which, using the triangle inequality, is bounded by

$$L = |\sum_{i=1}^{N} L_i| \leq \sum_{i=1}^{N} |L_i| \tag{11}$$

$$L \leq \frac{1}{K}\sum_{i=1}^{N} |w_i|. \tag{12}$$

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for helping improve the quality of the presentation of this paper. They also thank Robust Analysis, Inc. both for funding this research and for providing a stimulating environment for questioning the *status quo*. Fast, numerically efficient implementations of the Branch-and-Bound search algorithms introduced in this paper have been successfully integrated as part of the commercial release of Robust Analysis's STABLE Signal Processing Toolbox [16].

REFERENCES

[1] J. G. Gonzalez and G. R. Arce, "Statistically-efficient filtering in impulsive environments: Weighted myriad filters," *EURASIP J. Appl. Signal Process.*, vol. 2002, no. 1, pp. 4–20, 2002.

[2] J. G. Gonzalez, "Robust techniques for wireless communnications in non-Gaussian environments," Ph.D. dissertation, Univ. Delaware, , 1997.

[3] G. R. Arce, *Nonlinear Signal Processing: A Statistical Approach*. New York: Wiley , 2005.

[4] E. Grün, M. S. Hanner, S. B. Peschke, T. Müller, H. Boehnhardt, T. Y. Brooke, H. Campins, J. Crovisier, C. Delahodde, I. Heinrichsen, H. U. Keller, R. F. Knacke, H. Krüger, P. Lamy, C. Leinert, D. Lemke, C. M. Lisse, M. Müller, D. J. Osip, M. Solc, M. Stickel, M. Sykes, V. Vanysek, and J. Zarnecki, "Broadband infrared photometry of comet Hale-Bopp with ISOPHOT," *Astron. Astrophys.*, vol. 377, pp. 1098–1118, Oct. 2001.

[5] M. E. Hassouni and H. Cherifi, "Spatio-temporal weighted myriad filter for alpha-stable noise removal in video sequences," in *ICSES04 Int. Conf. Signals and Electronic Syst.*, Poznan, Poland, Sep. 2004.

[6] J. G. Gonzalez, S. Manji, J. Paredes, and W. Ahmed, "Systems and methods for adaptively decoding transmitted frames," U. S. Patent Appl. 20030223512, Serial No. 161812, U.S. Class No. 375/316; 375/341, Int. Class No. H03K 009/00; H04L 027/06, Jun. 2002.

[7] R. Ganguli and V. P. Surender, "Adaptive myriad filter for improved gas turbine condition monitoring using transient data," in *ASME Turbo Expo 2004*, Vienna, Austria, Jun. 2004.

[8] A. Hamza and H. Krim, "Image denoising: A nonlinear robust statistical approach," *IEEE Trans. Signal Process.*, vol. 49, pp. 3045–3054, Dec. 2001.

[9] T. Pander, "Robust filtering of ECG signal with a myriad filter," *Polish J. Med. Phys. Eng.*, vol. 9, no. 1, pp. 17–30, 2003.

[10] R. C Nunez, "Novel techniques for the acceleration of non-Gaussian signal processing algorithms," M.S. thesis, Elect. Comput. Eng. Dept., Univ. of Delaware, Newark, DE, 2007.

[11] S. Kalluri and G. R. Arce, "Fast algorithms for weighted myriad computation by fixed-point search," *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 159–171, 2000.

[12] S. Kalluri and G. R. Arce, "Adaptive weighted myriad filter algorithms for robust signal processing in alpha-stable noise environments," *IEEE Trans. Signal Process.*, vol. 46, no. 2, pp. 322–334, 1998.

[13] S. Kalluri, "Nonlinear adaptive algorithms for robust signal processing and communications in impulsive environments," Ph.D. dissertation, Univ. Delaware, , 1998.

[14] S. Kalluri and G. R. Arce, "Robust frequency-selective filtering using weighted myriad filters admitting real-valued weights," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2721–2733, Nov. 2001.

[15] G. R. Arce, "A general weighted median filter structure admitting negative weights," *IEEE Trans. Signal Process.*, vol. 46, no. 12, pp. 3195–3205, Dec. 1998.

[16] User Manual for the STABLE 5.0 Library. Robust Analysis, Inc., 2007 [Online]. Available: http://www.robustanalysis.com

[17] J. D. Pinter, Global Optimization in Action. Norwell, MA, Kluwer Academic, 1996.

[18] R. Horst, P. Pardalos, and N. Thoai, *Introduction to Global Optimization*. Norwell, MA: Kluwer Academic, 1995.

[19] J. G. Gonzalez, D. L. Lau, and G. R. Arce, "Towards a general theory of robust nonlinear filtering: Selection filters," in *IEEE Int. Conf. Acoust., Speech Signal Process., ICASSP'97*, Munich, Germany, Apr. 1997, vol. 5, pp. 3837–3840.

On Kalman Smoothing With Random Packet Loss

Alex S. Leong, Subhrakanti Dey, and Jamie S. Evans

***Abstract*—This correspondence studies the performance of Kalman fixed lag smoothers with random packet losses and its comparison with the Kalman filter with packet loss. In terms of estimator stability via boundedness of the expectation of the error covariance, we show that smoothing does not provide any benefit over filtering. On the other hand, it is demonstrated that using a probabilistic notion of performance, smoothing can provide significant gains when compared to Kalman filtering. An analysis of Kalman filtering using two simple retransmission schemes and its comparison with Kalman smoothing is also made.**

***Index Terms*—Kalman filtering, Kalman smoothing, missing observations, retransmissions, stability.**

I. INTRODUCTION

Problems involving estimation over lossy communication networks have received considerable attention in recent years, due to their relevance in areas such as wireless sensor networks and networked control systems. When measurements from sensors are located at separate locations and have to be transmitted for processing through unreliable (e.g., wireless) channels, losses can occur, and how these packet losses affect the performance of the estimator is of significant interest.

Early work on state estimation with measurements losses include [1], where the optimal linear estimator for linear systems with independent identically distributed (i.i.d.) Bernoulli losses was derived, where the parameters of the loss process is known, but which of the individual measurements are lost/received is not explicitly known. This was later extended to the optimal linear smoother in [2]. More recently, in the case where we know which measurements are lost/received, it was shown in [3] that for an unstable system with i.i.d. Bernoulli losses there exists a critical threshold such that the expected value of the error covariance (which is randomly time varying due to the random losses) will be bounded if the packet arrival rate exceeds this threshold, but will diverge otherwise. Further avenues of research suggested in [3] include studying multiple sensors [4], correlated loss processes such as Markov [5], consideration of delays [6], and smoothing, which is the subject of

Manuscript received July 23, 2007; revised January 29, 2008. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Sergiy A. Vorobyov. This work was supported by the Australian Research Council.

The authors are with the ARC Special Research Centre for Ultra-Broadband Information Networks, Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, VIC 3010, Australia (e-mail: a.leong@ee.unimelb.edu.au; s.dey@ee.unimelb.edu.au; j.evans@ee.unimelb.edu.au).

Digital Object Identifier 10.1109/TSP.2008.920470

this correspondence. A different notion of estimator performance using probabilistic constraints was considered in [7]. A survey of these and other related results can be found in [8].

If the sensor has local computation ability, which is sometimes referred to as a "smart sensor," then an alternative scheme is to transmit the state estimate instead of the raw measurements [9], which allows estimator stability to be achieved with lower packet arrival rates. With computation available at the sensors, distributed Kalman filtering with quantized measurements [10] has also been considered. Other related work include state estimation with random measurement losses for jump linear systems [11], hidden Markov models [12], robust filtering [13], and the problem of control over packet dropping links; see [8], [14], and the references therein.

This correspondence considers the situation where we allow for some additional delay and computational complexity so that Kalman smoothing can be done on the measurements, and whether this provides any advantages over filtering. We assume that the sensors transmit the raw measurements directly as they do not have enough computation ability to be a smart sensor. We first derive in Section II the equations for the Kalman fixed lag smoother with random packet loss and use these equations to analyze its performance. While intuitively we might expect smoothing to perform better than filtering, in Section III, we show that with the stability notion via expected error covariance in [3], the use of Kalman smoothing does not actually provide any improvement over the Kalman filter. However, using instead the probabilistic notion of performance in [7], we will see in Section IV that the Kalman smoother can provide significant performance gains over the Kalman filter. In Section V, we analyze Kalman filtering using two simple retransmission strategies, which we find provides the same performance as a Kalman filter without retransmission, and hence Kalman smoothing outperforms these retransmission strategies with the probabilistic performance measure of [7].

II. Derivation of Kalman Fixed Lag Smoother With Random Packet Loss

Let the discrete time linear system be

$$\begin{aligned} x_{k+1} &= Ax_k + w_k \\ y_k &= Cx_k + v_k \end{aligned} \tag{1}$$

with w_k and v_k being Gaussian with zero mean and covariance matrices $Q \geq 0$ and $R > 0$, respectively. We will assume that the system is detectable and stabilizable. Let $\{\gamma_k\}$ be the random packet loss process that is equal to 1 if the measurement y_k is received at time k, and 0 otherwise. Define the state estimates and corresponding error covariances as

$$\begin{aligned} \hat{x}_{k\,|\,m} &= \mathbb{E}[x_k \,|\, \{Y_0, \dots, Y_m, \gamma_0, \dots, \gamma_m\}] \\ P_{k|m} &= \mathbb{E}\left[(x_k - \hat{x}_{k|m})(x_k - \hat{x}_{k\,|\,m})^T \,|\right. \\ &\qquad \left.\{Y_0, \dots, Y_m, \gamma_0, \dots, \gamma_m\}\right]. \end{aligned}$$

In [3], it is shown that the Kalman filter equations for the system (1) when there is packet loss can be written as

$$\begin{aligned} \hat{x}_{k+1\,|\,k} &= A\hat{x}_{k\,|\,k-1} + \gamma_k K_k (y_k - C\hat{x}_{k\,|\,k-1}) \\ P_{k+1\,|\,k} &= AP_{k\,|\,k-1}A^T - \gamma_k K_k C P_{k\,|\,k-1} A^T + Q \end{aligned} \tag{2}$$

where the Kalman gain $K_k = AP_{k\,|\,k-1}C^T(CP_{k\,|\,k-1}C^T + R)^{-1}$ (note that K_k here is defined slightly differently from [3]).

Now from [15, pp. 176–179], it is known that one way to derive a fixed lag Kalman smoother for the system (1), with smoothing lag N, is to consider the "augmented" model

$$\begin{bmatrix} x_{k+1} \\ x_{k+1}^{(1)} \\ x_{k+1}^{(2)} \\ \vdots \\ x_{k+1}^{(N+1)} \end{bmatrix} = \begin{bmatrix} A & 0 & \cdots & 0 & 0 \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} x_k \\ x_k^{(1)} \\ x_k^{(2)} \\ \vdots \\ x_k^{(N+1)} \end{bmatrix} + \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} w_k$$

$$y_k = [C \quad 0 \quad \cdots \quad 0 \quad 0] \begin{bmatrix} x_k \\ x_k^{(1)} \\ x_k^{(2)} \\ \vdots \\ x_k^{(N+1)} \end{bmatrix} + v_k$$

so that

$$x_{k+1}^{(1)} = x_k \quad x_{k+1}^{(2)} = x_{k-1}, \dots, x_{k+1}^{(N+1)} = x_{k-N}.$$

Define

$$\begin{aligned} \hat{x}_{k+1\,|\,k}^{(i)} &\equiv \mathbb{E}\left[x_{k+1}^{(i)} \,\middle|\, \{Y_0, \dots, Y_k, \gamma_0, \dots, \gamma_k\}\right] \\ &= \hat{x}_{k-(i-1)|k}, \qquad i = 1, \dots, N+1 \end{aligned}$$

and let

$$\hat{\mathbf{x}}_{k+1\,|\,k} = \begin{bmatrix} \hat{x}_{k+1\,|\,k} \\ \hat{x}_{k+1\,|\,k}^{(1)} \\ \vdots \\ \hat{x}_{k+1\,|\,k}^{(N+1)} \end{bmatrix} \quad \mathbf{K}_k = \begin{bmatrix} K_k \\ K_k^{(1)} \\ \vdots \\ K_k^{(N+1)} \end{bmatrix}$$

$$\mathbf{P}_{k+1\,|\,k} = \begin{bmatrix} P_{k+1\,|\,k}^{(0,0)} & P_{k+1\,|\,k}^{(0,1)} & \cdots & P_{k+1\,|\,k}^{(0,N+1)} \\ P_{k+1\,|\,k}^{(1,0)} & P_{k+1\,|\,k}^{(1,1)} & \cdots & P_{k+1\,|\,k}^{(1,N+1)} \\ \vdots & \vdots & \ddots & \vdots \\ P_{k+1\,|\,k}^{(N+1,0)} & P_{k+1\,|\,k}^{(N+1,1)} & \cdots & P_{k+1\,|\,k}^{(N+1,N+1)} \end{bmatrix}$$

where we make the identifications

$$\begin{aligned} P_{k+1\,|\,k}^{(0,0)} &= P_{k+1\,|\,k} \\ P_{k+1\,|\,k}^{(0,i)} &= P_{k+1\,|\,k}^{(i)} \\ P_{k+1\,|\,k}^{(i,i)} &= P_{k-(i-1)\,|\,k}. \end{aligned}$$

Then, applying the result of [3] [(2) in this correspondence] to the augmented model with $\hat{\mathbf{x}}_{k+1\,|\,k}, \mathbf{P}_{k+1\,|\,k}, \mathbf{K}_k$ in place of $\hat{x}_{k+1\,|\,k}, P_{k+1\,|\,k}, K_k$, the Kalman smoother equations with packet loss can be extracted after some computation as follows:

$$\begin{aligned} \hat{x}_{k+1\,|\,k} &= A\hat{x}_{k\,|\,k-1} + \gamma_k K_k (y_k - C\hat{x}_{k\,|\,k-1}) \\ \hat{x}_{k|k} &= \hat{x}_{k\,|\,k-1} + \gamma_k K_k^{(1)} (y_k - C\hat{x}_{k\,|\,k-1}) \\ \hat{x}_{k-i|k} &= \hat{x}_{k-i|k-1} + \gamma_k K_k^{(i+1)} (y_k - C\hat{x}_{k\,|\,k-1}), \\ &\qquad i = 1, \dots, N \\ K_{k+1} &= AP_{k+1\,|\,k}C^T(CP_{k+1\,|\,k}C^T + R)^{-1} \\ K_{k+1}^{(i)} &= P_{k+1\,|\,k}^{(i-1)} C^T (CP_{k+1\,|\,k}C^T + R)^{-1}, \\ &\qquad i = 1, \dots, N+1 \\ P_{k+1\,|\,k} &= AP_{k\,|\,k-1}A^T - \gamma_k K_k C P_{k\,|\,k-1} A^T + Q \\ P_{k+1\,|\,k}^{(i)} &= AP_{k\,|\,k-1}^{(i-1)} - \gamma_k K_k C P_{k\,|\,k-1}^{(i-1)}, \\ &\qquad i = 1, \dots, N+1 \end{aligned} \tag{3}$$

with the error covariances of the filtered and smoothed estimates given by

$$P_{k\,|\,k} = P_{k\,|\,k-1} - \gamma_k K_k^{(1)} C P_{k\,|\,k-1}$$
$$P_{k-i\,|\,k} = P_{k-i\,|\,k-1} - \gamma_k K_k^{(i+1)} C P_{k\,|\,k-1}^{(i)}, \qquad i=1,\ldots,N \quad (4)$$

and with initial conditions as stated in [15].

When there are no packet losses at time k, we have $\gamma_k = 1$ and we see that the equations reduce to the standard fixed lag smoother equations. However, when a measurement at time k is lost, i.e., $\gamma_k = 0$, we have $\hat{x}_{k-i\,|\,k} = \hat{x}_{k-i\,|\,k-1}$ and $P_{k-i\,|\,k} = P_{k-i\,|\,k-1}$ and the previous estimates and covariance matrices will get propagated.

III. Stability of Kalman Smoothing

In [3], the authors showed that for an unstable system (i.e., the matrix A has an eigenvalue with magnitude ≥ 1) with i.i.d. Bernoulli packet losses, there exists a critical threshold such that the expected value of the error covariance $P_{k\,|\,k-1}$ will be bounded if the packet arrival rate exceeds this threshold, but becomes unbounded otherwise. In this section, we will show the somewhat surprising result that for an unstable system with Bernoulli packet losses, smoothing does not improve the stability of the estimator, in the sense that the critical threshold for keeping the expected error covariances bounded in the smoothing case is the same as in the filtering case.

We first introduce some definitions. For $\{\gamma_k\}$ Bernoulli, let $\lambda = \mathbb{P}(\gamma_k = 1)$, then λ will also be the arrival rate of the measurements. For matrices P and Q, $P \leq Q$ will mean that $Q - P$ is positive semidefinite. Matrices P_k are said to be *bounded* if there exists a matrix $M < \infty$ such that $P_k \leq M, \forall k$, and *unbounded* if such an M does not exist. Then, we have the following.

Lemma 1: For unstable systems, as $k \to \infty$, $\mathbb{E}[P_{k\,|\,k+N}]$ is unbounded if and only if $\mathbb{E}[P_{k\,|\,k-1}]$ is unbounded.

Proof: First, from the smoother equations (4) and the definitions of $K_k^{(i)}$ in (3), it is not hard to see that $P_{k\,|\,k+N} \leq P_{k\,|\,k-1}, \forall k$, so the implication that $\mathbb{E}[P_{k\,|\,k-1}]$ bounded $\Rightarrow \mathbb{E}[P_{k\,|\,k+N}]$ bounded, or the equivalent statement $\mathbb{E}[P_{k\,|\,k+N}]$ unbounded $\Rightarrow \mathbb{E}[P_{k\,|\,k-1}]$ unbounded, immediately follows.

We now show that $\mathbb{E}[P_{k\,|\,k-1}]$ unbounded $\Rightarrow \mathbb{E}[P_{k\,|\,k+N}]$ unbounded. Let $\lambda < \lambda_c$, where λ_c is the critical arrival rate such that $\mathbb{E}[P_{k\,|\,k-1}]$ will be unbounded as $k \to \infty$ if and only if $\lambda < \lambda_c$, whose existence for unstable systems is shown in [3].[1] Let $\mathcal{A}$ be the event that measurements at time $k, k+1, \ldots, k+N$ are lost. Using some elementary properties of positive–semidefinite matrices and the smoother (4), we have

$$\begin{aligned}\mathbb{E}[P_{k\,|\,k+N}] &= \mathbb{E}[P_{k\,|\,k+N}|\mathcal{A}]\mathbb{P}(\mathcal{A}) + \mathbb{E}[P_{k\,|\,k+N}|\mathcal{A}^c]\mathbb{P}(\mathcal{A}^c)\\ &\geq \mathbb{E}[P_{k\,|\,k+N}|\mathcal{A}]\mathbb{P}(\mathcal{A})\\ &= \mathbb{E}[P_{k\,|\,k-1}]\mathbb{P}(\mathcal{A})\\ &= \mathbb{E}[P_{k\,|\,k-1}](1-\lambda)^{N+1}\end{aligned}$$

and $\mathbb{E}[P_{k\,|\,k-1}]$ by hypothesis is unbounded as $k \to \infty$ for $\lambda < \lambda_c$. Hence, $\mathbb{E}[P_{k\,|\,k+N}]$ is also unbounded as $k \to \infty$. ■

Since $\mathbb{E}[P_{k\,|\,k+N}]$ is unbounded if and only if $\mathbb{E}[P_{k\,|\,k-1}]$ is unbounded, the critical threshold on the arrival rate of packets for the Kalman smoother must, therefore, be the same as the critical threshold λ_c for the Kalman filter derived in [3]. Thus, from the stability point of view of keeping the expected error covariances bounded, there is no advantage to be gained in smoothing. One can compare this result with the work in [6] where packets can be both lost or delayed, which showed that stability of the Kalman filter (using constant gains) does not depend on packet delay but only on the probability that the packet eventually arrives.

[1]In the general vector case, λ_c cannot be determined even numerically, though upper and lower bounds can be derived [3], however exact knowledge of the critical rate will not be required in the proof.

IV. Kalman Smoothing With Probabilistic Constraints

Rather than studying the expected error covariance, an alternative notion of performance for Kalman filtering that has been considered is putting probabilistic constraints on the error covariance [7], [16]. The motivation for this is that low-probability events, such as a long sequence of measurement losses, can cause the expected error covariance to become unbounded, even when the "typical" behavior is such that the error covariance will lie below a certain value with high probability.

Given an upper bound M and an $\epsilon \in (0,1)$, one can ask the question as to what packet arrival rate λ is required in order to satisfy the constraint

$$\mathbb{P}(P_{k\,|\,k-1} \leq M) > 1 - \epsilon \quad (5)$$

where ϵ is usually chosen to be small so that the error covariance satisfies $P_{k\,|\,k-1} \leq M$ with probability close to one. This can be extended naturally to Kalman smoothing as the constraint

$$\mathbb{P}(P_{k\,|\,k+N} \leq M) > 1 - \epsilon. \quad (6)$$

While we showed in Section III that estimator stability in the sense of keeping expected error covariance bounded cannot be improved by smoothing, the situation is different when we consider probabilistic constraints. For instance, we will see that given M and ϵ the arrival rate which is sufficient to satisfy (6) will be smaller than what is required to satisfy (5).

Define

$$k_1 \equiv \min\{k \in Z^+ : h^k(\bar{M}) \not\leq M\}$$

where $h^k(X)$ means that the operator $h(X) \equiv AXA^T + Q$ is applied k times to X. For C invertible (e.g., a scalar system), $\bar{M}$ is defined as $\bar{M} \equiv AC^{-1}R(C^T)^{-1}A^T + Q$. When C is not invertible, the expression for $\bar{M}$ is more complicated and may be found in [17], though the arguments below will still hold.

It is shown in [7, Corollary 4] that if the losses are Bernoulli and the packet arrival rate λ satisfies the condition

$$\lambda \geq 1 - \epsilon^{\frac{1}{k_1}} \quad (7)$$

then (5) will also be satisfied. The quantity k_1 specifies the number of successive packet losses that can be tolerated before the error covariance updates for $P_{k\,|\,k-1}$ applied to the matrix $\bar{M}$ can no longer be bounded by the threshold M, which is then used to derive the condition (7). Using the property of the smoother equations (4) that previous estimates will get propagated when no measurements are received, it can be seen that an extra $N+1$ losses can be tolerated before the smoothing error covariances (with lag N) will exceed the threshold M. Hence, replacing k_1 with $k_1 + N + 1$ allows us to translate the results of [7] to the smoothing case, so that if

$$\lambda \geq 1 - \epsilon^{\frac{1}{k_1+N+1}} \quad (8)$$

then the condition (6) will be satisfied. From the condition (8), it may be seen that as the smoothing lag N increases, the packet arrival rate λ that is sufficient to guarantee (6) will be smaller, at the expense of additional delay and computational complexity.

TABLE I
ARRIVAL RATES λ SUFFICIENT TO SATISFY PROBABILISTIC CONSTRAINTS (5)–(6) AND SIMULATED PROBABILITIES, WITH $A = 1.3, C = 1, Q = 0.5, R = 1, M = 6.25, \epsilon = 0.05$

$P_{k\|k+N}$	$P_{k\|k-1}$	$P_{k\|k}$	$P_{k\|k+1}$	$P_{k\|k+2}$	$P_{k\|k+3}$	$P_{k\|k+4}$
λ	0.776	0.632	0.527	0.451	0.393	0.348
$\mathbb{P}(P_{k\|k+N} \leq M)$	0.975	0.970	0.963	0.959	0.958	0.959

As an illustration, consider the scalar system $A = 1.3, C = 1, Q = 0.5, R = 1$, where it can be determined that $k_1 = 2$. Choose $M = 6.25$ and $\epsilon = 0.05$, so that we want the error covariances to lie below 6.25 for at least 95% of the time. The second row of Table I shows the arrival rates required for several different smoothing lags N obtained using the condition (8). The third row contains simulated probabilities $\mathbb{P}(P_{k\,|\,k+N} \leq M)$ using the corresponding λ values, where we simulate the system over 100 000 time steps. We can see that in each case we have $\mathbb{P}(P_{k\,|\,k+N} \leq M) > 0.95$.

V. KALMAN FILTERING WITH RETRANSMISSIONS

In this section, we analyze the performance of Kalman filtering using some simple retransmission strategies. The packet losses will again be assumed to be Bernoulli with packet arrival rate λ.

A. Deterministic Retransmission Strategy

Consider the following retransmission strategy. If a measurement is lost at time k, ask for retransmission of this measurement up to N times (if the packet is still not received), while the measurements at times $k+1, k+2, \ldots$ are discarded when retransmission is occurring. Note that with this scheme only one packet is transmitted at any time instance and there is no queueing of packets, other than the single packet waiting to be transmitted. Assume that the probability of receiving a retransmitted packet is still λ. We are interested in the Kalman filtering performance (which could be delayed by up to N) of this retransmission scheme.

Recall the sequence $\{\gamma_k\}$ of 0's and 1's that specifies which measurements are lost and received. With retransmission we will look at the sequences of 0's and 1's together with whether the retransmitted packets are successful. To introduce notation, such a sequence for $N = 2$ might look like

$$1 \quad \not{0}^{\not{0}1} \quad \_ \quad \_ \quad \not{0}^{\not{0}\not{0}} \quad \_ \quad \_ \quad 1 \quad 1 \quad \not{0}^{1} \quad \_ \quad 1 \quad \cdots \tag{9}$$

where "$\not{0}^{\not{0}1}$" represents that the retransmission is successful on the second attempt, "$\not{0}^{\not{0}\not{0}}$" that both retransmissions are not successful, "$\not{0}^{1}$" that the retransmission succeeded on the first attempt, and "_" that the measurement is not sent (so will be assumed to be 0), so that (9) is equivalent to the sequence $110000011101\ldots$. It is clear that the number of 0's and 1's in both sequences are the same, hence the probabilities of each occurring will be the same. The key idea in analyzing the performance of this scheme is the following.

Lemma 2: Ignoring the first entry,[2] there is a bijection between sequences $\{\gamma_k\}$ that can be obtained without retransmission and sequences that can be obtained with retransmission. Moreover, they have the same probabilities of occurring.

Proof: First, for any valid sequence of retransmissions, we will clearly obtain a corresponding $\{\gamma_k\}$ that has the same probability of occurring.

[2]The reason we ignore the first entry is that cases like $01\ldots$ cannot be obtained using the retransmission strategy considered here since the second measurement will always be discarded when we ask for retransmission of the first measurement.

Now given a sequence $\{\gamma_k\}$, the following procedure will allow us to obtain a retransmission sequence that will match up with $\{\gamma_k\}$.

If $\{\gamma_k\}$ starts off with 0, go to step 1), otherwise go to step 2).

Step 1) Let m count the number of 0's before the first 1.
- If $m = 1$, then the first entry for the retransmission sequence is 1. Go to step 2).
- If $2 \leq m \leq N$, the sequence starts with 0, followed by retransmissions with success at the $(m-1)$th retransmission. Go to step 2).
- If $m > N$, write this as $m = a(N+1) + b$, with a and b being nonnegative integers and $b < N+1$.
  - — If $b = 0$, the sequence consists of 0's with all retransmissions failing repeated a times. Go to step 2).
  - — If $b = 1$, the sequence consists of a 1, followed by 0's with all retransmissions failing repeated a times. Go to step 2).
  - — Else the sequence is 0 followed by retransmissions with success at the $(b-1)$th retransmission, followed by 0's with all retransmissions failing repeated a times. Go to step 2).

Step 2) Let n count the number of 0's between two successive 1's.
- If $n = 0$, the next entry in the retransmission sequence is 1. Return to step 2).
- If $0 < n \leq N$, the next entries are a 0, followed by retransmissions with success at the nth retransmission. Return to step 2).
- If $n > N$, write this as $n = c(N+1) + d$, with c and d being nonnegative integers and $d < N+1$.
  - — If $d = 0$, the next entries are a 1, followed by 0's with all retransmissions failing repeated c times. Return to step 2).
  - — Else the next entries are a 0 followed by retransmissions with success at the dth retransmission, followed by 0's with all retransmissions failing repeated c times. Return to step 2).

Following this procedure, we can find a retransmission sequence that will match up with any given$\{\gamma_k\}$, and the probabilities of obtaining both sequences are the same. Uniqueness comes from the fact that lost measurements are always retransmitted in this scheme. ■

Step 1) in the proof of Lemma 2 takes care of the situation where there is a possible mismatch in the first entry due to $\{\gamma_k\}$ starting off with a 0. Step 2) of the proof is then repeatedly applied, and here we can always find a matching set of retransmissions. Table II shows some simple examples of $\{\gamma_k\}$ and the corresponding retransmission sequences, with $N = 3$. The first three columns involve applications of step 1), and we see that the only possible mismatch in the sequences in the first and third columns is in the first entry. The last three columns will involve step 2), and we can see that the sequences in the fourth and sixth column are matched up. As an example of the procedure in full, consider $\{\gamma_k\} = 0000000100001000001$, and we want to find a retransmission sequence equivalent to this, using $N = 3$. Then, in step 1), $m = 7 = 1 \times 4 + 3$, so $a = 1$ and $b = 3$. The first time we run

TABLE II
SOME SIMPLE $\{\gamma_k\}$ AND RETRANSMISSION SEQUENCES, WITH $N = 3$

$\{\gamma_k\}$	m	retransmission sequence	$\{\gamma_k\}$	n	retransmission sequence
01	1	11	11	0	11
001	2	$\emptyset^{1}$ _ 1	101	1	$\emptyset^{1}$ _ 1
0001	3	$\emptyset^{01}$ _ _ 1	1001	2	$\emptyset^{01}$ _ _ 1
00001	4	$\emptyset^{000}$ _ _ _ 1	10001	3	$\emptyset^{001}$ _ _ _ 1
000001	5	1 $\emptyset^{000}$ _ _ _ 1	100001	4	1 $\emptyset^{000}$ _ _ _ 1
0000001	6	$\emptyset^{1}$ _ $\emptyset^{000}$ _ _ _ 1	1000001	5	$\emptyset^{1}$ _ $\emptyset^{000}$ _ _ _ 1

step 2), we have $n = 4 = 1 \times 4 + 0$, so $c = 1$ and $d = 0$. The second time we run step 2), we have $n = 5 = 1 \times 4 + 1$, so $c = 1$ and $d = 1$. Following the procedure, the retransmission sequence is constructed as

$\emptyset^{01}$ _ _ $\emptyset^{000}$ _ _ _ 1 $\emptyset^{000}$ _ _ _ $\emptyset^{1}$ _
$\emptyset^{000}$ _ _ _ 1

which agrees with the original $\{\gamma_k\}$ apart from the first entry.

Thus, apart from a possible mismatch in the first entry, we know that there is always a sequence of retransmissions that will reproduce the same behavior as a sequence $\{\gamma_k\}$ for the Kalman filter without retransmission, with the same probability of occurring, so asymptotically, the probability distributions when doing Kalman filtering with retransmissions will be the same as that for Kalman filtering with no retransmissions. Hence, stability properties are the same as that of [3], i.e., the critical thresholds λ_c do not change. The probabilistic behavior will also be the same as for Kalman filtering with no retransmissions so the bounds in [7] will still apply. Therefore, this retransmission scheme provides no advantages over filtering using both of the performance measures considered in this correspondence, while at the same time possibly introducing a delay up to N. This agrees somewhat with the intuition that using new measurements for estimation is better than retransmitting old measurements, though here we actually showed that their distributions are essentially the same. Consequently, comparing smoothing with retransmissions, we find that Kalman smoothing will outperform this retransmission strategy using the probabilistic notion of performance as discussed in Section IV. We remark that this retransmission scheme and Lemma 2 can also be applied to other estimators such as the hidden Markov model filter [12], so long as the measurement losses are restricted to be an i.i.d. Bernoulli process.

B. Random Retransmission Strategy

An extension of the deterministic retransmission strategy is for retransmission requests to be random. Here, if a packet is not received, then with probability p it asks for a retransmission independently up to a maximum of N times (if retransmissions were unsuccessful), otherwise it will wait for the next measurement. It turns out that the distribution is again asymptotically the same as that for Kalman filtering without retransmissions, hence the same conclusions on its performance and comparison with Kalman smoothing applies.

To analyze this scheme, let us look at the case of a 1 followed by n successive 0's. Consider, for instance, $n = 3$ (i.e., the sequence 1000) and $N = 2$. In contrast to the deterministic strategy, there are now seven possible ways in which we can obtain this via retransmissions: $1\ 0^*\ 0^*\ 0^*$, $1\ 0^*\ \emptyset^{0^*}$ _, $1\ \emptyset^{0^*}$ _ 0^*, $1\ \emptyset^{00}$ _ _, $\emptyset^{1}$ _ $0^*\ 0^*$, $\emptyset^{1}$ _ $\emptyset^{0^*}$ _, $\emptyset^{01}$ _ _ 0^*, where the "$*$" here indicates that we did not ask for a retransmission. The corresponding probabilities of these events (here, we are ignoring the $\lambda(1-\lambda)^3$ term which results from receiving a 1 and three 0's) are $(1-p)^3, (1-p)p(1-p), p(1-p)^2, p^2, p(1-p)^2, p^2(1-p), p^2(1-p)$. However, one can check that the sum of these probabilities is equal to one, so that considering all these events as a whole gives a correspondence to the original sequence 1000, similar to Lemma 2. If we can show that this is true in general, then by independence and similar arguments as in the proof of Lemma 2 the probability distributions for this scheme will asymptotically be the same as for Kalman filtering without retransmissions.

To do this, we first look at the block of 0's by themselves. Let $f(n)$ be the sum of the probabilities of the different ways in which we can get n successive 0's. Then, $f(0) = 1, f(-1) = f(-2) = \cdots = 0$ and the following recursion holds:

$$\begin{aligned} f(n) = (1-p)f(n-1) &+ p(1-p)f(n-2) \\ &+ p^2(1-p)f(n-3) + \cdots \\ &+ p^{N-1}(1-p)f(n-N) + p^N f(n-N-1). \end{aligned} \quad (10)$$

This is because the term $(1-p)f(n-1)$ comes from not asking for a retransmission after the first 0, the term $p(1-p)f(n-2)$ comes from asking for retransmission once unsuccessfully, and then deciding not to ask again, and so on with the last term $p^N f(n-(N+1))$ the case where we asked for N retransmissions but all were unsuccessful.

Returning to the case of a 1 followed by n successive 0's, let $g(n)$ be the sum of the probabilities of the different ways in which we can do this. Then, $g(n)$ satisfies the relation

$$g(n) = f(n) + pf(n-1) + p^2 f(n-2) + \cdots + p^N f(n-N). \quad (11)$$

The term $f(n)$ represents the situation where we successfully received the first measurement, and the other terms represent cases where there is retransmission of the first measurement until a success, similar to how (10) is derived. Substituting (10) into (11), we find that

$$\begin{aligned} g(n) = f(n-1) + pf(n-2) + p^2 f(n-3) + \cdots \\ + p^N f(n-N-1). \end{aligned}$$

Comparing this with (11), we see that $g(n)$ must be constant, and hence $g(n) = \cdots = f(0) = 1$.

VI. CONCLUSION

In this correspondence, we have derived the Kalman smoother in the presence of random packet losses, and compared it against the Kalman filter with packet losses as well as a simple retransmission scheme. We found that the Kalman smoother provided gains over the Kalman filter in the probabilistic sense of [7], but not in terms of the stability notion of [3]. We also found that the simple retransmission schemes considered here will statistically perform in the same way as the Kalman filter without retransmission, while also introducing additional delay. Future work will involve analysis of retransmission strategies where more than one packet can be sent at a time.

REFERENCES

[1] N. E. Nahi, "Optimal recursive estimation with uncertain observation," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 4, pp. 457–462, Jul. 1969.

[2] R. A. Monzingo, "Discrete optimal linear smoothing for systems with uncertain observations," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 3, pp. 271–275, May 1975.

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

[4] X. Liu and A. J. Goldsmith, "Kalman filtering with partial observation losses," in *Proc. IEEE Conf. Decision Control*, Bahamas, Dec. 2004, pp. 1413–1418.

[5] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, no. 4, pp. 598–607, Apr. 2007.

[6] L. Schenato, "Optimal estimation in networked control systems subject to random delay and packet drop," in *Proc. IEEE Conf. Decision Control*, San Diego, CA, Dec. 2006, pp. 5615–5620.

[7] L. Shi, M. Epstein, A. Tiwari, and R. M. Murray, "Estimation with information loss: Asymptotic analysis and error bounds," in *Proc. IEEE Conf. Decision Control*, Seville, Spain, Dec. 2005, pp. 1215–1221.

[8] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proc. IEEE*, vol. 95, no. 1, pp. 138–162, Jan. 2007.

[9] Y. Xu and J. P. Hespanha, "Estimation under uncontrolled and controlled communications in networked control systems," in *Proc. IEEE Conf. Decision Control*, Seville, Spain, Dec. 2005, pp. 842–847.

[10] 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 Process.*, vol. 54, no. 12, pp. 4782–4795, Dec. 2006.

[11] A. K. Fletcher, S. Rangan, and V. K. Goyal, "Estimation from lossy sensor data: Jump linear modelling and Kalman filtering," in *Proc. Int. Conf. Inf. Process. Sensor Netw.*, Berkeley, CA, Apr. 2004, pp. 251–258.

[12] A. S. Leong, S. Dey, and J. S. Evans, "Probability of error analysis for hidden Markov model filtering with random packet loss," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 809–821, Mar. 2007.

[13] Z. Wang, F. Yang, D. W. C. Ho, and X. Liu, "Robust H_∞ filtering for stochastic time-delay systems with missing measurements," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2579–2587, Jul. 2006.

[14] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry, "Foundations of control and estimation over lossy networks," *Proc. IEEE*, vol. 95, no. 1, pp. 163–187, Jan. 2007.

[15] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[16] S. Adlakha, "Kalman filtering in correlated losses," Stanford Univ., Stanford, CA, Tech. Rep., 2005.

[17] M. Epstein, L. Shi, A. Tiwari, and R. M. Murray, "Probabilistic performance of state estimation across a lossy network," *Automatica*, 2008, to be published.

Modified Pisarenko Harmonic Decomposition for Single-Tone Frequency Estimation

Kenneth Wing Kin Lui and Hing Cheung So

***Abstract*—In this correspondence, based on an alternative derivation of the Pisarenko harmonic decomposition (PHD) method, a new asymptotically unbiased estimator for the frequency of a single real tone in white noise is devised with the use of novel sample covariance expressions. Furthermore, extension to sample covariances with higher lags for performance enhancement is investigated while a simple and effective scheme is suggested to resolve the corresponding frequency ambiguity problem. The variance of the modified Pisarenko's method is also derived, which is then utilized to find the best estimate among all admissible solutions from various sets of sample covariances. Computer simulations are included to corroborate the theoretical development and to demonstrate that the proposed approach outperforms several existing low-complexity frequency estimators in terms of nearly uniform performance and estimation accuracy.**

***Index Terms*—Frequency estimation, Pisarenko's method, sample covariance, single real sinusoid.**

I. INTRODUCTION

Frequency estimation of sinusoidal signals in noise is a frequently addressed problem in the signal processing literature [1]–[5] because of its wide applicability in control theory, digital communications, biomedical engineering, instrumentation and measurement, and so on. In this work, we address the fundamental problem of single sinusoidal frequency estimation, and its discrete-time signal model is

$$x(n) = s(n) + q(n), \qquad n = 1, 2, \dots N \tag{1}$$

where

$$s(n) = \alpha \cos(\omega n + \phi). \tag{2}$$

The α, $\omega \in (0, \pi)$, and $\phi \in [0, 2\pi)$ are unknown but deterministic constants that represent the tone amplitude, frequency, and phase, respectively, while the noise $q(n)$ is assumed to be a zero-mean white process with unknown variance σ^2. The task is to find ω given the N samples of $\{x(n)\}$.

Under Gaussian noise assumption, the maximum-likelihood (ML) estimate of frequency [6], with estimation accuracy of order $N^{-3/2}$ in standard error, is obtained by maximizing a highly nonlinear and multimodal cost function, and thus extensive computations are involved. Apart from the ML method, some relatively fast algorithms such as the discrete Fourier spectrum (DFS) interpolator [7], contraction mapping [8], [9], and weighted subspace fitting [10] can achieve this accuracy. It is worthy to note that the efficient methods for complex tone frequency estimation [11]–[14] generally cannot be employed to real-valued data. For applications where real-time estimation is required, more computationally efficient but suboptimal frequency estimators such as notch filtering, Capon methods, linear prediction [15]–[17], Yule-Walker methods [18] and subspace-based approaches [19] are widely used choices. In this correspondence, we focus on fast frequency estimation of a single real tone. Our main contributions are summarized as follows: 1) proposal of novel sample

Manuscript received June 28, 2007; revised February 4, 2008. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Brian M. Sadler.

The authors are with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong (e-mail: Kenneth.Lui@student.cityu.edu.hk; hcso@ee.cityu.edu.hk).

Digital Object Identifier 10.1109/TSP.2008.920486