International Journal of Control
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tcon20

On the role of subspace zeros in retrospective cost adaptive control of non-square plants

E. Dogan Sumer [a] & Dennis S. Bernstein [a]
[a] Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109-2140, USA

Accepted author version posted online: 08 Aug 2014.Published online: 18 Sep 2014.

To cite this article: E. Dogan Sumer & Dennis S. Bernstein (2015) On the role of subspace zeros in retrospective cost adaptive control of non-square plants, International Journal of Control, 88:2, 295-323, DOI: 10.1080/00207179.2014.951690

To link to this article: http://dx.doi.org/10.1080/00207179.2014.951690

International Journal of Control, 2015
Vol. 88, No. 2, 295–323, http://dx.doi.org/10.1080/00207179.2014.951690

On the role of subspace zeros in retrospective cost adaptive control of non-square plants

E. Dogan Sumer and Dennis S. Bernstein*

Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109-2140, USA

(*Received 11 April 2013; accepted 1 August 2014*)

We consider adaptive control of non-square plants, that is, plants that have an unequal number of inputs and outputs. In particular, we focus on retrospective cost adaptive control (RCAC), which is a direct, discrete-time adaptive control algorithm that is applicable to stabilisation, command following, disturbance rejection, and model reference control problems. Previous studies on RCAC have focused on control of square plants. In the square case, RCAC requires knowledge of the first non-zero Markov parameter and the non-minimum-phase (NMP) transmission zeros of the plant, if any. No additional information about the plant or the exogenous signals need be known. The goal of the present paper is to consider RCAC for non-square plants. Unlike the square case, we show that the assumption that the non-square plant is minimum phase does not guarantee closed-loop stability and signal boundedness. The main purpose of this paper is to establish the existence of time-invariant input and output subspaces corresponding to the adaptive controller. In particular, we show that RCAC implicitly squares down non-square plants through pre-/post-compensation of the non-square plant with a constant matrix. We show that, for wide plants, the control input generated by RCAC lies in a time-invariant 'input subspace', which is equivalent to pre-compensating the plant with a constant matrix. On the other hand, for tall plants, we show that the controller update is driven by the output of the plant post-compensated with a constant matrix. Accordingly, in either case, signal boundedness properties of the closed-loop system are determined by the transmission zeros of the squared system, which we call the 'subspace zeros'. To deal with NMP subspace zeros, we introduce a robustness modification, which prevents RCAC from cancelling the NMP subspace zeros.

Keywords: adaptive control; non-square systems; transmission zeros; input-subspace zeros; output-subspace zeros

1. Introduction

Non-minimum-phase (NMP) zeros limit achievable control-system performance in various ways. These limitations are manifested as constraints on the closed-loop frequency response, pole locations, and step response (Freudenberg & Looze, 1985; Hoagg & Bernstein, 2007). Analogous issues arise in discrete-time control (Tokarzewski, 2006), with the additional difficulty that sampling may give rise to NMP zeros (Astrom, Hagander, & Sternby, 1984).

Zeros in multiple-input, multiple-output (MIMO) plants can be defined in terms of either a state-space realisation (invariant zeros) or a transfer function (transmission zeros). The presence of zeros in MIMO plants implies blocking of certain input signals. Associated with the zeros of MIMO plants are zero directions, which determine the directions along which each zero affects the response of the system. The zero directions are grouped into two categories, namely input zero directions, which determine the direction along which certain inputs are blocked, and output zero directions, which give rise to directions along which the output may be difficult to control (Davison & Wang, 1974; Hoagg & Bernstein, 2007; Karcanias & Kouvaritakis, 1979; MacFarlane & Karcanias, 1976; Rosenbrock, 1970; Schrader & Sain, 1989; Skogestad & Postlethwaite, 2005).

Zeros of non-square (tall or wide) plants are considered in Davison and Wang (1974) and Kouvaritakis and MacFarlane (1976), where it is shown that non-square plants generically have no transmission zeros (Davison & Wang, 1974, Theorem 5). However, non-square plants may have zero-like properties that cannot be detected through transmission-zero blocking (Latawiec, Banka, & Tokarzewski, 2000). In addition, control techniques developed for square plants may not extend to, or may have poor performance in the non-square case. For example, since asymptotic command following is typically not achievable for tall plants (Goodwin & Sin, 1984), multi-variable adaptive control methods, including MIMO extensions of model reference adaptive control (MRAC), are formulated exclusively for square plants (Ioannou & Sun, 1996; Narendra & Annaswamy, 1989). Therefore, for non-square plants, it is often desirable to transform the plant through squaring, where the plant is pre- or post-compensated, or augmented by additional actuators/sensors, so as to create a square plant with a desired zero structure (Davison, 1983; Karcanias & Giannakopoulos, 1989; Kouvaritakis & MacFarlane, 1976;

*Corresponding author. Email: dsbaero@umich.edu

Lavretsky, 2011; Saberi & Sannuti, 1988, 1990; Vardulakis, 1980). Necessary and sufficient conditions for the solution of squaring down problem are established, and methodologies for computing static and dynamic pre- and post-compensators are given in Leventides and Karcanias (2008) and Saberi and Sannuti (1988, 1990). In Misra (1992, 1993), the opposite case of 'squaring up and zero assignment' is considered, where the underactuated plant is augmented with additional sensors (for wide plants) or actuators (for tall plants) to obtain a square, minimum-phase plant. It should be noted, however, that these squaring-based zero-assignment methods require partial or full knowledge of the plant dynamics.

In this paper, we focus on retrospective cost adaptive control (RCAC), which is a direct, discrete-time adaptive control algorithm for stabilisation, command following, and disturbance rejection problems. This approach was developed in Hoagg (2010), Hoagg, Santillo, and Bernstein (2008a), Santillo and Bernstein (2010), and Venugopal and Bernstein (2000), where it was shown that, in the square case, RCAC requires knowledge of the first nonzero Markov parameter and the NMP transmission zeros of the plant, if any. No additional information about the plant or the exogenous signals need be known. Extensions of RCAC were given in D'Amato, Sumer, and Bernstein (2011), D'Amato, Sumer, Mitchell, et al. (2011), Sumer, D'Amato, and Bernstein (2011), and Sumer and Bernstein (2012), where the need to know the NMP zeros of the square plant was removed by augmenting the retrospective cost function with a performance-dependent control penalty. As shown in D'Amato, Sumer, and Bernstein (2011) for the single-input, single-output (SISO) case, the price paid for this relaxed modelling requirement is the need to ensure that the Markov parameters used in RCAC provide a suitable approximation of the frequency response of the plant. Except for the limited investigation of RCAC for single-input, multiple-output (SIMO) and multiple-input, single-output (MISO) plants provided in Sumer and Bernstein (2012), RCAC for non-square plants has not been studied.

The goal of the present paper is to consider RCAC for non-square plants without explicitly squaring-down or squaring-up the plant. We show that, unlike the square case, the assumption that the non-square plant is minimum phase does not guarantee closed-loop stability and signal boundedness. Specifically, we show that, due to the nature of the update law, RCAC involves two implicit squaring operations: one performed by pre-compensating the plant, and the other performed by post-compensating the plant. Despite the nonlinear nature of the update law, and the fact that the controller gains are updated at every time step, we show that the pre- and post-compensation are performed through constant gains. In the wide case, pre-compensation leads to squaring-down, which incorporates additional zeros due to squaring, which we call 'input-subspace zeros'. Similarly, in the tall case, post-compensation changes the zero structure and incorporates additional zeros, which we call 'output-subspace zeros'. We show that, if the non-square plant has NMP input or output subspace zeros, then RCAC may attempt to cancel these zeros, which leads to unbounded control input in the wide case, and unbounded control input and performance output in the tall case. In light of these findings, we extend the retrospective cost function to include a leakage-modification-type (Astrom & Wittenmark, 1995; Ioannou & Sun, 1996; Narendra & Annaswamy, 1989) control penalty in order to prevent the controller from generating an unbounded control input.

We begin in Section 3 by presenting instantaneous and cumulative update laws for RCAC. In Section 4, we present numerical examples where we apply RCAC to non-square plants, and demonstrate that the convergence and stability properties shown in Hoagg (2010) and Hoagg et al. (2008a) for square plants may not hold for non-square plants. Sections 5–8 are devoted to studying the mechanisms for instability observed in Section 4. Specifically, in Section 5, we show that, if the plant is wide, then the control input generated by RCAC lies inside a time-invariant subspace that is contained within the input space. In Section 6, we derive sufficient conditions under which the adaptive controller converges. The results developed in Sections 5 and 6 indicate the presence of two static squaring operations implicitly performed by the RCAC update laws. In Sections 7 and 8, we build on the results of Sections 5 and 6 and define the subspace zeros, which are introduced due to the implicit static-squaring operations. In these sections, we revisit the examples of Section 4, and demonstrate that the plants exhibiting closed-loop instability have NMP subspace zeros that are not transmission zeros, and that the instability is caused by cancellation of these NMP subspace zeros. Finally, in Section 9, we introduce a robustness modification to prevent the unstable pole-zero cancellation. We revisit the examples of Section 4, and demonstrate that no instability is observed with the modified update laws, despite the presence of NMP subspace zeros.

2. Preliminaries and problem formulation

Consider the following MIMO discrete-time plant:

$$x(k+1) = Ax(k) + Bu(k) + D_1 w(k), \tag{1}$$

$$y(k) = Cx(k) + D_2 w(k), \tag{2}$$

$$z(k) = E_1 x(k) + E_0 w(k), \tag{3}$$

where (A, B, C) is minimal, $x(k) \in \mathbb{R}^n$ is the state vector, $z(k) \in \mathbb{R}^{l_z}$ is the measured performance variable to be minimised, $y(k) \in \mathbb{R}^{l_y}$ contains additional measurements that are available for feedback control, $u(k) \in \mathbb{R}^{l_u}$ is the control input, and $w(k) \in \mathbb{R}^{l_w}$ is the exogenous signal, which can

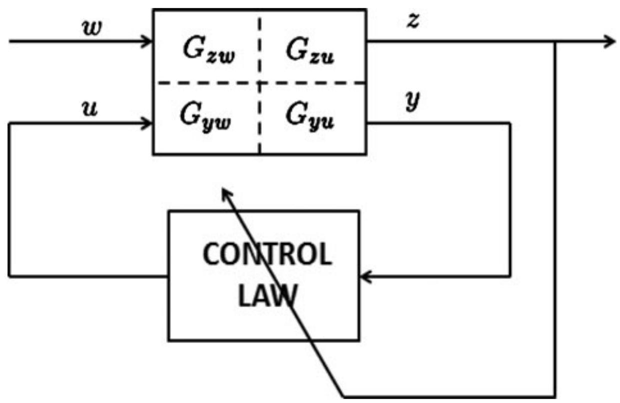

Figure 1. Adaptive output feedback control architecture.

be unknown. Throughout the paper, we write $\|\cdot\|$ to denote the Euclidean norm of a vector. The goal is to develop an adaptive output feedback controller of the form

$$u(k)=\sum_{i=1}^{n_c} M_i(k)u(k-i)+\sum_{i=1}^{n_c} N_i(k)y(k-i), \tag{4}$$

where, for all $i\in\{1,\dots,n_c\}$ and $k\geq 0$, $M_i(k)\in\mathbb{R}^{l_u\times l_u}$ and $N_i(k)\in\mathbb{R}^{l_u\times l_y}$, such that $\|z\|$ is minimised in the presence of the unknown exogenous signal w. The block diagram for Equations (1)–(4) is shown in Figure 1. The controller is activated at $k=1$ with initial conditions $M_i(k)=0$ and $N_i(k)=0$ for all $i\in\{1,\dots,n_c\}$ and $k\leq 0$. However, the controller coefficients could be initialised to the parameters of a stabilising baseline controller, if such a controller is known in advance. We rewrite the control law (4) in regressor form

$$u(k)=\theta^{\mathrm{T}}(k)\phi(k-1), \tag{5}$$

where

$$\theta(k)\triangleq\begin{bmatrix} M_1^{\mathrm{T}}(k)\\ \vdots \\ M_{n_c}^{\mathrm{T}}(k)\\ N_1^{\mathrm{T}}(k)\\ \vdots\\ N_{n_c}^{\mathrm{T}}(k)\end{bmatrix}\in\mathbb{R}^{n_c(l_u+l_y)\times l_u},$$

$$\phi(k-1)\triangleq\begin{bmatrix} u(k-1)\\ \vdots\\ u(k-n_c)\\ y(k-1)\\ \vdots\\ y(k-n_c)\end{bmatrix}\in\mathbb{R}^{n_c(l_u+l_y)}, \tag{6}$$

and $\theta(0)=0$.

The components of the exogenous signal w can represent either command signals to be followed, external disturbances to be rejected, or both. For instance, if $D_1=0$ and $D_2\neq 0$, then the objective is to have the output Cx follow the command signal $-D_2w$. On the other hand, if $D_1\neq 0$ and $D_2=0$, then the objective is to reject the disturbance w from the performance variable z. In this case, we say that w is a *matched disturbance* if $\mathcal{R}(D_1)\subseteq\mathcal{R}(B)$, where $\mathcal{R}(\cdot)$ denotes range, and w is an *unmatched disturbance* if it is not matched. Furthermore, if $D_1=[\,\hat{D}_1\ 0\,]$, $D_2=[\,0\ \hat{D}_2\,]$, and $w=[\,w_1\ w_2\,]^{\mathrm{T}}$, then the objective is to have Cx follow the command $-\hat{D}_2w_2$ while rejecting the disturbance $\hat{D}_1w_1$. Finally, if D_1 and D_2 are empty matrices, then the objective is to achieve $z(k)\to 0$ as $k\to\infty$ with no exogenous signals.

Throughout this paper, we consider a specialisation of Equations (1)–(3) at which the performance variable z is the only measurement available for feedback, that is, $y=z$. Therefore, from now on, y is used to denote the performance variable, which, together with the past values of u, is the only measurement used for feedback control. In light of this specialisation, the open-loop system (1)–(3) is characterised by the transfer matrix

$$y=G\begin{bmatrix}u\\ w\end{bmatrix}, \tag{7}$$

where

$$G\triangleq[\,G_{yu}\ \ G_{yw}\,]\in\mathbb{R}^{l_y\times(l_u+l_w)}(z) \tag{8}$$

and

$$G_{yu}(z)=C(zI-A)^{-1}B\in\mathbb{R}^{l_y\times l_u}(z), \tag{9}$$

$$G_{yw}(z)=C(zI-A)^{-1}D_1+D_2\in\mathbb{R}^{l_y\times l_w}(z). \tag{10}$$

We write $G_{yu}\sim\left[\begin{array}{c|c}A&B\\ \hline C&0\end{array}\right]$ to denote a *realisation* of G_{yu}, and $G_{yu}\overset{\min}{\sim}\left[\begin{array}{c|c}A&B\\ \hline C&0\end{array}\right]$ to denote a *minimal realisation* of G_{yu}. If $l_y=l_u$, then G_{yu} is *square*, whereas, if $l_y\neq l_u$, then G_{yu} is *non-square*. In particular, if $l_y>l_u$, then G_{yu} is *tall*, whereas, if $l_y<l_u$, then G_{yu} is *wide*.

Definition 2.1: Let $F\in\mathbb{R}^{l_y\times l_u}(z)$ be a rational transfer matrix, and, for all $i\in\{1,\dots,l_y\}$ and $j\in\{1,\dots,l_u\}$, let $F_{ij}(z)=p_{ij}(z)/q_{ij}(z)$, where q_{ij} is not the zero polynomial, and $p_{ij}(z),q_{ij}(z)\in\mathbb{R}(z)$ are coprime. Then, the poles of F are the elements of the set

$$\mathrm{poles}(F)\triangleq\bigcup_{i,j=1}^{l_y,l_u}\mathrm{roots}(q_{ij}(z)),$$

and the *normal rank* of F is the non-negative integer

$$\text{normal rank } F \triangleq \max_{z\in\mathbb{C}\backslash\text{poles}(F)} \text{rank } F(z).$$

Now, define the *Rosenbrock system matrix* of $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ by

$$\Sigma(z) \triangleq \begin{bmatrix} zI - A & B \\ C & 0 \end{bmatrix} \in \mathbb{R}^{(n+l_y)\times(n+l_u)}[z]. \tag{11}$$

The *transmission zeros* of G_{yu} are the elements of the set

$$\text{tzeros}(G_{yu}) \triangleq \{\zeta \in \mathbb{C} : \text{ rank } \Sigma(\zeta) < \text{normal rank } \Sigma\}. \tag{12}$$

Definition 2.2: Let $\zeta \in \text{tzeros}(G_{yu})$.

(i) If $|\zeta| \geq 1$, then ζ is an *NMP transmission zero* of G_{yu}.
(ii) If $|\zeta| < 1$, then ζ is a *minimum-phase transmission zero* of G_{yu}.
(iii) If G_{yu} has at least one NMP transmission zero, then G_{yu} is NMP.
(iv) If G_{yu} is not NMP, then G_{yu} is *minimum phase*.

Expanding G_{yu} for $|z| > \rho(A)$ yields the Laurent series

$$G_{yu}(z) = \sum_{i=1}^{\infty} H_i z^{-i} = \sum_{i=d}^{\infty} H_i z^{-i},$$

where $\rho(A)$ is the spectral radius of A, $H_i \triangleq CA^{i-1}B \in \mathbb{R}^{l_y\times l_u}$ is the ith Markov parameter of G_{yu}, and d is the *relative degree*, which is the smallest integer i such that H_i is non-zero. Throughout this paper, we assume that H_d has full rank, that is, rank $H_d = \min(l_u, l_y)$. For $j \in \{1, \ldots, l_y\}$ and $l \in \{1, \ldots, l_u\}$, let $H_{i,(j,l)}$ denote the (j, l) entry of H_i. Then, it follows from Equations (1) and (3) that $H_{i,(j,l)}$ is the impulse response at $k = i$ of the jth output of G_{yu} for a unit impulse at $k = 0$ applied to the lth input of G_{yu}. In practice, G_{yu} may be a sampled-data plant arising from a strictly proper continuous-time plant $T_{yu} \in \mathbb{R}^{l_y\times l_u}(s)$ under sample and hold operations with sampling period h. In this case, the relative degree of G_{yu} is generically equal to 1. In particular, for a zero-order hold, $d > 1$, if and only if, for all $i \in \{1, \ldots, l_y\}$ and $j \in \{1, \ldots, l_u\}$, the step response of $T_{yu,(i,j)}$ at $t = h$ is zero. Aside from a lower bound on the required controller order to facilitate internal model control (Hoagg et al., 2008a; Hoagg, Santillo, & Bernstein, 2008b), d and H_d are the only modelling information assumed to be available for controller synthesis. In particular, no modelling information about the poles and zeros of G_{yu} is required, and no knowledge of the amplitude, phase, or spectrum of the harmonic signal w is required. If additional Markov parameters are known, then this extra modelling information can be used to improve transient and steady-state performance, as shown in Santillo and Bernstein (2010). Furthermore, for square plants with NMP transmission zeros, knowledge of one Markov parameter is not sufficient to obtain the stability and convergence results shown in Hoagg et al. (2008a) and Hoagg (2010). However, in this paper, we focus on non-square plants, which typically have no transmission zeros. Therefore, we confine our analysis to the case where only one Markov parameter is used by the adaptive control algorithm.

In the non-adaptive case, a sufficient condition for command following and disturbance rejection is the output controllability condition (Goodwin & Sin, 1984, Lemma 5.2.2; Rosenbrock, 1970),

$$\text{normal rank } G_{yu} = l_y, \tag{13}$$

which is not satisfied if G_{yu} is tall. Furthermore, if G_{yu} is wide, it follows from Equation (13) that $l_u - l_y$ control inputs can be discarded without degrading command following and disturbance rejection performance. Therefore, multivariable MRAC is often formulated based on the assumption that G_{yu} is square (Goodwin & Sin, 1984; Hoagg et al., 2008a; Ioannou & Sun, 1996). In practice, this set-up may incorporate a 'squaring problem', where, given the input matrix B, the output matrix C (or vice versa) is chosen so that G_{yu} is minimum phase (Lavretsky, 2011). Solving the squaring problem may require partial or full knowledge about the matrices A, B, and C.

In the present paper, we focus on the case where G_{yu} is non-square. Since Equation (13) is not a necessary condition for achieving asymptotic disturbance rejection or command following, there exist special cases where asymptotic command following and/or disturbance rejection is achievable with tall G_{yu}. For example, in a matched disturbance rejection problem, since $\mathcal{R}(D_1) \subseteq \mathcal{R}(B)$, u can be chosen to cancel w from Equation (1). Another example is the case where the individual performance outputs are chosen compatibly for a command-following problem, for example, a step command-following problem with $y(k) = [\, y_0(k) - w \;\; y_0(k) - y_0(k-1)\,]^{\mathrm{T}}$. On the other hand, for wide plants, it follows from Equation (13) that both asymptotic command following and asymptotic disturbance rejection are achievable. Therefore, if a control technique is applicable to wide plants, there is no reason to discard control inputs to obtain a square plant.

3. Update laws based on retrospective cost optimisation

In this section, we present two RCAC update laws for the controller $\theta(k)$ in Equation (5). For convenience, we rewrite the control law (5) as

$$u(k) = \Phi(k-1)\Theta(k), \tag{14}$$

where

$$\Phi(k-1) \triangleq I_{l_u} \otimes \phi^{\mathrm{T}}(k-1) \in \mathbb{R}^{l_u \times l_u n_c (l_u + l_y)}, \tag{15}$$

$$\Theta(k) \triangleq \mathrm{vec}(\theta(k)) \in \mathbb{R}^{l_u n_c (l_u + l_y)}, \tag{16}$$

'⊗' denotes the Kronecker product, and 'vec' is the column-stacking operator (Bernstein, 2009). Note that $\Theta(0) = 0$.

3.1 *Retrospective performance*

For $k \geq 0$, we define the *retrospective performance variable*

$$\hat{y}(\hat{\Theta}, k) \triangleq y(k) + H_d \Phi(k-d-1)[\hat{\Theta} - \Theta(k-d)] \tag{17}$$

$$= y(k) - H_d u(k-d) + H_d \Phi(k-d-1)\hat{\Theta}, \tag{18}$$

where $\hat{\Theta}$ is an optimisation variable. To understand the meaning of $\hat{y}(\hat{\Theta}, k)$, note that it follows from Equations (1) and (3) that

$$\begin{aligned} y(k) &= CA^d x(k-d) + \sum_{i=1}^{d} H_i u(k-i) + D_2 w(k) \\ &\quad + \sum_{i=1}^{d} CA^{i-1} D_1 w(k-i) \\ &= CA^d x(k-d) + H_d \Phi(k-d-1)\Theta(k-d) \\ &\quad + D_2 w(k) + \sum_{i=1}^{d} CA^{i-1} D_1 w(k-i). \end{aligned} \tag{19}$$

Replacing $\Theta(k-d)$ by $\hat{\Theta}$ in Equation (19) yields

$$\begin{aligned} &CA^d x(k-d) + H_d \Phi(k-d-1)\hat{\Theta} + D_2 w(k) \\ &\quad + \sum_{i=1}^{d} CA^{i-1} D_1 w(k-i) \\ &= y(k) - H_d \Phi(k-d-1)[\Theta(k-d) - \hat{\Theta}] \\ &= \hat{y}(\hat{\Theta}, k). \end{aligned} \tag{20}$$

Note that $x(k-d)$ and $\Phi(k-d-1)$ in Equation (19) are independent of $\Theta(k-d)$, and thus are unchanged, if $\hat{\Theta}$ were used instead of $\Theta(k-d)$. Consequently, it follows from Equation (20) that the retrospective performance variable $\hat{y}(\hat{\Theta}, k)$ is the performance output that would have been obtained at time k if the controller $\hat{\Theta}$ had been used in place of $\Theta(k-d)$.

We now formulate two update laws based on $\hat{y}(\hat{\Theta}, k)$. In both cases, a quadratic cost function that depends on $\hat{y}(\hat{\Theta}, k)$ is minimised with respect to $\hat{\Theta}$.

3.2 *Instantaneous update law*

For each $k \geq 1$, we define the *instantaneous cost function*

$$\begin{aligned} J_{\mathrm{ins}}(\hat{\Theta}, k) &\triangleq \hat{y}^{\mathrm{T}}(\hat{\Theta}, k)\hat{y}(\hat{\Theta}, k) + \mu[\hat{\Theta} - \Theta(k-1)]^{\mathrm{T}} \\ &\quad \cdot [\hat{\Theta} - \Theta(k-1)], \end{aligned} \tag{21}$$

where $\mu > 0$ weighs the distance between $\hat{\Theta}$ and the controller $\Theta(k-1)$ used at step $k-1$. Substituting Equation (18) into Equation (21) yields

$$J_{\mathrm{ins}}(\hat{\Theta}, k) = \hat{\Theta}^{\mathrm{T}} \Gamma_1(k) \hat{\Theta} + \Gamma_2^{\mathrm{T}}(k)\hat{\Theta} + \Gamma_3(k), \tag{22}$$

where

$$\begin{aligned} \Gamma_1(k) &\triangleq \Phi^{\mathrm{T}}(k-d-1) H_d^{\mathrm{T}} H_d \Phi(k-d-1) \\ &\quad + \mu I \in \mathbb{R}^{l_u n_c (l_u + l_y) \times l_u n_c (l_u + l_y)}, \end{aligned} \tag{23}$$

$$\begin{aligned} \Gamma_2(k) &\triangleq 2\Phi^{\mathrm{T}}(k-d-1) H_d^{\mathrm{T}} [y(k) - H_d u(k-d)] \\ &\quad - 2\mu\Theta(k-1) \in \mathbb{R}^{l_u n_c (l_u + l_y)}, \end{aligned} \tag{24}$$

and $\Gamma_3(k) \in \mathbb{R}$. Since $\Gamma_1(k)$ is positive definite, $J_{\mathrm{ins}}(\hat{\Theta}, k)$ has the unique global minimiser

$$\Theta(k) = -\frac{1}{2}\Gamma_1^{-1}(k)\Gamma_2(k), \tag{25}$$

which is the instantaneous RCAC update law. Note that the only modelling information required to implement Equation (25) is the knowledge of d and H_d. We write Equation (25) in recursive form as follows.

Lemma 3.1: *For each $k \geq 1$, the unique global minimiser of the instantaneous cost function* (21) *is given by*

$$\begin{aligned} \Theta(k) &= \Theta(k-1) - \Phi^{\mathrm{T}}(k-d-1) H_d^{\mathrm{T}} \Psi^{-1}(k) \\ &\quad \cdot \hat{y}(\Theta(k-1), k), \end{aligned} \tag{26}$$

where

$$\Psi(k) \triangleq \mu I_{l_y} + H_d \Phi(k-d-1)\Phi^{\mathrm{T}}(k-d-1) H_d^{\mathrm{T}}. \tag{27}$$

Proof: See Appendix 1. □

3.3 *Cumulative update law*

For each $k \geq 1$, we define the *cumulative cost function*

$$J_{\mathrm{cum}}(\hat{\Theta}, k) \triangleq \sum_{i=1}^{k} \hat{y}^{\mathrm{T}}(\hat{\Theta}, i)\hat{y}(\hat{\Theta}, i) + \hat{\Theta}^{\mathrm{T}} P_0^{-1}\hat{\Theta}, \quad (28)$$

where $P_0 \in \mathbb{R}^{l_u n_c(l_u+l_y)\times l_u n_c(l_u+l_y)}$. Throughout the paper, we assume that $P_0 = \beta I$, where β is a positive constant. Substituting Equation (18) into Equation (28) yields

$$J_{\mathrm{cum}}(\hat{\Theta}, k) = \hat{\Theta}^{\mathrm{T}}\mathcal{C}_1(k)\hat{\Theta} + \mathcal{C}_2^{\mathrm{T}}(k)\hat{\Theta} + \mathcal{C}_3(k), \quad (29)$$

where

$$\mathcal{C}_1(k) \triangleq \sum_{i=1}^{k} \Phi^{\mathrm{T}}(i-d-1)H_d^{\mathrm{T}} H_d \Phi(i-d-1) + P_0^{-1}, \quad (30)$$

$$\mathcal{C}_2(k) \triangleq \sum_{i=1}^{k} 2\Phi^{\mathrm{T}}(i-d-1)H_d^{\mathrm{T}}[y(i) - H_d u(i-d)], \quad (31)$$

and $\mathcal{C}_3(k) \in \mathbb{R}$. Defining $\mathcal{C}_1(0) \triangleq P_0^{-1}$ and $\mathcal{C}_2(0) \triangleq 0$, we can rewrite Equations (30) and (31) in the recursive form

$$\mathcal{C}_1(k) = \mathcal{C}_1(k-1) + \Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}} H_d \Phi(k-d-1), \quad (32)$$

$$\mathcal{C}_2(k) = \mathcal{C}_2(k-1) + 2\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}} \cdot [y(k) - H_d u(k-d)]. \quad (33)$$

Since $\mathcal{C}_1(k)$ is positive definite, $J_{\mathrm{cum}}(\hat{\Theta}, k)$ has the unique global minimiser

$$\Theta(k) = -\frac{1}{2}\mathcal{C}_1^{-1}(k)\mathcal{C}_2(k), \quad (34)$$

which is the cumulative RCAC update law. As in the case of the instantaneous controller update (25), the only modelling information required to implement Equation (34) is the knowledge of d and H_d. We write Equation (34) in recursive form as follows. Recall that $\Theta(0) = 0$.

Lemma 3.2: *For all $k \geq 0$, define $P(k) \triangleq \mathcal{C}_1^{-1}(k)$. Then, for all $k \geq 1$, $P(k)$ satisfies*

$$P(k) = P(k-1) - P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k) \cdot H_d\Phi(k-d-1)P(k-1), \quad (35)$$

where

$$\Lambda(k) \triangleq I_{l_y} + H_d\Phi(k-d-1)P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}. \quad (36)$$

Furthermore, for each $k \geq 1$, let $\Theta(k)$ be the unique global minimiser of the cumulative cost function (28) *given by Equation* (34). *Then, for all $k \geq 1$,*

$$\Theta(k) = \Theta(k-1) - P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k) \cdot \hat{y}(\Theta(k-1), k). \quad (37)$$

Proof: See Appendix 2. □

4. Adaptive control of non-square plants: motivating examples

Under suitable assumptions on w and G_{yu}, it is shown in Hoagg et al. (2008a) for the instantaneous update law (26) that $\lim_{k\to\infty} y(k) = 0$, and u, Θ, and x are bounded. In particular, it is assumed in Hoagg et al. (2008a) that G_{yu} is minimum phase and square, H_d is non-singular, and w is a harmonic signal with unknown spectrum. These convergence results are extended to the cumulative update law (34) in Hoagg (2010). We now demonstrate that these properties may or may not hold, if G_{yu} is non-square.

4.1 *Examples with wide plants*

Example 4.1 (2×3 wide plant, convergent output, bounded control)**:** Consider Equations (1) and (2) with

$$A = \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 \\ 0 & 0 & 0.4 & -0.4 \\ 0 & 0 & 0.4 & 0.4 \end{bmatrix},$$

$$B = \begin{bmatrix} -0.8 & 1.35 & -0.85 \\ 1.02 & -0.22 & -1.12 \\ -0.13 & -0.59 & 2.53 \\ -0.71 & -0.29 & 1.66 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad (38)$$

$$C = \begin{bmatrix} 0.31 & -0.87 & 0.79 & -2.33 \\ -1.26 & -0.18 & -1.33 & -1.45 \end{bmatrix},$$

$$D_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad (39)$$

where (A, B, C) is minimal. We consider the harmonic disturbance $w(k) = \sin\frac{\pi}{5}k$. Since $[\,D_1\ B\,]$ is non-singular, it follows that D_1 is not an element of $\mathcal{R}(B)$, and thus w is unmatched. The plant G_{yu} has no transmission zeros. We let $n_c = 6$, and apply the cumulative update (34) with $P_0 = I$. As shown in Figure 2, y approaches zero, u is bounded, and Θ is bounded. □

Example 4.2 (2×3 wide plant, unbounded control)**:** Consider (1) and (2) where the matrices A, B, D_1, C, and D_2 are as in Equations (38) and (39) except that $B_{(1,1)} = -1.8$. Note that (A, B, C) is minimal. We consider the same harmonic disturbance as in Example 4.1. Since $[\,B\ D_1\,]$ is non-singular, it follows that D_1 is not an element of $\mathcal{R}(B)$,

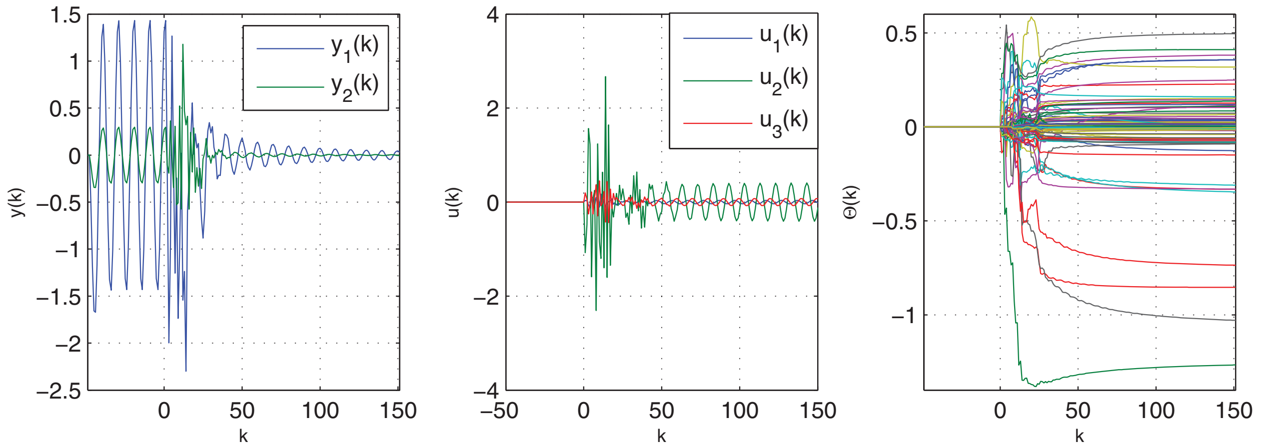

Figure 2. Example 4.1: unmatched disturbance rejection for the minimum-phase, 2×3 wide plant (38) and (39). The performance output y approaches zero, the control signal u is bounded, and the controller Θ converges.

and thus w is unmatched. The plant G_{yu} has no transmission zeros. We let $n_c = 6$ and apply the cumulative update (34) with $P_0 = I$. As shown in Figure 3, u grows without bound, while y approaches zero.

We revisit Examples 4.1 and 4.2 in Section 7.3.1 to investigate the mechanics behind the instability observed in Example 4.2. These examples are further revisited in Section 9 to demonstrate a modified RCAC algorithm which does not exhibit the instability observed in Example 4.2.

4.2 *Examples with tall plants*

Example 4.3 (3 × 1 tall plant, matched disturbance, no instability observed)**:** Consider Equations (1) and (2) with

$$A = \begin{bmatrix} 0.3 & -0.16 & 0 \\ 0.125 & 0 & 0 \\ 0 & 0.125 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 16 \\ 0 \\ 0 \end{bmatrix}, \; D_1 = B, \tag{40}$$

$$C = \begin{bmatrix} -0.0625 & -0.35 & -0.48 \\ 0.0625 & -1.5 & 9.16 \\ 0.1875 & -0.75 & -7.92 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{41}$$

where (A, B, C) is minimal. In this example, we consider the special case of matched disturbance. Since u can be used to directly cancel w from Equation (1), asymptotic disturbance rejection is achievable for tall plants in the case where w is matched with the input. We consider the two-tone harmonic disturbance $w(k) = \sin \frac{2\pi}{7} k + \sin \frac{\pi}{5} k$. The plant has no transmission zeros. We let $n_c = 7$ and apply the instantaneous update (25) with $\mu = 20$. As shown in Figure 4, all signals are bounded, Θ converges, and y approaches zero. Therefore, RCAC drives all three outputs to zero using only one control input, despite the sinusoidal disturbance.

Example 4.4 (3 × 1 tall plant, unmatched disturbance, no instability observed)**:** Consider Equations (1) and (2) where

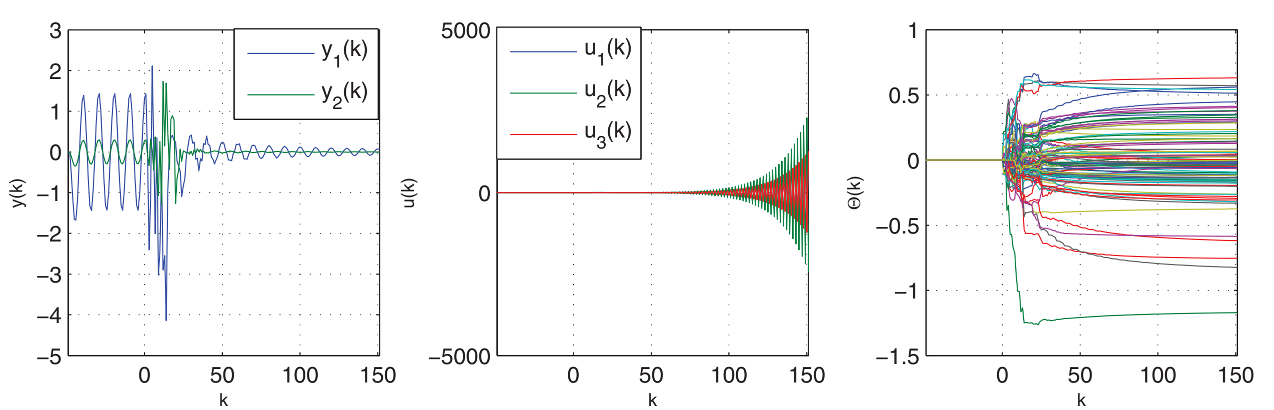

Figure 3. Example 4.2: unmatched disturbance rejection for the 2×3 wide plant (38) and (39) except that $B_{(1,1)} = -1.8$. Although G_{yu} is minimum phase, the control signal u grows without bound, while the performance output y approaches zero. The controller Θ converges.

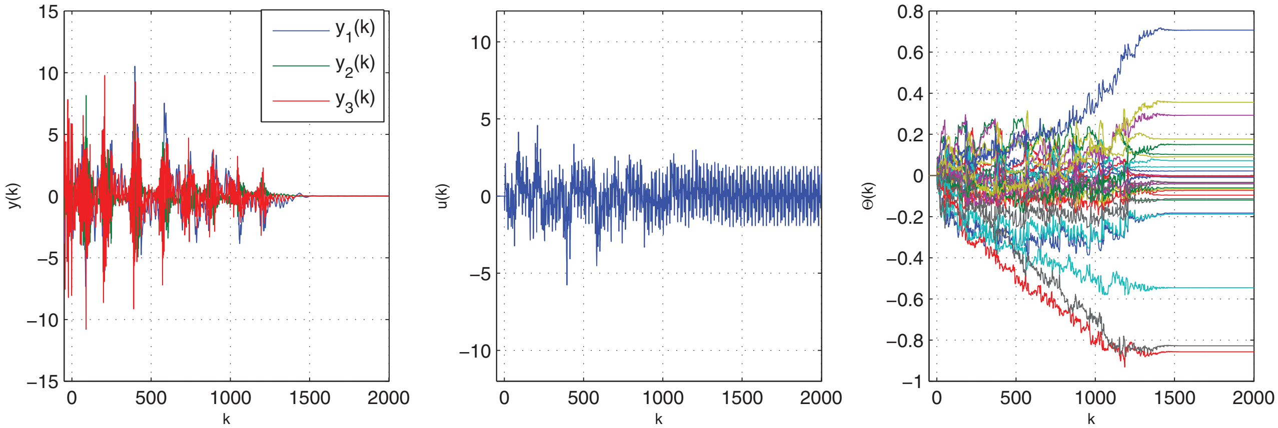

Figure 4. Example 4.3: Matched disturbance rejection for the minimum phase, 3×1 tall plant (40) and (41). The controller Θ converges, u is bounded, and y asymptotically approaches zero. RCAC cancels the matched sinusoidal disturbance from three outputs using only one actuator.

the matrices A, B, C, and D_2 are as in Equations (40) and (41), but D_1 is now given by $D_1 = [\,0\ 1\ 0\,]^{\mathrm{T}}$. We consider the same two-tone harmonic disturbance as in Example 4.3; however, since $[\,B\ \ D_1\,]$ is non-singular, it follows that D_1 is not an element of $\mathcal{R}(B)$, and thus w is now an unmatched disturbance. The plant G_{yu} has no transmission zeros. We let $n_{\mathrm{c}} = 7$ and apply the instantaneous update (25) with $\mu = 20$. As shown in Figure 5, all signals are bounded and Θ converges, but, since the plant is underactuated and the disturbance is unmatched, y does not converge to zero.

Example 4.5 (3×1 tall plant, unmatched disturbance, input and output grow without bound)**:** Consider Equations (1) and (2) where the matrices A, B, C, D_1, and D_2 are as in Example 4.4, except that $C_{(1,2)} = 0.6$. Note that (A, B, C) is minimal. We consider the same unmatched harmonic disturbance as in Example 4.4. The plant G_{yu} has no transmission zeros. We let $n_{\mathrm{c}} = 7$ and apply the instantaneous update (25) with $\mu = 20$. As shown in Figure 6, Θ converges, and both u and y grow without bound. □

We revisit Examples 4.4 and 4.5 in Section 8.3.1 to investigate the mechanics behind the instability observed in Example 4.5. These examples are further revisited in Section 9 to demonstrate a modified RCAC algorithm which does not exhibit the instability observed in Example 4.5.

5. Input subspace with retrospective cost adaptive control

We first consider the instantaneous update law (25), which is equivalent to Equations (26) and (27) as shown by Lemma 3.1. We require the following technical lemma.

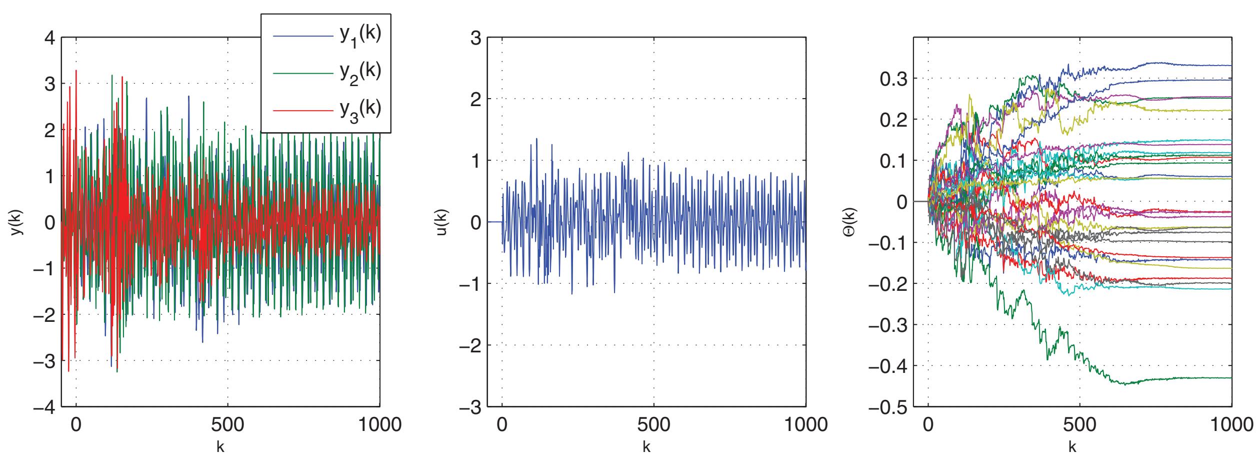

Figure 5. Example 4.5: Unmatched disturbance rejection for the minimum-phase, 3×1 tall plant of Example 4.3 with $D_1 = [0\ 1\ 0]^{\mathrm{T}}$. The controller Θ converges, and the signals u and y are bounded. The performance output y does not converge to zero due to the infeasibility of the unmatched disturbance rejection problem in the tall case.

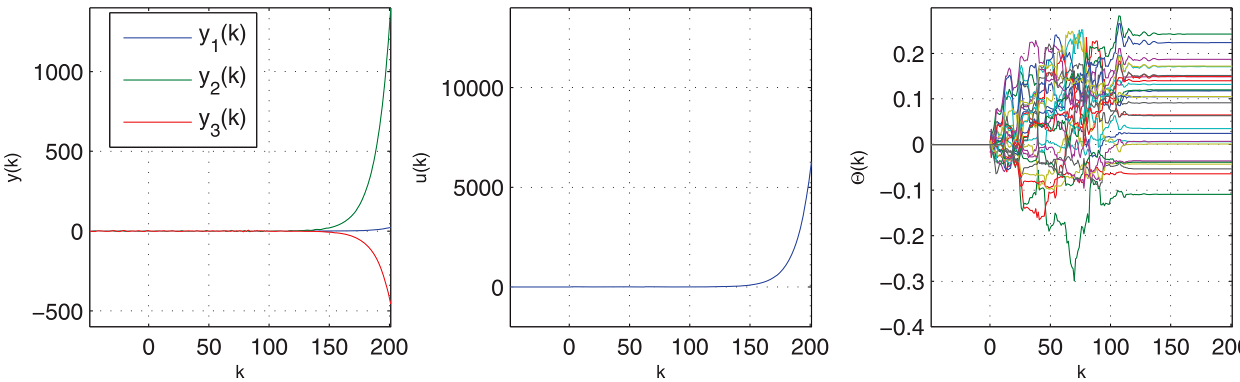

Figure 6. Example 4.5: Unmatched disturbance rejection for the 3×1 tall plant of Example 4.4 except that $C_{(1,2)} = 0.6$. Although G_{yu} is minimum phase, the signals u and y grow without bound. The controller Θ converges.

Lemma 5.1: *Let $\Theta(k)$ be given by the instantaneous update laws* (26) *and* (27), *let $\phi \in \mathbb{R}^{n_c(l_u+l_y)}$, and define $\Phi \stackrel{\triangle}{=} I_{l_u} \otimes \phi^{\mathrm{T}}$. Then, for all $k \geq 1$,*

$$\Phi\Theta(k) \in \mathcal{R}(H_d^{\mathrm{T}}). \tag{42}$$

Proof: See Appendix 3. □

We can now state the main result of this section for the instantaneous update law (25).

Theorem 5.2: *For all $k \geq 1$, let the control input $u(k)$ be given by the control law* (14) *with the instantaneous update laws* (26) *and* (27). *Then, for all $k \geq 1$,*

$$u(k) \in \mathcal{R}(H_d^{\mathrm{T}}). \tag{43}$$

Proof: For all $k \geq 1$, $\Phi(k-1) = I_{l_u} \otimes \phi^{\mathrm{T}}(k-1)$, where $\phi(k-1) \in \mathbb{R}^{n_c(l_u+l_y)}$. Using Lemma 5.1, it follows from Equations (14) and (42) that, for all $k \geq 1$, $u(k) \in \mathcal{R}(H_d^{\mathrm{T}})$. □

We now consider the cumulative update law (34), which is equivalent to Equations (35)–(37) as shown by Lemma 3.2. The following technical lemma is needed.

Lemma 5.3: *For all $k \geq 1$, let $P(k)$ and $\Theta(k)$ be given by the cumulative update laws* (35)–(37). *Then, for all $k \geq 1$, the following statements hold:*

(*i*) *Let $\phi_1, \phi_2 \in \mathbb{R}^{n_c(l_u+l_y)}$, and define $\Phi_1 \stackrel{\triangle}{=} I_{l_u} \otimes \phi_1^{\mathrm{T}}$, $\Phi_2 \stackrel{\triangle}{=} I_{l_u} \otimes \phi_2^{\mathrm{T}}$. Then, for all $k \geq 1$, there exists $N_{\Phi_1,\Phi_2}(k) \in \mathbb{R}^{l_y \times l_y}$ such that $\Phi_1 P(k) \Phi_2^{\mathrm{T}} H_d^{\mathrm{T}} = H_d^{\mathrm{T}} N_{\Phi_1,\Phi_2}(k)$.*

(*ii*) *Let $\phi \in \mathbb{R}^{n_c(l_u+l_y)}$, and define $\Phi \stackrel{\triangle}{=} I_{l_u} \otimes \phi^{\mathrm{T}}$. Then, $\Phi\Theta(k) \in \mathcal{R}(H_d^{\mathrm{T}})$.*

Proof: See Appendix 4. □

We can now state the main result of this section for the cumulative update laws (35)–(37).

Theorem 5.4: *For all $k \geq 1$, let the control input $u(k)$ be given by the control law* (14) *with the cumulative update laws* (35)–(37). *Then, for all $k \geq 1$,*

$$u(k) \in \mathcal{R}(H_d^{\mathrm{T}}). \tag{44}$$

Proof: The result follows from statement (ii) of Lemma 5.3. □

6. Convergence of Θ

Examples 4.4 and 4.5 show that the controller Θ with the instantaneous update (25) may converge despite the fact that y does not converge and may be unbounded. In this section, we provide sufficient conditions under which Θ converges. These convergence results involve the *zero-update output subspace* $\mathcal{S} \subseteq \mathbb{R}^{l_y}$, which has the property that, if y approaches $\mathcal{S}$ exponentially, then Θ converges. For the case where G_{yu} is tall, we show that $\mathcal{S}$ is non-zero, and thus Θ may converge despite the fact that y does not converge. The discussion is limited to the instantaneous update law (26), but similar results apply to the cumulative update laws (35) and (37). Define the *controller update vector* $\Delta\Theta(k) \stackrel{\triangle}{=} \Theta(k) - \Theta(k-1)$. Note that, since we assume that, for all $k \leq 0$, $\Theta(k) = 0$, it follows that, for all $k \leq 0$, $\Delta\Theta(k) = 0$.

6.1 *Case 1: d* = 1

Lemma 6.1: *Consider the instantaneous update law* (26), *and assume that $d = 1$. Then, $\Delta\Theta(k)$ satisfies*

$$\Delta\Theta(k) = \mathcal{B}(k) H_1^{\mathrm{T}} y(k), \tag{45}$$

where

$$\mathcal{B}(k) \stackrel{\triangle}{=} -[\mu I_{l_u n_c(l_u+l_y)} + \Phi^{\mathrm{T}}(k-2) H_1^{\mathrm{T}} H_1 \Phi(k-2)]^{-1} \cdot \Phi^{\mathrm{T}}(k-2). \tag{46}$$

Proof: Subtracting $\Theta(k-1)$ from both sides of Equation (26) and using the identity $Q(I+Q^{\mathrm{T}}Q)^{-1}=(I+QQ^{\mathrm{T}})^{-1}Q$ yields

$$\begin{aligned}\Delta\Theta(k) &= -\Phi^{\mathrm{T}}(k-2)H_1^{\mathrm{T}}[\mu I_{l_y}+H_1\Phi(k-2)\\&\quad\cdot\Phi^{\mathrm{T}}(k-2)H_1^{\mathrm{T}}]^{-1}\hat{y}(\Theta(k-1),k)\\&= [\mu I_{l_u n_c(l_u+l_y)}+\Phi^{\mathrm{T}}(k-2)H_1^{\mathrm{T}}H_1\Phi(k-2)]^{-1}\\&\quad\cdot\Phi^{\mathrm{T}}(k-2)H_1^{\mathrm{T}}\hat{y}(\Theta(k-1),k).\end{aligned}$$

Since $d=1$, it follows from Equation (17) that $\hat{y}(\Theta(k-1),k)=y(k)$. Therefore,

$$\begin{aligned}\Delta\Theta(k) &= -[\mu I_{l_u n_c(l_u+l_y)}+\Phi^{\mathrm{T}}(k-2)H_1^{\mathrm{T}}H_1\\&\quad\cdot\Phi(k-2)]^{-1}\Phi^{\mathrm{T}}(k-2)H_1^{\mathrm{T}}y(k)\\&= \mathcal{B}(k)H_1^{\mathrm{T}}y(k).\end{aligned}$$

□

It follows from Lemma 6.1 that, for $d=1$, controller update (45) is a memoryless process driven by $H_1^{\mathrm{T}}y(k)$. Note that, if $y(k)\in\mathcal{N}(H_1^{\mathrm{T}})$, then $\Delta\Theta(k)=0$. Furthermore, if G_{yu} is either square or wide, then $\mathcal{N}(H_1^{\mathrm{T}})=\{0\}$, and thus $H_1^{\mathrm{T}}y(k)=0$, if and only if $y(k)=0$. However, if G_{yu} is tall, then $\mathcal{N}(H_1^{\mathrm{T}})\neq\{0\}$, and thus $H_1^{\mathrm{T}}y(k)$ can be zero with non-zero $y(k)$. The next result shows that, if $H_1^{\mathrm{T}}y$ converges exponentially to zero, then Θ converges.

Theorem 6.2: *Consider the instantaneous update law* (26). *Assume that $\mathcal{B}$ is bounded and there exist $\alpha>0$ and $\gamma\in(0,1)$ such that, for all $k\geq 0$, $\|H_1^{\mathrm{T}}y(k)\|\leq\alpha\gamma^k$. Then, Θ converges.*

Proof: See Appendix 5. □

Assume that G_{yu} is tall or square. Then, $H_1^{\mathrm{T}}H_1$ is positive definite, and it follows from Equation (46) that $\mathcal{B}$ is bounded whether or not Φ is bounded. Therefore, if $H_1^{\mathrm{T}}y$ converges exponentially to zero, then Θ converges whether or not Φ is bounded.

Theorem 6.2 has a geometric interpretation. If G_{yu} is tall, and thus H_1^{T} has a non-zero null space, and if y converges to $\mathcal{N}(H_1^{\mathrm{T}})\subset\mathbb{R}^{l_y}$ exponentially, then Θ converges. Thus, Θ may converge whether or not y is bounded as long as y remains in, or exponentially approaches, $\mathcal{N}(H_1^{\mathrm{T}})$. In Section 8.3.1, we show that this is what happens in Examples 4.4 and 4.5.

6.2 *Case 2: d = 2*

Lemma 6.3: *For all $k\geq 1$, let $\Theta(k)$ be given by the instantaneous update law* (26) *with $d=2$. Then, $\Delta\Theta(k)$ satisfies*

$$\Delta\Theta(k)=\mathcal{A}(k)\Delta\Theta(k-1)+\mathcal{B}(k)H_2^{\mathrm{T}}y(k), \tag{47}$$

where

$$\begin{aligned}\mathcal{A}(k) \triangleq\; & -\Phi^{\mathrm{T}}(k-3)H_2^{\mathrm{T}}[\mu I_{l_y}+H_2\Phi(k-3)\\&\cdot\Phi^{\mathrm{T}}(k-3)H_2^{\mathrm{T}}]^{-1}H_2\Phi(k-3),\end{aligned}\tag{48}$$

$$\begin{aligned}\mathcal{B}(k) \triangleq\; & -[\mu I_{l_u n_c(l_u+l_y)}+\Phi^{\mathrm{T}}(k-3)H_2^{\mathrm{T}}H_2\Phi(k-3)]^{-1}\\&\cdot\Phi^{\mathrm{T}}(k-3).\end{aligned}\tag{49}$$

Proof: Substituting $d=2$ into Equation (26) and using the matrix identity $Q(I+Q^{\mathrm{T}}Q)^{-1}=(I+QQ^{\mathrm{T}})^{-1}Q$ yields Equation (47).

As in the case $d=1$, the controller update $\Delta\Theta(k)$ is driven by $H_d^{\mathrm{T}}y(k)$. However, unlike Equation (45), which is memoryless, Equation (47) is dynamic. Therefore, we need to consider the stability of Equation (47). In particular, for all $k\geq 1$, consider the free response of Equation (47), which is given by

$$\Delta\Theta(k)=\mathcal{A}(k)\Delta\Theta(k-1). \tag{50}$$

Definition 6.4: The zero solution of Equation (50) is *globally exponentially stable*, if, for all $\Delta\Theta(0)\in\mathbb{R}^{l_u n_c(l_u+l_y)}$ and $k\geq 0$, there exist $\alpha\geq 1$ and $\gamma\in(0,1)$ such that

$$\|\Delta\Theta(k)\|\leq\alpha\|\Delta\Theta(0)\|\gamma^k. \tag{51}$$

It follows from Equation (51) that, for all $\varepsilon>0$, if $\|\Delta\Theta(0)\|<\frac{\varepsilon}{\alpha}$, then, for all $k\geq 0$, $\|\Delta\Theta(k)\|<\varepsilon$. Therefore, if the zero solution of Equation (50) is globally exponentially stable, then it is Lyapunov stable.

Substituting the singular value decomposition $H_2\Phi(k-3)=U_k\Sigma_kV_k$ into Equation (48) yields

$$\begin{aligned}\mathcal{A}(k) &= V_k^*\Sigma_k^*U_k^*[U_k\mu IU_k^*+U_k\Sigma_kV_kV_k^*\Sigma_k^*U_k^*]^{-1}U_k\Sigma_kV_k\\&= V_k^*\Sigma_k^*[\mu I+\Sigma_k\Sigma_k^*]^{-1}\Sigma_kV_k,\end{aligned}$$

and thus,

$$\sigma_i(\mathcal{A}(k))=\frac{\sigma_i^2(H_2\Phi(k-3))}{\mu+\sigma_i^2(H_2\Phi(k-3))}, \tag{52}$$

where $\sigma_i(\mathcal{A}(k))$ denotes the ith singular value of $\mathcal{A}(k)$. Furthermore, define

$$\bar{\sigma}(\mathcal{A})\triangleq\sup_{k\geq 1}\sigma_{\max}(\mathcal{A}(k)),$$

where $\sigma_{\max}$ denotes the largest singular value. It follows from Equation (52) that $\bar{\sigma}(\mathcal{A})\in[0,1]$. Furthermore, if Φ is bounded, then $\bar{\sigma}(\mathcal{A})\in[0,1)$.

Proposition 6.5: *Let $\Theta(k)$ be given by the instantaneous update* (26) *with $d=2$, and assume that Φ is bounded. Then, the zero solution of Equation* (50) *is globally exponentially stable.*

Proof: See Appendix 6. □

Proposition 6.5 is restrictive in the sense that it requires that the regressor Φ be bounded. We now relax this requirement by introducing a persistency condition. For $k \geq 1$ and $m \geq 2$, define

$$Q_m(k) \triangleq \mathcal{A}(k) \cdots \mathcal{A}(k+m-1). \tag{53}$$

Proposition 6.6: *Let $\Theta(k)$ be given by the instantaneous update (26) with $d = 2$, and consider Equation (50). Assume that there exists $m \geq 2$ such that*

$$\bar{\sigma}(Q_m) \in (0, 1). \tag{54}$$

Then, the zero solution of Equation (50) is globally exponentially stable.

Proof: See Appendix 7. □

The next result shows that if the regressor is sufficiently persistent, in particular, persistently exciting of order two, then Equation (54) is satisfied.

For non-zero vectors $v_1, v_2 \in \mathbb{R}^{l_u n_c (l_u + l_y)}$, let

$$\Omega(v_1, v_2) \triangleq \cos^{-1} \frac{v_1^{\mathrm{T}} v_2}{\|v_1\| \|v_2\|} \in [0, \pi] \tag{55}$$

denote the angle between v_1 and v_2. The following technical lemma is needed.

Lemma 6.7: *Let $r, m \geq 2$, $v_1, \ldots, v_m \in \mathbb{R}^r$ be a set of vectors such that $\Omega(v_1, v_m) \in (0, \pi)$, and let $\Omega_0 \in (0, \pi/2)$ be such that*

$$\Omega_0 \leq \Omega(v_1, v_m) \leq \pi - \Omega_0. \tag{56}$$

Then, there exists $l \in \{1, \ldots, m-1\}$ such that

$$\frac{1}{m-1}\Omega_0 \leq \Omega(v_l, v_{l+1}) \leq \pi - \frac{1}{m-1}\Omega_0. \tag{57}$$

To illustrate Lemma 6.7, let $v_1, v_2, v_3 \in \mathbb{R}^2$, where $\Omega(v_1, v_3) \in [\Omega_0, \pi - \Omega_0]$. Then, it is geometrically obvious that at least one of the angles $\Omega(v_1, v_2)$ and $\Omega(v_2, v_3)$ must lie in the range $[\Omega_0/2, \pi - \Omega_0/2]$.

Proposition 6.8: *Assume that G_{yu} is either tall or square, and assume that there exist $m \geq 2$ and $\Omega_0 \in (0, \pi/2]$ such that, for all $k \geq 1$, there exist distinct $k_1, k_2 \in \{k, \ldots, k + m - 1\}$ such that $\Omega_0 \leq \Omega(\phi(k_1 - 3), \phi(k_2 - 3)) \leq \pi - \Omega_0$. Then, $\bar{\sigma}(Q_m) \leq \cos \frac{\Omega_0}{m-1} < 1$.*

Proof: See Appendix 8. □

It follows from Proposition 6.8 that, if there exist $m \geq 2$ and $\Omega_0 \in (0, \pi/2]$ such that, for all $k \geq 1$, the set $\{\phi(k-3), \ldots, \phi(k+m-4)\}$ contains two vectors, the angle between which is at least Ω_0 radians and at most $\pi - \Omega_0$ radians, then Equation (54) is satisfied. This condition implies that the regressor is persistently exciting of order two as defined in Goodwin and Sin (1984) and Soderstrom and Stoica (1989). It should be noted that this persistency condition cannot be guaranteed prior to design. Future work could explore the necessary and sufficient conditions under which this persistency condition is satisfied.

Now that we have established global exponential stability for (50), we consider $\Theta(k)$ generated by (47). Since $\mathcal{N}(H_2^{\mathrm{T}}) \neq \{0\}$ if and only if G_{yu} is tall, we consider only the case where G_{yu} is tall.

Theorem 6.9: *Consider the instantaneous update law (26) with $d = 2$, let G_{yu} be tall, and assume that there exist $\alpha > 0$ and $\gamma \in (0, 1)$ such that, for all $k \geq 0$, $\|H_2^{\mathrm{T}} y(k)\| \leq \alpha \gamma^k$. Then, the following statements hold:*

(*i*) *If Φ is bounded, then Θ converges.*
(*ii*) *If there exists $m \geq 2$ such that Equation (54) is satisfied, then Θ converges.*

Proof: See Appendix 9. □

Theorem 6.2 has a geometric interpretation similar to the case $d = 1$ with $\mathcal{N}(H_1^{\mathrm{T}})$ replaced by $\mathcal{N}(H_2^{\mathrm{T}})$, that is, if y converges exponentially to $\mathcal{N}(H_2^{\mathrm{T}})$, then Θ converges.

6.3 *Case 3: d ≥ 3*

We now briefly investigate the case $d \geq 3$. Consider the instantaneous update law (26) with arbitrary $d \geq 3$. First, from Equation (17),

$$\begin{aligned}
\hat{y}(\Theta(k-1), k) &= y(k) + H_d \Phi(k-d-1)[\Theta(k-1) - \Theta(k-d)] \\
&= y(k) + H_d \Phi(k-d-1)[\Theta(k-1) - \Theta(k-2) \\
&\quad + \Theta(k-2) - \cdots - \Theta(k-d+1) + \Theta(k-d+1) \\
&\quad - \Theta(k-d)] \\
&= y(k) + H_d \Phi(k-d-1) \sum_{i=1}^{d-1} \Delta\Theta(k-i).
\end{aligned} \tag{58}$$

Substituting Equation (58) into Equation (26), subtracting $\Theta(k-1)$ from Equation (26) and using the identity $Q(I + Q^{\mathrm{T}} Q)^{-1} = (I + Q Q^{\mathrm{T}})^{-1} Q$, we obtain

$$\begin{aligned}
\Delta\Theta(k) = \mathcal{M}(k)\Delta\Theta(k-1) + \cdots + \mathcal{M}(k)\Delta\Theta(k-d+1) \\
+ \mathcal{N}(k) H_d^{\mathrm{T}} y(k),
\end{aligned} \tag{59}$$

where

$$\begin{aligned}
\mathcal{M}(k) &\triangleq -\Phi(k-d-1)^{\mathrm{T}} H_d^{\mathrm{T}} [\mu I_{l_y} + H_d \Phi(k-d-1) \\
&\quad \cdot \Phi^{\mathrm{T}}(k-d-1) H_d^{\mathrm{T}}]^{-1} H_d \Phi(k-d-1), \\
\mathcal{N}(k) &\triangleq -[\mu I_{l_u n_c (l_u + l_y)} + H_d \Phi(k-d-1) \\
&\quad \cdot \Phi^{\mathrm{T}}(k-d-1) H_d^{\mathrm{T}}] \Phi(k-d-1).
\end{aligned}$$

Now, letting

$$\mathcal{X}(k) \triangleq \begin{bmatrix} \Delta\Theta(k) \\ \vdots \\ \Delta\Theta(k-d+2) \end{bmatrix},$$

we rewrite Equation (59) as

$$\mathcal{X}(k) = \mathcal{E}(k)\mathcal{X}(k-1) + \mathcal{F}(k)H_d^{\mathrm{T}}y(k), \tag{60}$$

$$\Delta\Theta(k) = \mathcal{C}\mathcal{X}(k), \tag{61}$$

where

$$\mathcal{E}(k) \triangleq \begin{bmatrix} \mathcal{M}(k) & \cdots & \mathcal{M}(k) & \mathcal{M}(k) \\ I & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix},$$

$$\mathcal{F}(k) \triangleq \begin{bmatrix} \mathcal{N}(k) \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{62}$$

$$\mathcal{C} \triangleq \begin{bmatrix} I \; 0 \cdots 0 \end{bmatrix}. \tag{63}$$

Thus, as in the cases $d = 1$ and $d = 2$, the controller update (60), (61) is driven by $H_d^{\mathrm{T}}y(k)$. Furthermore, in addition to $H_d^{\mathrm{T}}y(k)$, $\Delta\Theta(k)$ also depends on $d-1$ past controller updates. It follows from Equation (60) that, if, for all $k \geq 1$, $y(k) \in \mathcal{N}(H_d^{\mathrm{T}})$, then, for all $\mathcal{X}(0)$, $\mathcal{X}(k)$, and thus $\Delta\Theta(k)$, converges to zero if and only if the equilibrium $\mathcal{X} = 0$ of

$$\mathcal{X}(k) = \mathcal{E}(k)\mathcal{X}(k-1) \tag{64}$$

is globally attractive.

Since $\sigma_{\max}(\mathcal{E}(k))$ may be greater than 1, convergence results for $\Theta(k)$ in the case $d = 3$ are more complicated than in the cases $d = 1$ and $d = 2$. Numerical testing suggests that, if $y(k) \in \mathcal{N}(H_d)$ and Equation (54) is satisfied, then $\{\mathcal{X}(k)\}_{k=1}^{\infty}$ and $\sum_{i=1}^{\infty} \Delta\Theta(k)$ converge, and thus Θ converges.

7. Input-subspace zeros

In this section, we build on the results of Section 5, and introduce the notion of input-subspace zeros, which arise due to the fact that the control input is contained in $\mathcal{R}(H_d^{\mathrm{T}})$, so that there exists $v \in \mathbb{R}^{l_y}$ such that $u = H_d^{\mathrm{T}}v$. If G_{yu} is square or tall, then $\mathcal{R}(H_d^{\mathrm{T}}) = \mathbb{R}^{l_u}$; in this case, we show that the input-subspace zeros of G_{yu} are equal to the transmission zeros of G_{yu}. However, in the case where G_{yu} is wide, $\mathcal{R}(H_d^{\mathrm{T}})$ is a proper subspace of $\mathbb{R}^{l_u}$. In this case, we show that G_{yu} may be minimum phase but have NMP input-subspace zeros. Finally, in light of input-subspace zeros, we revisit Examples 4.1 and 4.2 and demonstrate that the instability observed in Example 4.2 is caused by unstable cancellation of an NMP input-subspace zero that is not a transmission zero of G_{yu}.

7.1 *Right-squared transfer matrix from v to y*

Consider $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$, and define the right-squared transfer matrix

$$G_{yu}^{\mathrm{R}} \triangleq G_{yu}H_d^{\mathrm{T}} \sim \left[\begin{array}{c|c} A & BH_d^{\mathrm{T}} \\ \hline C & 0 \end{array}\right]. \tag{65}$$

Theorems 5.2 and 5.4 imply that, for all $k \geq 1$, the control input $u(k)$ generated by the instantaneous and cumulative update laws lies in the subspace $\mathcal{R}(H_d^{\mathrm{T}}) \subseteq \mathbb{R}^{l_y}$, so that $u(k) = H_d^{\mathrm{T}}v(k)$, where $v(k) \in \mathbb{R}^{l_y}$. Hence, Equation (7) becomes

$$y = G_{yu}^{\mathrm{R}}v + G_{yw}w. \tag{66}$$

Note that $G_{yu}^{\mathrm{R}} \in \mathbb{R}^{l_y \times l_y}(z)$. If the realisation (65) is minimal, then the transmission zeros of G_{yu}^{R} are given by

$$\mathrm{tzeros}(G_{yu}^{\mathrm{R}}) = \{\zeta \in \mathbb{C} : \mathrm{rank}\ \Sigma^{\mathrm{R}}(\zeta) < \mathrm{normal\ rank}\ \Sigma^{\mathrm{R}}\}, \tag{67}$$

where

$$\Sigma^{\mathrm{R}}(z) \triangleq \begin{bmatrix} zI - A & BH_d^{\mathrm{T}} \\ C & 0 \end{bmatrix} \in \mathbb{R}^{(n+l_y)\times(n+l_y)}[z]. \tag{68}$$

The transmission zeros of G_{yu}^{R} are the *input-subspace zeros* of G_{yu}.

7.2 *Tall and square plants*

The following result concerns the minimality of Equation (65) for tall and square G_{yu}.

Proposition 7.1: *If* $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ *is tall or square, then* $(A, BH_d^{\mathrm{T}}, C)$ *is minimal.*

Proof: For all $\lambda \in \mathbb{C}$, we have

$$\begin{bmatrix} \lambda I - A & BH_d^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \lambda I - A & B \end{bmatrix}\mathcal{Q}, \tag{69}$$

where

$$\mathcal{Q} \triangleq \begin{bmatrix} I_n & 0_{n\times l_y} \\ 0_{l_u\times n} & H_d^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{(n+l_u)\times(n+l_y)}. \tag{70}$$

Since G_{yu} is tall or square and H_d has full rank, we have rank $H_d = l_u$, and thus, rank $\mathcal{Q} = n + l_u$. Therefore, it follows from Equation (69) that

$$\mathrm{rank}\,[\,\lambda I - A \;\; BH_d^{\mathrm{T}}\,] = \mathrm{rank}\,[\,\lambda I - A \;\; B\,]. \tag{71}$$

Since (A, B) is controllable, it follows from Equation (71) that (A, BH_d^{T}) is controllable. Furthermore, since (A, C) is observable, Equation (65) is minimal. □

Thus, if G_{yu} is tall or square, then the input-subspace zeros of G_{yu} are defined as in Equation (67). We now show that, if G_{yu} is tall or square, then its input-subspace zeros and transmission zeros are identical.

Proposition 7.2: *If* $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ *is tall or square, then* $\mathrm{tzeros}(G_{yu}^{\mathrm{R}}) = \mathrm{tzeros}(G_{yu})$.

Proof: It follows from Equations (11) and (68) that $\Sigma^{\mathrm{R}}(z) = \Sigma(z)\mathcal{Q}$, where $\mathcal{Q}$ is given by Equation (70). Since rank $\mathcal{Q} = n + l_u$, it follows that, for all $z \in \mathbb{C}$, rank $\Sigma(z) =$ rank $\Sigma^{\mathrm{R}}(z)$. It thus follows from Equations (12) and (67), and Proposition 7.1 that $\mathrm{tzeros}(G_{yu}^{\mathrm{R}}) = \mathrm{tzeros}(G_{yu})$. □

Therefore, for tall and square plants, the restriction $u(k) \in \mathcal{R}(H_d^{\mathrm{T}})$ has no effect on controllability, and does not alter the transmission zeros of the plant. This is expected because $\mathcal{R}(H_d^{\mathrm{T}}) = \mathbb{R}^{l_u}$ in the case where G_{yu} is tall or square.

7.3 *Wide plants*

It is reasonable to expect that the properties of G_{yu}^{R} for wide plants are dual to those of tall plants. However, as we now show, this is not the case. For example, although the realisation (65) is minimal for all tall plants G_{yu}, it turns out that Equation (65) for a wide plant G_{yu} may or may not be minimal, as illustrated by the following example.

Example 7.3 (Minimality of Equation (65))**:** Consider the 1×2 wide plant $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$, where

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0.5 & 0.5 \\ -0.5 & 0.5 \end{bmatrix}, \quad C = [\,0 \;\; 2\,],$$

$n = 2$, $d = 1$, and $H_1 = [\,-1 \;\; 1\,]$. Note that

$$\mathrm{rank}\,\left[\,BH_1^{\mathrm{T}} \;\; ABH_1^{\mathrm{T}}\,\right] = \mathrm{rank}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} < n,$$

which implies that Equation (65) is not minimal. □

Example 7.3 shows that the minimality of (A, B, C) does not imply that Equation (65) is minimal. However, throughout the rest of this section, we only consider plants for which Equation (65) is minimal.

Since $(A, BH_d^{\mathrm{T}}, C)$ is minimal, the input-subspace zeros of G_{yu} are defined as in Equation (67). The following example illustrates that the input-subspace zeros and the transmission zeros of a wide plant may be distinct.

Example 7.4 (Input-subspace zeros of a wide plant)**:** Consider $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ with

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -0.8 & -0.3 \\ 0.5 & 0.5 \end{bmatrix}, \quad C = [\,0 \;\; 2\,],$$

$n = 2$, $d = 1$, $H_1 = [\,1 \;\; 1\,]$. For this example, (A, BH_1^{T}) is controllable. It can be shown that $\mathrm{tzeros}(G_{yu}) = \varnothing$ and $\mathrm{tzeros}(G_{yu}^{\mathrm{R}}) = \{1.1\}$. Hence, this example shows that the transmission zeros and the input-subspace zeros of a wide plant may be distinct.

It follows from Davison and Wang (1974, Theorem 5) that wide $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ generically has no transmission zeros, whereas, since $G_{yu}^{\mathrm{R}} \in \mathbb{R}^{l_y \times l_y}(z)$, G_{yu} generically has $n - l_y$ input-subspace zeros. In particular, in the case $d = 1$, since rank $CB(CB)^{\mathrm{T}} = l_y$, it follows that G_{yu}^{R} has exactly $n - l_y$ zeros (Maciejowski, 1989). Therefore, if $n > l_y$, then G_{yu} generically has more input-subspace zeros than transmission zeros. Furthermore, if G_{yu} has at least one NMP input-subspace zero, then there exist infinitely many unbounded (MacFarlane & Karcanias, 1976; Tokarzewski, 2002; Tokarzewski, 2006, p. 21) *output-zeroing* input sequences $\{u(k)\}_{k=0}^{\infty} \subset \mathcal{R}(H_d^{\mathrm{T}})$, each of which is associated with an initial condition $x(0) \in \mathbb{R}^n$, such that, for all $k \geq 0$, the output $y(k)$ of G_{yu} due to $(x(0), \{u(k)\}_{k=0}^{\infty})$ is identically equal to zero. The next result characterises $(x(0), \{u(k)\}_{k=0}^{\infty})$, which produce identically zero output y.

Proposition 7.5: *Let* $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ *be wide with state* x *and output* y, *and let* ζ *be a non-zero input-subspace zero of* G_{yu}. *Then, the following statements hold:*

(*i*) *There exists non-zero* $\begin{bmatrix} x_0 \\ v_0 \end{bmatrix} \in \mathbb{C}^{n+l_y}$ *such that*

$$\Sigma^{\mathrm{R}}(\zeta)\begin{bmatrix} x_0 \\ v_0 \end{bmatrix} = 0. \tag{72}$$

(*ii*) *Let* $x(0) = -\mathrm{Re}(x_0)$, *and, for all* $k \geq 0$, *let*

$$u(k) = H_d^{\mathrm{T}}[\mathrm{Re}(\zeta^k)\mathrm{Re}(v_0) - \mathrm{Im}(\zeta^k)\mathrm{Im}(v_0)]. \tag{73}$$

Then, for all $k \geq 0$,

$$x(k) = -\mathrm{Re}(\zeta^k)\mathrm{Re}(x_0) + \mathrm{Im}(\zeta^k)\mathrm{Im}(x_0), \tag{74}$$

$$y(k) = 0. \tag{75}$$

(*iii*) *Let* $\alpha \in \mathbb{R}$, *let* $x(0) = -\alpha \mathrm{Re}(x_0)$, *and, for all* $k \geq 0$, *let*

$$u(k) = \alpha H_d^{\mathrm{T}}[\mathrm{Re}(\zeta^k)\mathrm{Re}(v_0) - \mathrm{Im}(\zeta^k)\mathrm{Im}(v_0)]. \quad (76)$$

Then, for all $k \geq 0$, $y(k) = 0$.

(*iv*) *Let* $\alpha \in \mathbb{R}$, *assume that* A *is discrete-time asymptotically stable, and let* $u(k)$ *be given by Equation* (76). *Then, for all* $x(0) \in \mathbb{R}^n$, $y(k) \to 0$ *as* $k \to \infty$ *with exponential convergence.*

Proof: See Appendix 10. □

Note that, if the input-subspace zero ζ satisfies $|\zeta| > 1$, then the output-zeroing input sequence (76) with $\alpha \neq 0$ is unbounded. Hence, if G_{yu} has at least one NMP input-subspace zero, then there exist infinitely many unbounded output-zeroing input sequences that are contained in the subspace $\mathcal{R}(H_d^{\mathrm{T}})$, even though G_{yu} itself is minimum phase. Since the retrospective cost functions (21) and (28) do not contain a control penalty or a constraint on the amplitude of u, RCAC may converge to a controller that produces an unbounded output-zeroing input sequence, namely an unstable controller with a pole (or poles) located at the NMP input-subspace zero(s) of G_{yu}. In the next subsection, we show that this is the cause of the instability in Example 4.2. In Section 9, we remedy this behaviour by modifying the retrospective cost (28).

7.3.1 *Examples* 4.1 *and* 4.2 *revisited*

In Examples 4.1 and 4.2, which are identical except for $B_{(1,1)}$, the cumulative adaptive controller (34) is applied to 2×3 plants in order to reject an unmatched harmonic disturbance. Both plants have no transmission zeros, the given realisations are minimal, and the open-loop systems have the same eigenvalues. However, as shown in Figure 3, the control signal u for the adaptive system in Example 4.2 is unbounded. We now demonstrate that the unbounded control signal is caused by the NMP input-subspace zero of G_{yu}.

Example 7.6 (Example 4.2 revisited)**:** We first confirm that Equation (44) holds. Note that $d = 1$ and

$$H_1 = \begin{bmatrix} 0.1062 & 0.8195 & -1.1582 \\ 3.2868 & -0.4562 & -4.4993 \end{bmatrix}.$$

Hence, $\mathcal{R}(H_1^{\mathrm{T}})$ is the plane described by $au_1 + bu_2 + cu_3 = 0$, where a, b, and c satisfy

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathcal{N}(H_1) = \mathrm{span}\left\{ \begin{bmatrix} 0.699 \\ 0.552 \\ 0.4547 \end{bmatrix} \right\}.$$

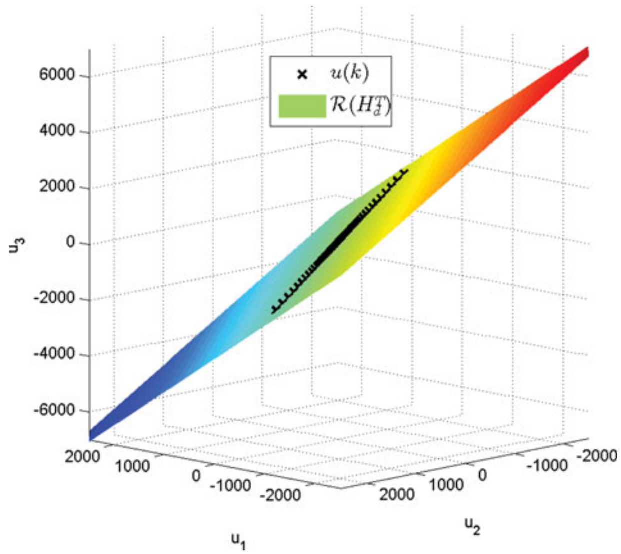

Figure 7. This figure illustrates the phase portrait of the unbounded control input u for Example 4.2 shown in Figure 6. For all $k \geq 1$, $u(k)$ is contained in the subspace $\mathcal{R}(H_d^{\mathrm{T}})$, which is the coloured plane in this figure. The control input is unbounded due to the fact that the input-subspace zeros of G_{yu} are NMP.

The phase portrait of $u(k)$ for $k \geq 1$ illustrated in Figure 7 shows that $u(k)$ is confined to the subspace $\mathcal{R}(H_1^{\mathrm{T}})$ for all $k \geq 1$.

We now investigate the input-subspace zeros of the plant. Since $(A, BH_d^{\mathrm{T}}, C)$ is minimal, Equation (67) can be used to obtain the input-subspace zeros of G_{yu}, which are given by tzeros$(G_{yu}^{\mathrm{R}}) = \{-1.0555,\ 0.7596\}$. Therefore, G_{yu} has an NMP input-subspace zero at -1.0555. Computing the controller poles at $k = 150$, Figure 8 shows that, as Θ converges, one controller pole is located near the NMP input-subspace zero location -1.0555. In effect, RCAC attempts to cancel the unmodelled NMP input-subspace zero. Thus, the results of Example 4.2 can be explained as follows. The unstable controller pole at the NMP input-subspace zero causes the control input to diverge, but the effect of the unbounded control input is blocked by the NMP input-subspace zero, and the performance output y converges to zero despite the fact that u is unbounded, as suggested by (iv) of Proposition 7.5. Furthermore, since G_{yu} is wide, $\mathcal{N}(H_d^{\mathrm{T}}) = \{0\}$, and thus Θ converges as y converges to zero.

Example 7.7 (Example 4.1 revisited)**:** We now revisit Example 4.1, where the control input u is bounded. The phase portrait of $u(k)$ for $k \geq 1$ illustrated in Figure 9 shows that $u(k)$ is contained in the subspace $\mathcal{R}(H_1^{\mathrm{T}})$ for all $k \geq 1$. The input-subspace zeros of the plant (38) and (39) are given by tzeros$(G_{yu}^{\mathrm{R}}) = \{-0.7334,\ 0.7679\}$. The input-subspace zeros for Example 4.1 are thus minimum phase, and, as shown in Figure 2, the control input u is bounded. Furthermore, Θ converges as y converges to zero.

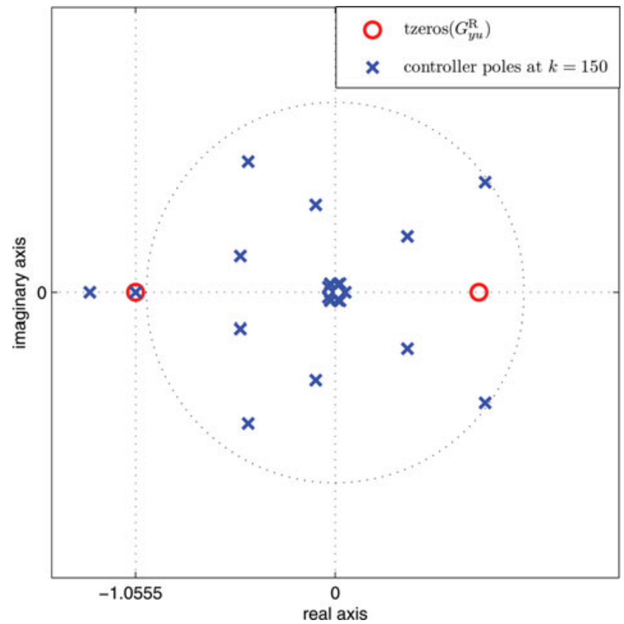

Figure 8. This figure illustrates the input-subspace zeros of the plant in Example 4.2 along with the poles of the adaptive controller at $k = 150$, whose time evolution is shown in Figure 3. The adaptive controller places a pole near the NMP input-subspace zero of G_{yu}, which is located at -1.0555. This unstable pole-zero cancellation is the cause of the unbounded control input shown in Figure 3. Note that the NMP input-subspace zero is not a transmission zero of G_{yu}.

8. Output-subspace zeros

In this section, we build on the convergence results of Section 6 and introduce the notion of output-subspace zeros, which are the zeros from the control input to the scaled performance variable $H_d^{\rm T} y$, which drives the update of Θ, as shown in Section 6. If G_{yu} is square or wide, then, since $\mathcal{N}(H_d^{\rm T}) = \{0\}$, $H_d^{\rm T} y = 0$, if and only if $y = 0$. In this case, it is reasonable to expect that zeros from u to y and zeros from u to $H_d^{\rm T} y$ are identical, which we show by proving that the output-subspace zeros and the transmission zeros of square and wide plants are identical. However, in the case where G_{yu} is tall, $\mathcal{N}(H_d^{\rm T})$ is a proper subspace of $\mathbb{R}^{l_y}$, and thus $H_d^{\rm T} y$ may be zero with non-zero y. In this case, we show that the output-subspace zeros and the transmission zeros of G_{yu} may be distinct. In particular, we show that G_{yu} may be minimum phase, but have NMP output-subspace zeros, and, in this case, we show that the control input may be unbounded despite the fact that $H_d^{\rm T} y$ is approaching zero exponentially fast, which in turn leads to converging Θ. At the end of the section, we revisit Examples 4.4 and 4.5 in light of output-subspace zeros, and demonstrate that the instability observed in Example 4.5 is caused by unstable cancellation of an NMP output-subspace zero, which leads to an unbounded control input and performance output, although the unbounded control input does not affect $H_d^{\rm T} y$ because of the NMP output-subspace zero.

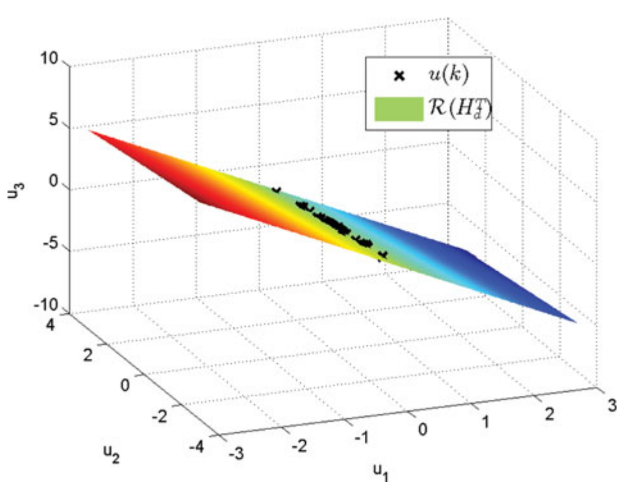

Figure 9. This figure illustrates the phase portrait of the bounded control input u for Example 4.1 shown in Figure 2. For all $k \geq 1$, $u(k)$ is contained in the subspace $\mathcal{R}(H_d^{\rm T})$, which is the coloured plane in this figure. The control input is bounded due to the fact that the input-subspace zeros of G_{yu} are minimum phase.

8.1 *Left-squared transfer matrix from u to $H_d^{\rm T} y$*

Consider $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$, and define the left-squared transfer matrix

$$G^{\rm L}_{yu} \triangleq H_d^{\rm T} G_{yu} \sim \left[\begin{array}{c|c} A & B \\ \hline H_d^{\rm T} C & 0 \end{array}\right]. \tag{77}$$

For plants with $d = 1$ or $d = 2$, Theorems 6.2 and 6.9 imply that, if $H_d^{\rm T} y$ converges to zero, then Θ converges. For tall plants, $\mathcal{N}(H_d^{\rm T}) \neq \{0\}$, and thus $H_d^{\rm T} y$ may converge to zero with possibly unbounded y. It follows from Equations (45) and (47) that $H_d^{\rm T} y$ drives the controller update $\Delta\Theta$. To investigate the zeros from u to $H_d^{\rm T} y$, we multiply Equation (7) by $H_d^{\rm T}$, and consider

$$H_d^{\rm T} y = G^{\rm L}_{yu} u + H_d^{\rm T} G_{yw} w. \tag{78}$$

Note that $G^{\rm L}_{yu} \in \mathbb{R}^{l_u \times l_u}(z)$. If Equation (77) is minimal, then the transmission zeros of $G^{\rm L}_{yu}$ are given by

$$\text{tzeros}(G^{\rm L}_{yu}) = \{\zeta \in \mathbb{C} : \text{ rank } \Sigma^{\rm L}(\zeta) < \text{normal rank } \Sigma^{\rm L}\}, \tag{79}$$

where

$$\Sigma^{\rm L}(z) \triangleq \begin{bmatrix} zI - A & B \\ H_d^{\rm T} C & 0 \end{bmatrix}. \tag{80}$$

The transmission zeros of $G^{\rm L}_{yu}$ are the *output-subspace zeros* of G_{yu}. We consider the output-subspace zeros of wide, square, and tall plants separately. Unlike Section

7, where we consider tall and square plants before wide plants, in this section, we consider wide and square plants before tall plants. As it turns out, there exists a duality between output-subspace zeros of wide and square (tall) plants and input-subspace zeros of tall and square (wide) plants.

8.2 *Wide and square plants*

First, as pointed out in Section 6, if G_{yu} is wide or square, then $\mathcal{N}(H_d^{\mathrm{T}}) = \{0\}$, and thus, $H_d^{\mathrm{T}} y(k) = 0$, if and only if $y(k) = 0$. It is, therefore, intuitive to expect that the output-subspace zeros of G_{yu} are equal to the transmission zeros of G_{yu}, that is, $\mathrm{tzeros}(G_{yu}^{\mathrm{L}}) = \mathrm{tzeros}(G_{yu})$. We now show that this is indeed the case.

First, we show that (77) is minimal for all wide and square plants G_{yu}.

Proposition 8.1: *If* $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ *is wide or square, then* $(A, B, H_d^{\mathrm{T}} C)$ *is minimal.*

Proof: Since G_{yu} is wide or square and H_d has full rank, we have $\mathrm{rank}\ H_d = l_y$, and thus, $\mathrm{rank}\ \mathcal{Q} = n + l_y$, where $\mathcal{Q}$ is given by Equation (70). Therefore, for all $\lambda \in \mathbb{C}$, we have

$$\begin{aligned} \mathrm{rank} \begin{bmatrix} \lambda I - A \\ H_d^{\mathrm{T}} C \end{bmatrix} &= \mathrm{rank} \left(\mathcal{Q} \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} \right) \\ &= \mathrm{rank} \begin{bmatrix} \lambda I - A \\ C \end{bmatrix}. \end{aligned} \tag{81}$$

Since (A, C) is observable, it follows from Equation (81) that $(A, H_d^{\mathrm{T}} C)$ is observable. Furthermore, since (A, B) is controllable, Equation (77) is minimal. □

Since Equation (77) is minimal, the output-subspace zeros of G_{yu} are defined as in Equation (79). We now show that if G_{yu} is wide or square, then its output-subspace zeros and transmission zeros are identical.

Proposition 8.2: *If* $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ *is wide or square, then* $\mathrm{tzeros}(G_{yu}^{\mathrm{L}}) = \mathrm{tzeros}(G_{yu})$.

Proof: It follows from Equations (11) and (80) that $\Sigma^{\mathrm{L}}(z) = \mathcal{Q}\Sigma(z)$, where $\mathcal{Q}$ is defined as in Equation (70). Since G_{yu} is wide or square and H_d has full rank, $\mathrm{rank}\ \mathcal{Q} = n + l_y$. Therefore, for all $z \in \mathbb{C}$, $\mathrm{rank}\ \Sigma(z) = \mathrm{rank}\ \Sigma^{\mathrm{L}}(z)$. It thus follows from Equations (12) and (79), and Proposition 8.1 that $\mathrm{tzeros}(G_{yu}^{\mathrm{L}}) = \mathrm{tzeros}(G_{yu})$. □

Therefore, for wide and square plants, G_{yu}^{L} is NMP, if and only if G_{yu} is NMP. Therefore, if G_{yu} is minimum phase, then y cannot converge to $\mathcal{N}(H_d^{\mathrm{T}}) = \{0\}$ with an unbounded input sequence, and thus Θ cannot converge to a controller that generates an unbounded input sequence.

8.3 *Tall plants*

We first investigate the minimality of the realisation (77) for tall plants. The following example illustrates that the minimality of (A, B, C) does not imply that Equation (77) is minimal.

Example 8.3 (Minimality of Equation (77))**:** Consider the 2×1 plant $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$, where

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad C = \begin{bmatrix} 0.5 & -0.5 \\ 0.5 & 0.5 \end{bmatrix}, \tag{82}$$

$n = 2$, $d = 1$, and $H_1 = \begin{bmatrix} -1 & 1 \end{bmatrix}^{\mathrm{T}}$. Note that

$$\mathrm{rank} \begin{bmatrix} H_1^{\mathrm{T}} C \\ H_1^{\mathrm{T}} C A \end{bmatrix} = \mathrm{rank} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} < n,$$

which implies that Equation (77) is not minimal. □

Example 8.3 shows that the minimality of (A, B, C) does not imply that Equation (77) is minimal. However, throughout the rest of this section, we only consider plants for which Equation (77) is minimal.

Since Equation (77) is minimal, the output-subspace zeros of G_{yu} are defined as in Equation (79). The following example illustrates that the output-subspace zeros and the transmission zeros of G_{yu} may be distinct.

Example 8.4 (Output-subspace zeros of a tall plant)**:** Consider the 3×2 plant $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$, where

$$A = \begin{bmatrix} 0.5 & 0.3 & 2.5 & 0.7 \\ 1.8 & -1.3 & 1.2 & 0 \\ -2.2 & -0.4 & -1.3 & 0.7 \\ 0.8 & 0.3 & 2 & -0.2 \end{bmatrix}, \quad B = \begin{bmatrix} -0.1 & 0.6 \\ 1.5 & -1.2 \\ 1.4 & 0.7 \\ 1.4 & 1.6 \end{bmatrix},$$
$$C = \begin{bmatrix} 0.4 & -0.3 & 0.8 & -0.8 \\ 1 & 0.3 & -1.1 & -1.9 \\ 0.7 & -0.8 & -1 & 1.4 \end{bmatrix}, \tag{83}$$

$n = 4$, $d = 1$, and $H_1 = \begin{bmatrix} -0.49 & -0.12 \\ -3.85 & -3.57 \\ -0.71 & 2.92 \end{bmatrix}$. For this example, $(A, H_1^{\mathrm{T}} C)$ is observable so that $(A, B, H_1^{\mathrm{T}} C)$ is minimal. It can be shown that $\mathrm{tzeros}(G_{yu}) = \varnothing$ and $\mathrm{tzeros}(G_{yu}^{\mathrm{L}}) = \{0.0969 + \jmath 0.8774, 0.0969 - \jmath 0.8774\}$. Hence, this example shows that the transmission zeros and the output-subspace zeros of a tall plant may be distinct. □

It turns out that the properties of G_{yu}^{L} for tall plants are dual to those of G_{yu}^{R} for wide plants. In particular, for almost all tall $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$, it follows from Davison and Wang (1974, Theorem 5) that $\mathrm{tzeros}(G_{yu}) = \varnothing$, whereas, since $G_{yu}^{\mathrm{L}} \in \mathbb{R}^{l_u \times l_u}(z)$, G_{yu} generically has $n - l_u$ output-subspace zeros. Furthermore, in the case $d = 1$, it follows that G_{yu} has exactly $n - l_u$ output-subspace zeros. Therefore, if $n > l_u$, then G_{yu} generically has more output-subspace zeros than transmission zeros. Furthermore, if G_{yu} has at least one NMP output-subspace zero, then there exist infinitely many unbounded input sequences $\{u(k)\}_{k=0}^{\infty}$, each of which associated with an initial condition $x(0) \in \mathbb{R}^n$, such that, for all $k \geq 0$, the scaled performance output $H_d^{\mathrm{T}} y(k)$ due to $(x(0), \{u(k)\}_{k=0}^{\infty})$ is identically equal to zero. The following result, which is the dual of Proposition 7.5, characterises $(x(0), u(k))$ that produce identically zero $H_d^{\mathrm{T}} y$. The proof is similar to the proof of Proposition 7.5 and is omitted.

Proposition 8.5: *Let* $G_{yu} \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ *be tall with state* $x(k)$ *and output* $y(k)$*, and let* ζ *be a non-zero output-subspace zero of* G_{yu}*. Then, the following statements hold:*

(*i*) *There exists non-zero* $\begin{bmatrix} x_0 \\ u_0 \end{bmatrix} \in \mathbb{C}^{n+l_u}$ *such that*

$$\Sigma^{\mathrm{L}}(z) \begin{bmatrix} x_0 \\ u_0 \end{bmatrix} = 0. \tag{84}$$

(*ii*) *Let* $x(0) = -\mathrm{Re}(x_0)$*, and, for all* $k \geq 0$*, let the control input be given by*

$$u(k) = \mathrm{Re}(\zeta^k)\mathrm{Re}(u_0) - \mathrm{Im}(\zeta^k)\mathrm{Im}(u_0). \tag{85}$$

Then, for all $k \geq 0$*,* $x(k)$ *and* $H_d^{\mathrm{T}} y(k)$ *satisfy*

$$x(k) = -\mathrm{Re}(\zeta^k)\mathrm{Re}(x_0) + \mathrm{Im}(\zeta^k)\mathrm{Im}(x_0), \tag{86}$$

$$H_d^{\mathrm{T}} y(k) = 0. \tag{87}$$

(*iii*) *Let* $\alpha \in \mathbb{R}$*, let* $x(0) = -\alpha\mathrm{Re}(x_0)$*, and, for all* $k \geq 0$*, assume that the control input be given by*

$$u(k) = \alpha[\mathrm{Re}(\zeta^k)\mathrm{Re}(u_0) - \mathrm{Im}(\zeta^k)\mathrm{Im}(u_0)]. \tag{88}$$

Then, for all $k \geq 0$*,* $H_d^{\mathrm{T}} y(k) = 0$*.*

(*iv*) *Let* $\alpha \in \mathbb{R}$*, assume that* A *is discrete-time asymptotically stable, and let* $u(k)$ *be given by Equation* (88)*. Then, for all* $x(0) \in \mathbb{R}^n$*,* $H_d^{\mathrm{T}} y(k) \to 0$ *as* $k \to \infty$ *with exponential convergence.*

It follows from Proposition 8.5 that, if G_{yu} has at least one NMP output-direction zero, then there exist infinitely many unbounded input sequences such that $y(k) \in \mathcal{N}(H_d^{\mathrm{T}})$ for all $k \geq 0$. It is shown in Section 6 that, if $y(k) \in \mathcal{N}(H_d^{\mathrm{T}})$, then Θ converges independently of y. Therefore, if G_{yu} has NMP output-direction zeros, then Θ may converge to a controller producing an unbounded input sequence which drives $y(k)$ to $\mathcal{N}(H_d^{\mathrm{T}})$, namely an unstable controller with a pole (or poles) located at the NMP output-subspace zero(s) of G_{yu}. In this case, unless G_{yu} and G_{yu}^{L} have the same NMP transmission zeros, since the control input is unbounded, the performance output y is also unbounded. In the next subsection, we verify that these heuristic arguments

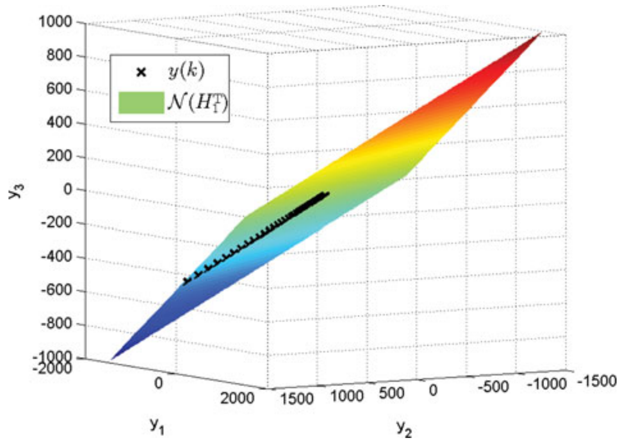

Figure 10. This figure illustrates the phase portrait for $k \in [100, 200]$ of the unbounded performance output y for Example 4.5 shown in Figure 6. Since the output-subspace zeros of G_{yu} are NMP, the unbounded output y grows without bound on the surface $\mathcal{N}(H_1^{\mathrm{T}})$. Since y is contained in $\mathcal{N}(H_1^{\mathrm{T}})$, the controller Θ in Figure 6 converges despite the fact that y is unbounded.

explain the closed-loop responses shown in Examples 4.4 and 4.5.

8.3.1 *Examples 4.4 and 4.5 revisited*

In Examples 4.4 and 4.5, the instantaneous adaptive controller (25) is applied to a 3×1 plant in order to reject the unmatched harmonic disturbance $w(k) = \sin\frac{2\pi}{7}k + \sin\frac{\pi}{5}k$. In both cases, the plant G_{yu} has no transmission zeros, the realisation (A, B, C) is minimal, and the eigenvalues of the open-loop system are equal. The only difference between Examples 4.4 and 4.5 is the entry $C_{(1,2)}$. In Example 4.4, Θ converges, and u and y are bounded. In Example 4.5, Θ converges, but u and y are unbounded. We now demonstrate that, in both cases, as k increases, $y(k)$ approaches $\mathcal{N}(H_d^{\mathrm{T}})$. Furthermore, we show that, in Example 4.5, the instability is due to the presence of an NMP output-direction zero. Note that, since G_{yu} is tall in Examples 4.4 and 4.5, $\mathcal{R}(H_d^{\mathrm{T}}) = \mathbb{R}^{l_u}$, which, since G_{yu} has no transmission zeros, implies that G_{yu} has no input-subspace zeros.

Example 8.6 (Example 4.5 revisited)**:** We first verify that $y(k)$ approaches $\mathcal{N}(H_d^{\mathrm{T}})$ as Θ converges. First, note that $d = 1$ and $H_1 = [-1\ 1\ 3]^{\mathrm{T}}$. Hence, $\mathcal{N}(H_1^{\mathrm{T}})$ is the plane described by $ay_1 + by_2 + cy_3 = 0$, where a, b, and c satisfy

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathcal{R}(H_1) = \mathrm{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix} \right\}.$$

The phase portrait of $y(k)$ for $k \in [100, 200]$ illustrated in Figure 10 shows that, as the controller converges, y

grows without bound on the surface $\mathcal{N}(H_1^{\rm T})$, which is the coloured surface in the figure. Therefore, even though y grows without bound in Figure 6, since $y \in \mathcal{N}(H_1^{\rm T})$, Θ converges.

We now investigate the output-subspace zeros of the plant. It is easy to verify that $(A, B, H_1^{\rm T}C)$ is minimal. Therefore, Equation (79) can be used to solve for the output-subspace zeros of G_{yu}, which are given by $\text{tzeros}(G_{yu}^{\rm L}) = \{-0.2954, 1.0863\}$. Therefore, G_{yu} has an NMP output-subspace zero at 1.0863. Computing the controller poles at $k = 200$, Figure 11 shows that, as Θ converges, one controller pole is located near the NMP output-subspace zero location 1.0863. Thus, the results of Example 4.5 can be evaluated as follows. The unstable controller pole at the NMP output-subspace zero location causes the input signal u to diverge. Since G_{yu} is minimum phase, the performance output y also diverges due to the unbounded input. However, since G_{yu} has an NMP output-subspace zero near the unstable controller pole location, it follows from Proposition 8.5 that y approaches $\mathcal{N}(H_1^{\rm T})$. Therefore, it follows from the results of Section 6 that Θ converges.

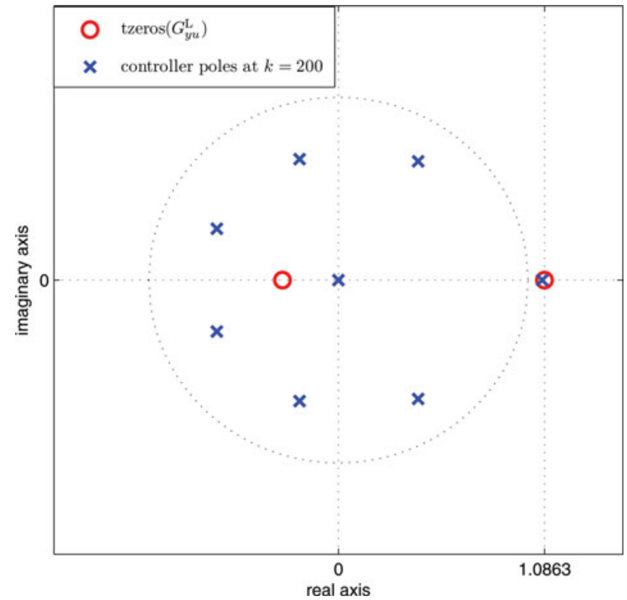

Figure 11. This figure illustrates the output-subspace zeros of the plant in Example 4.5 along with the poles of the adaptive controller at $k = 200$, whose time evolution is shown in Figure 6. The adaptive controller places a pole near the NMP output-subspace zero of G_{yu}, which is located at 1.0863. This unstable pole-zero cancellation is the cause of the unbounded control input shown in Figure 6. Note that the NMP output-subspace zero is not a transmission zero of G_{yu}, and thus the performance output shown in Figure 6 is also unbounded.

Example 8.7 (Example 4.4 revisited)**:** We now revisit Example 4.4, where u and y are bounded, y does not converge, and Θ converges. Figure 12 shows the phase portrait of $y(k)$ in $\mathbb{R}^3$ for $k \in [800, 1000]$. As shown in Figure 12, y oscillates on the surface $\mathcal{N}(H_1^{\rm T})$, which drives $\Delta\Theta$ to zero as k increases, as shown in Figure 5.

We now investigate the output-subspace zeros of G_{yu}. Since $(A, B, H_1^{\rm T}C)$ is minimal, we use Equation (79) to obtain $\text{tzeros}(G_{yu}^{\rm L}) = \{-0.3362, 0.9544\}$. Note that the output-subspace zeros for Example (4.4) are minimum phase, and, as shown in Figure 5, both u and y are bounded.

9. RCAC with η-modification for NMP subspace zeros

As shown in Examples 4.2 and 4.5, the update laws of Section 3 may converge to a controller that cancels NMP subspace zeros, leading to an unbounded control input and possibly unbounded performance output. We now modify the update laws of Section 3 in order to prevent the controller from generating an unbounded control input. This is done by extending the retrospective cost function to include a performance-dependent control penalty term. This approach is related to the leakage modification for robust adaptive control (Astrom & Wittenmark, 1995; Ioannou & Sun, 1996; Narendra & Annaswamy, 1989). We apply the modified RCAC update laws to Examples 4.2 and 4.5 to demonstrate this approach.

9.1 *Instantaneous update law with η-modification*

For each $k \geq 1$, we define the *modified instantaneous cost function*

$$\begin{aligned}\tilde{J}_{\rm ins}(\hat{\Theta}, k) &\triangleq \\ &\hat{y}^{\rm T}(\hat{\Theta}, k)\hat{y}(\hat{\Theta}, k) + \mu[\hat{\Theta} - \Theta(k-1)]^{\rm T}[\hat{\Theta} - \Theta(k-1)] \\ &\quad + \eta(k)\hat{\Theta}^{\rm T}\Phi^{\rm T}(k-d-1)\Phi(k-d-1)\hat{\Theta}, \qquad (89)\end{aligned}$$

where

$$\eta(k) \triangleq \eta_0 + \eta_1\|y(k)\|^2, \qquad (90)$$

$\eta_0 \geq 0$, and $\eta_1 \geq 0$. Substituting Equation (18) into Equation (89), we have

$$\tilde{J}_{\rm ins}(\hat{\Theta}, k) = \hat{\Theta}^{\rm T}\tilde{\Gamma}_1(k)\hat{\Theta} + \Gamma_2^{\rm T}(k)\hat{\Theta} + \Gamma_3(k), \qquad (91)$$

where

$$\begin{aligned}\tilde{\Gamma}_1(k) &\triangleq \Gamma_1(k) + \eta(k)\Phi^{\rm T}(k-d-1)\Phi(k-d-1) \\ &= \Phi^{\rm T}(k-d-1)[H_d^{\rm T}H_d + \eta(k)I_{l_u}]\Phi(k-d-1) \\ &\quad + \mu I_{l_u n_{\rm c}(l_u+l_y)}. \qquad (92)\end{aligned}$$

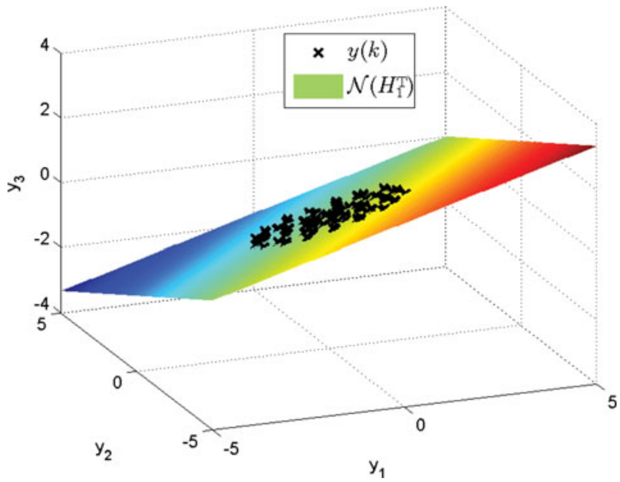

Figure 12. This figure illustrates the phase portrait for $k \in [800, 1000]$ of the output y for Example 4.4 shown in Figure 6. Since the output-subspace zeros of G_{yu} are minimum phase, the performance output y is bounded, and oscillates on the surface $\mathcal{N}(H_1^{\mathrm{T}})$. Since y is contained in $\mathcal{N}(H_1^{\mathrm{T}})$, the controller Θ converges.

The terms $\Gamma_2(k)$ and $\Gamma_3(k)$ in Equation (91) are identical to those in Equation (22). Since $\tilde{\Gamma}_1(k)$ is positive definite, $\tilde{J}_{\mathrm{ins}}(\hat{\Theta}, k)$ has the unique global minimiser

$$\Theta(k) = -\frac{1}{2}\tilde{\Gamma}_1^{-1}(k)\Gamma_2(k), \tag{93}$$

which is the instantaneous RCAC update law with η-modification.

The modified cost function (89) includes an additional term with the weighting $\eta(k)$, which penalises $\|\Phi(k-d-1)\hat{\Theta}\|$. This term tends to push the unique global minimiser of Equation (89) towards $\mathcal{N}(\Phi(k-d-1))$, which drives Θ towards a controller that would have generated $u(k-d) = 0$, if it had been used in place of $\Theta(k-d)$. The modified cost (89) thus indirectly penalises the control effort. Furthermore, note that, if $\eta_1 > 0$, then $\eta(k)$ is an increasing function of $\|y\|$. Therefore, if y diverges, then $\eta(k)\hat{\Theta}^{\mathrm{T}}\Phi^{\mathrm{T}}(k-d-1)\Phi(k-d-1)\hat{\Theta}$ dominates Equation (89), and the optimisation problem is approximately $\min_{\hat{\Theta}} \|\Phi(k-d-1)\hat{\Theta}\|$. Choosing $\eta_1 > 0$ thus prevents the situation in Example 4.5, where the adaptive controller destabilises an open-loop plant and leads to an unbounded performance variable y.

9.2 *Cumulative update law with η-modification*

For each $k \geq 1$, we define the *modified cumulative cost function*

$$\tilde{J}_{\mathrm{cum}}(\hat{\Theta}, k) \triangleq \sum_{i=1}^{k} \hat{y}^{\mathrm{T}}(\hat{\Theta}, i)\hat{y}(\hat{\Theta}, i) + \eta(i)\hat{\Theta}^{\mathrm{T}} \cdot \Phi^{\mathrm{T}}(i-d-1)\Phi(i-d-1)\hat{\Theta} + \hat{\Theta}^{\mathrm{T}}P_0^{-1}\hat{\Theta}, \tag{94}$$

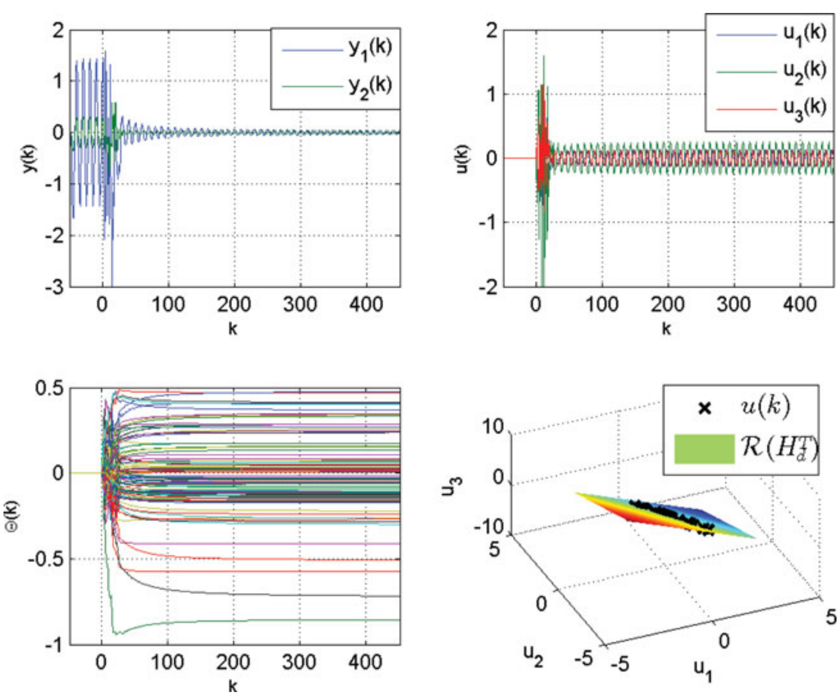

Figure 13. Example 4.2, RCAC with η-modification: We consider the same plant and disturbance as in Example 4.2, and apply the cumulative update law with η-modification. We use $n_{\mathrm{c}} = 6$, $P_0 = I$, $\eta_0 = 0.1$, and $\eta_1 = 0.05$. Despite the unmodelled NMP input-subspace zero, Θ converges, y is driven towards zero, and u is bounded.

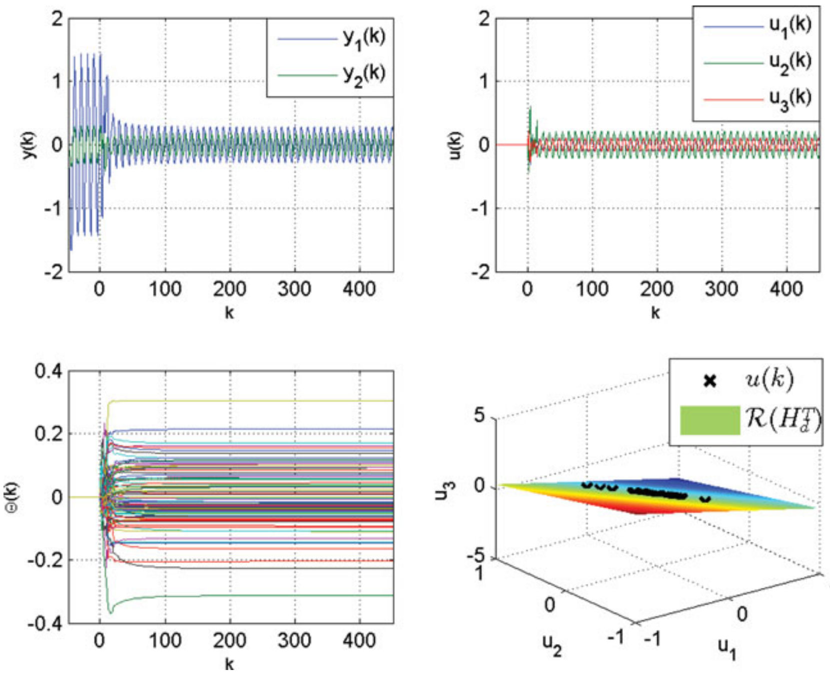

Figure 14. Example 4.2, RCAC with η-modification: Increasing the constant penalty term η_0 to 1 leads to a reduction in the control effort as well as a degradation in the steady-state performance level. Thus, η-modification introduces a trade-off between control effort and steady-state performance.

where $\eta(i)$ is as in Equation (90). Substituting Equation (18) into Equation (94), we have

$$J_{\mathrm{cum}}(\hat{\Theta}, k) = \hat{\Theta}^{\mathrm{T}}\tilde{\mathcal{C}}_1(k)\hat{\Theta} + \mathcal{C}_2^{\mathrm{T}}(k)\hat{\Theta} + \mathcal{C}_3(k), \qquad (95)$$

where

$$\tilde{\mathcal{C}}_1(k) \stackrel{\triangle}{=} \mathcal{C}_1(k) + \sum_{i=1}^{k} \eta(i)\Phi^{\mathrm{T}}(i-d-1)\Phi(i-d-1)$$
$$= \sum_{i=1}^{k} \Phi^{\mathrm{T}}(i-d-1)[H_d^{\mathrm{T}}H_d + \eta(i)I_{l_u}]\Phi(i-d-1) + P_0^{-1}. \qquad (96)$$

The terms $\mathcal{C}_2(k)$ and $\mathcal{C}_3(k)$ in Equation (95) are identical to those in Equation (29). Furthermore, defining $\tilde{\mathcal{C}}_1(0) \stackrel{\triangle}{=} P_0^{-1}$, we can rewrite Equation (96) in the recursive form

$$\tilde{\mathcal{C}}_1(k) = \tilde{\mathcal{C}}_1(k-1) + \Phi^{\mathrm{T}}(k-d-1)[H_d^{\mathrm{T}}H_d + \eta(k)I_{l_u}] \cdot \Phi(k-d-1). \qquad (97)$$

Since $\tilde{\mathcal{C}}_1(k)$ is positive definite, $\tilde{J}_{\mathrm{cum}}(\hat{\Theta}, k)$ has the unique global minimiser

$$\Theta(k) = -\frac{1}{2}\tilde{\mathcal{C}}_1^{-1}(k)\mathcal{C}_2(k), \qquad (98)$$

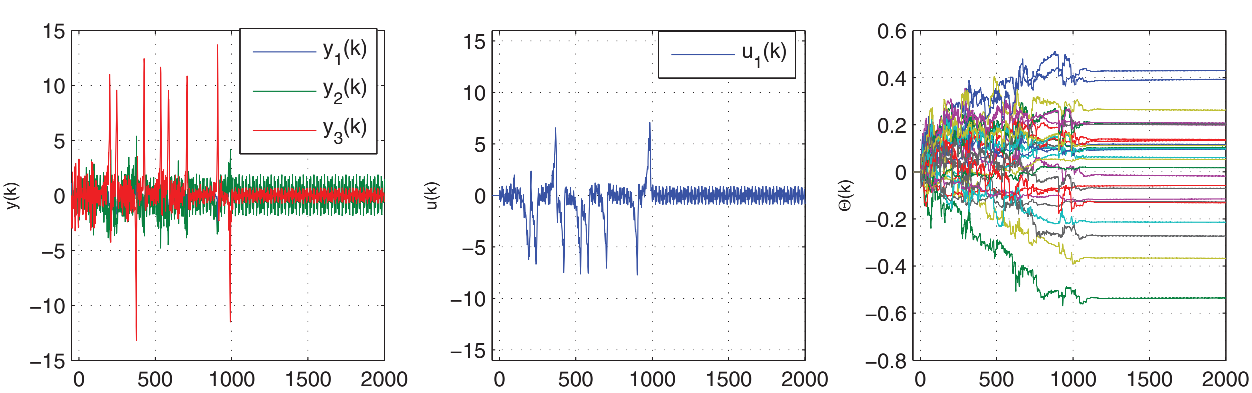

Figure 15. Example 4.5, RCAC with η-modification: We consider the same plant and disturbance as in Example 4.5, and apply the instantaneous update law with η-modification. We use $n_c = 7$, $\mu = 20$, $\eta_0 = 0$, and $\eta_1 = 0.01$. Despite the unmodelled NMP output-subspace zero, Θ converges, and u and y are bounded.

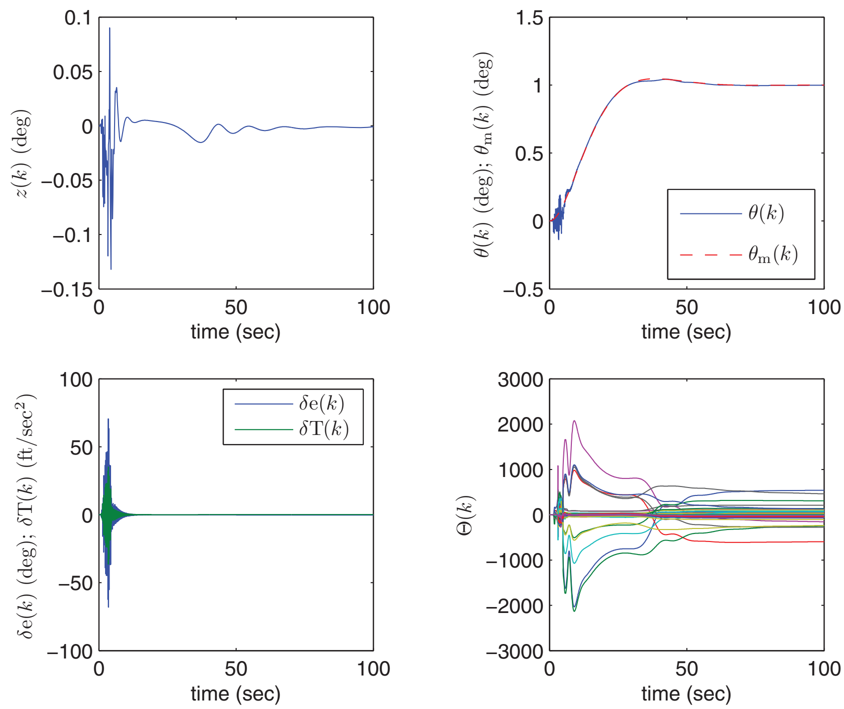

Figure 16. Boeing 747 longitudinal dynamics: We apply RCAC without η-modification. Although the pitch angle follows the output of the reference model, the elevator deflection and thrust excursion oscillate during the transient with peak magnitudes 71 deg and 36 ft/sec^2, respectively. The control inputs are not within practical saturation limits.

which is the cumulative RCAC update law with η-modification. The rationale for the η-modification is the same as for the instantaneous cost as discussed in Section 9.1.

9.3 *Examples 4.2 and 4.5 revisited with η-modification*

In this section, we apply the RCAC update laws with η-modification to Examples 4.2 and 4.5.

Example 9.1 (Example 4.2, cumulative RCAC with η-modification)**:** We consider the plant and unmatched harmonic disturbance in Example 4.2. We use the same tuning parameters $n_c = 6$, $P_0 = I$, let $\eta_0 = 0.1$ and $\eta_1 = 0.05$, and apply the cumulative update law with η-modification. Figure 13 shows that η-modification does not alter the input subspace, that is, u is still contained in $\mathcal{R}(H_d^{\mathrm{T}})$. Although the plant has an unmodelled NMP input-subspace zero near -1.0555, the control penalty prevents the control input from growing without bound, as shown in Figure 13.

As shown in Section 9.2, the modified cost function (94) has the additional control weighting $\eta(k)$. The term $\eta_1 \|y(k)\|^2$ in Equation (90) vanishes as y approaches zero, but the constant term η_0 does not vanish. Therefore, for $\eta_0 > 0$, we expect a trade-off between control effort and closed-loop performance. To demonstrate this trade-off, we keep n_c and P_0 the same, but increase η_0 to 1. Figure 14

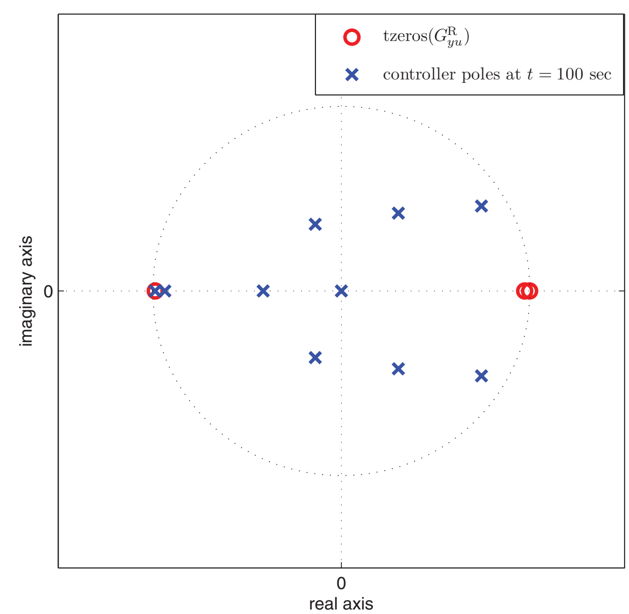

Figure 17. This figure illustrates the input-subspace zeros of the Boeing 747 longitudinal dynamics along with the poles of the adaptive controller at $t = 100$ sec. The adaptive controller places a pole near the input-subspace zero -0.9857. This pole-zero cancellation near the unit circle causes large transients in the elevator deflection and thrust excursion, as illustrated in Figure 16.

shows the closed-loop response with $\eta_0 = 1$. Comparing Figure 14 to Figure 13, we observe that, as η_0 increases, the control effort is reduced during transients as well as in steady state, but with a degradation in the steady-state performance level.

Example 9.2 (Example 4.5, instantaneous RCAC with η-modification)**:** We consider the plant and unmatched disturbance in Example 4.5. We use the same tuning parameters $n_c = 7, \mu = 20$, let $\eta_0 = 0$ and $\eta_1 = 0.01$, and apply the instantaneous update law with η-modification. Although the plant has an unmodelled NMP output-subspace zero near 1.0863, the control penalty prevents the control input u and the performance output y from growing without bound, as shown in Figure 15.

9.4 *Numerical example: Boeing 747 longitudinal dynamics*

Consider the longitudinal dynamics of a Boeing 747 aircraft, linearised about steady flight at 40, 000 ft and 774 ft/sec. The control inputs to the longitudinal dynamics are taken to be elevator deflection and thrust. The linearised equations of motion are thus given by Santillo and Bernstein (2010):

$$\begin{bmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} -0.003 & 0.039 & 0 & -0.322 \\ -0.065 & -0.319 & 7.74 & 0 \\ 0.02 & -0.101 & -0.429 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix} + \begin{bmatrix} 0.01 & 1 \\ -0.18 & -0.04 \\ -1.16 & 0.598 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta \mathrm{e} \\ \delta \mathrm{T} \end{bmatrix}, \tag{99}$$

where the state variables u, w, q, and θ are the forward-speed, vertical-speed, pitch-rate, and pitch perturbations, respectively. Furthermore, the control inputs δe and δT represent elevator deflection (deg) and thrust excursion (ft/sec^2), respectively. The control objective is to have the pitch perturbation follow the output θ_m of the reference model

$$G_\mathrm{m}(s) = \frac{0.0131}{s^2 + 0.16s + 0.0131}, \tag{100}$$

whose input is the exogenous model reference command r.

We discretise Equation (99) using a zero-order hold and a sampler with sampling period $h = 0.1$ sec/sample. We assume that samples of θ and r are measured. We thus

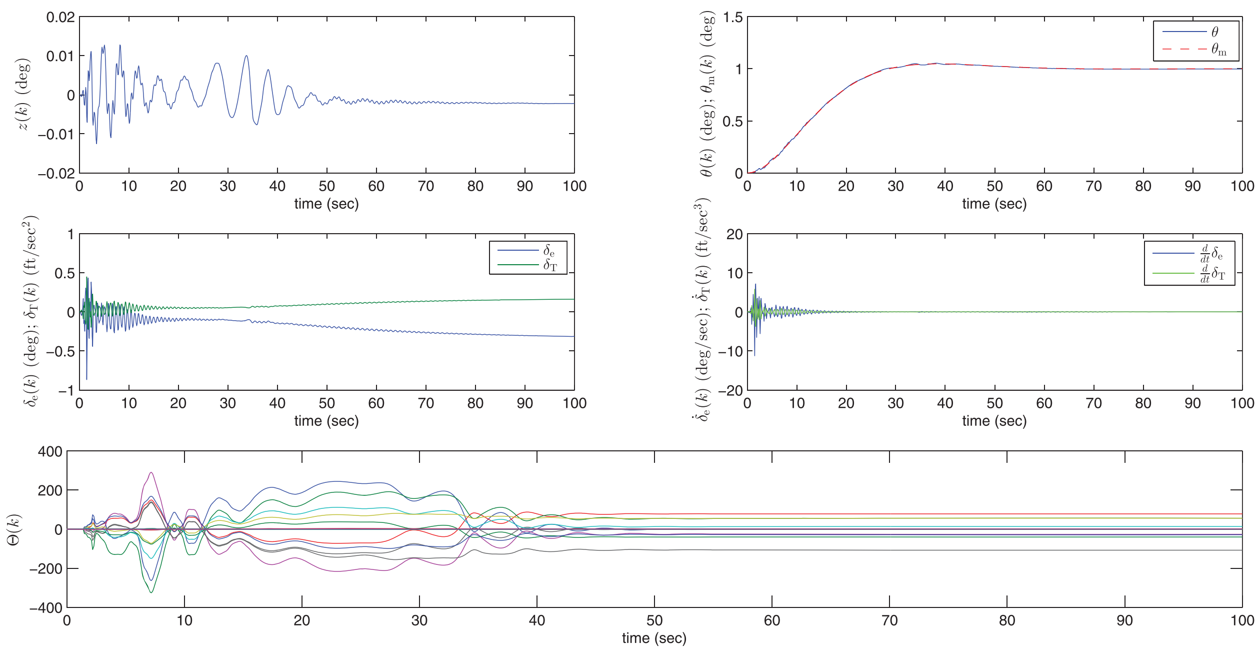

Figure 18. Boeing 747 longitudinal dynamics: We apply RCAC with η-modification. The pitch angle follows the output of the reference model, the peak elevator deflection magnitude is less than 0.9 deg, and the peak thrust excursion magnitude is less than 0.5 ft/sec^2. The absolute elevator deflection rate is less than 11.5 deg/sec, and the absolute thrust excursion rate is less than 6 ft/sec^3 throughout the manoeuvre. The command-following error $z(k)$ is less than 0.015 deg throughout the simulation.

have

$$y(k) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u(k) \\ w(k) \\ q(k) \\ \theta(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} r(k), \quad (101)$$

$$z(k) = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u(k) \\ w(k) \\ q(k) \\ \theta(k) \end{bmatrix} - \theta_{\mathrm{m}}(k). \quad (102)$$

Note that $y(k)$ includes a measurement of the command $r(k)$, and thus command feedforward is used. Therefore, we consider the 1×2 discretised plant G_{zu} with poles $\{0.9999 \pm \jmath 0.0067, 0.9594 \pm \jmath 0.0848\}$ and input-subspace zeros $\{-0.9857, 0.9714, 0.9972\}$. Throughout this example, the only modelling information used by the adaptive controller is the first non-zero Markov parameter $H_1 = \begin{bmatrix} -0.0057 & 0.0029 \end{bmatrix}$ of G_{zu}.

We take the model reference command to be a 1-deg step command in pitch angle, let $n_{\mathrm{c}} = 5$, and apply the cumulative update laws (35)–(37) with $P_0 = 10^{10} I$, $\eta_0 = 0$, and $\eta_1 = 0$. Figure 16 shows that the command-following error reduces to zero within 10 sec, but the elevator deflection and thrust inputs oscillate during the transient. In fact, the peak elevator deflection magnitude is 71 deg, and the peak thrust excursion magnitude is 36 ft/sec^2. Because of saturation limits, these values may be unacceptable in practice. The large transients in the control inputs are caused by the cancellation of the input-subspace zero -0.9857, as shown in Figure 17.

We now consider the same step reference command, keep n_{c} and P_0 the same, but introduce η-modification with $\eta_0 = 2000$ and $\eta_1 = 1$. The closed-loop response is shown in Figure 18. The resulting command-following error $z(k)$ does not exceed 0.015 deg throughout the simulation. Furthermore, the peak elevator deflection magnitude is less than 0.9 deg, and the peak thrust excursion magnitude is less than 0.5 ft/sec^2.

10. Conclusion

This paper provided a detailed analysis of RCAC for non-square plants. We have shown that, unlike the square case, closed-loop stability and signal boundedness are not guaranteed for minimum-phase, non-square plants. Specifically, we have shown that, due to the nature of the update law, RCAC involves two implicit squaring operations: one performed by pre-compensating the plant, the other performed by post-compensating the plant. In the wide case, pre-compensation leads to squaring-down, which incorporates additional zeros due to squaring, which we call input-subspace zeros. Similarly, in the tall case, post-compensation changes the zero structure and incorporates additional zeros, which we call output-subspace zeros. We have shown that, if the non-square plant has NMP subspace zeros, then RCAC attempts to cancel these zeros, which leads to an unbounded control input in the wide case and an unbounded control input and performance output in the tall case. Finally, in light of these findings, we extended the retrospective cost function to include a performance-dependent control penalty in order to prevent the controller from generating an unbounded control input. Establishing the boundedness and stability properties of this extension is left as future work.

Future research might focus on extending the stability proof in Hoagg et al. (2008a) to non-square plants with minimum-phase subspace zeros. Furthermore, similar to Hoagg et al. (2008a), the analysis presented in this paper is confined to the case where H_d is the only modelling information used by the adaptive control algorithm, and it is assumed that H_d has full rank. Future work might focus on extending the results developed in this paper to case where more than one Markov parameter is used, as in Hoagg (2010), Santillo and Bernstein (2010), Sumer et al. (2011), and Sumer and Bernstein (2012), as well as to the case where H_d does not have full rank. It should be noted that, in the case where H_d does not have full rank, input and zero-update output subspaces may play a role in square plants as well as non-square plants because, if H_d is square and rank deficient, then dim $\mathcal{R}(H_d^{\mathrm{T}}) < l_u$, which may lead to input-subspace zeros, and dim $\mathcal{N}(H_d^{\mathrm{T}}) > 0$, which may lead to output-subspace zeros in square plants.

References

Astrom, K.J., Hagander, P., & Sternby, J. (1984). Zeros of sampled systems. *IEEE Transactions on Automatic Control, 20*, 31–38.

Astrom, K.J., & Wittenmark, B. (1995). *Adaptive control* (2nd ed.). Reading, MA: Addison-Wesley.

Bernstein, D.S. (2009). *Matrix mathematics*. Princeton, NJ: Princeton University Press.

D'Amato, A.M., Sumer, E.D., & Bernstein, D.S. (2011). Frequency domain stability analysis of retrospective cost adaptive control for systems with unknown nonminimum-phase zeros. *Proceedings of Conference on Decision and Control* (pp. 1098–1103). Orlando, FL.

D'Amato, A.M., Sumer, E.D., Mitchell, K.S., Morozov, A.V., Hoagg, J.B, & Bernstein, D.S. (2011). Adaptive output feedback control of the NASA GTM model with unknown nonminimum-phase zeros. *Proceedings of AIAA Guidance, Navigation and Control Conference(AIAA-2011-6204)*. Portland, OR.

Davison, E.J. (1983). Some properties of minimum phase and squared-down systems. *IEEE Transactions on Automatic Control, 28*, 221–222.

Davison, E.J., & Wang, S.H. (1974). Properties and calculation of transmission zeros of linear multivariable systems. *Automatica, 10*, 643–658.

Freudenberg, J.S., & Looze, D.P. (1985). Right half plane poles and zeros and design tradeoffs in feedback systems. *IEEE Transactions on Automatic Control, 30*, 555–565.

Goodwin, G.C., & Sin, K.S. (1984). *Adaptive filtering prediction and control*. Englewood Cliffs, NJ: Prentice-Hall.

Hoagg, J.B. (2010). Retrospective cost adaptive control for nonminimum-phase discrete-time systems. Part 1: The ideal controller and error system. Part 2: The adaptive controller and stability analysis. *Proceedings of the 49th Conference on Decision and Control* (pp. 893–904). Atlanta, GA.

Hoagg, J.B., & Bernstein, D.S. (2007). Nonminimum-phase zeros: Much to do about nothing. *IEEE Control Systems Magazine, 27*, 45–57.

Hoagg, J.B., Santillo, M.A., & Bernstein, D.S. (2008a). Discrete-time adaptive command following and disturbance rejection with unknown exogenous dynamics. *IEEE Transactions on Automatic Control, 53*, 912–928.

Hoagg, J.B., Santillo, M.A., & Bernstein, D.S. (2008b). Internal model control in the shift and delta domains. *IEEE Transactions on Automatic Control, 53*, 1066–1072.

Ioannou, P.A., & Sun, J. (1996). *Robust adaptive control.* Upper Saddle River, NJ: Prentice Hall.

Karcanias, N., & Giannakopoulos, C. (1989). Necessary and sufficient conditions for zero assignment by constant squaring down. *Linear Algebra and its Applications, 122–124*, 415–446.

Karcanias, N., & Kouvaritakis, B. (1979). The output zeroing problem and its relationship to the invariant zero structure: A matrix pencil approach. *International Journal of Control, 30*, 395–415.

Kouvaritakis, B., & MacFarlane, A.G.J. (1976). Geometric approach to analysis and synthesis of system zeros. Part 1: Square systems. Part 2: Nonsquare systems. *International Journal of Control, 23*, 149–181.

Latawiec, K., Banka, S., & Tokarzewski, J. (2000). Control zeros and nonminimum phase LTI MIMO systems. *Annual Reviews in Control, 24*, 105–112.

Lavretsky, E. (2011). Adaptation along prescribed directions. *Proceedings of American Control Conference* (pp. 1748–1752). San Francisco, CA.

Leventides, J., & Karcanias, N. (2008). Structured squaring down and zero assignments. *International Journal of Control, 81*, 294–306.

MacFarlane, A.G.J., & Karcanias, N. (1976). Poles and zeros of linear multivariable systems: A survey of the algebraic, geometric and complex-variable theory. *International Journal of Control, 24*, 33–74.

Maciejowski, J.M. (1989). *Multivariable feedback design*. Wokingham: Addison-Wesley.

Misra, P. (1992). A computational algorithm for squaring-up. Part I: Zero input-output matrix. *Proceedings of the 31st Conference on Decision and Control* (pp. 149–150). Tucson, AZ.

Misra, P. (1993). Numerical algorithms for squaring-up non-square systems. Part II: General case. *Proceedings of the American Control Conference* (pp. 1573–1577). San Francisco, CA.

Narendra, K.S., & Annaswamy, A.M. (1989). *Stable adaptive systems*. Englewood Cliffs, NJ: Prentice Hall.

Rosenbrock, H.H. (1970). *State-space and multivariable theory*. New York, NY: Wiley.

Saberi, A., & Sannuti, P. (1988). Squaring down by static and dynamic compensators. *IEEE Transactions on Automatic Control, 45*, 1655–1704.

Saberi, A., & Sannuti, P. (1990). Squaring down of non-strictly proper systems. *International Journal of Control, 51*, 621–629.

Santillo, M.A., & Bernstein, D.S. (2010). Adaptive control based on retrospective cost optimization. *AIAA Journal of Guidance, Control and Dynamics, 33*, 289–304.

Schrader, C.B., & Sain, M.K. (1989). Research on system zeros: A survey. *International Journal of Control, 50*, 1407–1433.

Skogestad, S., & Postlethwaite, I. (2005). *Multivariable feedback control: Analysis and design* (2nd ed.). New York, NY: Wiley.

Soderstrom, T., & Stoica, P. (1989). *System identification*. Englewood Cliffs, NJ: Prentice Hall.

Sumer, E.D., & Bernstein, D.S. (2012). Retrospective cost adaptive control with error-dependent regularization for MIMO systems with uncertain nonminimum-phase transmission zeros. *Proceedings of AIAA Guidance Navigation and Control Conference (AIAA-2012-4670-123)*. Minneapolis, MN.

Sumer, E.D., D'Amato, A.M., & Bernstein, D.S. (2011). Robustness of retrospective cost adaptive control to Markov parameter uncertainty. In *Proceedings of Conference on Decision and Control* (pp. 6085–6090). Orlando, FL.

Tokarzewski, J. (2002). A general solution to the output-zeroing problem for MIMO LTI systems. *International Journal of Applied Mathematics and Computer Science, 12*, 161–171.

Tokarzewski, J. (2006). *Finite zeros in discrete time control systems*. Berlin: Springer-Verlag.

Vardulakis, A.I.G. (1980). Zero placement and the 'Squaring Down' problem: A polynomial approach. *International Journal of Control, 31*, 821–832.

Venugopal, R., & Bernstein, D.S. (2000). Adaptive disturbance rejection using ARMARKOV system representations. *IEEE Transactions on Control System Technology, 8*, 257–269.

Appendix 1: Proof of Lemma 3.1

Substituting Equation (24) into Equation (25), and using Equations (14) and (23) yields

$$\begin{aligned}
\Theta(k) &= \Gamma_1^{-1}(k)[\mu\Theta(k-1) - \Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}(y(k) \\
&\quad - H_d u(k-d))] \\
&= \Gamma_1^{-1}(k)[\mu\Theta(k-1) - \Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}(y(k) \\
&\quad - H_d u(k-d))] + \Gamma_1^{-1}(k)[\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}H_d\Phi \\
&\quad \cdot (k-d-1)]\Theta(k-1) - \Gamma_1^{-1}(k)[\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}H_d\Phi \\
&\quad \cdot (k-d-1)]\Theta(k-1) = \Gamma_1^{-1}(k)[\mu I + \Phi^{\mathrm{T}}(k-d-1) \\
&\quad \cdot H_d^{\mathrm{T}}H_d\Phi(k-d-1)]\Theta(k-1) \\
&\quad - \Gamma_1^{-1}(k)[\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}(y(k) - H_d u(k-d)) \\
&\quad + \Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}H_d\Phi(k-d-1)\Theta(k-1)] \\
&= \Theta(k-1) - \Gamma_1^{-1}(k)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\hat{y}(\Theta(k-1),k).
\end{aligned} \tag{A1}$$

Next, applying the matrix inversion lemma to Equation (23) and using Equation (27) yields

$$\Gamma_1^{-1}(k) = \frac{1}{\mu}[I - \Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Psi^{-1}(k)H_d\Phi(k-d-1)]. \tag{A2}$$

Now, substituting Equation (A2) into Equation (A1) yields

$$\begin{aligned}
\Theta(k) &= \Theta(k-1) - \frac{1}{\mu}\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Psi^{-1}(k)\Psi(k) \\
&\quad \cdot \hat{y}(\Theta(k-1),k) \\
&\quad + \frac{1}{\mu}\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Psi^{-1}(k)H_d\Phi(k-d-1) \\
&\quad \cdot \Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\hat{y}(\Theta(k-1),k)
\end{aligned}$$

$$= \Theta(k-1) - \frac{1}{\mu}\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Psi^{-1}(k)[\mu\hat{y}(\Theta(k-1),k)$$
$$+H_d\Phi(k-d-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\hat{y}(\Theta(k-1),k)$$
$$-H_d\Phi(k-d-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\hat{y}(\Theta(k-1),k)]$$
$$= \Theta(k-1) - \Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Psi^{-1}(k)\hat{y}(\Theta(k-1),k).$$

Appendix 2: Proof of Lemma 3.2

It follows from Equation (32) that

$$P^{-1}(k) = P^{-1}(k-1) + \Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}H_d\Phi(k-d-1). \tag{B1}$$

Applying the matrix inversion lemma to Equation (B1) and using Equation (36) yields

$$\begin{aligned}P(k) &= P(k-1) - P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\\ &\quad\cdot[I_{l_y} + H_d\Phi(k-d-1)P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}]^{-1}\\ &\quad\cdot H_d\Phi(k-d-1)P(k-1)\\ &= P(k-1) - P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k)H_d\\ &\quad\cdot\Phi(k-d-1)P(k-1).\end{aligned}$$

Hence, Equation (35) holds. Next, since $P(k) = \mathcal{C}_1^{-1}(k)$, it follows from Equations (33), (34), and (35) that

$$\begin{aligned}\Theta(k) &= -\frac{1}{2}P(k)\mathcal{C}_2^{\mathrm{T}}(k)\\ &= -\frac{1}{2}P(k-1)\mathcal{C}_2^{\mathrm{T}}(k-1) - P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\\ &\quad\cdot[y(k) - H_du(k-d)]\\ &\quad+\frac{1}{2}P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k)H_d\Phi(k-d-1)\\ &\quad\cdot P(k-1)\mathcal{C}_2^{\mathrm{T}}(k-1)\\ &\quad+P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k)H_d\Phi(k-d-1)\\ &\quad\cdot P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}[y(k) - H_du(k-d)]\\ &= \Theta(k-1) - P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k)\Lambda(k)\\ &\quad\cdot[y(k) - H_du(k-d)]\\ &\quad-P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k)H_d\\ &\quad\Phi(k-d-1)\Theta(k-1)\\ &\quad+P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k)H_d\Phi(k-d-1)\\ &\quad\cdot P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}[y(k) - H_du(k-d)]\\ &= \Theta(k-1) - P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k)\\ &\quad\cdot[H_d\Phi(k-d-1)\Theta(k-1)\\ &\quad+(I_{l_y} + H_d\Phi(k-d-1)P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\\ &\quad- H_d\Phi(k-d-1)P(k-1)\\ &\quad\cdot\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}})[y(k) - H_du(k-d)]]\\ &= \Theta(k-1) - P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k)\\ &\quad\cdot\hat{y}(\Theta(k-1),k).\end{aligned}$$

Appendix 3: Proof of Lemma 5.1

Since $\Theta(0) = 0$, it follows from Equation (26) that

$$\begin{aligned}\Phi\Theta(1) &= -\Phi\Phi^{\mathrm{T}}(-d)H_d^{\mathrm{T}}\Psi^{-1}(1)\hat{y}(\Theta(0),1)\\ &= -(I_{l_u}\otimes\phi^{\mathrm{T}})(I_{l_u}\otimes\phi(-d))H_d^{\mathrm{T}}\Psi^{-1}(1)\hat{y}(\Theta(0),1)\\ &= -H_d^{\mathrm{T}}\phi^{\mathrm{T}}\phi(-d)\Psi^{-1}(1)\hat{y}(\Theta(0),1)\\ &\in \mathcal{R}(H_d^{\mathrm{T}}).\end{aligned}$$

Hence, Equation (42) holds for $k = 1$. Next, suppose that Equation (42) holds for $k-1$. Then, there exists $v_\Phi(k-1) \in \mathbb{R}^{l_y}$ such that $\Phi\Theta(k-1) = H_d^{\mathrm{T}}v_\Phi(k-1)$. Multiplying Equation (26) on the left by Φ yields

$$\begin{aligned}\Phi\Theta(k) &= H_d^{\mathrm{T}}v_\Phi(k-1) - \Phi\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Psi^{-1}(k)\\ &\quad\cdot\hat{y}(\Theta(k-1),k)\\ &= H_d^{\mathrm{T}}v_\Phi(k-1) - H_d^{\mathrm{T}}\phi^{\mathrm{T}}\phi(k-d-1)\Psi^{-1}(k)\\ &\quad\cdot\hat{y}(\Theta(k-1),k)\\ &= H_d^{\mathrm{T}}[v_\Phi(k-1) - \phi^{\mathrm{T}}\phi(k-d-1)\Psi^{-1}(k)\\ &\quad\cdot\hat{y}(\Theta(k-1),k)]\\ &\in \mathcal{R}(H_d^{\mathrm{T}}).\end{aligned}$$

By induction, Equation (42) holds for all $k \geq 1$.

Appendix 4: Proof of Lemma 5.3

To show (i), note that it follows from Equation (35) that

$$\begin{aligned}\Phi_1P(1)\Phi_2^{\mathrm{T}}H_d^{\mathrm{T}} &= \Phi_1P(0)\Phi_2H_d^{\mathrm{T}} - \Phi_1P(0)\Phi^{\mathrm{T}}(-d)H_d^{\mathrm{T}}\Lambda^{-1}(1)H_d\\ &\quad\cdot\Phi(-d)P(0)\Phi_2^{\mathrm{T}}H_d^{\mathrm{T}},\end{aligned}$$

and, since $P(0) = \beta I$,

$$\begin{aligned}\Phi_1P(1)\Phi_2^{\mathrm{T}}H_d^{\mathrm{T}} &= \beta(I_{l_u}\otimes\phi_1^{\mathrm{T}})(I_{l_u}\otimes\phi_2)H_d^{\mathrm{T}} + \beta(I_{l_u}\otimes\phi_1^{\mathrm{T}})\\ &\quad\cdot(I_{l_u}\otimes\phi(-d))H_d^{\mathrm{T}}\Lambda^{-1}(1)H_d\Phi(-d)\beta\Phi_2H_d^{\mathrm{T}}\\ &= H_d^{\mathrm{T}}N_{\Phi_1,\Phi_2}(1),\end{aligned}$$

where $N_{\Phi_1,\Phi_2}(1) \triangleq \beta\phi_1^{\mathrm{T}}\phi_2I_{l_y} + \beta^2\phi_1^{\mathrm{T}}\phi(-d)\Lambda^{-1}(1)H_d\Phi(-d)\Phi_2^{\mathrm{T}}H_d^{\mathrm{T}} \in \mathbb{R}^{l_y\times l_y}$. Thus, (i) holds for $k = 1$. Now, suppose (i) holds for $k-1 \geq 1$. Multiplying Equation (35) on the left by Φ_1 and on the right by $\Phi_2H_d^{\mathrm{T}}$ yields

$$\begin{aligned}\Phi_1P(k)\Phi_2H_d^{\mathrm{T}} &= \Phi_1P(k-1)\Phi_2H_d^{\mathrm{T}}\\ &\quad-\Phi_1P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k)H_d\\ &\quad\cdot\Phi(k-d-1)P(k-1)\Phi_2H_d^{\mathrm{T}}\\ &= H_d^{\mathrm{T}}N_{\Phi_1,\Phi_2}(k-1) - H_d^{\mathrm{T}}N_{\Phi_1,\Phi(k-d-1)}(k-1)\\ &\quad\cdot\Lambda^{-1}(k)H_dH_d^{\mathrm{T}}N_{\Phi(k-d-1),\Phi_2}(k-1)\\ &= H_d^{\mathrm{T}}N_{\Phi_1,\Phi_2}(k),\end{aligned}$$

where

$$\begin{aligned}N_{\Phi_1,\Phi_2}(k) &\triangleq N_{\Phi_1,\Phi_2}(k-1) - N_{\Phi_1,\Phi(k-d-1)}(k-1)\\ &\quad\cdot\Lambda^{-1}(k)H_dH_d^{\mathrm{T}}N_{\Phi(k-d-1),\Phi_2}(k-1) \in \mathbb{R}^{l_y\times l_y},\end{aligned}$$

and thus, if (i) holds with k replaced by $k-1$, then (i) holds for k. Therefore, by induction, (i) holds for all $k \geq 1$.

Next, note that, since $\Theta(0) = 0$ and $P(0) = \beta I$, it follows from Equation (37) that $\Theta(1) = H_d^{\mathrm{T}}v_\Phi(1)$, where $v_\Phi(1) \triangleq$

$-\beta\phi^{\mathrm{T}}\phi(-d)\Lambda^{-1}(1)\hat{y}(\Theta(0),1)$. Therefore, $\Phi\Theta(1)\in\mathcal{R}(H_d^{\mathrm{T}})$, and thus (ii) holds for $k=1$. Next, suppose (ii) holds for $k-1\geq 1$ so that $\Phi\Theta(k-1)\in\mathcal{R}(H_d^{\mathrm{T}})$. Then, there exists $v_\Phi(k-1)\in\mathbb{R}^{l_y}$ such that $\Phi\Theta(k-1)=H_d^{\mathrm{T}}v_\Phi(k-1)$. Multiplying Equation (37) on the left by Φ and using (i) yields

$$\begin{aligned}\Phi\Theta(k) &= \Phi\Theta(k-1)-\Phi P(k-1)\Phi^{\mathrm{T}}(k-d-1)H_d^{\mathrm{T}}\Lambda^{-1}(k)\\ &\quad\cdot\hat{y}(\Theta(k-1),k)\\ &= H_d^{\mathrm{T}}v_\Phi(k-1)-H_d^{\mathrm{T}}N_{\Phi,\Phi(k-d-1)}(k-1)\Lambda^{-1}(k)\\ &\quad\cdot\hat{y}(\Theta(k-1),k)\\ &= H_d^{\mathrm{T}}v_\Phi(k),\end{aligned}$$

where

$$v_\Phi(k)=v_\Phi(k-1)-N_{\Phi,\Phi(k-d-1)}(k-1)\Lambda^{-1}(k)\hat{y}(\Theta(k-1),k).$$

Hence, if (ii) holds for $k-1$, then (ii) holds for k. Therefore, by induction, (ii) holds for all $k\geq 1$.

Appendix 5: Proof of Theorem 6.2

We show that $\{\Theta(k)\}_{k=0}^{\infty}$ is a Cauchy sequence. Let N, m_1, m_2 be positive integers such that $m_2>m_1>N$. Then, it follows from Lemma 6.1 that

$$\begin{aligned}&\|\Theta(m_1)-\Theta(m_2)\|\\ &\quad=\left\|\Theta(0)+\sum_{i=1}^{m_1}\Delta\Theta(i)-\sum_{i=1}^{m_2}\Delta\Theta(i)-\Theta(0)\right\|\\ &\quad=\left\|\sum_{i=m_1+1}^{m_2}\Delta\Theta(i)\right\|\\ &\quad\leq\sum_{i=m_1+1}^{m_2}\|\Delta\Theta(i)\|\leq\sum_{i=m_1+1}^{m_2}\sigma_{\max}(\mathcal{B}(i))\left\|H_1^{\mathrm{T}}y(i)\right\|.\end{aligned}$$

Since $\mathcal{B}$ is bounded, it follows that $\bar{\sigma}(\mathcal{B})\triangleq\sup_{k\geq1}\sigma_{\max}(\mathcal{B}(k))\in[0,\infty)$. Note that, if $\bar{\sigma}(\mathcal{B})=0$, then, for all $k\geq 1$, $\Delta\Theta(k)=0$, and thus $\lim_{k\to\infty}\Theta(k)=\Theta(0)$. Now, assume that $\bar{\sigma}(\mathcal{B})>0$. Then,

$$\begin{aligned}\|\Theta(m_1)-\Theta(m_2)\|&\leq\bar{\sigma}(\mathcal{B})\sum_{i=m_1+1}^{m_2}\left\|H_1^{\mathrm{T}}y(i)\right\|\\ &\leq\bar{\sigma}(\mathcal{B})\sum_{i=m_1+1}^{m_2}\alpha\gamma^i\leq\bar{\sigma}(\mathcal{B})\sum_{i=N}^{\infty}\alpha\gamma^i\\ &=\bar{\sigma}(\mathcal{B})\gamma^N\sum_{i=0}^{\infty}\alpha\gamma^i=\bar{\sigma}(\mathcal{B})\gamma^N\frac{\alpha}{1-\gamma}.\end{aligned}\tag{E1}$$

Hence, since $0<\gamma<1$, it follows from Equation (E1) that, for all $\varepsilon>0$, there exists N such that, for all $m_1,m_2>N$, $\|\Theta(m_1)-\Theta(m_2)\|<\varepsilon$. Therefore, $\{\Theta(k)\}_{k=0}^{\infty}$ is a Cauchy sequence and thus Θ converges.

Appendix 6: Proof of Proposition 6.5

For all $k\geq 0$, it follows from Equation (50) that

$$\begin{aligned}\|\Delta\Theta(k)\|&=\|\mathcal{A}(k)\cdots\mathcal{A}(1)\Delta\Theta(0)\|\\ &\leq\sigma_{\max}(\mathcal{A}(k))\cdots\sigma_{\max}(\mathcal{A}(1))\|\Delta\Theta(0)\|.\end{aligned}\tag{F1}$$

Since Φ is bounded, it follows that $\sigma_{\max}(H_2\Phi)$ is bounded, and thus, it follows from Equation (52) that $\bar{\sigma}(\mathcal{A})<1$. Therefore, it follows from Equation (F1) that

$$\|\Delta\Theta(k)\|\leq\|\Delta\Theta(0)\|\gamma^k,\tag{F2}$$

where $\gamma\triangleq\bar{\sigma}(\mathcal{A})<1$. Therefore, the zero solution of Equation (50) is globally exponentially stable.

Appendix 7: Proof of Proposition 6.6

For all $k\geq 0$, define

$$V(\Delta\Theta(k))\triangleq\|\Delta\Theta(k)\|^2.\tag{G1}$$

It follows from Equation (50) that

$$V(\Delta\Theta(k+m))=\Delta\Theta^{\mathrm{T}}(k)Q_m(k+1)Q_m^{\mathrm{T}}(k+1)\Delta\Theta(k),\tag{G2}$$

and thus, for all $k\geq 0$,

$$V(\Delta\Theta(k+m))\leq\bar{\sigma}^2(Q_m)V(\Delta\Theta(k)).$$

Hence, for all non-negative integers k, N, it follows that

$$\|\Delta\Theta(k+mN)\|\leq\bar{\sigma}^N(Q_m)\|\Delta\Theta(k)\|.\tag{G3}$$

Rewriting $k=mN+r$, where $0\leq r\leq m-1$, and N is a non-negative integer, it follows from Equation (G3) that, for all $k\geq 0$,

$$\begin{aligned}\|\Delta\Theta(k)\|&=\|\Delta\Theta(r+mN)\|\leq\bar{\sigma}^N(Q_m)\|\Delta\Theta(r)\|\\ &\leq\bar{\sigma}^N(Q_m)\max_{0\leq i\leq m-1}\|\Delta\Theta(i)\|.\end{aligned}\tag{G4}$$

Since $0\leq r<m$, it follows that $k/m-1<N$. Therefore, since $\bar{\sigma}(Q_m)\in(0,1)$, it follows from Equation (G4) that

$$\|\Delta\Theta(k)\|\leq\bar{\sigma}^{-1}(Q_m)\bar{\sigma}^{k/m}(Q_m)\max_{0\leq i\leq m-1}\|\Delta\Theta(i)\|.\tag{G5}$$

Since $\bar{\sigma}(\mathcal{A})\leq 1$, it follows from Equation (50) that $\max_{0\leq i\leq m-1}\|\Delta\Theta(i)\|=\|\Delta\Theta(0)\|$. Therefore, it follows from Equation (G5) that, for all $k\geq 0$,

$$\|\Delta\Theta(k)\|\leq\alpha\|\Delta\Theta(0)\|\gamma^k,$$

where $\alpha\triangleq\bar{\sigma}^{-1}(Q_m)$, $\gamma\triangleq\bar{\sigma}^{1/m}(Q_m)$. Since $\bar{\sigma}(Q_m)\in(0,1)$ it follows that $\alpha\geq 1$ and $\gamma<1$, which implies that the zero solution of Equation (50) is globally exponentially stable.

Appendix 8: Proof of Proposition 6.8

Let $r\triangleq l_u n_c(l_u+l_y)$, and, without loss of generality, assume that $k_1<k_2$. Since $\bar{\sigma}(\mathcal{A})\in[0,1]$, it follows from Equation (53) that

$$\begin{aligned}\sigma_{\max}(Q_m(k))&=\sigma_{\max}(\mathcal{A}(k)\cdots\mathcal{A}(k_1)\cdots\mathcal{A}(k_2)\cdots\mathcal{A}(k+m-1))\\ &\leq\sigma_{\max}(\mathcal{A}(k_1)\cdots\mathcal{A}(k_2)).\end{aligned}\tag{H1}$$

Since $\Omega_0\leq\Omega(\phi(k_1-3),\phi(k_2-3)\leq\pi-\Omega_0$, it follows from Lemma 6.7 that there exists $l\in\{k_1,\ldots,k_2-1\}$ such that

$$\tilde{\Omega}_0\leq\Omega(\phi(l-3),\phi(l-2))\leq\pi-\tilde{\Omega}_0,\tag{H2}$$

where

$$\tilde{\Omega}_0 \triangleq \frac{1}{k_2 - k_1}\Omega_0.$$

Since $k_1, k_2 \in \{k, \ldots, k+m-1\}$, it follows that $\tilde{\Omega}_0 \geq \frac{\Omega_0}{m-1} > 0$. Furthermore, since $\Omega_0 \leq \pi/2$ and $k_1 < k_2$, it follows that $\tilde{\Omega}_0 \leq \pi/2$. Therefore, $\tilde{\Omega}_0 \in [\frac{\Omega_0}{m-1}, \pi/2]$.

Now, it follows from Equation (H1) that

$$\begin{aligned}\sigma_{\max}(Q_m(k)) &\leq \sigma_{\max}(\mathcal{A}(k_1)\cdots\mathcal{A}(l)\mathcal{A}(l+1)\cdots\mathcal{A}(k_2))\\ &\leq \sigma_{\max}(\mathcal{A}(l)\mathcal{A}(l+1)),\end{aligned}$$

and thus,

$$\begin{aligned}\sigma_{\max}(Q_m(k)) &= \max_{v\in\mathbb{R}^r\setminus\{0\}} \frac{\|Q_m(k)v\|}{\|v\|}\\ &\leq \max_{v\in\mathbb{R}^r\setminus\{0\}} \frac{\|\mathcal{A}(l)\mathcal{A}(l+1)v\|}{\|v\|}.\end{aligned} \tag{H3}$$

Next, it follows from Bernstein (2009, Fact 3.9.5) that there exists an orthogonal matrix $R \in \mathbb{R}^{n_c(l_u+l_y)\times n_c(l_u+l_y)}$ such that

$$\phi(l-3) = \alpha R\phi(l-2), \tag{H4}$$

where $\alpha \triangleq \frac{\|\phi(l-3)\|}{\|\phi(l-2)\|}$. It follows from Equations (15), (48), and (H4) that

$$\begin{aligned}\mathcal{A}(l) &= (I_{l_u}\otimes \alpha R\phi(l-2))H_2^{\mathrm{T}}[\mu I + H_2(I_{l_u}\otimes \alpha\phi^{\mathrm{T}}(l-2)R^{\mathrm{T}})\\ &\quad\cdot(I_{l_u}\otimes \alpha R\phi(l-2))]^{-1}H_2(I_{l_u}\otimes \alpha\phi^{\mathrm{T}}(l-2)R^{\mathrm{T}})\\ &= \alpha^2(I_{l_u}\otimes R)(I_{l_u}\otimes\phi(l-2))H_2^{\mathrm{T}}[\mu I + \alpha^2 H_2\\ &\quad\cdot(I_{l_u}\otimes\phi^{\mathrm{T}}(l-2))(I_{l_u}\otimes R^{\mathrm{T}})\\ &\quad\cdot(I_{l_u}\otimes R)(I_{l_u}\otimes\phi(l-2))]^{-1}H_2(I_{l_u}\otimes\phi^{\mathrm{T}}(l-2))\\ &\quad\cdot(I_{l_u}\otimes R^{\mathrm{T}})\\ &= (I_{l_u}\otimes R)\Phi^{\mathrm{T}}(l-2)H_2^{\mathrm{T}}\left[\frac{\mu}{\alpha^2}I + H_2\Phi(l-2)\right.\\ &\quad\left.\cdot\Phi^{\mathrm{T}}(l-2)H_2^{\mathrm{T}}\right]^{-1}H_2\Phi(l-2)(I_{l_u}\otimes R^{\mathrm{T}}) \qquad \text{(H5)}\\ &= (I_{l_u}\otimes R)\tilde{\mathcal{A}}(l+1)(I_{l_u}\otimes R^{\mathrm{T}}), \qquad \text{(H6)}\end{aligned}$$

where $\tilde{\mathcal{A}}(l+1)$ is given by Equation (48) with μ replaced by μ/α^2. Note that $\mathcal{R}(\tilde{\mathcal{A}}(l+1)) = \mathcal{R}(\mathcal{A}(l+1))$, $\mathcal{N}(\tilde{\mathcal{A}}(l+1)) = \mathcal{N}(\mathcal{A}(l+1))$, and, since $\tilde{\mathcal{A}}(l+1)$ is symmetric, $\mathcal{N}^{\perp}(\tilde{\mathcal{A}}(l+1)) = \mathcal{R}(\tilde{\mathcal{A}}(l+1))$. Furthermore, since G_{yu} is tall or square, H_2^{T} is right invertible. It thus follows from Equation (48) that

$$\begin{aligned}\mathcal{R}(\mathcal{A}(l+1)) &= \mathcal{R}(\tilde{\mathcal{A}}(l+1)) = \mathcal{R}(\Phi^{\mathrm{T}}(l-2))\\ &= \operatorname{span}\left\{\begin{bmatrix}\phi(l-2)\\0\\\vdots\\0\end{bmatrix},\ldots,\begin{bmatrix}0\\\vdots\\0\\\phi(l-2)\end{bmatrix}\right\},\end{aligned}$$

and thus, for all $v \in \mathbb{R}^r$,

$$\mathcal{A}(l+1)v = \sum_{i=1}^{l_u}\alpha_i\begin{bmatrix}0_{(i-1)n_c(l_u+l_y)\times 1}\\ \phi(l-2)\\ 0_{(l_u-i)n_c(l_u+l_y)\times 1}\end{bmatrix}. \tag{H7}$$

Therefore, it follows from Equations (H6) and (H7) that, for all non-zero $v \in \mathbb{R}^r$,

$$\begin{aligned}&\frac{\|\mathcal{A}(l)\mathcal{A}(l+1)v\|_2}{\|v\|}\\ &= \frac{1}{\|v\|}\left\|(I_{l_u}\otimes R)\tilde{\mathcal{A}}(l+1)\sum_{i=1}^{l_u}\alpha_i\begin{bmatrix}0_{(i-1)n_c(l_u+l_y)\times 1}\\ R^{\mathrm{T}}\phi(l-2)\\ 0_{(l_u-i)n_c(l_u+l_y)\times 1}\end{bmatrix}\right\|\\ &= \frac{1}{\|v\|}\left\|\tilde{\mathcal{A}}(l+1)\sum_{i=1}^{l_u}\alpha_i\begin{bmatrix}0_{(i-1)n_c(l_u+l_y)\times 1}\\ R^{\mathrm{T}}\phi(l-2)\\ 0_{(l_u-i)n_c(l_u+l_y)\times 1}\end{bmatrix}\right\|\\ &= \frac{1}{\|v\|}\left\|\tilde{\mathcal{A}}(l+1)\left(\sum_{i=1}^{l_u}\cos(\Omega(\phi(l-3),\right.\right.\\ &\qquad \phi(l-2)))\alpha_i\begin{bmatrix}0_{(i-1)n_c(l_u+l_y)\times 1}\\ \phi(l-2)\\ 0_{(l_u-i)n_c(l_u+l_y)\times 1}\end{bmatrix}\\ &\qquad \left.\left. + \sum_{i=1}^{l_u}\sin(\Omega(\phi(l-3),\phi(l-2)))\alpha_i\begin{bmatrix}0_{(i-1)n_c(l_u+l_y)\times 1}\\ \phi_{\mathrm{perp}}(l-2)\\ 0_{(l_u-i)n_c(l_u+l_y)\times 1}\end{bmatrix}\right)\right\|,\end{aligned} \tag{H8}$$

where $\phi_{\mathrm{perp}}(l-2) \in \mathbb{R}^{n_c(l_u+l_y)}$ is orthogonal to $\phi(l-2)$. Since $\phi_{\mathrm{perp}}(l-2)$ is orthogonal to $\phi(l-2)$, it follows that

$$\begin{aligned}&\sum_{i=1}^{l_u}\sin(\Omega(\phi(l-3),\\ &\qquad \phi(l-2)))\alpha_i\begin{bmatrix}0_{(i-1)n_c(l_u+l_y)\times 1}\\ \phi_{\mathrm{perp}}(l-2)\\ 0_{(l_u-i)n_c(l_u+l_y)\times 1}\end{bmatrix} \in \mathcal{N}(\tilde{\mathcal{A}}(l+1)),\end{aligned} \tag{H9}$$

and thus, it follows from Equations (H2), (H7), (H8), and (H9) that

$$\begin{aligned}&\frac{\|\mathcal{A}(l)\mathcal{A}(l+1)v\|}{\|v\|}\\ &= \frac{1}{\|v\|}\left\|\cos\Omega(\phi(l-3),\phi(l-2))\tilde{\mathcal{A}}(l+1)\mathcal{A}(l+1)v\right\|\\ &= \frac{|\cos\Omega(\phi(l-3),\phi(l-2))|}{\|v\|}\left\|\tilde{\mathcal{A}}(l+1)\mathcal{A}(l+1)v\right\|\\ &\leq |\cos\Omega(\phi(l-3),\phi(l-2))|\,\sigma_{\max}(\tilde{\mathcal{A}}(l+1))\sigma_{\max}(\mathcal{A}(l+1))\\ &\leq \cos\tilde{\Omega}_0.\end{aligned}$$

Since $\tilde{\Omega}_0 \in [\frac{\Omega_0}{m-1}, \pi/2]$, it follows from Equation (H3) that $\sigma_{\max}(Q_m(k)) \le \cos\frac{\Omega_0}{m-1}$, and it thus follows that $\bar{\sigma}(Q_m) \le \cos\frac{\Omega_0}{m-1} < 1$. □

Appendix 9: Proof of Theorem 6.9

Let N, k_1, k_2 be positive integers such that $k_2 > k_1 > N$. Then,

$$\|\Theta(k_1) - \Theta(k_2)\| = \left\|\Theta(0) + \sum_{i=1}^{k_1}\Delta\Theta(i) - \Theta(0) - \sum_{i=1}^{k_2}\Delta\Theta(i)\right\| = \left\|\sum_{i=k_1+1}^{k_2}\Delta\Theta(i)\right\|. \tag{I1}$$

For all $k \ge 2$, it follows from Lemma 6.3 that

$$\Delta\Theta(k) = \sum_{i=1}^{k-1}\mathcal{A}(k)\cdots\mathcal{A}(i+1)\mathcal{B}(i)H_d^{\mathrm{T}}y(i) + \mathcal{B}(k)H_d^{\mathrm{T}}y(k). \tag{I2}$$

Substituting Equation (I2) into Equation (I1), and using Equation (E1) yields

$$\begin{aligned}
&\|\Theta(k_1) - \Theta(k_2)\| \\
&= \left\|\sum_{i=k_1+1}^{k_2}\left(\sum_{j=1}^{i-1}\mathcal{A}(i)\cdots\mathcal{A}(j+1)\mathcal{B}(j)H_d^{\mathrm{T}}y(j) + \mathcal{B}(i)H_d^{\mathrm{T}}y(i)\right)\right\| \\
&= \left\|\sum_{i=k_1+1}^{k_2}\sum_{j=1}^{i-1}\mathcal{A}(i)\cdots\mathcal{A}(j+1)\mathcal{B}(j)H_d^{\mathrm{T}}y(j) + \sum_{i=k_1+1}^{k_2}\mathcal{B}(i)H_d^{\mathrm{T}}y(i)\right\| \\
&\le \left\|\sum_{i=k_1+1}^{k_2}\sum_{j=1}^{i-1}\mathcal{A}(i)\cdots\mathcal{A}(j+1)\mathcal{B}(j)H_d^{\mathrm{T}}y(j)\right\| + \left\|\sum_{i=k_1+1}^{k_2}\mathcal{B}(i)H_d^{\mathrm{T}}y(i)\right\| \\
&\le \alpha\sum_{i=k_1+1}^{k_2}\sum_{j=1}^{i-1}\|\mathcal{A}(i)\cdots\mathcal{A}(j+1)\mathcal{B}(j)\|\,\gamma^j + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N,
\end{aligned} \tag{I3}$$

where the second term in Equation (I3) can be obtained by the same procedure that is used to obtain Equation (E1). Note that, since G_{yu} is tall, $H_2^{\mathrm{T}}H_2$ is positive definite. Therefore, it follows from Equation (49) that $\mathcal{B}$ is bounded and thus $\bar{\sigma}(\mathcal{B})$ is finite.

Assume that Φ is bounded so that $\bar{\sigma}(\mathcal{A}) < 1$. Defining $\tilde{\gamma} \triangleq \gamma/\bar{\sigma}(\mathcal{A})$ and applying Cauchy–Schwarz inequality to Equation (I3) yields

$$\begin{aligned}
&\|\Theta(k_1) - \Theta(k_2)\| \\
&\le \alpha\sum_{i=k_1}^{k_2}\sum_{j=1}^{i-1}\bar{\sigma}^{i-j}(\mathcal{A})\bar{\sigma}(\mathcal{B})\gamma^j + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N \\
&\le \alpha\sum_{i=N}^{\infty}\sum_{j=1}^{i-1}\bar{\sigma}^{i-j}(\mathcal{A})\bar{\sigma}(\mathcal{B})\gamma^j + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N \\
&= \alpha\bar{\sigma}(\mathcal{B})\sum_{i=N}^{\infty}\bar{\sigma}^i(\mathcal{A})\sum_{j=1}^{i-1}\tilde{\gamma}^j + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N.
\end{aligned} \tag{I4}$$

First, consider Equation (I4) with $\tilde{\gamma} = 1$. In this case, we have

$$\begin{aligned}
&\|\Theta(k_1) - \Theta(k_2)\| \\
&\le \alpha\bar{\sigma}(\mathcal{B})\sum_{i=N}^{\infty}(i-1)\bar{\sigma}^i(\mathcal{A}) + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N \\
&\le \alpha\bar{\sigma}(\mathcal{B})\sum_{i=N}^{\infty}i\bar{\sigma}^i(\mathcal{A}) + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N \\
&= \alpha\bar{\sigma}(\mathcal{B})\bar{\sigma}^N(\mathcal{A})\sum_{i=0}^{\infty}(i+N)\bar{\sigma}^i(\mathcal{A}) + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N \\
&= \alpha\bar{\sigma}(\mathcal{B})\bar{\sigma}^N(\mathcal{A})\left[\frac{\bar{\sigma}(\mathcal{A})}{(1-\bar{\sigma}(\mathcal{A}))^2} + \frac{N}{1-\bar{\sigma}(\mathcal{A})}\right] \\
&\quad + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N.
\end{aligned} \tag{I5}$$

Since $\bar{\sigma}(\mathcal{A}) \in (0,1)$ and $\gamma \in (0,1)$, it follows from Equation (I5) that, in the case $\tilde{\gamma} = 1$, for all $\varepsilon > 0$, there exists N such that, for all $k_1, k_2 > N$, $\|\Theta(k_1) - \Theta(k_2)\| < \varepsilon$. Therefore, in the case $\tilde{\gamma} = 1$, $\{\Theta(k)\}_{k=0}^{\infty}$ is Cauchy, and thus Θ converges. Next, consider Equation (I4) with $\tilde{\gamma} \ne 1$. In this case, we have

$$\begin{aligned}
&\|\Theta(k_1) - \Theta(k_2)\| \\
&\le \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\tilde{\gamma}}\sum_{i=N}^{\infty}\bar{\sigma}^i(\mathcal{A})(\tilde{\gamma} - \tilde{\gamma}^i) + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N \\
&= \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\tilde{\gamma}}\left(\sum_{i=N}^{\infty}\tilde{\gamma}\bar{\sigma}^i(\mathcal{A}) - \sum_{i=N}^{\infty}\gamma^i\right) + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N \\
&= \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\tilde{\gamma}}\left(\frac{\tilde{\gamma}}{1-\bar{\sigma}(\mathcal{A})}\bar{\sigma}^N(\mathcal{A}) - \frac{1}{1-\gamma}\gamma^N\right).
\end{aligned} \tag{I6}$$

Since $\bar{\sigma}(\mathcal{A}) < 1$ and $\gamma < 1$, it follows from Equation (I6) that, in the case $\tilde{\gamma} \ne 1$, for all $\varepsilon > 0$, there exists N such that $\|\Theta(k_1) - \Theta(k_2)\|_2 < \varepsilon$, and thus Θ converges. Thus, we have verified (i).

Now, assume that there exists $m \ge 2$ such that Equation (54) is satisfied. Define $\kappa \triangleq \bar{\sigma}(Q_m)$ and $\tilde{\kappa} \triangleq \kappa^{1/m} \in (0,1)$. For $N > m$, it follows from Equation (I3) that

$$\begin{aligned}
&\|\Theta(k_1) - \Theta(k_2)\| \\
&\le \alpha\sum_{i=k_1+1}^{k_2}\sum_{j=i-m+1}^{i-1}\|\mathcal{A}(i)\cdots\mathcal{A}(j+1)\mathcal{B}(j)\|\,\gamma^j \\
&\quad + \sum_{j=1}^{i-m}\|\mathcal{A}(i)\cdots\mathcal{A}(j+1)\mathcal{B}(j)\|\,\gamma^j + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N \\
&\le \alpha\sum_{i=N}^{\infty}\sum_{j=i-m+1}^{i-1}\|\mathcal{A}(i)\cdots\mathcal{A}(j+1)\mathcal{B}(j)\|\,\gamma^j \\
&\quad + \sum_{j=1}^{i-m}\|\mathcal{A}(i)\cdots\mathcal{A}(j+1)\mathcal{B}(j)\|\,\gamma^j + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N \\
&\le \alpha\sum_{i=N}^{\infty}\sum_{j=i-m+1}^{i-1}\bar{\sigma}(\mathcal{B})\gamma^j + \sum_{j=1}^{i-m}\kappa^{\frac{i-j}{m}-1}\bar{\sigma}(\mathcal{B})\gamma^j + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N \\
&= \alpha\bar{\sigma}(\mathcal{B})\sum_{i=N}^{\infty}\gamma^{i-m+1}\sum_{j=0}^{m-1}\gamma^j + \gamma\kappa^{-1}\tilde{\kappa}^{-1}\tilde{\kappa}^i
\end{aligned}$$

$$\cdot \sum_{j=0}^{i-m-1} \left(\frac{\gamma}{\tilde{\kappa}}\right)^j + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N$$

$$= \alpha\bar{\sigma}(\mathcal{B}) \sum_{i=N}^{\infty} \frac{1-\gamma^{m+1}}{1-\gamma}\gamma^{1-m}\gamma^i + \alpha\bar{\sigma}(\mathcal{B})\gamma\kappa^{-1}\tilde{\kappa}^{-1}$$

$$\cdot \sum_{i=N}^{\infty} \tilde{\kappa}^i \frac{1-\left(\frac{\gamma}{\tilde{\kappa}}\right)^{i-m}}{1-\frac{\gamma}{\tilde{\kappa}}} + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N. \tag{I7}$$

Define $c_1 \triangleq \frac{\gamma^{1-m}-\gamma^2}{(1-\gamma)^2}$, $c_2 \triangleq \frac{\alpha\bar{\sigma}(\mathcal{B})\gamma}{\kappa\tilde{\kappa}\left(1-\frac{\gamma}{\tilde{\kappa}}\right)}$. It follows from Equation (I7) that

$$\begin{aligned}
&\|\Theta(k_1) - \Theta(k_2)\| \\
&\le \alpha\bar{\sigma}(\mathcal{B})c_1\gamma^N + c_2\left(\sum_{i=N}^{\infty}\tilde{\kappa}^i - \sum_{i=N}^{\infty}\left(\frac{\tilde{\kappa}}{\gamma}\right)^m \tilde{\kappa}^i \left(\frac{\gamma}{\tilde{\kappa}}\right)^i\right) \\
&\quad + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N \\
&= \alpha\bar{\sigma}(\mathcal{B})c_1\gamma^N + \frac{c_2}{1-\tilde{\kappa}}\tilde{\kappa}^N - \frac{c_2\tilde{\kappa}^m}{\gamma^m(1-\gamma)}\gamma^N + \frac{\alpha\bar{\sigma}(\mathcal{B})}{1-\gamma}\gamma^N.
\end{aligned} \tag{I8}$$

Since $\gamma \in (0, 1)$ and $\tilde{\kappa} \in (0, 1)$, it follows from Equation (I8) that $\{\Theta(k)\}_{k=0}^{\infty}$ is Cauchy, and thus Θ converges. Thus, we have verified (ii).

Appendix 10: Proof of Proposition 7.5

To show (i), suppose $\zeta \in \operatorname{tzeros}(G_{yu}^{\mathrm{R}})$. Then, it follows from Equations (67) and (68) that rank $\Sigma^{\mathrm{R}}(\zeta) <$ normal rank $\Sigma^{\mathrm{R}} \le n + l_y$. Therefore, $\mathcal{N}(\Sigma^{\mathrm{R}}(\zeta)) \ne \{0\}$, and thus (i) is satisfied.

To show (ii), suppose that $x(0) = -\mathrm{Re}(x_0)$ and $u(k) = H_d^{\mathrm{T}}[\mathrm{Re}(\zeta^k)\mathrm{Re}(v_0) + \mathrm{Im}(\zeta^k)\mathrm{Im}(v_0)]$. Since $\zeta^0 = 1$, (74) holds for $k = 0$, and, from Equations (68) and (72), $y(0) = -C\mathrm{Re}(x_0) = 0$. Thus, Equations (74) and (75) hold for $k = 0$. Now, assume that Equations (74) and (75) hold for some $k > 0$. We thus have

$$\begin{aligned}
x(k+1) &= Ax(k) + Bu(k) \\
&= -\mathrm{Re}(\zeta^k)A\mathrm{Re}(x_0) + \mathrm{Im}(\zeta^k)A\mathrm{Im}(x_0) \\
&\quad + \mathrm{Re}(\zeta^k)BH_d^{\mathrm{T}}\mathrm{Re}(v_0) - \mathrm{Im}(\zeta^k)BH_d^{\mathrm{T}}\mathrm{Im}(v_0).
\end{aligned} \tag{J1}$$

Next, it follows from Equations (68) and (72) that

$$\zeta x_0 - Ax_0 + BH_d^{\mathrm{T}}v_0 = 0,$$

and thus,

$$BH_d^{\mathrm{T}}\mathrm{Re}(v_0) = A\mathrm{Re}(x_0) + \mathrm{Im}(\zeta)\mathrm{Im}(x_0) - \mathrm{Re}(\zeta)\mathrm{Re}(x_0) \tag{J2}$$

and

$$BH_d^{\mathrm{T}}\mathrm{Im}(v_0) = A\mathrm{Im}(x_0) - \mathrm{Re}(\zeta)\mathrm{Im}(x_0) - \mathrm{Im}(\zeta)\mathrm{Re}(x_0). \tag{J3}$$

Substituting Equations (J2) and (J3) into Equation (J1), we obtain

$$\begin{aligned}
x(k+1) &= -\mathrm{Re}(\zeta^k)A\mathrm{Re}(x_0) + \mathrm{Im}(\zeta^k)A\mathrm{Im}(x_0) \\
&\quad + \mathrm{Re}(\zeta^k)[A\mathrm{Re}(x_0) + \mathrm{Im}(\zeta)\mathrm{Im}(x_0) - \mathrm{Re}(\zeta)\mathrm{Re}(x_0)] \\
&\quad - \mathrm{Im}(\zeta^k)[A\mathrm{Im}(x_0) - \mathrm{Re}(\zeta)\mathrm{Im}(x_0) - \mathrm{Im}(\zeta)\mathrm{Re}(x_0)] \\
&= [-\mathrm{Re}(\zeta^k)\mathrm{Re}(\zeta) + \mathrm{Im}(\zeta^k)\mathrm{Im}(\zeta)]\mathrm{Re}(x_0) \\
&\quad + [\mathrm{Re}(\zeta^k)\mathrm{Im}(\zeta) + \mathrm{Im}(\zeta^k)\mathrm{Re}(\zeta)]\mathrm{Im}(x_0) \\
&= -\mathrm{Re}(\zeta^{k+1})\mathrm{Re}(x_0) + \mathrm{Im}(\zeta^{k+1})\mathrm{Im}(x_0),
\end{aligned} \tag{J4}$$

which shows that Equation (74) holds for $k+1$. Furthermore, since $Cx_0 = 0$ from Equation (72), it follows from Equation (J4) that

$$y(k+1) = Cx(k+1) = 0.$$

Thus, Equations (74) and (75) hold for $k+1$ if they hold for k. By induction, it follows that (ii) holds. Statement (iii) follows from the homogeneity of linear systems.

Finally, to show (iv), consider

$$y(k) = CA^k x(0) + \sum_{i=1}^{k} H_i u(k-i), \tag{J5}$$

where $u(k)$ is given by Equation (76), and $x(0) \in \mathbb{R}^n$. Adding and subtracting $CA^k(-\alpha\mathrm{Re}(x_0))$ from Equation (J5) and using (iii), we have

$$\begin{aligned}
y(k) &= CA^k[x(0) + \alpha\mathrm{Re}(x_0)] + CA^k[-\alpha\mathrm{Re}(x_0)] \\
&\quad + \sum_{i=1}^{k} H_i u(k-i) \\
&= CA^k\tilde{x}(k),
\end{aligned} \tag{J6}$$

where $\tilde{x}(k) \triangleq x(0) + \alpha\mathrm{Re}(x_0) \in \mathbb{R}^n$. Since A is discrete-time asymptotically stable, it follows from Equation (J6) that $y(k) \to 0$ as $k \to \infty$ with exponential convergence.