Title
State estimation with imperfect communications: escape time formulation and exact quantized-innovations filtering

Author
Huang, Chun-Chia

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

State estimation with imperfect communications: escape time formulation and exact quantized-innovations filtering

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy

in

Engineering Science (Mechanical Engineering)

by

Chun-Chia Huang

Committee in charge:

Professor Robert R. Bitmead, Chair
Professor Patrick J. Fitzsimmons
Professor Massimo Franceschetti
Professor William M. McEneaney
Professor Mauricio de Oliveira

2015

The Dissertation of Chun-Chia Huang is approved and is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2015

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge Part I is from the paper, “Escape time formulation of state estimation and stabilization with quantized intermittent communication”, *Automatica,* , Chun-Chia Huang, Robert R. Bitmead 2015. Part II is from the paper, “Exact formulation of quantized-innovations state estimation using Bayesian filtering,” submitted to IEEE transaction on signal processing , Chun-Chia Huang, Robert R. Bitmead 2015.

VITA

2005	Bachelor of Engineering, National Chung Cheng University, Chiayi Taiwan
2005–2007	Master of Engineering, National Chung Cheng University, Chiayi Taiwan
2007–2008	The Republic of China Army
2010–2013	Research Assistant, Department of Mechanical and Aerospace Engineering University of California, San Diego
2013-2015	Teaching Assistant, Department of Mechanical and Aerospace Engineering University of California, San Diego
2015	Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

"Escape time formulation of state estimation and stabilization with quantized intermittent communication", *Automatica,* Chun-Chia Huang, Robert R. Bitmead,2015

"Escape time formulation of state estimation and stabilization with intermittent communication", *European Control Conference*, Chun-Chia Huang, Robert R. Bitmead, 2013

"Exact formulation of quantized-innovations state estimation using Bayesian filtering," submitted to IEEE transaction on signal processing, Chun-Chia Huang, Robert R. Bitmead.

FIELDS OF STUDY

Major Field: Mechanical and Aerospace Engineering (Specialization in control theory)

Studies in Control theory
Professors Robert R. Bitmead

ABSTRACT OF THE DISSERTATION

State estimation with imperfect communications: escape time formulation and exact quantized-innovations filtering

by

Chun-Chia Huang

Doctor of Philosophy in Engineering Science (Mechanical Engineering)

University of California, San Diego, 2015

Professor Robert R. Bitmead, Chair

The problem of state estimation for a linear, time-varying, gaussian system from measurements which are communicated over an imperfect channel is considered from several perspectives. The communication imperfections include intermittency, channel noise, quantization, etc. The first part of the thesis examines the stochastic behavior of the state estimation error and of the regulated state itself in situations of intermittent and quantized measurements via the formulation of an escape time problem dealing with the cumulative distribution function of the probability of escape of these signals from a given set. This is compared to and contrasted with earlier analyses which considered the behavior of Kalman filters with intermittent data based on moments and conditional moments, and the evaluation of the minimal number of bits required for mean square stabilization. The main result shows the escape time is characterized by a Markov chain which is amenable to explicit analysis through the calculation of the its cumulative distribution function. The second part of the thesis focuses on developing an exact formulation of the conditional probability density function of the system state given quantized innovations signals communicated from a linear Kalman filter at the transmitter. This is based on Bayesian filtering and extends previous works on the subject but without the requirement for simplifying assumptions. This latter result follows from a simple observation concerning the correct choice of *state* for the transmitter, which includes the transmitters' Kalman filter estimate. This leads to an exact and recursive approach.

Chapter 1

Introduction

1.1 Framework

Wireless sensor networks have been in application for several decades because of the advantages stemming from remote communication of process data, such as the absence of cabling, grounding issues, remote installation, etc. However, imperfect communication can cause severe problems such as intermittency, channel noise, quantization, delay, etc.

We consider state estimation for the linear, time-varying, gaussian system at one terminal of a communications link – the transmitter.

$$x_{k+1} = A_k x_k + B_k u_k + w_k, \tag{1.1}$$

$$y_k = C_k x_k + v_k, \tag{1.2}$$

where $w_k \sim \mathcal{N}(0, Q_k)$, $v_k \sim \mathcal{N}(0, R_k)$, $x_0 \sim \mathcal{N}(\hat{x}_{0|-1}, \Sigma_{0|-1})$ with each of these signals mutually independent and with $\{w_k\}$ and $\{v_k\}$ white. Here, as usual, x_k, u_k, y_k, w_k, v_k are the system state, input, output, process noise and measurement noise signals of dimensions n, p, m, n, m respectively and $[A_k, B_k, C_k]$ are the system matrices of conformable dimensions. State estimation for this system given the measurements $\{y_k\}$ is performed by the Kalman predictor or filter, which we describe shortly. We will consider two variations of this problem determined by the data communicated from the transmitter through the imperfect channel to a receiver at which state estimation is performed.

In the first part of the thesis we discuss an intermittent and quantized communication channel providing data

$$z_k = \gamma_k \mathscr{Q}_d(y_k), \tag{1.3}$$

to the receiver. Function $\mathscr{Q}_d$ describes a subtractive dithered quantizer and signal $\gamma_k \in \{0,1\}$ captures the intermittency. The receiver given this data then computes a state estimate, $\hat{x}_{k|k-1}$, and possibly also a feedback control, u_k, which is applied to the system directly without passing through an additional communication channel. Such a single-link problem might arise in the control of a geographically distributed system with remote sensing and local actuation.

Under this structure, we examine the escape time of the state estimation error, $\tilde{x}_{k|k-1} = x_k - \hat{x}_{k|k-1}$, and, if controlled, of the state itself, x_k, from a given set $\mathscr{D} \subset \mathbb{R}^n$. By escape time we mean the first time in which the signal lies outside the set. Since the system is stochastic, this time is a random variable and we develop procedures to computes its cumulative distribution function as it evolves over time. This is compared and contrasted with earlier works which focused on the moments of the errors or of the controlled state. We believe that this escape time formulation is more appropriate and helpful an analysis than the moment-based approach, since for gaussian systems escape is almost sure no matter then bounds on moments, as we show shortly.

In the second part of thesis, we consider a clever and more communications oriented approach in which a Kalman predictor is operated at the transmitter to yield the least-squares optimal state estimate, $\hat{x}^{\mathrm{KF}}_{k|k-1}$, and the data communicated to the receiver is the transmitter-side innovations sequence, $\varepsilon_k = y_k - C_k\hat{x}^{\mathrm{KF}}_{k|k-1}$. The communication channel is both intermittent and quantized, although now through an arbitrary quantizer, $\mathscr{Q}$. This innovations sequence has the property that it is white, zero mean and low variance. So the bitrate associated with the digital channel is more efficiently used. The data arriving at the receiver is described by

$$\bar{\varepsilon}_k = \gamma_k \mathscr{Q}(\varepsilon_k). \tag{1.4}$$

This problem has been studied recently by a number of researchers who have needed to make simplifying assumptions in order to derive a receiver-side recursive state estimator based on Bayesian filtering. Our core observation is that the appropriate transmitter *state* for this Bayesian filter comprises the system state, x_k, augmented by the transmitter's Kalman predictor state, $\hat{x}^{\mathrm{KF}}_{k|k-1}$. This is used directly in the Bayesian filter to derive our results without further assumption. This does not mitigate the computational complexity of the Bayesian filter probability density function(PDF) calculations but does affect their accuracy.

1.2 Kalman Filter

The theme in this thesis is state estimation problem in an imperfect communication environment as described above. Before stepping into the main problems of the thesis, we would like to describe and unify the notation of Kalman filter (KF). Optimal state estimation for linear time varying gaussian systems with a least conditional mean squares criterion is provided by the KF. The initial condition of the Kalman predictor is the mean value $\hat{x}_{0|-1}$ with initial covariance $\Sigma_{0|-1}$. We present the Kalman filter here based on [1].

The system model is (2.1-2.2) with the assumptions concerning gaussian initial condition, process noise $\{w_k\}$, and measurement noise $\{v_k\}$ given immediately below. Since the KF is linear, the state estimate and the state estimation error are both gaussian and hence possess pdfs and conditional pdfs completely described by their first two moments. Denote: the output measurement sequence $Y_k = \{y_0, y_1, \cdots y_k\}$; and estimates:

$$\hat{x}_{k|k-1} = E[x_k \Big| Y_{k-1}] \qquad \hat{x}_{k|k} = E[x_k \Big| Y_k].$$

The estimate $\hat{x}_{k|k-1}$ minimizes

$$\Sigma_{k|k-1} = E\left[|x_k - \hat{x}_{k|k-1}|^2 \Big| Y_{k-1} \right],$$

and $\hat{x}_{k|k}$ minimizes

$$\Sigma_{k|k} = E\left[|x_k - \hat{x}_{k|k}|^2 \Big| Y_k \right],$$

We express the recursive algorithm Kalman filter as following.

measurement update

$$L_k = \Sigma_{k|k-1} C_k^T (C_k \Sigma_{k|k-1} C_k^T + R_k)^{-1},$$

$$K_k = A_k L_k, \qquad \textbf{(Kalman gain)}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k (y_k - C_k \hat{x}_{k|k-1}), \qquad \textbf{(Filtered state estimate)}$$

$$\Sigma_{k|k} = \Sigma_{k|k-1} - L_k C_k \Sigma_{k|k-1}. \qquad \textbf{(Filter covariance)}$$

$$\tilde{x}_{k|k} = x_k - \hat{x}_{k|k}, \qquad \textbf{(Filtered error)}$$

time update

$$\hat{x}_{k+1|k} = A_k \hat{x}_{k|k} + B_k u_k, \qquad \textbf{(Predicted state estimate)}$$

$$\Sigma_{k+1|k} = A_k \Sigma_{k|k} A_k^T + Q_k, \qquad \textbf{(Predicted covariance)}$$

$$\tilde{x}_{k+1|k} = x_{k+1} - \hat{x}_{k+1|k}, \qquad \textbf{(Prediction error)}$$

The KF recursion produces the conditional mean and the conditional variance of the state given the measurements. The following formula is the recursive formula for predicted error covariance, Riccati Difference Equation (RDE):

$$\Sigma_{k+1|k} = A_k\Sigma_{k|k-1}A_k^T - A_k\Sigma_{k|k-1}C_k^T(C_k\Sigma_{k|k-1}C_k^T + R_k)^{-1}C_k\Sigma_{k|k-1}A_k^T + Q_k,$$

and Lyapunov version of the RDE

$$\Sigma_{k+1|k} = (A_k - K_kC_k)\Sigma_{k|k-1}(A_k - K_kC_k)^T + K_kR_kK_k^T + Q_k.$$

There are several things worth mentioning, which we list as follows.

1. The recursive covariance formulas are precomputable and do not depend on the data while the but state formulas do.

2. The innovations signal $\varepsilon_k = y_k - C\hat{x}_{k|k-1}$ is the term in the filtered state estimate measurement update formula. It is white, gaussian, zero-mean, with covariance $P_{\varepsilon_k} = C_k\Sigma_{k|k-1}C_k^k + R_k$. We will use these properties in the second part of thesis.

3. For time invariant systems, a steady state solution of RDE exists (subject to conditions [1]) and is the limiting value of the RDE solution and satisfies the Algebraic Riccati Equation.

$$\Sigma = A\Sigma A^T - A\Sigma C^T(C\Sigma C^T + R)^{-1}C\Sigma A^T + Q,$$

4. The Information Filter is a variant of the KF recursion and yields the same solutions but has computational advantages in some circumstances. The derivation from the Kalman Filter to the Information Filter is based on the Matrix Inversion Lemma

and is shown in [1]. The covariance recursions of the Information Filter are

$$\Sigma_{k|k}^{-1} = \Sigma_{k|k-1}^{-1} + C_k^T R_k^{-1} C_k \tag{1.5}$$

$$\Sigma_{k+1|k}^{-1} = Q_k^{-1} - Q_k^{-1} A_k (\Sigma_{k|k}^{-1} + A_k^T Q_k^{-1} A_k)^{-1} A_k^T Q_k^{-1} \tag{1.6}$$

The importance of the Kalman filtering recursion to the work in this thesis is that it provides a background against which to compare the results derived. We study in Part I the formulation of the escape time for the KF prediction error from a specific domain in the circumstances where the measurement data are subject to intermittency and quantization as described by (1.3). These problems are no longer gaussian, because of the nonlinear quantization, and the escape time problem, even for gaussian processes, does not admit a straightforward recursive analysis. Our research presented here derives a recursive formula for the cumulative distribution function of escape as a function of time. This derivation is exact for known packet-dropping sequence $\{\gamma_k\}$. The distribution of escape time can then be computed by averaging over this sequence.

The escape time analysis of the Kalman predictor error is extended to the consideration of escape of the controlled output, y_k, of system (2.1-2.2), where the control signal is given by state estimate feedback

$$u_k = -K\hat{x}_{k|k}.$$

These results provide an alternative to the mean-square stabilization results of [20], which provide no upper bound on the controlled output and assume reliable communications. On the other hand, they are able to determine minimal communication requirements.

In Part II of the thesis, we focus entirely on the construction of the conditional density of x_k given the intermittent, quantized measurements $\bar{E}_k = \{\bar{\varepsilon}_\ell : \ell = 0, ..., k\}$ given by (1.4). This is conducted using the concept from Bayesian filtering presented next. The innovations sequence, ε_k, in the KF above is gaussian, zero mean with prescribed covariance. By contrast, the quantized innovations, $\bar{\varepsilon}_k$, has compact support and need not be zero mean depending on the quantizer. The Bayesian filter allows us to develop a recursive approach to the computation of the conditional density of the state, from which any desired statistic, not just the conditional mean and variance, can be computed. This has been a longstanding technical problem.

1.3 Bayesian filter

The KF gives the conditional density for gaussian signals, as would apply at the transmitter side. Quantization is nonlinear and data signals at the receiver are not gaussian. So the KF no longer yields the correct conditional density. Bayesian filtering is a more general nonlinear technique to produce the conditional probability density function of the state. The content of this section is based on Bayesian state estimation as presented in [3].

Suppose we have a state-space system as follows.

$$x_{k+1} = f_k(x_k, w_k), \tag{1.7}$$

$$y_k = h_k(x_k, v_k), \tag{1.8}$$

where process noise $\{w_k\}$ and measurement noise $\{v_k\}$ are independent and white. The functions $f_k(\cdot,\cdot)$ and $h_k(\cdot,\cdot)$ are not necessarily linear time-varying system equations. The goal of a Bayesian estimator is to compute the predicted pdf $p(x_k \Big| Y_{k-1} = \{y_0, y_1 \ldots y_{k-1}\})$ and the filtered pdf $p(x_k \Big| Y_k = \{y_0, y_1, \ldots y_k\})$ starting from the initial state pdf $p(x_0)$.

The first probability density function we compute is the filtered pdf $p(x_0 \Big| y_0)$

using Bayes Rule,

$$p(x_0 \Big| y_0) = \frac{p(y_0 \Big| x_0)}{p(y_0)} p(x_0).$$

The term $p(y_0 \Big| x_0)$ can be computed in terms of the distribution of v_k. Then, we step into the algorithm to find predicted pdf $p(x_{k+1} \Big| Y_k)$ and filtered pdf $p(x_{k+1} \Big| Y_{k+1})$. The predicted pdf can be computed as follows.

$$\begin{aligned} p(x_{k+1} \Big| Y_k) &= \int_{x_k} p[(x_{k+1}, x_k) \Big| Y_k] dx_k \\ &= \int_{x_k} p[x_{k+1} \Big| (x_k, Y_k)] p(x_k \Big| Y_k) dx_k \\ &= \int_{x_k} p[x_{k+1} \Big| x_k] p(x_k \Big| Y_k) dx_k \end{aligned} \tag{1.9}$$

The second equality follows by Bayes' Rule and the third equality follows from the Markovian property of the state. The $p(x_{k+1} \Big| x_k)$ can be calculated from the system state equation (1.7) and the distribution of the process noise w_k. Relation (1.9) is the computation of predicted state conditional pdf from the filtered state conditional pdf — equivalent to the time update phase of the KF.

Next we compute the filtered conditional pdf of x_k from the predicted conditional pdf and the new measurement. This coincides with the measurement update phase of the

KF. Start by applying Bayes Rule.

$$
\begin{aligned}
p(x_{k+1}\Big|Y_{k+1}) &= \frac{p(Y_{k+1}\Big|x_{k+1})}{p(Y_{k+1})}p(x_{k+1}) \\
&= \frac{p(y_{k+1},Y_k\Big|x_{k+1})}{p(y_{k+1},Y_k)}\frac{p(x_{k+1}\Big|Y_k)p(Y_k)}{p(Y_k\Big|x_{k+1})} \\
&= \frac{p(x_{k+1},y_{k+1},Y_k)}{p(x_{k+1}),p(y_{k+1},Y_k)}\frac{p(x_{k+1},Y_k)p(Y_k)}{p(Y_k)p(Y_k\Big|x_{k+1})} \\
&= \frac{p(x_{k+1},y_{k+1},Y_k)p(x_{k+1},Y_k)p(Y_k)}{p(x_{k+1}),p(y_{k+1},Y_k)p(Y_k)p(Y_k\Big|x_{k+1})}\frac{p(x_{k+1},y_{k+1})}{p(x_{k+1},y_{k+1})} \\
&= \frac{p(Y_k\Big|x_{k+1},y_{k+1})p(y_{k+1}\Big|x_{k+1})p(x_{k+1}\Big|Y_k)}{p(y_{k+1}\Big|Y_k)p(Y_k\Big|x_{k+1})} \qquad (1.10)
\end{aligned}
$$

Next apply the Markovian property of the state in the term $p(Y_k\Big|x_{k+1},y_{k+1})$.

$$
p(x_{k+1}\Big|Y_{k+1}) = \frac{p(y_{k+1}\Big|x_{k+1})p(x_{k+1}\Big|Y_k)}{p(y_{k+1}\Big|Y_k)} \qquad (1.11)
$$

All of these terms are available from the output equation (1.8) and the density of the measurement noise v_k. The denominator can be computed as follows.

$$
\begin{aligned}
p(y_{k+1}\Big|Y_k) &= \int_{x_{k+1}} p(y_{k+1},x_{k+1}\Big|Y_k)dx_{k+1} \\
&= \int_{x_{k+1}} p(y_{k+1}\Big|x_{k+1},Y_k)p(x_{k+1}\Big|Y_k)dx_{k+1} \\
&= \int_{x_{k+1}} p(y_{k+1}\Big|x_{k+1})p(x_{k+1}\Big|Y_k)dx_{k+1} \qquad (1.12)
\end{aligned}
$$

So, we obtain the filtered pdf from the predicted pdf and the new measurement.

$$p(x_{k+1} \Big| Y_{k+1}) = \frac{p(y_{k+1} \Big| x_{k+1})p(x_{k+1} \Big| Y_k)}{\int_{x_{k+1}} p(y_{k+1} \Big| x_{k+1})p(x_{k+1} \Big| Y_k)dx_{k+1}} \tag{1.13}$$

The equations (1.9-1.13) form the Bayesian filter.

As will be demonstrated in Chapter 5, our contribution to quantized innovations state estimation is to apply the Bayesian filter to: system equations 2.1 and (2.2); the KF equations of Section 1.2; and the quantized innovations measurement equation (1.4) to yield an exact recursion for the predicted and filtered state conditional densities. The (surprising to us) novelty compared to earlier work in this area is that we recognize that the appropriate state for the quantized innovations filter consists of system state x_k augmented by KF estimator state $\hat{x}_{k|k-1}$. The Bayesian filter calculations then proceed directly. While the Bayesian filter is structurally very similar to the Kalman filter — indeed the Kalman filter is the Bayesian filter for linear gaussian systems — the Bayesian filter is significantly more complicated computationally than is the KF, because of the representation of the pdf. For the KF gaussian densities one may propagate the conditional mean and covariance, while the Bayesian filter requires a more complex representation, usually a sampling of the density function. An alternative approach for nonlinear state estimation is the Particle filter, where samples are propagated through the dynamics and then resampled in the measurement update. The Particle filter is computationally significantly more challenging than the Bayesian filter.

1.4 Background and literature survey

1.4.1 Background and literature survey of escape time

For the transmitter-side system (2.1-2.2) with communicated data $\{z_k\}$ given by (1.3), we define the escape time.

Definition 1 (Escape time) *Given a closed domain $\mathscr{D} \subset \mathbb{R}^d$ and a stochastic process $\{\xi_k : k = 1, \dots\}$ on $\mathbb{R}^d$, the escape time is defined to be*

$$\tau_e = \begin{cases} \arg\min_k \xi_k \notin \mathscr{D}, & \\ \infty, & \textit{if } \xi_k \in \mathscr{D} \ \forall k. \end{cases}$$

Sometimes the escape time is called the 'first exit time,' 'stopping time,' 'hitting time' or 'residence time.' We shall be concerned with the escape time for the state process, x_k, or the output process, y_k, of (2.1-2.2) when the control input is causally computed. Only for simplicity in this part, we take the system to be time-invariant: $A_k = A$, $B_k = B$, $C_k = C$.

Figure 1.1 shows two simulations of the Kalman state estimation error $\tilde{x}_{k|k-1}$ for a scalar system with a magnitude bound placed at 60 and differing rates of packet dropping. In the left graph, the expected value of the conditional error covariance – the focus of [25] – is finite, while in the right it is not. Note the similarities between the figures except for the time scales. Escape time would correspond to the first achievement of the bound 60. Our interest is in characterizing this escape time distribution function. Clearly the escape time is a random variable provided the infinite value has zero probability. We have the following result applicable in the linear Gaussian case and independent of the system matrices $[A, B, C]$.

Lemma 1 *For the linear system* (2.1)*, with noise process $\{w_k\}$ Gaussian, white, possessing full-rank covariance and independent from x_1, with control u_k causally computed,*

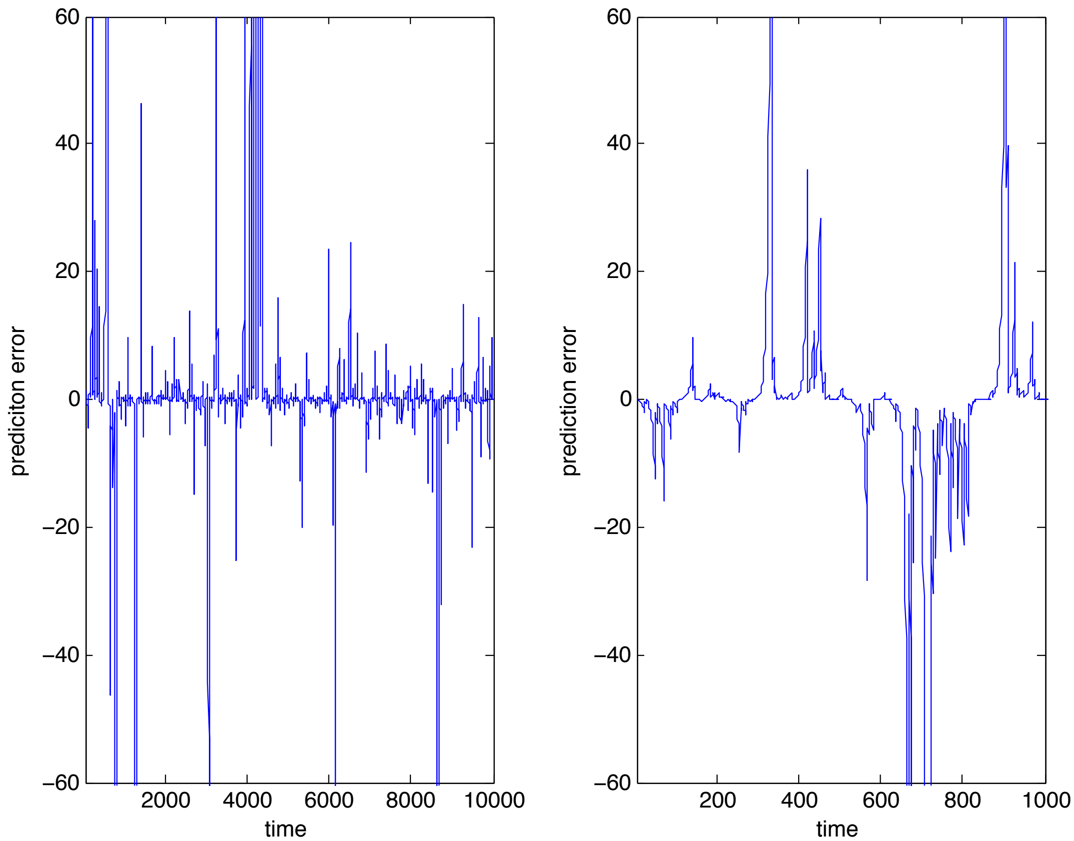

Figure 1.1. Simulation of Kalman predictor error with A=1.2, $C = 1$, $Q = 0.005$, $R = 0.001$, $P_\gamma = 0.15$ (left) and 0.1 (right), and bound level 60.

and for $\mathscr{D}$ compact, the escape time of x_k is almost surely finite.

Proof: Without loss of generality take $\mathscr{D}$ to be a hypersphere in $\mathbb{R}^n$, i.e. $|x|^2 \leq r$ for some $r > 0$. If this is not the case then, since $\mathscr{D}$ is compact, we can replace $\mathscr{D}$ by another compact set $\mathscr{D}^+$ which is a hypersphere containing $\mathscr{D}$ and perform the same analysis. Escape from $\mathscr{D}^+$ certainly implies escape from $\mathscr{D}$.

At time k with state x_k and control u_k, both of which are independent from w_k under the conditions of the lemma, the condition for x_{k+1} to lie outside $\mathscr{D}$ is that

$$Ax_k + Bu_k + w_k \in \mathscr{D}^c,$$

the complement of $\mathscr{D}$. Since $\mathscr{D}$ is a hypersphere and w_k is zero mean Gaussian

$$\Pr(w_k + v \in \mathscr{D}) \leq \Pr(w_k \in \mathscr{D}) = \beta < 1,$$

for any vector v. This argument can be extended to multiple steps as follows.

$$\begin{aligned}
&\Pr(w_k + v_k \in \mathscr{D}, w_{k+1} + v_{k+1} \in \mathscr{D}) \\
&= \Pr(w_{k+1} + v_{k+1} \in \mathscr{D} \Big| w_k + v_k \in \mathscr{D}) \Pr(w_k + v_k \in \mathscr{D}), \\
&\leq \beta \Pr(w_k + v_k \in \mathscr{D}), \\
&\leq \beta^2.
\end{aligned}$$

So the probability of escape by time k is bounded by β^k and, by the Borel-Cantelli Lemma, we have the probability of no escape is zero. ■

The import of Lemma 1 is that it ensures that, in the linear Gaussian case or equivalent problems able to be transformed to linear Gaussian, using say Girsanov's Theorem, the finite escape of the state and/or output from any compact domain is ensured. The analysis of such processes then ought to concentrate on the description of the escape time rather than attempting to establish almost sure confinement to a compact set or characterize moment properties. This hearkens back to the escape time or residence time analysis of, say, [13, 31, 39, 16, 8]. These earlier treatments focus on stable continuous-time systems with small stochastic perturbations and use the Theory of Large Deviations to develop escape time characterizations as the noise power tends to zero. Our approach will maintain discrete time and deal with both stable and unstable systems with non-infinitesimal perturbations. This will not draw on Large Deviations Theory other than for comparison.

Our treatment of (2.1-1.3) endeavors to blend two distinct trains of research. The first is associated with the behavior of state estimators for such systems as treated in, say, [25, 24, 12, 19] with or without control being applied. Since the system is linear and if the applied control is known, the controlled state behavior is derivable from the estimator. The second class of problems, characterized by results such as [30, 29, 20, 38] concentrates on the stabilization aspects of the feedback control. The distinction between the two sets of problems in the literature rests with the description of the communications channel. The work in [18, 21] studies the earlier work in a more general case and yields necessary conditions for stabilization, recovering results of some previous works in the two approaches. In the estimator problem, the communication is taken to be intermittent — that is the stochastic process $\{\gamma_k\}$ operates in a persistent fashion to cause arbitrarily long outages of communications — but the communication is not limited in bitrate (there is no quantizer) and full state reconstruction occurs with any successful communication packet. In earlier work on the stabilization problem, the emphasis is on the quantizer and its associated bitrate limit and the channel is assumed not intermittent, i.e. $\gamma_k = 1$ for all k, with a deterministic maximal delay and possible additive channel noise. The approach adopted in this paper is to permit both intermittency and limited bitrate, since the Markov model describing escape time applies to both. We also pose a different set of questions dealing with escape time, which we regard as being more apropos for these problems. These focus not on limiting behaviors or mean-square stabilization but on characterizing the cumulative probability distribution function (cdf) of the escape time of the system state, output or state estimate error, since in general there is no almost sure bound on these, as stated in Lemma 1.

Before launching into the analysis, it is pertinent to examine some practical sources of estimation and control problems associated with systems described by (2.1-1.3), since the presence of a single communications link rules out teleoperation-styled feedback control problems. Utility management of a geographically distributed system, such as a power grid or radar network, where the sensors, but not the actuators, are remotely placed and linked back to base by communications networks, is the clearest application of state estimation operating with communications limits. The study of sensor fusion and its sibling area of sensor scheduling [6, 7] has a long history in these arenas. Schweppe [23] was a pioneer in the application of such methods in power system state estimation using data of variable reliability. More generally, the study of missing data has been longstanding in statistics [15] and in array beamforming [28] with studies of estimation in high noise going back to Wiener [33], whose work was connected with the origins of radar.

Closing the loop on a system to achieve stabilization using communicated data would appear to be a more recent problem. Interestingly, Wong and Brockett study first the state estimation problem [36] and then the feedback stabilization problem [37] for systems with reliable but bandlimited communications. This extends earlier results due to Williamson [34, 35] and Delchamps [4] on finite-wordlength effects on estimation and control in deterministic contexts. Sensor scheduling is also a feedback control or decision problem with a solution achievable via dynamic programming. The unifying aspect of these earlier analyses of the estimation and the control problems is that the covariance function of the state estimate error or of the state itself is the target of the analysis. Thus, mean-square stabilization is the objective in Nair and Evans [20].

For communications problems with intermittence, the conditional covariance, $\Sigma_{k+1|k}$, of the state prediction error is a random process adapted to the $\{\gamma_k\}$ sequence. Recent works have been concerned with the distributional and moment properties of this $\{\Sigma_{k+1|k}\}$ process. Thus: Sinopoli et al. [25] consider the convergence of the expectation $\mathrm{E}[\Sigma_{k+1|k}]$ as $k \to \infty$; [11] models the packet dropping channel as the a two state Markov chain and arrives at the stabilization condition as in [25] for the scalar case;

Shi, Epstein and Murray [24] quantify the probability $\Pr\left(\Sigma_{k+1|k} > G\right)$ for given matrix G; [12] analyzes conditions for the weak (in distribution) convergence of $\Sigma_{k+1|k}$; and, [19] treats the *tail* distribution properties of this covariance. (In this latter reference 'tail' refers to the distribution on the tail σ-algebra as $k \to \infty$ and not the tail probabilities in the sense of $\Sigma_{k+1|k}$ taking on large values for any specific value of k.)

1.4.2 Background and literature survey of Bayesian filtering approach to quantized innovations

Consider the system (2.1-2.2), with the assumptions concerning gaussian initial condition and noises, operated on the transmitter-side and sending the complete innovations, $\varepsilon_k = y_k - C_k \hat{x}_{k|k-1}$ to the receiver. Denote the innovations sequence, $E_k = \{\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_k\}$. The receiver can compute the pdf by Bayesian filtering (1.9-1.13). Then, we have following formulas:

$$p(x_{k+1} \Big| E_k) = \int_{x_k} p(x_{k+1} \Big| x_k) p(x_k \Big| E_k) dx_k \tag{1.14}$$

$$p(x_{k+1} \Big| E_{k+1}) = \frac{p(\varepsilon_{k+1} \Big| x_{k+1}, E_k) p(x_{k+1} \Big| E_k)}{p(\varepsilon_{k+1} \Big| E_k)} \tag{1.15}$$

where $p(\varepsilon_{k+1} \Big| E_k) = p(\varepsilon_{k+1}) = \mathscr{N}(0, C_k \Sigma_{k|k-1} C_k^T + R_k)$ by the innovations signal properties of whiteness and zero-mean. The term $p(\varepsilon_{k+1} \Big| x_{k+1}, E_k)$ can be expressed as follows

$$\begin{aligned} p(\varepsilon_{k+1} \Big| x_{k+1}, E_k) &= p(C_{k+1}(x_{k+1} - \hat{x}_{k+1|k}) + v_{k+1} \Big| x_{k+1}, E_k) \\ &= \mathscr{N}(C_{k+1}(x_{k+1} - \hat{x}_{k+1|k}), R_{k+1}) \end{aligned} \tag{1.16}$$

The second equality is based on that fact: $\hat{x}_{k+1|k}$ is a linear combination of the history of innovations signals E_k. Then, all of the terms in (1.14-1.15) are computable. However, this recursion does not hold for quantized innovations signals as the received information because of their non-gaussian nature. The term $p(\varepsilon_{k+1} \Big| x_{k+1}, \bar{E}_k)$ is hard to compute because the quantized innovations signals contribute information for to $\hat{x}_{k+1|k}$ but in a complicated way.

The estimation problem with quantized measurements can be traced back to an early work [27]. The paper deals with a special case that the initial state density must be uniform by solving the Fokker-Planck equation and Bayes' rule. The paper [2] uses the sign of innovations (SOI) signals in the estimation problem with the Kalman filter recursion modified in the measurement update term through the replacement of the measurement noise covariance by an expression reflecting the effect of quantization — we demonstrate later that this assumption does not hold, although the method does yield both an approximate conditional mean and conditional variance. To proceed from this point with analysis, these authors are required to make an assumption that the predicted conditional pdf of the state is gaussian. The work [14] extends [2] to more complex multi-bit quantizers. Both of the above works appeal to the property that that innovations signals have smaller covariance than the output measurements. However, the work in [2] and [14] uses the Kalman filtering recursion based on quantized measurement innovations with the assumption that the prior conditional state density is Gaussian. The assumption only makes sense when the bitrate is high and saturation is absent. This assumption is questioned in [26] and the authors provide a Lemma in [22] which shows that the state given quantized innovations is the sum of two independent random variables. In [5], the authors provide another analysis to approach the exact solution but with approximate computation.

Sukhavasi and Hassibi provide the following lemma for the gaussian system above with transmitter-side measurements y_k, their quantized variants q_k and $Y_k = \{y_0, \ldots, y_k\}$, $Q_k = \{q_0, \ldots, q_k\}$.

Lemma 2 ([22]) *The state conditioned on the quantized measurements Q_k can be expressed as a sum of two independent random variables as follows*

$$x_k \Big| Q_k \sim Z_k + R_{x_k,Y_k} R_{Y_k}^{-1} [Y_k \Big| Q_k], \tag{1.17}$$

$$\textit{where} \qquad Z_k \sim \mathcal{N}\left(0, R_{x_k} - R_{x_k Y_k} R_{Y_k}^{-1} R_{Y_k x_k}\right) \tag{1.18}$$

where R_{x_k} is the covariance of x_k and $R_{x_k Y_k}$ is the cross-covariance of x_k and the previous history of measurements Y_k. The first term Z_k is a product of the Kalman filter at the transmitter. The second term is a correction term accommodating the quantization. Form the perspective of this thesis, the second term is problematic in that it is not computable in a recursive fashion. Rather, it relies on the entire measurement history ay each time.

The second paper, [5], provides the following idea. The predicted pdf $p(x_{k+1}|\bar{E}_k)$ can be computed from the traditional Bayesian filtering formula. The filtered pdf can be compute by the following formula.

$$p(x_{k+1} \Big| \bar{E}_{k+1}) = \int_{E_{k+1}} p(x_{k+1} \Big| E_{k+1}) p(E_{k+1} \Big| \bar{E}_{k+1}) dx_k \tag{1.19}$$

where $p(x_{k+1} \Big| E_{k+1})$ is a gaussian distribution function and $p(E_{k+1} \Big| \bar{E}_{k+1})$ is a truncated gaussian distribution function whose dimension increases with time, i.e. this is not a recursive formula. The integration operation in (1.19) is approximated with a recursive but approximate formula based on a mid-point approximation to the integrals. The above works either do not provide an exact solution or rely on the whole history and cannot be implemented recursively. The second part of thesis provides simple recursive formulas of Bayesian filtering for the exact computation of state estimation pdfs given quantized measurements/innovations.

1.5 Contribution and Thesis Organization

1.5.1 Contribution

The first part of thesis provides a new perspective for intermittent communication systems. We list several points of the contribution.

1. The behavior of escape time can be described as a Markov Chain process, which observation provides access to analytical tools.

2. Theorem 3 provides the algorithm to compute the exact probability distribution function of escape time when we know the $\{\gamma_k\}$ sequence of packet drops.

3. We compare the analysis of estimator escape time performance, which is applicable to both stable and unstable systems, to the conditional moment-based approaches of [25] and [24] which are of interest for unstable systems.

4. We extend the escape time analysis to include the state prediction error and the controlled system output. This invites comparison to [20].

5. Our analysis includes the treatment of quantization effects in output escape time problems and shows that bitrate choice can have a profound effect on the output escape time properties of a controlled system.

The second part of thesis focuses on the computation of the conditional probability density function of the state given quantized innovations signals. This is based on Bayesian filtering.

1. We provide an approach to calculate the exact conditional probability density functions of the predictor and of the filter.

2. We derive a recursive formulation of the pdf calculation based on the Bayesian filter and demonstrate its application.

3. The core innovation permitting the appeal to Bayesian filtering is in the identification of the appropriate underlying transmitter-side state process.

4. These new methods subsume earlier work and permit consideration of more complicated quantizers than the set of subtractive dithered quantizers used in Part I.

5. These methods, since they produce the exact state pdfs, obviate the approach to the estimation problem using particle filters.

1.5.2 Thesis Organization

The thesis is divided into two parts. The first part is organized as follows: The escape time problem formulation is presented in Chapter 2 and the concepts of escape and survival times are introduced for the estimation and stabilization problems. Chapter 3 develops the central Markov chain description in a general context before examining this for Gaussian and quantized Gaussian linear systems. These results pertain when the packet arrival sequence, $\{\gamma_k\}$, is known. Chapter 4 presents numerical examples to demonstrate: the method, heuristics of escape time, and the quality of approximations. The analysis is then extended to explore approaches to quantization and retransmission in the output stabilization escape time.

The second part of thesis is organized as follows: The problem formulation and derivation of the exact solution for the predictor and filter pdfs of the quantized innovations state estimator is in Chapter 5. It includes comparison with other approaches to dealing with this problem by others. The formal derivation of the new result here is brief, which is part of the importance of the work. Chapter 6 shows several simulations to demonstrate the implementation of the method and to compare these computed densities with the pdfs from both the Kalman filter and [2], which latter work is based on a gaussian assumption of the prior density.

Chapter 1, in part, has been submitted for publication of the material as it may appear in the introductions of the following two works. “Escape time formulation of state estimation and stabilization with quantized intermittent communication”, 2015 *Automatica,* Chun-Chia Huang, Robert R. Bitmead; “Exact formulation of quantized-innovations state estimation using Bayesian filtering,” 2015 submitted to IEEE transaction on signal processing, Chun-Chia Huang, Robert R. Bitmead.

Part I

Escape time formulation of state estimation and stabilization with quantized intermittent communication

Chapter 2

Problem formulation

2.1 Assumptions

We commence with the communications-linked control system as follows

$$x_{k+1} = Ax_k + Bu_k + w_k, \tag{2.1}$$

$$y_k = Cx_k + v_k, \tag{2.2}$$

$$z_k = \gamma_k \mathscr{Q}_d(y_k). \tag{2.3}$$

with the same description in Section 1.1 and make the following assumptions.

Assumption 1 *1.1* x_1*, the initial state, is Gaussian with mean* $\bar{x}_1$*, and covariance* $\Sigma_{1|0}$.

1.2 $u_k = -K\hat{x}_{k|k}$ *constant linear state filter feedback is applied.*

1.3 w_k *and* v_k *are independent, white, zero-mean Gaussian processes independent from* x_1 *and with covariance matrices Q and R, which we take to be positive definite.*

1.4 $\{\gamma_k\}$ *is a Bernoulli random process independent from* x_1, $\{w_k\}$*, and* $\{v_k\}$ *and taking values 0 or 1 to describe the non-arrival or arrival of a data packet, respectively.*

1.5 $\mathscr{Q}_d$ *is a subtractive, dithered, b-bit-per-channel, mid-rise, symmetric, linear quantizer with saturation values* $\pm\zeta$*. The subtractive dither signal is white, triangularly* *distributed,* $\mathrm{tr}\left(-\zeta/2^{b-1}, \zeta/2^{b-1}\right)$*, and known exactly to both transmitter and receiver.*

A number of extensions are possible at the expense of clarity of development. (i) Time-variation can be included into the noise covariance matrices and into the feedback control gain. Stationarity of the problem does permit the calculation of expectations which converge over time as in [25]. However given Lemma 1, our focus is not on limiting behaviors but explicitly on transient properties, where time-variation is readily accommodated. (ii) In Assumption 1.4 we assume the process $\{\gamma_k\}$ is Bernoulli, as in [25]. It might equally well be taken as Markov as in [12]. (iii) Extension to nonlinear forms of quantization is straightforward. Extension to adaptive quantizers, while it is at the heart of Nair's and Evans' [20] demonstration of the minimal feedback bitrate required for mean-square stabilization, comes at the cost of communications and is not part of our analysis.

2.2 Incorporating quantization

Subtractive dithered b-bit quantizer, $\mathscr{Q}_d(\cdot)$, is depicted in Figure 2.1. It includes the standard quantizer $\mathfrak{Q}$ and subtractive dither signal, d_k.

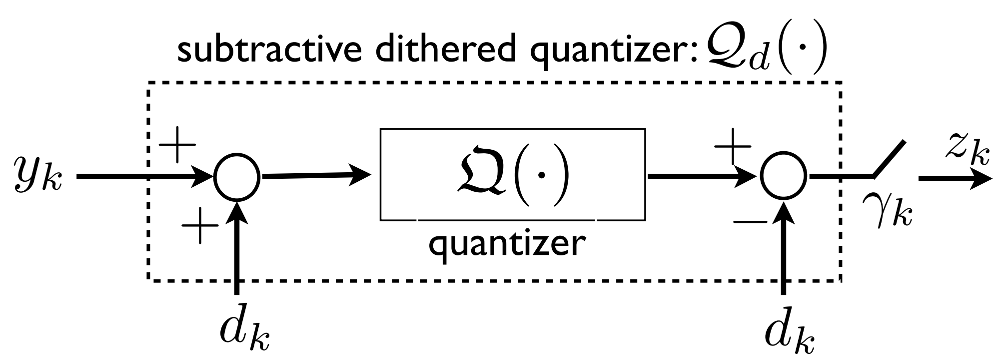

Figure 2.1. Subtractive dithered quantization including intermittent gain γ_k.

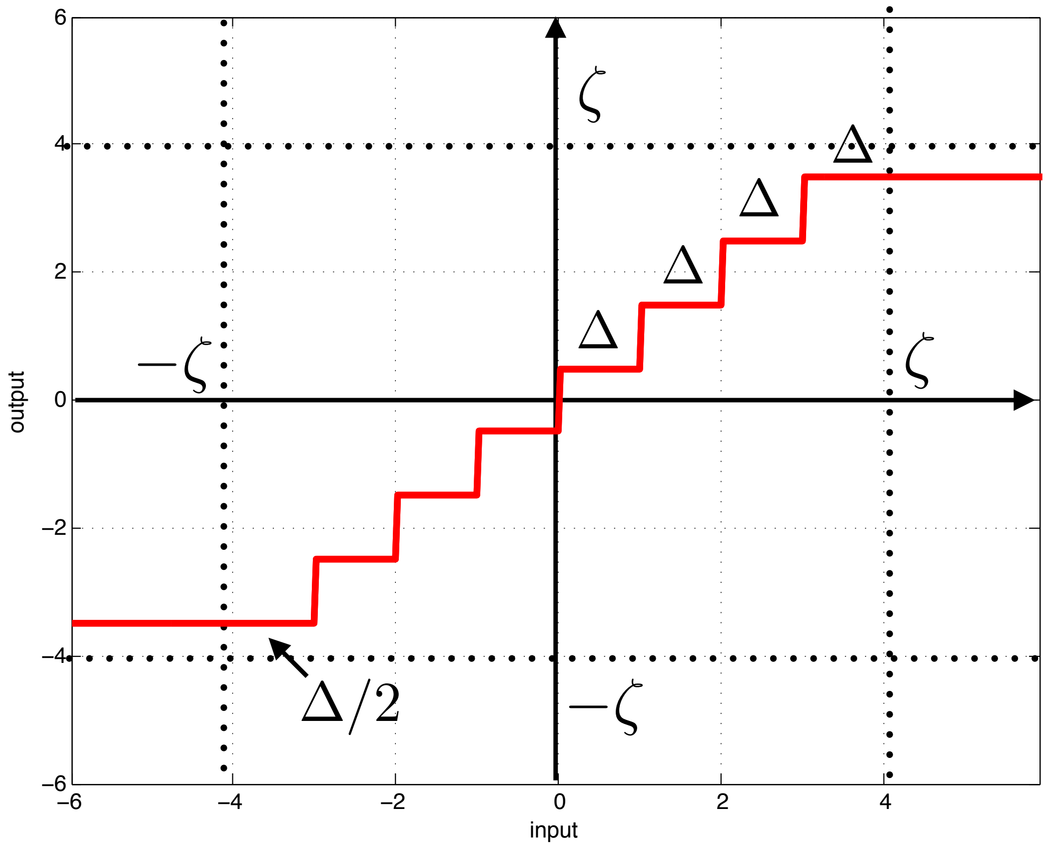

Figure 2.2. The linear quantizer $\mathfrak{Q}$

The fixed quantizer function $\mathfrak{Q}(\cdot)$ is a mid-rise, symmetric, linear quantizer with 2^b levels within the range $[-\zeta,\zeta]$. Figure 2.2 shows a 3-bit quantizer $\mathfrak{Q}$, the basic unit of the quantizer is $\Delta = \zeta/2^{b-1}$. We note that subtractive dithered quantization requires careful handling of the wordlength of the subtraction operation, since d_k is recorded at a higher bitrate than b bits, see [32].

Theorem 1 *Under the conditions of Assumption 1.5, and provided $y_k \in [-\zeta,\zeta]$, the quantization noise,*

$$n_k = \mathcal{Q}_d[y_k] - y_k = [\mathfrak{Q}(y_k + d_k) - d_k] - y_k, \tag{2.4}$$

is white, independent from $\{y_k\}$, *and uniformly distributed* $U(-\zeta/2^b, \zeta/2^b)$.

The independence and uniform distribution of the theorem statement follow directly from Theorem QTSD of [32]. The whiteness of $\{n_k\}$ follows from the whiteness of $\{d_k\}$.

The effective measurement noise in z_k is the sum of two independent terms, one Gaussian (measurement noise) and the other uniform (quantization noise).

$$v_{\text{eff},k} = v_k + n_k, \tag{2.5}$$

In our subsequent analysis, we show how to compute the escape times with any measurement noise process but for clarity specialize to either a strictly Gaussian measurement noise or a Gaussian-plus-uniform measurement noise. With independent m-vectors $v_k \sim N(0_m, R)$ and $n_k \sim U^m(-\zeta/2^b, \zeta/2^b)$ the pdf of $v_k + n_k$ is given by

$$f_{v+n}(z) = \frac{2^{b-1}}{\zeta}\left\{\Phi\left[R^{-1/2}\left(z+\frac{\zeta}{2^b}\right)\right] - \Phi\left[R^{-1/2}\left(z-\frac{\zeta}{2^b}\right)\right]\right\}, \tag{2.6}$$

where $\Phi(\cdot)$ is the multivariate standard normal cdf of appropriate dimension. This is the convolution of the Gaussian pdf of v_k with the uniform pdf of n_k.

When a quantizer is present and for $y_k \in [-\zeta, \zeta]$, the effective measurement noise covariance for z_k increases from R with perfect reconstruction of y_k to

$$R_{\text{eff}} = R + \frac{\zeta^2}{3 \times 2^{2b}} I_m. \tag{2.7}$$

which is the measurement noise, v_k, covariance plus the quantizer noise, n_k, since these two noises are additive in (2.5) and are independent. With quantization, this quantity replaces R in the Kalman filter recursion, which relies solely on second-order statistics and yields the least mean squares linear unbiased estimator.

Outside of the range $[-\zeta, \zeta]$ the quantization error becomes potentially unbounded and correlated. In our situation of escape time analysis with fixed quantization, once the outer levels are breached the state or state estimation error has escaped. Nair and Evans in [20] develop an ingenious approach to adaptive quantization amenable to mitigating the effects of saturation errors by expanding ζ faster than the state can escape and, further, to achieving this within the allocated bitrate. In 4.3, strategies are studied to mitigate the effects of intermittency through the allocation of a fraction of the b bits in the communicated signal to the coarse retransmission of earlier values of y_k. These studies extend earlier studies of signal reconstruction with missing data [9] and can deal with strategies for bit assignment in both TCP/IP and UMDP settings depending on the transmitter's knowledge of the arrival of packets.

2.3 Kalman state estimation with intermittent quantized observations

The Kalman state estimation equations with intermittent quantized observations described by $\{\gamma_k\}$ and $\mathscr{Q}$, commencing from initial values $\hat{x}_{1|0} = \bar{x}_1$ and $\Sigma_{1|0}$, are as follows. These equations are identical to those of [25, 24] and subsequent works modulo the incorporation of quantization via R_{eff} and n_k. It is similar to the formulas of Section 1.2

except the γ_k and R_{eff}.

measurement update

$$L_k = \gamma_k \left[\Sigma_{k|k-1} C^T (C\Sigma_{k|k-1}C^T + R_{\text{eff}})^{-1} \right], \tag{2.8}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k(z_k - C\hat{x}_{k|k-1}),$$

$$\Sigma_{k|k} = \Sigma_{k|k-1} - L_k C \Sigma_{k|k-1}. \tag{2.9}$$

time update

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k,$$

$$\Sigma_{k+1|k} = A\Sigma_{k|k}A^T + Q, \tag{2.10}$$

filtered error

$$\tilde{x}_{k|k} = x_k - \hat{x}_{k|k},$$

prediction error

$$\tilde{x}_{k+1|k} = x_{k+1} - \hat{x}_{k+1|k},$$

$$\tilde{x}_{k+1|k} = (A - AL_kC)\tilde{x}_{k|k-1} + w_k - AL_k(v_k + n_k). \tag{2.11}$$

Note the γ_k appears explicitly solely in the Kalman gain L_k and that, conditioned on $\{\gamma_k\}$, the state prediction error system (2.11) is linear, time-varying due to γ_k, and driven by not-necessarily-Gaussian white noise. The Kalman predictor error equation (2.11) is central to both the estimator and the output escape time formulations below. We note that, when $\gamma_k = 0$, then $L_k = 0$ and (2.11) will be unstable if A has eigenvalues outside the unit circle and driven by w_k alone. If $\gamma_k = 0$ for all k — a case where we refer to the escape time as the *survival time* — (2.11) is time-invariant, which simplifies the analysis.

2.4 Output feedback control with intermittent and quantized observations

The output feedback control problem with intermittent and quantized observations differs from the estimator problem since it involves both the filter error, $\tilde{x}_{k|k}$, and the controlled plant state, x_k. It is described by the linear system:

$$\begin{bmatrix} x_{k+1} \\ \tilde{x}_{k+1|k} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} A-BK & BK(I-L_kC) & 0 \\ 0 & A(I-L_kC) & 0 \\ C(A-BK) & CBK(I-L_kC) & 0 \end{bmatrix} \begin{bmatrix} x_k \\ \tilde{x}_{k|k-1} \\ y_k \end{bmatrix} + \begin{bmatrix} -BKL_k(v_k+n_k)+w_k \\ -AL_k(v_k+n_k)+w_k \\ -CBKL_k(v_k+n_k)+Cw_k+v_{k+1} \end{bmatrix}, \quad (2.12)$$

This system's state covariance matrix is shown as follows,

$$cov\left(\begin{bmatrix} x_k \\ \tilde{x}_{k|k-1} \\ y_k \end{bmatrix}\right) = \begin{bmatrix} P_k & \Sigma_{k|k-1} & P_kC^T \\ \Sigma_{k|k-1} & \Sigma_{k|k-1} & \Sigma_{k|k-1}C^T \\ CP_k & C\Sigma_{k|k-1} & CP_kC^T+R_{\text{eff}} \end{bmatrix} \quad (2.13)$$

The Kalman filtering recursion for intermittent quantized data, (2.8), (2.9), (2.10), yields the controlled state covariance, P_k, recursion

$$\begin{aligned} P_{k+1} = {} & (A-BK)P_k(A-BK)^T + (A-BK)\Sigma_{k|k-1}(I-L_kC)^TK^TB^T \\ & + BK(I-L_kC)\Sigma_{k|k-1}(A-BK)^T + BK(I-L_kC)\Sigma_{k|k-1}(I-L_kC)^TK^TB^T \\ & + BKL_kR_{\text{eff}}L_k^TK^TB^T + Q. \end{aligned} \quad (2.14)$$

The effect of increasing the measurement noise through quantization is two-fold; the error covariance, $\Sigma_{k+1|k}$, increases and then, as a consequence, the controlled state covariance, P_k, also increases. Additionallly, R_{eff} directly drives (2.14). During periods of packet loss, measurement noise and quantization noise do not affect the system directly, since $L_k = 0$. At times of packet arrival however, the quantization noise effect in increasing the underlying state estimate covariance and controlled state covariance is evident.

2.5 Escape times, conditional escape times and survival times

We have established that for output signals, y_k, within the quantization range $[-\zeta, \zeta]$, both the estimator error and the output feedback controlled system are governed by linear systems driven by white, not-necessarily-Gaussian noises: {(2.8), (2.9), (2.10), (2.11)} for the estimator; and additionally (2.12) for the controlled system. This will form the basis for our calculation of escape times.

We make the following definitions.

Definition 2 (Estimator escape time τ_e) *For given positive scalar bound G, we define the estimator escape time to be the escape time for the random sequence $\{\tilde{x}_{k+1|k}\}$ starting from the initial condition $\tilde{x}_{1|0} \sim N(0, \Sigma_{1|0})$ and with $\mathscr{D} = \{\left\|\tilde{x}_{k+1|k}\right\|_\infty \leq G\}$. The estimator escape time cdf is the cdf of τ_e given these initial conditions.*

Definition 3 (Output escape time τ_o) *Given the output quantizer magnitude upper bound, ζ from Assumption 1.5, and given initial state covariance matrix P_1 and state estimate error covariance matrix $\Sigma_{1|0}$, we define the output escape time to be the escape time from $\mathscr{D} = \{\|y_k\|_\infty \leq \zeta\}$ from initial condition satisfying $y_1 \sim N(0, CP_1C^T + R_{\text{eff}})$.*

The estimator and output escape times are random variables depending on: the bound ζ, the initial covariances $\Sigma_{1|0}$ and P_1, the realizations of the noise processes x_1, $\{w_k\}$, $\{v_k + n_k\}$, and the realization of $\{\gamma_k\}$, the sequence of packet arrival successes or failures. Our aim is to compute and to characterize the probability cumulative distribution functions (cdfs) of these escape times as the communication link's arrival probability, P_γ, changes. The results of [25] establish a lower bound on P_γ for the expected value of the conditional estimator covariance to be finite for all time, i.e. $\mathrm{E}[\Sigma_{k|k}] < \infty$, which can occur even though the estimate error itself escapes any bounded domain with probability one.

The inclusion of the output escape time problem differs from the analysis of [25, 24], which considers estimator behavior only, and brings us more into contact with works such as [20] which study stochastic stabilization. We note that for the state-estimate-feedback controlled system, the covariance of output y_k depends on the covariance of the controlled state x_k, which in turn depends on the covariance of the state estimation error, $\tilde{x}_{k|k}$.

For specific γ_k sequences, we introduce the following definition:

Definition 4 (Conditional escape time) *For bounded domain $\mathscr{D}$ and stochastic process $\{\xi_k\}$ adapted to the sequence $\{\gamma_k\}$, the conditional escape time cdf, $\Phi_e(k)$, is the escape time cdf from $\mathscr{D}$ for a given sequence, $\{\gamma_k\}$, of successful and dropped communications packets.*

In particular, we shall identify a special set of conditional escape times associated with unsuccessful communication from time 1 onwards.

Definition 5 (Survival time) *We define the survival time cdf for $\{\xi_k\}$, $\Phi_s(k)$, as the conditional escape time cdf for the domain $\mathscr{D}$ with sequence $\{\gamma_k = 0, k = 1, 2, \ldots\}$.*

Chapter 2, in full, is a reprint of the material as it appears in the section 2 of the work, "Escape time formulation of state estimation and stabilization with quantized intermittent communication", 2015 *Automatica,* Chun-Chia Huang, Robert R. Bitmead.

Chapter 3

EscapeTime Markov Chain Analysis

This section provides the fundamental theoretical support for these calculations, which later will be applied to examples.

3.1 General Markov Chain model

Following Definition 1, given a closed domain $\mathscr{D} \subset \mathbb{R}^d$ and a stochastic process $\{\xi_k : k = 1, \ldots\}$ on $\mathbb{R}^d$, the escape time of ξ_k from $\mathscr{D}$ is described by a Markov chain. This is a general result concerning adapted processes and is not limited to linear Gaussian systems nor to hyperspherical domains.

Theorem 2 (Principal theoretical result [10]) *For stochastic process* $\{\xi_k : k = 1, \ldots\}$ *the random variable*

$$J_{k+1} = \begin{cases} 1, & \textit{if } \xi_k \in \mathscr{D} \textit{ and } J_k = 1, \\ 0, & \textit{otherwise,} \end{cases} \tag{3.1}$$

is a Markov process and, denoting

$$\Pi_k = \begin{bmatrix} \Pr(J_k = 1) \\ \Pr(J_k = 0) \end{bmatrix},$$

is described by the Markov chain.

$$\Pi_{k+1} = \begin{bmatrix} \alpha_k & 0 \\ 1-\alpha_k & 1 \end{bmatrix} \Pi_k,$$

where

$$\alpha_k = \Pr(J_{k+1} \Big| J_k = 1) \tag{3.2}$$

The escape time of the stochastic process $\{\xi_k\}$ *is the value of* k *when* $J_k = 0$ *for the first time.*

Proof: From the definition (3.1), J_k satisfies the Markov property.

$$\Pr(J_{k+1} \,\Big|\, J_k, \ldots J_1) = \Pr(J_{k+1} \,\Big|\, J_k),$$

and so $\{J_k\}$ is described by a Markov chain. The transition matrix of this chain is given by

$$\begin{bmatrix} \Pr(J_{k+1} = 1 | J_k = 1) & \Pr(J_{k+1} = 1 | J_k = 0) \\ \Pr(J_{k+1} = 0 | J_k = 1) & \Pr(J_{k+1} = 0 | J_k = 0) \end{bmatrix}$$

The (1,2)-element of this matrix is zero, because there is no possibility of moving from $J_k = 0$ to $J_{k+1} = 1$.

Corollary 1 *For a given* $\{\gamma_k\}$ *sequence, the corresponding* $\{\alpha_k\}$ *sequence describes the* *evolution of the conditional escape time cdf,* $\Phi_e(k)$*,*

$$\Pr(J_{k+1}=0) = (1-\alpha_k)\Pr(J_k=1) + \Pr(J_k=0),$$
$$\Phi_e(k+1) = (1-\alpha_k)[1-\Phi_e(k)] + \Phi_e(k),$$

or,

$$\Phi_e(k+1) = (1-\alpha_k) + \alpha_k \Phi_e(k), \tag{3.3}$$

whence

$$\Phi_e(k) = 1 - \prod_{i=0}^{k-1} \alpha_i. \tag{3.4}$$

Proof: The final result (3.4) follows by rewriting (3.3) as

$$[1-\Phi_e(k+1)] = \alpha_k[1-\Phi_e(k)],$$

and recognizing that $1-\Phi_e(1) = \alpha_0$. ■

3.2 Computation of α_k for linear systems

As developed in Theorem 2, the escape time of stochastic process $\{x_k\}$ is governed by a Markov chain and the probability transition matrix is determined by a sole scalar value α_k. For any escape time problem, α_k might be computed as

$$\alpha_k = \frac{\Pr(x_{k+1} \in \mathscr{D}, x_k \in \mathscr{D}, \ldots, x_1 \in \mathscr{D})}{\Pr(x_k \in \mathscr{D}, \ldots, x_1 \in \mathscr{D})}.$$

This requires the computation of a $(k+1)n$-dimensional and a kn-dimensional multivariate Gaussian cdf, as could in the Gaussian case in principle be carried out using `mvncdf` in `matlab`. However, this is problematic due to the growth of dimension of the argument. We now present an alternative recursion for the computation of α_k for linear systems.

Theorem 3 *Consider the time-varying stochastic linear system*

$$\xi_{k+1} = F_k \xi_k + \omega_k, \tag{3.5}$$

with: ξ_1 possessing a pdf, $f_{\xi_1}(z)$, and $\{\omega_k\}$, independent from ξ_1, white and with pdfs $f_{\omega_k}(z)$.
Take

$$h_0(z) = f_{\xi_1}(z), \tag{3.6}$$

and define the sequence of (pdf) functions $\{h_k : \mathbb{R}^n \to \mathbb{R}_+\}$

$$h_k(z) = \frac{1}{\int_{\mathscr{D}} h_{k-1}(w)\,dw} \int_{\mathscr{D}} f_{\omega_k}(z - F_k w) h_{k-1}(w)\,dw. \tag{3.7}$$

Then

$$\alpha_k = \int_{\mathscr{D}} h_k(w)\,dw. \tag{3.8}$$

Proof: The proof hinges on the careful application of the Markov property. Denote by p the general pdf where the argument specifies the pdf and define

$$h_k(\xi_{k+1}) = p(\xi_{k+1} \Big| J_k = 1) \tag{3.9}$$

$$= \frac{p(\xi_{k+1}, J_k = 1)}{\Pr(J_k = 1)}. \tag{3.10}$$

Now note that the numerator

$$\begin{aligned} p(\xi_{k+1}, J_k = 1) &= \int_{\xi_k \in \mathscr{D}} p(\xi_{k+1} \Big| \xi_k, J_{k-1} = 1) p(\xi_k, J_{k-1} = 1)\, d\xi_k, \\ &= \int_{\xi_k \in \mathscr{D}} p(\xi_{k+1} \Big| \xi_k) p(\xi_k, J_{k-1} = 1)\, d\xi_k, \\ &= \Pr(J_{k-1} = 1) \int_{\xi_k \in \mathscr{D}} p(\xi_{k+1} \Big| \xi_k) p(\xi_k \Big| J_{k-1} = 1)\, d\xi_k, \\ &= \Pr(J_{k-1} = 1) \int_{\xi_k \in \mathscr{D}} f_{\omega_k}(\xi_{k+1} - F_k \xi_k) h_{k-1}(\xi_k)\, d\xi_k, \end{aligned} \tag{3.11}$$

where f_{ω_k} is the probability density function of ω_k. [The Markov property is invoked in replacing $p(\xi_{k+1} \Big| \xi_k, J_{k-1} = 1)$ by $p(\xi_{k+1} \Big| \xi_k)$.] Substituting (3.11) into (3.10) yields

$$\begin{aligned} h_k(\xi_{k+1}) &= \frac{\Pr(J_{k-1} = 1)}{\Pr(J_k = 1)} \int_{\xi_k \in \mathscr{D}} f_{\omega_k}(\xi_{k+1} - F_k \xi_k) h_{k-1}(\xi_k)\, d\xi_k, \\ &= \frac{1}{\alpha_{k-1}} \int_{\xi_k \in \mathscr{D}} f_{\omega_k}(\xi_{k+1} - F_k \xi_k) h_{k-1}(\xi_k)\, d\xi_k. \end{aligned}$$

Finally, integrating (3.9) yields

$$\begin{aligned} \alpha_k = \Pr(\xi_{k+1} \in \mathscr{D} | J_k = 1) &= \int_{\xi_{k+1} \in \mathscr{D}} p(\xi_{k+1} \Big| J_k = 1)\, d\xi_{k+1}, \\ &= \int_{\xi_{k+1} \in \mathscr{D}} h_k(\xi_{k+1})\, d\xi_{k+1}. \end{aligned}$$

Theroem 3 permits calculation of $\{\alpha_k\}$ for (3.5) for any process noise pdfs, even time-varying. In the sequel for comparison with earlier results, we shall limit investigation to either Gaussian or Gaussian-plus-uniform noise

Corollary 2 *For linear Gaussian system* (3.5) *with initial state* $\xi_1 \sim N(0,\rho_1)$ *with* $\rho_1 > 0$*, and* $\omega_k \sim N(0,\Omega_k)$*, white and independent from* ξ_1*,* α_k *is given by the recursion (3.6-3.8) with*

$$
\begin{aligned}
f_{\xi_1}(z) &= \frac{1}{(2\pi)^{n/2}|\rho_1|^{1/2}} \exp\left(-\frac{1}{2}z^t\rho_1^{-1}z\right), \\
f_{\omega_k}(z-F_kw) &= \frac{1}{(2\pi)^{n/2}|\Omega_k|^{1/2}} \exp\left[-\frac{1}{2}(z-F_kw)^T\Omega_k^{-1}(z-F_kw)\right].
\end{aligned} \tag{3.12}
$$

For the same system and initial state but with $\omega_k = v_k + n_k$ *where* $v_k \sim N(0,R_k)$ *and* $n_k \sim U\left(-\zeta/2^b, \zeta/2^b\right)$ *with* $\{v_k\}$ *and* $\{n_k\}$ *independent and white,* α_k *is given the recursion (3.6-3.8) with* $f_{\xi_1}(z)$ *as above and* $f_{\omega_k}(z)$ *given by* $f_{v+n}(z)$ *from* (2.6).

Corollary 3 *For linear n-dimension Gaussian system* (3.5) *with any sets as* $\mathscr{D}$ *shown below*

$$
\mathscr{D} = \bigcap_{i=1}^{n}\{||\xi_k(i)||_\infty \leq G_i\},\ G_i \geq 0
$$

$\xi_k(i)$ *is i-th element of* ξ_k

the probability quantities $\Pr(\xi_{k+1} \in \mathscr{D} \Big| \xi_k \in \mathscr{D}, \ldots, J_{i-1} = 1)$ *and* $\Pr(\xi_{k+1} \in \mathscr{D} \Big| \xi_k \in \mathscr{D}, \ldots, J_{i-1} = 0)$ *have an lower bound* $\Pr(\xi_{k+1} \in \mathscr{D} \Big| \xi_k$ *is a zero vector*$)$.

Proof:

$$\begin{aligned}
&\Pr(\xi_{k+1} \in \mathscr{D} \Big| \xi_k \in \mathscr{D}, \ldots, \xi_i \in \mathscr{D}, J_{i-1} = 1) \\
&= \frac{\int_{\xi_{k+1} \in \mathscr{D}} \cdots \int_{\xi_i \in \mathscr{D}} p(\xi_{k+1} \Big| \xi_k) \ldots p(\xi_i \Big| J_{i-1} = 1) d\xi_i \ldots d\xi_{k+1}}{\int_{\xi_k \in \mathscr{D}} \cdots \int_{\xi_i \in \mathscr{D}} p(\xi_k \Big| \xi_{k-1}) \ldots p(\xi_i \Big| J_{i-1} = 1) d\xi_i \ldots d\xi_k}
\end{aligned} \tag{3.13}$$

Denote

$$G(\xi_i) = \int_{\xi_k \in \mathscr{D}} \cdots \int_{\xi_{i+1} \in \mathscr{D}} p(\xi_k \Big| \xi_{k-1}) \ldots p(\xi_{i+1} \Big| \xi_i) d\xi_{i+1} \ldots d\xi_k$$

By Fubini's theorem, we can change the order of integration in (3.13). Then, we have

$$\begin{aligned}
&\Pr(\xi_{k+1} \in \mathscr{D} \Big| \xi_k \in \mathscr{D} \ldots \xi_{i+1} \in \mathscr{D}, J_{i-1} = 1) \\
&= \frac{\int_{\xi_i \in \mathscr{D}} \int_{\xi_{k+1} \in \mathscr{D}} p(\xi_{k+1} \Big| \xi_k) G(\xi_i) p(\xi_i \Big| J_{i-1} = 1) d\xi_{k+1} d\xi_i}{\int_{\xi_i \in \mathscr{D}} G(\xi_i) p(\xi_i \Big| J_{i-1} = 1) d\xi_i}
\end{aligned} \tag{3.14}$$

For the upper bound probability, we have

$$\begin{aligned}
&\Pr(\xi_{k+1} \in \mathscr{D} \Big| \xi_k \text{ zero vector}) \\
&= \int_{\xi_{k+1} \in \mathscr{D}} p(\xi_{k+1} \Big| \xi_k \text{ zero vector}) d\xi_{k+1} \\
&= \int_{\xi_{k+1} \in \mathscr{D}} p(\xi_{k+1} \Big| \xi_k \text{ zero vector}) d\xi_{k+1} \times \frac{\int_{\xi_i \in \mathscr{D}} G(\xi_i) p(\xi_i \Big| J_{i-1} = 1) d\xi_i}{\int_{\xi_i \in \mathscr{D}} G(\xi_i) p(\xi_i \Big| J_{i-1} = 1) d\xi_i} \\
&= \frac{\int_{\xi_i \in \mathscr{D}} \int_{\xi_{k+1} \in \mathscr{D}} p(\xi_{k+1} \Big| \xi_k \text{ zero vector }) G(\xi_i) p(\xi_i \Big| J_{i-1} = 1) d\xi_{k+1} d\xi_i}{\int_{\xi_i \in \mathscr{D}} G(\xi_i) p(\xi_i \Big| J_{i-1} = 1) d\xi_i}
\end{aligned} \tag{3.15}$$

Compare (3.14) and (3.15), if we can prove

$$\int_{\xi_{k+1}\in\mathscr{D}} p(\xi_{k+1}\Big|\xi_k \text{ zero vector})d\xi_{k+1} \geq \int_{\xi_{k+1}\in\mathscr{D}} p(\xi_{k+1}\Big|\xi_k)d\xi_{k+1}, \forall\xi_{k+1}\in\mathscr{D} \quad (3.16)$$

,we are done. $p(\xi_{k+1}\Big|\xi_k) = f_{\omega_k}(\omega_k = \xi_{k+1} - F_k\xi_k)$ is a normal distribution and $p(\xi_{k+1}\Big|\xi_k \text{ zero vector}) = f_{\omega_k}(\omega_k = \xi_{k+1})$ is also a normal distriubtion. Since

$$\int_{\xi_{k+1}\in\mathscr{D}} p(\xi_{k+1}\Big|\xi_k \text{ zero vector})d\xi_{k+1} = \int_{\xi_{k+1}\in\mathscr{D}} f_{\omega_k}(\omega_k = \xi_{k+1})d\xi_{k+1}$$

is the area whose center is origin. Then, we can conclude inequality (3.16) is correct. ■

In both the estimator and the output escape problems, a given $\{\gamma_k\}$ sequence and initial state estimate covariance determines the associated Kalman gain sequence $\{L_k\}$, which captures the time-variability of the problem. Both cases can be subsumed into the study of the linear system (2.12) with the corresponding escape set. For example, the corresponding escape set of output escape time $\mathscr{D}$ and initial covariance $\mathscr{C}$:

$$\mathscr{D} = \left\{ \begin{bmatrix} x_k \\ \tilde{x}_{k|k-1} \\ y_k \end{bmatrix} : ||y_k||_\infty \leq \zeta \right\}; \mathscr{C} = \begin{bmatrix} P_1 \\ \Sigma_{1|0} \\ CP_1C^T + R \end{bmatrix} \quad (3.17)$$

The escape set of estimation escape time uses $\{||\tilde{x}_{k|k-1}||_\infty \leq G\}$ instead of $\{||y_k||_\infty \leq \zeta\}$ in above equation.

For the estimator or for the output, the results of this section permit the calculation of the conditional escape time cdf for any given packet arrival sequence, $\{\gamma_k\}$. Since in this case, the system matrices, F_k, are known.

With $\gamma_k = 0, k = 1, 2, \ldots$, the survival time analysis for both the estimator escape and for the output escape with linear systems is a time-invariant analysis, since the Kalman gain $L_k = 0$. Accordingly, the study of survival time can be explicitly conducted, which provides insight into the processes underlying escape times, which usually require simulation for evaluation.

3.3 Properties of output escape time for linear time-invariant Gaussian systems

Denote by $\Phi_e(t)$ the cdf of the output escape time for the linear time-invariant Gaussian system

$$x_{k+1} = Ax_k + Bu_k + w_k, \tag{3.18}$$

$$y_k = C_k + v_k, \tag{3.19}$$

with

$$u_k = -K\hat{x}_{k|k};\ x_1 \sim N(0, P_1);$$
$$w_k \sim N(0, Q);\ v_k \sim N(0, R).$$

By Lemma 1 since $\mathscr{D} = \{\|y_k\|_\infty \leq \zeta\}$ is bounded, $\Phi_e(t) \to 1$ as $t \to \infty$. For the same γ-sequence (which means the same F_k, ω_k in Theorem 3), Φ_e is a function of t, P_1, Q, R and ζ. From the Theorem 3, we further have the following properties holding for all $t > 0$ where the superscript delineates each of two different cases:

$$\{P_1^1 \geq P_1^2\} \cup \{Q^1 \geq Q^2\} \cup \{R^1 \geq R^2\} \cup \{\zeta^1 \leq \zeta^2\}$$
$$\Longrightarrow \Phi_e(t, P_1^1, Q^1, R^1, \zeta^1) \geq \Phi_e(t, P_1^2, Q^2, R^2, \zeta^2), \tag{3.20}$$

Proof: Consider the time-varying, linear, gaussian system (3.18)(3.19) written as (2.12) with zero-mean initial state.

$$\xi_{k+1} = F_k \xi_k + \chi_k$$

with $\xi_0 \sim \mathscr{N}(0, P_0)$ and $\chi_k \sim \mathscr{N}(0, \upsilon_k)$. Further, consider escape from the region $\mathscr{D} = \{|D\xi_k| \leq g\}$ for some constant matrix D. At time t, escape of this system coincides with

$$\Xi_t = \begin{bmatrix} \xi_0 \\ \xi_1 \\ \vdots \\ \xi_t \end{bmatrix} \notin \mathscr{D}^{t+1} = \overbrace{\mathscr{D} \times \mathscr{D} \times \cdots \times \mathscr{D}}^{t+1 \text{ times}},$$

and Ξ_t is gaussian and zero mean. Now, if either $\text{cov}(\xi_0)$ increases or $\text{cov}(\chi_k)$ increases then the covariance of ξ_t increases and escape from the set $\mathscr{D}^t$ as defined above increases. By the same token, if the parameter g decreases then the escape probability increases. This proves (3.20).

Theorem 4 *For the linear time-invariant Gaussian system (3.18-3.19), the cdf of the survival time, Φ_s, admits the following implication*

$$P_1^1 \geq A^n P_1^2 A^{n^T} + A^{n-1} Q A^{n-1^T} + \cdots + AQA^T + Q \tag{3.21}$$

$$\implies \Phi_s(t, P_1^1, Q, R, \zeta) \geq \Phi_s(t+n, P_1^2, Q, R, \zeta). \tag{3.22}$$

Theorem 4 will be used when studying the effect of quantization errors on the escape time.

Chapter 3, in full, is a reprint of the material as it appears in the section 3 of the work, “Escape time formulation of state estimation and stabilization with quantized

intermittent communication", 2015 *Automatica,* Chun-Chia Huang, Robert R. Bitmead.

Chapter 4

Escape time computational examples

4.1 Example 1: Estimator survival time

Recall that the survival time is the escape time when no data packets are received at times $k = 1, 2, \ldots$. Because Theorem 3 provides an explicit formula for α_k, we study first the survival time of the estimator. This example, while simple, permits revealing comparisons between the escape/survival time analysis and the moment-based methods of [25, 24].

Consider the system

$$\begin{aligned} &A = 1.5, B = 1, C = 1, Q = 0.005, R = 0.001, \zeta = 10, \\ &\text{with } \Sigma_{1|0} = \bar{M} = AR_{\text{eff}}A^T + Q = 0.00725, \end{aligned}$$

which corresponds to the largest possible conditional prediction covariance immediately after receiving a single data sample. We compute the survival time cdf using Theorems 2 and 3 which is a time-invariant system transient analysis yielding explicitly the survival cdf. The probability density function is plotted in Figure 4.1.

The calculation of the density function of τ_s was verified, needlessly, by simulation. Figure 4.2 displays three sample $\tilde{x}_{k+1|k}$ trajectories of differing survival times: 8,

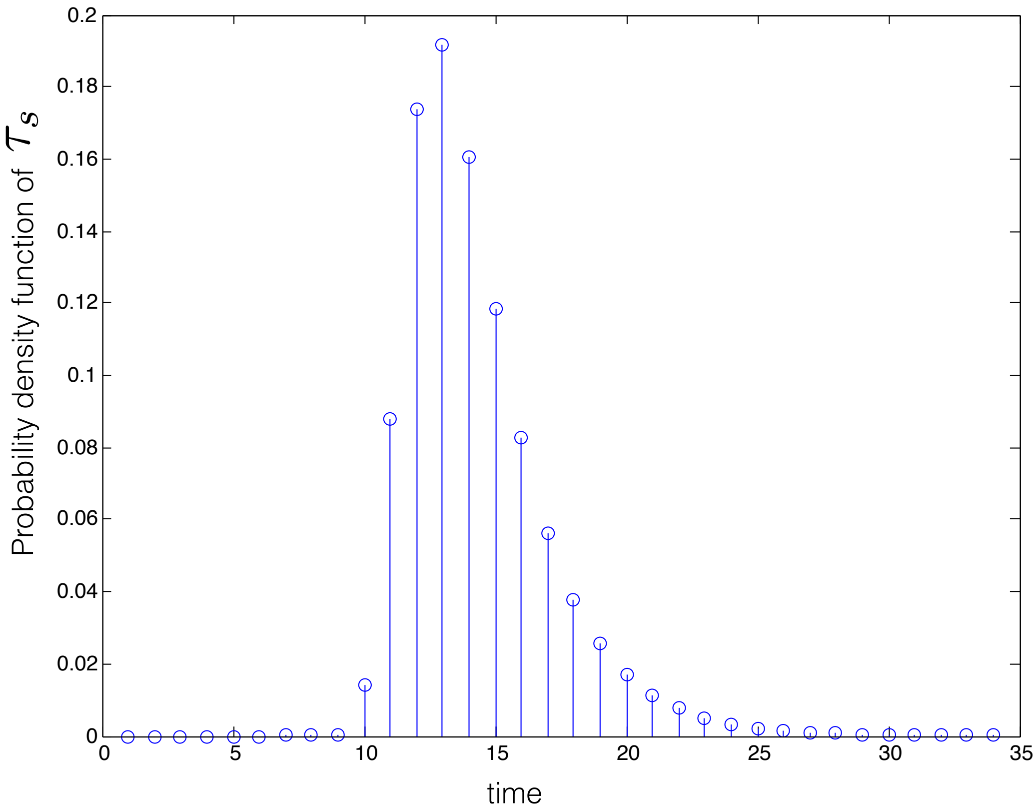

Figure 4.1. Probability density function of estimator survival time for Example 1.

12, and 19, from a Monte Carlo simulation with initial covariance $\Sigma_{1|0} = 0.00725$. We have selected three trajectories which escaped with positive state-estimate error values. From the problem description, the state-estimate error satisfies,

$$\tilde{x}_{k+1|k} = (1.5)^k \tilde{x}_{1|0} + \sum_{j=1}^{k} (1.5)^{k-j} w_j,$$

which displays the relative importance of the random terms. Since $w_k \sim N(0, 0.005)$, once the state exceeds roughly 0.2 (three σ) in magnitude it becomes most unlikely that w_k will arise to bring the error back to a small magnitude. Once the three trajectories exceed this value, we see that they all escape with roughly similar behavior, i.e. at a rate of 1.5^k. The difference between the trajectories lies in their residence time in the

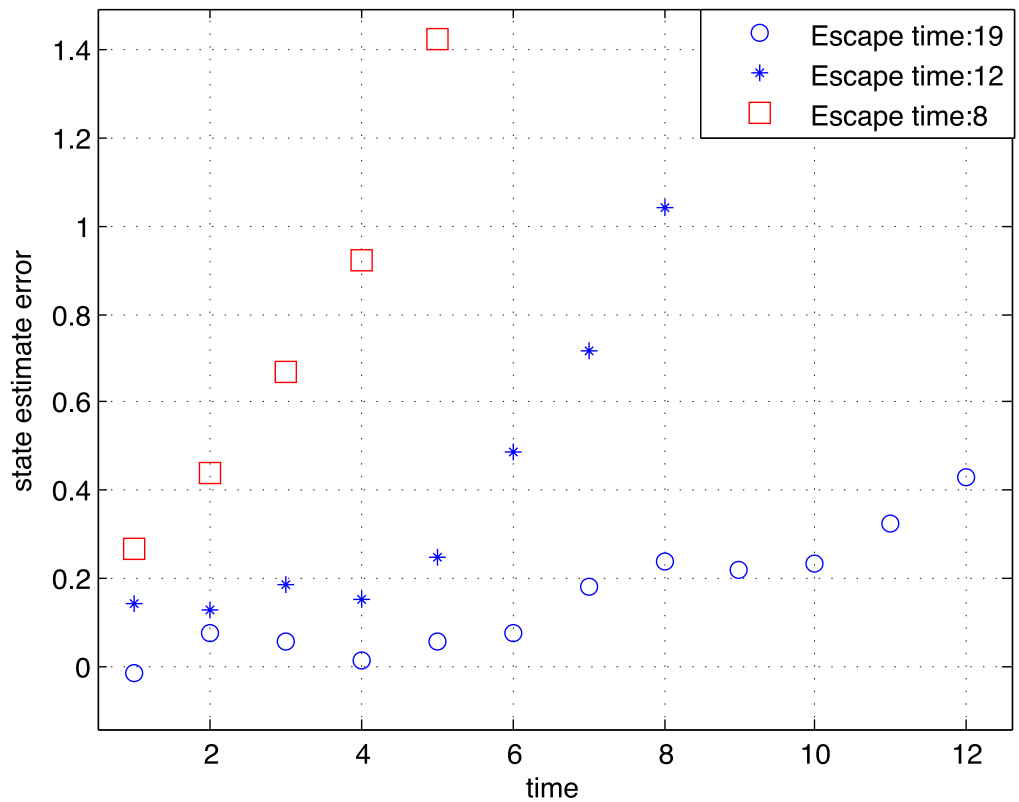

Figure 4.2. Three different estimator survival time trajectory samples for Example 1 with initial error covariance $\Sigma_{1|0} = 0.00725$.

neighborhood $|\tilde{x}| < 0.2$. Here the disturbance process is as likely to drag the error back towards small magnitudes as it is to increase the magnitude, resulting in some trajectories remaining close to zero for an extended time. This is captured by the Markov analysis.

4.2 Markov versus covariance comparison

Comparing the current Markov results to the covariance calculations of [25], we see that, in both survival time analyses, the conditional covariance is given by

$$\Sigma_{k+1|k} = A^k \Sigma_{1|0} A^{k^T} + \sum_{j=0}^{k-1} A^j Q A^{j^T}.$$

In apposition to the Markov analysis, the covariance $\Sigma_{k+1|k}$ may be used to compute directly the probabilities $\Pr(\|\tilde{x}_{k+1|k}\|_2 \leq \zeta)$ and $\Pr(\|\tilde{x}_{k+1|k}\|_2 > \zeta)$ using the normal cdf. Whence, the escape probability may be approximated via

$$\Pr(\left\|\tilde{x}_{k+1|k}\right\|_2 > \zeta) \times \prod_{j=1}^{k} \Pr(\left\|\tilde{x}_{k+1-j|k-j}\right\|_2 < \zeta). \tag{4.1}$$

This calculation may be performed using the stationary expected conditional covariance value or using the data-dependent covariance, $\Sigma_{k+1|k}$, derived from the $\{\gamma_k\}$ sequence.

Figure 4.3 displays three curves computed for Sinopoli's example from [25] in the region where the expected conditional covariance is finite, i.e. $P_\gamma > 0.36$. Here,

$$A = \begin{bmatrix} 1.25 & 0 \\ 1 & 1.1 \end{bmatrix}, \; Q = \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}, \; P_\gamma = 0.4$$

$$\bar{M} = \begin{bmatrix} 67.36 & -.052 \\ -.052 & 22.73 \end{bmatrix}, \; C = \begin{bmatrix} 1 & 1 \end{bmatrix}, \; R = 2.5, \zeta = 100.$$

This calculation demonstrates that attention solely to the covariance is problematic and, further, that the stationary expected conditional covariance is pessimistic in its implication regarding escape time, while the time-varying covariance is overly optimistic. This corroborates the results of [24], where the expected value of the conditional covariance is replaced by analysis of its probabilistic behavior.

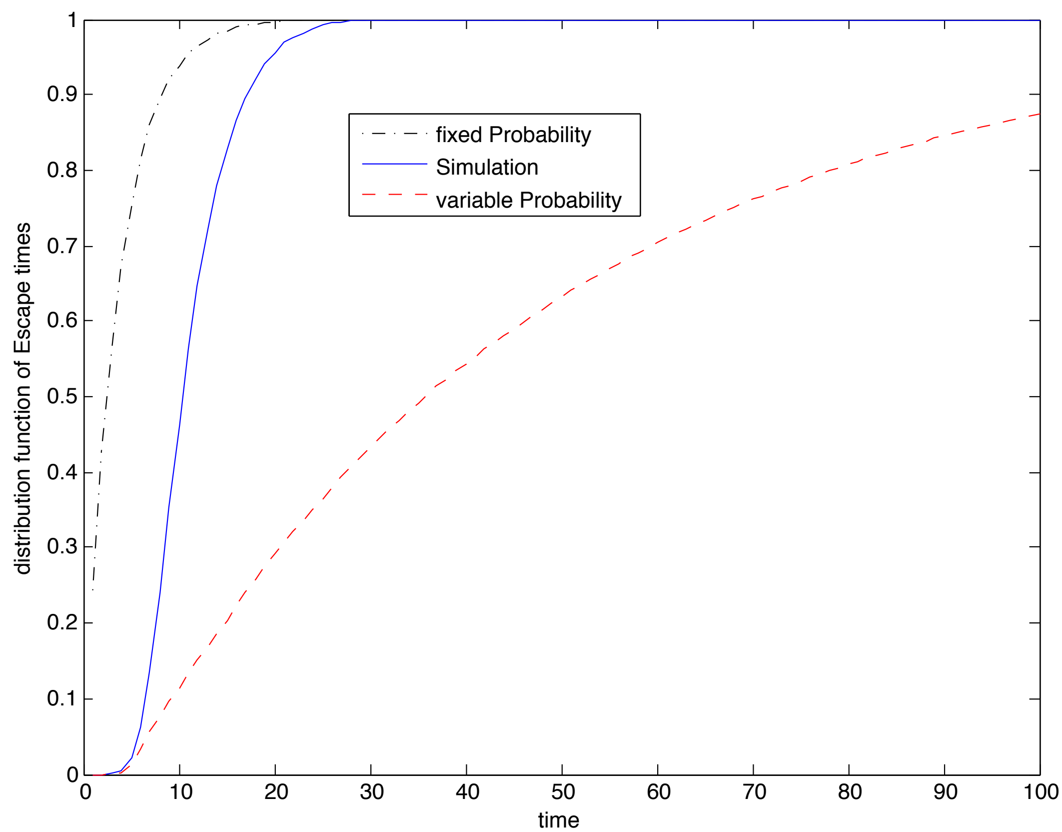

Figure 4.3. Comparison of escape time cdf for Sinopoli's Example: unbroken curve – simulation, dash-dot curve and dashed curve are both using (4.1) with fixed expected condition covariance or time-vary conditional covariance then averaged over 1000 trials.

4.2.1 Example 1 redux: Output survival times with quantized data

We return to the scalar-state Example 1, but now with state-estimate feedback control. The parameters of the controlled system and output signal bound, ζ, are:

$$A = 1.5, B = 1, C = 1, D = 0, K = 1.2,$$

$$Q = 0.005, R = 0.001, \zeta = 10.$$

We consider two quantization scenarios; 5-bit quantization of y_k, and 16-bit quantization.

Figure 4.4 depicts the cdfs for 16-bit (blue) and 5-bit (red) for P_γ taking values in

$\{0.2, 0.3, 0.4, 0.5\}$. It is noteworthy that the explicitly quantified difference in survival

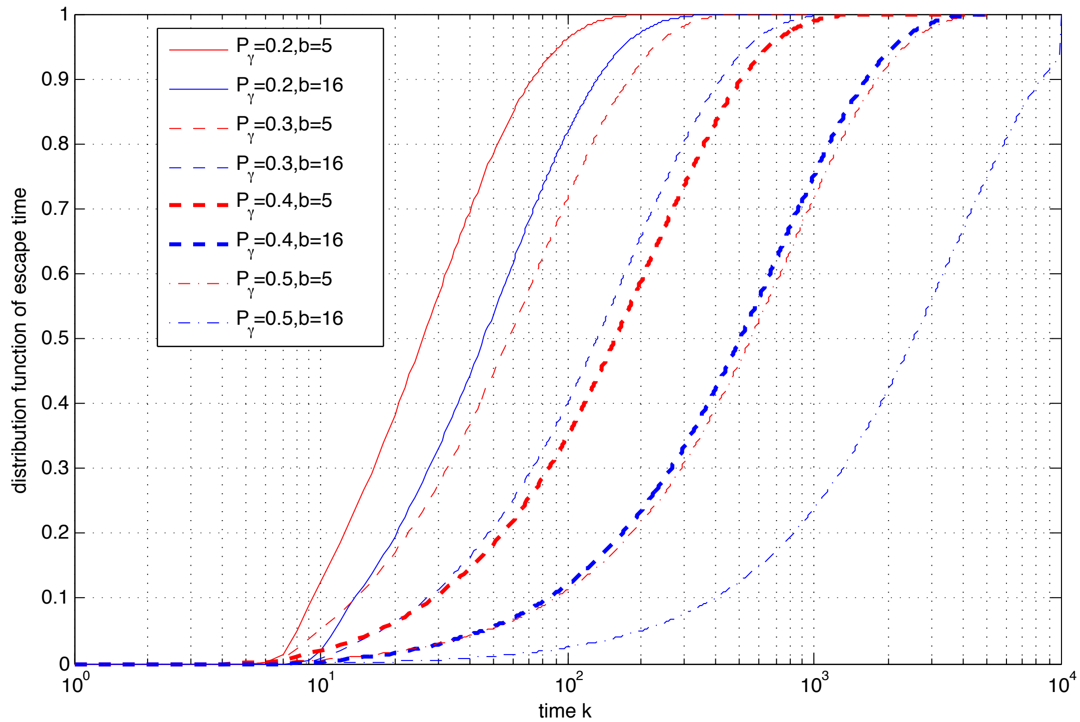

Figure 4.4. Semilog plot of cumulative distribution functions of output escape times for 5-bit (red) and 16-bit (blue) quantized measurements as the probability of successful packet arrival, P_γ, varies.

times of the differing bitrates, described by Theorem 4 in the case where $P_\gamma = 0$, is exacerbated when P_γ increases.

Figure 4.5 shows the variation in escape times for increasing values of the bound, ζ, defining the escape domain $\mathscr{D}$, and for fixed $P_\gamma = 0.2$. The bound ζ takes values in the set $[1, 10, 100, 500]$. This affects the quantization intervals and, in turn, affects: the effective measurement noise, R_{eff}, as in (2.7); the estimator covariance value $\bar{M}$ or $\Sigma_{k+1|k}$; and the initial controlled state covariance, P_1.

It is apparent that, for small numbers of bits, the quantization error becomes a significant factor when its effect on P_0 dominates that of the process noise Q. When the initial state covariance P_1 becomes large enough, the dependence of escape time on ζ

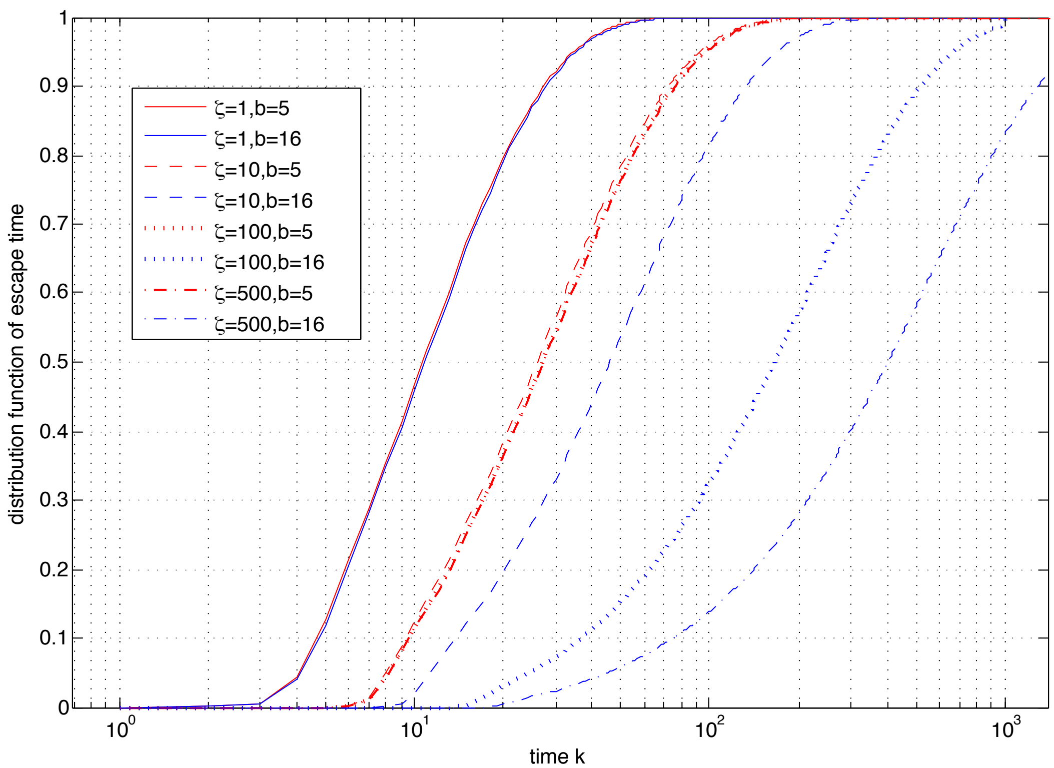

Figure 4.5. Cumulative distribution functions of controlled output escape times for Example 1 with $\zeta = 1, 10, 100, 500$, and with $P_\gamma = 0.2$. Results are shown for 5-bit quantization (red) and 16-bit quantization (blue).

depends primarily on growth in ζ versus the maximal eigenvalue of A. Thus, for these ζ values and 5-bit quantization, we see little apparent variation in the escape time statistics past $\zeta = 10$, while the 16-bit quantization escape times continue to grow dramatically with ζ.

4.3 Bitrate assignment and escape time

Communications issues of intermittency and limited bitrate may be combined for joint analysis and design, particularly in the realm of partitioning the available bitrate as a hedge against the deleterious effects of dropped packets. Evidently this is a rudimentary coding question, since one can sacrifice some bits to accommodate error recovery. This represents a departure from [25, 24], whose analysis is restricted to intermittency alone. Although, [24] countenances retransmission of data within each packet, albeit at no cost to the communication rate. More recently, [17] has studied the application of redundant parallel channels in the amelioration of packet loss effects, again without addressing quantization issues arising due to limited bitrates.

Sinopoli *et al.* [25] assume that the measurement matrix $C = I$ in (2.2) while [24] assumes that each packet contains r successive measurements, where r is the the least value for which

$$\text{range}\begin{bmatrix} C^T & A^T C^T & \ldots & A^{(r-1)^T} C^T \end{bmatrix} \supset \text{range}\, A^{r^T}.$$

Further, these data packets are not bitrate limited. In either case, the arrival of any single packet suffices for reconstruction of the state estimate with conditional covariance less than a bounded quantity derived from the Information Filter, see [1].

With a limited bitrate and $r > 1$, the assignment of bits within a measurement packet to retransmitted data comes at a cost in terms of the state estimate error, because multiply arriving copies of data carry no information and the associated bits might better have been used for the new data. We next analyze by simulation the effect of such bitrate assignment on escape time.

Suppose that the bits at any time are assigned to the most recent data and also to some fraction of the earlier data, as is illustrated in Figure 4.6. This redundancy in

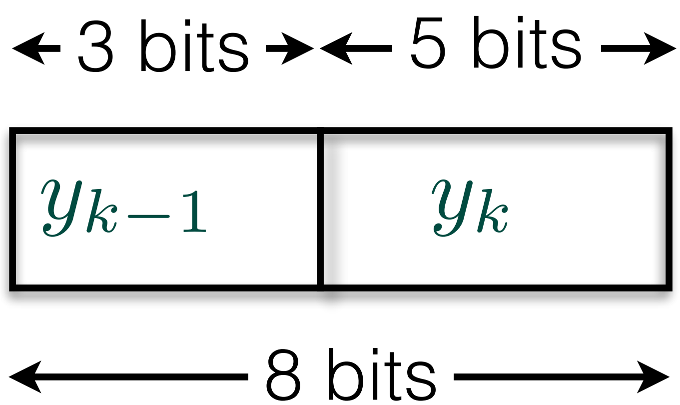

Figure 4.6. Bit allocation scheme. Here 3 of 8 bits at time k are assigned to y_{k-1} and 5 bits to y_k.

data transmission should provide some protection from packet loss at the expense of estimation accuracy. This impact can be subtle, since the reduction in bits for y_k affects both the limiting filtered estimate and controlled state covariances.

Our example takes

$$A = \begin{bmatrix} 0 & 1.5 \\ 1.5 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; C = \begin{bmatrix} 1 & 0 \end{bmatrix}, D = 0,$$

$$Q = \begin{bmatrix} 0.05 & 0 \\ 0 & 0.05 \end{bmatrix}, R = 0.01, \zeta = 5, \; K = [-1.1, 1.66].$$

Case A: We take 8 bits for y_k at each time k. Then the appropriate initial covariances are

$$\Sigma_{0|0} = \begin{bmatrix} 0.010 & 0 \\ 0 & 0.0073 \end{bmatrix}; P_0 = \begin{bmatrix} 1.027 & 1.353 \\ 1.353 & 2.361 \end{bmatrix}.$$

Case B: We take 4 bits each for y_{k-1} and for y_k with corresponding

$$\Sigma_{0|0} = \begin{bmatrix} 0.075 & 0 \\ 0 & 0.219 \end{bmatrix}, P_0 = \begin{bmatrix} 2.404 & 3.099 \\ 3.099 & 5.459 \end{bmatrix}.$$

Escape time cdfs are shown in Figure 4.7 for the case where $P_\gamma = 0.3$. We see that, in this case, the retransmission strategy serves to extend the time for escape attributable to nefarious sequences $\{\gamma_k\}$ at the cost of increasing the likelihood of early escape ascribed to the concomitant increase in initial covariances.

4.4 Comments and conclusions

We have presented an approach to the study of state estimation and state-estimate feedback stabilization with intermittent and quantized measurements. This is based on escape time analysis and computation, which is compared to earlier works in estimation and stabilization with differing descriptions of the properties – finite expected conditional covariance and mean-square stabilization – and of the communication system – intermittent but exact and quantized but certain. The central result is that the escape time can be described by a Markov chain, which in the linear case is easily computed. This yields much more precise evaluation of behavior than covariance calculations, even in the Gaussian case. Our analysis has included the treatment of quantization effects in output escape time problems and shown that bitrate choice can have a profound effect on the output escape time properties of a controlled system.

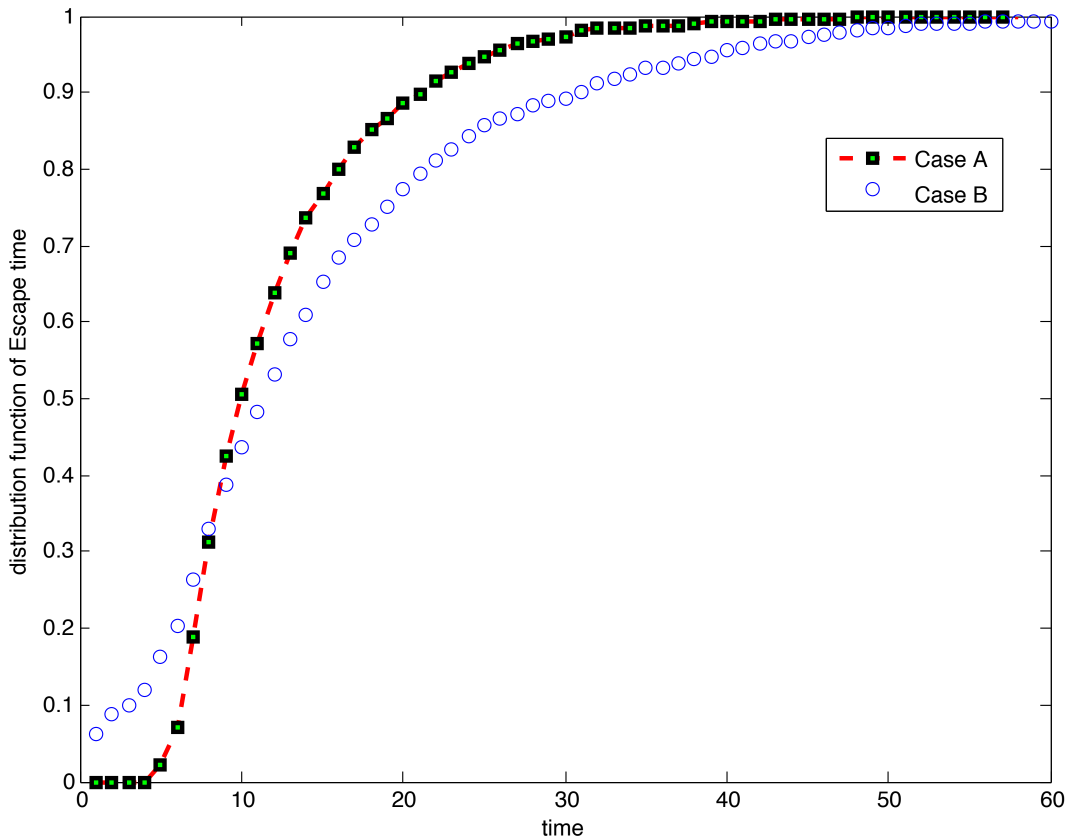

Figure 4.7. Output escape time cdfs with single transmission (Case A) and split bitrate signal retransmission (Case B) when $P_{\gamma} = 0.3$

The thrust of these problems was to replicate in the escape time framework the resetting of covariance-based analyses once certain sequences of successful packets arrive. In the Kalman filtering context, this is used to guarantee the return to an upper bounded conditional covariance matrix once a successful packet or sequence of packets arrives – the actual bound follows from the Information Filter formulation of the problem. This resetting permits the separation of the analysis of future behavior from that prior to the reset. In the escape time formulation this separation is no longer possible and it is instructive to understand why.

In the Gaussian case, one (including the authors) might be expected to use the following relation following from (3.2).

$$\begin{aligned}\alpha_k &= \Pr(J_{k+1}=1 \Big| J_k=1),\\ &= \Pr(J_{k+1}=1 \Big| J_k=1, J_{k-1}=1),\\ &= \frac{\Pr(J_{k+1}=1, J_k=1 \Big| J_{k-1}=1)}{\Pr(J_k=1 \Big| J_{k-1}=1)}.\end{aligned}$$

This latter quantity is familiar from Kalman filtering and computations such as

$$\Pr(x_{k+1}|y_k, y_{k-1}, \dots, y_1) \quad = \frac{\Pr(x_{k+1}, y_k | y_{k-1}, \dots, y_1)}{\Pr(y_k | y_{k-1}, \dots, y_1)}.$$

However for escape times, the conditioning is not on the specific value taken by a signal such as y_k. Instead, the conditioning is over the residence of $x_k \in \mathscr{D}$ and the probabilities above need to be computed as integrals over all of $\mathscr{D}$. The finite-dimensionality of the Kalman filter is lost since, even though (x_{k+1}, x_k) are jointly Gaussian, the conditional probability $\Pr(x_{k+1}|x_k \in \mathscr{D})$ is not Gaussian. This is evident in the precise calculations of Theorem 3.

Chapter 4, in full, is a reprint of the material as it appears in the section 4 and 5 of the work, "Escape time formulation of state estimation and stabilization with quantized intermittent communication", 2015 *Automatica,* Chun-Chia Huang, Robert R. Bitmead.

Part II

Bayesian filtering approach to quantized innovations Kalman filtering

Chapter 5

Problem formulation and derivation of the exact solution

5.1 Problem formulation

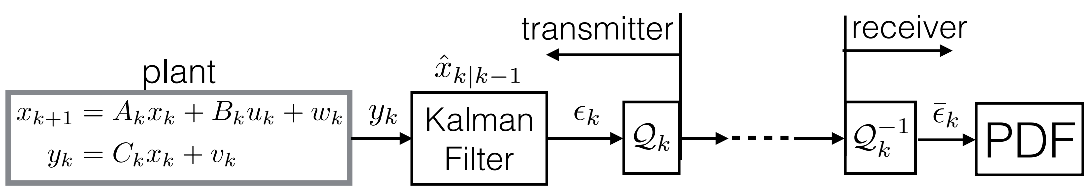

Figure 5.1. Block diagram of quantized innovations system.

We consider the system depicted in Figure 5.1, where the innovations, $\{\varepsilon_k\}$, signal from a Kalman filter is quantized and then transmitted across a communications channel to a receiver, where it is dequantized to received signal $\{\bar{\varepsilon}_k\}$ from which one seeks to compute predicted and/or filtered estimates of the state, x_k, of the plant. We make the following plant assumptions.

A.1: The plant is linear and satisfies

$$x_{k+1} = A_k x_k + B_k u_k + w_k, \tag{5.1}$$

$$y_k = C_k x_k + v_k. \tag{5.2}$$

A.2: Process and measurement noises $\{w_k\}$ and $\{v_k\}$ correlated white gaussian with

$$\begin{bmatrix} w_k \\ v_k \end{bmatrix} \sim \mathscr{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q_k & S_k \\ S_k & R_k \end{bmatrix} \delta_{kj}\right) \tag{5.3}$$

A.3: Initial state $x_0 \sim \mathscr{N}(\hat{x}_{0|-1}, \Sigma_{0|-1})$ and is independent from $\{w_k\}$ and $\{v_k\}$.

A.4: The innovations realization of the Kalman filter is run at the transmitter from initial values $\hat{x}_{0|-1}$ and $\Sigma_{0|-1}$.

$$\varepsilon_k = y_k - C_k \hat{x}_{k|k-1}, \tag{5.4}$$

$$K_k = (A_k \Sigma_{k|k-1} C_k^T + S_k)(C_k \Sigma_{k|k-1} C_k^T + R_k)^{-1} \tag{5.5}$$

$$\hat{x}_{k+1|k} = A_k \hat{x}_{k|k-1} + B_k u_k + K_k \varepsilon_k, \tag{5.6}$$

$$\Sigma_{k+1|k} = A_k \Sigma_{k|k-1} A_k^T - K_k (A_k \Sigma_{k|k-1} C_k^T + S_k)^T + Q_k. \tag{5.7}$$

Our derivation in the next section is a general formula for any quantizer function, as opposed to the specialized quantizer of the earlier chapters. We use an m-level quantizer, $\mathscr{Q}_k(\cdot)$, consisting of a collection of m intervals $(z_{k,l}, z_{k,u}]$, which form a disjoint covering of the real line and a corresponding rule, $\mathscr{Q}_k^{-1}(\cdot)$, for dequantizing the received signal into a real number — for a p-channel vector quantizer, we take a scalar quantizer in each channel for simplicity but could consider any p-channel-in/p-channel-out quantizer-dequantizer pair. When the innovations signal ε_k lies within the range $(z_{k,l}, z_{k,u}]$ the output of the quantizer is transmitted as one of m symbols which is then dequantized at the receiver as $\bar{\varepsilon}_k$. We take $\bar{\varepsilon}_k$ to be a value within the range $(z_{k,l}, z_{k,u}]$ including for the two edge saturation levels, where one of the limits is infinite. The quantizer range ζ should be related to the covariance of innovations signals. We do not focus on how to design an appropriate quantizer.

In order to clarify the quantizer, consider a 3-bit uniform quantizer-dequantizer cascade example shown in Figure 5.2 and used in Example 6.1 in Chapter 6. Evidently, $\bar{\varepsilon}_k = \mathscr{Q}_k^{-1}\mathscr{Q}_k(\varepsilon_k)$. The input is the innovations signal, ε_k, and the output is the recovered quantized innovations signal, $\bar{\varepsilon}_k$. The values $\pm\zeta$ denote the upper and lower saturation limits of the quantizer, which we take for simplicity to be symmetric. We take five-times steady state innovation covariance as the quantizer range in Figure 5.2.

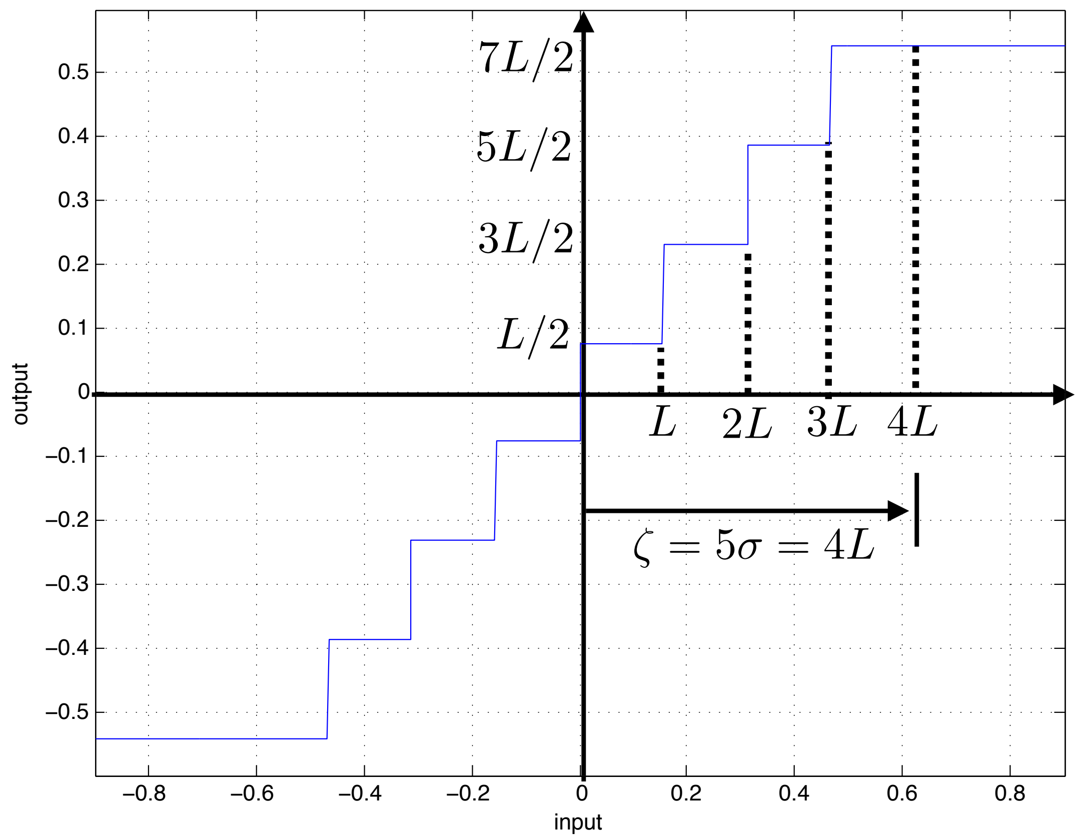

Figure 5.2. Quantizer-dequantizer function for Example 6.1. Three-Bit ($m = 8$) uniform quantizer-dequantizer pair input-output relation. The values for this quantizer are based on a steady-state innovations covariance $\sigma^2 = 0.0155$.

We make the additional assumptions.

A.5: For every value of $k \geq 0$ the receiver knows: A_k, B_k, u_k, C_k, Q_k, R_k, $\mathscr{Q}_k(\cdot)$, $\mathscr{Q}_k^{-1}(\cdot)$.

A.6: The receiver knows $\hat{x}_{0|-1}$ and $\Sigma_{0|-1}$.

5.2 Derivation of exact solution

The recursive formula of predicted pdf $p(x_{k+1}\Big|\bar{E}_k)$ and filtered pdf $p(x_{k+1}\Big|\bar{E}_{k+1})$ will be presented in this section based on the Bayesian filter presented in Section 1.3. Since the system is based on the transmission of the quantized innovations sequence $\{\bar{\varepsilon}_t, t=0.,1,\ldots,k\}$, we rely on properties of the innovations, $\{\varepsilon_t, t=0.,1,\ldots,k\}$, at the transmitter.

Property 1 *Under the linear gaussian system assumptions above, the innovations sequence, $\{\varepsilon_t, t=0.,1,\ldots,k\}$, is white, gaussian, zero-mean, with covariance $P_{\varepsilon_k} = C_k^T\Sigma_{k|k-1}C_k+R_k$.*

Property 2 *The least-squares optimal state estimates at the transmitter, $\hat{x}_{k+1|k}$ and $\hat{x}_{k|k}$, are linear combinations of $\{\varepsilon_t, t=0.,1,\ldots,k\}$. [For clarity, we take $\{u_k\}=\{0\}$.]*

$$\hat{x}_{k+1|k} = A_k\hat{x}_{k|k-1}+K_k\varepsilon_k, \tag{5.8}$$

$$= K_k\varepsilon_k + A_kK_{k-1}\varepsilon_{k-1}+\cdots+\left[\prod_{j=0}^{k+1}A_j\right]\hat{x}_{0|-1}. \tag{5.9}$$

$$L_k = \Sigma_{k|k-1}C_k^T\left(C_k\Sigma_{k|k-1}C_k^T+R_k\right)^{-1} \tag{5.10}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1}+L_k\varepsilon_k,$$

$$= L_k\varepsilon_k + K_{k-1}\varepsilon_{k-1}+A_{k-2}K_{k-2}\varepsilon_{k-2}+\cdots+\left[\prod_{j=0}^{k}A_j\right]\hat{x}_{0|-1}. \tag{5.11}$$

We set up a new system to help us compute the pdfs. The ε_k-generating system has a dimension-$2n$ state.

$$\mathscr{Z}'_{k+1} = \begin{bmatrix} x_{k+1} \\ \hat{x}_{k+1|k} \end{bmatrix} = \begin{bmatrix} A_k & 0 \\ K_kC_k & A_k - K_kC_k \end{bmatrix} \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix} + \begin{bmatrix} B_k \\ B_k \end{bmatrix} u_k + \begin{bmatrix} I & 0 \\ 0 & K_k \end{bmatrix} \begin{bmatrix} w_k \\ v_k \end{bmatrix}, \tag{5.12}$$

$$\varepsilon_k = \begin{bmatrix} C_k & -C_k \end{bmatrix} \begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix} + \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} w_k \\ v_k \end{bmatrix}. \tag{5.13}$$

Or equivalently,

$$\mathscr{Z}_{k+1} = \begin{bmatrix} x_{k+1} \\ \tilde{x}_{k+1|k} \end{bmatrix} = \begin{bmatrix} A_k & 0 \\ 0 & A_k - K_kC_k \end{bmatrix} \begin{bmatrix} x_k \\ \tilde{x}_{k|k-1} \end{bmatrix} + \begin{bmatrix} B_k \\ 0 \end{bmatrix} u_k + \begin{bmatrix} I & 0 \\ I & -K_k \end{bmatrix} \begin{bmatrix} w_k \\ v_k \end{bmatrix}, \tag{5.14}$$

$$\varepsilon_k = \begin{bmatrix} 0 & -C_k \end{bmatrix} \begin{bmatrix} x_k \\ \tilde{x}_{k|k-1} \end{bmatrix} + \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} w_k \\ v_k \end{bmatrix}. \tag{5.15}$$

If we regard either of these systems as the state equation whose state is being estimated by observation of the $\{\varepsilon_k\}$ sequence, then we would reconstruct the Innovations Filter.

We should then see:

$$\hat{\mathscr{Z}}_{k+1|k} = \left.\widehat{\begin{bmatrix} x_k \\ \tilde{x}_{k|k-1} \end{bmatrix}}\right|_{k-1} = \begin{bmatrix} \hat{x}_{k|k-1} \\ 0 \end{bmatrix},$$

$$\hat{\mathscr{Z}}'_{k+1|k} = \left.\widehat{\begin{bmatrix} x_k \\ \hat{x}_{k|k-1} \end{bmatrix}}\right|_{k-1} = \begin{bmatrix} \hat{x}_{k|k-1} \\ \hat{x}_{k|k-1} \end{bmatrix}.$$

The structure of the systems allow us to use the Bayesian filter which consists of two parts.

$$p(\mathscr{Z}_{k+1}\Big|\bar{E}_k) = \int_{\mathscr{Z}_k} p(\mathscr{Z}_{k+1}\Big|\mathscr{Z}_k)p(\mathscr{Z}_k\Big|\bar{E}_k)d\mathscr{Z}_k, \tag{5.16}$$

$$p(\mathscr{Z}_{k+1}\Big|\bar{E}_{k+1}) = \frac{p(\bar{\varepsilon}_{k+1}\Big|\mathscr{Z}_{k+1})p(\mathscr{Z}_{k+1}\Big|\bar{E}_k)}{p(\bar{\varepsilon}_{k+1})}. \tag{5.17}$$

The above equations form a recursive algorithm to compute the predicted pdf and filtered pdf for each time. We show the detail of each term:

1. $p(\mathscr{Z}_{k+1}\Big|\mathscr{Z}_k) = p\left([w_k = x_{k+1} - A_k x_k - B_k u_k] \cap [K_k v_k = (A_k - K_k C_k)\tilde{x}_{k|k-1}]\right)$
   This probability can be computed from the joint distribution of w_k and v_k.

2. The second equation, (5.17), applies the whiteness of the innovations signal in the denominator,

$$p(\bar{\varepsilon}_{k+1}\Big|\bar{E}_k) = p(\bar{\varepsilon}_{k+1}) = \int_{z_{k+1,l}}^{z_{k+1,u}} p(\varepsilon_{k+1})d\varepsilon_{k+1}$$
$$= \int_{z_{k+1,l}}^{z_{k+1,u}} \mathscr{N}(0, C_k\Sigma_{k+1|k}C^T + R_k)d\varepsilon_{k+1}$$

3. The term $p(\bar{\varepsilon}_{k+1}\Big|\mathscr{Z}_{k+1})$ can be computed as follows:

$$\begin{aligned} p(\bar{\varepsilon}_{k+1}\Big|\mathscr{Z}_{k+1}) &= \int_{z_{k+1,l}}^{z_{k+1,u}} p(\varepsilon_{k+1}\Big|\mathscr{Z}_{k+1})d\varepsilon_{k+1} \\ &= \int_{z_{k+1,l}}^{z_{k+1,u}} p(\varepsilon_{k+1} = C_k x_{k+1} - C_k\hat{x}_{k+1|k} + v_{k+1})d\varepsilon_{k+1} \\ &= \int_{z_{k+1,l}}^{z_{k+1,u}} \mathscr{N}(C_k x_{k+1} - C_k\hat{x}_{k+1|k}, R_{k+1})d\varepsilon_{k+1} \end{aligned} \tag{5.18}$$

4. The initial probability density function in the receiver

$$p(\mathscr{Z}_0\Big|\mathscr{Z}_{-1}) = p(\mathscr{Z}_0) = p\left(\begin{bmatrix} x_0 \\ \tilde{x}_{0|-1} \end{bmatrix}\right) = \mathscr{N}\left(\begin{bmatrix} \hat{x}_{0|-1} \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{0|-1} & \Sigma_{0|-1} \\ \Sigma_{0|-1} & \Sigma_{0|-1} \end{bmatrix}\right). \tag{5.19}$$

The second equality corresponds to Assumption **A.6:** that the transmitter and receiver are synchronized at time zero.

Then, we can compute all terms in formulas (5.16) and (5.17).This is conducted in the numerical computations in the next chapter.

5.3 Comparison with prior work

The central observation in our development is that the Bayesian filter state and measurement equations derive from the system

$$\mathscr{Z}_{k+1} = \begin{bmatrix} x_{k+1} \\ \tilde{x}_{k+1|k} \end{bmatrix} = \begin{bmatrix} A_k & 0 \\ 0 & A_k - K_k C_k \end{bmatrix} \begin{bmatrix} x_k \\ \tilde{x}_{k|k-1} \end{bmatrix} + \begin{bmatrix} B_k \\ 0 \end{bmatrix} u_k + \begin{bmatrix} I & 0 \\ I & -K_k \end{bmatrix} \begin{bmatrix} w_k \\ v_k \end{bmatrix}, \tag{5.20}$$

$$\bar{\varepsilon}_k = \mathscr{Q}_k^{-1} \mathscr{Q}_k \left(\begin{bmatrix} 0 & -C_k \end{bmatrix} \begin{bmatrix} x_k \\ \tilde{x}_{k|k-1} \end{bmatrix} + \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} w_k \\ v_k \end{bmatrix} \right), \tag{5.21}$$

which is directly in the form amenable to Bayesian filtering without further modification. This should be compared and contrasted to the earlier approaches of [2] and [22], which commence from the state equation of the system alone and the measurement. Both of these works use the methods of estimation theory to derive their algorithms but without accounting for the transmitter KF state, $\tilde{x}_{k|k-1}$. This is the source of the difficulty in developing the recursion for the exact pdfs. We next discuss in more detail each of these works and make comparisons.

5.3.1 Ribeiro, Giannakis and Roumeliotis

In [2], the authors use the sign of innovations (SOI) as the transmitted packet and derive the Kalman-filter-like recursive formulas from Bayesian filtering with a Gaussian assumption for the prediction pdf. In developing their recursive algorithm, the authors first approach the problem via Bayesian filtering to yield new expressions for the pdfs. Block diagram Figure 5.3 depicts their technique and shows how their innovations signal is computed; this signal is distinct from the KF innovations signal at the transmitter, since

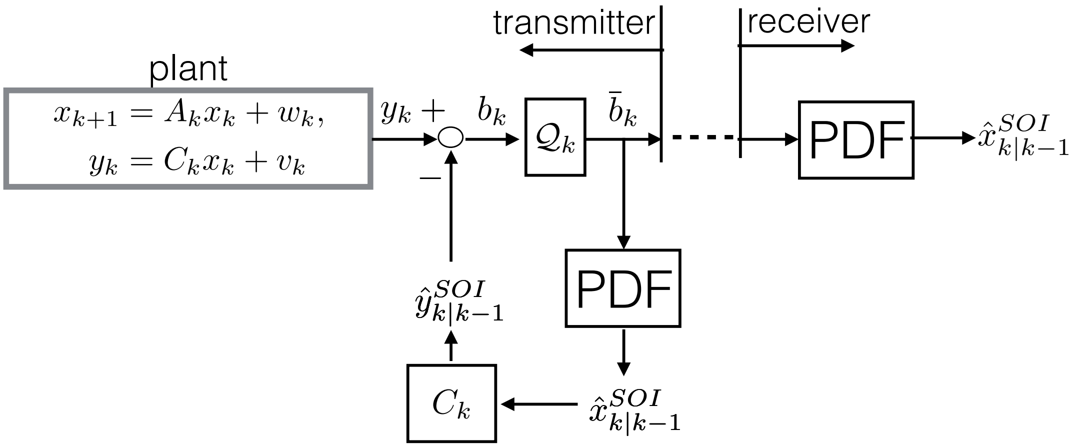

Figure 5.3. Block diagram representation of the Bayesian filtering approach of [2], where the transmitter replicates the receiver's computation of the predicted states conditional pdf to calculate an innovations signal for quantized transmission.

it is based on the data available to the receiver and the receiver's computation of the prediction of x_k, $\hat{x}^{SOI}_{k|k-1}$, the mean value of the prediction conditional pdf of the receiver. The signal transmitted to the receiver, $\bar{b}_k$, is the quantized value of b_k.

$$\bar{b}_k = \mathcal{Q}[b_k] = \mathcal{Q}_k\left[y_k - \hat{y}^{SOI}_{k|k-1}\right],$$

$$= \mathcal{Q}_k\left[y_k - C_k\hat{x}^{SOI}_{k|k-1}\right], \tag{5.22}$$

$$= \mathcal{Q}_k\left[C_k(x_k - \hat{x}^{SOI}_{k|k-1}) + v_k\right]. \tag{5.23}$$

Since b_k is not the innovations signal from the KF, it does not necessarily possess the whiteness, zero-mean, and known covariance properties of the KF innovations.

Denoting $B_k = \{b_0, b_1, \ldots, b_k\}$ and $\bar{B}_k = \{\bar{b}_0, \bar{b}_1 \ldots, \bar{b}_k\}$, the Bayesian filtering equations from [2] are as follows.

$$p(x_k \Big| \bar{B}_{k-1}) = \int p(x_k \Big| x_{k-1}) p(x_k \Big| \bar{B}_{k-1}) dx_k, \tag{5.24}$$

$$p(x_k \Big| \bar{B}_k) = \frac{p(\bar{b}_k \Big| x_k, \bar{B}_{k-1}) p(x_k \Big| \bar{B}_{k-1})}{\int_{x_k} p(\bar{b}_k \Big| x_k, \bar{B}_{k-1}) p(x_k \Big| \bar{B}_{k-1}) dx_k}. \tag{5.25}$$

The estimate $\hat{x}^{SOI}_{k|k-1}$ is defined as

$$\hat{x}^{SOI}_{k|k-1} = E[x_k \Big| \bar{B}_{k-1}] = \int x_k p(x_k \Big| \bar{B}_{k-1}) dx_k. \tag{5.26}$$

These formulas are akin to the traditional Bayesian filtering. The structure has a computational advantage in computing the term $p(\bar{b}_k \Big| x_k, \bar{B}_{k-1})$ in the filtered pdf because $\hat{x}^{SOI}_{k|k-1}$ is computed by (5.26) and admits the simplification

$$\begin{aligned} p(\bar{b}_k \Big| x_k, \bar{B}_{k-1}) &= p(\mathscr{Q}_k[C_k(x_k - \hat{x}^{SOI}_{k|k-1}) + v_k] \Big| x_k, \bar{B}_{k-1}), \\ &= \int_{z_{k,l}}^{z_{k,u}} p(C_k(x_k - \hat{x}^{SOI}_{k|k-1}) + v_k \Big| x_k, \bar{B}_{k-1}) db_k \\ &= \int_{z_{k,l}}^{z_{k,u}} \mathscr{N}\left(C_k(x_k - \hat{x}^{SOI}_{k|k-1}), R_k\right) db_k. \end{aligned}$$

We note, however, that [2] starts with the plant system state equation (5.1) with measurement signal b_k and then appeals to the Bayesian filtering equations to arrive at a recursion (5.24-5.25). Since the measurement sequence, b_k is not a function of x_k alone but of x_k and B_{k-1}, x_k is not longer the signal generating state in the formal sense. Rather the state is $x_k \cup B_{k-1}$. However, the insight that the authors bring is that x_k is the unknown part of the state, since B_{k-1} is available to both transmitter and receiver. While this is a clever observation, there remains the possibility that the b_k sequence is neither white nor zero-mean and therefore is not as efficient in the use of the channel capacity. Further research needs to be performed to assess the properties of this approach.

The authors then move from this Bayesian filtering based formulation to develop a recursion similar in form to the KF in that it propagates only conditional means and variances as if the densities were gaussian. Their calculations from the earlier analysis allow them to develop a replacement for the measurement noise covariance, R, in the KF recursion which reflects the effective measurement noise introduced by the signum function in a fashion alike that used to define R_{eff} in (2.7) in Chapter 2.

In our work, we preserve the formal BF analysis, since we identify the appropriate state of the quantized innovations signal and work with this and the transmission of the quantized innovations signal from the Kalman filter at the transmitter side. This permits the derivation of an exact Bayesian filter in recursive form. But it comes at a computational cost compared with the work of [2].

5.3.2 Sukhavasi and Hassibi

Sukhavasi and Hassibi provide the following lemma for the gaussian system above with: transmitter-side measurements y_k, their quantized variants q_k and $Y_k = \{y_0, \ldots, y_k\}$, $Q_k = \{q_0, \ldots, q_k\}$.

Lemma 3 ([22]) *The state conditioned on the quantized measurements Q_k can be expressed as a sum of two independent random variables as follows*

$$x_k \Big| Q_k \sim Z_k + R_{x_k,Y_k} R_{Y_k}^{-1} [Y_k \Big| Q_k], \tag{5.27}$$

$$\textit{where} \qquad Z_k \sim \mathcal{N}\left(0, R_{x_k} - R_{x_k Y_k} R_{Y_k}^{-1} R_{Y_k x_k}\right) \tag{5.28}$$

where R_{x_k} is the covariance of x_k and $R_{x_k Y_k}$ is the cross-covariance of x_k and the previous history of measurements Y_k and $Y_k \Big| Q_k$ is a random variable distributed as Y_k given Q_k. The first term Z_k is filtered state estimate error of the Kalman filter at the transmitter. The second term is an additive term accommodating the quantization and increasing the error covariance. From the perspective of this thesis, the second term is problematic in that it is not computable in a recursive fashion. Rather, it relies on the entire measurement history at each time.

The paper [5] provides the following idea. The predicted pdf $p(x_{k+1}|\bar{E}_k)$ can be computed from the traditional Bayesian filtering formula. The filtered pdf can be compute by the following formula.

$$p(x_{k+1}\Big|\bar{E}_{k+1}) = \int_{E_{k+1}} p(x_{k+1}\Big|E_{k+1})p(E_{k+1}\Big|\bar{E}_{k+1})dx_k \tag{5.29}$$

where $p(x_{k+1}\Big|E_{k+1})$ is a gaussian distribution function and $p(E_{k+1}\Big|\bar{E}_{k+1})$ is a truncated gaussian distribution function whose dimension increases with time, i.e. this is not a recursive formula. The integration operation in (5.29) is approximated with a recursive but approximate formula based on a mid-point approximation to the integrals. The above works either do not provide an exact solution or rely on the whole history and cannot be implemented recursively. The second part of thesis provides simple recursive formulas of Bayesian filtering for the exact computation of state estimation pdfs given quantized measurements/innovations.

Chapter 5, in full, is the section II (background and other literature) and III (analysis of techniques) of the submitted paper “Quantized Innovaitons Bayesian Filtering”, 2015 submitted to *IEEE transaction on signal processing* , Chun-Chia Huang, Robert R. Bitmead.

Chapter 6

Computational examples

In this chapter the Bayesian filtering formulas (5.16-5.17) are implemented for specific simulation examples. We study three time-invariant scalar examples with system,

$$x_{k+1} = Ax_k + w_k \tag{6.1}$$

$$y_k = Cx_k + v_k \tag{6.2}$$

1. The first example has scalar state and system matrices $A = C = 1$. The initial state is $\mathcal{N}(0, .02)$ and the process and measurement noise covariances are $Q = 0.0001$ and $R = 0.00001$. The quantizer is linear, symmetric with three bits or eight levels. The saturation level of the quantizer is ± 0.7. That is, the state remains close to its initial value and the measurements at the transmitter are very accurate. So ε_k is small with high probability.

2. The second case is the sign-of-innovations example which corresponds to the work [2]. The state process is scalar with $A = 0.95$, $C = 1$, $x_0 \sim (0, 0.0155)$, $Q = R = 0.01$. Since we transmit the signs of the innovations, the quantizer has one bit or two levels. In this case, we compare our computed pdfs with both those of the Kalman filter and the pdfs associated with [2], which takes the predicted pdf to be Gaussian.

6.1 Scalar examples

6.1.1 Case 1: almost fixed state

For the system

$$x_{k+1} = x_k + w_k,$$

$$y_k = x_k + v_k.$$

with $w_k \sim \mathcal{N}(0, 0.0001)$, $v_k \sim \mathcal{N}(0, 0.00001)$, $x_0 \sim \mathcal{N}(0, .02)$ and 3-bit quantizer (as displayed in Figure 5.2) with saturation value $\zeta = 0.7$ the Bayesian filter described in Section 5.2 is calculated in matlab using 101 samples points for each pdf. The resultant approximate pdfs are shown in Figures 6.1-6.2. For them, The upper figure is the filtered pdf and predicted pdf on receiver-side computed from Bayesian filtering with data $\bar{\varepsilon}_0 \in (0, 0.1750]$. The second figure shows the predicted and filtered pdfs on the transmitter-side from the Kalman filter. Since the state remains effectively at the same point, we have the same quantized innovations, $\bar{\varepsilon}_k \in (0, 0.1750]$ for these figures for different times. The next few steps do not significantly alter the figures further. The pdfs are close to truncated gaussians with slight smoothing at the edges according to the narrow densities of v_k in the filter and w_k in the predictor. These appear to be sensible estimated pdfs given the data. This example illustrates that the approach is sound.

6.1.2 Case 2: sign of innovations

For the system

$$x_{k+1} = 0.95 x_k + w_k,$$

$$y_k = x_k + v_k.$$

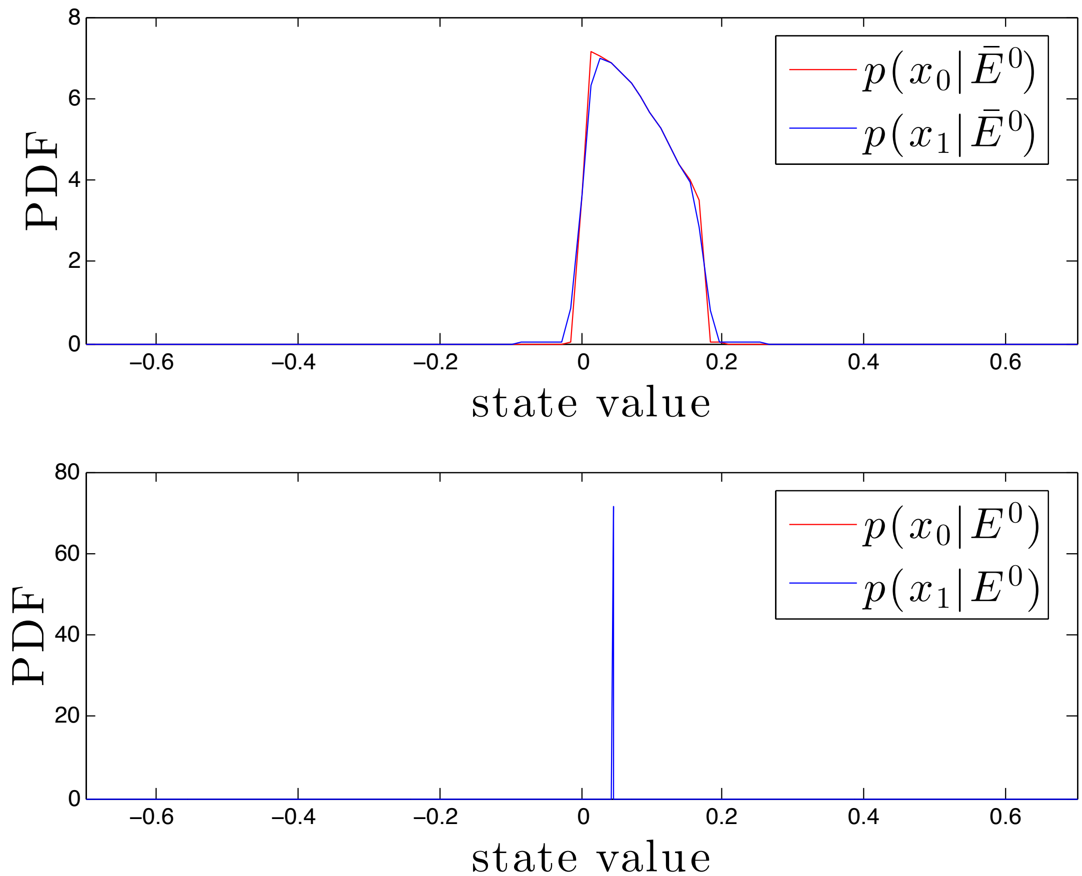

Figure 6.1. The comparison of predicted state pdf and filtered state pdf with exact measurement or quantized measurement.(first time step)

with $w_k \sim \mathcal{N}(0, 0.01)$, $v_k \sim \mathcal{N}(0, 0.01)$, $x_0 \sim \mathcal{N}(0, .0155)$ and one-bit/two-level quantizer the Bayesian filter is again applied in matlab with 101 samples for each pdf. This example appears in [2] where they apply their Kalman-filter-like algorithm. Figures 6.3-6.5 show the evolving probability density functions when the sign-of-innovations signal is negative at each of three successive times. The upper figures display the filtered pdfs and predicted pdfs on the receiver-side computed using our Bayesian filter. The middle figures show the predicted and filtered pdfs on the transmitter-side from the Kalman filter The lower figures show the gaussian predicted pdf using SOI statistic quantities from the algorithm of [2].

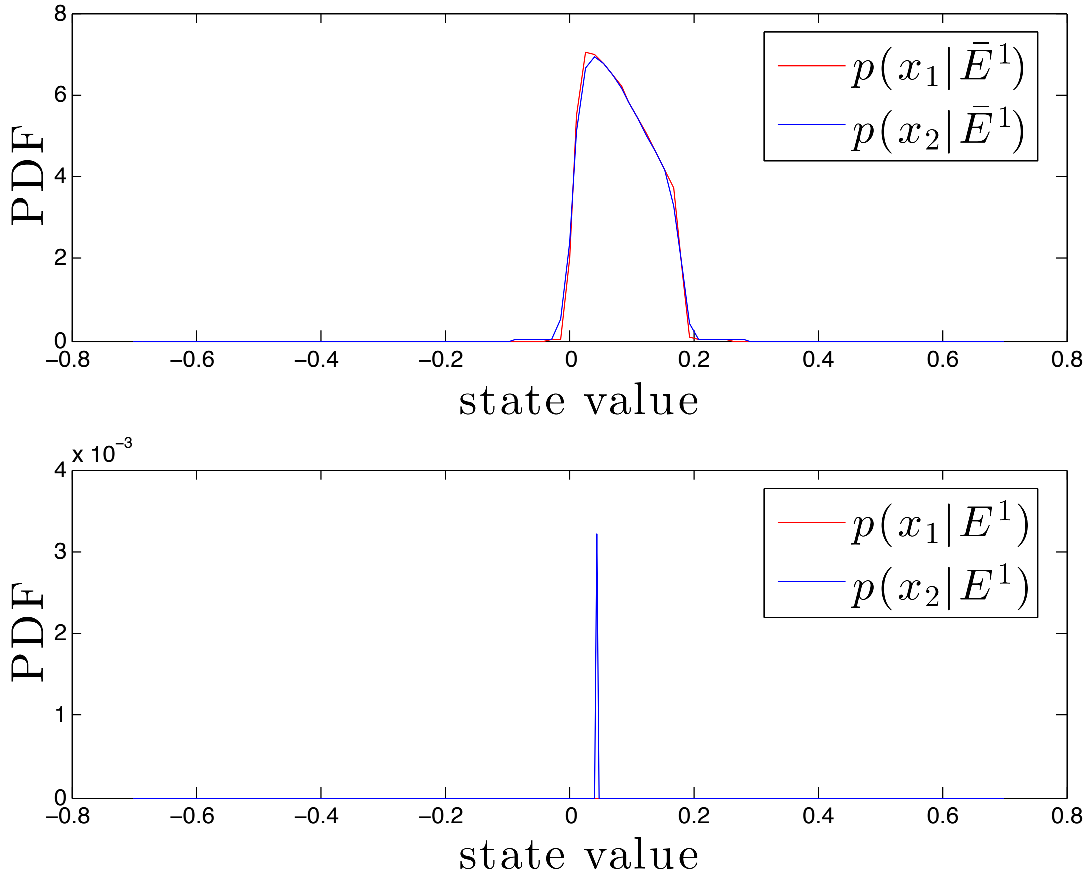

Figure 6.2. The comparison of predicted state pdf and filtered state pdf with exact measurement or quantized measurement.(second time step)

We compare the results from the Kalman filter at the transmitter and the prior pdf assumed in the derivation of [2]. This example indicates that the calculation in [2] might be overly optimistic about the estimate quality.

6.2 Conclusions and future work

We have presented an approach of state estimation problem for computing the probability density functions with the quantized innovations signals. The solution we provides is exact and also is recursive. The receiver-side computation of predicted and filtered state pdfs is performed using the Bayesian filter with a system state model containing the plant system state, x_k, augmented by the transmitter-side Kalman filter state error, $\tilde{x}_{k|k-1}$. We compare the work with [2] whose Kalman-filter-like algorithm is derived based on the Bayesian filter and uses synchronized transmitter and receiver state estimates and, therefore, a different notion of innovations from the KF. We have provided a more general algorithm which maintains the calculation of the KF and its innovations at the transmitter. On the other hand, our algorithm has significantly greater computational burden at the receiver side because of calculating full pdfs. By the same token, the workload is greatly reduced compared with, say, the Particle filter.

The future work could be: We would introduce more complicate communication channel, including packet-drop, delay in this work. Bit rate resource allocation as in Section 4.3 of escape time can be implemented in computing pdf from our Bayesian filtering by recomputing from previous pdfs whose new measurement just arrived. TCP/IP system with packet-dropping known to both ends is worth researching while applying our Bayesian filtering on it.

The full comparison with [2] and [22] to understand the computational comparisons versus the performance is worth researching. Especially, the structure of [2] presented in 5.3 provides another approach to compute exact pdf.

Chapter 6, in full, is the section IV (comparison and examples) of the submitted paper “Quantized Innovaitons Bayesian Filtering”, 2015 submitted to *IEEE transaction on signal processing* , Chun-Chia Huang, Robert R. Bitmead.

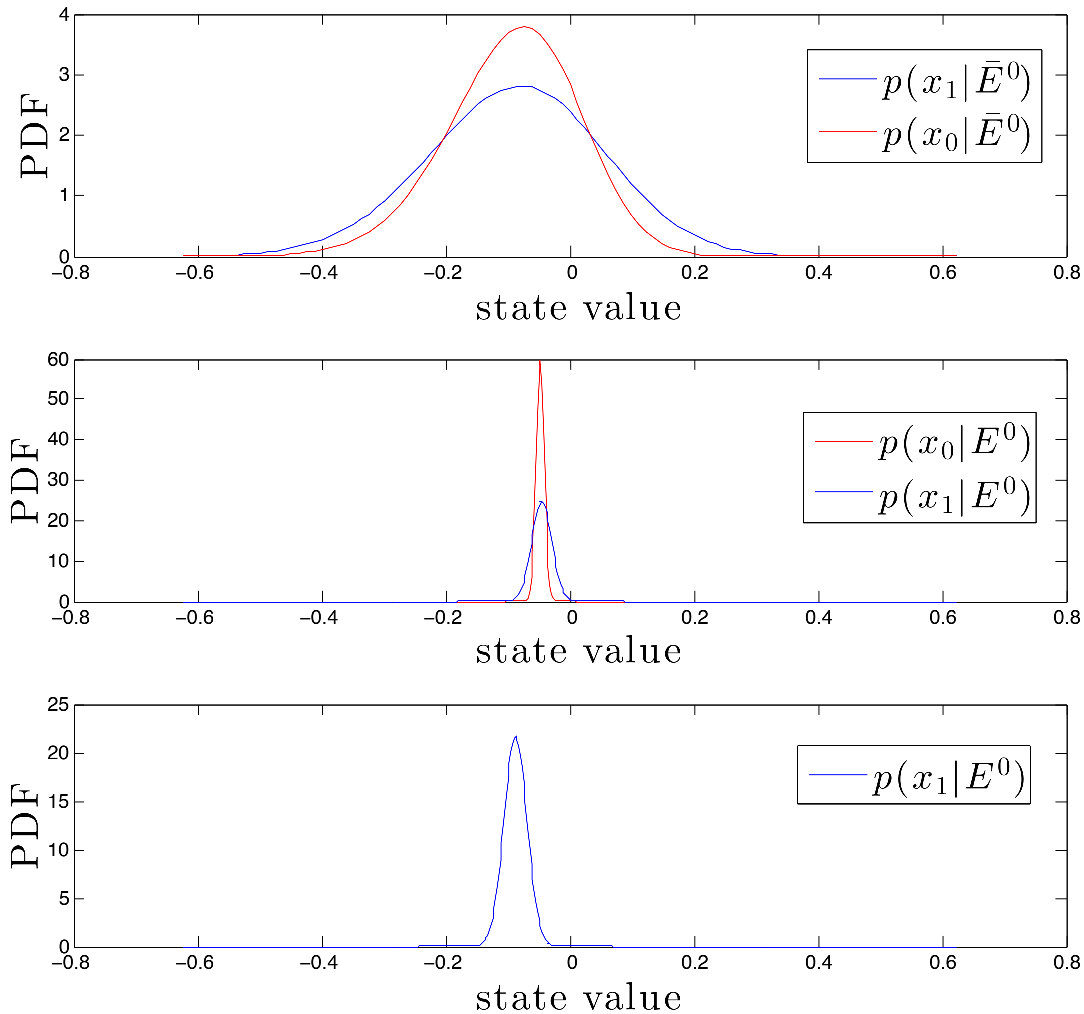

Figure 6.3. The estimation quantities: $\hat{x}_{1|0} = -0.0454$, $\Sigma_{1|0} = 0.016$, $\hat{x}^{SOI}_{1|0} = -0.0875$, $\Sigma^{SOI}_{1|0} = 0.018$. The signal values in the transmitter-side system are: $x_0 = -0.1088$, $x_1 = -0.1260$, $w_0 = -0.0226$, $v_0 = 0.0371$.

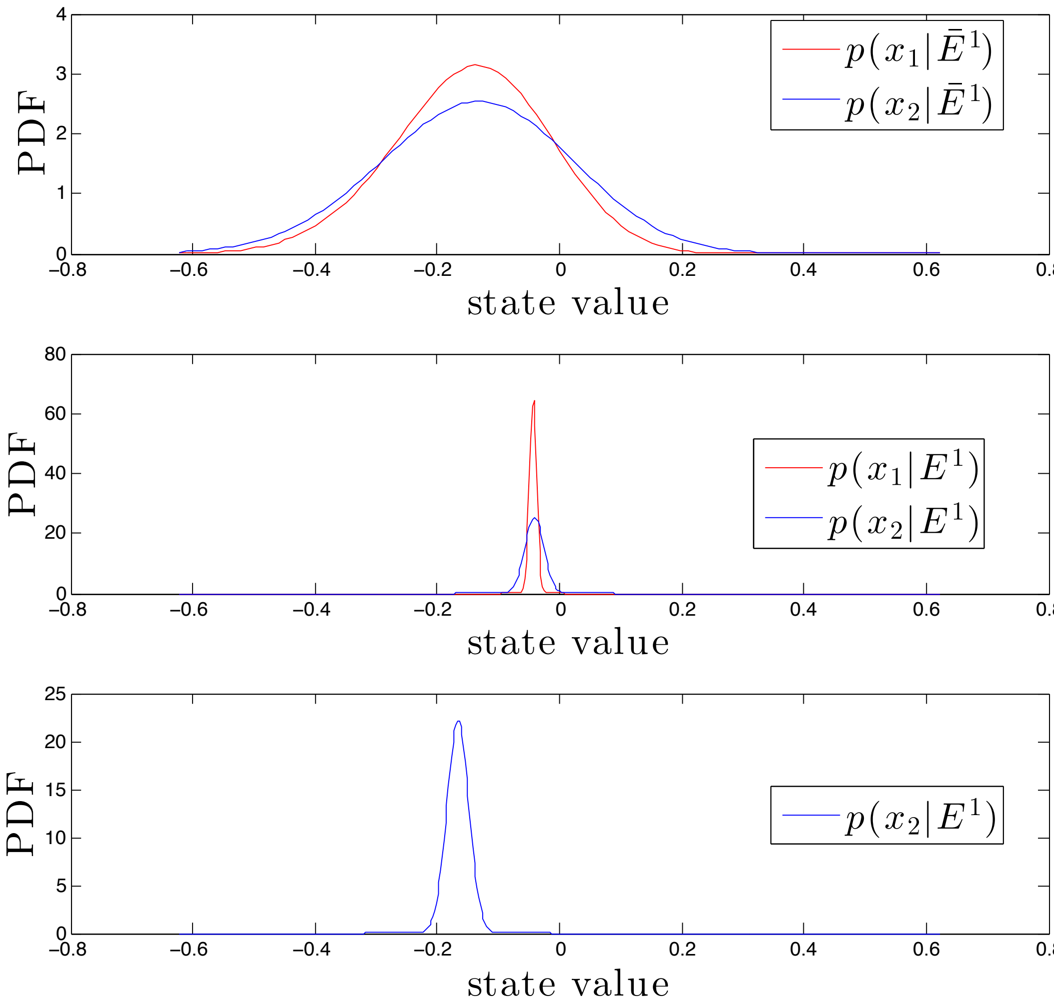

Figure 6.4. The estimation quantities: $\hat{x}_{2|1} = -0.0410$, $\Sigma_{2|1} = 0.0156$, $\hat{x}_{2|1}^{SOI} = -0.1661$, $\Sigma_{2|1}^{SOI} = 0.018$. The signal values in the transmitter-side system are: $x_1 = -0.1260$, $x_2 = -0.2286$, $w_1 = -0.1089$, $v_1 = 0.1117$.

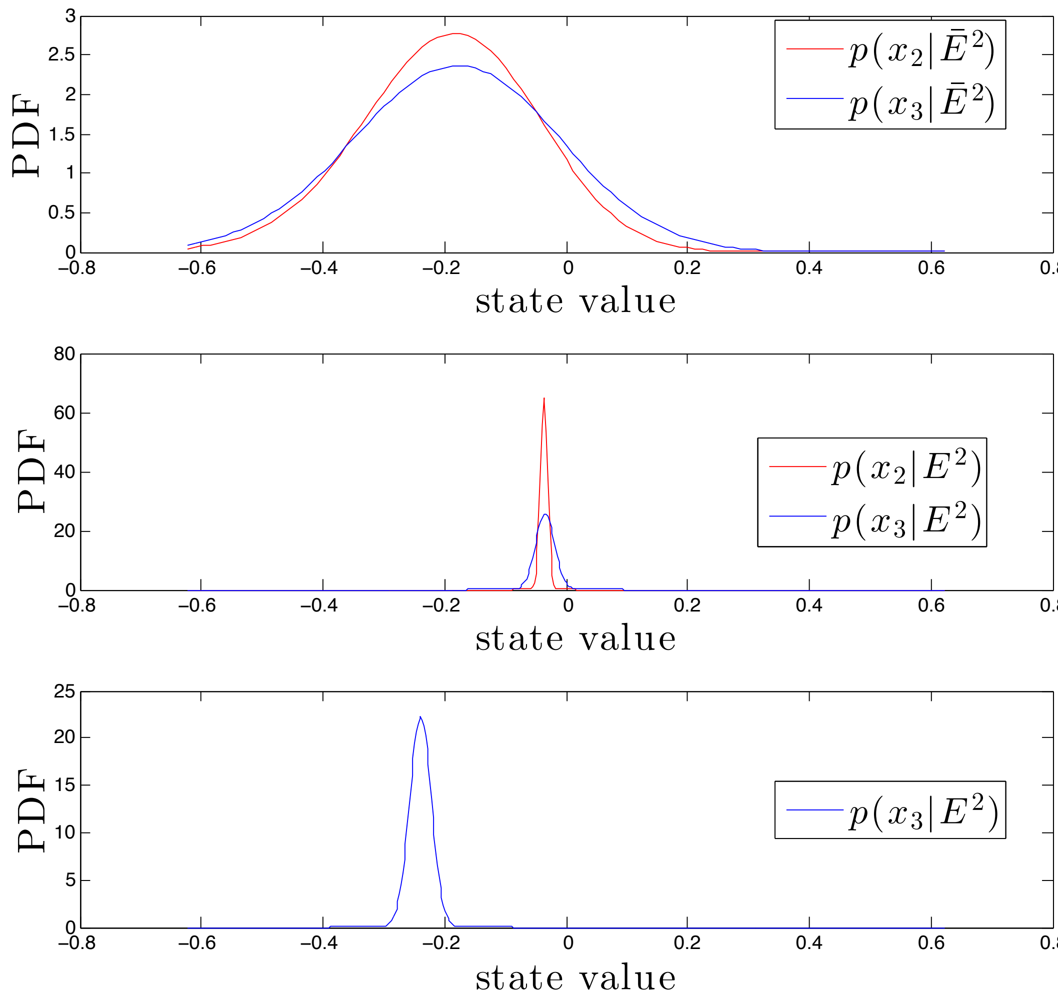

Figure 6.5. The estimation quantities: $\hat{x}_{3|2} = -0.0345$, $\Sigma_{3|2} = 0.0155$, $\hat{x}^{SOI}_{3|2} = -0.2394$, $\Sigma^{SOI}_{3|2} = 0.0179$. The real value : $x_2 = -0.2286, x_3 = -0.1619, w_2 = 0.0553, v_2 = 0.0033$.

Bibliography

[1] B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, New Jersey, 1979.

[2] A.Ribeiro, G.B. Giannakis, and S.I. Roumeliotis. Soi-kf: Distributed kalman filtering with low-cost communications using the sign of innovations. *IEEE Trans. Signal Processing*, 54:4782–4795, 2006.

[3] Simon Dan. *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[4] D.F. Delchamps. Stabilizing a linear system with quantized state feedback. *IEEE Transactions on Automatic Control*, 35(8):916 –924, August 1990.

[5] Vesselin P. Jilkov Duan, Zhansheng and X. Rong Li. State estimation with quantized measurements: Approximate mmse approach. *Information Fusion, 2008 11th International Conference on. IEEE*, 2008.

[6] J. Evans and V. Krishnamurthy. Optimal sensor scheduling for hidden markov models. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, 1998*, volume 4, pages 2161 –2164 vol.4, May 1998.

[7] R. Evans, V. Krishnamurthy, G. Nair, and L. Sciacca. Networked sensor management and data rate control for tracking maneuvering targets. *IEEE Transactions on Signal Processing*, 53(6):1979 – 1991, june 2005.

[8] M.I. Freidlin and A.D. Wentzell. *Random perturbations of dynamical systems*. Springer Verlag, Berlin, 1998.

[9] J.A. Fulton, R.R. Bitmead, and R.C. Williamson. Smoothing approaches to reconstruction of missing data in array processing. In D. Cochran, W. Moran, and L.B. White, editors, *Defence Applications of Signal Processing*, pages 87–94. Elsevier, 2001.

[10] C.C. Huang and R.R. Bitmead. Escape-time formulation of state estimation and stabilization with intermittent communication. In *European Control Conference*, pages 99–104, Zürich, 2013.

[11] M. Huang and S. Dey. Stability of kalman filtering with markovian packet losses. *Automatica*, 43(4):598–607, 2007.

[12] S. Kar, B. Sinopoli, and J.M.F. Moura. Kalman filtering with intermittent observations: Weak convergence to a stationary distribution. *IEEE Transactions on Automatic Control*, 57(2):405 –420, February 2012.

[13] R.Z. Khasminskii. *Stochastic stability of differential equations*. Sijthoff & Noordhoff, Alphen aan den Rijn, 1980.

[14] Y. Zhao K.You and L. Xie. Recursive quantized state estimation of discret-time linear stochastic systems. *Proceeedings of the 7th Asian Control Conference Hong Kong*, 2009.

[15] R.J.A. Little and D.B. Rubin. *Statistical Analysis with Missing Data*. John Wiley & Sons, New York, 1987.

[16] S.M. Meerkov and T. Runolfsson. Residence time control. *IEEE Transactions on Automatic Control*, 33(4):323 –332, April 1988.

[17] A.R. Mesquita, J.P. Hespanha, and G.N. Nair. Redundant data transmission in control/estimation over lossy networks. *Automatica*, 48:1612–1620, 2012.

[18] P. Minero, M. Franceschetti, S. Dey, and G. N. Nair. Data rate theorem for stabilization over time-varying feedback channels. *IEEE Trans Automatic Control*, 54(2):243–255, February 2009.

[19] Yilin Mo and B. Sinopoli. Kalman filtering with intermittent observations: Tail distribution and critical value. *Automatic Control, IEEE Transactions on*, 57(3):677 –689, March 2012.

[20] G. Nair and R.J. Evans. Stabilizability of stochastic linear systems with finite feedback data rates. *SIAM Journal on Control and Optimization*, 43(2):413–436, 2004.

[21] M. Franceschetti P. Minero, L. Coviello. Stabilization over markov feedback channels: The general case. *IEEE Trans Automatic Control*, 58(2):349–362, February 2013.

[22] R.Sukhavasi and B. Hassibi. The kalman-like particle filter: Optimal estimaiton with quantized innovaitons/meqasurements. *IEEE Transactions on Signal Processing*, 61(1), January 2013.

[23] F.C. Schweppe and E.J. Handschin. Static state estimation in electric power systems. *Proceedings of the IEEE*, 62(7):972 – 982, July 1974.

[24] Ling Shi, M. Epstein, and R.M. Murray. Kalman filtering over a packet-dropping network: a probabilistic perspective. *IEEE Trans. Automatic Control*, 55(3):594–604, 2010.

[25] B. Sinopoli, L. Schenato, M. Francheschetti, K. Poolla, M.I. Jordan, and S.S. Sastry. Kalman Filtering with intermittent observations. *IEEE Trans Automatic Control*, 49(9):1453–1464, September 2004.

[26] R. Sukhavasi and B. Hassibi. Particle filtering for quantized innovaitons. *IEEE Int. Conf. Acoust., Speech Signal Process.(ICASSP 2009)*, pages 2220–2232.

[27] Egils Sviestins and Torbjorn Wigren. Optimal recursive state estimation with quantized measurements. *IEEE Transactions on Automatic Control*, 45(4), April 2000.

[28] D.N. Swingler and R.S. Walker. Line-array beamforming using linear prediction for aperture interpolation and extrapolation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(1):16 –30, January 1989.

[29] S. Tatikonda and S. Mitter. Control over noisy channels. *IEEE Transactions on Automatic Control*, 49(7):1196 – 1201, July 2004.

[30] S. Tatikonda and S. Mitter. Control under communication constraints. *IEEE Transactions on Automatic Control*, 49(7):1056 – 1068, July 2004.

[31] S.R.S. Varadhan. *Large deviations and applications*. SIAM, Philadelphia, 1984.

[32] B. Widrow and I. Kollar. *Quantization Noise Roundoff Error in Digital Computation, Signal Processing, Control, and Communications*. Cambridge Univ. Press, 2008.

[33] N. Wiener. *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*. Wiley, New York, 1949.

[34] D. Williamson. Finite wordlength design of digital kalman filters for state estimation. *IEEE Transactions on Automatic Control*, 30(10):930 – 939, October 1985.

[35] D. Williamson and K. Kadiman. Optimal finite wordlength linear quadratic regulation. *IEEE Transactions on Automatic Control*, 34(12):1218 –1228, December 1989.

[36] Wing Shing Wong and R.W. Brockett. Systems with finite communication bandwidth constraints. i. state estimation problems. *IEEE Transactions on Automatic Control*, 42(9):1294 –1299, September 1997.

[37] Wing Shing Wong and R.W. Brockett. Systems with finite communication bandwidth constraints. ii. stabilization with limited information feedback. *IEEE Transactions on Automatic Control*, 44(5):1049 –1053, May 1999.

[38] Keyou You and Lihua Xie. Minimum data rate for mean square stabilizability of linear systems with markovian packet losses. *IEEE Trans Automatic Control*, 56(4), 2011.

[39] J. Zabzcyk. Exit problems and control theory. *Systems & Control Letters*, 6:165–172, 1985.