System Bounds for Multisensor Fusion With Intermittent Observations

Robert Distel
Department of Electrical and Electronic Engineering
University of Melbourne
Melbourne Australia
rdistel@pgrad.unimelb.edu.au

Abstract – *A closed form solution for an upper bound to the multi sensor Riccati equation with intermittent observations is derived. This solution is then used as a means to analyse system performance when the message loss process is a function of the number of sensors and to determine a simple analytic closed form method to determine when a message loss process causes a measure of the covariance to decrease as more sensors are added.*

Keywords: missing observations, sensor networks

1 Introduction

The continued increase in the applications for ubiquitous sensor systems requires better analytic tools to design and analyse the performance of these systems. Frequently sensor systems are not able to reliably communicate between the sensors and the nodes where the sensor information is required. A common scenario is that of the wireless sensor network where packets are deleted according to some Bernoulli process, or where the parameter to be observed is intermittently observable.

The goal of this paper is to study the effects of the system model, i.e., the model used to describe the movement of a target and its measurement by sensors, and the rate of communication failures on estimation performance. We consider a centralised fusion system with linear dynamic and measurement equations. Observations transmitted to the fusion centre are dropped according to a Bernoulli process with a rate dependent on the number of sensors. The relationship between the drop probability and the number of sensors is referred to as the observation loss function. The main contribution of this paper is to establish conditions under which estimation performance will be improved by the addition of further sensors to a network with a given system model and observation loss function. The utility of this result is that it can easily be used to guarantee a desired level of performance for a range of observation loss functions. This is an important consideration in practical systems which must operate reliably in varying conditions.

Despite the recent focus on analysing the properties of systems subject to communications failures the particular problem considered here does not appear to have been solved previously. Papers by Sinopoli [1], [2] focus on providing bounds on the drop probability when the drop probability is independent of other parameters for a single sensor case assuming the drop process is a Bernoulli process. The bound is established is the lower bound of the drop probability required to maintain a bounded expected covariance. This single sensor system was extended to the two sensor case by Liu and Goldsmith in [3] where the drop probabilities are different for each sensor (although still constant). A further development was the development of equivalent estimator stability criteria for a group of N sensors including both dropping and bounded delay where the drop probabilities are constant by Mateev and Savkin in [4].

Other investigations have focused on specific multisensor scenarios where the observation process is related to the number of sensors. In [5] the relationship between sensors and the transmission system is investigated in order to mitigate transmission losses through increasing the number of samples taken. In further papers coauthored by Sinopoli [6] and [7] a more complex sensor communications behaviour was modelled and methods of optimization of the sensor system are discussed. What is apparent in the discussion is that incorporating the specific sensor loss relationship into the problem is a complex exercise which prevents an analysis of the sensor system when subject to a generalised loss process.

The results in this paper greatly extend on [6] and [7] by analytically establishing the effects of a general observation loss function on estimation performance.

2 Problem Formulation

We consider a continuous time linear system governed by the differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) + \mathbf{w}(t)$$

where $\mathbf{x}(t) \in \mathbb{R}^m$ is the system state vector and $\mathbf{w}(t)$ is a continuous time additive process noise with zero mean and autocorrelation $\boldsymbol{Q}_c$. The state of the system is coupled to observations through the relationship

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t)$$

where $\mathbf{v}(t)$ is a continuous time additive process with zero mean with autocorrelation $\mathbf{R}$.

The process is unreliably observed by N observers and the observations are used to generate a fused estimator. Let the event $\gamma_{k,n}$ be the loss of the observation $y_{k,n}$ (the n th observer at time t_k) with a sequence of observation events denoted $\gamma_{1:k,1:N} \in \Gamma_{k,N} = \{0,1\}^{k\times N}$. The event is modelled as an i.i.d Bernoulli random variable with mean a.The sensors transmit at equispaced intervals such that for a fixed sampling interval t_s at time $\frac{(j-1)}{N}t_s : j = 1,\ldots,N$ the jth sensor produces an observation. In further discussion we will refer to a as the observation loss probability and when the observation loss probability is a function of the number of sensors we shall refer to this as an observation loss function $a(N)$.

Using the sensors and the unreliable communications system discrete observations are taken and a discrete state estimate $\hat{\mathbf{x}}(t_k)$ is formed with covariance $\mathbf{P}(t_k)$. As the observation process is a stochastic process the properties of the estimator are random variables. An important property of the estimator $\hat{\mathbf{x}}(t_k)$ is $\lim_{k\to\infty} \mathsf{E}(\mathbf{P}(t_k))$, the expected variance of the estimator over all sequences in $\Gamma_{k,N}$ as $k \to \infty$, i.e. the value that the variance of the estimator converges to after a large number of observations. This paper is concerned with the relationship between the objective measure function $\mathcal{M}(\lim_{k\to\infty} \mathsf{E}(\mathbf{P}(t_k)))$ and the observation loss function, noting that $\lim_{k\to\infty} \mathsf{E}(\mathbf{P}(t_k))$ is a function of both a and N. Imposing the requirement that $\mathcal{M}(\lim_{k\to\infty} \mathsf{E}(\mathbf{P}(t_k)))$ be constant requires that the observation loss function be of a particular form. We will refer to this form as the characteristic observation loss function $\alpha(N,\kappa)$ where κ is some parametric value. By determining the structure of $\alpha(N,\kappa)$ it becomes possible to determine if an arbitrary observation loss function $a(N)$ will cause the objective function to increase or decrease by simply considering the properties of $\alpha(N,\kappa)$ rather than $\mathcal{M}(\lim_{k\to\infty} \mathsf{E}(\mathbf{P}(t_k)))$. This is particularly useful in network design problems as the form of $\alpha(N,\kappa)$ is solely set by the linear system while $a(N)$ is solely set by the communications system under consideration. The study of these two equations allows the optimal number of sensors to be determined, and the sensitivity of the system to variations on the number of sensors, network parameters and observability.

3 Analysis

The analysis is broken into three sections. Firstly the upper bound of the sensor system of N sensors is determined in the presence of a Bernoulli loss process and derivatives of this bound are taken. Secondly the resulting derivatives are used to form differential equations based on the objective function and the differential equation solved to form a general expression for $\alpha(N,\kappa)$ that bounds the constant objective function contour. Finally these results are used to provide a method for determining the behaviour of a general observation loss function $a(N)$ by comparing it against the characteristic observation loss function $\alpha(N,\kappa)$. In order to progress this analysis the following conditions are imposed.

Notation 1 *$\lambda_{\max}(\mathbf{X})$ means the algebraically largest eigenvalue of $\mathbf{X}$. $\lambda_{\min}(\mathbf{X})$ means the algebraically smallest eigenvalue of $\mathbf{X}$. Consequently $-\lambda_{\max}(-\mathbf{X}) = \lambda_{\min}(\mathbf{X})$. $\operatorname{vec}\mathbf{X} > \operatorname{vec}\mathbf{Y}$ means that $\lambda_{\max}(\mathbf{X}-\mathbf{Y}) > 0$ when $\mathbf{X},\mathbf{Y} \in \mathbb{S}^{m\times m}$. $\mathbb{S}^{m\times m}$ is a symmetric real matrix of dimension m.*

Condition 1 *That the sensors make observations at constant intervals t_s such that at time $\frac{(j-1)}{N}t_s : j = 1,\ldots,N$ the jth sensor makes an observation.*

As more sensors are added the sampling times of the sensor set are adjusted so that the interval between any two sensors measurements is not less than $\frac{t_s}{N}$ while maintaining the individual sensor transmission interval of t_s. This allows the process of increasing the number of sensors to be equivalent to simply increasing the rate at which the system is observed. It can be argued that ensuring that all observations occur at the same time is no more or less difficult than ensuring all observations occur at different times, and there is the added benefit of reduced correlation between observations.

Condition 2 *The sensor system is homogenous. $\left(\mathbf{H}_j\mathbf{R}_j^{-1}\mathbf{H}_j'\right)^{-1}$ is the same for all sensors and exists.*

Condition 3 *The process parameter $\mathbf{F}$ have $\lambda_j(\mathbf{F}) > 0$. i.e. $-\mathbf{F}$ is asymptotically stable*

The principal results of this paper are stated as theorems below. Omitted proofs are found in the appendix at the end of the paper.

Theorem 1 *An upper bound $\overline{\mathbf{P}(a,N)} : \overline{\mathbf{P}(a,N)} > \lim_{k\to\infty} \mathsf{E}(\mathbf{P}(t_k)) = \mathbf{P}(a,N)$ on the covariance of a system of N sensors subject to an observation loss process* *characterised by* $\mathsf{E}(\gamma) = a$ *is given by*

$$\operatorname{vec}\overline{\mathbf{P}(a,N)} = (1-a)\left(\mathcal{A}^{1-\frac{1}{N}} - a\mathcal{A}\right)^{-1} \times \left(\operatorname{vec}\mathbf{P} + \mathcal{A}\operatorname{vec}(\mathbf{B})^{-1} + \left(\frac{a}{1-a}\mathcal{A} + \mathbf{I}\right)\mathbf{G}^{-1}\operatorname{vec}\boldsymbol{Q}_c\right) \quad (1)$$

where $\otimes$ *is the Kronecker product.*

$$\begin{aligned} \mathbf{B} &= \mathbf{H}\mathbf{R}^{-1}\mathbf{H}' \\ \mathbf{Q}_N &= \int_0^{\frac{t_s}{N}} \exp(\mathbf{F}t)\,\boldsymbol{Q}_c \exp(\mathbf{F}t)'\,dt \\ \mathbf{A}_N &= \mathbf{A}^{\frac{1}{N}} = \exp(\mathbf{F}t_s/N) \\ \mathbf{G} &= (\mathbf{F}\otimes\mathbf{I} + \mathbf{I}\otimes\mathbf{F}) \\ \mathcal{A}_N &= \mathcal{A}^{\frac{1}{N}} = \mathbf{A}_N\otimes\mathbf{A}_N = \mathbf{A}^{\frac{1}{N}}\otimes\mathbf{A}^{\frac{1}{N}} \end{aligned}$$

and

$$\mathbf{P} = \mathbf{A}\left(\mathbf{P}^{-1} + \mathbf{B}\right)^{-1}\mathbf{A}' + \mathbf{Q} \quad (2)$$

is the solution of the Riccati equation for a single sensor system with reliable observations.

The advantages of Theorem 1 is that it provides a direct extension to [1], [3] and [4] by providing an upper bound to the expected covariance for a fixed N and fixed a. The disadvantages of this upper bound are that it is derived assuming the invertibility of $\mathbf{B}$ and $\mathbf{F}$. Both assumptions do not hold in a number of useful and interesting cases, in particularly in target tracking. The case of $\mathbf{F}$ invertibility is easily dealt with by separating the integration of $\boldsymbol{Q}_c$ into a linear part and exponential part. The invertibility of $\mathbf{B}$ can be dealt with through an expansion such as that found in [8] which can be used to generate invertible parameters $\mathbf{A}, \mathbf{B}$ and $\mathbf{Q}$ by considering multiple observations as a single observation. A non-homogenous sensor network where the observation $\mathbf{B}$ is described in a statistical sense is more complex as the use of $\mathbf{E}(\mathbf{B})$ directly yields a lower bound of $\mathbf{P}$. The requirement that sensor transmissions be equispaced can also be relaxed to merely describing the sensor observation process as probabilistic process with an exponential distribution between observations and the observation function incorporated directly.

Theorem 2 *Consider as an objective function the norm* $\mathcal{M}\left(\overline{\mathbf{P}(a,N)}\right) : \overline{\mathbf{P}(a,N)} \in \mathbb{S}^{m\times m}$ *. The characteristic observation loss function*

$$\alpha(N,\kappa) = \left(1+\frac{1}{c}\right)\left(\frac{\kappa}{N}+1\right)^{-c} - \frac{1}{c} \quad (3)$$

where

$$c = \frac{\operatorname{tr}(\mathbf{F}\mathbf{C}_2 + \mathbf{C}_2\mathbf{F}' - \mathbf{C}_1)}{\operatorname{tr}\mathbf{C}_1 - (m-1)\lambda_{\min}(\mathbf{C}_1)} - \frac{(m-1)\lambda_{\min}(\mathbf{F}\mathbf{C}_2 + \mathbf{C}_2\mathbf{F}' - \mathbf{C}_1)}{\operatorname{tr}\mathbf{C}_1 - (m-1)\lambda_{\min}(\mathbf{C}_1)} \quad (4)$$

and

$$\begin{aligned} \mathbf{C}_1 &= \mathbf{F}^2\mathbf{P} + 2\mathbf{F}\mathbf{P}\mathbf{F}' + \mathbf{P}(\mathbf{F}')^2 + \mathbf{F}\boldsymbol{Q}_c + \boldsymbol{Q}_c\mathbf{F}' \\ &\quad + \mathbf{A}\left(\mathbf{F}^2\mathbf{B}^{-1} + 2\mathbf{F}\mathbf{B}^{-1}\mathbf{F}' + \mathbf{B}^{-1}(\mathbf{F}')^2\right)\mathbf{A}' \\ \mathbf{C}_2 &= \mathbf{A}\boldsymbol{Q}_c\mathbf{A}' \end{aligned} \quad (5)$$

describes a contour $\alpha(N,\kappa)$ *on* $\mathcal{M}\left(\overline{\mathbf{P}(a,N)}\right)$ *that satisfies the inequality*

$$\mathcal{M}\left(\overline{\mathbf{P}(\alpha(N,\kappa),N)}\right) \le \psi(\kappa) \quad (6)$$

where κ *is a positive constant that parameterises the solutions of (6) and* $\mathbf{C}_1$ *and* $\mathbf{C}_2$ *are positive (semi) definite and* c *is non-negative.*

The characteristic observation loss function summarises the behaviour of the linear system with increasing sensor count and provides an analytical tool to determine the behaviour of the system under alternative loss functions and determine if there is a loss function that satisfies (6) for all $N \in (1,\infty]$. Note that κ parameterises the characteristic observation loss function. By setting κ a particular value of the objective function is selected and particular observation loss function that passes through $\psi(\kappa)$ generated. Eq. (6) is a particularly important relationship for the design of a system as it ensures that increasing the number of sensors always improves the system performance as determined by the objective function $\mathcal{M}\left(\overline{\mathbf{P}(a,N)}\right)$. The following theorem uses the previous two theorems to determine if an observation system is guaranteed to have "nice" behaviour.

Theorem 3 *Satisfying the inequality*

$$N\frac{da(N)}{dN} < (ca(N)+1)\left(1 - \left(\frac{ca(N)+1}{(c+1)}\right)^{\frac{1}{c}}\right) \quad (7)$$

for some observation loss function $a(N)$ *where* c *is defined in (4) is sufficient to satisfy*

$$\frac{d\mathcal{M}\left(\overline{\mathbf{P}(a(N),N)}\right)}{dN} \le 0$$

and consequently indicates an improvement in the objective function for increasing N*. Furthermore any observation loss function that satisfies*

$$a(N) \le \left(1+\frac{1}{c}\right)\left(\frac{\sqrt[c]{(c+1)}-1}{N}+1\right)^{-c} - \frac{1}{c} \quad (8)$$

describes a region where $\mathcal{M}\left(\overline{\mathbf{P}(a(N),N)}\right) < \overline{\mathbf{P}(0,1)}$ *and assures that the system benefits from additional sensors.*

This theorem allows two important practical questions to be answered. Firstly is there a point where increasing the number of sensors is not of benefit to the system, and secondly, at this point does the system perform better than a single sensor system without losses. Note that the bound in [1] is extensible under Condition 1 to yield an equivalent bound to (8)

$$a(N) < \frac{1}{\lambda_{\max}\left(\mathbf{A}^2\right)^{\frac{1}{N}}} \tag{9}$$

but this bound is looser as it only satisfies $\mathcal{M}\left(\lim_{k\to\infty} \mathsf{E}\left(\mathbf{P}\left(t_k\right)\right)\right) < \infty$ where the expectation is taken over $\Gamma_{n,k}$ and the observation loss process is described by an observation loss function $a(N)$. Using the bound defined by (8) assures that not only is the expected covariance bounded, but that it is better than the expected covariance of the ideal single sensor system.

4 Discussion

Theorem 3 provides a mechanism for evaluating the sensitivity of a system to variations in the number of sensors and the systems observation loss function. Two questions are frequently raised when designing a system. Firstly, under what conditions does the system perform better than a single sensor system without communications, i.e. under what conditions is it useful to have the complexity of a multi-sensor system. Secondly, under what conditions does adding or removing sensors increase or decrease some performance metric such as the mean squared estimator error.

Theorem 3 is used to certify that a particular system will always improve in performance for increasing N provided that the observation loss function $a(N)$ satisfies (7). In many systems the characteristic observation loss function is a property not within the control of the designer, all that is available is the control of the communications system. Each communications system structure (queues, channel bandwidths, error rates) imposes a particular $a(N)$ that is characteristic of that communication system. The use of (7) allows the area of safe operation to be evaluated for a particular communications system and be expressed as an analytic expression, rather than through the use of numerical methods. This allows sensitivity and failure mode effects analysis to be performed.

The final advantage Theorem 3 provides is that it allows the evaluation of a system against a criteria of a finite metric, rather than just demonstrating that the system is bounded. This is more useful as it provides information on how well the system performs, rather than just the presence or absence of convergence. As such it is more useful than the bounds provided in [1] and [4] that merely demonstrate convergence. This is particularly useful in comparing between single and multi-sensor systems by demonstrating under what conditions the distributed multi-sensor system fails to outperform the local single sensor system.

The use of Theorem 3 allows an extension to the concept introduced by Adlakha et. al. in [5]. In Adlakha's paper the concept of increasing the sampling rate is introduced as a method of reducing the estimator covariance when using a fixed capacity link. Unfortunately this paper only considers the scalar case and a particular round robin packet loss model. The interesting outcome from [5] was the observation that in some cases sampling more finely and then sending a subset of those samples is sufficient to improve system performance. While Adlakha derived analytic results they are only applicable to the scalar system. Theorem 3 provides an alternative method to obtain similar results. Examination of (8) with $a(N) = 1 - k/N$ reveals that as $N \to \infty$ the performance of the system described by Adlakha approaches that of the single sensor system from below when the system satisfies Conditions 2 and 3. It is also apparent that fewer sensors is still better than a large number, ideally $k/N = 1$.

An important point to note on these methods is that they require access only to the basic system parameters $\{\mathbf{F}, \mathbf{B}, \mathbf{Q}_c\}$ and the solution of the Riccati Equation $\mathbf{P}$ for the ideal system. By forming the bounds in terms of the known Riccati equation solution the equations are considerably more tractable than other approaches that are written directly in terms of the basic system parameters.

5 Example

To demonstrate the utility of results an example will be considered. The goal is to determine if two different observation processes will causes improvement in the systems objective function if additional sensors are added. The first observation process is described by $1 - k/N : N \geq k$ and is consistent with a form of network congestion. The second observation process is $1 - \exp(-N/k)$. The system properties are

$$\mathbf{F} = \begin{bmatrix} .1 & 1 \\ 0 & .1 \end{bmatrix} \quad \mathbf{Q}_c = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Firstly the parameters of the characteristic observation loss function are generated using (9) yielding $c = .0048$. The upper bound of this system is plotted in Figure 1 using (1) and is compared with the bound obtained by solving $\mathcal{M}\left(\lim_{k\to\infty} \mathsf{E}\left(\mathbf{P}\left(t_k\right)\right)\right)$ where $\mathcal{M}(P) = \|P\|_2$ .Note that the upper bound nicely envelopes $\mathcal{M}\left(\lim_{k\to\infty} \mathsf{E}\left(\mathbf{P}\left(t_k\right)\right)\right)$, capturing the behaviour at the extremes of a and N. In Figure 2 the contour lines of the MARE are plotted and compared with the contours of the characteristic observation loss function bound generated by (3) for various κ. Note that the contours generated by the application of (3) always cross the true contours in a direction that

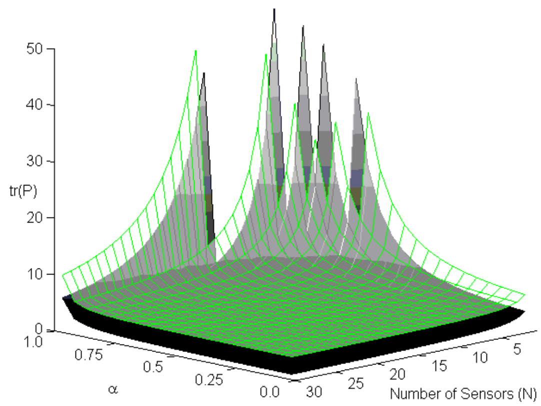

Figure 1: Comparison of Upper Bound and MARE Solution

causes the objective function to decrease. This is due to (3) being an inequality which can only approach an equality as $N \to \infty$. The Sinopoli Bound is given by (9) and indicates the upper bound for a stable system and is shown in red. The bound given by Theorem 3 (8) is given in green. It constrains the observation loss function more tightly as it bounds functions that actually improve upon the single sensor system, rather than are merely stable. Consider the observation process $1 - k/N$. Applying inequality (7) and re-arranging yields

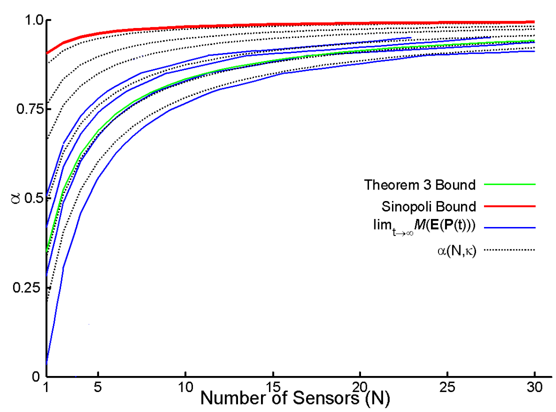

Figure 2: Contours of $\mathcal{M}\left(\lim_{k\to\infty} \mathsf{E}\left(\mathbf{P}\left(t_k\right)\right)\right)$, $\alpha\left(N, \kappa\right)$, Theorem 3 and the Sinopoli Bound

$$\left(1 - \frac{k}{N}\right) > \left(1 - \frac{c}{c+1}\frac{k}{N}\right)^{\frac{c+1}{c}} \tag{10}$$

(10) has the form $(1 - X) > (1 - X/Y)^{Y}$ and is not satisfied for all $Y \geq 1, 0 < X \leq 1$. This indicates that there is no $c > 0$ or $k/N \in (0,1)$ that will cause an increase in N to be beneficial. This is demonstrated in Figure 3 by evaluating the MARE numerically and plotting $\lim_{k\to\infty} \mathcal{M}\left(\mathsf{E}\left(\mathbf{P}\left(t_k\right)\right)\right)$ for differing k for the system above. As can be seen increasing N reduces system performance, albeit by diminishing amounts as N becomes larger. A final remark is that the obser-

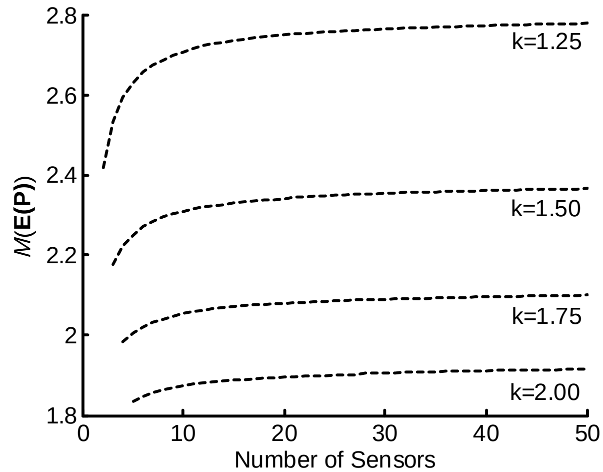

Figure 3: $\lim_{k\to\infty} \mathcal{M}\left(\mathsf{E}\left(\mathbf{P}\left(t_k\right)\right)\right)$ where $a\left(N\right) = 1 - k/N : N \geq k$ for $k = \{1, 1.25, 1.5, 1.75, 2\}$

vation loss process described is equivalent to a link with capacity k messages per scan interval with a single message buffer. It demonstrates that the system performance will always be reduced if the link capacity does not exceed the number of sensors.

Consider as a second example the observation process $1 - \exp\left(-N/k\right)$. Applying inequality (7) yields

$$\frac{N}{k}e^{-N/k} < \left(\left(1 - \exp\left(-N/k\right)\right)c + 1\right) \times \left(1 - \left(\frac{c\left(1 - \exp\left(-N/k\right)\right) + 1}{c+1}\right)^{1/c}\right) \tag{11}$$

Note that c has little impact on the solution region, which is dominated by the N/k term. For $c = 0$ (11) is satisfied when $0 < N/k < 0.806$ while as $c \to \infty$ then (11) is satisfied for $0 < N/k < 0.693$. This indicates that there is no benefit in increasing the number of sensors beyond $k \simeq N/0.7$ when using the observation process $1 - \exp\left(-N/k\right)$. Using (8) we can determine that consequently the minimum number of sensors required to achieve performance better than a single ideal sensor is 3. This is demonstrated in Figure 4 by plotting $\lim_{k\to\infty} \mathcal{M}\left(\mathsf{E}\left(\mathbf{P}\left(t_k\right)\right)\right)$ for differing k for the system. Note the initial improvement achieved by increasing the number of sensors, as predicted by (11) followed by deterioration. The trace of $k = 4$ shows a minimal improvement until 3 sensors are present. Of interest is that the behaviour is largely independent of c, the characteristics of the system being observed.

This investigation would have been impractical without using analytic tools, requiring large numbers of numerical plots. As the results are analytic we can

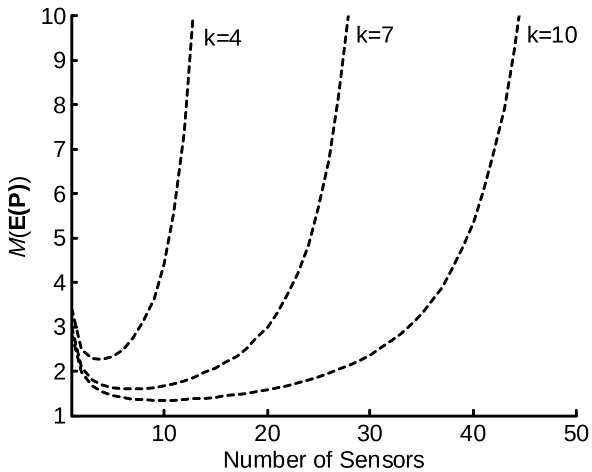

Figure 4: $\lim_{k\to\infty} \mathcal{M}\left(\mathsf{E}\left(\mathbf{P}\left(t_k\right)\right)\right)$ where $a\left(N\right) = 1 - \exp\left(-n/k\right)$ for $k = \{4, 7, 10\}$

draw conclusions about general linear systems using the same communications system. Finally we can study the system behaviour subject to variations in c and deduce the system sensitivity to changes in some system parameters.

6 Conclusion

Using the characteristic observation function to determining performance bounds for a multi-sensor system under arbitrary correlated lost observations provides a useful analytic method for determining the operational characteristics of a sensor system without resorting to numerical methods. Performance bounds and sensitivity to network parameters can be established when an analytic expression for the message loss function is available.

Further work is required to sharpen the bound and to relax the analytic requirements. Both the requirement for $\mathbf{B}$ invertibility and $\mathbf{F}$ having non-zero eigenvalues are under investigation. Furthermore it is desirable to describe the sensor characteristics in terms of statistical parameters rather than requiring all sensors to be identical. All extensions should preserve the analytic nature of the bounds to allow easier understanding of the system characteristics.

7 Appendix

7.1 Proof of Theorem 1

Proof. The upper bound is determined by noting that the system can be divided into two sections. Firstly an observation process that is performed by the N sensors, and then a Bernoulli loss process which deletes some fraction of the observations. It is shown in [1] that the MARE (Modified Algebraic Riccati Equation) provides an upper bound to $\mathsf{E}\left(\mathbf{P}\right)$

$$\begin{aligned} \mathsf{E}\left(\mathbf{P}\right) &< \mathbf{P}_a \\ &= \mathbf{A}\left(\left(1-a\right)\left(\mathbf{P}_a^{-1}+\mathbf{B}\right)^{-1} + a\mathbf{P}_a\right)\mathbf{A}' + \mathbf{Q} \end{aligned} \tag{12}$$

for a linear system that is sampled in the presence of a Bernoulli loss process. Let $\mathbf{P}$ be the solution of the equivalent system without loss, the Riccati equation (2). Assume that the solution of (12) can be written in the form

$$\mathbf{P}_a = \mathbf{P} + \boldsymbol{\epsilon} \tag{13}$$

where $\boldsymbol{\epsilon}$ is some positive definite error term. By substituting (13) into (12) and taking advantage of the properties of $\mathbf{P}$ (2) an expression for $\boldsymbol{\epsilon}$ can be obtained

$$\begin{aligned} \operatorname{vec}\epsilon &= \left(1-a\right)\left(\mathbf{I} - a\left(\mathcal{A}\right)\right)^{-1}\left(\mathcal{A}\right)\times \\ &\operatorname{vec}\begin{pmatrix} \left(\mathbf{I}+\mathbf{BP}\right)\left(\boldsymbol{\epsilon}\right)^{-1}\left(\mathbf{I}+\mathbf{PB}\right) \\ +\mathbf{B}+\mathbf{BPB} \end{pmatrix}^{-1} \\ &+ a\left(\mathbf{I} - a\left(\mathcal{A}\right)\right)^{-1}\left(\left(\mathcal{A}-\mathbf{I}\right)\operatorname{vec}\mathbf{P} + \operatorname{vec}\boldsymbol{Q}_c\right) \end{aligned} \tag{14}$$

In [1] it is shown that $\mathbf{P}_a$ is finite provided

$$a < \frac{1}{\lambda_{\max}\left(\mathbf{A}\right)^2} \tag{15}$$

Assuming that (15) is satisfied then $\mathbf{P}_a$ is finite and by (13) $\boldsymbol{\epsilon}$ must be finite as well. i.e. $\lambda_{\max}\left(\boldsymbol{\epsilon}\right) < \infty$. By Condition 2 then

$$\begin{aligned} 0 &< \begin{pmatrix} \left(\mathbf{I}+\mathbf{BP}\right)\left(\boldsymbol{\epsilon}\right)^{-1}\left(\mathbf{I}+\mathbf{PB}\right) \\ +\mathbf{B}+\mathbf{BPB} \end{pmatrix}^{-1} \\ &< \left(\mathbf{B}+\mathbf{BPB}\right)^{-1} < \mathbf{B}^{-1} \end{aligned} \tag{16}$$

Using (16) and substituting back into (13) and (14) the relationship for bounds on $\mathbf{P}_a$

$$\begin{aligned} \operatorname{vec}\mathbf{P}_a &< \operatorname{vec}\overline{\mathbf{P}_a} \\ \operatorname{vec}\overline{\mathbf{P}_a} &= \left(\mathbf{I} - a\left(\mathcal{A}\right)\right)^{-1}\times \\ &\quad \left(\left(1-a\right)\operatorname{vec}\mathbf{P} + a\operatorname{vec}\boldsymbol{Q}_c \right. \\ &\quad \left. + \left(\mathcal{A}\right)\left(1-a\right)\operatorname{vec}\left(\mathbf{B}\right)^{-1}\right) \end{aligned} \tag{17}$$

is obtained. Now consider the multiple sensor system. By Condition 1 and 2 the case of N sensors is equivalent to a single sensor sampling at N times the rate. We shall write the expression for the covariance (not expected covariance as the sampling process is not stochastic) in terms of the simpler single sensor case. Under the previously outlined conditions the expression for the estimator covariance is a Riccati equation

$$\mathbf{P}_N = \mathbf{A}_N\left(\mathbf{P}_N^{-1}+\mathbf{B}\right)^{-1}\mathbf{A}_N' + \mathbf{Q}_N \tag{18}$$

Label $\mathbf{P}$ as the solution to (2), the single sensor case of the Riccati equation. Assume that $\operatorname{vec}\mathbf{P}_N$ can be written in the form

$$\begin{aligned} \operatorname{vec}\mathbf{P}_N &= \mathcal{A}_N \operatorname{vec}\left(\left(\mathbf{P}^{-1}+\mathbf{B}\right)^{-1} + \boldsymbol{\epsilon}\right) \\ &+ \operatorname{vec}\left(\mathbf{Q}_N\right) \end{aligned} \tag{19}$$

Substitution of (19) into (18) forms the equation

$$\begin{aligned}&\mathcal{A}_N \operatorname{vec}\left(\left(\mathbf{P}^{-1}+\mathbf{B}\right)^{-1}+\boldsymbol{\epsilon}\right)+\operatorname{vec}\left(\mathbf{Q}_N\right)\\&=\mathcal{A}_N \operatorname{vec}\left(\mathbf{P}_N^{-1}+\mathbf{B}\right)^{-1}+\operatorname{vec}\left(\mathbf{Q}_N\right)\end{aligned}$$

but noting that $\mathbf{P}_N \leq \mathbf{P}$ and hence $\left(\mathbf{P}_N^{-1}+\mathbf{B}\right)^{-1} \leq \left(\mathbf{P}^{-1}+\mathbf{B}\right)^{-1}$ then $\left(\left(\mathbf{P}_N^{-1}+\mathbf{B}\right)^{-1}-\left(\mathbf{P}^{-1}+\mathbf{B}\right)^{-1}\right) \leq \boldsymbol{\epsilon}$ and leads to the conclusion that $\boldsymbol{\epsilon}$ must not be positive definite and hence (18) must be an upper bound when $\lambda_{\max}(\boldsymbol{\epsilon})$ is set to 0. Substitution of the expression for $\mathbf{Q}_N$ yields the upper bound for $\mathbf{P}_N$ in terms of the solution of (2) and the system parameters.

$$\begin{aligned}\mathbf{P}_N<\overline{\mathbf{P}_N}=\mathcal{A}^{\frac{1}{N}-1}&\left(\mathbf{G}^{-1}\operatorname{vec}\boldsymbol{Q}+\operatorname{vec}\mathbf{P}\right)\\&-\mathbf{G}^{-1}\operatorname{vec}\boldsymbol{Q}_c\end{aligned}$$

Substituting $\overline{\mathbf{P}_N}$, $\mathbf{A}_N$ and $\mathbf{Q}_N$ in place of $\mathbf{P}, \mathbf{A}$ and $\mathbf{Q}$ respectively into (17) provides $\overline{\mathbf{P}(a, N)}$ the upper bound of the expected covariance of the state estimator $\mathbf{x}$ for N sensors with dropping probability a and concludes the proof. ■

7.2 Proof of Theorem 2

Firstly some results on vec inequalities are required.

Lemma 1 *A sufficient condition to satisfy*

$$\left(\operatorname{vec}\mathbf{A}'\right)'\mathbf{B}\mathbf{C}\operatorname{vec}\mathbf{D}<0 \tag{20}$$

where $\operatorname{vec}^{-1}\left(\mathbf{B}'\operatorname{vec}\mathbf{A}'\right)$ *is positive semidefinite and* $\lambda_{\min}\left(\operatorname{vec}^{-1}\left(\mathbf{C}\operatorname{vec}\mathbf{D}\right)\right)<0$ *is given by*

$$(\operatorname{vec} I)'\mathbf{C}\operatorname{vec}\mathbf{D} \leq (m-1)\lambda_{\min}\left(\operatorname{vec}^{-1}\left(\mathbf{C}\operatorname{vec}\mathbf{D}\right)\right) \tag{21}$$

Proof. Use Theorem 3 in [9] and the properties of the vec operator. ■

Lemma 2 *If*

$$\operatorname{vec}\mathbf{X}=(\mathbf{I}-\mathbf{A}\otimes\mathbf{A})^{-1}\operatorname{vec}\mathbf{Y} \tag{22}$$

then $\mathbf{X}$ *is positive (semi)definite if* $\mathbf{Y}$ *is positive (semi)definite and* $\left|\lambda_{\max}(\mathbf{A})\right|<0$.

If

$$\operatorname{vec}\mathbf{X}=(\mathbf{F}\otimes\mathbf{I}+\mathbf{I}\otimes\mathbf{F})^{-1}\operatorname{vec}\mathbf{Y} \tag{23}$$

then $\mathbf{X}$ *is positive (semi)definite if* $\mathbf{Y}$ *is positive (semi)definite and* $\lambda_{\min}(\mathbf{F})>0$

Proof. Note that (22) is the vec of the discrete time Lyapunov equation while (23) is the vec of the continuous time form. The lemma follows from the standard results for Lyapunov equations. ■

The main proof of Theorem 2 follows.

Proof. A constant metric contour can be written as

$$\mathcal{M}\left(\overline{\mathbf{P}(a(N,\kappa), N)}\right)=\psi(\kappa)$$

where $a(N,\kappa)$ is a parametric function of parameters N and κ and $\psi(\kappa)$ is constant with respect to κ. To obtain the parametric function $a(N,\kappa)$ that satisfies a contour we require that

$$\begin{aligned}0=\frac{d\psi(\kappa)}{dN}&=\frac{d\mathcal{M}\left(\overline{\mathbf{P}(a(N,\kappa), N)}\right)}{dN}\\&=\frac{d\mathcal{M}(\overline{\mathbf{P}})}{d\overline{\mathbf{P}}}\left(\begin{array}{c}\frac{\partial\overline{\mathbf{P}(a,N)}}{\partial a}\frac{\partial a(N,\kappa)}{\partial N}\\+\frac{\partial\overline{\mathbf{P}(a,N)}}{\partial N}\end{array}\right)\end{aligned} \tag{24}$$

Take the partial matrix derivatives $\frac{\partial\overline{\mathbf{P}(a(N,\kappa),N)}}{\partial a(N,\kappa)}$ and $\frac{\partial\overline{\mathbf{P}(a(N,\kappa),N)}}{\partial N}$. As we are free to choose any structure of $\psi(\kappa)$ we can select the structure of the parameterisation of a and so select $a(N/\kappa)$. Starting with (1), taking the derivatives as in (24) and substituting $\mathbf{K}=\left(\mathcal{A}^{\frac{\kappa}{N}}-\mathbf{I}\right)$ we obtain the first order differential equation in a and N.

$$\begin{aligned}&\left(\frac{d\mathcal{M}(\overline{\mathbf{P}})}{d\operatorname{vec}\overline{\mathbf{P}}}\right)'\left(\mathcal{A}^{-\frac{\kappa}{N}}-a\mathbf{I}\right)^{-2}\mathcal{A}^{-\frac{\kappa}{N}-1}\mathbf{G}^{-1}\times\\&\left(\left(\begin{array}{c}\mathbf{KG}\operatorname{vec}\mathbf{P}\\+(\mathcal{A}+\mathbf{K})\operatorname{vec}\boldsymbol{Q}_c\\+\mathbf{K}\mathcal{A}\mathbf{G}\operatorname{vec}\mathbf{B}^{-1}\end{array}\right)\frac{da}{dN}\right.\\&\left.-\frac{t_s\kappa}{N^2}\left(\begin{array}{c}(1-a)\mathbf{G}^2\operatorname{vec}\mathbf{P}\\+(a\mathcal{A}+(1-a)\mathbf{I})\mathbf{G}\operatorname{vec}\boldsymbol{Q}_c\\+(1-a)\mathcal{A}\mathbf{G}^2\operatorname{vec}\mathbf{B}^{-1}\end{array}\right)\right)=0\end{aligned} \tag{25}$$

Note that by the definition of $\mathbf{A}$ that all the eigenvalues of $\mathbf{A}$ must be non-negative. By the definition of a measure that $\mathcal{M}(\overline{\mathbf{P}})$ must be a convex (not concave) function and hence $d\mathcal{M}(\overline{\mathbf{P}})/d\mathbf{P} : \mathbf{P}\in\mathbb{S}_+^{m\times m}$ must form a positive (semi)definite matrix. By Lemma 2 the common terms of (25) are the vec of a positive (semi) definite term (if $\lambda_{\min}(\mathbf{F})>0$). By Lemma 1 a sufficient condition to satisfy (25) where one term is positive (semi) definite is to set

$$\begin{aligned}&\operatorname{vec}(\mathbf{I})'\left(\left(\begin{array}{c}\mathbf{KG}\operatorname{vec}\mathbf{P}\\+(\mathcal{A}+\mathbf{K})\operatorname{vec}\boldsymbol{Q}_c\\+\mathbf{K}\mathcal{A}\mathbf{G}\operatorname{vec}\mathbf{B}^{-1}\end{array}\right)\frac{da}{dN}\right.\\&\left.-\frac{t_s\kappa}{N^2}\left(\begin{array}{c}(1-a)\mathbf{G}^2\operatorname{vec}\mathbf{P}\\+(a\mathcal{A}+(1-a)\mathbf{I})\mathbf{G}\operatorname{vec}\boldsymbol{Q}_c\\+(1-a)\mathcal{A}\mathbf{G}^2\operatorname{vec}\mathbf{B}^{-1}\end{array}\right)\right)\\&\leq(m-1)\times\\&\lambda_{\min}\left(\operatorname{vec}^{-1}\left(\begin{array}{c}\mathbf{KG}\operatorname{vec}\mathbf{P}\\+(\mathcal{A}+\mathbf{K})\operatorname{vec}\boldsymbol{Q}_c\\+\mathbf{K}\mathcal{A}\mathbf{G}\operatorname{vec}\mathbf{B}^{-1}\end{array}\right)\frac{da}{dN}\right.\\&\left.-\frac{t_s\kappa}{N^2}\operatorname{vec}^{-1}\left(\begin{array}{c}(1-a)\mathbf{G}^2\operatorname{vec}\mathbf{P}\\+(a\mathcal{A}+(1-a)\mathbf{I})\mathbf{G}\operatorname{vec}\boldsymbol{Q}_c\\+(1-a)\mathcal{A}\mathbf{G}^2\operatorname{vec}\mathbf{B}^{-1}\end{array}\right)\right)\end{aligned} \tag{26}$$

Noting that

$$\mathbf{K} = \lim_{\kappa\to\infty} \frac{\kappa}{N} \ln \mathcal{A} \quad (27)$$
$$= \lim_{\kappa\to\infty} \frac{\kappa t_s}{N}\mathbf{G} > \frac{\kappa t_s}{N}\mathbf{G}$$

and substituting (27) , (5) and (4) into (26) and using eigenvalue inequalities yields the ordinary differential inequation

$$N\left(1 + N\frac{k}{\kappa}\right)\frac{da}{dN} \leq (1 + ac) \quad (28)$$

where (28) can be used to provide a set of solutions $a \leq \alpha(N/\kappa)$ that satisfy (26) and hence inequality (6). The upper bound of (28) is of the form

$$j\left(1 + k\frac{N}{\kappa}\right)\frac{d\alpha}{dj} = (1 + \alpha c) \quad (29)$$

and has solutions

$$\alpha(j, \kappa) = \left(c_B\left(\frac{\kappa}{Nk} + 1\right)^{-c} - \frac{1}{c}\right) \quad (30)$$

where c_B is a constant set by boundary conditions. Hence the family of curves described by (30) represent bounds for constant objective function contours. To establish boundary conditions consider the behaviour of (1) as $N \to \infty$. It is clear that $\mathcal{M}(\mathbf{P}(a, N))$ is discontinuous at $a = 1$ as $N \to \infty$ and correspondingly all contours must converge at this point. This implies a boundary condition $\alpha(N, \kappa) \to 1$ as $N \to \infty$. Solving this boundary condition gives $c_B = 1 + \frac{1}{c}$. Pulling the constant k into the parametric value κ as $\Psi(\kappa)$ is an arbitrary function allows us to write

$$\mathcal{M}\left(\overline{\mathbf{P}(a(N, \kappa), N)}\right) < \psi(\kappa)$$

is given by

$$\alpha(N, \kappa) = \left(1 + \frac{1}{c}\right)\left(\frac{\kappa}{N} + 1\right)^{-c} - \frac{1}{c} \quad (31)$$

Finally note that $\mathbf{C}_2$ is clearly positive definite as $\boldsymbol{Q}_c$ must be positive definite. $\mathbf{C}_1$ is positive definite by Condition 3. c is positive as $\operatorname{tr}(X) - (m-1)\lambda_{\min}(X)$ must be a positive value for all $X \in \mathbb{S}^{m\times m}$. ■

7.3 Proof of Theorem 3

Proof. Firstly note that $\exists\kappa \in \mathbb{R}^+ : \mathcal{M}\left(\overline{\mathbf{P}(a(N), N)}\right) = \mathcal{M}\left(\overline{\mathbf{P}(a(N, \kappa), N)}\right) = \Psi(\kappa)$. where φ is a function of a and N. Hence $\frac{d\mathcal{M}\left(\overline{\mathbf{P}(a(N),N)}\right)}{dN} = \frac{d\Psi(\kappa)}{d\kappa}\frac{d\kappa}{dN}$. Note that $\frac{d\Psi(\kappa)}{d\kappa} = \frac{\mathcal{M}\left(\overline{\mathbf{P}(\alpha(N,\kappa),N)}\right)}{d\kappa} = \frac{\partial\mathcal{M}\left(\overline{\mathbf{P}(\alpha(N,\kappa),N)}\right)}{\partial \operatorname{vec}\mathbf{P}'}\frac{\partial \operatorname{vec}\mathbf{P}}{d\alpha}\frac{d\alpha}{d\kappa}$. Using the results in Theorem (2) and the properties of an objective function $\frac{\partial\mathcal{M}\left(\overline{\mathbf{P}(\alpha(N,\kappa),N)}\right)}{d \operatorname{vec}\mathbf{P}'}\frac{\partial \operatorname{vec}\mathbf{P}}{d\alpha}$ must be non-negative for all $N \geq 1$, $\kappa \geq 0$. Taking the derivative of (31) with respect to κ indicates that $\frac{d\alpha}{d\kappa}$ must be negative. Hence $\frac{d\Psi(\kappa)}{d\kappa}$ must be negative. Consequently $\frac{d\kappa}{dN} > 0 \Longrightarrow \frac{d\Psi(\kappa)}{d\kappa}\frac{d\kappa}{dN} < 0 \Longrightarrow \frac{d\mathcal{M}\left(\overline{\mathbf{P}(a(N),N)}\right)}{dN} < 0$. The intuitive understanding is that κ describes a constant contour line with $\Psi(0) = \infty$ and $\Psi(\kappa)$ decreasing as κ increases. The second bound follows by solving κ for $\alpha(1, \kappa) = 0$. Note that for the ideal single sensor system that the probability of losing a message is always 0. ■

References

[1] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[2] Y. Mo and B. Sinopoli, "A characterization of the critical value for Kalman filtering with intermittent observations," in *Proceedings of IEEE Conference on Decision and Control*, 2008, pp. 2692–2697.

[3] L. Xiangheng and A. Goldsmith, "Kalman filtering with partial observation losses," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, vol. 4, 2004, pp. 4180–4186 Vol.4.

[4] A. S. Matveev and A. V. Savkin, "The problem of state estimation via asynchronous communication channels with irregular transmission times," *IEEE Transactions on Automatic Control*, vol. 48, no. 4, pp. 670–676, 2003.

[5] S. Adlakha, B. Sinopoli, and A. Goldsmith, "Optimal sensing rate for estimation over shared communication links," in *Proceedings of the American Control Conference*, 2007, pp. 5043–5045.

[6] B. Zhu, B. Sinopoli, K. Poolla, and S. Sastry, "Estimation over wireless sensor networks," in *Proceedings of the American Control Conference*, 2007, pp. 2732–2737.

[7] R. Ambrosino, B. Sinopoli, K. Poolla, and S. S. Sastry, "Optimal sensor density for remote estimation over wireless sensor networks," in *Proceedings of the Forty Sixth Anual Allerton Conference*, vol. 142, 2008.

[8] H. Choi, "A closed form solution of discrete time Riccatti equations," in *Proceedings of ISL Winter Workshop*, 1989, pp. 65–76.

[9] F. Yuguang, K. A. Loparo, and F. Xiangbo, "Inequalities for the trace of matrix product," *IEEE Transactions on Automatic Control*, vol. 39, no. 12, pp. 2489–2490, 1994, 0018-9286.