15th Triennial World Congress, Barcelona, Spain

STATE ESTIMATION FOR SYSTEMS HAVING RANDOM MEASUREMENT DELAYS USING ERRORS IN VARIABLES

Graham C. Goodwin * **Juan Carlos Agüero** * **Arie Feuer** **

* *The University of Newcastle, Australia 2308*
** *Technion, Israel 32000*

Abstract: We address the problem of state estimation in linear time invariant systems when the measurements are subject to unknown random delays. In cases where the measurements are "time stamped" the delays can be computed on-line. In such cases, the estimation problem reduces to a standard Kalman Filtering problem. Here we will study the more challenging case when the measurements are not time stamped. We show that the latter case can be formulated as an errors in variables problem.

Keywords: delay compensation, state estimation, communication networks.

1. INTRODUCTION

The standard formulation of the linear state estimation problem assumes that the measurements arrive having a known fixed delay. This case has an elegant solution via the Kalman filter (Anderson and Moore, 1979). Here, we will consider the more challenging case where the measurements suffer an unknown random delay. We assume throughout that synchronous sampling is used and that data arriving between sampling intervals is buffered until the next clock pulse. This allows us to use a discrete formulation. Under these conditions it is straightforward to extend the standard Kalman filter to the case when the data experiences a random delay provided the data is "time stamped". Essentially all one needs to do to cover this case is to set up a delay line which covers all possible delays (assuming, of course, that there exists a known upper bound on the delay). When processing a given measurement on arrival, one can read the time-stamp and assign the appropriate entry to the **C** in the delay chain read-out. To explain this idea more fully, say that the known upper bound on the delay is L sampling periods. The system model is taken to be

$$x[k+1] = Ax[k] + Bu[k] + w[k] \tag{1}$$

$$y[k] = Cx[k-d] + \upsilon[k] \tag{2}$$

where $u[k]$, $w[k]$, $\upsilon[k]$ denote the known system input, process noise and measurement noise respectively. The process and measurement noise have zero mean and covariance matrices Γ_v, and Γ_w respectively. In the sequel we take $u[k] = 0$ for simplicity in the presentation. Comments are made regarding the case $u[k] \neq 0$ later. To account for variable delay, d, we write the model as:

$$\begin{aligned} z[k+1] &= \mathbf{A}z[k] + e[k] \\ y[k] &= \mathbf{C}[k]z[k] + v[k] \end{aligned} \tag{3}$$

$$\mathbf{A} = \begin{bmatrix} A & 0 & \cdots & 0 \\ I & 0 & \cdots & 0 \\ 0 & \ddots & \cdots & 0 \\ 0 & \cdots & I & 0 \end{bmatrix} \in \mathbb{R}^{n\times n}; \tag{4}$$

$$e[k] = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} w[k] = \mathbf{G}w[k] \tag{5}$$

$$\mathbf{C}[k] = C\left[\, \alpha_0[k]\, I \;\cdots\; \alpha_L[k]\, I \,\right] \in \mathbb{R}^{1\times n} \tag{6}$$

and $n = n_o(L+1)$ where n_o is dimension of the original state $x[k]$. Notice that we set $\alpha_i[k] = 0$ for $i = 0, \ldots, L$ save for $\alpha_d[k] = 1$ if we know that the sample $y[k]$ has experienced a 'd' sample delay. The latter information can be readily determined from time stamped data. Note that the state estimation problem is then a standard (time-varying) Kalman filtering problem (Anderson and Moore, 1979).

Here, we will examine the case when the data is not time-stamped and can experience nondeterministic delays. One potential application of these ideas is to process control where the “sampling system” for some process variables requires the use of off-line analyzers which can lead to random delays in the measurements. Another potential application is to Networked Control Systems where the data is transmitted over a communication network. Due to the asynchronous time-division multiplexing of common network protocols, time varying and possibly stochastic delays are introduced in the control system. (See (Branicky and Zhang, 2000; Chan and Őzgűner, 1995; Hong, 1995; Kaplan, 2001; Krtolica and Liubakka, 1990; Luck and Ray, 1990; Nilsson and Wittenmark, 1998; Ray, 1987; Walsh and Bushnell, 1999; Bushnell, 2001; Nair and Evans, 2000)). Our formulation is relevant to the case of random delays between sensors and controller.

There are many ways that one could formulate the estimation problem where data is subject to random delays. For example, if the delays themselves satisfy a discrete time Markov model, then one could set this up as a Hidden Markov Model estimation problem (Krtolica and Liubakka, 1990). The approach that we adopt will be based on “Errors in Variables” or “Total Least Squares” (Van Huffel and Vandewalle, 1991; De Moor, 1993; Golub and Loan, 1980). It is known, that such methods do not readily lend themselves to recursive solutions. We will thus adopt a “receding horizon” approach (Muske K. R. and Lee, 1993; Michalska and D.Q.Mayne, 1995; Rao and Lee, 2001) in which the data is processed in blocks. Thus, if we consider the situation at time k, we have accumulated data in the time interval $[k-N, k]$ and this data block will be used to generate the estimate we need. The reason for using a fixed block length for the errors in variables analysis is to constrain the “size” of the associated computational problem.

2. REPARAMETERIZATION OF THE PROBLEM

Equation (3) represents a set of constraints for our estimation problem. This constraint can be written in matrix form as follows (for simplicity, in this paper, we only treat the scalar observation case):

$$Y = M\theta_1 + \eta \tag{7}$$

where

$$Y = \begin{bmatrix} y[0] \\ 0 \\ y[1] \\ 0 \\ y[2] \\ \vdots \\ y[N] \\ 0 \end{bmatrix} \in \mathbb{R}^{(n_o+1)(N+1)} \tag{8}$$

$$M = \begin{bmatrix} \mathbf{C}[0] & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -\mathbf{G}^T\mathbf{A} & \mathbf{G}^T & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \mathbf{C}[1] & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -\mathbf{G}^T\mathbf{A} & \mathbf{G}^T & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \mathbf{C}[2] & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \mathbf{C}[N] & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & -\mathbf{G}^T\mathbf{A} & \mathbf{G}^T \end{bmatrix}$$

$$\in \mathbb{R}^{(n_o+1)(N+1)\times n(N+2)} \tag{9}$$

$$\theta_1 = \begin{bmatrix} z[0] \\ \vdots \\ z[N] \\ z[N+1] \end{bmatrix} \in \mathbb{R}^{n(N+2)} \tag{10}$$

$$\eta = \begin{bmatrix} v[0] \\ w[0] \\ v[1] \\ w[1] \\ v[2] \\ \vdots \\ v[N] \\ w[N] \end{bmatrix} \tag{11}$$

where η has zero mean and known covariance Γ (we assume for simplicity that both $w[k]$ and $v[k]$ are stationary i.i.d sequences).

Remark 1. In the above definition of θ_1 there is substantial redundancy. In fact, we can write

$$\theta_1 = L_1\theta_1^o \tag{12}$$

where

$$\theta_1^o = \begin{bmatrix} x[-L] \\ \vdots \\ x[0] \\ x[1] \\ \vdots \\ x[N+1] \end{bmatrix} \in \mathbb{R}^{n_o(N+L+2)}$$

$$L_1 = \sum_{k=0}^{N+1}\sum_{l=0}^{L}\sum_{m=1}^{n_o} e^{n(N+2)}_{kn+ln_o+m}\left(e^{n_o(N+L+2)}_{(L-l)n_o+kn_o+m}\right)^T \tag{13}$$

and e^i_j is the j^{th} column of an $i \times i$ identity matrix.

When the sequence $\{\mathbf{C}[k]\}$ is known, i.e., when the matrix M is known, then the Best Linear Unbiased Estimate (BLUE) of θ_1^o is

$$\hat{\theta}_1^o = \left[L_1^T M^T \Gamma^{-1} M L_1\right]^{-1}\left[L_1^T M^T \Gamma^{-1} Y\right] \tag{14}$$

Indeed, it can be readily shown that this estimate corresponds precisely to the Kalman filter solution (De Moor, 1993). Our interest here, however, focuses on the case where M is not known exactly. In this case, the elements of M are themselves random variables with some known mean value (which we take to be a “pseudo measurement”) and some known covariance structure. In this case, we see that the estimation problem implicit in equation (7) is in the form of an Errors in Variables problem (Van Huffel and Vandewalle, 1991), since both Y and M are measured with errors. The idea of using Errors in Variables methods for Kalman filtering problems where the model (i.e., $[C, A]$) contains errors has been previously discussed in (De Moor, 1993) where a first-order state space system was analyzed. Novel aspects of our problem are the physical motivation via delayed measurements and the implications this has on the special structure of the system matrices and the error covariance structure associated with the Total Least Square problem. Indeed, this leads to the class of problems referred to as Structured Total Least Squares (STLS). These key aspects will be further developed in the next section.

3. THE STLS PROBLEM

We next recast our estimation problem as a STLS problem. To do that, let

$$\theta = \begin{bmatrix} -1 \\ \theta_1 \end{bmatrix} \in \mathbb{R}^{n(N+2)+1}$$
$$\Phi = \begin{bmatrix} Y & M \end{bmatrix} \tag{15}$$

Since, as seen in equations (6) and (9), M (hence Φ) depends on values of $\{\alpha[k]\}_{i=0}^{L}$ which are not available, we will instead, use "pseudo measurements", based on their mean values $\{\overline{\alpha}_i = E\{\alpha_i\}\}$. The calculation of these means, as well as the covariance matrix of $\{\alpha[k]\}_{i=0}^{L}$ is described in the Appendix.

We are ready now to formulate the optimization problem as that of minimizing

$$J = \sum_{k=0}^{N} \Gamma_v^{-1} \left(y[k] - \widehat{y}[k]\right)^2 + \left(\underline{\overline{\alpha}} - \underline{\widehat{\alpha}}[k]\right)^T \Gamma_\alpha^{-1} \left(\underline{\overline{\alpha}} - \underline{\widehat{\alpha}}[k]\right) + \widehat{w}[k]^T \Gamma_w^{-1} \widehat{w}[k] \tag{16}$$

with respect to $\{\widehat{y}[k], \underline{\widehat{\alpha}}[k], \widehat{w}[k]\}_{k=0}^{N}$ and $\widehat{\theta}$, subject to the constraints

$$\widehat{\Phi}\widehat{\theta} = 0 \tag{17}$$
$$\left(e_1^{n(N+2)+1}\right)^T \widehat{\theta} + 1 = 0 \tag{18}$$

Note that (18) constrains the first entry of $\widehat{\theta}$ to be nonzero. We also impose other structural constraints implicit in the way $\widehat{\Phi}$ depends on $\{\widehat{y}[k], \underline{\widehat{\alpha}}[k], \widehat{w}[k]\}_{k=0}^{N}$. To make this dependence concrete, let us define the following set of matrices:

$$B_y[k] = e_{k(n_o+1)+1}^{(n_o+1)(N+1)} \left(e_1^{n(N+2)+1}\right)^T \tag{19}$$

$$B_w^i[k] = e_{k(n_o+1)+1+i}^{(n_o+1)(N+1)} \left(e_1^{n(N+2)+1}\right)^T \tag{20}$$

$$B_\alpha^l[k] = e_{k(n_o+1)+1}^{(n_o+1)(N+1)} C \sum_{r=1}^{n_o} e_r^{n_o} \left(e_{1+kn+r+n_o l}^{n(N+2)+1}\right)^T \tag{21}$$

and

$$\Phi_o = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & -\mathbf{G}^T\mathbf{A} & \mathbf{G}^T & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & -\mathbf{G}^T\mathbf{A} & \mathbf{G}^T & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & -\mathbf{G}^T\mathbf{A} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & -\mathbf{G}^T\mathbf{A} & \mathbf{G}^T \end{bmatrix} \tag{22}$$

It can then be shown using (5)-(9), (15) and (19)-(21) that:

$$\Phi = \Phi_o + \sum_{k=0}^{N} \left[y[k] B_y[k] + \sum_{l=0}^{L} \overline{\alpha}_l[k] B_\alpha^l[k] \right] \tag{23}$$

The structural constraint we impose on $\widehat{\Phi}$ is expressed as:

$$\widehat{\Phi} = \Phi_o + \sum_{k=0}^{N} \left[\widehat{y}[k] B_y[k] + \sum_{i=1}^{n_o} \left(\widehat{w}[k]\right)_i B_w^i[k] + \sum_{l=0}^{L} \widehat{\alpha}_l[k] B_\alpha^l[k] \right] \tag{24}$$

In summary, our optimization problem is described by equations (16)-(18) and (24). We note that the structural constrained does not fully utilize the prior information. Specifically, the fact that $\underline{\alpha}[k]$, for every k, has at most one entry equal to one and all other entries are zero is not used. An attempt to impose this prior as well, would lead to a mixed integer-continuous, optimization problem which would be computationally formidable. Instead, we relax the problem by treating $\underline{\alpha}[k]$ as an i.i.d. sequence with mean $\underline{\overline{\alpha}}$ and covariance Γ_α. This will give a superior result to that obtained via a traditional Kalman filter with the choice $\underline{\alpha}[k]$ equal to the nominal value. However, it will be inferior to incorporating the full structure of $\underline{\alpha}[k]$.

4. SOLUTION OF THE STLS PROBLEM

Different methods have been proposed in the literature to solve the **STLS** optimization problem outlined above: Structured Total Least Norm (**STLN**), Riemannian Singular Value Decomposition (**RiSVD**), Constrained Total Least Squares (**CTLS**). The **STLN** approach solves the problem as a nonlinear constrained optimization problem (Van Huffel and Lemmerlimg, 2002). The other two approaches first transform the problem into an equivalent form prior to solving it. The **RiSVD** transforms the nonlinear constrained optimization problem into an algebraic non-linear problem called The Riemannian singular value problem. The **CTLS** approach transforms the problem into a nonlinear unconstrained problem. The latter approach is the one used in this paper. The details are given in the following theorem:

Theorem 1. The non-linear constrained optimization problem (16)-(18) and (24) can be transformed into the following non-linear unconstrained optimization problem:

$$\min \begin{bmatrix} -1 \\ \widehat{\theta_1^o} \end{bmatrix}^T L_2^T \Phi^T D_\theta^{-1} \left(L_2 \begin{bmatrix} -1 \\ \widehat{\theta_1^o} \end{bmatrix} \right) \Phi L_2 \begin{bmatrix} -1 \\ \widehat{\theta_1^o} \end{bmatrix} \tag{25}$$

Proof. The proof of this theorem follows the same general line as that presented in (Lemmerling, 1999).

Using Lagrange multipliers Λ and λ we introduce the following cost:

$$\mathcal{L} = J + 2\Lambda^T \widehat{\Phi}\widehat{\theta} + 2\lambda \left(\left(e_1^{n(N+2)+1} \right)^T \widehat{\theta} + 1 \right)$$

where J is as in (16). The necessary conditions for a constrained minimum of J are then:

$$\frac{\partial \mathcal{L}}{\partial \widehat{y}[k]} = 0 \Rightarrow \widehat{y}[k] = y[k] - \Gamma_v \Lambda^T B_y[k]\widehat{\theta} \tag{26}$$

$$\frac{\partial \mathcal{L}}{\partial \widehat{\alpha}_l[k]} = 0 \Rightarrow \underline{\widehat{\alpha}}[k] = \underline{\bar{\alpha}}[k] - \Gamma_\alpha \begin{bmatrix} \Lambda^T B_\alpha^0[k]\widehat{\theta} \\ \Lambda^T B_\alpha^1[k]\widehat{\theta} \\ \vdots \\ \Lambda^T B_\alpha^L[k]\widehat{\theta} \end{bmatrix} \tag{27}$$

$$\frac{\partial \mathcal{L}}{\partial (\widehat{w}[k])_i} = 0 \Rightarrow \widehat{w}[k] = -\Gamma_w \begin{bmatrix} \Lambda^T B_w^1[k]\widehat{\theta} \\ \Lambda^T B_w^2[k]\widehat{\theta} \\ \vdots \\ \Lambda^T B_w^{n_0}[k]\widehat{\theta} \end{bmatrix} \tag{28}$$

$$\frac{\partial \mathcal{L}}{\partial \widehat{\theta}} = 0 \Rightarrow \widehat{\Phi}^T \Lambda + \lambda e_1^{(N+2)n+1} = 0 \tag{29}$$

together with equations (17) and (18). Premultiplying (29) by $\widehat{\theta}^T$ and using equations (17) and (18) we immediately conclude that

$$\lambda = 0 \tag{30}$$

From (17), (24), and (26)-(28) we have

$$\widehat{\Phi}\widehat{\theta} = \Bigg(\Phi_o + \sum_{k=0}^{N} \Bigg[\left(y[k] - \Gamma_v \Lambda^T B_y[k]\widehat{\theta} \right) B_y[k] - \sum_{i=1}^{n_0} (e_i^n)^T \Gamma_w \begin{bmatrix} \Lambda^T B_w^1[k]\widehat{\theta} \\ \Lambda^T B_w^2[k]\widehat{\theta} \\ \vdots \\ \Lambda^T B_w^{n_0}[k]\widehat{\theta} \end{bmatrix} B_w^i[k] + \sum_{l=0}^{L} \Bigg(\bar{\alpha}_l[k\} - \left(e_{l+1}^{L+1} \right)^T \Gamma_\alpha \begin{bmatrix} \Lambda^T B_\alpha^0[k]\widehat{\theta} \\ \Lambda^T B_\alpha^1[k]\widehat{\theta} \\ \vdots \\ \Lambda^T B_\alpha^L[k]\widehat{\theta} \end{bmatrix} \Bigg) B_\alpha^l[k] \Bigg] \Bigg) \widehat{\theta} = 0$$

Then, using equation 23:

$$\Phi\widehat{\theta} = \sum_{k=0}^{N} \Bigg(\Gamma_v \Lambda^T B_y[k]\widehat{\theta} B_y[k]\widehat{\theta} + \sum_{i=1}^{n_0} (e_i^n)^T \Gamma_w \begin{bmatrix} \Lambda^T B_w^1[k]\widehat{\theta} \\ \Lambda^T B_w^2[k]\widehat{\theta} \\ \vdots \\ \Lambda^T B_w^{n_0}[k]\widehat{\theta} \end{bmatrix} B_w^i[k]\widehat{\theta} + \sum_{l=0}^{L} \left(e_{l+1}^{L+1} \right)^T \Gamma_\alpha \begin{bmatrix} \Lambda^T B_\alpha^0[k]\widehat{\theta} \\ \Lambda^T B_\alpha^1[k]\widehat{\theta} \\ \vdots \\ \Lambda^T B_\alpha^L[k]\widehat{\theta} \end{bmatrix} B_\alpha^l[k]\widehat{\theta} \Bigg)$$

or

$$\Phi\widehat{\theta} = D_\theta(\widehat{\theta})\Lambda \tag{31}$$

where

$$D_\theta(\widehat{\theta}) = \sum_{k=0}^{N} B_y(k)\widehat{\theta}\Gamma_v \widehat{\theta}^T B_y^T(k) + \begin{bmatrix} B_\alpha^0(k)\widehat{\theta} & \cdots & B_\alpha^L(k)\widehat{\theta} \end{bmatrix} \Gamma_\alpha \begin{bmatrix} \widehat{\theta}^T B_\alpha^0(k)^T \\ \vdots \\ \widehat{\theta}^T B_\alpha^L(k)^T \end{bmatrix} + \begin{bmatrix} B_w^1(k)\widehat{\theta} & \cdots & B_w^{n_0}(k)\widehat{\theta} \end{bmatrix} \Gamma_w \begin{bmatrix} \widehat{\theta}^T B_w^1(k)^T \\ \vdots \\ \widehat{\theta}^T B_w^{n_0}(k)^T \end{bmatrix} \tag{32}$$

Similarly, using equations (29), (30),we have

$$\Phi^T \Lambda = D_\Lambda(\Lambda)\widehat{\theta} \tag{33}$$

where

$$D_\Lambda(\Lambda) = \sum_{k=0}^{N} B_y(k)^T \Lambda \Gamma_v \Lambda^T B_y^(k) + \begin{bmatrix} B_\alpha^0(k)^T \Lambda & \cdots & B_\alpha^L(k)^T \Lambda \end{bmatrix} \Gamma_\alpha \begin{bmatrix} \Lambda^T B_\alpha^0(k) \\ \vdots \\ \Lambda^T B_\alpha^L(k) \end{bmatrix} + \begin{bmatrix} B_w^1(k)^T \Lambda & \cdots & B_w^{n_0}(k)^T \Lambda \end{bmatrix} \Gamma_w \begin{bmatrix} \Lambda^T B_w^1(k) \\ \vdots \\ \Lambda^T B_w^{n_0}(k) \end{bmatrix} \tag{34}$$

As noted earlier (see Remark 1) we are in fact interested in $\theta^o \in \mathbb{R}^{n_o(N+L+2)+1}$. Thus, we define

$$L_2 = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & L_1 \end{array} \right] \tag{35}$$

where L_1 is as in (13). We then reexpress the problem in terms of the reduced parameter vector

$$\theta = L_2 \theta^o \tag{36}$$

In summary, to solve our estimation problem we need to solve the following equations.

$$\begin{aligned} \Phi L_2 \widehat{\theta}^o &= D_\theta (L_2 \widehat{\theta}^o)\Lambda \\ L_2^T \Phi^T \Lambda &= L_2^T D_\Lambda(\Lambda) L_2 \widehat{\theta}^o \\ (\widehat{\theta}^o)_1 &= -1 \end{aligned} \tag{37}$$

Using the definition of the cost function (16), and equations (26)-(28), we have:

$$J = \Lambda^T D_\theta(\widehat{\theta})\Lambda \tag{38}$$

Then using (37) we have:

$$\begin{aligned} J &= \Lambda^T \Phi L_2 \widehat{\theta}^o \\ &= (\widehat{\theta}^o)^T L_2^T \Phi^T \Lambda \\ &= (\widehat{\theta}^o)^T L_2^T \Phi^T D_\theta^{-1}(L_2\widehat{\theta}^o)\Phi L_2 \widehat{\theta}^o \\ &= \begin{bmatrix} -1 \\ \widehat{\theta}_1^o \end{bmatrix}^T L_2^T \Phi^T D_\theta^{-1} \left(L_2 \begin{bmatrix} -1 \\ \widehat{\theta}_1^o \end{bmatrix} \right) \Phi L_2 \begin{bmatrix} -1 \\ \widehat{\theta}_1^o \end{bmatrix} \end{aligned} \tag{39}$$

This is the result given in (25). $\square$

Note that the result of the above theorem is that the original cost function (16), which is a function of the variables $(\hat{y}(k), \underline{\hat{\alpha}}(k), \hat{w}(k))$ has been transformed into a function of the variable $(\hat{\theta})$ which appears in the constraints (17), and (18) of the original problem.

Finally, since a receding horizon approach will be used to estimate the states $\widehat{\theta}_1^o$, the above optimization problem has to be solved for every sample, and only the last value of this vector ($\hat{x}[N+1]$) will be used.

5. A SIMPLE ILLUSTRATIVE EXAMPLE

In order to examine the advantages of the method developed in this paper, we consider state estimation for the following system:

$$\begin{aligned} x(k+1) &= 0.8x(k) + w(k) \\ y(k) &= x(k-d) + v(k) \end{aligned} \tag{40}$$

where $E\{w(k)w(k)^T\} = 1$, $E\{v(k)v(k)^T\} = 10^{-2}$, and d is an uniform random variable, $d \sim Unif(0,4)$.

We will compare the estimates obtained with the **STLS** algorithm with those obtained by a standard Kalman Filter based on the average delay.

Using $N = 10$ as the prediction horizon, the results are shown in figure 1.

Table 1 gives the total quadratic error for the two estimators. We see that the performance achieved by the state estimator proposed here is significantly better than the performance achieved using the Kalman Filter.

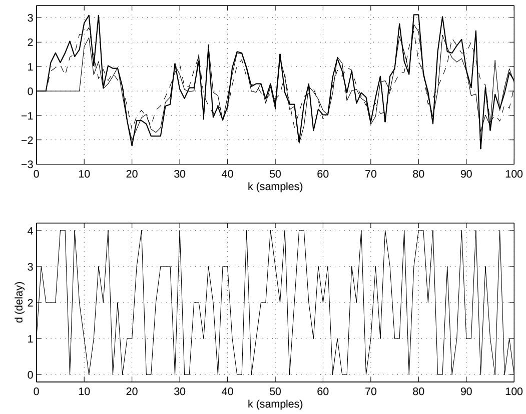

Fig. 1. State estimation: $x(k)$ (continuous thick line), $\hat{x}_{kalman}^{d=2}(k)$ (dash line), $\hat{x}_{stls}(k)$ (continuous thin line)

Table 1. Comparison between the traditional Kalman Filter and the technique proposed in this paper.

	$\sum_{k=10}^{100} (x(k) - \hat{x}(k))^2$
Kalman Filter	73.9
STLS Filter	35.1

Remark 2. For simplicity of exposition, this paper has considered the case when the input $u[k]$ is equal to zero. However, it is possible to extend the results to cover the case $u[k] \neq 0$. Indeed, it is anticipated that having a non-zero known input will significantly enhance the performance of the algorithm.

Remark 3. Again for simplicity, the paper has considered the case when the vector $\alpha(k)$ (which is a selector for the delay) is non-correlated at different samples, i.e. $E\{[\underline{\alpha}(k) - \underline{\bar{\alpha}}][\underline{\alpha}(k-\tau) - \underline{\bar{\alpha}}]^T\} = 0$ for $\tau \neq 0$. In practical problems, it is more likely that successive delays will be correlated. For example the delay encountered in a networked communication system is likely to change slowly depending on network loading. These more general cases can be treated by including a correlated delay structure in the cost function (16).

6. CONCLUSIONS

This paper has shown how the problem of state estimation with random (and unknown) delays can be formulated as an errors in variables problem. Particular emphasis has been given to issues arising from the special structure of the problem including prior knowledge regarding the error covariance structure. In the simple example presented, it has been shown that the mean quadratic error reached for the state estimator proposed here is smaller than that reached by the traditional Kalman Filter.

7. REFERENCES

Anderson, B. and J. Moore (1979). *Optimal Filtering*. Prentice Hall. Englewood Cliffs, N.J.

Branicky, M. S., S. M. Philips and W. Zhang (2000). Stability of network control systems: Explicit analysis of delay. In: *Proceedings of the American Control Conference*. Chicago, Illinois. pp. 2352–2357.

Bushnell, L. (2001). Ed. special section on networks and control. *IEEE Control Systems Magazine* **21**(1), 57–99.

Chan, H. and Ű. Őzgűner (1995). Closed-loop control of systems over a communications network with queues. *International Journal of Control* **62**(3), 493–510.

De Moor, B. (1993). Structured total least squares and l_2 approximation problems. *Linear Algebra and its Application* **188,189**, 163–205.

Golub, G.H. and C. Van Loan (1980). An analysis of the total least squares problem. *SIAM Journal of Numerical Analysis*.

Hong, S.H. (1995). Scheduling algorithm of data sampling times in the integrated communication and control systems. *IEEE Trans. on Control Systems Technology* **3**(2), 225–230.

Kaplan, G. (2001). Ethernet's winning ways. *IEEE Spectrum* pp. 113–115.

Krtolica, R., Ű. Őzgűner H. Chan H. GőM. Liubakka (1990). Stability of linear feedback systems with random communication delays. *ASME Journal of Dynamics, Measurement and Control* **59**(94), 925–953.

Lemmerling, P. (1999). Stuctured total least squares: Analysis, algortitms and applications. PhD thesis. Katholieke Universiteit Leuven. Belgium.

Luck, R. and A. Ray (1990). An observer-based compensator for distributed delays. *Automatica* **26**(5), 903–908.

Michalska, H. and D.Q.Mayne (1995). Moving horizon observers and observer based control. *IEEE Transactions on Automatic Control* **40**(6), 995–1006.

Muske K. R., J. B. Rawlings and K.H. Lee (1993). Receding horizon recursive estimation. In: *Proceedings of the American Control Conference*. pp. 900–904.

Nair, G.N. and R.J. Evans (2000). Stabilization of data-rate-limited feedback: Tightest attainable bounds. *Systems & Control Letters* **41**, 49–56.

Nilsson, J., B. Bernhardsson and B. Wittenmark (1998). Stochastic analysis and control of real-time systems with random time delays. *Automatica* **34**(1), 57–64.

Rao, C.V., J.B. Rawlings and J.H. Lee (2001). Constrained linear state estimation - a moving horizon approach. *Automatica* **37**(10), 1619–1628.

Ray, A. (1987). Performance evaluation of medium access control protocols for distributed digital avionics. *Transactions of the ASME* (109), 370–377.

Van Huffel, S. and J. Vandewalle (1991). *The total least squares problem: Computational aspects and analysis*. SIAM Frontiers Applied Math. Philadelphia.

Van Huffel, S. and P. Lemmerlimg (2002). *Errors in variables modelling: Analysis, Algorithms and Applications*. Kluwer Academic Publishers. Philadelphia. to appear.

Walsh, G.C., H.Ye and L. Bushnell (1999). Stability analysis of networked control systems. In: *Proceedings of the American Control Conference*. pp. 2876–2880.

Appendix A. APPENDIX A

Recall that the data is received with a random delay taking values in $[0, L]$ with equal probability $\frac{1}{L+1}$. At each time instance k, there are two possibilities. Either no data is received, in which case we have $\alpha_i[k] = 0$ for all i, or, a number of data values arrive simultaneously. In the latter case, one data point corresponding to say, delay $0 \le d \le L$ is picked at random. This means that $\alpha_d[k] = 1$ and $\alpha_i[k] = 0$ for $i \neq d$. It can readily be shown that the probability that all $\alpha_i[k] = 0$ is given by $p = \left(\frac{L}{L+1}\right)^{L+1}$ and the probability that any $\alpha_d[k] = 1$ is the same for all d. Hence the latter probability is given by $q = \frac{1-p}{L+1}$. Thus,

$$\begin{aligned} \overline{\alpha}_i &= E\{\alpha_i[k]\} \\ &= q \end{aligned} \tag{A.1}$$

and

$$\begin{aligned} \Gamma_\alpha &= \operatorname{cov}\left\{\begin{bmatrix} \alpha_0[k] \\ \vdots \\ \alpha_L[k] \end{bmatrix}\right\} = \operatorname{cov}\{\underline{\alpha}[k]\} \\ &= q\left(I - q\underline{1}\underline{1}^T\right)^2 \end{aligned} \tag{A.2}$$

Note that, since $q < \frac{1}{L+1}$, Γ_α is nonsingular.