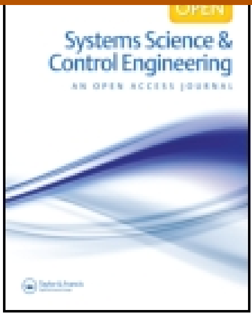

Systems Science & Control Engineering

An Open Access Journal

ISSN: (Print) 2164-2583 (Online) Journal homepage: http://www.tandfonline.com/loi/tssc20

Simultaneous state and input estimation with partial information on the inputs

Jinya Su, Baibing Li & Wen-Hua Chen

To cite this article: Jinya Su, Baibing Li & Wen-Hua Chen (2015) Simultaneous state and input estimation with partial information on the inputs, Systems Science & Control Engineering, 3:1, 445-452, DOI: 10.1080/21642583.2015.1082512

To link to this article: http://dx.doi.org/10.1080/21642583.2015.1082512

Accepted author version posted online: 26 Aug 2015.
Published online: 11 Sep 2015.

Systems Science & Control Engineering: An Open Access Journal, 2015
Vol. 3, 445–452, http://dx.doi.org/10.1080/21642583.2015.1082512

Taylor & Francis
Taylor & Francis Group

Simultaneous state and input estimation with partial information on the inputs

Jinya Su[a], Baibing Li[b*] and Wen-Hua Chen[a]

[a] *Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough LE11 3TU, UK;*
[b] *School of Business and Economics, Loughborough University, Loughborough LE11 3TU, UK*

(*Received 28 August 2014; accepted 10 August 2015*)

This paper investigates the problem of simultaneous state and input estimation (SSIE) for discrete-time linear stochastic systems when the information on the inputs is partially available. To incorporate the partial information on the inputs, matrix manipulation is used to obtain an equivalent system with reduced-order inputs. Then Bayesian inference is drawn to obtain a recursive filter for both state and input variables. The proposed filter is an extension of the recently developed state filter with partially observed inputs to the case where the input filter is also of interest, and an extension of the SSIE to the case where the information on the inputs is partially available. A numerical example is given to illustrate the proposed method. It is shown that, due to the additional information on the inputs being incorporated in the filter design, the performances of both state and input estimation are substantially improved in comparison with the conventional SSIE without partial input information.

Keywords: Bayesian inference; partial information; state filter; unknown input filter

1. Introduction

State estimation for discrete-time linear stochastic systems with unknown inputs has been receiving increasing attention (see, e.g. Cheng, Ye, Wang, & Zhou, 2009; Darouach & Zasadzinski, 1997; Darouach, Zasadzinski, & Boutayeb, 2003; Hsieh, 2000; Kitanidis, 1987, among many others) due to its widespread applications in the fields of weather forecasting (Kitanidis, 1987), fault diagnosis (Mann & Hwang, 2013), etc.

In some applications such as population estimation, traffic management (Li, 2013), and chemical engineering (Mann & Hwang, 2013), however, information on the input variables is not completely unknown; rather, it is available at an aggregate level. Li (2013) has recently proposed a unified filtering approach to incorporate this kind of information. It is shown that this approach includes two extreme scenarios as its special cases, that is, the filter where all the inputs are unknown (i.e. the scenario investigated in Kitanidis, 1987; Gillins & De Moor, 2007, etc.) and the filter where the inputs are completely available (i.e. the classical Kalman filter can be applied). Later, Su, Li, and Chen (2015a) further investigated some properties of the aforementioned unified filter such as existence, optimality and asymptotic stability. However, Li (2013) and Su et al. (2015a) only considered the problem of sole state estimation; the problem of simultaneous state and input estimation (SSIE) with partial information on the inputs has not been investigated.

Gillins and De Moor (2007) developed a SSIE method using the approach of minimum-variance unbiased estimation (MVUE), then Fang and Callafon (2012) further investigated its asymptotic stability. Potentially, the SSIE can be applied to a wide range of problems such as fault diagnosis (see, e.g. Gao & Ding, 2007; Patton, Clark, & Frank, 1989), fault-tolerant control (Jiang & Fahmida, 2005), disturbance rejection control (Profeta, Joseph, William, & Marin, 1990). In the field of fault detection and fault-tolerant control, for example, actuator, sensor and/or structure faults are usually modelled as inputs to the system with unknown dynamics. One can monitor system status by estimating the inputs for fault diagnosis purposes where the estimated inputs can provide valuable information for the fault-tolerant control system; see, for example, Jiang and Fahmida (2005) and Su, Chen, and Li (2014). In the field of disturbance rejection control, the uncertainties in system model are usually modelled as lumped system inputs (which may include system mismatches, parameter uncertainties, external disturbances); see Chen, Ballance, Gawthrop, and O'Reilly (2000), Yang, Li, Su, and Yu (2013), and Yang, Su, Li, and Yu (2014) for a detailed discussion. When inputs are approximately obtained based on disturbance estimation algorithms, one can attenuate their effects on dynamic systems by directly feedthrough of the estimated value.

In this paper, we investigate the problem of SSIE. Unlike Gillins and De Moor (2007) where the inputs are

*Corresponding author. Email: b.li2@lboro.ac.uk

assumed to be completely unknown, we consider the scenario where the information on the inputs is partially available. To incorporate the partial information on the inputs, the original inputs are decoupled into two parts, where the first part is completely known based on the available information on the original inputs, whereas the second part is completely unknown which will serve as the unknown inputs of the new system. On this basis, we draw Bayesian inference (see, e.g. Li, 2009, 2013) and obtain simultaneous estimates of the state and new unknown inputs. According to the Bayesian theory, the obtained estimates are optimal in the sense of minimum mean square estimation under the assumption of Gaussian noise terms (Li, 2013). Finally, the estimates of the original inputs can be worked out by pooling together all the available information on the inputs.

Compared with the filter in Li (2013) where only state estimate is of interest, the proposed method obtains SSIE, and hence the estimated inputs can be used in fault detection and other applications. Compared to the results in Gillins and De Moor (2007), this paper takes into account the additional information on the inputs, and hence it results in a better estimate of state and input vectors. In addition, we show that Bayesian inference can provide an alternative derivation of the filter in Gillins and De Moor (2007) for the SSIE problem. We further investigate the relationships of the proposed filter with some existing approaches. In particular, we show in this paper that (a) when the inputs are completely available, the proposed filter reduces to the classical Kalman filter (Simon, 2006); (b) when no information on the unknown inputs is available, it reduces to the results of Gillins and De Moor (2007) where both state and input estimation are concerned; and (c) if only state estimation is of interest, it is equivalent to the filter for partially available inputs developed in Li (2013).

The rest of paper is structured as follows. Section 2 formulates the considered problem. The main results of the paper are provided in Section 3. In Section 4, a simulation study is carried out to illustrate the proposed filter. Finally, Section 5 concludes the paper.

2. Problem formulation

Consider a discrete-time linear stochastic time-varying system with unknown inputs:

$$\begin{aligned} x_{k+1} &= A_k x_k + G_k d_k + \omega_k, \\ y_k &= C_k x_k + \upsilon_k, \end{aligned} \tag{1}$$

where $x_k \in R^n, d_k \in R^m, y_k \in R^p$ are the state vector, input vector, and measurement vector at each time step k with $p \geq m$ and $n \geq m$. Following Li (2013), the process noise $\omega_k \in R^n$ and the measurement noise $\upsilon_k \in R^p$ are assumed to be mutually independent, and each follows a Gaussian distribution with zero mean and a known covariance matrix, $Q_k = E[\omega_k \omega_k^{\mathrm{T}}] > 0$ and $R_k = E[\upsilon_k \upsilon_k^{\mathrm{T}}] > 0$, respectively. A_k, G_k, C_k are known matrices. Following the existing researches (e.g. Gillins & De Moor, 2007; Kitanidis, 1987; Li, 2013; Su et al., 2015a), G_k is assumed to have a full column-rank; otherwise, the redundant input variables can be removed.

We consider the scenario where the input vector d_k is not fully observed at the level of interest but rather it is available only at an aggregate level. Specifically, let D_k be a $q_k \times m$ known matrix with $0 \leq q_k \leq m$ and F_{0k} an orthogonal complement of D_k^{T} such that $D_k F_{0k} = O_{q_k \times (m-q_k)}$. It is assumed that the input data are available only on some linear combinations:

$$r_k = D_k d_k, \tag{2}$$

where r_k is available at each time step k, whereas no information on $\delta_k = F_{0k}^{\mathrm{T}} d_k$ is available. Hence, δ_k is assumed to have a non-informative probability density function $f(\delta_k)$ such that all possible values of δ_k are equally likely to occur:

$$f(\delta_k) \propto 1. \tag{3}$$

Without loss of generality, we assume that D_k has a full row-rank; otherwise, the redundant rows of D_k can be removed from the analysis (see Su et al., 2015a).

As pointed out in Li (2013), the matrix D_k characterizes the availability of input information at each time step k. It includes two extreme scenarios that are usually considered: (a) $q_k = 0$, that is, no information on the input variables is available; this is the problem investigated in Kitanidis (1987) and Gillins and De Moor (2007); (b) $q_k = m$ and D_k is an identity matrix, that is, the complete input information is available. This is the case that the classical Kalman filter can be applied (Simon, 2006).

The objective of this paper is to simultaneously estimate the state and input vectors based on Equations (1) and (2).

3. Main results

In this section, the main results of the paper will be given. To incorporate the partial information on the inputs d_k, $G_k d_k$ is decoupled into two parts based on a decoupling matrix, that is, the known part given by the prior information (2) and unknown part δ_k. Note that δ_k has a lower dimension than the original input vector d_k and it plays the role of unknown inputs in the new system. Next, Bayesian inference is drawn to obtain recursive estimates of both state variables x_k and unknown inputs δ_k, upon which the estimate of the original input vector d_k can be worked out. Finally, the relationships between the proposed method and the relevant existing filters are discussed. The diagram of the system and the proposed filter structure is shown in Figure 1.

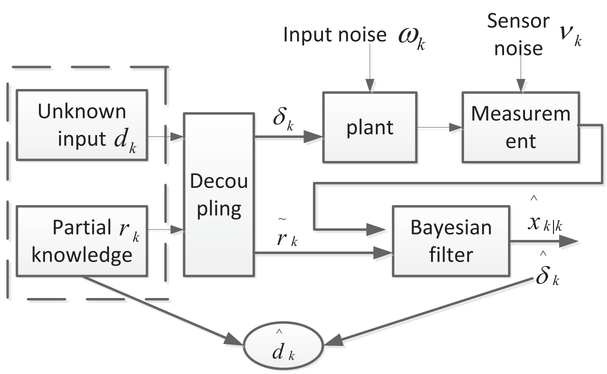

Figure 1. Diagram of the system and filter structure.

3.1. Transformation

To incorporate the information $r_k = D_k d_k$, a decoupling method is used here (see Su et al., 2015a). Define a non-singular decoupling matrix M_k of appropriate dimension as follows:

$$M_k = \begin{bmatrix} D_k & O \\ O & I \\ F_{0k}^{\mathrm{T}} & O \end{bmatrix} [G_k, \ \ G_k^{\perp}]^{-1},$$

where $G_k^{\perp}$ denotes an orthogonal complement of G_k, O and I represent the zero matrix and identity matrix of appropriate dimensions, respectively. F_{0k} is the orthogonal complement of D_k^{T} such that $D_k F_{0k} = O$ and $F_{0k}^{\mathrm{T}} F_{0k} = I$.

Then, $M_k G_k d_k$ can be expressed as follows:

$$\begin{aligned} M_k G_k d_k &= [D_k^{\mathrm{T}}, O, F_{0k}]^{\mathrm{T}} d_k \\ &= [(D_k d_k)^{\mathrm{T}}, O, O]^{\mathrm{T}} + [O, O, I]^{\mathrm{T}} F_{0k}^{\mathrm{T}} d_k \\ &= \tilde{r}_k + \tilde{G}_k \delta_k, \end{aligned} \tag{4}$$

where $\tilde{r}_k := [r_k^{\mathrm{T}} \ \ O \ \ O]^{\mathrm{T}}$ is completely available due to the available information on the inputs, $\tilde{G}_k := [O \ \ O \ \ I]^{\mathrm{T}}$ and $\delta_k := F_{0k}^{\mathrm{T}} d_k$.

Multiplying M_k^{-1} (the explicit form of M_k^{-1} is given in the appendix) on both sides of Equation (4), $G_k d_k$ can be decoupled into two parts:

$$G_k d_k = M_k^{-1} \tilde{r}_k + M_k^{-1} \tilde{G}_k \delta_k. \tag{5}$$

Consequently, the dynamics of x_{k+1} can be written as

$$\begin{aligned} x_{k+1} &= A_k x_k + M_k^{-1} \tilde{r}_k + M_k^{-1} \tilde{G}_k \delta_k + \omega_k \\ &= A_k x_k + M_k^{-1} \tilde{r}_k + F_k \delta_k + \omega_k, \end{aligned}$$

where $F_k := M_k^{-1} \tilde{G}_k = [G_k, \ \ G_k^{\perp}] \left[\begin{smallmatrix} F_{0k} \\ O \end{smallmatrix} \right] = G_k F_{0k}$.

Hence, the linear system (1) with the additional information on the inputs, $r_k = D_k d_k$, can equivalently be represented by the following system:

$$\begin{aligned} x_{k+1} &= A_k x_k + M_k^{-1} \tilde{r}_k + F_k \delta_k + \omega_k, \\ y_k &= C_k x_k + \upsilon_k. \end{aligned} \tag{6}$$

Remark 1 An alternative approach to incorporating the unknown input information is to use pseudo-inverse theory. From Equation (2), one can obtain the general solution of d_k

$$d_k = D_k^{+} r_k + F_{0k} \bar{\delta}_k, \tag{7}$$

where $D_k^{+} = D_k^{\mathrm{T}} (D_k D_k^{\mathrm{T}})^{-1}$ and $\bar{\delta}_k$ is completely unknown. If we select $\bar{\delta}_k := \delta_k = F_{0k}^{\mathrm{T}} d_k$, we can show that this approach is equivalent to the decoupling matrix based method.

Remark 2 It should be noted that the partial information on the inputs $r_k = D_k d_k$ has been fully incorporated into the system (6). We also note that the dimension of the inputs has been reduced from m to $m - q_k$.

3.2. Filter design

It can be seen from Equation (6) that y_k is a function of x_k, and x_k is related to the inputs δ_{k-1}. Hence, the input estimate of δ_k is delayed by one time unit (Gillins and De Moor, 2007). The objective of filter design is to obtain the estimate of x_k and δ_{k-1} based on the available measurement sequence $Y_k = \{y_1, y_2, \ldots, y_k\}$. For the new system (6), we can either solve the filtering problem based on the approach of MVUE (e.g. Gillins & De Moor, 2007) or Bayesian inference (e.g. Li, 2009, 2013). In the paper, we use the Bayesian method that can be seen as an alternative approach to that of Gillins and De Moor (2007).

In the context of Bayesian inference, the first step is to predict the dynamics of x_k and δ_{k-1} based on the available measurement sequence $Y_{k-1} = \{y_1, y_2, \ldots, y_{k-1}\}$. Since we do not assume that the unknown input vector δ_k satisfies any transition dynamics, prediction is only performed to determine the dynamics of x_k, that is, $p(x_k|Y_{k-1})$. The likelihood function can be determined based on the observation equation of system (6). The second step is to obtain the posterior distribution of the concerned variables after the measurement vector y_k is received based on Bayes' chain rule:

$$p(x_k, \delta_{k-1}|Y_k) \propto p(y_k|x_k) p(x_k, \delta_{k-1}|Y_{k-1}). \tag{8}$$

The main results on filtering design are summarized in Theorem 1.

THEOREM 1 *For state space model* (6), *suppose the matrix $C_k F_{k-1}$ has a full column-rank, then the prior and posterior distributions for x_k and δ_{k-1} at any time step k can be obtained sequentially as follows:*

(i) *Posterior of x_{k-1} for given Y_{k-1}* :

$$x_{k-1} \sim N(\hat{x}_{k-1|k-1}, P^{x}_{k-1|k-1}).$$

(ii) *Prediction for x_k :*

$$N(\hat{x}_{k|k-1}, P^x_{k|k-1}),$$

with $\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1|k-1} + M^{-1}_{k-1}\tilde{r}_{k-1}$,

$$P^x_{k|k-1} = A_{k-1}P^x_{k-1|k-1}A^{\mathrm{T}}_{k-1} + Q_{k-1}. \tag{9}$$

(iii) *Posterior of δ_{k-1} for given Y_k :*

$$\delta_{k-1} \sim N(\hat{\delta}_{k-1}, P^{\delta}_{k|k}),$$

where the posterior mean is given by

$$\hat{\delta}_{k-1} = P^{\delta}_{k|k}(C_kF_{k-1})^{\mathrm{T}}\tilde{R}^{-1}_k(y_k - C_k\hat{x}_{k|k-1}), \tag{10}$$

and the posterior covariance matrix is given by

$$P^{\delta}_{k|k} = (F^{\mathrm{T}}_{k-1}C^{\mathrm{T}}_k\tilde{R}^{-1}_kC_kF_{k-1})^{-1}, \tag{11}$$

while the posterior of x_k for given Y_k is

$$x_k \sim N(\hat{x}_{k|k}, P^x_{k|k}),$$

where the posterior mean is given by

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + P^x_{k|k-1}C^{\mathrm{T}}_k\tilde{R}^{-1}_k(y_k - C_k\hat{x}_{k|k-1}) \\ &\quad + (F_k - P^x_{k|k-1}C^{\mathrm{T}}_k\tilde{R}^{-1}_kC_kF_k)\hat{\delta}_{k-1},\end{aligned} \tag{12}$$

and the posterior covariance matrix is given by

$$\begin{aligned}P^x_{k|k} &= P^x_{k|k-1} - P^x_{k|k-1}C^{\mathrm{T}}_k\tilde{R}^{-1}_kC_kP^x_{k|k-1} \\ &\quad + (F_{k-1} - P^x_{k|k-1}C^{\mathrm{T}}_k\tilde{R}^{-1}_kC_kF_{k-1})(P^{\delta}_{k|k})^{-1}()^{\mathrm{T}},\end{aligned} \tag{13}$$

where $\tilde{R}_k = C_kP^x_{k|k-1}C^{\mathrm{T}}_k + R_k$, $()^{\mathrm{T}}$ *in* $(*)A()^{\mathrm{T}}$ *stands for the transpose of* $*$.

Proof From Equation (8), the posterior distribution $p(x_k, \delta_{k-1}|Y_k)$ is governed by

$$\begin{aligned}p(x_k, \delta_{k-1}|Y_k) \propto \exp\{&-(y_k - C_kx_k)^{\mathrm{T}}R^{-1}_k() \\ &- (x_k - \hat{x}_{k|k-1} - F_{k-1}\delta_{k-1})^{\mathrm{T}}(P^x_{k|k-1})^{-1}()\}.\end{aligned}$$

By completing the square on $[x^{\mathrm{T}}_k, \delta^{\mathrm{T}}_{k-1}]^{\mathrm{T}}$, the exponent can be rewritten as $-([x^{\mathrm{T}}_k, \delta^{\mathrm{T}}_{k-1}] - [\hat{x}^{\mathrm{T}}_{k|k}, \hat{\delta}^{\mathrm{T}}_{k-1}])P^{-1}_{k|k}()^{\mathrm{T}}$, where

$$\begin{bmatrix}\hat{x}_{k|k} \\ \hat{\delta}_{k-1}\end{bmatrix} = P_{k|k}\begin{bmatrix}C^{\mathrm{T}}_kR^{-1}_ky_k + (P^x_{k|k-1})^{-1}\hat{x}_{k|k-1} \\ -F^{\mathrm{T}}_{k-1}(P^x_{k|k-1})^{-1}x_{k|k-1}\end{bmatrix}$$

and

$$P_{k|k} = \begin{bmatrix}C^{\mathrm{T}}_kR^{-1}_kC_k + (P^x_{k|k-1})^{-1} & -(P^x_{k|k-1})^{-1}F_{k-1} \\ -F^{\mathrm{T}}_{k-1}(P^x_{k|k-1})^{-1} & F^{\mathrm{T}}_{k-1}(P^x_{k|k-1})^{-1}F_{k-1}\end{bmatrix}^{-1}.$$

This indicates that the posterior distribution is a Gaussian distribution with mean $[\hat{x}^{\mathrm{T}}_{k|k}, \hat{\delta}^{\mathrm{T}}_{k-1}]^{\mathrm{T}}$ and covariance matrix $P_{k|k}$. When C_kF_{k-1} is of full row-rank, based on the inverse of partitioned matrix, we can obtain the recursive estimation of both x_k and δ_{k-1} as shown in Equations (9)–(13).

So far, we have obtained the state estimate $\hat{x}_{k|k}$ and estimate $\hat{\delta}_{k-1}$ for the transformed system. When $F^{\mathrm{T}}_{0k-1}d_{k-1} = \hat{\delta}_{k-1}$ is obtained, based on Equation (5), we can further obtain the estimate of the original inputs d_{k-1} as follows:

$$\hat{d}_k = (G^{\mathrm{T}}_kG_k)^{-1}G^{\mathrm{T}}_k(M^{-1}_k\tilde{r}_k + M^{-1}_k\tilde{G}_k\hat{\delta}_k).$$

It can be verified that the obtained unknown input estimate satisfies the unknown input information (Equation (2)), that is,

$$D_k\hat{d}_k = r_k. \tag{14}$$

The proof is given in the appendix. ■

3.3. Relationships with the existing results

In this section, we investigate the relationships between the proposed approach and the relevant results in the existing literature. This is summarized in the following theorem.

THEOREM 2 *The set of recursive formulas* (9)–(13) *reduces to*

(1) *the classical Kalman filter when all entries of the input vector d_k are available;*
(2) *the filter in* Gillins and De Moor (2007) *when no information on the unknown inputs d_k is available;*
(3) *the filter in* Li (2013) *when only state estimation is concerned.*

Proof For the case where all the input variables are available at the level of interest, D_k becomes an $m \times m$ identity matrix, and F^{T}_{0k} becomes an zero-by-zero empty matrix. Consequently, the last term on the right-hand side of Equations (12) and (13) vanishes, and Equations (12) and (13) reduce to

$$P^x_{k|k} = P^x_{k|k-1} - P^x_{k|k-1}C^{\mathrm{T}}_kH^{-1}_kC_kP^x_{k|k-1}.$$

Since $M^{-1}_{k-1}\hat{r}_{k-1} = G_{k-1}d_{k-1}$, Equation (12) becomes

$$\begin{aligned}\hat{x}_{k|k} &= A_{k-1}\hat{x}_{k-1|k-1} + G_{k-1}d_{k-1} \\ &\quad + P^x_{k|k-1}C^{\mathrm{T}}_kH^{-1}_k(y_k - C_k(A_{k-1}\hat{x}_{k-1|k-1} \\ &\quad + G_{k-1}d_{k-1})).\end{aligned}$$

Clearly, these recursive formulas are identical to the classical Kalman filter equations (see Simon, 2006).

Next, we consider the case where no input information is available. Clearly $\tilde{r}_k$ in Equation (4) is an empty vector, F_k becomes G_k, and $\delta_k = d_k$. Hence, Equation (13) reduces to

$$P^x_{k|k} = P^x_{k|k-1} - P^x_{k|k-1} C_k^{\mathrm{T}} \tilde{R}_k^{-1} C_k P^x_{k|k-1} + [G_k - P^x_{k|k-1} C_k^{\mathrm{T}} H_k^{-1} C_k G_{k-1}] P^{\delta}_{k|k} []^{\mathrm{T}}$$

and the unknown input covariance matrix (11) becomes

$$P^{\delta}_{k|k} = (G_{k-1}^{\mathrm{T}} C_k^{\mathrm{T}} \tilde{R}_k^{-1} C_k G_{k-1})^{-1}.$$

In addition, Equation (12) becomes

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + P^x_{k|k-1} C_k^{\mathrm{T}} \tilde{R}_k^{-1} (y_k - C_k \hat{x}_{k|k-1}) + (G_k - P^x_{k|k-1} C_k^{\mathrm{T}} \tilde{R}_k^{-1} C_k G_k) \hat{\delta}_{k-1}$$

and the unknown input estimation Equation (10) becomes

$$\hat{\delta}_{k-1} = P^{\delta}_{k|k} (C_k G_{k-1})^{\mathrm{T}} \tilde{R}_k^{-1} (y_k - C_k \hat{x}_{k|k-1}).$$

These recursive formulas are identical to (a) the results in Kitanidis (1987) when only state filtering is of interest; and (b) the results in Gillins and De Moor (2007) for both unknown input and state estimations obtained using the approach of MVUE.

Finally, if only state estimation is concerned, the proposed method leads to the same results as those in Li (2013). To show this, we note that the state estimation error covariance matrix Equation (13) is the same as the one in Li (2013). In addition, inserting Equations (10) and (11) into Equation (12), Equation (12) can be rewritten in the following form:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (y_k - C_k \hat{x}_{k|k-1}),$$

where the gain matrix K_k is defined as

$$K_k = P^x_{k|k-1} C_k^{\mathrm{T}} \tilde{R}_k^{-1} + [F_{k-1} - P^x_{k|k-1} C_k^{\mathrm{T}} \tilde{R}_k^{-1} C_k F_{k-1}] (P^{\delta}_{k|k})^{-1} F_{k-1}^{\mathrm{T}} C_k^{\mathrm{T}} \tilde{R}_k^{-1}.$$

We can further show that (see the appendix for details)

$$M_{k-1}^{-1} \tilde{r}_{k-1} - K_k C_k M_{k-1}^{-1} \tilde{r}_{k-1} = P_{k|k} \bar{M}_{k-1}^{\mathrm{T}} (\bar{M}_{k-1} P_{k|k-1} \bar{M}_{k-1}^{\mathrm{T}})^{-1} \bar{r}_{k-1}, \qquad (15)$$

where the left-hand side of Equation (15) is the term associated with the prior information of the proposed filter, whereas the right-hand side of Equation (15) is the term associated with the prior information of the filter in Li (2013). This completes the proof. ■

4. Simulation study

In this section, we use a simple numerical example to illustrate the developed filter. First, we will show that, when only state estimation is of interest, the proposed filter can obtain the same result as that of Li (2013). Next we further demonstrate that incorporating the partially available information on the inputs can effectively improve both state estimation and unknown input estimation in comparison with the one without using the input information (Gillins & De Moor, 2007).

The system for the simulation is chosen the same as that of Su, Li, and Chen (2015b) that has been widely used in many previous studies (see, e.g. Cheng et al., 2009). However, to better assess the performance of the proposed filter under uncertainties, we considered a system subject to larger random variation: the covariance matrices Q_k and R_k of the system and measurement noises were taken 10 times as those of Cheng et al. (2009). The initial values of system model is chosen as $x_0 = [3,\ 1,\ 2,\ 2,\ 1]^{\mathrm{T}}$, the initial state and covariance matrix of filter are chosen as $\hat{x}_0 = 0_{5\times 1}$ and $P^x_{0|0} = 0.2 \times I_5$.

We applied the recursive formulas in this paper to estimate the state and unknown input vectors at each time step. To evaluate the quality of the state estimate and unknown input estimate obtained using the developed filter, we calculated the trace of the error covariance matrix $P^x_{k|k}$ and the trace of the error covariance matrix $P^{\delta}_{k|k}$ at each time step, as displayed in Figure 2(a) and Figure 2(b) (real line), respectively. For comparison, we also considered the state estimation using the filter in Li (2013) (only state estimation is concerned) and Gillins and De Moor (2007) (assuming that the inputs were completely unknown). The traces of $P^x_{k|k}$ are superimposed in Figure 2(a) (dotted line for Li, 2013 and dashed line for Gillins & De Moor, 2007), and the trace of $P^{\delta}_{k|k}$ is superimposed in Figure 2(b) (dashed line for Gillins & De Moor, 2007).

It can be seen from Figure 2(a) that the trace of state estimation error covariance using the proposed filter is the same as that of Li (2013). Both the method in Li (2013) and the proposed method have a smaller trace of the covariance matrix than that of Gillins and De Moor (2007).

In addition, Figure 2(b) shows that the trace of the error covariance matrix of the unknown input estimate using the proposed filter is smaller in comparison with that of Gillins and De Moor (2007). This is because more information on the unknown inputs was used by the filter developed in this paper. This demonstrates that when the unknown inputs are of practical interest, the proposed method in this paper will have a better performance than Gillins and De Moor (2007) if there is additional information available on the unknown inputs for filtering.

We also compared the state estimates obtained using the three filters, that is, the filter in Li (2013) (Figure 3), the proposed filter (Figure 4) and the filter in Gillins and De Moor (2007) (Figure 5). The upper graphs of Figures 3–5

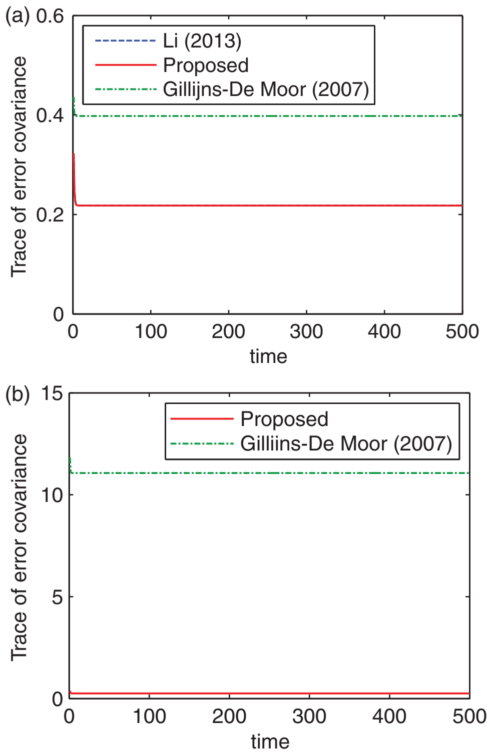

Figure 2. (a) Traces of the covariance matrix $P^{x}_{k|k}$ for three different filters; (b) traces of the covariance matrix $P^{\delta}_{k|k}$ for the proposed approach and the filter in Gillins and De Moor (2007).

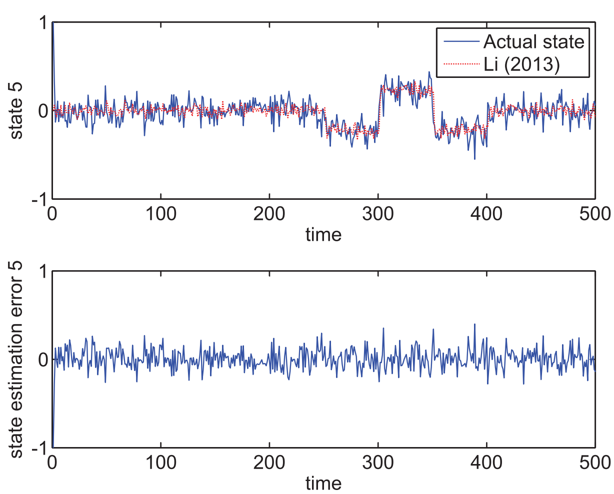

Figure 3. State estimation of the filter in Li (2013) and its estimation error.

display the simulated true values of the fifth state variable (real line) and the estimated state using the filters (dotted line), while the lower graphs plot the corresponding state estimation error for each filter.

It can be seen from Figures 3–5 that the three methods can provide a reasonably good estimate of the state vector. However, overall the state estimation errors using the

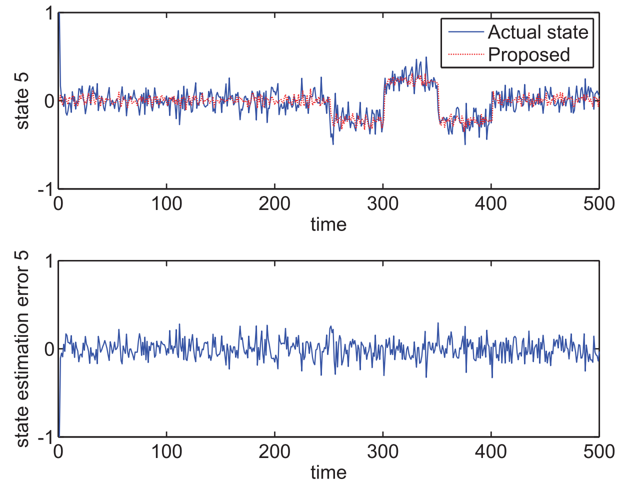

Figure 4. State estimation of the proposed filter and its estimation error.

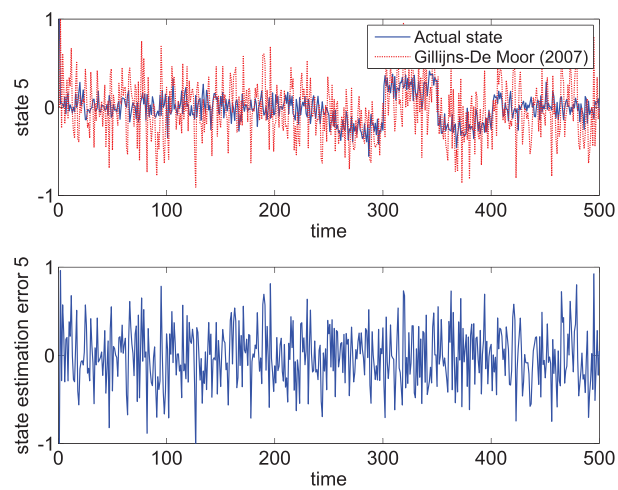

Figure 5. State estimation of the filter in Gillins and De Moor (2007) and its estimation error.

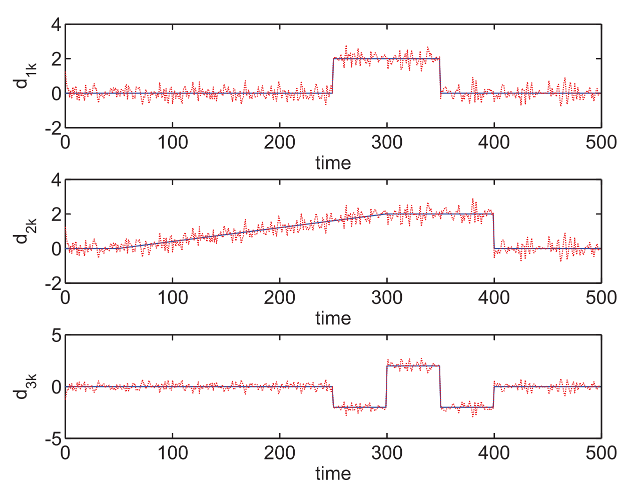

Figure 6. Unknown input estimation based on the proposed filter.

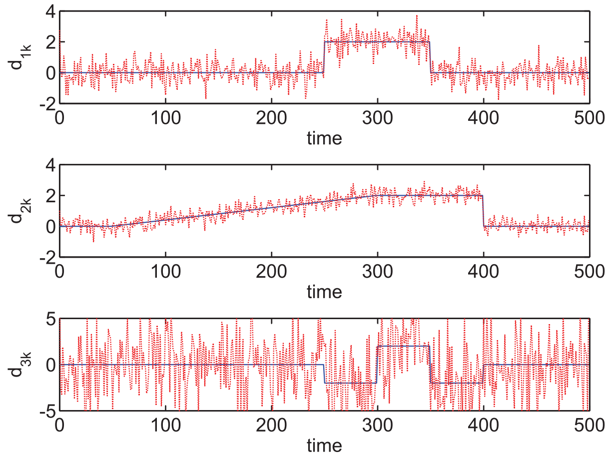

Figure 7. Unknown input estimation based on the filter in Gillins and De Moor (2007).

proposed filter and the filter in Li (2013) are smaller compared with that of Gillins and De Moor (2007) because the additional unknown input information was incorporated into the proposed filter and that of Li (2013).

Finally, we further compared our proposed method with the results in Gillins and De Moor (2007) for the purpose of unknown inputs estimation. The comparison results are shown in Figure 6 (the proposed method) and Figure 7 (the method in Gillins and De Moor (2007)), where real unknown inputs are depicted by real lines, and the unknown input estimations are depicted by the dotted lines.

We can see from Figures 6 and 7 that, by incorporating the information on the unknown inputs, the proposed method can obtain a much better performance for the unknown input estimation.

5. Conclusions

In this paper, the problem of SSIE has been investigated when partial information on the unknown inputs is available at an aggregate level. A decoupling approach is used to incorporate the unknown input information into the system dynamics. Then Bayesian inference is drawn to obtain the recursive state and input filter. The relationships of the proposed approach with the existing results are also discussed. Finally, the numerical example shows that, in comparison with the filter without using any input information, the proposed filter that makes use of the input information available at an aggregate level can substantially improve on the quality of both the state and input estimations. Future research can be done to extend the result to the case where there exists direct feedthrough of the partially observed inputs.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was jointly funded by UK Engineering and Physical Sciences Research Council (EPSRC) and BAE System (EP/H501401/1).

References

Chen, W., Ballance, D., Gawthrop, P., & O'Reilly, J. (2000). A nonlinear disturbance observer for robotic manipulators. *IEEE Transaction on Industrial Electronics*, *47*(4), 932–938.

Cheng, Y., Ye, H., Wang, Y. Q., & Zhou, D. H. (2009). Unbiased minimum-variance state estimation for linear systems with unknown input. *Automatica*, *45*(2), 485–491.

Darouach, M., & Zasadzinski, M. (1997). Unbiased minimum variance estimation for systems with unknown exogenous inputs. *Automatica*, *33*(4), 717–719.

Darouach, M., Zasadzinski, M., & Boutayeb, M. (2003). Extension of minimum variance estimation for systems with unknown inputs. *Automatica*, *39*(6), 867–876.

Fang, H.-Z., & Callafon, R. A. D. (2012). On the asymptotic stability of minimum-variance unbiased input and state estimation. *Automatica*, *48*(12), 3183–3186.

Gao, Z. W., & Ding, S. X. (2007). State and disturbance estimator for time-delay systems with application to fault estimation and signal compensation. *IEEE Transaction on Signal Processing*, *55*(12), 5541–5551.

Gillijns, S., & De Moor, B. (2007). Unbiased minimum-variance input and state estimation for linear discrete-time systems. *Automatica*, *43*(1), 111–116.

Hsieh, C. S. (2000). Robust two-stage Kalman filters for systems with unknown inputs. *IEEE Transaction on Automatic Control*, *45*(2), 2374–2378.

Jiang, B., & Fahmida, N. C. (2005). Fault estimation and accommodation for linear MIMO discrete-time systems. *IEEE Transaction on Control System Technology*, *13*(3), 493–499.

Kitanidis, P. K. (1987). Unbiased minimum-variance linear state estimation. *Automatica*, *23*(6), 775–778.

Li, B. (2009). A non-Gaussian Kalman filter with application to the estimation of vehicular speed. *Technometrics*, *51*(2), 162–172.

Li, B. (2013). State estimation with partially observed inputs: A unified Kalman filtering approach. *Automatica*, *49*(3), 816–820.

Mann, G., & Hwang, I. (2013). State estimation and fault detection and identification for constrained stochastic linear hybrid systems. *IET Control Theory and Application*, *7*(1), 1–15.

Patton, P., Clark, T., & Frank, P. M. (1989). *Fault diagnosis in dynamic systems: Theory and applications*. Upper Saddle River, NJ: Prentice-Hall International Series in Systems and Control Engineering, Prentice-Hall.

Profeta, III., Joseph, A., William, G. Vogt, & Marin, H. Mickle (1990). Disturbance estimation and compensation in linear systems. *IEEE Transaction on Aerospace and Electronics Systems*, *26*(2), 225–231.

Simon, D. (2006). *Optimal state estimation: Kalman, H_∞, and non-linear approaches*. New York: Wiley.

Su, J., Chen, W., & Li, B. (2014). Disturbance observer based fault diagnosis. *Proceedings of the 33rd Chinese control conference*, Nanjing, China, pp. 3024–3029.

Su, J., Li, B., & W.-H. Chen (2015a). On existence, optimality and asymptotic stability of the Kalman filter with partially observed inputs. *Automatica*, *53*, 149–154.

Su, J., Li, B., & W.-H. Chen (2015b). Recursive filter with partial knowledge on inputs and outputs. *International Journal of Automation and Computing*, *12*(1), 35–42.

Yang, J., Li, S., Su, J., & Yu, X. (2013). Continuous nonsingular terminal sliding mode control for systems with mismatched disturbances. *Automatica*, *49*(7), 2287–2291.

Yang, J., Su, J., Li, S., & Yu, X. (2014). High-order mismatched disturbance compensation for motion control systems via a continuous dynamic sliding-mode approach. *IEEE Transactions on Industrial Informatics*, *10*(1), 604–614.

Appendix

A.1. Proof of Equation (14)

First, we can obtain the inverse of M_k as follows:

$$M_k^{-1} = [G_k,\ \ G_k^{\perp}]\begin{bmatrix}(I - F_{0k}F_{0k}^{\mathrm{T}})D_k^{\mathrm{T}}(D_kD_k^{\mathrm{T}})^{-1} & O & F_{0k}\\ O & I & O\end{bmatrix}.$$

Then, Equation (14) can be obtained as follows:

$$\begin{aligned}D_k\hat{d}_k &= D_k(G_k^{\mathrm{T}}G_k)^{-1}G_k^{\mathrm{T}}M_k^{-1}[r_k\ \ O\ \ O]^{\mathrm{T}}\\ &\quad + D_k(G_k^{\mathrm{T}}G_k)^{-1}G_k^{\mathrm{T}}M_k^{-1}[O\ \ O\ \ I]^{\mathrm{T}}\hat{\delta}_k\\ &= D_k(G_k^{\mathrm{T}}G_k)^{-1}G_k^{\mathrm{T}}G_k(I - F_{0k}F_{0k}^{\mathrm{T}})D_k^{\mathrm{T}}(D_kD_k^{\mathrm{T}})^{-1}r_k\\ &\quad + D_k(G_k^{\mathrm{T}}G_k)^{-1}G_k^{\mathrm{T}}G_kF_{0k}\hat{\delta}_k\\ &= r_k.\end{aligned}$$

A.2. Proof of Equation (15)

Define $M_{k-1}^P = P_{k|k}\bar{M}_{k-1}^{\mathrm{T}}(\bar{M}_{k-1}P_{k|k-1}\bar{M}_{k-1}^{\mathrm{T}})^{-1}$. Then, we have

$$\begin{aligned}&M_{k-1}^{-1}\tilde{r}_{k-1} - K_kC_kM_{k-1}^{-1}\tilde{r}_{k-1}\\ &= (I - K_kC_k)M_{k-1}^{-1}\tilde{r}_{k-1}\\ &= M_{k-1}^P\bar{M}_{k-1}M_{k-1}^{-1}\tilde{r}_{k-1}\\ &= M_{k-1}^P\bar{M}_{k-1}G_{k-1}D_{k-1}^{\mathrm{T}}(D_{k-1}D_{k-1}^{\mathrm{T}})^{-1}D_{k-1}d_{k-1}\\ &= M_{k-1}^P\begin{bmatrix}D_{k-1} & O\\ O & I\end{bmatrix}[G_{k-1},\ \ G_{k-1}^{\perp}]^{-1}\\ &\quad \times G_{k-1}D_{k-1}^{\mathrm{T}}(D_{k-1}D_{k-1}^{\mathrm{T}})^{-1}D_{k-1}d_{k-1}\\ &= M_{k-1}^P\begin{bmatrix}r_{k-1}\\ O\end{bmatrix} = M_{k-1}^P\bar{M}_{k-1}G_{k-1}d_{k-1}\\ &= P_{k|k}\bar{M}_{k-1}^{\mathrm{T}}(\bar{M}_{k-1}P_{k|k-1}\bar{M}_{k-1}^{\mathrm{T}})^{-1}\bar{r}_{k-1},\end{aligned}$$

where in the above derivation, we have used the following identities:

$$M_{k-1}^{-1}\tilde{r}_{k-1} = G_{k-1}D_{k-1}^{\mathrm{T}}(D_{k-1}D_{k-1}^{\mathrm{T}})^{-1}D_{k-1}d_{k-1}, \tag{A1}$$

$$I - K_kC_k = M_{k-1}^P\bar{M}_{k-1}. \tag{A2}$$

Now we show Equation (A2):

$$\begin{aligned}&I - K_kC_k - M_{k-1}^p\bar{M}_{k-1}\\ &= I - K_kC_k - P_{k|k}\bar{M}_{k-1}^{\mathrm{T}}(\bar{M}_{k-1}P_{k|k-1}\bar{M}_{k-1}^{\mathrm{T}})^{-1}\bar{M}_{k-1}\\ &= I - P_{k|k}C_k^{\mathrm{T}}R_k^{-1}C_k - P_{k|k}[\bar{M}_{k-1}^{\mathrm{T}}(\bar{M}_{k-1}P_{k|k-1}\bar{M}_{k-1}^{\mathrm{T}})^{-1}\\ &\quad \times \bar{M}_{k-1} + C_k^{\mathrm{T}}R_k^{-1}C_k - C_k^{\mathrm{T}}R_k^{-1}C_k]\\ &= I - P_{k|k}C_k^{\mathrm{T}}R_k^{-1}C_k - [I - P_{k|k}C_k^{\mathrm{T}}R_k^{-1}C_k]\\ &= O,\end{aligned}$$

where $\bar{M}_k = \begin{bmatrix}D_k & O\\ O & I\end{bmatrix}[G_k,\ \ G_k^{\perp}]^{-1}$.