2018 IEEE Conference on Decision and Control (CDC)
Miami Beach, FL, USA, Dec. 17-19, 2018

Data-Driven Parameter Estimation for Models with Nonlinear Parameter Dependence

Ankit Goel and Dennis S. Bernstein

Abstract— **Many models have known structure but unknown parameters. Nonlinear estimation methods, such as the extended Kalman filter (EKF), unscented Kalman filter (UKF), and ensemble Kalman filter (EnKF) are typically applied to these problems by viewing the unknown parameters as constant states. An alternative approach is provided by retrospective cost model refinement (RCMR), which uses an error signal given by the difference between the output of the physical system and the output of the model to update the parameter estimate. The parameter update is based on the retrospective cost function, whose minimizer updates the coefficients of the estimator. The present paper extends RCMR to the case where the model depends nonlinearly on multiple unknown parameters.**

I. Introduction

Many models have known structure but unknown parameters. The usual approach to estimating these parameters is to view them as constant states and then apply nonlinear estimation methods to estimate the state of the augmented system, thereby providing estimates of the unknown parameters along with the dynamic states. The extended Kalman filter (EKF), unscented Kalman filter (UKF), and ensemble Kalman filter (EnKF) are typically applied to these problems [1]–[7]. Yet another approach to parameter estimation is the variational method [8]–[10]. This approach requires an adjoint formulation of the dynamics and is computationally expensive due to the need for multiple iterations of the forward model and backward adjoint.

The special case of a linear system with uncertain entries in the state space model typically occurs in applications. Since the parameter states multiply the original states, the resulting estimation dynamics are nonlinear despite the fact that the original dynamics are linear. For this problem, a two-step procedure is used in [11], where a black-box model is first constructed based on the input-output data, and a similarity transformation is used to recover the structured unknown parameters. In [12], a sequential convex relaxation method is used to estimate unknown entries in a state space realization.

The present paper focuses on retrospective cost model refinement (RCMR) developed in [13]–[16]. RCMR is applicable to parameter estimation in linear or nonlinear gray-box models, with possibly nonlinear parameterization. RCMR uses an error signal given by the difference between the output of the physical system and the output of the model to update the parameter estimate. The parameter update is based on the retrospective cost function, whose minimizer updates the coefficients of the estimator.

The authors are with the Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109.

As in the case of UKF, RCMR uses the structure of the model with the current parameter estimates to propagate the states, but, unlike EKF, which requires the Jacobian of the dynamics, RCMR does not use knowledge of the model for the parameter updates. Also, unlike UKF and EnKF, RCMR does not require an ensemble of models, and, unlike adjoint-based methods, RCMR does not require an adjoint model.

The contribution of the present paper is an extension of RCMR as presented in [16]. In particular, RCMR was demonstrated in [16] for the case of a single parameter that may appear nonlinearly in a linear or nonlinear gray-box model. The present paper extends the algorithm presented in [16] to the case of multiple parameters that may appear nonlinearly in a gray-box model. The main focus of the present paper is thus the filter $G_{\rm f}$, which defines the retrospective cost function. In particular, we show that, for the case of multiple parameter estimation, the coefficients of $G_{\rm f}$ determine the search directions of RCMR. Consequently, a necessary condition for reaching the unknown parameters is to ensure that the range of the coefficients of $G_{\rm f}$ spans the parameter space. Numerical examples illustrate the performance of RCMR.

The paper is structured as follows. In Section II, we formulate the problem of estimating unknown parameters in a gray-box model. In Section III, we present the retrospective cost parameter estimator structure. The RCMR algorithm is presented in Section IV. Numerical examples are presented in Section V, included a Wiener system and a Hammerstein system. Finally, we conclude the paper with a discussion of the results and future work.

II. Problem formulation

Consider the discrete-time physical system model

$$x(k+1) = f(x(k),u(k),\mu) + w_1(k), \tag{1}$$

$$y(k) = h(x(k),u(k),\mu) + w_2(k), \tag{2}$$

where $x \in \mathbb{R}^{l_x}$ is the state, $u \in \mathbb{R}^{l_u}$ is the input, $y \in \mathbb{R}^{l_y}$ is the measured output, $w_1 \in \mathbb{R}^{l_x}, w_2 \in \mathbb{R}^{l_y}$ are the process and measurement noise, respectively, and $\mu \in \mathbb{R}^{l_\mu}$ is the unknown parameter vector. The functional forms of f and h are assumed to be known and may be nonlinear functions of μ. For example, in the case where μ is a scalar, the system

$$f(x,u,\mu) = \begin{bmatrix} \sin\mu & \cos\frac{\mu}{3} \\ \dfrac{e^{-\mu}}{3} & \dfrac{1}{1+\mu} \end{bmatrix} x + \begin{bmatrix} \log(1+\mu^2) \\ 1+\cos\mu \end{bmatrix} u, \tag{3}$$

$$h(x,u,\mu) = \begin{bmatrix} \mu & \mu^2 \end{bmatrix} x \tag{4}$$

is considered in [16]. In the present paper, μ may represent a vector of unknown parameters, which extends the approach of [16] to models involving multiple unknown parameters.

Next, we consider the estimation model

$$\hat{x}(k+1) = f(\hat{x}(k), u(k), \hat{\mu}(k)), \tag{5}$$
$$\hat{y}(k) = h(\hat{x}(k), u(k), \hat{\mu}(k)), \tag{6}$$

where $\hat{x}(k)$ is the computed state, $\hat{y}(k)$ is the computed output, and $\hat{\mu}(k)$ is the output of the parameter estimator at step k. The parameter estimator is updated by minimizing a cost function based on the performance variable

$$z(k) \triangleq \hat{y}(k) - y(k) \in \mathbb{R}^{l_y}. \tag{7}$$

The problem objective is to estimate μ using measurements of u and y. The parameter-estimation problem is represented by the block diagram in Figure 1.

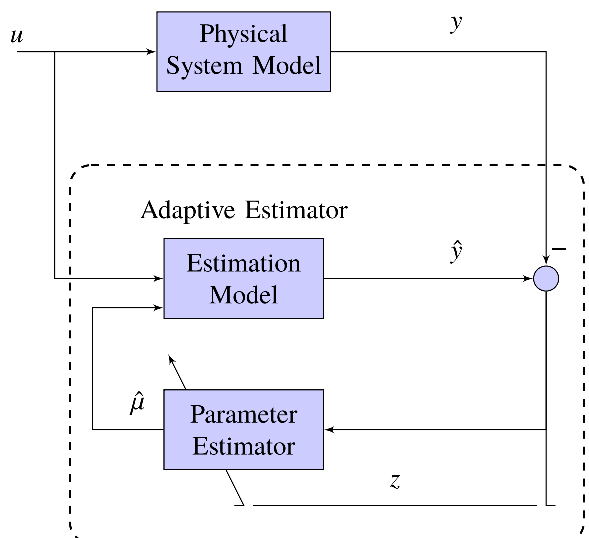

Fig. 1: Parameter-estimation architecture. The physical system, which is modeled by the physical system model (1), (2), is driven by u and produces measurements y. The adaptive estimator consists of the estimation model (5), (6), which is driven by measurements of u. The parameter estimate $\hat{\mu}$ is updated by the parameter estimator, which minimizes the error signal z.

In this paper, we make the following assumptions.

1) The parameter μ in (1), (2) is identifiable [17].
2) The input $u(k)$ is persistently exciting.

III. Parameter Estimator

We consider a parameter estimator represented by an ARMA model with a built-in integrator. The parameter estimate $\hat{\mu}$ is thus given by

$$\hat{\mu}(k) = \sum_{i=1}^{n_c} P_i(k)\hat{\mu}(k-i) + \sum_{i=1}^{n_c} Q_i(k)z(k-i) + R(k)g(k), \tag{8}$$

where

$$g(k) = g(k-1) + z(k-1), \tag{9}$$

and $P_i(k) \in \mathbb{R}^{l_\mu \times l_\mu}$, $Q_i(k), R(k) \in \mathbb{R}^{l_\mu \times l_y}$ are the coefficient matrices, which are updated by the RCMR algorithm. The integrator is embedded in the estimator to ensure that $z(k) \to 0$ as $k \to \infty$ and thus, assuming parameter identifiability and data persistency, that $\hat{\mu}(k) \to \mu$ as $k \to \infty$.

We rewrite (8) as

$$\hat{\mu}(k) = \Phi(k)\theta(k), \tag{10}$$

where the regressor matrix $\Phi(k)$ is defined by

$$\Phi(k) \triangleq I_{l_\mu} \otimes \phi^{\mathrm{T}}(k) \in \mathbb{R}^{l_\mu \times l_\theta},$$

where

$$\phi(k) \triangleq$$
$$[\hat{\mu}(k-1)^{\mathrm{T}} \ \cdots \ \hat{\mu}(k-n_c)^{\mathrm{T}} \ z(k-1)^{\mathrm{T}} \ \cdots \ z(k-n_c)^{\mathrm{T}} \ g(k)^{\mathrm{T}}]^{\mathrm{T}},$$

$$\theta(k) \triangleq \mathrm{vec}\left[\ P_1(k)\cdots P_{n_c}(k) \ Q_1(k)\cdots Q_{n_c}(k) \ R(k)\ \right] \in \mathbb{R}^{l_\theta},$$

$l_\theta \triangleq l_\mu^2 n_c + l_\mu l_y(n_c+1)$, "$\otimes$" is the Kronecker product, and "vec" is the column-stacking operator.

IV. RCMR Algorithm

In this section, we present the RCMR algorithm used to update the parameter estimator. RCMR is a specialized adaptation of the retrospective cost adaptive control (RCAC) algorithm [18].

A. Retrospective Performance Variable

We define the retrospective performance variable

$$\hat{z}(k) \triangleq z(k) + G_f(\mathbf{q})(\Phi(k)\hat{\theta} - \hat{\mu}(k)), \tag{11}$$

where $\mathbf{q}$ is the forward-shift operator, $\hat{\theta} \in \mathbb{R}^{l_\theta}$ contains the parameter estimator coefficients to be optimized,

$$G_f(\mathbf{q}) = \sum_{i=1}^{n_f} \frac{N_i}{\mathbf{q}^i}, \tag{12}$$

and, for all $i = 1, \ldots, n_f$, $N_i \in \mathbb{R}^{l_y \times l_\mu}$. G_f is an FIR filter of order n_f whose choice is discussed below. We rewrite (11) as

$$\hat{z}(k) = z(k) + N\Phi_b(k)\hat{\theta} - NU_b(k), \tag{13}$$

where

$$N \triangleq \begin{bmatrix} N_1 & \cdots & N_{n_f} \end{bmatrix} \in \mathbb{R}^{l_y \times n_f l_\mu},$$

$$\Phi_b(k) \triangleq \begin{bmatrix} \Phi(k-1) \\ \vdots \\ \Phi(k-n_f) \end{bmatrix} \in \mathbb{R}^{l_\mu n_f \times l_\theta},$$

$$U_b(k) \triangleq \begin{bmatrix} \hat{\mu}(k-1) \\ \vdots \\ \hat{\mu}(k-n_f) \end{bmatrix} \in \mathbb{R}^{l_\mu n_f}.$$

The vector $\hat{\theta}$, which contains the coefficients of the parameter estimator, is determined by minimizing the retrospective cost function, as described next.

B. Retrospective Cost Function

Using the retrospective performance variable $\hat{z}(k)$, we define the retrospective cost function

$$J(k,\hat{\theta}) \triangleq \sum_{i=1}^{k-1} \lambda^{k-i} \hat{z}^{\mathrm{T}}(i) R_z(i) \hat{z}(i) + \lambda^k \hat{\theta}^{\mathrm{T}} R_\theta \hat{\theta}, \tag{14}$$

where R_z and R_θ are positive-definite matrices, and $\lambda \le 1$ is the forgetting factor. The following result uses recursive least squares (RLS) to minimize (14).

Proposition IV.1. Let $P(0) = R_\theta^{-1}$, $\theta(0) = 0$. Then, for all $k \ge 1$, the retrospective cost function (14) has a unique global minimizer $\theta(k)$, which is given by

$$\theta(k) = \theta(k-1) - P(k)\Phi_{\mathrm{b}}(k-1)^{\mathrm{T}} N^{\mathrm{T}} R_z(k-1) \cdot (N\Phi_{\mathrm{b}}(k-1)\theta(k-1) + z(k-1) - NU_{\mathrm{b}}(k-1)), \tag{15}$$

$$P(k) = \lambda^{-1} P(k-1) - \lambda^{-1} P(k-1)\Phi_{\mathrm{b}}(k-1)^{\mathrm{T}} N^{\mathrm{T}} \Gamma(k)^{-1} \cdot N\Phi_{\mathrm{b}}(k-1)P(k-1), \tag{16}$$

where

$$\Gamma(k) \triangleq \lambda R_z(k-1)^{-1} + N\Phi_{\mathrm{b}}(k-1)P(k-1)\Phi_{\mathrm{b}}(k-1)^{\mathrm{T}} N^{\mathrm{T}}. \tag{17}$$

Furthermore, the parameter estimate at step k is given by

$$\hat{\mu}(k) = \Phi(k)\theta(k). \tag{18}$$

C. The filter G_{f}

The cost function (14) can be written as

$$J(k,\hat{\theta}) = \hat{\theta}^{\mathrm{T}} A_\theta(k)\hat{\theta} + 2b_\theta(k)^{\mathrm{T}}\hat{\theta} + c_\theta(k), \tag{19}$$

where

$$A_\theta(k) \triangleq \sum_{i=1}^{k-1} \lambda^{k-i} \Phi_{\mathrm{b}}(i)^{\mathrm{T}} N^{\mathrm{T}} R_z(i) N \Phi_{\mathrm{b}}(i) + \lambda^k R_\theta, \tag{20}$$

$$b_\theta(k) \triangleq \sum_{i=1}^{k-1} \lambda^{k-i} \Phi_{\mathrm{b}}(i)^{\mathrm{T}} N^{\mathrm{T}} R_z(i) (z(i) - NU_{\mathrm{b}}(i)). \tag{21}$$

The batch least squares minimizer $\theta(k)$ of (14) is given by

$$\theta(k) = -A_\theta(k)^{-1} b_\theta(k). \tag{22}$$

The following result shows that the estimate $\hat{\mu}(k)$ of μ is constrained to lie in a subspace determined by the coefficients of G_{f}.

Theorem IV.1. Let $R_\theta = \beta I_{l_\theta}$, and let $\theta(k)$ be given by (22). Let $\Phi \triangleq I_{l_u} \otimes \phi^{\mathrm{T}}$, where $\phi \in \mathbb{R}^{l_\phi}$, and $l_\phi \triangleq l_\theta / l_u$. Then, for all $k \ge 1$,

$$\Phi\theta(k) = \frac{-1}{\lambda^k \beta} \begin{bmatrix} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{bmatrix} \sum_{i=1}^{k-1} \lambda^{k-i} \Xi(\phi,i) R_z(i) \cdot (z(i) + N\Phi_{\mathrm{b}}(i)\theta(k) - NU_{\mathrm{b}}(i)) \in \mathcal{R}\left(\begin{bmatrix} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{bmatrix} \right), \tag{23}$$

where

$$\Xi(\phi,i) \triangleq \begin{bmatrix} \phi^{\mathrm{T}}\phi(i-1) \otimes I_{l_z} \\ \vdots \\ \phi^{\mathrm{T}}\phi(i-n_{\mathrm{f}}) \otimes I_{l_z} \end{bmatrix}. \tag{24}$$

The following proposition follows from Theorem IV.1.

Proposition IV.2. Let $R_\theta = \beta I_{l_\theta}$, and let $\theta(k)$ be given by (22). Then, for all $k \ge 1$,

$$\hat{\mu}(k) = \Phi(k)\theta(k) = \frac{-1}{\lambda^k \beta} \begin{bmatrix} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{bmatrix} \sum_{i=1}^{k-1} \lambda^{k-i} \Xi(\phi(k),i) \cdot R_z(i)(z(i) + N\Phi_{\mathrm{b}}(i)\theta(k) - NU_{\mathrm{b}}(i)) \in \mathcal{R}\left(\begin{bmatrix} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{bmatrix} \right). \tag{25}$$

Proposition IV.2 shows that, in order for RCMR to estimate μ, the coefficients of G_{f} must be chosen such that

$$\mu \in \mathcal{R}\left(\begin{bmatrix} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{bmatrix} \right). \tag{26}$$

Consequently, a first-order G_{f} suffices in the case where μ is scalar as in [16]. In the case where μ is vector with two components, the order of G_{f} must be at least two and it coefficients must be chosen to satisfy $\mathcal{R}\left(\begin{bmatrix} N_1^{\mathrm{T}} & N_2^{\mathrm{T}} \end{bmatrix} \right) = \mathbb{R}^2$. A similar observation applies in case of three or more uncertain parameters.

In addition to the range condition (IV.2), numerical examples show that the parameter estimates produced by RCMR lie initially along the direction $\sum_{i=1}^{n_{\mathrm{f}}} N_i^{\mathrm{T}}$. This property is due to the fact that the regressor vector $\phi(k)$ is approximately constant for a limited time interval after the start of parameter estimation. Secondly, after initially evolving in the direction $\sum_{i=1}^{n_{\mathrm{f}}} N_i^{\mathrm{T}}$, the parameter estimate $\hat{\mu}(k)$ moves toward the subspace spanned by $N_1^{\mathrm{T}}, \ldots, N_{n_{\mathrm{f}}-1}^{\mathrm{T}}$, then toward the subspace spanned by $N_1^{\mathrm{T}}, \ldots, N_{n_{\mathrm{f}}-2}^{\mathrm{T}}$, and so forth until, eventually, they tend toward the subspace spanned by N_1^{T}. Convergence may also depend on the choice of $\pm G_{\mathrm{f}}$. These properties are demonstrated in the numerical examples given below.

V. Numerical examples

In this section, we use RCMR to estimate multiple unknown parameters that appear nonlinearly in a system parameterization. Note that, like UKF, the functional form of the parameter dependence is used to propagate the estimation model; however, unlike EKF, explicit knowledge of the parameter dependence need not be known. In other words, the model must be computable, but the details of the computation need not be known by the user.

Consider the LTI physical system model

$$x(k+1) = A(\mu_1)x(k) + B(\mu_2)u(k) + D_1 w(k), \tag{27}$$

$$y(k) = C(\mu_3)x(k) + D_2 w(k), \tag{28}$$

where

$$A(\mu_1) = \begin{bmatrix} \sin\mu_1 & \cos\frac{\mu_1}{3} \\ \frac{e^{-\mu_1}}{3} & \frac{0.5}{1.1+\mu_1^2} \end{bmatrix}, \tag{29}$$

$$B(\mu_2) = \begin{bmatrix} \log(1+\mu_2^2) \\ 1+\sin\mu_2 \end{bmatrix}, \tag{30}$$

$$C(\mu_3) = \begin{bmatrix} \mu_3 & 4\mu_3^2 \end{bmatrix}, \tag{31}$$

the true values of the parameters μ_1, μ_2, μ_3 are $0.3, 0.2, 0.4$, respectively, and $D_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}^{\mathrm{T}}$ and $D_2 = 1$. The estimation model is

$$\hat{x}(k+1) = A(\hat{\mu}_1(k))\hat{x}(k) + B(\hat{\mu}_2(k))u(k), \tag{32}$$
$$\hat{y}(k) = E(\hat{\mu}_3(k))\hat{x}(k), \tag{33}$$

where $\hat{\mu}(k)$ is the output of the parameter estimator updated by RCMR.

We generate the measurement $y(k)$ using the input $u(k) = -2 + \sin\left(\frac{2\pi}{40}k\right) + \sin\left(\frac{2\pi}{80}k - 0.3\right) + \sin\left(\frac{2\pi}{160}k - 0.5\right)$, the initial state $x(0) = [10 \ \ 10]^{\mathrm{T}}$, the noise signal $w \sim \mathcal{N}(0, 10^{-6})$. To reflect the absence of additional information, the initial state $\hat{x}(0)$ of the estimation model and the initial estimate $\hat{\mu}(0)$ of the unknown parameter μ are both set to zero.

Example V.1. In this example, we estimate the unknown parameters μ_1 and μ_2 parameterizing A and B. Note that $\hat{\mu}_3(k) = \mu_3$ is assumed to be known in this example. We set

$$G_{\mathrm{f}}(\mathbf{q}) = \frac{\begin{bmatrix} 1 & 0 \end{bmatrix}}{\mathbf{q}} + \frac{\begin{bmatrix} 0 & 1 \end{bmatrix}}{\mathbf{q}^2},$$

and $n_{\mathrm{c}} = 2$ so that $l_\theta = 14$. Furthermore, let $R_\theta = 10^6 I_{l_\theta}$ and $\lambda = 0.999$. Figure 2 shows the estimates of μ_1 and μ_2. Note that $\hat{\mu}(k)$ initially evolves along $\sum_{i=1}^{2} N_i^{\mathrm{T}}$ and then moves toward N_1^{T}. $\diamond$

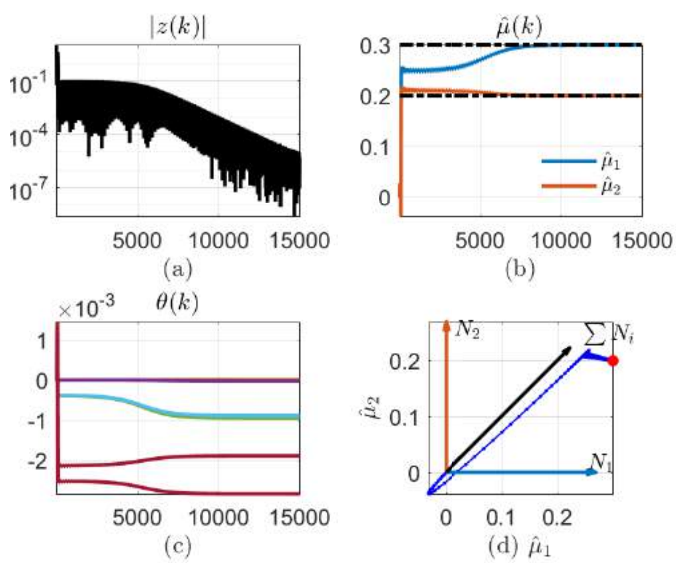

Fig. 2: RCMR estimates of the unknown parameters μ_1 and μ_2 parameterizing the system (27), (28) with the nonlinear parameter dependence (29), (30). (a) shows the performance z on a log scale; (b) shows the parameter estimates $\hat{\mu}_1$ and $\hat{\mu}_2$; (c) shows the adapted coefficients θ of the parameter estimator; (d) shows the trajectory of the estimates computed by RCMR along with the directions of filter coefficients N_i^{T} (scaled to focus on the estimates), where the true parameters are indicated by the red dot. Note that the estimate initially evolves along $\sum_{i=1}^{2} N_i^{\mathrm{T}}$, and then moves toward N_1^{T}.

Example V.2. Next, we estimate the unknown parameters μ_1 and μ_3 parameterizing A and C. Note that $\hat{\mu}_2(k) = \mu_2$ in this example. We set

$$G_{\mathrm{f}}(\mathbf{q}) = \frac{\begin{bmatrix} 0 & 1 \end{bmatrix}}{\mathbf{q}} + \frac{\begin{bmatrix} 1 & 0 \end{bmatrix}}{\mathbf{q}^2},$$

and $n_{\mathrm{c}} = 2$ so that $l_\theta = 14$. Furthermore, let $R_\theta = 10^6 I_{l_\theta}$ and $\lambda = 0.999$. Figure 3 shows the estimates of μ_1 and μ_3. Note that $\hat{\mu}(k)$ initially evolves along $\sum_{i=1}^{2} N_i^{\mathrm{T}}$ and then moves toward N_1^{T}. $\diamond$

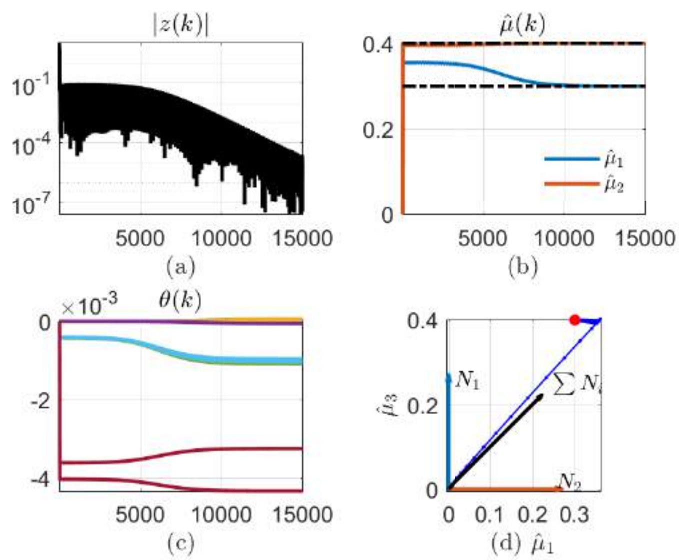

Fig. 3: RCMR estimate of the unknown parameters μ_1 and μ_2 parameterizing the system (27), (28) with the nonlinear parameter dependence (29), (31). (a) shows the performance z on a log scale; (b) shows the parameter estimate $\hat{\mu}_1$ and $\hat{\mu}_2$; (c) shows the adapted coefficients θ of the parameter estimator; (d) shows the trajectory of the estimates computed by RCMR, along with the filter coefficient N_i^{T} directions (scaled to focus on the estimates); and the true value is shown by the red dot. Note that $\hat{\mu}(k)$ initially evolves along $\sum_{i=1}^{2} N_i^{\mathrm{T}}$ and then moves toward N_1^{T}.

Example V.3. In this example, we estimate the unknown parameters μ_1, μ_2, and μ_3 parameterizing A, B, and C. We set

$$G_{\mathrm{f}}(\mathbf{q}) = \frac{\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}}{\mathbf{q}} + \frac{\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}}{\mathbf{q}^2} + \frac{\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}}{\mathbf{q}^3},$$

and $n_{\mathrm{c}} = 1$ so that $l_\theta = 15$. Furthermore, let $R_\theta = 10^6 I_{l_\theta}$ and $\lambda = 0.999$. In estimating three unknown parameters, we observe that $P(k)$ becomes ill-conditioned in (16) with forgetting. To prevent this, we use a variation of directional forgetting [19] to update $P(k)$. Instead of dividing the RHS of (16) by λ, we propagate $P(k)$ as follows. At step k, we compute

$$R(k) \triangleq \sum_{i=1}^{2l_\theta} \Phi_{\mathrm{b}}(k-i)^{\mathrm{T}} N^{\mathrm{T}} N \Phi_{\mathrm{b}}(k-i). \tag{34}$$

The rank of $R(k)$ indicates the persistency of the regressor matrix used to update θ. Without the forgetting factor, that is, with $\lambda = 1$, $\mathrm{rank}(R(k))$ singular values of $P(k)$ decrease since the data contains new information along the directions of the corresponding singular vectors. To prevent $P(k)$ from becoming ill-conditioned, we divide by λ only the singular values that correspond to the singular vectors receiving data with new information. At step k, we thus compute

$$\begin{aligned}\bar{P}(k) &= P(k-1) - P(k-1)\Phi_{\mathrm{b}}(k-1)^{\mathrm{T}} N^{\mathrm{T}} \Gamma(k)^{-1} \\ &\quad \cdot N\Phi_{\mathrm{b}}(k-1)P(k-1), \quad (35)\\ \Sigma(k) &= U(k)^{\mathrm{T}} \bar{P}(k) U(k), \quad (36)\\ \bar{\Sigma}(k)(i,i) &= \begin{cases} \Sigma(k)(i,i), & i < l_\theta - \operatorname{rank}(R(k)),\\ \Sigma(k)(i,i)/\lambda, & i \geq l_\theta - \operatorname{rank}(R(k)),\end{cases} \quad (37)\\ P(k) &= U(k)\bar{\Sigma}(k)U(k)^{\mathrm{T}}, \quad (38)\end{aligned}$$

where $\Sigma(k)$ contains the singular values of $P(k)$, and $U(k)$ contains the singular vectors of $P(k)$. Note that $\Sigma(k)$ and $\bar{\Sigma}(k)$ are diagonal matrices, and $\Sigma(k)(i,i)$ denotes the (i,i) entry of $\Sigma(k)$. Figure 4 shows the estimates of μ_1, μ_2, and μ_3. ◇

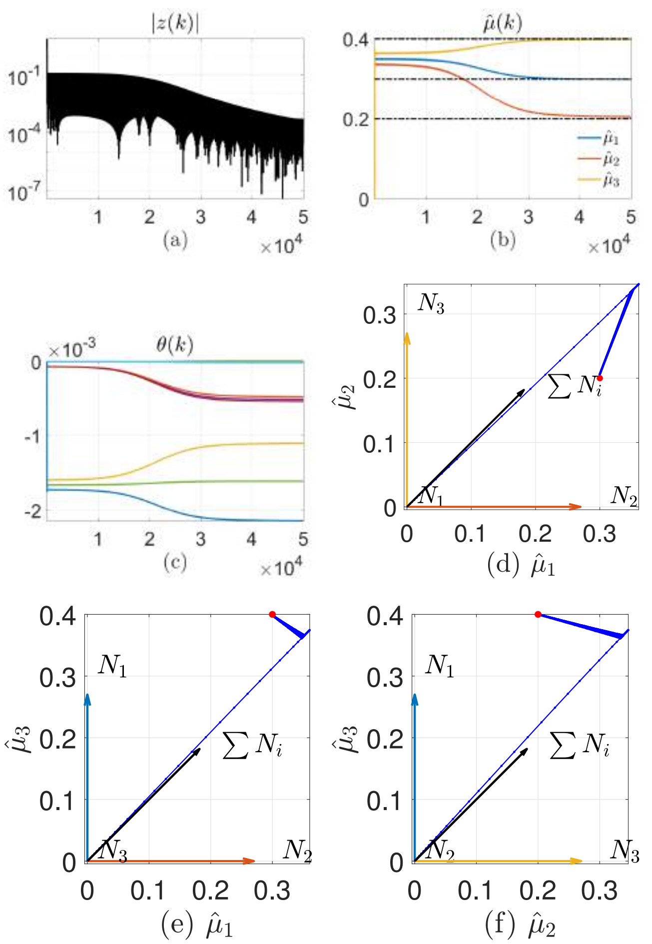

Fig. 4: RCMR estimate of the unknown parameters μ_1 and μ_2 parameterizing the system (27), (28) with the nonlinear parameter dependence (29)–(31). (a) shows the performance z on a log scale; (b) shows the parameter estimate $\hat{\mu}_1$ and $\hat{\mu}_2$; (c) shows the adapted coefficients θ of the parameter estimator; (d), (e) and (f) show the trajectory of the estimates computed by RCMR projected on the canonical planes, along with the filter coefficient N_i^{T} directions (scaled to focus on the estimates), and the true value is shown by the red dot.

Example V.4. In this example, we consider a nonlinear physical system model where the state evolves according to (27), and

$$y(k) = \cos x_1(k) + \sin x_2(k) + D_2 w(k). \quad (39)$$

Note that this is a Wiener system since the output map is nonlinear. We estimate the unknown parameters μ_1 and μ_2 parameterizing A and B. We set

$$G_{\mathrm{f}}(\mathbf{q}) = \frac{\begin{bmatrix} 0 & 1 \end{bmatrix}}{\mathbf{q}} + \frac{\begin{bmatrix} 1 & 0 \end{bmatrix}}{\mathbf{q}^2},$$

and $n_{\mathrm{c}} = 2$ so that $l_\theta = 14$. Furthermore, let $R_\theta = 10^3 I_{l_\theta}$ and $\lambda = 0.999$. Similar to Example V.3, we use (34)–(38) to update $P(k)$. Figure 5 shows the estimates of μ_1 and μ_2. Note that $\hat{\mu}(k)$ initially evolves along $\sum_{i=1}^{2} N_i^{\mathrm{T}}$. However, Unlike previous examples, $\hat{\mu}$ then moves toward N_2^{T}. ◇

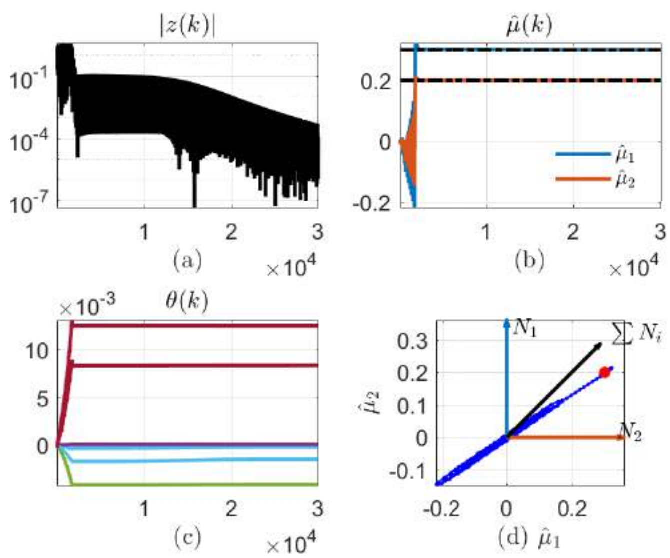

Fig. 5: RCMR estimate of the unknown parameters μ_1 and μ_2 parameterizing the system (27), (39) with the nonlinear parameter dependence (29), (30). (a) shows the performance z on a log scale; (b) shows the parameter estimate $\hat{\mu}_1$ and $\hat{\mu}_2$; (c) shows the adapted coefficients θ of the parameter estimator; (d) shows the trajectory of the estimates computed by RCMR, along with the filter coefficient N_i^{T} directions (scaled to focus on the estimates); and the true value is shown by the red dot. Note that $\hat{\mu}(k)$ initially evolves along $\sum_{i=1}^{2} N_i^{\mathrm{T}}$, and unlike previous examples, moves toward N_2^{T}.

Example V.5. In this example, we consider a nonlinear physical system model

$$x(k+1) = Ax(k) + B\operatorname{sat}(u(k)) + D_1 w(k), \quad (40)$$
$$y(k) = Cx(k) + D_2 w(k), \quad (41)$$

where

$$A = \begin{bmatrix} 0.3 & 0.2 \\ 0.1 & 0.6 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$
$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad D_2 = 1,$$

and $\operatorname{sat}(u(k)) = \min(\max(u(k), \mu_1), \mu_2)$. Note that this is a Hammerstein system since the input map is nonlinear. Note that the input map is a saturation function sat parameterized by unknown parameters μ_1 and μ_2. The true values of μ_1, μ_2 are $-1, 0.5$ respectively.

We generate the measurements $y(k)$ using the input $u(k) = -1+\sin\left(\frac{2\pi}{40}k\right)+\sin\left(\frac{2\pi}{80}k-0.3\right)+0.5\sin\left(\frac{2\pi}{160}k-0.5\right)$, the initial state $x(0) = [1\ 1]^{\mathrm{T}}$, the noise signal $w \sim \mathcal{N}(0, 10^{-6})$. To reflect the absence of additional information, the initial state $\hat{x}(0)$ of the estimation model and the initial estimate $\hat{\mu}(0)$ of the unknown parameter μ are both set to zero.

We set

$$G_{\mathrm{f}}(\mathbf{q}) = \frac{\left[\begin{array}{cc} -1 & 0 \end{array}\right]}{\mathbf{q}} + \frac{\left[\begin{array}{cc} 0 & 1 \end{array}\right]}{\mathbf{q}^2},$$

and $n_{\mathrm{c}} = 2$ so that $l_\theta = 14$. Furthermore, let $R_\theta = 10^3 I_{l_\theta}$ and $\lambda = 0.999$. Similar to Example V.3, we use (34)–(38) to update $P(k)$. Figure 6 shows the estimates of μ_1 and μ_2. Note that $\hat{\mu}(k)$ initially evolves along $\sum_{i=1}^{2} N_i^{\mathrm{T}}$ and then moves toward N_1^{T}. ◇

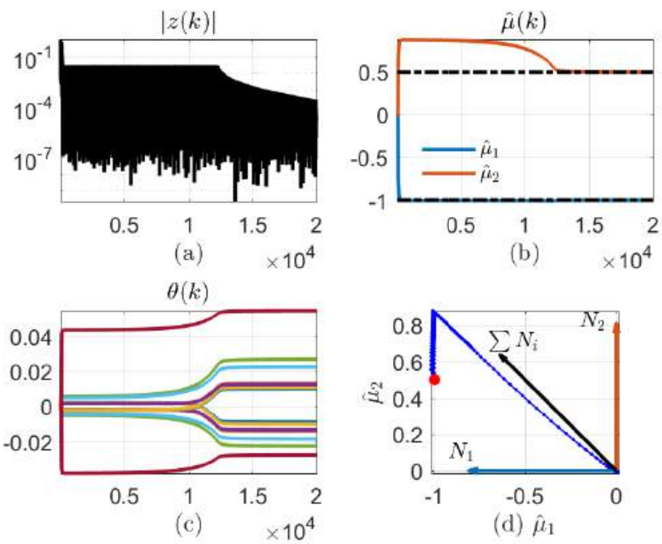

Fig. 6: RCMR estimate of the unknown parameters μ_1 and μ_2 parameterizing the system (27), (28) with the nonlinear parameter dependence (29), (31). (a) shows the performance z on a log scale; (b) shows the parameter estimate $\hat{\mu}_1$ and $\hat{\mu}_2$; (c) shows the adapted coefficients θ of the parameter estimator; (d) shows the trajectory of the estimates computed by RCMR, along with the filter coefficient N_i^{T} directions (scaled to focus on the estimates); and the true value is shown by the red dot. Note that $\hat{\mu}(k)$ initially evolves along $\sum_{i=1}^{2} N_i^{\mathrm{T}}$ and then moves toward N_1^{T}.

VI. Conclusions and Future Work

Retrospective cost model refinement (RCMR) was extended to the problem of estimating multiple unknown parameters in a linear or nonlinear model. It was shown that the parameter estimates of the unknown parameters are confined to the subspace spanned by the coefficients of the filter G_{f}, which is chosen by the user. It was also shown that the parameter estimates tend toward a sequence of subspaces spanned by subsets of the coefficients. These properties were illustrated by linear and nonlinear examples involving two or three unknown parameters in the presence of unknown state initial condition and unknown plant disturbance and sensor noise.

VII. Acknowledgments

This research was supported by AFOSR under DDDAS (Dynamic Data-Driven Applications Systems http://www.1dddas.org/) grant FA9550-16-1-0071.

References

[1] L. Ljung, "Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems," *IEEE Trans. Autom. Contr.*, vol. 24, pp. 36–50, 1979.

[2] G. L. Plett, "Extended kalman filtering for battery management systems of lipb-based hev battery packs: Part 3. state and parameter estimation," *J. Power Sources*, vol. 134, no. 2, pp. 277–292, 2004.

[3] R. Van Der Merwe and E. A. Wan, "The square-root unscented kalman filter for state and parameter-estimation," in *Int. Conf. Acoustics, Speech, Sig. Proc.*, 2001, pp. 3461–3464.

[4] E. A. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation," in *Adaptive Sys. Sig. Proc. Comm. Contr. Symp.*, 2000, pp. 153–158.

[5] G. Evensen, "The ensemble kalman filter for combined state and parameter estimation," *IEEE Contr. Sys. Mag.*, vol. 29, no. 3, 2009.

[6] H. Moradkhani, S. Sorooshian, H. V. Gupta, and P. R. Houser, "Dual state–parameter estimation of hydrological models using ensemble kalman filter," *Adv. Water Resources*, vol. 28, pp. 135–147, 2005.

[7] R. Madankan, P. Singla, T. Singh, and P. D. Scott, "Polynomial-chaos-based bayesian approach for state and parameter estimations," *AIAA J. Guid. Contr. Dyn.*, vol. 36, pp. 1058–1074, 2013.

[8] T. S. Jaakkola and M. I. Jordan, "Bayesian parameter estimation via variational methods," *Stat. Computing*, vol. 10, pp. 25–37, 2000.

[9] R. L. Raffard, K. Amonlirdviman, J. D. Axelrod, and C. J. Tomlin, "An adjoint-based parameter identification algorithm applied to planar cell polarity signaling," *IEEE Trans. Autom. Contr.*, vol. 53, pp. 109–121, 2008.

[10] M. Eknes and G. Evensen, "Parameter estimation solving a weak constraint variational formulation for an ekman model," *J. Geophys. Res.: Oceans*, vol. 102, pp. 12 479–12 491, 1997.

[11] G. Mercère, O. Prot, and J. A. Ramos, "Identification of parameterized gray-box state-space systems: From a black-box linear time-invariant representation to a structured one," *IEEE Trans. Autom. Contr.*, vol. 59, pp. 2873–2885, 2014.

[12] C. Yu, L. Ljung, and M. Verhaegen, "Identification of structured state-space models," *Automatica*, vol. 90, pp. 54–61, 2018.

[13] A. M. D'Amato, A. J. Ridley, and D. S. Bernstein, "Retrospective-cost-based adaptive model refinement for the ionosphere and thermosphere," *Stat. Anal. Data Mining*, pp. 446–458, 2011.

[14] X. Zhou, D. S. Bernstein, J. L. Stein, and T. Ersal, "Battery state of health monitoring by estimation of side reaction current density via retrospective-cost subsystem identification," *J. Dyn. Sys. Meas. Contr*, vol. 139, pp. 91 007–1–15, 2017.

[15] A. Goel, K. Duraisamy, and D. S. Bernstein, "Parameter estimation in the burgers equation using retrospective-cost model refinement," in *Proc. Amer. Contr. Conf.*, Boston, MA, 2016, pp. 6983–6988.

[16] A. Goel and D. S. Bernstein, "Parameter estimation for nonlinearly parameterized gray-box models," in *Proc. Amer. Contr. Conf.*, Milwaukee, WI, June 2018.

[17] M. Grewal and K. Glover, "Identifiability of linear and nonlinear dynamical systems," *IEEE Trans. Autom. Contr.*, vol. 21, no. 6, pp. 833–837, 1976.

[18] Y. Rahman, A. Xie, and D. S. Bernstein, "Retrospective Cost Adaptive Control: Pole Placement, Frequency Response, and Connections with LQG Control," *IEEE Contr. Sys. Mag.*, vol. 37, no. 5, pp. 28–69, October 2017.

[19] R. Kulhavỳ and M. Kárnỳ, "Tracking of slowly varying parameters by directional forgetting," *IFAC Proc. Vol.*, vol. 17, no. 2, pp. 687–692, 1984.