Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Receding Horizon Estimation of Arbitrarily Changing Unknown Inputs ⋆

Ruxandra Mustata * **Michel Verhaegen** * **Henrik Ohlsson** ** **Fredrik Gustafsson** **

* *Delft Center for Systems and Control, Delft University of Technology, Delft, 2628 CD, The Netherlands (e-mails: r.i.mustata@tudelft.nl, m.verhaegen@tudelft.nl).*
** *Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sewden (e-mails: ohlsson@isy.liu.se, fredrik@isy.liu.se)*

Abstract: In this paper we present a receding horizon estimation method for linear time invariant systems, subject to unknown inputs. The proposed approach is based on the idea of asymptotically decoupling the state estimation problem from the unknown input estimation problem. Consequently, the latter is formulated as a weighted least squares problem in a receding horizon manner. The proposed method does not assume a dynamic model for the unknown input, but it allows to incorporate prior knowledge about its abrupt nature by adding ℓ_1-regularization terms to the cost function of the weighted least squares problem. The receding horizon input estimation method and the necessary conditions for it to hold are outlined. The proposed method is illustrated in simulation for the case of abruptly changing piece-wise constant unknown inputs.

Keywords: Unknown input observer, fault estimation, receding horizon estimation, regularized least squares.

1. INTRODUCTION

This paper considers the estimation of the unknown input of a linear state space model. Such a problem has a wide interest and a long research history. In the area of fault detection, the unknown input may represent faults that occur due to failing components or changing external load conditions (Willsky and Jones, 1974; Patton et al., 1989; Basseville and Nikiforov, 1993; Gustafsson, 2001). In target tracking, the unknown input may represent unknown forces exerted on the moving object.

A first class of methods aim for jointly estimating the state and the unknown input. Examples are Hsieh (2000), Gillijns and De Moor (2007). These solutions are Riccati based, making them valid for arbitrary changing unknown inputs, but they do not allow the important specialization towards abrupt changing inputs.

When the unknown inputs change abruptly, e.g. as jumps, a pioneering contribution was made by Willsky and Jones (1974). The philosophy of this approach was to build a secondary moving horizon estimator on top of a Kalman filter, tuned for the situation when the unknown input is equal to zero. The major drawback of this approach is that the receding horizon estimation requires that no unknown input has occurred prior to the time window w on which the receding horizon estimation problem is defined. Such drawback restricts the practical usefulness of the approach. The drawback is a consequence of the inability to jointly estimate both the state and the unknown (abrupt) input sequence within the time window w.

A recent attempt to simultaneously estimate the state sequence and the unknown abrupt input was presented by Ohlsson et al. (2012). This paper considers a fixed time interval in the context of sum-of-norms (ℓ_1) regularization. The regularization is necessary since the system is underdetermined and, as a result, the state and the unknown input can not be determined uniquely. Moreover, the ℓ_1 regularization term penalizes the absolute values of the variables, thus inducing a bias towards zero in the solution.

In order to overcome the drawback of both existing pioneering contributions to deal with abrupt changing unknown inputs and hence come up with a procedure that is valid for both abrupt and arbritrary changing unknown inputs, a new receding horizon input estimation method is presented in this paper. The method is initially presented using the same philosophy as in Willsky and Jones (1974), on top of an unknown-input free Kalman filter. The method enables to “approximatively” decouple the estimation of the unknown input sequence over the time window w from that of the estimation of the state sequence in that same window.

The main advantage of the decoupling is that the unknown input estimation can be defined both for the case of arbitrary changing unknown inputs as well as for constrained unknown inputs. That is, the constraints on the input are not essential in order to be able to estimate the unknown input, as was necessary in Ohlsson et al. (2012).

⋆ This research is partially supported by the Dutch Ministry of Economic Affairs in the frame of the Smart Optics Systems STW program.

Copyright © 2014 IFAC

The paper is organized as follows. In Section 2 we present the problem of unknown input estimation in a receding horizon framework, given an LTI model in innovation form. Based on the idea of approximate decoupling between the unknown input estimation problem from the state reconstruction problem, in Section 3 we outline the receding horizon input (RHI) estimation method and the necessary conditions for it to hold. The RHI estimation method is extended to constrained unknown inputs in Section 4. Finally, the RHI estimation is illustrated in simulation in Section 5 and concluding remarks are drawn in Section 6.

2. PROBLEM FORMULATION

Starting from the same philosophy as used by Willsky and Jones (1974), we initially formulate the RHI estimation problem for the LTI system given in innovation from:

$$\begin{aligned} x(k+1) &= Ax(k) + Bb(k) + Ke(k), \\ y(k) &= Cx(k) + e(k), \end{aligned} \tag{1}$$

where $x(k) \in \mathbb{R}^{n_x}$ is the state vector, $b(k) \in \mathbb{R}^{n_b}$ is the unknown input vector and $y(k) \in \mathbb{R}^{n_y}$ is the output vector at time k. The innovation $e(k)$ is a zero-mean white noise sequence with a given covariance matrix Σ_e. Control input and feed-through terms, i.e. $Bu(k)$ and $Du(k)$, can be added to the state and output equations in (1), but for the sake of compactness they are left out in the exposure of this paper. We show in Section 5 that a deterministic control signal can be included in the RHI estimation.

If we denote by $\Phi = A - KC$, then we can write the dynamic equation of the model (1) as:

$$x(k+1) = \Phi x(k) + Bb(k) + Ky(k), \tag{2}$$

and separate the unknown-input free Kalman filter:

$$\begin{aligned} \hat{x}(k+1) &= \Phi \hat{x}(k) + Ky(k), \\ \hat{y}(k) &= C\hat{x}(k). \end{aligned} \tag{3}$$

Let us define the error in the state vector $x_e(k) = x(k) - \hat{x}(k)$ and the output vector $r(k) = y(k) - \hat{y}(k)$. Then the following dynamic model results:

$$\begin{aligned} x_e(k+1) &= \Phi x_e(k) + Bb(k), \\ r(k) &= Cx_e(k) + e(k). \end{aligned} \tag{4}$$

Over the time window $[k-L+1, k]$, the output residual $r(k)$ defined in (4) can be written as:

$$\begin{aligned} \mathbf{r}_{k,L} = \underbrace{\begin{bmatrix} C\Phi \\ \vdots \\ C\Phi^L \end{bmatrix}}_{\mathcal{O}_L} x_e(k-L) + \\ \underbrace{\begin{bmatrix} CB & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C\Phi^{L-1}B & C\Phi^{L-2}B & \cdots & CB \end{bmatrix}}_{\mathcal{T}_L} \mathbf{b}_{k-1,L} + \mathbf{e}_{k,L}, \end{aligned} \tag{5}$$

and denoted compactly as:

$$\mathbf{r}_{k,L} = \mathcal{O}_L x_e(k-L) + \mathcal{T}_L \mathbf{b}_{k-1,L} + \mathbf{e}_{k,L}, \tag{6}$$

with

$$\begin{aligned} \mathbf{r}_{k,L} &= \left[r(k-L+1)^T \ \cdots \ r(k)^T \right], \\ \mathbf{b}_{k-1,L} &= \left[b(k-L)^T \ \cdots \ b(k-1)^T \right], \\ \mathbf{e}_{k,L} &= \left[e(k-L+1)^T \ \cdots \ e(k)^T \right]. \end{aligned}$$

This equation clearly highlights the fundamental problem in estimating the unknowns $x_e(k-L)$ and $\mathbf{b}_{k-1,L}$, i.e., when the matrix $[\mathcal{O}_L \ \mathcal{T}_L]$ is singular, the system of equations (6) is under-determined, the unknowns $x_e(k-L)$ and $\mathbf{b}_{k-1,L}$ can not be uniquely determined. Various restrictions have been imposed to circumvent this problem. An example is the parity space analysis (PSA), which cancels the observability matrix $\mathcal{O}_L$ by multiplying (5) from the left by its orthogonal complement (Patton et al., 1989). Willsky and Jones (1974) put $x_e(k-L)$ to zero, corresponding to the hypothesis that no unknown input occurred prior to the moving window. The work of Ohlsson et al. (2012) solves for $x_e(k-L)$ and $\mathbf{b}_{k-1,L}$ simultaneously, thereby requiring the unknown input sequence $\mathbf{b}_{k-1,L}$ to be "sparse".

In this paper we build on the recent work of Dong and Verhaegen (2012) that splits the time window $[k-L+1, k]$ into a *past* time window of length p and a *future* time window of length f, with $p+f = L$. With these notations, we partition the output residual term $\mathbf{r}_{k,L}$, the unknown input term $\mathbf{b}_{k-1,L}$ and the noise term $\mathbf{e}_{k,L}$ as follows:

$$\mathbf{r}_{k,L} = \begin{bmatrix} \mathbf{r}_{k-f,p} \\ \mathbf{r}_{k,f} \end{bmatrix}, \ \mathbf{b}_{k-1,L} = \begin{bmatrix} \mathbf{b}_{k-f-1,p} \\ \mathbf{b}_{k-1,f} \end{bmatrix}, \ \mathbf{e}_{k,L} = \begin{bmatrix} \mathbf{e}_{k-f,p} \\ \mathbf{e}_{k,f} \end{bmatrix}.$$

In the sequel, we will use the following matrix partitions:

$$\mathcal{O}_L = \begin{bmatrix} \mathcal{O}_p \\ \mathcal{O}_f \Phi^p \end{bmatrix}, \quad \mathcal{T}_L = \begin{bmatrix} \mathcal{T}_p & 0 \\ \mathcal{H}_{f,p} & \mathcal{T}_f \end{bmatrix}, \tag{7}$$

where $\mathcal{O}_p \in \mathbb{R}^{pn_y \times n_x}$ and $\mathcal{O}_f \in \mathbb{R}^{fn_y \times n_x}$ are defined similarly to $\mathcal{O}_L$ in (5), $\mathcal{T}_p \in \mathbb{R}^{pn_y \times pn_b}$ and $\mathcal{T}_f \in \mathbb{R}^{fn_y \times fn_b}$ are defined similarly to $\mathcal{T}_L$ in (5), and $\mathcal{H}_{f,p} \in \mathbb{R}^{fn_y \times pn_b}$ is a block Toeplitz matrix of the form:

$$\mathcal{H}_{f,p} = \begin{bmatrix} C\Phi^p B & \cdots & C\Phi B \\ \vdots & \ddots & \vdots \\ C\Phi^{L-1}B & \cdots & C\Phi^f B \end{bmatrix}. \tag{8}$$

Applying this partitioning to (5), we get:

$$\begin{aligned} \begin{bmatrix} \mathbf{r}_{k-f,p} \\ \mathbf{r}_{k,f} \end{bmatrix} &= \begin{bmatrix} \mathcal{O}_p \\ \mathcal{O}_f \Phi^p \end{bmatrix} x_e(k-L) \\ &+ \begin{bmatrix} \mathcal{T}_p & 0 \\ \mathcal{H}_{f,p} & \mathcal{T}_f \end{bmatrix} \begin{bmatrix} \mathbf{b}_{k-f-1,p} \\ \mathbf{b}_{k-1,f} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_{k-f,p} \\ \mathbf{e}_{k,f} \end{bmatrix}. \end{aligned} \tag{9}$$

In this paper we address the following problem. Here, use is made of the notation $M^\dagger$ to represent the (left) pseudo-inverse of a matrix M that satisfies $M^\dagger M = I$.

Problem 1. Assume that: the system matrix Φ is asymptotically stable; the matrix CB is full column rank, such that the matrices $\mathcal{T}_p$ and $\mathcal{T}_f$ are full column rank and the left pseudo-inverses $(CB)^\dagger$ and $\mathcal{T}_p^\dagger$ exist; the system matrix $(I - B(CB)^\dagger C)\Phi$ of the inverse system from $r(k)$ to $b(k-1)$ is asymptotically stable.
Then, for p and f given, with $p+f = L$, transform the output residual equation (9) into,

$$\begin{bmatrix} \star \\ \bar{\mathbf{r}}_{k,f} \end{bmatrix} = \begin{bmatrix} \star \\ \Gamma_f \end{bmatrix} x_e(k-L) + \begin{bmatrix} \mathcal{T}_p & 0 \\ 0 & \mathcal{T}_f \end{bmatrix} \begin{bmatrix} \mathbf{b}_{k-f-1,p} \\ \mathbf{b}_{k-1,f} \end{bmatrix} + \begin{bmatrix} \star \\ \bar{\mathbf{e}}_{k,f} \end{bmatrix} \tag{10}$$

with $\star$ representing quantities of no direct relevance and $\Gamma_f x_e(k-L)$ vanishes in the limit for $p \to \infty$ for a bounded $x_e(k-L)$. Furthermore, find an estimate of the unknown input in the future time window, denoted by $\hat{\mathbf{b}}_{k-1,f}$, based on the transformed data equation (10).

$$\mathcal{T}_p^{\dagger} = \begin{bmatrix} (CB)^{\dagger} & 0 & 0 & \dots & 0 \\ -(CB)^{\dagger}C\Phi B(CB)^{\dagger} & (CB)^{\dagger} & 0 & \dots & 0 \\ -(CB)^{\dagger}C\Phi\Psi B(CB)^{\dagger} & -(CB)^{\dagger}C\Phi B(CB)^{\dagger} & (CB)^{\dagger} & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ -(CB)^{\dagger}C\Phi\Psi^{p-2}B(CB)^{\dagger} & -(CB)^{\dagger}C\Phi\Psi^{p-3}B(CB)^{\dagger} & & \dots & (CB)^{\dagger} \end{bmatrix} \tag{11}$$

3. RECEDING HORIZON INPUT ESTIMATION

We start by solving the problem of approximately decoupling the unknown input estimation problem from the state reconstruction problem. The decoupling consists in removing, in an asymptotic (for $p \to \infty$) manner, the influence of the unknown input over the past time window, $\mathbf{b}_{k-f-1,p}$, and that of the initial state, $x_e(k-L)$, from the output residuals of the future time window, $\mathbf{r}_{k,f}$.

Following the work in of Dong and Verhaegen (2012), it is noted that, since the matrix Φ is asymptotically stable, it can be assumed that the product $\Phi^p x_e(k-L)$ can be made arbitrary small by an appropriate selection of the past horizon length p. The approximate decoupling requires finding a transformation matrix that, upon multiplying the partitioned data equation (9) to the left, block diagonalizes the matrix $\mathcal{T}_L$ as in (10), and, in addition, yields a transformed matrix Γ_f which asymptotically vanishes as $p \to \infty$. Let us consider the (left) transformation matrix that block diagonalizes $\mathcal{T}_L$ given as:

$$T = \begin{bmatrix} I & 0 \\ -\mathcal{H}_{f,p}\mathcal{T}_p^{\dagger} & I \end{bmatrix}. \tag{12}$$

Applying the transformation matrix T to the partitioned data equation (9) on the left leads to the modified data equation (10). This results into the transformed output residual and the transformed noise sequence, respectively,

$$\bar{\mathbf{r}}_{k,f} = \begin{bmatrix} -\mathcal{H}_{f,p}\mathcal{T}_p^{\dagger} & I \end{bmatrix} \mathbf{r}_{k,L} = \mathbf{r}_{k,f} - \mathcal{H}_{f,p}\mathcal{T}_p^{\dagger}\mathbf{r}_{k-f,p}, \tag{13}$$

$$\bar{\mathbf{e}}_{k,f} = \begin{bmatrix} -\mathcal{H}_{f,p}\mathcal{T}_p^{\dagger} & I \end{bmatrix} \mathbf{e}_{k,L} = \mathbf{e}_{k,f} - \mathcal{H}_{f,p}\mathcal{T}_p^{\dagger}\mathbf{e}_{k-f,p}. \tag{14}$$

Moreover, it defines the matrix Γ_f as:

$$\Gamma_f = \begin{bmatrix} -\mathcal{H}_{f,p}\mathcal{T}_p^{\dagger} & I \end{bmatrix} \mathcal{O}_L = \mathcal{O}_f\Phi^p - \mathcal{H}_{f,p}\mathcal{T}_p^{\dagger}\mathcal{O}_p. \tag{15}$$

With these notations, we resume here only the bottom block line of the modified data equation (10):

$$\bar{\mathbf{r}}_{k,f} = \Gamma_f x_e(k-L) + \mathcal{T}_f \mathbf{b}_{k-1,f} + \bar{\mathbf{e}}_{k,f}. \tag{16}$$

Now we can show that the transformation (12) yields the desired asymptotic results in the following theorem.

Theorem 2. Let the system matrix Φ of the observer (3) and the system matrix Ψ of the inverse of the system (4), which represents the the transfer from $r(k)$ to $b(k-1)$, i.e.,

$$\Psi = (I - B(CB)^{\dagger}C)\Phi, \tag{17}$$

be asymptotically stable. Then, the matrix Γ_f defined in (15) can be written as:

$$\Gamma_f = \mathcal{O}_f\Psi^p. \tag{18}$$

Moreover, for any bounded vector $x(k)$, the matrix Γ_f satisfies:

$$\lim_{p\to\infty} \Gamma_f x(k) = 0. \tag{19}$$

Proof. Denote by $\mathcal{C}_p = \begin{bmatrix} \Phi^{p-1}B & \dots & \Phi B & B \end{bmatrix}$. Then, for $\mathcal{H}_{f,p}$ given by (8), it holds that: $\mathcal{H}_{f,p} = \mathcal{O}_f\mathcal{C}_p$. With this notation and given the matrix Γ_f as in (15), it holds that:

$$\Gamma_f = \mathcal{O}_f\left(\Phi^p - \mathcal{C}_p\mathcal{T}_p^{\dagger}\mathcal{O}_p\right). \tag{20}$$

We shall now compute the term $\mathcal{C}_p\mathcal{T}_p^{\dagger}\mathcal{O}_p$ in (20). Given the matrix $\mathcal{T}_p^{\dagger}$ as in (11), it follows that:

$$\mathcal{T}_p^{\dagger}\mathcal{O}_p = \begin{bmatrix} (CB)^{\dagger}C\Phi \\ (CB)^{\dagger}C\Phi\Psi \\ \vdots \\ (CB)^{\dagger}C\Phi\Psi^{p-1} \end{bmatrix}.$$

Furthermore, when left multiplying $\mathcal{T}_p^{\dagger}\mathcal{O}_p$ by $\mathcal{C}_p$, one gets:

$$\begin{aligned} \mathcal{C}_p\mathcal{T}_p^{\dagger}\mathcal{O}_p &= \sum_{j=0}^{p-1} \Phi^{p-j-1}B(CB)^{\dagger}C\Phi\Psi^j, \\ &= \sum_{j=0}^{p-1} \Phi^{p-j-1}(\Phi - \Psi)\Psi^j. \end{aligned} \tag{21}$$

The second inequality holds due to the substitution $B(CB)^{\dagger}C\Phi = \Phi - \Psi$ given by (17). The terms of the sum in (21) cancel each other out, leaving that:

$$\mathcal{C}_p\mathcal{T}_p^{\dagger}\mathcal{O}_p = \Phi^p - \Psi^p,$$

which, when replaced in (20), leads to (18). Based on (18) and on the asymptotic stability of the matrix Ψ, the limit (19) holds. This completes the proof. ■

Remark 3. Note that the annihilation of the influence of the past unknown input, $\mathbf{b}_{k-f-1,p}$, and that of the initial state, $x_e(k-L)$, in the output residual of the future time window, $\mathbf{r}_{k,f}$, is only achieved asymptotically, i.e., for $p \to \infty$. This result holds for a finite value of the future window size f. As such, the asymptotic decoupling is a generalization of the work of Dong and Verhaegen (2012), where the asymptotic annihilation of the term related to $x_e(k-L)$ from the residual $\mathbf{r}_{k,f}$ required both $p, f \to \infty$. Finite values of p lead to an approximate decoupling and, consequently, to a small bias in the unknown input estimation. This bias can be analysed and treated with the methodology introduced by Dong et al. (2012). For that reason and for the sake of compactness, we restrict this paper in the sequel to the case when $\Gamma_f x_e(k-L)$ in (16) is equal to zero. □

To complete the analysis, one needs to have knowledge on the stochastic properties of the noise sequence $\bar{\mathbf{e}}_{k,f}$. Let the covariance matrices of the temporally uncorrelated noise sequences $\mathbf{e}_{k,f}$ and $\mathbf{e}_{k-f,p}$ be:

$$\begin{aligned} \Sigma_{\mathbf{e}_{k,f}} &= \mathbf{E}[\mathbf{e}_{k,f}\mathbf{e}_{k,f}^T] &&= I_f \otimes \Sigma_e, \\ \Sigma_{\mathbf{e}_{k-f,p}} &= \mathbf{E}[\mathbf{e}_{k-f,p}\mathbf{e}_{k-f,p}^T] &&= I_p \otimes \Sigma_e, \end{aligned}$$

where $\otimes$ is the Kronecker product and $\Sigma_e = \mathbf{E}[e(k)e(k)^T]$ is the covariance matrix of the noise sequence $e(k)$. Then, the covariance matrix of the noise sequence $\bar{\mathbf{e}}_{k,f}$ will be:

$$\Sigma_{\bar{\mathbf{e}}_{k,f}} = \Sigma_{\mathbf{e}_{k,f}} + \mathcal{H}_{f,p}\mathcal{T}_p^{\dagger}\Sigma_{\mathbf{e}_{k-f,p}}(\mathcal{T}_p^{\dagger})^T\mathcal{H}_{f,p}^T. \tag{22}$$

Moreover, let $\Sigma_{\bar{\mathbf{e}}_{k,f}}$ be factorized into $\Sigma_{\bar{\mathbf{e}}_{k,f}} = \Sigma_{\bar{\mathbf{e}}_{k,f}}^{\frac{1}{2}}\Sigma_{\bar{\mathbf{e}}_{k,f}}^{\frac{T}{2}}$.

A weighted least squares problem can now be formulated for estimating the unknown input $\mathbf{b}_{k-1,f}$ from the transformed data equation (16), after neglecting the term $\Gamma_f x_e(k-L)$:

$$\min_{\mathbf{b}_{k-1,f}} \|\Sigma_{\bar{\mathbf{e}}_{k,f}}^{-\frac{1}{2}}(\bar{\mathbf{r}}_{k,f} - \mathcal{T}_f \mathbf{b}_{k-1,f})\|_2^2. \qquad (23)$$

If the matrix $\mathcal{T}_f$ has full column rank, the minimum variance solution to this problem is:

$$\hat{\mathbf{b}}_{k-1,f} = \left[\mathcal{T}_f^T \Sigma_{\bar{\mathbf{e}}_{k,f}}^{-1} \mathcal{T}_f\right]^{-1} \mathcal{T}_f^T \Sigma_{\bar{\mathbf{e}}_{k,f}}^{-1} \cdot \bar{\mathbf{r}}_{k,f}, \qquad (24)$$

and the covariance matrix of this solution is given by:

$$\Sigma_{\hat{\mathbf{b}}_{k-1,f}} = \left[\mathcal{T}_f^T \Sigma_{\bar{\mathbf{e}}_{k,f}}^{-1} \mathcal{T}_f\right]^{-1}. \qquad (25)$$

Finally, the unknown input estimate $\hat{b}(k-1)$ can be determined from the last n_b elements of $\hat{\mathbf{b}}_{k-1,f}$, while the covariance matrix of $\hat{b}(k-1)$ can be determined from the last $n_b \times n_b$ diagonal blocks of $\Sigma_{\hat{\mathbf{b}}_{k-1,f}}$.

Remark 4. When $n_b \leq n_y$, the least squares problem (23) can deal with arbitrary unknown inputs. Furthermore, if prior information is available on the temporal nature of the bias, such as it being piecewise constant or impulsive, the problem can be extended to a constrained least squares problem, as outlined in Section 4. Such flexibility is not present in the existing family of joint state and unknown input estimation algorithms, such as in, e.g., Gillijns and De Moor (2007). The constraints are, on the other hand, a necessary requirement in the sum-of-norms solution given by Ohlsson et al. (2012), which can also be reformulated in a receding horizon framework. □

4. ℓ_1-CONSTRAINED RHI CALCULATION.

The RHI estimation problem (23) allows for adding constraints or regularization terms to consider the case of abrupt changing unknown inputs. As in the work Ohlsson et al. (2010), this can improve the estimation when the prior information about the unknown input is correct. A general formulation of the regularized least squares problem is:

$$\min_{\mathbf{b}_{k-1,f}} \|\Sigma_{\bar{\mathbf{e}}_{k,f}}^{-\frac{1}{2}}(\bar{\mathbf{r}}_{k,f} - \mathcal{T}_f \mathbf{b}_{k-1,f})\|_2^2 + \lambda \|\mathcal{P}\mathbf{b}_{k-1,f}\|_1, \qquad (26)$$

where λ is a positive constant used to make a trade-off between the data fitting (the first term) and the temporal variation of the unknown input sequence within the time window $[k-f, k-1]$ (the second term). The matrix $\mathcal{P} \in \mathbb{R}^{fn_b \times fn_b}$ will introduce a pattern in the temporal variation of the unknown input, based on a priori knowledge.

For impulse type of inputs, where the input is equal to zero most of the times and different from zero at only a small number of times, it is convenient to consider $\mathcal{P} = I_{fn_b}$. This will result into a regularization term which favours that "many" of the unknown input components come out as exactly zero in the solution.

For piece-wise constant type of inputs, where the input is constant at most of the times and jumping to a new value at only a small number of time instants, it is convenient to consider $\mathcal{P} = I_f \otimes P$, with:

$$P = \begin{bmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{n_b \times n_b}.$$

This choice for the matrix $\mathcal{P}$ favours that "many" of the unknown input components remain constant from one time instant to the next in the time window $[k-f, k-1]$.

Remark 5. The ℓ_1-constraints tend to bias the estimates towards zero. However, the fact that the RHI estimation problem can be solved without the need of regularization allows us, in our future research, to use the constrained problem (26) in order to find the sparsity pattern of $\hat{\mathbf{b}}_{k-1,f}$ and, afterwards, to re-estimate the non-zero elements by using the unregularized criterion (23) over only the non-zero elements. This has the potential to remove the bias while preserving the sparsity pattern in the estimates. □

5. SIMULATION RESULTS

In the simulations, we demonstrate the application of the new RHI estimation method for generic systems with deterministic inputs, process and measurement noise. For the sake of compactness, the procedure of designing an observer gain K such that the matrices Φ and Ψ are stabilized simultaneously is not given here.

For the evaluation of the RHI estimation, we consider a standard benchmark given by the linearized VTOL (vertical take-off and landing) aircraft model:

$$\dot{x}(t) = \begin{bmatrix} -0.0366 & 0.0271 & 0.0188 & -0.4555 \\ 0.0482 & -1.01 & 0.0024 & -4.0208 \\ 0.1002 & 0.3681 & -0.707 & 1.42 \\ 0 & 0 & 1 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0.4422 & 3.5446 & -5.52 & 0 \\ 0.1761 & -7.5922 & 4.49 & 0 \end{bmatrix}^T (u(t) + b(t)), \qquad (27)$$

$$y(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} x(t). \qquad (28)$$

We reproduce the same conditions in terms of discretization time, noise characteristics, and unknown inputs as used by Dong and Verhaegen (2012). The continuous time model (27)-(28) is discretized with a sampling time equal to 0.5 s. Furthermore, since the open loop plant is unstable, a stabilizing output feedback controller is used, such that:

$$u(k) = -\begin{bmatrix} 0 & 0 & -0.5 & 0 \\ 0 & 0 & -0.1 & -0.1 \end{bmatrix}^T y(k). \qquad (29)$$

A fault in the actuators has been added. More specifically, a constant jump has been introduced at the second input after time instant $k = 500$, such that:

$$b(k) = \begin{cases} [0\ 0] & \text{for } k \leq 500, \\ [0\ 0.5] & \text{for } 500 < k \leq 1000. \end{cases} \qquad (30)$$

Process and measurement noise have been considered, denoted by $w(k)$ and $v(k)$, respectively, with covariance matrices $Q = E[w(k)w(k)^T] = 10^{-4} I_4$ and $R = E[v(k)v(k)^T] = 10^{-2} I_4$.

The input $u(k)$ is measurable and can be included in the observer form (3), such that its influence is subtracted from

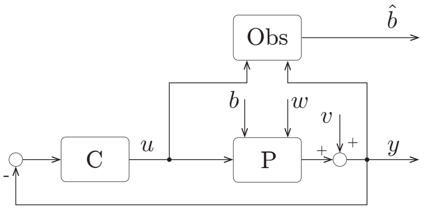


Fig. 1. Block scheme of the observer integration in closed loop: the plant "P", the stabilizing controller "C" and the unknown input observer "Obs".

the measurements $y(k)$, as described in Section 2, leaving the output residual $r(k)$ only dependent on the unknown input $b(k)$. Thus, the RHI estimation can be then applied to a dynamic model of type (4).

The system has been simulated in closed loop, as shown in Figure 1. The measurement and the control signals, $y(k)$ and $u(k)$, have been recorded, for $k = 1, \ldots, 1000$, and have been used for the estimation of the unknown input signal $b(k)$, as if the system was simulated in open loop.

5.1 Results for the unconstrained RHI case

It has been shown in Section 3 how the influence of the unknown inputs of the past window, $\mathbf{b}_{k-f-1,p}$, can be removed asymptotically from the output residuals of the future window, $\mathbf{r}_{k,f}$, via the transformation matrix (12). This particular feature is the main contribution of the unconstrained RHI estimation method with respect to the approach proposed by Dong and Verhaegen (2012), where the following constrained least squares problem is considered:

$$\min_{\mathbf{b}_{k-1,f}} \quad \left\| \Sigma_{\mathbf{v}_{k,f}}^{-\frac{1}{2}} \left(\mathbf{r}_{k,f} - [\mathcal{H}_{f,p} \;\, \mathcal{T}_f] \begin{bmatrix} \mathbf{b}_{k-f-1,p} \\ \mathbf{b}_{k-1,f} \end{bmatrix} \right) \right\|_2^2 \qquad (31)$$
$$\text{s.t.} \qquad \mathbf{b}_{k-f-1,p} = 0,$$

which has the following solution:

$$\hat{\mathbf{b}}_{k-1,f} = \left[\mathcal{T}_f^T \Sigma_{\mathbf{v}_{k,f}}^{-1} \mathcal{T}_f \right]^{-1} \mathcal{T}_f^T \Sigma_{\mathbf{v}_{k,f}}^{-1} \cdot \mathbf{r}_{k,f}. \qquad (32)$$

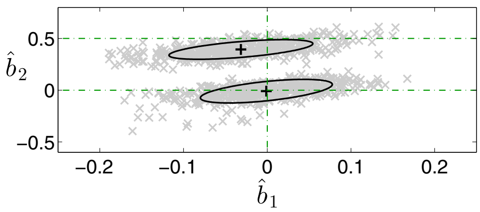
where $\Sigma_{\mathbf{v}_{k,f}} = I_f \otimes \Sigma_\nu$ and $\Sigma_\nu = diag(Q, R)$.

Here we show that the bias in the unknown input estimates is significantly reduced when solving the unconstrained RHI estimation problem (23) with respect to solving the constrained least squares problem (31).

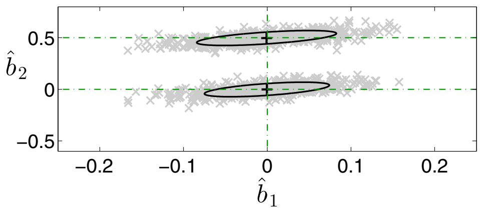
For the simulation of both methods we have used past and future time windows of lengths $p = 15$ and $f = 5$. The estimates $\hat{b}_1(k)$ and $\hat{b}_2(k)$, obtained using both methods for 1000 time instants, are plotted in Figure 2a-2b. The estimates obtained using the method in Dong and Verhaegen (2012) show a bias for the time interval $500 \leq k \leq 1000$, after the jump occurs in the second input.

5.2 Results for the ℓ_1-constrained RHI case

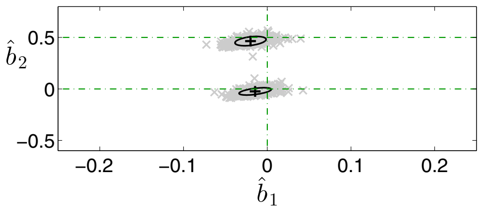
In this section we show that, when prior information about the unknown input class is included in the estimation problem formulation, the ℓ_1-constrained RHI estimation presented in Section 4 yields reduced variance estimates when compared to the unconstrained RHI estimation. For the sake of brevity, we only show here the results obtained for estimating piece-wise constant inputs, such as (30).



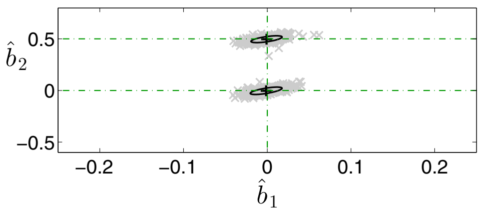
(a) Dong and Verhaegen (2012)



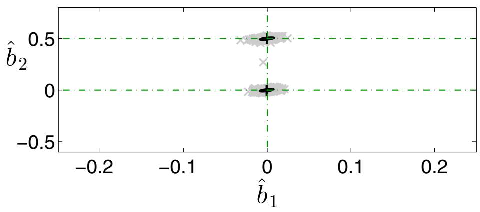
(b) Unconstrained RHI ($p = 15,\ f = 5$)



(c) StateSON ($L = 20$)



(d) ℓ_1-Constrained RHI ($p = 15,\ f = 5$)



(e) ℓ_1-Constrained RHI ($p = 15,\ f = 15$)

Fig. 2. Two-dimensional unknown input estimates for 1000 time instants with 1σ-ellipsoid joint distribution.

We compare the ℓ_1-constrained RHI estimation method (26) with the StateSON method proposed by Ohlsson et al. (2012) reformulated in a receding horizon framework:

$$\min_{\substack{x(k-L+1),\, b(t),\, w(t) \\ k-L+1 \leq t \leq k}} \sum_{t=k-L+1}^{k} \left\| \Sigma_\nu^{-\frac{1}{2}} \left(y(t) - Cx(t) \right) \right\|_2^2 + \lambda \sum_{t=k-L+1}^{k-1} \left\| b(t) - b(t-1) \right\|_1 \qquad (33)$$
$$\text{s.t.} \quad x(t+1) = Ax(t) + B\left[u(t) + b(t) \right] + w(t).$$

The StateSON method is recursively performing the optimization over a time horizon of length $L = 20$. For the constrained RHI estimation, we have used a past time

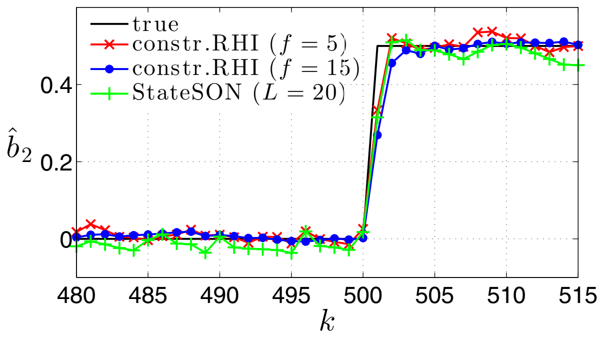


Fig. 3. Temporal evolution of the estimate $\hat{b}_2$ around the time instant corresponding to the jump.

window length $p = 15$ and we have increased the future time window length from $f = 5$ to $f = 15$. The weighting parameters λ have been empirically fixed to 20 for both methods. The estimates obtained using both methods are shown in Figure 2c-2e.

As the future time window length f increases, the variance of the ℓ_1-constrained RHI estimates (Figures 2d-2e) decreases with respect to that of the the unconstrained RHI estimates (Figure 2b). On the other hand, for small values of the window length L, the StateSON estimation method shows a bias in the estimates (Figure 2c). This is due to the fact that the problem of simultaneously estimating the state and the unknown input sequence is singular and adding the ℓ_1-regularization term yields a unique solution, but can not correct for the bias introduced by the singularity. This is not an issue for the RHI estimation method, where the unknown input estimation problem is asymptotically decoupled from the state estimation problem, yielding a well-posed problem with unique solution for the unknown input estimate.

A temporal evolution of the estimate $\hat{b}_2(k)$ around the time instant corresponding to the jump is shown in Figure 3. Note that, around that time instant, the estimates of both the constrained RHI and of StateSON methods are biased. This is due to the presence of the ℓ_1 regularization term in the constrained problems (26) and (33).

5.3 Computational complexity

We report in Table 1 the computational time necessary to compute one iteration for each of the algorithms introduced. For the simulations, we have implemented the algorithms in the MATLAB environment, on a computer with a 1.7 GHz processor and 4 GB RAM. The unconstrained RHI and the unknown input observer proposed by Dong and Verhaegen (2012) are computed based on the closed form solutions of the least squares problems (24) and (32). The ℓ_1-constrained RHI problem (26) and the StateSON problem (33) are solved using the CVX package (Grant and Boyd, 2013).

6. CONCLUDING REMARKS

In this paper we propose a receding horizon input estimation method for linear time invariant systems. The proposed approach asymptotically decouples the state estimation problem from the input estimation problem. The latter is consequently formulated as a weighted least squares problem in a receding horizon manner. The proposed method does not assume a dynamic model for the unknown input, but it allows for incorporating prior knowledge about the dynamics of the unknown input via regularization terms.

Method	Computational time (s)
Dong and Verhaegen (2012)	$6 \cdot 10^{-3}$
Unconstrained RHI ($p = 15$, $f = 5$)	$7.9 \cdot 10^{-3}$
StateSON ($L = 20$)	0.6850
Constrained RHI ($p = 15$, $f = 5$)	0.3832
Constrained RHI ($p = 15$, $f = 15$)	0.4372

Table 1. Computational time necessary to compute one iteration.

The asymptotic decoupling generalizes the Dong and Verhaegen (2012) approach by removing the influence of the past inputs on the output residuals of the receding horizon time window. Moreover, it generalizes the Ohlsson et al. (2012) approach in the sense that it removes the singularity in the least squares problem. A comparison between the three approaches is presented via simulation results.

REFERENCES

Basseville, M. and Nikiforov, I.V. (1993). Detection of abrupt changes: Theory and application.

Dong, J. and Verhaegen, M. (2012). Identification of fault estimation filter from I/O data for systems with stable inversion. *Automatic Control, IEEE Transactions on*, 57(6), 1347 –1361.

Dong, J., Verhaegen, M., and Gustafsson, F. (2012). Robust fault detection with statistical uncertainty in identified parameters. *IEEE Transactions on Signal Processing*, 60(10), 5064–5076.

Gillijns, S. and De Moor, B. (2007). Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough. *Automatica*, 43, 934–937.

Grant, M. and Boyd, S. (2013). CVX: Matlab software for disciplined convex programming, version 2.0 beta. `http://cvxr.com/cvx`.

Gustafsson, F. (2001). *Adaptive filtering and change detection*. Wiley, New York.

Hsieh, C.S. (2000). Robust two-stage Kalman filters for systems with unknown inputs. *Automatic Control, IEEE Transactions on*, 45(12), 2374–2378.

Ohlsson, H., Gustafsson, F., Ljung, L., and Boyd, S.P. (2010). State smoothing by sum-of-norms regularization. In *Proceedings of the 49th IEEE Conference on Decision and Control*, 2880–2885.

Ohlsson, H., Gustafsson, F., Ljung, L., and Boyd, S.P. (2012). Smoothed state estimates under abrupt changes using sum-of-norms regularization. *Automatica*, 48(4), 595–605.

Patton, R.J., Frank, P.M., and Clarke, R.N. (eds.) (1989). *Fault diagnosis in dynamic systems: theory and application*. Prentice-Hall, Inc.

Willsky, A. and Jones, H. (1974). A generalized likelihood ratio approach to state estimation in linear systems subjects to abrupt changes. In *Decision and Control including the 13th Symposium on Adaptive Processes, 1974 IEEE Conference on*, volume 13, 846–853.