Noname manuscript No.
(will be inserted by the editor)

Networked Control Systems: Estimation and Control over Lossy Networks

João P. Hespanha[†] **Alexandre R. Mesquita**[‡]

the date of receipt and acceptance should be inserted later

Abstract This article discusses optimal estimation and control for lossy networks. Conditions for stability are provided both for two-link and multiple-link networks. The online adaptation of network resources (controlled communication) is also considered.

Keywords

Networked control systems, automatic control, communication networks, stability, estimation, controlled communication

1 Introduction

Network Control Systems (NCSs) are spatially distributed systems in which the communication between sensors, actuators and controllers occurs through a shared band-limited digital communication network. In this article we consider the problem of estimation and control over such networks.

A significant difference between NCSs and standard digital control is the possibility that data may be lost while in transit through the network. Typically, *packet dropouts* result from transmission errors in physical network links (which is far more common in wireless than in wired networks) or from buffer overflows due to congestion. Long transmission delays sometimes result in packet re-ordering, which essentially amounts to a packet dropout if the receiver discards “outdated” arrivals. Reliable transmission protocols, such as TCP, guarantee the eventual delivery of packets. However, these protocols are not appropriate for NCSs since the retransmission of old data is generally not useful. Another important difference between NCSs and standard digital control systems is that, due to the nature of network traffic, delays in the control loop may be time varying and non-deterministic.

In this article we concentrate on the problem of control and estimation in the presence of packet losses, leaving other important features of NCSs (such as quantization and random delays) to be addressed in other articles of this encyclopedia. Consequently, we assume that the network can be viewed as a channel that can carry real numbers without distortion, but that some of the messages may be lost. This network model is appropriate when the number of bits in each data packet is sufficiently large so that quantization effects can be ignored, but packet dropouts cannot. For more general channel models, see for example [Imer and Basar, 2005].

This article also does not address network transmission delays explicitly. In general, network delays have two components: one that is due to the time spent transmitting packets and another due to the time packets wait in buffers waiting to be transmitted. Delays due to packet transmission present little variation and may be modeled as constants. For control design purposes, these delays may be incorporated into the plant model. Delays due to buffering depend on the network traffic and are typically random; they can be analyzed using the techniques developed in [Antunes et al, 2012].

Notation and basic definitions. Throughout the article, $\mathbb{R}$ stands for real numbers and $\mathbb{N}$ for nonnegative integers. For a given matrix $A \in \mathbb{R}^{n\times n}$ and vector $x \in \mathbb{R}^n$, $\|x\| := \sqrt{x'x}$ denotes the Euclidean norm of x, and $\lambda(A)$ the set of eigenvalues of A. Random variables are generally denoted in boldface. For a random variable $\mathbf{y}$, $\mathrm{E}[\mathbf{y}]$ stands for the expectation of $\mathbf{y}$.

[†] Center for Control, Dynamical Systems and Computation, University of California, Santa Barbara, CA 93106, U.S.A. ·
[‡] Dept. of Electronics Eng., Federal Univ. of Minas Gerais, Av. Antonio Carlos 6627, 31270-010, Belo Horizonte, Brazil.

2 Two-Link Networks

Here we consider a control/estimation problem when all network effects can be modeled using two erasure channels: one from the sensor to the controller, the other from the controller to the actuator (see Fig. 1).

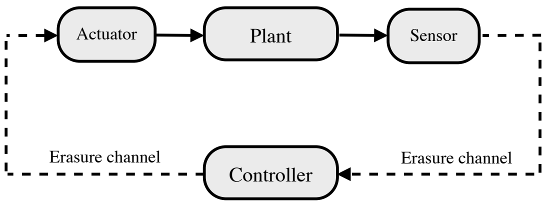

Fig. 1 Control system with two network links

We restrict our attention to a linear time-invariant (LTI) plant with intermittent observation and control packets:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + \boldsymbol{\nu}_k B\mathbf{u}_k + \mathbf{w}_k, \tag{1a}$$

$$\mathbf{y}_k = \boldsymbol{\theta}_k C\mathbf{x}_k + \mathbf{v}_k, \tag{1b}$$

$\forall k \in \mathbb{N},\ \mathbf{x}_k, \mathbf{w}_k \in \mathbb{R}^n,\ \mathbf{y}_k, \mathbf{v}_k \in \mathbb{R}^p$, where $(\mathbf{x}_0, \mathbf{w}_k, \mathbf{v}_k)$ are mutually independent, zero-mean Gaussian with covariance matrices (P_0, R_w, R_v), and $\boldsymbol{\theta}_k, \boldsymbol{\nu}_k \in \{0, 1\}$ are i.i.d. Bernoulli random variables with $\Pr\{\boldsymbol{\theta}_k = 1\} = \bar{\theta}$ and $\Pr\{\boldsymbol{\nu}_k = 1\} = \bar{\nu}$. The variable $\boldsymbol{\theta}_k$ models the packet loss between sensor and controller, whereas $\boldsymbol{\nu}_k$ models the packet loss between controller and actuator. When there is a packet drop from controller to actuator, we set the actuator's output to zero. Different strategies, such as holding the control input, could still be modeled using (1) by augmenting of the state vector.

The information available to the controller up to time k is given by the information set:

$$\mathcal{I}_k = \{P_0\} \cup \{\mathbf{y}_\ell, \boldsymbol{\theta}_\ell : \ell \le k\} \cup \{\boldsymbol{\nu}_\ell : \ell \le k-1\}.$$

Here we make an important assumption that acknowledgment packets from the actuator are always received by the controller so that $\boldsymbol{\nu}_\ell$, $\ell \le k-1$ is available at time k to the remote estimator.

2.1 Optimal estimation with remote computation

The optimal mean-square estimate of $\mathbf{x}_k$, given the information known to the remote estimator at time k is given by

$$\hat{\mathbf{x}}_{k|k} := \mathrm{E}\left[\mathbf{x}_k \,\middle|\, \mathcal{I}_k\right].$$

This estimate can be computed recursively using the following time-varying Kalman filter (TVKF) [Sinopoli et al, 2004]:

$$\hat{\mathbf{x}}_{0|-1} = 0, \tag{2a}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \boldsymbol{\theta}_k F_k(\mathbf{y}_k - C\hat{\mathbf{x}}_{k|k-1}), \tag{2b}$$

$$\hat{\mathbf{x}}_{k+1|k} = A\hat{\mathbf{x}}_{k|k} + \boldsymbol{\nu}_k B\mathbf{u}_k, \tag{2c}$$

with the gain matrix F_k calculated recursively as follows

$$F_k = P_k C'(CP_kC' + R_v)^{-1},$$
$$P_{k+1} = AP_kA' + R_w - \boldsymbol{\theta}_k AF_k(CP_kC' + R_v)F_k'A'.$$

Each P_k corresponds to the estimation error covariance matrix

$$P_k = \mathrm{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})'].$$

For this estimator, there exists a critical value θ_c for the dropout rate $\bar{\theta}$, above which the estimation error covariance becomes unbounded:

Theorem 1 ([Sinopoli et al, 2004]) *Assume that $(A, R_w^{1/2})$ is controllable, (A, C) is observable and A is unstable. Then there exists a critical value $\theta_c \in (0, 1]$ such that*

$$\mathrm{E}[P_k] \le M, \forall k \in \mathbb{N} \quad \Leftrightarrow \quad \bar{\theta} \ge \theta_c$$

where M is a positive definite matrix that may depend on P_0. Furthermore, the critical value θ_c satisfies $\theta_{\min} \le \theta_c \le \theta_{\max}$, where the lower-bound is given by

$$\theta_{\min} = 1 - \frac{1}{(\max\{|\lambda(A)|\})^2}, \tag{3}$$

and the upper-bound is given by the solution to the following (quasi-convex) optimization problem

$$\theta_{\max} = \min\big\{\theta \ge 0 : \Psi_\theta(Y, Z) > 0,\ 0 \le Y \le I \text{ for some } Y, Z\big\},$$

where

$$\Psi_\theta(Y, Z) = \begin{bmatrix} Y & \sqrt{\theta}(YA+ZC) & \sqrt{1-\theta}YA \\ \sqrt{\theta}(A'Y+C'Z') & Y & 0 \\ \sqrt{1-\theta}A'Y & 0 & Y \end{bmatrix}.$$

□

Remark 1 In some special cases, the upper-bound in (3) is tight in the sense that $\theta_c = \theta_{\min}$. The largest class of systems known for which this occurs is that of *non-degenerate systems* defined in [Mo and Sinopoli, 2012]. Examples of systems in this class include: 1) those for which the matrix C is invertible; 2) those with a detectable pair (A, C) and such that the matrix A is diagonalizable with unstable eigenvalues having distinct absolute values.

2.2 Optimal control with remote computation

From a control perspective, one may also be interested in finding control sequences $\mathbf{u}^N = \{\mathbf{u}_1, \ldots, \mathbf{u}_{N-1}\}$, as functions of the information set $\mathcal{I}_N$, which minimize cost functions of the form

$$J = \lim_{N\to\infty} \frac{1}{N} \mathrm{E}\Big[\sum_{k=0}^{N-1} (\mathbf{x}_k' W \mathbf{x}_k + \boldsymbol{\nu}_k \mathbf{u}_k' U \mathbf{u}_k)\Big|\mathcal{I}_k\Big].$$

Theorem 2 ([Schenato et al, 2007]) *Assume that (A, B) and $(A, R_w^{1/2})$ are controllable, (A, C) and $(A, W^{1/2})$ are observable and A is unstable. Then, finite control costs J are achievable if and only if $\bar{\theta} > \theta_c$ and $\bar{\nu} > \nu_c$, where the critical value ν_c is given by the (quasi-convex) optimization problem*

$$\nu_c = \min \big\{\nu \geq 0 : \Psi_\nu(Y, Z) > 0, \\ 0 \leq Y \leq I \textit{ for some } Y, Z\big\},$$

where

$$\Psi_\nu(Y, Z) = \\ \begin{bmatrix} Y & Y & \sqrt{\nu} Z U^{1/2} & \sqrt{\nu}(YA'+ZB') & \sqrt{1-\nu} Y A' \\ Y & W^{-1} & 0 & 0 & 0 \\ \sqrt{\nu} U^{1/2} Z' & 0 & I & 0 & 0 \\ \sqrt{\nu}(AY+BZ') & 0 & 0 & Y & 0 \\ \sqrt{1-\nu} AY & 0 & 0 & 0 & Y \end{bmatrix}.$$

□

Moreover, under the above conditions, the separation principle holds in the sense that the optimal control is given by

$$\mathbf{u}_k = -(B'SB + U)^{-1} B'SA\hat{\mathbf{x}}_{k|k},$$

where $\hat{\mathbf{x}}_{k|k}$ is an optimal state estimate given by (2) and the matrix S is the solution to the modified algebraic Riccati (MARE) equation

$$S = A'SA + W - \bar{\nu} A'SB(B'SB + U)^{-1} B'SA.$$

Solutions to the MARE may be obtained iteratively when $\bar{\nu} > \nu_c$.

2.3 Estimation with local computation

To reduce the gap between the bounds $\theta_{\min}$ and $\theta_{\max}$ on the critical value of the drop probability in Theorem 1 and to allow for larger probabilities of drop, one may choose to compute state estimates at the sensor and transmit those to the controller/actuator. This scheme is motivated by the growing number of *smart sensors* with embedded processing units that are capable of local computation. For the LTI plant

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k + \mathbf{w}_k, \\ \mathbf{y}_k = C\mathbf{x}_k + \mathbf{v}_k,$$

the smart sensor can compute locally an optimal state estimate using a standard stationary Kalman filter and transmits this estimate to the controller. We model packet dropouts as before using the process $\boldsymbol{\theta}_k$ and assume that the process $\boldsymbol{\theta}_k$ is known to the smart sensor by means of an perfect acknowledgment mechanism. This allows the sensor to know $\mathbf{u}_k$ exactly and to use it in the Kalman filter.

Let $\tilde{\mathbf{x}}_{k|k} = \mathrm{E}[\mathbf{x}_k | \mathbf{y}_\ell, \boldsymbol{\theta}_\ell, \ell \leq k]$ denote the local estimates transmitted by the sensor. Using the messages successfully received up to time k, the remote estimator computes the optimal estimate

$$\hat{\mathbf{x}}_{k|k-1} = \mathrm{E}\left[\mathbf{x}_k \mid \boldsymbol{\theta}_\ell, \tilde{\mathbf{x}}_{\ell|\ell}, \ell \leq k-1\right].$$

recursively by

$$\hat{\mathbf{x}}_{0|-1} = 0, \\ \hat{\mathbf{x}}_{k|k} = (1 - \boldsymbol{\theta}_k)\hat{\mathbf{x}}_{k|k-1} + \boldsymbol{\theta}_k \tilde{\mathbf{x}}_{k|k}, \quad k \in \mathbb{N}, \\ \hat{\mathbf{x}}_{k+1|k} = A\hat{\mathbf{x}}_{k|k} + B\mathbf{u}_k$$

Notice that now we are applying the (TVKF) to estimate $\tilde{\mathbf{x}}_k$, which is fully observable. Since $\theta_{\min}$ and $\theta_{\max}$ in Theorem 2 are equal for fully observable processes [Schenato et al, 2007], the local computation scheme grants a minimal critical value θ_c as stated in the theorem below.

Theorem 3 *Assume that $(A, R_w^{1/2})$ is controllable, (A, C) is observable and A is unstable. Then the critical value θ_c is given by $\theta_{\min}$ in (3), i.e.,*

$$\mathrm{E}[P_k] \leq M, \forall k \in \mathbb{N} \quad \Leftrightarrow \quad \bar{\theta} \geq \theta_{\min}$$

where M is a positive definite matrix that may depend on P_0.

2.4 Drops in the acknowledgement packets

When there are drops in the acknowledgement channel from the actuator to the controller, the controller does not always know $\boldsymbol{\nu}_k$ and therefore it might not always have access to the control inputs that are actually applied to the plant. In this case the posterior state probability becomes a Gaussian mixture distribution with infinitely many components and the separation principle no longer holds [Schenato et al, 2007]. This makes the estimation and control problems computationally more difficult and, due to the smaller information set, some performance degradation in the control performance should be expected. For this reason, it is generally a good design choice to keep controller and actuator collocated when drops in the acknowledgement channels are significant.

2.5 Buffering

As an alternative to the approach described in Section 2.3 to use local computation at a smart sensor to allow for larger probabilities of drop, the designer may also consider the transmission of a sequence of previous measurements $\mathbf{y}_k, \mathbf{y}_{k-1}, \ldots, \mathbf{y}_{k-N}$ in each packet. This approach is motivated by the fact that often data packets can carry much more than one vector of measured outputs. When N is reasonably large, one should expect similar estimation/control performances as in the approach described in Section 2.3, but with a reduced computational effort at the sensor.

Analogously, an improvement to zeroing or simply holding the control input in case of packet drops between controller and actuator is for the controller to transmit a control sequence $\mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_{k+N}$ that contains not only the control $\mathbf{u}_k$ to be used at the current time instant, but also a few future controls $\mathbf{u}_{k+1}, \mathbf{u}_{k+2}, \ldots, \mathbf{u}_{k+N}$. In the case of packet drops between controller of actuator, the actuator can use previously received "future" control inputs in lieu of the one contained in the lost packet. The sequence of future control inputs may be obtained, e.g., by an optimal receding horizon control strategy [Gupta et al, 2006].

2.6 Estimation with Markovian Drops

When $\boldsymbol{\theta}_k$ is a Markov process, we no longer have a separation principle and the optimal controller may depend on the drops sequence. Yet, optimal state estimates are obtained using the same TVKF presented earlier. Below we give conditions for the stability of the error covariance when drops are governed by the Gilbert-Elliot model: $\Pr\{\boldsymbol{\theta}_{k+1} = j \mid \boldsymbol{\theta}_k = i\} = p_{ij},\ i, j \in \{0, 1\}$.

Theorem 4 *[Mo and Sinopoli, 2012] Assume that $(A, R_w^{1/2})$ is controllable, A is unstable and the system given by the pair (A, C) is non-degenerate as discussed in Remark 1. Moreover, suppose that the transition probabilities for the Gilbert-Elliot model satisfy $p_{01}, p_{10} > 0$. Then the expected error covariance $\mathrm{E}[P_k]$ is uniformly bounded if*

$$p_{01} > \theta_{\min}$$

and it is unbounded for some initial condition if $p_{01} < \theta_{\min}$.

3 Networks with Multiple Links

We now consider feedback loops that are closed over a network of communication links, each of which drops packets according to a Bernoulli process. The sensor communicates with a controller across the network and we assume that controller and actuator are collocated. The network may be represented by a graph $\mathcal{G}$ with nodes in the set $\mathcal{V}$ and edges in the set $\mathcal{E}$, where edges are drawn between two communicating nodes. We denote by p_{ij} the probability of a drop when node i transmits to node j. Drops are assumed to be independent across links and time.

To maximize robustness with respect to drops, sensors use a Kalman filter to compute an optimal estimate for the state of the process based on their measurements and transmit this estimate across the network. When the sensors do not have access to the process input, they can take advantage of the linearity of the Kalman filter: as the output of a Kalman filter is the sum of a term due to measurements with another term due to control inputs, sensors may compute only the contribution due to measurements and transmit it to the controller, which can subsequently add the contribution due to the control inputs. This guarantees that optimal state estimates can still be computed at the control node, even when the sensors do not know the control input [Gupta et al, 2009].

The communication in the network goes as follows. Sensors time stamp their estimates and broadcast them to all nodes in their communication ranges. After receiving information from their neighbors, nodes compare time stamps and keep only the most recent estimates. These estimates are broadcasted to all neighboring nodes. When the controller receives new information, the optimal Kalman estimate is reconstructed, taking into account the total transmission delay (learned from the packet time stamps), and a standard LQG control can be used [Gupta et al, 2009].

To determine whether or not this procedure results in a stable closed loop, one defines a *cut* $\mathcal{C} = (\mathcal{S}, \mathcal{T})$ to be a partition of the node set $\mathcal{V}$ such that the sensor node is in $\mathcal{S}$ and the controller node is in $\mathcal{T}$. The *cut-set* is then defined as the set of edges $(i, j) \in \mathcal{E}$ such that $i \in \mathcal{S}$ and $j \in \mathcal{T}$, i.e., the set of edges that connect the sets $\mathcal{S}$ and $\mathcal{T}$. The *max-cut probability* is then defined as

$$p_{max-cut} = \max_{\text{all cuts } (\mathcal{S}, \mathcal{T})} \prod_{(i,j) \in \mathcal{S} \times \mathcal{T}} p_{ij}.$$

The above maximization can be rewritten as a minimization over the sums of $-\log p_{ij}$, which leads to a linear program known as the minimum cut problem in network optimization theory [Cook, 1995].

Theorem 5 *[Gupta et al, 2009] Assume that $R_w, R_v > 0$, that (A, B) is stabilizable, that (A, C) is observable and that A is unstable. Then the control and communication policy described above is optimal for quadratic costs and the expected state covariance is bounded if and only if*

$$p_{max-cut} \cdot (\max\{|\lambda(A)|\})^2 < 1.$$

4 Estimation with controlled communication

To actively reduce network traffic and power consumption, sensor measurements may not be sent to the remote estimator at every time step. In addition, one may have the ability to somewhat control the probability of packet drops by varying the transmit power or by transmitting copies of the same message through multiple channel realizations. This is known as *controlled communication* and it allows the designer to establish a trade-off between communication and estimation performance.

We consider the local estimation scenario described in Section 2.3 with the difference that the Bernoulli drops are now modulated as follows

$$\boldsymbol{\theta}_k = \begin{cases} 1 & \text{with prob. } \Lambda_k \\ 0 & \text{with prob. } 1 - \Lambda_k \end{cases}$$

where the sensor is free to choose $\Lambda_k \in [0, p_{\max}]$ as a function of the information available up to time k. With its choice, the sensor incurs on a communication cost $c(\Lambda_k)$ at time k, where $c(\cdot)$ is some increasing function that may represent, for example, the energy needed in order to transmit with a probability of drop equal to Λ_k. Note that transmission scheduling, where Λ_k is either 0 or $p_{\max}$, is a special case of this framework.

In order to choose Λ_k the sensor considers the estimation error $\tilde{\mathbf{e}}_k := \tilde{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k-1}$ between the local and the remote estimators. This error evolves according to

$$\tilde{\mathbf{e}}_{k+1} = \begin{cases} \mathbf{d}_k & \text{with prob. } \Lambda_k \\ A\tilde{\mathbf{e}}_k + \mathbf{d}_k & \text{with prob. } 1 - \Lambda_k, \end{cases}$$

where $\mathbf{d}_k$ is the innovations process arising from the standard Kalman filter in the smart sensor.

Our objective is to find a "communication policy" that minimizes the long-term average cost

$$\tilde{J} := \lim_{K\to\infty} \frac{1}{K} \mathrm{E}\left[\sum_{k=0}^{K-1} \|\tilde{\mathbf{e}}_k\|^2 + \lambda c(\Lambda_k)\right], \quad \lambda > 0, \qquad (4)$$

which penalizes a linear combination of the remote estimation error variance $\mathrm{E}[\|\tilde{\mathbf{e}}_k\|^2]$ and the average communication cost $\mathrm{E}[c(\Lambda_k)]$. In this context, a *communication policy* should be understood as a rule that selects Λ_k as a function of the information available to the sensor.

When

$$(1 - p_{\max}) \max\{|\lambda(A)|\}^2 < 1,$$

there exists an optimal communication policy that chooses Λ_k as a function of $\tilde{e}_k$, which may be computed via dynamic programming and value iteration [Mesquita et al, 2012]. While this procedure can be computationally difficult, it is often possible to obtain sub-optimal but reasonable performance with rollout policies such as the following one:

$$\Lambda_k = \arg \min_{\Lambda\in[0,p_{\max}]} [(p_{\max} - \Lambda)\tilde{\mathbf{e}}_k' A' H A \tilde{\mathbf{e}}_k + \lambda c(\Lambda)] \qquad (5)$$

where H is the positive semidefinite solution to the Lyapunov equation $(1 - p_{\max})A'HA - H = -I$ [Mesquita et al, 2012].

When computing $\tilde{\mathbf{e}}_k$ and Λ_k in (5) is computationally too costly for the sensor, one may prefer to make Λ_k a function of the number of consecutive dropped packets $\boldsymbol{\ell}_k$. In this case, minimizing $\tilde{J}$ in (4) is equivalent to minimizing the cost

$$\bar{J} := \lim_{K\to\infty} \frac{1}{K} \mathrm{E}\left[\sum_{k=0}^{K-1} \mathrm{trace}(\Sigma_{\boldsymbol{\ell}_k}) + \lambda c(\Lambda_k)\right],$$

where

$$\Sigma_\ell := \sum_{m=0}^{\ell} A'^m R_w A^m.$$

Since $\boldsymbol{\ell}_k$ belongs to a countable set, one can very efficiently solve this optimization using dynamic programming [Mesquita et al, 2012].

5 Summary and Future Directions

Most positive results in the subject rely on the assumption of perfect acknowledgments and on actuators and controllers being collocated. Future research should address ways of circumventing these assumptions.

Recommended Reading

Antunes D, Hespanha JP, Silvestre C (2012) Volterra integral approach to impulsive renewal systems: Application to networked control. IEEE Transaction on Automatic Control 57:607–619

Cook WJ (1995) Combinatorial optimization: papers from the DIMACS Special Year, vol 20. Amer Mathematical Society

Gupta V, Sinopoli B, Adlakha S, Goldsmith A, Murray R (2006) Receding horizon networked control. In: Proc. Allerton Conf. Commun., Control Comput

Gupta V, Dana AF, Hespanha JP, Murray RM, Hassibi B (2009) Data transmission over networks for estimation and control. Automatic Control, IEEE Transactions on 54(8):1807–1819

Imer OC, Basar T (2005) Optimal estimation with limited measurements. In: Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on, IEEE, pp 1029–1034

Mesquita AR, Hespanha JP, Nair GN (2012) Redundant data transmission in control/estimation over lossy networks. Automatica 48(8):1612 – 1620

Mo Y, Sinopoli B (2012) Kalman filtering with intermittent observations: tail distribution and critical value. Automatic Control, IEEE Transactions on 57(3):677–689

Schenato L, Sinopoli B, Franceschetti M, Poolla K, Sastry SS (2007) Foundations of control and estimation over lossy networks. Proceedings of the IEEE 95(1):163–187

Sinopoli B, Schenato L, Franceschetti M, Poolla K, Jordan MI, Sastry SS (2004) Kalman filtering with intermittent observations. IEEE Trans on Automat Contr 49(9):1453–1464