Paper No.I-3a-14-4
Proceedings of the 14th IFAC, ISBN 0 08 043248 4

A FINITE-DIMENSIONAL CODER-ESTIMATOR FOR RATE-CONSTRAINED STATE ESTIMATION

Girish N. Nair; Robin J. Evans*

* *Centre for Sensor Signal and Information Processing Department of Electrical and Electronic Engineering University of Melbourne Parkville VIC 3052* Australia email: gnair@ee.mu.oz.au, robevans@ee.mu.oz.au

EXTENDED ABSTRACT

This paper considers the problem of estimating the state of a linear system via a digital link with a finite bit-rate R. Classical estimation theory does not apply here since the estimator only observes the transmitted sequence of finite-valued symbols. A new coder-estimator for linear systems with process and measurement noise is given. Provided that a certain inequality linking the bit-rate to the dynamical parameters is satisfied, and under very mild assumptions on the noise distributions, the expected absolute estimation error remains of the same order as in the classical situation with infinite bit-rate. In particular, if the classical estimation error approaches zero, then the finite bit-rate error goes to zero at exactly the same rate.

KEYWORDS

• State estimation • Kalman filter • Finite bit-rate

14th Triennial World Congress, Beijing, P.R. China

I-3a-14-4

A FINITE-DIMENSIONAL CODER-ESTIMATOR FOR RATE-CONSTRAINED STATE ESTIMATION

Girish N. Nair and Robin J. Evans *

* *Centre for Sensor Signal and Information Processing*
Department of Electrical and Electronic Engineering
University of Melbourne
Parkville VIC 3052
Australia
email: gnair@ee.mu.oz.au, robevans@ee.mu.oz.au

Abstract: This paper considers the problem of estimating the state of a linear system via a digital link with a finite bit-rate R. Classical estimation theory does not apply here since the estimator only observes the transmitted sequence of finite-valued symbols. A new coder-estimator for linear systems with process and measurement noise is given. Provided that a certain inequality linking the bit-rate to the dynamical parameters is satisfied, and under very mild assumptions on the noise distributions, the expected absolute estimation error remains of the same order as in the classical situation with infinite bit-rate. In particular, if the classical estimation error approaches zero, then the finite bit-rate error goes to zero at exactly the same rate. *Copyright © 1999 IFAC*

Keywords: state estimation, Kalman filter, finite bit-rate

1. INTRODUCTION

In many problems in state estimation, the implicit assumption is that the estimator has direct access to a sequence of possibly noise-corrupted measurements. However, in certain situations there is a constraint on the amount of information available to the estimator. In particular, if the estimator is not at the same location as the measurement sensor and receives information via a digital communication channel with a finite capacity, then it can only have partial knowledge of the measurements, since they must be mapped to a finite set of symbols before transmission to the estimator. This causes a loss of resolution and also introduces an additional delay into the system, since the bits that constitute each symbol require a finite amount of time for complete transmission. The question is how to construct a joint coding and estimation scheme so that the estimation error is small in some defined sense.

Although this problem is related in a fundamental way to those addressed in in rate distortion theory (Berger, 1971) and estimation via compressed information (Zhang and Berger, 1989; Ahlswede *et al.*, 1997), the approach taken here is causal, whereas the results obtained in those areas typically use long block codes. In the context of online estimation this implies long delays, which is impractical when estimates are required immediately. The problem of estimation using quantized measurements is also related (Delchamps, 1989; Curry, 1970), however instead of a fixed quantizer the coder here is permitted to depend in a time-varying manner on all available system outputs. As such the estimator effectively has the freedom to choose what to measure, which makes the problem significantly different.

The causal approach taken in this paper follows on from the seminal work in (Wong and Brockett, 1995; Li and Wong, 1996; Wong and Brock-

 ISBN: 0 08 043248 4

ett, 1997) on *recursive coder-estimators.* Conditions were obtained there for the mean-square estimation error to remain bounded, under various restrictions on the distributions of the initial condition and process and measurement noise. These conditions were shown to lead to an aymptotic mean-square estimation error of zero for the special case when there was no process or measurement noise. In (Nair and Evans, 1997), a coder-estimator was derived that yielded an aymptotic estimation error of zero under less restrictive conditions, in the absence of process and measurement noise. In (Borkar *et al.*, 1997), the problem of causally coding and estimating a stable, Markovian process under entropy and distortion constraints was formulated and the existence of an optimal scheme under a combined cost on the entropy and distortion was established. In this paper, systems with process and measurement noise are considered under a hard bit-rate constraint. The problem is first formulated for a general D-dimensional linear system and then a coder-estimator is explicitly constructed for the particular case of a one-dimensional system. No assumptions are made about the distributions of the noise terms, other than some very weak ones on the average deviations of the innovations and the thickness of the tails of their distributions. It is shown that, provided the bit-rate and dynamic constants satisfy a certain inequality, the expected absolute estimation error is still of the same order as that obtained in the classical case with an infinite bit-rate.

2. FORMULATION

The formulation given below is a special case of a more general one discussed in (Nair and Evans, 1999). Consider the linear process

$$X_{k+1} = A_k X_k + V_k, \quad \forall k \geq 0, \tag{1}$$

observed via the output process

$$Y_k = H_k X_k + W_k, \tag{2}$$

where $X_k \in \mathbf{R}^D$, $Y_k \in \mathbf{R}^{D'}$ and $\{V_k, W_k\}_{k\geq 0}$ are mutually independent. Suppose estimates of the current state are required at a distant location, to be transmitted via a digital communication channel with a capacity of R bits of information per sampling interval (Cover and Thomas, 1991). For simplicity, we convert this into the more restrictive condition that only R bits of data may be sent per sampling period, where R is now constrained to be an integer. This implies that at each time k, one symbol from an alphabet of size $M = 2^R$ is sent to the estimator. Assuming for now that there is no restriction on the complexity of the coder, the transmitted symbol in general depends on all the past and present outputs of the system, i.e.

$$S_k = \gamma_k(\tilde{Y}_k, \tilde{S}_{k-1}), \quad \forall k \geq 0, \tag{3}$$

where $\gamma_k : \mathbf{R}^{D' \times (k+1)} \times \mathbf{Z}_M^k \to \mathbf{Z}_M$ is to be chosen and the notation $\tilde{z}_k$ denotes a sequence $\{z_j\}_{j=0}^k$, with the convention that $\tilde{z}_{-1} \triangleq \{\}$. Due to the finite bit-rate, each transmitted symbol takes one sampling interval to reach the estimator in entirety, neglecting any propagation delay. Hence at time k the estimator has $\tilde{s}_{k-1}$ available and estimates the current state x_k via

$$\hat{X}_k = \delta_k(\tilde{S}_{k-1}), \quad \forall k \geq 0, \tag{4}$$

where $\delta_k : \mathbf{Z}_M^k \to \mathbf{R}^D$ is to be chosen.

The objective is to choose a coder-estimator, defined by the joint sequence of mappings $\{\gamma_k, \delta_k\}_{k\geq 0}$, such that the estimation error is small in some defined sense. Unlike the classical situation, corresponding to $R \to \infty$, minimizing $\mathrm{E}\{\|X_k - \hat{X}_k\|^2\}$ for each $k \geq 0$ is not appropriate. Since $\hat{X}_k$ depends on all the past and present choices of γ_j, $j \leq k$, a choice of $\tilde{\gamma}_k$ that minimizes $\mathrm{E}\{\|X_k - \hat{X}_k\|^2\}$ will not in general minimize $\mathrm{E}\{\|X_{k+1} - \hat{X}_{k+1}\|^2\}$, so the coder will be optimal only at time k. In this respect this resembles a control problem, since the control inputs γ_k and δ_k affect the outputs $\hat{X}_j$, $\forall j \geq k$. If estimates are needed only up to time N, then a more suitable measure of error is a weighted average such as

$$J_N = \mathrm{E}\{\sum_{k=0}^{N} \alpha_k \|X_k - \hat{X}_k\|^2\}, \tag{5}$$

where $\{\alpha_k\}_{k=0}^N$, are nonnegative numbers. If we are interested in the aymptotic performance of the coder-estimator, an appropriate measure would be $\limsup_{N\to\infty} J_N$, or even just $\limsup_{k\to\infty} \mathrm{E}\{\|X_k - \hat{X}_k\|\}$, assuming they exist.

In (Nair and Evans, 1999), a key result was the following theorem:

Theorem 1. For the system described by (1) and (2) with the additional assumption that X_0, V_k, W_k, are Gaussian $\forall k \geq 0$, there exists an optimal coder-estimator under the finite-time, quadratic error criterion (5). Furthermore, it has a coding equation of the form

$$S_k = \gamma_k^*(\bar{X}_{k+1}, \tilde{S}_{k-1}),$$

where $\bar{X}_{k+1} = \mathrm{E}\{X_{k+1}|\tilde{Y}_k\}$. That is, the first stage of any optimal coder consists of a Kalman one-step-ahead predictor which is followed by a stage that uses all R bits available at time k to

 ISBN: 0 08 043248 4

encode only the most recent Kalman estimate, according to the symbols alreadt transmitted.

This says that the optimal coder-estimator first filters the measurements, so as to obtain the best possible state estimate before transmission on the digital link, and then encodes only the most recent filter output, according to the symbols already transmitted. Expressions for the minimum error and optimal coder-estimator are derived in (Nair and Evans, 1999). However, as these expressions are recursive, it is difficult to solve them to obtain explicit equations for the optimal cost and coder-estimator. Techniques such as the Generalized Lloyd Algorithm (Gersho and Gray, 1993) could be adapted to obtain numerical expressions for them, but only for specific dynamical parameters, noise densities, bit-rate and a fixed, finite time N. As such, in the next section we will use the structural result given above to obtain a suboptimal coder-estimator that depends explicitly on the system dynamics, noise statistics and bit-rate.

3. SUBOPTIMAL CODER-ESTIMATOR

In this section, we first construct a coder-estimator for the one-dimensional linear system

$$\begin{aligned} X_{k+1} &= a_k X_k + V_k, \\ Y_k &= X_k + W_k, \quad \forall k \geq 0, \end{aligned} \tag{6}$$

where $\{V_k, W_k\}_{k\geq 0}$ are mutually independent. The extension to multidimensional systems then follows by using diagonalization. Motivated by the results of the previous section, we let the first stage of the coder be a Kalman filter and then find a coder-estimator for a directly observed, scalar process of the form

$$X'_k = a_k X'_{k-1} + Z_k, \quad \forall k \geq 0, \tag{7}$$

where X'_{-1} is equal to a known constant x'_{-1}. In the context of coding and estimating the system (6), the state X'_k is the Kalman estimate $\mathrm{E}\{X_{k+1}|\tilde{Y}_k\}$ and $\{Z_k\}_{k\geq 0}$ are the innovations. If the process and measurement noise are Gaussian, then the innovations are mutually independent, but this is not needed for the error bounds of the next section. The only requirements imposed on the noise in the system are that there exist positive constants A, B, a known positive number n and a known nonnegative sequence $\{\epsilon_k\}_{k\geq 0}$ such that

$$\epsilon_k \leq \mathrm{E}\{|Z_k|\} \leq B\epsilon_k, \quad \forall k \geq 0, \tag{8}$$

and $\forall u > 0$,

$$\int\limits_{|z|>u\epsilon_k} |z| p_{Z_k}(z) dz \leq \frac{A}{u^n} \mathrm{E}\{|Z_k|\}, \quad \forall k \geq 0. \tag{9}$$

The first condition requires the average deviations of the innovations to be known to within some uniform, multiplicative constant, so that the coder-estimator can compensate for them. The second condition places a bound on the probability of obtaining arbitrarily large innovations, which have the effect of slowing down the convergence of the coder-estimator. Both these requirements are easily satisfied in most instances and admit a very large class of distributions. In particular, when Z_k is Gaussian with standard deviation σ_k, they hold for any $n > 0$ if ϵ_k is set equal to σ_k.

In (Nair and Evans, 1997), a coder-estimator was given for D-dimensional linear systems with no process or measurement noise. An inequality linking the dynamics of the system to the data rate was derived which guaranteed that the estimation error went to zero as time progressed. However, if process and measurement noise are present, the situation is somewhat different. At each time k the system (7) is perturbed by a new term Z_k, which must be compensated for by the coder-estimator. Since Z_k is not bounded, i.e. p_{Z_k} has support on the entire real line, and $\mathrm{E}\{|Z_k|\}$ does not necessarily approach zero as k increases, it is not immediately obvious how this might be done. As such, before stating the coder-estimator equations we give a heuristic, step-by-step explanation of how they are obtained. The convergence results will then be stated without proof in the next section.

We first consider how the coder-estimator proceeds if after time k the innovations noise disappears and the process become static, i.e. $X'_j = X'_{j-1}$, $\forall j > k$. In this situation it divides the real line into successively finer intervals, so that at time $j > k$ there are $M^{j-k} - 2$ finite intervals and two semi-infinite ones. Let q_j be the estimate of $x'_j = x'_k$ at time $j \geq k$. At time $j > k$, denote the interval which contains x'_k by I_j, defined in terms of q_j and two additional positive numbers l_j^+ and l_j^- by

$$I_j \triangleq \begin{cases} (-\infty, q_j + l_j^+] & \text{if } s_{k+1} = \cdots = s_j = 0, \\ (q_j - l_j^-, \infty) & \text{if } s_{k+1} = \cdots = s_j = M-1, \\ (q_j - l_j^-, q_j + l_j^+] & \text{otherwise.} \end{cases}$$

The manner in which the interval I_j is further subdivided depends critically on which of these three classes it falls into. As we desire the coder-estimator to be finite-dimensional and not to keep track of the entire sequence of transmitted symbols, the conditions on the R.H.S. of the above equation will be encapsulated by two nonnegative numbers λ_j^+ and λ_j^- which evolve according to

ISBN: 0 08 043248 4

$$\lambda_{j+1}^{+} = \begin{cases} \phi\lambda_j^{+} & \text{if } s_{j+1} = M-1 \\ 0 & \text{otherwise} \end{cases}, \quad (10)$$

$$\lambda_{j+1}^{-} = \begin{cases} \phi\lambda_j^{-} & \text{if } s_{j+1} = 0 \\ 0 & \text{otherwise} \end{cases}, \quad (11)$$

where $\phi \geq 1$. Hence if $\lambda_k^{+}, \lambda_k^{-} > 0$, $\lambda_j^{+} > 0$ iff $s_{k+1} = \cdots = s_j = M-1$ and $\lambda_j^{-} > 0$ iff $s_{k+1} = \cdots = s_j = 0$. The next question is how to generate s_{j+1}, q_{j+1}, l_{j+1}^{+} and l_{j+1}^{-} at time $j+1$, given x'_{j+1} and the previous state of the coder-estimator,

$$\psi_j \triangleq (q_j, l_j^{+}, l_j^{-}, \lambda_j^{+}, \lambda_j^{-})^{\mathrm{T}} \in \mathbf{R}^5.$$

A natural strategy is to construct an increasing sequence of M points, $\{\pi(\psi_j, s)\}_{s=0}^{M-1}$ in the previous interval I_j and use the nearest-neighbour rule to select the next symbol and estimate,

$$s_{j+1} = \arg\min_{s\in\mathbf{Z}_M} |x'_{j+1} - \pi(\psi_j, s)|, \quad (12)$$
$$= \arg\min_{s\in\mathbf{Z}_M} |x'_k - \pi(\psi_j, s)|,$$
$$q_{j+1} = \pi(\psi_j, s_{j+1}), \quad (13)$$

with the convention that if there are two minimizing values of s the lesser is chosen. As we would like the absolute error $|x'_k - q_j|$ to be continuous with respect to x'_k, $\forall j \geq k$, q_j and the quantization points nearest to it on the on the left and right, if any, must be equidistant from the common boundary of their corresponding containing intervals. Keeping this constraint in mind, and lacking any precise knowledge of $p_{X'_k|\tilde{S}_k}$, a reasonable thing to do is to distribute the candidate points $\{\pi(\tilde{s}_k, s)\}_{s\in\mathbf{Z}_M}$ linearly about q_k using

$$\pi(\psi_j, s) = q_j - \frac{M-1}{M} l_j^{-} + \frac{s}{M}(l_j^{+} + l_j^{-}). \quad (14)$$

The values of l_{j+1}^{+} and l_{j+1}^{-} are then updated according to the symbol s_{j+1} transmitted, via

$$l_{j+1}^{+} = \begin{cases} \frac{1}{M} l_j^{+} + \lambda_j^{+} & \text{if } s_{j+1} = M-1 \\ \frac{1}{2M}(l_j^{+} + l_j^{-}) & \text{otherwise} \end{cases}, \quad (15)$$

$$l_{j+1}^{-} = \begin{cases} \frac{1}{M} l_j^{-} + \lambda_j^{-} & \text{if } s_{j+1} = 0 \\ \frac{1}{2M}(l_j^{+} + l_j^{-}) & \text{otherwise} \end{cases}, \quad (16)$$

It can be verified with some effort that these update equations satisfy the continuity requirement on $|x'_k - q_j|$. In addition, by choosing the expansion constant ϕ to be greater than 1, the quantization points q_j furthest out from the middle spread exponentially fast with j over the real line while those in the middle still continue to pack together like $M^{-(j-k)}$, thereby enabling the coder-estimator to deal with distributions $p_{X'_k|\tilde{S}_k}$ that have unbounded support.

The next step is to account for the dynamics of the process. Note that if $a_{k+1} < 0$, the left and right tails of $p_{X'_k|\tilde{S}_k}$ become the right and left tails respectively of $p_{a_{k+1}X'_k|\tilde{S}_k}$. As such (l_k^{+}, λ_k^{+}) and (l_k^{-}, λ_k^{-}) will have to be swapped, which can be done concisely by means of a matrix

$$T_k = \begin{cases} \mathrm{I} & \text{if } a_k \geq 0, \\ \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} & \text{otherwise.} \end{cases} \quad (17)$$

The update equations (10) - (16) are then modified by making the substitutions $x'_k \to a_{k+1}x'_k$ and $\psi_k \to |a_{k+1}|T_{k+1}\psi_k$. The innovations noise Z_{k+1} is finally compensated for by including an additive term ϵ_{k+1}, defined in (8) and (9), to yield the coder-estimator equations below:

Coder 1. At time $k \geq -1$, let the state of the coder-estimator be defined by $\psi_k \in \mathbf{R}^5$, where

$$\psi_k \triangleq (q_k, l_k^{+}, l_k^{-}, \lambda_k^{+}, \lambda_k^{-})^{\mathrm{T}}, \quad (18)$$
$$\psi_{-1} = (x'_{-1}, 0, 0, 0, 0).$$

Define

$$\nu_k \triangleq (0, \epsilon_k, \epsilon_k, \epsilon_k, \epsilon_k), \quad (19)$$
$$\bar{\psi}_k \triangleq |a_{k+1}|T_{k+1}\psi_k + \nu_{k+1}, \quad (20)$$
$$= (\bar{q}_k, \bar{l}_k^{+}, \bar{l}_k^{-}, \bar{\lambda}_k^{+}, \bar{\lambda}_k^{-}),$$

where ϵ_k is defined by (8) - (9), T_{k+1} by (17) and a_{k+1} by (7). At time $k+1$, the coder constructs M numbers

$$\pi(\bar{\psi}_k, s) \triangleq \bar{q}_k - \frac{M-1}{M}\bar{l}_k^{-} + \frac{s}{M}(\bar{l}_k^{+} + \bar{l}_k^{-}), \quad (21)$$

$\forall s \in \mathbf{Z}_M$ and encodes x'_{k+1} using the nearest-neighbour rule

$$s_{k+1} = \arg\min_{s\in\mathbf{Z}_M} |x'_{k+1} - \pi(\bar{\psi}_k, s)|, \quad (22)$$
$$q_{k+1} = \pi(\bar{\psi}_k, s_{k+1}), \quad (23)$$

with the convention that if there are two minimizing values of s the lesser is chosen. The remaining components of the coder state are then updated as follows:

$$l_{k+1}^{+} = \begin{cases} \frac{1}{M}\bar{l}_k^{+} + \bar{\lambda}_k^{+} & \text{if } s_{k+1} = M-1 \\ \frac{1}{2M}(\bar{l}_k^{+} + \bar{l}_k^{-}) & \text{otherwise} \end{cases}, \quad (24)$$

$$l_{k+1}^{-} = \begin{cases} \frac{1}{M}\bar{l}_k^{-} + \bar{\lambda}_k^{-} & \text{if } s_{k+1} = 0 \\ \frac{1}{2M}(\bar{l}_k^{+} + \bar{l}_k^{-}) & \text{otherwise} \end{cases}, \quad (25)$$

ISBN: 0 08 043248 4

$$\lambda^+_{k+1} = \begin{cases} \phi\bar{\lambda}^+_k & \text{if } s_{k+1} = M-1 \\ 0 & \text{otherwise} \end{cases}, \quad (26)$$

$$\lambda^-_{k+1} = \begin{cases} \phi\bar{\lambda}^-_k & \text{if } s_{k+1} = 0 \\ 0 & \text{otherwise} \end{cases}, \quad (27)$$

where

$$\phi \triangleq \left(\frac{M}{2} + \frac{1}{4}\right)^{\frac{1}{1+n}} > 1, \quad \forall M \geq 2. \quad (28)$$

Estimator 1. After receiving the symbol s_k at time $k+1$, the estimator updates its copy of the state ψ_k via the same equations as Coder 1 and estimates x'_k using

$$\hat{x'}_k = q_k. \quad (29)$$

More succintly, the coder-estimator state evolves according to a dynamic equation of the form

$$\psi_{k+1} = g(s_{k+1}, \psi_k) \quad (30)$$

and has an output equation of the form

$$s_k = h(x'_k, \psi_k), \quad (31)$$

Having defined a coder-estimator for processes of the form (7), a coder-estimator for the system (6) can now be constructed:

Coder 2. At time $k \geq 0$, the first stage of the coder filters the measurements $\tilde{y}_k$ from the system (6) to form the Kalman estimate

$$\bar{x}_{k+1} = \mathrm{E}\{X_{k+1}|\tilde{y}_k\}.$$

The second stage is just Coder 1, which encodes $x'_k = \bar{x}_{k+1}$ and transmits a symbol s_k as defined above.

Estimator 2. At time k, Estimator 1 is used to yield

$$\hat{x}_k = \hat{x'}_{k-1} = q_{k-1}.$$

We briefly comment that multidimensional systems can be handled by diagonalizing the filter process and then using the scalar coder-estimator on each component, since each then obeys a recursion of the form (7) that is decoupled from the others. However, there remains the question of how many bits to allocate to each dimension. Using a constant number of bits to encode each component for all times k is one possible approach, but if D exceeds the number of bits R then some components will not be tracked at all. A better idea is use all R bits to encode only one component at a time in such a manner that the i-th component is selected T_i times in every duty cycle of length T. The frequency T_i/T should be chosen according to the relative stability of each component, so that unstable components are sampled and coded more frequently than stable ones (Nair, 1998).

In the next section upper bounds on the average absolute errors of these coder-estimators are given, which indicate that despite being suboptimal the errors they generate remain of the same order as in the classical, infinite bit-rate situation.

4. CONVERGENCE RESULTS

In this section, we present a number of results which give upper bounds on the the estimation errors of the coder-estimators given above. Proofs are omitted to save space, but it is worth mentioning that the additive manner in which Coder-Estimator 1 accounts for the tails of $p_{X'_k|\bar{S}_k}$ as well as the innovations Z_{k+1} is crucial, since it permits the construction of an upper-bounding function that splits subadditively to separate out the effect of Z_k. In addition, by looking at absolute rather than square errors, terms dependent on Z_k do not multiply others, hence it is unnecessary for $\{Z_k\}_{k\geq 0}$ to be mutually independent.

Theorem 2. If Coder-Estimator 1 is used on the process (7) under the condition

$$\limsup_{k\to\infty} \frac{|a_k|}{F} < 1, \quad (32)$$

where F is a certain number dependent on the bit-rate R, then $\exists\eta \in (0,1), \zeta > 0$ such that $\forall k \geq 0$,

$$\mathrm{E}\{|X'_k - \hat{X'}_k|\} \leq \zeta \sum_{j=0}^{k} \eta^{k-j}\mathrm{E}\{|Z_j|\}. \quad (33)$$

In particular, if $\exists\zeta' > 0$ such that $\forall k \geq 0$,

$$\sum_{j=0}^{k} \eta^{k-j}\mathrm{E}\{|Z_j|\} \leq \zeta'\mathrm{E}\{|Z_k|\}, \quad (34)$$

then

$$\mathrm{E}\{|X'_k - \hat{X'}_{k|k+1}|\} \leq \zeta\zeta'\mathrm{E}\{|Z_k|\}. \quad (35)$$

This result states that, provided (32) is met, the expected absolute estimation error of Coder-Estimator 1 is upper-bounded by a geometric average of the past noise deviations. The number η can be regarded as a measure of the rate at which Coder-Estimator 1 converges with no noise, i.e. when $Z_k = 0, \forall k \geq 1$. The inequality (34) is roughly equivalent to requiring that $\mathrm{E}\{|Z_k|\} \gg \eta^k$ for large k. If the noise deviations go to zero faster than η^k, then it can be shown that the R.H.S of (33) is of the order of $\mathrm{E}\{|Z_0|\}\eta^k$. That is, the uncertainty introduced by later noise terms

ISBN: 0 08 043248 4

is negligible compared to the uncertainty of the initial noise and so the system remains essentially deterministic. However, if (34) is met, then the current innovation slows down the rate of convergence so that the average absolute estimation error is of the same order as its deviation.

The exact expression for F is somewhat complicated and is given in (Nair, 1998). However, for highly unstable systems the sufficient condition (32) can be simplified using $M = 2^R$ to yield the inequality

$$R > 1 + \frac{n+1}{n} \log_2 \left(\limsup_{k \to \infty} |a_k| \right).$$

This is a slightly stronger requirement than the corresponding one in (Nair and Evans, 1997) for systems with no process or measurement noise, which is simply $R > \log_2 (\limsup_{k\to\infty} |a_k|)$. We conjecture that (32) may be relaxed so that it is the same as for deterministic systems.

The preceding theorem leads directly to the following result:

Theorem 3. If Coder-Estimator 2 is used on the system (6) under condition (32), then $\exists \eta \in (0, 1), \zeta > 0$ such that $\forall k \geq 0$,

$$\mathrm{E}\{|X_k - \hat{X}_k|\} \leq \mathrm{E}\{|X_k - \bar{X}_k|\} + \zeta \sum_{j=0}^{k-1} \eta^{k-1-j} \mathrm{E}\{|Z_j|\}, \quad (36)$$

where $\bar{X}_k = \mathrm{E}\{X_k | \tilde{Y}_{k-1}\}$. In particular, if the system is time-invariant and the process and measurement noise are stationary, then $\exists \bar{\zeta} > 0$ such that

$$\mathrm{E}\{|X_k - \hat{X}_k|\} \leq \bar{\zeta} \mathrm{E}\{|X_k - \bar{X}_k|\}. \quad (37)$$

In other words, under the stated conditions, the expected absolute estimation error of Coder-Estimator 2 remains of the same order as the error that would be obtained if there was no bit-rate constraint. For instance if there is no process noise and the measurement noise is stationary, then both the classical and data rate constrained errors go to zero at the same speed. This implies that a finite bit-rate is not a significant impediment to state estimation, provided an appropriate coder-estimator is used. Analogous results can be shown to hold for the multidimensional case (Nair, 1998).

5. CONCLUSION

In this paper we have considered the effect of a bit-rate constraint on the estimation of the state of a remote linear system. A coder-estimator for linear systems with process and measurement noise was explicitly constructed in terms of the system dynamics, noise statistics and bit-rate. Provided that certain conditions are satisfied, the estimation error remains of the same order as in the classical situation when $R \to \infty$. This suggests that a bit-rate constraint is not a significant impediment to state estimation, provided that the coding is done appropriately. Work is currently in progress on extending the results in this paper to multisensor systems with multiple bit-rate constraints.

6. REFERENCES

Ahlswede, R., E. Yang and Z. Zhang (1997). Identification via compressed data. *IEEE Trans. Info. Theory* **43**(1), 48–70.

Berger, T. (1971). *Rate Distortion Theory.* Prentice-Hall. Englewood Cliffs, New Jersey.

Borkar, V. S., S. K. Mitter and S. Tatikonda (1997). Optimal sequential vector quantization of markov sources. *IEEE Trans. Info. Theory (submitted).*

Cover, T. M. and J. A. Thomas (1991). *Elements of Information Theory.* Wiley-Interscience. New York.

Curry, R. E. (1970). *Estimation and Control with Quantized Measurements.* MIT Press. Cambridge, Massachusetts.

Delchamps, D. F. (1989). Extracting state estimation from a quantized output record. *Syst. Contr. Lett.* (13), 365–372.

Gersho, A. and R. M. Gray (1993). *Vector Quantization and Signal Compression.* Kluwer. Boston.

Li, X. and W. S. Wong (1996). State estimation with communication constraints. *Syst. Contr. Lett.* (28), 49–54.

Nair, G. N. (1998). *PhD thesis in progress.* University of Melbourne.

Nair, G. N. and R. J. Evans (1997). State estimation via a capacity-limited communication channel. In: *Proc. 36th IEEE Conf. Decision and Control.* pp. 866–71.

Nair, G. N. and R. J. Evans (1999). Structural results for finite bit-rate state estimation. In: *Information, Decision and Control '99 (accepted).*

Wong, W. S. and R. W. Brockett (1995). State estimation with finite communication bandwidth constraints. In: *Proc. 34th IEEE Conf. Decision and Control.* pp. 1400–1401.

Wong, W. S. and R. W. Brockett (1997). Systems with finite communication bandwidth constraints. *IEEE Trans. Automat. Contr.* **42**(9), 1294–1299.

Zhang, Z. and T. Berger (1989). Estimation via compressed information. *IEEE Trans. Info. Theory* **34**(2), 198–211.

ISBN: 0 08 043248 4