Check for updates

Retrospective Cost Subsystem Estimation and Smoothing for Linear Systems with Structured Uncertainty

Ming-Jui Yu* and Dennis S. Bernstein†
University of Michigan, Ann Arbor, Michigan 48109

DOI: 10.2514/1.I010581

Two algorithms for the combined state and parameter estimation (CSPE) of a linear, time-invariant system are presented. Retrospective cost subsystem estimation is formulated under the assumption that the initial state is known. A smoother algorithm based on this formulation is developed for the case in which the initial state is unknown. It is numerically demonstrated that these algorithms are more accurate for CSPE than the extended Kalman filter and the unscented Kalman filter.

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.I010581

I. Introduction

THE classical Kalman filter is applicable to state estimation for systems with known linear dynamics [1,2]. In the case in which some of the entries of the dynamic matrix are uncertain, however, the simultaneous estimation of both the states and uncertain parameters entails a nonlinear estimation problem. This is the combined state and parameter estimation (CSPE) problem. A fundamental question concerning CSPE is to determine whether or not the unmeasured states and unknown parameters can be estimated. This problem can be recast as an identifiability problem, and conditions are given in [3–5].

The classical approach to CSPE is to apply the extended Kalman filter (EKF) [6]. Alternatively, the unscented Kalman filter (UKF) can be used [7] and applied to CSPE in [8]. Yet, another approach to CSPE is based on the polynomial-chaos series expansion [9,10].

The present paper has two objectives. The first objective was to apply the EKF and UKF to numerical examples, and assess the effectiveness of these methods for CSPE. Within the context of the present paper, the EKF and UKF are used as nonlinear observers in the absence of process and sensor noise. The goal was thus to determine their effectiveness on the specific quadratic nonlinearity that arises in the state and parameter estimation problem. The effect of disturbances and sensor noise on these estimates is also of interest in practice, but is outside the scope of this investigation. The second objective was to apply retrospective cost subsystem estimation (RCSE) [11,12] to CSPE, and compare the performance to the EKF and UKF. The idea behind RCSE is to view the uncertain entries of the dynamic matrix as an uncertain subsystem connected in feedback. A performance metric based on the difference between the outputs of the true system and the model is then recursively optimized to update the estimates of the unknown parameters.

The first main contribution of the present paper is the formulation of RCSE under the assumption that the initial state of the system is known. Although this assumption is often unrealistic in practice, the accuracy of the resulting estimates is notable because knowledge of the initial state does not improve the performance of either EKF or UKF. In any event, this initial formulation motivates the development of the RCSE smoother (RCSES), which estimates both the unknown entries of the dynamic matrix and the unknown components of the initial state. Consequently, the second main contribution of the present paper is the development and assessment of RCSES.

The contents of the paper are as follows. The CSPE problem is stated in Sec. II. Next, the EKF and UKF are applied to second-order CSPE problems in Secs. III and IV. Section V shows that the accuracy of the UKF for CSPE can be enhanced by using a state-dependent coefficient formulation of the system dynamics. This approach is taken in [13] for nonlinear systems, and applied to second- and third-order CSPE problems. RCSE is formulated in Secs. VI and VII, and is applied to third- and eighth-order CSPE problems in Sec. VIII. RCSES is formulated in Sec. IX, and applied to second- and third-order CSPE problems in Sec. X. Finally, in Sec. XI, the UKF and RCSES are applied to a CSPE problem based on a linearized fourth-order longitudinal aircraft dynamic model, and the results are compared.

II. Combined State and Parameter Estimation

Consider the discrete-time, linear time-invariant system:

$$x(k+1) = Ax(k) \tag{1}$$

$$x(0) = x_0 \tag{2}$$

$$y_0(k) = Ex(k) \tag{3}$$

in which $n \geq 2$

$$x(k) = \begin{bmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{bmatrix} \in \mathbb{R}^n, \qquad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in \mathbb{R}^{n\times n} \tag{4}$$

The variable $y_0(k) \in \mathbb{R}$ is the measurement, and

Received 10 June 2017; revision received 27 February 2018; accepted for publication 7 July 2018; published online 6 September 2018. Copyright © 2018 by the American Institute of Aeronautics and Astronautics, Inc. All rights reserved. All requests for copying and permission to reprint should be submitted to CCC at www.copyright.com; employ the ISSN 2327-3097 (online) to initiate your request. See also AIAA Rights and Permissions www.aiaa.org/randp.

*Graduate Student, Aerospace Engineering Department, 1320 Beal Ave.
†Professor, Aerospace Engineering Department, 1320 Beal Ave.

$$E = [e_1 \ \cdots \ e_n] \in \mathbb{R}^{1\times n} \tag{5}$$

We assume that E is known, but A has structured uncertainty in the sense that some entries of A are known and others are unknown. We can thus write

$$A = A_0 + \Delta A \tag{6}$$

in which A_0 is the nominal dynamic matrix, and ΔA models the uncertain entries of A. Note that the assumption that E is known fixes the basis, in which A_0 and the uncertainty ΔA are represented. Also, the fact that some entries in A are unknown makes it impossible to transform A into a canonical form. The objective is to use the measurement $y_0(k)$, in which $k \geq 0$, to estimate the unknown entries of A and the components $x_1(k), \ldots, x_n(k)$ of the state $x(k)$. This is the CSPE problem.

If the state $x(k)$ is known for all $k \geq 0$, then it is straightforward to estimate the uncertain entries of A. Likewise, if all of the entries of A are known, then standard techniques can be used to estimate the state. The difficulty of the CSPE problem stems from the specific quadratic nonlinearity arising from the fact that both states and parameters are unknown. In fact, this problem is solvable only in certain special cases, as discussed in the previous section. Note that this problem formulation does not include either process noise or sensor noise, and thus, the problem is deterministic. The focus is thus on nonlinear observers; extensions to nonlinear estimation are mentioned in the Conclusions.

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.1010581

III. Extended Kalman Filter

To provide a baseline for later developments, in this section, we apply the EKF to the CSPE problem.

Example 1: $n = 2$ and Two Unknown Entries in a Single Row

Consider Eqs. (1–3) with

$$A = \begin{bmatrix} 0.27 & 1.17 \\ -0.8 & 0.2 \end{bmatrix}, \qquad x_0 = \begin{bmatrix} -23 \\ 17 \end{bmatrix}, \qquad E = [1 \quad 0] \tag{7}$$

and assume that the entries $a_{11} = 0.27$ and $a_{12} = 1.17$ of A are unknown. To apply the EKF, we first augment the dynamics (1) with additional equations that represent the fact that the unknown parameters are constant. The augmented system has the form

$$X(k+1) = \tilde{A}X(k) \tag{8}$$

$$X(0) = X_0 \tag{9}$$

$$y_0(k) = \tilde{E}X(k) \tag{10}$$

in which

$$\tilde{A} \triangleq \begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ -0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \tilde{E} \triangleq [E \quad 0_{1\times 2}], \qquad X(k) \triangleq \begin{bmatrix} x_1(k) \\ x_2(k) \\ a_{11} \\ a_{12} \end{bmatrix} \tag{11}$$

Forming the Jacobian matrix of Eq. (8) yields the augmented estimator system:

$$\hat{X}(k+1) = \hat{\tilde{A}}(k)\hat{X}(k) \tag{12}$$

$$\hat{X}(0) = \hat{X}_0 \tag{13}$$

$$\hat{y}_0(k) = \tilde{E}\hat{X}(k) \tag{14}$$

in which

$$\hat{\tilde{A}}(k) \triangleq \begin{bmatrix} \hat{a}_{11}(k) & \hat{a}_{12}(k) & \hat{x}_1(k) & \hat{x}_2(k) \\ -0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \hat{X}(k) \triangleq \begin{bmatrix} \hat{x}_1(k) \\ \hat{x}_2(k) \\ \hat{a}_{11}(k) \\ \hat{a}_{12}(k) \end{bmatrix} \tag{15}$$

The variables $\hat{x}_1(k), \hat{x}_2(k)$ denote estimates of $x_1(k), x_2(k)$, and the variables $\hat{a}_{11}(k), \hat{a}_{12}(k)$ denote estimates of a_{11}, a_{12}. The KF is then applied to Eqs. (12–14).

To evaluate the accuracy of the EKF, we define the relative initial estimation errors:

$$\xi_x \triangleq \frac{\|\hat{x}_2(0) - x_2(0)\|}{\|x_2(0)\|}, \qquad \xi_a \triangleq \frac{\|\hat{a}(0) - a\|}{\|a\|} \tag{16}$$

in which the true parameter vector a and its estimate $\hat{a}$ are defined as

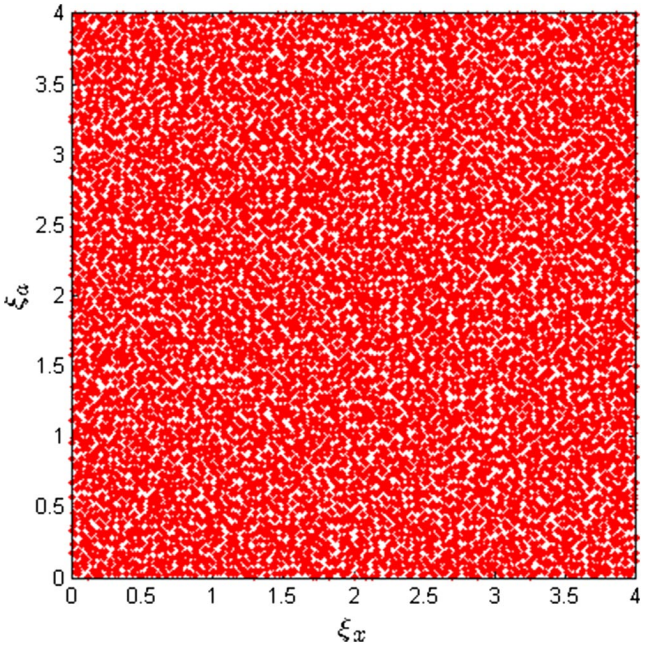


Fig. 1 Application of EKF to example 1; EKF is applied with 10,000 randomly generated initial estimates $(\hat{x}_2(0),\ \hat{a}_{11}(0),\ \hat{a}_{12}(0))$ using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 1000]$; trials in which EKF estimates both components of a within 10% relative error at step $k = 1000$ are labeled with cyan; trials in which EKF estimates exactly one component of a within 10% relative error at step $k = 1000$ are labeled with black; and trials in which EKF estimates neither of the components of a within 10% relative error at step $k = 1000$ are labeled with red; 100% of the trials are red; note: in all subsequent examples, cyan, black, and red indicate, respectively, trials in which all, at least one, and none of the components of a satisfy the accuracy specification.

$$a \triangleq \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix}, \qquad \hat{a} \triangleq \begin{bmatrix} \hat{a}_{11} \\ \hat{a}_{12} \end{bmatrix} \tag{17}$$

Note that $\xi_x = 0$ if and only if $\hat{x}_2(0) = x_2(0)$, and $\xi_a = 0$ if and only if $\hat{a}_{11}(0) = a_{11}$ and $\hat{a}_{12}(0) = a_{12}$.

To assess the performance of the EKF, we consider 10,000 randomly generated initial estimates of the unmeasured state and the uncertain entries of A. Because x_1 is measured, we set $\hat{x}_1(0) = x_1(0)$, and we choose initial estimates $(\hat{x}_2(0), \hat{a}_{11}(0), \hat{a}_{12}(0))$, such that $\xi_x, \xi_a \in (0, 4)$. Using the notation of [6], we set the initial covariance matrix to be $P(0) = 10{,}000I_4$, and choose the tuning parameters $Q = 10^{-2}I_{2l_x}$ and $R = 0$. Figure 1 shows that, for all 10,000 initial estimates, none of the estimates $\hat{a}(1000)$ are within 10% of the true parameter a.

IV. Unscented Kalman Filter

In this section, we apply the UKF to the CSPE problem.

Example 2: Example 1 Revisited

In example 1, the first row of the Jacobian matrix (15) gives an erroneous factor of 2 as compared to $\tilde{A}$, which is consistent with the resulting poor performance. Therefore, we revisit example 1 by defining

$$\tilde{A}(k) \triangleq \begin{bmatrix} \hat{a}_{11}(k) & \hat{a}_{12}(k) & 0 & 0 \\ -0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{18}$$

for Eqs. (12–14) and applying the UKF to the augmented system. To assess the performance of the UKF with Eq. (18), we consider 10,000 randomly generated initial estimates of the unmeasured state and the uncertain entries of A. Because x_1 is measured, we set $\hat{x}_1(0) = x_1(0)$, and we choose initial estimates $(\hat{x}_2(0), \hat{a}_{11}(0), \hat{a}_{12}(0))$, such that $\xi_x, \xi_a \in (0, 4)$. Using the notation of [7], we set the initial covariance matrix to be $P(0) = 10{,}000I_4$, and choose the tuning parameters $\alpha = 1$, $\kappa = 0$, $\beta = 2$, $Q = 10^{-2}I_{2l_x}$, and $R = 0$. Figure 2 shows that, for all 10,000 initial estimates, none of the estimates $\hat{a}(1000)$ are within 10% of the true parameter a.

V. UKF with State-Dependent Coefficients

In this section, we consider an extension of the UKF.

A. Example 3: Example 1 Revisited

We revisit example 1 by defining the state-dependent matrix:

$$\tilde{A}(k) \triangleq \begin{bmatrix} \alpha_1\hat{a}_{11}(k) & \alpha_2\hat{a}_{12}(k) & (1-\alpha_1)\hat{x}_1(k) & (1-\alpha_2)\hat{x}_2(k) \\ -0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{19}$$

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.1010581

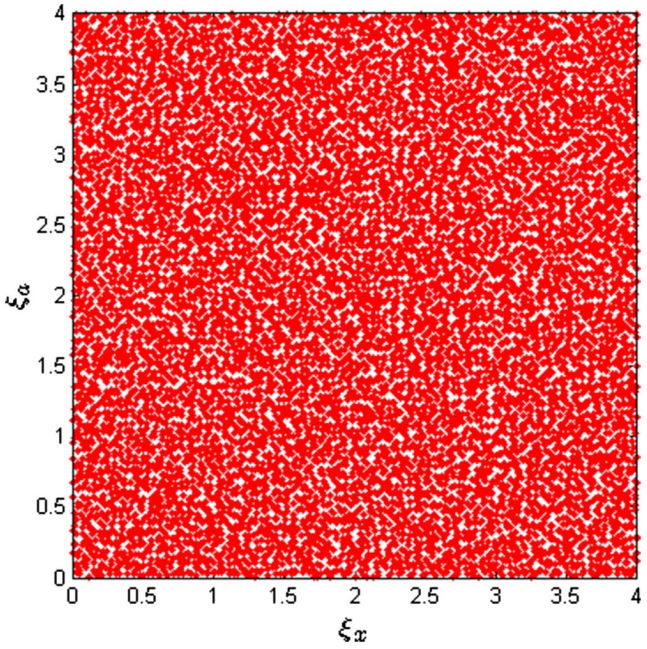


Fig. 2 Application of UKF with Eq. (18) to example 2; UKF with Eq. (18) is applied with 10,000 randomly generated initial estimates $(\hat{x}_2(0), \hat{a}_{11}(0), \hat{a}_{12}(0))$ using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 1000]$; 100% of the trials are red.

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.1010581

for Eqs. (12–14), in which $\alpha \in \mathbb{R}$, and applying UKF with Eq. (19). Note that Eq. (18) corresponds to setting $\alpha_1 = \alpha_2 = 1$. The use of the state-dependent matrix (19) based on the additional parameter a is a standard technique in the literature on the state-dependent Riccati equation [14,15]. However, this technique is ad hoc, and there are no guarantees that it may be effective for a given problem.

To assess the performance of the UKF with Eq. (19), we reconsider the 10,000 randomly generated initial estimates, initial covariance, and tuning parameters as in example 2. Setting $\alpha_1 = \alpha_2 = 0.5$, Fig. 3 shows that, for all 10,000 initial estimates, all of the estimates $\hat{a}(1000)$ are within 10% of the true parameter a.

To test the effect of α_1 and α_2, we consider 11 linearly spaced values of $\alpha_1 \in [-3, 3]$ and 11 linearly spaced values of $\alpha_2 \in [-3, 3]$. For each choice of α_1, α_2, we record the number of 10,000 trials for which the UKF with Eq. (19) estimates a within 10% relative error. Figure 4 shows that, generally, if $\alpha_1 < 1$ and $\alpha_2 < 1$, all of the estimates $\hat{a}(1000)$ are within 10% of the true parameter a. Otherwise, none of the estimates $\hat{a}(1000)$ are within 10% of the true parameter a. This example shows that, compared to example 2, the state-dependent coefficient can significantly improve the performance of the UKF depending on the choice of α_1 and α_2.

In all subsequent UKF examples, we set $\alpha_1 = \cdots = \alpha_p = 0.5$, in which p is the number of unknown entries in A.

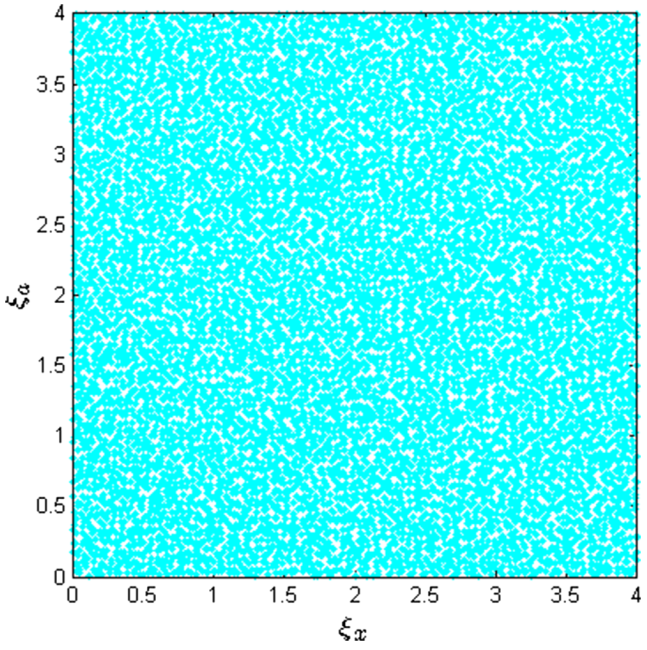


Fig. 3 Application of UKF with Eq. (19) to example 3; UKF with Eq. (19) is applied with 10,000 randomly generated initial estimates $(\hat{x}_2(0), \hat{a}_{11}(0), \hat{a}_{12}(0))$ using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 1000]$; 100% of the trials are cyan.

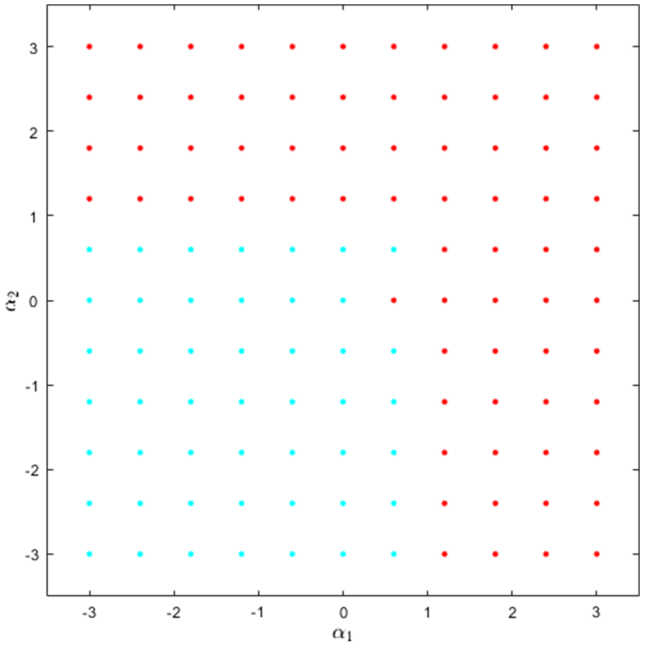


Fig. 4 Application of UKF with Eq. (19) to example 4; α_1 and α_2 are varied from −3 to 3 in increments of 0.6; for each pair of values (α_1, α_2) and 10,000 randomly generated initial estimates $(\hat{x}_2(0), \hat{a}_{11}(0), \hat{a}_{12}(0))$, UKF with Eq. (19) is applied using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 1000]$; a cyan dot indicates that UKF estimates both components of a within 10% relative error at step $k = 1000$ in 100% of the trials, and a red dot indicates that UKF estimates neither of the components of a within 10% relative error in 100% of the trials; all of the trials are either cyan or red.

B. Example 4: $n = 3$ and One Unknown Entry

Consider Eqs. (1–3) with

$$A = \begin{bmatrix} 0.51 & -0.285 & 0.05 \\ -0.012 & 0.34 & 1 \\ 0.03 & -0.88 & 0.34 \end{bmatrix}, \qquad x_0 = \begin{bmatrix} -23 \\ 67 \\ -31 \end{bmatrix}, \qquad E = [1 \quad 0 \quad 0] \tag{20}$$

and assume that one entry in the first row of A is unknown. To apply the UKF, we define the augmented system (12–14) with $\tilde{A}$ constructed as in Eq. (19) and X, $\tilde{E}$ constructed as in Eq. (15). Let $\hat{x}_1(k)$, $\hat{x}_2(k)$, $\hat{x}_3(k)$ be estimates of $x_1(k)$, $x_2(k)$, $x_3(k)$, and, for $i \in \{1, 2, 3\}$, let $\hat{a}_{1i}(k)$ be an estimate of a_{1i}. Define

$$\xi_x \triangleq \frac{\|\hat{x}_u(0) - x_u(0)\|}{\|x_u(0)\|}, \qquad \xi_a \triangleq \frac{\|\hat{a}_{1i}(0) - a_{1i}\|}{\|a_{1i}\|} \tag{21}$$

in which the unmeasured states and their estimates are defined by

$$x_u \triangleq \begin{bmatrix} x_2 \\ x_3 \end{bmatrix}, \qquad \hat{x}_u \triangleq \begin{bmatrix} \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} \tag{22}$$

Using the same UKF tuning parameters as in example 3, we consider 10,000 randomly generated initial estimates $(\hat{x}_2(0), \hat{x}_3(0), \hat{a}_{1i}(0))$ such that $\xi_x, \xi_a \in (0, 2)$. Figure 5a shows that 72.53% of the estimates $\hat{a}_{11}(1000)$ are within 10% of the true parameter a_{11}. In contrast, Figs. 5b and 5c show that 4.68 and 6.09% of the estimates $\hat{a}_{12}(1000)$ and $\hat{a}_{13}(1000)$ are within 10% of the true parameters a_{12} and a_{13}, respectively.

Note that both examples 3 and 4 involve a total of three unknown quantities in A and x_0. It is thus reasonable to expect that the performance of the UKF would be similar for both examples. However, example 3 involves two unknown constants and one unmeasured state, whereas example 4 involves one unknown constant and two unmeasured states. This distinction is consistent with the fact that the UKF performs worse for example 4 than for example 3.

C. Example 5: $n = 3$ and Three Unknown Entries in a Single Row

We revisit example 4 by assuming that all of the entries in the first row of A are jointly unknown. To apply the UKF, we define the augmented system (12–14) with $\tilde{A}$ constructed as in Eq. (19) and X, $\tilde{E}$ constructed as in Eq. (15). Let $\hat{x}_1(k)$, $\hat{x}_2(k)$, $\hat{x}_3(k)$ denote estimates of $x_1(k)$, $x_2(k)$, $x_3(k)$, and let $\hat{a}_{11}(k)$, $\hat{a}_{12}(k)$, $\hat{a}_{13}(k)$ denote estimates of a_{11}, a_{12}, a_{13}. Define the true parameter vector a, its estimate $\hat{a}$, the unmeasured state x_u, and its estimate $\hat{x}_u$ as

$$x_u \triangleq \begin{bmatrix} x_2 \\ x_3 \end{bmatrix}, \qquad \hat{x}_u \triangleq \begin{bmatrix} \hat{x}_2 \\ \hat{x}_3 \end{bmatrix}, \qquad a \triangleq \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \end{bmatrix}, \qquad \hat{a} \triangleq \begin{bmatrix} \hat{a}_{11} \\ \hat{a}_{12} \\ \hat{a}_{13} \end{bmatrix} \tag{23}$$

As in the case of example 4 and using the same tuning parameters for the UKF with Eq. (19), we consider 10,000 randomly generated initial estimates $(\hat{x}_2(0), \hat{x}_3(0), \hat{a}_{11}(0), \hat{a}_{12}(0), \hat{a}_{13}(0))$, such that $\xi_x, \xi_a \in (0, 4)$.

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.I010581

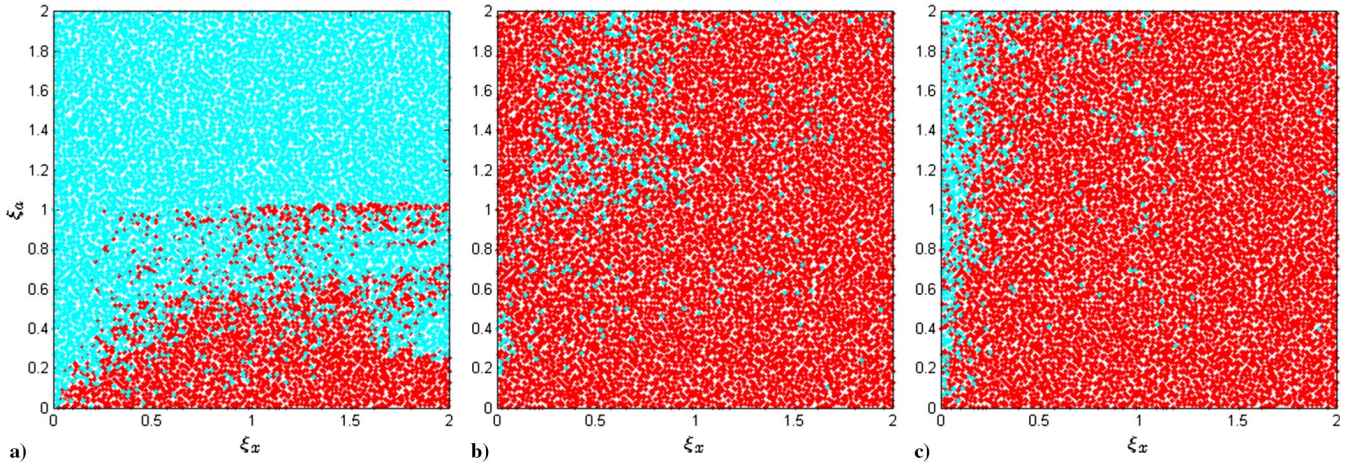


Fig. 5 Application of UKF with Eq. (19) to example 4: a) UKF with Eq. (19) is applied with 10,000 randomly generated initial estimates ($\hat{x}_2(0)$, $\hat{x}_3(0)$, $\hat{a}_{11}(0)$); b) UKF with Eq. (19) is applied with 10,000 randomly generated initial estimates ($\hat{x}_2(0)$, $\hat{x}_3(0)$, $\hat{a}_{12}(0)$); c) UKF with Eq. (19) is applied with 10,000 random initial estimates ($\hat{x}_2(0)$, $\hat{x}_3(0)$, $\hat{a}_{13}(0)$); in all three cases, KF with Eq. (19) uses the measurements $y_0(k) = x_1(k)$ for $k \in [0, 1000]$; 72.53, 4.68, and 6.09% of the trials in a), b), and c), respectively, are cyan.

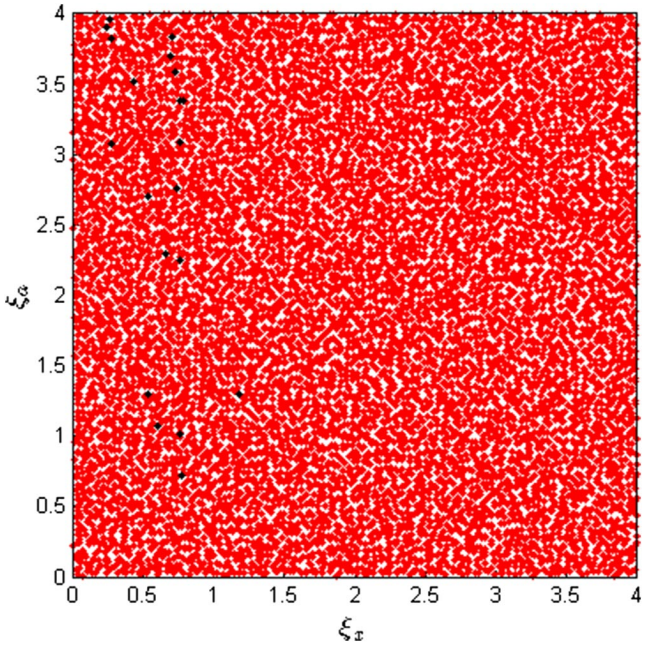


Fig. 6 Application of UKF with Eq. (19) to example 5; KF with Eq. (19) is applied with 10,000 randomly generated initial estimates ($\hat{x}_2(0)$, $\hat{x}_3(0)$, $\hat{a}_{11}(0)$, $\hat{a}_{12}(0)$, $\hat{a}_{13}(0)$) using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 1000]$; 0.20% of the trials are black and 99.80% of the trials are red.

Figure 6 shows that the UKF with Eq. (19) estimates at least one component of a within 10% error in 0.20% of the trials and none of the components of a within 10% error in 99.80% of the trials. Note that, whereas examples 4 and 5 concern the same unknown entries, the three entries in example 5 are estimated concurrently, whereas the three entries in example 4 are estimated separately, assuming the remaining entries are known. This distinction is consistent with the fact that the UKF with Eq. (19) performs worse for example 5 than for example 4.

D. Deficiencies of the UKF

Examples 3–5 suggest that, although the UKF with Eq. (19) can achieve a reasonably accurate parameter estimation for CSPE with $n = 2$, the performance deteriorates drastically for CSPE with $n = 3$. This motivates the need to develop parameter estimation algorithms that are more effective for CSPE problems with $n \geq 3$.

VI. Subsystem Estimation Framework

Consider the main system G shown in Fig. 7 with the realization:

$$x(k+1) = A_0 x(k) + Bu(k) + Dw(k) \tag{24}$$

$$y(k) = Cx(k) \tag{25}$$

$$y_0(k) = Ex(k) \tag{26}$$

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.1010581

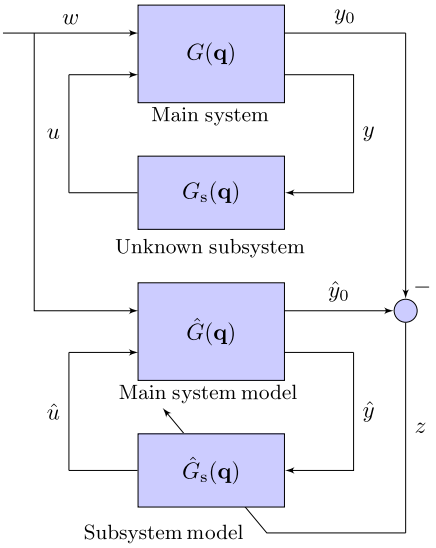


Fig. 7 Subsystem estimation framework for RCSE.

in which $x(k) \in \mathbb{R}^{l_x}$ is the main system state, $y(k) \in \mathbb{R}^{l_y}$ is the main system output, $u(k) \in \mathbb{R}^{l_u}$ is the main system input, $w(k) \in \mathbb{R}^{l_w}$ is the known excitation signal, and $y_0(k) \in \mathbb{R}^{l_z}$ is the main system measurement. The matrix A_0 is the nominal dynamic matrix. The main system (24–26) is interconnected with the unknown subsystem G_s modeled by

$$x_s(k+1) = A_s x_s(k) + B_s y(k) \tag{27}$$

$$u(k) = C_s x_s(k) + D_s y(k) \tag{28}$$

in which $x_s(k) \in \mathbb{R}^{l_{x_s}}$ is the unknown subsystem state. Together, Eqs. (24–28) represent the true system.

Next, the main system model $\hat{G}$ has the realization:

$$\hat{x}(k+1) = A_0\hat{x}(k) + B\hat{u}(k) + Dw(k) \tag{29}$$

$$\hat{y}(k) = C\hat{x}(k) \tag{30}$$

$$\hat{y}_0(k) = E\hat{x}(k) \tag{31}$$

in which $\hat{x}(k) \in \mathbb{R}^{l_x}$ is the main system model state, $\hat{y}(k) \in \mathbb{R}^{l_y}$ is the main system model output, $\hat{u}(k) \in \mathbb{R}^{l_u}$ is the main system model input, and $\hat{y}_0(k) \in \mathbb{R}^{l_z}$ is the main system model measurement. The main system model is interconnected with the subsystem model:

$$\hat{u}(k) = \hat{G}_s(\boldsymbol{q})\hat{y}(k) \tag{32}$$

in which $\boldsymbol{q}$ is the forward shift operator. Equations (29–32) together represent the modeled system. The subsystem estimation problem is represented by the block diagram in Fig. 7, in which the goal is to estimate the subsystem model $\hat{G}_s$ by minimizing a cost function based on the performance variable:

$$z(k) \triangleq \hat{y}_0(k) - y_0(k) \in \mathbb{R}^{l_z} \tag{33}$$

For the subsystem estimation problem, we assume that the unknown subsystem input y and the unknown subsystem output u are not measured, and thus, G_s is inaccessible. The input $\hat{y}$ of the subsystem model $\hat{G}_s$ is computed, and the input $\hat{u}$ of the main system model $\hat{G}$ is estimated. Then, $\hat{u}$ and $\hat{y}$ are used to construct $\hat{G}_s$, which is an estimate of G_s.

For parameter estimation, we assume that $G_s = D_s$ is static, and thus, Eqs. (27) and (28) become

$$u(k) = D_s y(k) \tag{34}$$

In this case, x satisfies

$$x(k+1) = Ax(k) + Dw(k) \tag{35}$$

in which the dynamic matrix of the true system is given by

$$A = A_0 + BD_sC \tag{36}$$

Note that the decomposition (36) represents the matrix A in examples 1–5, in which the uncertain entries of A are the entries of D_s and the corresponding entries of A_0 are set to zero. However, Eq. (36) can be used to model uncertain entries in A with nonzero nominal values, in which case each entry of D_s represents an offset from the nominal value. Consequently, the nominal values of the uncertain entries of A, which are given

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.I010581

by the corresponding entries of A_0, can be viewed as estimates of the uncertain entries of A that correspond to $D_s = 0$. Finally, if $w = 0$, then Eqs. (24–28) are equivalent to the CSPE problem (1–3). We thus assume for the remainder of this paper that $w = 0$.

Let p be the number of uncertain entries in A; let q be the number of rows of A, in which they appear; and let r be the number of columns of A, in which they appear. The expression (36) can be used to represent uncertain entries in A if and only if $p = qr$. This condition is equivalent to saying that, by reordering the rows and columns of A, the uncertain entries of A form a square or rectangular block of A. For example, uncertainty in a_{11} and a_{13} for a third-order system, which corresponds to $p = 2$, $q = 1$, and $r = 2$, can be represented using

$$B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad D_s \in \mathbb{R}^{1\times 2} \tag{37}$$

However, uncertainty in a_{11} and a_{23}, which corresponds to $p = 2$, $q = 2$, and $r = 2$, cannot be represented by Eq. (36). In the case in which $p \neq qr$, Eq. (36) can be replaced by

$$A = A_0 + [\,B_1 \quad \cdots \quad B_l\,] \begin{bmatrix} D_{s,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & D_{s,l} \end{bmatrix} \begin{bmatrix} C_1 \\ \vdots \\ C_l \end{bmatrix} = A_0 + \mathcal{B}\mathcal{D}_s\mathcal{C} \tag{38}$$

in which $l \geq 2$. Note that $\mathcal{D}_s$ in Eq. (38) has a block-diagonal structure, and thus, the estimation of $\mathcal{D}_s$ entails the estimation of $D_{s,1}, \ldots, D_{s,l}$ and all of the off-block-diagonal zero entries. In this case, we treat the block-diagonal matrix as fully populated, and we ignore the estimates of the off-block-diagonal entries, which are known to be zero.

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.I010581

VII. Retrospective Cost Subsystem Estimation

In this section, we formulate the RCSE algorithm for parameter estimation.

A. Subsystem Model

For static parameter estimation, the subsystem model is given by

$$\hat{u}(k) = \hat{D}_s(k)\hat{y}(k) \tag{39}$$

We rewrite Eq. (39) as

$$\hat{u}(k) = \Phi(k)\hat{\theta}(k) \tag{40}$$

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

$$\Phi(k) \triangleq \hat{y}(k)^T \otimes I_{l_u} \in \mathbb{R}^{l_u \times l_\theta} \tag{41}$$

and the unknown entries of A are written as

$$\hat{\theta}(k) \triangleq \mathrm{vec}(\hat{D}_s(k)) \in \mathbb{R}^{l_\theta} \tag{42}$$

in which $l_\theta \triangleq l_u l_y$, $\otimes$ is the Kronecker product, and vec is the column-stacking operator.

B. Retrospective Performance Variable

We define the retrospective input:

$$\tilde{u}(k-1) = \Phi(k-1)\hat{\theta} \tag{43}$$

and the corresponding retrospective performance variable

$$\hat{z}(k) \triangleq z(k) + \Phi_f(k-1)\hat{\theta} - \hat{u}_f(k-1) \tag{44}$$

in which $\hat{\theta} \in \mathbb{R}^{l_\theta}$ is determined by optimization, and $\Phi_f(k-1) \in \mathbb{R}^{l_z \times l_\theta}$ and $\hat{u}_f(k-1) \in \mathbb{R}^{l_z}$ are filtered versions of $\Phi(k-1)$ and $\hat{u}(k-1)$, respectively, defined by

$$\Phi_f(k-1) \triangleq G_f(\boldsymbol{q})\Phi(k-1), \qquad \hat{u}_f(k-1) \triangleq G_f(\boldsymbol{q})\hat{u}(k-1) \tag{45}$$

The filter G_f has the form:

$$G_f(\boldsymbol{q}) \triangleq D_f^{-1}(\boldsymbol{q})N_f(\boldsymbol{q}) \tag{46}$$

in which D_f and N_f are polynomial matrices, and D_f is monic. The choice of these filters is discussed next.

C. Retrospective Cost Function

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

$$J(k,\hat{\theta}) \triangleq \sum_{i=1}^{k} \hat{z}^T(i) R_z \hat{z}(i) + \Big(\hat{\theta} - \theta(0)\Big)^T R_\theta \Big(\hat{\theta} - \theta(0)\Big) \tag{47}$$

in which R_z and R_θ are positive definite. The following result is a restatement of standard recursive least-squares optimization. The update equation (49) can be viewed as a Riccati equation for the discrete-time KF in the case, in which the dynamic matrix is the identity and the output matrix is a data regressor.

Proposition: Let $P(0) = R_\theta^{-1}$. Then, for all $k \geq 1$, the retrospective cost function (47) has a unique global minimizer $\theta(k)$, which is given by

$$\hat{\theta}(k) = \hat{\theta}(k-1) - P(k-1)\Phi_f^T(k-1)\Gamma^{-1}(k-1)\Big[\Phi_f(k-1)\hat{\theta}(k-1) + z_f(k) - u_f(k-1)\Big] \tag{48}$$

$$P(k) = P(k-1) - P(k-1)\Phi_f^T(k-1)\Gamma^{-1}(k-1)\Phi_f(k-1)P(k-1) \tag{49}$$

in which

$$\Gamma(k-1) \triangleq R_z^{-1} + \Phi_f(k-1)P(k-1)\Phi_{f-1}^T(k-1) \tag{50}$$

D. Online Update of G_f

Note that the retrospective performance variable (44) can be rewritten as

$$\hat{z}(k) = z(k) - G_f(\boldsymbol{q})\hat{\mu}(k-1) \tag{51}$$

in which

$$\hat{\mu}(k-1) \triangleq \hat{u}(k-1) - \tilde{u}(k-1) \tag{52}$$

The signal $\hat{\mu}$ can be viewed as a virtual exogenous input, as shown in Fig. 8.

It can be seen from Eq. (51) that $\hat{z}$ is the residual of the fit between z and the output of G_f with input $\hat{\mu}$. However, the actual transfer function from $\hat{\mu}$ to z is given by

$$\tilde{G}_{\hat{y}_0\hat{\mu}}(\boldsymbol{q}) \sim \left[\begin{array}{cc} A_0 + B\hat{D}_s C & B \\ E & 0 \end{array}\right] \tag{53}$$

Consequently, minimizing $\hat{z}$ produces the value of $\hat{\theta}$, and thus, the value of $\hat{D}_s$ that optimally fits $\tilde{G}_{\hat{y}_0\hat{\mu}}$ to G_f. Therefore, a desirable choice of G_f is

$$\tilde{G}^*(\boldsymbol{q}) \sim \left[\begin{array}{cc} A_0 + BD_s C & B \\ E & 0 \end{array}\right] \tag{54}$$

Because D_s is unknown, however, Eq. (54) cannot be implemented in practice. Thus, in all subsequent applications of RCSE, we use the time-varying filter:

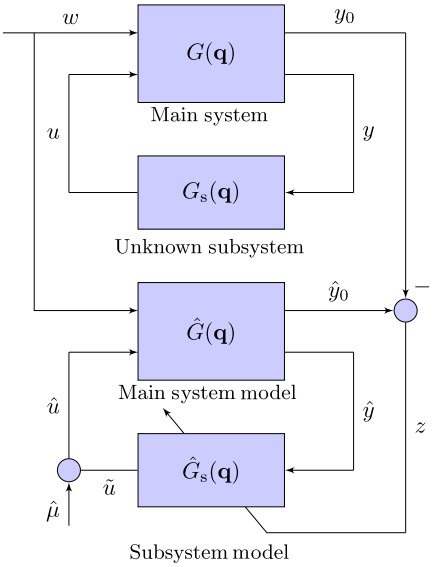


Fig. 8 Subsystem estimation framework showing the virtual exogenous input $\hat{\mu}$.

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.I010581

$$G_f(\boldsymbol{q}, \hat{D}_s(k-1)) \sim \begin{bmatrix} A_0 + B\hat{D}_s(k-1)C & B \\ E & 0 \end{bmatrix} \tag{55}$$

Note that, if $\hat{D}_s(k-1) = D_s$, then $G_f(\boldsymbol{q}, D_s) = \tilde{G}^*(\boldsymbol{q})$.

E. Data-Window Reiteration

To enhance the accuracy of the estimate $\hat{D}_s(k)$ of D_s, RCSE is applied multiple times to a given data set consisting of k_f data points. In the first iteration, we apply RCSE with $G_f(\boldsymbol{q}, \hat{D}_s(k-1))$ given by Eq. (55) initialized with $\hat{D}_s(0) = 0$. In addition, the entries of the nominal dynamic matrix A_0 in both Eq. (55) and the model (29) are set to the initial estimates of the unknown parameters. In subsequent iterations, we apply RCSE to the same data set with $\hat{D}_s(0)$ given by $\hat{D}_s(k_f)$ from the previous iteration and with A_0 replaced by $A_0 + B\hat{D}_s(k_f)C$.

VIII. RCSE with a Known Initial State

In this section, we apply RCSE to the CSPE problem assuming the initial state is known. This assumption is removed in the next section.

A. Example 6: $n = 3$ and Three Unknown Entries in a Single Row

We revisit example 5 with RCSE assuming that the initial state is known. We thus set $\hat{x}(0) = x_0$, which implies $\xi_x = 0$, and we choose 100 initial estimates $(\hat{a}_{11}(0), \hat{a}_{12}(0), \hat{a}_{13}(0))$, such that $\xi_a \in (0, 2)$. For all trials, the tuning parameters are $R_\theta = I_{l_\theta}$, $k_f = 100$, and $n_u = 4$. Figure 9 shows that, in all trials, the RCSE estimates of both components of a are within 10% error. Figure 10 shows how the RCSE estimates evolve for the case, in which $\xi_a = 0.53$. As described in Sec. VII.E, RCSE is applied to the same $k_f = 100$ data points multiple times. By the second application, RCSE is able to accurately estimate all three unknown parameters.

Next, the tuning parameters are changed to $R_\theta = 0.1I_{l_\theta}$ and $R_\theta = 10I_{l_\theta}$, thus spanning two orders of magnitude. Figures 11 and 12 show that the accuracy of RCSE is unchanged.

B. Example 7: $n = 8$ and Eight Unknown Entries in a Single Row

Consider Eqs. (1–3) with

$$A = \begin{bmatrix} 0.29 & 0.43 & 0.26 & 1.6 & 0.22 & -1.02 & -0.35 & -1.31 \\ 0.04 & 0.57 & 0.56 & 0.92 & -0.81 & -0.12 & 0.13 & -0.9 \\ 0.14 & 0.49 & 1.43 & 0.55 & -0.22 & -0.71 & -0.53 & -1.05 \\ -0.33 & -0.12 & -0.31 & -1.18 & 0.77 & 0.34 & 0.72 & 1.3 \\ -0.59 & 0.51 & 0.32 & 0.97 & 0.31 & -0.06 & -0.45 & -0.89 \\ 0.49 & -0.48 & -1.19 & -2.08 & 0.55 & 1.36 & 0.43 & 1.88 \\ 0.16 & -0.48 & -1.39 & -1.68 & 0.58 & 0.8 & 1.12 & 1.98 \\ 0 & 0.6 & 0.2 & 0.27 & -0.21 & -0.27 & -0.83 & 0.27 \end{bmatrix}, \quad x_0 = \begin{bmatrix} -23 \\ 67 \\ -31 \\ 5 \\ 44 \\ -81 \\ 41 \\ -17 \end{bmatrix} \quad E = [1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0] \tag{56}$$

and assume that the entries $a_{11} = 0.29$, $a_{12} = 0.43$, $a_{13} = 0.26$, $a_{14} = 1.6$, $a_{15} = 0.22$, $a_{16} = -1.02$, $a_{17} = -0.35$, and $a_{18} = -1.31$ of A are unknown. Define

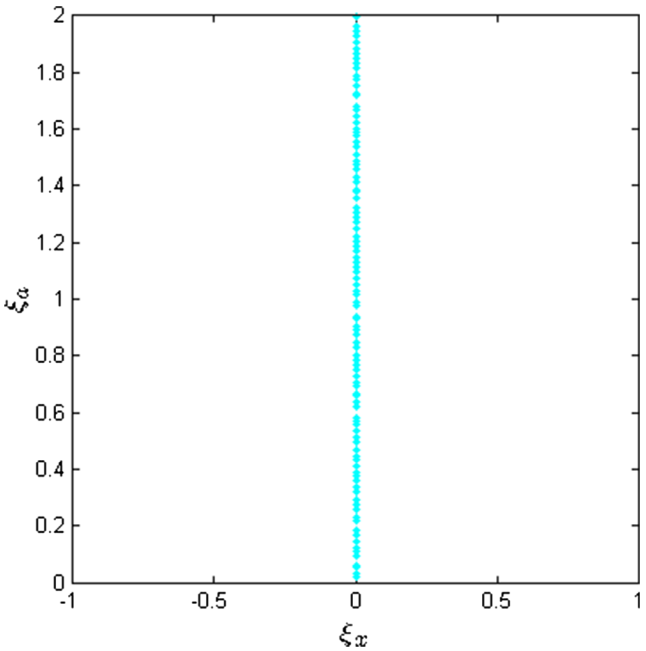


Fig. 9 Application of RCSE to example 6 assuming the initial state is known; RCSE is applied with 100 randomly generated initial estimates $(\hat{a}_{11}(0), \hat{a}_{12}(0)), \hat{a}_{13}(0))$ using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 100]$, and setting $\hat{x}(0) = x_0$ and $R_\theta = I_{l_\theta}$, respectively; 100% of the trials are cyan.

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.1010581

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.1010581

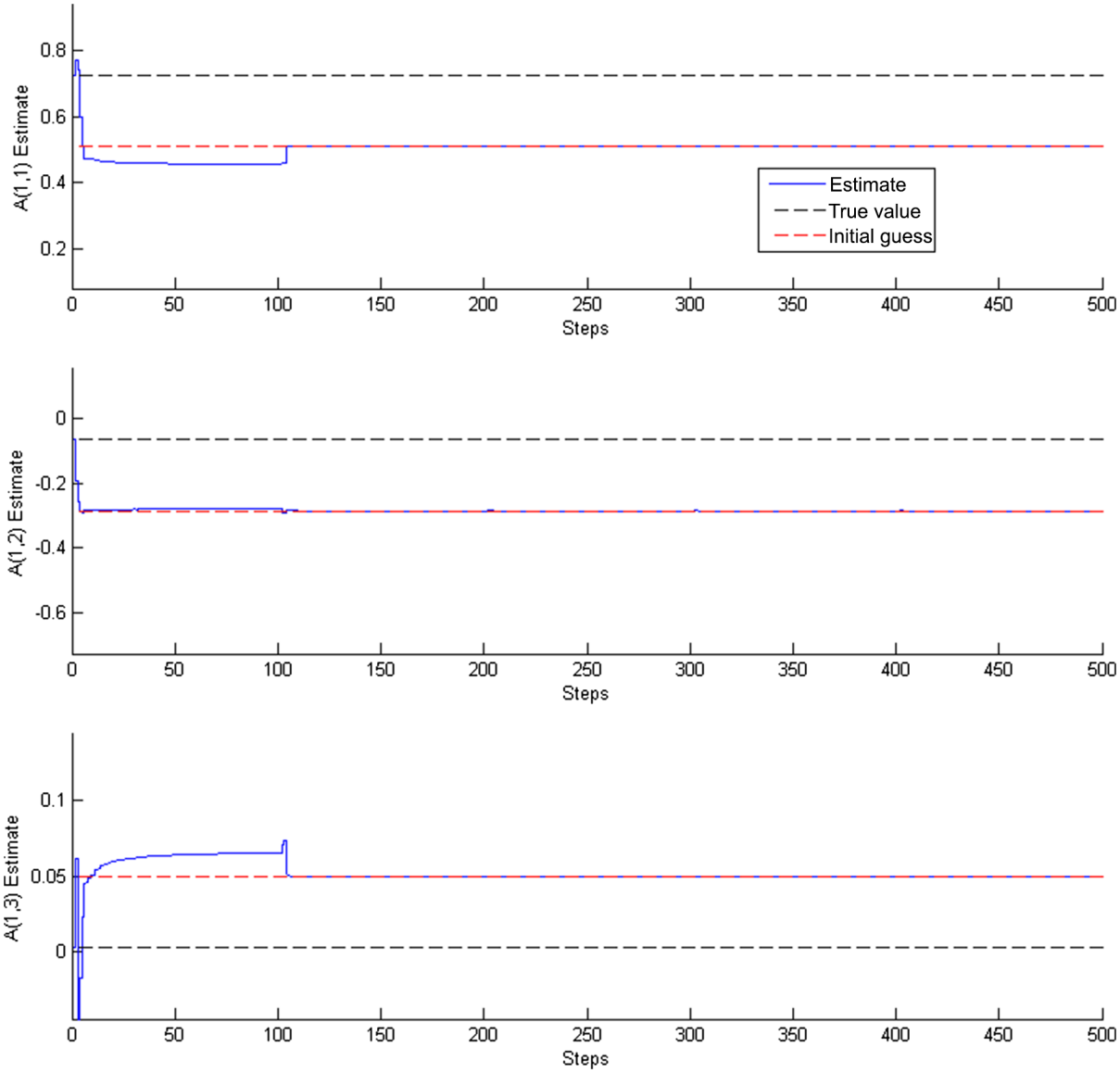


Fig. 10 Evolution of parameter estimates using RCSE for example 6 assuming the initial state is known; RCSE is applied to a case, in which $\xi_a = 0.53$ using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 100]$, and setting $\hat{x}(0) = x_0$ and $R_\theta = I_{l_\theta}$, respectively; RCSE is able to accurately estimate every unknown parameter.

$$x_u \triangleq \begin{bmatrix} x_2 \\ \vdots \\ x_8 \end{bmatrix}, \qquad \hat{x}_u \triangleq \begin{bmatrix} \hat{x}_2 \\ \vdots \\ \hat{x}_8 \end{bmatrix}, \qquad a \triangleq \begin{bmatrix} a_{11} \\ \vdots \\ a_{18} \end{bmatrix}, \qquad \hat{a} \triangleq \begin{bmatrix} \hat{a}_{11} \\ \vdots \\ \hat{a}_{18} \end{bmatrix} \tag{57}$$

Assuming the initial state is known, we set $\hat{x}(0) = x_0$, which implies $\xi_x = 0$, and choose 100 initial estimates $(\hat{a}_{11}(0), \hat{a}_{12}(0), \hat{a}_{13}(0), \hat{a}_{14}(0), \hat{a}_{15}(0), \hat{a}_{16}(0), \hat{a}_{17}(0), \hat{a}_{18}(0))$, such that $\xi_a \in (0, 2)$. For all trials, we use the tuning parameters $R_\theta = 10{,}000 I_{l_\theta}$, $k_f = 25$, and $n_u = 80$. Figure 13 shows that, in 97% of the trials, the RCSE estimates of all of the components of a are within 10% error.

C. Example 8: $n = 2$, Two Unknown Entries in a Single Row, and an Unknown Initial Condition

We revisit example 6 with RCSE assuming the initial conditions are unknown. Because x_1 is measured, we set $\hat{x}_1(0) = x_1(0)$ and choose 10,000 randomly generated initial estimates $(\hat{x}_2(0), \hat{a}_{11}(0), \hat{a}_{12}(0))$, such that $\xi_x, \xi_a \in (0, 2)$. For all trials, we use the tuning parameters $R_\theta = 1$, $k_f = 100$, and $n_u = 4$. Figure 14 shows that, as ξ_x increases, the performance of RCSE degrades. In addition, 31.75% of the estimate $\hat{a}$ is within 10% of both components of the true parameter a, 36.42% is within 10% of exactly one component of a, and 31.83% is within 10% of none of the components of a.

D. Deficiencies of RCSE

Figure 6 suggests that the UKF is not improved using knowledge of the initial conditions. In contrast, examples 6–8 show that, if x_0 is known and we set $\hat{x}(0) = x_0$, then RCSE performs well. Note, however, that the accuracy of RCSE degrades as the uncertainty in x_0 increases. This motivates the development of a variation of RCSE that simultaneously estimates the unknown initial state and the unknown parameters.

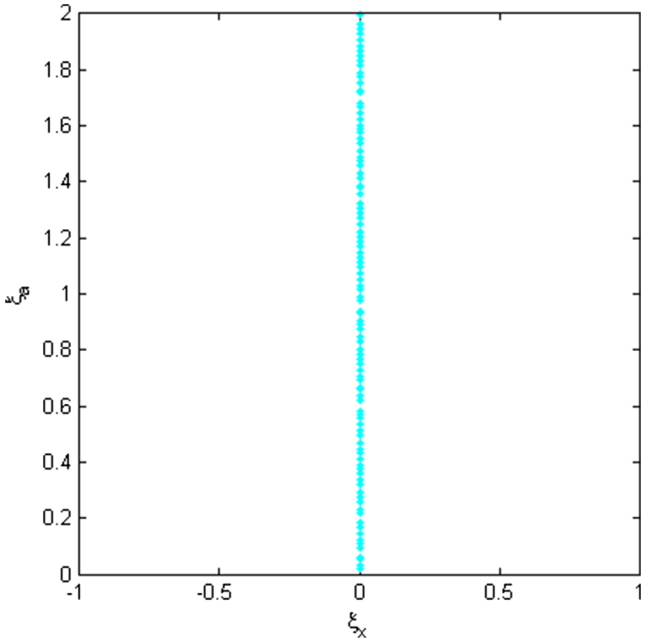


Fig. 11 Application of RCSE to example 6 assuming the initial state is known; RCSE is applied with 100 randomly generated initial estimates $(\hat{a}_{11}(0),\ \hat{a}_{12}(0)),\ \hat{a}_{13}(0))$ using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 100]$, and setting $\hat{x}(0) = x_0$ and $R_\theta = 0.1I_{l_\theta}$, respectively; as in the case of Fig. 9, in which $R_\theta = I_{l_\theta}$, 100% of the trials are cyan.

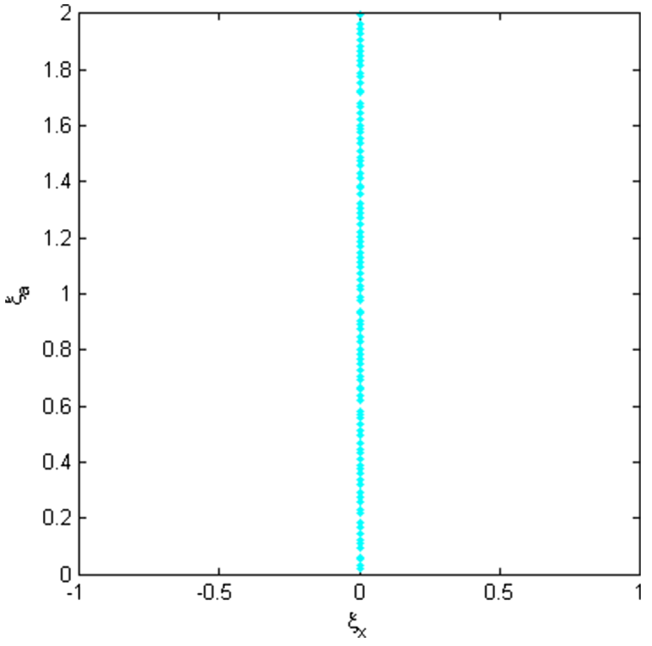


Fig. 12 Application of RCSE to example 6 assuming the initial state is known; RCSE is applied with 100 randomly generated initial estimates $(\hat{a}_{11}(0),\ \hat{a}_{12}(0)),\ \hat{a}_{13}(0))$ using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 100]$, and setting $\hat{x}(0) = x_0$ and $R_\theta = 10I_{l_\theta}$, respectively; as in the case of Fig. 9, in which $R_\theta = I_{l_\theta}$, and Fig. 11, in which $R_\theta = 0.1I_{l_\theta}$, 100% of the trials are cyan.

IX. RCSE Smoother

It was shown in the previous section that the RCSE estimates are reasonably accurate in the case, in which the initial state is known. To take advantage of this observation, we now formulate the RCSES algorithm for simultaneously estimating the unknown parameters and the initial state.

A. Augmented Subsystem Estimation Framework

Let $\delta(k)$ be the unit impulse function, and define $\delta_0(k) \triangleq \delta(k+1)$, which represents a unit impulse at step $k = -1$. Furthermore, define

$$\mathcal{A} \triangleq \begin{bmatrix} A & x_0 \\ 0_{1\times l_x} & 0 \end{bmatrix}, \qquad \mathcal{D} \triangleq \begin{bmatrix} 0_{l_x\times 1} \\ 1 \end{bmatrix}, \qquad \mathcal{E} \triangleq [E \quad 0], \qquad X \triangleq \begin{bmatrix} x \\ \delta_0 \end{bmatrix} \tag{58}$$

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.1010581

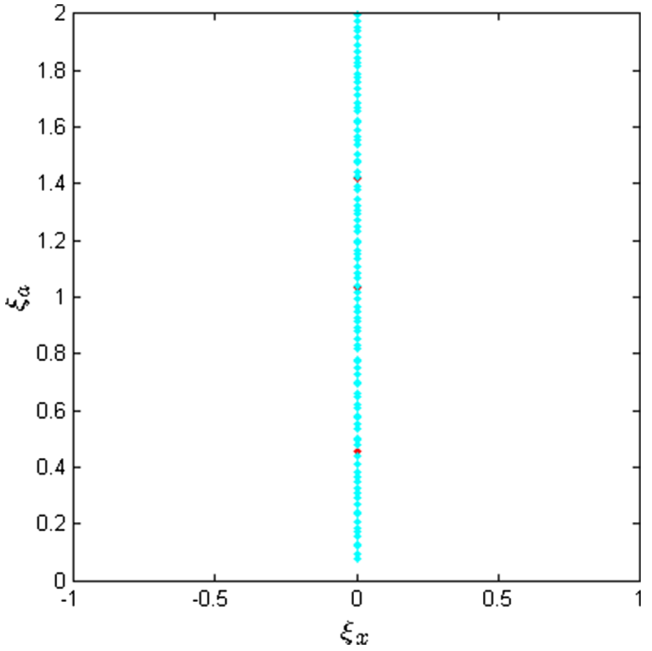


Fig. 13 Application of RCSE to example 7 assuming the initial state is known; RCSE is applied with 100 randomly generated initial estimates ($\hat{a}_{11}(0)$, $\hat{a}_{12}(0)$, $\hat{a}_{13}(0)$, $\hat{a}_{14}(0)$, $\hat{a}_{15}(0)$, $\hat{a}_{16}(0)$, $\hat{a}_{17}(0)$, $\hat{a}_{18}(0)$) using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 25]$ and setting $\hat{x}(0) = x_0$; 97% of the trials are cyan.

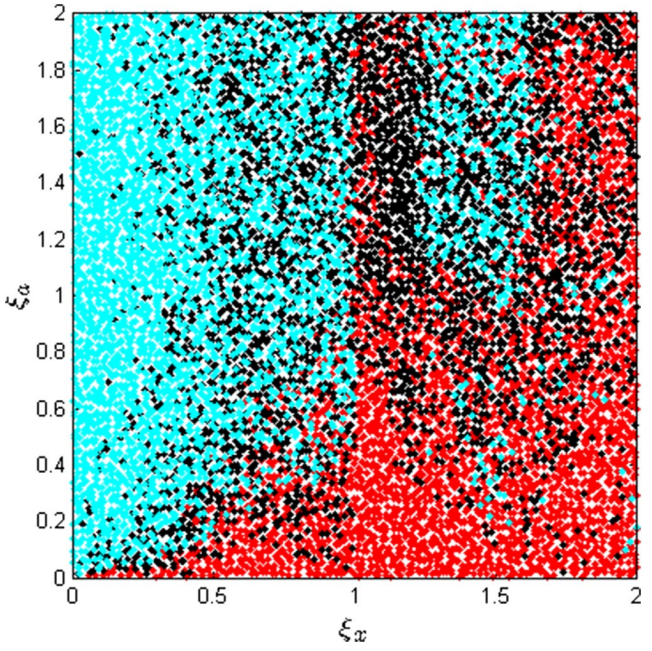


Fig. 14 Application of RCSE to example 8; RCSE is applied with 10,000 randomly generated initial estimates ($\hat{x}_2(0)$, $\hat{a}_{11}(0)$, $\hat{a}_{12}(0)$) using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 100]$; 31.75, 36.42, and 31.83% of the trials are cyan, black, and red, respectively.

Then, for all $k \geq -2$, Eqs. (1–3) can be rewritten as the augmented system:

$$X(k+1) = \mathcal{A}X(k) + \mathcal{D}\delta(k+2) \tag{59}$$

$$X(-2) = 0 \tag{60}$$

$$y_0(k) = \mathcal{E}X(k) \tag{61}$$

Note that k is chosen to begin at step -2 so that $X(0) = [x_0 \quad 0]^T$. Equations (63–65) are therefore a representation of Eqs. (1–3) with known zero initial state and an augmented dynamic matrix $\mathcal{A}$, which includes the uncertain entries of A as well as the unknown components of the initial state x_0.

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.1010581

Now, assume that the mth component of E is 1 and all other components of E are zero, let $\hat{x}_0$ be the nominal initial state, and define

$$\mathcal{A}_0 \triangleq \begin{bmatrix} A_0 & \hat{x}_0 \\ 0_{1\times l_x} & 0 \end{bmatrix}, \qquad \mathcal{B} \triangleq \begin{bmatrix} B & B_0 \\ 0_{1\times l_u} & 0_{1\times(n-1)} \end{bmatrix}, \qquad \mathcal{C} \triangleq \begin{bmatrix} C & 0_{l_y\times 1} \\ 0_{1\times l_x} & 1 \end{bmatrix} \tag{62}$$

in which B_0 is I_n with the mth column removed. Then, Eqs. (59–61) can be written in the form of Eqs. (24–26) as

$$X(k+1) = \mathcal{A}_0 X(k) + \mathcal{B}u(k) + \mathcal{D}\delta(k+2) \tag{63}$$

$$y(k) = \mathcal{C}X(k) \tag{64}$$

$$y_0(k) = \mathcal{E}X(k) \tag{65}$$

with a known initial state $X(-2) = 0$. Using Eq. (34), it follows that the augmented dynamic matrix of the true system is given by

$$\mathcal{A} = \mathcal{A}_0 + \Delta\mathcal{A} = \mathcal{A}_0 + \mathcal{B}\mathcal{D}_s\mathcal{C} \tag{66}$$

Note that Eq. (66) has the same form as Eq. (38) with A replaced by $\mathcal{A}$ and A_0 replaced by $\mathcal{A}_0$. Furthermore, the matrix $\mathcal{D}_s$ in $\Delta\mathcal{A} = \mathcal{B}\mathcal{D}_s\mathcal{C}$ models the uncertain entries of $\mathcal{A}$, which include the uncertain entries of A as well as the unknown components of the initial state x_0 of Eq. (1). Consequently, the estimation of $\mathcal{D}_s$ is a smoothing problem.

To construct an estimator based on Eqs. (63–65), we define

$$\hat{X} \triangleq \begin{bmatrix} \hat{x} \\ \hat{\delta}_0 \end{bmatrix} \tag{67}$$

and rewrite Eqs. (29–31) as

$$\hat{X}(k+1) = \mathcal{A}_0\hat{X}(k) + \mathcal{B}\hat{u}(k) + \mathcal{D}\delta(k+2) \tag{68}$$

$$\hat{y}(k) = \mathcal{C}\hat{X}(k) \tag{69}$$

$$\hat{y}_0(k) = \mathcal{E}\hat{X}(k) \tag{70}$$

in which the initial state $\hat{X}(-2) = 0$. For example, consider the case, in which $n = 2$, a_{11} is unknown, and $y(k) = x_1(k)$, and thus, $x_2(0)$ is unknown. Let the (1, 1) entry of A_0 be zero and set $\hat{x}_2(0) = 0$. Then

$$B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad C = [\,1 \quad 0\,], \qquad B_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \hat{\mathcal{D}}_s = \begin{bmatrix} \hat{\theta}_1 & \hat{\theta}_2 \\ \hat{\theta}_3 & \hat{\theta}_4 \end{bmatrix} \tag{71}$$

in which $\hat{\theta}_1$ and $\hat{\theta}_4$ are estimates of a_{11} and $x_2(0)$, respectively, and $\hat{\theta}_2$ and $\hat{\theta}_3$ are estimates of zero entries that will be ignored. Note that, for this smoother problem, $\mathcal{D}_s$ has the block-diagonal structure shown in Eq. (38).

B. Data Update

For concurrent parameter and initial state estimation, we apply RCSE to Eqs. (68–70). At each step k, RCSE produces $\hat{\mathcal{D}}_s$, which contains estimates of the unknown components of A and x_0. Next, $\hat{y}(k)$ and $\hat{y}_0(k)$ are computed using

$$\hat{y}(k) = \mathcal{C}(\mathcal{A}_0 + \mathcal{B}\hat{\mathcal{D}}_s(k)\mathcal{C})^{k+1}\mathcal{D}, \qquad \hat{y}_0(k) = \mathcal{E}(\mathcal{A}_0 + \mathcal{B}\hat{\mathcal{D}}_s(k)\mathcal{C})^{k+1}\mathcal{D} \tag{72}$$

Because the values of $\hat{y}$ and $\hat{y}_0$ at previous steps are computed from prior estimates of $\hat{A}$ and $\hat{x}(0)$, there may be a mismatch between $P(k-1)$ and $\Phi_f(k-1)$ in Eq. (49). To rectify this, at each step k, we use constant values of $\theta = \theta(k-1)$ to recompute $\hat{y}$, $\hat{y}_0$, $\hat{u}$, and Φ_f from steps -2 to $k-1$. Then, we rerun Eqs. (48) and (49) from steps -2 to k with these updated values to obtain $\theta(k)$ and $P(k)$.

X. RCSES with an Unknown Initial State

In this section, we apply RCSES to the CSPE problem in the case, in which the initial state is unknown.

A. Example 9: $n = 2$ and Two Unknown Entries in a Single Row

We reconsider example 1 using RCSES. In this case, $\mathcal{B}$ and $\mathcal{C}$ are given by Eq. (62) with

$$B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad B_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \hat{\mathcal{D}}_s = \begin{bmatrix} \hat{\theta}_1 & \hat{\theta}_2 & \hat{\theta}_3 \\ \hat{\theta}_4 & \hat{\theta}_5 & \hat{\theta}_6 \end{bmatrix} \tag{73}$$

In this case, $l_\theta = 6$, in which $\hat{\theta}_1$ and $\hat{\theta}_2$ are estimates of a_{11} and a_{12}, respectively; $\hat{\theta}_6$ is an estimate of $x_2(0) - \hat{x}_2(0)$; and $\hat{\theta}_3$, $\hat{\theta}_4$, and $\hat{\theta}_5$ are estimates of zero entries that will be ignored. Note that, for this smoothing problem, $\mathcal{D}_s$ has the block-diagonal structure shown in Eq. (38).

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.I010581

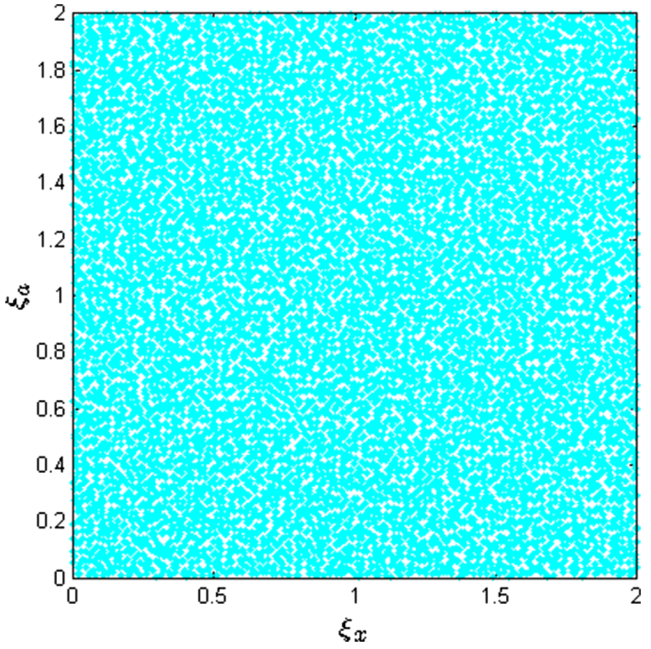


Fig. 15 Application of RCSES to example 9; RCSES is applied with 10,000 randomly generated initial estimates ($\hat{x}_2(0)$, $\hat{a}_{11}(0)$, $\hat{a}_{12}(0)$) using the measurements $y_0(k) = x_1(k)$ for $k \in [0,50]$ and setting $R_\theta = \mathrm{diag}(1,1,10^8,10^8,10^8,1)$; 100% of the trials are cyan.

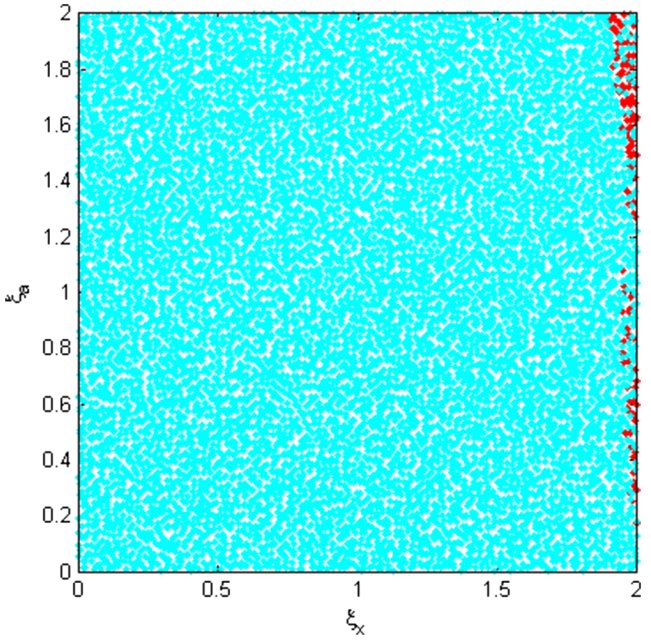


Fig. 16 Application of RCSES to example 9; RCSES is applied with 10,000 randomly generated initial estimates ($\hat{x}_2(0), \hat{a}_{11}(0), \hat{a}_{12}(0)$) using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 50]$ and setting $R_\theta = \mathrm{diag}(0.1, 0.1, 10^7, 10^7, 10^7, 0.1)$; 98.87 and 1.13% of the trials are cyan and red, respectively.

For all trials, we use the tuning parameters $k_f = 50$ and $n_u = 10$, and we choose $R_\theta = \mathrm{diag}(1, 1, 10^8, 10^8, 10^8, 1)$, in which the large entries correspond to the components of θ that are known to be zero. Figure 15 shows that, in all trials, RCSES estimates both components of a within 10% error.

Next, consider the same example with $R_\theta = \mathrm{diag}(0.1,\ 0.1,\ 10^7,\ 10^7,\ 10^7,\ 0.1)$ and $R_\theta = \mathrm{diag}(10,\ 10,\ 10^9,\ 10^9,\ 10^9,\ 10)$, respectively. Figures 16 and 17 demonstrate the insensitivity of RCSES to variations in R_θ.

B. Example 10: $n = 3$ and Three Unknown Entries in a Single Row

We revisit example 6 with RCSES. Once again, the uncertain entries in $\mathcal{A}$ must be represented with Eq. (38). In this case, $l_\theta = 12$, in which three components of θ are estimates of unknown parameters, two components are estimates of the unknown components of the initial state, and seven components are estimates of the known value zero, and thus, are ignored. For all trials, we use the tuning parameters $k_f = 50$, $n_u = 10$, and set $R_\theta = \mathrm{diag}(100, 100, 100, 10^8, 10^8, 10^8, 10^8, 10^{-4}, 10^8, 10^8, 10^8, 10^{-4})$. As in example 9, the largest diagonal entries of R_θ correspond to the components of θ that are zero. Figure 18 shows that 70.92% of the estimate $\hat{a}$ is within 10% of all three components of the true parameter a, 11.62% of $\hat{a}$ is within 10% of at least one component of a, and 17.46% of $\hat{a}$ is within 10% of none of the components of a.

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.1010581

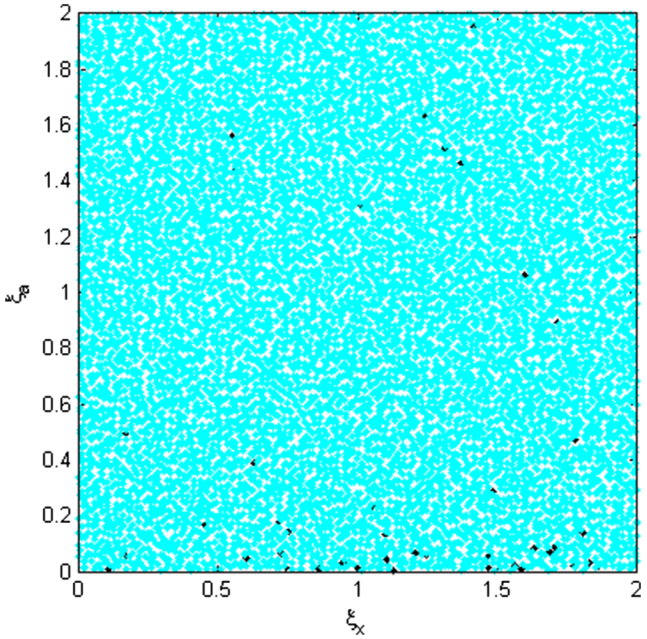


Fig. 17 Application of RCSES to example 9; RCSES is applied with 10,000 randomly generated initial estimates $(\hat{x}_2(0), \hat{a}_{11}(0), \hat{a}_{12}(0))$ using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 50]$ and setting $R_\theta = \mathrm{diag}(10, 10, 10^9, 10^9, 10^9, 10)$; 99.58 and 0.12% of the trials are cyan and black, respectively.

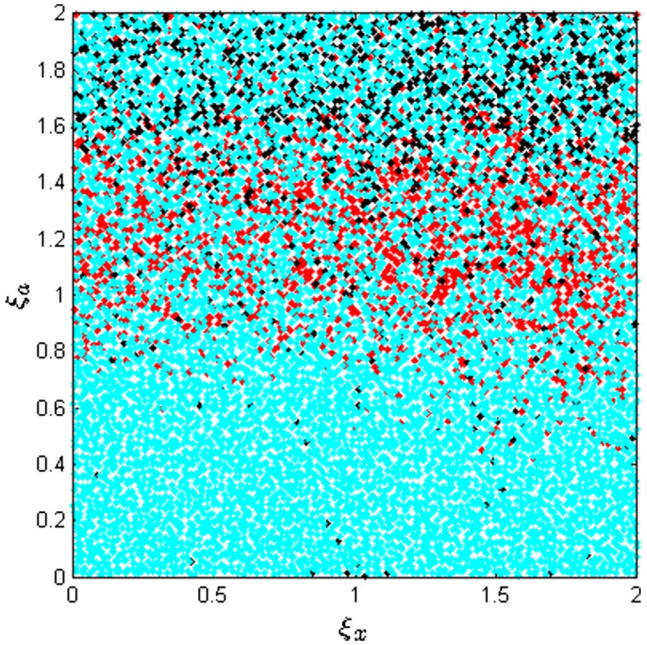


Fig. 18 Application of RCSES to example 10; RCSES is applied with 10,000 randomly generated initial estimates $(\hat{x}_2(0), \hat{x}_3(0), \hat{a}_{11}(0), \hat{a}_{12}(0), \hat{a}_{13}(0))$ using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 50]$; 70.92% of the trials are cyan, 11.62% of the trials are black, and 17.46% of the trials are red.

Examples 10 and 3 show that RCSES performs as well as the UKF with Eq. (19) in the case, in which $n = 2$ and one row of A is unknown. Examples 11 and 5 show that RCSES performs better than the UKF with Eq. (19) for the case, in which $n = 3$ and one row of A is unknown.

XI. Application to Linearized Longitudinal Aircraft Dynamics

In this section, we consider the CSPE problem for linearized longitudinal aircraft dynamics. Consider the continuous-time linearized longitudinal aircraft dynamic matrix:

$$A_c = \begin{bmatrix} -0.0505 & -9.49 & -0.0127 & -32.2 \\ -0.00236 & -2.45 & 0.962 & 0 \\ 0.0179 & -42.0 & -3.44 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{74}$$

Discretizing the dynamics with the time step $T_s = 0.01$ s yields the discrete-time linearized longitudinal aircraft dynamic matrix:

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.1010581

$$A = \begin{bmatrix} 0.999 & -0.0934 & -0.00216 & -0.322 \\ -1.49 \times 10^{-6} & 0.974 & 0.00933 & 2.86 \times 10^{-7} \\ 0.000176 & -0.408 & 0.964 & -2.85 \times 10^{-5} \\ 8.86 \times 10^{-7} & -0.00206 & 0.00982 & 1 \end{bmatrix} \tag{75}$$

Uncertain entries of Eq. (75) can be viewed as arising from uncertain stability derivatives [16].

A. Example 11: UKF with Two Unknown Entries

Consider Eqs. (1–3) with

$$x_0 = \begin{bmatrix} -50 \\ 30 \\ -10 \\ 95 \end{bmatrix}, \qquad E = [1 \quad 0 \quad 0 \quad 0] \tag{76}$$

and assume that the entries $a_{11} = 0.999$ and $a_{12} = -0.0934$ of A are unknown. To apply the UKF, we define the augmented system (12–14) with $\tilde{A}$ constructed as in Eq. (19) and X, $\tilde{E}$ constructed as in Eq. (15). Furthermore, define the true parameter vector a, its estimate $\hat{a}$, the unmeasured state x_u, and its estimate $\hat{x}_u$ as

$$x_u \triangleq \begin{bmatrix} x_2 \\ x_3 \\ x_4 \end{bmatrix}, \qquad \hat{x}_u \triangleq \begin{bmatrix} \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix}, \qquad a \triangleq \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix}, \qquad \hat{a} \triangleq \begin{bmatrix} \hat{a}_{11} \\ \hat{a}_{12} \end{bmatrix} \tag{77}$$

As in example 4, we consider 10,000 randomly generated initial estimates $(\hat{x}_2(0), \hat{x}_3(0), \hat{x}_4(0), \hat{a}_{11}(0), \hat{a}_{12})$ with the UKF, such that ξ_x, $\xi_a \in (0, 2)$. Using the notation of [7], we set the initial covariance matrix to be $P(0) = 10^{-4} I_4$, and choose the tuning parameters $\alpha = 1, \kappa = 0$, $\beta = 2, Q = 10^{-2} I_{2l_x}$, and $R = 0$. Figure 19 shows that 0.04% of the estimate $\hat{a}$ is within 10% of both components of the true parameter a, 2.76% of $\hat{a}$ is within 10% of at least one component of a, and 97.20% of $\hat{a}$ is within 10% of none of the components of a. In most of the trials, in which the estimation of the unknown entries is successful, the estimates converge within approximately 500 time steps, that is, 5 s.

B. Example 12: RCSES with Two Unknown Entries

We revisit example 11 with RCSES. Once again, the uncertain entries in $\mathcal{A}$ must be represented with Eq. (38). In this case, $l_\theta = 12$, two components of θ are estimates of the unknown parameter, three components are estimates of the unknown components of the initial state, and seven components are estimates of the known value zero, and thus, are ignored. For all trials, we use the tuning parameters $k_f = 25$, $n_u = 20$, and set $R_\theta = \text{diag}(0.1, 0.1, 10^8, 10^8, 10^8, 0.1, 10^8, 10^8, 0.1, 10^8, 10^8, 0.1)$, in which the largest entries correspond to the components of θ that are zero. Figure 20 shows that 2.23% of the estimate $\hat{a}$ is within 10% of both components of the true parameter a, 58.79% of $\hat{a}$ is within 10% of at least one component of a, and 38.98% of $\hat{a}$ is within 10% of none of the components of a.

Examples 11 and 12 show that RCSES performs better than the UKF for the case, in which two entries in the dynamic matrix of a linearized longitudinal aircraft model are unknown.

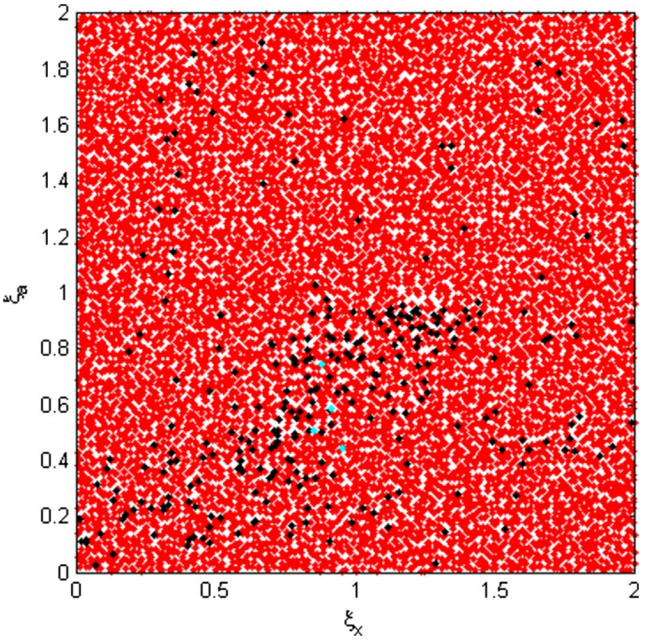


Fig. 19 Application of KF with Eq. (19) to example 11; KF with Eq. (19) is applied with 10,000 randomly generated initial estimates $(\hat{x}_2(0), \hat{x}_3(0), \hat{x}_4(0), \hat{a}_{11}(0), \hat{a}_{12}(0))$ using the measurements $y_0(k) = x_1(k)$ for $k \in [0,1000]$; 0.04% of the trials are cyan, 2.76% of the trials are black, and 97.20% of the trials are red.

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.I010581

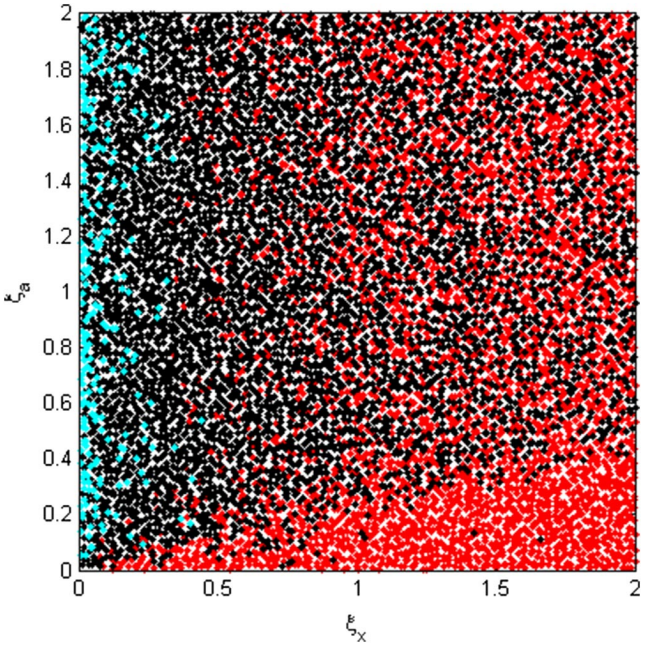


Fig. 20 Application of RCSES to example 12; RCSES is applied with 10,000 randomly generated initial estimates $(\hat{x}_2(0), \hat{x}_3(0), \hat{x}_4(0), \hat{a}_{11}(0), \hat{a}_{12}(0))$ using the measurements $y_0(k) = x_1(k)$ for $k \in [0, 25]$; 2.23% of the trials are cyan, 58.79% of the trials are black, and 38.98% of the trials are red.

XII. Conclusions

Combined state and parameter estimation (CSPE) is a specialized problem in nonlinear estimation, and thus, it is amenable to the standard extended Kalman filter (EKF) and unscented Kalman filter (UKF). However, low-order numerical examples show that the performance of the EKF and UKF is unsatisfactory. With this motivation, retrospective cost subsystem estimation (RCSE) was applied, and it was found that, in the case in which the initial condition is known, it is possible to obtain highly accurate estimates of the unknown entries of the dynamic matrix for both low- and high-order cases. Because the initial condition is usually unknown in practice, the retrospective cost subsystem estimation smoother (RCSES) was developed to estimate the unknown parameters as well as the unknown components of the initial state. RCSES was shown numerically to outperform the EKF and UKF. It is clear, however, that this estimation problem remains challenging, and there is a significant opportunity to refine RCSES and develop alternative methods that can offer an improved performance. Extensions of this problem to include process and sensor noise as considered in [10] are also of interest.

Acknowledgments

This research was supported by NASA under contract number NNX12AM54A, National Science Foundation under grant CMMI 1536834, and U.S. Air Force Office of Scientific Research under grant FA9550-16-1-0071.

References

[1] Crassidis, J. L., and Junkins, J. L., *Optimal Estimation of Dynamic Systems*, 2nd ed., Chapman and Hall/CRC, 2011.
[2] Simon, D., *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*, Wiley-Interscience, New York, 2006, pp. 123–145.
[3] Yu, M.-J., and Bernstein, D. S., "Combined State and Parameter Estimation and the Identifiability of State Space Realizations," *Proceedings of Conference on Decision and Control*, IEEE Publ., Piscataway, NJ, 2016, pp. 3054–3059.
[4] Grewal, M. S., and Glover, K., "Identifiability of Linear and Nonlinear Dynamical Systems," *IEEE Transactions on Automatic Control*, Vol. 21, No. 6, 1976, pp. 833–837.
doi:10.1109/TAC.1976.1101375
[5] Ljung, L., and Glad, T., "On Global Identifiability for Arbitrary Model Parametrizations," *Automatica*, Vol. 30, No. 2, 1994, pp. 265–276.
doi:10.1016/0005-1098(94)90029-9
[6] Ljung, L., "Asymptotic Behavior of the Extended Kalman Filter as a Parameter Estimator for Linear Systems," *IEEE Transactions on Automatic Control*, Vol. 24, No. 1, 1979, pp. 36–50.
doi:10.1109/TAC.1979.1101943
[7] Wan, E. A., and Merwe, R. V. D., "The Unscented Kalman Filter for Nonlinear Estimation," *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control (AS-SPCC) Symposium*, IEEE Publ., Piscataway, NJ, 2000, pp. 153–158.
[8] van der Merwe, R., and Wan, E. A., "The Square-Root Unscented Kalman Filter for State and Parameter-Estimation," *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, IEEE Publ., Piscataway, NJ, 2001, pp. 3461–3464.
[9] Madankan, R., Singla, P., Singh, T., and Scott, P. D., "Polynomial-Chaos-Based Bayesian Approach for State and Parameter Estimations," *Journal of Guidance, Control, and Dynamics*, Vol. 36, No. 4, 2013, pp. 1058–1074.
doi:10.2514/1.58377
[10] Konda, U., Singla, P., Singh, T., and Scott, P. D., "State Uncertainty Propagation in the Presence of Parametric Uncertainty and Additive White Noise," *Journal of Dynamic Systems, Measurement, and Control*, Vol. 133, No. 5, 2011, Paper 051009.
doi:10.1115/1.4004072
[11] D'Amato, A. M., Ridley, A. J., and Bernstein, D. S., "Retrospective-Cost-Based Adaptive Model Refinement for the Ionosphere and Thermosphere," *Statistical Analysis and Data Mining*, Vol. 4, No. 4, 2011, pp. 446–458.
doi:10.1002/sam.v4.4
[12] Ali, A. A., Goel, A., and Bernstein, D. S., "Retrospective-Cost-Based Adaptive Input and State Estimation for the Ionosphere–Thermosphere," *Journal of Aerospace Information Systems*, Vol. 12, No. 12, 2015, pp. 767–783.
doi:10.2514/1.I010286

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.I010581

[13] Chandrasekar, J., Ridley, A. J., and Bernstein, D. S., "An SDRE-Based Asymptotic Observer for Nonlinear Discrete-Time Systems," *Proceedings of the 2005 American Control Conference*, IEEE Publ., Piscataway, NJ, 2005, pp. 3630–3635.
[14] Erdem, E., and Alleyne, A., "Design of a Class of Nonlinear Controllers via State Dependent Riccati Equations," *IEEE Transactions on Control Systems Technology*, Vol. 12, No. 1, 2004, pp. 133–137.
doi:10.1109/TCST.2003.819588
[15] Mracek, C., and Cloutier, J., "Control Designs for the Nonlinear Benchmark Problem via the State-Dependent Riccati Equation Method," *International Journal of Robust and Nonlinear Control*, Vol. 8, Nos. 4–5, 1998, pp. 401–433.
doi:10.1002/(ISSN)1099-1239
[16] Stevens, B. L., and Lewis, F. L., *Aircraft Control and Simulation*, 3rd ed., Wiley, 2015.

E. W. Frew
Associate Editor

Downloaded by UNIVERSITY OF MICHIGAN on September 23, 2018 | http://arc.aiaa.org | DOI: 10.2514/1.I010581