

Retrospective-Cost-Based Model Refinement for System Emulation and Subsystem Identification

Alexey V. Morozov, Asad A. Ali, Anthony M. D'Amato, Aaron J. Ridley, Sunil L. Kukreja, and Dennis S. Bernstein

Abstract—We consider the problem of *data-based model refinement*, where we assume the availability of an initial model, which may incorporate both physical laws and empirical observations. With this initial model as a starting point, our goal is to use additional measurements to refine the model. In particular, components of the model that are poorly modeled can be updated, thereby resulting in a higher fidelity model. We consider two special cases, namely, system emulation and subsystem identification. In the former case, the main system is assumed to be uncertain and we seek an estimate of the unknown subsystem that allows the overall model to approximate the true system. In this case, there is no expectation that the constructed subsystem model approximates the unknown subsystem. In the latter case, we assume that the main system is accurately modeled and we seek an estimate of the unknown subsystem that approximates the unknown subsystem.

I. INTRODUCTION

In the present paper we consider the problem of *data-based model refinement*, where we assume the availability of an initial model, which may incorporate both physical laws and empirical observations. The components of the initial model may have varying degrees of fidelity, reflecting knowledge or ignorance of the relevant physics as well as the availability of data. With this initial model as a starting point, our goal is to use additional measurements to refine the model. In particular, we wish to update the components of the model that are poorly modeled, thereby resulting in a higher fidelity model [1–5].

System identification is typically concerned with the construction of a model of the entire system from measured inputs to measured outputs. In contrast, our goal is to identify only a subsystem of the model, where the remainder of the model is not modified. One motivation for this objective is to improve understanding of the physics of the poorly modeled subsystem despite its low accessibility. Here, accessibility refers to the availability of measurements or estimates of the inputs and outputs of the unknown subsystem. This lack of accessibility leads to a nonstandard system identification problem.

This work was supported in part by NASA GSRP, NASA grants NNX09AO55H and NNX0BA57A, and NSF grants CDI-1027192 and CNS-1035236

A. V. Morozov, A. A. Ali, A. M. D'Amato, and D. S. Bernstein are with the Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI, USA. {morozova, asadali, amdamoto, dsbaero}@umich.edu

A. J. Ridley is with the Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, MI, USA. ridley@umich.edu

S. L. Kukreja is with the Structural Dynamics Group at NASA Dryden Flight Research Center. sunil.l.kukreja@nasa.gov

The present paper goes beyond [1–5] in two ways. First, the model refinement algorithm described in Section II is based on the extension of the retrospective cost adaptive control (RCAC) algorithm described in [6]. The algorithm in [6] requires knowledge of a limited number of Markov parameters of the plant, and thus simplifies earlier versions of RCAC described in [7–9]. Therefore, the algorithm in [6] improves the model refinement technique described in [1, 4, 10]. Furthermore, the present paper encompasses multiple versions of the model refinement problem, including system emulation and subsystem identification. In the former case, we seek an estimate of the unknown subsystem that allows the overall model to approximate the true system. In this case, there is no expectation that the constructed subsystem model approximates the unknown subsystem. In contrast, in the latter case, we seek an estimate of the unknown subsystem that approximates the unknown subsystem.

II. PROBLEM FORMULATION

Consider the MIMO discrete-time main system

$$x(k+1) = Ax(k) + Bu(k) + D_1w(k), \quad (1)$$

$$y(k) = Cx(k), \quad (2)$$

$$y_0(k) = E_1x(k) + v(k), \quad (3)$$

where $x(k) \in \mathbb{R}^n$, $y(k) \in \mathbb{R}^{l_y}$, $y_0(k) \in \mathbb{R}^{l_{y_0}}$, $u(k) \in \mathbb{R}^{l_u}$, $w(k) \in \mathbb{R}^{l_w}$, and $k \geq 0$. The main system (1)–(3) is interconnected with the unknown subsystem modeled by

$$u(k) = G_s(q)y(k), \quad (4)$$

where q is the forward shift operator. The system (1)–(4) represents the true system. We assume that the excitation signal $w(k)$ is known. $v(k)$ denotes measurement noise.

Next, we assume a model of the main system of the form

$$\hat{x}(k+1) = \hat{A}\hat{x}(k) + \hat{B}\hat{u}(k) + \hat{D}_1w(k), \quad (5)$$

$$\hat{y}(k) = \hat{C}\hat{x}(k), \quad (6)$$

$$\hat{y}_0(k) = \hat{E}_1\hat{x}(k), \quad (7)$$

where $\hat{x}(k) \in \mathbb{R}^{\hat{n}}$, $\hat{y}(k) \in \mathbb{R}^{\hat{l}_y}$, $\hat{y}_0(k) \in \mathbb{R}^{\hat{l}_{y_0}}$, $\hat{u}(k) \in \mathbb{R}^{\hat{l}_u}$. The model of the main system is interconnected with the subsystem model $\hat{u}(k) = \hat{G}_s(q)\hat{y}(k)$.

The goal is to estimate a subsystem model $\hat{G}_s(q)$ that minimizes a cost function based on the performance variable

$$z(k) \triangleq \hat{y}_0(k) - y_0(k) \in \mathbb{R}^{l_z} \quad (9)$$

We estimate $\hat{G}_s(q)$ by retrospectively reconstructing the signal $\hat{u}(k)$ that minimizes the performance at the current time step. The reconstruction of $\hat{u}(k)$ uses minimal modeling

information about the true system (1)–(3), namely, a limited number of Markov parameters. We then use $\hat{u}(k)$ and $\hat{y}(k)$ to construct $\hat{G}_s(q)$. Figure 1 illustrates the model-refinement architecture, which includes system emulation and subsystem identification as special cases. Table II indicates the switch positions for various model-refinement architectures.

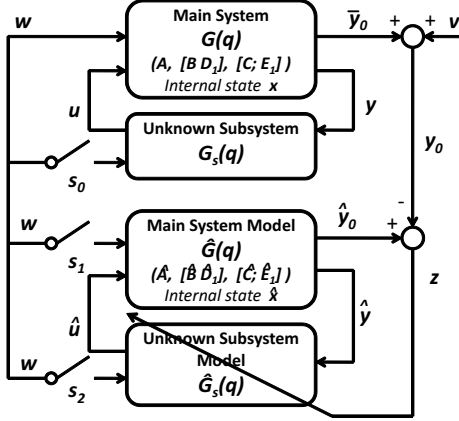


Fig. 1. Model-refinement architectures. The switches s_0 , s_1 and s_2 are used to define different architectures.

Case	s_0	s_1	s_2	Remarks
1	0	0	1	System emulation without subsystem excitation
2	0	1	0	System emulation without subsystem excitation. If A , B , C are known, this case is subsystem identification
3	1	0	1	System emulation with subsystem excitation
4	1	1	1	System emulation with subsystem excitation. If A , B , C are known, this case is subsystem identification

Table II. Switch positions for various model-refinement architectures. A switch in position 1 indicates the switch is closed, whereas a switch in position 0 indicates it is open.

The goal of system emulation is to determine a subsystem model $\hat{G}_s(q)$ such that the closed-loop frequency response of the true system (from w to y_0) matches the closed-loop frequency response of the system model (from w to \hat{y}_0). Since the matrices A, B, C are unknown, the matrices $\hat{A}, \hat{B}, \hat{C}$ in the main system model are approximations of A, B, C . The accuracy of this approximation determines how well the constructed subsystem model approximates the unknown subsystem. In the idealized case of subsystem identification, where A, B, C are known exactly, we set $\hat{A} = A$, $\hat{B} = B$, and $\hat{C} = C$ and use architectures 2 and 4 from Table II to obtain a subsystem model $\hat{G}_s(q)$ that approximates the unknown subsystem $G_s(q)$. However, the less stringent objective of system emulation is to obtain a model of the unknown subsystem such that the closed-loop model approximates the true closed-loop system.

III. RETROSPECTIVE SURROGATE-COST-BASED SIGNAL CONSTRUCTION

We begin by defining Markov parameters of the main system model $\hat{G}(q)$. For $i \geq 1$, let

$$H_i \triangleq \hat{E}_1 \hat{A}^{i-1} \hat{B}. \quad (10)$$

Therefore, $H_1 = \hat{E}_1 \hat{B}$ and $H_2 = \hat{E}_1 \hat{A} \hat{B}$. Let r be a positive integer. Then, for all $k \geq r$,

$$\hat{x}(k) = \hat{A}^r \hat{x}(k-r) + \sum_{i=1}^r \hat{A}^{i-1} \hat{B} \hat{u}(k-i) + \sum_{i=1}^r \hat{A}^{i-1} \hat{D}_1 w(k-i), \quad (11)$$

and thus

$$z(k) = \hat{E}_1 \hat{A}^r \hat{x}(k-r) + \sum_{i=1}^r \hat{E}_1 \hat{A}^{i-1} \hat{D}_1 w(k-i) - y_0(k) + \bar{\mathcal{H}} \bar{U}(k-1), \quad (12)$$

where $\bar{\mathcal{H}} \triangleq [H_1 \ \cdots \ H_r] \in \mathbb{R}^{l_z \times r l_{\hat{a}}}$, and

$$\bar{U}(k-1) \triangleq [\hat{u}^T(k-1) \ \cdots \ \hat{u}^T(k-r)]^T.$$

Next, we rearrange the columns of $\bar{\mathcal{H}}$ and the components of $\bar{U}(k-1)$ and partition the resulting matrix and vector so that

$$\bar{\mathcal{H}} \bar{U}(k-1) = \mathcal{H}' U'(k-1) + \mathcal{H} U(k-1), \quad (13)$$

where $\mathcal{H}' \in \mathbb{R}^{l_z \times (r l_{\hat{a}} - l_U)}$, $\mathcal{H} \in \mathbb{R}^{l_z \times l_U}$, $U'(k-1) \in \mathbb{R}^{r l_{\hat{a}} - l_U}$, and $U(k-1) \in \mathbb{R}^{l_U}$. Then, we can rewrite (12) as

$$z(k) = \mathcal{S}(k) + \mathcal{H} U(k-1), \quad (14)$$

where

$$\mathcal{S}(k) \triangleq \hat{E}_1 \hat{A}^r \hat{x}(k-r) + \sum_{i=1}^r \hat{E}_1 \hat{A}^{i-1} \hat{D}_1 w(k-i) - y_0(k) + \mathcal{H}' U'(k-1). \quad (15)$$

For example, $\bar{\mathcal{H}} = [H_1 \ H_2 \ H_3]$,

$$\mathcal{H}' = [H_1 \ H_2], \quad U'(k-1) = \begin{bmatrix} \hat{u}(k-1) \\ \hat{u}(k-2) \end{bmatrix},$$

and $\mathcal{H} = H_3$, $U(k-1) = \hat{u}(k-3)$. Next, we rewrite (14) with a delay of k_j time steps, where $0 \leq k_1 \leq k_2 \leq \cdots \leq k_s$, in the form

$$z(k - k_j) = \mathcal{S}_j(k - k_j) + \mathcal{H}_j U_j(k - k_j - 1), \quad (16)$$

where (15) becomes

$$\begin{aligned} \mathcal{S}_j(k - k_j) &\triangleq \hat{E}_1 \hat{A}^r \hat{x}(k - k_j - r) \\ &+ \sum_{i=1}^r \hat{E}_1 \hat{A}^{i-1} \hat{D}_1 w(k - k_j - i) - y_0(k - k_j) + \mathcal{H}'_j U'_j(k - k_j - 1) \end{aligned}$$

and (13) becomes

$$\bar{\mathcal{H}} \bar{U}(k - k_j - 1) = \mathcal{H}'_j U'_j(k - k_j - 1) + \mathcal{H}_j U_j(k - k_j - 1), \quad (17)$$

where $\mathcal{H}'_j \in \mathbb{R}^{l_z \times (r l_{\hat{a}} - l_{U_j})}$, $\mathcal{H}_j \in \mathbb{R}^{l_z \times l_{U_j}}$, $U'_j(k - k_j - 1) \in \mathbb{R}^{r l_{\hat{a}} - l_{U_j}}$, and $U_j(k - k_j - 1) \in \mathbb{R}^{l_{U_j}}$. Now, by stacking $z(k - k_1), \dots, z(k - k_s)$, we define the *extended performance*

$$Z(k) \triangleq [z^T(k - k_1) \ \cdots \ z^T(k - k_s)]^T \in \mathbb{R}^{s l_z}. \quad (18)$$

Therefore,

$$Z(k) \triangleq \tilde{\mathcal{S}}(k) + \tilde{\mathcal{H}} \tilde{U}(k-1), \quad (19)$$

where $\tilde{S}(k) \triangleq [\mathcal{S}^T(k-k_1) \cdots \mathcal{S}^T(k-k_s)]^T \in \mathbb{R}^{sl_z}$, $\tilde{\mathcal{H}} \in \mathbb{R}^{sl_z \times l_{\tilde{v}}}$, and $\tilde{U}(k-1) \in \mathbb{R}^{l_{\tilde{v}}}$. The vector $\tilde{U}(k-1)$ is formed by stacking $U_1(k-k_1-1), \dots, U_s(k-k_s-1)$ and removing repetitions of components. For example, with $k_1 = 0$ and $k_2 = 1$, stacking $U_1(k-1) = \begin{bmatrix} \hat{u}(k-1) \\ \hat{u}(k-2) \end{bmatrix}$ and $U_2(k-2) = \hat{u}(k-2)$ results in $\tilde{U}(k-1) = \begin{bmatrix} \hat{u}(k-1) \\ \hat{u}(k-2) \end{bmatrix}$. The coefficient matrix $\tilde{\mathcal{H}}$ consists of the entries of $\mathcal{H}_1, \dots, \mathcal{H}_s$ arranged according to the structure of $\tilde{U}(k-1)$. Furthermore, we assume that the last entry of $\tilde{U}(k-1)$ is a component of $\hat{u}(k-r)$.

Next, we define the *surrogate performance*

$$\hat{z}(k-k_j) \triangleq \mathcal{S}_j(k-k_j) + \mathcal{H}_j U_j^*(k-k_j-1), \quad (20)$$

where the actual past subsystem outputs $U_j(k-k_j-1)$ in (16) are replaced by the surrogate subsystem outputs $U_j^*(k-k_j-1)$. The *extended surrogate performance* for (20), which is defined as

$$\hat{Z}(k) \triangleq [\hat{z}^T(k-k_1) \cdots \hat{z}^T(k-k_s)]^T \in \mathbb{R}^{sl_z}, \quad (21)$$

is given by

$$\hat{Z}(k) = \tilde{S}(k) + \tilde{\mathcal{H}}\tilde{U}^*(k-1), \quad (22)$$

where the components of $\tilde{U}^*(k-1) \in \mathbb{R}^{l_{\tilde{v}}}$ are components of $U_1^*(k-k_1-1), \dots, U_s^*(k-k_s-1)$ ordered in the same way as the components of $\tilde{U}(k-1)$. Subtracting (19) from (22) yields

$$\hat{Z}(k) = Z(k) - \tilde{\mathcal{H}}\tilde{U}(k-1) + \tilde{\mathcal{H}}\tilde{U}^*(k-1). \quad (23)$$

Finally, we define the *retrospective cost function*

$$J(\tilde{U}^*(k-1), k) \triangleq \hat{Z}^T(k)R(k)\hat{Z}(k), \quad (24)$$

where $R(k) \in \mathbb{R}^{sl_z \times sl_z}$ is a positive-definite performance weighting. The goal is to determine refined subsystem outputs $\tilde{U}^*(k-1)$ that would have provided better performance than the subsystem outputs $U(k)$ that were applied to the system. The refined subsystem outputs values $\tilde{U}^*(k-1)$ are subsequently used to update the subsystem estimate.

A. Cost Function Optimization with Adaptive Regularization

To ensure that (24) has a global minimizer, we consider the regularized cost

$$\bar{J}(\tilde{U}^*(k-1), k) \triangleq \hat{Z}^T(k)R(k)\hat{Z}(k) + \eta(k)\tilde{U}^{*\top}(k-1)\tilde{U}^*(k-1), \quad (25)$$

where $\eta(k) = \bar{\eta}z^T(k)z(k)$ and $\bar{\eta} \geq 0$. Substituting (23) into (25) yields

$$\bar{J}(\tilde{U}^*(k-1), k) = \tilde{U}^*(k-1)^T \mathcal{A}(k) \tilde{U}^*(k-1) + \mathcal{B}(k)\tilde{U}^*(k-1) + \mathcal{C}(k), \quad (26)$$

where

$$\mathcal{A}(k) \triangleq \tilde{\mathcal{H}}^T R(k) \tilde{\mathcal{H}} + \eta(k)I_{l_{\tilde{v}}}, \quad (27)$$

$$\mathcal{B}(k) \triangleq 2\tilde{\mathcal{H}}^T R(k)[Z(k) - \tilde{\mathcal{H}}\tilde{U}(k-1)], \quad (28)$$

$$\mathcal{C}(k) \triangleq Z^T(k)R(k)Z(k) - 2Z^T(k)R(k)\tilde{\mathcal{H}}\tilde{U}(k-1) + \tilde{U}^T(k-1)\tilde{\mathcal{H}}^T R(k)\tilde{\mathcal{H}}\tilde{U}(k-1). \quad (29)$$

If either $\tilde{\mathcal{H}}$ has full column rank or $\eta(k) > 0$, then $\mathcal{A}(k)$ is positive definite. In this case, $\bar{J}(\tilde{U}^*(k-1), k)$ has the unique global minimizer

$$\tilde{U}^*(k-1) = -\frac{1}{2}\mathcal{A}^{-1}(k)\mathcal{B}(k). \quad (30)$$

B. Subsystem Modeling

The subsystem output $\hat{u}(k)$ is given by the exactly proper time-series model of order n_c given by

$$\hat{u}(k) = \sum_{i=1}^{n_c} M_i(k)\hat{u}(k-i) + \sum_{i=0}^{n_c} N_i(k)\hat{y}(k-i) + \sum_{i=0}^{n_c} O_i(k)w(k-i), \quad (31)$$

where, for all $i = 1, \dots, n_c$, $M_i(k) \in \mathbb{R}^{l_{\hat{u}} \times l_{\hat{u}}}$, $N_i(k) \in \mathbb{R}^{l_{\hat{u}} \times l_{\hat{y}}}$ and $O_i(k) \in \mathbb{R}^{l_{\hat{u}} \times l_w}$. The subsystem output (31) can be expressed as $\hat{u}(k) = \theta(k)\phi(k-1)$, where $\theta(k) \in \mathbb{R}^{l_{\hat{u}} \times n_c(l_{\hat{u}} + l_{\hat{y}} + l_w)}$ is

$$\theta(k) \triangleq [M_1(k) \cdots M_{n_c}(k) N_1(k) \cdots N_{n_c}(k) O_1(k) \cdots O_{n_c}(k)], \text{ and} \\ \phi(k-1) \triangleq [\hat{u}^T(k-1) \cdots \hat{u}^T(k-n_c) \hat{y}^T(k-1) \cdots \hat{y}^T(k-n_c) \hat{w}^T(k-1) \cdots \hat{w}^T(k-n_c)]^T \\ \in \mathbb{R}^{n_c(l_{\hat{u}} + l_{\hat{y}} + l_w)}. \quad (32)$$

Note if $s_2 = 0$ then $w(k)$ and O_i are removed from $\hat{u}(k)$, $\theta(k)$, and $\phi(k-1)$.

C. Recursive Least Squares Update

Let d be a positive integer such that $\tilde{U}^*(k-1)$ contains $u^*(k-d)$. We define the cumulative cost function

$$J_R(\theta(k)) \triangleq \sum_{i=1}^k \lambda^{k-i} \|u^*(k-d) - \phi^T(k-d-1)\theta^T(k)\|^2,$$

where $\phi(k-d)$ is given by (32) and $\lambda(k) \in (0, 1]$ is the forgetting factor. Minimizing the cumulative cost function yields retrospective cost optimization (RCO)

$$\theta^T(k) = \theta^T(k-1) + P(k-1)\phi(k-d-1) \cdot [\phi^T(k-d)P(k-1)\phi(k-d-1) + \lambda(k)]^{-1} \cdot (u^*(k-d) - \phi^T(k-d-1)\theta^T(k-1)). \quad (33)$$

The error covariance is updated by

$$P(k) = \lambda^{-1}(k)P(k-1) - \lambda^{-1}(k)P(k-1)\phi(k-d-1) \cdot [\phi^T(k-d-1)P(k-1)\phi(k-d) + \lambda(k)]^{-1} \cdot \phi^T(k-d-1)P(k-1). \quad (34)$$

We initialize the error covariance matrix as $P(0) = \beta I$, where $\beta > 0$.

IV. NUMERICAL EXAMPLES

We now consider numerical examples with various model-refinement architectures to illustrate the effect of noise and model uncertainty on the emulation of the closed-loop system and, where applicable, the identification of the unknown subsystem. For all examples in this section, RCO is turned on after 100 steps. The level of measurement noise varies for each example, where $v = \mathcal{N}(\mu_v, \sigma_v^2)$ means that the

output noise signal v is Gaussian white noise with mean μ_v and variance σ_v^2 . We define $\text{SNR} \triangleq \frac{\sigma_{\bar{y}_0}^2}{\sigma_v^2}$, where $\sigma_{\bar{y}_0}^2$ is the variance of the output signal \bar{y}_0 . The case number in each example refers to the positions of the switches in Figure 1 as described in Table II. For all examples, the subsystem model parameters $\theta(k)$ are initialized at zero. For convenience, let $G(q)$ represent the main system, let $G_{\text{cl}}(q)$ represent the closed-loop system from w to \bar{y}_0 , and let $G_s(q)$ represent the unknown subsystem.

We consider the spring-mass-damper system shown in Figure 2. For $i = 1, 2, 3$, let q_i be the position of i^{th} mass, and let m_i be the mass of the i^{th} block. For $i = 1, 2, 3, 4$, let k_i be the stiffness of the i^{th} spring, and let c_i be the damping coefficient of the i^{th} damper. Finally, let w be the force applied to the second block.

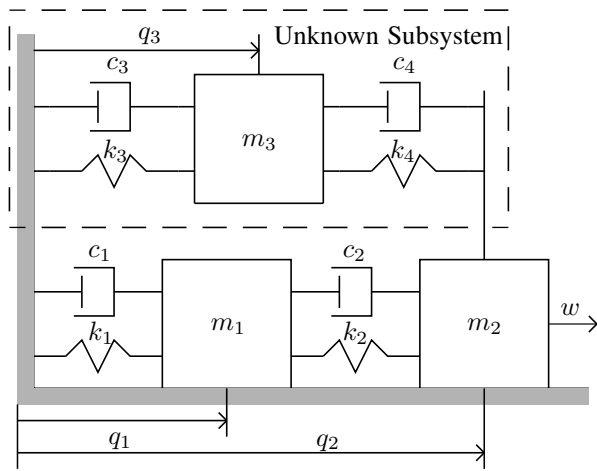


Fig. 2. Spring-mass-damper system with main system and unknown subsystem.

The discretized equations of motion of the main system are

$$x(k+1) = Ax(k) + Bu(k) + D_1w(k), \quad (35)$$

$$y(k) = Cx(k), \quad (36)$$

$$y_0(k) = E_1x(k) + \nu(k), \quad (37)$$

where $x(k) = [q_1(k) \ q_2(k) \ q_4(k) \ q_5(k)]^T$,

$$A = \begin{bmatrix} 1 & 0 & T_s & 0 \\ 0 & 1 & 0 & T_s \\ -\frac{T_s(k_1+k_2)}{m_1} & \frac{T_s k_2}{m_1} & 1 - \frac{T_s(c_1+c_2)}{m_1} & \frac{T_s c_2}{m_1} \\ \frac{T_s k_2}{m_2} & -\frac{T_s(k_2+k_4)}{m_2} & \frac{T_s c_2}{m_2} & 1 - \frac{T_s(c_2+c_4)}{m_2} \end{bmatrix},$$

$$D_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{T_s}{m_2} \end{bmatrix}, C = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T, E_1 = \begin{bmatrix} 0 \\ \frac{T_s k_4}{m_3} \\ 0 \\ \frac{T_s c_4}{m_3} \end{bmatrix}^T.$$

The discretized equations of motion of the unknown subsystem are

$$x_s(k+1) = A_s x_s(k) + B_s y(k), \quad (38)$$

$$u(k) = C_s x_s(k), \quad (39)$$

where

$$x_s(k) = \begin{bmatrix} q_3(k) \\ q_6(k) \end{bmatrix}, A_s = \begin{bmatrix} 1 & T_s \\ -\frac{T_s(k_3+k_4)}{m_3} & 1 - \frac{T_s(c_3+c_4)}{m_3} \end{bmatrix},$$

$$B_s = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C_s = \begin{bmatrix} \frac{T_s k_4}{m_2} \\ \frac{T_s c_4}{m_2} \end{bmatrix}.$$

Furthermore, $T_s = 0.25$, $m_1 = 4$, $m_2 = 2$, $m_3 = 10$, $k_1 = 12$, $k_2 = 2$, $k_3 = 4$, $k_4 = 6$, $c_1 = 4$, $c_2 = 2$, $c_3 = 5$, and $c_4 = 3$.

Example 4.1: (Case 1, A, B, C unknown, $\text{SNR} = 100$). Since A , B , and C are unknown, we choose \hat{A} , \hat{B} , and \hat{C} such that $\hat{G}(q)$ is stable and minimum phase, but otherwise arbitrarily. More specifically, we choose

$$\hat{A} = \begin{bmatrix} -0.039 & -0.029 \\ 0.023 & 0.0023 \end{bmatrix}, \hat{B} = \begin{bmatrix} 0.003 \\ 0.098 \end{bmatrix}, \quad (40)$$

$$\hat{C} = [2.5 \quad -0.78].$$

Moreover, $\mu_w = 0$ and $\sigma_w^2 = 5$. For this example we take $n_c = 20$, $\bar{\eta} = 0$, $\beta = 0.01$, and $\tilde{\mathcal{H}} = H_1$, which is the first Markov parameter of $\hat{G}(q)$. The parameters of the model refinement algorithm are chosen such that $z(k)$ is minimized. Figure 3 shows that the estimated frequency response of the closed-loop system $\hat{G}_{\text{cl}}(q)$ approximates the closed-loop frequency response of the true system $G_{\text{cl}}(q)$. Next, we run this example with three different SNR values for 5000 time steps. Figure 4 shows that, as the SNR increases, the frequency response of \hat{G}_{cl} provides an improved approximation of the frequency response of G_{cl} .

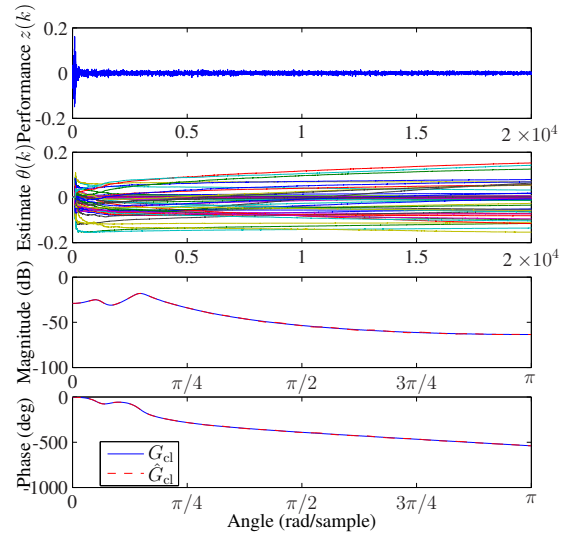


Fig. 3. The RCO algorithm is turned on at $k = 100$ steps. The closed-loop frequency response of \hat{G}_{cl} is indistinguishable from the frequency response of G_{cl} .

Example 4.2: (Case 3, A, B, C unknown, $\text{SNR} = 100$). The architecture for this example is different from the architecture of Case 1 only in that the unknown subsystem

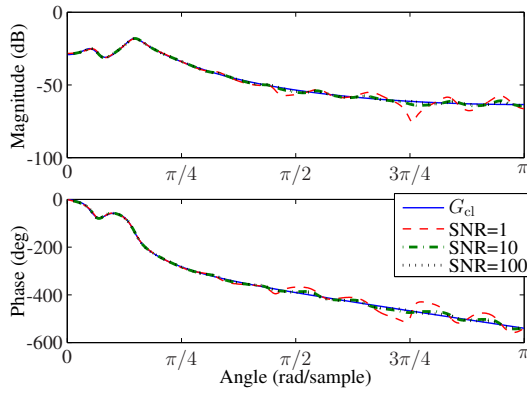


Fig. 4. As the SNR increases, the accuracy of the frequency response of \hat{G}_{cl} improves.

$G_s(q)$ has the additional input w , and hence

$$x_s(k+1) = A_s x_s(k) + B_s [y(k) \ w(k)]^T,$$

where,

$$B_s = \begin{bmatrix} 0 & 0 \\ 1 & T_s/m_3 \end{bmatrix}.$$

Furthermore, we let $\mu_w = 0$ and $\sigma_w^2 = 5$. Since A, B, C are unknown, we choose $\hat{A}, \hat{B}, \hat{C}$ as in (40). For this example we take $n_c = 20$, $\bar{\eta} = 0$, $\beta = 0.01$, and $\tilde{H} = H_1$. Figure 5 shows that, as the SNR increases, the accuracy of the frequency response of \hat{G}_{cl} improves.

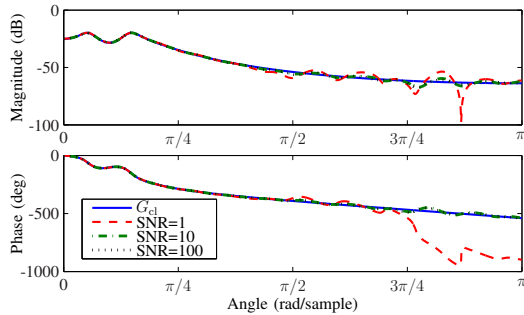


Fig. 5. As the SNR increases, the frequency response of \hat{G}_{cl} more closely approximates the frequency response of G_{cl} .

Example 4.3: (Case 2, A, B and C known). First, we investigate the effect of the amount of data on the identification of G_{cl} and G_s using Case 2 architecture when A, B , and C are known. For this example, $\mu_w = 0$, $\sigma_w^2 = 10$ and there is no noise. Furthermore, we let $n_c = 12$, $\bar{\eta} = 0$, $\beta = 0.01$, and $\tilde{H} = H_3$. Figures 6 and 7 show that as the amount of data increases, the accuracy of the frequency responses of \hat{G}_{cl} and \hat{G}_s improve. Note that the frequency response of \hat{G}_s cannot approximate G_s above 0.75 radians/sample because the transfer function that multiplies G_s in G_{cl} rolls off above this frequency. Next, we investigate the effect of SNR on Case 2 architecture when A, B , and C are known. The parameters are the same as in the previous example. Figures 8 and 9 show that as the SNR increases, the accuracy of the frequency responses of \hat{G}_{cl} and \hat{G}_s improve.

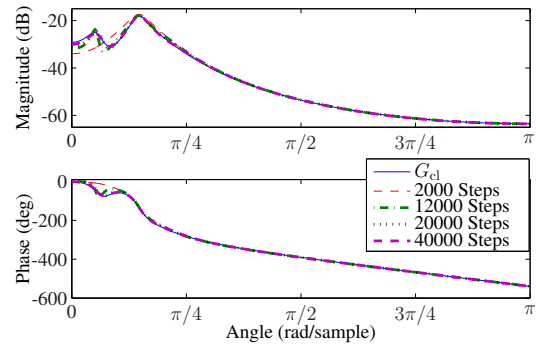


Fig. 6. As the amount of data increases, the frequency response of \hat{G}_{cl} more closely approximates the frequency response of G_{cl} .

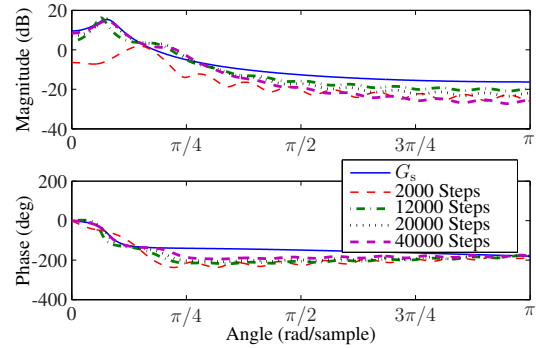


Fig. 7. As the amount of data increases, the frequency response of \hat{G}_s more closely approximates the frequency response of G_s .

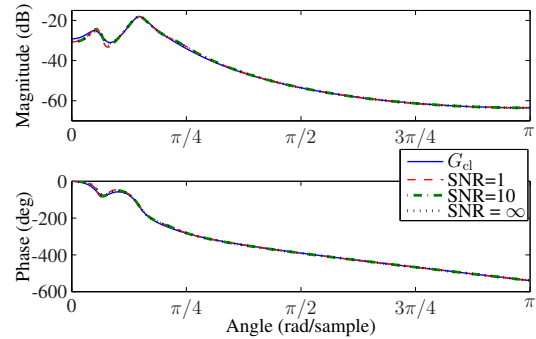


Fig. 8. As the SNR increases, the frequency response of \hat{G}_{cl} more closely approximates the frequency response of G_{cl} .

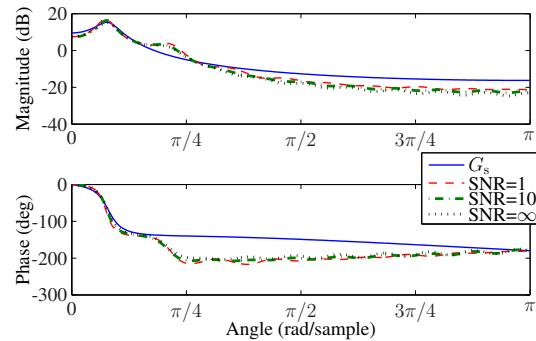


Fig. 9. As the SNR increases, the frequency response of \hat{G}_s more closely approximates the frequency response of G_s .

Example 4.4: (Case 2, A uncertain, B and C known, SNR=100). In this example we investigate the effect of uncertainty in A . Uncertainty in A is introduced by scaling the damping coefficient c_2 by an unknown scale factor α . Thus, \hat{A} is obtained by replacing c_2 in A by αc_2 . For this example, $\mu_w = 0$ and $\sigma_w^2 = 5$. Furthermore, we let $n_c = 12$, $\bar{\eta} = 0$, $\beta = 0.01$, and $\tilde{\mathcal{H}} = H_3$. Figures 10 and 11 show that as the uncertainty in A decreases (that is, α approaches 1), the frequency responses of \hat{G}_{cl} and \hat{G}_s more closely approximate the frequency responses of G_{cl} and G_s , respectively.

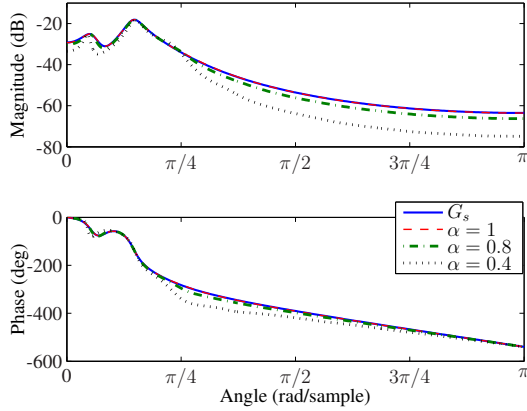


Fig. 10. Estimate of G_{cl} with uncertain $\hat{c}_2 = \alpha c_2$. As α approaches 1, the frequency response of \hat{G}_{cl} more closely approximates the frequency response of G_{cl} .

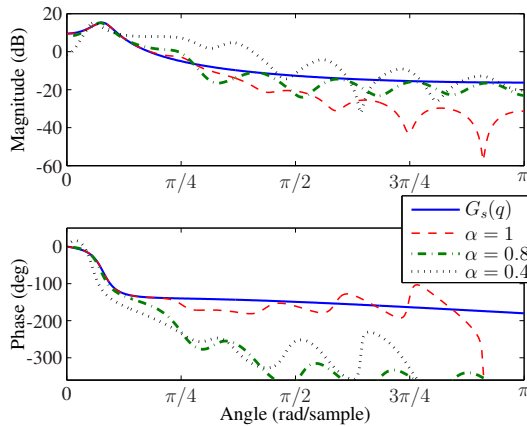


Fig. 11. Estimate of G_s with uncertain $\hat{c}_2 = \alpha c_2$. As α approaches 1, the frequency response of \hat{G}_s more closely approximates the frequency response of G_s .

V. CONCLUSIONS

This paper focused on the problem of model refinement, where data are used to improve the accuracy of a subsystem model connected by feedback to a given main system model. In particular, the objective is system emulation, where the goal is to estimate a subsystem model in order to provide a combined system model that has improved accuracy relative to the main system alone. The inputs and outputs of the unknown subsystem are not assumed to be accessible, and thus standard system identification techniques are

not applicable. We applied retrospective cost optimization, which reconstructs the input to the main system from the unknown subsystem. The main system may be well known or uncertain. In the latter case, there is no expectation that the estimated subsystem model approximates the unknown subsystem. However, if the main system is known exactly, then the estimated subsystem may provide a useful estimate of the unknown subsystem. Several numerical examples were used to illustrate the approach. The performance of the algorithm was assessed in terms of the closeness of the frequency response plots. The ultimate goal of this work is to provide a tool that engineers and scientists can use to improve the accuracy of large-scale models and estimate unknown subsystems that are difficult to model due to the inaccessibility of their inputs and outputs.

REFERENCES

- [1] H. Palanhandalam-Madapusi, E. L. Renk, and D. S. Bernstein, "Data-based model refinement for linear and Hammerstein systems using subspace identification and adaptive disturbance rejection," in *Proc. Conf. Contr. Appl.*, Toronto, Canada, August 2005, pp. 1630–1635.
- [2] M. A. Santillo, A. M. D'Amato, and D. S. Bernstein, "System identification using a retrospective correction filter for adaptive feedback model updating," in *Proc. Amer. Contr. Conf.*, St. Louis, MO, June 2009, pp. 4392–4397.
- [3] A. M. D'Amato and D. S. Bernstein, "Linear fractional transformation identification using retrospective cost optimization," in *Proc. SYSID*, Saint-Malo, France, July 2009, pp. 450–455.
- [4] A. M. D'Amato, B. J. Arritt, J. A. Banik, E. V. Ardelean, and D. S. Bernstein, "Structural health determination and model refinement for a deployable composite boom," in *AIAA SDM Conf.*, Palm Springs, CA, April 2009, AIAA-2009-2373.
- [5] A. M. D'Amato, A. R. Wu, K. S. Mitchell, S. L. Kukreja, and D. S. Bernstein, "Damage localization for structural health monitoring using retrospective cost model refinement," in *AIAA SDM Conf.*, Orlando, FL, April 2010, AIAA-2010-2628.
- [6] A. M. D'Amato, E. D. Sumer, and D. S. Bernstein, "Retrospective cost adaptive control for systems with unknown nonminimum-phase zeros," in *AIAA Guid. Nav. Contr. Conf.*, Portland, OR, August 2011, AIAA-2009-2373.
- [7] J. B. Hoagg, M. A. Santillo, and D. S. Bernstein, "Discrete-time adaptive command following and disturbance rejection with unknown exogenous dynamics," *IEEE Trans. Autom. Contr.*, vol. 53, pp. 912–928, 2008.
- [8] J. B. Hoagg and D. S. Bernstein, "Retrospective cost adaptive control for nonminimum-phase discrete-time systems part 1: The ideal controller and error system; part 2: The adaptive controller and stability analysis," in *Proc. Conf. Dec. Contr.*, Atlanta, GA, 2010.
- [9] —, "Retrospective cost model reference adaptive control for nonminimum-phase discrete-time systems, part 1: The ideal controller and error system; part 2: The adaptive controller and stability analysis," *Proc. Conf. Dec. Contr.*, pp. 2927–2932, 2011.
- [10] A. M. D'Amato, A. J. Ridley, and D. S. Bernstein, "Retrospective-cost-based adaptive model refinement for the ionosphere and thermosphere," *Statistical Analysis and Data Mining*, vol. 4, pp. 446–458, 2011.