International Journal of Control

ISSN: 0020-7179 (Print) 1366-5820 (Online) Journal homepage: http://www.tandfonline.com/loi/tcon20

A numerical investigation of direct and indirect closed-loop architectures for estimating nonminimum-phase zeros

Frantisek M. Sobolic, Khaled F. Aljanaideh & Dennis S. Bernstein

To cite this article: Frantisek M. Sobolic, Khaled F. Aljanaideh & Dennis S. Bernstein (2018): A numerical investigation of direct and indirect closed-loop architectures for estimating nonminimum-phase zeros, International Journal of Control, DOI: 10.1080/00207179.2018.1501609

To link to this article: https://doi.org/10.1080/00207179.2018.1501609

Accepted author version posted online: 18 Jul 2018.
Published online: 05 Aug 2018.

INTERNATIONAL JOURNAL OF CONTROL
https://doi.org/10.1080/00207179.2018.1501609

Taylor & Francis
Taylor & Francis Group

A numerical investigation of direct and indirect closed-loop architectures for estimating nonminimum-phase zeros

Frantisek M. Sobolic[a], Khaled F. Aljanaideh[b] and Dennis S. Bernstein[a]

[a]Aerospace Engineering, University of Michigan, Ann Arbor, MI, USA; [b]Faculty of Engineering, Aeronautical Engineering, Jordan University of Science and Technology, Irbid, Jordan

ABSTRACT

This paper presents a numerical investigation of three direct architectures and three indirect architectures for identifying a plant operating in closed loop. Motivated by adaptive control of systems with nonminimum-phase (NMP) zeros and taking advantage of the fact that zeros are not moved by feedback, the performance metric is the accuracy of the estimates of the NMP zeros of the plant. Assuming known plant order, single-input, single-output, infinite-impulse-response models are constructed in the presence of process and sensor noise. Least squares provides the baseline estimation technique, and prediction error methods are used to account for correlation between the model input and noise. The goal is to compare the accuracy of the NMP-zero estimates obtained from each method and for each architecture.

ARTICLE HISTORY
Received 28 October 2017
Accepted 12 July 2018

KEYWORDS
System identification; closed-loop identification; auxiliary input; prediction error method; least squares; consistency

1. Introduction

During closed-loop operation, it is often necessary to monitor the plant in order to detect changes that can degrade stability and performance. At the same time, more-accurate modelling of the plant can facilitate the ability to enhance closed-loop performance through on-line controller modification (Landau, 1999; Langer & Landau, 1996; von Wangenheim, 2012; Zhu & Butoyi, 2002). Accordingly, closed-loop identification has been extensively studied (Forssell & Ljung, 1999; Gustavsson, Ljung, & Soderstrom, 1977; Hjalmarsson, 2005; Hjalmarsson, Gevers, & De Bruyne, 1996; Van den Hof, 1998; Ljung, 2010; Shen, Ding, Alsaedi, & Hayat, 2017), and various applications are discussed in Epperlein, Bamieh, and Åström (2015), Engelhart, Boonstra, Aarts, Schouten and van der Kooij (2016), de Vlugt, Schouten, and van der Helm (2003), and Smith (1998).

One of the benefits of feedback control is the ability to move plant poles. In contrast, zeros are not moved by linear time-invariant (LTI) feedback. In fact, nonminimum-phase (NMP) zeros are one of the most challenging aspects of feedback control in terms of limiting achievable performance (Freudenberg, Hollot, Middleton, & Toochinda, 2003). As discussed in Hoagg and Bernstein (2007), knowledge of the locations of the NMP zeros can impact the performance of adaptive control due to the fact that the adaptive controller may attempt to cancel an erroneously modelled NMP zero.

From the point of view of system identification, the fact that zero locations are unchanged by LTI feedback can be viewed as advantageous. In particular, both open-loop identification (Martensson & Hjalmarsson, 2005b,2009; Martensson, Jansson, & Hjalmarsson, 2005; Rojas, Hjalmarsson, Gerencser, & Martensson, 2011) and closed-loop identification (Martensson & Hjalmarsson, 2005a) can be used to estimate plant zeros. The present paper takes advantage of this fact by using closed-loop identification methods to estimated NMP zeros. In particular, the goal of this paper is to compare the accuracy of various architectures for estimating the NMP zeros of a plant operating in closed loop. After constructing a model of the open- or closed-loop plant, the metric used to assess the performance of the identification architecture is the accuracy of the estimate of the NMP zeros.

The starting point for the present paper is the survey paper (Forssell & Ljung, 1999), which analyses various architectures for closed-loop identification. That work emphasises the practical importance of the problem and demonstrates the richness of the subject in terms of the diverse architectures that can be employed. The presentation in Forssell and Ljung (1999), however, does not include a numerical investigation of the relative merits of candidate architectures and identification algorithms. Consequently, as a complement to Forssell and Ljung (1999), the contribution of the present paper is a detailed numerical study that compares the accuracy of several closed-loop identification architectures in estimating NMP zeros.

In contrast to open-loop identification, closed-loop identification presents unique challenges in system identification (Van den Hof, 1998). For example, the model input may lack sufficient persistency due to the limited spectral content of the command and disturbances. In addition, because of closed-loop operation, the model input may be correlated with the exogenous noise, potentially leading to parameter bias.

Unlike open-loop identification, closed-loop identification can be performed with a variety of architectures. The basic architecture for closed-loop identification is *direct closed-loop (DCL) identification*, where the control signal and measurement, that is, the plant input and output, are used to construct

CONTACT Khaled F. Aljanaideh khaledfj@umich.edu

a plant model. The control signal, which serves as the model input, arises in response to disturbance and sensor noise, and thus is correlated with these noise signals, potentially leading to bias in the parameter estimates.

A variation of DCL identification is to add an auxiliary signal to the control signal and then use the sum of these signals as the model input. This approach, called *auxiliary direct closed-loop (ADCL) identification*, uses the auxiliary signal to enhance identification accuracy at the cost of disrupting closed-loop performance. The challenge of designing minimally disruptive signals for closed-loop identification is discussed in Rivera, Lee, Mittelmann, and Braun (2007).

An alternative approach to DCL identification is to estimate a closed-loop transfer function rather than the plant itself. This approach requires an external signal. The simplest approach of this type, called *indirect closed-loop (ICL) identification*, uses the command signal as the model input. To facilitate the method, the command signal can be chosen to be sufficiently persistent albeit at the cost of degrading closed-loop operation. A model of the plant is subsequently extracted from the estimated closed-loop model.

A variation of ICL identification is to add an auxiliary signal to the control signal and then use only the auxiliary signal as the model input. This approach, called *auxiliary indirect closed-loop (AICL) identification*, identifies the overall closed-loop system. As in the case of ICL identification, a model of the plant is subsequently extracted from the estimated closed-loop model. In order to simplify the identification, the auxiliary signal is chosen to be a multiple of the command signal.

In addition to DCL, ADCL, ICL, and AICL identification, we consider a nonstandard way in which the external signal used for ADCL identification and AICL identification are injected into the loop. In the standard approach, the auxiliary signal is added to the control signal. In contrast, in *intercalated injection*, the auxiliary signal is added to an internal signal in the feedback controller. This architecture arises in retrospective cost adaptive control (Rahman, Xie, & Bernstein, 2017), where the effect of the controller update at each step is equivalent to intercalated injection of a control-input perturbation. For closed-loop identification, the *intercalated auxiliary indirect closed-loop* and *AICL* architectures are characterised by the fact that the resulting transfer functions have restricted numerators. This paper thus considers six architectures for identifying a plant operating in closed loop, namely DCL identification, standard auxiliary direct closed-loop (ADCL/S) identification, intercalated auxiliary direct closed-loop (ADCL/I) identification, ICL identification, standard auxiliary indirect closed-loop (AICL/S) identification, and intercalated auxiliary indirect closed-loop (AICL/I) identification.

In order to estimate NMP zeros, we apply identification techniques with infinite impulse response (IIR) models. In order to compare architectures for closed-loop identification, we make the simplifying assumption that the plant order is known. In the case where the plant order is unknown, additional techniques for order estimation are needed. This can be done, for example, by overestimating the plant order, impulsing the estimated model, and then applying Ho-Kalman realisation theory (Ho & Kalman, 1966; Juang & Pappa, 1985) along with heuristic nuclear-norm minimisation techniques (Recht, Fazel, & Parrilo, 2010; Smith, 2014) to estimate the model order and construct a reduced-order model. Alternatively, this can be done by using a finite impulse response (FIR) model structure to directly estimate the Markov parameters for use in the Hankel matrix. This is the approach used in Aljanaideh and Bernstein (2017), where noncausal FIR models are used to account for the possibility that the plant may be open-loop unstable.

We consider two techniques for identification, namely, least squares (LS) and prediction error methods (PEM). LS provides a baseline technique for estimation. In most cases, LS estimates are biased (Ljung & Wahlberg, 1992) and, assuming that the model input is white, Theorem 1 provides an analytical expression for the bias. A related expression is given in Section 6.5.2 of Eykhoff (1974). Setting the bias to zero yields a necessary and sufficient condition under which the model estimate is consistent. In cases where Theorem 1 is applicable and the bias is nonzero, the expression for the LS estimate is used in this paper to provide a check on the correctness of the NMP-zero-error estimate based on the numerical data.

A more effective approach to obtaining consistent estimates within the context of closed-loop identification is PEM. PEM is described in Ljung (1976, 2002) and applied to DCL and ICL identification architectures in Van den Hof (1998), Forssell and Ljung (1999), Van den Hof and Schrama (1993), and Ljung and Forssell (1999). PEM can be viewed as an extension of LS that fits the dynamics and noise model parameters over a data window and a prediction horizon. The minimisation of the prediction error gives rise to a nonconvex optimisation problem in terms of the model parameters (Ljung, 1999; Söderström & Stoica, 1988). The consistency of PEM for closed-loop identification is analysed in Forssell and Ljung (1999); Ljung (2002), where it is shown that, if the command is persistently exciting, the output noise has zero-mean and is uncorrelated with the command, and the output noise model is asymptotically stable and minimum phase, then the estimates of the model parameters are consistent.

The contents of the paper are as follows. Section 3 describes the estimation algorithms, including LS and prediction error methods. Sections 4 and 5 describe the DCL and ICL identification architectures, respectively, with and without an auxiliary signal, including standard and intercalated injection. Section 6 presents a numerical comparison of the DCL identification architectures. Two examples are considered. In the first example, the plant is open-loop stable; in the second example, the plant is open-loop unstable. Section 7 presents similar results for the case of ICL identification architectures. Conclusions and directions for future research are given in Section 8.

Notation: I_n denotes the $n \times n$ identity matrix, $0_{n\times m}$ denotes the $n \times m$ zero matrix, and $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm.

2. Model structure

Consider the discrete-time n_0th-order single-input, single-output transfer function G_0 defined by

$$G_0(\mathbf{q}) \triangleq \frac{N_0(\mathbf{q})}{D_0(\mathbf{q})}, \tag{1}$$

where $\mathbf{q}$ is the forward shift operator and

$$N_0(\mathbf{q}) = \sum_{i=0}^{n_0} b_i \mathbf{q}^{n_0-i}, \quad D_0(\mathbf{q}) = \mathbf{q}^{n_0} + \sum_{i=1}^{n_0} a_i \mathbf{q}^{n_0-i}. \tag{2}$$

Letting u_0 and y_0 denote the input and output, respectively, of G_0, that is,

$$y_0(k) = G_0(\mathbf{q})u_0(k), \tag{3}$$

it follows that, for all $k \geq 0$,

$$\begin{aligned} y_0(k) = &-a_1 y_0(k-1) - \cdots - a_{n_0} y_0(k-n_0) + b_0 u_0(k) \\ &+ \cdots + b_{n_0} u_0(k-n_0), \end{aligned} \tag{4}$$

which can be written as

$$y_0(k) = \phi_{y_0,u_0}(k)\theta, \tag{5}$$

where

$$\begin{aligned} \phi_{y_0,u_0}(k) \triangleq [&-y_0(k-1) \ \cdots \ -y_0(k-n_0) \ u_0(k) \\ &\cdots \ u_0(k-n_0)], \end{aligned} \tag{6}$$

$$\theta \triangleq [a_1 \ \cdots \ a_{n_0} \ b_0 \ \cdots \ b_{n_0}]^{\mathrm{T}}. \tag{7}$$

The measured output y of G_0 is defined as

$$y(k) \triangleq y_0(k) + w_0(k) = G_0(\mathbf{q})u_0(k) + w_0(k), \tag{8}$$

where w_0 is a noise signal. Since w_0 is added to y_0 in (8), it appears to represent sensor noise. However, w_0 can represent either sensor noise or process noise or both by viewing w_0 as the output of a system G_w, that is,

$$w_0(k) = G_w(\mathbf{q})w(k) = \frac{N_w(\mathbf{q})}{D_w(\mathbf{q})}w(k), \tag{9}$$

where G_w is a proper transfer function of order $\bar{n}$ and w is zero-mean white noise with standard deviation σ_w. In the case of sensor noise, G_w represents a colouring filter. In the case of process noise, G_w represents the effect of w depending on how w enters the plant. In the special case where w and u_0 are colocated inputs, G_w is a copy of G_0, where the initial conditions of G_0 and G_w are assigned so that (8) provides the correct forced response. Note that the free response of G_w and G_0 is accounted for by the use of the forward shift operator $\mathbf{q}$ rather than the complex Z-transform variable $\mathbf{z}$ (Aljanaideh & Bernstein, 2018). It follows that, for all $k \geq 0$,

$$\begin{aligned} w_0(k) = &-\bar{a}_1 w_0(k-1) - \cdots - \bar{a}_{\bar{n}} w_0(k-\bar{n}) + \bar{b}_0 w(k) \\ &+ \cdots + \bar{b}_{\bar{n}} w(k-\bar{n}), \end{aligned} \tag{10}$$

which can be written as

$$w_0(k) = \phi_{w_0,w}(k)\theta_w, \tag{11}$$

where

$$\begin{aligned} \phi_{w_0,w}(k) \triangleq [&-w_0(k-1) \ \cdots \ -w_0(k-\bar{n}) \ w(k) \cdots w(k-\bar{n})], \end{aligned} \tag{12}$$

$$\theta_w \triangleq [\bar{a}_1 \ \cdots \ \bar{a}_{\bar{n}} \ \bar{b}_0 \ \cdots \ \bar{b}_{\bar{n}}]^{\mathrm{T}}. \tag{13}$$

3. Identification algorithms

In this section, we briefly review LS and PEM for estimating the coefficients of G_0. For LS, Theorem 1 provides an explicit expression for the bias of the estimates of the coefficients of a transfer function whose model order is known. For simplicity, we assume throughout the paper that the order n_0 of G_0 is known. For the DCL identification algorithms, n_0 represents the plant order, whereas, for the ICL identification algorithms, n_0 represents the order of the closed-loop transfer function. We also assume that all random processes in the paper are ergodic so that ensemble averages are equal, with probability 1, to time averages of realisations.

3.1 LS estimation

In order to estimate G_0, we use the model structure

$$G_{\mathrm{m}}(\mathbf{q}) \triangleq \frac{N_{\mathrm{m}}(\mathbf{q})}{D_{\mathrm{m}}(\mathbf{q})}, \tag{14}$$

where $N_{\mathrm{m}}(\mathbf{q}) \triangleq \sum_{i=0}^{n_0} \hat{b}_i \mathbf{q}^{n_0-i}$ and $D_{\mathrm{m}}(\mathbf{q}) \triangleq \mathbf{q}^{n_0} + \sum_{i=1}^{n_0} \hat{a}_i \mathbf{q}^{n_0-i}$. Define

$$\begin{aligned} \phi_y(k) &\triangleq [-y(k-1) \ \cdots \ -y(k-n_0)], \\ \phi_{u_0}(k) &\triangleq [u_0(k) \ \cdots \ u_0(k-n_0)], \end{aligned} \tag{15}$$

and $\tilde{n} \triangleq \max\{n_0, \bar{n}\}$. Let $\ell \geq \tilde{n}$ be the number of samples of y and u_0, and define

$$\begin{aligned} \Phi_{y,\ell} &\triangleq \begin{bmatrix} \phi_y(\tilde{n}) \\ \vdots \\ \phi_y(\ell) \end{bmatrix} \in \mathbb{R}^{(\ell-\tilde{n}+1)\times n_0}, \\ \Phi_{u_0,\ell} &\triangleq \begin{bmatrix} \phi_{u_0}(\tilde{n}) \\ \vdots \\ \phi_{u_0}(\ell) \end{bmatrix} \in \mathbb{R}^{(\ell-\tilde{n}+1)\times (n_0+1)}, \end{aligned} \tag{16}$$

$$\begin{aligned} \Phi_{y,u_0,\ell} &\triangleq \begin{bmatrix} \Phi_{y,\ell} & \Phi_{u_0,\ell} \end{bmatrix} \in \mathbb{R}^{(\ell-\tilde{n}+1)\times (2n_0+1)}, \\ \Psi_{y,\ell} &\triangleq \begin{bmatrix} y(\tilde{n}) \\ \vdots \\ y(\ell) \end{bmatrix} \in \mathbb{R}^{\ell-\tilde{n}+1}. \end{aligned} \tag{17}$$

Then the LS estimate $\hat{\theta}_\ell^{\mathrm{LS}}$ of θ, which is defined by

$$\hat{\theta}_\ell^{\mathrm{LS}} \triangleq \arg\min_{\bar{\theta}\in\mathbb{R}^{2n_0+1}} \|\Phi_{y,u_0,\ell}\bar{\theta} - \Psi_{y,\ell}\|_{\mathrm{F}}, \tag{18}$$

is given by

$$\hat{\theta}_\ell^{\mathrm{LS}} \triangleq (\Phi_{y,u_0,\ell}^{\mathrm{T}}\Phi_{y,u_0,\ell})^{-1}\Phi_{y,u_0,\ell}^{\mathrm{T}}\Psi_{y,\ell}. \tag{19}$$

Next, define the $(2n_0+1)\times(2n_0+1)$ positive-semidefinite covariance matrices

$$\Gamma \triangleq \mathbb{E}\begin{bmatrix} \mathcal{Y}^2(k) & \cdots & \mathcal{Y}(k)\mathcal{Y}(k+n_0-1) \\ \vdots & \ddots & \vdots \\ \mathcal{Y}(k)\mathcal{Y}(k+n_0-1) & \cdots & \mathcal{Y}^2(k) \\ -\mathcal{Y}(k)\mathcal{U}_0(k+1) & \cdots & -\mathcal{Y}(k)\mathcal{U}_0(k+n_0) \\ \vdots & \ddots & \vdots \\ -\mathcal{Y}(k)\mathcal{U}_0(k-n_0+1) & \cdots & -\mathcal{Y}(k)\mathcal{U}_0(k) \end{bmatrix}$$
$$\begin{bmatrix} -\mathcal{Y}(k)\mathcal{U}_0(k+1) & \cdots & -\mathcal{Y}(k)\mathcal{U}_0(k-n_0+1) \\ \vdots & \ddots & \vdots \\ -\mathcal{Y}(k)\mathcal{U}_0(k+n_0) & \cdots & -\mathcal{Y}(k)\mathcal{U}_0(k) \\ \mathcal{U}_0^2(k) & \cdots & \mathcal{U}_0(k)\mathcal{U}_0(k+n_0) \\ \vdots & \ddots & \vdots \\ \mathcal{U}_0(k)\mathcal{U}_0(k+n_0) & \cdots & \mathcal{U}_0^2(k) \end{bmatrix}, \quad (20)$$

$$\Gamma_0 \triangleq \mathbb{E}\begin{bmatrix} \mathcal{Y}_0^2(k) & \cdots & \mathcal{Y}_0(k)\mathcal{Y}_0(k+n_0-1) \\ \vdots & \ddots & \vdots \\ \mathcal{Y}_0(k)\mathcal{Y}_0(k+n_0-1) & \cdots & \mathcal{Y}_0^2(k) \\ -\mathcal{Y}_0(k)\mathcal{U}_0(k+1) & \cdots & -\mathcal{Y}_0(k)\mathcal{U}_0(k+n_0) \\ \vdots & \ddots & \vdots \\ -\mathcal{Y}_0(k)\mathcal{U}_0(k-n_0+1) & \cdots & -\mathcal{Y}_0(k)\mathcal{U}_0(k) \end{bmatrix}$$
$$\begin{bmatrix} -\mathcal{Y}_0(k)\mathcal{U}_0(k+1) & \cdots & -\mathcal{Y}_0(k)\mathcal{U}_0(k-n_0+1) \\ \vdots & \ddots & \vdots \\ -\mathcal{Y}_0(k)\mathcal{U}_0(k+n_0) & \cdots & -\mathcal{Y}_0(k)\mathcal{U}_0(k) \\ \mathcal{U}_0^2(k) & \cdots & \mathcal{U}_0(k)\mathcal{U}_0(k+n_0) \\ \vdots & \ddots & \vdots \\ \mathcal{U}_0(k)\mathcal{U}_0(k+n_0) & \cdots & \mathcal{U}_0^2(k) \end{bmatrix}. \quad (21)$$

It thus follows that

$$\Gamma \overset{\text{wp1}}{=} \lim_{\ell\to\infty}\frac{1}{\ell}\Phi_{y,u_0,\ell}^{\mathrm{T}}\Phi_{y,u_0,\ell}, \quad \Gamma_0 \overset{\text{wp1}}{=} \lim_{\ell\to\infty}\frac{1}{\ell}\Phi_{y_0,u_0,\ell}^{\mathrm{T}}\Phi_{y_0,u_0,\ell}. \quad (22)$$

Define the $n_0\times\bar{n}$ matrix $\mathcal{H}_{G_w}$, where, if $\bar{n}\le n_0$, then

$$\mathcal{H}_{G_w} \triangleq \begin{bmatrix} \sum_{i=0}^{\infty}H_{w,i}^2 & \cdots & \sum_{i=0}^{\infty}H_{w,i}H_{w,\bar{n}-1+i} \\ \vdots & \ddots & \vdots \\ \sum_{i=0}^{\infty}H_{w,i}H_{w,\bar{n}-1+i} & \cdots & \sum_{i=0}^{\infty}H_{w,i}^2 \\ \vdots & \cdots & \vdots \\ \sum_{i=0}^{\infty}H_{w,i}H_{w,n_0-1+i} & \cdots & \sum_{i=0}^{\infty}H_{w,i}H_{w,n_0-\bar{n}+i} \end{bmatrix} \quad (23)$$

and, if $\bar{n}\ge n_0$, then

$$\mathcal{H}_{G_w} \triangleq \begin{bmatrix} \sum_{i=0}^{\infty}H_{w,i}^2 & \cdots & \sum_{i=0}^{\infty}H_{w,i}H_{w,n_0-1+i} \\ \vdots & \ddots & \vdots \\ \sum_{i=0}^{\infty}H_{w,i}H_{w,n_0-1+i} & \cdots & \sum_{i=0}^{\infty}H_{w,i}^2 \end{bmatrix}$$
$$\begin{bmatrix} \cdots & \sum_{i=0}^{\infty}H_{w,i}H_{w,\bar{n}-1+i} \\ \cdots & \vdots \\ \cdots & \sum_{i=0}^{\infty}H_{w,i}H_{w,\bar{n}-n_0+i} \end{bmatrix}, \quad (24)$$

where $H_{w,i}$ is the ith Markov parameter of G_w. Note that, if $\bar{n}=n_0$, then (23) and (24) coincide and are given by the $n_0\times n_0$ matrix

$$\tilde{\mathcal{H}}_{G_w} \triangleq \begin{bmatrix} \sum_{i=0}^{\infty}H_{w,i}^2 & \cdots & \sum_{i=0}^{\infty}H_{w,i}H_{w,n_0-1+i} \\ \vdots & \ddots & \vdots \\ \sum_{i=0}^{\infty}H_{w,i}H_{w,n_0-1+i} & \cdots & \sum_{i=0}^{\infty}H_{w,i}^2 \end{bmatrix}. \quad (25)$$

Lemma 1: *Let u_0 and w be realisations of zero-mean, uncorrelated stationary white random processes $\mathcal{U}_0$ and $\mathcal{W}$, respectively, with finite second moments, and let w_0 be given by* (9)*. Then*

$$\Gamma = \Gamma_0 + \begin{bmatrix} \sigma_w^2\tilde{\mathcal{H}}_{G_w} & 0_{n_0\times(n_0+1)} \\ 0_{(n_0+1)\times n_0} & 0_{(n_0+1)\times(n_0+1)} \end{bmatrix}. \quad (26)$$

Proof: See Appendix 1. ■

Note that, if $\ell > 3\tilde{n} = 3\max\{n_0,\bar{n}\}$, then $\Phi_{y,u_0,\ell}$ is a tall matrix. The following result is a consequence of the fact that $\ell\mapsto\sigma_{\min}(\Phi_{y,u_0,\ell}^{\mathrm{T}}\Phi_{y,u_0,\ell})$ is nondecreasing.

Lemma 2: *Let $\tilde{n} \geq 1$. Then Γ is positive definite if and only if, with probability 1, there exists $\ell \geq 3\tilde{n}$ such that* rank $\Phi_{y,u_0,\ell} = 2n_0 + 1$.

Next, define the $n_0 \times (\bar{n}+1)$ matrix $\mathcal{H}_w$, where, if $\bar{n} < n_0$, then

$$\mathcal{H}_w \triangleq \begin{bmatrix} 0 & H_{w,0} & \cdots & H_{w,\bar{n}-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & H_{w,0} \\ 0_{(n_0-\bar{n})\times 1} & \cdots & \cdots & 0_{(n_0-\bar{n})\times 1} \end{bmatrix} \in \mathbb{R}^{n_0\times(\bar{n}+1)}, \tag{27}$$

if $\bar{n} = n_0$, then

$$\mathcal{H}_w \triangleq \begin{bmatrix} 0 & H_{w,0} & \cdots & H_{w,n_0-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & H_{w,0} \end{bmatrix} \in \mathbb{R}^{n_0\times(n_0+1)}, \tag{28}$$

and, if $\bar{n} > n_0$, then

$$\mathcal{H}_w \triangleq \begin{bmatrix} 0 & H_{w,0} & \cdots & H_{w,\bar{n}-n_0} & \cdots & H_{w,\bar{n}-1} \\ \vdots & \ddots & \ddots & \vdots & \ddots & \ddots \\ 0 & \cdots & 0 & H_{w,0} & \cdots & H_{w,\bar{n}-n_0} \end{bmatrix} \in \mathbb{R}^{n_0\times(\bar{n}+1)}. \tag{29}$$

Also, define

$$N_w \triangleq \begin{bmatrix} \bar{b}_0 \\ \vdots \\ \bar{b}_{\bar{n}} \end{bmatrix}, \quad D_w \triangleq \begin{bmatrix} \bar{a}_1 \\ \vdots \\ \bar{a}_{\bar{n}} \end{bmatrix}, \quad \theta_w \triangleq \begin{bmatrix} D_w \\ N_w \end{bmatrix} \in \mathbb{R}^{2\bar{n}+1}, \tag{30}$$

where $N_w(\mathbf{q}) = \sum_{i=0}^{\bar{n}} \bar{b}_i \mathbf{q}^{\bar{n}-i}$, and $D_w(\mathbf{q}) = \mathbf{q}^{\bar{n}} + \sum_{i=1}^{\bar{n}} \bar{a}_i \mathbf{q}^{\bar{n}-i}$.

The following result gives an explicit expression for the bias in LS estimation of the coefficients of (1) in the case where the input is white noise and the measurement noise is coloured and uncorrelated with the input. This result is an equivalent version of a result given in Section 6.5.2 of Eykhoff (1974).

Theorem 1: *Consider the problem of estimating the transfer function* (1) *with noisy measurements* (8)*, where the measurement noise is given by* (9)*. Assume that u_0 and w are realisations of zero-mean, uncorrelated stationary white random processes $\mathcal{U}_0$ and $\mathcal{W}$, respectively, with finite second moments. Then*

$$\Gamma \lim_{\ell\to\infty} \hat{\theta}_\ell^{\mathrm{LS}} \overset{\mathrm{wp1}}{=} \Gamma_0\theta + \begin{bmatrix} \sigma_w^2\mathcal{H}_{G_w}D_w - \sigma_w^2\mathcal{H}_wN_w \\ 0_{(n_0+1)\times 1} \end{bmatrix}. \tag{31}$$

Now assume that Γ is positive definite. Then

$$\lim_{\ell\to\infty} \hat{\theta}_\ell^{\mathrm{LS}} - \theta \overset{\mathrm{wp1}}{=} (\Gamma^{-1}\Gamma_0 - I_{2n_0+1})\theta + \Gamma^{-1}\begin{bmatrix} \sigma_w^2\mathcal{H}_{G_w}D_w - \sigma_w^2\mathcal{H}_wN_w \\ 0_{(n_0+1)\times 1} \end{bmatrix}. \tag{32}$$

Consequently, the following statements are equivalent:

(i) $\tilde{\mathcal{H}}_{G_w}D_0 = \mathcal{H}_{G_w}D_w - \mathcal{H}_wN_w$.

(ii) $\lim_{\ell\to\infty} \hat{\theta}_\ell^{\mathrm{LS}} \overset{\mathrm{wp1}}{=} \theta$.

Proof: See Appendix 2. ■

In the case where Γ is positive definite, Theorem 1 shows that the asymptotic bias $b \triangleq \lim_{\ell\to\infty} \hat{\theta}_\ell^{\mathrm{LS}} - \theta$ of the LS estimate $\hat{\theta}_\ell^{\mathrm{LS}}$ is also given by (32). Setting $\Gamma b = 0$ yields

$$\mathcal{H}_{G_w}D_w - \mathcal{H}_wN_w = \frac{1}{\sigma_w^2}[I_{n_0}\ 0_{n_0\times(n_0+1)}](\Gamma - \Gamma_0)\theta = \tilde{\mathcal{H}}_{G_w}D_0, \tag{33}$$

which is a necessary and sufficient condition under which the LS estimate $\hat{\theta}_\ell^{\mathrm{LS}}$ is consistent. As a special case, assume that $\bar{n} = n_0$, $N_w = 1$, and $D_w = D_0$. Then $\tilde{\mathcal{H}}_{G_w} = \mathcal{H}_{G_w}$ and, since the relative degree of G_w is n_0, it follows that $\mathcal{H}_w = 0_{n_0\times(n_0+1)}$. Therefore, (i) is satisfied, and thus $b = 0$. This case is considered in Söderström and Stoica (1988, p. 186) and Ljung (1999, p. 205).

Theorem 1 assumes that the random processes $\mathcal{U}_0$ and $\mathcal{W}$ are uncorrelated. As shown below, for the DCL identification architectures, u_0 represents the control signal, whereas, for the ICL identification architectures, u_0 represents the command signal. Consequently, these random processes are correlated for the DCL identification architectures considered in Section 4, but they are uncorrelated for the ICL architectures considered in Section 5.

3.2 Prediction error methods

For PEM, y is written as

$$y(k) = G_0(\mathbf{q})u_0(k) + w_0(k), \tag{34}$$

where

$$w_0(k) = G_w(\mathbf{q})w(k) \tag{35}$$

and G_w represents the noise dynamics, the order of G_w is $\bar{n}$, and w is zero-mean white noise. The one-step-ahead predictor of (34) is defined by Ljung (1999)

$$\hat{y}(k \mid \hat{\theta}_\ell, \hat{\theta}_{w,\bar{n},\ell}) \triangleq (1 - G_w(\mathbf{q}, \hat{\theta}_{w,\bar{n},\ell})^{-1})u_0(k) + G_w(\mathbf{q}, \hat{\theta}_{w,\bar{n},\ell})^{-1}\hat{G}_0(\mathbf{q}, \hat{\theta}_\ell)y(k), \tag{36}$$

where $\hat{G}_0(\mathbf{q}, \hat{\theta}_\ell)$ and $G_w(\mathbf{q}, \hat{\theta}_{w,\bar{n},\ell})$ are models of $G_0(\mathbf{q})$ and $G_w(\mathbf{q})$, respectively, and $G_w(\mathbf{q})$ is minimum phase. Note that $\hat{G}_0(\mathbf{q}, \theta) = G_0(\mathbf{q})$.

Using the one-step-ahead predictor model (36), the prediction error is defined as

$$\varepsilon(k \mid \hat{\theta}_\ell, \hat{\theta}_{w,\bar{n},\ell}) \triangleq y(k) - \hat{y}(k \mid \hat{\theta}_\ell, \hat{\theta}_{w,\bar{n},\ell}). \tag{37}$$

The prediction error estimate $\hat{\theta}_\ell^{\mathrm{PEM}}$ of θ is thus given by

$$\hat{\theta}_\ell^{\mathrm{PEM}} = \underset{\bar{\theta}\in\mathbb{R}^{2n_0+1}, \bar{\theta}_{w,\bar{n}}\in\mathbb{R}^{2\bar{n}+1}}{\arg\min} V_\ell(\bar{\theta}_\ell, \bar{\theta}_{w,\bar{n},\ell}), \tag{38}$$

where

$$V_\ell(\bar{\theta}_\ell, \bar{\theta}_{w,\bar{n},\ell}) \triangleq \frac{1}{\ell}\sum_{k=1}^{\ell} \|\varepsilon(k \mid \bar{\theta}_\ell, \bar{\theta}_{w,\bar{n},\ell})\|^2. \tag{39}$$

In Sections 6 and 7, we use the Matlab System Identification Toolbox (Ljung, 2001) to obtain PEM-based estimates of the plant NMP zeros. Note that the predictor filters in (36) are functions of both G_0 and the noise model G_w. Unlike LS, PEM estimates a model of the noise dynamics G_w. As discussed in Söderström and Stoica (1988, p. 209), if the order of the model is underparameterised, the parameter estimates will converge to a minimiser of the asymptotic loss function, and thus will not give consistent estimates. Therefore, for all examples in this paper, we choose the order of the model of G_w to be equal to order n_0 of G_0. When utilising PEM within the System Identification Toolbox, we use the Box–Jenkins model structure.

4. DCL architectures

Consider the discrete-time closed-loop system in Figure 1 consisting of the single-input, single-output transfer functions, plant G of order n, and controller G_c of order n_c. Note that the plant and controller are defined as

$$G(\mathbf{q}) \triangleq \frac{N(\mathbf{q})}{D(\mathbf{q})}, \quad G_c(\mathbf{q}) \triangleq \frac{N_c(\mathbf{q})}{D_c(\mathbf{q})}. \tag{40}$$

The signal r is the command, y is the measured output, the error $e = r - y$, and w_0 represents sensor noise. DCL identification is used to identify the open-loop transfer function within a closed-loop system by using measurements of the input and output of the plant. In this section, we define three architectures associated with DCL identification. For all three DCL architectures, the estimated transfer function G_m is of order n, the plant $G = G_0$, and the noise transfer function $G_w = 1$. As stated in Section 3.2, the order of the PEM noise model estimate is chosen to be of order n.

4.1 DCL identification

In DCL identification, the plant $G = G_0$ is identified from the control input u to the output y as shown in Figure 1. The measured output y is given by

$$y(k) = y_0(k) + w_0(k), \tag{41}$$

where

$$y_0(k) \triangleq G(\mathbf{q})u(k). \tag{42}$$

4.2 ADCL/S identification

In ADCL identification, we specify an auxiliary signal v_0 that is added to the controller output. The auxiliary signal v_0 can be used to enhance persistency for identification, and we identify the transfer function from u to y. We consider two variations of ADCL identification, namely standard and intercalated auxiliary closed-loop identification.

In ADCL/S identification, the auxiliary signal v_0 is added to the controller output u_c, and $G = G_0$ is identified by using u and y, where

$$u(k) = u_c(k) + v_0(k), \tag{43}$$

as shown in Figure 2. Note that, if $v_0 = 0$, then ADCL/S is equivalent to DCL identification.

4.3 ADCL/I identification

In ADCL/I identification, the auxiliary signal v_0 is added to the controller output u_{int} between the numerator and denominator of the controller, where

$$u_{\text{int}}(k) = \frac{N_c(\mathbf{q})}{D_c(\mathbf{q})}e(k) + \frac{\mathbf{q}^{n_c} - D_c(\mathbf{q})}{D_c(\mathbf{q})}v_0(k), \tag{44}$$

and $G = G_0$ is identified by using u and y as shown in Figure 3, where, by using (43) and (44),

$$u(k) = \frac{N_c(\mathbf{q})}{D_c(\mathbf{q})}e(k) + \frac{\mathbf{q}^{n_c}}{D_c(\mathbf{q})}v_0(k). \tag{45}$$

Note that, if $v_0 = 0$, then $u = u_{\text{int}}$ and ADCL/I is equivalent to DCL identification.

5. ICL architectures

ICL identification can be used to estimate the open-loop transfer function within a closed-loop system by first estimating the

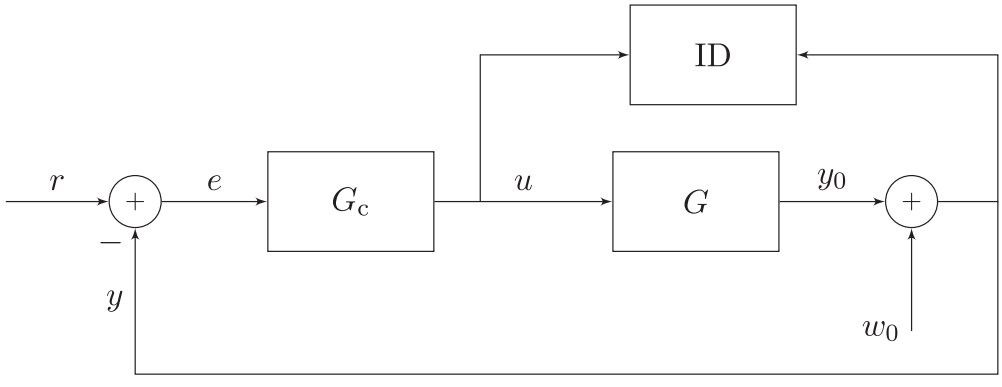

Figure 1. DCL identification from u to y. No auxiliary signal is used, and the persistency of u depends on the command r and the noise w_0.

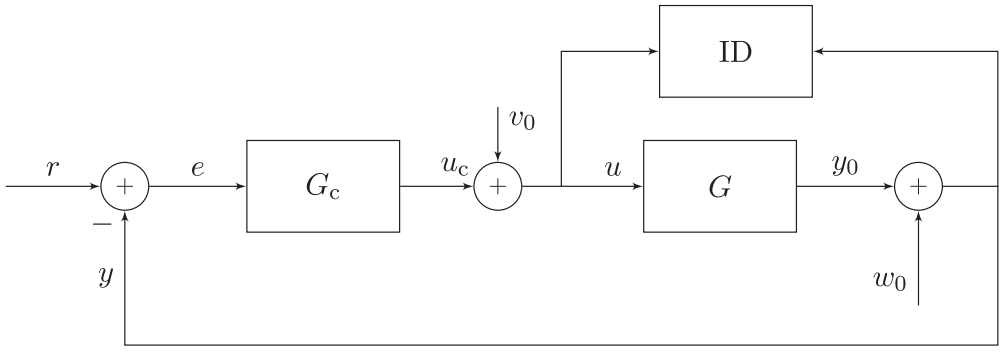

Figure 2. ADCL/S identification from u to y. To enhance persistency, the auxiliary signal v_0 is added to the controller output u_c.

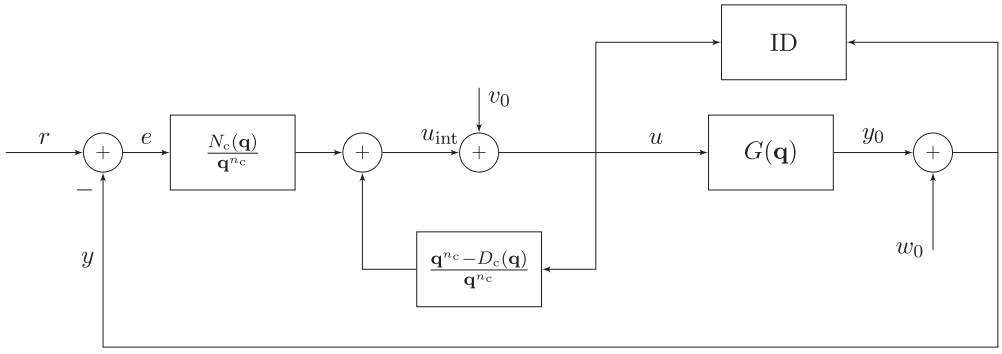

Figure 3. ADCL/I identification from u to y. To enhance persistency, the auxiliary signal v_0 is added to u_{int}, which is an internal signal in the controller. Intercalated injection arises in adaptive control (Rahman et al., 2017).

closed-loop system and then determining the open-loop plant by using knowledge of the controller. However, since feedback does not move the zeros of the plant, it is not necessary to extract a model of the open-loop plant in order to estimate the NMP plant zeros. In this section, we define three architectures associated with ICL identification. For all three ICL architectures, the true model G_0 varies for each architecture but the order for each is the same, namely, $n_0 = n + n_c$. The noise model G_w for each architecture is the sensitivity function, which is of order $\bar{n} = n + n_c$. Hence, $n_0 = \bar{n}$. Therefore, the estimated transfer function G_m is of order $n + n_c$, and, as stated in Section 3.2, the order of the PEM noise model estimate is chosen to be of order $n_0 = \bar{n}$.

5.1 ICL identification

In ICL identification, the closed-loop transfer function from r to y is identified as shown in Figure 4. The measured output y is given by

$$y(k) = T(\mathbf{q})r(k) + S(\mathbf{q})w_0(k), \tag{46}$$

where

$$T(\mathbf{q}) \triangleq \frac{G(\mathbf{q})G_c(\mathbf{q})}{1 + G_c(\mathbf{q})G(\mathbf{q})} = \frac{N(\mathbf{q})N_c(\mathbf{q})}{D(\mathbf{q})D_c(\mathbf{q}) + N(\mathbf{q})N_c(\mathbf{q})}, \tag{47}$$

$$S(\mathbf{q}) \triangleq \frac{1}{1 + G_c(\mathbf{q})G(\mathbf{q})} = \frac{D(\mathbf{q})D_c(\mathbf{q})}{D(\mathbf{q})D_c(\mathbf{q}) + N(\mathbf{q})N_c(\mathbf{q})}. \tag{48}$$

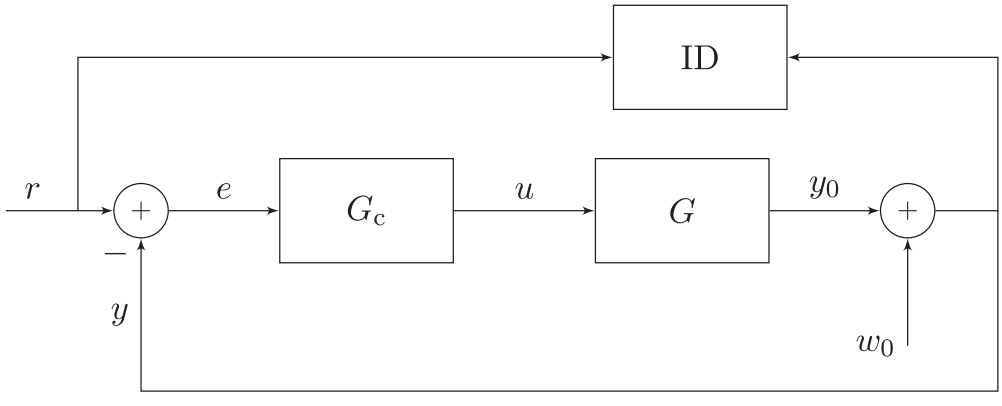

Figure 4. ICL identification from r to y. No auxiliary signal is used, and the persistency of u depends on the command r and the noise w_0.

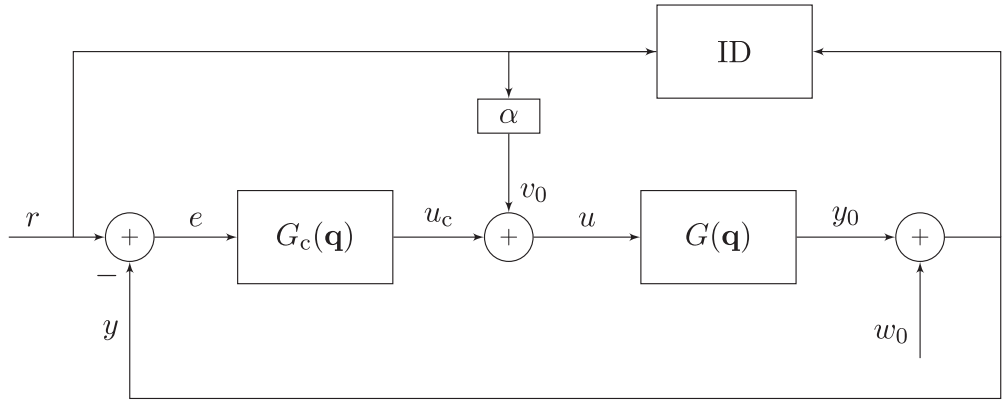

Figure 5. AICL/S identification from r to y. To enhance persistency, the auxiliary signal v_0, which is added to the controller output u_c, is proportional to the command r.

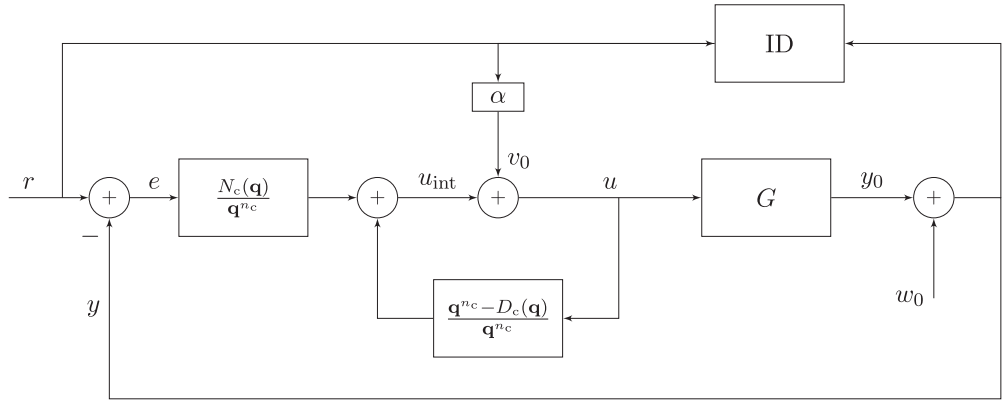

Figure 6. AICL/I identification from v_0 and r to y. To enhance persistency, the auxiliary signal v_0, which is added to u_{int}, is proportional to the command r.

Note that for ICL, $G_0(\mathbf{q}) = T(\mathbf{q})$ with $n_0 = n + n_c$, and $G_w(\mathbf{q}) = S(\mathbf{q})$.

5.2 AICL/S identification

In AICL identification as well as in ADCL identification, we specify an auxiliary signal v_0 that is added to the controller output. Unlike ADCL identification, however, we identify the closed-loop transfer function from r to y. In AICL/S, the auxiliary signal v_0 is added to the controller output u_c, as shown in Figure 5, where $v_0(k) = \alpha r(k)$ and α is a constant scalar value. The measured output y is given by

$$\begin{aligned} y(k) &= T(\mathbf{q})r(k) + \Gamma(\mathbf{q})v_0(k) + S(\mathbf{q})w_0(k) \\ &= (T(\mathbf{q}) + \alpha\Gamma(\mathbf{q}))r(k) + S(\mathbf{q})w_0(k), \end{aligned} \tag{49}$$

where T and S are defined in (47) and (48), respectively, and

$$\Gamma(\mathbf{q}) \triangleq \frac{G(\mathbf{q})}{1 + G_c(\mathbf{q})G(\mathbf{q})} = \frac{N(\mathbf{q})D_c(\mathbf{q})}{D(\mathbf{q})D_c(\mathbf{q}) + N(\mathbf{q})N_c(\mathbf{q})}. \tag{50}$$

Note that both of the zeros of the transfer functions T and Γ include the zeros of G. Furthermore, the transfer function from r to y is

$$G_0(\mathbf{q}) \triangleq T(\mathbf{q}) + \alpha\Gamma(\mathbf{q}) = \frac{N(\mathbf{q})(N_c(\mathbf{q}) + \alpha D_c(\mathbf{q}))}{D(\mathbf{q})D_c(\mathbf{q}) + N(\mathbf{q})N_c(\mathbf{q})}, \tag{51}$$

and the noise transfer function $G_w(\mathbf{q}) = S(\mathbf{q})$.

5.3 AICL/I identification

In AICL/I identification, the auxiliary signal v_0 is added to the controller output u_{int} between the numerator and denominator of the controller written as in (44) with the plant input u specified in (45). The transfer function from r to y is identified as shown in Figure 6, where $v_0(k) = \alpha r(k)$ and α is a constant scalar.

The plant output y is given by

$$\begin{aligned} y(k) &= T(\mathbf{q})r(k) + \Gamma_{\text{int}}(\mathbf{q})v_0(k) + S(\mathbf{q})w_0(k) \\ &= (T(\mathbf{q}) + \alpha\Gamma_{\text{int}}(\mathbf{q}))r(k) + S(\mathbf{q})w_0(k), \end{aligned} \tag{52}$$

where

$$\Gamma_{\text{int}}(\mathbf{q}) \triangleq \Gamma(\mathbf{q})\frac{\mathbf{q}^{n_c}}{D_c(\mathbf{q})} = \frac{N(\mathbf{q})\mathbf{q}^{n_c}}{D(\mathbf{q})D_c(\mathbf{q}) + N(\mathbf{q})N_c(\mathbf{q})}, \tag{53}$$

and T and S are defined in (47) and (48), respectively. Note that Γ_{int} given by (53) is similar to (50) except that the numerator of (53) includes n_c zeros at 0 rather than the controller denominator. Furthermore, the transfer function from r to y is given by

$$G_{0,\text{int}}(\mathbf{q}) \triangleq T(\mathbf{q}) + \alpha\Gamma_{\text{int}}(\mathbf{q}) = \frac{N(\mathbf{q})(N_c(\mathbf{q}) + \alpha\mathbf{q}^{n_c})}{D(\mathbf{q})D_c(\mathbf{q}) + N(\mathbf{q})N_c(\mathbf{q})}, \tag{54}$$

Table 1. Summary of DCL and ICL identification architectures.

Architecture	Input of G_m	G_m	Order of G_m
DCL	$G_c(\mathbf{q})e(k)$	$G(\mathbf{q})$	n
ADCL/S	$G_c(\mathbf{q})e(k) + v_0(k)$	$G(\mathbf{q})$	n
ADCL/I	$G_c(\mathbf{q})e(k) + \frac{\mathbf{q}^{n_c}}{D_c(\mathbf{q})}v_0(k)$	$G(\mathbf{q})$	n
ICL	$r(k)$	$T(\mathbf{q})$	$n+n_c$
AICL/S	$r(k)$	$T(\mathbf{q})+\alpha\Gamma(\mathbf{q})$	$n+n_c$
AICL/I	$r(k)$	$T(\mathbf{q})+\alpha\Gamma_{\text{int}}(\mathbf{q})$	$n+n_c$

For each architecture, the input, estimated transfer function, and order of the estimated transfer function are shown.

and the noise transfer function $G_w(\mathbf{q}) = S(\mathbf{q})$. For each identification architecture, Table 1 summarises the input to the estimated transfer function G_m, the estimated transfer function G_m, and the order of G_m.

6. Numerical investigation of DCL identification architectures

In this section, we investigate the accuracy of the NMP-zero estimates obtained from LS and PEM for the DCL architectures. Two examples are considered, both of which involve a third-order plant with one minimum-phase zero and one NMP zero. The first plant is asymptotically stable with a pole near the unit circle, while the second plant is unstable.

Table 2. DCL architecture exogenous-signal standard deviations for Example 6.1.

Architecture	Command r	Auxiliary signal v_0	Noise w_0
DCL	1.2	–	0.024
ADCL/S	1.2	0.5	0.034
ADCL/I	1.2	0.5	0.032

The standard deviation of the noise w_0 is adjusted so that the signal-to-noise ratio (57) is 31.5 dB for all simulations.

Example 1 (Asymptotically stable, NMP plant): Consider the asymptotically stable, NMP plant

$$G(\mathbf{q}) = \frac{(\mathbf{q}-0.6)(\mathbf{q}-1.5)}{(\mathbf{q}-0.1)(\mathbf{q}-0.5)(\mathbf{q}-0.98)}, \tag{55}$$

where the NMP zero is $z_{\text{NMP}} = 1.5$, with the controller

$$G_c(\mathbf{q}) = \frac{-0.5174\mathbf{q}^2 + 0.3315\mathbf{q} - 0.02795}{\mathbf{q}^3 - 0.2265\mathbf{q}^2 + 0.6855\mathbf{q} - 0.5142}. \tag{56}$$

The exogenous inputs r, w_0, and v_0 (where applicable) are independent zero-mean white noise signals whose standard deviations are listed in Table 2. For each architecture, the signal w_0 is adjusted so that the signal-to-noise ratio defined by

$$\text{SNR}_{\text{dB}} \triangleq 20\log_{10}\frac{\text{RMS}(y)}{\text{RMS}(w_0)}, \tag{57}$$

where $\text{RMS}(x)$ represents the root mean square of the sampled signal x. For all simulations, the signal-to-noise ratio is fixed to

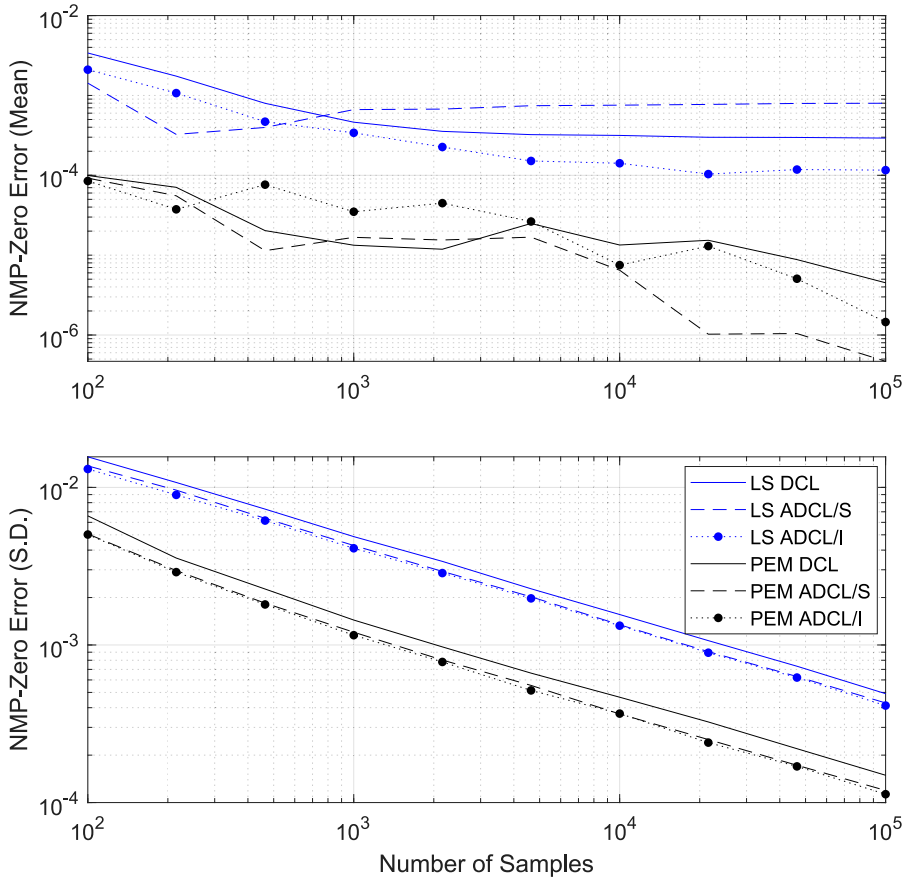

Figure 7. Simulation statistics for Example 6.1 for the DCL architectures for LS and PEM. For 2000 independent realisations of the inputs, the upper plot shows the absolute value of the mean NMP-zero error as a function of the number of samples. The lower plot shows the corresponding standard deviation of the NMP-zero error.

be 40, which corresponds to 31.5 dB. The motivation for maintaining a constant signal-to-noise ratio across all simulations is to ensure that comparisons between different architectures are meaningful.

For the DCL architectures using both identification algorithms, Figure 7 shows the mean absolute value of the error in the estimated NMP zero and the corresponding standard deviation averaged over 2000 independent realisations of the inputs for an increasing number of samples ℓ. To account for the presence of complex zeros in the estimated model and to determine the distance to the closed NMP zero, the NMP-zero estimate is found by taking the minimum of the difference between the real part of the roots of the estimated transfer function numerator and the true zero z_{NMP}. Note that, since the input u is correlated with the noise w_0, the LS estimates are not consistent for any of the three architectures. In contrast, the accuracy of the PEM estimates improves as the number of samples increases, especially for large numbers of samples, thus suggesting consistency. For both identification algorithms, the standard deviation decreases in a similar fashion, with PEM at a lower value.

Table 3. Comparison for the DCL architectures of the mean NMP-zero-error estimate based on LS with $\ell = 10^5$ and the NMP-zero error based on (32) for Example 6.1.

Architecture	Mean NMP-zero-error estimate based on LS	NMP-zero error based on (32)
DCL	2.92×10^{-4}	4.54×10^{-3}
ADCL/S	7.98×10^{-4}	3.61×10^{-3}
ADCL/I	1.16×10^{-4}	3.63×10^{-3}

Table 4. DCL architecture input signal standard deviations for Example 6.2.

Architecture	Command r	Auxiliary signal v_0	Noise w_0
DCL	0.134	–	0.033
ADCL/S	0.134	0.5	0.147
ADCL/I	0.134	0.5	0.040

The noise w_0 is scaled so that the signal-to-noise ratio (57) is 31.5 dB for all simulations.

In all three *DCL* architectures, the input r is correlated with the noise signal w_0, and thus Theorem 1 is not applicable. Table 3 compares the accuracy of the LS estimates of the NMP zero obtained from data from 2000 independent realisations of the inputs averaged over $\ell = 10^5$ samples with the estimate of the NMP-zero based on the analytical expression for $\lim_{\ell\to\infty} \hat{\theta}_\ell^{\text{LS}}$ given by (32). Note that, because of the correlation between the input u and the noise w_0, the NMP-zero-error bias given by (32) is inaccurate, as expected.

Example 2 (Unstable, NMP plant): Consider the unstable, NMP plant

$$G(\mathbf{q}) = \frac{(\mathbf{q} - 0.6)(\mathbf{q} - 1.5)}{(\mathbf{q} - 0.1)(\mathbf{q} - 0.5)(\mathbf{q} - 1.3)}, \tag{58}$$

where the NMP zero is $z_{\text{NMP}} = 1.5$, with the stabilising controller

$$G_c(\mathbf{q}) = \frac{-6.357\mathbf{q}^2 + 3.827\mathbf{q} - 0.3191}{\mathbf{q}^3 + 0.1312\mathbf{q}^2 + 6.943\mathbf{q} - 4.419}. \tag{59}$$

The inputs r, w_0, and v_0 (where applicable) are independent zero-mean white noise signals whose standard deviations are

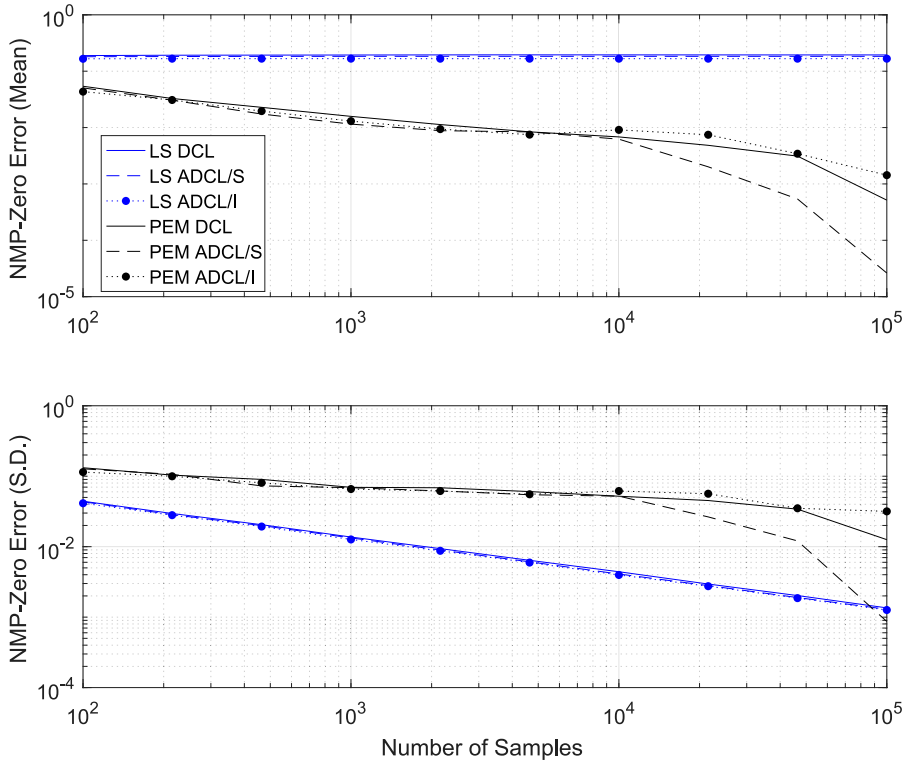

Figure 8. Simulation statistics for Example 6.2 for all three DCL architectures using LS and PEM. For 2000 independent realisations of the inputs, the upper plot shows the absolute value of the mean NMP-zero error as a function of an increasing number of samples. The lower plot shows the corresponding standard deviation of the NMP-zero error.

listed in Table 4. As in the previous example, w_0 is adjusted so that the signal-to-noise ratio is 31.5 dB for all simulations.

For all three DCL architectures using LS and PEM, Figure 8 shows the absolute value of the mean errors in the NMP-zero estimates and the standard deviations based on 2000 independent realisations of the inputs for various values of ℓ. Note that, as in the previous example, the LS estimates are not consistent for any of the three architectures since the input u is correlated with the noise w_0; in contrast the PEM estimates appear to be consistent. For both identification algorithms, the standard deviation decreases as the number of samples increases.

In all three *DCL* architectures, the input r is correlated with the noise signal w_0, and thus Theorem 1 is not applicable. Table 5 compares the accuracy of the LS estimates of the NMP zero obtained from data from 2000 independent realisations of the inputs averaged over $\ell = 10^5$ samples with the estimate of the NMP-zero based on the analytical expression for $\lim_{\ell\to\infty}\hat{\theta}_\ell^{\mathrm{LS}}$ given by (32). Note that, because of the correlation between the input u and the noise w_0, the NMP-zero-error bias given by (32) is inaccurate, as expected.

7. Numerical investigation of ICL identification architectures

In this section, Examples 6.1 and 6.2 are used to compare the consistency of the NMP-zero estimates obtained from LS and PEM for all three ICL architectures.

Table 5. Comparison for the DCL architectures of the mean NMP-zero-error estimate based on LS with $\ell = 10^5$ and the NMP-zero error based on (32) for Example 6.2.

Architecture	Mean NMP-zero-error estimate based on LS	NMP-zero error based on (32)
DCL	0.194	0.353
ADCL/S	0.183	0.333
ADCL/I	0.167	0.309

Table 6. ICL architecture input signal standard deviations for Example 7.1.

Architecture	Command r	Noise w_0
ICL	1.2	0.0235
AICL/S	1.2	0.0341
AICL/I	1.2	0.0267

The noise w_0 is scaled so that the signal-to-noise ratio (57) remains at 31.5 dB for all simulations.

Example 3 (Asymptotically stable, NMP plant): We consider the asymptotically stable NMP plant and controller in Example 6.1. The inputs r and w_0 are independent zero-mean white noise signals and, for the AICL/S and AICL/I architectures, we choose $\alpha = 0.5$ so that $v_0(k) = 0.5r(k)$. As in the previous examples, the standard deviation of w_0 is adjusted so that the signal-to-noise ratio is 31.5 dB for all simulations shown in Table 6.

For the ICL architectures using both identification algorithms, Figure 9 shows the absolute value of the mean errors in the estimated NMP zero and the standard deviations based on

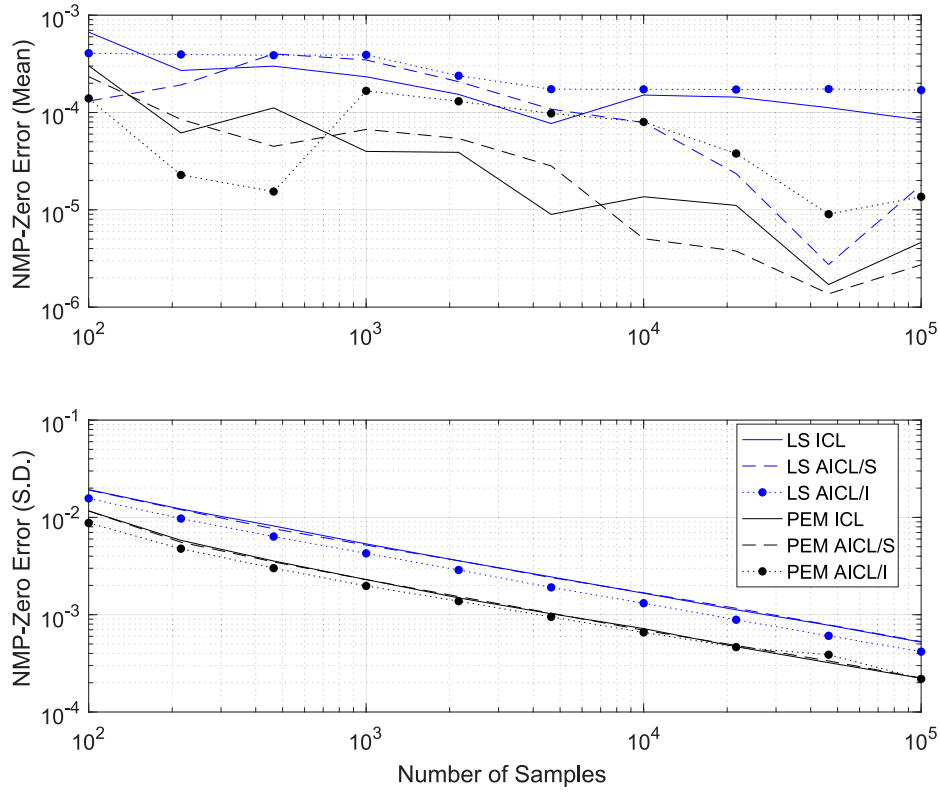

Figure 9. Simulation statistics for Example 7.1 for all three ICL architectures using both identification algorithms. For 2000 independent realisations of the inputs, the upper plot shows the absolute value of the mean NMP-zero error as a function of an increasing number of samples. The lower plot shows the corresponding standard deviation of the NMP-zero error.

Table 7. Comparison for the ICL identification architectures of the mean NMP-zero-error estimate based on LS with $\ell = 10^5$ and the NMP-zero error based on (32) for Example 7.1.

Architecture	Mean NMP-zero-error estimate based on LS	NMP-zero error based on (32)
ICL	0.84×10^{-4}	1.28×10^{-4}
AICL/S	1.80×10^{-5}	2.76×10^{-5}
AICL/I	1.70×10^{-4}	1.63×10^{-4}

Table 8. ICL architecture input signal standard deviations for Example 7.2.

Architecture	Command r	Noise w_0
ICL	0.134	0.0334
AICL/S	0.134	0.0383
AICL/I	0.134	0.0322

The noise w_0 is scaled so that the signal-to-noise ratio (57) is 31.5 dB for all simulations.

2000 independent realisations of the inputs for various values of ℓ. Note that, for all three architectures, the LS estimates are not consistent. Although the input signals r and w_0 are uncorrelated and the model and noise transfer functions have the same denominator, the matrix $\mathcal{H}_w N_w$ is not zero, and thus (i) of Theorem 1 is not satisfied. For both identification algorithms, the standard deviation decreases as the number of samples increases.

In all three *ICL* architectures, the input r is uncorrelated with the noise signal w_0, and thus Theorem 1 is applicable. For 2000 independent realisations of the inputs and $\ell = 10^5$ samples, Table 7 compares the averages of the NMP-zero errors using LS with the analytical expression (32) for the bias given by Theorem 1. Compared with the cases considered in Figures 3 and 5, where r and w_0 are correlated, the NMP-zero error based on (32) is significantly closer to the numerical estimates, which demonstrates agreement between Theorem 1 and the numerical data.

Example 4 (Unstable, NMP plant): To compare the three indirect architectures, we consider the unstable, NMP plant (58) with the stabilising controller (59). The inputs r and w_0 are independent zero-mean white noise signals and, for the AICL/S and AICL/I architectures, we set $\alpha = 0.5$ so that $v_0(k) = 0.5r(k)$. As in the previous examples, the standard deviation of w_0 is adjusted so that the signal-to-noise ratio is 31.5 dB for all simulations shown in Table 8.

For all three ICL architectures using both identification algorithms, Figure 10 shows the absolute value of the mean errors in the estimated NMP zero and standard deviations based on 2000 independent realisations of the inputs for various values of ℓ. Note that, for all three architectures, the LS estimates are not consistent. Although the input signals r and w_0 are uncorrelated and the model and noise transfer functions have the same denominator, the matrix $\mathcal{H}_w N_w$ is not zero, and thus (i) of Theorem 1 is not satisfied. For both identification algorithms, the standard deviation decreases as $O(\ell^{-1/2})$ as the number of samples increases.

In all three ICL architectures, the input r is uncorrelated with the noise signal w_0, and thus Theorem 1 is applicable. For 2000 independent realisations of the inputs and $\ell = 10^5$ samples, Table 9 compares the averages of the NMP-zero errors

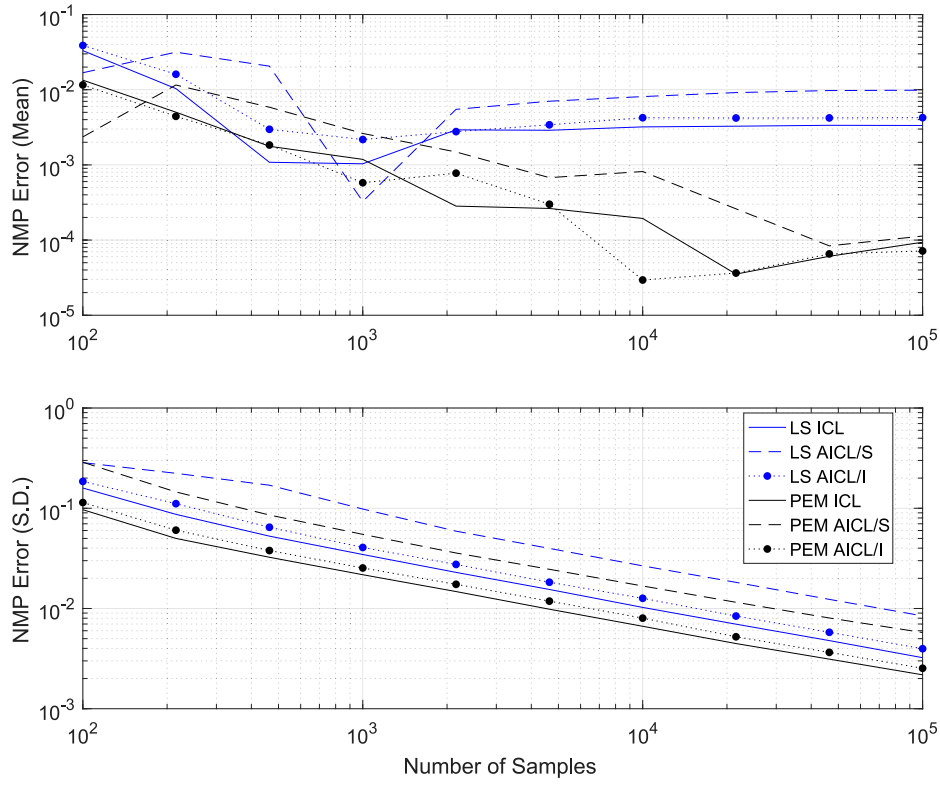

Figure 10. Simulation statistics for Example 7.2 for all three ICL architectures using both identification algorithms. For 2000 independent realisations of the inputs, the upper plot shows the absolute value of the mean NMP-zero error as a function of an increasing number of samples. The lower plot shows the corresponding standard deviation of the NMP-zero error.

Table 9. ICL architecture comparison between the LS NMP-zero-error estimates with $\ell = 10^5$ and the NMP-zero error based on (32) for Example 7.2.

Architecture	Mean NMP-zero-error estimate based on LS	NMP-zero error based on (32)
ICL	3.05×10^{-3}	3.26×10^{-3}
AICL/S	9.19×10^{-3}	9.44×10^{-3}
AICL/I	3.94×10^{-3}	3.95×10^{-3}

using LS with the analytical expression (32) for the bias given by Theorem 1. Compared with the cases considered in Figures 3 and 5, where r and w_0 are correlated, the NMP-zero error based on (32) is significantly closer to the numerical estimates, which demonstrates agreement between Theorem 1 and the numerical data.

8. Conclusions and future research

This paper numerically investigated the effectiveness of various architectures for closed-loop identification, including three DCL identification architectures and three ICL identification architectures. These architectures included standard cases with and without auxiliary signals as well as two novel architectures involving intercalated injection of the auxiliary signal. IIR models were fit using LS estimation, which provided a baseline method, and PEM, which account for noise correlation. To simplify the study, the plant order was assumed to be known; for each indirect architecture, the auxiliary signal was chosen to be a multiple of the command; and errors-in-variables (EIV) noise was not considered. To allow a meaningful comparison, the signal-to-noise ratio was normalised across all architectures. Motivated by adaptive control of systems with NMP zeros and taking advantage of the fact that zeros are not moved by feedback, the performance metric is the accuracy of the estimates of the NMP zero of the plant. Two examples were considered, both of which were third-order and NMP. One plant was asymptotically stable and one was unstable.

As expected, for all of the architectures and for both examples, the LS estimates exhibited bias, whereas the PEM estimates indicated consistency. In fact, Figures 7 and 8 show that PEM is highly accurate for large numbers of samples. In comparing architectures, some observations can be made. Based on PEM for DCL identification, standard injection of the auxiliary signal appears to be advantageous; DCL and ADCL with intercalated injection provide roughly the same accuracy. For ICL identification, the trends are mixed. For the asymptotically stable plant, standard injection of the auxiliary signal appears to be advantageous, whereas, for the unstable plant, intercalated injection appears advantageous.

In view of the practical importance of closed-loop identification, further investigation is warranted. A more detailed study would include a comparison with instrumental variables (Söderström & Stoica, 2002; Taylor, Pedregal, Young, & Tych, 2007; Young, Jakeman, & McMurtrie, 1980) as an alternative to PEM. An extension of practical value is the case where noise corrupts the signals used for identification; this is an *EIV* identification problem (Diversi, 2008; Hjalmarsson, Martensson, Rojas, & Soderstrom, 2011; Söderström, 2007; Van Huffel & Lemmerling, 2013). The cases of unknown plant order, MIMO plants, and higher order plants with multiple NMP zeros are of interest. Finally, experimental application of the various architectures and identification methods may shed more light on the relative advantages of the competing architectures.

Acknowledgments

The authors are grateful to the reviewers for their careful reading of the paper and numerous helpful suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported in part by National Science Foundation (NSF) [grant number CMMI 1536834].

References

Aljanaideh, K. F., & Bernstein, D. S. (2017). Closed-loop identification of unstable systems using noncausal FIR models. *International Journal of Control*, *90*, 168–185.

Aljanaideh, K. F., & Bernstein, D. S. (2018). Initial conditions in time- and frequency-domain system identification: Implications of the shift operator versus the $\mathcal{Z}$ and discrete fourier transforms. *Control Systems Magazine*, *38*, 80–93.

de Vlugt, E., Schouten, A. C., & van der Helm, F. C. (2003). Closed-loop multivariable system identification for the characterization of the dynamic arm compliance using continuous force disturbances: A model study. *Journal of Neuroscience Methods*, *122*, 123–140.

Diversi, R. (2008). A bias-compensated identification approach for noisy FIR models. *IEEE Signal Processing Letters*, *15*, 325–328.

Engelhart, D., Boonstra, T., Aarts, R., Schouten, A., & van der Kooij, H. (2016). Comparison of closed-loop system identification techniques to quantify multi-joint human balance control. *Annual Reviews in Control*, *41*, 58–70.

Epperlein, J. P., Bamieh, B., & Åström, K. J. (2015). Thermoacoustics and the Rijke tube: Experiments, identification, and modeling. *IEEE Control Systems Magazine*, *35*, 57–77.

Eykhoff, P. (1974). *System identification: Parameter and state estimation*. London, UK: Wiley.

Forssell, U., & Ljung, L. (1999). Closed-loop identification revisited. *Automatica*, *35*, 1215–1241.

Freudenberg, J. S., Hollot, C. V., Middleton, R. H., & Toochinda, V. (2003). Fundamental design limitations of the general control configuration. *IEEE Transactions on Automatic Control*, *48*, 1355–1370.

Gustavsson, I., Ljung, L., & Soderstrom, T. (1977). Identification of process in closed loop – Identifiability and accuracy aspects. *Automatica*, *13*, 59–75.

Hjalmarsson, H. (2005). From experiment design to closed-loop control. *Automatica*, *41*, 393–438.

Hjalmarsson, H., Gevers, M., & De Bruyne, F. (1996). For model-based control design, closed-loop identification gives better performance. *Automatica*, *32*, 1659–1673.

Hjalmarsson, H., Martensson, J., Rojas, C. R., & Soderstrom, T. (2011). On the accuracy in errors-in-variables identification compared to prediction-error identification. *Automatica*, *47*, 2704–2712.

Ho, B., & Kalman, R. E. (1966). Effective construction of linear state-variable models from input/output functions. *Regelungstechnik*, *14*, 545–548.

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

Juang, J. N., & Pappa, R. S. (1985). An Eigensystem realization algorithm for modal parameter identififfcation and model reduction. *Journal of Guidance, Control, and Dynamics*, *8*, 620–627.

Landau, I. (1999). From robust control to adaptive control. *Control Engineering Practice*, *7*, 1113–1124.

Langer, J., & Landau, I. (1996). Improvement of robust digital control by identification in the closed loop. Application to a 360 flexible arm. *Control Engineering Practice*, *4*, 1637–1646.
Ljung, L. (1976). On the consistency of prediction error identification methods. *Mathematics in Science and Engineering*, *126*, 121–164.
Ljung, L. (1999). *System identification: Theory for the user*. 2nd ed.Upper Saddle River, NJ: Prentice-Hall Information and Systems Sciences.
Ljung, L. (2001). System identification toolbox: For use with Matlab: Computation, visualization, programming. User's guide, version 5. The Mathworks.
Ljung, L. (2002). Prediction error estimation methods. *Circuits, Systems, and Signal Processing*, *21*, 11–21.
Ljung, L. (2010). Perspectives on system identification. *Annual Reviews in Control*, *34*, 1–12.
Ljung, L., & Forssell, U. (1999). An alternative motivation for the indirect approach to closed-loop identification. *IEEE Transactions on Automatic Control*, *44*, 2206–2209.
Ljung, L., & Wahlberg, B. (1992). Asymptotic properties of the least-squares method for estimating transfer functions and disturbance spectra. *Advances in Applied Probability*, *24*, 412–440.
Martensson, J., & Hjalmarsson, H. (2005a). *Closed loop identificaton of unstable poles and non-minimum phase zeros*. Proceedings of the IFAC 16th triennial world congress (pp. 518–523), Prague, Czech Republic.
Martensson, J., & Hjalmarsson, H. (2005b). *Exact quantification of the variance of estimated zeros*. Proceedings of the conference on decision and control (pp. 4275–4280), Seville, Spain.
Martensson, J., & Hjalmarsson, H. (2009). Variance-error quantification for identified poles and zeros. *Automatica*, *45*, 2512–2525.
Martensson, J., Jansson, H., & Hjalmarsson, H. (2005). *Input design for identification of zeros*. Proceedings of the IFAC 16th triennial world congress (pp. 470–475), Prague, Czech Republic.
Rahman, Y., Xie, A., & Bernstein, D. S. (2017). Retrospective cost adaptive control: Pole placement, frequency response, and connections with LQG control. *IEEE Control Systems Magazine*, *35*, (pp. 28–69).
Recht, B., Fazel, M., & Parrilo, P. A. (2010). Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review*, *52*, 471–501.
Rivera, D. E., Lee, H., Mittelmann, H. D., & Braun, M. W. (2007). High-purity distillation. *IEEE Control Systems Magazine*, *27*, 72–89.
Rojas, C. R., Hjalmarsson, H., Gerencser, L., & Martensson, J. (2011). An adaptive method for consistent estimation of real-valued non-minimum phase zeros in stable LTI systems. *Automatica*, *47*, 1388–1398.
Shen, B., Ding, F., Alsaedi, A., & Hayat, T. (2017). Gradient-based recursive identification methods for input nonlinear equation error closed-loop systems. *Circuits, Systems, and Signal Processing*, *36*, 2166–2183.
Smith, R. S. (1998). Closed-loop identification of flexible structures: An experimental example. *Journal of Guidance, Control, and Dynamics*, *21*, 435–440.
Smith, R. S. (2014). Frequency domain subspace identification using nuclear norm minimization and Hankel matrix realizations. *IEEE Transactions on Automatic Control*, *59*, 2886–2896.
Söderström, T. (2007). Errors-in-variables methods in system identification. *Automatica*, *43*, 939–958.
Söderström, T., & Stoica, P. (1989). *System identification*. London, UK: Prentice-Hall.
Söderström, T., & Stoica, P. (2002). Instrumental variable methods for system identification. *Circuits, Systems, and Signal Processing*, *21*, 1–9.
Taylor, C. J., Pedregal, D. J., Young, P. C., & Tych, W. (2007). Environmental time series analysis and forecasting with the captain toolbox. *Environmental Modelling & Software*, *22*, 797–814.
Van den Hof, P. (1998). Closed-loop issues in system identification. *Annual Reviews in Control*, *22*, 173–186.
Van den Hof, P., & Schrama, R. J. (1993). An indirect method for transfer function estimation from closed loop data. *Automatica*, *29*, 1523–1527.
Van Huffel, S., & Lemmerling, P. (2013). *Total least squares and errors-in-variables modeling: Analysis, algorithms and applications*. Dordrecht, The Netherlands: Springer.
von Wangenheim, L. (2012). Phase margin determination in a closed-loop configuration. *Circuits, Systems, and Signal Processing*, *31*, 1917–1926.
Young, P., Jakeman, A., & McMurtrie, R. (1980). An instrumental variable method for model order identification. *Automatica*, *16*, 281–294.
Zhu, Y., & Butoyi, F. (2002). Case studies on closed-loop identification for MPC. *Control Engineering Practice*, *10*, 403–417.

Appendices

Appendix 1. Proof of Lemma 1

First, note that

$$\Phi_{y,\ell} = \Phi_{y_0,\ell} + \tilde{\Phi}_{w_0,\ell}, \tag{A1}$$

where

$$\Phi_{y_0,\ell} = \begin{bmatrix} -y_0(\tilde{n}-1) & \cdots & -y_0(\tilde{n}-n_0) \\ \vdots & \cdots & \vdots \\ -y_0(\ell-1) & \cdots & -y_0(\ell-n_0) \end{bmatrix} \in \mathbb{R}^{(\ell-\tilde{n}+1)\times n_0}, \tag{A2}$$

$$\tilde{\Phi}_{w_0,\ell} = \begin{bmatrix} -w_0(\tilde{n}-1) & \cdots & -w_0(\tilde{n}-n_0) \\ \vdots & \cdots & \vdots \\ -w_0(\ell-1) & \cdots & -w_0(\ell-n_0) \end{bmatrix} \in \mathbb{R}^{(\ell-\tilde{n}+1)\times n_0}. \tag{A3}$$

Next, note that

$$\lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y_0,\ell}^{\mathrm{T}}\Phi_{y,\ell} \overset{\text{wp1}}{=} \lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y_0,\ell}^{\mathrm{T}}\Phi_{y_0,\ell},$$
$$\lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{y,\ell} \overset{\text{wp1}}{=} \lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{y_0,\ell}, \tag{A4}$$

$$\lim_{\ell\to\infty} \frac{1}{\ell}\tilde{\Phi}_{w_0,\ell}^{\mathrm{T}}\Phi_{y,\ell} \overset{\text{wp1}}{=} \lim_{\ell\to\infty} \frac{1}{\ell}\tilde{\Phi}_{w_0,\ell}^{\mathrm{T}}\tilde{\Phi}_{w_0,\ell} \overset{\text{wp1}}{=} \sigma_w^2\tilde{\mathcal{H}}_{G_w}. \tag{A5}$$

Using (22), (A1), (A4), and (A5) yields

$$\begin{aligned}
\Gamma &\overset{\text{wp1}}{=} \lim_{\ell\to\infty} \frac{1}{\ell}\begin{bmatrix} \Phi_{y,\ell}^{\mathrm{T}}\Phi_{y,\ell} & \Phi_{y,\ell}^{\mathrm{T}}\Phi_{u_0,\ell} \\ \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{y,\ell} & \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{u_0,\ell} \end{bmatrix} \\
&= \lim_{\ell\to\infty} \frac{1}{\ell}\begin{bmatrix} \Phi_{y_0,\ell}^{\mathrm{T}}\Phi_{y_0,\ell} + \tilde{\Phi}_{w_0,\ell}^{\mathrm{T}}\tilde{\Phi}_{w_0,\ell} & \Phi_{y_0,\ell}^{\mathrm{T}}\Phi_{u_0,\ell} \\ \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{y_0,\ell} & \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{u_0,\ell} \end{bmatrix}, \\
&= \lim_{\ell\to\infty} \frac{1}{\ell}\begin{bmatrix} \Phi_{y_0,\ell}^{\mathrm{T}}\Phi_{y_0,\ell} & \Phi_{y_0,\ell}^{\mathrm{T}}\Phi_{u_0,\ell} \\ \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{y_0,\ell} & \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{u_0,\ell} \end{bmatrix} \\
&\quad + \lim_{\ell\to\infty} \frac{1}{\ell}\begin{bmatrix} \tilde{\Phi}_{w_0,\ell}^{\mathrm{T}}\tilde{\Phi}_{w_0,\ell} & 0_{n_0\times(n_0+1)} \\ 0_{(n_0+1)\times n_0} & 0_{(n_0+1)\times(n_0+1)} \end{bmatrix}, \\
&= \Gamma_0 + \begin{bmatrix} \sigma_w^2\tilde{\mathcal{H}}_{G_w} & 0_{n_0\times(n_0+1)} \\ 0_{(n_0+1)\times n_0} & 0_{(n_0+1)\times(n_0+1)} \end{bmatrix}.
\end{aligned}$$

Appendix 2: Proof of Theorem 1

Note that (19) implies that

$$\Phi_{y,u_0,\ell}^{\mathrm{T}}\Phi_{y,u_0,\ell}\hat{\theta}_\ell^{\mathrm{LS}} = \Phi_{y,u_0,\ell}^{\mathrm{T}}\Psi_{y,\ell}. \tag{A6}$$

Dividing (A6) by ℓ and letting $\ell \to \infty$ yields

$$\lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y,u_0,\ell}^{\mathrm{T}}\Phi_{y,u_0,\ell}\hat{\theta}_\ell^{\mathrm{LS}} = \lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y,u_0,\ell}^{\mathrm{T}}\Psi_{y,\ell}. \tag{A7}$$

Using (22) we can write

$$\lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y,u_0,\ell}^{\mathrm{T}}\Phi_{y,u_0,\ell}\hat{\theta}_\ell^{\mathrm{LS}} \overset{\text{wp1}}{=} \Gamma \lim_{\ell\to\infty} \hat{\theta}_\ell^{\mathrm{LS}}. \tag{A8}$$

Next, note that

$$\Psi_{y,\ell} = \Psi_{y_0,\ell} + \Psi_{w_0,\ell}, \tag{A9}$$

where

$$\Psi_{y_0,\ell} = \Phi_{y_0,u_0,\ell}\theta, \quad \Psi_{w_0,\ell} = \Phi_{w_0,w,\ell}\theta_w, \tag{A10}$$

and

$$\Phi_{y_0,u_0,\ell} \triangleq \begin{bmatrix} \phi_{y_0,u_0}(\tilde{n}) \\ \vdots \\ \phi_{y_0,u_0}(\ell) \end{bmatrix} \in \mathbb{R}^{(\ell-\tilde{n}+1)\times(2n_0+1)},$$

$$\Phi_{w_0,w,\ell} \triangleq \begin{bmatrix} \phi_{w_0,w}(\tilde{n}) \\ \vdots \\ \phi_{w_0,w}(\ell) \end{bmatrix} \in \mathbb{R}^{(\ell-\tilde{n}+1)\times(2\tilde{n}+1)}.$$

Using (A9) and (A10) we can write

$$\begin{aligned} \Phi_{y,u_0,\ell}^{\mathrm{T}}\Psi_{y,\ell} &= \Phi_{y,u_0,\ell}^{\mathrm{T}}\Psi_{y_0,\ell} + \Phi_{y,u_0,\ell}^{\mathrm{T}}\Psi_{w_0,\ell} \\ &= \Phi_{y,u_0,\ell}^{\mathrm{T}}\Phi_{y_0,u_0,\ell}\theta + \Phi_{y,u_0,\ell}^{\mathrm{T}}\Phi_{w_0,w,\ell}\theta_w, \end{aligned} \tag{A11}$$

where

$$\Phi_{y_0,u_0,\ell} \triangleq [\Phi_{y_0,\ell}\ \Phi_{u_0,\ell}], \quad \Phi_{w_0,w,\ell} \triangleq [\Phi_{w_0,\ell}\ \Phi_{w,\ell}], \tag{A12}$$

$$\Phi_{y_0,\ell} \triangleq \begin{bmatrix} -y_0(\tilde{n}-1) & \cdots & -y_0(\tilde{n}-n_0) \\ \vdots & \cdots & \vdots \\ -y_0(\ell-1) & \cdots & -y_0(\ell-n_0) \end{bmatrix},$$

$$\Phi_{w,\ell} \triangleq \begin{bmatrix} w(\tilde{n}) & \cdots & w(\tilde{n}-\bar{n}) \\ \vdots & \cdots & \vdots \\ w(\ell) & \cdots & w(\ell-\bar{n}) \end{bmatrix}, \tag{A13}$$

$$\Phi_{w_0,\ell} \triangleq \begin{bmatrix} -w_0(\tilde{n}-1) & \cdots & -w_0(\tilde{n}-\bar{n}) \\ \vdots & \cdots & \vdots \\ -w_0(\ell-1) & \cdots & -w_0(\ell-\bar{n}) \end{bmatrix}. \tag{A14}$$

Moreover, note that

$$\Phi_{y,\ell} = \Phi_{y_0,\ell} + \tilde{\Phi}_{w_0,\ell},$$

where

$$\tilde{\Phi}_{w_0,\ell} \triangleq \begin{bmatrix} -w_0(\tilde{n}-1) & \cdots & -w_0(\tilde{n}-n_0) \\ \vdots & \cdots & \vdots \\ -w_0(\ell-1) & \cdots & -w_0(\ell-n_0) \end{bmatrix} \in \mathbb{R}^{(\ell-\tilde{n}+1)\times n_0}. \tag{A15}$$

Using (A12), (A11) can be written as

$$\begin{aligned} \Phi_{y,u_0,\ell}^{\mathrm{T}}\Psi_{y,\ell} &= \begin{bmatrix} \Phi_{y,\ell}^{\mathrm{T}}\Phi_{y_0,\ell} & \Phi_{y,\ell}^{\mathrm{T}}\Phi_{u_0,\ell} \\ \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{y_0,\ell} & \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{u_0,\ell} \end{bmatrix}\theta \\ &\quad + \begin{bmatrix} \Phi_{y,\ell}^{\mathrm{T}}\Phi_{w_0,\ell} & \Phi_{y,\ell}^{\mathrm{T}}\Phi_{w,\ell} \\ \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{w_0,\ell} & \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{w,\ell} \end{bmatrix}\theta_w. \end{aligned} \tag{A16}$$

Since $\mathcal{U}$ and $\mathcal{W}$ are uncorrelated zero-mean random processes, it follows that

$$\lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y,\ell}^{\mathrm{T}}\Phi_{y_0,\ell} \overset{\mathrm{wp1}}{=} \lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y_0,\ell}^{\mathrm{T}}\Phi_{y_0,\ell}, \tag{A17}$$

$$\lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y,\ell}^{\mathrm{T}}\Phi_{u_0,\ell} \overset{\mathrm{wp1}}{=} \lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y_0,\ell}^{\mathrm{T}}\Phi_{u_0,\ell}, \tag{A18}$$

$$\lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y,\ell}^{\mathrm{T}}\Phi_{w_0,\ell} \overset{\mathrm{wp1}}{=} \lim_{\ell\to\infty} \frac{1}{\ell}\tilde{\Phi}_{w_0,\ell}^{\mathrm{T}}\Phi_{w_0,\ell} \overset{\mathrm{wp1}}{=} \sigma_w^2\mathcal{H}_{G_w}, \tag{A19}$$

$$\lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y,\ell}^{\mathrm{T}}\Phi_{w,\ell} \overset{\mathrm{wp1}}{=} \lim_{\ell\to\infty} \frac{1}{\ell}\tilde{\Phi}_{w_0,\ell}^{\mathrm{T}}\Phi_{w,\ell} \overset{\mathrm{wp1}}{=} -\sigma_w^2\mathcal{H}_w, \tag{A20}$$

$$\lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{w_0,\ell} \overset{\mathrm{wp1}}{=} 0_{n_0\times\tilde{n}}, \quad \lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{w,\ell} \overset{\mathrm{wp1}}{=} 0_{n_0\times\tilde{n}+1}. \tag{A21}$$

Using (A17)–(A21), dividing (A16) by ℓ, and letting $\ell \to \infty$ yields

$$\begin{aligned} \lim_{\ell\to\infty} \frac{1}{\ell}\Phi_{y,u_0,\ell}^{\mathrm{T}}\Psi_{y,\ell} &\overset{\mathrm{wp1}}{=} \lim_{\ell\to\infty} \frac{1}{\ell}\begin{bmatrix} \Phi_{y_0,\ell}^{\mathrm{T}}\Phi_{y_0,\ell} & \Phi_{y_0,\ell}^{\mathrm{T}}\Phi_{u_0,\ell} \\ \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{y_0,\ell} & \Phi_{u_0,\ell}^{\mathrm{T}}\Phi_{u_0,\ell} \end{bmatrix}\theta \\ &\quad + \begin{bmatrix} \sigma_w^2\mathcal{H}_{G_w} & -\sigma_w^2\mathcal{H}_w \\ 0_{(n_0+1)\times\tilde{n}} & 0_{(n_0+1)\times(\tilde{n}+1)} \end{bmatrix}\theta_w \\ &= \Gamma_0\theta + \begin{bmatrix} \sigma_w^2\mathcal{H}_{G_w}D_w - \sigma_w^2\mathcal{H}_wN_w \\ 0_{(n_0+1)\times 1} \end{bmatrix}. \end{aligned} \tag{A22}$$

Using (A8) and (A22), (A7) yields

$$\Gamma \lim_{\ell\to\infty} \hat{\theta}_\ell^{\mathrm{LS}} = \Gamma_0\theta + \begin{bmatrix} \sigma_w^2\mathcal{H}_{G_w}D_w - \sigma_w^2\mathcal{H}_wN_w \\ 0_{(n_0+1)\times 1} \end{bmatrix}.$$