



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO
EM ENGENHARIA QUÍMICA



Controle Tolerante com Reconfiguração Estrutural Acoplado a Sistema de Diagnóstico de Falhas

Lucas Lacerda Gomes Reis
Uberlândia
2008



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO
EM ENGENHARIA QUÍMICA



Controle Tolerante com Reconfiguração Estrutural Acoplado a Sistema de Diagnóstico de Falhas

Lucas Lacerda Gomes Reis

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Química da Universidade Federal de Uberlândia como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Química, Área de Concentração em Desenvolvimento de Processos Químicos.

Uberlândia
2008

Dados Internacionais de Catalogação na Publicação (CIP)

R375c Reis, Lucas Lacerda Gomes, 1982-

Controle tolerante com reconfiguração estrutural acoplado a sistema de diagnóstico de falhas / Lucas Lacerda Gomes Reis. - 2008. 199 f. : il.

Orientador: Luís Cláudio Oliveira Lopes.

Dissertação (mestrado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Engenharia Química. Inclui Bibliografia.

1. Controle de processo - Teses. 2. Controladores PID - Teses. I. Lopes, Luís Cláudio Oliveira. II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Engenharia Química. III. Título.

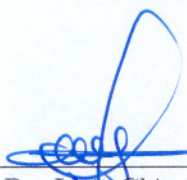
CDU 681.51

Elaborada pelo Sistema de Bibliotecas da UFU / Setor de Catalogação e Classificação

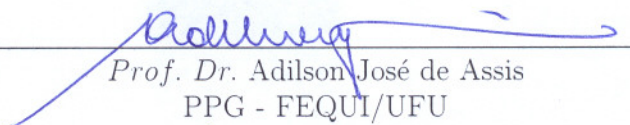
Controle Tolerante com Reconfiguração Estrutural Acoplado a Sistema de Diagnóstico de Falhas

Dissertação de mestrado submetida ao corpo docente do Programa de Pós-Graduação em Engenharia Química da Universidade Federal de Uberlândia como parte dos requisitos necessários para obtenção do grau de mestre em Engenharia Química em 26 de Março de 2008.

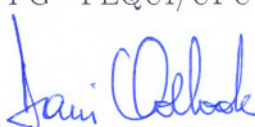
BANCA EXAMINADORA



Prof. Dr. Luis Cláudio Oliveira Lopes
Orientador - PPG - FEQUI/UFU



Prof. Dr. Adilson José de Assis
PPG - FEQUI/UFU



Prof. Dr. Darci Odloak
DEQ/EQ/USP

Agradecimentos

Aos meus pais, Luiz Carlos e Vânia pelo apoio que eu e meus irmãos tivemos desde os primeiros dias na escola.

A minha namorada Camila Emilia Figueira pelo apoio nos momentos difíceis e também pelos momentos divertidos durante este período tão importante de minha vida.

Ao Professor Luís Cláudio Oliveira Lopes pela orientação deste trabalho e de tantos outros, pela compreensão nos momentos críticos e principalmente pela amizade.

Aos meus amigos e companheiros de curso Diogo, Róbson, Daniel, Cássia, Cícero, Davi, David, Thiago, Fran Sérgio, Ricardo Corrêa, Luciano, pela força e incentivo ao longo do meu trabalho.

Aos professores da Faculdade de Engenharia Química.

Aos membros da banca, Prof. Dr. Adilson José de Assis e Prof. Dr. Darci Odloak, pelo enriquecimento deste trabalho.

À CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pela concessão da bolsa de estudos.

Aos meus Pais Luiz Carlos e Vânia.

SUMÁRIO

Lista de Figuras	xiii
Lista de Tabelas	xvii
Lista de Abreviaturas	xix
Notação Matemática	xxi
Glossário	xxiii
Resumo	xxvi
Abstract	xxix
1 Introdução	1
2 Monitoramento de Processos	5
2.1 Monitoramento e Controle Tolerante	5
2.2 Análise dos Componentes Principais	6
2.2.1 Introdução	6
2.2.2 Descrição da Técnica PCA	7
2.2.3 Análise de Redução de Ordem	10
2.2.4 Estatísticas Multivariáveis para Detecção de Falhas	11

2.2.5	Diagnóstico de Falhas para Estatísticas Multivariáveis	14
2.2.6	PCA Dinâmica	17
2.2.7	Resumo das Técnicas PCA e DPCA	18
2.3	Análise Discriminante de Fisher	18
2.3.1	Introdução	19
2.3.2	Descrição da Técnica FDA	19
2.3.3	Detecção e Diagnóstico de Falhas	22
2.3.4	Comparação entre PCA e FDA	24
2.3.5	Resumo da Técnica FDA	26
2.4	Análise Estrutural	26
2.4.1	Introdução	26
2.4.2	Construção Estrutural do Modelo	27
2.4.3	Pareamento: definições e aplicações	30
2.4.4	Decomposição de Grafos	36
2.4.5	Observabilidade estrutural	36
2.4.6	Monitorabilidade	37
3	Propostas de Controle Tolerante a Falhas	39
3.1	Introdução ao Problema de Controle Tolerante	39
3.1.1	Falhas na lei de controle	40
3.1.2	Proposta de controle tolerante	41
3.1.3	Controle tolerante ativo versus passivo	42
3.1.4	Utilização de informações sobre falhas	43
3.1.5	Estratégias de controle tolerante ativo	43
3.2	Controle Preditivo Tolerante a Falhas	44
3.3	Controle Ótimo Tolerante a Falhas	45
3.3.1	Controle ótimo regulador	46
3.3.2	Controle ótimo servo	47
3.3.3	Reconfiguração do controle	50

3.4	Controle Tolerante com <i>Model-Matching</i>	51
3.4.1	O método da pseudo-inversa	53
3.4.2	Projeto <i>model-matching</i> para falhas em sensores	54
3.4.3	Projeto <i>model-matching</i> para falhas em atuadores	56
3.5	Controle PID Tolerante a Falhas	57
3.6	Análise Estrutural e o Controle Tolerante a Falhas	58
3.6.1	Tolerância a falhas estruturais	59
3.6.2	Tolerância a falhas não estruturais	59
4	Resultados e Discussões	61
4.1	Controle Tolerante para Sistemas Discretos	62
4.1.1	Introdução	62
4.1.2	Exemplo 1: Controle Ótimo Servo	62
4.1.3	Exemplo 2: Controle Preditivo	70
4.1.4	Discussão dos Resultados	75
4.2	Coluna de Destilação Binária Ideal	77
4.2.1	Introdução	77
4.2.2	Modelagem matemática e controle	77
4.2.3	Estudo de falhas em processos de destilação	82
4.2.4	Deteção, identificação e diagnóstico de falhas	85
4.2.5	Controle tolerante para colunas de destilação	89
4.2.6	Discussão dos resultados	89
4.3	O Processo de Produção de Ciclopentanol	92
4.3.1	Descrição do Problema	92
4.3.2	Características do Processo	94
4.3.3	Simulação do Processo	95
4.3.4	Deteção e Diagnóstico de Falhas	99
4.3.5	Análise Estrutural	104
4.3.6	Controle Tolerante do Processo de Produção de Ciclopentanol . . .	110

4.3.7	Resultados e Discussões	115
4.4	O Reator de Neutralização	119
4.4.1	Modelo Físico	119
4.4.2	Modelagem do Reator de Neutralização	121
4.4.3	Controle do Processo	124
4.4.4	Falhas no Reator de Neutralização	126
4.4.5	Resultados e Discussão	131
5	Conclusões e Sugestões	133
A	Linearização Sucessiva	135
B	Controle Preditivo Baseado em Modelo	139
B.1	Introdução	139
B.2	Controle Preditivo para Sistema Estáveis	141
B.3	Função Objetivo do Controlador Preditivo	141
B.3.1	Conexão entre modelo e função objetivo	144
B.4	Restrições Aplicadas ao Problema de Otimização	146
B.4.1	Restrição Aplicada à Entrada do Processo	147
B.4.2	Restrição Aplicada à Variação na Entrada do Processo	147
B.4.3	Restrição Aplicada para a Saída do Processo	148
B.5	MPC Regulador com Restrições	151
B.6	MPC Servo com Restrições	153
B.7	MPC Regulador Sucessivamente Linearizado	155
B.7.1	Restrições aplicadas ao modelo sucessivamente linearizado	158
B.7.2	Conexão entre modelo e função objetivo	159
B.8	MPC Servo Sucessivamente Linearizado	161
	Referências Bibliográficas	163

LISTA DE FIGURAS

1.1	Situação do controle tolerante a falhas em 1997 (PATTON, 1997) (a) e proposta de contribuição deste trabalho (b).	3
1.2	Esquema de um processo com controle tolerante a falhas.	4
2.1	Revisão sobre monitoramento e controle tolerante a falhas.	6
2.2	Comparação entre PCA e FDA para uma transformação linear dos dados nas classes $\times$ e $\circ$ sobre os primeiros vetores característicos FDA e PCA. . .	25
2.3	Grafo bipartido do sistema linear do Exemplo 2.1.	29
3.1	Proposta de um sistema de controle tolerante a falhas utilizada neste trabalho.	41
3.2	Proposta de um sistema de controle preditivo tolerante a falhas.	45
3.3	Princípio da reconfiguração do controle.	50
3.4	Controlador tolerante a falhas por reconfiguração (BLANKE et al., 2006). .	52
3.5	Estrutura de alteração lógica entre um banco de controladores.	58
4.1	Resposta dinâmica do processo de separação de metanol da água devido a uma mudança de <i>set-point</i>	65
4.2	Simulação de uma falha no atuador da vazão de refluxo (R_F) para o exemplo da coluna de separação metanol-água.	66
4.3	Simulação de uma falha no atuador da vazão de refluxo (R_F) com controle tolerante ativo para o exemplo da coluna de separação metanol-água. . . .	67
4.4	Simulação de uma falha no sensor de concentração de metanol no produto de fundo para o exemplo da coluna de separação metanol-água.	68

4.5	Controle tolerante para falha no sensor de concentração de metanol no produto de fundo.	69
4.6	Resposta dinâmica do sistema devido a uma mudança de <i>set-point</i> do Exemplo 2.	71
4.7	Resposta do Exemplo 2 com um atuador extra utilizando MPC.	72
4.8	Simulação da perda do primeiro atuador (primeira coluna da matriz B) para o Exemplo 2.	73
4.9	Controle tolerante ativo para falha em um atuador do processo do Exemplo 2.	74
4.10	Simulação de um atuador com agarramento para o Exemplo 2.	75
4.11	Controle tolerante para falha em um atuador com agarramento (Exemplo 2).	76
4.12	Pratos em uma coluna de destilação.	78
4.13	Esquema de uma coluna de destilação.	78
4.14	Estruturas controle da coluna de destilação binária ideal.	81
4.15	Resultados dos problemas de controle servo e regulador para a coluna de destilação binária.	83
4.16	Resultados para operação em controle regulatório na estrutura de controle RV da coluna de destilação binária ideal.	84
4.17	Simulação de uma falha na estrutura DB (atuador B).	85
4.18	Simulação de uma falha na estrutura RB (atuador B) para a coluna de destilação binária ideal.	88
4.19	Deteção e diagnóstico utilizando FDA para a estrutura de controle RB com falha no atuador B.	88
4.20	Simulação de uma falha em D para o processo controlado pela estrutura DV.	90
4.21	Controle tolerante para falha em D na estrutura DV. Troca de estrutura de controle (de DV para RV).	91
4.22	Esquema do processo investigado.	93
4.23	Superfície de resposta para a concentração de ciclopentanol.	95
4.24	Simulação da Falha 01 para o reator de produção de ciclopentanol.	96
4.25	Simulação da Falha 02 para o reator de produção de ciclopentanol.	97
4.26	Simulação da Falha 03 para o reator de produção de ciclopentanol.	98
4.27	Simulação da Falha 04 para o reator de produção de ciclopentanol.	98

4.28	Interface gráfica do simulador desenvolvido em linguagem Tcl/Tk para o processo de produção de ciclopentanol.	99
4.29	Operação normal do processo de produção de ciclopentanol.	100
4.30	Estatísticas T^2 e Q para operação normal do reator de produção de ciclopentanol.	102
4.31	Estatísticas T^2 e Q para as Falhas 01 e 02 no reator de produção de ciclopentanol.	103
4.32	Estatísticas T^2 e Q para as Falhas 03 e 04 no reator de produção de ciclopentanol.	103
4.33	Comportamento do processo com controle tolerante para o reator de produção de ciclopentanol.	111
4.34	Estatísticas T^2 e Q para a Falha 04 no reator de produção de ciclopentanol.	112
4.35	Vazão da válvula redundante após ativação do controle tolerante.	112
4.36	Comportamento do processo com controle tolerante para o reator de produção de ciclopentanol (Falha 02).	115
4.37	Estatísticas T^2 e Q para a falha no sensor de concentração no reator de produção de ciclopentanol (Falha 02).	116
4.38	Comportamento do processo com controle tolerante para o reator de produção de ciclopentanol (Falha 03).	117
4.39	Estatísticas T^2 e Q para a falha no sensor de temperatura do reator de produção de ciclopentanol (Falha 03).	118
4.40	Desenho esquemático do reator de neutralização.	122
4.41	Mapeamento entradas (ácido e base) saída (pH) para o processo de neutralização. Vazão de solução tampão nula.	123
4.42	Controle do processo de neutralização próximo à região de pH=7.	124
4.43	Controle do processo de neutralização utilizando como variáveis manipuladas a vazão de ácido e de solução tampão (vazão de base constante e igual a 5 ml/s).	125
4.44	Vazão de solução tampão como variável manipulada em um esquema em que vazão de base é constante e igual a 5 ml/s.	125
4.45	Simulação de uma falha no atuador que alimenta base ao reator de neutralização.	127
4.46	Resíduos entre o sinal de entrada calculado pelo MPC e o sinal lido através do sensor de vazão de base (ml/s).	127

4.47	Controle tolerante para o reator de neutralização com falha na válvula de alimentação de base.	128
4.48	Esquema de controle tolerante proposto para acomodação de falhas em atuadores para o reator de neutralização.	129
4.49	Manipulação da vazão de solução tampão no esquema de controle tolerante a falhas na válvula de alimentação de base.	129
4.50	Controle tolerante a falhas com mudança completa dos objetivos de controle (nível e pH).	130
4.51	Vazão de solução tampão utilizada como variável manipulada devido a alteração de objetivos de controle.	131
B.1	Horizonte móvel do controle preditivo baseado em modelo.	152

LISTA DE TABELAS

1	Notação matemática utilizada.	xxi
2.1	Resumo das técnicas PCA e DPCA aplicadas para detecção e diagnóstico de falhas.	19
2.2	Dados originais e pré-tratados utilizados no exemplo comparativo entre PCA e FDA.	25
2.3	Resumo da técnica FDA para diagnóstico de falhas.	26
4.1	Operação em estado estacionário da coluna para separação metanol-água do Exemplo 1.	62
4.2	Parâmetros utilizados pelo o controlador PID servo e regulador da coluna de destilação binária.	82
4.3	Variáveis medidas na coluna de destilação binária.	83
4.4	Falhas nas estruturas de controle.	85
4.5	Taxa de classificações incorretas (TCI) percentual para os dados do conjunto de treinamento para cada modelo FDA e redução de ordem para a coluna de destilação binária.	86
4.6	Resultados do algoritmo FDA proposto para monitoramento em tempo real para dados independentes do conjunto de treinamento para a coluna de destilação binária.	87
4.7	Parâmetros do modelo do reator de produção de ciclopentanol.	94
4.8	Parâmetros do controlador PI regulador para o processo de produção de ciclopentanol.	95
4.9	Matriz de incidência do modelo do processo de produção de ciclopentanol.	108

- 4.10 Matriz de incidência do modelo com pareamento obtido através do algoritmo de posicionamento para o processo de produção de ciclopentanol. . . 109
- 4.11 Relações de redundância analíticas deduzidas para o processo de produção de ciclopentanol. 114
- 4.12 Resultados das simulações sem o modo de controle tolerante ativo. 40 horas de simulação, tempo de amostragem de 3 minutos. 116

LISTA DE ABREVIATURAS

AIC - *Akaike's Information Criterion*, Critério de Informação de Akaike.

AR - Modelo Multivariável Autoregressivo.

CUSUM *Cumulative Sum*, Soma Acumulativa.

CVA *Canonical Variate Analysis*, Análise Canônica Variada.

DPCA - *Dynamic PCA*, PCA Dinâmica.

EWMA *Exponentially Weighted Moving Average*, Ponderação Exponencial de Média Móvel.

FDA - *Fisher Discriminant Analysis*, Análise Discriminante de Fisher.

FDD *Fault Detection and Diagnosis*, Detecção de Falhas e Diagnóstico.

FDID *Fault Detection, Identification and Diagnosis*, Detecção de Falhas, Identificação e Diagnóstico.

ICA *Independent Component Analysis*, Análise dos Componentes Independentes.

MPC - *Model Predictive Control*, Controle Preditivo Baseado em Modelo.

PCA - *Principal Component Analysis*, Análise dos Componentes Principais.

PID - Proporcional-Integral-Derivativo.

PRESS - *Prediction Residual Sum of Squares*, Predição da Soma Residual dos Quadrados.

SPE - *Squared Prediction Error*, Erro de Predição ao Quadrado.

SVD - *Singular Value Decomposition*, Decomposição dos Valores Singulares.

NOTAÇÃO MATEMÁTICA

Neste trabalho, a representação de matrizes, vetores e escalares bem como outros tipos de variáveis nas formulações matemáticas está de acordo com a Tabela (1):

Tabela 1: Notação matemática utilizada.

Tipo	Exemplos
Matrizes	$\mathbf{A}$ ou $\mathbf{\Gamma}$
Matrizes ou Vetores de Números	$\mathbf{0}$ (matriz ou vetor de zeros)
Vetores ou Funções Vetoriais	$\mathbf{a}$ ou $\boldsymbol{\gamma}$
Escalares ou Funções	a ou γ
Conjuntos	$\mathbb{R}^*$ ou $\mathbb{Z}^+$
Grafos (Análise Estrutural)	$\mathcal{R}$ ou $\mathcal{V}$
Variáveis	A ou Γ
Índices	i ou l

GLOSSÁRIO

O Comitê Técnico de Processos Seguros (*Safe-process Technial Committee*) do IFAC e a Federação Internacional de Controle Automático (*International Federation of Automatic Control*) compilaram uma lista de definições sugeridas quando se trata de controle tolerante a falhas e monitoramento de processos (ISERMANN; BALLÉ, 1996).

Acomodação de falha: A ação de alterar uma lei de controle em resposta a uma falha, sem alterar qualquer componente do sistema. Na acomodação de falha, componentes com falhas são mantidos em operação devido a uma lei de controle adaptada.

Ação corretiva: Uma ação de correção (reconfiguração ou uma mudança na operação de um sistema) que previne que uma certa falha se propague para situações indesejadas.

Análise de propagação de falha: Análise que permite determinar como certos efeitos de falhas se propagam através de um sistema.

Análise estrutural: Análise das propriedades estruturais de um modelo, isto é, propriedades que são independentes dos valores reais dos parâmetros.

Detecção de falha: Determinação de falhas presentes em um sistema e o tempo de detecção.

Detector de falha: Um algoritmo que desempenha detecção e isolamento de falha.

Diagnóstico de falha: Determinação do tipo, tamanho, local e tempo de ocorrência de uma falha. Diagnóstico de falha inclui detecção, isolamento e estimação.

Efeito de falha: A consequência de um modo de falha sobre a operação, função, ou estado de um item.

Erro: Desvio entre um valor medido ou calculado (de uma variável de saída) e o valor real especificado ou valor teoricamente correto.

Estimação de falha: Determinação de um modelo do sistema com falha.

Falha: Desvio não permitido de no mínimo uma propriedade característica variável ou parâmetro de um sistema de sua condição aceitável, usual ou padrão. Uma falha é a ocorrência de um modo de falha.

Identificação de falha: Determinação do tamanho e comportamento com o tempo de uma falha.

Isolação de falha: Determinação da localização de uma falha.

Limite: Valor limite de um desvio de resíduo do zero, se excedido, uma falha é declarada como detectada.

Modelagem de falha: Determinação de um modelo matemático para descrever um efeito de falha específico.

Modo de falha: Modo particular que uma falha pode ocorrer.

Objetivo: Uma especificação de controle, a parte principal de um sistema de controle.

Reconfiguração de objetivo: A ação de alterar o objetivo do sistema de controle. A reconfiguração de objetivo é prioritária quando falhas não recuperáveis para os objetivos correntes ocorrem.

Reconfiguração do sistema: A ação de trocar componentes com falhas de acordo com a lei de controle, em resposta a uma falha. Na reconfiguração do sistema, componentes com falhas não são mais empregados.

Recuperação da falha: O resultado de um sistema de acomodação de falha bem sucedido.

Redundância Analítica: Uso de dois ou mais, mas não necessariamente os mesmos, caminhos para determinar uma variável em que um caminho usa um modelo matemático do processo na forma analítica.

Redundância física: Uso de mais de um instrumento independente para conseguir uma dada função.

Resíduo: Informações de falhas carregas por sinais, baseados no desvio entre medidas e cálculos por modelos.

Restrição: Limitação imposta pela natureza (leis físicas) ou pelo homem. Permite que as variáveis recebam somente valores determinados no espaço variável.

Sistema tolerante a falhas: Um sistema em que uma falha é recuperada com ou sem degradação de desempenho.

Sistema de controle tolerante ativo: Um sistema tolerante a falhas em que as falhas são explicitamente detectadas e acomodadas. Oposto de um sistema tolerante a falhas passivo.

Sistema de controle tolerante a falha passivo: Um sistema tolerante a falhas em que falhas não são explicitamente detectadas e acomodadas, mas o controlador é projetado para ser insensível a certos conjuntos de falhas restritos. Contrário de um sistema de controle tolerante ativo.

Supervisão: Monitoração de um sistema físico e tomada de ações apropriadas para manter a operação no caso de falhas.

Supervisor: Uma função que realiza supervisão usando resultados de diagnóstico de falhas, determina ações corretivas quando necessário e executa ações corretivas para manipular falhas.

Resumo

A exigência do aumento da disponibilidade de plantas industriais é cada vez mais presente no parque industrial atual. A disponibilidade representa a manutenção satisfatória da operação da planta mesmo após o aparecimento de falhas. Um sistema com essas características é definido como um sistema tolerante a falhas. Para a concretização de um sistema tolerante a falhas é importante a implementação de técnicas de monitoramento automático capazes de fornecer informações precisas sobre o estado do processo e formar a base necessária para a recuperação da condição diagnosticada. Essa estrutura de monitoramento funciona como a fonte principal de informações para sistemas de supervisão responsáveis por ações corretivas que acomodem efeitos causados por falhas. Neste trabalho, técnicas de detecção e diagnóstico de falhas baseadas em dados de operação da planta e em modelos matemáticos dos processos são utilizadas para o monitoramento do estado de processos industriais. As técnicas de controle tolerante propostas neste trabalho baseiam-se na união de diferentes estratégias de monitoramento, na reconfiguração do processo ou controladores envolvidos e adaptação dos objetivos de controle para a obtenção de novas estruturas de controle. Assim, o sistema proposto adiciona características de tolerância a falhas através do acoplamento da detecção e diagnóstico de falhas a um sistema de reconfiguração estrutural do sistema de controle. Nesse contexto, este trabalho apresenta propostas de controle tolerante a falhas ilustrando com sucesso a sua aplicação através do estudo de casos em estruturas de controle PID, controle ótimo e controle preditivo.

Palavras-chave: *Controle Tolerante, Monitoramento, Detecção de Falhas, Diagnóstico, MPC.*

Abstract

The requirement of increasing the availability of plants is present in the current industrial park. The availability represents the satisfactory operation of a plant even during process failures. A system with these characteristics is defined as a system fault-tolerant. For the implementation of a fault-tolerant system is important to use automatic monitoring techniques capable of providing accurate information on the state of the process and form the basis for the recovery of the diagnosed condition. This structure functions as the main source of information for supervisory systems responsible for corrective actions to accommodate effects caused by failures. In this work, techniques for fault detection and diagnosis based on data from the plant operation and mathematical models of processes are used to monitor the state of industrial processes. The tolerant control techniques proposed in this work is based on the union of different strategies for process monitoring, a process and control reconfiguration and control goals adaptation to obtain new control structures. Thus, the proposed system will add features for fault tolerance by coupling fault detection and diagnosis with the structural reconfiguration of the control system. In this context, this work presents a novel fault-tolerant control scheme illustrating their successful application through case studies on PID control, optimal control and predictive control structures.

Keywords: *Fault-Tolerant Control, Monitoring, Fault Detection and Diagnosis, MPC.*

CAPÍTULO 1

Introdução

O crescente aumento da necessidade de investimento em segurança, melhoria da confiabilidade e elevação dos lucros de processos químicos depende diretamente de sistemas de controle completamente automatizados. A automação, apesar de necessária em processos em crescimento e expansão da capacidade de monitoramento e controle, pode aumentar a quantidade de problemas em instrumentos eletrônicos devido a falhas nos mesmos (defeitos em equipamentos como sensores, por exemplo). Entre os problemas relacionados a falhas tem-se: dano físico aos equipamentos do processo, desperdício de matéria-prima, de fontes de energia e aumento no tempo de desligamento, resultando em perdas significantes da produção e colocando em risco a segurança de pessoas e do meio ambiente. Estas considerações fornecem uma justificativa para a dedicação de esforços no desenvolvimento de métodos que possibilitem o projeto de sistemas de controle capazes de tolerar falhas e que ao mesmo tempo garantam uma resposta eficiente e uma recuperação do processo em tempo adequado, prevenindo a propagação de falhas.

A grande integração presente nos processos industriais os tornam cada vez mais complexos e as falhas que neles ocorrem são problemas (detecção, identificação e diagnóstico das falhas) que não são prontamente resolvidos pela estatística monovariável. A ineficácia das cartas de controle monovariáveis para detectar falhas em processos multivariáveis tem levado ao surgimento de uma pesquisa concentrada no desenvolvimento de melhores métodos de monitoramento (CHIANG et al., 2001). O crescimento desta atividade de pesquisa pode também ser explicado pelo fato da grande quantidade de dados disponível para uso no monitoramento de processos e pelo aumento da capacidade de pro-

cessamento de informações e resolução de problemas com elevado esforço computacional durante operações em linha (*on-line*). A disponibilidade de dados obtidos durante várias operações, incluindo as com falhas é essencial ao monitoramento, visto que as métricas provenientes de estatísticas multivariáveis utilizam essas informações para obter modelos úteis para detecção e diagnóstico de falhas.

Dessa forma, pode-se constatar que, para buscar elevados padrões de qualidade, os processos industriais atuais contêm um grande número de variáveis operando em laço fechado e utilizam para isto diferentes tipos de estratégias de controle. Os controladores de processos convencionais (controladores do tipo PID (Proporcional-Integral-Derivativo), MPC (*Model Predictive Control*, Controle Preditivo Baseado em Modelo) *etc.*) são projetados para manter de forma satisfatória as condições operacionais através da compensação de efeitos de perturbações e mudanças que acontecem no processo. Enquanto esses controladores podem ser eficazes para muitos tipos de perturbações, existem mudanças que eles não podem lidar adequadamente. Dentre essas mudanças encontram-se as falhas, que são definidas como desvios não permitidos de pelo menos uma propriedade característica ou variável do sistema (ISERMMAN; BALLÉ, 1996). Assim, a implementação de sistemas de controle que estejam acoplados com gerenciadores de detecção e diagnóstico de falhas é uma tendência atual para o projeto de controladores, fornecendo maior segurança às operações controladas.

Para se obter processos controlados que operem conjuntamente com sistemas de monitoramento é necessário a união de diferentes temas em controle de processos, monitoramento de eventos anormais e controle robusto. Em meados dos anos 80 esta união se intensificou e resultou numa nova área de pesquisa, o controle tolerante a falhas (ZHANG; JIANG, 2003). Verifica-se que grande parte da literatura inicial é destinada à solução de problemas relacionados com sistemas de controle de vôos de aeronaves. No entanto, verifica-se que atualmente existe um número crescente de aplicações de controle tolerante a falhas para processos da Engenharia Química. Isto ocorre devido fato da requisição elevada pela qualidade dos produtos e segurança de operação com a rápida evolução da capacidade de processamento e armazenamento de informações.

A Figura (1.1)(a) mostra como a combinação de diferentes áreas de pesquisa resultou no controle tolerante a falhas (representado pela área hachurada). O estado da arte do controle tolerante a falhas no ano de 1997 (PATTON, 1997) mostrou que à medida que as demais áreas (Controle Robusto, Controle Reconfigurável e Detecção, Identificação e Diagnóstico de Falhas) se desenvolviam e exploravam cada vez mais problemas complexos e diferentes, tornava-se interessante unir diferentes estratégias de controle e monitoramento de processos para habilitar processos para a tolerância a falhas. A Figura (1.1)(b) mostra a proposta de contribuição deste trabalho, que além de explorar técnicas de detecção e

diagnóstico de falhas, utiliza-as com diferentes tipos de controladores, tais como, PID, ótimo e principalmente com o controlador preditivo linear baseado em modelo.

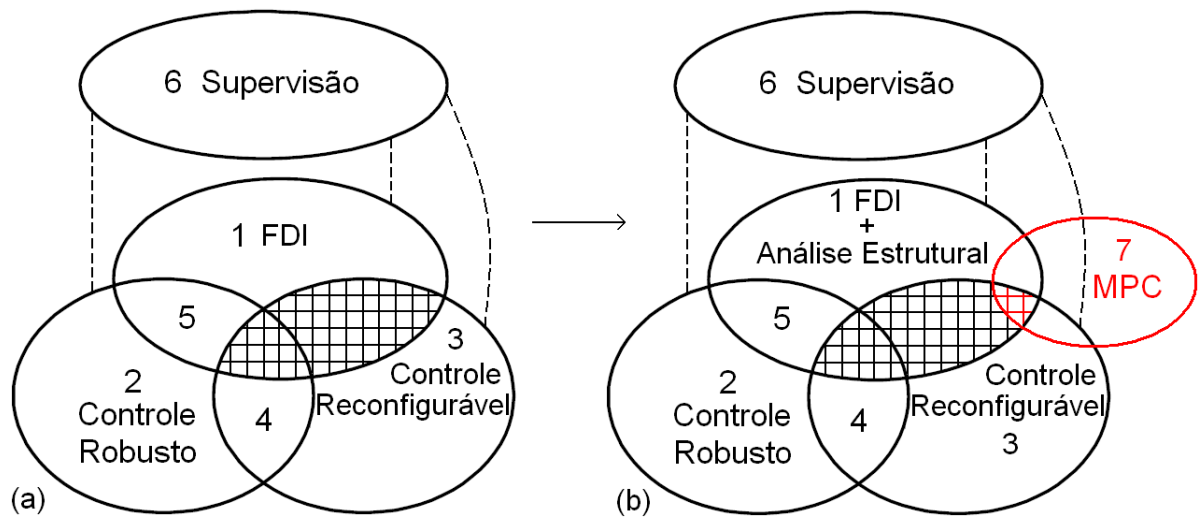


Figura 1.1: Situação do controle tolerante a falhas em 1997 (PATTON, 1997) (a) e proposta de contribuição deste trabalho (b).

A prioridade precedente ao controle tolerante a falhas é o projeto de um gerenciador de eventos anormais, também conhecido como sistema supervisor, cuja função é detectar, identificar (isolar) e diagnosticar falhas e enviar informações para algum algoritmo de reconfiguração. A Figura (1.2) mostra de forma sucinta a proposta de implementação de estratégias de controle tolerante a falhas. Assim, medidas de entrada e saída alimentam um bloco denominado FDD (*Fault Detection and Diagnosis*, Detecção de Falhas e Diagnóstico) capaz de decidir sobre o estado do processo e enviar informações para um sistema de supervisão, que é capaz de tomar ações corretivas sobre as falhas que atuam no processo, nos medidores, nos atuadores ou até mesmo no próprio controlador. Estas estratégias são usualmente referidas como estratégias de reconfiguração.

Apesar do esquema de reconfiguração dos parâmetros do controlador ser uma alternativa eficiente para tratar alguns tipos de falhas, tais como perturbações e mudanças não permitidas de algum parâmetro, deve-se dar atenção especial para a etapa de recuperação do processo. Esta etapa é mais importante, já que envolve mudanças físicas da operação, tal como a retirada de linha de uma bomba danificada e colocação em funcionamento de uma que esteja conectada em paralelo ao processo para que o mesmo objetivo de controle seja alcançado.

O objetivo principal desta dissertação é o de apresentar algumas alternativas para obtenção de controle tolerante a falhas unindo diferentes técnicas de detecção e diagnósticos de falhas com as estratégias de controle. A dissertação segue a seguinte estrutura: no Capítulo 2 apresentam-se técnicas de monitoramento, tais como detecção e diagnósticos



de falhas utilizando dados históricos e modelos. O Capítulo 3 apresenta diferentes formas de se obter controle tolerante a falhas utilizando controladores do tipo PID, ótimo e preditivo linear baseado em modelo, além de abordagens específicas para reconfiguração de controladores ou plantas. Resultados dessas explorações são mostrados no Capítulo 4 utilizando quatro exemplos: controle ótimo de uma coluna de destilação para separação de metanol da água e um sistema discreto invariante no tempo controlado por MPC e sujeito a falhas; uma coluna de destilação binária na qual são exploradas a troca entre diferentes estruturas de controle para acomodar falhas em atuadores; um sistema reacional que utiliza funções de monitoramento multivariáveis para detectar e diagnosticar falhas e análise estrutural para possibilitar a reconstrução de medidas incorretas; um reator de neutralização controlado por MPC sucessivamente linearizado em que falhas em atuadores são investigadas. Conclusões e sugestões para novos trabalhos são apresentadas no Capítulo 5.

CAPÍTULO 2

Monitoramento de Processos

Este capítulo apresenta as principais técnicas de detecção e diagnóstico de falhas em processos que possuem dados históricos ou modelos disponíveis. O intuito é demonstrar como diferentes técnicas podem ser utilizadas para detectar e diagnosticar falhas em processos. Para o monitoramento utilizando dados históricos, duas técnicas são apresentadas: a Análise dos Componentes Principais e a Análise Discriminante de Fisher, sendo que ao final de cada seção apresenta-se um resumo de cada técnica e suas devidas aplicações, seja para detecção ou diagnóstico, referindo-se diretamente às métricas que devem ser aplicadas para cada situação específica. Esta estrutura baseia-se naquela proposta por Chiang et al. (2001). A utilização de modelos para monitoramento tem como base a análise estrutural, que é capaz de escrever o modelo em uma maneira de fácil compreensão e explora com facilidade a realização de análises de observabilidade e monitorabilidade utilizando apenas estruturas de grafos e matriz de incidência (BLANKE et al., 2006).

2.1 Monitoramento e Controle Tolerante

A proposta deste trabalho se baseia na análise do desenvolvimento dos temas de monitoramento e controle tolerante a falhas pela literatura disponível. A Figura (2.1) mostra as principais contribuições realizadas ao longo de pouco mais de duas décadas.

O contribuição de Noura et al. (1993) reside na união de estratégias de monitoramento e controle reconfigurável para obtenção de um sistema de controle tolerante a

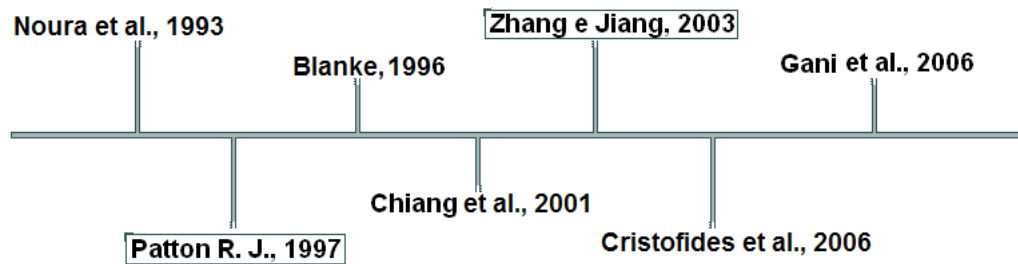


Figura 2.1: Revisão sobre monitoramento e controle tolerante a falhas.

falhas. Patton (1997) mostra o estado da pesquisa em controle tolerante a falhas até o ano de 1997, destacando as contribuições mais importantes realizadas até então. Além disso, indica direções futuras para a pesquisa nesta área. Um ano antes, Blanke et al. (1996) apresentam contribuições relevantes na área de controle tolerante utilizando análise estrutural. Chiang et al. (2001) apresentam várias ferramentas para detecção, identificação e diagnóstico de falhas, que é base para o projeto de sistemas de controle tolerante. Zhang e Jiang (2003) apresentam uma revisão bastante completa da literatura de controle tolerante e constataam que o número de contribuições nesta área é crescente. Patwardhan et al. (2006) exploram o controle preditivo tolerante a falhas com sucesso e Gani et al. (2006) desenvolvem um sistema de controle não linear tolerante a falhas para um reator de produção de polietileno.

2.2 Análise dos Componentes Principais

Esta técnica baseia-se em dados históricos de processos para construir modelos estatísticos que são capazes de discernir entre diferentes tipos de operações previamente categorizadas. A utilização desta em etapas precedentes à elaboração de um esquema de controle tolerante é uma alternativa, visto que é bastante versátil na monitoração das variações sistemáticas e de ruídos de forma independente ou conjunta. Assim, torna-se possível detectar e diagnosticar falhas de forma eficiente. Para unidades industriais de grande porte que possuem instrumentação para as variáveis mais importantes para propósitos de controle, é possível capturar dados fora de linha (*off-line*) e gerar modelos estatísticos que são as bases dos algoritmos de detecção e diagnóstico desenvolvidos nesta seção.

2.2.1 Introdução

A técnica estatística conhecida como PCA (PCA, *Principal Component Analysis*, Análise dos Componentes Principais) é capaz de realizar um monitoramento estatístico de processos multivariáveis através da projeção dos dados num espaço de dimensão reduzida, ou

seja, reduz-se a dimensão de um conjunto de dados preservando a estrutura de correlação entre eles para obter representações que realcem suas variabilidades.

A PCA, como uma ferramenta de monitoramento para processos industriais e redução de dimensionalidade de um conjunto de dados tem sido alvo de estudos por pesquisadores do meio acadêmico e industrial devido ao fato de sua fácil implementação e utilização de dados obtidos durante diferentes tipos de operações, incluindo as com falha. Um exemplo da aplicação da PCA para uma unidade industrial é da DuPont, com os resultados publicados em congressos e revistas específicas da área (KOSANOVICH et al., 1994, PIOVOSO; KOSANOVICH, 1992, PIOVOSO; KOSANOVICH, 1994, WISE; GALLAGHER, 1996). Quando uma unidade industrial real não está disponível, é comum utilizar simulações computacionais de unidades inteiras. Em ambos os casos, unidades reais e simuladas, tem-se processos multivariáveis, e a PCA torna-se útil pela redução da dimensão e explicação da variabilidade de todo o processo em apenas uma variável.

De fato, a PCA necessita de um conjunto de dados de treinamento, dados obtidos por operações passadas de algum processo. Estes dados passam então por um tratamento estatístico para produzir novas representações num espaço dimensional menor que se generalizam melhor para dados independentes do conjunto de treinamento (dados obtidos em linha). A utilização de uma dimensão reduzida melhora a eficiência de detecção e diagnóstico de falhas. Uma outra característica importante da PCA é a sua utilização para identificar variáveis responsáveis ou afetadas por algum comportamento anormal (WISE; RICKER; VELTKAMP, 1989). O monitoramento por PCA pode se tornar mais completo quando utiliza-se, além do monitoramento das variações de estado, a do ruído aleatório, que é a parte descartada pelo processo de redução de dimensionalidade da PCA. Tal fato justifica-se devido ao conhecimento de que alguns tipos de falhas afetam primeiro um dos dois subespaços (estados ou ruídos), logo, desenvolver medidas específicas para cada um deles pode significar em um aumento da sensibilidade do monitoramento.

As características da PCA ressaltadas anteriormente são discutidas ao longo desta seção. Apresenta-se também uma incorporação de correlação temporal para a PCA, resultando na DPCA (DPCA, *Dynamic PCA*, PCA Dinâmica). Destaca-se ainda a inserção de algoritmos para implementação de técnicas de monitoramento por PCA e de redução de dimensionalidade por análise paralela.

2.2.2 Descrição da Técnica PCA

Como enfatizado anteriormente, a PCA é uma técnica de redução de dimensão que captura a variabilidade dos dados de forma ótima, no entanto, é adequada apenas para dados gaussianos. Para dados não gaussianos é conveniente utilizar a ICA (*Independent Com-*

ponent Analysis, Análise dos Componentes Independentes) (REIS; OLIVEIRA-LOPES, 2006). A PCA advém da resolução de um problema de valor característico no qual um conjunto de vetores característicos ortogonais são associados com os maiores valores característicos, ordenados assim pela quantidade de variância explicada pelas direções destes vetores. Durante a aplicação da PCA e outras técnicas de monitoramento baseadas em dados históricos de algum processo, ocorrerá com frequência os termos “conjunto de treinamento” e “conjunto de teste”, o primeiro refere-se ao conjunto de dados do processo utilizados para treinar (obter modelos utilizando apenas dados de treinamento) a técnica em questão enquanto o último conjunto são dados independentes que são utilizados para validação das técnicas.

Considerando um conjunto de treinamento com n observações e m variáveis de processo dispostas em uma matriz $\mathbf{X}$, por exemplo,

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}. \quad (2.1)$$

Os vetores característicos são calculados através da solução dos pontos estacionários do problema de otimização

$$\max_{\mathbf{v} \neq 0} \frac{\mathbf{v}^T \mathbf{X}^T \mathbf{X} \mathbf{v}}{\mathbf{v}^T \mathbf{v}}, \quad (2.2)$$

em que $\mathbf{v} \in \mathbb{R}^m$. Os pontos estacionários da Equação (2.2) podem ser calculados através de uma SVD (SVD, *Singular Value Decomposition*, Decomposição dos Valores Singulares) de $\frac{1}{\sqrt{n-1}}\mathbf{X}$, resultando em:

$$\frac{1}{\sqrt{n-1}}\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \quad (2.3)$$

em que $\mathbf{U} \in \mathbb{R}^{n \times n}$ e $\mathbf{V} \in \mathbb{R}^{m \times m}$ são matrizes unitárias, e a matriz $\mathbf{\Sigma} \in \mathbb{R}^{n \times m}$ contém valores singulares reais não negativos com magnitude decrescente ao longo da diagonal principal ($\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(m,n)} \geq 0$), com os elementos não pertencentes à diagonal principal iguais a zero. Os vetores característicos são vetores coluna ortonormais na matriz $\mathbf{V}$, e a variância do conjunto de treinamento projetada ao longo da i -ésima coluna de $\mathbf{V}$ é igual a σ_i^2 (CHIANG et al., 2001). Resolver a Equação (2.3) é equivalente a resolver uma decomposição em valores característicos da matriz de covariância amostral $\mathbf{S}$,

$$\mathbf{S} = \frac{1}{n-1}\mathbf{X}^T \mathbf{X} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T, \quad (2.4)$$

em que a matriz diagonal $\mathbf{\Lambda} = \mathbf{\Sigma}^T \mathbf{\Sigma} \in \mathbb{R}^{m \times m}$ contém os valores característicos reais não negativos de magnitude decrescente ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$) e o i -ésimo valor

característico (λ_i) é igual ao quadrado do i -ésimo valor singular (σ_i^2).

O objetivo principal da PCA é capturar de forma ótima as variações dos dados e ao mesmo tempo minimizar o efeito da presença de possíveis ruídos aleatórios, visto que estes prejudicam a representação por PCA. Para que o objetivo exposto seja alcançado, é muito comum reter os ' a ' maiores valores característicos. Esta redução de dimensão é motivada para que se projete sistemas de detecção de falhas que separem as partes do processo que contém ruído aleatório, monitorando assim, de forma independente, esses dois tipos de situações.

A partir da matriz $\mathbf{V}$, seleciona-se as colunas que correspondem aos a maiores valores característicos e obtém-se a matriz $\mathbf{P} \in \mathbb{R}^{m \times a}$. A matriz $\mathbf{T}$ contém a projeção das observações em $\mathbf{X}$ num espaço de menor dimensão

$$\mathbf{T} = \mathbf{XP}, \quad (2.5)$$

e a projeção de $\mathbf{T}$, novamente no espaço de observação m -dimensional, é

$$\hat{\mathbf{X}} = \mathbf{TP}^T. \quad (2.6)$$

A diferença entre $\mathbf{X}$ e $\hat{\mathbf{X}}$ gera a matriz residual $\mathbf{E}$:

$$\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}}, \quad (2.7)$$

que captura as variações que ocorrem dentro do espaço de observações expandidas pelos vetores característicos associados com os $m - a$ menores valores característicos. Os subespaços expandidos por $\hat{\mathbf{X}}$ e $\mathbf{E}$ são comumente chamados de espaço de pontos e espaço residual, respectivamente. O subespaço contido na matriz $\mathbf{E}$ possui uma relação sinal-para-ruído bastante pequena, ou seja, possui maior quantidade de ruído devido ao fato do subespaço $\hat{\mathbf{X}}$ conter pouco ruído.

Definindo $\mathbf{t}_i$ como a i -ésima coluna de $\mathbf{T}$ no conjunto de treinamento, as seguintes propriedades podem ser demonstradas (PIOVOSO; KOSANOVICH, 1994).

1. $\text{Variância}(\mathbf{t}_1) \geq \text{Variância}(\mathbf{t}_2) \geq \dots \geq \text{Variância}(\mathbf{t}_a)$.
2. Média de $\mathbf{t}_i = 0$; $\forall i$.
3. $\mathbf{t}_i^T \mathbf{t}_k = 0$; $\forall i \neq k$.
4. Não existe outra expansão ortogonal de a componentes que capture mais variações dos dados.

Quando um novo vetor coluna de observações no conjunto de teste, $\mathbf{x} \in \mathbb{R}^m$, é projetado dentro do espaço de pontos de menor dimensão, $\mathbf{t}_i = \mathbf{x}^T \mathbf{p}_i$, em que $\mathbf{p}_i$ é o i -ésimo vetor característico, obtém-se o i -ésimo componente principal de $\mathbf{x}$, $\mathbf{t}_i$.

A descrição anterior é valiosa devido ao fato da análise se basear em apenas um número reduzido de componentes do modelo, chamados de neste texto por componentes principais. Tal fato possibilita utilizar técnicas estatísticas multivariáveis para monitorar o processo como um todo.

2.2.3 Análise de Redução de Ordem

Está teoricamente comprovado (WISE; RICKER; VELTKAMP; KOWALSKI, 1990) que a porção PCA correspondente aos maiores valores característicos descreve a maior parte das variações sistemáticas ou de estado do processo e que o espaço relacionado aos menores valores característicos descreve o ruído aleatório. As técnicas de redução de ordem visam desacoplar as variações de estados das variações aleatórias, determinando um número apropriado de valores característicos a que devem ser mantidos no modelo PCA. Assim, monitora-se de forma independente os dois tipos de variações. Existem várias técnicas para determinar o valor da redução de ordem a (HIMES; STORER; GEORGAKIS, 1994, JACKSON, 1991, QIN; DUNIA, 2000, VALLE-CERVANTES; LI; QIN, 1999), mas nenhuma delas é dominante. Alguns dos métodos para se determinar a são:

1. o teste da variância percentual,
2. o teste *scree*,
3. a análise paralela e
4. a estatística PRESS (*Prediction Residual Sum of Squares*, Soma dos Quadrados da Predição Residual).

Utiliza-se, neste trabalho, o teste da variância percentual e a análise paralela como técnicas de redução de ordem. A análise paralela é discutida aqui devido a sua maior utilização e testes realizados por Ku et al. (1995), que a recomendam devido ao desempenho superior quando comparada com outras técnicas. Já o teste da variância percentual determina a dimensão a através do cálculo do menor número de valores característicos que explicam uma percentagem mínima específica da variância total. Como esta percentagem é escolhida arbitrariamente, ela pode ser inferior, superior ou até mesmo igual a redução obtida pela análise paralela (Chiang et al., 2001).

A análise paralela determina a redução de dimensão através da comparação do perfil de variância dos dados reais com dados obtidos pela hipótese de variáveis independentes. A redução de ordem é determinada no ponto em que os dois perfis se cruzam. Esta aproximação garante que correlações significantes são capturadas no espaço de pontos, sendo bastante utilizada por ser intuitiva e de fácil implementação. O algoritmo para a análise paralela é apresentado a seguir (REIS; OLIVEIRA-LOPES, 2006):

Algoritmo 2.1: Análise Paralela

1. Gerar um conjunto de dados normalmente distribuídos com média zero e variância unitária com a mesma dimensão do conjunto de treinamento (m variáveis e n observações).
2. Realizar uma PCA sobre estes dados.
3. Obter os valores característicos classificados em ordem decrescente.
4. Plotar os valores característicos dos dados originais juntamente com o dos dados normalmente distribuídos.
5. Obter a redução de ordem através do cruzamento entre os perfis.

2.2.4 Estatísticas Multivariáveis para Detecção de Falhas

A detecção de falhas em processos multivariáveis que dispõe de dados históricos é feita neste trabalho através da estatística T^2 . Supondo que existem dados para treinamento dispostos em uma matriz, assim como em (2.1), ou seja, m variáveis observadas e n observações para cada variável.

É comum a normalização dos dados antes da aplicação da técnica, assim, para os dados do conjunto de treinamento, subtrai-se de todas n observações, a média e divide-se pelo desvio padrão de cada m variável. A partir dos dados normalizados calcula-se a matriz de covariância amostral do conjunto de treinamento por

$$\mathbf{S} = \frac{1}{n-1} \mathbf{X}^T \mathbf{X}. \quad (2.8)$$

Uma decomposição em valores característicos da matriz $\mathbf{S}$,

$$\mathbf{S} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T, \quad (2.9)$$

revela a estrutura de correlação para a matriz de covariância, em que $\mathbf{\Lambda}$ é diagonal e $\mathbf{V}$ é ortogonal ($\mathbf{V}^T \mathbf{V} = \mathbf{I}$, em que $\mathbf{I}$ é a matriz identidade) (GOLUB; VAN LOAN, 1996). A projeção $\mathbf{y} = \mathbf{V}^T \mathbf{x}$ de um vetor de observação $\mathbf{x} \in \mathbb{R}^m$ desacopla o espaço de observação em um conjunto de variáveis não correlacionadas correspondentes aos elementos de $\mathbf{y}$. A variância do i -ésimo elemento de $\mathbf{y}$ é igual ao i -ésimo valor característico na matriz $\mathbf{\Lambda}$. Supondo que $\mathbf{S}$ admite inversa e com a definição de

$$\mathbf{z} = \mathbf{\Lambda}^{-1/2} \mathbf{V}^T \mathbf{x}, \quad (2.10)$$

a estatística T^2 de Hotelling é dada por (JACKSON, 1959)

$$T^2 = \mathbf{z}^T \mathbf{z}. \quad (2.11)$$

A utilização da estatística T^2 como uma ferramenta no processo de detecção de falhas em processos multivariáveis se baseia na existência de um conjunto de dados de treinamento como referência. Aplica-se PCA para revelar a estrutura de correlação dos dados e quando novas informações, independentes do conjunto de treinamento, $\mathbf{x}$, são inseridas e supondo que $\mathbf{\Lambda} = \mathbf{\Sigma}^T \mathbf{\Sigma}$ admite inversa, a estatística T^2 é calculada diretamente pela representação PCA (Chiang et al., 2001)

$$T^2 = \mathbf{x}^T \mathbf{V} (\mathbf{\Sigma}^T \mathbf{\Sigma})^{-1} \mathbf{V}^T \mathbf{x}. \quad (2.12)$$

O monitoramento através da estatística T^2 deve ser analisado do ponto de vista da quantidade de dados disponíveis, que quando é relativamente pequena tende a ser uma representação imprecisa do comportamento do processo monitorado, fato este ocasionado por imprecisões nos menores valores característicos, que por sua vez, corrompem a estatística T^2 devido a pequena razão sinal-para-ruído. Assim, apenas os vetores característicos associados com os maiores valores característicos devem ser retidos no cálculo da estatística T^2 .

A utilização de alguma técnica de redução de ordem apresentada na seção 2.2.3 seleciona os maiores valores característicos e seus respectivos vetores, matriz $\mathbf{P}$. A estatística T^2 para o monitoramento das variações de estado (maiores valores característicos) é

$$T^2 = \mathbf{x}^T \mathbf{P} \mathbf{\Sigma}_a^{-2} \mathbf{P}^T \mathbf{x}, \quad (2.13)$$

em que $\mathbf{\Sigma}_a$ contém as primeiras a linhas e colunas de $\mathbf{\Sigma}$, ou seja, os valores característicos dispostos na diagonal principal. Se a média amostral e a matriz de covariância reais são

conhecidas, o limite para a estatística T^2 é dado por

$$T_\alpha^2 = \chi_\alpha^2(a). \quad (2.14)$$

Quando a matriz de covariância real para processos controlados não é conhecida, mas estimada através da matriz de covariância amostral, Equação (2.9), o limite para a estatística T^2 é

$$T_\alpha^2 = \frac{a(n-1)(n+1)}{n(n-a)} F_\alpha(a, n-a), \quad (2.15)$$

em que $F_\alpha(a, n-a)$ é a distribuição de Fisher com a e $n-a$ graus de liberdade.

Para detectar *outliers* no conjunto de treinamento, o seguinte limite pode ser usado

$$T_\alpha^2 = \frac{(n-1)^2 \left(\frac{a}{n-a-1} \right) F_\alpha(a, n-a-1)}{n \left(1 + \left(\frac{a}{n-a-1} \right) \right) F_\alpha(a, n-a-1)}. \quad (2.16)$$

A violação de algum limite pode indicar que as variações sistemáticas estão fora de controle.

A estatística T^2 na Equação (2.13) não é afetada por imprecisões nos menores valores característicos, visto que os mesmos não fazem parte da análise. Por este aspecto, pode-se concluir que ela é capaz de representar o comportamento normal do processo e é uma medida menos afetada por ruídos, que não é o caso da Equação (2.12), que por sua vez considera todos valores característicos. Para contornar este tipo de problema, a porção do espaço de observação correspondendo aos $m-a$ menores valores característicos pode ser monitorada mais precisamente através do uso da estatística Q

$$\begin{aligned} Q &= \mathbf{r}^T \mathbf{r}, \\ \mathbf{r} &= (\mathbf{I} - \mathbf{P}\mathbf{P}^T) \mathbf{x}, \end{aligned} \quad (2.17)$$

em que $\mathbf{r}$ é o vetor de resíduos, uma projeção do vetor de observação $\mathbf{x}$ dentro do espaço residual. Como a estatística Q não mede diretamente as variações ao longo de cada vetor característico, mas a soma total das variações no espaço residual, não sofre de sensibilidade devido a imprecisões nos menores valores característicos. A estatística Q é conhecida também como SPE (SPE, *Squared Prediction Error*, Erro de Predição ao Quadrado), e é uma norma quadrada que mede o desvio das observações do espaço de menor dimensão da representação PCA.

A distribuição para a estatística Q foi proposta por Jackson e Mudholkar (1979).

$$Q_\alpha = \theta_1 \left[\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0}, \quad (2.18)$$

em que $\theta_i = \sum_{j=a+1}^n \sigma_j^{2i}$, $h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}$, e c_α é o desvio normal correspondendo ao percentil $(1 - \alpha)$. Dado um nível de significância, α , o limite para a estatística Q pode ser calculado pela Equação (2.18) para detectar falhas.

A estatística Q mede as variações aleatórias do processo, por exemplo, aquelas associadas com medições de ruído. O limite dado pela Equação (2.18) pode ser aplicado para definir as variações normais para o ruído aleatório, e uma violação do limite poderia indicar que o ruído aleatório mudou significativamente. As estatísticas T^2 e Q , juntamente com seus limites, detectam diferentes tipos de falhas, e as vantagens de ambas estatísticas podem ser utilizadas através do emprego das duas medidas em um único esquema de monitoramento (REIS; OLIVEIRA-LOPES, 2007).

2.2.5 Diagnóstico de Falhas para Estatísticas Multivariáveis

Define-se como diagnóstico de falhas o procedimento de reconhecimento de falhas baseado em análise discriminatória. O diagnóstico de falhas utilizando dados históricos de um processo exige um conjunto de treinamento para cada falha, ou seja, as funções discriminatórias classificam as falhas baseadas num treinamento prévio. Diferentemente da identificação de falhas, que monitora um determinado conjunto de variáveis e através de um modelo PCA é capaz de contabilizar a contribuição de cada variável observada quando uma falha está presente no sistema, o diagnóstico será capaz de condensar as medições do processo e comparar com a base de dados disponível (um modelo PCA para cada falha, por exemplo) e dessa forma classificar as falhas.

Na prática, a grande vantagem da implementação de sistemas de diagnósticos é que se pode ter uma resposta mais específica sobre a falha do que um sistema de identificação, que apenas identifica variáveis que estão mais associadas com uma determinada falha, localizando assim subsistemas no qual a falha está presente. A união das técnicas de identificação e diagnóstico é o ideal para uma estratégia de monitoramento de processo em linha completa, pois estas informações podem ser direcionadas para estruturas de controle tolerante, capazes de analisar e reconfigurar o processo devido a ocorrência de uma falha.

Existem trabalhos que utilizam como base para o diagnóstico de falhas a análise dos componentes principais, construindo apenas um modelo PCA e definindo regiões do espaço

de dimensão reduzida que classificam uma determinada falha. Apesar de implementação simples, não é efetiva quando o número de falhas é grande, quando isto ocorre, pode-se desenvolver um modelo PCA para cada conjunto de dados de falhas coletadas durante operações passadas (dados fora de linha) e aplicar as estatísticas T^2 (KU; STORER; GEORGAKIS, 1993) e Q (RAICH; ÇINAR, 1996) para cada modelo para predizer qual falha ocorre. Esta técnica é uma combinação de análise dos componentes principais e análise discriminatória.

Para utilizar a PCA como ferramenta para o diagnóstico de falhas deve-se dispor em uma única matriz $\mathbf{X}$, dados de todas as classes de falhas e aplicar a PCA. A matriz de vetores característicos $\mathbf{P}$ é calculada pela Equação (2.3) ou Equação (2.4). A máxima probabilidade de classificação para uma observação $\mathbf{x}$ independente do conjunto de treinamento é a classe de falha i quando o valor máximo da função discriminante de pontos

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}}_i)^T \mathbf{P}(\mathbf{P}^T \mathbf{S}_i \mathbf{P})^{-1} \mathbf{P}^T (\mathbf{x} - \bar{\mathbf{x}}_i) + \ln(p_i) - \frac{1}{2} \ln[\det(\mathbf{P}^T \mathbf{S}_i \mathbf{P})] \quad (2.19)$$

é obtido. Em que $\bar{\mathbf{x}}_i$ é o vetor médio para classe i ,

$$\bar{\mathbf{x}}_i = \frac{1}{n_i} \sum_{\mathbf{x}_j \in \chi_i} \bar{\mathbf{x}}_j, \quad (2.20)$$

n_i é o número de observações na classe de falha i , χ_i é o conjunto de vetores $\mathbf{x}_j$ que pertencem à classe de falha i , e $\mathbf{S}_i \in \mathbb{R}^{m \times m}$ é a matriz de covariância amostral para a classe de falha i , definida na Equação (2.8).

Se a matriz $\mathbf{P}$ da Equação (2.19) incluir todas as dimensões dos dados, tem-se $\mathbf{P} = \mathbf{V} \in \mathbb{R}^{m \times m}$, e a probabilidade global para toda classe de falha é a mesma, a Equação (2.19) se reduz a função discriminante para a estatística multivariável. A estatística multivariável seleciona a classe de falha mais provável baseando-se na função discriminante

$$g_i(\mathbf{x}) = -(\mathbf{x} - \bar{\mathbf{x}}_i)^T \mathbf{S}_i^{-1} (\mathbf{x} - \bar{\mathbf{x}}_i) - \ln[\det(\mathbf{S}_i)]. \quad (2.21)$$

Esta estatística serve também como um ponto de referência para as outras estatísticas, pois a dimensionalidade somente deve ser reduzida (transformação de $\mathbf{V}$ em $\mathbf{P}$) se ocorre um decréscimo da taxa de classificações incorretas para o conjunto de treinamento.

A utilização de modelos baseados em análises PCA descritos neste trabalho confiam na separação entre dos dados para projetar funções discriminatórias para diagnosticar falhas (RAICH; ÇINAR, 1996). Assim, um vetor de observação $\mathbf{x}$ independente do conjunto de treinamento é classificado como pertencente à classe de falha i com o máximo

discriminante de pontos

$$g_i(\mathbf{x}) = -\frac{1}{2}\mathbf{x}^T \mathbf{P}_i \Sigma_{a,i}^{-2} \mathbf{P}_i^T \mathbf{x} - \frac{1}{2} \ln[\det(\Sigma_{a,i}^2)] + \ln(p_i), \quad (2.22)$$

em que $\mathbf{P}_i$ é a matriz de vetores característicos para a classe de falha i , $\Sigma_{a,i}$ é a matriz diagonal Σ_a como mostrado pela Equação (2.13) para a classe de falha i ($\Sigma_{a,i}^2$ é a matriz de covariância de $\mathbf{P}_i \mathbf{x}$), e p_i é a probabilidade global da classe de falha i . Supõe que o vetor de observação $\mathbf{x}$ na Equação (2.22) é normalizado de acordo com a média e desvio padrão do conjunto de treinamento para a classe de falha i . A Equação (2.22) baseia-se na função discriminatória (CHIANG et al., 2001)

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \mu_i)^T \Sigma_i^{-1}(\mathbf{x} - \mu_i) - \frac{m}{2} \ln 2\pi - \frac{1}{2} \ln[\det(\Sigma_i)] + \ln P(\omega_i). \quad (2.23)$$

As matrizes $\mathbf{P}_i$, $\Sigma_{a,i}$ e o escalar p_i na Equação (2.22) dependem unicamente da classe de falha i . O ponto fraco desse tipo de aproximação é que como a análise é feita para cada classe separadamente, informações úteis de outras classes não são levadas em consideração e além disso cada modelo para cada classe possui uma redução de ordem diferente. Dessa forma, como os dados independentes são projetados sobre diferentes reduções de ordem ($\mathbf{P}_i^T \mathbf{x}$), é possível a ocorrência de uma elevada taxa de classificações incorretas.

Diferentemente da função discriminante, Equação (2.22), a matriz de vetores característicos $\mathbf{P}$ na Equação (2.19) utiliza informações de todas classes de falhas bem como projeta os dados sobre as mesmas dimensões para cada classe. Devido a estas propriedades, a função discriminatória (Equação (2.19)), pode ter um desempenho melhor que a Equação (2.22) para diagnosticar falhas. Seguindo a mesma proposta apresentada por Chiang et al. (2001), para distinguir a PCA baseada em um único modelo da PCA com múltiplos modelos, refere-se àquela como PCA1 e à esta como PCAm.

Um caso particular da Equação (2.22) ocorre se a suposição de probabilidade global para toda classe de falha for a mesma bem como a matriz de covariância amostral de $\mathbf{P}_i \mathbf{x}$ para cada classe de falha. O uso da função discriminatória dada pela Equação (2.22), reduz-se então a usar a estatística T_i^2 . Neste caso, o discriminante de pontos selecionará as classes de falhas como aquelas que correspondem à estatística T_i^2 mínima,

$$T_i^2 = \mathbf{x}^T \mathbf{P}_i \Sigma_{a,i}^{-2} \mathbf{P}_i^T \mathbf{x}. \quad (2.24)$$

Da mesma forma que a função discriminatória dada pela Equação (2.22) leva em consideração apenas os maiores valores característicos de cada classe, monitorando apenas as variações de estados, é possível avaliar também as variações do espaço residual com o

discriminante mínimo

$$\frac{Q_i}{(Q_\alpha)_i}. \quad (2.25)$$

Assim, se as variações no espaço residual e no espaço de pontos são importantes, então uma observação é representada pelo mínimo discriminante combinado

$$c_i \left(\frac{T_i^2}{(T_\alpha^2)_i} \right) + (1 - c_i) \left(\frac{Q_i}{(Q_\alpha)_i} \right), \quad (2.26)$$

em que c_i é um fator de ponderação que varia entre 0 e 1 para cada classe de falha i . Essa ponderação c_i fornece maior ou menor importância para as estatísticas utilizadas na discriminação, T_i^2 e Q_i , para cada classe de falha i . Supondo que uma observação fora de controle não represente uma nova falha, cada uma destas técnicas de análises discriminatórias dadas pelas Equações (2.22), (2.25) e (2.26) podem ser utilizadas para diagnosticar falhas.

Quando uma falha é diagnosticada como falha i , não representa provavelmente uma nova falha quando

$$\frac{T_i^2}{(T_\alpha^2)_i} \ll 1 \quad (2.27)$$

e

$$\frac{Q_i^2}{(Q_\alpha)_i} \ll 1. \quad (2.28)$$

Estas condições indicam que a observação é uma boa combinação para a falha i . Se estas condições não são satisfeitas (por exemplo, $T_i^2/(T_\alpha^2)_i$ ou $Q_i^2/(Q_\alpha)_i$ são maiores do que 1), então a observação não é precisamente representada pela classe de falha i e é provável que as observações representem uma nova falha.

2.2.6 PCA Dinâmica

Os métodos de monitoramento baseados em PCA discutidos nas seções anteriores admitem que as observações em um determinado instante de tempo são estatisticamente independentes das observações passadas. Entretanto, para processos industriais, essa hipótese não é válida, visto que os intervalos de amostragem são bastante pequenos, medidos em segundos para a maioria das variáveis. A independência estatística é obtida apenas para intervalos de amostragem que variam entre 2 e 12 horas (CHIANG et al., 2001). Para contabilizar o efeito desta dependência para processos com curtos intervalos de amostragem desenvolveu-se um método de monitoramento que leva em consideração as correlações temporais. Existem técnicas específicas para determinar se existe correlação entre os dados do conjunto treinamento, tal como a carta de auto-correlação dos componentes

principais.

O modelo PCA levará em consideração as correlações temporais quando cada vetor de observação for ampliado com as h observações anteriores, dispondo numa matriz

$$\mathbf{X}(k-h,:) = [\mathbf{x}_k^T \ \mathbf{x}_{k-1}^T \ \mathbf{x}_{k-2}^T \ \cdots \ \mathbf{x}_{k-h}^T], \quad k = 1, 2, \dots, n, \quad (2.29)$$

em que $\mathbf{x}_k^T$ é o vetor de observação de dimensão m e pertence ao conjunto de treinamento no instante de amostragem k .

O método de aplicar PCA na Equação (2.29) para diagnóstico de falhas será chamado de DPCAm. O método para determinar automaticamente h descrito em Ku et al. (1995), não é usado aqui. Experiências indicam que $h = 1$ ou 2 é usualmente apropriado quando DPCA é usada para monitoramento de processos (KU et al., 1995). As medidas de detecção e diagnóstico de falhas para a PCA generalizam diretamente para a DPCA. Para identificação de falhas, as medidas para cada variável observada podem ser calculadas através do somatório dos valores das medidas correspondendo aos h atrasos anteriores.

Observa-se, na prática, que a presença de correlações seriais nos dados não compromete a efetividade do PCA quando existe dados suficientes para representar todas as variações normais do processo (CHIANG et al., 2001). Independente deste fato, incluir atrasos na matriz de dados como feito na Equação (2.29) pode resultar na representação PCA que correlaciona mais informações.

2.2.7 Resumo das Técnicas PCA e DPCA

A técnica PCA apresentada nesta seção pode ser aplicada para detecção e diagnóstico de falhas. Assim, mostra-se em uma tabela todas informações sobre como utilizar PCA para tais aplicações. Com uma quantidade suficiente de dados em um conjunto de treinamento e dados relativos a falhas em processos (fora de linha), implementa-se estratégias de monitoramento, possibilitando a detecção e diagnóstico em tempo adequado para acionar as estratégias de controle tolerante.

2.3 Análise Discriminante de Fisher

A análise discriminante de Fisher pode ser entendida como uma técnica desenvolvida especificamente para o diagnóstico de falhas. O objetivo desta seção é o de apresentar esta técnica de redução de dimensionalidade que é ótima em termos de classificação de informações entre diferentes classes de dados (falhas, operação normal *etc.*) (DUDA;

Tabela 2.1: Resumo das técnicas PCA e DPCA aplicadas para detecção e diagnóstico de falhas.

Detecção de Falhas			
Método	Base	Equação	Distribuição
PCA	T^2	(2.13)	(2.15)
PCA	Q	(2.17)	(2.18)
DPCA	T^2	(2.13) [✕]	(2.15) [✕]
DPCA	Q	(2.17) [✕]	(2.18) [✕]
Diagnóstico de Falhas			
Método	Base	Equação	
PCAm	T^2	(2.22)	
PCA1	T^2	(2.19)	
PCAm	Q	(2.25)	
PCAm	T^2 e Q	(2.26)	
DPCAm	T^2	(2.22) com (2.29)	
DPCAm	Q	(2.25) com (2.29)	
DPCAm	T^2 e Q	(2.26)	

[✕] Aplicada à matriz de dados com atraso.

HART; STORK, 1973).

2.3.1 Introdução

A técnica conhecida como FDA (*Fisher Discriminant Analysis*, Análise Discriminante de Fisher) é mais utilizada para diagnóstico de falhas, visto que leva em consideração as informações entre as classes quando determina a representação dos dados no espaço de dimensão reduzida. Dessa forma, a FDA é uma técnica de redução de dimensionalidade, assim como a PCA, que apesar de conter propriedades ótimas em termos de detecção de falhas, não é muito adequada para o diagnóstico, pois não avalia as informações entre as diferentes classes, fato este considerado pela FDA.

2.3.2 Descrição da Técnica FDA

Quando dados coletados durante operações com falhas são organizados em classes, em que cada classe contém um determinado tipo de falha, pode-se utilizar estas informações para projetar sistemas de diagnóstico de falhas. Numericamente, a FDA determina um conjunto linear de vetores ordenados de acordo com a maximização da dispersão entre as classes, assim eles são responsáveis pela transformação dos dados, maximizando a dispersão entre as classes e minimizando a dispersão dentro de cada classe.

Supondo que existem n observações de todas as classes, m variáveis medidas, p classes e n_j representa o número de observações da j -ésima classe, o vetor de variáveis obtidas através de instrumentos medidores é $\mathbf{x}_i$ para a i -ésima observação. Se todos os dados relacionados com falhas forem dispostos em uma única matriz $\mathbf{X} \in \mathbb{R}^{n \times pm}$ como na Equação (2.1) e lembrando que $n = \sum_j n_j$, então a transposta da i -ésima linha de $\mathbf{X}$ é o vetor coluna $\mathbf{x}_i$.

Para melhorar a compreensão da técnica FDA é necessário definir algumas matrizes que mensuram a dispersão total, a dispersão dentro das classes e a dispersão entre as classes. O procedimento aqui adotado é mesmo proposto por Chiang et al. (2001). A matriz de dispersão total é definida por (DUDA; HART; STORK, 1973, HUDLET; JOHNSON, 1977):

$$\mathbf{S}_t = \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T, \quad (2.30)$$

em que $\bar{\mathbf{x}}$ é o vetor médio total

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i. \quad (2.31)$$

Se χ_j é definido como o conjunto de vetores $\mathbf{x}_i$ que pertencem à classe j , então a matriz de dispersão dentro da classe j é

$$\mathbf{S}_j = \sum_{\mathbf{x}_i \in \chi_j} (\mathbf{x}_i - \bar{\mathbf{x}}_j)(\mathbf{x}_i - \bar{\mathbf{x}}_j)^T, \quad (2.32)$$

em que $\bar{\mathbf{x}}_j$ é o vetor médio para a classe j :

$$\bar{\mathbf{x}}_j = \frac{1}{n_j} \sum_{\mathbf{x}_i \in \chi_j} \mathbf{x}_i. \quad (2.33)$$

A matriz de dispersão dentro de cada classe é

$$\mathbf{S}_w = \sum_{j=1}^p \mathbf{S}_j, \quad (2.34)$$

e a matriz de dispersão entre as classes é

$$\mathbf{S}_b = \sum_{j=1}^p n_j (\bar{\mathbf{x}}_j - \bar{\mathbf{x}})(\bar{\mathbf{x}}_j - \bar{\mathbf{x}})^T. \quad (2.35)$$

A matriz de dispersão total é igual a soma da matriz de dispersão dentro das classes

com a matriz de dispersão dentro da classe (DUDA; HART; STORK, 1973),

$$\mathbf{S}_t = \mathbf{S}_b + \mathbf{S}_w. \quad (2.36)$$

A FDA objetiva maximizar a dispersão entre as classes ($\mathbf{S}_b$) para que se obtenha uma melhor classificação entre diferentes classes de falhas e minimizar a dispersão dentro das classes ($\mathbf{S}_w$):

$$\max_{\mathbf{v} \neq 0} \frac{\mathbf{v}^T \mathbf{S}_b \mathbf{v}}{\mathbf{v}^T \mathbf{S}_w \mathbf{v}}, \quad (2.37)$$

supondo que $\mathbf{S}_w$ admite inversa e $\mathbf{v} \in \mathbb{R}^m$. Sabe-se que os vetores FDA são obtidos em ordem decrescente da maximização da dispersão entre as classes. O cálculo dos pontos estacionários do problema de otimização da Equação (2.37) é uma transformação linear para os vetores FDA, que são iguais aos vetores característicos $\mathbf{w}_k$ do problema de valor característico generalizado

$$\mathbf{S}_b \mathbf{w}_k = \lambda_k \mathbf{S}_w \mathbf{w}_k, \quad (2.38)$$

em que os valores característicos λ_k indicam o grau de separabilidade global entre as classes pela projeção dos dados em $\mathbf{w}_k$. A maioria dos *softwares* que possuem manipulações matriciais, tais como MATLAB (HANSELMAN; LITTLEFIELD, 1997) ou SCI-LAB (CAMPBELL; CHANCELIER; NIKOUKHAH, 2006), possuem rotinas para o cálculo dos vetores e valores característicos generalizados. Devido ao fato da direção e não a magnitude de $\mathbf{w}_k$ ser importante, a norma Euclidiana (raiz quadrada da soma dos quadrados de cada elemento) de $\mathbf{w}_k$ pode ser escolhida como sendo igual a 1 ($\|\mathbf{w}_k\| = 1$).

O cálculo dos vetores FDA é feito através da resolução do problema de valor característico generalizado quando $\mathbf{S}_w$ possui inversa. Na maioria dos casos isso será válido, pois o número de observações totais n é consideravelmente maior que o número de medições m . Para o caso de diagnóstico de falhas, espera-se que $\mathbf{S}_w$ admita inversa devido à constatação anterior, no entanto, existem métodos específicos para calcular os vetores FDA quando $\mathbf{S}_w$ não admite inversa e são referidas por (CHENG; ZHUANG; YANG, 1992, HONG; YANG, 1991, TIAN, 1988).

Os vetores FDA são classificados de acordo com a magnitude dos valores característicos calculados pelo problema de valor característico generalizado. Então, o primeiro vetor FDA está associado com o maior valor característico e isto é aplicado para os vetores restantes. Um valor característico λ_k com magnitude elevada indica que quando os dados nas classes são projetados sobre os vetores característicos associados com $\mathbf{w}_k$, existe uma grande separação global das médias das classes relacionadas às variâncias, e conseqüentemente, um elevado grau de separação entre as classes ao longo da direção $\mathbf{w}_k$.

Para escrever o objetivo da FDA (maximização da dispersão entre as classes e mini-

mização da dispersão dentro da classe) define-se a matriz $\mathbf{W}_p \in \mathbb{R}^{m \times (p-1)}$ com os $p - 1$ vetores FDA como colunas. Então, transformam-se os dados de um espaço de dimensão m para $p - 1$ da seguinte forma:

$$\mathbf{z}_i = \mathbf{W}_p^T \mathbf{x}_i, \quad (2.39)$$

em que $\mathbf{z}_i \in \mathbb{R}^{p-1}$. O resultado da FDA está contido na matriz $\mathbf{W}_p$, na qual os dados $\mathbf{x}_1, \dots, \mathbf{x}_n$ são projetados em um espaço de dimensão $p - 1$. Se $p = 2$, o resultado desta projeção ocorre em uma linha.

2.3.3 Detecção e Diagnóstico de Falhas

A FDA pode ser aplicada para detectar ou diagnosticar falhas. Quando aplicada para o diagnóstico, a técnica de redução de dimensão é realizada sobre os dados de todas as classes simultaneamente. A matriz $\mathbf{W}_a \in \mathbb{R}^{m \times a}$ contém os vetores característicos $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_a$ calculados pela Equação (2.38) (a representa a redução de ordem). A função discriminante pode ser obtida da Equação (2.23) como

$$\begin{aligned} g_j(\mathbf{x}) = & -\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}}_j)^T \mathbf{W}_a \left(\frac{1}{n_j - 1} \mathbf{W}_a^T \mathbf{S}_j \mathbf{W}_a \right)^{-1} \mathbf{W}_a^T (\mathbf{x} - \bar{\mathbf{x}}_j) + \\ & + \ln(p_i) - \frac{1}{2} \ln \left[\det \left(\frac{1}{n_j - 1} \mathbf{W}_a^T \mathbf{S}_j \mathbf{W}_a \right) \right], \end{aligned} \quad (2.40)$$

em que $\mathbf{S}_j$, $\bar{\mathbf{x}}_j$, e n_j estão definidos nas Equações (2.32) e (2.33). Diferentemente da PCA1, a FDA usa a informação da classe para calcular o espaço de dimensão reduzida, tal que a função discriminante, Equação (2.40), explora a informação da classe em um grau bem mais elevado do que na técnica PCA. Em contraste com a PCAm, a FDA utiliza todas as informações das p classes de falhas quando avalia a função discriminante para cada classe.

A FDA também pode ser utilizada como uma ferramenta para detecção de falhas, para isto basta definir uma classe de dados adicional que contenha informações do processo obtidas durante operação uma considerada normal. A utilização da função discriminante dada pela Equação (2.40) é vantajosa quando existe uma semelhança entre os dados de condição de operação normal e os dados das classes de falha para os conjuntos de treinamento. Assim, quando existe uma transformação $\mathbf{W}$ que separa a classe de operação normal das outras classes de forma eficiente, a utilização da FDA para detecção de falhas é favorecida e produzirá uma pequena taxa de alarmes falsos para as classes de falha conhecidas, ou seja, a FDA passa a considerar e classificar a operação normal como uma classe de falha, da mesma forma que quando apenas classes de falhas são utilizadas. A Equação (2.40) não leva em consideração falhas desconhecidas associadas com dados de fora do espaço de menor dimensão definido pelos vetores FDA. Assim como desenvolvido

para a PCA (estatísticas T^2 e Q), é mais adequado monitorar as variações de estado (maiores valores característicos) e de ruído aleatório (menores valores característicos) de forma independente. A vantagem de usar a estatística FDA em vez da PCA é que a informação da classificação da falha pode ser levada em consideração para melhorar a capacidade de detectar falhas. A desvantagem é que estatísticas FDA requerem informação da classificação da falha para definir seu espaço de dimensão reduzida (definido por $\mathbf{W}_a$), ou seja, a redução baseia-se na taxa de classificações incorretas.

Somente os primeiros $p-1$ vetores característicos FDA maximizam a dispersão entre as classes enquanto minimizam a dispersão dentro de cada classe. Os vetores característicos restantes ($m-p-1$) que correspondem aos valores característicos nulos não são ordenados pelo objetivo FDA, que é o de maximizar $\mathbf{S}_b$ e minimizar $\mathbf{S}_w$. Se mais que $p-1$ vetores são úteis para a classificação, dois métodos podem ser utilizados para a seleção de vetores adicionais e são propostos por Chiang et al. (2001):

- Utilizar a FDA para o espaço definido pelos primeiros $p-1$ vetores característicos e os vetores PCA1 para o restante dos vetores $m-p+1$, ordenados dos vetores PCA associados com a mais alta variabilidade para os vetores associados com a menor variabilidade. Se a redução de ordem é $a \leq p-1$, a Equação (2.40) é usada diretamente. Se $a \geq p$, a função discriminante alternativa é usada:

$$g_j(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}}_j)^T \mathbf{W}_{mix,a} \left(\frac{1}{n_j - 1} \mathbf{W}_{mix,a}^T \mathbf{S}_j \mathbf{W}_{mix,a} \right)^{-1} \mathbf{W}_{mix,a}^T (\mathbf{x} - \bar{\mathbf{x}}_j) + \\ - \frac{1}{2} \ln \left[\det \left(\frac{1}{n_j - 1} \mathbf{W}_{mix,a}^T \mathbf{S}_j \mathbf{W}_{mix,a} \right) \right] + \ln(p_i), \quad (2.41)$$

em que $\mathbf{W}_{mix,a} = [\mathbf{W}_{p-1} \ \mathbf{P}_{a-p+1}]$ são as primeiras $a-p+1$ colunas da matriz característica $\mathbf{P}$ do PCA1 (definida na Seção 2.2.2). Quando este método for usado para diagnosticar falhas, será chamado de método FDA/PCA1. Dado que os vetores do PCA1 podem ser úteis em um procedimento de classificação (ver Seção 2.2.4), a incorporação dos primeiros $a-p+1$ vetores característicos do PCA1 no método FDA/PCA1 pode fornecer informações adicionais para discriminar entre as classes.

- Um outro método para definir vetores adicionais $m-p+1$ é aplicar PCA1 para o espaço residual da FDA, definido por

$$\mathbf{R} = \mathbf{X}(\mathbf{I} - \mathbf{W}_{p-1} \mathbf{W}_{p-1}^T). \quad (2.42)$$

Como antes, se a redução de ordem $a \leq p-1$, a Equação (2.40) é usada diretamente. Se $a \geq p$, então a função discriminante alternativa para a Equação (2.41) é usada com $\mathbf{W}_{mix,a} = [\mathbf{W}_{p-1} \ \bar{\mathbf{P}}_{a-p+1}]$, em que $\bar{\mathbf{P}}_{a-p+1}$ são as primeiras $a-p+1$ colunas da matriz

característica do PCA1 quando PCA é aplicado a $\mathbf{R}$. Este método de diagnosticar falhas é conhecido por FDA/PCA2.

2.3.4 Comparação entre PCA e FDA

As técnicas PCA e FDA são aqui comparadas através de análises teóricas, ressaltando as diferenças geométricas entre os dois procedimentos de redução de dimensionalidade. O modo em que a FDA pode resultar em um diagnóstico de falha superior ao buscado pela PCA é também mostrado.

Os problemas de otimização para a PCA e para a FDA foram apresentados matematicamente nas Equações (2.2) e (2.37), respectivamente, mostrando que os vetores característicos PCA e FDA são obtidos através do cálculo dos pontos estacionários dos problemas de otimização

$$\max_{\mathbf{v} \neq 0} \frac{\mathbf{v}^T \mathbf{S}_t \mathbf{v}}{\mathbf{v}^T \mathbf{v}} \quad (2.43)$$

e

$$\max_{\mathbf{v} \neq 0} \frac{\mathbf{v}^T \mathbf{S}_t \mathbf{v}}{\mathbf{v}^T \mathbf{S}_w \mathbf{v}}, \quad (2.44)$$

respectivamente. As Equações (2.43) e (2.44) indicam que os vetores PCA e FDA são idênticos para o caso em que $\mathbf{S}_w = \sigma \mathbf{I}$, e $\sigma > 0$. Diferenças entre as duas técnicas podem ocorrer somente se existe prolongamento dos dados usados para descrever qualquer uma das classes. Essas formas prolongadas ocorrem para conjunto de dados altamente correlacionados, típico para dados coletados de processos industriais. Além disso, quando PCA e FDA são aplicadas da mesma maneira para dados do processo, espera-se que os vetores característicos a PCA e a FDA sejam significativamente diferentes, e os objetivos diferentes, Equações (2.43) e (2.44), sugerem que a FDA será significativamente melhor para discriminar entre classes de falhas.

A Figura (2.2) ilustra uma diferença entre PCA e FDA que ocorre quando a distribuição dos dados em uma das classes é um pouco alongada, ou seja, a classe ‘ $\times$ ’ possui um ponto a mais que a classe ‘ $\circ$ ’. A transformação linear dos dados sobre o primeiro vetor FDA é muito mais capaz de separá-los nas duas classes do que a transformação linear dos dados sobre o primeiro vetor característico PCA. Os dados utilizados nesta ilustração são aproximações dos dados apresentados por Chiang et al. (2001). Os vetores FDA e PCA são obtidos através da implementação dos algoritmos mostrados nesta e nas seções anterior. A Tabela (2.2) apresenta os dados utilizados.

Tabela 2.2: Dados originais e pré-tratados utilizados no exemplo comparativo entre PCA e FDA.

Dados originais				Dados pré-tratados			
Classe ×		Classe ○		Classe ×		Classe ○	
x	y	x	y	x	y	x	y
0,10	1,0	1,00	0,0	-1,64	-0,48	-0,20	-2,04
0,12	1,5	1,25	0,5	-1,61	0,290	0,20	-1,26
0,35	1,1	1,50	0,6	-1,24	-0,33	0,60	-1,11
0,70	1,8	1,60	1,2	-0,68	0,75	0,75	-0,17
0,80	1,4	1,70	1,5	-0,52	0,135	0,91	0,29
1,10	2,3	2,00	1,0	-0,04	1,53	1,40	-0,48
1,20	2,0	2,10	1,6	0,115	1,07	1,55	0,44
1,40	2,2	-	-	0,435	1,38	-	-

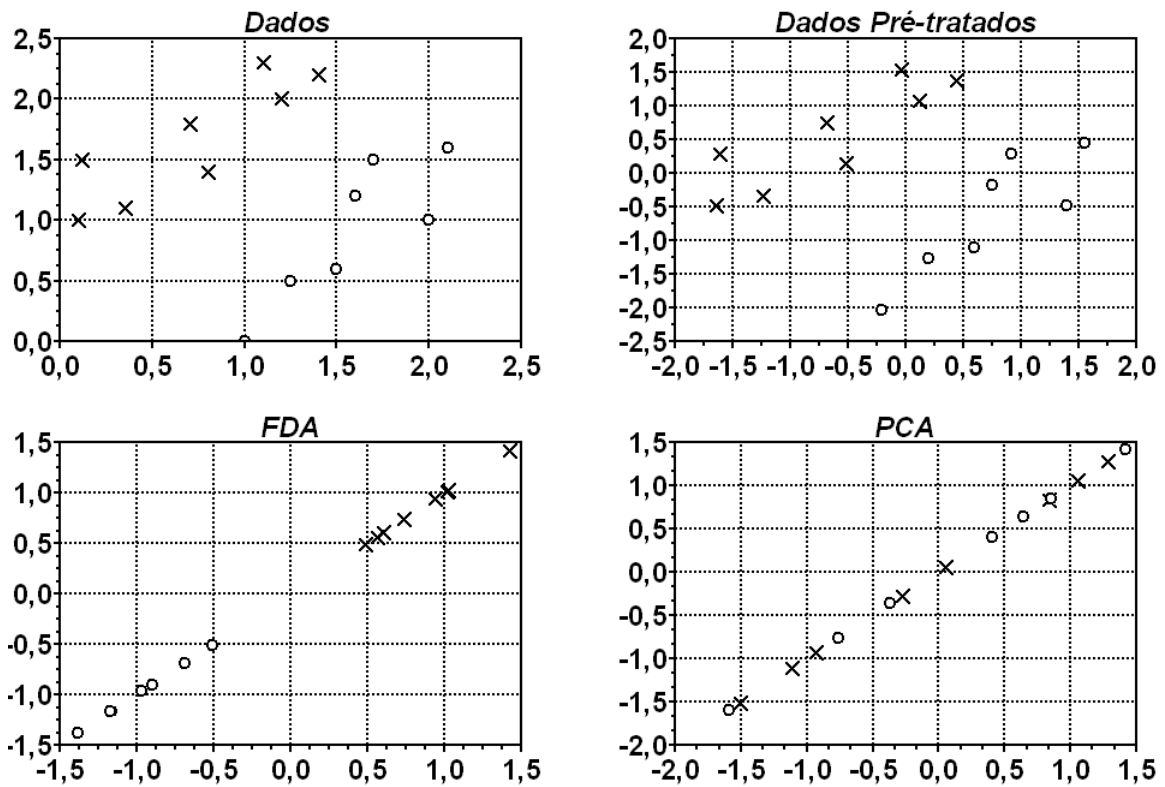


Figura 2.2: Comparação entre PCA e FDA para uma transformação linear dos dados nas classes × e ○ sobre os primeiros vetores característicos FDA e PCA.

Os primeiros vetores FDA e PCA obtidos foram:

$$\mathbf{w}_1 = \begin{bmatrix} -0,763669 \\ 0,6456070 \end{bmatrix} \quad \text{e} \quad \mathbf{p}_1 = \begin{bmatrix} 0,7071068 \\ 0,7071068 \end{bmatrix}.$$

2.3.5 Resumo da Técnica FDA

A Análise Discriminante de Fisher e suas variações estão resumidas na Tabela (2.3) para aplicações diretas de diagnóstico de falhas.

Tabela 2.3: Resumo da técnica FDA para diagnóstico de falhas.

Diagnóstico de Falhas		
Método	Base	Equação
FDA	T^2	(2.40)
FDA/PCA1	T^2	(2.41) e (2.40)
FDA/PCA2	T^2	(2.41) e (2.40)

2.4 Análise Estrutural

A análise estrutural usa estruturas de grafo e matriz de incidência para descrever as interações diretas entre as variáveis de um processo utilizando modelos. Os grafos obtidos podem ser usados para explorar redundâncias, que são úteis para a detecção de falhas e reconfiguração de controle. Na análise estrutural, falhas são vistas como violações de restrições. As análises mostram como componentes com falhas podem ser encontrados através da definição e utilização de relações de redundância apropriadas. Esta seção segue a estrutura proposta por Blanke et al. (2006), que apresenta a desde a construção do modelo estrutural até a utilização da técnica no contexto do controle tolerante a falhas.

2.4.1 Introdução

A disponibilidade de um modelo linear ou não linear torna possível que o mesmo seja reescrito, fornecendo assim uma interpretação útil do ponto de vista estrutural. As variáveis do processo são divididas em variáveis conhecidas e desconhecidas e apresentadas na forma de restrições. O modelo estrutural de um sistema é uma abstração do comportamento do modelo real, em que as ligações entre variáveis e parâmetros são consideradas. Essas ligações são representadas por um grafo bipartido (DULMAGE; MENDELSON, 1958, DULMAGE; MENDELSON, 1959), que independe do número de restrições, variáveis e dos valores dos parâmetros do modelo. Dessa forma, são exaltadas as características do modelo, que por sua vez são representadas numa forma facilitada de compreensão.

A utilização de modelos estruturais neste trabalho se deve ao fato de que as informações fornecidas por este tipo de representação são de extrema importância para a

detecção e diagnóstico de falhas, além de ser útil no projeto de estruturas de controle tolerante a falhas, visto que o resultado da análise estrutural mostra as medidas necessárias para monitorar o sistema como um todo. A análise estrutural é capaz de revelar relações de redundância analítica para geração de resíduos. Logo, este tipo de técnica possibilita a implementação de estruturas para validação de medições e até mesmo reconfiguração (GEHIN et al., 2000).

A reconstrução de um modelo qualquer na forma de um modelo estrutural se baseia na hipótese de que cada um de seus componente pode ser descrito por uma ou várias restrições. Uma falha pode ser detectada quando alguma dessas restrições é violada. Dessa forma, somente o modelo nominal do processo é necessário, que quando reescrito na forma estrutural, resulta numa forma de apresentação por grafo ou matriz de incidência. Utilizando estas representações é possível utilizar algoritmos capazes de propagar o conhecimento associado com as variáveis conhecidas para monitorar toda estrutura (variáveis desconhecidas e restrições).

2.4.2 Construção Estrutural do Modelo

A representação de um sistema por modelo estrutural mostra as ligações entre um conjunto de variáveis e o de restrições através de grafos (DULMAGE; MENDELSON, 1958) ou matrizes de incidência. Estes, são abstrações do comportamento do modelo original pois descrevem apenas as conexões entre as variáveis e as restrições e não mostram como as restrições realmente são. Dessa forma, um modelo estrutural é independente dos parâmetros.

Seguindo uma definição semelhante à aquela proposta por Blanke et al. (2006), com respeito ao comportamento de um sistema $(\mathcal{R}, \mathcal{V})$, em que $\mathcal{V} = \{v_1, v_2, \dots, v_N\}$ é um conjunto de variáveis e parâmetros, e $\mathcal{R} = \{r_1, r_2, \dots, r_M\}$ é um conjunto de restrições. As restrições são expressas de acordo com o tipo de modelo, por exemplo, equações algébricas, diferenciais, de diferenças, regras (*fuzzy*) etc.

Quando um conjunto de restrições formam um modelo, são feitas duas suposições para que se tenha um modelo estrutural bem construído.

Suposição 2.1 (BLANKE et al., 2006)

- (a) *Todas as restrições em $\mathcal{R}$ são compatíveis.*
- (b) *Todas as restrições em $\mathcal{R}$ são independentes.*

A Suposição 2.1(a) significa que o conjunto de restrições está associado com um modelo cujo conjunto solução exista. A Suposição 2.1(b) sobre a independência das restrições

significa que o modelo é mínimo no sentido que nenhuma restrição define (no mínimo localmente) o mesmo conjunto de soluções como uma outra define.

A representação estrutural de um sistema é feita através de um grafo bipartido. Este é uma representação gráfica das variáveis e restrições, ou seja, esta representação ocorre quando existe a separação em $\mathcal{R}$ e em $\mathcal{V}$ de tal forma que cada conexão possui um ponto final em $\mathcal{R}$ e um outro em $\mathcal{V}$.

Definição 2.1 (Modelo Estrutural) (BLANKE et al., 2006): *O modelo estrutural (ou a estrutura) do sistema $(\mathcal{R}, \mathcal{V})$ é um grafo bipartido $(\mathcal{R}, \mathcal{V}, \mathcal{C})$ em que $\mathcal{C} \subset \mathcal{R} \times \mathcal{V}$ é um conjunto de conexões definido por:*

$(r_i, v_j) \in \mathcal{C}$ se a variável v_j aparece na restrição r_i .

Um grafo bipartido pode ser interpretado como um conjunto de variáveis e parâmetros que se ligam de forma não direcionada às restrições, chamadas de vértices. Assim, todas as variáveis e parâmetros devem satisfazer a equação ou regra que representa este vértice. Estas representações englobam estrutura de modelos gerais, incluindo diferentes tipos de restrições, por exemplo, diferenciais ou algébricas. A matriz de incidência de um grafo bipartido é a matriz cujas linhas e colunas representam um conjunto de restrições e variáveis, respectivamente. Cada conexão $(r_i, v_j) \in \mathcal{C}$ é representada por um ‘ $\square$ ’ na intersecção da linha r_i com a coluna v_j .

Exemplo 2.1: Seja o modelo de um sistema linear descrito por quatro restrições (r_1 , r_2 , r_3 e r_4) e cinco variáveis (u , y_1 , y_2 , $\dot{y}_1$ e $\dot{y}_2$), dado por:

$$r1 : \dot{y}_1 = \frac{dy_1}{dt},$$

$$r2 : \dot{y}_1 = ay_2,$$

$$r3 : \dot{y}_2 = \frac{dy_2}{dt},$$

$$r4 : \dot{y}_2 = by_1 + cy_2 + du.$$

A matriz de incidência é representada por:

$\nearrow$	u	y_1	y_2	$\dot{y}_1$	$\dot{y}_2$
r_1		$\square$		$\square$	
r_2			$\square$	$\square$	
r_3			$\square$		$\square$
r_4	$\square$	$\square$	$\square$		$\square$

O grafo bipartido do sistema é mostrado na Figura (2.3):

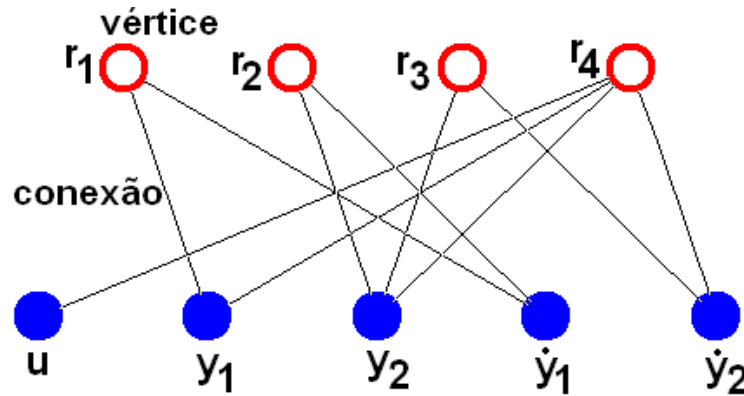


Figura 2.3: Grafo bipartido do sistema linear do Exemplo 2.1.

Para sistemas descritos por equações lineares de estado e medições, uma representação estrutural comum é na forma de grafo direcionado (dígrafo), cujo conjunto de vértices é o conjunto de variáveis de entrada (u), saída (y) e de restrições (r) e cujas conexões são definidas pelas seguintes regras baseadas no Exemplo 2.1 e que podem ser estendidas para qualquer sistema:

- uma conexão existe do vértice y_1 para o vértice r_1 se e somente se a variável de estado y_1 ocorre na função $\dot{y}_1$.

Na representação de dígrafos, conexões são interpretadas como “influências mutuais” entre variáveis: uma conexão de y_1 para r_1 significa que o tempo de evolução da derivada $\dot{y}_1(t)$ depende da evolução de $y_1(t)$ (BLANKE et al., 2006).

Uma visão geral sobre variáveis conhecidas, desconhecidas e restrições

O sistema estrutural como um todo é composto por variáveis conhecidas, desconhecidas e restrições. Variáveis de entrada e saída são medidas em processos reais. Da mesma forma, parâmetros que compõe o modelo são previamente identificados e assim conhecidos. As variáveis medidas são conhecidas em linha e são úteis para esquemas de detecção e diagnóstico de falhas ou algoritmos de controle tolerante, visto que podem ser projetadas sobre modelos estatísticos para a verificação de algum comportamento anormal. Já as variáveis desconhecidas não são medidas diretamente e a análise estrutural busca alguma forma de calcular seus valores através dos valores das variáveis conhecidas (observação).

O conjunto de variáveis é então dividido em

$$\mathcal{V} = \mathcal{K} \cup \mathcal{U},$$

em que $\mathcal{K}$ é o subconjunto de variáveis conhecidas e parâmetros e $\mathcal{U}$ é o subconjunto de variáveis desconhecidas. Similarmente, o conjunto de restrições é particionado em

$$\mathcal{R} = \mathcal{R}_{\mathcal{K}} \cup \mathcal{R}_{\mathcal{U}},$$

em que $\mathcal{R}_{\mathcal{K}}$ é o subconjunto de restrições que se ligam apenas com variáveis conhecidas e $\mathcal{R}_{\mathcal{U}}$ inclui restrições em que no mínimo uma variável desconhecida aparece.

Em alguns desenvolvimentos futuros, será de interesse focar no subsistema $(\mathcal{R}_{\mathcal{U}}, \mathcal{V})$ que é um chamado de grafo de estrutura reduzida. Neste grafo, existem restrições que referem a no mínimo uma variável desconhecida $v_i \in \mathcal{U}$. A observabilidade de um sistema (monitoramento das variáveis desconhecidas utilizando apenas as conhecidas) depende do sub-grafo $(\mathcal{R}_{\mathcal{U}}, \mathcal{U}, \mathcal{C}_{\mathcal{U}})$, resultante do grafo completo no qual as variáveis conhecidas $v_i \in \mathcal{K}$ foram apagadas. Neste trabalho, variáveis desconhecidas são representadas por um círculo cheio, restrições por um círculo vazio e variáveis conhecidas por um quadrado nas representações na forma de grafos (ver Figura (2.3)).

2.4.3 Pareamento: definições e aplicações

Tem-se o pareamento, que é uma escolha proposital de associação de variáveis desconhecidas do processo com as restrições que podem calcular tais variáveis, como uma ferramenta essencial para a análise estrutural. Sendo assim, variáveis desconhecidas que não participam de qualquer pareamento não podem ser calculadas. Se as variáveis podem ser pareadas por diferentes maneiras, é possível então gerar relações de redundância analítica para detecção e diagnóstico de falhas, além de informações para a reconfiguração (BERGE, 1957).

Definições

Seja $(\mathcal{R}, \mathcal{V}, \mathcal{C})$ um grafo bipartido, $c \in \mathcal{C}$, $c = (\alpha, \beta)$ uma conexão entre a restrição α e a variável β , e $m_{\mathcal{R}}$ e $m_{\mathcal{V}}$ as duas projeções (BLANKE et al., 2006)

$$\begin{aligned} m_{\mathcal{R}} : \mathcal{C} &\rightarrow \mathcal{R} \\ c &\mapsto m_{\mathcal{R}}(c) = \alpha, \end{aligned}$$

$$\begin{aligned} m_{\mathcal{V}} : \mathcal{C} &\rightarrow \mathcal{V} \\ c &\mapsto m_{\mathcal{V}}(c) = \beta. \end{aligned}$$

A projeção da conexão sobre o conjunto de restrições é $m_{\mathcal{R}}(c) = \alpha$ (o nó de restrição

da conexão c) e a projeção da conexão sobre o conjunto de variáveis é $m_{\mathcal{V}}(c) = \beta$ (o nó de variável da conexão c).

Definição 2.2 (Pareamento): *Um pareamento $\mathcal{P}$ é um subconjunto de $\mathcal{C}$ tal que as restrições de $m_{\mathcal{R}}$ e $m_{\mathcal{V}}$ em $\mathcal{P}$ são injetivas, isto é (BERGE, 1957)*

$$\forall c_1, c_2 \in \mathcal{P} : c_1 \neq c_2 \Rightarrow m_{\mathcal{R}}(c_1) \neq m_{\mathcal{R}}(c_2) \wedge m_{\mathcal{V}}(c_1) \neq m_{\mathcal{V}}(c_2)$$

Esta expressão mostra que um pareamento é um subconjunto de conexões tal que quaisquer duas conexões não possuem nó em comum ($\mathcal{R}$ nem $\mathcal{V}$).

Definição 2.3 (Pareamento máximo): *Um pareamento máximo é um pareamento $\mathcal{P}$ tal que $\forall \mathcal{N} \in 2^{\mathcal{C}}$ com $\mathcal{P} \subset \mathcal{N}$, $\mathcal{N}$ não é um pareamento (BERGE, 1957, CHARTRAND; OELLERMANN, 1993).*

Assim, um pareamento máximo é um pareamento que não permite adição de conexões sem violar a propriedade de nó comum, ou seja, a própria definição de pareamento. Como o conjunto de pareamentos P é parcialmente ordenado, segue que existe, em geral, mais que um pareamento máximo.

Seja $P^* \subseteq P$ o conjunto de pareamentos máximos. Estendendo a definição de projeções $m_{\mathcal{R}}$ e $m_{\mathcal{V}}$ para os conjuntos de conexões (em vez de um única conexão),

$$\begin{aligned} \pi_P : P &\rightarrow 2^{\mathcal{R}} \\ \mathcal{P} &\mapsto \pi_{\mathcal{R}}(\mathcal{P}) = \{r \in \mathcal{R}; \exists c \in \mathcal{P} \text{ tal que } r = m_{\mathcal{R}}(c)\}, \end{aligned}$$

$$\begin{aligned} \pi_P : P &\rightarrow 2^{\mathcal{V}} \\ \mathcal{P} &\mapsto \pi_{\mathcal{V}}(\mathcal{P}) = \{v \in \mathcal{V}; \exists c \in \mathcal{P} \text{ tal que } v = m_{\mathcal{V}}(c)\}, \end{aligned}$$

cada pareamento $\mathcal{P}$ está associado com o subconjunto $\pi_{\mathcal{R}}(\mathcal{P})$ de suas restrições pareadas, e com o subconjunto $\pi_{\mathcal{V}}(\mathcal{P})$ de suas variáveis pareadas. Como não mais que uma restrição ou variável pode estar associada com cada conexão do pareamento, segue que qualquer pareamento satisfaz a seguinte propriedade

$$\forall \mathcal{P} \in P^* \quad \begin{aligned} \pi_{\mathcal{R}}(\mathcal{P}) &\subseteq \mathcal{R} \\ \pi_{\mathcal{V}}(\mathcal{P}) &\subseteq \mathcal{U} \end{aligned}$$

da qual segue a seguinte relação válida

$$|\mathcal{P}| \leq \min\{|\mathcal{R}|, |\mathcal{V}|\}$$

onde as $|\cdot|$ denotam a cardinalidade dos conjuntos (que significa o número de elementos). Pareamentos em que os sinais de igualdade são válidos são chamados de pareamentos

completos.

Definição 2.4 (Pareamento completo): *Um pareamento é completo com respeito a $\mathcal{R}$ se a igualdade $|\mathcal{P}| = |\mathcal{R}|$ é válida. Dessa forma, um pareamento é completo com respeito a $\mathcal{V}$ se $|\mathcal{P}| = |\mathcal{V}|$ é válida (BERGE, 1957, CHARTRAND; OELLERMANN, 1993).*

Para um pareamento completo $\mathcal{P}$ em $\mathcal{R}$ ou em $\mathcal{V}$, cada restrição ou variável pertence a exatamente uma conexão do pareamento:

$$\forall r \in \mathcal{R} : \exists v \in \mathcal{V} \text{ tal que } (r, v) \in \mathcal{P},$$

$$\forall v \in \mathcal{V} : \exists r \in \mathcal{R} \text{ tal que } (r, v) \in \mathcal{P}.$$

Um pareamento é representado neste trabalho pela seleção de no máximo um símbolo “□” em cada linha e em cada coluna na matriz de incidência do grafo bipartido, que o transforma em “■” nos exemplos. Cada “□” selecionado representa uma conexão do pareamento. Nenhum outro pareamento deve conter a mesma variável (somente um em cada linha) ou a mesma restrição (somente um em cada coluna).

A definição de pareamentos, pareamentos completos e máximos pode levar em consideração toda estrutura do sistema ou apenas sub-grafos, ou seja, subconjuntos de restrições e variáveis. Sabe-se que as variáveis desconhecidas não precisam ser determinadas por qualquer restrição. Então, participam do pareamento apenas os sub-grafos que contêm as variáveis desconhecidas e restrições ligadas a elas. No entanto, na apresentação de pareamentos na forma de matriz de incidência é comum considerar todo o sistema estrutural.

Pareamento e grafos orientados

Quando um pareamento é definido numa estrutura de grafo, orientações para as conexões são feitas. Restrições não possuem direção, ou seja, antes do pareamento não é possível saber qual variável será determinada por qual restrição. Por exemplo, seja a seguinte restrição genérica

$$r_1 : y_1(t) - y_2(t) - \dot{m}(t) = 0. \quad (2.45)$$

Cada um das três variáveis pode ser calculada por r_1 quando outras duas variáveis forem conhecidas. A forma acima é não orientada, logo não existe preferência para o cálculo de qualquer uma das três variáveis.

Quando um pareamento é escolhido, ocorre então uma quebra da igualdade de pre-

ferência de cada variável, pois se uma restrição é pareada, conseqüentemente ocorre a determinação de uma variável por esta restrição. Variáveis pareadas e não pareadas são identificadas na matriz de incidência do grafo por “■” ou “□”, respectivamente.

A representação de um modelo na forma de um grafo pode incluir as não simetrias associadas com um pareamento através da transformação de conexões originalmente não orientadas em conexões orientadas, gerando assim um dígrafo. Para que isto seja possível, as regras a seguir podem ser aplicadas:

- **Restrições pareadas e não pareadas:** As conexões adjacentes a um pareamento são feitas com uma orientação: da variável não pareada para a restrição (entrada) e da restrição para a variável pareada (saída).

Essas regras são entendidas quando considera-se um pareamento $\mathcal{P}$ qualquer e escolhe-se uma conexão $(r, x) \in \mathcal{P}$. Se a variável x é considerada como saída da restrição r , as outras variáveis são consideradas entradas. Com isso, o pareamento representa uma relação de causa e efeito, ou seja, a restrição r é utilizada para calcular a variável x . Com isso, tem-se

$$x = f(r)$$

Restrições que não sofrem pareamento possuem todas variáveis como entrada e a variável de saída deve ser igual a zero.

Relações de causa e efeito

Como comentando anteriormente, a escolha de um pareamento implica numa relação de causa e efeito, pois quando se escolhe um pareamento (r, v) , a restrição r será utilizada para calcular a variável v , supondo que as outras variáveis pertencentes à restrição são conhecidas. Ao grafo resultante dessas atribuições (pareamentos) dá-se o nome de grafo causal. Grafos causais são usados em razões qualitativas, filtragem de alarmes ou no fornecimento de cadeias de cálculo necessárias para a determinação formal de algumas variáveis de interesse. As análises anteriores devem ser analisadas mais cuidadosamente quando ocorre a presença de laços e restrições diferenciais (BLANKE et al., 2006).

Do ponto de vista da interpretação causal, pelo menos uma variável pode ser pareada em uma dada restrição, mas não necessariamente pode-se afirmar que qualquer variável possua esta propriedade. Por exemplo, quando não é possível determinar x utilizando r devido ao fato de r não ser inversível com relação a x .

O caso de restrições diferenciais deve ser considerado mais cuidadosamente. Lem-

brando que as restrições diferenciais são representadas neste trabalho por:

$$r : x_2(t) - \frac{d}{dt}x_1(t) = 0. \quad (2.46)$$

As funções $x_1(t)$ e $x_2(t)$ não podem ser escolhidas de forma independente. Com isso, o conhecimento da trajetória de $x_1(t)$ implica no conhecimento de sua derivada (de um ponto de vista analítico, supõe-se neste trabalho que as derivadas existem, e de um ponto de vista numérico, pode existir problema de cálculo devido à presença de ruídos). Assim, essa restrição pode ser pareada para $x_2(t)$, este fato é conhecido como causalidade derivativa. Já quando apenas $x_2(t)$ é conhecido, parear r para $x_1(t)$ (fenômeno conhecido como causalidade integral), leva a:

$$x_1(t) = x_1(0) + \int_0^t x_2(\sigma) d\sigma, \quad (2.47)$$

que não determina de forma única $x_1(t)$, ao menos que a condição inicial $x_1(0)$ seja conhecida. Num contexto de simulação, condições iniciais são sempre conhecidas.

Algoritmos de pareamento

A realização de um pareamento neste trabalho é representada na matriz de incidência do grafo bipartido pela seleção de pelo menos um “□” em cada linha e em cada coluna. Cada “□” selecionado (que se transforma em ■) representa uma conexão do pareamento, e esta deve ser única, ou seja, não é permitido que uma conexão contenha mais do que uma variável. Pareamentos e pareamentos máximos podem ser definidos para qualquer grafo e tal fato é explorado para solução de problemas de diferentes tipos.

Ao analisar as relações de causa e efeito resultantes de um pareamento em grafo bipartido, nota-se que um pareamento completo nas variáveis mostra os cálculos que devem ser feitos para determinar seus valores utilizando os valores das variáveis conhecidas. A existência de restrições não pareadas leva diretamente à geração de relações de redundância analítica, ou seja, restrições não pareadas que apenas recebem variáveis (conhecidas e desconhecidas) e retornam uma saída igual a zero. A idéia principal para o algoritmo de posicionamento, que se baseia na propagação do conhecimento e que gera apenas grafos orientados e sem laços, é iniciar o algoritmo com alguma variável conhecida e propagar este conhecimento, passo a passo, realizando pareamentos, a cada etapa, das variáveis que influenciam nas restrições em que todas outras variáveis envolvidas são pareadas ou conhecidas (BERGE, 1957, CHARTRAND; OELLERMANN, 1993).

Algoritmo 2.2: Posicionamento das restrições (BLANKE et al., 2006)

Dado: Matriz de incidência ou grafo.

1. Marcar todas variáveis conhecidas, $i = 0$.
2. Encontrar todas restrições na matriz de incidência com exatamente uma variável não marcada. Associar a posição i com estas restrições e marcá-las, bem como a variável correspondente.
3. Faça: $i = i + 1$.
4. Se existem restrições não marcadas cujas variáveis estão todas marcadas, associá-las com a posição i , marcá-las e conectá-las com uma falsa variável ZERO (saída igual a zero).
5. Se existem variáveis ou restrições não marcadas, volte para a etapa 2.

Resultado: Ordem de cálculo das variáveis desconhecidas.

Analisando o algoritmo anterior nota-se que a primeira etapa é destinada à marcação de todas variáveis conhecidas $\mathcal{K}$. Pelo contrário, todas variáveis desconhecidas permanecem não marcadas. Então, cada restrição que contém no máximo uma variável não marcada recebe atribuição 0 para posição. Restrições não pareadas sofrem pareamento a cada etapa do algoritmo e são imediatamente incluídas no conjunto de variáveis conhecidas. Dessa forma, toda estrutura é percorrida e o algoritmo termina quando não é mais possível adicionar mais variáveis desconhecidas ao conjunto de variáveis conhecidas devido ao pareamento.

Para cada variável pareada é fornecido um número, que significa a posição. Devido a esta atribuição, o algoritmo é denominado de algoritmo de posicionamento. A posição é interpretada como o número de etapas necessárias para calcular uma variável desconhecida através das variáveis conhecidas, ou seja, a posição com que esta variável será calculada dentro de um conjunto de variáveis desconhecidas $\mathcal{U}$.

Em resumo, a utilização de algoritmos de pareamentos completos sobre as variáveis geram relações de redundância analíticas, que são restrições não pareadas que recebem informações apenas de variáveis conhecidas e retornam um valor igual a zero, ou seja, são resíduos estruturados (GERTLER; SINGER, 1990) obtidos através de grafos bipartidos. Relações de redundância são compostas por cadeias alternadas, que se iniciam com variáveis conhecidas e que terminam com restrições não pareadas.

2.4.4 Decomposição de Grafos

De acordo com a teoria dos grafos, afirma-se que qualquer grafo pode ser decomposto em três sub-grafos com propriedades particulares que são associadas com um subsistema super restrito, um exatamente restrito e um sub-restrito. Esta decomposição é única e sua importância é que através delas as propriedades estruturais de observabilidade, controlabilidade, monitorabilidade, reconfigurabilidade são reveladas.

Definição 2.5 (super restringindo, exatamente restringido e sub-restringido)

Um grafo $(\mathcal{R}, \mathcal{V}, \mathcal{C})$ é chamado de (BLANKE et al., 2006)

- *super restringido*, se existe um pareamento completo sobre as variáveis $\mathcal{V}$ mas não sobre as restrições $\mathcal{R}$,
- *exatamente restringido*, se existe um pareamento completo sobre as variáveis $\mathcal{V}$ e sobre as restrições $\mathcal{R}$,
- *sub-restringido*, se existe um pareamento completo sobre as restrições $\mathcal{R}$ mas não sobre as variáveis $\mathcal{V}$.

2.4.5 Observabilidade estrutural

Teorema 2.1 (Observabilidade estrutural) (MUROTA, 1987)

Uma condição necessária e suficiente para que um sistema de equações seja estruturalmente observável é que, sob causalidade derivativa

1. *todas variáveis desconhecidas são alcançadas pelas conhecidas,*
2. *os subsistemas super e exatamente restringidos são causais,*
3. *o subsistema sub-restringido é vazio.*

A condição 1 expressa que não existe qualquer subsistema cujo comportamento não é refletido no comportamento das variáveis conhecidas, enquanto que as condições 2 e 3 expressam que todas variáveis podem ser pareadas usando pareamentos causais.

Observabilidade baseada na teoria dos grafos

O pareamento completo de um sistema de restrições e variáveis resulta em um dígrafo, assim, analisa-se a observabilidade do sistema utilizando a teoria dos grafos. Para que

x , que é uma variável desconhecida, seja uma variável observável, é necessário que exista no mínimo um conjunto de cadeias alternadas que conecte esta variável ao conjunto de variáveis conhecidas.

Para que a variável x seja calculada é necessário percorrer uma seqüência de cálculos ao longo de uma cadeia alternada. Existe a possibilidade de análise automática para fornecer uma expressão formal para x . Além disso, é possível implementar estas seqüências de cálculos em linha para se obter uma simulação estrutural. É importante observar que esta interpretação expressa que x pertence a um subsistema exatamente ou super restrito e causal. De outra forma, se x pertencesse a um subsistema sub restrito, o sub-grafo correspondente poderia ter menos restrições que variáveis. Assim, algumas variáveis desconhecidas poderiam ser sinais de entrada para restrições e não de saída, provando assim uma não observabilidade.

2.4.6 Monitorabilidade

Um sistema é monitorável quando é possível usar apenas variáveis conhecidas para sua determinação, independente das restrições serem satisfeitas. A análise de monitorabilidade de sistemas estruturais é feita nesta seção para projetar algoritmos de detecção e diagnóstico de falhas baseado em relações de redundância analítica. O diagnóstico de falhas utilizando redundância analítica tenta identificar falhas através da comparação do comportamento real do sistema, que é observado através da evolução do tempo das variáveis conhecidas, com o comportamento teórico descrito pelas restrições do sistema (STAROSWIECKI; DECLERCK, 1989, DECLERCK; STAROSWIECKI, 1991).

Sistemas FDI baseados em redundância analítica

Entende-se por relações de redundância analítica as restrições que ligam a evolução do tempo das variáveis conhecidas quando o sistema opera em concordância com seu modelo. Uma vez projetadas, o procedimento de detecção de falhas verifica a cada instante se elas são satisfeitas, quando não, procedimentos de identificação ou diagnóstico verificam quais partes do sistema estão sob efeito de falhas ou mais especificamente, qual falha atua sobre o sistema. A existência de relações de redundância analítica é um pré-requisito para projetar sistemas de diagnóstico de falhas. Para o funcionamento correto destes sistemas, as relações devem ter as seguintes propriedades:

- Robustas: insensíveis a entradas e parâmetros desconhecidos. Isto garante que as relações são satisfeitas quando nenhuma falha está presente, prevenindo alarmes

falsos.

- Sensíveis a falhas: garante que as relações são satisfeitas apenas na presença de falhas.
- Estruturadas: garante que na presença de uma dada falha, somente um subconjunto de relações de redundância analítica não é satisfeito, permitindo assim reconhecer (de um subconjunto de relações satisfeitas e outro não), que a falha que ocorreu.

Projeto de resíduos robustos e estruturados

A realidade prática mostra que a existência de perturbações, entradas desconhecidas e parâmetros incertos faz com que a resposta nominal dada por um modelo nunca seja igual àquela apresentada pelo comportamento real. Essas características produzem resíduos que podem resultar em alarmes falsos.

Uma propriedade desejada para os resíduos é o não fornecimento de alarmes falsos, ou seja, utiliza-se o termo robustez para o projeto de resíduos para atribuir a propriedade de insensibilidade a parâmetros incertos ou entradas desconhecidas no que diz respeito a alarmes falsos. Esses resíduos ainda são satisfeitos quando nenhuma falha está presente. Para contornar o efeito de parâmetros incertos no projeto de resíduos, é possível considerá-los como variáveis desconhecidas para serem determinadas, no entanto, ocorre uma diminuição no número de relações de redundância disponíveis, já que o número de variáveis aumenta.

Um resíduo é uma estrutura composta por restrições, quando estas e os valores que as formam são conhecidos e inseridos em um resíduo, é possível verificar se o mesmo é nulo ou não. Se existe correspondência com operação normal ou com falha. Para falhas detectáveis, espera-se resíduos não nulos. Da mesma forma, para restrições que são modificadas devido a falhas, espera-se que resíduos que contenham em sua estrutura essas restrições também apresentem resíduos não nulos. Existem falhas que afetam todos os resíduos, logo, esses tipos de falhas não são isoláveis. Por outro lado, existem falhas que afetam apenas alguns resíduos, e estes são úteis para sistemas de diagnóstico de falhas (GERTLER; SINGER, 1980).

CAPÍTULO 3

Propostas de Controle Tolerante a Falhas

Neste capítulo faz-se uma apresentação dos métodos disponíveis para lidar com falhas em sensores, atuadores e sistemas de controle. Ao longo deste ficará claro que o controle tolerante não é apenas uma estrutura teórica, mas emprega idéias específicas para tratar diferentes tipos de problemas. Uma outra contribuição deste capítulo é a apresentação formal do problema de controle tolerante e como a presença de falhas em processos podem ser acomodadas. As propostas para obtenção de controle mostradas se aplicam para três tipos de controladores explorados: PID, ótimo e preditivo baseado em modelo. Dessa forma, parte-se de uma definição padrão para um problema de controle tolerante e desenvolve-se técnicas que utilizam as informações apresentadas no Capítulo 2 para tratar determinados tipos de problemas. A generalização das técnicas propostas é apresentada e discutida através de exemplos no Capítulo 4.

3.1 Introdução ao Problema de Controle Tolerante

Um problema de controle padrão é definido de acordo com um dado objetivo, um conjunto de restrições e um conjunto de leis de controle possíveis. Este problema tenta encontrar uma lei de controle num conjunto de leis de controle U , de tal forma que o sistema que está sendo controlado alcance os objetivos de controle O enquanto seu comportamento satisfaz um conjunto de restrições R . Assim, o problema de controle é formalmente definido.

Uma explicação mais elaborada é apresentada por Blanke et al. (2006) com o intuito

de explicar os componentes principais de um problema de controle bem posto.

- O conjunto U de leis de controle possíveis definem os algoritmos que podem ser implementados, por exemplo, controle em malha aberta (um mapeamento do domínio do tempo para o espaço de controle), controle em malha fechada (um mapeamento da saída *versus* o espaço de referências para o espaço de controle).
- O objetivo O define o que se espera que o sistema alcance, quando controlado pela lei de controle mencionada acima.
- Restrições R são relações funcionais que o comportamento dos sistemas controlados devem satisfazer com o tempo. Elas são expressas por restrições algébricas, diferenciais ou de diferenças. Restrições de desigualdade expressam que algumas saturações atuam sobre as soluções admissíveis do sistema.

3.1.1 Falhas na lei de controle

Para transformar um controle de processo convencional em um controle tolerante a falhas é necessário que o mesmo adquira a habilidade de controlar na presença de falhas. Para alcançar este objetivo, algumas técnicas podem ser utilizadas, por exemplo, alterar a lei de controle (adaptação e acomodação são termos comuns na literatura), ou alterar ambos, a lei de controle e o sistema (reconfiguração), são alternativas eficientes. Como os algoritmos de controle apenas implementam a solução do problema de controle para um dado sistema, a alteração do controle ou do sistema significa que o problema de controle também se altera devido a presença de falhas (BLANKE et al., 2006). A aplicação de diferentes estratégias de controle tolerante está de acordo com o tipo de falha presente no sistema. Supondo que a falha seja no conjunto de restrições, então problema passa de (O, R, U) para resolver $(O, R(\kappa), U)$. Sendo que, $R(\kappa)$ denota a dependência da restrição R sobre os parâmetros κ , que por sua vez dependem de falhas.

O aparecimento de falhas não deve alterar os objetivos do sistema. Assim, o controle tolerante tenta alcançar esses objetivos na presença de falhas. Isto pode ser possível ou não, definindo assim, dois casos:

1. Quando existe uma forma de continuar buscando os objetivos propostos para o sistema na presença de certas falhas o sistema é dito ser tolerante a falhas, com respeito aos objetivos e às falhas. O objetivo do engenheiro é projetar alguma lei de controle que seja capaz de realizar tal tarefa.
2. Quando os objetivos não podem ser alcançados na presença das falhas consideradas o sistema não é tolerante a falhas com respeito àqueles objetivos e falhas. Como

os objetivos correntes não podem ser alcançados, transforma-se o problema em encontrar novos objetivos que são de interesse da situação atual e projeta-se a lei de controle que é capaz de buscar esses novos objetivos.

De acordo com Blanke et al. (2006), a presença de falhas pode alterar as restrições $R(\kappa)$ do problema de controle de formas distintas:

- As restrições podem continuar as mesmas mas os parâmetros podem mudar, transformando o problema de controle $(O, R(\kappa_n), U)$ em $(O, R(\kappa_f), U)$, em que κ_f e κ_n denotam os parâmetros do sistema no caso nominal e com falhas, respectivamente.
- As restrições por si próprias podem se alterar, transformando o problema de controle $(O, R_n(\kappa_n), U)$ no problema $(O, R_f(\kappa_f), U)$, em que R_n é o conjunto de restrições nominais, e $R_f(\kappa_f)$ é o conjunto de restrições com novos parâmetros associados.

O ocorrência de falhas também pode alterar o conjunto de leis de controle admissíveis, visto que falhas podem ocorrer diretamente nos instrumentos responsáveis pela comunicação e fechamento da malha de controle. Dessa forma, denota-se o novo conjunto de leis de controle admissíveis por U_f , enquanto a nominal é U_n .

3.1.2 Proposta de controle tolerante

A proposta de um sistema de controle tolerante a falhas utilizada neste trabalho se baseia na aplicação das técnicas de detecção e diagnóstico proposta no Capítulo 2, visto que quando anexadas ao sistema de controle principal, este adquire a capacidade de tolerância falhas.

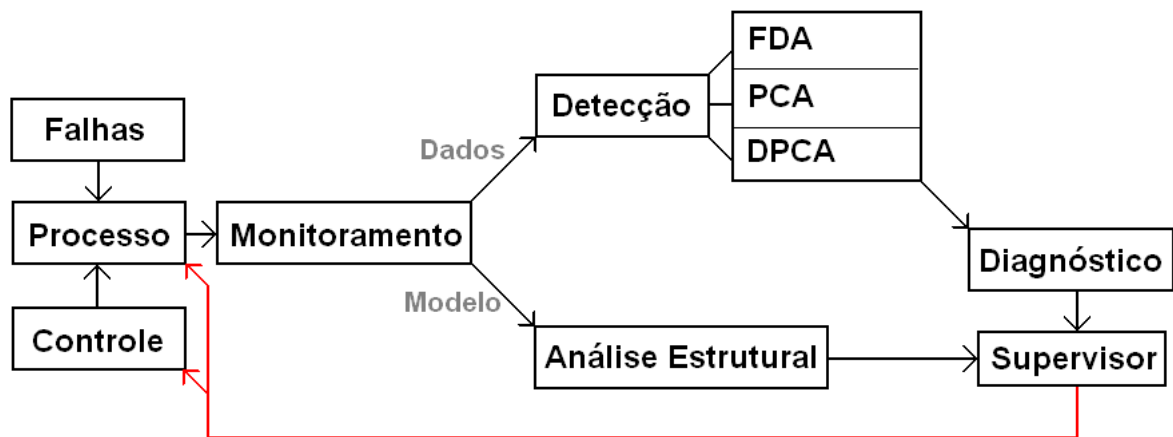


Figura 3.1: Proposta de um sistema de controle tolerante a falhas utilizada neste trabalho.

A etapa inicial que deve ser percorrida para que o sistema de controle tolerante proposto pela Figura (3.1) seja obtido é a implementação de sistemas de detecção e diagnóstico de falhas baseados em dados históricos ou modelo de um processo. Para isto, utiliza-se as técnicas de análise dos componentes principais, análise discriminante de Fisher, quando dados estão disponíveis e análise estrutural quando um modelo para o processo é utilizado. Ao final, as informações obtidas pelo monitoramento são enviadas para um sistema supervisor, responsável por atuar no processo, no controlador ou em ambos, para que a operação controlada na presença de falhas seja aceitável.

O esquema proposto pela Figura (3.1) é geral com respeito ao tipo de controlador e sistema supervisor, para falhas que aparecem no processo. No entanto, as técnicas de monitoramento são específicas para a PCA, FDA e análise estrutural. A técnica CVA (*Canonical Variate Analysis*, Análise Canônica Variada) é bastante explorada para detecção de falhas, assim como apresentado por Lee et al. (2006). Dessa forma, a ampliação do esquema com técnicas similares pode ser feita sem a descaracterização do controle tolerante a falhas proposto.

O bloco supervisor mostrado na Figura (3.1) tem como finalidade principal a introdução de um conjunto de possíveis objetivos de controle ao problema de controle corrente. É um problema de decisão que verifica a cada instante de tempo as possibilidades do sistema e as determina dependendo do estado corrente. Um problema de supervisão é um problema de controle tolerante a falhas associado com um problema de decisão, por exemplo, se ocorre uma falha que não pode ser tolerada pelo sistema de controle, os objetivos do sistema devem ser alterados.

Para algumas falhas, pode ocorrer a impossibilidade de alcançar qualquer objetivo. É comum então a manutenção da operação com falha para eventos com pouca possibilidade de ocorrência. Entretanto, operações seguras garantem o desligamento seguro de processos quando falhas que não podem ser toleradas ocorrem.

3.1.3 Controle tolerante ativo versus passivo

Na aproximação passiva, o algoritmo de controle é projetado para que o sistema seja capaz de buscar os objetivos propostos, em situações normais bem como em situações com falhas, sem qualquer alteração da lei de controle. Além disso, a habilidade do sistema ser capaz de alcançar seus objetivos é preservada, usando a mesma lei de controle, independente da situação do sistema.

No controle tolerante ativo a lei de controle se altera na presença de falhas, mantendo a capacidade do sistema de alcançar os objetivos propostos. Assim, uma lei de controle

é utilizada em cada situação. Como o resultado de falhas, o problema de controle é transformado de $(O, R_n(\kappa_n), U_n)$ em $(O, R_f(\kappa_f), U_f)$. Para se ter um controle tolerante a falhas ativo é necessário ter conhecimento sobre $R_f(\kappa_f)$ e U_f , isto pode ser feito por algoritmos de detecção e diagnóstico de falhas apresentados no Capítulo 2.

Supondo que $R_f(\kappa_f)$ e U_f são conhecidos, então a lei de controle tolerante a falhas deve resolver o problema $(O, R_f(\kappa_f), U_f)$. Se existe uma solução, o sistema é tolerante a falhas com respeito aos objetivos O e à situação de falha $R_f(\kappa_f), U_f$. No entanto, se o problema $(O, R_f(\kappa_f), U_f)$ não possui solução, então o sistema não é tolerante a falhas e o objetivo de reconfiguração deve ser explorado.

3.1.4 Utilização de informações sobre falhas

Fornecer informações sobre o impacto de falhas é a tarefa principal do sistema de diagnóstico. Entretanto, esse algoritmos possuem eficiência limitada. A detecção de falhas informa que o problema a ser resolvido não é mais $(O, R_n(\kappa_n), U_n)$. A identificação de falhas mostra o subconjunto de restrições $R_n(\kappa_n)$ que não sofreram alteração e o subconjunto $U_f \subseteq U_n$ de leis de controle que ainda podem ser utilizadas. O conhecimento sobre as restrições alteradas é feito através da estimação de falha, que é uma nova função a ser considerada para o projeto de sistema de controle tolerante. De acordo com as três etapas: detecção, identificação e diagnóstico, três casos são considerados:

1. O algoritmo de diagnóstico é capaz de fornecer uma estimativa $\hat{R}_f(\hat{\kappa}_f), \hat{U}_f$ do impacto da falha. Então, o problema a ser resolvido é o problema de controle padrão $(O, \hat{R}_f(\hat{\kappa}_f), \hat{U}_f)$.
2. O algoritmo de diagnóstico de falha fornece uma estimativa $\hat{\Gamma}_f(\hat{\kappa}_f), \hat{U}_f$ do impacto da falha, em que $\hat{\Gamma}_f$ é conjunto de possíveis restrições e $\hat{\kappa}_f$ é o conjunto de parâmetros associados. Quando existe uma solução, o sistema real com falha satisfaz os objetivos O dado as restrições reais $\hat{R}_f(\hat{\kappa}_f) \in \hat{\Gamma}_f(\hat{\kappa}_f)$.
3. O algoritmo de diagnóstico de falha detecta e isola falhas, mas não pode fornecer qualquer estimativa do impacto delas. Assim, deve-se projetar o controle de um sistema completamente desconhecido, ou seja, uma tarefa impossível.

3.1.5 Estratégias de controle tolerante ativo

Nas estratégias de controle tolerante, a acomodação de falhas está associada com os casos 1 (resolver o problema $(O, \hat{R}_f(\hat{\kappa}_f), \hat{U}_f)$) e 2 (resolver o problema $(O, \hat{\Gamma}_f(\hat{\kappa}_f), \hat{U}_f)$).

Estes estão associados com o controle do sistema com falha. A situação de falha pode ser acomodada com respeito aos objetivos O quando o problema possui solução.

O termo acomodação de falha é utilizada em situações em que se deve resolver o problema de controle $(O, \hat{R}_f(\hat{\kappa}_f), \hat{U}_f)$, em que $\hat{R}_f(\hat{\kappa}_f)$ é uma estimativa das restrições reais dada pelo algoritmo de diagnóstico de falha. O objetivo da acomodação de falha é manter o controle do processo na presença de falhas, assim como sistema nominal.

Diferentemente, o termo reconfiguração é utilizado quando é necessário encontrar um novo conjunto de restrições $\hat{R}_f(\hat{\kappa}_f)$ tal que o problema de controle $(O, R_f(\kappa_f), U_f)$ possua uma solução e implementar esta nova solução.

A escolha de um novo conjunto de restrições faz com que as relações entrada-saída entre o controlador e processo se altere. Isso ocorre devido ao fato de que em um esquema de reconfiguração é conveniente em alguns casos utilizar apenas a parte “saudável” do sistema para o controle global do processo. Além disso, é importante ressaltar que as estratégias de acomodação e reconfiguração podem precisar de novas leis de controle para compensar o efeito de falhas.

3.2 Controle Preditivo Tolerante a Falhas

Baseando-se nas informações das Subseções 1.1.3 a 1.1.6, propõe-se utilizar o controlador preditivo baseado em modelos lineares e sucessivamente linearizados apresentados no Apêndice B para estudar o comportamento desse tipo de controlador na presença de falhas.

A Figura (3.2) mostra um procedimento proposto para obtenção de uma estrutura de controle preditivo tolerante a falhas.

De acordo com a Figura (3.2), a transformação do controlador preditivo em um controlador tolerante a falhas é feita neste trabalho através da união de um esquema de monitoramento convencional com um sistema de atualização (utilização do conhecimento disponível), capaz de quantificar o efeito de uma falha através de técnicas de monitoramento e traduzir essas informações em termos de atualização do modelo do processo, das restrições e dos objetivos (*set-points*, por exemplo). Dessa forma, torna-se possível obter ações de controle que utilizem apenas a parte saudável do sistema, ou seja, um esquema de controle tolerante ativo é obtido. A implementação prática desse tipo de sistema é demonstrada no Capítulo 4, para um processo de neutralização.

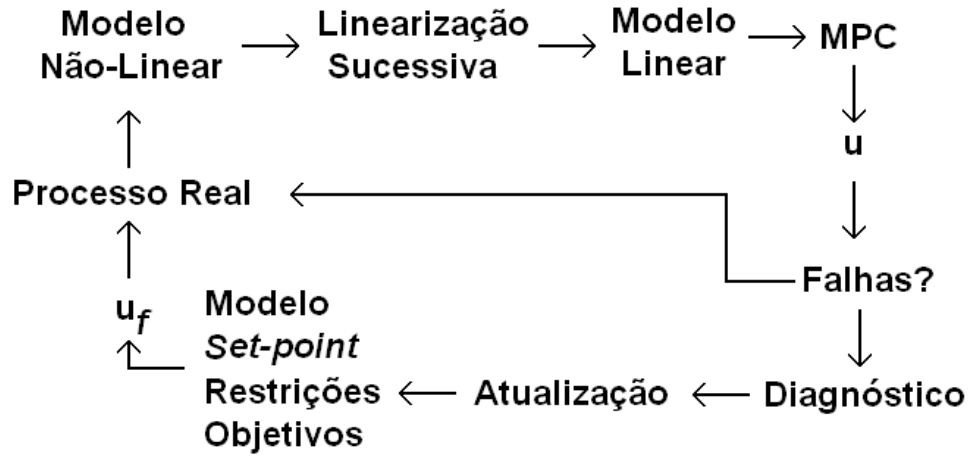


Figura 3.2: Proposta de um sistema de controle preditivo tolerante a falhas.

3.3 Controle Ótimo Tolerante a Falhas

O problema de controle linear quadrático (LQ) constitui um projeto de controle muito popular. Nesta seção, o problema LQ é analisado com respeito à possível ocorrência de falhas em sensores e atuadores. Mostra-se que a tolerância a falhas somente pode ser utilizada se o problema possui solução. A tolerância de falhas em atuadores é dada por estratégias de reconfiguração e acomodação.

Considere o sistema cujo comportamento nominal é modelado por

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \quad (3.1)$$

em que $\mathbf{x} \in \mathbf{X} \subset \mathbb{R}^n$ é o vetor dos estados e $\mathbf{u} \in \mathbf{U} \subset \mathbb{R}^m$ é o vetor contendo ações de controle. Supõe-se que o par $\mathbf{A}, \mathbf{B}$ é controlável.

O problema de controle ótimo é definido com respeito aos objetivos e restrições. Assim:

Objetivo: O : transferir o sistema de $\mathbf{x}_0 = \gamma$ até $\mathbf{x}_\infty = 0$, em que $\gamma \in \mathbb{R}^n$, e $\mathbf{x}_\infty$ é o limite para $\lim_{k \rightarrow \infty} \mathbf{x}_k$ enquanto minimiza a função objetivo

$$\min_{\mathbf{u}} \sum_{k=0}^{\infty} (\mathbf{u}_k^T \mathbf{R} \mathbf{u}_k + \mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k), \quad (3.2)$$

em que $\mathbf{R}$ e $\mathbf{Q}$ são matrizes simétricas, positiva definida e positiva semi-definida, respectivamente.

Restrições: R : a Equação (3.1) é satisfeita para $\forall k \in [0, \infty)$, $\mathbf{x}_k$ e $\mathbf{u}_k$ são funções discretas do tempo.

3.3.1 Controle ótimo regulador

A solução do problema linear quadrático infinito é transformado num problema de otimização quadrático finito

$$J = \frac{1}{2} \mathbf{x}_N^T \mathbf{S} \mathbf{x}_N + \frac{1}{2} \sum_{k=i}^{N-1} (\mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k),$$

em que $\mathbf{S}$ é uma matriz simétrica e positiva semi-definida.

A solução do problema de controle ótimo é bem conhecida da teoria de controle ótimo (LEWIS; SYRMOS, 1995). Seja a função Hamiltoniana do sistema:

$$H = \frac{1}{2} (\mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k) + \boldsymbol{\lambda}_{k+1}^T (\mathbf{A} \mathbf{x}_k + \mathbf{B} \mathbf{u}_k), \quad (3.3)$$

então, as condições de otimalidade são:

$$\begin{aligned} \mathbf{x}_{k+1} &= \frac{\partial H}{\partial \boldsymbol{\lambda}_{k+1}} = \mathbf{A} \mathbf{x}_k + \mathbf{B} \mathbf{u}_k, \\ \boldsymbol{\lambda}_k &= \frac{\partial H}{\partial \mathbf{x}_k} = \mathbf{Q} \mathbf{x}_k + \mathbf{A}^T \boldsymbol{\lambda}_{k+1}. \end{aligned} \quad (3.4)$$

A condição de estacionariedade define as equações de estado e co-estado:

$$\mathbf{0} = \frac{\partial H}{\partial \mathbf{u}_k} = \mathbf{R} \mathbf{u}_k + \mathbf{B}^T \boldsymbol{\lambda}_{k+1},$$

ou

$$\mathbf{u}_k = -\mathbf{R}^{-1} \mathbf{B}^T \boldsymbol{\lambda}_{k+1}. \quad (3.5)$$

Substituindo (3.5) em (3.1) tem-se:

$$\mathbf{x}_{k+1} = \mathbf{A} \mathbf{x}_k - \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \boldsymbol{\lambda}_{k+1}. \quad (3.6)$$

Para obter a solução do problema de otimização é necessário resolver as Equações (3.4) e (3.6) simultaneamente. Note que para as equações do sistema (3.1), a condição inicial, $\mathbf{x}_0$ é conhecida, enquanto que para a equação dos multiplicadores de Lagrange (3.4), a condição final $\boldsymbol{\lambda}_N$ é especificada. Assim o problema torna-se um problema de valor de contorno em dois pontos.

O controle ótimo em estado estacionário é utilizado neste trabalho para o desenvolvimento das técnicas de controle tolerante. Assim, a seqüência de controle ótimo é

determinada se é possível encontrar a sequência de co-estados, λ_k . Os dois casos mais comuns consideram que $\mathbf{Q} = \mathbf{0}$ e $\mathbf{S} = \mathbf{0}$ resolvem o problema de controle ótimo por duas abordagens: uma com o final livre e outra com o final fixo. A solução existe se o par $(\mathbf{A}, \mathbf{B})$ é controlável. Para discussão detalhada da resolução de problemas de controle ótimo ver Lewis e Syrmos (1995).

3.3.2 Controle ótimo servo

O desenvolvimento do controle ótimo servo utiliza o mesmo modelo discreto no espaço de estado apresentado anteriormente

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k,$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k,$$

com o controle servo dado por

$$\mathbf{u}_k = -\mathbf{K}\mathbf{x}_k + \mathbf{K}\mathbf{I}\mathbf{v}_k,$$

$$\mathbf{v}_k = \mathbf{v}_{k-1} + \mathbf{r}_k - \mathbf{y}_k.$$

Como

$$\mathbf{v}_{k+1} = \mathbf{v}_k + \mathbf{r}_{k+1} - \mathbf{C}(\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k),$$

$$\mathbf{v}_{k+1} = \mathbf{v}_k + \mathbf{r}_{k+1} - \mathbf{y}_{k+1},$$

$$\mathbf{v}_{k+1} = -\mathbf{C}\mathbf{A}\mathbf{x}_k + \mathbf{v}_k - \mathbf{C}\mathbf{B}\mathbf{u}_k + \mathbf{r}_{k+1},$$

pode-se escrever

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{v}_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{CB} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{v}_k \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ -\mathbf{CB} \end{bmatrix} \mathbf{u}_k + \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix} \mathbf{r}_{k+1}.$$

Supondo que $\mathbf{r}$ é constante:

$$\mathbf{r}_k = \mathbf{r}_{k+1} = \mathbf{r},$$

então, quando k que tende ao infinito,

$$\begin{bmatrix} \mathbf{x}_\infty \\ \mathbf{v}_\infty \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{CB} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_\infty \\ \mathbf{v}_\infty \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ -\mathbf{CB} \end{bmatrix} \mathbf{u}_\infty + \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix} \mathbf{r}_\infty.$$

Com a definição de

$$\mathbf{x}_{ek} = \mathbf{x}_k - \mathbf{x}_\infty,$$

$$\mathbf{v}_{ek} = \mathbf{v}_k - \mathbf{v}_\infty.$$

Então a equação do erro torna-se:

$$\begin{bmatrix} \mathbf{x}_{e(k+1)} \\ \mathbf{v}_{e(k+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{CB} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{ek} \\ \mathbf{v}_{ek} \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ -\mathbf{CB} \end{bmatrix} \mathbf{u}_{ek}.$$

Note que

$$\mathbf{u}_{ek} = -\mathbf{K}\mathbf{x}_{ek} + \mathbf{K}\mathbf{I}\mathbf{v}_{ek} = -[\mathbf{K} \quad -\mathbf{K}\mathbf{I}] \begin{bmatrix} \mathbf{x}_{ek} \\ \mathbf{v}_{ek} \end{bmatrix}.$$

Defini-se:

$$\hat{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{CB} & \mathbf{1} \end{bmatrix}, \quad \hat{\mathbf{B}} = \begin{bmatrix} \mathbf{B} \\ -\mathbf{CB} \end{bmatrix}, \quad \hat{\mathbf{K}} = [\mathbf{K} \quad -\mathbf{K}\mathbf{I}], \quad \mathbf{w}_k = \mathbf{u}_{ek}, \quad \boldsymbol{\xi}_k = \begin{bmatrix} \mathbf{x}_{ek} \\ \mathbf{v}_{ek} \end{bmatrix}.$$

Dessa forma, tem-se:

$$\boldsymbol{\xi}_{k+1} = \hat{\mathbf{A}}\boldsymbol{\xi}_k + \hat{\mathbf{B}}\mathbf{w}_k \quad (3.7)$$

com o controle:

$$\mathbf{w}_k = -\hat{\mathbf{K}}\boldsymbol{\xi}_k. \quad (3.8)$$

O problema é determinar a matriz $\hat{\mathbf{K}}$ tal que o seguinte índice de desempenho quadrático seja minimizado:

$$J = \frac{1}{2} \sum_{k=0}^{\infty} (\boldsymbol{\xi}^T \mathbf{Q} \boldsymbol{\xi} + \mathbf{w}^T \mathbf{R} \mathbf{w}),$$

em que $\mathbf{Q}$ e $\mathbf{R}$ são escolhidos adequadamente tal que a resposta dinâmica do sistema possua características desejadas.

As técnicas de tolerância a falhas apresentadas para o controlador ótimo regulador são facilmente estendidas para o controle ótimo servo utilizando as Equações (3.7) e (3.8).

A solução do problema de controle ótimo servo utilizada é a solução em estado estacionário proposta por Ogata (1995). Assim, para resolver a equação de Riccati em estado estacionário:

$$\mathbf{P} = \mathbf{Q} + \mathbf{A}^T \mathbf{P} \mathbf{A} - \mathbf{A}^T \mathbf{P} \mathbf{B} (\mathbf{R} + \mathbf{B}^T \mathbf{P} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{P} \mathbf{A},$$

inicia-se com a seguinte equação de Riccati em estado não estacionário

$$\mathbf{P}_{k+1} = \mathbf{Q} + \mathbf{A}^T \mathbf{P}_k \mathbf{A} - \mathbf{A}^T \mathbf{P}_k \mathbf{B} (\mathbf{R} + \mathbf{B}^T \mathbf{P}_k \mathbf{B})^{-1} \mathbf{B}^T \mathbf{P}_k \mathbf{A},$$

com a solução $\mathbf{P}_0 = \mathbf{0}$ e realiza-se iterações até uma solução satisfatória ser obtida.

Assim, o controlador e compensador são encontrados diretamente por:

$$\hat{\mathbf{K}} = (\mathbf{R} + \hat{\mathbf{B}}^T \mathbf{P} \hat{\mathbf{B}})^{-1} \hat{\mathbf{B}}^T \mathbf{P} \hat{\mathbf{A}}.$$

O desenvolvimento do controle ótimo regulador tolerante a falhas para sistemas contínuos é apresentado por Blanke et al. (2006). Neste trabalho, formaliza-se o desenvolvimento do controle ótimo tolerante a falhas para sistemas discretos. Explorações são feitas no Capítulo 4.

Tolerância a falhas em atuadores

A situação em que o sistema opera sem falhas até o instante t_f e que após este, uma falha em um ou vários atuadores ocorre, é abordada nesta subseção. Após a ocorrência da falha existirá atuadores que operam com falhas (válvulas emperradas, por exemplos) e atuadores que operam normalmente. Assim, após o tempo t_f tem-se

$$\boldsymbol{\xi}_{k+1} = \hat{\mathbf{A}} \boldsymbol{\xi}_k + \hat{\mathbf{B}}_f \mathbf{w}_k \quad (3.9)$$

em que $\hat{\mathbf{B}}_f$ representa a nova matriz $\hat{\mathbf{B}}$, estimada diretamente do processo ou de um modelo. O problema de controle ótimo servo tolerante a falhas em atuadores é o de continuar operando o processo na presença de atuadores com problemas ou até mesmo substituí-los quando necessário. A identificação dos atuadores com problemas é comumente realizada por algoritmos específicos. Uma alternativa proposta neste trabalho é a realização de linearizações a cada instante de amostragem do modelo não linear e obter um modelo linear atualizado que representam o comportamento atual do processo. Com isso é possível obter novos valores de alvo para o problema. Uma outra forma é a verificação entre o sinal calculado pelo controlador e o sinal medido por sensores, que geram resíduos e são capazes de detectar falhas. No caso de perda de atuadores, colunas inteiras da matriz $\mathbf{B}$ tornam-se nulas.

3.3.3 Reconfiguração do controle

Em muitos casos, a reconfiguração é vista como uma substituição da parte do processo que apresenta falha. Para falhas em atuadores, por exemplo, significa que alguns atuadores não estavam em uso antes da ocorrência da falha e que podem ser postos em operação após a falha.

A utilização da estratégia de reconfiguração usando controle ótimo deve ser feita de tal forma que as relações entre os sinais de entrada e saída não se alterem. Por exemplo, um processo em que existem duas correntes de entrada e uma corrente redundante para suprir alguma necessidade não atendida devido a falhas em quaisquer das duas correntes faz com que durante a operação normal do processo exista uma coluna com elementos nulos na matriz $\mathbf{B}$, que representa a entrada que admite manipulação mas não é utilizada durante operação normal. Quando uma falha ocorre em uma das duas correntes principais é possível então não utilizar a entrada que apresenta problemas, reajustando alguns parâmetros do controlador relacionados às restrições e ativar o uso da manipulada redundante.

A Figura (3.3) mostra a idéia principal dos métodos de reconfiguração. Em vez de adaptar o controlador à planta com falha, um bloco de reconfiguração é usado para adaptar a planta com falha ao controlador nominal. A planta com falha juntamente com o bloco de reconfiguração deve produzir, para uma dada entrada u_c , a mesma ou aproximadamente a mesma saída y_c da planta nominal. Assim, o controlador “enxerga” a mesma planta como antes da falha, reagindo assim, da mesma forma.

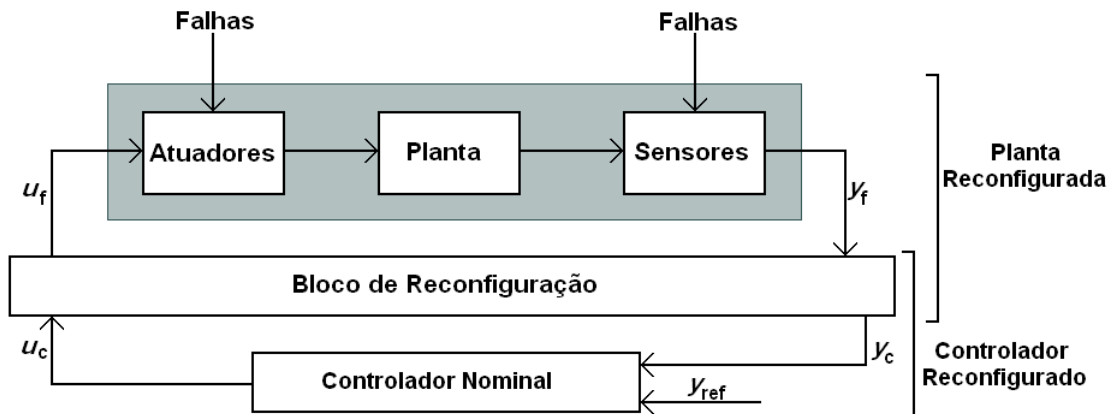


Figura 3.3: Princípio da reconfiguração do controle.

Este tipo de solução tenta aplicar uma mudança mínima ao laço de controle. Em particular, o controlador nominal permanece como um bloco inalterado do laço de controle. A razão para manter o controlador nominal é que a lei de controle existente inclui conhecimentos implícitos valiosos sobre o processo e um possível desempenho do sistema

em malha fechada. Este conhecimento é adquirido durante o ciclo de projeto e não é representado pelo modelo do processo. Por exemplo, durante o projeto, torna-se óbvio quais objetivos de controle (pico máximo, tempo de assentamento, *etc.*) podem ser compatibilizados com os esforços de controle e quais não, e isto é representado pelo controlador nominal.

3.4 Controle Tolerante com *Model-Matching*

Seja um controlador com realimentação das saídas

$$\mathbf{u}_k = \mathbf{K}(\mathbf{y}_k, \mathbf{y}_k^{stp}),$$

usado para atenuar perturbações e garantir o funcionamento da estrutura responsável para que o valor alvo $\mathbf{y}_k^{stp}$ seja alcançado. A lei de controle $\mathbf{K}$ é projetada tal que o sistema em malha fechada satisfaça algumas condições de desempenho para a planta sem falhas. Antes de uma falha f ocorrer, o nível de supervisão mostrado na Figura (3.1) não está ativo.

De acordo com a Figura (3.1), o nível de supervisão é responsável pela adaptação do controlador com o sistema com falha. Assim, dessa adaptação tem-se como resultado a obtenção de novos parâmetros para o controlador e uma nova configuração de controle. Se os sensores e atuadores trabalham de forma diferente depois da ocorrência da falha, mas a planta com falha continua observável e controlável, a configuração de controle pode permanecer a mesma, mas os parâmetros do controlador devem ser ajustados para o sistema com falha. Este processo é chamado de acomodação de falha. Diferentemente do controle adaptativo clássico, que é projetado para ser capaz de lidar com modificações internas do processo, a utilização da reconfiguração em sistemas de controle tolerante a falhas incide na tomada de decisões baseada em sistemas de monitoramento independentes.

Instrumentos que quebram o laço de controle precisam ser substituídos. Assim, o controlador deve ser configurado no sentido de repetir todo procedimento necessário para a escolha de parâmetros do controlador, bem como a configuração de controle adequada devem ser feitas após a falha estar presente no sistema. Não sendo suficiente alterar alguns parâmetros do controlador, o problema de controle deve partir do zero pela escolha adequada das variáveis controladas e manipuladas, da lei de controle $\mathbf{K}$ e do *set-point* $\mathbf{y}_k^{stp}$.

No contexto deste trabalho, a reconfiguração de controle pode ser entendida como uma reparação analítica do laço fechado do sistema, em que o controlador é alterado para explorar medidas redundantes ou sinais de controle que satisfaçam as especificações de

controle na presença de falhas (Figura (3.4)).

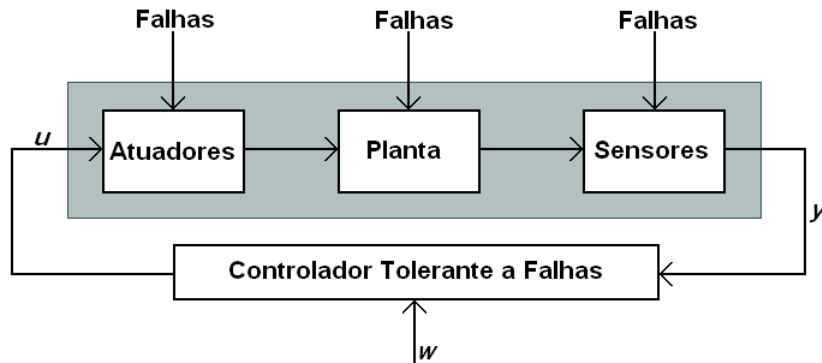


Figura 3.4: Controlador tolerante a falhas por reconfiguração (BLANKE et al., 2006).

Para resolver o problema de controle tolerante a falhas supõe-se que o modelo discreto no espaço de estado, assim como em (3.9), está disponível e também descreve as dinâmicas do processo com respeito a falhas, sendo estas previamente identificadas por algoritmos de detecção e diagnóstico.

De forma geral, o problema de controle tolerante a falhas pode ser resumido assim:

Problema 3.2 (O problema de controle tolerante)

Dados:

O modelo da planta (3.1).

O controlador nominal $\mathbf{K}$.

Especificações de controle.

Falha f .

Encontrar:

Configuração de controle e nova lei de controle $\mathbf{K}_f$.

O método descrito posteriormente é geral e inclui acomodação e reconfiguração de falhas. Dependendo dos sensores e atuadores utilizados, o controlador é simplesmente adaptado às novas dinâmicas da planta ou usa novos sensores e atuadores. Quando a falha ocorre no controlador, linhas inteiras da matriz $\mathbf{K}$ tornam-se nulas e devem ser substituídas por elementos não nulos, ou seja, explorando atuadores ainda inutilizados para obter uma nova configuração para o controlador.

3.4.1 O método da pseudo-inversa

Um dos primeiros métodos para a reprojeção de controladores se baseia no *model-matching* ou *model following* (KAUFMAN et al., 1998). Utiliza-se o conhecimento disponível do controlado nominal, tais como modelo e lei de controle, para descrever as propriedades dinâmicas que um novo controlador deverá produzir quando é possível traduzir o efeito de falhas para o modelo do processo.

O laço fechado nominal do sistema é composto pela planta linear nominal

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \quad (3.10)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k, \quad (3.11)$$

e o controlador $\mathbf{u}_k = -\mathbf{K}\mathbf{x}_k$, que fornece o laço fechado

$$\mathbf{x}_{k+1} = (\mathbf{A} - \mathbf{B}\mathbf{K})\mathbf{x}_k,$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k.$$

Se o controlador não usa todas entradas do vetor de entradas $\mathbf{u}_k$, a matriz $\mathbf{K}$ possui linhas inteiras nulas. Quando uma falha f ocorre, a planta com falha é dada por

$$\mathbf{x}_{k+1} = \mathbf{A}_f\mathbf{x}_k + \mathbf{B}_f\mathbf{u}_k, \quad (3.12)$$

$$\mathbf{y}_k = \mathbf{C}_f\mathbf{x}_k, \quad (3.13)$$

em que a falha f altera as propriedades do sistema, que são agora descritas pelas matrizes $\mathbf{A}_f$, $\mathbf{B}_f$ e $\mathbf{C}_f$. Se o conjunto de entradas disponíveis ou sinais de saída se altera, linhas ou colunas das matrizes $\mathbf{B}_f$ e $\mathbf{C}_f$ desaparecem, respectivamente. Um novo controlador

$$\mathbf{u}_k = -\mathbf{K}_f\mathbf{x}_k,$$

deve ser encontrado tal que o laço fechado do sistema

$$\mathbf{x}_{k+1} = (\mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f)\mathbf{x}_k,$$

$$\mathbf{y}_k = \mathbf{C}_f\mathbf{x}_k,$$

se comporte como o laço nominal. Isto é, a relação

$$\mathbf{A} - \mathbf{B}\mathbf{K} = \mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f, \quad (3.14)$$

deve ser válida, significando que ambos laços fechados possuem dinâmicas similares.

A Equação (3.14) não pode ser satisfeita ao menos que $\mathbf{B}$ e $\mathbf{B}_f$ possuam a mesma imagem, (como no caso de um atuador redundante). Além disso, o novo controlador $\mathbf{K}_f$ é escolhido para minimizar a diferença

$$\|(\mathbf{A} - \mathbf{BK}) - (\mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f)\|. \quad (3.15)$$

A solução deste problema é dada por

$$\mathbf{K}_f = (\mathbf{B}_f^T \mathbf{B}_f)^{-1} \mathbf{B}_f^T (\mathbf{A}_f - \mathbf{A} + \mathbf{BK}), \quad (3.16)$$

em que $\mathbf{B}_f^+$ denota a pseudo-inversa de $\mathbf{B}_f$.

O novo controlador (Equação (3.16)) é adaptado para o sistema com falha e minimiza a diferença dada pela Equação (3.15) entre as propriedades dinâmicas do laço nominal e do laço fechado do sistema com a planta com falha. Entretanto, isto não garante a estabilidade do laço fechado do sistema. Além disso, a estabilidade de $\mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f$ deve ser testada separadamente.

3.4.2 Projeto *model-matching* para falhas em sensores

O desenvolvimento da técnica de *model-matching* para falhas em sensores e atuadores é feita para o sistema discreto dado pela Equação (3.1). A aplicação para um processo com controle ótimo servo é diretamente estendida a partir das análises apresentadas nesta subseção devido à semelhança na construção do controlador.

Esta seção considera o caso de falhas completas em sensores. Se o i -ésimo sensor falha, a saída $\mathbf{y}_i$ é igual a zero. Tal fato é ocasionado pela alteração da matriz $\mathbf{C}$ para $\mathbf{C}_f$, cuja i -ésima linha é igual a zero. O problema de reconfiguração correspondente será investigado com um controlador com realimentação das saídas

$$\mathbf{u}_k = -\mathbf{K}\mathbf{y}_k,$$

logo, o laço fechado do sistema é descrito por

$$\mathbf{x}_{k+1} = (\mathbf{A} - \mathbf{BKC})\mathbf{x}_k,$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k.$$

Para a planta com falha, o novo controlador

$$\mathbf{u}_k = -\mathbf{K}_f \mathbf{y}_{fk},$$

deve ser encontrado tal que o laço fechado

$$\mathbf{x}_{k+1} = (\mathbf{A} - \mathbf{B}\mathbf{K}_f\mathbf{C}_f)\mathbf{x}_k,$$

$$\mathbf{y}_{fk} = \mathbf{C}_f\mathbf{x}_k,$$

possua as mesmas dinâmicas do laço nominal.

O controlador deve satisfazer a versão simplificada da Equação (3.14)

$$\mathbf{K}_f\mathbf{C}_f = \mathbf{K}\mathbf{C}.$$

Para encontrar uma matriz adequada $\mathbf{K}_f$ é possível somente se a condição

$$\text{Kern}(\mathbf{C}_f) \subseteq \text{Kern}(\mathbf{C}), \quad (3.17)$$

é satisfeita, em que Kern denota o kernel (núcleo) da matriz (conjunto de vetores $\mathbf{x}_k$ que fazem $\mathbf{C}\mathbf{x}_k = \mathbf{0}$). A condição significa que a medida de informação obtida pelo vetor de saída completo $\mathbf{y}_k$ é o mesmo que a informação obtida pelos sensores remanescentes através de $\mathbf{y}_{fk}$. A condição dada pela Equação (3.17) pode ser escrita na forma equivalente como

$$\text{posto } \mathbf{C}_f = \begin{pmatrix} \mathbf{C} \\ \mathbf{C}_f \end{pmatrix}.$$

Lema 3.1: *No caso de falhas em sensores, model-matching exato pode ser alcançado se a relação dada pela Equação (3.17) é válida. Então, o controlador*

$$\mathbf{u}_k = -\mathbf{K}\mathbf{P}\mathbf{y}_k, \quad (3.18)$$

resolve o problema de reconfiguração, em que

$$\mathbf{P} = \mathbf{C}\mathbf{C}'_f(\mathbf{C}_f\mathbf{C}'_f)^{-1}, \quad (3.19)$$

satisfaz a relação

$$\mathbf{C} = \mathbf{P}\mathbf{C}_f. \quad (3.20)$$

O controlador reconfigurado $\mathbf{K}_f = \mathbf{K}\mathbf{P}$ produz um sistema em malha fechada que possui exatamente as mesmas propriedades do sistema em malha fechada sem falhas.

Existem algumas situações em que a exigência imposta pela Equação (3.17) é satisfeita:

- Quando a falha muda a sensibilidade do sensor, mas o sinal não é completamente perdido, ou seja, $\mathbf{y}_f = a\mathbf{y}$ vale para algum escalar a .
- Quando um sensor que está com falha possui no mínimo um sensor paralelo redundante.
- Quando uma relação analítica entre a saída com falha e vários outros valores de saída existe, que podem ser reformulados usando a matriz $\mathbf{P}$.

Os dois últimos casos somente são possíveis se $\mathbf{C}$ não possui posto completo.

3.4.3 Projeto *model-matching* para falhas em atuadores

Para falhas em atuadores, a matriz $\mathbf{B}$ é substituída pela matriz $\mathbf{B}_f$ com uma coluna nula para o atuador com falha. O controlador com realimentação das saídas é

$$\mathbf{u}_k = -\mathbf{K}\mathbf{y}_k,$$

que leva ao sistema em malha fechada

$$\mathbf{x}_{k+1} = (\mathbf{A} - \mathbf{B}_f\mathbf{K}\mathbf{C})\mathbf{x}_k,$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k,$$

que deve ser substituído por um novo controlador

$$\mathbf{u}_k = -\mathbf{K}_f\mathbf{y}_{fk},$$

tal que o laço fechado

$$\mathbf{x}_{k+1} = (\mathbf{A} - \mathbf{B}_f\mathbf{K}_f\mathbf{C})\mathbf{x}_k,$$

$$\mathbf{y}_{fk} = \mathbf{C}\mathbf{x}_k,$$

possua as mesmas dinâmicas do laço nominal.

O controlador deve satisfazer a versão simplificada da Equação (3.14)

$$\mathbf{B}_f\mathbf{K}_f = \mathbf{B}\mathbf{K}. \quad (3.21)$$

A solução $\mathbf{K}_f$ desta equação existe somente se a condição

$$\text{Im}(\mathbf{B}_f) \supseteq \text{Im}(\mathbf{B}) \quad (3.22)$$

é válida, e Im denota a imagem de uma matriz (conjunto de vetores $\mathbf{y}_k$ para que um vetor $\mathbf{x}_k$ exista tal que $\mathbf{y}_k = \mathbf{C}\mathbf{x}_k$).

$$\text{posto}(\mathbf{B}_f) = \text{posto}(\mathbf{B}\mathbf{B}_f).$$

Lema 3.2: *No caso de falhas em atuadores, o model-matching exato pode ser alcançado se a Equação (3.22) é válida. Então, o controlador reconfigurado é dado por*

$$\mathbf{u}_k = -\mathbf{N}\mathbf{K}\mathbf{y}_k \quad (3.23)$$

em que

$$\mathbf{N} = (\mathbf{B}_f' \mathbf{B}_f)^{-1} \mathbf{B}_f' \mathbf{B} \quad (3.24)$$

é uma matriz que satisfaz a relação

$$\mathbf{B}_f \mathbf{N} = \mathbf{B} \quad (3.25)$$

O novo controlador $\mathbf{K}_f = \mathbf{N}\mathbf{K}$ gera um sistema em malha fechada com exatamente as mesmas propriedades que a malha fechada nominal.

3.5 Controle PID Tolerante a Falhas

Como a grande maioria dos processos são controlados por controladores do tipo PID, a implementação de estratégias de controle tolerante a falhas para estes é importante e explorada neste trabalho.

Os controladores do tipo PID implementados não sofrem qualquer tipo de reconfiguração. Duas alternativas são abordadas: os controladores trabalham sempre no caso nominal e busca-se a reconfiguração do processo que possibilite o seu uso ou existe um banco de controladores que altera o controle do processo baseado em uma lógica de supervisão. A Figura (3.5) mostra a alteração lógica entre um banco de controladores previamente sintonizados para acomodar efeitos de falhas.

A estrutura proposta pela Figura (3.5) é implementada em uma coluna de destilação binária controlada por quatro estruturas de controle para acomodar falhas em atuadores (Capítulo 4). Uma outra forma de construir controladores PID tolerante a falhas é

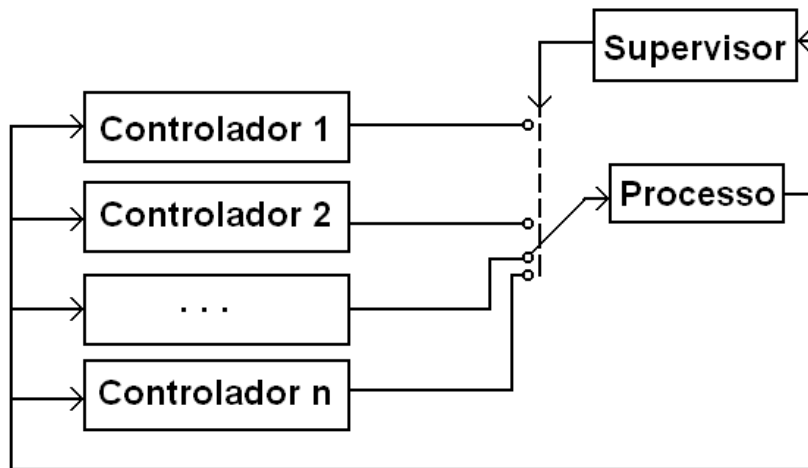


Figura 3.5: Estrutura de alteração lógica entre um banco de controladores.

a implementação de estratégias de monitoramento e sistemas de supervisão capazes de reconfigurar o processo. Essa idéia é apresentada pela Figura (3.3).

3.6 Análise Estrutural e o Controle Tolerante a Falhas

Um sistema é dito ser tolerante a falhas se seus objetivos podem ser alcançados na presença de falhas. O controle ou estimação de um subconjunto de variáveis são exemplos de objetivos de um processo, sendo que apenas serão alcançados com sucesso se pelo menos dois procedimentos podem ser alcançados com sucesso: acomodação ou reconfiguração. Este está ligado à troca de componentes com falha para buscar os objetivos projetados e aquele significa que deve-se estimar variáveis de interesse no sistema com falha.

A análise estrutural do sistema em operação normal e com falha é realizada neste trabalho para de mostrar na forma de grafo ou matriz de incidência quais restrições podem ser utilizadas para estimar um subconjunto de variáveis do sistema, definindo assim diferentes modos de operação. As falhas são inseridas no modelo estrutural pela exclusão ou não utilização de uma ou mais restrições, ou seja, são falhas consideradas estruturais pois alteram a natureza das relações entre variáveis e restrições, resultando assim em grafos modificados. Como exemplo de falhas estruturais tem-se a perda de variáveis medidas em um processo, que excluem completamente as restrições associadas com as respectivas medições. Assim, o objetivo torna-se estimar as variáveis perdidas, se possível, pelo subconjunto de restrições restantes. As falhas não estruturais são aquelas que interferem na estrutura do sistema, ou seja, falhas que não resultam na exclusão ou alteração de qualquer restrição, por exemplo, a alteração de parâmetros do modelo.

3.6.1 Tolerância a falhas estruturais

Falhas estruturais alteram a estrutura do grafo. As mudanças consideradas são: mudança do estado de variáveis de conhecida para desconhecida ou vice-versa, suprimindo alguns nós de restrições e de variáveis. Dois casos são apresentados de acordo com a informação fornecida pelo algoritmo de diagnóstico de falha.

- Quando o modelo do sistema com falha é conhecido, verifica-se as propriedades de observabilidade e controlabilidade estrutural para confirmar a existência de uma solução para o problema de acomodação.
- Quando o modelo do sistema com falha não é conhecido (as mudanças nas restrições podem ser somente detectadas e isoladas), a acomodação não é possível, e somente a reconfiguração do sistema pode ser usada para fornecer quais componentes podem ser retirados ou colocados em operação.

A questão principal para tolerância a falhas devido a alterações na estrutura de um grafo é se mesmo após a remoção de alguma restrição ou a troca de estado de alguma variável (de conhecida para desconhecida, por exemplo) ainda é possível obter um grafo com um pareamento completo com relação às variáveis. Além disso, deve-se verificar se estas alterações nas estruturas dos grafos podem ser feitas em linha e estarem prontamente disponíveis pelo sistema de diagnóstico. Se isto é possível, pode-se decidir reconfigurar todo o sistema ou estimar variáveis.

3.6.2 Tolerância a falhas não estruturais

Pela definição, falhas não estruturais não alteram o grafo estrutural do sistema. Assim, nenhum problema de análise estrutural específico precisa ser formalizado, pois as propriedades de observabilidade e controlabilidade são as mesmas durante operação normal e com falha.

Quando o modelo da falha do sistema com falha é conhecido (indicando que o sistema de diagnóstico também forneceu uma nova estimativa das novas restrições e parâmetros do sistema), o problema de acomodação pode ser realizado, se existe uma solução.

CAPÍTULO 4

Resultados e Discussões

Este capítulo utiliza as estruturas teóricas propostas nos Capítulo 2 e 3 para a obtenção de controle tolerante a falhas, demonstrando através de exemplos a efetividade das mesmas. O primeiro exemplo utiliza dois sistemas descritos por matrizes de funções de transferência para testar algumas técnicas de controle tolerante apresentadas no Capítulo 3, além disso, utilizam-se duas estratégias de controle diferentes para abordar os mesmos aspectos de falhas. O segundo exemplo mostra como uma coluna de destilação binária controlada por quatro estruturas de controle pode adquirir tolerância a falhas quando implementa-se uma estratégia de monitoramento por FDA (ver Capítulo 2) para detectar e diagnosticar falhas e alternar entre as estruturas de controle para acomodar efeitos de falhas. O terceiro exemplo trata-se de um reator de produção de ciclopentanol controlado por duas malhas PI e monitorado por DPCA e com características de tolerância a falhas com base em análise estrutural. O quarto utiliza um processo de neutralização com elevado grau de não linearidade e controlado por um controlador preditivo baseado em modelo sucessivamente linearizado para acomodar efeitos de falha em atuadores.

4.1 Controle Tolerante para Sistemas Discretos

4.1.1 Introdução

O objetivo desta subseção é o de demonstrar como o controle ótimo servo e o controle preditivo podem adquirir características de tolerância a falhas em sensores e atuadores. A vantagem da utilização deste tipo de controlador é que a sua estrutura multivariável intrínseca possibilita uma fácil alteração de *set-points* e restrições do problema de otimização para atender as mudanças ocasionadas por falhas. Neste exemplo, com o intuito de demonstrar alternativas para controle tolerante baseadas em sistemas discretos na representação no espaço de estado, supõe-se que as falhas inseridas são detectadas e diagnosticadas corretamente com um pequeno atraso.

4.1.2 Exemplo 1: Controle Ótimo Servo

A utilização do controle ótimo servo é feita neste exemplo para um processo de destilação binária para a separação metanol-água. Este processo tem sido alvo de estudo por vários autores em aplicações de estratégias de vários tipos de controladores (WOOD; BERRY, 1973).

O modelo linear do processo é escrito na forma de matriz de funções de transferência como

$$\begin{bmatrix} X_D(s) \\ X_B(s) \end{bmatrix} = \begin{bmatrix} \frac{12,8e^{-s}}{16,7s+1} & \frac{-18,9e^{-3s}}{21s+1} \\ \frac{6,6e^{-7s}}{10,9s+1} & \frac{-19,4e^{-3s}}{14,4s+1} \end{bmatrix} \begin{bmatrix} R_F(s) \\ Q_V(s) \end{bmatrix},$$

em que $X_D(s)$ representa a fração molar de metanol no destilado, $X_B(s)$ a fração molar de metanol no produto de fundo, $R_F(s)$ a quantidade de refluxo e $Q_V(s)$ a taxa de vapor que ascende a coluna em unidades de lb/min e no domínio de Laplace. Todos as unidades que contém a variável tempo estão em minutos. A Tabela (4.1) apresenta os valores de operação em estado estacionário da coluna.

Tabela 4.1: Operação em estado estacionário da coluna para separação metanol-água do Exemplo 1.

Variável	Descrição	Estado Estacionário
x_D	Fração de metanol no produto de topo	0,9625
x_B	Fração de metanol no produto de fundo	0,005
R_F	Vazão de refluxo	1,95 lb/min
Q_V	Vazão de vapor	2,45 lb/min

A aproximação para o tempo morto da função de transferência é feita por Padé e é

utilizada antes da discretização da função de transferência. Assim, utiliza-se

$$e^{-Ts} = \frac{2 - Ts}{2 + Ts}.$$

A representação no espaço de estado com mínima realização e com um tempo de amostragem de 3 minutos (SHRIDHAR; COOPER, 1998), o seguinte sistema no espaço de estado em mínima realização é obtido:

$$\mathbf{A} = \begin{bmatrix} 0,86530 & 0,07281 & -0,1220 & -0,0172 & 0,00963 & -0,0044 & 0,00056 \\ -0,0882 & 0,65473 & 0,08571 & 0,20843 & 0,16503 & -0,0207 & 0,01551 \\ 0,12113 & 0,26688 & 0,19959 & 0,07472 & -0,1751 & -0,0418 & 0,01169 \\ -0,0131 & 0,17173 & -0,1511 & 0,66550 & -0,1139 & 0,18127 & -0,0712 \\ 0,00064 & 0,15741 & 0,09359 & -0,3109 & 0,64264 & 0,00860 & -0,0899 \\ 0,00408 & 0,03361 & 0,03986 & 0,03404 & -0,2091 & -0,0096 & -0,1124 \\ 0,00258 & -0,0097 & -0,0135 & 0,06232 & 0,10351 & -0,1030 & 0,81787 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} -1,68361 & 3,170290 \\ 0,85212 & 1,626430 \\ 1,51635 & -0,93511 \\ -0,5774 & -0,19219 \\ -0,3940 & -0,23162 \\ -0,0215 & -0,06067 \\ 0,03620 & -0,01767 \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} -0,77596 & 0,91060 & -0,58379 & -0,39451 & -0,09413 & 0,85793 & 0,07984 \\ -0,77874 & 0,28316 & -1,42757 & -0,00307 & -0,26058 & -0,2614 & -0,0394 \end{bmatrix}.$$

Com a matriz $\mathbf{D}$ nula, projeta-se um sistema de controle ótimo servo para o sistema:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k,$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k,$$

da seguinte forma:

$$\mathbf{u}_k = -\mathbf{K}\mathbf{x}_k + \mathbf{K}\mathbf{I}\mathbf{v}_k$$

com

$$\mathbf{v}_k = \mathbf{v}_{k-1} + \mathbf{r}_k - \mathbf{y}_k$$

em que $\mathbf{r}_k$ representa a referência ou *set-point* desejado para o sistema. As simulações com este controlador utilizam as seguintes ponderações para os estados e para as entradas

(ver Capítulo 3, Seção 3.3.3):

$$\mathbf{Q} = \text{diag}([0, 001; 0, 001; 0, 001; 0, 001; 0, 001; 0, 001; 0, 001]),$$

e

$$\mathbf{R} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

A simulação para a coluna de destilação binária para separação de metanol da água se inicia em estado estacionário, com $x_D = 0,9625$ e $x_B = 0,005$. Supondo que é desejável um aumento na separação entre os dois componentes, é necessário alterar a vazão de refluxo e de vapor (R_F e Q_V) de tal forma que esse objetivo seja alcançado. A lei de controle ótimo projetada para o sistema composto pelas matrizes $\mathbf{A}$, $\mathbf{B}$ e $\mathbf{C}$ é:

$$\mathbf{K} = \begin{bmatrix} -0,0560 & 0,04160 & 0,01616 & 0,01640 & 0,00534 & -0,00126 & -0,00123 \\ 0,08938 & 0,02948 & -0,0048 & 0,01107 & -0,0028 & -0,00017 & -0,00056 \end{bmatrix},$$

e

$$\mathbf{K}\mathbf{I} = \begin{bmatrix} 0,02714 & -0,00969 \\ -0,00530 & -0,02587 \end{bmatrix}.$$

A resposta dinâmica devido a mudança de *set-point* ($x_D = 0,99$ e $x_B = 0,001$) é mostrada na Figura (4.1).

Falhas em atuadores e controle tolerante

Falhas em atuadores são problemas comuns nos processos industriais, por exemplo, uma válvula com agarramento ou emperrada são problemas que ocorrem com maior probabilidade devido ao desgastes dos equipamentos ao longo do tempo (THORNHILL et al., (2003)).

Apesar de problemas comuns em atividades práticas, a simulação de problemas em atuadores em sistemas discretos pode ser feita através da modificação da matriz dos atuadores $\mathbf{B}$, que é responsável por transmitir para o processo as entradas calculadas pela lei de controle. Quando as informações para os elementos finais de controle não podem ser exatamente transmitidas para o processo, supõe-se que alguma falha ocorre. Assim, um exemplo de falha em atuadores é a modificação de alguns elementos da matriz $\mathbf{B}$, transformando-a em $\mathbf{B}_f$, com o índice f representando falha. Supondo que a válvula de refluxo apresenta alguma característica que impeça que a entrada calculada pela lei de

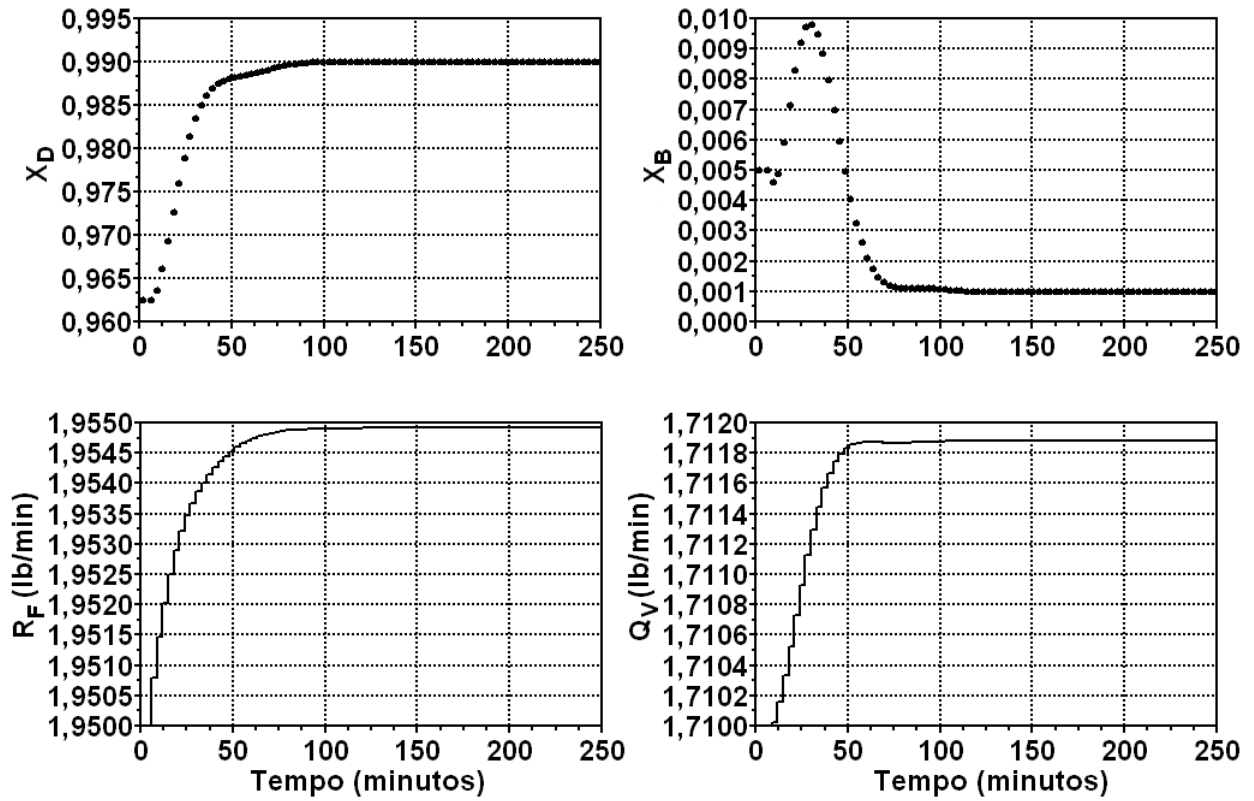


Figura 4.1: Resposta dinâmica do processo de separação de metanol da água devido a uma mudança de *set-point*.

controle possa ser exatamente enviada para o processo, escreve-se:

$$\mathbf{B}_f = \begin{bmatrix} -1,68361 & 3,170290 \\ 0,00000 & 1,626430 \\ 1,00000 & -0,93511 \\ -0,57742 & -0,19219 \\ 0,00000 & -0,23162 \\ 0,00000 & -0,06067 \\ 0,00000 & -0,01767 \end{bmatrix}.$$

A interpretação física desse processo permite avaliar previamente o comportamento quando esta falha é inserida durante uma operação controlada. Neste caso, tem-se uma alteração na corrente de refluxo, em que a válvula de refluxo perde a capacidade de atuação devido aos elementos nulos que aparecem na matriz $\mathbf{B}_f$. Com isso, o controlador tenta suprir esta insuficiência através de ações de controle mais agressivas, aumentando ou diminuindo o valor da variável manipulada.

A simulação do processo de destilação com este tipo de falha é mostrada na Figura (4.2) antes do estabelecimento no estado estacionário (a partir do tempo 36 minutos). A Figura (4.2) mostra que as variáveis manipuladas (R_F e Q_V) são alteradas para que

o *set-point* seja restabelecido após uma alteração no atuador responsável pela vazão de refluxo. O *set-point* ainda pode ser alcançado mesmo com o atuador com problemas, no entanto, sugere-se que lei de controle não adaptada para este tipo de situação seja reformulada utilizando a técnica de *model-matching* proposta no Capítulo 3 para se obter uma resposta dinâmica melhor.

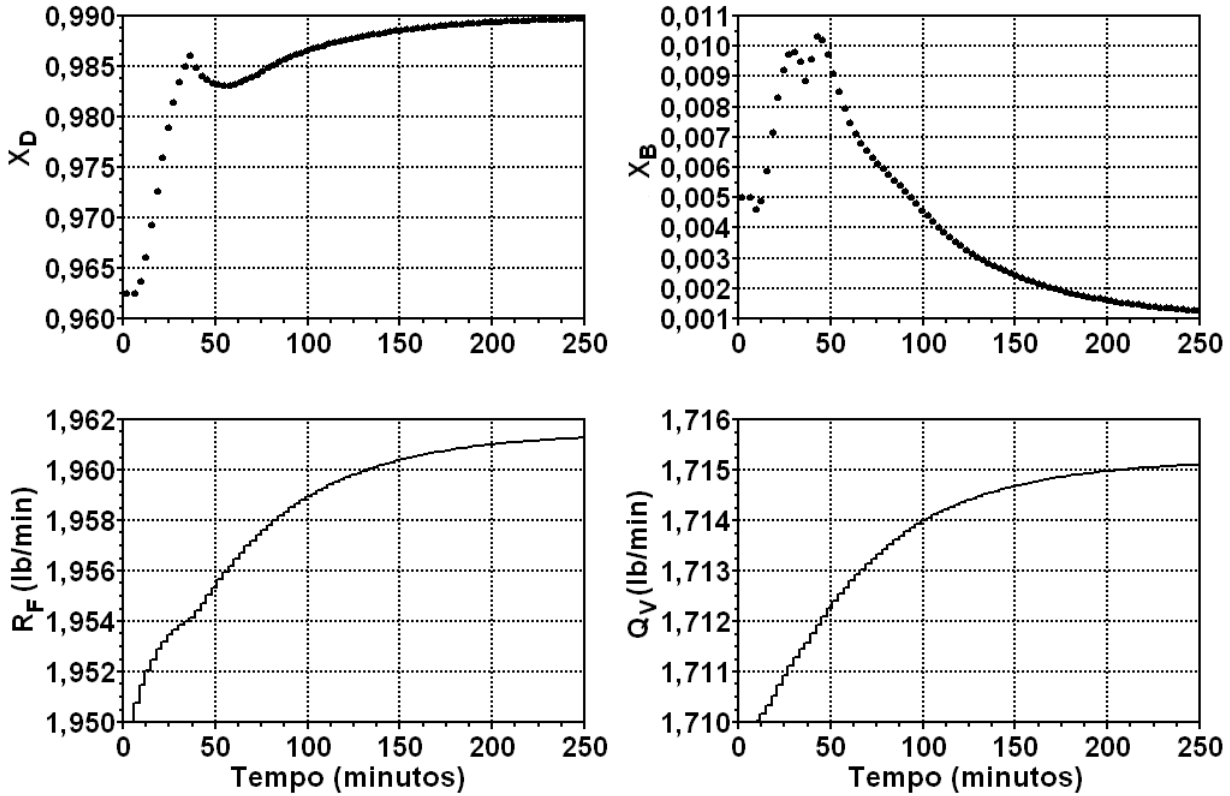


Figura 4.2: Simulação de uma falha no atuador da vazão de refluxo (R_F) para o exemplo da coluna de separação metanol-água.

A técnica de *model-matching* para falhas em atuadores apresentada no Capítulo 3 para o controlador ótimo servo é diretamente aplicada. Assim, ocorre um novo cálculo dos parâmetros do controlador com a planta com falha, possibilitando a obtenção de um novo controlador $\mathbf{K}_f$ que considera o efeito da falha para controlar o processo. Dessa forma, deve-se verificar que as dinâmicas do laço fechado para ambos controladores (planta sem falha e com falha) devem ser semelhantes.

O novo controlador e compensador calculados levando em consideração a alteração na matriz $\mathbf{B}$ para $\mathbf{B}_f$ devido a uma falha utilizando a técnica de *model-matching* é uma estratégia de controle tolerante para este processo. Como uma mudança no processo faz com que a eficiência do controlador seja reduzida devido a oscilações, isto pode ser demonstrado através da elaboração de índices de desempenho para o processo (SOUZA, 2007).

O novo controlador e compensador obtidos pela atualização da matriz $\mathbf{B}$ para $\mathbf{B}_f$ é:

$$\mathbf{K}_f = \begin{bmatrix} -0,09552 & 0,07088 & 0,02754 & 0,02794 & 0,00910 & -0,00215 & -0,00210 \\ 0,06730 & 0,04587 & 0,00150 & 0,01753 & -0,00073 & -0,00067 & -0,00104 \end{bmatrix},$$

e

$$\mathbf{K}\mathbf{I}_f = \begin{bmatrix} 0,04624 & -0,01651 \\ 0,00538 & -0,02969 \end{bmatrix}.$$

Com os parâmetros do controlador atualizados é possível obter uma resposta à falha melhorada, já que estes estão adaptados à alteração sofrida pelo atuador que regula a quantidade de refluxo de produto de topo para a coluna. Assim, busca-se com esse controlador o mesmo comportamento dinâmico do processo nominal. A Figura (4.3) mostra uma melhora sensível da resposta transiente após a inserção da falha quando comparada com o processo sem controle tolerante dado pela Figura (4.2), ou seja, os *set-points* são alcançados mais rapidamente.

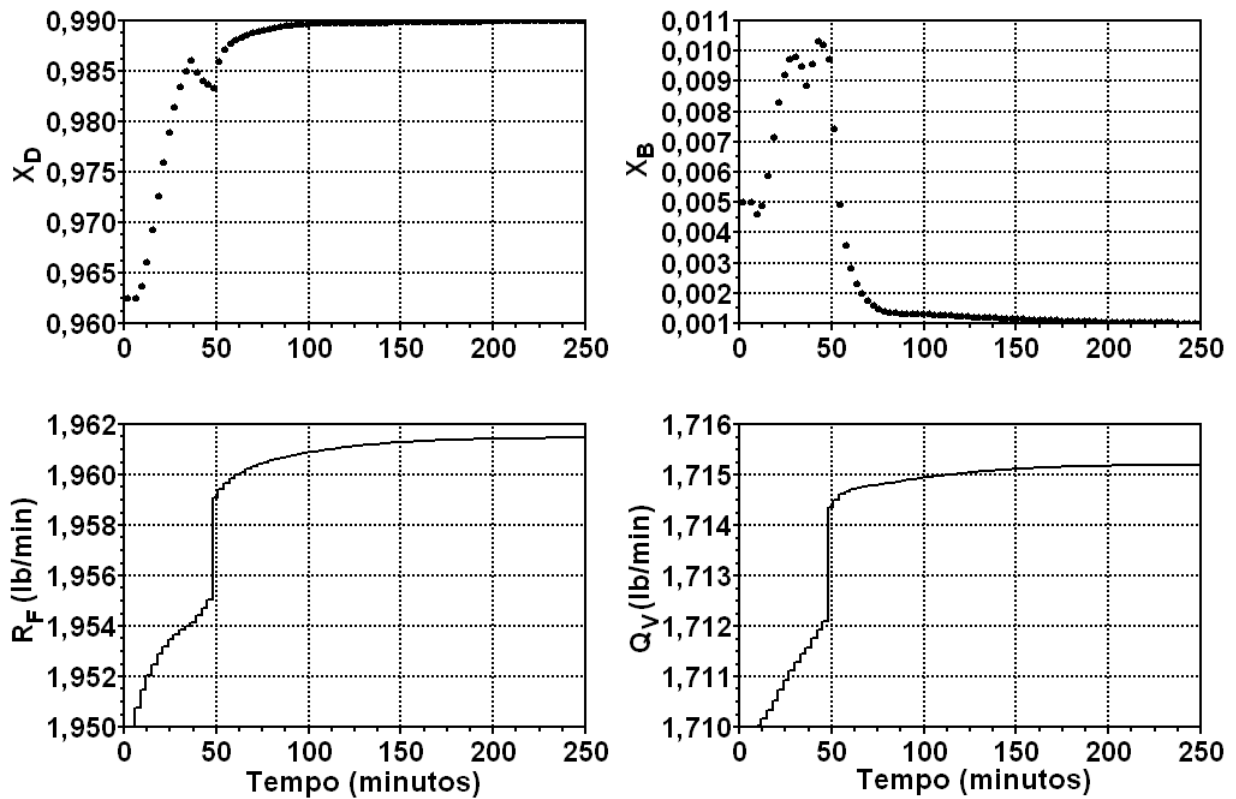


Figura 4.3: Simulação de uma falha no atuador da vazão de refluxo (R_F) com controle tolerante ativo para o exemplo da coluna de separação metanol-água.

Falhas em sensores

Falhas em sensores ocorrem com bastante frequência em processos industriais. Esses instrumentos podem produzir medidas com ruídos ou medidas congeladas, ou seja, a partir de um determinado instante uma medida torna-se fixa e incompatível com a medida real. Durante estas falhas o processo físico real não sofre nenhuma alteração. Ocorre que o controlador atua no processo de tal forma a compensar o erro gerado pela medida incorreta produzida, uma consequência imediata deste tipo de falha é o aumento ou diminuição das variáveis manipuladas até a saturação na tentativa de minimizar o erro entre o *set-point* e o valor medido, que é incorreto.

Com isso, uma possível falha para o processo de destilação é uma medida incorreta de qualquer uma das composições (metanol no produto de topo ou no produto de fundo). Supondo então que a partir do tempo 81 min uma falha ocorre no sensor de concentração de metanol no produto de fundo (congelamento da medida em $x_B = 0,005$, em contraste com o valor de *set-point* no estado estacionário que é $x_B = 0,001$). A resposta gráfica é apresentada na Figura (4.4):

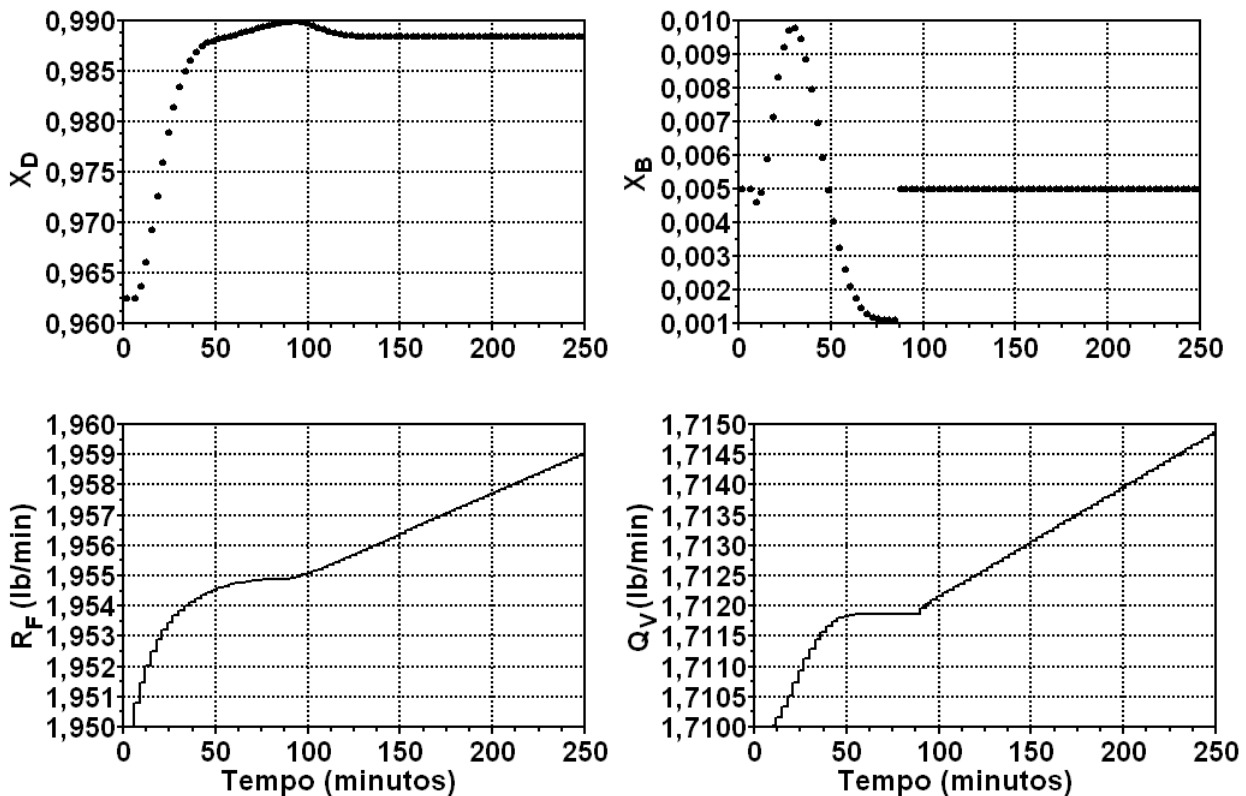


Figura 4.4: Simulação de uma falha no sensor de concentração de metanol no produto de fundo para o exemplo da coluna de separação metanol-água.

A análise da Figura (4.4) mostra que esta falha faz com que o processo se desvie completamente do alvo projetada ($x_D^{sp} = 0,99$ e $x_B^{sp} = 0,001$). Além disso, para compensar o

erro da medida, o controlador compensa este erro aumentando ambas variáveis manipuladas indefinidamente. Tal processo continua até que as saturações sejam alcançadas.

Nesta falha espera-se que a matriz $\mathbf{C}$ se modifique para $\mathbf{C}_f$, de tal forma que se produza sempre uma mesma medida, dado que a saída é uma combinação linear dos estados. Para este exemplo, explora-se como é possível atualizar a matriz $\mathbf{C}$ na tentativa de após a ocorrência deste tipo de falha continuar uma operação controlada de pelo menos uma variável, que neste caso é a parte “saúdável” do sistema, aquela que ainda não foi afetada por falhas. Assim, durante uma falha em sensor, o controle da variável que é medida e apresenta falhas é desligado, e a estratégia de *model-matching* não pode ser utilizada devido ao fato da relação (3.17) não ser válida. Com isso, uma estratégia adequada para este tipo de falha é a atualização da matriz $\mathbf{C}$ com as medidas do processo para detectar falhas. Assim que detectadas deve-se excluir do controle o atuador com defeito através da substituição de uma linha da matriz $\mathbf{C}$ por elementos nulos e em seguida obter um novo controlador com a matriz atualizada. Supondo que a falha ocorra no tempo igual a 93 min e que algoritmos capazes de detectar e identificar este tipo de falha decidem que o sensor de concentração de metanol no produto de fundo não deve mais ser utilizado, atualizando assim a matriz $\mathbf{C}$ para $\mathbf{C}_f$. A partir desse instante, o sistema de controle tolerante recalcula novos controladores e compensadores, $\mathbf{K}$ e $\mathbf{KI}$, respectivamente, do controle ótimo servo, capazes de adequar a esta mudança do processo e manter a parte saudável do sistema controlada.

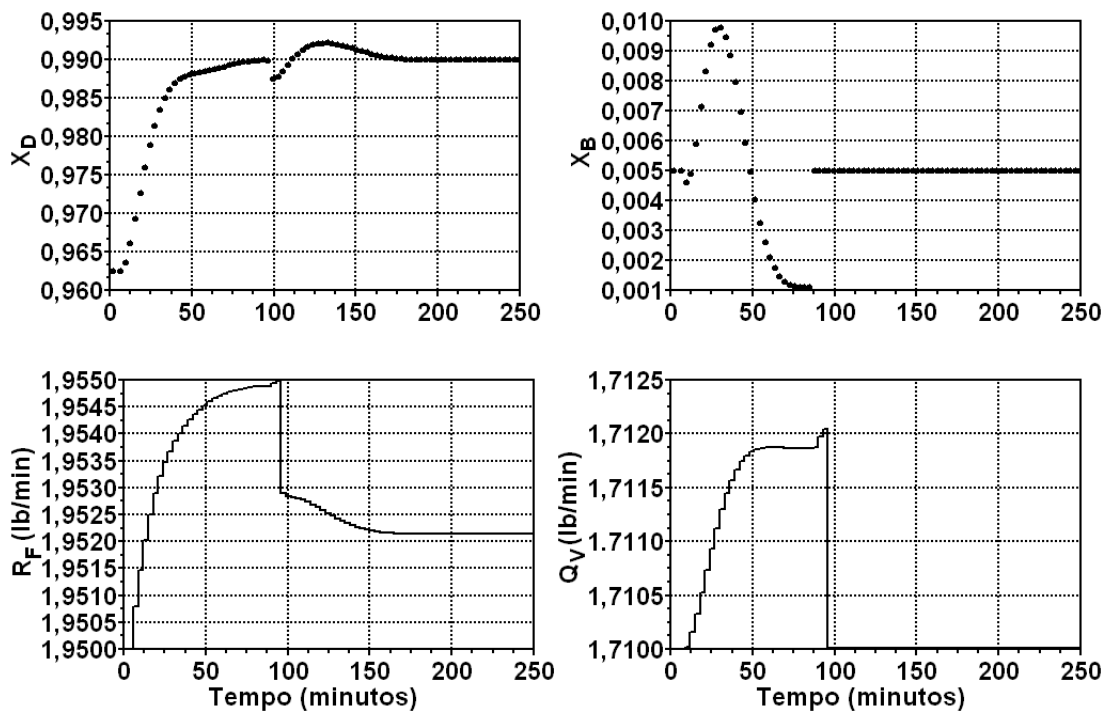


Figura 4.5: Controle tolerante para falha no sensor de concentração de metanol no produto de fundo.

4.1.3 Exemplo 2: Controle Preditivo

A estrutura de controle preditivo encontra-se descrita em detalhes no Apêndice B. Esta ilustração utiliza como exemplo o sistema proposto por Muske e Rawlings (1993):

$$G(z) = \begin{bmatrix} \frac{z}{2s-1} & \frac{z}{2,5z-1,5} \\ \frac{0,5z}{2z-1} & \frac{1,5z}{2,5z-1,5} \end{bmatrix}.$$

Sendo que uma realização no espaço de estado desta matriz de funções de transferência gera

$$\mathbf{A} = \begin{bmatrix} 0,5 & 0 & 0 & 0 \\ 0 & 0,6 & 0 & 0 \\ 0 & 0 & 0,5 & 0 \\ 0 & 0 & 0 & 0,6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0,5 & 0 \\ 0 & 0,4 \\ 0,25 & 0 \\ 0 & 0,6 \end{bmatrix}, \quad \text{e} \quad \mathbf{C} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

A representação no espaço de estados em variável desvio para esse sistema é então:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \quad k = 0, 1, 2, \dots$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k.$$

Supondo que o processo em questão opere em condições normais com as matrizes $\mathbf{A}$, $\mathbf{B}$ e $\mathbf{C}$ e uma mudança no *set-point* é feita para um novo ponto de operação (saídas no estado estacionário $\mathbf{y}_{ee} = [1, -1]^T$). Assim, os estados e a entrada no estado estacionário são:

$$\mathbf{x}_{ee} = \begin{bmatrix} 2,5 \\ -1,5 \\ 1,25 \\ -2,25 \end{bmatrix} \quad \text{e} \quad \mathbf{u}_{ee} = \begin{bmatrix} 2,5 \\ -1,5 \end{bmatrix}.$$

Então, o controlador preditivo apresentado no Apêndice B, Seção B.3, é capaz de controlar o processo quando os seguintes parâmetros são utilizados: $N = 10$; $j_1 = 1$; $j_2 = 5$; $\mathbf{u}_{k-1} = [0, 0]^T$; $\Delta\mathbf{u}_{max} = [1, 1]^T$; $\Delta\mathbf{u}_{min} = [2, 2]^T$; $\mathbf{u}_{max} = [4, 4]^T$; $\mathbf{u}_{min} = [-4, -4]^T$; $\mathbf{y}_{min} = [-2, -2]^T$ e $\mathbf{y}_{max} = [2, 2]^T$.

Utilizando os parâmetros do controlador a seguir:

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}.$$

o resultado mostrado pela Figura (4.6) é obtido:

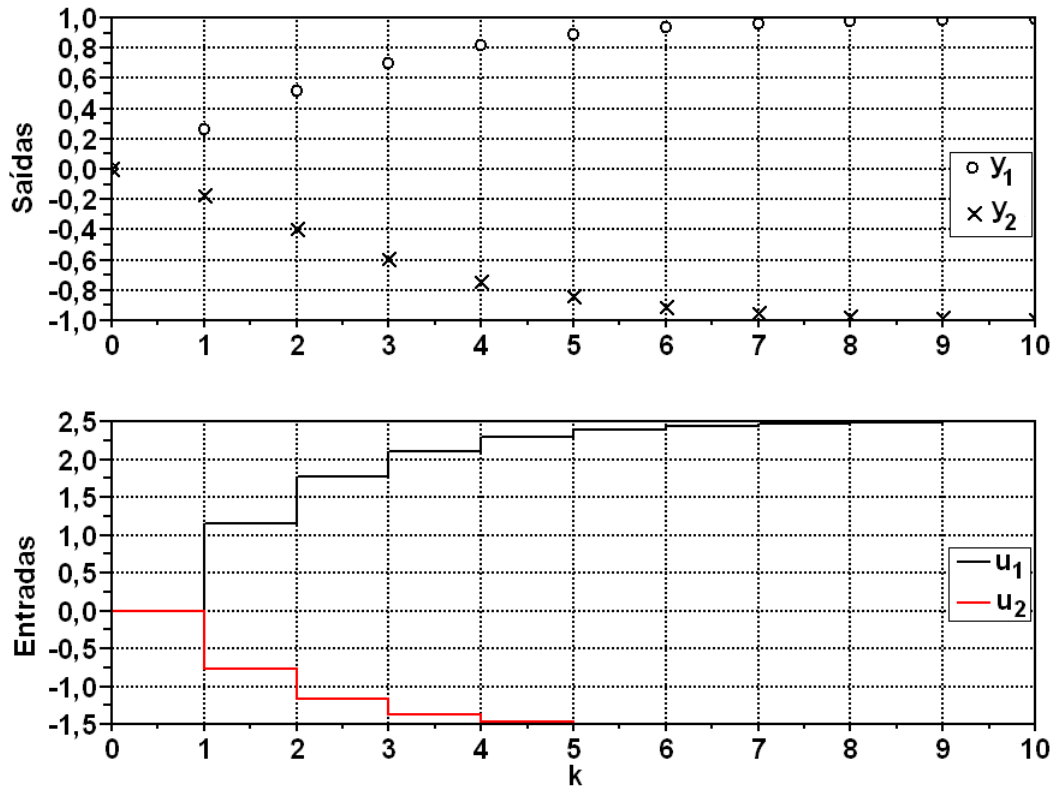


Figura 4.6: Resposta dinâmica do sistema devido a uma mudança de *set-point* do Exemplo 2.

Pode-se notar que o controlador preditivo consegue levar o sistema de um ponto de operação para outro em aproximadamente 8 instantes de amostragem.

Falhas em atuadores

O controlador preditivo não é eficaz em situações em que falhas aparecem. Estas são exploradas neste exemplo são supostamente identificadas e ocorrem na planta física, nos atuadores e nos sensores. Utiliza-se aqui a vantagem da estrutura multivariável do MPC para implementar estratégias de controle tolerante para o controlador MPC nominal. Por exemplo, se para o processo em questão existe um atuador redundante, que deve ser usado apenas em situações de falhas, tem-se com isso uma nova matriz $\bar{\mathbf{B}}$:

$$\bar{\mathbf{B}} = \begin{bmatrix} 0,5 & 0 & 0,3 \\ 0 & 0,4 & 0 \\ 0,25 & 0 & 0 \\ 0 & 0,6 & 0,7 \end{bmatrix}$$

Assim, a comparação entre $\mathbf{B}$ e $\bar{\mathbf{B}}$ mostra que existe uma nova coluna, ou seja, um novo atuador capaz de atuar no processo e manter as mesmas condições operacionais de projeto utilizando todos atuadores ou apenas dois deles. O gráfico mostrado pela Figura (4.6) é o caso principal e utiliza apenas dois atuadores. Com a ampliação do número de atuadores, a seguinte resposta para a mesma referência $\mathbf{y}_{ee} = [1, -1]^T$, mas com novas condições de estado estacionário é obtida para os parâmetros: $N = 10$; $j_1 = 1$; $j_2 = 5$; $\mathbf{u}_{k-1} = [0, 0, 0]^T$; $\Delta \mathbf{u}_{max} = [1, 1, 1]^T$; $\Delta \mathbf{u}_{min} = [2, 2, 2]^T$; $\mathbf{u}_{max} = [4, 4, 4]^T$; $\mathbf{u}_{min} = [-4, -4, -4]^T$; $\mathbf{y}_{min} = [-2, -2]^T$ e $\mathbf{y}_{max} = [2, 2]^T$; $\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{S} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

$$\mathbf{x}_{ee} = \begin{bmatrix} 0,86993 \\ 0,13006 \\ 0,77222 \\ -1,7722 \end{bmatrix} \quad \text{e} \quad \mathbf{u}_{ee} = \begin{bmatrix} 1,54444 \\ 0,13006 \\ -1,1241 \end{bmatrix}.$$

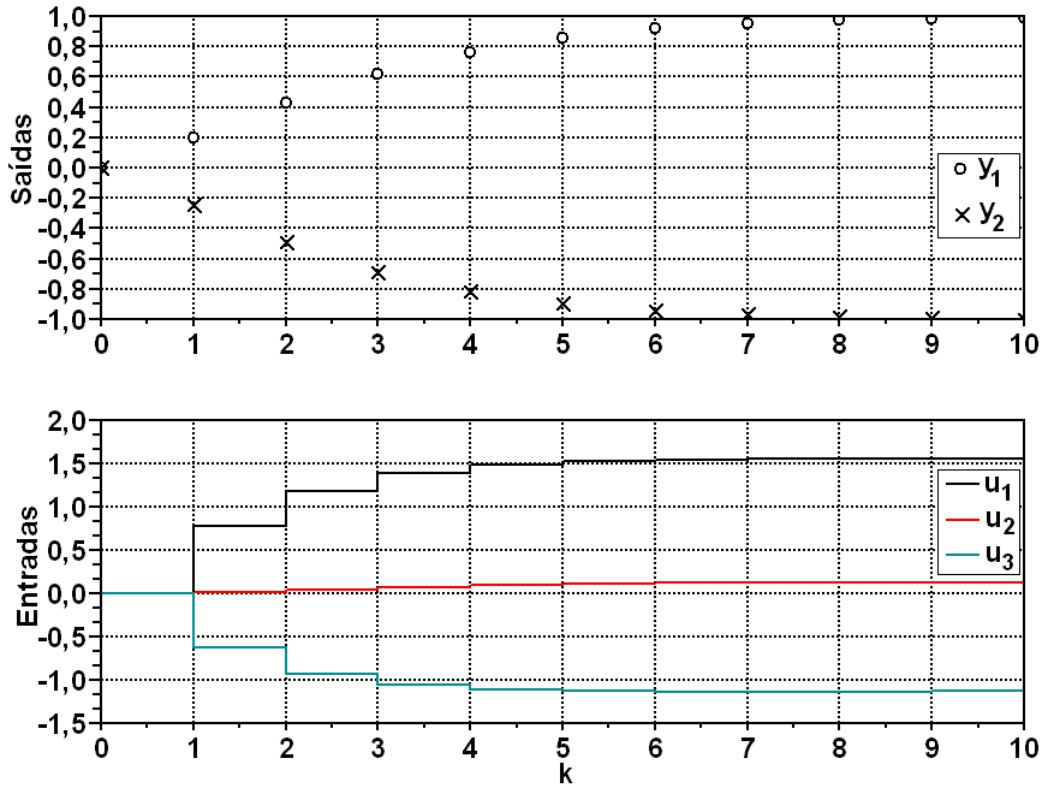


Figura 4.7: Resposta do Exemplo 2 com um atuador extra utilizando MPC.

Nota-se que a informação de inutilização de algum atuador é facilmente atendida pelo MPC, visto que esta configuração é feita nas restrições do processo. As falhas consideradas em atuadores fazem com que colunas inteiras da matriz $\mathbf{B}$ se tornem nulas, ou seja,

atuadores não utilizados ou não mais disponíveis.

Supondo que a atualização da matriz $\mathbf{B}$ é possível para este processo, a estratégia de controle tolerante mais apropriada é a adequação em linha dos *set-points* de operação do projeto, ou seja, objetiva-se sempre o mesmo *set-point* para as variáveis de saída enquanto que modificações no conjunto de atuadores ocorre devido a falhas. Assim, para este processo, mostra-se que uma estratégia de controle tolerante conveniente é a atualização das matrizes que constituem o processo bem como dos *set-points* para o controlador preditivo, visto que qualquer alteração nas matrizes e para um mesmo *set-point* os valores das variáveis de entrada e estados no estado estacionário se alteram e precisam ser atualizados para o controlador. A Figura (4.8) mostra como a perda de um atuador em um esquema sem controle tolerante ativo introduz modificações nas variáveis controladas a partir do instante de amostragem igual a 10.

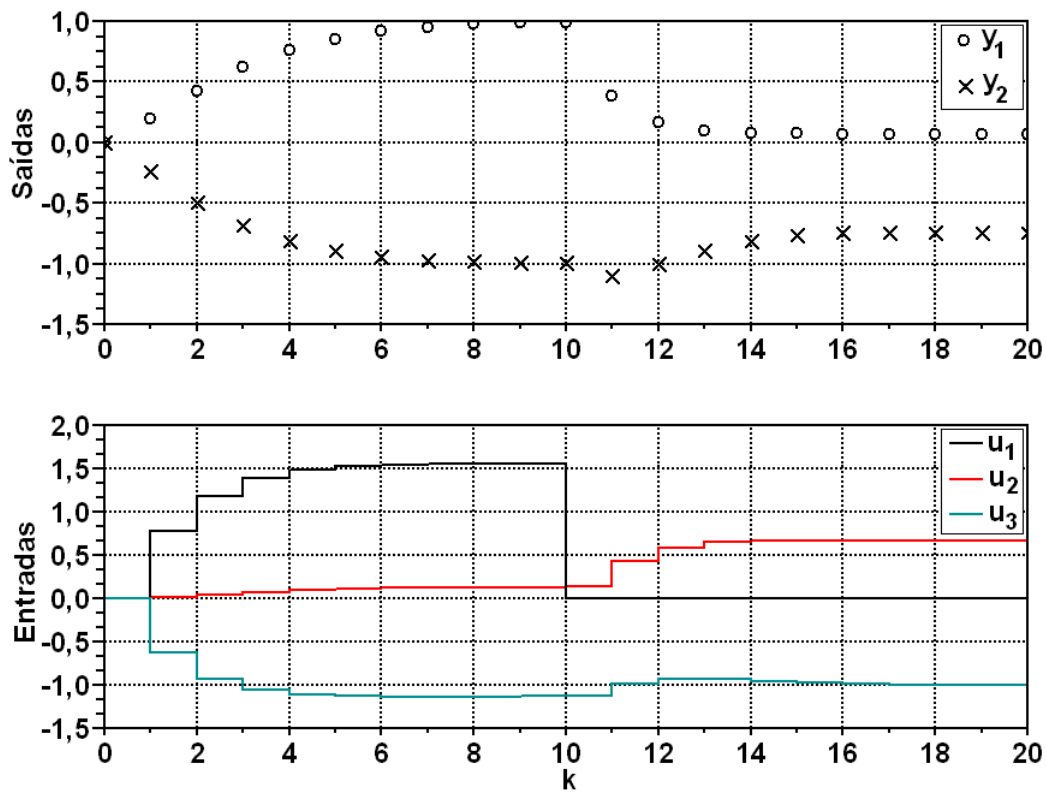


Figura 4.8: Simulação da perda do primeiro atuador (primeira coluna da matriz $\mathbf{B}$) para o Exemplo 2.

Para iniciar algum procedimento de correção desta falha é necessário que as seguintes etapas sejam percorridas: detecção, identificação ou diagnóstico da falha e atualização da matriz dos atuadores, necessária para o projeto do controle preditivo. Sendo realizadas estas etapas, a estratégia proposta é a de substituição da matriz do sistema nominal pela atualizada e com falha no desenvolvimento do controlador, que por sua vez interpreta as colunas nulas da matriz como uma não possibilidade de utilização do atuador referente

àquela coluna e desloca ações de controle para a parte do sistema que ainda é “saudável” (não afetada por falhas). Para efeito de demonstração da técnica proposta, supõe-se que as etapas de detecção, diagnóstico e atualização são realizadas com sucesso e com 2 instantes de amostragem de atraso desde o aparecimento da falha no sistema (instante de amostragem igual a 15). Logo, a Figura (4.9) mostra como o sistema de controle tolerante substitui um atuador perdido na presença de falha. Neste caso, existe a redução de uma entrada no processo, e isto é indiferente para o MPC, que ainda consegue manter o sistema na referência utilizando ainda duas entradas.

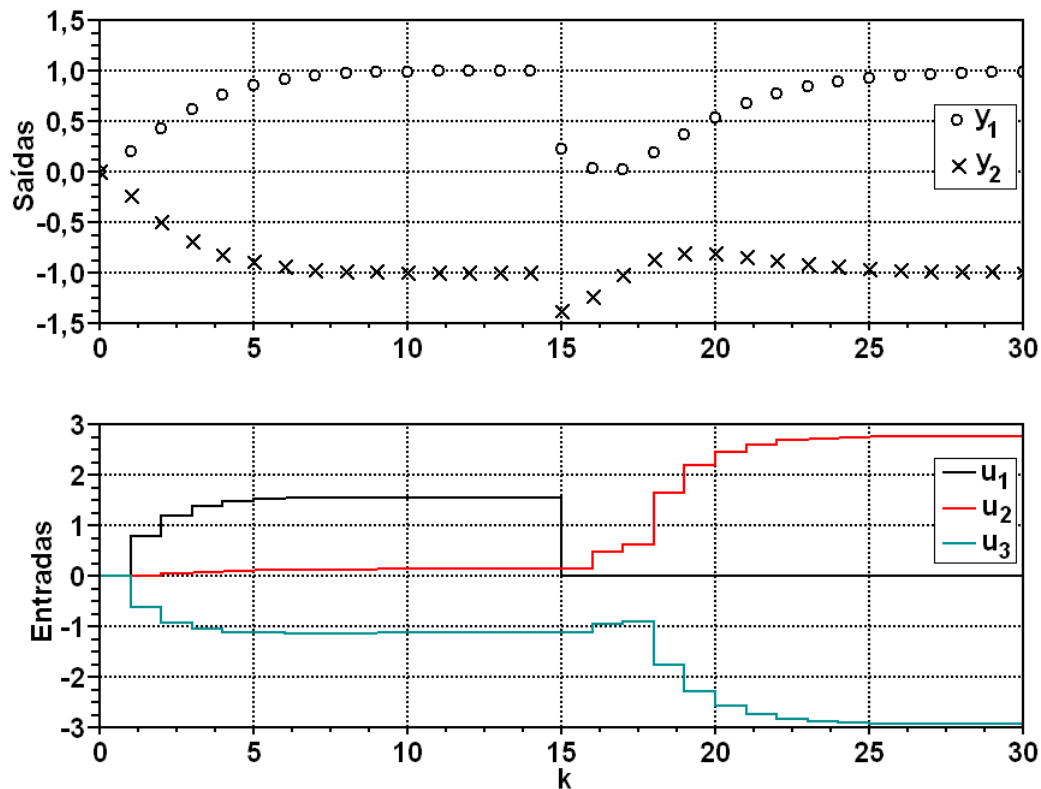


Figura 4.9: Controle tolerante ativo para falha em um atuador do processo do Exemplo 2.

Um outro tipo de falha bastante comum com respeito a atuadores (válvulas) de processos e já comentada anteriormente é o fenômeno conhecido como agarramento. Neste caso o atuador torna-se incapaz de atender as entradas calculadas pelo controlador para comandos de abrir e fechar válvulas. Na prática, como o atuador passa a não atender o controlador, este entende que o processo precisa de ações de controle cada vez mais bruscas até que o atuador é capaz de atender uma delas, ocorrerá então uma perturbação nas variáveis controladas por parte do atuador com agarramento. Assim que este fenômeno é detectado e diagnosticado, utiliza-se a estrutura multivariável do MPC para atualizar a matriz de atuadores para retirar o atuador danificado e continuar operando com os atuadores restantes.

A Figura (4.10) mostra o agarramento no atuador 3 com o respectivo comportamento dinâmico do processo frente a esta alteração.

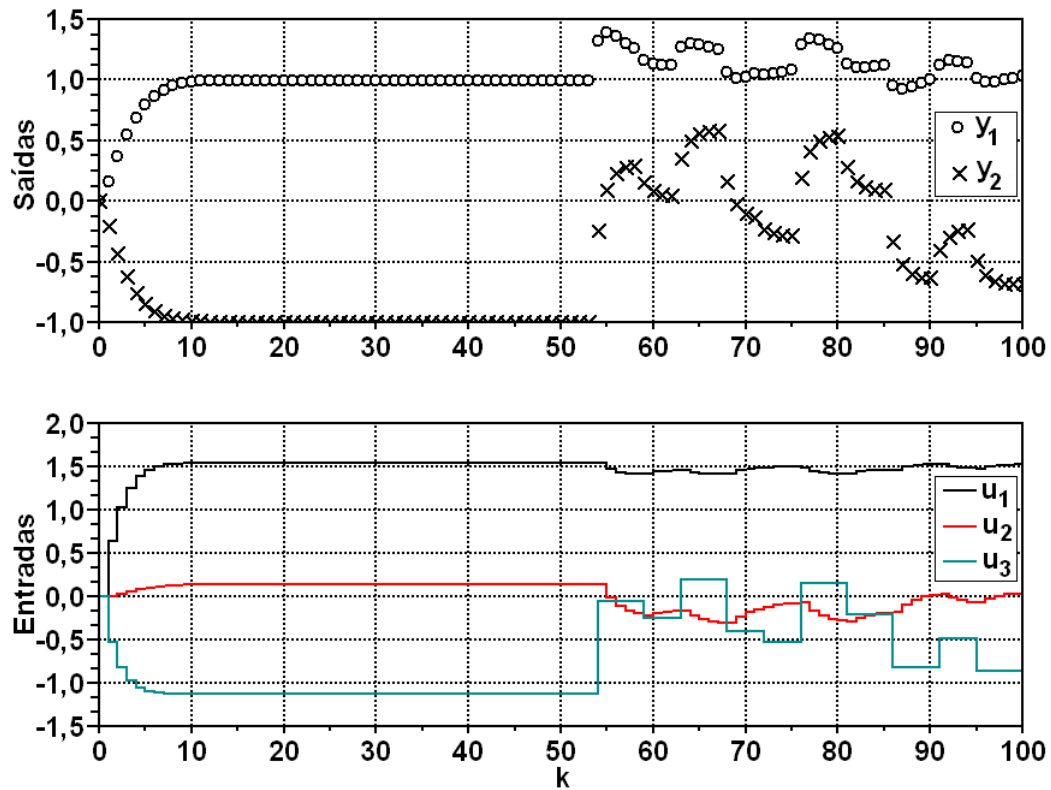


Figura 4.10: Simulação de um atuador com agarramento para o Exemplo 2.

Assim, a não utilização de atuadores com problemas é facilmente programada pelo MPC, já que atualiza-se os pontos estacionários requeridos pelo controlador para que o mesmo *set-point* seja seguido (ver Figura (4.11)).

4.1.4 Discussão dos Resultados

Os dois exemplos apresentados utilizam diferentes controladores para buscar os mesmos objetivos, ou seja, a tolerância a falhas. O controle ótimo servo utiliza as técnicas de *model-matching* propostas no Capítulo 3 com sucesso.

Apesar de não existir nenhuma técnica específica para a acomodação de falhas em processos controlados por controladores preditivos, utiliza-se como alternativa principal a atualização do modelo do processo para o controlador na tentativa de buscar tolerância a falhas.

No segundo exemplo, em que um novo atuador foi inserido no processo deve ser avaliada em termos financeiros para processos reais. No entanto, existem processos que

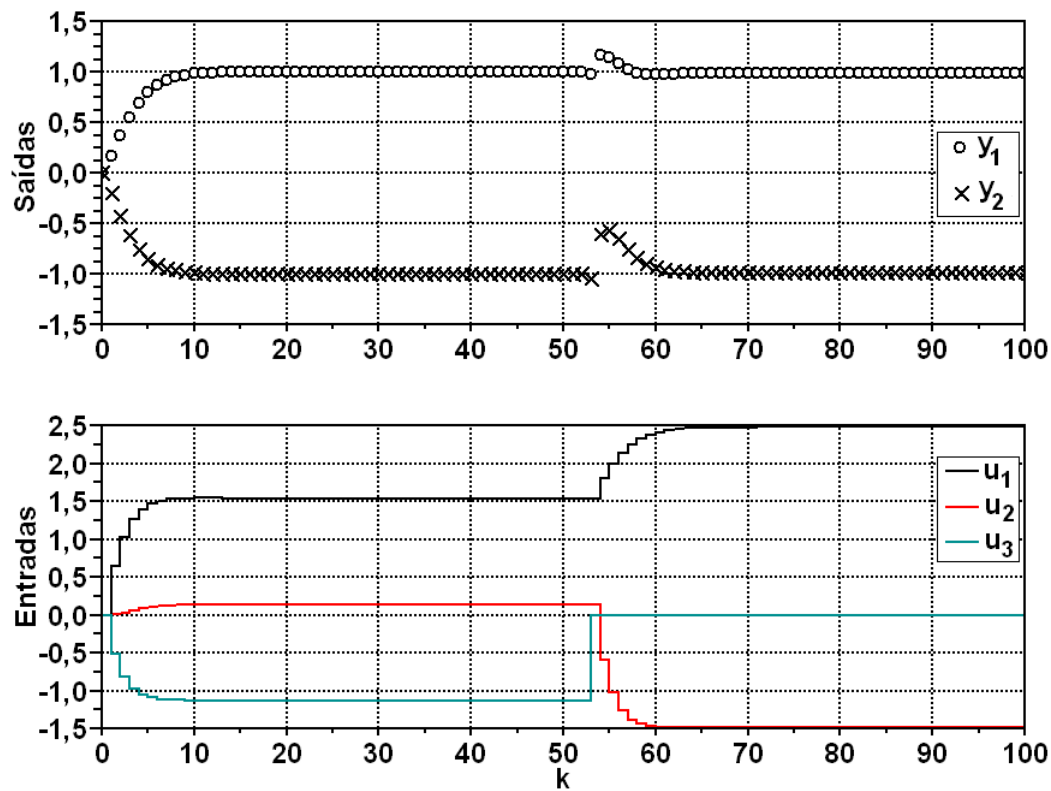


Figura 4.11: Controle tolerante para falha em um atuador com agarramento (Exemplo 2).

possuem mais variáveis de entrada do que saídas, possibilitando assim a implementação desse tipo de estratégia.

4.2 Coluna de Destilação Binária Ideal

Neste exemplo explora-se como uma coluna de destilação pode se tornar tolerante a falhas em atuadores através da utilização de técnicas estatísticas que se baseiam apenas em dados históricos (ver Capítulo 2 para mais informações). Utiliza-se a vantagem das estatísticas multivariáveis para detectar, diagnosticar e projetar um gerenciador de eventos anormais capaz de alternar entre estruturas de controle para manter a operação controlada, assim como proposto pela Figura (3.1) do Capítulo 3, para obtenção de um sistema de controle tolerante a falhas.

4.2.1 Introdução

Destilação é uma técnica de separação comum para correntes de fluidos contendo dois ou mais componentes e é uma das operações unitárias mais importantes da Engenharia Química. Projetar e controlar uma coluna de destilação é importante de forma a produzir correntes com pureza especificada, seja ela para a venda ou para o uso em outros processos.

A destilação se baseia na separação de componentes de uma mistura utilizando a diferença de ponto de ebulição. Para o exemplo estudado neste trabalho a separação é de uma mistura binária de uma corrente líquida. O componente com a menor temperatura de ebulição será chamado de componente leve e o componente com maior temperatura de ebulição de pesado.

Uma mistura líquida saturada de dois componentes a uma dada concentração está em equilíbrio com a fase vapor que possui uma concentração mais elevada do componente mais volátil do que na fase líquida. Utiliza-se x para representar a fração molar do componente leve na fase líquida e y para representar a fração molar do componente leve na fase vapor.

Para misturas ideais, é comum modelar a relação da fase de equilíbrio baseada volatilidade relativa (α) constante

$$y = \frac{\alpha x}{1 + (\alpha - 1)x} \quad (4.1)$$

4.2.2 Modelagem matemática e controle

Descrever matematicamente a operação de destilação, tendo como base os fenômenos de transferência de massa e energia diretamente, é considerada uma tarefa complexa. Tal dificuldade é contornada quando se adota o modelo de estágios de equilíbrio, no qual as correntes de líquido e de vapor que saem de um dado prato se encontram em perfeito equilíbrio entre si. A Figura (4.12) mostra o comportamento das correntes de líquido e

de vapor em um prato de contato e um prato genérico i utilizado para a modelagem da coluna.

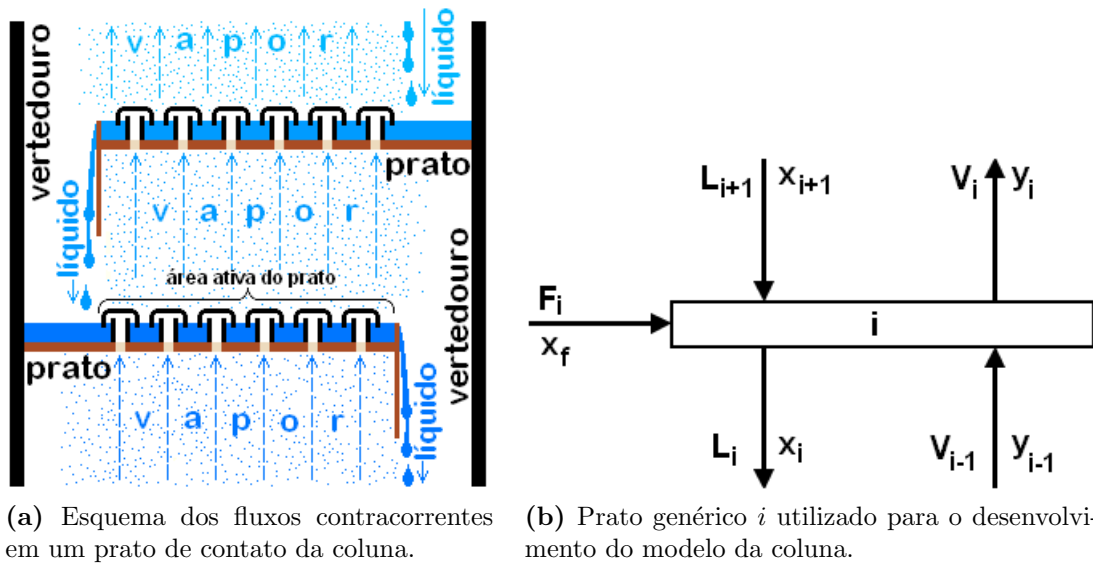


Figura 4.12: Pratos em uma coluna de destilação.

Assim, relações termodinâmicas podem ser usadas para determinar a temperatura e as concentrações das espécies envolvidas nos fluxos a uma dada pressão de operação.

Na Figura (4.13) tem-se um esquema simples de uma coluna de destilação.

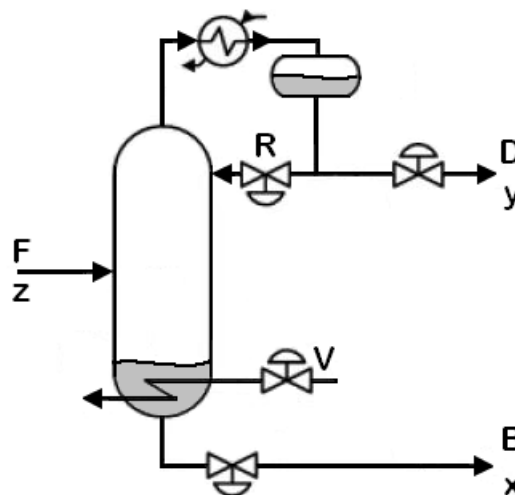


Figura 4.13: Esquema de uma coluna de destilação.

O modelo estudado representa matematicamente a dinâmica de uma coluna de destilação contínua utilizada para separar uma mistura binária ideal. A volatilidade relativa (α) entre os componentes da mistura binária é considerada constante e igual a 2. A corrente molar de vapor é admitida constante ao longo de toda coluna. Líquido saturado é

alimentado na coluna no prato F . A coluna estuda não considera balanços de energia e também não trata temperatura e pressão como variáveis. A numeração dos pratos é feita da base para o topo da coluna. O vapor que deixa o prato de topo é instantaneamente e totalmente condensado e forma um nível constante de líquido no tambor de refluxo, que se encontra perfeitamente misturado.

O refeedor é considerado um estágio de equilíbrio, e o *hold-up* constante de líquido no refeedor encontra-se perfeitamente misturado. Os níveis de líquido dos pratos também possuem mistura perfeita mas podem variar de valor quando as taxas de líquido que entram e saem dos pratos variam.

Utilizando a Equação (4.1) como relação termodinâmica de equilíbrio e com o auxílio da Figura (4.12) temos que as equações dinâmicas de balanço de massa e de equilíbrio de fase para todos os pratos são descritas da seguinte forma:

$$\begin{aligned}\frac{dM_i}{dt} &= F_i + L_{i+1} - L_i \\ M_i(0) &= \bar{M}_i\end{aligned}\quad (4.2)$$

$$\begin{aligned}M_i \frac{dx_i}{dt} &= F_i x_f + L_{i+1} x_{i+1} + V y_{i-1} - L_i x_i - V y_i - x_i \frac{dM_i}{dt} \\ x_i(0) &= \bar{x}_i\end{aligned}\quad (4.3)$$

$$L_i = \bar{L}_i + \frac{M_i - \bar{M}_i}{\beta} \quad (4.4)$$

F_i é diferente de zero apenas quando i é igual ao número que corresponde ao prato de alimentação F . As frações molares x e y se referem ao componente mais leve. As frações do componente mais pesado podem ser obtidas por $(1-x)$ e $(1-y)$. $\bar{L}_i$ e $\bar{M}_i$ correspondem a valores iniciais de estado estacionário (PERRY, 1997), e β é uma constante que pondera a hidráulica do prato. Para a combinação condensador-tambor de refluxo, tem-se que:

$$\begin{aligned}\frac{dx_D}{dt} &= \frac{V y_{N_T} - V x_D}{M_D} \\ x_D(0) &= \bar{x}_D\end{aligned}\quad (4.5)$$

$$V - (D + R) = 0 \quad (4.6)$$

$$\begin{aligned}\frac{dx_B}{dt} &= \frac{L_1 x_1 - V y_B - B x_B}{M_B} \\ x_B(0) &= \bar{x}_B\end{aligned}\quad (4.7)$$

$$L_1 - (B + V) = 0 \quad (4.8)$$

Estruturas de controle típicas

Existem várias alternativas de estruturas utilizadas no controle de colunas de destilação. A nomenclatura padrão de uma estrutura de controle é dada pela junção dos símbolos das duas variáveis manipuladas empregadas para o controle das frações molares de produto topo e de fundo. Desse modo, a estrutura de controle “**R-V**” indica que a vazão de refluxo é manipulada para controlar a composição do produto de topo e que a vazão de vapor ascendente é manipulada para controlar a fração molar do produto de fundo. A estrutura denominada “**D-B**” indica o uso da vazão de destilado e de produto de fundo para o controle das frações de produto de topo e fundo, respectivamente.

O controle simultâneo das frações molares proporciona a separação desejada com um consumo energético minimizado. Contudo, muitas colunas operam apenas com controle de uma das frações molares (produto de topo ou de fundo) e não duas. Isto porque o uso de dois controladores para composição geralmente propicia interação entre as malhas, o que faz com que a sintonia se torne complexa. Ainda, devido às dificuldades com medidores de composição em linha, temperaturas são freqüentemente utilizadas para inferir composições. E em muitas colunas, o controle de temperatura em um prato é suficiente para a manutenção de todo um perfil de temperatura na coluna.

Segue a descrição das estruturas estudadas neste trabalho:

- **R-V**: A corrente de refluxo controla a fração molar do destilado. A vazão de vapor, ou seja, o calor de entrada no refeedor controla a fração molar do produto de fundo. A vazão de destilado D e de produto de fundo B não sofre atuação. Essa estrutura é provavelmente a mais utilizada de todas. Uma de suas vantagens é que, usualmente, a resposta a mudanças de composição da alimentação é muito satisfatória.
- **D-V**: A corrente de destilado controla a fração molar do destilado. A vazão de vapor (o calor de entrada no refeedor) controla a composição do produto de fundo enquanto a vazão de refluxo R e de produto de fundo B não sofrem atuação. Sabe-se que para altas taxas de refluxo ($R/D > 4$), essa estrutura é fortemente recomendada visto que a vazão da corrente de destilado é baixa para controlar o nível do tambor de refluxo.
- **R-B**: A corrente de refluxo controla a fração molar do destilado e a vazão de produto de fundo controla a fração molar do produto de fundo. A vazão de destilado D e o

calor de entrada no refeedor não sofrem atuação. Quando a taxa de vaporização (V/B) é alta, essa estrutura deve ser utilizada, contudo algumas colunas podem apresentar resposta inversa, o que pode ocasionar em consideráveis problemas de controle.

- **D-B:** A corrente de destilado controla a fração molar do destilado. A vazão de produto de fundo controla a fração molar do produto de fundo. A vazão de refluxo R e o calor de entrada no refeedor não sofre atuação.

A Figura (4.14) é um resumo de todas estruturas de controle implementadas nesse trabalho.

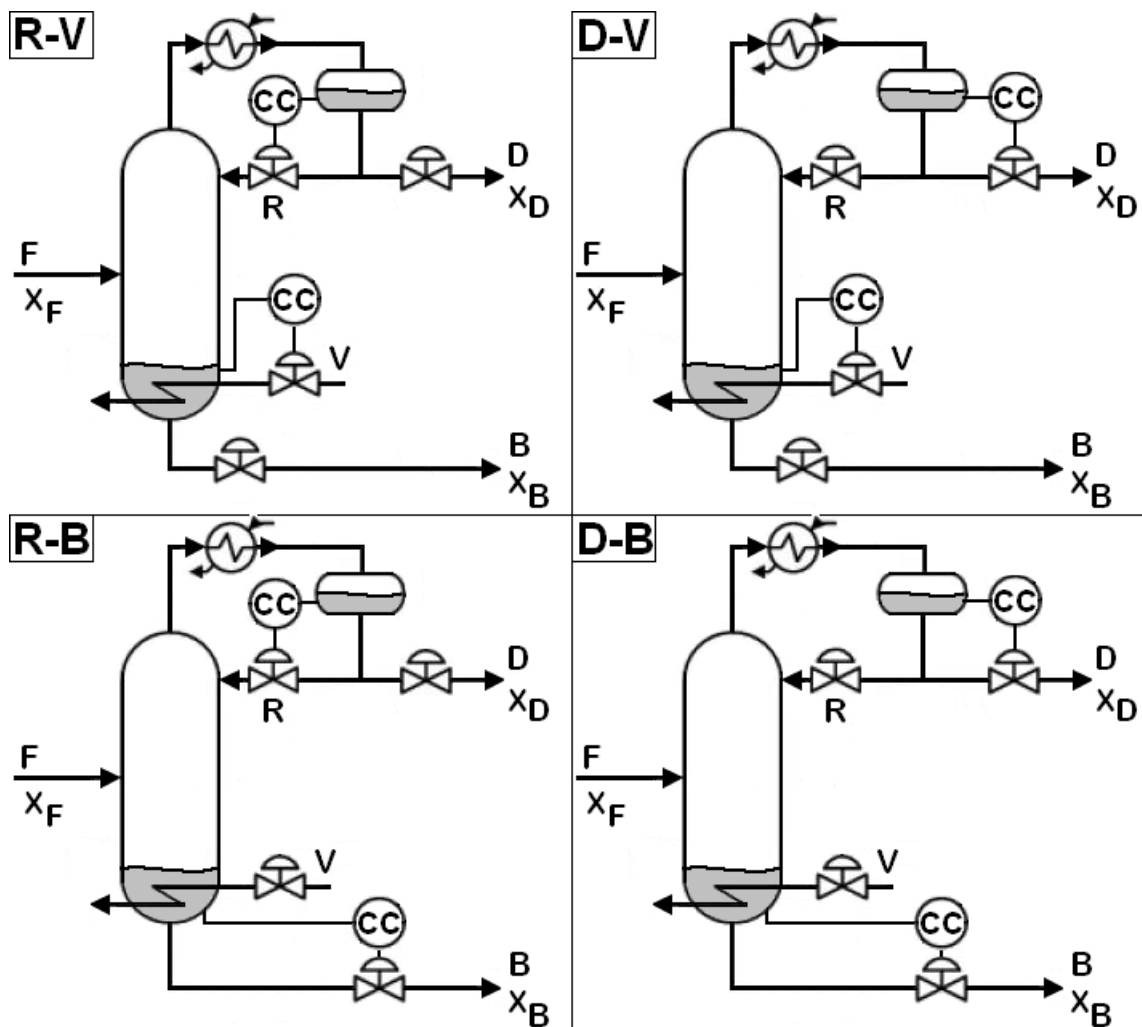


Figura 4.14: Estruturas controle da coluna de destilação binária ideal.

O controle PID discreto com amostragem das medidas do processo em intervalos de 15 segundos é implementado para o controle das frações molares do produto de topo e do produto de fundo. Os níveis de líquido do tambor de refluxo e do refeedor são

considerados constantes. Normalmente, controladores proporcionais conseguem manter os níveis de modo satisfatório.

Simulando a operação da coluna para os problemas do tipo servo ($x_B^{sp} = 0,010$ e $x_D^{sp} = 0,99$) e regulador ($x_f = 0,50$) com os parâmetros apresentados na Tabela (4.2), obtém-se o resultado mostrado pela Figura (4.15). As condições iniciais para a simulação da coluna são encontradas em Perry et al. (1997).

Tabela 4.2: Parâmetros utilizados pelo o controlador PID servo e regulador da coluna de destilação binária.

Parâmetros	RV	DB	RB	DV
Problema Servo				
Controlador de x_B				
k_c (lbmol/min)	-1350	1500	850	-5000
τ_i (-)	4	70	20	3,5
τ_d (-)	0	0,5	0	0
Controlador de x_D				
k_c	800	-3500	800	-1800
τ_i	4	50	10	200
τ_d	1	0	0	0
Problema Regulador				
Controlador de x_B				
k_c	-1350	600	850	-2000
τ_i	4	15	20	40
τ_d	0	0	0	0,6
Controlador de x_D				
k_c	800	-550	800	-800
τ_i	4	30	10	30
τ_d	1	5	0	1,20

4.2.3 Estudo de falhas em processos de destilação

Neste exemplo, explora-se de forma independente, falhas nos quatro atuadores utilizados. Dependendo da estrutura de controle utilizada é conveniente identificar qual atuador está com defeito para que o sistema de controle tolerante verifique a possibilidade de alterar entre as estruturas de controle disponíveis para que os mesmos objetivos de controle sejam alcançados, ou seja, controlar a composição do componente leve no topo e no produto de fundo. Quando não for possível alcançar os mesmos objetivos de controle na presença de falhas é conveniente manter um dos objetivos, continuar operando o processo e seguir a orientação do sistema de diagnóstico para tomar outra ação corretiva.

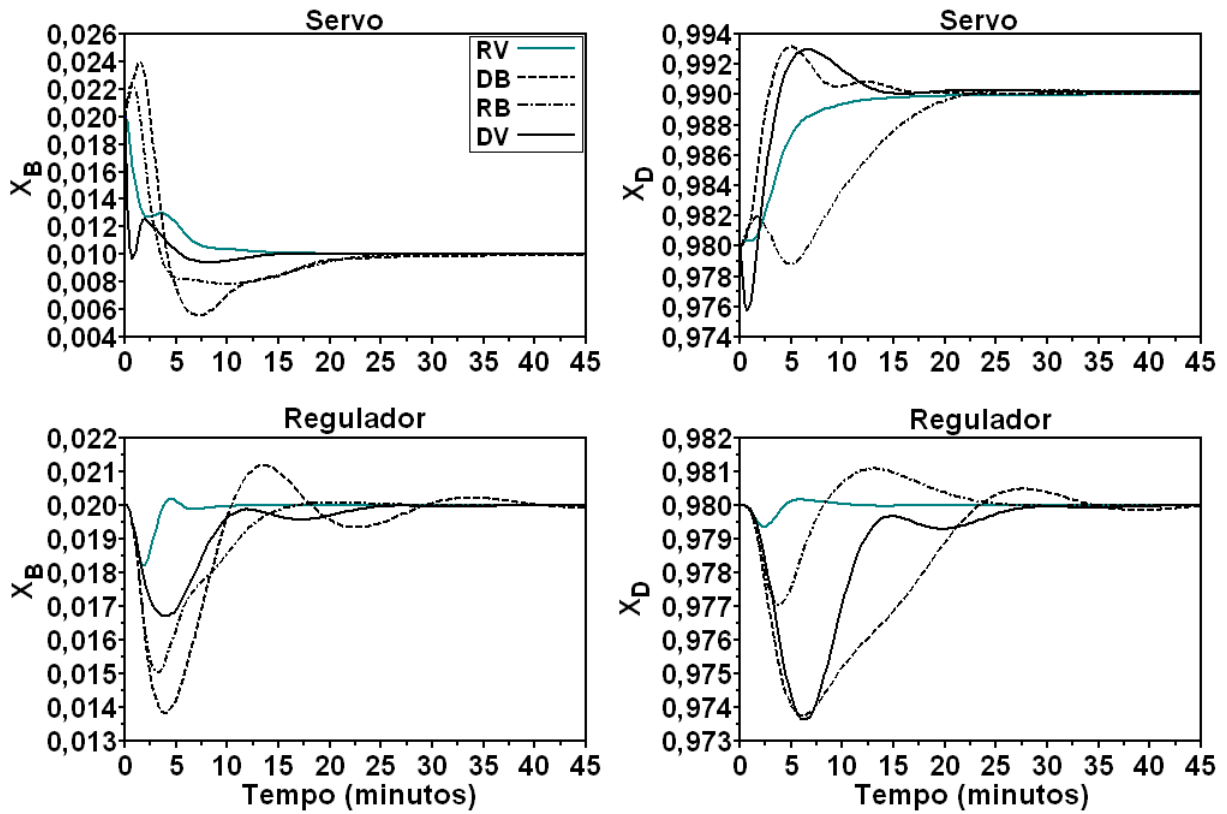


Figura 4.15: Resultados dos problemas de controle servo e regulador para a coluna de destilação binária.

Como a coluna de destilação possui quatro atuadores e quatro estruturas de controle disponíveis, é possível a ocorrência de falhas em cada atuador que recebe ação de controle para cada estrutura, resultando assim em um total de oito falhas. O processo opera em condições normais nas quatro estruturas de controle e as simulações para o controle regulador num tempo de 2 horas formam a base de dados de operação normal. Durante as simulações, ruídos com média 0 e variância de 10^{-6} são inseridos nas variáveis medidas listadas na Tabela (4.3).

Tabela 4.3: Variáveis medidas na coluna de destilação binária.

Variável	Significado	Unidade
x_B	Fração molar (leve) no fundo da coluna.	-
x_D	Fração molar (leve) no topo da coluna.	-
x_5	Fração molar (leve) no prato número 5.	-
M_5	Hold-up de líquido no prato 5.	lbmol
x_{15}	Fração molar (leve) no prato número 15.	-
M_{15}	Hold-up de líquido no prato 15.	lbmol
R	Vazão de refluxo.	lbmol/min
V	Vazão de vapor ascendente.	lbmol/min
D	Vazão de produto de topo.	lbmol/min
B	Vazão de produto de fundo	lbmol/min

Dados de operação normal da estrutura de controle **RV** para as variáveis medidas R , V , x_B e x_D são mostrados na Figura (4.16).

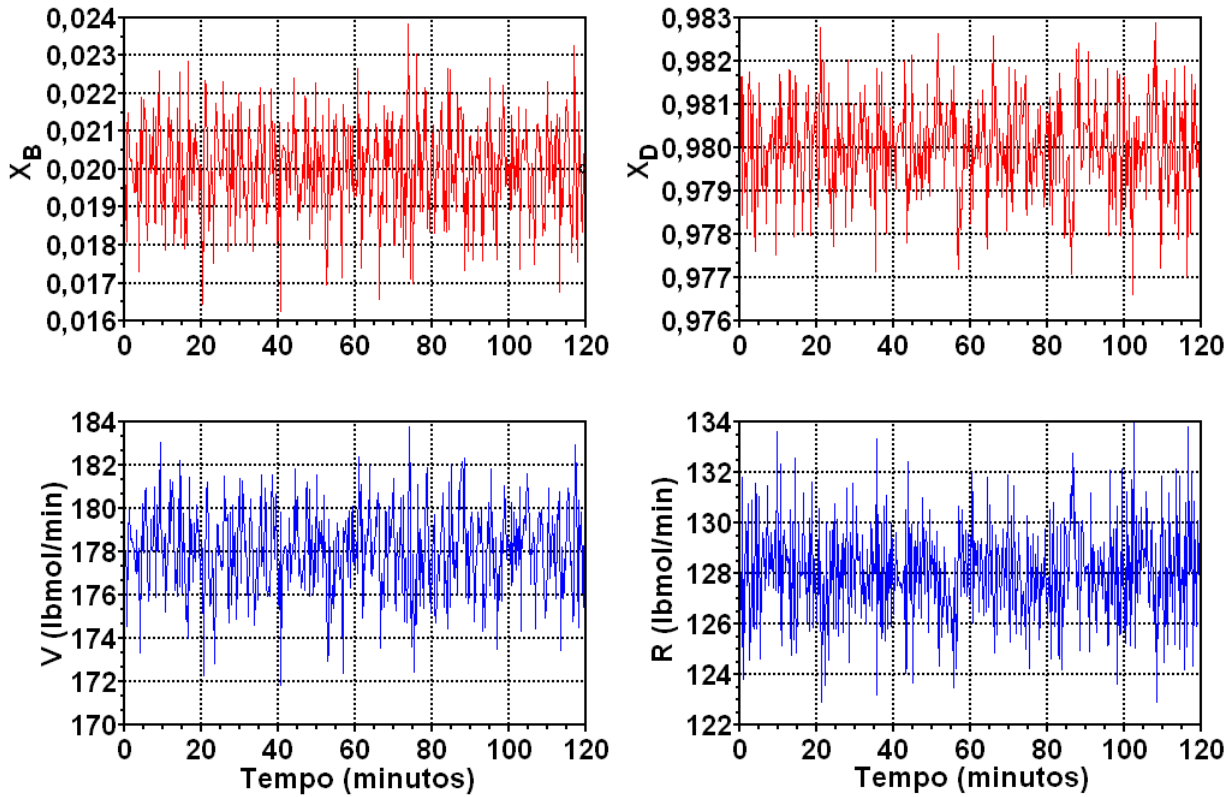


Figura 4.16: Resultados para operação em controle regulatório na estrutura de controle RV da coluna de destilação binária ideal.

As demais estruturas são simuladas em modo regulador e os dados são armazenados para compor a base de dados requerida para a estatística FDA responsável pela detecção e diagnóstico mostrados no Capítulo 2.

O resumo das falhas em atuadores exploradas neste trabalho é mostrado na Tabela (4.4). A simulação de falhas nos atuadores da coluna é feita através da fixação de valores para R , V , D e B , ou seja, supõe-se que a partir do instante da falha os atuadores permanecem em posições fixas (emperrados) cujos valores são mostrados na Tabela (4.4) e correspondem aos valores máximos permitidos para a operação dessa coluna. A Tabela (4.4) mostra quatro tipos de falhas em atuadores em cada uma das quatro estruturas. A construção da base de dados de falhas é feita através de simulações de 30 minutos de duração com a falha inserida imediatamente após o primeiro instante de amostragem. A Figura (4.17) mostra a coluna de destilação sendo controlada pela estrutura **DB** e com uma falha na válvula que controla a vazão de produto de fundo.

O comportamento obtido era esperado de acordo com o modelo, já que o aumento da retirada de produto de fundo faz com que x_B aumente, assim o efeito compensador

Tabela 4.4: Falhas nas estruturas de controle.

Estrutura	Falha	Reconfiguração
R	Atuador R - refluxo ($R_f=180$ lbmol/min)	Outra malha de controle
V	Atuador V - vapor ($V_f=200$ lbmol/min)	Outra malha de controle
D	Atuador D - destilado ($D_f=80$ lbmol/min)	Usar estrutura RB
B	Atuador B - fundo ($B_f=60$ lbmol/min)	Usar estrutura DV
R	Atuador R - refluxo ($R_f=180$ lbmol/min)	Outra malha de controle
B	Atuador B - fundo ($B_f=60$ lbmol/min)	Usar estrutura RV
D	Atuador D - destilado ($D_f=100$ lbmol/min)	Usar estrutura RV
V	Atuador V - vapor ($V_f=200$ lbmol/min)	Outra malha de controle

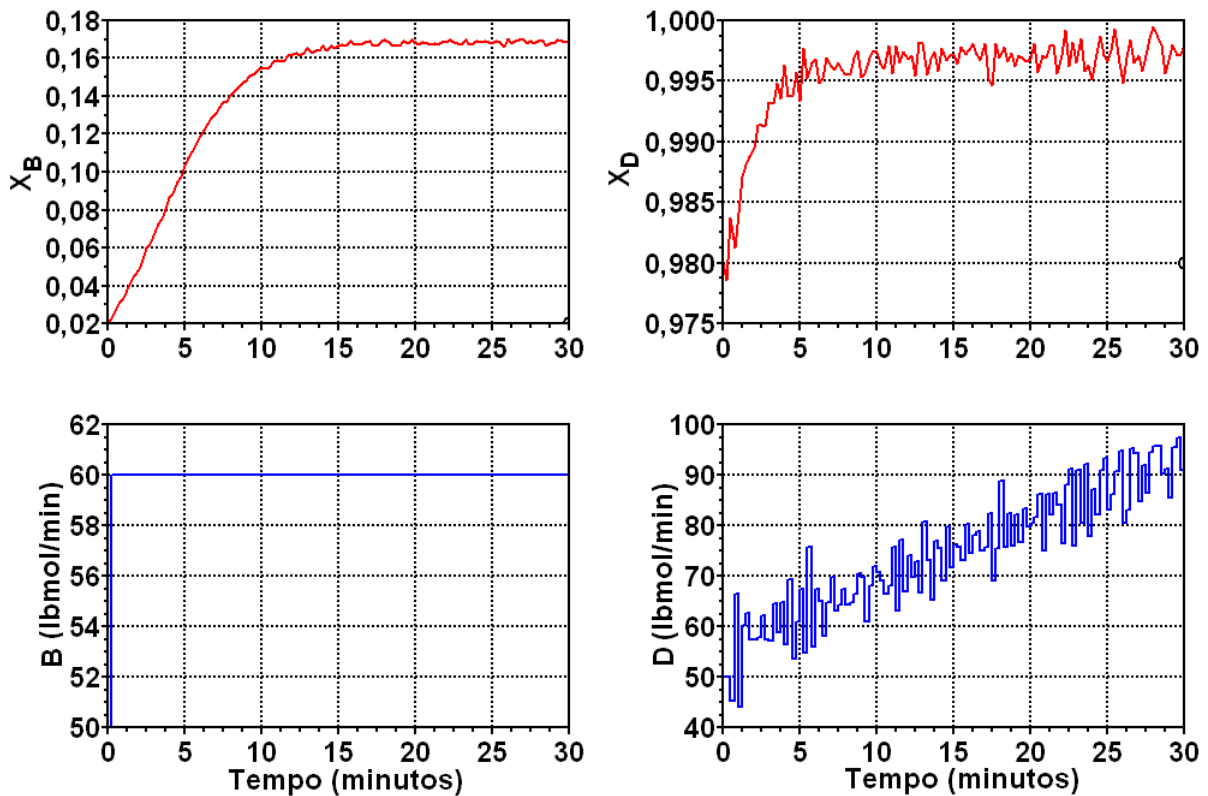


Figura 4.17: Simulação de uma falha na estrutura DB (atuador B).

no topo da coluna é o de aumentar a vazão de destilado para tentar compensar a vazão constante de B devido a falha. Nota-se também que o processo é levado a operar em outro ponto de operação e que esta falha proporciona um aumento constante da vazão de destilado.

4.2.4 Detecção, identificação e diagnóstico de falhas

O sistema de detecção e diagnóstico proposto para este exemplo baseia-se na utilização da Análise Discriminante de Fisher (FDA). Para a construção deste sistema FDD (*Fault*

Detection and Diagnosis, Detecção de Falha e Diagnóstico) é necessário criar uma base de dados de operação normal e com falhas para a coluna de destilação. De acordo com a Tabela (4.4) simula-se oito tipos de falhas para as quatro estruturas de controle disponíveis.

A utilização da FDA para este exemplo é vantajosa pelo fato da estrutura de controle utilizada sempre ser conhecida, possibilitando assim a construção de modelos FDA para cada estrutura de controle. Dessa forma, para a estrutura de controle RV, por exemplo, tem-se 3 classes de dados, uma para a operação normal, uma para a falha na válvula de refluxo e outra para falha na válvula de vapor. O mesmo é válido para as outras estruturas de controle com as respectivas falhas que nelas ocorrem. Assim, o sistema de diagnóstico sempre classifica entre três condições operacionais para cada estrutura selecionada. Essa divisão é uma alternativa à consideração convencional de utilização de todos os dados (operação normal e com falhas) em um único sistema de diagnóstico.

Como proposto no Capítulo 2, é comum a redução de dimensionalidade dos dados antes da implementação em linha de qualquer sistema de detecção e diagnóstico. Os dados do conjunto de treinamento (12 classes) geram quatro modelos FDA, um para cada estrutura de controle e a escolha da dimensão que deve ser mantida se baseia na redução da taxa de classificações incorretas do conjunto de treinamento. Os resultados obtidos são mostrados na Tabela (4.5).

Tabela 4.5: Taxa de classificações incorretas (TCI) percentual para os dados do conjunto de treinamento para cada modelo FDA e redução de ordem para a coluna de destilação binária.

Estrutura	Falha	TCI	Redução de ordem [†]
RV	Operação normal	0,00%	2
	Atuador R	0,82%	
	Atuador V	0,82%	
DB	Operação normal	0,00%	2
	Atuador D	1,65%	
	Atuador B	0,82%	
RB	Operação normal	0,00%	2
	Atuador R	0,82%	
	Atuador B	0,82%	
DV	Operação normal	0,00%	2
	Atuador D	0,82%	
	Atuador V	1,65%	

[†]A ordem total dos modelos FDA é 10 (número de variáveis medidas).

Para todos os casos mostrados na Tabela (4.5), a taxa de classificações incorretas é bastante pequena e é uma indicação que um procedimento de classificação implementado em tempo real terá sucesso. Além disso, a escolha da ordem igual a 2 para todos os casos se baseia na minimização da taxa de classificações incorretas, que para este caso foi obtida

pelo acréscimo da ordem do modelo e avaliação da TCI para cada estrutura de controle.

A partir dos resultados apresentados na Tabela (4.5) projeta-se o sistema de detecção e diagnóstico de falhas em linha para o processo. Durante as simulações, a classificação final somente é concluída após a persistência do diagnóstico por mais de 3 instantes de amostragem.

A Tabela (4.6) mostra todos resultados referentes ao diagnóstico e detecção utilizando modelos FDA para cada estrutura de controle proposta. A classificação em linha é feita através de uma avaliação periódica do resultado do algoritmo de diagnóstico. As simulações das falhas ocorrem a partir do tempo de 10 minutos de um total de 30 minutos de simulação.

Tabela 4.6: Resultados do algoritmo FDA proposto para monitoramento em tempo real para dados independentes do conjunto de treinamento para a coluna de destilação binária.

Estrutura	Falha	TCI	AD [‡] (minutos)
RV	Operação normal	0,00%	-
	Atuador R	0,80%	0,50
	Atuador V	0,80%	0,50
DB	Operação normal	0,00%	-
	Atuador D	1,60%	0,75
	Atuador B	0,80%	0,50
RB	Operação normal	0,00%	-
	Atuador R	0,80%	0,50
	Atuador B	0,80%	0,50
DV	Operação normal	0,00%	-
	Atuador D	0,80%	0,50
	Atuador V	0,80%	0,50

[‡]AD significa atraso de detecção em frações de tempo de 1 minuto.

Para exemplificar a técnica de detecção e diagnóstico utilizando Análise Discriminante de Fisher utiliza-se a estrutura RB. Supondo que uma falha (no atuador responsável pela retirada de produto de fundo) ocorre a partir de 10 minutos, mostra-se a resposta do processo ao longo do tempo na Figura (4.18) e a resposta do esquema de persistência de classificação para a detecção e diagnóstico de falhas na Figura (4.19).

Fica claro pela Figura (4.19) que a detecção e diagnóstico podem ser feitas com um atraso máximo de 3 instantes de amostragem, que é a quantidade de classificações iguais que devem ocorrer para que o sistema produza um diagnóstico final.

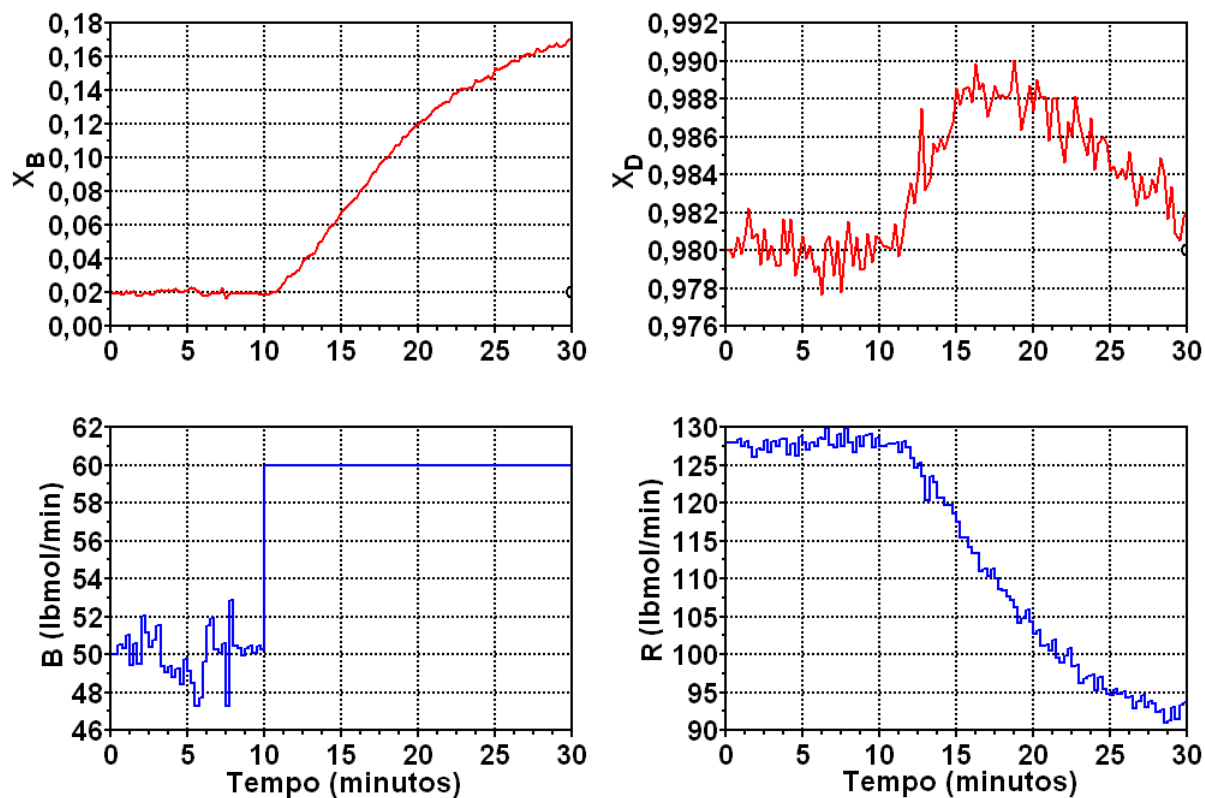


Figura 4.18: Simulação de uma falha na estrutura RB (atuador B) para a coluna de destilação binária ideal.

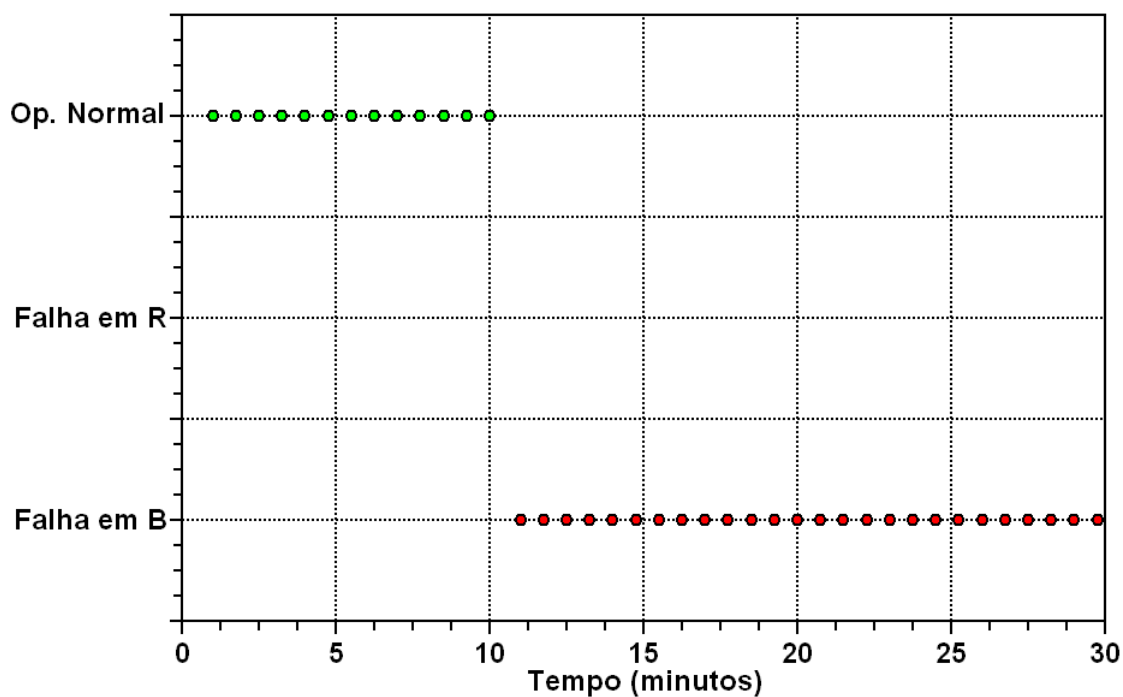


Figura 4.19: Detecção e diagnóstico utilizando FDA para a estrutura de controle RB com falha no atuador B.

4.2.5 Controle tolerante para colunas de destilação

A proposta de controle tolerante a falhas mostrada neste exemplo utiliza as estruturas de controle disponíveis para a operação convencional e transforma toda estrutura tolerante a falhas em atuadores, através do conhecimento do estado do processo (falha ou operação normal) através da FDA e alteração entre estruturas de controle para acomodar efeitos de falhas.

Falhas em atuadores que requerem o uso de outras malhas de controle não são explorados neste trabalho. Assim, o esquema de controle tolerante implementado aplica as principais idéias do Capítulo 3, ou seja, a utilização de um banco de controladores previamente sintonizados (estruturas RV, DB, RB e DV) aliado com um sistema de detecção e diagnóstico de falhas baseado em FDA eficiente e capaz de detectar e diagnosticar falhas em tempo adequado.

Para exemplificar a estratégia de controle tolerante proposta, utiliza-se a estrutura DV para controlar o processo em modo regulador e simula-se uma falha no atuador D, destinado à retirada de produto de topo. Assim, a partir de 30 minutos de simulação (total de 1 hora) a falha é inserida. De acordo com a Tabela (4.4), é possível utilizar a estrutura de controle RV para acomodar seu efeito. A troca de estruturas somente será realizada quando o sistema de diagnóstico e detecção por FDA for capaz de decidir sobre o estado do processo. A Figura (4.20) mostra o comportamento da coluna de destilação para a falha em D, controlada pela estrutura DV.

A Figura (4.21) mostra o comportamento do processo durante uma troca de estrutura de controle (de DV para RV) com três instantes de amostragem de atraso devido ao sistema de diagnóstico por FDA. Observa-se que esta troca é uma alternativa eficaz para controlar o processo.

4.2.6 Discussão dos resultados

A utilização da FDA como uma técnica de detecção e diagnóstico de falhas é vantajosa devido ao fato de ser ideal quando as classes são prolongadas, ou seja, dados do conjunto de treinamento e das falhas possuem dimensões diferentes. Além disso, o controle de uma coluna de destilação favorece a utilização de um modelo FDA para cada estrutura de controle. Como a estrutura de controle é sempre conhecida, o projeto de um sistema de diagnóstico desta maneira é menos susceptível a classificações incorretas, visto que cada modelo FDA para cada estrutura de controle possui apenas 3 classes (uma para operação normal e duas para falhas), melhor do que utilizar todas classes de falhas de todas estruturas de controle para classificação (12 classes).

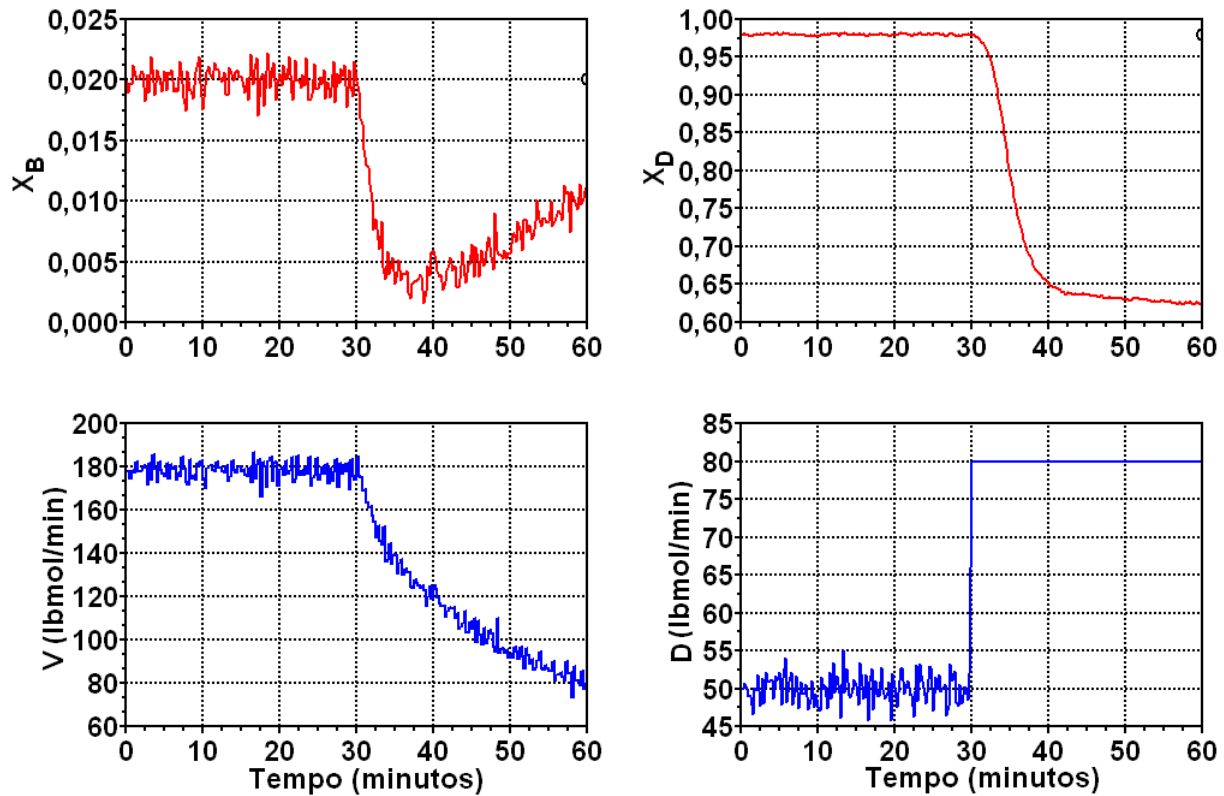


Figura 4.20: Simulação de uma falha em D para o processo controlado pela estrutura DV.

Os resultados obtidos mostram que a alternância entre estruturas de controle numa coluna de destilação é uma alternativa para acomodar efeitos de falhas, no entanto, existem falhas que não podem ser acomodadas pela simples troca de estrutura e requerem uma investigação mais detalhada. Mesmo assim, mostra-se com este exemplo a habilidade de combinar técnicas de detecção e diagnóstico baseadas em dados com uma estrutura de controle PID para tornar o sistema de controle tolerante a falhas.

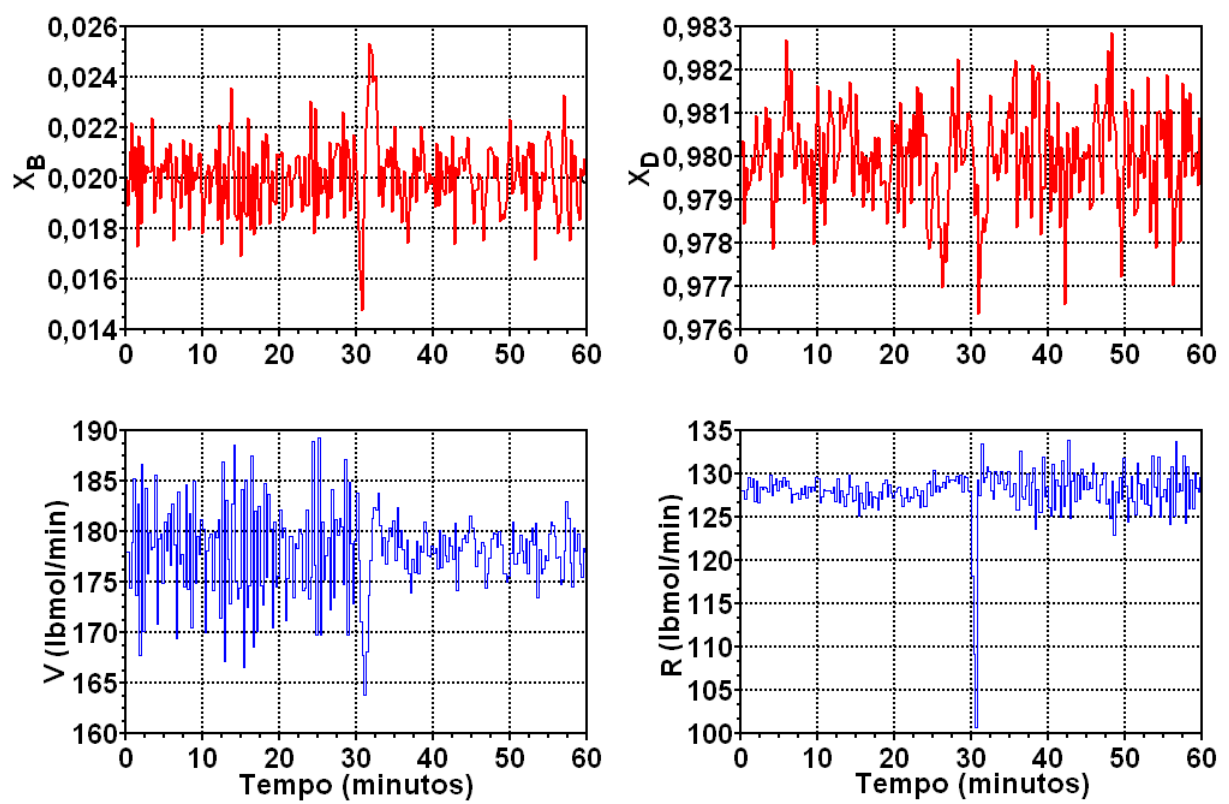


Figura 4.21: Controle tolerante para falha em D na estrutura DV. Troca de estrutura de controle (de DV para RV).

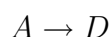
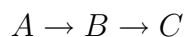
4.3 O Processo de Produção de Ciclopentanol

O processo de produção de ciclopentanol é utilizado aqui como exemplo no qual as técnicas de detecção e diagnóstico de falhas baseadas na análise dos componentes principais é utilizada como fonte de informações para um sistema supervisor responsável por determinar quais ações devem ser tomadas para que ocorra alguma alteração no modo de operação do processo para remover o efeito de uma falha.

A idéia principal para a simulação deste processo com controle tolerante é utilizar as técnicas apresentadas nos Capítulos 2 e 3, detecção e diagnóstico baseados em dados e análise estrutural, respectivamente, para acrescentar a característica de tolerância a falhas para o controlador PI regulador projetado.

4.3.1 Descrição do Problema

Considera-se a produção de ciclopentanol (B) partindo de ciclopentadieno (A) por adição eletrofílica de água catalisada por ácido em solução diluída como mostrado por Klatt e Engell (1993). Devido a forte reatividade do produto e dos reagentes, di-ciclopentadieno (D) é produzido por reação Diels-Alder como um produto lateral, e ciclopentanodiol (C) como um produto consecutivo através da adição de uma outra molécula de água. O esquema reacional completo é



A Figura (4.22) mostra esquematicamente o reator de mistura perfeita com a respectiva malha de controle PI proposta. A alimentação ao reator é somente reagente A em baixa concentração C_{A0} .

Durante a modelagem do reator contínuo e agitado supõe-se que a densidade do fluido é constante e o tempo de residência é ideal. Assim, as equações de balanço são escritas da seguinte forma:

$$\frac{dC_A}{dt} = \frac{\dot{W}_R}{V_R}(C_{A0} - C_A) - k_1(T_R)C_A - k_3(T_R)C_A^2 \quad (4.9)$$

$$\frac{dC_B}{dt} = -\frac{\dot{W}_R}{V_R}C_B + k_1(T_R)C_A - k_2(T_R)C_B \quad (4.10)$$

$$\frac{dT_R}{dt} = \frac{\dot{W}_R}{V_R}(T_{R0} - T_R) + \frac{U_R A_R}{\rho_R C_{pR} V_R}(T_{Rj} - T_R) +$$

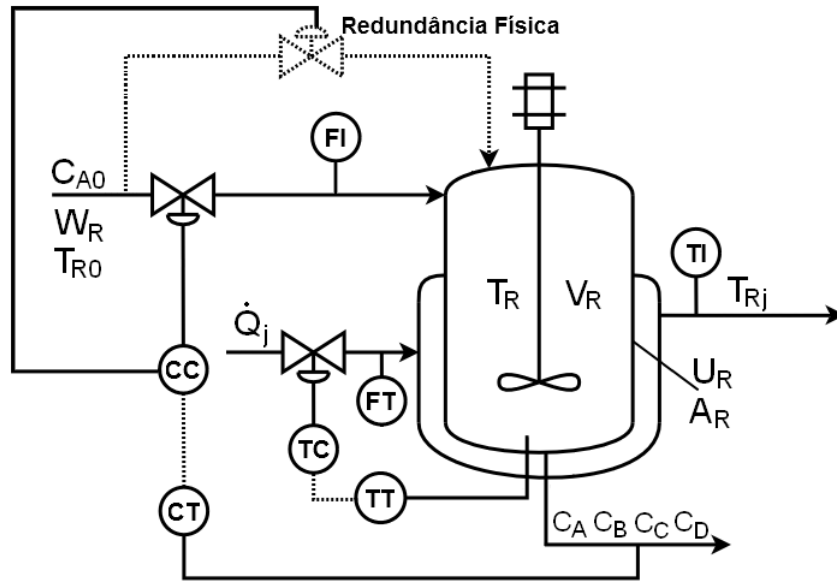


Figura 4.22: Esquema do processo investigado.

$$- \frac{k_1(T_R)C_A\Delta H_R^{AB} + k_2(T_R)C_B\Delta H_R^{BC} + k_3(T_R)C_A^2\Delta H_R^{AD}}{\rho_R C_{pR}} \quad (4.11)$$

$$\frac{dT_{Rj}}{dt} = \frac{\dot{Q}_j}{m_j C_{pj}} + \frac{U_R A_R (T_R - T_{Rj})}{m_j C_{pj}} \quad (4.12)$$

constituem um sistema não linear de quarta ordem. Aqui, C_A e C_B representam as concentrações do reagente A e do produto desejado B dentro do reator, T_R denota a temperatura do reator e T_{Rj} da camisa de resfriamento. As constantes da taxa k_1 , k_2 e k_3 dependem exponencialmente da temperatura do reator T_R via lei de Arrhenius

$$k_i(T_R) = k_{0i} \exp\left(\frac{-E_{Ai}}{R(T_R + 273,15)}\right) \quad (4.13)$$

Supõe-se que a temperatura do reator, a concentração de ciclopentadieno e ciclopentanol na saída reator, a temperatura da camisa de resfriamento, a vazão de alimentação de reagente e de líquido refrigerante são obtidas através de instrumentos de medição (ver Figura (4.22)). Para efeito de simulação, acrescenta-se um ruído gaussiano com média 0 e variância 10^{-5} para as concentrações e 10^{-3} para as demais medidas. A vazão de entrada de reagente A , $\dot{W}_R$, e a taxa de calor removido, $\dot{Q}_j$, são as variáveis manipuláveis e estão sujeitas às seguintes restrições:

$$50 \frac{L}{h} \leq \dot{W}_R = u_1 \leq 350 \frac{L}{h}$$

$$-8500 \frac{kJ}{h} \leq \dot{Q}_j = u_2 \leq 0 \frac{kJ}{h}$$

O controle da concentração do produto de interesse, C_B , é especificado de tal forma que qualquer valor na faixa de $0,7 \text{ mol/L} \leq C_B \leq 0,95 \text{ mol/L}$ pode ser alcançado sem erro do estado estacionário para todos valores de concentrações de entrada não medidas, que é nominalmente $C_{A0} = 5,1 \text{ mol/L}$, mas pode variar em uma faixa de $1 \text{ mol/L} \leq C_{A0} \leq 5,7 \text{ mol/L}$.

Os valores dos parâmetros do modelo são apresentados na Tabela (4.7):

Tabela 4.7: Parâmetros do modelo do reator de produção de ciclopentanol.

Parâmetros		
$k_{01} = 1,287 \cdot 10^{12} \text{ h}^{-1}$	$-E_{A3}/R = -8560 \text{ K}$	$C_{pR} = 3,01 \text{ kJ/kg K}$
$k_{02} = 1,287 \cdot 10^{12} \text{ h}^{-1}$	$\Delta H_R^{AB} = 4,2 \text{ kJ/mol A}$	$C_{pj} = 2,0 \text{ kJ/kg K}$
$k_{03} = 9,043 \cdot 10^9 \text{ h}^{-1}$	$\Delta H_R^{BC} = -11,0 \text{ kJ/mol B}$	$A_R = 0,215 \text{ m}^2$
$-E_{A1}/R = -9758,3 \text{ K}$	$\Delta H_R^{AD} = -41,85 \text{ kJ/mol A}$	$V_R = 10,0 \text{ L}$
$-E_{A2}/R = -9758,3 \text{ K}$	$\rho_R = 0,9342 \text{ kg/L}$	$m_j = 5,0 \text{ kg}$
$C_{A0} = 5,1 \text{ mol/L}$	$U_R = 4032 \text{ kJ/h m}^2 \text{ K}$	

4.3.2 Características do Processo

Estudos mostram que para uma taxa de calor removida constante e uma variação da alimentação de reagente entre 50 e 1500 L/h, existe para este processo 6 regiões com diferentes graus de não linearidades (OGUNNAIKE; RAY, 1994).

No presente trabalho, restringiu-se a vazão de alimentação entre 50 e 350 L/h para não ultrapassar os limites físicos proporcionais ao volume do reator simulado. Sendo assim, o mapeamento do espaço de saída para a concentração de ciclopentanol quando as entradas manipuladas u_1 e u_2 são alteradas resultou na superfície de resposta mostrada na Figura (4.23). Nota-se claramente pelo formato desta que a não linearidade do comportamento da concentração de saída de ciclopentanol do reator quando o mesmo opera nas extremidades das variáveis manipuladas, por exemplo, calor removido, $\dot{Q}_j$, entre -7000 e -8500 kJ/h e vazão de alimentação de reagente, $\dot{W}_R$, entre 50 e 100 L/h.

Para controlar o processo foram projetados dois controladores PI (Proporcional-Integral), um destinado a controlar a concentração de saída de ciclopentanol e outro para a temperatura do reator. Para demonstrar a idéia principal da técnica de controle tolerante utilizada neste trabalho, sem que fosse necessário projetar um controlador capaz de lidar com todas as regiões de elevada não linearidade. Assim, realizou-se um mapeamento entre as entradas e saídas do processo para escolher uma região de operação controlável por controladores do tipo PI. O valor de *set-point* para a concentração de ciclopentanol é igual a $0,69883 \text{ mol/L}$ e $407,031 \text{ K}$ para a temperatura do reator. Esses valores são

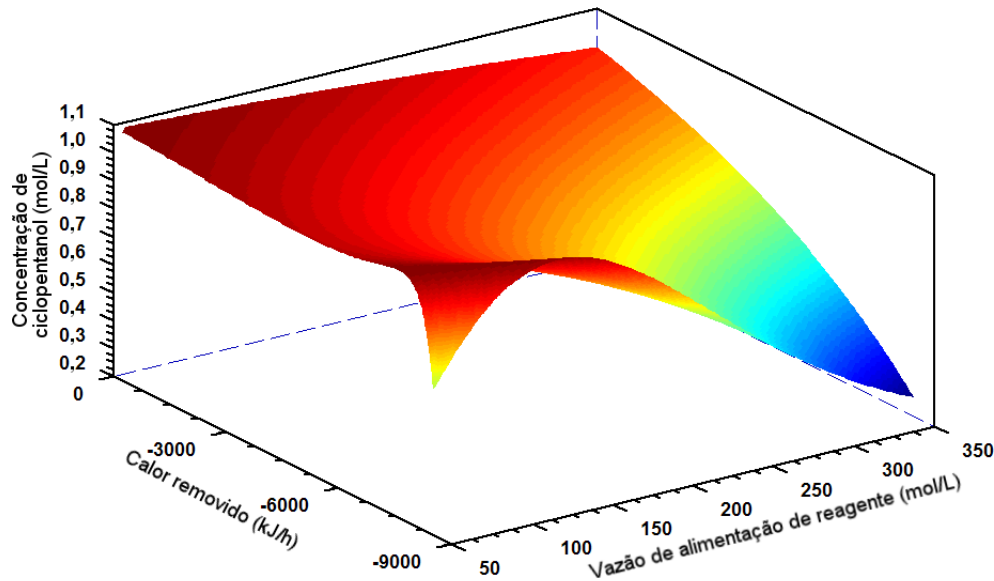


Figura 4.23: Superfície de resposta para a concentração de ciclopentanol.

relativos à operação em estado estacionário do modelo, gerados através da escolha de uma vazão de alimentação de reagente de 112 L/h ($\dot{W}_R = u_1$) e uma taxa de calor removido de $-2856,91$ kJ/h ($\dot{Q}_j = u_2$).

A Tabela (4.8) mostra os parâmetros do controlador PI regulador utilizados neste trabalho.

Tabela 4.8: Parâmetros do controlador PI regulador para o processo de produção de ciclopentanol.

CC		TC	
k_c	8 L ² /h mol	k_c	10 kJ/h K
τ_i	0,5 h ⁻¹	τ_i	0,1 K ⁻¹

4.3.3 Simulação do Processo

O estudo do processo de produção de ciclopentanol foi realizado através da implementação de um simulador. A proposta de simulação apresentada neste trabalho condiz com a estrutura proposta por Downs e Vogel (1993), quando apresentaram a simulação de um processo real, conhecido como Tennessee Eastman (TE).

O simulador do reator para produção de ciclopentanol foi implementado em *software* livre Scilab (Versão 4.1.1). Para ilustração, considera-se 4 cenários de falhas. A Falha 01, que é considerada uma perturbação na carga, ou seja, uma mudança de grau na alimentação de reagente A , C_{A0} , que inicialmente possui um valor de 5,1 mol/L, e após a

perturbação torna-se 4,7 mol/L. A Figura (4.24) mostra o comportamento das variáveis controladas e manipuladas sob a ação do controle regulador PI quando a Falha 01 é inserida. O simulador escolheu aleatoriamente o tempo de ocorrência de falha igual à fração de hora de 1,45. Nota-se que o controlador é capaz de lidar com esta mudança, no entanto é preciso que as variáveis manipuladas operem em outra região para que os objetivos de controle ainda sejam alcançados.

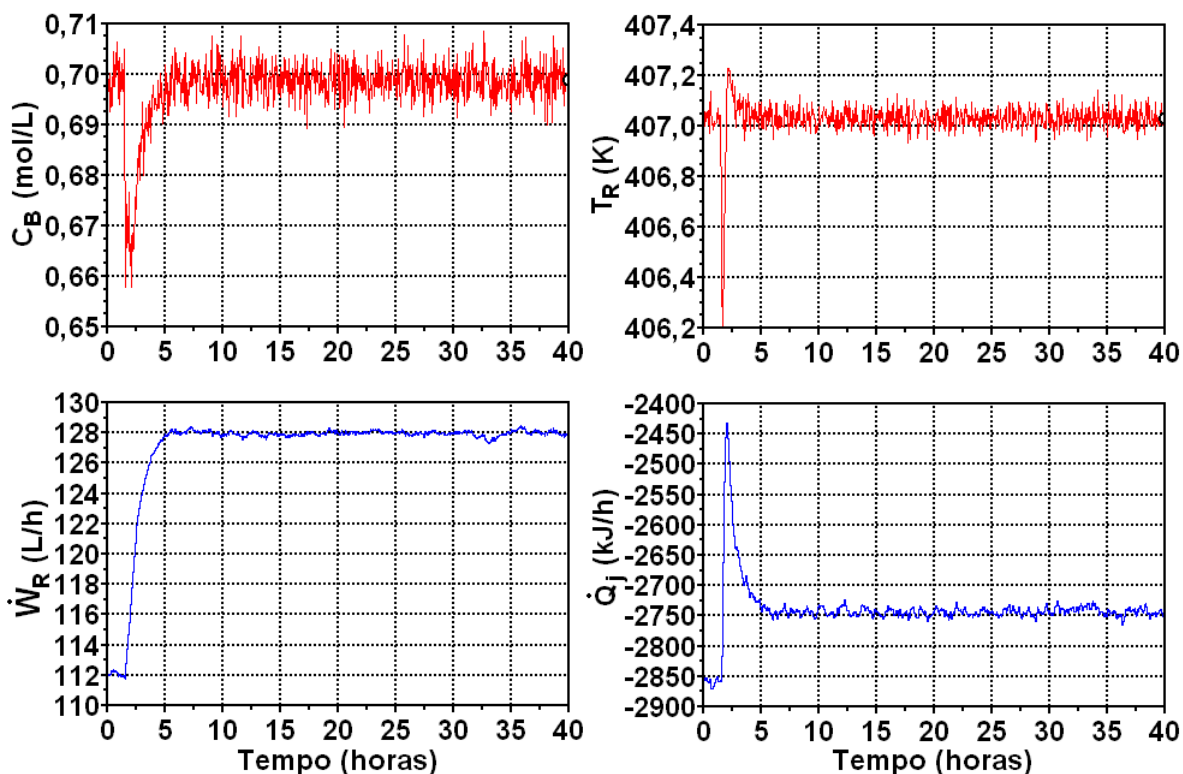


Figura 4.24: Simulação da Falha 01 para o reator de produção de ciclopentanol.

A segunda falha considerada, denominada Falha 02, trata de uma leitura incorreta da concentração de ciclopentanol, C_B , ou seja, a partir do momento em que a falha é inserida, simula-se uma medida de concentração corrompida, toma-se a última medida antes do aparecimento da falha multiplicada por um fator que varia entre 0,9 e 1,2, escolhido aleatoriamente através de uma distribuição uniforme e acrescida de um ruído aleatório gerado por uma distribuição gaussiana com média 0 e variância 10^{-5} . De acordo com a Figura (4.25) a falha ocorre na fração de hora igual a 13,6. Nesta falha, a medida de concentração que realimenta continuamente o controlador PI apresenta uma falha, fazendo com que a vazão de alimentação de reagente atinja o limite inferior, ocasionando mudanças drásticas nas variáveis controladas.

A Falha 03 considera que o sensor de temperatura do reator danifica-se a partir de um determinado instante, produzindo uma medida incorreta, que é simulada como sendo 1% maior que a última medida correta produzida pelo sensor acrescida de um ruído aleatório

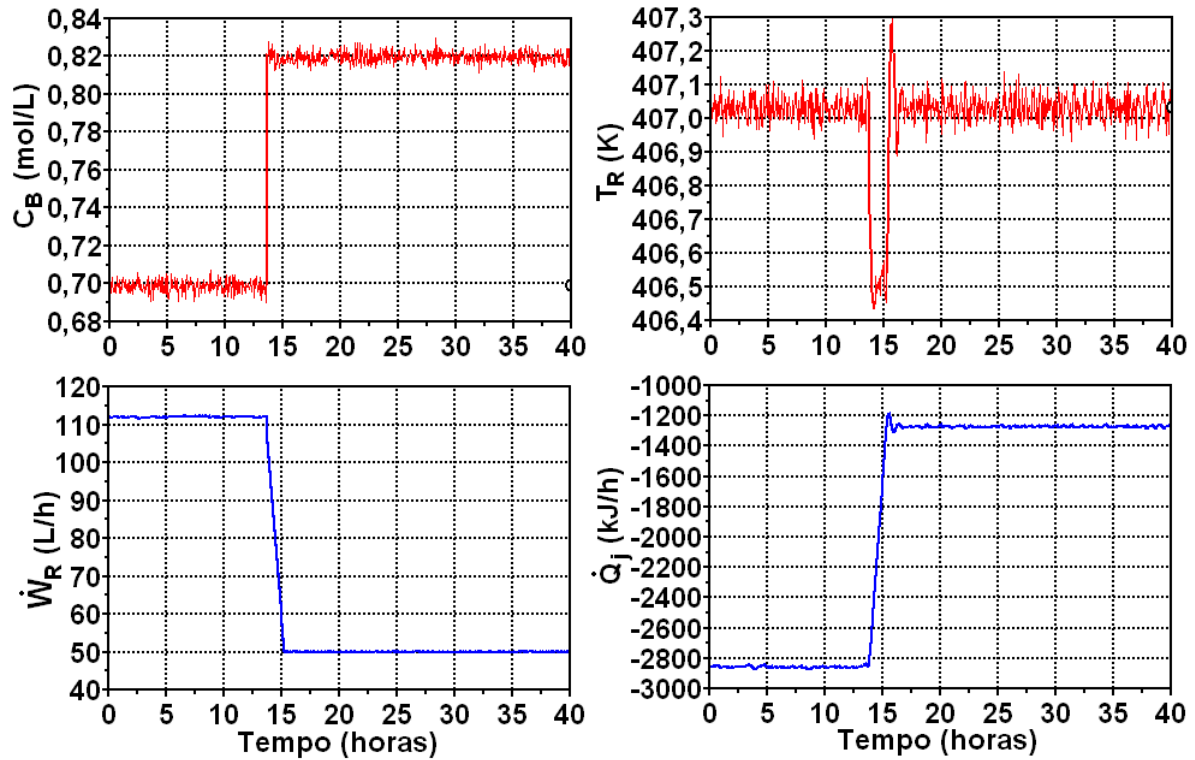


Figura 4.25: Simulação da Falha 02 para o reator de produção de ciclopentanol.

gerado por distribuição gaussiana com média 0 e variância 10^{-3} . A Figura (4.26) mostra que uma medida incorreta do sensor de temperatura que é enviada para o controlador PI faz com que ambas as variáveis controladas sejam levadas a operar em seus limites extremos. Observa-se por este gráfico o acoplamento das variáveis controladas e ainda a oscilação da concentração de ciclopentanol.

A última falha contida na base de dados do simulador é a Falha 04, que simula o emperramento da válvula de alimentação de reagente A numa posição que produz uma vazão 30% menor que a do estado estacionário (ver Figura (4.27)).

A Figura (4.28) apresenta a interface gráfica construída em Scilab para o simulador do processo de produção de ciclopentanol. A região de detecção reserva dois espaços destinados às métricas de detecção de falhas por T^2 e Q . Durante a simulação, aparecem nestes campos o estado instantâneo do processo, ou seja, “Op. normal” quando o sistema estiver funcionando na ausência de falhas, “Alarme”, quando as métricas indicarem que uma determinada observação se encontra entre a faixa de $\pm 10\%$ dos valores de T_α^2 e Q_α , e “Falha” quando uma determinada observação ultrapassar o limite superior da estatística.

Para a região de diagnóstico são reservados dois campos, no superior mostra-se em linha (*online*) a falha diagnosticada quando o limite superior da estatística é violado. O quadro inferior destina-se a mostrar uma resposta final sobre qual falha realmente ocorre. Isto é feito através de uma avaliação de persistência de diagnóstico, ou seja,

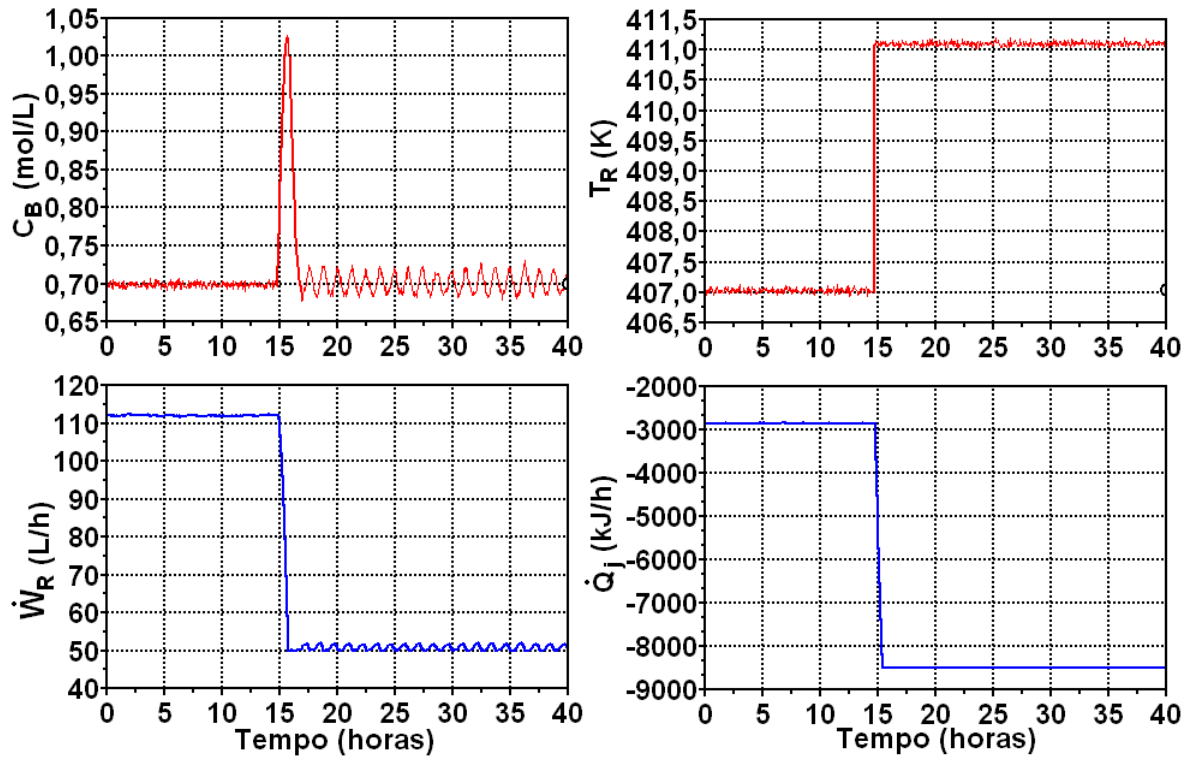


Figura 4.26: Simulação da Falha 03 para o reator de produção de ciclopentanol.

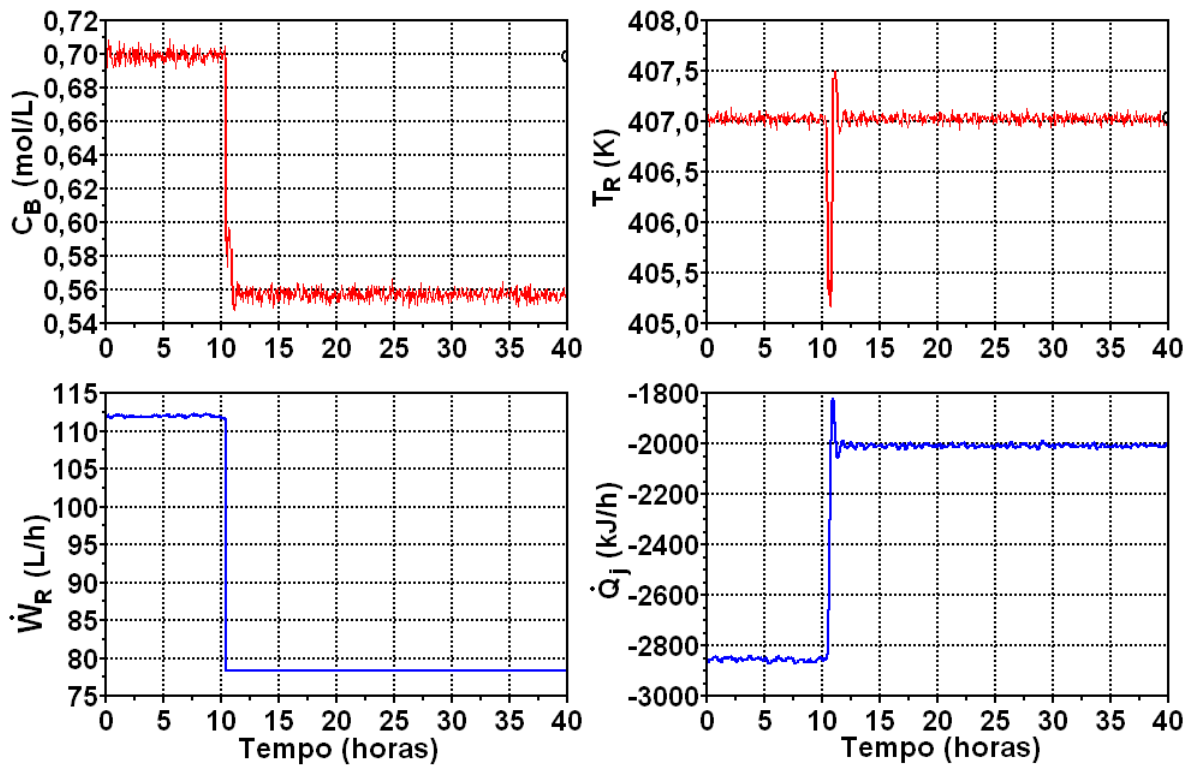


Figura 4.27: Simulação da Falha 04 para o reator de produção de ciclopentanol.

quando a resposta do sistema de diagnóstico se repetir por mais de três vezes após os limites de quaisquer estatísticas serem violados, o diagnóstico final se relaciona a esta

falha. Insere-se ainda neste quadro as informações relacionadas ao que deve ser reparado quando danificações de equipamentos ocorrerem.



Figura 4.28: Interface gráfica do simulador desenvolvido em linguagem Tcl/Tk para o processo de produção de ciclopentanol.

4.3.4 Detecção e Diagnóstico de Falhas

Detecção

A técnica de detecção de falhas utilizada neste trabalho é a técnica estatística denominada Análise dos Componentes Principais Dinâmica (DPCA, *Dynamic Principal Component Analysis*) e depende de dados de treinamento. Esses dados são oriundos de condições operacionais normais e são tratados fora de linha para se extrair informações essenciais para a detecção em tempo real.

As informações do conjunto de treinamento incluem dados medidos diretamente do processo ao longo do tempo relacionados à operação normal, utilizados pela DPCA. Para este exemplo, dados de operação normal são obtidos do processo com controle regulador operando em torno dos pontos estacionários projetados, então, as variáveis que formam a matriz de dados são apenas as variáveis que são medidas. A Figura (4.29) mostra a operação normal do reator para produção de ciclopentanol, operando em torno de um ponto estacionário especificado para as variáveis controladas (C_B e T_R) e manipuladas ($\dot{W}_R$ e $\dot{Q}_j$). O conjunto de dados de operação normal é formado pelas seguintes variáveis: temperatura do reator (T_R), concentração de ciclopentanol (C_B) na saída reator, a temperatura da camisa de resfriamento (T_{Rj}), a vazão de alimentação de reagente ($\dot{W}_R$) e de líquido refrigerante que é diretamente traduzida para a quantidade de calor removido

($\dot{Q}_j$). A medida da concentração do reagente A na saída do reator é utilizada apenas como medida redundante para o cálculo da concentração de ciclopentanol no esquema da análise estrutural, assim, não participa do conjunto de dados.

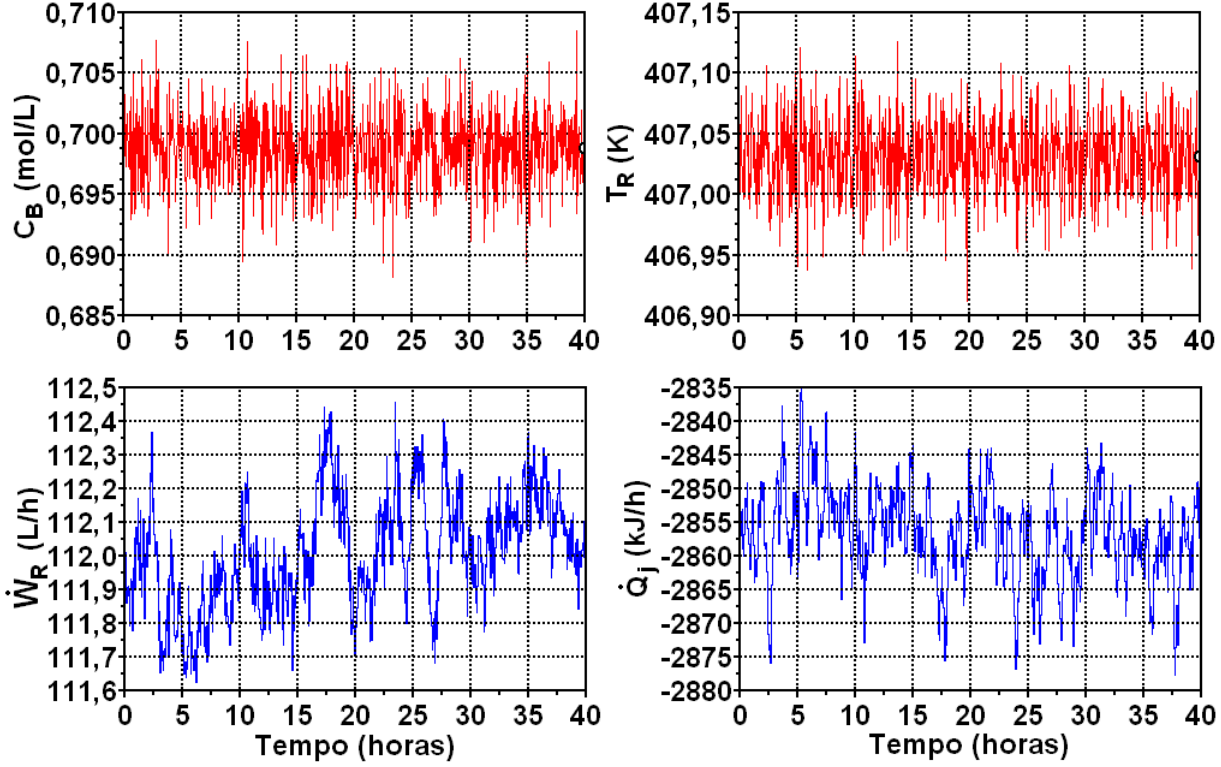


Figura 4.29: Operação normal do processo de produção de ciclopentanol.

Inicialmente, tomam-se os dados de operação normal como referência, estes são dispostos em uma matriz com h atrasos. Esta matriz é construída de acordo com (2.29):

$$\mathbf{X}(k-h,:) = [\mathbf{x}_k^T \ \mathbf{x}_{k-1}^T \ \mathbf{x}_{k-2}^T \ \cdots \ \mathbf{x}_{k-h}^T], \quad k = 1, 2, \dots, n.$$

A Equação (2.29) é válida para $k > h$, $\mathbf{x}_k^T$ é o vetor de observação que possui m variáveis medidas no instante de amostragem k .

Aplica-se à Equação (2.29) a técnica de redução de ordem conhecida como análise dos componentes principais, que é capaz de capturar a variabilidade dos dados de forma ótima. Esta variabilidade é revelada através da ordenação decrescente dos valores característicos calculados através da resolução de um problema de decomposição dos valores característicos de uma matriz de covariância amostral (Equação (2.3)):

$$\frac{1}{\sqrt{n-1}}\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T.$$

É comumente aceitável, e com certas suposições teoricamente justificáveis, que a porção do espaço PCA correspondendo aos maiores valores característicos descreve a maior parte das variações sistemáticas ou de estado que aparecem no processo, e a porção do espaço da PCA correspondente aos menores valores característicos descreve o ruído aleatório (CHIANG et al., 2001). Neste trabalho utiliza-se a análise paralela para determinar o número de dimensões que deve ser retida para o modelo PCA, que para este exemplo foi $a = 4$.

A utilização da DPCA faz com que a dimensão da matriz $\mathbf{X}$ cresça, por exemplo, para os dados de operação normal, são 5 variáveis medidas com 3 instantes de atraso para a construção da matriz de dados e 400 observações (20 horas de simulação ou 400 amostras de 3 minutos) para cada variável no conjunto de treinamento, faz com que a matriz possua dimensão $\mathbf{X} \in \mathbb{R}^{399 \times 15}$.

O monitoramento das variações sistemáticas do processo é realizado através da estatística T^2 , semelhante à Equação (2.13). Esta métrica leva em consideração apenas os a maiores valores característicos e seus vetores característicos associados.

$$T^2 = \mathbf{x}_h^T \mathbf{P} \mathbf{\Sigma}_a^{-2} \mathbf{P}^T \mathbf{x}_h$$

em que $\mathbf{x}_h$ (medidas independentes do conjunto de treinamento) representa um vetor de observação com h atrasos, $\mathbf{P}$ é a matriz de vetores característicos associados aos a maiores valores característicos e $\mathbf{\Sigma}_a$ contém os a maiores valores característicos dispostos em ordem decrescente ao longo da diagonal principal.

O limite para a estatística T^2 que é capaz de detectar falhas para observações tomadas de fora do conjunto de treinamento é dado pela Equação (2.15)

$$T_\alpha^2 = \frac{a(n-1)(n+1)}{n(n-a)} F_\alpha(a, n-a)$$

em que $F_\alpha(a, n-a)$ é a distribuição de Fisher com a e $n-a$ graus de liberdade.

O monitoramento das variações aleatórias leva em consideração a porção do espaço de observação excluído pela estatística T^2 , ou seja, os valores característicos descartados pela análise paralela. Sendo assim, este espaço pode ser mais robustamente monitorado utilizando a estatística Q (JACKSON; MUDHOLKAR, 1979)

$$Q = \mathbf{r}^T \mathbf{r}, \quad \mathbf{r} = (\mathbf{I} - \mathbf{P} \mathbf{P}^T) \mathbf{x},$$

em que $\mathbf{r}$ é o vetor residual, uma projeção da observação $\mathbf{x}_h$ no espaço residual. Como a estatística Q não mede diretamente as variações ao longo de cada vetor característico, mas

a soma total das variações no espaço residual, ela não sofre de sensibilidade a imprecisões nos menores valores característicos.

A distribuição para a estatística Q foi proposta por Jackson e Mudholkar (1979), e é apresentada no Capítulo 2:

$$Q_\alpha = \theta_1 \left[\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0}$$

Para dados do conjunto de treinamento a Figura (4.30) é obtida. Na parte superior desta figura tem-se a estatística T^2 e na parte inferior a estatística Q , ambas relativas à operação normal do processo com controle regulador. Nota-se que as estatísticas encontram-se bem abaixo do limite especificado para indicação de falhas (acima da última linha contínua) ou de alarmes falsos, entre as linhas. Construindo o modelo DPCA com dados de operação normal e testando este modelo com novos dados de operação normal também resultou na obediência dos limites, como esperado. O limite calculado pelas Equações (2.15) e (2.18) é a própria linha contínua inferior. Define-se como a região de alarme um limite 10% superior ao calculado por estas distribuições (T_α e Q_α).

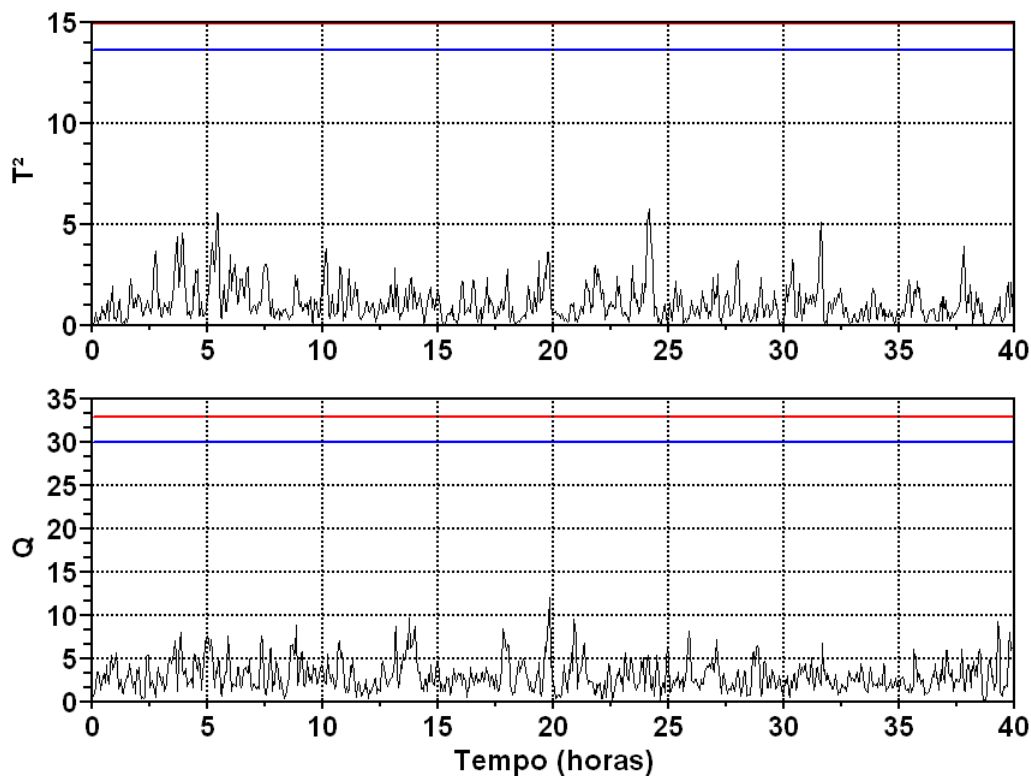


Figura 4.30: Estatísticas T^2 e Q para operação normal do reator de produção de ciclopentanol.

As Figuras (4.31) e (4.32) mostram os resultados para as estatísticas T^2 e Q para as

Falhas 01-04.

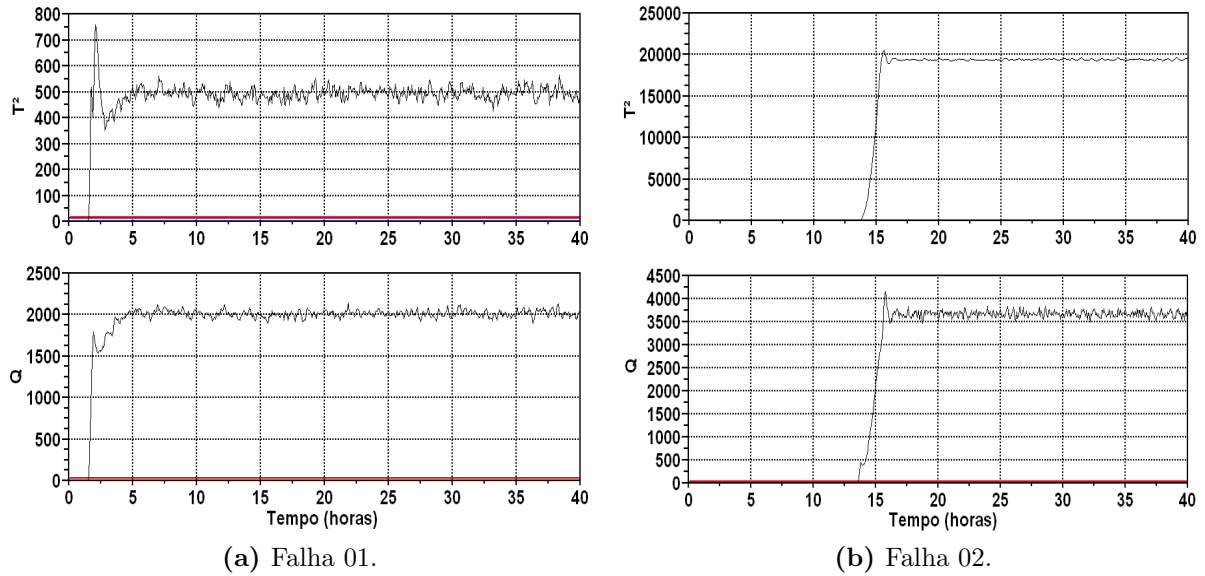


Figura 4.31: Estatísticas T^2 e Q para as Falhas 01 e 02 no reator de produção de ciclopentanol.

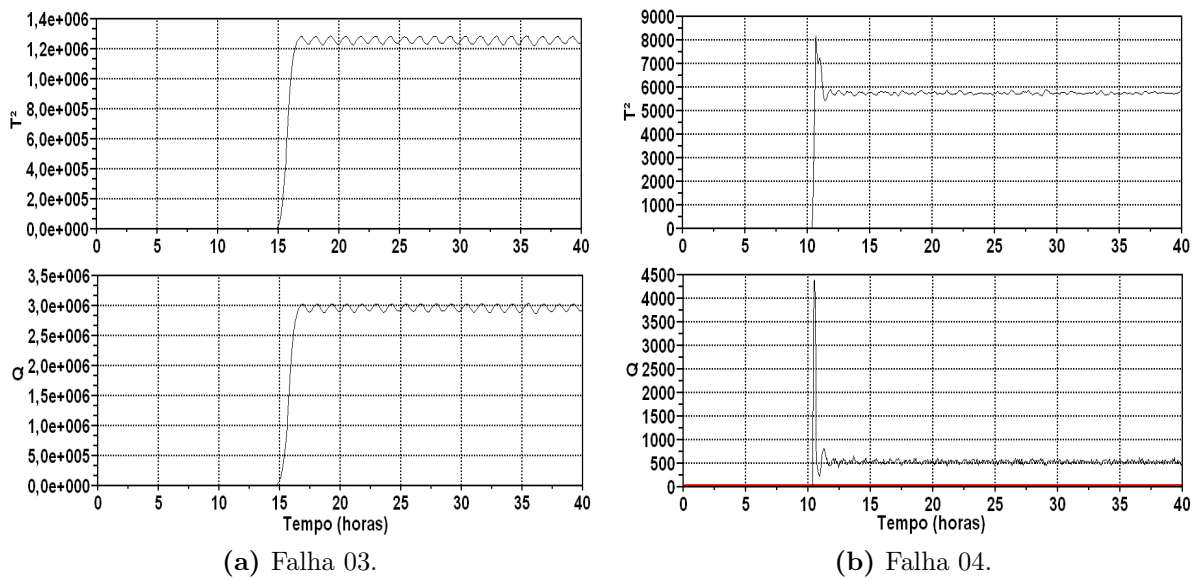


Figura 4.32: Estatísticas T^2 e Q para as Falhas 03 e 04 no reator de produção de ciclopentanol.

Diagnóstico

O objetivo de um sistema de diagnóstico é reconhecer o estado do sistema baseado em um banco de dados de operações com falhas. Em operações industriais é bastante comum a existência de um banco de dados com situações operacionais do processo, incluindo

as com falhas. Quando um processo real não está disponível, uma alternativa é utilizar um modelo do processo para gerar operações com falhas para treinar um sistema de diagnóstico. Uma aproximação DPCA é uma combinação de análise dos componentes principais e análise discriminante que é capaz de manipular um grande número de falhas e que usa um único modelo DPCA, e baseia-se em desenvolver um modelo separado utilizando dados coletados durante cada situação de falha específica. Aplica-se então as estatísticas T^2 e Q para predizer qual falha ocorreu.

Supondo que a probabilidade global para toda classe de falha i é a mesma e a matriz de covariância amostral de $\mathbf{P}_i \mathbf{x}_h$ para toda classe de falha i também é a mesma, tem-se:

$$T_i^2 = \mathbf{x}_h^T \mathbf{P}_i \Sigma_{a,i}^{-2} \mathbf{P}_i^T \mathbf{x}_h,$$

em que $\mathbf{P}_i$ é a matriz de vetores característicos para a falha i associados aos “ a ” maiores valores característicos e a matriz $\Sigma_{a,i}$ contém os a maiores valores característicos da classe de falha i .

Supondo que as variações importantes na discriminação entre as falhas estão contidas no espaço residual para cada classe de falha, é mais provável que uma observação seja representada pela classe de falha i com o mínimo discriminante residual

$$\frac{Q_i}{(Q_\alpha)_i}.$$

Se variações importantes na discriminação entre as falhas estão contidas no espaço residual, então uma observação é representada pela classe de falha i com o mínimo discriminante combinado, dado pela Equação (2.26)

$$c_i \left(\frac{T_i^2}{(T_\alpha^2)_i} \right) + (1 - c_i) \left(\frac{Q_i}{(Q_\alpha)_i} \right).$$

em que c_i é um fator de ponderação que fornece maior e menor importância para as estatísticas utilizadas na discriminação, T_i^2 e Q_i , para cada classe de falha i . Supondo que uma nova observação que exceda os limites impostos pelas estatísticas T^2 e Q não represente uma nova falha, a Equação (2.26) é utilizada para diagnosticar falhas.

4.3.5 Análise Estrutural

Conforme apresentado pelo Capítulo 3, pode-se reescrever o modelo de equações diferenciais de quarta ordem (Equações (4.9)-(4.13)) na forma estrutural. Para isto, basta dividir todas as variáveis em dois grupos principais, o de variáveis conhecidas ou medidas, $\mathcal{K}$, o

de variáveis desconhecidas $\mathcal{U}$ e o conjunto de restrições $\mathcal{R}$. No modelo estrutural a seguir (Equações (4.14)-(4.39)), $r1$ significa restrição de número 1, $m5$ é a restrição de número 5 que está ligada à uma variável medida e $d6$ é a restrição de número 6 que envolve derivada.

$$r1 : \dot{C}_A = m_1 - c_1 - c_2 \quad (4.14)$$

$$r2 : m_1 = \frac{W_R}{V_R}(C_{A0} - C_A) \quad (4.15)$$

$$r3 : c_1 = k_{01} \exp\left(\frac{-E_{A1}}{R(T_R + 273, 15)}\right) C_A \quad (4.16)$$

$$r4 : c_2 = k_{03} \exp\left(\frac{-E_{A3}}{R(T_R + 273, 15)}\right) C_A^2 \quad (4.17)$$

$$m5 : W_{Rm} = W_R + \omega \quad (4.18)$$

$$d6 : \dot{C}_A = \frac{dC_A}{dt} \quad (4.19)$$

$$r7 : \dot{C}_B = -m_2 + c_1 - c_3 \quad (4.20)$$

$$r8 : m_2 = -\frac{W_R}{V_R} C_B \quad (4.21)$$

$$r9 : c_3 = k_{02} \exp\left(\frac{-E_{A2}}{R(T_R + 273, 15)}\right) C_B \quad (4.22)$$

$$m10 : C_{Bm} = C_B + \omega \quad (4.23)$$

$$d11 : \dot{C}_B = \frac{dC_B}{dt} \quad (4.24)$$

$$r12 : \dot{T}_R = e_1 + e_2 - f_1 - f_2 - f_3 \quad (4.25)$$

$$r13 : e_1 = \frac{W_R}{V_R}(T_{R0} - T_R) \quad (4.26)$$

$$r14 : e_2 = \frac{U_R A_R}{\rho_R C_{pR} V_R}(T_{Rj} - T_R) \quad (4.27)$$

$$r15 : f_1 = k_{01} \exp\left(\frac{-E_{A1}}{R(T_R + 273, 15)}\right) \frac{C_A \Delta H_R^{AB}}{\rho_R C_{pR}} \quad (4.28)$$

$$r16 : f_2 = k_{02} \exp\left(\frac{-E_{A2}}{R(T_R + 273, 15)}\right) \frac{C_B \Delta H_R^{BC}}{\rho_R C_{pR}} \quad (4.29)$$

$$r17 : f_3 = k_{03} \exp\left(\frac{-E_{A3}}{R(T_R + 273, 15)}\right) \frac{C_A^2 \Delta H_R^{AD}}{\rho_R C_{pR}} \quad (4.30)$$

$$m18 : T_{Rm} = T_R + \omega \quad (4.31)$$

$$r19 : \dot{T}_R = \frac{dT_R}{dt} \quad (4.32)$$

$$r20 : \dot{T}_{Rj} = e_3 + e_4 \quad (4.33)$$

$$r21 : e_3 = \frac{\dot{Q}_j}{m_j C_{pj}} \quad (4.34)$$

$$r22 : e_4 = \frac{U_R A_R (T_R - T_{Rj})}{V_{Rj} \rho_j C_{pj}} \quad (4.35)$$

$$m23 : T_{Rjm} = T_{Rj} + \omega \quad (4.36)$$

$$m24 : W_{Rjm} = W_{Rj} + \omega \quad (4.37)$$

$$m25 : C_{Am} = C_A + \omega \quad (4.38)$$

$$d26 : \dot{T}_{Rj} = \frac{dT_{Rj}}{dt} \quad (4.39)$$

em que ω é o ruído aleatório gerado por uma distribuição gaussiana adicionado às variáveis medidas, que neste caso é conhecido e possui média 0 e variância 10^{-5} para as concentrações e média 0 e variância 10^{-3} para as demais medidas.

Supondo que não exista causalidade integral nem derivativa, não existirão problemas relacionados com o pareamento e o algoritmo de posicionamento pode ser utilizado. A apresentação do modelo estrutural também pode ser feita através da construção de uma matriz de incidência, que relaciona em suas linhas as restrições, em suas colunas as variáveis do processo e dentro da matriz utiliza-se o símbolo ‘ $\square$ ’ para atribuir uma determinada variável a uma restrição. A matriz de incidência para este processo é apresentada na Tabela (4.9) e não mostra o conjunto de variáveis conhecidas $\mathcal{K}$, pois estas não participam do pareamento.

A execução do algoritmo de posicionamento sobre esta matriz de incidência (Tabela (4.9)) gera um grafo direcionado, que é uma matriz de incidência com símbolos ‘ $\blacksquare$ ’, que indicam que determinada restrição (linha) será utilizada exclusivamente para calcular a variável da coluna correspondente. Assim, obtém-se a ordem em que as restrições devem ser percorridas para que a partir das variáveis conhecidas, seja possível monitorar todo o conjunto de variáveis desconhecidas $\mathcal{U}$. O resultado é apresentado na Tabela (4.10).

Tabela 4.9: Matriz de incidência do modelo do processo de produção de ciclopentanol.

	$\mathcal{U}$																					
	$\dot{C}_A$	m_1	c_1	c_2	W_R	C_A	T_R	$\dot{C}_B$	m_2	c_3	C_B	$\dot{T}_R$	e_1	e_2	f_1	f_2	f_3	T_{Rj}	$\dot{T}_{Rj}$	e_3	e_4	W_{Rj}
$r1$	☐	☐	☐	☐																		
$r2$		☐			☐	☐																
$r3$			☐			☐	☐														☐	
$r4$				☐		☐	☐															
$m5$					☐																	
$d6$	☐					☐																
$r7$			☐					☐	☐	☐												
$r8$					☐				☐		☐											
$r9$							☐			☐	☐											
$m10$											☐											
$d11$								☐			☐											
$r12$												☐	☐	☐	☐	☐	☐					W_{Rj}
$r13$					☐		☐						☐									
$r14$							☐							☐				☐				
$r15$						☐	☐								☐							
$r16$							☐				☐					☐						
$r17$						☐	☐										☐					
$m18$							☐															
$d19$							☐															
$r20$																			☐	☐	☐	
$r21$																		☐		☐		☐
$r22$						☐												☐			☐	
$m23$																			☐			
$m24$																						☐
$m25$						☐																
$d26$																		☐	☐			

Tabela 4.10: Matriz de incidência do modelo com pareamento obtido através do algoritmo de posicionamento para o processo de produção de ciclopentanol.

$\mathcal{U}$		$\dot{C}_A$	m_1	c_1	c_2	W_R	C_A	T_R	$\dot{C}_B$	m_2	c_3	C_B	$\dot{T}_R$	e_1	e_2	f_1	f_2	f_3	T_{Rj}	$\dot{T}_{Rj}$	e_3	e_4	W_{Rj}
r_1		☐	☐	☐	☐																		
r_2			☒			☐																	
r_3				☒			☐	☐															
r_4					☒		☐	☐															
m_5						☒																	
d_6		☒					☐																
r_7				☐					☐	☐	☐												
r_8						☐				☒		☐											
r_9								☐			☒	☐											
m_{10}												☒											
d_{11}									☒			☐											
r_{12}													☐	☐	☐	☐	☐	☐					
r_{13}						☐		☐						☒	☒				☐				
r_{14}								☐							☒	☒							
r_{15}							☐	☐								☒							
r_{16}								☐				☐					☒						
r_{17}							☐	☐										☒					
m_{18}								☒															
d_{19}								☐					☒										
r_{20}																			☐	☐	☐		
r_{21}																			☐		☒		
r_{22}							☐												☐			☒	
m_{23}																				☒			
m_{24}																							☒
m_{25}							☒																
d_{26}																			☐	☒			

4.3.6 Controle Tolerante do Processo de Produção de Ciclopentanol

A implementação da estratégia de controle tolerante para este processo se baseia na correção de falhas em sensores e utilização de uma redundância física, uma válvula ar-para-abrir que atua no processo quando a principal é considerada como emperrada pelo sistema de diagnóstico.

O objetivo do controle tolerante é continuar operando o processo na presença de falhas, no entanto, essas falhas devem ser reconhecidas pelo processo, ou seja, detectadas pelas estatísticas e diagnosticadas corretamente pelo sistema de diagnóstico. As informações resultantes das etapas de detecção e diagnósticos são enviadas para um sistema supervisor, também conhecido como gerenciador de eventos anormais, considerado como cérebro do controle tolerante, que por sua vez baseia-se nessas informações para decidir qual ação deve ser tomada para que o processo retorne a uma condição de operação normal.

O gerenciador de eventos anormais é implementado no simulador e possibilita um ambiente de simulação e correção das falhas inseridas no banco de dados PCA. Para aplicações reais é possível aumentar o banco de dados de falhas mantendo os dados de operação normal quando dados de uma nova falha estão disponíveis. Assim, para detectar falhas em processos, é necessário somente os dados de operação normal. Para o diagnóstico, utiliza-se um modelo PCA para cada conjunto de dados de falhas. Uma análise discriminatória é capaz de decidir qual falha está presente no processo.

Controle tolerante para as Falhas 01 e 04

A Falha 01 representa uma perturbação degrau na concentração da alimentação de reagente ao reator e não requer uma estratégia de controle tolerante específica, visto que o controlador PI projetado é capaz de compensar adequadamente essa mudança no processo. No entanto, a presença deste comportamento deve ser observada pelo sistema de detecção e diagnóstico para que alguma medida possa ser tomada, visto que a corrente de entrada é proveniente de outra unidade. Com isso, a detecção e diagnóstico fornecem informações para uma reconfiguração não implementada pela estratégia de controle tolerante proposta, que seria a investigação de alguma falha no processo anterior ao reator.

Na Falha 04, a válvula ar-para-abrir que manipula a vazão de alimentação de reagente emperra numa posição 30% inferior àquela do estado estacionário. A estratégia de controle tolerante destinada a corrigir esta falha é uma redundância física, como pode ser visto na Figura (4.22), ou seja, uma outra válvula que aguarda uma ação de controle para abrir e assim controlar a quantidade complementar de reagente que deve ser alimentada quando

a válvula principal fornece uma vazão fixa e menor que a desejada em projeto. A Figura (4.33) mostra o comportamento do processo quando o controle tolerante é ativado após a detecção e classificação correta da falha.

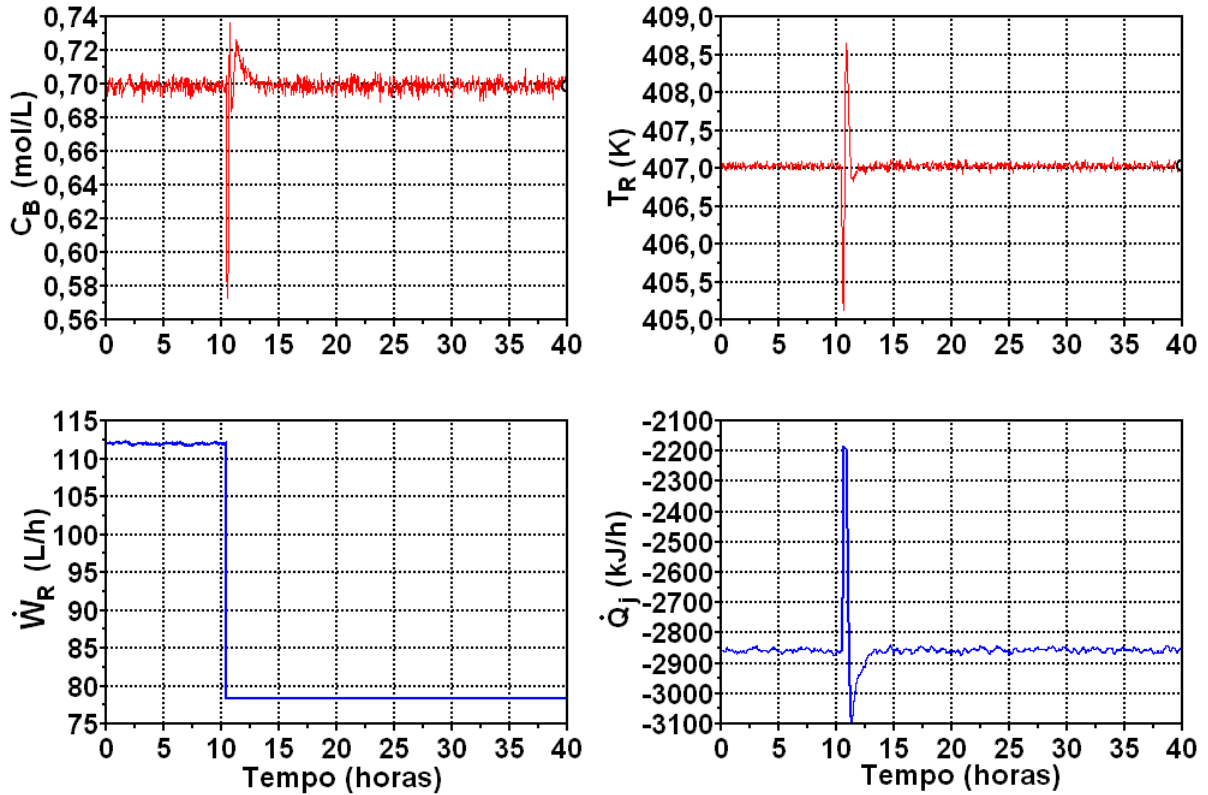


Figura 4.33: Comportamento do processo com controle tolerante para o reator de produção de ciclopentanol.

Nota-se pela Figura (4.33) que após a ocorrência da falha, o sensor de vazão localizado após a válvula com defeito indica sempre uma vazão 30% inferior a do estado estacionário. Além disso, as estatísticas T^2 e Q detectam a falha com dois instantes de amostragem de atraso (ver Tabela (4.12) e Figura (4.34) para informações mais detalhadas sobre o tempo de ocorrência de cada falha e atraso na detecção). Assim, a válvula redundante completa a quantidade necessária para o processo que não pode ser atendida pela válvula principal devido ao emperramento. A Figura (4.35) mostra as ações de controle enviadas para a válvula redundante para complementar a vazão não fornecida ao reator devido ao problema de emperramento.

Correção de falhas em sensores usando análise estrutural

As Falhas 02 e 03 representam falhas em sensores. Como ambos controladores necessitam das informações dos sensores para fechar a malha de controle PI, informações corrompidas prejudicam a operação do processo e não desaparecem até que um novo instrumento seja

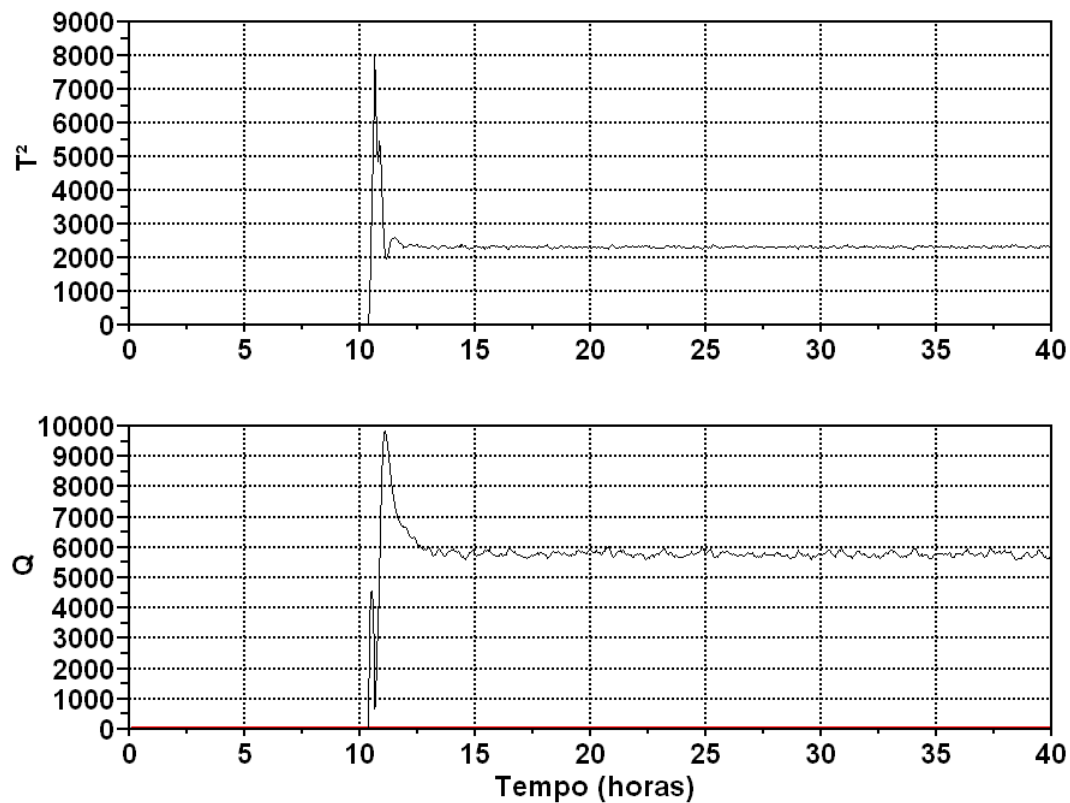


Figura 4.34: Estatísticas T^2 e Q para a Falha 04 no reator de produção de ciclopentanol.

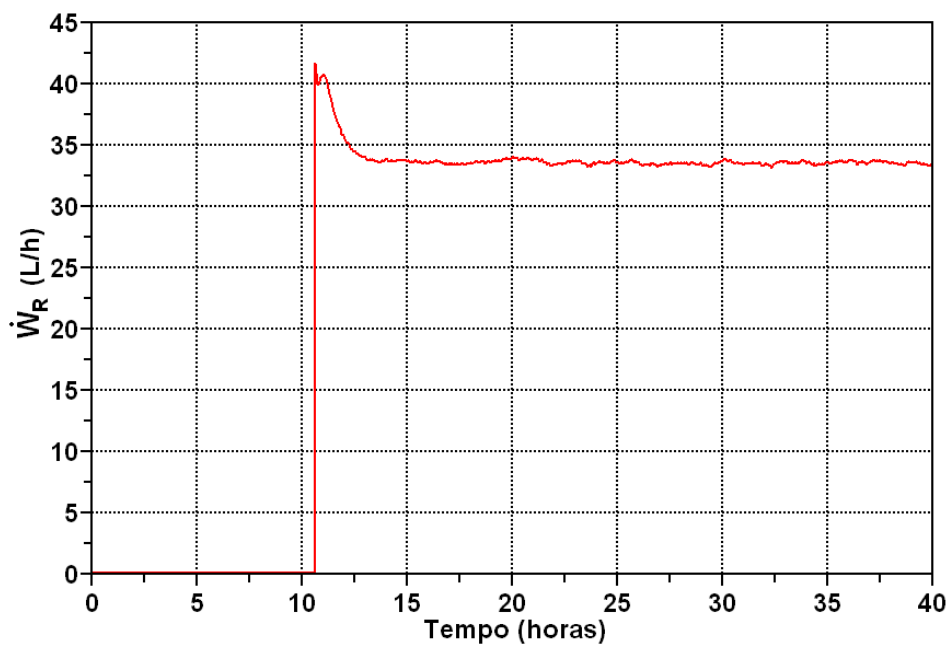


Figura 4.35: Vazão da válvula redundante após ativação do controle tolerante.

instalado. A estratégia de controle tolerante implementada para este caso é a estimativa das medidas com base no modelo não linear disponível e utilizando recursos da análise estrutural.

A Tabela (4.10) resulta da execução de um algoritmo de pareamento sobre a matriz de incidência original (Tabela (4.9)), ou seja, este algoritmo mostra a seqüência de cálculo que deve ser realizada para que todas variáveis desconhecidas sejam monitoradas a partir das variáveis conhecidas. A análise de falhas em sensores utilizando análise estrutural é feita através da utilização do algoritmo de posicionamento, que gera uma nova seqüência de cálculo para as variáveis desconhecidas quando variáveis medidas (conhecidas) se tornam desconhecidas devido a falhas. Assim, para estimar o valor de uma variável medida, basta excluir a restrição associada a esta do algoritmo de pareamento para obter a nova seqüência de cálculo que resulta no cálculo da variável excluída. Por exemplo, se ocorre uma falha com o sensor de temperatura do reator, a restrição $m18$ não é válida e T_R não pode ser determinada utilizando-a, com isso, a temperatura do reator medida não é mais associada com a temperatura real do reator e deve ser calculada por outra restrição e utilizando apenas as outras variáveis medidas.

A Tabela (4.11) resume três condições operacionais e a ordem de cálculo das variáveis desconhecidas: a condição de operação normal e a ordem que as restrições devem ser percorridas para que todas variáveis desconhecidas sejam monitoradas. A Falha 02, em que o sensor de concentração de ciclopentanol apresenta falhas, logo a restrição $m10$ não participa do algoritmo de posicionamento. A falha em que o sensor $m18$ não participa do algoritmo de posicionamento (Falha 03).

A técnica utilizada para projetar as relações de redundância analítica é conhecida como relações de redundância deduzidas e é apresentada no Capítulo 2.

De acordo com os resultados do algoritmo de posicionamento, é possível estimar as medidas de concentração de ciclopentanol (C_B) e temperatura do reator (T_R) pelas posições 19 e 13, respectivamente. Assim, para falhas nesses sensores é possível, após detecção e diagnóstico, percorrer as seqüências de cálculos para estimar os valores das variáveis de interesse, ou seja, dos sensores com problemas. Note que para a estimação da concentração de ciclopentanol, o valor da variável medida (C_{Bm}) é excluído da análise, logo para a Falha 02 consta um x em $m10$ e o valor de C_B é estimado usando apenas as outras variáveis medidas, sendo o mesmo válido para a medida de temperatura do reator, dado pela restrição $m18$.

A Figura (4.36) apresenta os resultados da simulação de uma falha na medição da concentração de ciclopentanol na saída do reator com controle tolerante e a Figura (4.37) os resultados das estatísticas T^2 e Q para esta falha.

Uma característica importante e esperada pela análise estrutural durante a estimação existem ruídos nas demais variáveis é a ampliação dos mesmos devido ao cálculo de derivadas por aproximações de diferenças finitas. Isto pode ser visto na Figura (4.36),

Tabela 4.11: Relações de redundância analíticas deduzidas para o processo de produção de ciclopentanol.

	Normal		Falha 02		Falha 03	
	Variável	Ordem	Variável	Ordem	Variável	Ordem
$r1$						
$r2$	m_1	10°	m_1	8°	m_1	8°
$r3$	c_1	11°	c_1	9°	c_1	14°
$r4$	c_2	12°	c_2	10°	c_2	15°
$m5$	W_R	1°	W_R	1°	W_R	1°
$d6$	$\dot{C}_A$	13°	$\dot{C}_A$	11°	$\dot{C}_A$	9°
$r7$					c_3	16°
$r8$	m_2	14°	m_2	20°	m_2	10°
$r9$	c_3	15°	c_3	21°		
$m10$	C_B	2°	$\mathbf{x}$	$\mathbf{x}$	C_B	2°
$d11$	$\dot{C}_B$	3°	$\dot{C}_B$	22°	$\dot{C}_B$	3°
$r12$			f_2	18°		
$r13$	e_1	16°	e_1	12°	e_1	17°
$r14$	e_2	17°	e_2	13°	e_2	18°
$r15$	f_1	18°	f_1	14°	f_1	19°
$r16$	f_2	19°	C_B	19°	f_2	20°
$r17$	f_3	20°	f_3	15°	f_3	21°
$m18$	T_R	4°	T_R	2°	$\mathbf{x}$	$\mathbf{x}$
$d19$	$\dot{T}_R$	5°	$\dot{T}_R$	3°	$\dot{T}_R$	22°
$r20$					e_4	12°
$r21$	e_3	21°	e_3	16°	e_3	11°
$r22$	e_4	22°	e_4	17°	T_R	13°
$m23$	T_{Rj}	6°	T_{Rj}	4°	T_{Rj}	4°
$m24$	W_{Rj}	7°	W_{Rj}	5°	W_{Rj}	5°
$m25$	C_A	8°	C_A	6°	C_A	6°
$d26$	$\dot{T}_{Rj}$	9°	$\dot{T}_{Rj}$	7°	$\dot{T}_{Rj}$	7°

pois após o tempo 13,6 horas a estimativa da concentração de ciclopentanol oscila com bastante ruído em torno do ponto estacionário. Este fato é esperado e foi observado por Blanke et al. (2006), que propuseram um filtro para reduzir a intensidade do ruído da variável estimada.

A Figura (4.38) mostra o comportamento do processo com controle tolerante e estimação da variável T_R após o tempo 14,65 horas. A Figura (4.39) mostra as estatísticas T^2 e Q para esta falha, e são particularmente diferentes das outras falhas quando o controle tolerante é ativado, pois para os outros casos, após a correção, as estatísticas são novamente levadas para níveis inferiores aos limites propostos. Esta falha em particular, apresenta o mesmo comportamento após a correção, ou seja sempre valores superiores aos limites devido ao fato da medição da vazão que alimenta o reator ocorrer após a válvula e

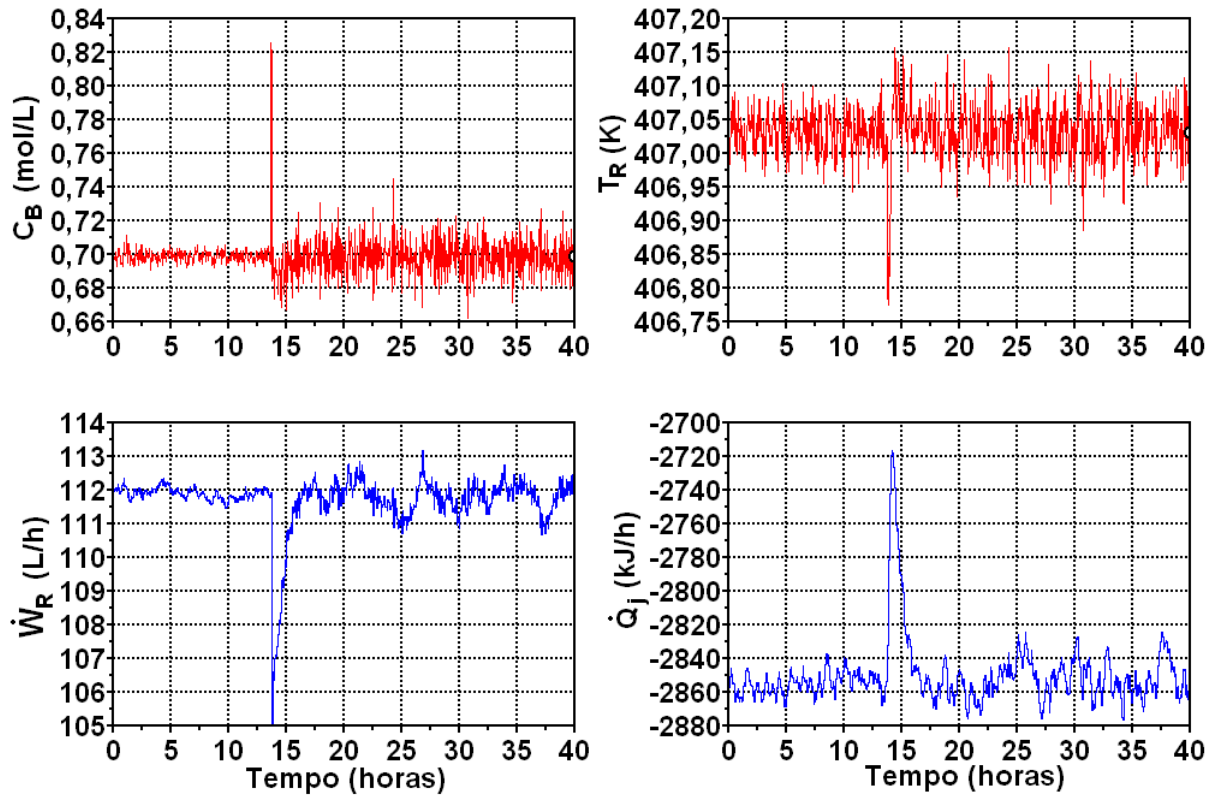


Figura 4.36: Comportamento do processo com controle tolerante para o reator de produção de ciclopentanol (Falha 02).

este valor é sempre o mesmo durante a falha. Logo os algoritmos de detecção e diagnóstico sempre indicam a presença de falhas através das estatísticas T^2 e Q .

4.3.7 Resultados e Discussões

Os resultados das simulações de todas as condições de operação do reator CSTR são resumidos na Tabela (4.10). Nesta tabela estão contidos os resultados das métricas de detecção e diagnóstico na presença de falhas. Para medir a qualidade das estatísticas propostas para a detecção de falhas calcula-se o atraso de detecção (AD) em fração de horas, que representa a quantidade de tempo decorrida desde o instante em que a falha foi introduzida no sistema.

Uma medida que quantifica a qualidade do diagnóstico por discriminante combinado é a taxa de classificações incorretas (TCI), que relaciona o número de classificações incorretas dividida pelo número total de classificações, sempre que aplicado para dados do conjunto de teste, ou seja, independentes do conjunto de treinamento.

A utilização de um processo não linear com duas malhas de controle PI indica que as técnicas de detecção e diagnóstico utilizadas neste trabalho são promissoras. Por se

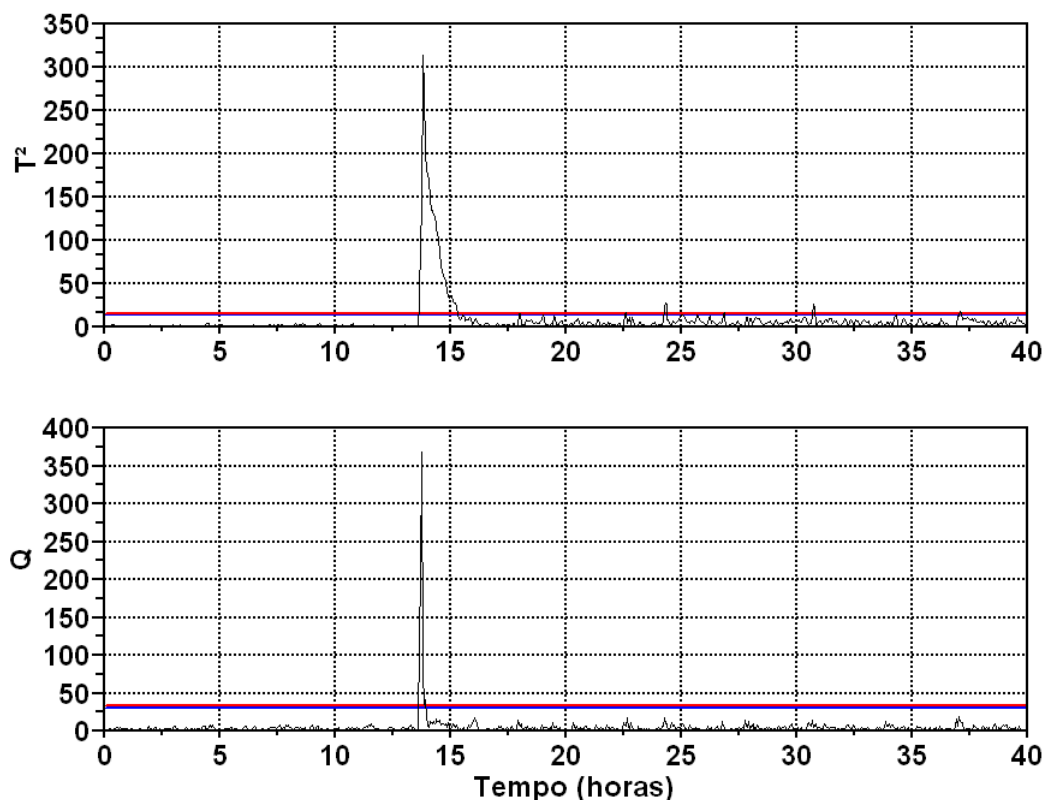


Figura 4.37: Estatísticas T^2 e Q para a falha no sensor de concentração no reator de produção de ciclopentanol (Falha 02).

Tabela 4.12: Resultados das simulações sem o modo de controle tolerante ativo. 40 horas de simulação, tempo de amostragem de 3 minutos.

Tipo de Simulação	Tempo (horas)	T^2	Q	TCI	AD	
					T^2	Q
Falha 01	1,45	1,55	1,55	1,427%	0,10	0,10
Falha 02	13,6	13,7	13,65	4,924%	0,10	0,05
Falha 03	14,65	14,7	14,7	0,000%	0,05	0,05
Falha 04	10,35	10,4	10,4	0,169%	0,05	0,05

tratarem de técnicas estatísticas e requerem a utilização de dados para treinamento, sejam eles de operação normal e com falhas, foi necessário gerar uma base de dados e realizar todos os cálculos de treinamento de forma fora de linha. Isto faz com que o esforço computacional durante a simulação em linha seja sensivelmente reduzido, possibilitando assim uma implementação industrial.

As medidas de qualidade das estatísticas de detecção e diagnóstico mostram que apenas a Falha 02 apresentou uma taxa de classificações incorretas elevada, no entanto, isto não interferiu na decisão do sistema em acionar o controle tolerante, pois observou-se que as detecções incorretas ocorreram em instantes de amostragens distantes do tempo exato de ocorrência da falha. Com respeito ao atraso de detecção, ambas estatísticas

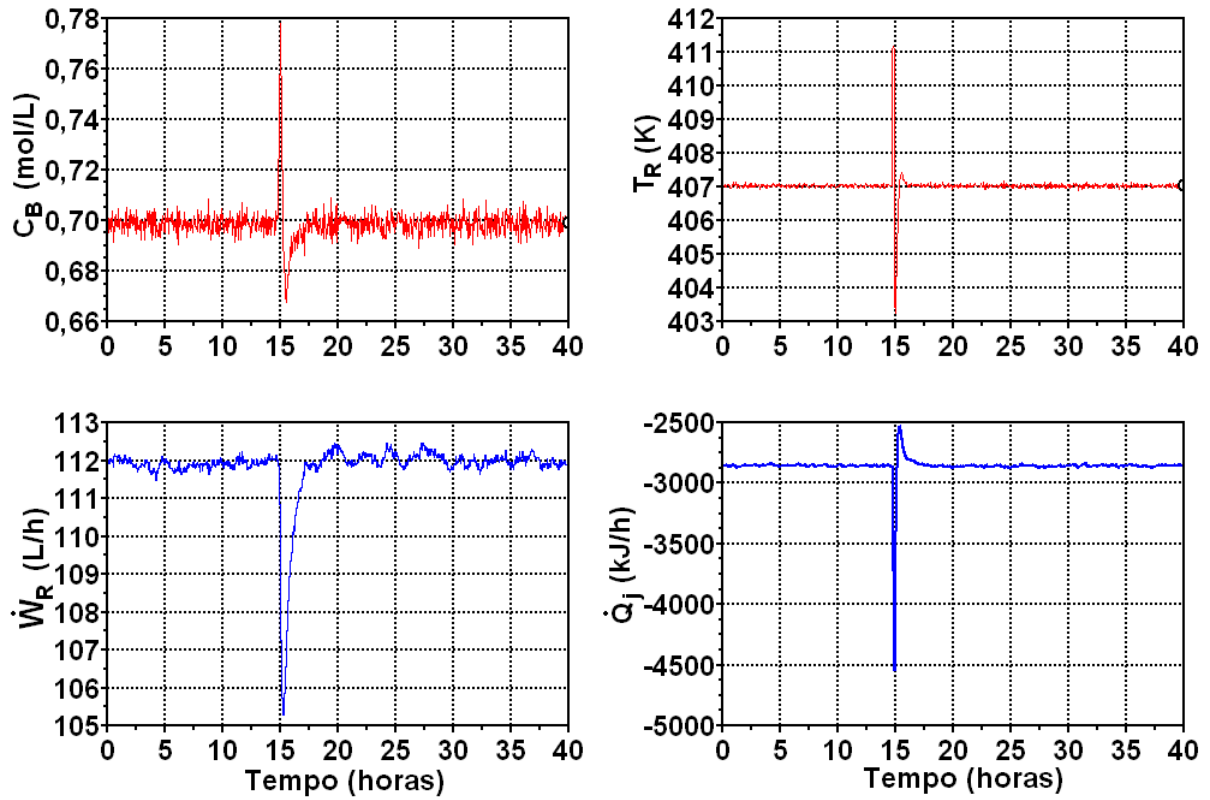


Figura 4.38: Comportamento do processo com controle tolerante para o reator de produção de ciclopentanol (Falha 03).

apresentaram bons resultados, máximo 2 instantes de amostragem após o aparecimento da falha.

As estratégias de controle tolerante para as Falhas 02, 03 e 04 foram capazes de conduzir novamente o processo para a operação normal, no entanto, observa-se a necessidade da implementação de um filtro para eliminação de ruídos da variável estimada C_B .

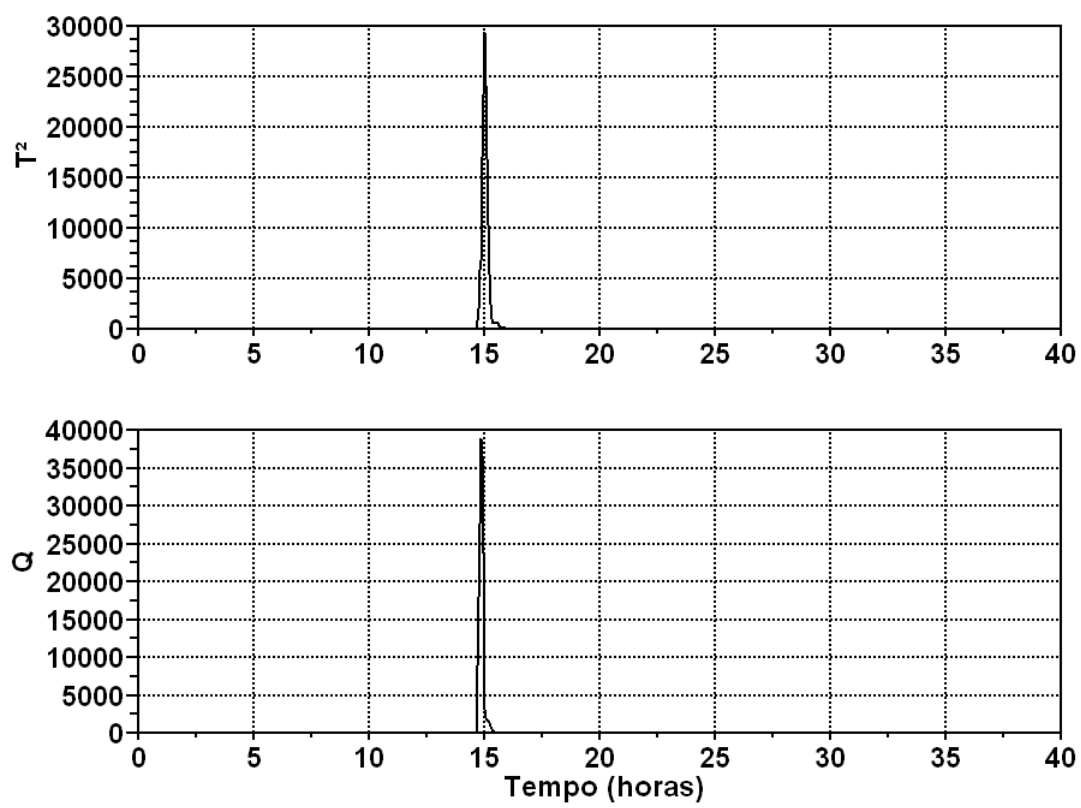


Figura 4.39: Estatísticas T^2 e Q para a falha no sensor de temperatura do reator de produção de ciclopentanol (Falha 03).

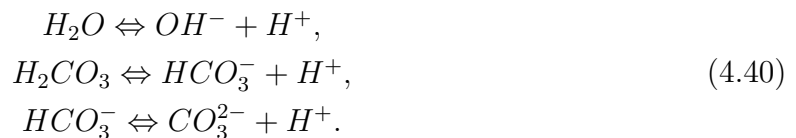
4.4 O Reator de Neutralização

O processo de neutralização tem um destacado papel no âmbito da biotecnologia. A elaboração de modelos para sistemas de neutralização e, especificamente, para o controle de pH tem se tornado mais relevante nesta área devido à necessidade de melhorar a qualidade do produto ou de otimizar o processo de produção. Muito embora os fundamentos físico-químicos e a natureza eletroquímica tenham sido bem estabelecidos, o controle de tais processos ainda não está bem resolvido do ponto de vista industrial. Como o problema de sistemas de neutralização não é simplesmente um problema de controle, haja vista que envolve equilíbrio químico e cinética, o desenvolvimento de modelos dinâmicos para esses sistemas é um fator determinante na elaboração de estratégias de controle eficientes (GUSTAFSSON et al., 1995).

Este exemplo utiliza um controlador preditivo baseado em modelo sucessivamente linearizado para controlar o processo de neutralização. Com este tipo de controlador torna-se possível uma aproximação precisa do modelo não linear por modelos sucessivamente linearizados. Assim, utiliza-se a proposta de controle tolerante mostrada na Figura (3.2) do Capítulo 3 para tornar este processo tolerante a falhas em atuadores.

4.4.1 Modelo Físico

O modelo físico baseia-se na hipótese de mistura perfeita, densidade constante e solubilidade completa dos ions envolvidos. As reações químicas envolvidas no processo de neutralização são:



Segundo Gustaffson e Waller (1983), dois invariantes da reação são definidos para a i -ésima corrente:

$$\begin{aligned} W_{ai}(t) &= [H^+]_i(t) - [OH^-]_i(t) - [HCO_3^-]_i(t) - 2[CO_3^{2-}]_i(t) \\ W_{bi}(t) &= [H_2CO_3]_i(t) + [HCO_3^-]_i(t) + [CO_3^{2-}]_i(t) \end{aligned} \quad (4.41)$$

As quantidades $W_{ai}(t)$ e $W_{bi}(t)$ são chamadas de invariantes da reação pois suas concentrações não são afetadas ao longo da reação, apesar das concentrações das espécies que as compõem variarem. Para uma corrente de alimentação de ácido, por exemplo, a concentração do íon hidrogênio somente será alterada se nesta corrente ocorrem perturbações, quando a corrente é proveniente de um outro processo, por exemplo. Considera-se ainda

que as reações são rápidas o bastante para permitir que o sistema seja considerado em equilíbrio. Logo, as relações de equilíbrio podem ser usadas para determinar a concentração do íon hidrogênio através da concentração dos invariantes das reações. As concentrações de equilíbrio são dadas por:

$$K_{a1} = \frac{[HCO_3^-](t)[H^+](t)}{[H_2CO_3](t)}, \quad (4.42)$$

$$K_{a2} = \frac{[CO_3^{2-}](t)[H^+](t)}{[HCO_3^-](t)}, \quad (4.43)$$

$$K_w = [H^+](t)[OH^-](t). \quad (4.44)$$

De acordo com as Equações (4.42) a (4.44), mostra-se que o que ocorre realmente é a variação das concentrações das espécies com o tempo, por exemplo, dentro do reator a quantidade de íons varia continuamente visto que as quantidades de ácidos e base alimentadas ao reator variam a cada instante para que o objetivo de controle do pH seja alcançado. Apesar disso, as constantes de dissociação mostram independência dessa variação, pois as outras espécies envolvidas também variam proporcionalmente.

Utilizando a Equação (4.41) para calcular os invariantes da reação dentro do reator, verifica-se a dependência com o tempo de todas concentrações da Equação (4.41), que é reescrita na seguinte forma:

$$\begin{aligned} W_a(t) &= [H^+](t) - [OH^-](t) - [HCO_3^-](t) - 2[CO_3^{2-}](t), \\ W_b(t) &= [H_2CO_3](t) + [HCO_3^-](t) + [CO_3^{2-}](t). \end{aligned} \quad (4.45)$$

Substituindo em $W_b(t)$ a Equação (4.42) e (4.43) tem-se:

$$W_b(t) = \left(\frac{[HCO_3^-](t)[H^+](t)}{K_{a1}} \right) + \left(\frac{[CO_3^{2-}](t)[H^+](t)}{K_{a2}} \right) + [CO_3^{2-}](t). \quad (4.46)$$

Substituindo a Equação (4.43) na Equação (4.46) chega-se a:

$$W_b(t) = \left(\frac{[CO_3^{2-}](t)[H^+]^2(t)}{K_{a1}K_{a2}} \right) + \left(\frac{[CO_3^{2-}](t)[H^+](t)}{K_{a2}} \right) + [CO_3^{2-}](t). \quad (4.47)$$

Isolando $[CO_3^{2-}](t)$ na Equação (4.47):

$$W_b(t) = [CO_3^{2-}](t) \left(\frac{[H^+]^2(t)}{K_{a1}K_{a2}} + \frac{[H^+](t)}{K_{a2}} + 1 \right). \quad (4.48)$$

Utilizando $W_b(t)$ da Equação (4.45) e substituindo as Equações (4.42) e (4.43), mas evidenciando agora o termo $[HCO_3^-](t)$ tem-se:

$$W_b(t) = \left(\frac{[HCO_3^-](t)[H^+](t)}{K_{a1}} \right) + [HCO_3^-](t) + \left(\frac{K_{a2}[HCO_3^-](t)}{[H^+](t)} \right), \quad (4.49)$$

ou ainda,

$$W_b(t) = [HCO_3^-](t) \left(\frac{[H^+](t)}{K_{a1}} + 1 + \frac{K_{a2}}{[H^+](t)} \right). \quad (4.50)$$

Com $W_a(t)$ da Equação (4.45) pode-se substituir as Equações (4.44), (4.48) e (4.50) para obter uma expressão que relaciona $W_a(t)$, $[H^+](t)$ e $W_b(t)$:

$$W_a(t) = [H^+](t) - \left(\frac{K_w}{[H^+](t)} \right) - \left(\frac{\frac{W_b(t)}{[H^+](t)}}{\frac{K_{a1}K_{a2}}{[H^+](t)} + \frac{[H^+](t)}{K_{a2}} + 1} \right) - \left(\frac{\frac{2W_b(t)}{[H^+](t)}}{\frac{[H^+](t)}{K_{a1}} + 1 + \frac{K_{a2}}{[H^+](t)}} \right). \quad (4.51)$$

A Equação (4.51) pode ser simplificada para uma forma mais compacta:

$$W_a(t) = [H^+](t) - \left(\frac{K_w}{[H^+](t)} \right) - W_b(t) \left(\frac{K_{a1}[H^+](t) + 2K_{a1}K_{a2}}{[H^+]^2(t) + K_{a1}[H^+](t) + K_{a1}K_{a2}} \right). \quad (4.52)$$

4.4.2 Modelagem do Reator de Neutralização

A Figura (4.40) mostra de forma esquemática o reator de neutralização. O sistema possuiu três correntes de entrada, uma de ácido (HNO_3), uma de base ($NaOH$) e uma de solução tampão ($NaHCO_3$). As vazão das correntes Q_1 e Q_3 são as manipuladas e da corrente Q_2 é considerada como perturbação. A corrente de saída Q_4 possui uma válvula com determinada resistência.

Dessa forma, o nível de produto dentro do reator possui uma dinâmica determinada, devido ao fato de existir uma restrição na corrente de saída.

$$\frac{dh}{dt} = \frac{Q_1 + Q_2 + Q_3 - Q_4}{Ah}, \quad (4.53)$$

com $Q_4 = c_v h$, em que c_v representa a constante da válvula, h o nível de produto dentro do reator e A a área da seção transversal do reator.

O modelo dinâmico do processo de neutralização consiste somente dos balanços de moles dos invariantes reacionais:

$$\frac{dW_a}{dt} = \frac{Q_1 W_{a1} + Q_2 W_{a2} + Q_3 W_{a3} - Q_4 W_a}{Ah}, \quad (4.54)$$

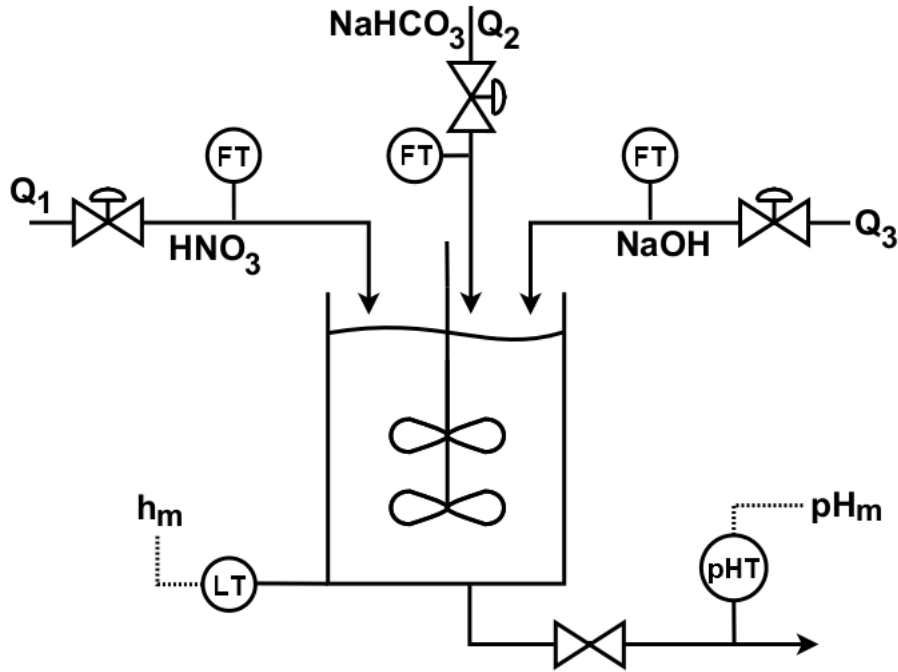


Figura 4.40: Desenho esquemático do reator de neutralização.

e

$$\frac{dW_b}{dt} = \frac{Q_1 W_{b1} + Q_2 W_{b2} + Q_3 W_{b3} - Q_4 W_b}{Ah} \quad (4.55)$$

em que V_R é o volume do reator, Q_1 a vazão de ácido, Q_2 a vazão de solução tampão, Q_3 a vazão de base, W_a é o invariante do ácido e W_b é o invariante da base. As Equações (4.54) e (4.55) permitem calcular as concentrações dos invariantes como uma função do tempo, supondo que a vazão de ácido, base e tampão e as concentrações iniciais sejam conhecidas. Uma equação que representa a variação do pH com o tempo é obtida através da Equação (4.52). Sabendo que:

$$[H^+](t) = 10^{-pH(t)}. \quad (4.56)$$

Substituindo (4.56) em (4.52):

$$W_a(t) = 10^{-pH(t)} - \left(\frac{K_w}{10^{-pH(t)}} \right) - W_b(t) \left(\frac{K_{a1} 10^{-pH(t)} + 2K_{a1} K_{a2}}{10^{-2pH(t)} + K_{a1} 10^{-pH(t)} + K_{a1} K_{a2}} \right). \quad (4.57)$$

Diferenciando a Equação (4.57) em relação ao tempo obtém-se uma equação que representa a variação do pH com o tempo. Para simplificar a visualização desta expressão os símbolos α e β representam denotam as seguintes expressões. Além disso, oculta-se a dependência com o tempo com o mesmo motivo de simplificação da apresentação.

$$\alpha = 10^{-2pH} + K_{a1} 10^{-pH} + K_{a1} K_{a2},$$

e

$$\beta = K_{a1}10^{-pH} + 2K_{a1}K_{a2}.$$

Dessa forma, tem-se:

$$\frac{dpH}{dt} = \frac{-\frac{dW_a}{dt} - \frac{\beta}{\alpha} \frac{dW_b}{dt}}{\ln(10) \left(10^{-pH} + \frac{K_w}{10^{-pH}} - \frac{W_b K_{a1} 10^{-pH}}{\alpha} - \frac{\beta W_b (-2 \cdot 10^{-2pH} - K_{a1} 10^{-pH})}{\alpha^2} \right)} \quad (4.58)$$

A Equação (4.58) representa a variação do pH com o tempo e juntamente com as Equações (4.53), (4.54) e (4.55) constituem o modelo do reator de neutralização.

Um mapeamento das variáveis de entrada para as de saída é capaz de revelar o elevado grau de não linearidade do modelo. A Figura (4.41) mostra como o pH se comporta quando diferentes combinações entre as variáveis de entrada (vazão de ácido e base) variam. Nesta figura, considera-se que a vazão de solução tampão é nula. No entanto, para as simulações deste exemplo, utiliza-se uma pequena quantidade de solução tampão para que o modelo processo não apresente ganho infinito em algumas regiões.

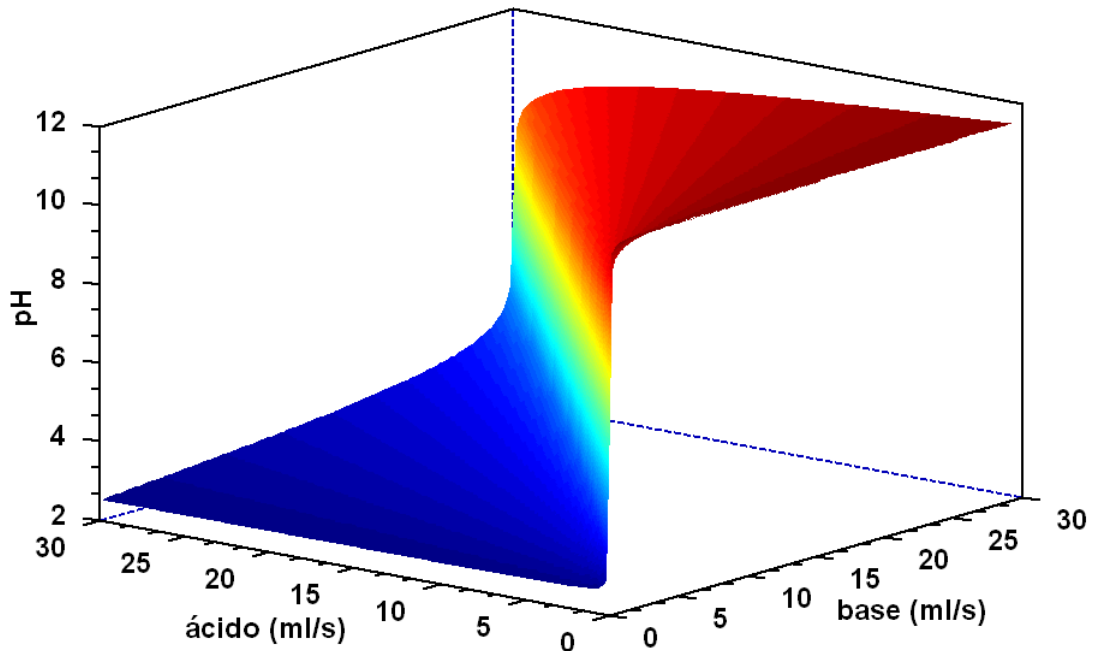


Figura 4.41: Mapeamento entradas (ácido e base) saída (pH) para o processo de neutralização. Vazão de solução tampão nula.

4.4.3 Controle do Processo

O controlador implementado para este processo é apresentado no Apêndice B. Utiliza-se a vazão de solução tampão como uma possível entrada para controlar o processo. Em operação normal (manipuladas: vazão de ácido e vazão de base) esta entrada é considerada como uma perturbação e possui valor de 0,01 ml/s e comumente não é capaz de receber atuação. A Figura (4.42) mostra o controle na região de pH igual a 7 utilizando a vazão de ácido e a de base como variáveis manipuladas.

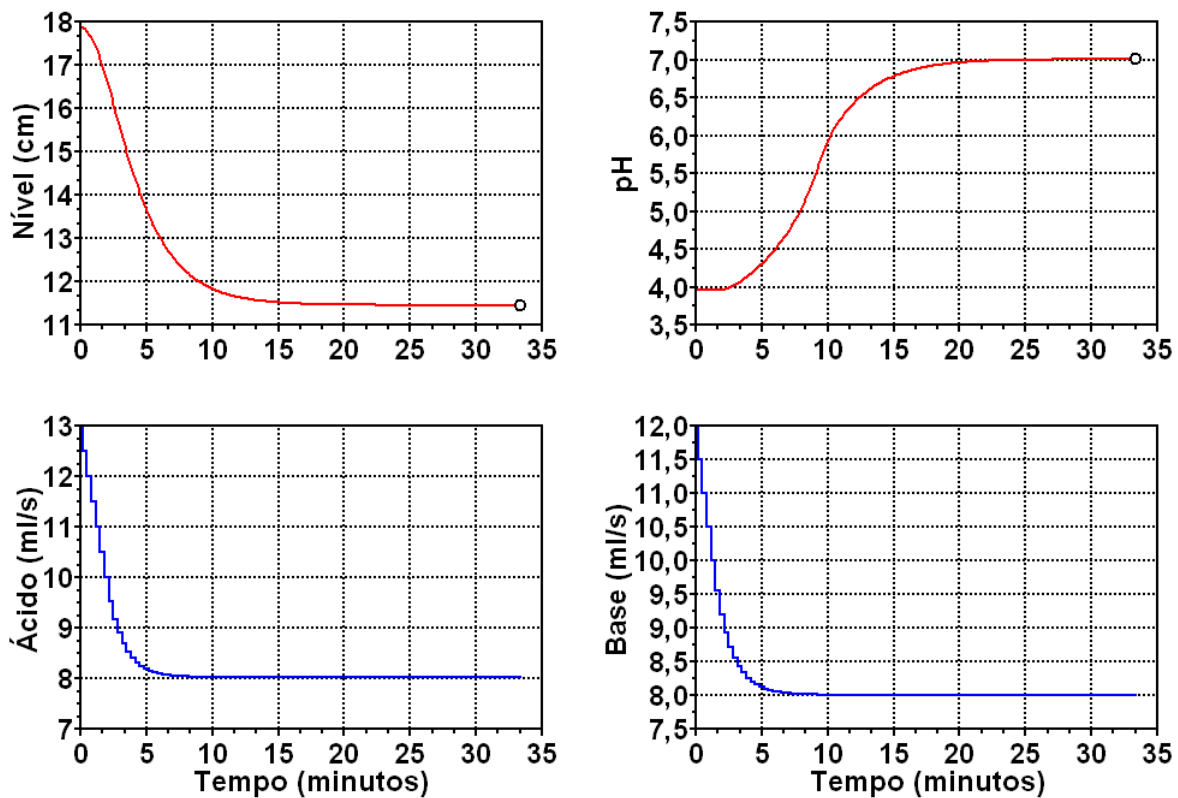


Figura 4.42: Controle do processo de neutralização próximo à região de pH=7.

De outra forma, é possível controlar o processo manipulando a vazão de solução tampão quando a vazão de base permanece constante e passa então a ser considerada como uma perturbação ($Q_3=5$ ml/s).

Assim, a utilização da vazão de solução tampão como variável manipulada para controlar o nível e o pH no reator de neutralização resultou na resposta mostrada pela Figura (4.44).

Essas explorações mostram que é possível controlar o processo com pelo menos duas variáveis de entrada, no entanto, é necessário a instalação de um instrumento alternativo para manipulação. Assim, quando o propósito é tornar um sistema tolerante a falhas, a instalação de sensores ou atuadores extras é essencial e deve ser considerada no projeto.

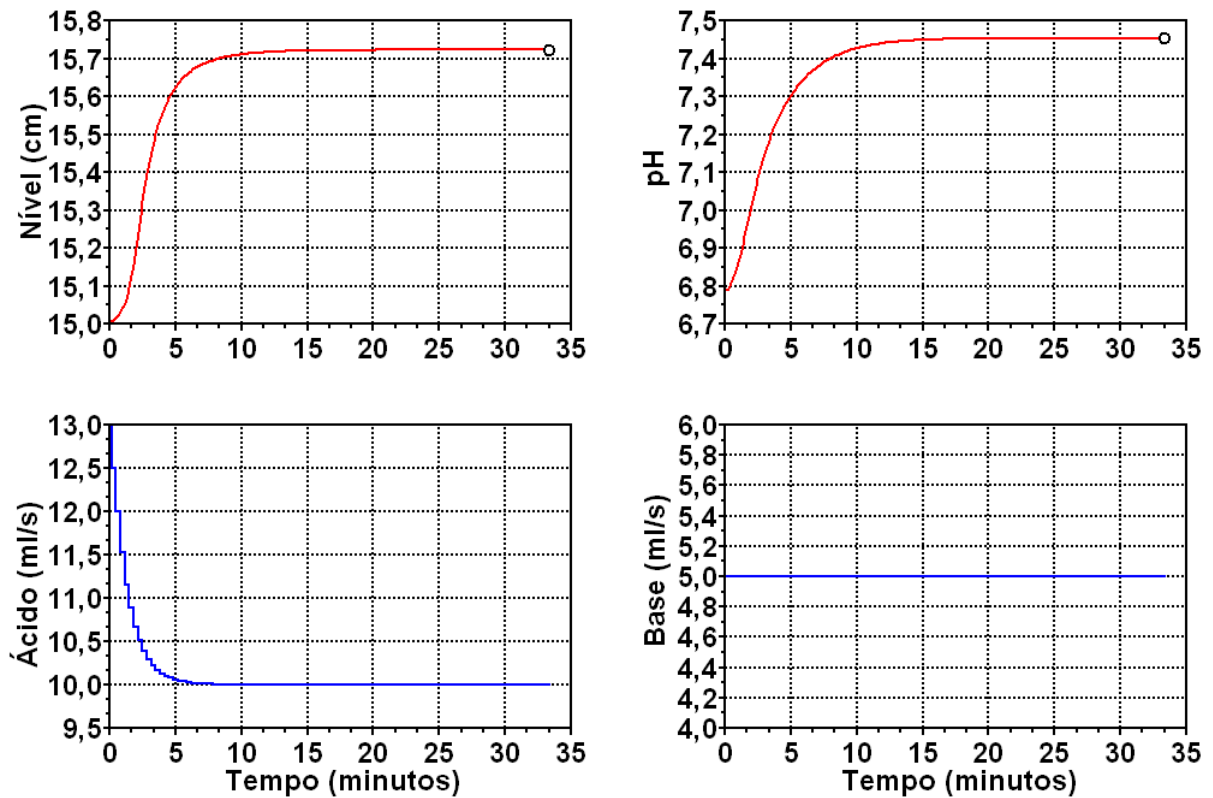


Figura 4.43: Controle do processo de neutralização utilizando como variáveis manipuladas a vazão de ácido e de solução tampão (vazão de base constante e igual a 5 ml/s).

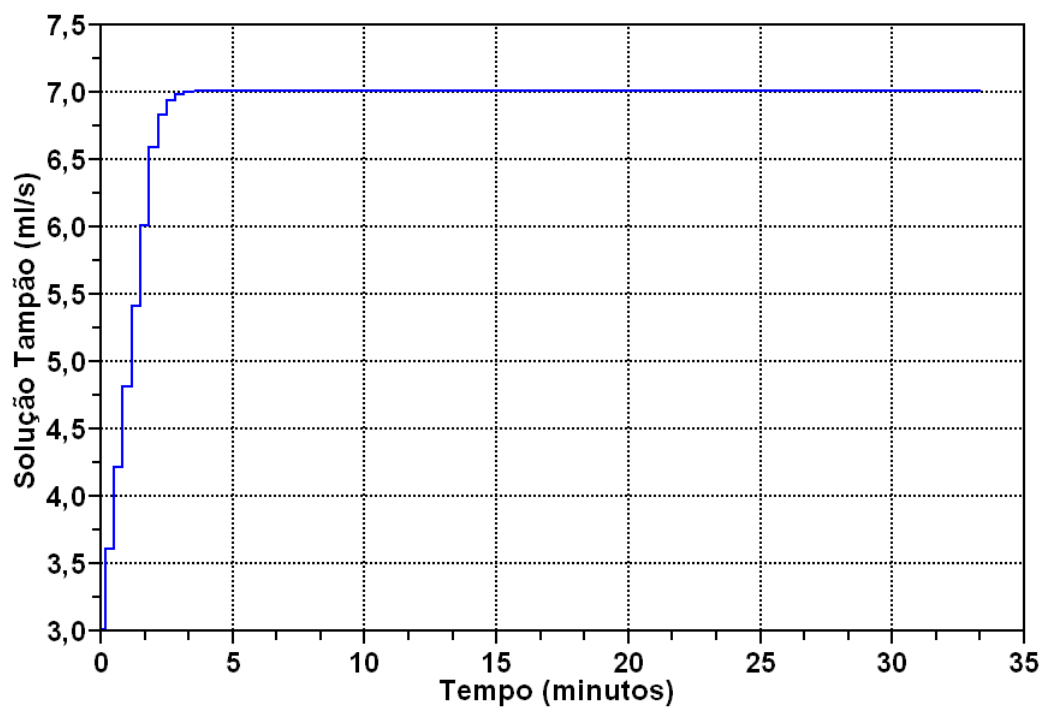


Figura 4.44: Vazão de solução tampão como variável manipulada em um esquema em que vazão de base é constante e igual a 5 ml/s.

4.4.4 Falhas no Reator de Neutralização

Durante a simulação do reator de neutralização alguns tipos de falhas são inseridos. Equipamentos eletrônicos estão sujeitos a defeitos, como é o caso de um sensor. No caso de equipamentos que utilizam uma comunicação entre uma interface eletrônica e uma física, como as válvulas com atuadores, pode ocorrer desgaste mecânico destes atuadores fazendo com que não haja mais correspondência entre o sinal enviado para o movimento do atuador e o movimento realmente realizado. Este fenômeno é comum nas operações industriais e é conhecido como emperramento ou agarramento de válvulas.

Para este processo utiliza-se apenas falhas em atuadores, ou seja, ocorre uma não correspondência entre o sinal enviado para o movimento do atuador e o movimento realizado gera perturbações ao processo, dificultando o controle e aumentando de maneira significativa a sua variabilidade. Para este exemplo, simula-se uma falha no atuador responsável pela alimentação de base, que partir de um determinado instante fornece uma vazão fixa.

Falha na válvula: alimentação de base

A simulação de uma falha na válvula que alimenta solução com pH básico é mostrada na Figura (4.45). Para este exemplo, a vazão de base torna-se constante e igual a 7 ml/s a partir do tempo 18 min. Nota-se que o processo se desvia dos pontos de operação desejados ($h^{sp} = 12,86$ cm e $pH^{sp} = 3,50$). Além disso, o controlador tenta compensar o erro gerado pela diminuição da vazão de base através da diminuição da vazão de ácido. No entanto, o processo não é capaz de retornar para o ponto de operação desejado.

Para detectar este tipo de falha implementa-se uma estratégia de verificação entre o sinal calculado pelo controlador preditivo para os atuadores e o sinal medido através de instrumentos medidores localizados imediatamente após as válvulas de alimentação de ácido e base. A Figura (4.46) mostra que a partir do tempo de falha de 18 min a correspondência entre o sinal calculado pelo controlador e o sinal medido pelo sensor de vazão não é mais verificada e a falha pode ser detectada e diagnosticada.

Controle tolerante

A estratégia de controle tolerante proposta para este tipo de falha é apresentada em detalhes no Capítulo 3 e se baseia na atualização do modelo sucessivamente linearizado devido ao aparecimento de falhas bem como a dos *set-points*, que se alteram para buscar os objetivos de controle previamente especificados (neste caso $h^{sp} = 12,86$ cm e $pH^{sp} =$

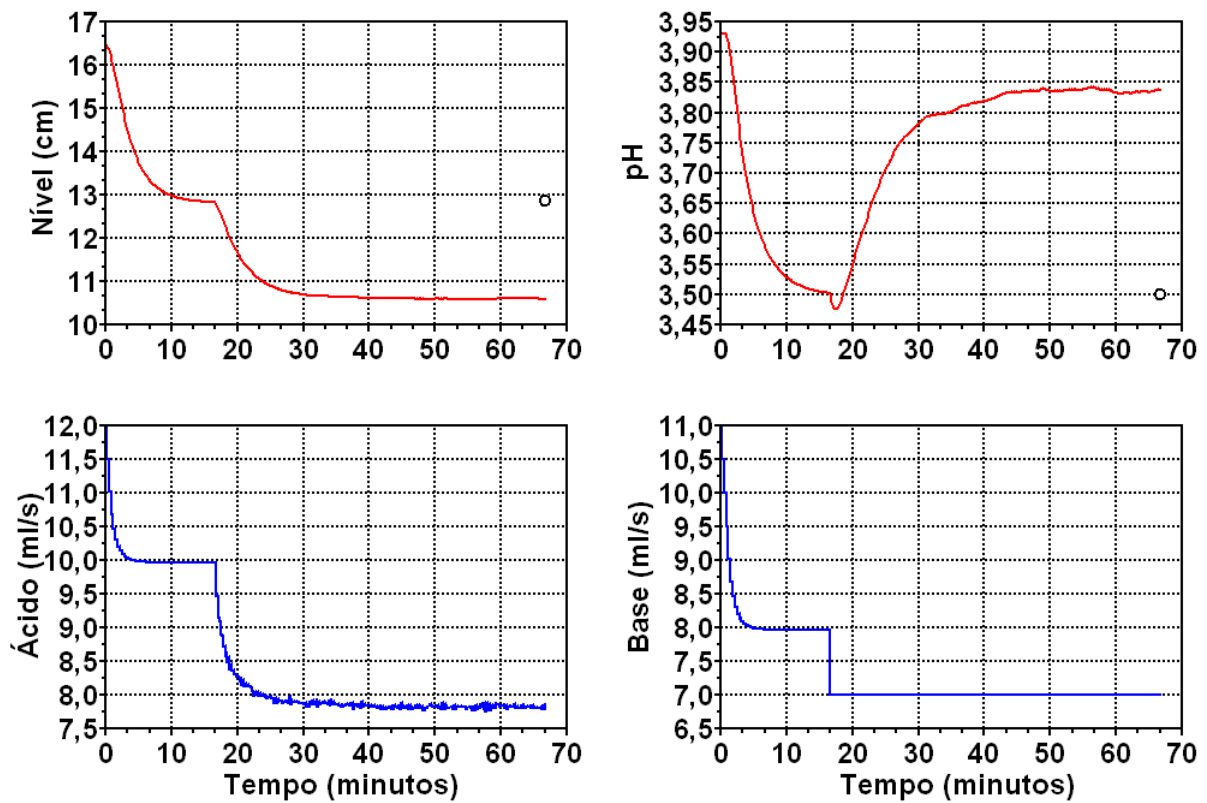


Figura 4.45: Simulação de uma falha no atuador que alimenta base ao reator de neutralização.

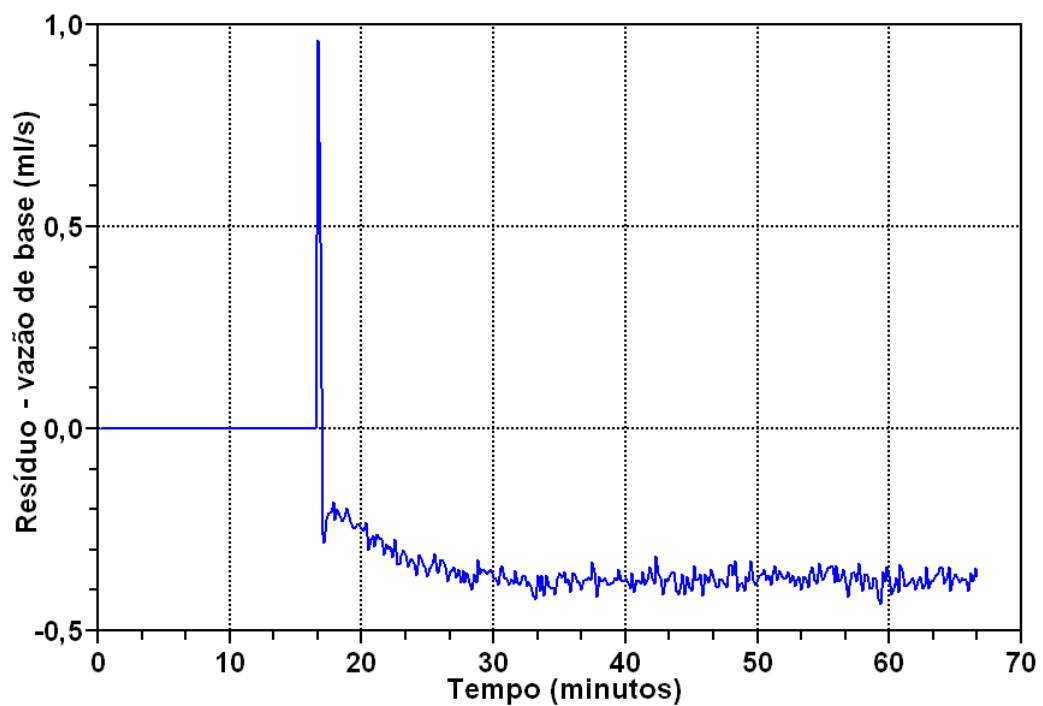


Figura 4.46: Resíduos entre o sinal de entrada calculado pelo MPC e o sinal lido através do sensor de vazão de base (ml/s).

3, 50). Assim, verifica-se a troca de variáveis manipuladas.

A Figura (4.47) mostra o comportamento do processo com controle tolerante. A falha é detectada através da geração de resíduos e para este caso, ocorre uma atualização da matriz dos atuadores e dos pontos de operação. Para uma vazão constante de base a partir do tempo 33 min, que neste caso é um problema do atuador que encontra-se travado nesta posição. O controle tolerante deve continuar a operação controlada do processo. Para isto, define-se como objetivo principal a manutenção do pH na saída do tanque e permite-se que o nível sofra alterações para que o efeito da falha seja acomodado. Com isso, a variável vazão de base não sofre atuação enquanto que a vazão de solução tampão, que antes da operação com falha era considerada uma perturbação passa a ser uma variável manipulada para manter o objetivo de controle do pH no tanque.

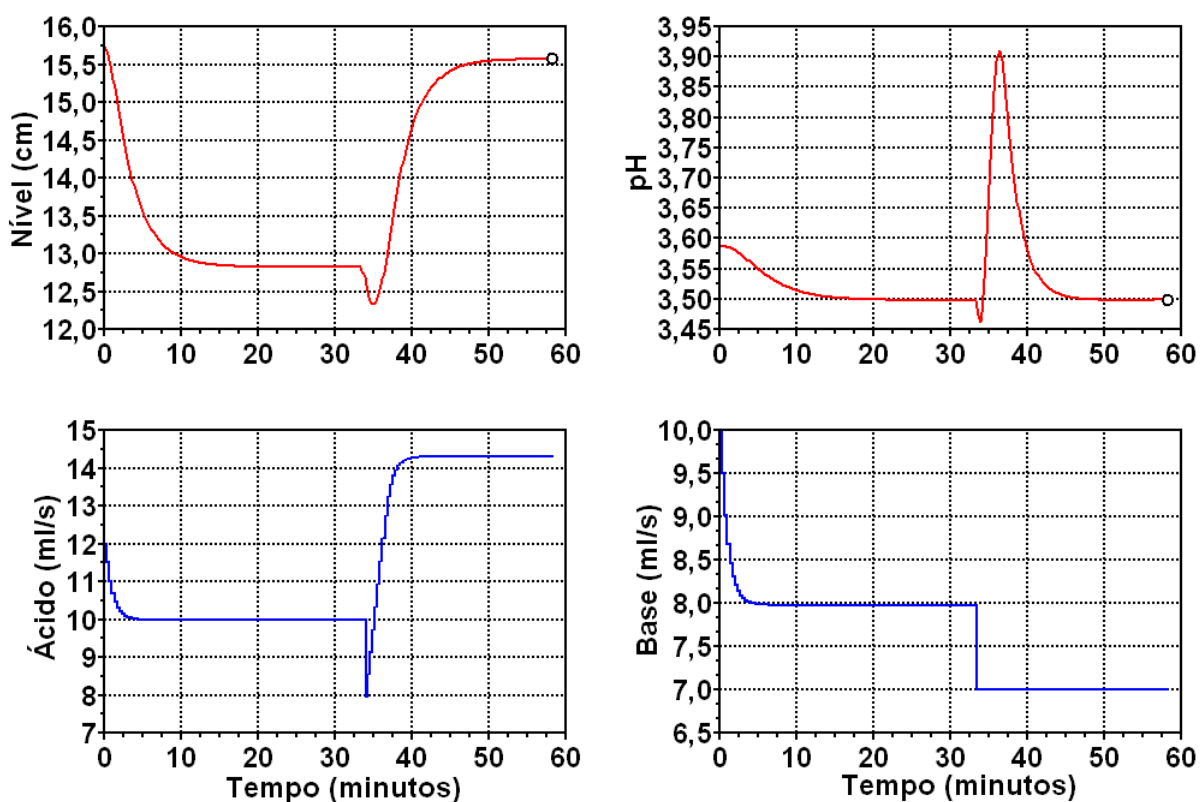


Figura 4.47: Controle tolerante para o reator de neutralização com falha na válvula de alimentação de base.

O processo de controle tolerante para este processo é melhor exemplificado através da Figura (4.48), que mostra como a atualização do modelo é feita. A linearização sucessiva é sempre realizada em torno de todas as variáveis de entrada. Para cada operação controlada escolhe-se as variáveis que deverão ser manipuladas. As variáveis que não serão manipuladas devem receber valor zero ao final de cada linearização. Em seguida um procedimento de verificação de falhas é feito para analisar a possibilidade de atualização dos valores de *set-point*, restrições e principalmente do modelo.

Utilizando o esquema proposta pela Figura (4.49) para este processo, é possível utili-

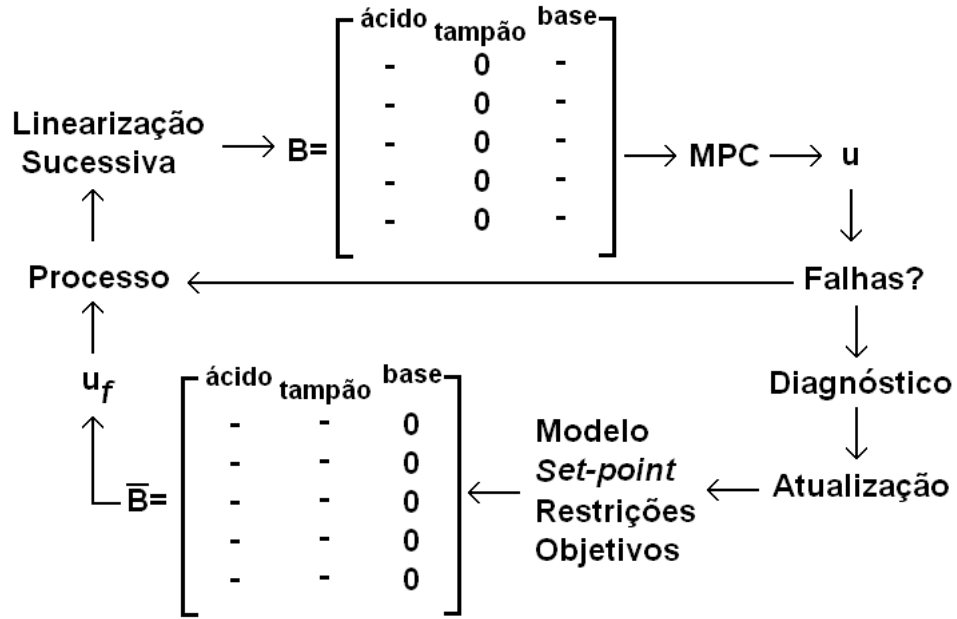


Figura 4.48: Esquema de controle tolerante proposto para acomodação de falhas em atuadores para o reator de neutralização.

zar como variável manipulada alternativa a vazão de solução tampão, que em operações normais é considerada como uma perturbação, passa a ser utilizada como variável manipulada e a vazão de base que se torna constante devido à fixação no valor de 7 ml/s devido a uma falha no respectivo atuador passa a ser considerada como uma perturbação para o processo.

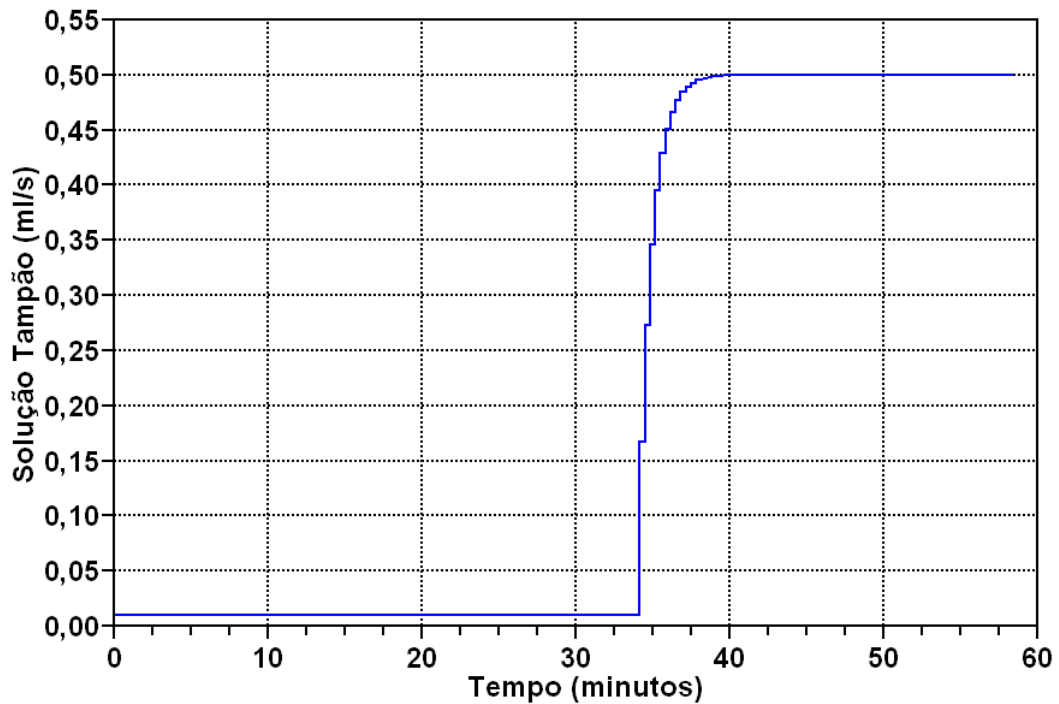


Figura 4.49: Manipulação da vazão de solução tampão no esquema de controle tolerante a falhas na válvula de alimentação de base.

A proposta de controle tolerante para este processo é eficiente e consegue recuperar o processo e traçar novos objetivos, que implementados em linha, conseguem levar o processo a operar no ponto desejado. A exploração de falhas nos outros atuadores (ácido e solução tampão) também é possível.

Mudança de objetivo

Assim como mostrado anteriormente, a mudança de objetivo em linha é capaz de recuperar operações controladas que são afetadas por falhas. Objetiva-se demonstrar que a alteração de objetivos totais é possível e deve ser avaliada. A Figura (4.50) mostra como a alteração total dos objetivos de controle ($h^{sp} = 15,73$ cm e $pH^{sp} = 7,45$) é feita com sucesso.

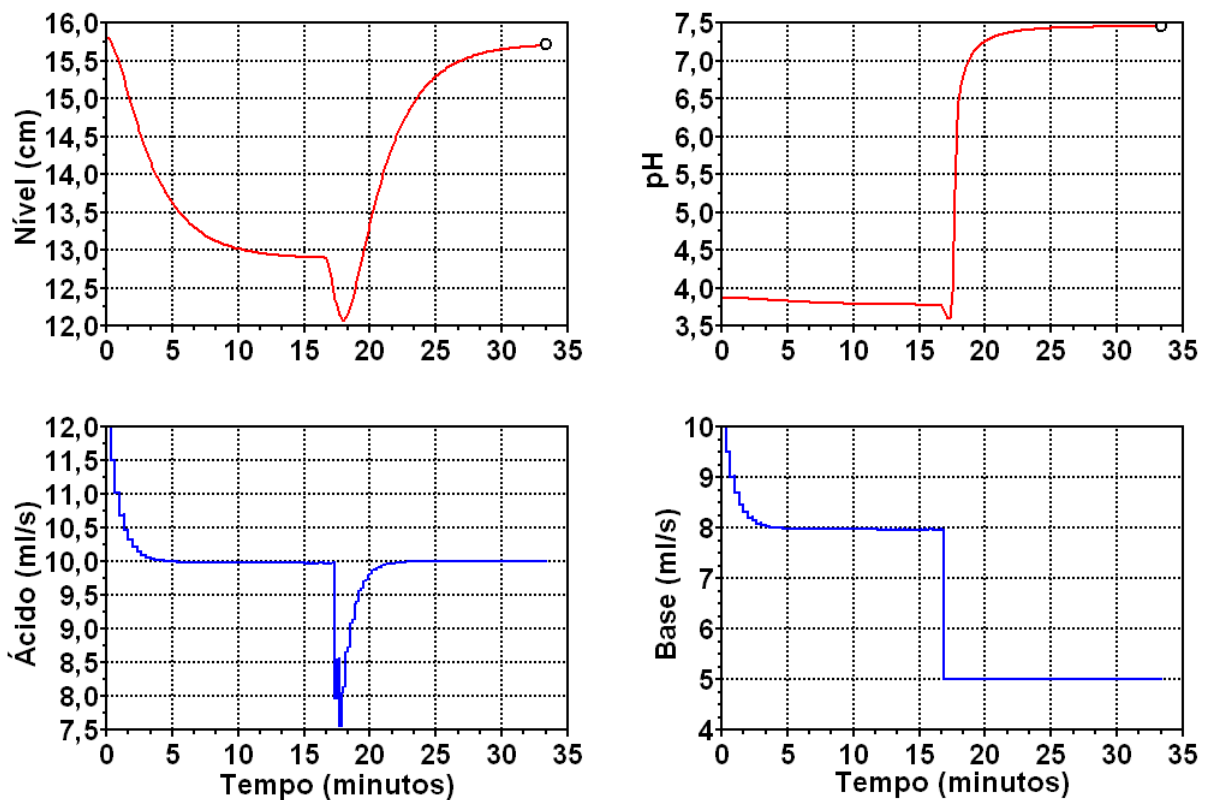


Figura 4.50: Controle tolerante a falhas com mudança completa dos objetivos de controle (nível e pH).

Para que isso seja possível ocorre uma maior utilização de solução tampão, assim como mostrado pela Figura (4.51).

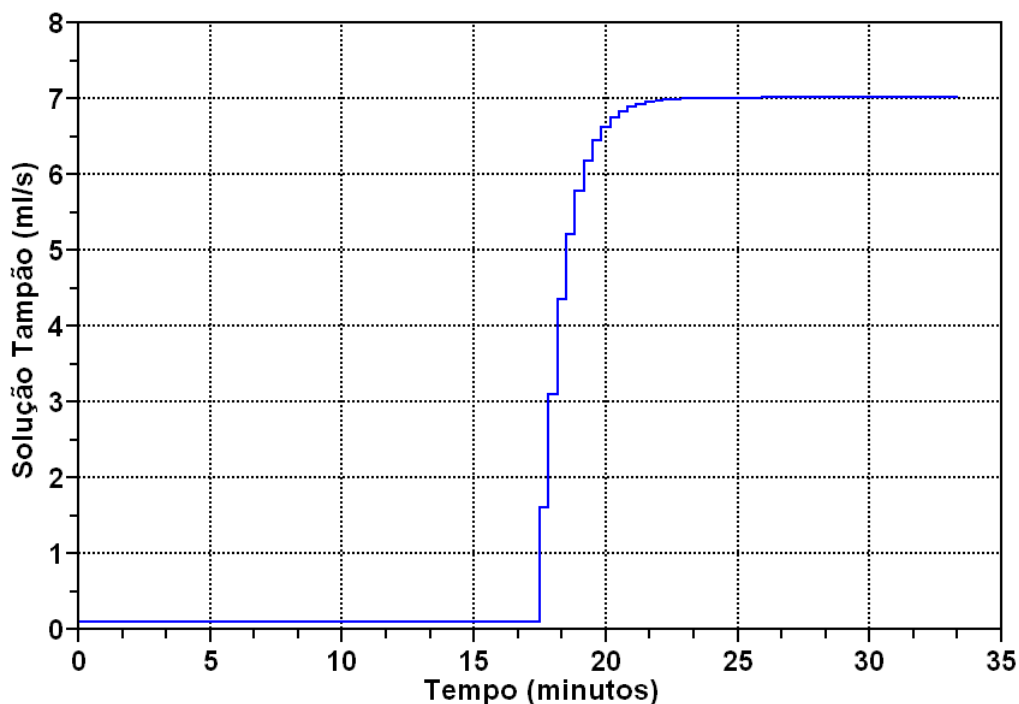


Figura 4.51: Vazão de solução tampão utilizada como variável manipulada devido a alteração de objetivos de controle.

4.4.5 Resultados e Discussão

O reator de neutralização é controlado de maneira eficiente com um controlador preditivo sucessivamente linearizado. Verificou-se a necessidade de adição de uma quantidade mínima de solução tampão para possibilitar o controle do processo nas proximidades do $\text{pH}=7$, devido ao elevado grau de não linearidade desta região.

Explorou-se a utilização de uma terceira variável manipulada (vazão de solução tampão) na presença de falhas na válvula de alimentação de base. As respostas com controle tolerante mostram que sua utilização é viável para contornar efeitos de falhas, no entanto, estudos devem ser feitos para compatibilizar a aquisição de tolerância a falhas versus o investimento em mais um atuador para o processo.

CAPÍTULO 5

Conclusões e Sugestões

O investimento em pesquisa e desenvolvimento em técnicas de monitoramento e controle tolerante a falhas é crescente como pode ser constatado pela seção de revisão 2.1 do Capítulo 2.

O controle tolerante a falhas não apresenta uma estrutura específica para a solução de todos problemas relacionados com falhas em processos. Dessa forma, foi necessário explorar e unir diferentes técnicas de monitoramento de processos para se obter controle tolerante a falhas. Essas informações podem ser acessadas diretamente pelo Capítulo 3.

As propostas de controle tolerante a falhas apresentadas neste trabalho (Capítulo 3) e testadas através de exemplos (Capítulo 4) fornecem uma indicação para o desenvolvimento e investigação de técnicas de controle tolerante a falhas.

Como a observação de alguns casos ilustrativos da técnica proposta, pode-se concluir que a metodologia sugerida é adequada para lidar com alguns tipos de falhas em processo de Engenharia Química. Apesar de não existir nenhuma técnica específica para a acomodação de falhas em processos controlados por controladores preditivos, utiliza-se como alternativa principal a atualização do modelo do processo para o controlador na tentativa de buscar tolerância a falhas. Além disso, as estratégias de manutenção do controlador nominal e reconfiguração da planta foram utilizadas com sucesso, sendo possível de estimar medidas incorretas para impedir a quebra de uma malha de controle, troca de instrumentos atuadores ou medidores. No entanto, a reconstrução de medidas através da análise estrutural sugere a implementação de filtros para a redução de ruídos.

A proposta de utilização de um banco de controladores mostrou-se viável para o exemplo da coluna de destilação binária. Os resultados obtidos mostram que a alternância entre estruturas de controle é uma alternativa para acomodar efeitos de falhas. No entanto, falhas no atuadores de vazão de refluxo e alimentação de vapor não podem ser acomodadas pela simples troca de estrutura e requerem uma investigação mais detalhada. Mesmo assim, mostra-se com este exemplo a habilidade de combinar técnicas de detecção e diagnóstico baseadas em dados com uma estrutura de controle PID para tornar o sistema de controle tolerante a falhas.

A abordagem de diferentes técnicas de detecção e diagnóstico de falhas, além das próprias técnicas de controle tolerantes específicas mostrou através de quatro exemplo principais a possibilidade de unir diferentes estratégias de monitoramento para fazer com que um processo controlado adquira novas características de tolerância a falhas. As técnicas de detecção e diagnóstico por PCA, FDA e Análise Estrutural foram implementadas com êxito, demonstradas pelos exemplos do processo de produção de ciclopentanol e da coluna de destilação binária.

O desenvolvimento de um controlador preditivo baseado em modelo sucessivamente linearizado é uma alternativa para a obtenção de atualizações do modelo do processo para propósitos de controle tolerante. Além disso, é possível obter representações mais precisas de processos lineares e controlar processos com elevado grau de não linearidade.

Sugestões

As seguintes sugestões para próximos trabalhos são relevantes:

- Exploração de técnicas de detecção e diagnóstico de falhas quando dados não gaussianos são disponíveis.
- Utilização do controlador preditivo baseado em modelo sucessivamente linearizado para processos com elevado grau de não linearidade juntamente com a exploração de falhas.
- Explorar a detecção, diagnóstico e controle tolerante através da lógica *Fuzzy*.
- Unir estratégias de monitoramento de controladores com estratégias de detecção, diagnóstico e controle tolerante.
- Explorar as técnicas de controle tolerante apresentas neste trabalho para processos reais.

APÊNDICE A

Linearização Sucessiva

Seja um sistema não linear e dinâmico descrito pela seguinte equação geral:

$$\frac{d\mathbf{x}}{dt} = \mathbf{m}(\mathbf{x}, \mathbf{u}), \quad (\text{A.1})$$

$$\mathbf{y} = \mathbf{n}(\mathbf{x}), \quad (\text{A.2})$$

em que $\mathbf{u} \in \mathbb{R}^n$ é o vetor de variáveis manipuladas do sistema, $\mathbf{x} \in \mathbb{R}^m$ é o vetor de estados estimados e $\mathbf{y} \in \mathbb{R}^n$ é o vetor de variáveis controladas. A linearização da Equação (A.1) através da expansão de Taylor em torno do ponto $(\mathbf{x}_{k-1}, \mathbf{u}_{k-1})$ ou seja, pontos anteriores ao ponto de observação atual $(\mathbf{x}_k, \mathbf{u}_k)$ do sistema resulta em:

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}_{k-1}\mathbf{x}' + \mathbf{B}_{k-1}\mathbf{u}' + \mathbf{m}_{k-1}, \quad (\text{A.3})$$

como $\frac{d\mathbf{x}}{dt} = \frac{d\mathbf{x}'}{dt}$, a Equação (A.3) pode ser escrita da seguinte forma:

$$\frac{d\mathbf{x}'}{dt} = \mathbf{A}_{k-1}\mathbf{x}' + \mathbf{B}_{k-1}\mathbf{u}' + \mathbf{m}_{k-1} \quad (\text{A.4})$$

e

$$\mathbf{y}' = \mathbf{C}_{k-1}\mathbf{x}', \quad (\text{A.5})$$

com as matrizes $\mathbf{A}_{k-1} \in \mathbb{R}^{m \times m}$, $\mathbf{B}_{k-1} \in \mathbb{R}^{m \times n}$ e $\mathbf{C}_{k-1} \in \mathbb{R}^{n \times m}$ dadas por:

$$\mathbf{A}_{k-1} = \left(\frac{\partial \mathbf{m}}{\partial \mathbf{x}} \right)_{\mathbf{x}=\mathbf{x}_{k-1}, \mathbf{u}=\mathbf{u}_{k-1}}, \quad (\text{A.6})$$

$$\mathbf{B}_{k-1} = \left(\frac{\partial \mathbf{m}}{\partial \mathbf{u}} \right)_{\mathbf{x}=\mathbf{x}_{k-1}, \mathbf{u}=\mathbf{u}_{k-1}}, \quad (\text{A.7})$$

$$\mathbf{C}_{k-1} = \left(\frac{\partial \mathbf{n}}{\partial \mathbf{x}} \right)_{\mathbf{x}=\mathbf{x}_{k-1}}, \quad (\text{A.8})$$

e as variáveis $\mathbf{x}'$, $\mathbf{u}'$ e $\mathbf{y}'$ como variáveis desvio em relação ao ponto imediatamente anterior ao ponto de operação atual. Supondo que o ponto de operação atual se encontra no instante k , tem-se:

$$\begin{aligned} \mathbf{x}'_k &= \mathbf{x}_k - \mathbf{x}_{k-1}, \\ \mathbf{u}'_k &= \mathbf{u}_k - \mathbf{u}_{k-1}, \\ \mathbf{y}'_k &= \mathbf{y}_k - \mathbf{y}_{k-1}. \end{aligned} \quad (\text{A.9})$$

Multiplicando a Equação (A.4) por $e^{-\mathbf{A}_{k-1}t}$:

$$e^{-\mathbf{A}_{k-1}t} \frac{d\mathbf{x}'}{dt} = e^{-\mathbf{A}_{k-1}t} (\mathbf{A}_{k-1}\mathbf{x}' + \mathbf{B}_{k-1}\mathbf{u}' + \mathbf{m}_{k-1}). \quad (\text{A.10})$$

Reorganizando os termos da Equação (A.10):

$$e^{-\mathbf{A}_{k-1}t} \frac{d\mathbf{x}'}{dt} - e^{-\mathbf{A}_{k-1}t} \mathbf{A}_{k-1}\mathbf{x}' = e^{-\mathbf{A}_{k-1}t} \mathbf{B}_{k-1}\mathbf{u}' + e^{-\mathbf{A}_{k-1}t} \mathbf{m}_{k-1}. \quad (\text{A.11})$$

A Equação (A.11) é reduzida à expressão (A.12):

$$\frac{d}{dt} (e^{-\mathbf{A}_{k-1}t} \mathbf{x}') = e^{-\mathbf{A}_{k-1}t} (\mathbf{B}_{k-1}\mathbf{u}' + \mathbf{m}_{k-1}), \quad (\text{A.12})$$

já que

$$e^{-\mathbf{A}_{k-1}t} \mathbf{A}_{k-1} = \mathbf{A}_{k-1} e^{-\mathbf{A}_{k-1}t}.$$

É possível integrar a Equação (A.12) de um tempo $t = t_k$ até um tempo $t = t_{k+1}$ apenas se a hipótese de que neste intervalo o valor de $\mathbf{u}'$ é constante for válida, utilizando para isto um *hold* de ordem zero, que significa que dentro do intervalo de amostragem ($t_k \rightarrow t_{k+1}$), $\mathbf{u}'$ é mantido constante. Assim:

$$\int_{t_k}^{t_{k+1}} d(e^{-\mathbf{A}_{k-1}t} \mathbf{x}') = \int_{t_k}^{t_{k+1}} e^{-\mathbf{A}_{k-1}t} (\mathbf{B}_{k-1}\mathbf{u}' + \mathbf{m}_{k-1}) dt. \quad (\text{A.13})$$

Como $\mathbf{B}_{k-1}$, $\mathbf{u}'$ e $\mathbf{m}_{k-1}$ são constantes no intervalo de integração (igual ao de amostragem, $t_k \rightarrow t_{k+1}$), a Equação (A.13) pode ser integrada, resultando em:

$$\begin{aligned} e^{-\mathbf{A}_{k-1}t_{k+1}} \mathbf{x}'_{k+1} - e^{-\mathbf{A}_{k-1}t_k} \mathbf{x}'_k = \\ -\mathbf{A}_{k-1}^{-1} e^{-\mathbf{A}_{k-1}t_{k+1}} (\mathbf{B}_{k-1}\mathbf{u}'_k + \mathbf{m}_{k-1}) + \mathbf{A}_{k-1}^{-1} e^{-\mathbf{A}_{k-1}t_k} (\mathbf{B}_{k-1}\mathbf{u}'_k + \mathbf{m}_{k-1}). \end{aligned} \quad (\text{A.14})$$

Multiplicando ambos os lados da Equação (A.14) por $\mathbf{A}_{k-1}$:

$$\mathbf{A}_{k-1}e^{-\mathbf{A}_{k-1}t_{k+1}}\mathbf{x}'_{k+1} - \mathbf{A}_{k-1}e^{-\mathbf{A}_{k-1}t_k}\mathbf{x}'_k = (e^{-\mathbf{A}_{k-1}t_k} - e^{-\mathbf{A}_{k-1}t_{k+1}})(\mathbf{B}_{k-1}\mathbf{u}'_k + \mathbf{m}_{k-1}). \quad (\text{A.15})$$

Multiplicando novamente ambos os lados da Equação (A.15) por $e^{\mathbf{A}_{k-1}t_{k+1}}$:

$$e^{\mathbf{A}_{k-1}t_{k+1}}\mathbf{A}_{k-1}e^{-\mathbf{A}_{k-1}t_{k+1}}\mathbf{x}'_{k+1} - e^{\mathbf{A}_{k-1}t_{k+1}}\mathbf{A}_{k-1}e^{-\mathbf{A}_{k-1}t_k}\mathbf{x}'_k = (e^{\mathbf{A}_{k-1}(t_{k+1}-t_k)} - \mathbf{I})(\mathbf{B}_{k-1}\mathbf{u}'_k + \mathbf{m}_{k-1}). \quad (\text{A.16})$$

Isolando o termo $\mathbf{x}'_{k+1}$ e sabendo que $\Delta t = t_{k+1} - t_k = \tau$ representa o tempo de amostragem, tem-se:

$$\mathbf{x}'_{k+1} = e^{\mathbf{A}_{k-1}\tau}\mathbf{x}'_k + (e^{\mathbf{A}_{k-1}t_{k+1}}\mathbf{A}_{k-1}e^{-\mathbf{A}_{k-1}t_{k+1}})^{-1}(e^{\mathbf{A}_{k-1}\tau} - \mathbf{I})(\mathbf{B}_{k-1}\mathbf{u}'_k + \mathbf{m}_{k-1}). \quad (\text{A.17})$$

Como $e^{-\mathbf{A}_{k-1}t}\mathbf{A}_{k-1} = \mathbf{A}_{k-1}e^{-\mathbf{A}_{k-1}t}$ é válida

$$\mathbf{x}'_{k+1} = \underbrace{e^{\mathbf{A}_{k-1}\tau}}_{\Phi_{k-1}}\mathbf{x}'_k + \underbrace{\mathbf{A}_{k-1}^{-1}(e^{\mathbf{A}_{k-1}\tau} - \mathbf{I})\mathbf{B}_{k-1}}_{\Gamma_{k-1}}\mathbf{u}'_k + \underbrace{\mathbf{A}_{k-1}^{-1}(e^{\mathbf{A}_{k-1}\tau} - \mathbf{I})}_{\Omega_{k-1}}\mathbf{m}_{k-1}. \quad (\text{A.18})$$

Finalmente:

$$\mathbf{x}'_{k+1} = \Phi_{k-1}\mathbf{x}'_k + \Gamma_{k-1}\mathbf{u}'_k + \Omega_{k-1}\mathbf{m}_{k-1}, \quad (\text{A.19})$$

$$\mathbf{y}'_k = \mathbf{C}_{k-1}\mathbf{x}'_k. \quad (\text{A.20})$$

As Equações (A.19) e (A.20) estão em variável desvio em relação ao ponto imediatamente anterior ao de operação k . As relações para as variáveis desvio mostradas na Equação (A.9) são agora estendidas para serem substituídas nas referidas equações:

$$\mathbf{x}'_{k+1} = \mathbf{x}_{k+1} - \mathbf{x}_k, \quad (\text{A.21})$$

$$\mathbf{u}'_k = \mathbf{u}_k - \mathbf{u}_{k-1},$$

$$\mathbf{y}'_k = \mathbf{y}_k - \mathbf{y}_{k-1}. \quad (\text{A.22})$$

Substituindo (A.21) em (A.19):

$$\mathbf{x}_{k+1} - \mathbf{x}_k = \Phi_{k-1}(\mathbf{x}_k - \mathbf{x}_{k-1}) + \Gamma_{k-1}(\mathbf{u}_k - \mathbf{u}_{k-1}) + \Omega_{k-1}\mathbf{m}_{k-1}. \quad (\text{A.23})$$

Isolando $\mathbf{x}_{k+1}$ e agrupando os termos, chega-se a:

$$\mathbf{x}_{k+1} = (\Phi_{k-1} + \mathbf{I})\mathbf{x}_k + \Gamma_{k-1}\mathbf{u}_k - [\Phi_{k-1}\mathbf{x}_{k-1} + \Gamma_{k-1}\mathbf{u}_{k-1} - \Omega_{k-1}\mathbf{m}_{k-1}]. \quad (\text{A.24})$$

Da mesma forma, substituindo (A.22) em (A.20):

$$\mathbf{y}_k - \mathbf{y}_{k-1} = \mathbf{C}_{k-1}(\mathbf{x}_k - \mathbf{x}_{k-1}). \quad (\text{A.25})$$

Isolando $\mathbf{y}_k$:

$$\mathbf{y}_k = \mathbf{C}_{k-1}\mathbf{x}_k + [\mathbf{y}_{k-1} - \mathbf{C}_{k-1}\mathbf{x}_{k-1}]. \quad (\text{A.26})$$

APÊNDICE B

Controle Preditivo Baseado em Modelo

B.1 Introdução

O controlador desenvolvido neste trabalho utiliza um modelo do processo na forma discreta, apresentado na forma de espaço de estados. Neste modelo, $\mathbf{y}$ representa o vetor contendo variáveis de saída, ou seja, medidas do processo; $\mathbf{u}$, o vetor de entradas e $\mathbf{x}$ o vetor dos estados. Assim:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \quad k = 0, 1, 2, \dots$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k.$$

Este modelo discreto do processo pode ser escrito de forma conveniente para representar seu comportamento em qualquer instante de amostragem quando horizontes de predição e controle são fornecidos:

$$\mathbf{x}_{k+1+h}|_k = \mathbf{A}\mathbf{x}_{k+h}|_k + \mathbf{B}\mathbf{u}_{k+h}|_k, \quad h = 0, 1, 2, \dots, N \quad (\text{B.1})$$

$$\mathbf{y}_{k+h}|_k = \mathbf{C}\mathbf{x}_{k+h}|_k, \quad h = 0, 1, 2, \dots, N \quad (\text{B.2})$$

em que, $\mathbf{x}_{k+1+h}|_k$ representa o estado $\mathbf{x}$ calculado em um instante de amostragem qualquer quando h recebe qualquer valor entre 0 e N . Com isso, são realizadas N predições do comportamento do processo em cada instante de amostragem k .

O horizonte de controle móvel se baseia na minimização do seguinte horizonte quadrático infinito em malha aberta, cuja função objetivo calculada num instante de amostragem qualquer k vale:

$$\min_{\mathbf{u}^N} \sum_{h=0}^{\infty} (\mathbf{y}_{k+h|k}^T \mathbf{Q} \mathbf{y}_{k+h|k} + \mathbf{u}_{k+h|k}^T \mathbf{R} \mathbf{u}_{k+h|k} + \Delta \mathbf{u}_{k+h|k}^T \mathbf{S} \Delta \mathbf{u}_{k+h|k}), \quad (\text{B.3})$$

em que $\mathbf{Q}$ é uma matriz de ponderação simétrica positiva semi-definida para as saídas $\mathbf{y}_{k+h|k}$ calculadas pela Equação (B.2). $\mathbf{R}$ é uma matriz simétrica positiva definida que penaliza as entradas e $\mathbf{u}_{k+h|k}$ é o vetor de entrada no tempo k e no horizonte de predição h na função objetivo em malha aberta. $\mathbf{S}$ é uma matriz simétrica positiva semi-definida que penaliza a taxa de mudança das entradas, com $\Delta \mathbf{u}_{k+h|k} = \mathbf{u}_{k+h|k} - \mathbf{u}_{k+h-1|k}$ sendo a variação no vetor de entrada no tempo k e no horizonte h . O vetor $\mathbf{u}^N$ contém as N ações de controle futuras em malha aberta para um dado instante k .

$$\mathbf{u}^N = \begin{bmatrix} \mathbf{u}_k|_k \\ \mathbf{u}_{k+1|k} \\ \vdots \\ \mathbf{u}_{k+N-1|k} \end{bmatrix}. \quad (\text{B.4})$$

No tempo k e no horizonte N , o vetor de entrada $\mathbf{u}_{k+h|k}$ é igualado a zero e mantido neste valor para todo $h \geq N$ no valor de cálculo da função objetivo.

O regulador de horizonte de controle móvel calcula o vetor $\mathbf{u}^N$ que otimiza a função objetivo em malha aberta na Equação (B.3). O primeiro valor de entrada em $\mathbf{u}^N$, $\mathbf{u}_k|_k$, é então fornecido para o processo. Este procedimento é repetido a cada intervalo sucessivo de controle com uma realimentação com as medidas da planta para atualizar o vetor de estados no tempo k .

A função objetivo do horizonte de controle infinito em malha aberta na Equação (B.3) é expressa como uma função objetivo de horizonte de controle finito em malha aberta como é mostrado na Equação (B.5). Para efeito de simplificação da notação, oculta-se $|_k$ a partir deste ponto.

$$\begin{aligned} \min_{\mathbf{u}^N} J_k &= \mathbf{x}_{k+N}^T \bar{\mathbf{Q}} \mathbf{x}_{k+N} + \Delta \mathbf{u}_{k+N}^T \mathbf{S} \Delta \mathbf{u}_{k+N} + \\ &+ \sum_{h=0}^{N-1} (\mathbf{x}_{k+h}^T \mathbf{C}^T \mathbf{Q} \mathbf{C} \mathbf{x}_{k+h} + \mathbf{u}_{k+h}^T \mathbf{R} \mathbf{u}_{k+h} + \Delta \mathbf{u}_{k+h}^T \mathbf{S} \Delta \mathbf{u}_{k+h}). \end{aligned} \quad (\text{B.5})$$

O termo que penaliza a saída, na Equação (B.3) foi substituído por um termo de penalidade de estado correspondente, na Equação (B.5). A determinação da matriz de estado terminal de penalidade, $\bar{\mathbf{Q}}$, depende da estabilidade do modelo da planta. Neste

trabalho utiliza-se apenas modelos de processos estáveis, como mostrado na seção seguinte.

B.2 Controle Preditivo para Sistema Estáveis

Para sistemas estáveis, a matriz $\bar{\mathbf{Q}}$ na Equação (B.5) é definida como a soma dada pela Equação (B.6).

$$\bar{\mathbf{Q}} = \sum_{i=0}^{\infty} \mathbf{A}^{Ti} \mathbf{C}^T \mathbf{Q} \mathbf{C} \mathbf{A}^i. \quad (\text{B.6})$$

Esta soma infinita pode ser determinada pela solução da seguinte equação discreta de Lyapunov:

$$\bar{\mathbf{Q}} = \mathbf{C}^T \mathbf{Q} \mathbf{C} + \mathbf{A}^T \bar{\mathbf{Q}} \mathbf{A}. \quad (\text{B.7})$$

B.3 Função Objetivo do Controlador Preditivo

A dedução da função objetivo que calcula as ações de controle ótimas baseadas em um horizonte de predição é demonstrada a seguir utilizando um valor de $N = 3$, revelando assim a estrutura das matrizes que compõe esta função objetivo. Feito isso, a estrutura de cada uma dessas matrizes é mostrada, tornando evidente a generalização das mesmas.

Sendo assim, substituindo $N = 3$ na Equação (B.5):

$$\begin{aligned} \min_{\mathbf{u}^3} J_k &= \mathbf{x}_{k+3}^T \bar{\mathbf{Q}} \mathbf{x}_{k+3} + \Delta \mathbf{u}_{k+3}^T \mathbf{S} \Delta \mathbf{u}_{k+3} + \\ &+ \sum_{h=0}^2 (\mathbf{x}_{k+h}^T \mathbf{C}^T \mathbf{Q} \mathbf{C} \mathbf{x}_{k+h} + \mathbf{u}_{k+h}^T \mathbf{R} \mathbf{u}_{k+h} + \Delta \mathbf{u}_{k+h}^T \mathbf{S} \Delta \mathbf{u}_{k+h}). \end{aligned} \quad (\text{B.8})$$

Expandindo o somatório dado pela Equação (B.8) e agrupando os termos que contém $\mathbf{x}$, $\mathbf{u}$ e $\Delta \mathbf{u}$, tem-se:

$$\begin{aligned} \min_{\mathbf{u}^3} J_k &= \mathbf{x}_{k+1}^T \mathbf{C}^T \mathbf{Q} \mathbf{C} \mathbf{x}_{k+1} + \mathbf{x}_{k+2}^T \mathbf{C}^T \mathbf{Q} \mathbf{C} \mathbf{x}_{k+2} + \mathbf{x}_{k+3}^T \bar{\mathbf{Q}} \mathbf{x}_{k+3} + \\ &+ \Delta \mathbf{u}_k^T \mathbf{S} \Delta \mathbf{u}_k + \Delta \mathbf{u}_{k+1}^T \mathbf{S} \Delta \mathbf{u}_{k+1} + \Delta \mathbf{u}_{k+2}^T \mathbf{S} \Delta \mathbf{u}_{k+2} + \Delta \mathbf{u}_{k+3}^T \mathbf{S} \Delta \mathbf{u}_{k+3} + \\ &+ \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k + \mathbf{u}_{k+1}^T \mathbf{R} \mathbf{u}_{k+1} + \mathbf{u}_{k+2}^T \mathbf{R} \mathbf{u}_{k+2}, \end{aligned} \quad (\text{B.9})$$

como $\mathbf{x}_k$ é um valor conhecido, este não influencia o processo de minimização de J_k e é então retirado.

A Equação (B.9) pode ser representada na forma matricial:

$$\begin{aligned}
\min_{\mathbf{u}^3} J_k = & [\mathbf{x}_{k+1} \ \mathbf{x}_{k+2} \ \mathbf{x}_{k+3}]^T \begin{bmatrix} \mathbf{C}^T \mathbf{Q} \mathbf{C} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^T \mathbf{Q} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \bar{\mathbf{Q}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{x}_{k+2} \\ \mathbf{x}_{k+3} \end{bmatrix} + \\
& + [\mathbf{u}_k \ \mathbf{u}_{k+1} \ \mathbf{u}_{k+2}]^T \begin{bmatrix} \mathbf{R} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \mathbf{u}_{k+2} \end{bmatrix} + \\
& + \underbrace{\Delta \mathbf{u}_k^T \mathbf{S} \Delta \mathbf{u}_k + \Delta \mathbf{u}_{k+1}^T \mathbf{S} \Delta \mathbf{u}_{k+1} + \Delta \mathbf{u}_{k+2}^T \mathbf{S} \Delta \mathbf{u}_{k+2} + \Delta \mathbf{u}_{k+3}^T \mathbf{S} \Delta \mathbf{u}_{k+3}}_F. \tag{B.10}
\end{aligned}$$

Os termos marcados com F são tratados separadamente.

$$F = \Delta \mathbf{u}_k^T \mathbf{S} \Delta \mathbf{u}_k + \Delta \mathbf{u}_{k+1}^T \mathbf{S} \Delta \mathbf{u}_{k+1} + \Delta \mathbf{u}_{k+2}^T \mathbf{S} \Delta \mathbf{u}_{k+2} + \Delta \mathbf{u}_{k+3}^T \mathbf{S} \Delta \mathbf{u}_{k+3}. \tag{B.11}$$

Dessa forma, cada um desses termos representam:

$$\begin{aligned}
\Delta \mathbf{u}_k &= \mathbf{u}_k - \mathbf{u}_{k-1}, \\
\Delta \mathbf{u}_{k+1} &= \mathbf{u}_{k+1} - \mathbf{u}_k, \\
\Delta \mathbf{u}_{k+2} &= \mathbf{u}_{k+2} - \mathbf{u}_{k+1}, \\
\Delta \mathbf{u}_{k+3} &= \mathbf{u}_{k+3} - \mathbf{u}_{k+2}. \tag{B.12}
\end{aligned}$$

De acordo com a hipótese apresentada anteriormente, $\mathbf{u}_{k+h} = 0$ para $h \geq N$. Nesta dedução, $N = 3$ e conseqüentemente $\mathbf{u}_{k+3} = 0$. Dessa forma, analisa-se cada um dos termos da Equação (B.11):

$$\begin{aligned}
\Delta \mathbf{u}_k^T \mathbf{S} \Delta \mathbf{u}_k &= (\mathbf{u}_k - \mathbf{u}_{k-1})^T \mathbf{S} (\mathbf{u}_k - \mathbf{u}_{k-1}) = \\
& \mathbf{u}_k^T \mathbf{S} \mathbf{u}_k - \mathbf{u}_k^T \mathbf{S} \mathbf{u}_{k-1} - \mathbf{u}_{k-1}^T \mathbf{S} \mathbf{u}_k + \mathbf{u}_{k-1}^T \mathbf{S} \mathbf{u}_{k-1}. \tag{B.13}
\end{aligned}$$

$$\begin{aligned}
\Delta \mathbf{u}_{k+1}^T \mathbf{S} \Delta \mathbf{u}_{k+1} &= (\mathbf{u}_{k+1} - \mathbf{u}_k)^T \mathbf{S} (\mathbf{u}_{k+1} - \mathbf{u}_k) = \\
& \mathbf{u}_{k+1}^T \mathbf{S} \mathbf{u}_{k+1} - \mathbf{u}_{k+1}^T \mathbf{S} \mathbf{u}_k - \mathbf{u}_k^T \mathbf{S} \mathbf{u}_{k+1} + \mathbf{u}_k^T \mathbf{S} \mathbf{u}_k. \tag{B.14}
\end{aligned}$$

$$\begin{aligned}
\Delta \mathbf{u}_{k+2}^T \mathbf{S} \Delta \mathbf{u}_{k+2} &= (\mathbf{u}_{k+2} - \mathbf{u}_{k+1})^T \mathbf{S} (\mathbf{u}_{k+2} - \mathbf{u}_{k+1}) = \\
& \mathbf{u}_{k+2}^T \mathbf{S} \mathbf{u}_{k+2} - \mathbf{u}_{k+2}^T \mathbf{S} \mathbf{u}_{k+1} - \mathbf{u}_{k+1}^T \mathbf{S} \mathbf{u}_{k+2} + \mathbf{u}_{k+1}^T \mathbf{S} \mathbf{u}_{k+1}. \tag{B.15}
\end{aligned}$$

$$\Delta \mathbf{u}_{k+3}^T \mathbf{S} \Delta \mathbf{u}_{k+3} = \mathbf{u}_{k+2}^T \mathbf{S} \mathbf{u}_{k+2}. \tag{B.16}$$

Substituindo (B.13), (B.14), (B.15) e (B.16) em (B.11):

$$\begin{aligned}
 F = & \mathbf{u}_k^T \mathbf{S} \mathbf{u}_k - \mathbf{u}_k^T \mathbf{S} \mathbf{u}_{k-1} - \mathbf{u}_{k-1}^T \mathbf{S} \mathbf{u}_k + \mathbf{u}_{k-1}^T \mathbf{S} \mathbf{u}_{k-1} + \mathbf{u}_{k+1}^T \mathbf{S} \mathbf{u}_{k+1} + \\
 & - \mathbf{u}_{k+1}^T \mathbf{S} \mathbf{u}_k - \mathbf{u}_k^T \mathbf{S} \mathbf{u}_{k+1} + \mathbf{u}_k^T \mathbf{S} \mathbf{u}_k + \mathbf{u}_{k+2}^T \mathbf{S} \mathbf{u}_{k+2} - \mathbf{u}_{k+2}^T \mathbf{S} \mathbf{u}_{k+1} + \\
 & - \mathbf{u}_{k+1}^T \mathbf{S} \mathbf{u}_{k+2} + \mathbf{u}_{k+1}^T \mathbf{S} \mathbf{u}_{k+1} + \mathbf{u}_{k+2}^T \mathbf{S} \mathbf{u}_{k+2}.
 \end{aligned} \tag{B.17}$$

Como a matriz $\mathbf{S}$ é simétrica, é possível agrupar os termos da Equação (B.17):

$$F = [\mathbf{u}_k \ \mathbf{u}_{k+1} \ \mathbf{u}_{k+2}]^T \begin{bmatrix} 2\mathbf{S} & -\mathbf{S} & \mathbf{0} \\ -\mathbf{S} & 2\mathbf{S} & -\mathbf{S} \\ \mathbf{0} & -\mathbf{S} & 2\mathbf{S} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \mathbf{u}_{k+2} \end{bmatrix} - 2\mathbf{u}_k^T \mathbf{S} \mathbf{u}_{k-1} + \mathbf{u}_{k-1}^T \mathbf{S} \mathbf{u}_{k-1}. \tag{B.18}$$

Com isso, substituindo (B.18) em (B.10) e generalizando para um horizonte N qualquer:

$$\begin{aligned}
 \min_{\mathbf{u}^N} J_k = & \underbrace{[\mathbf{x}_{k+1} \ \mathbf{x}_{k+2} \ \cdots \ \mathbf{x}_{k+N}]^T}_{\mathbf{x}^{NT}} \underbrace{\begin{bmatrix} \mathbf{C}^T \mathbf{Q} \mathbf{C} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^T \mathbf{Q} \mathbf{C} & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \bar{\mathbf{Q}} \end{bmatrix}}_{\Theta} \begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{x}_{k+2} \\ \vdots \\ \mathbf{x}_{k+N} \end{bmatrix} + \\
 & + \underbrace{[\mathbf{u}_k \ \mathbf{u}_{k+1} \ \cdots \ \mathbf{u}_{k+N-1}]^T}_{\mathbf{u}^{NT}} \underbrace{\begin{bmatrix} \mathbf{R} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R} & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{R} \end{bmatrix}}_{\Xi} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \vdots \\ \mathbf{u}_{k+N-1} \end{bmatrix} + \\
 & + \underbrace{[\mathbf{u}_k \ \mathbf{u}_{k+1} \ \cdots \ \mathbf{u}_{k+N-1}]^T}_{\Upsilon} \underbrace{\begin{bmatrix} 2\mathbf{S} & -\mathbf{S} & \cdots & \mathbf{0} \\ -\mathbf{S} & 2\mathbf{S} & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & -\mathbf{S} \\ \mathbf{0} & \mathbf{0} & -\mathbf{S} & 2\mathbf{S} \end{bmatrix}}_{\Upsilon} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \vdots \\ \mathbf{u}_{k+N-1} \end{bmatrix} + \\
 & - [\mathbf{u}_k \ \mathbf{u}_{k+1} \ \cdots \ \mathbf{u}_{k+N-1}]^T 2 \underbrace{\begin{bmatrix} \mathbf{S} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}}_{\mathbf{g}} \mathbf{u}_{k-1},
 \end{aligned} \tag{B.19}$$

em que o termo $\mathbf{u}_{k-1}^T \mathbf{S} \mathbf{u}_{k-1}$ presente na Equação (B.18) pode ser desconsiderado para

efeitos de otimização, visto que o mesmo é constante no instante de tempo corrente k .

Considerando esta nova notação, a Equação (B.19) é expressa em uma forma mais condensada.

$$\min_{\mathbf{u}^N} J_k = \mathbf{x}^{NT} \mathbf{\Theta} \mathbf{x}^N + \mathbf{u}^{NT} \mathbf{\Xi} \mathbf{u}^N + \mathbf{u}^{NT} \mathbf{\Upsilon} \mathbf{u}^N - \mathbf{u}^{NT} 2\mathbf{g} \mathbf{u}_{k-1}. \quad (\text{B.20})$$

B.3.1 Conexão entre modelo e função objetivo

O controlador preditivo apresentado neste trabalho propõe-se a controlar um determinado processo através de seu modelo na forma discreta e em espaço de estado utilizando para isto a minimização de uma função objetivo em relação à variável de entrada do processo $\mathbf{u}$. Como pode-se notar, a Equação (B.20) contém além de termos relacionados à entrada do processo, termos relacionados aos estados, $\mathbf{x}$. Para efeito de minimização da função objetivo em relação à entrada do processo para um dado horizonte de predição N , os termos que não contém esta variável são considerados constantes durante o processo de otimização.

É a partir da inserção do modelo do processo na função objetivo apresentada na Equação (B.20) que o controle preditivo baseado em modelo torna-se completo e conhecido como MPC (*Model Predictive Control*, Controle Preditivo Baseado em Modelo).

A eficiência do controle preditivo baseado em modelo está diretamente associada com a representatividade do processo pelo modelo, ou seja, a validade do modelo em um amplo espaço de variáveis manipuladas.

O objetivo, a partir de agora, é relacionar o modelo do processo com a função objetivo (Equação (B.20)). Para isto utiliza-se como exemplo um horizonte de predição $N = 3$, ou seja, $h = 0, 1, 2, 3$. Utilizando a Equação (B.1) e supondo que $\mathbf{x}_{k+h}|_k$ e $\mathbf{u}_{k+h}|_k$ em $h = 0$ sejam valores conhecidos para o processo em qualquer instante de amostragem k , a Equação (B.1) pode ser reescrita da seguinte forma:

$$\mathbf{x}_{k+h}|_k = \mathbf{A} \mathbf{x}_{k+h-1}|_k + \mathbf{B} \mathbf{u}_{k+h-1}|_k. \quad (\text{B.21})$$

Para $h = 0$, a Equação (B.21) se reduz a:

$$\mathbf{x}_k|_k = \mathbf{A} \mathbf{x}_{k-1}|_k + \mathbf{B} \mathbf{u}_{k-1}|_k, \quad (\text{B.22})$$

que por sua vez representa um valor constante para qualquer instante k e não é considerado pelo problema de otimização.

Substituindo $h = 1$ na Equação (B.21):

$$\mathbf{x}_{k+1}|_k = \mathbf{A}\mathbf{x}_k|_k + \mathbf{B}\mathbf{u}_k|_k. \quad (\text{B.23})$$

Da mesma forma, substituindo $h = 2$ na Equação (B.21) e substituindo (B.23) neste resultado chega-se a:

$$\mathbf{x}_{k+2}|_k = \mathbf{A}^2\mathbf{x}_k|_k + \mathbf{A}\mathbf{B}\mathbf{u}_k|_k + \mathbf{B}\mathbf{u}_{k+1}|_k. \quad (\text{B.24})$$

Para $h = 3$ e seguindo o mesmo procedimento anterior tem-se:

$$\mathbf{x}_{k+3}|_k = \mathbf{A}^3\mathbf{x}_k|_k + \mathbf{A}^2\mathbf{B}\mathbf{u}_k|_k + \mathbf{A}\mathbf{B}\mathbf{u}_{k+1}|_k + \mathbf{B}\mathbf{u}_{k+2}|_k. \quad (\text{B.25})$$

Representando as Equações (B.23), (B.24) e (B.25) na forma matricial e ocultando $|_k$ para simplificar a notação, tem-se:

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{x}_{k+2} \\ \mathbf{x}_{k+3} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{A}^2 \\ \mathbf{A}^3 \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} \mathbf{B} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}\mathbf{B} & \mathbf{B} & \mathbf{0} \\ \mathbf{A}^2\mathbf{B} & \mathbf{A}\mathbf{B} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \mathbf{u}_{k+2} \end{bmatrix}. \quad (\text{B.26})$$

Generalizando para um horizonte N qualquer:

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{x}_{k+2} \\ \vdots \\ \mathbf{x}_{k+N} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{A} \\ \mathbf{A}^2 \\ \vdots \\ \mathbf{A}^N \end{bmatrix}}_{\mathbf{\Lambda}} \mathbf{x}_k + \underbrace{\begin{bmatrix} \mathbf{B} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{A}\mathbf{B} & \mathbf{B} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{A}^{N-1}\mathbf{B} & \mathbf{A}^{N-2}\mathbf{B} & \cdots & \mathbf{B} \end{bmatrix}}_{\mathbf{\Psi}} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \vdots \\ \mathbf{u}_{k+N-1} \end{bmatrix}. \quad (\text{B.27})$$

A Equação (B.28) representa a Equação (B.27) em uma forma compacta com uma notação compatível com àquela proposta pela Equação (B.20).

$$\mathbf{x}^N = \mathbf{\Lambda}\mathbf{x}_k + \mathbf{\Psi}\mathbf{u}^N. \quad (\text{B.28})$$

Finalmente, o controle preditivo baseado em modelo fica evidente quando na Equação (B.20) substitui-se o modelo do processo expandido para um mesmo horizonte de predição N e um instante de amostragem qualquer k . Sendo assim, substituindo (B.28) em (B.20):

$$\min_{\mathbf{u}^N} J_k = (\mathbf{x}_k^{NT} \mathbf{\Lambda}^T + \mathbf{u}^{NT} \mathbf{\Psi}^T) \mathbf{\Theta} (\mathbf{\Lambda} \mathbf{x}_k^N + \mathbf{\Psi} \mathbf{u}^N) + \mathbf{u}^{NT} \mathbf{\Xi} \mathbf{u}^N + \mathbf{u}^{NT} \mathbf{\Upsilon} \mathbf{u}^N - \mathbf{u}^{NT} 2\mathbf{g} \mathbf{u}_{k-1}. \quad (\text{B.29})$$

Expandindo a Equação (B.29):

$$\begin{aligned} \min_{\mathbf{u}^N} J_k = & (\mathbf{x}_k^{NT} \mathbf{\Lambda}^T \mathbf{\Theta} \mathbf{\Lambda} \mathbf{x}_k^N + \underbrace{\mathbf{x}_k^{NT} \mathbf{\Lambda}^T \mathbf{\Theta} \mathbf{\Psi} \mathbf{u}^N}_{\dagger} + \underbrace{\mathbf{u}^{NT} \mathbf{\Psi}^T \mathbf{\Theta} \mathbf{\Lambda} \mathbf{x}_k^N}_{\dagger} + \underbrace{\mathbf{u}^{NT} \mathbf{\Psi}^T \mathbf{\Theta} \mathbf{\Psi} \mathbf{u}^N}_{\ddagger}) + \\ & + \underbrace{\mathbf{u}^{NT} \mathbf{\Xi} \mathbf{u}^N}_{\ddagger} + \underbrace{\mathbf{u}^{NT} \mathbf{\Upsilon} \mathbf{u}^N}_{\ddagger} - \mathbf{u}^{NT} 2\mathbf{g} \mathbf{u}_{k-1}. \end{aligned} \quad (\text{B.30})$$

Na Equação (B.30), os termos que não contém o vetor $\mathbf{u}^N$ não influenciam o processo de otimização, visto que a minimização ocorre em $\mathbf{u}^N$. Logo, todas as outras variáveis são consideradas constantes. Nesta mesma equação, os termos marcados com o símbolo $\dagger$ podem ser agrupados já que a matriz $\mathbf{\Theta}$ é simétrica, e os termos marcados com o símbolo $\ddagger$ são diretamente agrupados. O resultado deste agrupamento é:

$$\min_{\mathbf{u}^N} J_k = \mathbf{u}^{NT} (\mathbf{\Psi}^T \mathbf{\Theta} \mathbf{\Psi} + \mathbf{\Xi} + \mathbf{\Upsilon}) \mathbf{u}^N + 2\mathbf{u}^{NT} (\mathbf{\Psi}^T \mathbf{\Theta} \mathbf{\Lambda} \mathbf{x}_k - \mathbf{g} \mathbf{u}_{k-1}). \quad (\text{B.31})$$

B.4 Restrições Aplicadas ao Problema de Otimização

A representação final do controle preditivo baseado em modelo é um problema de otimização. A Equação (B.31) deve ser resolvida como um problema de programação quadrática. Tal fato possibilita anexar ao problema de encontrar de forma ótima o conjunto de variáveis manipuladas $\mathbf{u}^N$, restrições. Neste trabalho, apenas restrições nas variáveis de entrada e saída são consideradas:

$$\mathbf{u}_{min} \leq \mathbf{u}_{k+h}|_k \leq \mathbf{u}_{max}, \quad h = 0, 1, \dots, N-1 \quad (\text{B.32})$$

$$\Delta \mathbf{u}_{k+h}|_k \leq |\Delta \mathbf{u}_{max}|, \quad h = 0, 1, \dots, N-1 \quad (\text{B.33})$$

$$\mathbf{y}_{min} \leq \mathbf{y}_{k+h}|_k \leq \mathbf{y}_{max}. \quad h = h_1, h_1 + 1, \dots, h_2 \quad (\text{B.34})$$

As restrições de saída são aplicadas do instante k e horizonte h_1 , $h_1 \geq 1$, até o horizonte h_2 , $h_2 \geq h_1$ e para o mesmo instante k . O valor de h_2 é escolhido tal que a probabilidade da existência de solução com as restrições de saída após o instante $k + h_2$ implica na probabilidade de existência de solução dessas restrições no horizonte infinito. O valor de h_1 é escolhido tal que as restrições de saída sejam obedecidas em k . O controle regulador com restrição removerá as restrições de saída no início do horizonte após $k + h_1$, de tal forma a obter restrições realizáveis e uma solução da programação quadrática.

B.4.1 Restrição Aplicada à Entrada do Processo

A Equação (B.32) é expandida para um valor de N qualquer

$$\begin{aligned} \mathbf{u}_{min} &\leq \mathbf{u}_k \leq \mathbf{u}_{max} \\ \mathbf{u}_{min} &\leq \mathbf{u}_{k+1} \leq \mathbf{u}_{max} \\ &\vdots \\ \mathbf{u}_{min} &\leq \underbrace{\mathbf{u}_{k+N-1}}_{\mathbf{u}^N} \leq \mathbf{u}_{max} \end{aligned} \quad . \quad (\text{B.35})$$

As desigualdades da Equação (B.35) são divididas da seguinte forma:

$$\begin{aligned} \mathbf{u}_k &\leq -\mathbf{u}_{min} & \mathbf{u}_k &\leq \mathbf{u}_{max} \\ \mathbf{u}_{k+1} &\leq -\mathbf{u}_{min} & \mathbf{u}_{k+1} &\leq \mathbf{u}_{max} \\ &\vdots & &\vdots \\ \underbrace{\mathbf{u}_{k+N-1}}_{\mathbf{u}^N} &\leq -\mathbf{u}_{min} & \underbrace{\mathbf{u}_{k+N-1}}_{\mathbf{u}^N} &\leq \mathbf{u}_{max} \\ &\mathbf{i}_2 & &\mathbf{i}_1 \end{aligned} \quad \text{e} \quad (\text{B.36})$$

As duas partes da Equação (B.36) são representadas na forma matricial:

$$\begin{bmatrix} \mathbf{I} \\ -\mathbf{I} \end{bmatrix} \mathbf{u}^N \leq \begin{bmatrix} \mathbf{i}_1 \\ \mathbf{i}_2 \end{bmatrix}. \quad (\text{B.37})$$

B.4.2 Restrição Aplicada à Variação na Entrada do Processo

A Equação (B.33) considera apenas restrição para a variação máxima entre ações de controle em instantes diferentes. Este tipo de restrição é necessária para qualquer sistema real, pois com isso é possível impedir ações de controle agressivas que podem danificar elementos atuadores do processo. Dessa forma, a mesma equação é estendida para um horizonte N qualquer:

$$\begin{aligned} \Delta \mathbf{u}_k &\leq |\Delta \mathbf{u}_{max}| \\ \Delta \mathbf{u}_{k+1} &\leq |\Delta \mathbf{u}_{max}| \\ &\vdots \\ \Delta \mathbf{u}_{k+N} &\leq |\Delta \mathbf{u}_{max}| \end{aligned} \quad . \quad (\text{B.38})$$

Sabendo que $\Delta \mathbf{u}_{k+N} = \mathbf{u}_{k+N} - \mathbf{u}_{k+N-1}$, a desigualdade da Equação (B.38) é expandida

da seguinte forma:

$$\begin{aligned} \mathbf{u}_k &\leq |\Delta \mathbf{u}_{max}| + \mathbf{u}_{k-1} \\ \mathbf{u}_{k+1} - \mathbf{u}_k &\leq |\Delta \mathbf{u}_{max}| \\ &\vdots \\ \mathbf{u}_{k+N} - \mathbf{u}_{k+N-1} &\leq |\Delta \mathbf{u}_{max}| \end{aligned} \quad . \quad (\text{B.39})$$

O lado esquerdo da desigualdade da Equação (B.39) deve ser representado na forma matricial para que se reduza essa expressão a uma similar àquela desenvolvida na seção anterior (Equação (B.37)). Uma demonstração bastante clara da construção dessa matriz é exemplificada para um horizonte de predição fixo, $h = 0, 1, 2, 3$.

$$\begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} - \mathbf{u}_k \\ \mathbf{u}_{k+2} - \mathbf{u}_{k+1} \\ -\mathbf{u}_{k+2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \mathbf{u}_{k+2} \end{bmatrix}. \quad (\text{B.40})$$

A Equação (B.40) é generalizada e utilizada na Equação (B.39) para formar:

$$\underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ -1 & 1 & \cdots & 0 \\ \vdots & -1 & \ddots & \vdots \\ 0 & \cdots & \ddots & 1 \\ 0 & \cdots & 0 & -1 \end{bmatrix}}_{\mathbf{T}} \underbrace{\begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \mathbf{u}_{k+2} \\ \vdots \\ \mathbf{u}_{k+N-1} \end{bmatrix}}_{\mathbf{u}^N} \leq \underbrace{\begin{bmatrix} |\Delta \mathbf{u}_{max}| + \mathbf{u}_{k-1} \\ |\Delta \mathbf{u}_{max}| \\ |\Delta \mathbf{u}_{max}| \\ \vdots \\ |\Delta \mathbf{u}_{max}| \end{bmatrix}}_{\mathbf{t}_1}. \quad (\text{B.41})$$

ou

$$\mathbf{T} \mathbf{u}^N \leq \mathbf{t}_1. \quad (\text{B.42})$$

B.4.3 Restrição Aplicada para a Saída do Processo

A restrição na saída do processo para este tipo de controlador é apresentada na Equação (B.34). Como comentado naquela ocasião, as restrições são aplicadas do instante k e horizonte h_1 , $h_1 \geq 1$, até o horizonte h_2 para um mesmo instante k , $h_2 \geq h_1$. Os valores de h_1 e h_2 são escolhidos de tal forma que o problema de otimização possua solução. Uma discussão mais completa sobre a escolha desses parâmetros pode ser encontrada em Muske e Rawlings (1993).

Diferentemente das Equações (B.32) e (B.33), o horizonte de predição h , que nas outras equações variava de 0 até N , é agora modificado para variar de h_1 até h_2 . Isto é feito para que o problema de otimização admita solução, visto que dificilmente as restrições de saída são atendidas a partir do instante inicial. O objetivo a partir de agora é relacionar

a saída do processo com a entrada através do modelo. Para isto, utiliza-se como exemplo, $h_1 = 2$, $h_2 = 4$ e $N = 6$. Neste, apesar de se ter um horizonte de predição $h = 0, \dots, 6$ para as outras restrições, os horizontes em que as restrições na saída irão realmente atuar serão $h = 2, 3, 4$, para cada instante k . Tal fato gera uma inconsistência entre as matrizes do conjunto de restrições, devido a isso, deve-se adequar o tamanho do horizonte de predição do problema (N) com o horizonte em que a restrição na saída ($h_1 \rightarrow h_2$) é aplicada, fazendo $\mathbf{A}^{j-i} = \mathbf{0}$ para $j - i < 0$.

Para o horizonte de predição considerado ($N = 6$) e utilizando a Equação (B.4), tem-se:

$$\mathbf{u}^6 = \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \mathbf{u}_{k+2} \\ \mathbf{u}_{k+3} \\ \mathbf{u}_{k+4} \\ \mathbf{u}_{k+5} \end{bmatrix}. \quad (\text{B.43})$$

Utilizando os valores de $h_1 = 2$ e $h_2 = 4$ do exemplo e aplicando-os na Equação (B.34) tem-se:

$$\begin{aligned} \mathbf{y}_{min} &\leq \mathbf{y}_{k+2} \leq \mathbf{y}_{max} \\ \mathbf{y}_{min} &\leq \mathbf{y}_{k+3} \leq \mathbf{y}_{max} \\ \mathbf{y}_{min} &\leq \mathbf{y}_{k+4} \leq \mathbf{y}_{max} \end{aligned} \quad (\text{B.44})$$

Para relacionar a saída $\mathbf{y}_{k+2}$ com a entrada é necessário utilizar a Equação (B.1) iniciando com $h = 0$. Sendo assim, das Equações (B.1) e (B.2) tem-se:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k. \end{aligned} \quad (\text{B.45})$$

Prosseguindo para $h = 1$, e utilizando (B.45) tem-se:

$$\mathbf{y}_{k+1} = \mathbf{C}\mathbf{A}\mathbf{x}_k + \mathbf{C}\mathbf{B}\mathbf{u}_k. \quad (\text{B.46})$$

A partir de $h = 2$ até $h = 4$ o horizonte de predição será igual ao horizonte em que a restrição de saída será aplicada. Então:

$$\mathbf{x}_{k+2} = \mathbf{A}\mathbf{x}_{k+1} + \mathbf{B}\mathbf{u}_{k+1}. \quad (\text{B.47})$$

Substituindo (B.45) em (B.47) chega-se a:

$$\mathbf{x}_{k+2} = \mathbf{A}^2 \mathbf{x}_k + \mathbf{A} \mathbf{B} \mathbf{u}_k + \mathbf{B} \mathbf{u}_{k+1}. \quad (\text{B.48})$$

Finalmente,

$$\mathbf{y}_{k+2} = \mathbf{C} \mathbf{A}^2 \mathbf{x}_k + \mathbf{C} \mathbf{A} \mathbf{B} \mathbf{u}_k + \mathbf{C} \mathbf{B} \mathbf{u}_{k+1}. \quad (\text{B.49})$$

Analogamente chega-se a:

$$\mathbf{y}_{k+3} = \mathbf{C} \mathbf{A}^3 \mathbf{x}_k + \mathbf{C} \mathbf{A}^2 \mathbf{B} \mathbf{u}_k + \mathbf{C} \mathbf{A} \mathbf{B} \mathbf{u}_{k+1} + \mathbf{C} \mathbf{B} \mathbf{u}_{k+2} \quad (\text{B.50})$$

e

$$\mathbf{y}_{k+4} = \mathbf{C} \mathbf{A}^4 \mathbf{x}_k + \mathbf{C} \mathbf{A}^3 \mathbf{B} \mathbf{u}_k + \mathbf{C} \mathbf{A}^2 \mathbf{B} \mathbf{u}_{k+1} + \mathbf{C} \mathbf{A} \mathbf{B} \mathbf{u}_{k+2} + \mathbf{C} \mathbf{B} \mathbf{u}_{k+3}. \quad (\text{B.51})$$

Substituindo (B.49)-(B.51) em (B.44):

$$\begin{aligned} \mathbf{y}_{min} &\leq \mathbf{C} \mathbf{A}^2 \mathbf{x}_k + \mathbf{C} \mathbf{A} \mathbf{B} \mathbf{u}_k + \mathbf{C} \mathbf{B} \mathbf{u}_{k+1} \leq \mathbf{y}_{max} \\ \mathbf{y}_{min} &\leq \mathbf{C} \mathbf{A}^3 \mathbf{x}_k + \mathbf{C} \mathbf{A}^2 \mathbf{B} \mathbf{u}_k + \mathbf{C} \mathbf{A} \mathbf{B} \mathbf{u}_{k+1} + \mathbf{C} \mathbf{B} \mathbf{u}_{k+2} \leq \mathbf{y}_{max} \\ \mathbf{y}_{min} &\leq \mathbf{C} \mathbf{A}^4 \mathbf{x}_k + \mathbf{C} \mathbf{A}^3 \mathbf{B} \mathbf{u}_k + \mathbf{C} \mathbf{A}^2 \mathbf{B} \mathbf{u}_{k+1} + \mathbf{C} \mathbf{A} \mathbf{B} \mathbf{u}_{k+2} + \mathbf{C} \mathbf{B} \mathbf{u}_{k+3} \leq \mathbf{y}_{max} \end{aligned} \quad (\text{B.52})$$

As desigualdades da Equação (B.52) são analisadas individualmente para $\mathbf{y}_{min}$ e $\mathbf{y}_{max}$, respectivamente:

$$\begin{aligned} \mathbf{C} \mathbf{A} \mathbf{B} \mathbf{u}_k + \mathbf{C} \mathbf{B} \mathbf{u}_{k+1} &\geq \mathbf{y}_{min} - \mathbf{C} \mathbf{A}^2 \mathbf{x}_k \\ \mathbf{C} \mathbf{A}^2 \mathbf{B} \mathbf{u}_k + \mathbf{C} \mathbf{A} \mathbf{B} \mathbf{u}_{k+1} + \mathbf{C} \mathbf{B} \mathbf{u}_{k+2} &\geq \mathbf{y}_{min} - \mathbf{C} \mathbf{A}^3 \mathbf{x}_k \\ \mathbf{C} \mathbf{A}^3 \mathbf{B} \mathbf{u}_k + \mathbf{C} \mathbf{A}^2 \mathbf{B} \mathbf{u}_{k+1} + \mathbf{C} \mathbf{A} \mathbf{B} \mathbf{u}_{k+2} + \mathbf{C} \mathbf{B} \mathbf{u}_{k+3} &\geq \mathbf{y}_{min} - \mathbf{C} \mathbf{A}^4 \mathbf{x}_k \end{aligned} \quad (\text{B.53})$$

Utilizando (B.43), escreve-se (B.53) na forma matricial:

$$\begin{bmatrix} \mathbf{C} \mathbf{A} \mathbf{B} & \mathbf{C} \mathbf{B} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{C} \mathbf{A}^2 \mathbf{B} & \mathbf{C} \mathbf{A} \mathbf{B} & \mathbf{C} \mathbf{B} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{C} \mathbf{A}^3 \mathbf{B} & \mathbf{C} \mathbf{A}^2 \mathbf{B} & \mathbf{C} \mathbf{A} \mathbf{B} & \mathbf{C} \mathbf{B} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \mathbf{u}_{k+2} \\ \mathbf{u}_{k+3} \\ \mathbf{u}_{k+4} \\ \mathbf{u}_{k+5} \end{bmatrix} \geq \begin{bmatrix} \mathbf{y}_{min} - \mathbf{C} \mathbf{A}^2 \mathbf{x}_k \\ \mathbf{y}_{min} - \mathbf{C} \mathbf{A}^3 \mathbf{x}_k \\ \mathbf{y}_{min} - \mathbf{C} \mathbf{A}^4 \mathbf{x}_k \end{bmatrix}. \quad (\text{B.54})$$

Observando a primeira matriz do lado esquerdo da desigualdade (Equação (B.54)), nota-se que o preenchimento de algumas linhas com matrizes de elementos nulos se deve ao fato da diferença entre os horizontes utilizados, ou seja, a presença de matrizes de

elementos nulos ($\mathbf{0}$) identifica os valores de horizonte de predição em que a restrição na saída não estará ativa. Da mesma forma, a desigualdade para $\mathbf{y}_{max}$ fornece:

$$\begin{bmatrix} \mathbf{CA}\mathbf{B} & \mathbf{CB} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{CA}^2\mathbf{B} & \mathbf{CA}\mathbf{B} & \mathbf{CB} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{CA}^3\mathbf{B} & \mathbf{CA}^2\mathbf{B} & \mathbf{CA}\mathbf{B} & \mathbf{CB} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \mathbf{u}_{k+2} \\ \mathbf{u}_{k+3} \\ \mathbf{u}_{k+4} \\ \mathbf{u}_{k+5} \end{bmatrix} \leq \begin{bmatrix} \mathbf{y}_{max} - \mathbf{CA}^2\mathbf{x}_k \\ \mathbf{y}_{max} - \mathbf{CA}^3\mathbf{x}_k \\ \mathbf{y}_{max} - \mathbf{CA}^4\mathbf{x}_k \end{bmatrix}. \quad (\text{B.55})$$

As Equações (B.44) e (B.45) podem ser generalizadas para quaisquer valores de $h_1 \geq 1$, $h_2 \geq h_1$ e N :

$$\underbrace{\begin{bmatrix} \mathbf{CA}^{j_1-1}\mathbf{B} & \dots & \mathbf{CA}^{j_1-N}\mathbf{B} \\ \vdots & & \vdots \\ \mathbf{CA}^{j_2-1}\mathbf{B} & \dots & \mathbf{CA}^{j_2-N}\mathbf{B} \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} \mathbf{u}_k \\ \vdots \\ \mathbf{u}_{k+N-1} \end{bmatrix}}_{\mathbf{u}^N} \leq \underbrace{\begin{bmatrix} \mathbf{y}_{max} - \mathbf{CA}^{j_1}\mathbf{x}_k \\ \vdots \\ \mathbf{y}_{max} - \mathbf{CA}^{j_2}\mathbf{x}_k \end{bmatrix}}_{\mathbf{d}_1} \quad (\text{B.56})$$

e

$$-1 \underbrace{\begin{bmatrix} \mathbf{CA}^{j_1-1}\mathbf{B} & \dots & \mathbf{CA}^{j_1-N}\mathbf{B} \\ \vdots & & \vdots \\ \mathbf{CA}^{j_2-1}\mathbf{B} & \dots & \mathbf{CA}^{j_2-N}\mathbf{B} \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} \mathbf{u}_k \\ \vdots \\ \mathbf{u}_{k+N-1} \end{bmatrix}}_{\mathbf{u}^N} \leq \underbrace{\begin{bmatrix} -\mathbf{y}_{min} + \mathbf{CA}^{j_1}\mathbf{x}_k \\ \vdots \\ -\mathbf{y}_{min} + \mathbf{CA}^{j_2}\mathbf{x}_k \end{bmatrix}}_{\mathbf{d}_2}. \quad (\text{B.57})$$

As Equações (B.56) e (B.57) são representadas da seguinte forma:

$$\begin{bmatrix} \mathbf{D} \\ -\mathbf{D} \end{bmatrix} \mathbf{u}^N \leq \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}. \quad (\text{B.58})$$

B.5 MPC Regulador com Restrições

O desenvolvimento do controle preditivo baseado em modelo e das restrições de forma independente é feito com o objetivo primordial enfatizar o desenvolvimento matemático requerido para a construção do problema de controle. Nesta seção, a dedução do problema de controle preditivo baseado em modelo com restrições é explorado para mostrar que o controlador é obtido pela solução de um problema de otimização formulado com o modelo linear e discreto de um processo. A eficiência desse tipo de controlador é dependente de uma representação precisa do processo pelo modelo.

Formalmente, o problema de controle preditivo regulador baseado em modelo é apre-

sentado como um problema de otimização:

$$\min_{\mathbf{u}^N} J_k = \mathbf{u}^{NT} (\Psi^T \Theta \Psi + \Xi + \Upsilon) \mathbf{u}^N + 2\mathbf{u}^{NT} (\Psi^T \Theta \Lambda \mathbf{x}_k - \mathbf{g} \mathbf{u}_{k-1}). \quad (\text{B.59})$$

Sujeito a:

$$\begin{bmatrix} \mathbf{I} \\ -\mathbf{I} \\ \mathbf{D} \\ -\mathbf{D} \\ \mathbf{T} \end{bmatrix} \mathbf{u}^N \leq \begin{bmatrix} \mathbf{i}_1 \\ \mathbf{i}_2 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \\ \mathbf{t}_1 \end{bmatrix}. \quad (\text{B.60})$$

A caracterização do controle preditivo baseado em modelo é resumida em uma única figura, na qual demonstra-se o comportamento deste tipo de controlador. Para horizontes de controle e predição fixos e iguais e utilizando o modelo do processo, são feitas predições utilizando o modelo a partir do instante presente k até o instante de predição $k + N$, onde N é definido como o horizonte de predição e controle. Dessa forma, utilizando como exemplo $k = 10$ e $N = 7$, tem-se o processo no décimo instante de amostragem e utiliza-se como horizonte de predição e controle um valor de sete unidades.

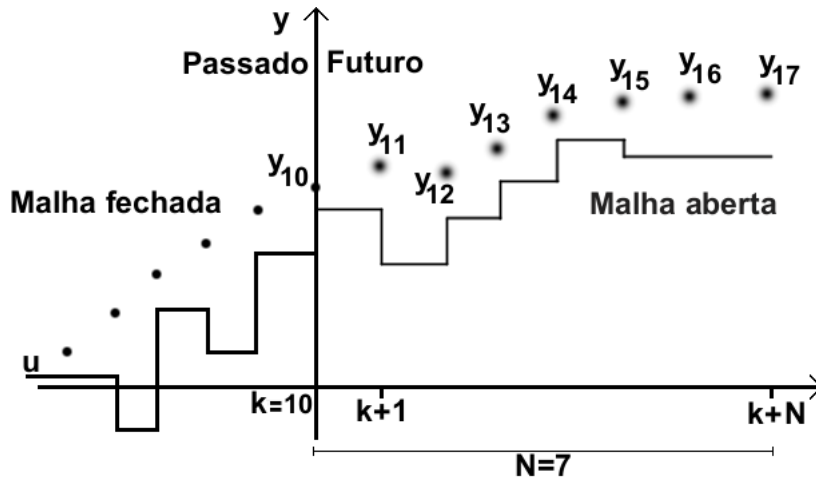


Figura B.1: Horizonte móvel do controle preditivo baseado em modelo.

A Figura (B.1) mostra como o controlador preditivo atua no processo a partir do décimo instante de amostragem. Neste caso, a medida do processo neste instante (y_{10}) de amostragem é obtida através dos instrumentos de medição. Entre os instantes 10 e 11 ocorrerá uma série de cálculos que utilizam o modelo do processo para prever o comportamento da planta no instante 11, ou seja, encontrar as variáveis manipuladas ótimas que levam o processo do instante 10 para o 11. Todo horizonte de predição é percorrido e apenas a primeira ação de controle predita pelo horizonte é enviada para o processo. Esse procedimento é realizado a cada instante de amostragem. Com isso, uma

interpretação que pode ser feita desse tipo de controlador é a de uma janela que se move com o tempo, prevendo N saídas da planta no futuro utilizando o modelo para cada novo instante implementado efetivamente apenas a primeira ação de controle de todo horizonte. De fato, o problema de otimização será resolvido ao longo de todo o horizonte de predição e controle, minimizando o erro entre a saída predita pelo modelo e o *set-point*.

B.6 MPC Servo com Restrições

O controle preditivo regulador apresentado anteriormente foi desenvolvido para um alvo zero. Se o controlador deve seguir um determinado valor de *set-point* diferente de zero, $\mathbf{y}_t$, então os vetores de estado e de entrada, $\mathbf{x}_s$ e $\mathbf{u}_s$, no estado estacionário são necessários para levar o sistema para $\mathbf{y}_t$ no estado estacionário. Dessa forma, a seguinte função objetivo quadrática é utilizada pelo regulador:

$$\begin{aligned} \min_{\mathbf{u}^N} J_k &= (\mathbf{x}_{k+N} - \mathbf{x}_s)^T \bar{\mathbf{Q}} (\mathbf{x}_{k+N} - \mathbf{x}_s) + \Delta \mathbf{u}_{k+N}^T \mathbf{S} \Delta \mathbf{u}_{k+N} + \\ &+ \sum_{h=0}^{N-1} [(\mathbf{x}_{k+h} - \mathbf{x}_s)^T \mathbf{C}^T \mathbf{Q} \mathbf{C} (\mathbf{x}_{k+h} - \mathbf{x}_s) + (\mathbf{u}_{k+h} - \mathbf{u}_s)^T \mathbf{R} (\mathbf{u}_{k+h} - \mathbf{u}_s) + \Delta \mathbf{u}_{k+h}^T \mathbf{S} \Delta \mathbf{u}_{k+h}]. \quad (\text{B.61}) \end{aligned}$$

Os vetores de estado estacionário $\mathbf{x}_s$ e $\mathbf{u}_s$ são calculados através da resolução de um problema de otimização quadrática.

As ponderações de estado e entrada nesta função objetivo penalizam desvios dos valores de alvo no estado estacionário. Além disso, o termo de penalidade $\mathbf{u} - \mathbf{u}_s$ é requerido juntamente com o termo $\mathbf{x} - \mathbf{x}_s$ de penalidade de estado para prevenir *offset* do regulador. Esta estratégia é equivalente a alterar a origem do sistema para o estado estacionário, descrito por $\mathbf{x}_s$ e $\mathbf{u}_s$ e usar o regulador apresentado anteriormente. De maneira diferente à apresentada na Seção B.1, o vetor $\mathbf{u}_{k+j}$ é igualado a $\mathbf{u}_s$ para todo $j \geq N$ nos cálculos da função objetivo em laço aberto.

Utilizando o mesmo desenvolvimento do controle tipo regulador dado pela Equação (B.5) para a Equação (B.61):

$$\min_{\mathbf{u}^N} J_k = [\mathbf{x}_{k+1} - \mathbf{x}_s \quad \mathbf{x}_{k+2} - \mathbf{x}_s \quad \cdots \quad \mathbf{x}_{k+N} - \mathbf{x}_s]^T \boldsymbol{\Theta} \begin{bmatrix} \mathbf{x}_{k+1} - \mathbf{x}_s \\ \mathbf{x}_{k+2} - \mathbf{x}_s \\ \vdots \\ \mathbf{x}_{k+N} - \mathbf{x}_s \end{bmatrix} +$$

$$\begin{aligned}
& + [\mathbf{u}_k - \mathbf{u}_s \quad \mathbf{u}_{k+1} - \mathbf{u}_s \quad \cdots \quad \mathbf{u}_{k+N-1} - \mathbf{u}_s]^T \Xi \begin{bmatrix} \mathbf{u}_k - \mathbf{u}_s \\ \mathbf{u}_{k+1} - \mathbf{u}_s \\ \vdots \\ \mathbf{u}_{k+2} - \mathbf{u}_s \end{bmatrix} + \\
& + [\mathbf{u}_k \quad \mathbf{u}_{k+1} \quad \cdots \quad \mathbf{u}_{k+N-1}]^T \Upsilon \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \vdots \\ \mathbf{u}_{k+2} \end{bmatrix} - [\mathbf{u}_k \quad \mathbf{u}_{k+1} \quad \cdots \quad \mathbf{u}_{k+N-1}]^T 2\tilde{\Upsilon} \begin{bmatrix} \mathbf{u}_s \\ \mathbf{u}_s \\ \vdots \\ \mathbf{u}_s \end{bmatrix} + \\
& - [\mathbf{u}_k \quad \mathbf{u}_{k+1} \quad \cdots \quad \mathbf{u}_{k+N-1}]^T \mathbf{g} \mathbf{u}_{k-1}, \tag{B.62}
\end{aligned}$$

em que:

$$\tilde{\Upsilon} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{S} \end{bmatrix}.$$

Na Equação (B.62), utiliza-se a mesma notação para as matrizes apresentadas anteriormente. Assim:

$$\begin{aligned}
\min_{\mathbf{u}^N} J_k &= (\mathbf{x}^N - \mathbf{x}_s^N)^T \Theta (\mathbf{x}^N - \mathbf{x}_s^N) + (\mathbf{u}^N - \mathbf{u}_s^N)^T \Xi (\mathbf{u}^N - \mathbf{u}_s^N) + \\
& + \mathbf{u}^{NT} \Upsilon \mathbf{u}^N - \mathbf{u}^{NT} 2\mathbf{g} \mathbf{u}_{k-1} - \mathbf{u}^{NT} 2\tilde{\Upsilon} \mathbf{u}_s^N. \tag{B.63}
\end{aligned}$$

A conexão da Equação (B.63) com o modelo é feita através da Equação (B.28). A expansão de (B.63) quando o modelo (B.63) é inserido fornece:

$$\begin{aligned}
\min_{\mathbf{u}^N} J_k &= \mathbf{x}_k^T \Lambda^T \Theta \Lambda \mathbf{x}_k + \mathbf{x}_k^T \Lambda^T \Theta \Psi \mathbf{u}^N - \mathbf{x}_k^T \Lambda^T \Theta \mathbf{x}_s^N + \mathbf{u}^{NT} \Psi^T \Theta \Lambda \mathbf{x}_k + \\
& + \mathbf{u}^{NT} \Psi^T \Theta \Psi \mathbf{u}^N - \mathbf{u}^{NT} \Psi^T \Theta \mathbf{x}_s^N - \mathbf{x}_s^{NT} \Theta \Lambda \mathbf{x}_k - \mathbf{x}_s^{NT} \Theta \Psi \mathbf{u}^N + \mathbf{x}_s^{NT} \Theta \mathbf{x}_s^N + \\
& + \mathbf{u}^{NT} \Xi \mathbf{u}^N - \mathbf{u}^{NT} \Xi \mathbf{u}_s^N - \mathbf{u}_s^{NT} \Xi \mathbf{u}^N + \mathbf{u}_s^{NT} \Xi \mathbf{u}_s^N + \mathbf{u}^{NT} \Upsilon \mathbf{u}^N - \mathbf{u}^{NT} 2\mathbf{g} \mathbf{u}_{k-1} - \mathbf{u}^{NT} 2\tilde{\Upsilon} \mathbf{u}_s^N. \tag{B.64}
\end{aligned}$$

Cancelando alguns termos da Equação (B.64) chega-se à função objetivo quadrática que dever ser minimizada para o problema de controle preditivo servo baseado em modelo, que é diferente da Equação (B.31) por apresentar termos referentes ao estado estacionário que o controlador deve levar o sistema.

Restrições da forma apresentada por (B.32)-(B.34) podem ser anexadas ao problema

de otimização (B.64), formando assim:

$$\min_{\mathbf{u}^N} J_k = \mathbf{u}^{NT} (\Psi^T \Theta \Psi + \Xi + \Upsilon) \mathbf{u}^N + 2\mathbf{u}^{NT} (\Psi^T \Theta \Lambda \mathbf{x}_k - \Psi^T \Theta \mathbf{x}_s^N - \Xi \mathbf{u}_s^N - \mathbf{g} \mathbf{u}_{k-1} - \tilde{\Upsilon} \mathbf{u}_s^N). \quad (\text{B.65})$$

Sujeito a:

$$\begin{bmatrix} \mathbf{I} \\ -\mathbf{I} \\ \mathbf{D} \\ -\mathbf{D} \\ \mathbf{T} \end{bmatrix} \mathbf{u}^N \leq \begin{bmatrix} \mathbf{i}_1 \\ \mathbf{i}_2 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \\ \mathbf{t}_1 \end{bmatrix}. \quad (\text{B.66})$$

B.7 MPC Regulador Sucessivamente Linearizado

Na busca por uma representação mais precisa de modelos não lineares utilizando modelos linearizados, utiliza-se a linearização local sucessiva, que em contraste com a metodologia bastante utilizada, em que o modelo não linear é linearizado em um ponto estacionário, utiliza-se agora uma descrição do comportamento do processo de forma local e realizada a cada instante de amostragem. Em cada instante de amostragem o procedimento de linearização é realizado baseado no ponto de amostragem imediatamente anterior ao corrente e um modelo linear atualizado do processo não linear é obtido. Este é válido apenas para aquele instante e é utilizado no projeto do controlador.

Para fins de projeto do controlador preditivo, algumas deduções realizadas nas seções anteriores devem ser revisadas para adequá-las ao modelo utilizado, que é escrito no espaço de estados e possui forma conveniente para representar o comportamento do processo em cada instante de amostragem k e ao longo de um horizonte de predição h . Para modelos sucessivamente linearizados, demonstra-se no Apêndice A, que estes possuem a seguinte forma:

$$\mathbf{x}_{k+1} = \Phi_{k-1} \mathbf{x}_k + \Gamma_{k-1} \mathbf{u}_k + \Omega_{k-1} \mathbf{f}_{k-1}, \quad (\text{B.67})$$

$$\mathbf{y}_k = \mathbf{C}_{k-1} \mathbf{x}_k. \quad (\text{B.68})$$

Para adequar o modelo ao projeto de um controlador preditivo desenvolvido, realizam-se as seguinte modificações:

$$\mathbf{x}_{k+1+h|k} = \mathbf{A}_{k-1} \mathbf{x}_{k+h|k} + \mathbf{B}_{k-1} \mathbf{u}_{k+h|k} + \mathbf{v}_{k-1}, \quad h = 0, 1, 2, \dots, N \quad (\text{B.69})$$

$$\mathbf{y}_{k+h|k} = \mathbf{C}_{k-1} \mathbf{x}_{k+h|k}, \quad h = 0, 1, 2, \dots, N \quad (\text{B.70})$$

em que $\mathbf{A}_{k-1} = \Phi_{k-1}$, $\mathbf{B}_{k-1} = \Gamma_{k-1}$ e $\mathbf{v}_{k-1} = \Omega_{k-1} \mathbf{f}_{k-1}$.

Nas equações (B.69) e (B.70), $\mathbf{x}$ representa o vetor dos estados, $\mathbf{u}$ o de variáveis manipuladas e $\mathbf{y}$ o de saídas em variável desvio com relação ao ponto anterior ao ponto corrente, ou seja, $k - 1$.

A Equação (B.69) ainda pode ser escrita da seguinte forma:

$$\mathbf{x}_{k+1+h}|_k = \mathbf{A}_{k-1}\mathbf{x}_{k+h}|_k + [\mathbf{B}_{k-1} \ \mathbf{v}_{k-1}] \begin{bmatrix} \mathbf{u}_{k+h}|_k \\ 1 \end{bmatrix}, \quad h = 0, 1, 2, \dots, N \quad (\text{B.71})$$

ou

$$\mathbf{x}_{k+1+h}|_k = \mathbf{A}_{k-1}\mathbf{x}_{k+h}|_k + \mathbf{H}_{k-1}\mathbf{w}_{k+h}|_k, \quad h = 0, 1, 2, \dots, N. \quad (\text{B.72})$$

A Equação (B.72) é utilizada para o controle preditivo baseado em modelo, sendo válida apenas quando o último elemento de $\mathbf{w}$, calculado para cada instante de amostragem e horizonte de predição, é igual a 1. Para que isto seja válido, utiliza-se a vantagem da resolução do problema de otimização sujeito a restrições do próprio controlador para limitar este valor para 1.

O horizonte de controle móvel se baseia na minimização do seguinte horizonte quadrático infinito em malha aberta, cuja função objetivo calculada em um instante de amostragem qualquer k é

$$\min_{\mathbf{w}^N} \sum_{h=0}^{\infty} (\mathbf{y}_{k+h}|_k^T \mathbf{Q} \mathbf{y}_{k+h}|_k + \mathbf{w}_{k+h}|_k^T \tilde{\mathbf{R}} \mathbf{w}_{k+h}|_k + \Delta \mathbf{w}_{k+h}|_k^T \tilde{\mathbf{S}} \Delta \mathbf{w}_{k+h}|_k). \quad (\text{B.73})$$

A função objetivo do horizonte de controle infinito em malha aberta na Equação (B.73) pode ser expressa como uma função objetivo de horizonte de controle finito em malha aberta. Para isto é necessário dividir o somatório em duas partes e substituir a Equação (B.70) na (B.72) para obter:

$$\begin{aligned} \min_{\mathbf{w}^N} J_k &= \mathbf{x}_{k+N}^T \bar{\mathbf{Q}} \mathbf{x}_{k+N} + \Delta \mathbf{w}_{k+N}^T \tilde{\mathbf{S}} \Delta \mathbf{w}_{k+N} + \\ &+ \sum_{h=0}^{N-1} (\mathbf{x}_{k+h}^T \mathbf{C}_{k-1}^T \mathbf{Q} \mathbf{C}_{k-1} \mathbf{x}_{k+h} + \mathbf{w}_{k+h}^T \tilde{\mathbf{R}} \mathbf{w}_{k+h} + \Delta \mathbf{w}_{k+h}^T \tilde{\mathbf{S}} \Delta \mathbf{w}_{k+h}). \end{aligned} \quad (\text{B.74})$$

Assim como apresentado anteriormente, para N horizontes de predição, a seguinte

função objetivo deve ser minimizada em $\mathbf{w}$:

$$\begin{aligned}
\min_{\mathbf{w}^N} J_k = & \underbrace{\begin{bmatrix} \mathbf{x}_{k+1} & \mathbf{x}_{k+2} & \cdots & \mathbf{x}_{k+N} \end{bmatrix}^T}_{\mathbf{x}^{NT}} \underbrace{\begin{bmatrix} \mathbf{C}_{k-1}^T \mathbf{Q} \mathbf{C}_{k-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{k-1}^T \mathbf{Q} \mathbf{C}_{k-1} & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \bar{\mathbf{Q}}_{k-1} \end{bmatrix}}_{\boldsymbol{\Theta}_{k-1}} \begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{x}_{k+2} \\ \vdots \\ \mathbf{x}_{k+N} \end{bmatrix} + \\
& + \underbrace{\begin{bmatrix} \mathbf{w}_k & \mathbf{w}_{k+1} & \cdots & \mathbf{w}_{k+N-1} \end{bmatrix}^T}_{\mathbf{w}^{NT}} \underbrace{\begin{bmatrix} \tilde{\mathbf{R}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{R}} & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \tilde{\mathbf{R}} \end{bmatrix}}_{\boldsymbol{\Sigma}} \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k+1} \\ \vdots \\ \mathbf{w}_{k+N-1} \end{bmatrix} + \\
& + \underbrace{\begin{bmatrix} 2\tilde{\mathbf{S}} & -\tilde{\mathbf{S}} & \cdots & \mathbf{0} \\ -\tilde{\mathbf{S}} & 2\tilde{\mathbf{S}} & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & -\tilde{\mathbf{S}} \\ \mathbf{0} & \mathbf{0} & -\tilde{\mathbf{S}} & 2\tilde{\mathbf{S}} \end{bmatrix}}_{\boldsymbol{\Pi}} \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k+1} \\ \vdots \\ \mathbf{w}_{k+N-1} \end{bmatrix} + \\
& - \underbrace{\begin{bmatrix} \tilde{\mathbf{S}} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}}_{\mathbf{f}} \begin{bmatrix} \mathbf{w}_k & \mathbf{w}_{k+1} & \cdots & \mathbf{w}_{k+N-1} \end{bmatrix}^T 2, \tag{B.75}
\end{aligned}$$

em que a solução da equação discreta de Lyapunov toma a forma:

$$\bar{\mathbf{Q}}_{k-1} = \mathbf{C}_{k-1}^T \mathbf{Q} \mathbf{C}_{k-1} + \mathbf{A}_{k-1}^T \bar{\mathbf{Q}} \mathbf{A}_{k-1}. \tag{B.76}$$

Apresentando a Equação (B.75) de acordo com a notação adotada acima, tem-se:

$$\min_{\mathbf{w}^N} J_k = \mathbf{x}^{NT} \boldsymbol{\Theta}_{k-1} \mathbf{x}^N + \mathbf{w}^{NT} \boldsymbol{\Sigma} \mathbf{w}^N + \mathbf{w}^{NT} \boldsymbol{\Pi} \mathbf{w}^N - \mathbf{w}^{NT} 2\mathbf{f} \mathbf{w}_{k-1}. \tag{B.77}$$

O vetor $\mathbf{w}^N$ contém as N ações de controle futuras em malha aberta.

$$\mathbf{w}^N = \begin{bmatrix} \mathbf{w}_k|_k \\ \mathbf{w}_{k+1}|_k \\ \vdots \\ \mathbf{w}_{k+N-1}|_k \end{bmatrix}. \tag{B.78}$$

A única restrição é que o último elemento de cada vetor $\mathbf{w}_k$ seja igual a 1 para que a Equação (B.71) seja satisfeita.

B.7.1 Restrições aplicadas ao modelo sucessivamente linearizado

As restrições que devem ser aplicadas ao problema de otimização obtido quando são utilizados modelos sucessivamente linearizados (Equação (B.77)) são sensivelmente modificadas. Considera-se restrições para as variáveis de entrada, variação da entrada e saída do processo:

$$\mathbf{w}_{min} \leq \mathbf{w}_{k+h}|_k \leq \mathbf{w}_{max}, \quad h = 0, 1, \dots, N-1 \quad (\text{B.79})$$

$$\Delta \mathbf{w}_{k+h}|_k \leq |\Delta \mathbf{w}_{max}|, \quad h = 0, 1, \dots, N-1 \quad (\text{B.80})$$

$$\mathbf{y}_{min} \leq \mathbf{y}_{k+h}|_k \leq \mathbf{y}_{max}. \quad h = h_1, h_1 + 1, \dots, h_2 \quad (\text{B.81})$$

De forma análoga àquela apresenta na Seção B.4, demonstra-se que a restrição na variável de entrada ($\mathbf{w}$) gera o seguinte resultado:

$$\begin{array}{ccc} \mathbf{w}_k \leq -\mathbf{w}_{min} & & \mathbf{w}_k \leq \mathbf{w}_{max} \\ \mathbf{w}_{k+1} \leq -\mathbf{w}_{min} & \text{e} & \mathbf{w}_{k+1} \leq \mathbf{w}_{max} \\ \vdots & & \vdots \\ \underbrace{\mathbf{w}_{k+N-1}}_{\mathbf{w}^N} \leq \underbrace{-\mathbf{w}_{min}}_{\mathbf{k}_2} & & \underbrace{\mathbf{w}_{k+N-1}}_{\mathbf{w}^N} \leq \underbrace{\mathbf{w}_{max}}_{\mathbf{k}_1} \end{array} \quad (\text{B.82})$$

Representando na forma matricial as duas partes da Equação (B.82) chega-se a:

$$\begin{bmatrix} \mathbf{I} \\ -\mathbf{I} \end{bmatrix} \mathbf{w}^N \leq \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \end{bmatrix}. \quad (\text{B.83})$$

Restringir a variação da entrada (Equação (B.80)) é semelhante ao que foi feito na Equação (B.33), logo os resultados anteriormente obtidos podem ser diretamente genera-

lizados para este caso.

$$\underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ -1 & 1 & \cdots & 0 \\ \vdots & -1 & \ddots & \vdots \\ 0 & \cdots & \ddots & 1 \\ 0 & \cdots & 0 & -1 \end{bmatrix}}_{\mathbf{T}} \underbrace{\begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k+1} \\ \mathbf{w}_{k+2} \\ \vdots \\ \mathbf{w}_{k+N-1} \end{bmatrix}}_{\mathbf{w}^N} \leq \underbrace{\begin{bmatrix} |\Delta \mathbf{w}_{max}| + \mathbf{w}_{k-1} \\ |\Delta \mathbf{w}_{max}| \\ |\Delta \mathbf{w}_{max}| \\ \vdots \\ |\Delta \mathbf{w}_{max}| \end{bmatrix}}_{\mathbf{l}_1}. \quad (\text{B.84})$$

ou

$$\mathbf{T} \mathbf{w}^N \leq \mathbf{l}_1. \quad (\text{B.85})$$

Restrições aplicadas para a saída do processo são mais difíceis de serem atendidas a partir do instante inicial pois pode não existir solução para o problema de otimização. Procedendo de maneira semelhante a apresentada na Seção B.4 chega-se a:

$$\underbrace{\begin{bmatrix} \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_1-1} \mathbf{H}_{k-1} & \cdots & \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_1-N} \mathbf{B}_{k-1} \\ \vdots & & \vdots \\ \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_2-1} \mathbf{H}_{k-1} & \cdots & \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_2-N} \mathbf{H}_{k-1} \end{bmatrix}}_{\mathbf{M}} \underbrace{\begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+N-1} \end{bmatrix}}_{\mathbf{w}^N} \leq \underbrace{\begin{bmatrix} \mathbf{y}_{max} - \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_1} \mathbf{x}_k \\ \vdots \\ \mathbf{y}_{max} - \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_2} \mathbf{x}_k \end{bmatrix}}_{\mathbf{m}_1} \quad (\text{B.86})$$

e

$$-1 \underbrace{\begin{bmatrix} \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_1-1} \mathbf{H}_{k-1} & \cdots & \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_1-N} \mathbf{H}_{k-1} \\ \vdots & & \vdots \\ \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_2-1} \mathbf{H}_{k-1} & \cdots & \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_2-N} \mathbf{H}_{k-1} \end{bmatrix}}_{\mathbf{M}} \underbrace{\begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+N-1} \end{bmatrix}}_{\mathbf{w}^N} \leq \underbrace{\begin{bmatrix} -\mathbf{y}_{min} + \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_1} \mathbf{x}_k \\ \vdots \\ -\mathbf{y}_{min} + \mathbf{C}_{k-1} \mathbf{A}_{k-1}^{j_2} \mathbf{x}_k \end{bmatrix}}_{\mathbf{m}_2}. \quad (\text{B.87})$$

As Equações (B.86) e (B.87) podem ser representadas da seguinte forma:

$$\begin{bmatrix} \mathbf{M} \\ -\mathbf{M} \end{bmatrix} \mathbf{w}^N \leq \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{bmatrix}. \quad (\text{B.88})$$

B.7.2 Conexão entre modelo e função objetivo

O modelo do processo dado pela Equação (B.1) se relaciona com a função objetivo (B.20) da mesma maneira, mas utiliza-se agora o modelo sucessivamente linearizado dado pela Equação (B.72). Por exemplo, para um horizonte de predição $N = 3$, o seguinte resultado

é obtido:

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{x}_{k+2} \\ \mathbf{x}_{k+3} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{k-1} \\ \mathbf{A}_{k-1}^2 \\ \mathbf{A}_{k-1}^3 \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} \mathbf{H}_{k-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_{k-1}\mathbf{H}_{k-1} & \mathbf{H}_{k-1} & \mathbf{0} \\ \mathbf{A}_{k-1}^2\mathbf{H}_{k-1} & \mathbf{A}_{k-1}\mathbf{H}_{k-1} & \mathbf{H}_{k-1} \end{bmatrix} \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k+1} \\ \mathbf{w}_{k+2} \end{bmatrix}. \quad (\text{B.89})$$

Generalizando para um horizonte N qualquer:

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{x}_{k+2} \\ \vdots \\ \mathbf{x}_{k+N} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{A}_{k-1} \\ \mathbf{A}_{k-1}^2 \\ \vdots \\ \mathbf{A}_{k-1}^N \end{bmatrix}}_{\mathbf{\Lambda}_{k-1}} \mathbf{x}_k + \underbrace{\begin{bmatrix} \mathbf{H}_{k-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{A}_{k-1}\mathbf{H}_{k-1} & \mathbf{H}_{k-1} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{A}_{k-1}^{N-1}\mathbf{H}_{k-1} & \mathbf{A}_{k-1}^{N-2}\mathbf{H}_{k-1} & \cdots & \mathbf{H}_{k-1} \end{bmatrix}}_{\mathbf{\Psi}_{k-1}} \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k+1} \\ \vdots \\ \mathbf{w}_{k+N-1} \end{bmatrix}. \quad (\text{B.90})$$

A Equação (B.90) é representada na forma:

$$\mathbf{x}^N = \mathbf{\Lambda}_{k-1}\mathbf{x}_k + \mathbf{\Psi}_{k-1}\mathbf{w}^N, \quad (\text{B.91})$$

que representa o modelo do processo expandido para um horizonte de predição N para um instante de tempo qualquer k . Sendo assim, substituindo (B.91) em (B.77) chega-se a:

$$\begin{aligned} \min_{\mathbf{w}^N} J_k &= (\mathbf{x}_k^T \mathbf{\Lambda}_{k-1}^T + \mathbf{w}^{NT} \mathbf{\Psi}_{k-1}^T) \mathbf{\Theta}_{k-1} (\mathbf{\Lambda}_{k-1} \mathbf{x}_k + \mathbf{\Psi}_{k-1} \mathbf{w}^N) + \\ &+ \mathbf{w}^{NT} \mathbf{\Sigma} \mathbf{w}^N + \mathbf{w}^{NT} \mathbf{\Pi} \mathbf{w}^N - \mathbf{w}^{NT} 2\mathbf{f} \mathbf{w}_{k-1}. \end{aligned} \quad (\text{B.92})$$

Expandindo a Equação (B.92):

$$\begin{aligned} \min_{\mathbf{w}^N} J_k &= (\mathbf{x}_k^T \mathbf{\Lambda}_{k-1}^T \mathbf{\Theta}_{k-1} \mathbf{\Lambda}_{k-1} \mathbf{x}_k + \underbrace{\mathbf{x}_k^T \mathbf{\Lambda}_{k-1}^T \mathbf{\Theta}_{k-1} \mathbf{\Psi}_{k-1}}_{\Delta} \mathbf{w}^N + \underbrace{\mathbf{w}^{NT} \mathbf{\Psi}_{k-1}^T \mathbf{\Theta}_{k-1} \mathbf{\Lambda}_{k-1} \mathbf{x}_k}_{\Delta} + \\ &+ \underbrace{\mathbf{w}^{NT} \mathbf{\Psi}_{k-1}^T \mathbf{\Theta}_{k-1} \mathbf{\Psi}_{k-1} \mathbf{w}^N}_{\Delta} + \underbrace{\mathbf{w}^{NT} \mathbf{\Sigma} \mathbf{w}^N}_{\Delta} + \underbrace{\mathbf{w}^{NT} \mathbf{\Pi} \mathbf{w}^N}_{\Delta} - \mathbf{w}^{NT} 2\mathbf{f} \mathbf{w}_{k-1}. \end{aligned} \quad (\text{B.93})$$

Na Equação (B.93), os termos que não contém o vetor $\mathbf{w}^N$ não influenciam o processo de otimização, visto que a minimização ocorre em $\mathbf{w}^N$. Assim, todas as outras variáveis são consideradas constantes. Nesta mesma equação, os termos marcados com o símbolo Δ podem ser agrupados já que a matriz $\mathbf{\Theta}_{k-1}$ é simétrica, e os termos marcados com o símbolo Δ são diretamente agrupados. O resultado deste agrupamento resulta na minimização da seguinte função objetivo quadrática:

$$\min_{\mathbf{w}^N} J_k = \mathbf{w}^{NT} (\mathbf{\Psi}_{k-1}^T \mathbf{\Theta}_{k-1} \mathbf{\Psi}_{k-1} + \mathbf{\Sigma} + \mathbf{\Pi}) \mathbf{w}^N + 2\mathbf{w}^{NT} (\mathbf{\Psi}_{k-1}^T \mathbf{\Theta}_{k-1} \mathbf{\Lambda}_{k-1} \mathbf{x}_k - \mathbf{f} \mathbf{w}_{k-1}). \quad (\text{B.94})$$

sujeito a:

$$\begin{bmatrix} \mathbf{I} \\ -\mathbf{I} \\ \mathbf{M} \\ -\mathbf{M} \\ \mathbf{T} \end{bmatrix} \mathbf{w}^N \leq \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \\ \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{l}_1 \end{bmatrix}. \quad (\text{B.95})$$

B.8 MPC Servo Sucessivamente Linearizado

Assim como apresentado anteriormente, utiliza-se o modelo sucessivamente linearizado do processo para projetar um controlador preditivo servo. O modelo utilizado neste trabalho é o mesmo proposto pela Equação (B.69) e (B.70), sendo que as mesmas simplificações são feitas para que se obtenha um modelo discreto, como na Equação (B.72).

A função objetivo a ser minimizada (Equação (B.96)) é semelhante à Equação (B.61):

$$\begin{aligned} \min_{\mathbf{w}^N} J_k &= (\mathbf{x}_{k+N} - \mathbf{x}_{ks})^T \bar{\mathbf{Q}}_{k-1} (\mathbf{x}_{k+N} - \mathbf{x}_{ks}) + \Delta \mathbf{w}_{k+N}^T \tilde{\mathbf{S}} \Delta \mathbf{w}_{k+N} + \\ &+ \sum_{h=0}^{N-1} (\mathbf{x}_{k+h} - \mathbf{x}_{ks})^T \mathbf{C}_{k-1}^T \mathbf{Q}_{k-1} \mathbf{C}_{k-1} (\mathbf{x}_{k+h} - \mathbf{x}_{ks}) + (\mathbf{w}_{k+h} - \mathbf{w}_{ks})^T \tilde{\mathbf{R}} (\mathbf{w}_{k+h} - \mathbf{w}_{ks}) + \Delta \mathbf{w}_{k+h}^T \tilde{\mathbf{S}} \Delta \mathbf{w}_{k+h}. \end{aligned} \quad (\text{B.96})$$

Utilizando o mesmo desenvolvimento do regulador dado pela Equação (B.62) para a Equação (B.96) e levando em consideração que para $j \geq N$ o vetor $\mathbf{w}_{k+j}$ é igual a $\mathbf{w}_s$, tem-se:

$$\begin{aligned} \min_{\mathbf{w}^N} J_k &= [\mathbf{x}_{k+1} - \mathbf{x}_s \quad \mathbf{x}_{k+2} - \mathbf{x}_s \quad \cdots \quad \mathbf{x}_{k+N} - \mathbf{x}_s]^T \boldsymbol{\Theta}_{k-1} \begin{bmatrix} \mathbf{x}_{k+1} - \mathbf{x}_s \\ \mathbf{x}_{k+2} - \mathbf{x}_s \\ \vdots \\ \mathbf{x}_{k+N} - \mathbf{x}_s \end{bmatrix} + \\ &+ [\mathbf{w}_k - \mathbf{w}_s \quad \mathbf{w}_{k+1} - \mathbf{w}_s \quad \cdots \quad \mathbf{w}_{k+N-1} - \mathbf{w}_s]^T \boldsymbol{\Sigma} \begin{bmatrix} \mathbf{w}_k - \mathbf{w}_s \\ \mathbf{w}_{k+1} - \mathbf{w}_s \\ \vdots \\ \mathbf{w}_{k+N-1} - \mathbf{w}_s \end{bmatrix} + \end{aligned}$$

$$\begin{aligned}
& + [\mathbf{w}_k \ \mathbf{w}_{k+1} \ \cdots \ \mathbf{w}_{k+N-1}]^T \mathbf{\Pi} \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k+1} \\ \vdots \\ \mathbf{w}_{k+2} \end{bmatrix} - [\mathbf{w}_k \ \mathbf{w}_{k+1} \ \cdots \ \mathbf{w}_{k+N-1}]^T 2\tilde{\mathbf{\Pi}} \underbrace{\begin{bmatrix} \mathbf{w}_s \\ \mathbf{w}_s \\ \vdots \\ \mathbf{w}_s \end{bmatrix}}_{\mathbf{w}_s^N} + \\
& - [\mathbf{w}_k \ \mathbf{w}_{k+1} \ \cdots \ \mathbf{w}_{k+N-1}]^T 2\mathbf{f}\mathbf{w}_{k-1}, \tag{B.97}
\end{aligned}$$

com:

$$\tilde{\mathbf{\Pi}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \tilde{\mathbf{S}} \end{bmatrix}.$$

A Equação (B.97) pode ser escrita da seguinte forma:

$$\begin{aligned}
\min_{\mathbf{w}^N} J_k &= (\mathbf{x}^N - \mathbf{x}_s^N)^T \mathbf{\Theta}_{k-1} (\mathbf{x}^N - \mathbf{x}_s^N) + (\mathbf{w}^N - \mathbf{w}_s^N)^T \mathbf{\Sigma} (\mathbf{w}^N - \mathbf{w}_s^N) + \\
& + \mathbf{w}^{NT} \mathbf{\Pi} \mathbf{w}^N - \mathbf{w}^{NT} 2\mathbf{f}\mathbf{w}_{k-1} - \mathbf{w}^{NT} 2\tilde{\mathbf{\Pi}} \mathbf{w}_s^N. \tag{B.98}
\end{aligned}$$

Após a substituição de (B.91) em (B.98) e a realização de algumas manipulações algébricas chega-se a seguinte minimização da função objetivo em $\mathbf{w}^N$:

$$\begin{aligned}
\min_{\mathbf{w}^N} J_k &= \mathbf{w}^{NT} (\mathbf{\Psi}_{k-1}^T \mathbf{\Theta}_{k-1} \mathbf{\Psi}_{k-1} + \mathbf{\Sigma} + \mathbf{\Pi}) \mathbf{w}^N + \\
& + 2\mathbf{w}^{NT} (\mathbf{\Psi}_{k-1}^T \mathbf{\Theta}_{k-1} \mathbf{\Lambda}_{k-1} \mathbf{x}_k - \mathbf{\Psi}_{k-1}^T \mathbf{\Theta}_{k-1} \mathbf{x}_s^N - \mathbf{\Sigma} \mathbf{w}_s^N - \mathbf{f}\mathbf{w}_{k-1} - \tilde{\mathbf{\Pi}} \mathbf{w}_s^N). \tag{B.99}
\end{aligned}$$

sujeito a:

$$\begin{bmatrix} \mathbf{I} \\ -\mathbf{I} \\ \mathbf{M} \\ -\mathbf{M} \\ \mathbf{T} \end{bmatrix} \mathbf{w}^N \leq \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \\ \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{l}_1 \end{bmatrix}. \tag{B.100}$$

REFERÊNCIAS BIBLIOGRÁFICAS

- BERGE, C. Two Theorems in Graph Theory. Princeton University, 1957.
- BLANKE, M. Consistent design of dependable control systems. *Control Engineering Practice*, v. 4, n. 9, p. 1305-1312, 1996.
- BLANKE, M.; KINNAERT, M.; LUNZE, J.; STAROSWIECKI. Fault Diagnosis and Fault-Tolerant Control, 2ª Edição, Springer, 2006.
- CAMPBELL S. L; CHANCELIER, J-P; NIKOUKHAH, R. Modeling and Simulation in Scilab/Scicos, Springer, 2006.
- CHARTRAND, G.; OELLERMANN, O. R. Applied and algorithmic graph theory. Pure and applied mathematics. McGraw-Hill Inc., 1993
- CHENG, Y. Q.; ZHUANG, Y. M.; YANG, J. Y. Optimal fisher discriminant analysis using the rank decomposition. *Pattern Recognition*, v. 25, p. 101-111, 1992.
- CHIANG, L. H.; RUSSEL, E. L.; BRAATZ, R. D. Fault Detection and Diagnosis in Industrial Systems, Springer, 2001.
- GANI, A.; MHASKAR, P.; CHRISTOFIDES, P. D. Fault-tolerant control of a polyethylene reactor. *Journal of Process Control*, 2006.
- DECLERCK, P.; STAROSWIECKI, M. Characterisation of the canonical components of a structural graph for fault detection in large scale industrial plants. PROCEEDINGS OF EUROPEAN CONTROL CONFERENCE, Grenoble, 1991.
- DOWNS, J.J.; VOGEL, E. A plant-wide industrial process control problem. *Computers and Chemical Engineering*, v. 17, p. 245-255, 1993.

- DUDA, R. O.; HART, P. E.; STORK, D. G. Pattern Classification, John Wiley & Sons, New York, 1973.
- DULMAGE, A. L.; MENDELSON, N. S. Covering of bi-partite graphs. *Canada J. Math.*, v. 10, p. 517-534, 1958.
- DULMAGE, A. L.; MENDELSON, N. S. A structure theory of bi-partite graphs of finite exterior dimension. *Trans. of Royal Soc. Canada*, Section III, v. 53, p. 1-13, 1959.
- GEHIN, A. L.; ASSAS, M.; STAROSWIECKI, M. Structural analysis of system reconfigurability. IFAC SYMPOSIUM ON FAULT DETECTION SUPERVISION AND SAFETY FOR TECHNICAL PROCESSES, p. 292-297, Budapest, 2000.
- GERTLER, J.; SINGER, D. A new structural framework for parity space equation based failure detection and isolation. *Automatica*, v. 26, p. 381-388, 1990.
- GOLUB, G. H.; VAN LOAN, C.F. Matrix Computations, 3^a Edição., The Johns Hopkins University Press, Baltimore, 1996.
- GUSTAFSSON, T. K. et al. Modeling of pH for control. *Ind. Eng. Chem. Res.* v. 34, p. 820-827, 1995.
- GUSTAFSSON, T. K.; WALLER, K. V. Dynamics modeling and reaction invariant control of pH. *Chemical Engineering Science*, v. 38, n. 3, p. 389-398, 1983.
- HANSELMAN, D.; LITTLEFIELD, B. The Student Edition of MATLAB: Version 5, User's guide. Prentice Hall, New Jersey, 1997.
- HIMES, D.M.; STORER, R.H.; GEORGAKIS, C. Determination of the Number of Principal Components for Disturbance Detection and Isolation. PROCEEDINGS OF THE 1994 AMERICAN CONTROL CONFERENCE, Baltimore, MD, 1994.
- HONG, Z. Q.; YANG, J. Y. Optimal discriminant plane for a small number of samples and design method of classifier on the plane, *PR(24)*, n. 4, pp. 317-324, 1991.
- HUDLET R.; JOHNSON, R. Linear Discrimination and Some Further Results on Best Lower Dimensional Representation. *In Classification and Clustering*, editado por J. Van Ryzin, p. 371-394, Academic Press, New York, 1977.
- ISERMANN, R.; BALLÉ, P. Trends in the application of model based fault detection and diagnosis of technical processes. IN PROCEEDINGS OF THE 13th IFAC WORLD CONGRESS, v. N, p. 112, Piscataway, New Jersey, 1996.
- JACKSON, J.E. A user's guide to principal components. New York: Wiley, 1991.

-
- JACKSON, J.E. Quality Control Methods for Several Related Variables, *Technometrics*, v. 1, n. 4, p. 359-377, 1959.
- JACKSON, J. E.; MUDHOLKAR, G. S. Control procedures for residuals associated with principal component analysis. *Technometrics*, v. 21, p. 341-349, 1979.
- KAUFMAN, H.; BARKANA, I; SOBEL, K. Direct adaptative control algorithms, 2^a Edição, Springer, 1998.
- KLATT, K.U.; ENGELL, S. Nonlinear Control of a Non-Minimum-Phase CSTR. PROCEEDINGS OF AMERICAN CONTROL CONFERENCE, p. 2041-2045, Los Angeles, 1993.
- KOSANOVICH et al. Multi-Way PCA Applied to an Industrial Batch Process. PROCEEDINGS OF AMERICAN CONTROL CONFERENCE, p. 1294-1298, 1994.
- KU, W.; STORER R. H.; GEORGAKIS, C. Isolation of Disturbances in Statistical Process Control by Use of Approximate Models. PRESENTED AT THE 1993 AIChE ANNUAL MEETING, St. Louis, MO, 1993.
- KU, W.; STORER, R.H.; GEORGAKIS, C. Disturbance Detection and Isolation by Dynamic Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems*, 1995.
- LEE, C.; Choi, S. W.; LEE, I. Variable reconstruction and sensor fault identification using canonical variate analysis. *Journal of Process Control*, v. 16, p. 747-761, 2006.
- LEWIS, F.L.; SYRMOS, V. Optimal Control. John Wiley & Sons, 1995.
- MUROTA, K. Systems analysis by graphs and matroids. Structural solvability and controllability. Springer Verlag, 1987.
- MUSKE, K. R.; RAWLINGS, J. B. Model Predictive Control with Linear Models. *AIChE Journal*, v. 39, n. 2, p. 262-287, 1993.
- NOURA, H.; AUBRUN C.; SAUTER, D.; ROBERT, M. A Fault Diagnosis and Reconfiguration Method Applied to a Thermal Plant. TOOLDIAG'93, p. 995-999, Toulouse, 1993.
- OGATA, K. Discrete-Time Control Systems, 2nd ed, PTR Prentice Hall, Upper Saddle River, N.J., 1995.
- OGUNNAIKE, B.A.; RAY, W.H. Process Dynamics, Modeling and Control, Oxford University Press, New York, NY, 1994.

- PATTON, R. J. Fault-Tolerant Control Systems: The 1997 Situation. PROCEEDINGS IFAC SAFEPROCESS, 1997.
- PATWARDHAN et al. From data to diagnosis and control using generalized orthonormal basis filters. Part II: Model predictive and fault tolerant control. *Journal of Process Control*. v. 16, p. 157-175, 2006.
- PERRY, R. H et al. Perry's Chemical Engineers Handbook. McGraw-Hill Publishing Company, 7^a edição, 1997.
- PIOVOSO, M. J.; KOSANOVICH, K.A. Applications of multivariate statistical methods to process monitoring and controller design. *International Journal of Control*, v. 59, p. 743-765, 1994.
- PIOVOSO, M. J.; KOSANOVICH, K.A. Monitoring process performance in real time, PROCEEDINGS OF THE AMERICAN CONTROL CONFERENCE, Piscataway, NJ, IEEE, p. 2359-2363, 1992.
- PUIG, V.; QUEVEDO, J. Fault-tolerant PID controllers using a passive robust fault diagnosis approach. *Control Engineering Practice*. v. 9, p. 1221-1234, 2001.
- QIN, S.J; DUNIA, R. Determining the number of principal components for best reconstruction. *Journal of Process Control*, v. 10, p. 245-250, 2000.
- RAICH, A.; ÇINAR, A. Statistical process monitoring and disturbance diagnosis in multivariable continuous processes. *American Institute of Chemical Engineering Journal*, v. 42, p. 995-1009, 1996.
- RAICH, A. C.; ÇINAR, A. Multivariate statistical methods for monitoring continuous processes: assessment of discrimination power of disturbance models and diagnosis of multiple disturbances. *Chemometrics and Intelligent Laboratory Systems*, v. 30, p. 37-48, 1995.
- REIS, L. L. G.; OLIVEIRA-LOPES, L. C. Process Monitoring Based on Independent Component Analysis. SICOP 2006 - WORKSHOP ON SOLVING INDUSTRIAL CONTROL AND OPTIMIZATION PROBLEMS, Gramado, Brazil, 2006.
- REIS, L. L. G.; OLIVEIRA-LOPES, L. C. Controle PID Tolerante a Falhas de um CSTR. V SNCA - SEMINÁRIO NACIONAL DE CONTROLE E AUTOMAÇÃO INDUSTRIAL, ELÉTRICA E DE TELECOMUNICAÇÕES, Salvador, Brasil, 2007.
- SHRIDHAR, R.; COOPER, D. J. A Tuning Strategy for Unconstrained Multivariable Model Predictive Controllers. *Industrial And Engineering Chemistry Research*, v. 37, p. 4003, 1998.

-
- SOUZA, D. L. Análise de desempenho de sistemas de controle. Dissertação (Mestrado), Universidade Federal de Uberlândia, 2007.
- STAROSWIECKI, M.; DECLERCK, P. Analytical redundancy in non-linear interconnected systems by means of structural analysis. IFCA/IMACS/IFORS CONFERENCE AIPAC' 89, Nancy, France, 1989.
- TIAN, Q. et. al. Comparison of statistical pattern-recognition algorithms for hybrid processing, II: eigenvector-based algorithm. *J. Opt. Soc. Am.* v. 5, n. 10, p. 1670-1682, 1988.
- THORNHILL, N. F.; HUANG, B.; ZHANG, H. Detection of multiple oscillations in control loops. *Journal of Process Control*, v. 13, p. 91-100, 2003.
- VALLE-CERVANTES, S.; LI, W.; QIN, S.J. Selection of the number of principal-components: A new criterion with comparison to existing methods. *Industrial and Chemical Research*, v. 38, p. 4389-4401, 1999.
- WISE, B.M.; GALLAGHER, N.B. The process chemometrics approach to process monitoring and fault-detection. *Journal of Process Control*, v. 6, p. 329-348, 1996.
- WISE, B. M.; RICKER, N. L.; VELTKAMP, D. F.; KOWALSKI, B. R. A theoretical basis for the use of principal component models for monitoring multivariate processes. *Process Control and Quality*, v. 1, p. 41-51, 1990.
- WISE, B.M.; RICKER, N.L.; VELTKAMP, D.J. Upset and Sensor Failure Detection in Multivariate Processes, AIChE ANNUAL MEETING, San Francisco, CA, November, 1989.
- WOOD, R. K.; BERRY, M. W. Terminal composition control of a binary distillation column. *Chemical Engineering Science*, v. 28, p. 1707-1717, 1973.
- ZHANG, Y.; JIANG, J. Bibliographical Review on Reconfigurable Fault-tolerant Control Systems. PROCEEDINGS OF THE 5th IFAC SYMPOSIUM ON FAULT DETECTION, SUPERVISION AND SAFETY FOR TECHNICAL PROCESSES, p. 265-276, Washington, D.C., USA, June 9-11, 2003.