A New Framework for Data Reconciliation and Measurement Bias Identification in Generalized Linear Dynamic Systems

Hua Xu and Gang Rong
Dept. of Control Science and Engineering, Institute of Advanced Process Control,
State Key Laboratory of Industrial Control Technology,
Zhejiang University, Hangzhou 310027, People's Republic of China

DOI 10.1002/aic.12102
Published online November 5, 2009 in Wiley InterScience (www.interscience.wiley.com).

This article describes a new framework for data reconciliation in generalized linear dynamic systems, in which the well-known Kalman filter (KF) is inadequate for filtering. In contrast to the classical formulation, the proposed framework is in a more concise form but still remains the same filtering accuracy. This comes from the properties of linear dynamic systems and the features of the linear equality constrained least squares solution. Meanwhile, the statistical properties of the framework offer new potentials for dynamic measurement bias detection and identification techniques. On the basis of this new framework, a filtering formula is rederived directly and the generalized likelihood ratio method is modified for generalized linear dynamic systems. Simulation studies of a material network present the effects of both the techniques and emphatically demonstrate the characteristics of the identification approach. Moreover, the new framework provides some insights about the connections between linear dynamic data reconciliation, linear steady state data reconciliation, and KF. © 2009 American Institute of Chemical Engineers *AIChE J,* 56: 1787–1800, 2010
Keywords: dynamic data reconciliation, linear equality constrained least squares, kalman filter, bias detection, bias identification

Introduction

Dynamic data reconciliation (DDR) is a process of estimating variables on the basis of measurements and dynamic material or energy balance constraints. DDR can provide more reliable information about the current state of a process. An implicit assumption in DDR is that measurement errors are independent with time and normally distributed with an expected value zero. However, the assumption is nullified when process data contain measurement biases. A measurement bias is made when a measurement of a variable deviates far from its true value due to malfunction on a device or improper use of a device. It can be described as a step function which has a large unknown magnitude.[1] Measurement biases seriously distort the results of DDR. Thus before DDR is performed, the statistical properties of measurement data should be checked to detect and identify measurement biases. Both DDR and measurement bias identification exploit the same information contained in measurements and balance constraints.

Many studies have been conducted on linear dynamic data reconciliation (LDDR) techniques. Some of the most widely used techniques are rooted on Kalman filter (KF). These techniques are well-developed for a class of linear dynamic systems. The deterministic part of these systems can be represented by a standard state space equation $x_{k+1} = Ax_k + Bu_k$ where x_k is an unknown vector at time instant k and u_k is a vector representing known time varying parameters. As far as we know, this system representation is restricted to two kinds of situations. One is the nominal steady state condition (i.e., variation around nominal point) for the purpose of control. The other is the quasi steady state process. In a quasi steady state

Correspondence concerning this article should be addressed to G. Rong at grong@iipc.zju.edu.cn.

process, variables hold in almost any time except for slow drifts. Thus the matrices in the standard state space equation can be described as $A = I, B = 0$. Stanly and Mah applied discrete KF to estimate flows and temperatures in a heat exchanger network which is in a quasi steady state.[2]

On the basis of the standard state space equation and KF, Willsky and Jones first developed the generalized likelihood ratios (GLR) method for identifying abrupt changes in the state variables and measurements.[3] Narasimhan and Mah made several significant extensions to this method and proposed a general framework for gross error detection (GED).[1] In their framework, once a gross error occurred, its effects on the innovation (i.e., the discrepancy between actual measurements and priori estimates) could be modeled by signature matrices at every sampling time. Then the GLR method was adopted to identify the location of the gross error as well as to estimate its magnitude. This framework was constructed under the assumption of Gaussian white noise. When it came to serially correlated measurement noise, Kao et al. applied autoregressive (AR) moving average time series (ARMA) to model measurement noise.[4] They adopted a prewhitening procedure before using the GLR method. Prakash et al. integrated the GLR method with conventional control and proposed a fault-tolerant control scheme.[5]

Although KF is elaborate and theoretical attractive, it inherently supposes the self-correlated property of state variables. It means that if KF is extended to obtain simultaneous state and parameter estimates (The exact values of the parameters are unknown but can be estimated from measurements and models), the parameter vector u_k is often modeled as a random walk process $u_{k+1} = u_k + w_k$.[6,7] The covariance of w_k should be manually tuned and its inappropriate value will deteriorate the performance of KF. In a chemical process, as the changes in some parameters (e.g., flow variables in a material balance process network) are never exactly known to the operator, it is hard to decide the covariance of w_k especially in the case of transient processes. When a process is in a transient state (i.e., a period of change from one steady state to a new steady state), parameters change so dramatically (e.g., abrupt jumps) that there is not any priori knowledge about the variations. Thus the random walk model is inappropriate to describe the dynamic characteristics of parameters. A more reliable method is to abandon the random walk process assumption and only adopt models with confidence for LDDR. In this situation, the standard state space equations and KF-based methods are not applicable. Also, as a LDDR strategy, KF is unable to include the linear steady state data reconciliation (LSSDR) as the limit for steady state conditions. A more detailed discussion about the self-correlated parameters model and KF can be seen in Almasy's work.[8]

To solve the above general LDDR problem, researchers extended the least squares principals of LSSDR for dynamic systems. The basis of the new solution technique is the concept "generalized linear dynamic model," which was first introduced to the field of data reconciliation by Darouach and Zasadzinski.[9] In their definition, the generalized dynamic model (i.e., a singular model) refers to the model formulation $Dx_{k+1} = Bx_k$, which contains more state variables than constraints and cannot be written in a standard state space form (i.e., matrix D is singular or rectangular). A unique feature of this modeling approach is that KF is not applicable for such systems (Also note that when $D = I$ the singular model returns to the standard state space equation). On the basis of this new model form, they developed a recursive filtering algorithm for dynamic systems. However, one has to work with a huge model and the derivation procedure seems rather cumbersome.[10] To compute effectively, Rollins and Devanathan adopted the same model formulation and proposed a constrained least squares estimator that utilized measurements at two consecutive time instants.[11] They also addressed the statistical properties of the estimates. Ramamurthi et al. presented a successively linearized horizon-based estimation (SLHE) approach for nonlinear dynamic data reconciliation (NDDR).[12] The SLHE approach was directly derived from a least squares criterion constrained by the standard state space equations. The expressions of SLHE were similar to KF when process noise was not taken into account.

So far, little work has been carried out on GED strategies for linear transient systems. Bagajewicz and Jiang presented an integral approach for LDDR and a dynamic integral measurement test (DIMT) for GED.[13] However, it was hard to obtain the standard deviation of estimates analytically in their integral approach. Thus the sampled standard deviation was used in DIMT. Rollins et al. developed a dynamic global test (DGT) to detect measurement biases in generalized linear dynamic systems.[14] Their method also could identify the types of a biased measurement (e.g., flow or level) but without the location of the measurement bias. They suggested that the steady-state GED techniques could be introduced and adapted to a singular system.

In this article, we propose a reduced least squares framework to integrate information from both measurements and models for generalized linear dynamic systems. We use simple mathematical tools, which have been developed in LSSDR, to construct the new framework and to analyze its statistical properties. On the basis of the new framework, we present a new derivation of Darouach's closed-form filtering formulas and extend the GLR method from a state space model to a singular model. The results of this study also help to reveal the links between LSSDR, LDDR, and KF from a systematic view. These relationships are explained in detail in the discussion.

Model

Consider a process network which is composed of n nodes and v streams. The material balance relationship in the process network can be represented by the following differential equations:

$$\frac{dw}{dt} = Mf \tag{1}$$

where f is a $v \times 1$ vector of instantaneous flow rates, w is a $n \times 1$ vector of instantaneous inventory, M is a $n \times v$ incidence matrix of full row rank.

By integrating Eq. 1 from time instant i to $i + 1$ and denoting $\int_i^{i+1} f dt = Q_{i+1}^*$, we obtain the discrete equivalent of the above equation:

$$\int_i^{i+1} \frac{dw}{dt} dt = M \int_i^{i+1} f dt \quad \Rightarrow \quad W_{i+1}^* - W_i^* = MQ_{i+1}^* \tag{2}$$

where Q_{i+1}^* is a $v \times 1$ total mass flow vector over a time period $[i, i + 1]$ and W_i^* is a $n \times 1$ inventory vector at time instant i.

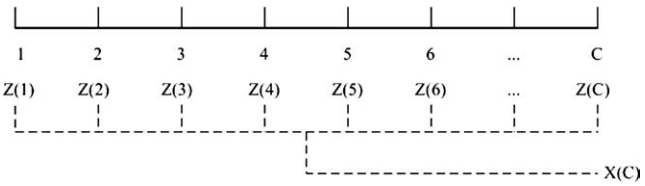

Figure 1. Using the CLSF to obtain the estimates.

As only material conversation relationships are invoked, the model uncertainty is not in question.

Then we cast Eq. 2 into a generalized dynamic model formulation:

$$DX_{i+1} = BX_i$$
$$X_i = \begin{bmatrix} W_i^* \\ Q_i^* \end{bmatrix}, \quad D = \begin{bmatrix} I \vdots -M \end{bmatrix}, \quad B = \begin{bmatrix} I \vdots 0 \end{bmatrix} \tag{3}$$

where X_i is a $(n + v) \times 1$ vector of the true values of unknown variables at time instant i, I is a $n \times n$ identity matrix, D is a $n \times (n + v)$ incidence matrix of full row rank, and B is a $n \times (n + v)$ incidence matrix of full row rank.

Assume that all process variables are measured. We obtain the measurement model:

$$Z_i = X_i + \varepsilon_i$$
$$Z_i = \begin{bmatrix} W_i \\ Q_i \end{bmatrix}, \quad \varepsilon_i = \begin{bmatrix} w_i \\ q_i \end{bmatrix} \tag{4}$$

where Z_i is a $(n + v) \times 1$ vector of measured values of unknown variables at time instant i, W_i is a $n \times 1$ vector of inventory measurements, Q_i is a $v \times 1$ vector of total mass flow measurements, ε_i is a $(n + v) \times 1$ vector of measurement errors at time instant i, w_i is a $n \times 1$ vector of inventory measurement errors at time instant i and q_i is a $v \times 1$ vector of flow measurement errors at time instant i.

Here we assume measurement errors are independent in time and normally distributed with zero mean. Thus the statistical properties of ε_i can be given by

$$\varepsilon_i \sim N(0, V) \tag{5}$$

$$\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0 \quad i \neq j$$
$$V = \begin{bmatrix} V_w & 0 \\ 0 & V_q \end{bmatrix} \tag{6}$$

where V is a $(n + v) \times (n \times v)$ known covariance matrix of measurement errors, V_w is a $n \times n$ known covariance matrix of inventory measurement errors and V_q is a $v \times v$ known covariance matrix of flow measurement errors.

Handling unmeasured process variables

In practice, a situation where all the states are measured is quite rare. Some of them are not measured and must be estimated. In this situation, we decouple the unmeasured variables from measured ones using variable classification. We can obtain a redundant system whose measurement equations can be described as Eq. 4.

The redundant part of the dynamic system can be written in the following equation:

$$B_R \frac{dw_R}{dt} = A_R f_R \tag{7}$$

Its discrete counterpart is as follow:

$$B_R W_{i+1} - A_R Q_{i+1} = B_R W_i \tag{8}$$

We can use Eq. 3 with $D = [B_R - A_R]$, $B = [B_R, 0]$ to model the redundant system.

The unmeasured but observable variables can be calculated using the reconciled values. We refer readers to Bagajewicz's work[13] for detail explanation about variable classification for LDDR.

Classical Least Squares Framework

Suppose that we have collected all measurements in the time horizon $[1,c]$. We aim to estimate the true vector X_c at the current time instant c (i.e., the filtering problem). Traditionally, this problem can be formulated as the following LSSDR problem:

$$\min_{\varepsilon_1, \varepsilon_2, \varepsilon_3 \ldots \varepsilon_c} \quad \frac{1}{2} \sum_{i=1}^{c} \varepsilon_i^T V^{-1} \varepsilon_i$$
$$s.t. \quad \begin{bmatrix} B & -D & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & B & -D & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 & B & -D \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_c \end{bmatrix} = 0$$
$$\begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_c \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_c \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_c \end{bmatrix} \tag{9}$$

The procedure is showed in Figure 1.

From this classical framework, if we want to estimate process variables at the current time instant, we have to use all information from the initial time to the current time. Thus the dimensions of the least-squares optimization problem will increase with the number of sampling periods. It seems not practical in the application.

Simplified Least Squares Framework

Here we propose a simplified least squares formulation to represent the filtering problem in generalized linear dynamic systems.

$$\begin{aligned} \min_{\varepsilon_{i+1}, \zeta_i} \quad & \tfrac{1}{2} \varepsilon_{i+1}^T V^{-1} \varepsilon_{i+1} + \tfrac{1}{2} \zeta_i^T \Sigma_{i/i}^{-1} \zeta_i \\ s.t. \quad & DX_{i+1} = BX_i \\ & Z_{i+1} = X_{i+1} + \varepsilon_{i+1} \\ & \hat{X}_{i/i} = X_i + \zeta_i \end{aligned} \tag{10}$$

where $\hat{X}_{i/i}$ is a $(n + v) \times 1$ vector of estimates at time instant i given measurements up to the time instant i, $\Sigma_{i/i}$ is a $(n + v) \times$

Figure 2. Using the SLSF to obtain the estimates.

$(n + v)$ covariance matrix of the estimates vector $\hat{X}_{i/i}$, ζ_i is a $(n + v) \times 1$ vector of the estimate residuals at time instant i. Note that ζ_i is normally distributed with an expected value of zero and the covariance matrix $\Sigma_{i/i}$.

This procedure is showed in Figure 2.

From this figure, we can see that this new formulation reduces the computational effort considerably because it requires only the information at the current and the preceding time instants. Furthermore, in Appendix A, we will use the sequential processing method[15,16] to prove that the simplified formulation is equivalent to the classical framework in the view of filtering.

Moreover, according to the new framework, we can find that the filter estimates (i.e., $\hat{X}_{i+1/i+1}$) are correlated with the measurements at time instant $i + 1$ (i.e., Z_{i+1}), the estimates at the preceding time instant i (i.e., $\hat{X}_{i/i}$), and the model structure at time instant $i + 1$ (i.e., $DX_{i+1} = BX_i$). By integrating the above information, we construct a residual formula $DZ_{i+1} - BX_{i/i}$. Its statistical properties can reflect abnormal changes in the measurements at time instant $i + 1$. This is the foundation of an approach for measurement bias detection and identification in generalized dynamic systems.

Derivation of recursion formulas for filtering

We use the Lagrange multipliers method to solve the optimization problem Eq. 10. The estimates can be represented in the following equation:

$$\hat{X}_{i+1/i+1} = Z_{i+1} + VD^T\Omega_{i+1}\left(B\hat{X}_{i/i} - DZ_{i+1}\right) \tag{11}$$

where $\Omega_{i+1} = \left(B\Sigma_{i/i}B^T + DVD^T\right)^{-1}$.

Assume that we have already known the covariance matrix $\Sigma_{i/i}$. We apply the propagation rules of the covariance and obtain the following equation:

$$\Sigma_{i+1/i+1} = V - VD^T\Omega_{i+1}DV \tag{12}$$

Equations 11 and 12 constitute the recursive formulas for LDDR. Given that the initial estimate vector $\hat{X}_{1/1}$(e.g., Z_1) and its covariance matrix $\Sigma_{1/1}$ (e.g., V), we can estimate the process variables at every time instant in a recursive way. These formulas agree with the Darouach's algorithm.[9] But our derivation is more intuitive and direct.

Statistical properties of residuals

Here we define a $n \times 1$ residual vector r_i by the following equation:

$$r_i = DZ_i - B\hat{X}_{i-1/i-1} \tag{13}$$

The statistical properties of the residual vector can be described as follows.

Suppose that we want to estimate the expected value of r_i. We can compute it as follows:

$$\begin{aligned} E[r_i] &= E\left[DZ_i - B\hat{X}_{i-1/i-1}\right] \\ &= DE[Z_i] - BE\left[\hat{X}_{i-1/i-1}\right] \\ &= DX_i - BX_{i-1} = 0 \end{aligned} \tag{14}$$

Next we consider the variance matrix of r_i:

$$\begin{aligned} &\mathrm{Var}[r_i] = \mathrm{Cov}(r_i, r_i) \\ &= \mathrm{Cov}\left(DZ_i - B\hat{X}_{i-1/i-1}, DZ_i - B\hat{X}_{i-1/i-1}\right) \\ &= D\underbrace{\mathrm{Cov}(Z_i, Z_i)}_{V}D^T - D\underbrace{\mathrm{Cov}\left(Z_i, \hat{X}_{i-1/i-1}\right)}_{0}B^T \\ &\quad - B\underbrace{\mathrm{Cov}\left(\hat{X}_{i-1/i-1}, Z_i\right)}_{0}D^T + B\underbrace{\mathrm{Cov}\left(\hat{X}_{i-1/i-1}, \hat{X}_{i-1/i-1}\right)}_{\Sigma_{i-1/i-1}}B^T \\ &= DVD^T + B\Sigma_{i-1/i-1}B^T \end{aligned} \tag{15}$$

Another important property is that

$$\mathrm{Cov}\left[r_i, r_j\right] = 0 \quad i \neq j \tag{16}$$

Equation 16 indicates that the residual at time instant i is uncorrelated with the residual at time instant j. The proof of Eq. 16 can be seen in Appendix B.

A GED strategy for singular systems

On the basis of the residual vector r_i and its statistical properties, we extend the GLR method[1,17] for measurement bias identification in generalized linear dynamic systems. Note that r_i is different from the innovation defined in KF, which constitutes the basis for the original GLR method. In our work, the residual vector r_i is closely related to the simplified least squares framework (SLSF) and its filtering formulas.

Consider a case that a bias of an unknown magnitude b occurs in measurement m at time instant k and its magnitude remains constant for all subsequent times. We modify Eq. 4 to model for a measurement bias of this type:

$$Z_i = X_i + \varepsilon_i + bf_m \tag{17}$$

where f_m is a $(n + v) \times 1$ vector in which the mth component is one and other components are all zeros. The unit vector f_m denotes the location of the biased measurement.

When a measurement bias occurs, its effects on the residual r_i can be described as follows:

$$r_i = r_i^1 + r_i^2 \tag{18}$$

$$r_i^2 = bG_{i,m}f_m \tag{19}$$

where r_i^1 is a residual vector obtained when no measurement bias occurs at time instant i, r_i^2 denotes the effects of the bias m

at time instant i. Formulas for computing the signature matrix $G_{i,m}$ can be seen in Appendix C. Note that the signature matrix $G_{i,m}$ is an extension of that defined in Narasimhan's work.[1,17] It can be used for both generalized linear dynamic models and the standard state space models.

From Appendix C, we see that $G_{i,m}$ depends only on the system matrix D, B and r_i^2 is a nonrandom vector. Thus we can infer that Eqs. 15 and 16 remain true no matter a measurement bias occurs or not. The expected value of r_i, namely, Eq. 14, deviates from zeros when a measurement bias occurs.

Now we formulate the hypothesis for measurement bias detection:

$$\begin{aligned} H_0 &: E[r_i] = 0 \\ H_1 &: E[r_i] = r_i^2 = bG_{i,m} f_m \neq 0 \end{aligned} \tag{20}$$

Then we use the global test (GT) statistic to test the above hypothesis (i.e., to find whether a gross error occurs at time instant i):

$$\psi_1 = r_i^T \mathrm{Var}_i^{-1} r_i \tag{21}$$

where Var_i is the covariance matrix of the residual r_i at time instant i and ψ_1 is a test statistic which follows a χ^2 distribution with degrees of freedom to n.

Moreover, in linear dynamic systems, it is reasonable to take into account the historical data in some way. Thus as an alternative way, we can use the cumulative global test (CGT) statistic for bias detection:

$$\psi_2 = \sum_{i=k}^{i=k+N-1} r_i^T \mathrm{Var}_i^{-1} r_i \tag{22}$$

where N is the length of a time window. Here ψ_2 follows a χ^2 distribution with degrees of freedom to Nn. It is because that the residual r_i follows the normal distribution and is uncorrelated with each other. For a given level of significance α, we choose a test criterion $\chi^2_{Nn,1-\alpha}$. We declare a measurement bias if ψ_2 exceeds the criterion.

Next we use the GLR method to estimate the location and the magnitude of the measurement bias. For simplicity, we reduce the GLR method to the following equations:

$$\begin{aligned} b_m &= d_m / c_m \\ T_m &= d_m^2 / c_m \end{aligned} \tag{23}$$

$$\begin{aligned} d_m &= f_m^T \sum_{i=k}^{i=k+N-1} G_{i,m}^T \mathrm{Var}_i^{-1} r_i \\ c_m &= f_m^T \left(\sum_{i=k}^{i=k+N-1} G_{i,m}^T \mathrm{Var}_i^{-1} G_{i,m} \right) f_m \end{aligned} \tag{24}$$

Note that statistical properties of r_i are the prerequisite of the above equations.

If a measurement bias has been detected, we calculate T_m for every measurement. We consider the measurement corresponding to the maximum value of T_m as biased. We can also estimate the magnitude of the bias by computing b_m.

Discussion

An optimization view of the new framework

When constraints are state space equations (i.e., $D = I$), the classical least squares framework (CLSF) and the SLSF can be regarded as special linear forms of the moving-horizon-based least-squares estimation (MHE),[18] which may be the most widely used approach for NDDR. The CLSF is a complete information formulation and the SLSF is a moving horizon estimator with horizon size of one. Under the MHE framework, Rao et al. showed that the above two optimization models were equivalent in filtering.[19] They analyzed the relationship using forward dynamic programming technique.

Here we consider the same relationship in another way. In Appendix A, we use the sequential processing method to calculate the filtering results of the CLSF and obtain the SLSF, which is an intermediate result from the CLSF. The equivalency between the CLSF and the SLSF is due to the fact that the equation $\hat{X}_{k-1/k-1} = X_{k-1} + \zeta_{k-1}$ can fully summarize past information. From the derivation procedure, we can see that the emergence of the equation depends solely on the sequential properties of the linear equality constrained least squares solution and the structure of the filtering problem formulation. Moreover, we have shown that the equivalency of the two optimization models is not limited to the standard space equations constraints. It can be extended to a more general case.

A statistical view of the new framework

Here we want to explain the optimality of the SLSF solution from the viewpoint of statistics (i.e., why we choose such a deterministic optimization problem formulation for the linear filtering problem). An elegant and widely used approach for filtering problem is the stochastic formulation, which generates the famous KF. Note that $\hat{X}_{k/k}$ in KF is the linear unbiased minimum variance estimator. It minimizes the stochastic loss function $E[(x-\hat{x})^T(x-\hat{x})]$ and can be represented in the form of conditional mean $\hat{x} = E[x|y]$.

However, we have noticed that there also has been well-defined statistical interpretation of the state estimates generated by the SLSF. Given $D = I$ in our case, several researchers have noticed that the solution of the SLSF is equivalent to KF.[16,17,20] Furthermore, the reason why the SLSF can also provide the "most probable" estimates can be explained in terms of the maximum a posteriori (MAP) estimate under the Bayesian framework.[21,22] Note that the MAP estimator requires the priori knowledge about the unknown state variables (i.e., the probability destiny function $p(x_{k+1}|y_{1:k})$), which is similar to the priori estimate $\hat{x}_{k+1/k}$ defined in KF.

When matrix D is singular, we find difficulty in directly choosing an appropriate statistical criterion (e.g., the MAP criterion), which leads to the SLSF, for determining the "best" estimates. Note that the key difference between a singular model and a standard state space model is that the matrix D is singular. This singular matrix fundamentally changes an important property of a state space model: it cannot be used to predict the future. That is to say, assume that we have known the current state estimate $\hat{x}_{k/k}$ we can obtain a priori estimate $\hat{x}_{k+1/k}$ through the state space equations $\hat{x}_{k+1/k} = A\hat{x}_{k/k}$. But we cannot do the same thing by a

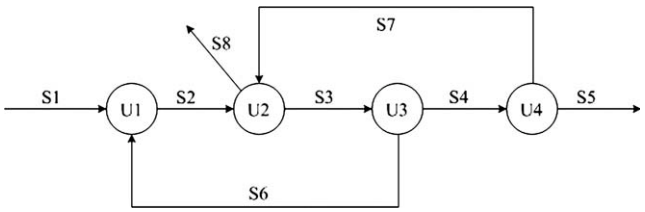

Figure 3. The process network.

singular model. This deficiency of a singular model hinders the MAP criterion and Bayesian interpretation.

However, as we have proved that the SLSF is an intermediate outcome of the batch least squares formulation, we can still explain the optimality of the estimates from the CLSF. The statistical interpretation for CLSF is obvious. For linear systems (i.e., regardless of a steady state model or a dynamic model) with uncorrelated Gaussian noise (i.e., the assumption made previously), the CLSF corresponds to the maximum likelihood estimator (MLE).

Interpretation of LSSDR from the viewpoint of filtering

Now let us consider the LSSDR problem. In a steady-state system, the state vector in the current time instant has no relationship with that in the past time. That is to say, the matrix B in the dynamic model keeps zero at every time instant. Therefore, without changing any assumption of the SLSF, we can obtain the basic results of LSSDR from the new framework:

Estimates:

$$\hat{X}_{i/i} = Z_i - VD^T\left(DVD^T\right)^{-1}DZ_i \tag{25}$$

Variance matrix of estimates:

$$\Sigma_{i/i} = V - VD^T\left(DVD^T\right)^{-1}DV \tag{26}$$

Residuals:

$$r_i = DZ_i \tag{27}$$

Variance matrix of residuals:

$$\mathrm{Var}[r_i] = DVD^T \tag{28}$$

Framework under process noise and rederivation of KF

In the preceding section, we have not taken model uncertainty into consideration. Now we consider process noise in our new framework. We assume the process model and the measurement model have the following structure:

$$DX_{k+1} = BX_k + w_{k+1} \tag{29}$$

$$Z_{k+1} = X_{k+1} + v_{k+1} \tag{30}$$

where X_{k+1} is a $n \times 1$ state vector, Z_{k+1} is a $n \times 1$ measurement vector, w_{k+1} is a $n \times 1$ process noise vector, and v_{k+1} is a $n \times 1$ measurement noise vector.

Here we assume that the measurement and process noise are zero-mean, normally distributed, mutually uncorrelated white noise sequences. The statistical properties of w_{k+1} and v_{k+1} can be represented by the following equations:

$$\begin{aligned} &\mathrm{Cov}(w_k, w_k) = R \\ &\mathrm{Cov}(v_k, v_k) = Q \\ &\mathrm{Cov}\left(w_k, w_j\right) = \mathrm{Cov}\left(v_{k,} v_j\right) = 0 \quad k \neq j \\ &\mathrm{Cov}\left(w_k, v_j\right) = 0 \end{aligned} \tag{31}$$

If we treat the process noise as an additional state vector, we can rewrite Eqs. 29 and 30 as follows:

$$DX_{k+1} = BX_k - S_{k+1} \tag{32}$$

$$Z_{k+1} = X_{k+1} + v_{k+1} \tag{33}$$

$$0 = S_{k+1} + w_{k+1} \tag{34}$$

The above equations also can take the following form when we use an augmented state vector:

$$D' X'_{k+1} = B' X'_k \tag{35}$$

$$Z'_{k+1} = X'_{k+1} + \varepsilon'_{k+1} \tag{36}$$

where

$$X'_{k+1} = \begin{bmatrix} X_{k+1} \\ S_{k+1} \end{bmatrix}, D' = [D : I], B' = [B \dot{:} 0], Z'_{k+1} = \begin{bmatrix} Z_{k+1} \\ 0 \end{bmatrix}, \varepsilon'_{k+1} = \begin{bmatrix} v_{k+1} \\ w_{k+1} \end{bmatrix}$$

Note that Eq. 35 can be regarded as a process model without model uncertainty. Meanwhile, according to the assumption represented by Eq. 31, new measurement noise ε_{k+1} is independent in time and normally distributed with zero mean. Thus the CLSF and the SLSF can be adopted for filtering in the system described by Eqs. 35 and 36. Now consider the filtering problem in the dynamic systems which can be represented by the state space equations. We have two optimal estimation techniques to deal with the problem. One is the famous KF. The other is the SLSF with $D' = [I,I]$, $B' = [A,0]$. The next problem is to examine whether they are consistent with each other. In Appendix D, we will rederive KF from the SLSF with generalized linear dynamic models constraints.

Simulation study

Comparative study of two DDR formulations

The objective of the present experiment is to compare the filtering accuracy and computational time between the simplified problem formulation and the classical problem formulation. The simplified problem formulation has the closed-form solutions represented by Eqs. 11 and 12. The classical problem formulation is directly solved by numerical optimization algorithms.

Figure 3 shows the material network used in the experiment. Measurement data were generated from true values in addition of normally distributed noises with zero mean. The variances of measurement data were taken from Daraouch's work.[9]

Figure 4 shows the true, measured, and two types of estimated values of inventories W1 and W4 of nodes U1 and

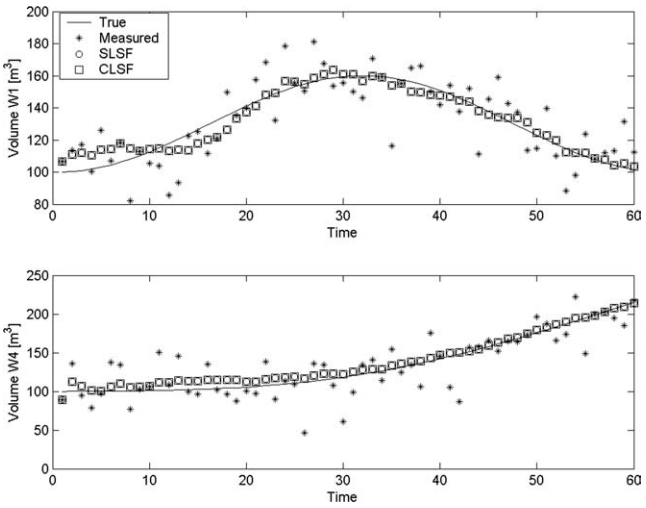

Figure 4. Filtering results of the SLSF and the CLSF.

U4, respectively. As shown in Figure 4, the filtering accuracy of two DDR formulations is almost same at any time instant. This result confirms the theoretical analysis that the simplified DDR formulation is equivalent to the classical formulation in the view of filtering.

The experiment was carried out for 60 time instants. The CPU time of each approach at a single time instant t_i ($i = 1 \ldots 60$) was recorded. We define the maximum processing time at a single time instant T_m and the average processing time T_{av} in the following equations:

$$T_m = \max(t_i) \tag{37}$$

$$T_{\text{av}} = \frac{\sum_{i=1}^{n} t_i}{n} \tag{38}$$

We found $T_m = 2.218$ s and $T_{\text{av}} = 0.6632$ s for the CLSF whereas $T_{\text{m}} = 0.016$ s and $T_{\text{av}} = 0.0003$ s for the SLSF. As expected, the SLSF is computationally more efficient than the CLSF.

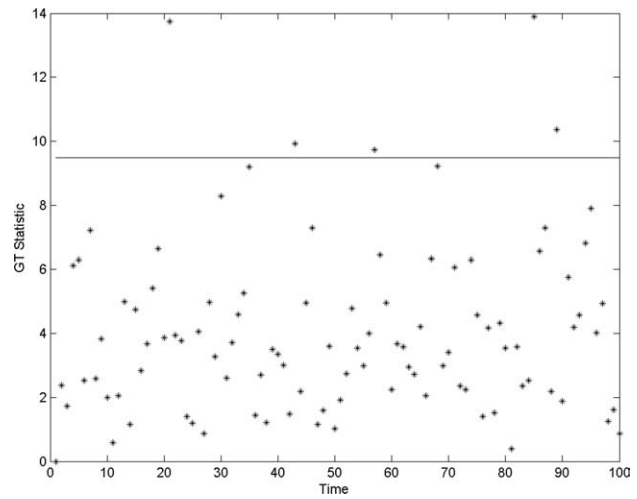

Figure 5. GT statistic (no measurement bias).

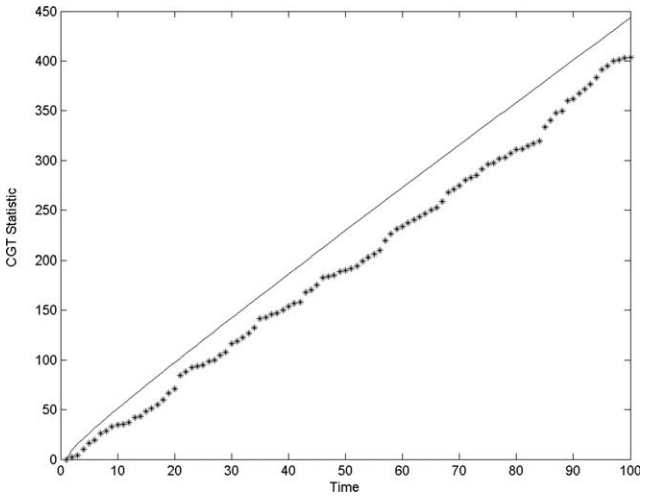

Figure 6. CGT statistic (no measurement bias).

Performance study of the bias detection methods

The objective of this experiment is to provide fundamental information about GT and CGT, thus to help us to choose one as the bias detection method in generalized linear dynamic systems. The process network shown in Figure 3 is used in this experiment. GT and CGT are calculated by Eqs. 21 and 22. The measurement variances are set to 1. The significance level α is set equal to 0.05 for both the test criterion. GT and CGT are compared in three scenarios: no measurement bias, only one flow measurement bias and only one inventory measurement bias.

Case 1: No measurement bias

Figures 5 and 6 show the values of GT and CGT when no measurement bias occurs, respectively. From Figure 6, it is observed that CGT does not reject the null hypothesis at any time instant. This is because the test criterion of CGT increases with time as the degree of freedom increases. This comparison study suggests that CGT is less apt to commit type I error than GT.

Case 2: One flow measurement bias

Here the flow measurement of stream S6 becomes biased from time instant 20 to time instant 60. As shown in Figures 7 and 8, GT and CGT both reject the null hypothesis from time 20 to 60. When the flow measurement returns to normal, GT no longer declares a measurement bias whereas CGT remains above the test criterion. It is because that CGT is an accumulation quantity. It uses all the measurement data in a time window to judge whether a bias appears or not at the current time instant. That is to say, most probably, if CGT has detected a bias at a time instant in the past, it may give an alarm until the current time even the bias has disappeared. It seems that CGT has a "memory" of past bias information thus has a delay time before accepting the null hypothesis. We should manually reset the CGT statistic to zero when we recalibrate the biased sensors. Compared with GT statistic, CGT seems a one-time GED strategy (i.e., it can be used to detect biases only once). However, in the practice,

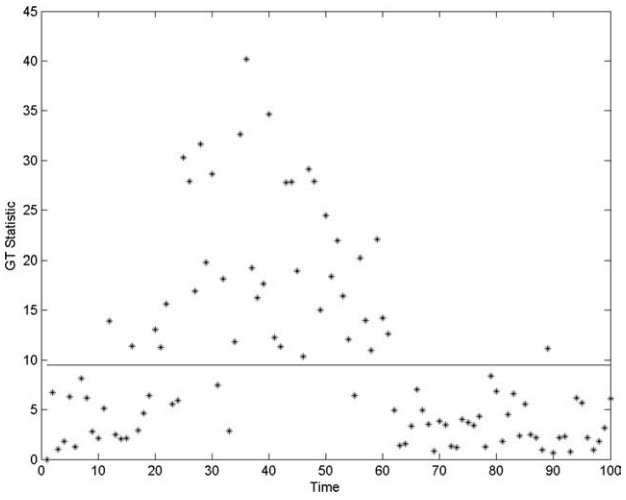

Figure 7. GT statistic (flow measurement bias).

we are often confronted with the following scenario: at first, all the sensors are under normal conditions. Then some biases occur at a time instant and a GED strategy provides an alarm. Then staff members recalibrate the biased instrument according to the information provided by GED. Under such a case, the key point of a GED strategy is its ability to detect biases rather than its postprocessing technology (i.e., it can return to normal when all the biases has disappeared). Thus in the current work, we did not pay much attention to the recovery ability of a GED strategy.

Case 3: One inventory measurement bias

In this experiment, the inventory measurement of node U1 was set biased from time instant 20. Figure 9 shows that GT only declares a bias at time instant 20 and 21; Afterward GT remains below the test criterion. This performance is totally different from that depicted in Figure 7. In this experiment, the distinction seems to be possible because an inventory bias' effects on the residuals can be rapidly canceled out over consecutive time instants. Note that it does not mean an inventory bias' effects on the estimates can disappear with an increase of time. The analysis is as follows.

According to the system model Eq. 3, we expand the residual formula Eq. 13 and obtain the following equation:

$$r_i = W_i - MQ_i - W_{i-1} + V_w\Omega_i r_{i-1} \tag{39}$$

Assume that an inventory measurement bias occurs at time instant k with a magnitude b. Thus we can directly obtain the following equation:

$$E[W_i] = W_i^* + bf'_m \quad i \geq k \tag{40}$$

$$E[r_k] = bf'_m \tag{41}$$

where f'_{m} is a $n \times 1$ vector in which the $m(m \leq n)$ th component is one and other components are all zeros.

Suppose that we want to estimate the expected value of r_{i+1} We perform the integration of Eqs. 39 and 40 and obtain the following equation:

$$\begin{aligned} E[r_{i+1}] &= E[W_{i+1}] - ME[Q_{i+1}] - E[W_i] + V_w\Omega_{i+1}E[r_i] \\ &= \underbrace{W_{i+1}^* + bf'_m - MQ_{i+1}^* - W_i^* - bf'_m}_{0} + V_w\Omega_{i+1}E[r_i] \\ &= V_w\Omega_{i+1}E[r_i] \end{aligned} \tag{42}$$

Daraouch[9] proves that Ω_i converges to a constant matrix Ω. Thus we rewrite Eq. 42 as follows:

$$E[r_{i+1}] = V_w\Omega E[r_i] \tag{43}$$

Then we can obtain the following equation:

$$E[r_{i+1}] = (V_w\Omega)^{i-k+1}bf'_m \quad i \geq k \tag{44}$$

Through calculation, we find that $(V_w\Omega)^{i-k+1}$ rapidly decreases to 0 and thus $E[r_i]$ converges to 0 with an increase in time. Therefore, in this experiment, as shown in Figure 9, GT only rejects the null hypothesis at few time instants and returns to normal rapidly although the inventory bias still exists. In contrast, as shown in Figure 10, CGT remains above the test criterion for a period since its memory of past information.

In summary, these three case studies indicate that CGT performs relatively better than GT as CGT exploits temporal redundancy of dynamic systems. Thus we choose CGT as the bias detection method for generalized linear dynamic systems. As shown in Figures 8 and 10, once a bias occurs, its abnormal effects can be exhibited by CGT for a period of time. To exploit this characteristic of CGT, we declare a measurement bias only if CGT rejects null hypothesis at two or three consecutive time instants. This empirical scheme may further lower the probability of type I error when α has been fixed.

Performance study of the CGT-GLR approach

The objective of the present experiment is to demonstrate the effects of the CGT-GLR approach and to investigate how a bias' characteristics influence the performance of this

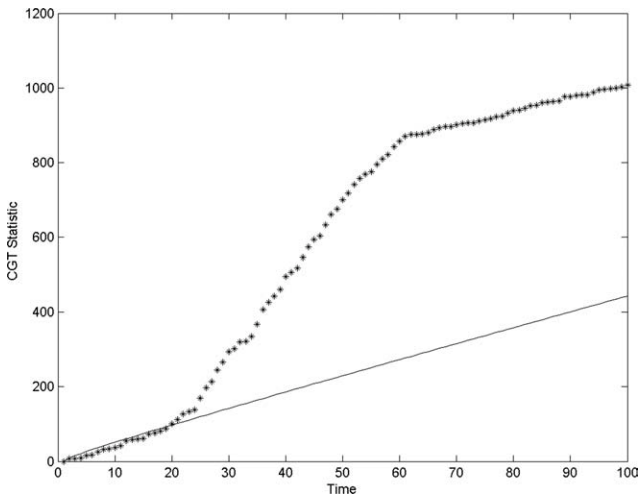

Figure 8. CGT statistic (flow measurement bias).

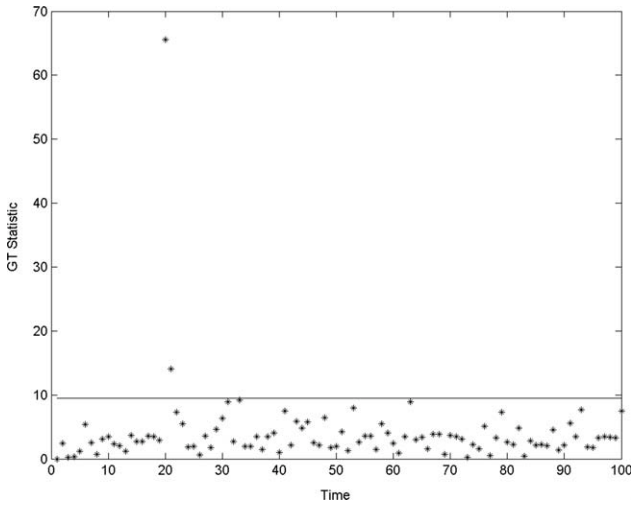

Figure 9. GT statistic (inventory measurement bias).

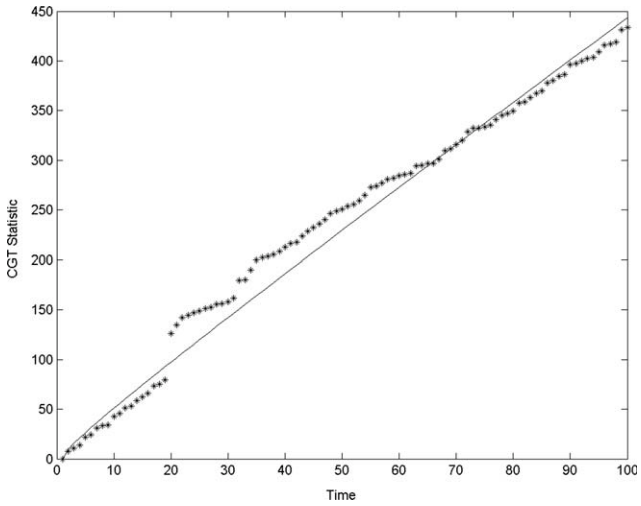

Figure 10. CGT statistic (inventory measurement bias).

approach. The experiment scheme is similar to that described by Rollins.[14]

The process network shown in Figure 3 is used in this study. This network consists of four inventory measurements and eight flow measurements. In a simulation run, only one measurement becomes biased. The magnitude, the location, and the true occurrence time of a bias (TTOC) are assigned in advance and remain unchanged during a simulation run. The magnitude of a measurement bias is chosen arbitrarily through the following inequality[17]:

$$l(x_i + \varepsilon_i) \leq \lambda_i \leq u(x_i + \varepsilon_i) \tag{45}$$

where l is a lower fraction and u is an upper fraction, x_i is the true value of the ith process variable and λ_i is the magnitude of the bias of the ith variable. Here let $l = 0.1$, $u = 0.5$ and x_i be the true value of the ith variable at the initial time. The bias' location is chosen from the 12 measurements. The measurement bias is generated at any time instant.

As a bias detection and identification procedure is stochastic in nature, performance of an identification approach should be evaluated through a suitably large number of simulation trials. In this study, one simulation run consists of 1000 simulation trials. Each trial consists of a new set of stochastic measurement data. These data are generated under the same conditions (i.e., same characteristics of the bias). In each trail, the CGT-GLR approach is performed in the following procedure. The cumulative GT is applied at every sampling period until it rejects the null hypothesis at several consecutive time instants (e.g., from time instant k to time instant $k + 3$). This time instant k is labeled as the estimated time of occurrence of a bias (ETOC). Then GLR method is used to identify the location and estimate the magnitude of the measurement bias. The CGT-GLR method in each simulation trial finally provide the following results: whether a measurement bias is presented or not; if a measurement bias exist, what its location, magnitude, and the occurrence time are.

When one simulation run is over, five measures are used to evaluate the performance of the CGT-GLR approach.

(1) The overall power (OP) is a measure for correct identification of a biased variable. It is defined as follows:

$$\text{OP} = \frac{\text{number of bias correctly identified}}{\text{number of bias simulated}} \tag{46}$$

(2) The average type I error (AVTI) is a measure for false identification of a biased variable. It is defined as follows:

$$\text{AVTI} = \frac{\text{number of bias wrongly identified}}{\text{number of simulation trials made}} \tag{47}$$

(3) Selectivity (SELE) is a measure for the probability of correct identification of a biased variable when a bias is detected. It is defined as follows:

$$\text{SELE} = \frac{\text{number of bias correctly identified}}{\text{number of bias declared}} \tag{48}$$

(4) The standard error of the average difference (STDD) is a measure for correct estimation of the occurrence time of a bias. It is defined as follows:

Table 1. Effect of Flow Measurement Bias Location (TTOC = 10)

Magnitude	Location	OP	AVTI	SELE	STDD	AEE
6	1	0.995	0.005	0.995	1.7099	0.0759
6	2	0.998	0.002	0.998	1.6657	0.0795
6	3	0.997	0.003	0.997	1.4882	0.0736
6	4	0.996	0.004	0.996	1.4434	0.0722
6	5	0.995	0.005	0.995	1.6787	0.0695
6	6	0.999	0.001	0.999	1.5117	0.0743
6	7	0.996	0.004	0.996	1.5305	0.0731
6	8	0.999	0.001	0.999	1.8861	0.082
10	1	1	0	1	1.1992	0.0533
10	2	0.999	0.001	0.999	1.246	0.0556
10	3	1	0	1	1.1992	0.0544
10	4	1	0	1	1.1393	0.0534
10	5	1	0	1	1.1832	0.0509
10	6	1	0	1	1.2087	0.0538
10	7	1	0	1	1.3255	0.0575
10	8	1	0	1	1.1358	0.0527

Table 2. Effect of Inventory Measurement Bias Location (TTOC = 10)

Magnitude	Location	OP	AVTI	SELE	STDD	AEE
13	1	0.912	0.066	0.9325	0.7994	0.1220
13	2	0.869	0.072	0.9245	0.9392	0.1330
13	3	0.916	0.066	0.9337	0.5984	0.1180
13	4	0.905	0.061	0.9378	0.5462	0.1161
15	1	0.937	0.056	0.9436	0.3426	0.0989
15	2	0.921	0.065	0.935	0.0466	0.1133
15	3	0.942	0.054	0.9458	0.2234	0.1032
15	4	0.938	0.061	0.9380	0.0385	0.1011

$$\mathrm{STDD} = \sqrt{\frac{\sum_{i=1}^{p} (\mathrm{ETOC}_i - \mathrm{TTOC})^2}{p}} \tag{49}$$

where p is the number of trials in which a bias is correctly identified.

(5) The average error of estimation (AEE) is a measure for accuracy of estimation of the bias magnitude. It is defined as follows:

$$\mathrm{AEE} = \frac{1}{p} \sum_{i=1}^{p} \left| \frac{\text{estimated value} - \text{actual value}}{\text{actual value}} \right| \tag{50}$$

In this study, parameters in the CGT-GLR method remain constant. The level of significance α used in CGT is chosen equal to 0.05 and the value of N used in GLR is set equal to 10.

The effect of the bias location is investigated in Tables 1 and 2. As shown in Tables 1 and 2, the location does not cause a significant difference in flow measurements or inventory measurements, respectively. However, it is observed that the measures OP and SELE in Table 2 are less than those in Table 1. It may be due to the fact that an inventory bias' effects on the residuals can be canceled out with an increase of time, which is explained in the preceding section. Thus CGT-GLR does not have enough information for detecting and identifying an inventory measurement bias. These results suggest that an inventory measurement bias is more difficult to be correctly identified than a flow measurement bias.

The effect of the magnitude of a bias is shown in Table 3. It is observed that measures OP and SELE increases with an increase of the magnitude. These results are obtained as expected because it is known that a bias which has a large magnitude is easily detected and correctly identified. Specially, due to its "disappearing effects," an inventory measurement bias which has a small magnitude may be ignored by CGT-GLR.

Table 3. Effect of the Bias Magnitude (An Inventory Measurement Bias at U1)

Magnitude	TTOC	OP	AVTI	SELE	STDD	AEE
8	5	0.501	0.15	0.7828	4.176	0.2374
10	5	0.77	0.079	0.9112	0.6894	0.1468
12	5	0.907	0.057	0.9428	0.1557	0.1257
14	5	0.9510	0.047	0.9529	0.1486	0.1135
16	5	0.9590	0.04	0.96	0.1769	0.1035
18	5	0.97	0.03	0.97	0.068	0.0824

Table 4. Effect of Measurement Bias TTOC (A Flow Measurement Bias at S1)

Magnitude	TTOC	OP	AVTI	SELE	STDD	AEE
6	10	0.995	0.005	0.995	1.6901	0.0769
6	20	0.952	0.048	0.952	2.0042	0.0709
6	30	0.929	0.071	0.929	2.338	0.0728
6	40	0.918	0.082	0.918	3.082	0.0752
10	10	1	0	1	1.1854	0.0512
10	20	0.951	0.049	0.951	1.2246	0.0478
10	30	0.93	0.07	0.93	1.2952	0.0517
10	40	0.92	0.08	0.92	1.236	0.0445

Table 4 shows that the measures OP and SELE decrease whereas STDD increases with an increase of TTOC. These results may be explained by the accumulation property of CGT. An increase of TTOC indicates that more normal residuals, which do not contain the effects of the bias, are used for computing CGT. If a bias is small and occurs late, its abnormal effects occupy only a small proportion of the accumulation amount. Thus the test statistic may not be sensitive to the measurement bias. We hope this disadvantage can be improved in the future work.

Conclusion

In this article, we have rigorously established a SLSF for LDDR in generalized linear dynamic systems. Besides naturally generating recursive filtering formulas, the new framework also gives rise to dynamic GED strategies. The SLSF is an extension of KF-based LDDR technique from state space models to singular systems. It also includes traditional LSSDR as the static limit case. We hope that the new framework can build a bridge between LDDR and LSSDR, and help to understand KF from a new perspective. We also hope that the new framework can provide a foundation for the research of simultaneous multiple gross errors detection and identification in linear dynamic systems in the future work.

Acknowledgments

The authors would like to acknowledge the support of the National Nature Science Foundation of China (No.60421002) and the National High Technology R&D Program (No.2007AA40702).

Literature Cited

1. Narasimhan S, Mah RSH. Generalized likelihood ratios for gross error identification in dynamic processes. *AIChE J*. 1988;34:1321–1331.
2. Stanley GM, Mah RSH. Estimation of flows and temperatures in process networks. *AIChE J*. 1977;23:642–650.
3. Willsky AS, Jones HLA. generalized likelihood ratio approach to state estimation in linear systems subject to abrupt changes. *Proc IEEE*. 1974;13:846–853.
4. Kao CS, Tamhane AC, Mah RSH. Gross error detection in serially correlated process data. II. Dynamic Systems. *Ind Eng Chem Res*. 1992;31:254–262.
5. Prakash J, Patwardhan SC, Narasimhan S. A supervisory approach to fault-tolerant control of linear multivariable systems. *Ind Eng Chem Res*. 2002;41:2270–2281.
6. Musulin E, Benqlilou C, Bagajewicz M, Puigjaner L. Instrumentation design based on optimal Kalman filtering. *J Process Control*. 2005;15:629–638.

7. Huang HP, Luo KY. On-line wavelets filtering with application to linear dynamic data reconciliation. *Ind Eng Chem Res*. 2007;46:8746–8755.
8. Almasy GA. Principles of dynamic balancing. *AIChE J*. 1990;36:1321–1330.
9. Darouach M, Zasadzinski M. Data reconciliation in generalized linear dynamic systems. *AIChE J*. 1991;37:193–200.
10. Veverka VV, Madron F. *Material and Energy Balancing in the Process Industries: From Microscopic Balances to Large Plants*. Amsterdam: Elsevier, 1997.
11. Rollins DK, Devanathan S. Unbiased estimation in dynamic data reconciliation. *AIChE J*. 1993;39:1330–1334.
12. Ramamurthi Y, Sistu PB, Bequette BW. Control-relevant dynamic data reconciliation and parameter estimation. *Comput Chem Eng*. 1993;17:41–59.
13. Bagajewicz M, Jiang QY. Integral approach to plant linear dynamic reconciliation. *AIChE J*. 1997;43:2546–2558.
14. Rollins DK, Devanathan S, Victoria M, Bascuñana B. Measurement bias detection in linear dynamic systems. *Comput Chem Eng*. 2002;26:1201–1211.
15. Romagnoli JA, Stephanopoulos G. Rectification of process measurement data in the presence of gross errors. *Chem Eng Sci*. 1981;36:1849–1963.
16. Romagnoli JA, Sanchez MC. *Data Processing and Reconciliation for Chemical Process Operations*. San Diego: Academic Press, 2000.
17. Narasimahn S, Jordache C. *Data reconciliation And Gross Error Detection: An Intelligent Use of Process Data*. Houston: Gulf Publishing Company, 2000.
18. Robertson DG, Lee JH, Rawlings JB. A moving horizon-based approach for least-squares estimation. *AIChE J*. 1996;42:2209–2224.
19. Rao CV, Rawlings JB, Lee JH. Constrained linear state estimation—a moving horizon approach. *Automatica*. 2001;37:1619–1628.
20. Vachhani P, Rengaswamy R, Gangwal V, Narasimahan S. Recursive estimation in constrained nonlinear dynamical systems. *AIChE J*. 2005;51:946–959.
21. Bai SH, Thibault J, McLean DD. Dynamic data reconciliation: alternative to Kalman filter. *J Process Control*. 2006;16:485–498.
22. Chen WS, Bakshi BR, Goel PK, Ungarala S. Bayesian estimation via sequential monte carlo sampling: unconstrained nonlinear dynamic system. *Ind Eng Chem Res*. 2004;43:4012–4025.

Appendix A: Equivalence Between the Filtering Results of CLSF and SLSF

Sequential processing of the linear data reconciliation solution

Consider the CLSF (Eq. 9). We rewrite this framework in a compact form:

$$\begin{aligned} \min_{\delta} \quad & \tfrac{1}{2}\delta^T\psi\delta \\ s.t. \quad & \phi\hat{X} = 0 \\ & Z = \hat{X} + \delta \end{aligned} \tag{A1}$$

$$\hat{X} = \begin{bmatrix} \hat{X}_1 \\ \hat{X}_2 \\ \vdots \\ \hat{X}_c \end{bmatrix}, \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_c \end{bmatrix},$$

$$\phi = \begin{bmatrix} B & -D & 0 & . & . & . & 0 \\ 0 & B & -D & 0 & . & . & 0 \\ . & . & . & . & . & . & . \\ 0 & . & . & . & 0 & B & -D \end{bmatrix} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_c \end{bmatrix}, Z = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_c \end{bmatrix}$$

$$\psi = \begin{bmatrix} V^{-1} & 0 & \cdots & 0 \\ 0 & V^{-1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & V^{-1} \end{bmatrix}$$

where $\hat{X}$ is a vector of estimates from time instant 1 to c, Z is a vector of measurements from time instant 1 to c, δ is a vector of measurement errors from time instant 1 to c, ϕ is a constraint matrix which can be divided into c submatrices $\phi_1,\phi_2,\ldots\phi_c$ and ψ is a weight matrix of measurement errors.

According to the sequential processing method proposed by Romagnoli,[15,16] the overall constrained least squares problem (A1) can be equivalently replaced by the following c subproblems.

Problem 1.

$$\begin{aligned} \min_{\delta^{(1)}} \quad & \frac{1}{2}\delta^{(1)T}\psi^{(1)}\delta^{(1)} \\ s.t. \quad & \phi_1\hat{X}^{(1)} = 0 \\ & Z^{(1)} = \hat{X}^{(1)} + \delta^{(1)} \end{aligned} \tag{A2}$$

where $\psi^{(1)} = \psi$, $Z^{(1)} = Z$, $\delta^{(1)} = \delta$, and $E[\delta^{(1)}] = 0$.

Solve this preceding problem. We obtain the estimates $\hat{X}^{(1)}$ and its covariance matrix $\Sigma^{(1)}$.

Problem 2.

In the second problem, the adjustments involve in it are not only the original measurements but also some of the new estimates from the previous stage. This problem can be formulated as follows:

$$\begin{aligned} \min_{\delta^{(2)}} \quad & \frac{1}{2}\delta^{(2)T}\psi^{(2)}\delta^{(2)} \\ s.t. \quad & \phi_2\hat{X}^{(2)} = 0 \\ & Z^{(2)} = \hat{X}^{(2)} + \delta^{(2)} \end{aligned} \tag{A3}$$

where $\psi^{(2)} = (\Sigma^{(1)})^{-1}$, $Z^{(2)} = \hat{X}^{(1)}$, and $E[\delta^{(2)}] = 0$.

The rest problems' formulation (*problem 3* to *problem c-1*) can be deduced by analogy. Finally, we arrive at the last problem.

Problem c

$$\begin{aligned} \min_{\delta^{(c)}} \quad & \frac{1}{2}\delta^{(c)T}\psi^{(c)}\delta^{(c)} \\ s.t. \quad & \phi_c\hat{X}^{(c)} = 0 \\ & Z^{(c)} = \hat{X}^{(c)} + \delta^{(c)} \end{aligned} \tag{A4}$$

where $\psi^{(c)} = (\Sigma^{(c-1)})^{-1}$, $Z^{(c)} = \hat{X}^{(c-1)}$ and $E[\delta^{(c)}] = 0$.

Note that the estimate vector $\hat{X}^{(c)}$ is equal to $\hat{X}$, which is the result of the overall problem (A1).

Derive the SLSF in the view of filtering

Here our aim is to obtain the estimate $\hat{X}_c$ on the basis of Z_1, Z_2, Z_3,……Z_{c-1}, Z_c. From *Problem* 1 to *Problem* c-1, we have:

$$\hat{X}_c^{(k)} = Z_c, \quad k = 1, 2\ldots c-1 \tag{A5}$$

where $\hat{X}_c^{(k)}$ is the cth element of $\hat{X}^{(k)}$.

The reason is that the corresponding part of $\hat{X}_c^{(i)}$ in the incidence matrix (i.e., $\phi_i(1 \le i \le c-1)$) remains zero. Namely, model constraints in the past time instants do not provide any redundant information for estimating process variables at the current time instant c.

The *Problem c* (A4) can be formulated in detail as follows:

$$\begin{aligned} \min_{\delta_c\,\beta} \quad & \frac{1}{2}\delta_c^T V^{-1}\delta_c + \frac{1}{2}\beta^T\Psi^{-1}\beta \\ s.t. \quad & BX_{c-1} - DX_c = 0 \\ & \begin{bmatrix} \hat{X}_{1/c-1} \\ \hat{X}_{2/c-1} \\ \vdots \\ \hat{X}_{c-1/c-1} \\ Z_c \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{c-1} \\ X_c \end{bmatrix} + \begin{bmatrix} \hat{\delta}_1 \\ \hat{\delta}_2 \\ \vdots \\ \hat{\delta}_{c-1} \\ \delta_c \end{bmatrix} \end{aligned} \tag{A6}$$

$$\beta = \begin{bmatrix} \hat{\delta}_1 \\ \hat{\delta}_2 \\ \vdots \\ \hat{\delta}_{c-1} \end{bmatrix}, \Psi = \begin{bmatrix} \Sigma_{11} & \dots & \Sigma_{1(c-1)} \\ \dots & \dots & \dots \\ \Sigma_{(c-1)1} & \dots & \Sigma_{(c-1)(c-1)} \end{bmatrix}, \hat{X}_{i/c-1}$$

is identical to $\hat{X}_i^{(c-1)}$.

Note that Σ_{ij} is the (i, j) block of dimensions $(n+v) \times (n+v)$, which represents the covariance matrix of the vectors $\hat{X}_{i/c-1}$ and $\hat{X}_{j/c-1}$.

The solution of problem (A6) is obtained by means of the Lagrange multipliers method. The formulation is as follows:

$$\begin{aligned} L &= \frac{1}{2}\delta_c^T V^{-1}\delta_c + \frac{1}{2}\beta^T\Psi^{-1}\beta - \lambda^T(BX_{c-1} - DX_c) \\ \frac{\partial L}{\partial X_c} &= V^{-1}\delta_c + D^T\lambda = 0 \\ \begin{bmatrix} \frac{\partial L}{\partial X_1} \\ \frac{\partial L}{\partial X_2} \\ \dots \\ \frac{\partial L}{\partial X_{c-1}} \end{bmatrix} &= \Psi^{-1}\begin{bmatrix} \hat{\delta}_1 \\ \hat{\delta}_2 \\ \dots \\ \hat{\delta}_{c-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ B^T\lambda \end{bmatrix} \Rightarrow \begin{bmatrix} \hat{\delta}_1 \\ \hat{\delta}_2 \\ \dots \\ \hat{\delta}_{c-1} \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_{1(c-1)}B^T\lambda \\ \Sigma_{2(c-1)}B^T\lambda \\ \dots \\ \Sigma_{(c-1)(c-1)}B^T\lambda \end{bmatrix} \\ \frac{\partial L}{\partial \lambda^T} &= BX_{c-1} - DX_c = 0 \end{aligned} \tag{A7}$$

We only consider the estimates vector X_c and X_{c-1}. From Eq. A7, we extract the following equations to obtain these two vectors:

$$\begin{aligned} & V^{-1}\delta_c + D^T\lambda = 0 \\ & \Sigma_{(c-1)(c-1)}^{-1}\hat{\delta}_{c-1} - B^T\lambda = 0 \\ & BX_{c-1} - DX_c = 0 \\ & Z_c = X_c + \delta_c \\ & \hat{X}_{c-1/c-1} = X_{c-1} + \hat{\delta}_{c-1} \end{aligned} \tag{A8}$$

We can also derive Eq. A8 from the following optimization formulation:

$$\begin{aligned} \min_{\delta_c,\hat{\delta}_{c-1}} \quad & \frac{1}{2}\delta_c^T V^{-1}\delta_c + \frac{1}{2}\hat{\delta}_{c-1}^T\Sigma_{(c-1)(c-1)}^{-1}\hat{\delta}_{c-1} \\ s.t. \quad & DX_c = BX_{c-1} \\ & Z_c = X_c + \delta_c \\ & \hat{X}_{c-1/c-1} = X_{c-1} + \hat{\delta}_{c-1} \end{aligned} \tag{A9}$$

Eq. A9 is the SLSF defined by Eq. 10. Thus our conclusion is that Eq. 9 is equivalent to Eq. 10 in the view of filtering.

Appendix B: The Residuals at Different Time Instants are Uncorrelated

Problem statement

Given the following equations:

$$\hat{X}_{i/i} = Z_i + VD^T\Omega_i(B\hat{X}_{i-1/i-1} - DZ_i) \tag{B1}$$

$$\Omega_i = (B\Sigma_{i-1/i-1}B^T + DVD^T)^{-1} \tag{B2}$$

$$\Sigma_{i/i} = V - VD^T\Omega_i DV \tag{B3}$$

$$r_i = DZ_i - B\hat{X}_{i-1/i-1} \tag{B4}$$

$$\begin{aligned} \mathrm{Var}[r_i] &= \mathrm{Cov}(r_i, r_i) \\ &= DVD^T + B\Sigma_{i-1/i-1}B^T \\ &= \Omega_i^{-1} \end{aligned} \tag{B5}$$

where Eqs. B1–B3 constitute the filtering formulas, Eq. B4 is the definition of the residual vector, Eq. B5 is the variance matrix of residual.

We want to prove the following equation:

$$\mathrm{Cov}(r_i,r_j) = 0 \quad i \ne j \tag{P1}$$

An equivalent problem of P1 can be stated as follows: given an residual sequence $[r_1\ r_2\ r_3 \dots\dots r_i\ r_{i+1}]$ the following i equations are true:

$$\mathrm{Cov}(r_{i+1}, r_k) = 0 \quad \forall k, 1 \le k \le i \tag{P2}$$

Preparation

For preparation, we obtain the following recursive formulas through substituting Eq. B1 into Eq. B4:

$$\begin{aligned} r_{i+1} &= DZ_{i+1} - B\hat{X}_{i/i} \\ &= DZ_{i+1} - B[Z_i - VD^T\Omega_i \underbrace{(DZ_i - B\hat{X}_{i-1/i-1})}_{r_i}] \\ &= DZ_{i+1} - BZ_i + BVD^T\Omega_i r_i \end{aligned} \tag{B6}$$

Then we rewrite Eq. B1 as follows:

$$\hat{X}_{i/i} = (I - VD^T\Omega_i D)Z_i + VD^T\Omega_i B\hat{X}_{i-1/i-1} \tag{B7}$$

Next we want to prove the following three lemmas:

Lemma 1.

$$Cov(\hat{X}_{k/k}, Z_k) = \Sigma_{k/k}$$

Proof:

$$\begin{aligned}
&\mathrm{Cov}(\hat{X}_{k/k}, Z_k)\\
&= Cov((I - VD^T\Omega_k D)Z_k + VD^T\Omega_k B\hat{X}_{k-1/k-1}, Z_k)\\
&= (I - VD^T\Omega_k D)\underbrace{Cov(Z_k, Z_k)}_{v} + VD^T\Omega_k B\underbrace{Cov(\hat{X}_{k-1/k-1}, Z_k)}_{0}\\
&= \underbrace{V - VD^T\Omega_k DV}_{\text{Eq.B3}}\\
&= \Sigma_{k/k}
\end{aligned}$$

■

Lemma 2.

$$\mathrm{Cov}(\hat{X}_{k/k}, DZ_k - BZ_{k-1}) = 0$$

Proof:

$$\begin{aligned}
&\mathrm{Cov}(\hat{X}_{k/k},\ DZ_k - BZ_{k-1})\\
&= \mathrm{Cov}((I - VD^T\Omega_k D)Z_k + VD^T\Omega_k B\hat{X}_{k-1/k-1}, DZ_k - BZ_{k-1})\\
&= (I - VD^T\Omega_k D)\underbrace{\mathrm{Cov}(Z_k, Z_k)}_{v} D^T - (I - VD^T\Omega_k D)\\
&\quad \times \underbrace{\mathrm{Cov}(Z_k, Z_{k-1})}_{0} B^T + VD^T\Omega_k B\underbrace{\mathrm{Cov}(\hat{X}_{k-1/k-1}, Z_k)}_{0} D^T\\
&- VD^T\Omega_k B\underbrace{\mathrm{Cov}(\hat{X}_{k-1/k-1}, Z_{k-1})}_{\Sigma_{k-1/k-1}:\text{from lemma1}} B^T\\
&= (I - VD^T\Omega_k D)VD^T - VD^T\Omega_k B\Sigma_{k-1/k-1}B^T\\
&= VD^T(I - \Omega_k DVD^T - \Omega_k B\Sigma_{k-1/k-1}B^T)\\
&= VD^T\Omega_k\underbrace{(\Omega_k^{-1} - DVD^T - B\Sigma_{k-1/k-1}B^T)}_{\text{Eq.B5}} = 0
\end{aligned}$$

■

Lemma 3.

$$\mathrm{Cov}(\hat{X}_{i/i}, DZ_k - BZ_{k-1}) = 0 \quad 1 < k \le i - 1$$

Proof:

$$\begin{aligned}
&\mathrm{Cov}(\hat{X}_{k+1/k+1}, DZ_k - BZ_{k-1})\\
&= \mathrm{Cov}((I - VD^T\Omega_{k+1}D)Z_{k+1} + VD^T\Omega_{k+1}B\hat{X}_{k/k}, DZ_k\\
&- BZ_{k-1}) = (I - VD^T\Omega_{k+1}D)\underbrace{\mathrm{Cov}(Z_{k+1}, Z_k)}_{0} D^T\\
&- (I - VD^T\Omega_{k+1}D)\underbrace{\mathrm{Cov}(Z_{k+1}, Z_{k-1})}_{0} B^T\\
&+ VD^T\Omega_{k+1}B\underbrace{\mathrm{Cov}(\hat{X}_{k/k}, DZ_k - BZ_{k-1})}_{0:\text{from Lemma2}}\\
&= 0
\end{aligned}$$

Then from $\mathrm{Cov}(\hat{X}_{k+1/k+1}, DZ_k - BZ_{k-1}) = 0$, we can obtain the following equation by the same manners:

$$\mathrm{Cov}(\hat{X}_{k+2/k+2}, DZ_k - BZ_{k-1}) = 0$$

From $\mathrm{Cov}(\hat{X}_{k+2/k+2}, DZ_k - BZ_{k-1}) = 0$, we can also get the following equation:

$$\mathrm{Cov}(\hat{X}_{k+3/k+3}, DZ_k - BZ_{k-1}) = 0$$

Through this recursive manner, we can finally obtain the target equation:

$$\mathrm{Cov}(\hat{X}_{i/i}, DZ_k - BZ_{k-1}) = 0$$

■

Proof of the proposition P2

Here we use mathematical induction to validate the proposition P2.

(1) The basis:

$$\mathrm{Cov}(r_{i+1}, r_k) = 0 \quad k = i \tag{P3}$$

Proof:

$$\begin{aligned}
&\mathrm{Cov}(r_{i+1}, r_i) = \mathrm{Cov}(DZ_{i+1} - BZ_i + BVD^T\Omega_i r_i, r_i)\\
&= \mathrm{Cov}(DZ_{i+1} - BZ_i, \underbrace{DZ_i - B\hat{X}_{i-1/i-1}}_{r_i})\\
&+ BVD^T\Omega_i\underbrace{\mathrm{Cov}(r_i, r_i)}_{\Omega_i^{-1}} = D\underbrace{\mathrm{Cov}(Z_{i+1}, Z_i)}_{0} D^T\\
&- D\underbrace{\mathrm{Cov}(Z_{i+1}, \hat{X}_{i-1/i-1})}_{0} B^T - B\underbrace{\mathrm{Cov}(Z_i, Z_i)}_{V} D^T\\
&+ B\underbrace{\mathrm{Cov}(Z_i, \hat{X}_{i-1/i-1})}_{0} B^T + BVD^T\Omega_i\Omega_i^{-1}\\
&= -BVD^T + BVD^T = 0
\end{aligned}$$

■

(2) The inductive step:

$$\text{If } \mathrm{Cov}(r_{i+1}, r_k) = 0, \text{then } \mathrm{Cov}(r_{i+1}, r_{k-1}) = 0, \quad k \le i - 1 \tag{P4}$$

Proof:

From Eq. B6, we can obtain the equation:

$$r_k = DZ_k - BZ_{k-1} + BVD^T\Omega_{k-1}r_{k-1} \tag{B8}$$

Eq. B8 equivalently becomes:

$$\underbrace{\mathrm{Cov}(r_{i+1}, r_k)}_{0:\text{from assumption}} = Cov(r_{i+1,}DZ_k - BZ_{k-1} + BVD^T\Omega_{k-1}r_{k-1}) \tag{B9}$$

We expand the right-hand side of Eq. B9:

$$\begin{aligned}
&\mathrm{Cov}(r_{i+1,}DZ_k - BZ_{k-1} + BVD^T\Omega_{k-1}r_{k-1})\\
&= \mathrm{Cov}(r_{i+1}, DZ_k - BZ_{k-1}) + \mathrm{Cov}(r_{i+1}, BVD^T\Omega_{k-1}r_{k-1})\\
&= \mathrm{Cov}(DZ_{i+1} - B\hat{X}_{i/i}, DZ_k - BZ_{k-1}) + \mathrm{Cov}(r_{i+1}, r_{k-1})\Omega_{k-1}DVB^T\\
&= -B\underbrace{\mathrm{Cov}(\hat{X}_{i/i}, DZ_k - BZ_{k-1})}_{0:\text{from Lemma3}} + \mathrm{Cov}(r_{i+1}, r_{k-1})\Omega_{k-1}DVB^T\\
&= 0
\end{aligned}$$

Finally we obtain the equation:

$$\mathrm{Cov}(r_{i+1}, r_{k-1})\Omega_{k-1}DVB^T = 0 \tag{B10}$$

Since Ω_{k-1},V are invertible matrices and D,B are matrices of full row rank specified in Eq. 3 or 8, $\Omega_{k-1}DVB^T$ is an invertible matrix. Thus from Eq. B10, we can obtain the following equation:

$$\mathrm{Cov}(r_{i+1}, r_{k-1}) = 0 \tag{B11}$$

■

Appendix C: Recursion Formulas for Computing the Signature Matrix

According to the procedure described by Narasimhan,[1] we will compute the matrix $G_{i,m}$.

Now we divide the measurements, the estimates and the residuals into two parts, respectively:

$$\begin{aligned} Z_i &= Z_i^1 + Z_i^2 \\ \hat{X}_{i/i} &= \hat{X}_{i/i}^1 + \hat{X}_{i/i}^2 \\ r_i &= r_i^1 + r_i^2 \end{aligned} \tag{C1}$$

where $Z_i^1, \hat{X}_{i/i}^1, r_i^1$ are free of measurement biases and $Z_i^2, \hat{X}_{i/i}^2, r_i^2$ lie on the effects of a measurement bias. The separation can be done since the system models and the estimates are linear.

Next we define $Z_i^2, \hat{X}_{i/i}^2, r_i^2$ by using the following equations:

$$\begin{aligned} Z_i^2 &= bf_m \\ \hat{X}_{i/i}^2 &= bJ_{i,m}f_m \\ r_i^2 &= bG_{i,m}f_m \end{aligned} \tag{C2}$$

On substituting Eq. C1 into Eq. 11, we obtain the following equation:

$$\begin{aligned} \hat{X}_{i/i}^1 + \hat{X}_{i/i}^2 = Z_i^1 + Z_i^2 + VD^T\Omega_i\{B(\hat{X}_{i-1/i-1}^1 \\ + \hat{X}_{i-1/i-1}^2) - D(Z_i^1 + Z_i^2)\} \Rightarrow \hat{X}_{i/i}^2 = Z_i^2 \\ + VD^T\Omega_i(B\hat{X}_{i-1/i-1}^2 - DZ_i^2) \end{aligned} \tag{C3}$$

Through integrating Eq. C2 into Eq. C3, we have the following equations:

$$\begin{aligned} bJ_{i,m}f_m &= bf_m + VD^T\Omega_i(BbJ_{i-1,m}f_m - Dbf_m) \\ \Rightarrow J_{i,m} &= I + VD^T\Omega_i(BJ_{i-1,m} - D) \end{aligned} \tag{C4}$$

By analogous, we integrate Eq. 10, Eqs. C1 and C2:

$$\begin{aligned} r_i^2 &= DZ_i^2 - B\hat{X}_{i-1/i-1}^2 \\ \Rightarrow G_{i,m} &= D - BJ_{i-1,m} \end{aligned} \tag{C5}$$

Eqs. C4 and C5 constitute the recursion formulas for computing $G_{i,m}$.

Appendix D: Rederivation of KF from the New DDR Framework

Here we want to obtain KF when the measurement matrix is identity.

We directly use Eqs. 10, 11 and extract the recursive formulas for state estimate $\hat{x}_{k+1/k+1}$ (note that the definition of $\Sigma_{k+1/k}$, K_k, $\hat{x}_{k+1/k}$ in KF):

$$\begin{aligned} \hat{x}_{k+1/k+1} &= y_{k+1} + Q(A\Sigma_{k/k}A^T + R + Q)^{-1}(A\hat{x}_{k/k} - y_{k+1}) \\ &= Q\underbrace{(A\Sigma_{k/k}A^T + R}_{\Sigma_{k+1/k}} + Q)^{-1}\underbrace{A\hat{x}_{k/k}}_{\hat{x}_{k+1/k}} \\ &\quad + \underbrace{(I - Q(A\Sigma_{k/k}A^T + R + Q)^{-1})}_{K_k}y_{k+1} \\ &= (I - K_k)\hat{x}_{k+1/k} + K_k y_{k+1} \\ &= \hat{x}_{k+1/k} + K_k(y_{k+1} - \hat{x}_{k+1/k}) \end{aligned} \tag{D1}$$

$$\begin{aligned} \Sigma_{k+1/k+1} &= Q - Q\underbrace{(A\Sigma_{k/k}A^T + R}_{\Sigma_{k+1/k}} + Q)^{-1}Q = (\Sigma_{k+1/k}^{-1} + Q^{-1})^{-1} \\ &= (I - \underbrace{\Sigma_{k+1/k}(\Sigma_{k+1/k} + Q)^{-1}}_{K_k})\Sigma_{k+1/k} \\ &= (I - K_k)\Sigma_{k+1/k} \end{aligned} \tag{D2}$$

Manuscript received July 13, 2008, and revision received Sept. 13, 2009.