a nova entrada pro processo através da realimentação do desacoplador em conjunto com a entrada do desacoplador ( $c$ ). A matriz  $D_d$  deve ter exatamente  $n$  elementos não nulos, pois deve existir apenas uma conexão direta para cada entrada do processo, essa restrição não se aplica a  $D_o$ . A partir do diagrama de bloco apresentado na Figura 4.6, calculando a função de transferência, a expressão para a matriz de desacoplamento é

$$D(s) = Dd(s) \cdot (I - D_o(s) \cdot Dd(s))^{-1}, \quad (4.26)$$

devido a complexidade da Equação (4.26), é mais simples trabalhar com a sua inversa, dada por

$$D(s)^{-1} = (I - D_o(s) \cdot Dd(s)) \cdot D_d^{-1}(s) = D_d^{-1} - D_o(s), \quad (4.27)$$

invertendo a Equação (4.25) e substituindo em (4.27), tem-se

$$D_d^{-1} - D_o(s) = Q^{-1}(s)G(s). \quad (4.28)$$

A matriz  $Q(s)$  é diagonal, e a diferença entre as inversas de  $D_d(s)$  e  $D_o(s)$  resulta em uma matriz de transferência com apenas um elemento a ser calculado em cada posição. É importante destacar que  $D_d(s)$  deve ser não singular. Por isso, ao selecionar seus elementos, apenas um elemento por linha e por coluna pode ser escolhido. Para um processo  $n \times n$ , isso resulta em  $n!$  combinações possíveis para  $D_d(s)$ . Para identificar as combinações, utiliza-se uma notação que indica a coluna correspondente ao elemento escolhido. Por exemplo,

em um sistema  $3 \times 3$ , a configuração 1-2-3 significa que os elementos  $D_d(1, 1)$ ,  $D_d(2, 2)$  e  $D_d(3, 3)$  foram selecionados. Já a configuração 3-1-2 indica que os elementos  $D_d(1, 3)$ ,  $D_d(2, 1)$  e  $D_d(3, 2)$  foram escolhidos, e assim por diante. Cada configuração define um conjunto distinto de elementos para o desacoplador (GARRIDO; VÁZQUEZ; MORILLA, 2011).

Para determinar a matriz  $D_d(s)$ , é necessário garantir que ela seja não singular e diagonal. Dessa forma, a matriz  $D_d(s)$  é expressa por

$$D_d(s) = \begin{bmatrix} dd_{11} & 0 & 0 \\ 0 & dd_{22} & 0 \\ 0 & 0 & dd_{33} \end{bmatrix}, \quad (4.29)$$

por outro lado, a matriz  $D_o(s)$  deve ter elementos nulos em sua diagonal, de forma a garantir que cada saída do sistema desacoplado não influencie diretamente a sua entrada correspondente. A estrutura de  $D_o(s)$  é dada por

$$D_o(s) = \begin{bmatrix} 0 & d_{o12} & d_{o13} \\ d_{o21} & 0 & d_{o23} \\ d_{o31} & d_{o32} & 0 \end{bmatrix}, \quad (4.30)$$

substituindo as Equações (4.29), (4.30) e os valores de  $Q(s)$  e  $G(s)$  na Equação (4.28) de forma escalar, obtém-se

$$\begin{bmatrix} \frac{1}{dd_{11}} & -d_{o12} & -d_{o13} \\ -d_{o21} & \frac{1}{dd_{22}} & -d_{o23} \\ -d_{o31} & -d_{o32} & \frac{1}{dd_{33}} \end{bmatrix} = \begin{bmatrix} \frac{g_{11}}{q_{11}} & \frac{g_{12}}{q_{11}} & \frac{g_{13}}{q_{11}} \\ \frac{g_{21}}{q_{22}} & \frac{g_{22}}{q_{22}} & \frac{g_{23}}{q_{22}} \\ \frac{g_{31}}{q_{33}} & \frac{g_{32}}{q_{33}} & \frac{g_{33}}{q_{33}} \end{bmatrix}, \quad (4.31)$$

logo,

$$\begin{aligned} dd_{11} &= \frac{q_{11}}{g_{11}} & do_{12} &= \frac{-g_{12}}{q_{11}} & do_{13} &= \frac{-g_{13}}{q_{11}}, \\ do_{21} &= \frac{-g_{21}}{q_{22}} & dd_{22} &= \frac{q_{22}}{g_{22}} & do_{23} &= \frac{-g_{23}}{q_{22}}, \\ do_{31} &= \frac{-g_{31}}{q_{33}} & do_{32} &= \frac{-g_{32}}{q_{33}} & dd_{33} &= \frac{q_{33}}{g_{33}}. \end{aligned}$$

Assumindo os elementos da diagonal principal da matriz  $Dd(s)$  como unitários e assumindo que a matriz  $Q(s)$  deve ter os mesmos valores na diagonal principal de  $G(s)$ , ou seja

$$dd_{11} = dd_{22} = dd_{33} = 1, \quad (4.32)$$

$$q_{11} = g_{11}, \quad q_{22} = g_{22}, \quad q_{33} = g_{33}, \quad (4.33)$$

substituindo na Equação (4.31), tem-se que

$$\begin{bmatrix} 1 & -d_{o12} & -d_{o13} \\ -d_{o21} & 1 & -d_{o23} \\ -d_{o31} & -d_{o32} & 1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{g_{12}}{g_{11}} & \frac{g_{13}}{g_{11}} \\ \frac{g_{21}}{g_{22}} & 1 & \frac{g_{23}}{g_{22}} \\ \frac{g_{31}}{g_{33}} & \frac{g_{32}}{g_{33}} & 1 \end{bmatrix}. \quad (4.34)$$

Uma vez que os elementos da matriz  $G(s)$ , as funções de transferência da planta do sistema, são conhecidos, é possível calcular os elementos da matriz  $D_o(s)$  por meio da expressão fornecida pela Equação (4.34).

Assim que as matrizes  $D_d$  e  $D_o$  são determinadas, os elementos do desacoplador total  $D(s)$  podem ser obtidos pela Equação (4.26). O desacoplador  $D(s)$ , composto pelas matrizes  $D_d$  e  $D_o$ , é responsável por minimizar as interações entre as variáveis de entrada e saída do sistema.

Por fim, com a matriz  $D(s)$  determinada, a matriz de função de transferência do sistema desacoplado  $Q(s)$  pode ser calculada. A função de transferência  $Q(s)$  descreve o comportamento dinâmico do sistema após a aplicação do desacoplamento e é dada pela Equação (4.24). A matriz  $Q(s)$  é diagonal, permitindo que cada variável de saída seja controlada independentemente das demais.

Substituindo os valores da matriz de função de transferência do *Cybership I* (Equação (4.23)) na Equação (4.34) para determinar a matriz  $D_o(s)$ , tem-se que

$$D_o(s) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & d_{o23} \\ 0 & d_{o32} & 0 \end{bmatrix}, \quad (4.35)$$

sendo

$$d_{o23} = \frac{-0,0101s^3 - 0,0068s^2 - 0,0009s}{0,0284s^3 + 0,0334s^2 + 0,0121s + 0,0012},$$

$$d_{o32} = \frac{-0,0101s^3 - 0,0068s^2 - 0,0009s}{0,5051s^3 + 0,4266s^2 + 0,1015s + 0,0074},$$

logo, a matriz  $D(s)^{-1}$  é dada por

$$D(s)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{0,0101s^3+0,0068s^2+0,0009s}{0,0284s^3+0,0334s^2+0,0121s+0,0012} \\ 0 & \frac{0,0101s^3+0,0068s^2+0,0009s}{0,5051s^3+0,4266s^2+0,1015s+0,0074} & 1 \end{bmatrix}, \quad (4.36)$$

as matrizes  $D(s)$  e  $Q(s)$  são apresentadas no Apêndice A. A implementação do controlador PID-LMS em diagrama de blocos é apresentada na Figura 4.7.

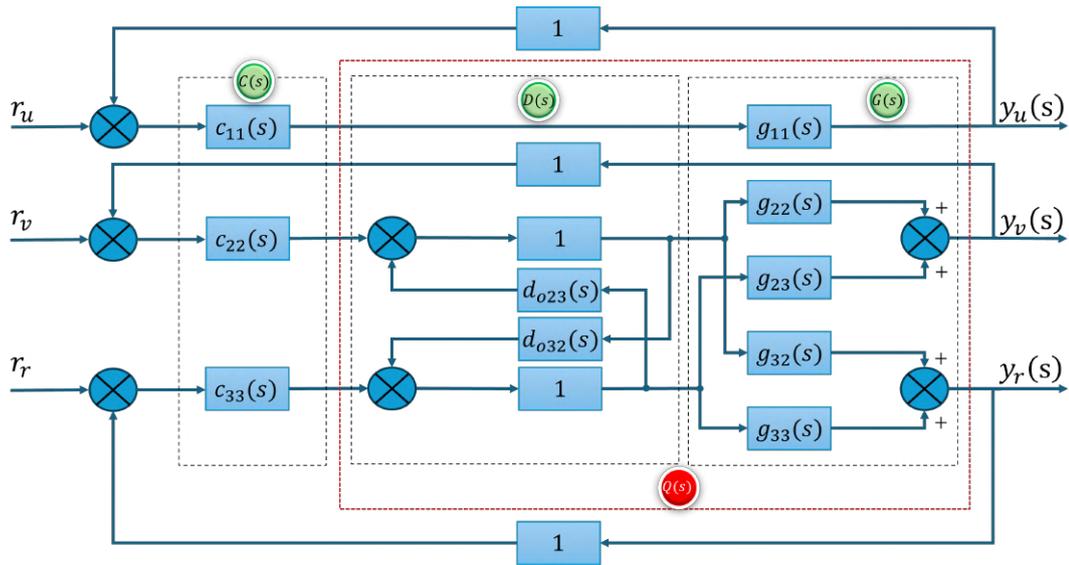


Figura 4.7 – Diagrama de bloco do modelo completo do *Cybership I* em malha fechada com controlador.

Alterando a apresentação do Diagrama de blocos da Figura 4.7 para a representação com  $Q(s)$ , o novo diagrama é apresentado na Figura 4.8.

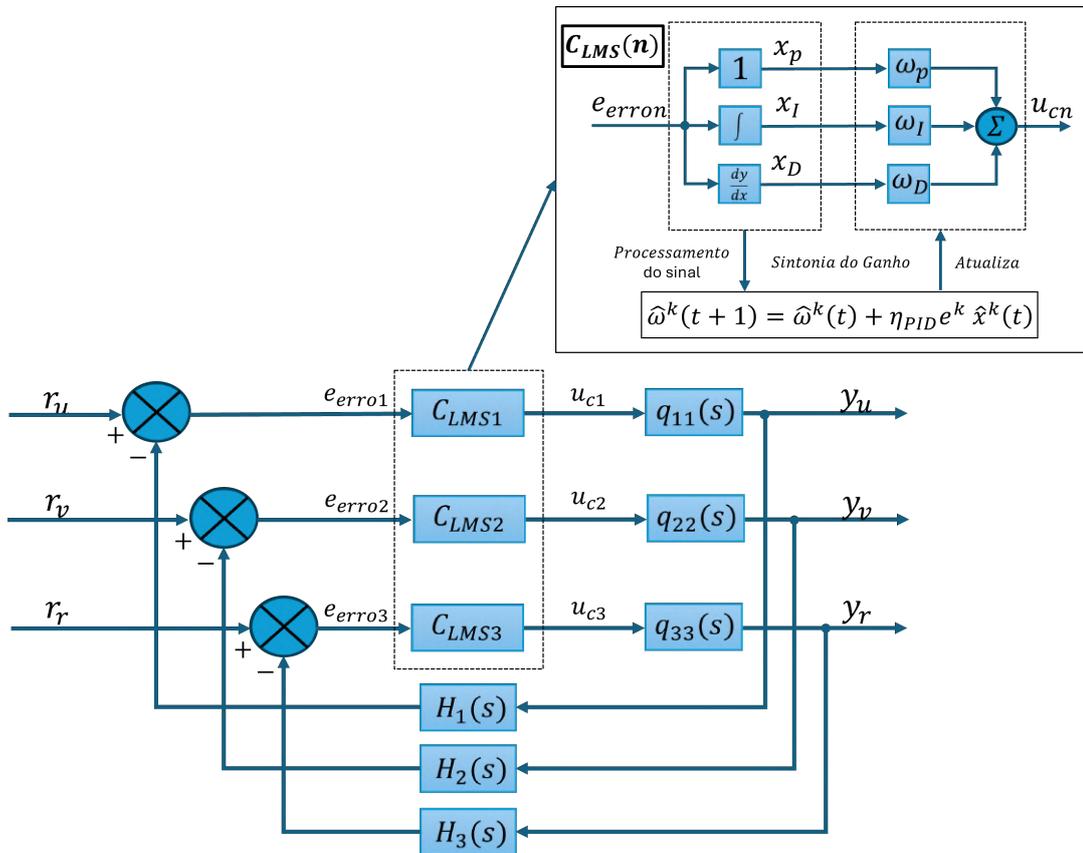


Figura 4.8 – Diagrama de bloco do sistema desacoplado em malha fechada com controlador PID-LMS.

onde  $q_{11}$ ,  $q_{22}$  e  $q_{33}$  são as Funções de Transferência do modelo desacoplado.

## CAPÍTULO 5

# RESULTADOS E ANÁLISE DO CONTROLADOR DE MANOBRABILIDADE

Os experimentos computacionais são realizados em MATLAB® para análise da convergência do algoritmo ADHDP, ou seja, seu desempenho em calcular a solução da Equação HJB de forma *online*, e da capacidade do DLQT-I em rastrear a entrada de referência e eliminar o erro de regime. A análise de convergência da ADHDP é efetuada de forma comparativa ao método de Schur (LAUB, 1979). Os dados de entrada para o controlador ADHDP-DLQT-I são obtidos pela simulação computacional do modelo matemático do USV *Cybership* I. O projeto de controle ADHDP-DLQT-I desenvolvido atua de forma *online* regulando as velocidades  $u$ ,  $v$  e  $r$  para os valores de entrada de referência desejados por meio das forças e momento angular que devem ser gerados pelos propulsores.

As matrizes do modelo completo linear do sistema, discretizadas pelo método do segurador de ordem zero (ZOH) com intervalo de amostragem 0,05 são dadas por

$$A_d = \begin{bmatrix} 0.9895 & 0 & 0 \\ 0 & 0.9914 & -0.0005 \\ 0 & -0.0030 & 0.9748 \end{bmatrix}, B_d = \begin{bmatrix} 0.0026 & 0 & 0 \\ 0 & 0.0014 & 0.0005 \\ 0 & 0.0005 & 0.0251 \end{bmatrix} \text{ e } C_d = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (5.1)$$

As matrizes de ponderação de estados  $Q$  e de entrada  $R$  são definidas por tentativa e erro e escolhidas como matrizes identidade. Os experimentos são realizados com estados iniciais nulos e 7000 iterações.

## 5.1 Resultados do Controlador ADHDP-DLQT-I no Modelo $xg_0$

A primeira análise realizada é para uma entrada degrau unitário e matrizes de ponderação  $Q$  e  $R$  como matrizes identidade. A Figura 5.1 apresenta a resposta das velocidades  $u$  e  $v$ .

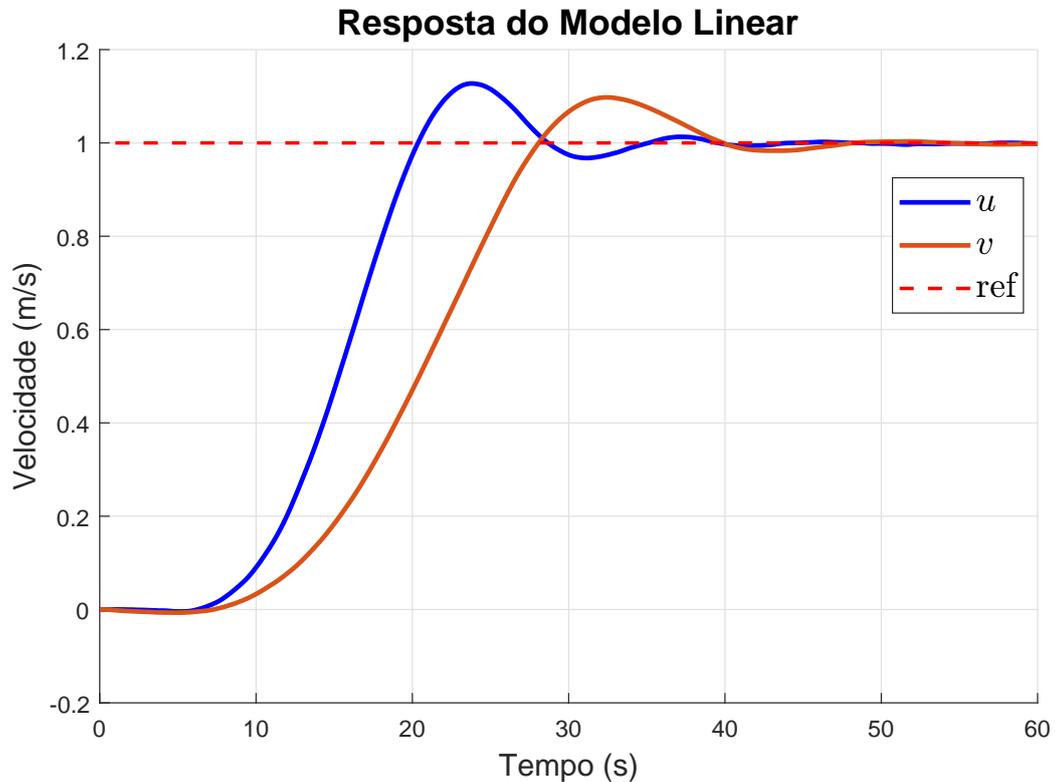


Figura 5.1 – Resposta das velocidades  $u$ -surge e  $v$ -sway do USV para o controlador ADHDP-DLQT-I

Pela Figura 5.1, observa-se que o controlador rastreou a entrada de referência degrau unitário. As velocidades  $u$  e  $v$  seguiram a entrada, onde  $v$  apresenta um comportamento mais lento. As velocidades  $u$  e  $v$  estabilizam seus valores em aproximadamente 40 e 50 segundos respectivamente. Os dois estados apresentaram sobressinal de aproximadamente 10% e não apresentaram erro de regime. Durante as primeiras iterações os estados permanecem em valores próximos a 0 pois é a primeira inicialização do sistema onde os ganhos são pequenos e até a iteração 45, o sistema é conduzido somente pelo ruído de controle.

A Figura 5.2 apresenta a resposta do controlador para a velocidade  $r$ .

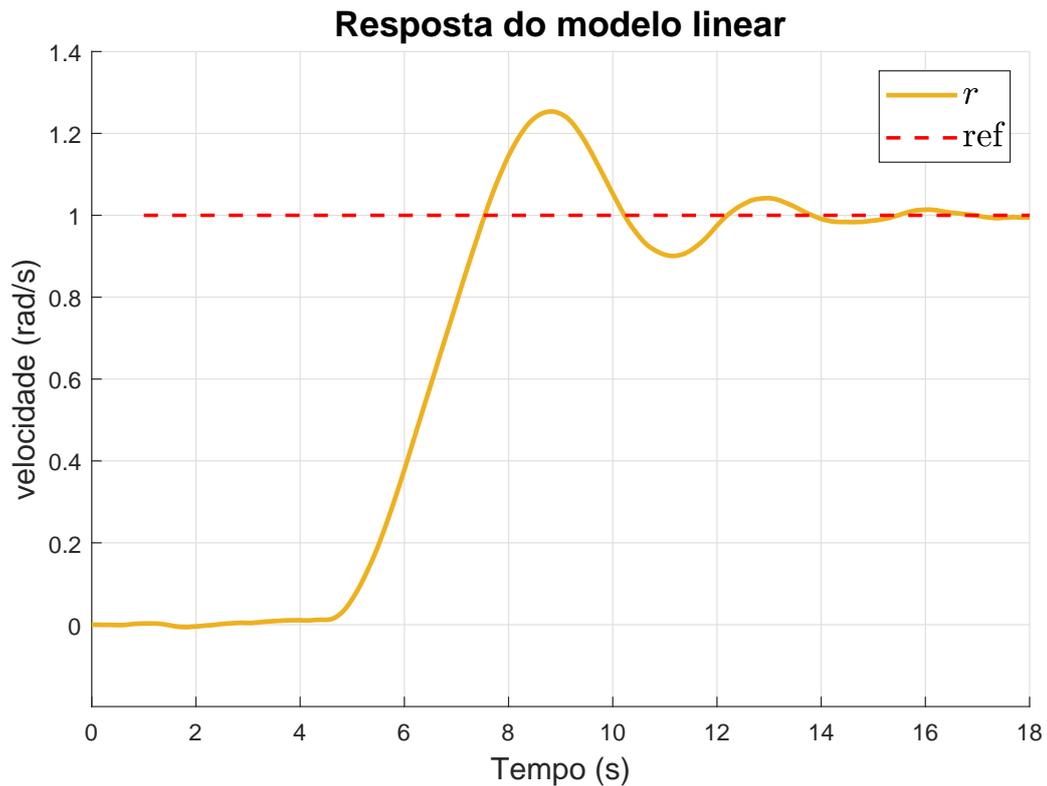


Figura 5.2 – Resposta da velocidade  $r$ -yaw do USV para o controlador ADHDP-DLQT-I.

Na Figura 5.2 é verificada a capacidade do controlador projetado em rastrear a entrada de referência degrau unitário para a velocidade  $r$ . A saída estabiliza com o valor da referência em aproximadamente 16 segundos, tendo condições iniciais nulas. A resposta apresenta um sobressinal aproximadamente 30% e erro de regime nulo.

As ações de controle  $\tau_1$  e  $\tau_2$  são apresentadas na Figura 5.3. As ações apresentam um regime transitório oscilante e estabilizam com valor de regime 4N e 6N respectivamente, a partir de 40 segundos e 50 segundos. As ações de controle estabilizam em um tempo aproximado aos que seus respectivos estados pois seus valores dependem dos estados e à medida que eles se aproximam da referência, as ações estabilizam seus valores.

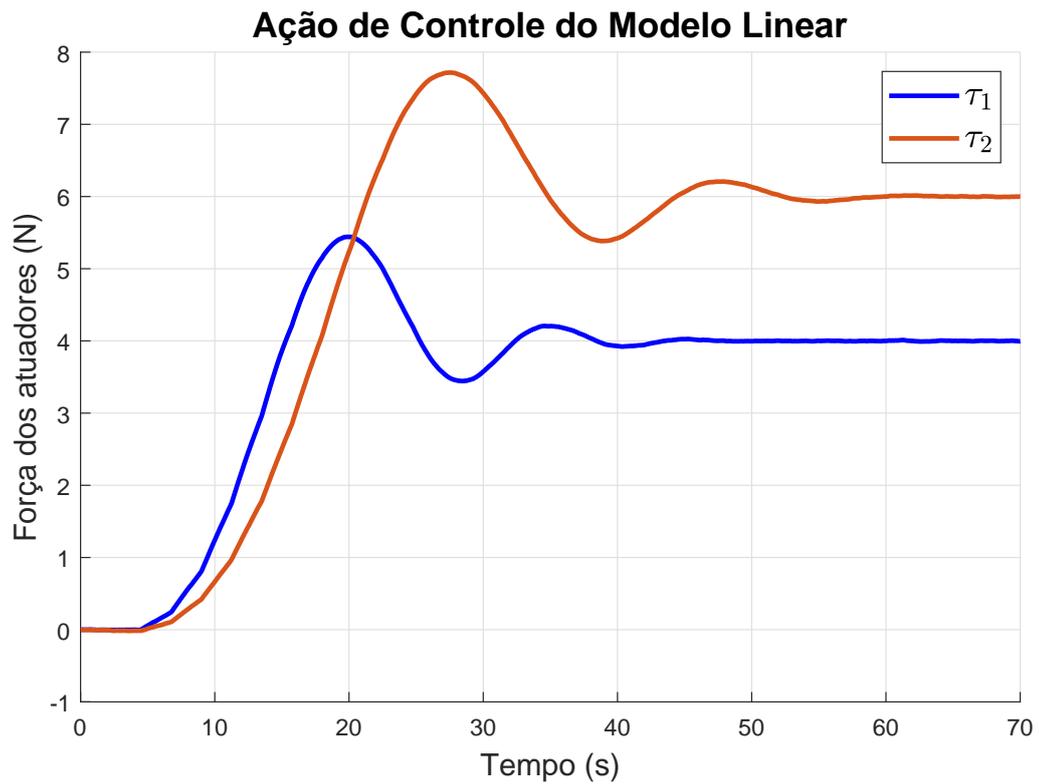


Figura 5.3 – Ações de controle  $\tau_1$  e  $\tau_2$  do controlador ADHDP-DLQT-I.

A Figura 5.4 apresenta a ação de controle  $\tau_3$  para a velocidade  $r$ . A ação de controle, possui uma oscilação inicial até estabilizar em 1N.m a partir de 16 segundos.

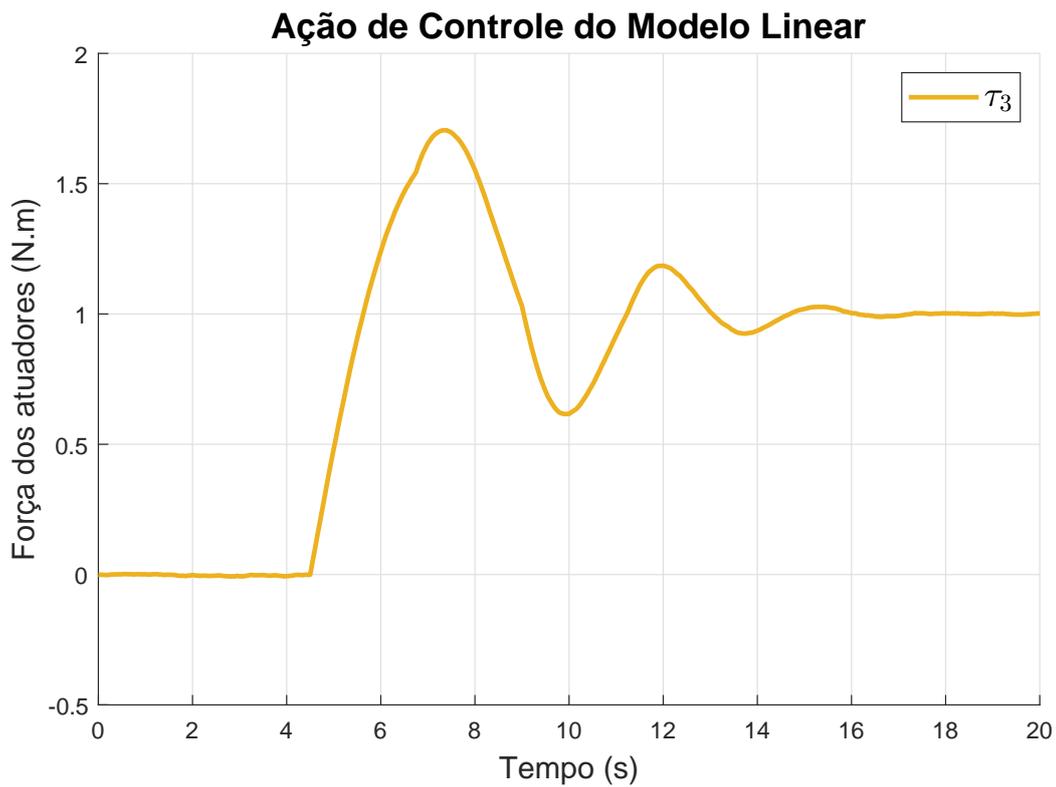


Figura 5.4 – Ação de controle  $\tau_3$  do controlador ADHDP-DLQT-I.

As ações de controle apresentadas nas Figuras 5.3 e 5.4, apresentam valor aproximadamente nulo até cerca de 5 segundos pois o ganho  $K_{ADHDP}$  inicializa com valor nulo, o que leva o esforço de controle para aproximadamente zero e a dinâmica do sistema é conduzida pelo ruído do sistema. Durante as primeiras iterações do processo os estados, os ganhos e ruído de controle possuem um valor pequeno que torna as ações de controle aproximadamente zero.

A convergência dos ganhos calculados para a matriz  $K_{ADHDP}$  é apresentada nas Figuras 5.5 e 5.6. A Figura 5.5 apresenta a convergência dos ganhos de realimentação de estados  $K_1$ .

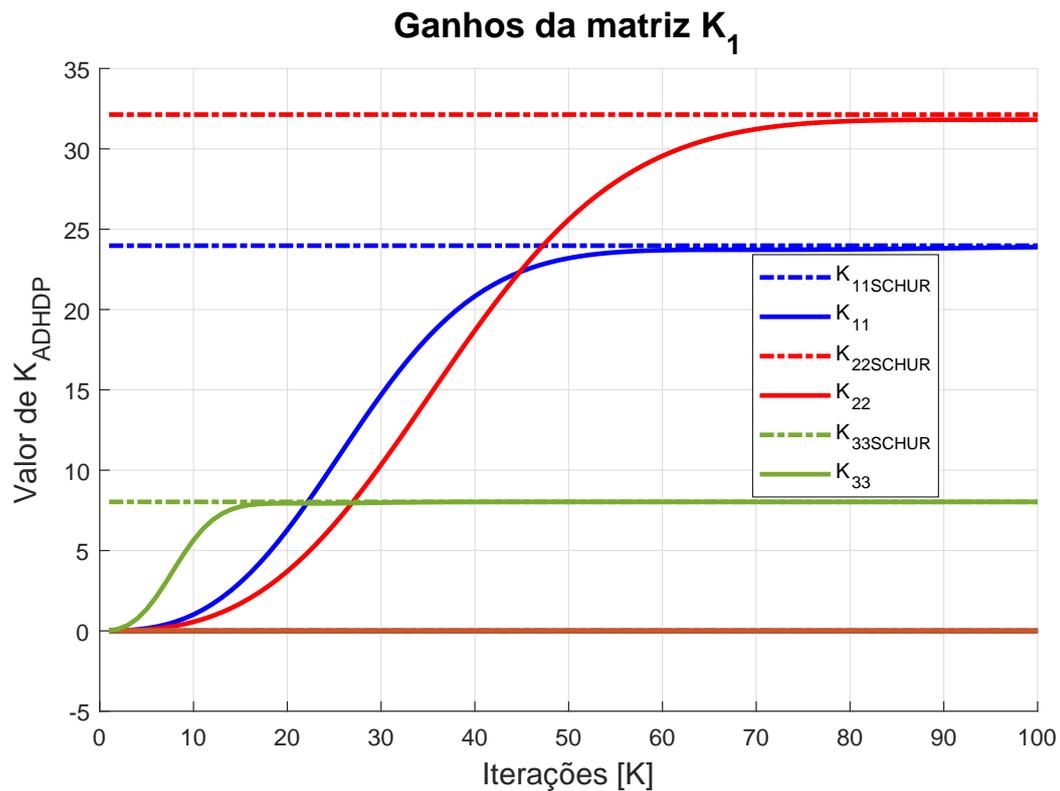


Figura 5.5 – Convergência da matriz de ganhos de realimentação de estados  $K_1$ .

Os nove elementos da matriz  $K_1$  são apresentados na Figura 5.5, entretanto, todos os elementos fora da diagonal principal apresentam valores próximos a 0. A legenda da Figura 5.5 apresenta os valores da diagonal principal e a comparação com os valores obtidos pelo método de Schur.

As iterações de  $K$  representam a quantidade de atualizações do valor do ganho durante o processo. A quantidade de iterações do processo que compõe uma iteração de  $K$ , é determinada pelo valor da variável  $n_{kron}$ , ou seja, 45 iterações do processo. Cada iteração do processo corresponde a uma unidade de tempo  $T_s = 0,05s$ , tal que uma iteração de  $K$  corresponde a 2,25s.

Na Figura 5.5, é verificado que o ganho  $K_{11}$  converge para o valor de Schur por volta de 60 iterações de  $K$ , equivalente a 135s. O ganho  $K_{22}$  converge para o valor obtido pelo método de Schur em aproximadamente 80 iterações de  $K$ , equivalente a 180s. O ganho  $K_{33}$  converge em aproximadamente 15 iterações de  $K$ , equivalente a 33,75s.

A Figura 5.6 apresenta a convergência da matriz de ganhos de realimentação de saída  $K_2$ .

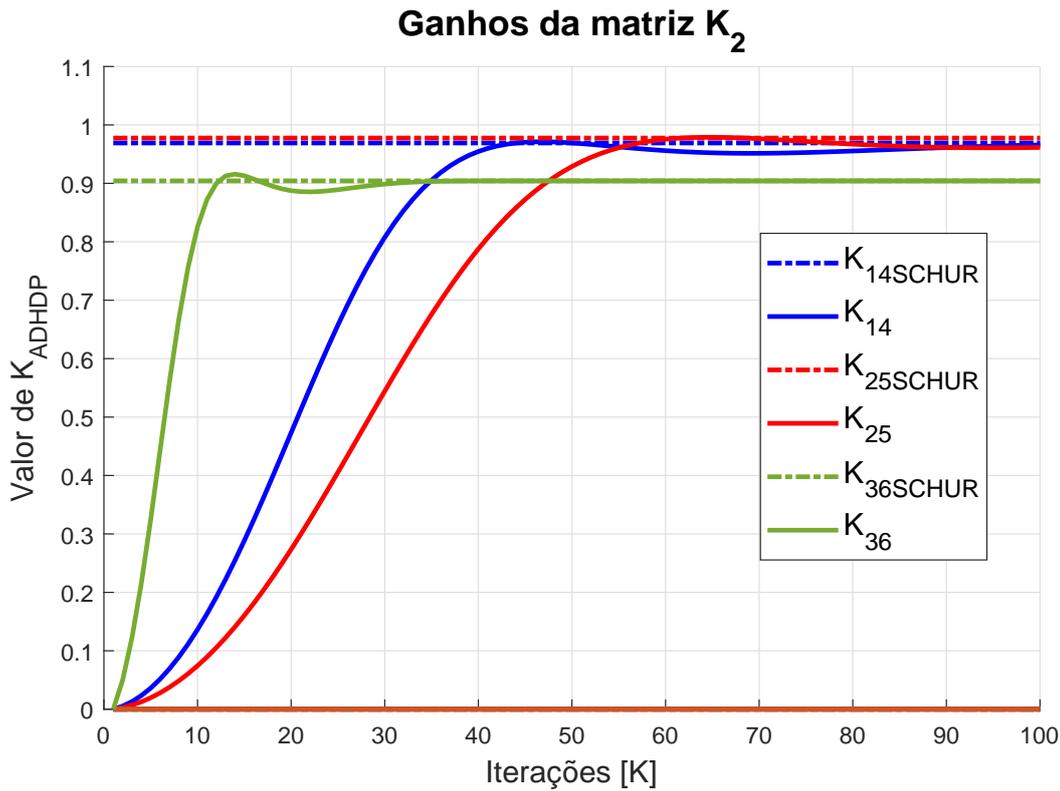


Figura 5.6 – Convergência da matriz de ganhos de realimentação de saída  $K_2$ .

Na Figura 5.6 é exibido que os elementos fora da diagonal principal de  $K_2$  apresentam valores próximos a 0. Os índices dos elementos apresentados na legenda da Figura 5.6 são referentes à matriz  $K_{ADHDP}$ , tal que, os ganhos  $K_{14}$ ,  $K_{25}$  e  $K_{36}$  formam a diagonal principal da matriz  $K_2$ . O ganho  $K_{14}$  converge para o valor de Schur em aproximadamente 90 iterações de  $K$  equivalente a 202,5s, entretanto, a partir de 45 iterações de  $K$ , apresenta valor próximo ao de Schur. o ganho  $K_{25}$  converge em aproximadamente 60 iterações de  $K$ . O ganho  $K_{36}$  converge em proximadamente 30 iterações de  $K$ .

As Figuras 5.5 e 5.6 demonstram que o método implementado possui a capacidade de calcular a solução de *Riccati* para obter o ganho ótimo  $K_{ADHDP}$ . A matriz de ganhos  $K_{ADHDP}$  para 7000 iterações é dada por

$$K_{ADHDP} = \begin{bmatrix} 23,9733 & 0,0000 & 0,0000 & 0,9691 & 0,0000 & 0,0000 \\ 0,0000 & 32,0928 & 0,0000 & 0,0000 & 0,9769 & 0,0000 \\ 0,0000 & 0,0000 & 8,0285 & 0,0000 & 0,0000 & 0,9043 \end{bmatrix}, \quad (5.2)$$

A matriz de ganhos calculada pelo método de Schur é dada por

$$K_{Schur} = \begin{bmatrix} 23,9737 & 0,0000 & 0,0000 & 0,9691 & 0,0000 & 0,0000 \\ 0,0000 & 32,1301 & 0,0000 & 0,0000 & 0,9775 & 0,0000 \\ 0,0000 & 0,0000 & 8,0285 & 0,0000 & 0,0000 & 0,9043, \end{bmatrix} \quad (5.3)$$

A partir das Equações (5.2) e (5.3) é possível notar a capacidade do método ADHDP-DLQT-I em calcular os ganhos ótimos.

## 5.2 Resultados do Controlador PID-LMS para o modelo $xg_0$

A resposta para o estado  $u$ , dada pelo controlador PID-LMS para uma entrada degrau unitário é apresentada na Figura 5.21.

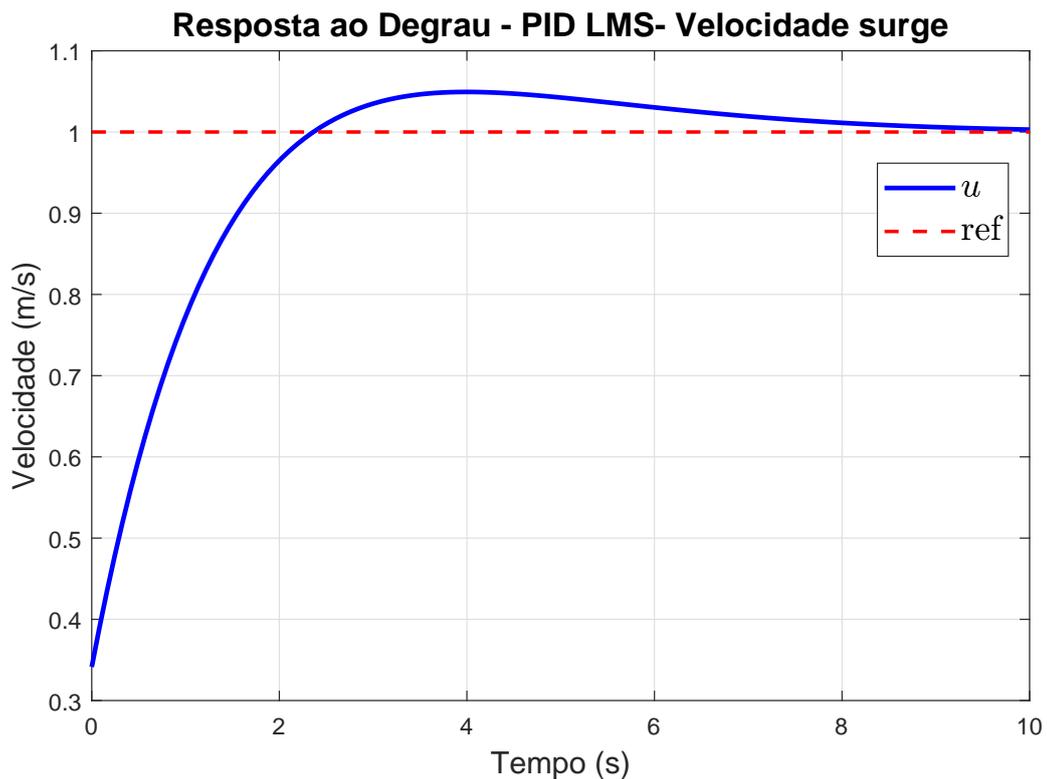


Figura 5.7 – Resposta ao degrau unitário do controlador PID-LMS da velocidade  $u$  do USV.

A resposta ao degrau apresentada na Figura 5.21 exibe a capacidade do controlador PID-LMS em rastrear a entrada de referência para velocidade  $u$ . A resposta apresenta um tempo de subida 1,55 segundos, sobressinal de 4,93% com tempo de pico de 3,98 segundos, tempo de acomodação de 6,92 segundos uma estabilização a partir de 8 segundos e erro de regime nulo. A Figura 5.8 apresenta o erro do sistema para velocidade  $u$ .

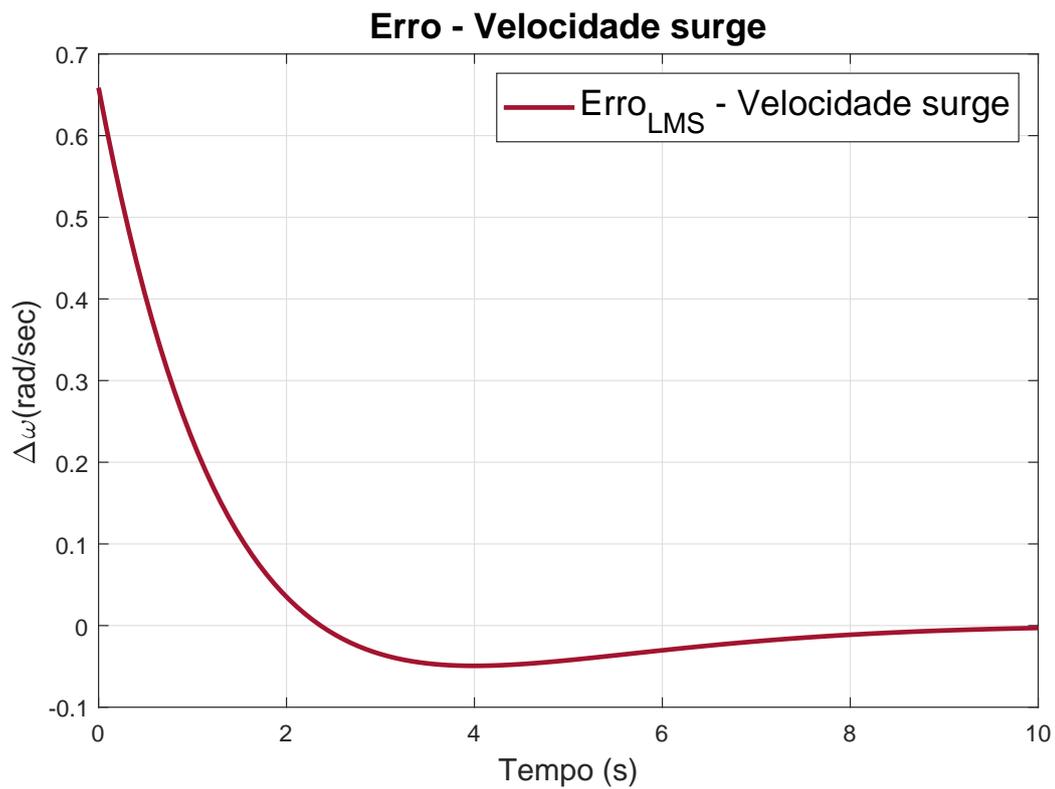


Figura 5.8 – Erro do sistema com controlador PID-LMS para velocidade  $u$ .

O erro apresentado na Figura 5.8 exibe a progressão da redução do erro até 0, causada pelo controlador. O erro apresenta valor máximo de 0,65 no instante inicial do processo, um subssinal de aproximadamente 5% no instante de 3 segundos e estabiliza em zero a partir de 8 segundos.

A Figura 5.9 exibe os sinais de entrada do combinador linear de treinamento LMS  $x_P$ ,  $x_I$  e  $x_D$  do controlador.

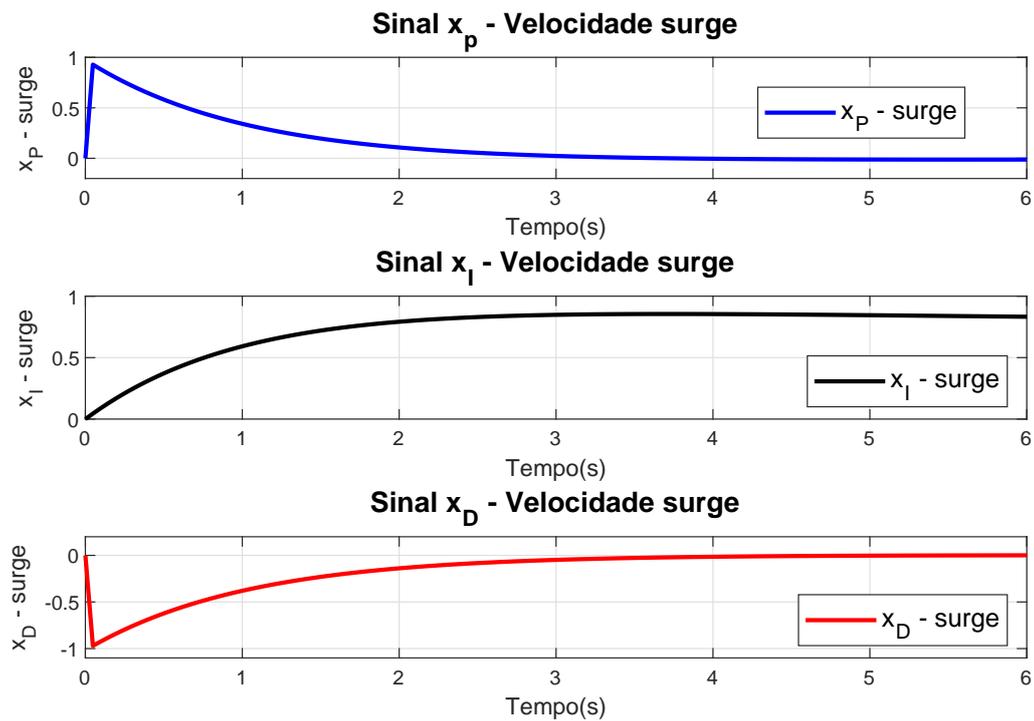


Figura 5.9 – Sinais de entrada do combinador linear do treinamento LMS.

Os sinais apresentados na Figura 5.9, demonstram a estrutura proporcional, integral e derivativa do pré-processamento do erro para a entrada do combinador linear do treinamento LMS para a velocidade  $u$ . Todos os sinais apresentam valor inicial zero como inicialização do processo. O sinal  $x_p$  atinge seu maior valor, aproximadamente 1, nos primeiros instantes, e estabiliza em 0 a partir de 3 segundos. O sinal  $x_I$  converge para 0,78 a partir de 3 segundos. O sinal  $x_D$  apresenta valor mínimo de -1 nos primeiros instantes, e estabiliza em 0 a partir de 3 segundos.

A progressão da convergência dos ganhos do controlador são exibidos na Figura 5.10.

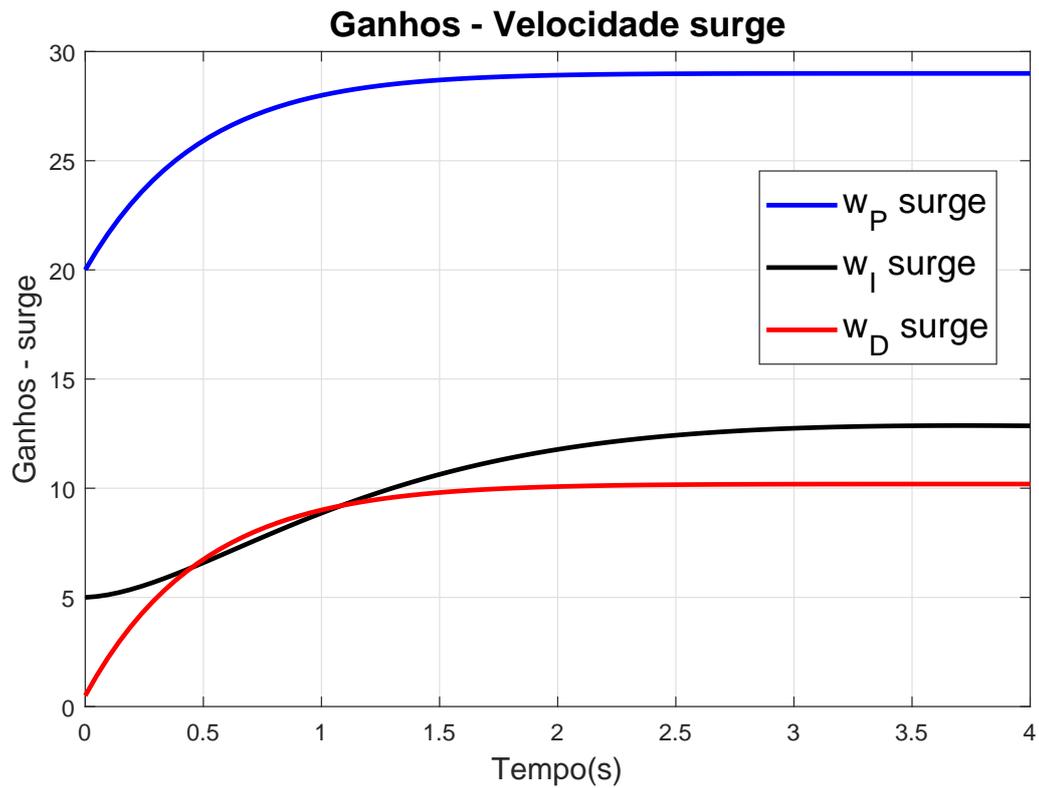


Figura 5.10 – Convergência dos ganhos do controlador PID-LMS para a velocidade  $u$ .

O ganho  $w_P$  inicial é 20 e converge em regime para 29. Para o ganho  $w_I$  o valor inicial e de convergência é 5 e 11,64, respectivamente. O ganho  $w_D$  inicia com valor 0,5 e converge para 10,19 em regime. Os ganhos  $w_P$  e  $w_D$  apresentam uma estabilidade em sua convergência a partir de 1,5 segundos, e  $w_I$  estabiliza a partir de 2,5 segundos.

O gráfico do sinal de controle  $U_c$  da saída do combinador linear é exibido na Figura 5.11.

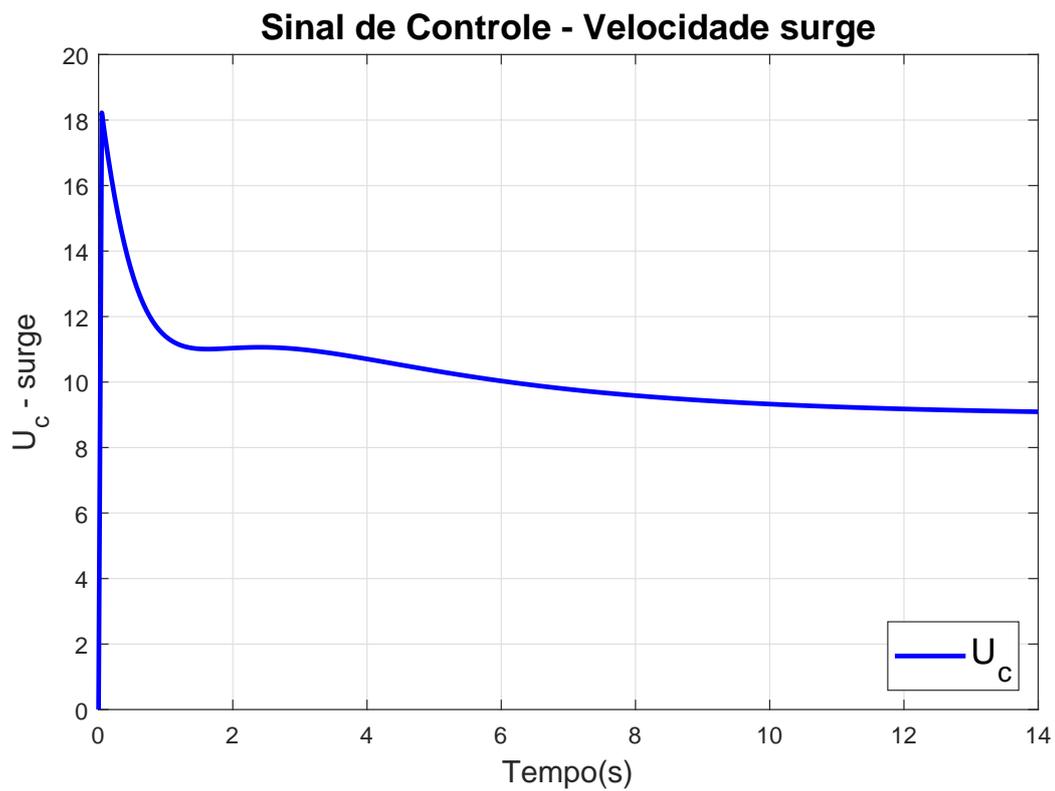


Figura 5.11 – Sinal de controle  $U_c$  para a velocidade  $u$ .

O valor inicial do sinal de controle é zero e atinge valor máximo 18 nos primeiros instantes do processo, como pode ser visto na Figura 5.11. O sinal de controle converge para o seu valor de regime 9, a partir de 12 segundos.

A resposta da velocidade  $v$ , dada pelo controlador para uma entrada degrau unitário, é apresentada na Figura 5.12.

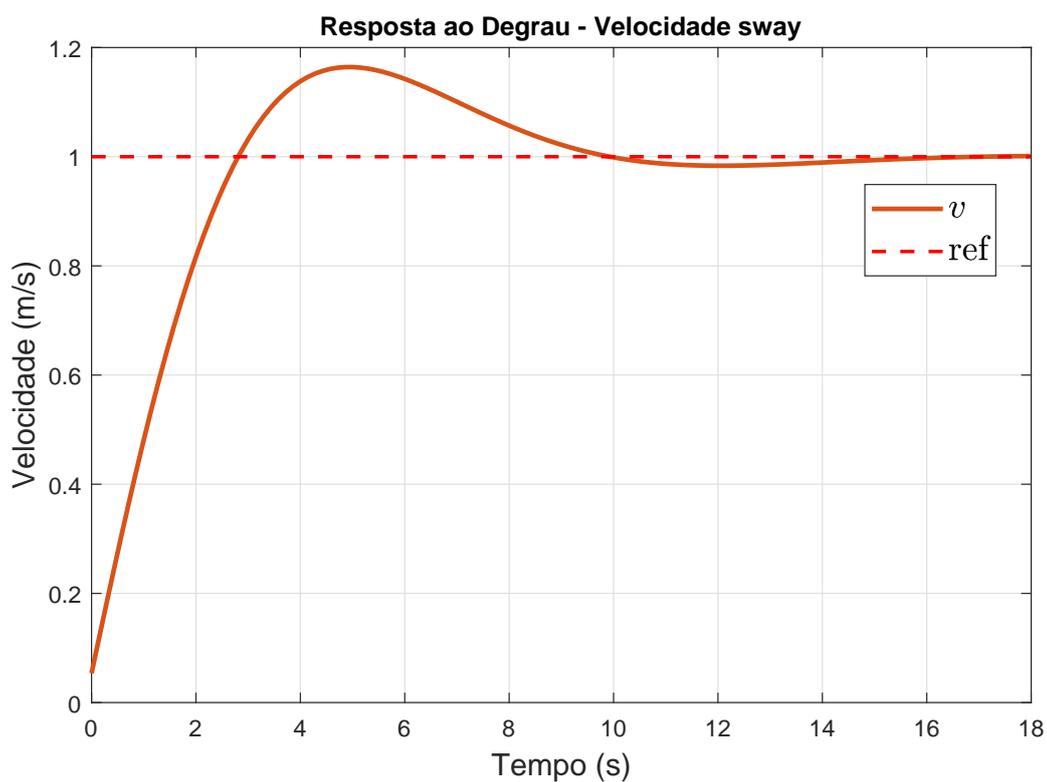


Figura 5.12 – Resposta ao degrau unitário do sistema com controlador PID-LMS para velocidade  $v$ .

A velocidade  $v$  possui valor inicial aproximadamente 0,12, sobressinal de 13,91 % no instante de 4,64 segundos. Alcança a estabilidade em aproximadamente 16 segundos e apresenta erro de regime nulo.

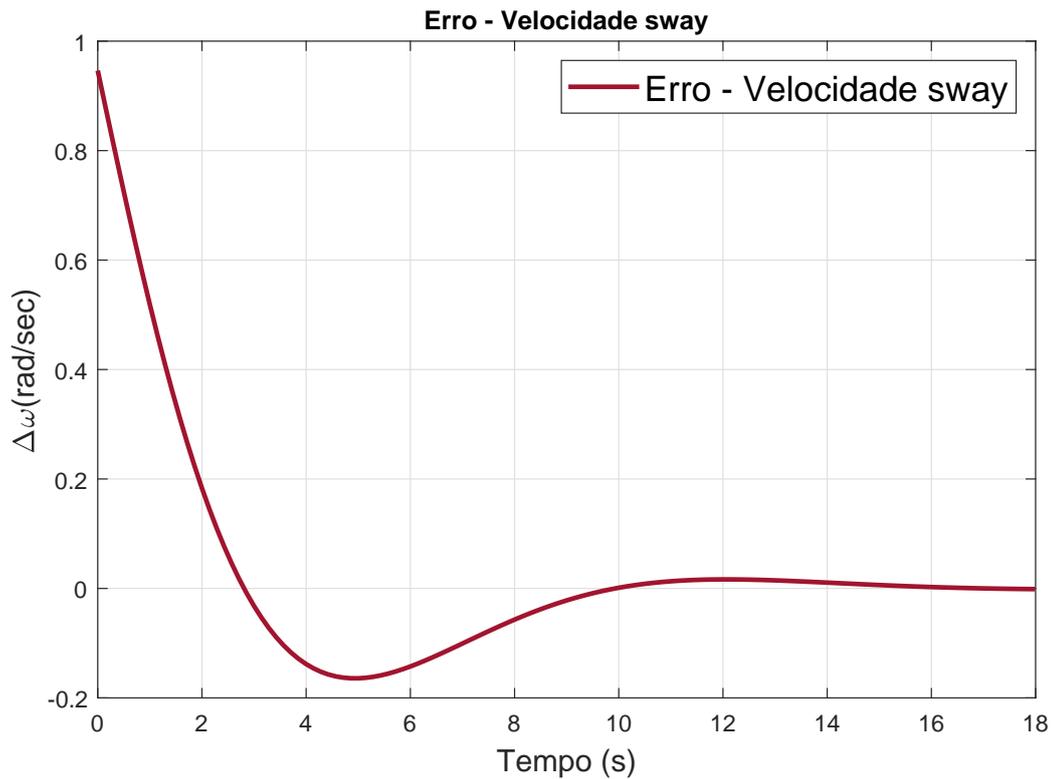


Figura 5.13 – Erro do sistema com controlador PID-LMS para velocidade  $v$ .

O erro apresentado na Figura 5.13 exibe a progressão da redução do erro até 0, causada pelo controlador. O erro apresenta valor máximo aproximadamente 0,9 no instante inicial do processo, um subssinal de aproximadamente 20 % no instante de 4 segundos e estabiliza em zero a partir de 12 segundos. A Figura 5.14 exibe os sinais de entrada do combinador linear de treinamento LMS  $x_P$ ,  $x_I$  e  $x_D$  para a velocidade  $v$ .

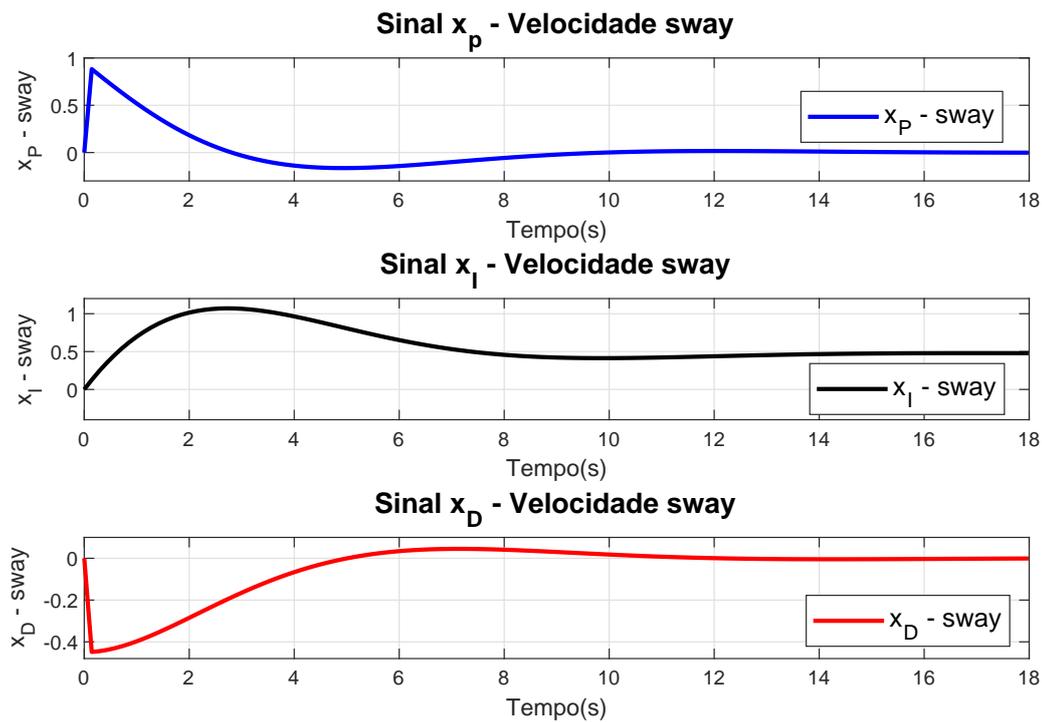


Figura 5.14 – Sinais de entrada do combinador linear de treinamento LMS para velocidade  $v$ .

Os sinais apresentados na Figura 5.14, demonstram a estrutura proporcional, integral e derivativa do pré-processamento do erro para a entrada do combinador linear do treinamento LMS para a velocidade  $v$ . Os três sinais apresentam valor inicial 0. O sinal  $x_p$  atinge seu valor máximo 0,88 nos primeiros instantes, demonstra um subssinal com máximo em aproximadamente 5 segundos e estabiliza em 0 a partir de 9 segundos. O sinal  $x_I$  apresenta valor máximo 1 no instante de 3 segundos e apresenta valor em regime 0,5. O sinal  $x_D$  tem valor mínimo aproximado -0,45 nos primeiros instantes, pequeno sobressinal no instante de 7 segundos e estabiliza em 0 a partir de 10 segundos.

A progressão da convergência dos ganhos do controlador são exibidos na Figura 5.15.

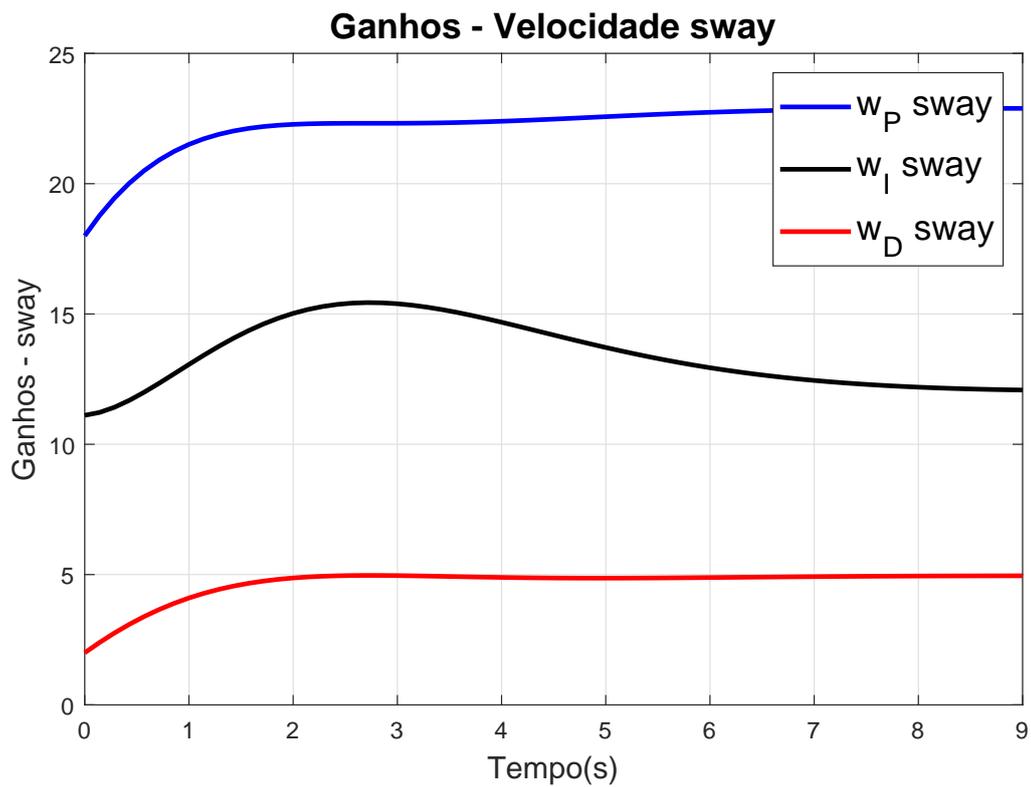


Figura 5.15 – Convergência dos ganhos do controlador PID-LMS para a velocidade  $v$ .

Os ganhos apresentam valores iniciais e mínimos entre 3 e 18, e os valores de convergência escalares entre 5 e 23 em 3 segundos. O ganho  $w_P$  inicial é 18 e converge em regime para 23. Para o ganho  $w_I$  o valor inicial e de convergência é 11 e 12, respectivamente. O ganho  $w_D$  inicia com valor aproximadamente 3 e converge para 5.

O gráfico do sinal de controle  $U_c$  da saída do combinador linear para velocidade  $v$  é exibido na Figura 5.11.

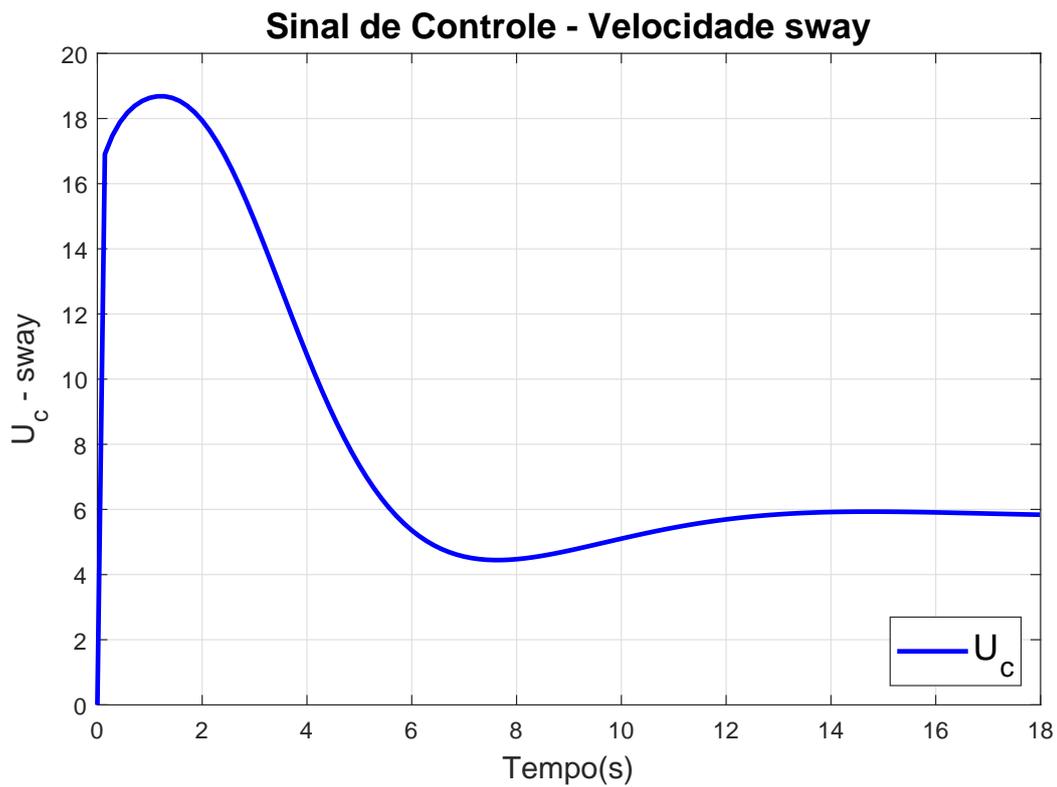


Figura 5.16 – Sinal de controle  $U_c$  para a velocidade  $v$ .

O valor inicial do sinal de controle é zero e atinge valor máximo de aproximadamente 19 nos primeiros instantes do processo, como pode ser visto na Figura 5.11. O sinal converge para 6 em regime a partir de 14 segundos.

A resposta da velocidade  $r$  para uma entrada degrau unitário, é exibida na Figura 5.17.

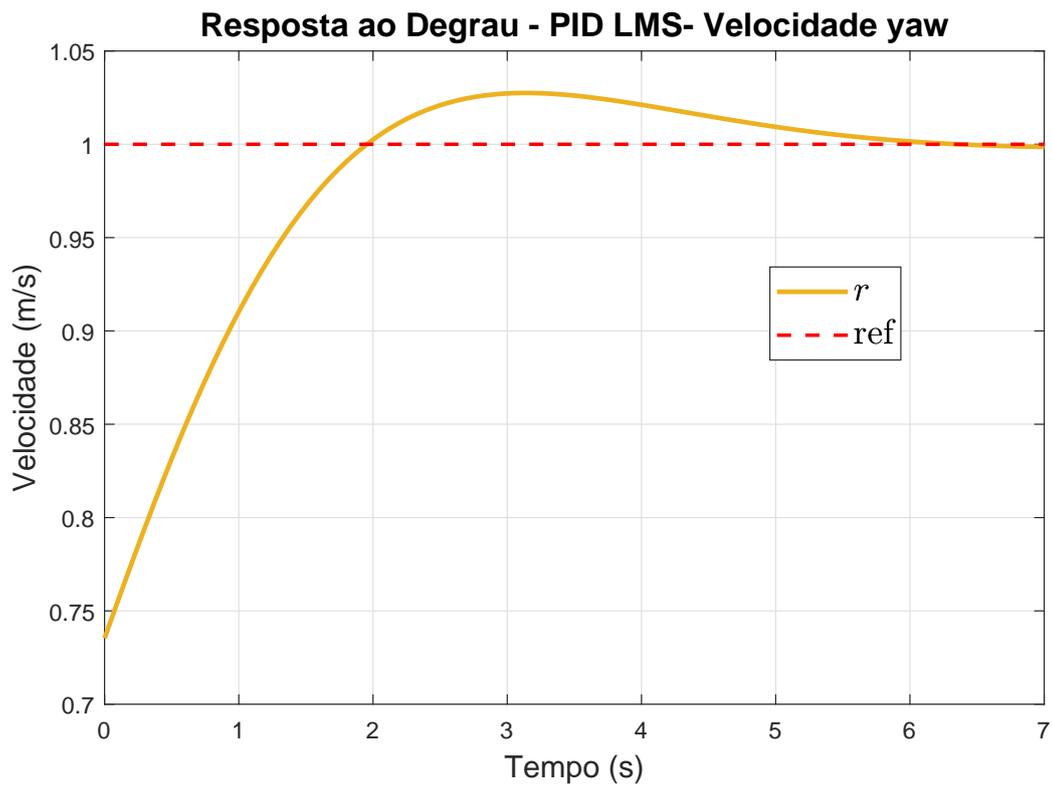


Figura 5.17 – Resposta ao degrau unitário do sistema com controlador PID-LMS para velocidade  $r$ .

A velocidade  $r$  possui valor inicial de aproximadamente 0,75, tempo de subida de 0,92 segundos, sobressinal de 2,74% com tempo de pico 3,12 segundos, tempo de acomodação de 4 segundos e erro de regime nulo. A Figura 5.18 apresenta o erro do sistema para velocidade  $r$ .

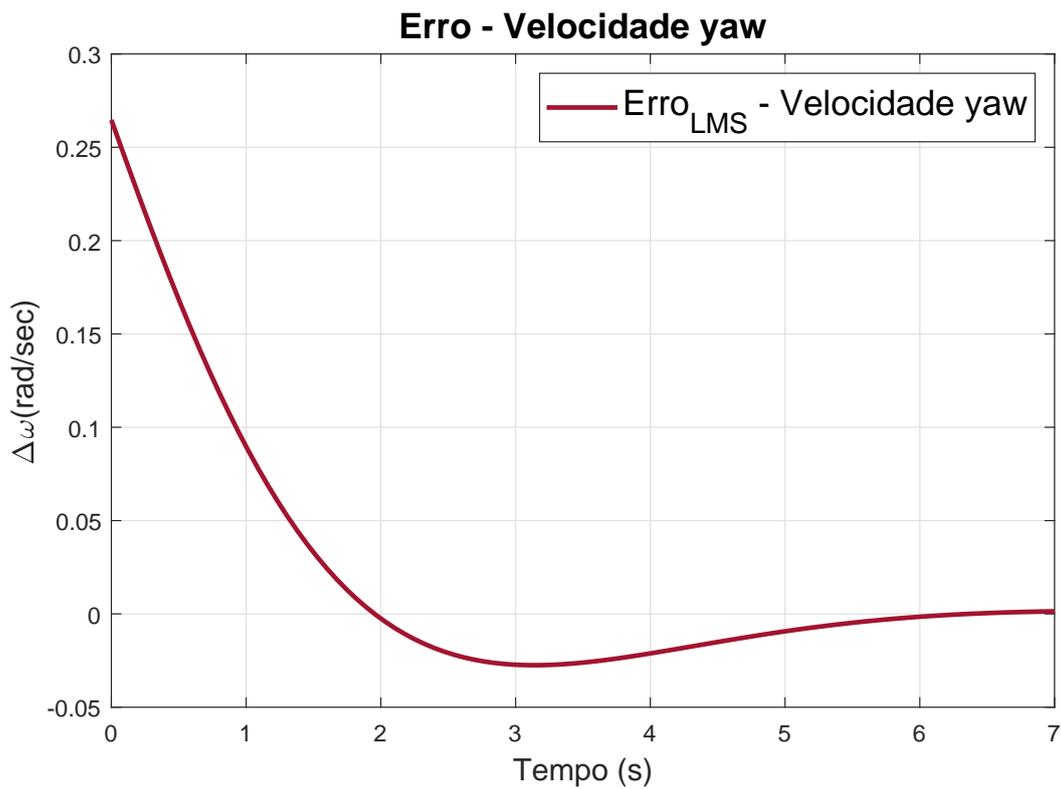


Figura 5.18 – Erro do sistema com controlador PID-LMS para velocidade  $r$ .

O erro apresentado na Figura 5.18 exibe a progressão da redução do erro até zero em regime, causada pelo controlador. O erro apresenta valor máximo de aproximadamente 0,26 no início do processo, um subssinal de aproximadamente 2,5 % no instante de 3 segundos e estabiliza em zero a partir de 6 segundos. A Figura 5.19 exibe os sinais de entrada do combinador linear de treinamento LMS  $x_P$ ,  $x_I$  e  $x_D$  para a velocidade  $r$ .

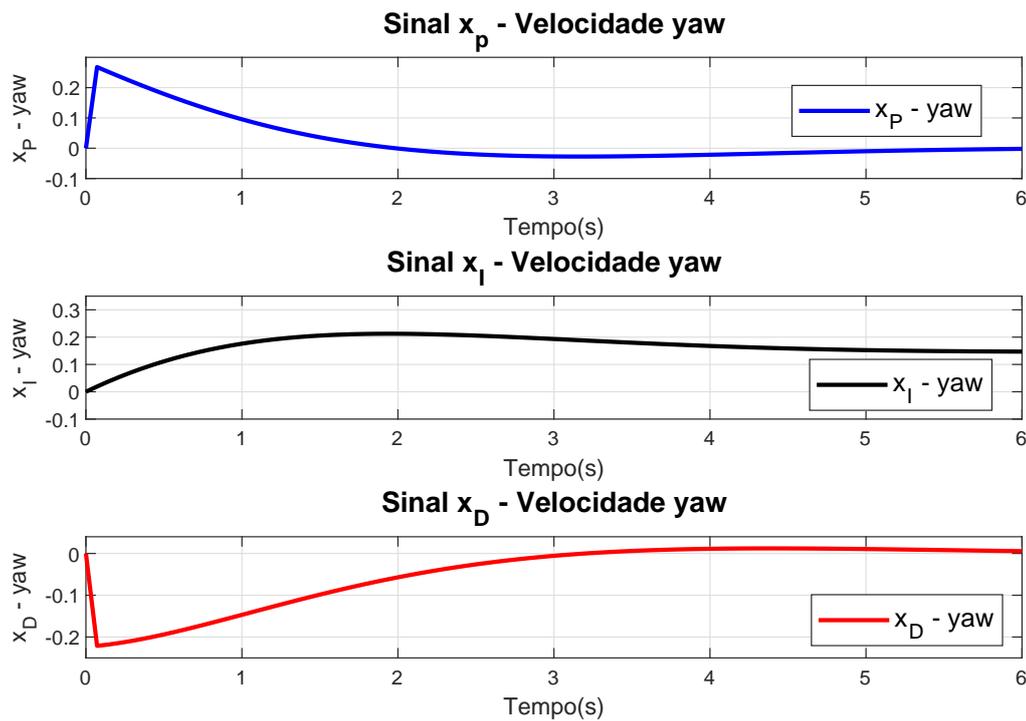


Figura 5.19 – Sinais de entrada do combinador linear de treinamento LMS para velocidade  $r$ .

Os sinais apresentados na Figura 5.19, exibem o sinais proporcional, integral e derivativo do erro como pré-processamento para a entrada do combinador linear do treinamento LMS para a velocidade  $r$ . O sinal  $x_p$  inicia em zero e atinge seu maior valor, aproximadamente 0,3, nos primeiros instantes, demonstra um pequeno subssinal no instante de 3 segundos e estabiliza em 0 a partir de 5 segundos. O sinal  $x_i$  tem valor inicial 0, valor máximo 0,2 em 2 segundos e converge para o valor de regime aproximadamente 0,15. O sinal  $x_D$  apresenta valor inicial 0, valor mínimo aproximado de -0,25 nos primeiros instantes, pequeno sobressinal no instante de 4 segundos e estabiliza em 0 a partir de 5 segundos.

A progressão da convergência dos ganhos do controlador são exibidos na Figura 5.15.

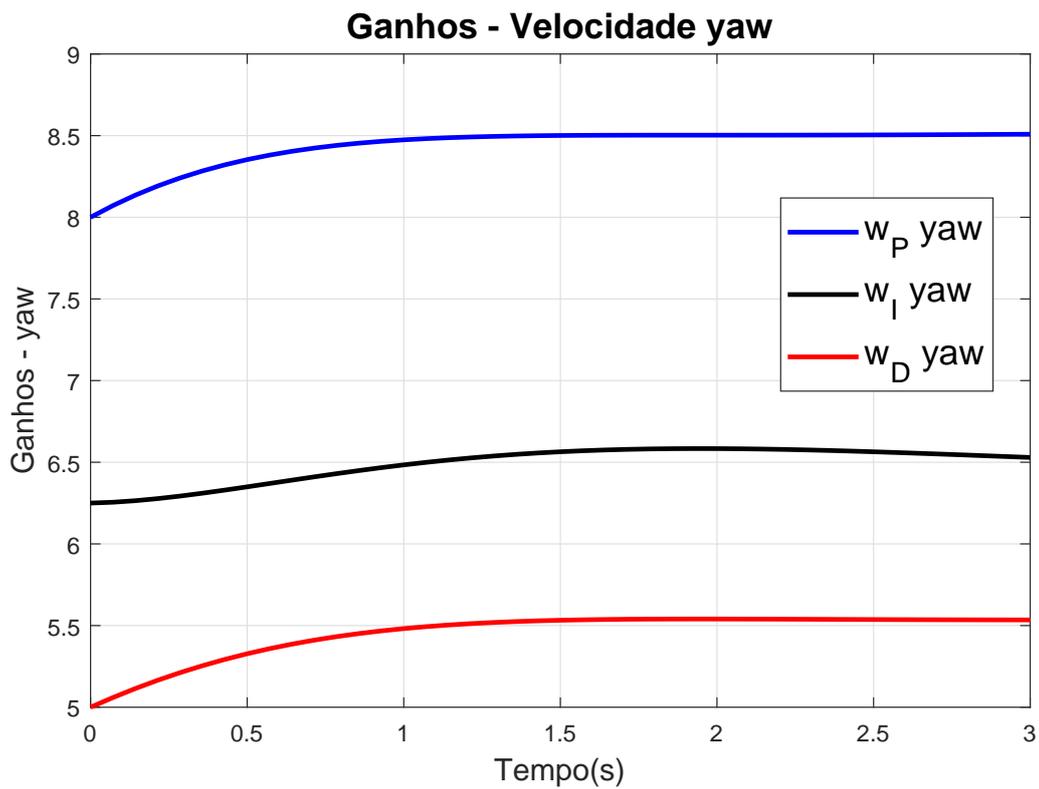
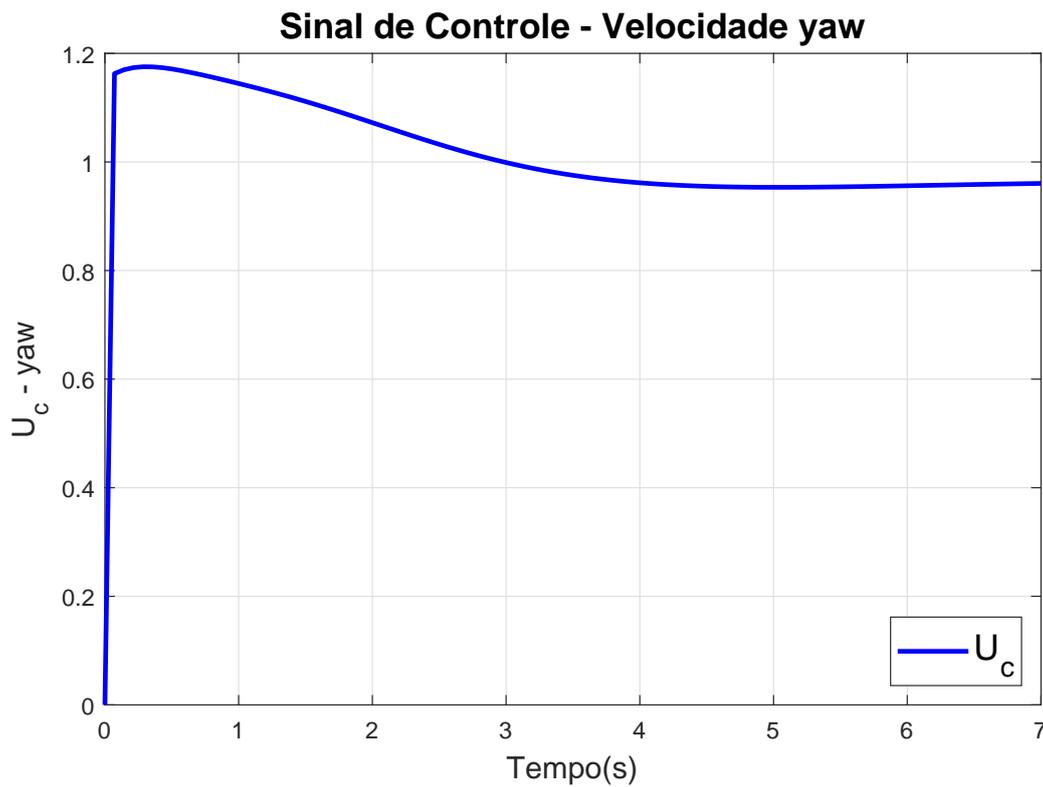


Figura 5.20 – Convergência dos ganhos do controlador PID-LMS para a velocidade  $r$ .

O ganho  $w_P$  inicial é 8 e converge em regime para 8,5. Para o ganho  $w_I$  o valor inicial e de convergência é 6,25 e 6,42, respectivamente. O ganho  $w_D$  inicia com valor 5 e converge para 5,54.

O gráfico do sinal de controle  $U_c$  da saída do combinador linear para velocidade  $r$ , exibido na Figura 5.11 é aproximadamente 0,95 em regime a partir de 5 segundos.

Figura 5.21 – Sinal de controle  $U_c$  para a velocidade  $r$ .

O valor inicial do sinal de controle é zero e atinge valor máximo de aproximadamente 3 nos primeiros instantes do processo, como pode ser visto na Figura 5.11.

O sinal  $x_P$  é o próprio erro do sistema, logo, é zero quando a resposta do sistema não apresenta erro. O  $x_I$  é a integral do erro, então, se torna constante quando o erro se torna zero. O  $x_D$  se torna zero pois é a variação do erro, logo, quando o valor do erro estabiliza a sua variação é nula. Os ganhos iniciais para todas as velocidades são determinados por tentativa e erro.

### 5.3 Resultados do Modelo Não Linear

Após a implementação do controlador em modelo linear, as ações de controle geradas são avaliadas no modelo não linear por meio do simulador apresentado na Seção (4.3). A Figura 5.22 apresenta a resposta do modelo não linear para as velocidades  $u$  e  $v$ .

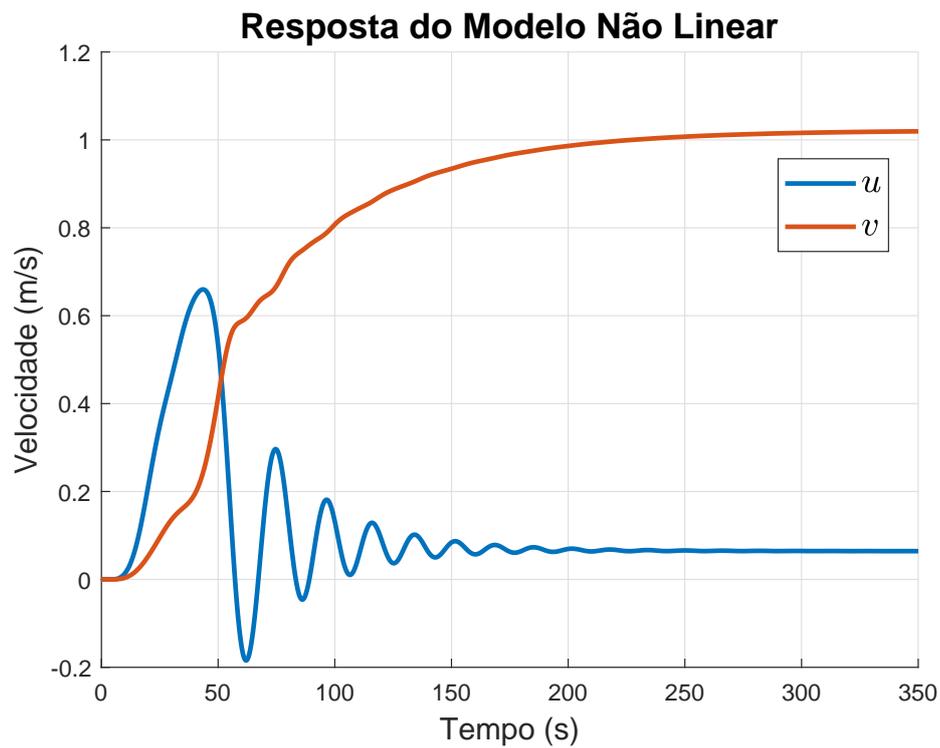


Figura 5.22 – Avaliação das velocidades  $u$  e  $v$  para o modelo não linear.

A partir da Figura 5.22, nota-se que a velocidade  $v$  apresenta uma resposta em regime permanente próxima ao valor de referência. A velocidade  $u$  apresenta uma oscilação até estabilizar em aproximadamente 0,12.

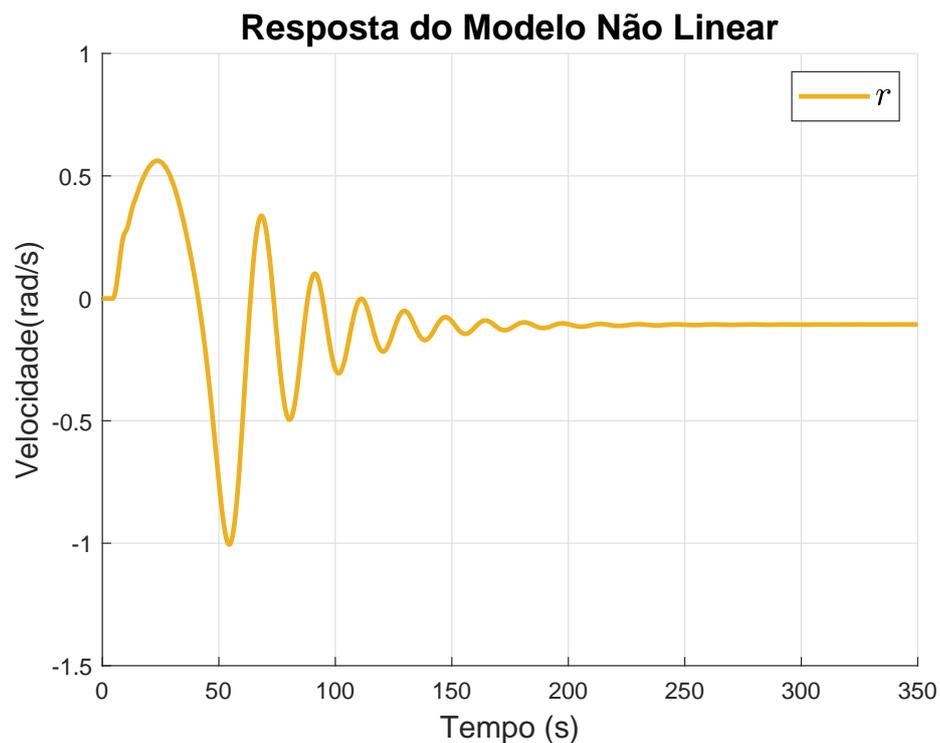


Figura 5.23 – Avaliação das velocidades  $u$  e  $v$  para o modelo não linear.

A resposta da velocidade  $r$ , exibida na Figura 5.23, apresenta oscilação por conta da influência das outras velocidades, até estabilizar em aproximadamente -0,15.

As respostas aplicadas ao modelo não linear demonstra a relação com o modelo linear e sua linearização no ponto de operação.

## 5.4 Resultados para o Modelo Completo com Controlador ADHDP-DLQTI

Os resultados para o modelo completo apresentam variações desprezíveis em seus resultados em comparação ao modelo  $x_{g_0}$ . Os resultados para as respostas, ações de controle e ganhos são apresentados nas Figuras 5.24, 5.25, 5.26 e 5.27.

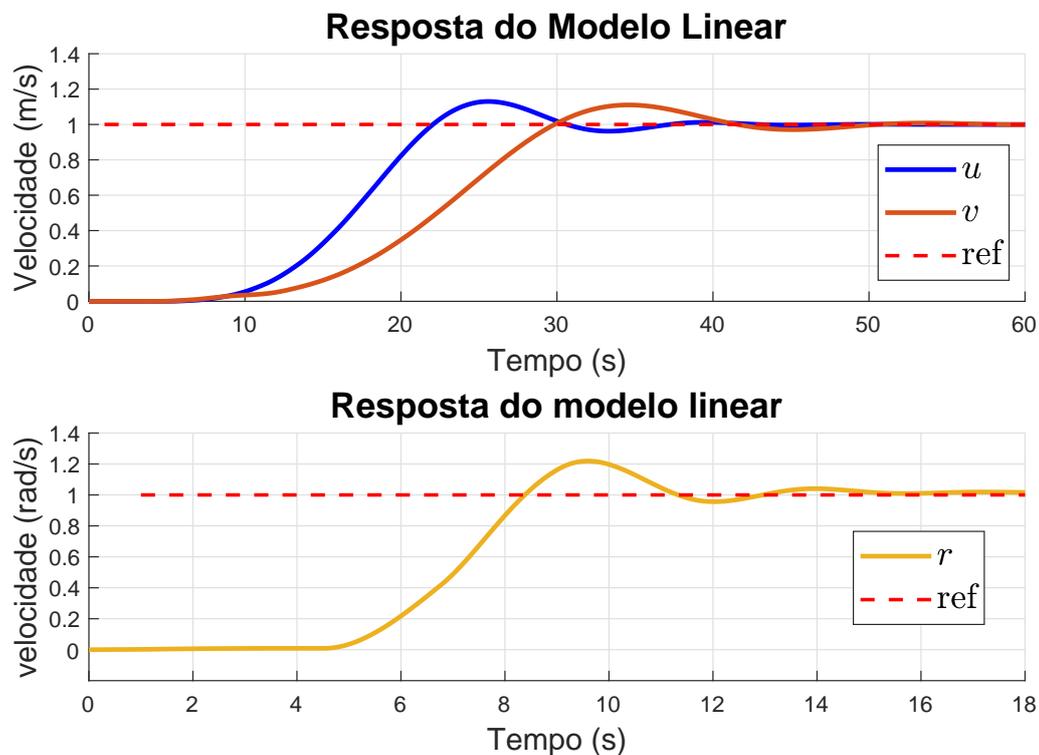


Figura 5.24 – Respostas do controlador ADHDP-DLQTI-I para as velocidades  $u$ ,  $v$  e  $r$  utilizando o modelo completo do Cybership I.

Avaliando as respostas apresentadas na Figura 5.24, observa-se que variação da dinâmica causada por  $x_g$  é nula para a velocidade  $u$ , pouco influente em  $v$ , e apesar de afetar a velocidade  $r$ , a sua resposta apresentou variação irrelevante em comparação ao modelo  $x_{g_0}$ . Nesse caso, as análises das respostas apresentadas na Figura 5.24 são as mesmas apresentadas na Seção 5.1.

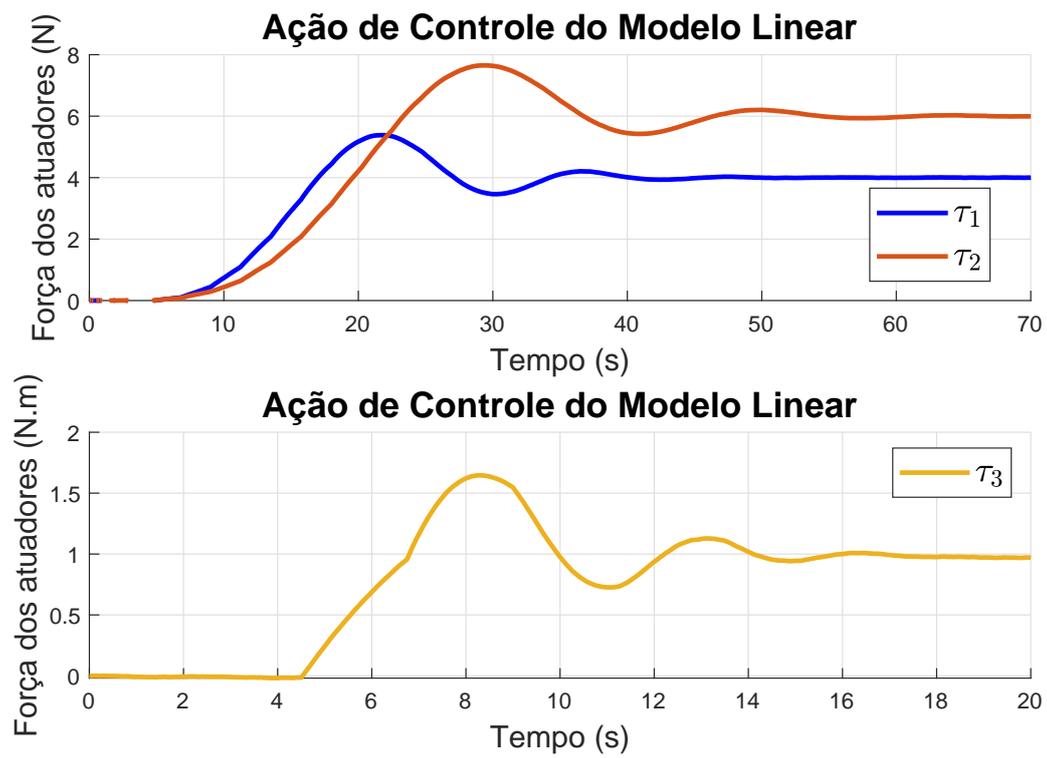


Figura 5.25 – Ação de controle do controlador ADHDP-DLQT-I para as velocidades  $u$ ,  $v$  e  $r$  utilizando o modelo completo do Cybership I.

As ações de controle  $\tau_1$ ,  $\tau_2$  e  $\tau_3$ , exibidas na Figura 5.25, apresentam comportamentos semelhantes ao modelo  $xg_0$ . A convergência dos ganhos  $K_1$  e  $K_2$  é apresentada nas Figuras 5.26 e 5.27.

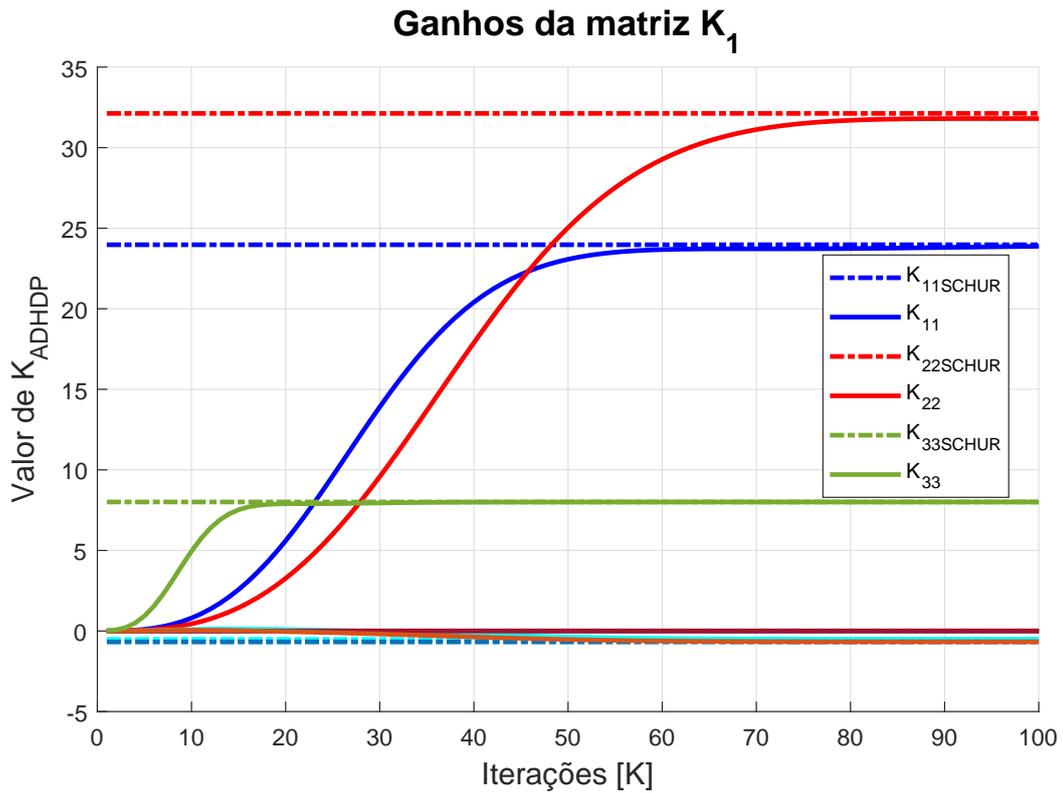


Figura 5.26 – Convergência dos ganhos da matriz  $K_1$  para o modelo completo do *Cybership* I.

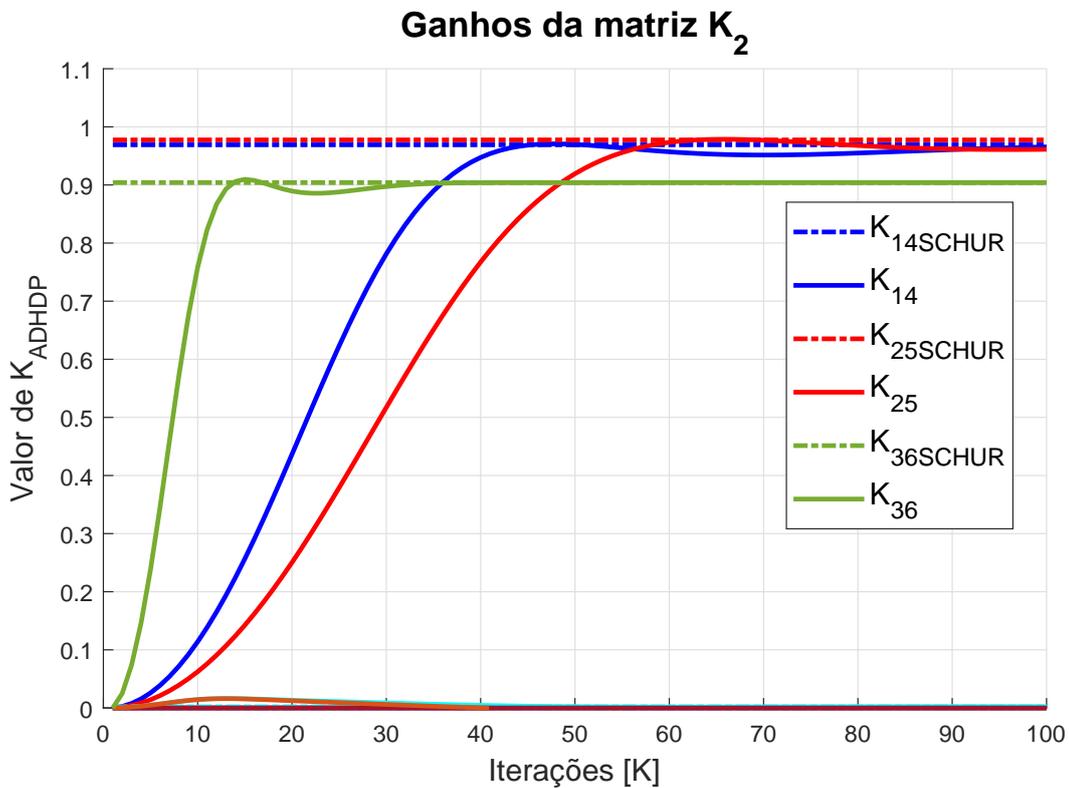


Figura 5.27 – Convergência dos ganhos da matriz  $K_2$  para o modelo completo do *Cybership* I.

A partir da análise das Figuras 5.26 e 5.27, nota-se que a convergência dos ganhos é similar à apresentada para o modelo  $x_{g_0}$ , a principal diferença na matriz de ganhos são os elementos fora das diagonais principais como pode ser notado pelas matrizes  $K_{ADHDP}$  e  $K_{Schur}$  dadas por

$$K_{ADHDP} = \begin{bmatrix} 23,9733 & 0,0000 & 0,0000 & 0,9691 & 0,0000 & 0,0000 \\ 0,0000 & 32,0852 & -0,5176 & 0,0000 & 0,9768 & 0,0020 \\ 0,0000 & -0,6755 & 8,0081 & 0,0000 & 0,0046 & 0,9041 \end{bmatrix}, \quad (5.4)$$

$$K_{Schur} = \begin{bmatrix} 23,9737 & 0,0000 & 0,0000 & 0,9691 & 0,0000 & 0,0000 \\ 0,0000 & 32,1253 & -0,5184 & 0,0000 & 0,9775 & 0,0020 \\ 0,0000 & -0,6764 & 8,0081 & 0,0000 & -0,0046 & 0,9041 \end{bmatrix}. \quad (5.5)$$

Os elementos fora da diagonal principal representam o acoplamento entre a velocidade  $v$  e  $r$ , pode-se notar que os ganhos para a velocidade  $u$  permaneceram iguais por conta do seu desacoplamento no modelo. Os parâmetros das simulações foram os mesmos para o modelo  $x_{g_0}$ .

## 5.5 Resultados para o Modelo Completo com Controlador PID-LMS

A resposta para a velocidade  $u$  é apresentada na Figura 5.28.

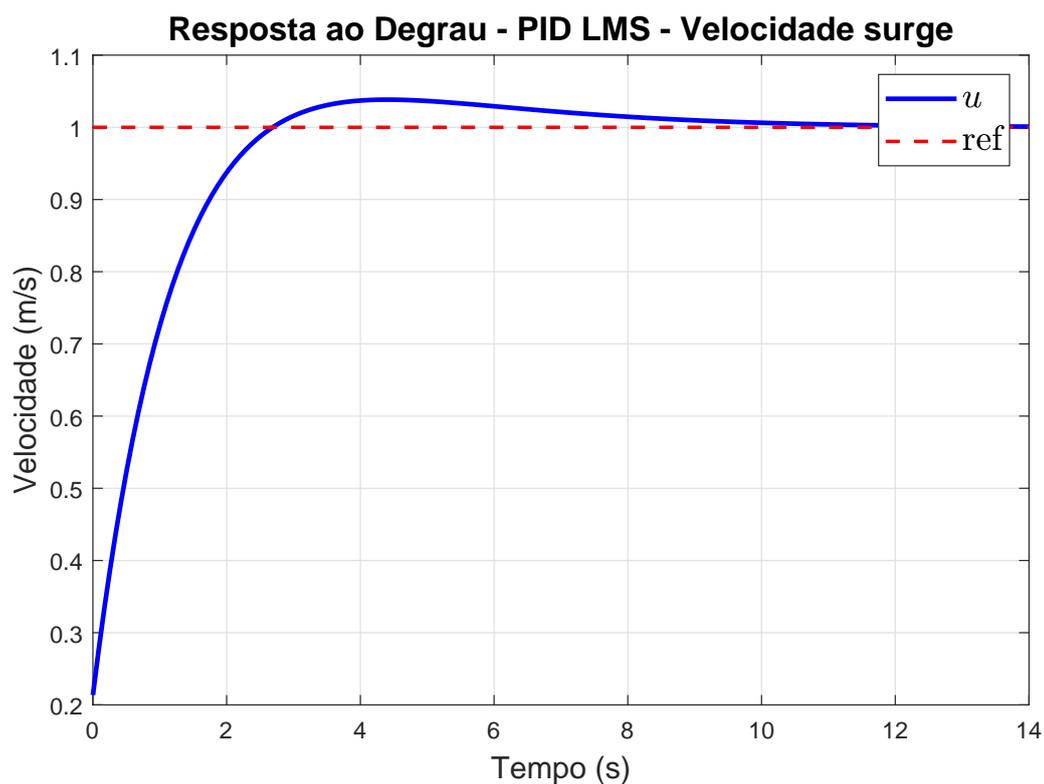


Figura 5.28 – Resposta ao degrau unitário do controlador PID-LMS da velocidade  $u$  do USV.

A resposta apresenta tempo de subida 1,74s, sobressinal de 3,83 % com tempo de pico de 4,4 segundos, tempo de acomodação 7,17 segundos e estabiliza a partir de 11 segundos. O erro do sistema é apresentado na Figura 5.29.

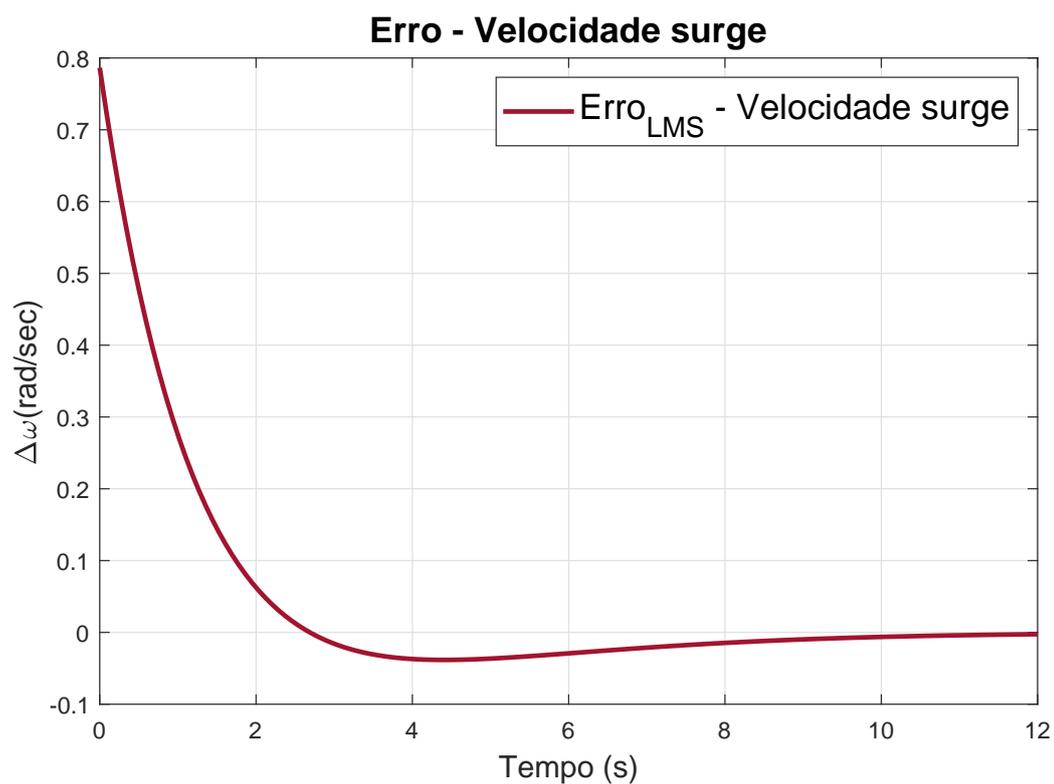


Figura 5.29 – Erro do sistema com controlador PID-LMS para velocidade  $u$ .

O erro do sistema apresenta valor inicial aproximadamente 0,8, e valor nulo em regime. Os sinais de pré processamento do erro são apresentados na Figura 5.30.

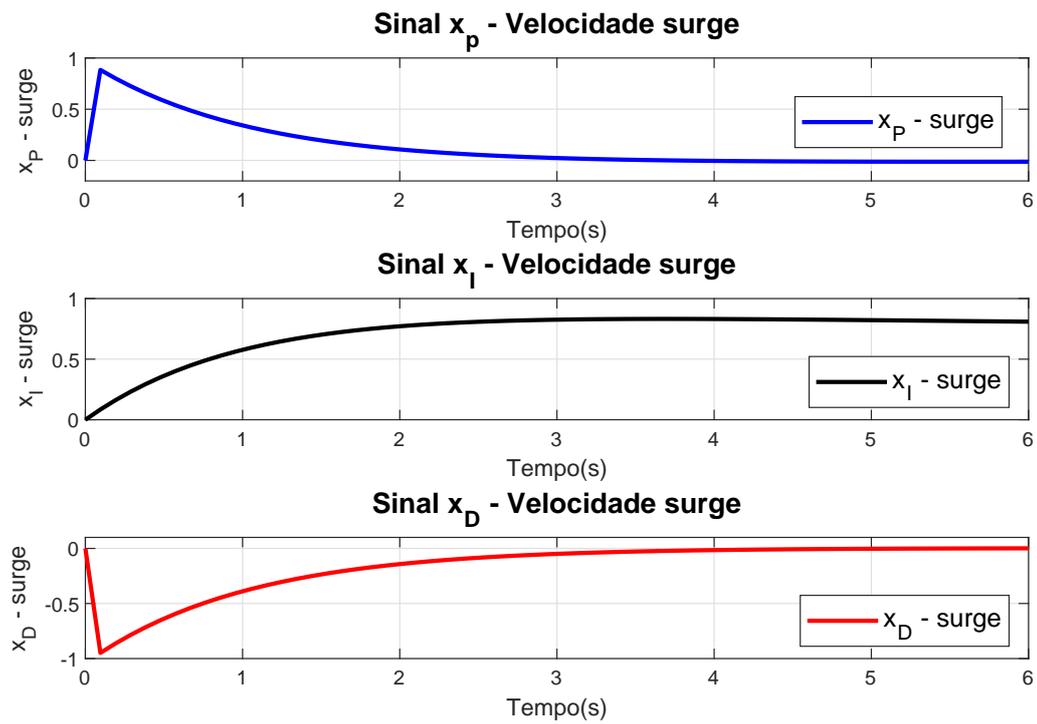


Figura 5.30 – Sinais de entrada do combinador linear do treinamento LMS.

Os sinais  $x_P$  e  $x_D$  apresentam comportamentos similares em sentidos opostos, com valores iniciais zero, maior valor de aproximadamente 1 nos primeiros instantes e convergência para zero a partir de 3 segundos. O sinal  $x_I$  apresenta valor inicial zero e converge para 0,75 a partir de 2,5 segundos.

A convergência dos ganhos é apresentada na Figura 5.31.

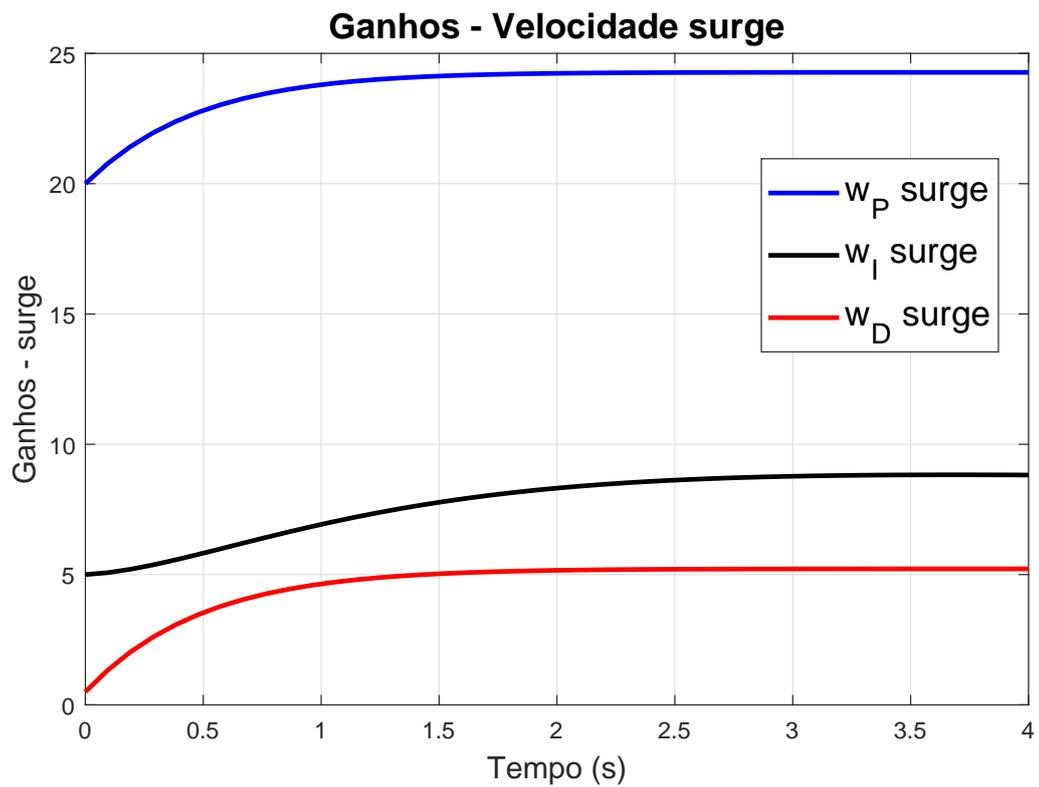


Figura 5.31 – Convergência dos ganhos do controlador PID-LMS para a velocidade  $u$ .

Os ganhos  $w_P$ ,  $w_I$  e  $w_D$  apresentam valores iniciais 20, 5 e 0,5 respectivamente.  $w_P$  e  $w_D$  convergem a partir de 1,5 com valores de regime 24,27 e 5, respectivamente. O ganho  $w_I$  converge para 8,18 a partir de 2,5 segundos.

O sinal de controle é apresentado na Figura 5.32.

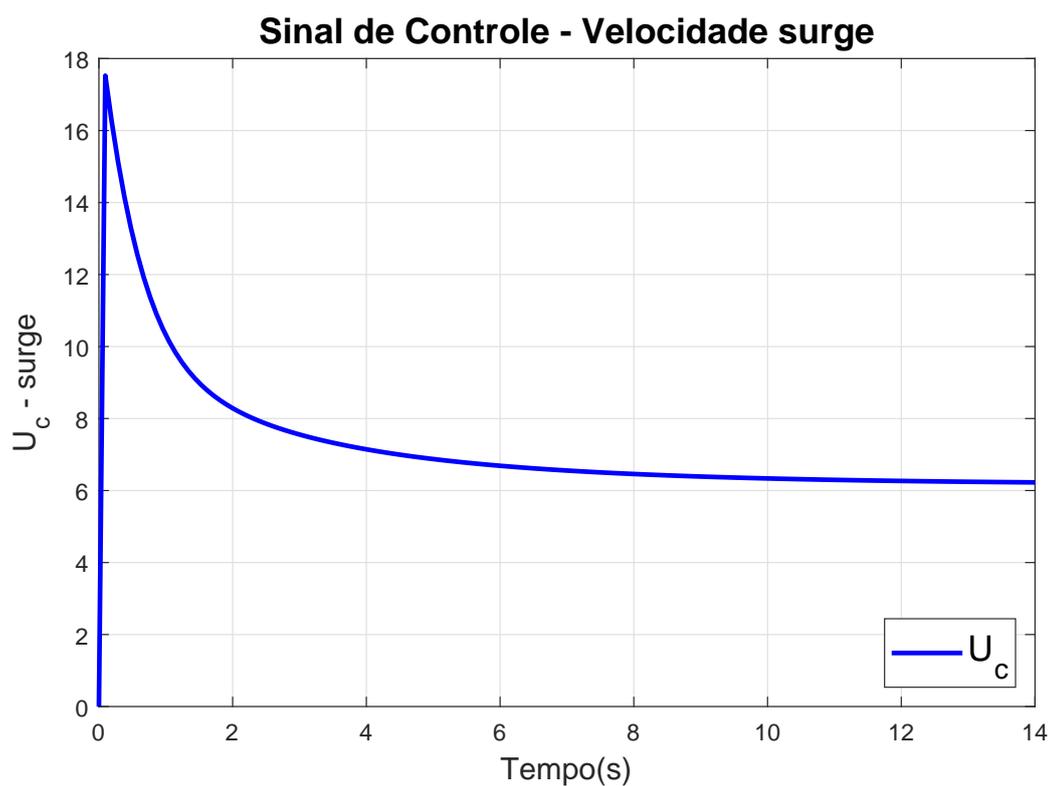


Figura 5.32 – Sinal de controle  $U_c$  para a velocidade  $u$ .

O sinal de controle inicia em zero e apresenta valor máximo de aproximadamente 17,5 nos primeiros instantes. A ação de controle apresenta valor de regime aproximadamente 6 a partir de 11 segundos.

A resposta para a velocidade  $u$  é apresentada na Figura 5.33

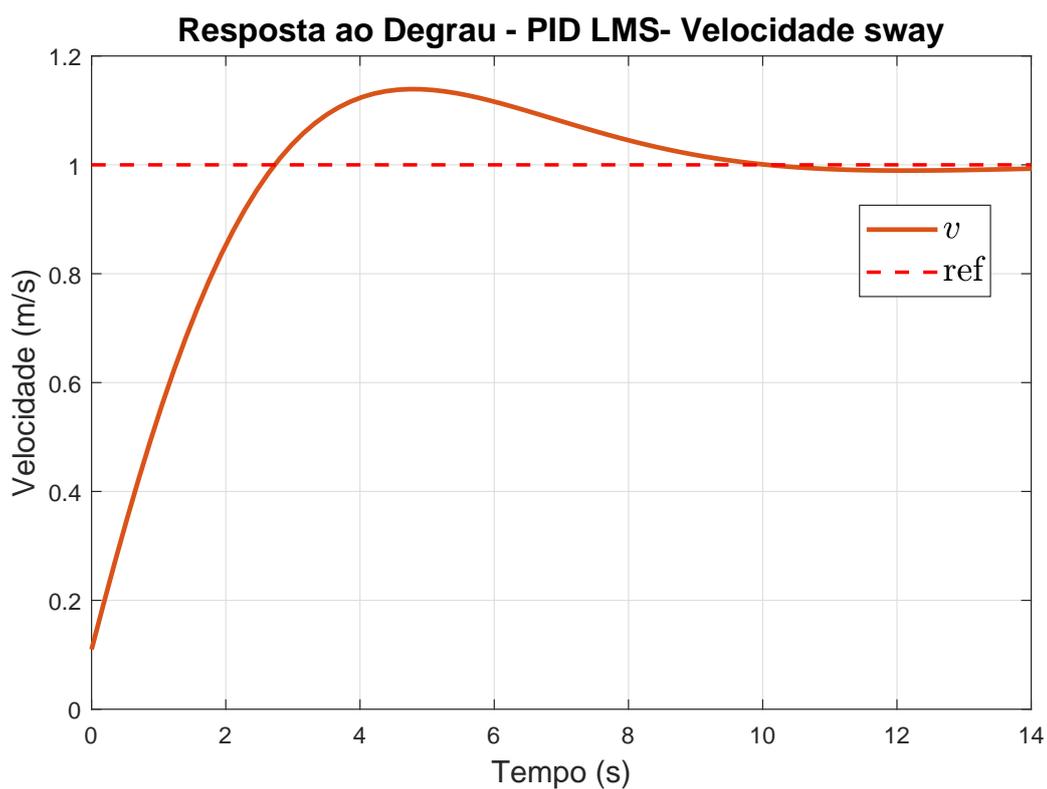


Figura 5.33 – Resposta ao degrau unitário do sistema com controlador PID-LMS para velocidade  $v$ .

A resposta apresentou tempo de subida 2,2s, sobressinal de 13,9 % com tempo de pico de 4,7 segundos, tempo de acomodação 8,9 segundos e estabiliza a partir de 12 segundos com erro de regime nulo. O erro do sistema é apresentado na Figura 5.34.

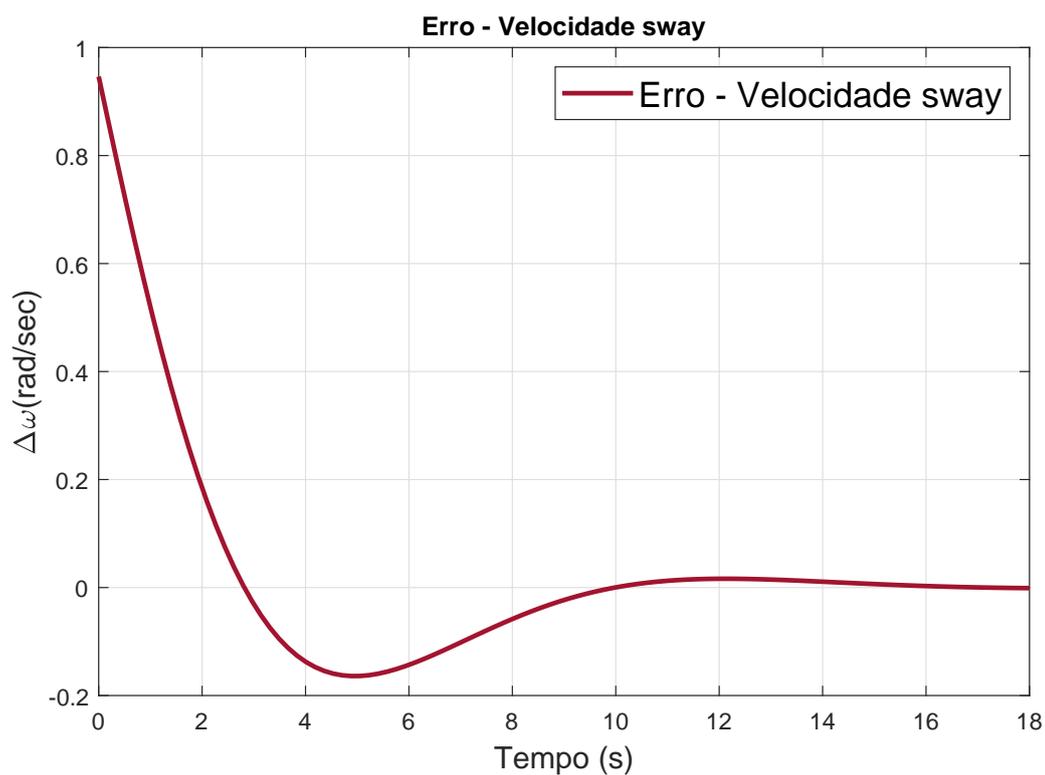


Figura 5.34 – Erro do sistema com controlador PID-LMS para velocidade  $v$ .

O erro do sistema apresenta valor inicial aproximadamente 0,9, e valor nulo em regime. Os sinais de pré processamento do erro são apresentados na Figura 5.35.

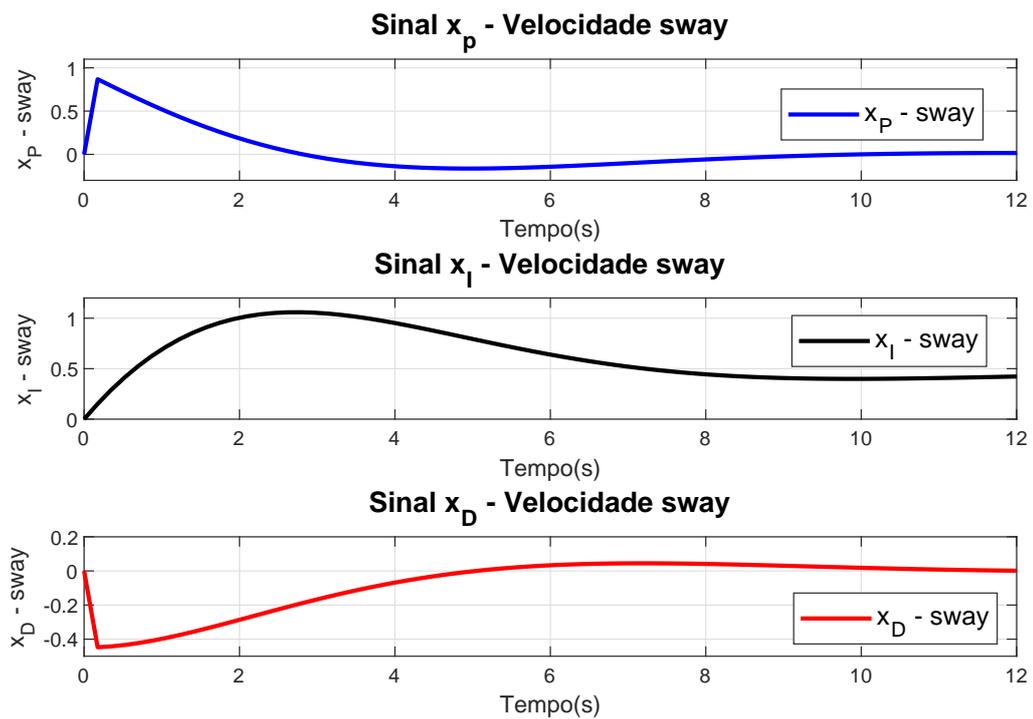


Figura 5.35 – Sinais de entrada do combinador linear de treinamento LMS para velocidade  $v$ .

Os sinais  $x_P$  e  $x_D$  apresentam comportamentos similares em sentidos opostos, com valores iniciais zero, convergência para zero a partir de 8 segundos e valores máximos aproximadamente nos primeiros instantes são aproximadamente 0,9 e -0,45 respectivamente. O sinal  $x_I$  apresenta valor inicial zero valor máximo de aproximadamente 1 no instante de 3 segundos e converge para 0,45 a partir de 9 segundos.

A convergência dos ganhos é apresentada na Figura 5.36.

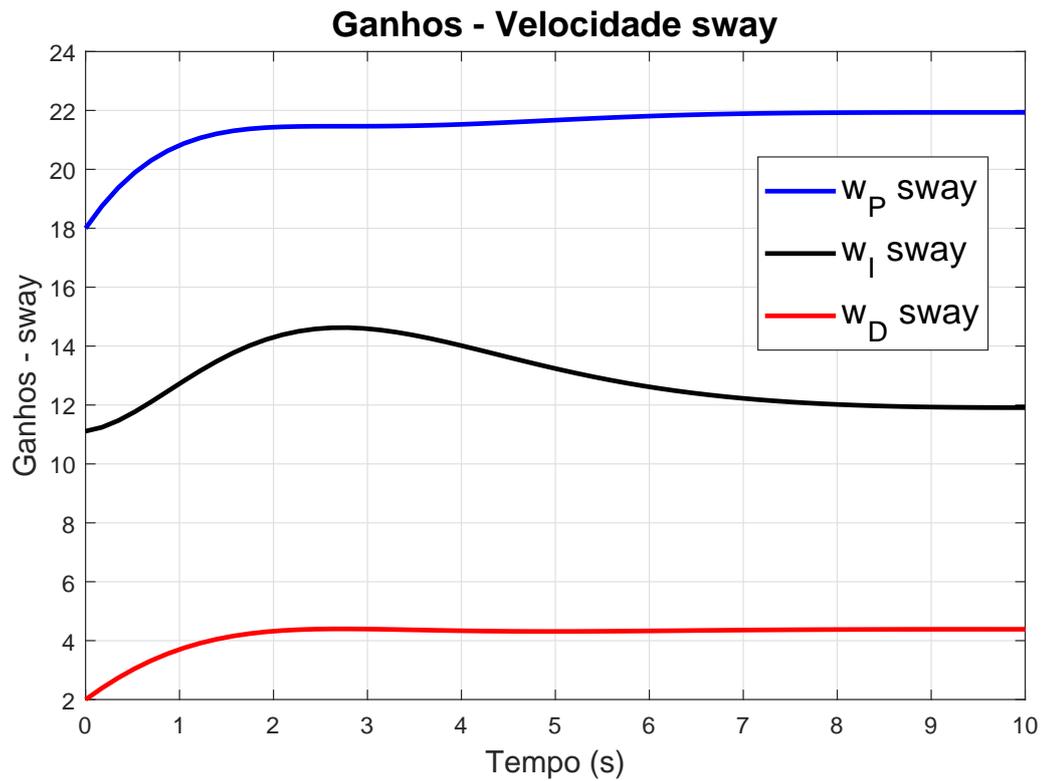


Figura 5.36 – Convergência dos ganhos do controlador PID-LMS para a velocidade  $v$ .

Os ganhos  $w_P$ ,  $w_I$  e  $w_D$  apresentam valores iniciais 18, 11,11 e 0,5 respectivamente.  $w_P$  converge para 22 a partir de 7 segundos,  $w_D$  converge a partir de 2 segundos para o valor de regime 4,38. O ganho  $w_I$  converge para 12 a partir de 8 segundos.

O sinal de controle é apresentado na Figura 5.37.

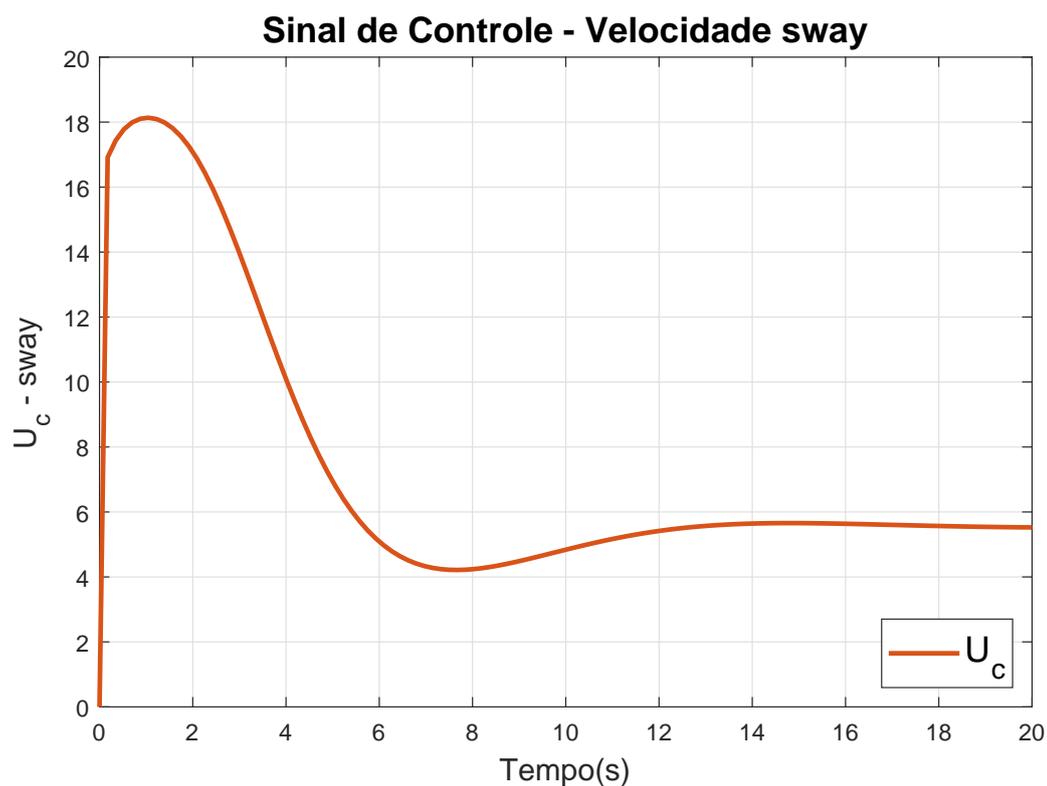


Figura 5.37 – Sinal de controle  $U_c$  para a velocidade  $v$ .

O sinal de controle inicia em zero e apresenta valor máximo de 18 no instante de 1 segundo. A ação de controle apresenta valor de regime aproximadamente 5,5 a partir de 14 segundos.

A resposta para a velocidade  $r$  é apresentada na Figura 5.38

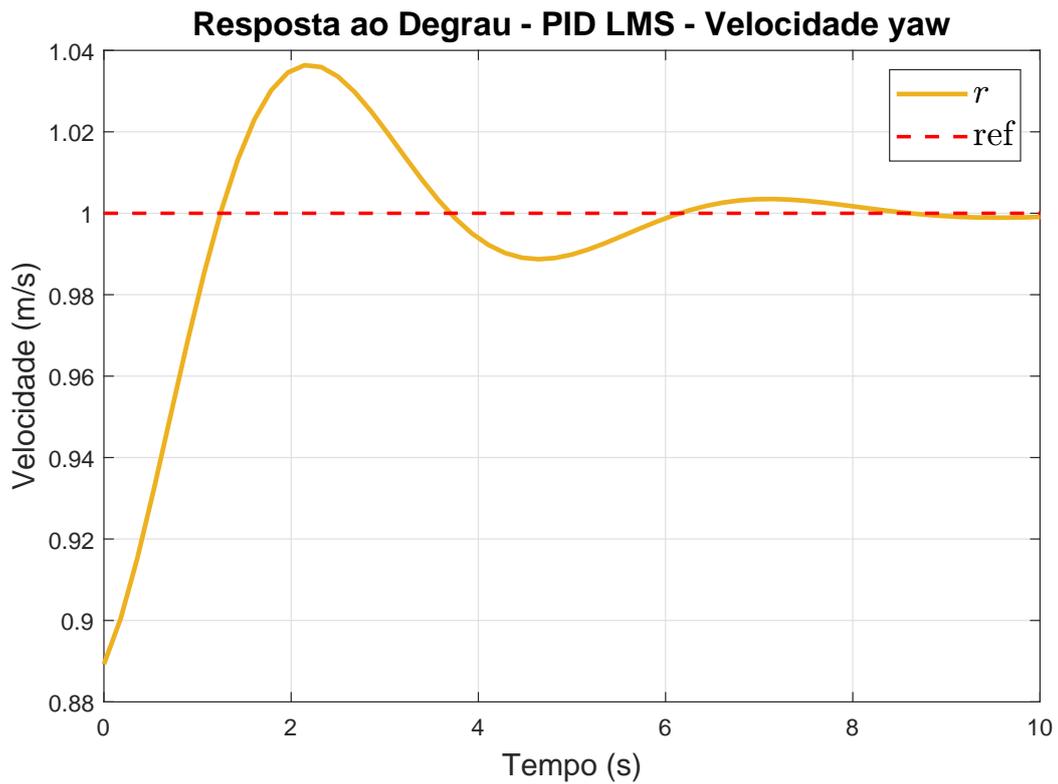


Figura 5.38 – Resposta ao degrau unitário do sistema com controlador PID-LMS para velocidade  $v$ .

A resposta apresentou tempo de subida 0,17s, sobressinal de 3,63 % com tempo de pico de 2,14 segundos, tempo de acomodação 3,02 segundos e entra em regime a partir de 9 segundos com erro de regime nulo. O erro do sistema é apresentado na Figura 5.39.

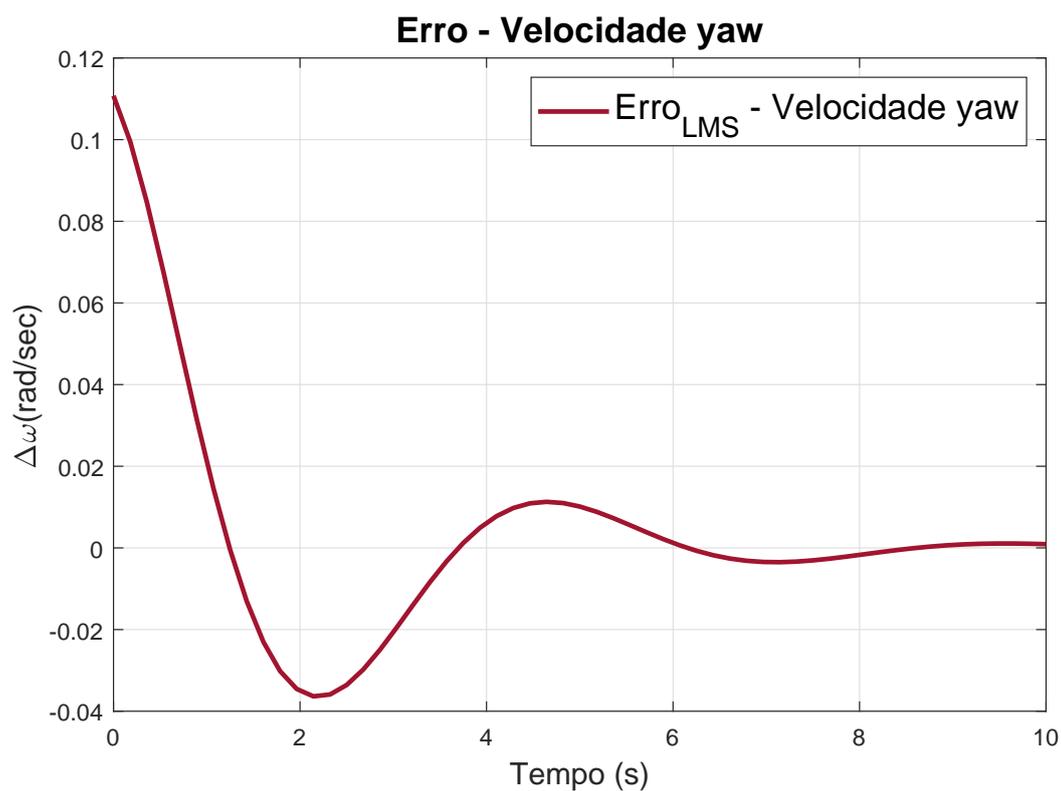


Figura 5.39 – Erro do sistema com controlador PID-LMS para velocidade  $v$ .

O erro do sistema apresenta valor inicial aproximadamente 0,11, e valor nulo em regime. Os sinais de pré processamento do erro são apresentados na Figura 5.40.

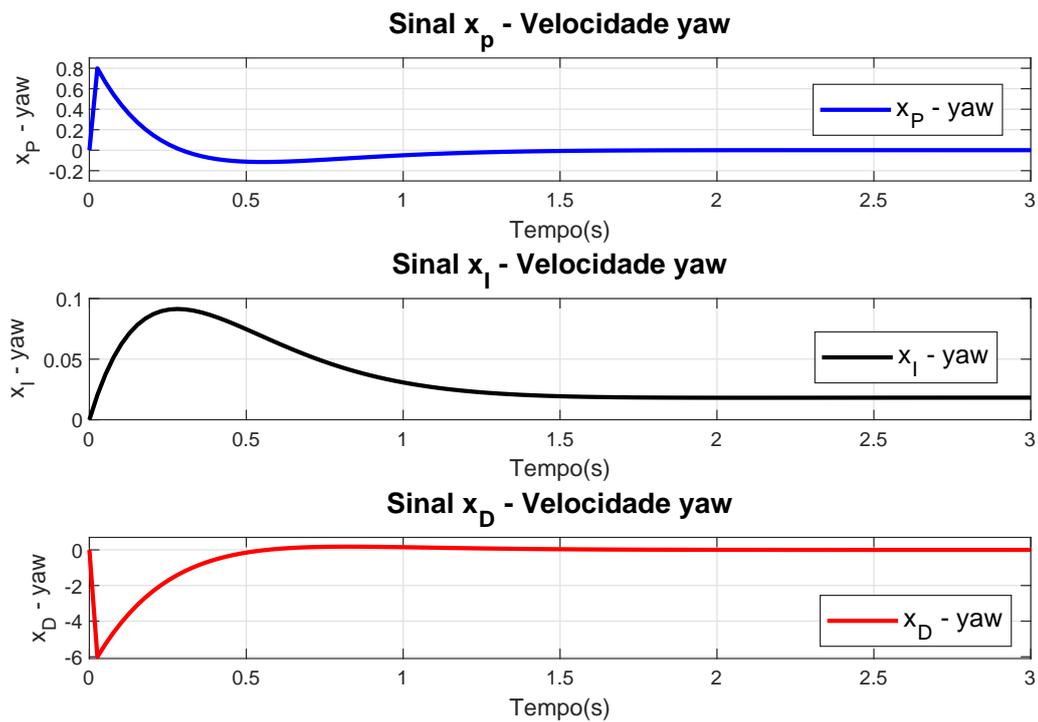


Figura 5.40 – Sinais de entrada do combinador linear de treinamento LMS para velocidade  $r$ .

Os sinais  $x_P$  e  $x_D$  apresentam comportamentos similares em sentidos opostos, com valores iniciais zero, convergência para zero a partir de 0,6 segundos e os valores máximos apresentados nos primeiros instantes são 0,8 e -6 respectivamente. O sinal  $x_I$  apresenta valor inicial zero, valor máximo de aproximadamente 0,1 no instante de 0,25 segundos e converge para 0,018 a partir de 1,5 segundos.

A convergência dos ganhos é apresentada na Figura 5.41.

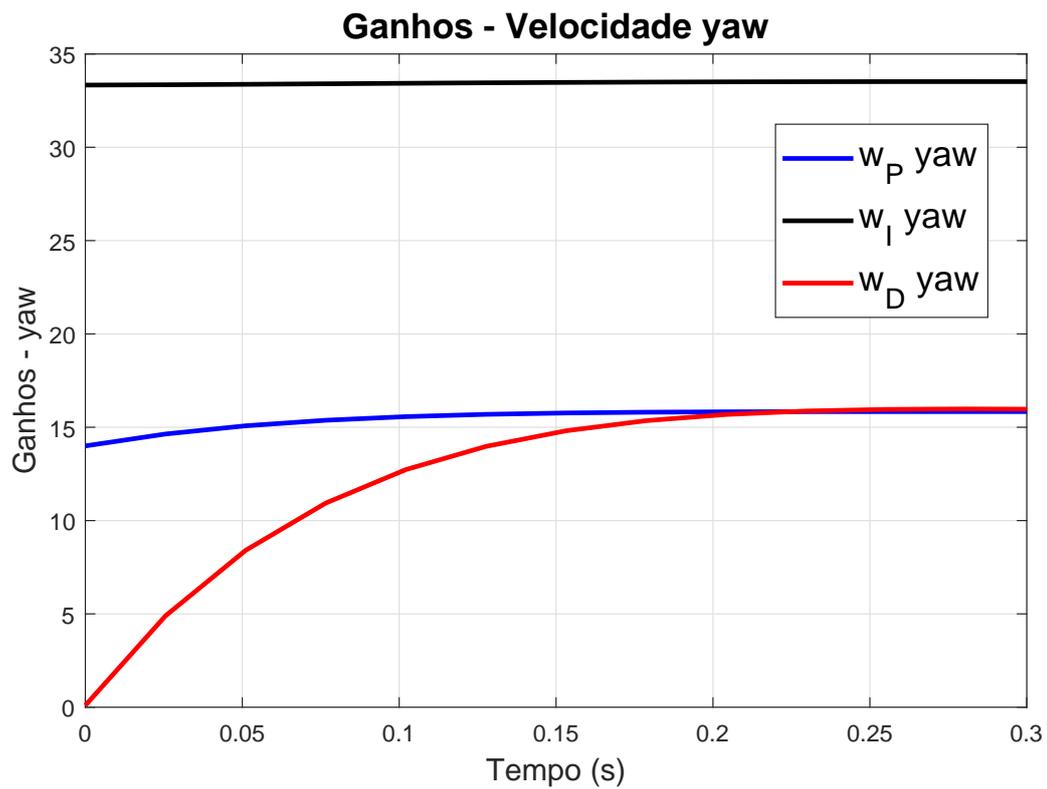


Figura 5.41 – Convergência dos ganhos do controlador PID-LMS para a velocidade  $r$ .

Os ganhos  $w_P$ ,  $w_I$  e  $w_D$  apresentam valores iniciais 14, 33,33 e 0,1 respectivamente.  $w_P$  converge para 16 a partir de 0,15 segundos,  $w_D$  converge a partir de 0,25 segundos para o valor de regime 15,92. O ganho  $w_I$  converge para 33,36 desde o início do processamento pois seu ganho inicial é muito próximo do valor de convergência.

O sinal de controle para a velocidade  $r$  é apresentado na Figura 5.42.

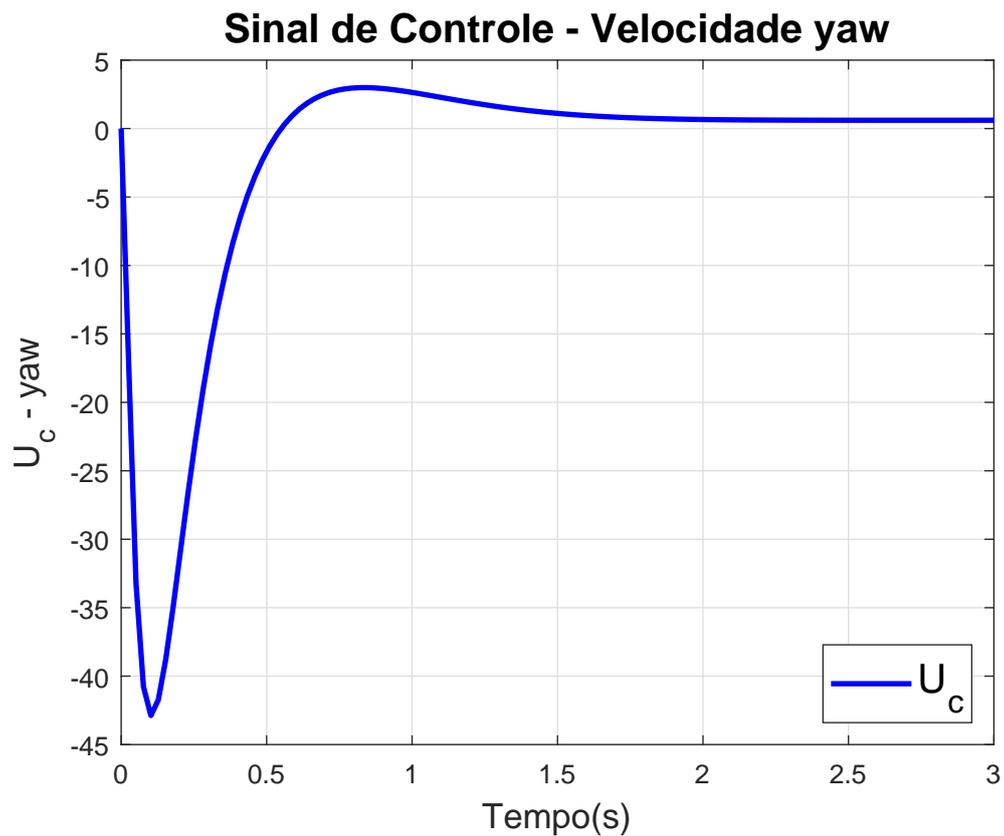


Figura 5.42 – Sinal de controle  $U_c$  para a velocidade  $r$ .

O sinal de controle apresentado na Figura 5.42 inicia em zero e apresenta um pico negativo de -42,87 no instante de 0,1 segundo. A ação de controle apresenta valor de regime aproximadamente 2,8 a partir de 2 segundos.

## 5.6 Comparação entre os métodos ADHDP-DLQT-I e PID-LMS

Os resultados obtidos para os dois controladores são comparados quanto às suas respostas, convergência dos ganhos e ações de controle. Avaliando a semelhança dinâmica entre os dois modelos da planta adotados, os resultados são comparados somente para o modelo  $x_{g_0}$ . A Figura 5.43 apresenta a comparação para a saídas do sistema  $u$ .

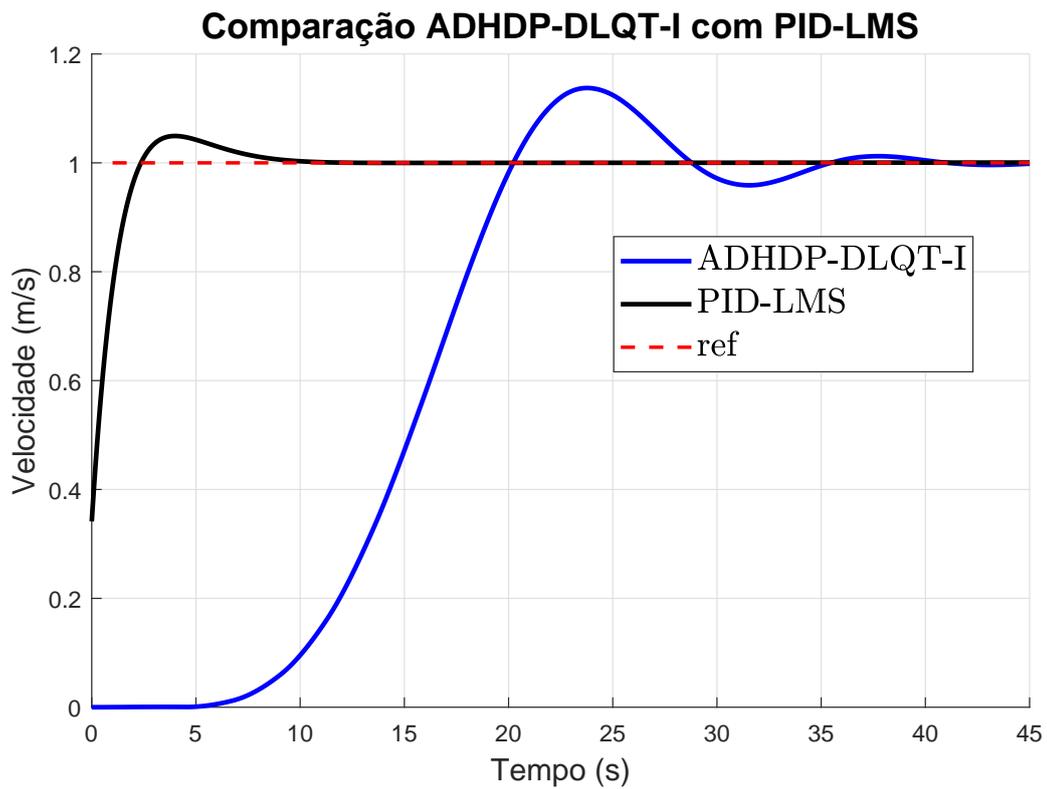


Figura 5.43 – Resposta comparativa dos controladores para a velocidade  $u$ .

A resposta para o controlador ADHDP-DLQT-I apresentado na Figura 5.43 apresenta um sobressinal maior, uma oscilação maior antes de estabilizar e um atraso maior na resposta. A Figura 5.44 apresenta a comparação para a saída do sistema  $v$ .

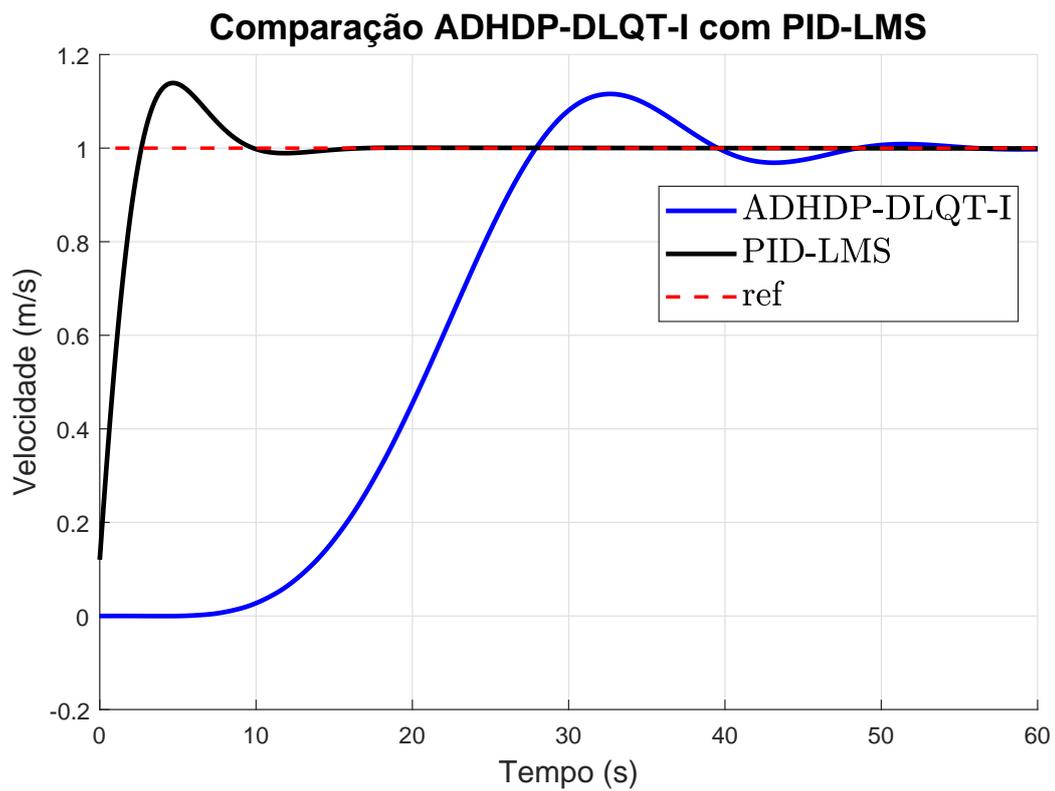


Figura 5.44 – Resposta comparativa dos controladores para a velocidade  $v$ .

A resposta do controlador ADHDP-DLQT-I apresentado na Figura 5.44 apresenta um sobressinal menor, uma oscilação maior antes de estabilizar e um atraso maior na resposta. A Figura 5.45 apresenta a comparação para a saída do sistema  $r$ .

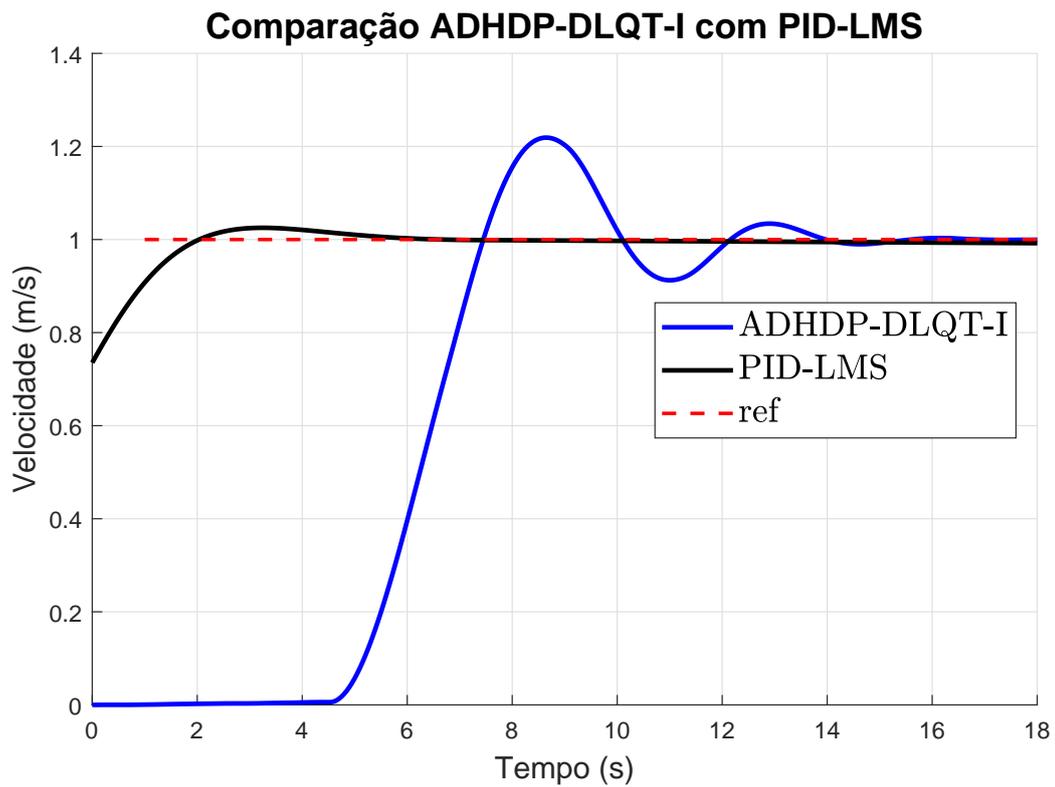


Figura 5.45 – Resposta comparativa dos controladores para a velocidade  $r$ .

A resposta do controlador ADHDP-DLQT-I apresentado na Figura 5.45 apresenta um sobressinal maior, uma oscilação maior antes de estabilizar e um atraso maior na resposta.

A comparação entre os sinais de controle é apresentada na Figura 5.46.

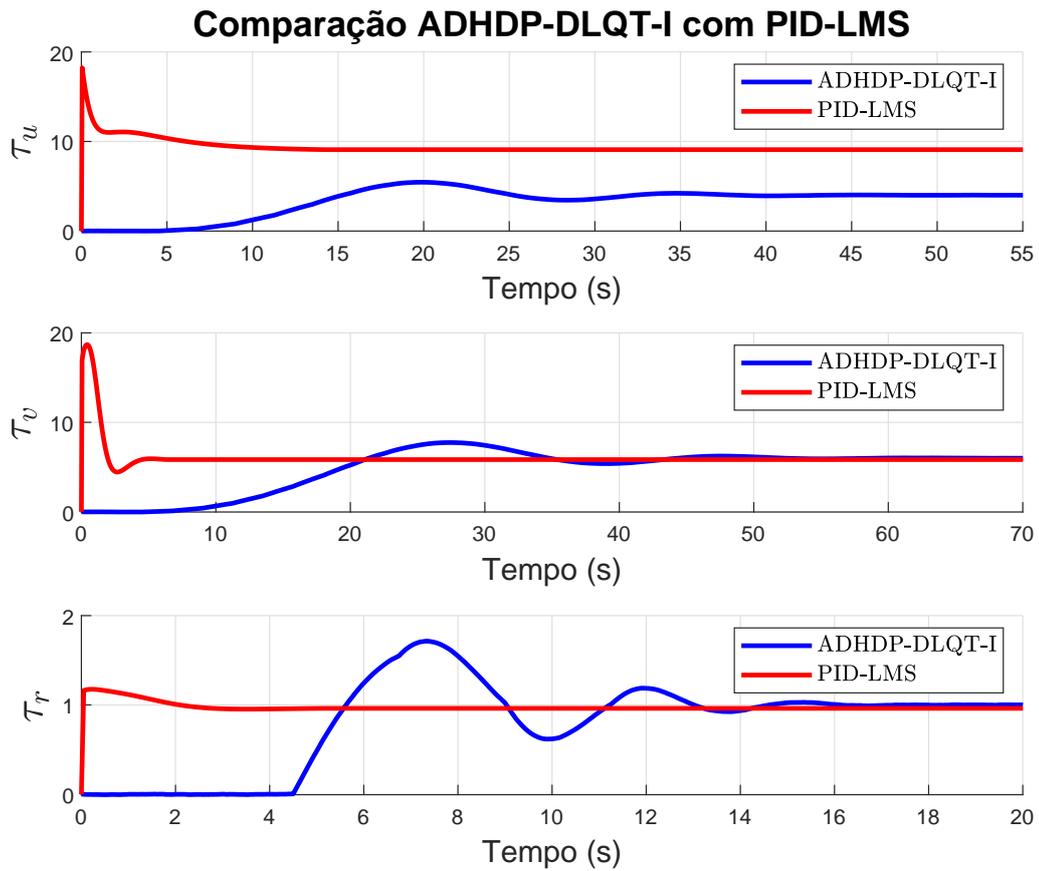


Figura 5.46 – Resposta comparativa das ações de controle para as velocidades  $u$ ,  $v$  e  $r$ .

As ações de controle apresentam algumas semelhanças, principalmente quantos aos seus valores de regime permanente para as velocidades  $v$  e  $r$ . As ações do controlador ADHDP-DLQT-I apresentam maiores oscilações e um tempo maior para sua convergência, principalmente pelo início do processo iterativo que leva um certo tempo para atualizar os ganhos, momento onde o sistema é controlado pelo ruído de controle.

A convergência dos ganhos do método PID-LMS acontece em tempo inferior ao método ADHDP-DLQT-I pela demora da ADHDP em atualizar os ganhos utilizando o método dos mínimos quadrados, a ADHDP realiza uma identificação do sistema a cada iteração o que torna o processo mais complexo. Apesar do menor tempo de convergência, o PID-LMS não garante uma resposta ótima para o sistema e seu objetivo é o desempenho do sistema. Uma resposta ótima leva em consideração uma característica essencial para o sistema, a energia. Uma análise do ponto de vista energético é mostra que, para uma mesma resposta em regime permanente, o PID-LMS apresentou um esforço de controle superior na Figura 5.46.

## CAPÍTULO 6

## CONCLUSÃO

Nesta dissertação foi apresentada uma metodologia para o projeto *online* de controladores ótimo DLQT-I baseado em Programação Dinâmica Heurística Dependente de Ação, a metodologia proposta foi avaliada no controle da manobrabilidade de um USV, tendo em vista aplicações no monitoramento aquático, como recurso, para mitigar os efeitos de derramamentos de petróleo e derivados.

A partir dos resultados obtidos pelos experimentos baseados em modelo para a simulação do controlador, constatou-se através da comparação com o método de Schur, a capacidade do algoritmo ADHDP para computar de forma aproximada a solução da equação de *Riccatti*, garantindo a solução do controlador ótimo. O projeto do controlador ADHDP-DLQT-I mostrou a viabilidade para aplicação *online* em manobrabilidade de USV's pela capacidade de rastrear a entrada de referência e eliminar o erro de regime, garantindo otimalidade mesmo com variações paramétricas.

A avaliação do controlador para o modelo não linear exibiu a discrepância no sentido de atingir o ponto de operação para uma aplicação do mundo real. A comparação entre os métodos ADHDP-DLQT-I e PID-LMS demonstrou o atraso na convergência dos ganhos para ADHDP, no entanto, apresentou uma resposta ótima após a convergência ponderando desempenho e custo energético.

### 6.1 Trabalhos Futuros

- Aplicação do controlador ADHDP-DLQT-I desenvolvido para rastreamento de trajetórias para complementar e validar o sistema de manobrabilidade em deslocamentos.
- Comparação do controlador com outros tipos de controladores adaptativos para avaliar a capacidade do controlador de forma mais ampla.

- Análise da sintonia QR para o controlador para verificar a relevância das matrizes no desempenho e consumo de energia.
- Desenvolvimento do controlador ADHDP-DLQT-I com observador de estados para se aproximar de uma aplicação onde não é possível medir as saídas.
- Análise da variação paramétrica para o controlador para comprovar a sua capacidade em compensá-las.
- Validação experimental em plataforma real e ambiente simulado com distúrbios.
- Aplicação de redes neurais profundas para aprimorar o crítico na estrutura ator-crítico.

## 6.2 Publicações

### 6.2.1 Principal Publicação

- FURTADO, V. G.; NETO, J. V. da F. Controlador ator-crítico para manobrabilidade de um USV baseado em DLQT-I e Programação Dinâmica Heurística Dependente de Ação. XXV Congresso Brasileiro de Automática. 2024.

### 6.2.2 Outras Publicações Relacionadas ao Tema

- SILVA, Y. F.; FONSECA NETO, J. V. ; FURTADO, V. G. . Smart Sensor Based On The IEEE 1451 Standard For Monitoring Coastal Areas Sensitive To Fuel Leakage. International Journal Of Advances In Engineering And Technology, v. 17, p. 303-316, 2024.

## REFERÊNCIAS

ABREU, F.; NETO, J. ao F. Controlador pid adaptativo via redes neurais-lms para controle de altitude de quadricoptero. In: FERNANDES, B. J. T.; {Pereira Júnior}, A. (Ed.). *Anais do 14 Congresso Brasileiro de Inteligência Computacional*. Curitiba, PR: ABRICOM, 2019. p. 1–6. ISBN 978-856997201-3. Citado na página 59.

AHMED, M. et al. Seal pipeline: Enhancing dynamic object detection and tracking for autonomous unmanned surface vehicles in maritime environments. *Drones*, MDPI, v. 8, n. 10, p. 561, 2024. Citado na página 17.

ALFARO-CID, E.; MCGOOKIN, E. W.; MURRAY-SMITH, D. J. Genetic algorithm optimisation of a supply ship propulsion and navigation systems. In: IEEE. *MTS/IEEE Oceans 2001. An Ocean Odyssey. Conference Proceedings (IEEE Cat. No. 01CH37295)*. [S.l.], 2001. v. 4, p. 2645–2652. Citado na página 60.

ALIFERIS, D. An analysis of potential use of unmanned surface vehicles (usv) in sea control activities. *Combined Joint Operations from the Sea Centre of Excellence*, 2021. Citado 2 vezes nas páginas 17 e 31.

AN, S.; HE, Y.; WANG, L. Adaptive neural fixed-time tracking control of underactuated usvs with external disturbances. In: IEEE. *2022 China Automation Congress (CAC)*. [S.l.], 2022. p. 357–362. Citado na página 27.

ANNAMALAI, A.; MOTWANI, A. A comparison between lqg and mpc autopilots for inclusion in a navigation, guidance and control system. *Marine and Industrial Dynamic Analysis*, School of Marine Science and Engineering, Plymouth University Plymouth, England, v. 1, p. 1–19, 2013. Citado na página 18.

ANNAMALAI, A. S. et al. Innovative adaptive autopilot design for uninhabited surface vehicles. IET, 2014. Citado 2 vezes nas páginas 18 e 27.

ASIF, Z. et al. Environmental impacts and challenges associated with oil spills on shorelines. *Journal of Marine Science and Engineering*, MDPI, v. 10, n. 6, p. 762, 2022. Citado na página 20.

ÅSTRÖM, K. J.; WITTENMARK, B. *Adaptive control* (2 rev. Dover Publications, 2008. Citado na página 19.

- BAE, I.; HONG, J. Survey on the developments of unmanned marine vehicles: intelligence and cooperation. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 23, n. 10, p. 4643, 2023. Citado na página 31.
- BELLMAN, R. Dynamic programming and stochastic control processes. *Information and control*, Elsevier, v. 1, n. 3, p. 228–239, 1958. Citado na página 51.
- BREIVIK, M. Topics in guided motion control of marine vehicles. NTNU, 2010. Citado 2 vezes nas páginas 17 e 18.
- CERQUEIRA, M.; LOPES, L.; FONSECA, J. Sintonia qr do regulador linear quadrático lqr discreto e programação dinâmica aproximada baseada em ação-estado para aplicações online do projeto de sistemas de controle ótimo. *Trends in Computational and Applied Mathematics*, SciELO Brasil, v. 25, p. e01686, 2024. Citado 2 vezes nas páginas 24 e 26.
- CHEN, L.; DAI, S.-L.; DONG, C. Adaptive optimal tracking control of an underactuated surface vessel using actor–critic reinforcement learning. *IEEE Transactions on Neural Networks and Learning Systems*, IEEE, 2022. Citado 2 vezes nas páginas 25 e 26.
- CID, M. E. A. *Optimisation of time domain controllers for supply ships using genetic algorithms and genetic programming*. Tese (Doutorado) — University of Glasgow, 2003. Citado na página 60.
- DIN, M. U. et al. Benchmarking vision-based object tracking for usvs in complex maritime environments. *arXiv preprint arXiv:2412.07392*, 2024. Citado na página 32.
- DO, K. D.; JIANG, Z.-P.; PAN, J. Robust global stabilization of underactuated ships on a linear course. In: IEEE. *Proceedings of the 2002 American control conference (IEEE Cat. No. CH37301)*. [S.l.], 2002. v. 1, p. 304–309. Citado na página 18.
- ELMAKIS, O. et al. Os-breeze: Oil spills boundary red emission zone estimation using unmanned surface vehicles. *Sensors*, MDPI, v. 24, n. 2, p. 703, 2024. Citado na página 20.
- EUZEBIO, C. S.; RANGEL, G. da S.; MARQUES, R. C. Derramamento de petróleo e seus impactos no ambiente e na saúde humana. *Revista Brasileira de Ciências Ambientais (RBCIAMB)*, v. 52, p. 79–98, 2019. Citado 2 vezes nas páginas 17 e 20.
- FALTINSEN, O. M.; SORTLAND, B. Slow drift eddy making damping of a ship. *Applied ocean research*, Elsevier, v. 9, n. 1, p. 37–46, 1987. Citado na página 39.
- FERREIRA, E. F. M. et al. Melhorias de estabilidade numérica e custo computacional de aproximadores de funções valor de estado baseados em estimadores rls para projeto online de sistemas de controle hdp-dlqr. Universidade Federal do Maranhão, 2016. Citado 2 vezes nas páginas 19 e 43.
- FOSSEN, T. I. *Nonlinear modelling and control of underwater vehicles*. [S.l.]: Universitetet i Trondheim (Norway), 1991. Citado na página 37.
- FOSSEN, T. I. Guidance and control of ocean vehicles. *Search in*, John Wiley & Sons, 1994. Citado na página 37.
- FOSSEN, T. I. Marine control systems—guidance, navigation, and control of ships, rigs and underwater vehicles. *Marine Cybernetics, Trondheim, Norway, Org. Number NO 985 195 005 MVA*, [www.marinecybernetics.com](http://www.marinecybernetics.com), ISBN: 82 92356 00 2, 2002. Citado 2 vezes nas páginas 33 e 35.

- FOSSSEN, T. I. *Handbook of marine craft hydrodynamics and motion control*. [S.l.]: John Wiley & Sons, 2011. Citado 10 vezes nas páginas 8, 33, 34, 35, 36, 37, 38, 40, 41 e 62.
- FOSSSEN, T. I.; FJELLSTAD, O.-E. Nonlinear modelling of marine vehicles in 6 degrees of freedom. *Mathematical Modelling of Systems*, Taylor & Francis, v. 1, n. 1, p. 17–27, 1995. Citado na página 42.
- FURTADO, V. G.; NETO, J. V. da F. Controlador ator-crítico para manobrabilidade de um usv baseado em dlqt-i e programação dinâmica heurística dependente de ação. In: *XXV Congresso Brasileiro de Automática*. [S.l.: s.n.], 2024. Citado na página 19.
- GADRE, A. S. et al. Guidance and control of an unmanned surface vehicle exhibiting sternward motion. In: IEEE. *2012 Oceans*. [S.l.], 2012. p. 1–9. Citado na página 18.
- GARRIDO, J.; VÁZQUEZ, F.; MORILLA, F. An extended approach of inverted decoupling. *Journal of process control*, Elsevier, v. 21, n. 1, p. 55–68, 2011. Citado 2 vezes nas páginas 73 e 75.
- GHAZALI, M. H. M.; SATAR, M. H. A.; RAHIMAN, W. Unmanned surface vehicles: From a hull design perspective. *Ocean Engineering*, Elsevier, v. 312, p. 118977, 2024. Citado na página 31.
- GILAT, A.; SUBRAMANIAM, V. *Métodos numéricos para engenheiros e cientistas: uma introdução com aplicações usando o MATLAB*. [S.l.]: Bookman Editora, 2009. Citado na página 57.
- HALVORSEN, H. *Dynamic positioning for unmanned surface vehicles*. Dissertação (Mestrado) — Institutt for teknisk kybernetikk, 2008. Citado 3 vezes nas páginas 27, 32 e 37.
- JEONG, J. S. et al. Modeling and control allocation for ship berthing system design. In: IEEE. *ICCAS 2010*. [S.l.], 2010. p. 195–200. Citado na página 60.
- KRAGELUND, S. et al. Adaptive speed control for autonomous surface vessels. In: IEEE. *2013 OCEANS-San Diego*. [S.l.], 2013. p. 1–10. Citado na página 18.
- LAUB, A. A Schur method for solving algebraic Riccati equations. *IEEE Transactions on automatic control*, IEEE, v. 24, n. 6, p. 913–921, 1979. Citado na página 78.
- LEFEBER, E.; PETTERSEN, K. Y.; NIJMEIJER, H. Tracking control of an underactuated ship. *IEEE transactions on control systems technology*, IEEE, v. 11, n. 1, p. 52–61, 2003. Citado na página 18.
- LEWIS, F. L.; VRABIE, D. Reinforcement learning and adaptive dynamic programming for feedback control. *IEEE circuits and systems magazine*, IEEE, v. 9, n. 3, p. 32–50, 2009. Citado 5 vezes nas páginas 8, 19, 52, 53 e 56.
- LEWIS, F. L.; VRABIE, D.; SYRMOS, V. L. *Optimal control*. [S.l.]: John Wiley & Sons, 2012. Citado 12 vezes nas páginas 23, 42, 43, 44, 45, 46, 47, 48, 51, 53, 54 e 55.
- LI, J. et al. A survey of maritime unmanned search system: Theory, applications and future directions. *Ocean Engineering*, Elsevier, v. 285, p. 115359, 2023. Citado na página 17.

- LIU, Z. et al. Unmanned surface vehicles: An overview of developments and challenges. *Annual Reviews in Control*, Elsevier, v. 41, p. 71–93, 2016. Citado 8 vezes nas páginas 8, 17, 18, 27, 30, 31, 32 e 33.
- MARTÍNEZ, J. R. et al. Design process and advanced manufacturing of an aquatic surface vehicle hull for the integration of a hydrogen power plant propulsion system. *Journal of Marine Science and Engineering*, MDPI, v. 12, n. 2, p. 268, 2024. Citado na página 32.
- MCGOOKIN, E. W. *Optimisation of Sliding Mode Controllers for Marine Applications: A study of methods and implementation issues*. Tese (Doutorado) — University of Glasgow, 1997. Citado 2 vezes nas páginas 60 e 61.
- MEINSMA, G.; SCHAFT, A. Van der. *A Course on Optimal Control*. [S.l.]: Springer, 2023. Citado na página 43.
- MENDOZA-CHOK, J. et al. Hybrid control architecture of an unmanned surface vehicle used for water quality monitoring. *IEEE Access*, IEEE, v. 10, p. 112789–112798, 2022. Citado na página 17.
- MOURA, J. P. D.; REGO, P. H. M.; NETO, J. V. da F. Online discrete-time LQR controller design with integral action for bulk bucket wheel reclaimer operational processes via action-dependent heuristic dynamic programming. *ISA transactions*, Elsevier, v. 90, p. 294–310, 2019. Citado 3 vezes nas páginas 18, 25 e 26.
- MU, C. et al. Air-breathing hypersonic vehicle tracking control based on adaptive dynamic programming. *IEEE transactions on neural networks and learning systems*, IEEE, v. 28, n. 3, p. 584–598, 2017. Citado 2 vezes nas páginas 23 e 26.
- NAVY, U. S. The navy unmanned surface vehicle (usv) master plan. 2007. Acessado em: 8 mar. 2025. Citado na página 18.
- NETO, J. V. D. F.; LOPES, L. R. On the convergence of DLQR control system design and recurrences of Riccati and Lyapunov in dynamic programming strategies. In: IEEE. *2011 UkSim 13th International Conference on Computer Modelling and Simulation*. [S.l.], 2011. p. 26–31. Citado na página 63.
- NI, Z.; HE, H.; WEN, J. Adaptive learning in tracking control based on the dual critic network design. *IEEE transactions on neural networks and learning systems*, IEEE, v. 24, n. 6, p. 913–928, 2013. Citado 2 vezes nas páginas 23 e 26.
- NIKOLA, T. *Method of and apparatus for controlling mechanism of moving vessels or vehicles*. [S.l.]: Google Patents, 1898. US Patent 613,809. Citado na página 17.
- NUNES, V. P. et al. Adaptive pid controllers tuning: Lms gain scheduling training and industrial programmable logic controllers. In: *Anais do 8 Congresso Brasileiro de Redes Neurais*. Florianópolis, SC: SBRN, 2007. p. 1–8. Citado na página 59.
- OGATA, K. *Engenharia de controle moderno. 5ª*. [S.l.]: São Paulo: Pearson, 2011. Citado na página 58.
- PAULO, T. F. Controle adaptativo com desacoplamento aplicado a um sistema de tanques acoplados mimo. 2015. Citado na página 72.

- PENG, Y.; HAN, J.; HUANG, Q. Adaptive ukf based tracking control for unmanned trimaran vehicles. *International Journal of Innovative Computing, Information and Control*, v. 5, n. 10, p. 3505–3516, 2009. Citado na página 18.
- PETTERSEN, K. Y.; FOSSEN, T. I. Underactuated ship stabilization using integral control: Experimental results with cybership i. In: CITESEER. *Proc. 1998 IFAC Symp. Nonlinear Control Systems Design*. [S.l.], 1998. p. 127–132. Citado na página 60.
- RÊGO, P. H. M. Aprendizagem por reforço e programação dinâmica aproximada para controle ótimo: Uma abordagem para o projeto online do regulador linear quadrático discreto com programação dinâmica heurística dependente de estado e ação. Universidade Federal do Maranhão, 2014. Citado 5 vezes nas páginas 19, 42, 51, 52 e 53.
- RUSCIO, D. L. D. Discrete LQ optimal control with integral action: A simple controller on incremental form for mimo systems. Norwegian Society of Automatic Control, 2012. Citado 3 vezes nas páginas 48, 49 e 50.
- SELMIC, R. R.; PHOHA, V. V.; SERWADDA, A. *Wireless Sensor Networks*. [S.l.]: Springer, 2016. Citado na página 17.
- SETIAWAN, F. et al. Dynamic modelling and controlling unmanned surface vehicle. In: IOP PUBLISHING. *IOP Conference Series: Earth and Environmental Science*. [S.l.], 2021. v. 649, p. 012056. Citado na página 40.
- SHAO, G. et al. A novel cooperative platform design for coupled usv–uav systems. *IEEE Transactions on Industrial Informatics*, IEEE, v. 15, n. 9, p. 4913–4922, 2019. Citado na página 17.
- SHUAI, L.; DAQIAN, W.; BIN, X. Adaptive tracking control for hypersonic flight vehicle using adhdhp. In: IEEE. *2022 4th International Conference on Data-driven Optimization of Complex Systems (DOCS)*. [S.l.], 2022. p. 1–6. Citado 2 vezes nas páginas 24 e 26.
- SILVA, F. N. D.; NETO, J. V. D. F. Data driven state reconstruction of dynamical system based on approximate dynamic programming and reinforcement learning. *IEEE Access*, IEEE, v. 9, p. 73299–73306, 2021. Citado 2 vezes nas páginas 24 e 26.
- SILVA, F. N. da; NETO, J. V. da F. Sintonia heurística e análise de convergência de algoritmo de aprendizagem por reforço para projeto de controle ótimo baseado em dados tuning heuristics and convergence analysis of reinforcement learning algorithm for online data-based optimal control design heurística de ajuste y análisis de convergencia del algoritmo de aprendizaje por. *Research, Society and Development*, v. 9, n. 2, p. e188922128, 2020. Citado 2 vezes nas páginas 25 e 26.
- SILVA, L. R. C. d. et al. Derramamento de petróleo no litoral brasileiro:(in) visibilidade de saberes e descaso com a vida de marisqueiras. *Ciência & Saúde Coletiva*, SciELO Brasil, v. 26, p. 6027–6036, 2021. Citado na página 17.
- SILVA, Y. F. d. Nó sensor inteligente para detecção e classificação de derivados de petróleo em ambiente aquático. Universidade Federal do Maranhão, 2024. Citado 2 vezes nas páginas 17 e 20.
- SNAME, T. Nomenclature for treating the motion of a submerged body through a fluid. *The Society of Naval Architects and Marine Engineers, Technical and Research Bulletin*, n. 1950, p. 1–5, 1950. Citado na página 34.

- SONNENBURG, C. R. *Modeling, Identification, and Control of an Unmanned Surface Vehicle*. Tese (Doutorado) — Virginia Polytechnic Institute and State University, dezembro 2012. Disponível em: <https://vtechworks.lib.vt.edu/server/api/core/bitstreams/f4264098-6cf7-431a-b959-e88595d827dd/content>. Citado na página 18.
- SONNENBURG, C. R.; WOOLSEY, C. A. Modeling, identification, and control of an unmanned surface vehicle. *Journal of Field Robotics*, Wiley Online Library, v. 30, n. 3, p. 371–398, 2013. Citado na página 18.
- SØRENSEN, A. J. Marine cybernetics, towards autonomous marine operations and systems. *Department of Marine Technology, NTNU*, 2018. Citado na página 40.
- STRAND, J. P. Nonlinear position control systems design for marine vessels. Citeseer, 1999. Citado na página 60.
- TOMERA, M.; PODGÓRSKI, K. Control of dynamic positioning system with disturbance observer for autonomous marine surface vessels. *Sensors*, MDPI, v. 21, n. 20, p. 6723, 2021. Citado na página 60.
- WERBOS, P. Approximate dynamic programming for real-time control and neural modeling. *Handbook of intelligent control*, Van Nostrand, 1992. Citado na página 19.
- WERBOS, P. J. Reinforcement learning and approximate dynamic programming (rladp)—foundations, common misconceptions, and the challenges ahead. *Reinforcement learning and approximate dynamic programming for feedback control*, Wiley Online Library, p. 1–30, 2012. Citado na página 52.
- WONDERGEM, M. et al. Output feedback tracking of ships. *IEEE transactions on control systems technology*, IEEE, v. 19, n. 2, p. 442–448, 2010. Citado na página 18.
- ZANOLI, S. M.; CONTE, G. Remotely operated vehicle depth control. *Control engineering practice*, Elsevier, v. 11, n. 4, p. 453–459, 2003. Citado na página 31.
- ZHANG, J.; YANG, P.; REN, X. Detection of oil spill in sar image using an improved deeplabv3+. *Sensors*, MDPI, v. 24, n. 17, p. 5460, 2024. Citado na página 20.
- ZHONG, Y. et al. Modeling and control of unmanned surface vehicles: An integrated approach. In: IEEE. *2023 8th International Conference on Automation, Control and Robotics Engineering (CACRE)*. [S.l.], 2023. p. 156–161. Citado na página 60.
- ZHOU, G. et al. Adaptive high-speed echo data acquisition method for bathymetric lidar. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, 2024. Citado na página 17.
- ZHOU, Z. et al. An offline reinforcement learning approach for path following of an unmanned surface vehicle. *Journal of Marine Science and Engineering*, MDPI, v. 12, n. 12, p. 2173, 2024. Citado na página 32.

# Apêndices

## APÊNDICE A

### MATRIZES DO DESACOPLAMENTO

As matrizes de desacoplamento  $D(s)$  e  $Q(s)$  apresentadas na Seção 4.5 são exibidas neste Apêndice. A matriz do desacoplador  $D(s)$  é dada por

$$D(s) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & d_{22}(s) & d_{23}(s) \\ 0 & d_{32}(s) & d_{33}(s) \end{bmatrix}, \quad (\text{A.1})$$

sendo

$$d_{22}(s) = \frac{1.007s^6 + 2.036s^5 + 1.632s^4 + 0.6586s^3 + 0.1404s^2 + 0.01508s + 0.0006418}{s^6 + 2.026s^5 + 1.627s^4 + 0.6577s^3 + 0.1404s^2 + 0.01508s + 0.0006418} \quad (\text{A.2})$$

$$d_{23}(s) = \frac{-0.3582s^6 - 0.5445s^5 - 0.3069s^4 - 0.0797s^3 - 0.009692s^2 - 0.0004484s + 8.945 \times 10^{-20}}{s^6 + 2.026s^5 + 1.627s^4 + 0.6577s^3 + 0.1404s^2 + 0.01508s + 0.0006418} \quad (\text{A.3})$$

$$d_{32}(s) = \frac{-0.02014s^6 - 0.03731s^5 - 0.0263s^4 - 0.008687s^3 - 0.001323s^2 - 7.477 \times 10^{-5}s - 9.976 \times 10^{-21}}{s^6 + 2.026s^5 + 1.627s^4 + 0.6577s^3 + 0.1404s^2 + 0.01508s + 0.0006418} \quad (\text{A.4})$$

$$d_{33}(s) = \frac{1.007s^6 + 2.036s^5 + 1.632s^4 + 0.6586s^3 + 0.1404s^2 + 0.01508s + 0.0006418}{s^6 + 2.026s^5 + 1.627s^4 + 0.6577s^3 + 0.1404s^2 + 0.01508s + 0.0006418} \quad (\text{A.5})$$

Observa-se que  $d_{22}$  e  $d_{33}$  são iguais e aproximadamente 1,  $d_{22} \approx 1$  e  $d_{33} \approx 1$ .

A matriz do sistema desacoplado  $Q(s)$  é dada por

$$Q(s) = \begin{bmatrix} \frac{0.0526}{s+0.2104} & 0 & 0 \\ 0 & q_{22} & q_{23} \\ 0 & q_{32} & q_{33} \end{bmatrix} \quad (\text{A.6})$$

sendo

$$q_{22} = \frac{q_{a22} + q_{b22} + q_{c22} + q_{d22}}{q_{pc}}, \quad (\text{A.7})$$

$$q_{23} = \frac{q_{a23} + q_{b23} + q_{c23} + q_{d23}}{q_{pc}}, \quad (\text{A.8})$$

$$q_{33} = \frac{q_{a32} + q_{b32} + q_{c32} + q_{d32} + q_{e32}}{q_{pc}}, \quad (\text{A.9})$$

$$q_{33} = \frac{q_{a33} + q_{b33} + q_{c33} + q_{d33}}{q_{pc}}, \quad (\text{A.10})$$

sendo  $q_{pc}$  o polinômio característico dado por

$$q_{pc} = q_{pc1} + q_{pc2} + q_{pc3} + q_{pc4} \quad (\text{A.11})$$

sendo

$$q_{pc1} = s^{16} + 5.404s^{15} + 13.46s^{14} + 20.52s^{13} + 21.39s^{12} \quad (\text{A.12})$$

$$q_{pc2} = 16.16s^{11} + 9.143s^{10} + 3.953s^9 + 1.319s^8 + 0.3408s^7 \quad (\text{A.13})$$

$$q_{pc3} = 0.06795s^6 + 0.01035s^5 + 0.001181s^4 + 9.769 \times 10^{-05}s^3 \quad (\text{A.14})$$

$$q_{pc4} = 5.526 \times 10^{-06}s^2 + 1.912 \times 10^{-07}s + 3.05 \times 10^{-09} \quad (\text{A.15})$$

e os demais elementos são dados por

$$q_{a22} = 0.0284s^{15} + 0.1485s^{14} + 0.3566s^{13} + 0.5209s^{12} + 0.5172s^{11}, \quad (\text{A.16})$$

$$q_{b22} = 0.3695s^{10} + 0.196s^9 + 0.07854s^8 + 0.02396s^7 + 0.005566s^6, \quad (\text{A.17})$$

$$q_{c22} = 0.000976s^5 + 0.0001269s^4 + 1.185 \times 10^{-05}s^3 + 7.507 \times 10^{-07}s^2, \quad (\text{A.18})$$

$$q_{d22} = 2.887 \times 10^{-08}s + 5.086 \times 10^{-10} \quad (\text{A.19})$$

$$q_{a23} = -1.735 \times 10^{-18}s^{15} + 4.163 \times 10^{-17}s^{13} + 1.11 \times 10^{-16}s^{12} \quad (\text{A.20})$$

$$q_{b23} = 8.327 \times 1.11 \times 10^{-16}s^{11} + 10^{-17}s^{10} + 4.163 \times 10^{-17}s^9 + 1.735 \times 10^{-17}s^8 \quad (\text{A.21})$$

$$q_{c23} = 3.903 \times 10^{-18}s^7 + 9.758 \times 10^{-19}s^6 + 1.491 \times 10^{-19}s^5 + 1.694 \times 10^{-20}s^4 \quad (\text{A.22})$$

$$q_{d23} = 1.376 \times 10^{-21}s^3 + 8.934 \times 10^{-23}s^2 + 3.671 \times 10^{-24}s + 7.088 \times 10^{-26} \quad (\text{A.23})$$

$$q_{a32} = -3.469 \times 10^{-18}s^{15} + 1.388 \times 10^{-17}s^{14} + 1.11 \times 10^{-16}s^{13} \quad (\text{A.24})$$

$$q_{b32} = 2.498 \times 10^{-16}s^{12} + 3.192 \times 10^{-16}s^{11} + 2.637 \times 10^{-16}s^{10} \quad (\text{A.25})$$

$$q_{c32} = 1.457 \times 10^{-16}s^9 + 5.725 \times 10^{-17}s^8 + 1.691 \times 10^{-17}s^7 \quad (\text{A.26})$$

$$q_{d32} = 3.361 \times 10^{-18}s^6 + 4.472 \times 10^{-19}s^5 + 3.642 \times 10^{-20}s^4 \quad (\text{A.27})$$

$$q_{e32} = 1.482 \times 10^{-21}s^3 - 3.309 \times 10^{-24}s^2 - 2.585 \times 10^{-24}s - 4.741 \times 10^{-26} \quad (\text{A.28})$$

$$q_{a33} = 0.5051s^{15} + 2.474s^{14} + 5.548s^{13} + 7.554s^{12} + 6.978s^{11} \quad (\text{A.29})$$

$$q_{b33} = 4.628s^{10} + 2.276s^9 + 0.8449s^8 + 0.2387s^7 + 0.05136s^6 \quad (\text{A.30})$$

$$q_{c33} = 0.008346s^5 + 0.001007s^4 + 8.741 \times 10^{-05}s^3 + 5.157 \times 10^{-06}s^2 \quad (\text{A.31})$$

$$q_{d33} = 1.851 \times 10^{-07}s + 3.05 \times 10^{-09} \quad (\text{A.32})$$

Os elementos  $q_{23}$  e  $q_{32}$ , apresentam valores muito pequenos, logo,  $q_{23} \approx 0$  e  $q_{32} \approx 0$ , tal que  $Q(s)$  é dada por

$$Q(s) = \begin{bmatrix} \frac{0.0526}{s+0.2104} & 0 & 0 \\ 0 & q_{22} & 0 \\ 0 & 0 & q_{33} \end{bmatrix}. \quad (\text{A.33})$$