A Stacked Model Structure for Off-line Parameter Variation Estimation in Multi-equilibria Nonlinear Systems

Emmanuel G. Collins, Jr.,
Center for Intelligent Systems, Control and Robotics
Department of Mechanical Engineering
Florida A&M University- Florida State University, Tallahassee, FL 32310
email: ecollins@eng.fsu.edu

Majura F. Selekwa,
Department of Mechanical Engineering and Applied Mechanics
North Dakota State University, Fargo, ND 58105
email: majura@eng.fsu.edu

Rhashan B. Walker
4506 Greenside Walk
Austell, GA 30316
email: rhashan@yahoo.com

Srinivas Palanki
Center for Intelligent Systems, Control and Robotics
Department of Chemical Engineering
Florida A&M University- Florida State University, Tallahassee, FL 32310
email: palanki@eng.fsu.edu

September 14, 2005

Abstract

Preventive diagnostics and maintenance of complex dynamic systems such as chemical plants and aircraft engines benefit from knowledge of the system components. The conditions of the components are often represented by the parameters of the system, in which case the current conditions of the components can be deduced from the variations of the system parameters about their nominal values. This paper considers nonlinear dynamic systems that tend to operate about equilibrium points and focuses on off-line estimation of the variations of important parameters about their nominal values. It uses the common practice of treating the parameter variation estimation vector as a subset of a large state vector, which here is formed by stacking state-space plant descriptions corresponding to selected equilibrium points. A fundamental problem for parameter estimation, no matter which parameter estimation method is used, is that the parameters of interest may not be observable unless data from an appropriate number and combination of equilibrium points is used in the estimation process. Using analysis based on linearized models, this paper establishes criteria for selecting the combinations of equilibrium points that, when stacked, render the parameter vector observable in the sensor data collected about these equilibrium points. A nonlinear filter is then applied to estimate the parameter variation vector as part of the augmented (or stacked) state vector. Numerical simulation results of a Continuously Stirred Tank Reactor verify the theory. In particular, it is seen, as predicted by the theory, that data from one, two or three equilibrium points is not sufficiently rich to yield information about each parameter variation. It is also seen that not all arbitrary combinations of four equilibrium points can provide the information required for estimation of the parameter variation vector; only combinations that satisfy the proposed criteria were able to yield successful estimation of all parameters.

Keywords: Parameter estimation, multi-equilibria systems, parameter observability.

1 Introduction

Preventive diagnostics of complex dynamic systems, for example chemical plants or aircraft engines, require full knowledge of the conditions of the system components. The conditions of the components are often represented by the parameters of the system, in which case the current conditions of the components can be deduced from the variations of the system parameters about their nominal values. Several methods of estimating parameter variations in nonlinear dynamic systems have been established; the most popular fall in two main groups: (a) methods that are based on minimization of the model errors based on past data such as the autoregressive recursive least squares methods [10, 16, 18, 9], and (b) methods that are based on maximization of the model prediction accuracy such as the statistical filtering and maximum likelihood methods [10, 17, 12, 4, 19, 11]. Methods in both groups can be applied on-line or off-line and assume that the parametric structure of the model is known so that the problem is to determine the model parameters themselves.

On-line parameter estimation techniques attempt to match a mathematical model to past and present process measurements while the process is in operation; and off-line parameter estimation techniques attempt to match a mathematical model to the process measurements after the process has run completely. If all parameters vary slowly then it suffices to use off-line estimation for system maintenance.

Methods that minimize the model errors do so by solving for parametric values that minimize the sum of the square of errors between the actual system and the model taken over a certain period of time. Depending on the size of the model and the time frame, this approach can sometimes be computationally intensive [13]. Most of the parameter estimation approaches are those based on maximizing the accuracy of the model prediction by stochastically filtering out the effects of stochastic noise. Assuming that a state space structure of the model is known, the parameters to be estimated are augmented with the state vector; when the system is driven under persistent noise, the state and parameter vectors are estimated by using a filter on the measured outputs of the system. This approach is normally referred to as parameter estimation by state estimation. Any of the available filtering methods can be applied on the model although the most popular methods are based on the standard linear Kalman filter [7] and its derivatives such as the Extended Kalman Filter (EKF) [5, 11] and the Unscented Kalman Filter (UKF) [6, 19].

This paper considers nonlinear dynamic systems that tend to operate about their equilibrium points. A fundamental problem, no matter which parameter estimation method is used, is that the parameters of interest may not be observable unless data from an appropriate number and combination of equilibrium points is used in the estimation process. Hence, using linearized models of the nonlinear system about the various equilibrium points, this paper establishes conditions for selecting combinations of equilibrium points in the stacked system such that the parameter vector is observable. (Preliminary results of this work appear in [2, 3].) It is seen that it is possible to estimate the parameter variations vector using stored measurement data from selected combinations of equilibrium points, provided these points satisfy certain conditions. The plant descriptions about these few selected equilibrium points are concatenated to create one “stacked” model of the plant. Using this stacked model, parameter estimation can be performed using an appropriate nonlinear filter such as the EKF or the UKF. The theory is applied to a Continuously Stirred Tank Reactor (CSTR) using the EKF and is found to work very well in estimating large parameter variations.

The remainder of the paper is as follows. Section 2 describes the fundamental parameter variation estimation problem for a nonlinear system that operates about various equilibrium points. Then, it uses linearized models about the various equilibrium points to provide conditions under which the data about a single equilibrium point or combination of equilibrium points is adequate for parameter variation estimation. Section 3 describes the stacked nonlinear model for parameter variation estimation based on data from a combination of the equilibrium points. Section 4 illustrates application of the analysis technique developed in Section 2 to a CSTR and then demonstrates parameter variation estimation using the stacked model of Section 3 in conjunction with the EKF. Finally, Section 5 presents the conclusions of this study.

Notation

$\mathbb{R}$	Real numbers
$\sigma_i(M)$	i^{th} singular value of M ($\sigma_1(M) \geq \sigma_2(M) \geq \ldots$)
$\sigma_{max}(M)$	Maximum singular value of M
$\sigma_{min}(M)$	Minimum singular value of M (if M is $n \times p$, $\sigma_{min}(M) = \sigma_p(M)$)
$\mathrm{nrow}(M)$	Number of rows of the matrix M
$\mathrm{ncol}(M)$	Number of columns of the matrix M

2 Estimation Problem Statement and Conditions for Solution

The problem addressed by this paper considers a nonlinear dynamic system of the form

$$\dot{\mathcal{X}}(t) = f(\mathcal{X}(t), \mathcal{U}(t), P) + w(t), \quad \mathcal{X}(0) = \mathcal{X}_0 \tag{1}$$

$$\mathcal{Z}(t) = h(\mathcal{X}(t), \mathcal{U}(t), P) + v(t), \tag{2}$$

where $\mathcal{X} \in \mathbb{R}^{n_x}$ is the state, $\mathcal{U} \in \mathbb{R}^{n_u}$ is the control input, $\mathcal{Z} \in \mathbb{R}^{n_z}$ is the measurement, w is the process disturbance, v is the measurement noise, and $P \in \mathbb{R}^{n_p}$ is a constant parameter vector, such that

$$\dot{P} = 0, \tag{3}$$

P_0 is the nominal parameter vector. It is assumed that there exists a set of constant vectors $\{\bar{\mathcal{U}}^{(i)}\}_{i=1}^{m} \subset \mathbb{R}^{n_u}$ where $m \geq 1$, such that $0 = f(\bar{\mathcal{X}}^{(i)}, \bar{\mathcal{U}}^{(i)}, P_0)$ has a solution $\bar{\mathcal{X}}^{(i)}$ for each for each $\bar{\mathcal{U}}^{(i)}$ and and $\frac{\partial f}{\partial \mathcal{X}}|_{\mathcal{X}=\bar{\mathcal{X}}^{(i)}}$ is stable. Furthermore, it is assumed that for each $i = 1, 2, \ldots, m$ the solution $\bar{\mathcal{X}}^{(i)}$ can assume more than one value, i.e., $\bar{\mathcal{X}}^{(i)} \in \left\{ {\bar{\mathcal{X}}_*}^{(i,1)}, {\bar{\mathcal{X}}_*}^{(i,2)}, \ldots, {\bar{\mathcal{X}}_*}^{(i,r_i)} \right\}$. Hence, each $\bar{\mathcal{X}}_*^{(i,j)}, j = 1, 2, \ldots, r_i$ is a stable equilibrium point corresponding to the constant input $\bar{\mathcal{U}}^{(i)}$ and the nominal parameter vector P_0. The nonlinear system (1)-(2) is said to be a *multi-equilibria system*. The number of stable equilibrium points for this systems is defined by $r \triangleq \sum_{i=1}^{m} r_i$.

2.1 Statement of the Estimation Problem

Let the equilibrium points be indexed by the set $\mathcal{L} \triangleq \{1, 2, \ldots, r\}$ such that the set of stable equilibrium points becomes $\{\mathcal{X}_{eq}^{(\ell)}\}_{\ell=1}^{r}$, and the set of corresponding constant inputs becomes $\{\mathcal{U}_c^{(\ell)}\}_{\ell=1}^{r}$. Note that for each $\ell \in \mathcal{L}$ there exists an $i \in \{1, 2, \ldots, m\}$ and a $j \in \{1, \ldots, r_i\}$ such that $\mathcal{U}_c^{(\ell)} = \bar{\mathcal{U}}^{(i)}$ and $\mathcal{X}_{eq}^{(\ell)} = \bar{\mathcal{X}}_*^{(i,j)}$. If the parameter vector P differs significantly from its nominal value P_0, the equilibrium points will also change. The problem studied by this paper is that of using stored data collected about the equilibrium points to estimate the parameter change $\Delta P \triangleq P - P_0$ off-line.

2.2 Linear Representation of the Equilibrium Points

To solve the problem posed in the previous subsection, it is necessary first to establish conditions under which the parameter variation vector ΔP can be determined from the measurements $\mathcal{Z}$. This analysis will be based on linearized models of the system about the various stable equilibrium points.

For $\ell \in \mathcal{L}$ the linearization of (1)-(3) about the equilibrium point denoted by the triple$(\mathcal{X}_{eq}^{(\ell)}, \mathcal{U}_c^{(\ell)}, P_0)$ is given by

$$\begin{bmatrix} \dot{x}^{(\ell)}(t) \\ \dot{p}(t) \end{bmatrix} = \begin{bmatrix} A_c^{(\ell)} & B_{c,P}^{(\ell)} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x^{(\ell)}(t) \\ p(t) \end{bmatrix} + \begin{bmatrix} B_c^{(\ell)} \\ 0 \end{bmatrix} u^{(\ell)}(t) + \begin{bmatrix} I \\ 0 \end{bmatrix} w(t), \quad \begin{bmatrix} x^{(\ell)}(0) \\ p(0) \end{bmatrix} = \begin{bmatrix} 0 \\ \Delta P \end{bmatrix}, \tag{4}$$

$$z^{(\ell)}(t) = \begin{bmatrix} C_c^{(\ell)} & D_{c,P}^{(\ell)} \end{bmatrix} \begin{bmatrix} x^{(\ell)}(t) \\ p(t) \end{bmatrix} + D_c^{(\ell)} u^{(\ell)}(t) + v(t) \tag{5}$$

where

$$x^{(\ell)}(t) \triangleq \mathcal{X}(t) - \mathcal{X}_{eq}^{(\ell)}, \tag{6}$$

$$u^{(\ell)}(t) \triangleq \mathcal{U}(t) - \mathcal{U}_c^{(\ell)}, \tag{7}$$

$$p \triangleq P - P_0 = \Delta P \text{ (a constant)}, \tag{8}$$

and

$$A_c^{(\ell)} \triangleq \frac{\partial f}{\partial \mathcal{X}}\Big|_{(\mathcal{X}_{eq}^{(\ell)}, \mathcal{U}_c^{(\ell)}, P_0)}, \quad B_c^{(\ell)} \triangleq \frac{\partial f}{\partial \mathcal{U}}\Big|_{(\mathcal{X}_{eq}^{(\ell)}, \mathcal{U}_c^{(\ell)}, P_0)}, \tag{9}$$

$$B_{c,P}^{(\ell)} \triangleq \frac{\partial f}{\partial P}\Big|_{(\mathcal{X}_{eq}^{(\ell)}, \mathcal{U}_c^{(\ell)}, P_0)}, \quad C_c^{(\ell)} \triangleq \frac{\partial h}{\partial \mathcal{X}}\Big|_{(\mathcal{X}_{eq}^{(\ell)}, \mathcal{U}_c^{(\ell)}, P_0)}, \tag{10}$$

$$D_c^{(\ell)} \triangleq \frac{\partial h}{\partial \mathcal{U}}\Big|_{(\mathcal{X}_{eq}^{(\ell)}, \mathcal{U}_c^{(\ell)}, P_0)}, \quad D_{c,P}^{(\ell)} \triangleq \frac{\partial h}{\partial P}\Big|_{(\mathcal{X}_{eq}^{(\ell)}, \mathcal{U}_c^{(\ell)}, P_0)}. \tag{11}$$

Assume that the system is sampled at a period T_s. Then, if $w(t)$ and $v(t)$ are white Gaussian random processes, discretization using the Euler-scheme yields [8]

$$\begin{bmatrix} x^{(\ell)}(k+1) \\ p(k+1) \end{bmatrix} = \begin{bmatrix} A^{(\ell)} & B_p^{(\ell)} \\ 0 & I \end{bmatrix} \begin{bmatrix} x^{(\ell)}(k) \\ p(k) \end{bmatrix} + \begin{bmatrix} B_c^{(\ell)} \\ 0 \end{bmatrix} u^{(\ell)}(k) + \sqrt{T_s} \begin{bmatrix} I \\ 0 \end{bmatrix} w(k),$$

$$\begin{bmatrix} x^{(\ell)}(0) \\ p(0) \end{bmatrix} = \begin{bmatrix} 0 \\ \Delta P \end{bmatrix}, \tag{12}$$

$$z^{(\ell)}(k) = \begin{bmatrix} C^{(\ell)} & D^{(\ell)} \end{bmatrix} \begin{bmatrix} x^{(\ell)}(k) \\ p(k) \end{bmatrix} + v(k). \tag{13}$$

where

$$A^{(\ell)} \triangleq I + T_s A_c^{(\ell)}, \quad B_p^{(\ell)} \triangleq T_s B_{c,P}^{(\ell)}, \quad C^{(\ell)} \triangleq C_c^{(\ell)}, \quad D^{(\ell)} \triangleq D_{c,P}^{(\ell)}. \tag{14}$$

Furthermore, if $w(t)$ and $v(t)$ have zero means, then taking the expected value of (12) and (13) and defining $\overline{x}^{(\ell)}(k) \triangleq \mathbf{E}[x^{(\ell)}(k)]$ and $\overline{z}^{(\ell)}(k) \triangleq \mathbf{E}[z^{(\ell)}(k)]$ and assuming that $\mathbf{E}[u^{(\ell)}(k)] = 0$ yields

$$\left[\begin{array}{c} \overline{x}^{(\ell)}(k+1) \\ p(k+1) \end{array}\right] = \left[\begin{array}{cc} A_c^{(\ell)} & B_p^{(\ell)} \\ 0 & (1+T_s)I \end{array}\right] \left[\begin{array}{c} \overline{x}^{(\ell)}(k) \\ p(k) \end{array}\right], \quad \left[\begin{array}{c} x^{(\ell)}(0) \\ p(0) \end{array}\right] = \left[\begin{array}{c} 0 \\ \Delta P \end{array}\right], \tag{15}$$

$$\overline{z}^{(\ell)}(k) = \left[\begin{array}{cc} C^{(\ell)} & D_p^{(\ell)} \end{array}\right] \left[\begin{array}{c} \overline{x}^{(\ell)}(k) \\ p(k) \end{array}\right]. \tag{16}$$

Definition 2.1 *The linear, stochastic system (12)-(13) is called* ***mean value steady-state observable*** *if given* $\overline{z}^(\ell)_{ss} \triangleq \lim_{k\to\infty} \overline{z}^(\ell)(k)$*, it is possible to uniquely determine the steady-state, state vector of (15), i.e.,*

$$\left[\begin{array}{c} \overline{x}_{ss}^{(\ell)} \\ p \end{array}\right] \triangleq \lim_{k\to\infty} \left[\begin{array}{c} \overline{x}^{(\ell)}(k) \\ p \end{array}\right] = \left[\begin{array}{c} \overline{x}^{(\ell)}(k) \\ \Delta P \end{array}\right]. \tag{17}$$

If the system (12)-(13) is mean value steady-state observable, then it is possible to reconstruct $p \triangleq \Delta P$ from the steady-state mean value of the sensor reading. In what follows the intial focus will be on establishing a necessary and sufficient condition for the linearized model about a single operating point to be mean value steady-state observable. This condition may not be satisfied; for example, it is never satisfied if the number of measurement is less than the number of parameters $(n_z < n_p)$. However, in this case it is still possible that by considering the steady-state mean values of the sensor readings at multiple equilibrium points, ΔP can be reconstructed. Hence, a condition is established for mean value steady-state observability of a stacked model consisting of the linearized models about multiple equilibrium points.

2.3 Conditions for Mean Value Steady-State Observability

The conditions presented and proved below are concisely presented in two theorems. The first theorem states necessary and sufficient conditions for mean value steady-state observability for one equilibrium point while the second theorem is a generalization of the first theorem when measurements from several equilibrium points are considered. Before stating the first theorem, define

$$\mathcal{A}^{(\ell)} \triangleq \left[\begin{array}{cc} (A^{(\ell)} - I) & B_p^{(\ell)} \\ C^{(\ell)} & D_p^{(\ell)} \end{array}\right]. \tag{18}$$

Theorem 2.1 *For any $\ell \in \mathcal{L}$ the system (12)-(13) is mean value steady-state observable only if*

$$\mathrm{nrow}(\mathcal{A}^{(\ell)}) \geq \mathrm{ncol}(\mathcal{A}^{(\ell)}), \tag{19}$$

Furthermore, (12)-(13) is mean value steady-state observable if and only if

$$\mathrm{rank}(\mathcal{A}^{(\ell)}) = \mathrm{ncol}(\mathcal{A}^{(\ell)}). \tag{20}$$

Proof. Equations (12) and (13) yield the following steady-state equations

$$x_{ss}^{(\ell)} = A^{(\ell)} x_{ss}^{(\ell)} + B_p^{(\ell)} \Delta P \tag{21}$$

$$z_{ss}^{(\ell)} = C^{(\ell)} x_{ss}^{(\ell)} + D_p^{(\ell)} \Delta P. \tag{22}$$

These are equivalent to

$$\mathcal{A}^{(\ell)} y^{(\ell)} = b^{(\ell)}, \tag{23}$$

where

$$y^{(\ell)} = \begin{bmatrix} x_{ss}^{(\ell)} \\ \Delta P \end{bmatrix}, \;\; b^{(\ell)} = \begin{bmatrix} 0 \\ z_{ss}^{(\ell)} \end{bmatrix}. \tag{24}$$

It is known [1] that (23) can have a solution only if there exists ϕ such that

$$y^{(\ell)} = (\mathcal{A}^{(\ell)})^{\dagger} b^{(\ell)} + [I - (\mathcal{A}^{(\ell)})^{\dagger} \mathcal{A}^{(\ell)}] \phi, \tag{25}$$

where $(\mathcal{A}^{(\ell)})^{\dagger}$ is the Moore-Penrose inverse of $\mathcal{A}^{(\ell)}$. Properties of the Moore-Penrose inverse show that the product $(\mathcal{A}^{(\ell)})^{\dagger} \mathcal{A}^{(\ell)}$ is possible only if (19) is satisfied. Furthermore, the solution is unique if and only if the null space of $\mathcal{A}^{(\ell)}$ is void and hence

$$I - (\mathcal{A}^{(\ell)})^{\dagger} \mathcal{A}^{(\ell)} = 0. \tag{26}$$

In that case,

$$\mathrm{rank}(\mathcal{A}^{(\ell)}) = \min(\mathrm{ncol}(\mathcal{A}^{(\ell)}), \mathrm{nrow}(\mathcal{A}^{(\ell)})), \tag{27}$$

which implies that if (19) is satisfied and (26) is valid, then (20) is also satisfied. Similarly, if (20) is satisfied and (26) is valid, then (19) is also satisfied. □

Remark 2.1 *Note that the condition (20) is equivalent to the basic algebraic condition,*

$$\mathrm{rank}(\mathcal{A}^{(\ell)}) = \mathrm{rank} \begin{bmatrix} \mathcal{A}^{(\ell)} & b^{(\ell)} \end{bmatrix}. \tag{28}$$

Remark 2.2 *The necessary condition (19) is satisfied only if $n_z \geq n_p$. Hence, in theory it is possible to satisfy (19) by adding more sensor outputs. Unfortunately, due to engineering limitations it is often not possible to sufficiently increase the number of sensors such that (19) is satisfied. Even if this were possible, (20) may not be satisfied. When $n_z < n_p$ or (20) is not satisfied, it becomes necessary to use off-line data from several equilibrium points to determine the parameter variation by estimation.*

Now, for some integer $N > 1$ define $\mathcal{L}_N \subseteq \mathcal{L}$ such that $\mathcal{L}_N \triangleq \{\ell_1, \dots, \ell_N\}$. Consider the system (12)-(13) stacked at N operating points $\ell_1, \dots, \ell_N$, i.e.,

$$\begin{bmatrix} x^{(\ell_1)}(k+1) \\ x^{(\ell_2)}(k+1) \\ \vdots \\ x^{(\ell_N)}(k+1) \\ p(k+1) \end{bmatrix} = \begin{bmatrix} A^{(\ell_1)} & 0 & \dots & 0 & B_p^{(\ell_1)} \\ 0 & A^{(\ell_2)} & \dots & 0 & B_p^{(\ell_2)} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & A^{(\ell_N)} & B_p^{(\ell_N)} \\ 0 & 0 & \dots & 0 & I \end{bmatrix} \begin{bmatrix} x^{(\ell_1)}(k) \\ x^{(\ell_2)}(k) \\ \vdots \\ x^{(\ell_N)}(k) \\ p(k) \end{bmatrix}, \quad \begin{bmatrix} x^{(\ell_1)}(0) \\ x^{(\ell_2)}(0) \\ \vdots \\ x^{(\ell_N)}(0) \\ p(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \Delta P \end{bmatrix} \tag{29}$$

$$\begin{bmatrix} z^{(\ell_1)}(k) \\ z^{(\ell_2)}(k) \\ \vdots \\ z^{(\ell_N)}(k) \end{bmatrix} = \begin{bmatrix} C^{(\ell_1)} & 0 & \dots & 0 & D_p^{(\ell_1)} \\ 0 & C^{(\ell_2)} & \dots & 0 & D_p^{(\ell_2)} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & C^{(\ell_N)} & D_p^{(\ell_N)} \end{bmatrix} \begin{bmatrix} x^{(\ell_1)}(k) \\ x^{(\ell_2)}(k) \\ \vdots \\ x^{(\ell_N)}(k) \\ p(k) \end{bmatrix} \tag{30}$$

Next, define

$$\mathcal{A}^{(\mathcal{L}_N)} = \begin{bmatrix} (A^{(\ell_1)} - I) & 0 & \cdots & 0 & B_p^{(\ell_1)} \\ C^{(\ell_1)} & 0 & \cdots & 0 & D_p^{(\ell_1)} \\ 0 & (A^{(\ell_2)} - I) & \cdots & 0 & B_p^{(\ell_2)} \\ 0 & C^{(\ell_2)} & \cdots & 0 & D_p^{(\ell_2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & (A^{(\ell_N)} - I) & B_p^{(\ell_N)} \\ 0 & 0 & \cdots & C^{(\ell_N)} & D_p^{(\ell_N)} \end{bmatrix}. \tag{31}$$

Then the next theorem generalizes Theorem 2.1 to multiple equilibrium points.

Theorem 2.2 *The system (29)-(30) is mean value steady-state observable only if*

$$\mathrm{nrow}(\mathcal{A}^{(\mathcal{L}_N)}) \geq \mathrm{ncol}(\mathcal{A}^{(\mathcal{L}_N)}). \tag{32}$$

Furthermore, (29)-(30) is mean value steady-state observable if and only if

$$\mathrm{rank}(\mathcal{A}^{(\mathcal{L}_N)}) = \mathrm{ncol}(\mathcal{A}^{(\mathcal{L}_N)}). \tag{33}$$

Proof. The system (29)-(30) yields the following steady state equations:

$$\begin{aligned} x_{ss}^{(\ell_1)} &= A^{(\ell_1)} x_{ss}^{(\ell_1)} + B_p^{(\ell_1)} \Delta P, \\ z_{ss}^{(\ell_1)} &= C^{(\ell_1)} x_{ss}^{(\ell_1)} + D_p^{(\ell_1)} \Delta P, \\ &\vdots \\ x_{ss}^{(\ell_N)} &= A^{(\ell_N)} x_{ss}^{(\ell_N)} + B_p^{(\ell_N)} \Delta P \\ z_{ss}^{(\ell_N)} &= C^{(\ell_N)} x_{ss}^{(\ell_N)} + D_p^{(\ell_N)} \Delta P. \end{aligned}$$

Concatenating the above equations yields the linear system

$$\mathcal{A}^{(\mathcal{L}_N)} y^{(\mathcal{L}_N)} = b^{(\mathcal{L}_N)}. \tag{34}$$

where

$$y^{(\mathcal{L}_N)} = \begin{bmatrix} x_{ss}^{(\ell_1)} \\ \vdots \\ x_{ss}^{(\ell_N)} \\ \Delta P \end{bmatrix}, \; b^{(\mathcal{L})} = \begin{bmatrix} 0 \\ z_{ss}^{(\ell_1)} \\ \vdots \\ 0 \\ z_{ss}^{(\ell_N)} \end{bmatrix}. \tag{35}$$

Since (34) is identical in form to (23) of Theorem 2.1, the remainder of the proof is identical to the proof of Theorem 2.1. □

Remark 2.3 *The results of Theorem 2.2 use analysis of the linearized models to provide a condition under which the parameter variation vector p can be estimated by using measurements taken from one or more equilibrium points. The condition essentially ensures that $\mathcal{A}^{(\mathcal{L}_N)}$ is invertible. From standard results in linear algebra, it is known that the invertibility of $\mathcal{A}^{(\mathcal{L}_N)}$ is related by its inverse condition number $\mu(\mathcal{A}^{(\mathcal{L}_N)})$, defined as*

$$\mu(\mathcal{A}^{(\mathcal{L}_N)}) \triangleq \frac{\sigma_{min}(\mathcal{A}^{(\mathcal{L}_N)})}{\sigma_{max}(\mathcal{A}^{(\mathcal{L}_N)})} \leq 1. \tag{36}$$

The value of $\kappa(\mathcal{A}^{(\mathcal{L}_N)})$ is related to the degree of invertibility of $\mathcal{A}^{(\mathcal{L}_N)}$. In particular, if $\mu(\mathcal{A}^{(\mathcal{L}_N)}) = 0$, then $\mathcal{A}^{(\mathcal{L}_N)}$ is not invertible. Furthermore, in practice $\mu(\mathcal{A}^{(\mathcal{L}_N)}) = 0$ may provide a measure of the "estimability" of the parameter variation vector ΔP using data from the corresponding equilibrium points. That is, the smaller the value of $\mu(\mathcal{A}^{(\mathcal{L}_N)}) = 0$, the less accurate may be the estimate of ΔP obtained using the corresponding data. That this is the case, is demonstrated in the numerical example of Section 4.

3 Stacked Nonlinear Model Structure for Parameter Variation Estimation

The analysis of the previous section is based upon the linearized models of the system about the stable equilibrium points. In practice, due to the plant noise and parameter variations, even when operating around an equilibrium point, the system will exhibit some nonlinear behavior. Hence, it is important to use a nonlinear filter when performing the estimation. In the example of the next section, the EKF [5] is used. However, the UKF [6, 19] may also be used and may yield better results. However, it is important to recognize that, as discussed in the previous section, a key issue for the success of the filter in estimating ΔP, no matter which filtering method is used, is the observability of ΔP in the data.

The stacked nonlinear model needed for nonlinear estimation is derived directly from the nonlinear system description (1)-(2). The control input $\mathcal{U}_c^{(\ell)}$ at each operating point ℓ is assumed to be constant although the presence of noise will always cause it to vary from time to time; and the state vector at each equilibrium point is denoted by the state vector $\mathcal{X}_{eq}^{(\ell)}(t)$. The parameter vector and the state vector for N different operating points are concatenated to form the augmented state vector,

$$x_{\text{aug}}(t) = [\mathcal{X}_{eq}^{(1)}(t)^T, \mathcal{X}_{eq}^{(2)}(t)^T, \cdots, \mathcal{X}_{eq}^{(N)}(t)^T, P^T]^T. \tag{37}$$

The augmented control input $u_{\text{aug}}(t)$ that corresponds to this augmented state vector is defined as

$$u_{\text{aug}}(t) = [\mathcal{U}_c^{(1)}(t)^T, \mathcal{U}_c^{(2)}(t)^T, \cdots, \mathcal{U}_c^{(N)}(t)^T]^T. \tag{38}$$

Suppose that $w^{(\ell)}(t)$ and $v^{(\ell)}(t)$ are the plant and measurement noise processes at each operating point ℓ, then the measurement and noise processes corresponding to the augmented vector $x_{\text{aug}}(t)$ are defined as

$$w_{\text{aug}}(t) = \text{block-diag}(w^{(1)}(t), w^{(2)}(t), \cdots, w^{(N)}(t), 0_{n_p}), \tag{39}$$

$$v_{\text{aug}}(t) = \text{block-diag}(v^{(1)}(t), v^{(2)}(t), \cdots, v^{(N)}(t)). \tag{40}$$

Note that if each of augmented plant and measurement noise processes $w^{(\ell)}(t)$ and $v^{(\ell)}(t)$ are Gaussian with zero mean, then the augmented noise processes $w_{\text{aug}}(t)$ and $w_{\text{aug}}(t)$ are also Gaussian with zero mean.

The dynamics of the parameter variation is still assumed to be the same as in (3). Therefore, the dynamics of the augmented system becomes

$$\dot{x}_{\text{aug}}(t) = F(x_{\text{aug}}(t), u_{\text{aug}}(t)) + w_{\text{aug}}(t), \tag{41}$$

$$z(t) = H(x_{\text{aug}}(t), u_{\text{aug}}(t)) + v_{\text{aug}}(t), \tag{42}$$

where

$$F(x_{\text{aug}}(t), u_{\text{aug}}(t)) = \begin{bmatrix} f(\mathcal{X}_{eq}^{(1)}(t), \mathcal{U}_c^{(1)}, P) \\ f(\mathcal{X}_{eq}^{(2)}(t), \mathcal{U}_c^{(2)}, P) \\ \vdots \\ f(\mathcal{X}_{eq}^{(N)}(t), \mathcal{U}_c^{(N)}, P) \\ \gamma(t) \end{bmatrix}, \tag{43}$$

$$H(x_{\text{aug}}(t), u_{\text{aug}}(t)) = \begin{bmatrix} h(\mathcal{X}_{eq}^{(1)}(t), \mathcal{U}_c^{(1)}, P) \\ h(\mathcal{X}_{eq}^{(2)}(t), \mathcal{U}_c^{(2)}, P) \\ \vdots \\ h(\mathcal{X}_{eq}^{(N)}(t), \mathcal{U}_c^{(N)}, P) \end{bmatrix}, \tag{44}$$

$$z(t) = [\mathcal{Z}^{(1)}(t)^T, \mathcal{Z}^{(2)}(t)^T, \cdots, \mathcal{Z}^{(N)}(t)^T]^T, \tag{45}$$

where $\mathcal{Z}^{(i)}(t)$ corresponds to the measurement at the i^{th} equilibrium point.

Discretization of this process by using the Euler-scheme yields

$$\begin{aligned} x_{\text{aug}}(k+1) &= x_{\text{aug}}(k) + T_s\left[F(x_{\text{aug}}(k), u_{\text{aug}}(k))\right] + \sqrt{T_s}\, w_{\text{aug}}(k), \\ &= \mathrm{F}_d(x_{\text{aug}}(k), u_{\text{aug}}(k)) + \overline{w}_{\text{aug}}(k), \end{aligned} \tag{46}$$

$$z(k) = \mathrm{H}_d(x_{\text{aug}}(k), u_{\text{aug}}(k)) + v_{\text{aug}}(k). \tag{47}$$

The filter for estimating the parameter variation vector ΔP is developed based on this augmented nonlinear system along with noise models $\overline{w}_{\text{aug}}$ and v_{aug}. This filter estimates the augmented state vector x_{aug}, which includes the states at the chosen equilibrium points $\mathcal{X}_{eq}^{(1)}(t)$, $\mathcal{X}_{eq}^{(2)}(t), \cdots, \mathcal{X}_{eq}^{(N)}(t)$ along with the parameter vector P. Therefore, if P_{est} denotes the estimate of P obtained from the EKF, the UKF or some other nonlinear filter, the estimate of ΔP is given by

$$\Delta P_{est} = P_{est} - P_0. \tag{48}$$

4 Parameter Variation Estimation for the CSTR Using Stacked Model Structures with the EKF

This section presents applications of the conditions developed for mean valued steady-state observability to a Continuously Stirred Tank Reactor (CSTR) along with parameter variation estimation results. First, a description the CSTR is given along with a discussion of the analysis that was performed in order to find the equilibrium points. The necessary conditions (19) and (32) of Theorems 2.1 and 2.2, respectively, along with the inverse condition number 36 are used to evaluate the suitability for parameter variation estimation of various combinations of equilibrium points. This analysis predicts that a combination of at least four equilibrium points is needed for parameter variation estimation. The actual estimation is performed using the EKF and the results confirm the validity of the analysis based on the linearized models.

4.1 Description of the CSTR System

The model of a standard 2-state continuous stirred tank reactor (CSTR) [14, 15], describing an exothermic diabatic irreversible first-order reaction (A → B), is a set of two nonlinear ordinary differential equations obtained from dynamic material and energy balances (with assumptions of constant volume, perfect mixing, negligible cooling jacket dynamics, and constant physical parameters):

$$\tfrac{dC_a}{dt} = \tfrac{Q}{V}(C_{af} - C_a) - k_o exp(\tfrac{-E_a}{RT})C_a, \tag{49}$$

$$\begin{aligned}\tfrac{dT}{dt} &= \tfrac{Q}{V}(T_f - T) - \tfrac{UA}{\rho(VC_p)}(T - T_c) \\ &\quad + \tfrac{-\Delta(H)}{C_p} k_o exp(\tfrac{-E_a}{RT})C_a,\end{aligned} \tag{50}$$

where C_a and T are the concentration of component A and the temperature in the reactor, respectively. An additional energy balance around the cooling jacket, assuming perfect mixing, yields:

$$\frac{dT_c}{dt} = \frac{Q_c}{V_c} + \frac{UA}{V_c \rho_c C_p c}(T - T_c), \tag{51}$$

where T_c is the cooling jacket temperature.

Adding plant noise, the nonlinear system (49)-(51) can be expressed in the dimensionless form,

$$\frac{dx_1}{d\tau} = q(x_{1f} - x_1) - \phi x_1 \kappa(x_2) + w_1, \tag{52}$$

$$\frac{dx_2}{d\tau} = q(x_{2f} - x_2) - \delta(x_2 - x_3) + \beta\phi x_1 \kappa(x_2) + w_2, \tag{53}$$

$$\frac{dx_3}{d\tau} = \delta_1[q_c(x_{3f} - x_3) + \delta\delta_2(x_2 - x_3)] + w_3, \tag{54}$$

where x_1, x_2, x_3, and q_c are, respectively, the dimensionless concentration of component A, reactor temperature, cooling jacket temperature, and cooling jacket flowrate, and $w = [w_1, w_2, w_3]^T$ denotes the plant disturbance. The dimensionless cooling jacket flowrate, q_c, is the manipulated input to the system, while the dimensionless reactor temperature, x_2, and dimensionless cooling jacket temperature, x_3, are the measured outputs. Definitions of the dimensionless parameters (ϕ, β, etc...) are given as

$$\left.\begin{array}{ll}
x_1 = \frac{C_a}{C_{af}}, & x_2 = \frac{(T - T_{fo})\gamma}{T_{fo}}, \\
x_3 = \frac{(T_c - T_{fo})\gamma}{T_{fo}}, & q_c = \frac{Q_c}{Q_o}, \\
\phi = \frac{V k_o e^{-\gamma}}{Q_o}, & \beta = \frac{(-\Delta H) C_{afo}\gamma}{C_p T_{fo} \rho}, \\
\delta = \frac{UA}{C_p Q_o \rho} & \gamma = \frac{E_a}{R T_{fo}}, \\
\kappa(x_2) = exp(\frac{x_2}{1 + \frac{x_2}{\gamma}}), & q = \frac{Q}{Q_o}, \\
\tau = \frac{Q_o t}{V}, & \delta_1 = \frac{V}{V_c}, \\
\delta_2 = \frac{\rho(C_p)}{\rho_c C_{pc}}, & x_{1f} = \frac{C_{af}}{C_{afo}}, \\
x_{2f} = \frac{(T_f - T_{fo})\gamma}{T_{fo}}, & x_{3f} = \frac{(T_{cf} - T_{fo})\gamma}{T_{fo}}.
\end{array}\right\} \tag{55}$$

The nominal values of these dimensionless parameters are follows: $\phi = 0.11$, $\beta = 7.0$, $\delta = 0.5$, $\gamma = 20$, $q = 1.0$, $\delta_1 = 10$, $\delta_2 = 1.0$, $x_{1f} = 1.0$, $x_{2f} = 0.0$, and $x_{3f} = -1.0$ It is assumed that each state vector can be measured such that the measurement vector z is given by

$$z = x + v \tag{56}$$

where $x = [x_1,\ x_2,\ x_3]^T$ is the state vector and v denotes the measurement noise.

The system's equilibrium points were determined by using the software package MATLAB/SIMULINK when the system parameters are assumed to be at their nominal values. The values of the manipulated input q_c were chosen from a finite region of the system's entire range of operation. In particular, q_c was chosen to be in $\{0.0, 0.25, 0.50, \ldots, 3.0\}$, a set of 13 elements. Higher values of q_c did not yield significant changes in the system equilibria.

Degradation in the reactor parts can cause variations in the physical parameters k_o, E_a, $\Delta(H)$, and U. The first set of experiments assumed that these parameters varied $\pm 15\%$ from their nominal values, while the second set of experiments assumed that the parameters varied $\pm 25\%$ from their nominal values. The above variations, as can be seen in (55), lead to variations in the dimensionless parameters ϕ, β, δ, γ, x_{2f}, and x_{3f}. The remaining parameters i.e., q, δ_1, δ_2, and x_{1f}, are considered independent of the reactor components and remain constant with time. Thus,

$$\Delta P = \begin{bmatrix} \Delta\phi \\ \Delta\beta \\ \Delta\delta \\ \Delta\gamma \\ \Delta x_{2f} \\ \Delta x_{3f} \end{bmatrix}. \tag{57}$$

4.2 The Suitability of Combinations of Equilibrium Points for Parameter Variation Estimation

The system was linearized about the stable equilibrium point for each value of q_c and discretized using a sample period, $T_s = 2.0$ sec. All equilibrium points and their combinations were examined to see if they satisfy (32) and (33), the conditions of Theorem 2.2. Table 1 relates the number of columns and rows in $\mathcal{A}^{(\mathcal{L}_N)}$ and the number of equilibrium points for estimating ΔP. It can be seen that the information from at least three equilibrium points is necessary for the satisfaction of (32). When three or more equilibrium points were used, the invertibility of $\mathcal{A}^{(\mathcal{L}_N)}$ was measured by the inverse of its condition number, i.e., $\mu(\mathcal{A}^{(\mathcal{L}_N)})$ given by (36). In order to satisfy (33) well, it is desired for $\mu(\mathcal{A}^{(\mathcal{L}_N)}) \in [0, 1]$ to be as large as possible. Thus, the combinations of equilibrium points that will be referred to as the "best" combinations are considered so in the sense that they have the largest $\mu(\mathcal{A}^{(\mathcal{L}_N)})$ of all respective combinations. Experimental estimation results indicate that the smallest value of $\mu(\mathcal{A}^{(\mathcal{L}_N)})$ for which parameter estimation for CSTR using stacked models can yield satisfactory results is approximately 10^{-7}.

Figures 1 through 3 show the variation of the inverse condition number (on a logarithmic scale) for different combinations of three, four and five equilibrium points, each of which satisfies the column-row condition of (32). Since combinations of two or less do not satisfy the column-row condition of (32) as indicated in Table 1, the condition numbers corresponding to these combinations do not have to be computed. Figure 1 shows that all 286 combinations of three equilibrium points have a very bad condition number. The "best" combination of three equilibrium points has $\mu(\mathcal{A}^{(\mathcal{L}_N)}) = 1.0511 \times 10^{-17} \approx 0$. Hence, (33) is not necessarily satisfied by any combination of three equilibrium points although (32) is satisfied. Figures 2 and 3 show that there are some combinations of four and five equilibrium points that have acceptable condition numbers, i.e., $\mu(\mathcal{A}^{(\mathcal{L}_N)}) \geq 10^{-7}$. The "best" combination of the 715 possible combinations of four equilibrium points has $\mu(\mathcal{A}^{(\mathcal{L}_N)}) = 1.6642 \times 10^{-4}$, and the "best" combination of the 1287 possible combinations of five equilibrium points has $\mu(\mathcal{A}^{(\mathcal{L}_N)}) = 1.6868 \times 10^{-4}$. This indicates that good estimation results should

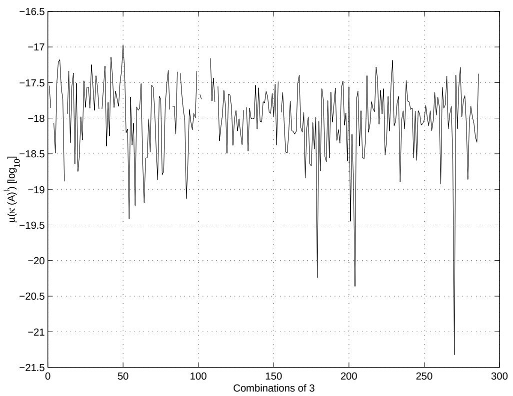

Figure 1: Variation of $\mu(\mathcal{A}^{(\mathcal{L}_N)})$ for Different Combinations of Three Equilibrium Points

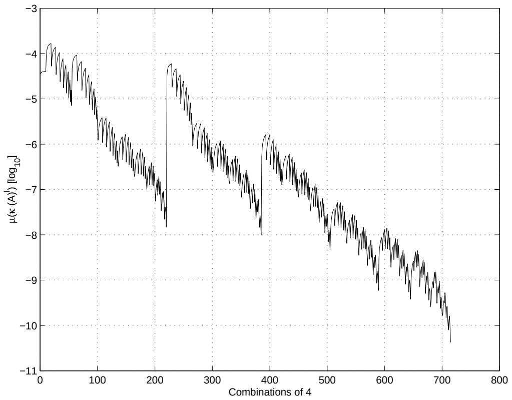

Figure 2: Variation of $\mu(\mathcal{A}^{(\mathcal{L}_N)})$ for Different Combinations of Four Equilibrium Points

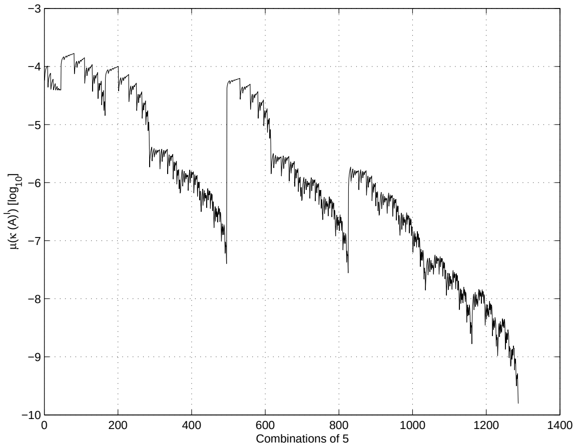

Figure 3: Variation of $\mu(\mathcal{A}^{(\mathcal{L}_N)})$ for $\mu(\mathcal{A}^{(\mathcal{L}_N)})$ for Different Combinations of Five Equilibrium Points

be expected from some combinations of four and five equilibrium points. The following parameter variation estimation results confirm this indication. Although the corresponding results for the “best” combination of five equilibrium points are not given here, they are very similar to those for the “best” combination of four equilibrium points.

In both sets of experiments described below, the EKF was used for parameter estimation. It was initialized with zero values and identity covariance. The zero mean, white, plant disturbance $w(k)$ had covariance $W = \mathrm{diag}(0.05, 0.05, 0.05)$ while the zero mean, white, sensor noise $v(k)$ had covariance $V = \mathrm{diag}(0.05, 0.05, 0.05)$.

4.3 First Set of Experiments

In the first set of experiments the parameter variations were limited to $\pm 15\%$ of their nominal values. In the below discussions, a parameter variations p_i is considered to be accurately estimated if the percentage relative error $\varepsilon(p_i)$ is less than 5%, where

$$\varepsilon(p_i) = \begin{cases} \left|\frac{p_i - p_{i(est)}}{p_i}\right| \times 100 & \text{if } p_i \neq 0, \\ \\ \left|\frac{p_i - p_{i(est)}}{p_i(est)}\right| \times 100 & \text{if } p_i = 0. \end{cases} \tag{58}$$

CASE 1: A Single Equilibrium Point

The first case of estimation involved the use of data from the stable equilibrium point of the CSTR, corresponding to q_c=3.00. Since for all single equilibrium points, the size of the $\mathcal{A}^{(\mathcal{L}_N)}$ matrix is 5×9 and does not satisfy either (32) or (33), the results were not expected to be satisfactory. However, this particular equilibrium point was chosen because it corresponded to the maximum value of the pseudo inverse condition number,

$$\mu_{ps}(\mathcal{A}^{(\mathcal{L}_N)}) \triangleq \frac{\sigma_5(\mathcal{A}^{(\mathcal{L}_N)})}{\sigma_{max}(\mathcal{A}^{(\mathcal{L}_N)})}. \tag{59}$$

(Note that the actual inverse condition number (36) is zero for each single equilibrium point.) Table 2 shows that none of the parameter variations were accurately estimated.

CASE 2: A Combination of 2 Equilibrium Points

For all combinations of 2 equilibrium points the size of the $\mathcal{A}^{(\mathcal{L}_N)}$ matrix is 10×12 and does not satisfy either (32) or (33) an hence the inverse condition number (36) is zero for each combination of two. The combination of two equilibrium points used corresponded to q_c=0.25, 3.0 and was chosen to maximize the pseudo inverse condition number,

$$\mu_{ps}(\mathcal{A}^{(\mathcal{L}_N)}) \triangleq \frac{\sigma_{10}(\mathcal{A}^{(\mathcal{L}_N)})}{\sigma_{max}(\mathcal{A}^{(\mathcal{L}_N)})}. \tag{60}$$

Table 2 shows that the information provided by this combination of equilibrium points was sufficient to accurately estimate $\Delta\beta$, $\Delta\delta$, and Δx_{3f}. However, the remaining parameter variations are not accurately estimated. As with Case 1, good estimation results were not expected.

CASE 3: The "Best" Combination of 3 Equilibrium Points

For the case that involve the use of the information from three equilibrium points, the "best" combination corresponds to q_c=0.0, 1.5, 2.25. For each combination of 3 equilibrium points the size of the $\mathcal{A}^{(\mathcal{L}_N)}$ matrix is 15×15 and satisfies (32). For the "best" combination considered here, the $\mu(\mathcal{A}^{(\mathcal{L}_N)})$=$1.0511 \times 10^{-17}$ $\approx$ 0, indicating that (33) is not effectively satisfied. Hence, it was not expected that each of the parameter variations would be accurately estimated. Table 5 shows that, like Case 2, $\Delta\beta$, $\Delta\delta$, and Δx_{3f} were estimated accurately but the remaining parameter variations were not accurately estimated.

CASE 4: The "Best" Combination of 4 Equilibrium Points

The fourth case of estimation involved the use of the information from the "best" combination of 4 equilibrium points of the CSTR, corresponding to q_c=0.0, 0.25, 0.75, 3.0. For each combination of 4 equilibrium points the size of the $\mathcal{A}^{(\mathcal{L}_N)}$ matrix is 20×18 and satisfies (32). For the "best" combination $\mu(\mathcal{A}^{(\mathcal{L}_N)})$=$1.1915 \times 10^{-4}$, which indicates that (33) is satisfied and it may be possible to accurately estimate each parameter variation. Table 5 shows that the information provided in this case is significantly richer than all of the previous cases and each of the parameter variations is accurately estimated. The next case is considered in order to emphasize the importance of choosing an "appropriate" combination of 4 equilibrium points to use for estimation.

CASE 5: An "Arbitrary" Combination of 4 Equilibrium Points

The final case of estimation involved the use of information from the "arbitrarily" chosen combination of 4 equilibrium points, corresponding to q_c=1.25, 1.5, 1.75, 2.0. For this case, $\mu(\mathcal{A}^{(\mathcal{L}_N)})$=$1.0087 \times 10^{-10}$, which is about 6 orders of magnitude smaller than the "best" combination of 4. Hence, although (33) is satisfied, the conditioning of $\mathcal{A}^{(\mathcal{L}_N)}$ is poor and it is unclear that accurate parameter variation estimation can be achieved. Table 2 shows that the information provided by this combination of equilibrium points is only sufficient to accurately estimate $\Delta\delta$ and Δx_{3f}. Each of the remaining parameter variations is not accurately estimated . Thus, there is a significant amount of information about the parameter variations that is not provided by this combination of equilibrium points. Hence, the size of $\mu(\mathcal{A}^{(\mathcal{L}_N)})$ was effective in revealing the possible inadequacy of this combination of equilibrium points.

4.4 Second Set of Experiments

In the first set of experiments the parameter variations were small enough that a Kalman filter was capable of giving essentially the same results as the EKF. That is, the nonlinear dynamics of the CSTR were not substantial since the system did not deviate substantially from the linear region about the various operating points. Hence, in the second set of experiments the parameters were allowed to vary within $\pm 25\%$ of their nominal values. For this case, the Kalman filter failed and the use of the EKF was essential to the success of the estimation due to the more substantial nonlinear effects seen in the simulations. The focus was on combinations of four equilibrium points since it was shown in the analysis and first set of experiments that combinations of fewer numbers of equilibrium points would not be adequate for paramter variation estimation.

Table 4 gives a summary of the parameter estimation results using 6 combinations of four equilibrium points. As seen in Table 4, Cases 1, 2, 3, 4 and 5 yielded accurate estimation of the parameter variations whereas Case 6 did not. Case 1 corresponds to the "best" combination that was obtained using the analysis based on the linearized models, and Cases 2 through 5 have inverse condition numbers $\mu(\mathcal{A}^{(\mathcal{L}_N)})$ that are greater than 10^{-7}. Case 6 corresponds to an "arbitrary" combination of four points, which has an inverse condition number less than 10^{-7}. Table 3 shows the condition numbers for these cases. It is also noted that in all combinations of 3 or 4 equilibrium points, the parameter changes $\Delta\delta$ and Δx_{3f} were estimated fairly accurately irrespective of the particular choice of the equilibrium points. This in agreement with the results obtained from the first set of experiments and indicates that some of the parameters can be strongly observable from all equilibrium points while others are not.

5 Conclusion

This paper has studied the limitations of estimating the state and parameter variations of a nonlinear system using data from equilibrium points of the system. It was seen that if the number of sensors is less than the number of parameter variations to be estimated, then it is always necessary to use data from more than one equilibrium point to perform the desired estimation. Criteria for calculating the minimum number of equilibrium points needed and for determining an appropriate combination of equilibrium points were developed in terms of the dimensions and rank of a matrix $\mathcal{A}^{(\mathcal{L}_N)}$ that is a function of the linearized system matrices at each of the equilibrium points. In addition, a "stacked" model structure was proposed as the basis of the estimation. Practical implementation of the estimation would require storing data from several equilibrium points.

The analysis results based on the linearized models were applied to parameter variation estimation for a CSTR. It was shown that data from any combination of 3 or less equilibrium points is inadequate to estimate the state and parameter variations. However, data from an appropriately chosen combination of 4 equilibrium points is adequate to accurately estimate the state and parameter variations. An EKF was used in combination with the stacked nonlinear models to perform the parameter variation estimation. The results verified the theory.

This research was limited to nonlinear systems that operate about various equilibrium points and considered the data necessary for parameter observability using linear analysis. However, the fundamental issue of parameter observability is present for nonlinear systems that do not necessarily operate at equilibrium points. Hence, future research should consider developing parameter observability conditions based on nonlinear analysis. This is an important, but difficult problem.

Acknowledgements

Authors acknowledge the funding by the National Aeronautics and Space Administration under Grant NAG3-2193 and the National Science Foundation under Grant CMS-9802197.

References

[1] A. Ben-Israel and A. Charnes. Contributions to the Theory of Generalized Inverses. *SIAM J. Appl. Math*, 17(2):434–440, March 1963.

[2] E. G. Collins, Jr., M. F. Selekwa, R. B. Walker, and S. Palanki. Parameter and State Variation Estimation for Multi-Equilibria Non-Linear Systems Using a “Stacked” Model Structure. In *Proc. Amer. Cont. Conf.*, pages 2530–2535, Arlington–Virginia, USA, June 2001.

[3] E. G. Collins, Jr., R. B. Walker, and S. Palanki. Parameter Variation Estimation for Nonlinear Systems Using a “Stacked” Linear Model Structure. In *Proc. IEEE 39rd Conf. on Dec. and Cont.*, pages 3265–3270, Phoenix–Arizona, USA, Dec 1999.

[4] J. Garcia-Velo and B. K. Walker. Aerodynamic and parameter estimation for high performance aircraft using extended kalman filtering. *Journal of Guidance, Control and Dynamics*, 20(6):1257–1259, 1997.

[5] M. S. Grewal and A. P. Andrews. *Kalman Filtering: Theory and Practice*. Prentice Hall, New Jersey, 1993.

[6] S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, March 2004.

[7] R. E. Kalman and R. Bucy. New Results in Linear Filtering and Prediction Theory. *ASME Journal of Basic Engineering*, 83:95–108, 1961.

[8] P. E. Kloeden and E. Platen. *Numerical Solutions of Stochastic Differential Equations*, volume 23 of *Applications of Mathematics*. Springer, New York, 1992.

[9] Yassine Koubaa. Recursive identification of induction motor parameters. *Simulation Modelling Practice and Theory*, 12(5):363–381, 2004.

[10] L. Ljung. *System Identification: Theory for the User*. Prentice Hall, Englewood Cliffs, New Jersey 07632, 1987.

[11] M. Nørgaard, N. K Poulsen, and O. Ravn. New developments in state estimation for nonlinear systems. *Automatica*, 36:1627–1638, 2000.

[12] J. B. Roberts, J. C. Peyton Jones, and K. J. Landsborough. Stochastic modelling and estimation for cyclic pressure variations in spark ignition engines. *Mechanical Systems and Signal Processing*, 15(2):419–438, 2001.

[13] D. Robertson, J. Lee, and J. Rawlings. A Moving-Horizon Based Approach for Least Squares State Estimation. *AIChE Journal*, 42:2337, 1996.

[14] L. P. Russo and B. W. Bequette. Impact of Process Design on Multiplicity Behavior of a Jacketed Exothermic CSTR. *AIChE Journal*, 41:135–147, January 1995.

[15] L. P. Russo and B. W. Bequette. Effect of Process Design on the Open-Loop Behavior of a Jacketed Exothermic CSTR. *Computers in Chemical Engineering*, 20(4):417–426, 1996.

[16] A. Sirohi and K. Choi. On-Line Parameter Estimation in a Continuous Polymerization Process. *Ind. Eng. Chem. Res.*,, 35:1332–1343, 1996.

[17] M. Sortino. Application of statistical filtering for optical detection of tool wear. *International Journal of Machine Tools & Manufacture*, 43:493–497, 2003.

[18] S. Tatiraju and M. Soroush. Parameter Estimator Design With Applications to a Chemical Reactor. *Ind. Engng. Chem. Res.*, 37(2):455–463, 1998.

[19] R. van der Merwe and E. A. Wan. The square-root unscented kalman filter for state and parameter estimation. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3461–3464, Salt Lake City, Utah, May 2001.

N	nrow($\mathcal{A}^{(\mathcal{L}_N)}$)	ncol($\mathcal{A}^{(\mathcal{L}_N)}$)
1	5	9
2	10	12
3	15	15
4	20	18
5	25	21

Table 1: Number of Rows and Columns of $\mathcal{A}^{(\mathcal{L}_N)}$ for the CSTR

	$\Delta\phi$ =0.000	$\Delta\beta$ =2.258	$\Delta\delta$ =0.075	$\Delta\gamma$ =3.000	Δx_{2f} =0.000	Δx_{3f} =-0.150
Case 1	-0.0114	0.0270	-0.1220	-0.09040	0.1634	-0.0656
Case 2	0.0057	2.2338	0.0764	3.6311	-0.0362	-0.1499
Case 3	-0.0116	2.2769	0.0815	0.4458	0.1032	-0.1531
Case 4	-0.0005	2.2764	0.0748	2.8288	-0.0025	-0.1504
Case 5	0.1353	-3.4511	0.0743	-7.445	-0.2341	-0.1475

Table 2: The Performance of the Kalman Filter for the Stacked Linear System Using Different Combinations of Equilibrium Points

Case	Equilibrium Points	$\kappa(\mathcal{A}^{(\mathcal{L}_4)})$
1	$q_c = \left\{ \begin{array}{cc} 3.00, & 0.00, \\ 0.25, & 0.75 \end{array} \right\}$	1.6642×10^{-4}
2	$q_c = \left\{ \begin{array}{cc} 3.00, & 0.00, \\ 0.25, & 1.00 \end{array} \right\}$	1.379159×10^{-4}
3	$q_c = \left\{ \begin{array}{cc} 0.00, & 0.25, \\ 1.00, & 1.75 \end{array} \right\}$	$1.011608e \times 10^{-4}$
4	$q_c = \left\{ \begin{array}{cc} 1.75, & 2.25, \\ 3.00, & 0.00 \end{array} \right\}$	1.943065×10^{-7}
5	$q_c = \left\{ \begin{array}{cc} 2.25, & 3.00, \\ 0.00, & 0.25 \end{array} \right\}$	2.610218×10^{-5}
6	$q_c = \left\{ \begin{array}{cc} 1.00, & 1.75, \\ 2.25, & 3.00 \end{array} \right\}$	4.573870×10^{-9}

Table 3: Condition Numbers for 6 Combinations of Four Equilibrium Points.

		$\Delta\phi$	$\Delta\beta$	$\Delta\delta$	$\Delta\gamma$	Δx_{2f}	Δx_{3f}
Case		=0.027500	=1.750000	=0.125000	=5.000000	=0.000000	=-0.250000
1	$q_c = \left\{ \begin{array}{cc} 3.00, & 0.00, \\ 0.25, & 0.75 \end{array} \right\}$	0.027496	1.752109	0.125102	4.989462	0.000394	-0.249967
2	$q_c = \left\{ \begin{array}{cc} 3.00, & 0.00, \\ 0.25, & 1.00 \end{array} \right\}$	0.027440	1.747488	0.125011	4.945375	0.002820	-0.249785
3	$q_c = \left\{ \begin{array}{cc} 0.00, & 0.25, \\ 1.00, & 1.75 \end{array} \right\}$	0.027573	1.740892	0.124981	4.965056	0.004682	-0.249842
4	$q_c = \left\{ \begin{array}{cc} 1.75, & 2.25, \\ 3.00, & 0.00 \end{array} \right\}$	0.027247	1.753882	0.124981	4.987719	0.001583	-0.249672
5	$q_c = \left\{ \begin{array}{cc} 2.25, & 3.00, \\ 0.00, & 0.25 \end{array} \right\}$	0.027779	1.755808	0.124999	4.937086	-0.001073	-0.249728
6	$q_c = \left\{ \begin{array}{cc} 1.00, & 1.75, \\ 2.25, & 3.00 \end{array} \right\}$	0.035271	1.605195	0.124989	0.212175	0.214146	-0.249978

Table 4: The Performance of the Extended Kalman Filter for the Stacked Nonlinear System with Four Equilibrium points