

Subspace Identification of Linear Systems with Partial Eigenvalue Constraints

R. Ricco, A. Verly, M. de Paula, and B. Teixeira

Abstract—For subspace identification methods with eigenvalue constraints, the constraints are enforced by means of an optimization problem subject to LMI constraints. First principals or step response tests could be used as a source of auxiliary information in order to build LMI regions. In these cases, all the eigenvalues of the identified state-space model are subject to the same constraints. However, it often happens that the non-dominant eigenvalues have larger real part or larger natural frequencies. In this paper, we propose a two-step method in order to constrain the dominant dynamics of SISO models into LMI regions. In virtue of this result, in addition, the model eigenvalues could be constrained into disjoint LMI regions. Numerical examples illustrate the effectiveness of our proposed method.

Index Terms—Subspace identification, Partial eigenvalue constraint, Gray-box identification, Linear systems, LMI, SISO systems.

I. INTRODUÇÃO

A informação auxiliar pode ser utilizada em identificação de sistemas para incorporar determinadas propriedades do sistema sob investigação, tais como, ganho em estado estacionário [1]–[3], restrições de igualdade nos estados [4], curva estática não linear [5], entre outras. O uso dessas informações auxiliares se faz necessário, principalmente, quando os dados são limitados em termos de persistência de excitação, número de amostras ou, no caso de sistemas não lineares, limitação de pontos de operação nos dados [6]. Na maioria dos casos, essas limitações ocorrem devido a restrições de testes na operação do sistema, principalmente no caso de dados originários de registros históricos do processo [7]. Como mostrado em [8], o uso apropriado de informação auxiliar pode contribuir na solução do problema de modelos obtidos por meio de dados limitados de entrada e saída.

Recentemente, o desenvolvimento de métodos de subespaços caixa-cinza tem chamado a atenção da comunidade científica de modelagem e controle [1]–[3], [9]–[12], mesmo sendo um desafio [13]. Tal desafio se deve à dificuldade dos modelos estimados terem significado físico, tendo em vista que são estimadosa menos de uma transformação de

similaridade desconhecida. Em [9], uma metodologia para garantir a estabilidade dos modelos obtidos em espaço de estados é estabelecida. O caso recursivo foi tratado em [12]. No trabalho de [1], o ganho em estado estacionário é utilizado como informação auxiliar em uma das etapas do método de identificação por subespaços em batelada. Em um procedimento similar, [10] inserem a informação auxiliar da constante de tempo e ganho em estado estacionário em modelos estimados de primeira ordem. Em [2], é proposta uma metodologia que permite utilizar a informação auxiliar do ganho em estado estacionário e de termos nulos na matriz de transferência em modelos multivariáveis variantes do tempo. Em consequência dos trabalhos de [10] e [2], em [3] as informações auxiliares de constante de tempo em sistemas de primeira ordem, ganho em estado estacionário e de termos nulos na matriz de transferência em modelos multivariáveis são utilizadas em um procedimento de identificação por subespaços em malha fechada. Dentre esses trabalhos, o trabalho de [11] se destaca pela proposta de uma abordagem que permite inserir informação auxiliar da localização dos autovalores (polos) do sistema por meio de regiões LMIs (do inglês, *linear matrix inequalities*).

Aspectos práticos relacionados ao mapeamento de regiões LMI em tempo discreto para restringir os autovalores de modelos, como proposto em [11], foram recentemente discutidos em [14]. Nesse último trabalho, informações da dinâmica dominante em tempo contínuo, tais como o sobressinal, período entre oscilações e tempo de acomodação são utilizados como base para formar regiões LMI convexas em tempo discreto. Uma das lacunas discutidas em [14], no uso de regiões LMIs por meio dessa abordagem, é a impossibilidade de restringir somente os autovalores relacionados à dinâmica dominante do sistema. Nesse sentido, sistemas com autovalores não dominantes reais ou de alta frequência estariam sujeitos às mesmas restrições dos autovalores dominantes. No melhor do conhecimento dos autores, restringir os autovalores de forma parcial em identificação por subespaços ainda é uma questão em aberto. Embora em alguns casos o efeito da dinâmica não dominante de alta frequência possa ser negligenciado devido a restrições nos testes, como apontado no início desta seção, ter a possibilidade de restringir somente os autovalores conhecidos previamente também é um resultado relevante. Nesse caso, tem-se em mente o conhecimento prévio de autovalores obtidos por meio das equações que regem o sistema dinâmico.

Dessa forma, a contribuição deste trabalho é propor uma metodologia em duas etapas para identificação de modelos SISO (do inglês, *single-input single-output*) no espaço de estados com restrição parcial nos autovalores. Esta proposta

Rodrigo A. Ricco, Universidade Federal de Ouro Preto, Departamento de Engenharia Elétrica, João Monlevade, Minas Gerais, Brasil e Universidade Federal de Minas Gerais, Programa de Pós-Graduação em Engenharia Elétrica, Belo Horizonte, Minas Gerais, Brasil, ricco@deelt.ufop.br

Anny Verly, Universidade Federal de Ouro Preto, Departamento de Engenharia Elétrica, João Monlevade, Minas Gerais, Brasil, annyverly@ufop.edu.br

Marcus V. de Paula, Universidade Federal de Minas Gerais, Programa de Pós-Graduação em Engenharia Elétrica, Belo Horizonte, Minas Gerais, Brasil, marcusdepaula@ufmg.br

Bruno O. S. Teixeira, Universidade Federal de Minas Gerais, Programa de Pós-Graduação em Engenharia Elétrica, Belo Horizonte, Minas Gerais, Brasil, brunoot@ufmg.br

é baseada em parte no desenvolvimento proposto por [15] em alocação parcial de polos de sistemas de controle robusto operando em malha fechada.

Este artigo é organizado como se segue. Na Seção II é formulado o problema sob investigação. Nas Seções III e IV, são apresentadas as ferramentas metodológicas essenciais para o desenvolvimento deste trabalho. Na Seção V é apresentada uma nova metodologia em duas etapas para identificação por subespaços com restrição parcial nos autovalores de sistemas SISO. Os exemplos numéricos da Seção VI ilustram a efetividade do método proposto. Por fim, na Seção VII são apresentadas as conclusões.

II. FORMULAÇÃO DO PROBLEMA

Considere o sistema linear, estritamente estável, discreto e invariante no tempo representado por

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k, \\ y_k &= Cx_k + Du_k + \nu_k, \end{aligned} \quad (1)$$

em que $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times 1}$, $C \in \mathbb{R}^{1 \times n}$ e $D \in \mathbb{R}^{1 \times 1}$. Os vetores $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}$ e $y_k \in \mathbb{R}$ representam, respectivamente, os estados desconhecidos, a entrada conhecida e a saída medida. As sequências $w_k \in \mathbb{R}^n$ e $\nu_k \in \mathbb{R}$ são os ruídos de processo e de medição, respectivamente, ambos assumidos Gaussianos com média nula.

Assumindo que a matriz de observabilidade estendida

$$\hat{O}^T \triangleq [C^T \quad (CA)^T \quad \dots \quad (CA^k)^T] \quad (2)$$

seja obtida por meio de algum método de identificação por subespaços da classe MOESP (do inglês, *multivariable output error state space*) [13]. Assumindo também que informação auxiliar dos autovalores (polos) dominantes de A em (1) está disponível.

Este artigo tem como objetivo geral impor restrições parciais nos autovalores da matriz \hat{A} do modelo estimado para (1) em uma metodologia de duas etapas. Como objetivo específico, propõe-se estimar a matriz \hat{A} , a menos de uma transformação de similaridade desconhecida, sujeita a restrições nos autovalores dominantes.

III. IDENTIFICAÇÃO POR SUBESPAÇOS COM RESTRIÇÃO NO GANHO EM ESTADO ESTACIONÁRIO

Em identificação por subespaços, o método MOESP com restrição no ganho em estado estacionário utiliza a matriz de observabilidade estendida \hat{O} para estimar as matrizes \hat{A} e \hat{C} do modelo (1). Logo após, por meio de \hat{A} e \hat{C} , da informação auxiliar do ganho em estado estacionário K do sistema e dos dados de entrada u_k e saída y_k , as matrizes \hat{B} e \hat{D} são estimadas [1]. Neste trabalho, a informação do ganho em estado estacionário K está disponível. Ressalta-se que a restrição de ganho em estado estacionário não é fundamental para a aplicação da metodologia proposta neste artigo. Essa escolha tem como objetivo a comparação das metodologias para o mesmo ganho em estado estacionário, tendo em vista que, com dados restritos, os modelos podem

falhar na estimação do ganho. Esse procedimento é brevemente revisitado em seguida.

Segundo [13], da primeira linha da matriz (2) obtém-se $\hat{C} = \hat{O}_{(1, :)}$, em que os subscritos de \hat{O} denotam o estilo de indexação do Matlab. Por meio da propriedade de invariância no tempo, definem-se $\hat{O}_0 \triangleq \hat{O}_{(1:k-1, :)}$ e $\hat{O}_1 \triangleq \hat{O}_{(2:k, :)}$. Essas matrizes deslocadas permitem estimar a matriz \hat{A} solucionando o problema de mínimos quadrados linear

$$J_O(A) \triangleq \left\| \hat{O}_0 A - \hat{O}_1 \right\|_F, \quad (3)$$

cujas solução analítica é dada por

$$\hat{A} = (\hat{O}_0^T \hat{O}_0)^{-1} \hat{O}_0^T \hat{O}_1. \quad (4)$$

De posse das matrizes \hat{A} e \hat{C} , a equação de saída (1) pode ser escrita como

$$y_k = \left(\sum_{\tau=0}^{k-1} u_\tau^T \otimes \hat{C} \hat{A}^{k-\tau-1} \right) \text{vec}(B) + (u_k^T \otimes I_1) \text{vec}(D) + e_k, \quad (5)$$

em que e_k é o termo relacionado à contribuição do ruído e \otimes é o operador produto de Kronecker [13]. A partir de (5) deseja-se encontrar \hat{B} e \hat{D} . Dessa forma, define-se

$$J(\theta) \triangleq \|Y - Z^T \theta\|^2, \quad (6)$$

onde Y representa os vetores y_k empilhados, $\theta^T = [\text{vec}(B) \quad \text{vec}(D)]^T$ e $Z = (\psi_1, \dots, \psi_N)$, em que

$$\psi_k^T \triangleq \left[\sum_{\tau=0}^{k-1} u_\tau^T \otimes \hat{C} \hat{A}^{k-\tau-1} \quad (u_k^T \otimes I_1) \right]. \quad (7)$$

A solução que minimiza a função custo (6) é dada por

$$\hat{\theta} = (Z^T Z)^{-1} Z^T Y, \quad (8)$$

caso o ruído que contamina o sistema (1) for Gaussiano, pode-se afirmar que o estimador (8) não é polarizado [1].

Em [1] assume-se que a soma dos elementos da resposta ao impulso de um sistema assintoticamente estável é igual ao valor em estado estacionário desse sistema

$$D + CB + CAB + CA^2B + \dots + CA^k B = K. \quad (9)$$

De (9), após algumas manipulações, obtém-se

$$S\theta = G, \quad (10)$$

em que $S \triangleq [I_1 \otimes \hat{C}(I_n - \hat{A})^{-1} \quad I_1]$ e $G \triangleq \text{vec}(K)$. A solução do problema (6), sujeito a restrição de ganho em estado estacionário (10), é dada por

$$\hat{\theta}_c = \hat{\theta} - (Z^T Z)^{-1} S^T [S(Z^T Z)^{-1} S^T]^{-1} (S\hat{\theta} - G). \quad (11)$$

IV. IDENTIFICAÇÃO POR SUBESPAÇOS COM RESTRIÇÃO NOS AUTOVALORES

O problema de otimização (3) pode ser reformulado como um problema de otimização sujeito a restrições LMI com o objetivo de impor restrições nos autovalores de \hat{A} em um conjunto convexo. Nesta seção, esses resultados são brevemente revisitados seguindo o desenvolvimento de [11].

A. Utilizando a Matriz de Observabilidade Estendida

Com o objetivo de resolver o problema convexo de otimização sujeito a restrições nos autovalores, multiplica-se (3), à direita, pela matriz $P = P^T \in \mathbb{R}^{n \times n}$, tal que

$$\mathcal{J}_O(A, P) = \|\hat{O}_0 A P - \hat{O}_1 P\|_F. \quad (12)$$

A função obtida em (12) deve ser linearizada, tal que $AP = Q$ e $Q \in \mathbb{R}^{n \times n}$, assim

$$\mathcal{J}_O(Q, P) = \|\hat{O}_0 Q - \hat{O}_1 P\|_F. \quad (13)$$

A função custo linearizada (13) está escrita em um formato adequado para inserir restrições de \mathcal{D} -estabilidade nos autovalores de \hat{A} , conforme será mostrado nas próximas subseções.

B. Restrição nos Autovalores

O conceito de regiões de estabilidade LMI, introduzido inicialmente em [16], define regiões convexas do plano complexo como LMIs. De acordo com [16], um sistema dinâmico é denominado \mathcal{D} -estável se todos os autovalores da matriz A estão contidos na região \mathcal{D} . Desse modo, A é denominada \mathcal{D} -estável.

Uma região LMI é uma região convexa \mathcal{D} do plano complexo, definida em termos de uma matriz simétrica α e uma matriz quadrada β [11], tal que

$$\mathcal{D} = \{z \in \mathbb{C} : f_{\mathcal{D}}(z) \geq 0\}, \quad (14)$$

em que

$$f_{\mathcal{D}}(z) = \alpha + \beta z + \beta^T z^*, \quad (15)$$

é a função característica cujos parâmetros α e β podem ser utilizados para formar desigualdades do tipo Lyapunov. As funções características que definem as regiões LMIs utilizadas neste trabalho para a escolha dos parâmetros α e β são apresentadas na Subseção IV-C e seguem as abordagens propostas em [11] e [14].

Empregando o teorema de [11], adaptado em [16], tem-se:

Teorema 4.1 ([16]): Os autovalores de uma matriz A estão no interior de uma região LMI dada por (14), se e somente se, existir uma matriz simétrica $P \in \mathbb{R}^{n \times n}$ tal que

$$P = P^T > 0, \quad \mathcal{M}_{\mathcal{D}}(A, P) \geq 0, \quad (16)$$

em que

$$\mathcal{M}_{\mathcal{D}}(A, P) = \alpha \otimes P + \beta \otimes (AP) + \beta^T \otimes (AP)^T. \quad (17)$$

Segundo [11], o problema de identificação por subespaços com restrição nos autovalores (MOESP-R) é dado por

$$\begin{aligned} & \text{minimize} && J_O(Q, P), \\ & \text{sujeito a} && \mathcal{M}(Q, P) \geq 0, \\ & && P = P^T > 0, \\ & && \text{tr}(P) = n, \end{aligned} \quad (18)$$

em que $J_O(Q, P)$ é dado por (13). Desse modo, uma vez estimadas as matrizes Q e P , tem-se que $\hat{A} = QP^{-1}$.

Observação 1: Vale notar que, com o objetivo de melhorar o condicionamento numérico do problema, [11] introduz a condição $\text{tr}(P) = n$ em (18), forçando uma possível solução $P = I$.

Observação 2: Destaca-se que, no problema de otimização com restrição (18) proposto por [11], todos os autovalores estão sujeitos às mesmas restrições LMI. Assim, mesmo que alguma informação auxiliar de cada um dos autovalores do sistema estivesse disponível, todos os autovalores do modelo estariam sujeitos às mesmas restrições LMI. Desse modo, de acordo com as informações auxiliares, deve-se avaliar qual a melhor restrição para o sistema a ser identificado.

A interseção de diferentes regiões LMI convexas também forma uma região LMI convexa [16]. Nesse sentido, o Lema 4.2 é de interesse.

Lema 4.2 ([16]): Dadas N regiões LMIs $\{\mathcal{D}_1, \dots, \mathcal{D}_N\}$, a interseção dessas regiões $\mathcal{D} = \mathcal{D}_1 \cap \dots \cap \mathcal{D}_N$ tem como função característica

$$f_{\mathcal{D}}(z) = \text{diag}\{f_{\mathcal{D}_1}(z), \dots, f_{\mathcal{D}_N}(z)\} \quad (19)$$

C. Regiões LMI úteis em Identificação por Subespaços

As subseções a seguir apresentam alguns tipos de regiões LMI úteis em identificação por subespaços. A Fig. 1 exemplifica cada uma dessas regiões.

1) *Região $\text{Re}(z) \geq \sigma$ (Fig. 1a):*

Fato 4.1 ([11]): O conjunto

$$\mathcal{S} = \{z \in \mathbb{C} : \text{Re}(z) \geq \sigma, \sigma \geq 0\},$$

é equivalente à região LMI $f_{\mathcal{S}}(z) \geq 0$,

$$f_{\mathcal{S}}(z) = \begin{bmatrix} 1 & 0 \\ 0 & -2\sigma \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} z + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} z^*. \quad (20)$$

Essa região pode ser utilizada quando se tem a informação auxiliar aproximada da localização dos autovalores dominantes.

2) *Região $\text{Im}(z) = 0$ (Fig. 1b):*

Fato 4.2 ([11]): O conjunto

$$\mathcal{P} = \{z \in \mathbb{C} : \text{Im}(z) = 0\},$$

é equivalente à região LMI $f_{\mathcal{P}}(z) \geq 0$,

$$f_{\mathcal{P}}(z) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} z + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} z^*. \quad (21)$$

Essa região é útil quando a resposta do sistema é de primeira ordem ou sobrearmortecida. Além disso, quando a ordem do sistema é ímpar, sabe-se que ao menos um dos pólos do sistema é puramente real.

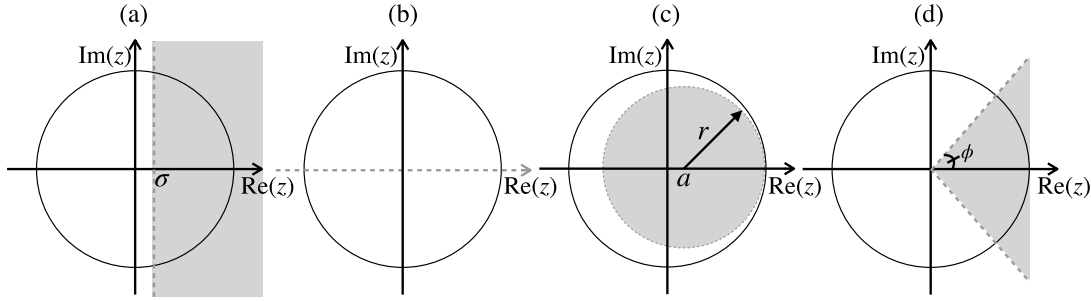


Fig. 1. Regiões LMI no plano- z . Os autovalores dos modelos identificados são restritos às regiões em destaque: (a) à direita de σ ($\text{Re}(z) \geq \sigma$); (b) no eixo real ($\text{Im}(z) = 0$); (c) dentro do círculo centrado em a de raio r ($|z| \leq r$) e (d) à direita do cone de ângulo ϕ ($|\text{Im}(z)| < \text{Re}(z)\tan(\phi)$).

3) Região do disco centrado em a e raio r (Fig. 1c):

Fato 4.3 ([11]): O conjunto

$$\mathcal{C} = \{z \in \mathbb{C} : |z| \leq r, 0 \leq r \leq 1\},$$

é equivalente à região LMI $f_{\mathcal{C}}(z) \geq 0$,

$$f_{\mathcal{C}}(z) = \begin{bmatrix} 1 & 0 \\ 0 & -(aa^* - r^2) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & a \end{bmatrix} z + \begin{bmatrix} 0 & 1 \\ 0 & a^* \end{bmatrix} z^*. \quad (22)$$

Se $a \in \mathbb{R}$, então o disco é simétrico em relação ao eixo real. Para $r = 1$ e $a = 0$, tem-se a condição de estabilidade de Lyapunov para sistemas discretos. Essa região é útil para formar regiões LMI quando se tem o sobressinal ou tempo de acomodação obtidos de testes de resposta ao degrau do sistema sob investigação [14].

4) Região à direita do cone (Fig. 1d):

Fato 4.4 ([14]): O conjunto

$$\mathcal{C}_{\mathcal{M}} = \{z \in \mathbb{C} : |\text{Im}(z)| - \text{Re}(z)\tan(\phi) < 0\},$$

em que $\text{tg}(\phi) = |\text{Im}(z)|/\text{Re}(z)$ é equivalente à região LMI $f_{\mathcal{C}_{\mathcal{M}}}(z) \geq 0$,

$$f_{\mathcal{C}_{\mathcal{M}}}(z) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \text{sen}(\phi) & \text{cos}(\phi) \\ -\text{cos}(\phi) & \text{sen}(\phi) \end{bmatrix} z + \begin{bmatrix} \text{sen}(\phi) & -\text{cos}(\phi) \\ \text{cos}(\phi) & \text{sen}(\phi) \end{bmatrix} z^*. \quad (23)$$

Segundo [14], esse tipo de região mapeia as bandas das frequências amortecidas ω_d do plano- s para o plano- z .

Maiores detalhes sobre o mapeamento de informação auxiliar de regiões do plano- s para o plano- z podem ser consultados em [14].

V. IDENTIFICAÇÃO POR SUBESPAÇOS COM RESTRIÇÃO PARCIAL NOS AUTOVALORES

A metodologia desenvolvida a seguir difere-se da proposta por [11], uma vez que é possível utilizar informação auxiliar para restringir a estimação dos autovalores de (1) separadamente. Com essa finalidade, inspirando-se no procedimento proposto por [15], propõe-se resolver o problema em duas etapas.

A. Estrutura da Matriz de Transição de Estados A

Primeiramente, escreve-se equação característica $\Lambda(z) \in \mathbb{C}^n$ de A na forma fatorada como

$$\Lambda(z) \triangleq \prod_{i=1}^{n_p} \left[\prod_{j=1}^{m_i} (z + p_{ij}) \right], \quad (24)$$

em que $n_p \in \mathbb{Z}_+^*$ é o número de partições do polinômio, m_i é o número de autovalores por partição, $i \in \mathbb{Z}_+^*$ é a posição de cada partição e $j \in \mathbb{Z}_+^*$ é a posição de cada autovalor em cada uma das partições. Considerando que não há cancelamento entre polos e zeros do sistema, os autovalores de A são idênticos aos polos da equação característica (24). Assumindo-se que os autovalores reais e complexos de (24) são diferentes, cada partição pode ser apresentada na sua forma canônica de Jordan, dada por

$$A_i \triangleq \text{diag}\{p_{i1}, p_{i2}, \dots, p_{im_i}\}. \quad (25)$$

Deve-se observar que a partição representada em (25) também pode ser escrita na forma canônica modal.

Se a informação auxiliar sobre n_{co} partições estão disponíveis, então a equação característica da matriz A (24) pode ser reescrita como

$$A \triangleq \text{diag}\{A_1, \dots, A_{n_{\text{co}}}, A_{\text{de}}\}, \quad (26)$$

em que A_{de} é o conjunto de partições desconhecidas. Considerando o conjunto de partições conhecidas, escritas em bloco, como $A_{\text{co}} \triangleq \text{diag}\{A_1, \dots, A_{n_{\text{co}}}\}$. Por simplicidade, no restante do manuscrito, considera-se que

$$A \triangleq \text{diag}\{A_{\text{co}}, A_{\text{de}}\}. \quad (27)$$

Observação 3: O particionamento de A proposto em (27) força o desacoplamento dos estados e a estrutura do modelo a ser estimado. Dessa forma, assume-se que, ao usar a estrutura (27), o sistema a ser identificado é desacoplado.

Observação 4: Nota-se que, dependendo do conhecimento prévio que se tem, a posição das submatrizes de blocos A_{co} e A_{de} podem ser trocadas. Dado que a posição dessas matrizes influenciam no desenvolvimento das próximas subseções, por questões simplificadoras, assume-se que a informação prévia

será utilizada para restringir os autovalores do bloco A_{co} . O resultado contrário pode ser facilmente obtido com algumas manipulações.

Definindo-se $P = P^T \in \mathbb{R}^{n \times n}$ e

$$P \triangleq \text{diag}\{P_{co}, I\}, \quad (28)$$

em que P_{co} é um conjunto de matrizes simétricas $P_i \forall i \in \{1, \dots, n_{co}\}$, de modo que cada uma dessas matrizes são associadas a cada uma das submatrizes de A_{co} .

Observação 5: Definindo-se a matriz simétrica $P = P^T$ como $P \triangleq \text{diag}\{I, P_{de}\}$, os resultados a seguir podem ser facilmente estendidos para as matrizes associadas a A_{de} . Na prática, como se tratam de parâmetros dos quais não se possui nenhuma informação prévia, é de interesse ao menos inserir uma restrição de estabilidade nas submatrizes associadas a esses parâmetros.

Nas próximas subseções a metodologia MOESP com restrição parcial (RP) nos autovalores é proposta.

B. MOESP com Restrição Parcial (RP) nos Autovalores

Seguindo um procedimento similar ao exposto na subseção IV-A, multiplicando-se (3) pela matriz (28) à direita, tem-se

$$\mathcal{J}_{ORP}(A, P) = \|\hat{O}_0 A P - \hat{O}_1 P\|_F. \quad (29)$$

Reparametrizando (29) com $AP = Q$, tem-se

$$\mathcal{J}_{ORP}(Q, P) = \|\hat{O}_0 Q - \hat{O}_1 P\|_F, \quad (30)$$

em que Q é dada por

$$Q \triangleq \text{diag}\{Q_{co}, A_{de}\}, \quad (31)$$

e $Q_{co} = A_{co} P_{co}$.

Dado que a região \mathcal{D} -estável é definida levando-se em consideração apenas os autovalores a serem estimados, obtém-se

$$\mathcal{M}_{\mathcal{D}}(Q_{co}, P_{co}) = \alpha \otimes P_{co} + \beta \otimes Q_{co} + \beta^T \otimes Q_{co}^T. \quad (32)$$

C. Identificação por Subespaços com Restrição Parcial dos Autovalores

O problema de *identificação por subespaços com restrição parcial dos autovalores* pode ser formulado da seguinte maneira

$$\begin{aligned} & \text{minimize} && \mathcal{J}_{ORP}(Q, P), \\ & \text{sujeito a} && \mathcal{M}_{\mathcal{D}}(Q_{co}, P_{co}) \geq 0, \\ & && P_{co} = P_{co}^T > 0, \\ & && \text{tr}(P_{co}) = n - m, \end{aligned} \quad (33)$$

em que m é o número de polos desconhecidos, $\mathcal{J}_{ORP}(Q, P)$ é dada por (30), $\mathcal{M}_{\mathcal{D}}(Q_{co}, P_{co})$ é dada por (32), Q é dada por (31), P é dada por (28) e A é dada por (25), (26) e (27). Se o problema (33) for factível, tem-se que $\hat{A}_{co} = \hat{Q}_{co}(\hat{P}_{co})^{-1}$.

Observação 6: O problema de otimização (33) pode ser formulado adicionando as restrições $\mathcal{M}_{\mathcal{D}}(Q_{de}, P_{de}) \geq 0$, $P_{de} = P_{de}^T > 0$ e $\text{tr}(P_{de}) = m$ de forma a garantir estabilidade nos autovalores de A_{de} . Nesse sentido, Q_{de} deve ser formulada à exemplo de (31) com os termos da diagonal principal invertidos.

D. Procedimento em Duas Etapas

O Algoritmo 5.1 apresenta a solução de (33) em duas etapas. Na primeira etapa, utiliza-se a informação *a priori* do sistema, que pode ser relativa à posição dos autovalores dominantes. Como proposto em [14], tal informação pode ser obtida por meio de ensaios ao degrau ou a partir das equações fenomenológicas do sistema dinâmico. De posse da informação *a priori*, a estimação dos demais autovalores desconhecidos é realizada por meio do método MOESP utilizando-se de restrições de \mathcal{D} -estabilidade para garantir que o modelo estimado seja estável.

A segunda etapa tem como objetivo aprimorar a estimativa dos autovalores pré-estabelecidos na primeira etapa. Dessa forma, inverte-se a ordem das soluções. Utilizam-se os autovalores estimados, via MOESP, na etapa anterior como autovalores fixos e realiza-se uma nova estimação, também via MOESP, por meio de restrições de \mathcal{D} -estabilidade para estimar os autovalores pré-estabelecidos na primeira etapa. Uma questão importante desta abordagem é a possibilidade de restringir regiões \mathcal{D} -estáveis de formas disjuntas. Desse modo, os autovalores podem ser estimados em regiões LMIs distintas sem a necessidade de haver interseção entre essas.

Algoritmo 5.1: MOESP com restrição parcial (RP) nos autovalores.

1° Passo: Por meio dos dados de entrada e saída, estime \mathcal{O} .
2° Passo: Estime C e avalie se alguma informação prévia do sistema está disponível. Se estiver, defina as partições n_p , o número de autovalores m_i de cada uma das n_p partições e as regiões LMI no qual os autovalores das n_{co} partições serão estimados.

1ª Etapa MOESP-RP

3° Passo: Fixe os autovalores de A_{co} oriundos da informação auxiliar e estime por meio de (33) os autovalores de A_{de} com restrição de estabilidade.

2ª Etapa MOESP-RP

4° Passo: Fixe os autovalores de A_{de} estimados no Passo 3 e estime por meio de (33) os autovalores de A_{co} com restrição de estabilidade de acordo com as regiões definidas no Passo 2.

5° Passo: Verifique se a informação de K está disponível. Se estiver, com os valores de \hat{A} e \hat{C} estimados nos passos precedentes, estime B e D com restrição no ganho em estado estacionário por meio de (11).

6° Passo: Valide o modelo estimado com restrições parciais nos autovalores e avalie a necessidade de voltar ao Passo 2, redefinindo as regiões obtidas por meio da informação auxiliar. \square

VI. RESULTADOS NUMÉRICOS

Nesta seção, aplica-se a metodologia proposta na Seção V em dois sistemas simulados, um de terceira ordem e outro de quarta ordem. Os resultados do MOESP-RP são comparados aos resultados dos métodos MOESP irrestrito e MOESP-R [11]. Para a resolução dos problemas de otimização com restrições convexas, as ferramentas YALMIP e MOSEK [17], ambos pacotes do MATLAB, são utilizadas.

A. Sistema de Terceira Ordem

Considere o sistema representado por

$$F(s) = \frac{10^3}{s^3 + 30s^2 + 400s + 4000}, \quad (34)$$

cujos autovalores da matriz A correspondente do espaço de estados são $s_{1,2} = -5 \pm 13,23j$ e $s_3 = -20$. O sinal de saída do sistema é medido com uma taxa de amostragem $T_s = 0,03s$ e contaminado com ruído Gaussiano ν com média nula e desvio padrão $\sigma_\nu = 0,1$. Para realizar a identificação do sistema, utiliza-se um sinal de entrada binário pseudo aleatório (PRBS) de 16 bits com período entre bits de 1 amostra, com tempo de simulação de 5s. São investigadas 200 execuções com diferentes realizações de ruído.

A informação *a priori* é extraída de um procedimento de média simples de cinco respostas ao degrau contaminadas com ruído Gaussiano ν com média nula e desvio padrão $\sigma_\nu = 0,1$, semelhante ao procedimento adotado em [14]. De posse das informações de máximo sobressinal, tempo de pico e valor em regime permanente, obtém-se o ganho $K = 0,25$ e os autovalores dominantes $s_{1,2}^{priori} = -6,25 \pm 10,65j$. Com a finalidade de aplicar o método MOESP-RP, os autovalores são mapeados para o plano- z utilizando a relação $z = e^{s_{1,2}T_s}$ [14]; ver legenda da Fig. 2.

Nesse sentido, definem-se duas regiões LMI diferentes entre as etapas 1 e 2 do método MOESP-RP na estimação de autovalores para (34); ver Algoritmo 5.1. Na etapa 1, como a informação sobre a posição do terceiro autovalor é desconhecida, restringe-se a estimação ao intervalo de números reais entre $[0,01 \ 0,99]$. Esta região é definida pela interseção do eixo real e o disco de raio $r = 0,49$ e centro $a = 0,5$; ver subseções IV-C2 e IV-C3. Na etapa 2, que trata da estimação de autovalores complexos conjugados com informação *a priori*, a região LMI é definida como sendo a interseção de um cone amortecido com $\phi = 35^\circ$, ver Subseção IV-C4, e um disco com especificações $a = 0,5$ e $r = 0,49$. Essa última região é escolhida como restrição do método MOESP-R.

A Fig. 2 ilustra os resultados das estimativas dos três métodos comparados entre si com a posição real dos autovalores (\star). Nota-se que o MOESP irrestrito (+) produz modelos instáveis. Além do mais, observa-se que o MOESP-RP (\ast) produz melhores estimativas do que o MOESP-R (x). Estes resultados sugerem que o uso de informação *a priori* em regiões disjuntas contribui para estimativas dos autovalores dominantes com maior precisão em comparação ao MOESP-R.

Analisando os histogramas das 200 simulações de Monte-Carlo dos autovalores estimados na Fig. 3, observa-se o par de

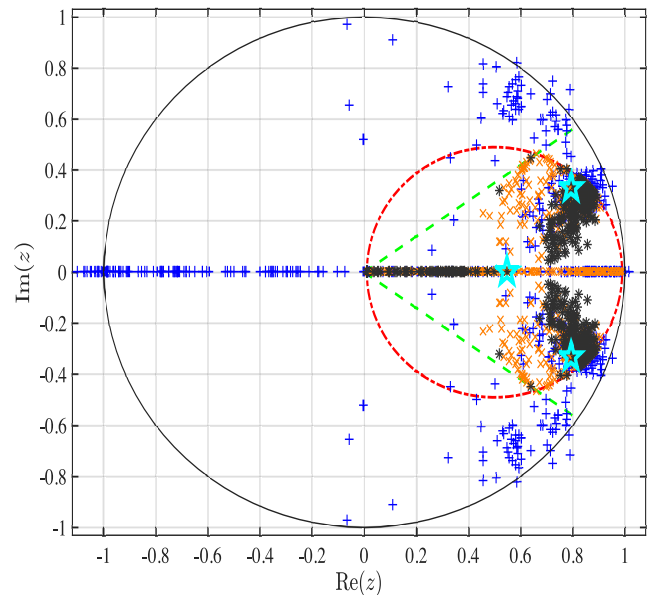


Fig. 2. Posições dos autovalores no plano- z . A posição verdadeira dos autovalores (\star) é $z_{1,2} = 0,79 \pm 0,33j$ e $z_3 = 0,55$. Em (+) têm-se as estimativas do MOESP irrestrito. A região LMI do MOESP-R é representada pela interseção do disco (—) com o cone amortecido (—), cujos autovalores estimados são dados por (x). A região LMI da etapa 1 do MOESP-RP é limitada pelo disco (—), enquanto a região LMI da etapa 2 do MOESP-RP é a mesma do MOESP-R. As estimativas do MOESP-RP são dadas por (\ast).

autovalores conjugados ($z_{1,2}$) e o polo real z_3 estimados pelo MOESP-RP à esquerda e pelo MOESP-R à direita. Em ambos os métodos, os autovalores complexos dominantes estão mais próximos dos autovalores de (34). Por outro lado, no MOESP-RP os autovalores complexos apresentam menor dispersão e maior precisão se comparado ao MOESP-R. Observando-se a resposta ao degrau obtida pela média e desvio padrão dos modelos estimados pelo MOESP-PR (—) e MOESP-R (—) em comparação com a resposta do sistema (34) (—) na Fig. 4, fica evidente a menor dispersão e maior precisão do MOESP-PR observada na Fig. 3. Apesar da melhora na estimação da parcela dominante, o autovalor real não é bem estimado tanto pelo MOESP-RP quanto pelo MOESP-R. Em especial, o MOESP-RP estima o autovalor real próximo a zero. Conjectura-se que, ao impor restrição nos autovalores dominantes por meio do MOESP-RP, a estimação do autovalor real é polarizada. Acredita-se que esse é o preço que se paga ao melhorar a estimação dos autovalores dominantes com restrição.

B. Sistema de Quarta Ordem

Considere o sistema representado por

$$F(s) = \frac{100}{s^4 + 10s^3 + 50s^2 + 80s + 100}, \quad (35)$$

cujos autovalores da matriz A correspondente do espaço de estados são $s_{1,2} = -4,17 \pm 3,97j$ e $s_{3,4} = -0,83 \pm 1,53j$; ver legenda da Fig. 5 com os autovalores do plano- s mapeados no plano- z . O sinal de saída do sistema é medido com uma taxa de amostragem $T_s = 0,1s$ e contaminado com ruído Gaussiano

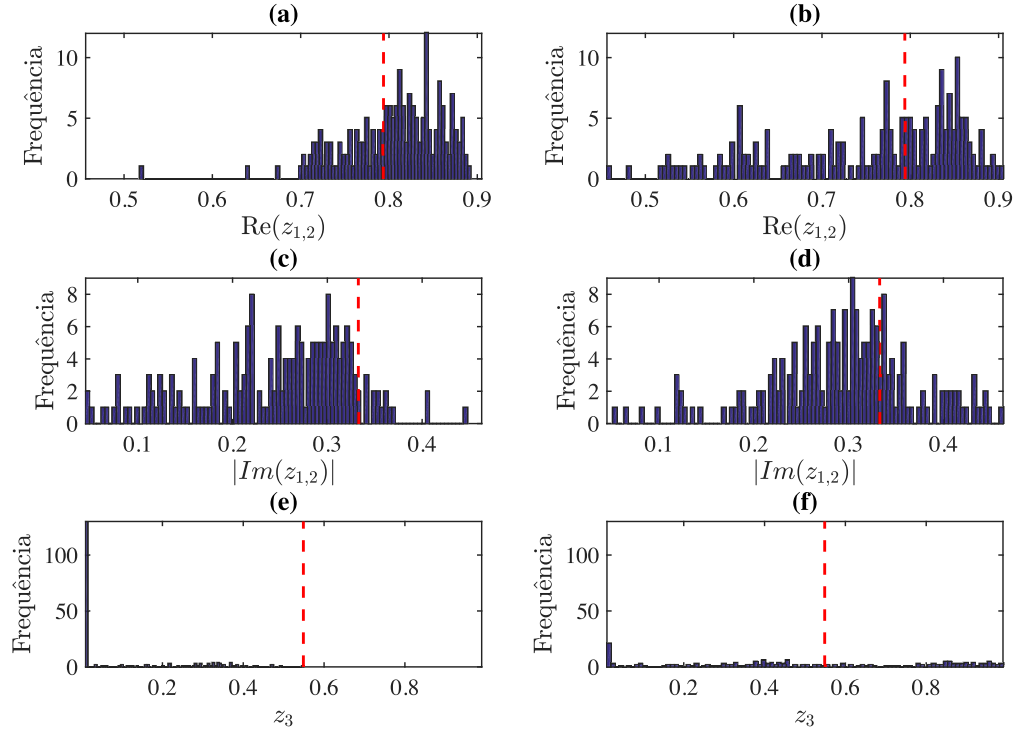


Fig. 3. Histogramas dos autovalores dos modelos estimados pelo MOESP-RP (à esquerda) e MOESP-R (à direita) para todas as 200 simulações de Monte Carlo. Em (a), (c) e (e) são apresentadas a parte real e imaginária do autovalor complexo conjugado e o autovalor real estimado pelo MOESP-RP. Em (b), (d) e (f) são apresentadas a parte real e imaginária do autovalor complexo conjugado e o autovalor real estimado pelo MOESP-R. Em vermelho apresentam-se os autovalores verdadeiros do sistema simulado (34).

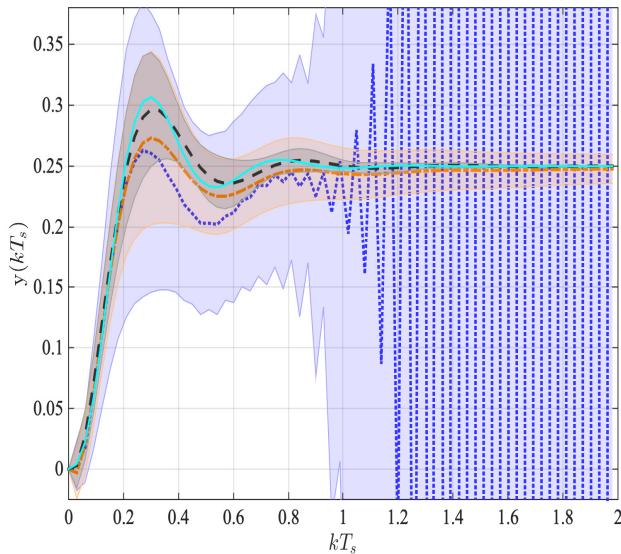


Fig. 4. Em (-) apresenta-se a saída $y(kT_s)$ obtida por meio da simulação para uma entrada ao degrau unitário em (34) sem ruído. A média de 200 realizações das respostas ao degrau são apresentadas em (- · -) para o MOESP-R, em (- -) para o MOESP-RP e em (· ·) para o MOESP irrestrito. Os respectivos desvios padrão são mostrados pelas regiões sombreadas.

ν com média nula e desvio padrão $\sigma_\nu = 0,1$. Na identificação do sistema, aplica-se um sinal de entrada PRBS de 16 bits

com período entre bits de 6 amostras durante 15s. São estudadas 200 execuções com realizações de ruído diferentes. A informação *a priori* é obtida com o mesmo procedimento apresentado na Subseção VI-A, resultando em ganho $K = 1$ e autovalores dominantes $s_{1,2}^{priori} = -0,73 \pm 1,36j$ [14].

Para a estimação dos modelos, a metodologia MOESP-RP utiliza regiões LMIs distintas em cada uma das etapas. Na etapa 1, a região LMI é definida à direita do plano- z para $\sigma = 0,05$, ver Subseção IV-C1, e limitada por uma circunferência com $a = 0$ e $r = 0,99$, ver Subseção IV-C3. A região escolhida para a etapa 2 utiliza informação *a priori* na construção das restrições LMI. Dessa forma, a região referente a informação auxiliar disponível é dada pela interseção de uma circunferência com $a = 0$ e $r = 0,99$, um cone amortecido com $\phi = 15^\circ$, ver Subseção IV-C4, e a região à direita do plano- z em $\sigma = 0,8$. Essa última região LMI é escolhida como restrição do método MOESP-R.

A Fig. 5 apresenta os resultados das estimativas dos três métodos comparados entre si com a posição verdadeira dos autovalores (\star). O método MOESP irrestrito (+) apresenta estimativas instáveis. Como pode ser notado, na etapa 1 do MOESP-RP (\ast) as estimativas dos autovalores não dominantes ficam dispersas. Contudo, as estimativas da etapa 2, autovalores dominantes, são menos dispersas ao redor do autovalor dominantes de (35) em comparação ao MOESP-R (\times).

Esse resultado é confirmado pela análise das repostas ao degrau unitário do sistema (35) (-) na Fig. 6 obtidas para os 200 modelos estimados por cada um dos três métodos,

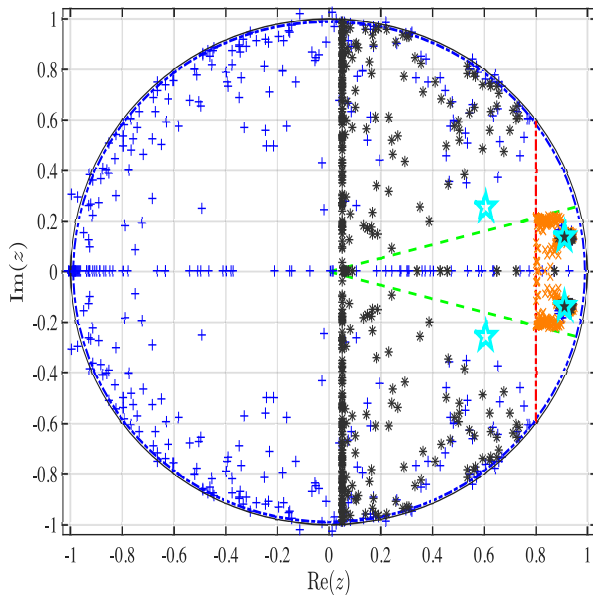


Fig. 5. Posições dos autovalores no plano- z . A posição verdadeira dos autovalores ($*$) é $z_{1,2} = 0,91 \pm 0,14j$ e $z_{3,4} = 0,61 \pm 0,25j$. Em (+) têm-se as estimativas do MOESP irrestrito. A região LMI do MOESP-R é representada pela interseção do plano a direita da reta (---) com o interior do cone amarelo (---) e limitada pelo interior do disco (---), cujos autovalores estimados são dados por (x). A região LMI da etapa 2 do MOESP-RP é a mesma do MOESP-R. As estimativas do MOESP-RP são dadas por (*).

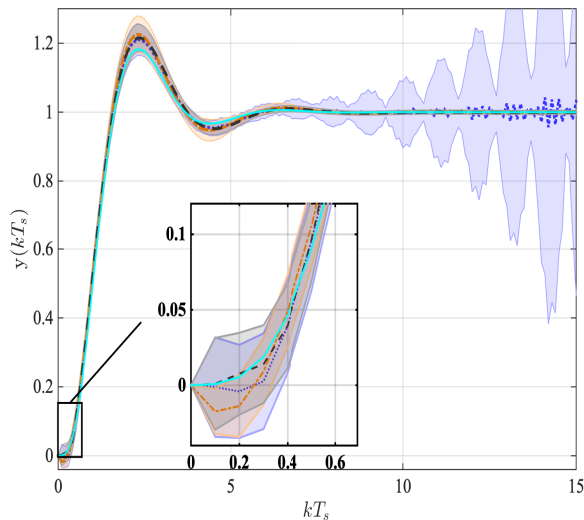


Fig. 6. Em (-) apresenta-se a saída $y(kT_s)$ obtida por meio da simulação para uma entrada ao degrau em (35) sem ruído. A média de 200 realizações das respostas ao degrau são apresentadas em (---) para o MOESP-R, em (---) para o MOESP-RP e em (---) para o MOESP irrestrito. Os respectivos desvios padrão são mostrados pelas regiões sombreadas.

a saber: MOESP (---), MOESP-R (---) e MOESP-RP (---). Apesar do resultado do MOESP-RP em média ser equivalente ao do MOESP-R, uma característica importante do modelo é preservada com a metodologia MOESP-RP. Analisando a parte destacada na Fig. 6, observa-se que o MOESP-RP em média apresenta modelos de fase mínima, como em (35). Por outro lado, as estimativas do MOESP-R e o MOESP em média

apresentaram modelos de fase não-mínima. Estes resultados sugerem que a metodologia de duas etapas confere maior liberdade na estimativa do autovalores não dominantes, o que acaba contribuindo para a estimação dos zeros dos modelos. Vale ressaltar que se o modelo estimado for utilizado para aplicações em controle, a inversão de fase é característica indesejável. Um outro ponto a se destacar é que a dispersão e polarização dos autovalores não dominantes não influenciaram significativamente nas respostas ao degrau unitário obtidas pelos modelos estimados.

VII. CONCLUSÕES

Neste artigo o problema de identificação por subespaços com restrição parcial nos autovalores é tratado. Assumindo que os estados do sistema a ser identificado são desacoplados, uma metodologia de duas etapas é proposta.

Apesar de forçar a estrutura do modelo ao considerar que os estados são desacoplados, os resultados numéricos indicam que a metodologia proposta contribui para a melhora do desempenho dos modelos estimados em comparação com as metodologias irrestrita e com restrição em todos os autovalores.

É importante mencionar que o presente artigo, por propor a estimação de modelos SISO com restrição nos autovalores de forma parcial, possibilita a criação de regiões LMI que podem ser disjuntas entre si. Essa proposta tem o objetivo de lançar uma luz para novos estudos e aprofundamento de propostas na área. Com efeito, envia esforços no futuro para propor uma metodologia para estimar modelos de sistemas MIMO (do inglês, *multiple-input multiple-output*) com restrição parcial nos autovalores é de interesse.

REFERÊNCIAS

- [1] S. Prívará, J. Cigler, Z. Váňa, and L. Ferkl, "Incorporation of system steady state properties into subspace identification algorithm," *International Journal of Modelling, Identification and Control*, vol. 16, no. 2, pp. 159–167, June 2012.
- [2] A. Alenany and H. Shang, "Recursive subspace identification with prior information using the constrained least squares approach," *Computers & Chemical Engineering*, vol. 54, no. 7, pp. 174–180, 2013.
- [3] Y. Wang, L. Zhang, and Y. Zhao, "Improved closed-loop subspace identification with prior information," *International Journal of Systems Science*, vol. 49, no. 9, pp. 1821–1835, 2018.
- [4] R. A. Ricco and B. O. S. Teixeira, "Least-squares parameter estimation for state-space models with state equality constraints," *arXiv e-prints*, arXiv:1904.05178, Apr 2019.
- [5] M. V. Paula, R. A. Ricco, and B. O. S. Teixeira, "Identificação de modelos de Hammerstein em malha fechada por meio de preditores baseados em subespaços," in *XXII Congresso Brasileiro de Automática 2018*, September 2018.
- [6] A. G. Costa Júnior, J. A. Riul, and P. H. M. Montenegro, "Application of the subspace identification method using the N4SID technique for a robotic manipulator," *IEEE Latin America Transactions*, vol. 14, no. 4, pp. 1588–1593, April 2016.
- [7] L. A. Aguirre, B. O. Teixeira, B. H. Barbosa, A. F. Teixeira, M. C. Campos, and E. M. Mendes, "Development of soft sensors for permanent downhole gauges in deepwater oil wells," *Control Engineering Practice*, vol. 65, pp. 83 – 99, 2017.
- [8] B. O. S. Teixeira and L. A. Aguirre, "Using uncertain prior knowledge to improve identified nonlinear dynamic models," *Journal of Process Control*, vol. 21, no. 1, pp. 82–91, 2011.
- [9] S. L. Lacy and D. S. Bernstein, "Subspace identification with guaranteed stability using constrained optimization," *IEEE Transactions on Automatic Control*, vol. 48, no. 7, pp. 1259–1263, July 2003.

- [10] A. Alenany, H. Shang, M. Soliman, and I. Ziedan, "Improved subspace identification with prior information using constrained least squares," *IET Control Theory Applications*, vol. 5, no. 13, pp. 1568–1576, Sept 2011.
- [11] D. N. Miller and R. A. de Callafon, "Subspace identification with eigenvalue constraints," *Automatica*, vol. 49, no. 8, pp. 2468 – 2473, 2013.
- [12] L. Shang, J. Liu, K. Turksoy, Q. M. Shao, and A. Cinar, "Stable recursive canonical variate state space modeling for time-varying processes," *Control Engineering Practice*, vol. 36, no. 3, pp. 113–119, 2015.
- [13] M. Verhaegen and V. Verdult, *Filtering and system identification - a least squares approach*. USA: Cambridge University Press, 2007.
- [14] R. A. Ricco and B. O. S. Teixeira, "Mapping prior information onto LMI eigenvalue-regions for discrete-time subspace identification," *arXiv e-prints*, arXiv:1904.05959, Apr 2019.
- [15] S. Datta, D. Chakraborty, and B. Chaudhuri, "Partial pole placement with controller optimization," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 1051–1056, April 2012.
- [16] M. Chilali and P. Gahinet, "H-infinity design with pole placement constraints: an LMI approach," *Automatica*, vol. 43, no. 3, pp. 358–367, 1996.
- [17] *The MOSEK optimization toolbox for MATLAB manual. Version 7.1 (Revision 28)*, ApS MOSEK, Copenhagen, Denmark, 2015).



Bruno O. S. Teixeira é bacharel em engenharia de controle e automação (2004) e doutor em engenharia elétrica (2008) pela Universidade Federal de Minas Gerais. Atualmente é professor do Departamento de Engenharia Eletrônica na mesma instituição. Entre 2006 e 2007 realizou doutorado sanduíche no Aerospace Engineering Department na University of Michigan. Trabalha com estimação de estados e identificação de sistemas.



Rodrigo A. Ricco é bacharel em engenharia elétrica pela Universidade Federal de São João del-Rei (2010). Obteve o título de mestre em engenharia elétrica (2012) na Universidade Federal de Minas Gerais, onde atualmente é doutorando. Também é professor do Departamento de Engenharia Elétrica da Universidade Federal de Ouro Preto. Tem interesse nas áreas de dinâmica linear e não linear, controle robusto e identificação de sistemas.



Anny Verly é bacharel em engenharia elétrica pela Universidade Federal de São João del-Rei (2010). Obteve o título de mestre engenharia elétrica (2012) na Universidade Federal de Minas Gerais, onde atualmente é doutoranda. Também é professora do Departamento de Engenharia Elétrica da Universidade Federal de Ouro Preto. Tem interesse nas áreas de estimação de estados, identificação de sistemas e redes complexas.



Marcus V. Paula é Bacharel em engenharia elétrica pela Universidade Federal de Ouro Preto (2014). Obteve o título de mestre engenharia elétrica (2016) na Universidade Federal de Minas Gerais, onde atualmente é doutorando. Trabalha, desde a graduação, com temas relacionados à identificação de sistemas não lineares multivariáveis no espaço de estados.