

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
Escola de Engenharia  
Programa de Pós-Graduação em Engenharia Elétrica



Luís Henrique dos Santos

---

**IDENTIFICAÇÃO DE MODELOS DE HAMMERSTEIN  
MULTIVARIÁVEIS COM NÃO LINEARIDADES ESTÁTICAS OU  
QUASE ESTÁTICAS FORTES**

---

Belo Horizonte  
2024

Luís Henrique dos Santos

---

**IDENTIFICAÇÃO DE MODELOS DE HAMMERSTEIN  
MULTIVARIÁVEIS COM NÃO LINEARIDADES ESTÁTICAS OU  
QUASE ESTÁTICAS FORTES**

---

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Prof. Dr. Bruno Otávio Soares Teixeira  
Coorientador: Prof. Dr. Rodrigo Augusto Ricco

Belo Horizonte  
2024

S237i

Santos, Luís Henrique dos.

Identificação de modelos de Hammerstein multivariáveis com não linearidades estáticas ou quase estáticas fortes [recurso eletrônico] / Luís Henrique dos Santos. - 2024.

1 recurso online (122 f. : il., color.) : pdf.

Orientador: Bruno Otávio Soares Teixeira.

Coorientador: Rodrigo Augusto Ricco.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Bibliografia: f. 116-122.

Exigências do sistema: Adobe Acrobat Reader.

1. Engenharia elétrica - Teses. 2. Modelos matemáticos - Teses. 3. Modelos não lineares (Estatística) - Teses. 4. Histerese - Teses. 5. Modelagem - Teses. I. Teixeira, Bruno Otávio Soares. II. Ricco, Rodrigo Augusto. III. Universidade Federal de Minas Gerais. Escola de Engenharia. IV. Título.

CDU: 621.3(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS  
ESCOLA DE ENGENHARIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

### FOLHA DE APROVAÇÃO

"IDENTIFICAÇÃO DE MODELOS DE HAMMERSTEIN MULTIVARIÁVEIS COM NÃO LINEARIDADES ESTÁTICAS OU QUASE ESTÁTICAS FORTES"

**LUÍS HENRIQUE DOS SANTOS**

Dissertação de Mestrado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do grau de Mestre em Engenharia Elétrica. Aprovada em 23 de fevereiro de 2024. Por:

Prof. Dr. Bruno Otávio Soares Teixeira  
DELT - UFMG - Orientador

Prof. Dr. Rodrigo Augusto Ricco  
DEELT - UFOP - Coorientador

Prof. Dr. Víctor Costa da Silva Campos  
DELT - UFMG

Prof. Dr. Samir Angelo Milani Martins  
DEE - UFSJ

Dr. Petrus Emmanuel Oliveira Gomes Brant Abreu  
Bolsista de Desenvolvimento Tecnológico em TICs do CNPq



Documento assinado eletronicamente por **Bruno Otavio Soares Teixeira, Professor do Magistério Superior**, em 23/02/2024, às 12:48, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Victor Costa da Silva Campos, Professor do Magistério Superior**, em 23/02/2024, às 12:52, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Samir Angelo Milani Martins, Usuário Externo**, em 26/02/2024, às 11:35, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Rodrigo Augusto Ricco, Usuário Externo**, em 26/02/2024, às 13:32, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Petrus Emmanuel Oliveira Gomes Brant Abreu, Usuário Externo**, em 28/02/2024, às 17:26, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufmg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **3048222** e o código CRC **0285396F**.

---

# Agradecimentos

---

Em primeiro lugar, expresso minha gratidão à UFMG por ter me proporcionado toda uma estrutura de excelência, para a realização dos meus estudos. É através de instituições públicas de ensino superior como ela que podemos evoluir como sociedade. Estendo meus agradecimentos ao PPGEE-UFMG e aos seus funcionários, professores e alunos que contribuíram para o desenvolvimento deste trabalho. À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES, pelo suporte financeiro.

Agradeço ao meu orientador, Prof. Bruno Teixeira, pela confiança e atenção, pelos ensinamentos e por possibilitar essa oportunidade ímpar de crescimento pessoal e profissional. Em especial, agradeço ao meu coorientador Prof. Rodrigo Ricco por ser a ponte para o Mestrado. Muito obrigado por todas oportunidades, conselhos, pelo conhecimento e por me orientar com tanto zelo desde a graduação.

Por fim, menciono aqueles sem os quais nada faria sentido nem seria possível. Aos meus pais, Jesus e Rosária, aos meus irmãos e cunhados, Paulo, Jussara, José Cláudio, Glaucus e Iaponyra, agradeço imensamente pela formação, pela confiança, pelos conselhos e pelo amor incondicional. À minha amada namorada Évelly, meu porto seguro e inspiração para todo o empenho.

---

# Resumo

---

Os modelos descrevem características essenciais de sistemas, possibilitando análises e previsão de comportamentos que seriam dispendiosos, arriscados ou até mesmo impraticáveis. Tendo em vista a complexidade dos processos, que por vezes são multivariáveis, uma representação adequada é o modelo linear em espaço de estados que pode lidar com várias entradas e várias saídas de forma compacta. No entanto, todo sistema prático possui algum grau de não linearidade, de modo que quanto mais não linear o efeito, menor a utilidade do modelo linear em termos de representatividade global. À vista disso, nota-se que modelos não lineares são necessários quando se deseja operar o processo em faixas amplas, sobretudo quando as não linearidades são fortes. No escopo desta dissertação, não linearidades fortes são aquelas que possuem em seu domínio ao menos um ponto onde a derivada não está definida. Apesar de apresentar avanços consideráveis nas últimas décadas, a literatura relacionada ao uso de métodos de subespaços com a finalidade de obter modelos em espaço de estados para sistemas não lineares multivariáveis é escassa, sobretudo quando se tratam de não linearidades fortes. Motivado por tal contexto, essa dissertação objetiva apresentar duas metodologias distintas de identificação de sistemas com não linearidades fortes: uma para não linearidades estáticas e outra para não linearidades quase estáticas, mais especificamente a histerese. Uma estrutura interessante para representar processos não lineares é o modelo de Hammerstein, composto por um subsistema estático não linear seguido de outro dinâmico linear. Sob a perspectiva de processos cujas não linearidades sejam estáticas, define-se no trabalho uma rede neural que tem capacidade de ajuste acurada e um sistema em espaço de estados adequado ao caso multivariável para compor o modelo de Hammerstein. Levando em conta tais características, foi desenvolvida uma abordagem para identificar um modelo de Hammerstein multivariável *neuro-fuzzy* por meio de um procedimento não iterativo, associado aos métodos de subespaços. Uma das vantagens é que não é necessário escolher a estrutura como ocorre nos modelos autorregressivos. Por sua vez, sob a perspectiva dos processos que possuem a não linearidade na forma de histerese, foi apresentada uma metodologia em duas etapas para fazer a identificação de sistemas multivariáveis. Diferentemente de outros trabalhos encontrados na literatura, primeiro estima-se a não linearidade para depois estimar a parcela dinâmica linear. Assim, evitam-se restrições adicionais, a exemplo da impossibilidade de inversão direta que ocorre quando o sistema identificado é de fase não mínima. O modelo de Hammerstein proposto é composto por um modelo generalizado

dependente da taxa de Prandtl-Ishlinskii seguido de outro em espaço de estados. A funcionalidade de ambas metodologias de identificação foram verificadas em quatro simulações, todas multivariáveis, apresentando um exemplo introdutório e outro mais complexo para cada uma das metodologias. Os modelos obtidos acompanharam o processo, mesmo com a presença de ruído de medição e acoplamento dinâmico.

Palavras-chave: modelos de Hammerstein multivariáveis; não linearidades fortes; histerese; sistemas *neuro-fuzzy*; modelo de Prandtl-Ishlinskii.

---

# Abstract

---

Models describe essential characteristics of systems, enabling analysis and prediction of behaviors that would be expensive, risky or even impractical. Given the complexity of the processes, which are sometimes multivariable, a suitable representation is the linear state space model that can handle multiple inputs and multiple outputs in a compact way. However, every practical system has some degree of nonlinearity, so the more nonlinear the effect, the less useful the linear model is in terms of global representativeness. In view of this, it is noted that nonlinear models are necessary when one wishes to operate the process in wide ranges, especially when non-linearities are strong. In the scope of this dissertation, strong nonlinearities are those that have at least one point in their domain where the derivative is not defined. Despite presenting considerable advances in recent decades, the literature related to the use of subspace methods for the purpose of obtaining state space models for multivariable nonlinear systems is scarce, especially when dealing with strong nonlinearities. Motivated by this context, this dissertation aims to present two distinct methodologies for identifying systems with strong nonlinearities: one for static nonlinearities and the other for quasi-static nonlinearities, more specifically hysteresis. An interesting structure to represent nonlinear processes is the Hammerstein model, composed of a nonlinear static subsystem followed by a linear dynamic one. From the perspective of processes whose nonlinearities are static, the work defines a neural network that has accurate adjustment capacity and a state space system suitable for the multivariable case to compose the Hammerstein model. Taking these characteristics into account, an approach was developed to identify a neuro-fuzzy multivariable Hammerstein model through a non-iterative procedure, associated with subspace methods. One of the advantages is that it is not necessary to choose the structure as in autoregressive models. In turn, from the perspective of processes that have nonlinearity in the form of hysteresis, a two-step methodology was presented to identify multivariable systems. Unlike other works found in the literature, the nonlinearity is first estimated and then the linear dynamic portion is estimated. Thus, additional restrictions are avoided, such as the impossibility of direct inversion that occurs when the identified system has a non-minimum phase. The proposed Hammerstein model is composed of a generalized rate-dependent Prandtl-Ishlinskii model followed by another in state space. The functionality of both identification methodologies was verified in four simulations, all multivariable, presenting an introductory example and a more complex example for

each of the methodologies. The models obtained followed the process, even with the presence of measurement noise and dynamic coupling.

Keywords: multivariable Hammerstein models; strong nonlinearities; hysteresis; neuro-fuzzy systems; Prandtl-Ishlinskii model.

---

---

# Lista de Figuras

---

1.1	Curva estática não linear representada por retas, sendo cada reta a representação de um modelo diferente. . . . .	21
1.2	Exemplos de não linearidades fortes. . . . .	22
1.3	Resultado da busca “ <i>block-oriented models</i> ”, “ <i>Hammerstein models</i> ” e “ <i>Wiener models</i> ” no <i>Web of Science</i> em 18 de janeiro de 2024. . . . .	26
1.4	Estruturas clássicas dos modelos de blocos interconectados em malha aberta. . . . .	27
1.5	Esquema da maneira geral como a identificação de sistemas funciona. . . . .	29
2.1	Estruturas de Hammerstein e Wiener monovariáveis. . . . .	34
2.2	Possíveis variações do modelo de Hammerstein multivariável, inspiradas nos modelos MISO apresentados em (Ribeiro e Aguirre, 2014). . . . .	35
2.3	Esquema do processo que pode ser utilizado para estimar os parâmetros de um modelo NF. . . . .	38
2.4	Operador relé elementar com dois disparos $p_0$ e $p_1$ e dois níveis -1 e +1. . . . .	40
2.5	Diagrama do funcionamento do modelo de Prandtl-Ishlinskii para reproduzir o fenômeno da histerese. . . . .	40
2.6	Exemplos ilustrativos dos operadores <i>stop</i> , <i>play</i> e <i>play</i> generalizado. . . . .	41
2.7	Dois tipos de funções que alteram as bordas dos laços de histerese para compor os operadores <i>play</i> de modo que eles se tornem operadores <i>play</i> generalizados. . . . .	43
2.8	Representação de um laço de histerese por meio do modelo GPI. . . . .	44
2.9	Exemplo do comportamento de processos com histerese independente e dependente da taxa quando a frequência da entrada é variada. . . . .	45
2.10	Esquema do algoritmo MOESP. . . . .	48
3.1	Estrutura da rede NF de 4 camadas, empregada para estimar a curva estática não linear de cada entrada do modelo de Hammerstein. . . . .	53

3.2	Modelo de Hammerstein multivariável em malha aberta com $p$ entradas e $m$ saídas. . . . .	53
3.3	Processo com dois subsistemas SUB 1 e SUB 2 e um sinal interno $v(k)$ desconhecidos. . . . .	56
3.4	Esquema da equivalência entre o sinal de entrada e o sinal intermediário, para identificação, quando as condições estabelecidas nas Suposições 10 a 12 são satisfeitas. . . . .	57
3.5	Esquema do algoritmo MOESP que, a partir de dados de entrada e saída coletados de um processo, estima a ordem do modelo e as matrizes em espaço de estados, a menos de uma transformação de similaridade $\mathcal{T}$ . . . . .	58
3.6	Diagrama dos testes feitos com sinal multinível para identificar as curvas estáticas não lineares. . . . .	60
3.7	Diagramação do exemplo a respeito de como são dadas as nomenclaturas aos dados necessários para implementação da <i>clusterização</i> no algoritmo NF. . . . .	61
3.8	Diagrama que ilustra a avaliação de cada dado $u_i(k)$ a fim de verificar a qual <i>cluster</i> ele irá pertencer, ou ainda, se irá formar um novo <i>cluster</i> . . . . .	62
3.9	Sinais relacionados ao subíndice 1 do ensaio dinâmico realizado no processo do Exemplo 1 para implementar a primeira etapa da metodologia de identificação do modelo de Hammerstein MIMO NF. . . . .	67
3.10	Sinais empregados para a identificação das curvas estáticas não lineares do Exemplo 1 do Capítulo 3. . . . .	69
3.11	Validação da característica estática do Exemplo 1 do Capítulo 3. . . . .	70
3.12	Validação completa do modelo de Hammerstein MIMO NF identificado no Exemplo 1 do Capítulo 3. . . . .	71
3.13	Sinais da primeira realização utilizados para identificação. . . . .	72
3.14	Validação da nuvem de curvas estáticas no Exemplo 2 do Capítulo 3. . . . .	74
3.15	Relação entre o aumento de número de amostras na identificação da curva estática e o RMSE da validação dinâmica do Exemplo 2 do Capítulo 3. . . . .	74
3.16	Validação dinâmica por simulação livre, partindo de condições iniciais nulas em 1000 realizações no Exemplo 2 do Capítulo 3. . . . .	75
4.1	Teste do comportamento das funções de borda polinomial e tangente hiperbólica quando a entrada é variada de -5 a 5. . . . .	81
4.2	Representação do modelo de Hammerstein KU multivariável cuja parcela não linear é composta de um subsistema GRDPI. . . . .	83

4.3	Resposta ao degrau de um exemplo ilustrativo para justificar a escolha de parâmetros a serem empregados no sinal de degraus crescentes e decrescentes. . . . .	86
4.4	Esquema da influência da excitação de uma entrada por vez em todas as saídas do modelo de Hammerstien GRDPI. . . . .	86
4.5	Teste com degraus consecutivos realizado para esclarecer a justificativa da escolha da quantidade total de degraus após definir a amplitude total de excursão. . . . .	87
4.6	Algoritmo MPSO com destaque de modificações em relação ao PSO clássico. . . . .	89
4.7	Saídas de uma função estática em resposta a um sinal PRBS. . . . .	94
4.8	Saídas de uma função quase estática em resposta a uma entrada PRBS. . . . .	94
4.9	Esquema do algoritmo MOESP para identificar a parcela LIT do modelo de Hammerstein GRDPI. . . . .	95
4.10	Diagramas do passo a passo a ser executado na segunda etapa da identificação dos modelos de Hammerstein GRDPI. . . . .	97
4.11	Dados do ensaio estático coletados do processo composto por (4.33) e (4.34). . . . .	99
4.12	Parâmetros dos modelos GRDPI estimados para as duas entradas do Exemplo 1 do Capítulo 4. . . . .	100
4.13	Nuvem de respostas para validação dos 100 modelos de Hammerstein GRDPI comparados aos sinais do processo composto por (4.33) e (4.34). . . . .	102
4.14	Dados de entrada e saída do ensaio estático empregado para identificar o modelo GRDPI no Exemplo 2 do Capítulo 4. . . . .	103
4.15	Todos os parâmetros dos 1000 modelos GRDPI estimados para as duas entradas do Exemplo 2 do Capítulo 4. . . . .	104
4.16	Dados de validação da parcela dinâmica linear do Exemplo 2 do Capítulo 4. . . . .	105
4.17	Validação dos 1000 modelos de Hammerstein GRDPI obtidos para identificar o processo composto por (4.36) e (4.37). . . . .	106
4.18	Validação do melhor modelo, composto por (4.38) e (4.39), dentre os 1000 modelos obtidos na identificação do modelo de Hammerstein GRDPI do Exemplo 2 do Capítulo 4. . . . .	107
5.1	Diagramas ilustrativos da diferença entre empregar a <i>clusterização</i> com abrangência constante ou adaptativa. . . . .	114

---

# Lista de Abreviações

---

<b>ARX</b>	modelo autorregressivo com entradas exógenas ( <i>AutoRegressive model with eXogenous inputs</i> )
<b>CVA</b>	análise de variáveis canônicas ( <i>Canonical Variate Analysis</i> )
<b>EIA-PSO</b>	otimização por enxame de partículas com estratégia adaptativa, informada e eficaz ( <i>Effective Informed Adaptive - Particle Swarm Optimization</i> )
<b>EJL</b>	estrutura do modelo de Hammerstein proposta por Eskinat, Johnson e Luyben
<b>GPI</b>	modelo de Prandtl-Ishlinskii generalizado ( <i>Generalized Prandtl-Ishlinskii</i> )
<b>GRDPI</b>	modelo de Prandtl-Ishlinskii generalizado e dependente da taxa ( <i>Generalized Rate-Dependent Prandtl-Ishlinskii</i> )
<b>KU</b>	estrutura do modelo de Hammerstein proposta por Kortmann e Unbehauen
<b>LIT</b>	sistema Linear e Invariante no Tempo
<b>LMI</b>	desigualdade matricial linear ( <i>Linear Matrix Inequality</i> )
<b>MATLAB</b>	<i>MATrix LABoratory</i>
<b>MBI</b>	Modelo de Blocos Interconectados
<b>MIMO</b>	múltiplas entradas e múltiplas saídas ( <i>Multiple Input Multiple Output</i> )
<b>MISO</b>	múltiplas entradas e uma saída ( <i>Multiple Input Single Output</i> )
<b>MOESP</b>	modelo multivariável de erro na saída em espaço de estados ( <i>Multivariable Output-Error State sPace</i> )
<b>MOESP-PO</b>	modelo multivariável de erro na saída em espaço de estados com entradas passadas atuando como variáveis instrumentais ( <i>Multivariable Output-Error State sPace - Past Output</i> )

<b>MPSO</b>	otimização por enxame de partículas modificado ( <i>Modified Particle Swarm Optimization</i> )
<b>MQ</b>	Mínimos Quadrados
<b>N4SID</b>	Algoritmos numéricos para identificação de sistema em espaço de estados empregando técnicas de subespaço ( <i>Numerical algorithms for Subspace State-Space System IDentification</i> )
<b>NARMAX</b>	modelo média móvel autoregressivo não linear com entrada exógena ( <i>Nonlinear AutoRegressive Moving Average with eXogenous input</i> )
<b>NARX</b>	modelo autoregressivo não linear com entrada exógena ( <i>Nonlinear AutoRegressive with eXogenous input</i> )
<b>NF</b>	<i>Neuro-Fuzzy</i>
<b>PBSID</b>	método de identificação por subespaço baseado em preditor ( <i>Predictor-Based Subspace IDentification method</i> )
<b>PEM</b>	método de predição de erro ( <i>Prediction Error Method</i> )
<b>P-I</b>	Prandtl-Ishlinskii
<b>PRBS</b>	sinal binário pseudoaleatório ( <i>Pseudo Random Binary Signal</i> )
<b>PSO</b>	otimização por enxame de partículas ( <i>Particle Swarm Optimization</i> )
<b>RDPI</b>	modelo de Prandtl-Ishlinskii dependente da taxa ( <i>Rate-Dependent Prandtl-Ishlinskii</i> )
<b>RMSE</b>	raiz quadrada do erro médio quadrático ( <i>Root Mean Square Error</i> )
<b>SBAI</b>	Simpósio Brasileiro de Automação Inteligente
<b>SIMO</b>	uma entrada e múltiplas saídas ( <i>Single Input Multiple Output</i> )
<b>SISO</b>	uma entrada e uma saída ( <i>Single Input Single Output</i> )
<b>SNR</b>	relação sinal ruído ( <i>Signal Noise Ratio</i> )
<b>SVD</b>	decomposição em valores singulares ( <i>Singular Value Decomposition</i> )

---

# Lista de Símbolos

---

$\in$	Pertence;
$\mathbb{R}$	Conjunto dos número reais;
$p$	Quantidade de entradas de processos ou modelos de blocos interconectados;
$k$	Variável de tempo discreto;
$u(k)$	Entrada de sistemas não autônomos no instante $k$ , $u(k) \in \mathbb{R}^p$ ;
$\varrho$	Quantidade de sinais intermediários de processos ou modelos de blocos interconectados;
$v(k)$	Sinal intermediário dos modelos de blocos interconectados no instante $k$ , $v(k) \in \mathbb{R}^\varrho$ ;
$m$	Quantidade de saídas de processos ou modelos de blocos interconectados;
$y(k)$	Sinal de saída dos modelos no instante $k$ , $y(k) \in \mathbb{R}^m$ ;
$\mathcal{N}$	Não linearidades do processo ou do modelo de blocos interconectados;
$\mathcal{L}$	Parcela linear do processo ou do modelo de blocos interconectados;
$x(k)$	Vetor de estados no instante $k$ , $x(k) \in \mathbb{R}^n$ ;
$A$	Matriz da dinâmica do sistema $\in \mathbb{R}^{n \times n}$ ;
$B$	Matriz de entradas $\in \mathbb{R}^{n \times \varrho}$ ;
$C$	Matriz de saída $\in \mathbb{R}^{m \times n}$ ;
$D$	Matriz de transmissão direta $\in \mathbb{R}^{m \times \varrho}$ ;
$w(k)$	Vetor de ruído de medição $\in \mathbb{R}^n$ ;
$\mathcal{T}$	Transformação de similaridade;
$(\hat{\bullet})$	Estimativa de $(\bullet)$ ;
$i$	Índice que assume valores de 1 a $p$ ;
$f(\cdot)$	Função crisp;
$\vartheta$	Parâmetro consequente do modelo Takagi-Sugeno de ordem zero;
$c$	Centro de uma função de pertinência gaussiana;
$\sigma$	Dispersão de uma função de pertinência gaussiana;
$t$	Variável de tempo contínuo;

- $b$ • Parâmetros do modelo de Bouc-Wen;
- Derivada de • no tempo contínuo;
- $d$ • Parâmetros do modelo de Duhem;
- $p$ • Disparo do operador elementar de Preisach ;
- $N_P$  Quantidade de operadores no modelo de Prandtl-Ishlinskii;
- $S$  Operador elementar *stop*;
- $r$  Disparo de um operador elementar de Prandtl-Ishlinskii;
- $j$  Índice que assume valores de 1 a  $N_P$ ;
- $\theta$  Parâmetro dos operadores do modelo de Prandtl-Ishlinskii;
- $\mathcal{P}$  Operador elementar *play*;
- $\mathcal{B}$  Função de borda do operador *play*;
- $\mathcal{G}$  Operador *play* generalizado;
- $b$  Parâmetros da função de borda;
- $\check{\theta}$  Parâmetro dos disparos dinâmicos do modelo de Prandtl-Ishlinskii;
- $T_s$  Período de amostragem;
- $p_n$  Ordem de persistência de excitação;
- $P_N$  Número de colunas da matriz  $\mathcal{V}$ ;
- $h$  Quantidade de blocos linha de uma matriz em blocos de Hankel;
- $\Gamma$  Matriz de observabilidade estendida;
- $T$  Transposto;
- $N_L$  Quantidade de dados para identificar dinâmica linear;
- $\mathcal{V}$  Matriz em bloco de Hankel equacionada com  $v(k)$ ;
- $\Upsilon$  Matriz em bloco de Hankel equacionada com  $y(k)$ ;
- $R$  Matriz à esquerda da fatoração RQ;
- $Q$  Matriz à direita da fatoração RQ;
- $U_o$  Matriz à esquerda da decomposição em valores singulares;
- $S_o$  Matriz com valores singulares da decomposição em valores singulares;
- $V_o$  Matriz à direita da decomposição em valores singulares;
- $l$  Índice que assume valores de 1 a  $m$ ;
- $N_N$  Quantidade de dados para identificar a não linearidade;
- $G_N$  Ganho em regime de uma parcela estática não linear do modelo de Hammerstein;
- $G_{\mathcal{L}}$  Ganho em regime de uma parcela dinâmica linear do modelo de Hammerstein;
- $G_{N\mathcal{L}}$  Ganho em regime de um modelo de Hammerstein;
- $\beta$  Ganho de uma função estática não linear no ponto  $u_\beta$ ;

- $I$  Matriz identidade de ordem  $\bullet$ ;
- $N_N^{(i)}$  Quantidade de dados do ensaio estático realizado na entrada  $i$ ;
- $\bar{N}_{i,j}$  Quantidade de dados formando o  $j$ -ésimo *cluster* da entrada  $i$ ;
- $N_i$  Quantidade total de *clusters* para a entrada  $i$ ;
- $\mathcal{J}$  Indica a posição do *cluster* mais próximo do dado  $u_i(k)$ ;
- $S_{i,0}$  Abrangência do *cluster*  $i$ ;
- $\lambda$  Fator de deslocamento do centro do *cluster*;
- $\rho$  Controle da dispersão da gaussiana;
- $E$  Função custo da metodologia *neuro-fuzzy*;
- $\Phi$  Ponderação utilizada na defuzificação;
- $\mu$  Função gaussiana;
- $n_d$  Ordem a partir da qual a derivada parcial de  $E$  é nula;
- $n_p$  Número de estágios do registrador para gerar um sinal binário pseudoaleatório;
- $\mathbb{N}$  Conjunto dos números naturais;
- $\mathcal{U}$  Distribuição uniforme;
- $\mu_u$  Média da distribuição uniforme;
- $\sigma_u$  Desvio padrão da distribuição uniforme;
- $\mathcal{N}$  Distribuição normal;
- $\mu_n$  Média da distribuição normal;
- $\sigma_n$  Desvio padrão da distribuição normal;
- $\iota$  Índice que assume valores de 1 a  $\varrho$ ;
- $\mathcal{GR}$  Operador *play* generalizad dependente da taxa;
- $\Theta$  Parâmetros do operador  $\mathcal{GR}$ ;
- $N_V$  Quantidade total de parâmetros de  $\Theta$ ;
- $F$  Função de aptidão das partículas de uma nuvem;
- $V$  Matriz com a velocidade das partículas de uma nuvem;
- $\tilde{\theta}$  Parâmetros que controlam a função de borda do operador generalizado dependente da taxa de Prandtl-Ishlinkii;
- $\epsilon$  Parâmetro interno do modelo GRDPI;
- $\Delta u$  Variação entre amostras do sinal de entrada;
- $N_G$  Quantidade de dados usados para ajustar o ganho total do modelo de Hammerstein MIMO GRDPI;
- $\xi$  Quantidade de partículas selecionadas para compor o grupo de melhores partículas no PSO;

$K_{\mu}$	Quantidade de iteração que iterações consecutivas sem alteração da função custo $F(\Theta)$ a partir da qual podem ocorrer mutações;
$P$	Quantidade de partículas do PSO;
$\ell$	Índice que assume valores de 1 a $P$ ;
$\kappa$	Índice da atual iteração do PSO;
$K_{\max}$	Quantidade de iterações que o PSO é executado;
$\alpha$	Probabilidade de ocorrer mutação no PSO;
$\omega$	Taxa de aprendizado cognitivo e social do PSO;
$U$	Matriz com amostras da distribuição uniforme;
$\odot$	Multiplicação elemento a elemento;
$M^{(p)}$	Média ponderada das melhores posições pessoais;
$M^{(g)}$	Média ponderada das melhores posições globais;
$\varsigma$	Média ponderada das melhores funções de aptidão;
$q$	Índice que assume valores de 1 a $\xi$ independente de $g$ ;
$g$	Índice que assume valores de 1 a $\xi$ independente de $q$ ;
$s$	Índice que assume valores de 2 a $P$ ;
$\chi$	Fator de constrição do PSO;
$\nu$	Amostra de uma distribuição uniforme;
$\gamma$	Uma dimensão aleatória de velocidade $V$ ;
$\eta$	Índice aleatório que indica uma das $P$ partículas da nuvem;
$\zeta$	Fator aleatório que altera a velocidade de uma partícula;
$\bar{b}$	Nível superior de um sinal PRBS;
$\underline{b}$	Nível inferior de um sinal PRBS;
$\bar{\theta}$	Ganhos da parte superior do sinal intermediário;
$\underline{\theta}$	Ganhos da parte inferior do sinal intermediário;
$\Theta$	Matriz de parâmetros com ganhos do sinal intermediário a serem otimizados;
$\underline{v}$	sinal intermediário compensado com os ganhos $\underline{\theta}$ ;
$\underline{y}$	sinal intermediário compensado com os ganhos $\underline{\theta}$ ;
$\underline{\mathcal{L}}$	Estimativa inicial do bloco dinâmico linear.

---

---

# Sumário

---

<b>1</b>	<b>Introdução</b>	<b>19</b>
1.1	Motivação e Justificativa . . . . .	19
1.2	Formulação Geral do Problema . . . . .	26
1.3	Objetivos da Pesquisa . . . . .	29
1.4	Estrutura da Dissertação . . . . .	30
<b>2</b>	<b>Revisão sobre Identificação de Sistemas Não Lineares</b>	<b>32</b>
2.1	Introdução . . . . .	32
2.2	Modelos de Blocos Interconectados . . . . .	32
2.2.1	Modelo de Hammerstein multivariável . . . . .	34
2.3	Modelos Neuro-Fuzzy . . . . .	36
2.4	Modelos de Histerese Monovariáveis . . . . .	38
2.4.1	Modelo de histerese de Preisach . . . . .	39
2.4.2	Modelo de histerese de Prandtl-Ishlinskii . . . . .	40
2.4.3	Modelo GRDPI . . . . .	44
2.5	Modelo para o Bloco Dinâmico Linear e Invariante no Tempo . . . . .	46
2.5.1	Identificação do modelo em espaço de estados . . . . .	46
2.6	Conclusão . . . . .	49
<b>3</b>	<b>Identificação de Não Linearidades Fortes Empregando a Metodologia NF</b>	<b>51</b>
3.1	Introdução . . . . .	51
3.2	Rede Neuro-Fuzzy Multivariável . . . . .	52
3.3	Formulação do Problema para Não Linearidades Estáticas . . . . .	53
3.4	Associação da Metodologia Neuro-Fuzzy e de Subespaços . . . . .	55
3.4.1	Identificação da dinâmica linear . . . . .	57

3.4.2	Identificação da curva estática não linear . . . . .	59
3.5	Exemplo Numérico . . . . .	64
3.5.1	Exemplo 1 . . . . .	65
3.5.2	Exemplo 2 . . . . .	70
3.6	Conclusão . . . . .	74
<b>4</b>	<b>Identificação de Histerese MIMO Empregando MPSO e Métodos de Subespaço</b>	<b>77</b>
4.1	Introdução . . . . .	77
4.2	Modelo GRDPI Multivariável . . . . .	78
4.3	Formulação do Problema para Não Linearidades Quase Estáticas . . . . .	82
4.4	Associação entre o MPSO e o Algoritmo de Subespaço . . . . .	84
4.4.1	Identificação do modelo GRDPI . . . . .	85
4.4.2	Identificação do modelo LIT . . . . .	92
4.5	Exemplos Numéricos . . . . .	96
4.5.1	Exemplo 1 . . . . .	98
4.5.2	Exemplo 2 . . . . .	101
4.6	Conclusão . . . . .	107
<b>5</b>	<b>Conclusões e Propostas para Trabalhos Futuros</b>	<b>109</b>
5.1	Considerações Finais . . . . .	109
5.2	Publicações . . . . .	111
5.3	Propostas de Trabalhos Futuros . . . . .	112
	<b>Bibliografia</b>	<b>122</b>

# Introdução

---

## 1.1 Motivação e Justificativa

Descrever processos reais por meio de modelos é uma tarefa de engenharia fundamental na indústria e que pode afetar diretamente o planejamento e impactar negativamente a produção. Para exemplificar, considere uma planta de lingotamento de alumínio que opera continuamente e tem etapas sequenciais a serem executadas. Além disso, considere que é comum que o espaço para estoque tanto de matéria-prima quanto de produtos sejam limitados, para evitar ociosidade em condições normais de operação (Jha et al., 2020). Desse modo, interromper o funcionamento a fim de aplicar testes em processos como o de lingotamento é dispendioso ou até mesmo financeiramente inviável devido à operação contínua acrescentado ao planejamento de funcionários, fornecedores e clientes (Bowers et al., 1995). No entanto, mesmo com as aproximações inerentes a existência de um modelo que represente a planta introduz a possibilidade de aplicar os testes com liberdade sobre o modelo em lugar de efetuarlos sobre a planta. Além disso, as condições de operação e análise necessárias para ambientar o teste podem ser prejudiciais ao processo, de modo que se torna impraticável aplicá-lo diretamente no sistema físico. Diante dessa conjuntura, o modelo apresenta-se como uma opção essencial para realizar análises e previsões com relação à planta em situações que não poderiam ser aplicadas fisicamente por limitação de tempo, segurança, entre outros. Como a capacidade do modelo pode ser um limitante na qualidade da solução final do problema, existe uma constante demanda por esquemas avançados de modelagem e identificação (Nelles, 2020).

Com o avanço tecnológico e industrial, a identificação de sistemas tem recebido atenção nas últimas décadas, dada a necessidade inerente aos processos de compreensão, análise, predição e controle. Essa dinâmica de pesquisas e descobertas cria a necessidade de desenvolver cada vez mais estudos de concepção e aprimoramento para suprir as lacunas que surgem na área. A evolução da inteligência computacional, por exemplo, apresenta diversos algoritmos que propiciam novos campos de melhorias para a identificação de sistemas. Os algoritmos podem permitir que modelos sejam ajustados automaticamente com base em dados experimentais. Podem ser usados para identificar sistemas complexos e não lineares, como o clima ou o mercado financeiro. Podem, ainda, ajudar a identificar padrões em grandes conjuntos de dados heterogê-

neos que poderiam ser impossíveis com outros métodos matemáticos (Ruano, 2005; Quaranta et al., 2020).

A identificação de sistemas dedica-se a obter um modelo matemático para um sistema dinâmico por meio de dados experimentais. Em geral, esses dados são observados com ruído em um sistema físico completamente desconhecido ou com determinadas características desconhecidas. O objetivo do modelo é aproximar propriedades específicas de um sistema e não necessariamente criar um modelo genérico que o represente completamente em qualquer situação. Como escrito por George Box, “essencialmente, todos os modelos estão errados, mas alguns são úteis” (Box e Draper, 1987, p. 424, tradução própria<sup>1</sup>), ou seja, o propósito de construir o modelo é que ele sirva para determinada tarefa e não que ele seja um modelo de tudo. Desse modo, é pertinente atestar se um modelo linear descreve o sistema da maneira desejada antes de partir para a abordagem não linear.

Uma das principais etapas do procedimento de identificação de um sistema é a seleção da estrutura que o modelo apresentará, na qual é essencial considerar a linearidade do elemento sob investigação (Aguirre, 2015; Nelles, 2020). Na prática, nenhum processo ou componente é linear, no entanto, quando os sinais de interesse que o descrevem permanecem em torno de um ponto de operação, o sistema pode ser analisado e, eventualmente, modelado sob a suposição de linearidade. Ainda assim, por mais que o modelo linear seja uma aproximação local, ele pode ser eficaz em regiões mais amplas à medida que as não linearidades presentes e a precisão exigida são reduzidas. Desse modo, os modelos lineares constituem uma abordagem amplamente utilizada (Sastry, 2013). Tendo em vista a complexidade dos sistemas, que, por vezes, são multivariáveis, uma representação adequada é o modelo linear em espaço de estados, conveniente para estimação, filtragem, predição e controle. Eles podem ser obtidos empregando os métodos de identificação por subespaços que podem estimar modelos no espaço de estados para sistemas multivariáveis por meio de dados de entrada e saída (Katayama, 2005; Van Overschee e De Moor, 2012).

Apesar da simplicidade obtida como consequência de assumir que determinado processo é linear em torno de um ponto de operação, métodos de análise e modelagem lineares podem não ser adequados para sistemas que apresentem não linearidades fortes. Caso essa natureza seja negligenciada, a acurácia do modelo diminuirá à medida que a não linearidade se torna mais severa (Nijmeijer e Van der Schaft, 2016; Grimble e Majecki, 2020). Ou seja, no melhor cenário o modelo perde a capacidade de representar o processo de forma acurada à medida que o modelo se distancia do comportamento do processo. Entretanto, ainda há a possibilidade de representar cada faixa que se comporta de maneira aproximadamente linear por um modelo diferente.

Observe as ilustrações da Figura 1.1 que retratam duas curvas não lineares aproximadas por retas, em que cada reta é um modelo diferente. Dentre as duas curvas, aquela exibida em a) possui menor variação, por isso pode ser representada por três retas. Por outro lado, a curva apresentada em b) precisou de treze retas no

<sup>1</sup>No original: “Essentially, all models are wrong, but some are useful.”

mesmo intervalo de entrada para ter uma aproximação qualitativamente próxima a b). À vista disso, a quantidade de modelos a serem gerenciados no caso apresentado na imagem b) é maior e mais custosa de implementar. Nessas circunstâncias, nota-se que modelos lineares podem ser ineficientes quando se deseja operar o processo em faixas amplas. A informação transmitida é que, para determinados sistemas, a representação por modelos lineares pode não ser uma alternativa atraente. Portanto, descartada a possibilidade de utilizar um modelo linear, pode-se adotar representações não lineares para analisar, prever o comportamento ou controlar o sistema sob estudo. Como os problemas não lineares costumam ser específicos, é comum que os métodos tratem de maneira particular cada tipo de não linearidade, aumentando os obstáculos de modelagem.

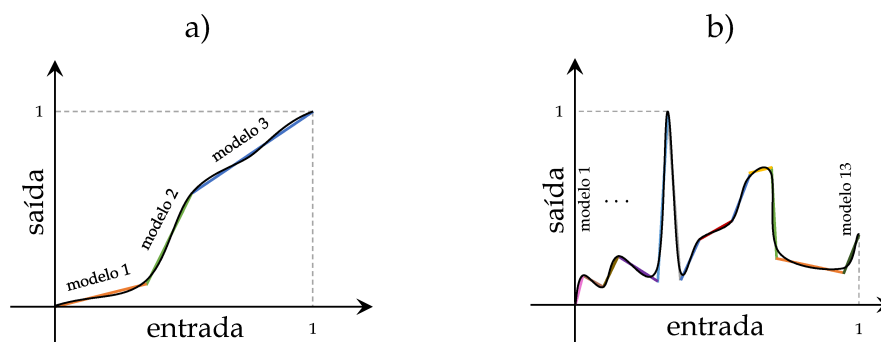


Figura 1.1: Curva estática não linear representada por retas, sendo cada reta a representação de um modelo diferente. Em a) é desenhada uma curva não linear que foi aproximada por três retas e em b) está uma curva com diversas sinuosidades de modo que foram necessárias treze retas para representar a curva b) de maneira qualitativamente próxima ao que foi feito com a curva em a).

A busca por uma representação matemática para sistemas não lineares levou ao desenvolvimento de técnicas para a modelagem baseadas em séries funcionais, como as séries de Volterra. Porém, elas são complexas do ponto de vista computacional e geram dificuldade na inserção de informação *a priori* (Borjas e Garcia, 2013). Estruturas que surgiram posteriormente incluem, por exemplo: modelos NARMAX (do inglês *Nonlinear AutoRegressive Moving Average with exogenous input*) (Billings, 2013), modelos *Neuro-Fuzzy* (NF) (Nelles, 2020), modelos de blocos interconectados (Mzyk, 2014), modelos não paramétricos (Ljung, 2010), entre outros.

No decorrer deste trabalho o foco será dado aos modelos de blocos interconectados (MBIs) que consistem na junção de subsistemas dinâmicos lineares invariantes no tempo (LIT) e elementos estáticos não lineares. As estruturas clássicas de blocos interconectados consistem na conexão em série entre os blocos linear e não linear. Elas recebem o nome de modelo de Hammerstein ou Wiener a depender da disposição desses blocos. Apesar da sua simplicidade, os modelos de Hammerstein e Wiener provaram poder descrever com precisão uma ampla variedade de sistemas não lineares, por exemplo, processos químicos (Zou et al., 2015; Roy et al., 2016; Li e Li, 2016), biológicos (Jalaleddini e Kearney, 2013; Fraçz, 2016), biomédicos (Jalaleddini e Kearney,

2013; Najafabadi e Shahrokhi, 2016) e também em identificação para controle (Biagiola et al., 2016; Zhang et al., 2017; Rayouf et al., 2018; Li et al., 2022).

No fim da década de 1980, surgiram os métodos de identificação por subespaços. Esses métodos foram formulados sob diversas perspectivas, tais como, *CVA* (do inglês *Canonical Variate Analysis*) (Larimore, 1990), *N4SID* (do inglês *Numerical algorithms for Subspace State Space System IDentification*) (Van Overschee e De Moor, 1994) e *MOESP* (do inglês *Multivariable Output Error State sPace*) (Verhaegen e Dewilde, 1992a,b). Essas técnicas puderam ser aplicadas para identificação da dinâmica linear dos modelos de blocos interconectados em espaço de estados. Trabalhos como os de (Gómez e Baeyens, 2005), (de Paula et al., 2015) e (Santos et al., 2021), empregam o *MOESP* para identificar as dinâmicas lineares associadas aos modelos de Hammerstein e Wiener em processos com não linearidades suaves.

Uma distinção que pode ser feita entre as curvas não lineares diz respeito à suavidade. Uma função suave possui derivada contínua em qualquer ponto do seu domínio, como é o caso daquelas apresentadas em preto contínuo na Figura 1.1a). Por outro lado, quando há ao menos um ponto em que a derivada não está definida, tem-se uma não linearidade definida como não linearidade forte.

**Definição 1.1.1** *Uma função não linear forte é aquela em que a derivada não está definida em pelo menos um ponto no seu domínio.*

Exemplos de não linearidades fortes que se manifestam frequentemente em sistemas práticos incluem histerese, saturação, zona morta e folga (Grimble e Majecki, 2020). Três delas são apresentadas na Figura 1.2.

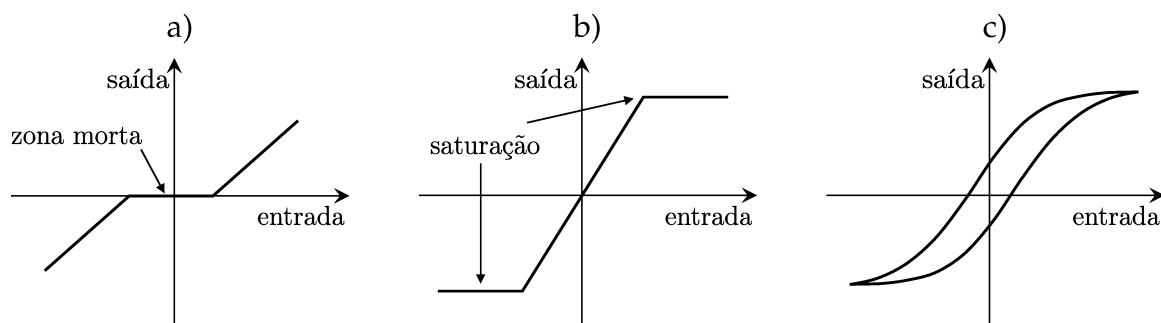


Figura 1.2: Exemplos de não linearidades fortes. Em a) apresenta-se a zona morta, em b) a saturação e em c) a histerese.

Sistemas físicos, sejam eles mecânicos, elétricos, fluídicos, entre outros, podem apresentar não linearidades fortes. Por exemplo, as máquinas elétricas e transformadores possuem não linearidades na forma de saturação. Tais efeitos surgem quando núcleos ferromagnéticos estão presentes, uma vez que eles são descritos por curvas e equações de magnetização não lineares (Fitzgerald et al., 2003). Por sua vez, os sistemas mecânicos podem incluir não linearidades como zona morta, histerese e folga. A presença dessas não linearidades em sistemas mecânicos deve-se ao desgaste ou

até mesmo aos limites físicos de percurso dos componentes. As não linearidades das válvulas são frequentes em plantas industriais, porém, um exemplo de não linearidade severa em sistemas fluídicos é a ação de relé comumente encontrada em processos de controle de nível (Garcia, 2013).

Devido à capacidade de modelar uma função estática não linear com precisão, alguns trabalhos utilizam redes neurais e sistemas *fuzzy* para representar o elemento não linear dos modelos de Hammerstein. Em geral, as investigações dedicam-se à identificação de um mapeamento não linear global a partir de dados, porém, o resultado apresentado por Jia et al. (2005) foi obtido ao propor identificar separadamente a curva estática com uma rede NF e a dinâmica linear utilizando o método clássico de mínimos quadrados. O sucesso da separação está relacionado ao emprego do sinal de teste proposto por Sung (2002) que desacopla a identificação das partes lineares e não lineares.

A metodologia apresentada em (Jia et al., 2005) foi desenvolvida para o caso monovariável. Posteriormente, o caso multivariável foi desenvolvido em (Jia et al., 2016), usando-se técnicas de inteligência computacional para identificar a curva não linear sob a forma de uma rede NF e um algoritmo de análise de correlação para estimar os parâmetros da dinâmica linear sob a forma de um modelo autorregressivo com entradas exógenas (ARX, do inglês *AutoRegressive with eXogenous input*). Entretanto, para sistemas multivariáveis, a representação em espaço de estados é a mais adequada.

O comportamento não linear estudado em (Jia et al., 2005) e (Jia et al., 2016) é estático e forte dada a presença de descontinuidades nas curvas identificadas. Apesar dos avanços nas últimas décadas, a literatura relacionada ao uso de métodos de subespaços para obtenção de modelos representados no espaço de estados para sistemas não lineares multivariáveis é escassa, sobretudo quando se tratam de não linearidades fortes. Dessa maneira, propõe-se, nesta dissertação, uma nova abordagem para identificar modelos de Hammerstein de várias entradas e várias saídas (MIMO, do inglês *Multiple Input Multiple Output*) com não linearidades fortes, empregando técnicas de inteligência computacional aliadas ao algoritmo MOESP (Verhaegen e Dewilde, 1992a). Busca-se identificar uma rede NF que represente a curva estática não linear e um modelo em espaço de estados que exprima a dinâmica linear sem a necessidade de qualquer conhecimento prévio. Isso é possível, pois o MOESP é capaz de retornar as matrizes e a ordem do sistema, sem exigir a escolha prévia da estrutura do modelo (Verhaegen e Dewilde, 1992a). Além disso, a rede NF evita restrições à função não linear estática, encontradas usando a abordagem polinomial (Jia et al., 2005), e não apresenta problemas de inicialização, como outros algoritmos iterativos (Chen e Chen, 2011).

Outro comportamento não linear forte de interesse é a histerese. Ela está presente em diversos sistemas mecânicos industriais, manifestada em sensores e atuadores como os piezoelétricos e válvulas de controle pneumático (Abreu et al., 2020). Essa não linearidade é um fenômeno quase estático, ou seja, suas propriedades são reveladas quando a variação da entrada é de baixa frequência. Alguns trabalhos recentes abordam

a identificação e compensação dessa não linearidade utilizando modelos polinomiais **NARX** (do inglês, *Nonlinear AutoRegressive models with exogenous input*) e em espaço de estados (Noël et al., 2017; Abreu et al., 2020; Tavares et al., 2022).

A histerese é um fenômeno não linear de interesse presente em diversos processos, sejam eles físicos, biológicos e até mesmo econômicos (Tan e Iyer, 2009; Hane, 2022). Uma classe de sistemas de interesse que empregam esse fenômeno em sua modelagem são os atuadores baseados em materiais inteligentes, como os piezo cerâmicos e magnetostritores. Esses atuadores são particularmente úteis para aplicações nas quais se deseja resposta rápida e com precisão micrométrica ou até mesmo nanométrica, como a varredura com microscópios de força atômica, dispositivos de microusinagem de precisão e manipulação de objetos em microambientes (Kwac e Kim, 2010; Aljanaideh et al., 2018; Tao et al., 2020). No entanto, esses materiais inteligentes apresentam algumas características inerentes que são empecilhos para operar com precisão, como a deformação, a influência da temperatura e a histerese, sendo essa última uma das mais degradantes (Gan e Zhang, 2019). Dessa forma, uma representação matemática adequada dessa propriedade é essencial para evitar erros de posicionamento.

Uma das formas de representar matematicamente os atuadores baseados em materiais inteligentes são os modelos de blocos interconectados, como, por exemplo, o modelo de Hammerstein (Wang et al., 2012; Aljanaideh et al., 2019). A não linearidade dos materiais inteligentes pode ser representada por modelos de histerese como os de Prandtl-Ishlinskii (**P-I**), Preisach, Bouc-Wen, Duhem, entre outros, que podem ser revisados em (Gan e Zhang, 2019) e nas referências nele contidas. Note que como a não linearidade é quase estática, a rigor, tem-se um modelo de blocos interconectados com duas parcelas dinâmicas. Então alguns artigos da literatura apresentam a nomenclatura pseudo-Hammerstein, no entanto, no decorrer deste trabalho, bem como na maioria de suas referências mantêm-se a nomenclatura modelo de Hammerstein

Dentre os modelos citados, o **P-I** tem sido utilizado recentemente para modelar não linearidades de materiais inteligentes com histerese (Gan et al., 2016; Qin et al., 2017; Aljanaideh et al., 2018; Wang et al., 2020; Aljanaideh et al., 2022). Ele é de interesse nesse trabalho devido à quantidade de parâmetros exigidos para sua construção, que é menor que outros modelos baseados em operadores, como o modelo de Preisach. Outra propriedade relevante é que seu modelo inverso pode ser obtido analiticamente, de modo a facilitar a tarefa de compensação em aplicações de controle, por exemplo. Além disso, há características que podem ser adicionadas ao modelo **P-I** para que ele se torne mais geral. Por exemplo, em (Aljanaideh et al., 2023), apresentam-se duas funções contínuas diferentes que alteram as bordas dos operadores para tornar o modelo de Prandtl-Ishlinskii generalizado (**GPI**, do inglês *Generalised Prandtl-Ishlinskii*), ou seja, para apresentar laços de histerese assimétricos. Por sua vez, em (Aljanaideh et al., 2007), apresentam-se várias situações em que a dependência da taxa pode ser modelada caracterizando o modelo de Prandtl-Ishlinskii dependente da taxa (**RDPI**, do inglês *Rate-Dependent Prandtl-Ishlinskii*). Por fim, os dois efeitos podem ser combinados em um modelo mais generalista que pode captá-los caso estejam presentes em determinados processos. Nessa situação, o sistema de Hammerstein tem sua parcela não linear

representada por um modelo de Prandtl-Ishlinkii generalizado e dependente da taxa (**GRDPI**, do inglês *Generalized Rate-Dependent Prandtl-Ishlinskii*) comumente aplicado para caracterizar atuadores baseados em materiais inteligentes; ver Subseção 2.4.3.

Uma série de trabalhos empregam a estrutura de Hammerstein para modelar processos com a presença de histerese, seja com o modelo P-I ou **GRDPI**, no entanto, a maioria se aplica a sistemas de uma entrada e uma saída (**SISO**, do inglês *Single Input Single Output*) (Aljanaideh et al., 2018, 2019, 2022; Al Janaideh et al., 2023). Poucos artigos abordam o caso multivariável, como em (Rakotondrabe, 2017) e mesmo assim não o empregam para modelos de Hammerstein, ou seja, não consideram a parte dinâmica linear. Uma extensão direta dos casos existentes na literatura não é interessante, uma vez que focam em modelos de difícil abordagem para o caso **MIMO**, como é o modelo **ARX**. Uma representação adequada para a parte linear do caso multivariável é o modelo em espaço de estados (Verhaegen e Dewilde, 1992a,b). Dessa forma, essa representação será utilizada para modelar a parcela dinâmica linear.

Outro problema das metodologias existentes, mesmo para o caso **SISO**, é que elas estimam o modelo dinâmico linear primeiro, de modo que uma aproximação da inversa deve ser obtida para estimar o sinal intermediário. Nos casos em que o sistema é de fase não mínima o processo para obtenção da inversa pode ser complexo, como pode ser verificado em (Aljanaideh et al., 2018, 2022) que empregam uma expansão da série de Laurent para aproximar a inversa e, a partir da aproximação, estimar o sinal intermediário. À vista disso, apresenta-se nessa dissertação uma metodologia para identificar sistemas com não linearidades quase estáticas multivariáveis empregando os **MBIs**.

Como citado anteriormente, esse texto se concentra em abordar a identificação de sistemas não lineares empregando os **MBIs**, em particular, o modelo de Hammerstein **MIMO**. Os subsistemas dinâmicos lineares do modelo de Hammerstein podem ser paramétricos ou não paramétricos, enquanto os elementos estáticos não lineares podem ser paramétricos ou não, com ou sem memória. Finalmente, os dois componentes do sistema podem ser interligados de diferentes maneiras, como série, paralelo ou realimentação (Bai, 2010). Essa flexibilidade fornece a esses modelos uma capacidade notável de capturar uma classe de sistemas complexos e não lineares variados e motiva o interesse por eles. Por exemplo, encontram-se variados estudos e aplicações dos modelos de Hammerstein para identificação não iterativa em uma e duas etapas (Gómez e Baeyens, 2005; Jaleddini e Kearney, 2013; de Paula et al., 2015; Li et al., 2017; Santos et al., 2021; Hou et al., 2022).

O estudo de ferramentas para identificação desses modelos iniciou-se no trabalho feito por Narendra e Gallman (1966) no qual os parâmetros dos blocos estático não linear e dinâmico linear foram estimados separadamente. A partir desse estudo diversos outros métodos foram desenvolvidos e a área de pesquisa continua ativa nas últimas décadas como pode ser observado na Figura 1.3. Os dados apresentados foram obtidos em uma busca realizada no *Web of Science* em janeiro de 2024. Os termos empregados na pesquisa foram “*block-oriented models*”, “*Hammerstein models*” e “*Wiener*

*models*". As restrições foram de que os termos deveriam aparecer no título de artigos publicados entre os anos de 1990 e 2023.

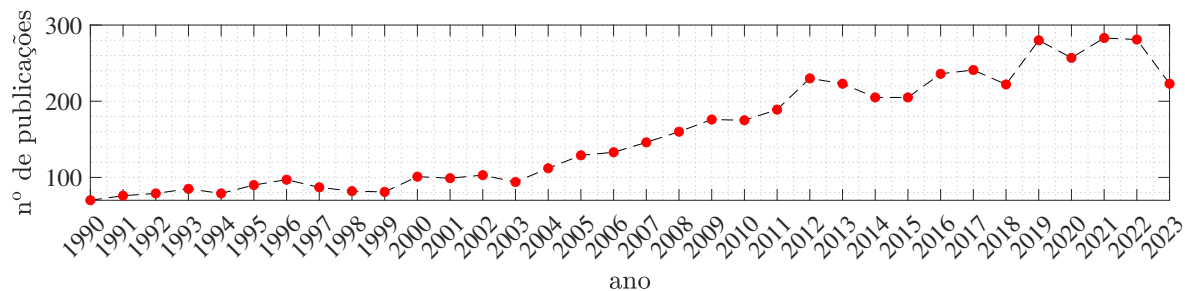


Figura 1.3: Resultado de busca no *Web of Science* em 18 de janeiro de 2024 para artigos com os termos “*block-oriented models*”, “*Hammerstein models*” e “*Wiener models*” presentes no título.

Neste trabalho a dinâmica linear será representada em espaço de estados devido à sua harmonia com casos multivariáveis e ao número de ferramentas para projeto de controladores disponíveis para estes tipos de modelos (Bai, 2010; Borjas e Garcia, 2012). Embora as técnicas modernas de projeto de controle tenham evoluído com base em uma abordagem no espaço de estados, os métodos clássicos de identificação de sistemas foram desenvolvidos empregando a estrutura de entrada e saída que caracteriza as funções de transferência. Na década de 1990 que o conceito de estado foi popularizado na identificação de sistemas, desenvolvendo assim muitos métodos de subespaço baseados na teoria clássica estocástica de realização (Verhaegen e Dewilde, 1992a,b; Van Overschee e De Moor, 1994; Katayama, 2005).

Embora tenham ocorrido avanços relevantes nas últimas décadas, pesquisas que utilizam técnicas de identificação por subespaço para identificar não linearidades combinadas com modelos em espaço de estados são escassas, particularmente quando se trata de não linearidades fortes (Aissaoui et al., 2016; Noël et al., 2017). Sendo assim, propõe-se como objeto de pesquisa desta dissertação a identificação de modelos de blocos interconectados por meio de técnicas de subespaços no cenário de não linearidades fortes. Tendo como horizonte a adequação do modelo em espaço de estados ao caso multivariável, os modelos de Hammerstein serão investigados, cujas curvas não lineares podem ser estáticas ou quase estáticas. Inicialmente, serão estudadas as formas de representação das curvas não lineares estáticas e quase estáticas com o intuito de modelar não linearidades fortes. O passo seguinte é aplicar essas técnicas junto aos métodos de subespaços, que no escopo deste trabalho será o MOESP.

## 1.2 Formulação Geral do Problema

Os dois modelos clássicos de blocos interconectados consistem na associação em cascata de uma parcela dinâmica linear e outra estática não linear. A ordem relativa

dos blocos define se a estrutura é de Hammerstein ou de Wiener, como pode ser observado de forma gráfica na Figura 1.4.

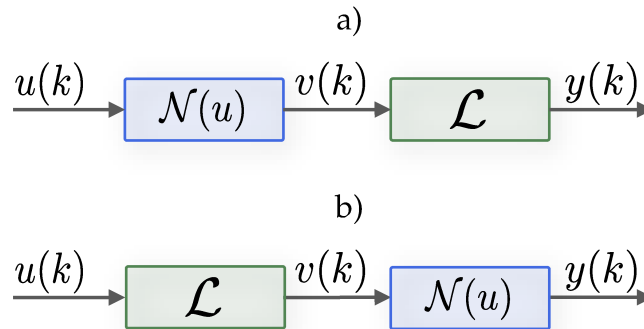


Figura 1.4: Estruturas clássicas dos modelos de blocos interconectados em malha aberta com  $p$  entradas  $u(k)$ ,  $\varrho$  sinais intermediários  $v(k)$  e  $m$  saídas  $y(k)$ . Em a) apresenta-se o modelo de Hammerstein composto por um subsistema não linear  $\mathcal{N}$  seguido de um LIT  $\mathcal{L}$ . Na parte b) do diagrama o modelo de Wiener composto pelos mesmos blocos em ordem invertida é apresentado.

Ambos blocos dos sistemas, de Hammerstein ou de Wiener, podem ser representados por modelos paramétricos ou não paramétricos. Exemplos amplamente conhecidos na literatura para representar sistemas LIT são os modelos autorregressivos, em espaço de estados e a resposta em frequência (Bai, 2010). Por sua vez, a não linearidade pode ser retratada por variadas representações, sobretudo devido à especificidade do problema não linear. Alguns exemplos incluem funções polinomiais e sistemas *fuzzy*.

Considere a estrutura apresentada na Figura 1.4 a) que configura o modelo de Hammerstein. Os sinais escritos no diagrama são:  $u(k) \in \mathbb{R}^p$ , que porta os dados da entrada do sistema,  $v(k) \in \mathbb{R}^{\varrho}$ , que representa os sinais intermediários e  $y(k) \in \mathbb{R}^m$ , que informam acerca dos sinais medidos na saída do processo.

Quanto aos subsistemas, o primeiro a aparecer, na forma de um retângulo azul, é o não linear que pode ser estático ou quase estático, monovariável ou multivariável, forte ou suave, representado pela notação  $\mathcal{N}(u(k))$ . Ou seja, é uma nomenclatura geral para uma função que represente a não linearidade. O fluxo de sinais intercorre conforme indicado pelas setas. Portanto, as amostras de entrada do processo  $u(k)$  são mapeadas pela não linearidade  $\mathcal{N}(u(k))$  produzindo o sinal intermediário  $v(k)$ . Esse sinal  $v(k)$ , em geral, não está disponível para medição. Note que  $\mathcal{N}$  é uma transformação não linear que leva do espaço de dimensão  $p$  da entrada para o espaço de dimensão  $\varrho$  do sinal intermediário.

Na formulação geral do problema apresentada nesta seção, a não linearidade permanecerá generalizada e sem representação definida, uma vez que ela será diferente para cada um dos cenários a serem detalhados no decorrer do texto, ou seja, não linearidade com e sem memória.

O segundo subsistema observado na Figura 1.4 a), representado por  $\mathcal{L}$  dentro do retângulo verde, é o modelo LIT. Ele, por sua vez, é responsável por estabelecer uma relação dinâmica entre os dados intermediários e os dados de saída do processo. Neste trabalho o estudo será feito com base nos modelos em espaço de estados devido à sua harmonia com casos multivariáveis. Dito isso, a descrição matemática do modelo de Hammerstein é

$$v(k) = \mathcal{N}(u(k)), \quad (1.1a)$$

$$x(k+1) = Ax(k) + Bv(k), \quad (1.1b)$$

$$y(k) = Cx(k) + Dv(k) + w(k), \quad (1.1c)$$

em que  $u(k)$ ,  $v(k)$  e  $y(k)$  são os sinais de entrada, intermediário e de saída do processo, respectivamente. O vetor  $x(k) \in \mathbb{R}^n$  contém os  $n$  estados do processo,  $w(k) \in \mathbb{R}^m$  é uma sequência de ruído branco de medição, estacionária e com média zero que contamina  $y(k)$ .  $A \in \mathbb{R}^{n \times n}$  é a matriz dinâmica do sistema, completamente caracterizada por seus autovalores.  $B \in \mathbb{R}^{n \times \rho}$  é a matriz de entrada que representa a transformação linear pela qual os sinais intermediários influenciam o próximo estado.  $C \in \mathbb{R}^{m \times n}$  é a matriz de saída que descreve como o estado interno é transferido para o mundo exterior nas medições  $y(k)$ . E, por fim, a matriz  $D \in \mathbb{R}^{m \times \rho}$  é denominada termo de transmissão direta por ser responsável por determinar a intensidade com que cada entrada afeta diretamente a saída.

É importante notar que os blocos do modelo podem não corresponder a componentes físicos do processo, uma vez que o sinal intermediário também é desconhecido e pode não existir. A inacessibilidade de tais medições, juntamente com as não linearidades do sistema, tornam a identificação dos modelos de blocos interconectados um desafio.

A dinâmica linear é descrita em espaço de estados pelas matrizes  $A$ ,  $B$ ,  $C$  e  $D$  e a função não linear por  $\mathcal{N}$  que pode ser descrita por funções de base radial, polinômios, e sistemas *fuzzy* no caso estático ou modelos de Bouc-Wen, Preisach ou P-I no caso da histerese (Noël et al., 2017). Além disso, as seguintes suposições são estabelecidas acerca do processo e seus sinais.

**Suposição 1** O processo está disponível para realizar testes.

**Suposição 2** O processo a ser identificado é estável em malha aberta.

**Suposição 3** Os dados são coletados do processo em malha aberta.

**Suposição 4** O sinal de entrada  $u(k)$  é limitado em amplitude.

**Suposição 5** O sinal de ruído  $w(k)$  é um processo aleatório i.i.d de amplitude limitada.

**Suposição 6** O sinal intermediário  $v(k)$  é inacessível, mas  $u(k)$  e  $y(k)$  são acessíveis.

**Suposição 7** A não linearidade  $\mathcal{N}$  é forte.

Desse modo, pode-se estabelecer o seguinte problema geral no contexto desta dissertação, cujo esquema é ilustrado na Figura 1.5.

**Problema 1** Obter estimativas da ordem  $n$  do sistema, das matrizes  $[A, B, C, D]_{\mathcal{T}}$ , em que  $\mathcal{T}$  é uma transformação de similaridade desconhecida e dos parâmetros das funções não lineares  $\mathcal{N}$  por meio dos dados de entrada e saída medidos do processo não linear em tempo discreto  $\{u(k), y(k)\}$  de modo a reduzir o erro entre os sinais produzidos pela saída do modelo e do processo quando a entrada dos dois é idêntica.

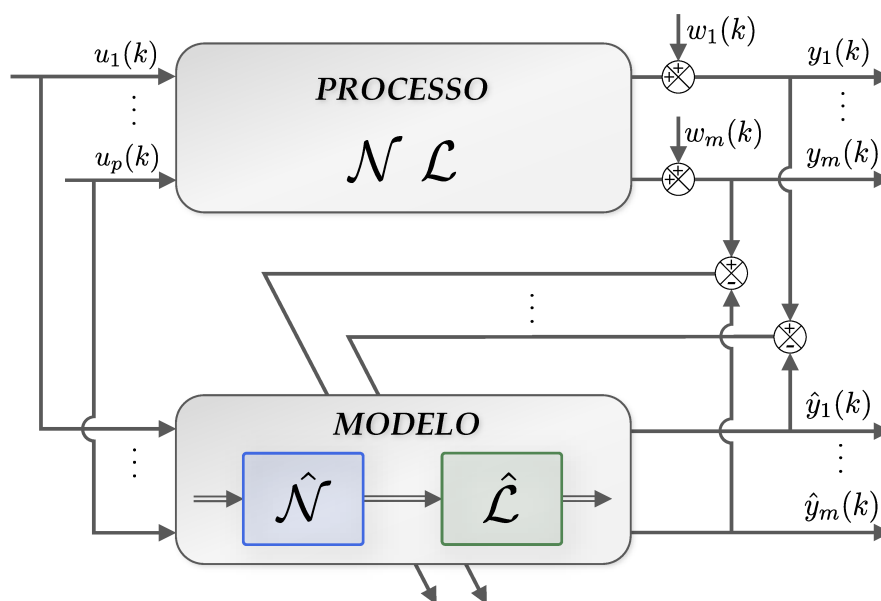


Figura 1.5: Esquema da maneira geral como a identificação de sistemas funciona. Os parâmetros do modelo são geralmente ajustados em função do erro entre a saída estimada pelo modelo e o saída do processo.

## 1.3 Objetivos da Pesquisa

O objetivo geral desta dissertação é investigar metodologias de identificação de modelos de Hammerstein por métodos de subespaços a fim de caracterizar processos com não linearidades fortes.

Têm-se como objetivos específicos neste trabalho:

- estudar as representações de não linearidades fortes na estrutura de blocos interconectados de Hammerstein contemplando não linearidades estáticas e quase estáticas;

- estudar os métodos de subespaços e os sinais de teste para identificar corretamente um modelo em espaço de estados linear na estrutura de Hammerstein **MIMO** sem ter acesso ao sinal intermediário;
- propor um método de identificação para estimar uma rede **NF MIMO** que represente as curvas estáticas não lineares e um modelo em espaço de estados multivariável que exprima a dinâmica linear sem a necessidade de qualquer conhecimento prévio a respeito do processo;
- identificar um modelo **GRDPI MIMO** que descreva laços de histerese assimétricos e dependentes da taxa em conjunto com um modelo em espaço de estados multivariável;
- implementar exemplos numéricos que possam validar a possibilidade de aplicação dos métodos de identificação propostos em processos com não linearidades fortes multivariáveis.

## 1.4 Estrutura da Dissertação

Esta dissertação está organizada em cinco capítulos, incluindo este, em que se motivou o estudo a partir de uma breve análise da literatura. Também neste primeiro capítulo foi apresentado o problema de pesquisa, bem como os objetivos a serem atingidos com o desenvolvimento do estudo.

Para inteirar o leitor quanto aos assuntos tratados na dissertação, apresenta-se, no Capítulo 2, a revisão de métodos e conceitos fundamentais que permeiam a identificação de modelos de blocos interconectados, especificamente, trata-se das representações de cada uma das parcelas linear e não linear que compõe os **MBIs**.

No Capítulo 3, apresenta-se uma proposta de identificação de não linearidades estáticas fortes, como uma curva sinuosa com descontinuidades, empregando uma metodologia **NF** junto com os métodos de subespaços para identificação. A metodologia **NF** proposta é capaz de identificar os parâmetros do modelo de Hammerstein cuja não linearidade seja representada por uma rede neural e a dinâmica linear em espaço de estados. Por fim apresentam-se alguns exemplos a fim de verificar a funcionalidade das metodologias.

O Capítulo 4 é destinado a uma segunda proposta para identificação de não linearidades quase estáticas, a histerese. Nesta situação emprega-se um algoritmo de otimização amplamente conhecido, o de otimização por enxame de partículas (**PSO**, do inglês *Particle Swarm Optimization*), com algumas modificações para identificar os parâmetros da parcela não linear do modelo de Hammerstein. A parcela dinâmica linear é identificada separadamente por meio de um algoritmo de subespaços. Nesse capítulo também apresentam-se alguns exemplos para verificar as metodologias.

Por fim, no Capítulo 5, retoma-se a atenção às principais contribuições do

---

trabalho e sugere-se perspectivas de temas para que sejam dadas continuidades a esta pesquisa.

# Revisão sobre Identificação de Sistemas Não Lineares

---

## 2.1 Introdução

A identificação de modelos não lineares é uma tarefa desafiadora devido às características únicas de cada processo não linear. Como eles raramente compartilham propriedades devido às particularidades de cada não linearidade, as estratégias de identificação precisam ser desenvolvidas com vistas a determinados aspectos de um processo específico e acabam não sendo eficazes para outros. Neste capítulo são abordados dois caminhos para identificação de processos com não linearidades fortes: caso estritamente estático, ou seja, sem memória, e caso de não linearidade quase estática. Ambos são desenvolvidos empregando o modelo de Hammerstein.

Apesar das particularidades dos processos, a modelagem feita neste trabalho é desenvolvida com base em uma só estrutura, flexível em termos de possibilidades de representação, que é a estrutura dos [MBIs](#). Por isso eles são introduzidos de maneira geral na Seção [2.2](#), sendo que, na Subseção [2.2.1](#), aprofunda-se a discussão do modelo de Hammerstein para o caso multivariável. Destaca-se nessa seção a existência de pelo menos duas diferentes estruturas para o modelo de Hammerstein multivariável. A diferença entre as duas estruturas é a independência entre as funções que modelam a não linearidade na entrada do processo. As duas seções seguintes [2.3](#) e [2.4](#) apresentam as opções para modelar a parcela não linear do modelo de Hammerstein que são: a rede [NF](#) para modelar a não linearidade estática forte e a histerese [GRDPI](#) para modelar a não linearidade quase estática forte. Por fim, na Seção [2.5](#), apresenta-se o método de subespaços que será utilizado para identificar a parte dinâmica linear dos processos.

## 2.2 Modelos de Blocos Interconectados

Os [MBIs](#) são constituídos de dois subsistemas elementares que são os dinâmicos [LITs](#) e os elementos estáticos ou quase estáticos não lineares. Esses modelos foram introduzidos na área de identificação de sistemas por meio do trabalho de ([Narendra e Gallman, 1966](#)). No entanto, por volta da década de 1980, [Billings e Fakhouri](#)

(1982) publicaram um artigo apresentando os MBI como uma alternativa atraente às expansões por meio da série de Volterra (Schetzen, 1981). A partir desse trabalho, os MBI receberam atenção devido a sua capacidade de aproximação e sobretudo devido à complexidade menor em relação à série de Volterra. Por exemplo, considerando a vantagem da separação entre os blocos do modelo de Hammerstein pode-se desenvolver uma função inversa da não linearidade e posicioná-la à frente do modelo para compensar somente esse comportamento. Caso a inversa seja exata, a estabilidade do modelo de Hammerstein pode ser determinada considerando somente o subsistema LIT (Nelles, 2020). Ademais, essa inversa pode ser posicionada à frente do processo para compensar a não linearidade estimada de forma aproximada (Santos et al., 2021).

Os modelos que compõem ambos os subsistemas dos MBI podem ser paramétricos ou não. Com respeito à parcela linear, caso seja definida como paramétrica, podem ser utilizadas em sua descrição as funções de transferência, representações em espaço de estados, modelos polinomiais ARX, entre outros. De outra forma, quando a parcela dinâmica é definida como não paramétrica, tem-se a possibilidade de se utilizar a resposta ao impulso, resposta em frequência, ou outras representações afins. Por sua vez, os elementos não lineares podem ser, além de paramétricos e não paramétricos, com ou sem memória. Representações frequentemente encontradas na literatura incluem polinômios, redes neurais, sistemas *fuzzy*, *look-up table*, modelos de histerese baseados em operadores e outros correlacionados (Nelles, 2020). Tal flexibilidade fornece aos MBI uma capacidade notável de representar diversas classes de sistemas não lineares complexos (Bai, 2010).

A divisão matemática dos blocos é um artifício matemático simplificador que não é necessariamente inspirado na característica física do processo, ou seja, não há separação completa entre a parcela linear e a não linear, na prática. Como consequência, não é adequado supor que os sinais intermediários, que conectam os subsistemas, sejam acessíveis para medição. Ainda que os obstáculos de modelagem sejam reduzidos pela estrutura de disjunção dos blocos interdependentes, a inacessibilidade do ponto intermediário, juntamente com as não linearidades do sistema, tornam a identificação do MBI um problema complexo (de Paula et al., 2015). Portanto, a maioria das soluções disponíveis dizem respeito a estruturas relativamente simples. Ainda assim, mesmo nos casos em que a separação não exista nos processos e seja virtualizada pelo modelo, a relação entrada-saída ainda pode ser representada de maneira conveniente por algum dos diversos MBI como os modelos de Hammerstein, de Wiener e a combinação deles, modelo de Hammerstein-Wiener e Wiener-Hammerstein (Mzyk, 2014).

O modelo de Hammerstein apresentado na Figura 2.1a), construído a partir de uma não linearidade estática e uma dinâmica linear conectadas em cascata, é a estrutura mais básica da classe dos MBI e, portanto, em sua maior parte considerada na literatura de identificação de sistemas (Mzyk, 2014). Inverter a ordem dos blocos resulta no modelo de Wiener que está apresentado na Figura 2.1b). Neste último modelo, o elemento não linear de saída pode representar adequadamente não linearidades do sensor, bem como quaisquer efeitos não lineares que possam ser incluídos na saída do sistema (Bai, 2010).

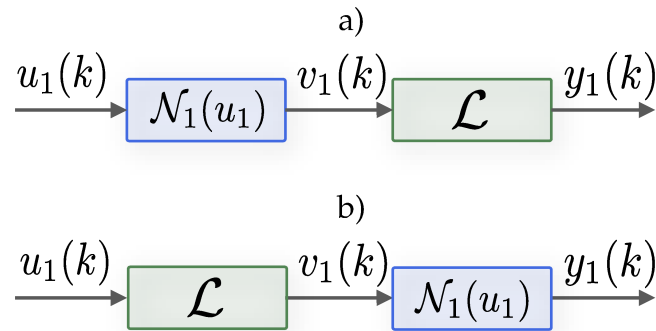


Figura 2.1: Estruturas dos modelos de blocos interconectados monovariáveis em malha aberta. Em a) apresenta-se o modelo de Hammerstein monovariável composto por um subsistema não linear  $\mathcal{N}$  seguido de um LIT  $\mathcal{L}$ . Na parte b) do diagrama o modelo de Wiener monovariável composto pelos mesmos blocos em ordem invertida é apresentado.

Este trabalho concentra-se no problema de identificação do modelo de Hammerstein que será apresentado com mais detalhes na Subseção 2.2.1 a seguir. Antes de iniciar a subseção é válido destacar que em todo o texto, assume-se que as características não lineares são mensuráveis do ponto de vista de entrada e saída (Mzyk, 2014). Além disso, assume-se que os blocos lineares são assintoticamente estáveis em malha aberta.

### 2.2.1 Modelo de Hammerstein multivariável

O modelo de Hammerstein é uma das estruturas de MBI que descreve adequadamente sistemas cujas não linearidades dominantes estejam presentes nos atuadores. Cita-se, por exemplo, a saturação de um motor eletromagnético e a presença de uma válvula de entrada em um sistema fluídico, mas ressalva-se que não se limita somente a essas situações (Nelles, 2020). Como apresentado na Figura 2.1a), o primeiro bloco do modelo de Hammerstein pode ser interpretado como uma função não linear capaz de mapear a entrada  $u_1(k)$  para um sinal intermediário  $v_1(k)$ . Logo, o sinal  $u_1(k)$  recebe um ganho, especificado pela não linearidade  $\mathcal{N}_1(u_1(k))$ , que depende do ponto de operação em que ocorre a excitação. Por sua vez, o sinal intermediário é utilizado na sequência como entrada do bloco dinâmico linear, o qual é responsável por agregar a dinâmica do sistema modelado ao sinal  $v_1(k)$  e produzir o sinal de saída  $y_1(k)$ . Assim,  $y_1(k)$  compreende toda a característica não linear estática e a dinâmica que o modelo detém (Bai, 2010).

Do mesmo modo que ocorre com os demais MBI, o sinal intermediário do modelo de Hammerstein é desconhecido, mas permite a separação entre os subsistemas estático não linear e dinâmico linear. A disposição dos subsistemas que compõe o modelo de Hammerstein monovariável é apresentada na Figura 2.1a). No entanto, os MBI não são restritos ao caso monovariável, ou seja, podem assumir o formato

multivariável. Devido à possibilidade de cruzamento das não linearidades e às escolhas em relação a quais entradas e saídas utilizar para identificação, o cenário multivariável é, em geral, mais desafiador que o monovariável.

Em (Ribeiro e Aguirre, 2014) são apresentadas diferentes estruturas de blocos interconectados para o modelo de Hammerstein com várias entradas e uma saída (MISO, do inglês *Multiple Input Single Output*). Duas dessas opções são apresentadas na Figura 2.2, entretanto, estendidas nesse texto para o caso MIMO. A representação de Kortmann e Unbehauen (KU) (Kortmann e Unbehauen, 1987), retratada na Figura 2.2a) é a menos geral das duas, uma vez que cada uma das  $p$  entradas, denotadas por  $u_i(k)$ , é inserida em apenas uma não linearidade  $\mathcal{N}_i(\cdot)$  resultando no sinal intermediário  $v_i(k)$ , para  $i = 1, 2, \dots, p$ .

**Observação 1** Note que o índice do sinal intermediário foi assumido como sendo  $i$ , o mesmo da entrada. Para isso é necessário assumir que as funções estáticas  $\mathcal{N}_i$  recebem como entrada somente  $u_i(k)$  de modo que as não linearidades são independentes como ocorre na Figura 2.2a) eliminando a possibilidade de  $\rho \neq p$ . Nesse caso utiliza-se  $p$  não somente para a quantidade de entradas, mas também para a quantidade de sinais intermediários daqui em diante.

O outro diagrama, que está apresentado na Figura 2.2b), é baseado na representação de Eskinat, Johnson e Luyben (EJL) (Eskinat et al., 1991).

**Observação 2** Na representação EJL não é necessário que o número de entradas seja igual ao número de sinais intermediários. No entanto, por não haver motivos para considerá-los diferentes no escopo deste texto, supõe-se que são iguais.

Essa segunda representação é a mais geral das duas, uma vez que pode descrever possíveis cruzamentos de não linearidades na entrada. Ao comparar os gráficos, é possível notar que a representação KU pode ser recuperada de EJL ao considerar que cada entrada  $u_i(k)$  tenha efeito apenas sobre a não linearidade  $\mathcal{N}_i(\cdot)$ .

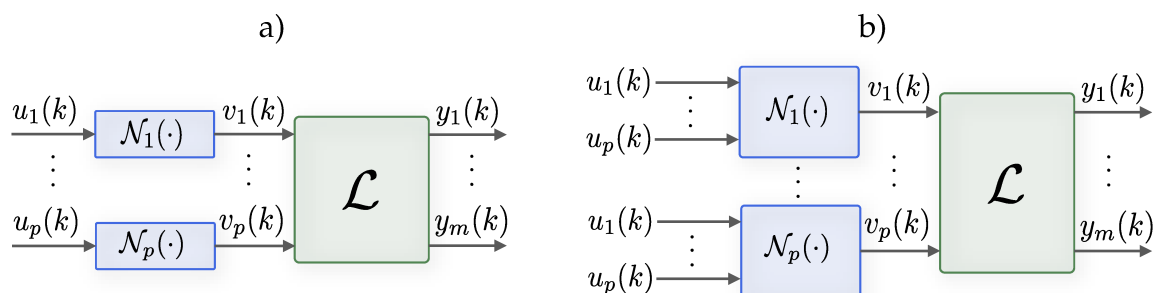


Figura 2.2: Possíveis variações do modelo de Hammerstein multivariável, inspiradas nos modelos MISO apresentados em (Ribeiro e Aguirre, 2014).

## 2.3 Modelos Neuro-Fuzzy

Nesta seção os modelos **NF** são introduzidos como uma possibilidade para representar a parcela não linear sem memória do modelo de Hammerstein. A representação para a qual o desenvolvimento é feito é inspirada na estrutura monovariável, ou seja, aquela apresentada em azul na Figura 2.1a). Desse modo, a função estática não linear  $\mathcal{N}_1$  é retratada matematicamente como  $v_1(k) = \mathcal{N}_1(u_1(k))$ .

Os modelos **NF** surgem como uma mescla de redes neurais e sistemas *fuzzy* na busca de somar as vantagens dessas representações e construir um modelo híbrido capaz de representar não linearidades variadas. Os modelos **NF** descrevem sistemas por meio de regras *fuzzy* SE – ENTÃO organizadas em uma estrutura de rede, nas quais algoritmos de aprendizagem conhecidos da área de redes neurais artificiais podem ser aplicados (Nelles, 2020). Para entender esses pontos, é apresentada uma breve revisão dos modelos *fuzzy* e das redes neurais.

Segundo Babuška e Verbruggen (2003), os conjuntos *fuzzy* são baseados em uma forma de pensar mais próxima à do ser humano, que transmite incerteza ou imprecisão ao se expressar em uma linguagem natural. O conceito é traduzido em um raciocínio lógico suave, de modo que valores intermediários são definidos entre avaliações convencionais como alto ou baixo, quente ou frio, entre outros correlatos.

Um modelo matemático que, de alguma forma, usa conjuntos nebulosos é chamado de modelo *fuzzy* (Babuška e Verbruggen, 2003). Na identificação de sistemas, geralmente, aplicam-se modelos *fuzzy* baseados em regras. Neles, as relações entre variáveis são representadas por meio de regras SE – ENTÃO com valores linguísticos, tais como

SE o fluxo é grande ENTÃO o reservatório enche rápido.

Esse exemplo relaciona, por meio de uma linguagem característica do ser humano, que é imprecisa, a relação entre o fluxo de entrada de líquido e o nível de um reservatório. Uma maneira de dar significados matemáticos aos termos “grande” e “rápido” é por meio dos conjuntos *fuzzy*, mais especificamente, por meio das funções de pertinência que mapeiam os elementos do universo considerado para o intervalo unitário  $[0, 1]$ . Os valores extremos 0 e 1 denotam não associação e adesão completa, respectivamente, enquanto um grau entre 0 e 1 significa associação parcial ao conjunto *fuzzy*. Segundo a estrutura das regras SE – ENTÃO, pelo menos dois tipos de modelos *fuzzy* podem ser construídos: o modelo Mamdani e o modelo Takagi-Sugeno. No modelo Mamdani, o antecedente (SE) e o consequente (ENTÃO) são conjuntos *fuzzy*, enquanto, no modelo Takagi-Sugeno, a parte consequente é definida por uma função estabelecida sobre o conjunto *crisp*. Ressalta-se que a nomenclatura *crisp* se refere ao conjunto numérico clássico no qual se aplica o raciocínio nítido. Em tal entendimento nítido, um objeto, ou pertence, ou não pertence a determinado grupo, sendo excluída a possibilidade de pertencer com diferentes graus (Sumathi e Panerselvam, 2010).

Diferente do modelo Mamdani, que é tipicamente usado em sistemas base-

ados em conhecimento, o modelo Takagi-Sugeno tornou-se popular na identificação orientada por dados (Sumathi e Paneerselvam, 2010). Nesse modelo, o antecedente é definido como um conjunto *fuzzy*, enquanto o consequente é uma função linear afim das variáveis de entrada, do seguinte modo

$$\boxed{\text{SE antecedente é satisfeito ENTÃO saída} = f(u_1(k)),}$$

em que  $f(u_1(k))$  é uma função *crisp*. Quando  $f(u_1(k))$  é especificada como uma constante, por exemplo  $\vartheta$ , tem-se o modelo Takagi-Sugeno de ordem zero

$$\boxed{\text{SE antecedente é satisfeito ENTÃO saída} = \vartheta.}$$

No contexto da área de inteligência computacional, existem diversas técnicas inspiradas em sistemas biológicos, como o comportamento social de pássaros, a teoria da evolução ou até mesmo mecanismos do cérebro humano. As redes neurais artificiais, originalmente inspiradas na funcionalidade de redes neurais biológicas, podem aprender relações complexas generalizando a partir de um determinado conjunto de dados de treinamento. As redes neurais podem, portanto, servir como modelos caixa preta de sistemas estáticos e dinâmicos não lineares, multivariáveis e podem ser treinadas usando dados de entrada e saída observados no sistema. As redes neurais artificiais mais comuns consistem em várias camadas de elementos de processamento simples chamados neurônios, eles são interconectados em redes em que cada conexão recebe um peso diferente. As informações relevantes para o mapeamento feito entre a entrada e saída da rede são armazenadas nos pesos (Babuška e Verbruggen, 2003).

Redes **NF** são modelos *fuzzy* projetados com uma estrutura escolhida pelo usuário, mas são treinados a partir de dados, de modo que não é necessário ter conhecimento prévio sobre o processo a ser identificado. As estreitas ligações entre modelos *fuzzy* e as redes neurais motivaram as primeiras abordagens para modelagem *fuzzy* baseada em dados (Nelles, 2020).

Em redes **NF**, diferentes componentes podem ser otimizados a partir dos dados e um deles são os parâmetros consequentes. Os parâmetros consequentes da regra correspondem aos pesos de saída que, no modelo Takagi-Sugeno de ordem zero, são as constantes  $\vartheta$ . Portanto, elas podem ser otimizadas, por exemplo, seguindo o esquema apresentado na Figura 2.3. Os parâmetros antecedentes são fixados e apenas os parâmetros da camada de saída são otimizados. No desenvolvimento a ser detalhado no Capítulo 3 desta dissertação, os dados que compõem a entrada  $u(k)$  são classificadas em *clusters* que consistem em conjuntos *fuzzy*, ou seja, após a classificação cada dado pertencerá a pelo menos um conjunto *fuzzy* que são chamados de *clusters*. Logo, até esse ponto os desenvolvimentos é como apresentado ainda na Figura 2.3. Na sequência, os parâmetros consequentes são computados de forma analítica, sem a necessidade de aplicar a otimização, ou seja, diferente do que apresenta-se no diagrama da Figura 2.3.

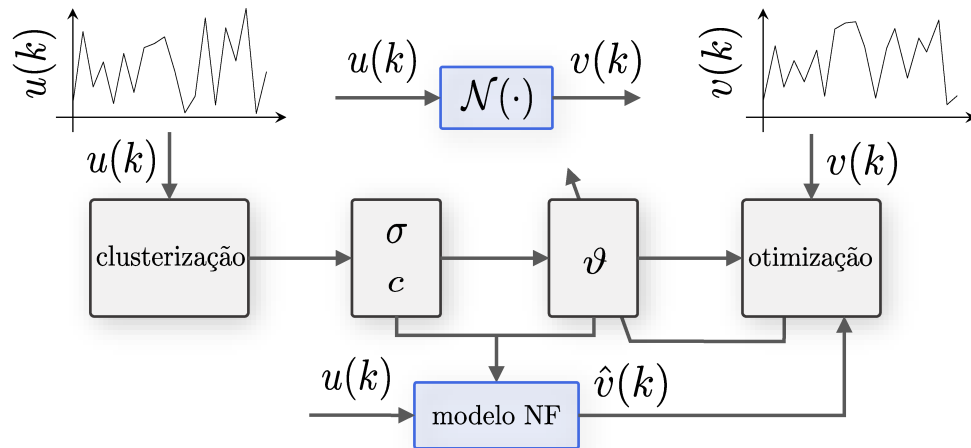


Figura 2.3: Esquema do processo que pode ser utilizado para estimar os parâmetros de um modelo NF. As variáveis  $c$  e  $\sigma$  representam o centro e a dispersão de uma função de pertinência empregada na etapa de classificação. Por sua vez, um processo de otimização é aplicado para aproximar o parâmetro consequente.

## 2.4 Modelos de Histerese Monovariáveis

A seguir, os modelos de histerese são introduzidos como uma possibilidade para representar a parcela não linear do modelo de Hammerstein. De maneira complementar ao desenvolvimento para não linearidades estáticas sem memória, nesta seção, a construção é feita para modelos com não linearidades estáticas com memória, também chamada de não linearidade quase estática. Nesta seção, a representação do modelo de Hammerstein também será inspirada na estrutura monovariável apresentada em azul na Figura 2.1a), tal que a função estática não linear será retratada matematicamente como  $v_1(k) = \mathcal{N}_1(u_1(k))$ .

Diversos modelos fundamentados em operadores e equações diferenciais têm sido utilizados para representar a histerese que é um fenômeno não linear forte que se manifesta em sistemas físicos. Dentre os desenvolvimentos embasados em operadores, pode-se citar os modelos de Preisach e P-I. Por outro lado, os modelos de Bouc-Wen e Duhem são exemplos de representações matemáticas desenvolvidas com base em equações diferenciais (Gan e Zhang, 2019). As principais vantagens dos modelos P-I são o fato de utilizarem um número menor de parâmetros do que outros modelos de histerese baseados em operadores e maior simplicidade quando comparados à abordagem de equações diferenciais. Adicionalmente, um modelo que corresponda à inversa da não linearidade pode ser obtido analiticamente, como foi estudado com detalhes em (Aljanaideh et al., 2018). Por fim, vale ressaltar que eles são mais apropriados para aplicações em tempo real do que os modelos de Preisach (Oubellil et al., 2015).

Uma das estruturas capazes de exprimir o fenômeno não linear da histerese por meio de equações diferenciais é o modelo de Bouc-Wen, que foi originalmente proposto por Robert Bouc em 1967 e, posteriormente, generalizado para descrever

comportamentos como a histerese de atuadores piezoelétricos por Yi-Kwei Wen em 1976 (Bouc R. 1967; Wen K. I. 1976 *apud* Song e Der Kiureghian 2006). A proposta é que a relação não linear entre a entrada  $u_1(t)$  a saída  $v_1(t)$  seja reproduzida no instante  $t$  pela seguinte equação

$$\dot{x}(t) = b_1 \dot{u}_1(t) - b_2 \dot{u}_1(t) |x(t)|^{b_4} - b_3 \dot{u}_1(t) |x(t)|^{b_4-1} x(t), \quad (2.1)$$

sendo  $b_1$  a  $b_4$  os parâmetros do modelo de Bouc-Wen que podem ser alterados de modo a formar diferentes laços de histerese, todos simétricos. Como a implementação envolve uma derivada, ela é considerada mais complexa que os casos baseados em operadores (Gan e Zhang, 2019).

Semelhante a representação proposta por Bouc e Wen, o modelo de Duhem é construído com o auxílio de derivadas para explicar a relação não linear entre a entrada  $u_1(t)$  e a saída  $v_1(t)$ . Segundo o modelo de Duhem essa relação é

$$\dot{v}_1(t) = d_0 |\dot{u}_1(t)| [f_1(u_1(t)) - v_1(t)] + \dot{u}_1(t) f_2(u_1(t)), \quad (2.2)$$

sendo  $d_0$  uma constante positiva e  $f_1(\cdot)$  e  $f_2(\cdot)$  funções aplicadas sobre a entrada  $u_1$  no instante  $t$ . Esse modelo também é complexo devido às derivadas e pode ser utilizado para representar histereses dependentes da taxa. Ao observar o sinal resultante em (2.2), nota-se que ele deve ser integrado para obter a saída  $v_1(t)$ . Logo, as representações matemáticas de Bouc-Wen e Duhem necessitam de algoritmos de integração numérica que podem ser computacionalmente mais custosos que os modelos baseados em operadores (Gan e Zhang, 2019).

### 2.4.1 Modelo de histerese de Preisach

Por sua vez, o modelo de Preisach é uma representação matemática fundamentada em operador, proposta inicialmente para estudar materiais ferromagnéticos com histerese. O modelo básico de Preisach consiste em uma combinação ponderada de operadores relés elementares como aquele apresentado na Figura 2.4 (Mayergoyz, 1991). Matematicamente, o modelo de Preisach pode ser descrito como

$$v_1(t) = \int \int f_{1,p_0p_1} f_{2,p_0p_1}(u_1(t)) dp_1 dp_0, \quad (2.3)$$

sendo  $f_{1,p_0p_1}$  uma função de densidade que descreve a relação entrada e saída e pode ser obtida experimentalmente (Mayergoyz, 1991).

O modelo de Preisach consegue reproduzir variadas formas de laços de histereses, no entanto, pode ter um custo computacional mais elevado se comparado ao modelo P-I que emprega um número reduzido de parâmetros (Oubellil et al., 2015). Tendo em vista esses aspectos, a subseção a seguir apresenta os fundamentos e a descrição matemática do modelo P-I, que serão empregados na metodologia desenvolvida neste trabalho.

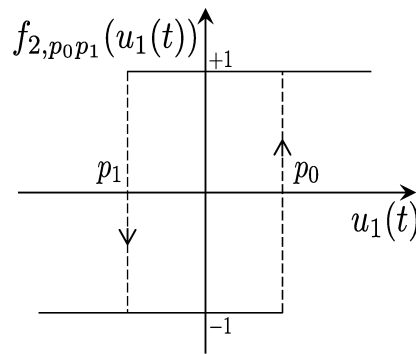


Figura 2.4: Operador relé elementar com dois disparos  $p_0$  e  $p_1$  e dois níveis  $-1$  e  $+1$ . Supondo que, inicialmente,  $f_{2,p_0p_1}(u_1(t)) = -1$ , o valor da função  $f_{2,p_0p_1}(u_1(t))$  é mantido em  $-1$  enquanto  $u_1(t) < p_0$ . Por outro lado, quando  $u_1(t)$  se torna igual ou superior a  $p_0$  o valor da função  $f_{2,p_0p_1}(u_1(t))$  é mantido em  $+1$  enquanto  $u_1(t) > p_1$  e em  $u_1(t)$  igual ou menor que  $p_1$  o valor volta a ser  $-1$ .

### 2.4.2 Modelo de histerese de Prandtl-Ishlinskii

A saída do modelo **P-I** é definida como uma soma ponderada de  $N_P$  operadores elementares aplicados a uma entrada  $u$  no instante de tempo  $k$ , ver Figura 2.5. Diferentes operadores são descritos na literatura para compor o modelo **P-I**, como os

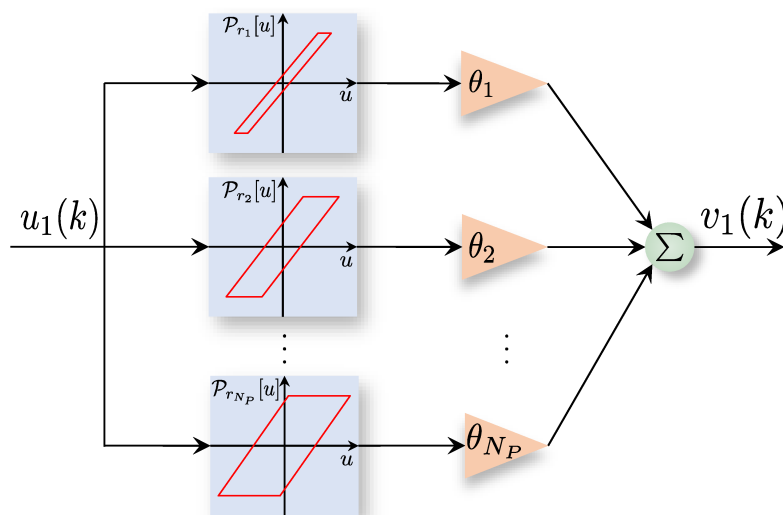


Figura 2.5: Diagrama do funcionamento do modelo **P-I** para reproduzir o fenômeno da histerese. Apresenta-se a soma, ponderada pelos parâmetros  $\theta_j$ , de  $N_P$  operadores  $\mathcal{P}_{r_j}$ . Adaptado de (Al Janaideh et al., 2023).

operadores *stop*, *play* e *play* generalizado. Eles podem ser entendidos como os elementos fundamentais do modelo, que em conjunto, formam a característica geral não linear que é a histerese, tal como ocorre no modelo de Preisach. Um desses operadores é o

*stop*. Ele pode ser descrito de maneira matemática por

$$\mathcal{S}_{r_j}[u](k) = \min\{r_j, \min\{-r_j, u_1(k) - u_1(k-1) + \mathcal{S}_{r_j}[u](k-1)\}\}, \quad (2.4)$$

sendo  $k$  o número da amostra atual na qual a equação é avaliada,  $u_1(k)$  o valor da  $k$ -ésima amostra de entrada e  $r_j$  o valor de disparo ou raio do operador  $j$ , tal que  $j = 1, 2, \dots, N_p$ . A representação gráfica desse operador está mostrada em um exemplo ilustrativo na Figura 2.6a). Os sinais que caracterizam essa forma são apresentados na Figura 2.6d). Em preto está o sinal de entrada  $u$ , constituído de uma senoide cuja amplitude tem o valor 2 e em vermelho tracejado está a saída correspondente,  $\mathcal{S}_{r_j}[u](k)$ . Nesse operador, os disparos modificam a região horizontal, tal que o valor escolhido para a Figura 2.6 é  $r = 1$ .

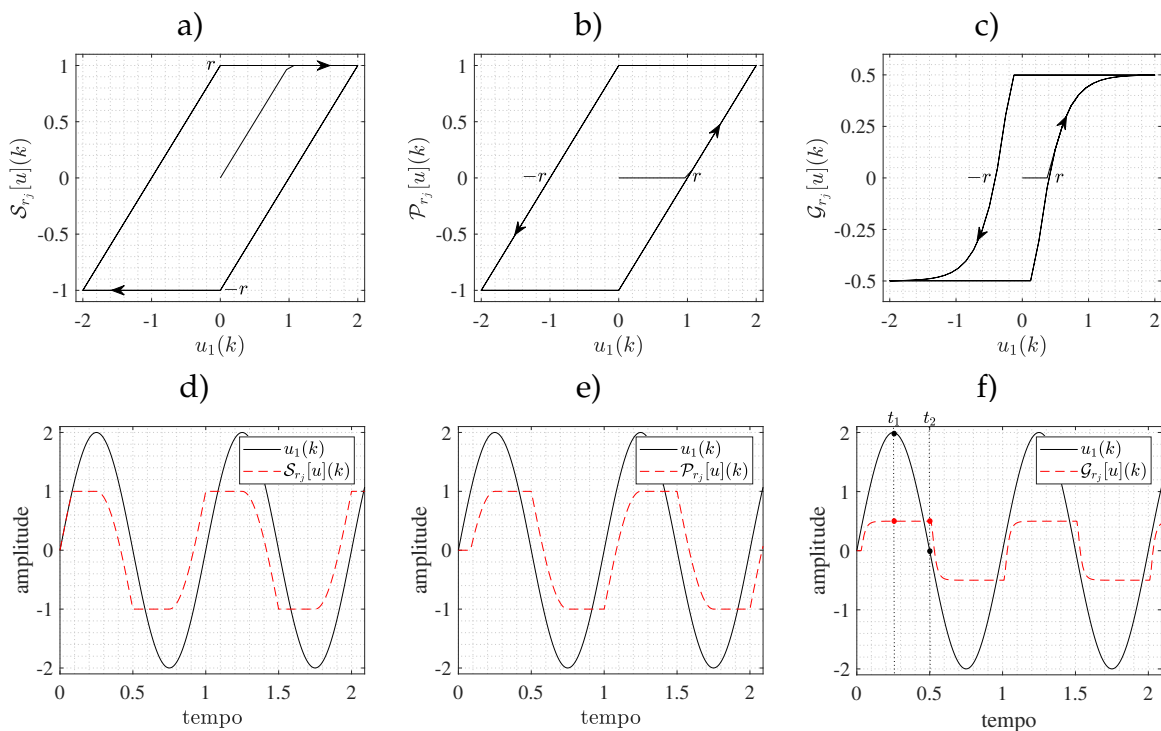


Figura 2.6: Exemplos ilustrativos dos operadores *stop*, *play* e *play* generalizado. Nos gráficos da parte superior a), b) e c) esses três operadores são apresentados em sequência. Logo abaixo, em d), e) e f), estão os dados de entrada e saída que geram o formato de cada um dos operadores ao qual estão alinhados verticalmente. O gráfico f) tem dois instantes em destaque que são discutidos no texto.

Outro tipo de operador utilizado com frequência para compor o modelo P-I é o *play*, utilizado para esquematizar a estrutura do modelo P-I na Figura 2.5. Por meio de uma lógica parecida com a do operador *stop*, o operador *play* pode ser descrito por

$$\mathcal{P}_{r_j}[u](k) = \max\{u_1(k) - r_j, \min\{u_1(k) + r_j, \mathcal{P}_{r_j}[u](k-1)\}\}, \quad (2.5)$$

que é uma forma considerada compacta. Outra forma intuitiva de descrever matematicamente o operador *play*, que, com modificações, também se aplica ao operador *stop*,

é mostrada a seguir

$$\mathcal{P}_{r_j}[u](k) = \begin{cases} u_1(k) + r_j, & \text{se } u_1(k) < u_1(k-1) \text{ e } u_1(k) + r_j < \mathcal{P}_{r_j}[u](k-1), \\ u_1(k) - r_j, & \text{se } u_1(k) > u_1(k-1) \text{ e } u_1(k) - r_j > \mathcal{P}_{r_j}[u](k-1), \\ \mathcal{P}_{r_j}[u](k-1), & \text{c.c.} \end{cases} \quad (2.6)$$

A perspectiva intuitiva que essa maneira de descrever o operador *play* apresenta pode ficar mais clara ao observar a Figura 2.6e). Nela é possível constatar que o valor da saída é mantido constante entre o instante no qual o valor da derivada de  $u_1(k)$  inverte o sinal e o momento mais próximo adiante em que  $u_1(k)$  supera a faixa de  $\max(u) - 2r$  ou  $\min(u) + 2r$ . Por exemplo, verifique que no instante  $t_1 = 0,25$  a derivada do sinal de entrada  $u_1(k)$  deixa de ser positiva e passa a ser negativa. Isso é marcado pelo ponto preto no sinal  $u_1(k)$ , mas também nesse momento o sinal de saída, marcado em vermelho, torna-se constante. Esse sinal permanece inalterado até o momento mais próximo em que o sinal de entrada passa por  $u_1(k) = \max(u) - 2r = 2 - 2 = 0$ . Isso ocorre no instante  $t_2 = 0,5$  em que se marca um ponto em preto o cruzamento com o zero e em vermelho o fim da constância do sinal de saída. O padrão se repete na sequência. Isso é intuitivo em (2.6) ao notar que os valores nos quais o sinal de saída se estagnam são  $\max(u) - 2r$  e  $\min(u) - 2r$ .

Por fim, a mesma representação (2.6) pode ser modificada para formar o operador *play* generalizado que tem sua entrada transformada por uma função  $\mathcal{B}$ .

$$\mathcal{G}_{r_j}[u](k) = \begin{cases} \mathcal{B}(u_1(k)) + r_j, & \text{se } u_1(k) < u_1(k-1) \text{ e } \mathcal{B}(u_1(k)) + r_j < \mathcal{G}_{r_j}[u](k-1), \\ \mathcal{B}(u_1(k)) - r_j, & \text{se } u_1(k) > u_1(k-1) \text{ e } \mathcal{B}(u_1(k)) - r_j > \mathcal{G}_{r_j}[u](k-1), \\ \mathcal{G}_{r_j}[u](k-1), & \text{c.c.} \end{cases} \quad (2.7)$$

Essa função altera as bordas do operador possibilitando que a borda esquerda  $\mathcal{B}_e$  difira da borda direita  $\mathcal{B}_d$ . Se as funções que alteram as bordas esquerda e direita, forem diferentes, o laço de histerese modelado pode assumir formas assimétricas, diferente dos operadores anteriores que eram simétricos.

**Suposição 8** *As funções de borda devem ser contínuas e não decrescentes, além de obedecer a seguinte relação*

$$\mathcal{B}_e \leq \mathcal{B}_d. \quad (2.8)$$

Algumas funções de borda — esquerda ou direita — que podem ser encontradas na literatura são as polinomiais, como por exemplo

$$\mathcal{B}_p = b_0 u_1(k)^3 + b_1 u_1(k), \quad (2.9)$$

em que  $b_0$  e  $b_1$  são constantes que alteram a forma da curva  $\mathcal{B}_p$  e as funções tangente hiperbólica

$$\mathcal{B}_h = b_0 \tanh(b_1 u_1(k) + b_2) + b_3, \quad (2.10)$$

sendo os parâmetros de  $b_0$  a  $b_3$  constantes a serem definidas a fim de moldar o formato

das curvas (Al Janaideh et al., 2023).

**Observação 3** Apesar da notação idêntica, os parâmetros  $b_0$  e  $b_1$  em (2.9) e (2.10) não têm o compromisso de serem iguais. Como no decorrer deste trabalho aplica-se somente a função de borda (2.10) evita-se o conflito de variáveis. A notação é mantida igual entre as duas funções para facilitar a descrição da Figura 2.7.

Graficamente essas funções podem ser observadas nas Figuras 2.7a) e b) para algumas variações de parâmetros. De forma específica,  $b_0$  e  $b_1$  são alterados como combinações do conjunto  $\{0,1, 1\}$ , ou seja, assumem o valor 0,1 ou 1. Os parâmetros  $b_2$  e  $b_3$  de (2.10) são considerados nulos no restante deste trabalho, inclusive na Figura 2.7. Apesar de ser apresentada apenas uma curva para cada situação nas Figuras 2.7a) e b), para compor o modelo de histerese assimétrico, são utilizadas duas funções de borda, a direita, empregada quando a derivada do sinal de entrada é positiva e a esquerda, aplicada quando a derivada do sinal de entrada é negativa. Para que as bordas sejam diferentes e caracterizem a assimetria, basta escolher parâmetros  $b_0$  diferentes entre si que obedeçam (2.8) para compor cada uma das curvas de borda direita e esquerda e seguir o mesmo procedimento para  $b_1$ . O operador *play* generalizado é apresentado na Figura 2.6c) considerando a função de borda apresentada na Figura 2.7b) em traçado preto contínuo. Tendo em vista qualquer um dos operadores *play*, *stop* ou *play*

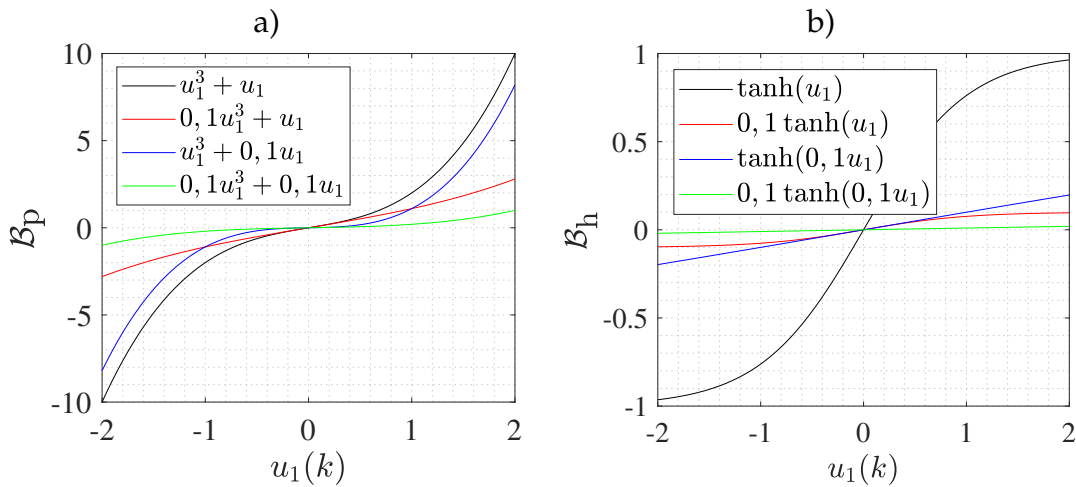


Figura 2.7: Dois tipos de funções que alteram as bordas dos laços de histerese para compor os operadores *play* de modo que eles se tornem operadores *play* generalizados. No gráfico a) apresentam-se 4 variações para a função de borda polinomial  $\mathcal{B}_p$  (2.9) e em b) mais 4 variações para função tangente hiperbólico  $\mathcal{B}_h$  (2.10).

generalizado, a saída do modelo P-I pode ser esquematizada como foi apresentado na Figura 2.5 para o operador *play*. Para os demais operadores basta substituir  $\mathcal{P}_{r_j}[u]$  por  $\mathcal{S}_{r_j}[u]$  ou  $\mathcal{G}_{r_j}[u]$ . De maneira matemática, o esquema pode ser descrito como

$$v_1(k) = \mathcal{N}_1[u](k) = \sum_{j=1}^{N_p} \theta_j \mathcal{P}_{r_j}[u](k), \quad (2.11)$$

em que  $v(k)$  é a saída do modelo e  $\theta_j$  é o peso relacionado ao operador elementar  $\mathcal{P}_{r_j}[u](k)$  que molda o formato do laço de histerese.

### 2.4.3 Modelo GRDPI

Uma estrutura mais geral do modelo **P-I** capaz de representar laços de histerese assimétricos e que se modificam à medida que a frequência da entrada se altera é a do modelo **GRDPI**. A fim de compreender uma das adições que tornam esse modelo mais abrangente, considere um exemplo em que  $\theta_j = 1$  no modelo **P-I** (2.11). Considere ainda que o sinal  $u(k)$  é modificado por uma função de borda como a tangente hiperbólica  $\mathcal{B}_h$  antes de ser inserido nos operadores *play*. Portanto, as bordas deles são alteradas, como se pode observar na Figura 2.8a). Essa é exatamente a descrição do operador *play* generalizado, assim a soma desses 3 operadores sem ponderação resulta no laço de histerese apresentado na Figura 2.8b) que configura o modelo **GPI**

$$v(k) = \mathcal{N}[u](k) = \sum_{j=1}^{N_p} \theta_j \mathcal{G}_{r_j}[u](k). \quad (2.12)$$

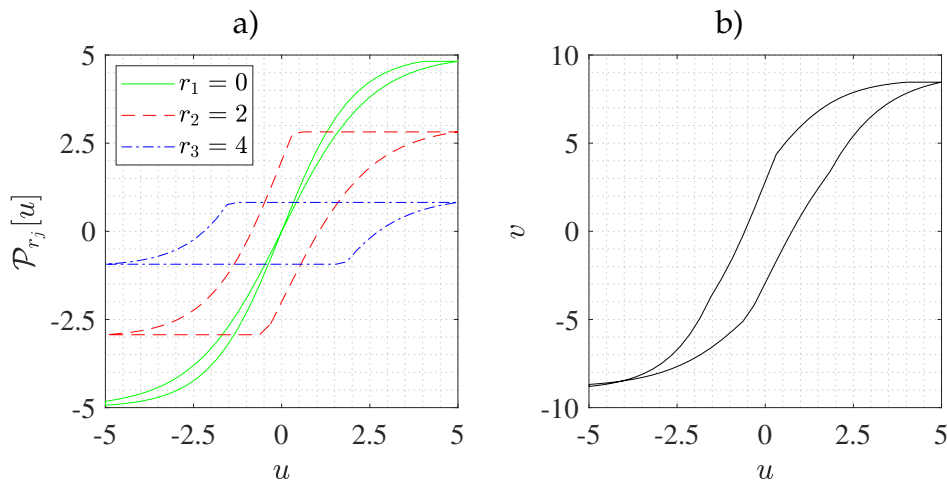


Figura 2.8: Representação de um laço de histerese por meio do modelo **GPI**. Em a) apresentam-se os operadores *play* generalizados, que somados sem ponderação, resultam no laço de histerese assimétrico apresentado em b).

Por sua vez, a dependência da taxa é uma propriedade que processos com histerese podem apresentar de alterar o formato do laço em resposta a uma variação em certa faixa de frequência de excitação. Para verificar esse efeito, aplica-se um sinal senoidal cuja frequência varia de 1 a 10 Hz como pode ser notado na Figura 2.9a). Em processos independentes da taxa, que é o caso apresentado na Figura 2.9b), o laço de histerese permanece inalterado à medida que a excitação sofre alterações na frequência. Esse é precisamente o caso apresentado até este ponto do texto, em que os disparos  $r_j$

são valores constantes computados, em geral, da seguinte maneira

$$r_j = \check{\theta}_1(j - 1). \quad (2.13)$$

Sob outro ponto de vista, os processos dependentes da taxa sofrem alteração no formato do laço de histerese à medida que a frequência de excitação é variada. Esse efeito pode ser observado na Figura 2.9c) que indica que, para o processo simulado, à medida que a frequência aumenta, a largura do laço de histerese também aumenta na faixa testada. Esse efeito é resultado de disparos dinâmicos, que dependem da variação do sinal de entrada.

**Observação 4** Para representar essa dependência com a variação do sinal de entrada emprega-se o símbolo da derivada, mas, a rigor, não é necessário computar a derivada e, sim, uma aproximação proporcional à variação do sinal de entrada a cada iteração.

Como se trabalha com dados amostrados, emprega-se a aproximação da diferença regressiva da derivada. Desse modo, uma descrição básica do disparo dinâmico é a seguinte

$$r_j(\dot{u}) = \check{\theta}_1(j - 1) + \check{\theta}_2 \frac{|u_1(k) - u_1(k - 1)|}{T_s}. \quad (2.14)$$

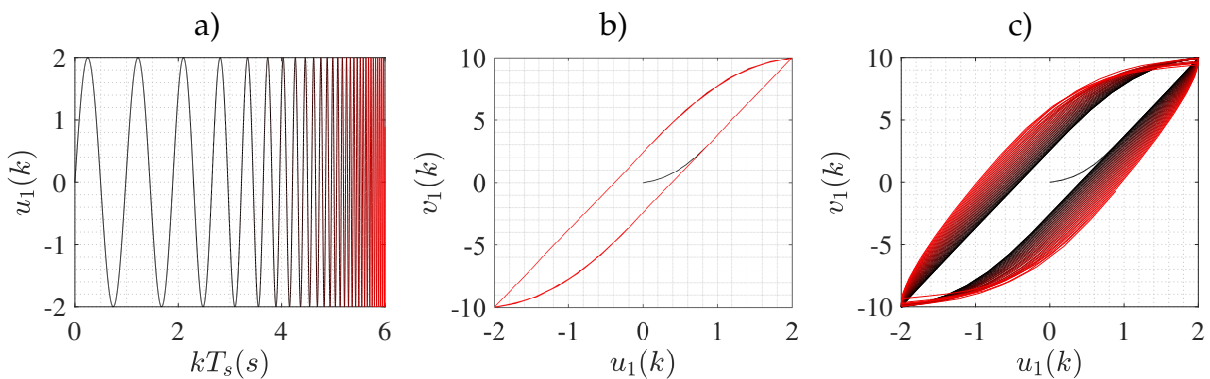


Figura 2.9: Exemplo do comportamento de processos com histerese independente e dependente da taxa quando a frequência da entrada é variada. Em a) apresenta-se o sinal de entrada senoidal  $u_1(k)$ , cuja frequência varia de 1 Hz a 10 Hz de maneira contínua representada pela variação de cor de preto a vermelho. Em b) apresenta-se o laço resultante de um processo independente da taxa quando o sinal  $u_1(k)$  é aplicado em sua entrada e, em c), o laço resultante de aplicar  $u_1(k)$  em um processo com histerese dependente da taxa.

Com a adição das bordas generalizadas e a dependência da taxa, o modelo **P-I** clássico adquire a capacidade de representar mais variedades de laços de histerese. Por exemplo, o modelo **P-I** pode ser usado para caracterizar não-linearidades de histerese que são simétricas e independentes da taxa, enquanto o modelo **GRDPI** pode ser usado para caracterizar não-linearidades de histerese que são assimétricas e dependentes da taxa. Observe que se nenhuma informação estiver disponível sobre a simetria ou

dependência da taxa da histerese, um modelo **GRDPI** é uma escolha interessante, pois os outros modelos, como o **P-I** e **GPI** representam um caso especial que pode ser obtido a partir do modelo **GRDPI**, por exemplo, escolhendo  $\check{\theta}_2 = 0$  e  $\mathcal{B}(u_1(k)) = u_1(k)$  (Aljanaideh et al., 2022).

## 2.5 Modelo para o Bloco Dinâmico Linear e Invariante no Tempo

Como destacado anteriormente, o elemento dinâmico linear  $\mathcal{L}$  do sistema de Hammerstein, pode ser representado por modelos paramétricos ou não paramétricos. Exemplos amplamente conhecidos na literatura são os modelos autorregressivos, a resposta em frequência e os modelos em espaço de estados (Bai, 2010). Alinhado com os objetivos deste trabalho, empregam-se, na representação do elemento dinâmico linear, os modelos paramétricos em espaço de estados tal como apresentado em (1.1) devido à sua harmonia com casos multivariáveis (Bai, 2010; Borjas e Garcia, 2012). A seguir, descreve-se a metodologia de identificação dessa classe de modelos por meio dos métodos de subespaços.

### 2.5.1 Identificação do modelo em espaço de estados

Para estimar um modelo a partir de medições feitas na entrada e saída do processo, os dados devem conter informações suficientes. No caso extremo em que a sequência de entrada  $v(k) = 0$  para toda amostra  $k$ , nenhuma informação sobre a transferência de  $v(k)$  para  $y(k)$  pode ser recuperada. Portanto, a entrada deve ser diferente de zero em certo sentido, para ser possível identificar um modelo consistente. Esta propriedade da entrada em relação à identificação do sistema é geralmente indicada pela noção de persistência de excitação (Verhaegen e Verdult, 2007).

Para evitar tal singularidade, a sequência de entrada  $v(k)$  deve ser “rica” o suficiente. A noção matemática utilizada para expressar esta propriedade de riqueza da entrada é chamada de persistência de excitação. A definição é a seguinte.

**Definição 2.5.1** (Verhaegen e Verdult, 2007) *A sequência  $v(k)$ ,  $k = 0, 1, 2, \dots$  é persistentemente excitante de ordem  $p_n$  se e somente se existe um inteiro  $P_N$  tal que a matriz*

$$\mathcal{V}_{0,p_n,P_N} = \begin{bmatrix} v(0) & v(1) & \dots & v(P_N - 1) \\ v(1) & v(2) & \dots & v(P_N) \\ \dots & \dots & \ddots & \dots \\ v(p_n - 1) & v(p_n) & \dots & v(P_N + p_n - 2) \end{bmatrix}, \quad (2.15)$$

tem posto completo  $p_n$ .

Nos métodos de subespaços, de maneira geral, o primeiro passo é construir as estimati-

vas de estado a partir de observações de dados de entrada e saída do processo com base em ferramentas de álgebra linear. Nesse desenvolvimento, um modelo em espaço de estados é obtido resolvendo um problema de mínimos quadrados. Assim, um ponto importante do estudo dos métodos de subespaços é compreender o ponto-chave de como os vetores de estado e a matriz de observabilidade estendida são obtidos usando ferramentas de álgebra linear numérica. Recapitulando, a vantagem dos métodos de subespaços, por serem baseados em algoritmos numéricos confiáveis de decomposição QR e decomposição em valores singulares (SVD, do inglês *Singular Value Decomposition*), é que não são necessárias técnicas de otimização não lineares, nem impor alguma forma canônica ao sistema. Isto implica que algoritmos de subespaço podem ser igualmente aplicáveis à identificação de sistemas MIMO e SISO. Em outras palavras, os métodos de subespaços não sofrem as inconveniências encontradas na aplicação de métodos de predição de erro (PEM, do inglês *Prediction Error Method*) quando se trata da identificação de sistemas MIMO (Katayama, 2005).

Com o propósito de identificar um modelo em espaço de estados, o algoritmo MOESP proposto em (Verhaegen e Dewilde, 1992a) é utilizado neste trabalho. Tal algoritmo opera os dados de entrada e saída de determinado processo para estimar a ordem e as matrizes de estado de um modelo que assimile a dinâmica linear captada. As matrizes são determinadas em duas etapas distintas: na primeira etapa,  $A$  e  $C$  são obtidas a partir da matriz de observabilidade estendida  $\Gamma_h = [C^T (CA)^T \dots (CA^{h-1})^T]^T$  e, na segunda,  $B$  e  $D$  são computadas utilizando mínimos quadrados (MQ).

A implementação do MOESP está esquematizada na Figura 2.10, em que a entrada é composta pelos dados de ensaio  $\{v(k), y(k)\}_{k=1}^{N_L}$  e uma escolha de projeto, a qual é a quantidade  $h$  de blocos linha da matriz de Hankel. Empregando uma projeção oblíqua de matrizes de Hankel formada pelos dados de entrada e saída, obtêm-se a matriz  $O_h$ , da qual, pode-se extrair as estimativas da ordem do sistema  $\hat{n}$  e de  $\Gamma_h$  por meio da SVD. Manipulando  $\Gamma_h$  estimam-se as matrizes  $\tilde{A}$  e  $\tilde{C}$ . Por fim, por meio dos dados de entrada e saída, de  $\Gamma_h$  e seu complemento ortogonal, estimam-se as matrizes  $\tilde{B}$  e  $\tilde{D}$ .

Os passos para implementação são esquematizados na Figura 2.10 e detalhados a seguir (Verhaegen e Dewilde, 1992a).

- **Passo 1:** construir as matrizes de Hankel  $\mathcal{V}_{h,N_L}$  e  $Y_{h,N_L}$  segundo a forma geral

$$\Delta_{h,N_L} = \begin{bmatrix} \delta_1 & \delta_2 & \dots & \delta_{N_L} \\ \delta_2 & \delta_3 & \dots & \delta_{1+N_L} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_h & \delta_{h+1} & \dots & \delta_{h-1+N_L} \end{bmatrix}, \quad (2.16)$$

em que  $\delta$  representa  $v$  para  $\mathcal{V}_{h,N_L}$  e  $y$  para  $Y_{h,N_L}$ .

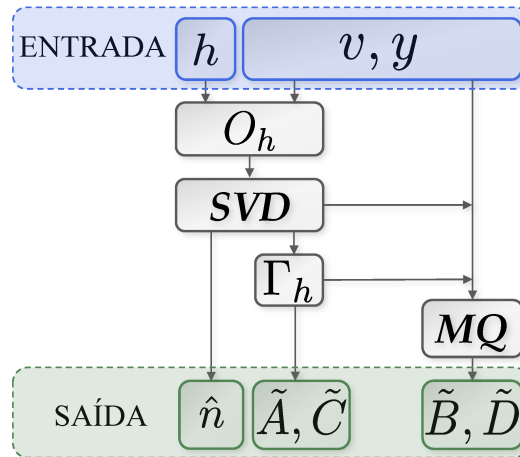


Figura 2.10: Esquema do algoritmo **MOESP** que, a partir de dados de do sinal intermediário  $v$  e saída  $y$ , estima a ordem do modelo e as matrizes em espaço de estados que é capaz de descrever a parte dinâmica linear dos **MBIs**.

- **Passo 2:** particionar os dados de acordo com uma fatoração RQ do seguinte modo

$$\begin{bmatrix} \mathcal{V}_{h,N_L} \\ \mathcal{Y}_{h,N_L} \end{bmatrix} = \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}, \quad (2.17)$$

em que  $R_{11} \in \mathbb{R}^{mh \times mh}$ ,  $R_{12} \in \mathbb{R}^{ph \times mh}$  e  $R_{22} \in \mathbb{R}^{ph \times ph}$ .

- **Passo 3:** nessa etapa a matriz  $R_{22} = O_h$  deve ser desmembrada empregando a **SVD** em  $O_h$

$$O_h = U_o S_o V_o^T = \begin{bmatrix} U_{o,1} & U_{o,2} \end{bmatrix} \begin{bmatrix} S_{o,1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{o,1}^T \\ V_{o,2}^T \end{bmatrix}, \quad (2.18)$$

em que  $S_{o,1} \in \mathbb{R}^{\hat{n} \times \hat{n}}$  é uma submatriz de  $S_o$  contendo os valores singulares não nulos de  $O_h$ . Note que a ordem dessa matriz não nula indica a ordem estimada do modelos em espaço de estados. Das matrizes resultantes da **SVD** pode-se extrair

$$\Gamma_h = U_{o,1} S_{o,1}^{1/2}. \quad (2.19)$$

A matriz de observabilidade estendida  $\Gamma_h$  é definida como a composição das matrizes  $A$  e  $C$  que se deseja estimar

$$\Gamma_h \triangleq \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{h-1} \end{bmatrix} \in \mathbb{R}^{mh \times n}. \quad (2.20)$$

As matrizes  $\tilde{A}$  e  $\tilde{C}$  podem ser obtidas utilizando a notação do *matrix laboratory*

(MATLAB<sup>®</sup>) como

$$\tilde{C} = \Gamma_h(1 : m, :), \quad (2.21)$$

$$\tilde{A} = \Gamma_h(1 : (h-1)m, :)^{\dagger} \Gamma_h(p+1 : hm, :), \quad (2.22)$$

em que o símbolo  $\dagger$  representa a pseudoinversa da matriz  $\Gamma_h$ .

- **Passo 4:** para estimar as matrizes  $B$  e  $D$  considere a definição

$$\Xi \triangleq (U_{o,2})^T R_{21} R_{11}^{-1}. \quad (2.23)$$

Assim, pode-se computar as matrizes utilizando novamente a notação do MATLAB<sup>®</sup> como

$$\left( \begin{bmatrix} U_{o,2}(1 : p, :)^T & \dots & U_{o,2}(p(h-1) + 1 : ph, :)^T \\ U_{o,2}(p+1 : 2p, :)^T & \dots & 0 \\ \vdots & \ddots & 0 \\ U_{o,2}(p(h-1) + 1 : ph, :)^T & 0 & 0 \end{bmatrix} \right) \left[ \begin{array}{c|c} I_p & 0 \\ \hline 0 & \Gamma_h(1 : (h-1)m, :) \end{array} \right]^{\dagger} \begin{bmatrix} \Xi(:, 1 : m) \\ \Xi(:, m+1 : 2m) \\ \vdots \\ \Xi(:, m(h-1) + 1 : mh) \end{bmatrix} = \begin{bmatrix} \tilde{D} \\ \tilde{B} \end{bmatrix}. \quad (2.24)$$

Após esse processo, tem-se as estimativas da ordem da representação em espaço de estados  $\hat{n}$  e das matrizes  $\tilde{A}$ ,  $\tilde{B}$ ,  $\tilde{C}$  e  $\tilde{D}$  que estão a uma transformação de similaridade  $\mathcal{T}$  desconhecida das matrizes originais do sistema, por isso indicadas com o símbolo ( $\sim$ ) em lugar de ( $\wedge$ ).

## 2.6 Conclusão

Devido à divisão da modelagem de um sistema não linear em subproblemas, os MBIs possuem dois subsistemas elementares que são combinados de diferentes maneiras para aproximar separadamente duas características de um processo. Uma delas é a não linearidade estática ou quase estática e a outra é a dinâmica linear. No decorrer deste capítulo, representações foram descritas para cada uma dessas parcelas. A primeira foi o modelo NF, usado para representar as não linearidades estáticas, como foi feito em (Jia et al., 2005) para sistemas SISO. Por outro lado, o modelo GRDPI é descrito como uma representação conveniente para expressar a histerese que é uma não linearidade quase estática. Desenvolvimentos nesse sentido foram feitos em (Aljanaideh et al., 2018, 2019), também para sistemas SISO. Por fim, descreve-se o método de identificação da parcela dinâmica linear representada por um modelo MIMO em espaço de estados. A metodologia detalhada é o MOESP, útil para identificar a parcela dinâmica linear do modelo de Hammerstein MIMO. Alinhado com os objetivos deste trabalho os próximos dois capítulos apresentam o desenvolvimento de extensão

---

dessas representações e metodologias para identificação dos modelos de Hammerstein **MIMO** com não linearidades estáticas ou quase estáticas fortes.

# Identificação de Não Linearidades Fortes Empregando a Metodologia NF

---

## 3.1 Introdução

No Capítulo 2, foram introduzidos conceitos básicos a respeito dos modelos NF. Tais modelos têm capacidade de aproximar uma função não linear com precisão arbitrária (Hornik et al., 1989; Zeng et al., 2000). Essa característica é de interesse para representar não linearidades fortes devido às variações abruptas, típicas em pontos de descontinuidades e que geralmente dificultam a identificação por meio de outras funções amplamente conhecidas, porém, consideradas simples, como a polinomial.

No meio científico, esforços foram direcionados a fim de identificar modelos de Hammerstein com não linearidades estáticas sinuosas e descontínuas empregando redes NF, como em (Jia et al., 2005), (Jia et al., 2016) e (Li et al., 2022). Porém, no melhor do conhecimento do autor, a combinação dessa rede com modelos em espaço de estados, que favorecem a extensão ao caso multivariável, não foi investigada. Tendo em conta esse aspecto não explorado, este capítulo objetiva apresentar um algoritmo para identificar não linearidades fortes para sistemas multivariáveis em tempo discreto empregando o modelo de Hammerstein, composto por uma rede NF e um modelo em espaço de estados.

Inicialmente, desenvolve-se, na Seção 3.2, os modelos NF multivariáveis. A rigor, a representação continua como aquela apresentada em (Jia et al., 2005), porém, com os índices necessários para apresentar um sistema *fuzzy* diferente para cada entrada  $i$ . O Problema 2, que é o problema central deste capítulo, é apresentado ao final da Seção 3.3. Em resumo, ele consiste em identificar processos com não linearidades estáticas fortes por meio de modelos de Hammerstein MIMO NF. Em seguida, na Seção 3.4, uma possibilidade de solução para esse obstáculo é desenvolvida em detalhes, o que resulta no algoritmo de identificação de modelos de Hammerstein MIMO NF. Para ilustrar a funcionalidade da metodologia proposta e discutir particularidades na escolha dos parâmetros, dois exemplos numéricos são estudados na Seção 3.5. Por fim, finaliza-se o capítulo com as considerações a respeito da metodologia e sua implementação na Seção 3.6.

## 3.2 Rede Neuro-Fuzzy Multivariável

Assuma uma rede **NF** constituída de 4 camadas, como ilustrado na Figura 3.1. O subíndice  $i = 1, 2, \dots, p$  evidencia que há uma rede com seus respectivos parâmetros para cada uma das  $p$  entradas. A primeira das camadas, que é a de entrada, encaminha os sinais diretamente para a segunda, onde ocorre a fuzificação. Em seguida, a segunda camada computa a pertinência da variável de entrada a cada *cluster*. Para essa tarefa, a função utilizada é a gaussiana

$$\mu_{\mathcal{F}_i^{(j)}}(u_i(k)) = \exp\left(-\frac{(u_i(k) - c_{i,j})^2}{\sigma_{i,j}^2}\right), \quad j = 1, 2, \dots, N_i, \quad (3.1)$$

na qual  $c_{i,j}$  e  $\sigma_{i,j}$  são, respectivamente, o centro e a dispersão da função de pertinência relacionada ao *cluster*  $\mathcal{F}_i^{(j)}$ , ou seja, o  $j$ -ésimo *cluster* referente à entrada  $i$ . Na terceira camada, o mecanismo de inferência Takagi-Sugeno é implementado (Takagi e Sugeno, 1985). O número de nós nesta camada representa o número  $N_i$  de regras *fuzzy*.

**Observação 5** Conforme observado em (Jia et al., 2005) e (Jia et al., 2016), devido aos resultados satisfatórios de adaptação a diversas curvas e ao custo computacional reduzido, as regras Takagi-Sugeno utilizadas serão de ordem zero. Tais regras são descritas por

$$\mathcal{R}_i^{(j)} : \text{SE } u_i(k) \text{ é } \mathcal{F}_i^{(j)}, \text{ ENTÃO } v_i(k) = \vartheta_{i,j},$$

em que  $\mathcal{R}_i^{(j)}$  denota a  $j$ -ésima regra *fuzzy* SE-ENTÃO relacionada à entrada  $i$ ,  $v_i(k)$  é a saída do modelo e  $\vartheta_{i,j}$  o parâmetro consequente.

Na última camada, a defuzificação é feita como a soma ponderada de todos os pesos consequentes, produzindo a saída

$$v_i(k) = \sum_{j=1}^{N_i} \Phi_{i,j}(u_i(k)) \vartheta_{i,j}, \quad (3.2)$$

em que

$$\Phi_{i,j}(u_i(k)) = \frac{\mu_{\mathcal{F}_i^{(j)}}(u_i(k))}{\sum_{l=1}^{N_i} \mu_{\mathcal{F}_i^{(l)}}(u_i(k))}, \quad j = 1, 2, \dots, N_i. \quad (3.3)$$

Portanto, a rede **NF** é descrita em termos de três parâmetros. Dois deles,  $c$  e  $\sigma$  são da função gaussiana utilizados na classificação dos dados de entrada em *clusters*. O terceiro é o parâmetro consequente da rede,  $\vartheta$ , utilizado na soma ponderada que gera a saída.

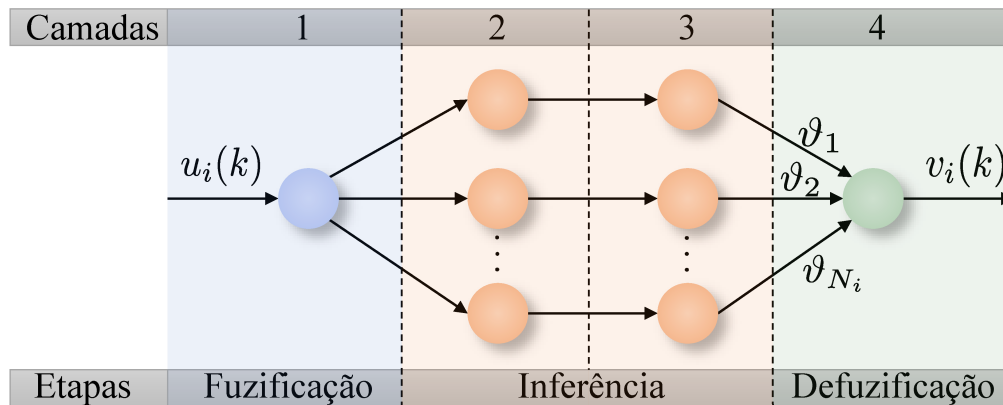


Figura 3.1: Estrutura da rede NF de 4 camadas, empregada para representar a curva estática não linear da  $i$ -ésima entrada do modelo de Hammerstein. A camada 1 recebe a entrada  $u_i(k)$  no instante  $k$ , enquanto a camada 4 retorna a saída  $v_i(k)$  da rede como uma soma ponderada dos parâmetros consequentes  $\vartheta$ . Por sua vez as camadas 2 e 3 realizam a classificação e o cálculo da pertinência dos dados recebidos.

### 3.3 Formulação do Problema para Não Linearidades Estáticas

Considere o modelo de Hammerstein multivariável em tempo discreto constituído de uma função estática não linear  $\mathcal{N} : \mathbb{R}^p \mapsto \mathbb{R}^p$  seguida de um sistema dinâmico LIT MIMO  $\mathcal{L}$ , ambos ilustrados na Figura 3.2. A parcela estática não linear é mostrada em azul sendo composta por  $p$  funções  $\mathcal{N}_i(\cdot) : \mathbb{R} \mapsto \mathbb{R}$ . Considere ainda que as funções  $\mathcal{N}_i(\cdot)$  são independentes entre si, de modo que cada não linearidade  $\mathcal{N}_i(\cdot)$  recebe como entrada o sinal  $u_i(k)$  e resulta em um sinal intermediário  $v_i(k)$ , para  $i = 1, 2, \dots, p$ . Por sua vez, o sistema dinâmico linear, representado em verde na Figura 3.2, recebe como entradas os sinais intermediários  $v_i(k)$ ,  $i = 1, 2, \dots, p$  e devolve como saídas os sinais  $y_l(k)$ ,  $l = 1, 2, \dots, m$ .

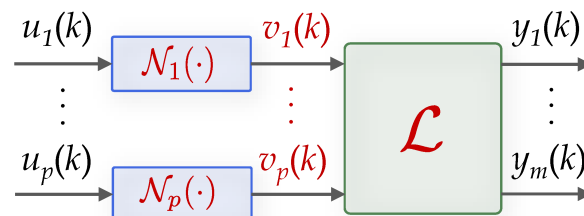


Figura 3.2: Modelo de Hammerstein multivariável em malha aberta com  $p$  entradas  $u_i(k)$ ,  $i = 1, 2, \dots, p$  e  $m$  saídas  $y_l(k)$ ,  $l = 1, 2, \dots, m$ . Em vermelho apresentam-se os elementos desconhecidos e em preto os que são conhecidos. Cada uma das não linearidades da entrada  $\mathcal{N}$  representam um subsistema NF.

Os subsistemas e sinais destacados em vermelho na Figura 3.2 são desconhe-

cidos e supõe-se que sejam inacessíveis. A combinação apresentada no diagrama de Hammerstein é genérica, pois ela não expressa quais as representações cada uma das duas parcelas deverá assumir. À vista disso, alinhado com os objetivos desta dissertação, informa-se que a parcela estática não linear será representada por redes **NF** e a parcela dinâmica linear será constituída por um modelo em espaço de estados de ordem  $n$ . Matematicamente essa composição é descrita como

$$v(k) = \mathcal{N}(u(k)), \quad (3.4a)$$

$$\mathcal{L} = \begin{cases} x(k+1) = Ax(k) + Bv(k), \\ y(k) = Cx(k), \end{cases} \quad (3.4b)$$

em que  $u(k) \in \mathbb{R}^p$ ,  $v(k) \in \mathbb{R}^p$ ,  $x(k) \in \mathbb{R}^n$  e  $y(k) \in \mathbb{R}^m$  são vetores compostos por sinais de entrada, intermediários, de estados e de saída, respectivamente. Cada  $\mathcal{N}_i(\cdot)$ , que compõe a não linearidade  $\mathcal{N}$ , é uma rede **NF** e  $A$ ,  $B$  e  $C$  são as matrizes da dinâmica, de entrada e de saída da representação em espaço de estados, respectivamente.

**Suposição 9** Em (3.4b) a matriz de transmissão direta  $D$  deve ser nula. Isso implica que os sinais inseridos na entrada levam pelo menos um instante de amostragem para produzir efeitos na saída, a qual é uma suposição realista para casos práticos em tempo discreto.

Seguindo a mesma estrutura apresentada em (Jia et al., 2005), a rede **NF**  $\mathcal{N}_i(u(k))$  é constituída de 4 camadas, como pode ser acompanhado na Figura 3.1, sendo que na primeira delas o sinal  $u_i(k)$  é classificado em conjuntos *fuzzy*. Na segunda e terceira camadas o mecanismo de inferência Takagi-Sugeno de ordem zero é implementado. Por fim, na última camada realiza-se a soma ponderada dos valores de pertinência da entrada a cada conjunto para formar o sinal intermediário  $v_i(k)$ .

Tendo em consideração um processo **MIMO** com não linearidades estáticas fortes, dois ensaios diferentes devem ser realizados. O primeiro é um ensaio dinâmico em que se aplicam sinais persistentemente excitantes nas entradas do processo  $u(k)$  e coletam-se os dados resultantes de saída  $y(k)$ . Esse conjunto é denotado por  $\{u(k), y(k)\}_{k=1}^{N_L}$ , sendo  $N_L$  a duração do teste em termos de amostras. Tais dados podem ser empregados para estimar as matrizes em espaço de estados do modelo de Hammerstein, ou seja, extrair as características dinâmicas lineares do sistema.

O segundo é um ensaio estático, nomeado assim por ser utilizado para estimar as curvas estáticas não lineares, descritas neste capítulo por redes **NF**. Nesse ensaio, um sinal que visa excitar diferentes pontos de operação dentro do range de interesse a ser explorado no sistema investigado é inserido na entrada  $\bar{u}_i(k)$ , de modo a excitar somente a não linearidade estática  $\mathcal{N}_i$ . Ao finalizar o ensaio com  $\bar{u}_i(k)$ , a próxima entrada  $\bar{u}_{i+1}(k)$ , é excitada. Assim, o sinal de entrada completo é denotado como  $\bar{u}(k)$ . Para compor o conjunto de dados do ensaio estático coleta-se todas saídas  $\bar{y}(k)$ . Portanto, o conjunto de dados do ensaio estático é denotado por  $\{\bar{u}(k), \bar{y}(k)\}_{k=1}^{N_N}$ , sendo  $N_N$  a duração do teste em termos de amostras. Esses dados podem ser aplicados para estimar os parâmetros  $c_{i,j}$ ,  $\sigma_{i,j}$  e  $\vartheta_{i,j,l}$ , com  $i = 1, 2, \dots, p$ ,  $l = 1, 2, \dots, m$ ,  $j = 1, 2, \dots, N_i$ , sendo  $N_i$  a quantidade

de *clusters* relacionados à entrada  $i$ .

Nesse contexto, define-se o Problema 2.

**Problema 2** Identificar o modelo de Hammerstein *MIMO NF*, apresentado em (3.4) por meio de duas etapas. A primeira etapa consiste em identificar as matrizes  $A$ ,  $B$ ,  $C$  e a ordem  $n$  da representação em espaço de estados (3.4b) por meio de uma sequência de  $N_L$  dados dinâmicos coletados do processo  $\{u(k), y(k)\}_{k=1}^{N_L}$ . A parte complementar do problema consiste em identificar os parâmetros da rede *NF* com base em outro ensaio, composto por  $N_N$  dados estáticos  $\{\bar{u}(k), \bar{y}(k)\}_{k=1}^{N_N}$ . Ou seja, deve-se estimar as variáveis que classificam as entradas em diferentes conjuntos fuzzy  $c_{i,j}$  e  $\sigma_{i,j}$  e os parâmetros consequentes da rede *NF*  $\vartheta_{i,j,l}$ .

### 3.4 Associação da Metodologia Neuro-Fuzzy e de Subespaços

As metodologias que constituem o nome desta seção são descritas na literatura de maneira dissociada, como em (Verhaegen e Dewilde, 1992a) e (Jia et al., 2005), no entanto, a associação delas para compor o modelo de Hammerstein *MIMO* requer considerações que ainda não foram investigadas. Essas ponderações serão feitas durante esta seção à medida que a descrição da metodologia for desenvolvida. Para iniciar, alguns elementos-chave de um modelo de Hammerstein serão discutidos com relação à identificação.

Uma observação importante a respeito dos modelos de blocos interconectados é que, em geral, o sinal intermediário é desconhecido. Um diagrama de um processo é apresentado na Figura 3.3 para exemplificar o problema. No interior desse processo, há dois subsistemas interconectados por um sinal  $v(k)$ , todos três desconhecidos e inacessíveis. No entanto, como dito na introdução do Capítulo 1, o objetivo da identificação de sistemas não são os dois subsistemas tampouco o sinal intermediário, mas sim explicar a relação entre a entrada e a saída do processo. Assim, na identificação de modelos de Hammerstein o objetivo continua o mesmo, porém, adiciona-se a restrição de separar a dinâmica linear e a estática não linear em parcelas diferentes, produzindo inevitavelmente um sinal intermediário possivelmente virtual. Como vantagem dessa separação, tem-se um modelo que pode ser tratado como se fosse composto apenas pela parcela dinâmica linear, desde que haja uma maneira de cancelar a curva estática não linear.

Suponha um sistema estático de modo que para qualquer sinal limitado inserido em sua entrada  $u(k)$  há uma resposta imediata  $v(k)$  de amplitude limitada expresso pela função não linear  $\mathcal{N}$ . Esse ganho pode ser diferente em cada ponto de operação  $u(k)$ , no entanto, suponha que ele é constante e vale  $G_N$ . Além disso, considere que na sequência adiciona-se um sistema dinâmico *LIT*  $\mathcal{L}$  estável em malha aberta, de modo que a resposta em regime permanente a um sinal de entrada constante também é uma constante que vale  $G_L$ . Por fim, adiciona-se novamente a mesma não linearidade  $\mathcal{N}$

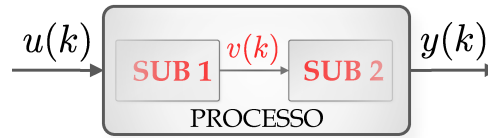


Figura 3.3: Processo com dois subsistemas SUB 1 e SUB 2 e um sinal interno  $v(k)$  desconhecidos. Somente as medidas de entrada e saída estão disponíveis.

como terceiro subsistema. Ou seja, um modelo ainda não apresentado graficamente neste texto, conhecido na literatura como modelo de Hammerstien-Wiener (Giri e Bai, 2010). Ao aplicar qualquer sinal constante na entrada do sistema, obtém-se em regime, na saída  $y(k)$ , uma resposta constante do processo que se pode nomear de  $G_{N\mathcal{L}N}$ . Note que não é possível saber o ganho de cada etapa sem medir os sinais intermediários. Caso o objetivo seja modelar o processo por meio de uma estrutura como a de Hammerstein, uma virtualização do sinal intermediário poderá ser feita mas não corresponderá a nenhum sinal do processo. Pois, observe que o processo tem os três ganhos  $G_N$ ,  $G_{\mathcal{L}}$  e  $G_{N\mathcal{L}}$ , enquanto o modelo de Hammerstein possuirá dois ganhos, um do primeiro e outro do segundo bloco. Ainda que no processo existissem somente os dois primeiros subsistemas totalizando o ganho  $G_{N\mathcal{L}}$ , a menos de um fator de escala  $\bar{\lambda}$  desconhecido, os subsistemas não poderiam ser determinados precisamente. Pois veja que  $G_N G_{\mathcal{L}} = G_{N\mathcal{L}}$  e  $(\bar{\lambda} G_N)(G_{\mathcal{L}}/\bar{\lambda}) = G_{N\mathcal{L}}$ , não se diferem do ponto de vista de entrada-saída (de Paula et al., 2015).

No entanto, essa discussão foi feita com base em um exemplo trivial envolvendo somente parcelas com ganhos constantes. No contexto do modelo de Hammerstein, com parte estática e parte dinâmica, uma saída interessante é estimar o sinal intermediário como proposto em (Sung, 2002). O autor afirma que uma estimativa do sinal intermediário pode ser recuperada se um sinal binário pseudoaleatório (PRBS, do inglês *PseudoRandom Binary Signal*) for utilizado para identificar a parcela dinâmica linear antes da curva estática não linear. Primeiramente, são feitas três considerações acerca do sinal.

**Suposição 10** O sinal de entrada  $u(k)$  deve ser binário, ou seja, composto por dois níveis.

**Suposição 11** Um dos níveis do sinal denotado  $u_\beta$  deve ser não nulo, enquanto o outro deve ser nulo.

**Suposição 12** O nível não nulo deve resultar em  $\mathcal{N}(u_\beta) = \beta \neq 0$ , enquanto o nulo deve resultar em  $\mathcal{N}(0) = 0$ .

Obedecidas essas condições, o sinal intermediário  $v(k)$  será uma versão escalonada do sinal de entrada. A conclusão pode ser obtida recordando que apesar do sinal intermediário receber um ganho distinto para cada ponto de operação no qual o sinal de entrada se encontra, o sinal  $u(k)$  é binário e só o nível não nulo terá um ganho

estático,  $\beta$ , não nulo associado. Um esquema que retrata essa lógica e pode esclarecer o seu funcionamento é exibido na Figura 3.4. Portanto, a menos de um fator de escala,  $\beta$  a ser compensado posteriormente pela rede NF, a parcela dinâmica linear  $\mathcal{L}$  pode ser identificada diretamente com base no conjunto de dados  $\{u(k), y(k)\}_{k=1}^{N_L}$ , ou seja, sem o conhecimento do sinal intermediário.

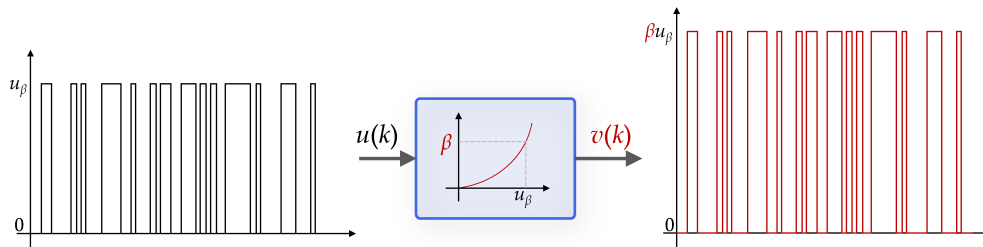


Figura 3.4: Esquema da equivalência entre o sinal de entrada e o sinal intermediário, para identificação, quando as condições estabelecidas nas Suposições 10 a 12 são satisfeitas. Em azul apresenta-se o bloco de uma curva estática não linear genérica. Os elementos em preto são conhecidos, enquanto os que estão em vermelho são desconhecidos.

Tendo em vista que no caso MIMO o modelo de Hammerstein assume a representação KU anunciada na Figura 3.2, não há cruzamento entre as curvas estáticas não lineares. Com a independência que esse modelo apresenta entre as não linearidades, o procedimento indicado na Figura 3.4 pode ser realizado separadamente para cada uma das entradas, todavia, há pelo menos 3 possibilidades para a aplicação do ensaio dinâmico no caso multivariável.

- 1) Aplicar um sinal diferente em cada entrada ao mesmo tempo.
- 2) Aplicar o mesmo sinal em todas as entradas ao mesmo tempo.
- 3) Aplicar um sinal diferente em cada entrada em tempos diferentes.

A opção 1) é geralmente adotada nos processos de identificação MIMO, pois, a 2) leva a problemas de mau condicionamento com as submatrizes obtidas na fatoração RQ, enquanto a 3) não excita todas as entradas ao mesmo tempo, o que não é interessante para o MOESP, que tem a proposta de estimar todas as relações em conjunto. Desse modo, será aplicada a opção 1) que consiste em aplicar sinais diferentes em cada uma das entradas do processo simultaneamente. A amplitude pode ser a mesma, no entanto, o que se altera são os pontos nos quais os sinais PRBS trocam de nível.

### 3.4.1 Identificação da dinâmica linear

Com o propósito de identificar um modelo em espaço de estados, o algoritmo MOESP será utilizado. Tal algoritmo opera os dados de entrada e saída de determinado

processo para estimar a ordem e as matrizes de estado de um modelo que assimile a dinâmica linear contida nos dados. Ou seja, só é possível identificar as características do processo que os dados conseguiram extrair durante o ensaio. Dessa forma, o sinal aplicado na entrada do sistema deve ser um sinal persistentemente excitante, tal que as particularidades dinâmicas do sistema sejam reveladas, como é o caso do sinal PRBS.

A implementação do MOESP está esquematizada novamente na Figura 3.5, no entanto, com as variáveis apropriadas para o modelo de Hammerstein MIMO NF. O algoritmo recebe como entrada os dados de ensaio dinâmico  $\{u(k), y(k)\}_{k=1}^{N_L}$  e uma escolha de projeto, a qual é a quantidade  $h$  de blocos linha da matriz de Hankel. Os sinais PRBS compõem o primeiro conjunto de dados a ser utilizado para a identificação da dinâmica linear. Posteriormente, na Subseção 3.4.2, outros  $p$  conjuntos, separados em faixas serão adicionados para fazer a identificação do restante do modelo de Hammerstein. A ideia é que apenas um teste composto de sinais diferentes seja realizado no processo, gerando uma faixa para cada conjunto de dados

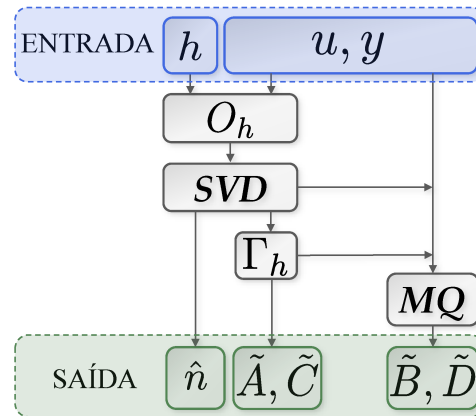


Figura 3.5: Esquema do algoritmo MOESP que, a partir de dados de entrada e saída coletados de um processo, estima a ordem do modelo e as matrizes em espaço de estados, a menos de uma transformação de similaridade  $\mathcal{T}$ .

Os passos para implementação do MOESP foram detalhados na Subseção 2.5.1. As matrizes resultantes são aplicadas na transformação de similaridade  $\mathcal{T} = \tilde{C}^+ I_{\hat{n}}$  de modo que a matriz  $\hat{C}$  seja uma identidade

$$(\hat{A}, \hat{B}, \hat{C}) = (\mathcal{T}^{-1} \tilde{A} \mathcal{T}, \mathcal{T}^{-1} \tilde{B}, \tilde{C} \mathcal{T}). \quad (3.5)$$

Tal transformação é empregada devido às manipulações necessárias para a obtenção do parâmetro  $\vartheta_{i,j}$ , mas restringem a aplicação do método a processos com número de entradas iguais ao número de saídas.

### 3.4.2 Identificação da curva estática não linear

Uma vez obtidos os parâmetros que compõe o bloco dinâmico linear na primeira etapa, pode-se implementar a segunda etapa, que consiste em estimar as não linearidades estáticas  $\mathcal{N}_i$ , para  $i = 1, 2, \dots, p$ . Os parâmetros estimados pela primeira etapa apresentada na Subseção 3.4.1, são as matrizes  $\hat{A}$ ,  $\hat{B}$  e  $\hat{C}$ . Supondo a existência das estimativas desses parâmetros, obtidas por (3.5), apresentada-se na sequência uma maneira de estimar também a saída  $\hat{y}$ . Assim sendo, têm-se as matrizes  $\hat{A}$ ,  $\hat{B}$  e  $\hat{C}$ , os sinais de entrada  $u(k)$ , de saída  $y(k)$  e uma estimativa do sinal de saída  $\hat{y}(k)$  para produzir a estimativa de  $\mathcal{N}_i$ .

Na primeira etapa, o algoritmo de subespaços **MOESP** foi aplicado para estimar um modelo capaz de descrever a relação dinâmica linear das entradas para as saídas do processo. Uma discussão foi feita em termos de três maneiras para aplicar o sinal de teste na entrada do sistema, na qual se decidiu por aplicar  $p$  entradas diferentes ao mesmo tempo e coletar todos os sinais de saída. Essa estratégia funciona bem no **MOESP** devido à sua compatibilidade com o cenário multivariável. Para a metodologia que será descrita no decorrer deste capítulo, o teste estático não pode seguir a mesma estratégia devido à característica independente de cada uma das curvas estáticas do modelo a ser obtido.

Como a abordagem adotada é caixa preta, não há informação a respeito do acoplamento entre as parcelas estáticas não lineares no processo, no entanto, a representação selecionada, que é a **KU**, é complemente desacoplada, o que assegura que esse será o comportamento do modelo (Kortmann e Unbehauen, 1987; Ribeiro e Aguirre, 2014). Para identificá-lo sob essa estrutura, o sinal de teste a ser aplicado obedece a opção 3) que consiste em aplicar um sinal diferente em cada entrada em tempos diferentes. Logo, com o intuito de captar o efeito da não linearidade relacionada somente a cada uma das entradas por vez, o sinal de teste é dividido virtualmente em  $(p + 1)$  faixas no tempo. A primeira faixa é dedicada ao ensaio com o sinal **PRBS** que foi apresentado na Figura 3.4 e, em cada faixa, a partir da segunda, uma entrada é excitada por um sinal multinível enquanto as demais permanecem nulas.

O diagrama da Figura 3.6 é útil para esclarecer como devem ser feitos os ensaios e como o algoritmo para obter o modelo **NF** se encaixa no processo de obtenção das curvas estáticas não lineares. Em virtude dos diversos elementos acoplados entre si no diagrama, para evitar dubiedade, volta-se a atenção para a identificação da curva estática não linear  $\mathcal{N}_1(\cdot)$ , que é apresentada na região pontilhada superior. Como o intuito dessa parte é estimar apenas a não linearidade que é excitada pela entrada  $u_1(k)$ , somente ela é excitada com um sinal multinível enquanto as demais entradas são anuladas. Isso pode ser confirmado ao observar a Faixa 2 pintada em lilás na Figura 3.6. Nota-se também que as demais faixas, até  $p + 1$ , estão opacas dentro da região pontilhada superior. Isso ocorre por essas faixas não interessarem para a identificação da não linearidade  $\mathcal{N}_1(\cdot)$ . Em seguida, coletam-se todos os sinais de saída do processo resultantes da aplicação do sinal  $u_1(k)$  e aplica-se o algoritmo **NF** para cada um dos conjuntos de entrada e saída  $\{u_1(k), y_{1,1}(k)\}, \{u_1(k), y_{1,2}(k)\}, \dots, \{u_1(k), y_{1,m}(k)\}$ . Cada conjunto

gera um candidato a  $\hat{N}_{1,l}(\cdot)$ ,  $l = 1, \dots, m$  e, por meio da validação dinâmica do modelo de Hammerstein completo, seleciona-se o  $\hat{N}_1(\cdot)$  que retorna o menor valor normalizado de índice **RMSE** (do inglês, *Root Mean Squared Error*). Na Figura 3.6, isso é simbolizado selecionando a curva estática que mais se aproxima de  $N_1(\cdot)$ .

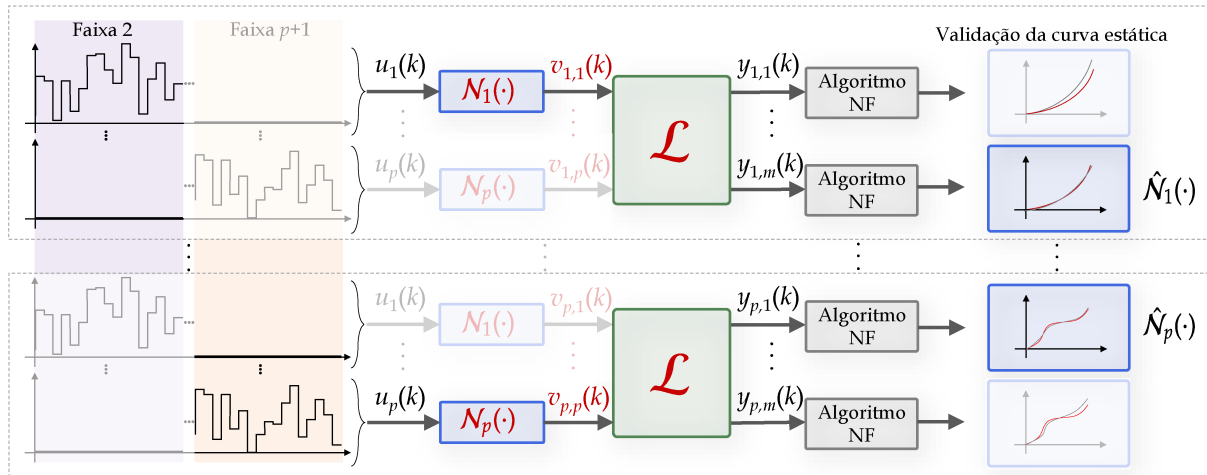


Figura 3.6: Diagrama dos testes feitos com sinal multinível para identificar as curvas estáticas não lineares. O que está em preto refere-se a informações conhecidas e em vermelho os elementos desconhecidos. Para identificar apenas a não linearidade  $N_1(\cdot)$ , aplica-se um sinal multinível em  $u_1(k)$ , coletam-se as saídas e aplica-se o resultado no algoritmo **NF**. No fim, por meio do **RMSE**, seleciona-se a melhor curva estática. A passagem do tempo e conseqüentemente a ordem nas quais os ensaios são realizados ocorre no diagrama de cima para baixo e da esquerda para a direita.

Alterando o foco para a região destacada dentro do retângulo pontilhado na parte inferior da Figura 3.6, ilustra-se como obter uma estimativa da curva estática  $N_p(\cdot)$ . O procedimento não se difere do que foi descrito no parágrafo imediatamente anterior, exceto na imposição de que a entrada excitada é somente a  $u_p(k)$ , uma vez que é a única que interfere na não linearidade que se deseja estimar nessa região do diagrama.

De maneira generalizada, como indicam as reticências entre os diagramas superior e inferior, em uma dada faixa de teste ( $i + 1$ ), a única entrada não nula  $u_i$ , afeta as  $m$  saídas denotadas  $y_{i,l}$ , para  $l = 1, 2, \dots, m$ . Visto que a não linearidade que afeta as saídas  $y_{i,l}$  é a mesma, são identificadas  $m$  candidatas  $\hat{N}_{i,l}(\cdot)$ . No fim, a decisão de qual das  $m$  candidatas utilizar é feita quantitativamente por meio do índice **RMSE** normalizado resultante da validação dinâmica do modelo de Hammerstein completo.

Dada a semelhança entre algumas notações do algoritmo **NF** que podem dificultar a reprodução, a diferença entre elas é esclarecida previamente.  $N_N^{(i)}$  indica a quantidade de dados resultantes do ensaio estático realizado na entrada  $i$ .  $c_{i,j}$  indica o centro do  $j$ -ésimo *cluster* formado pelos dados da entrada  $i$ .  $\bar{N}_{i,j}$  expressa a quantidade de dados que estão formando o  $j$ -ésimo *cluster* da entrada  $i$ . Por fim,  $N_i$  representa a

quantidade total de *clusters* para a entrada  $i$ .

Para exemplificar, suponha que um experimento tenha sido realizado em um processo com 2 entradas no qual se coletou um conjunto de 10 dados para cada entrada, como se pode observar na Figura 3.7. Suponha que, ao aplicar o procedimento de clusterização nos 10 dados dessa primeira entrada, foram formados 2 *clusters*, sendo que um dos *clusters* agregou 4 dos dados totais e os 6 restantes terminaram no segundo *cluster*. Por sua vez, os outros 10 dados relativos à segunda entrada formaram 3 *clusters*, tendo 3 dados no primeiro *cluster*, 5 no segundo e 2 no terceiro. Logo,  $N_1 = 2$  e  $N_2 = 3$ ,  $N_N^{(1)} = N_N^{(2)} = 10$ ,  $\bar{N}_{1,1} = 4$ ,  $\bar{N}_{1,2} = 6$ ,  $\bar{N}_{2,1} = 3$ ,  $\bar{N}_{2,2} = 5$ ,  $\bar{N}_{2,3} = 2$ .

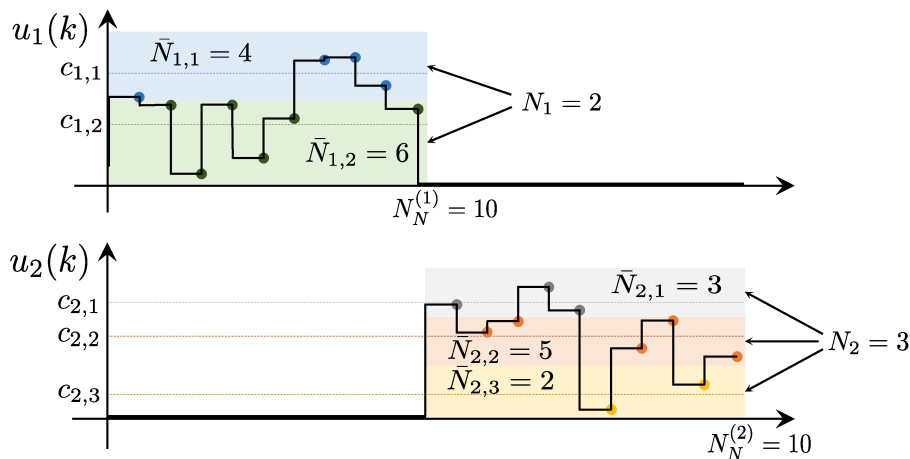


Figura 3.7: Diagramação do exemplo a respeito de como são dadas as nomenclaturas aos dados necessários para implementação da *clusterização* no algoritmo NF.

Em conformidade com a ideia de que o modelo estimado não terá termos cruzados entre as parcelas estáticas não lineares, os passos a seguir podem ser replicados para cada entrada  $i$ , com  $i = 1, 2, \dots, p$  a fim de obter uma estimativa das curvas estáticas não lineares (Jia et al., 2005).

- **Passo 1:** realiza-se um ensaio com o processo, no qual  $u_i(k)$  é definido como um sinal multinível enquanto as demais entradas permanecem zeradas, ou seja, refere-se ao conjunto de sinais que compõe a faixa  $(i + 1)$  de acordo com a lógica da Figura 3.6. Nesse momento, coleta-se o conjunto dos  $N_N^{(i)}$  dados de ensaio  $\{u_i(k), y_l(k)\}_{k=1}^{N_N^{(i)}}$ ,  $l = 1, 2, \dots, m$ . Considerando cada um desses conjuntos por vez, a primeira amostra de entrada coletada,  $u_i(1)$ , é definida como o centro do primeiro *cluster*, ou seja,  $c_{i,1} = u_i(1)$ . Inicializa-se a quantidade de dados nesse primeiro *cluster* como  $\bar{N}_{i,1} = 1$  e a quantidade total de *clusters* relacionados à entrada  $i$  como  $N_i = 1$ .
- **Passo 2:** para o  $k$ -ésimo dado de treinamento  $u_i(k)$ , verifica-se a sua proximidade com os centros dos *clusters* existentes e encontra-se o mais próximo. Para

mensurar essa proximidade quantitativamente emprega-se

$$S_{i,j} = \max_{1 \leq j \leq N_i} \left( \sqrt{(\exp(-|u_i(k) - c_{i,j}|^2))} \right). \quad (3.6)$$

Pela notação adotada é possível constatar que o  $j$  indica a posição do *cluster* mais próximo do  $k$ -ésimo dado de treinamento  $u_i(k)$ , ou seja, é o  $j$ -ésimo *cluster* que foi nomeado  $j$  por ser o mais próximo do atual dado de treinamento. Graficamente essa lógica é retratada na Figura 3.8. Como o  $j$  é apenas uma variável auxiliar, ele não é armazenado e será sobrescrito a cada iteração.

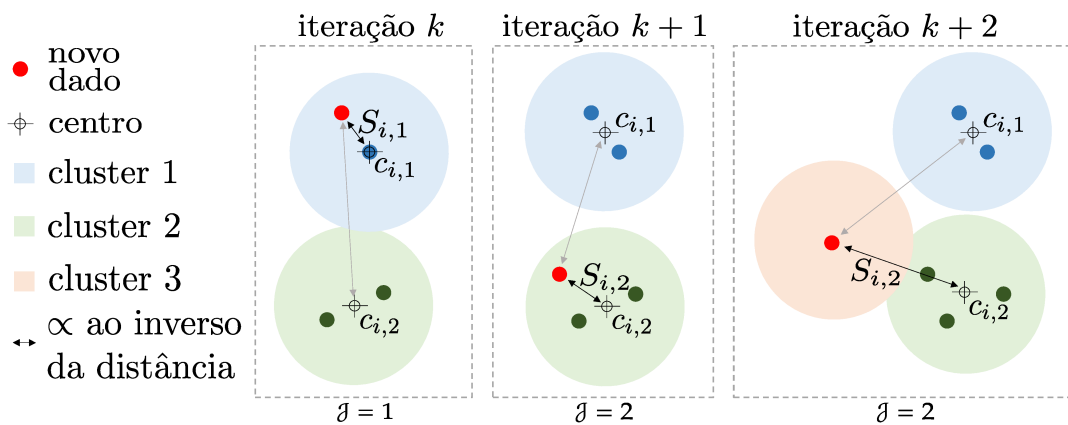


Figura 3.8: Diagrama que ilustra a avaliação de cada dado  $u_i(k)$  a fim de verificar a qual *cluster* ele irá pertencer, ou ainda, se irá formar um novo *cluster*. Na iteração  $k$  do diagrama o  $j = 1$  pois o novo dado  $u_i(k)$  está próximo do *cluster* 1.  $S_{i,j}$  é uma medida inversamente proporcional à distância entre o novo dado e o centro do *cluster*.

- **Passo 3:** na sequência inspeciona-se a necessidade de formar, ou não, um novo *cluster*, segundo o seguinte critério:
  - se  $S_{i,j} < S_{i,0}$ , em que  $S_{i,0} \in [0, 1)$  é um parâmetro escolhido pelo usuário (em geral, próximo de 1), então, o  $k$ -ésimo dado de treinamento não pertence a nenhum *cluster* e cria-se um novo, como é retratado na iteração  $k + 2$  da Figura 3.8. O  $S_{i,0}$  influencia o quanto os *clusters* da entrada  $i$  serão abrangentes. Em outras palavras, quanto maior o valor de  $S_{i,0}$  menor será o raio de cada *cluster* e conseqüentemente menos abrangentes eles serão. A conseqüência é que a curva estimada pode-se tornar menos suave, mas segue melhor as descontinuidades;
  - se  $S_{i,j} \geq S_{i,0}$ , constata-se que o  $k$ -ésimo dado de treinamento  $u_i(k)$  pertence ao  $j$ -ésimo *cluster*. Desse modo, a variável que exprime a quantidade de dados que formam tal *cluster* é acrescida em uma unidade  $\bar{N}_{i,j} = \bar{N}_{i,j} + 1$ . Com a

adição do dado  $u_i(k)$  o centro do  $j$ -ésimo *cluster* é deslocado de acordo com

$$c_{i,j} = c_{i,j} + \frac{\lambda_i}{N_{i,j} + 1}(u_i(k) - c_{i,j}), \quad (3.7)$$

em que  $\lambda_i \in [0, 1]$  é um parâmetro de projeto que controla o quanto cada dado que chega em determinado *cluster* altera a posição atual do centro. De maneira gráfica, a atualização do centro pode ser observada na Figura 3.8 entre as iterações  $k$  e  $k + 1$  para o primeiro *cluster*.

Os passos de 1 a 3 são aplicados para  $k = 1, 2, \dots, N_N^{(i)}$  em  $u_i(k)$ . Finalizada essa etapa, tem-se definida a quantidade de *clusters* e os respectivos centros para a  $i$ -ésima entrada.

- **Passo 4:** o passo seguinte é calcular as dispersões em torno dos centros como

$$\sigma_{i,j} = \min_{\tilde{j}=1,2,\dots,N_i; \tilde{j} \neq j} \frac{|c_{i,j} - c_{i,\tilde{j}}|}{\rho_i}, \quad (3.8)$$

com  $\rho_i \in [1, 2]$ . Sendo  $\rho_i$  o terceiro parâmetro de projeto que o usuário deve escolher. Ele tem a função de tornar o formato da gaussiana mais largo ou estreito. O valor que foi comumente utilizado em (Jia et al., 2005) e (Jia et al., 2016) foi 1,5.

- **Passo 5:** tendo obtido uma estimativa das matrizes  $\hat{A}, \hat{B}, \hat{C}$  e da ordem  $\hat{n}$  pertinentes ao modelo em espaço de estados (3.4b) e os parâmetros  $\vec{c}_i = [c_{i,1}, c_{i,2}, \dots, c_{i,N_i}]$  e  $\vec{\sigma}_i = [\sigma_{i,1}, \sigma_{i,2}, \dots, \sigma_{i,N_i}]$  relacionados ao agrupamento das entradas em *clusters*, pode-se estimar o vetor de parâmetros consequentes  $\vec{\vartheta}_{i,l} = [\vartheta_{i,1,l}, \vartheta_{i,2,l}, \dots, \vartheta_{i,N_i,l}]^T$ . Tal vetor será determinado para minimizar a função custo associada a cada saída  $l$

$$E_{i,l} = \frac{1}{2N_N^{(i)}} \sum_{k=1}^{N_N^{(i)}} (\hat{y}_{i,l}(k) - y_{i,l}(k))^2, \quad l = 1, 2, \dots, m, \quad (3.9)$$

com base nos dados de treinamento. Derivar a função custo (3.9) em relação ao parâmetro consequente  $\vec{\vartheta}_{i,l}$  resulta em

$$\frac{\partial E_{i,l}}{\partial \vec{\vartheta}_{i,l}} = \frac{1}{N_N^{(i)}} \sum_{k=1}^{N_N^{(i)}} (\hat{y}_{i,l}(k) - y_{i,l}(k)) \frac{\partial \hat{y}_{i,l}(k)}{\partial \vec{\vartheta}_{i,l}}. \quad (3.10)$$

Com base na Suposição 9 e ciente de que  $\hat{C} = I_{\hat{n}}$ , chama-se a atenção para a igualdade  $\hat{y}(k) = \hat{C}x(k) = x(k)$ . Assim, uma estimativa da primeira linha de (3.4b) pode ser escrita como

$$\hat{y}(k) = \hat{A}y(k-1) + \hat{B}\hat{\delta}(k-1). \quad (3.11)$$

Considerar cada uma das derivadas para  $l = 1, 2, \dots, N_i$  em (3.10) resulta em

$$\frac{\partial \hat{y}_{i,l}(k)}{\partial \vartheta_{i,j,l}} = \frac{\partial \hat{y}_{i,l}(k)}{\partial \hat{v}_i(k-1)} \frac{\partial \hat{v}_i(k-1)}{\partial \vartheta_{i,j,l}} = \hat{B}_{i,i} \frac{\partial \hat{v}_i(k-1)}{\partial \vartheta_{i,j,l}} = \hat{B}_{i,i} \Phi_{i,j}(u_i(k-1)). \quad (3.12)$$

Por sua vez, a derivada segunda de (3.9) pode ser obtida como

$$\frac{\partial^2 E_{i,l}}{\partial \vec{\vartheta}_{i,l}^2} = \frac{1}{N_N^{(i)}} \sum_{k=1}^{N_N^{(i)}} (\hat{y}_{i,l}(k) - y_{i,l}(k)) \frac{\partial^2 \hat{y}_{i,l}}{\partial \vec{\vartheta}_{i,l}^2} + \frac{1}{N_N^{(i)}} \sum_{k=1}^{N_N^{(i)}} \left[ \frac{\partial \hat{y}_l}{\partial \vec{\vartheta}_{i,l}} \right] \left[ \frac{\partial \hat{y}_l}{\partial \vec{\vartheta}_{i,l}} \right]^T. \quad (3.13)$$

Observando em (3.12) que a derivada segunda de  $\hat{v}_i(k-1)$  em relação a  $\vec{\vartheta}_{i,l}$  é nula, pode-se inferir que

$$\frac{\partial^{n_d} E_{i,l}}{\partial \vec{\vartheta}_{i,l}^{n_d}} = 0, \quad n_d \geq 3, \quad (3.14)$$

e como consequência constata-se que a função custo é quadrática. Segundo [Jia et al. \(2005\)](#), nessas condições pode-se estabelecer

$$\frac{\partial E_{i,l}}{\partial \vec{\vartheta}_{i,l}} = \frac{\partial E_{i,l}}{\partial \vec{\vartheta}_{i,l}} \Big|_{\vec{\vartheta}_{i,l}=0} + \frac{\partial^2 E_{i,l}}{\partial \vec{\vartheta}_{i,l}^2} \Big|_{\vec{\vartheta}_{i,l}=0} \vec{\vartheta}_{i,l}, \quad (3.15)$$

que pode ser rearranjada para estimar o parâmetro consequente da seguinte maneira

$$\vec{\vartheta}_{i,l} = - \left[ \frac{\partial^2 E_{i,l}}{\partial \vec{\vartheta}_{i,l}^2} \Big|_{\vec{\vartheta}_{i,l}=0} \right]^{-1} \frac{\partial E_{i,l}}{\partial \vec{\vartheta}_{i,l}} \Big|_{\vec{\vartheta}_{i,l}=0}. \quad (3.16)$$

Isso finaliza o processo de identificação da rede **NF** para representar a relação estática não linear do modelo de Hammerstein, mesmo para os cenários nos quais as não linearidades podem conter descontinuidades. A fim de conferir clareza ao procedimento o Algoritmo 1 resume as deliberações que possibilitam a estimativa do parâmetro  $\vec{\vartheta}_{i,l}$ .

**Observação 6** Destaca-se que a metodologia pode ser aplicada para os casos multivariáveis e também monovariáveis, bastando para isso ajustar as dimensões pertinentes.

## 3.5 Exemplo Numérico

Nesta seção, as etapas da metodologia de identificação dos modelos de Hammerstein com não linearidades estáticas fortes serão aplicadas a dois exemplos numéricos multivariáveis. O primeiro exemplo tem a função de esclarecer certas particularidades da escolha dos parâmetros nos ensaios e implementação do algoritmo, por isso ele é descrito com mais detalhes. Por sua vez, o segundo exemplo é construído em um cenário mais completo que o primeiro, apresentando ruído branco de medição e uma comparação com um modelo de Hammerstein cuja parcela não linear é uma

**Algoritmo 1:** Hammerstein MIMO NF**Entrada:**  $\hat{A}, \hat{B}, \hat{C}, \bar{c}_i, \bar{\sigma}_i, u_i, y_{i,l}, N_N^{(i)}, N_i$ .**Saída:**  $\bar{\vartheta}_{i,l}$ .**início**

$$\frac{\partial E_{i,l}}{\partial \bar{\vartheta}_{i,l}} \leftarrow \frac{\partial E_{i,l}^2}{\partial \bar{\vartheta}_{i,l}^2} \leftarrow 0$$

**para**  $k \in 2:N_N^{(i)}$  **faça**

$$\hat{y}_i(k) \leftarrow \hat{A}y_i(k-1)$$

**para**  $j \in 1:N_i$  **faça**

$$\bar{\mu}_i \leftarrow \exp\left(\frac{-(u_i(k) - \bar{c}_i)^2}{\bar{\sigma}_i^2}\right)$$

$$\phi_{i,j} \leftarrow \frac{\mu_i(j)}{\sum \bar{\mu}_i}$$

$$\frac{\partial \hat{y}_{i,l}(k)}{\partial \vartheta_{i,j,l}} \leftarrow \hat{B}_{l,i} \phi_{i,j}$$

$$\frac{\partial E_{i,l}}{\partial \bar{\vartheta}_{i,l}} \leftarrow \frac{\partial E_{i,l}}{\partial \bar{\vartheta}_{i,l}} + \frac{1}{N_N^{(i)}} (\hat{y}_{i,l}(k) - y_{i,l}(k)) \frac{\partial \hat{y}_{i,l}(k)}{\partial \bar{\vartheta}_{i,l}}$$

$$\frac{\partial^2 E_{i,l}}{\partial \bar{\vartheta}_{i,l}^2} \leftarrow \frac{\partial^2 E_{i,l}}{\partial \bar{\vartheta}_{i,l}^2} + \left[ \frac{\partial \hat{y}_{i,l}(k)}{\partial \bar{\vartheta}_{i,l}} \right] \left[ \frac{\partial \hat{y}_{i,l}(k)}{\partial \bar{\vartheta}_{i,l}} \right]^T$$

$$\bar{\vartheta}_{i,l} \leftarrow - \left[ \frac{\partial^2 E_{i,l}}{\partial \bar{\vartheta}_{i,l}^2} \right]^{-1} \frac{\partial E_{i,l}}{\partial \bar{\vartheta}_{i,l}}$$

**fim**

função polinomial. Visando possibilitar a comparação entre os sinais do processo e do modelo identificado, os sistemas serão estruturados conforme o modelo de Hammerstein (3.4). Ou seja, ao invés do processo ser livre com possíveis combinações entre parcelas lineares e não lineares, ele será composto por subsistemas independentes.

**3.5.1 Exemplo 1**

O exemplo a ser investigado é um processo de Hammerstein composto por  $p = 2$  entradas e  $m = 2$  saídas. Considere que as não linearidades estáticas  $\mathcal{N}_1(\cdot)$  e  $\mathcal{N}_2(\cdot)$  pelas quais os sinais de entrada são transformados nos sinais intermediários são completamente independentes uma da outra. Matematicamente a representação dessas não linearidades são escolhidas como

$$\begin{aligned} v_1(k) &= \mathcal{N}_1(u_1(k)) = |\sin(u_1(k)) - u_1(k)/5| \\ v_2(k) &= \mathcal{N}_2(u_2(k)) = \begin{cases} u_2(k)/2, & u_2(k) \leq 4, \\ 0,06(u_2(k) - 7)^2 + u_2(k)/10, & u_2(k) > 4. \end{cases} \end{aligned} \quad (3.17)$$

Considere também que, na sequência, o sinal intermediário é inserido em um subsistema dinâmico linear retratado pelo sistema de segunda ordem ( $n = 2$ ) em espaço de estados constituído pelas matrizes

$$A = \begin{bmatrix} 0,4 & 0,1 \\ 0,3 & 0,6 \end{bmatrix}, \quad B = \begin{bmatrix} 0,6 & 0,4 \\ 0,1 & 0,3 \end{bmatrix} \quad \text{e} \quad C = \begin{bmatrix} 0,6 & 0,1 \\ 0,4 & 0,2 \end{bmatrix}. \quad (3.18)$$

Conforme a metodologia desenvolvida, duas etapas devem ser implementadas a fim de identificar o modelo de Hammerstein. A primeira etapa consiste em obter uma estimativa de um modelo em espaço de estados a partir de dados de entrada e saída. Como apontado na Seção 3.4, esse objetivo é factível quando o sinal de entrada para obter os dados do processo é um sinal PRBS. Nessas circunstâncias, um PRBS é projetado para ser aplicado no processo e gerar os respectivos dados de saída, úteis para o MOESP. Observe na Figura 3.9a), em que se apresenta somente a não linearidade  $\mathcal{N}_1$ , que essa não linearidade estática possui descontinuidade exatamente no ponto de restrição em magenta. A restrição a que o destaque se refere tem fundamento na Suposição 11. A ideia é que se a entrada não nula que deveria retornar uma saída não nula, for na verdade nula, o sinal intermediário também é nulo e não levará informações dinâmicas do processo para que o MOESP possa extrair. Os demais pontos não têm restrição, por isso uma escolha arbitrária foi feita do ponto destacado em azul na Figura 3.9a) para a aplicação do nível não nulo do PRBS. Os sinais de entrada, intermediário e de saída são apresentados na Figura 3.9b). Notam-se variações da saída mesmo em instantes em que não há mudança no sinal de PRBS de entrada  $u_1(k)$ . O motivo é que nesses instantes há mudança na entrada  $u_2(k)$  e o acoplamento da dinâmica linear faz com que todas as saídas sejam afetadas pelas variações de qualquer entrada.

O sinal PRBS é uma sequência que simula aleatoriedade, no entanto, é periódica com período máximo  $2^{n_p} - 1$ , por isso é nomeada de pseudoaleatória. Como tal sequência varia somente entre dois níveis, ela é nomeada de sequência binária. Por fim, a variação entre os dois níveis respeita um tempo mínimo. Assim o sinal permanece em um mesmo nível por tempo igual ou superior ao mínimo e depois a troca de nível é feita de maneira pseudoaleatória. Esse tempo mínimo é importante para que o processo possa perceber corretamente a entrada. Por exemplo, variações rápidas podem ser imperceptíveis, pois quando o processo inicia a resposta a entrada se altera, por outro lado, variações muito lentas podem ser enxergadas como degraus, como pouca informação dinâmica. Esse problema é mais complexo para sistemas multivariáveis. Como nas implementações apresentadas ao longo deste trabalho n

A sua implementação prática pode ser feita por um registrador com  $n_p \in \mathbb{N} > 3$  estágios (Katayama, 2005). De modo que, quanto mais estágios, maior o tempo que a sequência pode ser coletada sem haver repetição. O projeto do sinal PRBS envolve, portanto, a escolha do nível, o período máximo e a quantidade mínima de iterações que o mesmo nível permanece inalterado.

Um dos direcionamentos para identificação de sistemas dinâmicos lineares é

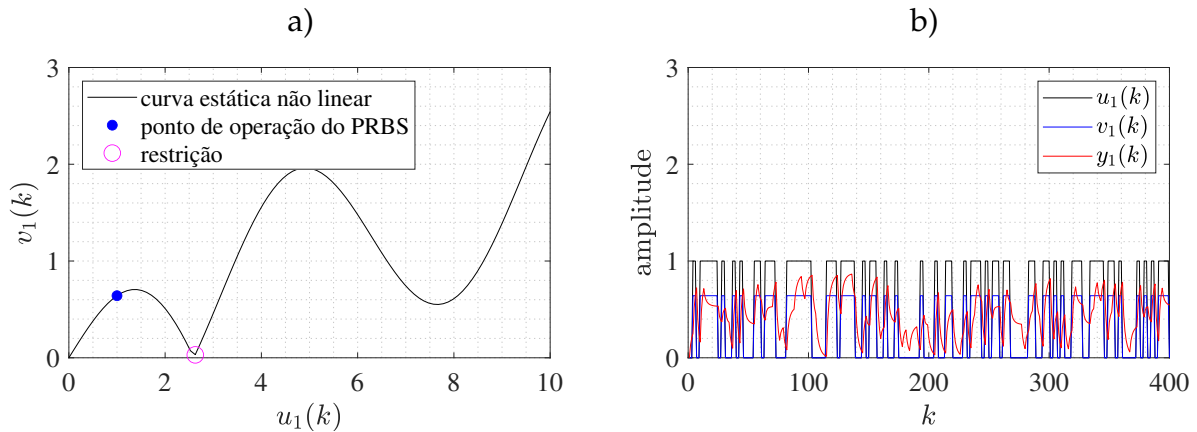


Figura 3.9: Sinais relacionados ao subíndice 1 do ensaio dinâmico realizado no processo do Exemplo 1 para implementar a primeira etapa da metodologia de identificação do modelo de Hammerstein **MIMO NF**. No gráfico a) apresenta-se a curva estática não linear  $\mathcal{N}_1$  do processo junto a duas marcações. ( $\circ$ ) indica onde não se tem interesse em excitar o processo, uma vez que o sinal intermediário é praticamente anulado nessa localidade. ( $\bullet$ ) indica o ponto no qual o sinal **PRBS** de fato foi excitado. Por sua vez, no gráfico b) pode-se consultar: os sinais do ensaio dinâmico a serem utilizados na primeira etapa da identificação; e o sinal intermediário cuja amplitude está relacionada ao ponto azul do gráfico a).

que o sinal **PRBS** deve permitir que a resposta atinja cerca de 70% do seu valor em regime permanente (Aguirre, 2015). Esse direcionamento pode ser seguido para o Exemplo 1 ao escolher o tempo mínimo que o nível fica fixo como 3 iterações. Quanto ao período máximo, é interessante que ele seja maior que a quantidade de amostras com as quais a identificação será feita. Como serão empregadas 400 amostras (ver Figura 3.9) é desejável que o período máximo se aproxime desse valor. Testar as possibilidades de valores de  $n_p$  em ordem crescente, ou seja,  $n_p = 4, 5, \dots$  leva a  $2^{n_p} - 1 = 2^7 - 1 = 127$ .

Uma discussão foi feita nos últimos dois parágrafos em relação ao nível e como cada nível fica fixo por pelo menos 3 iterações, o período máximo também é expandido em 3 vezes de modo que ele se torna  $3 \cdot 127 = 381$ , ou seja, 19 amostras são repetidas. Salienta-se que a escolha dos níveis foram realizadas conforme as Suposições 10 a 12.

Os sinais de entrada e saída apresentados na Figura 3.9b) constituem o ensaio dinâmico da primeira etapa. A finalização desta etapa é consumada pela execução do algoritmo **MOESP**, que requer os dados de entrada e saída do processo a ser identificado, além de um parâmetro que controla a quantidade de blocos linha das matrizes de Hankel. A variável responsável por isso é denotada  $h$ . O valor 7 foi escolhido para a variável  $h$  com base nos resultados da identificação. Ou seja, após identificar o modelo completo de Hammerstein **MIMO NF** e fazer a validação é possível verificar a melhor escolha desse parâmetro por meio do **RMSE**. No exemplo numérico, como se conhece completamente o sistema dinâmico linear, é possível se guiar também pelos autovalores da matriz  $\tilde{A}$  estimada.

As matrizes obtidas como resultado do **MOESP** foram

$$\tilde{A} = \begin{bmatrix} 0,6324 & 0,1378 \\ 0,1631 & 0,3676 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} -0,1327 & -0,0847 \\ 0,0830 & 0,0195 \end{bmatrix} \quad \text{e} \quad \tilde{C} = \begin{bmatrix} -1,4800 & 0,4940 \\ -1,3232 & -0,1057 \end{bmatrix}, \quad (3.19)$$

sendo os autovalores da matriz  $\tilde{A}$  iguais aos da matriz  $A$  apresentada em (3.18), 0,7 e 0,3. Para executar os passos que serão detalhados a seguir, a matriz  $\tilde{C}$  precisa se tornar uma matriz identidade de modo que se aplica a transformação apresentada em (3.5), cujo resultado é o modelo em espaço de estados que caracteriza a parcela dinâmica linear do modelo de Hammerstein **MIMO NF**

$$\hat{A} = \begin{bmatrix} 0,0750 & 0,5625 \\ -0,2500 & 0,9250 \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} 0,2373 & 0,1350 \\ 0,1668 & 0,1100 \end{bmatrix} \quad \text{e} \quad \hat{C} = I_2. \quad (3.20)$$

Finalizada a primeira etapa, direciona-se a atenção para a segunda etapa a fim de estimar o subsistema estático não linear. O ensaio completo realizado sobre o processo é composto por  $p + 1$  faixas, ou seja, 3 faixas, uma vez que  $p = 2$  no atual Exemplo 1. A primeira corresponde ao sinal **PRBS** e as outras duas são sinais multiníveis aplicados em cada entrada para fazer a identificação da parcela estática não linear. Nessa segunda etapa, busca-se captar o comportamento estático não linear do processo; por isso, é importante excursionar os sinais de entrada pela faixa mais ampla possível. Como na metodologia apresentada em (Jia et al., 2005), os sinais utilizados são os sinais multiníveis, que estão apresentados na Figura 3.10a). A cada iteração a amplitude é uma amostra diferente de uma distribuição uniforme definida como

$$\mathcal{U}(\mu_u, \sigma_u), \quad (3.21)$$

em que  $\mu_u$  é a média e  $\sigma_u$  é o desvio padrão. Para gerar o sinal da Figura 3.10a) a distribuição é caracterizada por  $\mathcal{U}(5,10/\sqrt{12})$ . No total foram geradas 800 amostras para cada entrada, sendo 400 para cada uma das duas faixas na Figura 3.10a) e as primeiras 400 amostras descartadas. Os sinais correspondentes de saída são apresentados na Figura 3.10b).

Os dados que completam os pré-requisitos para aplicar a metodologia **NF** de identificação da curva não linear são os parâmetros  $S_0$ ,  $\rho$  e  $\lambda$ . As escolhas dessas variáveis foram feitas segundo o valor final de **RMSE** da validação do modelo de Hammerstein **MIMO**. Os valores indicados foram  $S_{1,0} = 0,98$ ,  $S_{2,0} = 0,989$ ,  $\rho_1 = 1$ ,  $\rho_2 = 1$ ,  $\lambda_1 = 0,1$  e  $\lambda_2 = 0,1$ . A partir desses parâmetros e dos sinais de entrada e saída da Figura 3.10 o algoritmo **NF** apresentado na Subseção 3.4.2 foi implementado, resultando em  $N_1 = 36$  clusters para a primeira entrada e  $N_2 = 49$  clusters para a segunda entrada.

Como resultado de aplicar um sinal não nulo somente em  $u_1(k)$  o único sinal intermediário não nulo é  $v_1(k)$ . No entanto, todas as  $m$  saídas respondem ao estímulo. Por isso podem ser utilizadas  $p \cdot m$  combinações de entrada e saída para estimar as curvas estáticas não lineares. Tendo em conta tais combinações, a implementação do

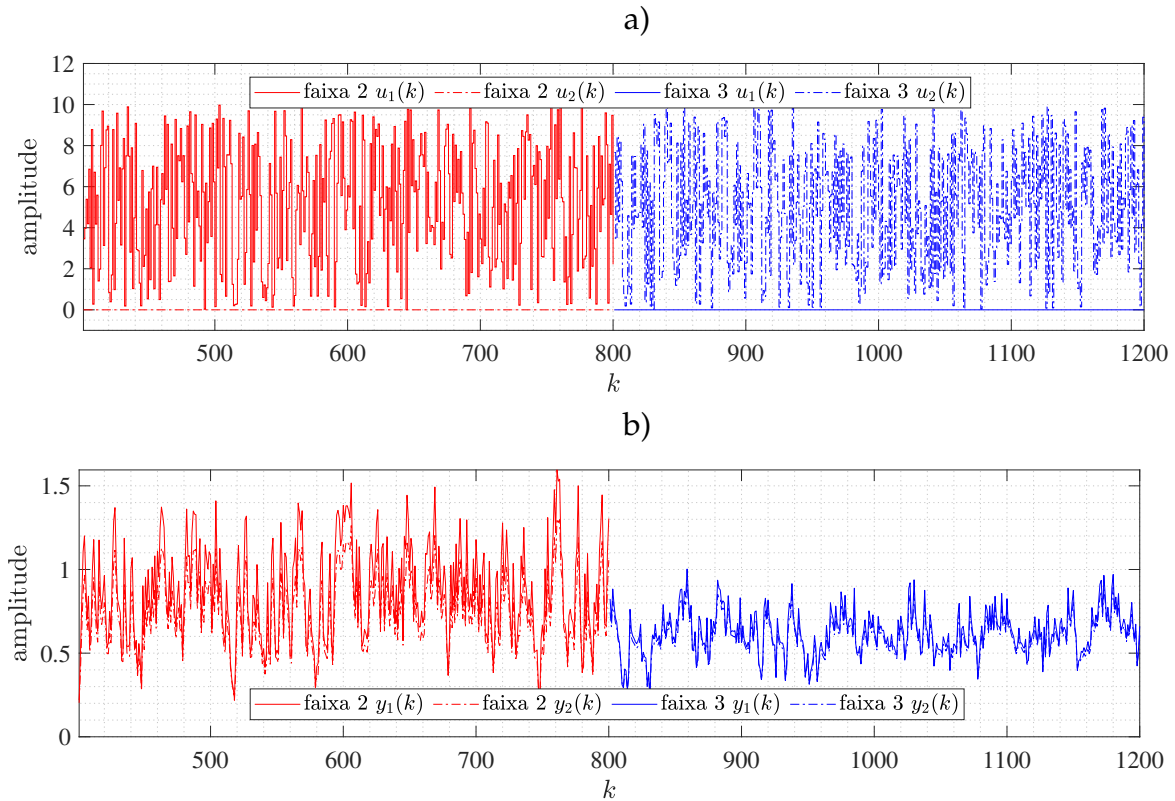


Figura 3.10: Sinais empregados para a identificação das curvas estáticas não lineares do Exemplo 1. Em a) apresentam-se os sinais multinível e em b) as respectivas respostas. Na primeira faixa que está à esquerda somente a entrada  $u_1(k)$  é não nula, enquanto na segunda faixa que está à direita apenas a entrada  $u_2(k)$  é não nula.

Algoritmo 1 resultou em 4 estimativas de parâmetros consequentes  $\hat{\vartheta}$ , sendo 2 relacionados à primeira entrada e consequentemente à  $\mathcal{N}_1(\cdot)$  e outros 2 relacionados à segunda entrada e a  $\mathcal{N}_2(\cdot)$ . Ou seja, os parâmetros estimados  $\hat{\vartheta}_{1,1}$  e  $\hat{\vartheta}_{1,2}$  geram estimativas da mesma não linearidade  $\mathcal{N}_1(\cdot)$  a partir do conjunto de dados  $\{u_1(k), y_1(k)\}_{k=1}^{N_N^{(1)}}$  e  $\{u_1(k), y_2(k)\}_{k=1}^{N_N^{(1)}}$ , respectivamente. O mesmo efeito ocorre para a segunda não linearidade a ser estimada. Desses resultados, seleciona-se um para cada entrada conforme o valor normalizado do **RMSE** de validação do modelo de Hammerstein que eles geram.

No exemplo desta subseção, os menores valores de **RMSE** normalizados foram 0,0144 e 0,0127 resultantes do emprego de  $\hat{\vartheta}_{1,1}$  para a primeira entrada e  $-\hat{\vartheta}_{2,2}$  para a segunda entrada. As curvas estáticas não lineares identificadas são apresentadas na Figura 3.11. Para exibir as curvas, o sinal de entrada é variado linearmente em degraus ascendentes por toda a faixa de excursão permitida, ou seja, de 0 a 10 e, em seguida, aplicado tanto no processo como no modelo. Salienta-se que o degrau é mantido por tempo suficiente para que os efeitos dinâmicos não mais interfiram e seja possível coletar o valor em regime permanente. Mostra-se na Figura 3.11 os valores em regime permanente da saída, uma vez que, na prática, não há acesso ao sinal intermediário. Devido ao valor em regime as variáveis de saída são denotadas com a barra sobre elas.

Esse foi o resultado estático do modelo de Hammerstein **MIMO NF**. Para validação

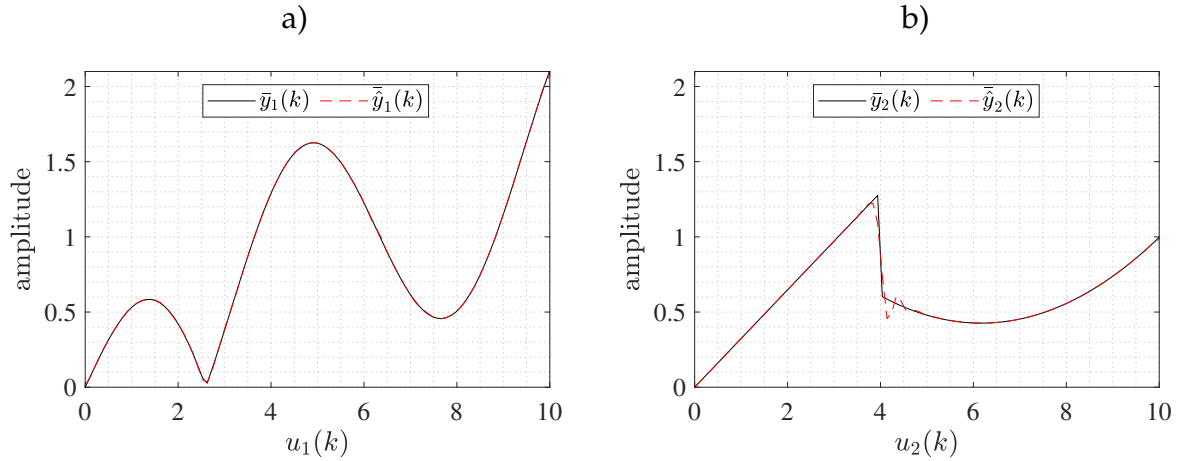


Figura 3.11: Validação da característica estática do Exemplo 1. O modelo de Hammerstein estimado por meio da metodologia **NF MIMO** é comparado ao processo por meio de dados estáticos da saída. Para realizar o ensaio foram empregados degraus ascendentes variando a amplitude de 0 a 10. Em a) apresenta-se a validação da primeira curva estática  $\hat{N}_1$  e em b) da segunda curva estática  $\hat{N}_2$ .

completa, incluindo a dinâmica, um sinal multinível diferente daquele utilizado para identificação é empregado. Assim realiza-se um novo ensaio no processo, cujos sinais de entrada são sinais multiníveis aplicados simultaneamente às entradas. Ou seja, são dois sinais multiníveis diferentes sendo aplicados ao mesmo tempo nas entradas do processo e do modelo. Os resultados são comparados em um mesmo gráfico na Figura 3.12.

### 3.5.2 Exemplo 2

Como estabelecido no Exemplo 1, considere também para o Exemplo 2 um processo de Hammerstein composto por  $p = 2$  entradas e  $m = 2$  saídas e que as não linearidades estáticas  $\mathcal{N}_1(\cdot)$  e  $\mathcal{N}_2(\cdot)$ , ambas com entradas limitadas entre 0 e 5, são independentes entre si. A representação matemática para as duas curvas não lineares são

$$v_1(k) = \mathcal{N}_1(u_1(k)) = \begin{cases} 2 - \cos(3u_1(k)) - \exp(-u_1(k)), & u_1(k) \leq 3, \\ 3, & u_1(k) > 3, \end{cases} \quad (3.22)$$

$$v_2(k) = \mathcal{N}_2(u_2(k)) = \begin{cases} \tanh(2u_2(k)), & u_2(k) \leq 1,5, \\ \frac{\exp(u_2(k))-1}{\exp(u_2(k))+1}, & u_2(k) > 1,5. \end{cases}$$

Ambas não linearidades foram utilizadas separadamente em exemplos em (Jia et al., 2005) e (Jia et al., 2016) e, aqui, serão identificadas em conjunto na estrutura de um modelo de Hammerstein **MIMO**. Em seguida, os sinais intermediários que carregam a informação não linear estática são empregados como entrada da dinâmica linear representada pelo modelo em espaço de estados de segunda ordem descrito pelas

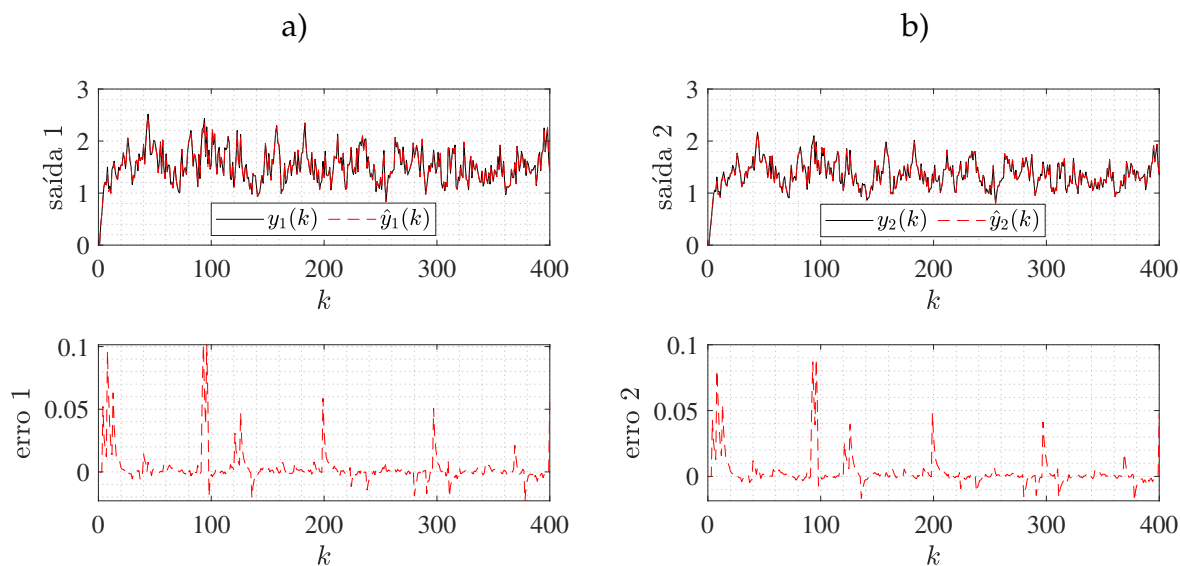


Figura 3.12: Validação completa do modelo de Hammerstein **MIMO NF** identificado no Exemplo 1. Os gráficos a) e b) apresentam o resultado da validação sobrepondo a comparação entre a saída do processo e do modelo. Os valores de **RMSE** obtidos foram de 0,0144 e 0,0127 para as saídas 1 e 2, respectivamente.

matrizes

$$A = \begin{bmatrix} 0,9608 & -0,0395 \\ 0,0607 & 0,9048 \end{bmatrix}, \quad B = \begin{bmatrix} 0,0699 & 0,1 \\ 0,1 & -0,1748 \end{bmatrix}, \quad \text{e} \quad C = I_2. \quad (3.23)$$

Diferente do desenvolvimento feito no Exemplo 1, o sistema apresentado em (3.23) será afetado por ruído branco de medição, caracterizado por uma distribuição normal definida como

$$\mathcal{N}(\mu_n, \sigma_n), \quad (3.24)$$

sendo  $\mu_n$  a média e  $\sigma_n$  o desvio padrão. O **MOESP** pode estimar as matrizes em espaço de estados mesmo com a presença desse tipo de ruído, de modo que as matrizes a serem aplicadas no Algoritmo 1 continuam sendo esperadas como representativas, tal como eram no caso sem ruído.

Conforme a metodologia desenvolvida, são necessárias ( $p+1 = 2+1 = 3$ ) faixas distintas do sinal de ensaio, das quais a primeira é constituída de um sinal **PRBS** e as subsequentes são sinais multiníveis que excitam apenas uma das entradas em cada faixa. Essas características gerais podem ser vistas com clareza na Figura 3.13. Na faixa 1, encontra-se o sinal **PRBS** a ser utilizado para a identificar o modelo dinâmico linear. A amplitude não nula do sinal foi escolhida como 1 para ambas entradas, excitando pontos que não anulam os sinais intermediários e conseqüentemente os de saída, ou seja, respeitando as Suposições 10 a 12. Além disso, o intervalo mínimo no qual um patamar permanece fixo foi definido como 3 iterações, de modo que a resposta pudesse atingir aproximadamente 70% do seu valor em regime permanente.

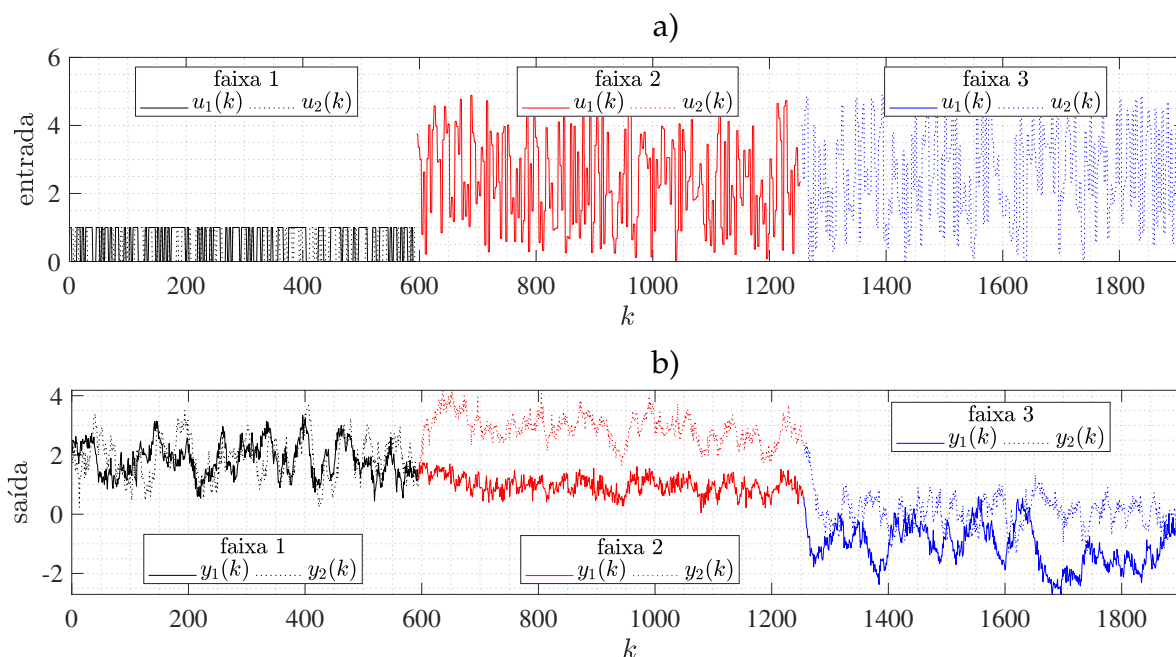


Figura 3.13: Sinais da primeira realização utilizados para identificação no Exemplo 2 do Capítulo 3. Em a), o sinal PRBS está na primeira faixa e os sinais multinível aplicados a uma entrada de cada vez estão nas duas faixas seguintes, 2 e 3. Em b), estão as respectivas respostas.

Enquanto para a identificação da dinâmica linear é mais interessante projetar um sinal de teste que permaneça em torno de um ponto de operação, para identificar a não linearidade estática é importante excursionar o sinal de teste por faixas amplas, a fim de captar os efeitos não lineares. Dito isso, permitiu-se que o sinal multinível excursionasse por uma faixa de valores entre 0 e 5, em busca de captar os efeitos não lineares. As amplitudes do sinal foram geradas seguindo uma distribuição uniforme com  $\mu_u = 5/2$  e  $\sigma_u = 5/\sqrt{12}$ , ou seja,  $u(5/2, 5/\sqrt{12})$ .

O ruído de medição adicionado às saídas simuladas foi gerado por meio de uma distribuição normal caracterizada por  $\mu_n = 0$  e  $\sigma_n = 1/6$ ,  $\mathcal{N}(0, 1/6)$  que resulta em uma relação sinal ruído (SNR, do inglês *Signal Noise Ratio*) de aproximadamente 21 dB para ambas saídas. Com a intenção de verificar os efeitos dessa adição de ruído na identificação, foram tomadas 1000 realizações de todo o processo, gerando 1000 modelos diferentes.

O conjunto de dados de entrada e saída exibidos na faixa 1 das Figuras 3.13 a) e b) foram aplicados no algoritmo MOESP descrito com mais detalhes na Subseção 2.5.1. Como parâmetro de projeto escolheu-se a quantidade de blocos linha da matriz de Hankel como  $h = 7$ . Uma vez que foram tomadas 1000 realizações de ruído de medição, as matrizes estimadas são ligeiramente diferentes em cada execução, de modo que se apresentam somente as matrizes resultantes da primeira realização. Ademais, o resultado completo será apresentado na forma gráfica.

A matriz  $\tilde{C}$  estimada pelo **MOESP** na primeira realização foi

$$\tilde{C} = \begin{bmatrix} -2,6252 & 1,5810 \\ -3,0241 & -1,8912 \end{bmatrix}. \quad (3.25)$$

Aplicar a transformação de similaridade (3.5) resultou no seguinte modelo em espaço de estados

$$\hat{A} = \begin{bmatrix} 0,9624 & -0,0409 \\ 0,0602 & 0,9023 \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} 0,1819 & 0,0977 \\ 0,2628 & -0,1793 \end{bmatrix} \quad \text{e} \quad \hat{C} = I_2, \quad (3.26)$$

que tem autovalores  $0,9323 \pm 0,0396i$ , próximos ao da matriz  $A$  apresentada em (3.23) que são  $0,9328 \pm 0,0402i$ , sendo  $i$  a notação da parte imaginária, ou seja,  $i = \sqrt{-1}$ .

Como parâmetros de projeto da identificação dos blocos estáticos não lineares pela metodologia **NF** foram selecionados  $S_{1,0} = 0,98$ ,  $S_{2,0} = 0,998$ ,  $\rho_1 = 1$ ,  $\rho_2 = 1,5$ ,  $\lambda_1 = 0,1$  e  $\lambda_2 = 0,1$ . A partir desses parâmetros e dos  $N_N^{(1)} + N_N^{(2)} = 1300$  dados de entrada e saída exibidos nas faixas 2 e 3 da Figura 3.13 foram obtidos  $N_1 = 16$  e  $N_2 = 43$  *clusters* em todas as realizações. Note que mesmo com a adição de ruído, a quantidade de *clusters* formados não foi alterada entre as realizações. Isso não é verdade sempre e torna-se menos provável à medida que se aumenta a variância do ruído, entretanto a variabilidade da quantidade de *clusters* não é um problema. O problema é a presença do ruído que pode interferir na acurácia dos modelos obtidos. O tratamento aplicado no contexto desta dissertação consiste em empregar o **MOESP** que pode tratar ruído branco de medição.

O Algoritmo 1 foi implementado para cada realização ao considerar todas as combinações de entradas e saídas, o que resultou em duas possibilidades de parâmetros consequentes para a primeira entrada  $u_1$ , sendo elas  $\bar{\vartheta}_{1,1}$  e  $\bar{\vartheta}_{1,2}$  e duas para a segunda entrada  $u_2$ , as quais são  $\bar{\vartheta}_{2,1}$  e  $\bar{\vartheta}_{2,2}$ . Segundo os índices **RMSE** obtidos dos dados de validação dinâmica, a melhor combinação foi entre  $\bar{\vartheta}_{1,1}$  cuja curva estática resultante está apresentada na Figura 3.14a) e  $\bar{\vartheta}_{2,2}$  que origina a curva estática da Figura 3.14b).

Destaca-se que ambas curvas obtidas com a metodologia **NF** proposta possuem desempenho superior ao caso polinomial, para um polinômio de ordem 10. Como citado anteriormente, as curvas estáticas do processo foram retiradas de (Jia et al., 2005). Nesse artigo os autores comparam as curvas obtidas por sua metodologia **NF SISO** com polinômios aproximadores de ordem 10. Esse valor é utilizado devido a tal ordem retornar o melhor modelo polinomial em termos quantitativos.

Na Figura 3.15 apresenta-se uma comparação entre a metodologia com a função polinomial e o sistema **NF**. O ensaio consiste em aumentar a quantidade de dados  $N_N^{(i)}$  utilizados para identificação da curva estática não linear  $N_i$  nas duas metodologia e verificar qual delas tem desempenho superior em termos de **RMSE**. Pelas retas ajustadas, nota-se que os índices **RMSE** obtidos para o modelo de Hammerstein **NF** estão abaixo dos índices do modelo polinomial em todas as faixas testadas.

A validação por simulação livre do modelo identificado pode ser observada

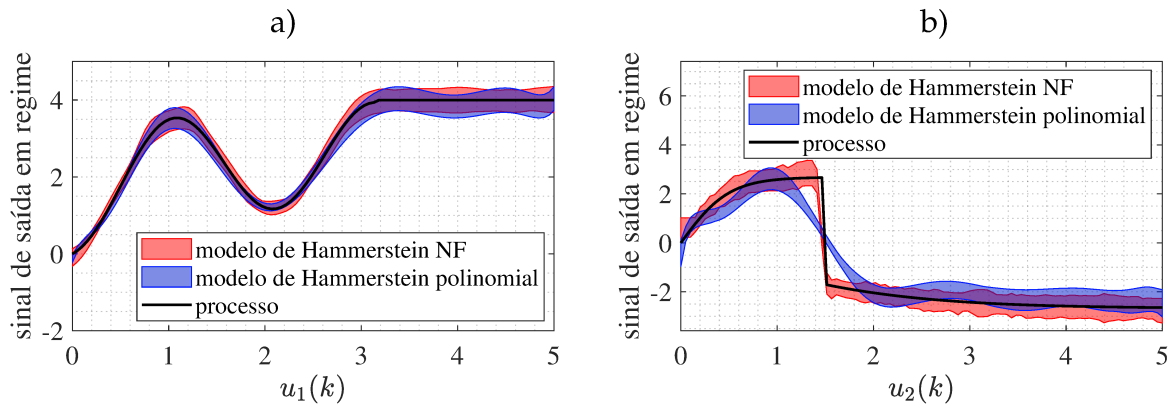


Figura 3.14: Validação da nuvem de curvas estáticas da entrada  $u_1$  em a) e  $u_2$  em b). *A priori* não se conhece o sinal intermediário, à vista disso, considera-se como dados formadores da curva estática não linear a relação em regime permanente entre a entrada e a saída em diversos pontos de operação.

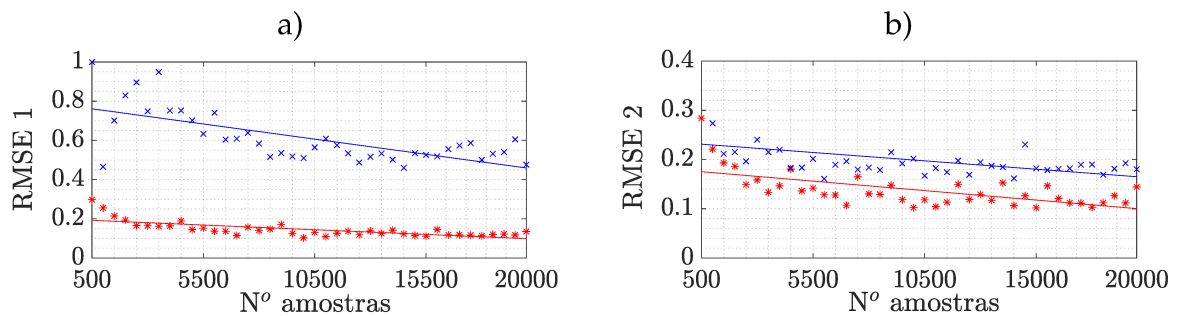


Figura 3.15: Relação entre o aumento de número de amostras na identificação da curva estática e o RMSE da validação dinâmica. Modelo Hammerstein NF (\*) e polinomial (x).

na Figura 3.16. Para produzi-la, foi gerada uma nova realização de sinal multinível conforme a distribuição uniforme  $\mathcal{U}(5/2, 5/\sqrt{12})$  com 400 amostras para cada entrada. As saídas do modelo de menor RMSE produzidas ao utilizar o sinal multinível como entrada são apresentadas nas Figuras 3.16 a) e b), bem como as diferenças entre a saída do processo e dos modelos em c) e d). O índice RMSE da primeira saída foi de 0,0943 para o caso de Hammerstein MIMO NF e 0,3660 para o caso polinomial. Por sua vez, os valores do RMSE da segunda saída foram de 0,11 e 0,1983 para os modelos Hammerstein MIMO NF e polinomial, respectivamente.

## 3.6 Conclusão

Neste capítulo, uma nova abordagem para identificar um modelo de Hammerstein MIMO baseado em sistemas NF e técnicas de subespaço foi apresentada. A não linearidade estática foi representada por uma rede NF com regras Takagi-Sugeno de

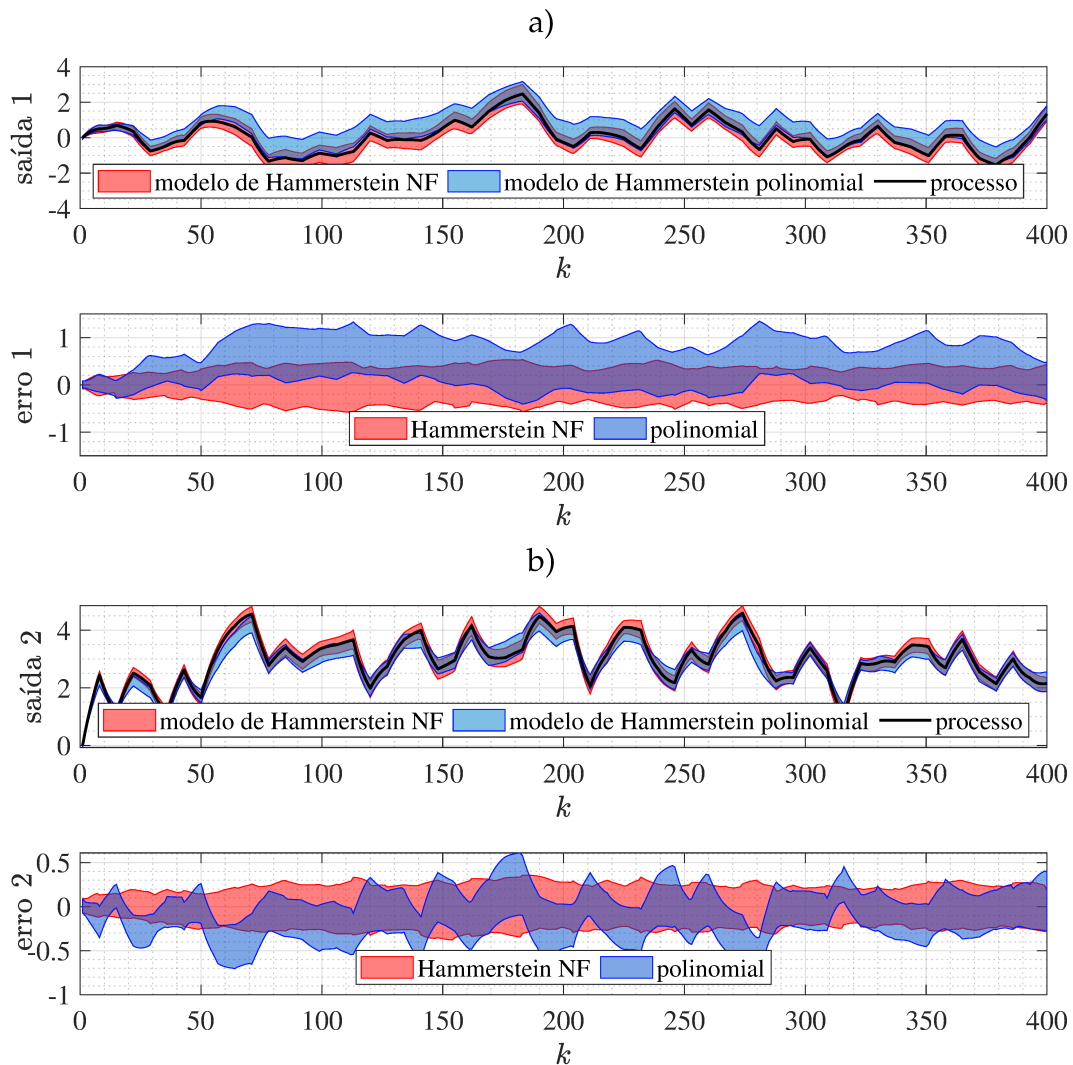


Figura 3.16: Validação dinâmica por simulação livre, partindo de condições iniciais nulas em 1000 realizações. Em a) apresenta-se a comparação da primeira saída do modelo de Hammerstein (o NF e o polinomial) com o processo e logo abaixo apresenta-se a diferença entre eles. Em b) comparam-se os sinais da segunda saída, na parte superior está a dinâmica no tempo e na parte inferior o erro de estimativa.

ordem zero e a dinâmica linear por um modelo MIMO em espaço de estados. A rede de 4 camadas mostrou-se capaz de captar e aproximar curvas com não linearidades fortes a partir de dados de entrada e saída (Jia et al., 2005). A dinâmica linear foi identificada utilizando o MOESP que é um algoritmo de subespaços. Dadas essas escolhas, a identificação herdou certas vantagens da representação em espaço de estados. Cita-se, por exemplo, a facilidade para extensão ao caso MIMO, a possibilidade de realizar o projeto dos controladores diretamente em espaço de estados e a dispensa da delicada etapa de escolha de estrutura de modelos como ocorre com o ARX em (Jia et al., 2016). O algoritmo desenvolvido foi aplicado em dois exemplos simulados com quantidade de entradas iguais às de saídas, nos quais se constatou a eficácia da abordagem proposta para identificar modelos MIMO de Hammerstein com não linearidades fortes, mesmo

---

na presença de ruído de medição branco.

# Identificação de Histerese **MIMO** Empregando **MPSO** e Métodos de Subespaço

---

## 4.1 Introdução

Uma particularidade que caracteriza a presença da histerese em um sistema é a persistência de um laço quando se traça um gráfico entre os dados de entrada e saída. Esse laço é formado à medida que o sinal de entrada tem sua frequência aproximada de zero. O sinal deve ser caracterizado por um período de carga (monotonicamente crescente) e outro de descarga (monotonicamente decrescente), como, por exemplo, uma onda senoidal ou triangular. Isso motiva o emprego do termo não linearidade quase estática, uma vez que a histerese é um fenômeno que depende das entradas passadas e, portanto, tem memória, mas se manifesta em frequências próximas de zero (Noël et al., 2017).

Neste capítulo o modelo de Hammerstein será utilizado para descrever esse fenômeno não linear que é a histerese e também a dinâmica linear do ponto de vista dos dados de entrada e saída. A parcela quase estática não linear do modelo de Hammerstein será representada por um modelo **GRDPI** enquanto a parcela dinâmica **LIT** será representada em espaço de estados.

Apesar de haver trabalhos que apresentam maneiras de modelar processos **SISO** com a presença de histerese, como (Aljanaideh et al., 2018, 2019, 2022) e (Al Janaideh et al., 2023), poucos artigos desenvolveram o caso **MIMO**, como foi feito em (Rakotondrabe, 2017). No entanto, mesmo em (Rakotondrabe, 2017) a extensão multivariável é feita somente para a não linearidade, ou seja, não é voltada para os **MBIs**. Nesse contexto, propõe-se neste capítulo uma metodologia em duas etapas para identificar sistemas **MIMO** com histerese por meio dos modelos de Hammerstein.

Para esse desenvolvimento, na Seção 4.2, apresenta-se a extensão ao caso **MIMO** para a representação do modelo **GRDPI**, introduzido na Subseção 2.4.3. Com relação à parcela dinâmica linear uma extensão direta dos casos existentes na literatura não é interessante, uma vez que focam em modelos de difícil abordagem para sistemas

**MIMO**, como é o modelo autorregressivo com entradas exógenas. Uma representação adequada para a parte linear multivariável é o modelo em espaço de estados (Verhaegen e Dewilde, 1992a). Dessa forma, essa representação será utilizada para modelar a parcela dinâmica linear.

Neste capítulo, o problema de pesquisa é anunciado na Seção 4.3. As representações do modelo de Hammerstein são definidas como um modelo GRDPI seguido de um modelo em espaço de estados, ambos orientados para o cenário MIMO. Na Seção 4.4 descreve-se o procedimento completo de identificação, incluindo os algoritmos para implementar uma versão modificada da otimização por enxame de partículas (MPSO, do inglês *Modified Particle Swarm Optimization*) e o MOESP, usados para identificar o modelo GRDPI e em espaço de estados, respectivamente. Por fim, a Seção 4.5 é dedicada a dois exemplos numéricos nos quais se verifica a eficácia da metodologia proposta e, a conclusão do capítulo é feita na Seção 4.6.

## 4.2 Modelo GRDPI Multivariável

A parcela não linear do modelo de Hammerstein MIMO foi representada na Figura 2.2 de duas maneiras distintas. Na primeira, mostrada no diagrama a), cada entrada é transformada em um sinal intermediário por uma única curva não linear, ou seja, as funções  $\mathcal{N}_i(u_i(k))$  são completamente independentes entre si. De outro modo, na segunda maneira, exibida no diagrama b), todas as entradas são processadas por cada uma das não linearidades, dando origem ao sinal intermediário  $v_i(k) = \mathcal{N}_i(u_1(k), u_2(k), \dots, u_p(k))$ . Nesse segundo caso, quando somente uma entrada é excitada, todos os sinais intermediários podem corresponder ao estímulo. Portanto, apesar da separação no diagrama entre as funções não lineares, os sinais podem ser cruzados, ou seja, a entrada  $u_i(k)$  pode afetar até mesmo um sinal intermediário  $v_{\neq i}(k)$ , caracterizando-a como mais geral que a primeira maneira. A notação  $\neq i$  no subíndice expressa qualquer sinal intermediário com o índice diferente de  $i$ .

Inteirado dessas nomenclaturas, considere a seguinte definição acerca da histerese multivariável com  $p$  entradas e  $q$  saídas intermediárias.

**Definição 4.2.1** (Rakotondrabe, 2017) *Uma histerese multivariável pode ser representada por um vetor com não linearidades  $\mathcal{N}_i$  — sendo pelo menos uma delas uma histerese — com  $p$  entradas  $u_i, i = 1, \dots, p$  e  $q$  saídas intermediárias  $v_i, i = 1, \dots, q$*

$$\vec{v} = \begin{bmatrix} v_1(k) \\ \vdots \\ v_i(k) \\ \vdots \\ v_q(k) \end{bmatrix} = \begin{bmatrix} \mathcal{N}_1[u_1, \dots, u_i, \dots, u_p](k) \\ \vdots \\ \mathcal{N}_i[u_1, \dots, u_i, \dots, u_p](k) \\ \vdots \\ \mathcal{N}_q[u_1, \dots, u_i, \dots, u_p](k) \end{bmatrix}. \quad (4.1)$$

Foi destacado por Rakotondrabe (2017) que a estrutura do modelo e, portanto,

o procedimento de identificação podem diferir se o número  $p$  de entradas é ou não igual ao número  $\varrho$  de saídas intermediárias em (4.1). Desse modo, nomenclaturas definidas em termos das quantidades de entradas e saídas podem ser úteis na sequência. Quando  $p = \varrho$ , a histerese multivariável é quadrada. Em contrapartida, quando  $p \neq \varrho$ , a histerese multivariável é retangular. À vista disso, para um sistema com histerese retangular, diz-se que ele é subatuado se  $p < \varrho$  e superatuado se  $p > \varrho$ . De outro modo, um sistema quadrado com histerese, ou seja,  $p = \varrho$ , é chamado de totalmente atuado.

Se a histerese apresentada em cada saída intermediária for interpretada como uma soma dos efeitos individuais, o modelo não linear multivariável na Definição 4.2.1 pode ser escrito como uma soma dos efeitos de cada não linearidade como

$$\vec{v} = \begin{bmatrix} v_1(k) \\ \vdots \\ v_l(k) \\ \vdots \\ v_\varrho(k) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^p \mathcal{N}_{1,i}[u_i](k) \\ \vdots \\ \sum_{i=1}^p \mathcal{N}_{l,i}[u_i](k) \\ \vdots \\ \sum_{i=1}^p \mathcal{N}_{\varrho,i}[u_i](k) \end{bmatrix}, \quad (4.2)$$

em que cada elemento  $\mathcal{N}_{l,i}[u_i](k)$  é uma não linearidade do tipo histerese ou um ganho linear, sendo necessariamente pelo menos um deles uma histerese. Com essa condição, o operador é, portanto, chamado de operador com histerese multivariável. Em diversos sistemas que utilizam atuadores piezoelétricos ou magnéticos, geralmente toda a diagonal é do tipo histerese (Habineza et al., 2015). A diagonal principal é caracterizada pelos operadores de transferência direta e os elementos fora da diagonal principal são os operadores de transferência cruzada. Se todos os acoplamentos cruzados são nulos, então a histerese multivariável se resume a várias histereses monovariáveis independentes, como foi apresentado na representação KU, que é um cenário mais básico e envolve o escopo desta dissertação.

Para compreender o caso multivariável a ser empregado neste texto, considere que a saída de cada um dos modelos GRDPI que compõe o modelo multivariável é dada por

$$v_i(k) = \mathcal{N}_i(\mathcal{B}_i[u_i](k)), \quad i = 1, \dots, p, \quad (4.3)$$

em que  $\mathcal{B}_i$  é um conjunto que descreve as duas funções generalizadas de borda, a esquerda e a direita. Conceitualmente isso significa que as histereses tratadas são independentes, de modo que as saídas podem ser organizadas vetorialmente da seguinte maneira

$$\begin{bmatrix} v_1(k) \\ \vdots \\ v_i(k) \\ \vdots \\ v_p(k) \end{bmatrix} = \begin{bmatrix} \mathcal{N}_1(\mathcal{B}_1[u_1](k)) \\ \vdots \\ \mathcal{N}_i(\mathcal{B}_i[u_i](k)) \\ \vdots \\ \mathcal{N}_p(\mathcal{B}_p[u_p](k)) \end{bmatrix}. \quad (4.4)$$

Note que como se trata de uma histerese totalmente atuada ( $p = \rho$ ) o índice  $\iota$  que em (4.1) e (4.2) indicavam o sinal intermediário, foi omitido, pois  $i = \iota$ . Para apresentar uma notação ainda mais reduzida, a igualdade (4.4) pode ser reapresentada como

$$\vec{v}(k) = \vec{N}(\vec{B}[\vec{u}](k)), \quad (4.5)$$

ou seja, as variáveis sem o subíndice e com a seta indicam variação em  $i$  de acordo com a quantidade de entradas de 1 a  $p$ .

**Observação 7** Com o objetivo de reduzir a notação, vetores como  $[u_1(k), \dots, u_i(k), \dots, u_p(k)]^T$  serão representados por variáveis simples sem subscrito da entrada  $i$ , mas com o sobrescrito de uma seta apontada para a direita, como  $\vec{u}(k)$ . O mesmo também é válido para  $\vec{v}$ ,  $\vec{N}$  e  $\vec{B}$  além de outras variáveis que ainda serão discutidas mais à frente como  $\vec{r}$ ,  $\vec{\theta}$ ,  $\vec{\Theta}$ ,  $\vec{F}$  e  $\vec{V}$ . Ou seja, todas são matrizes com  $p$  linhas e demais dimensões adequadas, de modo que qualquer subscrito adicionado não terá relação com a entrada a qual a variável pertence, mas sim às outras dimensões. Por outro lado, vetores com o índice dependente de  $l = 1, 2, \dots, m$  como  $[y_1(k), \dots, y_l(k), \dots, y_m(k)]^T$  podem ser denotados por  $\vec{y}(k)$ .

Para decidir entre as duas funções de borda apresentadas em (2.9) e (2.10) outras características da estrutura serão exploradas ao variar o sinal de entrada. O teste consiste em empregar uma amplitude diferente daquela aplicada nos testes da Figura 2.7, que servirá de comparação. O valor da amplitude empregada nos testes da Figura 2.7 eram unitários e passarão a valer 5. Com esse valor a parcela cúbica da função polinomial resulta em números elevados se comparados à ordem de grandeza da entrada, enquanto a função tangente hiperbólica mantém sua resposta na faixa de -1 a 1. Esse contraponto pode ser observado na Figura 4.1 ao comparar os testes em a) e b). Desse modo, o valor de  $b_0$  que faz o resultado de (2.9) se tornar mais próximo da ordem de grandeza da entrada pode ser desafiador de identificar por ser sensível a variações de entradas maiores que a unidade. Por outro lado o resultado de (2.10) fica limitado entre  $-b_0$  e  $b_0$  de modo que a sensibilidade não é um problema para esses parâmetros. Como um dos objetivos do exemplo a ser apresentado na Subseção 4.5.1 é o de estimar parâmetros próximos dos originais do processo conhecido, escolheu-se, de forma conservadora, manter os parâmetros do processo com pequenas variações. Desse modo, adota-se como envelope o seguinte conjunto de funções fundamentado em (2.10)

$$\vec{B} = \begin{cases} \vec{B}_e = \max(|\vec{u}|) \vec{\theta}_1 \tanh(\vec{\theta}_2 \vec{u}(k)), \\ \vec{B}_d = \max(|\vec{u}|) \vec{\theta}_3 \tanh(\vec{\theta}_4 \vec{u}(k)), \end{cases} \quad (4.6)$$

sendo  $\vec{B}_e$  e  $\vec{B}_d$  funções envelopes contínuas para as bordas esquerda e direita, respectivamente e,  $\vec{\theta}_1$  a  $\vec{\theta}_4$  vetores de constantes a serem estimados que controlam a forma das funções.

O fator  $\max(|\vec{u}|)$  foi adicionado para que os vetores  $\vec{\theta}_1$  e  $\vec{\theta}_3$  não tivessem que compensar o ganho associado ao sinal de entrada. Note que com ele o sinal de saída da função que altera a borda deixa de variar de  $-\vec{\theta}_3$  a  $\vec{\theta}_1$  e passa a variar de  $-\max(|\vec{u}|)\vec{\theta}_3$  a

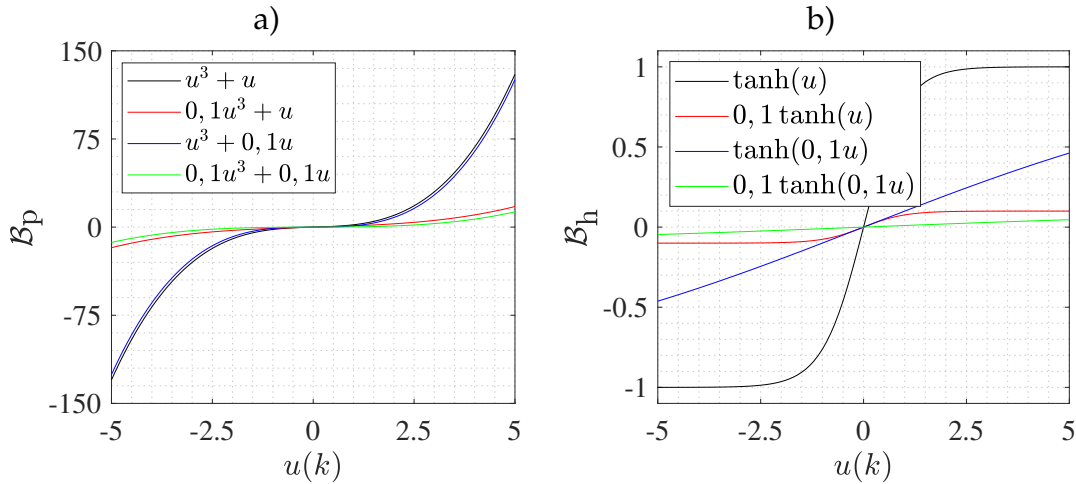


Figura 4.1: Teste do comportamento das funções de borda polinomial e tangente hiperbólica quando a entrada é variada de -5 a 5. Em a) apresenta-se o resultado para a função polinomial  $\mathcal{B}_p$  e em b) para a função tangente hiperbólica  $\mathcal{B}_h$ .

$\max(|\vec{u}|)\vec{\theta}_1$ , ou seja, acompanha o ganho do sinal de entrada, sem precisar alterar o  $\vec{\theta}_1$ . O mesmo raciocínio se aplica a  $\vec{\theta}_3$ .

A segunda questão a ser discutida na implementação do modelo GRDPI é com relação à dependência da taxa. Sejam  $\vec{r}_1(\vec{u}(k)), \dots, \vec{r}_{N_p}(\vec{u}(k))$  os vetores de disparos dinâmicos de cada um dos operadores do modelo GRDPI que podem ser descritos como

$$\vec{r}_j(\vec{u}(k)) = \frac{\vec{\theta}_1 \max(|\vec{u}|)(j-1)}{N_p} + \frac{\vec{\theta}_2 |\Delta \vec{u}(k)/T_s| T_s}{\max(|\vec{u}|)} = \frac{\vec{\theta}_1 \max(|\vec{u}|)(j-1)}{N_p} + \frac{\vec{\theta}_2 |\Delta \vec{u}(k)|}{\max(|\vec{u}|)}, \quad (4.7)$$

sendo  $\vec{\theta}_1$  e  $\vec{\theta}_2$  vetores de constantes a serem estimadas e  $\Delta \vec{u}(k)/T_s = (\vec{u}(k) - \vec{u}(k-1))/T_s$  uma aproximação discreta da derivada para um período de amostragem  $T_s$ .

Como foi desenvolvido para as funções de borda, algumas adições serão consideradas para remover a parte dos ganhos correspondentes a valores máximos das entradas ou saídas que recairiam sobre os parâmetros  $\vec{\theta}$  e deixá-los como competência da estrutura do processo definido, tornando o parâmetro a ser estimado menos sensível. A primeira alteração proposta em (2.14) é que a parcela mais à esquerda após o sinal de igualdade terá sua amplitude variada de acordo com um percentual da entrada de modo que se adiciona o fator  $\max(|\vec{u}|)/N_p$ . Destaca-se que  $N_p$  normaliza a primeira parcela de  $\vec{r}_j(\vec{u}(k))$  e  $\max(|\vec{u}|)$  o torna proporcional à amplitude da entrada. Com relação à segunda parcela, a participação e, portanto, a dependência que havia com relação ao período de amostragem é removida ao multiplicar a parcela pelo próprio período de amostragem. Note que isso poderia tornar sinais abruptos, como o sinal PRBS, um problema, uma vez que eles tem derivada com valor tanto maior quanto menor for o período de amostragem. Além disso, esse efeito também é normalizado ao dividir a parcela pela máxima excursão da entrada.

Tendo feito essas escolhas, o operador *play* do modelo GRDPI pode ser definido como

$$\vec{\mathcal{R}}_{\vec{r}_j(\dot{u}(k))}[\vec{\mathcal{B}}](k) = \begin{cases} \vec{\mathcal{B}}_e + \vec{r}_j(\dot{u}(k)), & \text{se } \Delta\vec{u}(k) < 0 \text{ e } \vec{\epsilon}_1 < 0, \\ \vec{\mathcal{B}}_d - \vec{r}_j(\dot{u}(k)), & \text{se } \Delta\vec{u}(k) > 0 \text{ e } \vec{\epsilon}_2 > 0, \\ \vec{\mathcal{R}}_{\vec{r}_j(\dot{u}(k))}[\vec{\mathcal{B}}](k-1), & \text{c.c.,} \end{cases} \quad (4.8)$$

em que  $\vec{\epsilon}_1 = \vec{\mathcal{B}}_e + \vec{r}_j(\dot{u}(k)) - \vec{\mathcal{R}}_{\vec{r}_j(\dot{u}(k))}[\vec{\mathcal{B}}](k-1)$  e  $\vec{\epsilon}_2 = \vec{\mathcal{B}}_d - \vec{r}_j(\dot{u}(k)) - \vec{\mathcal{R}}_{\vec{r}_j(\dot{u}(k))}[\vec{\mathcal{B}}](k-1)$ . Assim,

$$\vec{v}(k) = \vec{\theta}_0 \vec{\mathcal{B}} + \sum_{j=1}^{N_p} \vec{\theta}_j \vec{\mathcal{R}}_{\vec{r}_j(\dot{u}(k))}[\vec{\mathcal{B}}](k), \quad (4.9)$$

define a saída do modelo GRDPI. Os parâmetros a estimar podem ser organizados em uma matriz  $\vec{\Theta} \in \mathbb{R}^{p \times N_V}$ , sendo  $N_V$  a quantidade total de variáveis a serem estimadas, ou seja,  $N_p + 1$  parâmetros dos operadores *play*, 4 parâmetros das funções de borda e 2 parâmetros dos disparos dinâmicos. Nesse caso  $N_V = N_p + 7$ , ficando a matriz de parâmetros configurada do seguinte modo sem motivo para ordenação

$$\vec{\Theta} = \left[ \vec{\theta}_0 \mid \vec{\theta}_1 \quad \dots \quad \vec{\theta}_{N_p} \mid \vec{\theta}_1 \quad \dots \quad \vec{\theta}_4 \mid \vec{\theta}_1 \quad \vec{\theta}_2 \right]. \quad (4.10)$$

### 4.3 Formulação do Problema para Não Linearidades Quase Estáticas

Considere o modelo de Hammerstein multivariável em tempo discreto apresentado na Figura 4.2. Ele foi introduzido no Capítulo 2 e é reproduzido novamente aqui por conveniência para destacar determinados aspectos. Na Figura 4.2, nota-se que cada entrada  $u_i(k)$  se relaciona com uma e, apenas uma, variável intermediária  $v_i(k)$  por meio de uma função não linear  $\mathcal{N}_i$  que representa uma histerese, para cada  $i = 1, 2, \dots, p$ . Na sequência, uma função dinâmica linear multivariável  $\mathcal{L}$ , relaciona os sinais intermediários  $v_i(k)$  com as saídas  $y_l(k)$ ,  $l = 1, 2, \dots, m$ . Os elementos destacados em vermelho são desconhecidos, de modo que apenas os sinais de entrada e saída, em preto, constituem informação acessível.

Em um primeiro momento, suponha que os sinais intermediários — que recorda-se ao leitor que podem ser denotados por  $\vec{v}(k)$  (ver Observação 7) — sejam todos conhecidos. À vista disso, com um ensaio estático adequado pode-se coletar um conjunto de  $N_N$  dados não lineares  $\{\vec{u}(k), \vec{v}(k)\}_{k=1}^{N_N}$  e, em seguida, ajustar uma curva que aproxime a parcela não linear  $\vec{\mathcal{N}}$ . Adicionalmente, um segundo ensaio dinâmico pode ser executado resultando em outro conjunto com  $N_L$  dados  $\{\vec{v}(k), \vec{y}(k)\}_{k=1}^{N_L}$  para estimar a parcela linear  $\mathcal{L}$ . Como é comum que a suposição proposta não se aplique na prática, o desconhecimento do sinal intermediário eleva o grau de complexidade do problema. Todavia, empregar a separação dos subsistemas, como se observa na Figura 4.2, introduz a possibilidade de cancelar somente a não linearidade e obter um sistema dinâmico linear, no qual as técnicas clássicas de controle se aplicam (Dorf e Bishop, 2011). Em

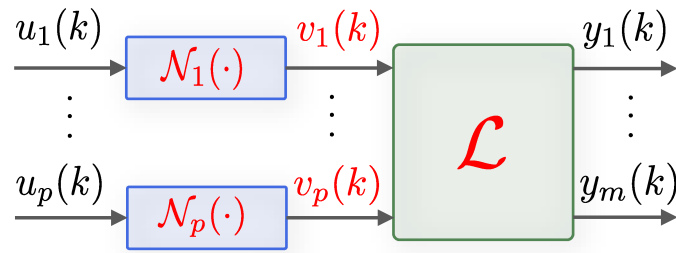


Figura 4.2: Representação do modelo de Hammerstein KU multivariável em malha aberta com distinção de cores para os dados conhecidos e desconhecidos. Os elementos pintados de vermelho são desconhecidos. Ou seja, somente os sinais de entrada  $u_i(k)$ ,  $i = 1, 2, \dots, p$  e saída  $y_l(k)$ ,  $l = 1, 2, \dots, m$ , são acessíveis para medição. Cada uma das não linearidades  $\mathcal{N}$  representam um subsistema GRDPI.

busca de preservar tal característica proporcionada pelo modelo de Hammerstein, a estratégia proposta neste capítulo consiste em estimar os sinais intermediários  $\vec{v}(k)$  e os subsistemas  $\mathcal{N}_i$ , para  $i = 1, 2, \dots, p$  e  $\mathcal{L}$ . Como foi feito, na Seção 3.4, destaca-se aqui, de maneira sucinta, as consequências do desconhecimento do sinal intermediário na metodologia.

**Observação 8** Do ponto de vista dos dados de entrada e saída, um sistema descrito por  $\{\bar{\lambda}\mathcal{N}_i, \mathcal{L}/\bar{\lambda}\}$  não difere de outro descrito por  $\{\mathcal{N}_i, \mathcal{L}\}$ . Ou seja, como o sinal intermediário é desconhecido, os subsistemas podem ser estimados a menos de um fator de escala  $\bar{\lambda}$ , também desconhecido.

A identificação de cada uma das parcelas é feita por meio de duas etapas diferentes, inclusive sendo a justificativa para o nome do método: identificação em duas etapas. Uma das etapas objetiva estimar os parâmetros do elemento não linear  $\mathcal{N}$  e, a outra, estimar os parâmetros da parte linear  $\mathcal{L}$ . Sem entrar em detalhes, o conceito do ensaio realizado na primeira etapa é aplicar um sinal lento, ou até mesmo constante, que permita ao processo entrar em regime permanente. Desse modo, a dinâmica não terá mais influência e, portanto, o sinal de saída irá se tornar uma versão escalonada do sinal intermediário. Por sua vez, na segunda etapa espera-se captar a dinâmica linear presente no processo, de tal modo que o sinal a ser aplicado precisa respeitar a premissa da persistência de excitação. Ainda no contexto da segunda etapa, devido à característica quase estática da histerese, um terceiro ensaio é realizado para ajustar os ganhos do modelo na perspectiva entrada-saída. Partindo do pressuposto que esses três ensaios tenham sido realizados corretamente, formula-se o seguinte problema.

**Problema 3** Dados três conjuntos de medições de entrada e saída do processo, o problema sob estudo é dividido em duas etapas. A primeira delas consiste em estimar os parâmetros (4.10) de  $\mathcal{N}_i$  do modelo de Hammerstein GRDPI apresentado em (4.9). Para executar essa etapa emprega-se os dados em regime permanente de um ensaio cuja entrada é uma sequência de degraus ascendentes e descendentes  $\{\bar{u}(k), \bar{y}(k)\}_{k=1}^{N_N}$ . E, de forma complementar, a outra etapa consiste em obter as estimativas da ordem e das matrizes do modelo em espaço de estados (1.1b)

e (1.1c) a partir de outros dois ensaios. Um deles é utilizado para identificar os parâmetros, cuja entrada é um sinal PRBS  $\{\vec{u}(k), \vec{y}(k)\}_{k=1}^{N_L}$  e, o outro, para captar o ganho total das entradas para as saídas, cuja entrada é um sinal que exibe as características estáticas e dinâmicas do sistema em uma ampla faixa como um sinal de carga e descarga  $\{\vec{u}(k), \vec{y}(k)\}_{k=1}^{N_G}$ .

## 4.4 Associação entre o MPSO e o Algoritmo de Subespaço

Em primeira análise, a identificação de modelos de Hammerstein necessita do conhecimento dos sinais intermediários  $\vec{v}(k)$ . Por exemplo, pressupondo que os sinais  $\vec{v}(k)$  fossem conhecidos, uma possível estratégia seria ajustar polinômios para estimar as funções não lineares  $\vec{N}$  por meio de dados coletados em regime permanente  $\{\vec{u}(k), \vec{v}(k)\}_{k=1}^{N_N}$ . Além disso, técnicas de identificação poderiam ser aplicadas para obter um modelo ARX, recorrendo aos dados dinâmicos  $\{\vec{v}(k), \vec{y}(k)\}_{k=1}^{N_L}$ . Via de regra  $\vec{v}(k)$  são sinais inacessíveis, ou até mesmo inexistentes, de maneira que o problema se torna mais complexo (de Paula et al., 2015). Todavia, parte da simplicidade pode ser recuperada utilizando pelo menos duas estratégias que serão nomeadas de direta e reversa. Na metodologia direta, primeiro identifica-se  $\vec{N}$  e depois  $\mathcal{L}$  enquanto na reversa a ordem de identificação é invertida (como foi feito no Capítulo 3). Aqui no Capítulo 4 será tratada a metodologia direta, uma vez que com ela evita-se a necessidade de inverter o sistema dinâmico linear, tarefa complexa para o caso em que o modelo identificado é de fase não mínima (Verhaegen e Westwick, 1996; Verhaegen, 1996). Também destaca-se que no Capítulo 3 essa inversão não foi necessária graças à escolha específica do sinal intermediário.

A aplicação de degraus em processos estáveis resulta em sinais de saída constantes em regime permanente. Desse modo é possível eliminar os efeitos da dinâmica e obter informações a respeito da parcela estática ao utilizar os dados em regime permanente (Dorf e Bishop, 2011). Explorando esses aspectos, na estratégia direta, degraus consecutivos crescentes e decrescentes são aplicados na entrada do processo, de modo que se torna possível estimar a não linearidade por meio do conjunto de dados em regime  $\{\vec{u}(k), \vec{y}(k)\}_{k=1}^{N_N}$ . Em concordância com a Observação 8 a estimativa é obtida a menos de um fator de escala desconhecido.

As duas próximas subseções descrevem a metodologia proposta neste trabalho para resolver o Problema 3, ou seja, estimar modelos de Hammerstein MIMO GRDPI a partir de dados de entrada e saída coletados do processo. À vista disso, sugerem-se adaptações na versão do MPSO desenvolvido por Yang et al. (2013), sem as quais não seria possível estimar o modelo da histerese multivariável considerando o acoplamento dinâmico que a parte linear  $\mathcal{L}$  insere nos sinais de saída disponíveis  $\vec{y}(k)$ . Como a metodologia é desenvolvida em duas etapas, apresenta-se cada uma delas em uma subseção. Em 4.4.1 descreve-se o passo a passo para identificar o modelo da histerese multivariável e em 4.4.2 as considerações sobre o emprego do MOESP elaborado em (Verhaegen e Dewilde, 1992a) para identificar a parcela dinâmica linear.

### 4.4.1 Identificação do modelo GRDPI

Uma vez definido o conceito do teste que será utilizado para estimar o modelo GRDPI, é importante também definir qual o método, a fim de direcionar a escolha dos parâmetros. Como o modelo de histerese é não linear e não diferencial, o cálculo do gradiente não pode ser formulado por expressões explícitas. Assim, os métodos convencionais de otimização baseados em gradiente não se aplicam. Embora este problema possa ser resolvido por um algoritmo de busca direta, como o algoritmo simplex de Nelder-Mead, ele depende fortemente de uma boa escolha de pontos iniciais e pode estagnar com frequência em ótimos locais. Portanto, um algoritmo de busca global é necessário, como o algoritmo genético, os evolutivos ou o PSO (Yang et al., 2013).

Seguindo a abordagem proposta por Yang et al. (2013) uma otimização de enxame de partículas modificada é aplicada para identificar os parâmetros do modelo GRDPI. Esse algoritmo recebe como entrada apenas os dados em regime permanente coletados do processo, não havendo a necessidade de manter o período de amostragem constante.

Com as informações apresentadas até o momento, pode-se explicar alguns conceitos para direcionar a escolha dos parâmetros do ensaio estático. Os conceitos são anunciados independentemente, ou seja, sem entrar em detalhes sobre a interação entre eles. Todavia, deve-se ter em vista que as escolhas podem influenciar umas nas outras, de modo que a ponderação de cada um dos conceitos deve ser feita pelo projetista. Para fins de esclarecimento, considere um processo com duas entradas e duas saídas que é composto de uma histerese construída pelo modelo GRDPI cujos parâmetros são ordenados como apresentado em (4.10)

$$\vec{\Theta} = \left[ \begin{array}{c|cccccccc|cccc|cc} 0,5 & 0,2 & 0,6 & 0,7 & 0,4 & 0,2 & 0,1 & 0,3 & 0,4 & 0,2 & 0,8 & 0,9 & 0,25 & 0,8 & 0,3 & 1 & 1 \\ \hline 0,2 & 1 & 0,9 & 0,1 & 0,7 & 0,1 & 0,8 & 0,4 & 0,1 & 0,5 & 0,6 & 1 & 0,6 & 1 & 0,4 & 1 & 1 \end{array} \right]' \quad (4.11)$$

e a parcela linear é descrita pelas matrizes do sistema em espaço de estados (1.1b) e (1.1c)

$$A = \begin{bmatrix} 0,8 & 0,3 \\ 0,1 & 0,7 \end{bmatrix}, \quad B = \begin{bmatrix} 0,1 & 0,7 \\ 0,5 & 0,9 \end{bmatrix}, \quad C = \begin{bmatrix} 0,5 & 0,4 \\ 0,7 & 0,4 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 0,4 & 0 \\ 0 & 0,1 \end{bmatrix}. \quad (4.12)$$

A resposta ao degrau unitário desse processo pode ser vista na Figura 4.3a). Os degraus são aplicados em cada uma das entradas por vez, mas quando ocorre a excitação de cada entrada, ambas saídas respondem simultaneamente, uma vez que o modelo dinâmico linear (4.12) é acoplado.

A primeira escolha é sobre qual das saídas coletar. Note que dentro do contexto puramente numérico no qual a dissertação é desenvolvida não há preferência sobre qualquer uma das saídas. Para perceber essa indiferença, pode-se recordar que o processo é definido por meio da estrutura de Hammerstein, ou seja, uma parcela não

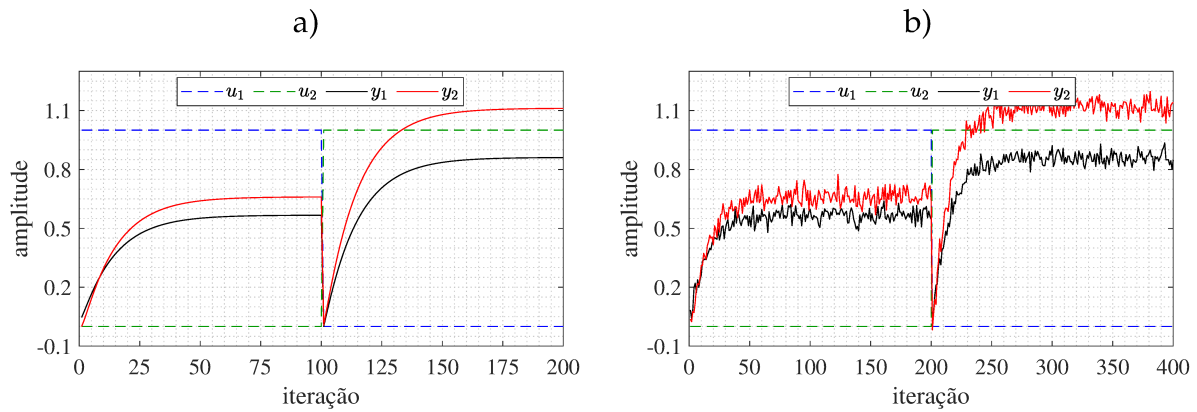


Figura 4.3: Resposta ao degrau de um exemplo ilustrativo para justificar a escolha de parâmetros a serem empregados no sinal de degraus crescentes e decrescentes. Em a) apresenta-se a resposta sem ruído e em b) com ruído que resulta em SNR de 27 dB.

linear seguida por uma outra parcela linear. Como se excita apenas uma das entradas por vez, é razoável assumir o caso de uma entrada e várias saídas (SIMO, do inglês, *Single Input Multiple Output*), apresentado na Figura 4.4. Uma vez que qualquer uma das não linearidades afetam todas as saídas e o modelo  $\mathcal{L}$  é linear, em regime permanente, as saídas serão cópias umas das outras com ganhos diferentes. Nesta etapa o ganho não é de interesse, uma vez que ele deverá ser compensado pela próxima etapa, por isso, não há preferência na escolha de qualquer uma das saídas. Para manter um padrão, durante o texto será selecionada a que excursiona pela maior faixa.

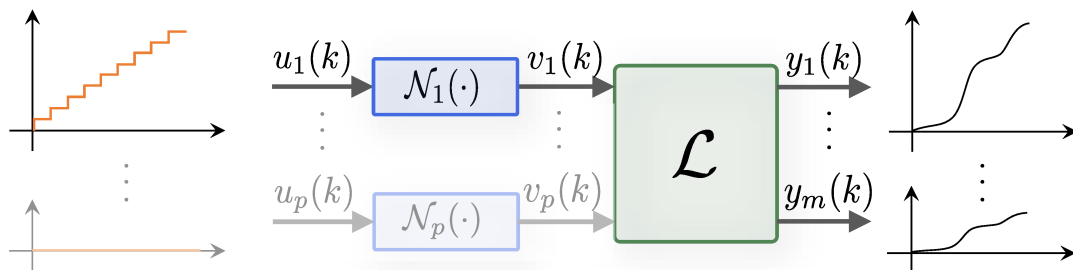


Figura 4.4: Esquema da influência da excitação de uma entrada por vez em todas as saídas do modelo de Hammerstien GRDPI. Todas as saídas respondem de forma idêntica, exceto por um fator de escala.

Ainda com relação ao teste apresentado na Figura 4.3a), é possível constatar que a resposta de cada degrau nas diferentes entradas leva aproximadamente 100 iterações para atingir o regime permanente. Ademais, na Figura 4.3b) apresenta-se a versão do teste afetada por ruído de medição branco que leva à SNR de 27 dB. Nessa ocasião, para apresentar um tratamento ao ruído, a duração do teste de cada degrau foi estendida por mais 100 iterações com a finalidade de substituir o que antes era um ponto em regime permanente pela média dos últimos 100 dados. Como o ruído de medição é branco, quanto mais pontos forem coletados em regime permanente, mais precisa pode ser a

resposta. Apesar disso, também é conhecido que quanto mais pontos forem coletados, mais tempo o processo deverá ficar inoperante para execução do ensaio, de modo que o projetista deve avaliar os recursos disponíveis e os benefícios decorrentes.

O último parâmetro a ser definido é a quantidade de degraus do ensaio estático. Supondo que a coleta de dados em regime permanente seja feita em intervalos regulares — o que não é necessário, mas facilita a visualização e análise dos gráficos — pode-se inferir que o tempo de ensaio aumenta de maneira proporcional à quantidade de degraus escolhida pelo projetista. Além disso, o número total de degraus é exatamente a quantidade de dados aplicados ao MPSO, portanto, exigirá maior tempo de processamento à medida que aumenta a massa de dados.

Outro atributo que se altera com a quantidade de degraus é a precisão desejada ou necessária. No entanto, foi indicado a dois parágrafos que o ruído de medição também afeta a precisão. Para esclarecer os efeitos do ruído de medição chama-se a atenção para o que ocorre na Figura 4.5, em que se apresenta os quatro primeiros degraus de dois testes realizados sobre o mesmo processo apresentado em (4.11) e (4.12). À esquerda, em a), apresenta-se um teste cujos degraus variam de 0,2 em 0,2. Por sua vez, em b), a variação dos degraus é dez vezes menor, ou seja, de 0,02 em 0,02, de modo que os pentagramas em preto, visivelmente, não apresentam um padrão monotonicamente crescente como ocorria com a variação de 0,2. Por isso, apresenta-se a observação a seguir.

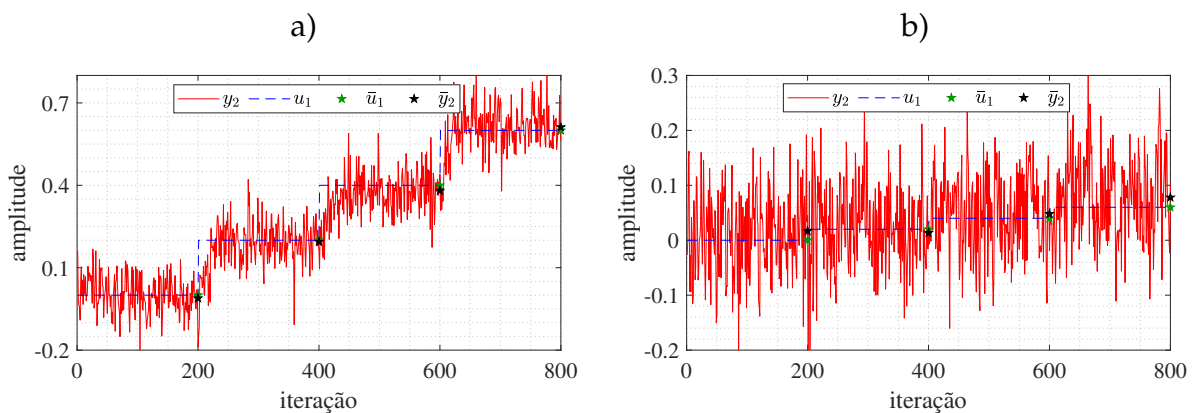


Figura 4.5: Teste com degraus consecutivos realizado para esclarecer a justificativa da escolha da quantidade total de degraus após definir a amplitude total de excursão. Em a) apresenta-se um cenário com passo 0,2 nos degraus, cuja média em regime permanente sempre acrescenta informação. Por sua vez, em b), apresenta-se um cenário oposto no qual o erro da média em regime permanente supera o passo dos degraus que é 0,02.

**Observação 9** Na Figura 4.5b) observa-se que a média dos dados em regime permanente  $\bar{y}_2$  do primeiro degrau é superior à do segundo degrau. Isso sugere que reduzir o passo dos degraus para 0,02 é desnecessário, uma vez que o erro de cálculo da média empregando 100 dados em regime permanente é superior ao passo dos degraus. Desse modo, cabe ao projetista avaliar o custo computacional e os ganhos de desempenho em termos de estimativa a fim de definir o passo

para realizar o teste. Caso se julgue importante reduzir o passo para 0,02, a quantidade de dados em regime permanente deve ser aumentada.

Após coletar os dados do ensaio estático pode-se proceder com a implementação do algoritmo de otimização que estimará os parâmetros do modelo GRDPI. Para essa tarefa, a função custo é definida como

$$\vec{F}(\vec{\Theta}) = \sum_{k=1}^{N_N} (\Psi(k) - \hat{\Psi}(k))^2, \quad (4.13)$$

na qual  $\Psi$  é um sinal coletado do processo, enquanto  $\hat{\Psi}$  é a estimativa do sinal  $\Psi$  produzida pelo modelo GRDPI. A variável  $\vec{\Theta}$  que aparece como argumento de  $\vec{F}$  indica que  $\hat{\Psi}$  é afetado pelos parâmetros do modelo GRDPI. Portanto, no contexto deste capítulo, o problema de otimização pode ser enunciado de maneira completa como

$$\min_{\vec{\Theta}} \vec{F}(\vec{\Theta}) = \sum_{k=1}^{N_N} (\bar{y}(k) - \hat{y}(k))^2, \quad (4.14)$$

sujeito a

$$\hat{B} = \begin{cases} \hat{B}_e = \max(|\bar{u}|) \hat{\theta}_1 \tanh(\hat{\theta}_2 \bar{u}(k)), \\ \hat{B}_d = \max(|\bar{u}|) \hat{\theta}_3 \tanh(\hat{\theta}_4 \bar{u}(k)), \end{cases} \quad (4.15)$$

$$\hat{r}_j(\hat{u}(k)) = \frac{\hat{\theta}_1(j-1)}{N_p} + \frac{\hat{\theta}_2 |\Delta \bar{u}(k)|}{\max(|\bar{u}|)}, \quad (4.16)$$

$$\hat{G}R_{\hat{r}_j(\hat{u}(k))}[\hat{B}](k) = \begin{cases} \hat{B}_e + \hat{r}_j(\hat{u}(k)), & \text{se } \Delta \bar{u}(k) < 0 \text{ e } \hat{\xi} < 0, \\ \hat{B}_d - \hat{r}_j(\hat{u}(k)), & \text{se } \Delta \bar{u}(k) > 0 \text{ e } \hat{\xi} > 0, \\ \hat{G}R_{\hat{r}_j(\hat{u}(k))}[\hat{B}](k-1), & \text{c.c.}, \end{cases} \quad (4.17)$$

$$\hat{v}(k) = \hat{\theta}_0 \hat{B} + \sum_{j=1}^{N_p} \hat{\theta}_j \hat{G}R_{\hat{r}_j(\hat{u}(k))}[\hat{B}](k), \quad (4.18)$$

sendo  $\hat{v}(k)$  uma estimativa da saída  $\hat{y}(k)$  desconsiderando os ganhos da parte linear que serão compensados mais adiante.

O algoritmo PSO clássico seria capaz de estimar os parâmetros definidos em (4.10) por meio do problema de otimização (4.14), no entanto, segundo Yang et al. (2013) a rápida convergência do PSO pode prendê-lo em ótimos locais com frequência. Para solucionar o impasse, eles propõem duas adições ao PSO clássico para auxiliar na convergência do algoritmo. Uma delas é adotar a estratégia adaptativa, informada e eficaz (EIA-PSO, do inglês *Effective Informed Adaptive PSO*) que consiste em compartilhar informações de um grupo ao invés da melhor partícula. Em termos de implementação, na atualização da velocidade, substitui-se as variáveis que representam as melhores

posições individuais e globais pelas médias ponderadas de duas partículas vizinhas e das  $\xi$  melhores partículas, respectivamente (ver (4.23)). A segunda mudança consiste em aplicar uma mutação na velocidade quando a função custo deixar de reduzir por  $K_\alpha$  iterações consecutivas. Essas duas alterações estão destacadas em verde na Figura 4.6.

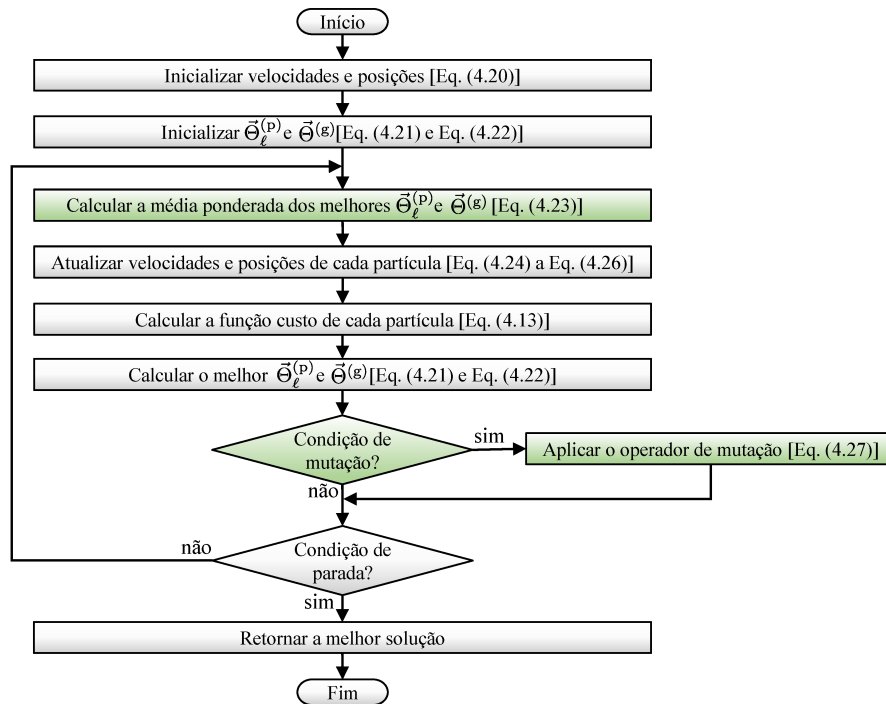


Figura 4.6: Algoritmo **MPSO** cujas modificações em relação ao **PSO** clássico recebem destaque em verde. Adaptado de (Yang et al., 2013).

O algoritmo **MPSO** foi desenvolvido para o caso monovariável em (Yang et al., 2013) com base em outros trabalhos como (Qi et al., 2009) e (Li et al., 2010). Desse modo, os passos a serem descritos em detalhes foram inspirados no conteúdo desses três trabalhos. No entanto, os aspectos abordados na dissertação contribuem para a elucidação da metodologia no contexto multivariável e também para esclarecimento de alguns parâmetros do caso monovariável, uma vez que nos três artigos citados neste parágrafo, há menos espaço para a descrição detalhada a ser feita a seguir.

A ideia básica do **MPSO** é armazenar em  $\vec{\Theta}_\ell^{(p)}$ , nomeada melhor posição pessoal, a localização no espaço de busca que produz o menor custo histórico de cada partícula  $\ell$  até a iteração atual  $\kappa$ . Por sua vez, o ponto do espaço de busca que resulta no menor custo de toda nuvem é nomeado de melhor posição global  $\vec{\Theta}^{(g)}$ . Ao fim do algoritmo, cujos passos são descritos a seguir, o resultado é o valor armazenado em  $\vec{\Theta}^{(g)}$ .

- **Passo 1:** inicializar os parâmetros do algoritmo **MPSO**, sendo eles o número máximo de iterações  $K_{\max}$ , a quantidade total de partículas  $P$ , a probabilidade de ocorrer mutação  $\alpha$ , a quantidade  $\xi$  de partículas utilizadas para calcular a média ponderada global, a quantidade de iterações a partir da qual podem ocorrer

mutações  $K_\alpha$ , as taxas de aprendizado cognitivo e social  $\omega_1$  e  $\omega_2$ , o número total de variáveis do modelo GRDPI  $N_V$  e, por fim, os valores limites de posição e velocidade  $\vec{\Theta}_{\min} \in \mathbb{R}^{p \times N_V}$ ,  $\vec{\Theta}_{\max} \in \mathbb{R}^{p \times N_V}$  e  $\vec{V}_{\max} \in \mathbb{R}^{p \times N_V}$ . Na sequência, para cada partícula  $\ell$ , deve-se inicializar as matrizes de velocidade e posição de acordo com

$$\begin{aligned}\vec{V}_\ell(1) &= -\vec{V}_{\max} + \vec{U} \odot 2\vec{V}_{\max}, \\ \vec{\Theta}_\ell(1) &= \vec{\Theta}_{\min} + \vec{U} \odot (\vec{\Theta}_{\max} - \vec{\Theta}_{\min}),\end{aligned}\quad (4.19)$$

sendo  $\odot$  um símbolo que indica a operação de multiplicação elemento a elemento e  $\vec{U} \in \mathbb{R}^{p \times N_V}$  uma matriz preenchida por amostras da distribuição uniforme (3.21)  $\mu_u = 1/2$  e  $\sigma_u = \sqrt{1/12}$ , ou seja  $u(1/2, \sqrt{1/12})$ .

- **Passo 2:** avaliar a função de aptidão de cada uma das  $P$  partículas de acordo com  $\vec{F}(\vec{\Theta}_\ell)$  apresentada em (4.14). Ainda neste passo, atualizar  $\vec{\Theta}_\ell^{(p)}$  por meio de

$$\vec{\Theta}_\ell^{(p)}(\kappa + 1) = \begin{cases} \vec{\Theta}_\ell^{(p)}(\kappa), & \text{se } \vec{F}(\vec{\Theta}_\ell(\kappa + 1)) > \vec{F}(\vec{\Theta}_\ell^{(p)}(\kappa)), \\ \vec{\Theta}_\ell(\kappa + 1), & \text{c.c.,} \end{cases}\quad (4.20)$$

e também a variável  $\vec{\Theta}_\ell^{(g)}$  de acordo com

$$\vec{\Theta}_\ell^{(g)}(\kappa + 1) = \min(\vec{\Theta}_\ell^{(p)}(\kappa + 1)), \quad \ell = 1, \dots, P, \quad (4.21)$$

ou seja, seleciona-se a melhor dentre todas as posições visitadas por todas as partículas até o instante atual.

- **Passo 3:** dispor as partículas em ordem crescente pelo valor da função de aptidão  $\vec{F}(\vec{\Theta}_\ell^{(p)})$  e armazená-las em uma matriz  $[\vec{F}_1, \vec{F}_2, \dots, \vec{F}_P]$ , além de associar o subíndice às devidas posições  $[\vec{\Theta}_1^{(p)}, \vec{\Theta}_2^{(p)}, \dots, \vec{\Theta}_P^{(p)}]$  que geram tais valores de  $\vec{F}_\ell$ . Em seguida, calcular as médias ponderadas das melhores posições global  $\vec{M}^{(g)}$  e pessoal  $\vec{M}^{(p)}$  por meio de

$$\begin{aligned}\vec{M}^{(g)} &= \sum_{q=1}^{\xi} \vec{c}_q \odot \vec{\Theta}_q^{(p)}, \quad \vec{c}_q = \frac{1/\vec{F}_q}{\sum_{g=1}^{\xi} 1/\vec{F}_g}, \\ \vec{M}_s^{(p)} &= \frac{\vec{F}_s \odot \vec{\Theta}_{s-1}^{(p)} + \vec{F}_{s-1} \odot \vec{\Theta}_s^{(p)}}{\vec{F}_s + \vec{F}_{s-1}}, \quad s = 2, \dots, P, \\ \vec{M}_1^{(p)} &= \vec{\Theta}_1^{(p)}.\end{aligned}\quad (4.22)$$

- **Passo 4:** atualizar a velocidade e a posição de cada partícula em concordância com

$$\begin{aligned}\vec{V}_\ell(\kappa + 1) &= \chi(\vec{V}_\ell(\kappa) + \omega_1 v(\vec{M}_\ell^{(p)} - \vec{\Theta}_\ell(\kappa)) + \omega_2 v(\vec{M}^{(g)} - \vec{\Theta}_\ell(\kappa))), \\ \vec{\Theta}_\ell(\kappa + 1) &= \vec{\Theta}_\ell(\kappa) + \vec{V}_\ell(\kappa + 1),\end{aligned}\quad (4.23)$$

em que  $\chi = 2/|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}|$ , sendo  $\varphi = \omega_1 + \omega_2 > 4$  e  $v \in \mathbb{R}$  representa, não uma variável que armazena um valor fixo, mas sim uma amostra diferente da

distribuição uniforme  $\mathcal{U}(1/2, \sqrt{1/12})$  a cada vez que ela é utilizada na codificação, como a função `rand` do **MATLAB**<sup>®</sup>. Para controlar o movimento excessivo do enxame, a velocidade de cada partícula é limitada entre  $-\vec{V}_{\max}$  e  $\vec{V}_{\max}$

$$\vec{V}_\ell(\kappa + 1) = \begin{cases} -\vec{V}_{\max}, & \text{se } \vec{V}_\ell(\kappa + 1) < -\vec{V}_{\max}, \\ \vec{V}_{\max}, & \text{se } \vec{V}_\ell(\kappa + 1) > \vec{V}_{\max}. \end{cases} \quad (4.24)$$

Por sua vez, quando os limites da matriz de posição são violados, ela é reinicializada e a velocidade é espelhada do seguinte modo

$$\begin{aligned} \vec{\Theta}_\ell(\kappa + 1) &= \vec{\Theta}_{\min} + \vec{U} \odot (\vec{\Theta}_{\max} - \vec{\Theta}_{\min}), \\ \vec{V}_\ell(\kappa + 1) &= -\vec{V}_\ell(\kappa + 1). \end{aligned} \quad (4.25)$$

- **Passo 5:** avaliar a função custo de cada partícula  $\ell$  de acordo com  $\vec{F}(\vec{\Theta}_\ell)$  descrita em (4.14).
- **Passo 6:** atualizar  $\vec{\Theta}_\ell^{(p)}$  e  $\vec{\Theta}^{(g)}$  de acordo com (4.20) e (4.21).
- **Passo 7:** se forem concluídas  $K_\alpha$  ou mais iterações sem evolução significativa no melhor custo da nuvem  $\vec{F}(\vec{\Theta}^{(g)})$ , o operador de mutação atua com probabilidade  $\alpha$ , como pode ser observado na estrutura de decisão em verde na Figura 4.6. De maneira objetiva, caso se passem  $K_\alpha$  ou mais iterações sem que  $\vec{F}(\vec{\Theta}^{(g)})$  reduza ao menos 0,1% do seu valor atual e a condição  $v < \alpha$  for atendida, deve-se sortear uma dimensão  $\gamma$  da velocidade de uma partícula aleatória  $\eta$  e alterá-la por um fator  $\vec{\zeta} = \mathcal{U}(1/2, \sqrt{1/12})$

$$\vec{V}_\eta^\gamma(\kappa + 1) = \begin{cases} \vec{V}_\eta^\gamma(\kappa + 1) + \vec{\zeta} \odot \vec{V}_{\max, \eta}^\gamma, & \text{se } \vec{\zeta} > 0,5, \\ \vec{V}_\eta^\gamma(\kappa + 1) - \vec{\zeta} \odot \vec{V}_{\max, \eta}^\gamma, & \text{c.c.,} \end{cases} \quad (4.26)$$

caso contrário, vá para o Passo 8.

- **Passo 8:** se a contagem de iterações exceder  $K_{\max}$ , armazenar  $\vec{\Theta}^{(g)}$  como a solução final, caso contrário, voltar para o Passo 3 como indicado na Figura 4.6.

**Observação 10** É relevante destacar que as operações apresentadas em todos os passos são realizadas em bloco para cada uma das não linearidades  $N_i$ ,  $i = 1, \dots, p$ , mas de forma desacoplada e independente. Isso significa que uma função  $N_i$  não influencia a resposta de  $N_{\neq i}$  e a razão é que os testes com os degraus excitam apenas uma das não linearidades por vez. Por exemplo, a

forma expandida da primeira linha de (4.23) é

$$\begin{aligned}
V_{1,\ell}(\kappa + 1) &= \chi(V_{1,\ell}(k) + \omega_1 v(M_{1,\ell}^{(p)} - \Theta_{1,\ell}(\kappa)) + \omega_2 v(M_1^{(g)} - \Theta_{1,\ell}(\kappa))), \\
\vdots &= \vdots \\
V_{i,\ell}(\kappa + 1) &= \chi(V_{i,\ell}(k) + \omega_1 v(M_{i,\ell}^{(p)} - \Theta_{i,\ell}(\kappa)) + \omega_2 v(M_i^{(g)} - \Theta_{i,\ell}(\kappa))), \\
\vdots &= \vdots \\
V_{p,\ell}(\kappa + 1) &= \chi(V_{p,\ell}(k) + \omega_1 v(M_{p,\ell}^{(p)} - \Theta_{p,\ell}(\kappa)) + \omega_2 v(M_p^{(g)} - \Theta_{p,\ell}(\kappa))),
\end{aligned} \tag{4.27}$$

cujas operações são feitas em conjunto, mas não há cruzamento de variáveis com índices  $i$  diferentes. O mesmo ocorre dentro do modelo GRDPI ao computar a função custo que possui  $i$  saídas independentes.

**Observação 11** Diferente do símbolo  $v$  que é uma amostra distinta da distribuição uniforme a cada vez que aparece,  $\vec{\zeta} \in \mathbb{R}^{p \times 1}$ , apresentada no passo 7, é uma variável gerada para verificar a parte condicional em (4.26) e é mantida inalterada por toda a extensão de (4.26) a cada iteração que for executada, por isso  $\vec{\zeta}$  recebe uma denominação diferente de  $v$ , mesmo sendo gerada da mesma forma.

#### 4.4.2 Identificação do modelo LIT

Após a parcela não linear  $\vec{N}$  ter sido identificada, torna-se possível estimar o sinal intermediário ao aplicar o sinal de entrada no modelo obtido  $\hat{\vec{N}}$  do seguinte modo:  $\hat{\vec{v}}(k) = \hat{\vec{N}}(\hat{\vec{B}}[u](k))$ . Destaca-se que no Problema 3, sob investigação, há três elementos desconhecidos,  $\vec{N}$ ,  $\vec{v}$  e  $\mathcal{L}$  de maneira que a estimativa do sinal intermediário  $\hat{\vec{v}}(k)$  viabiliza a aplicação do conjunto de dados  $\{\hat{\vec{v}}(k), \hat{\vec{y}}(k)\}_{k=1}^{N_L}$  para fazer a identificação do modelo dinâmico linear  $\mathcal{L}$  sem ter que recorrer a inversões aproximadas de modelos dinâmicos, como em (Aljanaideh et al., 2018) e (Aljanaideh et al., 2022).

Neste trabalho, a identificação da parcela dinâmica linear do modelo de Hammerstien MIMO GRDPI é feita por meio de um algoritmo de subespaços. O MOESP ordinário é um dos algoritmos mais básicos da família de métodos de subespaços, mas é capaz de identificar corretamente os modelos multivariáveis em espaço de estados mesmo na presença de ruído branco de medição (Verhaegen e Dewilde, 1992a; Katayama, 2005). Esse algoritmo opera sinais de entrada e saída para estimar um modelo como aquele apresentado em (1.1b) e (1.1c), que assimile a dinâmica linear presente nos dados do processo. Como o sinal utilizado no primeiro ensaio é projetado especificamente para coletar os dados em regime permanente, nessa segunda etapa, outro ensaio é realizado no processo com uma entrada apropriada para gerar dados dinâmicos para o MOESP. O sinal de teste a ser aplicado na entrada do processo deve ser persistentemente excitante, por exemplo, PRBS ou o sinal aleatório (Verhaegen e Verdult, 2007).

O PRBS é uma sequência periódica com período máximo  $2^{n_p} - 1$ , que varia

entre dois níveis  $+\bar{b}$  e  $-\underline{b}$ , sendo  $\bar{b} = \underline{b}$  e, pode ser gerada por um registrador com  $n_p \in \mathbb{N} > 3$  estágios (Katayama, 2005). Outra característica é a quantidade mínima de iterações que a sequência permanece fixa em um mesmo nível. A escolha desses três parâmetros: níveis, período máximo e quantidade mínima de iterações fixa pode ser guiada conforme o valor final de **RMSE** retornado pela validação final do modelo de Hammerstein **MIMO GRDPI**. Como os exemplos serão desenvolvidos em um ambiente de simulação computacional hipotético, um valor fixo e elevado de  $n_p$  que controla a periodicidade do sinal será empregado. Seu valor será fixo em  $n_p = 10$  no restante deste capítulo, mas vale destacar que como a quantidade de estágios necessária no registrador sobe proporcionalmente ao período máximo, na prática, esse valor deve ser ponderado para melhorar o custo de implementação.

No contexto de identificação de sistemas lineares, duas recomendações comuns para projetar os sinais de teste para identificação, mesmo que empíricas, são que o processo deve ser excitado em uma faixa de operação reduzida e que o sinal deve ser capaz de extrair informação dinâmica do sistema. A primeira recomendação objetiva evitar excitar as possíveis não linearidades que se ressaltam à medida que o sinal de entrada excita diferentes pontos de operação, logo o quão pequena será essa variação é uma análise do projetista de acordo com o sistema. A segunda é que os sinais de entrada devem permitir que a resposta atinja cerca de 70% do que é o seu valor em regime permanente, de modo que não seja tão rápida que o sistema não tenha tempo para responder, nem tão lenta que o sistema atinja o regime permanente (Aguirre, 2015). Em processos cuja resposta passa pela origem pode-se adotar um dos níveis como sendo nulo, desse modo é como se a não linearidade estivesse sendo excitada apenas no ponto de operação não nulo. Geralmente a histerese não atende a condição de passagem pela origem, desse modo alguns exemplos visuais podem ajudar a entender esses e outros problemas.

As recomendações citadas podem ser estendidas sem complicações para os modelos de Hammerstien cuja parcela não linear seja estática, o que não é o caso do problema apresentado nesse capítulo, que possui a não linearidade quase estática. Para esclarecer essas afirmações, observe a Figura 4.7, na qual se apresenta uma curva não linear de terceiro grau  $v(k) = u(k)^3$ . Ao inserir um sinal **PRBS** na entrada dessa função estática, seja usando os pares de nível  $\{0, 2\}$  ou  $\{-2, 2\}$  que estão marcados pelos pontos vermelhos na Figura 4.7a), os sinais  $v(k)$  gerados respeitam exatamente os ganhos da função não linear desses pontos (ver Figura 4.7b) e c) e comparar as amplitudes com os pontos vermelhos da Figura 4.7a)).

O mesmo não ocorre quando a não linearidade é quase estática, que é o caso da histerese. Por exemplo, observe na Figura 4.8 que, diferentemente do teste anterior, os pontos de operação excitados  $\{0, 5\}$  ou  $\{-5, 5\}$  não recebem o ganho daquele nível que se observa na curva não linear (compare as amplitudes dos pontos vermelhos na Figura 4.8a) com as amplitudes atingidas nas Figuras 4.8b) e c)). Isso ocorre devido à memória que o modelo **GRDPI** possui. Dado esse obstáculo, podem ser aplicadas diferentes maneiras de compensar o ganho que difere o modelo obtido dos dados coletados do processo. Uma maneira direta é ajustar o ganho final manualmente com

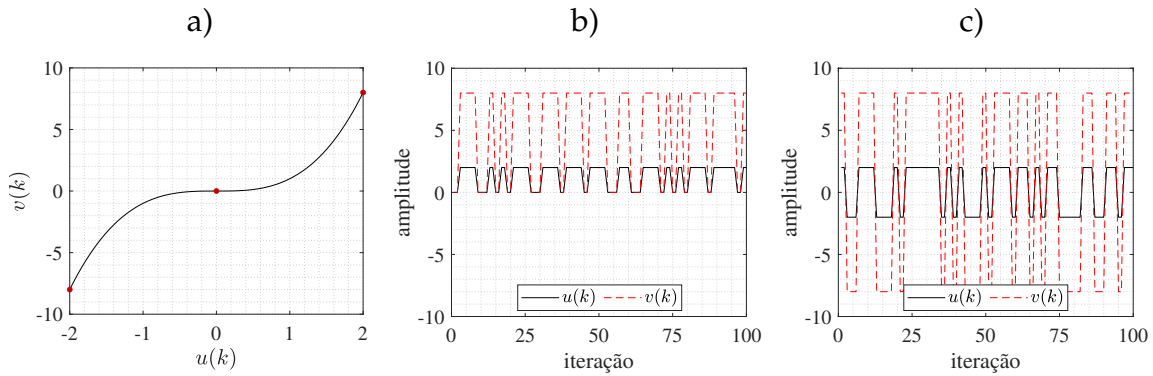


Figura 4.7: Saídas de uma função estática em resposta a um sinal PRBS. Em a) apresenta-se a função estática não linear que é cúbica, em b) apresenta-se a resposta à sequência PRBS com níveis  $\{0, 2\}$  e em c) com níveis  $\{-2, 2\}$ .

base em dados do processo.

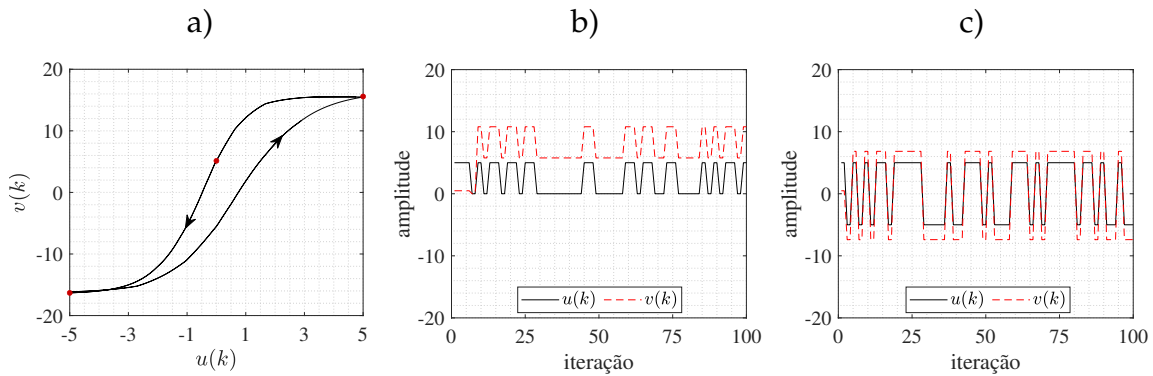


Figura 4.8: Saídas de uma função quase estática em resposta a uma entrada PRBS. Em a) apresenta-se a característica da função que é um laço de histerese, em b) a resposta a uma entrada PRBS com níveis  $\{0, 5\}$  e em c) a resposta para entradas com níveis  $\{-5, 5\}$ .

Outra abordagem será adotada a fim de obter resultados mais acurados com base em um processo de otimização. Ela consiste em aplicar dois ganhos em cada um dos sinais  $v_i(k)$  utilizados na identificação da parcela dinâmica linear, sendo um ganho na parte positiva do sinal e outro na parte negativa. No caso do sinal PRBS, isso se resume aos dois níveis do sinal binário, o  $+b_v$  e o  $-b_v$ . No caso do sinal aleatório, um dos ganhos é aplicado à toda parte positiva, independente da amplitude bem como o outro ganho é aplicado em toda parte negativa, independente da amplitude. Em qualquer um dos casos, os ganhos serão otimizados de modo que compensem a diferença entre o modelo e o processo. Para desenvolver o problema de maneira generalizada, admite-se que o  $i$ -ésimo sinal intermediário receba um ganho  $\underline{\theta}_i$  na parte negativa e  $\bar{\theta}_i$  na parte positiva. Os ganhos podem ser organizados em um vetor do seguinte modo

$$\underline{\Theta} = [\underline{\theta}_1, \bar{\theta}_1, \dots, \underline{\theta}_p, \bar{\theta}_p]. \quad (4.28)$$

Com tais ganhos gera-se uma versão escalonada  $\underline{\vec{v}}(k)$  do sinal intermediário  $\hat{\vec{v}}(k)$  para ser aplicada no MOESP.

O diagrama de implementação do MOESP é rerepresentado na Figura 4.9 considerando as variáveis para a formulação do problema específico da histerese. A entrada é composta pelos dados de ensaio dinâmico  $\{\underline{\vec{v}}(k), \tilde{\mathbf{y}}(k)\}_{k=1}^{N_L}$  e uma escolha de projeto  $h$ . Empregando a metodologia descrita na Subseção 2.5.1 pode-se fazer a identificação das matrizes  $\hat{A}$ ,  $\hat{B}$ ,  $\hat{C}$  e  $\hat{D}$ . Como não há aplicação da transformação de similaridade, note que o resultado do MOESP é denotado com o  $(\hat{\cdot})$  e não com o  $(\sim)$ .

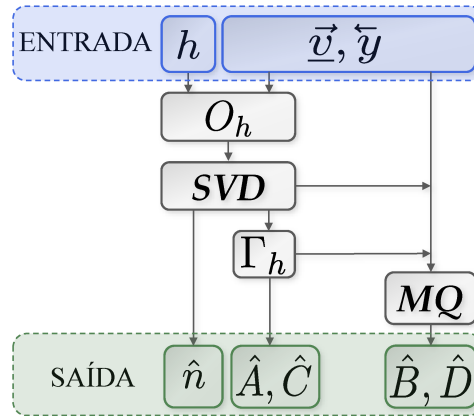


Figura 4.9: Esquema do algoritmo MOESP para identificar a parcela LIT do modelo de Hammerstein GRDPI a partir do conjunto de dados  $\{\underline{\vec{v}}(k), \tilde{\mathbf{y}}(k)\}_{k=1}^{N_L}$ .

Uma função custo deve ser definida para otimizar os ganhos intermediários  $\underline{\Theta}$  que dependa não somente da parte quase estática mas também da parte dinâmica para ajustar os ganhos em regime permanente da parcela LIT. Essa função é definida como

$$F(\underline{\Theta}) = \sum_{k=1}^{N_N} (\underline{\Psi}(k) - \hat{\underline{\Psi}}(k))^2, \quad (4.29)$$

na qual  $\underline{\Psi}$  é um sinal de saída afetado por  $\underline{\Theta}$  e pelas matrizes  $A$ ,  $B$ ,  $C$  e  $D$  e  $\hat{\underline{\Psi}}$  é o valor estimado pelo modelo de Hammerstein MIMO GRDPI. Vale destacar que, como no sistema real não há parâmetros  $\underline{\Theta}$ , então  $\underline{\Psi}$  é apenas uma saída medida. Um sinal diferente daquele utilizado na identificação das matrizes do modelo em espaço de estados deve ser empregado para computar essa função custo. Esse sinal deve excitar as características dinâmicas e quase estáticas do processo, por isso seleciona-se uma senoide  $\tilde{\mathbf{u}}(k)$  que excursionsse pela faixa mais ampla possível para calcular a função de aptidão. À vista disso, um terceiro ensaio deve ser realizado no processo a fim de coletar os dados de treinamento para guiar a otimização dos ganhos intermediários  $\{\tilde{\mathbf{u}}(k), \tilde{\mathbf{y}}(k)\}_{k=1}^{N_G}$ . Mais detalhes sobre esse teste serão especificados para cada um dos exemplos da Seção 4.5.

O problema de otimização pode ser enunciado como

$$\min_{\underline{\Theta}} F(\underline{\Theta}) = \sum_{k=1}^{N_N} (\underline{\hat{y}}(k) - \underline{\hat{y}}(k))^2, \quad (4.30)$$

sujeito a (4.6), (4.7), (4.8), (4.9),

$$\underline{v}_i(k) = \begin{cases} \bar{\theta}_i \hat{v}_i(k), & \text{se } \hat{v}_i(k) \geq 0, \\ \underline{\theta}_i \hat{v}_i(k), & \text{se } \hat{v}_i(k) < 0, \end{cases} \quad i = 1, \dots, p, \quad (4.31)$$

e a

$$x(k+1) = \hat{A}x(k) + \hat{B}\underline{\vec{v}}(k), \quad (4.32a)$$

$$\underline{\hat{y}}(k) = \hat{C}x(k) + \hat{D}\underline{\vec{v}}(k). \quad (4.32b)$$

Para solucionar esse segundo problema de otimização também será utilizado o **MPSO** proposto na Subseção 4.4.1. Os sinais empregados para o treinamento são multivariáveis, tal como ocorreu na parte não linear, no entanto, em lugar de gerar um sinal para cada entrada e aplicá-los separadamente no processo, os sinais são idênticos e aplicados simultaneamente em todas as entradas. Por fim, destaca-se que os parâmetros  $\underline{\Theta}$  do sinal **PRBS** e as matrizes  $\hat{A}$ ,  $\hat{B}$ ,  $\hat{C}$  e  $\hat{D}$  consequentes do **MOESP** serão ajustadas pelo **MPSO** de maneira a minimizar o valor da função custo (4.30).

Um diagrama que resume e ilustra o procedimento para obter uma estimativa inicial do modelo em espaço de estados  $\underline{\mathcal{L}}$  é apresentado na Figura 4.10a). Para simplificar utiliza-se o subíndice  $i$  para representar cada entrada, mas salienta-se que as operações são realizadas para todas as entradas e saídas simultaneamente. Em resumo, a entrada  $\underline{\vec{u}}(k)$ , que é um sinal aleatório ou **PRBS**, é aplicado no modelo  $\hat{N}$  e também no processo. Essa entrada resulta na estimativa do sinal  $\underline{\hat{v}}(k)$  e a saída  $\underline{\hat{y}}(k)$ , respectivamente. Em seguida, o **MPSO** irá prover os ganhos iniciais  $\underline{\Theta}$  e, a partir deles, os valores dos sinais reescalados  $\underline{\vec{v}}(k)$  poderão ser computados e utilizados no **MOESP**, na forma do conjunto de dados  $\{\underline{\vec{v}}(k), \underline{\hat{y}}(k)\}_{k=1}^{N_L}$ , para gerar uma estimativa provisória do sistema **LIT**  $\underline{\mathcal{L}}$ . De posse de  $\hat{N}$  e  $\underline{\mathcal{L}}$  pode-se aplicar o sinal senoidal  $\underline{\vec{u}}(k)$  no processo e no modelo de Hammerstein **GRDPI** e comparar as duas saídas por meio da função custo (4.30) para ajustar o **MPSO**. Supondo que o teste com o sinal **PRBS** ou aleatório tenha sido realizado, o procedimento iterativo que envolve o **MPSO** é esquematizado na Figura 4.10b). Observa-se que o **MPSO** ajusta os parâmetros  $\underline{\Theta}$  para estimar corretamente  $\hat{\mathcal{L}}$  (incluindo os ganhos estáticos) ao final do processo.

## 4.5 Exemplos Numéricos

Nesta seção, serão desenvolvidos dois exemplos numéricos hipotéticos. O primeiro deles, retratado na Subseção 4.5.1, tem o propósito de apresentar a metodologia, bem como discutir a respeito das limitações e a escolha de parâmetros. Ele é consti-

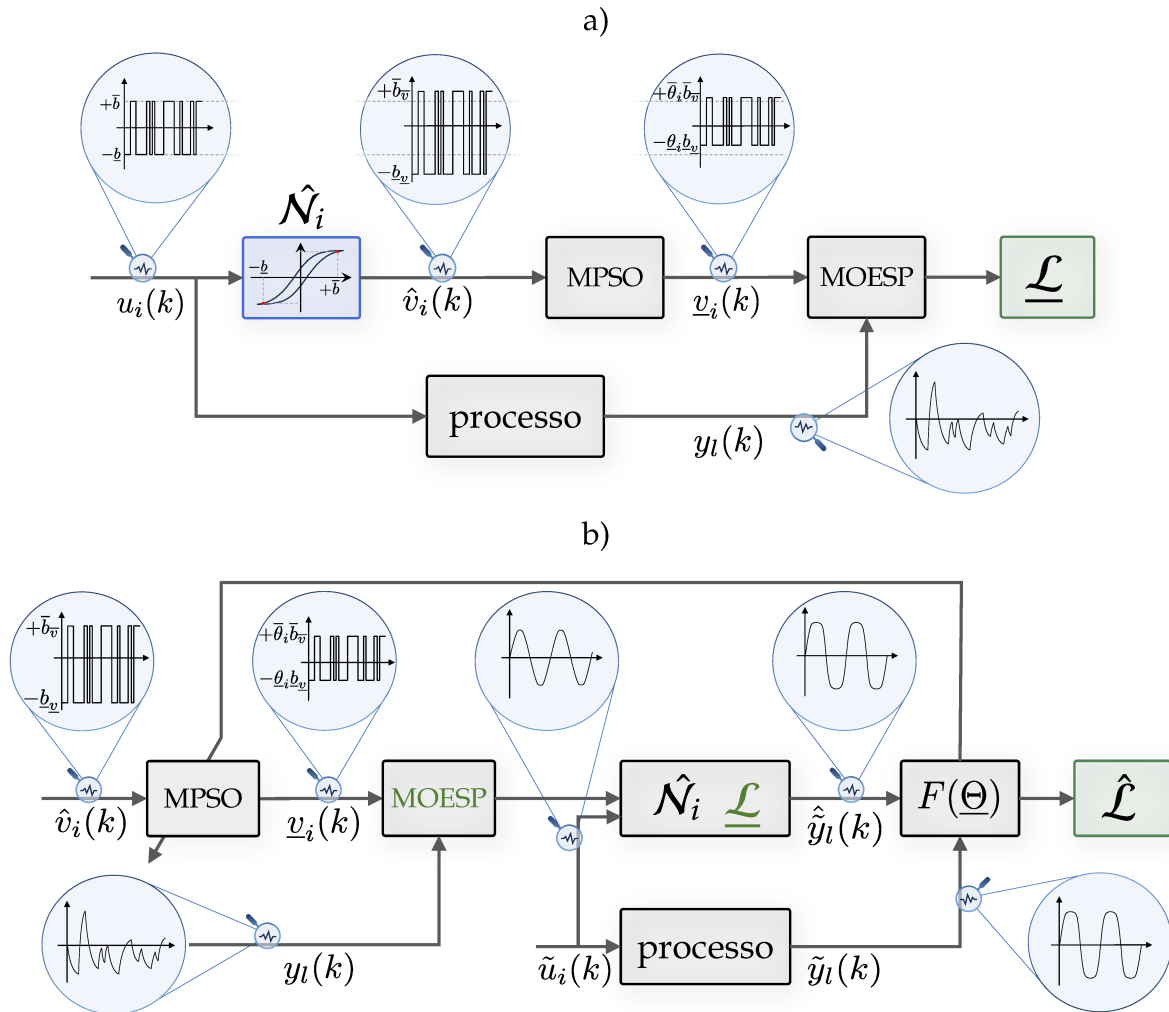


Figura 4.10: Diagramas do passo a passo a ser executado na segunda etapa da identificação dos modelos de Hammerstein GRDPI. Por simplicidade apresenta-se o diagrama com as variáveis em termos dos subíndices em lugar das vetoriais. Em a) apresentam-se os passos, bem como os sinais que estarão presentes durante a inicialização do primeiro modelo em espaço de estados  $\mathcal{L}$ . Por sua vez, na parte b) apresenta-se o procedimento iterativo feito no MPSO para ajustar os ganhos  $\Theta$  guiado pela função custo (4.30).

tuído por 3 operadores *play* para representar uma histerese simétrica. A parte linear é composta por um modelo em espaço de estados MIMO desacoplado e sem ruído. Por sua vez, o segundo exemplo apresenta um cenário mais desafiador que o anterior, com 10 operadores *play*, histerese assimétrica, modelo em espaço de estados MIMO acoplado, de fase não mínima e com a presença de ruído branco de medição cuja SNR é de 25 dB. Desse modo, a finalidade do Exemplo 2, apresentado na Subseção 4.5.2 é verificar a funcionalidade da metodologia sob as circunstâncias para as quais o modelo foi projetado.

### 4.5.1 Exemplo 1

Considere um sistema hipotético, multivariável e não linear formulado sob a estrutura de um processo de Hammerstein em tempo discreto, com período de amostragem  $T_s = 0,001$ . Assuma que a parcela não linear é dada por um modelo GRDPI com a estrutura apresentada em (4.10) e parâmetros

$$\Theta = \left[ \begin{array}{c|cc|cc|cccc|cc} 0,1 & & 0,9 & 0,4 & 0,8 & 0,9 & 0,6 & 0,9 & 0,6 & 1 & 1 \\ 0 & & 0,7 & 0,2 & 0,5 & 0,9 & 0,6 & 0,9 & 0,6 & 1 & 1 \end{array} \right]', \quad (4.33)$$

e a parcela linear é descrita pelas matrizes do sistema em espaço de estados (4.32a) e (4.32b) desacoplado e com autovalores 0,7 e 0,9

$$A = \begin{bmatrix} 0,9 & 0 \\ 0 & 0,7 \end{bmatrix}, \quad B = \begin{bmatrix} 0,4 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{e} \quad C = \begin{bmatrix} 0,5 & 0 \\ 0 & 0,6 \end{bmatrix}. \quad (4.34)$$

Por fim, assumo também que a entrada desse processo seja limitada entre -5 e 5 e a frequência de trabalho seja de 10 Hz.

Como apontado no Problema 3 a metodologia de identificação é dividida em duas etapas. Na primeira etapa o objetivo é identificar uma versão escalonada da histerese. Para isso, cada entrada do processo é submetida a uma sequência de degraus ascendentes e descendentes, cuja ordem relativa não importa, ou seja, pode ser primeiro crescente e depois decrescente ou o inverso. A amplitude total que o sinal de teste excursionará deve ser a maior suportada pela respectiva entrada e o tempo fixo em cada nível deve ser suficiente para captar a resposta em regime. Em contrapartida, a quantidade de degraus é uma escolha empírica do projetista, mas influencia diretamente o custo computacional e a acurácia do resultado. Para essas escolhas pode-se variar os valores dos parâmetros do sinal de teste manualmente ou por meio de algum algoritmo e escolher os melhores com base na função custo (4.29).

Fundamentado nesses direcionamentos e outros discutidos na Seção 4.4.1, as amplitudes dos sinais de entrada são variadas de -5 a 5, faixa em que se escolheu gerar  $N_N = 50$  degraus (ascendentes e descendentes) e permanecer 100 iterações em cada, como pode ser verificado na Figura 4.11. Como nesse exemplo não há ruído, em contraste com o que foi proposto na Subseção 4.4.1 pode-se coletar, sim, apenas o último valor da resposta de cada degrau em lugar de ter que fazer a média dos últimos dados.

Na primeira metade do teste apenas a entrada  $u_1$  está ativa e o oposto ocorre na segunda metade, em que apenas  $u_2$  está ativa. As primeiras amostras de cada uma das metades é descartada para eliminar os transientes. Como o sistema é desacoplado, é possível notar que a saída  $y_1$  responde apenas à entrada  $u_1$  e a saída  $y_2$  responde à entrada  $u_2$ . Ao final do ensaio armazena-se os dados em regime permanente da entrada ativa e da saída que excursiona pela maior faixa. Particularmente neste exemplo, coleta-se os conjuntos  $\{\bar{u}_1, \bar{y}_1\}$  que contém informações sobre  $\mathcal{N}_1$  e  $\{\bar{u}_2, \bar{y}_2\}$  que contém informação a respeito de  $\mathcal{N}_2$ . As barras sobre  $\bar{u}_1$  e  $\bar{u}_2$  indicam que os pontos da entrada

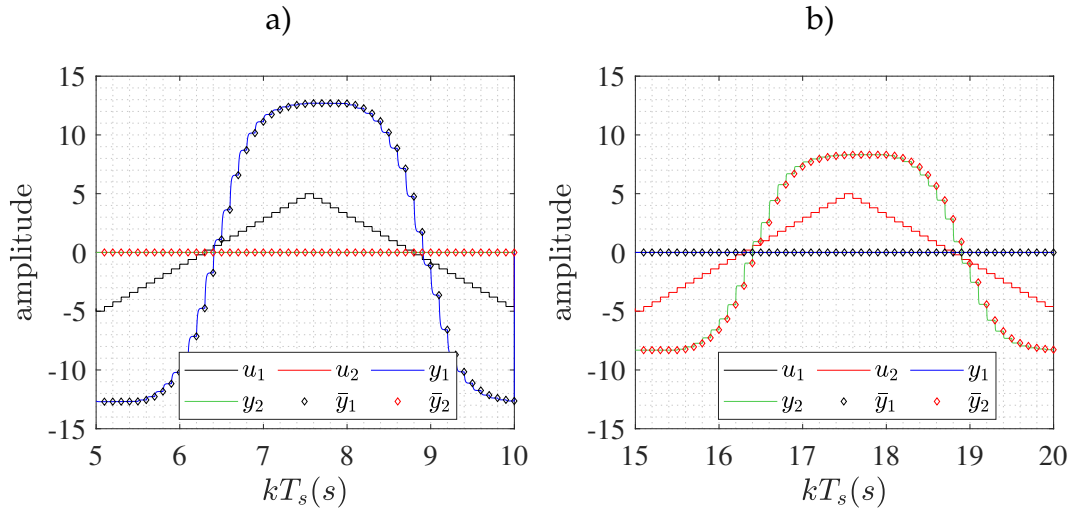


Figura 4.11: Dados do ensaio estático coletados do processo composto por (4.33) e (4.34). Para eliminar o transiente gerado pelo modelo GRDPI, os dados iniciais de cada teste são descartados. Em a) apresenta-se a excitação  $u_1$  e a respectiva resposta e em b) os mesmos sinais, porém, no momento em que a entrada  $u_2$  que está ativa.

são coletados em regime permanente. As saídas  $\bar{y}_1$  e  $\bar{y}_2$  são coletadas de forma síncrona com as entradas. Vale destacar que as especificações de amplitudes das entradas e o tempo mínimo que cada degrau permanece fixo não precisam ser iguais para as diferentes entradas, apesar de assim o serem neste exemplo.

Os conjuntos de dados coletados  $\{\bar{u}_1, \bar{y}_1\}$  e  $\{\bar{u}_2, \bar{y}_2\}$  são aplicados simultaneamente no algoritmo MPSO descrito na Subseção 4.4.1 para identificar um modelo GRDPI que aproxime a histerese combinada com o ganho em regime permanente da parcela linear. Com o propósito de favorecer a estimativa de parâmetros próximos aos originais, algumas escolhas foram realizadas tirando proveito do conhecimento dos parâmetros do processo. Foram elas,  $N_p = 3$  e  $\Theta_{\min}$  como uma matriz nula. Os demais parâmetros foram ajustados manualmente conforme a métrica de validação adotada que é o índice RMSE normalizado. São eles,  $V_{\max} \in \mathbb{R}^{2 \times 10}$  uma matriz preenchida com valores 0,8,  $\xi = 4$ ,  $\alpha = 0,6$ ,  $K_{\max} = 200$ ,  $P = 50$ ,  $K_\alpha = 10$ ,  $\omega_1 = \omega_2 = 2,05$  e

$$\Theta_{\max} = \begin{bmatrix} 3 & 3 & 3 & 3 & 1 & 1 & 1 & 1 & 1 & 1 \\ 3 & 3 & 3 & 3 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \quad (4.35)$$

O MPSO pode gerar resultados diferentes em cada execução a depender das condições iniciais aleatórias e também por haver fluxos a serem seguidos no algoritmo que dependem de valores não determinísticos, como o  $v$ ,  $\zeta$  entre outros. Desse modo, são feitas 100 realizações do MPSO, mantendo fixos todos os parâmetros descritos no parágrafo anterior. Como resultado obtém-se 100 modelos distintos cujos valores estimados de  $\hat{\Theta}$  podem ser observados na Figura 4.12. A respeito desses valores a primeira análise de interesse é com relação à consistência. Observando a Figura 4.12a) nota-se que há maior consistência nas estimativas das posições de 5 a 10, que pertencem

à função de borda e aos disparos dinâmicos. Isso ocorre devido à limitação feita por (4.35) que é mais restrita para esses parâmetros, mas também se deve a característica de maior sensibilidade desses parâmetros, de modo que menores desvios neles causam maior variação na saída.

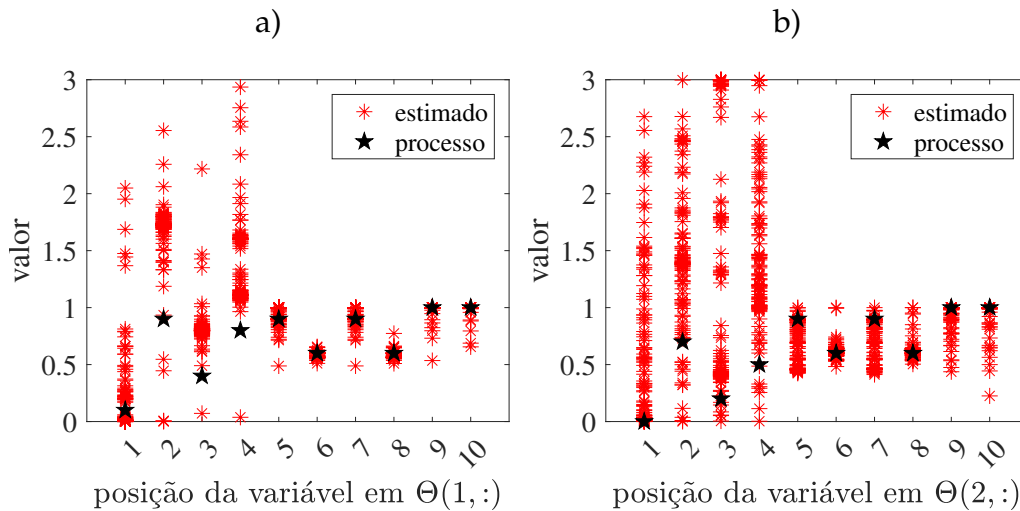


Figura 4.12: Parâmetros dos modelos GRDPI estimados para a primeira entrada em a) e para a segunda em b). Os parâmetros de 1 a 4 se referem ao modelo P-I, de 5 a 8 à função de borda do operador generalizado e 9 e 10 aos disparos dinâmicos.

Para os demais parâmetros ainda há consistência na estimativa, no entanto, ficaram mais distantes dos parâmetros originais. Isso era esperado, uma vez que, além de compensar o ganho da histerese, o modelo GRDPI absorve o ganho estático em regime permanente da parcela dinâmica linear. Portanto, os parâmetros são estimados, mas estão a um ganho ou a alguma transformação desconhecida dos parâmetros originais. Isso pode ser observado com mais clareza na Figura 4.12b), em que há menos consistência e maior distância dos parâmetros originais, mas a saída do modelo de Hammerstein completa que será apresentada mais adiante é consistente e próxima da original.

Finalizada a primeira etapa, procede-se para a segunda, que se traduz em estimar a parcela dinâmica linear  $\mathcal{L}$ . A princípio, dois sinais são gerados para fazer diferentes testes no processo. Um deles é um sinal PRBS, utilizado no MOESP para identificação das matrizes do modelo em espaço de estados. O outro é um sinal senoidal, de carga e descarga, utilizado para computar a função custo no MPSO e assim otimizar os ganhos intermediários (4.28). Dois parâmetros devem ser escolhidos para caracterizar o sinal PRBS, a amplitude e a quantidade de iterações mínimas que cada patamar permanece inalterado. Por sua vez, a senoide precisa da escolha de sua amplitude e frequência.

A fim de identificar corretamente a não linearidade na faixa em que o problema está definido, as amplitudes de cada um dos dois ensaios são escolhidas como uma variação de -5 a 5. Guiado pelo resultado da função custo (4.30), a quantidade mínima

de interações que cada patamar do PRBS permanece inalterado foi selecionada como 1 iteração. Por sua vez, a frequência da senoide deve ser decidida dentro de uma faixa operacional do processo e que não destaque somente a característica estática, ou seja, que apresente dinâmica. Tomando essas direções e sabendo que a frequência de operação é de 10 Hz, a frequência decidida para o teste foi de 5 Hz, assim reserva-se a frequência de 10 Hz para a validação.

O sinal PRBS monovariável é gerado e aplicado em todas as entradas do processo simultaneamente. Além disso, esse mesmo sinal é inserido na entrada do modelo GRDPI. As saídas dos dois ensaios são coletadas e usadas na estimativa da parcela dinâmica linear, como foi indicado na Figura 4.10. Como discutido anteriormente, o sinal de saída do modelo GRDPI é uma versão escalonada do sinal intermediário, portanto, como as entradas utilizadas no modelo e processo são as mesmas, pode-se utilizar o sinal de saída do modelo GRDPI em conjunto com a saída do processo para fazer a identificação da parcela dinâmica linear. Para ajustar as amplitudes de saída, são adicionados ganhos  $\Theta$  estimados por meio do MPSO.

Para a estimativa do modelo em espaço de estados, o projetista precisa decidir pela variável  $h$ . No exemplo do qual se trata essa subseção  $h$  é escolhido como 2, com base nos valores da função custo (4.30). Quanto ao MPSO, seguindo os mesmos direcionamentos do primeiro MPSO implementado nesse mesmo exemplo, os parâmetros escolhidos foram  $N_V = 2$ ,  $\Theta_{\min}$  como uma matriz nula,  $\Theta_{\max}$  como uma matriz preenchida com valores 2,  $V_{\max}$  como uma matriz preenchida com valores 0,2,  $\xi = 4$ ,  $\alpha = 0,6$ ,  $K_{\max} = 30$ ,  $P = 40$ ,  $K_\alpha = 10$  e  $\omega_1 = \omega_2 = 2,05$ .

Os autovalores do modelo em espaço de estados identificado ficaram próximos aos originais, com média 0,9 e 0,7 e desvio padrão de aproximadamente  $2 \times 10^{-5}$  e  $6 \times 10^{-5}$ , respectivamente. Graficamente, a validação do modelo de Hammerstein GRDPI completo para uma mesma entrada senoidal de amplitude 5 e frequência de 10 Hz aplicadas simultaneamente nas duas entradas podem ser observados na Figura 4.13. O contorno da região sombreada foi obtida pelo laço de histerese dos piores modelos, apontando que todas as estimativas ficaram dentro dessa região. Em média, os valores de RMSE normalizados pela amplitude de teste, 5, foram obtidos como 0,0137 e 0,0184 para as saídas 1 e 2 respectivamente. Portanto, dada a característica do exemplo de ser desacoplado, sem ruído e com apenas 3 parâmetros, foi possível constatar a capacidade que o modelo de Hammerstein GRDPI tem de representação em casos básicos. O próximo exemplo será desenvolvido em um ambiente com desafios adicionais, tais como acoplamento dinâmico e a presença de ruído de medição.

## 4.5.2 Exemplo 2

Considere um segundo processo multivariável, também idealizado na estrutura de Hammerstein, cuja parcela não linear é uma histerese definida pelo modelo

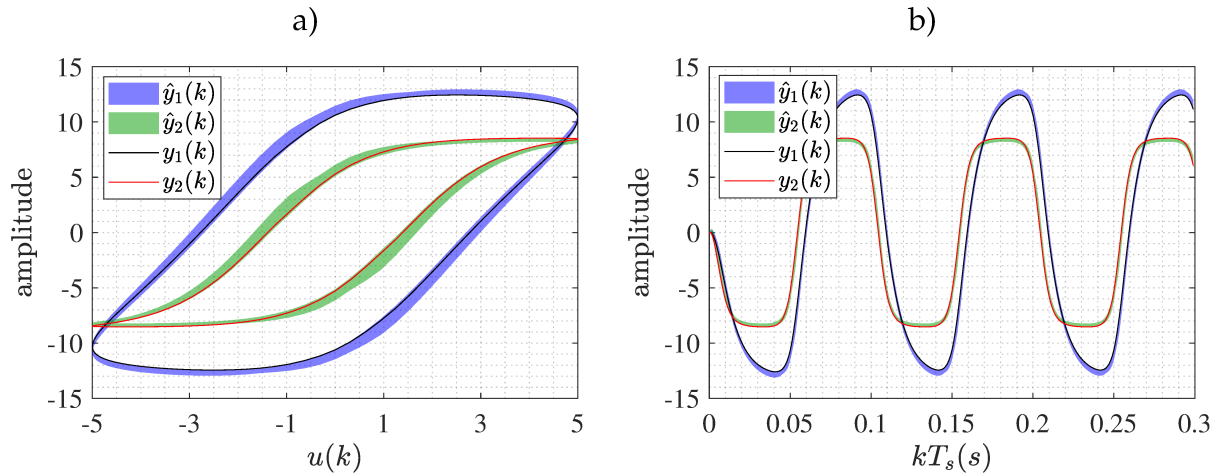


Figura 4.13: Nuvem de respostas para validação dos 100 modelos de Hammerstein GRDPI comparados aos sinais do processo composto por (4.33) e (4.34). Em a) encontra-se a validação do laço de histerese com influência da dinâmica, que dá o aspecto mais suave aos laços. Em b) exibe-se a saída no tempo. Nos dois gráficos apresentam-se em preto e vermelho os sinais de saída do processo para facilitar a comparação.

GRDPI com  $N_p = 10$  operadores *play*

$$\Theta = \left[ \begin{array}{c|cccccccccccc|cccc|cc} 0 & 0,2 & 0,6 & 0 & 0,4 & 0,2 & 0,1 & 0,3 & 0,4 & 0,2 & 0,8 & 0,9 & 0,25 & 0,8 & 0,3 & 1 & 1 \\ \hline 0,1 & 1 & 0,9 & 0,1 & 0,7 & 0,1 & 0,8 & 0,4 & 0,1 & 0,5 & 0,6 & 1 & 0,6 & 1 & 0,4 & 1 & 1 \\ \hline \end{array} \right], \quad (4.36)$$

e a parcela linear é representada pelo seguinte sistema dinâmico acoplado

$$A = \begin{bmatrix} 0,8 & 0,3 \\ 0,1 & 0,7 \end{bmatrix}, \quad B = \begin{bmatrix} 0,6 & 0,1 \\ 0,2 & 0,7 \end{bmatrix}, \quad C = \begin{bmatrix} 0,4 & 0,2 \\ 0 & 0,3 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 0,4 & 0 \\ 0 & 0,1 \end{bmatrix}, \quad (4.37)$$

com autovalores 0,9303 e 0,5697 e que é de fase não mínima. Os zeros para cada um dos pares de entrada e saída do modelo (4.37) são  $(0,6 \text{ e } 0,2)_{1,1}$ ,  $(0,5)_{1,2}$ ,  $(0,3)_{2,1}$ ,  $(-1,4 \text{ e } 0,8)_{2,2}$ , sendo os subscritos indicadores das respectivas entradas e saídas. O zero com módulo maior que 1 indica que esse processo é de fase não mínima, o que causaria problemas em inversões da dinâmica identificada (Verhaegen e Westwick, 1996; Aljanaideh et al., 2022). Como no Exemplo 1, assuma que a entrada seja limitada entre -5 e 5 e a frequência de trabalho seja de 10 Hz. Além disso, assuma que há um ruído de medição que leva à SNR de 25 dB corrompendo cada uma das saídas. O modelo em espaço de estados (4.37) é definido em tempo discreto, cujo período de amostragem é  $T_s = 0,001$ . Essa informação é útil para plotar os gráficos com o tempo correto e, apenas isso, uma vez que no disparo dinâmico a dependência com o período de amostragem foi removida (ver (4.7) e seu contexto).

O procedimento é dividido em duas etapas seguindo os mesmos direcionamentos apresentados na Seção 4.4 e descritos com maior riqueza de detalhes na Subseção 4.5.1. Na primeira etapa, os dados estáticos, que serão utilizados para implementar o algo-

ritmo **MPSO** são gerados por um ensaio em que se aplica 50 degraus com amplitudes variando de -5 a 5 nas duas entradas, uma por vez, como pode ser observado na Figura 4.14. Uma diferença para a implementação apresentada na Figura 4.11 é que há ruído de medição afetando as saídas, por isso a média dos dados em regime permanente deve ser aplicada. Em tal caso, para ter dados suficientes para implementar a média, os níveis dos degraus são variados a cada 500 iterações, sendo as 100 primeiras amostras consideradas transientes e as últimas 400, como saída em regime permanente sobre as quais será aplicada a média. Tanto na Figura 4.14a) quanto na Figura 4.14b) nota-se que o sinal que excursiona pela maior faixa é a saída  $y_1(k)$ . Dito isso, ela é escolhida para compor o conjunto de dados devido ao padrão adotado, informado na Subseção 4.4.1. Concluída a coleta e pré-processamento, as medidas do sinal de entrada  $\vec{u}(k)$  e a média da saída em regime permanente  $\bar{y}(k)$  podem ser utilizadas para ajustar os parâmetros do modelo **GRDPI** por meio do **MPSO**. Note que a quantidade de dados em regime permanente necessária para que o erro da média computada mantenha-se inferior ao passo dos degraus é diretamente proporcional ao nível de ruído que afeta o processo. Observando a Figura 4.14 é possível verificar que os 400 dados coletados em regime permanente estão próximos do limite, uma vez que nas partes planas, percebe-se uma flutuação dos losangos que representam as médias.

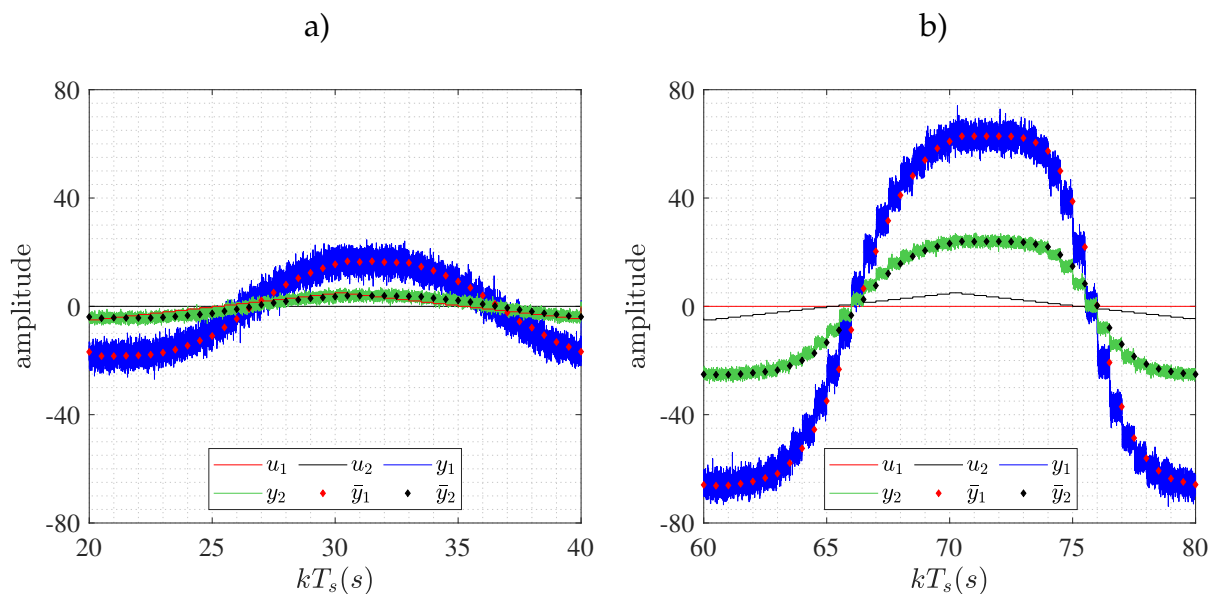


Figura 4.14: Dados de entrada e saída do ensaio estático empregado para identificar o modelo **GRDPI** no Exemplo 2. Em a) apresenta-se a excitação em degraus ascendentes e descendentes e a respectiva resposta à entrada  $u_1$  e em b) com relação à  $u_2$ . O transiente que a histerese insere foi ignorado, por isso, os dados iniciais em cada ensaio foram descartados.

Os parâmetros do **MPSO** foram selecionados com base no método de tentativa e erro segundo a melhoria da função custo obtida ao final da otimização. Ao final do processo o **MPSO** foi configurado com os parâmetros  $N_p = 13$ ,  $K_{\max} = 200$ ,  $P = 30$ ,  $\xi = 3$ ,  $K_\alpha = 10$ ,  $\alpha = 0,8$  e  $\omega_1 = \omega_2 = 2,05$ ,  $V_{\max}$  como uma matriz preenchida com o valor

0,1,  $\Theta_{\min}$  como uma matriz nula e  $\Theta_{\max}$  como uma matriz preenchida com o valor 2.

Aproveitando as possibilidades do ambiente de simulação foram feitas 1000 realizações do processo, de modo que foram obtidos 1000 modelos diferentes cujos parâmetros resultantes são apresentados na Figura 4.15a). Diferente do que foi feito no Exemplo 1, nenhuma informação *a priori* a respeito dos parâmetros foi utilizada, de modo que, orientar-se pelo valor da função custo (4.30) levou a uma quantidade de parâmetros diferente da original. Assim sendo, não há como comparar se todos os parâmetros do modelo GRDPI foram estimados próximos àqueles do processo. No entanto, ainda é possível comparar os parâmetros de (4.6) e (4.7). Como ocorreu no exemplo anterior, nota-se maior consistência nas estimativas dos parâmetros de (4.6), sobretudo nos que são argumento da função tangente hiperbólica. O motivo é o mesmo: a sensibilidade desses parâmetros a pequenas variações em seus valores. Além disso, eles não são responsáveis diretamente por absorver o ganho da parcela dinâmica linear, como é o caso dos parâmetros de 1 a 14.

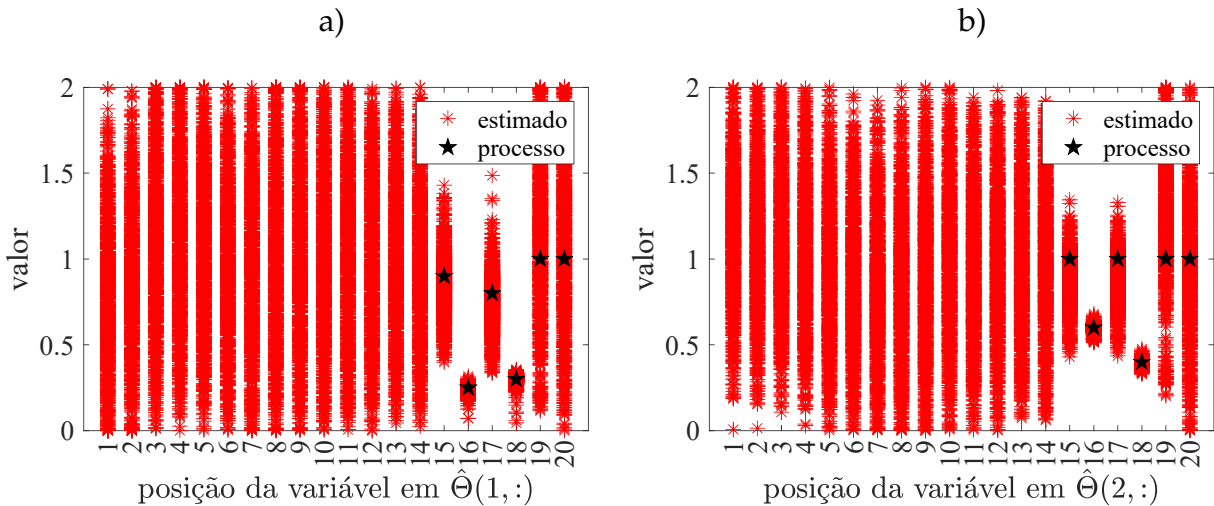


Figura 4.15: Todos os parâmetros dos 1000 modelos GRDPI estimados para a primeira entrada em a) e para a segunda em b). Como o número de parâmetros do processo e do modelo ficaram diferentes, mantém-se a comparação apenas da parcela relacionada às funções de borda e aos disparos dinâmicos.

Na segunda etapa, um sinal aleatório com amplitudes variando no máximo entre -5 a 5 é aplicado em todas entradas do modelo GRDPI simultaneamente. Além disso, esse mesmo sinal é aplicado também em todas as entradas do processo. Ao final da execução coleta-se a estimativa dos sinais intermediários  $\hat{\vec{v}}(k)$  e as saídas do processo  $\hat{\vec{y}}(k)$ . O sinal estimado,  $\hat{\vec{v}}(k)$ , é escalonado pelos parâmetros  $\underline{\Theta}$  inicializados pelo MPSO e gera  $\vec{v}(k)$  de modo que o MOESP, pode ser executado sobre o conjunto de dados  $\{\vec{v}(k), \hat{\vec{y}}(k)\}_{k=1}^{N_L}$ . O último ensaio é projetado com um sinal de entrada senoidal  $\tilde{u}(k)$  cujas amplitudes variam de -5 a 5 e tem frequência de 5 Hz. Esse conjunto de dados é útil para computar a função custo do MPSO (4.30). Por fim, empregam-se as iterações do MPSO descrito na Subseção 4.4.1 para refinar os ganhos  $\underline{\Theta}$  que estimam o modelo em espaço de estados  $\hat{\mathcal{L}}$  com os ganhos corretos a partir de  $\vec{v}(k)$  e  $\hat{\vec{y}}(k)$  e de  $\tilde{u}(k)$  e  $\tilde{\vec{y}}(k)$ .

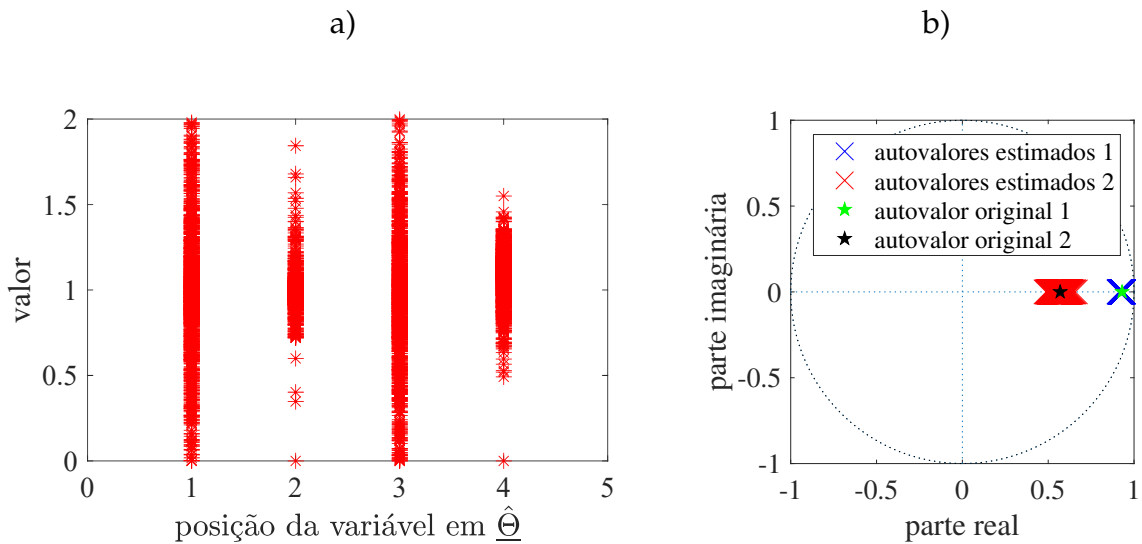


Figura 4.16: Dados de validação da parcela dinâmica linear do Exemplo 2. Em a) apresenta-se os ganhos resultantes do MPSO que foram aplicados aos sinais intermediários para identificar as matrizes em espaço de estados. Em b) comparam-se os autovalores estimados com os originais do processo.

Para identificar o modelo em espaço de estados foi utilizado o algoritmo MOESP, cujas matrizes de Hankel foram compostas por  $h = 2$  blocos. Os modelos estimados apresentaram autovalores consistentes como pode ser observado na Figura 4.16b), com média 0,9293 e 0,5626 e desvio padrão 0,0047 e 0,0313, respectivamente. Nota-se maior variabilidade no menor autovalor, pois este tem menor influência dinâmica na resposta. Os parâmetros utilizados no MPSO foram  $\Theta_{\min}$  como uma matriz nula,  $\Theta_{\max}$  uma matriz preenchida com o valor 2,  $V_{\max}$  uma matriz preenchida com 0,2,  $K_{\max} = 40$ ,  $P = 20$ ,  $\xi = 2$ ,  $K_{\alpha} = 10$ ,  $\alpha = 0,8$  e  $\omega_1 = \omega_2 = 2,05$ . Os parâmetros resultantes estão apresentados na Figura 4.16a) em que se observa consistência semelhante em todas as posições, uma vez que todos os parâmetros têm a mesma função, a de dar o ganho no sinal intermediário.

A validação é feita empregando um sinal senoidal de amplitude variando de -5 a 5 e frequência de 10 Hz distintos dos dados de identificação. O resultado é apresentado graficamente na Figura 4.17. Diversamente ao Exemplo 1 apresenta-se todos os modelos na Figura 4.17 ao invés da nuvem de respostas com o contorno dos piores modelos, assim, é possível notar que a faixa fica mais densa onde o modelo tem maiores chances de ser identificado. Nota-se que, em geral, os piores casos estão isolados, por exemplo, na Figura 4.17a) quando o laço está próximo de  $u(k) = +4$  com derivada positiva há apenas 1 modelo discrepante dos outros 999, que tem comportamento mais próximo ao original. Observa-se que os sinais dos modelos, ficaram próximos aos sinais do processo, representados pelas linhas contínuas. Isso ocorre nas relações quase estáticas representado pelo laço na Figura 4.17a) e nas relações dinâmicas, que é a resposta no tempo da Figura 4.17b). O valor do índice RMSE calculado para a saída no tempo foi de 0,0093 e 0,0123 para as saídas 1 e 2, respectivamente.

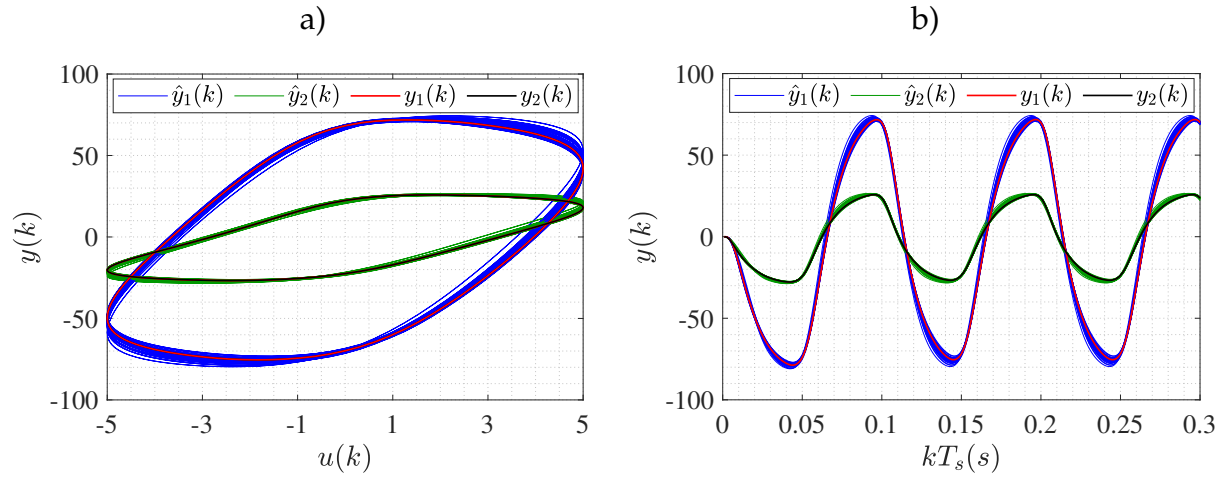


Figura 4.17: Validação dos 1000 modelos de Hammerstein GRDPI obtidos para identificar o processo composto por (4.36) e (4.37). Em a) encontra-se a validação dos laços de histerese e em b) a validação da saída no tempo. Em ambos gráficos a resposta é comparada com a saída do processo em vermelho e preto. Os dados são distintos dos dados de identificação.

Um teste foi implementado a fim de verificar a efetividade do modelo em outras regiões. O experimento consiste em, primeiro, selecionar a melhor estimativa, sob o ponto de vista da função custo, dentre todos os 1000 modelos gerados durante a aplicação da metodologia desenvolvida no Exemplo 2. Após essa seleção o modelo é submetido, junto com o processo, a uma entrada cuja amplitude varia de 1 a 5 enquanto a frequência também varia de 1 a 5 Hz. Destaca-se que não há necessidade de variar os dois parâmetros em conjunto, no entanto, seguir com a frequência além de 5 Hz tornaria o gráfico mais poluído para analisar. Com esse ensaio pode-se observar o comportamento do modelo em diversas condições diferentes daquela na qual foi feita a identificação.

O melhor modelo obtido durante as simulações tem a parcela não linear descrita pelos parâmetros a seguir

$$\hat{\Theta} = \begin{bmatrix} 0,37 & 0,71 & 0,92 & 0,88 & 0,82 & 1,14 & 0,77 & 0 & 0,91 & 1,01 & 1,31 & 0,3 \\ 0,95 & 0,85 & 1,17 & 1,5 & 0,17 & 0,72 & 0,64 & 1,17 & 0,66 & 1,55 & 1,13 & 1,56 \\ & & & & 0,71 & 0,53 & 1,04 & 0,24 & 0,93 & 0,28 & 1,28 & 0,71 \\ & & & & 1,02 & 0,39 & 0,99 & 0,6 & 0,99 & 0,41 & 1,49 & 1,25 \end{bmatrix}, \quad (4.38)$$

e a parcela dinâmica linear dada pelo modelo em espaço de estados de segunda ordem

$$\begin{aligned} \hat{A} &= \begin{bmatrix} 0,9438 & 0,0638 \\ -0,0940 & 0,5292 \end{bmatrix}, & \hat{B} &= \begin{bmatrix} -0,0051 & -0,0107 \\ 0,0015 & -0,0178 \end{bmatrix}, \\ \hat{C} &= \begin{bmatrix} -13,4884 & 2,1797 \\ -4,8003 & -4,5817 \end{bmatrix} & \text{e } \hat{D} &= \begin{bmatrix} 0,0744 & 0,0007 \\ 0,0071 & 0,0589 \end{bmatrix}, \end{aligned} \quad (4.39)$$

com autovalores 0,9288 e 0,5443 e que é de fase não mínima. Os zeros para cada um

dos pares de entrada e saída do modelo (4.37) são  $(0,56 \text{ e } -0,04)_{1,1}$ ,  $(-1,52 \text{ e } 0,51)_{1,2}$ ,  $(-147,26 \text{ e } 0,21)_{2,1}$ , e  $(-1,57 \text{ e } 0,79)_{2,2}$ , sendo os subscritos indicadores das respectivas entradas e saídas. Com isso é possível notar que, para esse melhor modelo, os autovalores foram identificados próximos aos valores originais, diferente dos zeros. O resultado do teste é apresentado na Figura 4.18. Pode-se observar no gráfico a) que, de maneira geral, o modelo acompanha o processo mesmo com as variações de amplitude e frequência impostas pelo teste. O resultado pode ser observado com mais clareza na Figura 4.18b) em que se apresenta a resposta do processo e do modelo no tempo. Um efeito visível nesse gráfico b) é que à medida que a frequência aumenta, a amplitude decai na faixa testada.

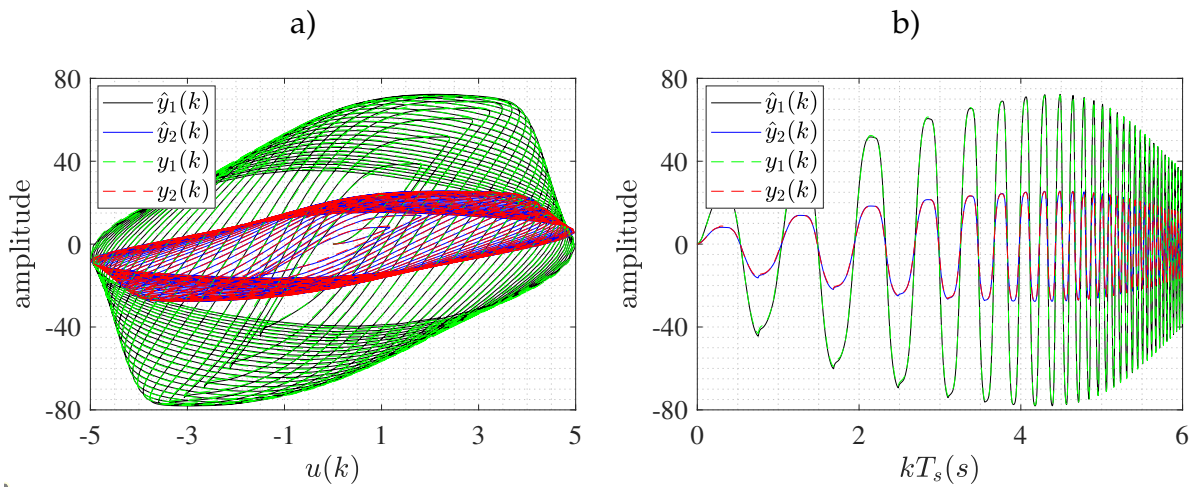


Figura 4.18: Validação do melhor modelo, composto por (4.38) e (4.39), dentre os 1000 modelos obtidos na identificação do modelo de Hammerstein GRDPI do Exemplo 2. Os testes são feitos com a amplitude e frequência do sinal de entrada variando juntas de 1 a 5. Em a) apresenta-se a comparação entre o laço de histerese do modelo e do processo. Os sinais de saída no tempo que geram esses laços são apresentados no gráfico b).

## 4.6 Conclusão

Nesse capítulo uma metodologia em duas etapas foi apresentada para fazer a identificação de sistemas multivariáveis que combinam histerese, que é um fenômeno não linear quase estático, com uma parcela dinâmica linear. O procedimento é direcionado a sistemas que possam ser representados por modelos de Hammerstein, ou seja, cuja não linearidade esteja na entrada. Diferentemente de outros trabalhos encontrados na literatura, como (Aljanaideh et al., 2018, 2022; Al Janaideh et al., 2023), primeiro estima-se  $\mathcal{N}$  para depois estimar  $\mathcal{L}$ . Assim, evitam-se restrições adicionais, mesmo quando o sistema a ser identificado é de fase não mínima. Para fazer a identificação propôs-se utilizar a estrutura de blocos interconectados de Hammerstein, composto por um modelo não linear GRDPI seguido por outro dinâmico linear em espaço de estados.

O modelo **GRDPI** foi capaz de captar as informações quase estáticas do processo por meio do **PSO** modificado. As modificações se referem ao emprego de mutações quando a solução deixa de evoluir por determinadas iterações e ao compartilhamento efetivo de informação entre as partículas da nuvem. A segunda parcela que compõe o modelo de Hammerstein é um sistema **MIMO** em espaço de estados, que também foi capaz de identificar com sucesso os autovalores do processo. Os exemplos numéricos permitiram verificar a eficácia do método em cenários desafiadores como histerese assimétrica com até 10 operadores *play* e sistemas com ruído de medição, multivariáveis, acoplados e de fase não mínima. Em todos os casos apresentados o índice **RMSE** normalizado foi inferior a 0,02 para qualquer saída. Uma limitação da metodologia é a necessidade de realizar três ensaios diferentes constituídos de uma sequência de degraus, sinais **PRBS** e senoidais usados para identificar a histerese, a dinâmica linear e para ajustar os ganhos do modelo em espaço de estados.

---

# Conclusões e Propostas para Trabalhos Futuros

---

## 5.1 Considerações Finais

As pesquisas desenvolvidas no decorrer deste trabalho surgiram como subproblemas da questão de estimar os parâmetros de um modelo de Hammerstein **MIMO** com não linearidades fortes. Um dos desdobramentos é o procedimento de identificação para não linearidades estáticas e o outro para não linearidades quase estáticas, mais especificamente a histerese. Como o modelo de Hammerstein pertence à classe dos **MBIs**, ele herda, dentre outras, a característica de que o sinal intermediário não corresponde a um sinal necessariamente acessível ou mesmo existente. Além disso, o cenário **MIMO**, no qual os procedimentos foram desenvolvidos, possui as suas dificuldades particulares. Considerando esses desafios, as metodologias foram restritas à estrutura mais simples dentre as duas representações relatadas para os modelos de Hammerstein. Essa representação é a de **KU**, que supõe que os elementos da parcela não linear são interdependentes, ou seja, a não linearidade associada a uma entrada não afeta o sinal intermediário de outra (ver Figura 2.2a)).

Por mais que essa suposição de interdependência feita para a representação de **KU** seja importante para aplicar as metodologias tal como foram desenvolvidas, a aplicação pode, a princípio, ser estendida a uma classe mais ampla de sistemas dinâmicos. Pois, na prática, os sistemas podem apresentar cruzamento entre as não linearidades, inclusive sem haver separação entre a não linearidade estática ou quase estática e a dinâmica linear. Desse modo, aplicações experimentais ou com significados físicos são apontadas como expectativas de trabalhos futuros.

Tendo em consideração os obstáculos que permeiam o problema principal, os desenvolvimentos foram direcionados para empregar a estrutura mais básica dos **MBIs**, que é o modelo de Hammerstein. Nesse sentido, de maneira geral, foi apresentada uma conexão entre as metodologias de identificação por subespaço, adequadas para modelagem de sistemas **MIMO** e outras metodologias existentes para estimar não linearidades fortes; apresentadas em (Jia et al., 2005) e (Jia et al., 2016) para o caso estático e em (Yang et al., 2013) e (Aljanaideh et al., 2018) para o cenário quase estático.

O primeiro progresso apresentado no Capítulo 3 envolveu a identificação de um modelo de Hammerstein **MIMO NF** para representar processos com não linearidades estáticas sinuosas e descontínuas. No desenvolvimento da metodologia, foram estabelecidas determinadas suposições, a saber: a matriz de transmissão direta deve ser nula; a não linearidade deve ser excitada em um ponto nulo e outro não nulo (tanto em termos de valores de entrada como de saída, ou seja, a saída resultante de uma entrada não nula deve ser uma resposta não nula). Como ocorre em (Jia et al., 2005), a possibilidade de aplicar a metodologia está condicionada a uma configuração específica do sinal de teste. Além disso, o sinal de teste deve ser aplicado separadamente em cada uma das entradas por vez, de modo que uma das entradas está ativa enquanto as demais estão nulas. Desse modo, foi possível estimar a relação não linear entre cada uma das entradas e as saídas. Portanto, o teste utiliza tantas faixas quantas forem as entradas, o que pode prolongar o ensaio, dependendo do tempo de resposta da planta.

No primeiro exemplo numérico, implementado na Subseção 3.5.1, foi possível estimar modelos para aproximar o comportamento de cada uma das duas não linearidades da entrada. O ponto de descontinuidade, como esperado, é a região onde a aproximação do modelo mais se distancia do processo. Isso ocorre devido à quantidade de dados e *clusters* utilizados. Nesse sentido, a aproximação pode ter um desempenho superior ao atual, no entanto, ao custo de aumentar a quantidade de *clusters*, o que pode deixar a execução cada vez mais lenta à medida que o número de *clusters* cresce e, conseqüentemente, o tempo de processamento para obtenção e simulação do modelo.

O segundo exemplo numérico, que deu origem à Subseção 3.5.2, tem como característica a presença de ruído branco de medição. O **MOESP** está preparado para tratar justamente este cenário, estimando eficientemente as matrizes da parcela dinâmica linear. Por sua vez, a estimativa da não linearidade estática não inclui tratamento direto ao ruído de medição. Apesar disso, ele emprega as matrizes da parcela dinâmica linear que receberam tratamento por meio do **MOESP**. Por outro lado, o sinal de saída contaminado ainda é empregado na metodologia apresentada no Algoritmo 1, impedindo o tratamento completo do ruído de medição. Em trabalhos futuros, utilizar o filtro média móvel antes de aplicar os dados no Algoritmo 1 pode contornar esse inconveniente. Mesmo com tais questões a serem aprimoradas, o desempenho do modelo de Hammerstein **MIMO NF** foi superior ao modelo de Hammerstein **MIMO** polinomial em uma ampla variação de quantidades de amostras empregadas para compor o ensaio estático.

A segunda metodologia desenvolvida neste trabalho foi apresentada no Capítulo 4 para identificar modelos de Hammerstein **MIMO GRDPI**. A ordem de identificação dos subsistemas é diferente daquela adotada no Capítulo 3. A identificação da não linearidade é feita primeiro por meio do **MPSO**. Em seguida, com a estimativa do sinal intermediário, a dinâmica linear é estimada empregando o **MOESP**. A implementação dos métodos em conjunto é possível sem a necessidade da suposição de nulidade da matriz de transmissão direta. Ademais, por estimar a curva estática primeiro, a dinâmica linear não fica restrita a sistemas de fase mínima.

A exemplo do procedimento adotado no Capítulo 3, um ensaio quase estático é realizado separadamente em cada entrada. Dado o acoplamento inserido pela dinâmica linear, as saídas respondem com sinais em regime permanente que são versões escalonadas entre si. Desse modo, como não é de interesse selecionar nenhuma saída específica, para manter um padrão emprega-se a saída que varia pela maior faixa. A efetividade da metodologia foi atestada por meio de outros dois exemplos numéricos. No primeiro exemplo numérico, apresentado na Subseção 4.5.1, desenvolve-se um cenário de testes com modelo GRDPI composto de 4 parâmetros, matriz de transmissão direta nula, sistema de fase mínima e sem ruído. A estimativa resultante foi consistente com as duas saídas do processo durante a validação em termos quase estáticos e dinâmicos, indicando que a metodologia desenvolvida foi útil ao produzir modelos de Hammerstein MIMO com acurácia.

Um segundo exemplo numérico, com características desafiadoras, foi implementado na Subseção 4.5.2. O sistema sob estudo foi constituído de 11 parâmetros com histerese assimétrica dependente da taxa, matriz de transmissão direta não nula, de fase não mínima e sob a influência de ruído branco de medição. O resultado, novamente, foi positivo com estimativas que acompanham os dados de validação quase estáticos e dinâmicos para qualquer uma das duas saídas.

Como a parcela dinâmica linear nas metodologias descritas nos Capítulos 3 e 4 são apresentadas em espaço de estados, evita-se a etapa de escolha de estrutura, que é um problema desafiador quando se trata de modelos do tipo ARX. De fato, o procedimento de identificação da parcela linear foi realizado sem conhecimento prévio a respeito do processo nas duas metodologias desenvolvidas, pois, somente as medições de dados de entrada e saída estavam disponíveis. Além disso, outro aspecto positivo do desenvolvimento é que os modelos representados em espaço de estados tem a possibilidade de ser aplicados diretamente em projetos de controle, por exemplo, utilizando desigualdades matriciais lineares (LMIs, do inglês *Linear Matrix Inequalities*). Com efeito, dada a separação dos blocos lineares e não lineares, tem-se a possibilidade de cancelar o efeito da não linearidade para aplicar técnicas de controle com base na parcela linear. Por outro lado, cada suposição feita torna a aplicabilidade do método menos abrangente. Por exemplo, supõe-se que seja possível abrir a malha do processo para realizar os ensaios e que se disponha de tempo suficiente para executar todos os ensaios nessa condição, o que restringe a aplicação a sistemas que atendam essas exigências.

## 5.2 Publicações

Os esforços dedicados para contribuir com o desenvolvimento da identificação dos modelos de blocos interconectados com não linearidades fortes resultaram na elaboração de duas produções científicas. A primeira delas foi apresentada no Simpósio Brasileiro de Automação Inteligente (SBAI) do ano de 2023. O trabalho implementa a metodologia descrita no Capítulo 3 desta dissertação. Ou seja, trata da identificação

de modelos de Hammerstein **MIMO** cujas não linearidades fortes sejam estáticas. Para aproximar as curvas estáticas não lineares da entrada emprega-se uma rede **NF**.

Uma segunda possível produção encontra-se em desenvolvimento e trata da metodologia desenvolvida no Capítulo 4. Neste artigo, aborda-se a identificação de um sistema de Hammerstein **MIMO**, composto por um modelo de histerese **GRDPI** seguido por uma representação dinâmica linear em espaço de estados. Com base nessa configuração, desenvolve-se uma metodologia para identificar sistemas multivariáveis com histerese empregando uma versão modificada do **PSO** juntamente com um algoritmo de subespaços.

As referências são apresentadas a seguir.

- (Santos et al., 2023) Santos, L. H., Ricco, R. A., e Teixeira, B. O. (2023). Identificação de modelos de Hammerstein multivariáveis com não linearidades estáticas fortes. In *XVI Simpósio Brasileiro de Automação Inteligente (SBAI 2023)*, Manaus, Brasil.
- (Santos et al., 2024) Santos, L. H., Ricco, R. A., e Teixeira, B. O. (2024). Identification of multivariable Hammerstein systems with hysteresis nonlinearities. *Em progresso*.

## 5.3 Propostas de Trabalhos Futuros

Durante a condução dessa pesquisa foram observadas possibilidades de contribuições relacionadas ao tema em estudo. Ademais, com essas contribuições, surgiram perspectivas de extensões que podem ser implementadas em trabalhos futuros. Enumera-se a seguir pontos importantes para a continuidade deste trabalho.

1. Tanto a metodologia do Capítulo 3 quanto a do Capítulo 4 são desenvolvidas aplicando o modelo em espaço de estados para representar a parcela dinâmica linear do modelo de Hammerstein. A identificação desse bloco foi implementada por meio do **MOESP** ordinário que é eficaz mesmo em situações com ruído de medição branco. No contexto de algoritmos de subespaços, existem outras opções com robustez a ruído colorido e ruído de processo, como o **MOESP-PO** (do inglês *Multivariable Output-Error State sPace - Past Output*) que merecem investigação futura.
2. As metodologias desenvolvidas foram implementadas de acordo com a Suposição 3. Todavia, nem sempre é possível abrir a malha do processo para aplicar testes. Portanto, sugere-se investigar a aplicação das metodologias em sistemas em malha fechada empregando, por exemplo, o método de identificação por subespaço baseado em preditor (**PBSID**, do inglês *Predictor-Based Subspace IDentification method*) (Chiuso, 2007).
3. Mesmo em um ambiente de simulação ideal, tendo em conta o desenvolvimento inicial da metodologia e o caráter flexível dos exemplos numéricos, aspectos

práticos foram citados em discussões durante o texto. Nesse sentido, sugere-se, como trabalhos futuros, aplicar as metodologias em exemplos com significados físicos e, sobretudo, sistemas experimentais, adicionando análises e discussões aprofundadas em termos práticos.

4. Neste trabalho a estruturação da parcela não linear do modelo de Hammerstein foi inspirada na representação **KU**, que é um caso especial da representação **EJL**, ambas evidenciadas na Figura 2.2. Dada a generalidade da segunda representação em relação à primeira, sugestiona-se o estudo da aplicação de **EJL** nas metodologias desenvolvidas nesta dissertação (Eskinat et al., 1991).
5. Os modelos de blocos interconectados possuem duas estruturas clássicas que são a de Hammerstein e a de Wiener. As duas são apresentadas na Figura 1.4. Enquanto o modelo de Hammerstein tem a estrutura mais adequada para modelar não linearidades na entrada do processo, o modelo de Wiener é mais adequado para modelar não linearidades na saída (Bai, 2010). Como o desenvolvimento do trabalho foi restrito aos modelos de Hammerstein, apresenta-se como opção para desenvolvimentos futuros a identificação, não somente dos modelos de Wiener, mas também da combinação entre eles, como Hammerstein-Wiener e Wiener Hammerstein.
6. Uma das vantagens dos modelos clássicos de blocos interconectados citadas ao longo do texto é a possibilidade de aplicar controle como se o processo fosse linear. Para que isso seja possível, pode-se utilizar o modelo estimado para compensar a não linearidade ao computar a função não linear inversa  $\mathcal{N}^{-1}$ . Posicionar  $\mathcal{N}^{-1}$  adjacente à não linearidade do modelo resulta em cancelamento da parcela não linear, de modo que somente a parcela dinâmica linear  $\mathcal{L}$  precisa ser considerada para projetar um controlador. Neste sentido, indica-se como possibilidade de trabalhos futuros a aplicação de controle sobre os modelos obtidos com as metodologias apresentadas nesta dissertação.
7. No contexto da metodologia desenvolvida no Capítulo 3, a versão monovariável da metodologia de classificação da entrada apresentada em (Jia et al., 2005) é estendida para o caso **MIMO**. A classificação ocorre considerando uma medida da distância entre as amostras do sinal de entrada. Como pode ser observado na Figura 5.1a), a largura dos *clusters*, definida pelo projetista como  $S_0$ , é mantida fixa. No entanto, nessa Figura 5.1b) nota-se que há uma parte constante que não precisa de vários *clusters* para ser representada. A implementação de um parâmetro  $S_0$  adaptativo é deixada como proposta de trabalho futuro, levando em consideração o sinal de entrada  $u(k)$  e de saída  $v(k)$ .
8. Com relação à metodologia desenvolvida no Capítulo 4, foi apresentada uma crítica que reside no fato de haver a necessidade de aplicar três ensaios para estimar o modelo de Hammerstein **MIMOGRDPI**. O sinal a ser aplicado no ensaio gerador de dados para o **MOESP** deve ser persistentemente excitante de ordem  $n$  (ver Definição 2.5.1), enquanto o sinal utilizado na identificação do modelo

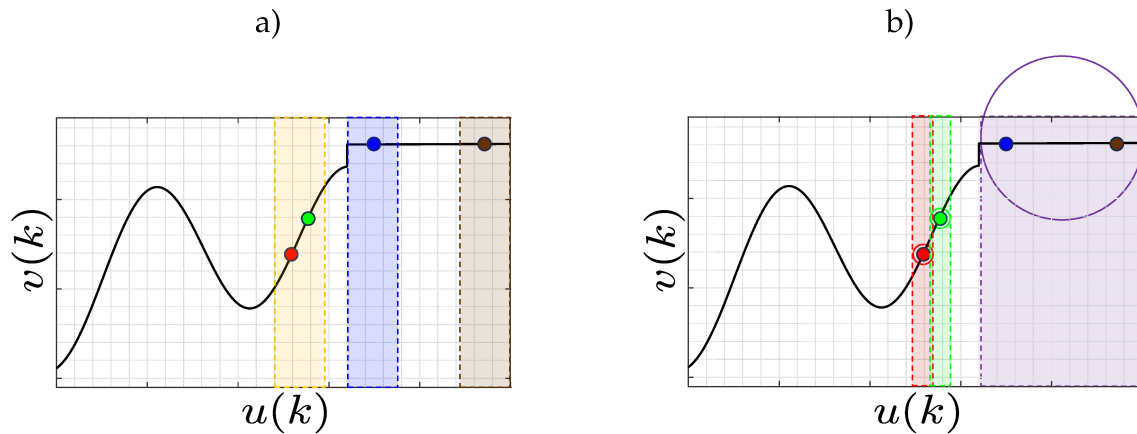


Figura 5.1: Diagramas ilustrativos da diferença entre empregar a *clusterização* com abrangência constante e dependente somente de  $u(k)$ , em a) ou adaptativa, que depende de  $u(k)$  e  $v(k)$ , em b).

GRDPI não é limitado aos degraus, mas deve ser um sinal com características de carga e descarga. Por exemplo, um sinal senoidal pode ser utilizado, como apresentado em (Yang et al., 2013). Nesse sentido, se o sinal de entrada for um sinal senoidal composto de  $n$  ou mais frequências, ele pode ser utilizado na identificação da parcela estática não linear e dinâmica linear, exceto pelo ganho total que ainda precisa de um outro ensaio para ajustes. À vista disso, propõe-se como desenvolvimento futuro reduzir a quantidade de ensaios necessários para implementar a metodologia.

9. Ainda com relação ao procedimento do Capítulo 4, observou-se que o MOESP cumpriu o seu papel de tratar ruído de medição branco, além de que o modelo de Hammerstein MIMO NF resultante alcançou desempenho superior ao polinomial em termos do índice RMSE. Apesar disso, suspeita-se que o resultado pode se tornar mais interessante, sobretudo nos casos práticos, se o sinal de saída coletado para o ensaio estático passar por tratamento antes de ser aplicado no Algoritmo 1. Portanto, apresenta-se como possibilidade de extensão deste trabalho o tratamento de ruídos no contexto da identificação da não linearidade estática do modelo de Hammerstein MIMO NF.
10. Com relação à parcela dinâmica linear estimada na metodologia do Capítulo 4, foi possível notar que há consistência na estimativa dos polos, no entanto, os zeros não são estimados corretamente. É comum o principal interesse ser dado aos polos, uma vez que eles são empregados para controle (Ricco, 2019). Todavia, quando o modelo dinâmico identificado precisa ser invertido, como ocorre em (Aljanaideh et al., 2018), mesmo nos casos em que o sistema é de fase mínima, os zeros incorretos podem resultar em modelos inversos instáveis. Desse modo, propõe-se como perspectiva de trabalhos futuros a estimativa de modelos em espaço de estados com restrição nos zeros.
11. Por fim, pode ser interessante estender as metodologias apresentadas para outros

---

algoritmos de identificação da parte não linear como o algoritmo genético ou mesmo outras representações como as redes neurais ([Krikelis et al., 2021](#)).

---

# Referências Bibliográficas

---

- Abreu, P. E. O. G. B., Tavares, L. A., Teixeira, B. O. S., e Aguirre, L. A. (2020). Identification and nonlinearity compensation of hysteresis using NARX models. *Nonlinear Dynamics*, 102(1):285–301.
- Aguirre, L. A. (2015). *Introdução à identificação de sistemas – Técnicas lineares e não lineares aplicadas a sistemas reais*. Editora UFMG, Belo Horizonte, Brasil, 4ª edition.
- Aissaoui, B., Soltani, M., Chaari, A., et al. (2016). Subspace identification of Hammerstein model with unified discontinuous nonlinearity. *Mathematical Problems in Engineering*, 2016.
- Al Janaideh, M., Al Saaideh, M., e Tan, X. (2023). The Prandtl–Ishlinskii hysteresis model: Fundamentals of the model and its inverse compensator [lecture notes]. *IEEE Control Systems Magazine*, 43(2):66–84.
- Al Janaideh, M., Rakheja, S., e Su, C. Y. (2007). A generalized Prandtl–Ishlinskii model for characterizing rate dependent hysteresis. In *2007 IEEE International Conference on Control Applications*, pages 343–348.
- Aljanaideh, K. F., Al Janaideh, M., Rakotondrabe, M., Al Saaideh, M., Almomani, A. M., Bani Hani, M. A., e Kundur, D. (2022). Identification of a class of precision motion systems with uncertain hysteretic nonlinearities. *International Journal of Control*, pages 1–18.
- Aljanaideh, K. F., Al Janaideh, M., Rakotondrabe, M., e Kundur, D. (2019). Identification of Hammerstein systems with rate-dependent hysteresis nonlinearities in a class of smart material-based actuators. In *2019 American Control Conference (ACC)*, pages 2495–2500.
- Aljanaideh, K. F., Rakotondrabe, M., Al Janaideh, M., e Kundur, D. (2018). Identification of precision motion systems with Prandtl–Ishlinskii hysteresis nonlinearities. In *2018 Annual American Control Conference (ACC)*, pages 5225–5230. IEEE.
- Babuška, R. e Verbruggen, H. (2003). Neuro-fuzzy methods for nonlinear system identification. *Annual reviews in control*, 27(1):73–85.
- Bai, E.-W. (2010). *Introduction to Block-oriented Nonlinear Systems*, pages 03–10. Springer.

- Biagiola, S. I., Agamennoni, O. E., e Figueroa, J. L. (2016). Robust control of Wiener systems: application to a pH neutralization process. *Brazilian Journal of Chemical Engineering*, 33:145–153.
- Billings, S. A. (2013). *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons.
- Billings, S. A. e Fakhouri, S. Y. (1982). Identification of systems containing linear dynamic and static nonlinear elements. *Automatica*, 18(1):15–26.
- Borjas, S. D. e Garcia, C. (2013). Identificação de sistemas não lineares usando métodos por subespaços. *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, 1(1).
- Borjas, S. D. M. e Garcia, C. (2012). Identificação determinística por subespaços. *TEMA (São Carlos)*, 13:207–218.
- Bowers, M., Kaplan, L., e Hooker, T. (1995). A two-phase model for planning the production of aluminum ingot. *European Journal of Operational Research*, 81(1):105–114.
- Box, G. E. P. e Draper, N. R. (1987). *Empirical model-building and response surfaces.*, volume 8. John Wiley & Sons.
- Chen, X.-M. e Chen, H.-F. (2011). Recursive identification for MIMO Hammerstein systems. *IEEE Transactions on Automatic Control*, 56(4):895 – 902.
- Chiuso, A. (2007). On the relation between CCA and predictor-based subspace identification. *IEEE Transactions on Automatic Control*, 52(10):1795–1812.
- de Paula, M. V., Ricco, R. A., e Teixeira, B. O. S. (2015). Identificação de modelos de Hammerstein e Wiener para sistemas não lineares multivariáveis via métodos de subespaços. *XII Simpósio Brasileiro de Automação Inteligente*.
- Dorf, R. C. e Bishop, R. H. (2011). *Modern control systems*. Prentice Hall, Upper Saddle River, NJ.
- Eskinat, E., Johnson, S. H., e Luyben, W. L. (1991). Use of Hammerstein models in identification of nonlinear systems. *AIChE Journal*, 37(2):255–268.
- Fitzgerald, A. E., Kingsley, C., e Umans, S. D. (2003). *Electric machinery.*, volume 1. McGraw-Hill, 6ª edition.
- Frącz, P. (2016). Nonlinear modeling of activated sludge process using the Hammerstein-Wiener structure. In *E3S Web of Conferences*, volume 10, page 00119. EDP Sciences.
- Gan, J. e Zhang, X. (2019). A review of nonlinear hysteresis modeling and control of piezoelectric actuators. *AIP Advances*, 9(4).

- Gan, J., Zhang, X., e Wu, H. (2016). A generalized Prandtl-Ishlinskii model for characterizing the rate-independent and rate-dependent hysteresis of piezoelectric actuators. *Review of Scientific Instruments*, 87(3).
- Garcia, C. (2013). *Modelagem e simulação de processos industriais e de sistemas eletromecânicos*. Edusp, São Paulo, Brasil, 2<sup>a</sup> edition.
- Giri, F. e Bai, E.-W. (2010). *Block-oriented nonlinear system identification*, volume 1. Springer.
- Gómez, J. C. e Baeyens, E. (2005). Subspace-based identification algorithms for Hammerstein and Wiener models. *European Journal of Control*, 11(2):127–136.
- Grimble, M. J. e Majecki, P. (2020). *Introduction to nonlinear systems modelling and control*.
- Habineza, D., Rakotondrabe, M., e Le Gorrec, Y. (2015). Bouc–Wen modeling and feedforward control of multivariable hysteresis in piezoelectric systems: Application to a 3-DoF piezotube scanner. *IEEE Transactions on Control Systems Technology*, 23(5):1797–1806.
- Hane, Y. (2022). Hysteresis modeling for power magnetic devices based on magnetic circuit method. *Journal of the Magnetics Society of Japan*, 46(2):22–36.
- Hornik, K., Stinchcombe, M., e White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366.
- Hou, J., Su, H., Yu, C., Chen, F., Li, P., Xie, H., e Li, T. (2022). Consistent subspace identification of errors-in-variables Hammerstein systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(4):2292–2303.
- Jalaleddini, K. e Kearney, R. E. (2013). Subspace identification of SISO Hammerstein systems: application to stretch reflex identification. *IEEE Transactions on Biomedical Engineering*, 60(10):2725–2734.
- Jha, P. R., Patil, S. M., e Manjunath (2020). An analysis of current supply chain best practices in die casting industry. *IJRASET International Journal for Research in Applied Science & Engineering Technology*, 8(1):633–644.
- Jia, L., Chiu, M.-S., e Ge, S. S. (2005). A noniterative neuro-fuzzy based identification method for Hammerstein processes. *Journal of Process Control*, 15(7):749–761.
- Jia, L., Li, X., e Chiu, M.-S. (2016). The identification of neuro-fuzzy based MIMO Hammerstein model with separable input signals. *Neurocomputing*, 174:530–541.
- Katayama, T. (2005). *Subspace methods for system identification*, volume 1. Springer.
- Kortmann, M. e Unbehauen, H. (1987). Identification methods for nonlinear MISO systems. *IFAC Proceedings Volumes*, 20(5):233–238.

- Krikelis, K., van Berkel, K., e Schoukens, M. (2021). Artificial neural network hysteresis operators for the identification of Hammerstein hysteretic systems. *IFAC-PapersOnLine*, 54(7):702–707.
- Kwac, L. K. e Kim, H. G. (2010). Nano positioning control for ultra-precision machining using inverse hysteresis model. *Advanced Materials Research*, 123:735–738.
- Larimore, W. E. (1990). Canonical variate analysis in identification, filtering, and adaptive control. In *29th IEEE Conference on Decision and control*, pages 596–604. IEEE.
- Li, C., Li, F., e Cao, Q. (2022). Nonlinear identification and control of the Hammerstein system with application to pH neutralization process. In *2022 34th Chinese Control and Decision Conference (CCDC)*, pages 323–327. IEEE.
- Li, F., Li, J., e Peng, D. (2017). Identification method of neuro-fuzzy-based Hammerstein model with coloured noise. *IET Control Theory & Applications*, 11(17):3026–3037.
- Li, Q., Chen, W., Wang, Y., Liu, S., e Jia, J. (2010). Parameter identification for PEM fuel-cell mechanism model based on effective informed adaptive particle swarm optimization. *IEEE Transactions on Industrial Electronics*, 58(6):2410–2419.
- Li, S. e Li, Y. (2016). Model predictive control of an intensified continuous reactor using a neural network Wiener model. *Neurocomputing*, 185:93–104.
- Ljung, L. (2010). Approaches to identification of nonlinear systems. In *Proceedings of the 29th Chinese Control Conference*, pages 1–5. IEEE.
- Mayergoyz, I. D. (1991). *Mathematical Models of Hysteresis*. Springer, New York, NY.
- Mzyk, G. (2014). *Combined parametric-nonparametric identification of block-oriented systems*, volume 454. Springer.
- Najafabadi, H. A. e Shahrokhi, M. (2016). Model predictive control of blood sugar in patients with type-1 diabetes. *Optimal Control Applications and Methods*, 37(4):559–573.
- Narendra, K. e Gallman, P. (1966). An iterative method for the identification of nonlinear systems using a Hammerstein model. *IEEE Transactions on Automatic control*, 11(3):546–550.
- Nelles, O. (2020). *Nonlinear dynamic system identification*. Springer.
- Nijmeijer, H. e Van der Schaft, A. (2016). *Nonlinear dynamical control systems*, volume 464. Springer.
- Noël, J.-P., Esfahani, A. F., Kerschen, G., e Schoukens, J. (2017). A nonlinear state-space approach to hysteresis identification. *Mechanical Systems and Signal Processing*, 84:171–184.

- Oubellil, R., Ryba, Ł., Voda, A., e Rakotondrabe, M. (2015). Experimental model inverse-based hysteresis compensation on a piezoelectric actuator. In *19th International Conference on System Theory, Control and Computing (ICSTCC)*, pages 186–191. IEEE.
- Qi, L., Weirong, C., Junbo, J., Thean, C. Y., e Ming, H. (2009). Proton exchange membrane fuel cell modeling based on adaptive focusing particle swarm optimization. *Journal of renewable and sustainable energy*, 1(1):013105.
- Qin, Y., Zhao, X., e Zhou, L. (2017). Modeling and identification of the rate-dependent hysteresis of piezoelectric actuator using a modified Prandtl-Ishlinskii model. *Micro-machines*, 8(4):114.
- Quaranta, G., Lacarbonara, W., e Masri, S. F. (2020). A review on computational intelligence for identification of nonlinear dynamical systems. *Nonlinear Dynamics*, 99(2):1709–1761.
- Rakotondrabe, M. (2017). Multivariable classical Prandtl–Ishlinskii hysteresis modeling and compensation and sensorless control of a nonlinear 2-DoF piezoactuator. *Nonlinear Dynamics*, 89(1):481–499.
- Rayouf, Z., Ghorbel, C., e Braiek, N. (2018). A new Hammerstein model control strategy: feedback stabilization and stability analysis. *International Journal of Dynamics and Control*, 7(4):1453–1461.
- Ribeiro, A. H. e Aguirre, L. A. (2014). Relações estáticas de modelos NARX MISO e sua representação de Hammerstein. In *XX Congresso Brasileiro de Automática*.
- Ricco, R. (2019). Identification of dynamical systems in state-space: Gray-box approaches.
- Roy, P., Kar, B., Kumar, A., e Roy, B. K. (2016). Level control of two tank system by fractional order integral state feedback controller tuned by PSO with experimental validation. In *2016 IEEE First International Conference on Control, Measurement and Instrumentation (CMI)*, pages 224–228. IEEE.
- Ruano, A. E. (2005). *Intelligent control systems using computational intelligence techniques*, volume 70. IET Institution of Engineering and Technology.
- Santos, L. H., Braga, M. F., e Ricco, R. A. (2021). Nova abordagem para linearização de modelos de Hammerstein identificados por métodos de subespaço. *XV Simpósio Brasileiro de Automação Inteligente (SBAI 2021)*.
- Santos, L. H., Ricco, R. A., e Teixeira, B. O. (2023). Identificação de modelos de Hammerstein multivariáveis com não linearidades estáticas fortes. In *XVI Simpósio Brasileiro de Automação Inteligente (SBAI 2023)*, Manaus, Brasil.
- Santos, L. H., Ricco, R. A., e Teixeira, B. O. (2024). Identification of multivariable Hammerstein systems with hysteresis nonlinearities. In *Não submetido*.

- Sastry, S. (2013). *Nonlinear systems: analysis, stability, and control*, volume 10. Springer Science & Business Media.
- Schetzen, M. (1981). Nonlinear system modeling based on the Wiener theory. *Proceedings of the IEEE*, 69(12):1557–1573.
- Song, J. e Der Kiureghian, A. (2006). Generalized Bouc–Wen model for highly asymmetric hysteresis. *Journal of engineering mechanics*, 132(6):610–618.
- Sumathi, S. e Paneerselvam, S. (2010). *Computational intelligence paradigms: theory & applications using MATLAB*. crc Press.
- Sung, S. W. (2002). System identification method for Hammerstein processes. *Industrial & Engineering Chemistry Research*, 41(17):4295–4302.
- Takagi, T. e Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-15(1):116–132.
- Tan, X. e Iyer, R. V. (2009). Modeling and control of hysteresis. *IEEE Control Systems Magazine*, 29(1):26–28.
- Tao, Y.-D., Li, H.-X., e Zhu, L.-M. (2020). Hysteresis modeling with frequency-separation-based gaussian process and its application to sinusoidal scanning for fast imaging of atomic force microscope. *Sensors and Actuators A: Physical*, 311:112070.
- Tavares, L. A., Abreu, P. E. O. G. B., e Aguirre, L. A. (2022). Nonlinearity compensation based on identified NARX polynomials models. *Nonlinear Dynamics*, 107(1):709–725.
- Van Overschee, P. e De Moor, B. (1994). N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, 30(1):75–93.
- Van Overschee, P. e De Moor, B. (2012). *Subspace identification for linear systems: Theory—Implementation—Applications*. Springer Science & Business Media.
- Verhaegen, M. (1996). A subspace model identification solution to the identification of mixed causal, anti-causal LTI systems. *SIAM Journal on Matrix Analysis and Applications*, 17(2):332–347.
- Verhaegen, M. e Dewilde, P. (1992a). Subspace model identification part 1: The output-error state-space model identification class of algorithms. *International Journal of Control*, 56(5):1187–1210.
- Verhaegen, M. e Dewilde, P. (1992b). Subspace model identification part 2: Analysis of the elementary output-error state-space model identification algorithm. *International journal of control*, 56(5):1211–1241.
- Verhaegen, M. e Verdult, V. (2007). *Filtering and system identification: a least squares approach*. Cambridge University Press.

- Verhaegen, M. e Westwick, D. (1996). Identifying MIMO Hammerstein systems in the context of subspace model identification methods. *IFAC Proceedings Volumes*, 29(1):4092–4097.
- Wang, W., Wang, R., Chen, Z., Sang, Z., Lu, K., Han, F., Wang, J., e Ju, B. (2020). A new hysteresis modeling and optimization for piezoelectric actuators based on asymmetric prandtl-ishlinskii model. *Sensors and Actuators A: Physical*, 316:112431.
- Wang, Z., Zhang, Z., Mao, J., e Zhou, K. (2012). A Hammerstein-based model for rate-dependent hysteresis in piezoelectric actuator. In *2012 24th Chinese Control and Decision Conference (CCDC)*, pages 1391–1396.
- Yang, M.-J., Gu, G.-Y., e Zhu, L.-M. (2013). Parameter identification of the generalized Prandtl–Ishlinskii model for piezoelectric actuators using modified particle swarm optimization. *Sensors and Actuators A: Physical*, 189:254–265.
- Zeng, K., Zhang, N.-Y., e Xu, W.-L. (2000). A comparative study on sufficient conditions for Takagi-Sugeno fuzzy systems as universal approximators. *IEEE Transactions on Fuzzy Systems*, 8(6):773–780.
- Zhang, B., Mao, Z., e Zhang, T. (2017). Intelligent control for Hammerstein nonlinear systems with arbitrary deadzone input. *Transactions of the Institute of Measurement and Control*, 39(4):567–578.
- Zou, Z., Zhao, D., Liu, X., Guo, Y., e Guan (2015). Pole-placement self-tuning control of nonlinear Hammerstein system and its application to pH process control. *Chinese Journal of Chemical Engineering*, 23(8):1364–1368.