2018 Annual American Control Conference (ACC)
June 27–29, 2018. Wisconsin Center, Milwaukee, USA

System Identification Using Composite FIR/IIR Models

Khaled F. Aljanaideh[1] and Dennis S. Bernstein[2]

Abstract— **Since physical systems usually have infinite-impulse-response (IIR) dynamics, IIR models are widely used for system identification. Using IIR models, however, requires knowledge of the system order, which, if chosen incorrectly, may yield poor parameter estimates. On the other hand, finite-impulse-response (FIR) models of sufficiently high order can approximate asymptotically stable systems without knowledge of the system order. However, FIR models cannot approximate systems with poles located outside the open unit disk. In this paper we introduce a composite FIR/IIR (CFI) model, which is the product of a possibly noncausal FIR model and a pole-only IIR model. CFI models are shown to more accurately approximate systems with arbitrary poles under the condition that a lower bound on the number of poles located outside the open unit disk is known.**

I. Introduction

Since physical systems usually have infinite-impulse-response (IIR) dynamics, IIR models are widely used for system identification. However, fitting an IIR model requires a choice of the model order, and over-estimation or under-estimation of the true order of the system can lead to substantial parameter errors [1].

An alternative approach to IIR model fitting is to use models with finite impulse response (FIR) [2]–[5]. As their name implies, FIR models have the property that their impulse response converges to zero in a finite number of steps; accordingly, all minimal state space realizations of FIR models have nilpotent dynamics matrices. For an asymptotically stable IIR system, the numerator coefficients of an FIR model can be chosen to approximate the impulse response of an IIR system; this approximation improves as the order of the FIR model increases. The coefficients of the FIR model can then be also used with the Ho-Kalman realization [6], usually in the form of ERA [7] with nuclear norm optimization [8], to estimate the system order and construct an IIR model.

Unfortunately, FIR models often provide a poor fit to systems that are not asymptotically stable. For unstable plants operating in an asymptotically stable closed loop, noncausal FIR models can provide a useful approximation of the asymptotically stable and unstable components of the system [1], [9]. However, noncausal FIR models cannot approximate poles that lie on the unit circle.

To remedy this problem, the present paper considers a composite FIR/IIR (CFI) model $\hat{G}_{l,\mu}$, which is the product of an FIR model $\hat{G}_{\mathrm{F},\mu}$ and an IIR model $\hat{G}_{\mathrm{I},l}$. The rationale for the form of $\hat{G}_{l,\mu}$ is to let $\hat{G}_{\mathrm{I},l}$ capture the outermost poles of the system, that is, all of the poles that lie outside of the open unit disk. This includes all poles on the unit circle as well as all poles with modulus greater than 1. All remaining poles of the system, that is, all poles in the open unit disk, are captured by $\hat{G}_{\mathrm{F},\mu}$. This method requires knowledge of the number poles that lie outside the open unit disk, that is, the number of poles whose modulus is 1 or greater.

The goal of the present paper is to investigate the effectiveness of CFI models for system identification in the absence of knowledge of the system order but in the presence of knowledge of the number of poles located outside the open unit disk. The contribution of this paper is a numerical assessment of the accuracy of CFI models relative to models that are either entirely FIR or entirely IIR. In particular, numerical examples demonstrate the ability of $\hat{G}_{\mathrm{I},l}$ to capture the outermost l poles. In addition, we also investigate the effect of underestimating or overestimating l.

II. A Factorization of G

Throughout the paper, G is a proper nth-order rational transfer function with real coefficients. Let $\mathbf{q}$ denote the forward-shift operator, which is used to represent time-domain dynamics. Pole-zero analysis is facilitated by replacing $\mathbf{q}$ with the complex Z-transform variable $\mathbf{z}$. Although $G(\mathbf{q})$ and $G(\mathbf{z})$ have the same functional form, the former captures both the free and forced response, whereas the latter captures only the forced response.

We write G as

$$G(\mathbf{q}) = \frac{(\mathbf{q}-z_1)\cdots(\mathbf{q}-z_m)}{(\mathbf{q}-p_1)\cdots(\mathbf{q}-p_n)}, \tag{1}$$

where n is the order of G, $m \leq n$, $z_1, \ldots, z_m \in \mathbb{C}$ are the zeros of G, and $p_1, \ldots, p_n \in \mathbb{C}$ are the poles of G. Let $l \in \{0, \ldots, n\}$ be the number of poles of G located outside the open unit disk. Hence, for all $i = 1, \ldots, n-l$, $|p_i| < 1$, and, assuming $l \geq 1$, for all $i = n-l+1, \ldots, n$, $|p_i| \geq 1$. Then, define

$$D_{\mathrm{I},l}(\mathbf{q}) \triangleq (\mathbf{q}-p_{n-l+1})\cdots(\mathbf{q}-p_n) = \mathbf{q}^l + \sum_{i=1}^{l-1} c_i \mathbf{q}^i, \tag{2}$$

where $c_1, \ldots, c_l \in \mathbb{R}$. Then G can be written as

$$G(\mathbf{q}) = \frac{1}{D_{\mathrm{I},l}(\mathbf{q})} G_l(\mathbf{q}), \tag{3}$$

where

$$G_l(\mathbf{q}) \triangleq \frac{(\mathbf{q}-z_1)\cdots(\mathbf{q}-z_m)}{(\mathbf{q}-p_1)\cdots(\mathbf{q}-p_{n-l})}. \tag{4}$$

Note that all of the poles of G_l are in the open unit disk.

Let $\mathbb{P}(\rho) \triangleq \{z \in \mathbb{C} : |z| > \rho\}$ denote the open punctured plane centered at the origin with inner radius $\rho \geq 0$, and let

[1]Dept. Aero. Eng., Jordan Univ. Science and Tech., Irbid, Jordan 22110
[2]Dept. Aerospace Engineering, Univ. Michigan, Ann Arbor, MI 48109.

$\rho_{G_l} \triangleq \max\{|p_1|, \ldots, |p_{n-l}|\} < 1$ denote the spectral radius of G_l. Then, for all $z \in \mathbb{P}(\rho_{G_l})$, the Laurent expansion of G_l is given by

$$G_l(z) = \sum_{i=n-m-l}^{\infty} h_i z^{-i}, \tag{5}$$

where $n-m-l$ is the relative degree of G_l and, for all $i \geq n-m-l$, $h_i \in \mathbb{R}$. Using (3) and (5) implies that, for all $z \in \mathbb{P}(\rho_{G_l})$,

$$G(z) = \frac{1}{D_{\mathrm{I},l}(z)} \sum_{i=n-m-l}^{\infty} h_i z^{-i}. \tag{6}$$

Truncating the sum in (5) yields the truncated model

$$G_{l,\mu}(\mathbf{q}) \triangleq \frac{1}{D_{\mathrm{I},l}(\mathbf{q})} G_{\mathrm{F},l,\mu}(\mathbf{q}), \tag{7}$$

where the FIR truncation $G_{\mathrm{F},l,\mu}$ of G_l is defined by

$$G_{\mathrm{F},l,\mu}(\mathbf{q}) \triangleq \sum_{i=n-m-l}^{\mu} h_i \mathbf{q}^{-i}. \tag{8}$$

The following result provides necessary and sufficient conditions for the boundedness of the coefficients $(h_i)_{i=n-m-l}^{\infty}$ of the Laurent expansion of G_l given by (5).

Theorem 2.1: [1], [10] Consider the Laurent expansion of G_l in $\mathbb{P}(\rho_{G_l})$ given by (5). Then, the following statements are equivalent:

i) $\rho_{G_l} < 1$.
ii) $\sum_{i=n-m-l}^{\infty} h_i$ is finite.
iii) $\lim_{i\to\infty} h_i = 0$.
iv) $(h_i)_{i=n-m-l}^{\infty}$ is bounded.

Furthermore, for all $z \in \mathbb{P}(\rho_{G_l})$,

$$\lim_{\mu\to\infty} G_{\mathrm{F},l,\mu}(z) = G_l(z). \tag{9}$$

III. Composite FIR/IIR Model

The factorization (3) and the limit (9) suggest an approximate model with components that can separately capture the poles of G located inside and outside the open unit disk. We thus consider the CFI model

$$\hat{G}_{l,\mu}(\mathbf{q}) = \frac{1}{\hat{D}_{\mathrm{I},l,\mu}(\mathbf{q})} \hat{G}_{\mathrm{F},l,\mu}(\mathbf{q}), \tag{10}$$

where $\frac{1}{\hat{D}_{\mathrm{I},l,\mu}}$ is IIR and $\hat{G}_{\mathrm{F},l,\mu}$ is FIR. In particular, $\frac{1}{\hat{D}_{\mathrm{I},l,\mu}}$ is an IIR transfer function of order l and relative degree l, i.e.,

$$\hat{D}_{\mathrm{I},l,\mu}(\mathbf{q}) \triangleq \mathbf{q}^l + \sum_{i=1}^{l-1} \hat{c}_i \mathbf{q}^i, \tag{11}$$

and $\hat{G}_{\mathrm{F},l,\mu}$ is an FIR transfer function of the form

$$\hat{G}_{\mathrm{F},l,\mu}(\mathbf{q}) \triangleq \sum_{i=n-m-l}^{\mu} \hat{h}_i \mathbf{q}^{-i}, \tag{12}$$

where $\mu \geq 1$ is sufficiently large to provide a desired approximation of G. The notation $\hat{D}_{\mathrm{I},l,\mu}$ reflects the fact that the choice of μ affects the estimated coefficients $\hat{c}_i$. Note that $\hat{D}_{\mathrm{I},l,\mu}$ and $\hat{G}_{\mathrm{F},l,\mu}$ depend on l, and thus we assume that l is known. The effect of mismodeled l on the accuracy of the estimates of $\hat{D}_{\mathrm{I},l,\mu}$ and $\hat{G}_{\mathrm{F},l,\mu}$ is considered in later sections.

Let $H = N/D$ be a rational function, where N and D are polynomials. Then the relative degree of H is defined by $\operatorname{reldeg} H \triangleq \deg D - \deg N$. Note that H is proper if and only if $\operatorname{reldeg} H \geq 0$, whereas H is improper (and thus noncausal) if and only if $\operatorname{reldeg} H < 0$.

Note that the relative degree of $\hat{G}_{\mathrm{F},l,\mu}$ is $n-m-l$. Hence, $\hat{G}_{\mathrm{F},l,\mu}$ is improper and thus noncausal if and only if $n < m+l$. However, since the relative degree of $1/\hat{D}_{\mathrm{I},l,\mu}$ is l, it follows that the relative degree of $\hat{G}_{l,\mu}$ is $n-m-l+l = n-m$, and thus $\hat{G}_{l,\mu}$ is proper.

IV. Least Squares Identification

For all $k \geq 0$, let $u(k)$ and $y(k)$ be the input and output of G at step k, respectively, and let ℓ denote the size of the data window. Furthermore, let $\theta_{l,\mu} \in \mathbb{R}^{1\times(2l+\mu-n+m+1)}$ denote the vector of parameters of the truncated model $G_{l,\mu}$ defined by (7), where

$$\theta_{l,\mu} = \left[\begin{array}{cc} \theta_{c,l} & \theta_{h,\mu} \end{array}\right], \tag{13}$$

$$\theta_{c,l} = \left[\begin{array}{ccc} c_1 & \cdots & c_l \end{array}\right], \tag{14}$$

$$\theta_{h,\mu} = \left[\begin{array}{ccc} h_{n-m-l} & \cdots & h_\mu \end{array}\right]. \tag{15}$$

The least squares estimate $\hat{\theta}_{l,\mu,\ell}$ of $\theta_{l,\mu}$ is given by

$$\hat{\theta}_{l,\mu,\ell} = \arg\min_{\bar{\theta}_{l,\mu}} \left\| \Psi_{y,l,\ell} - \bar{\theta}_{l,\mu} \Phi_{l,\mu,\ell} \right\|_{\mathrm{F}}, \tag{16}$$

where $\bar{\theta}_{l,\mu} \in \mathbb{R}^{1\times(2l+\mu-n+m+1)}$ and the components of $\hat{\theta}_{l,\mu,\ell}$ are the coefficients of the CFI model (10)–(12) given by

$$\hat{\theta}_{l,\mu,\ell} = \left[\begin{array}{cc} \hat{\theta}_{c,l,\mu,\ell} & \hat{\theta}_{h,l,\mu,\ell} \end{array}\right], \tag{17}$$

$$\hat{\theta}_{c,l,\mu,\ell} = \left[\begin{array}{ccc} \hat{c}_1 & \cdots & \hat{c}_l \end{array}\right], \tag{18}$$

$$\hat{\theta}_{h,l,\mu,\ell} = \left[\begin{array}{ccc} \hat{h}_{n-m-l,\ell} & \cdots & \hat{h}_{\mu,\ell} \end{array}\right], \tag{19}$$

$$\Psi_{y,l,\ell} \triangleq \left[\begin{array}{ccc} y(\mu) & \cdots & y(\ell) \end{array}\right],$$

$$\Phi_{l,\mu,\ell} \triangleq \left[\begin{array}{c} \Phi_{y,l,\ell} \\ \Phi_{u,\mu,\ell} \end{array}\right],$$

$$\Phi_{y,l,\ell} \triangleq \left[\begin{array}{ccc} \phi_{y,l}(\mu) & \cdots & \phi_{y,l}(\ell) \end{array}\right],$$

$$\Phi_{u,\mu,\ell} \triangleq \left[\begin{array}{ccc} \phi_{u,\mu}(\mu) & \cdots & \phi_{u,\mu}(\ell) \end{array}\right],$$

$$\phi_{y,l}(k) = \left[\begin{array}{ccc} -y(k-1) & \cdots & -y(k-l) \end{array}\right]^{\mathrm{T}},$$

$$\phi_{u,\mu}(k) = \left[\begin{array}{ccc} u(k) & \cdots & u(k-\mu-l) \end{array}\right]^{\mathrm{T}}.$$

V. CFI Models and Poles of Largest Modulus

A key feature of CFI models is the fact that the IIR part of the CFI model captures the poles of largest modulus. To explain why this occurs, note that, since G_l is asymptotically stable, the Markov parameters of G_l decay, whereas the Markov parameters of $\frac{1}{D_{\mathrm{I},l}}$ are bounded away from zero. Furthermore, since G_l is asymptotically stable, it follows that the coefficients of the Laurent expansion of G_l about infinity converge. In addition, since the modulus of every pole of $\frac{1}{D_{\mathrm{I},l}}$ is at least 1, it follows that the coefficients of the Laurent expansion of $\frac{1}{D_{\mathrm{I},l}}$ about infinity do not converge.

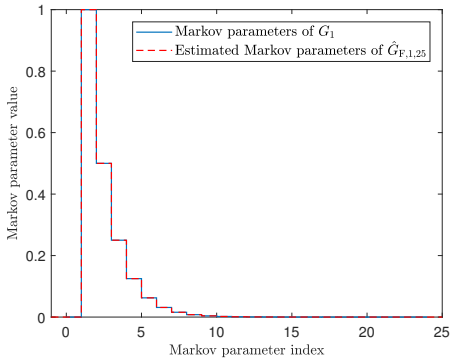

Fig. 1. Example 5.1. Markov parameters of $G_1(\mathbf{q}) = \frac{1}{\mathbf{q}-0.5}$ and the estimated Markov parameters of $\hat{G}_{\mathrm{F},1,25}$.

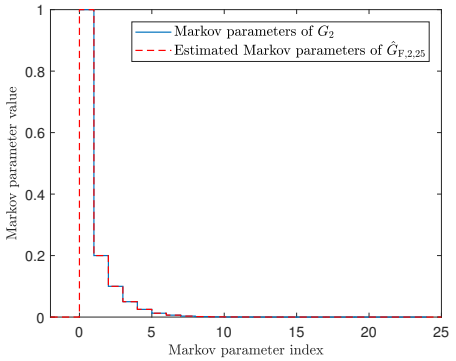

Fig. 2. Example 5.2. Markov parameters of $G_2(\mathbf{q}) = \frac{\mathbf{q}-0.3}{\mathbf{q}-0.5}$ and the estimated Markov parameters of $\hat{G}_{\mathrm{F},2,25}$.

Theorem 2.1 implies that the distance between G_l and the truncation of its Laurent expansion about infinity is much smaller than the distance between $\frac{1}{D_{\mathrm{I},l}}$ and the truncation of its Laurent expansion about infinity. The optimal solution of (16) thus places the poles with largest modulus in the IIR part of the estimated CFI model and places the remaining part of the model in the FIR part of the estimated CFI model. The following examples show that the IIR component $\frac{1}{D_{\mathrm{I},l}}$ of the CFI model captures the l poles of largest moduli.

Example 5.1: Consider the Lyapunov-stable

$$G(\mathbf{q}) = \frac{1}{(\mathbf{q}-1)(\mathbf{q}-0.5)}. \tag{20}$$

Let u be a realization of a stationary, Gaussian, white, zero-mean random process. We use noise-free measurements of u and y with a CFI model with orders $l = 1$ and $\mu = 25$ to identify (20). The estimated IIR component of the model is $\frac{1}{\hat{D}_{\mathrm{I},1}(\mathbf{q})} = \frac{1}{\mathbf{q}-1}$. Furthermore, Figure 1 shows that the estimated Markov parameters of the FIR component $\hat{G}_{\mathrm{F},1,25}$ of the model match the Markov parameters of $G_1(\mathbf{q}) = \frac{1}{\mathbf{q}-0.5}$.

Since the Markov parameters of $G_1(\mathbf{q}) = \frac{1}{\mathbf{q}-0.5}$ decay quickly and the Markov parameters of $\frac{1}{\mathbf{q}-1}$ are bounded away from zero, it follows that the distance between $\frac{1}{\mathbf{q}-0.5}$ and the truncation of its Laurent expansion about infinity is much smaller than the distance between $\frac{1}{\mathbf{q}-1}$ and the truncation of its Laurent expansion about infinity. Therefore, the optimal solution of (16) places the pole with largest modulus in the IIR part of the estimated CFI model and approximates the zeros and remaining poles of G by the FIR part of the estimated CFI model. ◇

Example 5.2: Consider the unstable transfer function

$$G(\mathbf{q}) = \frac{\mathbf{q}-0.3}{(\mathbf{q}-1)^2(\mathbf{q}-0.5)}. \tag{21}$$

Let u be a realization of a stationary, Gaussian, white, zero-mean random process. Assuming noise-free measurements of u and y, we use a CFI model with orders $l = 2$ and $\mu = 25$ to identify (21). The estimated IIR component of the model is $\frac{1}{\hat{D}_{\mathrm{I},l,\mu}(\mathbf{q})} = \frac{1}{(\mathbf{q}-1)^2}$. Figure 2 shows that the estimated Markov parameters of the FIR component $\hat{G}_{\mathrm{F},2,25}$ of the model match the Markov parameters of $G_2(\mathbf{q}) = \frac{\mathbf{q}-0.3}{\mathbf{q}-0.5}$.

Since the Markov parameters of $G_2(\mathbf{q}) = \frac{\mathbf{q}-0.3}{\mathbf{q}-0.5}$ decay quickly and the Markov parameters of $\frac{1}{(\mathbf{q}-1)^2}$ diverge, it follows that the distance between $G_2(\mathbf{q}) = \frac{\mathbf{q}-0.3}{\mathbf{q}-0.5}$ and the truncation of its Laurent expansion about infinity is much smaller than the distance between $\frac{1}{(\mathbf{q}-1)^2}$ and the truncation of its Laurent expansion about infinity. Therefore, the optimal solution of (16) places the poles with largest modulus in the IIR part of the estimated CFI model and approximates the zeros and remaining poles of G by the FIR part of the estimated CFI model. ◇

For the rest of the paper, all white noise signals used as either the input u or measurement noise added to u or y are assumed to be zero-mean Gaussian.

VI. Effect of μ

We now present numerical examples to investigate the effect of μ on the accuracy of the estimated CFI models.

Example 6.1: Consider the transfer function

$$G(\mathbf{q}) = \frac{\mathbf{q}-0.6}{(\mathbf{q}-1)(\mathbf{q}+1)(\mathbf{q}-0.3)}. \tag{22}$$

Then, $l = 2$ and $D_{\mathrm{I},l}(\mathbf{q}) = (\mathbf{q}-1)(\mathbf{q}+1) = \mathbf{q}^2 - 1$. Let u be white noise with unit variance. We add white input measurement noise to u with SNR 1000, 100, 10, and 1, where the output is noise free. We use $\ell = 10,000$ samples of u and y and a CFI model with $l = 2$ and $\mu = 1, \ldots, 50$. Figure 3 shows $\|\hat{D}_{\mathrm{I},l,\mu,\ell} - D_{\mathrm{I},l}\|_2$. Note from Figure 3 that increasing μ can help improve the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $D_{\mathrm{I},l}$. With noise-free data, the estimate $\hat{D}_{\mathrm{I},l,\mu}$ of D_{I_l} reaches machine precision for $\mu = 25$. Figure 4 shows $\|\hat{G}_{l,\mu,\ell} - G\|_2$. Figure 4 shows that increasing μ can help improve the estimate $\hat{G}_{l,\mu,\ell}$ of G. Figure 4 also shows that, with noise-free data, the estimate $\hat{G}_{l,\mu,\ell}$ of $G_{l,\mu}$ reaches machine precision for $\mu = 25$.

Next, we assume that u is noise free but add white output measurement noise to y with SNR 1000, 100, 10, and 1. We use least squares with a CFI model with $l = 2$ and $\mu = 1, \ldots, 50$ to identify (22). Figure 5 shows $\|\hat{D}_{\mathrm{I},l,\mu,\ell} - D_{\mathrm{I},l}\|_2$. Note from Figure 5 that increasing μ can help improve the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $D_{\mathrm{I},l}$. With noise-free data, the estimate $\hat{D}_{\mathrm{I},l,\mu}$ of D_{I_l} reaches machine precision for $\mu = 25$. Figure 6 shows $\|\hat{G}_{l,\mu,\ell} - G\|_2$. Figure 6 shows that increasing μ can

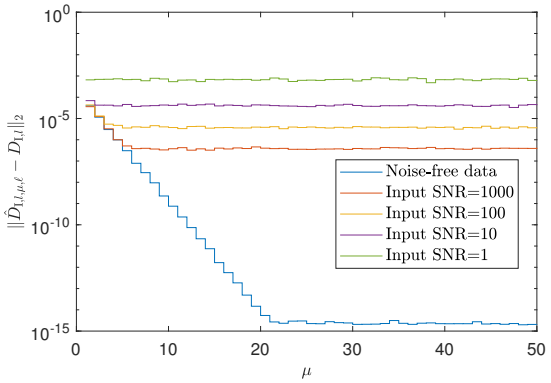

Fig. 3. Example 6.1. Effect of white output noise on the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $D_{\mathrm{I},l}$ for $l = 2$ and $\mu = 1, \ldots, 50$. Note that increasing μ can help improve the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $D_{\mathrm{I},l}$. For noise-free data, the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $D_{\mathrm{I},l}$ reaches machine precision for $\mu = 22$.

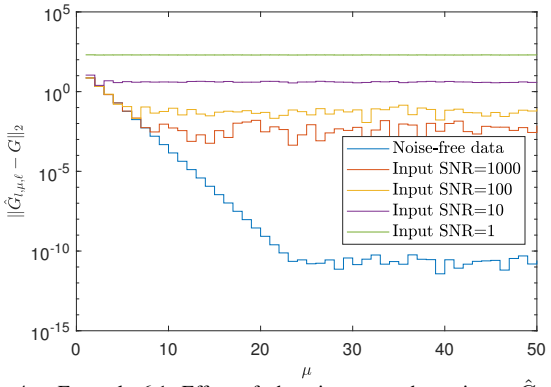

Fig. 4. Example 6.1. Effect of changing μ on the estimate $\hat{G}_{l,\mu,\ell}$ of G for $l = 2$ and $\mu = 1, \ldots, 50$ under white output measurement noise. Note that increasing μ can help improve the estimate $\hat{G}_{l,\mu,\ell}$ of G. For noise-free data, the estimate $\hat{G}_{l,\mu,\ell}$ of G reaches machine precision for $\mu = 25$.

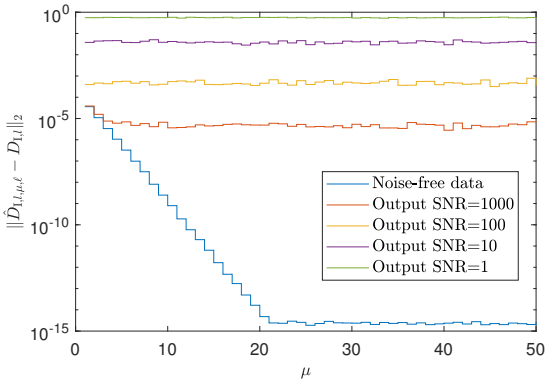

Fig. 5. Example 6.1. Effect of white output measurement noise on the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $D_{\mathrm{I},l}$ for $l = 2$ and $\mu = 1, \ldots, 50$. Note that increasing μ can help improve the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $D_{\mathrm{I},l}$. For noise-free data, the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $D_{\mathrm{I},l}$ reaches machine precision for $\mu = 22$.

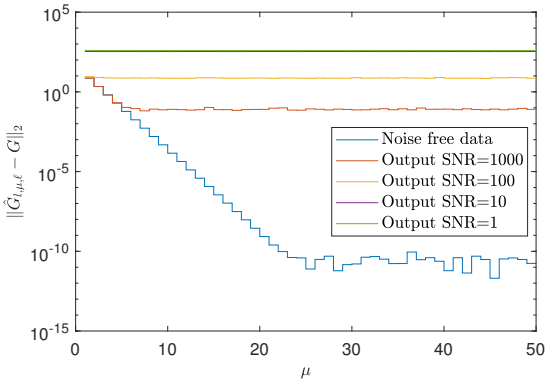

Fig. 6. Example 6.1. Effect of white input measurement noise on the estimate $\hat{G}_{l,\mu,\ell}$ of G for $l = 2$ and $\mu = 1, \ldots, 50$. Note that increasing μ can help improve the estimate $\hat{G}_{l,\mu,\ell}$ of G. For noise-free data, the estimate $\hat{G}_{l,\mu,\ell}$ of G reaches machine precision for $\mu = 25$.

help improve the estimate $\hat{G}_{l,\mu,\ell}$ of G. Figure 6 also shows that, with noise-free data, the estimate $\hat{G}_{l,\mu,\ell}$ of $G_{l,\mu}$ reaches machine precision for $\mu = 25$. ◇

VII. Effect of ℓ

We now present numerical examples to investigate the effect of ℓ on the accuracy of estimated CFI models.

Example 7.1: Consider the transfer function (22). Let u be white noise with unit variance. We add white input measurement noise to u with SNR 1000, 100, 10, and 1, and we use a CFI model with $l = 2$, $\mu = 30$, and $\ell = 10,000$ samples to identify G. Figure 7 shows $\|\hat{D}_{\mathrm{I},l,\mu,\ell} - D_{\mathrm{I},l}\|_2$, where ℓ ranges from 10^2 to 10^5 samples. Figure 7 shows that the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $\hat{D}_{\mathrm{I},l}$ is consistent. Moreover, Figure 8 shows $\|\hat{G}_{l,\mu,\ell} - G_{l,\mu}\|_2$ for $l = 2$ and $\mu = 30$, where ℓ ranges from 10^2 to 10^5 samples. Figure 8 shows that the estimate $\hat{G}_{l,\mu,\ell}$ of G is not consistent.

Next, we add white output measurement noise to y with SNR of 1000, 100, 10, and 1, and we use a CFI model with $l = 2$, $\mu = 30$, and $\ell = 10,000$ samples to identify G. Figure 9 shows $\|\hat{D}_{\mathrm{I},l,\mu,\ell} - D_{\mathrm{I},l}\|_2$, where ℓ ranges from 10^2 to 10^5 samples. Figure 9 shows that the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $D_{\mathrm{I},l}$ is not consistent. Figure 8 also shows $\|\hat{G}_{l,\mu,\ell} - G_{l,\mu}\|_2$ for $l = 2$ and $\mu = 30$, where ℓ ranges from 10^2 to 10^5 samples. Figure 8 shows that the estimate $\hat{G}_{l,\mu,\ell}$ of G is not consistent. ◇

VIII. Comparing CFI Models to IIR Models

In practice l may be unknown, and an estimate $\hat{l}$ of l may be used in solving the least squares problem (16). In this case, the estimate $\hat{\theta}_{l,\mu,\ell}$ is replaced by $\hat{\theta}_{\hat{l},\mu,\ell}$, where

$$\hat{\theta}_{\hat{l},\mu,\ell} = \begin{bmatrix} \hat{\theta}_{c,\hat{l},\mu,\ell} & \hat{\theta}_{h,\hat{l},\mu,\ell} \end{bmatrix}, \tag{23}$$

$$\hat{\theta}_{c,\hat{l},\mu,\ell} = \begin{bmatrix} \hat{c}_{1,\ell} & \cdots & \hat{c}_{\hat{l},\ell} \end{bmatrix}, \tag{24}$$

$$\hat{\theta}_{h,\hat{l},\mu,\ell} = \begin{bmatrix} \hat{h}_{n-m-\hat{l},\ell} & \cdots & \hat{h}_{\mu,\ell} \end{bmatrix}. \tag{25}$$

In this section, we present numerical examples to investigate the effect of unknown l on the estimates of CFI models. We also compare the accuracy of the estimates of CFI models with the accuracy of the estimates of IIR models under unknown l and n.

Example 8.1: Consider the 7th-order transfer function

$$G(\mathbf{q}) = \frac{(\mathbf{q}-0.8)(\mathbf{q}-0.3)(\mathbf{q}+0.5)(\mathbf{q}+0.3)(\mathbf{q}^2+0.16)}{(\mathbf{q}+0.6)(\mathbf{q}+0.7)(\mathbf{q}^2+0.25)(\mathbf{q}-1)(\mathbf{q}^2+1)}. \tag{26}$$

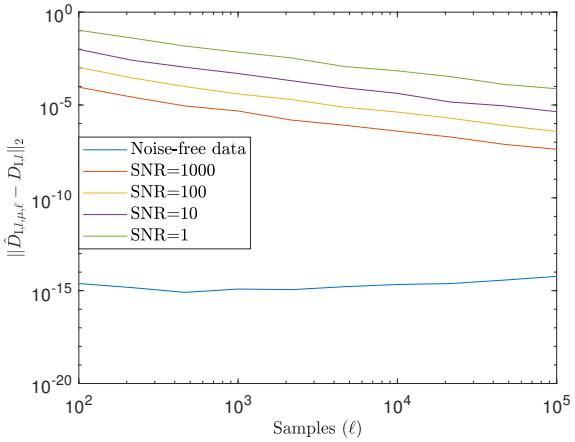

Fig. 7. Example 7.1. $\|\hat{D}_{\mathrm{I},l,\mu,\ell} - D_{\mathrm{I},l}\|_2$ for $l = 2$ and $\mu = 30$, for noise-free data and for noisy data with white input measurement noise with SNR of 1000, 100, 10, and 1, where the output is noise free. Note that the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $D_{\mathrm{I},l}$ is consistent.

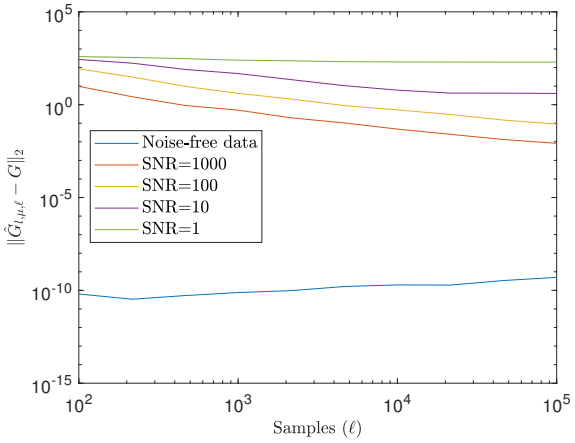

Fig. 8. Example 7.1. $\|\hat{G}_{l,\mu,\ell} - G_{\mu,l}\|_2$ for $l = 2$ and $\mu = 30$, for noise-free data and for noisy data with white input measurement noise with a SNR of 1000, 100, 10, 1, where the output is noise free. Note that the estimate $\hat{G}_{l,\mu,\ell}$ of $G_{\mu,l}$ is not consistent

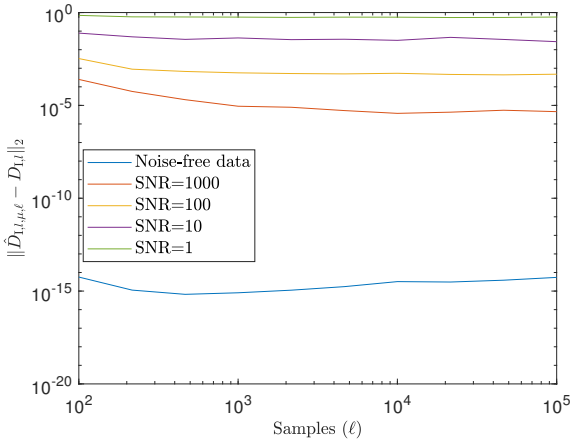

Fig. 9. Example 7.1. $\|\hat{D}_{\mathrm{I},l,\mu} - D_{\mathrm{I},l}\|_2$ for $l = 2$ and $\mu = 30$, for noise-free data and for noisy data with white output measurement noise with SNR 1000, 100, 10, and 1, where the input is noise free. Note that the estimate $\hat{D}_{\mathrm{I},l,\mu,\ell}$ of $D_{\mathrm{I},l}$ is not consistent.

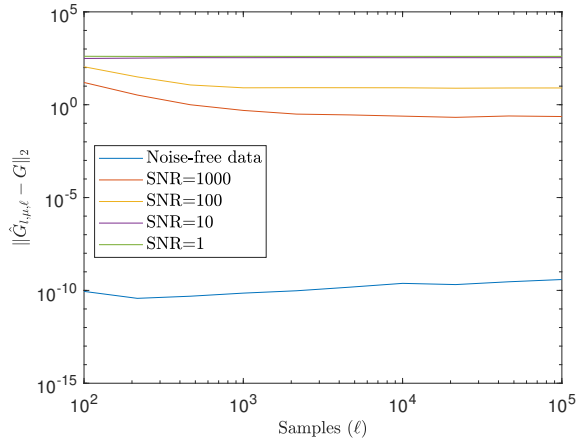

Fig. 10. Example 7.1. $\|\hat{G}_{l,\mu,\ell} - G_{\mu,l}\|_2$ for $l = 2$ and $\mu = 30$, for noise-free data and for noisy input data with white output measurement noise with SNR 1000, 100, 10, and 1, where the input is noise free. Note that the estimate $\hat{G}_{l,\mu,\ell}$ of $G_{\mu,l}$ is not consistent.

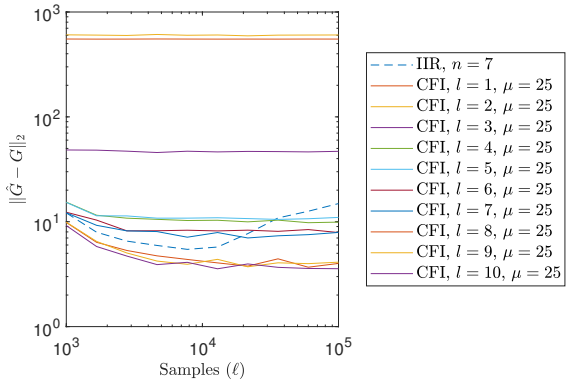

Fig. 11. Example 8.1. Distance between the frequency response of G and the frequency response of $\hat{G}$, where $\hat{G}$ is the estimated transfer function obtained using least squares with an IIR model of order $n = 7$ or a CFI model with $l = 1, \ldots, 10$ and $\mu = 25$. Note that CFI models with $l > 7$ give the most accurate estimate of G.

Let u be white noise with unit variance. We add white output measurement noise with SNR 100 to y, and we use $\ell = 10,000$ samples of u and y to identify G. We consider an IIR model of order $n = 7$, and a CFI model with $l = 1, \ldots, 10$ and $\mu = 25$. Figure 11 shows the distance between the frequency response of G and the frequency response of $\hat{G}$, where $\hat{G}$ is the estimated transfer function obtained using least squares with a CFI model or an IIR model. Figure 11 shows that the least squares estimates obtained with CFI models for $l > 7$ are more accurate than the least squares estimate obtained with an IIR model of order $n = 7$. $\diamond$

Example 8.2: Consider (26). Let u be white noise with unit variance. We add white output measurement noise with SNR 100 to y, and we use $\ell = 10,000$ samples of u and y to identify G. We consider a CFI model with order $l = 3$ and $\mu = 25$, as well as an IIR model with $n = 1, \ldots, 10$ and $\mu = 25$. Figure 12 shows the distance between the frequency response of G and the frequency response of $\hat{G}$, where $\hat{G}$ is the estimated transfer function obtained using least squares with either a CFI model or an IIR model. Note from Figure

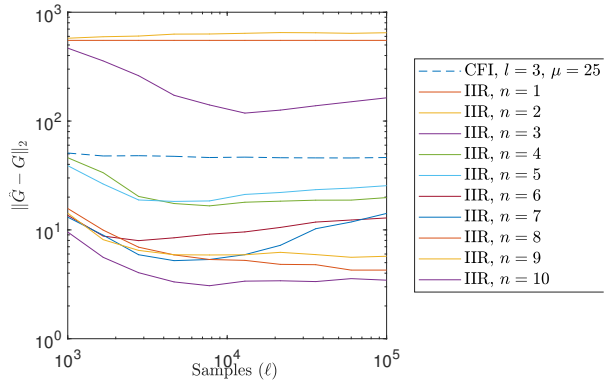

Fig. 12. Example 8.2. Distance between the frequency response of G and the frequency response of $\hat{G}$, where $\hat{G}$ is the estimated transfer function obtained using least squares with a CFI model with $l = 3$ and $\mu = 25$, or an IIR model with $n = 1, \ldots, 10$. Note that IIR models with $n > 3$ give more accurate estimates of G than the CFI model with $l = 3$ and $\mu = 25$.

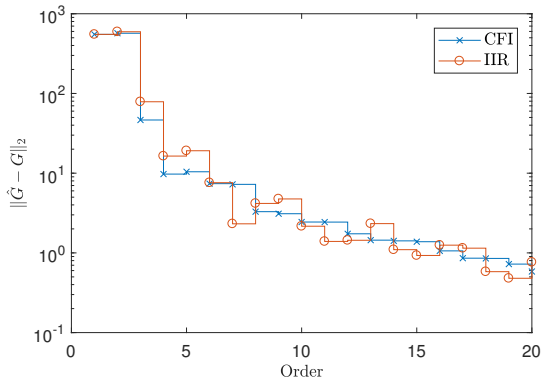

Fig. 13. Example 8.3. The distance between the frequency response of G and the frequency response of $\hat{G}$, where $\hat{G}$ is the transfer function estimate obtained using least squares with an IIR model of order $\hat{n} = 1, \ldots, 20$ and a CFI model with $\mu = 25$ and $\hat{l} = 1, \ldots, 20$. White input measurement noise with SNR 100 is added to y. Note that CFI models and IIR models give similar estimates of G.

12 that the least squares estimates obtained with IIR models of order $n > 3$ are more accurate than the least squares estimate obtained with a CFI model with $l = 3$. ◇

The following example compares IIR and CFI models for a range of values of n for IIR models and l for CFI models.

Example 8.3: Consider (26). Let u be white noise with unit variance. We add white output measurement noise with SNR 100 to y, and we use $\ell = 10,000$ samples of u and y to identify G. Figure 13 shows the distance between the frequency response of G and the frequency response of $\hat{G}$, where $\hat{G}$ is the transfer function estimate obtained using least squares with an IIR model of order $\hat{n} = 1, \ldots, 20$ and a CFI model with $\mu = 25$ and $\hat{l} = 1, \ldots, 20$. Note from Figure 13 that CFI models and IIR models give similar estimates of G.

Next, consider the case where white input measurement noise with SNR 100 is added to u, where y is noise free. Figure 14 shows $\|\hat{G} - G\|_2$, where $\hat{G}$ is the transfer function estimated using least squares with an IIR model with order $\hat{n} = 1, \ldots, 20$ or CFI model with $\mu = 25$ and $\hat{l} = 1, \ldots, 20$. Note from Figure 14 that CFI models give more accurate estimates of G. ◇

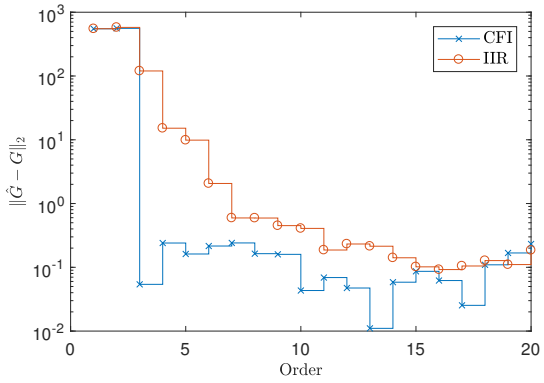

Fig. 14. Example 8.3. The distance between the frequency response of G and the frequency response of $\hat{G}$, where $\hat{G}$ is the estimate obtained using least squares with an IIR model of order $\hat{n} = 1, \ldots, 20$ or a CFI model with $\mu = 25$ and $\hat{l} = 1, \ldots, 20$. White input measurement noise with SNR 100 is added to u. Note that the CFI models give more accurate estimates of the transfer function than IIR models over a wide range of l.

IX. Conclusions

Using IIR models requires knowledge of the system order, which, if unknown, may yield poor parameter estimates. On the other hand, FIR models, which are truncations of the Laurent expansion of a transfer function in the annulus that contains the unit circle can approximate asymptotically stable systems without the need for knowledge of the system order. However, FIR models cannot approximate systems with poles on the unit circle. In this paper, we introduced composite FIR/IIR (CFI) models in order to approximate systems with or without poles on the unit circle and without the need to know the system order. We showed that the IIR component of the CFI model automatically estimates the poles with the largest moduli, where the FIR component of the CFI model estimates the remaining part of the transfer function.

References

[1] K. F. Aljanaideh and D. S. Bernstein, "Closed-loop identification of unstable systems using noncausal FIR models," *Int. J. Contr.*, vol. 90, no. 2, pp. 168–185, 2017.

[2] R. Diversi, "A bias-compensated identification approach for noisy FIR models," *IEEE Sig. Proc. Lett.*, vol. 15, pp. 325–328, 2008.

[3] V. Cerone, D. Piga, and D. Regruto, "Fixed-order FIR approximation of linear systems from quantized input and output data," *Sys. Contr. Lett.*, vol. 62, no. 12, pp. 1136–1142, 2013.

[4] L. Chai, J. Zhang, C. Zhang, and E. Mosca, "From IIR to FIR digital MIMO models: a constructive hankel norm approximation method," in *Proc. IEEE Conf. Dec. Contr.*, 2005, pp. 5893–5898.

[5] Y. Yamamoto, B. D. Anderson, M. Nagahara, and Y. Koyanagi, "Optimizing FIR approximation for discrete-time IIR filters," *IEEE Sig. Proc. Lett.*, vol. 10, no. 9, pp. 273–276, 2003.

[6] B. Ho and R. Kalman, "Efficient construction of linear state variable models from input/output functions," *Regelungstechnik*, vol. 14, pp. 545–548, 1966.

[7] J. N. Juang, *Applied System Identification*. Prentice-Hall, 1993.

[8] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.*, vol. 52, no. 3, pp. 471–501, 2010.

[9] U. Forssell and L. Ljung, "Closed-loop identification revisited," *Automatica*, vol. 35, no. 7, pp. 1215–1241, 1999.

[10] D. S. Bernstein, K. F. Aljanaideh, and A. E. Frazho, "Laurent series and ℓ^p sequences," *Amer. Math. Monthly*, vol. 123, no. 4, pp. 398–398, 2016.