

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

CRISTIANE SILVA GARCIA

**CONTRIBUIÇÕES PARA O MÉTODO
*VIRTUAL REFERENCE FEEDBACK
TUNING***

Porto Alegre
2021

CRISTIANE SILVA GARCIA

**CONTRIBUIÇÕES PARA O MÉTODO
*VIRTUAL REFERENCE FEEDBACK
TUNING***

Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Rio Grande do Sul como parte dos requisitos para a obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração: Controle e Automação

ORIENTADOR: Prof. Dr. Alexandre Sanfelici
Bazanella

Porto Alegre
2021

CRISTIANE SILVA GARCIA

**CONTRIBUIÇÕES PARA O MÉTODO
*VIRTUAL REFERENCE FEEDBACK
TUNING***

Esta tese foi julgada adequada para a obtenção do título de Doutor em Engenharia Elétrica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: _____
Prof. Dr. Alexandre Sanfelici Bazanella, UFRGS
Doutor pela Universidade Federal de Santa Catarina – Florianópolis, Brasil

Banca Examinadora:

Prof. Dr. Julio Elias Normey Rico, PGEAS — UFSC
Doutor pela Universidad de Sevilla – Sevilla, Espanha

Prof. Dr. Celso José Munaro, Deptº Engª Elétrica — UFES
Doutor pela Universidade Estadual de Campinas – Campinas, Brasil

Profª. Drª. Lucíola Campestrini, PPGEE — UFRGS
Doutora pela Universidade Federal do Rio Grande do Sul – Porto Alegre, Brasil

Coordenador do PPGEE: _____
Prof. Dr. Sérgio Luís Haffner.

Porto Alegre, maio de 2021.

DEDICATÓRIA

Dedico este trabalho ao meu companheiro Róger pelo apoio e carinho durante todos esses anos.

AGRADECIMENTOS

Agradeço primeiramente ao SENHOR, pois sem Ele jamais teria chegado tão longe.

Agradeço ao Róger pelo apoio e companheirismo durante essa jornada.

Também ao meu orientador, professor Alexandre Sanfelici Bazanella, obrigada por ser meu mentor nesses quatro anos.

Por fim, agradeço aos meus pais, José Delmar e Ilsa, obrigada por me incentivar a buscar o conhecimento.

RESUMO

Em métodos de controle baseados em dados o ajuste dos parâmetros do controlador é realizado diretamente a partir dos dados coletados, sem a necessidade de estimar um modelo para o processo. Dentre os métodos propostos na literatura, o *Virtual Reference Feedback Tuning* (VRFT) talvez seja o método não iterativo mais empregado para realizar o ajuste dos parâmetros de um controlador de modo a atingir um desempenho em malha fechada predeterminado. Contudo, a qualidade da estimativa dos parâmetros com este método é afetada quando: há presença de ruído nos sinais coletados, empregam-se controladores de ordem reduzida, ou os dados são coletados de um experimento pouco informativo. Assim, o presente trabalho propõe soluções para esses três problemas. Para o caso ruidoso, o problema do VRFT é formulado com as soluções de mínimos quadrados *Data Least Squares* (DLS) e *Constrained Total Least Squares* (CTLS). Na solução CTLS considerou-se não apenas controladores com parametrização linear, mas também controladores com polos e zeros livres, permitindo mais flexibilidade na escolha da estrutura do controlador. Para o caso de controladores de ordem reduzida, são propostos filtros para a solução CTLS, permitindo sua aplicação nesse caso. No caso em que os dados são coletados de um experimento pouco informativo, como dados de operação, por exemplo, a utilização de todo o conjunto de dados não necessariamente melhora a estimativa dos parâmetros, nem o desempenho em malha fechada. Portanto, no presente trabalho foram adaptados para o método VRFT dois critérios de seleção de subconjuntos muito informativos de dados, presentes na literatura de identificação de sistemas. No contexto do presente trabalho, subconjuntos muito informativos de dados são segmentos do conjunto de dados original que contêm informação relevante para identificação dos parâmetros do controlador. Todas as contribuições propostas são ilustradas através de simulações.

Palavras-chave: Controle baseado em dados, soluções de mínimos quadrados, problemas de erros nas variáveis, VRFT, critérios de seleção de intervalos muito informativos.

ABSTRACT

In the data-driven (DD) control framework, the task of tuning the controller's parameters is carried out directly from the collected data, without the knowledge of the process model. Among the DD methods proposed in the literature, the Virtual Reference Feedback Tuning (VRFT) is perhaps the non-iterative method most commonly employed to tune a controller's parameters aiming to achieve a prescribed closed-loop performance. However, the parameters estimate's quality with this method is compromised: when the collected data is affected by noise, in the reduced-order controllers case, and when the data is little informative. Therefore, the present work proposes solutions to these three problems. For the noisy case, the VRFT problem is formulated using the least squares solutions known as Data Least Squares (DLS) and Constrained Total Least Squares (CTLS). The CTLS solution is formulated not only using linearly parametrized controllers but also controllers with free poles and zeros, allowing more flexibility to the choice of the controller's structure. For the reduced-order controllers case, filters are proposed for the CTLS solution, allowing its application in this case. For the case of little informative data, such as routine operating data, for example, more data does not necessarily mean a better estimate or a better closed-loop performance. Therefore, in the present work, two data selection criteria to select strongly informative subsets, applied in the system identification framework, were adapted to the VRFT problem. In the present work, strongly informative subsets are defined as sections of the original dataset containing the information relevant to identify the controller's parameters. All the proposed contributions are illustrated through simulations.

Keywords: data-driven control, least squares solutions, errors-in-variables problems, VRFT, data selection criteria.

LISTA DE ILUSTRAÇÕES

Figura 1	— Experimento para coleta de dados em malha fechada.	23
Figura 2	— Experimento para coleta de dados em malha aberta.	23
Figura 3	— Experimento virtual do método VRFT.	33
Figura 4	— Sinal de entrada e uma realização do sinal de saída coletado no exemplo da Seção 2.4.1.3.	42
Figura 5	— Espectro de densidade de energia do sinal de saída para uma entrada do tipo PRBS.	48
Figura 6	— Espectro de densidade de energia do sinal de saída para um degrau unitário com 1152 amostras como sinal de entrada.	48
Figura 7	— E_c calculado com os parâmetros estimados.	49
Figura 8	— Espectro de densidade de energia do sinal de saída para um degrau com 288 amostras como sinal de entrada.	50
Figura 9	— Entrada $u(t)$ e uma realização da saída $y(t)$ coletada de um experimento pouco informativo no exemplo da Seção 2.5.2.	51
Figura 10	— Interpretação geométrica das soluções de mínimos quadrados.	53
Figura 11	— Sinais de entrada e saída para o projeto de um controlador com polos livres.	64
Figura 12	— Sinais de entrada e saída de um controlador genérico.	68
Figura 13	— Uma realização dos dados do experimento em malha fechada do exemplo da Seção 3.2.4.	75
Figura 14	— Primeira diferença do quadrado do menor valor singular de uma realização dos dados do exemplo da Seção 4.2.5.	93
Figura 15	— Evolução no tempo do número de condicionamento para uma realização dos dados do exemplo da Seção 4.3.3.1.	97
Figura 16	— Evolução do número do condicionamento $\kappa(P(\tau))$ e da função custo estimada $\hat{J}_y(\hat{\rho}(t))$ para um degrau do exemplo da Seção 4.3.3.2.	98
Figura 17	— Distribuição da função custo estimada $\hat{J}^{\text{MR}}(\hat{\rho})$ com os parâmetros estimados com OLS, CTLS e OCI da simulação da Seção 5.2.1.	102
Figura 18	— Sinal de entrada e uma realização do sinal de saída da simulação da Seção 5.2.2.	104
Figura 19	— Distribuição da função custo estimada $\hat{J}^{\text{MR}}(\hat{\rho})$ com os parâmetros estimados com as soluções OLS, VI, DLS e CTLS na simulação da Seção 5.2.2.	104
Figura 20	— Sinal de referência e uma realização dos sinais de entrada e saída da simulação da Seção 5.2.3.	106

Figura 21	— Histogramas das funções custo estimadas $\hat{J}^{\text{MR}}(\hat{\rho})$ calculadas com os controladores estimados com OLS e CTLS na simulação da Seção 5.2.3.	107
Figura 22	— Sinal de entrada e uma realização do sinal de saída da simulação da Seção 5.3.1.	108
Figura 23	— Curva de $\Delta\sigma^2(\Phi_L(t))$ para uma realização dos dados da simulação da Seção 5.3.1.	109
Figura 24	— Curva de $\kappa(P(t))$ para uma realização dos dados da simulação da Seção 5.3.1.	110
Figura 25	— Distribuição da função custo estimada $\hat{J}^{\text{MR}}(\hat{\rho})$ com os parâmetros estimados com as soluções OLS e DLS com todo o dado e com os critérios de seleção da simulação da Seção 5.3.1.	110
Figura 26	— Dados de entrada e saída da simulação da Seção 5.3.2.	112
Figura 27	— Curva de $\Delta\sigma^2(\Phi_L(t))$ para uma realização dos dados da simulação da Seção 5.3.2.	113
Figura 28	— Distribuição da função custo estimada $\hat{J}^{\text{MR}}(\hat{\rho})$ com os parâmetros estimados com as soluções OLS e DLS com todo o dado e com o critério SSV da simulação da Seção 5.3.2.	113
Figura 29	— Diagramas de Bode dos controladores estimados com a solução OLS, com os intervalos mais informativos, e completo (com todo o dado) da simulação da Seção 5.3.2.	114
Figura 30	— Diagramas de Bode dos controladores estimados com a solução DLS, com os intervalos mais informativos, e completo (com todo o dado) da simulação da Seção 5.3.2.	114
Figura 31	— Dados de entrada e saída da simulação da Seção 5.3.3.	116
Figura 32	— Curva de $\Delta\sigma^2(\Phi_L(t))$ para uma realização dos dados da simulação da Seção 5.3.3.	117
Figura 33	— Curva de $\kappa(P(t))$ para uma realização dos dados da simulação da Seção 5.3.3.	117
Figura 34	— Distribuição da função custo estimada $\hat{J}^{\text{MR}}(\hat{\rho})$ com os parâmetros estimados com a solução OLS para cada caso da simulação da Seção 5.3.3.	118
Figura 35	— Distribuição da função custo estimada $\hat{J}^{\text{MR}}(\hat{\rho})$ com os parâmetros estimados com a solução DLS para cada caso da simulação da Seção 5.3.3.	118

LISTA DE TABELAS

Tabela 1	— Média e desvio padrão dos parâmetros estimados com o VRFT e as soluções OLS e VI no exemplo da Seção 2.4.1.3.	42
Tabela 2	— Média e desvio padrão dos parâmetros estimados com o OCI no exemplo da Seção 2.4.2.2.	46
Tabela 3	— Média e desvio padrão dos parâmetros estimados com o VRFT e OLS com dados do experimento pouco informativo no exemplo da Seção 2.5.2.	50
Tabela 4	— Média e desvio padrão dos parâmetros estimados com o VRFT e as soluções CTLS e OLS do exemplo da Seção 3.2.4.	75
Tabela 5	— Média e desvio padrão dos parâmetros estimados com o VRFT e a solução DLS do exemplo da Seção 3.3.1.	78
Tabela 6	— Média e desvio padrão dos parâmetros estimados com o VRFT e as soluções OLS e DLS com todo o conjunto de dados do exemplo da Seção 4.2.5.	93
Tabela 7	— Média e desvio padrão dos parâmetros estimados com o VRFT e as soluções OLS e DLS com os subconjuntos muito informativos selecionados através do critério SSV para o exemplo da Seção 4.2.5.	94
Tabela 8	— Média e desvio padrão dos parâmetros estimados com o VRFT e as soluções OLS e DLS com os subconjuntos muito informativos obtidos com o critério CN para o exemplo da Seção 4.3.3.1.	97
Tabela 9	— Média e desvio padrão dos parâmetros estimados com OLS, CTLS e OCI para a simulação da Seção 5.2.1.	102
Tabela 10	— Parâmetros estimados com OLS e VI para a simulação da Seção 5.2.2.	105
Tabela 11	— Média e desvio padrão dos parâmetros estimados com DLS e CTLS para a simulação da Seção 5.2.2.	105
Tabela 12	— Média e desvio padrão dos parâmetros estimados com OLS e CTLS para a simulação da Seção 5.2.3.	107
Tabela 13	— Média e desvio padrão dos parâmetros estimados a solução OLS com todo o dado e com os subconjuntos mais informativos para a simulação da Seção 5.3.1.	111
Tabela 14	— Média e desvio padrão dos parâmetros estimados a solução DLS com todo o dado e com os subconjuntos mais informativos para a simulação da Seção 5.3.1.	111

Tabela 15 — Média e desvio padrão dos parâmetros estimados com a solução OLS com todo o dado e com os subconjuntos mais informativos para a simulação da Seção 5.3.2.	115
Tabela 16 — Média e desvio padrão dos parâmetros estimados com a solução DLS com todo o dado e com os subconjuntos mais informativos para a simulação da Seção 5.3.2.	115
Tabela 17 — Média e desvio padrão dos parâmetros estimados com as soluções OLS e DLS com os dados da simulação da Seção 5.3.3.	119

LISTA DE ABREVIATURAS

ARX	<i>AutoRegressive with eXogenous input</i>
BIBO	<i>Bounded-Input Bounded-Output</i>
BJ	Box-Jenkins
CbT	<i>Correlation-based Tuning</i>
CN	<i>Condition Number</i> (número de condicionamento)
CTLS	<i>Constrained Total Least Squares</i>
DCbT	<i>Disturbance Correlation-based Tuning</i>
DLS	<i>Data Least Squares</i>
FIR	<i>Finite Impulse Response</i>
IFT	<i>Iterative Feedback Tuning</i>
IIR	<i>Infinite Impulse Response</i>
LTI	<i>Linear Time-Invariant</i> (linear e invariante no tempo)
OCI	<i>Optimal Controller Identification</i>
OE	<i>Output Error</i>
OLS	<i>Ordinary Least Squares</i>
PID	Proporcional-Integral-Derivativo
PRBS	<i>Pseudo Random Binary Sequence</i>
SISO	<i>Single-Input Single-Output</i>
STLS	<i>Structured Total Least Squares</i>
SSV	<i>Smallest Singular Value</i> (menor valor singular)
TLS	<i>Total Least Squares</i>
VDFT	<i>Virtual Disturbance Feedback Tuning</i>
VI	Variável Instrumental
VRFT	<i>Virtual Reference Feedback Tuning</i>

LISTA DE SÍMBOLOS

$A(q, \rho)$	denominador da parte ajustável do controlador
$B(q, \rho)$	numerador da parte ajustável do controlador
$C(q, \rho)$	estrutura do controlador
$C_d(q)$	controlador ideal
$C_F(q)$	parte fixa do controlador
$C_I(q, \rho)$	parte identificável do controlador
$\bar{C}(q)$	vetor que representa a classe de controladores
\mathcal{C}	classe de controle
$\bar{e}(t)$	sinal de erro virtual
$E[x(t)]$	valor esperado
$\bar{E}[\cdot]$	$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E[x(t)]$
$G(q)$	modelo do processo
$H(q)$	modelo do ruído
$J^{\text{CTLS}}(\rho)$	função custo do problema do CTLS
$J^{\text{MR}}(\rho)$	função custo do critério de desempenho de seguimento de referência
$J^{\text{OCI}}(\theta)$	função custo do erro de predição do método OCI
$J^{\text{VR}}(\rho)$	função custo do erro de predição do método VRFT
$L(q)$	filtro para pré-filtragem dos dados
$P(t)$	matriz de informação
m	número de parâmetros do controlador
N	número de amostras
$r(t)$	sinal de referência
$\bar{r}(t)$	sinal de referência virtual
$S(q, \rho)$	função de sensibilidade
t	índice da amostra
$T(q, \rho)$	função de sensibilidade complementar

$T_d(q)$	modelo de referência
$u(t)$	sinal de entrada do processo
u	vetor de saída medido
u_0	vetor de saída sem ruído
$\hat{u}(t, \rho)$	preditor do sinal de entrada
$v(t)$	ruído branco
$w(t)$	ruído de medição
$y(t)$	sinal de saída medido
$y_d(t)$	sinal de saída desejado
δ_u	vetor com as contribuições de ruído do vetor de saída medido
Δ_Φ	matriz com as contribuições de ruído da matriz regressora calculada
$\Delta\sigma^2(\bullet)$	primeira diferença do quadrado do menor valor singular
$\varepsilon^*(t, \rho)$	erro de predição
η_{SSV}, η_{CN}	limiares dos critérios SSV e CN, respectivamente
θ	vetor com os parâmetros do controlador e do ruído
$\kappa(\bullet)$	número de condicionamento
$\xi(t)$	vetor de variáveis instrumentais
ρ	vetor de parâmetros do controlador
$\hat{\rho}$	vetor de parâmetros estimado
ρ_d	vetor de parâmetros ideal
ρ^*	vetor de parâmetros ótimo
$\sigma^2(\bullet)$	quadrado do menor valor singular
$\varphi(t)$	vetor regressor
$\Phi_x(e^{j\omega})$	espectro do sinal $x(t)$
Φ	matriz regressora calculada
Φ_0	matriz regressora sem ruído
$\bullet^{(i)}$	i -ésima coluna da matriz
\bullet_i	i -ésimo elemento do vetor

SUMÁRIO

1 INTRODUÇÃO	16
1.1 Motivação	16
1.2 Objetivos	18
1.2.1 Objetivo geral	18
1.2.2 Objetivos específicos	18
1.3 Lista de contribuições	19
1.4 Estrutura do trabalho	20
2 DEFINIÇÕES E MÉTODOS DE CONTROLE BASEADOS EM DADOS	21
2.1 Introdução	21
2.2 Definições do sistema	21
2.2.1 Problema de controle baseado em otimização	23
2.2.2 O controlador	25
2.3 Identificação de sistemas	27
2.3.1 Modelo autoregressivo com entrada exógena (ARX)	28
2.3.2 Modelo de resposta ao impulso finita (FIR)	29
2.3.3 Modelo de erro na saída (OE)	29
2.3.4 Modelo Box-Jenkins (BJ)	30
2.3.5 Nota sobre as estruturas de modelo	30
2.4 Métodos de controle baseados em dados	31
2.4.1 <i>Virtual Reference Feedback Tuning</i>	31
2.4.2 <i>Optimal Controller Identification</i>	42
2.5 Qualidade dos dados	46
2.5.1 Definições	46
2.5.2 Exemplo motivacional	50
2.6 O problema de mínimos quadrados	51
2.6.1 Formulação do problema	51
2.6.2 Soluções de mínimos quadrados	52
2.7 Nota	53
2.8 Considerações finais	54
3 AS SOLUÇÕES DE MÍNIMOS QUADRADOS CTLS E DLS APLICADAS AO MÉTODO VRFT	56
3.1 Introdução	56
3.2 Constrained Total Least Squares	57
3.2.1 Controlador com polos fixos	61

3.2.2	Controlador com polos livres	63
3.2.3	O caso em que o controlador ideal não pertence à classe de controle	67
3.2.4	Exemplo ilustrativo	74
3.3	Data Least Squares	75
3.3.1	Exemplo ilustrativo	78
3.4	Nota	78
3.5	Considerações finais	79
4	CRITÉRIOS DE SELEÇÃO DE SUBCONJUNTOS INFORMATIVOS DE DADOS	80
4.1	Introdução	80
4.2	Critério baseado no menor valor singular	82
4.2.1	Estimativa dos parâmetros	82
4.2.2	Análise estatística da matriz regressora	84
4.2.3	Análise estatística do vetor de parâmetros estimado	88
4.2.4	O critério de seleção dos dados	91
4.2.5	Exemplo ilustrativo	92
4.3	Critério baseado no número de condicionamento	93
4.3.1	O critério de seleção dos dados	94
4.3.2	O caso em que o controlador ideal não pertence à classe de controle	96
4.3.3	Exemplos ilustrativos	96
4.4	Nota	99
4.5	Considerações finais	99
5	RESULTADOS	100
5.1	Introdução	100
5.2	Soluções de mínimos quadrados	100
5.2.1	Caso de ordem cheia, dados de malha fechada e controlador com polos livres	100
5.2.2	Caso de ordem reduzida, dados de malha aberta e controlador com polos fixos	103
5.2.3	Caso de ordem cheia, dados de malha fechada e sequência de degraus como entrada	105
5.3	Crítérios de seleção de subconjuntos muito informativos de dados	107
5.3.1	Caso de ordem cheia, dados de malha aberta e entrada do tipo onda quadrada	107
5.3.2	Caso de ordem reduzida, dados de malha fechada e entrada com perturbação do tipo pulso	111
5.3.3	Caso de ordem cheia, dados de malha aberta e erro de medida no sinal de saída	115
5.4	Considerações finais	119
6	CONCLUSÕES E QUESTÕES EM ABERTO	121
6.1	Conclusões	121
6.2	Questões em aberto	122
	REFERÊNCIAS	124

1 INTRODUÇÃO

1.1 Motivação

O problema de ajuste dos parâmetros de um controlador de realimentação é usualmente solucionado através de métodos baseados em modelo, métodos de controle baseados em dados, ou uma mistura de ambos. No projeto do controlador por meio de métodos baseados em modelo, o modelo do processo a ser controlado precisa ser conhecido e essa informação é utilizada diretamente para calcular os parâmetros do controlador. Em contrapartida, em métodos de controle baseados em dados, o ajuste dos parâmetros do controlador é realizado *diretamente* a partir dos dados coletados do processo, sem a necessidade de um conhecimento profundo sobre o mesmo (BAZANELLA; CAMPESTRINI; ECKHARD, 2011).

Ao longo das últimas três décadas, diferentes métodos de controle baseados em dados foram propostos na literatura. Em sua maioria, os métodos baseados em dados são fundamentados no paradigma do modelo de referência, no qual escolhe-se *a priori* o comportamento desejado para o processo em malha fechada, ou seja, o *modelo de referência*, e a estrutura do controlador a ser projetado. Com base nessas escolhas, nos dados de entrada e saída coletados do processo, e em uma função objetivo, o propósito desse paradigma consiste em ajustar um controlador de modo que a resposta em malha fechada, obtida com o controlador projetado, seja tão próxima quanto possível da resposta desejada. Dentre os métodos baseados em dados propostos na literatura, existem os métodos iterativos e os não iterativos. Os métodos iterativos (HJALMARSSON et al., 1998; HJALMARSSON, 2002; KARIMI; MIŠKOVIĆ; BONVIN, 2004; KARIMI; MIŠKOVIĆ; BONVIN, 2016) são aqueles nos quais os parâmetros do controlador ótimo são obtidos através de uma sequência de experimentos no processo real, ou seja, o projeto do controlador é realizado iterativamente, com os dados coletados em malha fechada com o controlador obtido na iteração anterior, e esse procedimento é repetido até que a função objetivo atinja um valor mínimo. Em contrapartida, os métodos de controle baseados em dados não iterativos ou *one-shot* (CAMPI; LECCHINI; SAVARESI, 2002; KARIMI; VAN HEUSDEN; BONVIN, 2007; CAMPESTRINI et al., 2017) são aqueles em que os parâmetros do controlador são ajustados com os dados coletados de um único experimento no processo.

Dentre os métodos de controle baseados em dados, o *Virtual Reference Feedback Tuning* (VRFT) é um método que recebeu atenção considerável na literatura e se mostrou bastante eficaz em aplicações práticas. O método VRFT utiliza os dados coletados no processo para gerar os dados de entrada e saída *virtuais* do controlador ideal e então identifica o melhor controlador considerando a estrutura de controle disponível. Dessa forma, o problema original de modelar a resposta em malha fechada é reformulado como um problema de identificação dos parâmetros do controlador.

Apesar do VRFT não ser restrito ao caso de controladores linearmente parametrizados, essa suposição é favorecida em sua formulação usual e na maioria das aplicações na literatura. De fato, a estrutura de controle comumente escolhida possui um denominador fixo e um numerador parametrizado. No caso de controladores linearmente parametrizados, a estimativa dos parâmetros é facilmente encontrada através da solução de um problema de mínimos quadrados padrão, conhecido na literatura como *Ordinary Least Squares* (OLS) (GOLUB; LOAN, 1980). Contudo, essa solução produz uma estimativa polarizada para os parâmetros quando os dados são afetados por ruído. A polarização da estimativa é inerente à formulação do problema e é esperada até mesmo no caso de controladores de ordem cheia, ou seja, quando o controlador ideal pode ser representado com a estrutura de controle disponível, pois sua origem reside no fato de que o VRFT não é um problema de identificação padrão (SÖDERSTRÖM, 2018). Usualmente, o efeito do ruído é contornado pela utilização do método das variáveis instrumentais (VI) (CAMPI; LECCHINI; SAVARESI, 2002). Porém, essa abordagem possui o inconveniente de aumentar a variância da estimativa, podendo acarretar em uma significativa perda de desempenho e até mesmo em instabilidade em malha fechada (SÖDERSTRÖM, 2006; RALLO et al., 2016). Com isso em mente, outras abordagens foram propostas na literatura para lidar com dados ruidosos no método VRFT. Por exemplo, Van Heusden, Karimi e Söderström (2011) propuseram a utilização de um método de variáveis instrumentais estendidas, desenvolvido para dados periódicos. Rallo et al. (2016) propuseram a adição de um termo de regularização à função custo minimizada pelo método.

Além disso, na maioria dos métodos baseados em dados a qualidade da estimativa depende da informatividade dos dados coletados. Por conta disso, é uma prática comum coletar dados de um experimento específico onde um sinal suficientemente rico é aplicado à entrada do processo. Contudo, a execução de tal experimento pode ser uma tarefa custosa e indesejável em alguns casos (BITTENCOURT et al., 2015; SHARDT; HUANG, 2013). Para evitar essa tarefa, uma alternativa é estimar os parâmetros do controlador utilizando dados coletados durante a operação normal do processo. Em processos industriais, esses dados são comumente armazenados em um banco de dados, uma vez que os mesmos são normalmente empregados em monitoramento e detecção de faltas (MACGREGOR; CINAR, 2012). Porém, a simples aplicação desses dados em um método de controle baseado em dados pode levar a um ajuste inapropriado dos parâmetros, por causa da sua baixa informatividade e da presença de ruído. Na literatura de identificação de sistemas, existem alguns trabalhos que apresentam formas de identificar subconjuntos muito informativos desse tipo de dado, antes de sua aplicação na identificação do processo (CARRETTE et al., 1996; BITTENCOURT et al., 2015; SHARDT; HUANG, 2013; ARENGAS; KROLL, 2018; WANG et al., 2018). Por subconjuntos muito informativos de dados entende-se segmentos do conjunto de dados original, que são identificados como úteis através de algum critério de seleção. Contudo, esse assunto foi pouquíssimo pesquisado na área de controle baseado em dados. Por exemplo, Radrizzani et al. (2020) empregaram dados de operação no ajuste dos parâmetros do controlador com o VRFT e os intervalos muito informativos foram identificados manualmente pelo projetista.

Com tudo isso em mente, observou-se que ainda existe espaço para melhorias no método VRFT, para permitir seu emprego em situações comumente encontradas na prática. As situações abordadas no presente trabalho compreendem o caso ruidoso e o caso em que os dados disponíveis são pouco informativos. No caso em que os dados são afetados por ruído, a alternativa considerada à abordagem com VI consiste na utilização de outras soluções de mínimos quadrados na formulação do método. Para essas soluções, considera-

se controladores com o denominador parametrizado, o que adiciona flexibilidade à estrutura de controle. Além disso, para uma dessas soluções em específico, considera-se o caso de controladores de ordem reduzida, ou seja, com a estrutura do controlador disponível não é possível representar o controlador ideal. Por fim, para o caso de dados pouco informativos, a abordagem considerada consiste na adaptação para o VRFT de critérios de seleção de subconjuntos muito informativos de dados, inicialmente desenvolvidos para identificação de sistemas.

1.2 Objetivos

1.2.1 Objetivo geral

O principal objetivo dessa tese é propor melhorias para o método VRFT no caso de dados com ruído e pouco informativos. Primeiramente, buscou-se estudar o impacto da utilização de dados que não respeitam as condições de informatividade necessárias para garantir a qualidade da estimativa. Nesse tópico, estudaram-se alguns dos trabalhos propostos na literatura de identificação de sistemas (CARRETTE et al., 1996; PERETZKI et al., 2011; SHARDT; HUANG, 2013; SHARDT; SHAH, 2014; BITTENCOURT et al., 2015; HOLCOMB; BITMEAD, 2017; SHARDT; HUANG, 2017; ARENGAS; KROLL, 2017a; ARENGAS; KROLL, 2017b; ARENGAS; KROLL, 2018; WANG et al., 2018; ARENGAS; KROLL, 2019). Dessa pesquisa, foram escolhidos dois critérios de seleção de subconjuntos muito informativos de dados para serem aplicados ao método VRFT.

O segundo objetivo dessa tese é o estudo dos problemas de mínimos quadrados para encontrar uma solução alternativa à abordagem com VI. Observou-se que outras soluções poderiam ser aplicadas ao método, além das soluções originalmente propostas para o mesmo. A partir daí, procurou-se estudar quais dessas soluções poderiam representar melhor a estrutura do problema proposto pelo VRFT.

1.2.2 Objetivos específicos

Dos objetivos gerais apresentados anteriormente surgiram os seguintes objetivos específicos:

- **investigar os critérios de seleção existentes na literatura.** Na área de identificação de sistemas já existem trabalhos que propuseram critérios para selecionar os intervalos que possuem informação relevante a partir de um conjunto maior de dados classificados como pouco informativos. Esses trabalhos mostram que a qualidade da estimativa com os dados selecionados é superior à obtida com todo o conjunto de dados. Assim sendo, foi feita uma investigação sobre esses critérios, resultando na escolha de dois deles e permitindo que os mesmos fossem adaptados para o método VRFT;
- **aplicar os critérios de seleção ao VRFT.** O objetivo dessa etapa foi adaptar os critérios de seleção para serem utilizados com os sinais virtuais do método VRFT. Nesse caso, foi considerado que o controlador é de ordem cheia, ou seja, com a estrutura de controle disponível é possível alcançar exatamente o comportamento desejado em malha fechada;
- **analisar a aplicação dos critérios de seleção em casos mais realistas.** É muito comum que com a estrutura de controle escolhida/disponível não seja possível alcançar o comportamento desejado em malha fechada, ou seja, o controlador possui

ordem reduzida. Por conta disso, uma análise do comportamento dos critérios nesse caso específico também precisou ser realizada. Além disso, esses critérios de seleção são normalmente empregados considerando-se algumas condições ideais quanto aos dados (descritas detalhadamente no capítulo dedicado aos critérios). Como tais condições nem sempre podem ser atendidas em casos mais realistas, outro objetivo do presente trabalho foi analisar o comportamento desses critérios nos casos em que as condições ideais não são respeitadas;

- **investigar os métodos de mínimos quadrados.** Como já mencionado, quando dados ruidosos são empregados no método VRFT, a abordagem com VI reduz a polarização da estimativa com o inconveniente de aumentar sua variância. Por conta disso, foi realizada uma investigação na literatura de problemas de mínimos quadrados, permitindo que duas soluções diferentes fossem empregadas no método VRFT;
- **aplicar os métodos de mínimos quadrados.** Nessa etapa foram desenvolvidas as alterações necessárias para possibilitar a aplicação das duas soluções de mínimos quadrados ao método VRFT. Nesse caso, considerou-se que o controlador era de ordem cheia com denominador fixo;
- **estender a aplicação para controladores com denominador parametrizado.** Apesar de um controlador com denominador fixo, um PID com polos fixos por exemplo, ser uma escolha bastante comum na prática de sistemas de controle, existem casos em que deseja-se estimar um controlador com denominador parametrizado (polos livres). Por conta disso, outro objetivo dessa tese foi o desenvolvimento de uma das soluções de mínimos quadrados também para o caso de controladores com polos livres;
- **estender a aplicação para o caso de controladores de ordem reduzida.** Quando o comportamento desejado em malha fechada é um pouco distante do que pode ser alcançado com a estrutura de controle disponível, é possível melhorar o desempenho em malha fechada através da pré-filtragem dos sinais empregados na estimativa. Essa abordagem já foi desenvolvida para o método VRFT considerando-se a solução OLS. No presente trabalho, estudou-se o desenvolvimento dos filtros quando outra solução de mínimos quadrados é aplicada ao método VRFT;
- **combinar os critérios de seleção com outra solução de mínimos quadrados.** O último objetivo dessa tese foi estudar a possibilidade de aplicar os intervalos muito informativos identificados através dos critérios de seleção com outra solução de mínimos quadrados.

1.3 Lista de contribuições

A pesquisa realizada para atingir os objetivos descritos na seção anterior resultou nas seguintes publicações:

- a adaptação dos critérios de seleção para o método VRFT considerando-se o caso em que o controlador possui ordem cheia e que os parâmetros do mesmo são estimados através da solução original foi publicado em (SILVA GARCIA; BAZANELLA, 2019);

- a aplicação dos critérios de seleção para o método VRFT no caso em que o controlador possui ordem reduzida e a estimativa dos parâmetros é obtida com a solução original foi publicada em (SILVA GARCIA; BAZANELLA, 2020b). Naquele trabalho mostrou-se que apenas um dos critérios de seleção pode ser aplicado com sucesso no caso de controladores de ordem reduzida;
- a aplicação de outra solução de mínimos quadrados ao método VRFT foi publicada em (SILVA GARCIA; BAZANELLA, 2020a). Naquele trabalho, considerou-se um controlador linearmente parametrizado com polos fixos e o caso em que o controlador possui ordem cheia. Além disso, a aplicação da solução foi desenvolvida tanto para o caso de dados de malha aberta quanto para dados de malha fechada.

1.4 Estrutura do trabalho

O restante desse trabalho está estruturado como segue. No Capítulo 2 são realizadas algumas definições importantes que serão empregadas ao longo do trabalho. Esse capítulo também apresenta alguns conceitos de identificação de sistemas. Além disso, é realizada uma breve revisão bibliográfica sobre os métodos de controle baseados em dados e são descritos em detalhes dois dos métodos existentes. Finalmente, são realizadas algumas definições referentes à qualidade dos dados e ao final do capítulo é apresentada uma revisão sobre os métodos de mínimos quadrados. O Capítulo 3 apresenta duas soluções de mínimos quadrados aplicadas ao método VRFT. Além disso, nesse capítulo é realizado o desenvolvimento dos filtros para o caso em que o controlador possui ordem reduzida. Mais adiante, no Capítulo 4 são apresentados dois critérios de seleção de subconjuntos muito informativos de dados adaptados para o método VRFT. O Capítulo 5 apresenta alguns estudos de caso simulados mostrando a viabilidade das soluções propostas. Finalmente, no Capítulo 6 são apresentadas as conclusões e são discutidas as questões em aberto.

2 DEFINIÇÕES E MÉTODOS DE CONTROLE BASEADOS EM DADOS

2.1 Introdução

O objetivo deste capítulo consiste em apresentar os conceitos que serão empregados no presente trabalho. Primeiramente, são apresentadas algumas definições básicas. Em seguida, são descritos dois métodos de controle baseados em dados existentes na literatura, sendo eles: o método VRFT, para o qual foram desenvolvidas as melhorias apresentadas no presente trabalho e o método *Optimal Controller Identification* (OCI), utilizado para comparação de parte dos resultados. Além disso, são apresentadas algumas considerações quanto aos dados empregados na estimativa dos parâmetros do controlador.

Em seguida, são apresentadas algumas definições referentes à qualidade dos dados. Finalmente, o problema de mínimos quadrados será abordado. Uma breve descrição do problema será apresentada, juntamente com as soluções propostas na literatura.

2.2 Definições do sistema

Considere o sistema em tempo discreto, *Single-Input Single-Output* (SISO), linear e invariante no tempo (LTI, do inglês *Linear Time-Invariant*), dado por

$$y(t) = G(q)u(t) + w(t), \quad (1)$$

onde $G(q)$ é a função de transferência do processo, $y(t)$ é o sinal de saída, $u(t)$ é o sinal de entrada, e $w(t)$ é um ruído aditivo que afeta a saída do processo. O sinal $w(t)$ representa todo efeito estocástico que não é capturado pela relação de entrada-saída entre $u(t)$ e $y(t)$, podendo ser uma perturbação estocástica, erro de medida, etc (BAZANELLA; CAMPESTRINI; ECKHARD, 2011). Além disso, q é o operador de avanço definido como

$$qx(t) \triangleq x(t + 1),$$

para $x(t)$, um sinal qualquer.

Seja o operador $\bar{E}[\cdot]$ definido como

$$\bar{E}[f(t)] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E\{f(t)\},$$

onde $E\{\cdot\}$ denota o valor esperado (LJUNG, 1999). Um sinal é dito quasi-estacionário se o mesmo atende à definição abaixo.

Definição 2.1 (Sinal quasi-estacionário (LJUNG, 1999)). Um sinal $x(t)$ é dito um processo quasi-estacionário se

$$\begin{aligned} E\{x(t)\} &= m_x(t), & |m_x(t)| &\leq C, \quad \forall t \\ E\{x(t)x(n)\} &= R_x(t, n), & |R_x(t, n)| &\leq C, \quad \forall t \end{aligned}$$

e se o seguinte limite existe:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N R_x(t, t - \tau) = R_x(\tau) \quad \forall \tau,$$

onde $m_x(t)$ é o valor esperado do sinal $x(t)$ e $R_x(t, n)$ é a função de autocorrelação do sinal $x(t)$ nos instantes t e n .

O ruído $w(t)$ é um processo quasi-estacionário, que pode ser descrito por

$$w(t) = H(q)v(t), \quad (2)$$

onde $H(q)$ é o modelo do ruído e $v(t)$ é um ruído branco com média zero e variância σ_v^2 . Considera-se que $G(q)$ e $H(q)$ são funções de transferência racionais e causais. Além disso, assume-se que $H(q)$ é *Bounded-Input Bounded-Output* (BIBO) estável (BAZANELLA; CAMPESTRINI; ECKHARD, 2011).

O processo (1) é controlado por um controlador linear e invariante no tempo, cuja estrutura é fixa. A ação de controle é dada por

$$u(t) = C(q, \rho)(r(t) - y(t)), \quad (3)$$

onde $C(q, \rho)$ é a função de transferência do controlador, parametrizada em função de um vetor de parâmetros real $\rho \in \mathbb{R}^m$, onde m é o número de parâmetros a serem estimados. Além disso, $r(t)$ é o sinal de referência. Assume-se que $r(t)$ é um sinal determinístico e descorrelacionado com o ruído, ou seja,

$$\bar{E}[r(t)v(t - \tau)] = 0 \quad \forall \tau.$$

O caso em que os dados são coletados de um experimento em malha fechada é descrito através da Definição 2.2. Em contrapartida, o caso em que os dados são provenientes de um experimento em malha aberta é descrito pela Definição 2.3.

Definição 2.2 (Dados de malha fechada). O processo (1) – (3) em malha fechada é apresentado na Figura 1 e descrito por

$$y(t) = T(q, \rho)r(t) + S(q, \rho)w(t), \quad (4)$$

$$u(t) = S(q, \rho)C(q, \rho)r(t) - S(q, \rho)C(q, \rho)w(t), \quad (5)$$

$$S(q, \rho) = (1 + C(q, \rho)G(q))^{-1}, \quad (6)$$

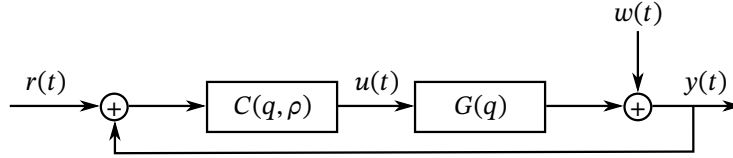
$$T(q, \rho) = S(q, \rho)G(q)C(q, \rho), \quad (7)$$

onde (4) é obtida substituindo-se (3) em (1) e (5) é calculada substituindo-se (1) em (3). Além disso, $S(q, \rho)$ e $T(q, \rho)$ são conhecidas como a *função de sensibilidade* e a *função de sensibilidade complementar* obtidas com o controlador $C(q, \rho)$, respectivamente. Os dados de malha fechada coletados formam o conjunto \mathcal{Z}^N dado por

$$\mathcal{Z}^N = \{u(1) \ y(1) \ \dots \ u(N) \ y(N)\}, \quad (8)$$

onde N é o número de amostras coletadas no experimento.

Figura 1 — Experimento para coleta de dados em malha fechada.

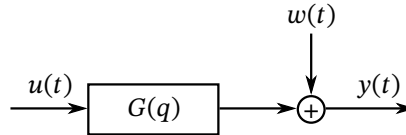


Fonte: elaborado pela autora.

Definição 2.3 (Dados de malha aberta). Quando os dados são coletados durante a execução de um experimento em malha aberta, a entrada $u(t)$ é um sinal exógeno e a saída é dada em (1), como apresentado na Figura 2. Neste caso, assume-se que $G(q)$ é BIBO estável. As N amostras de dados coletados em malha aberta formam o conjunto \mathcal{Z}^N definido como

$$\mathcal{Z}^N = \{u(1) \ y(1) \ \dots \ u(N) \ y(N)\}. \quad (9)$$

Figura 2 — Experimento para coleta de dados em malha aberta.



Fonte: elaborado pela autora.

2.2.1 Problema de controle baseado em otimização

Neste problema deseja-se encontrar um vetor de parâmetros ρ de modo que um determinado objetivo de controle seja minimizado, ou seja,

$$\min_{\rho} J(\rho),$$

onde ρ é a variável de decisão e $J(\rho)$ é uma função objetivo, ou função custo, que expressa o desempenho do processo em atingir um objetivo de controle predefinido. Alguns objetivos de controle usualmente empregados são: seguimento de referência, rejeição de perturbação e esforço de controle (BAZANELLA; CAMPESTRINI; ECKHARD, 2011).

Suponha que o objetivo de controle seja seguimento de referência. Dentre as diversas técnicas de projeto de controle existentes, no presente trabalho, o projeto de controle é realizado por modelo de referência, no qual o projetista define *a priori* o comportamento desejado para o sistema em malha fechada através da escolha de uma função de transferência $T_d(q)$, chamada de *modelo de referência*. Nesta função de transferência são especificadas as características desejadas para o sistema em malha fechada, como tempo de acomodação, sobrepasso, entre outras. A resposta em malha fechada desejada para o processo, dado um sinal de referência $r(t)$, é dada por

$$y_d(t) = T_d(q)r(t). \quad (10)$$

O projeto de controle por modelo de referência leva em consideração somente a resposta do processo em malha fechada referente ao sinal de referência, sem levar em consideração

o ruído no sinal de saída (BAZANELLA; CAMPESTRINI; ECKHARD, 2011). Portanto, a resposta obtida com o controlador $C(q, \rho)$ é definida como

$$y(t, \rho) = T(q, \rho)r(t), \quad (11)$$

onde $T(q, \rho)$ é definido em (7).

O objetivo é encontrar o controlador que fará com que a resposta $y(t, \rho)$ seja tão próxima quanto possível da resposta desejada $y_d(t)$. Para tanto, define-se a função custo do critério de desempenho de seguimento de referência como

$$J^{\text{MR}}(\rho) \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N |y(t, \rho) - y_d(t)|^2. \quad (12)$$

Substituindo-se (10) e (11) na equação anterior tem-se

$$J^{\text{MR}}(\rho) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N |\varepsilon_{\text{MR}}(t, \rho)|^2, \quad (13)$$

$$\varepsilon_{\text{MR}}(t, \rho) = [T(q, \rho) - T_d(q)]r(t), \quad (14)$$

onde $\varepsilon_{\text{MR}}(t, \rho)$ é o erro da resposta em malha fechada em relação à resposta do modelo de referência, dado um mesmo sinal de referência $r(t)$. O objetivo consiste em encontrar o vetor de parâmetros ótimo ρ_{MR}^* que minimiza a função custo $J^{\text{MR}}(\rho)$, ou seja,

$$\rho_{\text{MR}}^* = \underset{\rho}{\text{argmin}} J^{\text{MR}}(\rho).$$

A estrutura do controlador $C(q, \rho)$ delimita um subconjunto dentro do conjunto de controladores conhecido como classe de controle \mathcal{C} , definida como

$$\mathcal{C} = \{C(q, \rho) \mid \rho \in \Omega \subseteq \mathbb{R}^{m_1}\}, \quad (15)$$

onde Ω é o subconjunto de todos os vetores de parâmetros implementáveis. O *controlador ideal*, $C_d(q)$, é aquele que faz com que a função de sensibilidade complementar seja igual à desejada, ou seja,

$$T_d(q) = [1 + C_d(q)G(q)]^{-1} C_d(q)G(q).$$

A expressão acima pode ser rearranjada para obter o controlador ideal como:

$$C_d(q) = [G(q) - T_d(q)G(q)]^{-1} T_d(q). \quad (16)$$

Se o modelo do processo $G(q)$ fosse conhecido, se o modelo de referência fosse escolhido de modo apropriado e não houvesse nenhuma restrição quanto à estrutura do controlador, o controlador ideal poderia ser facilmente calculado de (16). Contudo, essas condições são raramente satisfeitas quando projeta-se controladores para processos reais. Na maioria dos casos, o modelo do processo $G(q)$ não é conhecido e precisa ser modelado matematicamente ou estimado através de alguma técnica de identificação, e a estimativa obtida apresentará variações devido ao ruído, normalmente presente nos dados coletados. Nesse caso, o controlador obtido em (16) não necessariamente fará com que o processo em malha fechada represente exatamente o modelo de referência previamente escolhido. Além disso, uma escolha inadequada de $T_d(q)$ pode fazer com que o controlador calculado em (16) possua características indesejadas, como zeros de fase não mínima, instabilidade e seja não

causal, por exemplo. Uma breve discussão sobre a escolha de $T_d(q)$ será apresentada mais adiante no texto. Finalmente, a estrutura do controlador escolhida/disponível raramente será flexível o suficiente para representar (16), pois comumente emprega-se um controlador com estrutura fixa onde apenas os ganhos do controlador são ajustados. Por conta disso, o projeto do controlador através de um método baseado em dados com modelo de referência não é realizado através de (16). Ao invés disso, o projeto do controlador é realizado através da minimização indireta da função custo $J^{\text{MR}}(\rho)$.

Quando o controlador ideal pode ser representado com a estrutura de controle escolhida, tem-se que $C_d(q) \in \mathcal{C}$. Isto implica que existe um vetor de parâmetros ideal ρ_d tal que $C(q, \rho_d) = C_d(q)$, como descrito na Suposição 2.1. Quando esta suposição é respeitada tem-se que $T(q, \rho_d) = T_d(q)$ e $J^{\text{MR}}(\rho_d) = 0$. Em contrapartida, quando a estrutura de controle não é capaz de representar o controlador ideal, ou seja, a Suposição 2.1 é violada, encontra-se o controlador ótimo ρ_{MR}^* e a função custo $J^{\text{MR}}(\rho)$ não é nula.

Suposição 2.1 (O controlador ideal pertence à classe de controle). Assumindo-se que o controlador ideal $C_d(q)$ pode ser representado com a classe de controladores escolhida tem-se que:

$$\exists \rho_d \in \Omega \mid C(q, \rho_d) = C_d(q),$$

onde $\rho_d \in \mathbb{R}^m$ é o vetor de parâmetros ideal. ◆

Observe que a Suposição 2.1 é respeitada dependendo da estrutura escolhida (ou disponível) para o controlador e da estrutura do controlador ideal. Como apresentado em (16), o controlador ideal depende do modelo de referência escolhido e da função de transferência do processo, que é desconhecida. Assim, pode-se garantir que a Suposição 2.1 seja respeitada ou moderadamente violada através da escolha de $T_d(q)$. Se uma função de transferência $T_d(q)$ muito exigente é escolhida, isso é refletido na estrutura do controlador ideal, que pode ser bem diferente da classe de controle escolhida/disponível. O que pode resultar em um comportamento em malha fechada muito diferente do desejado e até mesmo em instabilidade (BAZANELLA; CAMPESTRINI; ECKHARD, 2011). Portanto, o modelo de referência deve ser escolhido de modo apropriado. Diretrizes para escolha do modelo de referência para o caso monovariável e multivariável são apresentadas por Gonçalves da Silva, Campestrini e Bazanella (2014) e Gonçalves da Silva, Bazanella e Campestrini (2019). Contudo, a escolha do modelo de referência é um assunto que está fora do escopo do presente trabalho, e aqui apenas assume-se que $T_d(q)$ é escolhida de forma adequada de modo que a Suposição 2.1 seja moderadamente violada. Em outras palavras, escolhe-se $T_d(q)$ de modo que o comportamento desejado seja razoavelmente próximo do comportamento alcançável com a estrutura de controle escolhida/disponível.

2.2.2 O controlador

Como visto, a estrutura do controlador ideal $C_d(q)$ depende do modelo do processo $G(q)$ e do modelo de referência escolhido $T_d(q)$. Por conta disso, se a escolha da estrutura de $C(q, \rho)$ não é livre para o projetista, é bastante provável que a Suposição 2.1 seja violada. Por conta disso, o comportamento desejado pode não ser alcançado. Uma estrutura comumente empregada em métodos baseados em dados é a estrutura com parametrização linear, como descrito na Definição 2.4.

Definição 2.4 (Controlador linearmente parametrizado (BAZANELLA; CAMPESTRINI; ECKHARD, 2011)). O controlador é dito linearmente parametrizado se sua função de

transferência pode ser escrita como

$$C(q, \rho) = \rho^T \bar{C}(q),$$

onde $\bar{C}(q)$ é um vetor de m funções racionais independentes do vetor de parâmetros ρ .

Um exemplo de controlador linearmente parametrizado é um controlador Proporcional-Integral-Derivativo (PID) com polos fixos, definido como abaixo

$$C(q, \rho) = \underbrace{[K_p \quad K_i \quad K_d]}_{\rho^T} \underbrace{\left[1 \quad \frac{q}{q-1} \quad \frac{q-1}{q}\right]^T}_{\bar{C}(q)}. \quad (17)$$

Uma outra opção de estrutura para o mesmo controlador é dada por

$$C(q, \rho) = B(q, \rho)C_F(q),$$

onde

$$B(q, \rho) = b_1 + b_2 q^{-1} + \dots + b_{n_b} q^{-n_b+1},$$

$n_b = m$ é o número de parâmetros a serem estimados e $C_F(q)$ é a parte fixa do controlador. Como exemplo, considera-se um controlador PID com a seguinte estrutura

$$C(q, \rho) = [\rho_1 \quad \rho_2 \quad \rho_3] \left[\frac{q^2}{q(q-1)} \quad \frac{q}{q(q-1)} \quad \frac{1}{q(q-1)} \right]^T \quad (18)$$

$$= \underbrace{(\rho_1 + \rho_2 q^{-1} + \rho_3 q^{-2})}_{B(q, \rho)} \underbrace{\frac{q^2}{q(q-1)}}_{C_F(q)} \quad (19)$$

Nota 2.1. Observe que a estrutura (17) pode ser facilmente convertida para a estrutura (18) através da seguinte transformação linear

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} K_p \\ K_i \\ K_d \end{bmatrix}. \quad (20)$$

Por outro lado, quando não há restrição de que a estrutura do controlador possua polos fixos, pode-se escolher um controlador de modo que tanto seus polos quanto zeros sejam estimados. Tal estrutura pode ser representada por

$$C(q, \rho) = \frac{B(q, \rho)}{A(q, \rho)} C_F(q), \quad (21)$$

onde

$$B(q, \rho) = b_1 + b_2 q^{-1} + \dots + b_{n_b} q^{-n_b+1} \quad (22)$$

$$A(q, \rho) = 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a}, \quad (23)$$

com $m = n_a + n_b$ e $C_F(q)$ sendo a parte fixa do controlador, podendo conter qualquer informação previamente conhecida pelo projetista, como por exemplo, um polo em $q = 1$ para garantir erro nulo tanto em seguimento de referências quanto em rejeição de perturbações constantes.

2.3 Identificação de sistemas

Devido à existente relação entre a área de identificação de sistemas e métodos de controle baseado em dados, nessa seção, são apresentados alguns conceitos definidos para identificação de sistemas, que serão utilizados no decorrer do trabalho.

Considere o processo apresentado em (1). Este processo pode ser descrito pelo seguinte modelo paramétrico

$$y(t) = G(q, \theta)u(t) + H(q, \theta)v(t), \quad (24)$$

onde $G(q, \theta)$ e $H(q, \theta)$ são funções racionais em q , θ é o vetor de parâmetros a ser identificado, e $v(t)$ é ruído branco com média zero e variância σ_v^2 . O modelo generalizado de (24) é reescrito como (LJUNG, 1999)

$$A(q, \theta)y(t) = \frac{B(q, \theta)}{F(q, \theta)}u(t) + \frac{K(q, \theta)}{D(q, \theta)}v(t), \quad (25)$$

onde

$$A(q, \theta) = 1 + a_1q^{-1} + \dots + a_{n_a}q^{-n_a}, \quad (26)$$

$$B(q, \theta) = b_1q^{-1} + \dots + b_{n_b}q^{-n_b}, \quad (27)$$

$$K(q, \theta) = 1 + k_1q^{-1} + \dots + k_{n_k}q^{-n_k}, \quad (28)$$

$$F(q, \theta) = 1 + f_1q^{-1} + \dots + f_{n_f}q^{-n_f}, \quad (29)$$

$$D(q, \theta) = 1 + d_1q^{-1} + \dots + d_{n_d}q^{-n_d}. \quad (30)$$

O preditor para (25) é dado por

$$\hat{y}(t, \theta) = \frac{D(q, \theta)B(q, \theta)}{K(q, \theta)F(q, \theta)}u(t) + \left[1 - \frac{D(q, \theta)A(q, \theta)}{K(q, \theta)} \right] y(t). \quad (31)$$

A partir da escolha da estrutura do modelo para representar o sistema e do conjunto de dados coletados \mathcal{Z}^N , define-se o método utilizado para estimar o vetor de parâmetros θ com base nas escolhas realizadas. Talvez o método mais comumente empregado na literatura seja a minimização do erro de predição (LJUNG, 1999), $\varepsilon(t, \theta)$, definido como

$$\varepsilon(t, \theta) \triangleq y(t) - \hat{y}(t, \theta). \quad (32)$$

Seja, agora o erro de predição filtrado por um filtro linear e estável

$$\varepsilon_L(t, \theta) = L(q)[y(t) - \hat{y}(t, \theta)], \quad (33)$$

se o preditor for linear e invariante no tempo e $y(t)$ e $u(t)$ forem escalares, então, filtrar $\varepsilon(t, \theta)$ é o mesmo que filtrar os sinais de entrada e saída e depois calcular $\varepsilon_L(t, \theta)$ (LJUNG, 1999). O erro de predição pode ser visto como um vetor de dimensão \mathbb{R}^N , o “tamanho” deste vetor pode ser medido com qualquer norma em \mathbb{R}^N . Uma escolha padrão é uma norma quadrática como

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^N |\varepsilon_L(t, \theta)|^2. \quad (34)$$

A estimativa dos parâmetros é obtida através da minimização de (34):

$$\hat{\theta}_N = \arg \min_{\theta} V_N(\theta). \quad (35)$$

Veja que o erro de predição em (32) depende da estrutura do modelo escolhida. Segundo Ljung (1999), a maioria das aplicações práticas podem ser representadas com estruturas mais simples do que a apresentada em (25). As estruturas de modelo são classificadas dependendo de quais polinômios de (25) são fixados em 1. Na sequência, serão descritas apenas as estruturas de modelo, e seus respectivos preditores, pertinentes ao presente trabalho. Uma descrição completa dessas estruturas pode ser encontrada no livro de Ljung (1999).

2.3.1 Modelo autoregressivo com entrada exógena (ARX)

Para estruturas de modelo do tipo ARX (do inglês *AutoRegressive with eXogenous input*) considera-se que

$$F(q, \theta) = K(q, \theta) = D(q, \theta) = 1. \quad (36)$$

O modelo ARX é escrito substituindo-se (36) em (25), chegando-se a

$$A(q, \theta)y(t) = B(q, \theta)u(t) + v(t), \quad (37)$$

que pode ser reescrito através da equação de diferenças, substituindo-se (26) e (27) na equação acima, como

$$y(t) = b_1 u(t-1) + \dots + b_{n_b} u(t-n_b) - a_1 y(t-1) - \dots - a_{n_a} y(t-n_a) + v(t).$$

Este modelo é frequentemente chamado de *modelo de erro na equação*, pois o termo de ruído $v(t)$ afeta diretamente a equação de diferenças. Além disso, este modelo é batizado como ARX devido à parte autorregressiva $A(q, \theta)y(t)$ (*AutoRegressive* – AR) e à entrada externa $B(q, \theta)u(t)$ chamada de variável exógena em econometria (*eXogenous* – X) (LJUNG, 1999).

O preditor do modelo ARX é reescrito substituindo-se (36) em (31) como

$$\hat{y}(t, \theta) = B(q, \theta)u(t) + [1 - A(q, \theta)]y(t). \quad (38)$$

Substituindo-se (26) e (27) na equação acima obtém-se:

$$\begin{aligned} \hat{y}(t, \theta) &= b_1 u(t-1) + \dots + b_{n_b} u(t-n_b) - a_1 y(t-1) - \dots - a_{n_a} y(t-n_a) \\ &= \theta^T \varphi(t), \end{aligned} \quad (39)$$

onde o vetor de parâmetros $\theta^T \in \mathbb{R}^m$, com $m = n_a + n_b$, e o vetor regressor $\varphi(t)$ são dados por:

$$\theta^T = [b_1 \quad \dots \quad b_{n_b} \quad a_1 \quad \dots \quad a_{n_a}], \quad (40)$$

$$\varphi(t) = [u(t-1) \quad \dots \quad u(t-n_b) \quad -y(t-1) \quad \dots \quad -y(t-n_a)]^T. \quad (41)$$

Como pode ser visto em (39) o vetor regressor para o modelo ARX não depende dos parâmetros, sendo esta uma propriedade bastante importante, tornando-o uma escolha muito popular em diversas aplicações (LJUNG, 1999).

O erro de predição do modelo ARX é obtido substituindo-se (39) em (33) como

$$\varepsilon_L^{\text{arx}}(t, \theta) = L(q) [y(t) - \theta^T \varphi(t)]. \quad (42)$$

2.3.2 Modelo de resposta ao impulso finita (FIR)

Adicionalmente à restrição do modelo ARX, (36), no modelo do tipo resposta ao impulso finita considera-se que $A(q, \theta) = 1$. Assim, o modelo FIR (do inglês *Finite Impulse Response*) é dado por

$$y(t) = B(q, \theta)u(t) + v(t), \quad (43)$$

e seu preditor é escrito como:

$$\hat{y}(t, \theta) = b_1 u(t-1) + \dots + b_{n_b} u(t-n_b). \quad (44)$$

Considerando-se o vetor de parâmetros $\theta \in \mathbb{R}^m$, com $m = n_b$, e o vetor regressor $\varphi(t)$ definidos como

$$\theta^T = [b_1 \quad \dots \quad b_{n_b}], \quad (45)$$

$$\varphi(t) = [u(t-1) \quad \dots \quad u(t-n_b)]^T, \quad (46)$$

o preditor de (44) é dado por

$$\hat{y}(t, \theta) = \theta^T \varphi(t). \quad (47)$$

Assim como o vetor regressor do modelo ARX, o regressor do modelo FIR também não depende dos parâmetros, sendo também bastante empregado para identificação de sistemas (LJUNG, 1999). Substituindo-se (44) em (33) obtém-se o erro de predição do modelo FIR

$$\varepsilon_L^{\text{fir}}(t, \theta) = L(q) [y(t) - \theta^T \varphi(t)]. \quad (48)$$

2.3.3 Modelo de erro na saída (OE)

Em modelos do tipo erro na saída OE (do inglês *Output Error*) considera-se que a relação entre a entrada e a saída não perturbada pode ser escrita como uma equação de diferenças linear e que a perturbação é um ruído branco (LJUNG, 1999), ou seja,

$$x(t) + f_1 x(t-1) + \dots + f_{n_f} x(t-n_f) = b_1 u(t-1) + \dots + b_{n_b} u(t-n_b), \quad (49)$$

e

$$y(t) = x(t) + v(t).$$

Seja $F(q, \theta)$ e $B(q, \theta)$ dados por (29) e (27), então o modelo OE é definido como

$$y(t) = \frac{B(q, \theta)}{F(q, \theta)} u(t) + v(t), \quad (50)$$

ou seja, $A(q, \theta) = K(q, \theta) = D(q, \theta) = 1$ em (25).

Veja que $x(t)$ em (49) nunca é observado mas pode ser construído a partir de $u(t)$ (LJUNG, 1999). Por conta disso, escreve-se $x(t)$ em função do vetor de parâmetros θ da seguinte forma:

$$x(t, \theta) + f_1 x(t-1, \theta) + \dots + f_{n_f} x(t-n_f, \theta) = b_1 u(t-1) + \dots + b_{n_b} u(t-n_b),$$

assim o preditor do modelo OE é dado por

$$\hat{y}(t, \theta) = \frac{B(q, \theta)}{F(q, \theta)} u(t) = x(t, \theta). \quad (51)$$

Este é o preditor natural e pode ser reescrito como

$$\hat{y}(t, \theta) = \varphi(t, \theta)\theta^T, \quad (52)$$

onde

$$\begin{aligned} \theta^T &= [b_1 \quad \dots \quad b_{n_b} \quad f_1 \quad \dots \quad f_{n_f}], \\ \varphi(t, \theta) &= [u(t-1) \quad \dots \quad u(t-n_b) \quad -x(t-1, \theta) \quad \dots \quad -x(t-n_f, \theta)]^T. \end{aligned}$$

Observe que o vetor regressor $\varphi(t, \theta)$ depende dos parâmetros e que $x(t-1, \theta)$ é calculado a partir de (51), ou seja, $x(t-k, \theta) = \hat{y}(t-k, \theta)$ com $k = 1, 2, \dots, n_f$ (LJUNG, 1999). O erro de predição do modelo OE é obtido substituindo-se (52) em (33) e dado por

$$\varepsilon_L^{\text{oe}}(t, \theta) = L(q)[y(t) - \varphi(t, \theta)\theta^T]. \quad (53)$$

2.3.4 Modelo Box-Jenkins (BJ)

No caso de estruturas de modelo do tipo Box-Jenkins assume-se que as funções racionais $G(q, \theta)$ e $H(q, \theta)$ em (24) são parametrizadas independentemente (LJUNG, 1999). Para tanto, considera-se $A(q, \theta) = 1$ em (25), obtendo-se o seguinte modelo

$$y(t) = \frac{B(q, \theta)}{F(q, \theta)}u(t) + \frac{K(q, \theta)}{D(q, \theta)}v(t),$$

a mesma consideração é feita para obter o preditor em (31), que é dado por

$$\hat{y}(t, \theta) = \frac{D(q, \theta)B(q, \theta)}{K(q, \theta)F(q, \theta)}u(t) - \frac{K(q, \theta) - D(q, \theta)}{K(q, \theta)}y(t). \quad (54)$$

Calcula-se o erro de predição da estrutura de modelo Box-Jenkins substituindo-se (54) em (33) como

$$\varepsilon_L^{\text{bj}}(t, \theta) = L(q) \left[\left(1 + \frac{K(q, \theta) - D(q, \theta)}{K(q, \theta)} \right) y(t) - \frac{D(q, \theta)B(q, \theta)}{K(q, \theta)F(q, \theta)} u(t) \right].$$

2.3.5 Nota sobre as estruturas de modelo

Os modelos ARX e FIR possuem a vantagem de que o vetor regressor não depende do vetor de parâmetros a serem estimados e a estimativa é obtida através de uma solução analítica. A estrutura ARX possui a vantagem de que também é possível estimar os parâmetros do denominador do controlador, como será apresentado no Capítulo 3. Quanto ao modelo FIR, o mesmo pode ser empregado no presente trabalho na estimativa da amplitude da resposta em frequência da inversa do modelo do processo para realizar uma pré-filtragem dos dados, como também descrito no Capítulo 3. No modelo de erro na saída o vetor regressor depende do vetor de parâmetros e, assim como para um modelo Box-Jenkins, a estimativa é obtida através de um problema de otimização não convexo. Por conta disso, esses dois últimos modelos são utilizados apenas no método OCI, apresentado mais adiante no presente capítulo.

2.4 Métodos de controle baseados em dados

Em métodos de controle baseados em dados o ajuste dos parâmetros do controlador é realizado *diretamente* a partir dos dados coletados durante um experimento no processo, sem um conhecimento profundo do comportamento do mesmo. A maioria dos métodos baseados em dados emprega o paradigma do modelo de referência, no qual define-se *a priori* o comportamento desejado para o processo em malha fechada. Além disso, também escolhe-se *a priori* a estrutura do controlador a ser projetado. Com base nestas escolhas, nos dados coletados e em uma função custo, essa abordagem busca projetar um controlador de modo que a resposta em malha fechada seja tão próxima quanto possível da resposta do modelo de referência desejado.

Ao longo do tempo, diferentes métodos baseados em dados foram propostos na literatura. Tais métodos podem ser classificados como iterativos e não iterativos. Nos métodos iterativos, os parâmetros do controlador ótimo são obtidos através de uma sequência de experimentos no processo real. A cada iteração, projeta-se um controlador com os dados coletados do processo em malha fechada com o controlador projetado na iteração anterior. As iterações são realizadas até que a função custo atinja um valor mínimo. Como exemplos desta classe de métodos, pode-se citar: o *Iterative Feedback Tuning* (IFT) (HJALMARSSON et al., 1998; HJALMARSSON, 2002) e o *Correlation-based Tuning* (CbT) (KARIMI; MIŠKOVIĆ; BONVIN, 2016; KARIMI; MIŠKOVIĆ; BONVIN, 2004). Em contrapartida, em métodos não iterativos, os parâmetros do controlador são estimados a partir dos dados coletados de um único experimento. O *Virtual Disturbance Feedback Tuning* (VDFT) (ECKHARD; CAMPESTRINI; BOEIRA, 2018), e o *Disturbance Correlation-based Tuning* (DCbT) (PINTO DA SILVA; ECKHARD, 2019; PINTO DA SILVA; ECKHARD, 2020b) são exemplos desta classe de métodos voltados para rejeição de perturbações. Para seguimento de referências, pode-se citar: o VRFT (CAMPI; LECCHINI; SAVARESI, 2000; CAMPI; LECCHINI; SAVARESI, 2002), o *Noniterative Correlation-based Tuning* (NCbT) (KARIMI; VAN HEUSDEN; BONVIN, 2007) e o OCI (CAMPESTRINI et al., 2012; CAMPESTRINI et al., 2017).

Nas próximas subseções são descritos os métodos VRFT e OCI. Além disso, são apresentados exemplos de simulação ilustrando a aplicação destes métodos.

2.4.1 *Virtual Reference Feedback Tuning*

O VRFT é um método bastante conhecido e bastante empregado na literatura. A ideia geral do método foi inicialmente proposta por Guardabassi e Savaresi (2000), onde os autores propuseram um método chamado de *Virtual Reference Direct Design* (VRD²). Naquele trabalho, os autores analisaram o caso de controladores com estrutura linear considerando que o problema de controle é bem posto, ou seja, o controlador ideal pode ser representado com a estrutura de controle escolhida. Além disso, é apresentada uma análise de convergência dos parâmetros quando os dados são afetados por ruído. O método VRFT, na sua forma descrita neste trabalho, foi proposto por Campi, Lecchini e Savaresi (2000), Campi, Lecchini e Savaresi (2002), e é baseado na ideia original de Guardabassi e Savaresi (2000). Campi, Lecchini e Savaresi (2002) propuseram como lidar com o caso em que o controlador ideal não pode ser representado com a estrutura de controle escolhida. Além disso, aqueles autores sugeriram a utilização do método de variáveis instrumentais para lidar com o ruído. Guardabassi e Savaresi (2001) investigaram a aplicação do método VRFT em diferentes estudos de caso simulados.

Ao longo dos anos, diversas aplicações e extensões para o método VRFT foram

propostas na literatura. Campestrini et al. (2011) abordaram o caso em que o processo possui zeros de fase não mínima. Naquele trabalho, é proposta a utilização de um modelo de referência flexível, cujos polos são fixados e cujos zeros são livres. Os zeros de fase não mínima do processo são estimados como zeros do modelo de referência flexível juntamente com os parâmetros do controlador. Uma extensão do método VRFT para o caso de processos multivariáveis é proposta por Nakamoto (2004), onde considera-se que a matriz de modelo de referência possui parametrização diagonal. Campestrini et al. (2016) propuseram uma extensão para o caso multivariável considerando que a matriz de modelo de referência é totalmente parametrizada. Naquele trabalho, também é proposto um filtro para lidar com o caso em que o controlador ideal não pertence à classe de controle. Gonçalves da Silva, Campestrini e Bazanella (2018) propuseram uma extensão do trabalho desenvolvido por Campestrini et al. (2011), onde é abordado o caso de processos multivariáveis com zeros de fase não mínima, considerando-se um modelo de referência flexível. Silva Garcia e Bazanella (2017) implementaram a versão original e recursiva, com fator de esquecimento, da solução de mínimos quadrados do método VRFT em um aplicativo, cujo objetivo era realizar a coleta de dados e o ajuste dos parâmetros do controlador.

Uma extensão do método para controladores com estrutura não linear é proposta por Campi e Savaresi (2006). O caso de controladores não lineares também foi estudado por Bazanella e Neuhaus (2014), onde foram consideradas classes de controladores diferentes das estudadas por Campi e Savaresi (2006). Além disso, Bazanella e Neuhaus (2014) apresentaram uma análise estatística para o caso ruidoso.

De acordo com a literatura, diversos trabalhos aplicam o método VRFT em processos reais. Previdi et al. (2010) aplicaram o VRFT para realizar o controle de velocidade de um manipulador manual auto-balanceado utilizado no transporte de pequenas cargas. A extensão multivariável proposta por Nakamoto (2004) foi empregada por Rojas et al. (2012) para o controle de uma estação de tratamento de águas residuais. Lenon Remes et al. (2019) utilizaram a abordagem tradicional do método VRFT para estimar controladores para conversores CC-CC. Para os mesmos conversores, Lenon Remes et al. (2021) aplicaram a abordagem com modelo de referência flexível, desenvolvida por Campestrini et al. (2011). Além disso, os autores apresentaram diretrizes de como escolher o modelo de referência, a estrutura de controle, e como projetar o experimento para coleta de dados, para os conversores utilizados. Além dos trabalhos mencionados, o VRFT foi empregado em diversos outros casos (FORMENTIN et al., 2019; NASCIMENTO; BERNARDO; DOTTA, 2020; RADRIZZANI et al., 2020; BORDIGNON; CAMPESTRINI, 2018; RALLO et al., 2016).

2.4.1.1 *Descrição do método*

A ideia geral do método consiste em transformar o problema de otimização do critério de seguimento de referência $J^{\text{MR}}(\rho)$ em um problema de identificação por erro de predição dos parâmetros do controlador $C(q, \rho)$. Quando um controlador com forma autorregressiva é escolhido, a identificação por erro de predição consiste em minimizar $J^{\text{VR}}(\rho)$, uma função custo quadrática nos parâmetros do modelo do controlador. Sob certas condições ideais, será mostrado que o mínimo da função custo quadrática do VRFT, $J^{\text{VR}}(\rho)$, é igual ao mínimo global da função custo $J^{\text{MR}}(\rho)$ (BAZANELLA; CAMPESTRINI; ECKHARD, 2011). Na sequência será apresentado o caso ideal e mais adiante será apresentado o caso em que o controlador ideal não pertence à classe de controle. Além disso, serão descritas as soluções tradicionais empregadas para estimar os parâmetros do controlador com o

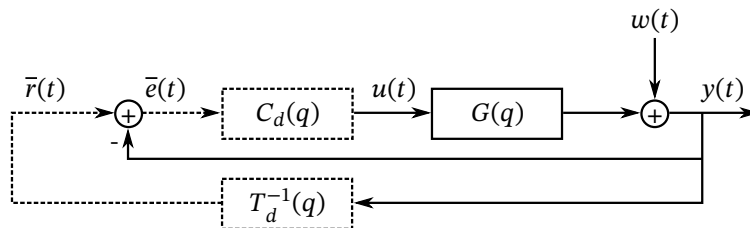
método VRFT.

Como mencionado anteriormente, o método VRFT é baseado no paradigma do modelo de referência, no qual escolhem-se *a priori* as especificações desejadas para o processo em malha fechada, ou seja, a função de transferência para o modelo de referência. Além disso, também escolhe-se a estrutura de controle, $C(q, \rho)$. Com base nestas escolhas e nos dados de entrada e saída $u(t)$ e $y(t)$ de um experimento em malha aberta ou fechada são gerados os sinais *virtuais* do método (CAMPI; LECCHINI; SAVARESI, 2002; BAZANELLA; CAMPESTRINI; ECKHARD, 2011). Para seguir a formulação do método, primeiramente, considere o caso ideal, caracterizado através da Suposição 2.2.

Suposição 2.2 (O caso ideal). Neste caso, as seguintes suposições são respeitadas: o controlador ideal pertence à classe de controle, Suposição 2.1, e o controlador é parametrizado linearmente, Definição 2.4. Além disso, assume-se que o processo não é afetado por ruído, ou seja, $\sigma_v^2 = 0$. ♦

Com base nas suposições do caso ideal, nas definições de $T_d(q)$ e $C(q, \rho)$ e nos sinais coletados, geram-se os sinais *virtuais* do método VRFT, como apresentado na Figura 3, referenciado como *experimento virtual*. Primeiramente, gera-se o sinal de *referência*

Figura 3 — Experimento virtual do método VRFT.



Fonte: adaptado de Bazanella, Campestrini e Eckhard (2011).

virtual $\bar{r}(t)$, a partir do sinal de saída $y(t)$ coletado no experimento, como descrito abaixo

$$\bar{r}(t) = T_d^{-1}(q)y(t), \quad (55)$$

sendo que este é o sinal ideal que deveria ser aplicado ao sistema em malha fechada desejado, $T_d(q)$, para gerar o sinal de saída coletado, $y(t)$. Em outras palavras, $\bar{r}(t)$ é o sinal ideal que deveria ser aplicado em um experimento em malha fechada considerando-se que o controlador ideal $C_d(q)$ estivesse operando na malha. Na sequência, gera-se $\bar{e}(t)$ que é o sinal de *erro virtual*:

$$\bar{e}(t) = \bar{r}(t) - y(t), \quad (56)$$

reescrito substituindo-se (55) em (56), como

$$\bar{e}(t) = [T_d^{-1}(q) - 1]y(t). \quad (57)$$

De posse dos dados de entrada $\bar{e}(t)$ e saída $u(t)$ do controlador ideal $C_d(q)$ do experimento virtual deseja-se estimar os parâmetros do controlador $C(q, \rho)$. Para tanto, considere que a saída do controlador a ser identificado é dada por

$$u(t) = C(q, \rho)\bar{e}(t) + \mu(t), \quad (58)$$

onde $\mu(t)$ é o erro do modelo. Considere também que um preditor $\hat{u}(t, \rho)$ para este modelo é dado por

$$\hat{u}(t, \rho) = C(q, \rho)\bar{e}(t).$$

Com base nisso, a função custo do erro de predição do método VRFT é definida como

$$J^{\text{VR}}(\rho) \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N |\varepsilon_{\text{VR}}(t, \rho)|^2, \quad (59)$$

$$\varepsilon_{\text{VR}}(t, \rho) = L(q) [u(t) - C(q, \rho)\bar{e}(t)], \quad (60)$$

onde $\varepsilon_{\text{VR}}(t, \rho)$ é o erro de predição da saída do controlador. Supondo que o controlador é parametrizado linearmente como na Definição 2.4, reescreve-se a função custo em (59) como

$$J^{\text{VR}}(\rho) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N |L(q) [u(t) - \rho^T \varphi(t)]|^2, \quad (61)$$

onde o vetor regressor é definido como

$$\varphi(t) = \bar{C}(q)\bar{e}(t). \quad (62)$$

Veja que aqui os sinais que formam o erro de predição são filtrados por um filtro $L(q)$ como no erro de predição da identificação de sistemas, apresentado em (33). Uma vez que o preditor é linear e invariante no tempo e os sinais $u(t)$ e $\rho^T \varphi(t)$ são escalares, a função custo (61) pode ser reescrita como

$$J^{\text{VR}}(\rho) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N |u_L(t) - \rho^T \varphi_L(t)|^2, \quad (63)$$

onde

$$u_L(t) = L(q)u(t), \quad (64)$$

$$\varphi_L(t) = L(q)\bar{C}(q) \left[T_d^{-1}(q) - 1 \right] y(t). \quad (65)$$

Na equação acima o vetor regressor é obtido substituindo-se (57) em (62). A estrutura do filtro $L(q)$ é ajustada para atingir alguns propósitos no método VRFT, que serão discutidos mais adiante no texto. O objetivo do método é obter o vetor de parâmetros ótimo ρ_{VR}^* que minimiza a função custo $J^{\text{VR}}(\rho)$:

$$\rho_{\text{VR}}^* = \underset{\rho}{\operatorname{argmin}} J^{\text{VR}}(\rho). \quad (66)$$

Note que o erro de predição pode ser reescrito substituindo-se (57) em (60) como

$$\varepsilon_{\text{VR}}(t, \rho) = L(q) \left[u(t) - C(q, \rho) \left(T_d^{-1}(q) - 1 \right) y(t) \right]$$

e novamente reescrito considerando-se que os dados são coletados em malha aberta (como definido em (1) assumindo-se $\sigma_v^2 = 0$)

$$\varepsilon_{\text{VR}}(t, \rho) = L(q) \left[u(t) - C(q, \rho) \left(T_d^{-1}(q) - 1 \right) G(q)u(t) \right]. \quad (67)$$

Considerando-se o experimento virtual mencionado anteriormente, onde deseja-se encontrar o controlador ideal $C_d(q)$, tem-se que o sinal $u(t)$ é a saída medida e $\bar{e}(t)$ sua entrada, logo, $u(t) = C_d(q)\bar{e}(t)$, que é reescrito como

$$u(t) = C_d(q) \left(T_d^{-1}(q) - 1 \right) G(q) u(t), \quad (68)$$

onde assume-se que os dados são de um experimento em malha aberta e não são afetados por ruído, ou seja, $\sigma_v^2 = 0$ em (1). Com base nessa definição, reescreve-se o erro de predição $\varepsilon_{VR}(t, \rho)$ substituindo-se (68) em (67), obtendo-se:

$$\begin{aligned} \varepsilon_{VR}(t, \rho) &= L(q) [C_d(q) - C(q, \rho)] \left[T_d^{-1}(q) - 1 \right] G(q) u(t) \\ &= L(q) [C_d(q) - C(q, \rho)] \frac{1 - T_d(q)}{T_d(q)} G(q) u(t) \\ &= L(q) [C_d(q) - C(q, \rho)] \frac{S_d(q)}{T_d(q)} G(q) u(t), \end{aligned} \quad (69)$$

onde $S_d(q)$ é a função de sensibilidade desejada dada por

$$S_d(q) = 1 - T_d(q). \quad (70)$$

Deseja-se agora comparar os dois critérios de desempenho $J^{VR}(\rho)$ e $J^{MR}(\rho)$. Para tanto, estas duas funções custo são escritas no domínio da frequência utilizando-se o teorema de Parseval, descrito abaixo.

Teorema 2.1 (Teorema de Parseval (OPPENHEIM; BUCK; SCHAFER, 1999)). *Se $x(t)$ e $X(e^{j\omega})$ são pares da transformada de Fourier, então*

$$\sum_{t=-\infty}^{\infty} |x(t)|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(e^{j\omega})|^2 d\omega.$$

De (69) e utilizando-se o teorema de Parseval reescreve-se a função custo $J^{VR}(\rho)$ no domínio da frequência da seguinte forma

$$J^{VR}(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|L(e^{j\omega})|^2 |C_d(e^{j\omega}) - C(e^{j\omega}, \rho)|^2 |S_d(e^{j\omega})|^2 |G(e^{j\omega})|^2}{|T_d(e^{j\omega})|^2} \Phi_u(e^{j\omega}) d\omega, \quad (71)$$

onde $\Phi_u(e^{j\omega})$ é o espectro do sinal $u(t)$.

Na sequência, o erro da resposta do modelo $\varepsilon_{MR}(t, \rho)$ em (14) pode ser reescrito como

$$\begin{aligned} \varepsilon_{MR}(t, \rho) &= \left[\frac{C_d(q)G(q)}{1 + C_d(q)G(q)} - \frac{C(q, \rho)G(q)}{1 + C(q, \rho)G(q)} \right] r(t) \\ &= \frac{[C_d(q) - C(q, \rho)]G(q)}{[1 + C_d(q)G(q)][1 + C(q, \rho)G(q)]} r(t) \\ &= [C_d(q) - C(q, \rho)]G(q)S_d(q)S(q, \rho)r(t), \end{aligned} \quad (72)$$

onde $S_d(q)$ é obtida de (6) com o controlador ideal. Utilizando-se o teorema de Parseval e substituindo-se (72) em (13) reescreve-se $J^{\text{MR}}(\rho)$ no domínio da frequência como

$$J^{\text{MR}}(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |C_d(e^{j\omega}) - C(e^{j\omega}, \rho)|^2 |S_d(e^{j\omega})|^2 \times |S(e^{j\omega}, \rho)|^2 |G(e^{j\omega})|^2 \Phi_r(e^{j\omega}) d\omega, \quad (73)$$

onde $\Phi_r(e^{j\omega})$ é o espectro do sinal $r(t)$.

Considerando-se o vetor de parâmetros ideal ρ_d tem-se que $C(e^{j\omega}, \rho_d) = C_d(e^{j\omega})$. Por conta disso, de (73) e (71) temos que $J^{\text{MR}}(\rho_d) = J^{\text{VR}}(\rho_d) = 0$. Desta forma, no caso ideal, confirma-se o que foi mencionado anteriormente: o mínimo da função custo do VRFT e o mínimo global da função custo do erro da resposta do modelo são iguais. Contudo, quando o controlador ideal não pertence à classe de controle, ou seja, quando a Suposição 2.1 não é respeitada, os mínimos dessas duas funções custo podem ser bem diferentes. Essa situação é mitigada ao escolher-se o filtro $L(e^{j\omega})$ de modo apropriado, como será apresentado nas próximas páginas. Veja que este não é o único caso em que o filtro $L(q)$ é utilizado no método VRFT. Um outro exemplo, é o caso em que o processo possui zeros de fase não mínima. Tais zeros devem aparecer como zeros do modelo de referência $T_d(q)$, o que é um problema, pois o sinal de erro virtual $\bar{e}(t)$ é gerado a partir da inversa de $T_d(q)$. Neste caso, uma opção é utilizar como filtro a própria função de transferência do modelo de referência (BAZANELLA; CAMPESTRINI; ECKHARD, 2011), ou seja,

$$L(q) = T_d(q),$$

desta forma o sinal de saída do controlador e o sinal de erro virtual são gerados como

$$\begin{aligned} u_L(t) &= T_d(q)u(t), \\ \bar{e}_L(t) &= T_d(q) \left[T_d^{-1}(q) - 1 \right] y(t) \\ &= [1 - T_d(q)]y(t). \end{aligned}$$

Esta é a alternativa favorecida no presente trabalho, embora não seja a única para lidar com este caso. Para identificar os zeros de fase não mínima automaticamente recomenda-se ao leitor interessado o trabalho de Campestrini et al. (2011).

Até o presente momento, assumiu-se que todas as suposições do caso ideal (Suposição 2.2) são respeitadas. Considere agora que o controlador ideal não pertence à classe de controle, ou seja, a Suposição 2.1 é violada. Em outras palavras, assume-se que o controlador escolhido é subparametrizado, i.e., a estrutura do controlador escolhida não é capaz de representar exatamente a estrutura do controlador ideal. Neste caso, não há garantia de que os vetores de parâmetros ótimos, ρ_{MR}^* e ρ_{VR}^* , com os quais obtém-se os mínimos de $J^{\text{MR}}(\rho_{\text{MR}}^*)$ e de $J^{\text{VR}}(\rho_{\text{VR}}^*)$, serão iguais. De modo a aproximar os mínimos das duas funções custo, observe que (71) e (73) seriam iguais se o filtro $L(e^{j\omega})$ tivesse a seguinte estrutura (CAMPI; LECCHINI; SAVARESI, 2002)

$$|L(e^{j\omega})|^2 = |T_d(e^{j\omega})|^2 |S(e^{j\omega}, \rho)|^2 \frac{\Phi_r(e^{j\omega})}{\Phi_u(e^{j\omega})}, \quad \forall \omega \in [-\pi; \pi]. \quad (74)$$

Porém, este filtro depende da função de sensibilidade $S(e^{j\omega}, \rho)$, (6), que é desconhecida pois depende do modelo do processo $G(q)$. Portanto, a implementação do filtro depende de

alguma aproximação da estrutura em (74). Quanto melhor a aproximação, mais próximos serão os mínimos das funções custo $J^{\text{VR}}(\rho)$ e $J^{\text{MR}}(\rho)$. Embora, provavelmente, essa não seja a única escolha sensata para uma aproximação de $S(q, \rho)$, o método VRFT utiliza a seguinte aproximação (CAMPI; LECCHINI; SAVARESI, 2002; BAZANELLA; CAMPESTRINI; ECKHARD, 2011):

$$S(q, \rho) \approx S_d(q),$$

ou seja, considera-se que o comportamento desejado em malha fechada $T_d(q)$ não é muito distante do comportamento que pode ser alcançado com o controlador $C(q, \rho)$. A partir desta aproximação e de (70) tem-se que:

$$|S(e^{j\omega}, \rho)|^2 \approx |S_d(e^{j\omega})|^2 = |1 - T_d(e^{j\omega})|^2. \quad (75)$$

Essa é uma aproximação plausível, uma vez que é esperado que as duas funções de sensibilidade $S(q, \rho)$ e $S_d(q)$ sejam bastante próximas ao redor do mínimo (BAZANELLA; CAMPESTRINI; ECKHARD, 2011). Com base nisso, reescreve-se o filtro $L(e^{j\omega})$, substituindo-se (75) em (74), como:

$$|L(e^{j\omega})|^2 \approx |T_d(e^{j\omega})|^2 |1 - T_d(e^{j\omega})|^2 \frac{\Phi_r(e^{j\omega})}{\Phi_u(e^{j\omega})}, \quad \forall \omega \in [-\pi; \pi] \quad (76)$$

Observe agora que o filtro depende dos espectros dos sinais $r(t)$ e $u(t)$. Veja que $\Phi_u(e^{j\omega})$ é conhecido apenas quando o sinal de entrada é escolhido pelo projetista, caso contrário, o mesmo precisa ser estimado (BAZANELLA; CAMPESTRINI; ECKHARD, 2011). Assumindo-se que o sinal de referência que é normalmente aplicado no processo é do mesmo tipo que o sinal de entrada em malha aberta tem-se que $\Phi_r(e^{j\omega}) \approx \Phi_u(e^{j\omega})$ (CAMPI; LECCHINI; SAVARESI, 2002), logo

$$\frac{\Phi_r(e^{j\omega})}{\Phi_u(e^{j\omega})} \approx 1.$$

Com base nesta aproximação, o filtro para o caso em que a Suposição 2.1 é violada, pode ser simplificado para:

$$|L(e^{j\omega})|^2 \approx |T_d(e^{j\omega})|^2 |1 - T_d(e^{j\omega})|^2, \quad \forall \omega \in [-\pi; \pi]. \quad (77)$$

Em suma, quando as suposições do caso ideal são respeitadas o vetor de parâmetros estimado com o VRFT é o ideal, que é o mínimo de $J^{\text{MR}}(\rho)$. Como apresentado acima, quando a suposição de que o controlador ideal não pertence à classe de controle é moderadamente violada, o mínimo da função custo do VRFT é aproximado do mínimo de $J^{\text{MR}}(\rho)$ por um filtro $L(q)$ cuja estrutura respeita a condição (77). Conforme a violação da Suposição 2.1 se torna mais pronunciada, o filtro ainda ajuda, mas perde gradualmente sua efetividade. Portanto, se essa suposição é fortemente violada, ou seja, se o melhor desempenho que pode ser obtido com a estrutura de controle disponível é muito diferente do comportamento desejado, então o método se torna ineficiente, mesmo com a utilização do filtro. Isso é esperado, pois a formulação do filtro (77) é baseada na aproximação (75). Se tal aproximação não for válida, a resposta em malha fechada obtida com os parâmetros estimados com o VRFT pode ser bastante diferente da desejada e até mesmo instável em alguns casos (BAZANELLA; CAMPESTRINI; ECKHARD, 2011). Por conta disso, como mencionado anteriormente, nos exemplos de simulação do presente trabalho considera-se que a Suposição 2.1 é moderadamente violada. Na sequência, serão apresentadas as soluções originalmente propostas na literatura para o método VRFT.

2.4.1.2 Soluções empregadas na literatura

A formulação para o método VRFT apresentada até agora considera que a saída do processo não é corrompida por ruído. Contudo, esse caso não é realista pois normalmente os dados coletados de processos reais são ruidosos. Esta seção tem como objetivo apresentar a solução original do método VRFT, que é polarizada na presença de ruído, e a alternativa proposta na literatura para reduzir a polarização da estimativa.

Considere a equação (61), este é um problema de mínimos quadrados, cuja solução é dada pela equação normal

$$\rho_* = \lim_{N \rightarrow \infty} \left[\sum_{t=1}^N \varphi_L(t) \varphi_L^T(t) \right]^{-1} \left[\sum_{t=1}^N \varphi_L(t) u_L(t) \right], \quad (78)$$

onde ρ_* é o valor assintótico de ρ . A solução da equação acima existe e é única se a matriz $\lim_{N \rightarrow \infty} \sum_{t=1}^N \varphi_L(t) \varphi_L^T(t)$ tem posto cheio, ou seja, se o vetor $\varphi_L(t)$ é persistentemente excitante.

Definição 2.5 (Persistência de excitação (BAZANELLA; CAMPESTRINI; ECKHARD, 2011)). Um vetor quasi-estacionário $\varphi_L(t)$ é dito persistentemente excitante se

$$\lim_{N \rightarrow \infty} \sum_{t=1}^N \varphi_L(t) \varphi_L^T(t) > 0.$$

Observe que $\varphi_L(t)$ é formado a partir de $u(t)$, logo, a persistência de excitação, que é necessária para que a solução da equação normal seja única, pode ser transferida para condições no sinal $u(t)$. Dessa forma, pode-se dizer que o vetor regressor $\varphi_L(t)$ é persistentemente excitante se o sinal $u(t)$ é suficientemente rico de ordem p (BAZANELLA; CAMPESTRINI; ECKHARD, 2011).

Definição 2.6 (Processo suficientemente rico de ordem p (BAZANELLA; CAMPESTRINI; ECKHARD, 2011)). Um processo quasi-estacionário é dito suficientemente rico de ordem p se seu espectro possui ao menos p componentes diferentes de zero.

Como (78) não é uma quantidade diretamente calculável, pois depende do valor esperado, o valor ρ_* é estimado para um conjunto de N amostras coletadas de um experimento, como

$$\begin{aligned} \hat{\rho} &= \left[\sum_{t=1}^N \varphi_L(t) \varphi_L^T(t) \right]^{-1} \sum_{t=1}^N \varphi_L(t) u_L(t) \\ &= P^{-1} \sum_{t=1}^N \varphi_L(t) u_L(t), \end{aligned} \quad (79)$$

onde P é a matriz de informação definida como

$$P = \sum_{t=1}^N \varphi_L(t) \varphi_L^T(t). \quad (80)$$

Na literatura, a solução (79) também é conhecida como OLS. O vetor de parâmetros estimado $\hat{\rho}$ só é igual ao valor assintótico ρ_* no caso ideal, ou seja, quando a Suposição 2.2

é respeitada. No caso em que os dados são afetados por ruído, a solução $\hat{\rho}$ não é mais um valor determinístico e sim uma variável aleatória, uma vez que o vetor regressor $\varphi_L(t)$ é construído a partir do sinal de saída coletado $y(t)$, que é afetado por ruído. Em suma, o método VRFT identifica o vetor de parâmetros ideal quando a Suposição 2.2 é respeitada e o sinal $u(t)$ é suficientemente rico de ordem m (BAZANELLA; CAMPESTRINI; ECKHARD, 2011).

O erro entre a variável aleatória $\hat{\rho}$ e o vetor de parâmetros ideal ρ_d é dado por (BAZANELLA; CAMPESTRINI; ECKHARD, 2011):

$$\hat{\rho} - \rho_d = (\hat{\rho} - \rho_*) + (\rho_* - \rho_d),$$

onde $\hat{\rho} - \rho_*$ é chamado de *erro de variância* e $\rho_* - \rho_d$ é chamado de *erro de polarização*. Uma estimativa é dita consistente se a mesma tende ao valor correto. Se uma estimativa é consistente o erro de polarização é zero, e pode-se fazer com que a mesma seja arbitrariamente próxima do valor ideal ao aumentar-se o número de amostras coletadas, uma vez que a variância da estimativa diminui conforme o número de amostras tende ao infinito. Em contrapartida, se uma estimativa não é consistente o termo de erro de polarização não é zero, assim sendo, aumentar o número de amostras não mudará este termo (BAZANELLA; CAMPESTRINI; ECKHARD, 2011). A estimativa $\hat{\rho}$ no caso do VRFT é caracterizada através do Teorema 2.2.

Teorema 2.2 (de Bazanella, Campestrini e Eckhard (2011)). *Seja a Suposição 2.1 respeitada. A estimativa (78) do controlador ideal obtida com o método VRFT é polarizada pela seguinte quantidade*

$$\text{polarização} = \bar{E}[\varphi_L(t)\varphi_L^T(t)]^{-1} \bar{E}[\varphi_L(t)G^{-1}(q)H(q)v(t)].$$

Com base neste teorema, conclui-se que a polarização da estimativa do método VRFT é zero quando $\bar{E}[\varphi_L(t)G^{-1}(q)H(q)v(t)] = 0$. Essa condição é dificilmente respeitada, pois $\varphi_L(t)$ é formado a partir de valores passados de $y(t)$, e por sua vez, $y(t)$ é afetado por amostras passadas do ruído $v(t)$. Assim, na maioria das situações reais, a estimativa obtida com a solução original do método VRFT é polarizada, a menos que a contribuição do ruído seja negligenciável.

Nota 2.2. O problema de identificação dos parâmetros do controlador com o método VRFT não é um problema de identificação padrão. Se os dados são coletados em malha aberta o ruído afeta a entrada do controlador a ser identificado e não sua saída. Em contrapartida, quando são empregados dados de malha fechada o ruído está presente nos sinais de entrada e saída do controlador.

Para mitigar o problema da polarização, Campi, Lecchini e Savaresi (2002) propuseram a utilização do método das variáveis instrumentais (VI). Neste caso, a VI $\zeta(t) \in \mathbb{R}^m$ é escolhida de modo a satisfazer as seguintes condições

$$\bar{E}[\zeta(t)\varphi^T(t)] > 0,$$

e

$$\bar{E}[\zeta(t)v(s)] = 0 \quad \forall t, s.$$

Assim, deseja-se um vetor $\zeta(t)$ que seja fortemente correlacionado com o vetor regressor $\varphi(t)$ e descorrelacionado com o ruído $v(t)$. A partir da escolha da VI, a estimativa dos parâmetros com o método VRFT é obtida como:

$$\hat{\rho} = \left[\sum_{t=1}^N \zeta_L(t) \varphi_L^T(t) \right]^{-1} \left[\sum_{t=1}^N \zeta_L(t) u_L(t) \right], \quad (81)$$

onde $\zeta_L(t) = L(q)\zeta(t)$. Campi, Lecchini e Savaresi (2002) sugerem duas opções para a VI do VRFT quando utiliza-se dados coletados em malha aberta: a primeira opção consiste em coletar dados de um segundo experimento no processo aplicando-se o mesmo sinal de entrada; a segunda opção é estimar um modelo para o processo e simular um segundo experimento. Abaixo estas duas opções são descritas de modo detalhado.

Segundo experimento: neste caso, o mesmo sinal de entrada $u(t)$ é utilizado para coletar os dados de um segundo experimento em malha aberta no processo. Assim, tem-se os sinais $u(t)$, $y_1(t)$ e $y_2(t)$ (onde o subscrito indica de qual experimento os dados de saída foram coletados). A partir destes dados gera-se $\varphi_L(t)$ com os dados do primeiro experimento, substituindo-se $y_1(t)$ em (65) e a VI é obtida utilizando-se o sinal de saída do segundo experimento, como descrito abaixo

$$\zeta_L(t) = L(q)\bar{C}(q) \left[T_d^{-1}(q) - 1 \right] y_2(t). \quad (82)$$

Em seguida, gera-se o sinal $u_L(t)$ como em (64). A estimativa dos parâmetros é obtida a partir de (81). Observe que os sinais $y_1(t)$ e $y_2(t)$ são diferentes, uma vez que são afetados por realizações de ruído diferentes. Além disso, assume-se que as realizações de ruído são descorrelacionadas. Portanto, a estimativa (81) não é polarizada (BAZANELLA; CAMPESTRINI; ECKHARD, 2011). Uma alternativa à coleta de dados de um segundo experimento consiste em realizar um único experimento com um sinal de entrada periódico e dividir os dados coletados em dois conjuntos, simulando dois experimentos. Embora a abordagem com VI forneça uma estimativa não polarizada, ela produz uma estimativa pouco precisa para os parâmetros (SÖDERSTRÖM, 2007; SÖDERSTRÖM, 2006; SÖDERSTRÖM, 2018). Esta característica pode levar a uma significativa perda de desempenho e até mesmo a instabilidade em malha fechada (RALLO et al., 2016).

Identificação do processo: nesta abordagem estima-se um modelo para o processo, $\hat{G}(q)$, a partir dos dados de entrada e saída coletados, $u(t)$ e $y(t)$, respectivamente. Em seguida, a VI é construída com o sinal de saída gerado a partir do modelo estimado, assim gera-se $\hat{y}(t) = \hat{G}(q)u(t)$ e a VI é obtida como

$$\zeta_L(t) = L(q)\bar{C}(q) \left[T_d^{-1}(q) - 1 \right] \hat{y}(t). \quad (83)$$

A estimativa dos parâmetros do controlador é obtida de (81), utilizando-se (83), $\varphi_L(t)$ obtido em (65) e $u_L(t)$ gerado como em (64). Neste caso, a qualidade da estimativa depende da qualidade da estimativa do modelo do processo $\hat{G}(q)$ (VAN HEUSDEN; KARIMI; SÖDERSTRÖM, 2011). Em outras palavras, quanto melhor for a estimativa de $\hat{G}(q)$ mais próximo $\hat{y}(t)$ estará da saída real do processo, que por ser simulada não sofre com o ruído. Por conta disso, $\zeta_L(t)$ será descorrelacionada do ruído e fortemente correlacionada com o vetor regressor $\varphi_L(t)$, como desejado.

É importante mencionar, que no presente trabalho será utilizada a VI que requer a execução de um segundo experimento no processo, pois esta é a VI comumente empregada na literatura do método VRFT. Na sequência, apresenta-se um exemplo de simulação ilustrando o efeito do ruído na estimativa dos parâmetros do controlador.

2.4.1.3 Exemplo motivacional

O objetivo desta subseção é apresentar um exemplo ilustrativo da estimativa dos parâmetros do controlador com o VRFT. A configuração deste exemplo é baseada em um dos estudos de caso de Campestrini et al. (2017). O processo possui tempo de acomodação de 35 amostras e tempo de subida de 19 amostras, sua função de transferência é dada por

$$G(q) = \frac{0,5(q-0,8)}{(q-0,7)(q-0,9)}. \quad (84)$$

Para obter uma resposta em malha fechada mais rápida do que a resposta em malha aberta e com tempo de subida mais rápido, escolheu-se a seguinte função de transferência para o modelo de referência

$$T_d(q) = \frac{0,21675(q-0,8)(q-0,8632)}{(q-0,6069)(q^2-1,776q+0,7914)}, \quad (85)$$

que possui tempo de acomodação de 18 amostras e tempo de subida de 11 amostras. O controlador ideal, calculado substituindo-se (84) e (85) em (16), é dado por

$$C_d(q) = \underbrace{[0,3742 \quad 0,0593]}_{\rho_d^T} \underbrace{\left[1 \quad \frac{q}{q-1}\right]^T}_{\bar{c}(q)} = \frac{0,4335(q-0,8632)}{q-1}. \quad (86)$$

O objetivo é projetar um controlador do tipo PI com a seguinte estrutura

$$C(q, \rho) = [K_p \quad K_i] \left[1 \quad \frac{q}{q-1}\right]^T. \quad (87)$$

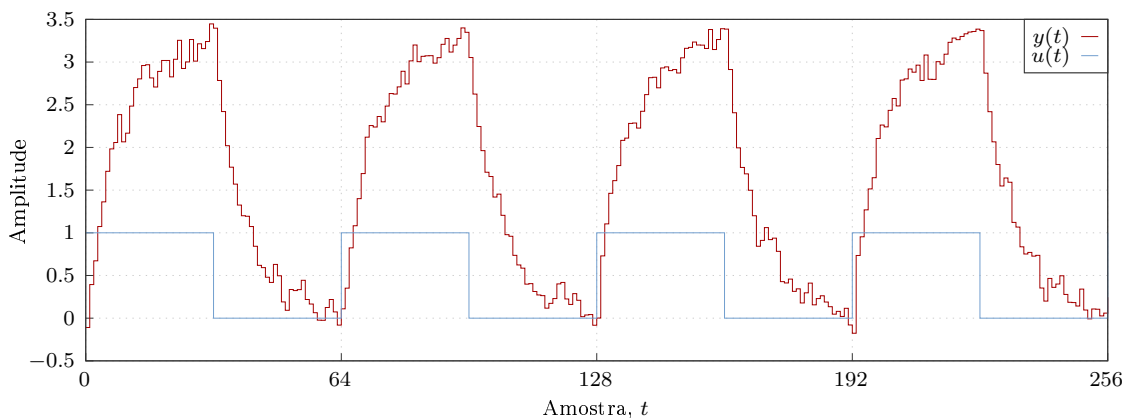
Como a estrutura do controlador a ser projetado é igual à estrutura do controlador ideal, a Suposição 2.1 é respeitada. A saída do processo é corrompida por um ruído colorido gerado a partir de um ruído branco com variância $\sigma_v^2 = 0,01$ filtrado através da seguinte função de transferência

$$H(q) = \frac{q}{q-0,3}. \quad (88)$$

Foram executadas 100 rodadas de Monte Carlo e os dados de entrada e saída são coletados de uma simulação em malha aberta. Para cada simulação gera-se uma realização diferente do ruído, que afeta o sinal de saída $y(t)$. Um sinal do tipo onda quadrada com níveis 0 e 1 e período de 64 amostras é aplicado como sinal de entrada $u(t)$. Neste caso, são coletados 16 períodos da onda quadrada, totalizando 1024 amostras. A Figura 4 apresenta as primeiras 256 amostras do sinal de entrada (linha contínua azul) e as respectivas amostras de uma realização do sinal de saída coletado (linha contínua vermelha).

A partir dos dados coletados, da escolha do modelo de referência e da estrutura de controle, estimam-se os parâmetros do controlador utilizando-se as soluções OLS e VI para cada realização. A estimativa OLS é obtida como em (79), com $L(q) = 1$, uma vez que o controlador ideal pertence à classe de controle. Para a estimativa com VI o sinal de saída é dividido em duas partes com o mesmo número de amostras, a primeira parte $y_1(t)$ possui as primeiras 512 amostras e a segunda parte $y_2(t)$ as demais amostras do sinal. A estimativa com VI é obtida como em (81), onde $u(t)$ possui as primeiras 512 amostras do sinal de entrada coletado e $L(q) = 1$. As estatísticas dos parâmetros estimados com as soluções OLS e VI e o vetor de parâmetros ideal são apresentados nas respectivas colunas

Figura 4 — Sinal de entrada e uma realização do sinal de saída coletado no exemplo da Seção 2.4.1.3.



Fonte: elaborado pela autora.

Tabela 1 — Média e desvio padrão dos parâmetros estimados com o VRFT e as soluções OLS e VI no exemplo da Seção 2.4.1.3.

ρ	ρ_d	Estimativas	
		OLS	VI
K_p	0,3742	0,2222 ($\pm 0,0053$)	0,3740 ($\pm 0,0145$)
K_i	0,0593	0,0592 ($\pm 0,0002$)	0,0593 ($\pm 0,0002$)

da Tabela 1. Observe que a solução OLS apresenta polarização e uma pequena variância. A solução com VI apresenta uma pequena polarização, mas uma variância mais acentuada do que com OLS, como esperado.

Como mencionado anteriormente, a abordagem com VI produz uma estimativa não polarizada, mas pouco precisa para os parâmetros (SÖDERSTRÖM, 2007), podendo levar a um comportamento em malha fechada bastante diferente do desejado e até mesmo a instabilidade (RALLO et al., 2016). Por conta disso, um dos temas do presente trabalho consiste em estimar os parâmetros do controlador com o método VRFT através de outras soluções de mínimos quadrados. Tais soluções são apresentadas no Capítulo 3.

2.4.2 *Optimal Controller Identification*

O método OCI, assim como o VRFT, é um método baseado em dados não iterativo (CAMPESTRINI et al., 2012; CAMPESTRINI et al., 2017). Neste método também escolhe-se *a priori* o comportamento desejado para o processo em malha fechada e um controlador com estrutura fixa. O problema de identificação dos parâmetros é formulado como a identificação por erro de predição da inversa do controlador ótimo. A partir da escolha do modelo de referência, da estrutura do controlador e dos dados de entrada e saída de um único experimento estimam-se os parâmetros do controlador. Tal estimativa é obtida através de um problema de otimização e, considerando algumas restrições quanto ao sinal de entrada, obtém-se uma estimativa não polarizada.

A formulação do método OCI para o caso monovariável é apresentada por Campestrini et al. (2017), onde os autores também mostram que a identificação *direta* dos parâmetros

do controlador (métodos baseados em dados) resulta em melhores resultados em malha fechada do que a identificação *indireta* do controlador (identificação do modelo e posterior cálculo do controlador), no caso de controladores subparametrizados, ou seja, quando o controlador ideal não pertence à classe de controle.

A extensão para o caso multivariável foi inicialmente proposta por Huff, Gonçalves da Silva e Campestrini (2018). Naquele trabalho, os autores aplicaram a versão multivariável do OCI em um sistema de refrigeração por compressão de vapor. Contudo, o método foi validado considerando-se que os dados não eram corrompidos com ruído. Huff et al. (2019) exploraram a utilização da formulação multivariável do método OCI considerando dados ruidosos. Neste caso, o problema de otimização se torna mais complexo, devido à estrutura da inversa do controlador, impossibilitando a utilização de ferramentas comerciais normalmente empregadas, como por exemplo, a toolbox `ident` do Matlab. Por conta disso, os autores propuseram uma solução específica para o problema de otimização para o caso multivariável. Segundo os autores, uma estimativa não polarizada é obtida, a partir dos dados de entrada e saída de um único experimento, quando o controlador ideal pertence à classe de controle. Além disso, o método foi validado considerando também o caso em que a Suposição 2.1 é violada e o caso em que são coletados dados em malha fechada.

Varriale da Silva e Campestrini (2019) investigaram o efeito de perturbações na estimativa dos parâmetros do controlador através de métodos baseados em dados. Naquele trabalho, foi considerado que os dados eram coletados em malha fechada, o processo não era afetado por ruído, e foram aplicados diferentes tipos de perturbações. A estimativa dos parâmetros foi realizada através dos métodos VRFT e OCI, onde os sinais de entrada e saída foram filtrados através do filtro ótimo do VRFT, (77), antes de serem utilizados nos métodos.

2.4.2.1 Descrição do método

A ideia principal do método consiste em escrever o processo em função do modelo de referência escolhido e do controlador a ser identificado e então realizar a identificação da inversa do controlador e do modelo do ruído (HUFF; GONÇALVES DA SILVA; CAMPESTRINI, 2018; CAMPESTRINI et al., 2017). Como mencionado anteriormente, o método OCI é baseado no paradigma do modelo de referência, e por conta disso, minimiza-se o critério de seguimento de referência, dado em (13)–(14).

Primeiramente, considere que a função de transferência do processo $G(q)$ em (1), é escrita em função do controlador ideal. Isto é obtido escrevendo-se (16) em função de $G(q)$, como abaixo

$$G(q) = \frac{1}{C_d(q)} \frac{T_d(q)}{(1 - T_d(q))}.$$

Na sequência, um modelo para o processo pode ser obtido em função do controlador a ser identificado:

$$G(q, \rho) = \frac{1}{C(q, \rho)} \frac{T_d(q)}{(1 - T_d(q))}, \quad (89)$$

e o objetivo é encontrar uma estimativa $C(q, \hat{\rho})$ para o controlador ideal $C_d(q)$ dentro da classe de controle escolhida \mathcal{C} , definida em (15) (CAMPESTRINI et al., 2017).

Considere agora que o controlador $C(q, \rho)$ é dividido em duas partes

$$C(q, \rho) = C_I(q, \rho)C_F(q), \quad (90)$$

onde $C_I(q, \rho)$ é a parte a ser identificada e $C_F(q)$ é a parte fixa. Neste caso, assume-se que o numerador de $C_I(q, \rho)$ é mônico, para tornar esta fatoração única e para facilitar a

escrita do problema no contexto da estimativa por erro de predição. O tipo de estrutura em (90) é particularmente conveniente pois permite a incorporação de qualquer informação previamente conhecida sobre o controlador, como por exemplo um polo em $q = 1$ para garantir tanto seguimento de referências quanto rejeição de perturbações constantes, ambas com erro nulo.

Com base nestas considerações, o modelo de entrada e saída em (1) é reescrito, substituindo-se (89) e (90), como

$$y(t, \theta) = \frac{1}{C_I(q, \rho)} \frac{T_d(q)}{C_F(q)(1 - T_d(q))} u(t) + H(q, \eta)v(t), \quad (91)$$

onde $\theta = [\rho^T \ \eta^T]^T$, $\eta \in \mathbb{R}^{n_h}$ é um vetor de parâmetros adicional aparecendo no modelo do ruído e $H(q, \eta)$ é uma estrutura de modelo paramétrica para $H(q)$. Veja que (91) pode ser reescrita como

$$y(t, \theta) = \tilde{C}(q, \rho)\tilde{u}(t) + H(q, \eta)v(t), \quad (92)$$

onde $\tilde{C}(q, \rho) = C_I^{-1}(q, \rho)$ e $\tilde{u}(t) = [C_F(q)(1 - T_d(q))]^{-1}T_d(q)u(t)$. Observe que o denominador de $\tilde{C}(q, \rho)$ é mônico, uma vez que assume-se que o numerador de $C_I(q, \rho)$ é mônico. Como o problema de identificação do controlador foi reescrito como um problema de identificação da inversa do controlador ideal, (91), o projeto do controlador segue como um problema de identificação padrão. Assim sendo, a partir de um conjunto de N dados coletados

$$\mathcal{Z}_{\text{oci}}^N = \{\tilde{u}(1) \ y(1) \ \dots \ \tilde{u}(N) \ y(N)\}$$

a estimativa dos parâmetros $\hat{\theta} = [\hat{\rho}^T \ \hat{\eta}^T]^T$ é dada por

$$\hat{\theta} = \arg \min_{\theta} J^{\text{OCI}}(\theta).$$

A função custo $J^{\text{OCI}}(\theta)$ é dada por

$$J^{\text{OCI}}(\theta) = \frac{1}{N} \sum_{t=1}^N |\varepsilon_{\text{OCI}}(t, \theta)|^2 \quad (93)$$

$$\varepsilon_{\text{OCI}}(t, \theta) \triangleq y(t) - \hat{y}(t, \theta), \quad (94)$$

onde $\varepsilon_{\text{OCI}}(t, \theta)$ é o erro de predição do controlador, e o preditor ótimo $\hat{y}(t, \theta)$ para o modelo (91) é definido como

$$\hat{y}(t, \theta) = H^{-1}(q, \eta)\tilde{C}(q, \rho)\tilde{u}(t) + [1 - H^{-1}(q, \eta)]y(t). \quad (95)$$

A partir de $\tilde{C}(q, \rho)$ e (90) obtém-se o controlador ótimo estimado, como apresentado abaixo

$$C(q, \hat{\rho}) = \frac{1}{\tilde{C}(q, \rho)} C_F(q).$$

Assim sendo, o problema de minimizar a função custo de desempenho de seguimento de referência $J^{\text{MR}}(\rho)$, que depende de $G(q)$, é transformado em um problema de minimizar a função custo $J^{\text{OCI}}(\theta)$, que depende puramente dos dados coletados e não depende do modelo do processo.

Agora, considere o erro de predição em (94) reescrito substituindo-se (95) e rearranjando

$$\varepsilon_{\text{OCI}}(t, \theta) = H^{-1}(q, \eta) [(\tilde{C}_d(q) - \tilde{C}(q, \rho))\tilde{u}(t) + w(t)], \quad (96)$$

onde $\tilde{C}_d(q) = C_F(q)C_d^{-1}(q)$. A função custo (93) pode ser reescrita no domínio da frequência substituindo-se (96) e então aplicando-se o teorema de Parseval (Teorema 2.1), chegando-se a

$$J^{\text{OCI}}(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H(e^{j\omega}, \eta)|^{-2} [(\tilde{C}_d(e^{j\omega}) - \tilde{C}(e^{j\omega}, \rho)) \Phi_{\tilde{u}}(\omega) + \Phi_w(\omega)] d\omega, \quad (97)$$

onde $\Phi_w(\omega)$ é o espectro do ruído e $\Phi_{\tilde{u}}(\omega)$ é o espectro de $\tilde{u}(t)$, ou seja,

$$\Phi_{\tilde{u}}(\omega) = |C_F(e^{j\omega})(1 - T_d(e^{j\omega}))|^{-2} |T_d(e^{j\omega})|^2 \Phi_u(\omega).$$

De modo a mostrar que o controlador estimado $\tilde{C}(q, \rho)$ com o método OCI tende assintoticamente ao controlador ideal, considere a suposição abaixo.

Suposição 2.3 (Riqueza do sinal de entrada (CAMPESTRINI et al., 2017)). O sinal de entrada filtrado $\tilde{u}(t)$ é suficientemente rico para tornar o experimento informativo com respeito à estrutura escolhida para $\tilde{C}(q, \rho)$. ◆

Considere que o controlador ideal pertence à classe de controle (Suposição 2.1), os dados são coletados em malha aberta, o sinal de entrada é suficiente rico (Suposição 2.3) e $\tilde{C}(q, \rho)$ e $H(q, \eta)$ são parametrizados independentemente, ou seja, $\frac{\partial H(q, \eta)}{\partial \rho} = 0$. Então, quando $N \rightarrow \infty$ (LJUNG, 1999; CAMPESTRINI et al., 2012):

$$C(q, \hat{\rho}_N) \rightarrow C_d(q). \quad (98)$$

Observe que, quando as considerações acima descritas são respeitadas, o controlador estimado converge para o controlador ideal, independente da parametrização do modelo do ruído. Em outras palavras, quando os dados são coletados em malha aberta, não se faz necessário estimar um modelo para o ruído, podendo-se utilizar um modelo do tipo OE para estimar os parâmetros, ou seja, $H(q, \eta) = 1$, como descrito na Subseção 2.3.3. Neste caso, a estimativa pode ser obtida, por exemplo, utilizando-se a função `oe` da toolbox `ident` do Matlab (CAMPESTRINI et al., 2017). Contudo, se o modelo do ruído possuir uma estrutura mais complexa, incluir uma estrutura de modelo para $H(q, \eta)$, capaz de representar $H(q)$, melhorará a variância da estimativa do controlador, mas a polarização continuará sendo nula (CAMPESTRINI et al., 2012).

Se a Suposição 2.1 é respeitada e dados de malha fechada são coletados então (98) é válida desde que $\exists \eta_d$ tal que $H(q, \eta_d) = H(q)$, onde η_d é o vetor de parâmetros ideal para o modelo do ruído. Neste caso, a estimativa dos parâmetros será polarizada se a estrutura de modelo para o ruído $H(q, \eta)$ não for capaz de representar $H(q)$ corretamente (CAMPESTRINI et al., 2012). Uma opção é utilizar um modelo do tipo Box-Jenkins, como descrito na Subseção 2.3.4, e a estimativa do parâmetros pode ser obtida utilizando-se a função `bj` da toolbox `ident` do Matlab, por exemplo.

Em suma, desde que as considerações descritas acima sejam respeitadas, para os casos de dados de malha aberta e fechada, identifica-se o vetor de parâmetros ideal θ_d com o método OCI, sendo que $\theta_d = [\rho_d^T \ \eta_d^T]^T$. De (97) e (73) tem-se que ambas funções custo $J^{\text{MR}}(\rho_d)$ e $J^{\text{OCI}}(\theta_d)$ são zero no vetor de parâmetros ideal. Contudo, quando as considerações acima não são respeitadas os mínimos das duas funções custo serão distintos. Conforme apresentado por Campestrini et al. (2017), no caso em que a Suposição 2.1 é

violada, a estimativa com o método OCI apresentará polarização e variância, e a polarização será significativa. Neste caso, aumentando o número de parâmetros a serem estimados, aumentará a complexidade e a flexibilidade do controlador, fazendo com que o controlador estimado consiga representar melhor o controlador ideal, reduzindo assim a polarização da estimativa. A flexibilidade do controlador pode ser aumentada ou reduzida através da escolha de $C_F(q)$, pois quanto mais informações forem adicionadas em $C_F(q)$ menos parâmetros serão estimados. Contudo, o aumento no número de parâmetros a serem estimados aumenta a variância da estimativa, pois a identificação de mais parâmetros resulta em uma variância maior da função de transferência estimada (CAMPESTRINI et al., 2017).

É importante mencionar, que no presente trabalho, não será abordada a análise do método OCI para o caso em que a Suposição 2.1 não é respeitada. Esse caso foi abordado por Varriale da Silva (2020) onde foi proposto um filtro para lidar com o caso de controladores subparametrizados. Além disso, o presente trabalho também não aborda a influência da escolha de $C_F(q)$, nem os casos em que o processo $G(q)$ possui polos instáveis ou zeros de fase não mínima. Diretrizes de como lidar com estes casos no método OCI são discutidas por Campestrini et al. (2017) e Campestrini et al. (2012).

2.4.2.2 Exemplo ilustrativo

A configuração do experimento deste exemplo é a mesma utilizada no primeiro exemplo ilustrativo do método VRFT (Subsubseção 2.4.1.3). Por conta disso, o modelo do processo, o modelo de referência, o controlador ideal e a estrutura de controle são descritos nas equações (84), (85), (86) e (87), respectivamente. Os mesmos dados de entrada e saída das 100 rodadas de Monte Carlo do exemplo do VRFT são utilizados para estimar os parâmetros do controlador com o método OCI. Uma vez que os dados são coletados em malha aberta, não é necessário estimar o modelo do ruído, por conta disso, a estimativa dos parâmetros foi obtida através da função `oe` da toolbox `ident` do Matlab para cada uma das 100 rodadas de Monte Carlo. Neste caso, o vetor de parâmetros ideal em (86) foi utilizado como ponto de inicialização do algoritmo de otimização.

Na Tabela 2 são apresentadas as estatísticas dos parâmetros estimados através do método OCI e o vetor de parâmetros ideal. Observa-se que a polarização da estimativa obtida com o método OCI é desprezível.

Tabela 2 — Média e desvio padrão dos parâmetros estimados com o OCI no exemplo da Seção 2.4.2.2.

ρ	ρ_d	OCI
K_p	0,3742	0,3741 (\pm 0,0027)
K_i	0,0593	0,0593 (\pm 0,0002)

2.5 Qualidade dos dados

2.5.1 Definições

Em métodos de controle baseados em dados, assim como em identificação de sistemas, a qualidade da estimativa dos parâmetros depende da qualidade dos dados coletados/disponíveis. No decorrer do texto, recorrentemente faremos uso dos termos dados

muito informativos e dados *pouco* informativos. Assim sendo, é importante definir o que são dados muito informativos no contexto do presente trabalho. Um conjunto de dados é classificado como muito informativo se o mesmo respeita as seguintes condições:

- (a) se os dados possuem uma boa relação sinal-ruído (SNR, do inglês *Signal-to-Noise Ratio*) nas frequências de interesse;
- (b) se a matriz de informação P formada a partir dos dados é bem condicionada.

Ou seja, um sinal muito informativo possui uma alta densidade de informação sobre o processo. Se o sinal de saída possui uma boa SNR, item (a), significa que a densidade do espectro de potência da saída, y , possui mais contribuição da entrada, u , do que do ruído, w . Considere o espectro de densidade de energia do sinal de saída, considerando dados coletados de um experimento em malha aberta, dado por (TANGIRALA, 2015)

$$\begin{aligned}\Phi_Y &= |Y(\omega)|^2 = |G(e^{j\omega})U(\omega) + W(\omega)|^2 \\ &= |G(e^{j\omega})|^2 \Phi_U + \Phi_W,\end{aligned}\quad (99)$$

onde Φ_U e Φ_W são os espectros dos sinais $u(t)$ e $w(t)$, respectivamente. Partindo de (99) conclui-se que uma boa SNR no sinal de saída implica que o espectro de energia do sinal de saída Φ_Y possui mais contribuição do termo da entrada $|G(e^{j\omega})|^2 \Phi_U$ do que do termo do ruído Φ_W . Quanto à matriz de informação, item (b), quando o experimento é longo, ou seja, $t \rightarrow \infty$, um bom número de condicionamento se traduz na persistência de excitação (Definição 2.5). Portanto, se um conjunto de dados não respeita estas condições o mesmo é classificado como pouco informativo. Na sequência, são apresentados alguns exemplos de dados utilizados no presente trabalho e suas classificações.

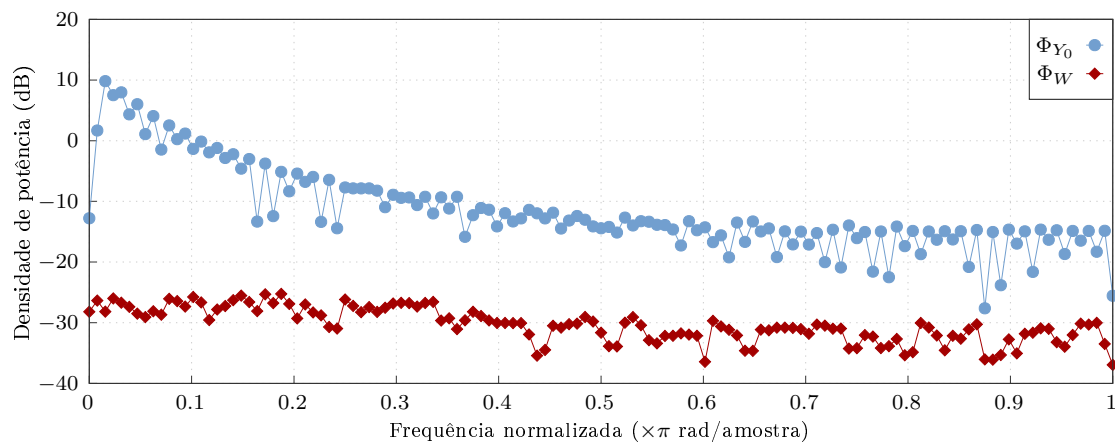
Nos exemplos desta seção, o processo é caracterizado pela seguinte função de transferência (BAZANELLA; CAMPESTRINI; ECKHARD, 2011)

$$G(q) = \frac{0,5}{q - 0,9}.\quad (100)$$

Considere que os dados são coletados de um experimento em malha aberta no qual um sinal do tipo PRBS (do inglês *Pseudo Random Binary Sequence*) com 9 períodos de 127 amostras com amplitude variando entre -1 e 1 é utilizado como entrada do processo. O tamanho total desse sinal é de 1143 amostras. À saída do processo é somado um ruído branco com variância $\sigma_v^2 = 0,0025$ filtrado por (88). A Figura 5 apresenta os espectros de densidade de energia do sinal proveniente somente do sinal de saída sem o ruído Φ_{Y_0} (pontos em azul) e do ruído $w(t)$, Φ_W (quadrados em vermelho). Observe que a contribuição proveniente da entrada é maior do que a contribuição do ruído em todas as frequências. O número de condicionamento para este exemplo é 4,71, calculado a partir da matriz de informação em (80) com os dados coletados. Com base no gráfico do espectro de potência e no valor do número de condicionamento, é possível classificar este dado como muito informativo, pois o mesmo respeita as condições apresentadas nos itens (a) e (b).

Considere agora que os dados são coletados de um experimento em malha aberta, no qual um degrau unitário com 1152 amostras excita a entrada do processo. Assim como no exemplo anterior, a saída do processo é corrompida por um ruído aditivo, gerado a partir de um ruído branco filtrado por (88). Neste caso, a variância do ruído branco é $\sigma_v^2 = 0,0625$. O espectro de densidade de energia do sinal de saída deste exemplo é apresentado na Figura 6. Observe que a maior contribuição da energia da entrada (em relação à energia do

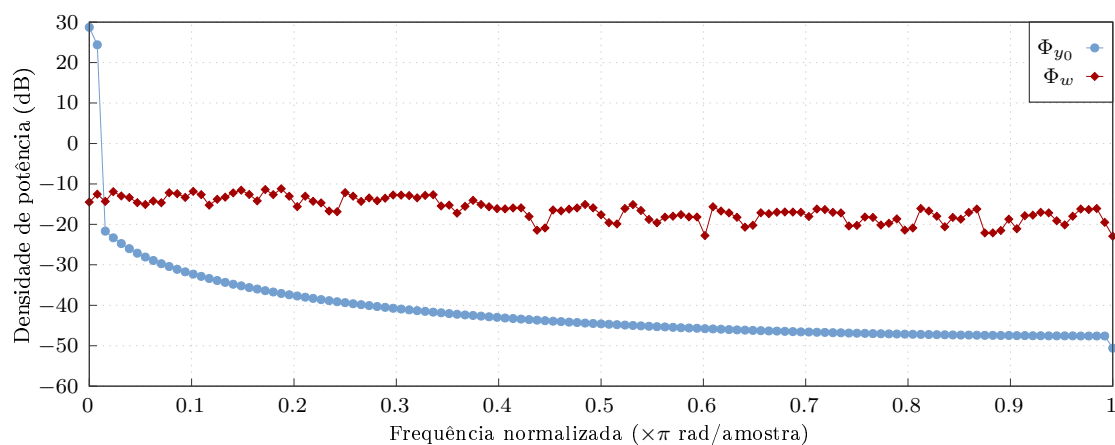
Figura 5 — Espectro de densidade de energia do sinal de saída para uma entrada do tipo PRBS.



Espectro do sinal proveniente da entrada $G(q)u(t)$ (círculos azuis) e espectro do ruído $w(t)$ (quadrados vermelhos). Fonte: elaborado pela autora.

ruído) está concentrada em uma estreita faixa de frequências, assim, se essa faixa coincidir com a faixa de frequências de interesse, a condição (a) será respeitada. O número de condicionamento neste caso é 206,89, calculado a partir da matriz de informação em (80) obtida com os dados coletados. Como a matriz de informação é mal condicionada, temos que a condição (b) não é respeitada. Por conta disso, os dados coletados de um experimento onde o sinal de entrada é um degrau longo são considerados pouco informativos.

Figura 6 — Espectro de densidade de energia do sinal de saída para um degrau unitário com 1152 amostras como sinal de entrada.



Espectro do sinal proveniente da entrada $G(q)u(t)$ (círculos azuis) e espectro do ruído $w(t)$ (quadrados vermelhos). Fonte: elaborado pela autora.

Devido à pouca informação relevante contida nesse dado, sua simples aplicação em um método baseado em dados pode levar a um ajuste inadequado dos parâmetros do controlador. Para mostrar o efeito desse dado na estimativa considere o processo e os dados de entrada e saída do exemplo anterior. O modelo de referência escolhido é dado por (BAZANELLA; CAMPESTRINI; ECKHARD, 2011)

$$T_d(q) = \frac{0,4}{q - 0,6}. \quad (101)$$

O controlador ideal, calculado a partir de (100) e (101), é

$$C_d(q) = [0,8 \quad -0,72] \begin{bmatrix} q & 1 \\ q-1 & q-1 \end{bmatrix}^T = \frac{0,8(q-0,9)}{(q-1)}. \quad (102)$$

O controlador a ser estimado é do tipo PI com a seguinte estrutura:

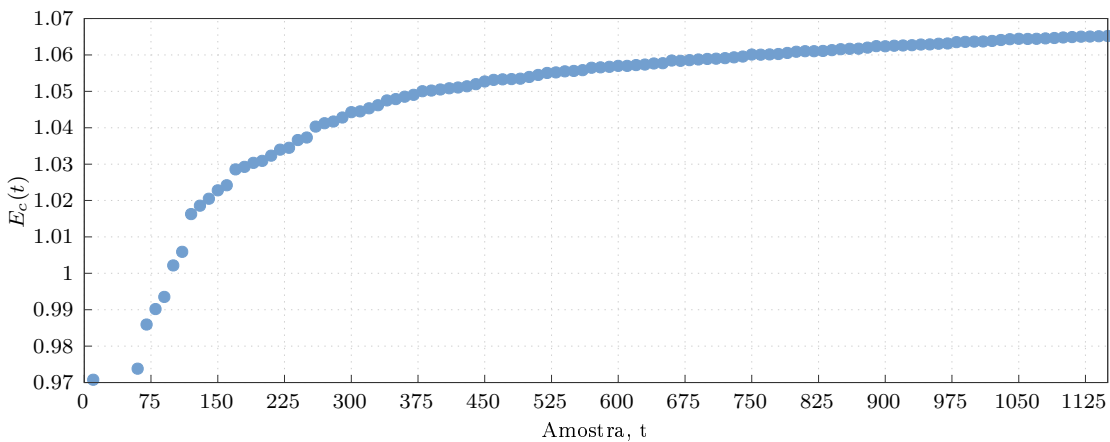
$$C(q, \rho) = [\rho_1 \quad \rho_2] \begin{bmatrix} q & 1 \\ q-1 & q-1 \end{bmatrix}^T. \quad (103)$$

Com estas definições e com os dados de entrada e saída do exemplo anterior estima-se os parâmetros do controlador como em (79), com $L(q) = 1$. Além disso, assuma que a estimativa dos parâmetros é obtida incrementando gradualmente a quantidade de amostras empregadas, ou seja, $t = 10, 20, \dots, 1150$. Para avaliar a qualidade das estimativas obtidas utiliza-se um erro de medida $E_c(t)$ (CAMPESTRINI et al., 2017) definido como

$$E_c(t) = \|(q-1)[C_d(q) - C(q, \hat{\rho}(t))]\|_2, \quad (104)$$

onde o polo instável do controlador é removido para obter um valor finito e t corresponde ao índice da amostra. A Figura 7 apresenta os valores obtidos de $E_c(t)$. Observe que aumentando-se o número de amostras aumenta-se a distância entre o controlador estimado e o controlador ideal.

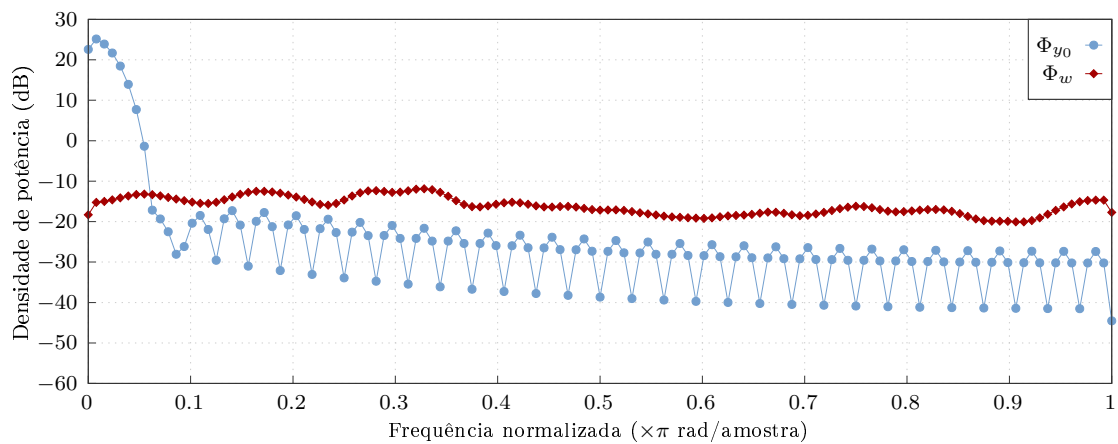
Figura 7 — E_c calculado com os parâmetros estimados.



Fonte: elaborado pela autora.

Contudo, embora dados pouco informativos não apresentem as características ideais que os dados muito informativos possuem, não é incomum que eles possuam segmentos contendo informações relevantes para a estimativa dos parâmetros, como será apresentado no Capítulo 4. Por exemplo, considere as primeiras 288 amostras dos dados do exemplo anterior. O espectro de densidade de energia do sinal de saída também foi calculado e é apresentado na Figura 8. Similarmente ao gráfico do exemplo anterior, a contribuição do dado proveniente da entrada é maior do que do ruído em uma faixa de frequências. Porém, essa faixa representa uma fração maior do espectro. O número de condicionamento para este caso é 174,20. Portanto, os dados coletados de um experimento onde um degrau curto é aplicado como sinal de entrada são mais informativos do que com um degrau longo. Com base nisso, reduzir o número de amostras pode melhorar a qualidade dos dados empregados.

Figura 8 — Espectro de densidade de energia do sinal de saída para um degrau com 288 amostras como sinal de entrada.



Espectro do sinal proveniente da entrada $G(q)u(t)$ (círculos azuis) e espectro do ruído $w(t)$ (quadrados vermelhos). Fonte: elaborado pela autora.

Na literatura de identificação de sistemas foram propostos critérios que visam selecionar os subconjuntos mais informativos de dados pouco informativos. Dois desses critérios foram aplicados ao método VRFT e são apresentados no Capítulo 4. Um dos critérios baseia-se diretamente no número de condicionamento da matriz de informação, enquanto o outro é baseado no menor valor singular da matriz regressora, podendo ser visto como uma medida indireta do número de condicionamento.

2.5.2 Exemplo motivacional

Neste exemplo, considere a configuração da simulação na Seção 2.4.1.3. Considere que estão disponíveis dados de entrada e saída de um experimento em malha aberta onde um sinal pouco informativo é utilizado como entrada do experimento. Neste caso, o sinal de entrada $u(t)$ é composto por um período de 1024 amostras de uma onda quadrada com níveis 0 e 1. 100 rodadas de Monte Carlo foram executadas, em cada rodada uma realização de ruído afeta a saída do processo. O sinal de ruído é gerado a partir de um ruído branco com variância $\sigma_v^2 = 0,01$ filtrado através de (88). O sinal de entrada e uma realização do sinal de saída são apresentados na Figura 9.

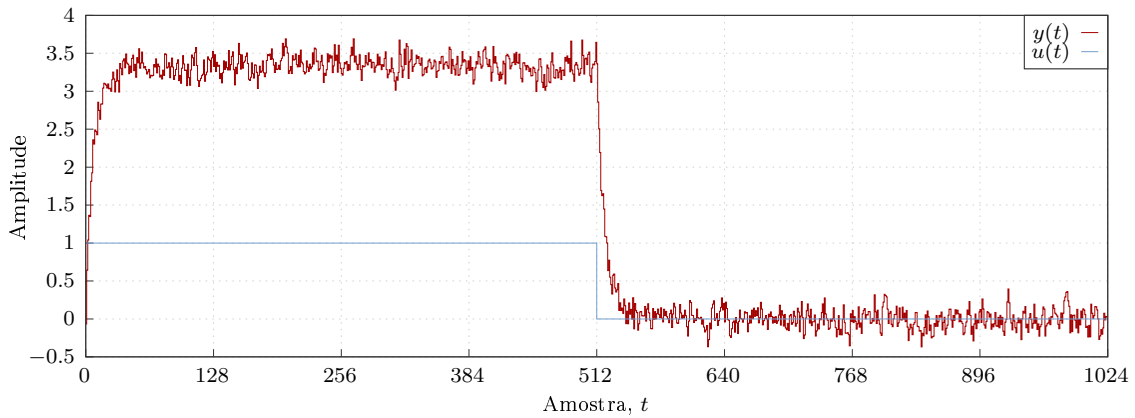
Na Tabela 3 são apresentados o vetor de parâmetros ideal ρ_d e as estatísticas dos

Tabela 3 — Média e desvio padrão dos parâmetros estimados com o VRFT e OLS com dados do experimento pouco informativo no exemplo da Seção 2.5.2.

ρ	ρ_d	OLS
K_p	0,3742	0,0088 ($\pm 0,0028$)
K_i	0,0593	0,0592 ($\pm 0,0001$)

parâmetros obtidas com a solução OLS e com os dados pouco informativos. A estimativa dos parâmetros apresenta uma polarização bastante acentuada no parâmetro proporcional. Contudo, esse efeito não ocorre no parâmetro integral, uma vez que a maior contribuição do sinal de entrada ocorre nas baixas frequências, que são as frequências de interesse

Figura 9 — Entrada $u(t)$ e uma realização da saída $y(t)$ coletada de um experimento pouco informativo no exemplo da Seção 2.5.2.



Fonte: elaborado pela autora.

para estimativa do ganho integral. Estes resultados concordam com a afirmação de que a simples aplicação de dados pouco informativos em um método baseado em dados pode levar a um ajuste inapropriado dos parâmetros do controlador.

2.6 O problema de mínimos quadrados

Essa seção tem como objetivo apresentar brevemente o problema de mínimos quadrados considerando a notação empregada no método VRFT. Além disso, são apresentadas as classes de problemas de mínimos quadrados, que são definidas de acordo com a influência do ruído no sinais empregados.

2.6.1 Formulação do problema

Os problemas de mínimos quadrados consistem em encontrar a solução de um conjunto sobredeterminado de equações (GOLUB; LOAN, 1980; HUFFEL; LEMMERLING, 2002). Primeiramente, considere que, para o caso sem ruído, a relação de entrada e saída do controlador ideal é definida como

$$\rho_d^T \varphi_0(t) = u_0(t), \quad (105)$$

onde o subscrito $_0$ indica que a quantidade é gerada sem ruído. Esta equação pode ser reescrita em forma vetorial como

$$\Phi_0 \rho_d = u_0, \quad (106)$$

onde o vetor de saída do controlador $u_0 \in \mathbb{R}^N$ e a matriz regressora $\Phi_0 \in \mathbb{R}^{N \times m}$ são definidos como

$$u_0 = [u_0(1) \ u_0(2) \ \dots \ u_0(N)]^T, \quad (107)$$

$$\Phi_0 = [\varphi_0(1) \ \varphi_0(2) \ \dots \ \varphi_0(N)]^T, \quad (108)$$

com N sendo o número de amostras coletadas no experimento.

Essa situação é dificilmente encontrada, pois em casos realistas os dados são sempre afetados por ruído, muitas vezes não negligenciável. Para lidar com o caso ruidoso, assume-se que a matriz regressora calculada Φ e o vetor de saída medido u são descritos como a

soma de dois termos: um termo que não é afetado por ruído e um termo que representa puramente o ruído. Assim sendo, Φ e u são representados como

$$u = u_0 + \delta_u, \quad (109)$$

$$\Phi = \Phi_0 + \Delta_\Phi, \quad (110)$$

onde $\delta_u \in \mathbb{R}^N$ representa as contribuições de ruído em u e $\Delta_\Phi \in \mathbb{R}^{N \times m}$ representa as contribuições de ruído em Φ . Além disso, u e Φ são definidos como

$$u = [u(1) \ u(2) \ \dots \ u(N)]^T, \quad (111)$$

$$\Phi = [\varphi(1) \ \varphi(2) \ \dots \ \varphi(N)]^T. \quad (112)$$

A partir de (109), (110) e (106) encontra-se a formulação geral do problema de estimar ρ . Assim, isola-se u_0 e Φ_0 em (109) e (110), que são substituídos em (106), chegando-se a

$$(\Phi - \Delta_\Phi)\rho = u - \delta_u. \quad (113)$$

As soluções de mínimos quadrados são classificadas na literatura dependendo de como o ruído afeta (113), ou seja, se Δ_Φ , δ_u ou ambos estão presentes (DEGROAT; DOWLING, 1993). Na sequência, serão descritas essas soluções.

2.6.2 Soluções de mínimos quadrados

Talvez o problema mais comumente empregado na literatura seja o OLS, no qual assume-se que a contribuição de ruído aparece no vetor de saída u , enquanto a matriz regressora Φ não é afetada por ruído. O problema OLS é formulado como apresentado abaixo, considerando-se $\Delta_\Phi = 0$ em (113),

$$\min_{\delta_u, \rho} \|\delta_u\|_2^2 \quad (114)$$

$$\text{sujeito a } \Phi\rho = u - \delta_u. \quad (115)$$

O objetivo é encontrar o vetor de parâmetros ρ enquanto compensa-se a influência do ruído em u . A solução para esse problema é bastante conhecida e dada por

$$\hat{\rho} = (\Phi^T\Phi)^{-1} \Phi^T u. \quad (116)$$

A estimativa com essa solução não é polarizada, contanto que apenas u seja afetado por ruído (ABATZOGLOU; MENDEL; HARADA, 1991). Observe que (116) é a forma vetorial de (79). No caso do VRFT, Φ é sempre afetado por ruído, logo, a estimativa OLS é polarizada.

Quando tanto Φ quanto u são afetados por ruído tem-se o problema conhecido como *Total Least Squares* (TLS). Na literatura este problema também é conhecido como: problema de erros nas variáveis, regressão ortogonal e erros de medida (MARKOVSKY; HUFFEL, 2007). O problema TLS é formulado da seguinte forma

$$\min_{\Delta_\Phi, \delta_u, \rho} \|[\Delta_\Phi \ \delta_u]\|_F^2 \quad (117)$$

$$\text{sujeito a } (\Phi - \Delta_\Phi)\rho = u - \delta_u, \quad (118)$$

onde $\|\cdot\|_F$ é a norma de Frobenius, ou seja, $\|X\|_F^2 = \sum_{i,j} |x_{i,j}|^2$. Assumindo-se que as fontes de ruído que afetam as colunas da matriz $[\Delta_\Phi \ \delta_u]$ são descorrelacionadas, e que a variância do ruído é igual em cada coluna, a solução para o problema TLS é dada por

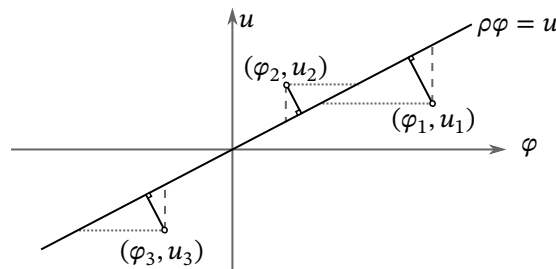
$$\hat{\rho} = (\Phi^T \Phi - \underline{\sigma}^2 I_m)^{-1} \Phi^T u, \quad (119)$$

onde $\underline{\sigma}^2$ é o quadrado do menor valor singular de $[\Phi \ u]^T [\Phi \ u]$. Como será visto no Capítulo 3 esta solução não se encaixa bem no problema do VRFT, uma vez que as realizações de ruído que afetam as colunas da matriz $[\Delta_\Phi \ \delta_u]$ não são independentes.

Ao longo dos anos diversos casos particulares (extensões) do problema TLS foram propostos na literatura. No presente trabalho, os casos de interesse são conhecidos como *Constrained Total Least Squares* (CTLS) (ABATZOGLOU; MENDEL, 1987) e *Data Least Squares* (DLS) (DEGROAT; DOWLING, 1993). No CTLS assume-se que o ruído presente em Φ e u é algebricamente correlacionado, sendo caracterizado pela adição de mais uma restrição no problema original do TLS descrito em (117) e (118). Já o problema DLS considera que apenas a matriz regressora é afetada por ruído. A aplicação destas duas soluções no método VRFT será apresentada em detalhes no Capítulo 3. Outras extensões do método TLS podem ser encontradas em outros trabalhos (HUFFEL; LEMMERLING, 2002; HUFFEL; VANDEWALLE, 1989).

A Figura 10 apresenta uma interpretação geométrica dos problemas de mínimos quadrados OLS, TLS e DLS, para a estimativa de um único parâmetro. Observe que no problema OLS são minimizadas as distâncias verticais à reta (linhas tracejadas), no problema DLS as distâncias horizontais à reta (linhas pontilhadas) são minimizadas e com TLS são minimizadas as distâncias perpendiculares à reta (linhas contínuas).

Figura 10 — Interpretação geométrica das soluções de mínimos quadrados.



OLS (linha vertical tracejada), TLS (linha perpendicular contínua) e DLS (linha horizontal pontilhada) para um parâmetro. Fonte: adaptado de Golub e Loan (1980).

2.7 Nota

A validação da estabilidade em malha fechada do controlador estimado não é abordada no presente trabalho pois não faz parte do escopo do mesmo. Contudo, cita-se em seguida alguns trabalhos que tratam este problema e podem servir como material de consulta ao leitor interessado. Van Heusden, Karimi e Söderström (2011) propuseram a utilização do teorema do pequeno ganho para determinar condições suficientes para garantir estabilidade em malha fechada, que são adicionadas como restrições no problema de otimização do método VRFT. Gonçalves da Silva, Bazanella e Campestrini (2020) propuseram testes de desempenho e de certificação de controladores baseados na norma H_∞ de uma matriz de transferência específica. Pinto da Silva e Eckhard (2020a) sugeriram um algoritmo para

certificação de controladores, considerando que o modelo do processo é desconhecido, através de comparação entre a distância de Vinnicombe e a margem de estabilidade generalizada do sistema.

2.8 Considerações finais

Em métodos de controle baseados em dados não iterativos a estimativa do controlador é obtida diretamente a partir dos dados de um único experimento, o que é bastante atrativo do ponto de vista prático. Outro atrativo é a possibilidade de utilização de dados pouco informativos, como dados de operação, por exemplo. Contudo, a simples aplicação de tais dados em um método baseado em dados pode levar a um ajuste inapropriado dos parâmetros.

Provavelmente o método baseado em dados não iterativo mais utilizado na literatura seja o método VRFT. Embora este método tenha sido empregado com sucesso em diversos trabalhos, ainda existem alguns problemas que requerem atenção. Abaixo cita-se os problemas encontrados:

- (a) a alternativa proposta na literatura para lidar com o ruído, o método de variáveis instrumentais, reduz a polarização da estimativa com o inconveniente de aumentar sua variância;
- (b) no caso em que dados pouco informativos estão disponíveis, a simples aplicação destes dados no método VRFT leva a um ajuste inapropriado dos parâmetros.

Dessa maneira, verifica-se a possibilidade de contribuições para o método VRFT para que o mesmo possa ser empregado em situações mais realistas. Por exemplo, no caso ruidoso, uma alternativa para reduzir a polarização da estimativa com VI é proposta por Rallo et al. (2016) para o caso monovariável e estendida por Boeira e Eckhard (2018) para o caso multivariável. Nesta abordagem, um termo de regularização é adicionado à função custo minimizada, resultando em uma redução na variância da estimativa com o custo de aumentar um pouco sua polarização. Contudo, esta técnica possui o inconveniente de ser necessário estimar a matriz de covariância dos parâmetros. Van Heusden, Karimi e Söderström (2011) também propuseram uma alternativa à utilização da VI comumente empregada no método VRFT. Naquele trabalho, o problema de identificação dos parâmetros do controlador é formulado como um problema de erros nas variáveis e a estimativa dos parâmetros é obtida através do método das variáveis instrumentais estendidas. Contudo, a solução foi desenvolvida considerando que o sinal de entrada é periódico. A alternativa proposta no presente trabalho consiste no emprego de outras soluções de mínimos quadrados para estimar os parâmetros do controlador com o método VRFT. Tais soluções e as alterações necessárias para aplicá-las ao método VRFT são apresentadas no Capítulo 3. Vale salientar que a abordagem proposta não requer a estimativa da matriz de covariância dos parâmetros ou a coleta de dados de um segundo experimento. Além disso, ela não requer a utilização de dados periódicos, independente se dados de malha aberta ou fechada são empregados.

A utilização de dados pouco informativos na estimativa dos parâmetros do controlador é outro quesito que requer atenção. Essa proposta permite, por exemplo, utilizar dados armazenados durante a operação normal do processo, ou seja, possibilita que o projetista não precise coletar dados de um experimento específico (que pode ser uma tarefa custosa). Apesar desta vantagem, como já mencionado, uma estimativa inadequada dos parâmetros é

obtida quando estes dados são simplesmente empregados em um método baseado em dados. Porém, isso não significa que dados pouco informativos não possuam nenhuma informação relevante para o projeto do controlador. De fato, não é incomum que hajam intervalos proveitosos em dados pouco informativos que, se detectados corretamente, resultam em um ajuste satisfatório dos parâmetros. Esta linha de pesquisa é um assunto já abordado na área de identificação de sistemas mas é um assunto muito pouco abordado na área de controle baseado em dados, como pode ser visto através das poucas contribuições na literatura. Por exemplo, Radrizzani et al. (2020) selecionam manualmente os subconjuntos de dados mais informativos que pertencem ao regime transitório do processo e utilizam esses dados na estimativa dos parâmetros com o VRFT. Outra solução para este problema é apresentada no Capítulo 4, onde propõe-se a utilização de dois critérios de seleção de subconjuntos muito informativos de dados aplicados ao método VRFT.

3 AS SOLUÇÕES DE MÍNIMOS QUADRADOS CTLS E DLS APLICADAS AO MÉTODO VRFT

3.1 Introdução

O método VRFT é um método baseado em dados não iterativo bastante conhecido para projeto dos parâmetros do controlador. Esse método baseia-se no paradigma do modelo de referência, no qual define-se previamente o desempenho em malha fechada almejado e a estrutura de controle. Essas informações são utilizadas, juntamente com os dados coletados, para gerar os sinais de entrada e saída (virtuais) empregados na identificação do melhor controlador com a estrutura de controle disponível (CAMPI; LECCHINI; SAVARESI, 2002; BAZANELLA; CAMPESTRINI; ECKHARD, 2011).

Embora a aplicação do método VRFT não seja restrita ao caso de controladores linearmente parametrizados, essa suposição é favorecida na formulação usual do método e na maioria das aplicações descritas na literatura. De fato, a estrutura de controle comumente escolhida possui um denominador fixo e um numerador parametrizado. No caso de controladores linearmente parametrizados, a solução é facilmente obtida através da solução de um problema OLS. Contudo, essa solução produz uma estimativa polarizada no caso em que os dados são corrompidos por ruído. A polarização é intrínseca à formulação do problema e é esperada mesmo no caso em que o controlador ideal pertence à classe de controle, uma vez que ela se deve ao fato de que o problema posto pelo VRFT não é um problema de identificação padrão (SÖDERSTRÖM, 2018). Considerando dados de malha aberta, apenas a entrada do controlador a ser identificado é afetada por ruído. Em contrapartida, quando considera-se dados de malha fechada o ruído afeta tanto a entrada quanto a saída do controlador a ser identificado. Usualmente, o efeito do ruído é atenuado através da utilização de um método de variáveis instrumentais (CAMPI; LECCHINI; SAVARESI, 2002). Contudo, essa abordagem possui a desvantagem de aumentar a variância da estimativa, o que pode levar à perda de desempenho e à instabilidade em malha fechada (SÖDERSTRÖM, 2018).

Com isso em mente, outras abordagens foram propostas na literatura para lidar com dados ruidosos no método VRFT. Por exemplo, Van Heusden, Karimi e Söderström (2011) apresentaram uma discussão sobre o problema e propuseram estimar os parâmetros do controlador com o VRFT usando um método de variáveis instrumentais estendidas. Contudo, este método foi desenvolvido para dados periódicos. Rallo et al. (2016) propuseram, para o caso monovariável, a adição de um termo de regularização à função custo minimizada, o que resulta na redução da variância da estimativa com a desvantagem de aumentar um pouco a polarização da mesma. Porém, essa abordagem possui o inconveniente de requerer uma estimativa para a matriz de covariância dos parâmetros. Uma extensão para o caso

multivariável desta abordagem é proposta por Boeira e Eckhard (2018). A fim de superar essa dificuldade no caso ruidoso, Silva Garcia e Bazanella (2020a) propuseram a aplicação de outra solução de mínimos quadrados ao método VRFT, conhecida como CTLS. O problema formulado pelo CTLS é uma extensão do problema TLS com a restrição de que a mesma realização de ruído afeta tanto a matriz regressora quanto o vetor de saída. As peculiaridades do problema CTLS fazem com que ele represente melhor o problema posto pelo VRFT. A solução CTLS produz uma estimativa com reduzida polarização e menor variância, quando comparada com a solução com VI. Contudo, Silva Garcia e Bazanella (2020a) consideraram apenas controladores com estrutura linearmente parametrizada e polos fixos.

No presente capítulo, apresenta-se a formulação do problema do método VRFT com duas soluções de mínimos quadrados. Primeiramente, apresenta-se a solução CTLS considerando-se controladores com parametrização linear com polos fixos. Em seguida, desenvolve-se a solução CTLS assumindo-se que na estrutura do controlador tanto o numerador quanto o denominador podem ser estimados. Para ambos os casos, são desenvolvidas as modificações necessárias considerando-se dados de malha aberta e fechada e que o controlador ideal pertence à classe de controle. Além disso, apresenta-se o desenvolvimento dos filtros para os casos descritos acima assumindo-se que o controlador ideal não pertence à classe de controle. Em seguida, apresenta-se o desenvolvimento da solução DLS para o método VRFT. Nesse caso, considera-se que a Suposição 2.1 é respeitada e que o controlador a ser estimado é linearmente parametrizado com polos fixos. Contudo, sua aplicação não é restrita ao caso de controladores de ordem cheia, como será apresentado no Capítulo 5 onde essa restrição é relaxada. Exemplos de simulação ilustram a aplicação das soluções empregadas.

3.2 *Constrained Total Least Squares*

O problema CTLS é uma técnica para resolver um conjunto sobredeterminado de equações (113) que tenta remover a contribuição de ruído da matriz regressora Φ e do vetor de saída u , enquanto simultaneamente explora a estrutura algébrica do ruído que afeta Φ e u (ABATZOGLOU; MENDEL; HARADA, 1991). Este problema foi proposto por Abatzoglou e Mendel (1987) como um caso particular do problema TLS (GOLUB; LOAN, 1980). Exemplos de aplicação da solução CTLS em processos reais podem ser encontrados na literatura. Por exemplo, Kim et al. (2007) aplicaram a solução CTLS em processos químicos. Tsai e Kao (2006) propuseram a aplicação da solução CTLS para melhorar a precisão da estimativa de um sistema de GPS (do inglês *Global Positioning System*). Além dos trabalhos mencionados, a solução CTLS foi aplicada em outros casos (ABATZOGLOU; SOON, 1988; ABATZOGLOU; HARADA; SHINE, 1988; ABATZOGLOU; MENDEL; HARADA, 1991).

Para formular o problema posto pelo CTLS, primeiramente, considere o problema de mínimos quadrados para o caso sem ruído descrito em (105) e o problema de mínimos quadrados generalizado (113), reescrito aqui por conveniência

$$(\Phi - \Delta_\Phi)\rho = u - \delta_u. \quad (113)$$

Suponha que a matriz Δ_Φ e o vetor δ_u são afetados pela mesma fonte de ruído, ou seja,

$$\Delta_\Phi = [P_1 v \quad \dots \quad P_m v], \quad (120)$$

$$\delta_u = P_{m+1} v, \quad (121)$$

onde $v \in \mathbb{R}^N$ é um vetor com as amostras de ruído que afetam Φ e u , ou seja, $v = [v(1) \dots v(N)]^T$ e $P_i \in \mathbb{R}^{N \times N}$ com $i = 1, \dots, m+1$ são matrizes Toeplitz cujas colunas possuem as respostas impulsivas dos filtros $P_i(q)$. Essas matrizes possuem a seguinte estrutura:

$$P_i = \begin{bmatrix} p_i(1) & 0 & \dots & 0 \\ p_i(2) & p_i(1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ p_i(N) & p_i(N-1) & \dots & p_i(1) \end{bmatrix} \quad (122)$$

O problema CTLS é formulado a partir do problema TLS (descrito em (117) e (118)) no qual adiciona-se a suposição de que a matriz regressora e o vetor de saída são afetados pela mesma fonte de ruído, descritas em (120) e (121), ou seja,

$$\min_{v, \rho} \|\Delta_\Phi \delta_u\|_F^2 \quad (123)$$

$$\text{sujeito a } (\Phi - \Delta_\Phi)\rho = u - \delta_u, \quad (124)$$

$$[\Delta_\Phi \delta_u] = [P_1 v \dots P_m v P_{m+1} v]. \quad (125)$$

onde $\|\Delta_\Phi \delta_u\|_F^2 = J^{\text{CTLS}}(v)$.

Considere que o controlador ideal pertence à classe de controle (Suposição 2.1) e que os sinais de entrada e saída não são afetados por ruído. Neste caso, tem-se que

$$\min_{v, \rho} J^{\text{CTLS}}(v) = 0$$

e ρ_d é o único vetor de parâmetros que satisfaz a equação (124), ou seja, $\Phi_0 \rho_d = u_0$, (106), desde que Φ_0 tenha posto cheio em colunas. Assim sendo, observa-se que o vetor de parâmetros ideal ρ_d é o único vetor com o qual obtém-se o mínimo da função custo $J^{\text{MR}}(\rho)$ e o mínimo do problema de otimização posto pelo CTLS. Para o caso em que a Suposição 2.1 não é respeitada, desenvolveu-se um filtro para aproximar o mínimo da função custo $J^{\text{MR}}(\rho)$ do mínimo de $J^{\text{CTLS}}(v)$ (considerando-se as restrições do problema), cujo desenvolvimento será apresentado mais adiante. A solução do problema CTLS é apresentada no Teorema 3.1.

Teorema 3.1 (de Abatzoglou e Mendel (1987)). *Seja $Z = [\Phi \ u]$ e $\Delta_Z = [\Delta_\Phi \ \delta_u]$. Então o vetor de parâmetros estimado $\hat{\rho}$ é a solução do seguinte problema de otimização*

$$\min_{\rho} \begin{bmatrix} \rho \\ -1 \end{bmatrix}^T Z^T (\Gamma_\rho K^{-1} \Gamma_\rho^T)^{-1} Z \begin{bmatrix} \rho \\ -1 \end{bmatrix}, \quad (126)$$

onde $\Gamma_\rho = \sum_{i=1}^m P_i \rho_i - P_{m+1}$, ρ_i é o i -ésimo elemento de ρ e $K = \sum_{i=1}^{m+1} P_i^T P_i$.

A prova deste teorema é apresentada abaixo e segue o mesmo desenvolvimento descrito por Abatzoglou e Mendel (1987).

Prova. Considere o problema de mínimos quadrados generalizado (113) reescrito como

$$\begin{aligned} [\Phi \ u] \begin{bmatrix} \rho \\ -1 \end{bmatrix} - [\Delta_\Phi \ \delta_u] \begin{bmatrix} \rho \\ -1 \end{bmatrix} &= 0, \\ (Z - \Delta_Z)\beta &= 0, \end{aligned} \quad (127)$$

onde $Z \in \mathbb{R}^{N \times (m+1)}$, $\Delta_Z \in \mathbb{R}^{N \times (m+1)}$ e $\beta \in \mathbb{R}^{m+1}$ são definidos como

$$Z = [\Phi \quad u], \quad (128)$$

$$\Delta_Z = [\Delta_\Phi \quad \delta_u], \quad (129)$$

$$\beta = [\rho^\top \quad -1]^\top. \quad (130)$$

A partir das definições em (128), (129) e (130) reescreve-se o problema CTLS como segue

$$\min_{v, \rho} \|\Delta_Z\|_F^2 \quad (131)$$

$$\text{sujeito a } (Z - \Delta_Z)\beta = 0, \quad (132)$$

$$\Delta_Z = [P_1 v \quad \dots \quad P_m v \quad P_{m+1} v]. \quad (133)$$

onde $\|\Delta_Z\|_F^2 = J^{\text{CTLS}}(v)$. Observe que o termo $\|\Delta_Z\|_F^2$ pode ser reescrito como

$$\|\Delta_Z\|_F^2 = \text{tr}\{\Delta_Z^\top \Delta_Z\} \quad (134)$$

$$= \sum_{i=1}^{m+1} v^\top P_i^\top P_i v$$

$$= v^\top \left(\sum_{i=1}^{m+1} P_i^\top P_i \right) v$$

$$= v^\top K v. \quad (135)$$

Agora, define-se o Lagrangeano para o problema:

$$\mathcal{L}(v, \rho, \lambda) = v^\top K v + \lambda^\top (Z - \Delta_Z)\beta + \beta^\top (Z - \Delta_Z)^\top \lambda,$$

reescrito substituindo-se (133) como

$$\mathcal{L}(v, \rho, \lambda) = v^\top K v + \lambda^\top (Z - [P_1 v \quad \dots \quad P_{m+1} v])\beta + \beta^\top (Z - [P_1 v \quad \dots \quad P_{m+1} v])^\top \lambda, \quad (136)$$

O termo $[P_1 v \quad \dots \quad P_{m+1} v]\beta$ pode ser expandido, lembrando de (130), da seguinte forma

$$\begin{aligned} [P_1 v \quad \dots \quad P_{m+1} v]\beta &= [P_1 v \quad \dots \quad P_{m+1} v][\rho^\top \quad -1]^\top \\ &= P_1 \rho_1 v + \dots + P_m \rho_m v - P_{m+1} \rho_{m+1} v \\ &= \left(\sum_{i=1}^m P_i \rho_i - P_{m+1} \right) v \\ &= \Gamma_\rho v. \end{aligned} \quad (137)$$

Substituindo-se (137) em (136) obtém-se

$$\mathcal{L}(v, \rho, \lambda) = v^\top v + \lambda^\top Z \beta - \lambda^\top \Gamma_\rho v + \beta^\top Z^\top \lambda - v^\top \Gamma_\rho^\top \lambda$$

Na sequência, calcula-se o gradiente de $\mathcal{L}(v, \rho, \lambda)$ com respeito a v :

$$\frac{\partial \mathcal{L}(v, \rho, \lambda)}{\partial v} = v^\top K - \lambda^\top \Gamma_\rho = 0.$$

Desta equação, tem-se que $Kv = \Gamma_\rho^\top \lambda$, uma vez que $K = K^\top$, e

$$v = K^{-1} \Gamma_\rho^\top \lambda. \quad (138)$$

Na sequência, deseja-se calcular o gradiente de $\mathcal{L}(v, \rho, \lambda)$ com respeito a ρ . Para tanto, considere que $Z = [Z_1 \dots Z_{m+1}]$ onde Z_i corresponde à i -ésima coluna da matriz Z . Agora, novamente lembrando de (130), tem-se

$$\begin{aligned} Z[\rho^T - 1]^T &= [Z_1 \dots Z_{m+1}][\rho^T - 1]^T \\ &= Z_1\rho_1 + \dots + Z_m\rho_m - Z_{m+1} \\ &= \sum_{i=1}^m Z_i\rho_i - Z_{m+1} \end{aligned} \quad (139)$$

O gradiente de $\mathcal{L}(v, \rho, \lambda)$ com respeito a ρ é calculado a partir da equação acima e da definição de Γ_ρ em (137), como segue

$$\begin{aligned} \frac{\partial \mathcal{L}(v, \rho, \lambda)}{\partial \rho} &= \frac{\partial}{\partial \rho} \left(\lambda^T \left[\sum_{i=1}^m Z_i\rho_i - Z_{m+1} \right] - \lambda^T \left(\sum_{i=1}^m P_i\rho_i - P_{m+1} \right) v \right) \\ &= \lambda^T ([Z_1 \dots Z_m] - [P_1 \dots P_m]v) = 0 \end{aligned} \quad (140)$$

Uma expressão para λ é obtida substituindo-se (133) em (127) e então (137) na expressão resultante, como abaixo

$$\begin{aligned} Z\beta - \Delta_Z\beta &= Z\beta - [P_1v \dots P_{m+1}v]\beta \\ &= Z\beta - \Gamma_\rho v, \end{aligned}$$

substituindo (138) e resolvendo para λ chega-se a

$$\lambda = \left(\Gamma_\rho K^{-1} \Gamma_\rho^T \right)^{-1} Z\beta. \quad (141)$$

Substituindo-se (141) em (140) obtém-se

$$\beta^T Z^T \left(\Gamma_\rho K^{-1} \Gamma_\rho^T \right)^{-1} ([Z_1 \dots Z_m] - [P_1 \dots P_m]v) = 0.$$

Observe que cada i -ésima equação é obtida como:

$$\begin{aligned} \beta^T Z^T \left(\Gamma_\rho K^{-1} \Gamma_\rho^T \right)^{-1} Z_i &= \beta^T Z^T \left(\Gamma_\rho K^{-1} \Gamma_\rho^T \right)^{-1} P_i v \quad \text{com } i = 1, \dots, m \\ &= \beta^T Z^T \left(\Gamma_\rho K^{-1} \Gamma_\rho^T \right)^{-1} P_i K^{-1} \Gamma_\rho^T \lambda \end{aligned} \quad (142)$$

$$= \beta^T Z^T \left(\Gamma_\rho K^{-1} \Gamma_\rho^T \right)^{-1} P_i K^{-1} \Gamma_\rho^T \left(\Gamma_\rho K^{-1} \Gamma_\rho^T \right)^{-1} Z\beta, \quad (143)$$

onde (142) e (143) são obtidas substituindo-se (138) e (141), respectivamente. O vetor de parâmetros estimado $\hat{\rho}$ é obtido resolvendo-se simultaneamente as m equações acima. Contudo, este é um problema de difícil solução devido as não linearidades das equações. Para simplificar o problema, substitui-se (141) em (138) (ABATZOGLOU; MENDEL, 1987), obtendo-se:

$$v = K^{-1} \Gamma_\rho^T \left(\Gamma_\rho K^{-1} \Gamma_\rho^T \right)^{-1} Z\beta.$$

Com base na equação anterior, minimizar $\|\Delta_Z\|_F^2$ em (134) corresponde a minimizar

$$\begin{aligned} v^T K v &= \beta^T Z^T \left(\Gamma_\rho K^{-1} \Gamma_\rho^T \right)^{-1} \Gamma_\rho K^{-1} K K^{-1} \Gamma_\rho^T \left(\Gamma_\rho K^{-1} \Gamma_\rho^T \right)^{-1} Z\beta \\ &= \begin{bmatrix} \rho \\ -1 \end{bmatrix}^T Z^T \left(\Gamma_\rho K^{-1} \Gamma_\rho^T \right)^{-1} Z \begin{bmatrix} \rho \\ -1 \end{bmatrix}, \end{aligned} \quad (144)$$

obtida lembrando de (130). Assim, a solução CTLS é estimada minimizando-se (144). ■

Para aplicar a solução CTLS ao método VRFT é preciso definir a resposta impulsiva dos filtros e gerar as matrizes P_i , com $i = 1, \dots, m + 1$, que serão utilizadas no problema de otimização. Esses filtros representam a forma como as contribuições de ruído afetam cada coluna da matriz Δ_Z . Por conta disso, a estrutura dos filtros é determinada por alguns fatores, como: o tipo de experimento, a estrutura escolhida para o controlador e se o controlador ideal pertence à classe de controle, por exemplo. As próximas subseções apresentam a estrutura dos filtros para cada caso considerado.

Nota 3.1. Observe que a estimativa dos parâmetros do controlador $\hat{\rho}$ com CTLS é obtida através da solução de um problema de otimização não convexo, como apresentado no Teorema 3.1. Contudo, uma solução fechada pode ser obtida quando a matriz Δ_Z possui colunas linearmente dependentes (ABATZOGLOU; MENDEL, 1987). Essa condição é alcançada, por exemplo, quando o vetor u não é corrompido por ruído, o que caracteriza o caso em que os dados são coletados em malha aberta no VRFT. Essa condição também será satisfeita caso uma ou mais colunas da matriz Φ sejam conhecidas exatamente, ou seja, não forem afetadas por ruído. Esse caso em particular não se aplica ao método VRFT, uma vez que a matriz regressora sempre é afetada por ruído. Essa solução será apresentada na Seção 3.3.

3.2.1 Controlador com polos fixos

Nesse caso, assume-se que o controlador é linearmente parametrizado com polos fixos, como descrito na Definição 2.4. Para dados de malha fechada $m + 1$ filtros são obtidos. Em contrapartida, obtém-se m filtros para o caso de dados de malha aberta. A seguir são apresentadas as estruturas dos filtros para cada um desses casos (SILVA GARCIA; BAZANELLA, 2020a).

Filtros para o caso de dados de malha fechada: Considere o conjunto de dados \mathcal{Z}^N coletado de um experimento em malha fechada como descrito na Definição 2.2. Nesse caso, os sinais de entrada $u(t)$ e saída $y(t)$ são afetados por ruído. Por conta disso, o ruído corrompe a matriz regressora Φ e o vetor de saída u . Consequentemente, o número de filtros deste caso é $m + 1$, pois todas as colunas de Δ_Z em (129) são diferentes de zero. Para definir a estrutura dos filtros, primeiramente, considere a matriz regressora Φ formada a partir dos vetores regressores, como em (112). Substituindo-se (4) e (2) em (65) (considerando-se $L(q) = 1$) obtém-se a seguinte expressão para o vetor regressor:

$$\begin{aligned} \varphi(t) &= \left[T_d^{-1}(q) - 1 \right] \bar{C}(q) [T(q, \rho_0)r(t) + S(q, \rho_0)H(q)v(t)] \\ &= \underbrace{\left[T_d^{-1}(q) - 1 \right] \bar{C}(q)T(q, \rho_0)r(t)}_{\varphi_0(t)} + \underbrace{\left[T_d^{-1}(q) - 1 \right] \bar{C}(q)S(q, \rho_0)H(q)v(t)}_{\Delta_\varphi(t)} \end{aligned} \quad (145)$$

onde $S(q, \rho_0)$ e $T(q, \rho_0)$ são a função de sensibilidade e a função de sensibilidade complementar obtidas como em (6) e (7) com o controlador $C(q, \rho_0)$, que é o controlador *inicial* operando na malha enquanto são coletados os dados do experimento. Observe que o vetor regressor $\varphi(t)$ foi dividido em duas componentes: uma puramente dependente do sinal de referência $\varphi_0(t)$ e outra somente com a contribuição de ruído $\Delta_\varphi(t)$.

A estrutura dos m primeiros filtros é derivada a partir da componente $\Delta_\varphi(t)$. Primeiramente, escreve-se (145) em forma vetorial como em (110), veja que Φ_0 é construído como em (108) com as amostras de $\varphi_0(t)$. Nosso interesse concentra-se em Δ_Φ , que é a forma matricial de $\Delta_\varphi(t)$. Cada uma das m colunas de Δ_Φ é construída a partir de uma matriz

Toeplitz P_i multiplicada pelo vetor v , com i sendo o índice da coluna. De (145) observa-se que essas matrizes são formadas a partir da resposta impulsiva dos filtros

$$P_i(q) = -\left[T_d^{-1}(q) - 1\right]\bar{C}_i(q)S(q, \rho_0)H(q) \quad \text{com } i = 1, \dots, m. \quad (146)$$

Uma vez que os m primeiros filtros foram definidos, resta definir o último filtro, que é obtido do vetor de saída. Assim sendo, a saída do controlador a ser identificado $u(t)$, é reescrita substituindo-se (5) e (2) em (64) (assumindo-se $L(q) = 1$), chegando em

$$u(t) = \underbrace{S(q, \rho_0)C(q, \rho_0)r(t)}_{u_0(t)} - \underbrace{C(q, \rho_0)S(q, \rho_0)H(q)v(t)}_{\delta_u(t)}, \quad (147)$$

onde $u(t)$ também foi dividido em duas partes: uma proveniente do sinal de referência $u_0(t)$ e outra apenas com a contribuição de ruído $\delta_u(t)$. Considere $u(t)$ escrito em forma vetorial como em (109), onde u_0 é obtido como em (107) com as amostras de $u_0(t)$ e δ_u é formado com as amostras de $\delta_u(t)$. Note que o vetor δ_u é construído como em (121), considerando-se que P_{m+1} é construída a partir da resposta impulsiva do filtro dado por

$$P_{m+1}(q) = C(q, \rho_0)S(q, \rho_0)H(q). \quad (148)$$

Observe agora que as expressões (146) e (148) possuem um termo em comum dado por $S(q, \rho_0)H(q)$, que é desconhecido, pois depende do modelo do ruído $H(q)$ e da função de sensibilidade $S(q, \rho_0)$ obtida a partir do modelo do processo $G(q)$, assumido ser desconhecido. Contudo, esse inconveniente pode ser contornado. Veja que na formulação do problema de otimização do CTLS, (123) a (125), não é feita nenhuma suposição quanto à estrutura do sinal de ruído, ou seja, v pode ser, por exemplo, um ruído branco ou colorido. A única suposição a ser garantida é que a mesma realização de ruído afeta as colunas de Φ e u (o que é o mesmo que dizer que a mesma realização de ruído afeta as colunas de Δ_Z em (128)). Portanto, se considerarmos que as colunas de Δ_Z são afetadas por um ruído colorido, $\eta = [\eta(1) \ \eta(2) \ \dots \ \eta(N)]^T$, com $\eta(t) = S(q, \rho_0)H(q)v(t)$, temos η ao invés de v em (133). Com base nisso, os filtros para o caso de dados de malha fechada com um controlador linearmente parametrizado e com polos fixos, são dados por

$$P_i(q) = \begin{cases} -\left[T_d^{-1}(q) - 1\right]\bar{C}_i(q), & i = 1, \dots, m \\ C(q, \rho_0), & i = m + 1 \end{cases} \quad (149)$$

que dependem apenas de informações conhecidas. Observe que, devido à formulação do problema em (124), foi preciso ajustar o sinal dos m primeiros filtros. Assim sendo, as respostas impulsivas dos $m + 1$ filtros definidos acima são utilizadas para gerar as matrizes Toeplitz P_i do problema de otimização do CTLS, e os parâmetros do controlador são estimados como descrito no Teorema 3.1.

Filtros para o caso de dados de malha aberta: Considere que os dados são coletados como descrito na Definição 2.3, ou seja, a entrada do controlador a ser identificado é afetada por ruído e sua saída é um sinal determinístico. Como consequência, apenas a matriz Φ é afetada por ruído. Seguindo os mesmos passos do caso de malha fechada, escreve-se o vetor regressor substituindo-se (1) e (2) em (65), obtendo-se

$$\begin{aligned} \varphi(t) &= \left[T_d^{-1}(q) - 1\right]\bar{C}(q)[G(q)u(t) + H(q)v(t)] \\ &= \underbrace{\left[T_d^{-1}(q) - 1\right]\bar{C}(q)G(q)u(t)}_{\varphi_0(t)} + \underbrace{\left[T_d^{-1}(q) - 1\right]\bar{C}(q)H(q)v(t)}_{\Delta_\Phi(t)}, \end{aligned} \quad (150)$$

onde o vetor regressor foi dividido em um termo $\varphi_0(t)$ dependendo apenas do sinal de entrada e outro termo $\Delta_{\Phi}(t)$ dependendo puramente do ruído.

Assim como no caso de dados de malha fechada, a estrutura dos filtros é obtida avaliando-se (150) em forma vetorial. Como apresentado em (120), as m colunas de Δ_{Φ} são formadas multiplicando-se as matrizes Toeplitz P_i pelo vetor v . Essas matrizes são construídas a partir das respostas impulsivas dos filtros definidos a partir de $\Delta_{\varphi}(t)$ em (150), que são dados por

$$P_i(q) = -\left[T_d^{-1}(q) - 1\right]\bar{C}_i(q)H(q), \quad \text{com } i = 1, \dots, m. \quad (151)$$

Como a saída do controlador a ser identificado $u(t)$ não é afetada por ruído, define-se

$$P_{m+1}(q) = 0. \quad (152)$$

Nesse caso, nota-se que os primeiros m filtros dependem do modelo do ruído $H(q)$, que é desconhecido. Assim como no caso anterior, esse inconveniente pode ser contornado se consideramos que as colunas de Δ_Z são afetadas por um ruído colorido. Com essa consideração, temos $w(t) = H(q)v(t)$, assim, pode-se dizer que ao invés de v em (120) e (121) tem-se $w = [w(1) \dots w(N)]^T$. Desta forma, os filtros são simplificados para

$$P_i(q) = \begin{cases} -\left[T_d^{-1}(q) - 1\right]\bar{C}_i(q), & i = 1, \dots, m \\ 0, & i = m + 1. \end{cases} \quad (153)$$

Observe que no caso de dados de malha aberta os filtros também dependem apenas de informações conhecidas. Além disso, novamente, ajustou-se o sinal dos m primeiros filtros para concordar com a formulação do problema em (124).

3.2.2 Controlador com polos livres

Considere que a estrutura do controlador é dividida em duas partes, como apresentado em (21) e reproduzido abaixo:

$$C(q, \rho) = \frac{B(q, \rho)}{A(q, \rho)}C_F(q) \quad (21)$$

$$= C_I(q, \rho)C_F(q), \quad (154)$$

onde $B(q, \rho)$ e $A(q, \rho)$ são dados por (22) e (23), respectivamente. Além disso, $C_I(q, \rho) = B(q, \rho)/A(q, \rho)$ é a parte identificável e $C_F(q)$ a parte fixa do controlador. Com esta estrutura, os sinais de entrada e saída do controlador a ser projetado com o método VRFT são apresentados na Figura 11. O sinal $\bar{e}_F(t)$ é definido como

$$\bar{e}_F(t) = C_F(q)\bar{e}(t) \quad (155)$$

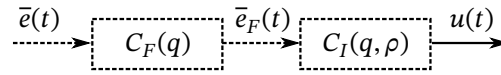
$$= C_F(q)\left[T_d^{-1}(q) - 1\right]y(t), \quad (156)$$

onde (156) é obtida substituindo-se (57) em (155).

O objetivo agora é identificar $C_I(q, \rho)$ a partir dos sinais de entrada $\bar{e}_F(t)$ e saída $u(t)$. Uma vez que $B(q, \rho)$ e $A(q, \rho)$ são lineares nos parâmetros considera-se que esta estrutura é do tipo ARX. Seguindo a descrição do modelo ARX na Subseção 2.3.1 tem-se que:

$$u(t) = B(q, \rho)\bar{e}_F(t) + [1 - A(q, \rho)]u(t) + x(t), \quad (157)$$

Figura 11 — Sinais de entrada e saída para o projeto de um controlador com polos livres.



Fonte: elaborado pela autora.

onde $x(t)$ é o erro do modelo. Um preditor para o sinal de saída é dado por

$$\hat{u}(t, \rho) = B(q, \rho)\bar{e}_F(t) + [1 - A(q, \rho)]u(t). \quad (158)$$

Esse preditor é reescrito a partir de (22) e (23) como

$$\hat{u}(t, \rho) = \rho^T \varphi(t), \quad (159)$$

onde o vetor de parâmetros ρ e o vetor regressor $\varphi(t)$ são definidos como

$$\rho^T = [b_1 \quad \dots \quad b_{n_b} \quad a_1 \quad \dots \quad a_{n_a}], \quad (160)$$

$$\varphi(t) = [\bar{e}_F(t) \quad \dots \quad \bar{e}_F(t - n_b + 1) \quad -u(t-1) \quad \dots \quad -u(t - n_a)]^T. \quad (161)$$

De $u(t)$, (157), e do preditor, (159), tem-se que o erro de predição é o mesmo definido em (60) (considerando-se $L(q) = 1$). Assim, a função custo a ser minimizada para este caso é dada por

$$J^{VR}(\rho) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N |u(t) - \rho^T \varphi(t)|^2. \quad (162)$$

Nota 3.2. Observe que (162) é a mesma função custo da formulação original do VRFT (61) (assumindo-se $L(q) = 1$). A diferença entre a formulação usual e ARX consiste na construção do vetor regressor. Considerando a estrutura ARX, o vetor regressor em (161) é formado a partir de $\bar{e}_F(t)$ e $u(t)$. Contudo, na formulação original $u(t)$ não aparece no vetor regressor, pois o denominador da estrutura do controlador é fixo.

Considere agora a matriz regressora Φ , (112), construída com o vetor regressor em (161), definida como

$$\Phi = [\Phi_e \quad \Phi_u], \quad (163)$$

onde as matrizes $\Phi_e \in \mathbb{R}^{N \times n_b}$ e $\Phi_u \in \mathbb{R}^{N \times n_a}$ são dadas por

$$\Phi_e = \begin{bmatrix} \bar{e}_F(1) & 0 & \dots & 0 \\ \bar{e}_F(2) & \bar{e}_F(1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{e}_F(N) & \bar{e}_F(N-1) & \dots & \bar{e}_F(N - n_b + 1) \end{bmatrix}, \quad (164)$$

$$\Phi_u = \begin{bmatrix} 0 & 0 & \dots & 0 \\ -u(1) & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -u(N-1) & -u(N-2) & \dots & -u(N - n_a) \end{bmatrix}. \quad (165)$$

Como pode ser observado em (164) e (165), as colunas das matrizes Φ_e e Φ_u são formadas pelos respectivos vetores

$$\bar{e}_F = [\bar{e}_F(1) \quad \bar{e}_F(2) \quad \dots \quad \bar{e}_F(N)]^T, \quad (166)$$

$$u = [-u(1) \quad -u(2) \quad \dots \quad -u(N)]^T, \quad (167)$$

e as versões atrasadas destes vetores. Da estrutura da matriz regressora acima definida são gerados os filtros $P_i(q)$ utilizados na solução do problema de otimização do CTLS. Nas próximas subseções são apresentadas as estruturas desses filtros para o caso de dados de malha fechada e aberta.

Filtros para o caso de dados de malha fechada: Considere que os dados de entrada $u(t)$ e saída $y(t)$ do processo são coletados como descrito na Definição 2.2. Como já mencionado, o ruído corrompe os dados de entrada e saída neste caso. Por conta disso, os sinais de entrada $\bar{e}_F(t)$ e saída $u(t)$ da parte identificável do controlador $C_I(q, \rho)$ são afetados por ruído. Assim, $m + 1$ filtros precisam ser determinados, ou seja, todas as colunas da matriz Δ_Z em (129) são diferentes de zero. As estruturas dos $n_b + n_a$ primeiros filtros são obtidas a partir da matriz Φ em (163). Primeiramente, leva-se em consideração o sinal $\bar{e}_F(t)$ que compõe a matriz Φ_e em (164). Para tanto, substitui-se (4) e (2) em (156), obtendo-se

$$\begin{aligned} \bar{e}_F(t) &= \left[T_d^{-1}(q) - 1 \right] C_F(q) (T(q, \rho_0)r(t) + S(q, \rho_0)H(q)v(t)) \\ &= \underbrace{\left[T_d^{-1}(q) - 1 \right] C_F(q) T(q, \rho_0)r(t)}_{\bar{e}_0(t)} + \underbrace{\left[T_d^{-1}(q) - 1 \right] C_F(q) S(q, \rho_0)H(q)v(t)}_{\delta_{\bar{e}}(t)}, \end{aligned} \quad (168)$$

onde $S(q, \rho_0)$ e $T(q, \rho_0)$ são a função de sensibilidade e a função de sensibilidade complementar obtidas como em (6) e (7) com o controlador $C(q, \rho_0)$, que é o controlador *inicial* operando na malha enquanto são coletados os dados do experimento. Além disso, o sinal de entrada do controlador a ser identificado foi dividido em um termo $\bar{e}_0(t)$ dependente do sinal de referência e outro termo $\delta_{\bar{e}}(t)$ dependente apenas da contribuição do ruído.

Note que Φ_e em (164) é formada pelo sinal $\bar{e}_F(t)$ e suas versões atrasadas. Portanto, de $\delta_{\bar{e}}(t)$ em (168) e (164), concluímos que as primeiras n_b colunas de Δ_Z em (129) são formadas a partir das matrizes Toeplitz P_i , com $i = 1, \dots, n_b$, multiplicadas por v . Essas matrizes são construídas a partir das respostas impulsivas dos filtros

$$P_i(q) = - \left[T_d^{-1}(q) - 1 \right] C_F(q) S(q, \rho_0) H(q) q^{-(i-1)}, \quad \text{com } i = 1, \dots, n_b. \quad (169)$$

Relembrando que a matriz Φ é formada por Φ_e e Φ_u , temos que os n_a filtros seguintes são determinados a partir de Φ_u . De (165) observa-se que as colunas de Φ_u são formadas a partir de versões atrasadas do vetor u em (167). Portanto, reescreve-se $u(t)$ substituindo-se (2) em (5):

$$u(t) = \underbrace{S(q, \rho_0)C(q, \rho_0)r(t)}_{u_0(t)} - \underbrace{C(q, \rho_0)S(q, \rho_0)H(q)v(t)}_{\delta_u(t)}, \quad (147)$$

onde o sinal de saída foi dividido em duas partes: um termo $u_0(t)$ com a contribuição do sinal de referência e outro termo $\delta_u(t)$ apenas com a contribuição do ruído. Da constatação acima mencionada de que as colunas de Φ_u são formadas por versões atrasadas do vetor u e de $\delta_u(t)$, determinamos que as n_a colunas seguintes de Δ_Z em (129) são formadas por matrizes Toeplitz P_i multiplicadas por v . Essas matrizes são formadas com as respostas impulsivas dos seguintes filtros

$$P_i(q) = -C(q, \rho_0)S(q, \rho_0)H(q)q^{-(i-n_b)} \quad \text{com } i = n_b + 1, \dots, n_b + n_a. \quad (170)$$

Finalmente, é preciso determinar o filtro que será empregado para gerar δ_u , que é a matriz P_{m+1} multiplicada por v , produzindo a última coluna de Δ_Z em (129). Esse filtro é obtido diretamente a partir de $\delta_u(t)$ em (147), ou seja,

$$P_{m+1} = C(q, \rho_0)S(q, \rho_0)H(q). \quad (171)$$

Os $m + 1$ filtros são descritos em (169), (170) e (171). Observe que essas expressões apresentam $S(q, \rho_0)H(q)$ como termo comum, que é desconhecido. Novamente, se considerarmos que as colunas de Δ_Z são afetadas por um ruído colorido, $\eta = [\eta(1) \ \eta(2) \ \dots \ \eta(N)]^T$ e que $\eta(t) = S(q, \rho_0)H(q)v(t)$, temos η ao invés de v na formulação do problema do CTLS (123) a (125). Nesse caso, os filtros para o caso de dados de malha fechada e controlador com polos livres são simplificados para

$$P_i(q) = \begin{cases} -\left[T_d^{-1}(q) - 1\right] C_F(q) q^{-(i-1)}, & i = 1, \dots, n_b \\ -C(q, \rho_0) q^{-(i-n_b)}, & i = n_b + 1, \dots, m = (n_b + n_a) \\ C(q, \rho_0), & i = m + 1 \end{cases} \quad (172)$$

que dependem apenas de quantidades conhecidas. Neste caso, o sinal dos n_b primeiros filtros foi ajustado para concordar com a formulação do problema em (124).

Filtros para o caso de dados de malha aberta: Assuma que os dados de entrada e saída do processo são coletados como na Definição 2.3. Nesse caso, apenas a entrada da parte identificável do controlador $C_I(q, \rho)$, $\bar{e}_F(t)$, é afetada por ruído, enquanto sua saída é um sinal determinístico. Por conta disso, apenas n_b filtros precisam ser determinados, pois somente Φ_e em (163) é afetada por um sinal ruidoso. Assim sendo, somente as primeiras n_b colunas de Δ_Z em (129) são formadas por vetores diferentes de zero.

Para determinar a estrutura dos n_b filtros, reescreve-se $\bar{e}_F(t)$ substituindo (1) e (2) em (156), obtendo:

$$\begin{aligned} \bar{e}_F(t) &= \left[T_d^{-1}(q) - 1\right] C_F(q) [G(q)u(t) + H(q)v(t)] \\ &= \underbrace{\left[T_d^{-1}(q) - 1\right] C_F(q) G(q) u(t)}_{\bar{e}_0(t)} + \underbrace{\left[T_d^{-1}(q) - 1\right] C_F(q) H(q) v(t)}_{\delta_{\bar{e}}(t)}, \end{aligned} \quad (173)$$

onde este sinal foi dividido em duas partes: uma parte $\bar{e}_0(t)$ contendo a contribuição da entrada e outra parte $\delta_{\bar{e}}(t)$ contendo somente a contribuição do ruído. De $\delta_{\bar{e}}(t)$ e lembrando que as colunas de Φ_e em (164) são formadas a partir de $\bar{e}_F(t)$ e suas versões atrasadas, temos que as primeiras n_b colunas de Δ_Z em (129) são formadas através de matrizes Toeplitz P_i multiplicadas por v , e que essas matrizes são formadas a partir da resposta impulsiva dos filtros

$$P_i(q) = -\left[T_d^{-1}(q) - 1\right] C_F(q) H(q) q^{-(i-1)}, \quad \text{com } i = 1, \dots, n_b. \quad (174)$$

Como o sinal de saída do controlador a ser identificado não é afetado por ruído, os demais filtros são definidos como

$$P_i(q) = 0, \quad \text{com } i = n_b + 1, \dots, m + 1. \quad (175)$$

Observe que os filtros em (174) dependem do modelo do ruído, o qual assume-se ser desconhecido. Novamente, se considerarmos que as colunas de Δ_Z são afetadas por um ruído colorido, $w = [w(1) \ w(2) \ \dots \ w(N)]^T$ onde $w(t) = H(q)v(t)$, temos w ao invés de v no problema de otimização do CTLS. Com essa consideração, os filtros são simplificados para

$$P_i(q) = \begin{cases} -\left[T_d^{-1}(q) - 1\right] C_F(q) q^{-(i-1)}, & i = 1, \dots, n_b \\ 0, & i = n_b + 1, \dots, m + 1, \end{cases} \quad (176)$$

que dependem somente de informações conhecidas. Neste caso, ajustou-se o sinal dos n_b primeiros filtros para concordar com a formulação do problema em (124).

3.2.3 O caso em que o controlador ideal não pertence à classe de controle

O objetivo desta subseção consiste em desenvolver os filtros para o problema CTLS considerando-se a violação da Suposição 2.1. Como será apresentado, a estrutura dos filtros é determinada pela escolha da estrutura do controlador e do tipo de dado (malha aberta ou fechada).

3.2.3.1 Dados de malha aberta

Assuma que os dados são coletados de um experimento em malha aberta como descrito na Definição 2.3. Neste caso, a estrutura do filtro que aproxima o mínimo da função custo $J^{MR}(\rho)$, (13), do mínimo do problema de otimização do CTLS, (123) a (125), é dada no Teorema 3.2.

Teorema 3.2. *Quando a Suposição 2.1 não é respeitada e os dados são coletados durante um experimento em malha aberta, o filtro $L(e^{j\omega})$ que aproxima o mínimo do problema de otimização do CTLS do mínimo da função custo $J^{MR}(\rho)$ é dado por*

$$|L(e^{j\omega})|^2 \approx \frac{|T_d(e^{j\omega})|^2 |1 - T_d(e^{j\omega})|^2}{\sum_{i=1}^m |P_i(e^{j\omega})|^2 |G(e^{j\omega})|^2}, \quad \forall \omega \in [-\pi; \pi] \quad (177)$$

onde a estrutura de $P_i(e^{j\omega})$ é determinada pela estrutura do controlador a ser projetado.

Prova. Considere que os dados provêm de um experimento em malha aberta, o problema CTLS de (123) a (125) é reescrito como

$$\min_{v, \rho} J^{\text{CTLS}}(v) = \min_{v, \rho} \|[P_1 v \dots P_m v]\|_F^2 \quad (178)$$

$$\text{sujeito a } (\Phi - \Delta_\Phi)\rho = u - \delta_u. \quad (179)$$

Observe que a função custo em (178) pode ser reescrita como

$$J^{\text{CTLS}}(v) = \sum_{i=1}^m \|P_i v\|_F^2 \quad (180)$$

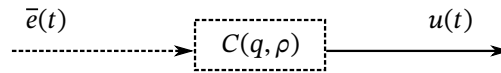
$$= \sum_{i=1}^m \|P_i v\|_2^2 \quad (181)$$

$$\approx \sum_{i=1}^m |P_i(q)v(t)|^2, \quad (182)$$

onde (181) é obtida lembrando que $P_i v$ é um vetor e que, dado um vetor x , $\|x\|_F^2 = \text{tr}\{x^T x\} = \|x\|_2^2$. Além disso, como P_i é uma matriz Toeplitz representando a resposta impulsiva do filtro $P_i(q)$, a equação em (181) pode ser aproximada por (182).

Assuma que o controlador a ser identificado através do método VRFT possui uma estrutura genérica, como mostra a Figura 12. Neste caso, os sinais de entrada e saída do controlador a ser identificado são $\bar{e}(t)$ e $u(t)$. O sinal de saída é dado por

Figura 12 — Sinais de entrada e saída de um controlador genérico.



Fonte: elaborado pela autora.

$$u(t) = C(q, \rho)\bar{e}(t) + x(t), \quad (183)$$

onde onde $x(t)$ é o erro do modelo. Um preditor para a saída é definido como

$$\hat{u}(t, \rho) = C(q, \rho)\bar{e}(t) \quad (184)$$

$$= C(q, \rho) \left[T_d^{-1}(q) - 1 \right] y(t), \quad (185)$$

onde (185) é obtida substituindo-se (57) em (184).

De $u(t)$ (183) e $\hat{u}(t, \rho)$ (185) obtém-se o erro de predição

$$\varepsilon(t, \rho) = \hat{u}(t, \rho) - u(t) \quad (186)$$

$$= C(q, \rho) \left[T_d^{-1}(q) - 1 \right] (G(q)u(t) + v(t)) - u(t), \quad (187)$$

onde $v(t)$ representa o ruído (podendo ser colorido) que afeta o sinal de saída coletado $y(t)$. O erro de predição considerando-se o vetor de parâmetros ρ_d é dado por

$$\varepsilon(t, \rho_d) = C_d(q) \left[T_d^{-1}(q) - 1 \right] G(q)u(t) + C_d(q) \left[T_d^{-1}(q) - 1 \right] v(t) - u(t). \quad (188)$$

Observe que o termo $C_d(q)[T_d^{-1}(q) - 1]G(q)$ pode ser desenvolvido como

$$C_d(q) \left[T_d^{-1}(q) - 1 \right] G(q) = C_d(q) \left[S_d^{-1}(q)G^{-1}(q)C_d^{-1}(q) - 1 \right] G(q) \quad (189)$$

$$= S_d^{-1}(q) - C_d(q)G(q) \quad (190)$$

$$= 1 + C_d(q)G(q) - C_d(q)G(q) = 1, \quad (191)$$

onde (189) é obtida lembrando que $T_d(q) = C_d(q)G(q)S_d(q)$ e obteve-se (191) a partir de (6) considerando-se o controlador ideal $C_d(q)$. A partir de (191) tem-se:

$$C_d(q) \left[T_d^{-1}(q) - 1 \right] = G^{-1}(q). \quad (192)$$

Com (191) reescreve-se (188) como

$$\varepsilon(t, \rho_d) = C_d(q) \left[T_d^{-1}(q) - 1 \right] v(t) + u(t) - u(t) \quad (193)$$

$$= G^{-1}(q)v(t), \quad (194)$$

obtida substituindo-se (192) em (193). Partindo de (194), obtém-se a seguinte expressão para o ruído

$$v(t) = G(q)\varepsilon(t, \rho_d),$$

Com base nisso, pode-se obter uma expressão para a estimativa do ruído da seguinte forma

$$\hat{v}(t, \rho) = G(q)\varepsilon(t, \rho). \quad (195)$$

Agora, utilizando-se (195) reescreve-se a função custo $J^{\text{CTLS}}(v)$ como

$$J^{\text{CTLS}}(v) \approx \sum_{i=1}^m \|P_i(q)\hat{v}(t, \rho)\|_2^2, \quad (196)$$

que é novamente reescrita substituindo-se (195) na equação anterior, obtendo-se

$$J^{\text{CTLS}}(\rho) \approx \sum_{i=1}^m \|P_i(q)G(q)\varepsilon(t, \rho)\|_2^2. \quad (197)$$

O erro de predição $\varepsilon(t, \rho)$ é escrito agora a partir de (187) assumindo-se que o ruído real que afeta o processo é zero e que os sinais são filtrados por um filtro $L(q)$

$$\begin{aligned} \varepsilon(t, \rho) &= C(q, \rho) \left[T_d^{-1}(q) - 1 \right] G(q)L(q)u(t) - L(q)u(t) \\ &= \left[C(q, \rho)T_d^{-1}(q)G(q) - (1 + C(q, \rho)G(q)) \right] L(q)u(t) \end{aligned} \quad (198)$$

$$= \left[C(q, \rho)T_d^{-1}(q)G(q) - S^{-1}(q, \rho) \right] L(q)u(t) \quad (199)$$

$$= [C(q, \rho)G(q)S(q, \rho) - T_d(q)] S^{-1}(q, \rho) T_d^{-1}(q) L(q)u(t) \quad (200)$$

$$= [T(q, \rho) - T_d(q)] (1 - T(q, \rho))^{-1} T_d^{-1}(q) L(q)u(t), \quad (201)$$

onde (199) é obtida substituindo-se (6) em (198), obtém-se (200) isolando-se $S^{-1}(q)T_d^{-1}(q)$ em (199) e (201) é obtida da relação $S(q, \rho) = 1 - T(q, \rho)$.

Uma nova expressão é obtida para a função custo do CTLS, substituindo-se (201) em (197), como descrito abaixo

$$J^{\text{CTLS}}(\rho) \approx \sum_{i=1}^m \left\| \frac{L(q)P_i(q)G(q) [T(q, \rho) - T_d(q)]}{T_d(q)(1 - T(q, \rho))} u(t) \right\|_2^2. \quad (202)$$

Aplicando-se o teorema de Parseval (Teorema 2.1) tem-se:

$$\begin{aligned} \lim_{N \rightarrow \infty} J^{\text{CTLS}}(\rho) &\approx \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\sum_{i=1}^m |P_i(e^{j\omega})|^2 |G(e^{j\omega})|^2 |T(e^{j\omega}, \rho) - T_d(e^{j\omega})|^2}{|T_d(e^{j\omega})|^2 |1 - T(e^{j\omega}, \rho)|^2} \\ &\quad \times |L(e^{j\omega})|^2 \Phi_u(e^{j\omega}) d\omega. \end{aligned} \quad (203)$$

Considere a função custo $J^{\text{MR}}(\rho)$ em (13) escrita no domínio da frequência com o teorema de Parseval com o erro de predição $\varepsilon_{\text{MR}}(t, \rho)$ dado por (14)

$$J^{\text{MR}}(\rho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |T(e^{j\omega}, \rho) - T_d(e^{j\omega})|^2 \Phi_r(e^{j\omega}) d\omega. \quad (204)$$

Observe que (204) e (203) seriam iguais se

$$|L(e^{j\omega})|^2 = \frac{|T_d(e^{j\omega})|^2 |1 - T(e^{j\omega}, \rho)|^2 \Phi_r(e^{j\omega})}{\sum_{i=1}^m |P_i(e^{j\omega})|^2 |G(e^{j\omega})|^2 \Phi_u(e^{j\omega})}. \quad (205)$$

Através das mesmas aproximações do filtro do VRFT, sendo elas: $S(q, \rho) \approx S_d(q) = 1 - T_d(q)$ e assumindo-se que o sinal de referência que é normalmente aplicado no processo

é do mesmo tipo que o sinal de entrada em malha aberta, ou seja, $\Phi_r(e^{j\omega}) \approx \Phi_u(e^{j\omega})$, reescreve-se o filtro (205) como

$$|L(e^{j\omega})|^2 \approx \frac{|T_d(e^{j\omega})|^2 |1 - T_d(e^{j\omega})|^2}{\sum_{i=1}^m |P_i(e^{j\omega})|^2 |G(e^{j\omega})|^2}. \quad (206)$$

■

Observe que a estrutura do filtro em (206) depende de quantidades conhecidas, exceto pela amplitude da inversa da resposta em frequência do processo. Esta pode ser estimada a partir de uma identificação de um modelo FIR de alta ordem. Vale salientar que quanto melhor for a estimativa da amplitude da inversa da resposta em frequência do processo, mais próximos serão os mínimos das funções custo $J^{\text{CTLS}}(\rho)$ e $J^{\text{MR}}(\rho)$. Na sequência, apresenta-se as estruturas dos filtros considerando primeiramente um controlador com polos fixos e em seguida com polos livres.

Filtro considerando um controlador com polos fixos: Suponha que o controlador seja linearmente parametrizado com polos fixos, Definição 2.4, e que a estrutura escolhida para $\bar{C}(q)$ seja dada por

$$\bar{C}(q) = \left[\frac{q^{m-1}}{q^{m-2}(q-1)} \quad \cdots \quad \frac{1}{q^{m-2}(q-1)} \right]^T. \quad (207)$$

Nesse caso, os filtros são dados por (153), e são utilizados para gerar o fator $\sum_{i=1}^m |P_i(e^{j\omega})|^2$ em (206), desenvolvido abaixo

$$\begin{aligned} \sum_{i=1}^m |P_i(e^{j\omega})|^2 &= \sum_{i=1}^{m-1} |e^{j\omega(m-i)}|^2 \frac{|T_d^{-1}(e^{j\omega}) - 1|^2}{|e^{j\omega(m-2)}|^2 |e^{j\omega} - 1|^2} \\ &= (m-1) \frac{|T_d^{-1}(e^{j\omega}) - 1|^2}{|e^{j\omega} - 1|^2}. \end{aligned} \quad (208)$$

Substituindo-se (208) em (206) obtém-se

$$\begin{aligned} |L(e^{j\omega})|^2 &\approx \frac{|T_d(e^{j\omega})|^2 |1 - T_d(e^{j\omega})|^2 |e^{j\omega} - 1|^2}{(m-1) |T_d^{-1}(e^{j\omega}) - 1|^2 |G(e^{j\omega})|^2} \\ &\approx \frac{|T_d(e^{j\omega})|^2 |T_d(e^{j\omega})|^2 |e^{j\omega} - 1|^2}{(m-1) |G(e^{j\omega})|^2}, \end{aligned} \quad (209)$$

lembrando que $[T_d^{-1}(e^{j\omega}) - 1] = [1 - T_d(e^{j\omega})]/T_d(e^{j\omega})$.

Nota 3.3. Observe que a estrutura de controle em (207), definida em (18), foi escolhida para simplificar o desenvolvimento do filtro. Contudo, imagine que deseja-se utilizar uma outra estrutura, como por exemplo a descrita em (17). Neste caso, projeta-se o controlador com o filtro desenvolvido considerando-se a estrutura em (18) e então o vetor de ganhos estimado é convertido para a estrutura desejada através da transformação linear em (20).

Filtro considerando um controlador com polos livres: Assuma que a estrutura do controlador é dada em (21). Os filtros para este caso foram determinados na Seção 3.2.2 e são descritos em (176). A resposta em frequência destes filtros é utilizada para gerar a parcela $\sum_{i=1}^m |P_i(e^{j\omega})|^2$ do filtro em (206) como apresentado abaixo

$$\begin{aligned} \sum_{i=1}^{n_b} |P_i(e^{j\omega})|^2 &= \sum_{i=1}^{n_b-1} |e^{j\omega(m-i)}|^{-2} |T_d^{-1}(e^{j\omega}) - 1|^2 |C_F(e^{j\omega})|^2 \\ &= (n_b - 1)^{-1} |T_d^{-1}(e^{j\omega}) - 1|^2 |C_F(e^{j\omega})|^2. \end{aligned} \quad (210)$$

Substituindo-se a equação acima em (206) tem-se

$$|L(e^{j\omega})|^2 \approx \frac{|T_d(e^{j\omega})|^2 |1 - T_d(e^{j\omega})|^2 (n_b - 1)}{|T_d^{-1}(e^{j\omega}) - 1|^2 |C_F(e^{j\omega})|^2 |G(e^{j\omega})|^2} \quad (211)$$

$$= \frac{|T_d(e^{j\omega})|^2 |T_d(e^{j\omega})|^2 (n_b - 1)}{|C_F(e^{j\omega})|^2 |G(e^{j\omega})|^2}, \quad (212)$$

lembrando novamente que $[T_d^{-1}(e^{j\omega}) - 1] = [1 - T_d(e^{j\omega})]/T_d(e^{j\omega})$.

3.2.3.2 Estrutura do filtro para dados de malha fechada

Assuma que os dados de entrada e saída são coletados durante um experimento em malha fechada, como descrito na Definição 2.2 e que a Suposição 2.1 não é atendida. Nesse caso, o filtro $L(e^{j\omega})$ que força que o mínimo de $J^{\text{CTLS}}(\rho)$ seja próximo do mínimo da função custo $J^{\text{MR}}(\rho)$ é descrito no Teorema 3.3.

Teorema 3.3. *Considerando que a Suposição 2.1 não é atendida e que os dados são coletados durante um experimento em malha fechada, o filtro $L(e^{j\omega})$ que faz com que o mínimo de $J^{\text{CTLS}}(\rho)$ seja próximo do mínimo de $J^{\text{MR}}(\rho)$ é dado por*

$$|L(e^{j\omega})|^2 \approx \frac{|T_d(e^{j\omega})|^2 |1 - T_d(e^{j\omega})|^2}{\sum_{i=1}^{m+1} |P_i(e^{j\omega})|^2 |T(e^{j\omega}, \rho_0)|^2}, \quad (213)$$

onde a estrutura de $P_i(e^{j\omega})$, com $i = 1, \dots, m + 1$, é determinada pela estrutura escolhida para o controlador.

Prova. Quando os dados provém de um experimento em malha fechada, o problema CTLS é formulado como

$$\min_{v, \rho} J^{\text{CTLS}}(v) = \min_{v, \rho} \|[P_1 v \dots P_{m+1} v]\|_F^2 \quad (214)$$

$$\text{sujeito a } (\Phi - \Delta_\Phi)\rho = u - \delta_u. \quad (215)$$

Veja que a diferença entre (178) e (214) é que $J^{\text{CTLS}}(v)$ possui $m + 1$ colunas. Por conta disso, pode-se fazer a mesma aproximação em (182). Considerando um controlador com estrutura genérica $C(q, \rho)$, com os sinais de entrada e saída como na Figura 12, e $u(t)$ e

seu preditor dados em (183) e (185), o erro de predição é definido, considerando-se dados de malha fechada (Definição 2.2), como

$$\varepsilon(t, \rho) = \hat{u}(t, \rho) - u(t) \quad (216)$$

$$= C(q, \rho) \left[T_d^{-1}(q) - 1 \right] [T(q, \rho_0)r(t) + S(q, \rho_0)v(t)] - S(q, \rho_0)C(q, \rho_0)[r(t) - v(t)]. \quad (217)$$

Nesse caso, assume-se que o ruído afetando a saída é $v(t)$. Contudo, o desenvolvimento do filtro poderia ser feito considerando-se $w(t)$ afetando a saída e, conseqüentemente, w ao invés de v em (214).

O erro de predição, considerando-se o vetor de parâmetros ideal, é dado por

$$\varepsilon(t, \rho_d) = \left[C_d(q) \left(T_d^{-1}(q) - 1 \right) G(q)C(q, \rho_0)r(t) - C(q, \rho_0)r(t) + C(q, \rho_0)v(t) + \left(T_d^{-1}(q) - 1 \right) v(t) \right] S(q, \rho_0) \quad (218)$$

$$= [C(q, \rho_0)G(q) + 1]S(q, \rho_0)G^{-1}(q)v(t) \quad (219)$$

$$= G^{-1}(q)v(t), \quad (220)$$

onde (219) é obtida substituindo-se (191) em (218) e encontra-se (220) lembrando que $S^{-1}(q, \rho_0) = 1 + C(q, \rho_0)G(q)$. A partir de (220) tem-se

$$v(t) = G(q)\varepsilon(t, \rho_d). \quad (221)$$

Desta equação, obtém-se um preditor para o ruído como

$$\hat{v}(t, \rho) = G(q)\varepsilon(t, \rho). \quad (222)$$

Agora, com o preditor do ruído da equação acima reescreve-se a função custo em (182) como

$$J^{\text{CTLS}}(\rho) \approx \sum_{i=1}^{m+1} |P_i(q)\hat{v}(t, \rho)|^2 \quad (223)$$

$$\approx \sum_{i=1}^{m+1} |P_i(q)G(q)\varepsilon(t, \rho)|^2. \quad (224)$$

Na sequência, escreve-se o erro de predição $\varepsilon(t, \rho)$ a partir de (217) assumindo-se que o ruído real que afeta o processo é zero ($v(t) = 0$) e que os sinais são filtrados por um filtro $L(q)$, como segue

$$\varepsilon(t, \rho) = \left[C(q, \rho) \left(T_d^{-1}(q) - 1 \right) G(q)S(q, \rho_0)C(q, \rho_0) - S(q, \rho_0)C(q, \rho_0) \right] L(q)r(t) \quad (225)$$

$$= \left[C(q, \rho)T_d^{-1}(q)G(q) - (1 + C(q, \rho)G(q)) \right] S(q, \rho_0)C(q, \rho_0)L(q)r(t) \quad (226)$$

$$= \left[C(q, \rho)T_d^{-1}(q)G(q) - S^{-1}(q, \rho) \right] S(q, \rho_0)C(q, \rho_0)L(q)r(t) \quad (227)$$

$$= [C(q, \rho)G(q)S(q, \rho) - T_d(q)] T_d^{-1}(q)S^{-1}(q, \rho)S(q, \rho_0)C(q, \rho_0)L(q)r(t) \quad (228)$$

$$= [T(q, \rho) - T_d(q)][1 - T(q, \rho)]^{-1} T_d^{-1}(q)S(q, \rho_0)C(q, \rho_0)L(q)r(t), \quad (229)$$

onde (225) é obtida da relação (7), (226) é obtida colocando $S(q, \rho_0)C(q, \rho_0)$ em evidência, enquanto calcula-se (227) lembrando de (6). Além disso, (228) é encontrada colocando $T_d^{-1}(q)$ e $S^{-1}(q, \rho)$ em evidência e, por fim, (229) é obtida lembrando das relações (7) e $S^{-1}(q, \rho) = [1 - T(q, \rho)]^{-1}$.

Substituindo-se o erro de predição (229) em (224) e lembrando da relação (7), reescreve-se a função custo $J^{\text{CTLS}}(\rho)$ como

$$J^{\text{CTLS}}(\rho) \approx \sum_{i=1}^{m+1} \left| \frac{P_i(q) [T(q, \rho) - T_d(q)] T(q, \rho_0) L(q) r(t)}{T_d(q) [1 - T(q, \rho)]} \right|^2, \quad (230)$$

que é reescrita no domínio da frequência com o teorema de Parseval da seguinte forma

$$\begin{aligned} \lim_{N \rightarrow \infty} J^{\text{CTLS}}(\rho) &\approx \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\sum_{i=1}^{m+1} |P_i(e^{j\omega})|^2 |T(q, e^{j\omega}) - T_d(e^{j\omega})|^2 |T(e^{j\omega}, \rho_0)|^2}{|T_d(e^{j\omega})|^2 |1 - T(q, e^{j\omega})|^2} \\ &\times |L(e^{j\omega})|^2 \Phi_r(e^{j\omega}) d\omega. \end{aligned} \quad (231)$$

Observe que a função custo em (231) seria igual a função custo $J^{\text{MR}}(\rho)$ em (204) se

$$|L(e^{j\omega})|^2 = \frac{|T_d(e^{j\omega})|^2 |1 - T(q, e^{j\omega})|^2}{\sum_{i=1}^{m+1} |P_i(e^{j\omega})|^2 |T(e^{j\omega}, \rho_0)|^2}. \quad (232)$$

Nesse caso, se considerarmos a aproximação $S(q, \rho) \approx S_d(q)$, ou seja, $1 - T(q, \rho) \approx 1 - T_d(q)$, obtemos a seguinte aproximação para a estrutura do filtro acima descrito

$$|L(e^{j\omega})|^2 \approx \frac{|T_d(e^{j\omega})|^2 |1 - T_d(e^{j\omega})|^2}{\sum_{i=1}^{m+1} |P_i(e^{j\omega})|^2 |T(e^{j\omega}, \rho_0)|^2}. \quad (233)$$

■

A estrutura do filtro em (233) depende de quantidades conhecidas, exceto pela amplitude da inversa da resposta em frequência da função de sensibilidade complementar. Esta pode ser estimada a partir de uma identificação de um modelo FIR de alta ordem. Observa-se que quanto melhor for a estimativa do modelo, mais próximos serão os mínimos das duas funções custo $J^{\text{CTLS}}(\rho)$ e $J^{\text{MR}}(\rho)$. Na sequência, apresenta-se as estruturas dos filtros considerando primeiramente um controlador com polos fixos e em seguida com polos livres.

Estrutura do filtro considerando um controlador com polos fixos: O caso em que os dados são coletados em malha fechada e a estrutura do controlador é linearmente parametrizada foi abordado na Seção 3.2.1. Portanto, considerando-se a resposta em frequência dos filtros em (149) tem-se

$$\sum_{i=1}^{m+1} |P_i(e^{j\omega})|^2 = \sum_{i=1}^{m+1} |\bar{C}_i(e^{j\omega})|^2 |T_d^{-1}(e^{j\omega}) - 1|^2 + |C(e^{j\omega}, \rho_0)|^2. \quad (234)$$

Uma aproximação para o filtro (233), neste caso, é obtida substituindo-se (234) em (233).

Estrutura do filtro considerando um controlador com polos livres: Neste caso, assume-se que os dados são obtidos de um experimento em malha fechada e o controlador possui polos livres, como abordado na Seção 3.2.2. Com base nisso, a resposta em frequência dos filtros é obtida considerando-se (172), obtendo-se

$$\begin{aligned} \sum_{i=1}^{m+1} |P_i(e^{j\omega})|^2 &= \sum_{i=1}^{n_b} |e^{-j\omega(i-1)}|^2 |T_d^{-1}(e^{j\omega}) - 1|^2 |C_F(e^{j\omega})|^2 \\ &+ \sum_{n_b+1}^m |e^{-j\omega(i-n_b)}|^2 |C(e^{j\omega}, \rho_0)|^2 + |C(e^{j\omega}, \rho_0)|^2. \end{aligned} \quad (235)$$

A partir da expressão acima obtém-se uma aproximação para o filtro em (233), ou seja, substituindo-se (235) em (233).

3.2.4 Exemplo ilustrativo

Neste exemplo, considera-se que a Suposição 2.1 é respeitada e que os dados são coletados durante um experimento em malha fechada. A configuração do experimento é a mesma empregada no exemplo do método VRFT e descrita novamente aqui por conveniência. Considere o seguinte modelo para o processo

$$G(q) = \frac{0,5(q-0,8)}{(q-0,7)(q-0,9)}. \quad (84)$$

Vale salientar que $G(q)$ é utilizada apenas para gerar os dados, assim sendo, o modelo do processo não é empregado no projeto do controlador. O modelo de referência escolhido é dado por

$$T_d(q) = \frac{0,21675(q-0,8)(q-0,8632)}{(q-0,6069)(q^2-1,776q+0,7914)} \quad (85)$$

A partir de (84) e (85) e (16) calcula-se o controlador ideal:

$$C_d(q) = \underbrace{[0,3742 \quad 0,0593]}_{\rho_d^T} \underbrace{\left[1 \quad \frac{q}{q-1}\right]^T}_{\bar{c}(q)} = \frac{0,4335(q-0,8632)}{q-1}, \quad (86)$$

Os dados são coletados de um experimento em malha fechada, com o seguinte controlador operando na malha

$$C(q, \rho_0) = [0,2263 \quad 0,0381] \left[1 \quad \frac{q}{q-1}\right]^T = \frac{0,2644(q-0,8559)}{q-1}. \quad (236)$$

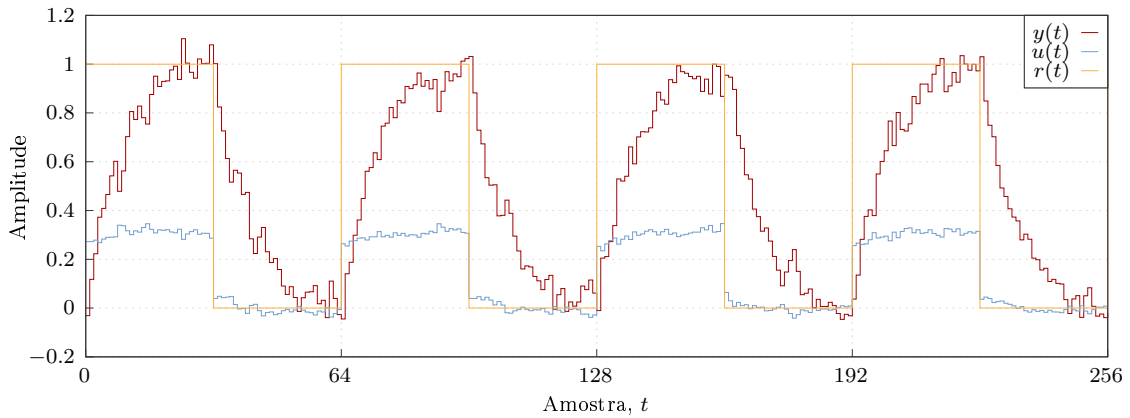
Como sinal de referência utilizou-se oito períodos de uma onda quadrada com níveis 0 e 1, cada período é composto por 64 amostras, totalizando 512 amostras. Os dados são obtidos através de 100 rodadas de Monte Carlo, que são geradas variando-se a realização do ruído colorido que afeta a saída do processo. As realizações de ruído são geradas a partir de realizações de ruído branco com variância $\sigma_v^2 = 0,0016$ filtradas através do seguinte modelo

$$H(q) = \frac{q}{q-0,3}. \quad (88)$$

As primeiras 256 amostras do sinal de referência e uma realização dos sinais de entrada $u(t)$ e saída $y(t)$ coletados são apresentadas na Figura 13.

Os parâmetros foram estimados com a solução CTLS descrita no Teorema 3.1, onde as matrizes Toeplitz P_i foram formadas a partir dos filtros para o caso de dados de malha fechada em (149). Para estimar os parâmetros utilizou-se o algoritmo de otimização iterativo `fminsearch` do Matlab. Para inicializar o algoritmo, para cada realização, utilizou-se a estimativa dos parâmetros com a solução OLS do VRFT (79), com $L(q) = 1$. A Tabela 4 apresenta média e desvio padrão dos parâmetros estimados com o método VRFT e a solução CTLS juntamente com o vetor de parâmetros ideal. Além disso, esta tabela apresenta as estatísticas dos parâmetros estimados com a solução OLS, utilizada como inicialização do problema de otimização do CTLS. Observe que a estimativa com CTLS praticamente não apresenta polarização, ao contrário da estimativa obtida com OLS.

Figura 13 — Uma realização dos dados do experimento em malha fechada do exemplo da Seção 3.2.4.



Fonte: elaborado pela autora.

Tabela 4 — Média e desvio padrão dos parâmetros estimados com o VRFT e as soluções CTLS e OLS do exemplo da Seção 3.2.4.

ρ	ρ_d	Estimativas	
		CTLS	OLS
K_p	0,3742	0,3734 ($\pm 0,0053$)	0,1643 ($\pm 0,0074$)
K_i	0,0593	0,0593 ($\pm 0,0006$)	0,0590 ($\pm 0,0003$)

3.3 Data Least Squares

O problema DLS foi formulado considerando-se que apenas a matriz regressora Φ é afetada por ruído. Esta matriz é conhecida na literatura como a matriz de *dados*, o que confere o nome ao problema. DeGroat e Dowling (1993) apresentaram o problema e uma solução fechada para o mesmo baseada no trabalho de Abatzoglou e Mendel (1987). Ao longo dos anos alguns trabalhos estenderam ou aplicaram a solução DLS. Por exemplo, Lim (2008) aplicou a solução DLS para o ajuste dos parâmetros de uma rede neural e propôs uma solução recursiva para o problema. Lim (2007) aplicou a solução DLS em uma rede neural, cujo objetivo era o ajuste de um equalizador de canal.

O problema DLS é formulado como

$$\min_{\Delta_\Phi, \rho} \|\Delta_\Phi\|_F^2 \quad (237)$$

$$\text{sujeito a } (\Phi - \Delta_\Phi)\rho = u, \quad (238)$$

obtido fixando-se $\delta_u = 0$ em (117) e (118). Quando o controlador ideal pertence à classe de controle, ou seja, a Suposição 2.1 é respeitada, e os sinais não são afetados por ruído tem-se que

$$\min_{\Delta_\Phi, \rho} \|\Delta_\Phi\|_F^2 = 0$$

e ρ_d é o único vetor de parâmetros que satisfaz (238). Portanto, tanto o problema de otimização formulado pelo DLS quanto a função custo $J^{\text{MR}}(\rho)$ são minimizados com o vetor de parâmetros ideal ρ_d .

Observe que esse problema é um caso particular do problema proposto por Abatzoglou e Mendel (1987), onde considera-se que uma ou mais colunas da matriz Z , (128), são conhecidas exatamente, ou seja, não são afetadas por ruído. Como exemplo, tem-se o caso em que apenas Φ é afetada por ruído, ou seja, o problema DLS. No método VRFT, este caso ocorre quando os dados são coletados durante um experimento em malha aberta, pois, como explicado anteriormente, o sinal de entrada do controlador a ser identificado é afetado por ruído e não sua saída. Na sequência, apresenta-se através do Teorema 3.4 a solução proposta por Abatzoglou e Mendel (1987), com as devidas adaptações para o método VRFT. Além disso, apresenta-se através do Teorema 3.5 a solução proposta por DeGroat e Dowling (1993), também com as adaptações para a estimativa dos parâmetros com o método VRFT.

Teorema 3.4 (de Abatzoglou e Mendel (1987)). *Suponha que a matriz Δ_Z possui colunas linearmente dependentes, ou seja, existe uma matriz X tal que $\Delta_Z X = 0$, então a solução do problema CTLS é obtida através de*

$$\begin{bmatrix} \rho \\ -1 \end{bmatrix} = \beta = -\frac{\bar{v}}{\bar{v}_{m+1}},$$

onde \bar{v} é o autovetor associado ao maior autovalor de $(Z^T Z)^{-1}[I - X(X^T X)^{-1}X^T]$ e \bar{v}_{m+1} é o último elemento de \bar{v} .

A prova desse teorema é apresentada abaixo e segue o desenvolvimento do artigo em que foi proposta (ABATZOGLOU; MENDEL, 1987).

Prova. Considerando-se que existe uma matriz X tal que $\Delta_Z X = 0$, o problema (131) a (133) é reformulado como

$$\min_{v, \rho} \|\Delta_Z\|_F^2 \quad (239)$$

$$\text{sujeito a } (Z - \Delta_Z)\beta = 0, \quad (240)$$

$$\text{e } \Delta_Z X = 0. \quad (241)$$

O Lagrangeano para o problema é calculado como

$$\mathcal{L}(\Delta_Z, \Lambda, \lambda) = \text{tr}\{\Delta_Z^T \Delta_Z\} - \lambda^T \Delta_Z \beta - \beta^T \Delta_Z^T \lambda + \text{tr}\{\Lambda^T \Delta_Z X\} + \text{tr}\{X^T \Delta_Z^T \Lambda\}$$

O gradiente de $\mathcal{L}(\Delta_Z, \Lambda, \lambda)$ com respeito a Δ_Z^T é obtida da seguinte maneira

$$\frac{\partial \mathcal{L}(\Delta_Z, \Lambda, \lambda)}{\partial \Delta_Z^T} = \Delta_Z - \lambda \beta^T + \Lambda X^T,$$

igualando a equação acima a zero e isolando Δ_Z tem-se

$$\Delta_Z = \lambda \beta^T - \Lambda X^T. \quad (242)$$

Multiplicando-se esta equação à direita por X e lembrando da restrição $\Delta_Z X = 0$, chega-se na seguinte expressão para Λ :

$$\Lambda = \lambda \beta^T X (X^T X)^{-1} \quad (243)$$

Agora, substitui-se (242) na restrição $(Z - \Delta_Z)\beta = 0$ e depois (243) obtém-se

$$(Z - \lambda\beta^T + \Lambda X^T)\beta = 0 \quad (244)$$

$$Z\beta - \lambda\beta^T\beta + \lambda\beta^T X (X^T X)^{-1} X^T \beta = 0, \quad (245)$$

e resolvendo para λ :

$$\lambda = \frac{Z\beta}{\|\beta\|_2^2 - \beta^T Q \beta}, \quad (246)$$

onde $Q = X(X^T X)^{-1} X^T$.

A expressão para Λ em (243) é reescrita substitui-se a equação acima e considerando que $\|\beta\|_2^2 = \beta^T I \beta$, chegando-se a

$$\Lambda = \frac{Z\beta\beta^T X (X^T X)^{-1}}{\beta^T (I - Q)\beta}. \quad (247)$$

Substituindo-se (247) e (246) em (242) tem-se

$$\Delta_Z = \frac{Z\beta\beta^T (I - Q)}{\beta^T (I - Q)\beta}$$

a partir da equação acima e sabendo-se que $(I - Q)$ é idempotente obtém-se uma expressão para $\Delta_Z \Delta_Z^T$ da seguinte forma

$$\Delta_Z \Delta_Z^T = \frac{Z\beta\beta^T Z^T}{\beta^T (I - Q)\beta}.$$

Finalmente, sabendo-se que $\text{tr}\{\Delta_Z \Delta_Z^T\} = \text{tr}\{\Delta_Z^T \Delta_Z\}$ e aplicando-se a propriedade cíclica do traço, tem-se

$$\text{tr}\{\Delta_Z \Delta_Z^T\} = \text{tr}\left\{ \frac{\beta^T Z^T Z \beta}{\beta^T (I - Q)\beta} \right\} = \frac{\beta^T Z^T Z \beta}{\beta^T (I - Q)\beta},$$

pois $(\beta^T Z^T Z \beta)/(\beta^T (I - Q)\beta)$ é um escalar. A equação acima é uma forma Hermitiana transposta quadrática em β e sua solução mínima é caracterizada pelo autovetor \bar{v} que corresponde ao maior autovalor de $(Z^T Z)^{-1}(I - Q)$ (ABATZOGLOU; MENDEL, 1987), ou seja,

$$\beta = \begin{bmatrix} \rho \\ -1 \end{bmatrix} = -\frac{\bar{v}}{\bar{v}_{m+1}}, \quad (248)$$

onde \bar{v}_{m+1} é o último elemento de \bar{v} . ■

Teorema 3.5 (de DeGroat e Dowling (1993)). *A solução para o problema DLS é dada por*

$$\hat{\rho} = \frac{u^T u}{(u^T \Phi u)} \underline{v}, \quad (249)$$

onde \underline{v} é o autovetor associado ao menor autovalor de $P_b \Phi$ e a matriz $P_b \in \mathbb{R}^{N \times N}$ é definida como $P_b = \left(I - u(u^T u)^{-1} u^T \right)$.

Prova. A prova deste teorema é apresentada por DeGroat e Dowling (1993). ■

Observe que a solução apresentada no Teorema 3.4 é mais abrangente do que a solução do Teorema 3.5. O motivo consiste no fato de que a primeira foi formulada considerando que uma ou mais colunas da matriz Δ_Z são linearmente dependentes. Isso é garantido, por exemplo, quando uma ou mais colunas de Z são conhecidas exatamente, ou seja, não são corrompidas por ruído. Em contrapartida, a solução do Teorema 3.5 (nomeada como DLS) é um caso particular do problema do Teorema 3.4, pois considera especificamente que a última coluna de Δ_Z é linearmente dependente. Essa restrição é atendida no caso do método VRFT quando os dados são coletados de um experimento em malha aberta, ou seja, a matriz regressora Φ é afetada por ruído, enquanto o vetor de saída u é conhecido exatamente. Por conta disso, no caso do VRFT com dados de malha aberta, as soluções são equivalentes, i.e., ambas podem ser empregadas na estimativa dos parâmetros do controlador. Além disso, veja que as soluções nos Teoremas 3.4 e 3.5 podem ser aplicadas tanto para controladores linearmente parametrizados com polos fixos quanto para controladores com polos livres.

Finalmente, nota-se que para o problema DLS o caso em que o controlador ideal não pertence à classe de controle não foi abordado. Portanto, o desenvolvimento de um possível filtro para este caso é um tema a ser pesquisado. As alternativas sugeridas para lidar com o caso em que a Suposição 2.1 é violada consistem em utilizar um filtro com a estrutura do filtro original do método VRFT, (77), ou utilizar o filtro do CTLS, Teorema 3.2, de acordo com a estrutura escolhida para o controlador.

3.3.1 Exemplo ilustrativo

Este exemplo tem como objetivo ilustrar a aplicação da solução DLS no método VRFT. A configuração do experimento é igual a empregada no exemplo motivacional descrito na Seção 2.4.1.3.

A partir dos dados de entrada e saída, do modelo de referência e da estrutura de controle, estima-se os parâmetros do controlador com a solução DLS como descrito no Teorema 3.4. A Tabela 5 apresenta as estatísticas dos parâmetros estimados com a solução DLS juntamente com os parâmetros do controlador ideal. Observe que a média da estimativa obtida com DLS é bem próxima do vetor de parâmetros ideal.

Tabela 5 — Média e desvio padrão dos parâmetros estimados com o VRFT e a solução DLS do exemplo da Seção 3.3.1.

ρ	ρ_d	DLS
K_p	0,3742	0,3907 (\pm 0,0052)
K_i	0,0593	0,0587 (\pm 0,0002)

3.4 Nota

Vale salientar que, embora no presente trabalho a estabilidade em malha fechada não seja validada durante a etapa de otimização, esse assunto já foi investigado por Van Heusden, Karimi e Söderström (2011) para o método VRFT. Naquele trabalho, o teorema do pequeno ganho é empregado para determinar condições suficientes para garantir a estabilidade em malha fechada, que são adicionadas como restrições no problema do VRFT. Apesar dessas restrições poderem também ser adicionadas ao problema de otimização do CTLS, esse assunto excede o escopo do presente trabalho.

3.5 Considerações finais

Nesse capítulo, foram apresentadas duas soluções de mínimos quadrados aplicadas ao método VRFT com o objetivo de lidar com o ruído presente nos dados coletados. Tais soluções são conhecidas na literatura como CTLS e DLS. Além disso, o método VRFT foi formulado considerando-se controladores com estrutura semelhante ao modelo ARX. Essa estrutura de controle fornece ao projetista mais flexibilidade para a escolha do controlador do que as estruturas comumente empregadas na literatura.

Quanto à solução CTLS, primeiramente, considerou-se o caso em que o controlador ideal pertence à classe de controle e desenvolveu-se as modificações necessárias para permitir sua aplicação ao método VRFT. Na sequência, foram desenvolvidas as modificações para o caso em que a Suposição 2.1 é violada. Em contrapartida, a solução DLS foi formulada considerando-se que a Suposição 2.1 é respeitada.

Como continuação desse tópico, tem-se a extensão da solução DLS para o caso em que o controlador ideal não pertence à classe de controle. Outro tópico a ser investigado é a avaliação da consistência das estimativas obtidas com as soluções CTLS e DLS. Abatzoglou, Mendel e Harada (1991) mostraram que quando as colunas de Δ_Z são afetadas por ruído branco a solução obtida com CTLS é próxima do limite de Cramér-Rao, desde que exista uma boa relação sinal-ruído. Contudo, a análise estatística das propriedades dessas soluções é um tema a ser investigado. Além disso, a extensão da solução CTLS para o VRFT no caso multivariável é outro tema a ser pesquisado. Finalmente, outro tópico de pesquisa consiste no emprego da solução *eXtended Least Squares* (XLS) (YEREDOR, 2001) no método VRFT. Essa solução visa compensar não apenas a contribuição de ruído mas também o erro do modelo.

No Capítulo 5, mais adiante, serão comparados os resultados obtidos com a solução CTLS com aqueles obtidos com o método OCI. Além disso, serão comparados os resultados encontrados com as estimativas dos parâmetros com OLS e DLS juntamente com os critérios de seleção de subconjuntos mais informativos de dados.

4 CRITÉRIOS DE SELEÇÃO DE SUBCONJUNTOS INFORMATIVOS DE DADOS

4.1 Introdução

Em métodos de controle baseados em dados, o projeto do controlador é realizado diretamente a partir dos dados coletados do processo, sem passar pela etapa de modelagem do mesmo, e a qualidade da estimativa dos parâmetros depende da informatividade dos sinais coletados. Visto que os dados armazenados ou coletados durante a operação normal do processo são usualmente pouco informativos para o projeto dos parâmetros do controlador, é uma prática comum realizar um experimento específico para coleta de dados, no qual um sinal suficientemente rico é aplicado na entrada do processo. Contudo, a execução desse experimento pode ser uma tarefa custosa, indesejável e até mesmo impossível em alguns casos (BITTENCOURT et al., 2015; SHARDT; HUANG, 2013).

Com isto em mente, chega-se à seguinte questão: *não seria possível utilizar dados de operação para realizar o projeto do controlador?* Em processos industriais, é comum que os dados coletados durante a operação normal do processo sejam armazenados em um banco de dados, uma vez que tais dados são utilizados para monitoramento e detecção de faltas (MACGREGOR; CINAR, 2012). Além disso, mesmo que estes dados não sejam armazenados, eles podem ser facilmente coletados, visto que não é necessário interromper a execução normal do processo. Contudo, simplesmente empregar este tipo de dado em um método baseado em dados pode levar a um ajuste inapropriado dos parâmetros do controlador. (Um exemplo do efeito da utilização de dados pouco informativos na estimativa dos parâmetros do controlador foi apresentado no Capítulo 2.) Por conta disso, outra questão emerge: *seria possível extrair os subconjuntos mais informativos deste tipo de dado?* Embora pouco abordado em controle baseado em dados, este tema já foi objeto de pesquisa na área de identificação de sistemas. Gevers et al. (2009) mostraram que a informatividade é determinada pela regularidade da matriz de informação de Fischer, e depende basicamente do espectro e magnitude dos sinais externos aplicados ao processo. Carrette et al. (1996) propuseram um critério para descartar dados que possuem forte contribuição de ruído. Este critério baseia-se na evolução ao longo do tempo da primeira diferença do menor valor singular da matriz regressora. Além disso, os autores mostraram que a degradação da qualidade da estimativa utilizando todo o conjunto de dados disponível deve-se ao aumento da polarização da mesma.

Outra métrica empregada na literatura é baseada na evolução ao longo do tempo do número de condicionamento da matriz de informação. Peretzki et al. (2011) desenvolveram um algoritmo para extrair os intervalos mais informativos em um banco de dados. Aquele algoritmo identifica degraus no sinal de entrada, aplica o critério do número de

condicionamento para selecionar os subconjuntos relevantes e os utiliza para identificar um modelo de Laguerre para o processo. Uma extensão do trabalho de Peretzki et al. (2011) foi proposta por Bittencourt et al. (2015), onde os autores adicionaram um fator de esquecimento à estimativa do modelo de Laguerre e também estimaram um modelo para o ruído. Shardt e Shah (2014) aplicaram o algoritmo proposto por Peretzki et al. (2011) para segmentação dos dados, no qual o algoritmo foi utilizado para detectar mudanças no ponto de operação do processo.

Esse mesmo critério também foi utilizado por Shardt e Huang (2013), onde os autores investigaram seu comportamento variando-se os valores de limiar. Arengas e Kroll (2017b) propuseram um algoritmo para detecção de mudanças no sinal de entrada e empregam o critério do número de condicionamento da matriz de informação para selecionar os subconjuntos mais informativos de dados. Shardt e Huang (2017) investigaram a aplicação deste critério considerando dados de malha fechada e, além disso, foram investigadas as condições para identificabilidade do modelo com este tipo de dado. Outros trabalhos abordam o caso multivariável. Por exemplo, Arengas e Kroll (2017a) empregaram este critério de seleção para o caso multivariável considerando que o processo pode ser modelado através de uma estrutura ARX. Arengas e Kroll (2019) extraem os subconjuntos muito informativos com o critério do número de condicionamento da matriz de informação e estimam os parâmetros do processo por meio de um método de variáveis instrumentais recursivo.

Além destes critérios, outras métricas foram propostas na literatura com o propósito de selecionar subconjuntos muito informativos de dados. No trabalho proposto por Holcomb e Bitmead (2017) os subconjuntos muito informativos são determinados utilizando-se um teste de posto e os mesmos são empregados para identificação do processo em espaço de estados. Um outro método foi proposto por Wang et al. (2018), onde a seleção de dados muito informativos baseia-se na detecção de mudanças na magnitude dos sinais de entrada e saída. O caso em que o processo possui múltiplas entradas e uma saída é avaliado por Arengas e Kroll (2018). Naquele trabalho, foi proposto um critério de seleção baseado na evolução, ao longo do tempo, de diferentes médias da matriz de informação. Os resultados foram comparados com os encontrados com o método baseado no menor valor singular da matriz regressora proposto por Carrette et al. (1996).

Como apresentado acima, a seleção de subconjuntos muito informativos de dados é um tema já pesquisado na área de identificação de sistemas, mas muito pouco abordado na área de controle baseado em dados, como pode ser visto através das poucas contribuições presentes na literatura. Radrizzani et al. (2020) selecionam os subconjuntos muito informativos de dados coletados durante a operação normal do processo para ajustar um controlador de um processo real. Contudo, a seleção dos subconjuntos é realizada manualmente, onde o projetista seleciona os intervalos em que o sinal de saída está em regime transitório. Os subconjuntos selecionados são empregados no ajuste dos parâmetros do controlador com o método VRFT. Por conta disso, um dos temas do presente trabalho consiste em adaptar para o método VRFT dois critérios de seleção de subconjuntos muito informativos de dados, originalmente desenvolvidos para identificação de sistemas. Parte desses resultados foram publicados em dois artigos de congresso. A adaptação desses critérios considerando o caso em que a Suposição 2.1 é respeitada é apresentado em (SILVA GARCIA; BAZANELLA, 2019). Já em (SILVA GARCIA; BAZANELLA, 2020b) os mesmos critérios são aplicados assumindo-se que o controlador ideal não pertence à classe de controle. Nas próximas seções são apresentados estes dois critérios de seleção aplicados ao método VRFT. Primeiramente, apresenta-se o critério baseado no menor valor

singular da matriz regressora e, na sequência, apresenta-se o critério baseado no número de condicionamento da matriz de informação.

4.2 Critério baseado no menor valor singular

O objetivo desta seção é apresentar o critério baseado no menor valor singular (do inglês *Smallest Singular Value* – SSV) da matriz regressora, proposto por Carrette et al. (1996), com as modificações necessárias para sua aplicação no método VRFT. Assim sendo, na sequência é apresentado o critério (aqui abreviado para critério SSV) e ao final da seção é apresentado um exemplo ilustrando a aplicação do mesmo.

4.2.1 Estimativa dos parâmetros

Considere o problema de identificação dos parâmetros de um controlador linearmente parametrizado (Definição 2.4) com o método VRFT. O sinal de saída do controlador a ser identificado $u_L(t)$ e um preditor para o mesmo $\hat{u}_L(t)$ são dados por:

$$u_L(t) = \varphi_L^T(t)\rho_d + \mu(t), \quad (250)$$

$$\hat{u}_L(t) = \varphi_L^T(t)\hat{\rho}, \quad (251)$$

onde $\mu(t)$ é o erro do modelo, $\hat{\rho}$ é o vetor de parâmetros estimado, ρ_d é o vetor de parâmetros ideal e o vetor regressor $\varphi_L(t)$ é definido em (65).

Assumindo que o controlador ideal pertence à classe de controle (Suposição 2.1), reescreve-se o sinal de saída como

$$u_L(t) = \rho_d^T \bar{C}(q) \left[T_d^{-1}(q) - 1 \right] L(q) y(t) \quad (252)$$

$$= C_d(q) \left[T_d^{-1}(q) - 1 \right] L(q) G(q) u(t) \quad (253)$$

$$= F(q) G(q) u(t), \quad (254)$$

onde (252) é calculada substituindo-se (65) em (250), (253) é obtida assumindo-se que os dados são coletados de um experimento em malha aberta e que a saída não é corrompida pelo ruído, ou seja, substituindo-se $y(t) = G(q)u(t)$ em (252), enquanto (254) é obtida definindo-se

$$F(q) = C_d(q) \left[T_d^{-1}(q) - 1 \right] L(q). \quad (255)$$

Suponha agora que os dados são coletados de um experimento em malha aberta, que o controlador ideal pertence à classe de controle e é linearmente parametrizado, mas que o sinal de saída é afetado por ruído. Neste caso, (250) é reescrita como

$$u_L(t) = \rho_d^T \bar{C}(q) \left[T_d^{-1}(q) - 1 \right] L(q) [G(q)u(t) + H(q)v(t)] + \mu(t) \quad (256)$$

$$= F(q)G(q)u(t) + F(q)H(q)v(t) + \mu(t), \quad (257)$$

onde (256) é obtida substituindo-se (65) e (1) em (250) e (257) é encontrada definindo-se $F(q)$ como em (255).

Uma expressão para o erro do modelo é obtida substituindo-se (254) em (257) e então resolvendo-se para $\mu(t)$, como desenvolvido abaixo

$$u_L(t) = u_L(t) + F(q)H(q)v(t) + \mu(t), \quad (258)$$

$$\mu(t) = -F(q)H(q)v(t). \quad (259)$$

Em forma vetorial, (250) e (251) são escritas como:

$$u_L = \Phi_L \rho_d + \mu, \quad (260)$$

$$\hat{u}_L = \Phi_L \hat{\rho}, \quad (261)$$

onde a matriz regressora $\Phi_L \in \mathbb{R}^{N \times m}$, o vetor de saída $u_L \in \mathbb{R}^N$, o vetor da saída predita $\hat{u}_L \in \mathbb{R}^N$, e o vetor do ruído não modelado $\mu \in \mathbb{R}^N$ são dados por

$$\Phi_L = [\varphi_L(1) \quad \varphi_L(2) \quad \dots \quad \varphi_L(N)]^T, \quad (262)$$

$$u_L = [u_L(1) \quad u_L(2) \quad \dots \quad u_L(N)]^T, \quad (263)$$

$$\hat{u}_L = [\hat{u}_L(1) \quad \hat{u}_L(2) \quad \dots \quad \hat{u}_L(N)]^T, \quad (264)$$

$$\mu = [\mu(1) \quad \mu(2) \quad \dots \quad \mu(N)]^T. \quad (265)$$

Considere agora a solução OLS (116) para o problema. Assumindo que Φ_L tem posto cheio em coluna, a solução pode ser escrita em termos da pseudoinversa, ou seja,

$$\hat{\rho} = \Phi^+ u_L. \quad (266)$$

onde Φ^+ é a pseudoinversa de Φ_L , definida como

$$\Phi^+ = (\Phi_L^T \Phi_L)^{-1} \Phi_L^T. \quad (267)$$

A matriz regressora, decomposta utilizando decomposição em valores singulares é dada por

$$\Phi_L = U \Sigma V^T, \quad (268)$$

onde a matriz dos valores singulares é $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_m)$ e $\sigma_i^2 = \lambda_i(\Phi_L^T \Phi_L)$, enquanto $U \in \mathbb{R}^{N \times m}$ e $V \in \mathbb{R}^{m \times m}$ são, respectivamente, as matrizes singulares à esquerda e à direita de Φ_L . Além disso, tem-se que

$$U^T U = I_m, \quad (269)$$

$$V V^T = I_m, \quad (270)$$

onde I_m é a matriz identidade de dimensão m . Veja que de (268) chega-se à seguinte relação

$$\Phi_L V = U \Sigma. \quad (271)$$

É bem conhecido que a pseudoinversa de Φ_L pode ser obtida através da decomposição em valores singulares, resultando na seguinte expressão

$$\Phi^+ = V \Sigma^{-1} U^T. \quad (272)$$

A partir da decomposição em valores singulares de Φ_L pode-se reescrever o vetor de saída u_L como:

$$u_L = U \Sigma V^T \rho_d + \mu \quad (273)$$

$$= \Phi_L V V^T \rho_d + \mu \quad (274)$$

$$= \Phi_V \rho_{dV} + \mu, \quad (275)$$

onde (273) é obtida substituindo-se (268) em (260), (274) é calculada substituindo-se (271) em (273), enquanto (275) é encontrada considerando-se $\Phi_V = \Phi_L V$ e $\rho_{dV} = V^T \rho_d$. Seguindo o mesmo raciocínio, obtém-se uma nova expressão para o vetor da saída predito em (261):

$$\hat{u}_L = \Phi_V \hat{\rho}_V, \quad (276)$$

onde $\hat{\rho}_V = V^T \hat{\rho}$.

Agora, uma nova expressão para os parâmetros estimados pode ser obtida como

$$\hat{\rho} = V \Sigma^{-1} U^T (U \Sigma V^T \rho_d + \mu) \quad (277)$$

$$= \rho_d + V \Sigma^{-1} U^T \mu, \quad (278)$$

onde calcula-se (277) substituindo-se (272) e (275) em (266) e (278) é obtida lembrando de (269), (270) e $\Sigma^{-1} \Sigma = I_m$. A estimativa também pode ser obtida em termos dos autoparâmetros, ou seja,

$$\hat{\rho}_V = V^T \hat{\rho} \quad (279)$$

$$= \rho_{dV} + \Sigma^{-1} U^T \mu, \quad (280)$$

onde (280) foi obtida substituindo-se (278) em (279) e usando novamente (270). Mais adiante no texto, essa expressão será utilizada para obter os dois primeiros momentos de probabilidade dos parâmetros originais do controlador.

4.2.2 Análise estatística da matriz regressora

Considere que a matriz regressora é escrita em função do tempo, ou seja,

$$\Phi_L(t) = [\varphi_L(1) \quad \varphi_L(2) \quad \dots \quad \varphi_L(t)]^T \quad (281)$$

onde o inteiro $t \leq N$ é o índice da amostra. Observe que a estimativa dos parâmetros $\hat{\rho}(t)$ também depende do número de amostras. Desta forma, considerando um subconjunto de t amostras, tem-se

$$\sigma_i^2(\Phi_L(t)) = \lambda_i(\Phi_L^T(t) \Phi_L(t)), \quad i = 1, \dots, m, \quad (282)$$

onde $\sigma_i^2(\Phi_L(t))$ é o quadrado do i -ésimo valor singular de $\Phi_L(t)$ calculado até a t -ésima amostra dos dados. Nota-se que $\sigma_i^2(\Phi_L(t))$ é uma quantidade estocástica pois a matriz regressora $\Phi_L(t)$ é construída a partir do sinal de saída coletado $y(t)$ que é afetado por ruído. Além disso, $\sigma_i^2(\Phi_L(t))$ é monotonicamente não decrescente (teorema de Weyl (HORN; JOHNSON, 2013)). Considerando que $\Phi_L(t)$ é decomposta em valores singulares como em (268) tem-se que

$$\begin{aligned} \Sigma^2(t) &= V^T(t) (\Phi_L^T(t) \Phi_L(t)) V(t) \\ &= \text{diag}(\sigma_i^2(\Phi_L(t))). \end{aligned}$$

Similarmente, pode-se definir os autovalores do valor esperado de $\Phi_L^T(t) \Phi_L(t)$ como

$$s_i^2(t) = \lambda_i(E\{\Phi_L^T(t) \Phi_L(t)\}), \quad i = 1, \dots, m. \quad (283)$$

Neste caso, também pode-se definir uma matriz com os autovalores na diagonal, como

$$\text{diag}(s_i^2(t)) = \mathcal{V}^T(t) E\{\Phi_L^T(t) \Phi_L(t)\} \mathcal{V}(t) \quad (284)$$

$$= E\{\Phi_V^T(t) \Phi_V(t)\}, \quad (285)$$

onde $\mathcal{V}(t)$ é a matriz de autovetores correspondente e $\Phi_{\mathcal{V}} = \Phi_L(t)\mathcal{V}(t)$.

Na sequência, deseja-se encontrar uma expressão para os elementos da diagonal principal de (285). Para tanto, considere $\Phi_{\mathcal{V}}(N)$ escrito como

$$\Phi_{\mathcal{V}}(N) = \Phi_L(N)\mathcal{V}(N) = \begin{bmatrix} \varphi_L^T(1) \\ \varphi_L^T(2) \\ \vdots \\ \varphi_L^T(N) \end{bmatrix} [\mathcal{V}^{(1)} \quad \mathcal{V}^{(2)} \quad \dots \quad \mathcal{V}^{(m)}], \quad (286)$$

onde $\mathcal{V}^{(i)}$ é a i -ésima coluna de $\mathcal{V}(N)$. Cada elemento da i -ésima coluna da t -ésima linha de (286) é dado por

$$\begin{aligned} \Phi_{\mathcal{V}}^{(i)}(t) &= \varphi_L^T(t)\mathcal{V}^{(i)} \\ &= \mathcal{V}^{(i)T}\varphi_L(t). \end{aligned} \quad (287)$$

A equação acima é reescrita assumindo-se que o controlador é linearmente parametrizado e que os dados são de um experimento em malha aberta, como

$$\mathcal{V}^{(i)T}\varphi_L(t) = \mathcal{V}^{(i)T}L(q)\bar{C}(q) \left[T_d^{-1}(q) - 1 \right] y(t) \quad (288)$$

$$= \underbrace{\mathcal{V}^{(i)T}L(q)\bar{C}(q) \left[T_d^{-1}(q) - 1 \right]}_{X(q)} [G(q)u(t) + H(q)v(t)], \quad (289)$$

onde (288) é obtida substituindo-se $\varphi_L(t)$ definido em (65) e (289) é calculada considerando-se que os dados são coletados como descrito na Definição 2.2. O termo $X(q)$ pode ser rearranjado como

$$\begin{aligned} X(q) &= \mathcal{V}^{(i)T}L(q)\bar{C}(q) \left[T_d^{-1}(q) - 1 \right] \\ &= \left[\mathcal{V}^{(i,1)}\bar{C}_1(q) + \dots + \mathcal{V}^{(i,m)}\bar{C}_m(q) \right] L(q) \left[T_d^{-1}(q) - 1 \right] \end{aligned} \quad (290)$$

$$= \mathcal{V}^{(i,1)}R_1(q) + \dots + \mathcal{V}^{(i,m)}R_m(q), \quad (291)$$

onde $\mathcal{V}^{(i,j)}$ é o j -ésimo elemento de $\mathcal{V}^{(i)}$, enquanto que $\bar{C}_i(q)$ em (290) é o i -ésimo componente do vetor $\bar{C}(q)$ e

$$R_i(q) = \bar{C}_i(q)L(q)[T_d^{-1}(q) - 1]. \quad (292)$$

Substituindo-se (291) em (289) chega-se a

$$\begin{aligned} \mathcal{V}^{(i)T}\varphi_L(t) &= \left[\mathcal{V}^{(i,1)}R_1(q) + \dots + \mathcal{V}^{(i,m)}R_m(q) \right] G(q)u(t) \\ &\quad + \left[\mathcal{V}^{(i,1)}R_1(q) + \dots + \mathcal{V}^{(i,m)}R_m(q) \right] H(q)v(t) \\ &= M_{Gi}(q)u(t) + M_{Hi}(q)v(t), \end{aligned} \quad (293)$$

onde

$$M_{Gi}(q) = \left[\mathcal{V}^{(i,1)}R_1(q) + \dots + \mathcal{V}^{(i,m)}R_m(q) \right] G(q), \quad (294)$$

$$M_{Hi}(q) = \left[\mathcal{V}^{(i,1)}R_1(q) + \dots + \mathcal{V}^{(i,m)}R_m(q) \right] H(q). \quad (295)$$

Observe que, assumindo-se $G(q)$ e $H(q)$ estáveis e causais, a estabilidade e causalidade dos filtros $M_{Gi}(q)$ e $M_{Hi}(q)$ é determinada pela escolha de $L(q)$ em (292). No caso em que o controlador ideal pertence à classe de controle, propõe-se utilizar $L(q) = T_d(q)$. Com esta escolha, a estrutura do filtro em (292) é dada por $R_i(q) = \bar{C}_i(q)[1 - T_d(q)]$. Assim, $R_i(q)$ é estável e causal, pois assume-se que $T_d(q)$ é causal e estritamente própria. Quando o controlador ideal não pertence à classe de controle utiliza-se o filtro ótimo do VRFT, cuja estrutura é definida em (77). Neste caso, ao escolher $L(q) = T_d(q)(1 - T_d(q))$ também é garantido que a restrição seja atendida, resultando em $R_i(q) = \bar{C}_i(q)[1 - T_d(q)]^2$. Por conta disso, o critério SSV não requer modificações para ser aplicado no caso em que o controlador ideal não pertence à classe de controle.

Na sequência, escreve-se (293) em forma vetorial como

$$\Phi_{\mathcal{V}}^{(i)} = M_{Gi}u + M_{Hi}v, \quad (296)$$

onde M_{Gi} e M_{Hi} são matrizes Toeplitz que representam os filtros $M_{Gi}(q)$ e $M_{Hi}(q)$ em (294) e (295), respectivamente. Estas matrizes possuem as seguintes estruturas

$$M_{Gi} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ m_{Gi}(1) & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ m_{Gi}(N-1) & m_{Gi}(N-2) & \dots & m_{Gi}(1) & 0 \end{bmatrix}, \quad (297)$$

$$M_{Hi} = \begin{bmatrix} m_{Hi}(0) & 0 & \dots & 0 & 0 \\ m_{Hi}(1) & m_{Hi}(0) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ m_{Hi}(N-1) & m_{Hi}(N-2) & \dots & m_{Hi}(1) & m_{Hi}(0) \end{bmatrix}. \quad (298)$$

Com base nisso, foi possível chegar a uma expressão para $\Phi_{\mathcal{V}}^{(i)}$. Considere agora $\Phi_{\mathcal{V}}^{(i)\text{T}} \Phi_{\mathcal{V}}^{(j)}$, escrito substituindo-se (296), como desenvolvido abaixo

$$\Phi_{\mathcal{V}}^{(i)\text{T}} \Phi_{\mathcal{V}}^{(j)} = [M_{Gi}u + M_{Hi}v]^{\text{T}} [M_{Gj}u + M_{Hj}v] \quad (299)$$

$$= u^{\text{T}} M_{Gi}^{\text{T}} M_{Gj}u + v^{\text{T}} M_{Hi}^{\text{T}} M_{Hj}v + u^{\text{T}} M_{Gi}^{\text{T}} M_{Hj}v + v^{\text{T}} M_{Hi}^{\text{T}} M_{Gj}u. \quad (300)$$

O valor esperado de $\Phi_{\mathcal{V}}^{(i)\text{T}} \Phi_{\mathcal{V}}^{(j)}$ é dado por

$$E\{\Phi_{\mathcal{V}}^{(i)\text{T}} \Phi_{\mathcal{V}}^{(j)}\} = u^{\text{T}} M_{Gi}^{\text{T}} M_{Gj}u + E\{v^{\text{T}} M_{Hi}^{\text{T}} M_{Hj}v\} \quad (301)$$

$$= u^{\text{T}} M_{Gi}^{\text{T}} M_{Gj}u + E\{\text{tr}(v^{\text{T}} M_{Hi}^{\text{T}} M_{Hj}v)\} \quad (302)$$

$$= u^{\text{T}} M_{Gi}^{\text{T}} M_{Gj}u + \text{tr}(M_{Hi}^{\text{T}} M_{Hj} E\{vv^{\text{T}}\}) \quad (303)$$

$$= u^{\text{T}} M_{Gi}^{\text{T}} M_{Gj}u + \eta_{ij} t \sigma_v^2, \quad (304)$$

lembrando que $u(t)$ é um sinal determinístico e que $u(t)$ e $v(t)$ são decorrelacionados obtém-se (301) diretamente de (300). Já (302) é obtida atentando-se para o fato de que $v^{\text{T}} M_{Hi}^{\text{T}} M_{Hj}v$ é um escalar. Calcula-se (303) através da propriedade cíclica do traço e lembrando que $E\{\cdot\}$ e $\text{tr}(\cdot)$ são operadores lineares. Finalmente, chega-se a (304) definindo-se a seguinte variável relacionada ao ganho de potência do ruído:

$$\eta_{ij} = \text{tr}(M_{Hi}^{\text{T}} M_{Hj}) / t, \quad (305)$$

e lembrando que $v(t)$ é um ruído branco com média zero e variância σ_v^2 , ou seja,

$$E\{vv^T\} = \sigma_v^2 I. \quad (306)$$

De (285), tem-se que os elementos da diagonal principal de (304) são dados por

$$s_i^2(t) = \|M_{Gi}u\|_2^2 + \eta_{ii}t\sigma_v^2. \quad (307)$$

Perceba que a energia em cada subespaço, $s_i^2(t)$, é formada por $\|M_{Gi}u\|_2^2$, que depende da energia da entrada naquele subespaço, e de $\eta_{ii}t\sigma_v^2$, que é proporcional à potência do ruído e ao número de amostras.

Lembrando que $s_i^2(t)$ é uma função não decrescente do número de amostras e que a potência do ruído foi assumida constante, se os dados forem muito informativos, o incremento de $s_i^2(t)$ para cada amostra é sempre grande por conta das contribuições da entrada. Em contrapartida, se os dados forem pouco informativos, o crescimento instantâneo de $s_i^2(t)$ será grande apenas enquanto as contribuições da entrada dominarem a contribuição de ruído. Com isso em mente, foi desenvolvido um critério para selecionar os dados que geram os maiores incrementos em $s_i^2(t)$.

Note que $s_i^2(t)$ não é uma quantidade computável pois depende do valor esperado de $\Phi_L(t)$. Porém, Carrette et al. (1996) mostraram que sob a seguinte suposição de excitação é possível utilizar $V^{(i)}$ e $\sigma_i^2(\Phi_L(t))$ ao invés de $\mathcal{V}^{(i)}$ e $s_i^2(t)$, como descrito no Teorema 4.1 mais adiante.

Suposição 4.1 (Suposição de excitação (CARRETTE et al., 1996)). A matriz regressora $\Phi_L(t)$ é tal que

$$\sigma_v^2 \ll \sigma_i^2(\Phi_L(t)) \quad i = 1, \dots, m. \quad (308)$$

◆

A Suposição 4.1 pode ser traduzida como: a energia de cada subespaço $\sigma_i^2(\Phi_L(t))$ é muito maior do que a potência do ruído σ_v^2 . Em outras palavras, considera-se uma boa relação sinal-ruído.

Teorema 4.1 (de Carrette et al. (1996)). *Sob a Suposição 4.1, tem-se:*

$$\frac{\sigma_i^2(\Phi_L(t))}{s_i^2(t)} \approx_p 1 \quad i = 1, \dots, m,$$

$$V^{(i)T} \mathcal{V}^{(i)} \approx_p 1 \quad i = 1, \dots, m,$$

onde \approx_p denota a aproximação no sentido amplo de probabilidade¹.

Prova. A prova deste teorema encontra-se em (CARRETTE et al., 1996). ■

¹Uma variável aleatória $x \in \mathbb{R}$ é dita aproximar-se da constante real $x_0 \neq 0$ no sentido amplo de probabilidade (i. e., $x \approx_p x_0$ ou $x_0 \approx_p x$) se e somente se $E\{(x/x_0 - 1)^2\} \ll 1$ (CARRETTE et al., 1996).

Corolário 4.1 (de Carrette et al. (1996)). *As seguintes relações de transitividade são válidas:*

$$E \left\{ \frac{[\cdot]}{\sigma_i^2(\Phi_L(t))} \right\} \approx_p \frac{E\{[\cdot]\}}{\sigma_i^2(\Phi_L(t))}, \quad i = 1, \dots, m,$$

$$E\{V^{(i)}[\cdot]\} \approx_p V^{(i)} E\{[\cdot]\}, \quad i = 1, \dots, m,$$

ou seja, sob a suposição de excitação, $V^{(i)}$ e $\sigma_i^2(\Phi_L(t))$ são aproximadamente determinísticos.

O Teorema 4.1 e o Corolário 4.1 são bastante importantes e serão utilizados na análise estatística dos parâmetros, apresentada na sequência.

4.2.3 Análise estatística do vetor de parâmetros estimado

O objetivo dessa seção consiste em encontrar expressões para os dois primeiros momentos de probabilidade do vetor de parâmetros estimado $\hat{\rho}$. Para tanto, primeiramente, serão avaliados os dois primeiros momentos dos autoparâmetros.

Teorema 4.2 (Valor esperado dos autoparâmetros (CARRETTE et al., 1996)). *Assuma que a Suposição 4.1 é respeitada. Neste caso, o valor esperado do i -ésimo elemento de $\hat{\rho}_V$ calculado até a t -ésima amostra dos dados é dado por*

$$E\{\hat{\rho}_{Vi}(t)\} \approx_p \hat{\rho}_{dVi} - \frac{\alpha_i t \sigma_v^2}{\sigma_i^2(\Phi_L(t))}, \quad (309)$$

onde $\alpha_i = \text{tr}(M_{Hi}^T P) / t$ com $P \in \mathbb{R}^{t \times t}$ sendo uma matriz Toeplitz, cujas colunas são formadas com a resposta impulsiva do filtro $P(q) = C_d(q)(T_d^{-1}(q) - 1)L(q)H(q)$ que filtra o ruído em (259).

Prova. Considere que cada i -ésimo componente de (280) pode ser escrito como

$$\hat{\rho}_{Vi} = \rho_{dVi} + \sigma_i^{-1} U^{(i)T} \mu, \quad (310)$$

onde $U^{(i)}$ é a i -ésima coluna de U e σ_i é o i -ésimo elemento da diagonal de Σ . A partir de (271) obtém-se:

$$\Sigma U^T = V^T \Phi_L^T, \quad (311)$$

cuja i -ésima coluna é dada por

$$U^{(i)T} \sigma_i = V^{(i)T} \Phi_L^T, \quad (312)$$

$$U^{(i)T} = \Phi_V^{(i)T} \sigma_i^{-1}, \quad (313)$$

onde (313) é obtida considerando-se $V^{(i)T} \Phi_L^T = \Phi_V^{(i)T}$. Substituindo-se (313) em (310), chega-se a

$$\hat{\rho}_{Vi} = \rho_{dVi} + \frac{\Phi_V^{(i)T} \mu}{\sigma_i^2}. \quad (314)$$

Observe que pode-se reescrever a equação anterior considerando-se cada elemento de $\hat{\rho}_V$ obtido até a t -ésima amostra do dado, ou seja,

$$\hat{\rho}_{Vi}(t) = \rho_{dVi} + \frac{\Phi_V^{(i)T} \mu}{\sigma_i^2(\Phi_L(t))}. \quad (315)$$

Aqui as dimensões de $\Phi_V^{(i)T}$ e μ dependem do número de amostras t , contudo, o índice é omitido nestas quantidades para simplificar a notação. A partir da equação acima, calcula-se o valor esperado de cada elemento de $\hat{\rho}_V$, obtido até a t -ésima amostra do dado, como descrito abaixo

$$E\{\hat{\rho}_{Vi}(t)\} = \rho_{dVi} + E\left\{\frac{(v^T M_{Hi}^T + u^T M_{Gi}^T) \mu}{\sigma_i^2(\Phi_L(t))}\right\} \quad (316)$$

$$\approx_p \rho_{dVi} - \frac{E\{(v^T M_{Hi}^T + u^T M_{Gi}^T) P v\}}{\sigma_i^2(\Phi_L(t))} \quad (317)$$

$$\approx_p \rho_{dVi} - \frac{E\{\text{tr}(v^T M_{Hi}^T P v)\}}{\sigma_i^2(\Phi_L(t))} \quad (318)$$

$$\approx_p \rho_{dVi} - \frac{\text{tr}(M_{Hi}^T P E\{v v^T\})}{\sigma_i^2(\Phi_L(t))} \quad (319)$$

$$\approx_p \rho_{dVi} - \frac{\alpha_i t \sigma_v^2}{\sigma_i^2(\Phi_L(t))} \quad (320)$$

onde (316) é obtida considerando-se o Teorema 4.1, que permite utilizar a definição em (296) para $\Phi_V^{(i)}$. Observe que as dimensões de v , u , μ , M_{Gi} e M_{Hi} também variam de acordo com o índice da amostra t (para simplificar a notação este índice foi omitido). O resultado em (317) é encontrado a partir do Corolário 4.1 e substituindo-se $\mu = -Pv$ (com $P \in \mathbb{R}^{t \times t}$ sendo uma matriz Toeplitz, cujas colunas são formadas com a resposta impulsiva do filtro $P(q) = C_d(q)(T_d^{-1}(q) - 1)L(q)H(q)$ que filtra o ruído em (259)). Já (318) é obtida lembrando que os sinais $u(t)$ e $v(t)$ são descorrelacionados e que o termo $v^T M_{Hi}^T P v$ é um escalar. Encontra-se (319) aplicando-se a propriedade cíclica do traço e o fato de que o traço e o valor esperado são operadores lineares. Finalmente, (320) é obtida de (306) e definindo-se $\alpha_i = \text{tr}(M_{Hi}^T P) / t$. ■

Em seguida, avalia-se a matriz de covariância dos autoparâmetros, apresentada no Teorema 4.3. A prova deste teorema é apresentada na sequência.

Teorema 4.3 (de Carrette et al. (1996)). *A matriz de covariância dos autoparâmetros é dada por*

$$\text{Cov}\{\hat{\rho}_V(t)\}_{ij} \approx_p \left[u^T M_{Gi}^T P P^T M_{Gj} u + (\kappa_{ij} - \bar{\kappa}_{ij}) t^2 \sigma_v^2 \right] \frac{\sigma_v^2}{\sigma_i^2(\Phi_L(t)) \sigma_j^2(\Phi_L(t))}, \quad (321)$$

onde $\kappa_{ij} = \text{tr}(M_{Hi}^T P P^T M_{Hj}) / t^2$, $\bar{\kappa}_{ij} = [\text{tr}(M_{Hi}^T P) \text{tr}(P^T M_{Hj})] / t^2$ e P é a matriz Toeplitz formada a partir da resposta impulsiva do filtro $P(q) = C_d(q)(T_d^{-1}(q) - 1)L(q)H(q)$, que filtra o ruído em (259).

Prova. A matriz de covariância dos autoparámetros é escrita da seguinte forma

$$\text{Cov}\{\hat{\rho}_V(t)\}_{ij} = E\left\{(\hat{\rho}_{Vi} - E\{\hat{\rho}_{Vi}\})(\hat{\rho}_{Vj} - E\{\hat{\rho}_{Vj}\})^T\right\} \quad (322)$$

$$\approx_p E\left\{\left(\frac{\Phi_V^{(i)T}\mu}{\sigma_i^2(\Phi_L(t))} + \frac{\alpha_i t \sigma_v^2}{\sigma_i^2(\Phi_L(t))}\right)\left(\frac{\Phi_V^{(j)T}\mu}{\sigma_j^2(\Phi_L(t))} + \frac{\alpha_j t \sigma_v^2}{\sigma_j^2(\Phi_L(t))}\right)^T\right\} \quad (323)$$

$$\approx_p \frac{E\left\{\Phi_V^{(i)T} P v v^T P^T \Phi_V^{(j)} - \Phi_V^{(i)T} P v \alpha_j t \sigma_v^2 - v^T P^T \Phi_V^{(j)} \alpha_i t \sigma_v^2\right\} + \alpha_i \alpha_j t \sigma_v^2}{\sigma_i^2(\Phi_L(t)) \sigma_j^2(\Phi_L(t))} \quad (324)$$

onde (323) é obtida substituindo-se (310) e (320) em (322). Além disso, encontra-se (324) substituindo-se $\mu = -Pv$ em (323). Agora, avalia-se individualmente cada termo de (324). Primeiramente, considere que o primeiro termo de (324) pode ser escrito como:

$$E\left\{\Phi_V^{(i)T} P v v^T P^T \Phi_V^{(j)}\right\} \approx_p E\left\{(u^T M_{Gi}^T + v^T M_{Hi}^T) P v v^T P^T (M_{Gj} u + M_{Hj} v)\right\} \quad (325)$$

$$\approx_p E\left\{\text{tr}(P^T M_{Gj} u u^T M_{Gi}^T P v v^T)\right\} + E\left\{\text{tr}(M_{Hi}^T P v v^T P^T M_{Hj} v v^T)\right\} \quad (326)$$

$$\approx_p u^T M_{Gi}^T P P^T M_{Gj} u \sigma_v^2 + \kappa_{ij} t^2 \sigma_v^4, \quad (327)$$

onde (325) é obtida considerando-se o Teorema 4.1 e de (296), (326) é encontrada lembrando que $u(t)$ e $v(t)$ são descorrelacionados. Por fim, obtém-se (327) usando a definição de κ_{ij} . Na sequência, veja que o segundo termo de (324) pode ser desenvolvido da seguinte maneira

$$E\left\{\Phi_V^{(i)T} P v \alpha_j t \sigma_v^2\right\} \approx_p E\left\{(u^T M_{Gi}^T + v^T M_{Hi}^T) P v \alpha_j t \sigma_v^2\right\} \quad (328)$$

$$\approx_p \text{tr}(M_{Hi}^T P E\{v v^T\}) \alpha_j t \sigma_v^2 \quad (329)$$

$$\approx_p \alpha_i \alpha_j t^2 \sigma_v^4, \quad (330)$$

onde (329) é obtida a partir do Teorema 4.1 e de (296) e lembrando que $u(t)$ e $v(t)$ são descorrelacionados, enquanto (330) é obtida de (306) e usando a variável auxiliar

$$\alpha_i = \text{tr}(M_{Hi}^T P) / t. \quad (331)$$

Similarmente, avalia-se o terceiro termo de (324), chegando-se a

$$E\left\{v^T P^T \Phi_V^{(j)} \alpha_i t \sigma_v^2\right\} \approx_p E\left\{v^T P^T (M_{Gj} u + M_{Hj} v) \alpha_i t \sigma_v^2\right\} \quad (332)$$

$$\approx_p \text{tr}(P^T M_{Hj} E\{v v^T\}) \alpha_i t \sigma_v^2 \quad (333)$$

$$\approx_p \alpha_j \alpha_i t^2 \sigma_v^4, \quad (334)$$

onde obtém-se (332) de (296), considerando-se o Teorema 4.1, e lembrando que $u(t)$ e $v(t)$ são descorrelacionados. Já, (334) é obtida de (306) e definindo-se α_j como em (331).

Finalmente, substituindo-se (327), (330) e (334) em (324), obtém-se:

$$\text{Cov}\{\hat{\rho}_V(t)\}_{ij} \approx_p \frac{u^T M_{Gi}^T P P^T M_{Gj} u \sigma_v^2 + \kappa_{ij} t^2 \sigma_v^2 - \alpha_i \alpha_j t^2 \sigma_v^4}{\sigma_i^2(\Phi_L(t)) \sigma_j^2(\Phi_L(t))} \quad (335)$$

$$\approx_p (u^T M_{Gi}^T P P^T M_{Gj} u + (\kappa_{ij} - \bar{\kappa}_{ij}) t^2 \sigma_v^2) \frac{\sigma_v^2}{\sigma_i^2(\Phi_L(t)) \sigma_j^2(\Phi_L(t))}, \quad (336)$$

com $\bar{\kappa}_{ij}$ como definido no enunciado do teorema. ■

Com base nos dois primeiros momentos de probabilidade dos autoparâmetros dos Teoremas 4.2 e 4.3 e do Corolário 4.1, pode-se avaliar os dois primeiros momentos de probabilidade dos parâmetros originais, ou seja, $\hat{\rho} = V\hat{\rho}_V$. Nesse caso, tem-se que:

$$E\{\hat{\rho}(t)\} \approx_p V E\{\hat{\rho}_V(t)\}, \quad (337)$$

$$\text{Cov}\{\hat{\rho}(t)\} \approx_p V \text{Cov}\{\hat{\rho}_V(t)\} V^T, \quad (338)$$

observe que V depende do índice da amostra t , que é omitido aqui para simplificar a notação.

Dos Teoremas 4.2 e 4.3 tem-se que se um valor singular for muito menor do que os outros valores singulares, o autoparâmetro associado a ele é o que possui maiores polarização e variância (CARRETTE et al., 1996). Como cada elemento de $\hat{\rho}$ depende do autoparâmetro que mais sofre com a polarização, referido aqui como $\hat{\rho}_{V_{i_{min}}}$, deduz-se que o quadrado do menor valor singular $\underline{\sigma}^2(\Phi_L(t))$ é o fator dominante que determina a precisão do vetor de parâmetros $\hat{\rho}$ (CARRETTE et al., 1996). Esta premissa foi utilizada como base para desenvolver o critério SSV, apresentado detalhadamente na próxima subseção.

4.2.4 O critério de seleção dos dados

Este critério consiste em monitorar a evolução da primeira diferença do quadrado do menor valor singular da matriz regressora, $\Delta\underline{\sigma}^2(\Phi_L(t))$, ao longo do tempo. A quantidade $\Delta\underline{\sigma}^2(\Phi_L(t))$ é calculada como

$$\Delta\underline{\sigma}^2(\Phi_L(t)) = \underline{\sigma}^2(\Phi_L(t)) - \underline{\sigma}^2(\Phi_L(t-1)), \quad (339)$$

onde $\underline{\sigma}^2(\Phi_L(t))$ é o quadrado do menor valor singular de $\Phi_L(t)$, calculado como em (282) e t é o número da amostra. O critério SSV é definido da seguinte forma

$$\Delta\underline{\sigma}^2(\Phi_L(t)) > \eta_{ssv}, \quad (340)$$

onde η_{ssv} é um valor de limiar escolhido pelo projetista. Na próxima subseção será sugerida uma forma de escolher um valor adequado para η_{ssv} através da análise visual da curva de $\Delta\underline{\sigma}^2(\Phi_L(t))$. Como escolher esse limiar automaticamente é um tópico a ser investigado. Neste critério, as linhas da matriz regressora, $\varphi_L(t)$, e as correspondentes linhas do vetor de saída, $u_L(t)$, são mantidas sempre que a desigualdade em (340) for atendida. Um resumo de como aplicar o critério SSV é apresentado no Algoritmo 1.

A matriz regressora Φ_{ssv} e o vetor de saída u_{ssv} classificados como mais informativos pelo critério de seleção, são empregados para estimar os parâmetros do controlador com o método VRFT com a estimativa original (OLS) apresentada em (266) e (267), substituindo-se Φ_L e u_L por Φ_{ssv} e u_{ssv} , respectivamente. Contudo, Φ_{ssv} e u_{ssv} também podem ser empregados na estimativa dos parâmetros do controlador com a solução DLS, apresentada nos Teoremas 3.4 e 3.5, substituindo-se Φ e u por Φ_{ssv} e u_{ssv} , respectivamente. Nesse caso, a qualidade da estimativa dos parâmetros também é melhorada, e essa melhora ocorre pois a matriz Φ_{ssv} é “mais informativa” do que a matriz original Φ_L , formada com todo o conjunto de dados. Observe que isso não significa que Φ_{ssv} possui mais informação do que Φ_L , mas sim que naquela a informação proveniente da entrada está mais concentrada, enquanto que nesta a informação útil está diluída em meio às contribuições de ruído. Os resultados de simulação apresentados na próxima subseção apresentam um exemplo onde pode ser observada a melhora na qualidade da estimativa.

Algoritmo 1: Seleção dos subconjuntos muito informativos de dados com o critério SSV

Dados: Modelo de referência $T_d(q)$, vetor $\bar{C}(q)$, filtro $L(q)$, limiar η_{SSV} , número de parâmetros a serem estimados m , vetor com o sinal de entrada u e vetor com o sinal de saída y

Resultado: Matriz regressora Φ_{SSV} e vetor de saída u_{SSV} contendo as componentes classificadas como informativas.

```

1 calcule  $\Phi_L$  como em (262) utilizando (65) ;
2 calcule  $u_L$  como em (263) utilizando (64);
3  $\Phi_{SSV} \leftarrow 0$  ;
4  $u_{SSV} \leftarrow 0$  ;
5  $t \leftarrow m + 1$  ;
6 enquanto  $t < N$  faça
7   calcule  $\Delta\sigma^2(\Phi_L(t))$  como em (339) ;
8   se  $\Delta\sigma^2(\Phi_L(t)) > \eta_{SSV}$  então
9      $\Phi_{SSV} \leftarrow [\Phi_{SSV} \ \varphi_L(t)]$  ;
10     $u_{SSV} \leftarrow [u_{SSV} \ u_L(t)]$  ;
11  fim
12   $t \leftarrow t + 1$  ;
13  próxima iteração;
14 fim

```

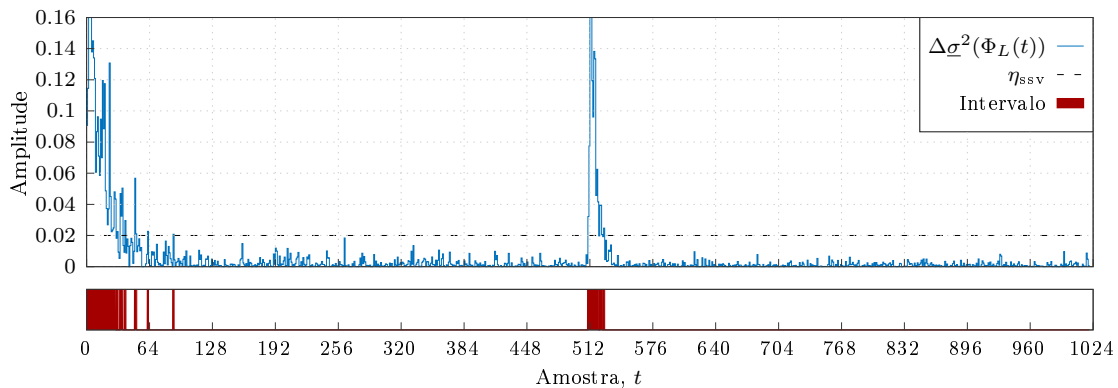
4.2.5 Exemplo ilustrativo

O objetivo desta subseção é apresentar um exemplo ilustrando a aplicação do critério SSV para seleção de subconjuntos muito informativos de dados com o método VRFT. Além disso, será apresentada uma forma de escolher o valor de limiar η_{SSV} através da análise visual da curva do quadrado do menor valor singular da matriz regressora. Neste caso, utilizou-se a mesma configuração do exemplo motivacional do método VRFT, apresentado na Seção 2.5.2.

Para cada rodada de Monte Carlo, obteve-se uma matriz regressora Φ_{SSV} e um vetor de saída u_{SSV} selecionados com o critério SSV, descrito no Algoritmo 1. A Figura 14 apresenta, para uma realização dos dados, a evolução ao longo do tempo da primeira diferença do quadrado do menor valor singular (linha azul contínua). Esta mesma figura também apresenta o valor de limiar escolhido η_{SSV} (linha preta pontilhada) e os subconjuntos muito informativos (linhas vermelhas verticais). Como mencionado anteriormente, este gráfico pode ser utilizado para escolher um valor adequado para o limiar η_{SSV} . Lembrando que valores baixos para $\Delta\sigma^2(\Phi_L(t))$ provém das linhas de $\Phi_L(t)$ que são muito afetadas por ruído, pode-se escolher o limiar de modo que tais linhas sejam removidas. Este comportamento pode ser observado na Figura 14. Neste exemplo, escolheu-se $\eta_{SSV} = 0,02$, para todas as realizações, com o qual obteve-se uma matriz Φ_{SSV} com aproximadamente 60 linhas informativas, em média, ou seja, aproximadamente 5,86% de todo o conjunto de dados foi utilizado na estimativa dos parâmetros do controlador.

A partir da matriz regressora e do vetor de saída selecionados, é realizado o ajuste dos parâmetros do controlador através das soluções OLS e DLS, para cada uma das rodadas de Monte Carlo. Além disso, os resultados obtidos são comparados com as estimativas calculadas utilizando-se as mesmas soluções empregando-se todo o conjunto de dados. A

Figura 14 — Primeira diferença do quadrado do menor valor singular de uma realização dos dados do exemplo da Seção 4.2.5.



Evolução do quadrado do menor valor singular $\Delta\sigma^2(\Phi_L(t))$ (linha contínua azul), limiar η_{ssv} (linha tracejada preta) e subconjuntos muito informativos (linhas verticais vermelhas). Fonte: elaborado pela autora.

Tabela 6 apresenta as estatísticas dos parâmetros estimados com todo o conjunto de dados. Já as estatísticas dos parâmetros estimados com os subconjuntos muito informativos são apresentadas na Tabela 7. Observe que as estimativas obtidas com todo o conjunto de dados apresentam uma polarização bastante acentuada, ou seja, estão distantes do valor ideal. Em contrapartida, as estimativas obtidas com os intervalos muito informativos possuem uma polarização bem menor do que com a correspondente solução e todo o dado. Além disso, nota-se que houve um aumento na variância das estimativas com os subconjuntos, comparando-se com os resultados obtidos com todo o conjunto de dados. Contudo, este aumento não é prejudicial, devido à expressiva redução na polarização. Finalmente, note que a estimativa obtida com OLS e todo o conjunto de dados deste caso é diferente da estimativa obtida no exemplo da Seção 2.5.2 (Tabela 3). Esta diferença é esperada, pois, neste exemplo, os sinais são filtrados através do filtro $L(q) = T_d(q)$ que por ser um filtro passa baixas, reduz a contribuição do ruído. Por conta disso, ocorreu a melhora na qualidade da estimativa obtida (considerando-se todo o conjunto de dados).

Tabela 6 — Média e desvio padrão dos parâmetros estimados com o VRFT e as soluções OLS e DLS com todo o conjunto de dados do exemplo da Seção 4.2.5.

ρ	ρ_d	Estimativas	
		OLS	DLS
K_p	0,3742	0,2257 ($\pm 0,0073$)	0,4114 ($\pm 0,0122$)
K_i	0,0593	0,0593 ($\pm 0,0001$)	0,0592 ($\pm 0,0001$)

4.3 Critério baseado no número de condicionamento

Esta seção tem como objetivo apresentar o critério do número de condicionamento (do inglês *Condition Number* – CN) da matriz de informação aplicado ao método VRFT. Na sequência, esse critério é apresentado (aqui referido como critério CN) e ao final da seção

Tabela 7 — Média e desvio padrão dos parâmetros estimados com o VRFT e as soluções OLS e DLS com os subconjuntos muito informativos selecionados através do critério SSV para o exemplo da Seção 4.2.5.

ρ	ρ_d	Estimativas	
		OLS	DLS
K_p	0,3742	0,3576 ($\pm 0,0111$)	0,3742 ($\pm 0,0117$)
K_i	0,0593	0,0605 ($\pm 0,0006$)	0,0605 ($\pm 0,0006$)

são mostrados dois exemplos de simulação ilustrando sua aplicação.

4.3.1 O critério de seleção dos dados

O número de condicionamento de uma matriz X qualquer, não singular, é dado por

$$\kappa_p(X) = \|X\|_p \|X^{-1}\|_p,$$

onde $\|\cdot\|_p$ corresponde à norma p da matriz (QUARTERONI; SACCO; SALERI, 2007). No presente trabalho, considera-se $p = 2$, e a matriz em questão é a matriz de informação do método VRFT. Assim, o número de condicionamento da matriz de informação é calculado como segue

$$\kappa(P(t)) = \frac{\bar{\sigma}(P(t))}{\underline{\sigma}(P(t))}, \quad (341)$$

onde o índice $p = 2$ é omitido para simplificar a notação, t é o índice da amostra e $P(t) \in \mathbb{R}^{m \times m}$ é a matriz de informação definida como

$$P(t) = \Phi_L^T(t) \Phi_L(t), \quad (342)$$

calculada até a t -ésima amostra, enquanto $\bar{\sigma}(P(t))$ é o maior valor singular e $\underline{\sigma}(P(t))$ o menor valor singular de $P(t)$.

O critério CN é aplicado para cada degrau presente nos dados coletados. Por conta disso, se o sinal possuir múltiplos degraus um algoritmo pode ser aplicado para detectar os intervalos onde ocorrem as mudanças no sinal de excitação (BITTENCOURT et al., 2015; ARENGAS; KROLL, 2017b). Contudo, no presente trabalho, assume-se que os intervalos onde ocorrem as mudanças no sinal de excitação são conhecidos.

Esse critério baseia-se no fato de que se o sinal de entrada $u(t)$ é pouco informativo e sua contribuição domina sobre a contribuição do ruído, a matriz de informação $P(t)$ se tornará mal condicionada com o tempo. Portanto, pode-se prevenir a degradação do número de condicionamento descartando-se os dados obtidos muito depois do início do degrau. Esse critério consiste em manter os vetores regressores $\varphi_L(t)$ e as respectivas amostras da saída $u_L(t)$ enquanto a seguinte desigualdade é respeitada:

$$\kappa(P(t)) < \eta_{cn}, \quad (343)$$

onde η_{cn} é um valor de limiar escolhido pelo projetista e t é o índice da amostra que varia do início do degrau até o final do intervalo (até o início do próximo degrau). Assim sendo, o subconjunto muito informativo é delimitado pelo início da mudança no sinal de

Algoritmo 2: Seleção dos subconjuntos muito informativos de dados com o critério CN

Dados: Modelo de referência $T_d(q)$, vetor $\bar{C}(q)$, filtro $L(q)$, limiar η_{cn} , número de parâmetros a serem estimados m , vetores com os sinais de entrada u e saída y .

Resultado: Matriz regressora Φ_{cn} e vetor de saída u_{cn} contendo as componentes classificadas como informativas.

```

1  $\Phi_{cn} \leftarrow 0$  ;
2  $u_{cn} \leftarrow 0$  ;
3  $t \leftarrow m$  ;
4 calcule  $\Phi_L$  como em (262) ;
5 calcule  $u_L$  como em (263) ;
6 enquanto  $t \leq N$  faça
7   | calcule  $P(t)$  como em (342) ;
8   | calcule  $\kappa(t)$  como em (341) ;
9   | se  $\kappa(t) \geq \eta_{cn}$  então fim da execução;
10  |  $\Phi_{cn} \leftarrow [\Phi_{cn} \ \varphi_L(t)]$  ;
11  |  $u_{cn} \leftarrow [u_{cn} \ u_L(t)]$  ;
12  |  $t \leftarrow t + 1$  ;
13 fim

```

referência ou de entrada até, mas não incluindo, a primeira amostra em que a desigualdade em (343) deixa de ser respeitada. A aplicação do critério CN é resumida no Algoritmo 2.

É importante mencionar que a escolha do valor de limiar é um tema abordado na literatura de identificação de sistemas. Shardt e Huang (2013) mencionam que $\eta_{cn} = 10^4$ é um valor apropriado para esse limiar. Naquele trabalho, os autores comentam que valores maiores ou menores podem ser empregados, cuja escolha é determinada pela estrutura do modelo a ser identificado, entre outros fatores. Embora sugerido para identificação de sistemas, no presente trabalho, utiliza-se a sugestão acima mencionada como valor inicial para η_{cn} e o mesmo é ajustado conforme a necessidade, identificada a partir da análise visual do gráfico da evolução ao longo do tempo do número de condicionamento. Por exemplo, se os valores de $\kappa(P(t))$ são sempre menores do que 10^4 , não adianta usar tal valor como limiar. Portanto, nesse caso é preciso escolher um valor para η_{cn} de modo que as amostras obtidas muito depois do degrau sejam descartadas. Uma forma de definir automaticamente um valor para esse limiar é uma questão a ser investigada.

A partir da matriz regressora Φ_{cn} e do vetor de saída u_{cn} selecionados pelo critério CN estimam-se os parâmetros do controlador com o método VRFT. A estimativa dos parâmetros pode ser calculada empregando a solução OLS, em (266) e (267), ou a solução DLS dos Teoremas 3.4 e 3.5. A única alteração necessária é substituir Φ_L e u_L (OLS) ou Φ e u (DLS) por Φ_{cn} e u_{cn} , respectivamente. O critério CN foi aplicado com sucesso no caso em que o controlador ideal pertence à classe de controle, ou seja, quando a Suposição 2.1 é respeitada. Contudo, este comportamento não se mantém quando a Suposição 2.1 é violada, como será apresentado na próxima subseção.

4.3.2 O caso em que o controlador ideal não pertence à classe de controle

Considere o vetor regressor $\varphi_L(t)$ reescrito como

$$\varphi_L(t) = \bar{C}(q)(1 - T_d(q))^2 y(t) \quad (344)$$

$$= \bar{F}(q)y(t), \quad (345)$$

onde (344) é obtida substituindo-se $L(q)$ em (65) por um filtro cuja estrutura respeita (77) e (345) é calculada definindo-se $\bar{F}(q) = \bar{C}(q)(1 - T_d(q))^2$. Observe que quando a estrutura do controlador escolhida é do tipo PID, o processo em malha fechada deve ter ganho DC unitário, o que deve ser refletido no modelo de referência $T_d(q)$. Por conta disso, a função de sensibilidade $(1 - T_d(q))$ possui um zero em um, consequentemente, $(1 - T_d(q))^2$ possui dois zeros em um.

Seja a matriz de informação $P(t + 1)$ escrita como

$$P(t + 1) = P(t) + \varphi_L(t + 1)\varphi_L^T(t + 1) \quad (346)$$

$$= P(t) + [\bar{F}(q)y(t + 1)][\bar{F}(q)y(t + 1)]^T \quad (347)$$

$$\begin{aligned} &= P(t) + [\bar{F}(q)G(q)u(t + 1)][\bar{F}(q)G(q)u(t + 1)]^T \\ &\quad + 2[\bar{F}(q)G(q)u(t + 1)][\bar{F}(q)H(q)v(t + 1)]^T \\ &\quad + [\bar{F}(q)H(q)v(t + 1)][\bar{F}(q)H(q)v(t + 1)]^T, \end{aligned} \quad (348)$$

onde (347) é obtida substituindo-se (345) em (346), enquanto (348) usa (1). Considerando-se que $\bar{F}(q)$ possui dois zeros em um e que $t \gg 1$, tem-se:

$$\bar{F}(q)G(q)u(t) \approx 0, \quad (349)$$

lembrando que $u(t)$ é um degrau. Portanto, após algum tempo a matriz de informação será afetada somente pela contribuição do ruído, que é um sinal muito informativo. Dessa forma, o número de condicionamento irá diminuir conforme aumenta-se o número de amostras sem nunca satisfazer (343). Por conta disso, no presente trabalho, este critério não é aplicado para o caso em que o controlador ideal não pertence à classe de controle.

4.3.3 Exemplos ilustrativos

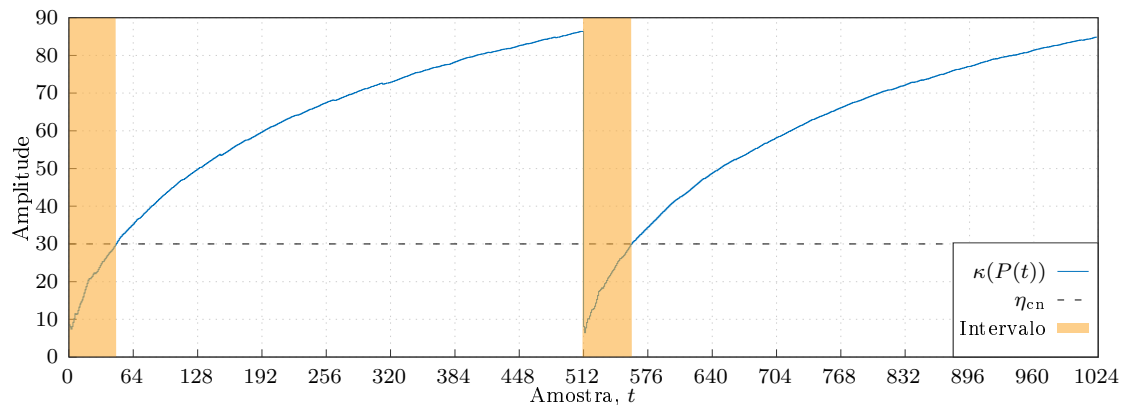
Nesta subseção são apresentados dois exemplos de simulação que ilustram a aplicação do critério CN no método VRFT. O primeiro exemplo considera que o controlador ideal pertence à classe de controle, ou seja, a Suposição 2.1 é respeitada. No segundo exemplo, considera-se o caso em que a Suposição 2.1 é violada.

4.3.3.1 O caso em que o controlador ideal pertence à classe de controle

O primeiro exemplo de simulação utiliza a mesma configuração da simulação apresentada na Seção 2.5.2. Em cada rodada de Monte Carlo, seleciona-se uma matriz regressora Φ_{cn} e um vetor de saída u_{cn} com as componentes informativas selecionadas com o critério CN, como descrito no Algoritmo 2. A Figura 15 apresenta, para uma realização, a evolução do número de condicionamento da matriz de informação $P(t)$ (linha contínua azul), para cada degrau presente no sinal de entrada. Essa figura também apresenta o valor de limiar escolhido η_{cn} (linha tracejada preta) e os intervalos muito informativos selecionados

(linhas verticais amarelas). Nesse caso, $\eta_{cn} = 30$ foi o valor escolhido para o limiar, com o qual foram selecionadas em média 97 amostras de todo o conjunto de dados. Desta forma, aproximadamente 9,47% dos dados são utilizados na estimativa dos parâmetros do controlador.

Figura 15 — Evolução no tempo do número de condicionamento para uma realização dos dados do exemplo da Seção 4.3.3.1.



Evolução do número de condicionamento $\kappa(P(t))$ para cada degrau (linha contínua azul), limiar η_{cn} (linha tracejada preta) e subconjuntos muito informativos (linhas verticais amarelas). Fonte: elaborado pela autora.

A partir da matriz regressora Φ_{cn} e do vetor de saída u_{cn} foram estimados os parâmetros do controlador através das soluções OLS e DLS. As estatísticas das estimativas obtidas são apresentadas na Tabela 8. Observe que a polarização da estimativa com OLS é um pouco maior do que a obtida com a solução DLS, contudo, a variância das estimativas é bem parecida. Considere agora os resultados obtidos com as soluções OLS e DLS e todo o conjunto de dados, Tabela 6. Veja que a polarização da estimativa é bem menor com os dados selecionados com o critério CN do que com todo o conjunto de dados, independente da solução empregada. Além disso, nota-se um aumento na variância das estimativas, contudo, este aumento é compensado pela grande redução na polarização.

Tabela 8 — Média e desvio padrão dos parâmetros estimados com o VRFT e as soluções OLS e DLS com os subconjuntos muito informativos obtidos com o critério CN para o exemplo da Seção 4.3.3.1.

ρ	ρ_d	Estimativas	
		OLS	DLS
K_p	0,3742	0,3488 ($\pm 0,0122$)	0,3796 ($\pm 0,0139$)
K_i	0,0593	0,0597 ($\pm 0,0005$)	0,0592 ($\pm 0,0005$)

4.3.3.2 O caso em que o controlador ideal não pertence à classe de controle

O objetivo do segundo exemplo de simulação consiste em apresentar o comportamento do critério CN no caso em que a Suposição 2.1 é violada. Considere que o modelo do processo é descrito em (84) e que o modelo de referência é dado por

$$T_d(q) = \frac{0,15443(q - 0,4779)(q - 0,8)(q - 0,905)}{(q - 0,8098)(q - 0,9075)(q^2 - 1,513q + 0,5997)}. \quad (350)$$

O controlador ideal, calculado de (16) a partir de (84) e (350), é dado por

$$C_d(q) = \frac{0,30885(q - 0,905)(q - 0,4779)}{(q - 1)(q - 0,7844)}. \quad (351)$$

Nesse caso, deseja-se projetar um controlador do tipo PI com a seguinte estrutura

$$C(q, \rho) = [\rho_1 \quad \rho_2] \begin{bmatrix} \frac{q}{q-1} & \frac{1}{q-1} \end{bmatrix}^T, \quad (352)$$

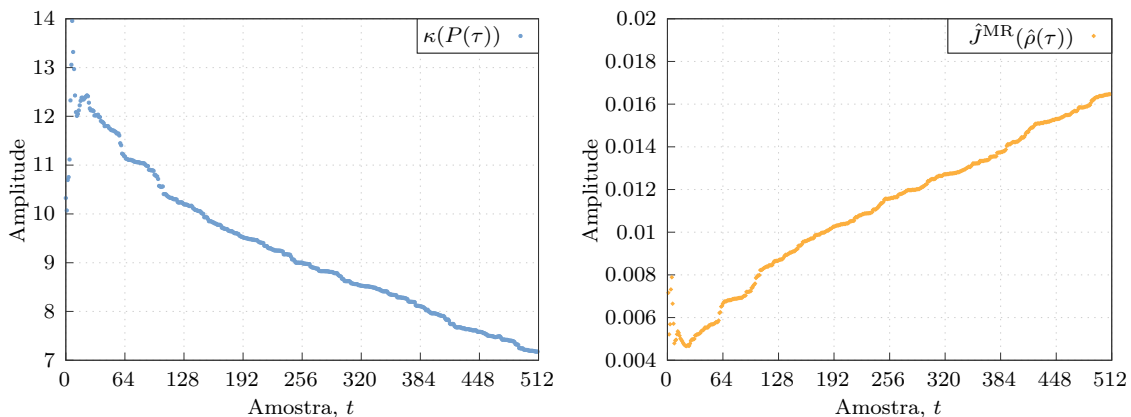
desta forma, temos o caso em que a Suposição 2.1 é violada. Para estimar os parâmetros do controlador, coletou-se os dados de entrada $u(t)$ e saída $y(t)$ do processo em um experimento em malha aberta. O sinal de entrada consiste em um degrau com 512 amostras e amplitude 1. O sinal de saída é contaminado com um ruído branco com variância $\sigma_v^2 = 0,01$ filtrado através de (88). Os sinais são filtrados através do filtro $L(q) = T_d(q)(1 - T_d(q))$, que respeita a estrutura (77). O número de condicionamento $\kappa(P(\tau))$ foi calculado variando-se o número da amostra, ou seja, $\tau \geq m + 1, \dots, N$, e é apresentado na Figura 16a. Observe que conforme o número de amostras aumenta, o número de condicionamento diminui. Este comportamento é esperado, e como mencionado anteriormente é devido ao ruído. Para mostrar que aumentando-se o número de amostras não melhora a qualidade da estimativa, foram estimados os parâmetros do controlador ao longo do tempo, ou seja, projetou-se $\hat{\rho}(\tau)$ com $\tau \geq m + 1, \dots, N$. Para cada controlador projetado obteve-se uma estimativa para a função custo do critério de desempenho de seguimento de referência (12), dada por

$$\hat{J}^{\text{MR}}(\hat{\rho}(\tau)) = \frac{1}{N} \sum_{t=1}^N [y(t, \hat{\rho}(\tau)) - y_d(t)]^2, \quad (353)$$

onde $y(t, \hat{\rho}(\tau)) = T(q, \hat{\rho}(\tau))r(t)$ é a resposta em malha fechada com o controlador $C(q, \hat{\rho}(\tau))$, $y_d(t) = T_d(q)r(t)$ é a resposta em malha fechada desejada, $r(t)$ é o sinal de referência empregado, neste exemplo, um degrau, e $N = 50$ é o número de amostras.

A evolução de $\hat{J}^{\text{MR}}(\hat{\rho}(\tau))$ é apresentada na Figura 16b. Perceba que sua curva apresenta um vale (que não parece ter relação com a curva de $\kappa(P(\tau))$ na Figura 16a) e que após

Figura 16 — Evolução do número do condicionamento $\kappa(P(\tau))$ e da função custo estimada $\hat{J}_y(\hat{\rho}(t))$ para um degrau do exemplo da Seção 4.3.3.2.



(a) Uma realização de $\kappa(P(\tau))$.

(b) Evolução de $\hat{J}_y(\hat{\rho}(\tau))$ ao longo do tempo.

Fonte: elaborado pela autora.

um certo número de amostras há um aumento no valor da função custo. Por conta disso, aumentar o número de amostras não melhora o desempenho em malha fechada do processo.

4.4 Nota

Observe que o desenvolvimento/aplicação dos critérios descritos anteriormente considera que algumas condições são respeitadas. Por exemplo, o critério SSV foi concebido considerando-se as seguintes condições ideais: os dados são coletados durante a execução de um experimento em malha aberta, o sinal de entrada é pouco informativo (um degrau), há uma boa relação sinal-ruído na saída, e o processo é BIBO estável. Quanto ao critério CN, o mesmo é aplicado em cada degrau. Veja que em ambos os critérios, as linhas da matriz regressora e respectivas linhas do vetor de saída selecionadas provêm de um degrau, que é um sinal pouco informativo. Contudo, com base nos desenvolvimentos apresentados, nota-se que a ideia de ambos os critérios consiste em tornar a matriz regressora mais informativa. Considere o critério SSV, por exemplo. Naquele caso, valores pequenos para $\Delta\sigma^2(\Phi_L(t))$ indicam que as correspondentes linhas da matriz regressora $\Phi_L(t)$ são pouco informativas, ou seja, elas são dominadas pela contribuição do ruído. Por isso, remover tais linhas e as correspondentes linhas de u_L melhora a qualidade da estimativa obtida para os parâmetros do controlador. Em contrapartida, o critério CN previne a degradação do número de condicionamento de $P(t)$ ao utilizar-se uma quantidade reduzida de amostras localizadas após uma mudança no sinal de entrada. Por conta disso, utilizar os subconjuntos muito informativos selecionados através de um dos critérios leva a uma melhora no desempenho em malha fechada. Conforme será apresentado no Capítulo 5, esta melhora é observada também nos casos em que as condições não são ideais.

4.5 Considerações finais

Esse capítulo apresentou a solução proposta no presente trabalho para suprir as dificuldades quanto à utilização de dados pouco informativos com o método VRFT. Tal solução consistiu em adaptar para o VRFT dois critérios de seleção de subconjuntos muito informativos de dados, originalmente aplicados em identificação de sistemas. Em cada critério são pré-selecionadas as linhas informativas da matriz regressora e as correspondentes linhas do vetor de saída, que são então utilizados na estimativa dos parâmetros do controlador.

As abordagens propostas nesse capítulo ainda apresentam algumas questões em aberto que valem a pena serem investigadas. A extensão da análise teórica do critério SSV para dados de malha fechada é um assunto interessante. Resultados de simulação, que serão apresentados no Capítulo 5, já indicam um bom desempenho desse critério nesse caso. Outra questão interessante consiste no ajuste automático dos valores de limiar. Além disso, a extensão do critério SSV para lidar com dados com perturbações também é um assunto atrativo. Uma última questão é a aplicação de outros critérios de seleção de intervalos muito informativos, como o proposto por Holcomb e Bitmead (2017), no método VRFT.

5 RESULTADOS

5.1 Introdução

O objetivo desse capítulo consiste em apresentar alguns estudos de caso simulados utilizados para validar as soluções propostas no presente trabalho. Para validar a aplicação de outras soluções de mínimos quadrados ao método VRFT foram realizados três estudos de caso, apresentados na Seção 5.2. O primeiro estudo de caso considera que o controlador ideal pertence à classe de controle, ou seja, a Suposição 2.1 é respeitada. No segundo estudo de caso assume-se que essa suposição é violada. Finalmente, o terceiro exemplo de simulação considera que os dados são coletados em um cenário mais realista.

Em seguida, na Seção 5.3, mais três estudos de caso são apresentados, com o objetivo de validar a aplicação dos critérios de seleção de subconjuntos muito informativos de dados ao método VRFT. O primeiro estudo de caso assume que as condições ideais, sob as quais o critério SSV foi desenvolvido, são respeitadas. Nos dois exemplos seguintes, considera-se que algumas das condições ideais são violadas. Além disso, o terceiro exemplo apresenta a combinação dos dois critérios de seleção.

Uma das métricas empregadas para comparar os resultados obtidos é uma estimativa da função custo do critério de desempenho de seguimento de referência $J^{\text{MR}}(\rho)$, dada por

$$\hat{J}^{\text{MR}}(\hat{\rho}) = \frac{1}{N} \sum_{t=1}^N [(T(q, \hat{\rho}) - T_d(q))r(t)]^2, \quad (354)$$

onde $\hat{\rho}$ é o vetor de parâmetros estimado em cada realização dos dados, $T(q, \hat{\rho})$ é a função de sensibilidade complementar, calculada como em (7) com o controlador estimado em cada realização $C(q, \hat{\rho})$, enquanto $r(t)$ é um sinal do tipo degrau e N é o número de amostras coletadas. É importante salientar que essa função é calculada para cada controlador estimado com cada método e/ou critério. Outras métricas empregadas são a média e o desvio padrão dos parâmetros estimados em cada caso. Nos casos em que o controlador ideal não pertence à classe de controle, a resposta em frequência dos controladores estimados também é utilizada como métrica para avaliar os resultados.

5.2 Soluções de mínimos quadrados

5.2.1 Caso de ordem cheia, dados de malha fechada e controlador com polos livres

O objetivo desse estudo de caso consiste em validar a solução CTLS com polos livres e comparar os resultados dessa abordagem com os gerados pelas soluções originais do método VRFT. Além disso, nesse exemplo mostra-se que o CTLS produz resultados similares ao OCI, cuja estimativa é não polarizada e estatisticamente eficiente. A configuração desse estudo de caso é baseada em um dos exemplos de simulação de Campestrini et al. (2017).

Nesse caso, 100 simulações de Monte Carlo foram executadas variando a realização de ruído. A função de transferência do processo $G(q)$ é descrita em (84) e reescrita abaixo por conveniência

$$G(q) = \frac{0,5(q-0,8)}{(q-0,7)(q-0,9)}. \quad (84)$$

O modelo de referência escolhido é dado por

$$T_d(q) = \frac{0,16q}{(q-0,6)^2}. \quad (355)$$

O controlador ideal é calculado substituindo-se (84) e (355) em (16) chegando-se a

$$C_d(q) = \frac{0,32(q-0,7)(q-0,9)}{(q-0,36)(q-0,8)} \underbrace{\frac{q}{q-1}}_{C_F(q)}, \quad (356)$$

onde a parte fixa contém um integrador e um zero na origem. Por conta disso, 5 parâmetros são estimados. Note que o modelo do processo não é empregado no projeto do controlador, o mesmo foi utilizado apenas para calcular o controlador ideal. O vetor de parâmetros ideal ρ_d é dado por

$$\rho_d = [0,3200 \quad -0,5120 \quad 0,2016 \quad -1,1600 \quad 0,2880]^T. \quad (357)$$

Os dados de entrada $u(t)$ e saída $y(t)$ são coletados de um experimento em malha fechada, com o seguinte controlador operando na malha:

$$C(q, \rho_0) = \frac{0,15(q-0,6667)}{q-1}. \quad (358)$$

Um sinal do tipo PRBS com 1023 amostras é aplicado como sinal de referência $r(t)$. Durante a simulação, um ruído colorido é adicionado à saída do processo. Esse sinal é gerado para cada realização através de um ruído branco com variância $\sigma_v^2 = 1 \times 10^{-4}$ filtrado através do modelo para o ruído descrito em (88):

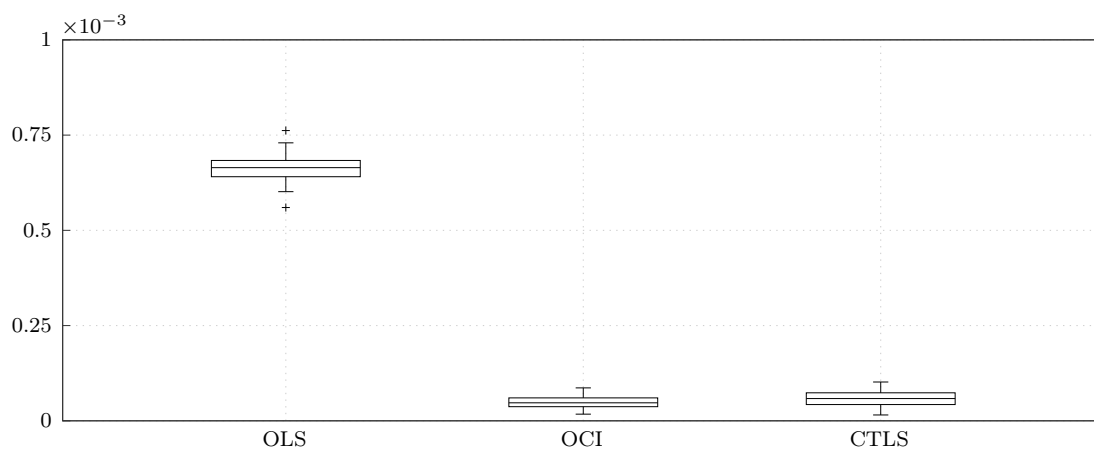
$$H(q) = \frac{q}{q-0,3}. \quad (88)$$

Para cada simulação de Monte Carlo os parâmetros do controlador são estimados utilizando-se: a solução CTLS como apresentada na Seção 3.2.2; as soluções OLS e VI como descritas na Seção 2.4.1.2; e o método OCI apresentado no Seção 2.4.2. O problema de otimização do CTLS é não convexo, por conta disso, sua solução requer um algoritmo iterativo e um valor inicial para os parâmetros. Nesses estudos de caso, a estimativa dos parâmetros é obtida através do algoritmo de otimização iterativo `fminsearch` do Matlab. Esse algoritmo é inicializado com o vetor $\rho_{ini} = 0,8\rho_d$. Observe que este é apenas um exemplo ilustrativo e outras escolhas de algoritmos e inicializações são possíveis. Em situações práticas, as estimativas obtidas com as soluções OLS ou VI podem ser consideradas como valores de inicialização do algoritmo de otimização. Contudo, essas escolhas devem ser cuidadosamente consideradas, pois elas podem fazer com que o algoritmo convirja para um mínimo local. Quanto ao método OCI, a ferramenta `ident` do Matlab é empregada para estimar os parâmetros do controlador, como sugerido por Campestrini et al. (2017). Veja que devido aos dados serem coletados em malha fechada,

é preciso estimar também um modelo para o ruído. Por conta disso, são estimados os parâmetros de modelo do tipo Box-Jenkins, como apresentado na Seção 2.3.4. Assim sendo, no método OCI 6 parâmetros são estimados. Como a estimativa dos mesmos também é obtida através de um problema de otimização não convexo utilizou-se como valores iniciais: $\rho_{\text{ini}} = 0,8\rho_d$ e $\eta_{\text{ini}} = 0,8\eta_d$, onde $\eta_d = 0,3$ é obtido de (88).

A Figura 17 apresenta a distribuição das funções custo estimadas $\hat{J}^{\text{MR}}(\hat{\rho})$, calculadas como em (354) com $N = 100$, com os parâmetros ajustados através do método VRFT com as soluções OLS e CTLS e através do método OCI. Observe que os valores encontrados

Figura 17 — Distribuição da função custo estimada $\hat{J}^{\text{MR}}(\hat{\rho})$ com os parâmetros estimados com OLS, CTLS e OCI da simulação da Seção 5.2.1.



Fonte: elaborado pela autora.

com a solução OLS são muito maiores do que os gerados pela solução CTLS e pelo método OCI. Esse resultado é esperado, uma vez que a estimativa obtida com OLS é polarizada na presença de ruído. De fato, as abordagens CTLS e OCI apresentam resultados aproximadamente 10 vezes menores do que os encontrados com OLS. Os resultados obtidos com VI não são apresentados nesse caso pois 27% dos controladores encontrados resultaram em um comportamento estável em malha fechada. Esse comportamento é esperado, pois a variância da estimativa com a abordagem de VI é bastante pronunciada.

As estatísticas dos parâmetros estimados com OLS, OCI e CTLS são apresentadas na Tabela 9. Observe que com OCI e CTLS a polarização da estimativa é bem menor do que com a solução OLS. Além disso, nota-se que as variâncias das estimativas com OCI e CTLS são muito parecidas e um pouco maiores do que a encontrada com a solução OLS.

Tabela 9 — Média e desvio padrão dos parâmetros estimados com OLS, CTLS e OCI para a simulação da Seção 5.2.1.

ρ	ρ_d	Estimativas		
		OLS	OCI	CTLS
ρ_1	0,3200	0,3116 ($\pm 0,0017$)	0,3201 ($\pm 0,0015$)	0,3195 ($\pm 0,0015$)
ρ_2	-0,5120	-0,1176 ($\pm 0,0073$)	-0,5120 ($\pm 0,0090$)	-0,5113 ($\pm 0,0099$)
ρ_3	0,2016	-0,1289 ($\pm 0,0064$)	0,2015 ($\pm 0,0075$)	0,2014 ($\pm 0,0084$)
ρ_4	-1,1600	0,0479 ($\pm 0,0224$)	-1,1587 ($\pm 0,0285$)	-1,1607 ($\pm 0,0320$)
ρ_5	0,2880	-0,2248 ($\pm 0,0115$)	0,2870 ($\pm 0,0161$)	0,2884 ($\pm 0,0176$)

Com base nos resultados apresentados e discutidos acima, pode-se observar que a solução CTLS apresenta resultados melhores do que as soluções originais do método VRFT (OLS e VI). De fato, com CTLS não foram encontrados controladores que instabilizaram o processo em malha fechada, como com VI, e as funções custo estimadas são aproximadamente 10 vezes menores do que com OLS. Além disso, o CTLS apresentou resultados similares quando comparado com o método OCI. Observa-se que a solução CTLS não requer o conhecimento de informações sobre o modelo do ruído, mesmo com dados de malha fechada. Já no método OCI, para dados de malha fechada, é preciso definir o número de parâmetros a serem estimados para $H(q)$, que é usualmente desconhecida. Por conta disso, temos que menos parâmetros são estimados com CTLS, o que pode ser uma vantagem, pois, como já mencionado nesse texto, a identificação de mais parâmetros aumenta a variância da estimativa.

5.2.2 Caso de ordem reduzida, dados de malha aberta e controlador com polos fixos

O objetivo desse exemplo é ilustrar a utilização das soluções CTLS e DLS no caso em que o controlador ideal não pertence à classe de controle, ou seja, a Suposição 2.1 não é respeitada. A configuração da simulação desse estudo de caso é baseada no exemplo da Seção 4.3.3.2. Portanto, o modelo que representa o processo é descrito em (84). Nesse caso, considere que o modelo de referência é dado por

$$T_d(q) = \frac{0,15443(q - 0,4779)(q - 0,8)(q - 0,905)}{(q - 0,8098)(q - 0,9075)(q^2 - 1,513q + 0,5997)}. \quad (350)$$

O controlador ideal, calculado de (16) a partir de (84) e (350), é dado por

$$C_d(q) = \frac{0,30885(q - 0,905)(q - 0,4779)}{(q - 1)(q - 0,7844)}. \quad (351)$$

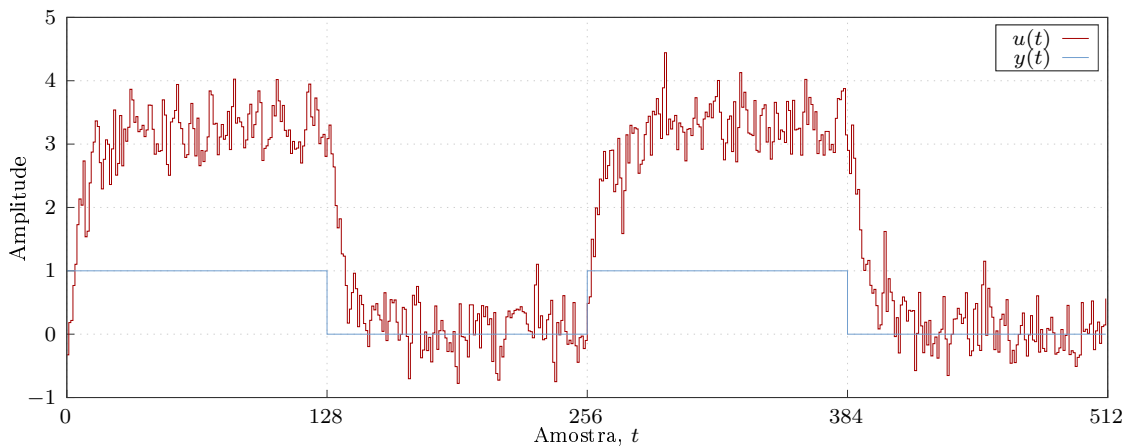
Nesse caso, o objetivo é projetar um controlador do tipo PI com a seguinte estrutura

$$C(q, \rho) = [\rho_1 \quad \rho_2] \begin{bmatrix} \frac{q}{q-1} & \frac{1}{q-1} \end{bmatrix}^T. \quad (352)$$

Como a estrutura escolhida não é capaz de representar exatamente a estrutura do controlador ideal temos o caso em que a Suposição 2.1 é violada. Para estimar os parâmetros do controlador, foram coletados dados de entrada $u(t)$ e saída $y(t)$ do processo em um experimento em malha aberta. O sinal de entrada compreende 2 períodos de onda quadrada com níveis 0 e 1, sendo que cada período é composto por 256 amostras. Assim sendo, foram coletadas 512 amostras dos sinais de entrada e saída. 100 rodadas de Monte Carlo foram executadas, em cada rodada o sinal de saída é contaminado por uma realização de um ruído colorido. Estas realizações são geradas através de realizações de ruído branco com variância $\sigma_v^2 = 0,09$ filtradas através do modelo descrito em (88). A Figura 18 apresenta o sinal de entrada e uma realização do sinal de saída.

Nesse exemplo, a estrutura de controle escolhida, (352), é igual à apresentada em (207). Portanto, a estrutura do filtro que aproxima as funções custo $J^{\text{CTLS}}(\rho)$ e $J^{\text{MR}}(\rho)$ é dada em (209). Note que o objetivo desse exemplo consiste em ilustrar a utilização do filtro, por conta disso, para simplificar sua aplicação, considera-se que o módulo de $G(q)$ é conhecido. Com o filtro escolhido, estimou-se os parâmetros do controlador com a solução DLS, apresentada no Teorema 3.4, para cada realização dos dados. As estimativas obtidas

Figura 18 — Sinal de entrada e uma realização do sinal de saída da simulação da Seção 5.2.2.

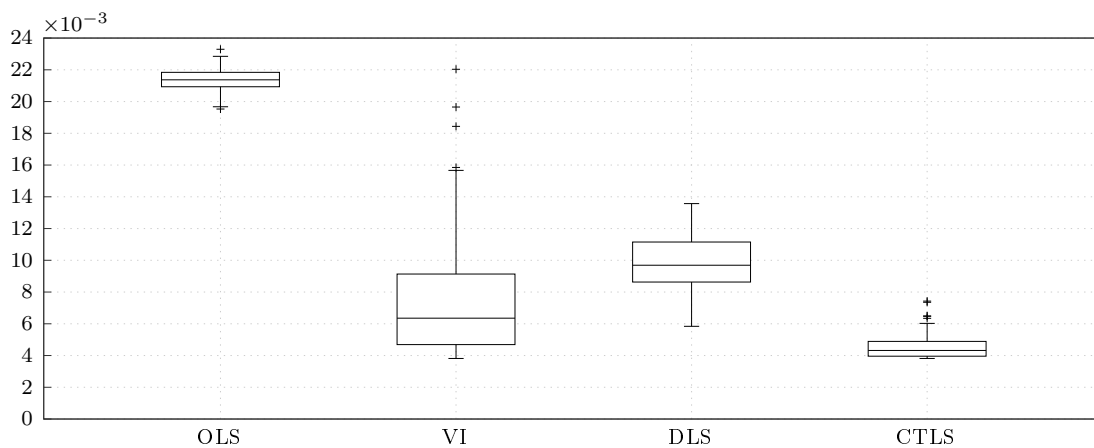


Fonte: elaborado pela autora.

com DLS foram utilizadas como inicialização do algoritmo de otimização empregado na solução do problema CTLS, descrito no Teorema 3.1, onde o filtro acima mencionado é utilizado na filtragem dos sinais $u(t)$ e $\bar{e}(t)$. Além disso, as soluções originais do método VRFT (OLS e VI) também foram empregadas na estimativa dos parâmetros do controlador, como apresentadas em (79) e (81), onde a estrutura de $L(q)$ é dada em (77).

A Figura 19 apresenta a distribuição das funções custo estimadas $\hat{J}^{\text{MR}}(\hat{\rho})$, como em (354) com $N = 50$, com os parâmetros ajustados com as soluções OLS, VI, DLS e CTLS. Observe que os maiores valores de $\hat{J}^{\text{MR}}(\hat{\rho})$ foram encontrados com OLS, um

Figura 19 — Distribuição da função custo estimada $\hat{J}^{\text{MR}}(\hat{\rho})$ com os parâmetros estimados com as soluções OLS, VI, DLS e CTLS na simulação da Seção 5.2.2.



Fonte: elaborado pela autora.

comportamento esperado, pois a mesma apresenta polarização na presença de ruído. Comparando-se os resultados das demais soluções com a OLS observa-se que a solução com VI apresenta valores aproximadamente 2,8 vezes menores, com a solução DLS encontrou-se resultados aproximadamente 2,2 vezes menores e a solução CTLS apresenta resultados aproximadamente 4,7 vezes menores do que os obtidos com OLS.

As estatísticas das estimativas encontradas com cada solução são apresentadas na Tabela 10, para OLS e VI, e na Tabela 11, para as soluções DLS e CTLS. Essas tabelas

também apresentam o vetor de parâmetros ótimo, que minimiza a função custo $J^{\text{MR}}(\rho_{\text{MR}}^*)$, obtido através da minimização de uma aproximação da função custo em (13)–(14), onde o sinal de referência empregado $r(t)$ é igual ao sinal $u(t)$ do experimento. A partir da

Tabela 10 — Parâmetros estimados com OLS e VI para a simulação da Seção 5.2.2.

ρ	ρ_{MR}^*	Estimativas	
		OLS	VI
ρ_1	0,4296	0,0826 ($\pm 0,0064$)	0,5076 ($\pm 0,3178$)
ρ_2	-0,3453	-0,0198 ($\pm 0,0080$)	-0,4261 ($\pm 0,3217$)

Tabela 10 observa-se que a média dos parâmetros estimados com VI é mais próxima da estimativa do vetor ótimo do que a média dos parâmetros com OLS. Além disso, a estimativa com VI apresenta uma variância bastante pronunciada e bem maior do que a encontrada com a solução OLS. Ao comparar esses resultados com os encontrados com DLS, na Tabela 11, observa-se que sua média é um pouco mais distante do vetor de parâmetros ótimo do que com VI, porém, com uma variância menor. Finalmente, veja que a média dos parâmetros com CTLS é bastante próxima do vetor ótimo e que a variância é bem menor do que com VI e DLS.

Tabela 11 — Média e desvio padrão dos parâmetros estimados com DLS e CTLS para a simulação da Seção 5.2.2.

ρ	ρ_{MR}^*	Estimativas	
		DLS	CTLS
ρ_1	0,4296	0,5971 ($\pm 0,1123$)	0,4318 ($\pm 0,0618$)
ρ_2	-0,3453	-0,5442 ($\pm 0,1120$)	-0,3448 ($\pm 0,0649$)

5.2.3 Caso de ordem cheia, dados de malha fechada e sequência de degraus como entrada

Nesse estudo de caso, um cenário mais realista é simulado, onde o processo em malha fechada é excitado com uma sequência de degraus com amplitudes e tamanhos diferentes. Os parâmetros do controlador são estimados com as soluções OLS, VI e CTLS.

Considere o modelo do processo em (84). O modelo de referência escolhido é descrito abaixo.

$$T_d(q) = \frac{0,27985(q - 0,8326)(q - 0,8)(q + 0,3502)}{(q - 0,2301)(q - 0,3906)(q^2 - 1,699q + 0,7264)}. \quad (359)$$

O controlador ideal foi calculado substituindo-se (84) e (359) em (16), obtendo-se:

$$C_d(q) = [0,5964 \quad 0,1265 \quad -0,1632] \left[1 \quad \frac{q}{q-1} \quad \frac{q-1}{q} \right]^T. \quad (360)$$

O objetivo consiste em projetar um controlador do tipo PID com a seguinte estrutura

$$C(q, \rho) = [K_p \quad K_i \quad K_d] \left[1 \quad \frac{q}{q-1} \quad \frac{q-1}{q} \right]^T. \quad (361)$$

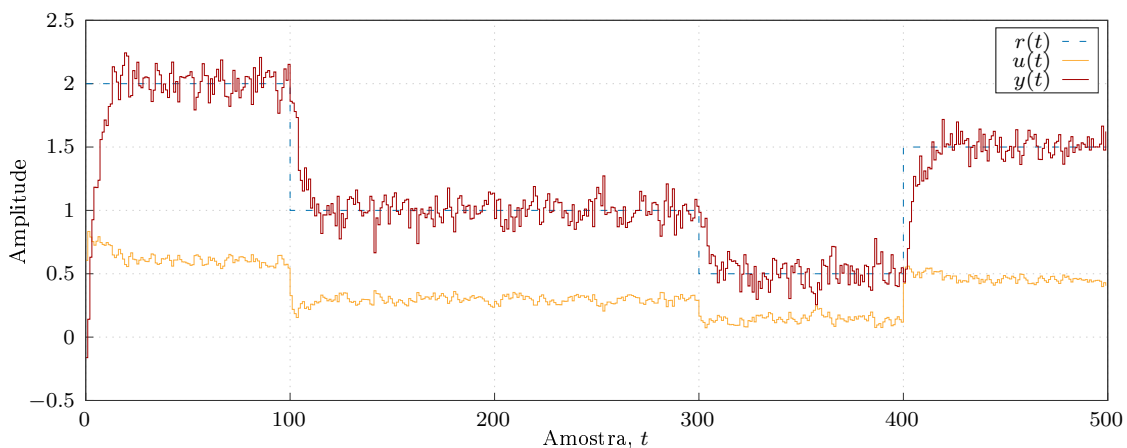
Uma vez que a estrutura escolhida é igual à estrutura do controlador ideal temos o caso em que o controlador ideal pertence à classe de controle.

Como mencionado, os dados são coletados de um experimento em malha fechada. Nesse caso, o controlador inicial operando na malha fechada enquanto os dados são coletados é dado por

$$C(q, \rho_0) = [0,2982 \quad 0,0633 \quad -0,0816] \begin{bmatrix} 1 & \frac{q}{q-1} & \frac{q-1}{q} \end{bmatrix}^T. \quad (362)$$

O sinal de referência excitando a entrada do processo em malha fechada é composto por uma sequência de degraus. Nesse caso, foram executadas 100 rodadas de Monte Carlo, onde em cada rodada a saída é corrompida por um ruído colorido, gerado a partir de um ruído branco com variância $\sigma_v^2 = 6,4 \times 10^{-3}$ filtrado através de (88). A Figura 20 apresenta o sinal de referência (linha tracejada azul) e uma realização dos sinais de entrada (linha contínua amarela) e do sinal de saída (linha contínua vermelha). Perceba que o sinal de

Figura 20 — Sinal de referência e uma realização dos sinais de entrada e saída da simulação da Seção 5.2.3.



Fonte: elaborado pela autora.

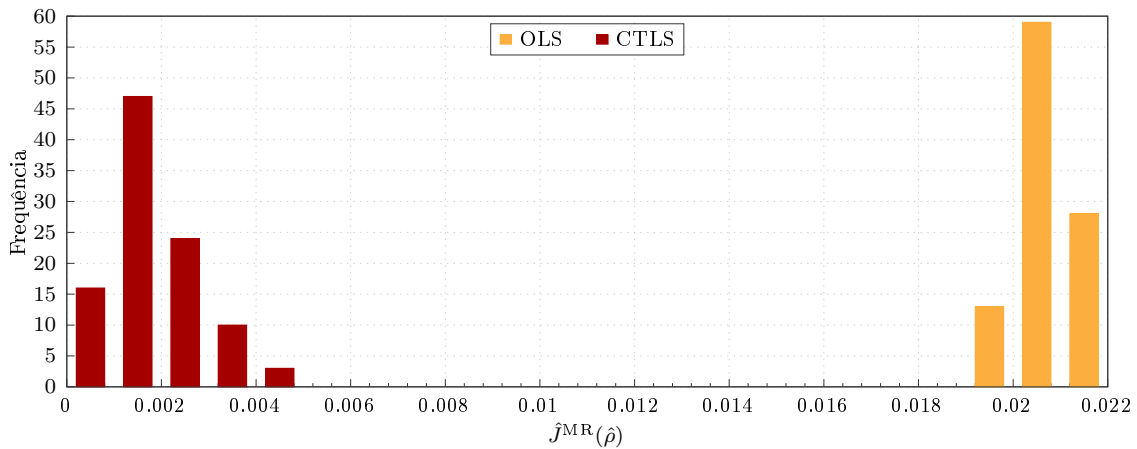
referência é composto por uma sequência de degraus com amplitudes e tamanhos (número de amostras) diferentes. Além disso, esse sinal possui apenas 500 amostras. Nesse caso, os parâmetros do controlador são ajustados empregando-se as soluções OLS, VI e CTLS. Vale mencionar que os parâmetros estimados com a solução OLS são utilizados para inicializar o problema de otimização do CTLS.

Para cada um dos controladores estimados calculou-se a função custo estimada em (354) com $N = 50$. Na Figura 21 são apresentados os histogramas das funções custo estimadas $\hat{J}^{\text{MR}}(\hat{\rho})$ calculadas com os controladores obtidos com OLS e CTLS. Nesse caso, os valores encontrados com a solução CTLS são aproximadamente 10 vezes menores do que os obtidos com OLS. Os resultados obtidos com a abordagem com VI não são apresentados pois 7% dos controladores resultaram em um comportamento instável em malha fechada.

A Tabela 12 apresenta o vetor de parâmetros ideal e as médias e desvios padrão dos parâmetros estimados com as soluções OLS e CTLS. Como pode ser notado, a solução CTLS apresenta uma polarização bem pequena, quando comparada com a solução OLS.

Como mencionado anteriormente, no caso de dados pouco informativos o efeito da polarização da estimativa OLS é mais pronunciado. Assim sendo, utilizar essa estimativa

Figura 21 — Histogramas das funções custo estimadas $\hat{J}^{\text{MR}}(\hat{\rho})$ calculadas com os controladores estimados com OLS e CTLS na simulação da Seção 5.2.3.



Fonte: elaborado pela autora.

Tabela 12 — Média e desvio padrão dos parâmetros estimados com OLS e CTLS para a simulação da Seção 5.2.3.

ρ	ρ_d	Estimativas	
		OLS	CTLS
K_p	0,5964	0,0326 ($\pm 0,0082$)	0,5933 ($\pm 0,0513$)
K_i	0,1265	0,1263 ($\pm 0,0005$)	0,1268 ($\pm 0,0016$)
K_d	-0,1632	-0,0237 ($\pm 0,0043$)	-0,1253 ($\pm 0,1531$)

para inicializar o algoritmo de otimização do problema do CTLS pode fazer com que a solução da otimização convirja para um mínimo local. Nesse caso específico, pode-se pré-selecionar os subconjuntos muito informativos utilizando um critério de seleção desses dados, como apresentado no Capítulo 4. Essa estimativa é menos polarizada e pode ser empregada para inicializar o algoritmo de otimização do CTLS.

5.3 Critérios de seleção de subconjuntos muito informativos de dados

5.3.1 Caso de ordem cheia, dados de malha aberta e entrada do tipo onda quadrada

Nesse estudo de caso, considere o seguinte processo em tempo contínuo (ÅSTRÖM; HÄGGLUND, 2000)

$$G(s) = \frac{1}{(s+1)^2}. \quad (363)$$

Esse modelo foi discretizado utilizando o método ZOH (do inglês *zero-order holder*) no Matlab com período de amostragem $T_s = 0,2s$, resultando no seguinte modelo em tempo discreto:

$$G(q) = \frac{0,087615(q+0,8752)}{(q-0,8187)^2}, \quad (364)$$

com tempo de acomodação de 30 amostras. Para alcançar erro zero em regime permanente para referências constantes e um tempo de acomodação um pouco menor do que em malha aberta, o seguinte modelo de referência foi escolhido, com tempo de acomodação de 25 amostras,

$$T_d(q) = \frac{0,071477(q + 0,8752)(q^2 - 1,634q + 0,6677)}{(q - 0,07276)(q - 0,8167)(q^2 - 1,677q + 0,7029)} \quad (365)$$

O controlador ideal, calculado substituindo-se (364) e (365) em (16), é dado por

$$C_d(q) = [0,2438 \quad 0,0273 \quad 0,5447] \begin{bmatrix} 1 & \frac{q}{q-1} & \frac{q-1}{q} \end{bmatrix}^T \quad (366)$$

$$= \frac{0,8158(q^2 - 1,634q + 0,6677)}{q(q-1)}. \quad (367)$$

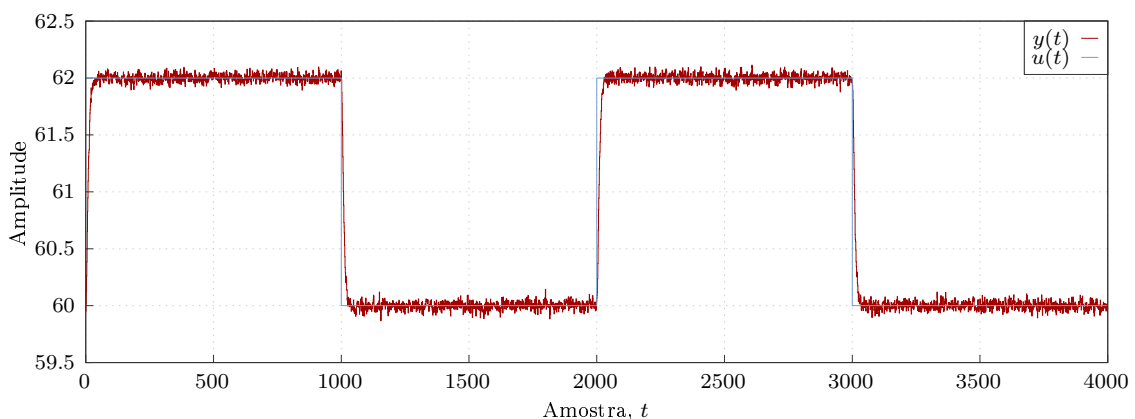
É importante mencionar que o modelo do processo não é empregado no projeto do controlador, sendo utilizado apenas para calcular o controlador ideal. Deseja-se estimar um controlador PID com a seguinte estrutura:

$$C(q, \rho) = [K_p \quad K_i \quad K_d] \begin{bmatrix} 1 & \frac{q}{q-1} & \frac{q-1}{q} \end{bmatrix}^T, \quad (368)$$

onde $[K_p \quad K_i \quad K_d]$ formam o vetor de parâmetros ρ a ser estimado. Veja que a estrutura de controle escolhida é igual à estrutura do controlador ideal, logo, a Suposição 2.1 é respeitada.

Nesse estudo de caso, 100 rodadas de Monte Carlo foram executadas, onde os dados de entrada e saída empregados no ajuste dos parâmetros do controlador são coletados de experimentos em malha aberta. O sinal de entrada $u(t)$ aplicado na entrada do processo é uma onda quadrada com período de 2000 amostras, os dados coletados compreendem 2 períodos dessa onda quadrada, totalizando 4000 amostras. A saída do processo $y(t)$ é corrompida por um ruído colorido aditivo, gerado através de um ruído branco com variância $\sigma_v^2 = 0,0225$ filtrado através de (88). A Figura 22 apresenta o sinal de entrada (linha contínua azul) e uma realização do sinal de saída (linha contínua vermelha). Observe

Figura 22 — Sinal de entrada e uma realização do sinal de saída da simulação da Seção 5.3.1.

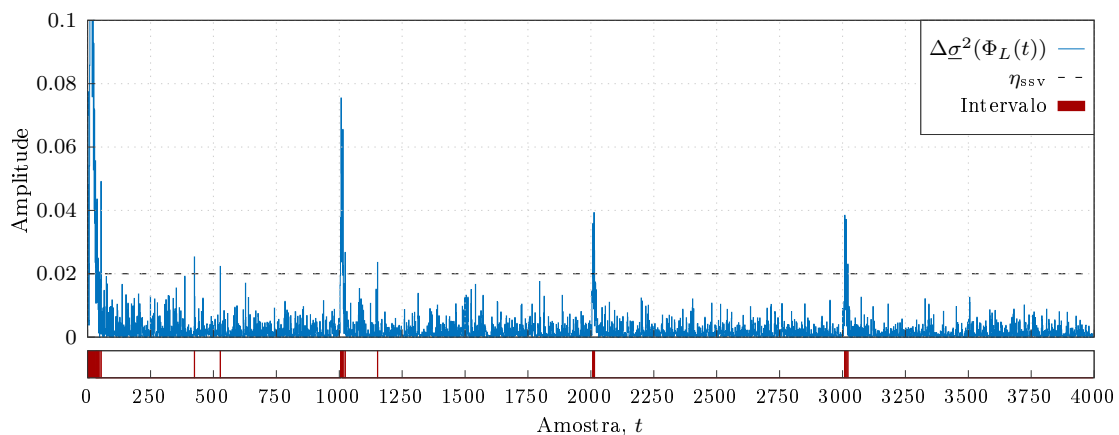


Fonte: elaborado pela autora.

que o sinal de saída foi dividido pelo ganho DC do processo para facilitar sua apresentação junto ao sinal de entrada.

Na Figura 23 é apresentada a curva da primeira diferença do quadrado do menor valor singular $\Delta\sigma^2(\Phi_L(t))$ (linha contínua azul) calculada utilizando os dados apresentados acima (mesma realização). Essa figura também apresenta o limiar escolhido η_{SSV} (linha tracejada preta) e os subconjuntos mais informativos (linhas verticais vermelhas). Como

Figura 23 — Curva de $\Delta\sigma^2(\Phi_L(t))$ para uma realização dos dados da simulação da Seção 5.3.1.



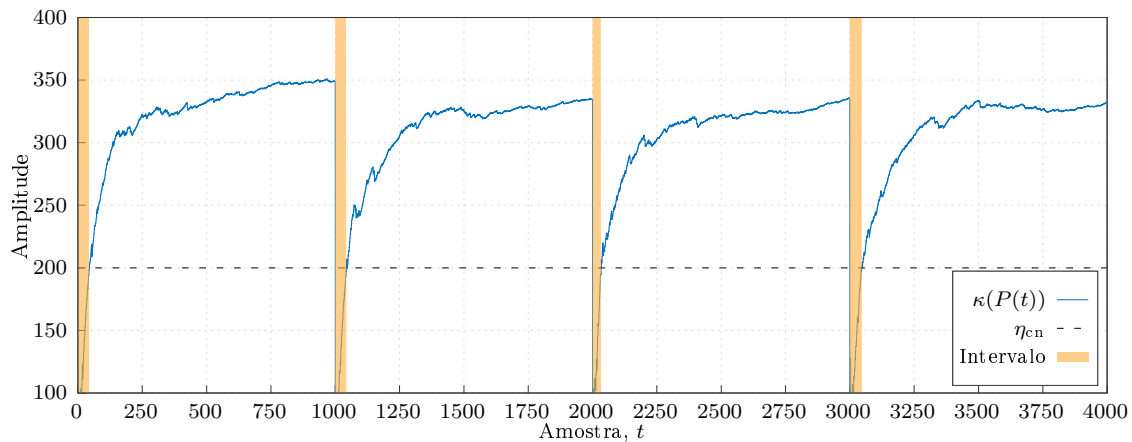
Fonte: elaborado pela autora.

mencionado anteriormente, esse gráfico pode ser utilizado para escolher um valor de limiar. Relembre que valores pequenos para $\Delta\sigma^2(\Phi_L(t))$ são obtidos para as linhas de Φ_L que são muito afetadas por ruído. Observe na Figura 23 como é possível escolher um limiar para remover essas linhas de Φ_L ao evitar os pequenos valores de $\Delta\sigma^2(\Phi_L(t))$. Nesse caso, escolheu-se $\eta_{SSV} = 0,02$, com esse valor de limiar a matriz resultante Φ_{SSV} obtida com o critério SSV possui 65 linhas informativas, em média. Ou seja, os parâmetros são estimados com aproximadamente 1,62% do conjunto total de dados.

Similarmente, o número de condicionamento $\kappa(P(t))$ foi calculado utilizando a mesma realização dos dados e é apresentado na Figura 24 (linha contínua azul). O valor de limiar escolhido $\eta_{cn} = 200$ e os intervalos selecionados são apresentados na mesma figura, através da linha tracejada preta e das linhas verticais amarelas, respectivamente. Em média, 210 amostras de Φ_L foram classificadas como informativas com o valor de limiar indicado. Nesse caso, os parâmetros do controlador são estimados com aproximadamente 5,25% de todo o conjunto de dados.

Para cada realização dos dados, selecionou-se a matriz regressora Φ_{SSV} e o vetor de saída u_{SSV} com o critério de seleção SSV, como apresentado no Algoritmo 1. Os parâmetros do controlador são estimados com as soluções OLS (116) e DLS (Teorema 3.4) empregando-se Φ_{SSV} e u_{SSV} . Similarmente, os dados de cada uma das 100 rodadas de Monte Carlo são empregados para selecionar a matriz regressora Φ_{cn} e o vetor de saída u_{cn} com o critério CN, descrito no Algoritmo 2. O ajuste dos parâmetros é obtido a partir de Φ_{cn} e u_{cn} as soluções OLS (116) e DLS (Teorema 3.4). Para comparar os resultados, também realizou-se a estimativa dos parâmetros com essas soluções e todo o conjunto de dados, para cada realização. Uma estimativa da função custo $J^{MR}(\rho)$ é calculada para cada controlador estimado em cada caso acima mencionado. Essa estimativa é calculada como em (354) onde $r(t)$ é um degrau com 50 amostras ($N=50$) e $T(q, \hat{\rho})$ é a função de sensibilidade

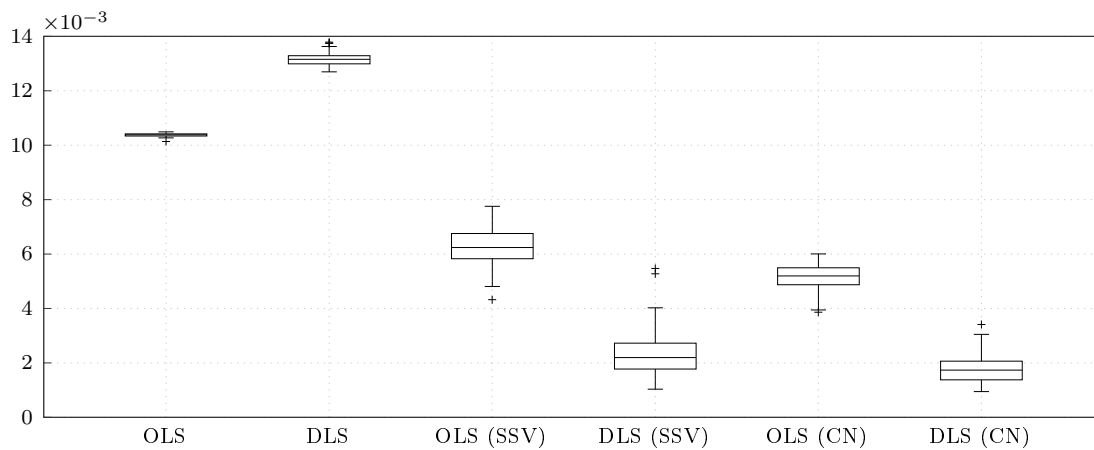
Figura 24 — Curva de $\kappa(P(t))$ para uma realização dos dados da simulação da Seção 5.3.1.



Fonte: elaborado pela autora.

complementar calculada como em (7) com o controlador $C(q, \hat{\rho})$ estimado em cada caso. A Figura 25 apresenta a distribuição da função custo estimada $\hat{J}^{MR}(\hat{\rho})$ com os controladores obtidos com as soluções OLS e DLS com o conjunto inteiro de dados e com os critérios de seleção. Primeiramente, observe que com os critérios de seleção foram obtidos resultados

Figura 25 — Distribuição da função custo estimada $\hat{J}^{MR}(\hat{\rho})$ com os parâmetros estimados com as soluções OLS e DLS com todo o dado e com os critérios de seleção da simulação da Seção 5.3.1.



Fonte: elaborado pela autora.

aproximadamente 2 vezes melhores do que utilizando todo o dado e a solução original do VRFT. Um comportamento similar é visto com a solução DLS, onde com os critérios de seleção apresentam resultados aproximadamente 5 e 7 vezes menores, critério SSV e CN, respectivamente, do que com todo o conjunto de dados. Comparando agora os resultados obtidos com os critérios, observa-se que a solução DLS resultou em valores menores do que OLS, independente do critério empregado.

Finalmente, as estatísticas dos parâmetros estimados são apresentadas nas Tabelas 13 e 14. Note que a média das estimativas obtidas com os critérios de seleção são mais próximas do controlador ideal do que com todo o conjunto de dados. Além disso, observa-se um aumento na variância das estimativas encontradas com os critérios, quando comparadas

Tabela 13 — Média e desvio padrão dos parâmetros estimados a solução OLS com todo o dado e com os subconjuntos mais informativos para a simulação da Seção 5.3.1.

ρ	ρ_d	Estimativas		
		OLS	OLS (SSV)	OLS (CN)
K_p	0,2438	0,2166 ($\pm 0,0017$)	0,2439 ($\pm 0,0044$)	0,2504 ($\pm 0,0027$)
K_i	0,0273	0,0272 ($\pm 0,0001$)	0,0265 ($\pm 0,0002$)	0,0267 ($\pm 0,0002$)
K_d	0,5447	-0,0753 ($\pm 0,0021$)	0,0971 ($\pm 0,0367$)	0,1649 ($\pm 0,0289$)

Tabela 14 — Média e desvio padrão dos parâmetros estimados a solução DLS com todo o dado e com os subconjuntos mais informativos para a simulação da Seção 5.3.1.

ρ	ρ_d	Estimativas		
		DLS	DLS (SSV)	DLS (CN)
K_p	0,2438	0,2779 ($\pm 0,0026$)	0,2614 ($\pm 0,0072$)	0,2359 ($\pm 0,0027$)
K_i	0,0273	0,0272 ($\pm 0,0001$)	0,0262 ($\pm 0,0003$)	0,0275 ($\pm 0,0001$)
K_d	0,5447	-0,2589 ($\pm 0,0096$)	0,4288 ($\pm 0,1273$)	0,6967 ($\pm 0,0476$)

com as estimativas com todo o dado. Esse aumento ocorre devido ao reduzido número de amostras empregadas. Porém, a qualidade da estimativa não é prejudicada, dada a grande redução na polarização. Isso pode ser visto nos resultados obtidos na Figura 25.

5.3.2 Caso de ordem reduzida, dados de malha fechada e entrada com perturbação do tipo pulso

Esse estudo de caso considera que o controlador ideal não pode ser representado com a estrutura de controle disponível. Por conta disso, como apresentado na Seção 4.3, somente o critério SSV pode ser aplicado para seleção dos subconjuntos muito informativos de dados. O processo em tempo contínuo é dado por (ÅSTRÖM; HÄGGLUND, 2000)

$$G(s) = \frac{(s+6)^2}{s(s+36)(s+1)^2},$$

discretizado utilizando o Matlab através do método ZOH, com tempo de amostragem de $T_s = 0,5s$, resultando no seguinte modelo em tempo discreto

$$G(q) = \frac{0,049562(q+0,7763)(q^2-0,1261q+0,00542)}{q(q-1)(q-0,6065)^2}. \quad (369)$$

Esse processo não é estável no sentido BIBO. Assim, os dados são coletados de uma simulação em malha fechada, com o seguinte controlador PD operando na malha fechada

$$C(q, \rho_0) = \frac{1,1451(q-0,7948)}{q}. \quad (370)$$

Nesse caso, o modelo de referência escolhido é dado por

$$T_d(q) = \frac{0,13}{q-0,87}. \quad (371)$$

O controlador ideal foi calculado com (369) e (371) em (16), obtendo-se

$$C_d(q) = \frac{2,623(q - 0,6065)^2 q}{(q + 0,7763)(q^2 - 0,1261q + 0,00542)}. \quad (372)$$

Os parâmetros de um controlador PD com estrutura fixa serão estimados. A parametrização do controlador é dada por

$$C(q, \rho) = [K_p \quad K_d] \left[1 \quad \frac{q-1}{q} \right]^T, \quad (373)$$

onde $[K_p \quad K_d]$ são os parâmetros ρ do controlador a ser estimado. Como a estrutura de controle escolhida não é capaz de representar exatamente o controlador ideal, os sinais são pré-filtrados através do filtro ótimo do VRFT, cuja estrutura é descrita em (77). O filtro empregado $L(q)$, calculado de (371), é apresentado abaixo

$$L(q) = \frac{0,13(q-1)}{(q-0,87)^2}. \quad (374)$$

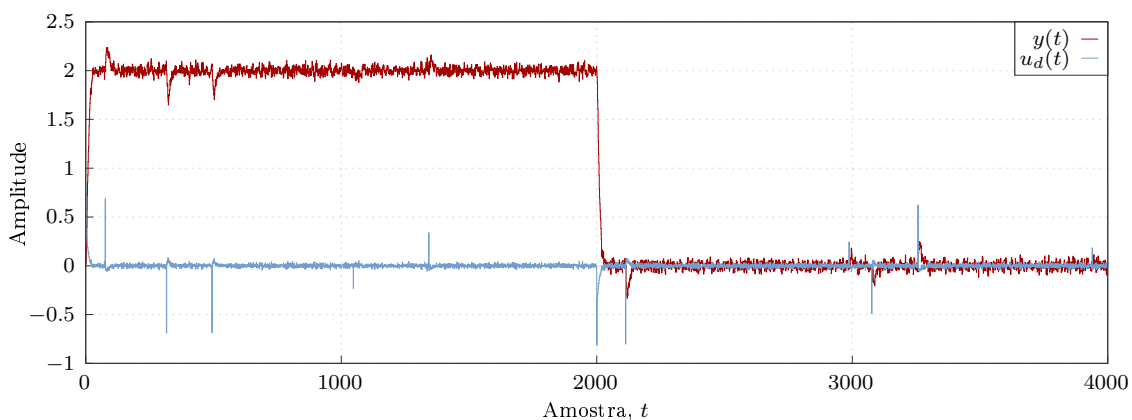
Nesse caso, 100 rodadas de Monte Carlo são executadas. Os dados de entrada $u_d(t)$ e saída $y(t)$ são dados por

$$y(t) = T(q, \rho_0)r(t) + S(q, \rho_0)H(q)v(t) + S(q, \rho_0)G(q)d(t), \quad (375)$$

$$u_d(t) = S(q, \rho_0)[C(q, \rho_0)r(t) - C(q, \rho_0)H(q)v(t) + d(t)], \quad (376)$$

onde $r(t)$ é uma onda quadrada com 2 períodos e amplitude variando entre 62 e 60, cada período possui 4000 amostras, $d(t)$ é uma perturbação estocástica, formada por pulsos de amplitudes diferentes aleatoriamente distribuídos nas amostras, enquanto $v(t)$ é ruído branco com variância $\sigma_v^2 = 9 \times 10^{-4}$. É importante mencionar que o sinal $u_d(t)$ é assumido desconhecido, e os parâmetros são estimados com $u(t)$ em (5) e $y(t)$ em (375). Ou seja, assume-se que a saída do controlador é conhecida e não a entrada do processo. A Figura 26 apresenta as primeiras 4000 amostras de uma realização do sinal perturbado $u_d(t)$ e do correspondente sinal de saída $y(t)$ (após a remoção do *offset*).

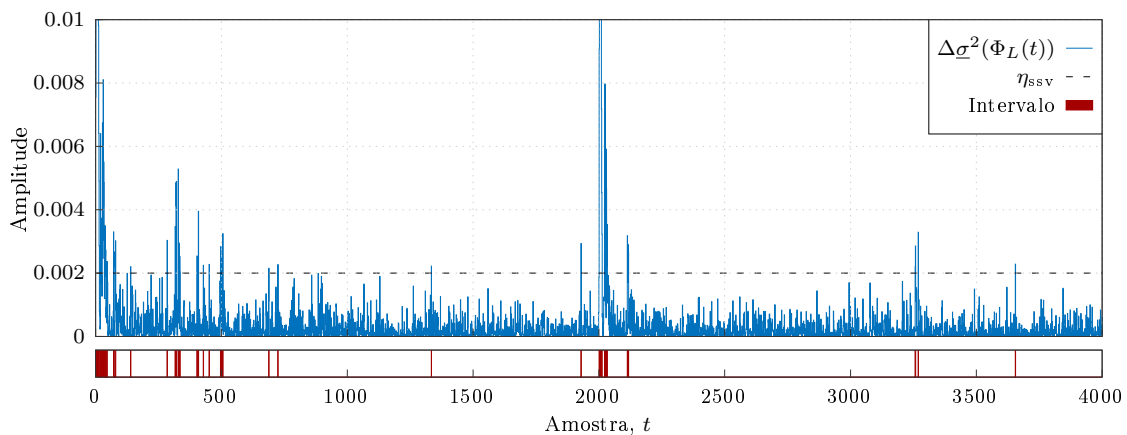
Figura 26 — Dados de entrada e saída da simulação da Seção 5.3.2.



Fonte: elaborado pela autora.

A Figura 27 apresenta as primeiras 4000 amostras da primeira diferença do quadrado do menor valor singular, $\Delta \underline{\sigma}^2(\Phi_L(t))$, calculado utilizando os dados da Figura 26. O limiar

Figura 27 — Curva de $\Delta\sigma^2(\Phi_L(t))$ para uma realização dos dados da simulação da Seção 5.3.2.

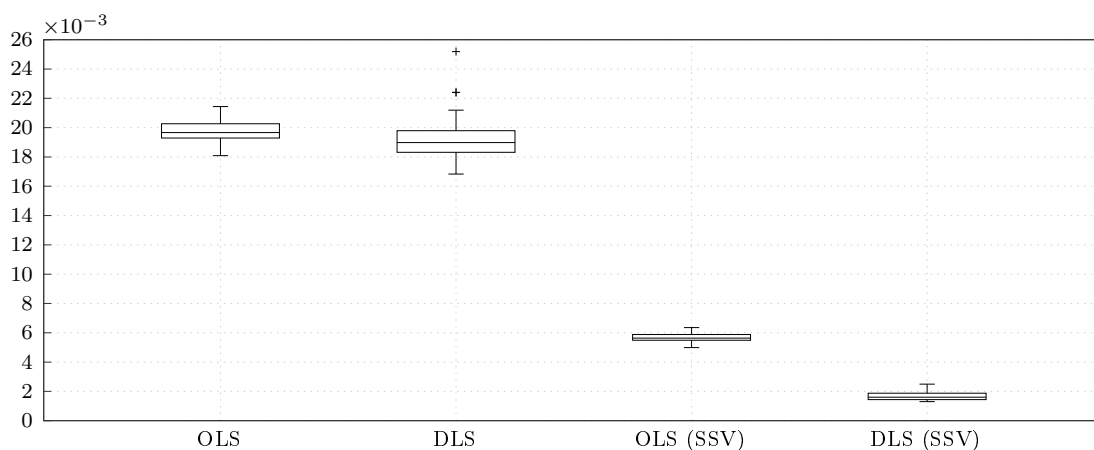


Fonte: elaborado pela autora.

$\eta_{SSV} = 0,02$ e os intervalos mais informativos também são apresentados nessa figura. Em média, a matriz resultante Φ_{SSV} possui 130 amostras, ou seja, os parâmetros do controlador são estimados com 1,57% do conjunto de dados inicial.

Novamente, para cada realização dos dados obteve-se uma matriz regressora Φ_{SSV} e um vetor de saída u_{SSV} , selecionados como descrito no Algoritmo 1. Essas informações foram utilizadas para ajustar os parâmetros do controlador com as soluções OLS e DLS, apresentadas em (116) e no Teorema 3.4, respectivamente. Para comparar os resultados foram estimados também os parâmetros do controlador com todo o conjunto de dados empregando-se as mesmas soluções. A Figura 28 apresenta a distribuição da função custo estimada $\hat{J}^{MR}(\hat{\rho})$ para cada um dos casos, calculada de (354) com $N = 100$. Observe que

Figura 28 — Distribuição da função custo estimada $\hat{J}^{MR}(\hat{\rho})$ com os parâmetros estimados com as soluções OLS e DLS com todo o dado e com o critério SSV da simulação da Seção 5.3.2.

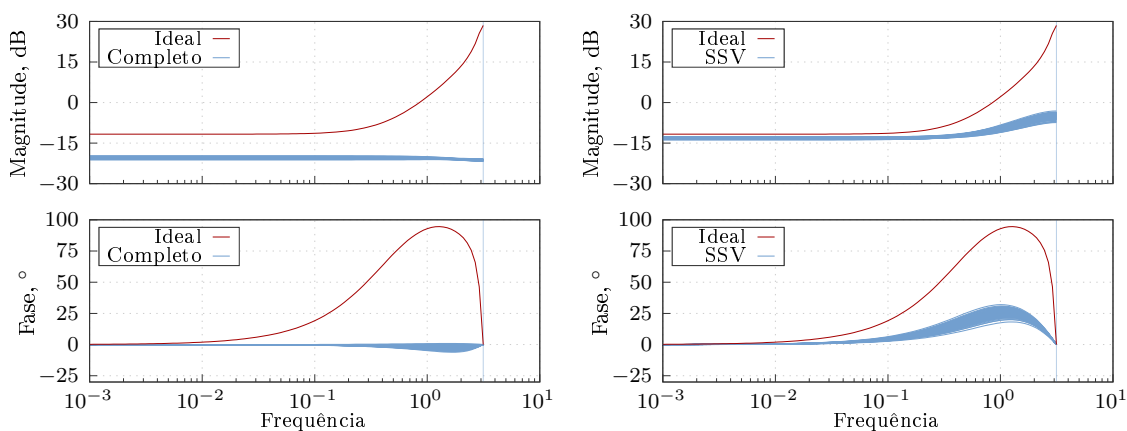


Fonte: elaborado pela autora.

o critério de seleção apresenta resultados aproximadamente 3,5 (OLS) e 11 (DLS) vezes menores do que com todo o conjunto de dados.

Na Figura 29 são apresentadas as respostas em frequência dos controladores estimados com a solução OLS. Claramente, não é esperado que as respostas obtidas sejam exatamente

Figura 29 — Diagramas de Bode dos controladores estimados com a solução OLS, com os intervalos mais informativos, e completo (com todo o dado) da simulação da Seção 5.3.2.



(a) OLS – Completo (todo o dado).

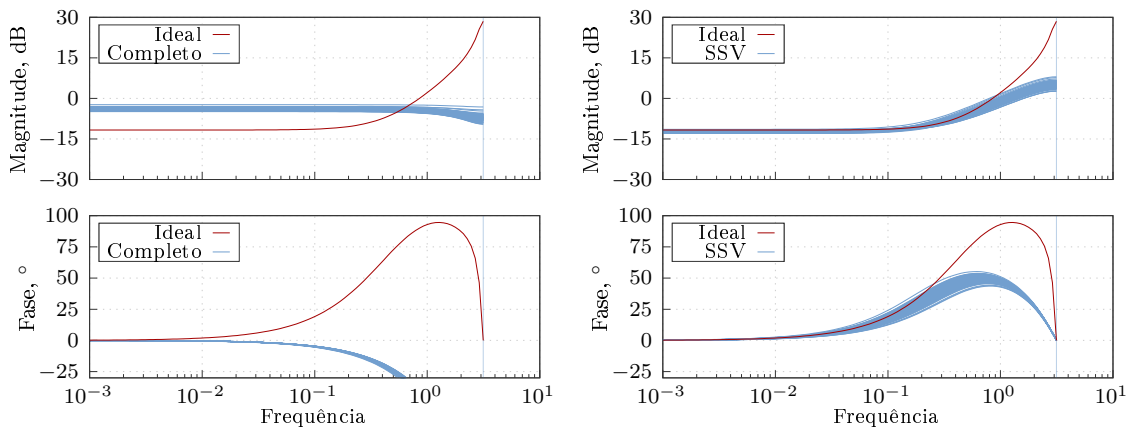
(b) OLS – SSV.

Fonte: elaborado pela autora.

iguais à resposta do controlador ideal, pois aqui considera-se o caso em que o controlador ideal não pertence à classe de controle. Contudo, observe que as respostas encontradas com os subconjuntos mais informativos, Figura 29b, são mais próximas da resposta ideal, do que utilizando todo o conjunto de dados, Figura 29a, principalmente quando compara-se os gráficos da magnitude. Observe também na Figura 29a que os controladores estimados são praticamente um ganho, pois o zero estimado dos controladores é muito próximo de zero.

Similarmente, a Figura 30 apresenta as respostas em frequência dos controladores estimados através da solução DLS. Nesse caso, a melhora obtida utilizando os subconjuntos

Figura 30 — Diagramas de Bode dos controladores estimados com a solução DLS, com os intervalos mais informativos, e completo (com todo o dado) da simulação da Seção 5.3.2.



(a) DLS – Completo (todo o dado).

(b) DLS – SSV.

Fonte: elaborado pela autora.

mais informativos é mais evidente, principalmente ao comparar os gráficos da fase nas Figuras 30a e 30b.

As médias e desvios padrão dos parâmetros estimados com as soluções OLS e DLS com todo o conjunto de dados e com os intervalos mais informativos selecionados com o critério SSV são apresentados nas Tabelas 15 e 16, respectivamente. Observe que as estimativas encontradas com os subconjuntos mais informativos de dados são mais próximas do vetor de parâmetros ótimo do que as estimativas obtidas com todo o conjunto de dados.

Tabela 15 — Média e desvio padrão dos parâmetros estimados com a solução OLS com todo o dado e com os subconjuntos mais informativos para a simulação da Seção 5.3.2.

ρ	ρ_{MR}^*	Estimativas	
		OLS (Completo)	OLS (SSV)
K_p	0,2326	0,0952 (\pm 0,0032)	0,2188 (\pm 0,0061)
K_d	0,9159	-0,0045 (\pm 0,0021)	0,1692 (\pm 0,0264)

Tabela 16 — Média e desvio padrão dos parâmetros estimados com a solução DLS com todo o dado e com os subconjuntos mais informativos para a simulação da Seção 5.3.2.

ρ	ρ_{MR}^*	Estimativas	
		DLS (Completo)	DLS (SSV)
K_p	0,2326	0,6296 (\pm 0,0337)	0,2442 (\pm 0,0076)
K_d	0,9159	-0,5355 (\pm 0,0456)	0,7776 (\pm 0,1089)

5.3.3 Caso de ordem cheia, dados de malha aberta e erro de medida no sinal de saída

Esse estudo de caso considera o caso em que ocorrem erros de medida no sinal de saída. Como será apresentado, nesse caso um ajuste melhor dos parâmetros do controlador é obtido quando os subconjuntos mais informativos são selecionados através da combinação dos critérios SSV e CN. Considere o seguinte modelo para o processo em tempo contínuo

$$G(s) = \frac{1}{s+1},$$

discretizado através do método ZOH no Matlab, com tempo de amostragem $T_s = 0,1$ s, obtendo-se o seguinte modelo em tempo discreto

$$G(q) = \frac{0,19033}{(q-0,9048)}, \quad (377)$$

com tempo de acomodação de 40 amostras. Para alcançar erro zero em regime permanente para referências constantes e um tempo de acomodação menor do que em malha aberta, o seguinte modelo de referência foi escolhido, com tempo de acomodação de 18 amostras,

$$T_d(q) = \frac{0,2}{q-0,8}. \quad (378)$$

O controlador ideal foi calculado substituindo-se (377) e (378) e em (16), obtendo-se:

$$C_d(q) = [0,9508 \quad 0,1000] \left[1 \quad \frac{q}{q-1} \right]^T = \frac{1,0508(q-0,9048)}{q-1} \quad (379)$$

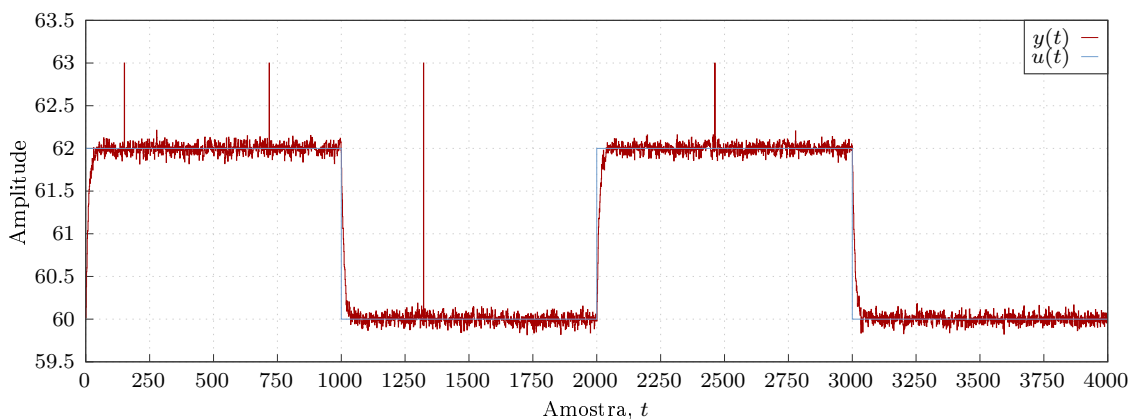
Nesse exemplo, deseja-se estimar um controlador PI com a seguinte estrutura

$$C(q, \rho) = [K_p \quad K_i] \left[1 \quad \frac{q}{q-1} \right]^T. \quad (380)$$

Por conta disso, tem-se o caso em que o controlador ideal pertence à classe de controle. Novamente, é importante mencionar que o modelo do processo é empregado apenas no cálculo do controlador ideal, não sendo utilizado no projeto do controlador.

Os dados de entrada $u(t)$ e saída $y(t)$ são coletados de um experimento em malha aberta e 100 rodadas de Monte Carlo são executadas. O sinal de entrada aplicado ao processo é uma onda quadrada com níveis 60 e 62 e período de 2000 amostras, os dados coletados compreendem 2 períodos dessa onda quadrada, totalizando 4000 amostras. O sinal de entrada e uma realização do sinal de saída são apresentados na Figura 31. Para

Figura 31 — Dados de entrada e saída da simulação da Seção 5.3.3.



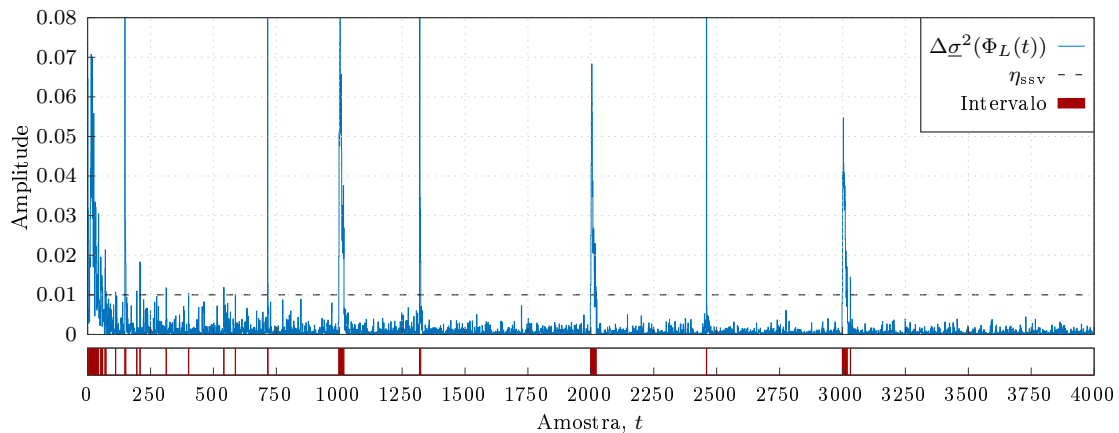
Fonte: elaborado pela autora.

cada realização esses erros de medida ocorrem em amostras aleatoriamente escolhidas. Veja que o sinal de saída é apresentado após a remoção do *offset*.

A primeira diferença do quadrado do menor valor singular, $\Delta\sigma^2(\Phi_L(t))$, foi calculada com os dados apresentados acima e é mostrada na Figura 32. Observe nessa figura que $\Delta\sigma^2(\Phi_L(t))$ apresenta picos que não provém do sinal de entrada $u(t)$, mas do erro de medida da saída. Por conta disso, essas linhas da matriz regressora Φ_L são equivocadamente classificadas como informativas pelo critério. A Figura 32 também apresenta o limiar escolhido $\eta_{ssv} = 0,01$ e os intervalos mais informativos. Em média, 110 linhas da matriz original Φ_L foram classificadas como informativas com o limiar indicado. Isso significa que os parâmetros do controlador são estimados com 2,83% dos dados disponíveis.

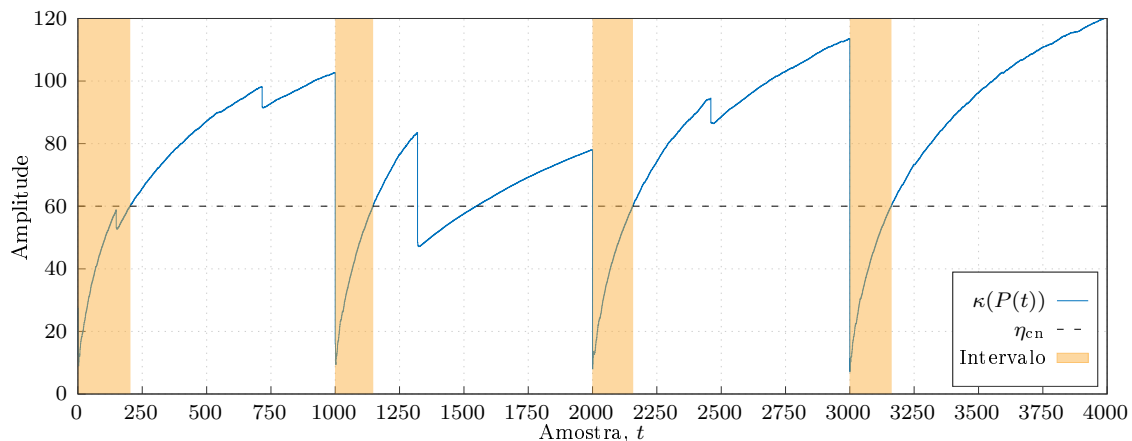
De maneira similar, a Figura 33 apresenta o número de condicionamento calculado utilizando o mesmo dado. O valor de limiar $\eta_{cn} = 60$ e os intervalos selecionados também são apresentados na mesma figura. Em média, a matriz resultante Φ_{cn} possui 790 linhas, ou seja, os parâmetros do controlador são estimados com aproximadamente 19,76% do conjunto inteiro de dados. Observe que o número de condicionamento não apenas aumenta com o número de amostras. De fato, há partes em que ele decresce. Esse comportamento

Figura 32 — Curva de $\Delta\sigma^2(\Phi_L(t))$ para uma realização dos dados da simulação da Seção 5.3.3.



Fonte: elaborado pela autora.

Figura 33 — Curva de $\kappa(P(t))$ para uma realização dos dados da simulação da Seção 5.3.3.



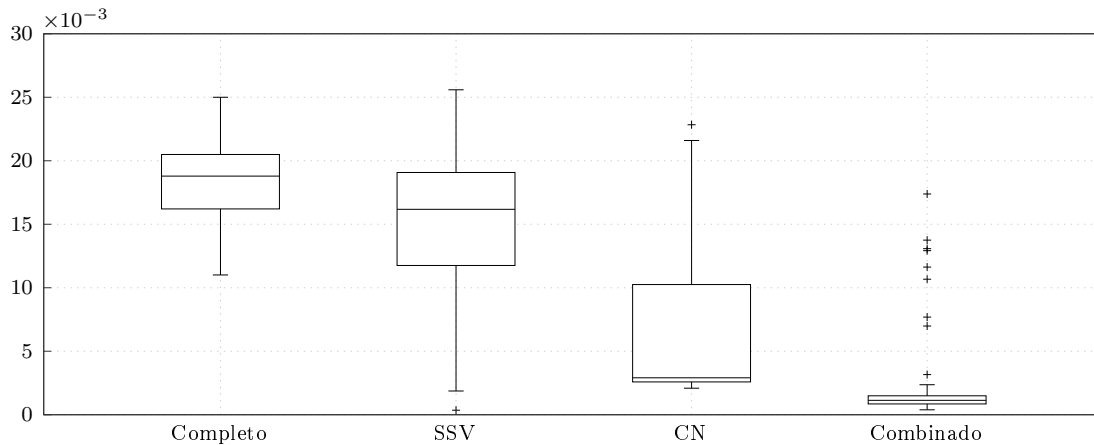
Fonte: elaborado pela autora.

é decorrente do efeito dos erros de leitura no sinal de saída. Contudo, note que essas amostras nem sempre são classificadas como informativas, uma vez que o critério é aplicado individualmente para cada degrau e seleciona as amostras enquanto o número de condicionamento é menor do que o limiar. Claramente, se os erros de leitura ocorrerem próximo ao início do degrau (veja o comportamento das primeiras 250 amostras no gráfico) as componentes da matriz regressora formadas por essas amostras serão classificadas como muito informativas.

Como pode ser observado, nos casos em que há erros de medida no sinal de saída o critério SSV sempre classifica os intervalos onde ocorrem esses erros, enquanto o critério CN é menos afetado pelos mesmos, como mencionado acima. Por conta disso, no presente trabalho, sugere-se que a escolha dos subconjuntos mais informativos nesse caso seja realizada através da combinação dos dois critérios. Isso é realizado da seguinte forma: as linhas de Φ_{SSV} são mantidas se também corresponderem às linhas de Φ_{CN} , dessa forma, evitando que os picos em $\Delta\sigma^2(\Phi_L(t))$ provenientes da saída perturbada sejam utilizados na estimativa dos parâmetros. O efeito benéfico da utilização dos critérios combinados

pode ser visto nas Figuras 34 e 35. Essas figuras apresentam as estimativas da função custo

Figura 34 — Distribuição da função custo estimada $\hat{J}^{\text{MR}}(\hat{\rho})$ com os parâmetros estimados com a solução OLS para cada caso da simulação da Seção 5.3.3.

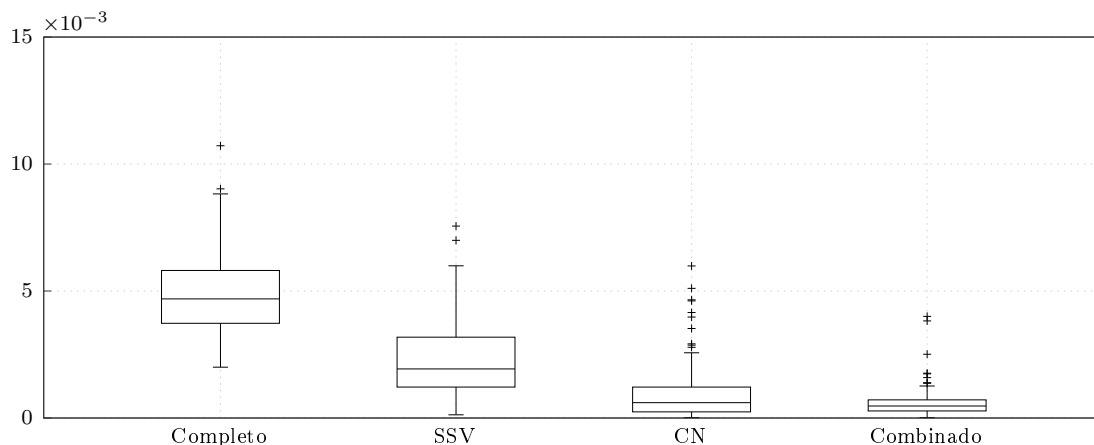


Fonte: elaborado pela autora.

$\hat{J}^{\text{MR}}(\hat{\rho})$ calculadas como em (354), com $N = 50$, para cada um dos controladores estimados em cada realização. Primeiramente, observe os resultados obtidos com a solução OLS da Figura 34. Note que o critério CN apresenta valores aproximadamente 3 vezes menores para a função custo estimada $\hat{J}^{\text{MR}}(\hat{\rho})$ do que com todo o conjunto de dados. Contudo, ao combinar os critérios obteve-se resultados aproximadamente 9 vezes melhores.

Esse comportamento também é observado com a solução DLS, cujos resultados são apresentados na Figura 35. Perceba que com os critérios SSV e CN foram obtidos

Figura 35 — Distribuição da função custo estimada $\hat{J}^{\text{MR}}(\hat{\rho})$ com os parâmetros estimados com a solução DLS para cada caso da simulação da Seção 5.3.3.



Fonte: elaborado pela autora.

resultados, respectivamente, 2 e 5 vezes menores para a função custo estimada, enquanto que a combinação dos critérios gerou resultados aproximadamente 8 vezes melhores do que os encontrados com o conjunto completo de dados. Finalmente, ao comparar os resultados das Figuras 34 e 35, observa-se que a solução DLS apresentou valores menores para $\hat{J}^{\text{MR}}(\hat{\rho})$ do que a solução OLS, mesmo no caso em que todo o conjunto de dados é utilizado na estimativa dos parâmetros do controlador.

Na Tabela 17 são apresentadas as médias e os desvios padrão dos parâmetros estimados com as soluções OLS e DLS dos parâmetros estimados com todo o conjunto de dados (completo), com os subconjuntos mais informativos selecionados pelos critérios SSV e CN, e com os subconjuntos encontrados através da combinação desses critérios. Observe que os parâmetros estimados com os intervalos selecionados através da combinação dos critérios apresentam a menor polarização, quando comparados com os demais resultados.

Tabela 17 — Média e desvio padrão dos parâmetros estimados com as soluções OLS e DLS com os dados da simulação da Seção 5.3.3.

ρ	K_p	K_i
ρ_d	0,9508	0,100
Estimativas OLS		
Completo	0,2492 (\pm 0,0813)	0,1000 (\pm 0,0006)
SSV	0,3556 (\pm 0,1847)	0,0959 (\pm 0,0020)
CN	0,6751 (\pm 0,1941)	0,1008 (\pm 0,0009)
Combinado	0,8515 (\pm 0,1296)	0,0998 (\pm 0,0025)
Estimativas DLS		
Completo	1,3088 (\pm 0,1495)	0,0998 (\pm 0,0006)
SSV	1,0955 (\pm 0,1192)	0,0957 (\pm 0,0039)
CN	1,0125 (\pm 0,0860)	0,1000 (\pm 0,0008)
Combinado	0,9783 (\pm 0,0435)	0,0992 (\pm 0,0023)

5.4 Considerações finais

Nesse capítulo foram apresentados os resultados de simulação gerados com o objetivo de validar as soluções propostas no presente trabalho. Tais soluções compreendem o emprego de outras soluções de mínimos quadrados ao método VRFT, para lidar com dados ruidosos, e o emprego de dois critérios de seleção de subconjuntos muito informativos de dados ao mesmo método, visando possibilitar sua aplicação quando dados pouco informativos estão disponíveis.

Assim sendo, primeiramente, o método VRFT foi formulado considerando-se um controlador representado com uma estrutura semelhante à estrutura ARX. Essa escolha confere mais flexibilidade ao controlador, do que as escolhas comumente encontradas na literatura. Além disso, foi proposta uma formulação que possibilita fixar parte do controlador, com as informações conhecidas *a priori*, reduzindo o número de parâmetros a serem estimados. Para lidar com o ruído de medida, foi proposto o emprego das soluções DLS e CTLS para estimar os parâmetros do controlador, ao invés das soluções originais (OLS e VI). A viabilidade das soluções propostas foi apresentada em três estudos de caso simulados. No primeiro estudo de caso, os resultados foram comparados com as soluções originais e com o método OCI, cuja estimativa é não polarizada e estatisticamente eficiente. Uma comparação entre aqueles resultados mostrou que o CTLS apresentou resultados significativamente melhores do que as soluções originais do VRFT, ou seja, com CTLS não foram encontrados controladores que instabilizaram o processo em malha fechada como com VI e os resultados foram aproximadamente 10 vezes melhores do que com OLS.

Além disso, a solução CTLS apresentou resultados similares quando comparada ao método OCI. O segundo estudo de caso foi desenvolvido com o objetivo de validar a abordagem proposta para a solução CTLS no caso em que o controlador ideal não pertence à classe de controle. Naquele exemplo, foi apresentado um comparativo dos resultados com DLS e CTLS e as soluções originais do método VRFT. Os resultados mostraram que com DLS e CTLS obteve-se resultados aproximadamente 2,2 e 4,7 vezes menores do que com OLS, respectivamente. Além disso, a média dos parâmetros estimados com a solução CTLS foi mais próxima do vetor de parâmetros ótimo e também apresentou menor variância do que com VI. Finalmente, no terceiro estudo de caso, foi simulado um cenário mais próximo do que pode ser encontrado em uma situação prática, ou seja, o sinal de referência era composto por um reduzido número de amostras e por uma sequência de degraus. Um comparativo dos resultados mostrou que o CTLS apresentou resultados muito melhores do que as soluções originais, comparado com OLS os resultados foram aproximadamente 10 vezes melhores e não obteve-se controladores que instabilizaram o processo, ao contrário da solução com VI.

Quanto à utilização de dados pouco informativos no método VRFT, foi proposto o emprego de dois critérios de seleção de subconjuntos mais informativos. Para estimar os parâmetros do controlador com os intervalos mais informativos selecionados, além da solução OLS, também foi proposta a utilização da solução DLS. As abordagens propostas foram validadas através de três estudos de caso simulando cenários parecidos com o que se pode encontrar em situações práticas. No primeiro estudo de caso considerou-se o caso em que o controlador ideal pertence à classe de controle. Daqueles resultados, observou-se que com os critérios de seleção obteve-se resultados aproximadamente 2 vezes melhores do que com todo o conjunto de dados para o OLS, enquanto que com a solução DLS os resultados foram entre 5 e 7 vezes melhores. O segundo estudo de caso considerou que os dados são coletados em malha fechada, perturbações do tipo pulso afetam a entrada de controle e que o controlador a ser projetado é de ordem reduzida. Naquele caso, foi empregado apenas o critério SSV e dos resultados observou-se que com o critério de seleção obteve-se valores aproximadamente 3,5 (OLS) e 11 (DLS) vezes melhores do que com as respectivas soluções e todo o conjunto de dados. No último estudo de caso foi considerado que havia erro de medida no sinal de saída do processo. Naquele caso, foi proposta a combinação dos critérios de seleção. Um comparativo dos resultados mostrou que com essa abordagem foram encontrados resultados aproximadamente 9 (OLS) e 8 (DLS) vezes melhores do que os obtidos com as mesmas soluções e todo o conjunto de dados. Finalmente, dos resultados conclui-se que é sempre vantajoso selecionar os subconjuntos mais informativos, pois com eles obteve-se melhores desempenhos em malha fechada do que com todo o conjunto de dados.

6 CONCLUSÕES E QUESTÕES EM ABERTO

No presente trabalho foram apresentadas duas modificações para lidar com problemas distintos no método VRFT. A primeira modificação consiste no emprego de outras soluções de mínimos quadrados ao método. A segunda modificação é a aplicação ao método VRFT de dois critérios de seleção de subconjuntos muito informativos de dados. Na sequência, são apresentadas algumas conclusões referentes a realização desse trabalho e algumas questões que ficaram em aberto.

6.1 Conclusões

No método VRFT, a estimativa dos parâmetros do controlador é facilmente encontrada através da solução do problema de mínimos quadrados OLS, contanto que a estrutura do controlador seja linearmente parametrizada. Contudo, essa estimativa é polarizada se há presença de ruído nos dados coletados. A polarização da estimativa é intrínseca à formulação do problema e está presente mesmo no caso de controladores de ordem cheia, uma vez que o problema posto pelo método não é um problema de identificação padrão. O efeito do ruído é usualmente contornado pela aplicação do método das variáveis instrumentais, que apesar de reduzir a polarização da estimativa possui o inconveniente de aumentar a variância da mesma, podendo acarretar em perda de desempenho e até mesmo em instabilidade em malha fechada. Além disso, a qualidade da estimativa depende da informatividade dos dados empregados no ajuste dos parâmetros. Por isso, é usual a execução de um experimento específico para coleta de dados, o que pode ser uma tarefa custosa. Para evitar essa tarefa, uma alternativa consiste na utilização de dados coletados durante a execução normal do processo, porém, a simples aplicação desses dados em um método baseado em dados pode levar a um ajuste inapropriado dos parâmetros.

Para lidar com o caso ruidoso, foi proposta a aplicação das soluções DLS e CTLS ao método VRFT. Como mencionado anteriormente, no caso de dados de malha aberta a solução DLS descreve melhor o problema posto pelo método, pois o ruído está presente apenas na entrada do controlador a ser identificado. Já no caso em que os dados são coletados em malha fechada a solução CTLS se adequa melhor ao problema. O desenvolvimento da solução CTLS foi realizado considerando controladores com polos livres e fixos. Além disso, foi considerado não apenas o caso de controladores de ordem cheia, mas também o caso de controladores de ordem reduzida, para o qual foram desenvolvidos os filtros empregados na pré-filtragem dos sinais. As soluções propostas foram comparadas com as soluções originais e com o método OCI, que fornece uma solução não polarizada e estatisticamente eficiente. Um comparativo dos resultados mostrou que as soluções propostas apresentaram resultados significativamente melhores do que as soluções originais e semelhantes aos obtidos com o OCI.

Quanto ao caso em que os dados disponíveis são pouco informativos, foi proposto o emprego no método VRFT de dois critérios de seleção de subconjuntos muito informativos de dados, inicialmente desenvolvidos para identificação de sistemas. Primeiramente, esses critérios foram adaptados para o caso em que o controlador possui ordem cheia. Em seguida, o comportamento dos mesmos foi avaliado no caso de controladores de ordem reduzida, no qual apenas um dos critérios pode ser aplicado com sucesso. O emprego da solução DLS com os dados selecionados pelos critérios também foi explorado. Além disso, a aplicação desses critérios também foi explorada nos casos em que as condições ideais referentes aos sinais coletados não são respeitadas, ou seja, dados de malha fechada, com perturbações e no caso em que o controlador possui ordem reduzida. Os resultados obtidos com as soluções propostas foram comparados com os resultados obtidos utilizando-se todo o conjunto de dados na estimativa dos parâmetros. Ao comparar os resultados, observou-se que as estimativas obtidas com os dados selecionados pelos critérios de seleção apresentaram melhores resultados do que as obtidas com todo o conjunto de dados, em todos os estudos de caso executados. Além disso, pode-se observar que no caso de perturbações no sinal de saída a combinação dos critérios apresentou melhores resultados que os obtidos com os critérios em separado.

6.2 Questões em aberto

Embora tenha-se tentado cobrir o maior número de pontos possíveis, alguns tópicos permanecem em aberto e podem ser assunto de interesse de pesquisas futuras. Esses tópicos são discutidos abaixo.

- (a) **escolha automática dos valores de limiar dos critérios de seleção.** Como mencionado no texto, para o critério SSV a escolha do limiar pode ser realizada através da análise gráfica da evolução da primeira diferença do quadrado do menor valor singular, enquanto que para o critério CN pode-se empregar o valor proposto na literatura e ajustá-lo dependendo do caso. Contudo, seria interessante que a escolha dos limiares fosse realizada automaticamente; Para o critério SSV, uma opção a ser investigada consiste em analisar qual a distribuição de densidade de probabilidade que melhor descreve a primeira diferença do quadrado do menor valor singular, e então utilizar as propriedades estatísticas desta distribuição para escolher um valor de limiar, por exemplo, o limiar poderia ser a média mais n vezes o desvio padrão. Já para o critério CN, uma opção seria calcular a primeira diferença do número de condicionamento e empregar a mesma sugestão descrita acima para o critério SSV;
- (b) **formulação matemática do critério SSV para dados de malha fechada.** Embora esse critério tenha apresentado bons resultados no caso de dados de malha fechada, seria interessante que a formulação matemática do mesmo levasse em consideração esse tipo de dado, isto é, uma prova formal;
- (c) **extensão dos critérios de seleção para lidar com perturbações nos sinais.** Em dados de processos reais é comum que ocorram perturbações nos sinais coletados. Portanto, seria bastante útil que os critérios de seleção lidassem naturalmente com esse tipo de dado;
- (d) **análise estatística das propriedades das soluções CTLS e DLS.** Essas soluções são empregadas considerando-se uma única realização dos dados, portanto, não foi

realizada uma análise estatística das estimativas. Seria interessante o desenvolvimento dessa análise estatística das propriedades dessas soluções;

- (e) **extensão da solução DLS para o caso em que o controlador possui ordem reduzida.** A solução DLS foi aplicada no caso de controladores de ordem reduzida utilizando-se como filtro o desenvolvido para a solução CTLS, apresentando bons resultados. Contudo, seria interessante que o desenvolvimento do filtro fosse realizado considerando as peculiaridades dessa solução;
- (f) **extensão da solução CTLS para o caso multivariável.** A solução CTLS foi aplicada no caso monovariável, onde considerou-se controladores de ordem reduzida e foram desenvolvidos os filtros para lidar com este caso. Porém, seria interessante estender tal solução para o caso multivariável considerando-se também o caso de controladores de ordem reduzida;
- (g) **analisar a aplicação das soluções considerando processos com tempo morto.** É comum que processos reais apresentem tempo morto. Portanto, seria interessante investigar a aplicação das soluções propostas nesse caso.

REFERÊNCIAS

ABATZOGLOU, T.; MENDEL, J. Constrained total least squares. *In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING*, 1987, Dallas, TX. *Proceedings[...]*. Piscataway, NJ: IEEE, 1987. p. 1485–1488. Disponível em: <https://doi.org/10.1109/ICASSP.1987.1169438>. Acesso em: 24 de junho de 2021.

ABATZOGLOU, T. J.; HARADA, G. A.; SHINE, M. J. Total least squares techniques for high resolution direction finding. *In: MILITARY COMMUNICATIONS CONFERENCE*, 1988, San Diego, CA. *Proceedings[...]*. Piscataway, NJ: IEEE, 1988. p. 405–409. Disponível em: <https://doi.org/10.1109/MILCOM.1988.13423>. Acesso em: 24 de junho de 2021.

ABATZOGLOU, T. J.; MENDEL, J. M.; HARADA, G. A. The constrained total least squares technique and its applications to harmonic superresolution. *IEEE Transactions on Signal Processing*, Piscataway, NJ, v. 39, n. 5, p. 1070–1087, May 1991. Disponível em: <https://doi.org/10.1109/78.80955>. Acesso em: 24 de junho de 2021.

ABATZOGLOU, T. T.; SOON, V. Constrained total least squares applied to frequency estimation of sinusoids. *In: ANNUAL ASSP WORKSHOP ON SPECTRUM ESTIMATION AND MODELING*, 4., 1988, Minneapolis. *Proceedings[...]*. Piscataway, NJ: IEEE, 1988. p. 250–252. Disponível em: <https://doi.org/10.1109/SPECT.1988.206201>. Acesso em: 24 de junho de 2021.

ARENGAS, D.; KROLL, A. A search method for selecting informative data in predominantly stationary historical records for multivariable system identification. *In: CONFERENCE ON SYSTEM THEORY, CONTROL AND COMPUTING*, 21., 2017, Sinaia. *Proceedings[...]*. Piscataway, NJ: IEEE, 2017. p. 100–105. Disponível em: <https://doi.org/10.1109/ICSTCC.2017.8107018>. Acesso em: 24 de junho de 2021.

ARENGAS, D.; KROLL, A. Searching for informative intervals in predominantly stationary data records to support system identification. *In: CONFERENCE ON INFORMATION, COMMUNICATION AND AUTOMATION TECHNOLOGIES*, 12., 2017, Sarajevo, Bosnia and Herzegovina. *Proceedings[...]*. Piscataway, NJ: IEEE, 2017. p. 1–6. Disponível em: <https://doi.org/10.1109/ICAT.2017.8171617>. Acesso em: 24 de junho de 2021.

ARENGAS, D.; KROLL, A. Removal of insufficiently informative data to support system identification in miso processes. *In: EUROPEAN CONTROL CONFERENCE*, 2018, Limassol, Cyprus. *Proceedings[...]*. Piscataway, NJ: IEEE, 2018. p. 2842–2847.

Disponível em: <https://doi.org/10.23919/ECC.2018.8550533>. Acesso em: 24 de junho de 2021.

ARENGAS, D.; KROLL, A. A data selection method for large databases based on recursive instrumental variables for system identification of MISO models. *In: EUROPEAN CONTROL CONFERENCE*, 18., 2019, Naples, Italy. *Proceedings[...]*. Piscataway, NJ: IEEE, 2019. p. 357–362. Disponível em: <https://doi.org/10.23919/ECC.2019.8796086>. Acesso em: 24 de junho de 2021.

ÅSTRÖM, K. J.; HÄGGLUND, T. Benchmark systems for PID control. *IFAC Proceedings Volumes*, Amsterdam, v. 33, n. 4, p. 165–166, Apr 2000. Trabalho apresentado no IFAC Workshop on Digital Control: Past, Present and Future of PID Control, 2000, [Terrassa, Espanha]. Disponível em: [https://doi.org/10.1016/S1474-6670\(17\)38238-1](https://doi.org/10.1016/S1474-6670(17)38238-1). Acesso em: 24 de junho de 2021.

BAZANELLA, A. S.; CAMPESTRINI, L.; ECKHARD, D. *Data-driven controller design: the H2 approach*. Dordrecht: Springer Science & Business Media, 2011. (Communications and Control Engineering). Disponível em: <https://doi.org/10.1007/978-94-007-2300-9>. Acesso em: 24 de junho de 2021.

BAZANELLA, A. S.; NEUHAUS, T. Tuning nonlinear controllers with the virtual reference approach. *IFAC Proceedings Volumes*, Amsterdam, v. 47, n. 3, p. 10269–10274, Aug 2014. Trabalho apresentado no 19th IFAC World Congress, 2014, [Cidade do Cabo, África do Sul]. Disponível em: <https://doi.org/10.3182/20140824-6-ZA-1003.00562>. Acesso em: 24 de junho de 2021.

BITTENCOURT, A. C. et al. An algorithm for finding process identification intervals from normal operating data. *Processes*, Basel, v. 3, p. 357–383, May 2015. Disponível em: <https://doi.org/10.3390/pr3020357>. Acesso em: 24 de junho de 2021.

BOEIRA, E. C.; ECKHARD, D. Multivariable virtual reference feedback tuning with bayesian regularization. *In: CONGRESSO BRASILEIRO DE AUTOMÁTICA*, 12., 2018, João Pessoa, PB. *Anais[...]*. Campinas: Sociedade Brasileira de Automática, 2018. v. 1. Disponível em: https://www.sba.org.br/open_journal_systems/index.php/sba/article/view/859/804. Acesso em: 24 de junho de 2021.

BORDIGNON, V.; CAMPESTRINI, L. Application of multivariable virtual reference feedback tuning with anti-windup to the benchmark PID 2018. *IFAC-PapersOnLine*, Amsterdam, v. 51, n. 4, p. 515–520, May 2018. Trabalho apresentado no 3rd IFAC Conference on Advances in Proportional-Integral-Derivative Control PID, 2018, [Ghent, Bélgica]. Disponível em: <https://doi.org/10.1016/j.ifacol.2018.06.147>. Acesso em: 24 de junho de 2021.

CAMPESTRINI, L. et al. Model reference control design by prediction error identification. *IFAC Proceedings Volumes*, Amsterdam, v. 45, n. 16, p. 1478–1783, Jul 2012. Trabalho apresentado no 16th IFAC Symposium on System Identification, 2012, [Brussels, Bélgica]. Disponível em: <https://doi.org/10.3182/20120711-3-BE-2027.00190>. Acesso em: 24 de junho de 2021.

CAMPESTRINI, L. et al. Data-driven model reference control design by prediction error identification. *Journal of the Franklin Institute*, Amsterdam, v. 354, n. 6, p. 2628–2647,

Apr. 2017. Disponível em: <https://doi.org/10.1016/j.jfranklin.2016.08.006>. Acesso em: 24 de junho de 2021.

CAMPESTRINI, L. et al. Unbiased MIMO VRFT with application to process control. *Journal of Process Control*, Amsterdam, v. 39, p. 35–49, Mar. 2016. Disponível em: <https://doi.org/10.1016/j.jprocont.2015.12.010>. Acesso em: 24 de junho de 2021.

CAMPESTRINI, L. et al. Virtual reference feedback tuning for non-minimum phase plants. *Automatica*, Amsterdam, v. 47, n. 8, p. 1778–1784, Aug. 2011. Disponível em: <https://doi.org/10.1016/j.automatica.2011.04.002>. Acesso em: 24 de junho de 2021.

CAMPI, M. C.; LECCHINI, A.; SAVARESI, S. M. Virtual reference feedback tuning (VRFT): a new direct approach to the design of feedback controllers. In: CONFERENCE ON DECISION AND CONTROL, 39., 2000, Sydney, Australia. *Proceedings[...]*. Piscataway, NJ: IEEE, 2000. p. 623–629. Disponível em: <https://doi.org/10.1109/CDC.2000.912835>. Acesso em: 24 de junho de 2021.

CAMPI, M. C.; LECCHINI, A.; SAVARESI, S. M. Virtual reference feedback tuning: a direct method for the design of feedback controllers. *Automatica*, Amsterdam, v. 38, n. 8, p. 1337–1346, Aug. 2002. Disponível em: [https://doi.org/10.1016/S0005-1098\(02\)00032-8](https://doi.org/10.1016/S0005-1098(02)00032-8). Acesso em: 24 de junho de 2021.

CAMPI, M. C.; SAVARESI, S. M. Direct nonlinear control design: the virtual reference feedback tuning (VRFT) approach. *IEEE Transactions on Automatic Control*, Piscataway, NJ, v. 51, n. 1, p. 14–27, Jan 2006. Disponível em: <https://doi.org/10.1109/TAC.2005.861689>. Acesso em: 24 de junho de 2021.

CARRETTE, P. et al. Discarding data may help in system identification. *IEEE Transactions on Signal Processing*, Piscataway, NJ, v. 44, n. 9, p. 2300–2310, Sep 1996. Disponível em: <https://doi.org/10.1109/78.536685>. Acesso em: 24 de junho de 2021.

DEGROAT, R. D.; DOWLING, E. M. The data least squares problem and channel equalization. *IEEE Transactions on Signal Processing*, Piscataway, NJ, v. 41, n. 1, p. 407–411, Jan 1993. Disponível em: <https://doi.org/10.1109/TSP.1993.193165>. Acesso em: 24 de junho de 2021.

ECKHARD, D.; CAMPESTRINI, L.; BOEIRA, E. C. Virtual disturbance feedback tuning. *IFAC Journal of Systems and Control*, Amsterdam, v. 3, p. 23–29, Mar. 2018. Disponível em: <https://doi.org/10.1016/j.ifacsc.2018.01.003>. Acesso em: 24 de junho de 2021.

FORMENTIN, S. et al. Deterministic continuous-time virtual reference feedback tuning (VRFT) with application to PID design. *Systems & Control Letters*, Amsterdam, v. 127, p. 25–34, May 2019. Disponível em: <https://doi.org/10.1016/j.sysconle.2019.03.007>. Acesso em: 24 de junho de 2021.

GEVERS, M. et al. Identification and the information matrix: How to get just sufficiently rich? *IEEE Transactions on Automatic Control*, Piscataway, NJ, v. 54, n. 12, p. 2828–2840, Dec. 2009. Disponível em: <https://doi.org/10.1109/TAC.2009.2034199>. Acesso em: 24 de junho de 2021.

GOLUB, G. H.; LOAN, C. F. V. An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis*, Philadelphia, v. 17, n. 6, p. 883–893, Dec 1980. Disponível em: <https://doi.org/10.1137/0717073>. Acesso em: 24 de junho de 2021.

GONÇALVES DA SILVA, G. R.; BAZANELLA, A. S.; CAMPESTRINI, L. On the choice of an appropriate reference model for control of multivariable plants. *IEEE Transactions on Control Systems Technology*, Piscataway, NJ, v. 27, n. 5, p. 1937–1949, Sept. 2019. Disponível em: <https://doi.org/10.1109/TCST.2018.2842178>. Acesso em: 24 de junho de 2021.

GONÇALVES DA SILVA, G. R.; BAZANELLA, A. S.; CAMPESTRINI, L. One-shot data-driven controller certification. *ISA Transactions*, Amsterdam, v. 99, p. 361–373, Apr. 2020. Disponível em: <https://doi.org/10.1016/j.isatra.2019.10.011>. Acesso em: 24 de junho de 2021.

GONÇALVES DA SILVA, G. R.; CAMPESTRINI, L.; BAZANELLA, A. S. Automating the choice of the reference model for data-based control methods applied to PID controllers. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, 20., 2014, Belo Horizonte, MG. *Anais[...]*. Campinas: Sociedade Brasileira de Automática, 2014. p. 1088–1095. Disponível em: <http://www.swge.inf.br/CBA2014/anais/PDF/1569933823.pdf>. Acesso em: 24 de junho de 2021.

GONÇALVES DA SILVA, G. R.; CAMPESTRINI, L.; BAZANELLA, A. S. Multivariable virtual reference feedback tuning for non-minimum phase plants. *IEEE Control Systems Letters*, Piscataway, NJ, v. 2, n. 1, p. 121–126, Jan. 2018. Disponível em: <https://doi.org/10.1109/LCSYS.2017.2763824>. Acesso em: 24 de junho de 2021.

GUARDABASSI, G. O.; SAVARESI, S. M. Virtual reference direct design method: An off-line approach to data-based control system design. *IEEE Transactions on Automatic Control*, Piscataway, NJ, v. 45, n. 5, p. 954–959, May 2000. Disponível em: <https://doi.org/10.1109/9.855559>. Acesso em: 24 de junho de 2021.

GUARDABASSI, G. O.; SAVARESI, S. M. The virtual input approach to direct data-based control system design: some simulation studies. In: CONFERENCE ON DECISION AND CONTROL, 40., 2001, Orlando, Florida. *Proceedings[...]*. Piscataway, NJ: IEEE, 2001. p. 1156–1161. Disponível em: <https://doi.org/10.1109/CDC.2001.981041>. Acesso em: 24 de junho de 2021.

HJALMARSSON, H. Iterative feedback tuning: an overview. *International Journal of Adaptive Control and Signal Processing*, Hoboken, NJ, v. 16, n. 5, p. 373–395, June 2002. Disponível em: <https://doi.org/10.1002/acs.714>. Acesso em: 24 de junho de 2021.

HJALMARSSON, H. et al. Iterative feedback tuning: Theory and applications. *IEEE Control Systems Magazine*, Piscataway, NJ, v. 18, n. 4, p. 26–41, Aug. 1998. Disponível em: <https://doi.org/10.1109/37.710876>. Acesso em: 24 de junho de 2021.

HOLCOMB, C. M.; BITMEAD, R. R. *Subspace Identification with Multiple Data Records: unlocking the archive*. 2017. Disponível em: <https://arxiv.org/abs/1704.02635>. Acesso em: 24 de junho de 2021.

HORN, R. A.; JOHNSON, C. R. *Matrix analysis*. 2. ed. New York: Cambridge University Press, 2013. Disponível em: <https://doi.org/10.1017/9781139020411>. Acesso em: 24 de junho de 2021.

HUFF, D. D. et al. Data-driven control design by prediction error identification for multivariable systems. *Journal of Control, Automation and Electrical Systems*, Berlin, v. 30, p. 465–478, Aug. 2019. Disponível em: <https://doi.org/10.1007/s40313-019-00468-9>. Acesso em: 24 de junho de 2021.

HUFF, D. D.; GONÇALVES DA SILVA, G. R.; CAMPESTRINI, L. Data-driven control design by prediction error identification for a refrigeration system based on vapor compression. *IFAC-PapersOnLine*, Amsterdam, v. 51, n. 4, p. 704–709, May 2018. Trabalho apresentado no 3rd IFAC Conference on Advances in Proportional-Integral-Derivative Control PID, 2018, [Ghent, Bélgica]. Disponível em: <https://doi.org/10.1016/j.ifacol.2018.06.186>. Acesso em: 24 de junho de 2021.

HUFFEL, S. V.; LEMMERLING, P. *Total least squares and errors-in-variables modeling: analysis, algorithms and applications*. Dordrecht: Springer Science & Business Media, 2002. Disponível em: <https://doi.org/10.1007/978-94-017-3552-0>. Acesso em: 24 de junho de 2021.

HUFFEL, S. V.; VANDEWALLE, J. Analysis and properties of the generalized total least squares problem $ax \approx b$ when some or all columns in a are subject to error. *SIAM Journal on Matrix Analysis and Applications*, Philadelphia, v. 10, n. 3, p. 294–315, Jul 1989. Disponível em: <https://doi.org/10.1137/0610023>. Acesso em: 24 de junho de 2021.

KARIMI, A.; MIŠKOVIĆ, L.; BONVIN, D. Iterative correlation-based controller tuning. *International Journal of Adaptive Control and Signal Processing*, Hoboken, NJ, v. 18, n. 8, p. 645–664, Oct 2004. Disponível em: <https://doi.org/10.1002/acs.825>. Acesso em: 24 de junho de 2021.

KARIMI, A.; MIŠKOVIĆ, L.; BONVIN, D. Convergence analysis of an iterative correlation-based controller tuning method. *IFAC Proceedings Volumes*, Amsterdam, v. 35, n. 1, p. 413–418, Apr. 2016. Trabalho apresentado no 15th IFAC World Congress, 2002, [Barcelona, Espanha]. Disponível em: <https://doi.org/10.3182/20020721-6-ES-1901.00150>. Acesso em: 24 de junho de 2021.

KARIMI, A.; VAN HEUSDEN, K.; BONVIN, D. Non-iterative data-driven controller tuning using the correlation approach. In: EUROPEAN CONTROL CONFERENCE (ECC), 2007., 2007, Kos, Grécia. *Proceedings[...]*. Piscataway, NJ: IEEE, 2007. p. 5189–5195. Disponível em: <https://doi.org/10.23919/ECC.2007.7068802>. Acesso em: 24 de junho de 2021.

KIM, J. et al. Least-squares methods for identifying biochemical regulatory networks from noisy measurements. *BMC Bioinformatics*, Berlin, v. 8, p. 15, Jan 2007. Disponível em: <https://doi.org/10.1186/1471-2105-8-8>. Acesso em: 24 de junho de 2021.

LENON REMES, C. et al. Virtual reference feedback tuning applied to DC-DC converters. *IEEE Transactions on Industrial Electronics*, Piscataway, NJ, v. 68, n. 1, p. 544–552, Jan 2021. Disponível em: <https://doi.org/10.1109/TIE.2020.2967729>. Acesso em: 24 de junho de 2021.

LENON REMES, C. et al. Data-driven approach for current control in DC-DC boost converters. *IFAC-PapersOnLine*, Amsterdam, v. 52, n. 1, p. 190–195, Apr 2019. Trabalho apresentado no 12th IFAC Symposium on Dynamics and Control of Process Systems,

2019, [Florianópolis, Brasil]. Disponível em: <https://doi.org/10.1016/j.ifacol.2019.06.059>. Acesso em: 24 de junho de 2021.

LIM, J. Recursive DLS solution for extreme learning machine-based channel equalizer. *Neurocomputing*, Amsterdam, v. 71, n. 4–6, p. 592–599, Jan 2008. Disponível em: <https://doi.org/10.1016/j.neucom.2007.07.022>. Acesso em: 24 de junho de 2021.

LIM, J.-S. A neural network solution on data least square algorithm and its application for channel equalization. In: LIU, D. et al. (Ed.). *Advances in Neural Networks – ISNN 2007*. Berlin, Heidelberg: Springer, 2007, (Lecture Notes in Computer Science, 4492). p. 678–685. Disponível em: https://doi.org/10.1007/978-3-540-72395-0_84. Acesso em: 24 de junho de 2021.

LJUNG, L. *System identification: theory for the user*. 2. ed. Upper Saddle River, NJ: Prentice Hall, 1999.

MACGREGOR, J.; CINAR, A. Monitoring, fault diagnosis, fault-tolerant control and optimization: Data driven methods. *Computers & Chemical Engineering*, Amsterdam, v. 47, p. 111–120, Dec 2012. Disponível em: <https://doi.org/10.1016/j.compchemeng.2012.06.017>. Acesso em: 24 de junho de 2021.

MARKOVSKY, I.; HUFFEL, S. V. Overview of total least-squares methods. *Signal Processing*, Amsterdam, v. 87, n. 10, p. 2283–2302, Oct 2007. Disponível em: <https://doi.org/10.1016/j.sigpro.2007.04.004>. Acesso em: 24 de junho de 2021.

NAKAMOTO, M. An application of the virtual reference feedback tuning for an MIMO process. In: SICE ANNUAL CONFERENCE, 2004., 2004, Sapporo, Japão. *Proceedings[...]*. Piscataway, NJ: IEEE, 2004. v. 3, p. 2208–2213. Disponível em: <https://ieeexplore.ieee.org/document/1491812>. Acesso em: 24 de junho de 2021.

NASCIMENTO, M. M.; BERNARDO, R. T.; DOTTA, D. Data-driven secondary voltage control design using PMU measurements. In: POWER & ENERGY SOCIETY INNOVATIVE SMART GRID TECHNOLOGIES CONFERENCE, 2020, Washington, DC. *Proceedings[...]*. Piscataway, NJ: IEEE, 2020. p. 1–5. Disponível em: <https://doi.org/10.1109/ISGT45199.2020.9087785>. Acesso em: 24 de junho de 2021.

OPPENHEIM, A. V.; BUCK, J. R.; SCHAFER, R. W. *Discrete-time signal processing*. 2. ed. Upper Saddle River, NJ: Prentice Hall, 1999.

PERETZKI, D. et al. Data mining of historic data for process identification. In: AICHE ANNUAL MEETING, 2011, Minneapolis. *Proceedings[...]*. New York: American Institute of Chemical Engineers, 2011. p. 16–21. Disponível em: <https://www.diva-portal.org/smash/get/diva2:650878/FULLTEXT01.pdf>. Acesso em: 24 de junho de 2021.

PINTO DA SILVA, R. W.; ECKHARD, D. Extension of the correlation-based tuning method for load disturbance rejection. In: IEEE CONFERENCE ON DECISION AND CONTROL (CDC), 58., 2019, Nice, França. *Proceedings[...]*. Piscataway, NJ: IEEE, 2019. p. 8386–8391. Disponível em: <https://doi.org/10.1109/CDC40024.2019.9029360>. Acesso em: 24 de junho de 2021.

PINTO DA SILVA, R. W.; ECKHARD, D. Data-driven controller certification through Vinnicombe test local rational models. *In: IEEE CONFERENCE ON CONTROL TECHNOLOGY AND APPLICATIONS (CCTA), 2020., 2020, Montreal, Canadá. Proceedings[...].* Piscataway, NJ: IEEE, 2020. p. 231–236. Disponível em: <https://doi.org/10.1109/CCTA41146.2020.9206310>. Acesso em: 24 de junho de 2021.

PINTO DA SILVA, R. W.; ECKHARD, D. Multivariable correlation-based tuning for load disturbance rejection. *In: IFAC WORLD CONGRESS, 21., 2020, Berlim. Proceedings[...].* Laxemburgo, Áustria: IFAC, 2020. p. 4043–4048. Meio digital.

PREVIDI, F. et al. Virtual reference feedback tuning (VRFT) of velocity controller in self-balancing industrial manual manipulators. *In: AMERICAN CONTROL CONFERENCE, 2010, Baltimore, MD. Proceedings[...].* Piscataway, NJ: IEEE, 2010. p. 1956–1961. Disponível em: <https://doi.org/10.1109/ACC.2010.5531358>. Acesso em: 24 de junho de 2021.

QUARTERONI, A.; SACCO, R.; SALERI, F. *Numerical Mathematics*. 2. ed. Berlin: Springer, 2007. (Texts in Applied Mathematics). Disponível em: <https://doi.org/10.1007/b98885>. Acesso em: 24 de junho de 2021.

RADRIZZANI, S. et al. A data-driven approach for fast controller calibration of brake-by-wire actuators. *In: IEEE CONFERENCE ON CONTROL TECHNOLOGY AND APPLICATIONS (CCTA), 2020., 2020, Montreal, Canadá. Proceedings[...].* Piscataway, NJ: IEEE, 2020. p. 561–566. Disponível em: <https://doi.org/10.1109/CCTA41146.2020.9206393>. Acesso em: 24 de junho de 2021.

RALLO, G. et al. Virtual reference feedback tuning with bayesian regularization. *In: EUROPEAN CONTROL CONFERENCE, 2016, Aalborg, Dinamarca. Proceedings[...].* Piscataway, NJ: IEEE, 2016. p. 507–512. Disponível em: <https://doi.org/10.1109/ECC.2016.7810335>. Acesso em: 24 de junho de 2021.

ROJAS, J. D. et al. Application of multivariate virtual reference feedback tuning for wastewater treatment plant control. *Control Engineering Practice*, Amsterdam, v. 20, n. 5, p. 499–510, May 2012. Disponível em: <https://doi.org/10.1016/j.conengprac.2012.01.004>. Acesso em: 24 de junho de 2021.

SHARDT, Y. A.; HUANG, B. Parameter-based conditions for closed-loop system identifiability of ARX models with routine operating data. *Journal of the Franklin Institute*, Amsterdam, v. 354, n. 2, p. 722–751, Jan 2017. Disponível em: <https://doi.org/10.1016/j.jfranklin.2016.10.027>. Acesso em: 24 de junho de 2021.

SHARDT, Y. A. W.; HUANG, B. Data quality assessment of routine operating data for process identification. *Computers & Chemical Engineering*, Amsterdam, v. 55, p. 19–27, Aug 2013. Disponível em: <https://doi.org/10.1016/j.compchemeng.2013.03.029>. Acesso em: 24 de junho de 2021.

SHARDT, Y. A. W.; SHAH, S. L. Segmentation methods for model identification from historical process data. *IFAC Proceedings Volumes*, Amsterdam, v. 47, n. 3, p. 2836–2841, Aug 2014. Trabalho apresentado no 19th IFAC World Congress, 2014, [Cidade do Cabo, África do Sul]. Disponível em: <https://doi.org/10.3182/20140824-6-ZA-1003.00594>. Acesso em: 24 de junho de 2021.

SILVA GARCIA, C.; BAZANELLA, A. S. Recursive and traditional VRFT method implemented in a mobile application. *In: SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE*, 8., 2017, Porto Alegre, RS. *Anais[...]*. Campinas: Sociedade Brasileira de Automática, 2017. p. 1281–1286. Disponível em: https://www.ufrgs.br/sbai17/papers/paper_361.pdf. Acesso em: 24 de junho de 2021.

SILVA GARCIA, C.; BAZANELLA, A. S. Extraction of informative subsets from routine operating data for use in data-driven control. *In: IEEE CONFERENCE ON DECISION AND CONTROL (CDC)*, 58., 2019, Nice, França. *Proceedings[...]*. Piscataway, NJ: IEEE, 2019. p. 8392–8397. Disponível em: <https://doi.org/10.1109/CDC40024.2019.9029265>. Acesso em: 24 de junho de 2021.

SILVA GARCIA, C.; BAZANELLA, A. S. The constrained total least squares solution for virtual reference feedback tuning. *In: IFAC WORLD CONGRESS*, 21., 2020, Berlim. *Proceedings[...]*. Laxemburgo, Áustria: IFAC, 2020. p. 4049–4054. Meio digital.

SILVA GARCIA, C.; BAZANELLA, A. S. Selection of informative intervals in routine operating data for use in data-driven control design. *In: IEEE CONFERENCE ON CONTROL TECHNOLOGY AND APPLICATIONS (CCTA)*, 2020., 2020, Montreal, Canadá. *Proceedings[...]*. Piscataway, NJ: IEEE, 2020. p. 225–230. Disponível em: <https://doi.org/10.1109/CCTA41146.2020.9206266>. Acesso em: 24 de junho de 2021.

SÖDERSTRÖM, T. Errors-in-variables methods in system identification. *IFAC Proceedings Volumes*, Amsterdam, v. 39, n. 1, p. 1–19, Mar 2006. Trabalho apresentado no 14th IFAC Symposium on Identification and System Parameter Estimation, 2006, [Newcastle, Austrália]. Disponível em: <https://doi.org/10.3182/20060329-3-AU-2901.00002>. Acesso em: 24 de junho de 2021.

SÖDERSTRÖM, T. Errors-in-variables methods in system identification. *Automatica*, Amsterdam, v. 43, n. 6, p. 939–958, Jun 2007. Disponível em: <https://doi.org/10.1016/j.automatica.2006.11.025>. Acesso em: 24 de junho de 2021.

SÖDERSTRÖM, T. *Errors-in-Variables Methods in System Identification*. Cham: Springer International Publishing, 2018. (Communications and Control Engineering). Acesso em: 24 de junho de 2021.

TANGIRALA, A. K. *Principles of System Identification: Theory and practice*. Boca Raton, FL: CRC Press, 2015.

TSAI, C.-L.; KAO, W.-W. Constrained total least-square solution for GPS compass attitude determination. *Applied Mathematics and Computation*, Amsterdam, v. 183, n. 1, p. 106–118, Dec 2006. Disponível em: <https://doi.org/10.1016/j.amc.2006.05.064>. Acesso em: 24 de junho de 2021.

VAN HEUSDEN, K.; KARIMI, A.; SÖDERSTRÖM, T. On identification methods for direct data-driven controller tuning. *International Journal of Adaptive Control and Signal Processing*, Hoboken, NJ, v. 25, n. 5, p. 448–465, May 2011. Disponível em: <https://doi.org/10.1002/acs.1212>. Acesso em: 24 de junho de 2021.

VARRIALE DA SILVA, E. *Contribuições para aplicabilidade de métodos não iterativos de projeto de controladores baseados em dados*. 2020. 102 f. Dissertação (Mestrado em

Engenharia Elétrica) — Universidade Federal do Rio Grande do Sul, Porto Alegre, 2020. Disponível em: <http://hdl.handle.net/10183/215334>. Acesso em: 24 de junho de 2021.

VARRIALE DA SILVA, E.; CAMPESTRINI, L. The effects of disturbances in non-iterative data-driven controller design. *In: SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE*, 14., 2019, Ouro Preto, MG. *Anais[...]*. Campinas: Sociedade Brasileira de Automática, 2019. v. 1, p. 2534–2539. Disponível em: <https://proceedings.science/sbai-2019/papers/the-effects-of-disturbances-in-non-iterative-data-driven-controller-design>. Acesso em: 24 de junho de 2021.

WANG, J. et al. Searching historical data segments for process identification in feedback control loops. *Computers & Chemical Engineering*, Amsterdam, v. 122, p. 6–16, Apr 2018. Disponível em: <https://doi.org/10.1016/j.compchemeng.2018.01.018>. Acesso em: 24 de junho de 2021.

YEREDOR, A. The extended least squares criterion: Minimization algorithms and applications. *IEEE Transactions on Signal Processing*, Piscataway, NJ, v. 49, n. 1, p. 74–86, Jan 2001. Disponível em: <https://doi.org/10.1109/78.890344>. Acesso em: 24 de junho de 2021.